The daemons of water times

Riccardo Rigon[1] and Marialaura Bancheri[2]

[1]University of Trento
[2]Affiliation not available

January 11, 2022

This is research material under development. If used, it should be cited properly. Citations to published papers is scanty and it should be fixed along the way.

To account for water movements in a catchment or any other control volume, let us assume that we can record the movements of small group of water molecules, called **parcels**, inside the control volume (**cv**) and at its boundaries. We do some assumptions first:

- When it rains over the control volume we have a differently coloured rainfall for any different (clock) time. Rain is whimsical and it does not fall at each time step but quite randomly.

- After falling, parcels move inside the control volume also with some randomness and gain the cv boundaries (and beyond).

All the rainfall is monitored by three little water daemons, He, She and They. All of them have a clock synchronised at the same time and date, and each one is able of an action, at any time step:

- **She** counts the parcels of any water color **inside the cv**

- **He** counts the parcels of any water of any color that **exit the cv**

- **They** made calculations with the data received in form of table from She and He

He and **She** annotate their action in two tables whose rows are marked by the current clock time and whose columns are the injection times, t_{in} the times when rain have fallen. Clock time is continuous but sampled at discrete time steps while t_{in} is discrete cause the nature of precipitation. The matter can be presented in a quite abstract manner, but here we rely on a simple example.

Therefore, assuming that we had three precipitations at time $t_{in} = \{0, 3, 5\}$ of mass (or volumes or moles, the units specification is unnecessary at moment) $P = \{20, 18, 15\}$ respectively , **She** and **He** can organise their record according to the following tables which grow one row any time step and which we show here in their definitive form, after all precipitations we are interested about exited the cv:

For She, it is:

$t_{in} \rightarrow$	0	3	5
$t \downarrow$			
0	20	-	-
1	17	-	-
2	13	-	-
3	10	8	-
4	8	6	-
5	3	3	15
6	1	2	12
7	0	2	8
8	-	0	5
9	-	-	3
10	-	-	0
11	-	-	-

For He, it is:

$t_{in} \rightarrow$	0	3	5
$t_{ex} \downarrow$			
0	0	-	-
1	3	-	-
2	4	-	-
3	3	0	-
4	2	2	-
5	5	3	0
6	2	1	3
7	1	0	4
8	0	2	3
9	-	0	2
10	-	-	3
11	-	-	0

When **He** marks a measure his time is by definition, t_{ex}. In their asymptotic form (ideally at $t = \infty$) the two tables contain the same information (say the **She** table in a "integral" form, and the **He** table in a "derivative" form), but for any time t, the tables contain mostly complementary information that **They** must use jointly to reconstruct where waters went.

They found useful to build the cumulative **He** table , i.e.

$t_{in} \rightarrow$	0	3	5
$t_{ex} \downarrow$			
0	0	-	-
1	3	-	-
2	7	-	-
3	10	0	-
4	12	2	-
5	17	5	0
6	19	6	3
7	20	6	7
8	-	8	10
9	-	0	12
10	-	-	15
11	-	-	0

