Information entropy of the carbon cycle

Holger Metzler · Carlos A. Sierra

Abstract Entropy and Friendlingstein!

Linear compartmental models are commonly used in different areas of science, particularly in modeling the cycles of carbon and other biogeochemical elements. The representation of these models as linear autonomous compartmental systems is useful for comparisons of different model structures and parameterizations on a macroscopic scale. The interpretation of such models as continuous-time Markov chains allows a deeper model analysis on a microscopic scale. In particular we can asses the uncertainty of a single particle's path as it travels through the system as described by path entropy and entropy rate. Path entropy measures the uncertainty of the entire path of a traveling particle from its entry into the system until its exit, whereas entropy rate measures the average uncertainty of the instantaneous future of a particle while it is in the system. We derive explicit formulas for these two types of entropy for compartmental systems in equilibrium based on Shannon information entropy and show how they can be used to assess the complexity of such models. Model complexity based on entropy can in turn be used to resolve the problems of equifinality and structural identification in the realm of model selection by means of the maximum entropy principle. We derive the entropy formulas by three/two different approaches, each one allowing different views on and insights into mass-balanced systems.

Keywords Entropy · Complexity · Carbon Cycle · Equifinality · Model selection · Compartmental system · Reservoir model

Holger Metzler
Department of Crop Production Ecology, Swedish University of Agricultural Sciences, Ulls väg 16, 756 51 Uppsala, Sweden, E-mail: holger.metzler@slu.de

Carlos A. Sierra
Max Planck Institute for Biogeochemistry, Hans-Knöll-Str. 10, 07745 Jena, Germany, E-mail: csierra@bgc-jena.mpg.de
Department of Ecology, Swedish University of Agricultural Sciences, Ulls väg 16, 756 51 Uppsala, Sweden

Mathematics Subject Classification (2000) 34A30 · 60J28 · 60K20 · 92B05

1 Introduction

− cite Agren paper on entropy, the papers Carlos sent by email on 2022-08-17
− RQ: Friendlingstein: model uncertainty, explainable by information entropy?
− CS models in C cycle
− compare models: A, TT
− take from entropy chapter motivation
− goal: entropy as information content and prediction uncertainty
− classical measures fail: also discussion?
− two/three entropy measures
− examples
− max. ent. principle?

Ages, transit times, and remaining lifetimes are diagnostic tools of open compartmental systems. They also help compare behavior and quality of different models. Nevertheless, structurally very different models might show very similar ages, transit times, and remaining lifetimes. If we are in the position to choose among such models, which is the one to select? By common sense, the answer is to select the least complex model and we can ask the question:

Can a model with fewer compartments reach the same complexity as a model with more compartments?
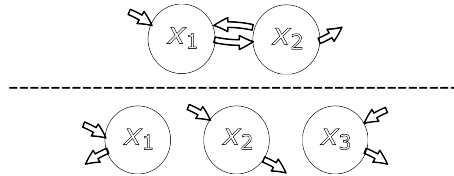

Fig. 1: Which model is more complex, the two-compartment model with feedback or the three-compartment model without feedback?

This leads to the problem of how to define complexity for compartmental systems in the first place (Figure 1). Walter and Contreras (1999, Chapter 23) ask a complexity measure/index to have at least the following natural properties:

(1) For a given structure, the index should have its greatest value when the flow rates are even (all the same).

(2) Given two structures with the same number of compartments and even flow rates, the index should have a larger value for the one with more nonzero flow rates.

Many common complexity measures of dynamical systems are closely related to the information content of the system and hence to some kind of entropy. Two examples are the topological entropy and the Kolmogorov-Sinai/metric entropy. However, linear autonomous compartmental systems are asymptotically stable. By Pesin's theorem (Pesin, 1977), both the metric- and the topological entropy vanish and cannot serve as a complexity measure here; we need a different concept.

Alternatively, we can interpret compartmental systems as weighted directed graphs. There are numerous different complexity measures for graphs. Dehmer and Mowshowitz (2011) provide a comprehensive overview of the history of graph entropy measures. Unfortunately, most of such entropy measures are based on the number of vertices, vertex degree, edges, or degree sequence (Trucco, 1956). Thus, they concentrate on only the structural information of the graph. The Markov chain maximization in the review paper does the same.

There are also graph theoretical measures that take edges and weights into account by using probability schemes. Their drawback is that the underlying meaning of complexity becomes difficult to interpret because the assigned probabilities seem somewhat arbitrary (Bonchev and Buck, 2005).

Since in the previous chapters we, amongst others, addressed the transit times of particles that travel through the system, we are naturally guided to a different approach. In terms of a single particle that moves through the system governed by a stochastic process, we can ask how difficult it is for us to guess the particle's current compartment, the particle's next compartment, and the particle's previous compartment. The more difficult it is to answer these three questions, the higher the complexity of the model should be. Consequently, a model's complexity should increase with the number of compartments, the number of fluxes leaving compartments, and the number of fluxes entering compartments. A weighted average of these numbers seems desirable. But how to choose the correct weights?

Since in open systems all material that enters the system also exits it later on, in this chapter we try to define a reasonable complexity measure for open compartmental systems based on the Shannon entropy (Shannon and Weaver, 1949) of the stochastic path covered by a particle from the moment of entering the system until the moment of leaving it. We further define a model's information content and touch the above mentioned problem of model selection, based on the concept of maximum entropy.

The manuscript is organized as follows. First we describe compartmental systems in equilibrium, then... and then we compute the three introduced entropy quantities for two carbon cycle models in equilibrium, depending on important parameter values.

2 Materials and methods

First, we introduce some basic notations and well-known properties of Shannon information entropy of random variables and stochastic processes. Then we present compartmental systems as a means to model material cycle systems that obey the law of mass balance. Then we consider such systems from a single-particle point of view and define the path of a single particle through the system along with its visited compartments, sojourn times, occupation times, and transit time. Based on these basic structures of a path, we compute three different types of entropy. For a better understanding, we provide a summary of the desirable relations among the three different types:

(1) As a particle travels through the system, it jumps a certain number of times to the next compartment until it finally jumps out of the system to the environmental compartment $d + 1$. Between two jumps, the particle resides in some compartment. Each jump comes with the uncertainties about which compartment will be next and how long will the particle stay there. The *entropy rate per jump* measures the average of these uncertainties with respect to the mean number of jumps.

(2) The travel of the particle takes a certain time. In each unit time interval before the particle leaves, uncertainties exist whether the particle jumps, where it jumps, and even how often it jumps. The mean of these uncertainties over the mean length of the travel interval is measured by the *entropy rate per unit time.*

(3) The *path entropy* measures the entire uncertainty about the particles travel through the system. We should be able to compute it if we multiply the mean entropy rate per jump by the mean number of jumps, and also if we multiply the entropy rate per unit time by the mean transit time.

2.1 Basic ideas of Shannon information entropy

We introduce basic concepts of information entropy along the lines of Cover and Thomas (2006). There are two concepts of entropy of a random variable, depending on whether the random variable has a discrete or a continuous distribution. Let Y be a discrete real-valued random variable with d distinct values y_i, $i = 1, 2, \ldots, d$ and probability mass function p such that $p(y_i) = p_i \geq 0$ and $\sum_{i=1}^{d} p_i = 1$ for a probability vector $\mathbf{p} = (p_i)_{i=1,2,\ldots,d}$. Then the *(Shannon information) entropy* of Y and $\mathbf{p}$ is defined by

$$\mathbb{H}(Y) = H(\mathbf{p}) = \mathbb{H}(p_1, p_2, \ldots, p_d)$$
$$= -\sum_{i=1}^{d} p(y_i) \log p(y_i) = -\sum_{i=1}^{d} p_i \log p_i = -\mathbb{E}\left[\log p(Y)\right],$$

where by convention $0 \log 0 := 0$. The *(differential) entropy* of a continuous real-valued random variable Y with probability density function f is defined

by

$$\mathbb{H}(Y) = - \int_{-\infty}^{\infty} f(y) \, \log f(y) \, \mathrm{d}y = -\mathbb{E}\left[\log f(Y)\right].$$

Throughout this manuscript we use Euler's number e as logarithmic base such that the unit of the entropy is nats, if not explicitly stated otherwise.

