Hydrol. Earth Syst. Sci. Discuss., 11, C1521–C1531, 2014
www.hydrol-earth-syst-sci-discuss.net/11/C1521/2014/

Hydrology and Earth System Sciences
Discussions
Open Access

HESSD
11, C1521–C1531, 2014

Interactive Comment

Interactive comment on "Modelling of snow processes in catchment hydrology by means of downscaled WRF meteorological data fields" *by* K. Förster et al.

K. Förster et al.

kristian.foerster@uibk.ac.at

Received and published: 26 May 2014

We wish to thank anonymous reviewer 1 for her/his detailed comments on the manuscript. This will help us to improve the manuscript.

General comments:

> *"Different snow models are driven by a set of parameters delivered by the WRF model. The investigation area is located in the Hartz Mountains (Germany). The general aim of the paper, to introduce a model chain which is independent of surface measurements, is an important research topic. The*

authors clearly describe the deficits of sometimes sparsely distributed point measurements with respect to spatially distributed models. Hence, they are using WRF fields for driving land surface models. This would be a favorable approach but because of deficits within the precipitation fields they are substituted by measurements in the course of the paper. This definitely limits the significance of the whole paper and stands in contradiction to title, abstract and introduction. Beside of this the calibration strategy remains unclear and is sometimes inexplicable. It is said that the snow models are calibrated by using WRF fields instead of measurements but precipitation measurements are used for driving the model afterwards. This is hard to understand. Finally, the availability of validation data is extremely limited and the data seems to be inappropriate for checking the quality of different key results of the snow models. If these deficits can be eliminated the paper could be worth for publication. "

Reply: We agree that the substitution of simulated precipitation by observations limits the application of the approach that has been presented. Our intention was to present that the meteorological fields, which were derived using downscaled WRF analysis data, are generally suitable datasets to drive snowmelt models. It was possible to derive reliable values of precipitation depth at the seasonal scale. However, at smaller time scales simulated precipitation intensities are less reliable. This finding holds especially for the period 25 Mar 2006 until 09 Apr 2006, during which a rain on snow event occurred (Fig. 4). Therefore, we decided to substitute simulated precipitation with observations. In this case, the simulations of runoff at both scales coincide well when compared to observations. Hence, it could be shown that the other variables, which are also relevant to snowmelt, are suitable to drive snowmelt models at both scales.

Alternatively, we will show that using simulated precipitation does not yield good performance, by showing results based on simulated precipitation. This approach proves that 1) the precipitation simulations are not as accurate as desired and 2) the other variables including temperature, humidity, wind speed, shortwave radiation, and longwave radiation are very useful in order to serve as input for snowmelt models.

Using observed precipitation could be used to improve the representation of precipitation. These findings restrict the applicability of this approach. However, in many cases all these variables are less widely available than precipitation. Hence, a substitute might be relevant in regions where only precipitation is available and some of the other variables are missing. This is a common situation. As a prospect for future research, a combined product of precipitation observed at stations, and simulated precipitation fields could also be useful to improve the spatial and temporal representation of precipitation at the catchment scale.

We regret that the explanations regarding to the calibration of the models have led to misunderstandings. As we will explain in the specific comments sections, we will clarify the text with respect to the calibration procedure. Unfortunately, snow water equivalent observations are not available at the Torfhaus meteorological station. We will provide a figure including both observed snow depth and snow water equivalent simulations in order to improve the validation of the models.

Specific comments:

> *"Line 9: Can you proof this assumption by data or by a citation. Does it matter at all?"*

If you refer to page 4067, line 9, we agree that it does not matter at all. A comparison between climatological means of temperature and precipitation among stations in the Harz and the Alps could underline this statement. However, as it is not particularly relevant for this study, we will delete this sentence.

> *"P4068 Line 1 pp.: What is the database for the mentioned values? How*

HESSD

11, C1521–C1531, 2014

many stations were available? Where are these stations located? What is the configuration of the stations?"

All values refer to monthly means as well as long-term recordings of Braunlage automatic weather station. We will briefly describe this and we will indicate the location of this station in the map (Fig. 1).

"P4068 Line10 pp.: What are the reasons for choosing these two winter seasons? Why have you chosen seasons which are significantly different from the average? Would it make sense to analyze a season which is close to the average as a benchmark?"

The basic idea behind choosing two different winter seasons is based on the concept of applying a differential split sample test, which is described at the end of this section. The question you have raised is interesting and it would be worth pursuing this issue. As explained in Sect. 2.2, observations with high temporal resolution were only available for the last decade. But this would also raise the question which winter seasons during the last decade are closest to average? We therefore decided to select two different winter seasons in order to test the models for a wide range of possible meteorological conditions.

"P4074 Line 3: On the basis of which criteria?"

The subdivision of sub-catchments was carried out using digital elevation data, as it is a straightforward approach. In a subsequent step, hydrological response units were derived using land use and soil information. This section will be completed with respect to that.

"P4074 Line 4: What information have you used for doing this?"

The local water supply company provides seasonal flow rates, which have also been observed using gauging stations. We will refer to these.

> *"P 4074 Line 6pp: Which parameters where calibrated? Which calibration scheme was used, what are the quality criteria's? What was the quality of the model after calibration? Why have you calibrated the model by using met-stations instead of using WRF fields, which are used for the successive model runs? What was calibrated in which snow model? Why have you used WRF fields here and no meteorological stations? Please show why a calibration scheme which id based on two input data sets (meteorological stations for the hydrological part/ WRF for the snow model part) is consistent and applicable."*

The model was calibrated manually by maximizing the Nash-Sutcliffe model efficiency. We will add the model efficiency for both calibration steps. Mainly the parameters of the soil, the runoff concentration as well as the routing model were altered. In addition the temperature-index model was calibrated for the long-term simulations. We will briefly describe this. We think that the description of every single parameter would be too detailed for the scope of this study. However, we will provide the number of adjusted parameters.

