

Interactive comment on “Runoff regime estimation at high-elevation sites: a parsimonious water balance approach” by E. Bartolini et al.

E. Bartolini et al.

elisa.bartolini@polito.it

Received and published: 15 April 2011

We would like to thank anonymous referee no.2 for the comprehensive review. We take the opportunity of this "interactive discussion" to reply to the general comments by this referee, clarifying some methodological aspects; in the second part of this reply we will address the specific comments outlined in the review. Some specific comments related to grammar issues have been directly amended in the revised version of the paper, while others, concerning methodological issues, are discussed in the following. Our replies are reported in italic, interspaced by the original comments by the reviewer.

General comments.

- 1 In both the abstract and the introduction, the original contribution should be men-
C921

tioned and emphasized. In doing this, the objective should be immediately placed in context of the ultimate need of that this research is designed to meet. Somewhere in the manuscript, the new water balance model should be compared and contrasted to the mentioned water balance models and a statement as to its originality should be made.

We agree. In order to highlight the objective of this study, the first sentence of the abstract will be changed to: "We develop a water balance model, parsimonious both in terms of parameterization and of required input data, to characterize the runoff regime of high-elevation and scarcely monitored basins." Several changes will be made also in the Introduction, with the aim of highlighting the originality of the methodological approach when compared with other water balance models conceived for high-elevation sites. For example, we will add the sentence: "The modeling framework adopted in this study differs from those of Zappa et al., 2003, Eder et al., 2005 and Schaeffli et al., 2005 among others, because it operates at the monthly time scale and it is conceived to be applied in scarcely monitored basins."

- 2 The authors make an argument in the discussion section that the second parameter is justified based on the improved model performance by including that parameter. However, no justification is made for not including some processes and parameters that may be very important. The most obvious example is the decision on whether or not to include soil moisture storage and associated parameter(s). It is possible that an additional parameter related to soil moisture storage and release may be as important as or more important than the second snowmelt-related parameter that was included. This should be tested. Not including soil moisture storage dynamics is a major weakness in this model that deserves more attention.

The problem of the consideration of the soil water storage has been raised not only here but also in several specific comments reported in the second part of this reply.

We will try to address all the comments here in a comprehensive way.

The model partially accounts for the effects of soil storage through the coefficient SR of storm runoff. In fact, this variable represents the fraction of precipitation that directly contributes to surface runoff, while the remaining precipitation is theoretically available for infiltration and constitutes the storage for evapotranspiration. As a consequence, a preliminary assessment of the role of the soil moisture storage is in Fig. 8b (P990) where the results with SR=0.3 are compared to those of a model with SR=0. Since the MAEs associated to the model with SR=0.3 are generally smaller than the ones with SR=0, we concluded that the assumptions made were adequate.

In the model, SR was not considered as a parameter to be calibrated, but as a fixed value, valid for the whole region. This value has been obtained as a result of several considerations. We initially considered SR as a third parameter to be calibrated, and varied it in the interval 0.2-0.6 with steps of 0.05. This formulation of the water balance model, characterized by the presence of three parameters (σ , c and SR), has been applied both at the local scale and at the regional scale. Analyzing the variability of the MAE on the parameter values, it was found that the MAE remains almost constant for changes in SR, both for the regional application of the water balance (see Fig. 1b of this reply) and for the local-scale application (see Fig. 1a of this reply), in the large majority of catchments. For this reason, we decided to consider the storm runoff as a fixed value, also considering that the regional parameter calibration returns a value of SR for the entire region equal to 0.3.

We adopted this as a balance solution between the need to account, at least in a very minimalistic way, for the soil storage effect, and the importance of maintaining a small number of parameters, which is an essential requisite for a model to be applied in scarcely gauged regions.

Note that no delay time for the soil moisture storage is considered because the introduction of a retention time (aimed at modelling the groundwater storage) would require a characterization of the soil properties of the basins, which is far beyond the scope of this paper. This is the reason why we stated, probably in a confusing way, that no soil

C923

moisture storage is considered in our model. Both in the Model description and in the Discussion sections we have now clarified this point.

