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 Hydrological Balance

This chapter describes the calculation of quantities related to hydrological balance. There are two levels of modeling involved:

- "Global" balance, *i.e.* the balance of the Albufera Lake. This step involves a reduced set of assumptions, and really boils down to the calculation of the lake's daily total inflow, from its measured outflows and precipitation/evapotranspiration levels, and from the measured water level changes.

- "Local" balance, that is the hydrological balance of rice paddy clusters in the Albufera National Park. This step uses the output of the global balance calculations, and involves a much wider set of assumptions about how the draining and irrigation of rice paddies is managed, to approximate as closely as possible the ideal water levels required by sowing, fertilization, *etc.* during the year.

2.1 Global balance of the Albufera Lake

This modeling layer computes the total daily inflow to the Albufera Lake, from the measured outflows and precipitation/evapotranspiration levels, together with the measured water level changes. The relevant equation expressing the hydrological balance is:

$$\text{Inflow} + \text{Outflow} = \text{Volume Change} - \text{Rain} + \text{Evapotranspiration} \tag{2.1}$$

where the unknown is Inflow. The subsections that follow explain in detail how the various volume terms appearing in Equation 2.1 are actually obtained from measured quantities.

2.1.1 Volume Change term

The lake volume change is obtained from the measured water height change, assuming volume and height are related through a linear storage curve:

$$\text{Volume} = a + b \cdot \text{Height} \tag{2.2}$$

The volume *change* that appears in Equation 2.1 is given on day d by:

$$\text{Volume Change}(d) = \text{Volume}(d+1) - \text{Volume}(d) \tag{2.3}$$

The default values used for a and b are discussed in Appendix Section A.4.

2.1.2 Rain and Evapotranspiration terms

The volume changes due to precipitation and evapotranspiration are also assumed to be linearly related to the corresponding measured water levels (usually expressed in mm):

$$\begin{aligned} \text{Rain}(\text{m}^3) &= \alpha \cdot \text{Rain}(\text{mm}) \\ \text{Evapotranspiration}(\text{m}^3) &= \beta \cdot \text{Evapotranspiration}(\text{mm}) \end{aligned} \tag{2.4}$$

The default values used by the software α and β are discussed in Appendix Section A.4.

2.1.3 Outflow term

The Albufera lake has two sources of outflow, namely estuaries (*golas*) and enclosed marshlands (*tancats*) [TODO: terms suggested by chatGPT, are these appropriate?], so that we may write:

$$\text{Outflow} = \text{Outflow}_{\text{estuaries}} + \text{Outflow}_{\text{marshlands}} \tag{2.5}$$

The outflows from estuaries are monitored by the *Confederación Hidrográfica del Júcar* (CHJ), which makes the measurements available on a daily basis (see Section A.1). This results in three additive terms:

$$\text{Outflow}_{\text{estuaries}} = \text{Outflow}_{\text{Pujol}} + \text{Outflow}_{\text{Perellonet}} + \text{Outflow}_{\text{Perelló}} \tag{2.6}$$

corresponding to the three estuaries of the Albufera lake: *Gola de Pujol*, *Gola del Perellonet* and *Gola del Perelló*.

In contrast, the outflow from marshlands is not a measured quantity, and can only be estimated through some additional assumption. The relevant constraint, with a small leap in logic, is that for each day in the considered time series, either $\text{Outflow}_{\text{marshlands}}$ or Inflow must be zero. [TODO: what is the underlying assumption here?]

2.2 Local balance of rice paddy clusters

As anticipated, the simulation of local hydrological balance of rice paddies involves an important set of assumptions regarding the management of water levels. We detail in this section the algorithm used in our simulations.

2.2.1 Model inputs and parameters

The two main observational inputs for this modeling layer are:

- The output of the "Global balance" model, described in the previous Section, and in particular the daily total inflow to the Albufera lake, obtained from Equation 2.1. As explained below in Section 2.2.2, this is used to estimate the inflow from each of the twenty-six ditches that flow into the Albufera.

- The measured amounts of precipitation and evapotranspiration, which is the same data-set involved in global balance calculations, and is described in Section A.2.

The subdivision of the rice fields surface of the Albufera into clusters of a specific rice variety, and the corresponding assignation of a single ditch to each cluster is described in Section A.5.

Our assumptions on the management of water levels are subsumed in the data-set described in Section A.3, that provides the irrigation and draining plans and the ideal water levels for a cluster of a given rice variety, for each day of the year.

