ERAHUMED DSS

Valerio Gherardi Pablo Amador Crespo Andreu Rico

Table of contents

Preface

The purpose of this book is to provide a comprehensive reference for the ERAHUMED Decision Support System. Here you can find the technical descriptions of the algorithms employed by the system, as well as the user manual for the accompanying software.

The Support System and, hence, this book are currently under development on Github. In particular, the `{erahumed}` R package is hosted here.

For general information on the ERAHUMED project, please refer to the official website. If you want to get in touch, you can contact any of us via e-mail:

- Andreu Rico (Coordinator)
- Pablo Amador (PhD Researcher)
- Valerio Gherardi (Software Developer)

1 Introduction

This is a book created from markdown and executable code.

See Martínez-Megías et al. (2024) for additional info.

Part I

Technical description

2 The ERAHUMED model: a bird's eye view

The ERAHUMED model for assessing the ecological status of the Albufera Natural Park consists of three key components:

- **Hydrology**: Water dynamics within the park
- **Exposure**: Estimating the exposure to toxic chemicals
- **Risk Assessment**: Evaluating the impact of exposure

From a spatial perspective, the natural park is divided into three types of water bodies: the Albufera lake, rice field clusters[1], and irrigation ditches, which hydrologically connect the lake to the fields. Each of the model's computational layers incorporates specific quantitative models to simulate the relevant processes across all water bodies. This is summarized in Figure 2.1, where arrows indicate downstream dependencies and define the logical computation order.

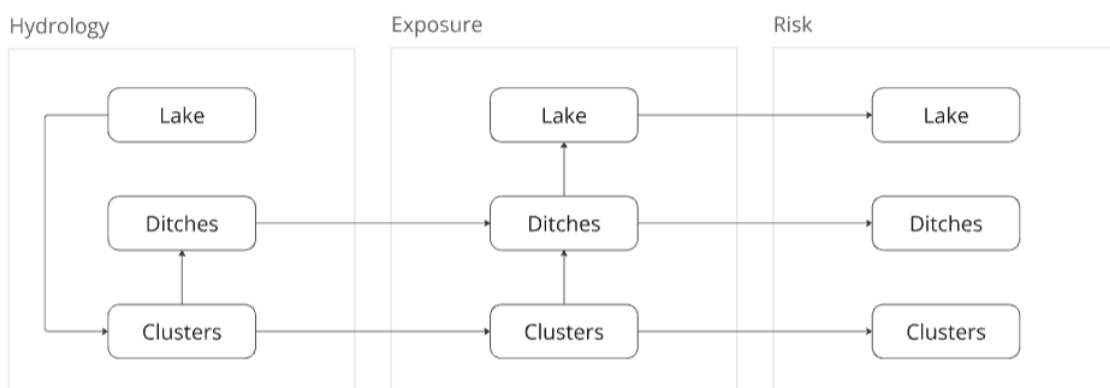

Figure 2.1: Scheme of ERAHUMED model components. Directional arrows indicate the downstream dependencies of the various simulation layers.

To clarify this structure, we can summarize the role of each simulation layer in Figure 2.1 as follows:

[1]The exact definition of "clusters" is discussed in Section 4.3. For the purposes of this high-level description, we can think of them simply as groups of rice fields.

1. The system's hydrology, including water volumes and flows for all hydrological elements, is derived from minimal input data: daily water levels and sea outlet outflows for the Albufera lake. This is achieved through a set of simplifying assumptions about the hydrology of rice fields and irrigation ditches. Details on this model are provided in Chapter 4.
2. Exposure to chemicals is calculated by first simulating their application to rice fields based on typical cultivation patterns. The dispersion of chemicals is then modeled using a simplified set of differential equations designed to capture the key physical processes driving their spread. These calculations are described in detail Chapter 5.
3. The impact of chemicals is evaluated across all water bodies using a simplified approach based on Species Sensitivity Distributions, utilizing publicly available toxicity data for their estimation. This is described in detail in Chapter 6.

3 Model inputs

This chapter serves as a central reference for all input parameters used in ERAHUMED simulations.

3.1 Landscape parameters

Landscape parameters are collected in Table 3.1. In this table, a `numeric(n)` type indicates a numeric vector of `n` components, while `data.frame` inputs have more complex formats, detailed below.

3.2 Chemical-specific parameters

In addition to landscape parameters, ERAHUMED includes a set of parameters for each supported chemical, defining their physico-chemical and toxicological properties. At this stage, these parameters are internal and not reported here. However, our roadmap includes future support for customizing these parameters and defining new chemicals.

3.3 Data frame inputs

We detail in the following sections the format of data frame inputs.

3.3.1 Lake outflows and levels data frame

Time-series dataset that provides the observational hydrological data on the Albufera lake, along the template of `albufera_outflows` (the default value).

Table 3.2: Lake outflows and levels data frame [one row per day in the desired study frame.]

Column	Description
date	Date of measurement
level	Lake level (in meters above sea level)

Column	Description
outflow_pujol	Outflow at Pujol (meters cube per second)
outflow_perellonet	Outflow at Perellonet (meters cube per second)
outflow_perello	Outflow at Perello (meters cube per second)
is_imputed_level	Whether the `level` value was imputed.
is_imputed_outflow	Whether (any of) the outflows were imputed.

3.3.2 Weather data frame

A dataset that provides the relevant metereological time series, along the template of `albufera_weather` (the default value).

Table 3.3: Weather data frame [one row per day in the desired study frame.]

Column	Description
date	Date of measurement
temperature_ave	Average temperature.
temperature_min	Minimum temperature.
temperature_max	Maximum temperature.
precipitation_mm	Daily precipitation in millimiters.
evapotranspiration_mm	Daily evapotranspiration in millimiters.

