Universidade Federal do Rio Grande do Sul
Instituto de Pesquisas Hidráulicas – IPH

Programa de Pós-Graduação em Recursos Hídricos e Saneamento Ambiental

Iporã Brito Possantti

A synthesis on fundamental principles and problems in the use of hydrological models for planning watershed conservation

Porto Alegre

2024

Iporã Brito Possantti

A synthesis on fundamental principles and problems in the use of hydrological models for planning watershed conservation

Thesis presented to the Graduate Program in Water Resources and Environmental Sanitation at the Federal University of Rio Grande do Sul, as a partial requirement for obtaining the degree of Doctor of Philosophy.

Advisor:
Prof. Dr. Guilherme Fernandes Marques

Porto Alegre
2024

CIP - Catalogação na Publicação

Possantti, Iporã
Uma síntese sobre princípios e problemas fundamentais no uso de modelos hidrológicos para planejar a conservação de bacias hidrográficas / Iporã Possantti. -- 2024.
225 f.
Orientador: Guilherme Marques.

Tese (Doutorado) -- Universidade Federal do Rio Grande do Sul, Instituto de Pesquisas Hidráulicas, Programa de Pós-Graduação em Recursos Hídricos e Saneamento Ambiental, Porto Alegre, BR-RS, 2024.

1. Hidrologia. 2. Recursos Hídricos. 3. Bacias Hidrográficas. 4. Serviços Ecossistêmicos. 5. Modelos Hidrológicos. I. Marques, Guilherme, orient. II. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os dados fornecidos pelo(a) autor(a).

Iporã Brito Possantti

A synthesis on fundamental principles and problems in the use of hydrological models for planning watershed conservation

Thesis presented to the Graduate Program in Water Resources and Environmental Sanitation at the Federal University of Rio Grande do Sul, as a partial requirement for obtaining the degree of Doctor of Philosophy.

Approved in: Porto Alegre, December 6th of 2024

Prof. Dr. Guilherme Fernandes Marques
Universidade Federal do Rio Grande do Sul
Advisor

Prof. Dr. Manuel Augusto Pulido Velázquez
Universitat Politècnica de València
Examiner

Prof. Dr. Márcia Maria Guedes Alcoforado de Moraes
Universidade Federal de Pernambuco
Examiner

Prof. Dr. Walter Collischonn
Universidade Federal do Rio Grande do Sul
Examiner

Source code

This document was produced using LaTeX. The source code, as well as the figures and English versions, are available in updates maintained in the following repository: https://doi.org/10.5281/zenodo.14560036

To friendships

Acknowledgments

I am grateful to Professor and advisor Guilherme Fernandes Marques for the teachings, encouragement, partnership, and, above all, the trust placed in me since 2018. This has been a seven-year journey. We are already seeing some fruits ripen, but it is only the beginning.

I am grateful to my academic colleagues from GESPLA, the research group on water resources planning and management at the IPH. For the lunches and dinners at the University Restaurant, the jokes, and the company during focused work afternoons. Among the seasoned and new members I had the chance to work with, my thanks go especially, though not exclusively, to Ana Paula Dalcin, Ariane Sigallis, Giúlia Carrard, Gláucio Souza, Ivo Melo, Júlia Daiello, Juliano Finck, Julieta Nhampossa, Luísa Lucchese, Márcio Inada, Paola Kuele, Rossano Belladona and Vicente Lutz.

I also thank my graduate program colleagues at IPH who tolerated my frequent intrusions, asking what they were doing and how much they *truly* believed in it. To those responsible for the Water Science Student Seminar, who consistently elected me as the go-to speaker for all editions. I extend my gratitude, in particular, to those who formed the Casa dos Brothers in Aracaju: Bruno Abatti, Carlos Ferrari, Cléber Gama, Juliana Andrade, Lara Nonnemacher, Marina Fagundes, and Priscila Kipper (this is, finally, the IpoTese). I am also deeply thankful to my friend Rafael Barbedo, *The Veredas*, for his partnership and for sharing his knowledge about HAND during bohemian nights in Vienna, Berlin, and Bom Fim. To my brilliant friends from USP (who are not from IPH), I appreciate the friendship and invaluable exchanges in Vienna, Belo Horizonte, and Aracaju: Dimaghi Schwamback, Julien Sone, and Rodrigo Perdigão.

I am grateful to the network of mentors who guided me before graduate school. At age 19, in 2011, I began my scientific initiation, assisting the now-professor Maria Cristina de Almeida with bio-hydrogen experiments. At that time, the late professor Luiz Monteggia's guidance left a lasting impression of inventiveness and curiosity. During this period, I also interacted with Ayan Fleischmann, whom I thank for many fruitful exchanges, even during graduate school. I thank the late professor Dieter Wartchow for his inspiration in the pursuit of socially inclusive engineering. I am also grateful to professor Rualdo Menegat, who inspired me, especially with the Environmental Atlas of Porto Alegre, an encyclopedic and spectacularly illustrated work. Motivated by the Atlas, I try to invoke the spirit of Humboldt and Lovelock here. I thank professor Tatiana Silva, as well as colleagues from LABMODEL and FURG, for the practical experiences beyond university walls – and for transforming me from a future microbiologist into "the map guy". I am grateful to my last advisor, professor Fernando Dornelles, who brought me back to IPH at the end of my undergraduate studies in 2016.

I also feel indebted to the citizens of the Federative Republic of Brazil, whose taxes sustain the academic infrastructure that I was fortunate enough to benefit from and, in several projects, to be funded by. I am therefore thankful to the National Water and Basic Sanitation Agency, the Municipality of São Leopoldo, and the Rio Grande do Sul Sanitation Company for funding my scholarships during my doctorate. In particular, I am retroactively grateful to the administration of President Dilma Rousseff, who, in 2014, created the now-defunct Science Without Borders academic exchange program. Without the opportunity to study in Philadelphia, USA, the challenges I faced with the English language would have hindered my deep dive into the literature for this thesis. Even in global quarantine, without leaving home, I could explore knowledge thanks to the internet, the CAPES system, and proficiency in English.

I am grateful to Marcelo Kronbauer and the entire UNISC team for their friendship and for the opportunity to apply my studies in the Arroio Castelhano basin, even before finishing the thesis. We have much work ahead.

I thank the professors at IPH who accompanied me through the doctorate. Especially professor Walter Collischonn, for the enriching exchanges along this journey and for the opportunity to teach about TOPMODEL in the Hydrological Simulation course, even as a graduate student. I am also grateful to professor Rodrigo Paiva, for his insights and for helping me break protocol at the SBRH conference in Belo Horizonte by asking a live question to professor Keith Beven. I also thank professor Anderson Ruhoff, who showed me the path to real evapotranspiration estimates, and professor Nilza de Castro, for trusting me to create a case study with the Potiribu River. I apologize publicly for not managing to incorporate this study here in time, as an extraordinary flood crossed all our paths.

The largest recorded flood in Rio Grande do Sul, caused by global climate instabilities, substantially delayed my progress in this text. I am therefore grateful to friends who were present during this traumatic time: Guilherme, *The Kura*; Laura Azeredo; and José Augusto Müller Neto. I am also grateful to all professors, researchers, and journalists who together formed the information network during the disaster.

I dedicate this thesis to the network of friendships that became the center of my life during these years of graduate studies. I am deeply grateful to my friends for their support, companionship, and shared joy. Silly laughter, jokes, and hugs that lit up the dark days of the pandemic and continued to be the reason not to move to the mountains. I would like to mention all those who made me happier during this period, but it would be unfair – I would inevitably forget someone. Still, I especially thank Joana Winckler, Luiza Tonial, and Luisa Sarmento for making IpoFest happen. To Lucia Torres, Thomas Silveira, and João Paulo Niedererauer, for the pandemic companionship. To Clara Martinez, for everything, from Imbé to Barcelona. To Mariana Vivian, mainly for the virtual discussions with Karl Popper, but not only that. To Rafaela Machado, for guiding me through Berlin and Tapera. To Júlia Kuse, for the trails that happened and the trips that didn't. To Ananda Casanova and little Alice, for many things. To the caring friends Bettina Rubin, Matheus Schia, and Giovana Corsetti. To Mateus Coimbra and Santiago Costa, for our spaceship. To Luísa Acauan, Luciana Ruy, and Karina Kerne, for the ocean mornings.

With affection, I thank dear Fernanda Prestes for her companionship during my retreat from the world.

I thank my father, Genuir, for giving me critical thinking, and my mother, Maria Lucia, for giving me creativity. Together, they also gave me all the material and immaterial support to get here.

Finally, admitting my Animist side, I thank Cantagalo, the fig trees, the howler monkeys, the aracuãs, the saracuras, the sound of the stream and the wind, the vultures and hawks, the preás and snakes. Loba, Bela, and Duque. I thank the Sun, the Moon, Venus, Mars, Jupiter, and Saturn, with whom I conversed for almost two years during quarantine. I miss that greatly. I thank the Atlantic Ocean and its breath. At the mouth of the Tramandaí River, I thank the dolphins and orcas, especially Geraldona. Stay strong. At Zimba, I thank the right whales and turtles, the rock that separates Luz from Barrinha, which is the center of the Universe. To the Diva's cliff, thank you very much. I thank the Santa Marta Lighthouse, the Sambaquis, and the indigenous spirits with whom I had to make an agreement. But most importantly, I thank a certain wave that opened to the left, at Malvina, and made me see some things *above and beyond*.

Abstract

This thesis is a synthesis that organizes the conceptual and philosophical foundations for the use of hydrological models in planning watershed conservation and expansion of green infrastructure. With the aim of creating an integrated map of ideas, the work not only connects theoretical and practical foundations but also offers a structured guide for future investigations and applications in the field. The initial chapter presents the epistemological bases, addressing the justifications and limitations of hydrological models. The second chapter develops an ontological approach with Systems Dynamics, exploring how model architecture defines hydrological responses and identifying points for strategic intervention. In the third chapter, the thesis reviews the evolution of hydrological paradigms, culminating in the Theory of Connectivity, which proposes an unification of surface and subsurface flow processes. The final chapter explores the role of Ecological Economics in watershed management, using the `PLANS` model in Payment for Ecosystem Services (PES) schemes to prioritize conservation areas at the operational scale of the farms. By synthesizing these concepts, the thesis establishes a structured body of knowledge that facilitates the advancement of new research, enabling the scientific community to articulate sustainable and adaptive solutions for watershed management with greater clarity and objectivity.

keywords — Hydrological Modeling; Watershed Conservation; Payments for Ecosystem Services.

Contents

List of Figures

List of Tables

Acronyms

`CICES` Commom International Classfication of Ecosystem Services. 133, 134

`CN` *Curve Number*. 73–75, 82, 159

`ET` evapotranspiration. 49, 51, 52, 69

`GLUE` *Generalized Likelihood Uncertantiy Estimation*. 12, 60

`HAND` Height Above Nearest Drainage. 102, 103

`MDE` Digital Elevation Model. 98, 102

`MEA` Millenial Ecosystem Assessment. 132, 133

`NBS` Nature-based solutions. 101, 141

`PES` Payments for Environmental Services. 146, 147

`SCS` *Soil Conservation Service*. 70, 73–75, 82

`TEEB` The Economics of Ecosystems and Biodiversity. 135–137

`TWI` Topographic Wetness Index. 97, 98, 101, 103

Symbols

Chapter 1 – The role of theories and evidences

$\mathcal{L}$	Informal likelihood
μ	Mean
Ω	Space of possibilities
σ	Standard deviation
σ^2	Variance
Θ	Parameter vector
Υ	Input data vector
ε	Error
E	Evidence
H	Hypothesis
M	Model
n	Sample size
O	Observation
P	Probability
S	Statement
s^2	Sample variance

Chapter 2 – The workings of systems and models

Δt	Time step
μ_M	Mean of simulated points
μ_O	Mean of observed points
Ω_Θ	Parametric space
σ_M	Standard deviation of simulated points
σ_O	Standard deviation of observed points
KGE	Kling-Gupta efficiency
MAE	Mean absolute error
RMSE	Root mean square error
R^2	Coefficient of determination
Θ	Set of parameters
Υ	Exogenous variables, input data
E	Evapotranspiration

I	Input material flow into a compartment
k	Mean residence time
O	Output material flow from a compartment
P	Precipitation, rain
Q	Runoff, discharge
R	Rapid surface runoff
r	Correlation coefficient
S	State or level of a compartment
$s_{\max}$	Storage capacity
s_a	Activation level
s_c	Connectivity level
$y_{M,i}$	Simulated point (value)
$y_{O,i}$	Observed point (value)

Chapter 3 – The understanding of hydrological processes

α	Drainage area per unit contour
β	Slope of the terrain
Δz	Difference in hydrostatic potential (Darcy's Law)
κ	Lateral hydraulic conductivity, transmissivity
$\kappa_{\max}$	Maximum lateral hydraulic conductivity
λ	Topographic wetness index
$\nabla\Phi$	Hydrostatic potential gradient (Darcy's Law)
ν	Hydrogram volume
ω	Scaling factor
$\mathbf{C}$	Vegetation canopy
$\mathbf{D}$	Gravitational deficit
$\mathbf{D}_{\mathrm{v}}$	Capillary deficit
$\mathbf{G}$	Phreatic zone
$\mathbf{O}$	Organic horizon
$\mathbf{S}$	Soil surface
$\mathbf{V}$	Vadose zone
HT	TWI enhanced by HAND
H	HAND - Height Above the Nearest Drainage
$\mathrm{H}_{\max}$	Upper threshold of HAND
T	TWI - Topographic Wetness Index
$\mathrm{T}_{\max}$	Upper threshold of TWI
V_{c}	Capillary water in the vadose zone
V_{g}	Gravitational water in the vadose zone

$\tilde{\mathrm{H}}$ Normalized HAND

$\tilde{\mathrm{T}}$ Normalized TWI

A Cross-sectional area of the pipe (Darcy's Law)

c_{max} Interception capacity

e_{c} Evaporation in the canopy

e_{g} Transpiration in the phreatic zone

e_{o} Transpiration in the organic horizon

e_{pot} Potential evapotranspiration

e_{s} Evaporation at the surface

e_{v} Transpiration in the vadose zone

f Infiltration

f_{max} Infiltration capacity

g Mean residence time of the aquifer

H_w Dominance factor of HAND

H_α Threshold for drainage initiation

K Hydraulic conductivity

k Mean detention time of linear reservoir

l Length of the pipe (Darcy's Law)

m Vertical uniformity of the soil

n Effective number of linear reservoirs

o_{max} Field capacity of the organic horizon

P Precipitation, rain

P_{s} Effective rain

P_{x} Excess rain

Q Runoff, flow, discharge

Q_{gt} Translational runoff

$q_{\mathrm{g},i}$ Lateral base flow per unit contour

$Q_{\mathrm{g,max}}$ Aquifer production capacity

Q_{g} Base flow

q_{o} Percolation between horizons

q_{se} Direct rain, excess saturation runoff

q_{si} Runoff, surface flow

q_{ss} Exfiltration, subsurface runoff

q_{v} Recharge

s_{max} Surface detention capacity

u Darcy velocity (Darcy's Law)

v_{max} Field capacity of the mineral horizon

Our mistake is not that we take our theories too seriously, but that we do not take them seriously *enough.*

Steven Weinberg

The rise of artificial intelligence might push most humans out of the job market (...) So if you want to study something that will guarantee a good job in the future, maybe philosophy is not such a bad gamble

Yuval Harari

Chapter 0

Introduction

— Iporã, think about the problem first, not the tool.

This is what Professor Guilherme Marques told me on a hot December afternoon in 2017 at the Hydraulic Research Institute of the Federal University of Rio Grande do Sul, in Porto Alegre. I had been accepted into the academic master's selection in the graduate program, establishing first contacts with the researcher who would guide me for the next two years, which, with the PhD, turned into seven. "*We'll consider the tool later; first, we need to deeply understand the relevant research problem questions.*", he continued, drinking coffee and explaining that choosing the method before the problem would imprison us irreparably. If necessary, he encouraged me, I would have to invent an innovative method to obtain an innovative solution. But first, I should formulate the questions well. I liked that.

With this guidance in mind, I went to Brasília for the World Water Forum in March 2018. At the event, the United Nations agenda (2018) [1] motivated me to study nature-based solutions for water resource management. Upon returning to Porto Alegre, I decided to investigate how to expand green infrastructure to improve urban water security. Faced with future climate pressures and water demand, numerous questions arose: what to do? when? at what cost? is it worth it? where to invest? These questions became even more pertinent following institutional initiatives such as the National Policy for Payment for Environmental Services (2021) [2] and the National Watershed Revitalization Program. The impacts of climate change, evident in 2024, also underscore the urgency of rigorously grounded adaptive strategies.

These questions led me to conceive and program a hydrological model, which I named `PLANS`. That is, rather than relying on generic, readily usable models, the problem forced me to design a customized tool. The first results of this journey were published in Possantti & Marques (2022) [3], where we evaluated the cost-benefit and timeline of actions for scenarios in the Sinos River watershed. In this case, the main need for a flexible model was coupling it to an optimization algorithm, Dynamic Programming. As thousands of simulations were required, I configured the model as a routine in the same programming language as the algorithm.

From then on, we moved on to the problem of spatial allocation: *where* to act? We studied the Arroio Castelhano basin in Venâncio Aires, where a pilot project for Payments for Environmental Services is underway. To generate the priority areas map, published in Possantti *et al.* (2023) [4], I delved into the spatial representation of hydrological processes on hillslopes, which led me to philosophical reflections on modeling. To obtain results on an operational scale, a sufficiently detailed model was needed. However, physically-based models, although useful for simulating vector fields of water velocity, are computationally prohibitive and have uncertainties similar to semi-distributed models. Thus, I concluded that the principles of `TOPMODEL` were more effective in addressing the spatial allocation problem, leading to programming a new version of the `PLANS` model.

Over time, thinking about the problem first became a complex task. I ended up opening many doors along the problem path, in an almost endless exploration. The qualification board correctly pointed out that I was "moving backward". In the case of modeling, theoretical deepening brought to light inevitable problems, such as empirical uncertainty, equifinality, and scale. Similar dilemmas arose in management, as setting priorities requires an economic principle, such as potential additionality in watersheds. Similarly, assessing the feasibility of investments in green infrastructure requires a system to value hydrological processes. Quickly, I encountered more theoretical questions, which Philosophy articulates better than Science.

The thesis presented is the final expression of this problem exploration. The following text provides a **synthesis** of fundamental principles and problems in the use of hydrological models for watershed conservation planning. The aim of the monograph is to establish conceptual connections to support the use of these models in this context. Unlike **analysis**, which performs conceptual breakdown, synthesis gathers distinct elements and organizes them into a coherent whole. Often, my colleagues focus on *analyzing* specific scientific questions in their theses, responding to them with specialized methods and strategies. This intellectual movement is essential for local depth, but someone eventually needs to make the opposite effort, unifying analytical approaches into a coherent global view. As we will see, Science is not just about fitting pieces into a puzzle but having a vision of the final image that the completed puzzle will reveal.

It is worth noting that synthesis work, especially when aimed at foundational aspects, is increasingly overlooked due to academic pressure for peer-reviewed publications that demand novel results. This can lead to hydrological model usage without a critical sense of their fundamental premises. Lieke Melsen (2022) [5], for instance, interviewed 14 water resources researchers and found that the main reason for choosing hydrological models is the influence of more experienced colleagues in the group. There is an evident efficiency in following colleagues' work, but this highlights the importance of synthesis work that revises foundations, preventing this continuity from becoming mere imitation. With the advent of Artificial Intelligence and language models, generalized imitation presents increasing challenges for truly human knowledge production. Thus, I suggest we draw inspiration from works like Keith Beven (2002) [6] in Hydrology and Herman Daly (2015) [7] in Economics, authors who seek greater coherence by making explicit underlying philosophical questions.

The innovative result of this synthesis, therefore, is a **map of ideas** articulated in four chapters, arranged in an ascending order of practical application and integration of systematized concepts. In addition to the chapters, a Glossary was created with the same spirit, serving as a map. Thus, the synthesis results in a structure that facilitates a comprehensive and integrated understanding of issues related to using hydrological models in integrated water resource management, especially in designing strategies to expand green infrastructure in zero-order basins. It is hoped that the scientific community will move firmly forward to articulate organized concepts on both pure and applied fronts, seeking new solutions, filling gaps, and revealing new problems and flaws that may eventually demand a complete reformulation of established foundations.

The first chapter emphasizes the epistemological foundations of model use, including hydrological models, which aim to convey theories about reality. The focus is on the philosophy of **instrumentalism**, which recognizes uncertainties when trying to "capture" reality with precise mathematical theories. The role of empirical evidence is explored from various perspectives, as well as the importance of paradigms in theory construction. The practical implication is that we can adopt a rejection criterion for hydrological models, the **encapsulation test** by the observational uncertainty band. How to conduct this test, covering from rating curve uncertainties to future scenarios, are open paths for investigation.

The second chapter addresses model ontology, developing a transition between theoretical and practical themes. The central point is to introduce **System Dynamics**, an approach that creates models from a network of reservoirs connected by input and output flows. Although these "building blocks" are simple, I illustrate how level and flow arrangements can quickly generate complex behaviors. The chapter also systematizes **model diagnostic** techniques, assessing their adequacy in various aspects. Many avenues for future research are opened, including agent-based models, exploratory modeling, sensitivity analysis, operational research, and crucially, the reproducibility problem — the difficulty of using models by others besides their creators.

The third chapter delves into the field of Hydrology, reviewing the evolution of this Science throughout the 20th century, marked by successive paradigm shifts precipitated by empirical evidence from experimental watersheds. The essential message is that, for watershed conservation, hydrological processes must be assessed on the hillslope scale, called **zero-order basins**. It is highlighted that hillslope runoff is part of a range of hydrological responses — fast and slow, underground and surface — that vary according to topography, soil, vegetation, climate, and season. It demonstrates that the difficulty in reconciling this complexity with hydrological models has generated debates and uncertainties, such as the problems of **equifinality** and **scale**. New promising paths emerge with connectivity theory, a recently proposed unifying paradigm.

Finally, the fourth chapter reaches applied watershed management, preceded by a theoretical exposition of the economic and ethical principles guiding decisions in this field. The chapter highlights the role of **Ecological Economics** in promoting a paradigm shift that sees watersheds as natural capital, providers of natural watershed services. Thus, managing watershed areas with economic instruments, such as Payment for Environmental Services (PES) schemes, is interesting to ensure water security for water users. In this context, the `PLANS` model presents itself as an adequate tool to estimate **potential additionality** at the farm level and prioritize conservation areas, considering uncertainties. New paths emerge, especially for evaluating trade-offs and synergies with other natural services and integrating the benefits of gray infrastructure.

On February 16, 2023, at 3:23 PM, two black swans were observed in the park behind the Palácio das Laranjeiras, Rio de Janeiro, Brazil. A single black swan would already be sufficient to prove that *not all swans are white*.

Chapter 1

The role of theories and evidences

> I had long observed that, with regard to customs, it is sometimes necessary to follow opinions that we know to be uncertain as if they were indubitable; but, as I then wished to concern myself solely with the search for truth, I thought it was necessary to do the opposite, and reject as absolutely false everything in which I could imagine the slightest doubt, in order to see if anything would remain in my beliefs that was entirely indubitable.
>
> René Descartes, *Discourse on the method*, p. 15 [8]

> It is clearly possible to develop and use environmental models without any underlying philosophy. Many practitioners do so, although most perhaps ultimately want to develop and use models that are *as realistic as possible*, given the constraints of current knowledge, computational capabilities, and observational technologies.
>
> Keith Beven, (2002, p. 2465) [6]

1.1 Electronic circuits

Simulating a hydrological model consists of applying electrical voltages to electronic circuits. This is *literally* what happens during a simulation. The manner and order in which the voltages are applied directly correspond to the instructions provided to the central processing unit (CPU) of a machine, usually a digital computer. In this case, we provide the operating system with code in a high-level language (such as `Fortran` or `Python`), which is interpreted into a lower-level version, processable by the machine. Thus, all information, including data and instructions, is converted into binary digits

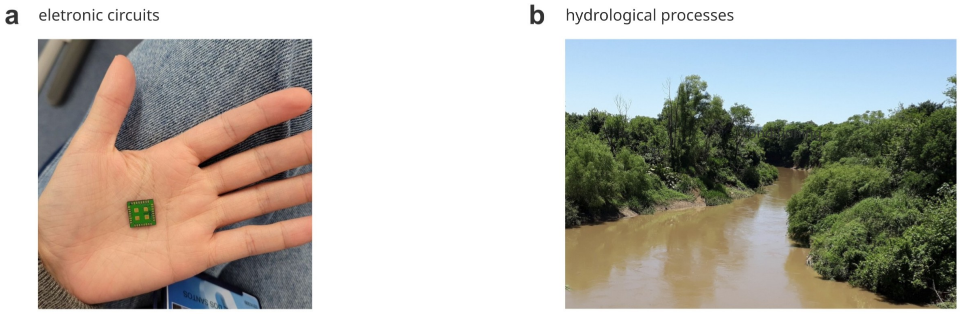

Figure 1.1 — From electronic circuits to hydrological processes. a. — Applying a hydrological model consists of *literally* applying voltages to digital electronic circuits. The simulation results refer literally and objectively to the processing of the binary states of transistors. **b.** — Hydrological model users generally accept that the electronic results are a *realistic representation*, albeit approximate, of various hydrological processes, whether it be rainfall infiltration, surface runoff q_{si}, or river discharge. How is this possible? The photograph in (**a**) was kindly provided by electrical engineer Karina Kerne; the photograph in (**b**) was taken by the author on the bridge over the Pardinho River south of Lago Dourado, Santa Cruz do Sul, Rio Grande do Sul, Brazil.

(*bits*) stored by states of electronic circuits called **transistors**. The processing, in turn, happens through logic gates that perform *boolean* operations of conjunction $\wedge$, disjunction $\vee$, and negation $\neg$ on the bits. All the graphs, maps, and animations we see after a simulation objectively refer to the recorded states of sections of the machine's digital memory—numbers represented in binary form by transistors. That said, how is it possible that binary patterns in digital circuits have *anything to do* with rainfall, river runoff, or soil saturation?

As highlighted by Keith Beven in the epigraph of this chapter, it is rare for applications of hydrological models[1] to ground this issue with an *explicit philosophy* [6]. Even so, model users generally believe that simulation results provide a *realistic representation* of processes and phenomena that indeed exist in the world, not just in electronic circuits, such as soil erosion, river runoff, plant transpiration, etc. Of course, no one thinks that hydrological model simulations provide an *exact* description of reality, but it is generally accepted that they offer an *approximate* description of reality and that this approximation can be improved as new computing and observation technologies become available. This belief is further accentuated when simulation results are directed to assist in important decision-making related to water resource management in State organizations or private companies. If material and human resources are allocated based on these results, they had better agree with reality!

Beven calls this common implicit philosophy among hydrological model users **pragmatic realism**. He classifies this stance as naïve, demonstrating that the much-desired realistic representation, when taken seriously, becomes an extraordinary goal in light of the **empirical uncertainty** that exists in hydrological modeling. If water resource management genuinely wishes to build **evidence-based policies**, scientific advice must adopt a *critical stance* toward the use of models and their results. As highlighted by Ongaro and Andreoletti, uncertainty is a pervasive attribute in the decision-making process [9]. In this context, they argue that the proper role of scientific authorities is to inform about empirical uncertainties so that other interest groups can deliberate on the ethical and political uncertainties at play. However, this critical stance is only possible from an explicit philosophical perspective, referred to here as **instrumentalism**, which brings to the fore several problems that need to be addressed before affirming the correspondence between electromagnetic fields in electronic circuits

[1] Although Beven actually refers to environmental models in general, including atmospheric and geochemical models, I will here limit the discussion to hydrological models.

and hydrological processes in watersheds.

In this spirit, the aim of this chapter is to establish the instrumentalist philosophical foundations that will be explicitly used throughout this thesis. I will articulate the following issues from the Philosophy of Science: the problem of justification, Bayesian epistemology, the falsifiability of theories, the concept of paradigms, and the underdetermination problem. The ontology of models, which is also a relevant topic, will be addressed in the next chapter. For the purposes of this chapter, one must consider that **a model is a symbolic vehicle of a theory**. The intention is not to exhaust the presented topics (as this would require writing many other theses!). On the contrary, this chapter should be understood as a panoramic view, as if we were standing atop a mountain. This analogy is particularly useful, as the summit provides a good understanding of the landscape stretching below our feet. At the same time, it is a hostile territory. The air is thin, and we feel dizzy. Movement is complicated, possible only through labyrinths of trails surrounded by cliffs and caves. Care must be taken not to get lost and never return.

1.2 The problem of justification

Epistemology is the branch of Philosophy that investigates the nature of human knowledge itself. The epistemological question is: *how is it possible to know something*? From this perspective, a **theory** is a *universal statement* that provides a definitive explanation about a particular phenomenon. Thus, creating a theory is relatively easy. Someone might profess, for example, the theory that forest fairies are responsible for making household objects disappear, especially socks. We have a phenomenon (the disappearance of things) and a definitive explanation (the fairies). The difficult part, however, is *justifying the truth* of a theory—the so-called **problem of justification**. Here, the epistemological question becomes a bit more complicated: *how is it possible to know something* ***true***? In the given example, if someone claims that, in fact, the dog is responsible for making the socks disappear, how can one defend the fairy theory? How do we separate the true from the false? The example of forest fairies might seem ridiculous at first glance. But just a few centuries ago, people were tortured and burned alive in public squares over issues like this (especially women accused of being witches). Even today, in fact, it is a very serious issue, with significant social and political implications. According to Daniel Kahneman, research in the field of psychology shows that humans exhibit numerous **cognitive biases** that compromise their ability to distinguish the true from the false [10].

The epistemological problem of justifying theories was assimilated by two different philosophical schools of thought during the birth of the **scientific method** in modernity [11]. One of these schools, **rationalism**, holds that *Logic* must be the justification for the truth of theories. After all, it is clear that an incoherent explanation cannot be true. This position is commonly associated with Descartes (1596-1650), Spinoza (1632-1677), and Leibniz (1646-1716) — the so-called continental rationalists. René Descartes, one of the greatest rationalists of his time, propagated the notion that the senses are deceptive and elected reason as the guide for his opinions, establishing a program he called the *method of doubt*. In his *Discourse on the Method*[2], he went so far as to doubt the existence of everything, claiming that reality could be indistinguishable from a mere dream [8]. With that, he concluded that the simple act of *doubting* existence guarantees at least one absolute certainty: the existence of oneself[3]. The other philo-

[2]Descartes delves deeper into his ideas in *Meditations*.

[3]Descartes' idea that "I think, therefore I am" (*cogito, ergo sum*) guarantees the existence of a *Self*

sophical school, **empiricism**, argues that a theory is justified by *empirical evidence*, that is, through a collection of direct observations of events and phenomena. This line of thought is associated with philosophers Locke (1632-1704), Hume (1711-1776), and Reid (1710-1796), the so-called British empiricists. Locke, for instance, became known for spreading the idea that all people are born equal, like a blank slate, a *tabula rasa*, and acquire knowledge through experience, by interacting with their environment [12].

Although not exactly opposing, what these philosophical schools favor, in essence, are different methods of *inference*. While rationalists prefer **deductive inference**, empiricists prefer **inductive inference**. In the first case, statements are justified when they logically follow from their premises, that is, they are deduced [13]. For example, considering the premises that "*all Gauchos like mate*" and that "*Clara is a Gaucho*", then we deduce that "*Clara likes mate*". Typically, the process of deductive inference follows a structure in which a conditional statement (if S_1 is true, then S_2 is also true) is followed by an *antecedent* sentence, which can be either an affirmation (*modus ponens*: S_1) or a negation (*modus tollens*: $\neg S_1$), leading to the conclusion of the *consequent* sentence. In the affirmative form (*modus ponens*):

$$\begin{array}{rl} S_1 \implies S_2 & \text{if someone is from the Pampas, then that person likes drinking mate} \\ S_1 & \text{Clara is from the Pampas} \\ \therefore\ S_2 & \text{therefore, Clara likes drinking mate} \end{array}$$

deductive inference guarantees the truth of the consequent sentence as long as its antecedent premises are true. In the affirmative mode illustrated, it deduces *singular statements* from *universal statements*. inductive inference, on the other hand, goes in the opposite direction. Empiricists seek to build theories (universal statements) by *generalizing* from evidence (singular statements) obtained through empirical experience. This gives rise to the notion of a **hypothesis**: a draft of a theory to be *confirmed* or *verified* by evidence. For example, if we observe that the Gauchos we know like to drink mate, we can infer, by induction, that *all* Gauchos must like mate. With a bit more caution and empirical rigor, we might eventually state that the *probability* of any given Gaucho liking mate is high, since 99% of respondents in a survey of a thousand Gauchos said that yes, they like to drink mate. For empiricists, the observation of phenomena justifies the construction of generalizations that are plausible or *probable*.

Both lines of reasoning are problematic. The strict use of Logic as a justification for theories produces the **infinite regress problem** [14]: if a statement S_1 is justified by another statement S_2 ($S_2 \implies S_1$), what justifies S_2? Perhaps statement S_3 justifies S_2, and statement S_4 justifies S_3, and so on, *ad infinitum* ($\infty \implies ... \implies S3 \implies S_2 \implies S_1$). An alternative is to establish a circular chain of statements ($S_1 \implies S_2 \implies S_3 \implies S_1$), but this is, logically speaking, even worse than the previous situation, as ultimately the statements end up self-justifying. Infinite regress is easily noticed when children, being naturally curious, learn to ask "*why?*". One can stop the regress (or at least *stop the questions*) by establishing fundamental truths, unquestionable axioms, as happens in Mathematics. But in Science, this only produces *dogmatic* theories based on relatively arbitrary conventions, which are easy targets for empiricist criticism. On the other hand, using empirical experience as a justification for theories brings with it an insoluble problem, described by David Hume (1711-1776) and known today as the **induction problem**. inductive inference is fragile because

that asks questions. Descartes uses this to attempt to prove the existence of God as well. However, Descartes' proof of God is questionable, which can easily lead us to **solipsism**: the thesis that the *Self* is alone, floating in the void, imagining a reality that fundamentally does not exist. After all, how can *you*, the one reading this text, be certain that *I*, the author, exist? How can you be sure that all your memories were not just created right now, at this very instant? And what if your entire life is nothing but an immaterial hallucination?

it is based on the **principle of uniformity**: the assumption that the regularities of nature observed in the past will be the same as those observed in the future. This assumption is the foundation of all Physics, after all, the laws observed today are often used to predict future events and reconstruct past events. Despite being intuitive, it is not possible to rationally justify the principle of uniformity. If we invoke the fact that it has always proven functional or correct, we fall into a circular, logically invalid argument that *evokes the very principle of uniformity to justify the principle of uniformity* [15]. In other words, one cannot defend inductive inference through an inductive argument! Thus, the principle of uniformity is ultimately the fundamental dogma of the empiricists[4].

The progress of modernity gradually reduced the intellectual rivalry between thinkers from the European continent and the British Isles. empiricism, although revised and moderated, triumphed over rationalism — especially after the work of Immanuel Kant (1724-1804). In his *Critique of Pure Reason* (1781), this thinker proposes a synthesis between rationalism and empiricism [16], [17]. In it, Kant agrees with the empiricists — that empirical experience justifies knowledge — but makes a concession to rationalism, establishing that theoretical concepts about the objects to be known are necessary *a priori*. Without what he called **transcendental categories**, we cannot do much with the perceptions we collect about the external world. The hegemony of empiricism in modernity culminated in the early 20th century with the philosophical movement of **logical positivism**. This movement established the common sense that scientific theories are *verified* or *confirmed* through the collection of observations and statistical analyses. However, this conception was surpassed in the mid-20th century, mainly due to the remarkable transformations in Physics. This reignited the debate about how Science justifies its theories and changes them in the long term, with new perspectives proposed by Karl Popper (1902-1994) and Thomas Kuhn (1922-1996), creating the conditions for the contemporary debate in the Philosophy of Science: scientific realism and the purpose of Science.

1.3 The process of confirmation

In the previous section, we saw that the empiricist school holds that a hypothesis represents a draft of a theory that must be *confirmed* by empirical experience. In this sense, the ideas of the philosopher and statistician Thomas Bayes (1701-1761) provide a particularly useful method of inductive inference for the confirmation of hypotheses through the mathematics of probabilities [18], [19]. The central idea of the so-called **Bayesian epistemology** professes that knowledge is not an all-or-nothing matter, black or white, but presents subtleties, with various shades of gray between true and false. The reason for this stems from the recognition that empirical observations are inevitably subject to *random noise*, meaning there is **statistical uncertainty** in the sampled data. For this empiricist approach, the subtleties of knowledge consist in the **credence** in the truth of a hypothesis. This credence should be updated as favorable or unfavorable evidence is obtained through empirical experience.

Before proceeding, it will be useful to establish an intuitive example.

[4] Hume's ideas inevitably lead us to **skepticism**. Beyond the induction problem, Hume argued that **causality** between phenomena is an imaginary concept, never *directly* experienced in reality. Imagine a billiard ball colliding with another ball at rest. Then, the ball at rest begins to move as well. Newton would say that this happened *because of* the law of conservation of motion. But this law requires several imagined theoretical concepts, such as mass and energy. What we can actually experience are events happening one after the other, but never the cause-and-effect connection between them. For Hume, *this connection does not exist.*

Consider that you are at the airport of a distant country, in the food court of a busy international terminal. You want to know if the person at the table in front of you will take the same flight as you to Brazil. With no information available, your credence in this hypothesis is relatively low, as there are dozens of flights scheduled from this terminal to various other countries. But if you suspect that the person is Brazilian, your credence increases — after all, your flight is likely full of other Brazilians. At the same time, caution is needed, as being Brazilian does not necessarily imply traveling to Brazil. Even if the person were not Brazilian, there would still be a small chance that they are traveling to Brazil on the same flight as you, for tourism or business. One thing is certain: *favorable evidence that the person is Brazilian will lead you to update your degree of conviction.* Bayesian epistemology articulates how to do this in a mathematically precise way.

A good starting point to formalize the concept of credence in terms of probabilities is to note that a given hypothesis H and its favorable evidence E are distributed in a **space of possibilities** Ω. From a Bayesian perspective, each of these possibilities is assigned a credence, which can be considered probabilistic as long as they are *mutually exclusive* (they cannot both be true at the same time) and *jointly exhaustive* (at least one of them is true). Given these conditions, the **principle of probabilism** is observed based on the following axioms:

- **Non-negativity**. The probability of a given possibility A must be a non-negative real number: $P(A) \geq 0$;
- **Normalization**. The probabilities of all possibilities must sum to one: $P(\Omega) = 1$, and;
- **Additivity**. Because they are incompatible, the probability of two possibilities A and B is the sum of their individual probabilities: $P(A \cup B) = P(A) + P(B)$.

In the airport example, there are four possibilities: the person is Brazilian and on the same flight (E is true and H is true); the person is Brazilian and not on the same flight (E is true and H is false); the person is not Brazilian and on the same flight (E is false and H is true); and the person is not Brazilian and not on the same flight (E is false and H is false). This space of possibilities can be more easily visualized through a table with illustrative numbers, as in Table 1.1. So, consider that in the international terminal, there are 200 scheduled flights (only one of them to Brazil) and that each plane carries 500 passengers. In total, there are 100,000 passengers circulating in the terminal. Additionally, there are 800 Brazilians in the terminal, with 450 on your flight and another 350 with other destinations.

	same flight (H)	**different flight ($\neg H$)**	totals
is Brazilian (E)	450	350	800
is not Brazilian ($\neg E$)	50	99150	99200
totals	500	99500	100000

Table 1.1: Possibility space for the airport example. The numbers represent the distribution of passengers in the different combinations between hypothesis H (on the same flight) and evidence E (nationality is Brazilian)

By inspecting the table, we can easily see that the probability of your hypothesis being true, that any given passenger is on your flight, is $P(H) = 0.005$ (0.5%), as $500/100000 = 0.005$. In Bayesian jargon, this is the **prior**. However, the evidence that this person might be Brazilian changes everything. Now, the probability of your hypothesis being true given that the evidence is true is $P(H|E) = 0.56$ (56%), as $450/800 = 0.56$. In Bayesian jargon, this is the **posterior**. The **Bayes' theorem**

Hypothesis *i*	$P(H_i)$	$P(E\|H_i)$	$P(H_i) \cdot P(E\|H_i)$	$P(H_i\|E)$
H_1: $X < x_1$	0.143	0.008	0.001	0.008
H_2: $x_1 \leq X < x_2$	0.143	0.026	0.004	0.026
H_3: $x_2 \leq X < x_3$	0.143	0.084	0.012	0.084
H_4: $x_3 \leq X < x_4$	0.143	0.171	0.024	0.171
H_5: $x_4 \leq X < x_5$	0.143	0.208	0.03	0.208
H_6: $x_5 \leq X < x_6$	0.143	0.275	0.039	0.275
H_7: $X \geq x_6$	0.143	0.227	0.032	0.227
totals	1.0	1.0	0.14	1.0

Table 1.2: Illustrative example of the conditioning of the probability distribution of a random variable X. In this case, the prior distribution $P(H)$ was defined as uniform by observing the principle of indifference (objective Bayesianism).

formalizes this reasoning:

$$P(H|E) = \frac{P(H) \cdot P(E|H)}{P(E)} \tag{1.1}$$

Where $P(H|E)$ is the probability of hypothesis H being true given that evidence E is true (posterior probability); $P(H)$ is the probability of hypothesis H being true (prior probability); $P(E|H)$ is the probability of evidence E being true given that hypothesis H is true, called the **likelihood**, and $P(E)$ is the probability of the evidence being true under all possibilities. In the case of the airport:

$$P(H|E) = \frac{\frac{500}{100000} \cdot \frac{450}{500}}{\frac{800}{100000}} = \frac{450}{800} = 0.56\,(56\%)$$

What Bayes' Theorem tells us is that the estimate of the probability of the hypothesis being true given favorable evidence must take into account both the space of possibilities reduced *by the favorable evidence* (in this case, the fact that there are few Brazilians in the international terminal) and the chance *that the evidence does not guarantee the hypothesis* (in this case, the fact that not everyone on your flight is Brazilian). Or simply:

$$\text{Posterior} = \text{Prior} \times \text{Likelihood} \div \text{Evidence}$$

The update of the credence is only possible *provided that* the evidence favorable to the hypothesis is true. For this reason, this is called the **principle of conditionalization** of Bayesian epistemology. In the airport example, everything is quite clear because the absolute numbers of passengers were provided in the table. The probabilities calculated this way are objective, based on the *frequency* of each set in the space of possibilities Ω. But in a practical situation, we usually only have access to the degrees of conviction in the space of possibilities Ω, which must change as new evidence is obtained. To remain consistent with the axioms of probabilism, conditioning or **conditioning** [5] with the evidence must *zero out*, *rescale*, and *normalize* the probability values. In the airport example, when we find out that the person in the airport is definitely Brazilian, we must zero out the probabilities that they are not Brazilian, which implies proportionally rescaling the probability of the other possibilities and *normalizing their values* so that their sum equals one, as $P(\Omega) = 1$. On the other hand, if we only infer that the person at the table ahead speaks Portuguese (perhaps by reading a book), this is not enough to zero out the probability that they are not Brazilian (after all, other nationalities also speak Portuguese!).

The principle of conditionalization allows for the application of Bayes' Theorem to a finite number of N hypotheses H_1, ..., H_N, as long as they are mutually exclusive

[5] The terms *conditioning* and *conditionalization* will be used equivalently here.

Hypothesis *i*	$P(H_i)$	$P(E\|H_i)$	$P(H_i) \cdot P(E\|H_i)$	$P(H_i\|E)$
H_1: $X < x_1$	0.227	0.008	0.002	0.022
H_2: $x_1 \leq X < x_2$	0.273	0.026	0.007	0.088
H_3: $x_2 \leq X < x_3$	0.229	0.084	0.019	0.236
H_4: $x_3 \leq X < x_4$	0.146	0.171	0.025	0.306
H_5: $x_4 \leq X < x_5$	0.082	0.208	0.017	0.208
H_6: $x_5 \leq X < x_6$	0.036	0.275	0.01	0.12
H_7: $X \geq x_6$	0.008	0.227	0.002	0.022
totals	1.0	1.0	0.08	1.0

Table 1.3: Another illustrative example of conditioning of the probability distribution of a random variable X. In this case, the prior distribution $P(H)$ was defined subjectively, by expert opinion (subjective Bayesianism).

and collectively exhaustive possibilities. In this case, Bayes' Theorem takes the following form:

$$P(H_i|E) = \frac{P(H_i) \cdot P(E|H_i)}{\sum_{j=1}^{N} P(E|H_j) \cdot P(H_j)} \tag{1.2}$$

The denominator in this equation is a constant that plays the role of normalizing the probability values, ensuring that $P(\Omega) = 1$. Here, it is noted that normalization provides an opportunity to assign *non-probabilistic* degrees of conviction to the likelihood $P(E|H)$ without violating the principle of probabilism, as long as they are non-negative values. In other words, the likelihood $P(E|H)$ can be interpreted as a *weight* given by the evidence. If this is the case, the notation for likelihood should be modified to an **informal likelihood function**, denoted by $\mathcal{L}(E|H)$[6]. Either way, the posterior distribution of a random variable must be conditioned by an empirically estimated probability (or weight) distribution. To do this, it is necessary to discretize the values of the variable into N intervals, which are the hypotheses H_1, ..., H_N in Equation (1.2). For each hypothesis H_i, its prior $P(H_i)$ must first be assigned. Next, empirical evidence must be obtained to derive the likelihood $P(E|H_i)$ for each hypothesis. The posterior of each hypothesis is then conditioned through the application of Equation (1.2). Finally, if we consider the posterior distribution obtained as the prior distribution for the next stage, the confirmation of the probability distribution of the variable occurs incrementally as new evidence accumulates. Table 1.2 shows an example of the conditioning of the probability distribution of a random variable X, which in this case was divided into seven discrete intervals defined by thresholds $x_1, ..., x_6$.

A controversial issue in the conditioning process is the definition of prior probabilities in the very first stage when there is still no evidence. In the case of Table 1.2, what justifies the prior distribution $P(H_i)$ being uniform? This is the so-called **problem of priors** in Bayesian epistemology, which divides the field into two main branches of Bayesianism: the *subjective* and the *objective*. To proceed, one must decide between the two. On the one hand, the **subjective Bayesianism** branch accepts any prior distribution as long as it does not violate the principle of probabilism. The obvious problem here is that the posterior distribution may become more sensitive to the values of the prior distribution than to the evidence itself, as demonstrated in Table 1.3. Figure 1.2 clearly shows the difference between the posterior results $P(H|E)$ between Tables 1.2 and 1.3, even though exactly the same likelihood distribution $P(E|H)$ was used for the conditioning. Defenders of this branch argue that it is legitimate for different subjective opinions to be compared objectively. Moreover, if the evidence is consistent, divergent opinions about the prior distribution become irrelevant in the long run, dissipating during successive conditioning stages. On the other hand, the **objective Bayesianism**

[6]This is the basis of the *Generalized Likelihood Uncertantiy Estimation* (GLUE) method for uncertainty analyses, introduced in the next chapter

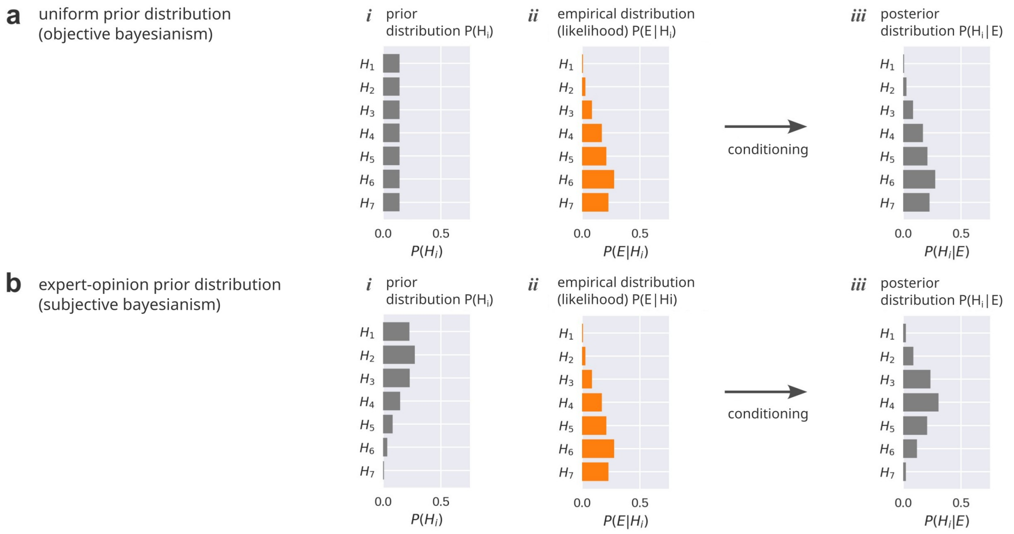

Figure 1.2 — Conditioning the prior distribution of a random variable. This visually presents the illustrative data from Tables 1.2 and 1.3. In both cases, the evidence is the same; what changes are the assumptions used in the prior distribution. **a** — Uniform distribution, considering the *principle of indifference* of objective Bayesianism. **b** — Distribution defined by opinion, a practice considered valid in subjective Bayesianism.

branch prefers to eliminate any bias when defining the prior distribution. To this end, the **principle of indifference** is recommended: the credence in two or more hypotheses should be equal in the absence of sufficient reasons to the contrary. In the case of Table 1.2, there may not be enough reasons to differentiate the values of $P(H_i)$, which is why a uniform distribution was adopted. This principle may seem natural, but it is nothing more than an arbitrary convention: in the face of complete ignorance, any prior distribution is equally probable, with the uniform distribution being, in fact, an extremely specific case.

Another open issue in the conditioning process is how to obtain the likelihood $P(E|H)$ from the evidence. After all, without precise details about the entire space of possibilities Ω, as in the airport example, the evidence usually consists of sampled data that does not automatically translate into probabilities. This issue becomes even more pronounced in the case of a continuous random variable when we seek to estimate a probability distribution from the available data. Similar to the problem of priors, the solution to this question requires a decision-making process, which involves establishing a set of assumptions about *the behavior of random noise*[7]. In this direction, a particularly useful example for our future discussion on hydrological models is the fitting of mathematical curves that relate two phenomena, such as in the case of the **rating curve** that describes discharge as a function of the observed level at a river section (see Highlight 1.3). In this case, the mathematical curve represents the hypothesis, or **model**, for which we are interested in knowing the credence. A general relationship for this problem is as follows [20]:

$$O(x,t) = M(x,t,\Theta) + \varepsilon(x,t) \tag{1.3}$$

In which $O(x,t)$ is the empirical observation obtained at the independent variable x and time t; $M(x,t,\Theta)$ are the predictions of the model M at x,t given the parameter

[7] An ironic fact about the Bayesian approach is that the hypothesis about the behavior of the noise can only be justified through logic, never by the evidence, under penalty of entering an infinite regression of decisions about the noise of the noise. After all, to justify the hypothesis about random noise based on evidence, we would need to repeat the Bayesian method *ad infinitum*, evaluating the noise of the noise, and the noise of the noise of the noise, and so on.

vector Θ, and; $\varepsilon(x,t)$ is the **error**[8] of the empirical observation at x,t. Of course, in this context, the random variables of interest for estimating the likelihood of the model M are the parameters Θ themselves. A typical set of assumptions[9] about the behavior of random noise is that the error ε exhibits:

1. zero mean, i.e., $\mu_\varepsilon = 0$;
2. constant (stable) variance σ_ε^2;
3. independence over time, and;
4. normal distribution: $p(\varepsilon) = \frac{1}{\sigma_\varepsilon\sqrt{2\pi}}e^{-(\varepsilon-\mu_\varepsilon)^2/(2\sigma_\varepsilon^2)}$, where μ_ε is the error mean (zero) and σ_ε is the standard deviation of the error.

The rational justification for considering the normal distribution is based on the **central limit theorem**, which states that the sample mean of any population is normally distributed[10]. Thus, the problem of the likelihood of the parameters Θ is resolved by estimating the population variance of the error σ_ε^2 by its sample variance, which is defined as $s_\varepsilon^2 = \frac{1}{df}\sum_{i=1}^{n}(\varepsilon_i - \bar{\varepsilon})^2$, where $\bar{\varepsilon}$ is the sample mean of the error; n is the sample size, and; df is the degrees of freedom[11] [21]. The error ε_i of each n observation is the difference between the observation O_i and the prediction of the model curve M_i defined by the parameters Θ adjusted with optimization techniques, such as the least squares method. Next, the uncertainty associated with the variance of the error must be assimilated in some way by the parameters Θ. For linear models, this can be done analytically, based on the principles of linear combination of random variables. A robust alternative, applicable to models in general, is the **Monte Carlo simulations** method. In this method, numerous *resamplings* of the error ε are performed, i.e., simulations of statistically equivalent realizations[12]. For each simulation, new values for the parameters Θ are adjusted using optimization techniques. Thus, the database generated by these simulations allows the estimation of the empirical probability distribution of the parameters Θ of the model $M(x,t,\Theta)$, which can finally be used in the conditioning process.

Figures 1.3 and 1.4 present an illustrative example for the conditioning process of a model. The goal was to condition a linear model of the type $M(x,\Theta) = c_1x + c_0$ using the available empirical evidence[13]. Note that $\Theta = \{c_0, c_1\}$, meaning the problem consists of obtaining the posterior distribution of the parameters c_0 and c_1. Let's first consider the case of the initial situation (Figure 1.3), when the first sample of empirical data was obtained ($n = 50$). Given the data, the model $M(x,\Theta)$ was fitted using the least squares method. The exact values obtained for the parameters in the initial fit are *irrelevant*, as we are interested in obtaining a probability distribution, not precise values. The initial fit of the model only serves to estimate the dispersion of the error

[8] Also referred to as **residual**.

[9] The introduction of bias, temporal autocorrelation, and other distributions can also be made within a mathematically formal, albeit more intricate, framework.

[10] What this theorem means is that sums of random numbers (note that the mean is a sum) tend to naturally produce the normal curve pattern simply by combining high values with low values. For example, consider rolling a fair six-sided die. The probability of each top face showing one of the values from the set $\{1,2,3,4,5,6\}$ is the same, 1/6. However, the probability of the *mean* of the values sampled in n rolls tends to be much higher among intermediate values as n increases, since high values are offset by low values. This is the fact that produces the bell-shaped pattern modeled by the normal curve.

[11] $df = n - 2$ for models with two parameters.

[12] In the case of small samples, with $n < 30$, the error ε should be simulated using the Student's t distribution with $n - 1$ degrees of freedom.

[13] The data here is synthetic, generated for the purpose of illustrating the Bayesian approach.

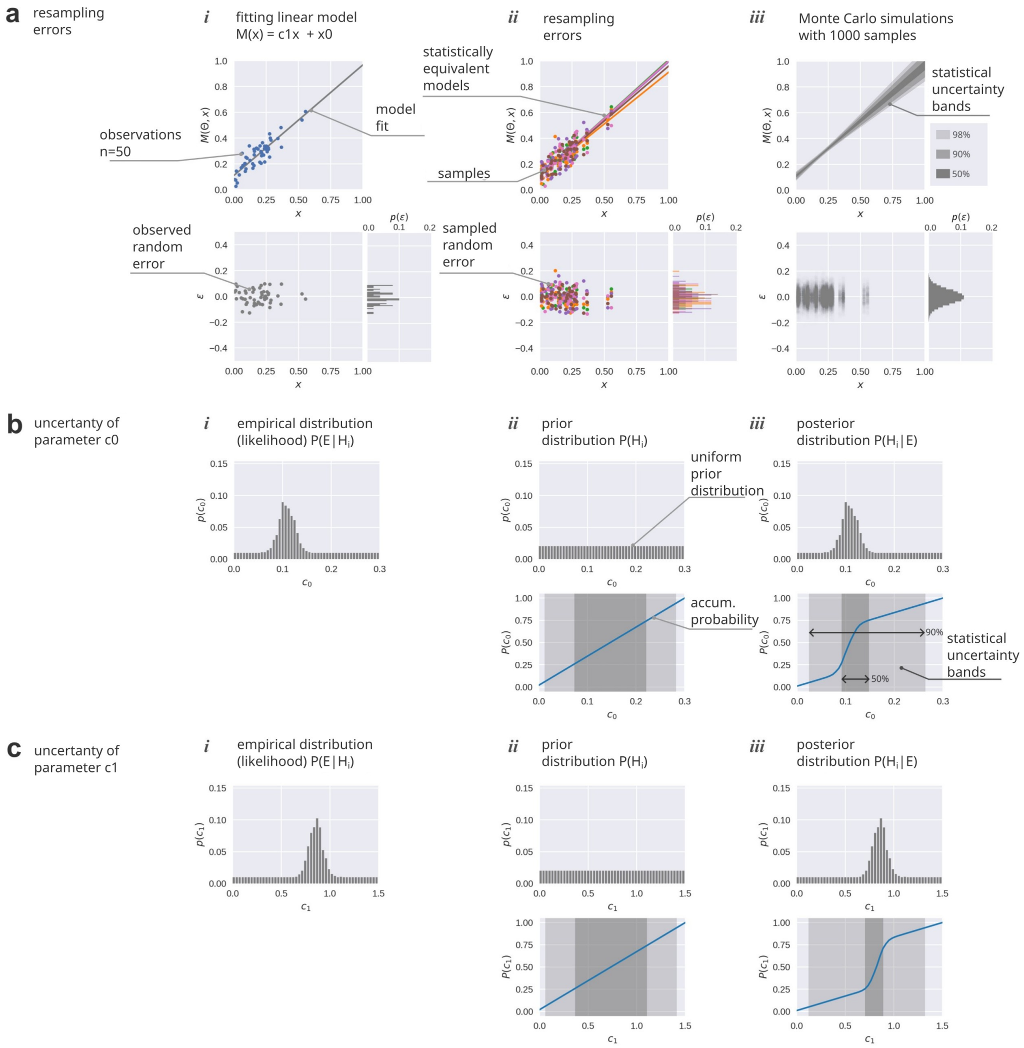

Figure 1.3 — First stage of conditioning of a linear model of the type $M(x, \Theta) = c_1 x + c_0$. **a** — Fitting a linear model using the least squares method. This first fit allows estimating the behavior of the error ε (detail a.i.). If the assumptions about the error are met, numerous error resamplings are performed (Monte Carlo simulations, details a.ii. and a.iii). For each resampling, new models are fitted. The resampling database allows the estimation of the empirical probability distribution of the parameters c_0 and c_1. **b** — Application of Bayes' Theorem to obtain the posterior distribution of parameter c_0 (detail b.iii.). The empirical distribution (detail b.i.) was obtained from the results of the Monte Carlo simulation. **c** — Application of Bayes' Theorem to obtain the posterior distribution of parameter c_1 (detail c.iii.). The empirical distribution (detail c.i.) was obtained from the results of the Monte Carlo simulation. In both cases, the prior distribution was considered uniform (b.ii.; c.ii.).

ε. By simple visual inspection, it can be seen that the distribution of the model's errors is well-symmetrical around zero with stable dispersion. Assuming that the error ε follows a normal distribution with zero mean and constant variance, the Monte Carlo method was applied with a thousand resamplings, which were done by approximating the population variance by the sample variance (i.e., $\sigma_\varepsilon^2 \approx s_\varepsilon^2$). In each simulation, new fits for the model were performed, allowing the estimation of the uncertainty bands for the model $M(x, \Theta)$. Finally, the likelihood distribution $P(E|H)$ of the parameters c_0 and c_1 was estimated from the histogram of the list of a thousand statistically equivalent values generated by the simulations. Since the prior distribution $P(H)$ was uniform, the posterior distribution pattern $P(H|E)$ was completely influenced by the likelihood. Now let's consider the second stage (Figure 1.4), when a new sample of empirical data was introduced ($n = 50$). In this stage, the new data was mixed with the old, and the same procedure was followed: a model was fitted to the data, and new error ε simulations were conducted by approximating its population variance with its sample

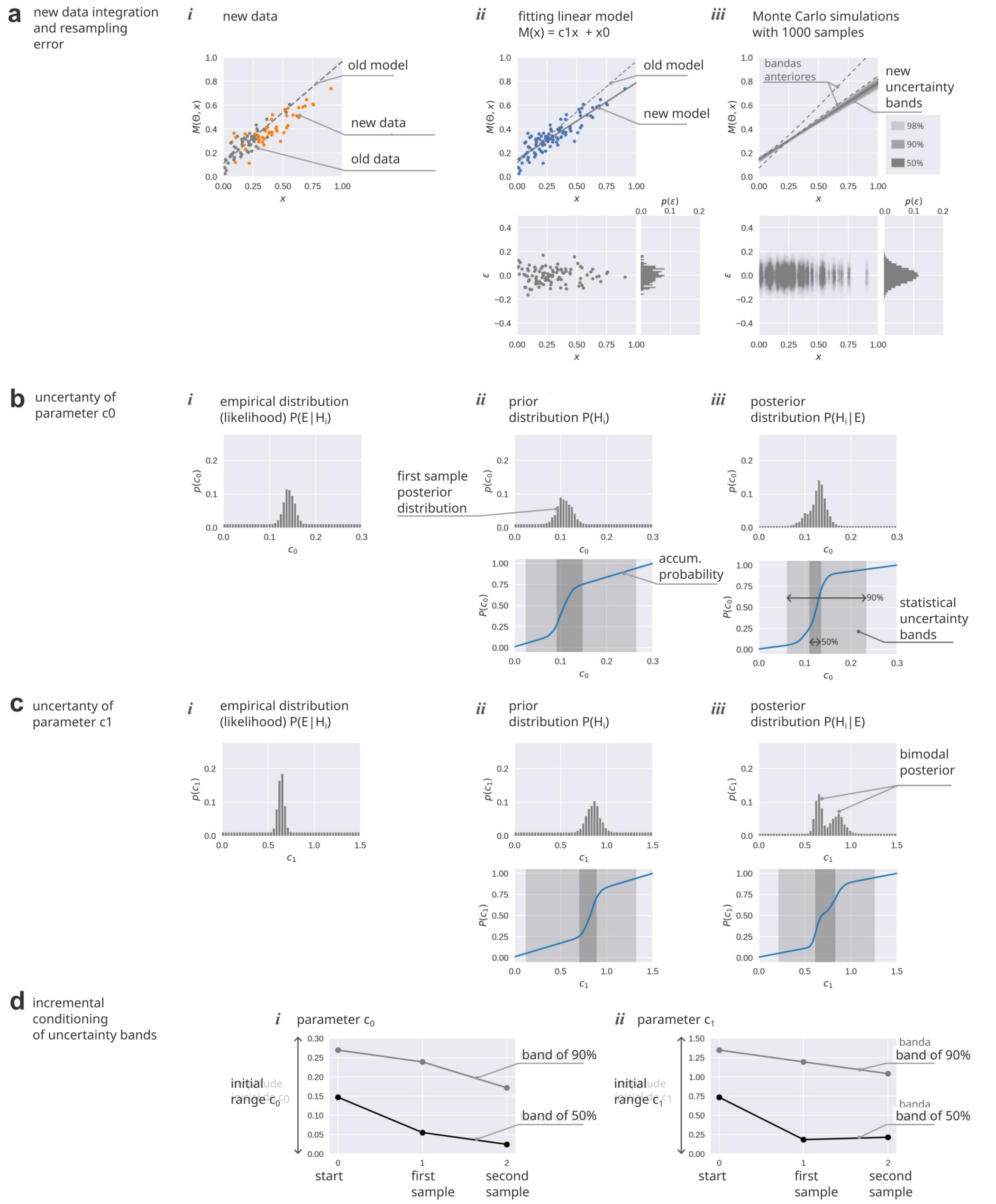

Figure 1.4 — Second stage of conditioning of a linear model of the type $M(x, \Theta) = c_1 x + c_0$. **a** — Second stage of conditioning. The same procedures are carried out as in the first stage, with the difference that new data obtained is integrated with previous observations and that the prior distribution used is the posterior distribution from the first stage. **b** — Application of Bayes' Theorem to obtain the posterior distribution of parameter c_0. **c** — Application of Bayes' Theorem to obtain the posterior distribution of parameter c_1. **d** — Analysis of the uncertainty bands of the parameters as new samples are taken. The uncertainty bands of the model's predictions and parameters have been reduced, except for the 50% uncertainty band in the second stage of parameter c_1. This occurred because the evidence in the second sampling is highly inconsistent with those obtained in the first sampling, resulting in a bimodal posterior distribution (detail c.*iii.*).

variance. The exception was that the prior distribution is now the posterior distribution obtained in the initial situation. With this, the arrival of new empirical observations can both *reinforce* or *weaken* the previously obtained credence for the parameters c_0 and c_1. In the illustrated case, it is clear that the new observations slightly weakened the credence for parameter c_1, slightly widening the 50% uncertainty band obtained in the first stage and generating a bimodal posterior distribution.

In Equation (1.3), the embedded assumption is that the error ε is *linearly additive* to the model. However, it could be *multiplicative*, which would make Equation

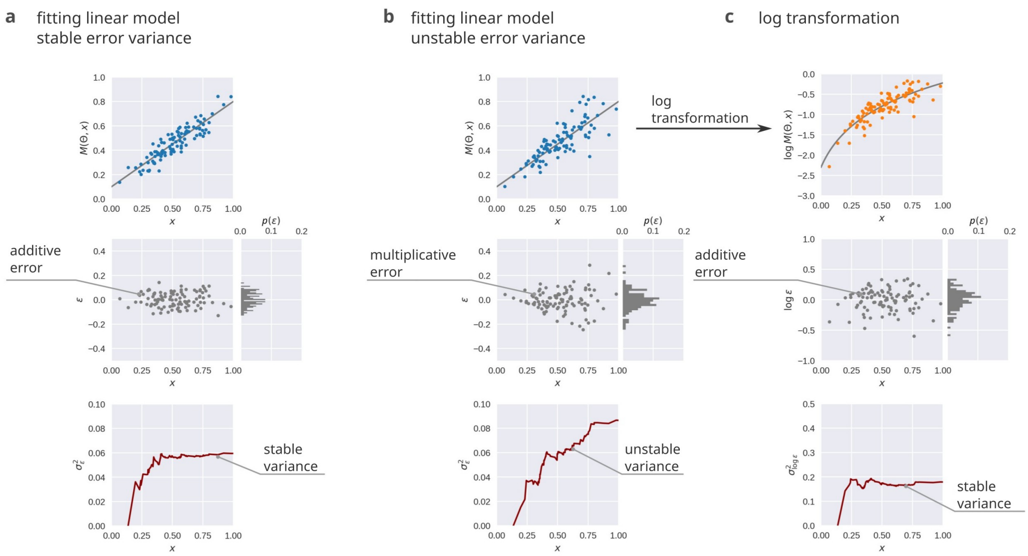

Figure 1.5 — Additive error and multiplicative error in the fitting of a linear model. **a** — Additive error in a linear model, with stable error variance (homoscedastic). **b** — Multiplicative error in a linear model, with unstable error variance (heteroscedastic). **c** — Stabilization of the multiplicative error variance through a logarithmic transformation. The logarithm of the error is additive.

(1.3) take the following form:

$$O(x,t) = M(x,t,\Theta) \cdot \varepsilon(x,t) \tag{1.4}$$

This assumption makes sense when the random noise increases as the independent variable grows. In the case of rating curves, it is reasonable to expect that proportionally larger errors are present in the measurement of high flows from the water level, especially (but not only) due to greater uncertainties in the geometry and roughness of the channel section. When this occurs, the error variance is not constant but gradually increases. In this case, the variance is heteroscedastic (unstable), as opposed to homoscedastic (stable). An alternative to address this case is through **variable transformation**, converting the problem to an additive case by taking the logarithm on both sides of Equation (1.4), since $\log(ab) = \log(a)+\log(b)$. Thus, the assumptions mentioned earlier can be evaluated over $\log(\varepsilon)$, which possibly shows stable variance. To apply the Monte Carlo method, the resampling of the error can be done directly on $\log(\varepsilon)$ and then converted back into Equation (1.4) for fitting the model $M(x,\Theta)$ through optimization techniques. Figure 1.5 illustrates this process.

1.4 Rejection of theories

Despite the success of empiricism as the hegemonic current in the Philosophy of Science during modernity, the remarkable changes in Physics in the 20th century gave another chance to rationalism, that is, the deductive approach to the problem of justification. The impact of Albert Einstein's (1879-1955) work is a good example of this historical moment. In this case, Einstein revolutionized Physics with what he called *thought experiments*. If theories are products of empirical experience, as empiricists claim, Einstein would never have written his first papers, since at the time he worked at a patent office and had no access to laboratories or other resources to collect empirical data. On the contrary, it was other scientists who, through observations and experiments, supported Einstein's theory *a posteriori*, that is, *after* his ideas were already published. Something was definitely wrong with the empiricist theory justification current.

Destaque 1.3.1– Flow uncertainty bands from rating curves

Flow in rivers is almost never measured directly, as it requires a specialized technical team. It is much easier and cheaper to observe the **water level** in rivers using staff gauges. In fact, the water level of many rivers in Brazil is observed twice a day at streamflow stations of the National Hydrometeorological Network. Thus, the rare feasible flow observations are used to construct a **rating curve**, which is usually a power-type model:

$$Q = a \cdot (h - h_0)^b$$

Where Q is the flow in m^3/s; h is the observed level, and; h_0, a, and b are the parameters of the model. This curve can then be used to estimate flow from routine level observations.

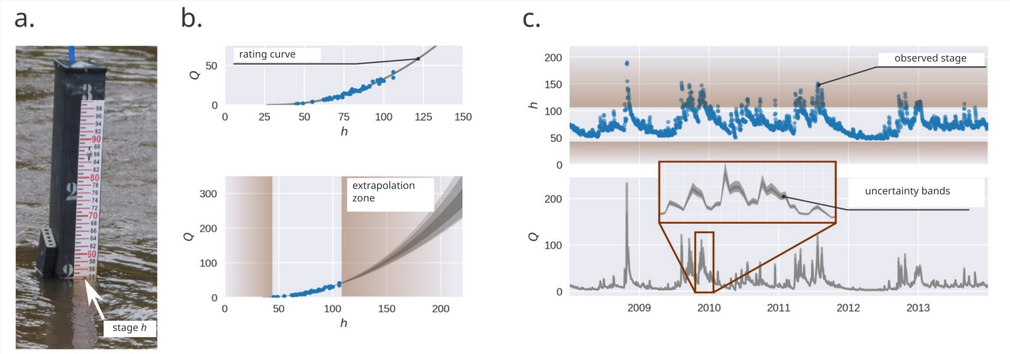

Figure 1.6 — **Flow uncertainty bands from rating curves. a** — Water level observation h on a staff gauge. **b** — Fitting a power-type model and estimating uncertainty through error resampling methods. **c** — Historical series of water level and flow uncertainty bands.

As illustrated in Figure 1.6, the confirmation of this model based on water level and flow evidence begins by fitting the parameters to the available data using optimization techniques. The behavior of random error can then be evaluated. If the error variance is stable, statistically equivalent resamplings of the data (Monte Carlo simulations) can be performed. Thus, the uncertainty bands of the rating curve reflect the uncertainty in the flow estimates in the historical series. In the presented example, **extrapolation zones** can be observed at the extremes, where the uncertainty expands disproportionately. This is expected, as extreme flow events are rare or difficult to measure. On the other hand, capturing a few extreme events can drastically reduce the uncertainty in these zones.
An interesting approach in this context is presented by Thomas Morlot and colleagues, who highlight that in addition to statistical uncertainties, rating curves exhibit temporal dependency associated with changes in the river section morphology [22]. Other complexities also exist, such as hydraulic hysteresis, which can manifest under different flow regimes.

The exponent of this new rationalist movement was the philosopher Karl Popper (1902-1994), especially with his work *The Logic of Scientific Discovery*. He introduced the current now known as **critical rationalism**. On one hand, Popper was aware of the seriousness of the induction problem, which remained (and still remains) unsolved since its formulation by Hume – for him, inductive empiricism clearly could not be sustained. On the other hand, the philosophical currents of his time, called Conventionalists, also did not appeal to him, even though they represented deductive approaches to theory justification. In this sense, Popper concluded that the great epistemic power of empirical evidence is to justify, by the deductive method, the *falsity* of a theory. In other words, it is through **rejection**[14] of theories by evidence that true knowledge is obtained.

A simple example conveys the strength of this argument.

Consider the universal statement that "*all swans are white*". To definitively establish the truth of this statement by the inductive method, it would be necessary to observe all the swans that exist in the universe, including swans in the past and future,

[14] Here, the terms *rejection*, *refutation*, and *falsification* are used interchangeably.

which is obviously impossible in practice. On the other hand, it only takes seeing a *single black swan* (or any other color) in time and space to definitively refute the theory that all swans are white. After all, if the singular statement "*a certain swan is black*" is true (because it was empirically verified), then we deduce that the universal statement "*all swans are white*" is false:

$$
\begin{array}{rl}
S_1 \implies S_2 & \text{if all swans are white, then a certain swan is white} \\
\neg S_2 & \text{a certain swan is } \textbf{not} \text{ white} \\
\therefore\ \neg S_1 & \text{therefore, not all swans are white}
\end{array}
$$

This mode of deductive logic, involving negation, is called *modus tollens*, in contrast to *modus ponens*, which involves positive affirmation. What Popper demonstrates is that there is a fundamental asymmetry between these two modes of logic when we want to deduce universal statements from singular statements (make generalizations from specific observations), with *modus tollens* being the only way to obtain secure knowledge (in this case, the falsity of a universal statement). Hence, observations and empirical experiences are important for refuting theories, not for confirming them. As long as a theory survives rigorous empirical tests, it is said that the theory is *corroborated* by empirical evidence – but never confirmed.

Armed with this argument, Popper turns to what he calls the **demarcation problem**, initially raised by Kant: the difficulty of distinguishing whether a theory is *scientific* or merely *metaphysical*, based solely on abstractions. Unlike empiricists, who claim that experience is the origin of all knowledge, for Popper the origin of where theories arise is irrelevant. As long as they are logically consistent, they can emerge either from some motivating empirical observation (like the example of white swans) or from creative intuition (such as Einstein's thought experiments). What matters is the theory's ability to not survive the tests of empirical experience, that is, its ability to be rejected. This ability, called **falsifiability**, is the *demarcation criterion* that categorizes a theory as scientific. In summary, from the perspective of critical rationalism, a scientific theory must be falsifiable. Here, it is important to emphasize that *being falsifiable* does not mean *being false*. Being falsifiable means that the structure of the theory allows for observations to be used to demonstrate its potential falsity. If *in principle* it is impossible to prove that a theory is false through observations or experiments, then that theory is not scientific. This generally implies that scientific theories must be precise enough to produce observable predictions. In Einstein's case, his theory was precise and made observable predictions that were, in Popper's jargon, *corroborated* by other scientists (but could have been perfectly refuted). A more intuitive example of a non-falsifiable theory is the Multiverse theory, a cosmological theory that postulates the existence of Universes parallel to the one we inhabit. As enticing as it may be to *explain* some cosmological mysteries, this theory is not scientific because it does not allow for any *test* with empirical observations – in principle, there is no way to observe beyond our own Universe. In Popper's words:

> (...) a theory is something the mind attempts to prescribe to nature; something nature often does not allow to be prescribed to her; a hypothesis we attempt to impose on nature, but which can be contradicted by her – Karl Popper [23].

Once falsifiability is designated as the demarcation criterion for scientific theories, Popper moves on to the so-called **problem of simplicity**. This epistemological problem consists of the difficulty in explaining why simpler theories should be preferred over more complex ones (also known as *Occam's Razor*). For example, consider a series of point observations of a given phenomenon plotted in a system of coordinates. If there is a theoretical law that describes this phenomenon, this law will be a curve connecting all the observed points. However, for a finite number of points, it is always possible to fit an infinite number of curves using various mathematical formulas. If a straight line provides a good fit, so can the asymptotic part of a hyperbola. As we saw in the previous section, while Bayesian empiricists have methods to update the credence in the fit of a given curve as new observations are collected, they have nothing to say about the justification for *choosing that curve in the first place*, except for its supposed *simplicity*. This is indeed a confusing problem, as it depends on what we mean by simplicity. For some, it means an aesthetic aspect, something related to mathematical elegance – like the fact that circular orbits for planets seem more beautiful than elliptical orbits. For others, it means a pragmatic aspect, something related to time and resource economy – a simpler method for solving a task should be preferred over a more intricate one. The statistician George Box (1919-2013), for instance, advocates for what he calls the **principle of parsimony** in mathematical models based on purely practical criteria, such as cognitive load, better precision, and objectivity [20].

In view of this, Popper removes any aesthetic or pragmatic aspect, equating the simplicity of a theory with its **degree of falsifiability**: the simpler it is, the more falsifiable. This logically solves the problem of simplicity, because, in his words:

> Simple statements (...) tell us more because they contain greater empirical content and are susceptible to more rigorous testing – Karl Popper [24].

In other words, as long as a simpler, more *restrictive* theory survives empirical tests, it makes no logical sense to adopt a less simple, more *flexible* theory. This becomes clearer in mathematical terms, since the number of parameters in a curve is inversely associated with its degree of falsifiability. Consider a curve with three parameters, such as a second-degree polynomial (a parabola): $f(x) = c_2x^2 + c_1x + c_0$. This curve is much more flexible for fitting data than a curve with two parameters, such as a first-degree polynomial (a straight line): $g(x) = c_1x + c_0$. After all, if we make the quadratic parameter c_2 small enough, we can fit the data equally well with the straight line $g(x)$ without falsifying the theory that the studied phenomenon is described by the parabola $f(x)$, because $\lim_{c_2 \to 0} f(x) = g(x)$. The same logic applies to circular and elliptical orbits: the theory of the circle, being a specific case of an ellipse, should be rejected before the theory of the ellipse not for its aesthetics, but for its ease of being demonstrated false by observed evidence.

The concept of simplicity in Popper makes logical sense, but it doesn't exactly answer how to proceed when faced with observed evidence that presents random noise. In practice, it is impossible to obtain data that perfectly adheres to a mathematical relationship based on some theoretical principle, such as a linear, quadratic, or power function. Before proceeding, it is important to differentiate a theory from a **statistical model**. Theories establish mathematically precise models about specific phenomena. Statistical models, on the other hand, are a very specific type of theory that precisely define *the mathematical pattern of a dataset*, without theoretical links to the underlying phenomena[15]. In the case of theories, the open question in Popper's rationalist approach is when anomalies in the data should be taken seriously enough to reject a simple theory in favor of a more complex one. That is, how much should the data deviate from the

[15] An intuitive example of this difference is to consider a population of, say, a thousand triangles with random sizes. If we look at the data for perimeter and area, we can easily create a statistical model between these two variables: large perimeters are generally accompanied by large areas. But this is not a theoretical law about triangles, as some very acute triangles have large perimeters and small areas. The mathematical theory is that the area of a triangle is its base times height divided by two – the perimeter is only partially and indirectly related.