- 3 Somewhere in the manuscript, possibly in discussion (and touched on in conclusions and abstract), a description of the types of basins that are appropriate for this water balance model should be included. This should include, basin size, type of climate, topography, etc.. Also, in determining regional-wide calibration parameters, how large of a region can this be approach be used for?

We agree. We now mention the basin characteristics in the discussion and in the conclusions of the paper respectively:

"The structure of the balance equation (Eq. (1)) implies that the runoff reaches the outlet within one month. As a consequence, the water balance is suitable to be applied in small and medium-size basins. Moreover, given that the model is conceived for catchments characterized by the presence of the snow storage, it is especially suitable to be used at high-elevation sites."

"The model is suitable to be applied in small and medium size mountain basins and proves to be able to discern between different mechanisms of runoff formation in the study domain, which is a transition region characterized by headwater basins governed by snow dynamics and middle-elevation catchments characterized by temperate regimes."

With regard to the proper size of the region when the regional application of the water balance model is pursued, we cannot univocally answer to this question. In fact, the proper size will depend on the climatological and physiographic characteristics of the area, on the number and location of basins with observations available for calibration, etc. . .

- 4.1 In using observed streamflow to both calibrate and evaluate the model, was the streamflow observation partitioned in two for each of these? In other words, is

C924

there an independent (of calibration) assessment of model performance?

Streamflow observations were not partitioned into a calibration and an evaluation period. This is because the model uses monthly average precipitation as an input to simulate monthly average runoff. In other words, the model is not aimed at simulating in continuous time the monthly time series of runoff. As a consequence, the traditional approach of model calibration and evaluation on different time periods cannot be followed.

4.2 I strongly suggest that the authors evaluate the snow dynamics module independently using snow course measurements or any snow depth or SWE measurements that might exist over their study basins. Because the emphasis is on snowpack, this type of evaluation would be an important component of this paper.

Unfortunately, the evaluation of the performances of the snow dynamics module using SWE or snow depth measurements is not an option for this study region. In fact, measurements of SWE are not available and only a limited number of snow depth observations, without information about snow density, are present. In order to obtain a reasonable estimate of the basin SWE – in order to compare it with the SWE simulated by the model - it would be necessary to spatially interpolate the observations, assuming reasonable values of snow density. This would introduce a high degree of uncertainty, also due to the complex orography of the study domain. For this reason we decided to assess the model performances by comparing observed and simulated runoff, which represents the overall response of the basin and, as such, integrates all the physical processes involved in the water balance.

Specific comments

P960L7: what is a nuisance parameter?

C925

In statistics, a nuisance parameter is any parameter which is not of immediate interest but which must be accounted for in the analysis of those parameters which are of interest. We intended to stress the disturbance effect of the presence of many parameters to calibrate when only few observations are available, but we agree that the sentence is confusing, so we will erase the word "nuisance".

P961 – Model description: Mention assumptions here for snowmelt model: no sublimation, no evap of liquid water content in snow,...

Done

Mention upfront that soil moisture storage is neglected.

See Answer no. 2

P 961L17: the authors state that SR is "unavailable for recharging soil moisture". Based on this, it would suggest that soil moisture storage is included in the model, so this is misleading.

See Answer no. 2

Mention right away that the reason for separating out SR from total liquid P is so that actual ET only occurs from the 0.7P+ portion.

Done

What is the justification for SR=30

See Answer no. 2

P962L5-8: the assumption is, therefore, that all runoff exits the basin within one month. This sets limitations as to the size of the basin being simulated. Earlier, the authors suggest that this approach should be applicable to a large range of basin sizes. Assumptions such as this one limit the size of the basin appropriate to be simulated by this model.

C926

This is right. In fact, the entire study domain is characterized by small and medium size catchments for which the mentioned assumption is reasonable. The model cannot be applied for large river basins. As previously reported, we will clarify this point in the Discussion section of the revised manuscript.

P962L20: is there an estimate of goodness of fit?

No, there is no goodness-of-fit test here, because the availability of daily temperature time series for the study region is limited, and this is the reason why we developed a model working at the monthly time scale. In any case, we believe the specific form of this probability distribution does not significantly impact the results.