3.3.3 Rice paddy management data frame

Dataset that provides the yearly schedule for irrigation and draining, along the template of `albufera_management` (the default value).

Table 3.4: Rice paddy management data frame [one row per day of year (29th of Feb. included) and per combination of the categorical variables `tancat` and `variety`.]

Column	Description
mm	numeric. Month of year (1 = January, 2 = February, *etc.*).
dd	numeric. Day of month.
tancat	logical. Whether the paddy is a tancat or not.
variety	character. Rice variety of the paddy under consideration.
sowing	logical. Whether `mm` and `dd` correspond to the sowing day.
ideal_irrigation	logical. Whether the paddy is scheduled to be irrigated on this day.
ideal_draining	logical. Whether the paddy is scheduled to be drained on this day.

Column	Description
ideal_height_eod_cm	numeric. Scheduled water level of the paddy *at the end of the day* (that is, after irrigation and draining).

3.3.4 Chemical application schedules data frame

A dataset that provides the list of scheduled chemical applications, along the template of `albufera_ca_schedules`.

Table 3.5: Chemical application schedules data frame [one row per scheduled application.]

Column	Description
day	numeric. Scheduled day, counted starting from the sowing day, for the application under consideration.
rice_variety	character. Rice variety for this specific application.
chemical	character. Name of applied chemical.
kg_per_ha	numeric. Amount of chemical applied, in kilograms per hectare.
application_type	either `"ground"` or `"aerial"`. Application mode of the chemical to rice paddies.

Table 3.1: ERAHUMED input parameters

Parameter	Name	Unit	Group
\texttt{outflows_df}	Lake outflows and levels data frame	N/A	Hydrolog
\texttt{weather_df}	Weather data frame	N/A	Meteoro
\texttt{variety_prop}	Rice variety proportion	N/A	Environ
\texttt{seed}	\texttt{seed}	N/A	Hyperpa
\texttt{storage_curve_slope_m2}	Storage curve slope	m^2	Hydrolog
\texttt{storage_curve_intercept_m3}	Storage curve intercept	m^3	Hydrolog
\texttt{petp_surface_m2}	PET surface	m^2	Hydrolog
\texttt{management_df}	Rice paddy management data frame	N/A	Environ
\texttt{ideal_flow_rate_cm}	Ideal flow rate	cm	Hydrolog
\texttt{height_thresh_cm}	Cluster Height Threshold	cm	Hydrolog
\texttt{ditch_level_m}	Ditch water level	m	Hydrolog
\texttt{ca_schedules_df}	Chemical application schedules data frame	N/A	Environ
\texttt{drift}	Drift	1	Environ
\texttt{covmax}	Max interception potential	1	Environ
\texttt{jgrow}	Maturation cycle length	day	Environ
\texttt{SNK}	SNK	1	Environ
\texttt{dact_m}	Depth of active sediment	m	Environ
\texttt{css_ppm}	Suspended sediment concentration	ppm	Environ
\texttt{foc}	Fraction of organic content	1	Environ
\texttt{bd_g_cm3}	Bulk density of sediment	$g \cdot cm^{2}$	Environ
\texttt{qseep_m_day}	Seepage rate	$m \cdot day^{1}$	Environ
\texttt{wilting}	Wilting point	1	Environ
\texttt{fc}	Field capacity	1	Environ

4 Hydrological model of the Albufera Natural Park

The first step in assessing toxicological risks in the Albufera Natural Park is to determine the system's hydrology, as water serves as the primary transport medium for the chemicals under study. Specifically, by "hydrology," we refer to the water volumes present in each water body at a given moment and the flow of water between them over a defined time frame (*e.g.*, one day). Since most of these quantities are not directly measurable, they must be estimated through simulation. This Chapter outlines the algorithms used for this process.

4.1 Definition of hydrological elements

Our hydrological model represents the Albufera Natural Park in terms of three main landscape elements (or "water bodies"): rice field clusters, irrigation ditches, and Albufera Lake.

The definition of rice field clusters and irrigation ditches used in our modeling is discussed in detail in Ref. [TODO: insert Pablo's paper reference]. For our purposes, we note that:

- The park's cultivation area is divided into rice field clusters, each comprising several rice fields that share the same hydrological management system (*tancat* or regular) and rice variety (*J.Sendra*, *Bomba* or *Clearfield*).
- Each cluster is assumed to drain into a single irrigation ditch, selected based on the shortest distance (see Ref. TODO for details).

4.2 Random assignation of rice variety

Since the actual rice variety cultivated in individual fields is unknown, a random variety is assigned to each cluster based on the following criteria:

- The proportion of cultivated surface allocated to each variety is determined by the `variety_prop` parameter (cf. Table 3.1).
- The *Bomba* variety is cultivated exclusively in *tancats*.
- The *Clearfield* variety is restricted to the northern part of the natural park, in clusters draining into ditches 1 to 19.

4.3 Scheme of the hydrological model

This schematic diagram represents the simplified hydrological model of the Albufera Natural Park employed by ERAHUMED. It highlights water flows across the three primary landscape elements defined in the previous Section. Key simplifying assumptions embedded in the model and visually summarized in the diagram are as follows:

- **Rice Clusters** Clusters are irrigated by external water sources and drain exclusively a single ditch. There is no direct hydrological interaction or exchange between individual clusters.
- **Ditches** Ditches collect water from the clusters and, potentially, from additional external sources, channeling all inflows directly into the Albufera lake.
- **The Albufera Lake** The lake receives water exclusively from the ditches. While two types of outflow are considered, namely direct discharge to the sea and water recirculation to the rice fields, the latter is typically negligible[1].