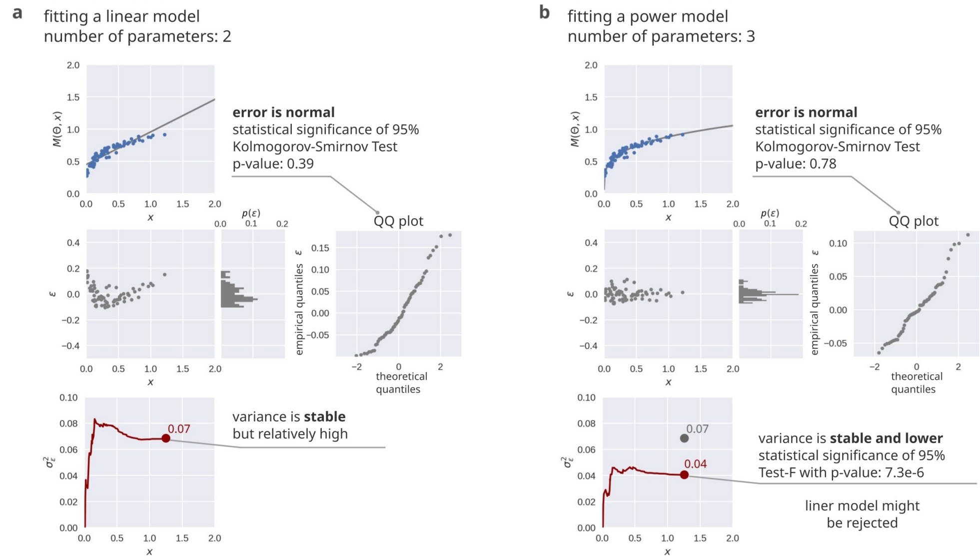

Figure 1.7 — Rejection criteria for model selection. a — Fit of a linear model of the type $M(x, \Theta) = c_1 x + c_0$. **b** — Fit of a power model of the type $M(x, \Theta) = c_1 x^k + c_0$. Both models are fitted to the same set of observed data. Both fits satisfy the assumptions of normal error and stable variance (homoscedasticity). Only by the principle of parsimony (simplicity), the linear model should be preferred as it has fewer parameters. But the power model shows significantly lower error variance with 95% statistical significance. This could be a rejection criterion for the linear model.

proposed model for the theory to be considered falsified by the evidence? This question inevitably introduces a decision, which is the prior definition of a **rejection criterion**. The decision on the rejection criterion needs to be *before* the evaluation of the theory (*a priori*) because if it is *after*, nothing prevents the theory from never being rejected – one just needs to establish a criterion that is known to be lenient. Popper himself criticizes theories that, in the face of clearly falsifying evidence, resort to subterfuges and *ad hoc* explanations to try to survive. This dilemma becomes evident in the case of the premises mentioned in Section 1.3 regarding the statistical model of random noise: zero mean, constant variance, independence over time, and normal distribution. For example, one can evaluate the premise of normality of the error ε using a Quantile-Quantile plot or hypothesis tests, such as the Kolmogorov–Smirnov test. In the case of the Quantile-Quantile plot, the interpretation is purely visual. If the plot deviates significantly from a straight line, the premise should be rejected. In the case of the hypothesis test, the confidence level for the null hypothesis is defined *a priori*. If the desired confidence level is set at 95%, a p-value less than or equal to 0.05 indicates that the probability of normality given the considered data is less than 5%, and the premise should be rejected. Is there a logical justification for the 95% level? There is none. Another example, illustrated in Figure 1.7, is when all the premises about the residuals are satisfied by both a simple model and a complex model, but the complex model is more *accurate* than the simple model, meaning the error dispersion is *less*. How much should the dispersion be less to reject the simpler model? Again, one could apply a hypothesis test for equal variances, the F-Test, and obtain an answer for a previously defined confidence level. If the variances are statistically significantly different, the simpler model should be rejected. One way or another, there is subjectivity involved in rejection, as criteria or thresholds defined *a priori* are necessary.

1.5 Paradigm shifts

In the Philosophy of Science, there is an important distinction between **context of justification** and **context of discovery**. The first context addresses the epistemological problem of how to justify the truth of a theory. The second context investigates the historical and sociological problem of how progress occurs in Science – if indeed progress exists. The different philosophical currents mentioned in the previous sections fit into the first context, as they provide solutions to the epistemological problem of justification. In one way or another, they assign an important role to empirical evidence. For Bayesian empiricists, evidence would be used inductively to confirm the credence in a theory based on the mathematics of probabilities. For critical rationalists, evidence would be essential to falsify theories through deductive logic, leaving theories in an eternal provisional state – they would be corroborated but never confirmed. *Confirmation* and *rejection*, thus, form a somewhat paradoxical dichotomy, as both make sense in the practice of Science, but contradict each other. This paradox is resolved by the perspective of the context of discovery. In his work *The Structure of Scientific Revolutions*, Thomas Kuhn (1922-1996) provides substantial contributions in this regard.

As a historian, Kuhn warns that Science does not exist in a vacuum: rather, it is composed of a *community* of human beings who interact over History, across successive generations, within a larger society. The existence of the **scientific community** implies that one must consider not only the History of Science but also the Sociology of Science to understand the context of discovery. This community is obviously not a single block, but a social network of smaller communities across different disciplines and fields of knowledge. In this light, Kuhn proposes that the dynamics of a given scientific community produce a cyclical historical pattern, structured into three interconnected phases: the period of **normal science**, the period of **crisis**, and the period of **revolution**. Thus, the confirmation and falsification of theories predominately occur at different stages of this historical pattern, with confirmation being a dominant process in the normal science period and rejection being an essential characteristic of the crisis period. However, the most important process for change in Science, occurring during the revolutionary period, is the **competition** between theories. In his words:

> (...) the competition between segments of the scientific community is the only historical process that actually results in the rejection of an previously acceptable theory or the adoption of another. – Thomas Kuhn [25].

This process is inevitably intergenerational, as new members of the community need to be reeducated to think within a new worldview. To illustrate this point, Kuhn cites a striking excerpt from the autobiography of physicist Max Planck (1858-1947):

> (...) a new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die, and a new generation grows up familiar with it. – Max Planck *apud* Thomas Kuhn [25].

A key idea for Kuhn is that a scientific community shares a common **paradigm**. By paradigm, he refers to a set of exemplary solutions for research problems, meaning a system of theories, instruments, and auxiliary practices that effectively solve certain widely accepted problems and are *promising* for addressing mysterious and controversial issues with great competitive appeal. This competitive appeal is crucial, as the attraction of segments of the community around a paradigm creates a positive feedback: the more segments adopt the paradigm, the more new segments become convinced that they need to adopt it as well, under the penalty of falling behind. The scientific community in the normal period thus operates on both theoretical and applied fronts to reaffirm and articulate the hegemonic paradigm. Scientific research during this period is not explicitly aimed at discovering unexpected novelties; instead, the success of normal research is defined precisely by not encountering any surprises. A successful research effort typically confirms what the prevailing paradigm has already promised by providing a more refined detailing or by expanding the range of applications (for example, through the invention of new technologies). Like a jigsaw puzzle, in normal science, it is presumed in advance what the complete picture that the pieces form looks like – the only challenge is to fit the pieces together. Here, the sense arises that Science is a *cumulative* endeavor: each new member introduced into the scientific community would have the humble mission of laying another small brick in the grand "edifice of human knowledge". In Kuhn's view, this impression of accumulation, besides being misleading in the larger context, is reinforced by the widespread use of textbooks in training new researchers. These textbooks serve as vehicles for the perpetuation of the hegemonic paradigm because, when they are not simply written in an anti-historical manner, they distort History to make it appear as a linear and inevitable process leading to current theories.

Kuhn argues, with various examples from the History of Science, that the pieces of the puzzle studied by normal science eventually do not fit together. While knowledge accumulates during the normal period, empirical and theoretical **anomalies** also accumulate. Typically avoided or ignored, at some point these anomalies begin to cause widespread discomfort within the scientific community, leading it into the crisis period. The most detailed example of crisis provided by Kuhn is that of geocentrism, but he also offers examples in chemistry, mechanics, and electromagnetism. From this perspective, history shows that some crises develop slowly, as in chemistry, while others are sudden, like the one caused by Einstein in Physics. A scientific community in crisis exhibits various symptoms, such as discord, discontent, philosophical debates, and, most importantly, the widespread proliferation of candidate theories to explain the anomalies.

The only way out of the crisis is the revolution brought about by the proposal of a paradigm that is irresistible to the scientific community. As previously noted, the new paradigm must be effective in solving already known problems and make enticing promises for solving open issues. The new ideas must, in some way, offer a **retrocompatibility** with the old ideals without being contaminated by the problems embedded in the fundamental principles of the old ideas. Scientific revolutions, therefore, are episodes in which the supposed edifice of knowledge is demolished so that a new structure can be erected on a new foundation, with a new blueprint. During this revolutionary period, which typically lasts a generation, the scientific community migrates en masse to the new paradigm. New textbooks are then written, and a new historical cycle of normal science is established. An important aspect of this process is that, for Kuhn, a new paradigm is so fundamentally different from the old one that they are *incommensurable*: intellectual communication between them is extremely precarious, as they represent different worldviews. Typical examples of **theoretical incommensurability** occur with the concepts of mass, space, and time in Newtonian physics and in Einstein's physics. Despite having the same name and symbol, these concepts have distinct meanings under the different paradigms, with distinct theoretical implications[16]. Thus, Kuhn brings forth a troubling conclusion: that there is *no absolute progress* in Science towards the

[16] In the Newtonian paradigm, gravity is an attractive force related to mass that acts instantaneously at a distance. In the Einsteinian paradigm, gravity is not a force but a *consequence* of the distortion of space itself, implying the existence of gravitational waves. For Kuhn, Einstein does not merely extrapolate the limits of Newton: he produces a new worldview that is irreconcilable with the previous one.

truth about reality, only *relative* to what we are concerned with explaining. More than that, with his thesis, Kuhn highlights the depth of the social dynamics surrounding Science, which is often portrayed as the most rational of human endeavors[17].

As previously outlined, Thomas Kuhn's thesis on the context of discovery eliminates the paradox between confirmation and rejection, which are contradictory solutions in the context of justification. However, a cautious look reveals that Kuhn's approach is essentially empiricist: he seeks to *confirm* the ideas of paradigms and scientific revolutions based on examples from the History of Science, that is, based on *empirical evidence*. Kuhn employs inductive inference to justify a theory about the context of discovery. From the perspective of critical rationalism, no matter how well corroborated, a single counterexample would be sufficient to falsify Kuhn's theory. The problem is that this fact, paradoxically, *resurrects the dichotomy between confirmation and falsification*. To make matters worse, if Kuhn's theory is scientific (i.e., falsifiable), wouldn't it itself be a *paradigm* for how to explain the context of discovery? Here arises a recursive loop of self-reference. Recursion in an argument is typically an indicator of the infinite regress problem mentioned earlier. This is a typical terrifying situation of being endlessly caught in circles that Philosophy provides. Karl Popper, perhaps because he was a philosopher and not a historian, seems to have foreseen these problems and pre-established that the theory on the scientific method cannot itself be scientific – falsifiable by evidence – but only a theory based on Logic.

1.6 The problem of underdetermination

What we have seen so far fits into the broader philosophical current known as **scientific realism**. This current essentially defends the thesis that the purpose of Science is to provide theories that are true descriptions of reality [26]. For example, we began this chapter by mentioning that Keith Beven classifies the philosophy of most users of hydrological models as *pragmatic realism*, which is the tacit understanding that models provide an approximate description of reality that can be improved with new technologies. The pragmatic realism, for Beven, would be a branch of scientific realism. The origins of scientific realism can be traced back to the ideas of René Descartes [27]. Here, it is important to establish that **realism** itself consists of the conception in Metaphysics that admits the existence of an *objective* reality, meaning that the world does not depend on anyone to observe it. In this sense, when a person enters a room and observes a table, it is assumed that the table was there before they entered. The table did not come into existence at the moment of observation. Objects, like tables, exist independently of subjects. This conception opposes **idealism**, a current that considers reality strictly as a product of subjects, that is, *subjective*. If we agree on the existence of a supposed object, like a table, it is because it manifests similarly in our minds, that is, *intersubjectively*. Descartes flirts with idealism when he questions his own existence in the *Discourse on the Method*, particularly with the branch of solipsism—the idea that the mind of the person reading this text is the only thing that truly exists. Descartes basically points out that, although we have absolute certainty about the ideas in our minds, it is difficult to guarantee that they correspond to external reality. In his terms, *likelihood does not imply truth*. To try to resolve this problem, Descartes describes the method of doubt, which inspired the formation of the modern scientific

[17] Kuhn's emphasis on the relativity of knowledge, historical contingency, and the presence of paradigms has invigorated the emergence of the philosophical current of **postmodernism**, bringing with it the notion that human knowledge is a **discourse**. Thus, postmodernists reject grand absolutist narratives and emphasize the linguistic, cultural, and especially political influences that permeate the production of knowledge.

method, contributing to the debate around the problem of justification that we have addressed so far. Ultimately, the problem of justification is inherently contaminated by the *assumption that objective reality exists*, with the concept of **truth** being precisely the *correspondence* between theories and reality.

The thesis of scientific realism seems obvious, but defending it is not so straightforward. In fact, Bas van Fraassen [26] and Nancy Cartwright [28] provide a profound critique, proposing a radically empiricist viewpoint known as **instrumentalism**[18] [29]. Both argue that the objective of Science is to produce theories that exhibit *empirical adequacy*—and nothing beyond that. Since empirical adequacy does not logically imply a true description of reality, the claim of scientific realism is too ambitious in epistemological terms. This viewpoint does not deny the existence of reality (it is not an idealist current): what it denies is the ambition of obtaining a true description of reality. Theories and their models would merely be *instruments* constructed by scientists to explain empirical evidence. One of the main reasons for this claim is the **underdetermination problem**, which is the difficulty of ensuring that the observed evidence determines the truth of a theory without there being empirically equivalent theories [30], [31]. Popper's orientation to always prefer the simplest theory works well only for theories that are completely falsifiable by empirical evidence. This is not the case for most theories, which almost always postulate the existence of *unobservable entities* to explain phenomena that are directly observable. For example, in Physics, the existence of electrons and electromagnetic fields (unobservable) is invoked to explain the lightning and thunder of a storm (observable). This complicates matters, as no matter how we detect unobservable entities, like electromagnetic fields, indirect evidence will always be contaminated with a *theoretical load* that establishes the existence of these entities in the first place. This type of theoretical approach involves a kind of reasoning that is non-deductive, called **inference to the best explanation**, or abduction. Because it is not deductive, this reasoning does not guarantee the truth of the consequent and is also subject to the induction problem postulated by Hume. Thus, a theory that instantiates unobservable entities pays the price of being underdetermined by observable empirical evidence.

One of the main defenses of scientific realism consists in evoking the success of Science as evidence that scientific theories, even when instantiating unobservable entities, progress toward describing reality in an increasingly true manner [32]. From the critical rationalist perspective, although the ultimate truth about reality remains permanently shielded, the rejection of theories allows for the incremental isolation of a set of potentially true ideas. Indeed, it is undeniable that the theoretical predictions and technological applications that Science has produced in recent centuries are impressive and unprecedented in historical terms. Given all this success, it even sounds somewhat absurd to consider that modern Science does not describe reality. Although inference to the best explanation does not guarantee a logical implication, as instrumentalists correctly point out, defenders of scientific realism argue that it would be *a miracle* extremely unlikely for current theories to achieve good results for the wrong reasons. However, Donald Hoffman introduces the possibility that scientific theories describe the behavior of a *cognitive interface* with remarkable empirical adequacy [33]. He argues that cognitive systems, when subjected to natural selection, are pressured to operate through **heuristic**. That is, those systems that condense the necessary information to make useful decisions gain a competitive advantage. The evolution of these systems results in a perceptual interface optimized for survival and reproduction, but whose probability of being equivalent to reality is *precisely zero*. As an analogy, consider the

[18] Instrumentalism is a broad and neutral term. For example, van Fraassen self-identifies his thesis as *empiricism constructivist*. Realists, on the other hand, classify instrumentalism as *anti-realism*.

graphical interface of a computer. In this case, we can easily observe the behavior of buttons and icons to identify patterns without knowing anything about the underlying electronic mechanisms. The information from the graphical interface tells us absolutely nothing about the *hardware*. For Hoffman, the truth about reality may simply have nothing to do with space, time, energy, matter, etc.—in Kantian terms, these would be the transcendental categories we use to condense and integrate perceptual information [19]

The underdetermination problem has direct and relevant implications for users of environmental models, including hydrological models. In this context, Naomi Oreskes and colleagues point out that underdetermination occurs because various processes represented by the models are not observable *in practice*, meaning that information about the modeled system is *incomplete* both in time and space [35]. This milder version of underdetermination is also referred to as the **equifinality problem** [36]. For example, consider the groundwater flow occurring in watersheds. It is evident that this process exists: a field expedition makes this clear by directly observing the springs of streams, the places where groundwater flows to the surface. In fact, piezometers can be installed to monitor the water table level, providing more direct evidence of this process. However, the extent and complete dynamics of these underground flows are practically impossible to monitor, being observable only in specific points. Oreskes *et al.* argue that the partiality of the information renders the modeled natural systems *logically open*. Unlike logically closed systems, such as algorithms and mathematical equations, they point out that it is impossible to verify or validate a logically open system in light of *extenuating circumstances* that often ensure empirically equivalent explanations, or *equifinal*. This is intuitive: when we do not have complete information about some event we observe, it is natural for rival and equally valid explanations to emerge. Indeed, it is precisely for this reason that scientific experiments are designed to reduce the logical openness of the evaluated system, that is, to lessen the influence of extenuating circumstances. Thus, models of natural systems present themselves as a **main hypothesis** that requires the assistance of **auxiliary hypotheses** — such as parameters, input data, the adopted scales, and, especially, the underlying theoretical assumptions. This gives rise to a paradox: it is precisely due to the lack of information that the application of models is sought in the first place. If the information were already completely available, it is unlikely that a model would be relevant for decision-making. But without complete information, a model becomes underdetermined by the available evidence—the inexorable underdetermination problem in model application.

For Keith Beven, recognizing the underdetermination problem in hydrological modeling brings radical consequences for the confirmation of models in light of observed evidence, specifically the need to evaluate the **total error** associated with a given hydrological model [37]. From this perspective, Equation (1.3) would be incomplete, as the error ε there represents only the random noise related to the observed evidence. It is necessary to include not only the **statistical uncertainty**, resulting solely from sampling noise, but also the epistemic uncertainty, which arises from the auxiliary hypotheses necessary to address the underdetermination problem [38]. Thus, the **total error equation** for hydrological models takes the following form:

$$O(x,t) + \varepsilon_O(x,t) + \varepsilon_\Delta(\Delta x, \Delta t, x, t) = M(\Theta, \Upsilon, \varepsilon_\Upsilon, x, t) + \varepsilon_M(\Theta, \Upsilon, \varepsilon_\Upsilon, x, t) + \varepsilon_r \quad (1.5)$$

[19] Donald Hoffman subverts the paradigm of material physicalism by proposing that reality is not fundamentally constituted of subatomic particles, but rather of an infinite network of interactions among conscious agents [34]. The interactions of these agents produce cognitive interfaces that eventually instantiate self-referential ties, that is, realize a *Self*, an "I". This hypothesis simultaneously explains why subjective experiences exist (note that they are not predicted within material physicalism) and why realism definitely does not hold at quantum scales (the supposedly physical properties are realized instantaneously at the moment of observation).

where $O(x,t)$ is the observation obtained at the independent variable x[20] and at time t; $\varepsilon_O(x,t)$ is the **measurement error** of the observation; $\varepsilon_\Delta(\Delta x, \Delta t, x, t)$ is the **commensurability error** at the modeling scale Δx and Δt; $M(\Theta, \Upsilon, \varepsilon_\Upsilon, x, t)$ is the prediction of the model at x, t based on the vector of parameters Θ, the vector of input data Υ, and the **input data error** ε_Υ; $\varepsilon_M(\Theta, \Upsilon, \varepsilon_\Upsilon, x, t)$ is the **model structural error**, and; ε_r is the **random error** remaining. The commensurability error ε_Δ results from the conversion between scales, representing the epistemic uncertainty of the difference in meaning between an observation obtained at x, t and the corresponding modeled variable at $\Delta x, \Delta t$. For example, while the observed flow of a river is instantaneous and refers to a specific section of the channel, the modeled flow integrates some time step and refers to a discrete spatial extent. The measurement error ε_O and the commensurability error ε_Δ are kept on the left side of Equation (1.5) to denote that together they constitute the **effective observational error**. The input data error ε_Υ originates from the aggregation of uncertainties in both boundary conditions (such as maps of topography, soil, vegetation, etc.) and the forcing variables of the model (such as rainfall, temperature, wind speed, etc.). In this case, the uncertainty is generally statistical, so representative samples tend to reduce its impact. However, it can also take on an epistemic nature when the input data correspond to **scenarios**, which adds a conceptual burden. Finally, the model structural error ε_M results from the epistemic uncertainty of the theoretical and numerical configuration of the hydrological model. This component is strongly influenced by the theoretical assumptions previously defined about the system and its hydrological processes.

With Equation (1.5), Keith Beven operationalizes an instrumentalist paradigm for hydrological modeling that, beyond confirmation, *allows for the rejection of models*. This approach follows the recommendations discussed by Albert Tarantola, to consider both the empiricist confirmation of Bayes and the rationalist rejection of Popper for a philosophically explicit approach to environmental modeling [39]. In this line, the critique of the hegemonic modeling paradigm, dominated by pragmatic realism, is that model confirmation occurs at the cost of underestimating epistemic uncertainties, masking them as random error minimized through optimization techniques, as seen in Equation (1.3). This leads to the **overfitting problem** of models to the available data used. Through the conventional **calibration process**[21], one reaches the (incorrect) conclusion that the adjusted model identified is the only empirically adequate representation. On the other hand, the new instrumentalist approach, in Beven's words:

> (...) There is, however, another approach. That is to accept that it is very unlikely that our current model structures are truly realistic descriptions of the enviromental systems of interest so that there may indeed be many different models that can be shown to provide predictions that are acceptably consistent with whatever observed data are available. That is to treat the problem of identifiability as one of equifinality of model structures and parameter sets in reproducing the known behaviour of the system. – Keith Beven (2009. p. 15) [40].

It is important to note that this new approach does not abandon the confirmation process through the Bayesian conditioning of the posterior distribution of the parameters Θ. Although the total error equation makes a formal treatment of the likelihood $P(O|M)$ impossible, it remains possible to assign different degrees of conviction, or weights, to the **empirically equivalent models** through informal likelihood measures $\mathcal{L}(O|M)$. The

[20] In hydrological models, the independent variable is usually two-dimensional space, meaning x should be replaced by x, y.

[21] Also referred to as the *inverse problem*

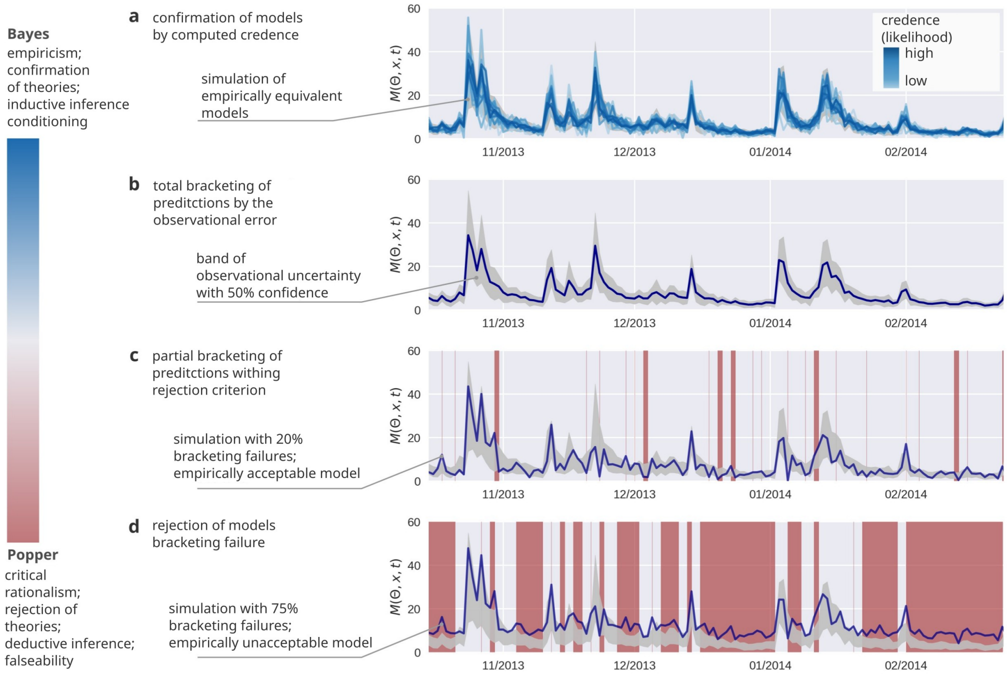

Figure 1.8 — Instrumentalist approach to hydrological modeling. In this approach, both Bayes' confirmation and Popper's rejection are employed under the recognition of the underdetermination problem (equifinality). **a** — Confirmation of models through Bayesian conditioning, where degrees of confirmation (informal likelihood measures) are assigned to empirically equivalent models. In this case, all models are encapsulated by the observational uncertainty band within the pre-established rejection threshold. **b** — Total (no failures) encapsulation of a simulation (time series) by the effective observational error of the empirical evidence. In the illustrated case, the uncertainty band has a confidence level of 50%. A more or less comprehensive band should be defined *a priori*. **c** — Partial encapsulation (with 20% failures) of a simulation by the effective observational error. As the band is at 50% confidence, the failures are within the rejection threshold, and the model can be considered empirically acceptable. **d** — Rejection of models due to insufficient encapsulation (75% failures). In this case, the failures exceed the 50% confidence level, and the model must be considered empirically unacceptable.

novelty of the approach lies in establishing that an **empirically acceptable model**[22] occurs when its structural error ε_M is less than its effective observational error $\varepsilon_O + \varepsilon_\Delta$. Otherwise, the model must be rejected. This implies that the predictions of any model must be *encapsulated* by a confidence interval defined *a priori* as a rejection criterion. In other words, for a confidence level of $\alpha\%$, the **bracketing inequality**:

$$O_{100-\alpha\%}(x,t) < M(\Theta, \Upsilon, \varepsilon_\Upsilon, x, t) < O_{\alpha\%}(x,t) \quad \forall \quad x, t \tag{1.7}$$

must hold true with a frequency of at least $\alpha\%$; where $O_{100-\alpha\%}$ and $O_{\alpha\%}$ are the lower and upper thresholds of the confidence interval of the effective observational error at each sample point x, t. This approach has exactly the same structure as statistical hypothesis tests: 1) a rejection criterion is pre-defined by a confidence level $\alpha\%$; 2) a test statistic is calculated, in this case the encapsulation rate of the simulation, and; 3) the p-value of the test is evaluated, which in this case is the failure rate of encapsulation. If the p-value is greater than the confidence level, the model should be rejected. On the other hand, all models that pass the minimum encapsulation test are considered empirically equivalent and can be confirmed based on likelihood measures. Figure 1.8 illustrates the approach for encapsulating a time series of any hydrological variable, but it is generally the discharge in a river section. It is noted that the pre-definition of the confidence level implies more or less comprehensive observational uncertainty bands. This fact introduces the following dilemma: when high certainty about predictions is desired, the observational band may be very wide, resulting in various empirically

[22] Keith Beven refers to empirically acceptable models as *behavioral models*.

Destaque 1.6.1– The impact of uncertainty on mapping priority actions

When applied under the instrumentalist paradigm, by admitting multiple empirically adequate solutions, hydrological models do not produce exact values of the simulated variables but rather intervals where confidence is higher or lower. For instance, from a population of suitable models, it is possible to simulate a **bundle of flow series**. Moving statistics on these time series can capture the central tendency, such as the mean, or the dispersion, such as the bands produced by lower and upper percentiles. However, when models simulate hydrological processes in a distributed manner in space, a **stack of maps** is generated, which can be somewhat difficult to visualize. One solution is to visualize centrality and dispersion on separate maps or synthesize an uncertainty index, such as a normalized coefficient of variation.

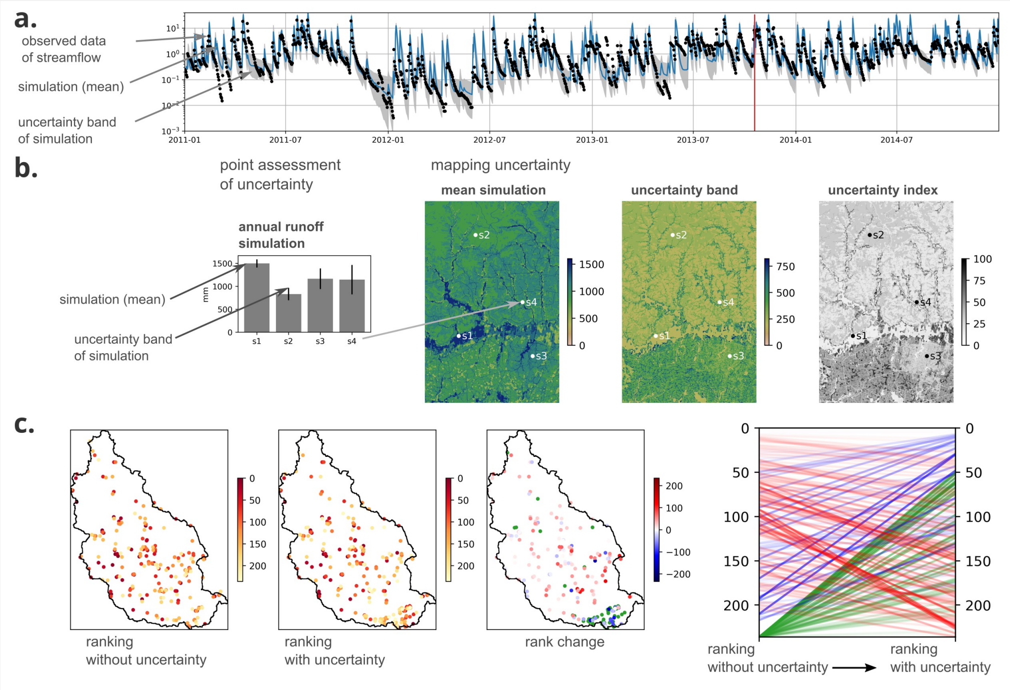

Figure 1.9 — Mapping priority areas considering uncertainties in modeling a — Expression of uncertainties over time through uncertainty bands. **b** — Expression of uncertainties in space by considering separate maps or a normalized uncertainty index **c** — Impact on the ranking of priority areas, comparison of ranking with and without uncertainties.

Since evidence-based policies need to account for empirical uncertainties, model uncertainty has been used as relevant information in identifying priority areas [41]. This approach was articulated by me and my colleagues in Possantti et *al.* (2023) [4], by employing a priority index weighted by uncertainty in regions i:

$$\mathsf{IP}_i = \frac{\sum_{i=1}^{N} \mathsf{IU}_i * \mathsf{IA}_i}{\sum_{i=1}^{N} \mathsf{IU}_i} \quad \forall i \tag{1.6}$$

Where IP_i is the **priority index**; IA is the **suitability index** chosen based on the central tendency, and; IU is the **uncertainty index**. In other words, when prioritizing actions where there is greater surface runoff, the ranking of areas will be weighted by the modeling uncertainty. The results demonstrate that model uncertainty is not uniform in space, causing substantial impacts on ranking when considered by the prioritization formula.

acceptable simulations and little precision in structural error. This creates the need to obtain more evidence, so that the observations themselves present narrow bands for high confidence levels. Another aspect that differs from the hegemonic paradigm, which operates solely through confirmation, is that nothing prevents the eventual *rejection of all tested models* by the bracketing inequality. If this is the case, Beven emphasizes, a valuable opportunity arises to transform modeling into a learning process, forcing users to review both the auxiliary hypotheses and the main hypothesis, that is, the very theoretical assumptions adopted in the conception of the modeled hydrological processes. Ultimately, total rejection imposes the need for new theories and explanations [42]. Without this, the scientific community in this field will be forever trapped in the same paradigms. ■

1.7 Chapter summary

In this chapter, I aimed to establish the foundations of an instrumentalist philosophy for the application of hydrological models. The distinction between rationalism, with its emphasis on deduction, and empiricism, which values induction, was highlighted. From the empiricist side, I presented Bayesian epistemology, which proposes a gradual confirmation of hypotheses based on probabilities. From the rationalist perspective, I articulated the deductive rejection of theories, a stance defended by Karl Popper. In his thesis on scientific paradigms, Thomas Kuhn explains the alternation between periods of normal science and crises. The problem of underdetermination, raised by critics of scientific realism, is applied to hydrological modeling, culminating in an instrumentalist proposal that addresses epistemic uncertainty in the acceptance of empirically adequate models.

- **The problem of justification**. There is a difficulty in justifying the truth of theories, in establishing definitive explanations for events and phenomena. On one hand, rationalists appeal to the use of deductive inference, which guarantees the truth of statements as long as their premises are true. On the other hand, empiricists prefer to use inductive inference, which employs empirical evidence to generalize observed patterns.
- **Inductive confirmation of hypotheses**. Bayesian epistemology describes the process of empirical conditioning to confirm hypotheses. By recognizing the existence of random noise in empirical observations, the truth of a hypothesis must be described as a credence, or probability. In this process, the probability distribution of hypotheses is incrementally adjusted through the application of Bayes' Theorem.
- **Deductive rejection of theories**. Karl Popper, when analyzing the Logic of scientific research, argues that the only safe way to acquire knowledge is through deductive refutation. In this sense, the role of empirical evidence is to test a hypothesis against counterexamples that prove its falsehood. For this reason, Popper claims that scientific theories must be falsifiable theories that allow for their own rejection.
- **Paradigms and the context of discovery**. Thomas Kuhn, by exploring the History of Science, eliminates the apparent contradiction between confirmation and rejection of theories. He argues that the dynamics of the scientific community plays a profound role in the production of knowledge, especially in the advent of paradigms. For him, confirmation occurs in periods of normal science, while rejection predominates in crisis periods. Crisis periods end only when the competition of new ideas leads to a new paradigm.
- **The problem of underdetermination**. scientific realism is deeply questioned by Bas van Fraassen and Nancy Cartwright. They establish an instrumentalist perspective, where the goal of Science is solely to produce empirically adequate theories. This mainly stems from the instance of unobservable entities, which renders theories underdetermined by the evidence. A version of this occurs in hydrological modeling due to many modeled processes being practically unobservable – the so-called equifinality problem. In this line, Keith Beven proposes an instrumentalist paradigm for the application of models, allowing for the rejection of models based on the encapsulation test of predictions by observational uncertainty. Models that pass the test are deemed empirically acceptable and equivalent.

The whole is not merely the sum of its parts. If it were, chairs could not exist. Similarly, people could not exist. The **form** unifies in **dynamic systems** the parts that, when isolated, bear no resemblance to the whole.

Chapter 2

The workings of systems and models

> Everything we think we know about the world is a model. Every word and every language is a model. All maps and statistics, books and databases, equations and codes are models. So are the ways I imagine the world in my head – my mental models. None of these is or ever will be the real world.
>
> Donella Meadows (2008, p. 86) [43]

> If validation is impossible and all models are wrong, why do we bother to build them? As a leader, you must recognize that you will be using a model – mental or formal – to make decisions. Your choice is never whether to use a model, but which model to use. Your responsibility is to use the best model available for the purpose at hand, despite its limitations. Delaying actions in the vain search for a perfect model is, in itself, a decision, with its own consequences.
>
> John Sterman (2000, p. 850) [44]

2.1 The Modeling Process

Donella Meadows (1941-2001) may have been the most brilliant environmental systems modeler to ever live, leading the ambitious initiative proposed by the book *Limits to Growth*, published in 1972 and revised in two subsequent editions. This book issued an unprecedented warning about the ecological scenarios that the current industrial society, based on non-renewable resources, may face by the year 2100, including the possibility of a catastrophic collapse [45]. Her argumentation was based on simulations of a comprehensive model of the world, the model `World3`, mapping the availability of numerous stocks and flows of natural resource consumption, from arable land to oil

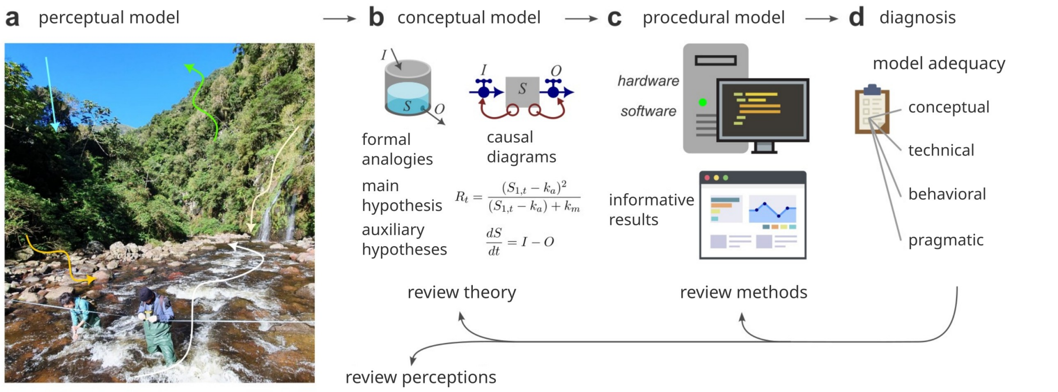

Figure 2.1 — The modeling process. Hydrological modeling can be understood as an iterative learning process. **a** — The first stage consists of the perceptual model (mental models), which is a collection of subjective and personal perceptions acquired through empirical experience (field expeditions) and theoretical experience (textbooks, lectures, classes, etc). **b** — The second stage consists of the conceptual model, which instantiates formal (mathematical) analogies and causal diagrams (structures) to obtain an objective main hypothesis in the form of equations. Various auxiliary hypotheses are generally required, making the conceptual model a logically open system (underdetermined). **c** — The third stage consists of the procedural model, which is the synthesis of the computational methods used (*hardware* and *software*) to simulate the conceptual model and produce results in symbolic forms such as tables, graphs, maps, animations, etc. **d** — Finally, the diagnostic stage applies various procedures to assess the adequacy of the models in conceptual (theoretical problems), technical (computational problems), behavioral (empirical justification), and practical (decision-making impacts) terms. The diagnosis is iterative, reviewing all created models and closing the learning cycle. The photograph in (**a**) was kindly provided by hydrologist Marina Fagundes, who is measuring the flow of a mountain river during a field expedition in Rio Grande do Sul, Brazil.

reserves. Despite the significant social, political, and economic impact of her work, Meadows contributed little toward the more philosophical direction, such as the epistemological problems addressed in Chapter 1. Still, as emphasized in the above epigraph, she left evidence of sharing the Kantian tradition, according to which pure reason has access only to transcendent categories or, in her terms, to **mental models**. These mental models would then be expressed in various forms, including diagrams, texts, equations, and computer programs. Her line of thought eventually suggests an instrumentalist view, in which we will never have the conditions to establish the truth about the world, but only empirically adequate theories:

> Our models usually have a strong congruence with the world. That is why we are such a successful species in the biosphere. Especially complex and sophisticated are the mental models we develop from direct, intimate experience of nature, people, and organizations immediately around us. However, and conversely, our models fall far short of representing the world fully. That is why we make mistakes and why we are regularly surprised. In our heads, we can keep track of only a few variables at one time. We often draw illogical conclusions from accurate assumptions, or logical conclusions from inaccurate assumptions. – Donella Meadows [43].

Regardless of Meadows' position on philosophical currents, her view is clear in that modeling is a *process* that begins in a *subjective and personal* manner with mental models, and then becomes increasingly *objective and impersonal* through texts, equations, and computer programs.

In the field of Hydrology, Keith Beven emphasizes Meadows' perspective, proposing that the modeling process consists of at least three stages represented by models of different natures: the *perceptual* stage, the *conceptual* stage, and the *procedural* stage[1] [46]. Figure 2.1 illustrates this conception, including a final diagnostic

[1] Two additional stages in the modeling process include the calibration and validation of the procedu-

stage. The **perceptual model** begins with the hydrologist's subjective and qualitative understanding of how a watershed responds to precipitation events. This model is profoundly influenced by individual experiences, studies, analyzed data, and the hydrologist's field experience. It is an inherently personal model and varies substantially from person to person. Moving to the **conceptual model**, Beven describes a transition to a more formalized and simplified representation of the processes identified in the perceptual model. This model involves creating hypotheses and adopting assumptions to *abstract* the complex processes of reality into tangible and objective forms, often utilizing mathematical formulations. Finally, the **procedural model** represents the practical implementation of the conceptual model in a computer program. At this stage, the equations and concepts from the conceptual model are translated into code, allowing simulations and predictions of flows and levels based on input data through the application of tensions in electronic circuits. In the case of digital computers, this process involves the application of numerical methods and may introduce additional errors or approximations, making precision and care in execution extremely important. It is these electronic computations that produce the supposedly informative results we see in tables, graphs, maps, etc. For Beven, the interaction and evolution between these three models are crucial in the modeling process in Hydrology. With several caveats, he includes two additional stages, which would be the *calibration* and *validation* of the model against empirical evidence. These are jargons of pragmatic realism. An instrumentalist nomenclature would be *conditioning* and *testing* against empirical evidence. One way or another, a final stage of **diagnosis** should lead to the review and refinement of the previously developed models, giving rise to an *iterative learning cycle* and potential *scientific revolutions* in understanding hydrological processes.

It is with this perspective that the objective of this chapter is to establish the necessary details about the modeling process so that we can soon discuss hydrological models properly. At a certain point in the previous chapter, it became essential to define a model as a **symbolic vehicle of a theory**, a typically instrumentalist conception that resonates with Nancy Cartwright's view [28] – which is effective in articulating the epistemological problems that underlie modeling practices. In this perspective, models are seen as mere translators of our theories or hypotheses about real phenomena, such as the hydrological cycle. However, this is still a generic and abstract definition that does not provide a concrete understanding of the exact nature of models. As emphasized at the beginning of the first chapter, hydrological models materialize in the states of electronic circuits in digital computers, but they are also other things before this materialization. To articulate the enigma of what exactly models are, this chapter will abandon the domain of Epistemology and the Philosophy of Science, delving into the field of Ontology of models. I will address the problem of representation, the paradigm of systems, Systems Dynamics, and model diagnostics. If in the previous chapter we were on a panoramic view with thin air, like at the top of a mountain, we are now certainly descending from the heights, following the valleys of the streams. The analogy remains interesting, as the path is still difficult and steep, but the landscape is becoming increasingly familiar. Hope grows that soon we will be on gentle and open ground.

2.2 The problem of representation

Models serve the function of representing a **target system**. That is, precisely because they symbolically convey a theory, models aim to re-interpret a given phenomenon or entity that supposedly exists and develops in the real world. The problem of justifying

ral model, but these stages are not models themselves; rather, they are steps of empirical justification.

the correspondence between the model and reality was the subject of the first chapter. Here, however, we have a new problem: *how is it possible to create the representations themselves*? The solution to this **representation problem** consists of establishing a process of **idealization** of the target system combined with the application of **analogical inference**, that is, the construction of a **analogy** between the target system and the model. In this line, Mary Hesse proposes that such analogies manifest both through *material models*, semantic structures realized by physical objects, and through *formal models*, syntactic structures expressed by mathematical equations implemented by computer programs [47].

The process of idealization is the foundation of all modeling and is characterized by *deliberate simplifications*, which make the model more tangible and understandable than the target system itself, emphasizing crucial aspects while ignoring supposedly less relevant details. According to R. Frigg and S. Hartmann [48], there are two forms of idealization that are not mutually exclusive: **Aristotelian idealization** and **Galilean idealization**. In the case of Aristotelian idealization, the key lies in the process of **abstraction**, when all supposed superficialities of the target system are removed, leaving only a supposed *essence*. In other words, abstraction aims to preserve the truth, albeit only the part that is relevant. In a hydrological model, for example, the vegetation canopy is usually treated as a single reservoir that intercepts rainwater. It is clear that each leaf and twig plays a role in interception, but this individual process is considered irrelevant and abstracted as a general process occurring throughout the canopy. Alan Musgrave, however, notes that abstraction can also result in falsehoods, especially when **negligibility premises** are introduced, that is, when a *knowingly true* causal factor is neglected [49]. He initially brings this critique to neoclassical economic theories, but it is also the case, for instance, when hydrological models ignore the importance of solar radiation and terrain shading on evaporative processes. The Galilean idealization, on the other hand, consists of applying a controlled experimental distortion, which can be incrementally reversed from the simple to the complex, from the ideal to the real [50]. In other words, idealization exhibits an *asymptotic behavior* that, in the limit, makes the model identical to the target system. The reference to Galileo Galilei (1564-1642) relates to his famous experiments with inclined planes, which led him to conclude that objects fall at the same time, regardless of their mass. In this case, the inclined plane idealized free fall, allowing for a better understanding of the physical process. In hydrological models, an example of this type of idealization is the spatial discretization into response units, sub-basins, or drainage networks – when taken to the extreme of small parcels, it asymptotically approaches the watershed.

Among the available forms of analogies, a somewhat direct alternative is to construct a *copy* of what is understood as the target system, at a scale suitable for manipulation by humans. These material models are referred to as **scale models** reduced or enlarged, illustrated in Figure 2.2a and Figure 2.2b. To some extent, we are all accustomed to models of this type, as the toys we played with as children are like reduced-scale models. A scale model of a building or a car in a wind tunnel, for example, is a reduced-scale model used for engineering applications. Atoms of chemical elements with fittings to form more complex molecules, on the other hand, are enlarged-scale models for educational purposes. In a highly technological age, scale models may seem crude or simplistic, but they are actually extremely interesting options for investigating, visualizing, and experimentally testing the implications of a given theory or hypothesis. A notable example in the History of Science that involved the contribution of an enlarged-scale model was the discovery of the structure of DNA by Watson and Crick in the early 1950s [51]. Despite their appeal, the **scale similarity** of representation is only feasible in special cases or for certain characteristics. For example, if a model of a city is built to observe the shading effects of buildings, the reduction of scale does not interfere

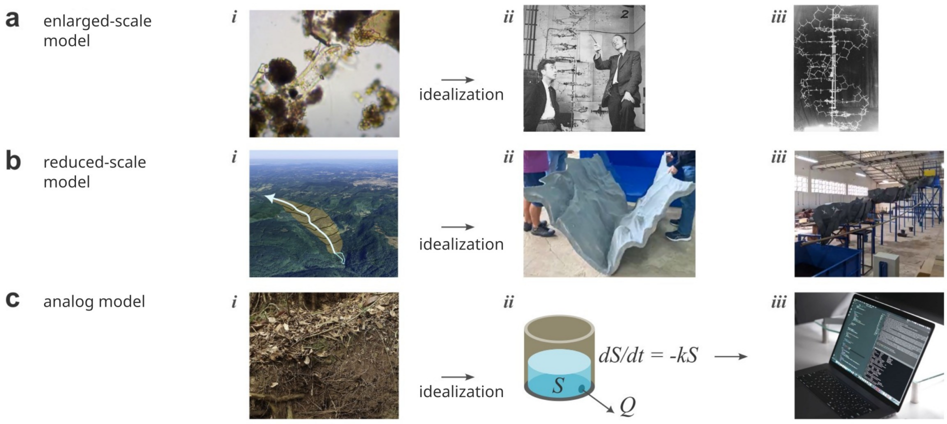

Figure 2.2 — Representation of systems by models. idealization is necessary to represent target systems in sufficiently tractable models. **a** — A famous enlarged-scale model in the History of Science was the double helix model for the DNA molecule, which stores the genetic code of organic cells (detail i); Francis Crick and James Watson handling the model (detail ii); the original DNA model (detail iii). **b** — A reduced-scale model for empirical studies of dam breakage. In this case, the model represents 5.5 km of the valley downstream from the Canastra dam, Canela, Rio Grande do Sul (detail i); the topobathymetric representation of the valley (detail iii) with cross-section modules and steel long beams, filled with fiberglass and resin (detail ii). **c** — A typical analog model of Hydrology for water storage in the soil and groundwater (detail i); the formal analogy (homology) is made with a linear reservoir, as if it were a bucket with a porous outlet at the bottom (detail iii); the model is realized in a digital computer, from the interaction of *hardware* with *software* (detail iii). Credits for the images: (**a**) the author (detail i) and from Chadarevian [51] (details ii and iii); (**b**) the author (detail i) and Flávia Pereira [52] (details ii and iii); (**c**) the author (detail i) and Pinterest (detail iii).

with the shadow patterns produced by light, as the geometry is completely preserved at both scales. However, a reduced-scale water channel or pipe may exhibit viscosity and surface tension effects much greater than those observed at the real scale, making the conversion between scales a non-trivial problem. In fluid mechanics problems like this, the conversion is usually solved through dimensional analysis, which seeks to establish a characterization of the target system that is scale-free, such as the Mach, Reynolds, and Froude numbers.

Depending on the target system in question, representation by reduced or enlarged scale models may not be possible due to some fundamental principle or simply due to a lack of material resources. An epidemiological model at a reduced scale is clearly not possible for ethical reasons, for example. Meanwhile, a reduced-scale model of an environmental system, such as a floodplain or the atmosphere itself, can be very expensive. Given this condition, it is necessary to resort to a form of analogical representation. In other words, it is essential to adopt a modeling approach that establishes a formal analogy, or **homology**, with the target system, that is, an equivalence between the *mathematical structures* of the target system and the model. In Hydrology, this is often achieved by establishing that the soil (or any other compartment of the hydrological cycle) functions *as if* it were a linear reservoir, like a bucket with a porous hole at the bottom, as illustrated in Figure 2.2c. The conjunctive phrase "as if" is crucial, as it establishes the analogy that underpins the idealization of modeling. The implementation of the formal analogy, that is, the realization of its mathematical structure, generally occurs through the programming of digital computers (which is the case for hydrological models), although it is also possible to create material models of the analogous system. In this sense, formal models make the symbolic conveyance of the theory or hypothesis about the target system much clearer than scale models, as they seek to test a mathematical structure based on a supposedly analogous system. Just like deduction, induction, and abduction mentioned in the context of justification of theories in the first chapter, the analogy also constitutes a form of inference, which presents the following logical structure [53]:

1. The objects $O_1, O_2, O_3, ...O_n$ share the properties $P_2, P_3, P_4, ...P_k$.
2. The objects $O_2, O_3, ...O_n$ share the property P_1.

3. Therefore, it is likely that the object O_1 possesses the property P_1.

Thus, analogical inference allows for multiple items to be evaluated, although generally the relationship is made between only two objects – in the case of modeling, the target system and the model. Another characteristic is that, unlike abduction, analogical inference is not a special form of induction, as it does not involve a universal generalization from singular statements. Still, it is also not as secure as deduction, as there is no guarantee of the truth of the consequent statement. For this reason, analogical inference is considered an independent form of inference.

In many cases of research and scientific investigation, obtaining an empirically adequate representation of a given target system is not necessarily the ultimate goal of model construction, but rather the *exploration* of the implications of the theory that the model conveys. In other words, instead of confronting the models with empirical evidence to test or confirm the hypotheses embedded in their structure, they can also serve the epistemic function of *articulating* the theory itself. In this sense, Axel Gelfert introduces the concept of **exploratory experimentation** with models, a process that has the potential to reveal various new hypotheses and elucidations in the theoretical field [54]. The advantage of **exploratory models**, often maintained as **minimalist models** to facilitate understanding, is that the analogy with the target system suggests that unexpected and surprising behaviors of the exploratory model may eventually be empirically observed in the target system, under limit conditions. One example that Axel Gelfert highlights from the History of Science is the experiments with the Lotka-Volterra ecological model, which explored predator-prey dynamics. Although the model did not provide empirically precise predictions, it offered several important qualitative *insights* about the interdependencies among different species in a situation of complete isolation: it was demonstrated that oscillations in populations can emerge even without external interference. In the environmental and hydrological field, for example, exploratory models can contribute to understanding the impacts of *scenarios* never observed in the historical record, such as the ongoing climate changes. In this conception, exploratory models are versatile tools in scientific research, serving various roles, from providing starting points for future investigations, demonstrations of *proof of principle*, formulating potential explanations, to evaluating the adequacy of the model. Furthermore, they are particularly valuable in situations where established theories are in crisis, allowing new paradigms to be proposed based on experimental explorations.

2.3 General systems theory

When we deal with models, a central concept arises, which is the notion of **system** – after all, models convey a theory by representing a target system. A system is defined as *a set of parts with relationships among themselves*. This definition may seem simple, but it carries with it a holistic worldview that instantiates things that transcend materiality. As has been pointed out, by exploring the essence of "things", we enter the field of **Ontology**, which is the study of what exists. The ontological question is: *what exists*? Consider, for example, a classic ontological question: the existence of chairs [55]. Whether chairs exist or not, the answer varies depending on the interpretation of the nature of fundamental elements. From a reductionist perspective, which considers atoms of matter as the only possible entities, chairs are merely collections of atoms and, therefore, *do not exist.* This bottom-up perspective leads to a disquieting conclusion: nothing exists, *not even people*, except matter being scattered in a great flow from nothing to nothing. However, it is evident that chairs do exist; otherwise, we would all be sitting on the floor. Even more evident is the fact that there are people; otherwise, I could not write this text and no one could read it. The solution to instantiate the existence of objects like chairs or people consists of adopting a holistic approach, that is, a top-down view. This perspective understands objects and subjects as entities that **emerge** from the relationship and interaction among their fundamental components, their elements, their parts. In this sense, a chair exists independently of its material, whether it is metal, wood, or plastic. At the same time, it is useless to obtain a pile of wood and expect a chair to emerge from it: **organization** is necessary. A chair would then be the system that emerges from an organized structure that serves the function of providing a seat.

The roots of systems thinking date back to Antiquity, especially in the ideas of Aristotle (384-322 B.C.). This Greek philosopher developed a concept known as **hylomorphism**, which permeates various aspects of his philosophy, ranging from natural science to politics. With this perspective, Aristotle argued that every existing object is composed of both **matter** and **form**, with the latter being essential for the *unification* of the object into a single entity [56]. For example, in living organisms, the body represents the matter and the soul, the form. In politics, citizens would be the matter and the constitution, the form. With the advent of the scientific method in modernity, primarily influenced by the ideas of Descartes, there was a decline in the conception of form as an ontological unifier. In his *Discourse on Method*, Descartes introduced an analytical, reductionist, and mechanistic approach to the world. For example, one of the essential steps of his method to dispel doubts involves isolating difficulties into as many parts as necessary for easier resolution, gradually building the complete vision of the whole, from the simplest to the most complex. In this regard, the focus should remain on the individual parts, with the whole being merely an overlay or linear concatenation. Here lies a **principle of additivity**, which allows understanding larger scales from smaller scales. Descartes illustrates this view by describing the human heart in terms of a hydraulic pump that functions to distribute blood, suggesting that the human body is actually a machine, with each organ performing a specific function. This movement gained traction from Newtonian physics, with a landmark of its peak being Laplace's celestial mechanics and later, in the 19th century, classical thermodynamics, which established blind and relentless laws that describe random and disorganized complexity.

The renaissance of systems thinking in the 20th century was substantially marked by the work of the biologist Ludwig von Bertalanffy (1901-1972). Criticizing the hegemonic mechanistic and reductionist paradigm, Bertalanffy initiated the **General Theory of Systems** starting in the 1920s, although it only consolidated in the 1960s. The influence of biology in this contemporary movement of systems thinking was partially related to the refutation of vitalist theories about living organisms. Given that living beings are completely composed of the same matter as their environment, the need arose to explain the enigma of how simple molecules can form cells, tissues, organs, individuals, and societies. But there were also influences from other theories and disciplines of the time, such as cybernetics and the theory of information, which introduced the concepts of **feedback** and signals among the components of a system. The generality of the theory lies in what Bertalanffy calls **structural isomorphism**, which is the formal analogy (homology) between completely different phenomena in material terms, but which present the same relationships among the parts, that is, the same form. In this sense, the proposal becomes somewhat ambitious, as Bertalanffy suggests that there is a unifying potential for a Science that was overly compartmentalized:

> The systemic point of view has penetrated and proved indispensable in a wide variety of scientific and technological fields. This ultimate fact that it represents an original paradigm in scientific thought (to use Thomas Kuhn's expression) has the consequence that the concept of system can be defined and developed in different ways as required by research objectives. – Ludwig von Bertalanffy [57].

Bertalanffy's theory, in essence, advocates for a holistic understanding of living organisms and systems in general, treating them as **open systems** that constantly interact with the environment and are subject to the flow of matter, energy, and information. This view contrasts with the perspective offered by classical thermodynamics, which focuses on closed systems governed by random disorganization. Open systems, on the other hand, allow for the emergence of homeostasis, metabolism, and steady states, phenomena that, according to Bertalanffy, help explain the apparent violations of the laws of thermodynamics in biology. In the mechanistic view of the world, the fate of any system is rigidly determined by blind laws and its initial conditions. But Bertalanffy emphasizes that this does not occur in open systems, exemplifying with the phenomenon of equifinality[2], which occurs when different initial conditions lead to the same final state, a process observed mainly in the embryonic development of living organisms. Darwinian evolution itself, Bertalanffy points out, also apparently violates the dictates of the second law of thermodynamics, as it allows for the accumulation of information and complexity over time. It is clear that the laws of thermodynamics are not violated in any of these cases, but it is the capacity to import free energy from degrading sources that allows open systems to remain stable against the natural flow of disorder, in a process of constant *self-organization*.

Although Bertalanffy admits that the General Systems Theory can be broadly applied from what he called *verbal models*, he illustrates that *formal models* of systems can be derived from a more or less general mathematical formulation. In this case, this formulation involves a system of simultaneous differential equations. Thus, for n elements characterized by a quantitative measure S:

$$\begin{aligned}
\frac{dS_1}{dt} &= f(S_1, S_2, ..., S_n) \\
\frac{dS_2}{dt} &= f(S_1, S_2, ..., S_n) \\
&... \\
\frac{dS_n}{dt} &= f(S_1, S_2, ..., S_n)
\end{aligned} \tag{2.1}$$

In other words, any variation in S_i is a function of the overall state of the system, which includes all other elements. This formulation also allows for the destruction of the relationship between the parts: it is enough to make the state S of an element solely a function of itself, that is, $dS_i/dt = f(S_i)$. In this case, the system as an ontological entity ceases to exist, and the final state of the whole is completely reduced to the superposition of the states of the individual elements. However, when there are relationships, no matter how trivial they may be, Bertalanffy shows that from the differential equations emerges a rich variety of *end behaviors*, such as exponential growth or decay and processes described by the logistic curve, like saturation and autocatalysis. With just two elements, the system of linear constant coefficients assumes the following

[2] Keith Beven adopted the term "equifinality" to describe the underdetermination problem in hydrological models based on Bertalanffy's General Systems Theory [58].

general form:

$$dS_1/dt = c_{11}S_1 + c_{12}S_2$$
$$dS_2/dt = c_{21}S_1 + c_{22}S_2$$

In this simple system, the Taylor series expansion allows solutions to be obtained for S_1 and S_2 through mathematical analysis. The different solutions demonstrate the emergence of different **stability conditions** (Figure 2.3 a). This can be visualized graphically through a **phase plane** in which the trajectories of the states of the two elements are drawn, as well as the evolution of the variables in the temporal domain. Thus, the multiple configurations of parameter values (coefficients) and also of initial conditions reveal the **attractors** that act on the system. For example, under certain conditions, the system is stable and migrates from a *source* to a final state (S_1^*, S_2^*), or node, in a *drain*. This can occur smoothly or through **damped oscillations** (Figure 2.3 b, details *i* and *ii*). Under other conditions, the system is unstable, migrating eternally, either in a fixed direction ($-\infty$ or $+\infty$) or through **amplified oscillations** (Figure 2.3 c, details *i* and *ii*). Alternatively, the system may exhibit **stable oscillations**, remaining eternally in a loop when visualized in the phase plane (Figure 2.3 b, detail *iii*). A famous example of stable oscillations is the non-linear system of Lotka-Volterra mentioned in Section 2.2, which simulates the interaction between the prey populations S_1 and the predator population S_2:

$$dS_1/dt = r_1S_1 - c_1S_1S_2$$
$$dS_2/dt = -r_2S_2 + c_2S_1S_2$$

where r_1 and r_2 are growth and decay rates, respectively. The product S_1S_2 aims to represent the rate of encounters between prey and predators, weighted by the coefficients c_1 and c_2, creating a feedback that balances the populations in cycles. Bertalanffy emphasizes that these are simple examples that help illustrate the versatility of systems in representing patterns observed in nature. If the system of interest has various relationships or even greater complexities, such as partial terms, the analytical solution of the model can be extremely difficult or even impossible, necessitating the application of numerical methods to solve the equations in any domain, whether in time or space (or both).

Although Bertalanffy admits that the General Systems Theory can be broadly applied from what he called *verbal models*, he illustrates that *formal models* of systems can be derived from a more or less general mathematical formulation. In this case, this formulation involves a system of simultaneous differential equations. Thus, for n elements characterized by a quantitative measure S:

$$\begin{aligned}
\frac{dS_1}{dt} &= f(S_1, S_2, ..., S_n) \\
\frac{dS_2}{dt} &= f(S_1, S_2, ..., S_n) \\
&... \\
\frac{dS_n}{dt} &= f(S_1, S_2, ..., S_n)
\end{aligned} \tag{2.2}$$

In other words, any variation in S_i is a function of the overall state of the system, which includes all other elements. This formulation also allows for the destruction of the relationship between the parts: it is enough to make the state S of an element solely a function of itself, that is, $dS_i/dt = f(S_i)$. In this case, the system as an

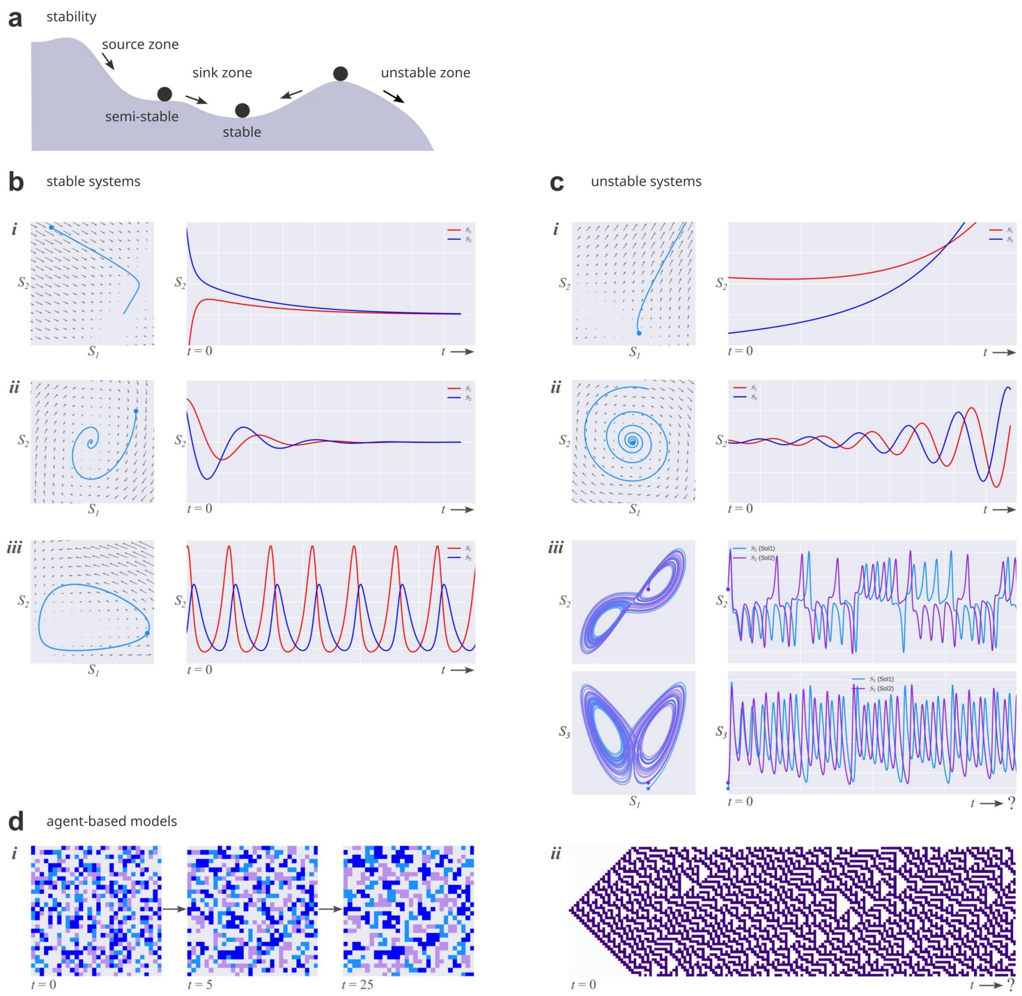

Figure 2.3 — Stability and behavior of systems. A system is defined by a set of parts with relationships among them. Since it is the relationships that unify the whole, similar final behaviors emerge in different scientific fields. **a** — The behavior of a system can be classified as stable or unstable, depending on its initial and boundary conditions. **b** — Stable systems with two elements (S_1 and S_2): exponential decay (detail i) and damped oscillations (detail ii) make an asymptotic movement toward an equilibrium point, being variations of the homogeneous linear system. A stable system can also exhibit eternal oscillations around the equilibrium point, as in the case of the Lotka-Volterra predator-prey model, a non-linear system (detail iii). **c** — Unstable systems with two elements (S_1 and S_2): exponential growth (detail i) and amplified oscillations (detail ii) make a movement toward $+\infty$ or $-\infty$ (or both), also being variations of the homogeneous linear system. Instability can also be chaotic, as in the Lorenz model, a non-linear system with three elements (S_1, S_2, and S_3; detail iii). In the case of the chaotic system, two solutions (in blue and purple) are visualized for very close initial conditions, but diverge in the long run (high sensitivity to initial conditions). **d** — Agent-based models illustrate that complex behaviors can emerge from simple interactions in the immediate neighborhoods of each agent. The Schelling model (detail i) illustrates the emergence of ordered clusters from random initial conditions. Wolfram's Rule 30 (detail i) illustrates computational irreducibility: the only way to understand the final behavior of the system is to simulate the model step-by-step.

ontological entity ceases to exist, with the final state of the whole completely reduced to the overlay of the states of the individual elements. But when there are relationships, however trivial, Bertalanffy demonstrates that from the differential equations emerges a rich variety of *final behaviors*, such as exponential growth or decay and processes described by the logistic curve, such as saturation and autocatalysis. With just two elements, the linear system of constant coefficients takes the following general form:

$$dS_1/dt = c_{11}S_1 + c_{12}S_2$$
$$dS_2/dt = c_{21}S_1 + c_{22}S_2$$

In this simple system, the Taylor series expansion allows solutions to be obtained for S_1 and S_2 through mathematical analysis. The different solutions demonstrate the emergence of different **stability conditions** (Figure 2.3 a). This can be visually represented graphically from a **phase plane** where the trajectories of the states of the two elements are drawn and also by the evolution of the variables in the time domain. Thus, the multiple configurations of values of the parameters (coefficients) and also the initial conditions reveal the **attractors** acting on the system. For example, under certain conditions, the system is stable and migrates from a *source* to a final state (S_1^*, S_2^*), or node, in a *drain*. This can occur smoothly or through **damped oscillations** (Figure 2.3 b, details i and ii). Under other conditions, the system is unstable, eternally migrating, either in a fixed direction ($-\infty$ or $+\infty$) or through **amplified oscillations** (Figure 2.3 c, details i and ii). Alternatively, the system may exhibit **stable oscillations**, remaining forever in a loop when viewed in the phase plane (Figure 2.3 b, detail iii). A famous example of stable oscillations is the non-linear system of Lotka-Volterra mentioned in Section 2.2, which simulates the interaction between the populations of prey S_1 and the population of predators S_2:

$$
\begin{aligned}
dS_1/dt &= r_1 S_1 - c_1 S_1 S_2 \\
dS_2/dt &= -r_2 S_2 + c_2 S_1 S_2
\end{aligned}
$$

Where r_1 and r_2 are growth and decay rates, respectively. The **strange attractor** of this system is illustrated in two phase planes in Figure 2.3c, in detail iii. The detail also illustrates two trajectories that started very close but take on different behaviors in the long run. On the other hand, the problem of computational irreducibility relates (mainly) to the application of **agent-based models**. The agent-based models represent systems through fundamental elements – the agents – that follow simple rules in their immediate neighborhood. When represented in a regular matrix, such as a board, these models are called **cellular automata**. An exemplary agent model is the **model of segregation by Schelling** [59]. In this model, the agents have qualitative categories. At each time step, the agents evaluate their immediate neighborhood and decide whether to move or stay, depending on their tolerance rate with agents of different categories. This system with simple rules spontaneously produces organized clusters, as illustrated in Figure 2.3d, detail i. In this line, Stephen Wolfram demonstrates that simple rules in certain systems can generate unpredictable complexity, accessible only through simulations that evaluate the system *step-by-step* [60]. This computational law arose from experiments with cellular automata that followed simple rules of Boolean conversion (from 0 to 1 and vice versa) based on binary representation. Some rules, such as **Rule 30**, exhibit computational irreducibility (Figure 2.3d, detail ii). Both deterministic chaos and computational irreducibility convey the same message: these are issues that cast doubt on the *predictive capacity* of theories when it comes to dynamic and non-linear systems. At the same time, they are concepts that reinforce the importance of empirical adequacy and estimation of epistemic uncertainties so that policies are based on *evidence*, not just on *theories*, as explored in the previous chapter.

2.4 Systems dynamics

The advent of the paradigm sistêmico in the 1960s allowed for the emergence of the discipline of **Systems Dynamics**, which is, in fact, a fusion of control engineering with management science and decision-making. Systems Dynamics, as the name indicates, studies the evolution of complex systems over time. Furthermore, John Sterman argues that Systems Dynamics is fundamentally a method for *learning* about the behavior of complex systems [44]. As emphasized in the second epigraph of this chapter, Sterman asserts that models ultimately allow for gaining *insights* into the structure and behavior of systems, exploring **leverage points** to achieve desired outcomes in policy formulation and decision-making. The ability of a model to accurately predict the state of a given system, from this perspective, is not as important as understanding its functioning and developing action strategies. The creation of this discipline is attributed to Jay Forrester (1918 - 2016), who sought to understand the behavior of systems from a technological and managerial perspective, that is, focused on solving problems and achieving pre-established goals, from capturing a market share by an industry to reducing the concentration of greenhouse gases in the atmosphere. This is illustrated by his account that the fundamental ideas of this discipline emerged from a challenging industrial problem at *General Electrics*, related to long-term fluctuations in jobs. After studying the decision-making processes of the industry, he used a simple simulation, with pencil and paper, which revealed a potential for oscillations in the internal organization of the system:

> (...) even with constant incoming orders, one would get employment instability as a consequence of commonly used decision-making policies within the supply chain. That first inventory-control system with pencil and paper simulation was the beginning of the System Dynamics field – Jay Forrester [61].

Despite its beginnings with pencil and paper, Systems Dynamics clearly requires the use of digital computers to simulate highly complex models in industrial, urban, social, economic, environmental, and global contexts. Today, the application of concepts and the construction of Systems Dynamics models are typically carried out using software such as `Stella` and `Vensim`, which utilize advanced graphical interfaces to facilitate the development of complex systems. A pioneering example of global-scale application is the model `World3`, whose simulations were explored by Donella Meadows in *Limits to Growth*, as part of Forrester's research group at MIT [45]. This application, not solely environmental, significantly marked a new era of economic modeling, with the emergence of Neoclassical Environmental Economics. This discipline moved away from considering the economic system as an isolated entity and began to include the environment and its interactions with the economic system (material balance model), a development that, a few years later, would result in the proposal of a new economic paradigm, Ecological Economics (presented in detail in Chapter 4).

Systems Dynamics formalizes the basic architecture observed in hydrological and environmental models. This architecture, in philosophical terms, is a singular ontology that consists of **compartment model**, illustrated in Figure 2.4a. In Hydrology, it corresponds to the reservoir or "bucket" model. This approach has consolidated in the environmental field, primarily due to the ease of abstraction and the (relative) low computational demand. Another contributing aspect is that empirical evidence about environmental processes often results from aggregated processes, such as the flow of a river or the concentration of a substance in water or air, a fact that is changing with the advent of high spatial and temporal resolution remote sensing technologies. However, Sterman argues that compartment models are not the only form of representation in Systems Dynamics. It also allows for architectures with disaggregated, heterogeneous, or even individualized parts, such as the previously mentioned agent-based models [62]. In light of this, Sterman establishes a pragmatic attitude, arguing that the decision regarding the architecture of the model should be considered from the perspective of the problem being evaluated, but without losing the ability to manage the model with ease. As an example, he mentions that the SIR epidemiological model[3] is a compartment model that exhibits practically the same aggregated final behavior as any other more detailed version. The justification for introducing heterogeneities, such as age groups,

[3] The acronym SIR stands for Susceptible, Infectious, and Recovered.

spatialization, or even agents that follow more or less social distancing rules, should reside in the final purposes of the study, within the scope of relevant recommendations for policy formulation and decision-making. Otherwise, one incurs in a practically infinite regression of details: after all, why model only the host agents if it is possible to model their organs, cells, and even the bacteria or viruses themselves? Another relevant issue regarding detailed architecture is its high computational demand. Although currently accessible and somewhat seductive, Sterman emphasizes that highly detailed simulations with long simulation times introduce cognitive biases in the modeling process, especially in the iterative component. In this case, there arises resistance both to reviewing deeper conceptual aspects and to diagnosing the model through sensitivity and uncertainty analyses, which require many simulations.

In the compartment architecture, the **causal structure** of the modeled system is defined by the arrangement of compartments connected by flows that can be material (transfer rates) or informational (positive and negative feedback loops). From the Aristotelian perspective, the *matter* of the system consists of the compartments, while the *form* of the system consists of the flows. Thus, two identical sets of compartments, when connected by different material and informational flows, reveal themselves to be completely different systems. In Systems Dynamics jargon, the emphasis on form is often expressed by the fact that *the causal structure of a system defines its behavior*. The model should initially be visualized through a **causal loop diagram**, as shown in Figure 2.4a. Here, it is crucial to properly establish the **system boundary** that the model represents, that is, which flows beyond which the subsequent compartments do not have significant causal effects on the modeled system[4]. A compartment consists of a *level* of a state variable S that accumulates over time, meaning it has *memory*. An easy way to identify a level is to consider what happens if the material flows cease: in this situation, the levels in the compartments continue to exist, inert. The only way to change the level is through the action of the material flows. The level in the compartments is governed by some **conservation principle**, generally the conservation of *mass*[5]. In practice, this implies the application of a **balance equation**, where any variation in the level of a compartment results from the net effect of the input rates (positive) minus the output rates (negative). Mathematically:

$$\frac{dS}{dt} = I - O \tag{2.3}$$

In which the material input flows I and output flows O are rates of change of the level S, and have units of S divided by the adopted time unit. A compartment can have multiple input and output flows, with Equation (2.3) being the simplest possible version. These flows are defined as *functions* of both the **state variable** S (when feedback exists) and of **exogenous variables** Υ (outside the system boundaries[6]) and a set of **parameters** Θ (constants adjusted to reproduce the expected behavior of the system). In general terms:

$$\begin{aligned} I &= f(S, \Upsilon_I, \Theta_I) \\ O &= g(S, \Upsilon_O, \Theta_O) \end{aligned} \tag{2.4}$$

By including feedback, the equations that define the material flows also capture the flows of *information* that connect the compartments. Ultimately, they capture the structure

[4] As it is a decision, the boundary design has the dangerous potential to be a premise of neglect, to use Musgrave's term.

[5] In environmental models, it is generally assumed that water is an incompressible fluid of constant density, which enables a simple volumetric balance of water.

[6] In environmental models, the exogenous variables are generally referred to as **external forcings** of the system. In a typical hydrological model, for example, precipitation is an exogenous variable.

of the system, and thus, its final behavior. The behavior of the system is so sensitive to them that, to some extent, the flow equations become intertwined with much of the hypotheses postulated by the theory that the model is conveying[7]

Equation (2.3) expresses the balance of a compartment as an instantaneous and *continuous* process over time, which generally corresponds to the expectations for the modeled target system. For example, the volume of water in a bathtub that is being filled by a faucet increases continuously, not in discrete jumps. Other systems, such as the population in an ecological model, exhibit discrete transitions over time as new generations replace the previous ones. In one way or another, it is impossible to program a digital computer to solve continuous differential equations directly, necessitating the application of numerical methods. This technological limitation of digital computers, while allowing significant advances in other aspects, such as multifunctionality, leads to the so-called **numerical integration problem**. In essence, this problem consists of the **truncation error** associated with the numerical scheme used in modeling. In the case of the balance, this problem involves the difficulty of accurately determining the level S_{t+1} from the known level S_t and the selection of a time interval Δt. After all, how do you calculate the average of the input and output flows during the time interval? Especially when there is feedback, any minimal variation in S directly influences the rates of input or output flow. In light of this issue, Jay Forrester advocates the need to sacrifice the numerical accuracy of simulated results in favor of gaining useful knowledge about the target system [63]. Forrester's guidance, which can be viewed as a *pragmatic convention*, suggests defining a time interval Δt that is sufficiently small relative to the time scale of the modeled flows and then applying the **Euler method** for numerical integration. Figure 2.4d illustrates the truncation error in the numerical solution of the differential equation $dS/dt = -kS$ (a linear reservoir), whose analytical solution $S = S_0 e^{kt}$ is easily obtained. In this case, the Euler method was applied with different time intervals Δt, demonstrating the improvement in integration with smaller intervals. For more complex systems without an analytical solution, it is expected that adopting a sufficiently short time step will ensure that the flow between one moment and another is approximately constant.

The choice of the Euler method for numerical integration is certainly controversial, as there are other numerical methods that are recognized as more efficient (such as the Runge-Kutta methods), but that require greater computational demand. John Sterman advances this debate by establishing the **temporal insensitivity principle**[8]: a crucial test that a model must pass is to demonstrate that different time intervals do not influence (for practical purposes) the results of the simulations [44]. After all, the results of a model sensitive to the time interval defined in numerical integration are devoid of theoretical meaning. As long as the model shows numerical instabilities depending on the time step, it is necessary to adopt progressively smaller time steps until reaching behavior that is independent of the chosen time interval. In the extreme case that the behavior of a modeled system fails to remain stable across the range of viable time intervals with the available technology, then one must consider using a more efficient numerical integration method. In the case of the Euler method, the numerical arrangement of finite differences from Equation (2.3) exhibits the following form:

$$S_{t+1} = S_t + I_t \Delta t - O_t \Delta t \quad \forall t \tag{2.5}$$

That is, it is assumed that the input flows I and output flows O are constant during the course of the time step Δt, with the value of the rates always computed at time t

[7]Evidently, the underlying theory is also represented by the instantiated compartments, by the boundary design, and even by the balance equations.

[8]The term principle of insensitivity is my own.

and then *extrapolated* to $t+1$. For a compartment with N input flows and M output flows:

$$S_{t+1} = S_t + \sum_{i}^{N} I_{t,i}\Delta t - \sum_{j}^{M} O_{t,j}\Delta t \quad \forall\, t \tag{2.6}$$

Starting from the indexing of t, the algorithm to simulate the system on a computer basically consists of inserting Equation (2.5) within a loop[9]. This loop then iterates through all values of t, incrementally calculating the states of the levels and subsequently updating the value of the flows based on the values from the previous step. In addition to the flow balance equations, Forrester suggests that a procedural model of a system (i.e., the computer code itself) should also include **auxiliary equations** and **supplementary equations**. The auxiliary equations are derived directly from the balance and flow equations, implemented to simplify the understanding of the computational steps by humans. The supplementary equations, on the other hand, define variables of interest that are not part of the modeled system, such as statistics accumulated over time or in moving time windows.

As previously mentioned, understanding the *structure* of a modeled system is key to predicting its *behavior*. In this regard, the structural isomorphism postulated by Bertalanffy becomes, within the realm of Systems Dynamics, what Donella Meadows refers to as a "zoological garden of systems": a set of systems that exhibit *archetypical behaviors*, which can be generalized across a wide range of real examples [43]. A good starting point in this context is to consider the simplest possible model, which is one that has a single compartment, governed by Equation (2.3) (Figure 2.4b). Although simple, different behaviors manifest depending on the *dominance* of one flow over another. In the case where the inflow I predominates over the outflow O, the level S of the compartment will tend to increase (detail *i* in Figure 2.4b). If there is positive feedback, the pattern will be that of **exponential growth curve**. Conversely, if the outflow predominates, the tendency will be for the level S to decrease (detail *ii* in Figure 2.4b). Here, the existence of feedback produces **exponential decay curve**. A concrete example of this archetypical system is a culture of cells (such as bacteria or fungi), growing on a Petri dish without significant nutritional limitations. The more microorganisms reproduce (inflow), the more new generations are added to the total population, which grows exponentially. Simultaneously, the increasing confinement of cells leads to the accumulation of toxic waste from their own metabolism, which also increases mortality (outflow). These two flows act as a **reinforcing loop** and a **balancing loop**, competing for dominance over time, producing patterns more intricate than mere growth or decay, such as the **logistic curve** (detail *iii* in Figure 2.4b).

A second step in this direction is to consider the behaviors that emerge from the *coupling* of two or more compartments. Meadows explores the basic model with two compartments, particularly when the level S_1 of the first acts as a source of inputs E for the level S_2 of the second (Figure 2.4c). This is the case, for example, when the previously mentioned cell culture has limited nutritional resources. In fact, this arrangement is the archetype of any system producer-consumer, which includes the global economy itself (natural resources and capital). A notable pattern that emerges from this system is the **overshoot and collapse curve**, which occurs when the balancing loop in the consumption of available resources does not exist or is very weak, causing the level of the second compartment to rise rapidly until it exhausts its own source, resulting in a similarly rapid decline (detail *i* in Figure 2.4c). The introduction of **double feedbacks** and **activation thresholds** to mitigate or even suspend the consumption of resources

[9]Note that, depending on the simplicity of the system, a typical spreadsheet can perform the computation, where each row of the table consists of a time step.

