Comment on egusphere-2024-3184

by J. Lerat

January 8, 2025

1 General comments

The technical note proposed by J. Lerat introduces a new method for the integration of a scalar Ordinary Differential Equation (ODE), describing the evolution of a conceptual component such as a reservoir in a lumped hydrological model. The rationale for the development of the method is that, given this ODE formulated in state-space, it is difficult to compute together both the evolution of the single state variable (denoted S) over a timestep $[t, t+\delta]$, as well as the different outputs from the system *integrated over the timestep* δ. The QuaSoARe method uses analytical (quadratic) substitutes for the instantaneous fluxes entering/leaving the reservoir.

The paper is written in a very straightforward way and is relatively easy to follow with some familiarity in analytical and numerical integration (apart from the very first part where continuity assumptions are stated, see comments below). I am overall supportive of the approach which has a great potential for speeding up current formulations of many lumped models, but I still have questions regarding the kind of solutions which should be put forward to cure the problems raised by Clark and Kavetski about the numerical soundness of many such models, in the case of multistate model structures.

2 Specific comments

2.1 Discrete time vs. state-space formulations and the problem of sub-timestep distribution of inputs / outputs

I think that the introduction somehow lacks a brief presentation of the context in which unbalanced, cumulative outputs from model reservoirs could happen. Equation 1

$$\dot{S} = \frac{dS}{dt} = \sum_i f_i(S, \tilde{V})$$

is straightforward to understand but it already constitutes a "zoom" on a single model element (reservoir), along with two assumptions:

(i) the right-hand-side of this state-space formulated equation does not depend on any other state (e.g. the storage R of an other reservoir),

(ii) in the absence of information about the sub-timestep distribution of forcing variables such as $\tilde{V}$, they are assumed constant over the timestep so that the equation does not explicitly depend on the independent time variable t: we can write $f_i(S, \tilde{V} = \text{cst})$ rather than $f_i(S(t), t)$. Restricting the analysis to such autonomous ODEs (Clark and Kavetski, 2010) is very common and is a lesser loss of generality than assumption (i), but it is still worth mentionning.

Assumption (i) can basically be met with operator splitting (OS) procedures, and this point will be discussed later. My first question would be: in validating model outputs against observations (e.g. runoff, actual evapotranspiration, recharge, etc.), why wouln't Eq. 3 be suitable for *flux* comparison, and why do we absolutely need to compute integrated quantities such as the $O_i's$ (runoff *depth* over δ, evapotranspiration *depth*, recharge *depth*, etc.) which have dimension L (length), rather than to their instantaneous counterparts $\frac{dO_i}{dt}$ (runoff *rate*, evapotranspiration *rate*, etc. with dimension $\mathrm{L} \cdot \mathrm{T}^{-1}$) at successive times $t, t+\delta, t+2\delta$, etc.

My understanding is that models targeted by the QuaSoARe procedure are essentially *balance* models whose purpose is mainly to ensure mass conservation at quite large time steps (e.g., daily), i.e. to yield a value for $S[\delta]$ and a set of values for the $\{O_i\}_{1 \leq i \leq n}$ such that:

$$(B) \begin{cases} S(\delta) - S(0) \approx \sum_{i=1}^{n} \int_0^{\delta} f_i(S, \tilde{V}) dt & \text{(approximately)} \\ S(\delta) - S(0) = \sum_{i=1}^{n} O_i & \text{(strict equality)} \end{cases}$$

QuaSoARe clearly meets these requirements, because it uses an analytical substitute $\hat{f}_i(S, \tilde{V})$ for each flux, such that

$$(\hat{B}) \begin{cases} S(\delta) - S(0) = \sum_{i=1}^{n} \int_0^{\delta} \hat{f}_i(S, \tilde{V}) dt \approx \sum_{i=1}^{n} \int_0^{\delta} f_i(S, \tilde{V}) dt & \text{(approximately)} \\ O_i = \int_0^{\delta} \hat{f}_i(S, \tilde{V}) dt \qquad \forall i \\ S(\delta) - S(0) = \sum_{i=1}^{n} \int_0^{\delta} \hat{f}_i(S, \tilde{V}) dt = \sum_i O_i & \text{(strict equality yielded by previous eqns)} \end{cases}$$