The entropy $\mathbb{H}(Y)$ of a random variable Y has two intertwined interpretations. On the one hand, $\mathbb{H}(Y)$ is a measure of uncertainty, i.e., a measure of how difficult it is to predict the outcome of a realization of Y. On the other hand, $\mathbb{H}(Y)$ is also a measure of the information content of Y, i.e., a measure of how much information we gain once we learn about the outcome of a realization of Y. It is important to note that, even though their definitions and information theoretical interpretations are quite similar, the Shannon- and the differential entropy have one main difference. The Shannon entropy is always nonnegative, whereas the differential entropy can have negative values. Consequently, the Shannon entropy is an absolute measure of information and makes sense in its own right, whereas the differential entropy is not an absolute information measure, is not scale-invariant, and makes sense only in comparison with the differential entropy of another random variable.

Panel A of Fig. 2 depicts the Shannon entropy of a Bernoulli random variable Y with $\mathbb{P}(Y = 1) = 1 - \mathbb{P}(Y = 0) = p$ with $p \in [0, 1]$. This random variable could represent the outcome of a coin toss. We can see that the entropy is low when p is close to 0 or 1. In these cases, we have some information that the coin is biased, and hence we have a preference if we guess the outcome. The entropy is maximum if the coin is fair ($p = 1/2$), since we have no additional information about the outcome of the coin toss. The Shannon entropy of Y is

$$\mathbb{H}(Y) = -p \, \log p - (1 - p) \, \log(1 - p).$$

Panel B of Fig. 2 shows the differential entropy of an exponentially distributed random variable $Y \sim \mathrm{Exp}(\lambda)$ with rate parameter $\lambda > 0$, probability density function $f(y) = \lambda \, e^{-\lambda y}$ for $y \geq 0$, and $\mathbb{E}[Y] = \lambda^{-1}$.

We can imagine it to represent the duration of stay of a particle in a well-mixed compartment in a linear autonomous compartmental system, where λ is the total outflow rate from the compartment. The higher the outflow rate is, the likelier is an early exit of the particle, and the easier it is to predict the moment of exit. Hence, the differential entropy decreases with increasing λ. It is given by

$$\mathbb{H}(Y) = 1 - \log \lambda.$$

According to Bad Dumitrescu (1988) and Girardin and Limnios (2003) we can extend the concept of entropy to continuous-time stochastic processes $Z = (Z_t)_{\geq 0}$. We first define the entropy of Z on a finite time interval $[0, T]$ by

$$\mathbb{H}_T(Z) = - \int f_T(z) \, \log f_T(z) \, \mathrm{d}\mu_T(z),$$

where f_T is the probability density function of $(Z_t)_{0 \leq t \leq T}$ with respect to some reference measure μ_T, if it exists. Note that by this definition we interpret the
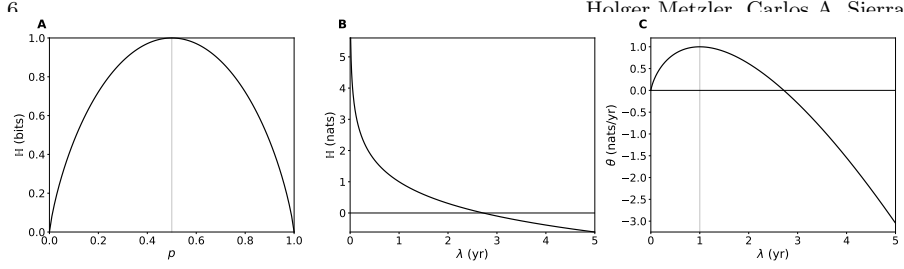
Fig. 2: A) Shannon entropy (logarithmic base 2) of a Bernoulli random variable depending on its success probability p. B) Differential entropy with logarithmic base e of an exponentially distributed random variable depending on its rate parameter λ. C) Entropy rate of a Poisson process with intensity rate λ.

entire stochastic process Z on the interval $[0, T]$ as a single random variable on the space

$$\{z = (z_t)_{t \in [0, T]} : z_t \in \mathbb{R}\}.$$

Then the *entropy rate* of Z is defined by

$$\theta(Z) = \lim_{T \to \infty} \frac{1}{T} \, \mathbb{H}_T(Z),$$

if the limit exists.

Let $Z \sim \text{Poi}(\lambda)$ be a Poisson process with intensity rate $\lambda > 0$ describing the moments of occurrence of certain events. The interarrival times of Z or the times between events are $\text{Exp}(\lambda)$-distributed, such that in the long run on average the time span between events has length λ^{-1}. The entropy of the interarrival times is given by $\mathbb{H}(\text{Exp}(\lambda)) = 1 - \log \lambda)$, and averaging it over the mean interarrival time gives the entropy rate of the Poisson process Z (Gaspard and Wang, 1993, Section 3.3), i.e.,

$$\mathbb{H}(Z) = \mathbb{H}(\text{Poi}(\lambda)) = \lambda \, (1 - \log \lambda).$$

This entropy rate increases with λ with values between 0 and 1, where it reaches its maximum, and then it decreases. This behavior is independent of the unit of λ, because it is based on the differential entropy of the exponential distribution and hence not scale-invariant. Consequently, it is no absolute measure of information content, but only useful in comparison to the entropy rate of other stochastic processes.

2.2 Linear autonomous compartmental systems

Mass-balanced flow of material into a system, within a system and out a a system that consists of several compartments can modeled by so-called compartmental systems (Anderson, 1983). Following Jacquez and Simon (1993), a *compartment* is an amount of some material that is kinetically homogeneous. By kinetically homogeneous we mean that the material of a compartment is

at all times homogeneous; any material entering the compartment is instantaneously mixed with the material already there. Hence compartments are always *well-mixed*. One way to describe compartmental systems is by the d-dimensional linear system of ordinary differential equations

$$\frac{\mathrm{d}}{\mathrm{d}\,t}\,\mathbf{x}(t) = \mathsf{B}\,\mathbf{x}(t) + \mathbf{u}, \quad t > 0, \tag{1}$$

with some initial condition $\mathbf{x}(0) = \mathbf{x}^0 \in \mathbb{R}^d$. The nonnegative vector $\mathbf{x}(t)$ describes the amount of material in the different compartments at time t, the nonnegative vector $\mathbf{u}$ is the vector of external inputs to the compartments, and the compartmental matrix $\mathsf{B} \in \mathbb{R}^{d \times d}$ describes the flux rates between the compartments and out of the system. To ensure that the system is mass balanced, we require the matrix B be compartmental, i.e.,

  (i) all off-diagonal entries are nonnegative;
  (ii) all diagonal entries are nonpositive;
(iii) all column sums are nonpositive.

The off-diagonal value B_{ij} is the flux rate from compartment j to compartment i, the absolute value of the negative diagonal value B_{jj} is the total rate of fluxes out of compartment j, and the nonnegative value $z_j = -\sum_{i=1}^{d} B_{ij}$ is the rate of the flux from compartment j out of the system. We require additionally that at least one column sum of B is strictly negative. This guarantees that the compartmental system is open in the sense that all material that enters the system will also leave the system at some point in time. The open compartmental system (1) has a unique steady-state or equilibrium compartment vector $\mathbf{x}^* = -\mathsf{B}^{-1}\,\mathbf{u}$ to which $\mathbf{x}(t)$ converges as $t \to \infty$, independently of the initial vector $\mathbf{x}^0$. In this manuscript, we are interested only in systems that have already reached the equilibrium such that $\mathbf{x}(t) = \mathbf{x}^*$ for all $t > 0$. Note that also nonlinear systems, in which $\mathsf{B}(\mathbf{x})$ or $\mathbf{u}(\mathbf{x})$ or both can depend on the system content $\mathbf{x}$, might reach a steady state $\mathbf{x}^* = -[\mathsf{B}(\mathbf{x}^*)]^{-1}\,\mathbf{u}(\mathbf{x}^*)$, in which case $\mathsf{B} = \mathsf{B}(\mathbf{x}^*)$ and $\mathbf{u} = \mathbf{u}(\mathbf{x}^*)$ are constant. A compartmental system in equilibrium given by Eq. (1) is fully characterized by $\mathbf{u}$ and B, and we denote it by $\mathcal{M} = \mathcal{M}(\mathbf{u}, \mathsf{B})$.