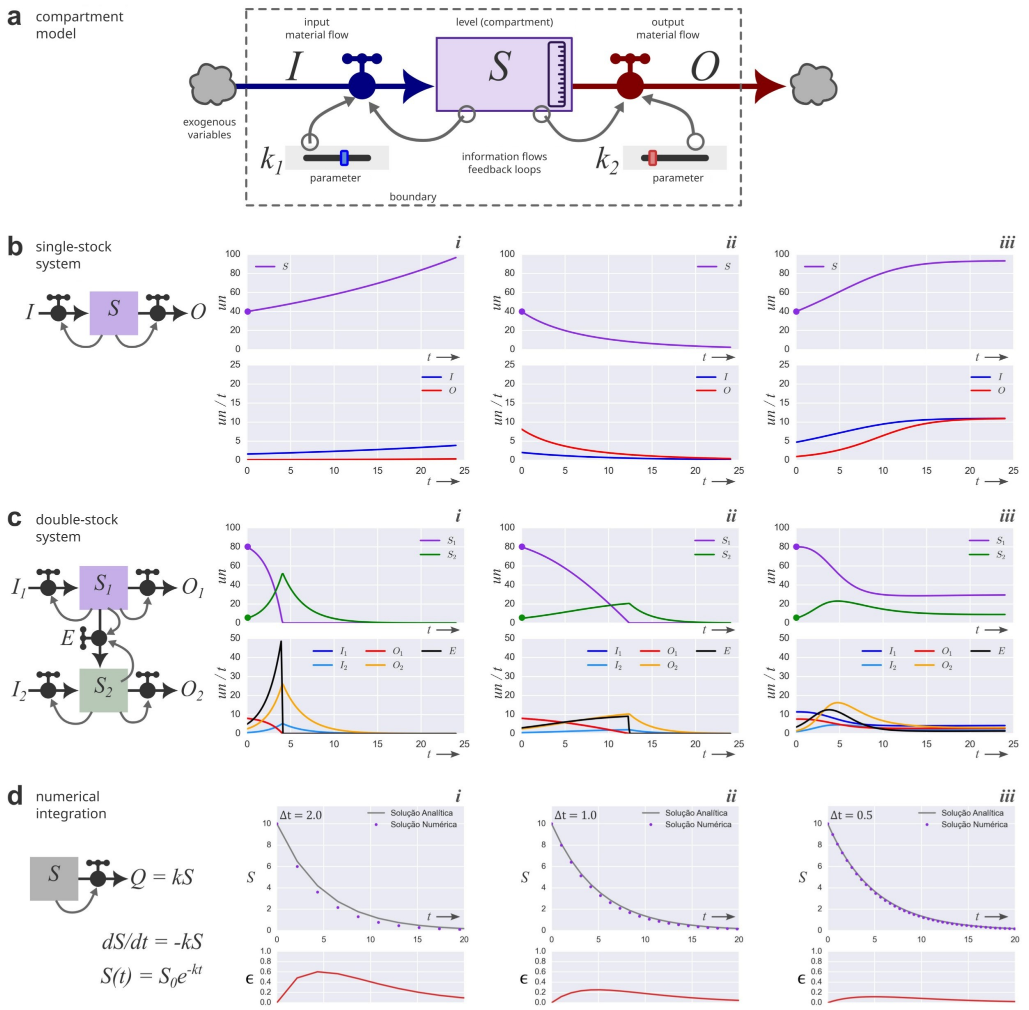

Figure 2.4 — Systems Dynamics and the compartment model. The compartment model consists of the basic architecture for constructing models within Systems Dynamics. The system is solved numerically, revealing complex patterns. **a** — Causal loop diagram: the level of the compartment S changes due to the action of material inflows I and outflows O. Information flows relate the level to the material flows through feedback, which is regulated by parameters of the model (such as k_1 and k_2). The **boundary** of the system should not neglect significant feedbacks with the exogenous variables. **b** — Simulations with the single stock model, with different dominances between material inflows and outflows: exponential growth (detail i); exponential decay (detail ii), and; logistic curve (detail iii); **c** — Simulations with the double stock model, where level S_2 depletes S_1 with an extraction flow E: overload and rapid collapse (detail i); overload and delayed collapse (detail ii), and; sustainable equilibrium (detail iii). **d** — Numerical integration introduces the truncation error ϵ, which can be minimized with sufficiently short time steps Δt (details i to iii). The overall behavior of the system should not be sensitive to the time step Δt.

can either delay the collapse (detail ii in Figure 2.4c) or establish a **sustainable equilibrium** in the long term[10] (detail iii in Figure 2.4c). Moreover, the introduction of **delays** in the flow of information can produce stable oscillations (naturally damped) or unstable oscillations (amplified and chaotic) in the levels. It is easy to see that the complexity and diversity of behaviors grow exponentially as new compartments and feedbacks are introduced into the models. The advantage of Systems Dynamics, with its strictly computational nature, lies in its ability to simulate the system step by step from the basic relationships between the compartments, eliminating the need to explicitly solve systems of non-linear differential equations. Thus, patterns of growth, decay, saturation, collapse, and oscillations simply emerge.

[10] In the case of the global economy, this optimistic scenario is referred to as *prosperous decline* by Odum and Odum [64]

2.5 A prototype of model

With what has been presented so far, we finally arrive at a suitable position to introduce a prototype of a model hidrológico. In this line, the aim here is to establish the basic implications that Systems Dynamics brings to hydrological modeling through an exploratory and minimalist model. Theoretical and practical deepening, both on hydrological processes and on more detailed models, will be articulated in the next chapter.

As highlighted in Section 2.1, every modeling process begins from a *model perceptual*. Thus, the minimalist model arises here from some perceptions, especially the one that a watershed has at least two different ways of *responding* to rainfall events: one faster and the other slower. The **hydrological response** fast manifests in the rises of rivers that occur after rains. The hydrological response slow, on the other hand, is evident during dry times, when rivers continue to flow even after many days or even months without rain. In this line, it is assumed that the fast response is more related to surface processes, while the slow response is associated with subsurface processes. The first relationship derives from the perception that highly impermeable watersheds or those with shallow soils produce large runoff (fast response). Conversely, the springs of streams and wet areas in the valley bottoms of more preserved watersheds reinforce the perception of the role of groundwater in sustaining base flow in rivers during dry times. An important perceptual detail that we can introduce here is that not all rain produces a fast response, as it is necessary to surpass a certain *activation level*, such as interception of water in the canopy or filling the depressions in the terrain that are not connected. Once this threshold is surpassed, the incremental *saturation* of the surface produces an increasingly greater fast response, which will only cease when the surface is again dry. This occurs due to the *level of fragmentation* of the surface – that is, more water is needed to connect isolated pockets of water with the available outlets (in channels or macropores). Regardless of the forms of response, water also needs to travel through a network of channels to reach the outlet of the watershed. This reinforces another relevant perception, that the movement of water implies an attenuation of the response pulses due to energy dissipation effects. Finally, one last perception refers to the outflow of evapotranspiration (`ET`). In this case, it is expected that transpiration from plants through the canopy is the initial flow, followed then by evaporation of water from the surface.

Once a perceptual model is established, we can move on to a conceptual model from the perspective of Systems Dynamics. The diagram of this model is displayed in Figure 2.5a. Thus, the first step in this process is to define the system boundary of the modeled system. In the case of typical hydrological models, the aim is to represent a watershed, which is an area of land traversed by the hydrological cycle. The watershed, therefore, is an open system to water flows that enter through atmospheric precipitation P (rain, snow, and dew) and to outflows, which can occur through both runoff Q and evapotranspiration, denoted here as E. Thus, the flows P and E are maintained as exogenous variables, obtained from **input data** and acting outside the boundary of the target system. It is assumed, therefore, that the internal state of the system does not exert causal influence over the value of these variables[11]. It is clear that, as an outflow, evapotranspiration depends on the water available in the system, but its *potential* flow is determined without causal links. Runoff Q, on the other hand, is a flow calculated by the model itself through the application of its flow equations. Models with this characteristic are often called "rain-runoff" models, although this designation obscures

[11] This assumption becomes increasingly fragile as the scale of the watershed evolves from a small area to larger regions or continents.

the fact that many other flows are calculated to estimate the outflow.

The second step in building a conceptual model involves configuring the reservoirs[12] of the system under study. For this, it is essential to mobilize the concept of hydrological response provided by the perceptual model. To keep the model in a minimalist condition, we define only two response reservoirs: one for the fast response, S_1, and another for the slow response, S_2. These reservoirs are vertically coupled, so that the water must pass through the fast response reservoir (upper) before reaching the slow response reservoir (lower). This scheme aims to represent the water balance in the soil, intuitively relating the fast response to surface processes and the slow response to subsurface processes. However, due to the simplification of the model, this interpretation should be considered with caution: the representation in only two reservoirs is essentially a synthesis of several subprocesses that could be more specifically defined in a more detailed model. In addition to the coupled reservoirs, a third reservoir, S_3, collects both the fast and slow flows, acting as a filter over the signal of both. The purpose of this reservoir is to model the effects of attenuation and storage during the propagation of runoff through the channel network before reaching the watershed outlet. Just as in the case of the water balance in the soil, this reservoir encompasses several subprocesses that, in a more complex model, could be explored in detail. Considering that the area of the watershed is constant, it is convenient, though not mandatory, to express the levels of the reservoirs S_i in mm of water column and the flows in $mm/\Delta t$. The reservoirs S_1 and S_3 (surface and channel network, respectively) have unlimited total capacity, while the reservoir S_2 (soil and subsurface) reaches its maximum capacity from a certain level, such that:

$$S_{2,t} \leq s_{2,\max} \quad \forall; i, t \tag{2.7}$$

Where $s_{2,\max}$ is the *maximum storage capacity* of S_2, a parameter expressed in level units (mm).

The third step, finally, is to define the flow equations that govern the water balance in each reservoir. In this sense, all three reservoirs operate as a **linear reservoir**, which implies that they exhibit an outflow Q_t directly proportional to the level S_t, that is:

$$Q_{i,t} = \frac{1}{k_i} \cdot S_{i,t} \quad \forall\, i, t \tag{2.8}$$

Where k_i is a parameter with time units that is equivalent to the mean residence time of the reservoir S_i. This implies that the *greater* the value of k_i, the more *slow* its emptying is, as illustrated in Figure 2.5b (details i and ii). A linear reservoir is analogous to a water tank with vertical walls and a porous outlet, which allows a laminar flow directly proportional to the water column. In the case of the fast response reservoir S_1, the outflow Q_1 is the flow of vertical transfer of water to the reservoir S_2, interpretable as the *infiltration* from the surface into the soil, with the necessary effectiveness caveats. In the case of reservoir S_2, the outflow Q_2 is the very slow response of the watershed, interpretable as the *base flow* that the soil produces directly into the drainage network. Finally, the outflow Q_3 is the final *outflow rate*, resulting from the attenuation of the runoff through the propagation process. The rapid outflow from reservoir S_1, denoted by R, has a specific formulation, corresponding to a hypothesis of how rapid processes, such as surface runoff q_{si}, develop in the watershed. Just like in the outflow of a linear reservoir, R is directly proportional to the stored level, with the difference that the level must exceed a minimum value s_a before it begins to overflow:

$$R_t = \begin{cases} 0 & \text{if} \quad S_{1,t} \leq s_a \\ c \cdot (S_{1,t} - s_a) & \text{if} \quad S_{1,t} > s_a \end{cases} \tag{2.9}$$

[12] In the context of hydrological models, I will use the term *reservoir* as a synonym for *compartment*.

Where s_a is the **activation level** of the fast response, a parameter of the model expressed in level units; and c is a *runoff coefficient*, with units of t^{-1}. However, unlike the linear reservoir, the value of c is not constant, but rather a function of the level S_1 itself. For the purposes of this chapter, we will simply establish that the fast response R results from a *saturation process* of the reservoir S_1, such that:

$$c = \frac{(S_{1,t} - s_a)}{(S_{1,t} - s_a) + s_c} \frac{1}{\Delta t} \quad \forall\, t \tag{2.10}$$

Where s_c is the **fragmentation level**, a parameter with the same units as the level S_1 that regulates the speed of the saturation process. The level s_c represents 50% connectivity, so that the greater the value of s_c, the more slowly the saturation of the reservoir occurs. As the reservoir level increases, the runoff coefficient c asymptotically approaches 1[13], as illustrated in Figure 2.5b (details *iii* and *iv*). The term $1/\Delta t$ was retained in the definition of c to clarify its units, even though it is effectively eliminated in the balance equation (Equation (2.5)). By substituting Equation (2.10) into Equation (2.9), we arrive at the flow equation for R (in the case of $S_{1,t} > s_a$)[14]:

$$R_t = \frac{(S_{1,t} - s_a)^2}{(S_{1,t} - s_a) + s_c} \quad \forall\, t \tag{2.11}$$

In this minimalist model, the external evapotranspiration flow E acts on the soil water balance, affecting reservoirs S_1 and S_2, such that $E = E_1 + E_2$. The drainage of water, in this case, occurs from the bottom up, meaning that the flow E only starts to act on the upper reservoir S_1 when the lower reservoir S_2 is empty. That is, the outflow E_2 in S_2 corresponds to the transpiration of plants, which removes water from the soil, and the outflow E_1 in S_1 corresponds to surface evaporation. Thus, it is noted that the reservoirs S_1 and S_2 are *simultaneously depleted* by more than one outflow. For S_1, the outflows are E_1, R, and Q_1. For S_2, the outflows are E_2 and Q_2. This is a good point to introduce two practical problems in compartment models within Systems Dynamics, which are the **congested output problem** and the **simultaneous depletion problem**. The first problem consists of the difficulty in determining a given outflow $O_{t,j}$ that *is also an input* in a compartment with limited storage capacity. In the case of the minimalist hydrological model, this occurs in the flow Q_1 (infiltration), which goes from S_1 (surface) to S_2 (soil and subsoil). However, since the reservoir S_2 is limited by $s_{2,\max}$ (Equation (2.7)), there is a negative feedback from S_2 on Q_1, which can be interpreted by the notion that the infiltration flow is *congestible* by the moisture present in the soil. After all, if the soil pores are already filled with water, it doesn't matter how much water is available for infiltration stored on the surface. The **actual flow** $Q_{1,t}$, thus, is obtained by confronting the **potential flow** outflow $Q^*_{1,t}$ with the **maximum flow** possible input to the next reservoir, which in the case of S_2 is defined by the **storage deficit** $D_{2,t}$ divided by the time step Δt:

$$Q_{1,t} = \begin{cases} Q^*_{1,t} & \text{if} \quad Q^*_{1,t} \leq D_{2,t}/\Delta t \\ D_{2,t}/\Delta t & \text{if} \quad Q^*_{1,t} > D_{2,t}/\Delta t \end{cases} \tag{2.12}$$

Where $D_{2,t}$ is calculated by:

$$D_{2,t} = s_{2,\max} - S_{2,t} \tag{2.13}$$

[13]Equation (2.10) presents a typical form of saturation processes found in distinct fields, such as the **Michaelis-Menten Equation** in the kinetics of chemical reactions.

[14]A cautious look reveals that Equation (2.11) has exactly the same structure as the empirical formula of the CN method proposed by the Soil Conservation Service to estimate effective runoff from rainfall events and types of land cover. This is an intriguing fact that deserves further study.

Componente	Nome	Dimensão	Unidade	Categoria
S_1	quick response reservoir (surface)	L	mm	level
S_2	slow response reservoir (subsurface)	L	mm	level
S_3	drainage network reservoir	L	mm	level
P	precipitation	L/T	mm/h	flow (exogenous)
E	potential evapotranspiration	L/T	mm/h	flow (exogenous)
R	quick runoff ($S_1 \to S_3$)	L/T	mm/h	flow
Q_1	infiltration ($S_1 \to S_2$)	L/T	mm/h	flow
Q_2	slow runoff ($S_2 \to S_3$)	L/T	mm/h	flow
Q_3	outflow from S_3	L/T	mm/h	flow
E_1	evaporation	L/T	mm/h	flow
E_2	transpiration	L/T	mm/h	flow
k_1	residence time of S_1 (surface)	T	h	parameter
k_2	residence time of S_2 (subsurface)	T	h	parameter
k_3	residence time of S_3 (drainage network)	T	h	parameter
s_a	activation level of quick response	L	mm	parameter
s_c	fragmentation level of S_1	L	mm	parameter
$s_{2,\text{max}}$	maximum capacity of S_2	L	mm	parameter

Table 2.1: Summary of the hydrological model prototype developed, listing the components of levels, flows, and parameters. Due to the high degree of aggregation of the model, the names and meanings of the components should be interpreted with caution, as they are in fact effective processes that could be better detailed and disaggregated in more complex versions.

This solution can be generalized for other congestible outflows[15]. When a compartment with limited capacity has *multiple* simultaneous inflows, then the solution needs to adopt an analogous (but inverse) approach to the other problem mentioned, which is the simultaneous depletion problem. This problem, in turn, consists of the difficulty in *preventing negative values* in a level subjected to multiple outflows that is numerically integrated using the Euler method. In a mathematically continuous system, the various outflow rates smoothly act on a given level S_t, so it tends asymptotically toward zero. However, the Euler method, by considering that the rates are constant during a discrete time interval Δt, introduces the risk that S_{t+1} may take on negative values. The solution to this problem is to compute the outflows in three steps. In the first step, the total *potential* outflow O_t^* is calculated by summing the individual potential outflows. More generically, for M potential outflows $O_{t,j}^*$:

$$O_t^* = \sum_j^M O_{t,j}^* \quad \forall\, t \tag{2.14}$$

Next, the *actual* total outflow O_t is determined by confronting the potential flow with the *maximum* possible outflow, which is the value of the level S_t of the reservoir divided by the time step Δt:

$$O_t = \begin{cases} O_t^* & \text{if} \quad O_t^* \leq S_t/\Delta t \\ S_t/\Delta t & \text{if} \quad O_t^* > S_t/\Delta t \end{cases} \tag{2.15}$$

Finally, the third step consists of calculating the values of the *individual* actual outflows. Since the Euler method assumes constant flow rates, the actual outflows are directly proportional to the *allocation* of the potential outflows:

$$O_{t,j} = \frac{O_{t,j}^*}{O_t^*} \cdot O_t \quad \forall\, t \tag{2.16}$$

[15] In this context, it is worth noting that John Sterman argues against using conditional structures like `IF... THEN... ELSE` in simulation code, suggesting that an alternative for Equation (2.12) of the form $Q_{1,t} = \texttt{MIN}(Q_{1,t}^*, D_{2,t}/\Delta t)$ is more robust and readable [44].

Such problems, inherent to the computational nature of Systems Dynamics, indicate the next step in the modeling process: the development of a *model procedural*, that is, a computer program. In the form proposed above, the conceptual model is simple enough to be implemented in a *spreadsheet*, where the columns represent different storage and flow variables and the rows represent the time steps of the simulation. To follow the Euler method, the formulas in the reservoir balance cells should refer to the previous row, in the flow columns. Some auxiliary columns should be created to implement the intermediate steps of the potential flows and maximum flows. Additionally, certain static cells should be kept isolated, such as the values of the parameters $\Theta = \{k_1, k_2, k_3, s_a, s_c, s_{2,\max}\}$, the initial conditions values $\mathbf{S}_{t=0} = \{S_{1,t=0}, S_{2,t=0}, S_{3,t=0}\}$, and the exogenous variables values $\Upsilon = \{P_t, E_t\}$. A more robust alternative, however, is to implement the procedural model from code, such as `C`, `Fortran`, `Python`, etc. A simple code structure should be based on the functional programming paradigm, with three interconnecting steps:

1. importing input data;
2. processing the model, and;
3. exporting the output data.

Each step has its typical characteristics and technical limitations, from defining the format of the data files to using efficient structures offered by the chosen programming language. In this regard, a code presents two fundamental advantages: one cognitive and the other operational. The cognitive advantage is that a code completely makes explicit the equations and the computational algorithm itself[16]. Spreadsheets and other graphical interfaces, in contrast, tend to obscure the formulas and the algorithm structure, making the procedural model somewhat inaccessible in cognitive terms. The operational advantage, on the other hand, is that a code allows the *nesting* of the simulation process in a larger hierarchy of processes, such as running batches (multiple simulations in series or parallel) and coupling with other models (where the output of one is used as input data in another). As we will see later, this operational advantage is essential for the diagnosis and research of models.

Before conducting any computational simulation with the developed hydrological model prototype, it is important to highlight some valuable aspects that the formalization of the perceptual model into a conceptual model brings. As mentioned earlier, Systems Dynamics has an exploratory spirit that seeks not only to make reliable predictions about a target system but also to use models to *learn* about the behavior of this system, thereby identifying useful leverage points in decision-making. That said, let us consider for a moment that the theory conveyed by the proposed model is *justified*, which frees us from the issues raised in Chapter 1. If that is the case, what can we deduce *a priori* about the behavior of the watershed? What are its critical leverage points when considering, for example, the problem of water security[17]?

In light of these issues, the first aspect to note is that **the system is stable**, dominated by feedbacks that cause the levels of the reservoirs to *tend towards zero*, as illustrated by the simulation in Figure 2.5c. In other words: the reservoirs S_1, S_2, and S_3 empty on their own[18], even without any external action (when $P = 0$ and $E = 0$). Unlike ecological, social, and economic systems, no form of exponential growth or oscillations is expected from the proposed system. A second aspect is to classify the

[16] Highly efficient codes often imply a sacrifice in readability, which reinforces the importance of supplementary materials such as documentation and comments.

[17] This problem will be explored in Chapter 4

[18] In physical terms: through the action of gravity.

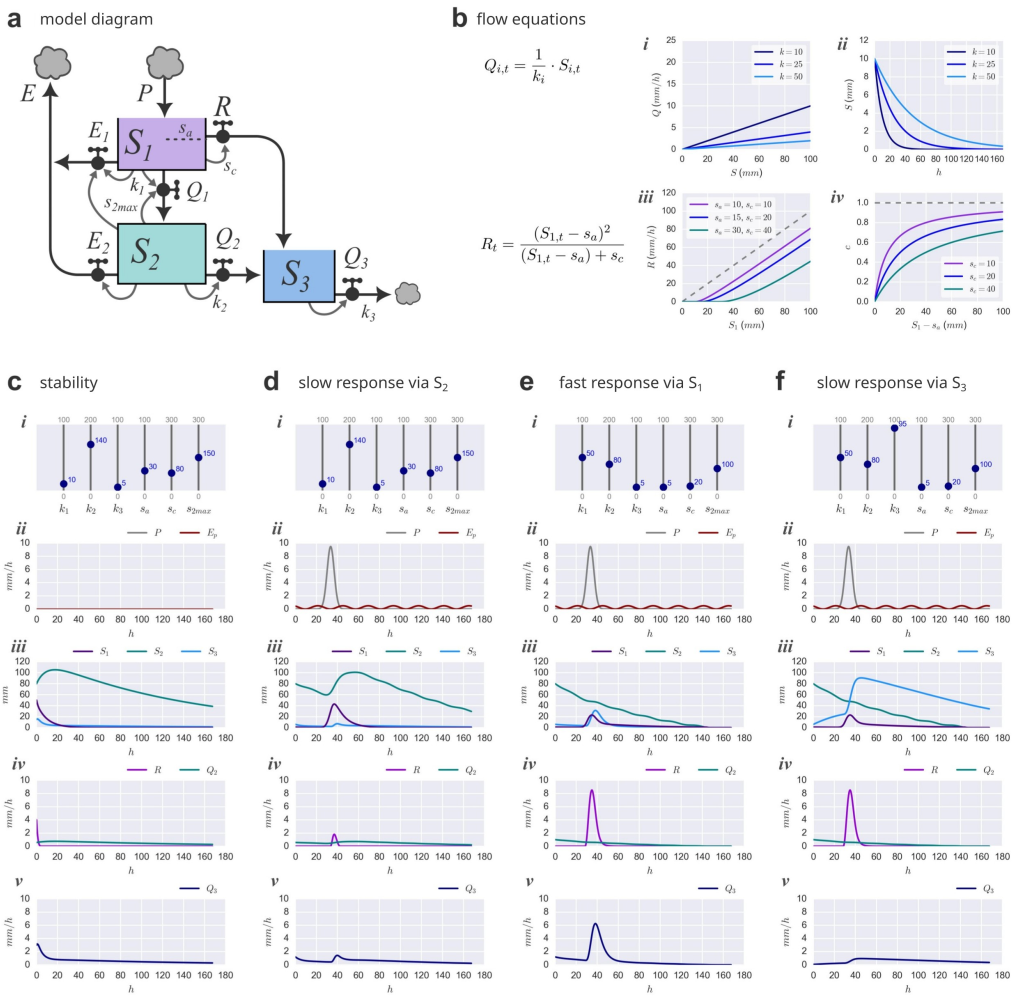

Figure 2.5 — A prototype of a hydrological model and its behavior. The model is kept minimalistic for exploratory purposes. **a** — The structure of the model is designed with three reservoirs (S_1, S_1, and S_3) to represent mechanisms of slow or fast hydrological response. S_1 represents the surface, S_2 represents the subsurface, and S_3 represents the drainage network. With six parameters regulating internal flows, the system is subjected to the exogenous flows of evapotranspiration E and precipitation P. The reservoir S_2 is the only one with limited capacity ($s_{2\text{max}}$). **b** — Two equations regulate the flows (Equation (2.8) and Equation (2.11)). The exponential decay equation for the reservoirs ($Q_i = S_i/k_i$, detail i), where k_i is the residence time. The larger the value of k, the more slowly the reservoir empties (detail ii). The quick response equation ($R = c \cdot (S_1 - s_a)$, detail i) is similar, but c results from a saturation process of the surface ($c = (S_1 - s_a)/(S_1 - s_a + s_c)$, detail ii). This process is regulated by the activation threshold s_a and the fragmentation level s_c. **c** — The system shows stable behavior – it empties by itself, even when P and E are null (detail ii). In this case, S_2 showed a peak as water from S_1 infiltrated (detail iii). But soon after, all reservoirs emptied. **d** — A configuration of parameters (detail i) that defines a slow response through the action of S_2 (subsurface). Here, a rainfall pulse P with a maximum of 9.5 mm/h and a daily oscillation of E (detail ii) result in an attenuated flow Q_3 (detail iv); R has a maximum of only 2 mm/h, while the base flow Q_2 sustains the flow during most of the simulated period (detail iii). **e** — A configuration of parameters (detail i) that defines a fast response through the action of S_2 (surface). The same pulse of P and E results in a typical hydrograph of Q_3, with a maximum of 6 mm/h (detail iv); R shows a maximum of 8 mm/h and the base flow Q_2 is extinguished in about 140 hours (detail iii). **f** — A configuration of parameters (detail i) similar to (**e**), but defining a slow response through the action of S_3 (drainage network). The same pulse of P and E results in an attenuated flow Q_3 (detail iv) – the flow is sustained by damping in the drainage network.

final behavior of the system in terms of sensitivity to the input flow P. At one extreme, we have low sensitivity behavior, characterized by the dominance of the slow response mechanism in S_2 and high residence time in S_3. The simulations in both Figure 2.5d and Figure 2.5f illustrate possibilities of this behavior. At the other extreme, we have high sensitivity behavior, characterized by the dominance of the rapid response mechanism in S_1 and low residence time in S_3, as shown in the simulation in Figure 2.5e. Maintaining the proposed structure and the same exogenous flows P and E, it is clear that behavior will strictly depend on the set of parameters $\Theta = \{k_1, k_2, k_3, s_a, s_c, s_{2,\text{max}}\}$, although different values may eventually result in similar behaviors (high or low sensitivity). For example, it is possible to reduce the sensitivity of the system by increasing both the activation threshold of the rapid response ($\uparrow s_a$) and the residence time in the drainage network ($\uparrow k_3$), or both simultaneously[19]. This brings us to the third and final aspect, which concerns the leverage points of the system in addressing practical problems, particularly ensuring water availability[20]. Therefore, it is concluded that any leverage strategy within the system must aim to reduce its sensitivity to incoming flows, striving for a **regularization effect**. In the soil water balance, this translates into reducing the dominance of the rapid response mechanism: increasing the activation threshold, reducing surface connectivity, and decreasing surface residence time. In the case of the drainage network, the only available alternative is to increase residence time through detention basins, ponds, or even dams.

The aspects raised above illustrate that the simple deductive logic applied to the developed conceptual model allows for the identification of important *insights* about the behavior and the leverage points in the system. Furthermore, if the model captures the essence of the target system, it is expected that more detailed versions do not eliminate the essence of the conclusions drawn, but rather introduce the necessary nuances in the decision-making process based on hydrological models, depending on the problem at hand. Heterogeneities in lithology, pedology, land cover, and topography should certainly expand the range of theoretical understandings and practical recommendations. Despite the leap observed between the perceptual model (mental models) and the conceptual model, it is clear that this reasoning can hide surprises from blind spots and non-intuitive interactions among the parts of the system. Moreover, the assumption that the conveyed theory is justified was a provisional move: it is necessary to test the model against empirical evidence. Thus, the only way to make stronger assertions is to simulate the model using diagnostic techniques, which are presented next.

2.6 Model diagnosis

The **model diagnostics** consists of a vast set of techniques applied to evaluate the *adequacy* of a model. Before empirical justification, which is a crucial test, a model needs to be adequate from the conceptual, technical, practical, and behavioral angles. After all, an extremely well-fitting statistical model in empirical terms can be directly obtained with optimization techniques, such as machine learning. But what can one *learn* about the target system with a statistical model? An overfitted statistical model to the available data, for example, may be useful for *interpolations*, but not for *extrapolations*. Such a model does not contribute much to understanding how the system would *behave* in response to a given leverage policy or a future scenario that has never been observed. The search for learning, in the context of Systems Dynamics, implies that modeling is a process of *deductive* inference, which requires a robust and reliable definition of the antecedent statements (main and auxiliary hypothesis) before producing their consequent statements (simulated results). In this regard, John Sterman suggests a list of twelve general strategies for diagnosing[21] these adequacies, displayed in Table 2.2, enhancing Jay Forrester's initial proposals [44]. However, given the highly practical nature of Systems Dynamics, perhaps the diagnosis *of zero order* of a model is to

[19] This ambiguity is directly associated with the equifinality problem. As presented in Chapter 1, the largest number of possible empirical evidences should be employed to *condition* the behavior of the model.

[20] This practical problem is part of a broader context, namely the issue of water security, which is discussed in more detail in Chapter 4.

[21] Originally, he uses the term "model testing", but here the term "test" will be reserved for when a *rejection criterion* is explicitly defined.

Diagnosis*	Purpose	Procedures
0. Problem Adequacy	Diagnose if the model is suitable for the practical problem being addressed.	Explicitly state the questions implied by the problem that need to be answered by the model.
1. Boundary Adequacy	Diagnose if the exogenous variables of the model do not imply serious causal neglects.	Explicitly define the system and its endogenous and exogenous variables in causal loop diagrams.
2. Structural Adequacy	Diagnose if the structure (including flow equations) aligns with the perceptual model and does not violate basic theoretical principles. The degree of aggregation also needs to be useful in practical terms.	Inspect equations and causal diagrams; clarify decision-making questions related to expected outcomes.
3. Dimensional Consistency	Diagnose if the equations and parameters are consistent and make sense concerning real phenomena.	Inspect equations; dimensional analysis; rationalize the underlying theory.
4. Parameter Distribution	Diagnose how distributions of values align with conceptual and empirical expectations.	Obtain prior distributions from expert opinion.
5. Comparative Studies (System Families)	Diagnose if the distribution of parameters is consistent across target systems of the same family. E.g., different watersheds;	Obtain distributions of parameters consistent with the largest possible number of system members (generalist model).
6. Integration Error	Diagnose if the results are not sensitive to the time step and numerical integration method.	Reduce the time step; change the numerical integration method.
7. Extreme Conditions	Diagnose how robust the procedural model is to very high or very low values.	Simulate the model under synthetic conditions with extreme shocks to values.
8. Sensitivity Analysis	Diagnose how variations in the parameters of the model affect the results, identifying critical parameters for the behavior of the system.	Apply exploratory techniques, such as the Monte Carlo Method, for random sampling in the parameters space. Use search techniques to identify critical scenarios and reveal unusual leverage policies.
9. Anomalous Behavior	Diagnose unexpected behaviors of the model that may indicate formulation errors or important *insights* about the modeled system.	Subject the model to encapsulation testing; discard.
10. Empirical Adequacy	Diagnose if the model can reproduce observed behaviors in the target system, adjusting parameters to improve adherence to empirical data.	Compare the outputs of the model with observed data, using metrics like MAE, RMSE, and coefficients such as determination and KGE.
11. Surprising Behavior	Diagnose if the modeling results surprise the target audience in some way, revising their mental models.	Effectively communicate results; reveal nuances and details; demonstrate non-intuitive mechanisms.
12. Positive Practical Impacts	Diagnose if modeling has brought positive practical impacts on the decision-making process.	Prepare impact indicators from the model; technical documentation; reproducibility.

Table 2.2: Model Diagnostics in the context of Systems Dynamics. — Summary of the twelve "Model Tests" proposed by John Sterman, following the tests suggested by Jay Forrester. Adapted from Sterman [44].

evaluate the **adequacy of the problem** it is being addressed. This primary diagnosis assesses whether the model was *tailored* to answer the truly relevant questions of the practical problems at hand. Although obvious, this diagnosis can pose significant challenges, even forcing modeling teams to *abandon* comfortable modeling strategies as the nature of the problem is understood more deeply. From the decision-making side, that is, the clients of model sellers, this type of diagnosis is essential to avoid uneconomical situations when the model sold is disproportionate.

Among the conceptual diagnostics, it is essential to evaluate the **boundary adequacy**. As illustrated in the prototype of the hydrological model, an important assumption is that the storage and flows of water in the watershed (the endogenous variables) do not influence precipitation and potential evapotranspiration (the exogenous variables). This assumption may be questionable for large watersheds, where evapotranspiration in one region converts into precipitation, either locally (cloud condensation nuclei [65], [66]) or in other areas (flying river effect [67], [68]). Thus, as one shifts from a local scale to a continental scale, neglecting the causal interactions of the meteorological and climatic system with terrestrial hydrological processes tends to become increasingly inadequate. Another basic conceptual evaluation is the **structural adequacy** of the model, both in theoretical terms (physical principles) and practical terms (decision-making). The structure of the model must ensure that physical principles, such as mass conservation, non-negativity of levels, and certain irreversible processes, are not violated. For example, Sterman describes an economic model that produced remarkable results in simulating the leather market, but this occurred because the model reverted produced leather *back* into cows whenever necessary [44]. Similarly, it is important to evaluate whether the level of aggregation of the model meets the pre-established needs of the decision-making process. In hydrological modeling, the aforementioned prototype would hardly be useful for identifying priority action areas within a watershed, due to its highly aggregated nature.

Two other related conceptual diagnostics involve **dimensional consistency** and **parameter distribution**. This evaluation relates to the *meaning* of the parameters in the flow equations and the distribution of their values. When starting from a deductive logic, the flow equations need to make theoretical sense (after all, *they convey a theory*), and their parameters should have consistent names and units that are equivalent (at least in effective terms) to real processes. According to Sterman, parameters and variables with names and units that do not make sense in the real world are *symptoms that the theory about the target system is poorly formulated.* Moreover, the values of the parameters themselves also need to meet conceptual expectations. Since different combinations of parameters can result in similar final behaviors (the equifinality problem), certain combinations of parameters may be empirically adequate but theoretically questionable. For example, in a mountainous watershed without artificial reservoirs, one would expect low attenuation of the flow pulse in the drainage network, or at least lower than in flatter watersheds with floodplains. Thus, **comparative studies** between different systems but of the same *family*, and the definition of prior distributions through **expert opinion**, help to discard inconsistent parameters values[22].

On the technical side, a crucial diagnostic is the **test of numerical integration**, which was described in Section 2.4. It is important to emphasize that *no simulation is informative if its result stems from numerical instabilities.* Thus, this test consists of evaluating whether the behavior of the system adheres to the temporal insensitivity principle. Another diagnostic in this vein is to subject the model to **extreme and boundary conditions**, such as very high values of input and output flows. This is an assessment of technical robustness, as it is under these unusual (but possible) conditions that problems in the procedural model may arise, such as errors in the numerical representation of data structures (*overflow* and *underflow*) and violations of non-negativity in the simulated levels. For example, a discrete level variable (like a population) can be instantiated by a data structure with positive, integer numbers of 16 bits—which offers interesting memory gains, unlike 64-bit floating-point numbers. However, this data structure has an upper numerical limit of 65535—values above this will wrap around to low numbers, causing an *overflow* error that can render simulation results meaningless. In this case, it is necessary to ensure that a more efficient code does not compromise the robustness of simulations under extreme but feasible conditions.

Diagnostics that evaluate the behavior of the model include **sensitivity analysis**, **anomaly detection**, and **empirical adequacy**. These evaluations form a spectrum in terms of justification. On one side, sensitivity analysis seeks to understand how the system responds to changes in its elements, such as input flows and parameters values. In this case, the approach is exploratory, without major commitments to justification. On the other side, empirical adequacy seeks to find the sets of parameters that condition the modeled system to reproduce the observed behavior. Clearly, this is the only test capable of pointing out the need for significant revisions in the modeling process, as it is here that the theory is directly confronted with the available evidence. Still, the rejection of the proposed model is only possible through the modeling paradigm discussed in Section 1.6 (Chapter 1), which applies an encapsulation test of a set of empirically adequate models. A purely confirmatory approach, in contrast, seeks to "calibrate" the parameters of the model in order to identify a single set of

[22]It is expected that expert opinion will *inform* about the prior distribution of parameters with qualitative empirical evidence from their perceptual models that have not been transformed into quantitative data. Still, care must be taken not to turn this process into a **confirmation bias**, keeping an openness for empirical anomalies to act as a means to refute the theories posited by the models. One way to achieve this is to maintain posterior probabilities within a minimum threshold.

parameters deemed adequate.

One way or another, behavior diagnostics generally require a broad assessment of the **parametric space** Ω_Θ, the mathematical space with N dimensions created by the N instantiated parameters in the conceptual model. Here, a technical barrier arises, which is the **dimensionality problem**: the difficulty of thoroughly evaluating the parametric space Ω_Θ in a reasonable simulation time. For example, consider a **exhaustive sampling**[23] approach with M regular intervals across the estimated range for each parameter Θ_i, with $i \in \{1, ..., N\}$. It follows that the number of simulations n_s in a exhaustive sampling approach is $n_s = M^N$. This number can grow exorbitantly quickly: with $M = 100$, one would need to simulate the proposed hydrological model ($N = 6$) one trillion times, $100^6 = 1,000,000,000,000$. With a simulation time of one second, this evaluation would take around *31 thousand years* to complete. A simulation time of one millisecond would require 31 years. To be practical, exploration in the parametric space Ω_Θ should last a maximum of a few days, preferably a few hours or minutes.

More powerful computers greatly help in overcoming the dimensionality problem, but in practice, the sampling strategy also employs more efficient methods than exhaustive sampling, which can be divided into **exploratory techniques** and **search techniques**, although they are largely related. Exploratory techniques are variations of the Monte Carlo Method, mentioned in Section 1.3, that perform random sampling over the expected value ranges for the parameters Θ. The goal here is merely to reveal regions of the parametric space. If a prior distribution for the parameters is available (from expert opinion, for example), this sampling can be *weighted by the density* of the distribution, which directs the exploration more towards certain regions than others. An efficient variation of the Monte Carlo Method consists of random sampling *without replacement*, known as **Latin Hypercube Sampling**, which ensures a more spaced distribution of the sampled parameter sets, avoiding the risk of redundant samples. Search techniques, in turn, consist of optimization techniques aimed at identifying regions of the parametric space that meet predefined specifications regarding the behavior of the system. In the jargon of Operations Research, these techniques maximize or minimize a given **objective function**. A wide variety of optimization algorithms can be implemented for this purpose, such as linear programming, dynamic programming, gradient ascent, evolutionary algorithms, Markov chains, etc. The criterion for selection, however, depends on various technical issues, especially in nonlinear models that exhibit objective functions with multiple local optima.

In sensitivity analysis, exploratory techniques are applied to quantify the numerical sensitivity of each parameter. This analysis can be either *local*, performed by varying one parameter while keeping the others constant, or *global*, conducted through a more comprehensive exploration of the parametric space Ω_Θ [69]. Andrea Saltelli and colleagues emphasize the importance of the latter, demonstrating that only global analysis has the potential to capture interactions and synergies that emerge when a given set of parameters changes simultaneously [70]. However, sensitivity analysis also employs search techniques in explorations designed to **discover critical scenarios** and reveal unusual leverage policies. To illustrate this approach, John Miller applied two optimization techniques (genetic algorithms and gradient ascent) to the model `World3` used by Donella Meadows and colleagues in *Limits to Growth*, in order to discover alternative scenarios for world population up to 2100 [71]. The results obtained from these searches indicated that it is possible to maximize the world population by up to six times the amount projected by Meadows' base scenario (4 billion, peaking at 9 billion in 2050), but it is also possible to minimize the population to half of the forecasted

[23] Also known as enumeration method or brute-force method.

amount. These different modes of behavior of the same system help to better understand the critical parameters and flows necessary for policy development. For instance, a global population of 2 billion by 2100 might be desirable, not due to wars, pollution, and famine, but because of economic, technological, and cultural changes that improve people's quality of life.

Destaque 2.6.1– Identifying the best land-use change routes with Dynamic Programming

Among the various search techniques offered by Operational Research, **Dynamic Programming** stands out as an interesting tool for the simulation and optimization of environmental systems under future scenarios. Unlike conventional optimization techniques, this method allows multiple system simulation models to be nested within a more comprehensive search algorithm. This is possible because the search space is discretized into states and stages that are simulated in simplified steps. A recursion relationship then allows the best route between stages and states to be deduced after the simulations.

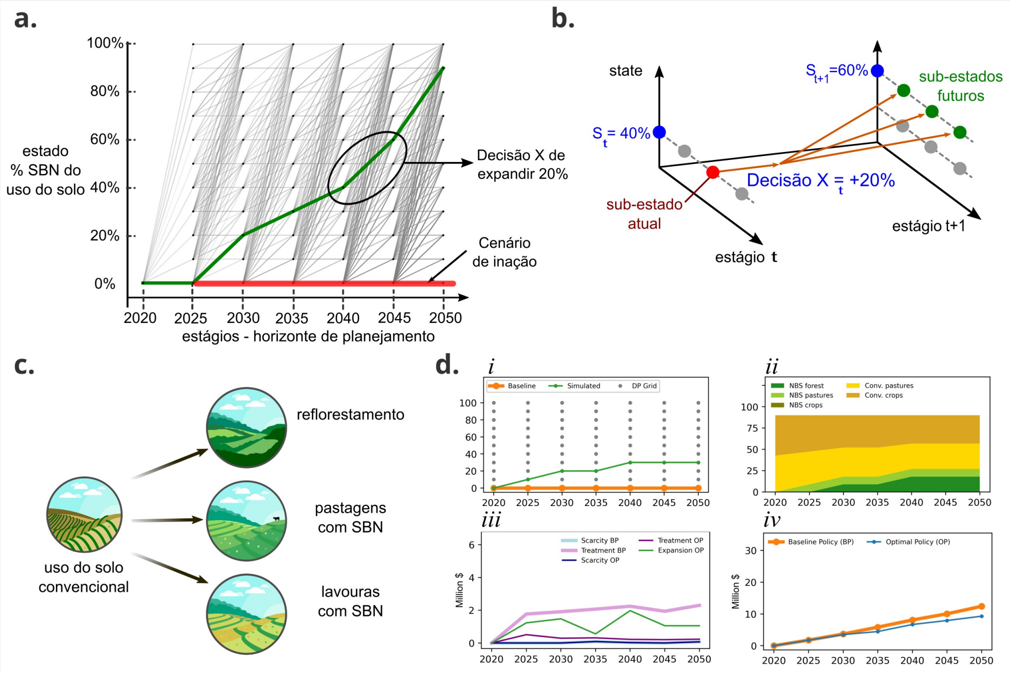

Figure 2.6 — **Exploration of NbS expansion scenarios in watersheds with Dynamic Programming a** — Search space discretized into states and stages. **b** — Representation of transitions between states and stages **c** — Simulated transitions between land-use typologies. **d** — Example of results obtained for a scenario with high initial degradation.

Recently, Guilherme Marques and I adapted the **infrastructure capacity expansion problem** to explore nature-based solutions (NbS) expansion scenarios, focusing on downstream water security (Possantti & Marques, 2022 [3]). The system states were represented by the land cover rate of a mix of natural solutions, discretely represented as sub-states. Each sub-state implies a cost C to the system:

$$C = SC + TC + XC \tag{2.17}$$

Where SC is the cost of water scarcity; TC is the cost of water treatment, and; XC is the cost of expanding natural solutions. Given scenarios of initial conditions and systemic pressure changes, such as climate change and population growth, the best expansion routes were identified. In initially well-conserved scenarios, it was concluded that doing nothing would be better, as the reductions in treatment and scarcity costs did not compensate the expansion costs. However, in more degraded initial conditions, the long-term expansion of NbS reduced the accumulated cost by up to 25%. These conclusions, however, depend heavily on the model representation employed in the simulations, highlighting the need for empirical adjustment.

On the side of empirical adequacy and anomaly detection, exploratory techniques are related to the uncertainty analysis of the parameters. A method of particular relevance proposed in hydrological modeling is the *Generalized Likelihood Uncertantiy Estimation* (`GLUE`) method, introduced by Keith Beven and Andrew Binley, which applies Bayes' Theorem (Equation (1.4)) with an informal likelihood function $\mathcal{L}(E|H)$ [72]. Thus, the posterior distribution of the parameters is obtained by conditioning the prior distribution with the normalized histogram of the likelihood $\mathcal{L}(E|H)$. This histogram, in turn, is calculated from a robust exploration of the parametric space Ω_Θ. Search techniques, on the other hand, act to "calibrate" the model by maximizing the likelihood $\mathcal{L}(E|H)$, resulting in a single set of parameters deemed empirically adequate[24]. In any case, when it comes to models in Systems Dynamics, the informal likelihood $\mathcal{L}(E|H)$ is usually equated to some point-to-point statistic or fit metric that seeks to measure the adherence of the simulated variables to the observed data. Therefore, the closer a simulated point $y_{M,i}$ is to its corresponding observed point $y_{O,i}$, the greater its empirical adequacy. A typical fit metric is the mean absolute error `MAE`:

$$\text{MAE} = \frac{1}{n}\sum_{i}^{N} |y_{M,i} - y_{O,i}| \tag{2.18}$$

An alternative that disproportionately penalizes larger errors more and smaller errors less is the root mean square error `RMSE`:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i}^{N}(y_{M,i} - y_{O,i})^2} \tag{2.19}$$

Both the `MAE` and `RMSE` metrics are positive and have the same units as the evaluated variable y, which makes them difficult to compare with other models or even other variables. A more universal metric is the coefficient of determination R^2 [73]:

$$R^2 = 1 - \frac{\sum_i^N (y_{M,i} - y_{O,i})^2}{\sum_i^N (y_{M,i} - \bar{y}_O)^2} \tag{2.20}$$

where $\bar{y}_O$ is the mean of the observed data. In this sense, the coefficient of determination R^2 can be interpreted as a measure of how much the model M is better at determining the observed values compared to the simple mean of the observed data O. In the context of hydrological simulation, the coefficient of determination is also referred to as the Nash-Sutcliffe Efficiency `NSE` [74]:

$$\text{NSE} = R^2 \tag{2.21}$$

A commonly used alternative in Hydrology is the Kling-Gupta Efficiency `KGE`, which establishes a decomposition among the correlation coefficient r, the mean μ, and the standard deviation σ of the modeled and simulated data [75]:

$$\text{KGE} = 1 - \sqrt{(r_{M,O} - 1)^2 + \left(\frac{\mu_M}{\mu_O} - 1\right)^2 + \left(\frac{\sigma_M}{\sigma_O} - 1\right)^2} \tag{2.22}$$

Finally, an important practical diagnosis involves assessing how the modeled system produces a **surprising behavior** against the mental models (perceptual) of the interest groups involved in the modeling process, changing opinions. If the use of Systems Dynamics models has a *null* effect on the ingrained preconceptions of the target audience (including scientists), then their use has no cognitive learning value. In the

[24] In multi-objective approaches, the set of parameters considered empirically adequate corresponds to a Pareto frontier.

Destaque 2.6.2– A hybrid approach in using the GLUE method

The separation between exploration and search in the parametric space of models is not necessarily an absolute dichotomy, and there are possibilities for **hybridization**. This is what my colleagues and I did to condition the `PLANS` model in Possantti et *al.* (2023) [4] (more details in the next chapter). In this case, we realized that there is a great opportunity to reduce the dimensionality problem by employing an **elitist evolutionary algorithm**. Elitism in an evolutionary algorithm essentially ensures that a population of high-performing solutions is cataloged over the simulated generations without losing them during successive crossovers and random variations [76]. This enables an intermediate, or hybrid, approach that both explores the parametric space by randomly generating solutions and seeks to obtain a high-performing population according to a likelihood measure.

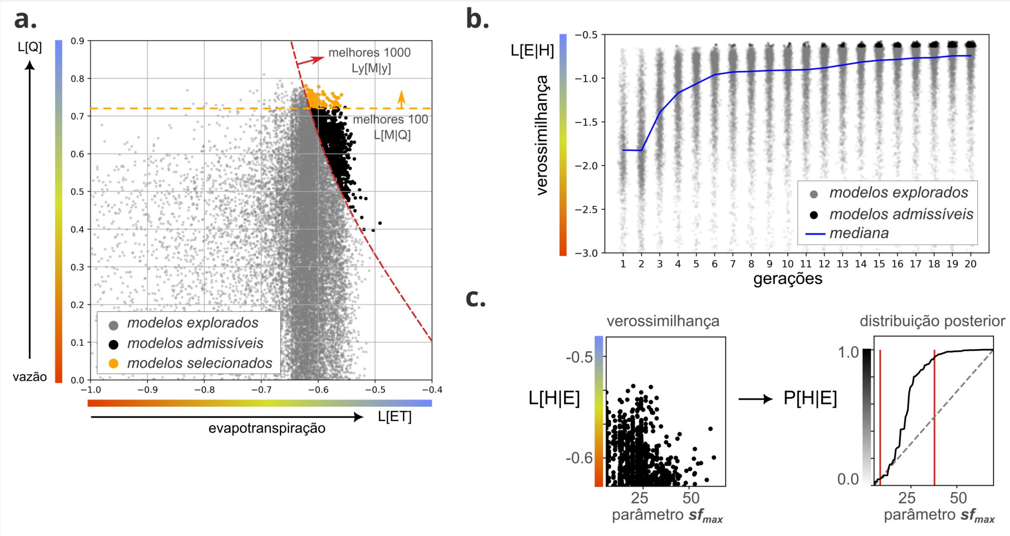

Figure 2.7 — Hybrid approach of the GLUE method in conditioning the `PLANS` model a — Two-dimensional space created by two informal likelihood measures, one for streamflow and another for evapotranspiration. **b** — Generations produced by an evolutionary algorithm with elitism, which preserves the best solutions found while exploring the parametric space **c** — Example of converting informal likelihood into the posterior distribution of a model parameter via Bayes' Theorem.

Another form of hybridization we used was a composite likelihood measure, based both on observed streamflow and evapotranspiration maps derived from remote sensing techniques. The metric, therefore, consisted of a Euclidean distance in the two-dimensional space created by the simulated variables:

$$\mathcal{L}[E|H] = 1 - \sqrt{(1-\mathcal{L}[Q])^2 + (1-\mathcal{L}[ET])^2} \tag{2.23}$$

Where $\mathcal{L}[E|H]$ is the informal likelihood of model H given the set of observations E; $\mathcal{L}[Q]$ is the informal likelihood of flow, and; $\mathcal{L}[ET]$ is the informal likelihood of evapotranspiration. Both likelihoods for each process were determined based on the Kling-Gupta Efficiency (KGE). As a result, maximizing the proposed formula requires that solutions with high likelihood be found in both dimensions of flow and evapotranspiration.

worst-case scenario, modeling becomes a **fallacious appeal to authority** that ultimately blinds its users from making evidence-based decisions (confirmation bias). Thus, the obtained results must be communicated effectively to surprise the target audience, revealing at least new *nuances and details* in the results and, at most, *non-intuitive* mechanisms of system behavior[25]. For example, considering the use of a hydrological model in the context of watershed revitalization, it may be that the target audience overestimates the capacity of the system to mitigate rapid response mechanisms through nature-based solutions. However, this same audience may underestimate the potential of these solutions to improve water quality. Along this line, another relevant diagnosis is to identify whether the model results actually produced **practical changes in**

[25] From the perspective of Thomas Kuhn's paradigms, this diagnosis assesses whether the modeling results are embedded in the *normal cycle* of science, articulating and enhancing the prevailing theory.

decision-making, in formulating strategies and action plans. Changing the personal opinions of involved actors is one thing; effectively altering the decision-making process is another. This category of diagnosis should include the evaluation of whether the developed model is *usable by third parties*, known as the **reproducibility problem**. A model accessible only by its developers, no matter how good, will likely leave a small legacy, producing low practical impact outside the scope of the project for which it was conceived. In this sense, Sterman emphasizes strategies to maximize reproducibility, such as low computational demand, ease of installation and operation, and maintaining accessible documentation, including *websites* with training, tutorials, etc. In the pragmatic spirit of Systems Dynamics, evaluating the positive impact of a model is the essence of the modeling endeavor and should be carefully planned with impact indicators from the outset of any project. The relevance of preparing the diagnosis *a priori* is to avoid **anecdotal evidence** and minimize the **retrospective bias** in assessing the success of the model. It is clear that important decisions always involve ethical and political issues at play, but the use of models in scientific advisory roles, including their uncertainties, should have at least some positive impact. ■

2.7 Chapter summary

In this chapter, I presented modeling as a learning process, where perceptions are refined into computational models. The discussion on representation highlighted idealizations and analogies as tools to make target systems understandable. Hydrological modeling was treated as an example, emphasizing the importance of identifying leverage points to influence hydrological responses. Finally, the chapter underscored the significance of diagnostics, ensuring that models are both technically sound and practically applicable.

- **Modeling is a learning process.** In Hydrology, this involves an iterative learning cycle that begins with the perceptual model (subjective impressions), moves through a conceptual model (objective expressions), and culminates in a procedural model (computation). A diagnostic step reviewing adequacies closes the cycle, revisiting methods, theories, and perceptions.

- **The problem of representation.** Idealizations are deliberate simplifications used to make the target system tangible. Scaled-down or scaled-up models are idealizations in the form of copies of target systems. Conversely, analog models are formal analogies. In any case, analogical inference is employed.

- **Systems are an ontological paradigm.** Systems are a set of parts with relationships among them. Stable or unstable behaviors emerge from these relationships. This Aristotelian paradigm sees *form* as the unifying element of the object. Ludwig von Bertalanffy proposed the General Systems Theory as a unification of Science. deterministic chaos and computational irreducibility pose challenges to the predictive capability of systemic theories.

- **Structure defines behavior.** Systems Dynamics is an applied discipline where the goal of modeling is to understand modes of behavior and identify leverage points in systems for better decision-making. The compartment architecture allows modeling of levels connected by material flows and feedback loops. Parameters act in the flow equations, regulating the levels. The system is solved numerically using the Euler method, allowing complex and nonlinear behaviors to emerge from simulations.

- **Rapid and slow hydrological responses.** A hydrological model is created for exploratory purposes. The concepts of slow and rapid hydrological response are introduced. Three reservoirs interact based on two basic flow equations regulated by six parameters. The model demonstrates equifinality, with slow responses produced by more than one mechanism. It is possible to evaluate where to act in the system to maximize water availability by reducing the dominance of rapid responses.

- **Adequacy diagnostics.** John Sterman asserts that model adequacy should be tested across conceptual, technical, behavioral, and practical aspects. The test of empirical adequacy, while crucial, should integrate other assessments: adequacy of the boundary; adequacy of the structure; dimensional consistency; parameter distribution; comparative studies; integration error; extreme conditions; sensitivity; anomalies; surprises; and practical impacts. An empirically inadequate model with practical impacts is harmful; but an empirically adequate model without practical impact is devoid of meaning.

The soil mantle acts as a natural reservoir for rainwater, storing it in its internal cavities. The fauna and flora, by excavating macropores, not only create pathways for water infiltration but also accelerate its release, especially in the upper horizons.

Chapter 3

The understanding of hydrological processes

> The assertion or assumption that all rises are caused by surface runoff has persisted in articles and even in some hydrology textbooks, despite much evidence to the contrary in forestry and agricultural research.
>
> Hewlett & Hibbert (1967, p. 275) [77]

> A box, objectively defined by distinct dynamics of groundwater, soil solution chemistry, or isotopic composition, with defined area, depth, and porosity, is a much better modeling building block than a myriad of elements in landscapes that are notoriously heterogeneous both vertically and laterally!
>
> Jeffrey McDonnell (2003, p. 1872) [78]

3.1 Zero-order Basins

Hydrology is the science that studies continental waters, seeking to understand how water is distributed on continents after precipitating from the atmosphere and before returning to the oceans [79]. In other words, Hydrology investigates how the **hydrological cycle** manifests in its terrestrial phase, in contrast to Meteorology (which focuses on the atmosphere) or Oceanography (which studies the oceans). When contemplating this, it is easy to imagine large rivers, such as the Amazon and Paraná, as well as other notable rivers like the Danube, Nile, Yellow, Indus, Ganges, and Mississippi. In the case of Brazil, images of lush nature arise, including the vast Amazonian and Pantanal floodplains, as well as the spectacular Iguaçu Falls. Visions related to large-scale human intervention also emerge, such as the hydroelectric complex, with its dams spread throughout the country, and the major transposition and irrigation projects, exemplified by the reservoirs in the Cantareira Mountains, the transposition of the São Francisco River, and the central pivots in the São Marcos River basin. The recent floods that devastated the cities located in the river valleys of Rio Grande do Sul also illustrate

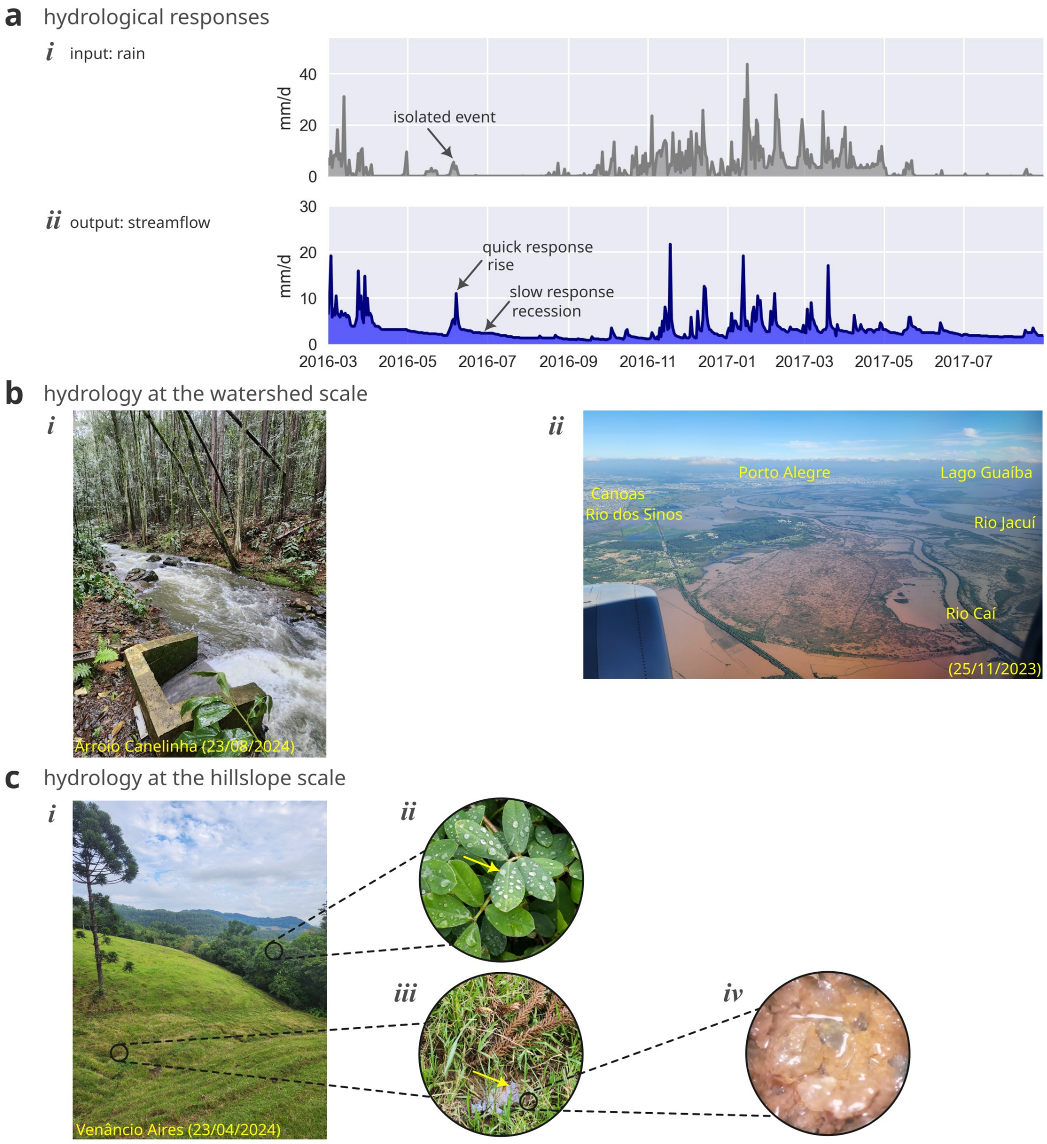

Figure 3.1 — Hillslopes: where it all begins. The most intuitive scale when thinking about hydrological processes is the flow of water in rivers, which are channels that drain water to the ocean. However, the hydrological responses to rainfall events originate in the slopes, in zero-order basins, where rain interacts with the landscape. **a** — The alternation between fast responses (rises, detail i) and slow responses (recessions, detail ii), observed in a medium-sized drainage area in the Paraíba do Sul basin (342 km², Rio de Janeiro). Daily rainfall obtained from `INMET 83738` Station (Resende) and daily flow obtained from `ANA 58287000` Station (Rialto). **b** — The evident hydrological processes at the basin scale include the propagation of flow through the drainage network, downhill (detail i) and the flooding of plains, when river flow exceeds the larger section of the channels, invading adjacent dry lands (detail ii). **c** — The water that supplies river flow comes from the interaction of rain with the slopes (detail i), resulting in surface processes (such as interception, in detail ii) and subsurface processes (such as soil saturation, details iii and iv).

that rivers are crucial not only for energy and food production but also for ensuring the basic health and physical safety of the inhabitants of vast urban metropolises. In this regard, it is not uncommon for textbooks on Hydrology to mention how the first city-states, which emerged in Mesopotamia and Egypt, developed an almost symbiotic relationship with large rivers and their floodplains. Water and society are closely linked.

This intuitive interpretation of Hydrology, however, results from two particular perceptions. The first is the **engineering bias** that permeates Hydrology, which has always been marked by a **science-management duality**. This duality implies that Hydrology exists in a fluid interface between theoretical investigation of nature (problems that are *important* for human knowledge) and practical solutions to social, environmental, and economic impasses (problems that are *urgent* for people). James

Dooge (1988) [80] argues that the field was born relatively distinct from other scientific disciplines, such as Physics or Biology, being essentially pragmatic in its formation. Throughout history, hydrological problems generally presented themselves directly from their application, leading to new data being obtained and ultimately some knowledge being produced. For example, Dooge illustrates that, according to Pliny the Elder (24-79 AD), the level of the Nile River was measured in antiquity not in terms of flow but on a scale of the socio-economic impact involved: famine (low level), safety (medium level), and disaster (above flood stage). In this regard, Murugesu Sivapalan and Günter Blösch suggest that the field evolved from a phase based on reductionist and pragmatic engineering methods (Empirical Era) to becoming, during the twentieth century, an Earth Science (Geoscience Era), holistic and integrative, merging with important branches of Physical Geography, Geology, Pedology, Ecology, and recently, Sociology (Co-evolution Era) [81], [82]. In other words, Hydrology is increasingly establishing itself as an interdisciplinary science that views the terrestrial phase of the hydrological cycle as a product of various ecological, economic, social, and other interactions and dynamics. The second perception is the **fluvialist bias** that dominates the field, especially in areas traversed by large rivers, such as Brazil, the United States, China, and Central Europe. This view directs hydrological studies toward essentially hydraulic problems at a continental scale, such as the propagation of flow in the channels of the drainage network, the flooding of adjacent plains, and, with the advent of remote sensing, continental water balance. The efficiency of economic activities such as electricity production, navigation, basic sanitation, irrigation, etc., fundamentally depends on this type of knowledge. This bias makes sense in much of the continents but carries less weight in regions dominated by small and medium rivers, such as the archipelagos of Japan, New Zealand, and the British Isles. The fluvial perspective is illustrated by the account reported by a Brazilian hydrologist during a visit to the United Kingdom:

> When I was in England, I went to visit a certain reference gauging station on a famous river in the region. They talked about that river all the time, but when I arrived at the site, I was somewhat perplexed and disappointed. That was not a river: it was a stream. With a little push, it was even possible to jump to the other bank. – Walter Collischonn (2023, in personal communication).

With the influence of these two biases, it is somewhat easy to forget that water only flows in rivers as a consequence of processes occurring on the hillslopes and, ultimately, in the vertical profile that starts at the vegetation canopy, passes through the surface and horizons of the soil, and ends in the underground rock foundation. The propagation of flow through the channels and the flooding of plains are nothing more than processes of transport and dissipation of rises produced by the interaction of rain with the higher and mountainous terrain of the landscape. This importance of the *slope scale* was initially highlighted in Japan by Tsukamoto (1973) [83] in the early 1970s, who expanded the systematic hierarchization of channels proposed by Strahler (1957) [84] by introducing the concept of **zero-order basin** (in Japanese: 0次谷). Although Tsukamoto's emphasis on the slopes and valleys of the terrain advances specific issues of erosion and sediment production, its primordial importance in the hydrological cycle is evident. As illustrated in Figure 3.1, *it is in zero-order basins where it all begins.* The alternation between rises and recessions, a primordial observation in Hydrology (Figure 3.1**a**), does not originate from the propagation of water downstream or from the flooding of plains (Figure 3.1**b**), but from the interaction of rain with the landscape, at the slope scale (Figure 3.1**c**). In this same regard, Mediondo & Tucci (1997) [85] use the term **drainage basin**, which they also argue is the starting point for understanding the diversity of hydrological processes, reflecting both at the micro and macro scale. I will use the term zero-order basin here, considering that, according to Godoy et al. (2021) [86], this term has become popular in the international literature.

Table 3.1 organizes the nomenclature regarding the hydrological processes that occur on hillslopes and valleys of the terrain, to be explored in more detail in this chapter. Although relevant in theory, is this entire diversity relevant *in practice*? When considering the application of hydrological models to assist in decision-making and strategy formulation, how much can the complexity present in zero-order basins be simplified or even neglected? After all, as we saw in Chapter 2, modeling systems needs to utilize idealizations, which are deliberate simplifications to make the target system more tangible. Moreover, in the face of rivers that travel continental distances, the minute details about hydrological processes in zero-order basins lose any practical sense. The mere confluence of two medium-sized rivers or a floodplain inundation can completely erase the hydrological signature left by some typical feature produced by processes on the hillslopes. The mass of water and sediments, energy, and momentum are necessarily preserved by the laws of conservation, but detailed information is progressively attenuated and mixed in the large flow moving toward the ocean. In this sense, as long as a model presents empirically adequate quantitative results, the details about the processes in zero-order basins would be irrelevant.

This *appeal for simplification* becomes a seductive objection, as it greatly facilitates the modeling process. But it is merely a reflection of the fluvialist bias: a perspective that frames questions to be understood and problems to be solved from upstream to downstream, riverward. In this regard, the most popular hydrological models, at least in Brazil, such as `SWAT`, `HEC-HMS`, `MGB`, `SWMM`, treat zero-order basins as hermetic units or black boxes, making it impossible to recover details about hydrological processes on the hillslopes and valleys of the terrain, except in average and aggregated terms. At the end of the computational simulations, the most informative visualization possible is a mosaic of sub-basins[1]. It is clear that this simplification is justified when the objective of a given study is to understand phenomena and solve downstream hydrological problems, i.e., fluvial ones. However, to address much of the issues related to water security, such as the revitalization of watersheds, it is necessary to take a look from downstream to upstream, hillslope above, representing the zero-order basins with sufficient detail, because it is at this scale that the relevant processes occur and actions need to be specified. Therefore, a useful model must take seriously what hydrological theories say about runoff generation in zero-order basins. Otherwise, there is a risk of instantiating a model that fails both the boundary adequacy test and the structural adequacy test (see Section 2.6).

That said, this chapter marks the point at which I will articulate how theories about hydrological responses in zero-order basins can be conveyed by hydrological models. This is a critical point because here we will encounter all the philosophical, scientific, and technical challenges and problems exposed in the previous two chapters, but now from a hydrological perspective. The topics will all be revisited directly or indirectly, such as the rise and fall of paradigms, the refutation and confirmation of hypotheses, the problems of structure, dimensionality, and underdetermination, etc. Essentially, it will be seen that the complexity of hydrological processes in the soil and on the hillslopes brought by empirical evidence, combined with the difficulty of obtaining direct observations in any given basin, makes any attempt at modeling based on continuous and spatially distributed mathematical formalizations a disproportionate effort. As we will see, a unifying solution to this problem, recently proposed by Jeffrey McDonnell (2021)

[1]It is also possible to recover information in the hydrological response units within each sub-basin. However, as we will see later, this representation is irretrievably static, while the processes at the scale of zero-order basins are dynamic in time and space.

Component	Name	Dimension	Unit	Category
C	vegetation canopy	L	mm	reservoir
S	soil surface	L	mm	reservoir
O	organic horizon	L	mm	reservoir
V	vadose zone, mineral horizon	L	mm	reservoir
$\mathbf{V}_\mathrm{c}$	capillary water in the vadose zone	L	mm	reservoir
$\mathbf{V}_\mathrm{g}$	gravitational water in the vadose zone	L	mm	reservoir
$\mathbf{D}_\mathrm{v}$	capillary deficit	L	mm	reservoir
G	groundwater zone	L	mm	reservoir
D	saturation deficit	L	mm	reservoir
p	precipitation, rain	L/T	mm/h	flow (exogenous)
p_s	effective rain	L/T	mm/h	flow
p_x	excess rain	L/T	mm/h	flow
Q	river flow, fluvial runoff	L^3/T	l/h	flow
q	specific river flow, fluvial runoff	L/T	mm/h	flow
f	infiltration	L/T	mm/h	flow
q_si	runoff, runoff by excess of infiltration	L/T	mm/h	flow
q_se	direct rain, surface runoff due to saturation excess	L/T	mm/h	flow
q_ss	exfiltration, subsurface runoff, lateral runoff	L/T	mm/h	flow
q_o	percolation between horizons	L/T	mm/h	flow
q_v	recharge, final percolation	L/T	mm/h	flow
Q_g	base flow, slow groundwater discharge	L^3/T	l/h	flow
Q_gt	translational flow, rapid groundwater discharge	L^3/T	l/h	flow
e_pot	potential evapotranspiration	L/T	mm/h	flow (exogenous)
e_c	evaporation in the canopy	L/T	mm/h	flow
e_s	evaporation at the surface	L/T	mm/h	flow
e_o	transpiration in the organic horizon	L/T	mm/h	flow
e_v	transpiration in the vadose zone	L/T	mm/h	flow
e_g	transpiration in the groundwater zone	L/T	mm/h	flow
f_max	infiltration capacity	L/T	mm/h	parameter
K	hydraulic conductivity	L/T	mm/h	parameter
g	aquifer detention time	T	h	parameter
c_max	interception capacity	L	mm	parameter
s_max	surface retention capacity	L	mm	parameter
o_max	field capacity of the organic horizon	L	mm	parameter
v_max	field capacity of the mineral horizon	L	mm	parameter
m	vertical uniformity constant of the soil	L	mm	parameter

Table 3.1: Hydrological processes in zero-order basins — Relation of reservoirs, flows, and parameters related to hydrological processes in zero-order basins.

[87], consists of adopting conceptual models that *effectively* represent the processes of **connectivity** at the scale we need to address to answer our research questions. Returning to the analogy of the landscape that I introduced at the beginning of this thesis, we are clearly moving out of the narrow valleys of abstract and philosophical subjects to enter a broader field of more tangible and applied questions. Here, the mountain streams converge, forming mighty rivers that flow through bars and banks.

3.2 The age of infiltration

In Section 2.5 of the previous chapter, I organized a prototype of a hydrological model aimed at illustrating and articulating how Systems Dynamics can be employed in the modeling process. The obtained conceptual model, maintained in a minimalist condition, was constructed primarily based on the perception that a watershed exhibits **fast and slow responses** to rainfall events, thus producing the phenomenon of alternation between **rises** and **recessions** in rivers [77]. This is a fundamental perception in Hydrology: when it rains, rivers become agitated, the water gets muddier, and levels rise (fast response); between rains, rivers calm down, the water becomes clearer, and levels fall (slow response); if too much time passes before it rains again, smaller streams begin

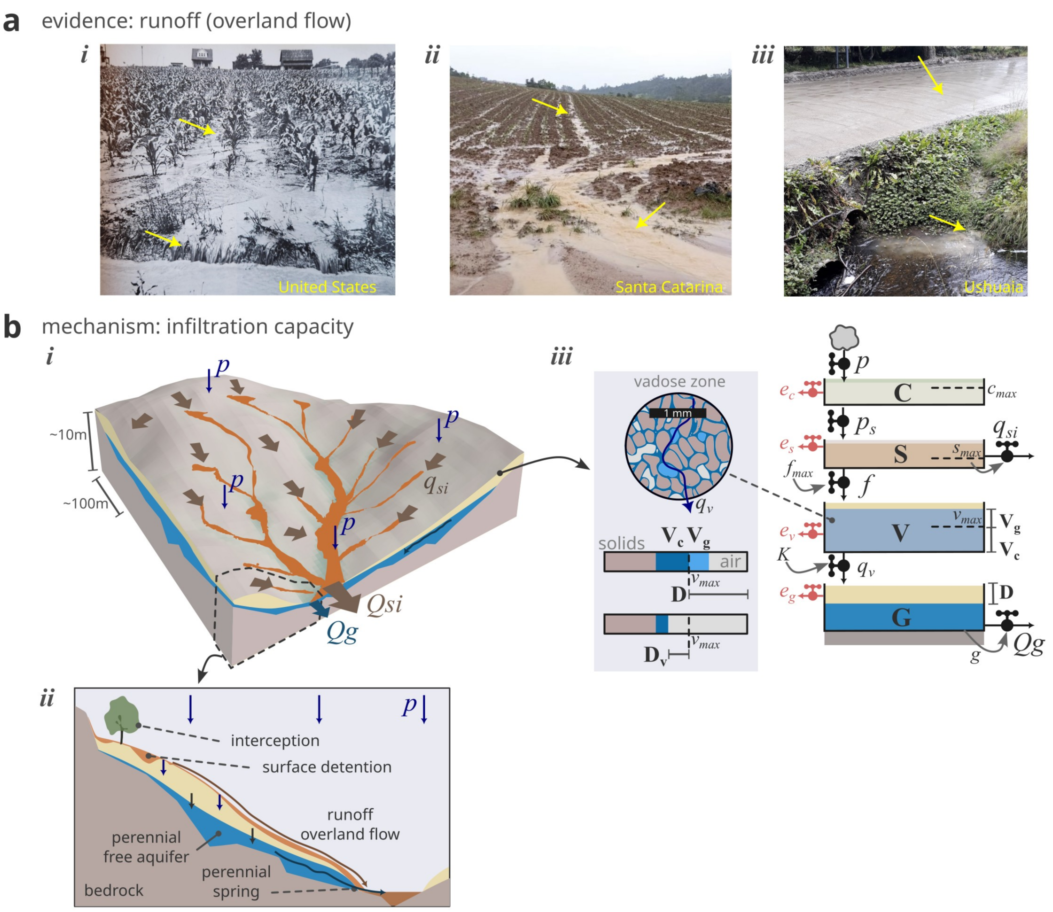

Figure 3.2 — **The Hortonian paradigm.** The Hortonian paradigm explains the alternation between rises and recessions based on the concept of the soil's infiltration capacity f_{max}. **a** — The motivating empirical evidence includes the surface runoff observed after rains when water fails to infiltrate: runoff in the United States reported by the SCS [88] (detail i); runoff in Santa Catarina reported by EPAGRI [89] (detail ii), and; rural road runoff in Ushuaia, authored by me (detail iii). **b** — The surface runoff q_{si} occurs generally in the basin (detail i) from the moment when the rainfall flow exceeds the interception capacity c_{max} and the surface retention capacity s_{max} (details ii and iii).

to dry up (the system tends to empty). With some empirical rigor, this phenomenon in any given watershed can be measured and reproduced in graphs with the aid of a rain gauge and a level stick. With a bit more empirical rigor, the perception of this phenomenon becomes sharper through field expeditions, observing the spatial and temporal dynamics of springs and puddles (where groundwater slowly surfaces) and the rapid runoff (caused by more intense rains).

In the context of zero-order basins, the dominant scientific theory today postulates that hydrological responses to rainfall events are the consequence of **multiple mechanisms of runoff generation**, both surface and subsurface, encompassing fast and slow responses, whether simultaneous or not, which will be described in the next section. These mechanisms were revealed and corroborated by successive experimental investigations in small basins, hillslopes with trenches, and soil plots during a scientific revolution in Hydrology that took place throughout the second half of the twentieth century. Before this revolution, however, the hegemonic scientific explanation for the fast and slow responses of hillslopes was primarily based on the hydrological theory of Robert Horton (1875-1945) [90], [91]. Once published, the perceptual model described by Horton solidified as a true paradigm in the following years and decades, marking the so-called **Infiltration Age** – a long period of normal science in which the scientific community developed research in both pure and applied fronts to articulate its implications [92], [93]. Although ultimately surpassed by a more complex explanation, Horton's theory, being scientific (i.e., falsifiable), greatly contributed to elevating Hydrology from its Empirical Era, focused on engineering applications, to being understood as a Geoscience, aimed at explaining natural phenomena.

The central idea of Horton's perceptual model is established in the paper *The role of infiltration in the hydrological cycle* (1933) [90], in which the soil is conceived as a *separating surface* for rain: a portion of the rainwater infiltrates into the hillslopes, lodging in the soil matrix, while another portion runs off superficially as **runoff**, causing dramatic increases in river flow downstream (Figure 3.2**a**). Thus, **infiltration** f would be the key process for understanding the hydrological cycle in its terrestrial phase:

> Infiltration divides rainfall into two parts, which thereafter pursue different courses through the hydrologic cycle. One part goes via overland flow and stream-channels to the sea as surface-runoff; the other goes initially into the soil and thence through ground-water flow again to the stream or else is returned to the air by evaporative processes. The soil therefore acts as a separating surface, and the author believes that various hydrologic problems are simplified by starting at this surface and pursuing the subsequent course of each part of the rainfall as so divided, separately. – Robert Horton (1933, p. 446–447) [90].

To articulate this perceptual model, Horton instantiates various flows, reservoirs, and important parameters of the system representing the zero-order basin. Figure 3.2**b** illustrates the modeled system (the evapotranspiration flows in each reservoir are denoted by E). The primordial flow consists of **effective rainfall** $p_\mathbf{s}$, that is, the flow of rain that actually reaches the soil after the rain p exceeds the interception capacity $c_{\max}$ in the vegetation canopy $\mathbf{C}$. The soil, in turn, consists of a porous matrix of solid minerals that stores water in films maintained by the surface tension of its particles, thus forming the **vadose zone V**. The accumulation of water in the films in this zone occurs up to a certain limit, which is the characteristic **field capacity** $v_\mathbf{max}$ of the soil. This water in the vadose zone $\mathbf{V}$, which is trapped in the pores, is referred to as **capillary water** $\mathbf{V_c}$. In addition to the field capacity $v_{\max}$, the films of water on the particles, when mixed, create a relatively mobile mass of water, termed **gravitational water** $\mathbf{V_g}$. This water then percolates vertically through the pores due to gravity, forming a **phreatic zone G** above the impermeable layer, which is the solid rock foundation (i.e., this zone forms an unconfined aquifer). Horton refers to **recharge** $q_\mathbf{v}$ as the vertical flow of water suspended in the vadose zone $\mathbf{V}$ to the aquifer of the phreatic zone $\mathbf{G}$, a process that can eventually raise the groundwater level (thus increasing the hydraulic load in this porous system). In this context, **gravitational deficit D** represents the amount of gravitational water $\mathbf{V}_\mathrm{g}$ in the vadose zone $\mathbf{V}$ necessary to achieve complete soil saturation and, consequently, the emergence of the groundwater table at the surface. Here, the maximum flux of recharge q_v is limited by the **hydraulic conductivity** K of the soil: the closer the vadose zone $\mathbf{V}$ is to saturation, the vertical percolation tends to be dominated by the hydraulic load, overcoming surface tension.