Finally, the model receives a few other numerical input parameters:

- *Ideal flow rate k.* Rate at which water flows through rice paddies when these are being simultaneously irrigated and drained, with the overall level being kept constant. Expressed in $\mathrm{cm} \cdot \mathrm{day}^{-1}$. [TODO: this must be some cumbersome paraphrase of a technical expression which I don't know about.].

- *Height threshold H_{thres}.* Maximum allowed water level for a cluster to be considered emptied, used in the calculation of draining/irrigation plan delays. Expressed in cm.

2.2.2 Ditch inflows

The first step consists in breaking down the total inflow to the lake, obtained through Equation 2.1, into inflows from individual ditches. Each paddy cluster is assumed to communicate with a single ditch (*cf.* Section A.5), and the flow through a ditch is assumed to be proportional to the total area of clusters belonging to said ditch. Therefore, denoting by I_i and A_i the inflow and area of ditch i, and by I the total inflow, we compute:

$$I_i = \frac{A_i}{\sum_j A_j} \cdot I \tag{2.7}$$

2.2.3 Main Algorithm

In the following, we focus on the set of clusters that communicate with a specific ditch, whose inflow (estimated according to Section 2.2.2) we denote by Q. The essence of the algorithm is to approximate as closely as possible the cluster's *ideal* inflows and outflows, with the constraint that the sum of the actual outflows from all clusters must equal Q.

Let us start by setting up some basic notation. We denote by:

$$h_c^{\mathrm{id}}(t), \quad I_c^{\mathrm{id}}(t), \quad O_c^{\mathrm{id}}(t), \tag{2.8}$$

the ideal water level, inflow and outflow of cluster c at time (*i.e.* day) t, and by:

$$h_c^{\mathrm{re}}(t), \quad I_c^{\mathrm{re}}(t), \quad O_c^{\mathrm{re}}(t), \tag{2.9}$$

the corresponding real quantities. The area of cluster c, that provides the conversion between water volumes and column heights, is denoted by A_c. Precipitation and evapotranspiration water column values are denoted by:

$$\mathrm{P}(t), \quad \mathrm{ETP}(t). \tag{2.10}$$

Finally, we will define below a *plan delay* accumulated for cluster c at time t, denoted $D_c(t)$. This is computed recursively along the iterations of the algorithm, and its role will be clarified below.

The local balance algorithm proceeds iteratively as follows. At day t, assume that $h_c^{\mathrm{re}}(t-1)$ and $D(t-1)$ have already been computed; a single iteration consists of the following steps, which we describe in full detail below:

- **Step 1.** Recovering the $h_c^{\mathrm{id}}(t)$, $I_c^{\mathrm{id}}(t)$ and $O_c^{\mathrm{id}}(t)$ from the irrigation and draining plan, applying the computed plan delay $D_c(t-1)$.

2. **Step 2.** Computing $h_c^{\mathrm{re}}(t)$, $I_c^{\mathrm{re}}(t)$ and $O_c^{\mathrm{re}}(t)$ enforcing the constraint $\sum_c O_c^{\mathrm{re}}(t) = Q(t)$.

3. **Step 3.** Computing the updated plan delays $D_c(t)$.

In order to initialize the iteration, we assume that at some initial time, say $t = 0$, we have $D_c(t = 0) = 0$ $h_c^{\mathrm{re}}(t = 0) = h_c^{\mathrm{id}}(t = 1)$.

2.2.3.1 Step 1: ideal balance

The management data-set described Section [A.3](#) provides the ideal water level for each cluster and for every day of the year. Denoting by $d(t)$ the day of the year corresponding to t, the relevant entry of the management data-set is that corresponding to day:

$$d_c'(t) = d(t) - D_c(t-1), \tag{2.11}$$

where $D_c(t-1)$ is the accumulated plan delay for thi cluster. Therefore we set:

$$h_c^{\mathrm{id}}(t) = \text{Planned water level on day of year } d_c'(t). \tag{2.12}$$