2.3 The one-particle perspective

While Eq. (1) describes the movement of bulk material through the system, compartmental systems in equilibrium can also be described probabilistically by considering the random path of a single particle through the system (Metzler and Sierra, 2018). If $X_t \in \mathcal{S} := \{1, 2, \ldots, d\}$ denotes the compartment in which the single particle is at time t, and $X_t = d+1$ if the particle has already left the system, then $X := (X_t)_{t \geq 0}$ is an absorbing continuous-time Markov chain (Norris, 1997) on $\widetilde{\mathcal{S}} := \mathcal{S} \cup \{d+1\}$. Its initial initial distribution is given by $\widetilde{\boldsymbol{\beta}} = (\beta_1, \beta_2, \ldots, \beta_d, 0)^T$, where $\boldsymbol{\beta} = \mathbf{u}/\|\mathbf{u}\|$ and $\beta_j = \mathbb{P}(X_0 = j)$ is the probability of the single particle to enter the system through compartment j. The

superscript T denotes the transpose of the vector/matrix and $\|\mathbf{u}\| = \sum_{i=1}^{d} |u_i|$ denotes the l_1-norm of the vector $\mathbf{u}$. The state-transition matrix of X is given by

$$\mathsf{Q} = \begin{pmatrix} \mathsf{B} & \mathbf{0} \\ \mathbf{z}^T & 0 \end{pmatrix}, \tag{2}$$

and thus

$$\mathbb{P}(X_t = i) = (e^{t\,\mathsf{Q}}\,\boldsymbol{\beta})_i = \sum_{j=1}^{d} (e^{t\,\mathsf{Q}})_{ij}\,\beta_j, \quad i \in \widetilde{\mathcal{S}},$$

is the probability of the particle to be in compartment i at time t if $i \in \mathcal{S}$ or that the particle has left the system if $i = d+1$. Here, $e^{t\,\mathsf{Q}}$ denotes the matrix exponential, and

$$\mathbb{P}(X_t = i \,|\, X_s = j) = (e^{(t-s)\,\mathsf{Q}})_{ij}, \quad s \le t, \quad i,j \in \widetilde{\mathcal{S}},$$

is the probability that X is in state i at time t given it was in state j at time s. Since the Markov chain X and the compartmental system in equilibrium given by Eq. (1) are equivalent, we can write

$$\mathcal{M} = \mathcal{M}(\mathbf{u}, \mathsf{B}) = \mathcal{M}(X).$$

2.4 The path of a single particle

A particle's path through the system from the moment of entering until the moment of exit can be described by

$$\mathcal{P}(X) = ((Y_1 = X_0, T_1), (Y_2, T_2), \ldots, (Y_{\mathcal{N}-1}, T_{\mathcal{N}-1}), Y_{\mathcal{N}} = d+1), \tag{3}$$

where X is the absorbing Markov chain associated to the particle's journey. The sequence $Y_1, Y_2, \ldots, Y_{\mathcal{N}-1} \in \mathcal{S}$ represents the successively visited compartments along with the associated sojourn times $T_1, T_2, \ldots, T_{\mathcal{N}-1}$, the random variable

$$\mathcal{N} := \inf \{n \in \mathbb{N} : Y_n = d+1\}$$

denotes the first hitting time of the *embedded jump chain* $Y := (Y_n)_{n=1,2,\ldots,\mathcal{N}}$ of X (Norris, 1997). With $\lambda_j := -Q_{jj}$ the one-step transition probabilities of Y are given by, for $i,j \in \widetilde{\mathcal{S}}$,

$$P_{ij} := \mathbb{P}(Y_{n+1} = i \,|\, Y_n = j) = \begin{cases} 0, & i = j \text{ or } \lambda_j = 0, \\ Q_{ij}/\lambda_j, & \text{else.} \end{cases} \tag{4}$$

We can also write $\mathsf{P} = \mathsf{Q}\,\mathsf{D}^{-1} + \mathsf{I}$, where

$$\mathsf{D} = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_d, \lambda_{d+1})$$

is the diagonal matrix with the diagonal entries of Q and I denotes the identity matrix of appropriate dimension. Defining the matrix $\mathsf{P}_{\mathsf{B}} = (P_{ij})_{i,j \in \mathcal{S}}$, then $\mathsf{M} := (\mathsf{I} - \mathsf{P}_{\mathsf{B}})^{-1}$ is the *fundamental matrix* of Y, with $\mathsf{I} \in \mathbb{R}^{d \times d}$ denoting the

identity matrix. The entry M_{ij} denotes the expected numbers of visits to compartment i given that the particle entered the system through compartment j. Consequently, the expected number of visits to compartment $i \in \mathcal{S}$ is given by

$$N_i = (\mathsf{M}\,\boldsymbol{\beta})_i = \left[(\mathsf{I} - \mathsf{P_B})^{-1}\,\boldsymbol{\beta}\right]_i = (\mathsf{D}\,\mathsf{B}^{-1}\,\boldsymbol{\beta})_i = \frac{\lambda_i\,x_i^*}{\|\mathbf{u}\|} \quad (5)$$

and the total expected number of jumps is given by

$$\mathbb{E}\left[\mathcal{N}\right] = \sum_{i=1}^{d}(\mathsf{M}\,\boldsymbol{\beta})_i + 1 = \sum_{i=1}^{d} N_i + 1,$$

where we take into account also the first jump into the system.

The last jump, $\mathcal{N}$, leads the particle out of the system such that at the moment of this last jump X takes on the value $d+1$. This last jump happens at the absorption time of the Markov chain X, which is defined as

$$\mathcal{T} := \inf\{t > 0 : X_t = d+1\}.$$

The absorption time is phase-type distributed (Neuts, 1981), $\mathcal{T} \sim \text{PH}(\boldsymbol{\beta}, \mathsf{B})$, with probability density function

$$f_{\mathcal{T}}(t) = \mathbf{z}^T\,e^{t\,\mathsf{B}}\,\boldsymbol{\beta}, \quad t \geq 0.$$

It can be shown (Metzler and Sierra, 2018, Section 3.2) that the mean or expected value of $\mathcal{T}$ equals the resident time (Sierra et al, 2016) of system (1) in equilibrium and is given by total stocks over total fluxes, i.e.,

$$\mathbb{E}\left[\mathcal{T}\right] = \frac{\|\mathbf{x}^*\|}{\|\mathbf{u}\|}.$$

Furthermore, it is obvious by construction that $\sum_{k=1}^{\mathcal{N}-1} T_k = \mathcal{T}$. If we denote by $\mathbb{1}_{\{A\}}$ the indicator function of the logical expression A, given by

$$\mathbb{1}_{\{A\}} = \begin{cases} 1, & A \text{ is true,} \\ 0, & \text{else,} \end{cases}$$

then $\sum_{k=1}^{\mathcal{N}-1} \mathbb{1}_{Y_k=j}\,T_k$ is the time that the particle spends in compartment j. This time is called *occupation time* of j and its mean is given by (Metzler and Sierra, 2018, Section 3.3)

$$\mathbb{E}\left[O_j\right] = \frac{x_j^*}{\|\mathbf{u}\|}, \quad (6)$$

which induces $\mathbb{E}\left[\mathcal{T}\right] = \sum_{j=1}^{d} \mathbb{E}\left[O_j\right]$.