At this point, Horton introduces a crucial parameter in his perceptual model: the **infiltration capacity** $f_\mathbf{max}$, that is, the maximum maximum flow of infiltration that the soil surface can offer at any given moment. It is important to highlight that this capacity is an attribute of the thin top layer of soil and, according to Horton, tends to be lower than the hydraulic conductivity K of the soil matrix, functioning as a critical limiting factor in the system. Thus, water from a effective rainfall p_s with an intensity less than or equal to infiltration capacity $f_{\max}$ is completely absorbed by the soil matrix. On the other hand, when a effective rainfall p_s has an intensity greater than infiltration capacity $f_{\max}$, the water from the **excess rainfall** p_x begins to fill the small surface depressions. If the excess rainfall p_x persists long enough, the **surface retention capacity** $s_{\mathbf{max}}$ is exceeded, and the process of surface runoff q_{si} or **runoff by excess of infiltration** begins, where the rainwater flows through furrows and ravines downhill until it reaches the channels of the streams.

The infiltration capacity $f_{\max}$ depends on soil type, texture, and management practices, which implies variability in the response of different basins, even when subjected to identical rainfall events. In addition to spatial variability, Horton argues that the infiltration capacity $f_{\max}$ of the soil varies over time, dynamically oscillating between extremes of minimum and maximum capacity, in a **decay and restoration cycle** (Figure 3.3**a**). In this cycle, the decay phase occurs during rainfall events due to the expansion of colloidal particles, clogging by fine particles, and compression caused by the impact of raindrops. On the other hand, the restoration phase occurs during dry periods as new cracks and pores open up due to the retraction of colloidal particles, thermal expansions, and the activity of soil fauna, such as insects and earthworms. With this conception, it is expected that a long rain, even of relatively low intensity, will eventually produce surface runoff q_{si} if the ongoing decay leads the soil's infiltration capacity $f_{\max}$ to a value *below* the intensity of the effective rainfall p_{s}. Furthermore, this concept introduces the effect of **antecedent moisture conditions**, such as the difference in responses between the beginning and the end of a rainy season or during the onset of a cold front with persistent rains. In this case, if the restoration speed of the soil's infiltration capacity $f_{\max}$ is relatively slow, subsequent rains, even if *less* intense, tend to produce *more* surface runoff q_{si} than the initial rains. In other words, the behavior of the system becomes highly non-linear.

With the theory regarding the role of infiltration in the hydrological cycle, Horton then advances to definitively explain the phenomenon of rises observed in rivers, proposing a method for separating the hydrograph (Figure 3.3**b**). To this end, he argues that total runoff consists of two separable flow components: (1) the groundwater flow, which is a slow response from the unconfined aquifer, and; (2) the surface runoff, which is a rapid response of runoff produced on the hillslopes. Both responses are controlled, in whole or in part, by the soil's infiltration capacity $f_{\max}$:

> In accordance with this theory, total runoff consists of two parts: (1) Suface-runoff, which is dependent on rainfall-amount, rain-intensity, and infiltration-capacity and is practically independent of evaporation-rate. (2) Ground-water runoff. This is dependent on (a) total infiltration and hence indirecty on the same factors which control surface-runoff and is also dependent on (b) vegetational activity and evaporation, which in part determine the water losses, and on (c) the complex interrelations between infiltration-capacity, field moisture-capacity, vegetational activity, and accretion to the water-table. – Robert Horton (1933, p. 454) [90].

Assuming that the phreatic zone **G**, the unconfined aquifer, functions as a linear reservoir, the **recession curve** of the ground flow, or flow of discharge, consists of a typical exponential decay curve of the type $Q_g = Q_{g,o}e^{-t/g}$. This curve is a physical characteristic, and the aquifer detention time g can be extracted from hydrographs during dry periods when water losses due to evapotranspiration are minimal (for example, in the colder months in temperate and subtropical climates). Once obtained, the recession curve can be horizontally shifted on the hydrograph, allowing for the separation of the surface component of river flow from the purely subsurface contribution. Thus, Horton proposes that there are four possible typologies of hydrological response to rainfall events. The **Type 0** response occurs when the intensity of effective rainfall p_{s} is less than infiltration capacity $f_{\max}$ and the total infiltrated water is below the capil-

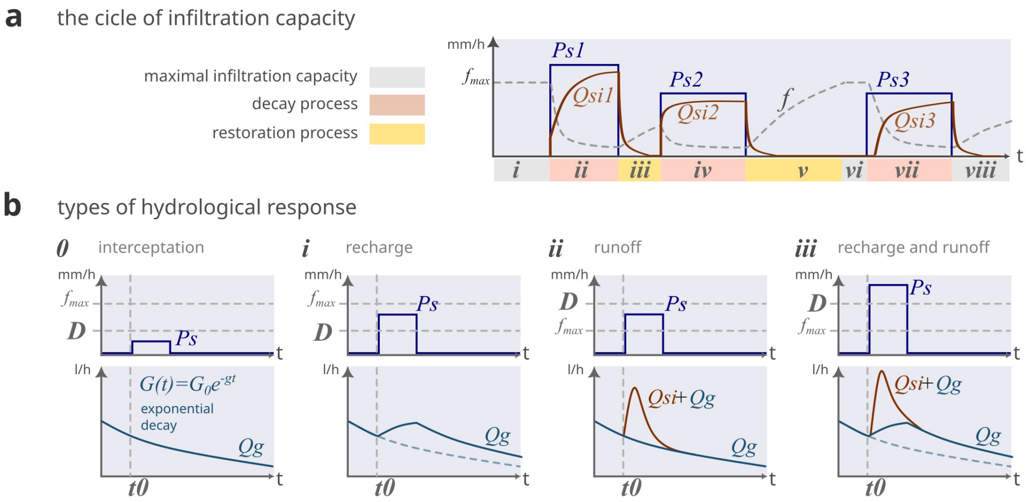

Figure 3.3 — Implications of the Hortonian model. a — The capacity for infiltration f goes through cycles of decay and restoration, generating non-linearities, such as the fact that identical rains ($P_{g,2}$ and $P_{g,3}$) produce different fast responses ($Q_{s,2}$ and $Q_{s,3}$) (details iv and vii). **b** — It is also possible to deduce different typologies of hydrological response: Type 0, when rain produces neither recharge nor runoff (no response, detail 0); Type 1, when rain produces only recharge (slow response, detail i); Type 2, when rain produces only runoff (fast response, detail ii), and; Type 3, when rain produces both a slow and a fast response (detail iii).

lary deficit $\mathbf{D}_\mathrm{v}$ (there is *no* surface runoff q_{si} and *no* recharge). In this situation, even though it has rained, there is no detectable change in the river's recession curve. The **Type 1** response occurs when the intensity of effective rainfall p_s is less than infiltration capacity f_max, but the total infiltrated water is greater than capillary deficit $\mathbf{D}_\mathrm{v}$ (there is *no* surface runoff q_{si} but *there is* recharge q_v from the aquifer). In this case, the recession curve is shifted, depending on how much the recharge q_v exceeds (or falls short of) the discharge flow, potentially producing a (relatively slow) pulse in the river flow purely from the increase in hydraulic load in the aquifer system. The **Type 2** response occurs when the intensity of effective rainfall p_s exceeds infiltration capacity f_max and the total infiltrated is very low, below capillary deficit $\mathbf{D}_\mathrm{v}$ (there is *sf-runoff* but *no* recharge q_v from the aquifer). This type of response consists of a rapid pulse of surface runoff water superimposed on the recession curve that was developing before the event. Finally, the **Type 3** response happens when both the fast and slow responses occur simultaneously: both surface runoff q_{si} and recharge q_v from the aquifer manifest as overlapping pulses. These four typologies illustrate the complexity that emerges from Horton's perceptual model, providing significant leeway for explanations of rises that vary according to the characteristics of the surface, soil, subsurface, rainfall, and antecedent moisture conditions.

In light of this perceptual model, various conceptual models have been developed through physical approaches (when physical principles are applied *a priori*) and empirical approaches (when equations are fitted to data *a posteriori*) [94]. In the realm of the physical approach, notable advancements include those by Philip (1957) [95], who laid the foundations for a mathematically formal theory of infiltration as a special case of the Darcy-Richards equation, that is, the movement of water in an unsaturated porous medium. On the empirical side, Robert Horton himself maintained an applied research line, proposing a conceptual model of exponential decay for the soil's infiltration capacity f_max [96]. Consequently, the production of experimentally standardized infiltration curves enabled a more sophisticated technique for estimating total surface runoff q_{si}, in contrast to the rational method, which is based on a simple runoff coefficient [92]. Another empirically influential method produced in this context was the **Curve Number Method (*Curve Number* (`CN`))**, developed by the *Soil*

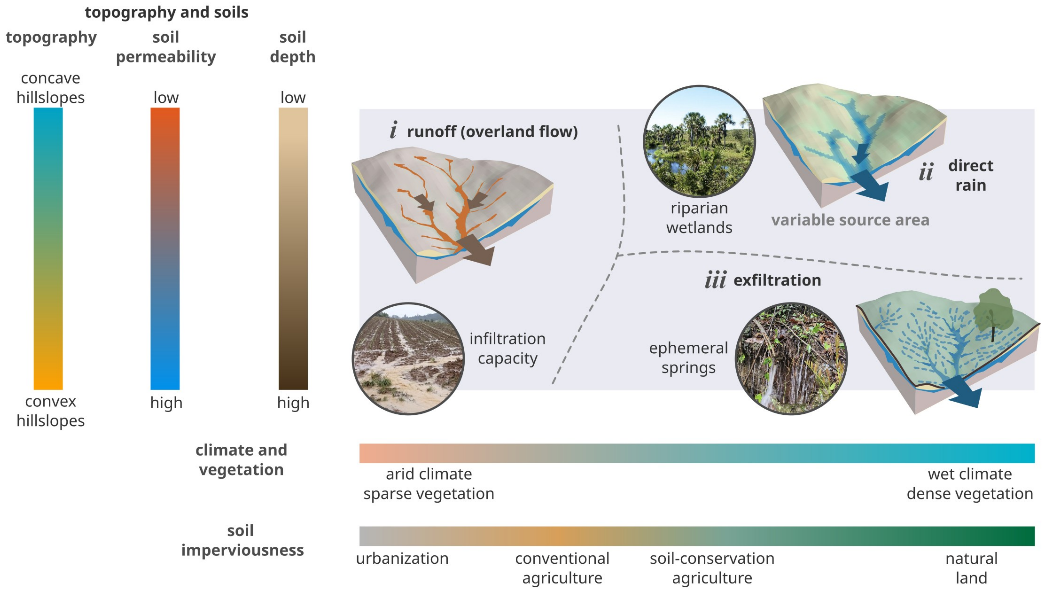

Figure 3.4 — Differentiation of rapid response mechanisms. Schematic view proposed by Thomas Dunne (1983) [99] on the new paradigm of rapid response mechanisms in zero-order basins. Runoff, considered the only mechanism in the Hortonian paradigm, has been reserved for special conditions in arid climates or anthropized environments, such as farms or cities, where the soil's infiltration capacity is very low (detail i). At least two new mechanisms are differentiated: rapid responses due to saturation excess, or direct rain on riparian wetlands (detail ii), and; rapid responses due to exfiltration in ephemeral springs, dominant in deeper, structured soils with macropores (detail iii).

Conservation Service (SCS) in 1954 and presented as a technical guideline in the following decades. According to Rallison & Miller (1981) [97], the CN method from the SCS emerged from the results of experimental research in small basins, but was primarily motivated by the passage of environmental protection legislation in the United States. It is noted that Horton was a consultant for the SCS, but his infiltration curve method did not gain much traction, yielding to the aggregated approach of Mockus (1949) [98], which evaluates the relationship between total rainfall and runoff for individual events. Justified by this evidence, the equation of the CN method seeks to express the supposed *transition* from a *non-linear* response (when infiltration and surface retention dominate the surface water balance) to a *linear* response (when the intensity of effective rainfall p_s exceeds infiltration capacity f_max and surface retention). The CN parameter, in this sense, calibrates the effect of non-linearity for different types of soil, land covers, management practices, and antecedent moisture conditions. Rallison & Miller point out that the choice of this approach by the SCS had a strong convenience bias, as the data used were readily available on a national scale (in the United States). Nevertheless, the essence of the CN method reproduces Horton's perceptual model, as surface runoff q_{si} is regarded as the only rapid response of the watershed and is determined by the estimate of excess rainfall p_x.

3.3 The age of differentiation

In Section 1.5, I highlighted that, according to Thomas Kuhn, the success of a scientific theory is primarily associated with its *competitiveness* in relation to other ideas circulating within the scientific community. Thus, a theory tends to establish itself as the hegemonic paradigm when it efficiently explains known phenomena and opens promising avenues for investigation for new generations of researchers. In the absence of structured competing theories in this sense, Horton's perceptual model for explaining rapid and slow hydrological responses provided these two attributes to a scientific community strongly marked by engineering bias and fluvialist bias. These two characteristics have a shielding effect on perceptions that are technically irrelevant for solving typical downstream problems, as they do not require details about what happens in zero-order basins. Ironically, this was precisely the objective of the `SCS` – to protect the soil on the hillslopes. The Hortonian model, whether partially or completely incorrect, did not prevent good estimates of design flows for bridges, dams, and water balance in large basins, among others. After all, the equifinality of systems allows similar behaviors to manifest from distinct mechanics (until the day everyone is surprised). In strictly technical terms, to some extent, the development of the `CN` method by the `SCS` *perpetuated* the ideas of the Infiltration Age, being generally the basic method required or accepted by technical guidelines in engineering projects and the main module for generating runoff in various distributed hydrological models in software such as `SWAT` and `SWMM`.

Although the paradigm of infiltration may have consolidated simply due to a lack of competitors, its crisis has been present since its formation in the 1930s and 1940s. For example, the rainfall intensity and soil infiltration capacity $f_{\max}$ data measured by Horton's own laboratory suggest that it is highly *unlikely* he observed widespread surface runoff q_{si} in his experimental basin at La Grange Brook (14.4 ha, New York) [100]. The laboratory data, evaluated by Keith Beven, indicate that the soil in the main covers of the basin (fields and orchards) exhibited a relatively high infiltration capacity $f_{\max}$ compared to the recorded rains. In fact, the reconciliation between Horton's infiltration capacity $f_{\max}$ measurements of the soil *in situ*, the **field value**, and estimates at the basin scale, the **effective value**, possibly never made much sense with each other, requiring numerous auxiliary hypotheses and negligibility premises (the **scale problem**, which we will address later). This was made clear in Betson's (1964) article [101], which achieved excellent fittings of a conceptual model to observed data only after *relaxing* the Hortonian model, implying a **partial contribution area** concept of surface runoff. From a rationalist perspective, Betson attempted to "save" Horton's theory by admitting *ad hoc* modifications to avoid its refutation. Even though statistically well-fitted, the results hinted that a small and relatively constant fraction of the basins in the Tennessee River valley produced surface runoff q_{si} (basins ranging from 500 ha to 8.4 km^2, North Carolina).

However, it was throughout the 1960s that the Infiltration Age faced its ultimate crisis in the scientific realm. Not coincidentally, this period witnessed the **International Hydrological Decade** (1965-1974), a United Nations (`UNESCO`) program developed to promote research in Hydrology. The revolution in perceptions within the scientific community occurred after a profusion of empirical evidence accumulated in the literature, reporting the observation of new mechanisms of hydrological response from hillslopes, which will be described in the following sections. At least three new rapid response mechanisms (or potentially rapid) can be summarized beyond those posited by Horton: (1) **exfiltration** $q_{\mathbf{ss}}$, (2) **direct rainfall** $q_{\mathbf{se}}$, and (3) **translational flow** $Q_{\mathbf{gt}}$. The first two refer to new rainwater contributing to rises. The last consists of old water stored underground, which also contributes to the rapid response (with surprisingly high prevalence rates). The nomenclature surrounding these mechanisms is somewhat confusing in the literature even today, perhaps due to the fact that water lacks a visible label, complicating differentiation out of context (for example, water emergence in springs can represent either a slow or fast response). In contrast to the Hortonian model, these rapid responses include surface and subsurface pathways controlled by other parts of the system beyond the topsoil, such as **macroporosity** (preferential lateral and vertical pathways in the soil) and topography (dynamic patterns of soil saturation). Essentially, empirical evidence in the second half of the twentieth century irreversibly exposed the complexity of processes in zero-order basins.

A landmark for the beginning of the end of the Infiltration Age was Mike Kirkby's (1969) article [102], which reviewed results from various experimental studies and presented a new way to understand and name hydrological processes in zero-order basins. At this point, Kirkby (1969) definitively marks the importance of rapid responses of water moving through the soil via a network of macropores (subsurface runoff). After more than a decade of accumulating new empirical evidence, a landmark for the rise of the new paradigm was Thomas Dunne's (1983) article [99], which definitively organized a new schematic view of rapid hydrological responses, illustrated in Figure 3.4, proposing a promising new research program in both pure and applied fronts. Dunne (1983) makes clear the idea that different climates, scales, and landscapes favor the *dominance* of one mechanism over another, even though they may occur simultaneously or alternate seasonally. For example, in semi-arid climates, factors such as long dry periods and the formation of soil crusts favor Horton's mechanism—the infiltration capacity $f_{\max}$ tends to be insufficient, and there are often no saturated areas in valley bottoms at the end of the dry season. In contrast, in humid tropical climates, the formation of deep soils or excess water during the rainy season may favor either mechanism. Ultimately, this revolution produced new understandings of the complex relationships between these mechanisms, as demonstrated by Jeffrey McDonnell (1990) [103] in the case of the MaiMai Experimental Basin (New Zealand). However, McDonnell (2013) [104] also critiques the entrenched paradigm: the hegemonic research program primarily focuses on *differentiating* the multiple hydrological responses, reaffirming the idea of complexity and uniqueness of each environment. While this attitude may continue *ad infinitum*, he argues that the true objective of hydrological science might be to produce *generalizations*, theories that are *unifying*. In this spirit, the prevailing hydrological paradigm might deserve the title of **Differentiation Age**.

3.3.1 The function of macropores

In the late 1930s and early 1940s, the scientific literature already recognized the weaknesses of Horton's model, asserting that the rapid response of small basins should include one or more *subsurface* mechanisms. Snyder (1939) [107], for example, suggests using the term direct runoff to denote the rainwater that contributes to the river flood *without ever having moved through the soil*. In this context, Barnes (1939) [108] divides the components of river flow into *three* (rather than *two*): (1) surface runoff; (2) storm seepage; and (3) base flow. By "storm seepage", Barnes referred to a rapid subsurface flow of rainwater that moves *laterally* through the vadose zone $\mathbf{V}$, feeding the stream channels at a much faster rate than expected from the aquifer detention time g:

> This consists of water which has penetrated only the upper soil-layers during a rainstorm or a thaw and has filtered more or less horizontally through the soil to discharge into the stream-system by seepage. It was observed by the writer in 1936 while analyzing discharge records of Zumbro River in Minnesota and called by him "secondary base-flow". – Bertram Barnes (1939, p. 721) [108].

Thus, a distinction is made between **perennial springs**, fed by the flow of true groundwater (unconfined aquifer), and **ephemeral springs**, fed by the subsurface flow of rainwater (suspended aquifer). This response mechanism was greatly reinforced by studies conducted by Charles Hursh at the Coweeta Experimental Forest (North Carolina, United States), with results reported for various forested basins ranging in size from 16 to 760 hectares [105], [109], generating the model illustrated in the details of Figure 3.5**c**. For example, Hursh & Brater (1941) [110] claim that surface runoff q_{si}

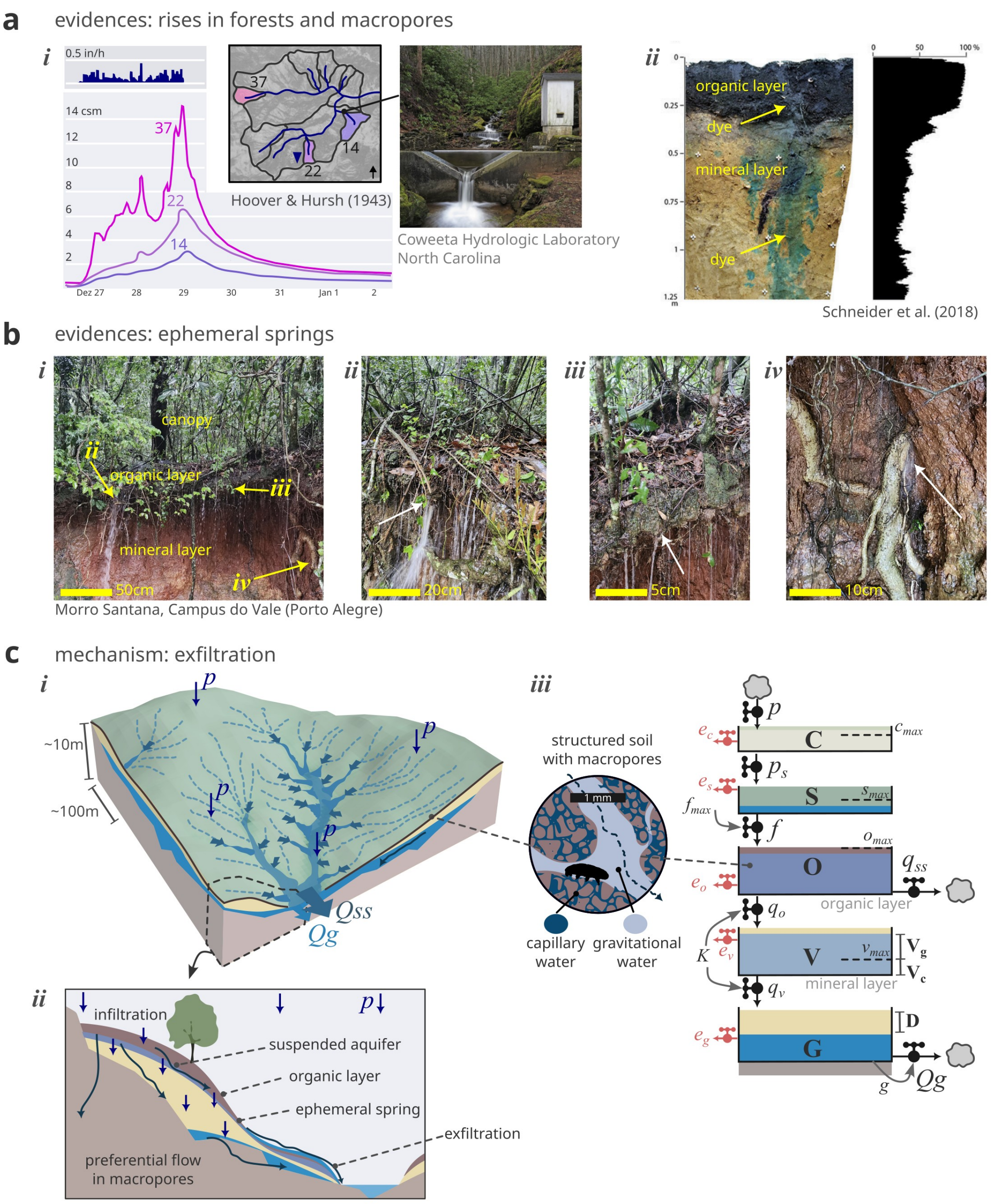

Figure 3.5 — **Rapid exfiltration through macropores.** Macropores and preferential subsurface pathways produce rapid exfiltration q_{ss} responses, especially in forests. **a** — Evidence: rises without runoff in basins at the Coweeta Experimental Forest (North Carolina, United States) reported by Hoover & Hursh (1943) [105] (detail i), and; the distribution of macropores in the soil profile highlighted by dyes, reported by Schneider et al. (2018) [106] (detail i). **b** — Evidence: ephemeral springs observed on June 16, 2023, at a road cut on the Vale Campus of the Federal University of Rio Grande do Sul (Morro Santana, Porto Alegre). Despite the extraordinary rain on that day (141.7 mm in 24 hours), no runoff was observed in the forest soil. i – profile of the bank; ii – preferential flow; iii – flow at the interface between horizons; iv – turbulent flow in a fracture in granite, created by a root. **c** — Systematization of the ephemeral springs mechanisms: rainwater infiltrates rapidly, creating a suspended aquifer at the transition between the organic and mineral horizons. This suspended aquifer emerges at various points in the basin, facilitated by macropores and preferential subsurface pathways, creating ephemeral springs during and shortly after rainfall. Source of the photograph in **a**: https://www.be-roberts.com/se/cwta/cwta1.htm .

was never observed on the hillslopes of one of the monitored basins, despite the rapid responses observed in river flow:

> Surface storm-runoff as overland-flow has not been observed on this drainage-area; nevertheless, characteristic flood-hydrographs are produced by heavy rains. – Hursh & Brater (1941, p. 863) [110].

This claim, followed by data in subsequent studies (see Figure 3.5**a**, in detail i), inevitably introduces a *counterexample* to the theory of infiltration capacity f_{max}. After all, if there is no surface runoff q_{si}, the only available response in Horton's perceptual model is the *slow* response caused by the recharge q_{v} from the aquifer, which is responsible for the recession curve. Among other mechanisms at work in the basin, detailed later, the authors point to the existence of *rapid subsurface responses* that include both water flow in highly permeable soil layers (unsaturated flow) and the formation of temporary suspended aquifers (saturated flow), which develop in different parts of the landscape during rainfall. In this regard, Hursh & Fletcher (1942) [111] emphasize the importance of soil macroporosity. Especially in forests, this property would help explain the dominance of preferential subsurface flows, significantly increasing the soil's gravitational water $\mathbf{V}_{\text{g}}$, in contrast to capillary water $\mathbf{V}_{\text{c}}$:

> The exact nature of this macro-pore space occurring in different horizons of the soil profile has yet to be described. It includes all large underground channels formed from decayed roots, fractured rock, insect and animal burrows, and larger spaces that may exist. It also includes macro-pore spaces formed through the complex structural patterns created by the aggregation of soil particles in the presence of organic materials. In the upper horizons of natural soils these biological openings and structural patterns built up from lattice-like aggregates are far more important in determining noncapillary porosity than the single grain soil particle size. Root channels and animal burrows are of particular significance in the detention storage and draining of gravitational water. A single earthworm burrow may be far more important in draining through a block of heavy soil than the entire cross sectional area of the pore space. In like manner, it is conceivable that a few continuous void spaces may give rise to rapid discharge of groundwater through a soil profile which, when viewed as a uniformly pervious medium, would be expected to transmit water slowly. – Hursh & Fletcher (1942, p. 485) [111].

Unlike recent studies with dyes, which clearly demonstrate the existence of macropores (as shown in detail ii of Figure 3.5**a**, with recent results from Schneider et al. (2018) [106]), the authors from Coweeta presented no evidence beyond observations of aggregated processes, such as rain, flow, and well levels. This research gap was permanently addressed in the 1960s when a new wave of studies provided more detailed quantitative results obtained through a more experimental approach than observational. Still in the context of the Coweeta Experimental Forest, Hewlett and Hibbert (1963) [112] used a lysimeter to demonstrate the critical role of water flow in the vadose zone $\mathbf{V}$ in sustaining base flow in streams. Whipkey (1965) [113] detailed lateral flows in the soil profile of a hillslope in Ohio (USA). The hillslope was monitored by a trench at its base, demonstrating the dynamics of exfiltration q_{ss}, especially in the upper organic layers, where high hydraulic conductivity K was measured due to the presence of macropores. Here, the function of **permeability transitions** between the soil horizons appears, especially between the organic horizon **O** (upper) and the mineral horizon (lower). This discontinuity generates a loss of hydraulic head that results in *lateral* flow in the vadose zone $\mathbf{V}$. I personally observed this process during a severe storm on June 16, 2023, at the Vale Campus of the Federal University of Rio Grande do Sul, at the base of Morro Santana, Porto Alegre (see details in Figure 3.5**b**). According to the `INMET` Station `83967`, June 16, 2023, recorded 141.7 mm of rain, which is above the monthly average for that month (around 130 mm). Despite such extreme rain and streams practically overflowing, I did not observe surface runoff in the forested areas of the campus, except where the terrain's channel forced the emergence of the suspended aquifer (i.e., through the expansion of riparian wet areas).

Other trench studies reaching similar conclusions to Whipkey's (1965) in the United States include: Ragan (1968) in Vermont [114]; Beasley (1976) in Mississippi [115], and; Harr (1977) in Oregon [116]. In the latter case, it was reported that exfiltration q_{ss} was 6 to 9 times greater in the upper soil layers than in the lower layers, corroborating the function of macroporosity. In the same study, the authors report that exfiltration q_{ss} was responsible for about 97% of the rapid response during rises. This was consistent with earlier results from Patric and Swanston (1968) [117], who cut down all the trees on a slope in Alaska and applied sprinkler irrigation. They observed no surface runoff q_{si} – the applied water traveled through preferential subsurface pathways, emerging rapidly at the base of the slope. In the British Isles, Weyman (1970) [118] reported that unsaturated subsurface runoff constitutes the main rapid response in an experimental basin, while Jones (1971) [119] noted that the widespread occurrence of the phenomenon known as *piping* – the formation of natural tunnels in the soil profile – contributes to high velocities in subsurface response. Consolidating this new generation of field research, results from the MaiMai experimental basin (New Zealand) established a new experimental research program, combining water balance, trenches, and the novelty of chemical tracers, dyes, and isotopes. In the case of MaiMai, Mosley (1979) [120] reaffirms (almost forty years after Hursh) the crucial role of macroporosity and natural tunnels in exfiltration q_{ss} while addressing some theoretical objections:

> Freeze [1972, p. 1282] considerou que um valor limiar de condutividade hidráulica saturada da ordem de 0,002 cm/s é necessário para que a exfiltration q_{ss} seja significativa, mas em um solo que contém canais de raízes, túneis e zonas de afloramento, a condutividade hidráulica saturada não é um fator limitante. O fluxo de corante traçador através de macroporos no solo foi observado a taxas até 3 ordens de magnitude maiores, e a resposta sensível e rápida da exfiltration q_{ss} às variações na precipitação sugere que o fluxo por macroporos, e não pela matriz do solo, contribui para as enchentes nos canais.[2] – Paul Mosley (1979, p. 806) [120].

3.3.2 The influence of topography

The Hortonian paradigm has not only been refuted by the recognition of exfiltration q_{ss}, as empirical evidence has also accumulated to support the existence of two additional, less intuitive rapid response mechanisms occurring in **riparian wetlands**. These mechanisms, illustrated in Figure 3.6, result from the interaction of rain with a shallow and dynamic groundwater table, one being direct rainfall q_{se}, and the other translational flow Q_{gt} (details in the next section). Both are interrelated, are strongly controlled by the terrain's topography, and also have consequences for the manifestation of exfiltration q_{ss} in macropores, as we will see later. The first of these emerges in the literature when direct precipitation in the area surrounding channels and springs is cited by Hursh & Brater (1941) [110] as one of the sources of river runoff in the basins of the Coweeta Experimental Forest:

> Contributions from areas of normally shallow water-tables located in close proximity to the stream, and occurring in soil-profiles which are

[2] Tradução livre de: *Freeze [1972, p. 1282] considered that a threshold value for saturated hydraulic conductivity of the order of 0,002 cm/s is necessary for subsurface stormflow to be significant, but in a soil that contains root channels, pipes, and seepage zones, saturated hydraulic conductivity is not a limiting factor. Flow of dye tracer through macropores in the soil was observed at rates up to 3 orders of magnitude greater, and the sensitive as rapid response of subsurface flow to variations in precipitation suggests that flow through macropores rather than through soil matrix contributes to channel stormflow.*

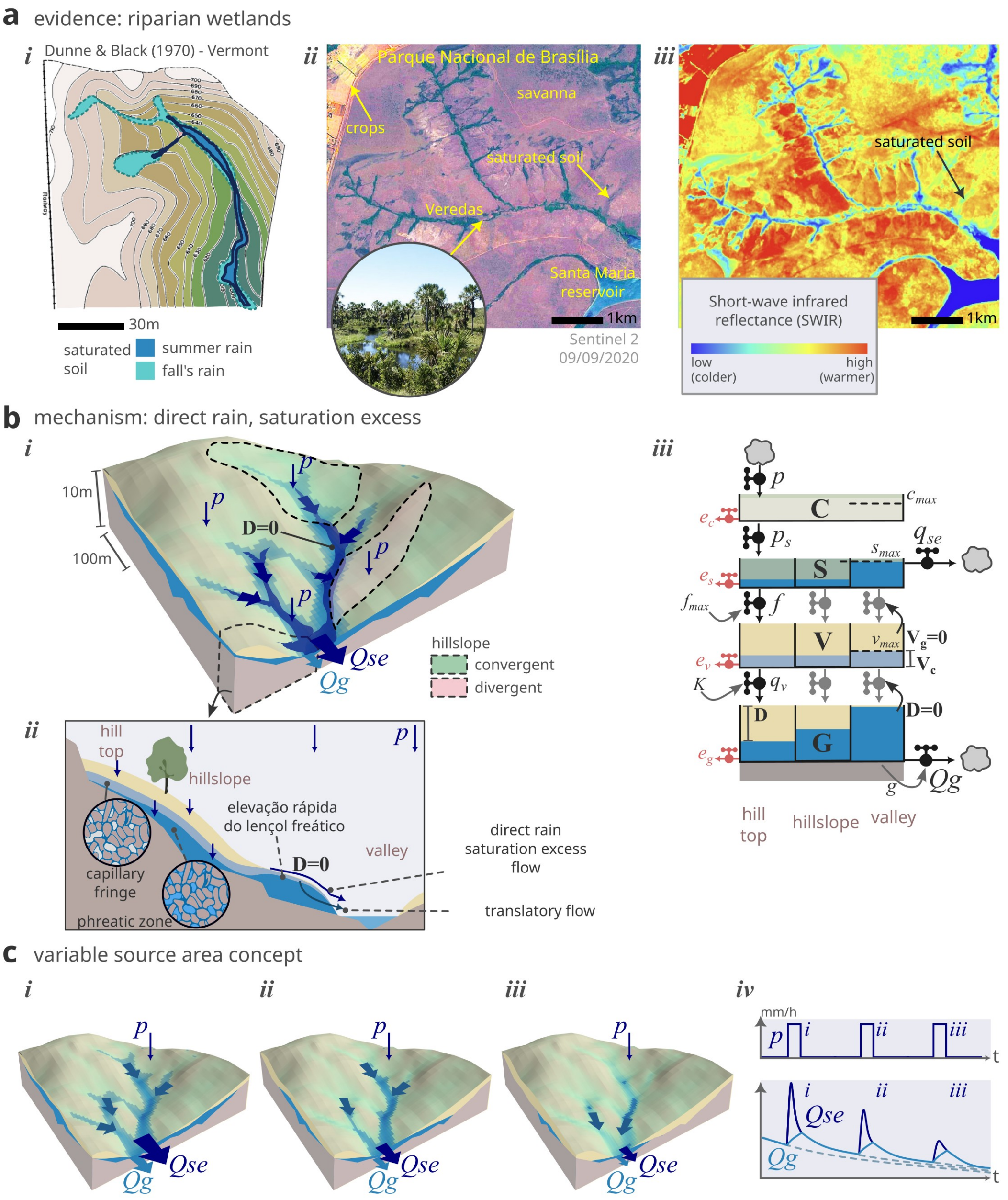

Figure 3.6 — Topography and the variable source area. Topography plays a crucial role in the formation of riparian wetlands that vary in extent during rainfall and throughout the seasons. These saturated soil areas thus produce direct rainfall q_{se}, as well as rapid subsurface responses from translational flow Q_{gt}. **a** — Evidence: map by Dunne & Black (1970) in Vermont (United States), demonstrating the extent of riparian wetlands at different times of the year (detail i), and; wet veredas in the dry season at the Brasília National Park, observed through short-wave infrared reflectance from a Sentinel-II scene on September 9, 2020 (details ii and iii). **b** — Systematization of the mechanism of runoff due to saturation excess. The soil in different parts of the basin (valley bottom, hillslope, and hilltop) saturates at different rates, generating rapid responses primarily in the valley bottoms. Convergent slopes tend to produce more direct rainfall q_{se} than divergent slopes, where recharge q_v and baseflow Q_g prevail. **c** — Schematic representation of the retraction of riparian wetlands during a dry spell (seasonal dynamics). This process also exhibits ephemeral dynamics during and shortly after rainfall events. Source of the photograph in **b**: https://commons.wikimedia.org/wiki/File:Veredas.jpg .

> quickly saturated. Where such conditions occur along a stream, it is expected that there will be an actual increase in the width of the channel and subsequent increase in the amount of channel-precipitation. Areas of high water tables adjacent to spring-heads would be expected to contribute similarly. – Hursh & Brater (1941, p. 870) [110].

This is certainly one of the pioneering descriptions of the concept of **variable source**

area: the generation of surface runoff q_{si} as a function of the *expansion and contraction of wet areas* in valley bottoms, adjacent to streams. This mechanism, illustrated in Figure 3.6**c**, allows any effective rainfall p_s to transform into excess rainfall p_x when it falls on saturated soil areas, which helps explain the prevalence of rises even in basins with high infiltration capacity $f_{\max}$ soils (as in the case of La Grange Brook, near where Robert Horton lived [100]). This concept was well organized by Cappus (1960) [121] in a study in the Alrance Experimental Basin (315 ha, France). The author claimed to have evidence for a "new theory of surface runoff", in which the basin area can be divided into a *surface runoff zone* and an *infiltration zone*. The former includes a *fixed* part of impermeable areas and a *variable* part of permeable areas that are nearly completely saturated with water:

> The experimental basin can be divided into two zones S_r and S_i of variable extents: — The runoff zone S_r of area A_r includes, on one hand, fixed extent impermeable zones (roads, paved paths, compacted dirt paths from repeated traffic of people or livestock, rocky surfaces, etc.) and, on the other hand, variable extent zones consisting of permeable land, but almost completely saturated with water. The rain that falls in the zone S_r entirely transforms into surface runoff q_{si} or subsurface runoff. — The infiltration zone S_i of area A_i consists of unsaturated permeable land. The sandy-textured soil, which forms the surface layers of the experimental basin, is characterized by a very high infiltration capacity $f_{\max}$ f that exceeds the intensity of all rains that may fall on this basin—except for those of extremely rare occurrence. Thus, except in very exceptional cases, the rain that falls in the zone S_i is completely absorbed by infiltration and consequently generates no runoff. [3] – Cappus (1960, p. 503) [121].

Tsukamoto (1963) [122] also structured a similar theory, based on results obtained in a basin at the University of Tokyo Forest (106.7 ha, Japan). In his paper, he points out that riparian areas exhibit rapid saturation responses due to the influence of the **capillary fringe** of groundwater, generating surface runoff q_{si} in this part of the slope, as opposed to the higher and well-drained parts of the terrain. The experimental results from Ragan (1968) [114], mentioned above, also demonstrated rapid rises in the water table near the stream during monitored rainfall events. Even Betson (1964) [101], while attempting to uphold the Hortonian perceptual model, noted that half of the runoff was possibly generated by a swamp area in one of the basins analyzed in his study.

To corroborate the theory of variable source area, Dunne & Black (1970) [123], [124] published detailed results for a small experimental basin in Vermont (United States). In this study, the authors present maps of saturated soil areas that follow the terrain's channel (detail i in Figure 3.6**a**), which showed variations both throughout the year (seasonal dynamics) and during rainfall events (ephemeral dynamics). The seasonal dynamics of these wet areas is explained by the increase in recharge q_v of groundwater during the wetter season, which expands the extent of spring emergence

[3] Freely translated from: *Le Bassin expérimental peut être partagé en deux zones S_r et S_i d'étendues variables: — La zone de ruissellement S_r de superficie A_r comporte, d'une part, des zones imperméables d'étendue fixe (routes, chemins empierrés, chemins de terre tassée par le passage répété des hommes ou du bétail, surfaces rocheuses, etc.) et, d'autre part, des zones d'étendue variable constituées de terrains perméables, mais à peu près complètement saturés d'eau. La pluie tombant sur la zone Sr se transforme entièrement en ruissellement superficiel ou hypodermique. — La zone d'infiltration S_i de superficie A_i est constituée par les terrains perméables non saturés. Le sol de texture sableuse, qui forme les couches superficielles du Bassin expérimental est caractérisé par une capacité d'infiltration f très forte qui dépasse l'intensité de toutes les pluies pouvant tomber sur ce bassin — à l'exception seulement de pluies d'une rareté extrême — ainsi, sauf en des cas très exceptionnels, la pluie tombant sur la zone S_i est entièrement absorbée par infiltration et ne donne lieu par conséquent à aucun ruissellement.*

zones in the valley bottoms. Detail *ii* in Figure 3.6**a** demonstrates evidence of seasonal dynamics from short-wave infrared (SWIR) reflectance with a Sentinel-II scene on September 9, 2020, in the Brasília National Park. This period is marked by dry conditions, but the riparian areas remain moist, forming the Veredas. On the other hand, the ephemeral dynamics are explained by a rapid rise of the phreatic zone **G** when the capillary fringe is very close to the surface (more details ahead).

The unequivocal observation of the dynamics of saturated areas by Dunne & Black (1970) solidified the perception that *topography* exerts an important control in the Hydrology of zero-order basins, and not just the *surface* of the soil, as postulated by the infiltration paradigm. In this sense, Anderson & Burt (1978) [125] demonstrated that, in a basin in the Quantock Hills (England), the rapid rise of the water table along the channels of the **convergent slopes** is much greater than in the **divergent slopes**. The former tend to generate relatively more direct rainfall q_{se} and also more exfiltration q_{ss}, as the sudden rise of groundwater activates the macropore network in the soil. In the divergent slopes, on the other hand, slower processes of recharge q_{v} and baseflow Q_{g} predominate.

In this context, the mechanism of surface runoff q_{si} defended in Horton's theory (infiltration excess) was not exactly refuted but reserved as a response mechanism limited to extraordinary precipitation events or areas with altered soil, whether in natural environments (such as rocky outcrops and arid regions) or anthropized environments (such as agriculture and urban areas). This implies that the use of the CN method from the SCS is justified when its application is directed towards extreme rainfall events or in urban and rural basins where the Hortonian mechanism is clearly dominant. However, this restriction is not explicitly stated in the official manuals of the method, which also include CN values for forests and other natural land covers. Additionally, simulation models such as SWAT and SWMM utilize continuous simulation (various rainfall events) and represent any land cover. It is worth noting that Horton (1936) [126] came close to analytically deducing this mechanism in one of his papers, as he highlights that sloping soil hillsides induce the water table to intercept the surface above the valley bottom level, causing the emergence of a saturated surface in this convergent part of the topography [127].

3.3.3 The water age paradoxes

Translational subsurface runoff, in turn, is conceptually speculated by Hewlett & Hibbert (1967) [77], in a clearly revolutionary article in the field of Hydrology [132]. While criticizing the hegemonic paradigm of the time (the infiltration capacity theory), the authors organize new relevant concepts, such as the terms "rapid and slow responses" and "variable contributing area", paving the way for the advent of the new paradigm. In this direction, the authors suggest a mechanism of *instantaneous* subsurface response that occurs when the field capacity v_{max} of the soil is exceeded by the infiltration of rainwater in the riparian zones, where there is greater influence from capillarity fringes. In summary, they postulate that this response, although rapid, would not exactly be rainwater but water that had settled in the soil matrix *before* the event occurred. In this process, the thickness of water films on soil particles in the vadose zone **V** suddenly reaches a limit where the pore network becomes pressurized by gravity. Therefore, the **new water** from the rain (event water) triggers a pulse, a pressure wave, that expels the water stored in the soil at the base of the slope, referred to as **old water** (pre-event water):

> However, of the part contributed to direct runoff, a fraction will be some of the actual drops that fell during the storm – that is, some new rain –

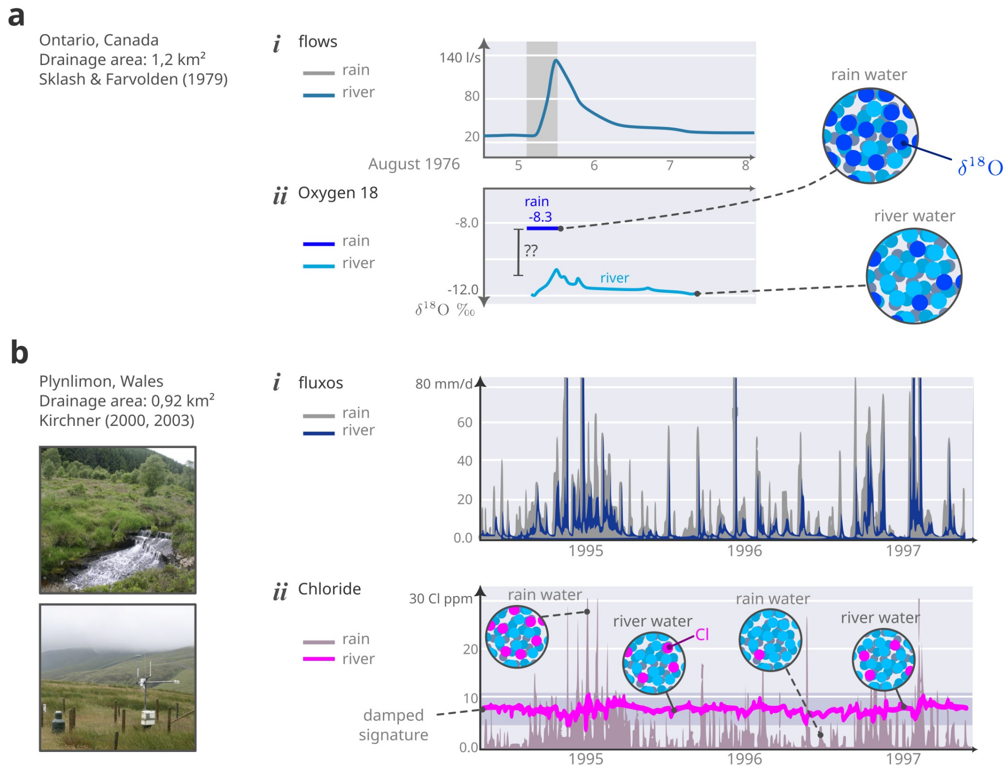

Figure 3.7 — The old water paradox. Analyses of isotopic signature and geochemistry of rainwater and river water during rises make it clear that they are waters of different ages, thus creating a paradox. **a** — Evidence provided by Sklash & Farvolden (1979) [128] in a rural basin in Ontario, Canada (1.2 km² drainage area). The flows clearly denote a rapid response from the basin (flood) to the rainfall event (detail i). However, the isotopic signature with Oxygen-18 shows that the water in the flood is not the same as the rainwater (detail ii). **b** — The same paradox observed with Chloride (marine aerosol) by Kirchner *et al.* (2000, 2003) [129], [130] in an experimental basin in Plynlimon, Wales (0.92 km² drainage area). The flows are typical rapid responses (detail i), but the river signature shows a pronounced damping over the years, suggesting a long-duration mix (detail ii). Source of the photographs: Ecological Continuity Trust [131].

> and the other fraction will be what we might call translatory flow, or flow produced by a process of displacement. This is a contribution to direct flow of water already stored in the soil mantle before rainfall began. It will be released in large quantities only when the soil is within field capacity range or wetter. Above the zone of saturation, we may regard such movement as due to thickening of the water films surrounding soil particles and a resulting pulse of water flux as the saturated zone is approached. – Hewlett & Hibbert (1967, p. 279) [77].

Although the theory makes sense and cites laboratory studies, the text by Hewlett & Hibbert (1967) does not provide empirical evidence obtained in the field to justify the reality of this mechanism. However, this gap was filled by Pinder and Jones (1969) [133], who evaluated the separation of flood hydrographs in three monitored basins in Nova Scotia (ranging from 647.5 ha to 13.5 km², Canada). Unlike conventional graphical methods, they inferred the separation between surface runoff q_{si} and subsurface flow using chemical tracers and a simple mass balance model[4]. In the presented case, sodium, calcium, bicarbonate, magnesium, and sulfate concentrations were monitored before and during flood events. The results indicated a substantial prevalence of groundwater flow during rises, accounting for 32% to 42% of the maximum hydrograph flow. However, this did not eliminate the alternative explanation of

[4]The developed model consists of two mixing compartments: $C_{tr}Q_{tr} = C_{dr}Q_{dr} + C_{gw}Q_{gw}$, where: C is the concentration of some conservative solute; Q is the flow; tr denotes total flow; dr denotes direct runoff, and; gw denotes groundwater flow.

a subsurface response from rainwater (new water) that dissolved the monitored solutes while rapidly moving through the soil. Evidence in favor of groundwater (old water) became much more robust with the advent of hydrogen and oxygen isotope monitoring, such as Deuterium (^{2}H), Tritium (^{3}H), and Oxygen-18 (^{18}O), which are ideal markers that are part of the water molecule itself[5]. In locations with significant variability in the **isotopic signature** of the precipitated water, it is possible to estimate how much of this new water is present during rises[6]. This strategy was suggested by Dinçer et al. (1970) [134] in a study in the mountains of Czechoslovakia that demonstrated the effect of **thermal fractionation**[7] on the concentrations of ^{3}H and ^{18}O in snow layers that precipitated and melted over the seasons. Subsequently, the results published by Martinec et al. (1974) [136] noted that river water in the Swiss mountains exhibited a relatively low variability in ^{18}O concentration, approaching the long-term average of seasonal oscillations observed in precipitation. The strategy then took on well-defined contours with the article by Sklash et al. (1976) [135], which, in addition to organizing the logic of the method, showed that in two monitored basins in Ontario (Canada), the contribution of groundwater to the maximum flood flow ranged from 55% (in upstream basins) to 70% (in downstream basins draining an area of 700 km^2). These results carry revolutionary implications:

> The most important finding is that the pre-storm component of storm runoff for the 16 May storm was large. For example, at peak total discharge, the pre-storm component of Big Otter Creek at Vienna was 70 ± 9% of storm runoff. These results substantiate the findings of Pinder & Jones (1969) and Fritz *et al.* (1974), even though the basins in the present study are one to two orders of magnitude larger in areal extent.These results are not consistent with the simulated results of Freeze (1972b), the field results of Dunne & Black (1970a,b), and Hewlett & Hibbert (1967), or the theoretical implications of Horton (1933). The results are particularly encouraging, though, in light of the large subsurface (prestorm) component of snowmelt noted by Dinçer *et al.* (1970).
> – Sklash et al. (1976, p. 276) [135].

Michael Sklash continued studies of this type, corroborating the existence of this process in basins in Canada, New Zealand, and the British Isles [128], [137], [138]. For example, the article by Sklash & Farvolden (1979) [128] in Canada presents similar results for a basin with intensive agriculture (1.2 km$^{\mathbf{2}}$, in the Hillman Creek Experimental Basin, Ontario, Figure 3.7**a**) and two highly forested basins (1.0 km$^{\mathbf{2}}$ and 3.9 km$^{\mathbf{2}}$, in the Ruisseau des Eaux Volées Experimental Basin, Québec). In addition to reporting the surprising prevalence of old water in rises (between 80% to 94% of total runoff), the authors contribute to the theory of rapid rises in groundwater in valley bottoms to explain the phenomenon. In the MaiMai Experimental Basin (New Zealand), Sklash et al. (1986) [137] present results that drastically revise the interpretations of Mosley (1979) [120]. As mentioned above, Mosley (1979) argues for the dominance of exfiltration q_{ss} (new water) in this basin. The unequivocal prevalence of old water in rises, obtained

[5]Unlike common solutes, the concentration of ^{18}O is measured by the difference in parts per thousand (δ^{18}O ‰) of the ratio of ^{18}O/^{16}O of a standard and the sample: $\delta^{18}\mathrm{O} = (\frac{^{18}\mathrm{O}/^{16}\mathrm{O\,sample}}{^{18}\mathrm{O}/^{16}\mathrm{O\,standard}} - 1) \times 1000$. The standard is usually the mean ocean water (`SMOW`). Waters depleted of ^{18}O relative to the standard exhibit negative δ^{18}O, and vice versa.

[6]Obviously, if the rainwater is isotopically indistinguishable from the river water just before the event, then it is impossible to extract any relevant information.

[7]The main cause of the fractionation of these isotopes in the atmosphere arises from the difference in vapor pressure between isotopically heavy and light water molecules: $H_2{}^{18}O$ has a lower vapor pressure than $H_2{}^{16}O$ and, therefore, $H_2{}^{16}O$ remains preferentially in the liquid phase during both evaporation and condensation processes. Thus, the observed concentrations of δ^{18}O in precipitation tend to be incrementally negative as moist air masses move over continents [135].

with isotopic markers, created a certain impasse in the scientific community, which since then has proposed plausible mechanisms [139]. In this context, McDonnell (1990) [103] synthesizes the response mechanisms in the MaiMai basin, introducing the concept of **activation of subsurface flow** from the entry of rainwater into the macropore network of the hillslopes, i.e., through exfiltration q_{ss}. In this scheme (illustrated in Figure 3.8**a**), he emphasizes the role of *vertical* shortcuts in the soil profile, created by macropores, which allow the new rainwater to rapidly lodge in the capillary fringes of the phreatic zone $\mathbf{G}$, thereby activating the hydraulic head necessary to quickly expel the old groundwater at the base of the slope. A new review by McGlynn et al. (2002) [140] also highlights the relevant effect of the **bedrock topography** of the slope (the underlying relatively impermeable bedrock). It is suggested that the irregularities of this layer can create **stagnation zones** or **pockets** that store water underground for much longer than expected. The eventual hydraulic activation of these relatively inaccessible parts expels old water at the base of the slope, facilitated by the presence of macropores. Indeed, the influence of underlying geological structures (fractures) on the emergence of groundwater had already been mentioned by Huff et al. (1982) [141], but without analyzing the age of the water using isotopes.

The evidence, impasses, and plausible mechanisms proposed to explain the dominance of old water in rapid responses further increase in complexity given the results of **geochemical signature**, which tend to exhibit high variability. In this sense, Burns et al. (2001) [143] suggest that the surface responses in the Panola Mountain Experimental Basin (Georgia, United States) end up mixing with groundwater in the riparian zone before entering the channels. Seibert et al. (2003) [144] also emphasize the difference in geochemical signature between the water in the riparian zone (anoxic conditions) and the water in well-drained hillslope soils (greater aeration). This complexity has led to some perplexity, expressed by Kirchner (2003) [130] in the so-called *double paradox of catchment hydrology and geochemistry*, or simply **old water paradox**. For him, this paradox has two components that, although related, are somewhat contradictory: (1) Hydrology: the rapid mobilization of old water—the quick replacement of old water by new, as postulated by translational flow Q_{gt}, and; (2) Geochemistry: the chemical variability of old water—the fact that old water assumes different chemical signatures depending on the flow velocity. In this sense, based on chloride concentrations[8] monitored in a basin in Wales, Kirchner et al. (2000) [129] propose a **hydro-geochemical compartmentalization** of the soil, where pores and fractures exhibit a fractal structure of residence times (see Figure 3.7**b**). This implies why the slopes of the basins transmit *hydrological signals* much faster than *geochemical signals*. This concept becomes clearer through Iorgulescu et al. (2007) [145], who reinforce the difference between the *wave* speed (**celerity**) of water and the *molecular* speed of water—beyond material, flood flow is an energy flow. In the same spirit, McDonnell (2014) [146] also draws a new perspective on evapotranspiration flows, particularly regarding the age of the water that plants consume, proposing the possibility of a **hydro-ecological compartmentalization** of the soil, which he refers to as **two-world hypothesis**. In this situation, the water consumed by the roots and **fine roots** of plants (known as *green* water) would be capillary water $\mathbf{V}_{\mathrm{c}}$, relatively older than gravitational water $\mathbf{V}_{\mathrm{g}}$ (see Figure 3.8**b**). Evaristo et al. (2015) [147] provide evidence supporting this hypothesis, showing that ecological separation is common in various biomes—plants use soil water with an isotopic signature distinct from the water contributing to the recharge q_{v} of groundwater and to river runoff.

[8] An analogous strategy to that used with isotopes is possible in basins with marine aerosol deposition, with chloride being an inert marker that can be analyzed in rainwater.

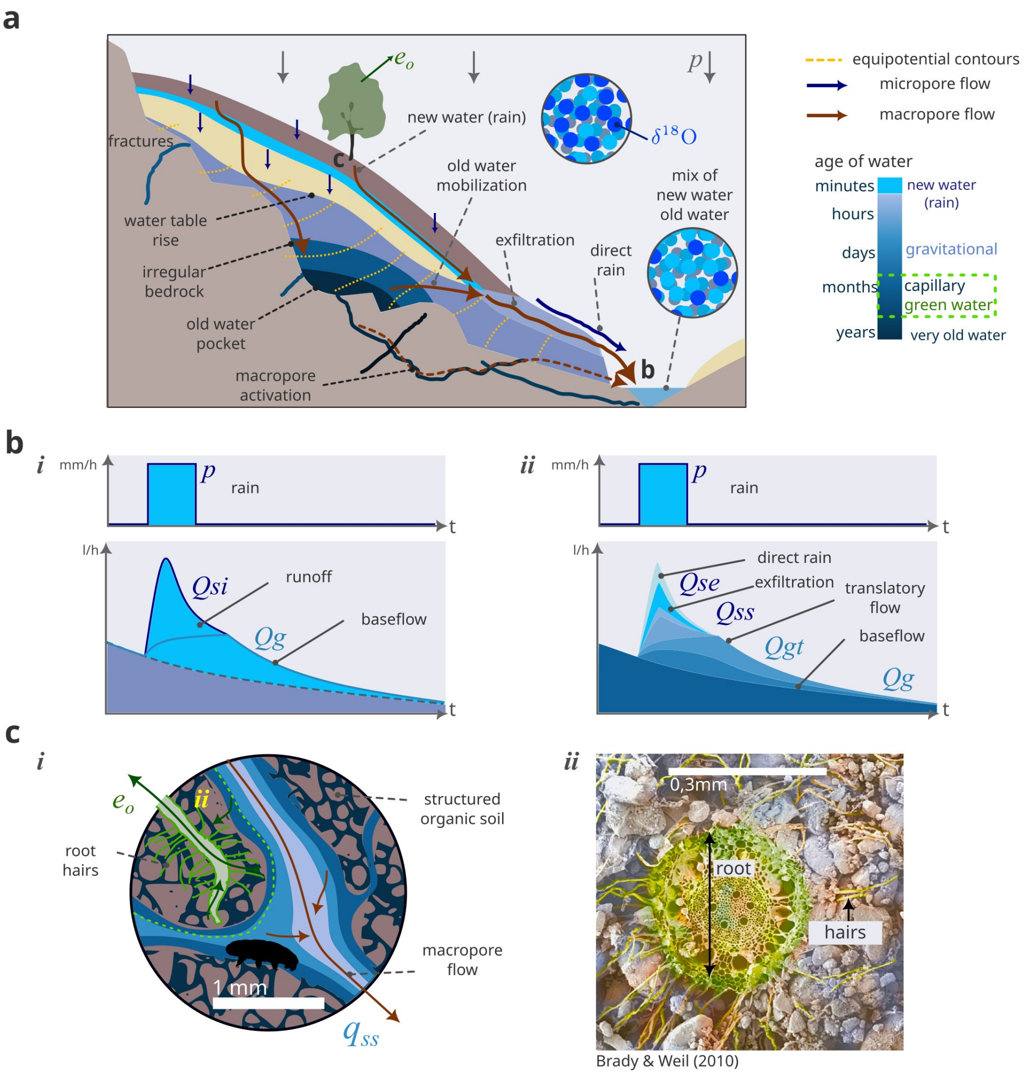

Figure 3.8 — Mobilization of old water. a — The rise of the water table combined with flow through macropores (fractures and preferential pathways) helps explain the dominance of old water in the rapid responses of rivers. Vertical shortcuts allow new water (rain) to mix with older water in stagnation zones (pockets) at the aquifer's threshold, activating hydraulic head in fractures and preferential pathways, resulting in the rapid expulsion of old water. **b** — Differences between the Hortonian model of rapid response (Type 3 flood) and the model obtained from isotopic analyses. **c** — Water in the soil also shows an age differentiation, which is perceived in the evapotranspiration flows of plants (detail i). The capillary water $\mathbf{V}_c$ is absorbed by fine roots, while gravitational water $\mathbf{V}_g$ drains to the base of the slope (detail ii). Photograph from electron microscopy adapted from Brady & Weil (2010) [142].

3.4 Limitations of hydrological models

Accompanying the scientific revolution brought about by empirical evidence regarding runoff mechanisms on hillslopes, the 1960s was also marked by the advent of the first hydrological models simulated on digital computers. This occurred largely due to a confluence of factors, such as the intellectual context of von Bertalanffy's General Systems Theory and the practices of Jay Forrester's Systems Dynamics, following the emergence of digital mainframe computers. Keith Beven reports that by 1971, he had counted over a hundred hydrological models in the literature, which were basically versions of the Stanford University model, the *Stanford Watershed Model IV* (`SWM`) [148]. The model was developed starting in 1959 as Norman Crawford's doctoral thesis, supervised by Ray Linsley, and later gave rise to a program called the *Hydrologic Simulation Program, Fortran* (`HSPF`), developed for and with the support of the U.S. Environmental Protection Agency (`USEPA`) [149]. This pioneering model clearly exemplifies the influence of the ontology offered by Systems Dynamics: a network of reservoirs connected by flows that is solved numerically. In this case, the SWM is only slightly more intricate than the minimalist model presented in the previous chapter, with four reservoirs (canopy, vadose zone **V**, phreatic zone **G**, and drainage channels) and three response mechanisms (including a subsurface response in addition to surface runoff q_{si} and groundwater flow). However, the model does not represent water storage on the surface, nor does it differentiate topographic aspects in a way that surface runoff can be separated into surface runoff q_{si} or direct rainfall q_{se}. But this is not exactly a problem in Systems Dynamics; it is sufficient to add a new compartment and connect the flows. The flexibility provided by Systems Dynamics, in this sense, introduced it as a conceptual and procedural paradigm in hydrological modeling, resulting in both a proliferation of models from this period and an increased theoretical understanding of the importance of **scale** throughout the modeling process.

3.4.1 On data and processes

Before the advent of hydrological simulations, however, the approach to obtaining flood hydrographs from rainfall data primarily relied on the concept of **Unit Hydrograph** of the basin, introduced by Sherman (1932) [150]. This concept is based on the theory that the hydrological response of a basin can be summarized as a linear process of kinematic propagation through the channel network from a rainfall pulse, which can be reduced to a unit pulse. According to the **principle of superposition**, more complex rainfall pulses can be integrated over time (convolution method). In this case, the fundamental parameter of a basin is its **time of concentration**, which is relatively larger in elongated basins than in rounded basins, even if they have exactly the same area. The systemic view allowed this process to be represented as a series of connected reservoirs, a "cascade," resulting in the parameterization of a Gamma distribution, or **Kalinin-Miyukov-Nash model**:

$$Q(t) = \frac{\nu}{k\,\Gamma(n)} e^{-t/k} (t/k)^{n-1} \tag{3.1}$$

Where ν [L^3] is the volume of the hydrograph; n [-] can be interpreted as the effective number of reservoirs; and k [T] can be interpreted as the average residence time of the reservoirs. Reportedly, this parameterization was first obtained independently by Kalinin & Miyukov (1957) [151] in the Soviet Union, and later by Nash (1958) in England [152]. From this, the notion emerged that the response of the basin is analogous to a *function* or *filter* that acts on the rainfall signal (or other input signals). This approach to obtaining hydrographs, which represents a modeling paradigm with its own ontology, evolved into what Todini [153] refers to as **data-driven models**, in contrast to **process-driven models**. The data-driven models today encompass a set of techniques that include, for example, artificial neural networks. This approach is visibly contaminated by the fluvialist bias; after all, it is impossible to *explain* exactly where and how runoff was generated based on a truly hydrological theory—the basin is treated as a black box. In this regard, Todini argues that this family of models sought to maximize **predictive capability** at the expense of **explanatory capability**, that is, to produce results that have "physical meaning." An attempt to re-establish the explanatory capability of data-driven models is the modeling approach termed *Data Based Mechanistic* (`DBM`), schematized by Young (2002) [154]. This technique results not only in predictions of flow but also identifies internal structures and parameters that possess explanatory capability. In this context, Todini argues:

> Although the DBM modelling approach recognises the importance of the physical coherence of the identified model structure, it derives it from the

> observations, thus disregarding de facto the results of at least 50 years of research efforts aimed at specifying the physical hydrological mechanisms that generate rises. This contrasts with the Bayes principle which would combine the observations with all possible a priori knowledge on the hydrological processes and possibly on the parameter values to obtain less uncertain a posteriori forecasts. – Todini (2007, p. 471) [153].

As highlighted in Chapter 1, *models are symbolic vehicles for theories*. In this sense, data-driven models are, in their essence, **statistical models**: they establish a theory about the data *themselves*, about their internal relationships. As mentioned in Chapter 2, such models tend to be *overfitted to the data*, allowing for good interpolations but making extrapolations problematic, not contributing to the learning process that Systems Dynamics offers. Process-based models, on the other hand, instantiate a representation of a *target system* that exists in an objective reality *beyond* the data—the watershed. Therefore, a truly hydrological model, based on processes, is capable of simulating the behavior of a basin *even without any empirical observation available* (a synthetic scenario, for example), as modeling is a process of **deductive inference**. The role of empirical evidence, in this sense, is to reject or corroborate the theory conveyed by the model.

3.4.2 The incommensurability dilemma

Despite the appeal in terms of explanatory capability, process-driven models, enabled by Systems Dynamics, have also begun to demonstrate their limitations, especially in light of the evidence supposedly associated with the parameters. Even when representing known hydrological response mechanisms, the highly *aggregated* nature of the instantiated compartments has made it increasingly clear that defining the parameters of a hydrological model to achieve good results is not a trivial practice, requiring a long process of trial and error, marked by many nuances[9]. To make matters worse, the parameters values that produced results adhering to empirical observations rarely coincided with field-measured values. For instance, Amorocho & Hart (1964) [155] draw attention to unrealistic results obtained internally within this type of model, due to **compensating effects** in the mass balance imposed on the compartments. For these reasons, Todini suggests that calibrating hydrological models with optimization methods without greater concern for the physical coherence of the parameters ultimately transforms a process-based model into a data-based model, as the focus tends to be on adjusting input data (rain) and output (flow), rather than explaining phenomena in an objective reality [153]. This limitation arises from two inexorable and inseparable problems in hydrological modeling: (1) the **equifinality problem** and, (2) the **scale problem**.

The equifinality problem was explored in Chapter 1 (Section 1.6), being a milder version of the underdetermination problem of theories that postulate unobservable entities. The term "equifinality" was introduced by von Bertalanffy in General Systems Theory (Chapter 2, Section 2.3), conveying the notion that open systems systems can converge to similar structures. In modeling, it is associated with the fact that systems with different structures or even parameters can exhibit similar *behaviors*, as in the case of slow responses illustrated in the prototype model from Section 2.5. Thus, the calibration process of a model with partial information about its processes (only

[9] According to Keith Beven, in the early days of rainfall-runoff modeling, there was a story that the only person who could truly calibrate the Stanford Model, with all its parameters, was Norman Crawford, who wrote the original version of the model as part of his doctoral thesis (Beven 2012, p. 233 [58]).

observed flow, for example) **does not** guarantee that other internal processes are adequately represented in empirical terms—hence the discrepancy between observed and adjusted parameters. But even if *complete* information exists, the scale problem, discussed below, implies that the differences between the scale represented by the model and the scale of observations are *incommensurable*, or incompatible, introducing the commensurability error ε_Δ in the results of the model (see Equation (1.5), the total error equation). It is noteworthy that the issue of scale similarity was a problem promptly recognized in the field of reduced scale models, but was only appreciated from the 1980s onward in hydrological modeling.

3.4.3 Approaches using vector fields

In light of the difficulties in reconciling field observations with adjustments of modeled systems and the increasing computational capacity available through *mainframes*, Freeze & Harlan (1969) inaugurated a new perspective on hydrological modeling, originating what they termed **physically based models**. This form of modeling, like that in Systems Dynamics, is based on the description of processes. The difference, however, is that the processes described by these models are derived *directly* from laws postulated by Physics: the conservation of mass, momentum, and energy. The article by Freeze & Harlan established a "design" of a physically based model that fundamentally differs from Systems Dynamics in its ontological aspects. Unlike the systemic paradigm, which is based on aggregated compartments connected by flows and feedbacks, the physical paradigm consists solely of **velocity vector fields** that act continuously, distributed in three-dimensional space $\mathbb{R}^3$ and modulated by initial and boundary conditions. With this, the authors aimed to provide a superior alternative to systemic models:

> With hydrologic systems models, it is possible to simulate streamflow hydrographs with a high degree of accuracy for a variety of hydrologic and geographic conditions. The Stanford Watershed Model IV (Crawford and Linsley), is the best-known and most successful model of this type. If the model we espouse is to offer promise for the future, it must be able to compete with the systems approach in terms of practical results and utility. A case could then be made for its superiority on the basis that a better understanding of the internal processes and their effects on the overall hydrologic system is desirable and could be beneficial to the solution of practical problems. – Freeze & Harlan (1969, p. 242) [156].

In other words, the authors believed that the solution to avoid the apparent problems in the calibration process of Systems Dynamics models was to apply the laws of Physics (Fluid Mechanics) directly to describe the hydrological cycle in the basins—after all, there was no need to reinvent the wheel. The only hindrance might have been the available computational capacity, although on the other hand, it would not be necessary to calibrate the models through any intensive search method, as the *truly* physical parameters could be defined *a priori*, such as channel roughness or hydraulic conductivity. Another promised advantage was the ability for continuous integration among the parts of the system, such as surface runoff q_{si} and subsurface flow. They pointed out that, although certain processes of the hydrological cycle at the time still lacked physically based studies (such as evaporation processes), unidimensional flow in channels and three-dimensional flow in porous media were already well established by the St. Venant and Darcy-Richards equations, respectively. Variations for different boundary conditions or negligibility premises could be developed, and solutions obtained in new theoretical studies.

A good example of the physically-based approach (and its problematic aspects) is the modeling of flow in porous media, specifically water in soil. In this case, the logic emerges from **Darcy's Law**. This law was experimentally derived by Henry Darcy (1803-1858) using a pipe filled with sand, where he observed that the flow of water in the pipe Q [L^3T^{-1}] is directly proportional to the cross-sectional area of the pipe A [L^2] and to the difference in hydrostatic potential between the inlet and outlet Δz [L] [157]. At the same time, the flow is inversely proportional to the length of the pipe l [L]. To transform these relationships into a dimensionally consistent equation, the **hydraulic conductivity** K [$L\,T^{-1}$] is introduced [10]:

$$Q = K\frac{A}{l}\Delta z \tag{3.2}$$

This is an analysis at the **global scale**, that is, evaluating the *aggregated* behavior of the system of the pipe. But then a crucial analytical move is made to migrate to the **local scale**. This is done by *assuming* that it is possible to represent *infinitesimal elements* of the soil, which leads to the definition of the gradient of hydrostatic potential $\nabla\Phi$ [LL^{-1}]:

$$\nabla\Phi = \frac{\Delta z}{l} \tag{3.3}$$

Therefore, from Equation (3.2) it follows that:

$$Q/A = K\nabla\Phi \quad \Rightarrow \quad u = K\nabla\Phi \tag{3.4}$$

Where u [$M\,T^{-2}$] is the **Darcy velocity**[11] of the fluid. For a three-dimensional spatial domain $\mathbb{R}^3 = \{x, y, z\}$:

$$u_x = -K\frac{\partial\Phi}{\partial x} \quad u_y = -K\frac{\partial\Phi}{\partial y} \quad u_z = -K\frac{\partial\Phi}{\partial z} \tag{3.5}$$

Which makes the Darcy's Law assume the following differential and vector notation[12]:

$$\mathbf{u} = -K\nabla\Phi \tag{3.6}$$

The maneuver to *collapse* the global scale into a local scale of infinitesimal elements is a typical **idealization** of the Galilean type, where a mathematical representation in a *limit condition* is deduced from the representation of an *observed condition* (see Section 2.2). Galileo used the inclined plane to then idealize the limit condition of the vertical angle for freely falling objects. In the case of flow in porous media, the Darcy's Law for a pipe filled with sand assumes the form of Equation (3.6) in the limit of infinitesimal soil elements. The complete formulation to describe the movement of water in soil, including flows in the vadose zone $\mathbf{V}$, is described by the Richards Equation (or Darcy-Richards). Richards (1931) [158] coupled the Darcy Equation with the mass balance in the local scale (in the assumed infinitesimal elements), producing a system of partial differential equations that need to be solved over time and in three-dimensional space[13].

[10] For any fluid and any porous medium, K is defined as: $K = \frac{c}{\mu}$, where c [$M\,T^{-2}$] is the permeability of the porous medium, and; μ [$M\,L^{-1}\,T^{-1}$] is the viscosity of the fluid.

[11] The **real velocity** of the fluid is higher since the fluid must flow through a relatively smaller section, where there are connected pores.

[12] The negative sign denotes that the direction of velocity is opposite to the gradient of hydrostatic potential.

[13] The Richards Equation can take different notations, but generally it establishes the expansion of hydrostatic potential to include, in addition to gravitational potential Δz, also the capillary potential of water, so that: $\Phi = \Delta z + \psi$. Thus, the hydraulic conductivity of the fluid becomes variable in saturated conditions, even exhibiting hysteresis effects.

The innovative modeling proposal made by Freeze & Harlan (1969) was explicitly termed a "project", as it was not readily operational. However, it already pointed to directions for new research in both theoretical and applied fronts so that a fully integrated model could eventually be realized beyond the equations. This process was, in part, led by Freeze himself, in a series of articles where he presents the results of various experimental simulations in the realm of groundwater flow [159]. In a typical demonstration of exploratory modeling, Freeze begins this movement by organizing the theoretical mathematical details (the differential equations) and numerical details (the solution methods) to simulate transient flow in unsaturated porous media within the three-dimensional domain of an idealized slope [160]. The result obtained by Freeze consists of a solution using the **finite difference method**, with a regular **computational grid** that can be applied to any surface geometry of slope and geological subsurface pattern (for example, impermeable beds and different soil horizons). Alternatively, Beven (1977) [161] demonstrated that it is also possible to implement numerical solutions using the **finite element method**, employing an **irregular computational grid**. With a focus on the plan and profile of the simulated variables, virtual experiments with models of this type show in detail the behavior of groundwater in response to spatial patterns of rainfall and water extraction by wells or channels. In subsequent advances, Freeze seeks to engage with the empirical evidence regarding the new flow mechanisms being reported by the scientific community at the end of the 1960s, emphasizing that the physically based model developed naturally produced such phenomena, depending only on the boundary conditions, that is, the geometry of the slope [162], [163]. In this context, the physical theory would indicate that exfiltration q_{ss} would be dominant in convex convergent slopes (incised valleys) while saturation excess would dominate in concave convergent slopes (amphitheater-shaped valleys).

3.4.4 El Dorado – crisis in Hydrology

The project envisioned by Freeze & Harlan (1969) was thus realized in various models, some more and some less integrated with the hydrological cycle, including models like `HEC-RAS` (focused on surface runoff) and `MODFLOW` (focused on groundwater) [164]. Among the pioneering and fully integrated models, the model Système Hydrologique Européen (SHE) stands out, developed starting in 1976 through a collaboration between the Danish Hydraulic Institute, the British Hydrology Institute, and the French consulting company SOGREAH. After ten years of development, operational results began to be released, and the structure of the model was published in a series of articles in 1986 [165], [166]. According to its authors, the model was explicitly inspired by the project of Freeze & Harlan (1969), although they implemented a simplified version of the flow in the vadose zone $\mathbf{V}$, with a unidimensional formulation of the Darcy-Richards equation. Despite all the effort allocated and the computational complexity compared to models based on Systems Dynamics, the authors of the model `SHE` readily acknowledge its limitations, especially the scale problem:

> In principle, because the parameter values are based on physical measurements, models such as the SHE should not require calibration. In practice, though, problems such as inadequate representation of the hydrological processes and the possible difference in scale between the measurement and the model grid square mean that some calibration is likely to continue to be required. In a SHE context this is regarded as a selective improvement of initial parameter estimates, by a comparison between observed and simulated hydrological variables, e.g. stream discharges or phreatic surface levels. At present this is carried out on a trial and error basis. – Abbott et al. (1986, p. 53) [165].

This fact clearly breaks the promise made by Freeze & Harlan (1969) that a physically-based model would be free from such nuances, with the definition of parameters made *a priori*, without the need for manual or automated adjustments *a posteriori*.

The practical limitations of the model SHE opened a gap for a crisis in hydrological modeling, providing inputs for a theoretical and philosophical discussion on scale and uncertainty issues in the 1990s and 2000s. This crisis was laid bare in the critical essay by Beven (1989) [167], who systematically organizes the problems of physically based models. At this point, Beven points out that, in practice, physically-based modeling applies a **scalability premise**, which is as idealizing as the other simplifications seen in Systems Dynamics, with the advantage that the latter is more intuitive. For example, Beven cites the application of the model SHE in a catchment in England that instantiated computational grid elements with a length of 250 meters, as if the physics of velocity fields were applicable at this scale. The variables simulated in a mesh element hundreds of meters long are clearly not commensurable with point empirical evidence. Furthermore, even with relatively small mesh elements (on the scale of centimeters), the models underrepresent the processes that are known to occur *below* this scale. Unlike free flow in channels or in extensive, homogeneous aquifers, which are well represented by velocity fields, empirical evidence about macroporosity in slopes with structured soils brings fundamental incompatibilities with the ontology of physically based models [148], [168]. As emphasized by Hursh & Fletcher (1942), cited above, "A single earthworm tunnel can be much more important in draining a mass of soil than the entire cross-sectional area of the porous space". By instantiating continuous velocity fields, the Darcy-Richards equation simply does not capture the local complexity of the soil's macropore structure (or, equivalently, fractures in a fractured aquifer). From a scientific standpoint, Kirchner (2006) [169] reminds us that elegant differential equations do not guarantee good results for good reasons—this is a role reserved for empirical evidence and hypothesis testing.

With the advent of strong criticisms and discussions, the defense of physically based models took on a pragmatic tone, with a much milder discourse compared to that articulated by Freeze & Harlan (1969). In this vein, Woolhiser (1996) [170] suggests that the development of models that realistically represent hydrological processes directly from physical theory may have been a great illusion of the scientific community, analogous to the search for "El Dorado". On the other hand, Simmons et al. (2020) [164] argue that the central spirit of Freeze & Harlan's project was to promote the *coupling* between the various compartments of the hydrological cycle, such as the atmosphere, soil, subsurface, and rivers—and not to obtain a supposedly true description of reality. Since most of the criticisms revolved around the representation of continuous fields (the ontology) and their philosophical consequences, the essence of the project remains alive and produces important *insights* by integrating various sciences, such as Hydrology, Climatology, and Ecology, into a modeling platform. For this reason, they emphasize that the term "physically-based" leads to a false interpretation of the ultimate purposes, with "integrated models" being a more appropriate designation. An undeniable pragmatic fact that contributes to this direction, brought by Fatichi et al. (2016) [171], is that complex problems often require complex solutions. That is, various practical applications need **distributed models**, which represent hydrological processes in sufficient detail in two- or three-dimensional space to aid in decision-making regarding water resource management, such as flood mapping and land use changes. Moreover, Clark et al. (2017) [172] argue that the philosophical problems of scale and uncertainty, while inevitable, are increasingly assimilated by integrated models, especially with **scaling techniques**, featuring nested parameterizations that can range from the finest mesh element of the local scale, through intermediate scales, to the global scale of the modeling domain.