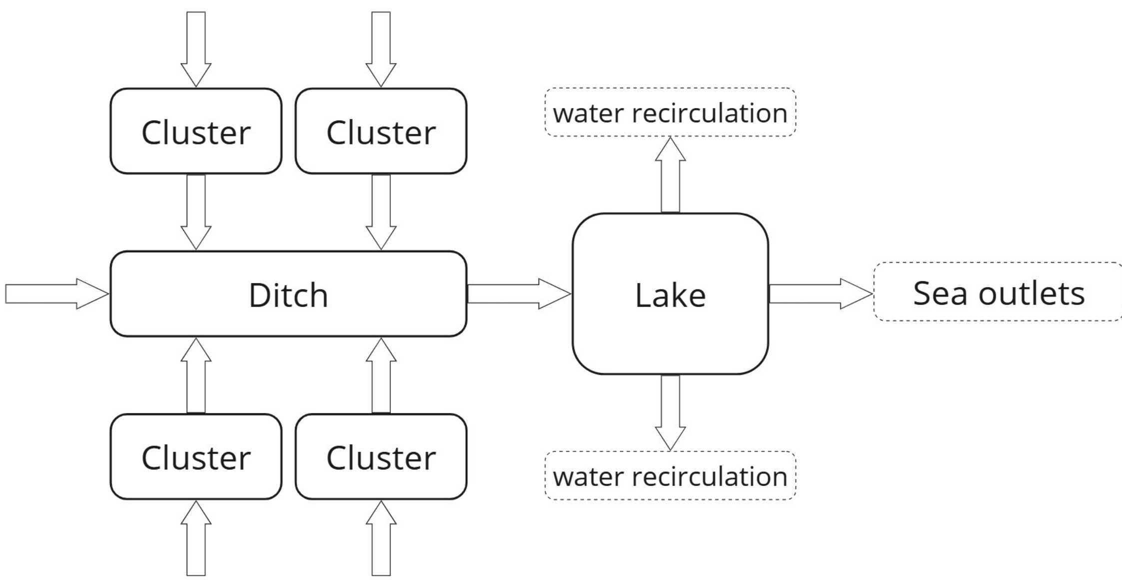

Figure 4.1: Scheme of ERAHUMED hydrological model of the Albufera Natural Park

[1]Further details on this are discussed in Section 4.4.1.

4.4 Water balance calculations

This section provides the details of the water balance calculations for (in the order of computation) the Albufera Lake, rice field clusters, and irrigation ditches of the park.

4.4.1 Albufera Lake

Water balance calculations for the Albufera Lake are relatively simple. The relevant equation expressing hydrological balance is:

$$\text{Volume Change} = \text{Inflow} - \text{Outflow} + \text{Precipitation} - \text{Evapotranspiration}. \tag{4.1}$$

The variables collected into the following table, that enter the balance equation 4.1, have direct correspondence with model input parameters listed in Chapter 3 (we use the notation `df$col` to indicate column `col` of data frame `df`).

Variable	Source	Units	Description
h_t	`outflows_df$level`	m	Lake water level daily time series
O_t^{Pujol}	`outflows_df$outflow_pujol`	m^3	*Pujol* outflow daily time series
$O_t^{\text{Perelló}}$	`outflows_df$outflow_perello`	m^3	*Perelló* outflow daily time series
$O_t^{\text{Perellonet}}$	`outflows_df$outflow_perellonet`	m^3	*Perellonet* outflow daily time series
P_t	`outflows_df$precipitation_mm`	mm	Precipitation (per unit area) daily time series
ET_t	`outflows_df$evapotranspiration_mm`	mm	Evapotranspiration (per unit area) daily time series
α	`storage_curve_intercept_m3`	m^3	Storage curve intercept
β	`storage_curve_slope_m2`	m^2	Storage curve slope
σ_{PET}	`petp_surface_m2`	m^2	PET surface

Calculated quantities are listed in the following table.

Variable	Units	Description
V_t	m^3	Lake water volume daily time series

Variable	Units	Description
$\Delta V_t \equiv V_{t+1} - V_t$	m^3	Lake water volume change daily time series
$\Delta V_t^{\mathrm{PET}}$	m^3	Lake water volume change due to precipitation and evapotranspiration daily time series
I_t	m^3	Lake total inflow daily time series
O_t	m^3	Lake total outflow daily time series

The volume time-series is computed as:

$$V_t = \alpha + \beta \cdot h_t, \tag{4.2}$$

while the volume changes due to precipitation and evapotranspiration are given by:

$$\Delta V_t^{\mathrm{PET}} = \sigma_{\mathrm{PET}}(\mathrm{P}_t - \mathrm{ET}_t), \tag{4.3}$$

Total inflow and outflow must satisfy Equation 4.1, which we may rewrite explicitly as:

$$\Delta V_t - \Delta V_t^{\mathrm{PET}} = I_t - O_t, \tag{4.4}$$

Strictly speaking, O_t is not merely the sum of O_t^{Pujol}, $O_t^{\mathrm{Perelló}}$ and $O_t^{\mathrm{Perellonet}}$, but is rather calculated as follows:

$$O_t = \max \left[O_t^{\mathrm{Pujol}} + O_t^{\mathrm{Perelló}} + O_t^{\mathrm{Perellonet}}, \Delta V_t^{\mathrm{PET}} - \Delta V_t \right]. \tag{4.5}$$

The rationale is that the simple sum of estuaries outflows omits potentially important contributions from *water recirculation*, that is to say, water being pumped out from the lake for rice-field irrigation, by the so-called *tancats*. Such amount of recirculated water is hard to estimate and, in the lack of a better model, we simply assume this to be negligible, *except* when a positive amount is required by Equation 4.4 itself, due the physical constraint that $I_t \geq 0$.