to observe that the sum element by element of the **She** table and the of the cumulative **He** table gives a constant in any column, equal to the the total amount of precipitation fallen at the time t_{in}. A more colorful representation of the tables above was given in (Hrachowitz et al., 2013), Figure 4, and in the more recent (Hrachowitz et al., 2016) which is rerpoduced here below
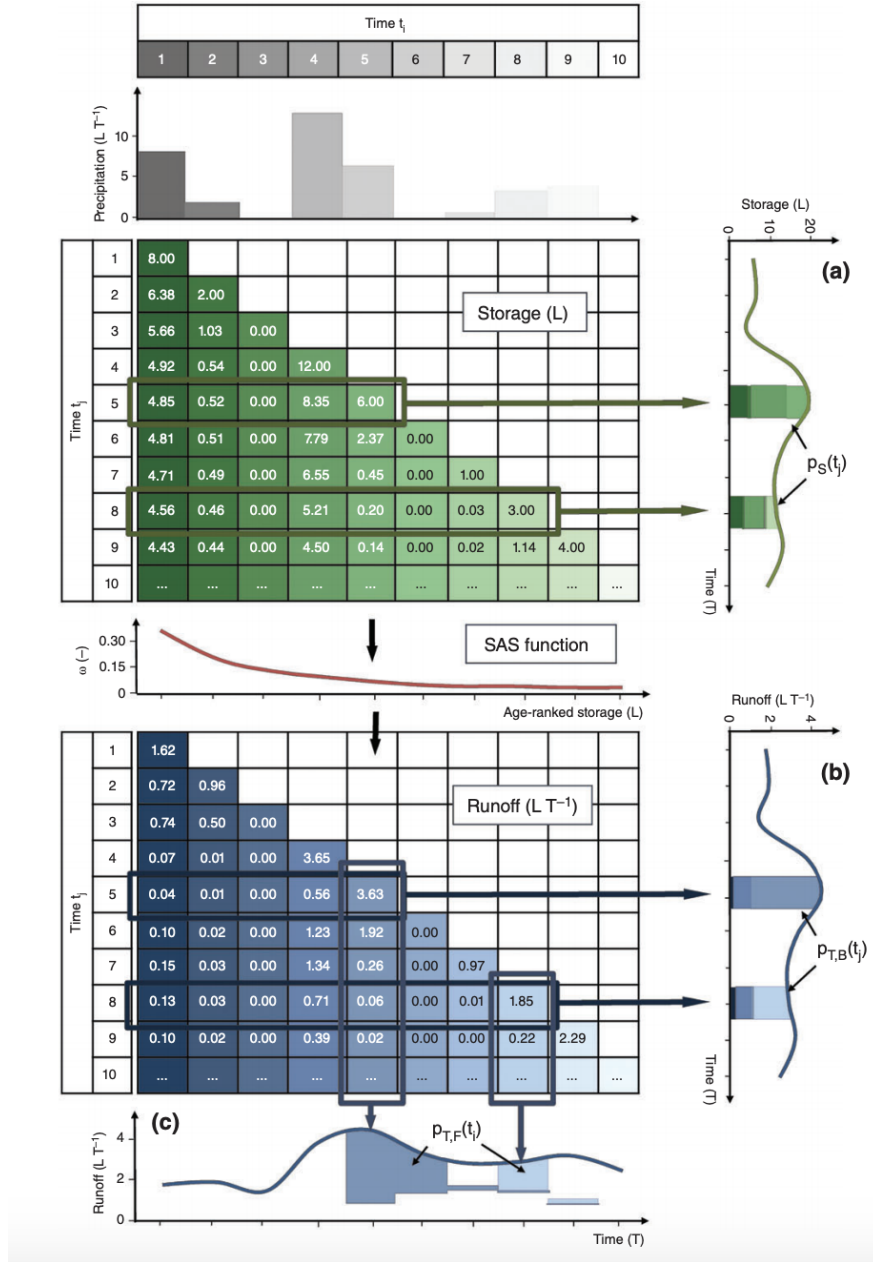
Figure 1: Figure from Hrachowitz et al, 2016. The top Figure is the She table, the bottom Figure is the He Table. The right side figures represent the storage at clock time t (top right) and the discharge (bottom right). On top there is the hyetograph, on bottom the discharge is reproduced also, but it is misleading because it would be more natural to represent a different function (see below). At the center the storage selection function (see below) is also presented. Note: this figure should be modified and moved elsewhere below.

Putting in formulas what **He**, **She** and made is, as **They** knows, requires some abstraction. **They** in fact defines:

- $v(t, t_{in}, t_{ex})$ as the mass of the water stored in the cv at time t which entered at time t_{in} and **will** exit at time t_{ex}.

The 3-dimensional function v is not actually accessible to **They**, because it reflects the whole story of water parcels but **They** found useful to start to build definitions from it. **They** first assumed to have just one rainfall to analyze. **They** observed then that, by definition, it is $t_{in} \leq t_{ex}$, while for the clock time there exists various possibilities, and, for any t_{in}, there is a sequence of $T_{ex} = \{t_{ex_1}, t_{ex_2}, ..., t_{ex_n}\}$, as many as the number of non zero entries, after t, for a t_{in} columns under scrutiny in **He** table (which could be infinite). Therefore we have the following cases:

- $t \leq t_{in} < t_{ex_1}$: in this case $v(t, t_{in}, t_{ex}) = 0$ because not any rain was received.

