

Opportunities and challenges for the use of scintillometer-based catchment-averaged evapotranspiration estimates as model forcing

Bruno Samain¹, and V. R. N. Pauwels²

¹Provincie Oost-Vlaanderen – Department Integrated Water Management, Ghent, Belgium

²Department of Civil Engineering, Monash University, Clayton, Victoria, Australia

Submitted to *Hydrological. Earth System Sciences*

In the reply to the issues raised by the reviewer, the remarks by the reviewer are written in italics, and the replies in normal font. In the citation of the remarks, the original page, paragraph, table, figure, and line numbers have been used.

1. Reply to the issues raised by Anonymous Referee #1

1.1. General comments

1. *The manuscript presents a study on the use of actual evapotranspiration (AE) estimates as inputs of a rainfall-runoff (RR) model, instead of the more conventional potential evapotranspiration (PE) inputs. Since AE measurements are becoming widely available, this issue is particularly relevant while relatively novel. Despite the clear theoretical advantage of AE measurements to constrain conceptual RR models, the results appear quite disappointing. While these disappointing results should not prevent publication, I am not sure the followed methodology is really appropriate. It is clear that the authors have the material to reach quite interesting conclusions, but more should be done on the RR model structure and parameterization/optimization.*

We thank the reviewer for his support of the novelty of our study. Regarding the remark concerning the methodology, please see below (specific comments).

1.2. Specific comments

2. *My main concern is on the use of AE within a pre-calibrated RR model. As pointed out by the authors RR model parameter values are largely influenced by the inputs used during calibration, and consequently, the modification in the inputs without recalibrating model parameters is very often followed by a decrease in model performance. However, the authors did not consider a re-calibration of RR model parameters while they change PE inputs. This is all the more the case when they modify the structure of the RR model AND the inputs. Given the accuracy of AE estimates and the fact that no recalibration of the model is allowed in the paper, the disappointing results appeared logic to me. To my opinion, a wiser approach at this stage could consist in calibrating the RR model with the same structure (i.e. by using PE inputs) so that AE estimates by the model fit AE measurements (and also flow within a multi-objective framework).*

We certainly agree with the reviewer that using different types of inputs without recalibration will lead to worse results. We would like to emphasize that we have performed a recalibration of the model with the different

evapotranspiration inputs (not shown or mentioned in the paper). This resulted in a model performance that was at least as good as the results from the initial calibration. However, since this can be expected, we did not discuss this in the paper.

The real objective of the presented paper is assessing to what extent the use of a model with different PET estimates will lead to worse results. This may seem evident, but in many cases PET inputs from different sources are used simultaneously in rainfall-runoff models. The first part of our study shows (again) that this leads to a deterioration in the model results.

Furthermore, Oudin et al. (2005b) have shown that the temporal resolution of the PET measurements is not important for the discharge simulations from rainfall-runoff models. However, what we have shown is that this also leads to a deterioration of the modeled actual evapotranspiration. In other words, if one only examines the modeled discharge rates, the coarsening of the temporal resolution of the evapotranspiration inputs may have only a very limited impact, but this may be obtained by a serious deterioration of other model results. To our knowledge, this has not been shown/described before .

The second part of the paper performed a similar analysis, but with real ET data as model input. Here, we prove that good discharge simulations can only be obtained if the volumes of actual ET are accurate. We did recalibrate the model with these actual inputs, but the results and conclusions were exactly the same (results not shown in the paper): because of the underestimation of the actual ET we did not obtain acceptable discharge simulations.

3. *I did not understand the focus on flood events, while RR models sensitivity to PE likely emphasized on low flows and water balance. It appears quite obvious that a modification on PE will not affect largely flood peaks, unless the model is calibrated on this modified PE inputs, which is not done in the paper.*

Rainfall-runoff models in the low countries are mainly used to predict high flows, because that is the problem that exists here. This is why we have focused on the impact of ET inputs on peak flows.

4. *The first part of the paper on the sensitivity of RR models to PE is interesting but not really novel. An interesting and novel add-in could be to analyse simulated AE and measured AE in terms of annual volume (bias) and dynamics (correlation, variance ratio) for all possible PE inputs. This is done at the end of the paper but for only one PE configuration..*

We did compare the simulated AE for different PET inputs to the (measured) scintillometer estimates of AE (statistics of ET_{act} in Table 2 where RMSE and total volumes are compared).