Once O_t is calculated through Equation 4.5, I_t can be immediately obtained from Equation 4.4. Notice that whenever the aforementioned compensating outflow term due to water recirculation is included (which happens when the maximum in Eq. 4.5 is given by the second term), the total inflow is always estimated to be zero.

4.4.2 Rice field clusters

Water balance calculations for rice field clusters are complex due to the lack of observational data. The simulation algorithm relies on several key components:

- The hydrology of the Albufera lake (Section 4.4.1).
- Assumptions about the hydrological connections between the various water bodies in the natural park, detailed in Section 4.3.
- An ideal yearly management plan for irrigation and drainage of the rice paddies.

The quantitative inputs for this calculation are collected in the following table; see the definition of the `management_df` input data frame, discussed in Section 3.3.3. Below, v denotes rice variety, and θ denotes the hydrological management system (*tancat* or *regular*).

Variable	Source	Units	Description
$\mathrm{Ir}_{d,v,\theta}$	`management_df$ideal_irrigation`	NA	Boolean expressing whether a cluster of variety v and management system θ is supposed to be irrigated on day of year d
$\mathrm{Dr}_{d,v,\theta}$	`management_df$ideal_draining`	NA	Boolean expressing whether a cluster of variety v and management system θ is supposed to be drained on day of year d
$\mathcal{H}_{d,v,\theta}$	`management_df$ideal_height_eod_cm`	cm	Ideal water depth (in cm) of a cluster of variety v and management system θ on day of year d
k_{flow}	`ideal_flow_rate_cm`	$\mathrm{cm} \cdot \mathrm{day}^{-1}$	Rate of water flow through rice paddies during simultaneous irrigation and drainage, while maintaining a constant overall water level.

Variable	Source	Units	Description
h_{thres}	`height_thresh_cm`	cm	Water height below which a cluster is considered emptied. Used to determine delays in the draining/irrigation plan. Expressed in cm.
A_c	Internal parameter	m^2	Surface area of cluster c

The outputs are collected below, where the indices c and t denotes the cluster and time, respectively:

Variable	Units	Description
$V_{c,t}$	m^3	Cluster's water volume
$I_{c,t}$	m^3	Cluster's inflow
$O_{c,t}$	m^3	Cluster's outflow

The fundamental equation for hydrological balance is:

$$V_{c,t+1} - V_{c,t} = I_{c,t} - O_{c,t} + (\mathrm{P}_t - \mathrm{ETP}_t) \times A_c \tag{4.6}$$

where P_t and ETP_t are the same as in Section 4.4.1.

In what follows, we will focus on the set of all clusters draining into a given ditch, and we will enumerate clusters through the index $c = 1, 2, \ldots, N_C$. On the other hand, clusters are assumed to be irrigated from sources external the Albufera system, *i.e.*, not through any of the ditches that eventually flow into the lake. According to our assumptions on the hydrology of clusters and ditches (*cf.* Section 4.3), the sum of cluster outflows is constrained by:

$$\sum_{c=1}^{N_C} O_{c,t} \leq Q_t \tag{4.7}$$

where Q_t denotes the inflow to the relevant ditch. With a small leap in logic, we anticipate from Section 4.4.3, that the water levels in irrigation ditches are assumed to be constant, so that Q_t can also be identified with the ditch outflow, that is in turn estimated as:

$$Q_t = \frac{\text{Area of clusters draining into ditch}}{\text{Area of all clusters}} \times \text{Lake's Inflow}_t \tag{4.8}$$

where the lake's inflow is computed as described in Section 4.4.1.

Qualitatively speaking, the simulation determines daily values of $I_{c,t}$ and $O_{c,t}$ that satisfy Eqs. 4.6 and 4.7, and such that the resulting hydrology aligns as closely as possible with the "ideal" conditions prescribed by the specified management plan. This is accomplished in three steps:

1. Computing ideal inflows and outflows based on current water levels and management plans.
2. Determining actual inflows and outflows, along with the corresponding actual water level changes.
3. Adjusting management plan delays if some clusters were scheduled to be drained but could not be due to insufficient flow through the common ditch (see below for details).

These steps are iterated on a daily basis, starting from some initial time (say $t = 0$) at which all cluster water levels match the ideal ones, and no plan delays are present.

Concerning the last step, a few words may serve to clarify the algorithm described below. The purpose of management plan delays is to ensure that during each year's sowing season, all cluster's are eventually emptied as required for ground applications of chemicals (modeled in subsequent layers of the simulation). This is achieved by postponing the management plan by one day whenever an emptying condition is not met. Outside of the sowing window, delays are reset to zero to prevent them from accumulating indefinitely, which would be unrealistic.

In what follows, the conversion between cluster water volumes and depths is provided by $V_{c,t} = A_c \cdot h_{c,t}$, and we denote by $\delta_{c,t}$ the time series of plan delays for cluster c, which is initialized by $\delta_{c,0} = 0$.

4.4.2.1 Step 1: ideal balance

Ideal balance quantities for each cluster c are obtained from the management plan data-set, whose relevant row is identified by the cluster's rice variety v and field type θ, and the delayed day:

$$d_{c,t+1} = d_{t+1} - \delta_{c,t}, \tag{4.9}$$

where d_{t+1} denotes the day of year corresponding to time $t+1$, and $\delta_{c,t}$ the accumulated plan delay.

Let $h^{\text{id}}_{c,t+1}$ denote the ideal depth for cluster c at time $t+1$ retrieved in this way, and $\text{Ir}_{c,t}$, $\text{Dr}_{c,t}$ the corresponding irrigation and draining plans. Furthermore, denote by $V^{\text{id}}_{c,t+1} = A_c \cdot h^{\text{id}}_{c,t+1}$ the corresponding ideal water volume.