2.5 Path entropy, entropy rate per unit time, entropy rate per jump

The path $\mathcal{P}$ given by Eq. (3) can be interpreted in three different ways. Each of these ways leads to a different interpretation of the path's entropy. First, we can look at $\mathcal{P}$ as the result of bookkeeping of the absorbing continuous-time Markov chain X, where as a sequence of pairs on the occasion of a jump we note down the old compartment of the traveling particle and the associated time the particle spent in this compartment. Second, we can consider the path as a discrete-time process. In each time step n, we choose randomly a new compartment Y_{n+1} and an associated sojourn time T_{n+1} of the particle in this compartment. Third, we can look at $\mathcal{P}$ as a single random variable with values in the space of all possible paths. Based on the latter interpretation we now derive the path entropy.

We are interested in the uncertainty/information content of the path $\mathcal{P}(X)$ of a single particle. Along the lines of Albert (1962), we construct a space $\wp$ that contains all possible paths that can be taken by a particle that runs through the system until it leaves. Let $\wp_n := (S \times \mathbb{R}_+)^n \times \{d + 1\}$ denote the space of paths that visit n compartments/states before ending up in the environmental compartment/absorbing state $d + 1$. By $\wp := \bigcup_{n=1}^{\infty} \wp_n$ denote the space of all eventually absorbed paths. Note that, since B is invertible, a path through the system is finite with probability 1. Let l denote the Lebesgue measure on $\mathbb{R}_+$ and c the counting measure on S. Furthermore, let σ_n be the sigma-finite product measure on $\wp_n$. It is defined by $\sigma_n := (c \otimes l)^n \otimes c$. Almost all sample functions of $(X_t)_{t \geq 0}$ can be represented as a point $p \in \wp$ (Doob, 1953, Chapter VI). Consequently, we can represent X by a finite-length path $\mathcal{P}(X) = ((Y_1, T_1), (Y, T_2), \ldots, (Y_n, T_n), Y_{n+1})$ for some $n \in \mathbb{N}$, where $Y_{n+1} = d + 1$.

For each set $W \subseteq \wp$ for which $W \cap \wp_n$ is σ_n-measurable for each $n \in \mathbb{N}$, we define $\sigma^*(W) := \sum_{n=1}^{\infty} \sigma_n(W \cap \wp_n)$. It is defined on the σ-field $\mathcal{F}^*$ which is the smallest σ-field containing all sets $W \subseteq \wp$ whose projection on $\mathbb{R}_+^n$ is a Borel set for each $n \in \mathbb{N}$. Let σ be a measure on *all* sample functions, defined for all subsets W whose intersection with $\wp$ is in $\mathcal{F}^*$. We define it by $\sigma(W) := \sigma^*(W \cap \wp)$.

Let $p = ((x_1, t_1), (x_2, t_2,), \ldots, (x_n, t_n), d + 1) \in \wp$ for some $n \in \mathbb{N}$. For $i \neq j$, we denote by $N_{ij}(p)$ the total number of path p's transitions from j to i and by $R_j(p)$ the total amount of time spent in j.

Lemma 1 *The probability density function of $\mathcal{P} = \mathcal{P}(X)$ with respect to σ is given by*

$$f_{\mathcal{P}}(p) = \beta_{x_1} \left(\prod_{j=1}^{d} \prod_{i=1, i \neq j}^{d+1} (Q_{ij})^{N_{ij}(p)} \right) \prod_{j=1}^{d} e^{-\lambda_j R_j(p)},$$
$$p = ((x_1, t_1), (x_2, t_2), \ldots, (x_n, t_n), d + 1) \in \wp.$$

Proof Let $x_1, x_2, \ldots, x_n \in S$, $x_{n+1} = d+1$, and $t_1, t_2, \ldots, t_n \in \mathbb{R}_p$. Since

$$\mathbb{P}((Y_1 = x_1, T_1 \leq t_1), \ldots, (Y_n = x_n, T_n \leq t_n), Y_{n+1} = d+1)$$

$$= \mathbb{P}(Y_{n+1} = d+1 \mid Y_n = x_n) \prod_{k=1}^{n} \mathbb{P}(Y_k = x_k, T_k \leq t_k \mid Y_{k-1} = x_{k-1})$$

$$= P_{d+1,x_n} \left[\prod_{k=2}^{n} P_{x_k x_{k-1}} \left(1 - e^{-\lambda_{x_k} t_k}\right) \right] \beta_{x_1} \left(1 - e^{-\lambda_{x_1} t_1}\right)$$

$$= \int_{\mathbb{T}_n} \beta_{x_1} \prod_{k=1}^{n} Q_{x_{k+1} x_k} e^{-\lambda_{x_k} \tau_k} \, \mathrm{d}\tau_1 \mathrm{d}\tau_2 \cdots \mathrm{d}\tau_n$$

with $\mathbb{T}_n = \{(\tau_1, \tau_2, \ldots, \tau_n) \in \mathbb{R}_+^n : 0 \leq \tau_1 \leq t_1, 0 \leq \tau_2 \leq t_2, \ldots, 0 \leq \tau_n \leq t_n\}$, the probability density function of $\mathcal{P} = \mathcal{P}(x)$ with respect to σ is given by

$$f_{\mathcal{P}}(p) = \beta_{x_1} \prod_{k=1}^{n} Q_{x_{k+1} x_k} e^{-\lambda_{x_k} t_k},$$

$$p = ((x_1, t_1), (x_2, t_2), \ldots, (x_n, t_n), d+1) \in \wp.$$

The term $Q_{x_{k+1} x_k} = Q_{ij}$ enters exactly $N_{ij}(p)$ times. Furthermore,

$$\prod_{k=1}^{n} e^{-\lambda_{x_k} t_k} = \prod_{k=1}^{n} \prod_{j=1}^{d} \mathbb{1}_{\{x_k = j\}} e^{-\lambda_j t_k} = \prod_{j=1}^{d} e^{-\lambda_j \sum_{k=1}^{n} \mathbb{1}_{\{x_k = j\}} t_k} = \prod_{j=1}^{d} e^{-\lambda_j R_j(p)}.$$

We make the according substitutions and the proof is finished.

The *entropy of the absorbing continuous-time Markov chain X* is equal to the entropy on the random but finite time horizon $[0, \mathcal{T}]$, which in turn equals the entropy of a single particle's path $\mathcal{P}$ through the system.

Theorem 1 *The entropy of the absorbing continuous-time Markov chain X is given by*

$$\mathbb{H}(X) = \mathbb{H}(\mathcal{P})$$

$$= -\sum_{i=1}^{d} \beta_i \log \beta_i$$

$$+ \sum_{j=1}^{d} \frac{x_j^*}{\|\mathbf{u}\|} \left[\sum_{i=1, i \neq j}^{d} B_{ij} \left(1 - \log B_{ij}\right) + z_j \left(1 - \log z_j\right) \right]. \tag{7}$$

Proof Let X have the finite path representation

$$\mathcal{P} = \mathcal{P}(X) = ((Y_1, T_1), (Y_2, T_2), \ldots, (Y_n, T_n), d+1)$$

for some $n \in \mathbb{N}$, and denote by $f_{\mathcal{P}}$ its probability density function. Then, by Lemma 1,

$$-\log f_{\mathcal{P}}(\mathcal{P}) = -\log \beta_{Y_1} - \sum_{j=1}^{d} \sum_{i=1, i\neq j}^{d+1} N_{ij}(\mathcal{P}) \log Q_{ij} + \sum_{j=1}^{d} \lambda_j R_j(\mathcal{P}).$$