To compute ideal inflow and outflow, we require:

$$h_c^{\mathrm{id}}(t) = \max(0,\ h_c^{\mathrm{re}}(t-1) + \mathrm{P}(t) - \mathrm{ETP}(t)) + I_c^{\mathrm{id}}(t) - O_c^{\mathrm{id}}(t), \tag{2.13}$$

where we assume $I_c^{\mathrm{id}}(t) > 0$ and $O_c^{\mathrm{id}}(t) > 0$. Clearly, Equation [2.13](#) alone does not individually specify $I_c^{\mathrm{id}}(t)$ and $O_c^{\mathrm{id}}(t)$, but only their difference $\Delta_c(t) = I_c^{\mathrm{id}}(t) - O_c^{\mathrm{id}}(t)$. In order to fix both these quantities, we first define[1]:

$$
\begin{aligned}
(I_c^{\mathrm{id}}(t))^{(0)} &= \begin{cases} k & \text{cluster planned to be in flux on day of year } d_c'(t) \\ 0 & \text{otherwise} \end{cases}, \\
(O_c^{\mathrm{id}}(t))^{(0)} &= (I_c^{\mathrm{id}}(t))^{(0)} - \Delta_c(t).
\end{aligned}
\tag{2.14}
$$

and, in order to ensure that flows are positive[2], we finally set:

$$
\begin{aligned}
O_c^{\mathrm{id}}(t) &= \max\{(O_c^{\mathrm{id}}(t))^{(0)},\ 0\} \\
I_c^{\mathrm{id}}(t) &= O_c^{\mathrm{id}}(t) + \Delta_c^{\mathrm{id}}(t)
\end{aligned}
\tag{2.15}
$$

---

[1]The condition is again evaluated using the the management data-set, where the relevant variables are the `irrigation` and `draining` columns. The cluster is understood to be in flux if both `irrigation` and `draining` are `TRUE`.

[2]The positivity of $I_c^{\mathrm{id}}(t)$ below is ensured by the fact that $(O_c^{\mathrm{id}}(t))^{(0)} \geq -\Delta_c(t)$ by construction.

2.2.3.2 Step 2: real balance

Real flows are obtained from ideal ones (Equation 2.15) in such a way to satisfy the constraint:

$$\sum_c O_c^{\text{re}}(t) = Q(t), \tag{2.16}$$

where the right-hand side is the total ditch flow computed earlier (Section 2.2.2). At each time-step t, the cluster's index set is randomly permuted[3], and the real flows are calculated as:

$$O_c^{\text{re}}(t) = \min\{O_c^{\text{re}}(t),\, Q(t) - \sum_{c'<c} O_{c'}^{\text{re}}(t)\} + \frac{\max\{0, Q(t) - \sum_{c'} O_{c'}^{\text{re}}(t)\}}{\text{number of clusters}}, \tag{2.17}$$
$$I_c^{\text{re}}(t) = \max\{I_c^{\text{id}}(t) - O_c^{\text{id}}(t) + O_c^{\text{re}}(t),\, 0\}$$

In words, clusters are emptied in a random order within the allowed capacity of the corresponding ditch (*i.e.* its actual total flow) - if the sum of ideal outflows is less than capacity, the remaining outflow is equally shared among clusters. Using Equation 2.17, we finally determine the real water level achieved as:

$$h_c^{\text{re}}(t) = \max(0,\, h_c^{\text{re}}(t-1) + \text{P}(t) - \text{ETP}(t)) + I_c^{\text{re}}(t) - O_c^{\text{re}}(t), \tag{2.18}$$

to be compared with Equation 2.13.

2.2.3.3 Step 3: updating the plan delay

The purpose of the plan delay $D_c(t)$ is to allow all clusters to be emptied as required by the ideal management plan, which may be hindered on the originally scheduled days by the first of Equation 2.17, since this sets to zero the real outflows for some clusters whenever the ditch's flow is saturated.

The updated value $D_c(t)$ is obtained as follows. If $d(t)$ (the *actual* day of year) is outside of the window $W = [20\text{th of April}, 15\text{th of October}]$, we reset all $D_c(t) = 0$. Otherwise, if $h_c^{\text{id}}(t) > 0$ or $h_c^{\text{re}}(t) < H_{\text{thres}}$, the plan delay is unchanged for this cluster: $D_c(t) = D_c(t-1)$. Finally, if $h_c^{\text{id}}(t) = 0$ but $h_c^{\text{re}}(t) > H_{\text{thres}}$, we set $D_c(t) = D_c(t-1) + 1$.

---

[3]With some abuse of notation, we assume the indexes c and c' in Equation 2.17 to be sorted according to this random permutation.

Part II

User Manual

3 The ERAHUMED DSS User Interface

4 The {erahumed} R package

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 Albufera Lake Hydrological Balance

A.2 Precipitation and Evapotranspiration

A.3 Albufera Rice Paddies Management

A.4 Storage Curve and P-ETP Surface Values

A.5 Definition of rice clusters