5. *In the introduction, there is no reference on previous studies using actual AE measurements for RR model simulations. Is this was done before?.*

We are positive that the use of long-term AE estimates from scintillometer data as model forcing for RR models has not yet been studied. We have thoroughly studied all literature data bases and did not find any references.

6. What about the spatial significance of AE measured by scintillometer at the catchment scale?.

For this question we need to refer to our previous paper in HESS, to which we also refer in this manuscript. In that paper, we compare remotely sensed ET data to the measurements from the scintillometer and a spatially distributed hydrologic model. During the Summer months good results were obtained. However, during the Winter months no validation was undertaken, since we did not have remote sensing data for that period. However, the results in this paper indicate that during the Winter a significant mass balance error is obtained with the scintillometer data, which is reflected in the results of the rainfall-runoff modeling.

2. Reply to the issues raised by Anonymous Referee #2

2.1. General comments

The objective of this paper is “to thoroughly examine to what extent the results of a rainfall-runoff model can be improved by forcing them with actual evapotranspiration data, obtained using a large aperture scintillometer, instead of using potential rates”. The title of the paper also eludes to this. However, the scintillometer derived ET is not dealt with until page 20 out of a total of 24 pages in the manuscript. A large part of the analysis is dedicated to other analyses, i.e. the performance of the RR-model with standard potential ET parameterisation (a modelled yearly cycle) and Penman and Penman-Monteith formulations at various timescales. The authors should either change the title and general research objective to better cover the actual analysis done or leave out all other analyses that do not include the scintillometer data. In addition, the structure of the paper needs to be revised to better follow either of the two approaches.

We agree that only about five pages in the paper are directly dealing with the scintillometer results, but we need the analysis before this in order to be able to draw our conclusions. In other words, the parts of the paper before the analysis of the scintillometer data are taking steps that allow us to thoroughly analyse the scintillometer data. For this reason, we would like to keep the outline and organization of the paper the way it is.

However, we changed the title of the paper as to represent more the content of the paper. Also, we slightly modified the abstract and introduction as to make the different objectives of the paper more clear..

We propose as new title: ‘Impact of potential and (scintillometer-based) actual evapotranspiration estimates on the performance of a lumped rainfall-runoff model.’

My main problem with this paper is that the authors are not convincing me that their main result is not a trivial one. They have a model that has been optimised based on ~10 variables and a climatological yearly ETp cycle. If I understand it correctly, out of this modelled yearly cycle a daily average ETp is taken into the optimisation, ignoring the daily cycle. Next, they use the model with ET inputs with different timescales and corresponding different amplitudes and they get a different result.

Our results are more complicated than this.

→ If the same method for the calculated PET is used, the temporal

- resolution of these PET data does not strongly impact the modeled discharge. The papers from Oudin come to the same conclusion.
- However, for different PET calculation methods, Oudin et al. [2005b] always performed a volume correction. For this reason the source of the PET data did not matter. However, in our paper, we show that different PET calculation methods lead to less accurate discharge simulations. This has not been shown before. Modifying the temporal resolution within each PET estimation method did not alter the results.
 - What we have also shown is that the internal model dynamics can very strongly change (soil water storage, evapotranspiration) with changing PET or AET inputs, even when the modeled discharge hardly changes at all.

To continue they use the LAS based actual ET, which has both a different time-scale and it bypasses the positional to actual ET step in the model, which was part of the original optimisation. Isn't it to be expected that that the model won't perform well using input parameters for which the model is not optimised?

We refer to the comments of the first reviewer for an answer to this question.

The paper is well written. However, the structure of the paper needs to be revised. Also, the paper is very full, there are too many messages. The analyses are generally sound, but the discussion of the results could be better.

All things considered I recommend publication after major revisions

We refer to our answers on the first general comment and the specific comments.