We compute the expectation and get

$$\begin{aligned}
\mathbb{H}(X) = \mathbb{H}(\mathcal{P}) &= -\mathbb{E}\left[\log f_{\mathcal{P}}(\mathcal{P})\right] \\
&= -\mathbb{E}\left[\log \beta_{Y_1}\right] - \sum_{j=1}^{d} \sum_{i=1, i\neq j}^{d+1} \mathbb{E}\left[N_{ij}(\mathcal{P})\right] \log Q_{ij} + \sum_{j=1}^{d} \lambda_j \mathbb{E}\left[R_j(\mathcal{P})\right] \\
&= \mathbb{H}(Y_1) + \sum_{j=1}^{d} \lambda_j \mathbb{E}\left[R_j(\mathcal{P})\right] - \sum_{j=1}^{d} \sum_{i=1, i\neq j}^{d+1} \mathbb{E}\left[N_{ij}(\mathcal{P})\right] \log Q_{ij}.
\end{aligned}$$

Obviously, $\mathbb{E}\left[R_j(\mathcal{P})\right] = \mathbb{E}\left[O_j\right] = x_j^*/\|\mathbf{u}\|$ is the mean occupation time of compartment $j \in S$ by X. Furthermore, for $i \in \widetilde{S}$ and $j \in S$ such that $i \neq j$, by Eqs. (5) and (4),

$$\mathbb{E}\left[N_{ij}(\mathcal{P})\right] = \mathbb{E}\left[N_j(\mathcal{P})\right] P_{ij} = \begin{cases} \frac{x_j^*}{\|\mathbf{u}\|} B_{ij}, & i \leq d, \\ \frac{x_j^*}{\|\mathbf{u}\|} z_j, & i = d+1. \end{cases}$$

Together with $\lambda_j = \sum_{i=1, i\neq j}^{d} B_{ij} + z_j$, we obtain

$$\begin{aligned}
\mathbb{H}(X) &= \mathbb{H}(Y_1) + \sum_{j=1}^{d} \frac{x_j^*}{\|\mathbf{u}\|} \left[\left(\sum_{i=1, i\neq j}^{d} B_{ij} + z_j \right) - \sum_{i=1, i\neq j}^{d} B_{ij} \log B_{ij} - z_j \log z_j \right] \\
&= -\sum_{i=1}^{d} \beta_i \log \beta_i + \sum_{j=1}^{d} \frac{x_j^*}{\|\mathbf{u}\|} \left[\sum_{i=1, i\neq j}^{d} B_{ij} \left(1 - \log B_{ij}\right) + z_j \left(1 - \log z_j\right) \right].
\end{aligned}$$

By some simple substitutions and rearrangements, we obtain two representations of $\mathbb{H}(X) = \mathbb{H}(\mathcal{P})$ that are easy to interpret.

Proposition 1 *We The entropy of the absorbing continuous-time Markov chain X is also given by*

$$\mathbb{H}(X) = \mathbb{H}(\boldsymbol{\beta}) + \sum_{j=1}^{d} \mathbb{E}\left[O_j\right] \left(\sum_{i=1, i\neq j}^{d} \theta(\mathrm{Poi}(B_{ij})) + \theta(\mathrm{Poi}(z_j)) \right) \qquad (8)$$

and

$$\begin{aligned}
\mathbb{H}(X) = \mathbb{H}(\boldsymbol{\beta}) \\
+ \sum_{j=1}^{d} \mathbb{E}\left[N_j\right] \left(\mathbb{H}(\mathrm{Exp}(\lambda_j)) + \mathbb{H}(P_{1,j}, P_{2,j}, \ldots, P_{d,j}, P_{d+1,j}) \right).
\end{aligned} \qquad (9)$$

Proof By virtue of Eq. (8) we replace $x_j^*/\|\mathbf{u}\|$ by $\mathbb{E}\left[O_j\right]$ in Eq. (7) and take into account that the entropy rate of a Poisson process with intensity rate λ equals $\lambda\left(1 - \log\lambda\right)$ to prove Eq. (8). To prove Eq. (9) we use Eq. (5) to replace $x_j^*/\|\mathbf{u}\|$ in Eq. (7) by $\mathbb{E}\left[N_j\right]/\lambda_j$ and obtain

$$\mathbb{H}(X) = -\sum_{i=1}^{d} \beta_i \, \log\beta_i$$
$$+ \sum_{j=1}^{d} \mathbb{E}\left[N_j\right]\left(\left(1 - \log\lambda_j\right) - \sum_{i=1,i\neq j}^{d} \frac{B_{ij}}{\lambda_j} \log\frac{B_{ij}}{\lambda_j} - \frac{z_j}{\lambda_j} \log\frac{z_j}{\lambda_j}\right).$$

Here, $(1 - \log\lambda_j)$ is the entropy of an exponential random variable with rate parameter λ_j. Using the definition (4) of P_{ij} we replace B_{ij}/λ_j for $i \in \mathcal{S}$ and z_j/λ_j by $P_{d+1,j}$ ad finish the proof.

We now define the *path entropy of the compartmental system in equilibrium* $\mathcal{M}$, given by Eq. (1), as the path entropy of its associated continuous-time Markov chain X, i.e.

$$\mathbb{H}(\mathcal{M}) := \mathbb{H}(X) = \mathbb{H}(\mathcal{P}).$$

For a one-dimensional compartmental system $\mathcal{M}_\lambda$ in equilibrium with rate $\lambda > 0$ and positive external input given by

$$\frac{\mathrm{d}}{\mathrm{d}t} x(t) = -\lambda\, x(t) + u, \quad t > 0, \tag{10}$$

the entropy of the initial distribution vanishes, and we obtain

$$\mathbb{H}(\mathcal{M}_\lambda) = \frac{x^*}{u}\,\lambda\left(1 - \log\lambda\right) = \frac{1}{\lambda}\,\lambda\left(1 - \log\lambda\right) = 1 - \log\lambda,$$

which which equals the differential entropy $1 - \log\lambda$ of the exponentially distributed mean transit time $\mathcal{T}_\lambda \sim \mathrm{Exp}(\lambda)$, reflecting that the only uncertainty of the particle's path in a one-pool system is the time of the particle's exit. The exponential distribution with rate parameter λ is the distribution of the interarrival time of a Poisson process wit intensity rate λ. Hence, we can interpret $\mathbb{H}(\mathcal{M}_\lambda) = \lambda^{-1}\,\lambda\left(1 - \lambda\right)$ as the instantaneous Poisson entropy $\lambda\left(1 - \lambda\right)$ multiplied with the expected duration $\mathbb{E}\left[\mathcal{T}\right] = \lambda^{-1}$ of the particle's stay in the system.

Migrating to a d-dimensional system, we can interpret $\mathbb{H}(\mathcal{M})$ as the entropy of a continuous-time process in the light of Eq. (8) and as the entropy of a discrete-time process in the light of Eq. (9). In both interpretation represents the first term $\mathbb{H}(\boldsymbol{\beta}) = \mathbb{H}(X_0) = \mathbb{H}(Y_1)$ the uncertainty of the first pool through which the particle enters the system. In the continuous-time interpretation, the uncertainty of the subsequent travel is the weighted average of the overlapping Poisson processes describing the instantaneous uncertainty of possible jumps of the particle inside the system $(\theta(\mathrm{Poi}(B_{ij})))$ and out of the system $(\theta(\mathrm{Poi}(z_j)))$, where the weights are the expected occupation times

of the different compartments. In the discrete-time interpretation, the subsequent travel's uncertainty is the average of uncertainties associated to each pool, weighted by the number of visits to the respective pools. The uncertainties associated to each pool comprises the uncertainty of the length of the stay in the pool ($\mathbb{H}(\mathrm{Exp}(\lambda_j))$) and the uncertainty of where to jump afterwards ($\mathbb{H}(P_{1,j}, \ldots, P_{d+1,j})$).