- $t_{in} < t_{ex_i} = t < t_{ex_j}$: then $v(t, t_{in}, t_{ex}) \neq 0$ and a function of t_{ex_i} of value given by the appropriate entry of asymptotic **He** table for any value of $t_{ex} > t$. (This means that the **They** algorithm ideally runs over all the t and over the t_{in} larger than t)

- $t_{in} < t_{ex_n} < t$: In this last case $v(t, t_{in}, t_{ex}) = 0$

If more than one t_{in} is present, the assignment of values must be repeated for any t_{in}. From the analysis it is apparent that all the $v(t, t_{in}, t_{ex})$ information can be derived from the asymptotic **He** table using t as an index.

For instance, if we set $t = 4, t_i n = 0$, $t_{ex} = 4$, the volume is $v(t = 4, t_{in} = 0, t_{ex} = 4) = 2$, as well as $v(t = 4, t_{in} = 4, t_{ex} = 4) = 2$.

We can also define:

- $v_q(t, t_{in}, t_{ex})$ as the mass of water which **entered** the cv at time at time t_{in} and exited at any time $t_{ex} < t$, (i.e. $t_{in} \leq t_{ex} < t$). Then, it results:

$$v(t, t_{in} = \tau, t_{ex}) + v_q(t, t_{in} = \tau, t_{ex}) = P(\tau) \tag{1}$$

where τ is the common water injection time (this is equivalent to sum the same column of the **She** and cumulative **He** tables).

If we integrate $v(t, t_{in}, t_{ex})$ over all the future exit time, we obtain:

$$s(t, t_{in}) = \int_t^\infty v(t, t_{in}, t_{ex}) dt_{ex} \tag{2}$$

which is the mass of water inside the cv that entered at t_{in} (independently of when they will go out. This $s(t, t_{in})$ is what She measures in her table. There is an apparent the conundrum that while integration in (2) is on all future times, $s(t, t_{in})$ it is known at time t : but this is due simply to mass conservation of which the future must be respectful. By definition:

$$\int_0^t v(t, t_{in}, t_{ex}) dt_{ex} = 0 \tag{3}$$

because at time t all the parcels that have $t_{ex} \leq t$ are already out of cv.

Symmetrically, we can integrate over t_{in}:

$$v_{ex}(t, t_{ex}) = \int_0^t v(t, t_{in}, t_{ex})dt_{in} \tag{4}$$

which represents the mass of the parcels that entered any time prior to t and will exit at t_{ex}. The $v_{ex}(t, t_{ex})$ is, for any row in the asymptotic **She** table, the sum over all the columns (all the existing $t_{in} \leq t$) containing data. They observes also that $v_{ex}(t, t_{ex})$ is essentially unknown at time t because so are the specifications of the future exit times. Therefore $v_{ex}(t, t_{ex})$ at present $t$remains a theoretical construct.

Besides, They suggests to further integrate (3) and (4) to obtain the total mass inside the cv:

$$S(t) = \int_t^\infty \int_0^t v(t, t_{in}, t_{ex})dt_{in}dt_{ex} \tag{5}$$

$S(t)$ is completely known at any time t because, as already remarked in comment to equation (2), the integral over future time gives $s(t, t_{in})$ which are the entries of **She** table at clock time t in the column t_{in}. The reader can observe that

$$p_S(t - t_{in}|t) := \frac{s(t, t_{in})}{S(t)} \tag{6}$$

as defined in (Rigon et al., 2016), is the conditional probability distribution known as **residence time pdf** , while

$$p_{ex}(t_{ex} - t|t) := \frac{v_{ex}(t, t_{ex})}{S(t)} \tag{7}$$

is the (**resident**) **parcels life expectancy** pdf, conditional on t. To better understand what the previous definitions mean, it is useful to understand how they can be visualisrd by **They**, using the **She** and **He** tables. Please first observe that being at a time t means to separate the tables into at least into the three regions in figure below