3.5 Scale issues

3.5.1 Upscaling and downscaling

The relevance of scale in hydrological modeling took clear shape in the 1990s, particularly following the review by Blöschl & Sivapalan (1995) [173]. These authors present a comprehensive conceptual paper that transforms the scale problem, although inevitable, into something manageable through a structured approach. The starting point of their analysis is the science-management duality, that is, the distinction between **predictive models**, used to solve specific practical issues, and **exploratory models**, aimed at formalizing and articulating theories about the system hydrological. The practical problems of water resource management, the main targets of the application of hydrological models, vary substantially in terms of temporal scale, from hours, for flood alerts, to decades, for the impacts of land use changes. In this context, challenges related to scale in modeling arise when models are configured to operate predictively, with their parameters conditioned by empirical observations under specific time and space conditions, and then applied to produce predictions in different situations. A classic example already mentioned is the difficulty Horton faced when using point measurements of infiltration capacity to make predictions at the watershed scale in La Grange Brooke [100]. This difference in conditions necessitates an information transfer between scales, or **scaling**, which is often non-trivial, as highlighted in the discussion about physically based models. Thus, the concept of scale is defined by the attributes of time and space, which can be summarized as a **characteristic velocity**. The scale problem, therefore, presents itself as a problem of scaling, that is, in the difficulty of transferring information between different velocities.

Blöschl & Sivapalan (1995) also advance to establish the crucial notion that there are three scales to be understood and reconciled in a modeling exercise: the **natural scale**, the **observational scale**, and the **conceptual scale** (Figure 3.9). The natural scale refers to the actual characteristic velocity exhibited by hydrological processes. It can be classified in different ways: as the lifespan of intermittent **events**, such as rises; by the **period** of annual events, such as snowmelt or the arrival of wet seasons; or by the duration of **trends** in long-duration stochastic processes, which exhibit some degree of autocorrelation, such as the depletion or filling of aquifers (Figure 3.9**a**). The authors also expand this idea to spatial scales, defined by extent and trends in space, depending on the nature of the process. Some processes, like precipitation, do not have a preferred scale, as they distribute across multiple scales due to the nesting of subprocesses of both small and large scales, often with **spectral gaps** between them, that is, intervals where certain scales are less frequent. River runoff also follows this nested process structure, with flood peaks resulting from overlapping rapid response mechanisms and slower response mechanisms, such as groundwater or the occupation of large floodplains. Even rapid responses occur at different nested scales, such as flash rises from small soil plots and the formation of riparian wetlands or ephemeral springs, manifesting at the slope scale. The observational scale, in turn, consists of the scale occupied by empirical evidence, arising from the need to manage a finite number of samples. It has three main aspects: the **extent** or coverage of the dataset, the **resolution** or spacing between samples, and the interval of **integration** of the sample (Figure 3.9**b**). If sampling were infinite (or infinitesimal), the observational scale would coincide with the natural scale, capturing even the sample noise. In contrast, a very sparse sampling captures only the trend of the process, at best. A typical example of this is a rain gauge that, when read daily, reports the accumulated rainfall over a one-day interval, which may be much larger than the natural duration of a rain that occurred for only a few minutes. Nevertheless, this reading captures the trend of rainfall

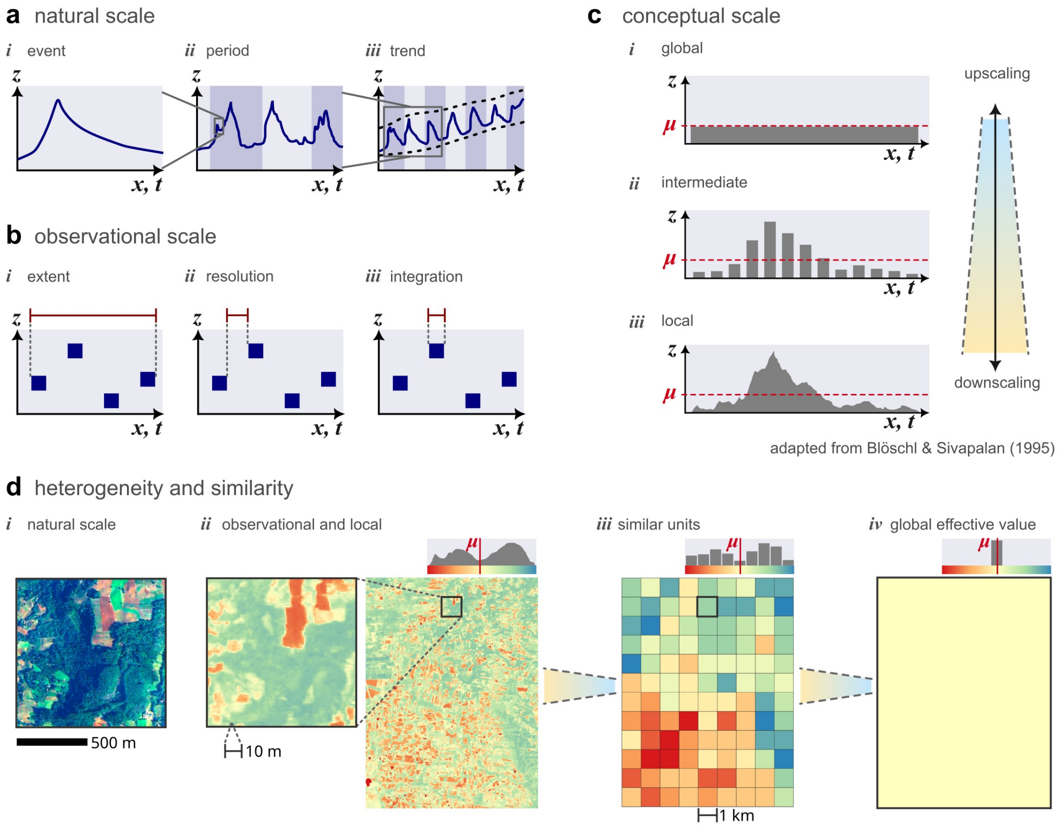

Figure 3.9 — Schematic representation of different scales. Organized system by Blöschl & Sivapalan (1995) [173] on the scales to be reconciled in time and space. **a** — The natural scale of processes varies in: (i) event scale; (ii) period scale, and; (iii) trend scale. **b** — The observational scale presents three aspects: (i) extent; (ii) resolution, and; (iii) integration. **c** — The conceptual scale mediates between the natural scale and observational through scaling methods. The model operates at nested scales: (i) global scale; (ii) intermediate scales, and; (iii) local scale. **d** — Effective values across scales: natural scale, where processes occur (i); the observational scale establishes the lower limit of the local conceptual scale (ii); intermediate scale of similar units (iii), and; effective process value at the global scale (iv).

on a weekly or monthly scale. On the other hand, detailed soil samples can provide information about extremely localized hydraulic conductivity, which does not reflect the actual effect of macropores and preferential pathways at the slope scale. In this case, the process scale is broader, making point samples incomparable. Ideally, observations should be compatible with the scale of the processes of interest, positioning the sampling at an optimal point between the noise range and the trend range.

The natural scale and the observational scale relate in hydrological modeling by being mediated by the conceptual scale, which is the scale of representation of the model itself (Figure 3.9**c**). Here lies the challenge of scaling, as the conceptual scale is often much larger or much smaller than the observational scale, introducing the inevitable commensurability error ε_{Δ}. As previously discussed, both the soil reservoir instantiated by Systems Dynamics and an element of computational grid in a physically-based model represent massive blocks that are incomparable with any point observation obtained in the field. Still, this error can be minimized by representing the target system simultaneously at scales close to the available observations. In practice, this means dividing the model into at least two nested levels: a more aggregated **global scale** and a more detailed **local scale**. Intermediate scales can also be incrementally instantiated, depending on the simulated hydrological processes and the available observations. At the global scale, for example, the highly aggregated processes of the watershed are represented, such as river flow at a river section and the final flow of evapotranspiration, accumulated results of various subprocesses at smaller scales. At the local scale, the details of these subprocesses are represented in small plots or mesh elements, such as rainfall input and runoff generation in different parts of the landscape. In this sense, the storage and flow variables, the parameters, and the input data need to be compatible at all levels, necessitating the transfer of information from one level to another, that is, they need to be scaled.

The solution to this situation, therefore, consists of defining a **scaling function** that is valid between the simulated levels. These functions perform both **upscaling**[14] of information (bottom-up transfer) and **downscaling**[15] of information (top-down transfer). For material levels and flows, which are conserved, the upscaling function can simply be the average or the sum over a given spatial or temporal extent. The global evapotranspiration flow of a watershed, therefore, would be the average of local flows (the integral). The downscaling, on the other hand, generally consists of a non-trivial process that strongly depends on the process in question. In the case of soil and rock properties, such as hydraulic conductivity, the only way to unpack this information is through maps that reveal its pattern or **spatial heterogeneity**. The same mapping strategy applies to parameters related to vegetation or land cover, such as interception capacity c_{max} and surface retention capacity s_{max}. In the absence of direct information, the use of **co-variables** or indicators can be employed with a downscaling function, or **distribution function**, which adds new auxiliary hypotheses to the theoretical framework of the model. For example, Collischonn *et al.* (2007) [174] assume the hypothesis that the local interception capacity c_{max} is directly proportional to the Leaf Area Index (LAI), that is: $c_{\text{max},i} = c \cdot \texttt{LAI}_i$, where c is a proportionality constant. On the other hand, co-variables can be applied to *group* spatial regions that theoretically exhibit **hydrological similarity**, that is, regions that are sufficiently homogeneous with respect to a given process at the assessed scale. In this context, the co-variable is referred to as the **hydrological similarity index**. The homogeneous regions resulting from this grouping, referred to as **hydrological response units**, significantly reduce computational cost as they execute block processing, as opposed to the mesh processing required at the local scale. Thus, models that apply this approach in downscaling are regarded as **semi-distributed models**, as they do not represent the local scale completely explicitly; the information is still compacted at the intermediate scale of the hydrological response units. Finally, another challenge at the local scale consists of the **regionalization** of values located at points or patches to their lateral neighborhoods, at the same scale, as in the case of point rainfall observations that are interpolated to represent a continuous field in space. Just like in the case of distribution function of parameters, this interpolation process introduces new auxiliary hypotheses (and their uncertainties) into the modeling process.

3.5.2 Scaling in `TOPMODEL`

An example of scaling that is worth presenting at this moment is the model `TOPMODEL`, initially articulated by Beven & Kirkby (1979) in a study of the Crimple Beck basin (England, 8 km^2) [175]. This model, instantiated in the paradigm of Systems Dynamics, despite exhibiting a relatively simple compartment structure, effectively represents the mechanism of variable source area, producing rapid hydrological responses both through flash rises and through saturation excess in wet areas. During the simulation, the model explicitly represents the expansion and contraction of riparian wetlands along the terrain's thalweg as the watershed receives more or less rain. The flow of effective rainfall p_{s} that directly impacts the saturated areas eventually becomes part of the rapid

[14] Translation of *upscaling* in English.
[15] Translation of *downscaling* in English.

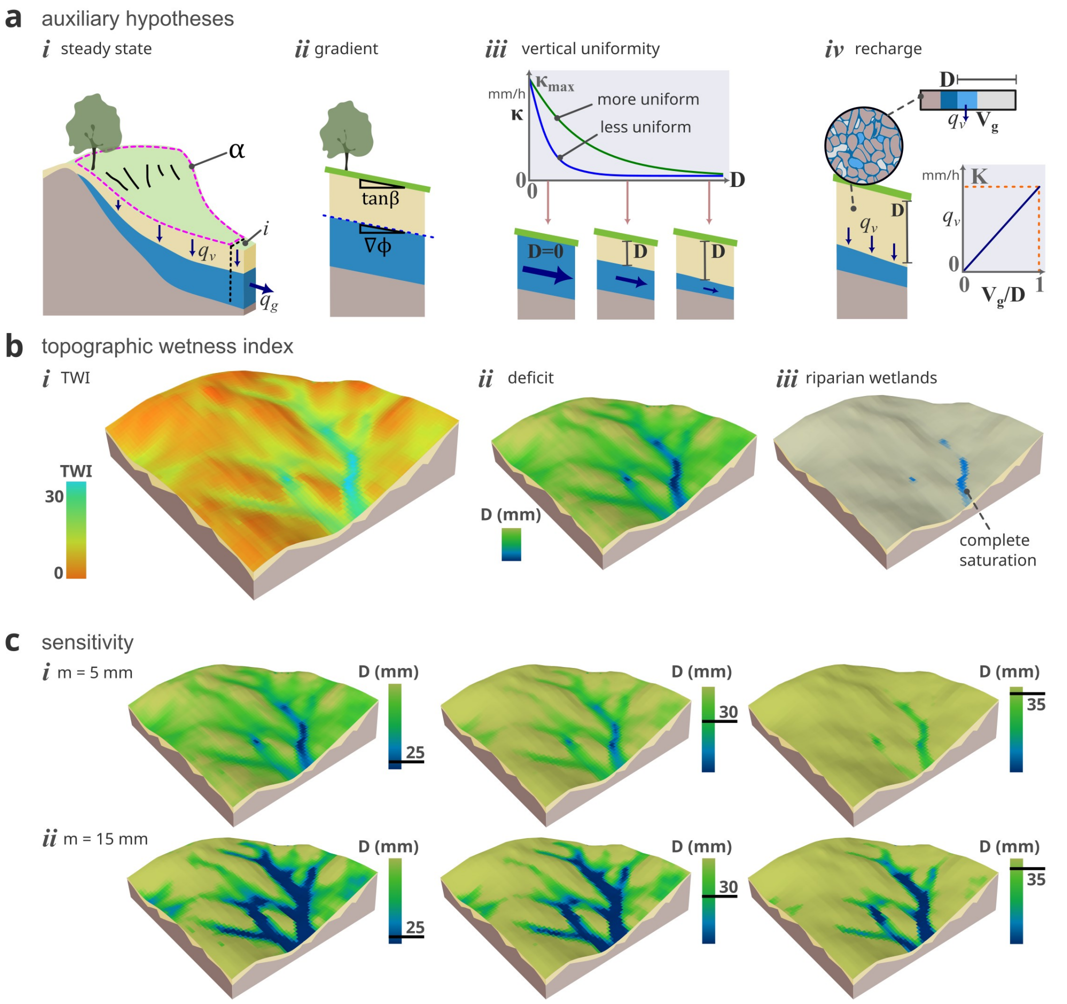

Figure 3.10 — Hypotheses and implications of `TOPMODEL`**.** The `TOPMODEL` is a model that performs downscaling of soil saturation from a topographic wetness index, the Topographic Wetness Index (`TWI`). **a** — The classic version of the `TWI` in the model is obtained by applying Darcy's Law with three auxiliary hypotheses: the hypothesis of steady-state (detail i); the hypothesis of shallow soils (detail ii), and; the hypothesis of vertical decay of transmissivity (detail iii). A fourth auxiliary hypothesis is the flow of recharge q_v as a linear function of the pressure in the vadose zone **V** (detail iv). **b** — The application of the scaling function of the model uses the distribution of the `TWI` (detail i) to determine the gravitational deficit **D** and the riparian wetlands at the local scale (detail ii and iii). **c** — The details i and ii compare the sensitivity of local deficit distribution for more or less uniform soils (parameter m). A basin with more uniform soil (m high) exhibits a more dispersed deficit distribution than a basin with less uniform soil (m low).

response of flood events, while the remaining portion of effective rainfall p_s falls on dry soil and can then infiltrate.

Unlike the high computational cost of physically based models, which need to numerically solve the Darcy-Richards Equation on a computational grid, the `TOPMODEL` approach identifies the spatial pattern of soil saturation at the local scale through a low-cost computational downscaling function (distribution). In the presence of empirical evidence solely from rainfall and discharge, both approaches are empirically equivalent, with the advantage that `TOPMODEL` is simpler (i.e., it has a greater degree of falsifiability). In this sense, the use of a physically-based model becomes justified only when more detailed evidence becomes available, such as piezometric levels, bedrock topography, and water quality parameters[16].

Soil saturation in `TOPMODEL` is expressed by the gravitational deficit $\mathbf{D}$ at the

[16] For detailed mappings of contamination plume evolution in the subsurface, for example, a physically-based model is the only alternative that offers the ontology compatible with the problem at hand.

local scale in the watershed, denoted by D_i [L], where i is any element of a mesh that divides the basin into N elements. That is, when $D_i = 0$, the soil is completely saturated in element i, and the water from the effective rainfall p_s that falls on this element cannot infiltrate, remaining accumulated on the surface until it reaches the surface retention capacity s_max. The compaction function of this variable transfers information across scales through the simple calculation of the average:

$$D = \frac{1}{N}\sum_{i}^{N} D_i \quad \forall\, i \in \{1, 2, ..., N\} \tag{3.7}$$

Where D [L] is the global deficit and; D_i [L] is the local deficit. Thus, the gravitational deficit $\mathbf{D}$ at the global scale consists of the average of the deficits at the local scale D_i. The distribution of gravitational deficit from global scale to local scale, on the other hand, is based on the use of a **saturation index**. This index is thus considered a *co-variable* of the gravitational deficit, such that the *deviations from the mean* between the local scale and global are linearly proportional:

$$D - D_i \propto \lambda_i - \lambda \quad \forall\, i \tag{3.8}$$

Where λ_i [$-$] is the saturation index at the local scale, and; λ [$-$] is the saturation index at the global scale, that is, the average obtained by:

$$\lambda = \frac{1}{N}\sum_{i}^{N} \lambda_i \quad \forall\, i \tag{3.9}$$

The Equation (3.8) becomes an equality when a proportionality constant is introduced:

$$D - D_i = \omega(\lambda_i - \lambda) \quad \forall\, i \tag{3.10}$$

Where ω [L] is the scaling factor. By rearranging the terms, the local gravitational deficit D_i is obtained through the following distribution function:

$$D_i = D + \omega(\lambda - \lambda_i) \quad \forall\, i \tag{3.11}$$

Where D_i must be truncated at zero, so as not to take negative values. In a hydrological model, Equation (3.11) aims to locally distribute the global deficit D at each time step, allowing other variables at the local scale to be specified, such as recharge q_v and surface runoff. In this case, the elements i where $D_i = 0$ correspond precisely to saturated soil areas, the variable source area that will invariably produce rapid runoff responses in the face of rainfall events.

Here, it is worth noting that Equation (3.11) is a generically applicable formulation for any saturation index λ_i. However, Beven & Kirkby (1979) originally deduced it theoretically from Darcy's Law and some auxiliary hypotheses (see Figure 3.10**a**), resulting in the saturation index λ_i then referred to as **Topographic Wetness Index (`TWI`)**, which is calculated by:

$$\mathrm{T}_i = \ln\left(\alpha_i / \tan\beta_i\right) \quad \forall\, i \tag{3.12}$$

Where α_i [$L^2 L^{-1}$] is the local drainage area per unit of contour, and; β_i[$-$] is the local slope of the terrain. That is, the local potential for soil saturation (1) is greater the larger the drainage area, and (2) is greater the smaller the slope of the terrain. Maps of `TWI` can thus be obtained directly from a **digital elevation model (`DEM`)** using geoprocessing techniques. The local slope β_i, for example, can be estimated using Horn's method (1981) [176] by computing the altitude differences in the west-east and north-south directions in the **mantissa**[17] of the mesh element. On the other hand, determining the local drainage area per unit of contour α_i requires a more computationally intensive analysis, as it is necessary to trace all upstream mesh elements for each given element. This cannot occur before removing spurious depressions in the `MDE`, which cause the method to truncate. Barnes *et al.* (2014) [177] introduce an efficient algorithm for this process, as well as review various other strategies available in the literature. Once a depression-free `MDE` is obtained, the drainage area is computed using flow accumulation methods, such as the unidirectional flow method by O'Callaghan & Mark (1984) [178] or the multidirectional flow method by Freeman (1991) [179]. Quinn *et al.* (1991) [180] demonstrate that there is substantial sensitivity in `TOPMODEL` regarding the choice of flow accumulation method, suggesting that the multidirectional method presents better empirical adequacy. Additionally, the authors also evaluate the possibility of *overlapping* methods to adjust the `TWI` between ephemeral (multidirectional) and perennial (unidirectional) drainage regions.

There are three auxiliary hypotheses that theoretically underpin Equation (3.12). The first is the hypothesis of steady-state (detail i in Figure 3.10**a**), which establishes that a local steady-state condition is achieved at each time step, such that the lateral base flow equals the recharge flow:

$$q_{\mathrm{g},i} = q_{\mathrm{v}} \cdot \alpha_i \quad \forall i \tag{3.13}$$

Where $q_{\mathrm{g},i}$ $[L^2T^{-1}]$ is the lateral base flow per unit of contour; q_{v} $[LT^{-1}]$ is the recharge flow, and; α_i $[L^2L^{-1}]$ is the local drainage area per unit of contour. The second auxiliary hypothesis (detail ii in Figure 3.10**a**) assumes that the soil is shallow enough for the local hydraulic gradient in the water table $\nabla\Phi_i$ $[LL^{-1}]$ to be approximated by the local slope of the terrain $\tan\beta_i$ $[LL^{-1}]$:

$$\nabla\Phi_i = \tan\beta_i \quad \forall i \tag{3.14}$$

The third auxiliary hypothesis (detail iii in Figure 3.10**a**) is that local hydraulic conductivity K_i $[LT^{-1}]$ decays exponentially with gravitational deficit, meaning that the drier the soil, the lower the hydraulic conductivity. This is a hypothesis consistent with empirical observations that the upper soil horizons, with organic layers and macropores, exhibit higher conductivity than the lower, more mineral parts. The hydraulic conductivity per unit of contour is expressed as **hydraulic transmissivity**, leading to the hypothesis taking the following form:

$$\kappa_i = \kappa_{\max} \cdot e^{-D_i/m} \quad \forall i \tag{3.15}$$

Where κ_i $[L^2T^{-1}]$ is the local transmissivity; $\kappa_{\max}$ $[L^2T^{-1}]$ is the maximum transmissivity under saturated conditions; D_i $[L]$ is the local deficit, and; m $[L]$ is the constant of **vertical uniformity of the soil**. The larger the value of m, the more gradual the change in transmissivity as a function of saturation, such that: $\lim_{m\to\infty} T = T_{\max}$. Considering the flows per unit of contour of the terrain, Darcy's Equation (3.6) takes the following structure:

$$u = K\nabla\Phi \Rightarrow q_{\mathrm{g,i}} = \kappa_{\max}\nabla\Phi_i \quad \forall i \tag{3.16}$$

Where the Darcy velocity u $[LT^{-1}]$ corresponds to the lateral base flow per unit of contour $q_{\mathrm{g},i}$ $[L^2T^{-1}]$. Connecting Equations (3.13), (3.14), and (3.15) in Darcy's Equation (3.16):

$$q_{\mathrm{v}}\alpha_i = \kappa_{\max} e^{-D_i/m} \tan\beta \quad \forall i \tag{3.17}$$

[17] In a rectangular mesh of elements, the mantissa consists of the eight neighboring elements surrounding any given element.

The local deficit D_i can be isolated, yielding:

$$D_i = -m \ln\left(q_\text{v}\alpha_i/\kappa_\text{max} \tan\beta_i\right) \quad \forall i \tag{3.18}$$

By logarithmic properties, the following relationship is also obtained, isolating the terms of static hydrological variables from dynamic hydrological variables and purely topographic terms:

$$\ln(q_\text{v}/\kappa_\text{max}) = -D_i/m - \ln(\alpha_i/\tan\beta_i) \quad \forall i \tag{3.19}$$

Now, considering that the global deficit D is the average of the local deficits D_i, Equation (3.18) can be applied in Equation (3.7):

$$D = \frac{1}{N}\sum_i^N -m \ln\left(q_\text{v}\alpha_i/\kappa_\text{max} \tan\beta_i\right) \quad \forall\ i \tag{3.20}$$

Using summation properties and assuming m and κ_max are spatially homogeneous, local terms can be isolated from global terms:

$$D = \left[-m\frac{1}{N}\sum_i^N \ln\left(\alpha_i/\tan\beta_i\right)\right] - \left[m \ln\left(q_\text{v}/\kappa_\text{max}\right)\right] \quad \forall\ i \tag{3.21}$$

Substituting (3.19) into (3.21), we arrive at:

$$D = \left[-m\frac{1}{N}\sum_i^N \ln\left(\alpha_i/\tan\beta_i\right)\right] + D_i + m\ln\left(\alpha_i/\tan\beta_i\right) \quad \forall\ i \tag{3.22}$$

Which is homologous to Equation (3.10), with $\lambda_i = \text{T}_i$ and $\omega = m$:

$$D_i = D + m\left[\left(\frac{1}{N}\sum_i^N \text{T}_i\right) - \text{T}_i\right] \quad \forall\ i \tag{3.23}$$

In total, the version of `TOPMODEL` articulated by Beven & Kirkby (1979) includes seven parameters regulating reservoirs and flows of the water balance in the soil, as well as a flow velocity parameter used in simulating the propagation of flow in the drainage network of channels. In particular, the parameters m and $Q_\text{g,max}$ can be estimated *a priori* from the recession curve of the river during observed recessions in cold weather (with low water loss due to evapotranspiration), as the integration of Equation (3.15) over all lateral stretches of channels determines the base flow:

$$Q_{\text{g},t} = Q_\text{g,max} \cdot e^{-D_t/m} \quad \forall\ t \tag{3.24}$$

Where Q_g $[L^3T^{-1}]$ is the base flow; $Q_\text{g,max}$ $[L^3T^{-1}]$ is the production capacity of the aquifer; D_t [L] is the global deficit at time t, and; m $[L]$ is the vertical uniformity of the soil. The details in Figure 3.10**b** demonstrate how sensitive the model is to changes in the vertical uniformity of the soil. The distribution of local deficit, in this regard, becomes increasingly more dispersed as the value of m increases. This occurs evidently because it acts as a multiplier in the scaling function (Equation (3.11)). The practical implication is that a basin with relatively more uniform soil will produce relatively more riparian wetlands. The physical interpretation of this implication is that, keeping the same production capacity of the aquifer $Q_\text{g,max}$, a more uniform soil transmits more water, draining the higher slopes more quickly as the global deficit of the basin increases.

The hydraulic conductivity K of the soil is employed in `TOPMODEL` in a revision of the model presented by Beven & Wood (1983) [181], under a fourth auxiliary hypotheses concerning the flow of recharge q_v (detail *iv* in Figure 3.10**a**). In this case, the authors assume that the vertical flow of recharge q_v at the local scale tends linearly toward the value of hydraulic conductivity K as the vadose zone $\mathbf{V}$ becomes pressurized by hydraulic load[18]:

$$q_{\mathrm{v},i} = K \cdot \frac{V_{\mathrm{g},i}}{D_i} \quad \forall\, i \tag{3.25}$$

Where $q_{\mathrm{v},i}$ $[LT^{-1}]$ is the local recharge; K $[LT^{-1}]$ is the hydraulic conductivity of the soil; $V_{\mathrm{g},i}$ $[L]$ is the local gravitational water in the vadose zone $\mathbf{V}$, and; D_i [L] is the local deficit. That is, the pressurization in the vadose zone $\mathbf{V}$ is represented by the ratio of gravitational water to the saturation deficit, which can also be interpreted as the *capacity* for storing gravitational water in the vadose zone $\mathbf{V}$. Thus, as the gravitational water in the vadose zone $\mathbf{V}$ is constrained by the deficit, the ratio $V_{\mathrm{g},i}/D_i$ tends to 1 when the deficit approaches zero, or $\lim_{D_i \to 0} q_v = K$.

The classic version of `TOPMODEL` is essentially summarized by the concepts and equations organized above, with its main hypothesis being the downscaling function represented by Equation (3.11). Its fundamental hallmark, therefore, is to represent rises produced both by overland runoff and by saturation excess, which is accomplished with simulated maps of riparian wetlands obtained from a topographic indicator (in this case, the `TWI`). On the other hand, the other flow equations, such as the evapotranspiration flow in each reservoir of the system and hydraulic propagation downstream, are couplings to the basic model that can be modified or not. Authors like Ambroise *et al.* (1996) [182] and Iorgulescu & Musy (1997) [183] have indeed implemented generalizations in the auxiliary hypotheses, deriving generic formulations to calculate the topographic wetness index. In the same vein, Beven & Freer (2001) [184] produced a more complex version of the model, called Dynamic `TOPMODEL`, where the local drainage area α_i is replaced by the local recharge area $\alpha_{v,i}$, which needs to be updated at each time step of the simulation (a procedure that makes this version computationally more intensive).

Besides the capacity for modifications, Beven [58] suggests that the classic version can be instantiated as a semi-distributed model, aggregating the mean of the saturation index into discrete ranges of its histogram in hydrological response units. That is, in small incremental ranges of the index, it is assumed that the mesh elements exhibit hydrological similarity. This way, relatively broad regions of elements in space are scaled to a relatively small number of homogeneous blocks, drastically reducing the computational cost of simulating the model. More detailed maps at the local scale can thus be recovered after processing—one just needs to use the source map of the saturation index to position the simulated processes in their respective mesh elements. Although it is a practically oriented strategy related to the procedural model, using a semi-distributed model has implications for the simulated results. For instance, mesh processing in a fully distributed approach allows for the representation of spatial and temporal spatial heterogeneity of other flows and parameters (such as rainfall distribution, for example), which is not possible in a semi-distributed approach. Since the intensive use of simulations in diagnostic techniques requires that the simulation time of the models not be a critical bottleneck (see Section 2.6), these issues must be weighed to arrive at an appropriate strategy for addressing the dimensionality problem.

[18] In fact, the authors introduced a peculiar term of *delay per unit deficit* t_d $[TL^{-1}]$, causing the recharge equation to take the following form: $q_v = V_g / t_d D$. Although identical, this notation does not make much hydrological sense, especially considering that hydraulic conductivity is a well-established concept. In observance of John Sterman's modeling principles (Chapter 2), I maintained a clearer notation.

3.5.3 Scaling in PLANS

The model PLANS is a version of TOPMODEL that exhibits strategies for both the generalization of the saturation index and semi-distributed modeling. Illustrated in Figure 3.11, the model was developed by myself and colleagues with the explicit purpose of establishing a tool to assist in the formulation of evidence-based policies in the context of expanding Nature-based solutions (NBS)[19] in watersheds in Brazil [3], [4]. The term "Nature-based solutions" serves as a conceptual umbrella for a collection of techniques and approaches at different scales that draw inspiration from or utilize natural processes. We will see later, in the next chapter, that this public policy movement can benefit from the use of modeling within a framework of basic principles. The first version of the model PLANS was a somewhat more complex model than the prototype presented in the previous chapter. This initial version was used in an exploratory modeling study that applied search techniques to optimally allocate the expansion of NBS over time under future scenarios [3]. On one hand, the application of the model successfully highlighted nuances and stimulated revisions of mental models. In this case, the model found that the expansion of the listed NBS yields scaling benefits in relatively more degraded watersheds—the incremental performance in more preserved areas likely does not justify the investment. On the other hand, the highly aggregated nature of this initial version did not allow for an exact assessment of *where* the expansion of NBS should occur, causing the model to fail in addressing the spatial allocation problem. This shortcoming forced me to abandon the initial structure and instead instantiate a version of TOPMODEL tailored to address the issue of expanding NBS both in time and space [4].

In the model PLANS, as in TOPMODEL, the local distribution of water deficit in the soil is done through the downscaling function defined in Equation 3.11, using a topographic wetness index λ_i as a co-variable. However, unlike TOPMODEL, the only hypothesis fundamentally defended by the model in this aspect is the linear relationship that the scaling function implies, which allows testing other topographic indices beyond the TWI (Figure 3.11**b**). This relaxation of the hypotheses of the classic model was primarily motivated by the practical need to address the problem of expanding NBS in Brazil, a country with a wide heterogeneity of soils and landscapes, including tropical soils that are much deeper than those observed in temperate or subtropical climates. Another motivation, of a conceptual nature, is grounded in the empirical observations presented by Crave & Gascuel-Odoux (1997) [185] regarding the distribution of saturated areas in a small watershed in France (1.3 km^2), with soils varying from 40 cm to 2 meters in depth. In the study, the authors report a weak correlation between local soil saturation and the TWI, as well as the relative immobility of the saturation patch at the valley bottom, largely refuting the underlying theory of TOPMODEL. On the other hand, they show that the soil saturation at sampled points i has an inverse relationship with the *altitude difference* $\Delta Z_{i,o}$, or height $\mathrm{H}_{i,o}$, relative to the nearest water outcrop o. This inverse relationship holds up to a certain height threshold $\mathrm{H}_{\max}$, beyond which saturation exhibits a relatively uniform range of values. Given these observations, the authors suggest that in basins with relatively deep and well-drained soils, the landscape divides into two parts: the drier headwater region and the wetter riparian region (Figure 3.11**b**, detail i). Keeping other variables constant, the separation between these two regions is reasonably delineated by the height threshold $\mathrm{H}_{\max}$ above the valley bottom.

This same topographic wetness index, referred to as **Height Above Nearest**

[19] The acronym PLANS stands for *Planning Nature-based Solutions*, meaning "Planning Nature-Based Solutions". The overarching goal of the initiative is to establish a toolkit of concepts and tools to address the challenges of expanding NBS. The model presented here might be designated as the hydrological module of the PLANS project.

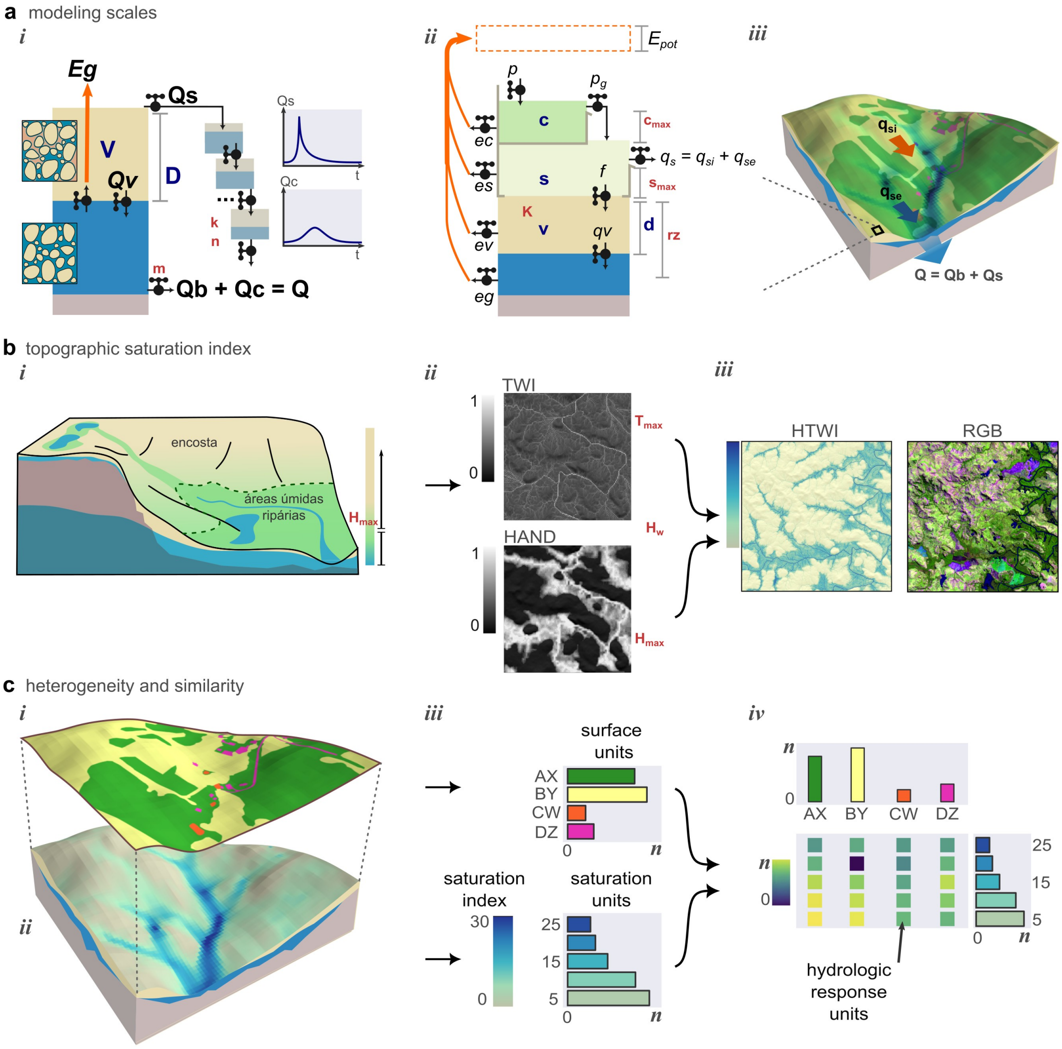

Figure 3.11 — **The model** `PLANS`. The model is a tailored version of `TOPMODEL` designed to address the problem of expanding Nature-Based Solutions. **a** — Modeling scales of the model: global scale (detail i); hydrological response units scale (detail ii), and; local scale, at the mesh element (detail iii). **b** — Topographic wetness index HTWI. The saturation index is based on the hypothesis of dual separation between slope areas and riparian areas (detail i). The TWI and HAND indices are normalized by fuzzy logic (detail ii). A weighting between the indices generates the HTWI (detail iii). **c** — Heterogeneity and spatial similarity: the surface layer of static variables is separated from the subsurface layer of dynamic variables (details i and ii); the variables at the local scale are grouped into surface units and saturation units (detail iii); a two-dimensional histogram, or frequency matrix, is computed to store the hydrological response units (detail iv).

Drainage (`HAND`)[20], was articulated ten years later by Rennó *et al.* (2008) [186], who demonstrated its effectiveness in mapping wet areas in the Amazon. However, the differentiating factor of the study by Rennó *et al.* (2008) is that the authors organized the computational method to obtain the `HAND` by applying geoprocessing techniques to `MDE`. In general terms, the technique involves initially establishing a map of the drainage network, which can be done using a drainage initiation threshold H_α [L^2], representing the minimum area for water to outcrop in the soil. Thus, for each mesh element o in the drainage network, their respective altitudes $Z_{o,i}$ and drainage areas (basins) are obtained. Finally, the local `HAND` H_i in each basin area is calculated by the difference between the local altitude Z_i and the altitude of the nearest drainage $Z_{o,i}$. The result is, therefore, a normalized Digital Elevation Model such that the zero altitude is always the level of the nearest river, stream, or valley bottom. The derivation of `HAND` through geoprocessing has led to new applications, such as mapping flood risk of major rivers,

[20]The acronym HAND stands for "Height Above the Nearest Drainage".

since the higher one is above a river channel, the greater the safety [187]. It becomes evident, however, that the value of HAND is highly sensitive to the initially established drainage map or area threshold H_α, making applications for mapping floods of major rivers (macro-drainage) very different from mapping soil saturation (micro-drainage).

Thus, the approach adopted in the model PLANS encourages that the topographic wetness index λ_i be obtained through fuzzy logic combinations between the TWI and the HAND, resulting in the index termed *HAND-enhanced TWI*, or the TWI enhanced by the HAND (HTWI), illustrated in detail *iii* of Figure 3.11**b**. This approach was actually suggested tangentially by Quinn *et al.* (1991) [180] to differentiate between ephemeral and perennial drainage regions, but concerning different flow accumulation methods. The proposed index, in this line, retains the characteristic of the TWI in increasing the saturation of the landscape from upstream to downstream, but makes this effect relatively more pronounced near the riparian wetlands than on drier slopes. The map is initially calculated by the fuzzy normalization of both variables, requiring the establishment of upper thresholds for each (Figure 3.11**b**, detail *ii*). For the TWI, the normalization is ascending:

$$\tilde{\mathrm{T}}_i = \mathtt{MIN}(\mathrm{T}_i/\mathrm{T}_{\max},\ 1) \quad \forall\, i \tag{3.26}$$

Where T_i $[-]$ is the local TWI; $\mathrm{T}_{\max}$ $[-]$ is the upper threshold of the TWI, and; $\tilde{\mathrm{T}}_i \in \{0,1\}$ $[-]$ is the normalized local TWI. In the case of the HAND, the normalization is descending:

$$\tilde{\mathrm{H}}_i = \mathtt{MAX}(1 - \mathrm{H}_i/\mathrm{H}_{\max},\ 0) \quad \forall\, i \tag{3.27}$$

Where H_i $[-]$ is the local HAND; $\mathrm{H}_{\max}$ $[-]$ is the upper threshold of the HAND, and; $\tilde{\mathrm{H}}_i \in \{0,1\}$ $[-]$ is the normalized local HAND. Thus, the HTWI is determined by the weighted average between these normalized variables and scaled back to the original range of the TWI:

$$\mathrm{HT}_i = \mathrm{T}_{\max}\frac{\tilde{\mathrm{T}}_i + H_w\tilde{\mathrm{H}}_i}{1 + H_w} \quad \forall\, i \tag{3.28}$$

Where $\mathrm{HT}_i \in \{0, \mathrm{T}_{\max}\}$ $[-]$ is the HTWI, and; H_w $[-]$ is a positive dimensionless factor or weight reflecting the dominance of the HAND over the TWI. The theory embedded in the derivation of HTWI is that there exists a spectrum of hydrological landscapes that extends from the total prevalence of shallow soils and dynamic saturated areas (total dominance of TWI over HAND) to the total prevalence of deep soils and static saturated areas (total dominance of HAND over TWI), with intermediate alternatives between these extreme situations. The dominance of one over the other is regulated by the weight H_w, while the mobility of saturated areas is regulated by the thresholds $\mathrm{T}_{\max}$ and $\mathrm{H}_{\max}$. In the special case where $H_w = 0$, the HTWI is identical to the TWI truncated at $\mathrm{T}_{\max}$. This generalized topographic wetness index, adjustable for any landscape, introduces three additional parameters into the model[21], an epistemological cost that the authors deemed acceptable given the practical need to model hydrological processes in the diverse environments of Brazil. One way to reduce uncertainties in the subsequent distribution of parameters might be to pre-condition the *prior* distribution with other spatial variables obtained from remote sensing, such as moisture indices, surface temperature, or simply shortwave infrared reflectance, as illustrated in Figure 3.6**a**.

Another difference of the model PLANS compared to TOPMODEL is the representation of surface heterogeneity (Figure 3.11**c**). The classic version of TOPMODEL was developed assuming that the surface is homogeneous, which makes sense in the Crimple Beck watershed in England, a rural area dominated by pastures for livestock. However,

[21] In fact, when considering the drainage area threshold H_α for defining the HAND, there are four additional parameters.

a model designed to address the problem of expanding Nature-Based Solutions must not only represent the soil and its cover in heterogeneous situations but also enable the simulation of alternative cover scenarios to assess the positive or negative impact of a given expansion policy. For instance, the model should be able to inform whether there is a difference in the behavior of the hydrological system between reforestation in different parts of the landscape. This requirement is recognized by Gao *et al.* (2015) [188], which motivated them to implement a distributed version of `TOPMODEL` to evaluate hydrological impacts of land use and cover change. The model `PLANS`, on the other hand, utilizes a semi-distributed approach, collapsing the local scale into an intermediate hydrological response units scale (Figure 3.11**a**, detail *ii*). This process is accomplished through the cross-tabulation of surface units (similar patches of soil and vegetation) with saturation units (regular intervals of the topographic wetness index)[22] (Figure 3.11**c**, details *i*, *ii*, and *iii*). This tabulation ultimately results in a two-dimensional histogram, or **frequency matrix**, which specifies the area prevalence of each hydrological response unit in the spatial region of interest[23]. As illustrated in detail *iv* of Figure 3.11**c**, the columns represent the histogram of the saturation index in each surface unit. The rows of the matrix, on the other hand, are similar in terms of saturation, which readily facilitates the determination of local deficit through the downscaling function of the model.

3.6 The paradigm of connectivity

In this chapter, I organized the current twists in Hydrology since the International Hydrological Decade, that is, from the 1960s onward. On one hand, in the experimental front, the Infiltration Age faced its ultimate crisis with the rise of a new paradigm that reaffirms the differentiation and uniqueness of hydrological response mechanisms as functions of climate, topography, soils, vegetation, etc. On the other hand, in the realm of modeling, the advent of digital computers paved the way for ontologically diverse methods, such as Systems Dynamics, vector fields, and statistical models, with the first two based on the theoretical description of processes and the last solely based on empirically obtained data.

As new scientific paradigms do not establish themselves by being perfect, but by being better than the competition, both approaches that have settled in the field are not free from problems. In the case of experimental research, the paradigm of differentiation encountered empirical paradoxes involving the rapid mobilization of old water and the diversity of geochemical signatures, which are difficult to explain by the mechanisms articulated by Dunne's systematization (1983) [99]. Furthermore, the experimental research program is essentially a catalog of the hydrological response mechanisms of each unique watershed. Even though the catalog can be detailed and endlessly expanded, this operational mode does not contribute to a unifying scientific theory [78]. In the case of modeling, crises have affected the various approaches, giving rise to two main inescapable epistemological problems in hydrological modeling: the equifinality problem and the scale problem [190]. Although different, these problems are interconnected, resulting in inexorable epistemic uncertainties that hang over model results. The im-

[22] Although the soil and vegetation maps are maintained as input data for the model, it is important to note that these maps are also the result of a scaling process of other local variables. For example, the vegetation and land use classes provided in the annual maps published by Souza *et al.* (2020) [189], used in the model `PLANS`, were derived from groupings of spectral reflectance and band indices from orbital scene data using machine learning.

[23] The propagation of flow to a given river section, therefore, must be conducted by specifying the prevalence of each response unit in the area of interest basin, which may or may not approximate the prevalence of the total region. This approach, thus, allows for evaluating the final flow in multiple basins of interest.

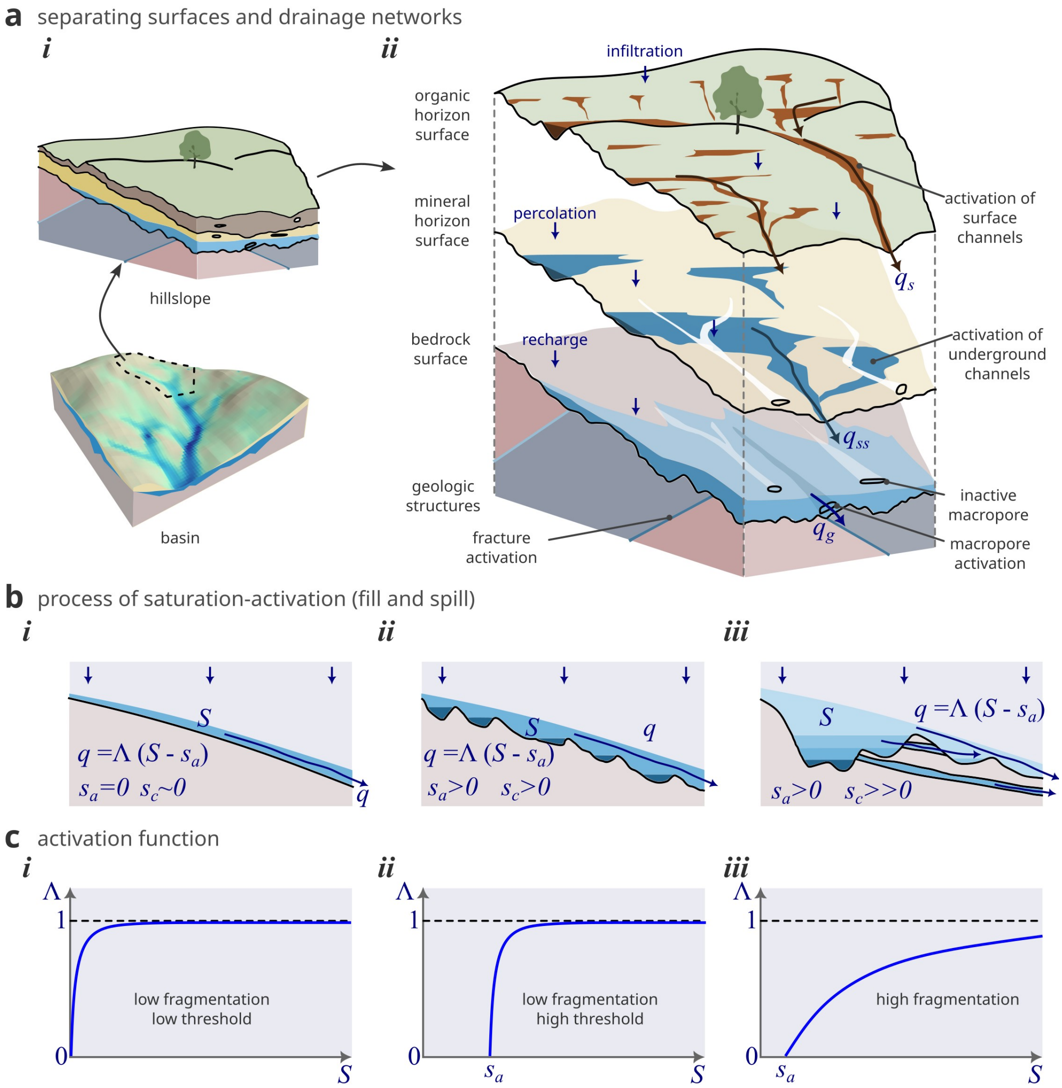

Figure 3.12 — The connectivity paradigm. The Connectivity Theory proposed by Jeffrey McDonnell and colleagues presents a unifying and revolutionary potential in Hydrology. **a** — The perceptual model is based on the principle that all hydrological responses are consequences of the same phenomenon. The vertical permeability transitions create separating surface. Thus, the network of channels on these surfaces can eventually become saturated and activated. Interactions also occur from bottom to top when a layer becomes saturated enough to interfere with the percolation process of the layer immediately above. **b** — The saturation-activation process occurs due to the topology of the channel network that drains the separating surface at various nested scales. The surface can be fully connected (detail i) or require an initial activation level s_a (detail ii). A surface with high fragmentation s_c offers multiple incrementally activated drainage networks, attenuating the response signal (detail iii). **c** — The activation function Λ can be modeled in terms of a saturation process that approaches the potential maximum flux $S - s_a$ as the surface saturates, that is, $\Lambda \to 1$.

perative of these epistemic problems, while overshadowing the vector field approach (forcing its proponents to appeal to pragmatic advantages), has shed new light on the ontology of Systems Dynamics, making it useful for estimating uncertainties through the application of low-cost computational semi-distributed models.

That said, I will conclude the chapter by articulating the recent revolutionary ideas of Jeffrey McDonnell and colleagues, who have sought over the past two decades to pave the way for a new unifying paradigm in Hydrology. These ideas have a direct impact on the application of hydrological models in the context of zero-order basins, making their assimilation and articulation in future versions of the model `PLANS` of utmost importance. Although widely published, McDonnell's synthesis can be traced through three articles separated by intervals of approximately ten years: McDonnell (2003) [78]; McDonnell (2013) [104], and; McDonnell *et al.* (2021) [87].

Initially, McDonnell (2003) demonstrates the increasing disconnect between hydrological models and evidence, highlighting that the mechanisms from the International Hydrological Decade are based on assumptions that fail to explain the old water paradox. Essentially, he argues that evidence points to both a greater separation between well-drained slopes and riparian wetlands, as well as a greater influence of bedrock topography on the outcropping of groundwater. Together, these two factors produce rapid translational flow of pre-event (old) water with diverse geochemical signatures. In the field of modeling, McDonnell (2003) asserts that the appropriate ontology for the new challenges is Systems Dynamics, not merely for convenience, but because it allows for the representation of the different compartments of the zero-order basin, primarily slopes and riparian zones, and guarantees exploratory experiments at a low computational cost. As highlighted in the chapter's epigraph, Systems Dynamics reaffirms itself as an ontological paradigm that is both intuitive and objective for understanding and learning about environmental systems.

The second step by McDonnell (2013) is to propose the **Connectivity Theory** as the definitive explanation for both the hydrological processes systematized by Dunne (1983) and the (not so) recent findings regarding the old water paradox. Initially reported in Meerveld & McDonnell (2006) [191] to explain processes in a specific experimental basin, McDonnell generalizes his concepts, treating it as a revolutionary and unifying theory that seeks to resolve the crisis established in the field. McDonnell proposes this theory from a question that sounds like heresy: could the rapid and slow responses of basins all be *the same phenomenon*?

> (...) the simple premise that all runoff processes are the same opens up new avenues to explore: Is there common emergent behaviour across all runoff types? – Jeffrey McDonnell (2013, p. 4110) [104].

This provocative question leads him to demonstrate that any system formed by networks of small channels can be modeled by the **Percolation Theory**, a mathematical branch of Network Theory[24] [192]. According to this theory, flow occurs through a network as long as there are connections between the nodes. In the case of a zero-order basin, in the physical world, the different parts of the system each function as a network of small reservoirs connected by small channels (open or closed, macroscopic or microscopic). Here, the relevance of two key concepts from the theory arises: (1) the **separating surface** between compartments, created by the transition of vertical permeability between soil horizons, and; (2) the **activation threshold** of the compartment, which primarily arises from the heterogeneity of the separating surface.

A clear example of this is the generation of flash floods when the soil's infiltration capacity is insufficient. Typically, the level of surface depressions needs to reach an activation threshold, at which point some surface depressions connect for the first time and start to spill water downhill. With more rain, the puddles continue to fill until a point is reached where all the incoming rainwater can flow downhill. The speed at which the connection occurs depends on the heterogeneity of the soil surface: a smooth surface is much more connected than a rough surface.

Although it may seem obvious, McDonnell suggests that the **saturation-activation process**[25] that generates flash floods occurs in all subsurface layers where vertical water flow encounters a transition in permeability, including the relatively impermeable bedrock layer (Figure 3.12**a**). The only difference in the subsurface environment is that the network of channels is composed of the micro and macropores of the

[24] Unlike scientific theories, which require empirical evidence to be corroborated, mathematical theories are based on axioms and deductive inference.

[25] A loose translation of the English term *fill and spill*.

immediately overlying horizon. In this context, the theory even allows for relatively small amounts of event water (new water) to activate the connection between pockets of stored water from before the event (old water). The pressurization in the saturated zone, along with the formation of natural siphons in the subsurface, eventually expels *more* water than what entered, generating a negative mass balance on the slope and a hysteretic behavior of flow pulses. Finally, when a separating surface becomes sufficiently saturated, it propagates this effect from bottom to top, causing saturation of the upper layer, thereby creating conditions analogous to the riparian wetlands we observe on the surface.

Finally, in their third move, McDonnell *et al.* (2021) articulate how to approach the Connectivity Theory in the context of modeling, considering the scale problem. Again, Systems Dynamics is presented as the appropriate ontology for representing the target system, emphasizing the explicit definition of the scale of interest, identifying the saturation-activation processes that manifest at the chosen conceptual scale. The authors' hypothesis is that saturation-activation processes occur at all scales; however, the signals emitted by smaller scales are progressively masked by saturation-activation at larger scales (Figure 3.12**b**). For example, while at the scale of zero-order basins the saturation-activation process is primarily dictated by topography, soil, and vegetation, this signal fades at the scale of higher-order basins, with the saturation-activation effects of the river drainage system and flooding of the plains becoming more dominant. Thus, experimental and modeling research must explicitly question what scale is being addressed and which critical saturation-activation processes are necessary to understand the target system. Relatively simple (yet objective) Systems Dynamics models can capture this knowledge, formalizing the main hypothesis of the **activation function**, which takes the following general form:

$$Q_a = \begin{cases} 0 & \text{if} \quad S \leq s_a \\ \Lambda \cdot (S - s_a) & \text{if} \quad S > s_a \end{cases} \tag{3.29}$$

Where Q_a [LT^{-1}] is the activation flow of the reservoir with level S [L]; s_a [L] is the **activation level** of the reservoir, and; Λ [T^{-1}] is the **activation function** of the reservoir, to be defined based on the auxiliary hypotheses of the model. In the previous chapter, during the development of the hydrological model prototype, I applied these exact principles for the quick response flow R of the surface reservoir S_1 (see Section 2.5). Equation (2.9) has exactly the same structure as Equation (3.29), with $\Lambda = c$, a runoff coefficient defined between 0 and 1 obtained from a activation function with the following structure (Figure 3.12**c**):

$$\Lambda = \frac{(S - s_a)}{(S - s_a) + s_c} \frac{1}{\Delta t} \tag{3.30}$$

Where s_a [L] is the activation level of the reservoir, and; s_c [L] is the **fragmentation level** of the reservoir. This function, when coupled into (3.29), implies that the outflow Q_a caused by activation asymptotically approaches the potential maximum flux $(S - s_a)$ as the reservoir level S increases, because higher levels increasingly activate the drainage network. In other words, $\lim_{S \to \infty} \Lambda = 1$ in (3.30) and $\lim_{S \to \infty} Q_a = (S - s_a)/\Delta t$ in (3.29). The fragmentation level s_c is a parameter that plays the role of regulating the speed of this process, serving as a measure of inverse connectivity (the higher, the less connected the reservoir is). The physical interpretation of the fragmentation level is the level S necessary to reach half of the potential maximum flux. The **Michaelis-Menten equation** [193], which describes an enzyme saturation process, coincidentally has an identical structure, making it a notable homology. Another identical structure is the equation from the **CN method**, empirically proposed based on the results of

Mockus (1949) [98]. The difference in this case is that the creators of the CN method express the level S in terms of accumulated rainfall P, which is only valid for a surface reservoir with unlimited capacity. The fragmentation level, in this regard, is expressed in terms of a dimensionless connectivity coefficient, such that $s_c = (1000/\mathrm{CN}) - 10$. Such homologies and theoretical explanations of old empirical adjustments outline the contours of a legitimately revolutionary scientific theory. ■

3.7 Chapter summary

In this chapter, I provided a historical and theoretical overview of the evolution of hydrological models, highlighting the main changes in approaches and scientific paradigms. Beginning with slopes, where hydrological processes start, I explored the transition from Horton's infiltration model to more complex concepts that incorporate climatic, geomorphological, and biological variability. I also analyzed the limitations faced by modern computational models, such as scale and equifinality problems, culminating in the Connectivity Theory, which proposes an innovative integration of surface and subsurface flow processes.

- **Slopes are where it all begins.** Rapid hydrological responses (floods) and slow responses (recessions) begin with rain on the slopes or zero-order basins. Simplifying this complexity can lead to inadequate models, especially in the context of watershed revitalization. Therefore, it is crucial to recognize the theories regarding runoff generation at this scale.

- **The Age of Infiltration.** During the mid-20th century, the hegemony of Horton's hydrological model was established, explaining hydrological responses through the soil's infiltration capacity, separating rainfall into runoff and recharge. Although surpassed, this paradigm elevated Hydrology from its empirical phase to a geoscience.

- **The Age of Differentiation.** With the International Hydrological Decade in the 1960s, new evidence and theories emerged that refuted Horton. This new paradigm explores how different hydrological responses arise due to climate, topography, soils, and vegetation. In addition to floods, the roles of macropores and riparian wetlands were highlighted. However, a crisis emerged with the old water paradox.

- **Inevitable Limitations.** Digital computers enabled hydrological models, divided into two families: data-driven (predictions) and process-based (explanations). Systems Dynamics exposed the limitations imposed by equifinality and scale problems, which persisted even with attempts to resolve them using vector field-based models.

- **Information Scaling.** The scale problem refers to the difficulty of reconciling the natural scales of hydrological processes with observational and conceptual scales. The solution is the scaling of information. The `TOPMODEL` achieves this by scaling soil saturation with the `TWI` index. The `PLANS` combines `HAND` and `TWI` to scale saturation across different landscapes and instantiate hydrological response units.

- **The Connectivity Theory.** Jeffrey McDonnell proposes a unifying and revolutionary theory suggesting that surface and subsurface flows are manifestations of a single phenomenon: the saturation-activation of channel networks. In light of the inescapable limitations, Systems Dynamics is deemed the best alternative to model this theory.

Mars has a rocky and sandy terrain. Its atmosphere is composed of approximately 95% carbon dioxide, with an average atmospheric pressure of about 600 Pa. Temperatures vary drastically, with average values around -60℃. The planet has no bodies of liquid water, and its surface is constantly exposed to cosmic and solar radiation due to the lack of a magnetic field.

Chapter 4

The ecological framework for planning

> When I worked at the World Bank, I often heard the statement: "There is no conflict between economy and ecology. We can and must grow the economy and protect the environment at the same time". I still hear this often today.
>
> Herman Daly (2015, p. 1) [7]

> We can break our dependence on fossil fuels, excessive consumption, and the current development model by creating a more sustainable and desirable future. It will not be easy; it will require a new vision, new measures, and new institutions. It will require a directed evolution of our entire society. But breaking this dependence is not a sacrifice of quality of life. On the contrary, it is a sacrifice not to do so.
>
> Robert Costanza (2016, p. 23) [194]

4.1 Choosing an economic paradigm

In the previous chapter, I explored the evolution of Hydrology throughout the 20th century, which has rapidly shifted paradigms since the advent of the International Hydrological Decade in the 1960s, arriving today at an era that acknowledges the diversity of hydrological processes, especially in zero-order basins. Furthermore, the use of hydrological models has also become more critical and aware of its limitations, reflecting the advance of the instrumentalist approach which, in the absence of complete information about the target system, accepts multiple empirically adequate realizations and the treatment of associated uncertainties. This trajectory raises a fundamental question for water resources management: how can this sophisticated theoretical and practical knowledge offered by Hydrology be applied to enhance the integrated management of water resources?

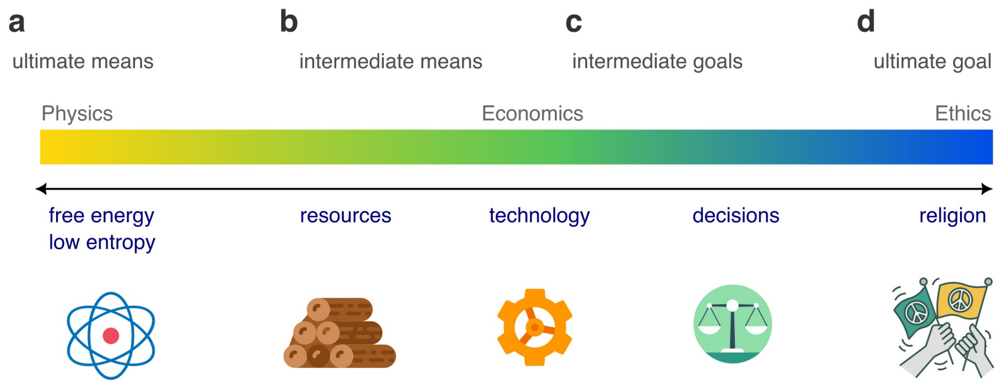

Figure 4.1 — The spectrum of Economics. Economics, as the allocation of scarce resources among different goals, connects resources (means) and objectives (ends). **a** — Free energy and low-entropy materials are the ultimate means. **b** — Energy and materials manifest through intermediate means: natural resources and resources transformed by technology. **c** — Decisions based on the application of technology guide the allocation of intermediate resources to meet intermediate goals. **d** — Decisions are oriented by some religion, a set of ethical values aimed at a supreme goal, whether explicit or not.

In the context we find ourselves in, where water resource management aims to ensure society's water security, an inevitable confluence with Economics — the science of allocating scarce resources — emerges. This challenge becomes even more critical in the face of climate change, which this year (2024) manifested alarmingly in various parts of the world, revealing society's vulnerability to extreme weather events. Simultaneously, recent initiatives in Brazil, such as the National Program for Watershed Revitalization, the National Policy for Payments for Environmental Services, and the consolidation of ANA's Water Producer Program, signal a movement being nurtured by federal institutions, creating momentum for other Union members and sectors as well. Therefore, more than a superficial application of hydrological models is needed. On the contrary, it becomes essential to deeply understand the **economic paradigm** that will guide these decisions. Only then will it be possible to advance in the use of hydrological knowledge and modeling techniques to design management policies that respond both to the demands of an uncertain climate context and the need to rationally allocate our natural resources.

Although this chapter advances to a more applied sphere, aimed directly at supporting water resource management, it will be necessary to take a few steps back and explore fundamental philosophical questions. This digression, however extensive it may seem, is indispensable to grounding the use of hydrological models within **Ecological Economics** — the economic paradigm that embodies both the concept of "sustainable development" and the concept of "environmental or natural services". From this understanding, we can more deeply interpret what it means to regenerate or conserve natural services in watersheds. In particular, we will see that to reach the application of models, I will need to scale a stack of concatenated concepts, from abstract ethical principles to the mapping of priority rural plots for green infrastructure expansion. In relation to the landscape analogy I have developed so far, if in the previous chapter we entered an open and tangible terrain, where rivers flowed abundantly, here the rivers enter avulsion, creating vast swamps as labyrinthine as the mountain tops. With some patience, however, it is possible to reach the sea.

4.2 Between means and ends: Economics

Economics is the science dedicated to studying the allocation of scarce resources among different objectives. For example, we might choose to use steel to build houses or missiles; allocate water for agricultural irrigation or energy generation; use land to maintain standing tropical forests or to raise and slaughter cattle. What is the best option? If these resources are truly scarce, these options are not mere fallacious false dichotomies but very serious questions, with drastically different consequences depending on the choice. In this sense, Economics is not only associated with the existence of scarce resources but is also intrinsically linked to the **decision-making** process, as choices must be made because we cannot have everything simultaneously. This raises the following central question:

– What are the objectives that guide choices?

This is clearly not an empirical question. It cannot be answered by undertaking a field expedition and testing hypotheses against observed evidence, as scientific questions are. In this case, the empirical world would simply offer a confusion of contradictory answers. Nor is it a Logical question, which could be answered from fundamental axioms, as in Geometry. The question, therefore, reveals itself to be essentially **Ethics**: it presupposes a moral value of what *should* be aimed at[1]. Ethics is the branch of Philosophy that explores the principles and values guiding human behavior, seeking to define what is considered right or wrong, just or unjust, what should and should not be done. Therefore, the central question of Economics depends on an ethical orientation, making this science a bridge that connects means (resources) with ends (objectives), as illustrated in Figure 4.1. In the strict realm of Biology, of course, all organisms must deal with the allocation of scarce resources to address the **existential imperative**, which has nothing to do with Ethics. This imperative consists of a clear objective: to exist, one must make decisions to *continue* existing. Thus, the ultimate goal of a living being is to metabolize and reproduce, by definition. In human terms, this implies obtaining the minimum conditions for survival, such as food, water, clothing, hygiene, physical shelters, and, above all, social relationships.

Unlike other living beings, historian Yuval Harari suggests that *Homo sapiens* instantiate **religions** that define existential goals beyond mere survival [195]. Religions are collective fictions, intersubjective realities that exist only in the shared imagination among people of a given group. In this sense, religions define the ethical framework by which humans guide their choices. The historical processes of the Renaissance and Modernity gave rise to **Humanism** in Western Europe, a religion that is widely influential today. In contrast to earlier religions that worshiped supernatural entities, Humanism is expressed as a form of **anthropocentrism**, as it centers on the human being as the focal point of all our objectives [196]. An example of the transformation implied by Humanism can be found in René Descartes' *Discourse on the Method*, where the philosopher proposes that the pursuit of truth should go beyond a theoretical understanding of reality to find practical applications that improve human conditions [8]. In the medieval mindset, however, this conception was inaccessible, as knowledge aimed to understand God's work, and the miserable human condition would be saved only after death. Other religions eventually assume that our misery is related to reincarnations, etc. Through a political lens, Humanism gave rise to various sub-religions, or doctrines, such as Liberalism, which worships the "individual", Socialism, which worships "society", and Fascism, which worships the "nation"[2]. But primarily, Humanism

[1]This type of question is also known as **normative**.

[2]These doctrines are obviously not rigid, existing on a gradient and featuring unusual combinations with one another and with other traditional religions. For example, Socialism may be influenced by

also gave rise to modern **Naturalism**, a theory grounded in realism that asserts that nothing exists outside the natural world and, therefore, any explanation of reality must be found within the Universe itself.

4.2.1 The ethics of Utilitarianism

In the field of Ethics, a philosophical branch within Humanism is Jeremy Bentham's **Utilitarianism** (1748-1832) [197]. This philosopher proposes that the morality of an action should be determined by its consequences, focusing on maximizing the **human well-being** of the largest number of individuals possible. For Bentham, the fundamental principle guiding moral evaluation is **utility**, understood as an action's capacity to increase or decrease the pleasure and pain of those affected by it. Bentham defines human well-being as the predominance of pleasure over pain and argues that all human actions are motivated by these two factors. He develops an evaluation method known as the **hedonic calculus**, which allows for measuring the moral impact of an action by considering aspects such as intensity, duration, certainty, proximity, and extent of the pleasures and pains involved. This calculation aims to quantify the well-being generated by actions and guide moral choices. Bentham's Utilitarianism thus adopts a consequentialist and impartial perspective, where the correctness of an action depends on its practical outcomes rather than the agent's intentions. Additionally, this theory considers that the well-being of all affected should be weighed equitably. Thus, the morally correct action is the one that promotes the maximum pleasure and minimum pain for the largest number of people, serving as the ultimate metric for judging the moral value of actions.

From a distance, Bentham's Utilitarianism makes sense. Within Humanism, the idea of maximizing human well-being, of guiding decisions to avoid pain and seek pleasure, is intuitive and natural. However, this ethical theory is vague enough to allow for diverse interpretations, leading to very different political and economic orientations in practice. Theoretically, Liberalism, Socialism, and Fascism are utilitarian in one form or another, though the materialization of their ideas during the 20th century resulted in vastly different economies and political regimes. In Liberalism, the premise is that if each individual is free to maximize their own well-being, the overall well-being will increase through a kind of self-regulation. In Socialism, it is assumed that collective well-being will grow through State regulation or management of the economy. In Fascism, it is believed that the general well-being of the nation will be maximized as long as individuals or groups considered traitors or impure are expelled (or worse).

To complicate matters, Utilitarianism faces challenges with a very fragile scientific foundation. Modern psychological and neurological theories suggest that it is impossible to indefinitely amplify a subjective sensation. **Adaptation Level Theory**, for instance, establishes that constant emotional stimuli shift to a mental background where they are then ignored, freeing mental resources to adequately process *novelties* [198]. Along these lines, the pursuit of more well-being leads humans into a **hedonic treadmill**, a form of chronic and addictive existential dissatisfaction where, even with constant improvements in material comfort and rewarding experiences, the level of well-being tends to return to a fixed baseline, frustrating the search for lasting and increasing happiness [199]. Empirical evidence supports this thesis, such as in situations where lottery winners (a positive event) and individuals disabled by accidents (a negative event) return to their baseline level after the event [200]. Since neuroscience supports that pain and pleasure are fundamentally material processes (synapses), these psychological and physiological states cannot grow indefinitely, complicating the idea of

Liberalism, as seen in the European Union, or by Fascism, as seen in the Soviet Union.

"maximizing human well-being"[3].

Economics has deeply incorporated the concept of maximizing utility as its **ultimate goal**, which underpins both **Microeconomics**, focused on explaining the functioning of specific markets and the interactions between consumers and producers, and **Macroeconomics**, which seeks to understand the economy as a whole, from the national level to the global economy. However, the classical doctrines of these theories, developed in the 19th century, although scientific, struggled to sustain themselves empirically over time. Theorists began questioning their fundamental assumptions as early as the 19th century, and these critiques intensified throughout the 20th century. In Microeconomics, the classical version was revised by modern approaches such as **Evolutionary Microeconomics**, which expands classical concepts by introducing bounded rationality, dynamic behavior, and institutional interactions [201], [202]. This approach aims to explain more complex and realistic phenomena, such as price dispersion and the diversity of industrial structures, which classical theory, based on perfect rationality and static equilibrium, found challenging to address.

Another transformation of substantial importance in Microeconomics was the advent of **Environmental Economics**[4]. According to Pearce (2002) [203], this microeconomic school originated in the 1950s in the United States with the creation of *Resources for the Future* (RFF). This institution was dedicated to investigating issues related to the scarcity of natural resources and their economic impacts. According to the author, the 1960s were marked by a strengthening of this discipline, driven by the environmental movement, notably through works such as *Silent Spring* (1962) by Rachel Carson and *Spaceship Earth* (1966) by Boulding. This context led economists to delve deeper into the study of **externalities** — unpriced impacts of economic activities on third parties — based on theoretical foundations such as those developed by Pigou (1920). Pearce highlights that, building on these foundations, economists began employing cost-benefit analyses to underpin environmental public policies, using concepts such as the Kaldor-Hicks compensation criterion. With this advancement, the traditional separation between natural resource economics (focused on the optimal use of resources) and environmental economics (centered on pollution) began to blur. This occurred particularly with the incorporation of economic growth theories that accounted for environmental limits. Additionally, numerous contributions emerged on how to address externalities, proposing solutions such as taxes, regulations, or private negotiations. These approaches fostered market-aligned environmental policies. Thus, Environmental Economics established itself as a subdiscipline of Neoclassical Microeconomics, integrating theories of economic welfare, sustainable growth, and instruments such as pollution taxes and tradable permits to promote economic efficiency and environmental sustainability. Despite its practical nature regarding natural resource management instruments, this school does not challenge the broader understanding of Macroeconomics. On the other hand, proponents of Ecological Economics, such as Herman Daly (1938–2022), completely reject the classical and neoclassical Macroeconomics doctrine, arguing that it overlooks the laws of thermodynamics. As we will see later, Ecological Economics proposes a genuine paradigm shift, placing sustainability at the center of macroeconomic analyses and emphasizing that unlimited economic growth is unfeasible on a planet with finite resources.

[3] Neuroscience establishes that thoughts are synapses. Without synapses, thoughts do not exist. However, neuroscience does not explain the *subjective manifestation* that accompanies synapses. This is known as the hard problem of consciousness.

[4] Also called Neoclassical Environmental Economics, to denote its distinction from Ecological Economics

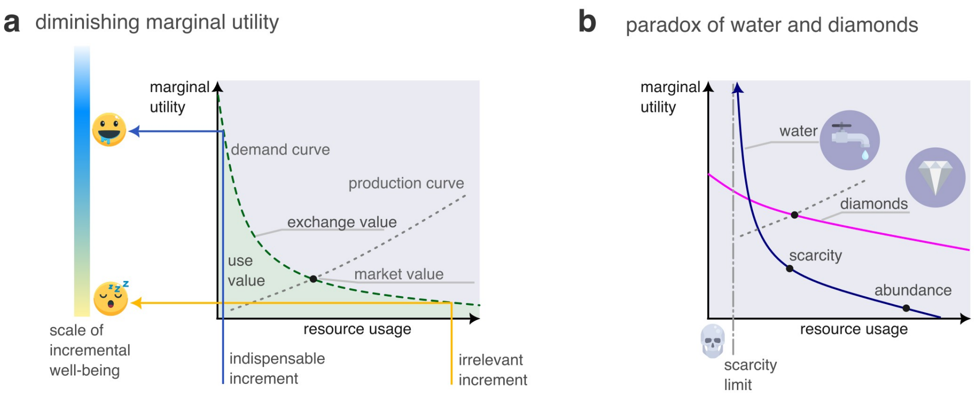

Figure 4.2 — Marginal utility and demand curves. The theory of marginal utility explains the observed price system in markets. Marginal utility, therefore, is equivalent to the exchange value of resources. **a** — The theory is founded on the principle of diminishing marginal utility, the concept that the economic agent's consumption gradually satisfies them as they consume a given resource, forming a downward-sloping demand curve. Initially, the first units consumed produce high marginal utility, an indispensable consumption increment. But marginal utility decreases as consumption increments become irrelevant. The production curve for this resource on the supply side establishes the point at which producers and consumers are equally satisfied, forming the market price. **b** — The theory of marginal utility explains exchange value in conditions of scarcity and abundance. Water, essential to life, has low marginal utility and market value when abundant, but in scarcity, its marginal utility rises, making it valuable. In contrast, diamonds, which are non-essential, maintain high market value due to their scarcity. The scarcity limit indicates the point where the marginal utility of essential resources like water reaches critical levels, raising its exchange value far above non-essential resources.

4.2.2 The Theory of Marginal Utility

In Microeconomics, Utilitarianism found its peak expression in the **Marginal Utility Theory** during the 19th century, articulated by thinkers such as William Stanley Jevons and Carl Menger [197]. This theory seeks to explain prices in markets, understood as the result of thousands of independent and decentralized decisions among consumers and producers. In a free-market system, information about a resource's scarcity is conveyed to consumers through its **price**. The Marginal Utility Theory holds that the price of a good or service constitutes its **market value** or **exchange value**, in conditions that satisfy both producers and consumers (see Figure 4.2**a**). This value corresponds to the additional utility an individual gains by consuming an extra unit of a good, which diminishes as consumption increases — the so-called **principle of diminishing marginal utility**. Practically, this means that a consumer tends to become *satiated* as they consume more of a given good or service, producing a downward-sloping demand curve.

The theory successfully explains the so-called **value paradox** paradox, the bizarre value difference between diamonds (of little utility) and water (of great utility) (see Figure 4.2**b**). In this case, the **use value** of a resource essential for survival, like water, can be extremely high, while its exchange value is relatively low, especially where water is abundant. Conversely, goods such as diamonds or jewelry have low use value but high exchange value due to their scarcity and the high demand from those who consider them status items (though the exchange value is low in societies that prefer feather headdresses, for example). In a situation of extreme water scarcity, on the other hand, the exchange value of water could easily surpass that of diamonds, as it becomes a matter of life or death. This contrast illustrates how use value and exchange value can diverge significantly, and how marginal utility helps explain the market price of a good or service, regardless of its functionality.[5] Based on this logic,

[5] Karl Marx's Labor Theory of Value contrasts with Marginal Utility Theory by asserting that the **labour value** of a good is determined by the socially necessary labor to produce it, rather than by the

the theory provides a model to explain (and predict) price formation in markets, and its concept of **Pareto efficiency** suggests that, under certain conditions, markets reach an equilibrium, allocating scarce resources optimally to satisfy everyone without improving one person's situation at the expense of another.

The explanatory power of Marginal Utility Theory seems to corroborate the concept of utility, as consumers and producers in real markets supposedly adjust their decisions according to marginal benefits — even without ever having heard of Jeremy Bentham or diminishing marginal utility. Along with other auxiliary theories, Marginal Utility Theory forms the foundation of Classical Microeconomics, which seeks to scientifically explain the general functioning of specific markets. This set of theories assumes that agents in the economy are fully rational, capable of maximizing their individual utility, and that markets operate efficiently to maximize total utility, reaching an equilibrium automatically, without external interventions. However, for the supposed efficiency in resource allocation to be achieved, it is necessary for there to be no **market distortions** and for ideal conditions to be present, such as the existence of many producers, full availability of information, and the absence of speculation and advertising — factors rarely observed in reality. In addition to market distortions, the irrational behavior of human beings, as demonstrated by Daniel Kahneman [10], challenges the idea of completely efficient markets. Consumers and producers, influenced by cognitive biases such as loss aversion and anchoring bias, often make decisions that are not entirely rational, generating inefficiencies and price distortions.

Evolutionary Microeconomics attempts to overcome these limitations by adopting a more dynamic approach, primarily including simulations with agent-based models [202]. In this case, agents have bounded rationality and local information, and their interactions occur over time, in explicitly defined processes. Moreover, various institutions, beyond the market, influence and sustain resource allocation. Thus, evolutionary microeconomics provides a more robust theoretical explanation for phenomena that classical microeconomics struggles to justify. The agent-based models of the evolutionary approach, naturally chaotic and computationally irreducible, make it clearer that, although they aim to make predictions, microeconomic theories generally remain limited to exploratory modeling. These models, therefore, offer a wide explanatory capability but low predictive capability, making empirical adequacy occur in very specific cases, typically favoring some of the auxiliary hypotheses but not the model in its entirety.

4.2.3 Macroeconomic paradigms

As previously mentioned, Macroeconomics differs from Microeconomics by seeking to explain the economy as a whole, going beyond specific markets and encompassing national and global economies [204]. The classical and neoclassical doctrines of Macroeconomics, however, tend to be a generalization of Microeconomics, expanding the view to all markets, producers, and consumers in a region of interest, such as a country. The classical macroeconomic model consists of the **circular flow of exchange value** between producers and consumers. In practice, this corresponds to the flow of exchange value between firms and households, which can be measured annually by the total revenue of companies and household income. Since it is a circular flow, the total value of

utility the good provides to the consumer. According to Marx, the value of a good is proportional to the amount of labor embedded in its production, including both direct labor and the labor necessary to produce the means of production. This concept underpins Marx's analysis of labor exploitation to accumulate capital, where goods are exchanged based on the labor value, but workers receive only a fraction of this value, creating surplus value, or profit, for the owners of the means of production. Although the labour value of a good or service may exist, the problem with Marx's thesis is that, in practice, goods produced by workers are sold in markets for the exchange value received by consumers.

one must be equivalent to the other.