P962L26: Is this assumption tested? Also, there is a typo at the end of this line.

The validity of this assumption can be seen analyzing the values of σ reported in Table 2 in the Reply to anonymous reviewer 1. In fact, the mean monthly σ_j obtained by averaging the station values, are:

| | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
|------------|------|------|------|------|------|------|------|------|------|------|------|------|
| σ_j | 3.62 | 3.42 | 3.07 | 2.91 | 2.97 | 2.83 | 2.65 | 2.51 | 2.71 | 3.35 | 3.00 | 3.41 |

It appears that the standard deviation of daily mean temperature has a slight increase during the cold months and a slight decrease during warm months, but these changes are very small and σ remains in the range of 3° C. For this reason, it is reasonable to assume σ not to significantly vary within the year.

P963L18: do the units work out here? mm = mm – mm/month + month*mm/month?

Yes, they do because the SWE, the actual snowmelt and the snowfall are expressed in mm. In particular:

$$M_{pot,j} = c \cdot (T_j^+)^2 \cdot PT_j, \quad (1)$$

C927

so that

$$[mm] = \left[\frac{mm}{^{\circ}C^2 \cdot month} \right] \cdot [^{\circ}C^2] \cdot [month]. \quad (2)$$

P965L6: “heat indices are supposed to take on values greater than 5” – is there a reference for this fact?

No, there is no reference for this. As mentioned in the paper, we found that the application of the Thornthwaite equation for high-elevation cells leads to an overestimation of potential evapotranspiration due to the values assumed by the heat indices. To demonstrate this effect, the average monthly temperature regime of the study domain at the sea level has been translated, by applying a fixed lapse rate equal to -5.5° C/km, at different elevation, namely 1000, 2000, 3000 and 4000 m asl (Fig. 2a). For each case, the indices I, a and the monthly potential evapotranspiration are computed. Figure 2b shows that, during summer months, the potential evapotranspiration simulated at 4000 m asl is equal or even larger than the one simulated at the sea level, even if the monthly temperature is significantly lower. This problem can be solved imposing that evapotranspiration continues to decrease also at higher elevations, as actually happens in real conditions. This condition is met by imposing that the heat index can take only values greater than 5.

P965L7-8: “the model proves to be able to predict realistic values of ET. . .” - where are these results and what metrics are used to evaluate this?

The values of evapotranspiration that are used to evaluate the capability of the water balance model in assessing the potential evapotranspiration are reported in the study of Henning and Henning, 1981, that is cited in the manuscript at P966 L 9. This study reports the mean annual amounts of potential evapotranspiration estimated using the Penman modified approach. These data, given as reference values for elevation bands, are used to qualitatively evaluate the model results.

P965L15: There should be a 0.7 in front of P_{j+} ;

Right, thanks for notice this mistake, we will correct it in the revised version of the manuscript.

Also there is no consideration in Thornthwaite ET for aerodynamic and vegetative resistances, correct? Also, there is no consideration for the non availability of deeply infiltrated water to ET.

Correct, none of these phenomena have been accounted for.

P967L2: The authors must mean that RI is V_i/V_w , or the ratio of volume of artificial lake divided by annual water volume at basin outlet. A value of 25% then would indicate that artificial lake storage is only 25% of annual water volume at basin outlet, which is reasonable. Is this correct?

Yes, this is correct. We will correct the mistake, thank you for pointing it out.

P967L26: The assumption here is that 100% of the model bias in R is due to underestimation of P. State this assumption here and discuss whether or not this is a reasonable assumption. Is there any reason to believe that ET may be biased high or low? Would not including soil moisture storage cause a systematic bias? This should be fully discussed and evaluated, if possible.

In the revised version of the paper, we will add, as suggested, that the precipitation correction is developed under the assumption that the model bias depends only on the precipitation undercatch. This is reasonable for several reasons: i) bias between the average annual precipitation and runoff are detected not only in the simulated results but also in the observations; ii) the simulated evapotranspiration is comparable with the potential evapotranspiration reported in the literature and computed with a different method; iii) even if no soil moisture storage is considered, at the annual scale the inflow should equal the outflow.