2.2. Specific comments

1. *Timescales. There are many time-scales used in the paper (hour-day-month-year) and it is not always clear which one is used in the current analysis. It becomes especially confusing when the ET rates are given in different units AND averaged over different times throughout the paper. ET rates are expressed as mm/hour, mm/day or mm/month and averaged over either a hour, day, or year. I understand that the RR-model is run with a time step of one hour? It would be helpful to have a short section at the beginning of the paper that explains timescale convections used for rates and averaging periods.*

We indeed did average the inputs over a number of time steps. However, the model always uses hourly inputs. So we always had to convert the PET or AET inputs into mm/h. We have clarified this in the section where we explain the averaging procedures (§4.1).

2. *Structure-Organization. The paper discusses the following ET input approaches:*

1. *standard ETp input (climatological yearly cycle, the same one used for the optimisation) taken as a daily average*
2. *P and PM ETp input taken as hourly, daily, monthly and yearly averages*
3. *LAS based ETa input taken as hourly average*

It would be very helpful if this lay-out was communicated to the reader at an early stage and organize the sections in such a way that this structure is recognizable throughout the manuscript.

Seeing that the analysis focuses both on varying the timescales and the type of ET model/measurements. Why not do all three ET approaches with all four of the timescales? In

that way you can distinguish between the effect of varying the timescales from the effect of varying the ET input. Now, the two are mixed. For instance, compare first the modelled discharges using hourly vs daily averaged ETp both from standard ETp input based on the climatological yearly cycle before including P, PM and LAS ET's in the analysis. Constructing the hourly averaged standard ETp input data may require some work, as one has to superimpose a daily cycle on the climatological yearly cycle taking into account the varying day-length over the year. Ignoring the daily cycle, as is done now in the model, means that ET rates are the same during the day and the night, which is not realistic. Once the climatological, truly hourly averaged yearly cycle has been constructed, it can also be used to optimise the model, which will hopefully yield better model parameters when working with the hourly averaged ETp from P and PM or the LAS ETa. The same could be done for the monthly or yearly average.

We have further clarified the paper structure at the end of the introduction. What we have added is that we not only want to confirm the results of Oudin et al. [2005a], but also assess to what extent their conclusions depend on the source of the PET data. Furthermore, we want to examine the impact of the PET source and temporal resolution on the internal model dynamics.

At this point, we want to avoid adding a second sine wave to the sine-wave annual cycle in PET that we have now. This annual cycle is already quite artificial. We do not think that adding a second artificial cycle to this, and comparing to real ETP-estimates, will give extra value to the paper.

3. *The discussing of the results displayed in the very crowded tables with regression statistics is poor. For instance on line 24 on page 19: "Comparing the model performance statistics of this approach in Table 3 .., it is clear that PDM performs worse ...". There are hundreds of number given in these Tables; where do you want me to look? Which statistic in particular show the poor performance?*

Indeed, there are a lot of statistics and they are all explained in paragraph 3.3. In Table 2 and Table 3, they are grouped according to different types of statistics: statistics about the performance of the PDM for total flow, high or peak flows, low flows and actual evapotranspiration. In the text it is always indicated which of these groups are considered for the discussion of the results.

The sentence on p. 19 has been rephrased in order to make it more clear.

4. *In addition to the item above, the statistics Tables 2 and 3 and Figures 3, 4, 6 7 are very busy and some of them are hardly discussed in the text. For instance the peak and low discharge analyses don't seem to add much.*

We refer to the previous comment.

We'd like to keep all statistics, because they illustrate that not all considered output (total flow, low flow, peak flow, internal ETact) improves when using other ET as model forcing: sometimes the overall modeled flow improves, while the performance for the low flows decreases etc.

5. *LAS based ETact. The LAS doesn't measure ETact, it measure the sensible heat flux, which is used to estimate ETact using the energy balance approach as is done in TOPOPLAST. It is well known issue in micro-meteorology that the energy balance doesn't close (eg Foken, 2008). With the approach followed the energy balance non-closure is*

accumulated in the LAS ETact estimate. As a result, the daily averaged ET rate in wintertime is even negative (8mm in total in November!). This is not realistic. It is better to keep ETact as zero under these circumstances.

Also, H will be very small during winter time, typically less 50W/m². It is questionable therefore, whether it can be considered a LAS based ET estimate as the other terms of the energy balance (i.e. the TOPOLAST algorithm) will be dominant in the ETact estimate.!