The two interpretations of the path entropy $\mathbb{H}(\mathcal{M})$ (as a time-continuous or time-discrete process) motivate two different entropy rates. The *entropy rate per unit time* is defined as

$$\theta(\mathcal{M}) = \frac{\mathbb{H}(\mathcal{M})}{\mathbb{E}\left[\mathcal{T}\right]}$$

and the *entropy rate per jump* as

$$\theta_J(\mathcal{M}) = \frac{\mathbb{H}(\mathcal{M})}{\mathbb{E}\left[\mathcal{N}\right]}.$$

2.6 Maximum entropy principle

2.7 Structural model identification

3 Results/Examples

3.1 Simple examples

From Table 1 we can see that depending on the connections between compartments smaller systems can have path entropy and entropy rates than bigger systems, even systems with more compartments can theoretically have higher entropy. Furthermore, we see that the system with the highest path entropy does neither have the highest entropy rate per unit time nor per jump. Adding connections to a system, one would expect higher path entropy, but the path entropy might increase because the new connection potentially provides a faster way out out the system.

3.2 A linear autonomous global carbon cycle model

We consider the global carbon cycle model introduced by Emanuel et al (1981) (Fig. 3). The model comprises five compartments: non-woody tree parts x_1, woody tree parts x_2, ground vegetation x_3, detritus/decomposers x_4, and active soil carbon x_5. We introduce an environmental rate modifier ξ which controls the speed of the system. This parameter could potentially increase and speed up the system with increasing global Earth surface temperature. Since the model is considered to be in equilibrium, the initial state is negligible and, the model is given by

$$\frac{\mathrm{d}}{\mathrm{d}\,t}\mathbf{x}(t) = \mathsf{B}\,\mathbf{x}(t) + \mathbf{u}, \quad t > 0,$$

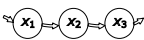| Structure	$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}(t)$	θ_J	$\mathbb{E}[\mathcal{N}]$	θ	$\mathbb{E}[\mathcal{T}]$	$\mathbb{H}(\mathcal{P})$
x_1	$-\lambda\,x + 1$	$0.5\,(1-\log\lambda)$	2.00	$\lambda\,(1-\log\lambda)$	$1/\lambda$	$1-\log$
$x_1 \to x_2$	$\begin{pmatrix} -1 & 0 \\ 1 & -1 \end{pmatrix} x + \begin{pmatrix} 1 \\ 0 \end{pmatrix}$	0.67	3.00	1.00	2.00	2.00
$x_1 \; x_2$	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} x + \begin{pmatrix} 1 \\ 1 \end{pmatrix}$	0.85	2.00	1.69	1.00	1.69
$x_1 \rightleftarrows x_2$	$\begin{pmatrix} -1 & 0.5 \\ 1 & -1 \end{pmatrix} x + \begin{pmatrix} 1 \\ 0 \end{pmatrix}$	1.08	5.00	1.35	<u>4.00</u>	<u>5.39</u>
$x_1 \rightleftarrows x_2$	$\begin{pmatrix} -1 & 0.5 \\ 0.5 & -1 \end{pmatrix} x + \begin{pmatrix} 1 \\ 1 \end{pmatrix}$	<u>1.36</u>	3.00	2.04	2.00	4.08
$x_1 \to x_2 \to x_3$	$\begin{pmatrix} -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} x + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$	0.75	4.00	1.00	3.00	3.00
$x_1 \; x_2 \; x_3$	$\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} x + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$	1.05	2.00	<u>2.10</u>	1.00	2.10

Table 1: Overview of different entropies of simple models with different structures. The columns from left to right represent a schematic of the model, its mathematical representation, its entropy rate per jump, its entropy rate per unit time, its mean transit time, and its path entropy. Underlined numbers are the highest values per column. The two gray rows emphasize the examples of the model structures considered at the beginning of the chapter in Fig. 1. include also mean number of jumps, to save space use $\dot{x}$, introduce in text
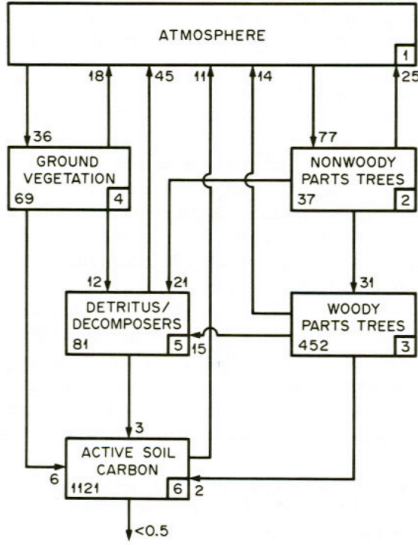

Fig. 3: Schematic of the linear autonomous global carbon cycle model in steady state introduced by Emanuel et al (1981). The model comprises five compartments: non-woody tree parts x_1 (2; $37\,\mathrm{Pg\,C}$), woody tree parts x_2 (3; $452\,\mathrm{Pg\,C}$), ground vegetation x_3 (4; $69\,\mathrm{Pg\,C}$), detritus/decomposers x_4 (5; $81\,\mathrm{Pg\,C}$), and active soil carbon x_5 (6; $1,121\,\mathrm{Pg\,C}$). The atmosphere (1) is considered to be outside of the modeled system but provides the system with external inputs and receives external outputs from it. Numbers next to arrows indicate fluxes between compartments in $\mathrm{Pg\,C\,yr^{-1}}$. (Figure extracted from Emanuel et al 1981)

where the input vector is given by

$$\mathbf{u} = (77.00;\ 0.00;\ 36.00;\ 0.00;\ 0.00)^T \,\mathrm{Pg\,C\,yr^{-1}}$$

and the compartmental matrix by

$$\mathsf{B} = \begin{pmatrix} -77/37 & 0 & 0 & 0 & 0 \\ 31/37 & -31/452 & 0 & 0 & 0 \\ 0 & 0 & -36/69 & 0 & 0 \\ 21/37 & 15/452 & 12/69 & -48/81 & 0 \\ 0 & 2/452 & 6/69 & 3/81 & -11/1121 \end{pmatrix} \mathrm{yr^{-1}},$$

where the numbers are chosen as in Thompson and Randerson (1999). The input vector is expressed in units of petagrams of carbon per year ($\mathrm{Pg\,C\,yr^{-1}}$) and the fractional transfer coefficients in units of per year ($\mathrm{yr^{-1}}$). Because B is a lower triangular matrix, the model contains no feedbacks. For every value of ξ the system has a different steady state (Fig.4, panel A). The higher the value of ξ, the faster is the system, which makes the mean transit time (panel B) decrease and because of shorter paths also the path entropy (panel D) decreases. Since ξ has no impact on the structure of the model, the mean number of jumps (panel C) remains unaffected. Nevertheless, the entropy rate per jump (panel F) decreases with increasing ξ because the path entropy of the system decreases. The entropy rate per unit time increases until $\xi \approx 6$, while the mean transit time decreases faster than the path entropy, and then trend turns around and the entropy rate per unit time decreases (panel E). Orange lines in panel D and E show the respective entropy values for a one-pool system with the same mean transit time. The blue and orange lines intersect at $\xi \approx 4.31$.

3.3 A nonlinear autonomous soil organic matter decomposition model

Consider the nonlinear autonomous compartmental system

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}(t) = \mathsf{B}(\mathbf{x}(t))\,\mathbf{x}(t) + \mathbf{u}, \quad t > 0, \tag{11}$$

where $\mathsf{B}: \mathbb{R}^d \to \mathbb{R}^{d \times d}$ is a matrix-valued mapping. In this setup, the fractional transfer coefficients are not constant but depend on the system's content.

Assume now that system (11) is in a steady state $\mathbf{x}^*$. From $\mathrm{d}\mathbf{x}^*/\mathrm{d}t = 0$ follows that the compartment contents x_j^* do not change in time, and the mapping B turns into a matrix with constant coefficients. Hence, if we assume the nonlinear autonomous compartmental system (11) to be in a steady state, we can treat it as a linear autonomous compartmental system.