Another assumption is that the bias is uniform among the months. We know that un-

C929

dercatch is largest when solid precipitation occurs during the winter. Can monthly T and P be used to determine a better division of bias among the months? This would not introduce any further parameters, but it should make the model more accurate. P977L3: "improved" instead of "ameliorated"? How do the authors suggest the generalized precipitation correction be improved?

We grouped these two comments because they refer to the same topic. Correction of precipitation could possibly be improved taking into account the seasonality of precipitation (i.e., considering a monthly precipitation correction factor that is proportional to the monthly measured precipitation and not constant), the higher undercatch due to solid precipitation (i.e., considering a correction factor that depends on temperature) or even a combination of these effects. We rapidly tested some of these strategies to distribute precipitation correction in different month but, at this stage of the study, we have not found conclusive evidence that other partition strategies can improve the performances of the model. We therefore decided to use the simplest possible strategy, i.e. to distribute the correction homogeneously in the 12 months.

P969L16: Is there a goodness of fit estimate?

Yes. The significance of the regression has been tested with the T student test on the slope of the regression line with a significance level α equal to 0.05. Since no mention to the T student test is present in the paper, we will add it.

P970L1: MAE is calculated using mean monthly R. This would produce a lower MAE then if it were calculated for all months for all years. By averaging monthly R for all years, the error is underestimated.

MAE is calculated using mean monthly runoff values because the model uses the average values of precipitation to simulate the average values of runoff.

P970L27: Is this altitude determined by looking at where glaciers occur in this region? How is it determined?

C930

The altitude of 3000 m asl, used as a constraint during the parameter calibration as the minimum elevation above which the snow can persist also during summer, has been chosen trying to find a compromise between a good reproduction of the observed runoff regime and the closure of the water balance. It is also the elevation where the front of most glaciers present in the study domain is located.

P971L6-9: If the parameter values with the lowest MAE are in an error of Figure 4 where the snow storage depletion condition is not met, is this an indication that there is something wrong with the model physics? What is going on here?

In our opinion it is not a problem in the model physics but rather a consequence of the precipitation correction and of the minimalistic framework of the model. In fact, increasing the precipitation causes an increase in the size of the snow storage that needs to be melt during the warm months. Moreover, the parameter calibration procedure based on the minimization of the error statistic, the MAE, does not require the closure of the water balance, that it is actually imposed by the constraint on the snow line elevation during summer.

P971L25 and P972L1: How are these confidence bands determined? How were these equations derived?

For the determination of the confidence bands we assumed that the monthly observed runoff time series follow a normal distribution. For each month, the runoff time series is used to compute the monthly runoff standard deviation, $\sigma_{R,j}$ while the monthly average constitutes the mean of the distribution, R_j . Then, using the quantile function of the normal distribution, one obtains that the 40% (80%) of the observations in the sample falls in the interval $R_j \pm 0.53\sigma_{R,j}$ ($R_j \pm 1.28\sigma_{R,j}$).

P972L6: This underestimation on summer could also be because the snowpack was underestimated, possibly because of uniformly applying the P bias among all months.

We agree, this could be a possibility and we thank the reviewer for this suggestion,

C931

which can eventually be tested developing different precipitation correction redistribution. However, our first analyses in this direction have not shown clear results.

P972L25-27: this nuisance is lost to me – why is it interesting to note this?

We wanted to point out that the melt factor obtained from the regional parameter calibration is different from that used in the reference run. We agree that it is not so interesting to be remarked, so we will change the sentences into: "In this case, the melting rate is smaller than. . ."

P972L29: Why can't QI be used to compare different model applications? This is stated both here and in Appendix, but this is not clear to me in either place. P978L27-29: The meaning of this sentence is not clear to me. Please clarify.