The energy balance for eddy covariance systems indeed does not close. We want to emphasize that TOPLATS closes the energy balance, the model (and any land surface model in general) has explicitly been designed for this.

It is not so much the energy balance non-closure that leads to the negative ET estimates during the winter, but uncertainties in each term, in addition to uncertainties in the scintillometer-estimates. However, this is still a better approach than assuming zero evapotranspiration during the winter. There are a number of papers that indicate that fluxes during the winter, even though they may be small in absolute terms, add up to significant values over time. Therefore, we would prefer to avoid using this approach. It is true that the other terms in the energy balance are dominated by TOPLATS, but these are the terms (most importantly the net radiation) which show the best agreement with in-situ measurements.

Furthermore, using the long-term estimates of AE (obtained by the energy balance approach from H-measurements from a scintillometer) in the catchment water balance and RR model, shows that there are indeed a number of uncertainties and difficulties in this energy balance approach that is always used to convert scintillometer based H measurements into evapotranspiration estimates. This has not been shown before and indicates that more research is required to obtain long-term (year-round) estimates of actual evapotranspiration based on scintillometer-H-measurements.

6. *LAS saturation. The authors claim that the LAS will not saturate over 9.5km with a path height of 15m. This is not true. Saturation is to be expected. The BLS2000 software corrects for this effect. This correction is adequate as long as the saturation level is weak. How was this correction for this experiment? (corrected and uncorrected Cn2 can be found in the BLS2000 software output).*

Indeed, saturation occurs and a saturation correction is applied when converting the rough measurements into Cn2 and further into H-estimates. This is fully explained in Samain et al (2011).

→ The sentence has been adjusted.

7. *Comparing Figs 6a and 7a: In 2007 (Fig6a) $Q_{obs} > Q_{sim_using_Ep}$ for the entire year, whereas in 2010 $Q_{obs} < Q_{sim_using_Ep}$. How is this possible?*

This can be explained by the source of the meteorological forcings, more importantly the rainfall. We are using point measurements as representation of catchment-averaged rainfall. The test site we are working on is relatively small (approximately 100 square kilometers). Despite this, there will always be discrepancies between the catchment-averaged rainfall and the point measurement, with sometimes overestimations and sometimes underestimations. This will be reflected in the discharge simulations.

2.3. Minor comments

8. Tables 2, 3 and 4: What are NS and CB? In general, the statistical variables in the tables are not explained in the caption nor in the text.

In paragraph 3.3, all statistical variables are explained. More specifically, NS (Nash Sutcliffe criterion) and CB (Cumulative Balance Error) are clarified by their equations (equations 25 and 26 respectively).

9. Tables 2, 3 and 4: Check the number of significant digits of the values given. The cumulative Q for instance is given as -2346.307mm. I doubt that the number behind decimal point are still significant.

We agree with this comment. The significant digits of the cumulative values have been adjusted in all these tables.

10. Figures:

*Figures 3-10: Increase the font-size of the tick-marks and axes labels

This has been adjusted.

*Fig6: part of the y-label is missing

This has been adjusted

*Timeseries plot of Figs 3 and 7 only have 1 or 2 tick-marks on time (x) axis; increase this number

This has been adjusted

*Timeseries plot of Fig 8 have 35 (!!!) tick-marks on time (x) axis; decrease this number

This has been adjusted

*Fig 4: what is the meaning of a negative discharge (y-axis)?

By applying the Box-Cox transformation (paragraph 3.3), the resulting BC-transformed flows can become negative

*Fig 10: too much information in one figure. Present your results in another way (eg Winter-Spring-Summer-Autumn

The results grouped per season are not as clear as per month. We prefer to present these results the way it was.

In order to make the figure less 'heavy', we deleted some statistics for all the scatter plots.

11. Page 3 line 17: "However, Oudin ..."; delete "However".

Adjusted as suggested.

12. Page 4: line 2: "actual into potential ET" should be "potential into actual ET"..

Adjusted as suggested.

13. Page 4 line 10-11: change "The impacts of ... are examined." to "The impacts of... are examined as well."

As stated earlier, this section of the introductions has been rewritten.

14. Page 7 line 17: A reference is made to Flanders, but in section 2 where the catchment is described it is not mentioned that the catchment is situated in Flanders.

| This has been added in section 2.