Interactive comment on "Multi-level spatiotemporal validation of snow/ice mass balance and runoff modeling in glacierized catchments" *by* Florian Hanzer et al.

B. Schaefli (Referee)

bettina.schaefli@unil.ch

This paper proposes a detailed validation for a physics-based model for highly glacierized catchments. It is certainly the most comprehensive validation study that I have come across for this kind of models / environments and represents a sort of review of all type of data that can be used for model validation. The paper is very well written and organized. It was a pleasure to read. It was also particularly nice to see that the paper discusses in an entirely honest way the weak points of the model. I recommend publication with moderate revisions without re-review.

The idea of visualizing the data set observation scales is very nice. Food for thought: the MODIS / Landsat data seems very complete in this representation even if they rep-

resent only a partial aspect of a state variable (absence presence). There is a dimension missing here of the observation scale (match between simulated and observed flux / state variable content). How could this be integrated?

Detailed comments:

- The runoff module has 12 parameters, which is a very high number; it is in particular difficult to justify to have so many different time scales for the fast component of the flow (not going into the soil). The discharge will never contain enough information to identify them. It is also rather unusual to have a constant fraction of water going into the soil and a single time scale for the soil. This should certainly by re-thought for future applications of this model.

- Section 3.5: How is the calibration done (algorithm? Manual? How many model runs?) Is this limiting or are you confident to have found a good solution? Could you have presented model ranges for the mass balance / discharge time series rather than the result of a single parameter set?

- Areal precipitation: the variation of the soil water storage cannot simply be neglected at the monthly scale; what about ground water? A short comment would be useful.

- How does the uncertainty on initial ice thickness distribution impact the results?

- Section 5.3.1: This section discusses in detail the differences between MODIS and Landsat (which is very useful) but lacks a concise discussion of / conclusion on what these data sets say (respectively can say) about how good the model simulates snow accumulation. As far as I see the model does a good job in winter, when the data does not contain a lot of information (there is snow everywhere) but a bad job in autumn / spring.

- It could be interesting to discuss the extreme year 2003 for the mass balance simulations; why is the model particularly wrong here?

- P. 17, line 6: it is stated here that there is no calibration on the glacier mass balance;

this statement should be re-thought in light of the fact that there is a snowfall correction factor calibrated on all available glacio-hydrologic data; I would also not include areal-precip in the set of validation data since it is actually used to estimate a key parameter, not for validation (reformulate abstract and conclusion)

- Does the implementation of the cold content / liquid water module for the snowpack improve discharge simulations? If yes: I would show it (perhaps in the supplementary material), otherwise: why did you choose to implement it?

- Section 5.5: your results suggest that the precipitation amounts are not compatible for glacier mass balance and discharge (discharge bias increase for corrected precipitation). This is not unusal but could you comment on it? Could the bias be explained by errors in other water balance terms (evapotranspiration, groundwater exchange)?

Even more detailed comments (no need to answer in the public discussion)

-Math notations: some variable names are used for different things (Q, E, W). k is used as parameter for very different things

-p. 2, line 4: conceptual models are not empirical models

-p. 2, line 9: conceptual model are run in a stationary mode because they cannot be run in a transient mode; so it is not the calibration that invokes the stationarity assumption; I suggest reformulating

-p. 2: line 10: I would be more careful with the statement of deriving parameters from field measurement, it omits completely the problem of scale

-p. 3: I would not say that areal precipitation is a data set, reformulate (say which data are available; areal precip is based on this)

p. 4, line 23: how are the lapse rates identified?

p. 7, line 21: "manual optimization for the best fit with model results"; which model results? Does this mean that there is a calibration going on here? Where do the

factors of eq. 8/9 come from? A bit unclear.

p. 8: line 5, I would specify which set is variable in time, fixed in time and space, variable in space

p. 8, 3.4: if not described elsewhere, should this module be described in more detail in the supplementary material

p. 8, line 24: should it read "the water leaving the snowpack" rather than precipitation and melt since the model has a snowpack retention capacity

p. 10, line 1: to which value is evaporation set?

p. 16: line 16: would it be possible to see these results in the supplementary material?

Figure 9: titles (obs, simulated) missing; would have been interesting to see SWE distributions (simulated / observed) as a function of elevation (e.g. for some selected elevation bands)

Fig. 10: SWE differences between what?

Figure 14: I would represent also storage changes and evapotranspiration

---