In order to compute ideal inflow and outflow, we require (*cf.* Equation 4.6):

$$V^{\mathrm{id}}_{c,t+1} = \max\{V^{\mathrm{id}}_{c,t} + (\mathrm{P}_t - \mathrm{ETP}_t) \times A_c, 0\} + I^{\mathrm{id}}_{c,t} - O^{\mathrm{id}}_{c,t} \tag{4.10}$$

Clearly, Equation 4.10 alone does not individually specify $I^{\mathrm{id}}_{c,t}$ and $O^{\mathrm{id}}_{c,t}$, but only their difference $\Delta^{\mathrm{id}}_{c,t} = I^{\mathrm{id}}_{c,t} - O^{\mathrm{id}}_{c,t}$. In order to fix both these quantities:

$$\begin{aligned}(I^{\mathrm{id}}_{c,t})^{(0)} &= \begin{cases} k_{\mathrm{flow}} & \text{if } \mathrm{Ir}_{c,t} = \mathrm{Dr}_{c,t} = 1 \\ 0 & \text{otherwise} \end{cases}, \\ (O^{\mathrm{id}}_{c,t})^{(0)} &= (I^{\mathrm{id}}_{c,t})^{(0)} - \Delta^{\mathrm{id}}_{c,t}.\end{aligned} \tag{4.11}$$

and, in order to ensure that flows are positive, we finally set:

$$\begin{aligned} O^{\mathrm{id}}_{c,t} &= \max\{(O^{\mathrm{id}}_{c,t})^{(0)}, 0\} \\ I^{\mathrm{id}}_{c,t} &= O^{\mathrm{id}}_{c,t} + \Delta^{\mathrm{id}}_{c,t}, \end{aligned} \tag{4.12}$$

which satisfy Equation 4.10 and give rise to positive $O^{\mathrm{id}}_{c,t}$ and $I^{\mathrm{id}}_{c,t}$.

4.4.2.2 Step 2: real balance

At each time-step t, the cluster's index set is randomly permuted [2], and the real flows are calculated as:

$$\begin{aligned} O_{c,t} &= \min\{O^{\mathrm{id}}_{c,t}, Q_t - \sum_{c'<c} O_{c',t}\}, \\ I_{c,t} &= \max\{I^{\mathrm{id}}_{c,t} - O^{\mathrm{id}}_{c,t} + O_{c,t}, 0\} \end{aligned} \tag{4.13}$$

In simple terms, clusters are emptied in a random order within the allowed capacity of the corresponding ditch. Using Equation 4.13, we finally determine the real water level achieved as:

$$V_{c,t+1} = \max\{V_{c,t} + (\mathrm{P}_t - \mathrm{ETP}_t) \times A_c, 0\} + I_{c,t} - O_{c,t} \tag{4.14}$$

to be compared with Equation 4.10.

[2] With some abuse of notation, we assume the indexes c and c' in Equation 4.13 to be sorted according to this random permutation.

4.4.2.3 Step 3: updating the plan delay

The updated value $\delta_{c,t+1}$ is obtained as follows. If d_{t+1} (the *actual* day of year) is outside of the window $W =$ [20th of April, 15th of October], then $\delta_{c,t+1} = 0$. Otherwise, if $h^{\text{id}}_{c,t} > 0$ or $h_{c,t} < h_{\text{thres}}$, the plan delay is unchanged: $\delta_{c,t+1} = \delta_{c,t}$. Finally, if $h^{\text{id}}_{c,t} = 0$ but $h_{c,t} > h_{\text{thres}}$, we add one day of delay: $\delta_{c,t+1} = \delta_{c,t} + 1$.

4.4.2.4 Step 3: updating the plan delay

The updated value $\delta_{c,t+1}$ is obtained as follows. If d_{t+1} (the *actual* day of year) is outside of the window $W =$ [20th of April, 15th of October], then $\delta_{c,t+1} = 0$. Otherwise, if $h^{\text{id}}_{c,t} > 0$ or $h_{c,t} < h_{\text{thres}}$, the plan delay is unchanged: $\delta_{c,t+1} = \delta_{c,t}$. Finally, if $h^{\text{id}}_{c,t} = 0$ but $h_{c,t} > h_{\text{thres}}$, we add one day of delay: $\delta_{c,t+1} = \delta_{c,t} + 1$.

4.4.3 Irrigation ditches

Our approach to the hydrology of irrigation ditches is simplified, with the main assumption being that all ditches have a common, constant water depth h_{ditch} corresponding to the input parameter `ditch_level_m`.

Using the index $D = 1, 2, \ldots, 26$ to enumerate the park's main ditches, ditch outflows to the Albufera lake $O_{D,t}$ are calculated according to Equation 4.8. These outflows also coincide with the total ditch inflows $I_{D,t}$ (due to the constant water volume assumption).

5 Exposure

5.1 Overview

5.2 Pesticide applications

The simulation of pesticide applications to rice fields follows a straightforward algorithm that takes as input the scheduled application list and the system's hydrology (Chapter 4). The algorithm outputs a list of actual applications, specifying the application day and the applied mass. Since the application of each pesticide is computed independently of others, we focus on a single pesticide for the present discussion.

The information on scheduled applications is contained in the `ca_schedules_df` data frame, whose structure is described in detail in Section 3.3.4. For each scheduled application, this data set specifies:

- An application day D, representing the day of the crop cycle when the pesticide is expected to be applied.
- An application amount, expressed in units of mass per unit area.
- An application type, which is either *aerial* or *ground*, that determines the required rice field conditions for the application.

To determine the actual application date, we proceed as follows (for notation related to cluster hydrology, see Section 4.4.2):

1. **Filtering by irrigation and draining states.** For ground applications, we select only those days in which both $\mathrm{Ir}_{c,t} = \mathrm{Dr}_{c,t} = 0$, where $\mathrm{Ir}_{c,t}$ and $\mathrm{Dr}_{c,t}$ are the true (as opposed to ideal) irrigation and draining states time-series, defined after Equation 4.9. Similarly, for aerial applications, we select days according to $\mathrm{Ir}_{c,t} = \mathrm{Dr}_{c,t} = 1$.