We aggregate these two comments because both of them are related to the Quality Index QI and its field of application. The QI index is obtained scoring some indicators, namely the monthly runoff standard deviation, the month of occurrence of the runoff peak, the bias and the MAE. In particular, to score the bias and the MAE, we used their respective quantiles, which hinders the use of the QI for comparing different models. For example, let us consider different model applications, namely M1 at the local scale and M2 at the regional scale. 39 MAE values, one for each basin, are associated to M1 and other 39 MAE values are found with M2. Let us assume that the MAE values of M1 fall in the range 10 to 20 mm and the ones of M2 fall in the range 20 to 40 mm. A basin with MAE=20 mm will be scored with a 0 in M1 (because it is the largest MAE value) and with a 1 in M2 (because it is the lowest MAE value). The comparison of the performances of different models based on QI would therefore be meaningless. In the paper, in order to clarify this aspect, we will stress the dependence of QI on sample quantiles as follow: "On the contrary, due to the mechanism of score assignment based on sample quantiles, that change depending on the model simulations, the QI cannot be used to compare the results of two different modeling frameworks (i.e., local versus regional scale)".

C932

P976L14: However, I would argue that not all the governing mechanisms are incorporated. Lack of soil moisture dynamics is the main missing mechanism. Therefore, it is not possible to identify the role of soil moisture dynamics without including it in one of the sensitivity test simulations.

See Answer no. 2

Supplementary Material: Include RI in this table.

Done.

Figures and Captions:

It would be good to show monthly plots of precipitation and temperature, either on average over the case study region, or by basin, as in figures 5 and 6. This would help readers who are not intimately familiar with these basins in interpreting the results.

The possible introduction in the paper of an additional figure showing the seasonality of the other variables involved in the water cycle has been raised also by the anonymous referee 1. However, due to the length of the paper, we believe that this would not be the right choice. At the same time, we are aware that some additional information on the climatology of the basins inside the study region could be really helpful for readers that are not familiar with this area. For this reason, we introduced the figure suggested by the referee here in the public discussion and we will recall it in the revised manuscript. In this way, it will always be publicly available.

Figure 1: Make the "+" and "-" larger in this figure. Define SR and R in the caption

Done.

Figure 2: Define what is meant by "budget" in this caption.

Done. We will change the caption into: "Study domain and catchments used for the model application. Hatched (cross-hatched) basins are characterized by a negative (positive) budget (i.e., the difference between annual precipitation and runoff). Gray-

C933

shaded areas indicate basins with positive budget that becomes negative when accounting for evapotranspiration losses (i.e., $P - R - ET_{act}$)."

Figures 5 and 6: Provide a qualitative definition of QI in the captions. Also, try to emphasize that figure 5 is for local and figure 6 is for regional in the captions, to make it really obvious how they are different.

Done. In order to emphasize that figure 5 is relative to the model application at the local scale, while figure 6 is relative to the model application at the regional scale, we will state it at the beginning of the caption. Moreover, we will add: "In the upper left corner of each panel a measure of the model performances, the Quality Index QI, is reported."

Figure 8 caption: Avoid the word "validation" as this is not a validation of the assumptions (that word is much too strong) but just a sensitivity study.

We agree. We will change the word "validation" with "evaluation".

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 8, 957, 2011.

C934

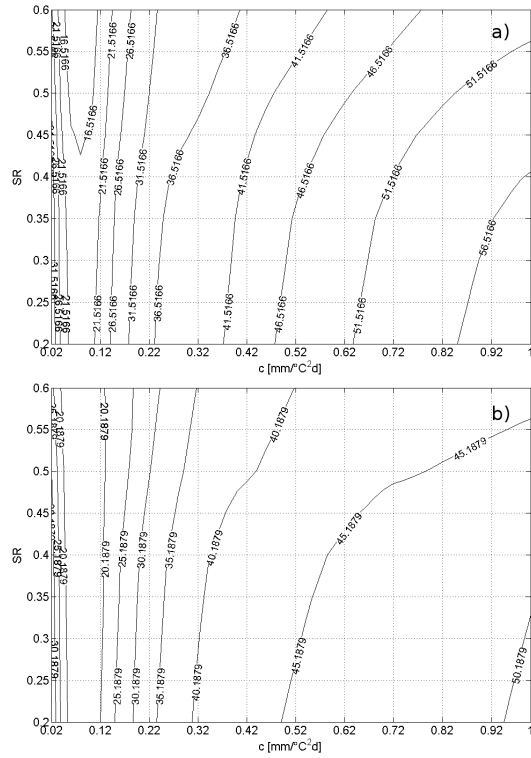


Fig. 1. MAE variability depending on the values assumed by SR and the melt coefficient c. a) local scale application, sigma=4°C; b) regional scale application, sigma=3°C (River Evancou at Champoluc).