The Neoclassical school revises the classical model by introducing injections and leakages in the circular flow of exchange value caused by private finances (individual investments and savings), public finances (tax collection and public investments), and international finances (imports and exports). The maximization of total utility, regarded as the ultimate goal, can be understood as the maximization of this exchange value flow in the circular system. Thus, the Neoclassical school focuses solely on the exchange value flow in markets and strategies to increase it, a concept known as **economic growth**. Differences among these strategies generally involve more or less regulation or state intervention in markets to reduce unemployment and stabilize prices. Fiscal and monetary policies aim to control the leakage of taxes and individual savings in some way. The success of a strategy, however, is tested by the increase in companies' revenues or household incomes, with this growth interpreted as a sign of progress toward the ultimate goal of maximizing total utility.

Ecological Economics, primarily articulated by Herman Daly, proposes a new paradigm that rejects the Neoclassical school [205]. This perspective proposes that macroeconomic theory be incorporated into **Ecology**, a science based on physical principles, recognizing that the target system to be modeled is, essentially, a material system [206]. Evidently, the material perspective on Economics was present in its modern origins but was slowly replaced by an interest in exchange value, which is immaterial [207]. A good example is that of Thomas Malthus (1766-1834), who identified potential limits to human population growth based on food production, which would grow at a relatively slower pace. In essence, Malthus clearly saw humans as material beings who need nutrients and energy to survive, like any other species on the planet. For example, if a human needs to consume 3 liters of water and 0.4 kg of food daily (approximately 2000 calories, divided among fats, carbohydrates, and proteins), it follows that today's global population of 8 billion people demands 24 million cubic meters of water and 50,000 tons of food daily. Without a material flow of this magnitude, the global population would not survive for long. Eight billion seems like a large number, but would it be possible to double or triple that number?

Since Malthus's predictions of widespread famine and population collapse did not materialize in his time, classical and neoclassical economists tend to dismiss his theory, confusing his circumstantial conclusions with the material foundations of his approach. In the heated debate that arose with the proposal of Ecological Economics, neoclassical economists accuse proponents of the new paradigm of being "neo-Malthusians", a term with a somewhat pejorative connotation [208]. The ecological perspective, in turn, accuses the Neoclassical school of being excessively focused on the exchange value flows of **commodities**, very specific goods and services that can be bought or sold [209], [210]. This focus on value flow is so dominant that the Neoclassical school does not even consider the materiality of commodities; they are merely *vehicles* that carry marginal utility, defined by their price. Thus, the Neoclassical school instantiates an immaterial and abstract ontology of exchange value. This representation may be an interesting model for understanding and explaining value flows in markets, but it is far from being a representation of material reality.

4.2.4 The ontology of physicalism

Ontologically, Ecological Economics upholds a physicalist reality. **Physicalism**[6] is a realist and naturalist line of thought that posits that objective reality exists, is governed

[6] Also known as Materialism.

by the laws of Physics, and ultimately comprises matter and energy [211]. As Physics is a scientific theory, Physicalism forms the primary worldview directly implied by Scientific Realism. From this perspective, even the concept of utility is considered material, as notions like "well-being" or "satisfaction" are, in reality, synapses in brains, an organ of the human body.

Despite its appeal due to strong empirical adequacy, Physicalism faces some critical structural problems. One major issue is the mind-body problem, or the **hard problem of consciousness** — the apparent impossibility of explaining subjective experiences from physical processes like synapses [212]. Another serious issue is the **free will problem**, which arises when accepting the ultimate consequences of Physicalism, namely Determinism [213]. In a deterministic reality, decisions would merely be sensations, devoid of real agency, rendering the idea of choices in resource allocation, central to Economics, meaningless[7]. However, both issues are symptomatic of Scientific Realism, which proposes that Science aims to describe the truth about reality. Instrumentalists, such as Bas van Fraassen and Nancy Cartwright, contest this view, arguing that scientific theories aim merely to produce *empirically adequate* descriptions without necessarily reflecting the ultimate truth about the world [26], [28]. In this sense, choosing a physicalist ontology for Macroeconomics provides a more empirically robust foundation than the Neoclassical school, whose explanation is limited to the flow of exchange value in markets. Whether true or not, a physical basis offers greater explanatory capability and predictive capability than the Neoclassical circular flow model.

One of the pioneers in accepting and promoting Physicalism within Economics, laying the groundwork for the ecological paradigm, was Nicholas Georgescu-Roegen (1906-1994), especially with his seminal work *The Entropy Law and the Economic Process* (1971). In this work, Georgescu-Roegen emphasizes the implications of the first and second laws of thermodynamics for an economy based on material resources. The **first law of thermodynamics**, known as the **principle of conservation**, states that matter and energy can neither be created nor destroyed, only transformed from one form to another. The **second law of thermodynamics**, on the other hand, establishes that the **entropy** of an isolated system tends to increase, indicating that natural processes follow a trajectory of increasing disorder. This means that **free energy**, capable of performing work, tends to degrade into forms less capable of doing work. Similarly, material structures tend to spontaneously disintegrate from an ordered (heterogeneous and less probable) state to a disordered (homogeneous and more probable) state. The second law also implies that no process of energy transfer or material conversion is 100% efficient, as there will always be losses in the form of heat and waste. Therefore, these laws establish that the maintenance of order in the Universe is only possible locally and at the cost of external sources of free energy. Even so, the creation of order through work inevitably results in the generation of material and energetic waste without utility.

Another pioneer of the physicalist view was Jay Forrester, the inventor of Systems Dynamics, a topic introduced in Chapter 2. Jay Forrester and his team of engineers and scientists at `MIT` demonstrated through the publications *World Dynamics* and *Limits to Growth* how compartmental models can be employed to obtain predictions about the state of the global economy over time [45], [214]. The simulations of the `World3` model introduced a physicalist approach, demonstrating the materiality of the economic system by integrating the levels and material flows of the global system, such as population, industrial capital, arable land, and resulting pollution. The results, although varied across different scenarios, highlight the physical limits imposed by the biosphere on the anthroposphere, such as the depletion of non-renewable energy

[7]Radical Physicalism leads to Nihilism, an ethical theory that denies the existence of moral values or ultimate objectives to guide decisions

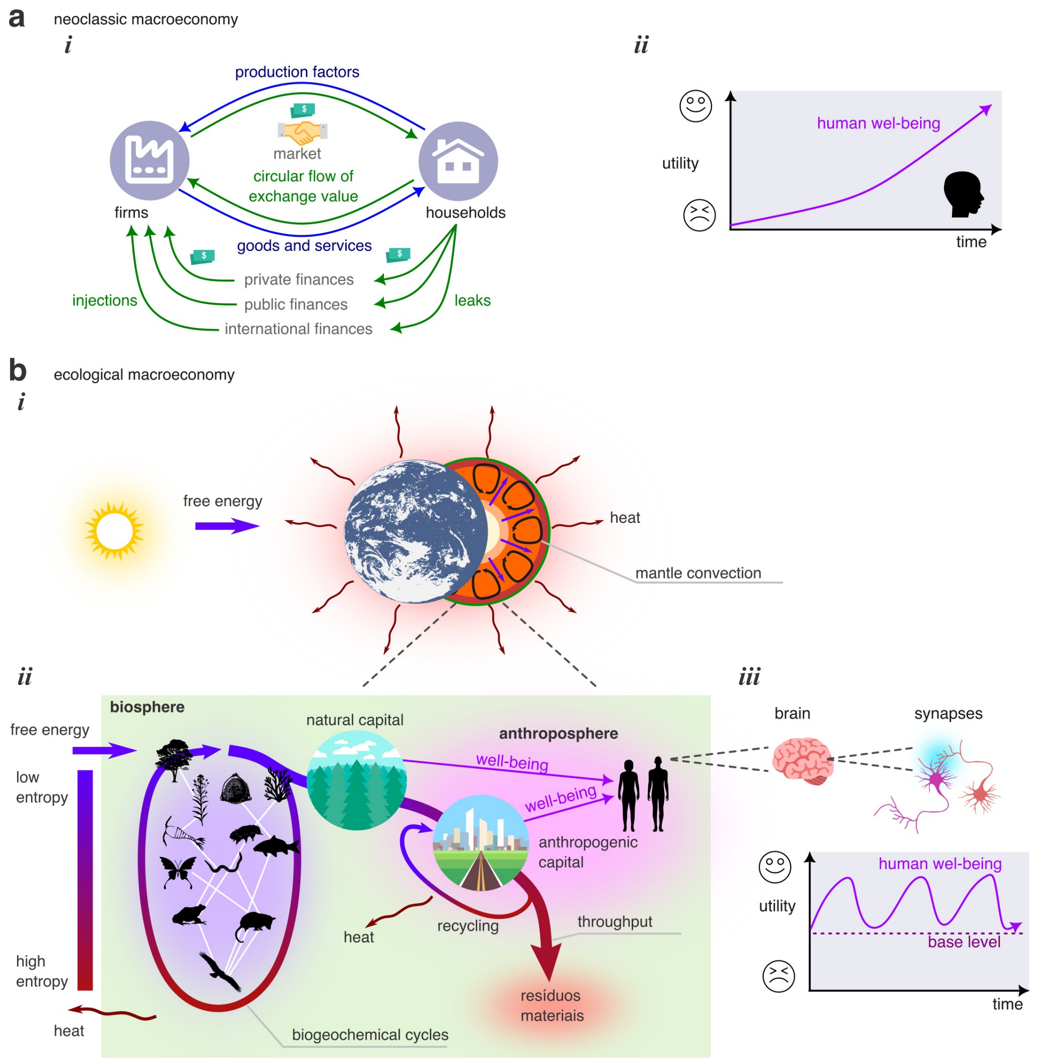

Figure 4.3 — Ecological Macroeconomics. Unlike neoclassical macroeconomics, ecological macroeconomics instantiates a model based on the laws of thermodynamics. **a** — The neoclassical macroeconomic model seeks to represent a circular flow of exchange value between firms and households, with injections and leakages generated by fiscal and financial policies of private and public agents (detail *i*). If the exchange value flow increases, utility increases, as exchange value is equivalent to marginal utility. Consequently, it is concluded that an increase in the exchange value flow is a direct indicator of increased human well-being (detail *ii*). **b** — The ecological macroeconomic model assumes a physicalist ontology, instantiating matter and energy flows in open systems. The biosphere is the Earth's upper layer that slowly exchanges matter with the mantle and receives free energy flows from the Sun and core, emitting heat to space (detail *i*). The free energy flow in the biosphere is processed by biogeochemical cycles, including ecosystems, circulating materials between high and low entropy states. The anthroposphere consists of the human habitat embedded within the biosphere and is traversed by a linear flow of matter, resulting in waste (detail *ii*). In this system, human well-being is also a material and finite process, derived both from anthropogenic capital, the structures of the anthroposphere, and directly from natural capital, the pre-existing structures in the biosphere (detail *iii*).

sources, agricultural land degradation, and the finite capacity of ecosystems to absorb pollutants. In the various scenarios explored, human population growth and industrial capital inevitably encounter a tipping point, where increasing resource extraction costs, agricultural productivity loss, and pollution impacts impose increasingly negative feedbacks, eventually leading to a decline in industrial capital growth. Thus, the explorations of `World3` illustrated the interdependence between the economic system and the physical limits of the biosphere.

4.2.5 The ecological model of macroeconomics

Based on thermodynamic laws and a systemic view, Ecological Economics represents macroeconomics through an ecological model, illustrated in Figure 4.3. In this model, the **anthroposphere** – the material environment inhabited by humans – is embedded within the **biosphere**, encompassing Earth's material system, from the atmosphere to the first kilometers of the lithosphere. Energetically, the biosphere is not an isolated system but an open one, receiving a continuous flux of free energy from solar radiation and Earth's inner core. This high-quality energy, after performing work, is emitted into space as less useful forms of heat. Materially, the biosphere is also connected to input and output flows driven by subduction, burial, uplift, and volcanism processes, which result from mantle convection, along with the loss of light gases into space. Although processes like volcanism are more abrupt exceptions, most material exchanges in the biosphere occur extremely slowly on geological scales. Thus, the biosphere can practically be considered an energetically open but materially closed system.

The energy openness of a system is essential for creating order within it. Photosynthesis, for example, is one of the main order-generating processes in the biosphere, employing free energy from solar radiation to produce low-entropy materials, such as sugars, from high-entropy compounds like carbon dioxide. The burial of these low-entropy materials over millions of years produced the fossil fuel reserves widely used today by the anthroposphere. Beyond photosynthesis, the driving force behind the carbon cycle, various other **biogeochemical cycles** occur in the biosphere, all powered by external energy sources. The hydrological cycle, for instance, is driven by solar energy, which evaporates water and raises it, imparting gravitational potential that can perform work in mills or turbines. These cycles maintain the dynamics and order of the biosphere, demonstrating the interdependence between external energy input and internal system organization.

The physical foundations of Ecological Economics result in a completely different interpretation of economic growth. The first law of thermodynamics, viewed through the lens of the ecological model, implies that no real material growth exists in the biosphere, as matter is conserved. The anthroposphere, as a human habitat within the biosphere, can expand its structural complexity by processing materials from the biosphere, but always at the cost of generating heat and other waste. The second law of thermodynamics, in turn, implies that merely *sustaining* the anthroposphere requires a flow of free energy and replacement materials, as it is necessary to continuously work against the spontaneous degradation of the material system. This inevitable flow of matter processed by the anthroposphere and generating useless waste dispersed into the biosphere is called **throughput flow** in Ecological Economics[8]. In this sense, the economic growth theorized by Neoclassical Economics is, above all, an illusion, an imaginary phenomenon instantiated by the intersubjective reality of exchange value. However, since goods and services in markets are material entities, increasing their consumption necessarily implies increasing throughput flow. The difference between economic paradigms becomes very clear here: while one instantiates a circular and immaterial flow (marginal utility and exchange value), the other envisions a linear and material flow (processing of free energy).

[8] Free translation of *throughput*.

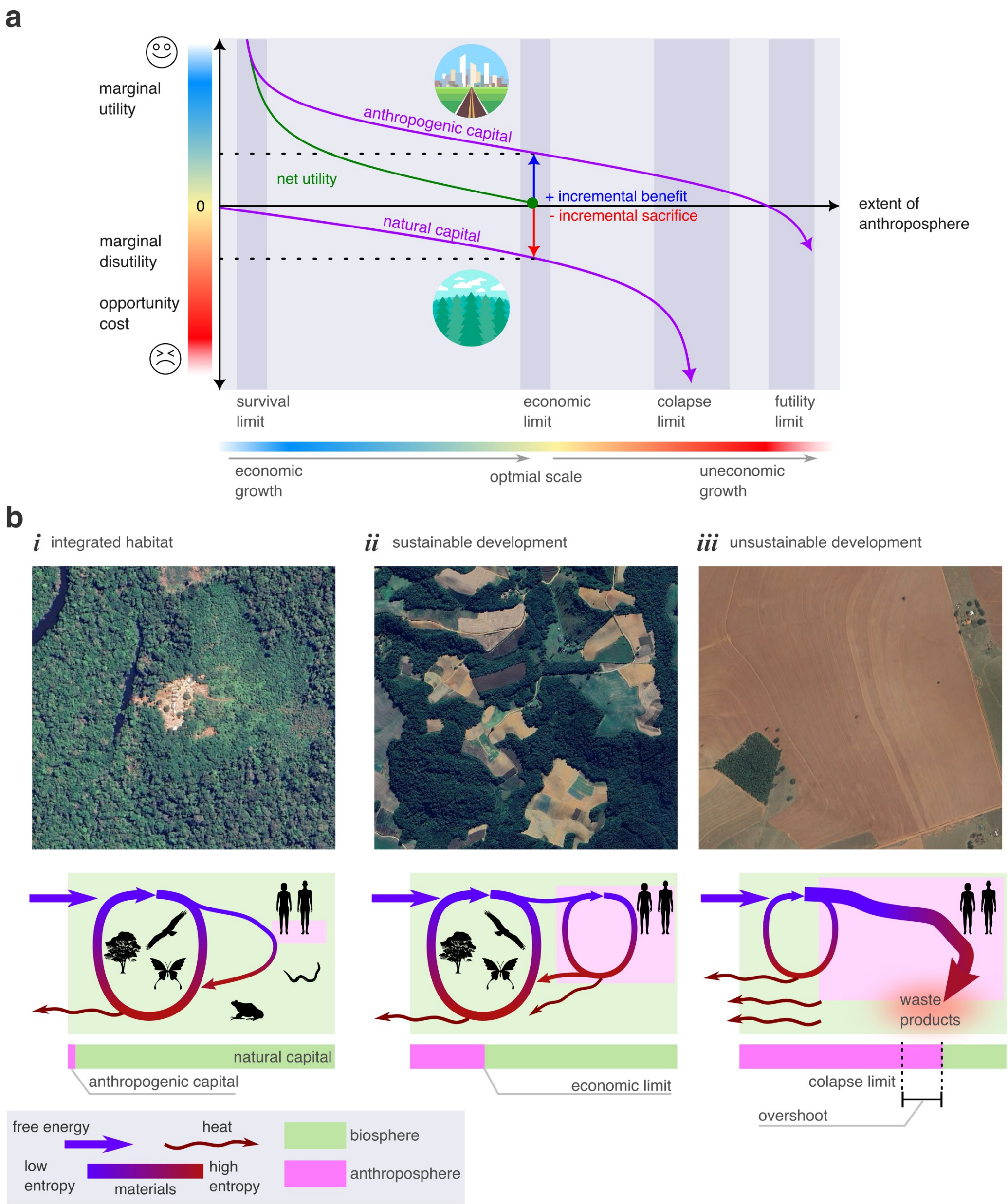

Figure 4.4 — The optimal scale of the anthroposphere. The main implication of the ecological macroeconomic model is that, being material, the anthroposphere imposes an economic extension over the biosphere, or an optimal scale. The research questions of the ecological paradigm revolve around identifying the optimal scale and other limits from different perspectives. **a** — The expansion of the anthroposphere incurs both diminishing marginal utility and an opportunity cost, or marginal disutility, due to the sacrifice of natural capital for the expansion of anthropogenic capital. The survival limit, therefore, is the minimally viable habitat that humans need to metabolize and reproduce. The economic limit is the scale at which the incremental benefit is equivalent to the incremental sacrifice. Beyond the optimal scale, the expansion of the anthroposphere is anti-economic, as it results in negative net marginal utility. In this region, there is the collapse limit, which varies according to the scale of analysis, and the futility limit. **b** — The optimal scale can be visualized by land use, showing the gradual transition from integrated habitat with small open clearings (detail i), through a mosaic of fragmented areas (detail ii), to complete anthropization (detail iii). Sustainable development represents the maximum size of the anthroposphere that can still be integrated into the ecological system, which requires a resource recycling flow.

4.3 Natural capital

4.3.1 The optimal scale of the anthroposphere

Although Ecological Economics rejects neoclassical ideas in Macroeconomics, it shares several fundamental concepts of Microeconomics, such as Marginal Utility Theory and the ultimate goal of maximizing human well-being. In this sense, Herman Daly [7] proposes that, as the anthroposphere is a material subsystem of the biosphere, there exists an **optimal scale** for the anthroposphere, the point at which human well-being reaches its maximum, as illustrated in Figure 4.4. Beyond the so-called **economic limit**, the growth of the anthroposphere becomes anti-economic, generating increasing detriments. The concept of **sustainable development** is directly tied to this idea, indicating that the material growth of the economy must be balanced within the limits of the biosphere. Thus, economic choices need to consider this balance, implementing mechanisms and policies that prevent the anthroposphere from exceeding its optimal point. The research agenda of this paradigm therefore aims to investigate how close we are to this limit. In an empty world, there is room to expand. In a crowded world, however, this limit may have been surpassed, requiring planning that prioritizes the **degrowth** of the anthroposphere.

Understanding the optimal scale is only possible by generalizing the concept of *capital* beyond its conventional conception. In the ecological view, **anthropogenic capital** consists of the material and social structures built by humans that produce a flow of goods and services, resulting in human well-being. The definition is vague and may be confused with the anthroposphere itself. A motorcycle factory, for example, is a typical example of industrial capital that produces goods (motorcycles). However, a motorcycle can be used to provide a delivery service, making it a good on one hand and capital on the other. The ability to ride a motorcycle makes the workers of a delivery company its human capital. The term "capitalism", in this sense, consists of the doctrine of accumulating anthropogenic capital or expanding the anthroposphere. It should be noted that, from the neoclassical perspective, capitalism explicitly dictates capital growth but only indirectly assumes that this will increase human well-being by increasing the flow of exchange value. As we have seen, this doctrine results in the expansion of the anthroposphere over the biosphere, increasing throughput flow.

Beyond anthropogenic capital, the ecological view also introduces the concept of **natural capital**. The idea was initially articulated by Robert Costanza and Herman Daly, in analogy to built capital [215]. Thus, natural capital is defined by the material structures of the biosphere itself that produce a *direct* flow of goods and services, resulting in human well-being. In addition to directly usable goods, such as timber or fish stocks, services like pollination, which supports food production, fertile soil formation, essential for agriculture, and biogeochemical cycles like carbon sequestration, which regulates the global climate, are also considered. These services stem from fundamental **environmental functions** essential to life on Earth, even though they are often overlooked by the conventional approach to natural resources [216]. A good exercise to illustrate their importance is to imagine what would be required to reproduce natural conditions in a human colony on Mars. Without the support of natural capital, all processes essential to life would need to be artificially replicated, highlighting the critical interdependence between the anthroposphere and the biosphere.

The core of the optimal scale concept, therefore, is to recognize that constructing anthropogenic capital requires dismantling the existing natural capital infrastructure. This necessarily results in a sacrifice of well-being provided directly by resources and natural services. In economic jargon, there is an **opportunity cost** to the expansion of the anthroposphere, which is the sacrifice of resources and natural services provided by natural capital. Marginal Utility Theory helps illustrate this dilemma through the utility diagram (Figure 4.4): the economic limit of anthropogenic capital growth occurs when its marginal utility equals its **marginal disutility**. marginal disutility consists of the incremental loss of utility on the natural capital side and is exactly the opportunity cost. This incremental sacrifice of utility is small in an empty world: a small clearing in an immense forest is practically imperceptible in terms of detriments, even at the local scale of the ecosystem. Indeed, it can be conceived that the anthroposphere requires a minimally viable limit, the **survival limit** of humans in the biosphere. But as the world becomes crowded, the losses in utility grow non-linearly. In many cases, it is possible for a **catastrophic limit** to be exceeded, when virtually irreversible degradation processes are triggered. The overload region is a dangerous situation, as crossing the collapse limit does not imply an immediate catastrophe due to the delay in the propagation of impacts within the ecological system. The diagram represents the entirety of the biosphere but can also be interpreted at other scales, such as countries, watersheds, etc., and in relation to specific natural processes. In a watershed, for example, the economic limit would be a balanced land occupation and water use that preserves various other goods and natural services. A more intuitive example is the global climate system: the economic limit would be a greenhouse gas pollution level balanced in terms of well-being, while the catastrophic limit represents a pollution level at which climate change processes become uncontrolled, triggering positive feedback loops.

The economic limit and catastrophic limit are generally lower than the **futility limit**, the level at which the marginal utility of anthropogenic capital is zero or even negative (the situation when more resources directly bring *less* well-being). This brings important political implications that challenge the social *status quo*. The Neoclassical economic school tends to be well-regarded by social elites, as it suggests that simply increasing resource consumption will raise collective well-being, regardless of the relative inequalities among consumers. With the so-called **promise of a bigger slice**, it is enough to make the "pie" grow for everyone so that the poorer will have a larger slice tomorrow than today — even if the slicing remains unfair. But by admitting that an optimal scale exists for the anthroposphere, we see that resources should be consumed at the economic limit to maximize human well-being, far below the futility limit. The only viable way for someone to consume resources above the economic limit is to ensure that others consume far below, one compensating for the loss of natural capital by the other. Since this unequal arrangement does not maximize collective well-being, Ecological Economics postulates that, in a finite material world, inequality is immoral.

4.3.2 Understanding natural resources

The classification of **natural resources** by Ecological Economics contrasts sharply with the approach of the Neoclassical school. Since Neoclassical Macroeconomics tends to be an extended view of Microeconomics, this paradigm treats resources homogeneously, reducing them to mere inputs, or **factors of production**, that contribute to maximizing goods and services through anthropogenic capital. In general, these factors are categorized as "land" (raw materials) and "labor" (workforce), along with capital itself (machinery and equipment). In the Neoclassical view, the idea of **substitutability** prevails, implying the possibility of substituting one input for another when it becomes scarce, which creates a relative insensitivity to resource depletion. By not recognizing natural capital, resource depletion is not seen as a fundamental problem, provided that technological innovations allow for substitutions.

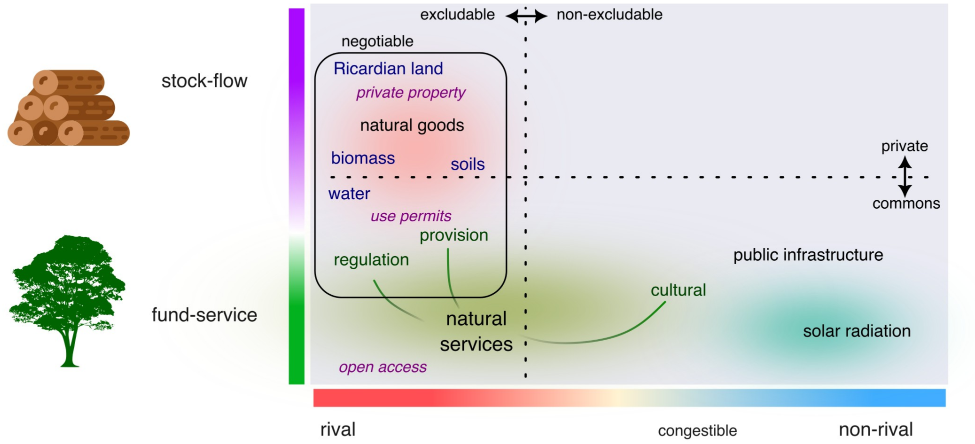

Figure 4.5 — Classification of natural resources. The ecological paradigm organizes scarce resources into stock-flow and fund-service typologies. Stock-flow resources provide utility when materially transformed during their consumption, while fund-service resources provide utility without being transformed. Stock-flow resources can be consumed at any rate but are depleted suddenly. Fund-service resources can only be consumed at a maximum possible rate and can eventually become congested. All stock-flow resources are rivals by definition, meaning multiple users cannot consume them simultaneously, and all non-rival resources, by definition, are fund-services. Natural services, unlike raw materials, are generally rival fund-services and therefore require some regulated form of exclusivity to avoid the tragedy of the commons.

Nonetheless, Herman Daly articulates a critical distinction between **stock-flow resources**, which can be consumed at any rate, and **fund-service resources**, whose productivity is time-limited and cannot be easily substituted [205]. natural capital can be of either type — or both, as in the case of forests. Forests are a stock-flow of timber, a raw material. But they are also a fund-service of various **natural services** (see Figure 4.5). Here, the scale problem becomes evident: the economic benefit of dismantling a forest comes with the opportunity cost of the sacrificed natural services. However, since forests are a system that regenerates with solar energy, it is possible to find a sustainable flow of timber extraction without entirely giving up natural services. This classification highlights the need for strategies that consider the limits and **irreplaceability** of certain resources, providing a more suitable framework for designing sustainable development policies.

stock-flow resources are those resources materially transformed during the production process, meaning they become part of the final product. In the context of natural capital, these resources are classified as **renewable resources** and **non-renewable resources**. A clear example is oil, which is refined into fuel, or trees, which are converted into boards and paper. When available, these resources can be consumed at nearly any rate. For instance, large quantities of oil can be quickly extracted if enough wells and refineries are in operation. Similarly, a forest can be cleared within days if sufficient machinery and workers are available. These resources can also be stored for future use; raw materials like grains, fuels, and minerals can be stockpiled, allowing for flexible management over time. Similarly, forest stands can be preserved for future harvesting.

When consumed at a rate exceeding the **natural renewal rate**, stock-flow resources resources become suddenly depleted. Since oil takes millions of years to form naturally, it is considered a non-renewable resource, as is the case with most **mineral resources**. Forests, however, when properly managed, can regenerate within a reasonably short time, making timber a renewable resource, as is the case with most **biotic resources**. Water, despite being an **abiotic resources**, is also renewable because the hydrological cycle constantly operates to eventually restore river and lake levels.

However, once these resources are used, they are destroyed and cannot be recovered in their original form. In some cases, the waste generated can be reused in other processes or recycled. Identifying **reusable resources** and **recyclable resources** is an essential strategy for sustainability as it reduces throughput flow in the anthroposphere. Nonetheless, it is important to remember that these processes do not reduce the need to import free energy for work.

In contrast, fund-service resources are those that participate in the production process *without* being physically transformed into the final product. They provide essential means and services for the production process but are not immediately exhausted. Their depletion occurs through wear or **depreciation** over time, following the second law of thermodynamics. In the case of anthropogenic capital, infrastructure such as highways, energy systems, communication networks, as well as machinery and labor, are examples of these resources. They do not become part of the final product; a sewing machine, for example, participates in clothing manufacturing but is not incorporated into the product. Unlike stock-flow resources, fund-service resources have a limited production capacity per unit of time. A machine or a worker can produce only a finite quantity of products within a given period, regardless of the available raw material. Additionally, these services cannot be stored. If a factory remains inactive for a week, that week's production capacity is lost and cannot be recovered later. Over time, these resources wear out; a machine, for example, rusts and requires maintenance, and workers need rest and continuous training to maintain productivity. Therefore, these resources do not exist statically but also participate in throughput flow, consuming free energy and low-entropy materials.

Unlike anthropogenic capital, which requires continuous maintenance through human intervention, natural capital possesses intrinsic **self-organization** mechanisms. Biogeochemical cycles, for example, are driven by sources of free energy such as solar radiation and Earth's core heat, making the biosphere a self-organized system that sustains life and regulates the flow of materials and energy. As we will see in more detail later, natural services are the fund-service resources we obtain directly from this self-regulated system. These services can be categorized into three major groups: **provisioning natural services**, which encompass the biosphere's ability to provide and regenerate stock-flow resources, such as food, water, and raw materials; **regulating and maintenance natural services**, which control vital processes like the water cycle, climate regulation, and pollination, balancing and organizing material flows within the biosphere; and **cultural natural services**, which refer to the immaterial well-being provided by human interaction with the natural environment, through recreational, scientific, religious, or spiritual activities. Concrete examples of these services include tropical forests' ability to sequester atmospheric carbon, regulating the global climate, or coral reefs, which not only provide habitats for a vast marine biodiversity but also serve as natural barriers against coastal storms. Similarly, interacting with natural landscapes, such as national parks, promotes mental health and human well-being, exemplifying the cultural dimension of these services.

4.3.3 The tragedy of the commons

Designing resource allocation policies to keep the anthroposphere within its economic limit invariably involves addressing the **open access problem**. This issue was articulated in the article *The Tragedy of the Commons* by ecologist Garret Hardin (1968) [217]. In this article, Hardin argues that environmental problems, such as land overuse and water pollution, are consequences of unrestricted access to goods and services that

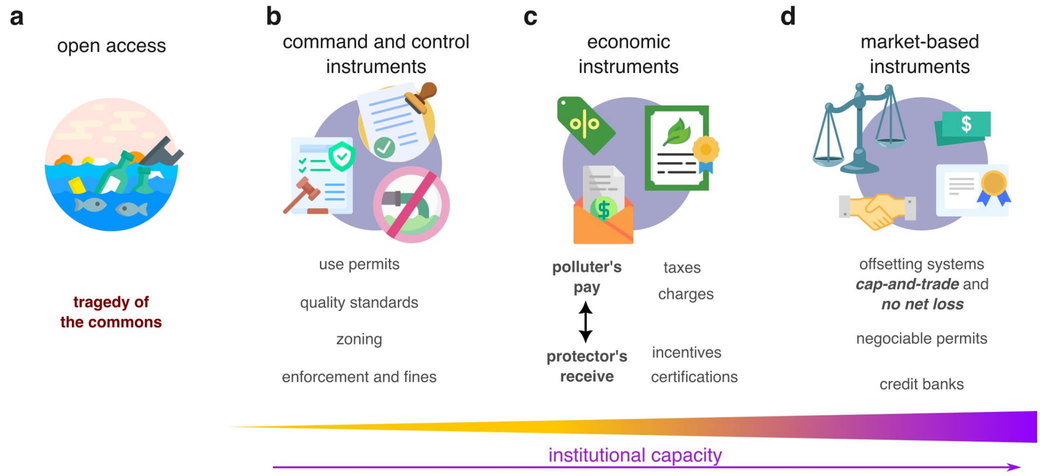

Figure 4.6 — Solutions to the open access problem. The tragedy of the commons arises with common resources that lack institutionalized exclusivity. The solution involves incremental layers of management instruments, requiring greater institutional capacity from the State. **a** — Without regulations on the use of common resources, individual action leads to a race to exploit natural services and raw materials, producing the tragedy of the commons — an environmental collapse. **b** — Command-and-control instruments are the first layer of management, setting general rules enforced by law on economic agents, such as usage permissions, quality standards, zoning, and penalties for infractions. **c** — Economic instruments form the second layer of management, encouraging economic agents to change their behavior beyond the mandatory general rules. Here, both the polluter-pays principle, with taxes and fees, and the protector-receives principle, with incentives and certifications, can be applied. **d** — Market-based instruments form a third layer of management, creating a negotiation space for usage permissions, allowing economic agents to find optimized arrangements on their own. In this approach, compensation systems are typically established, creating a market of credits and debits, such as cap-and-trade and no net loss. Banks are necessary to facilitate and reduce transaction costs.

are commonly used, or **common resources**[9]. As individuals and companies act solely to maximize their own well-being, unrestricted access ultimately creates a rush to exploit the resource, which ends up being congested or depleted. It becomes a tragedy, therefore, as the supposedly rational actions of individuals turn into irrational collective behavior. Hardin thus recognizes that treating public resources similarly to private goods can be an interesting solution, though it does not apply in all cases:

> The tragedy of the commons as a food basket is averted by private property, or something formally like it. But the air and waters surrounding us cannot readily be fenced, and so the tragedy of the commons as a cesspool must be prevented by different means, through coercive laws or taxing devices that make it cheaper for the polluter to treat their pollutants than to discharge them untreated. We have not progressed as far with the solution to this problem as we have with the first. Indeed, our particular concept of private property, which deters us from exhausting the positive resources of the earth, favors pollution. The owner of a factory on the bank of a stream—whose property extends to the middle of the stream—often has difficulty seeing why it is not their natural right to muddy the waters flowing past their door. The law, always behind the times, requires elaborate stitching and fitting to adapt it to this newly perceived aspect of the commons. – Garret Hardin (1968, p. 1245) [217].

The problem of water and air pollution can be addressed by ensuring that polluters take full responsibility for their waste. A central concept in Neoclassical Environmental Economics, in this regard, is **externality**, which can be either positive or negative. Externality refers to the impact of one economic agent's actions on the well-being of others [218]. In Hardin's example, where a factory pollutes the air and water, regulatory

[9]This term refers to both natural capital and anthropogenic capital.

mechanisms should be created to prevent the loss of collective well-being (the negative externality), forcing the factory to manage its waste, even if it incurs higher costs. In this context, Ecological Economics provides a broader interpretation of open access problem and externalities compared to Neoclassical Environmental Economics. The concept of negative externality, according to this perspective, is related to the opportunity cost of the expansion of the anthroposphere over the biosphere. Transforming a river into an effluent channel generates a series of negative externalities precisely due to the loss of natural capital — the resources and services directly provided by the river, ranging from the supply of drinking water to cultural activities.

Here, the concepts of **rivalry** and **exclusivity**, also from Microeconomics, are relevant for conceiving appropriate allocation policies. **rival resources** exhibit an intrinsic characteristic that prevents their use by multiple economic agents simultaneously. This interpretation is quite intuitive: if a bakery consumes flour from a silo, it automatically prevents another bakery from doing the same; if a fishing boat catches a school of fish, it automatically prevents another boat from doing the same; if a farmer plants corn on a plot of land, they prevent another from grazing sheep on the same plot. All rival resources are excludable, but not necessarily exclusive. This is because **exclusive resources** require a social guarantee of reserved use by a given economic agent, which does not always exist. exclusivity gives rise to markets, as it turns rival resources into **tradable resources**, encouraging economic agents to allocate resources among themselves based on their exchange value. The guarantee of exclusivity manifests essentially through *deterrence* among economic agents. In small-scale societies, like families and clans, this deterrence can occur through tacit agreements. In larger societies, where strangers interact, the State tends to guarantee the right of use through a monopoly on force and the diffusion of cultural norms of coexistence. The right to **private property** is an institution created by the State to confer exclusive use of a rival resource to the owner, through an official document. To negotiate the resource, one must also negotiate the State's official document that establishes exclusivity. The difference between private property and a public concession is that property is lifelong and hereditary, though both are exclusivity guarantees provided by the State. In large societies with weak States, exclusivity can also occur but often requires greater investment in private deterrence mechanisms, such as fences, weapons, and hiring mercenaries.

All stock-flow resources are rival, and all **non-rival resources** are fund-services. Since stock-flow resources are destroyed during their use, they are inherently rival — the same coal burned by one power plant cannot be burned by another. By definition, non-rival resources are those that can be consumed by more than one economic agent *simultaneously*. Logically, therefore, these resources cannot be stock-flow resources. They are common resources, for general use, accessible by all at the same time and not exhaustible or storable. A typical non-rival resource is solar radiation, a free energy flow that is relatively homogeneously distributed. If a plant photosynthesizes with sunlight, it does not automatically prevent other plants from also photosynthesizing at the same time. Another typical example is transportation routes, such as roads and highways. If a car drives on a road, it does not prevent another car from driving at the same time. Likewise, the scenic beauty of parks and beaches offers well-being to users indiscriminately.

The open access problem can be divided between its mild form, involving non-rival resources, and its severe form, involving rival resources. non-rival resources, being necessarily fund-service resources, are considered **congestible resources** because they can be consumed up to the limit of their provision rate. Congestion, as a mild form of open access problem, occurs when the resource is neither destroyed nor degraded during use but imposes a maximum number of users. Typical examples occur when very tall buildings block sunlight for shorter buildings; when the road system becomes overcrowded with the number of cars; or when parks and beaches are crowded with people seeking leisure. In all these cases, the services provided by the resources eventually lose their use value due to unrestricted access. In such cases, the solution to the problem inevitably involves **command and control instruments**, coercive impositions by the State through the action of regulatory authorities, along with auxiliary strategies such as reinforcing cultural norms.

The severe form of open access problem occurs in the context of rival common resources, as their uncontrolled consumption not only causes congestion but can lead to environmental collapse — Hardin's "tragedy of the commons" mentioned above. This is evident with stock-flow types of common resources, such as when river water is used as an input in production processes. In this situation, unrestricted access may lead economic agents to use up all available water resources, creating scarcity and associated negative externality. However, river water can also serve as a public fund-service, as in the case of pollution. Here, the river's capacity for dilution and self-purification of waste is an example of a *rival* fund-service. The river itself is not transformed into any new product, but the natural service of regulating environmental quality is progressively degraded as it is used. When an industry discharges waste into the river, the original dilution capacity is consumed, preventing subsequent industries downstream from using the service simultaneously.

The solution to the severe problem of unrestricted access also involves implementing some form of exclusivity through the official issuance of **use permit** for economic agents [219]. Usage permits for common resources are not private property, which would imply lifelong and hereditary rights of use; instead, they resemble public concessions, periodically reviewed and renewed by regulatory institutions. In the case of water, official permits for withdrawal establish exclusivity within predefined limits, compatible with the natural availability of water sources. Similarly, official permits for effluent discharge aim to regulate the use of dilution capacity, again limiting users. Thus, allocation policies can range from conventional command and control instruments, which establish general rules, to **economic instruments** that not only impose penalties for negative externalities (**polluter-pays principle**[todo:gls]) but also offer incentives for users who generate positive externalities (**protector-receives principle**)[todo:gls]. An advancement in this direction is **market-based instruments**, or cap-and-trade policies [220]. In this type of regulation, resources retain their public nature, but use permit are defined by exclusive quotas that can be freely traded among economic agents. A classic example of this instrument is the carbon market, where agents with credits (sequestered carbon) trade with agents with debits (emitted carbon), keeping the system in a neutral condition [221].

Here, it is worth noting that the development of economic instruments and market-based instruments requires a **institutional capacity** on the part of the State *greater* than simple command and control. This is because these management policies are not merely substitutes for one another but are incrementally more complex strategies. Command and control, by itself, cannot influence economic agents beyond ensuring compliance with *general rules*, such as quality standards, zoning, etc. If command and control instruments goes beyond general rules, economic agents cease to exist *by definition*, becoming extensions of the State itself. In societies that tolerate the existence of multiple economic agents, therefore, the State needs an extra institutional apparatus to operationalize the economic limit without making decisions for the agents. In this regard, economic instruments aim to encourage behavior changes that are not mandatory but *desirable*. market-based instruments go further, with greater potential to maximize the total economic benefit, as negotiations tend to find an optimized al-

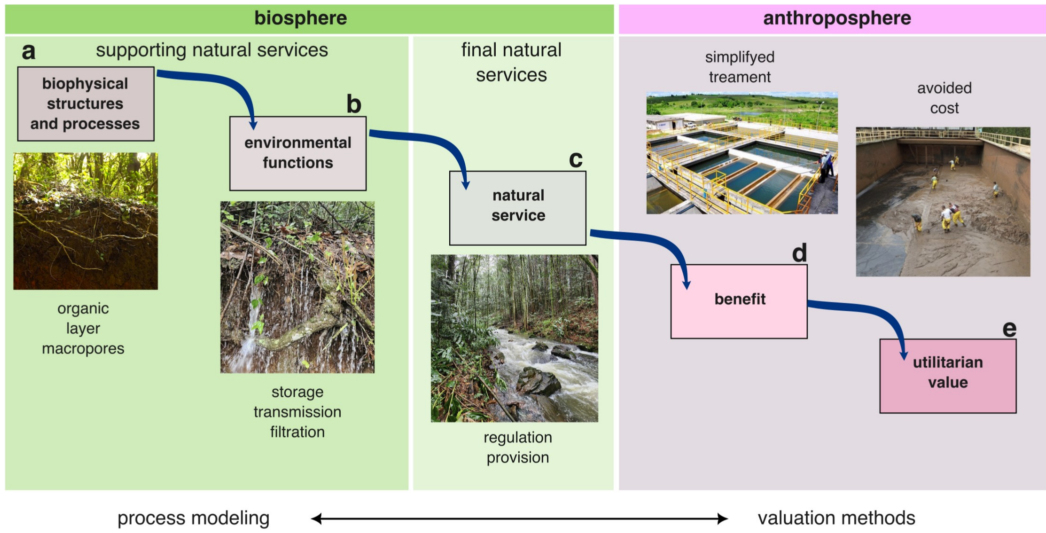

Figure 4.7 — The cascade model of natural services. The most modern conceptual framework for understanding natural services is the cascade model, proposed by Haines-Young & Potschin (2010) [222]. This model describes how natural services flow to generate benefits for society in five stages, with the first three stages in the biosphere and the last two stages in the antroposphere. Quantitative estimates can be developed at each end through process models and valuation methods. **a** — Structural base and biophysical processes: represent the fundamental ecosystem elements that support their functions, such as physical and chemical characteristics. For example, this includes the organic horizon of the soil and the presence of macropores. **b** — Environmental functions: correspond to ecological processes stemming from structural dynamics. In the soil example, these functions encompass water storage, transmission, and filtration, essential processes for aquifer recharge and natural water purification. **c** — Final natural services: benefits derived directly from environmental functions that can be used by the antroposphere, such as providing clean water or ensuring the perennial flow of downstream watercourses. **d** — Benefit: refers to the utilitarian value derived from natural services, such as improved health, environmental security, and support for economic activities. In the case of water, one of the benefits is simplified and less costly treatment for supply. **e** — Utilitarian value: represents the economic valuation of the benefits obtained. For example, one way to assess value is by estimating the costs avoided in sludge management and advanced water treatment, or the need to source high-quality water from a more distant location.

location solution within the predefined limits. In the case of pollution, the regulatory and enforcement apparatus forms a command-and-control base that cannot be abandoned, as it sets discharge standards, monitors, and applies penalties. On top of the command-and-control framework, an incentive system can be created to further reduce the use of natural services, such as encouraging water reuse. However, maintaining and updating a **credit bank** requires a higher administrative burden than just the previous enforcement apparatus. Finally, a permit market can be implemented, but not without the advent of an institution to enable and regulate its dynamics, reducing the embedded **transaction cost** of negotiating permits.

4.4 Natural services

The natural services implied by the concept of natural capital have presented both theoretical and practical challenges in designing sustainable development policies. The idea of natural service, without clear definitions, becomes intangible for influencing allocation decisions. In contrast, managing non-renewable resources is far more evident, as consuming these resources, with their irreversible destruction or transformation, makes them progressively scarce. Knowing that a coal deposit will inevitably be exhausted makes it logical to devise strategies for energy diversification. On the other hand, a tropical forest represents natural capital that provides a vast array of services, many of which are diffuse and benefit much more than just its inhabitants or visitors, at various scales. The renewable resources of the forest itself, such as timber or food, result from provisioning natural services, which can be degraded for various reasons, such as ecological collapse. To design management instruments for these natural services, such as control, incentives, and trade mechanisms, they need to be identified, measured, and assigned **economic value**. This is a major challenge in the new paradigm, even though recent advances have been made.

A more modern conceptual framework for understanding natural services is the **cascade model of natural services**, proposed by Haines-Young & Potschin (2010) [194], [222]. This model describes how ecosystem services flow from nature to benefit society in five main stages: (1) structure and process; (2) function; (3) service; (4) benefit; and (5) value. The first three stages belong to the biosphere, highlighting the primary importance of **biophysical structure**, such as forests and wetlands, which provide the foundation for the development of **ecological processes**, such as water infiltration and nutrient cycling. These processes, in turn, generate **environmental functions** that can become useful to humans. Thus, the model emphasizes that an ecosystem's environmental function, that is, its *capacity* to do something potentially useful for humans, is not automatically a natural service. Only when this function is considered *beneficial* to society does it become a service, such as the utility of flood regulation in densely populated areas. The last two stages of the cascade, on the other hand, belong to the anthroposphere. The **economic benefit**, therefore, corresponds to the tangible manifestation of human well-being derived from this natural service, such as a city being less vulnerable to flood impacts. But the perception of this well-being depends on context and specific human needs, leading to variability in the assessment of its **economic value**. For example, value varies according to supply and demand conditions: during a rodent population explosion, the economic value of pest control natural service is much higher than in normal periods. Furthermore, not all benefits can be assessed in terms of exchange value, exhibiting only use value, as in the case of cultural natural services, which directly impact subjective well-being without intermediate material products.

4.4.1 Classifications of natural services

The systematization of natural services began with Daily *et al.* (1997), especially in terms of identifying and characterizing natural services, but also regarding valuation issues [223]. The authors define these services as "the conditions and processes through which natural ecosystems, and the species that comprise them, sustain and enhance human life". They provide a pioneering categorization, though without including details on cultural aspects, which separates **general natural services** from **biome-specific natural services**. For general services, the list includes **climate regulation natural service** provided by various biogeochemical cycles (such as carbon, nitrogen, phosphorus, and sulfur) and sediment cycles in water. In this list, biodiversity is cited as a critical factor for ecosystem stability, ensuring greater resilience to disturbances. Also mentioned are **soil natural services**, such as the regulation of the hydrological cycle, physical support and nutrient provision for plants, as well as organic matter decomposition, which recycles nutrients and prevents pathogen proliferation. Other general services cited include **pollination natural service**, which supports agricultural production, and **pest control natural service**, provided by predators and parasites that reduce the need for pesticides, preserving food production stability. Biome services, on the other hand, perform all the mentioned general services but with specific uniqueness across different ecosystems. These include services from oceans, freshwater services, forest services, and natural grassland services.

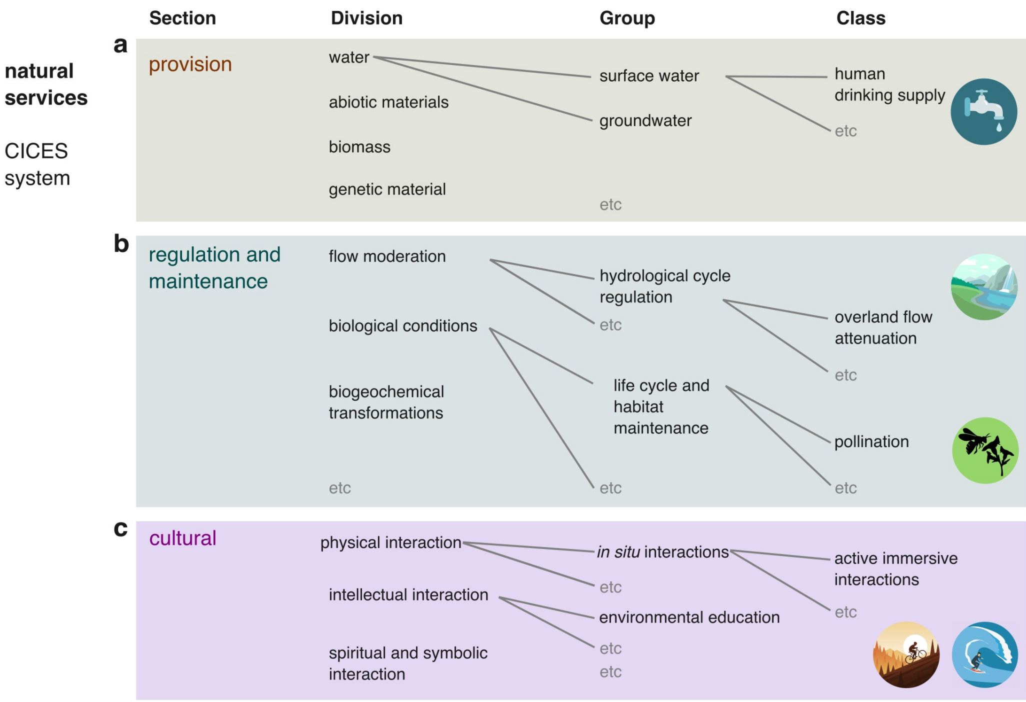

Figure 4.8 — The CICES Classification System for Natural Services The Common International Classification of Ecosystem Services (CICES) provides a system categorizing ecosystem services into three main sections: provisioning, regulation and maintenance, and cultural. Each section is further detailed into divisions, groups, and classes. **a** — Provisioning natural services are related to the direct supply of raw materials (stock-flows), such as water and biomass, essential for human supply and other economic activities. For example, water is classified into subgroups like surface water and groundwater, providing essential resources for human consumption, agriculture, etc. **b** — Regulation and maintenance natural services include fund-services from biogeochemical processes and functions that ensure environmental quality and the functioning of natural cycles. These services encompass regulation of the hydrological cycle, such as flood control, and habitat maintenance for various species, like pollination. **c** — Cultural natural services represent services that provide immaterial fund-services related to physical, intellectual, spiritual, and symbolic interactions between humans and nature. These services include recreational activities, such as interactions within the natural environment, and opportunities for environmental education, promoting well-being and cultural enrichment.

Another historical milestone for the theoretical and practical consolidation of the natural services concept was the *Ecosystems and Human Well-being* report, a United Nations initiative completed in 2005, also known as the Millenial Ecosystem Assessment (`MEA`) report[10] [224]. This report advanced, to some extent, the global modeling effort of natural resources published in 1974 by the Club of Rome in *Limits to Growth* [45], providing more detailed data on ecosystem distribution and changes. At the time, the assessment highlighted that approximately 60% of natural services were being degraded or used unsustainably, endangering the well-being of populations worldwide, especially the most vulnerable. Besides providing extremely relevant sizing, the `MEA` report helped consolidate the definition of natural services as *the direct benefits people obtain from nature*. It is worth noting that the `MEA` report also solidified the term **ecosystem service**, which is currently the most used, although there are variations like **environmental services**[11]. This terminological separation is not yet entirely clear in theory, but there are practical and regulatory differences. In Brazil, for instance, the National Policy on Payment for Environmental Services makes the separation as follows:

> (...) Art. 2º For the purposes of this Law, the following are considered:
>
> I - ecosystem: a dynamic complex of plant, animal, and mi-

[10] Millennium Ecosystem Assessment, in a free translation from English.

[11] Personally, I prefer the term "natural service", not only because it is more comprehensive but also because it is simpler for the general public to understand.

> croorganism communities and their inorganic environment interacting as a functional unit;
>
> II - ecosystem services: benefits relevant to society generated by ecosystems, in terms of maintaining, recovering, or improving environmental conditions, (...)
>
> III - environmental services: individual or collective activities that favor the maintenance, recovery, or improvement of ecosystem services.
>
> – Brazil (2021, p. 1) [2].

Despite this official definition, the term "environmental service" can also refer to benefits obtained from natural elements within built environments, such as cities. In this view, urban forestry, for example, would consist of a form of **green infrastructure** that provides environmental services.

The MEA report paved the way for the systematization of natural services currently adopted by the Commom International Classfication of Ecosystem Services (CICES), illustrated in Figure 4.8, which was mentioned in the previous section: provisioning natural services, regulating and maintenance natural services, and cultural natural services [225]. A fourth category was also articulated by the MEA, which would be **supporting natural services**, but in CICES, services in this class were either disaggregated or conceptually maintained as precursor processes of the services themselves. The systematization proposed by MEA and CICES adopts an approach opposite to that of Daily *et al.* (1997), as it takes an economic and anthropocentric perspective focused on the typology of services rather than their origin in the biosphere. This classification directly guides decisions and evaluations related to sustainable development, regardless of the ecosystems involved. In provisioning services, the main issue is whether the materials provided by the ecosystem are being consumed at a rate exceeding their regeneration capacity, as seen with soil erosion in agricultural lands or overfishing. For regulation and maintenance services, the focus is on assessing the ecosystem's capacity to regulate resources, identifying the extent to which this capacity can be used without requiring additional measures, such as soil water storage to regulate water availability or carbon absorption by oceans to mitigate global warming. For cultural services, the central question is to identify and assess the intangible benefits different groups obtain from the ecosystem, from leisure to scientific and educational value. This systematization thus emphasizes the economic aspect, facilitating policy formulation for sustainable development. Costanza (2008) suggests that this is an inevitable consequence of the transition between the biosphere and the anthroposphere, making it necessary to maintain more than one system of classification—some to map and identify services on the biosphere side and others to assist decision-making on the anthroposphere side [226].

The CICES system presents a hierarchical structure that organizes natural services into three levels: sections (highest level), divisions (intermediate level), and groups (lowest level), providing a standardization that facilitates environmental accounting and inventory. Without such a detailed system, it would be virtually impossible to develop effective policies for the management and valuation of these services. With it, projects and actions can be explicitly focused on maximizing one or more natural services. From a hydrological perspective, for example, water provision is a division within the provisioning service (code 4.2), which also covers the provision of biomass and other abiotic resources. This division is subdivided into groups, such as surface water provision (code 4.2.1) and groundwater provision (code 4.2.2). In the case of surface water, service classes include potable water provision (code 4.2.1.1), water provision as a material input (code 4.2.1.2), and water provision for energy production (code 4.2.1.3). These

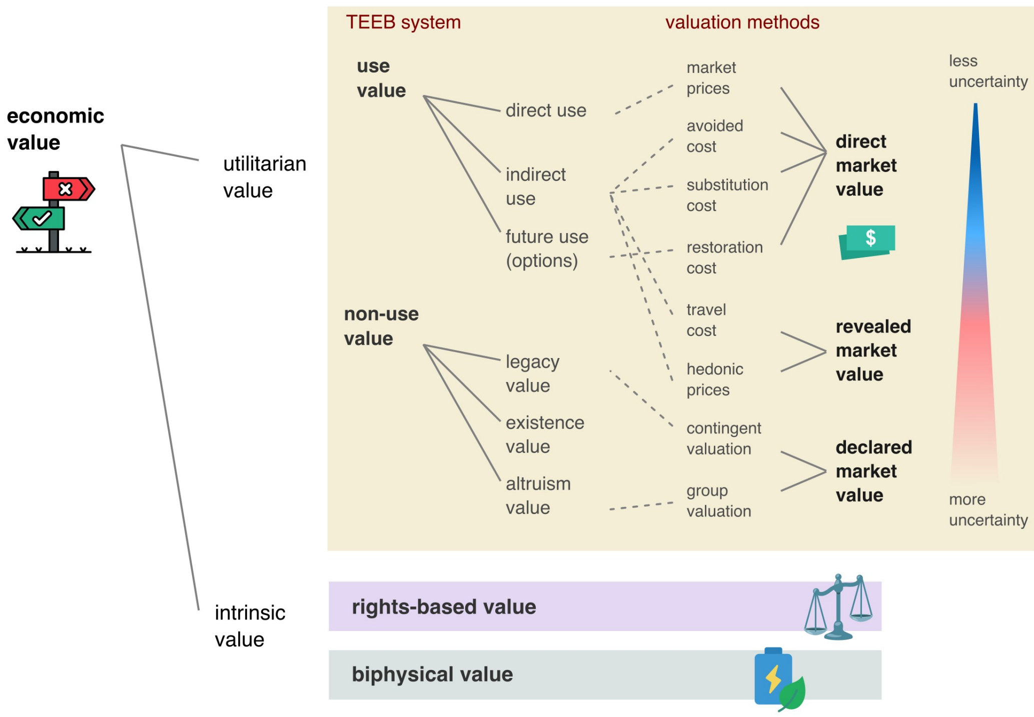

Figure 4.9 — Manifestations of utilitarian value. Utility value is the metric of the utilitarian ethical paradigm, which selects the maximization of utility as its ultimate goal. However, other forms of intrinsic value are possible for guiding economic decisions, such as rights-based (deontological) value and biophysical value, like carbon footprint or water footprint. **a** — In the TEEB system, utility value involves both use and non-use value. Use value includes direct use (consumptive and non-consumptive); indirect use (with intermediaries); and future use value, also known as option value. Non-use value consists of utility values that include legacy value, existence value, and altruism value. **b** — Valuation methods include direct market value, such as prices and costs; indirect market value, like the added value of scenic beauty and the opportunity cost of tourists' time; and declared market value, when surveys are conducted to infer value in hypothetical markets. The more direct the valuation, the lower the uncertainty associated with the method.

services can compete with each other, as in a reservoir supplying a hydroelectric plant, where use for energy generation may conflict with demand for potable or industrial water. Concerning regulation and maintenance services, attenuation of hydrological cycle flows in `CICES` results from both biotic (code 2.2.1.3) and abiotic divisions (code 5.2.1.2). Biotic regulation involves hydrological processes at the hillslope scale, where fauna and flora influence surface fragmentation, enhancing infiltration. Abiotic regulation occurs at the watershed scale, such as in floodplains. Both services require management, as they can be compromised by soil compaction on hillslopes or dike construction in floodplains.

4.4.2 Manifestations of utility value

The valuation of natural services involves the last two stages of the cascade model: benefit assessment and economic value. To fully understand this process, it is necessary to step back from practical aspects and revisit more theoretical issues. A *value* is a metric used in making choices, which are always guided by a supreme ethical objective. For instance, if honesty is an important value for human relationships, then it is *better* to interact with trustworthy people than with deceitful ones (and to be honest so as not to be *devalued* by others). Thus, valuation is the calculation of this metric based on predefined ethical foundations. In this sense, a value is considered *economic* when it is the metric for allocating scarce resources among different ends. In Ecological Economics, which is based on Utilitarianism, economic value is a **utilitarian value**—the utility or sense of human well-being. Thus, despite all the differences resulting from adopting a physicalist ontology, Ecological Economics aligns with the Neoclassical economic paradigm in its ethical dimension. Both consider utility as the ultimate goal, though the former recognizes its physical limit while the latter does not.

However, utilitarian value is not the only possible way to assess choices in an economic context, as there are also forms of **intrinsic value**, based on rights or biophysical properties [227]. The **rights-based value**[12] seeks to correct the anthropocentrism inherent to Utilitarianism, which, for example, does not explicitly ensure that other species are not driven to extinction, provided that this does not detract from the overall human well-being calculation. An alternative to this approach is **biocentrism**, inspired by Kantian ethics, which asserts that other species have the inalienable right to exist. In this context, Kant's **categorical imperative** serves as a moral principle that should be followed regardless of utilitarian consequences. This ethical principle states that decisions should be made *as if* they were derived from a *universalizable* law. Thus, this logic can be applied to establish the rights of other species, since if we dislike the idea of being used or exterminated by technologically superior species, we should not do this to less advanced species [223]. The **biophysical value** uses physical metrics, such as embodied energy and materials, to decide on resource allocation. This value is an extended version of **labor value** as identified in Karl Marx's economics. The concept of **ecological footprint**, for instance, assigns value based on the material load required to produce a good or service, such as carbon and water footprints. These valuations, though philosophically distinct, are not mutually exclusive and can complement each other through **multicriteria analysis** or multi-objective optimization. Biocentric rights, for example, can function as constraints on the extent of utilitarianism, while life cycle analysis can influence the sense of satisfaction or human well-being. After learning about the substantial carbon footprint of an airplane trip, you might be willing to pay an extra fee to offset emissions.

The possible manifestations of utilitarian value for natural services are systematically explored by the The Economics of Ecosystems and Biodiversity (`TEEB`) initiative, illustrated in Figure 4.4.2. The authors of this initiative define **Total Economic Value** as the sum of all forms of utilitarian value, organized into a hierarchy of values. It is worth noting that this is a broad conceptual system, applicable to both natural and anthropogenic capital. While obtaining the total value of a given resource is challenging and uncertain, this is often unnecessary for decision-making, as economic agents usually focus on specific components of this value. At the same time, from a natural capital management perspective, the Total Economic Value concept is crucial in reminding us that management solutions must consider the trade-offs between agents involved in different components.

The Total Economic Value is divided into **use value** and **non-use value**. use value, as previously mentioned, is a relatively intuitive and well-established concept in Economics. The Marginal Utility Theory suggests that marginal use value—the incremental benefit obtained per unit consumed—corresponds to the **exchange value**, which varies according to supply and demand in efficient markets. On the other hand, non-use value refers to the satisfaction and well-being derived from *not* using a given good or service. Although less tangible, these values cannot be ignored and include **legacy value**, reflecting concern for resource availability for future generations; **altruistic value**, which considers equity among members of the current generation; and **existence value**, relating to satisfaction derived simply from knowing that species and ecosystems continue to exist.

[12] Also known in philosophy as deontological value, as it is not consequentialist but based on rules defined *a priori*.

use value includes both **direct use value** and **indirect use value**. direct use value refers to explicit interaction with the resource, whether it is a natural service or a conventional consumption good. This direct use can be **consumptive** or **non-consumptive**. consumptive involves the irreversible transformation of the resource, as is the case for all stock-flow resources. non-consumptive, on the other hand, is more related to fund-service resources, which are not transformed. Going to the movies, for example, involves direct use value with consumptive (the popcorn) and direct use value with non-consumptive (the movie itself). In the context of natural capital, provisioning natural services offers resources with direct use value and consumptive, while cultural natural services provides direct use value and non-consumptive. A river supplying drinking water to a city, for instance, has direct use value with consumptive because the water is transformed in the process. However, recreational activities on the same river have direct use value with non-consumptive. indirect use value, on the other hand, refers to less explicit interactions with the resource, where the value is realized through intermediary processes often not immediately visible. Using the cinema example, we seldom interact directly with the cleaning staff, but a clean environment enhances our experience. Similarly, in the context of natural capital, indirect use is associated with regulating and maintenance natural services. Pollination by insects, for instance, is essential for agriculture, but this service does not involve direct interaction between the food consumer and the pollinators. Similarly, water infiltration on hillsides benefits downstream users by providing clean water, even if they do not interact directly with the natural process.

direct use value and indirect use value are both forms of use value *in the present*. However, the authors of TEEB also highlight a form of **future use value**, known as **option value**. This value, in the context of natural capital, is strongly tied to ecosystem resilience and recovery from disturbances. However, this resilience has limits and may be compromised when the ecosystem undergoes a significant disturbance, surpassing a **point of no return**. Beyond this critical threshold, the system attractor moves to a new stable or unstable state, resulting in the loss or significant degradation of previously available natural services. An example of a stable-stable regime shift is the introduction of invasive species, which quickly disrupt ecological balance and reduce biodiversity. In this situation, restoring the original ecosystem becomes extremely costly as the system attractor has shifted to another point or region of stability. A stable-unstable regime shift, on the other hand, occurs when soil erosion progresses from small rills to large gullies and ravines, with strong positive feedback intensifying erosion over time. option value reflects the **risk aversion** of regime shifts in the system, which would entail exorbitant costs. In this context, epistemic uncertainty — the lack of knowledge about the system — becomes a crucial factor in assessing this value, increasing as ignorance of the system behavior is recognized. future use value, therefore, is directly linked to the **precautionary principle**, which advises that, under high uncertainty about non-return points, policies should be guided by conservative safeguards.

4.4.3 Valuation methods

Determining each component of the Total Economic Value involves using various methods that proceed incrementally, according to available information on marginal use value (prices). Naturally, technical, methodological, and uncertainty limitations increase as available information becomes scarcer. When natural services are closely tied to real markets, **direct market valuation** techniques can be applied. In cases where natural services have an indirect relationship with real markets, information is sought in parallel real markets through **revealed preferences valuation** techniques. Finally, in the complete absence of a real or parallel market, valuation alternatives employ **stated**

preference valuation techniques, estimating the value in hypothetical markets.

In the most basic case, direct market valuation identifies market transactions directly linked to natural services. This method is divided into three main categories: **price-based valuation**, **cost-based valuation**, and **production function-based valuation**. Market price-based valuation relates to provisioning natural services whose products are traded in markets. This approach is as direct as possible: the direct, marginal use value with consumptive is the market price, which may be adjusted to remove potential distortions. Cost-based approaches, more associated with regulating and maintenance natural services, focus on estimating the expenses required to replicate the benefits of natural services through artificial means. These techniques include **avoided cost method**, which considers production costs that would be incurred in the absence of natural services; **replacement cost method**, which estimates the cost needed to replace the service itself with artificial technologies; and **mitigation or restoration cost method**, which evaluates the expenses required to mitigate or restore natural services. Production function-based approaches are the most integrated methods within this category when there is enough knowledge to establish the causal link between natural services performance and the production of a market-traded good or service.

In the intermediate case, revealed preference techniques rely on observing individual choices in existing markets indirectly related to the natural service in question. In other words, economic agents *reveal* the utility of the service through their choices. TEEB authors emphasize that the primary methodologies in this scope are **travel cost method** and **hedonic pricing method**, both strongly related to cultural natural services values with direct and non-consumptive use. Travel cost estimation is essential for valuing cultural services related to tourism and recreational activities, including direct expenses and opportunity cost of time. Thus, the value of a change in the quality or quantity of a recreational site can be inferred from the estimated demand function for visits. In contrast, hedonic pricing uses information about the implicit demand for an environmental attribute embedded in resources traded in markets. A typical example of this *added value* is the scenic beauty highly valued in property prices with views of natural landscapes. Installing wind turbines and other aspects of visual pollution in the landscape can lead to the loss of this cultural natural service, as evidenced by property devaluation.

Finally, in the absence of market information, stated preference approaches simulate a hypothetical market for natural services using surveys that consider changes in service provision. These techniques can estimate both use and non-use values, such as existence or legacy value. Among the available techniques is **contingent valuation method**, which uses questionnaires to infer economic agents' willingness to pay for improvements in a natural service or willingness to accept degradation in a natural service—two sides of the same coin. A more complex form involves the **choice modeling valuation** method, where survey respondents are presented with a network of possible choices, better indicating declared loss and gain relationships. An even more robust assessment includes **group valuation**, where different economic agent groups participate in a pluralistic process to deliberate the value of a given service or natural resource. This process enables, with multi-criteria approaches, integration with other value metrics and ethical systems mentioned above.

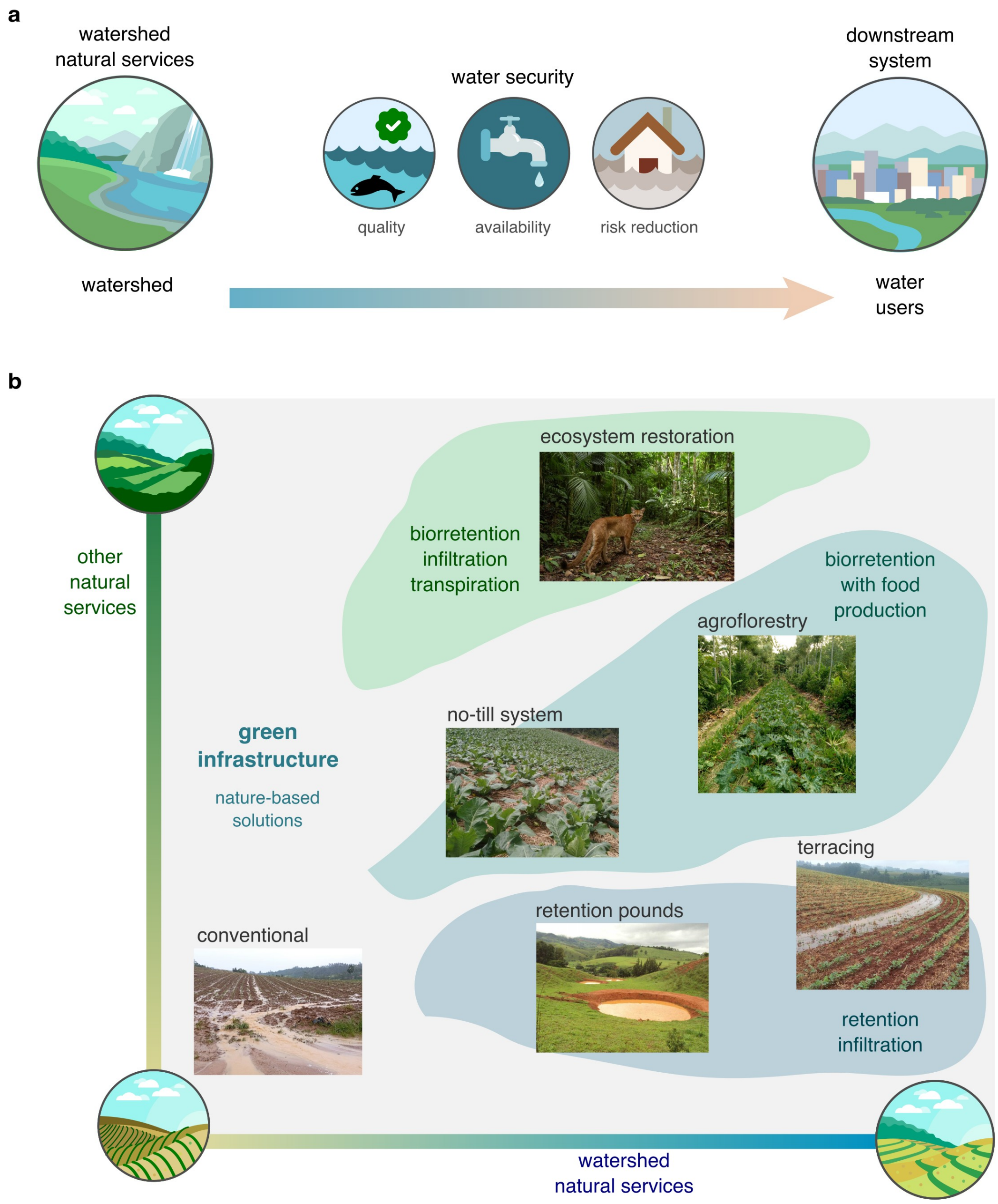

Figure 4.10 — Watershed natural services and the expansion of green infrastructure. Watershed natural services encompass a wide range of services related to land use, spanning from zero-order catchments to floodplains. **a** — The primary characteristic of watershed natural services is their scale of manifestation (the watershed) and their linear causal relationship, benefiting downstream water users. From an integrated water resource management perspective, watershed natural services contribute to water security by improving quality, increasing availability, and reducing risks associated with hydrological processes. **b** — Complementary to gray infrastructure, the expansion of various forms of green infrastructure, which are nature-based solutions, can enhance the provision of watershed natural services. In zero-order basins and slopes, these solutions include techniques to increase surface retention capacity and soil infiltration. Synergies and trade-offs with other natural services may arise here. For example, ecological restoration implies forgoing conventional food production. Agroforestry balances both water production and food production while also supporting natural services related to biodiversity.

4.5 Watershed natural services

4.5.1 Water security and green infrastructure

With the advent of Ecological Economics and its concept of natural capital, new perspectives have emerged to frame **Integrated Water Resources Management** (IWRM), defined by the Global Water Partnership as "a process that promotes the coordinated development and management of water, land, and related resources, aiming to maximize the resultant economic and social welfare in an equitable manner without compromising the sustainability of vital ecosystems" [228]. As a paradigm for watershed management and planning, IWRM largely aims to ensure **water security**. The United Nations defines water security as the capacity of a population to: ensure sustainable access to adequate quantities of water of acceptable quality, necessary to sustain its socioeconomic development; protect against water-related pollution and disasters; and preserve ecosystems in a context of peace and political stability [229]. In practice, ensuring water security requires planning and managing programs, projects, and actions that, within socioeconomic and geopolitical contexts, guarantee the availability, quality, and regularity of water in sufficient conditions to meet the demands of multiple water users, in addition to mitigating human risks associated with extreme events, such as major floods and landslides. Such users in a watershed represent different agents or economic sectors, ranging from traditional communities and water supply companies to economic sectors such as agriculture, industry, and energy production.

Under the ecological-economic paradigm, water ceases to be viewed solely as a production factor for these users and is instead recognized as a common resource provided by the natural capital of the watershed. From this perspective, it becomes clear that the allocation of water for different uses actually involves the consumption of a series of natural services provided by the basin's ecosystems, referred to as **hydrological natural services**. Smith *et al.* (2006) [230] demonstrate that these services go beyond the mere provision of water, also including the provision of food, regulation of water and sediment flows, water quality, mitigation of hydrological risks, and various cultural natural services (Table 4.1). Moreover, this interpretation suggests that the economic value of water, fundamental in allocation strategies and water use charging in Brazil[13], is actually a network of providers and consumers of the natural service of provision and regulation. If a city abstracts water from a source with good quality and availability, the value of that water is tied to the fact that the watershed is in a good state of conservation, for example. Similarly, a downstream irrigator owes part of their production to an upstream farmer who has invested heavily in soil conservation actions, such as the construction of infiltration trenches.

Evidently, the spatial extent of watershed natural services is directly linked to the watershed, distinguishing them from other ecosystem services, like climate regulation (which operates on a global scale). The upstream-to-downstream dynamics in watersheds generate a causal chain of hydrological responses, economically translating into externalities, both positive and negative. These effects vary according to the state of conservation of natural services in upstream areas and the usage of these services in downstream regions. Following the cascade model (Figure 4.7), it is evident that hydrological services arise from interactions between vegetation, soil, geology, and topography, the biophysical structures, and processes that generate environmental functions. As discussed in Chapter 3, hydrological processes occurring on slopes and in zero-order catchments are the original source of many watershed natural services. However, in large river basins, these natural services also manifest through processes like floodplain inundation and the ecological functioning of wetlands, which contribute to regulating the hydrological cycle. Recognizing this flow of externalities creates the opportunity to manage this natural capital to address the free-access problem and prevent the tragedy of the commons. Management policies, in this sense, should design and reinforce strategic connections between downstream water users and the economic agents responsible for conservation and management on upstream slopes. The ultimate goal, therefore, is to maximize human well-being by optimizing human actions within

[13] In this case, by the National Water Resources Policy (Brazil, 1997)

Service section	Watershed services	Service attributes	State indicator	Sustainable use indicator
Provision	Water supply	Precipitation, infiltration, soil retention, percolation, streamflow, groundwater flow, biotic and abiotic effects on water quality	Water storage capacity (m^3/m^2), Pollutant concentrations	Discharge (m^3/year)
	Food provision	Crop, fruit, livestock production, edible plants and animals (e.g., fish, algae, invertebrates)	Agricultural water use (m^3/ha), Fish stock (kg/m^3)	Maximum sustainable water use for irrigation (m^3/year), Net productivity (kg/ha/year)
	Non-food goods	Production of raw materials (e.g., timber, reeds), production of medicines	Amounts available (kg/ha/year)	Maximum sustainable harvest (kg/ha/year)
	Hydro-electric power	Flow for energy generation	Storage capacity of riverbeds and lakes (m^3/m^2), Slope (degree), Elevation (m)	Maximum sustainable energy production (kWh/year)
Regulation	Regulation of water flows	Retention of rainfall and release (especially by forests and wetlands), Water storage by rivers, lakes, and wetlands, Groundwater recharge and discharge	Infiltration capacity (mm/h), Water storage capacity of soils (m^3/m^2)	Baseflow volume (m^3/year)
	Hazard mitigation	Reduced flood peaks and storm damage, Coastal protection, Slope stability	Maximum natural water storage capacity (m^3/m^2)	Size (km^2) and economic value (US\$/km^2/year) are protected from flooding
	Control of soil erosion and sedimentation	Protection of soil by vegetation and soil biota	Infiltration capacity (mm/h), Slope length (m), Barren land (%)	Soil loss (kg/ha/year), Sediment storage (kg/m^2/year)
	Water purification	Reduced siltation of streams and lakes, Nutrient uptake and release by ecosystems, Removal of organic matter, salts, pollutants	Nitrogen amount (kg/ha), Total dissolved solids (kg/m^2), Electric conductivity (μS/cm)	Denitrification (kg/ha/year)
	Wildlife habitat	Wildlife and nursery habitats	Resident and endemic species (number), Surface area per ecosystem type (ha)	Increase or decline in species population size (number)
	Environmental flows	Maintenance of river flow regime	Area of critical habitats (ha), Discharge for each season (m^3/day)	Fish species and population, Total fish catch (t/year)
Cultural	Aesthetic and recreational services	Landscape quality and features, Recreational value	Stated appreciation, Recreational value (e.g., entrance fees, US$/visit)	Houses on lakeshore (number/km), Visitors (number/year)
	Heritage and identity	Landscape features or species	Cultural significance and sense of belonging	Visitors (number/year)
	Spiritual and artistic inspiration	Inspirational value of landscape features and species	Books and paintings using watershed as inspiration	Pilgrims (number/year)

Table 4.1: Serviços naturais hidrológicos — Relação dos serviços naturais hidrológicos, principais atributos, indicadores de estado e indicadores de uso sustentável. Adaptado de Smith et *al.* (2006) [230].

the watershed's economic limit.

In a world of rapid urbanization, water security primarily depends on managing the sources that supply cities [231]. In this regard, the World Bank report by Dudley *et al.* (2003) [232] highlighted the importance of managing watershed natural services in source areas, noting that protected areas with well-managed forests generally improve both the quality and quantity of water supplied. The study showed that natural forests, especially well-conserved ones, offer higher quality water with fewer sediments and pollutants compared to other catchment areas. Moreover, mature native tropical forests can increase water flow, although young forests and exotic plantations may reduce it. The report also revealed that about one-third (33 of 105) of the world's largest cities drew water directly from protected areas, while five others depended on distant watersheds that included protected areas, and at least eight cities drew water from forests managed specifically for water supply. The authors suggested that despite the benefits these areas provide for urban supply and biodiversity, the economic value of hydrological services was still underestimated. The report proposed creating financial mechanisms, such as user fees, to help cover the protection and management costs of these areas, a concept now aligned with Payments for Ecosystem Services incentives (details to follow).

With the role of watershed natural services well established, maximizing water security has increasingly been conceived within integrated water resource management through a combination of anthropogenic and natural capital [1]. Cities, for example, rely on **gray infrastructure** for potable water, including storage dams, pumping, and treatment stations, among others. However, the vegetation and soil in the watershed above the intake point act as **green infrastructure**, reducing the intensity of runoff on slopes—a quick-response process with high erosion potential. Hydrologically, this effect is mainly due to the high infiltration capacity of forest soils, as discussed in Chapter 3. The high degree of surface fragmentation in forests also reflects a relatively higher effective surface retention capacity, helping control rainfall runoff that exceeds soil infiltration capacity. Table 4.2 summarizes the range of actions that green infrastructure can offer under different natural services contexts, across various action scales, in comparison with their gray infrastructure counterparts.