2. **Filtering by water depth (ground applications only).** For ground applications, from the days selected in the previous step, we further filter those where the simulated water depth of the cluster satisfies $h_{c,t} < h_{\mathrm{thresh}}$.

3. **Ensuring separation between applications of the same pesticide.** We retain only those days that are at least five days apart from any previous application of the same chemical.

4. **Selecting the final application day.** From the remaining candidate days, we choose the one where the *delayed day* $d_{c,t}$ defined by Equation 4.9 is closest to D, the originally scheduled application day.

These steps are repeated for all subsequent applications of a given pesticide in a given cluster. It is worth noting that the algorithm for simulating the cluster hydrological balance (Section 4.4.2) guarantees that at least one candidate application day will always remain available in the final selection step.

5.3 Pesticide dispersion

As illustrated in our hydrological model (Figure 4.1), pesticides applied to rice field clusters are transported through water flow into ditches, then into the Albufera lake, and ultimately to the sea.

From a mathematical perspective, the fate of these chemicals is governed by a system of differential equations that captures key physico-chemical processes, including transport, diffusion between compartments, and degradation.

The remainder of this chapter describes our approach to pesticide dispersion. We define the simplified differential system used to model these processes in Section 5.3.2, and we present our semi-analytic solution method in Section 5.3.3.

5.3.1 Diagram of physical processes

The schematic diagram below illustrates the processes captured by our model for chemical dispersion. Directional arrows represent the transfer of chemical matter among the three compartments considered—Foliage, Water, and Soil—as well as exchanges with external water bodies (denoted as "Watercourse").

While the overall scheme applies to all landscape elements, specific details vary for clusters, ditches, and the lake (*cf.* the hydrological model of Figure 4.1):

- **Clusters:** Inflow waters originate outside the system and are assumed to be free of pesticides.
- **Ditches and the Albufera lake:** No direct pesticide applications occur in these areas, meaning the foliage compartment plays no role in their dynamics.
- **Ditches:** Receive inflows from two sources—clusters, which contribute pesticide-laden water, and external sources, which are assumed to be free of chemicals.
- **Albufera lake:** Its only inflows come from ditches, which contain nonzero chemical concentrations.

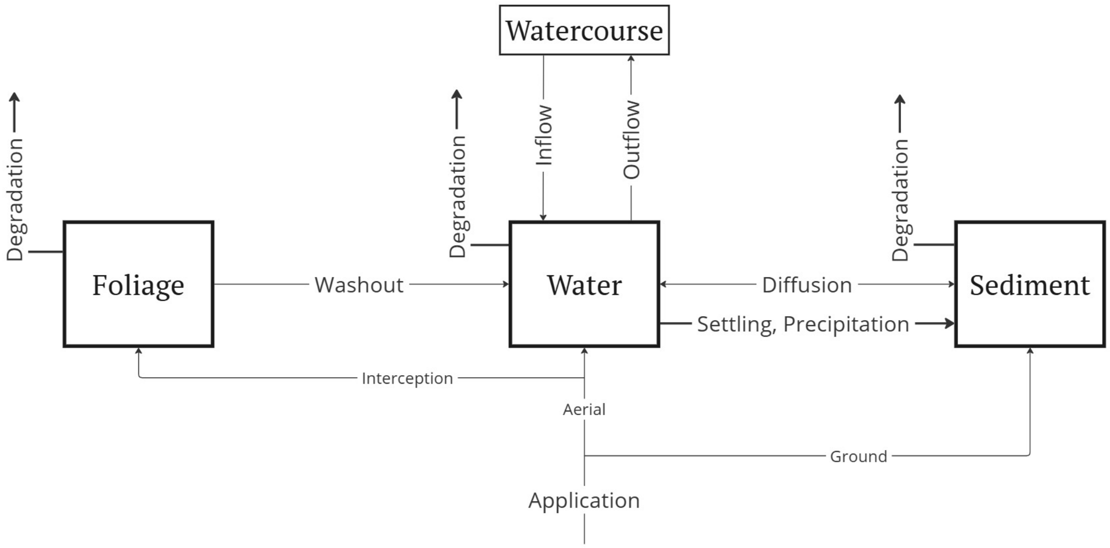

Figure 5.1: Scheme of physical processes considered in ERAHUMED model of chemical dispersion

5.3.2 Evolution Equations

As in Section 5.2, we focus on a single pesticide and a fixed hydrological element (*e.g.*, a cluster). We denote its masses in the three compartments—foliage, water, and sediment—by m_f, m_w, and m_s, respectively. We assume that, at any given time, these masses are fully diluted and homogeneously distributed. In particular, we can meaningfully define spatially constant concentrations c_f and c_s in the water[1] and sediment compartments.

We first describe the differential system that drives the temporal dynamics of m_f, m_w and m_s. The correspondence between the parameters used in this and the following Section is provided in Section 5.3.4.

The fundamental equations describing pesticide diffusion are given as follows:

$$\begin{aligned}
\frac{\mathrm{d}m_f}{\mathrm{d}t} &= -(k_f + w)m_f + a_f \\
\frac{\mathrm{d}m_w}{\mathrm{d}t} &= -(k_w + d_w + s + \frac{O}{V})m_w + d_s m_s + w m_f + a_w - \sigma(\frac{m_w}{\rho V}) \\
\frac{\mathrm{d}m_s}{\mathrm{d}t} &= (d_w + s)m_w - (k_s + d_s)m_s + a_s + \sigma(\frac{m_w}{\rho V})
\end{aligned} \tag{5.1}$$

where:

[1] Provided the water volume is nonzero.