As an example, consider the nonlinear two-compartment carbon cycle model described by Wang et al (2014) (Fig. 5). We denote by C_s and C_b soil organic carbon and soil microbial biomass ($\mathrm{g\,C\,m^{-2}}$), respectively, by ε the carbon use efficiency or fraction of assimilated carbon that is converted into microbial biomass (unit-less), by μ_b the turnover rate of microbial biomass per year
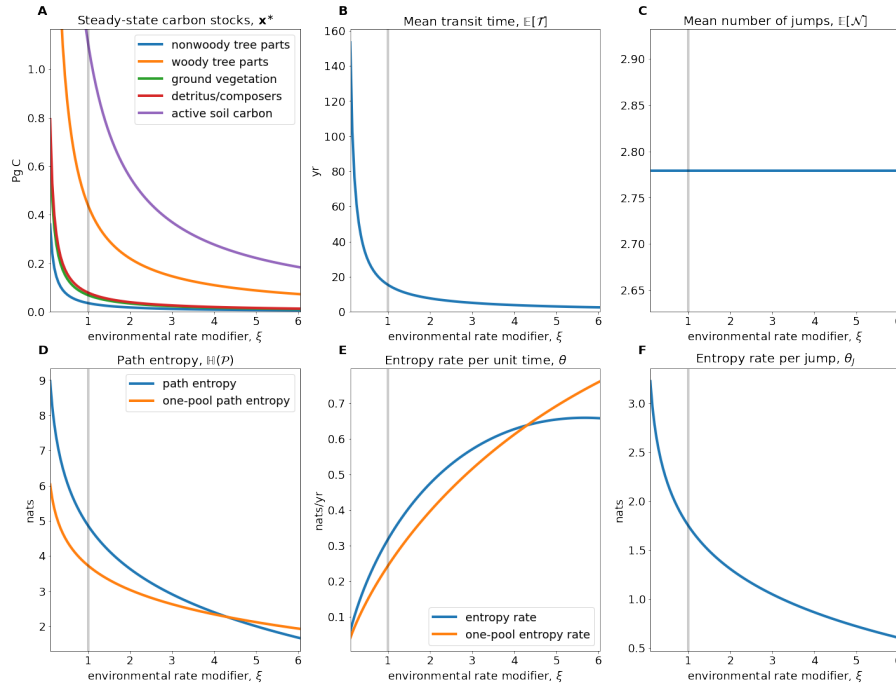
Fig. 4: A) Equilibrium carbon stocks. B)–F) Entropy related quantities of the global carbon cycle model introduced by Emanuel et al (1981) in dependence on the environmental rate coefficient ξ (blue lines). Orange lines correspond to the quantities derived from a one-pool model with the same speed. Vertical gray lines show $\xi = 1$, the original speed of the model.
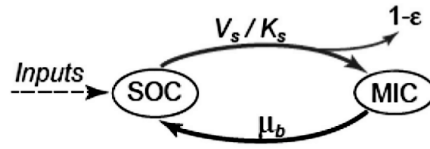

Fig. 5: Scheme of the nonlinear autonomous carbon cycle model introduced by Wang et al (2014). The two compartments C_s and C_b are here denoted by SOC (substrate organic carbon) and MIC (microbial biomass carbon), the external input flux F_{NPP} is denoted by *Inputs*, the maximum rate of soil carbon assimilation by V_s, the half saturation constant by K_s, the carbon use efficiency by ε, and the turnover rate of microbial biomass by μ_b, respectively. (Figure extracted from Wang et al (2014))

(yr^{-1}), by F_{NPP} the carbon influx into soil (g C m^{-2} yr^{-1}), and by V_s and K_s the maximum rate of soil carbon assimilation per unit microbial biomass per year (yr^{-1}) and the half-saturation constant for soil carbon assimilation by microbial biomass (g C m^{-2}), respectively. Then, we can describe the model by

$$\frac{\mathrm{d}}{\mathrm{d}\,t}\begin{pmatrix} C_s \\ C_b \end{pmatrix} = \begin{pmatrix} -\lambda(\mathbf{x}) & \mu_b \\ \varepsilon\lambda(\mathbf{x}) & -\mu_b \end{pmatrix}\begin{pmatrix} C_s \\ C_b \end{pmatrix} + \begin{pmatrix} F_{\text{NPP}} \\ 0 \end{pmatrix}.$$

The matrix B depends on $\mathbf{x} = (C_s, C_b)^T$ through λ's dependence on $\mathbf{x}$, which is given by

$$\lambda(\mathbf{x}) = \frac{C_b V_s}{C_s + K_s}. \tag{12}$$

Steady-state formulas for the compartment contents can be computed as

$$C_s^* = \frac{K_s}{\frac{V_s\varepsilon}{\mu_b} - 1} \quad \text{and} \quad C_b^* = \frac{F_{\text{NPP}}}{\mu_b\left(-1 + \frac{1}{\varepsilon}\right)}.$$

From Wang et al (2014) we take the parameter values $F_{\text{NPP}} = 345.00\,\text{g C m}^{-2}\,\text{yr}^{-1}$, $\mu_b = 4.38\,\text{yr}^{-1}$, and $K_s = 53{,}954.83\,\text{g C m}^{-2}$. Since the description of V_s is missing in the original publication, we let it be equal to $59.13\,\text{yr}^{-1}$ to, for $\varepsilon = 0.39$, approximately meet the given steady-state compartment contents $C_s^* = 12{,}650.00\,\text{g C m}^{-2}$ and $C_b^* = 50.36\,\text{g C m}^{-2}$. We leave the carbon use efficiency ε as a free parameter.

With the given parameters, the steady-state transfer matrix $\mathsf{B} = \mathsf{B}(\mathbf{x}^*)$ and the input vector $\mathbf{u}$ are given by

$$\mathsf{B} = \begin{pmatrix} -\frac{59.13\,C_b^*}{C_s^* + 53{,}954} & 4.38 \\ \varepsilon\frac{59.13\,C_b^*}{C_s^* + 53{,}954} & -4.38 \end{pmatrix}\text{yr}^{-1} \quad \text{and} \quad \mathbf{u} = \begin{pmatrix} 345.00 \\ 0.00 \end{pmatrix}\text{g C m}^{-2}\,\text{yr}^{-1},$$

In contrast to the system from the first example, this system exhibits a feedback. This feedback results from dead soil microbial biomass being considered as new soil organic matter. The feedback can also be recognized by noting that B is not triangular. For every value of ε the system has a different steady state (Fig. 6, panel A). The higher the value of ε, the lower the equilibrium substrate organic carbon and the higher the microbial biomass carbon. Caused by the model's nonlinearity expressed in Eq. (12), the system speed increases and the mean transit time goes down (panel B). At the same time, higher carbon use efficiency increases the probability of the C atom to be reused more often, hence the mean number of jumps increases (panel C), making the entropy rate per jump decrease (panel F). Even though, the average paths become shorter, with increasing carbon use efficiency the path entropy increases as well. Until $\varepsilon = 0.5$ this has two reasons. First, the uncertainty of where to jumping out of C_s increases, this uncertainty decreases then for $\varepsilon > 0.5$. Second, the rate $-B_{11}$ of leaving the substrate pool is increasing and smaller than 1. The corresponding Poisson process reaches its maximum entropy at an intensity rate equal to 1 (Fig. 2, panel C), here at $\varepsilon \approx 0.926$. This is also reflected in the

entropy rate per unit time (panel D). The maximum does not exactly occur at $\varepsilon = 0.926$, because the times that the particle stays in the different pools also depends on ε. For $\varepsilon > 0.926$ both the path entropy and the entropy rate rapidly decline as both the jump uncertainty and the Poisson entropy rate inf C_s decline sharply.