C935

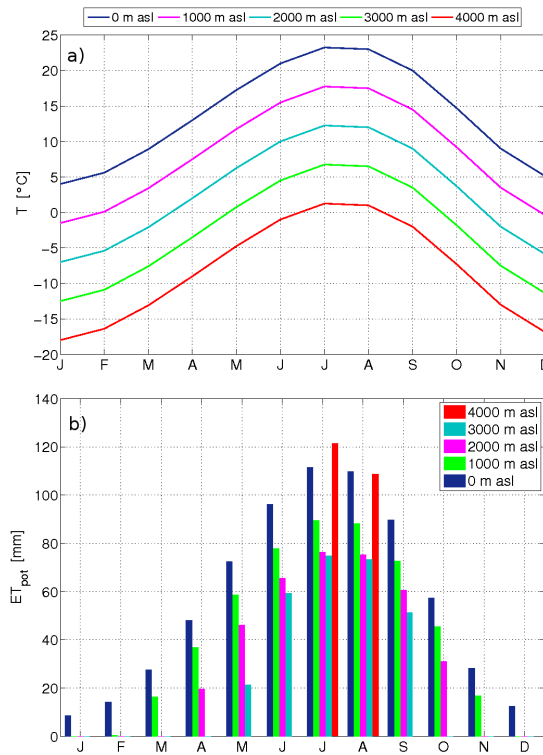


Fig. 2. Effects of the elevation-dependent variability of the average temperature regime on the Thornthwaite potential evapotranspiration

C936

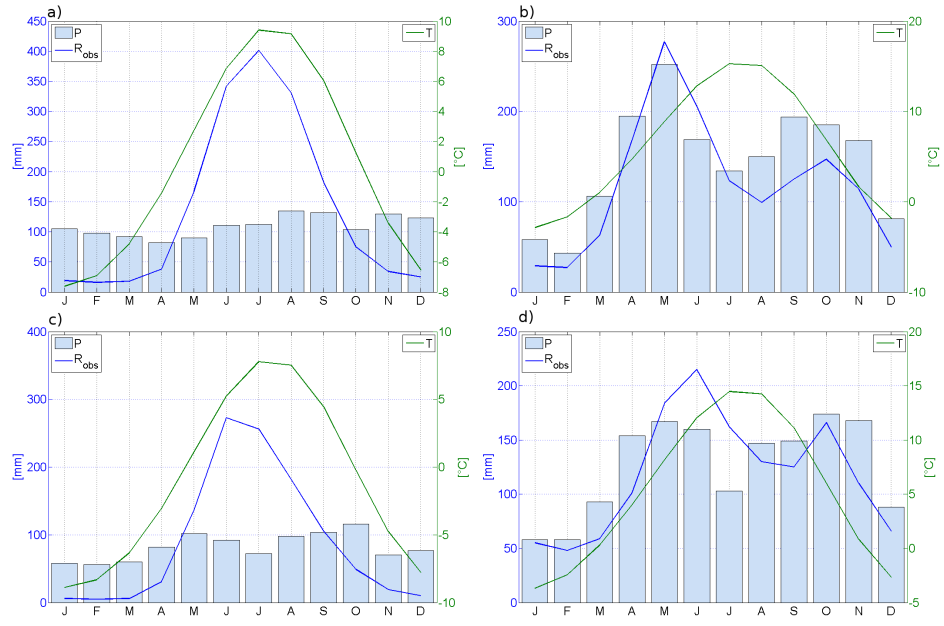


Fig. 3. Mean basin precipitation, P, temperature, T, and runoff, R_obs, observed regimes: a) Savara at Eau Rousse; b) Sesia at Ponte Aranco; c) Rutor at Promise; d) Toce at Candoglia.