- $a_{f,w,s}$ are the mass income rates of the chemical in the three compartments, that include direct pesticide application and mass coming from inflows.
- $k_{f,w,s}$ are the degradation rates of the chemical in the three compartments,
- d_w and d_s are the water-sediment diffusion rates,
- s is the settling rate,
- w is the washout rate,
- O is the outflow rate of the rice field,
- V is the volume of water in the rice field,
- ρ is the chemical solubility in water,
- $\sigma(x)$ is a function (not further specified, see below) that grows quickly for $x > 1$, and vanishes for $x \leq 1$. This function accounts for solubility.

Strictly speaking, all these terms have instantaneous time dependence[2]. Apart from making the system 5.1 hard to attack by analytic means, such a dependence is troubling because we don't have access to the exact time dependence of the majority of these terms (*e.g.* outflow, volume or chemical applications), our input consisting of simple daily average/cumulative values. On the other hand, if all the terms were constant, and if we could neglect σ, the solution of Equation 5.1 would be immediate, as the corresponding system becomes a linear ODE with constant coefficients, in addition whose eigenvalues and eigenvectors can be computed explicitly.

What we use in practice is an intermediate semi-analytic approach that allows us to compute daily values of $m_{f,w,s}$, which is described next.

5.3.3 Semi-analytic solution of Equation 5.1

The time evolution of $m_{f,w,s}$ in a daily time-step (from t to $t+1$, say, assuming t is measured in days) is obtained through the following four consecutive stages:

1. We compute the exact evolution of $m_{f,w,s}$ according to the linear ODE obtained by disregarding the processes of outflow, chemical application and solubility, and using daily constant values for all the remaining constants involved (described in more detail below).
2. We compute the m_w losses due to outflow as if they happened instantaneously after the processes computed in the previous step took place, with the water volume of the rice paddy varying from $V(t+1) + O(t)$ to $V(t+1)$.
3. We compute the mass applications again as instantaneous, after the losses in the water compartment due to outflows took place.
4. We compare the resulting m_w with the maximum amount allowed by the solubility in water, that is $m_w^{\max}(t) = \rho V(t+1)$ and transfer any excess to the sediment compartment (m_s).

[2]This observation also includes pure physico-chemical "constants" such as degradation rates, where the time dependence would stem from *temperature* dependence.

In the following, we describe in full detail the computations involved in these four steps.

5.3.3.1 Step 1: linear ODE evolution (physico-chemical processes)

Disregarding outflow, mass application and solubility, we get a linear ODE of the form:

$$\begin{aligned}\frac{\mathrm{d}m_f}{\mathrm{d}t} &= \gamma m_f,\\ \frac{\mathrm{d}m_w}{\mathrm{d}t} &= a_{ww}m_w + a_{ws}m_s + wm_f,\\ \frac{\mathrm{d}m_s}{\mathrm{d}t} &= a_{sw}m_w + a_{ss}m_s,\end{aligned} \tag{5.2}$$

with:

$$\begin{aligned}\gamma &= -(k_f + w),\\ a_{ww} &= -(k_w + d_w + s)\\ a_{ws} &= d_s\\ a_{sw} &= d_w + s\\ a_{ss} &= -(k_s + d_s)\end{aligned} \tag{5.3}$$

where the various physico-chemical parameters are assumed to be constant during a daily time-step (*cf.* Section 5.3.4).

The evolution of m_f is easily obtained:

$$m_f(t) = e^{\gamma t} m_f(0). \tag{5.4}$$

Plugging this into the three-dimensional system 5.2 we obtain a reduced two-dimensional system for the water-sediment compartments:

$$\begin{aligned}\frac{\mathrm{d}m_w}{\mathrm{d}t} &= a_{ww}m_w + a_{ws}m_s + we^{\gamma t}m_f(0),\\ \frac{\mathrm{d}m_s}{\mathrm{d}t} &= a_{sw}m_w + a_{ss}m_s,\end{aligned} \tag{5.5}$$

which is solved explicitly with standard methods[3].

[3]Letting $x = (m_w,\, m_s)^T$ and $b = w(m_f(0), 0)^T$, the system can be rewritten in the form:

$$\dot{x} = Ax + e^{\gamma t} b$$

5.3.3.2 Step 2: mass losses due to outflow

Mass losses due to outflow are computed as if the outflow process happened instantaneously after the processes discussed in the previous step took place. The amount of mass in the water compartment is affected as follows:

$$m_w \to \frac{V_f}{V_i} m_w, \tag{5.6}$$

where V_i and V_f denote the initial and final volume for the outflow process. Here V_f coincides with the simulated value $V(t+1)$ for the relevant water body and $V_i = V_f + O(t)$, where $O(t)$ denotes the total simulated outflow for day t.

5.3.3.3 Step 3: mass income from direct application and inflows

These contributions are computed as instantaneous and subsequent to outflow. Masses in the various compartments ($i = f,\ w,\ s$) are modified as follows:

$$m_i \to m_i + a_i \tag{5.7}$$

As mentioned above, a_i accounts for both direct pesticide applications and water inflow, that is:

$$a_i = a_i^{\text{app}} + a_i^{\text{inflow}}$$

The calculation of direct application rates from input parameters is described below. As to the second term, the amount of incoming mass is exactly given by the amount of outgoing mass in the outflow of the preceding element along the water course chain, which is computed as detailed in the previous Section.