faster is the system, which makes the mean transit time (panel B) decrease and because of shorter paths also the path entropy (panel D) decreases. Since ξ has no impact on the structure of the model, the mean number of jumps (panel C) remains unaffected. Nevertheless, the entropy rate per jump (panel F) decreases with increasing ξ because the path entropy of the system decreases. The entropy rate per unit time increases until $\xi \approx 6$, while the mean transit time decreases faster than the path entropy, and then trend turns around and the entropy rate per unit time decreases (panel E). Orange lines in panel D and E show the respective entropy values for a one-pool system with the same mean transit time. The blue and orange lines intersect at $\xi \approx 4.31$.
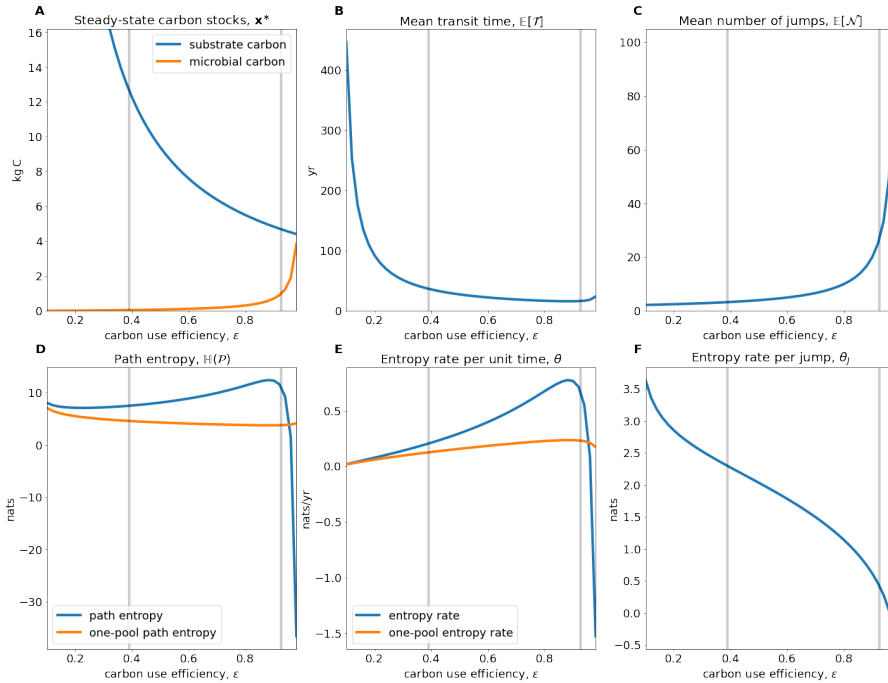

Fig. 6: A) Equilibrium carbon stocks. B)–F) Entropy related quantities of the global carbon cycle model introduced by Wang et al (2014) in dependence on the microbial carbon use efficiency ε (blue lines). Orange lines correspond to the quantities derived from a one-pool model with the same speed. The left vertical gray lines show $\varepsilon = 0.39$, the original carbon use efficiency of the model, the right $\varepsilon = 0.926$, the carbon use efficiency value with the maximum entropy rate of the Poisson process associated with C_s.

4 Discussion

- complexity
- gap to data, e.g. time series in Friedlingstein: slow system, high ET, high path entropy, ET, stock size, C uptake (net stock change)
- Can a CS represent model certain data, does it produce sufficient information?
- A one-pool model cannot... what?
- MaxEnt, MaxCal
- path entropy here and in MaxCal paper (Pressé et al, 2013)
- Markov model as Maxent (only jump chain) (Ge et al, 2012)

In panels D and E of Fig. 4 we see that before the break-even point of $\xi \approx 4.31$ the path of the Emanuel model is harder to predict than the path (i.e. the exit time of the particle) of a one-pool model with the same mean transit time. After this point of break even, the path of the Emanuel model with five compartments is easier to predict than only the transit time in a one-pool model. The reason is that as the system becomes faster, the differential entropy of the sojourn times in slow pools decreases so fast that at some point the sojourn times in slow pools visited by few particles becomes rather unimportant. In consequence this means that after the break-even point, i.e. for a sufficiently fast system, a one-pool model is too biased on the biased on slow-cycling paths, while fast paths are dominating the system. A more detailed model that separates fast from slow paths is then even easier to predict, even though the paths look more complicated.

5 Conclusions

- perspective complexity measure
- perspective: non-equilibrium

6 Acknowledgements

References

Albert A (1962) Estimating the Infinitesimal Generator of a Continuous Time, Finite State Markov Process. The Annals of Mathematical Statistics 33(2):727–753

Anderson DH (1983) Compartmental modeling and tracer kinetics, vol 50. Springer Science & Business Media

Bad Dumitrescu ME (1988) Some informational properties of markov pure-jump processes. Časopis pro pěstování matematiky 113(4):429–434

Bonchev D, Buck GA (2005) Quantitative Measures of Network Csomplexity. In: Complexity in Chemistry, Biology, and Ecology, Springer, pp 191–235

Cover TM, Thomas JA (2006) Elements of Information Theory, 2nd edn. Wiley

Dehmer M, Mowshowitz A (2011) A history of graph entropy measures. Information Sciences 181(1):57–78

Doob JL (1953) Stochastic Processes, vol 7. Wiley, New York

Emanuel WR, Killough GG, Olson JS (1981) Modelling the Circulation of Carbon in the World's Terrestrial Ecosystems. In: Carbon Cycle Modelling, SCOPE 16, John Wiley and Sons, pp 335–353

Gaspard P, Wang XJ (1993) Noise, chaos, and (ε, τ)-entropy per unit time. Physics Reports 235(6):291–343

Ge H, Pressé S, Ghosh K, Dill KA (2012) Markov processes follow from the principle of maximum caliber. The Journal of Chemical Physics 136(6):064,108

Girardin V, Limnios N (2003) On the Entropy for Semi-Markov Processes. Journal of Applied Probability 40(4):1060–1068

Jacquez JA, Simon CP (1993) Qualitative theory of compartmental systems. Siam Review 35(1):43–79

Metzler H, Sierra CA (2018) Linear autonomous compartmental models as continuous-time Markov chains: Transit-time and age distributions. Mathematical Geosciences 50(1):1–34, DOI 10.1007/s11004-017-9690-1

Neuts MF (1981) Matrix-geometric solutions in stochastic models: An algorithmic approach. The Johns Hopkins University Press

Norris JR (1997) Markov Chains. Cambridge University Press

Pesin YB (1977) Characteristic Lyapunov exponents and smooth ergodic theory. Uspekhi Matematicheskikh Nauk 32(4):55–112

Pressé S, Ghosh K, Lee J, Dill KA (2013) Principles of maximum entropy and maximum caliber in statistical physics. Reviews of Modern Physics 85(3):1115

Shannon CE, Weaver W (1949) The Mathematical Theory of Communication. The University of Illinois Press, Urbana

Sierra CA, Müller M, Metzler H, Manzoni S, Trumbore SE (2016) The muddle of ages, turnover, transit, and residence times in the carbon cycle. Global Change Biology in print, DOI 10.1111/gcb.13556

Thompson MV, Randerson JT (1999) Impulse response functions of terrestrial carbon cycle models: method and application. Global Change Biology 5(4):371–394, DOI 10.1046/j.1365-2486.1999.00235.x

Trucco E (1956) A note on the information content of graphs. Bulletin of Mathematical Biology 18(2):129–135

Walter GG, Contreras M (1999) Compartmental Modeling with Networks. Birkhäuser

Wang YP, Chen BC, Wieder WR, Leite M, Medlyn BE, Rasmussen M, Smith MJ, Agusto FB, Hoffman F, Luo YQ (2014) Oscillatory behavior of two nonlinear microbial models of soil carbon decomposition. Biogeosciences 11(7):1817–1831, DOI 10.5194/bg-11-1817-2014