6

$t_{in} \rightarrow$	0	3
t $\downarrow$		
0	20	–
1	17	–
2	13	–
3	10	8
4	8	6

Future

Future depending on external inputs

Past

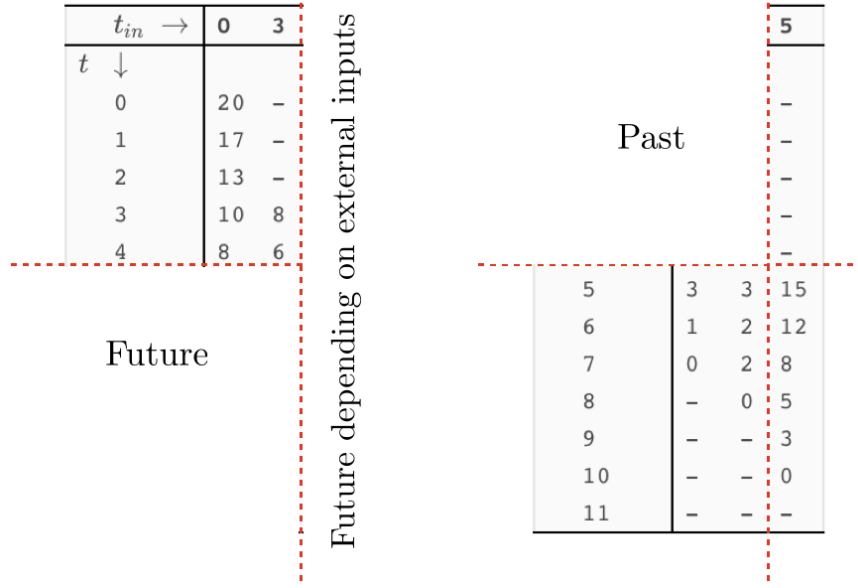|			5
			–
			–
			–
			–
5	3	3	15
6	1	2	12
7	0	2	8
8	–	0	5
9	–	–	3
10	–	–	0
11	–	–	–

Figure 2: This is **She** table subdivided with respect to time $t = 4$. What we exactly know at this time is the left figure. The right figure represents what we do not know at that time, which is divided into two parts: The future depending on past inputs to the system (say the known unknown) and the future depending on external inputs (say the unknown unknown).

Let us consider at first the case of the past according to **She** table. The sequence of masses $S(t)$ is obtained by They summing over all the columns belonging to the same row, obtaining the following pairs:

$$S(t) = \{(0, 20), (1, 17), (2, 13), (3, 18), (4, 14)\}$$

where the first number of each pair is the clock time and the second number is the sum over the the columns in the corresponding row. For instance, for $t = 3$, the sum is $S(3) = 10 + 8 = 18$.

Within the past knowledge, as we already said, They is able to estimate $s(t, t_{in})$ which is each single entry visible in the left table, and therefore the residence time pdf, which is the ratio between each entry and the row summation $S(t)$ above. For the way the normalization factor $S(t)$ is obtained, They gets a different probability function for each row.

To obtain the residence life expectancy we can relay on $S(t)$ (which is , for instance, 14 for $t = 4$, i.e. $S(4) = 14$) but They has to use the **He** table to get the numerator appearing in the definition of equation (7). Unfortunately the desired numerator numbers are in the future domain and it is unknown, unless some divination have been acquired or some assumption have been made.

Figure 3: This is the known unknown future of the He table at time $t = 4$

Therefore the entries of the numerator in equation (7) are given by the following $v_{ex}(4, t_{ex}) = \{8, 3, 1, 2\}$ with $t_{ex} = \{5, 6, 7, 8\}$ respectively.

They remarks that there is a difference between life expectation and age. In fact, parcels having different ages, i.e. $age = t - t_{in}$, can have the same life expectation $le = t_{ex} - t$. But obviously **They** can manage to make the life expectation, $le(t)$ conditional on the age, $age(t)$.

The average age of resident parcels, according to definitions is given by:

$$\overline{age} = \int_0^t (t - t_{in}) p_s(t - t_{in}|t) dt_{in} \tag{8}$$

while the average life expectation is:

$$\overline{le} = \int_t^\infty (t_{ex} - t) p_{ex}(t_{ex} - t|t) dt_{ex} \tag{9}$$

They also suggests that life expectancy can be further made conditional on age. i.e. on t_{in}. In this case a new doubly conditioned probability should be introduced, which results, by Bayes formula:

$$p_{ex}(t_{ex} - t|t, t_{in}) = \frac{p_{ex}(t_{ex} - t|t)}{p(t_{in}|t)} \tag{10}$$

8

By construction it is also,

$$p_{ex}\left(t_{ex} - t | t, t_{in}\right) := \frac{v_{ex}(t_{in}, t_{ex})}{s(t, t_{in})} \qquad (11)$$

because, among the parcels inside the cv, injected at t_{in}, measured by **She**, we are interested to those exiting the cv after t. By comparison from (11) and (10) it follows that:

$$p(t_{in}|t) = \frac{s(t, t_{in})}{S(t)} \qquad (12)$$

which is, coherently, the fraction of not exited parcels, over all the non exited parcels, or one element of the She table, divided the sum over all the elements in the same row of the same table.