The integration of green and gray infrastructure has increasingly been framed by the concept of Nature-based solutions (NBS), practices that are inspired by or directly utilize natural processes to improve water management, food production, and biodiversity conservation [1], [233]. The underlying idea of NBS integrates established concepts such as ecological engineering, soil conservation management, and environmental management approaches. Additionally, it aligns with the principles of Ecological Economics, which recognize the central importance of natural capital for sustainable development [234]. The distinctive feature of NBS, however, is that it does not necessarily require the preservation of pristine and untouched natural capital, such as native forests, but rather expands the scope of actions by *drawing inspiration* from natural processes, enabling an ecological *transition* that favors the *reproduction* of these processes, even in anthropized landscapes. For example, a crop field that adopts no-till techniques with terraces and bioretention and infiltration bands reproduces a biophysical structure analogous to the infiltration processes observed in native forest areas. Thus, NBS facilitate ecosystem restoration and the recovery of ecosystem services by integrating practices that mimic essential ecological functions in managed environments. Forestry or agroforestry with non-invasive exotic tree planting can also contribute to the recovery of degraded soils, regenerating the organic horizon and preparing the soil for native forest regeneration in subsequent stages. In this way, NBS promote a gradual transition toward the economic limit without relying solely on intact natural ecosystems.

Payment for Environmental Services (PES) schemes[14] are economic instruments for environmental management that encourage the sustainable use of natural services. In the context of watershed management, they are known as Watershed PES, promoting changes in land use, especially by expanding green infrastructure in zero-order basins in headwater areas. The water PES instrument can be seen as the formalization of **conservation leases**, agreements that maintain private land ownership but transfer the right to use it for natural resource conservation for a defined, renewable period. The negotiation of land use rights can vary from total prohibition in highly sensitive areas to implementing sustainable management practices, with the paying institution responsible for monitoring compliance by current and future landowners. In this context, the land itself should be understood under the concept of **Ricardian land**, which consists of the indestructible and inexhaustible characteristics that the land surface offers for human activities to develop [235]. Ricardian land is generally not a common resource, as it is often parceled into exclusive lots among economic agents, primarily as private property. Additionally, Ricardian land has a use value associated not only with soil fertility but also with access conditions and proximity to *other* resources. For instance, due

[14] A commonly used term in Brazil, as per the National Policy. Internationally, the term "Payments for Ecosystem Services" is used.

Water management issue/natural service	Green Infrastructure solution	Watershed	Floodplain	Urban	Corresponding Grey Infrastructure solution
Water supply (flow regulation)	Re/afforestation and forest conservation	x			Dams and groundwater pumping
	Reconnecting rivers to floodplains	x	x		Water distribution systems
	Wetlands restoration/conservation	x	x	x	
	Constructing wetlands		x	x	
	Water harvesting			x	
	Bioretention and infiltration	x		x	
	Permeable pavements			x	
Water purification (quality regulation)	Re/afforestation and forest conservation	x			Water treatment plant
	Reconnecting rivers to floodplains	x	x		
	Riparian buffers	x			
	Wetlands restoration/conservation	x	x	x	
	Constructing wetlands		x	x	
	Bioretention and infiltration	x		x	
	Permeable pavements			x	
Erosion control (quality regulation)	Re/afforestation and forest conservation	x			Reinforcement of slopes
	Riparian buffers	x			
	Reconnecting rivers to floodplains	x	x		
	Wetlands restoration/conservation	x	x	x	
	Constructing wetlands		x	x	
Water temperature control (quality regulation)	Re/afforestation and forest conservation	x			Dams
	Riparian buffers	x			
	Reconnecting rivers to floodplains	x	x		
	Wetlands restoration/conservation	x	x	x	
	Constructing wetlands		x	x	
	Green spaces (shading of water ways)			x	
Biological control (quality regulation)	Re/afforestation and forest conservation	x			Water treatment plant
	Riparian buffers	x			
	Reconnecting rivers to floodplains	x	x		
	Wetlands restoration/conservation	x	x	x	
	Constructing wetlands		x	x	
Riverine flood control (disturbance regulation)	Re/afforestation and forest conservation	x			Dams and levees
	Riparian buffers	x			
	Reconnecting rivers to floodplains	x	x		
	Wetlands restoration/conservation	x	x	x	
	Constructing wetlands		x	x	
	Establishing flood bypasses		x		
Urban stormwater (disturbance regulation)	Green roofs			x	Urban stormwater infrastructure
	Bioretention and infiltration			x	
	Water harvesting			x	
	Permeable pavements			x	

Table 4.2: Green infrastructure solution for enhancing watershed natural services — Systematized relationship by Cassin et *al.* (2021) [229] between natural hydrological services, green infrastructure, application scales (watershed, floodplain, or cities), and the corresponding grey infrastructure solution.

to its location, a hectare of land within a city's urban perimeter has a relatively higher use value than a hectare in rural areas. However, land use often generates negative externalities that degrade many common natural services, including watershed natural services. An example of this is the soil's vadose and groundwater zones' water storage function, which prolongs river baseflow during droughts—an essential service for water security. This service is severely degraded by soil compaction or impermeabilization due to rural and urban activities. Thus, as with other common resource issues, PES schemes aim to induce economic agents using land to neutralize the negative externalities they generate or even produce positive externalities.

This management instrument generally falls into three main categories based on funding agents and the type of economic incentive involved [230], [236]. In terms of funding, the first category is **user-funded PES schemes**, involving direct beneficiaries of watershed natural services. These users can be individuals, companies, NGOs, or public actors who directly benefit from the conservation, improvement, or reestablishment of natural services. A typical example is incentives paid by hydroelectric and sanitation companies to upstream landowners, making the companies intermediaries between end consumers and the initial producers of natural services [237]. The second category, **third-party-funded PES schemes**, involves economic agents acting on behalf of beneficiaries. Here, the buyer is a public or private entity, such as conservation organizations or government agencies, which do not directly use natural services. Governmental programs aimed at reducing deforestation and encouraging reforestation, particularly in China, exemplify this model. Finally, **regulation-funded PES schemes** include incentives paid by economic agents required to meet regulatory obligations, compensating for externalities in other locations. Such schemes, which create demand for offsets in a cap-and-trade format, fostering an **environmental credits market**, necessitate institutions to centralize and manage these transactions in a bank. In terms of economic incentives, the categories are divided into **direct economic incentives** and **indirect economic incentives**. The first involves direct payments to economic agents for conservation leasing, while the second involves **cost-sharing assistance** for establishing green infrastructure, including rural extension activities, as well as **ecological certification**, a label that adds market value to the agents' production. These classifications are not rigid, and combinations and overlaps between approaches are possible. In water PES schemes, regardless of the primary funding and incentive format, experience shows that PES schemes require a comprehensive socio-environmental program with multiple partner institutions and other sustainable development goals to become effective [238]. For instance, Richards *et al.* (2015) [158] illustrate that the "Water Conservationist" PES program in Extrema (MG), Brazil, was implemented through a collaboration of multiple funding and supporting organizations, including municipalities, state governments, watershed agencies, NGOs, rural extension agencies, educational institutions, and research entities.

It is worth noting that economic instruments are not substitutes for command-and-control measures but rather supplementary to them. Thus, PES schemes aim to *enhance* the conservation of natural services beyond legally mandated obligations. In Brazil, the primary regulation on land use at the national level is the **Federal Native Vegetation Protection Law**, informally called the **New Forest Code** (despite native vegetation in Brazil encompassing other vegetation types as well) [239]. Other relevant regulations at the national scale, though not uniformly applied, include the City Statute [2], which grants municipalities regulatory power over land use through the Master Plan, and the National System of Conservation Units [240], allowing federative entities to create and regulate protected areas. When enacted in 2012 to revise previous legislation, the New Forest Code introduced cadastral tools like the **Rural Environmental Registry (CAR)** and guidelines for deforestation compensation, supporting the establishment of PES schemes. These initiatives are even more necessary given the problematic changes brought by the new legislation. According to Soares-Filho *et al.* (2014) [241], the revision granted amnesty to illegal deforesters, potentially compromising conservation on private properties, where 53% of Brazil's native vegetation is located. Additionally, Brancalion *et al.* (2016) [242] point out that relaxed requirements for environmental recovery and the maintenance of agricultural activities in protected areas weakened mandatory protections for soil and water resources, increasing the risk of natural service loss. In this context, PES schemes serve as a valuable complement to legal regulations, offering financial incentives for landowners to voluntarily commit to conservation practices.

In summary, the Federal Native Vegetation Protection Law establishes zoning rules for private properties in Brazil, resulting in the **Individual Property Plan (PIP)**. In addition to optional zones like cultivation areas, fallow land, access easements, property headquarters, etc., the PIP includes two mandatory zones: the **Legal Reserve (RL)** and **Permanent Preservation Areas (APP)**. The Legal Reserve consists of the mandatory native vegetation zone that each lot must preserve, calculated as a fraction of the total area. In other words, it defines the maximum extent of economic activities that do not directly use native vegetation, such as agriculture. This zone's proportion varies by region. In the Legal Amazon, the RL fraction must be 85% in forest areas, 35% in cerrado areas, and 20% in general fields. In the rest of the country, the RL fraction is 20%, a questionable percentage in terms of sustainability, depending on land use in the remaining 80%. From an ecological perspective, Banks-Leite *et al.* (2014) [243] indicate that, in the Atlantic Forest biome, a viable limit for conserving vertebrate communities is approximately 40% of natural habitat. From a hydrological perspective, Caldwell *et al.* (2023) [244] demonstrate that 20% forest cover could lead to a degradation in water quality up to an order of magnitude worse (10 times poorer) than in fully conserved basins. Additionally, according to the law, the RL can be economically exploited through sustainable management, and if there is insufficient vegetation on the lot, the owner can restore the area or compensate for it on another property within the same biome. Permanent Preservation Areas, on the other hand, are areas that, even without native vegetation, must be preserved to protect water resources, biodiversity, soil, and slope stability. They include springs, steep slopes, hilltops, and watercourses, among others, forming buffer zones in riparian areas with widths defined according to the size of the watercourse. Legally, APPs can count towards the Legal Reserve percentage, though they are mandatory even if this percentage is exceeded. For instance, if a lot contains more than 20% of very steep slopes, the total APP area must expand to protect these slopes.

Although PES schemes are a relevant tool to set economic limits on human activities in a watershed, they are not the only solution. A possible, more robust alternative is simply to mandate stricter land use regulations by enforcing laws. In an extreme scenario, economic agents interested in preserving watershed natural services acquire exclusive land use rights in the watershed, eliminating the need for negotiations with other agents. This scenario is largely observed in the northeastern United States, a region with high demand for water to supply major cities like New York and Boston [245]. In this region, sanitation companies, both private and government-owned, secured exclusive access to critical watershed areas through direct purchase of rural properties. In this regard, New York's case study in the Catskills and Delaware basins has received some criticism from Blanchard *et al.* (2015) [246], mainly for having become a favorable narrative for economic incentives, while in reality, it was substantially grounded in command and control and invested more in rural sanitation (gray infrastructure) than in forest restoration (green infrastructure). A similar strategy applies to zoning for protected areas, known in Brazil as **Strict Protection Conservation Units**, such as Ecological Stations or Nature Parks. In these cases, the land is exclusively managed by the State, with compensation for former owners, either at the municipal, state, or federal level. An intermediate alternative in Brazil is the creation of **Sustainable Use Conservation Units**, such as Environmental Protection Areas (APA), where private land ownership is maintained, but land use is regulated by the unit's **Management Plan**. In headwater areas, for instance, conventional agriculture practices using pesticides may be prohibited. Municipal **Master Plans** can also serve this intermediate role by regulating permissible activities in headwater areas, as in the case of the Peri Lagoon Municipal Park in Florianópolis, where existing rural properties are subject to stricter land use regulations.

Principle	Objective	Basic Scientific Guidelines	Desirable Scientific Guidelines
Metrics	Robust, efficient, and versatile methods for data acquisition.	• Should be relevant, reliable, and scale-appropriate. • Should comply with voluntary standards, certifications, and regulations. • Should reflect spatial-temporal scales, as identified in Dynamics. • Optimize the balance between precision and simplicity. • Evaluate progress (alongside Baseline and Monitoring). • Establish benchmarks (alongside Baseline and Monitoring). • Measure both absolute changes and trend changes.	• Preferably selected to allow comparisons between service types. • Evaluate how services influence each other.
Baseline	Document initial and reference conditions.	• Measure influences of interventions on services. • Measure status and trends of non-target services. • Ensure measurements are viable given resources.	• Assess the initial state of exogenous and endogenous threats to services. • Measure important factors to predict service trends.
Monitoring	Track factors necessary for management, negotiations, evaluations, and forecasts.	• Quantify benefits associated with target services. • Identify spatial-temporal scales before implementation. • Use established methods/protocols and best practices for monitoring. • Monitoring should inform decision-making. • Monitoring should detect potential changes in baseline conditions.	• Monitor non-target services that influence target services.
Dynamics	Ensure project's ability to adapt to dynamic natural and human processes.	• Identify key natural services for each service class beyond target services. • Identify spatial-temporal scales of target services. • Identify data needs, resources, and gaps. • Identify stressors and their spatial-temporal variability. • Identify and predict trends of endogenous and exogenous threats.	• Determine how functional diversity influences resilience. • Identify service production functions and sensitivities. • Determine trade-offs and synergies between services.
Multiple Services	Recognize trade-offs and synergies between services.		• Assess how intervention influences other services. • Avoid double-counting services. • Assess intervention impacts on non-target services.
Sustainability	Ensure project durability and sustainability.		• Estimate short- and long-term performance of the project or program.

Table 4.3: Principles, objectives, and scientific guidelines for PES programs — Structured by Naeem *et al.* (2015) [247]

4.5.2 Planning guidelines

Over the last two decades, both academic and technical literature, based on theoretical and practical experiences, have extensively debated the foundational structure a PES scheme in watersheds must provide to be considered a successful strategy. Building on the Millennium Ecosystem Assessment report, the report by Smith *et al.* (2006) [230], an initiative of the *International Union for Conservation of Nature* (IUCN), presented a pioneering and comprehensive conceptual and practical framework for implementing PES schemes in watersheds. Among the key contributions was the term "watershed natural services," cataloged in Table 4.1. Moreover, Smith *et al.* (2006) emphasize that, to be effective, PES schemes must be founded on the strongest evidence possible. Therefore, developed actions should yield identifiable benefits for users, with clear connections between upstream land use and downstream watershed natural services for users.

Among the guidelines from Smith *et al.* (2006), the emphasis lies on estimating the **service supply capacity**, that is, the expected changes in upstream biophysical processes and, consequently, in downstream economic value. In essence, the authors advocate for measuring both ends of the Haines-Young & Potschin cascade model – processes and value. Thus, the performance of the **target service** should be estimated using specific, consensual, and measurable process indicators, enabling clear action priorities and continuous tracking aligned with the program's goals. Along these lines, **state indicators**, which reveal the observed state of processes underlying the services, and **sustainable use indicators**, which show how far usage is above or below the economic limit of resource use, are suggested. Examples of these indicators are also listed in Table 4.1. Regarding valuation, Smith *et al.* (2006) highlight that economic valuation efforts should focus on **cost-benefit analysis** among the involved parties. Costs correspond to the economic incentives offered by the paying entity, negotiated directly with economic agents. Ideally, these costs should be lower than the estimated economic benefits, evaluated using valuation methods based on improvements to watershed natural services, thus justifying the investment in the payment scheme.

The guidelines proposed by Smith *et al.* (2006) led to a wide-ranging discussion, with various authors contributing on different aspects. In this context, Luca Tacconi's work (2012) [248] aimed to unify key concepts for this field by reviewing the defining attributes and issues of a successful PES scheme. Tacconi points to three essential principles: additionality, conditionality, and transparency. The **additionality principle** seeks to ensure that economic incentives produce benefits that would not occur *without* the PES scheme. This principle, crucial in applying hydrological models, will be further discussed in the next section. The **conditionality principle**, on the other hand, stipulates that payments in a PES scheme are directly tied to practices that ensure the provision of watershed natural services. In other words, conservation leases must be honored by the contracted parties. This requires periodic monitoring of land use that conditions the continuity of the contract. For instance, if a landowner conducts a large burn on their plot, the PES contract should be terminated, and payments should cease. While strict application of this principle may increase indirect costs and impact economic agents' voluntary motivation, its complete absence can compromise the scheme's effectiveness, risking resource waste without guaranteed environmental outcomes. Lastly, the **transparency principle** involves providing clear and reliable information to all stakeholders, which is essential for the acceptance and integrity of PES schemes. Tacconi emphasizes that low transparency in negotiations and payment determination can foster distrust and perceptions of corruption, especially when third-party funding, such as environmental offsets, is involved. Transparency is also critical for program scalability, as it relates to trust and verifiability, both key components for cooperation among stakeholders. To promote transparency, public disclosure of information regarding the technical criteria for selecting priority areas, participant selection procedures, valuation methods, and benefit distribution is recommended. A strategy gaining traction to implement this principle, illustrated by Young & Bakker (2014) [249], involves PES **calculators**. This approach uses a scoring system to evaluate various attributes of land plots that qualify for payments, determining value in an objective and transparent manner. The system allows a landowner to increase their PES value by meeting additional environmental requirements, thereby earning more points in the calculator. For example, if they adopt soil conservation practices, the rural producer should receive more incentives compared to others.

The foundational cycle of guidelines was largely solidified by Naeem *et al.* (2015) [247] in their article *Get the science right when paying for nature's services*. With contributions from dozens of technical and academic institutions with experience in hundreds of case studies, the authors propose a structured set of scientific principles and guidelines to strengthen science integration in PES schemes, aiming to increase their efficacy and ensure they provide consistent environmental and social benefits. The principles, listed in Table 4.3, include defining robust metrics, documenting baselines, continuous monitoring, dynamic management, evaluating multiple services, and adopting a long-term perspective. Each principle includes basic and desirable guidelines: for example, the metrics principle guides the selection of methods balancing precision and simplicity, while continuous monitoring is essential for tracking benefits and adapting management effectively. The dynamics principle emphasizes adaptive management, attentive to ongoing natural and human processes and identifying necessary course corrections. The multiple services principle seeks to recognize and manage trade-offs and synergies among natural services beyond the target service, avoiding double counting and assessing impacts on non-target services. The sustainability principle, crucial for ensuring project longevity, encourages short- and long-term performance assessment. Evidently, Naeem *et al.* (2015)'s principles do not replace Tacconi's approach, as they have multiple connections with additionality, conditionality, and transparency.

4.5.3 The additionality principle

As mentioned above, the additionality principle aims to ensure that economic incentives yield benefits that would not occur *without* PES. Although this may seem evident, it was not the case in many planning studies involving hydrological models, as noted by Lele's review (2009) [250]. Lele identified a lack of a shared, rigorous methodological framework across the modeling and valuation extremes, particularly regarding evaluations under alternative land-use scenarios. Such methodological weaknesses prompted Sven Wunder (2005, 2007) [251], [252] to put forth pioneering considerations on this principle. Observing additionality requires assessing natural service indicators relative to a **baseline scenario**, which serves as a comparison reference for quantifying the natural service supply capacity. This principle can support the viability of a PES scheme even under unfavorable absolute conditions. For example, even if rainfall in a basin decreases due to climate change, expanding green infrastructure may still result in a comparatively less severe future scenario. Additionality is, therefore, the *difference* between the state indicator under current conditions and the state indicator in baseline conditions:

$$\Delta \mathrm{S}_t = \mathrm{S}_t - \mathrm{S}_{t^*} \quad \forall t \tag{4.1}$$

where $\Delta \mathrm{S}_t$ is the additionality at time t; S_t is the natural service state indicator at time t; and S_{t^*} is the natural service state indicator at time t^*, the time in the baseline scenario. Estimating **potential additionality** in planning studies is the first step in justifying the implementation of a PES scheme, as well as providing the foundation for mapping and prioritizing **priority areas**. The second step, proposed by Kroeger *et al.* (2013, 2019) [253], [254], involves integrating with valuation methods to support cost-benefit analysis and investment returns. In this regard, the exploration of future scenarios conducted by Possantti & Marques (2022) [3] demonstrates that this analytical approach is complex, relying heavily on auxiliary hypotheses, especially when estimating the expansion costs of NBS and the avoided costs verified downstream (see Highlight 2.6 in Chapter 2).

In watershed PES schemes, using hydrological models becomes essential for estimating potential additionality, as it is necessary to simulate the watershed's hydro-

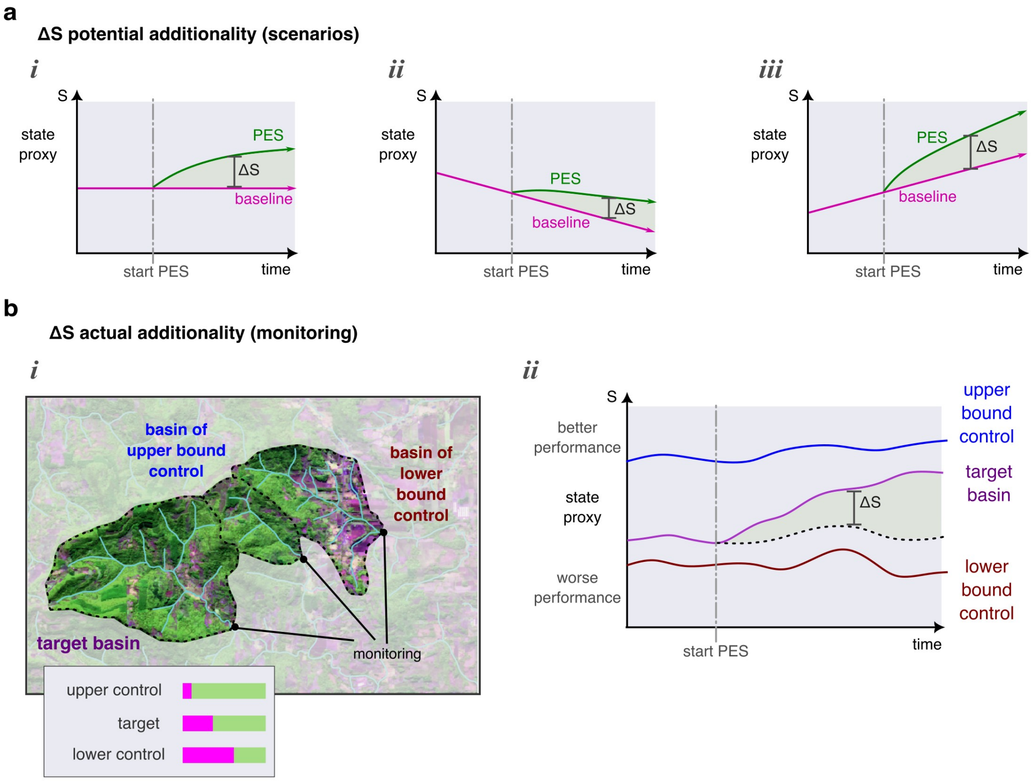

Figure 4.11 — Potential and actual additionality. The additionality principle aims to ensure that economic incentives yield benefits that would not occur without PES. **a** — In the planning phase of a PES scheme, efforts should be made to estimate potential additionality ΔS, calculated as the difference between a state indicator S in scenarios with and without PES. Scenarios can be stationary (detail i), unfavorable (detail ii), and favorable (detail iii). The scenario without PES serves as a baseline for comparison. In practice, one approach is to compare the current scenario (without PES) to a potential native vegetation scenario (100 **b** — In the operational phase of a PES scheme, actual additionality should be measured through monitoring. In watershed PES schemes, in particular, efforts should be made to isolate causality from exogenous variables. For example, paired watersheds near the target basin can be selected as controls (detail i). More degraded watersheds serve as a lower control, while better-preserved watersheds serve as an upper control (detail ii).

logical performance under alternative boundary condition scenarios, comparing results with and without planned actions (for example: Ullrich & Volk, 2009; Carvalho-Santos *et al.*, 2014; Martinez-Martinez *et al.*, 2014; Daneshi *et al.*, 2021 [255]–[258]). However, this evaluation is only feasible when applying conceptual models that explicitly represent various hydrological processes in zero-order watersheds. Data-driven models, regardless of their predictive capacity, are not applicable in such cases, as they do not permit deductive inference about scenarios. A key methodological decision, however, is defining the baseline scenario. Following a natural services restoration logic, the strategy proposed by Lima *et al.* (2017) [259] involves defining this scenario based on **potential native vegetation**, that is, a land cover scenario representing natural vegetation development without anthropic changes. In this context, potential additionality equates to **hydrological anomaly** since it refers to the hydrological disturbance caused by current anthropic activities, such as agriculture and urbanization. This approach is straightforward and evidence-based, as the posterior distribution of hydrological parameters associated with potential native vegetation is estimable in well-preserved watersheds. However, it does not provide a means to estimate additionality for other land cover types seeking good hydrological performance, like NBS, but with parameters that are not readily estimable based on empirical evidence.

The additionality principle should not, however, be limited to the planning phase; it should also be present in the continuous operation and management of a PES program. Thus, beyond potential additionality, it is essential to estimate **actual addi-**

tionality over time, enabling dynamic adaptation and course correction. Therefore, it is necessary to establish and monitor state indicators of the target service that demonstrate actual additionality. Ideally, these indicators are the hydrological variables that represent the underlying processes, such as baseflow, but can also include indirect and complementary metrics, such as land cover and land use, which can be monitored at a lower cost via remote sensing. An example of actual additionality monitoring is presented by Sone *et al.* (2019) [260], where the authors evaluated whether a watershed PES program in the Guariroba River Basin in central Brazil produced observable impacts. In the study, precipitation and flow were monitored from 2012 to 2016, during which soil and water conservation practices, such as contour terrace construction and riparian vegetation restoration, were implemented. The results, combined with trend analysis of time series, showed that while precipitation records exhibited a downward trend (1 mm per month), baseflow increased by 18 liters per second during the monitoring period. Such field observations reinforce the need to check not only target service indicators but also exogenous variables that create *extenuating circumstances*, like precipitation. If precipitation had *increased* during the monitoring period, Sone *et al.* (2019) would not have been as confident in asserting the actual additionality of the actions. In this regard, an ideal monitoring system in a watershed PES scheme should adopt a **paired watersheds** approach, where **control watersheds** where *no* actions are taken are also monitored. More degraded control watersheds thus provide a **real baseline**, while more preserved control watersheds offer a **real ceiling**. A monitoring system of this scale naturally consumes program financial resources, encouraging the exploration of less conventional strategies, such as using **bioindicators** – algae, plankton, and lotic ecosystem invertebrates. These organisms, as they reflect multiple aspects of ecological structure, can be periodically sampled, providing information on the recent history of environmental conditions. This approach allows ecosystem changes to be identified more economically and complements hydrological variables, offering an integrated view of conservation practices' impacts over time.

4.5.4 Applications with the `PLANS` model

The `PLANS` model, developed collaboratively with colleagues (see Section 3.5.3), emerges as an alternative to the widely-used `SWAT` and `InVEST` models for modeling watershed natural services (Francesconi *et al.*, 2016 [261]; Possantti *et al.*, 2023 [4]). The `SWAT` model simulates multiple hydrological processes by considering hydrological response units, drainage networks, and sub-watersheds [262]. In contrast, `InVEST`, particularly in its water balance module, utilizes high spatial resolution maps, though processes are simplified to an annual scale [258]. While `SWAT` provides greater conceptual detail of processes, `InVEST` offers finer spatial resolution [263]. The `PLANS` model combines these attributes: conceptual detail and high spatial resolution, achieved by expanding the classic `TOPMODEL` version, which defines hydrological response units based on topographical attributes, lending detailed spatial nuances to the model.

Regarding potential additionality mapping, illustrated in Figure 4.12, PLANS land-cover scenario simulations in the Arroio Castelhano watershed (383 km^2) in Venâncio Aires estimated an average surface runoff reduction capacity of around 20% (200 mm/year) and an increase in baseflow of approximately 25% (80 mm/year), meaning roughly a thousand liters per second of cleaner, relatively perennial water in the watercourse. Annualized maps show hydrological anomaly patterns, especially in anthropized areas where land cover has a greater influence. In the mountainous headwaters and more preserved Atlantic Forest areas, the anomaly is lower or even nonexistent. Among anthropized areas, one sandy-soil region demonstrated greater sensitivity, as surface runoff was nearly absent in the reference conditions due to high soil hydraulic conductivity.

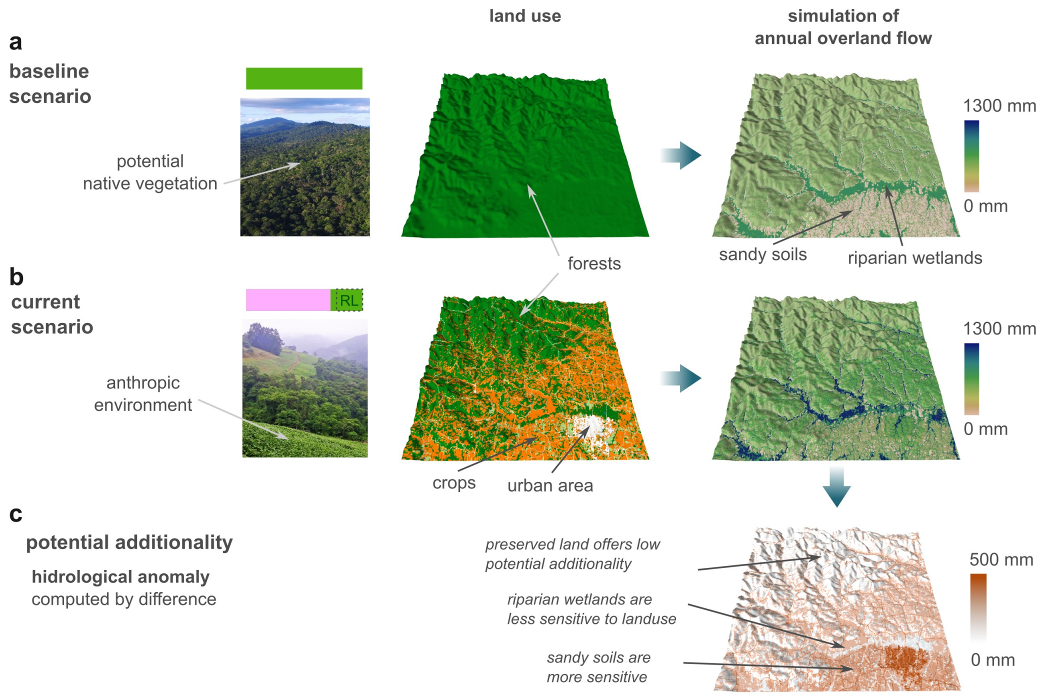

Figure 4.12 — Estimating potential additionality with the PLANS model. With the PLANS model, it is possible to estimate potential additionality by simulating watershed hydrology under alternative land-use scenarios. Based on TOPMODEL, it considers not only soils and cover but also the dynamic influence of topography and riparian wetlands, which distinguishes it from other models. This example illustrates the results for the Arroio Castelhano watershed in Venâncio Aires, RS. Daily surface runoff maps from 10 years of simulations were annualized and compared. **a** — Baseline scenario configured as potential native vegetation. In this uniform land-cover scenario, areas with sandy soils produce naturally less surface runoff due to hydraulic conductivity, while riparian wetlands naturally produce more runoff due to topography. **b** — Current land-cover scenario, with mosaics of cities, roads, and croplands, highlights increased surface runoff across the entire basin, primarily due to reduced infiltration capacity. **c** — The difference between scenarios then demonstrates the sensitivity of each landscape area, showing that already preserved areas offer no additionality. Likewise, riparian wetlands offer little additionality. Cropland areas with sandy soils, however, are far more sensitive, attracting action prioritization.

Additionally, the model showed that in thalweg and riparian wetland areas, where topography is more influential, hydrological anomalies are relatively lower. In these locations, surface runoff generation is less dependent on land cover and use, whether natural or anthropized. This sensitivity is not captured in hydrological models like `SWAT` and `InVEST`, which do not represent the variable contributing area phenomenon.

On the spatial resolution side, the `PLANS` model enables potential additionality assessment within each rural lot, i.e., at the operational scale suitable for differentiating among lots. This advantage results from representing topographical influence, which allows the precise spatial reconstruction of simulated hydrological process maps. Therefore, rural lots with the same soil type and cover class, but located in different parts of the landscape (e.g., hillside and floodplain), may exhibit distinct potential additionalities, depending on the predominant runoff generation mechanisms. It is worth noting that rural lot boundaries, available from the Rural Environmental Registry, are an essential data source for developing Individual Property Plans (PIP), applicable in both PES programs and command-and-control contexts, such as enforcing Forest Code regulations. Integrating with this registry in potential additionality modeling is thus a notable advantage of the `PLANS` model for mapping priority areas and objectively scoring candidate lots in PES calculators, also observing the transparency principle. Detailed maps showing rural lots and their various environmental attributes, including hydrological ones, strengthen trust among all parties involved.

Finally, applying the `PLANS` model in the Arroio Castelhano watershed incor-

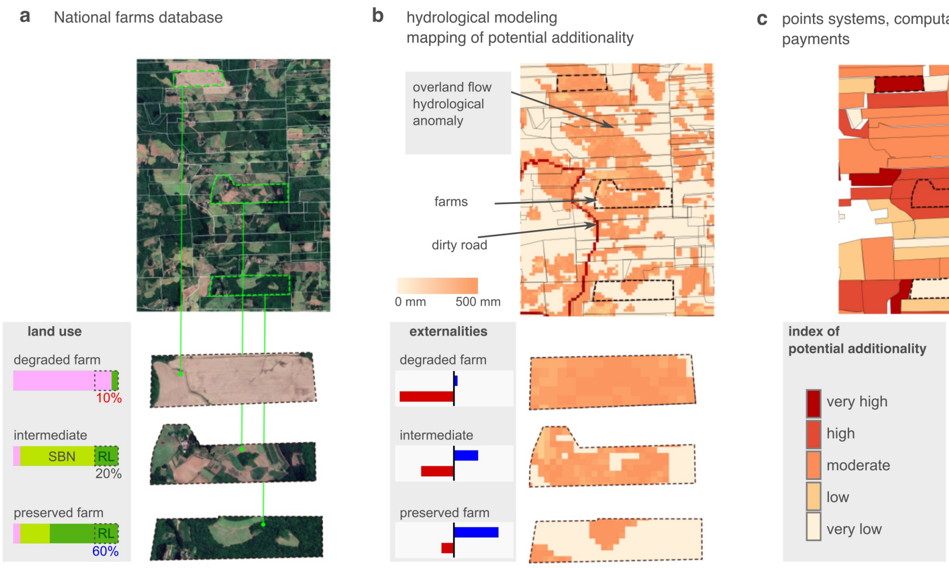

Figure 4.13 — Evaluating potential additionality at the operational scale of lots. Another unique aspect of the PLANS model is its capacity to produce high spatial resolution results, facilitating potential additionality estimates at the operational scale of rural lots. This advantage stems primarily from its organization into hydrological response units more detailed than conventional models. **a** — The Rural Environmental Registry provides the territorial division of rural lots, including the individual property plan. Combined with high-resolution remote sensing, it is possible to discern more or less degraded lots. **b** — PLANS model results help estimate each lot's potential additionality, providing insight into the magnitude of each lot's negative externality for downstream. **c** — Each lot's potential additionality can then be aggregated into priority indices, aiding planning actions in PES schemes.

porated the inherent uncertainties of hydrological modeling (detailed in Chapter 1; see Highlight 1.6). To do this, the priority index construction included weighting by the dispersion of empirically adequate models generated by the GLUE method, in combination with a search algorithm (see Highlight 2.6 for more information). Thus, priority lots are defined as those with the highest potential additionality and the lowest associated uncertainty. This approach also opens new paths, such as trade-off and multiple-target service evaluations, as recommended by Naeem *et al.* (2015). For example, Figure 4.14 also shows an assessment between potential hydrological and ecological additionality. Generally, these two indicators do not coincide, forming a **pareto frontier** between services. This occurs because, in ecological terms, potential additionality is higher where native ecosystem fragmentation is lower, favoring priority actions on *better-preserved* lots than degraded ones, as in the hydrological potential case. With these contributions, the `PLANS` model seeks to integrate hydrological modeling from its most philosophical foundations to its most operational consequences, providing a robust and scientifically rigorous methodological framework to guide the transition to sustainable development in the coming decades. ■

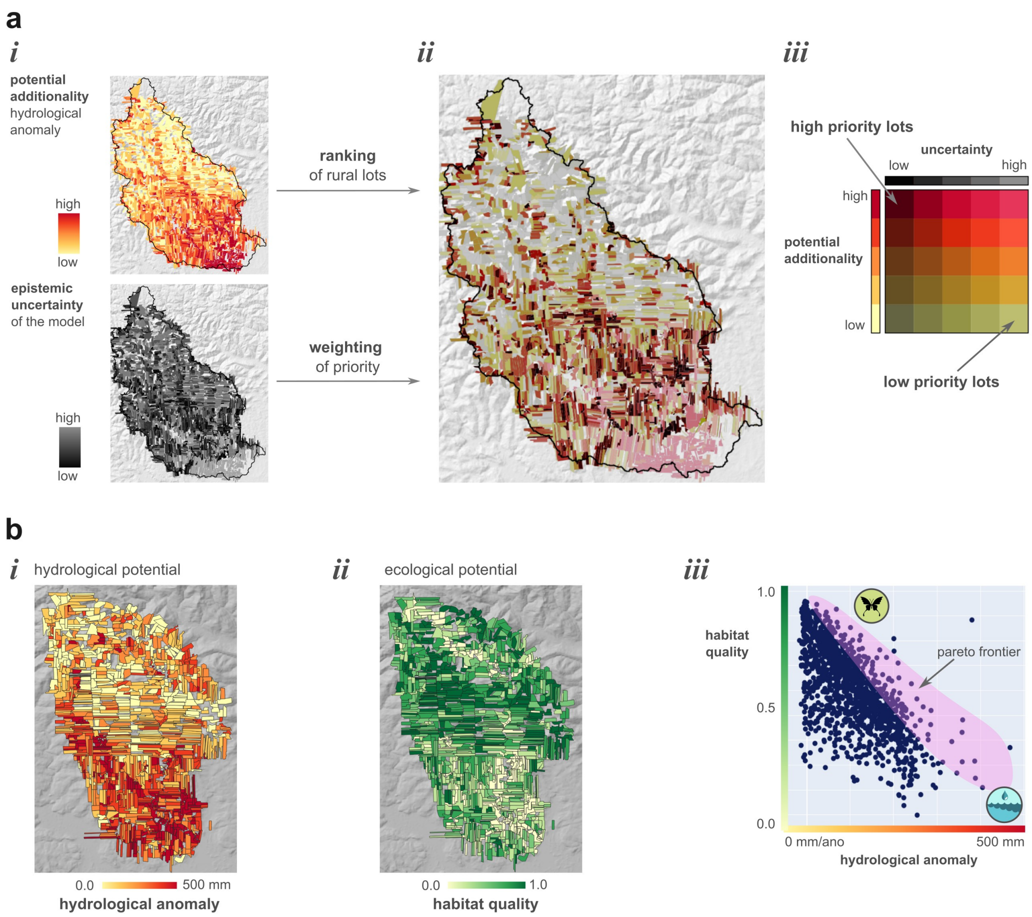

Figure 4.14 — Priority index and trade-off assessment. Applications with the PLANS model open pathways for incorporating uncertainties and evaluating trade-offs. **a** — A priority index for rural lots based on models like PLANS should consider the spatial uncertainties of simulations (detail *i*). Spatial uncertainty bands, thus, act as weighting factors for ranking, producing a priority index (detail *ii*) that reflects both performance attractiveness (higher potential additionality) and estimation reliability (lower uncertainty, detail *iii*). **b** — Results from ecological models like the Habitat Quality model in `InVEST` can be integrated into a comprehensive analysis that considers synergies and trade-offs. In the Arroio Grande watershed example (Venâncio Aires), potential hydrological additionality emerged in lower watershed areas, where rural lots are more degraded. On the other hand, habitat quality offers ecological additionality potential, as it is more worthwhile to restore ecosystem fragments where preservation is already higher. This dichotomy in priorities results in a trade-off analysis where the best lots are located on a pareto frontier.

4.6 Chapter summary

In this chapter, I address the application of ecological economic principles to watershed management, advocating for a paradigm shift that sees watersheds not merely as exploitable resources but as integral components of natural capital with inherent limits. The discussion centers on the sustainable balance of resource use, emphasizing the essential role of natural services, especially watershed services, in supporting both human well-being. By adopting economic valuation methods and promoting the use of land use economic instruments like Payment for Ecosystem Services (PES), this chapter aims to outline strategies for enhancing watershed management while highlighting the importance of scientific rigor in environmental planning.

- **Ecology implies a finite world.** The ecological economics model is introduced as a departure from traditional economic paradigms, emphasizing that nature should not only be viewed as a resource to exploit but as natural capital with inherent limits. Unlike neoclassical economics, which often neglects these boundaries, ecological economics places ecological sustainability at the forefront, encouraging strategies that maintain ecosystem functions while supporting human well-being. This approach includes the recognition of externalities, especially in watershed contexts where upstream land users impact downstream water users.

- **Natural capital implies an optimal scale.** Natural capital, such as forests in watersheds, is positioned as essential for providing natural services that support human life. The chapter argues that there is an optimal scale of resource use that ensures the regeneration of natural capital. When applied to watersheds, this concept urges a balance between water use for agriculture, urban needs, and conservation. Overstepping these natural limits risks degrading the ecosystem's ability to provide services like clean water, flood control, and sediment retention.

- **Escaping the tragedy.** Common goods, as open-access resources, are vulnerable to overuse and degradation—a dynamic described as *the tragedy of the commons.* In a watershed, upstream practices (e.g., deforestation, pollution) can create negative externalities for downstream users, leading to conflicts and degradation. Economic instruments like watershed PES programs can help advance conservation in upstream areas beyond the mandatory law, thereby protecting water quality and availability for downstream communities.

- **Natural accounting.** Using frameworks like the Common International Classification of Ecosystem Services (CICES) and the Economics of Ecosystems and Biodiversity (TEEB) allows for the systematic categorization and valuation of natural services. Watershed services are broken down into provisioning, regulating, and cultural sections, each with distinct utilitarian values. Valuation methods are emphasized as tools to demonstrate the economic importance of watershed services, providing stakeholders with information for decision-making and fostering public and private support for conservation efforts.

- **The application of scientific guidelines.** Among others, the principle of additionality is central to PES programs, ensuring that incentives lead to environmental benefits beyond what would occur without intervention. By comparing scenarios with and without conservation practices, potential additionality quantifies the impact of PES in the planning phase, ensuring transparency and effectiveness in achieving conservation goals. Explicit and measurable metrics allow the monitoring of processes to infer the actual additionality during the operational phase.

- **Advantages of the PLANS model.** The PLANS model, based on TOPMODEL, is highlighted for its capability to simulate watershed processes with spatial detail, incorporating topographic, soil, and land cover data. This high-resolution approach allows PLANS to assess additionality at the plot level, aiding in prioritizing areas for conservation. Unlike other models, PLANS integrates both conceptual detail and spatial accuracy, providing nuanced insights into the impacts of land-use changes, evaluating trade-offs, and assessing uncertainty, enabling better-targeted conservation actions within PES programs.

Glossary

abiotic resources Inorganic resources from nature, such as water and minerals. 125

abstraction An idealization process that seeks to remove supposedly irrelevant factors and aspects of the target system, leaving only its essence. 36, 44

activation function A formal hypothesis about how the activation of a reservoir develops as saturation occurs. 105, 107

activation level Equivalent to the activation threshold. 105, 107

activation threshold The minimum level required for the connectivity of a reservoir's drainage network (channels or pores) to begin transmitting water. 106

actual flow The flow that actually influences the level of a compartment in the simulation of a dynamic system. 51

Adaptation Level Theory A theory suggesting that constant stimuli are mentally adjusted, becoming normal and requiring new variations to generate the same sensation of pleasure or displeasure. 114

agent-based models Agent-based models are computational simulation systems that use autonomous entities, with individual behaviors and interactions,. 43, 44, 117

altruistic value Utilitarian value reflecting concern for the well-being of other members of the current generation. 135

analog models Representations based on an analogy with the target system, preferably involving systems with supposedly the same mathematical structure (formal analogy). 63

analogical inference Non-deductive and non-inductive reasoning that concludes that a given object O_1 has the property P_1 of object O_2 because they share other properties. 36, 38, 63

analogy A comparison between two or more objects, emphasizing their supposedly similar aspects. Analogy is used in modeling to idealize systems in terms of other, more tangible systems that supposedly have the same mathematical structure. 7, 25, 35–39, 69, 123

antecedent moisture conditions The set of initial conditions of the system on the eve of a rainfall event. 72

anthropocentrism The view that places humans at the center of all decisions and objectives, focusing on human needs and interests as fundamental in any analysis or action. 113, 135

anthropogenic capital Material and social structures built by humans to produce goods and services that generate well-being. 123, 124, 126, 127

anthroposphere The part of the biosphere encompassing the built environment inhabited by humans, where natural resources are transformed to meet societal needs. 119, 121, 123, 124, 126, 128, 131, 133

aquifer detention time g A characteristic parameter of the aquifer, assuming the phreatic zone behaves as a linear reservoir. 72, 76

Aristotelian idealization (See abstraction). A method of idealization that uses abstraction, a process aimed at removing supposedly irrelevant factors and aspects of the target system, leaving only its essence. 36

attractors A set of final, stable, or unstable behaviors observed in the solution of a given system of differential equations, depending on parameter values and initial conditions. 41, 43

auxiliary equations Equations used in the programming of dynamic systems (procedural model) to break down into steps that are easier for humans to understand. 47

auxiliary hypotheses A set of hypotheses necessary in addition to the main hypothesis of a model. 26, 29, 34, 75, 95–100, 107, 117

avoided cost method A valuation method that considers the costs avoided by the presence of a natural service, such as flood protection. 137

balance equation A differential equation that establishes that the variation in a level results from the net effect of the inflow and outflow rates. 45, 51

balancing loop Feedback that acts on both inflows and outflows, reducing the value of this flow, leading to a situation that tends towards a state of equilibrium with or without oscillations. 47

baseflow $Q_{\mathbf{g}}$ The outflow or drainage of the phreatic zone, usually considered a slow response of the watershed. 80, 82

Bayes' theorem Mathematical formulation for determining the posterior probability: $P(H|E) = P(H) \cdot P(E|H)/P(E)$, meaning Posterior = Prior × Likelihood ÷ Evidence. 10

bedrock topography Irregularities in the bedrock that establish the relatively impermeable bottom of the unconfined aquifer in the phreatic zone. These irregularities can create pockets or stagnant zones in the aquifer. 85, 96, 106

biocentrism An ethical approach that recognizes the intrinsic value of all forms of life, advocating for the right of species to exist. 135

biome-specific natural services Natural services with specific characteristics of different biomes, such as tropical forests or oceans. 131

biophysical structure Physical components of ecosystems that provide the basis for the provision of natural services, such as forests and wetlands. 131

biophysical value Non-utilitarian value of a natural resource calculated based on physical properties, such as energy or materials needed for its production. 135

biosphere The layer of Earth that encompasses all ecosystems and living beings, exchanging energy and matter with the physical environment and fundamental to ecological economics. 119–121, 123, 124, 126, 128, 131, 133

biotic resources Biological resources that depend on living processes for their formation and renewal, such as plants and animals. 125

bracketing inequality Inequality used to test whether a model is empirically acceptable. The model's simulated results must be bracketed by the observational uncertainty bands (total observational error) with a predefined level of confidence (rejection criterion). 28, 29

calibration process A procedure to adjust model parameters to increase its degree of confirmation against empirical evidence. 27, 88, 89

capillary deficit $\mathbf{D_v}$ The potential storage capacity of capillary water in the vadose zone. 72, 73

capillary fringe A region in the vadose zone with zero capillary deficit due to its proximity to the phreatic zone, from which water is suctioned. 81, 82

capillary water $\mathbf{V_c}$ Water stored in the vadose zone, held by the cohesive forces (surface tension) of soil particles. This water is more accessible for plant transpiration than for recharge to the phreatic zone. 71, 78, 85, 86

cascade model of natural services A model proposed to describe the flow of natural services from nature to their value for society, passing through stages such as function, service, benefit, and value. 130, 131

catastrophic limit The point at which environmental degradation leads to nearly irreversible impacts, potentially triggering an ecological system collapse. 124

categorical imperative A Kantian moral principle guiding ethical decisions, advocating actions to be taken as if they were universal laws. 135

causal loop diagram In systems dynamics, a causal loop diagram is a visual tool that represents cause-and-effect relationships between the variables of a system, highlighting how changes in one variable influence others through reinforcing and balancing loops. 45

causal structure In systems dynamics, the causal structure refers to the set of cause-and-effect relationships that determine a system's behavior over time, including feedback loops and flows between the system's compartments. 45

central limit theorem A theorem that establishes the mathematical fact that the sample mean of any population exhibits a normal distribution. Regardless of the population's distribution (uniform, normal, etc.), the mean obtained from samples will be normally distributed. This happens because the mean is a sum, and in sums of random numbers, low sampled values tend to offset high sampled values, resulting in a bell-shaped pattern similar to a normal distribution. 14

choice modeling valuation A method that models hypothetical choices to reveal preferences for different aspects of a natural resource. 137

circular flow of exchange value A neoclassical model representing the continuous circulation of goods and services between firms and households, sustained by the exchange of monetary value, with the objective of maximizing utility. 117, 118

climate regulation natural service An ecosystem service that contributes to climate regulation through processes such as carbon sequestration and plant transpiration. 131

command and control instruments Regulations imposed by the State to manage the use of natural resources, such as emission standards and zoning. 129

commensurability error Epistemic uncertainty resulting from the difference in scales of time and space between observed processes and modeled processes. 27, 89, 94

commodities Goods and services with exchange value in markets, whose demand and supply are determined by economic factors such as marginal utility and scarcity. 118

common resources Open-access resources, such as air and oceans, that everyone can use but require management to prevent overuse. 127–129

compartment model In systems dynamics, a compartment model is a modeling technique that divides a system into different sectors, where each compartment represents an accumulated quantity (level) of a specific variable, and the rates of flow between these compartments describe the changes over time. xii, 44, 48

compensating effects Internal discrepancies in hydrological models caused by mass balance constraints in model compartments. These effects can lead to unrealistic results, as mass balances can mask underlying physical processes, compromising the accuracy of simulations. 88

computational grid A spatial discretization structure used in numerical methods to simulate hydrological processes. It consists of dividing the spatial domain into elements or cells, which can be regular or irregular, facilitating the application of techniques such as the finite difference method or finite element method to solve the equations that describe physical processes. 91, 92, 94, 96

conceptual model A formalized and simplified representation of the processes identified in the perceptual model. This model involves creating hypotheses and adopting assumptions to abstract the complex processes of reality in a tangible and objective way, often using mathematical formulations. 34, 35, 49, 50, 53, 55, 58, 63, 69, 73, 75

conceptual scale A level of analysis that serves as a bridge between the natural scale of hydrological processes and the observational scale of empirical evidence, representing the processes and interactions in the model in a structured way. 93, 94, 107

conditioning [synonym of conditionalization] Application of Bayes' Theorem to a hypothesis to update its degree of conviction. It can be done in successive steps as new evidence is obtained. xii, xiv, 11–16, 27, 28, 31

congested output problem The difficulty of applying balance equations using the Euler method when outflows feed into other compartments that may eventually become saturated. One solution is to compute both the maximum and potential flows before defining the actual flow. 51

congestible resources Resources that allow simultaneous use up to a certain point, such as roads or beaches, where excessive use leads to congestion. 128

Connectivity Theory A unifying theory of Hydrology proposed by Jeffrey McDonnell and colleagues to overcome the problems of the differentiation paradigm and address the old water paradox. 105–107, 109

conservation principle A principle used in systems dynamics to apply balance equations to compartments, often referring to the conservation of mass or energy (in physical systems). 45

consumptive Use of resources that results in their irreversible transformation or consumption, as in fuel burning. 136, 137

context of discovery A perspective in the philosophy of science that deals with the problem of understanding the historical change of theories. 22, 24

context of justification A perspective in the philosophy of science that deals with the problem of justifying the truth of theories. 22, 24, 37

contingent valuation method A method that uses surveys to obtain individuals' willingness to pay for the conservation of a natural resource. 137

convergent slopes Slopes where surface and subsurface drainage converge toward a single region, generally producing riparian wetland areas. 82, 91

cost-based valuation A method that assesses the utilitarian value of a service or resource based on the costs required for its replacement or maintenance. 137

credence Measure of belief. Central concept in Bayesian epistemology derived from the idea that knowledge is not a matter of all-or-nothing, but exhibits subtleties between true and false. This concept can be considered a probability under certain circumstances. 9–11, 13, 16, 20, 22, 31

critical rationalism A rationalist philosophical approach proposed by Karl Popper, which establishes falsifiability as the criterion for demarcating scientific theories. In this view, the power of empirical evidence lies in justifying, through deductive logic, the falsity of theories (never their truth). For instance, it takes just one black swan to disprove the theory that all swans are white. While unrefuted, theories are merely corroborated by evidence. 18, 19

cultural natural services Intangible benefits provided by nature, such as recreation and cultural inspiration. 126, 131, 133, 136, 137

Curve Number Method (*Curve Number* (`CN`)) An empirical method for estimating the water balance in watersheds developed by the Soil Conservation Service in the 1950s. 73

Darcy's Law A principle that describes water flow through porous media. It establishes that the flow is directly proportional to the cross-sectional area of the conduit and the difference in hydrostatic potential, and inversely proportional to the length of the conduit. Mathematically, it is expressed as $\mathbf{u} = -K\nabla\Phi$, where $\mathbf{u}$ is the Darcy velocity, K is the hydraulic conductivity, and $\nabla\Phi$ is the hydrostatic potential gradient. 90, 96, 97

data-driven models Hydrological models that focus on observational data analysis and forecasting, using techniques such as artificial neural networks. They aim to maximize predictive capability (predictive capability), but may compromise the explanatory capability (explanatory capability) of hydrological processes, often treating the watershed as a black box. 87, 88

deductive inference Logical reasoning that establishes the truth of a given statement from antecedent premises. The truth of the consequent statement is guaranteed only if the antecedent premises are also true. 8, 31, 88, 106, 173

demarcation problem The difficulty of establishing the difference between a scientific theory and a merely metaphysical theory, which relies solely on pure abstractions. 19

depreciation The deterioration of resources or anthropogenic capital due to use and the passage of time. 126

deterministic chaos Extreme sensitivity produced by nonlinearities in dynamic systems, generally associated with initial conditions. A chaotic system evolves in a highly unstable manner, oscillating between various final states. Rounding errors can amplify this effect even more, though the origin of the process lies in the mathematical formulation itself. 43, 63

Differentiation Age The period from 1970 to the present in which the scientific community in Hydrology seeks to establish how different environments exhibit distinct response mechanisms, depending on climatic, topographic, and land use conditions. 76

digital elevation model (DEM) A digital representation of the Earth's surface that captures the topography of a specific area, used in geoprocessing and hydrological modeling to derive characteristics such as slope, drainage area, and topographic wetness indices. 97

dimensionality problem The difficulty of exploring high-dimensional parametric spaces, requiring exorbitant computational resources to execute simulations in a reasonable time frame. 58, 100

direct market valuation A method using actual market prices to estimate the utilitarian value of natural resources and services. 136, 137

direct rainfall $q_{\mathbf{se}}$ The surface flow produced due to rainfall on saturated soil. It can be generalized but occurs more frequently in riparian wetland areas. Common on convergent slopes, forcing the water table to surface. Considered a fast response. 75, 79, 80, 82, 87

direct use value Utilitarian value obtained from the direct use of a resource, such as timber extraction or park visitation. 136

distributed models Hydrological models that represent hydrological processes on a detailed spatial grid, allowing for the simulation of local variations in parameters and flows. These models can capture spatial heterogeneity and provide more accurate results at local scales, though they generally require greater computational power. 92

distribution function A method or function used in hydrological modeling to distribute or adjust parameters or variables from a larger scale to a smaller scale, or vice versa. This function helps transfer information between different levels of spatial detail, ensuring that local characteristics are adequately represented in the model. 95, 97

divergent slopes Slopes where surface and subsurface drainage spread out in different directions, preventing the formation of riparian wetland areas. 80, 82

downscaling The process of transferring information from larger scales to smaller scales in hydrological modeling, which generally involves non-trivial methods and auxiliary hypotheses that consider the heterogeneity of hydrological processes across scales, including the use of co-variables (indicators). 95, 96, 100, 101, 104

Ecological Economics An economic approach that integrates ecological and thermodynamic considerations, emphasizing sustainability and criticizing unlimited economic growth on a finite planet. 115, 118, 121, 123, 124, 128, 134, 135

ecological footprint A measure of environmental impact that calculates the amount of natural resources consumed and the space needed to absorb waste. 135

ecological processes Natural processes that sustain environmental functions, such as nutrient cycling and water infiltration. 131

Ecology The science studying the relationships between organisms and their environment, analyzing population dynamics, ecosystems, and biophysical processes, essential in ecological economics for assessing environmental impacts. 118

economic benefit Monetary or well-being benefits derived from natural services that contribute to human quality of life. 125, 129, 131

economic growth The expansion of value added to goods and services in an economy. In the neoclassical view, it implies an increase in utility and well-being; in ecological economics, it is limited by finite natural resources and their physical laws. 115, 118, 121

economic instruments Public policy tools that use financial incentives to influence conservation behaviors and sustainable use. 129

economic limit The point at which the incremental benefit of expanding the anthroposphere equals the incremental sacrifice of natural capital, beyond which expansion becomes uneconomic. 123, 124, 126

economic value The monetary assessment of a natural service or resource, reflecting its utility to society and its importance in economic decision-making. 131, 134

Economics The science that studies the allocation of scarce resources among different objectives, implying choices and decisions regarding the use of these resources. Example: directing water for agriculture or energy generation. 112, 113, 115, 118, 119, 121, 135

ecosystem service Benefits that ecosystems provide for human well-being, encompassing provisioning, regulating, supporting, and cultural services. 132

effective observational error The overlap of measurement error and commensurability error, representing the uncertainty that a model must be subjected to in order to assess its empirical adequacy. 27, 28

effective rainfall p_s The flow of rainfall that reaches the soil surface after the vegetation canopy is saturated. 71–74, 81, 95–97

empirical uncertainty The scientific component of uncertainties in the decision-making process based on evidence. It consists of both epistemic and statistical uncertainties about the state of the world. Other non-empirical components include ethical and political uncertainty. 6

empirically acceptable model A model that yields simulated results that satisfy empirical observations with a pre-established level of confidence. 28

empirically equivalent models Different models that yield simulated results with no significant deviations given the total observational error. In this case, there is no empirical reason to favor one model over another, at least concerning the main hypothesis of the models. 27, 28

empiricism Philosophical doctrine that argues that all knowledge originates from empirical experience, that is, observations of the external world. This doctrine opposes rationalism. 8, 9, 17, 18, 25

engineering bias A formative trend in Hydrology that seeks knowledge aimed at solving practical societal problems. 66, 75

environmental functions The capacities of an ecosystem to perform activities that can be useful to humans, such as flood regulation. 131

environmental services Human or natural activities that support the maintenance and enhancement of ecosystem services. 132, 133

ephemeral springs Points and patches in the landscape where small perched aquifers emerge, formed in more superficial organic soil horizons, contributing to exfiltration in the fast response of rivers (rises). 74, 76, 77, 93

epistemic uncertainty A general concept referring to the various non-statistical uncertainties present in the modeling process. Unlike statistical uncertainty, which refers to the available information, epistemic uncertainty is associated with unavailable information. 26, 27, 136

equifinality problem A term coined by Keith Beven for the mild version of the underdetermination problem in the case of environmental numerical models. The underdetermination of models occurs because the information about the modeled processes is incomplete, ensuring the existence of model structures that are empirically equivalent, or equifinal. 26, 31, 55, 88

Ethics A branch of Philosophy that investigates the moral principles and values guiding human behavior, addressing what is considered right, just, and morally appropriate. 113, 114

Euler method A simple numerical integration technique used to solve ordinary differential equations, where the solution is approximated by advancing in small steps, using the known derivative to estimate the value of the function at the next point from the current value. 46, 52, 53, 63

evidence-based policies A concept in public policies that seeks total or partial support from objective evidence to guide decision-making and resource allocation. 6, 101

excess rainfall p_x The excess rainfall that exceeds the soil's infiltration capacity. 71, 72, 74, 81

exchange value The value attributed to a good or service in a market context, reflecting what it can be traded for in monetary terms or for other goods. 116–119, 121, 123, 128, 131

exclusive resources Resources whose use is guaranteed for a single user through ownership or concession. 128

exclusivity A characteristic that allows only one economic agent to use the resource at a given time. 128, 129

exfiltration q_{ss} Flow from ephemeral springs located at the base of slopes, resulting from the rapid passage of water through organic horizons with large amounts of macropores. Common in forests. Considered a fast response. 75, 77–79, 82, 84, 85, 91

exhaustive sampling Also known as brute force, it is a strategy for sampling the parametric space by enumerating all possibilities after uniform discretization. 58

existence value Utilitarian satisfaction derived simply from knowing that a species or ecosystem continues to exist. 135

existential imperative The fundamental goal of living beings to maintain their own existence, involving decisions to ensure survival conditions such as obtaining food and shelter, regardless of ethical considerations. 113

exogenous variables In systems dynamics, exogenous variables are factors external to the modeled system that influence its behavior but are not affected by the system's internal dynamics. They are imposed from outside (external forces) and remain constant or follow a predetermined pattern during the simulation. 45, 48, 49, 53, 56

explanatory capability The ability of hydrological models to provide an understanding of the physical processes that generate runoff. This includes the ability to describe how and why observed hydrological phenomena occur, beyond merely predicting outcomes. 87, 88, 117, 119, 159

exploratory models Models used in scientific research as tools to investigate and develop new hypotheses, especially useful in areas where established theories are insufficient or non-existent, allowing the exploration of theoretical possibilities and potential explanations. 38, 93

exponential decay curve A graph that represents a rapid drop in a compartment's level over time, usually resulting from the dominance of outflows over inflows, associated with the presence of reinforcing loops (positive feedback) on these flows. 47, 72

exponential growth curve A graph that represents the accelerated increase in a compartment's level over time, usually resulting from the dominance of inflows over outflows, associated with the presence of reinforcing loops (positive feedback) on these flows. 47

externality The impact of an economic action on the well-being of others, without that impact being reflected in the agent's cost or benefit. 127–129

falsifiability The ability of a theory to be shown as false through empirical experience (observations and experiments). A falsifiable theory is not necessarily false but *can* be proven false by empirical evidence. In critical rationalism, this ability is the criterion for determining whether a theory is scientific. 7, 19, 20

feedback A recursive information loop that acts on a system. It can be positive, reinforcing a given process, or negative, stabilizing a given process. 22, 39, 41, 45–48, 51, 63

field capacity $v_{\mathbf{max}}$ The maximum storage level of capillary water in the vadose zone. 71, 82

finite difference method A numerical technique used to solve partial differential equations, employed in the simulation of transient flows in porous media. This method discretizes the spatial domain into a regular grid, allowing the approximation of the derivatives involved in the physical equations. 91

finite element method A numerical technique used to solve partial differential equations in complex domains, employing an irregular computational grid. It allows for more flexible representation of the domain's geometry and is particularly useful in simulating transient flows in porous media with varied geometries. 91

first law of thermodynamics The principle of conservation of energy and matter, stating that neither can be created nor destroyed, only transformed. This implies the impossibility of infinite material growth. 119, 121

fluvialist bias A hegemonic trend in Hydrology that focuses on essentially hydraulic problems at the watershed scale, such as flow propagation in river channels and flooding of adjacent plains. 67, 68, 75, 87

fragmentation level A parameter that regulates the speed of activation in the saturation equation. The higher the fragmentation, the more dampened the reservoir's activation becomes. The value of the fragmentation level corresponds to the reservoir value at half the maximum activation speed. 107, 108

free will problem A philosophical challenge questioning the existence of genuine choices in a deterministic universe; in physicalism, human decisions would be predetermined, eliminating the idea of agency. 119

fund-service resources Resources that provide utility without being materially consumed, such as fertile soil, infrastructure, and energy. 125, 126, 128

futility limit The point at which the expansion of anthropogenic capital generates zero or negative marginal utility, becoming directly harmful to well-being. 124

future use value Utilitarian value associated with the potential future use of a natural resource, such as biodiversity for future scientific discoveries. 136

Galilean idealization A method of idealization that applies controlled distortions that could be incrementally removed to asymptotically reach the final behavior of the target system. 36

general natural services Natural services common to different ecosystems, such as climate regulation and soil formation. 131

geochemical signature The concentration of solutes that allows identifying or tracing the origin of water from its diffusion process in the soil and rocks. 85

global scale A level of analysis that assesses the aggregated and integrated behavior of the hydrological system as a whole, considering all its parts and interactions from a macroscopic perspective. 90, 92, 94, 97, 102

gravitational deficit D The potential storage capacity of gravitational water in the vadose zone or the effective depth of the water table. 71, 96, 97

gravitational water $\mathbf{V_g}$ Water stored in the vadose zone that is free to percolate vertically to the phreatic zone under the influence of gravity (recharge). 71, 78, 85, 86

green infrastructure Urban vegetation systems, such as parks and wooded areas, that offer environmental and social benefits. 133

group valuation A method involving discussions and consensus among groups to assess the value of a natural service. 137

hard problem of consciousness The philosophical difficulty of explaining subjective experience (consciousness) solely through physical and synaptic processes, one of the main challenges to physicalism. 115, 119

hedonic calculus A method proposed by Bentham to morally evaluate an action by considering factors such as the intensity and duration of the pleasure or pain it may cause, aiming to quantify the well-being generated. 114

hedonic pricing method A method that uses variations in the prices of goods, such as real estate, to infer the value of environmental attributes. 137

hedonic treadmill A concept describing the tendency of humans to return to a stable level of well-being, even after positive or negative experiences, limiting sustainable increases in happiness. 114

Height Above Nearest Drainage (`HAND`) A topographic index representing the height of a point relative to the nearest drainage channel. It is calculated as the difference between the local altitude and the altitude of the nearest drainage point. `HAND` is used to map wetlands and identify flood risks, making it a valuable tool for hydrological modeling and geoprocessing by representing the topography in relation to drainage networks. 101

heuristic A set of problem-solving techniques that do not guarantee an optimal or rational solution but are sufficient to achieve practical decision-making purposes. The main example is solving problems through trial and error. 25

homology A formal analogy made in modeling, which is an equivalence between the mathematical structures of the target system and the model. 37, 107

human well-being A condition of satisfaction and quality of life for an individual or group, which utilitarianism aims to maximize through actions that promote pleasure and reduce pain. 114, 123, 124, 126, 131, 135

Humanism A philosophical movement that places human beings at the center of ethical and existential considerations, in contrast to belief in supernatural deities. Humanism is seen as the basis for doctrines like Liberalism, Socialism, and Fascism, which interpret human value in distinct ways. 113, 114

hydraulic conductivity K The maximum potential percolation flow of water in the phreatic zone of an aquifer. 71, 78, 99, 100

hydraulic transmissivity A soil property that represents the capacity of the porous medium to transmit water per unit of width and depth. Equivalent to hydraulic conductivity per unit of lateral contour. 98

hydro-ecological compartmentalization The concept of separating water in the vadose zone between water widely absorbed by plant rootlets, occupying soil micropores, and water that is rapidly drained (exfiltration and recharge) through soil macropores. 85

hydro-geochemical compartmentalization The concept of separation at multiple scales of water residence times in the pores of the phreatic zone, creating a diversity of geochemical water signatures. 85

hydrological cycle The circular flow of water on Planet Earth, energized by solar radiation. Evaporation transfers surface water to the atmosphere, and this water returns to the surface in the form of precipitation as rain, dew, and snow. 35, 49, 65, 67, 71, 72, 89, 91, 92, 121, 125, 131, 134

hydrological response The way a watershed's outflows (runoff) manifest in response to inflows (rainfall). Typically, there is a clear separation between rapid responses (rises) and slow responses (recessions). 49, 50, 54, 63, 68, 72, 73, 75, 87, 88, 104

hydrological response units Spatial segments or blocks in a watershed that represent hydrologically homogeneous regions in terms of hydrological response, facilitating semi-distributed modeling by grouping areas with similar hydrological behavior. 95, 100, 102, 104, 109

hydrological similarity A condition in which different regions or hydrological units exhibit similar hydrological behavior, allowing them to be grouped or treated uniformly in hydrological models. Hydrological similarity is used to simplify distributed modeling by grouping areas with homogeneous hydrological responses. 95, 100

Hydrology The natural science that studies the hydrological cycle in its terrestrial phase on continents. 34, 35, 37, 44, 60, 63, 65–67, 69–71, 75, 82, 85, 92, 104, 105, 109

hylomorphism A holistic ontological theory proposed by Aristotle, which states that all things are composed of both matter and form. 39

hypothesis A universal statement in a trial phase, aiming to be elevated to the status of theory after confirmation or corroboration. 8–13, 19, 26, 28, 29, 31, 34, 36, 37, 50, 55, 85, 95, 96, 98, 100–102, 107

idealism A metaphysical conception that opposes realism. In this perspective, which can have ontological or epistemological interpretations, reality is understood as a subjective product of the mind. 24

idealization A fundamental procedure used to construct models, making the representations more tangible and understandable than the target system itself. 36, 37, 90

indirect use value Utilitarian value indirectly obtained from a resource, such as the benefit of climate regulation by a forest. 136

induction problem Also known as **Hume's induction problem**. A circular invalid argument that arises from the justification of inductive knowledge through the principle of uniformity, as it invokes inductive knowledge to support itself. 8, 9, 18, 25

inductive inference Empirical reasoning based on generalization or extrapolation, establishing a universal statement from observations of singular statements. The truth of the universal statement is not guaranteed but presents degrees of probability. 8, 9, 24, 31

inference to the best explanation [synonym of abduction] Non-deductive reasoning that seeks to define the hypothesis that best explains empirical evidence. 25

infiltration f The flow of surface water into the soil matrix. 71, 73

Infiltration Age The period between 1930 and 1970 when the scientific community in Hydrology operated under the normality of the Hortonian paradigm, which established infiltration as the key process to explain the alternation between rises and recessions in rivers. 70, 75, 76, 104

infiltration capacity $f_{\mathbf{max}}$ The maximum potential infiltration flow determined by the characteristics of the soil surface. 70–76, 78, 81

infinite regress problem The challenge of establishing the ultimate origin of logical or rational knowledge, given that all premises must be deduced from more fundamental premises, leading to an infinite (or circular) chain of premises. 8, 24

input data Input data are the information or values provided to a model for processing or analysis, serving as the basis for generating results or simulating the behavior of the system under study. 26, 27, 35, 49, 53, 88, 95, 104

input data error Statistical and epistemic uncertainty associated with the data used to configure the model. For instance, rainfall data present statistical measurement uncertainty and the epistemic uncertainty of spatial interpolation. 27

institutional capacity The ability of an institution to manage resources and implement policies effectively and sustainably. 129

instrumentalism An empiricist radical philosophy of science that opposes scientific realism. This doctrine holds that the goal of science is to produce empirically adequate theories and nothing more. It argues that empirical adequacy does not imply a true description of reality. 6, 25

interception The initial flow that fills the vegetation canopy with rainwater. 36, 49, 66, 69

interception capacity $c_{\mathbf{max}}$ The maximum storage level of water in the vegetation canopy before effective rainfall is produced. 70, 71, 95

intrinsic value The value of a resource or service that exists independently of its utility to humans. 135

isotopic signature The concentration of isotopes that allows identifying or tracing the origin of water from its thermal fractionation process in the atmosphere. 83–85

Kalinin-Miyukov-Nash model A hydrological model representing the response of a watershed as a network of reservoirs arranged in series, known as a cascade. It uses a Gamma distribution to parameterize the hydrograph, incorporating parameters such as the hydrograph volume (ν), the effective number of reservoirs (n), and the mean residence time of the reservoirs (k). Developed independently by Kalinin & Miyukov (1957) and Nash (1958). 87

labour value A Marxist concept defining the value of a good based on the labor necessary to produce it, contrasting with marginal utility theory. 116, 117

Latin Hypercube Sampling A statistical sampling strategy used to generate sets of sample points in a high-dimensional space efficiently, ensuring that each dimension is equally represented in all parts of its interval, improving the coverage and representativeness of samples compared to simple random sampling methods. 58

legacy value Utilitarian value attributed to the preservation of resources for the benefit of future generations. 135, 137

leverage points In systems dynamics, leverage points are strategic locations within a complex system where a small change in one aspect can lead to significant changes in the system's behavior, making them crucial for effective interventions and systemic changes. 43, 53, 55, 63

likelihood The probability that the evidence E is true after considering the probability that the hypothesis H is true. Denoted as $P(E|H)$. 11–15, 24, 27, 28

linear reservoir A compartment that exhibits an outflow directly proportional to its level: $Q_t = S_t/k$, where Q is the reservoir's outflow at time t; S is the reservoir's level at time t, and k is the reservoir's mean residence time. 37, 50, 51, 72

local scale A level of analysis that focuses on infinitesimal elements of the soil, allowing for a detailed and accurate representation of hydrological processes in small areas or units, facilitating the description of hydrological phenomena from a microscopic perspective. 56, 90, 92, 94–97, 100, 102, 104, 124

logical positivism An empiricist philosophical movement from the early 20th century, also known as logical empiricism. 9

logistic curve A graph that represents the alternation between the dominance of reinforcing loops and balancing loops, showing initially rapid growth (or decay) that later stabilizes at a plateau due to balancing effects. 40, 42, 47, 48

Macroeconomics A branch of economics focused on analyzing large-scale economic phenomena such as economic growth, inflation, and unemployment, encompassing both national and global economies. xii, 115, 117, 119, 120, 123, 124

macroporosity A network of pores that stores and conducts water in much greater proportion than if considering the apparent porosity of the soil matrix. Higher in structured organic soils with the presence of fauna and flora (bioturbation). Common in weathered rocks, with the presence of fractures and other geological structures. 75, 78, 79, 92

marginal disutility The incremental loss of utility resulting from the reduction of natural capital as the anthroposphere expands. 124

Marginal Utility Theory An economic theory that explains market prices based on the additional (marginal) utility gained from consuming extra units of a good, which decreases as consumption increases. 116, 117, 123, 124, 135

market distortions Factors that prevent the market from allocating resources efficiently, such as lack of information, speculation, or monopolies, which divert the ideal functioning of markets. 117

market-based instruments Policies that create markets for credits and permits for natural resources, such as carbon trading, encouraging efficiency in resource allocation. 129

maximum flow The highest possible flow defined by the physical constraint of a given compartment. 51, 71

measurement error Statistical uncertainty resulting from the measurement of empirical evidence. 27

mental models A term from systems dynamics for subjective and personal models that are still in the early stages of the modeling process. 33, 34, 55, 56, 60, 101

Microeconomics A branch of economics that studies the behavior of individual consumers and producers, exploring how their decisions influence prices and resource allocation in specific markets. 115–117, 123, 124, 128

mineral resources Elements extracted from the Earth's crust that are non-renewable, such as gold, copper, and oil. 125

minimalist models Models that are simplified to the extreme, used to understand complex phenomena by reducing them to the essentials, focusing on fundamental aspects without the complication of excessive details. 38

mitigation or restoration cost method A method that assesses the utilitarian value of a natural service based on the costs required to restore it after degradation. 137

model A model is a simplified representation of a real-world phenomenon, often used to explain, predict, or simulate various processes. xii, xiv, 5–7, 13–18, 21, 26–28, 33, 35–38, 41–58, 60–63, 68, 69, 73, 75, 76, 83, 86–89, 91, 92, 94–97, 99–105, 107, 117–119, 121, 131

model diagnostics A broad set of techniques applied to assess the adequacy of a model in various aspects. In systems dynamics, John Sterman lists the following diagnostics: boundary adequacy; structural adequacy; dimensional consistency; parameter distribution; comparative studies; integration error; extreme conditions; sensitivity analysis; anomalous behaviors; empirical adequacy; surprises; and positive impacts. 35, 55

model structural error Epistemic uncertainty associated with the theoretical concepts and computational procedures employed in a given model. 27

Monte Carlo simulations A numerical method in which numerous statistically equivalent resamplings are performed to estimate the final behavior of a model involving random variables (i.e., when $n \to \infty$). The name Monte Carlo refers to a casino in Monaco, alluding to the idea of making numerous "rolls" to perform a robust statistical analysis. 14, 15

multicriteria analysis An evaluation method considering multiple criteria for decision-making, applicable to environmental policies and sustainability. 135

natural capital Natural resources and services of the biosphere that provide direct or indirect benefits for human well-being, such as water, clean air, fertile soils, and biogeochemical cycles. 123–128, 130, 135, 136

natural renewal rate The maximum rate at which a renewable resource can be extracted without compromising its regenerative capacity. 125

natural resources Elements of nature that can be used directly or indirectly to benefit humans, such as water, minerals, and air. 123–125, 132

natural scale A level of analysis referring to the actual characteristic speeds exhibited by hydrological processes in nature, including the lifespan of intermittent events, annual event periods, and trends in long-term stochastic processes. 93, 94

natural services Services provided by natural capital, such as pollination, climate regulation, and water purification. 123–126, 130–137

Naturalism A philosophical theory that holds that reality is composed only of natural elements, without supernatural elements, and that any explanation about the universe must come from the natural world itself. 114

negligibility premises A concept introduced by Musgrave (1980), referring to the process of ignoring known important causal factors during abstraction, i.e., when abstraction ends up presenting a model with known falsehoods. 36, 75, 89

non-consumptive Use of resources without direct transformation or consumption, such as recreational activities in natural areas. 136, 137

non-renewable resources Resources that form over geological timescales, such as oil and minerals. 125

non-rival resources Resources that can be used simultaneously by multiple agents, such as solar radiation. 128

non-use value Utilitarian value attributed to a natural resource or service without direct use, such as preservation value. 135

normal science A concept articulated by Thomas Kuhn referring to the historical period in which a given scientific community shares the same paradigm. Normal science tends to end in a crisis, followed by a revolution imposed by the advent of a new paradigm. 22, 23, 31, 70

numerical integration problem The difficulty of obtaining exact values when solving balance equations in the simulation of dynamic systems on digital computers. See truncation error. 46

objective Bayesianism An approach in Bayesian epistemology that argues that the prior distribution must be defined in such a way as to observe the principle of indifference. 12, 13

objective function A mathematical expression that defines the variable (or set of variables) to be maximized or minimized in an optimization problem. 58

observational scale A level of analysis related to the scale of empirical observations in hydrological modeling, including aspects such as data extent, sampling resolution, and sampling integration intervals. 93, 94

old water paradox The difficulty in explaining the rapid mobilization and high prevalence of old water in rivers after precipitation of new water, as well as the diversity of geochemical signatures of old water. 83, 85, 106, 109

open access problem An economic dilemma in which common resources are overexploited due to lack of regulation, leading to environmental degradation. 126–129

open systems Systems capable of processing an inflow and outflow of matter, energy, and information, in contrast to the closed systems described by classical thermodynamics. 40, 88

opportunity cost The sacrifice of natural services and resources due to the expansion of anthropogenic capital over the biosphere. 123–125, 128, 137

option value Utilitarian value associated with preserving a resource for future use, considering the uncertainty about its potential utilities. 136

organic horizon O A general term for the upper layer of soil with greater macroporosity due to the action of soil fauna and flora. 78

overfitting problem A problem that emerges in model calibration when a model is excessively fitted to the available empirical information, resulting in poorer performance when new empirical observations are evaluated. 27

overshoot and collapse curve A typical graph of systems with two main compartments, where one compartment is drained by the other, producing a pattern of accelerated growth followed by an abrupt drop in levels when resources are exhausted. 47

paradigm A concept articulated by Thomas Kuhn referring to the set of exemplary solutions to research problems, i.e., a system of theories, instruments, and auxiliary practices that solve certain widely accepted problems and are promising for resolving open controversial problems with great competitive appeal. xii, 22, 23, 26, 27, 29, 31, 35, 39, 40, 43, 53, 63, 70, 74–76, 79, 82, 87, 89, 95, 104–106, 109, 115, 118, 119, 123, 124, 131

parameters Fixed values of coefficients that define the characteristics and behaviors of elements and processes within a model. They are used to adjust the relationships and functions of the system, determining the system's response and dynamics under different conditions. 14, 18, 20, 21, 26, 27, 43, 45, 48, 52–54, 56–60, 63, 69, 71, 87–89, 92, 93, 95, 96, 99, 100, 103

parametric space The set of all possible combinations of a model's parameters, used to explore and analyze how different parameter values affect the system's behavior and results. In general, the parametric space has N dimensions, where N is the number of parameters. 58, 60