Clearly, solving flux equations individually is a *sufficient* condition to fullfill the conditions (B), but I rather disagree on the fact that it is a *necessary* one, as stated in lines 60–70 ("*Finally, jointly solving Eq. 1 and Eq. 2 using a numerical solver* ***requires*** *transforming the scalar equation Eq. 1 to a system of differential equations by adding one scalar equation for each flux*"). If we have a procedure to solve the single, state-space formulated ODE in S, attributing residual errors in cumulative balance to any of the outputs would yield the same result at lower computational cost. If we chose an output O_{i_0} for balance error compensation, there are several possible fixes allowing to fullfill conditions (B), such as:

$$(B')\begin{cases} O_i &= f_i\left(\dfrac{S(0)+S(\delta)}{2},\tilde{V}\right) \quad \forall i \neq i_0 \\ O_{i_0} &= S(\delta)-S(0)-\displaystyle\sum_{i\neq i_0} O_i \end{cases} \qquad (B'')\begin{cases} O_i &= \dfrac{1}{2}\left[f_i\left(S(0),\tilde{V}\right)+f_i\left(S(\delta),\tilde{V}\right)\right] \quad \forall i \neq i_0 \\ O_{i_0} &= S(\delta)-S(0)-\displaystyle\sum_{i\neq i_0} O_i \end{cases}$$

Using the language of atmospheric modeling we might say that the O_i's are some kind of *diagnostic* variables that could be computed at the end of the simulated time period, S being the only variable truly belonging to the category of *state variables*. Adding this distinction in the formulation of the problem could improve the paper and help modelers think about the way their models are written (either discrete time, or state-space). We must acknowledge that this issue about *individual fluxes* computations raised by J. Lerat is even less discussed in the hydrological litterature than that of numerical scheme adequacy. The issue of solution constraints (storage remaining non-negative, as well as fluxes such as evaporation, etc.) is discussed in Clark and Kavetski (2010) along with the semi-implicit Euler scheme applied to an autonomous system of N_s coupled ODEs in the form:

$$\dot{\mathbf{S}} = \mathbf{g}(\mathbf{S})$$

These authors propose to compute the k-th *individual flow volume* for store i using the end-of-time-step (index $n+1$) value of the instantaneous flux multiplied by the time step duration, i.e.

$$O_i^{(k)} = \Delta t \; g_i^{(k)}(\mathbf{S}^{n+1})$$

They acknowledge that "ad hoc fixes such as zeroing negative fluxes works in simple cases but are not a satisfactory general solution", and again the issue of multi-state models arises since "in multistate models where the states are couple via the fluxes, fixing the violation in one state can impact on the feasibility of another". This would be my main question about a possible extension of QuaSoARe to the case of multistate models: can we avoid operator splitting, and if no, shouldn't we still favor direct integration of the full, coupled system of ODEs describing the evolution of the state *vector* $\mathbf{S} \in \mathbb{R}^m$ using the semi-implicit Euler (SIE) scheme for example?

2.2 Mathematical assumptions on the flux functions $f_i(S,\tilde{V})$

The introduction of the paper quickly gets us into the swing of things with the mathematical assumptions stated p.2 (l. 36–41). It is a bit difficult to understand the practical implications of the Lipschitz-continuity assumption: does it mean that some functions are not admissible as flux functions? If we consider the classical equation of a draining tank, according to Bernoulli equation the spout exit velocity is given by:

$$u(t) = \sqrt{2gh(t)}$$

where $h(t)$ is the water depth in the reservoir. Reformulating this as a storage-output equation using $Q = a \cdot u$ with a the spout exit section and using A the base section of the reservoir such that $h(t) = S(t)/A$, we have

$$Q(t) = \frac{dS}{dt} = a\sqrt{2g\frac{S(t)}{A}}$$

This function is not Lipschitz-continuous on the possible range of storage values since the derivative of $\sqrt{x}$ is unbounded as x tends to zero; does it mean that we have to check each flux function?