What obtained is not the only information got from the past. **They** can consider in fact $q(t_{ex}, t_{in})$, i.e. the entries of the He table to be normalized by $S(t)$, the sum of the entries for $t = t_{ex}$ of the **She** table, being the total water inside the cv at time t:

$$p_Q(t - t_{int}|t) := \frac{q(t_{ex} = t, t_{in})}{S(t)} \qquad (13)$$

which is the backward probability of dicharges which is obviously known for any $t_{ex} \leq t$.

While not strictly necessary here, They finds useful also to define:

$$\omega(t, t_{in}) := \frac{p_Q(t - t_{in}|t)}{p_Q(t - t_{in}|t)} = \frac{s(t, t_{in})}{q(t, t_{in}} \qquad (14)$$

where the ωs are called StorAge Selection functions (SAS), which can be used for some analytical manipulation. Clearly, by definition, they are the element by element ratio of the **She** and the **He** tables.

As it is known, for instance from (Rigon et al., 2016), They can elaborate the He table in a different way. In Figure 3 a single column of the **He** table, $t_{in} = 0$ is highlighted.

$t_{in} \to$	0	3	5
$t_{ex} \downarrow$			
0	0	–	–
1	3	–	–
2	7	–	–
3	10	0	–
4	12	2	–
5	17	5	0
6	19	6	3
7	20	6	7
8	–	8	10
9	–	0	12
10	–	–	15
11	–	–	0

Figure 4: He's table with the situation subdivided in past and future with respect to clock time $t = 4$. The column for $t_{in} = 0$ is highlighted.

This column is the discharge at the cv boundaries conditional to the injection time $t_{in} = 0$, $q(t, 0)$. They can access also the additional information, coming from **She** table that the total input at time $t = t_{in} = 0$ is equal to $P(0) = s(t = 0, t_{in} = 0) = 20$ (and must be equal to the total sum of the terms appearing in column for $t_{in} = 0$, performed after the concentration time).

Thus, **They** can then be properly normalise the column by subdividing all by $P(0)$. In fact the response time distribution, is defined by:

$$p_q(t - t_{in}|t_{in}) := \frac{q(t, t_{in})}{J(t_{in})} \tag{15}$$

Besides They can argue that $q(t, t_{in}) = v_{ex}(t = t_{ex}, t_{in})$ if we are running the clock time t for any of the times larger than t_{in}. However, it is also, as we already mentioned $s(t_{in}, t_{in}) = P(t_{in})$ and therefore:

$$p_q(t - t_{in}|t_{in}) := \frac{q(t, t_{in})}{J(t_{in})} = \frac{v_{ex}(t, t_{in})}{P(t_{in})} = p(t_e x - t_{in}|t = t_{in}, t_{in}) \tag{16}$$

10

Therefore, the response time probability is the life expectation conditioned on t_{in} when $t = t_{in}$. Once having all the response pdf, for all the t_{in}, **They** can reconstruct the overall life expectancy probability, applying the definitions.

They remarks that the response time pdf is actually more general that the life expectancy, because having it, it is possible to get all the conditional life expectancies, just properly normalizing, for any time t, the future age-ranked discharges. When in possess of all the response functions, for all t_{in}, is finally possible to get the unrestricted life expectancy.

References

What can flux tracking teach us about water age distribution patterns and their temporal dynamics?. (2013). *Hydrology and Earth System Sciences*, *17*(2), 533–564. https://doi.org/10.5194/hess-17-533-2013

Transit times-the link between hydrology and water quality at the catchment scale. (2016). *Wiley Interdisciplinary Reviews: Water*, *3*(5), 629–657. https://doi.org/10.1002/wat2.1155

Age-ranked hydrological budgets and a travel time description of catchment hydrology. (2016). *Hydrology and Earth System Sciences*, *20*(12), 4929–4947. https://doi.org/10.5194/hess-20-4929-2016