---

where

$$A = \begin{pmatrix} a_{ww} & a_{ws} \\ a_{sw} & a_{ss} \end{pmatrix}$$

The general solution reads:

$$x(t) = e^{At}x(0) + (e^{At} - e^{\gamma t} I)(A - \gamma I)^{-1} b$$

where I is the identity matrix. Notice that the exponential can be computed explicitly from the eigenvalues of A:

$$e^{At} = e^{\lambda_+ t} P_+ + e^{\lambda_- t} P_-,$$

with:

$$\lambda_\pm = \frac{\mathrm{Tr} A}{2} \sqrt{\left(\frac{\mathrm{Tr} A}{2}\right)^2 - \det A}$$

and

$$P_\pm = \pm \frac{1}{\lambda_+ - \lambda_-}(A - \lambda_\mp)$$

5.3.3.4 Step 4: solubility

The chemical density in the water compartment after Steps 1, 2, and 3 is eventually compared with the chemical's solubility ρ, and any mass excess is instantaneously transferred to the sediment compartment. This implies the following modifications:

$$m_s \to m_s + \max(0,\, m_w - \rho V(t+1)), \quad m_w \to \max(m_w,\, \rho V(t+1)). \tag{5.8}$$

5.3.4 Input parameters (correspondence with Chapter 3)

This Section describes the sources of the various numerical inputs to Equation 5.1.

5.3.4.1 Pesticide application rates

Application rates are given by:

$$\begin{aligned} a_f^{\text{app}} &= A(t)\cdot(1-\text{drift})\cdot c(t), \\ a_s^{\text{app}} &= A(t)\cdot(1-\text{drift})(1-c(t))I(h(t+1)\neq 0), \\ a_w^{\text{app}} &= A(t)\cdot(1-\text{drift})\cdot(1-\text{SNK})(1-c(t))I(h(t+1)=0), \end{aligned}$$

Here, $A(t)$ is the applied amount on day t, whose computation was described in this chapter; $I(h(t+1)=0)$ is equal to one if the water level of the cluster $h(t+1)=0$, and vanishes otherwise[4]; the quantities drift, SNK correspond to the internal `drift` and `SNK` parameters (*cf.* Chapter 3); finally c is the coverage fraction, computed as:

$$c = \min\left[\frac{d_s(t)}{j_{\text{grow}}},\, 1\right]\cdot c_{\max},$$

where $d_s(t)$ is the number of days elapsed from seeding, assumed to be on the 20th of April, while j_{grow} and $c_{\max}$ correspond to the input parameters `jgrow` and `covmax`.

[4] The reason why the height of day $t+1$ is used here is because, as mentioned above, applications are computed using the final water volume of the cluster, as if all daily water flows already took place.

5.3.4.2 Degradation rates

Daily degradation rates $k_{f,w,s}(t)$ for the chemical are modeled through the Arrhenius equation, schematically:

$$k(t) = k_0 \cdot Q_{10}^{\frac{T(t)-T_0}{10}}$$

where k_0 is the degradation rate k_0 at a reference temperature T_0, Q_{10} is a numerical constant, and $T(t)$ is the average temperature on day t. The constants k_0 and Q_10 are internal model parameters, while $T(t)$ is given by `weather_df$temperature_ave`, see Section 3.3.1.

5.3.4.3 Diffusion rates

Diffusion rates are given by:

$$d_w(t) = \frac{k_{\text{dif}} \cdot f_{D,w}}{h(t)}, \quad d_s = \frac{k_{\text{dif}} \cdot f_{D,s}}{h_{\text{act}} \cdot \text{pos}}$$

where $h(t)$ is the water depth of the hydrological element; h_{act} is the depth of active sediment, corresponding to the input parameter `dact_m`; *textpos* is the porosity, and is expressed in terms of input parameters as `fc - wilting`; k_{dif} is given by the empirical formula (see TODO):

$$k_{\text{dif}} = \left(\frac{69.35}{365} - \text{pos} \cdot M^{-2/3} \right) \frac{\text{m}}{\text{day}}$$

where M is the molecular weight of the chemical (internal parameter); finally, $f_{D,w}$ and $f_{D,s}$ are given by:

5.3.4.4 Settling rate

The daily settling rate is computed as:

$$s(t) = \frac{k_s(1 - f_{D,w})}{h(t)}$$

where k_s is a (chemical-specific) internal parameter.

5.3.4.5 Washout rate

The daily washout rate is computed as:

$$w(t) = \mathrm{fet} \cdot P(t)$$

where $P(t)$ is the daily precipitation per unit area, while fet is a (chemical-specific) internal parameter.

5.3.4.6 Solubility

The solubility ρ of chemical is an internal parameter.

6 Risk assessment

6.1 Overview

6.2 Calculation of risk using SSDs

Part II

User Manual

7 The ERAHUMED DSS User Interface

This chapter should explain how to run ERAHUMED simulations using the Shiny app. It may contain screenshots taken from the app to exemplify the various points.

7.0.1 How to run the DSS?

Describe various options available (which at the moment of writing may as well be “download the package” only - and perhaps a basic deployment on shinyapps.io).

7.0.2 The “Output” tab

7.0.3 The “Input” tab

8 The {erahumed} R package

This should not be an exhaustive description of the R package, but rather mention its existence and giving basic instructions for its installation and refer to the package vignette's and documentation for more details.

References

Martínez-Megías, Claudia, Alba Arenas-Sánchez, Diana Manjarrés-López, Sandra Pérez, Yolanda Soriano, Yolanda Picó, and Andreu Rico. 2024. "Pharmaceutical and Pesticide Mixtures in a Mediterranean Coastal Wetland: Comparison of Sampling Methods, Ecological Risks, and Removal by a Constructed Wetland." *Environmental Science and Pollution Research* 31 (10): 14593–609.

A Input Data

A.1 Hydrological data

A.2 Meteorological data

A.3 Albufera Rice Paddies Management

A.4 Storage curve and P-ETP function

A.5 Definition of rice clusters