Pareto efficiency A situation in which resources are allocated so that no improvement is possible without worsening the condition of at least one person, indicating an efficient distribution according to marginalist theory. 117

perceptual model Also called a mental model, it consists of the subjective and highly personal representation of an individual about the target system (object). 34, 35, 49, 50, 53, 55, 56, 63, 70, 71, 73, 74, 78, 81, 105

perennial springs Points and patches in the landscape where the main unconfined aquifer of a watershed surfaces, contributing to the baseflow in the slow response of rivers (recessions). 76

permeability transitions Changes in effective hydraulic conductivity observed between different soil horizons. Generally, hydraulic conductivity increases in more superficial horizons due to macroporosity. 78, 105

pest control natural service Control of pest populations by natural predators and parasites, reducing the need for pesticides. 131

phreatic zone G A porous matrix of soil and rock that stores water in an unconfined aquifer under atmospheric pressure. Also known as the saturated zone. 71, 72, 82, 85, 87

Physicalism An ontological doctrine postulating that everything that exists is material and can be described by the laws of physics, a worldview aligned with Scientific Realism and ecological economics. 118, 119

physically based models Hydrological models based on fundamental physical laws, such as the conservation of mass, momentum, and energy. They differ from systemic models by using continuous representations of vector fields to simulate hydrological processes, providing a more detailed and theoretically consistent description of watershed behavior. 89, 92, 93, 96

point of no return A critical threshold where an environmental change becomes irreversible, leading to a significant transformation in the ecosystem. 136

pollination natural service An ecosystem service in which pollinators aid plant reproduction, essential for agricultural production. 131

posterior The probability that the hypothesis H is true after considering the probability that the favorable evidence E is true. Denoted as $P(H|E)$. 10, 12, 14

potential flow A calculated inflow or outflow that potentially alters the level of a compartment in the simulation of a dynamic system. The flow needs to be confronted with the imposed physical constraints (usually conservation and non-negativity). 51, 52

pragmatic realism Term proposed by Keith Beven to describe the implicit realism commonly held by environmental model users. In this philosophy, it is accepted that models provide approximate representations of reality and can improve as new technologies become available. 6, 24, 27, 35

precautionary principle A principle recommending preventive measures in the face of uncertainty about environmental risks. 136

predictive capability The ability of hydrological models to predict runoff behavior based on input data, such as precipitation. It refers to the accuracy and reliability of the model's forecasts. 63, 87, 117, 159

predictive models Hydrological models used to solve specific practical problems, focusing on predicting hydrological events under given conditions. These models apply parameters conditioned by empirical observations in specific temporal and spatial contexts to generate forecasts in different situations. 93

price The monetary amount assigned to a good or service in markets, indicating its scarcity and marginal utility. 116, 118, 137

price-based valuation A method that calculates the utilitarian value of a resource based on market prices, adjusting for distortions when necessary. 137

principle of conditionalization Principle used in Bayesian epistemology to update degrees of conviction in hypotheses based on evidence. To maintain consistency with the principle of probabilism, conditionalization involves zeroing, scaling, and normalizing the values of updated probabilities. 11

principle of diminishing marginal utility A concept indicating that the utility of consuming additional units of a good tends to decrease as total consumption increases. 116

principle of indifference A principle adopted by the objective Bayesianism approach, stating that the degree of conviction in two or more hypotheses should be equal unless there are reasons to the contrary. In the case of complete ignorance, the prior distribution must be uniform. 11, 13

principle of probabilism Principle used in Bayesian epistemology to treat degrees of conviction as probabilities. It has three axioms: non-negativity; normalization; and additivity. 10, 12

principle of uniformity Assumption that the same natural regularities observed empirically in the past will be the same in the future, i.e., that nature is predictable based on its past and that no arbitrary changes will occur in its laws (for example, the Earth suddenly stopping its rotation). 9

prior The probability that the hypothesis H is true before considering the probability that the favorable evidence E is true. Denoted as $P(H)$. 10, 12, 13

private property A right guaranteed by the State to an individual or organization, allowing exclusive use and trade of a resource. 128, 129

problem of justification The challenge of establishing the truth of a particular piece of knowledge or theory. 7, 17, 25

problem of priors The difficulty of justifying the initial definition of degrees of conviction in a hypothesis before any evidence is obtained. In Bayesian epistemology, solutions to this problem are proposed mainly through two approaches: objective and subjective. 12, 13

procedural model A practical representation of a conceptual model in a computer program, where the equations and concepts of the conceptual model are translated into code, allowing simulations and predictions of flows and levels based on input data. 34, 35, 47, 53, 56, 57, 63, 100

process-driven models Hydrological models that represent the physical processes occurring in a watershed. They allow the simulation of watershed behavior even without empirical observations, offering a theoretical basis for runoff generation and enabling deductive inference of hydrological processes. 87, 88

production function-based valuation A method that links the performance of natural services to the production of goods in markets. 137

provisioning natural services Direct material resources provided by nature, such as food and water. 126, 131, 133, 136, 137

rating curve Functional relationship between the level and flow of a river or channel at a specific section. Typically, the following power function is used: $Q = a(h-h_0)^b$, where Q is the flow; h is the level; and a, b, and h_0 are parameters adjusted by observed data. 13, 18

rationalism Philosophical doctrine that supports the superiority of deductive, intuitive, and innate logic to human knowledge, justifying the truth of theories. This doctrine opposes empiricism. 7, 9, 17

realism A metaphysical conception that admits the existence of objective reality, i.e., reality does not depend on anyone to observe it. 24, 26

recession curve The drainage curve of the phreatic zone displayed on a river's hydrograph during a drought (baseflow). 72, 73, 78, 99

recharge $q_\mathbf{v}$ The vertical water flow (percolation) that transfers water from the vadose zone to the phreatic zone. Also known as ultimate percolation. 71, 73, 78, 80–82, 85, 96, 97, 100

recyclable resources Materials that can be processed for reuse, reducing the consumption of virgin resources. 126

regionalization The process of adapting and applying hydrological models developed for a specific region to other regions with different characteristics, involving the generalization of model parameters and processes to fit new geographical and hydrological conditions. 95

regularization effect The stabilization of water flow over time, minimizing extreme variations and ensuring availability for longer periods. 55

regulating and maintenance natural services Natural processes that regulate environmental conditions, such as water purification and climate regulation. 126, 133, 136, 137

reinforcing loop Feedback that acts on both inflows and outflows, increasing the value of this flow, which can result in exponential behaviors (growth or decay). 47

renewable resources Resources that can regenerate naturally within a relatively short period, such as water and wood. 125, 131

replacement cost method A method that calculates the utilitarian value of a service based on the cost of replacing it with artificial alternatives. 137

representation problem The difficulty of constructing a model that performs the semantic or syntactic function of representing a target system. 36

reproducibility problem A typical problem of dynamic systems models pointed out by John Sterman, where models are difficult to use by anyone other than their developers. 62

reusable resources Resources that can be used more than once without significant loss of quality, such as recyclable metals. 126

revealed preferences valuation A method that deduces the utilitarian value of a resource based on behavior in indirectly related markets. 136

rights-based value Non-utilitarian value attributed to natural resources based on ethical principles or inherent rights of existence, not solely on utility. 135

riparian wetlands Zones in valley bottoms, near watercourses, where the water table frequently surfaces. 74, 79, 80, 93, 95, 96, 99, 100, 103, 106, 107, 109

risk aversion The tendency to avoid changes that could result in significant environmental or economic losses. 136

rival resources Resources that cannot be consumed by more than one user at the same time, such as food and fuels. 128

rivalry A characteristic of a resource that prevents simultaneous use by multiple users. 128

runoff by excess of infiltration The generalized surface flow produced due to the soil's relatively lower infiltration capacity compared to the effective rainfall. A synonym for **surface runoff**. 69, 72

saturation index A topographic indicator used in hydrological modeling to represent soil saturation in a given area. The saturation index relates topographic characteristics such as slope and drainage area to determine the propensity of an area to become saturated during rainfall events, and it is fundamental for the local distribution of soil water deficit in hydrological models. 97, 100–102, 104

saturation-activation process A hydrological process that manifests at all scales in Connectivity Theory. 105–107

scale models Representations that are literally copies of the target system at a scale suitable for human manipulation, whether scaled-down or enlarged. 36, 37, 89

scale problem The difficulties that arise when the scale represented by the hydrological model differs from the scale of empirical observations. This discrepancy can introduce errors into the model's results, making them incommensurable or incompatible with observed evidence. 88, 89, 91, 93, 104, 107, 109, 125

scale similarity The ability to convert between the real scale of a target system and the scale of a reduced or enlarged model. Similarity is usually not complete, being valid only in certain aspects (e.g., geometrically similar, but not in terms of density or strength). 36, 89

scaling The process of transferring information between different spatial and temporal scales in hydrological modeling, involving the adaptation of data and parameters from one scale to another to ensure model consistency and accuracy. 93–95, 97, 104, 109

scaling function A mathematical function that defines how information is transferred between different scales in hydrological modeling, determining the form of aggregation or distribution of parameters and variables to maintain model consistency and accuracy. 95, 96, 99, 101

science-management duality A characteristic of Hydrology, existing at the interface between theoretical investigation of nature and practical solutions for social, environmental, and economic issues. 66, 93

scientific community The people who practice science at a given period in history. It can refer to the entirety of scientists or a specific subset within a field. Thomas Kuhn argues that, during certain historical periods, the scientific community is characterized by sharing a paradigm. 22, 23, 30, 31, 70, 74, 75, 85, 91, 92

scientific realism A current in the philosophy of science that defends the thesis that the purpose of science is to provide theories that are true descriptions of reality. 9, 24, 25, 31

second law of thermodynamics A law stating that entropy in an isolated system tends to increase, implying that energy degrades and useful work becomes increasingly difficult to achieve. 119, 121, 126

self-organization The ability of a system to organize and regulate itself without external intervention, as occurs in natural ecosystems. 126

semi-distributed models Hydrological models that use an intermediate approach between fully distributed models and aggregated models. These models divide the watershed into hydrological response units, allowing for a more detailed representation than aggregated models but with less computational complexity than distributed models. 95, 105

sensitivity analysis A model diagnostic technique that seeks to understand how the modeled system responds to changes in its elements, such as inflows and parameter values. 57, 58

separating surface A surface created by the vertical permeability transition between soil horizons, separating vertical flow into a lateral component. The concept is generalized by the theory of connectivity. 105, 106

simultaneous depletion problem The difficulty of applying balance equations using the Euler method when multiple outflows drain the level of a compartment. A solution is to proportionally allocate between the individual flows if the total outflow exceeds the level in the simulated time step. 51, 52

soil natural services Services provided by soil, including plant support, water retention, and nutrient cycling. 131

space of possibilities The set of possibilities generated between hypotheses and evidence in Bayesian epistemology. For probability mathematics to apply to this set, the possibilities must be *mutually exclusive* (cannot be true at the same time) and *collectively exhaustive* (at least one is true). 10, 11, 13

spatial heterogeneity Variability or diversity in the spatial distribution of hydrological characteristics, such as soils, vegetation cover, topography, and hydrodynamic properties. Spatial heterogeneity significantly influences the hydrological response of a watershed, requiring models to accurately represent this diversity to simulate runoff and infiltration processes. 95, 100

stated preference valuation A method that estimates the utilitarian value of resources through surveys with individuals in simulated contexts. 136

statistical model A statistical model is a specific theory about the mathematical behavior of data, without theoretical links to underlying phenomena. 20, 21, 55

statistical uncertainty Uncertainty arising from random noise in the observed data. This type of uncertainty has stationary statistical characteristics that may or may not be structured with bias, heteroscedasticity, and autocorrelation. One way or another, this uncertainty can be modeled by probability distributions. 9, 26

stock-flow resources Resources that are materially consumed in the production process, such as oil and wood. 125, 126, 128

storage deficit The available storage of a compartment that has a maximum capacity. 51

strange attractor A set of states in a dynamic system that, despite being chaotic, possesses a defined geometric structure and attracts the system's trajectories, characterizing an ordered behavior within chaos. 43

structural isomorphism A concept articulated by Ludwig von Bertalanffy to support General Systems Theory, being a formal analogy (homology) observed in different phenomena. 39, 47

subjective Bayesianism An approach in Bayesian epistemology that argues that any prior distribution is valid as long as it does not violate the principle of probabilism. 12, 13

supplementary equations Equations used in the programming of dynamic systems (procedural model) to capture important information that is not part of the modeled system itself, such as statistics of variables. 47

supporting natural services Ecosystem services essential to the functioning of ecosystems, such as oxygen production and soil formation. 133

surface retention capacity $s_{\mathbf{max}}$ The maximum storage level in surface depressions. 70, 72, 95, 97

surface runoff q_{si} The generalized surface flow produced due to the soil's relatively lower infiltration capacity compared to the effective rainfall. 6, 50, 70, 72–76, 78, 79, 81–83, 87, 89

sustainable development Balanced economic growth within the limits imposed by the biosphere, aiming to preserve natural capital for future generations. 123, 125, 130, 133

system An emergent ontological entity defined by a set of fundamental parts that exhibit relationships with each other. 5, 19, 22, 26, 27, 34, 37–50, 52–61, 63, 70–73, 75, 89, 90, 93, 100, 104, 106, 107, 116, 118–121, 123–126, 129, 133, 135, 136

system boundary In systems dynamics, the system boundary defines the limits of what is included or excluded in a system analysis, specifying which compartments exert relevant causal effects on the system without being considered external factors. 45, 49

Systems Dynamics Systems dynamics is a modeling and analysis approach that uses feedback loops, levels, flows, and delays to understand the behavior of complex systems over time, helping to identify and predict behavior patterns and their underlying causes. xii, 35, 43–45, 47–49, 51, 53, 55, 56, 60, 62, 63, 69, 86–89, 91, 92, 94, 95, 104–107, 109, 119

target system A real system that a model supposedly seeks to represent, conveying a theory or hypothesis about this system. 35–38, 46, 49, 53, 55–57, 63, 68, 94, 107, 118

temporal insensitivity principle A guideline by John Sterman, within systems dynamics, that the results of model simulations should not be sensitive to the time step used in numerical integration, regardless of the method adopted. 46, 57

theoretical incommensurability A concept articulated by Thomas Kuhn, referring to the problem of intellectual communication between theories under different paradigms. Two paradigms are fundamentally different, making comparison between their concepts precarious (even if they use the same name and mathematical symbol). 23

theory A universal statement (or system of statements) that definitively establishes the truth of a phenomenon. 7–9, 17–22, 24, 25, 35–40, 46, 53, 55–57, 61, 68, 70–75, 78, 81–84, 87, 88, 91, 92, 101, 103, 104, 106–109, 114–119, 132

thermal fractionation A change in the concentration of isotopes caused by phase changes of water (evaporation and condensation). The water from a given rain event has a different isotopic signature from ocean water (and other rain events) due to its trajectory of thermal fractionation. 84

throughput flow A linear flow of matter and energy through the anthroposphere, consuming resources and generating waste; a central concept in ecological economics to illustrate the environmental impact of economic activities. 121, 123, 126

time of concentration An effective response time parameter used to determine the unit hydrograph. 87

Topographic Wetness Index (`TWI`) A topographic index used to estimate soil moisture based on terrain slope and drainage area. It is calculated by the formula $\mathrm{T}_i = \ln(\alpha_i / \tan\beta_i)$, where α_i is the local drainage area per unit contour length, and β_i is the local terrain slope. `TWI` helps identify areas prone to saturation and runoff, and is widely used in hydrological models such as `TOPMODEL`. 96, 97

Total Economic Value The sum of all utilitarian values of a resource, including use and non-use values. 135, 136

total error equation An equation that includes all sources of errors in a model, both statistical and epistemic. 26, 27, 89

tradable resources Resources that can be exchanged in the market, facilitating allocation according to supply and demand. 128

transaction cost Expenses associated with negotiating a resource, including information search, enforcement, and contract implementation. 130

translational flow $Q_{\mathbf{gt}}$ The water flow from the phreatic zone produced by the sudden pressurization of capillary fringes in riparian areas, where the water table is near the surface. Considered a fast response. 75, 79, 80, 85

travel cost method A method that estimates the utilitarian value of natural areas based on the costs incurred by visitors to access these areas. 137

truncation error The difference between the exact value of a function or analytical mathematical calculation and its approximation resulting from the numerical method employed to calculate the value in a computational environment. 46, 48

two-world hypothesis A testable hypothesis about hydro-ecological compartmentalization proposed by Jeffrey McDonnell. One world of water would be the water for plants (green water), and the other world would be the water for rivers (blue water). 85

ultimate goal The fundamental purpose, from a utilitarian perspective, of maximizing well-being and minimizing suffering for the greatest possible number of individuals. 113, 115, 118, 123, 135

underdetermination problem The difficulty of ensuring that the observed evidence determines the truth of a theory without there being other empirically equivalent theories. 7, 25, 26, 28, 40, 88

Unit Hydrograph The minimal linear response of a watershed. Complex responses can be constructed from the unit hydrograph through convolution. 87

upscaling The process of transferring information from smaller scales to larger scales in hydrological modeling, often performed through averaging or summation over a specific spatial or temporal extent. 95

use permit A formal authorization issued by the State for the use of a natural resource, regulating access and the permissible amount of use. 129

use value The intrinsic value of a resource based on its ability to meet direct needs or desires, such as water for human consumption. 116, 129, 131, 135–137

utilitarian value Value attributed to a resource based on the utility it provides to humans, typically measured in monetary terms. 134, 135

Utilitarianism An ethical theory that assesses the morality of actions based on their consequences, seeking to maximize well-being and reduce suffering for the greatest number of people. 114, 116, 134, 135

utility The capacity of an action, good, or service to satisfy the needs or desires of an individual, serving as a central measure in utilitarian theory to evaluate the moral value of choices. 114–119, 121, 124, 131, 135, 137

vadose zone V A porous matrix of solid soil minerals that stores water in films held by surface tension. Also known as the unsaturated zone. 71, 76, 78, 82, 87, 90, 91, 96, 100

value paradox The difference between the use value and exchange value of goods, such as water and diamonds, where essential items may have a lower market value than luxury items due to their abundance. 116

variable source area The phenomenon of expansion and contraction of riparian wetland areas that produces a variable contribution of runoff from direct rainfall. The variation of the source area can occur during a rainfall event or over the course of seasons. 80, 81, 95, 97

vegetation canopy C The leaves and branches of plants that act as a compartment or reservoir that stores water through surface tension before the rain reaches the soil surface. 71

zero-order basin Regions on slopes and higher grounds of the landscape where precipitation interacts with vegetation, soil, and rocks, producing the hydrological responses observed downstream in rivers. Also known as **slope basin**. 67, 68, 71, 106

Bibliography

[1] U. Nations, *The United Nations world water development report 2018: nature-based solutions for water* (The United Nations world water development report). United Nations Educational, Scientific and Cultural Organization, 2018, ISBN: 9789231002649.

[2] Brasil, *Lei Nº 14.119, de 13 de Janeiro de 2021. Institui a Política Nacional de Pagamento por Serviços Ambientais e altera as Leis n.os 8.212, de 24 de julho de 1991, 8.629, de 25 de fevereiro de 1993, e 6.015, de 31 de dezembro de 1973, para adequá-las à nova política.* 2021. [Online]. Available: http://www.planalto.gov.br/ccivil_03/_Ato2019-2022/2021/Lei/L14119.htm.

[3] I. Possantti and G. Marques, "A modelling framework for nature-based solutions expansion planning considering the benefits to downstream urban water users," *Environmental Modelling & Software*, vol. 152, p. 105 381, 2022, ISSN: 1364-8152. DOI: https://doi.org/10.1016/j.envsoft.2022.105381.

[4] I. Possantti, R. Barbedo, M. Kronbauer, W. Collischonn, and G. Marques, "A comprehensive strategy for modeling watershed restoration priority areas under epistemic uncertainty: A case study in the atlantic forest, brazil," *Journal of Hydrology*, vol. 617, p. 129 003, Feb. 2023, ISSN: 00221694. DOI: https://doi.org/10.1016/j.jhydrol.2022.129003.

[5] L. A. Melsen, "It takes a village to run a model—the social practices of hydrological modeling," *Water Resources Research*, vol. 58, no. 2, e2021WR030600, 2022, e2021WR030600 2021WR030600. DOI: https://doi.org/10.1029/2021WR030600.

[6] K. Beven, "Towards a coherent philosophy for modelling the environment," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 458, no. 2026, pp. 2465–2484, 2002. DOI: 10.1098/rspa.2002.0986.

[7] H. Daly, "Economics for a full world," *Great Transition Initiative*, Jun. 2015. [Online]. Available: https://greattransition.org/publication/economics-for-a-full-world.

[8] R. Descartes, P. Neves, and D. Rosenfield, *Discurso do método:* (Coleção L&PM pocket). L&PM, 2008, ISBN: 9788525410979.

[9] M. Ongaro and M. Andreoletti, "Non-empirical uncertainties in evidence-based decision making," *Perspectives on Science*, vol. 30, pp. 305–320, 2 Apr. 2022, - Evidence-based policy
- Scientific advice is limited to empirical uncertainty, ISSN: 15309274. DOI: 10.1162/posc_a_00416.

[10] D. Kahneman, *Thinking, Fast and Slow*. Farrar, Straus and Giroux, 2011, ISBN: 9781429969352.

[11] P. Markie and M. Folescu, "Rationalism vs. Empiricism," in *The Stanford Encyclopedia of Philosophy*, E. N. Zalta and U. Nodelman, Eds., Spring 2023, Metaphysics Research Lab, Stanford University, 2023.

[12] W. Uzgalis, "John Locke," in *The Stanford Encyclopedia of Philosophy*, E. N. Zalta and U. Nodelman, Eds., Fall 2022, Metaphysics Research Lab, Stanford University, 2022.

[13] P. Johnson-Laird, "Deductive reasoning," *WIREs Cognitive Science*, vol. 1, no. 1, pp. 8–17, 2010. DOI: https://doi.org/10.1002/wcs.20. eprint: https://wires.onlinelibrary.wiley.com/doi/pdf/10.1002/wcs.20. [Online]. Available: https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/wcs.20.

[14] A. D. Cling, "The epistemic regress problem," *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, vol. 140, pp. 401–421, 3 2008. DOI: 10.1007/sl. [Online]. Available: https://about.jstor.org/terms.

[15] L. Henderson, "The Problem of Induction," in *The Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., Spring 2020, Metaphysics Research Lab, Stanford University, 2020.

[16] F. L. d. Silveira, "A teoria do conhecimento de kant: O idealismo transcendental," *Caderno brasileiro de ensino de física*, vol. 19, pp. 28–51, 2002.

[17] L. C. Weiss and M. Bräscher, "Organização do conhecimento e kant: Uma análise do debate epistemológico sobre realismo e idealismo," *Informação, Memória e Tecnologia*, p. 48, 2017.

[18] H. Lin, "Bayesian Epistemology," in *The Stanford Encyclopedia of Philosophy*, E. N. Zalta and U. Nodelman, Eds., Fall 2022, Metaphysics Research Lab, Stanford University, 2022.

[19] J. Sprenger and S. Hartmann, *Bayesian Philosophy of Science*. Oxford University Press, Aug. 2019, ISBN: 9780199672110. DOI: 10.1093/oso/9780199672110.001.0001.

[20] G. Box, *Robustness in the Strategy of Scientific Model Building*. Academic Press, 1979, pp. 201–236, ISBN: 0124381502. DOI: 10.1016/b978-0-12-438150-6.50018-2. [Online]. Available: http://dx.doi.org/10.1016/B978-0-12-438150-6.50018-2.

[21] F. Graybill and H. Iyer, *Regression Analysis: Concepts and Applications* (An Alexander Kugushev book). Duxbury Press, 1994, ISBN: 9780534198695.

[22] T. Morlot, C. Perret, A. C. Favre, and J. Jalbert, "Dynamic rating curve assessment for hydrometric stations and computation of the associated uncertainties: Quality and station management indicators," *Journal of Hydrology*, vol. 517, pp. 173–186, Sep. 2014, ISSN: 00221694. DOI: 10.1016/j.jhydrol.2014.05.007.

[23] K. R. Popper, *Os dois problemas fundamentais da teoria do conhecimento*. Editora Unesp, 2013.

[24] K. R. Popper, *A lógica da pesquisa científica*. Editora Cultrix, 2004.

[25] T. Kuhn and I. Hacking, *The Structure of Scientific Revolutions*. University of Chicago Press, 2012, ISBN: 9780226458144.

[26] B. C. v. Fraassen, *The Scientific Image*. Oxford University Press, Dec. 1980, ISBN: 9780198244271. DOI: 10.1093/0198244274.001.0001. [Online]. Available: https://doi.org/10.1093/0198244274.001.0001.

[27] E. Agazzi, "The truth of theories and scientific realism," in *Varieties of Scientific Realism: Objectivity and Truth in Science*, E. Agazzi, Ed. Cham: Springer International Publishing, 2017, pp. 49–68, ISBN: 978-3-319-51608-0. DOI: 10.1007/978-3-319-51608-0_3. [Online]. Available: https://doi.org/10.1007/978-3-319-51608-0_3.

[28] N. Cartwright, *How the Laws of Physics Lie* (Clarendon paperbacks). Clarendon Press, 1983, ISBN: 9780198247043.

[29] B. Monton and C. Mohler, "Constructive Empiricism," in *The Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., Summer 2021, Metaphysics Research Lab, Stanford University, 2021.

[30] K. Stanford, "Underdetermination of Scientific Theory," in *The Stanford Encyclopedia of Philosophy*, E. N. Zalta and U. Nodelman, Eds., Summer 2023, Metaphysics Research Lab, Stanford University, 2023.

[31] D. Tulodziecki, "Underdetermination," in *The Routledge Handbook of Scientific Realism*, J. Saatsi, Ed. Taylor & Francis, 2017, pp. 60–71.

[32] J. Saatsi, *The Routledge Handbook of Scientific Realism* (Routledge Handbooks in Philosophy). Taylor & Francis, 2017, ISBN: 9781351362900.

[33] D. D. Hoffman, M. Singh, and C. Prakash, "The interface theory of perception," *Psychonomic Bulletin and Review*, vol. 22, pp. 1480–1506, 6 Dec. 2015, ISSN: 15315320. DOI: 10.3758/s13423-015-0890-8.

[34] D. D. Hoffman, C. Prakash, and R. Prentner, "Fusions of consciousness," *Entropy*, vol. 25, 1 Jan. 2023, ISSN: 10994300. DOI: 10.3390/e25010129.

[35] O. N., S.-F. K., and K. Belitz, "Verification, validation, and confirmation of numerical models in the earth sciences," *Science*, vol. 263, pp. 641–646, 5147 1994.

[36] K. Beven, "A manifesto for the equifinality thesis," *Journal of Hydrology*, vol. 320, pp. 18–36, 1-2 Mar. 2006, ISSN: 00221694. DOI: 10.1016/j.jhydrol.2005.07.007.

[37] K. Beven, "On the concept of model structural error," *Water Science and Technology*, vol. 52, pp. 167–175, 6 2005. [Online]. Available: https://iwaponline.com/wst/article-pdf/52/6/167/434056/167.pdf.

[38] K. Beven, "Facets of uncertainty: Epistemic uncertainty, non-stationarity, likelihood, hypothesis testing, and communication," *Hydrological Sciences Journal*, vol. 61, pp. 1652–1665, 9 Jul. 2016, ISSN: 21503435. DOI: 10.1080/02626667.2015.1031761.

[39] A. Tarantola, "Popper, bayes and the inverse problem," *Nature Physics*, vol. 21, pp. 492–494, 2006.

[40] K. Beven, *Environmental Modelling: An Uncertain Future? An introduction to techniques for uncertainty estimation in environmental prediction.* Routledge, 2009, p. 329, ISBN: 0-203-93248-X.

[41] G. R. Evenson, M. Kalcic, Y.-C. Wang, *et al.*, "Uncertainty in critical source area predictions from watershed-scale hydrologic models," *Journal of Environmental Management*, vol. 279, p. 111 506, 2021, ISSN: 0301-4797. DOI: https://doi.org/10.1016/j.jenvman.2020.111506.

[42] K. J. Beven, "On hypothesis testing in hydrology: Why falsification of models is still a really good idea," *WIREs Water*, vol. 5, no. 3, e1278, 2018. DOI: https://doi.org/10.1002/wat2.1278.

[43] D. H. Meadows, *Thinking in systems: A primer.* Chelsea Green Publishing, 2008.

[44] J. Sterman, *Business Dynamics: Systems Thinking and Modeling for a Complex World* (McGraw-Hill Higher Education). Irwin/McGraw-Hill, 2000, ISBN: 9780072311358.

[45] D. Meadows and C. of Rome, *The Limits to Growth: A Report for the Club of Rome's Project on the Predicament of Mankind* ((A Signet classic) pt. 1). Universe Books, 1974, ISBN: 9780876639016.

[46] K. J. Beven, *Rainfall-runoff modelling: the primer*. John Wiley & Sons, 2011.

[47] M. Hesse, "Models and analogies," in *A Companion to the Philosophy of Science*. John Wiley and Sons, Ltd, 2017, ch. 44, pp. 299–307, ISBN: 9781405164481. DOI: https://doi.org/10.1002/9781405164481.ch44.

[48] R. Frigg and S. Hartmann, "Models in science," in *The Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., Spring 2020, Metaphysics Research Lab, Stanford University, 2020.

[49] A. Musgrave, "'unreal assumptions'in economic theory: The f-twist untwisted," *Kyklos*, vol. 34, no. 3, pp. 377–387, 1981.

[50] E. McMullin, "Galilean idealization," *Studies in History and Philosophy of Science Part A*, vol. 16, no. 3, pp. 247–273, 1985, ISSN: 0039-3681. DOI: https://doi.org/10.1016/0039-3681(85)90003-2. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0039368185900032.

[51] S. de Chadarevian, "Portrait of a discovery. watson, crick, and the double helix," *Isis*, vol. 94, no. 1, pp. 90–105, 2003. [Online]. Available: https://doi.org/10.1086/376100.

[52] F. L. M. Pereira, *Elaboração e construção de modelo físico de representação do vale a jusante da barragem de canastra - rs*, Salão de Iniciação Científica (35. : 2023 nov. 6-10 : UFRGS, Porto Alegre, RS), Abstract published in event, Porto Alegre, RS, 2023. [Online]. Available: https://lume.ufrgs.br/handle/10183/269798.

[53] W. H. Shaw and L. R. Ashley, "Analogy and inference," *Dialogue*, vol. 22, no. 3, pp. 415–432, 1983. DOI: 10.1017/S0012217300046588.

[54] A. Gelfert, *How to do science with models: A philosophical primer*. Springer, 2016.

[55] B. Rettler, "Quantification in the ontology room," *Dialectica*, vol. 73, no. 4, pp. 563–585, 2019. DOI: https://doi.org/10.1111/1746-8361.12286.

[56] C. Witt, "Hylomorphism in aristotle," *The Journal of Philosophy*, vol. 84, no. 11, pp. 673–679, 1987, ISSN: 0022362X. [Online]. Available: http://www.jstor.org/stable/2026773 (visited on 10/28/2024).

[57] V. Bertalanffy, *Teoria geral dos sistemas: Fundamentos, desenvolvimento e aplicações*. Editora Vozes, 2014, ISBN: 9788532636904.

[58] K. J. Beven, *Rainfall-runoff modelling: the primer*. John Wiley & Sons, 2012.

[59] T. C. Schelling, "Dynamic models of segregation†," *The Journal of Mathematical Sociology*, vol. 1, no. 2, pp. 143–186, 1971. DOI: 10.1080/0022250X.1971.9989794.

[60] S. Wolfram, "Computer software in science and mathematics," *Scientific American*, vol. 251, no. 3, pp. 188–203, 1984.

[61] J. W. Forrester, "System dynamics—a personal view of the first fifty years," *System Dynamics Review: The Journal of the System Dynamics Society*, vol. 23, no. 2-3, pp. 345–358, 2007.

[62] J. Sterman, "System dynamics at sixty: The path forward," *System Dynamics Review*, vol. 34, no. 1-2, pp. 5–47, 2018.

[63] J. Forrester, *Industrial Dynamics* (Students' edition). M.I.T. Press, 1964, ISBN: 9780262060035.

[64] H. Odum and E. Odum, *A Prosperous Way Down: Principles and Policies*. University Press of Colorado, 2008, ISBN: 9781607320814.

[65] U. Pöschl, S. Martin, B. Sinha, *et al.*, "Rainforest aerosols as biogenic nuclei of clouds and precipitation in the amazon," *science*, vol. 329, no. 5998, pp. 1513–1516, 2010.

[66] V.-M. Kerminen, M. Paramonov, T. Anttila, *et al.*, "Cloud condensation nuclei production associated with atmospheric nucleation: A synthesis based on existing literature and new results," *Atmospheric Chemistry and Physics*, vol. 12, no. 24, pp. 12 037–12 059, 2012. DOI: 10.5194/acp-12-12037-2012. [Online]. Available: https://acp.copernicus.org/articles/12/12037/2012/.

[67] J. M. Arraut, C. Nobre, H. M. J. Barbosa, G. Obregon, and J. Marengo, "Aerial rivers and lakes: Looking at large-scale moisture transport and its relation to amazonia and to subtropical rainfall in south america," *Journal of Climate*, vol. 25, no. 2, pp. 543–556, 2012. DOI: 10.1175/2011JCLI4189.1.

[68] F. Pearce, "Rivers in the sky," *New Scientist*, vol. 244, no. 3254, pp. 40–43, 2019, ISSN: 0262-4079. DOI: https://doi.org/10.1016/S0262-4079(19)32070-6.

[69] A. Saltelli, M. Ratto, S. Tarantola, and F. Campolongo, "Sensitivity analysis practices: Strategies for model-based inference," *Reliability Engineering and System Safety*, vol. 91, no. 10, pp. 1109–1125, 2006, The Fourth International Conference on Sensitivity Analysis of Model Output (SAMO 2004), ISSN: 0951-8320. DOI: https://doi.org/10.1016/j.ress.2005.11.014.

[70] A. Saltelli, K. Aleksankina, W. Becker, *et al.*, "Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices," *Environmental Modelling and Software*, vol. 114, pp. 29–39, 2019, ISSN: 1364-8152. DOI: https://doi.org/10.1016/j.envsoft.2019.01.012.

[71] J. H. Miller, "Active nonlinear tests (ants) of complex simulation models," *Management Science*, vol. 44, no. 6, pp. 820–830, 1998.

[72] K. Beven and A. Binley, "The future of distributed models: Model calibration and uncertainty prediction," *Hydrological Processes*, vol. 6, no. 3, pp. 279–298, 1992. DOI: https://doi.org/10.1002/hyp.3360060305.

[73] S. A. Glantz, B. K. Slinker, and T. B. Neilands, *Primer of applied regression & analysis of variance, ed.* McGraw-Hill, Inc., New York, 2001, vol. 654.

[74] J. E. Nash and J. V. Sutcliffe, "River flow forecasting through conceptual models part i-a discussion of principles," *Journal of Hydrology*, vol. 10, pp. 282–290, 1970.

[75] H. V. Gupta, H. Kling, K. K. Yilmaz, and G. F. Martinez, "Decomposition of the mean squared error and nse performance criteria: Implications for improving hydrological modelling," *Journal of Hydrology*, vol. 377, pp. 80–91, 1-2 Oct. 2009, ISSN: 00221694. DOI: 10.1016/j.jhydrol.2009.08.003.

[76] A. Eiben and J. Smith, *Introduction to Evolutionary Computing* (Natural Computing Series). Springer Berlin Heidelberg, 2015, ISBN: 9783662448748.

[77] J. Hewlett and A. R. Hibbert, "Factors affecting the response of small watersheds to precipitation in humid areas," *Forest hydrology*, vol. 11, pp. 275–290, 1967.

[78] J. J. McDonnell, "Where does water go when it rains? moving beyond the variable source area concept of rainfall-runoff response," *Hydrological Processes*, vol. 17, pp. 1869–1875, 9 Jun. 2003, ISSN: 0885-6087. DOI: 10.1002/hyp.5132.

[79] V. T. Chow, *Handbook of Applied Hydrology* (Handbook of Applied Hydrology: A Compendium of Water-resources Technology v. 1). McGraw-Hill, 1964, ISBN: 9780070107748.

[80] J. C. Dooge, "Hydrology in perspective," *Hydrological Sciences Journal*, vol. 33, pp. 61–85, 1 1988, ISSN: 21503435. DOI: 10.1080/02626668809491223.

[81] M. Sivapalan and G. Blöschl, *The growth of hydrological understanding: Technologies, ideas, and societal needs shape the field*, Oct. 2017. DOI: 10.1002/2017WR021396.

[82] M. Sivapalan, "From engineering hydrology to earth system science: Milestones in the transformation of hydrologic science," *Hydrology and Earth System Sciences*, vol. 22, pp. 1665–1693, 3 2018, ISSN: 16077938. DOI: 10.5194/hess-22-1665-2018.

[83] Y. Tsukamoto, "Study on the growth of stream channel (i)," *Shin sabo*, vol. 25, no. 4, pp. 4–13, 1973. DOI: 10.11475/sabo1948.25.4_4.

[84] A. N. Strahler, "Quantitative analysis of watershed geomorphology," *Transactions American Geophysical Union*, vol. 38, no. 6, pp. 913–920, 1957.

[85] E. Mediondo and C. TUCCI, "Escalas hidrológicas. ii: Diversidade de processos na bacia vertente," *Revista Brasileira de Recursos Hídricos*, vol. 2, no. 1, pp. 81–100, 1997.

[86] J. V. Z. Godoy, M. Kobiyama, K. Campagnolo, and M. R. Fagundes, "Utilização do termo bacia de ordem zero no brasil," *Revista Geográfica Acadêmica. Boa Vista. Vol. 15, n. 2 (jul./dez. 2021), p. 89-107*, 2021.

[87] J. J. McDonnell, C. Spence, D. J. Karran, H. J. (van Meerveld, and C. J. Harman, "Fill-and-spill: A process description of runoff generation at the scale of the beholder," *Water Resources Research*, vol. 57, no. 5, e2020WR027514, 2021. DOI: https://doi.org/10.1029/2020WR027514.

[88] A. Strahler, *Geografia fisica*. Omega, 1986.

[89] EPAGRI, *Desafios e soluções para mitigar a erosão provocada pelas chuvas na agricultura catarinense*, https://blog.epagri.sc.gov.br/index.php/artigo-desafios-e-solucoes-para-mitigar-a-erosao-provocada-pelas-chuvas-na-agricultura-catarinense/, Accessed: 2024-08-26, 2024.

[90] R. E. Horton, "The role of infiltration in the hydrologic cycle," *Transactions American Geophysical Union*, pp. 446–460, Reports and papers - hydrology 1933, ISSN: 23249250. DOI: 10.1029/TR014i001p00446.

[91] K. Beven, "Robert e. horton's perceptual model of infiltration processes," *Hydrological Processes*, vol. 18, pp. 3447–3460, 17 Dec. 2004, ISSN: 08856087. DOI: 10.1002/hyp.5740.

[92] H. L. Cook, "The infiltration approach to the calculation of surface runoff," *Eos, Transactions American Geophysical Union*, vol. 27, pp. 726–747, 5 1946.

[93] K. Beven, "The era of infiltration," *Hydrology and Earth System Sciences*, vol. 25, pp. 851–866, 2 Feb. 2021, ISSN: 16077938. DOI: 10.5194/hess-25-851-2021.

[94] S. K. Mishra, J. V. Tyagi, and V. P. Singh, "Comparison of infiltration models," *Hydrological processes*, vol. 17, no. 13, pp. 2629–2652, 2003.

[95] J. Philip, "The theory of infiltration: 1. the infiltration equation and its solution," *Soil science*, vol. 83, no. 5, pp. 345–358, 1957.

[96] R. E. Horton, "Analysis of runoff-plat experiments with varying infiltration capacity," *Transactions American Geophysical Union*, pp. 693–711, Reports and papers - hydrology 1939.

[97] R. Rallison and N. Miller, "Past, present, and future scs runoff procedure," Water Resources Publications, 1981, pp. 353–364.

[98] V. Mockus, *Estimation of total (and peak rates of) surface runoff for individual storms*, 1949.

[99] T. Dunne, "Relation of field studies and modeling in the prediction of storm runoff," *Journal of Hydrology*, vol. 65, no. 1, pp. 25–48, 1983, Scale Problems in Hydrology, ISSN: 0022-1694. DOI: https://doi.org/10.1016/0022-1694(83)90209-3.

[100] K. Beven, "Infiltration excess at the horton hydrology laboratory (or not?)" *Journal of Hydrology*, vol. 293, pp. 219–234, 1-4 Jun. 2004, Horton may never observerd overland runoff at the LaGrange catchment!! ISSN: 00221694. DOI: 10.1016/j.jhydrol.2004.02.001.

[101] R. P. Betson, "What is watershed runoff?" *Journal of Geophysical research*, vol. 69, pp. 1541–1552. 8 1964. DOI: https://doi.org/10.1029/JZ069i008p01541.

[102] M. J. Kirkby, "Infiltration, throughflow, and overland flow," in Methuen, 1969, pp. 215–227. DOI: 10.4324/9781003170181.

[103] J. J. McDonnell, "A rationale for old water discharge through macropores in a steep, humid catchment," *Water Resources Research*, vol. 26, no. 11, pp. 2821–2832, 1990.

[104] J. J. Mcdonnell, "Are all runoff processes the same?" *Hydrological Processes*, vol. 27, pp. 4103–4111, 26 Dec. 2013, ISSN: 08856087. DOI: 10.1002/hyp.10076.

[105] M. D. Hoover and C. R. Hursh, "Influence of topography and soil-depth on runoff from forest land," *Transactions American Geophysical Union*, vol. 24, pp. 693–698, 2 1943. DOI: https://doi.org/10.1029/TR024i002p00693.

[106] A. Schneider, F. Hirsch, A. Raab, and T. Raab, "Dye tracer visualization of infiltration patterns in soils on relict charcoal hearths," *Frontiers in Environmental Science*, vol. 6, 2018, ISSN: 2296-665X. DOI: 10.3389/fenvs.2018.00143.

[107] F. F. Snyder, "A conception of runoff-phenomena," *Eos, Transactions American Geophysical Union*, vol. 20, pp. 725–738, 4 1939.

[108] B. S. Barnes, "The structure of discharge-recession curves," *Eos, Transactions American Geophysical Union*, vol. 20, pp. 721–725, 4 1939.

[109] C. R. Hursh, "Subsurface-flow," *Eos, Transactions American Geophysical Union*, vol. 25, pp. 743–746, 5 1944.

[110] C. R. Hursh and E. F. Brater, "Separating storm-hydrographs from small drainage-areas into surface-and subsurface-flow," *Eos, Transactions American Geophysical Union*, vol. 22, pp. 863–871, 3 1941.

[111] C. R. Hursh and P. W. Fletcher, "The soil profile as a natural reservoir," *Soil Science Society of America Journal*, vol. 7, pp. 480–486, C 1942.

[112] J. D. Hewlett and A. R. Hibbert, "Moisture and energy conditions within a sloping soil mass during drainage," *Journal of Geophysical Research*, vol. 68, pp. 1081–1087, 4 1963. DOI: https://doi.org/10.1029/JZ068i004p01081.

[113] R. Z. Whipkey, "Subsurface stormflow from forested slopes," *International Association of Scientific Hydrology. Bulletin*, vol. 10, pp. 74–85, 2 1965, ISSN: 00206024. DOI: 10.1080/02626666509493392.

[114] R. M. Ragan, "An experimental investigation of partial area contributions," *Hydrological Aspects of the Utilization of Water, Reports and Discussions (Proc. IAHS Assembly)*, pp. 241–251, 1968.

[115] R. S. Beasley, "Contribution of subsurface flow from the upper slopes of forested watersheds to channel flow," *Soil Science Society of America Journal*, vol. 40, pp. 955–957, 6 1976.

[116] R. D. Harr, "Water flux in soil and subsoil on a steep forested slope," *Journal of Hydrology*, vol. 33, p. 37, 1977.

[117] J. H. Patric and D. N. Swanston, "Hydrology of a slide-prone glacial till soil in southeast alaska," *Journal of Forestry*, vol. 66, no. 1, pp. 62–66, 1968.

[118] D. Weyman, "Throughflow on hillslopes and its relation to the stream hydrograph," *Hydrological sciences journal*, vol. 15, no. 3, pp. 25–33, 1970.

[119] A. Jones, "Soil piping and stream channel initiation," *Water Resources Research*, vol. 7, pp. 602–610, 3 1971.

[120] M. P. Mosley, "Streamflow generation in a forested watershed, new zealand," *Water Resources Research*, vol. 15, pp. 795–806, 4 1979.

[121] P. Cappus, "Etude des lois de l'écoulement-application au calcul et à la prévision des débits," *La houille blanche*, pp. 493–520, 1960.

[122] Y. Tsukamoto, "Storm discharge from an experimental watershed," *Journal of the Japanese Forestry Society*, vol. 45, pp. 186–190, 6 1963. DOI: https://doi.org/10.11519/jjfs1953.45.6_186.

[123] T. Dunne and R. D. Black, "Partial area contributions to storm runo in a small new, england watershed," *Water Resources Research*, vol. 6, 5 1970.

[124] T. Dunne and R. D. Black, "An experimental investigation of runoff production in permeable soils," *Water Resources Research*, vol. 6, no. 2, pp. 478–490, 1970. DOI: https://doi.org/10.1029/WR006i002p00478.

[125] M. G. Anderson and T. P. Burt, "The role of topography in controlling throughflow generation," *Earth Surface Processes*, vol. 3, pp. 331–344, 4 1978.

[126] R. Horton, "Maximum ground-water levels," *Transactions American Geophysical Union*, vol. 17, pp. 344–357, 2 1936. DOI: https://doi.org/10.1029/TR017i002p00344.

[127] K. Beven, "Robert e. horton and abrupt rises of ground water," *Hydrological Processes*, vol. 18, pp. 3687–3696, 18 Dec. 2004, ISSN: 08856087. DOI: 10.1002/hyp.5741.

[128] M. G. Sklash and R. N. Farvolden, "The role of groundwater in storm runoff," *Journal of Hydrology*, vol. 43, no. 1-4, pp. 45–65, 1979.

[129] J. W. Kirchner, X. Feng, and C. Neal, "Fractal stream chemistry and its implications for contaminant transport in catchments," *Nature*, vol. 403, no. 6769, pp. 524–527, 2000.

[130] J. W. Kirchner, "A double paradox in catchment hydrology and geochemistry," *Hydrological processes*, vol. 17, pp. 871–874, 4 2003.

[131] Ecological Continuity Trust, *Plynlimon research catchments*, Accessed: 2024-09-01, 2024. [Online]. Available: https://www.ecologicalcontinuitytrust.org/plynlimon-research-catchments.

[132] J. J. McDonnell, "Hewlett, jd and hibbert, ar 1967: Factors affecting the response of small watersheds to precipitation in humid areas. in sopper, we and lull, hw, editors, forest hydrology, new york: Pergamon press, 275—90," *Progress in Physical Geography*, vol. 33, pp. 288–293, 2 2009.

[133] G. F. Pinder and J. F. Jones, "Determination of the ground-water component of peak discharge from the chemistry of total runoff," *Water Resources Research*, vol. 5, no. 2, pp. 438–445, 1969.

[134] T. Dincer, J. Martinec, B. Payne, and C. Yen, "Variation of the tritium and oxygen-18 content in precipitation and snowpack in a representative basin in czechoslovakia," *Isotopes in hydrology: Vienna, International Atomic Energy Agency*, pp. 23–42, 1970.

[135] M. Sklash, R. Farvolden, and P. Fritz, "A conceptual model of watershed response to rainfall, developed through the use of oxygen-18 as a natural tracer," *Canadian Journal of Earth Sciences*, vol. 13, no. 2, pp. 271–283, 1976.

[136] J. Martinec, U. Siegenthaler, H. Oeschger, and E. Tongiorgi, "New insights into the run-off mechanism by environmental isotopes," in *Isotope techniques in groundwater hydrology 1974, Vol. I. Proceedings of a symposium*, 1974.

[137] M. Sklash, M. Stewart, and A. Pearce, "Storm runoff generation in humid headwater catchments: 2. a case study of hillslope and low-order stream response," *Water Resources Research*, vol. 22, no. 8, pp. 1273–1282, 1986.

[138] M. Sklash, K. Beven, K. Gilman, and W. Darling, "Isotope studies of pipeflow at plynlimon, wales, uk," *Hydrological Processes*, vol. 10, no. 7, pp. 921–944, 1996.

[139] J. Buttle, "Isotope hydrograph separations and rapid delivery of pre-event water from drainage basins," *Progress in physical geography*, vol. 18, no. 1, pp. 16–41, 1994.

[140] B. L. McGlynn, J. J. McDonnel, and D. D. Brammer, "A review of the evolving perceptual model of hillslope flowpaths at the maimai catchments, new zealand," *Journal of Hydrology*, vol. 257, no. 1-4, pp. 1–26, 2002.

[141] D. D. Huff, R. V. O'Neill, W. R. Emanuel, J. W. Elwood, and J. D. Newbold, "Flow variability and hillslope hydrology," *Earth Surface Processes and Landforms*, vol. 7, pp. 91–94, 1 1982.

[142] N. Brady and R. Weil, *Elements of the Nature and Properties of Soils*. Pearson Prentice Hall, 2010, ISBN: 9780135051955.

[143] D. A. Burns, J. J. McDonnell, R. P. Hooper, *et al.*, "Quantifying contributions to storm runoff through end-member mixing analysis and hydrologic measurements at the panola mountain research watershed (georgia, usa)," *Hydrological processes*, vol. 15, no. 10, pp. 1903–1924, 2001.

[144] J. Seibert, K. Bishop, A. Rodhe, and J. J. McDonnell, "Groundwater dynamics along a hillslope: A test of the steady state hypothesis," *Water Resources Research*, vol. 39, no. 1, 2003.

[145] I. Iorgulescu, K. J. Beven, and A. Musy, "Flow, mixing, and displacement in using a data-based hydrochemical model to predict conservative tracer data," *Water Resources Research*, vol. 43, 3 Mar. 2007, ISSN: 00431397. DOI: 10.1029/2005WR004019.

[146] J. J. McDonnell, "The two water worlds hypothesis: Ecohydrological separation of water between streams and trees?" *Wiley Interdisciplinary Reviews: Water*, vol. 1, no. 4, pp. 323–329, 2014.

[147] J. Evaristo, S. Jasechko, and J. J. McDonnell, "Global separation of plant transpiration from groundwater and streamflow," *Nature*, vol. 525, pp. 91–94, 7567 Sep. 2015, ISSN: 14764687. DOI: 10.1038/nature14983.

[148] K. Beven, "How to make advances in hydrological modelling," *Hydrology Research*, vol. 50, pp. 1481–1494, 6 2019, ISSN: 22247955. DOI: 10.2166/nh.2019.134.

[149] S. J. Burges, *History of the stanford watershed model*, 2004. [Online]. Available: https://www.researchgate.net/publication/242220953.

[150] L. Sherman, "The relation of hydrographs of runoff to size and character of drainage-basins," *Eos, Transactions American Geophysical Union*, vol. 13, no. 1, pp. 332–339, 1932. DOI: https://doi.org/10.1029/TR013i001p00332.

[151] G. Kalinin and P. Milyukov, "On the computation of unsteady flow in open channels," *Met. Gidrol*, vol. 10, pp. 10–18, 1957.

[152] J. E. Nash, "The form of the instantaneous unit hydrograph," *International Association of Hydrological Sciences*, vol. 45, pp. 114–121, 1958.

[153] E. Todini, "Hydrological catchment modelling: Past, present and future," *Hydrology and Earth System Sciences*, vol. 11, pp. 468–482, 1 2007, ISSN: 16077938. DOI: 10.5194/hess-11-468-2007.

[154] P. C. Young, "Advances in real–time flood forecasting," *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 360, no. 1796, pp. 1433–1450, 2002. DOI: https://doi.org/10.1098/rsta.2002.1008.

[155] J. Amorocho and E. Hart, "A critique of current methods in hydrologic systems investigation," *Eos, Transactions American Geophysical Union*, vol. 45, no. 2, pp. 307–321, 1964. DOI: https://doi.org/10.1029/TR045i002p00307.

[156] R. A. Freeze and R. L. Harlan, "Blueprint for a physically-based, digitally-simulated hydrologicl response model," *Journal of Hydrology*, vol. 9, pp. 237–258, 1969. DOI: 10.1016/0022-1694(69)90020-1.

[157] C. T. Simmons, "Henry darcy (1803–1858): Immortalised by his scientific legacy," *Hydrogeology Journal*, vol. 16, no. 6, pp. 1023–1038, 2008. DOI: https://doi.org/10.1007/s10040-008-0304-3.

[158] L. A. Richards, "Capillary conduction of liquids through porous mediums," *physics*, vol. 1, no. 5, pp. 318–333, 1931. DOI: 10.1063/1.1745010.

[159] R. A. Freeze, "Streamflow generation," *Reviews of Geophysics*, vol. 12, no. 4, pp. 627–647, 1974. DOI: https://doi.org/10.1029/RG012i004p00627.

[160] R. A. Freeze, "Three-dimensional, transient, saturated-unsaturated flow in a groundwater basin," *Water Resources Research*, vol. 7, no. 2, pp. 347–366, 1971. DOI: https://doi.org/10.1029/WR007i002p00347.

[161] K. Beven, "Hillslope hydrographs by the finite element method," *Earth Surface Processes*, vol. 2, no. 1, pp. 13–28, 1977. DOI: https://doi.org/10.1002/esp.3290020103.

[162] R. A. Freeze, "Role of subsurface flow in generating surface runoff 1. base flow contributions to channel flow," *Water Resources Research*, vol. 8, pp. 609–623, 3 1972. DOI: 10.1029/wr008i003p00609.

[163] R. A. Freeze, "Role of subsurface flow in generating surface runoff: 2. upstream source areas," *Water Resources Research*, vol. 8, no. 5, pp. 1272–1283, 1972. DOI: https://doi.org/10.1029/WR008i005p01272.

[164] C. T. Simmons, P. Brunner, R. Therrien, and E. A. Sudicky, "Commemorating the 50th anniversary of the freeze and harlan (1969) blueprint for a physically-based, digitally-simulated hydrologic response model," *Journal of Hydrology*, vol. 584, May 2020, ISSN: 00221694. DOI: 10.1016/j.jhydrol.2019.124309.

[165] M. Abbott, J. Bathurst, J. Cunge, P. O'Connell, and J. Rasmussen, "An introduction to the european hydrological system — systeme hydrologique europeen, "she", 1: History and philosophy of a physically-based, distributed modelling system," *Journal of Hydrology*, vol. 87, no. 1, pp. 45–59, 1986, ISSN: 0022-1694. DOI: https://doi.org/10.1016/0022-1694(86)90114-9. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0022169486901149.

[166] M. Abbott, J. Bathurst, J. Cunge, P. O'Connell, and J. Rasmussen, "An introduction to the european hydrological system — systeme hydrologique europeen, "she", 2: Structure of a physically-based, distributed modelling system," *Journal of Hydrology*, vol. 87, no. 1, pp. 61–77, 1986, ISSN: 0022-1694. DOI: https://doi.org/10.1016/0022-1694(86)90115-0.

[167] K. Beven, "Changing ideas in hydrology - the case of physically-based models," *Journal of Hydrology*, vol. 105, pp. 157–172, 1989. DOI: 10.1016/0022-1694(89)90101-7.

[168] K. Beven and P. Germann, "Macropores and water flow in soils revisited," *Water Resources Research*, vol. 49, no. 6, pp. 3071–3092, 2013. DOI: https://doi.org/10.1002/wrcr.20156.

[169] J. W. Kirchner, "Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology," *Water Resources Research*, vol. 42, no. 3, 2006. DOI: https://doi.org/10.1029/2005WR004362.

[170] D. A. Woolhiser, "Search for physically based runoff model—a hydrologic el dorado?" *Journal of Hydraulic Engineering*, vol. 122, no. 3, pp. 122–129, 1996. DOI: 10.1061/(ASCE)0733-9429(1996)122:3(122).

[171] S. Fatichi, E. R. Vivoni, F. L. Ogden, *et al.*, "An overview of current applications, challenges, and future trends in distributed process-based models in hydrology," *Journal of Hydrology*, vol. 537, pp. 45–60, 2016, ISSN: 00221694. DOI: 10.1016/j.jhydrol.2016.03.026.

[172] M. P. Clark, M. F. Bierkens, L. Samaniego, *et al.*, "The evolution of process-based hydrologic models: Historical challenges and the collective quest for physical realism," *Hydrology and Earth System Sciences*, vol. 21, pp. 3427–3440, 7 2017, ISSN: 16077938. DOI: 10.5194/hess-21-3427-2017.

[173] G. Bloschl and M. Sivapalan, "Scale issues in hydrological modelling: A review," *HYDROLOGICAL PROCESSES*, vol. 9, pp. 251–290, 1995. DOI: 10.1002/hyp.3360090305.

[174] W. COLLISCHONN, D. ALLASIA, B. C. D. SILVA, and C. E. M. TUCCI, "The mgb-iph model for large-scale rainfall—runoff modelling," *Hydrological Sciences Journal*, vol. 52, pp. 878–895, 5 2007, ISSN: 0262-6667. DOI: 10.1623/hysj.52.5.878.

[175] K. J. Beven and M. J. Kirkby, "A physically based, variable contributing area model of basin hydrology," *Hydrological Sciences Bulletin*, vol. 24, pp. 43–69, 1 1979, ISSN: 03036936. DOI: 10.1080/02626667909491834.

[176] B. K. Horn, "Hill shading and the reflectance map," *Proceedings of the IEEE*, vol. 69, pp. 14–47, 1 1981, ISSN: 15582256. DOI: 10.1109/PROC.1981.11918.

[177] R. Barnes, C. Lehman, and D. Mulla, "Priority-flood: An optimal depression-filling and watershed-labeling algorithm for digital elevation models," *Computers and Geosciences*, vol. 62, pp. 117–127, Jan. 2014, ISSN: 00983004. DOI: 10.1016/j.cageo.2013.04.024.

[178] J. F. O'Callaghan and D. M. Mark, "The extraction of drainage networks from digital elevation data," *Computer Vision, Graphics, and Image Processing*, vol. 27, no. 2, p. 247, 1984, ISSN: 0734-189X. DOI: https://doi.org/10.1016/S0734-189X(84)80047-X.

[179] T. Freeman, "Calculating catchment area with divergent flow based on a regular grid," *Computers & Geosciences*, vol. 17, no. 3, pp. 413–422, 1991, ISSN: 0098-3004. DOI: https://doi.org/10.1016/0098-3004(91)90048-I.

[180] P. Quinn, K. Beven, P. Chevallier, and O. Planchon, "The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models," *Hydrological Processes*, vol. 5, no. 1, pp. 59–79, 1991. DOI: https://doi.org/10.1002/hyp.3360050106.

[181] K. Beven and E. F. Wood, "Catchment geomorphology and the dynamics of runoff contributing areas," *Journal of Hydrology*, vol. 65, pp. 139–158, 1983. DOI: 10.1016/0022-1694(83)90214-7.

[182] B. Ambroise, K. Beven, and J. Freer, "Toward a generalization of the topmodel concepts: Topographic indices of hydrological similarity," *Water Resources Research*, vol. 32, pp. 2135–2145, 7 Jul. 1996, Relaxation of Topmodel A3 to exponential, parabolic and linear functions., ISSN: 00431397. DOI: 10.1029/95WR03716.

[183] I. Iorgulescu and A. Musy, "Generalization of topmodel for a power law transmissivity profile," *Hydrological Processes*, vol. 11, pp. 1353–1355, 1997, Relaxation of Topmodel A3 to generalized power function.

[184] K. Beven and J. Freer, "A dynamic topmodel," *Hydrological Processes*, vol. 15, no. 10, pp. 1993–2011, 2001. DOI: https://doi.org/10.1002/hyp.252.

[185] A. Crave and C. Gascuel-Odoux, "The influence of topography on time and space distribution of soil surface water content," *Hydrological Processes*, vol. 11, pp. 203–210, 1997.

[186] C. D. Rennó, A. D. Nobre, L. A. Cuartas, *et al.*, "Hand, a new terrain descriptor using srtm-dem: Mapping terra-firme rainforest environments in amazonia," *Remote Sensing of Environment*, vol. 112, pp. 3469–3481, 9 Sep. 2008, ISSN: 00344257. DOI: 10.1016/j.rse.2008.03.018.

[187] A. D. Nobre, L. A. Cuartas, M. R. Momo, D. L. Severo, A. Pinheiro, and C. A. Nobre, "Hand contour: A new proxy predictor of inundation extent," *Hydrological Processes*, vol. 30, pp. 320–333, 2 Jan. 2016, ISSN: 10991085. DOI: 10.1002/hyp.10581.

[188] J. Gao, J. Holden, and M. Kirkby, "A distributed topmodel for modelling impacts of land-cover change on river flow in upland peatland catchments," *Hydrological Processes*, vol. 29, pp. 2867–2879, 13 Jun. 2015, ISSN: 10991085. DOI: 10.1002/hyp.10408.

[189] C. M. Souza, J. Z. Shimbo, M. R. Rosa, *et al.*, "Reconstructing three decades of land use and land cover changes in brazilian biomes with landsat archive and earth engine," *Remote Sensing*, vol. 12, 17 Sep. 2020, ISSN: 20724292. DOI: https://doi.org/10.3390/RS12172735.

[190] K. Beven, "The limits of splitting: Hydrology," *Science of the Total Environment*, vol. 183, pp. 89–97, 1-2 1996, ISSN: 00489697. DOI: 10.1016/0048-9697(95)04964-9.

[191] H. J. Tromp-van Meerveld and J. J. McDonnell, "Threshold relations in subsurface stormflow: 2. the fill and spill hypothesis," *Water Resources Research*, vol. 42, no. 2, 2006. DOI: https://doi.org/10.1029/2004WR003800.

[192] D. Janzen and J. J. McDonnell, "A stochastic approach to modelling and understanding hillslope runoff connectivity dynamics," *Ecological Modelling*, vol. 298, pp. 64–74, 2015, Complexity of Soils and Hydrology in Ecosystems, ISSN: 0304-3800. DOI: https://doi.org/10.1016/j.ecolmodel.2014.06.024.

[193] K. A. Johnson and R. S. Goody, "The original michaelis constant: Translation of the 1913 michaelis–menten paper," *Biochemistry*, vol. 50, no. 39, pp. 8264–8269, 2011. DOI: 10.1021/bi201284u.

[194] M. Potschin, R. Haines-Young, R. Fish, and R. Turner, *Routledge Handbook of Ecosystem Services* (Routledge Environment and Sustainability Handbooks). Taylor & Francis, 2016, ISBN: 9781317687030.

[195] Y. Harari, *Sapiens: A Brief History of Humankind*. HarperCollins, 2015, ISBN: 9780062316103.

[196] C. Lamont, *The Philosophy of Humanism*. Humanist Press, 1997, ISBN: 9780931779077.

[197] H. Gordon, *The History and Philosophy of Social Science*. Taylor & Francis, 2002, ISBN: 9781134863075.

[198] J. Edwards, "Harry helson's adaptation-level theory, happiness treadmills, and behavioral economics," *Journal of the History of Economic Thought*, vol. 40, no. 1, pp. 1–22, 2018. DOI: 10.1017/S1053837216001140.

[199] E. Diener, R. E. Lucas, and C. N. Scollon, "Beyond the hedonic treadmill: Revising the adaptation theory of well-being," in *The Science of Well-Being: The Collected Works of Ed Diener*, E. Diener, Ed. Dordrecht: Springer Netherlands, 2009, pp. 103–118, ISBN: 978-90-481-2350-6. DOI: 10.1007/978-90-481-2350-6_5. [Online]. Available: https://doi.org/10.1007/978-90-481-2350-6_5.

[200] P. Brickman, D. Coates, and R. Janoff-Bulman, "Lottery winners and accident victims: Is happiness relative?" *Journal of Personality and Social Psychology*, vol. 36, no. 8, pp. 917–927, 1978, ISSN: 0022-3514. DOI: 10.1037//0022-3514.36.8.917. [Online]. Available: http://dx.doi.org/10.1037//0022-3514.36.8.917.

[201] R. Nelson, *An Evolutionary Theory of Economic Change*. Harvard University Press, 1985, ISBN: 9780674041431.

[202] P. Bourgine, J. Lesourne, A. Orléan, *et al.*, *Evolutionary Microeconomics*. Springer Berlin Heidelberg, 2006, ISBN: 9783540285373.

[203] D. Pearce, "An intellectual history of environmental economics," *Annual Review of Environment and Resources*, vol. 27, no. Volume 27, 2002, pp. 57–81, 2002, ISSN: 1545-2050. DOI: https://doi.org/10.1146/annurev.energy.27.122001.083429. [Online]. Available: https://www.annualreviews.org/content/journals/10.1146/annurev.energy.27.122001.083429.

[204] P. Samuelson and W. Nordhaus, *Economics*. McGraw-Hill Education, 2009, ISBN: 9780073511290.

[205] H. Daly and J. Farley, *Ecological Economics, Second Edition: Principles and Applications*. Island Press, 2011, ISBN: 9781597269919.

[206] H. E. Daly, "On economics as a life science," *Journal of Political Economy*, vol. 76, no. 3, pp. 392–406, 1968. DOI: https://doi.org/10.1086/259412.

[207] P. P. Christensen, "Classical roots for a modern materials-energy analysis," *Ecological Modelling*, vol. 38, no. 1, pp. 75–89, 1987, Ecological Economics, ISSN: 0304-3800. DOI: https://doi.org/10.1016/0304-3800(87)90045-7.

[208] J. W. Forrester, G. W. Low, and N. J. Mass, "The debate on world dynamics: A response to nordhaus," *Policy Sciences*, vol. 5, no. 2, pp. 169–190, Jun. 1974, ISSN: 1573-0891. DOI: 10.1007/BF00148039. [Online]. Available: https://doi.org/10.1007/BF00148039.

[209] H. E. Daly, "Georgescu-roegen versus solow/stiglitz," *Ecological Economics*, vol. 22, no. 3, pp. 261–266, 1997, ISSN: 0921-8009. DOI: https://doi.org/10.1016/S0921-8009(97)00080-3.

[210] H. E. Daly, "Reply to solow/stiglitz," *Ecological Economics*, vol. 22, no. 3, pp. 271–273, 1997, ISSN: 0921-8009. DOI: https://doi.org/10.1016/S0921-8009(97)00086-4.

[211] D. Stoljar, "Physicalism," in *The Stanford Encyclopedia of Philosophy*, E. N. Zalta and U. Nodelman, Eds., Spring 2024, Metaphysics Research Lab, Stanford University, 2024.

[212] R. Van Gulick, "Consciousness," in *The Stanford Encyclopedia of Philosophy*, E. N. Zalta and U. Nodelman, Eds., Winter 2022, Metaphysics Research Lab, Stanford University, 2022.

[213] G. Caruso, "Skepticism About Moral Responsibility," in *The Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., Summer 2021, Metaphysics Research Lab, Stanford University, 2021.

[214] J. Forrester and J. Forrester, *World Dynamics*. Wright-Allen Press, 1973, ISBN: 9780262560184.

[215] R. Costanza and H. E. Daly, "Natural capital and sustainable development," *Conservation Biology*, vol. 6, pp. 37–46, 1 1992.

[216] R. S. d. Groot, "Environmental functions as a unifying concept for ecology and economics," *The Environmentalist*, 1987, ISSN: 02511088. DOI: 10.1007/BF02240292.

[217] G. Hardin, "The tragedy of the commons," *Science*, vol. 162, pp. 1243–1248, 3859 1968.

[218] N. Mankiw, *Principles of Microeconomics*. Nelson Thomson Learning, 2002, ISBN: 9780030340673.

[219] E. Schlager and E. Ostrom, "Property-rights regimes and natural resources: A conceptual analysis," *Land Economics*, vol. 68, no. 3, pp. 249–262, 1992. [Online]. Available: http://www.jstor.org/stable/3146375.

[220] S. Borghesi, "Water tradable permits: A review of theoretical and case studies," *Journal of Environmental Planning and Management*, vol. 57, no. 9, pp. 1305–1332, 2014. DOI: 10.1080/09640568.2013.820175.

[221] M. Mazaheri, J. Bonnin Roca, A. Markus, and B. Walrave, "Market-based instruments and sustainable innovation: A systematic literature review and critique," *Journal of Cleaner Production*, vol. 373, p. 133 947, 2022, ISSN: 0959-6526. DOI: https://doi.org/10.1016/j.jclepro.2022.133947.

[222] R. Haines-Young and M. Potschin, "The links between biodiversity, ecosystem services and human well-being," in *Ecosystem Ecology: A New Synthesis*, D. Raffaelli and C. Frid, Eds., Cambridge: Cambridge University Press, 2010, pp. 110–139.

[223] G. Daily, S. Postel, K. Bawa, *et al.*, *Nature's Services: Societal Dependence On Natural Ecosystems*. Island Press, 1997, ISBN: 9781597267755.

[224] M. E. Assessment, *Ecosystems and Human Well-being - Synthesis*. Washington, DC: Island Press, 2005, ISBN: 1-59726-040-1.

[225] R. Haines-Young and M. B. Potschin, *Common international classification of ecosystem services (cices) v5.1*, 2018. [Online]. Available: https://www.cices.eu.

[226] R. Costanza, "Ecosystem services: Multiple classification systems are needed," *Biological conservation*, vol. 141, no. 2, pp. 350–352, 2008.

[227] P. Kumar, *The Economics of Ecosystems and Biodiversity: Ecological and Economic Foundations* (TEEB - The Economics of Ecosystems and Biodiversity). Taylor & Francis, 2012, ISBN: 9781136538797.

[228] G. W. Partnership, "Integrated water resources management," Global Water Partnership, SE-10525 Stockholm, Sweden, Tech. Rep. 4, 2000. [Online]. Available: https://iwrmactionhub.org/resource/integrated-water-resources-management.

[229] J. Cassin, J. Matthews, and E. Gunn, *Nature-Based Solutions and Water Security: An Action Agenda for the 21st Century*. Elsevier, 2021, ISBN: 9780128198988.

[230] M. Smith, D. de Groot, and G. Bergkamp, *Pay - Establishing payments for watershed services*. Gland, Switzerland: IUCN, 2006, p. 109, ISBN: 978-2-8317-0958-1.

[231] Z. Liu, J. Ying, C. He, *et al.*, "Scarcity and quality risks for future global urban water supply," *Landscape Ecology*, vol. 39, no. 2, p. 10, Jan. 2024, ISSN: 1572-9761. DOI: https://doi.org/10.1007/s10980-024-01832-0.

[232] N. Dudley and S. Stolton, *Running Pure - The Importance of Forest Protected Areas to Drinking Water*. World Bank/WWF Alliance for Forest Conservation and Sustainable Use: World Bank/WWF Alliance for Forest Conservation and Sustainable Use, 2003, With contributions by Rachel Asante Owusu, Ahmet Birsel, David Cassells, José Courrau, Lawrence Hamilton, Sedat Kalem, Wang Luan Keng, Leonardo Lacerda, Yıldıray Lise, Stefano Pagiola, Sara Scherr, and Claudio Sericchio, ISBN: 2-88085-262-5.

[233] F. Nations, *Nature-Based Solutions for agricultural water management and food security* (Land and water discussion paper). Food and Agriculture Organization of the United Nations, 2019, ISBN: 9789251311257.

[234] C. Nesshöver, T. Assmuth, K. N. Irvine, *et al.*, "The science, policy and practice of nature-based solutions: An interdisciplinary perspective," *Science of the total environment*, vol. 579, pp. 1215–1227, 2017.

[235] B. M. Haddad and B. D. Solomon, *Dictionary of Ecological Economics: Terms for the New Millennium*. Edward Elgar Publishing, 2023.

[236] J. Salzman, G. Bennett, N. Carroll, A. Goldstein, and M. Jenkins, "The global status and trends of payments for ecosystem services," *Nature Sustainability*, vol. 1, pp. 136–144, 3 2018, ISSN: 23989629. DOI: 10.1038/s41893-018-0033-0. [Online]. Available: http://dx.doi.org/10.1038/s41893-018-0033-0.

[237] R. Fuentes-Llanillo, T. S. Telles, D. Soares Junior, T. R. de Melo, T. Friedrich, and A. Kassam, "Expansion of no-tillage practice in conservation agriculture in brazil," *Soil and Tillage Research*, vol. 208, pp. 104–877, 2021, ISSN: 0167-1987. DOI: https://doi.org/10.1016/j.still.2020.104877.

[238] R. A. Viani, H. Bracale, and D. Taffarello, "Lessons learned from thewater producer project in the atlantic forest, brazil," *Forests*, vol. 10, 11 Nov. 2019, Discussion of PES programs implementations.
PES values was 12 smaller than other costs!!!!
, ISSN: 19994907. DOI: 10.3390/f10111031.

[239] Brasil, *Lei nº 12651 de 25 de maio de 2012. dispõe sobre a proteção da vegetação nativa; altera as leis nºs 6.938, de 31 de agosto de 1981, 9.393, de 19 de dezembro de 1996, e 11.428, de 22 de dezembro de 2006; revoga as leis nºs 4.771, de 15 de setembro de 1965, e 7.754, de 14 de abril de 1989, e a medida provisória nº 2.166-67, de 24 de agosto de 2001; e dá outras providências*, 2012.

[240] Brasil, *LEI No 9.985, DE 18 DE JULHO DE 2000. Institui o Sistema Nacional de Unidades de Conservação da Natureza e dá outras providências*, 2000.

[241] B. Soares-Filho, R. Rajão, M. Macedo, *et al.*, "Cracking brazil's forest code," *Science*, vol. 344, no. 6182, pp. 363–364, 2014. DOI: https://doi.org/10.1126/science.1246663.

[242] P. H. Brancalion, L. C. Garcia, R. Loyola, R. R. Rodrigues, V. D. Pillar, and T. M. Lewinsohn, "Análise crítica da lei de proteção da vegetação nativa (2012), que substituiu o antigo código florestal: Atualizações e ações em curso," *Natureza & Conservação*, vol. 14, e1–e16, Apr. 2016, ISSN: 16790073. DOI: 10.1016/j.ncon.2016.03.004.

[243] C. Banks-Leite, R. Pardini, L. R. Tambosi, *et al.*, "Using ecological thresholds to evaluate the costs and benefits of set-asides in a biodiversity hotspot," *Science*, vol. 345, no. 6200, pp. 1041–1045, 2014. DOI: https://doi.org/10.1126/science.1255768.

[244] P. V. Caldwell, K. L. Martin, J. M. Vose, *et al.*, "Forested watersheds provide the highest water quality among all land cover types, but the benefit of this ecosystem service depends on landscape context," *Science of The Total Environment*, vol. 882, p. 163 550, 2023, ISSN: 0048-9697. DOI: https://doi.org/10.1016/j.scitotenv.2023.163550.

[245] E. Alcott, M. Ashton, and B. Gentry, *Natural and Engineered Solutions for Drinking Water Supplies: Lessons from the Northeastern United States and Directions for Global Watershed Management*. Taylor & Francis, 2013, ISBN: 9781466551640.

[246] L. Blanchard, B. Vira, and L. Briefer, "The lost narrative: Ecosystem service narratives and the missing wasatch watershed conservation story," *Ecosystem Services*, vol. 16, pp. 105–111, 2015, ISSN: 2212-0416. DOI: https://doi.org/10.1016/j.ecoser.2015.10.019.

[247] S. Naeem, J. C. Ingram, A. Varga, *et al.*, "Get the science right when paying for nature's services," *Science*, vol. 347, no. 6227, pp. 1206–1207, 2015. DOI: https://doi.org/10.1126/science.aaa1403.

[248] L. Tacconi, "Redefining payments for environmental services," *Ecological Economics*, vol. 73, pp. 29–36, Jan. 2012, ISSN: 09218009. DOI: 10.1016/j.ecolecon.2011.09.028.

[249] C. E. F. Young and L. B. d. Bakker, "Payments for ecosystem services from watershed protection: A methodological assessment of the oasis project in brazil," *Natureza a Conservacao*, vol. 12, pp. 71–78, 1 2014, ISSN: 16790073. DOI: 10.4322/natcon.2014.013.

[250] S. Lele, "Watershed services of tropical forests: From hydrology to economic valuation to integrated analysis," *Current Opinion in Environmental Sustainability*, vol. 1, pp. 148–155, 2 Dec. 2009, ISSN: 18773435. DOI: 10.1016/j.cosust.2009.10.007.

[251] S. Wunder, "Payments for environmental services: Some nuts and bolts," *CIFOR*, 42 2005, ISSN: 0854-9818. [Online]. Available: http://www.cifor.cgiar.org.

[252] S. Wunder, "The efficiency of payments for environmental services in tropical conservation," *Conservation Biology*, vol. 21, pp. 48–58, 1 2007. DOI: 10.1111/j.

[253] T. Kroeger, "The quest for the "optimal" payment for environmental services program: Ambition meets reality, with useful lessons," *Forest Policy and Economics*, vol. 37, pp. 65–74, Dec. 2013, ISSN: 13899341. DOI: 10.1016/j.forpol.2012.06.007.

[254] "Returns on investment in watershed conservation: Application of a best practices analytical framework to the rio camboriú water producer program, santa catarina, brazil," *Science of the Total Environment*, vol. 657, pp. 1368–1381, Mar. 2019, ISSN: 18791026. DOI: 10.1016/j.scitotenv.2018.12.116.

[255] A. Ullrich and M. Volk, "Application of the soil and water assessment tool (swat) to predict the impact of alternative management practices on water quality and quantity," *Agricultural Water Management*, vol. 96, pp. 1207–1217, 8 Aug. 2009, ISSN: 0378-3774. DOI: 10.1016/J.AGWAT.2009.03.010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378377409000808.

[256] C. Carvalho-Santos, J. P. Honrado, and L. Hein, "Hydrological services and the role of forests: Conceptualization and indicator-based analysis with an illustration at a regional scale," *Ecological Complexity*, vol. 20, pp. 69–80, Dec. 2014, ISSN: 1476-945X. DOI: 10.1016/J.ECOCOM.2014.09.001. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1476945X14000932.

[257] E. Martinez-Martinez, A. P. Nejadhashemi, S. A. Woznicki, and B. J. Love, "Modeling the hydrological significance of wetland restoration scenarios," *Journal of Environmental Management*, vol. 133, pp. 121–134, Jan. 2014, ISSN: 0301-4797. DOI: 10.1016/J.JENVMAN.2013.11.046. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0301479713007421.

[258] A. Daneshi, R. Brouwer, A. Najafinejad, M. Panahi, A. Zarandian, and F. F. Maghsood, "Modelling the impacts of climate and land use change on water security in a semi-arid forested watershed using invest," *Journal of Hydrology*, vol. 593, Feb. 2021, ISSN: 00221694. DOI: 10.1016/j.jhydrol.2020.125621.

[259] J. E. F. W. Lima, F. de Gois Aquino, T. A. Chaves, and C. Lorz, "Development of a spatially explicit approach for mapping ecosystem services in the brazilian savanna – mapes," *Ecological Indicators*, vol. 82, pp. 513–525, 2017, ISSN: 1470-160X. DOI: https://doi.org/10.1016/j.ecolind.2017.07.028.

[260] J. S. Sone, G. C. Gesualdo, P. A. Zamboni, *et al.*, "Water provisioning improvement through payment for ecosystem services," *Science of the Total Environment*, vol. 655, pp. 1197–1206, Mar. 2019, Evidence of hydrologic impacts of PES, ISSN: 18791026. DOI: 10.1016/j.scitotenv.2018.11.319.

[261] W. Francesconi, R. Srinivasan, E. Pérez-Miñana, S. P. Willcock, and M. Quintero, "Using the soil and water assessment tool (swat) to model ecosystem services: A systematic review," *Journal of Hydrology*, vol. 535, pp. 625–636, 2016, ISSN: 0022-1694. DOI: https://doi.org/10.1016/j.jhydrol.2016.01.034.

[262] M. Strauch, J. E. Lima, M. Volk, C. Lorz, and F. Makeschin, "The impact of best management practices on simulated streamflow and sediment load in a central brazilian catchment," *Journal of Environmental Management*, vol. 127, S24–S36, 2013, Integrated land-use and regional resource management – A cross-disciplinary dialogue on future perspectives for a sustainable development of regional resources, ISSN: 0301-4797. DOI: https://doi.org/10.1016/j.jenvman.2013.01.014.

[263] W. Cong, X. Sun, H. Guo, and R. Shan, "Comparison of the swat and invest models to determine hydrological ecosystem service spatial patterns, priorities and trade-offs in a complex basin," *Ecological Indicators*, vol. 112, p. 106 089, June 2019 2020, ISSN: 1470160X. DOI: 10.1016/j.ecolind.2020.106089. [Online]. Available: https://doi.org/10.1016/j.ecolind.2020.106089.