You ask for proof of the consistency and applicability of the calibration procedure, since it is based on two input datasets. We cannot prove this in a more rigorous way, e.g., by using a cross-validation. However, the results of the passively coupled atmospheric and hydrological model are still good. Since we have also considered a validation period for each calibration step, we expect the model to be valid.

> *"P4074 Line 21: Please show the network."*

The network will be added to the map (Fig. 1).

HESSD

11, C1521–C1531, 2014

"P4074 Line 24: Why have you chosen such a generalized illustration of precipitation? If 19 stations area available a spatially distributed illustration would be possible by e.g. showing the Nash Sutcliff coefficient for any station. This would also allow a more precise estimation of the quality of precipitation within the study area. Moreover, it would be also necessary to see if WRF is able to calculate the correct phase of precipitation."

We could provide a table including the precipitation depth for both the entire winter season and the snowmelt event in March/April 2006 for all stations. We could also present Nash-Sutcliffe model efficiencies for different time steps and both periods. However, we think that this information would go too far within the context of this paper. Such a table could be enclosed as an appendix.

"P4075 Line 12: The usage of Nash Sutcliff would be again more meaningful than r^2."

We will also provide this information for the temperature simulations.

"P4075 Line 16pp: This is critical and leads to a negative evaluation of the whole paper. As you said precipitation is a key parameter with respect to snow cover modelling. You mention the problem of the areal representativeness of point measurements in the introduction and mentioning that this is the reason for using models like WRF. But right now you are going in the opposite direction because you are argue that point scale precipitation is better than WRF."

As outlined in the general comments section, we will provide results based on WRF precipitation. These results will clearly underline that is not necessarily possible to simulate single events reliably. By substituting simulated with observed precipitation we show that in this case, point scale information is better than the model output because the snowmelt models perform well using observed precipitation (we will carry out a comparative study as suggested by Anonymous Referee 2). Alternatively, this configuration underlines the applicability of temperature, humidity, radiation, and wind speed. We believe that these results are worth being reported, regardless of the fact that it somewhat constitutes a failure story. As recommended by Andréassian et al. (2010, p. 855), "hydrologists should be encouraged to dedicate part of their publication efforts to reporting their mistakes or what can be called negative results." However, there is still a significant contribution to hydrological modelling in a positive sense, since temperature, humidity, radiation and wind speed can be downscaled with sufficient accuracy for snowmelt modelling at both scales. These results may be of interest to the scientific community since these observations are sometimes hardly available.

> *"P4076 Line 1pp: The results are hard to interpret. First of all you are discussing the results for the calibration period by using no quantitative measures. Moreover, it is not clear if you have used WRF for the calibration of the snow models (as it was mentioned before) or WRF and measured precipitation which would be more meaningful as you are driving the models in this configuration."*

We will add further quantitative measures including Nash-Sutcliffe model efficiencies. We agree that the current explanations with respect to calibration and the precipitation input are ambiguous and unclear. In all cases, we used observed precipitation for the calibration as well as the validation. We will clarify all statements that refer to the calibration of the snowmelt models.

> *"P4077 Line11pp: You are validating your models mainly on the basis of snow melt. It would be good to know which kind of Lysimeter you have*

used. Moreover, it is well known that the melt rates measured by Lysimeters can be significantly biased. Hence, additional parameters would be needed for estimating the quality of the models (e.g. snow depth, SWE)."

The snowmelt lysimeter located at Torfhaus is an unenclosed type with a small rim above the collector, which covers an area of $2m^2$. A tipping bucket is installed below the collector in order to continuously record melt rates. We will complete this information about the lysimeter in Sect. 2.1. We agree that lysimeters may be biased, e.g., by lateral flow. However, we carefully checked the data by comparing total runoff depth with precipitation measurements and the timing of the melt rates. In our opinion, the lysimeter seems to provide reliable melt rates. Moreover, the evaluation against melt rates enables a comparison of the models across scales when considering stream flow observations at the catchment scale.

Since snow water equivalent measurements are not available, we will provide a figure including snow water equivalent simulations and the corresponding snow depth observations at Torfhaus.

"P4078 Line 3: How? Which parameters of which models? What are the input parameters?"

As explained earlier, we will clarify the text with respect to precipitation input. Moreover, we will briefly explain which parameters were adjusted (mainly concerning the calibration of forest effects).

"P4078 Line 7pp: I don't see your calibration strategy. How have you altered the parameters? How have you defined parameter ranges?"

The parameters, which will be added to the text as stated in the previous comment,

HESSD

11, C1521–C1531, 2014

were also altered manually, as were the parameters for the hydrological model of the Sieber catchment.

> *"P4078 Line 29pp: This is problematic because in the abstract you are talking about an approach able to simulate snow pack and snow melt processes on the catchment scale by using WRF fields. First of all you have excluded the precipitation field and now you are not able to give a measure for the accuracy of the snow pack evolution."*

We agree that snowmelt alone does not necessarily provide a comprehensive measure of snow pack evaluation. Since we will add snow depth observations, which also enable a model evaluation throughout the entire winter season, this statement will be proven in more comprehensive way.

> *"P4079 Line 13pp: I don't think that this kind of evaluation is adequate. The model should be quantitatively validated by available and meaningful parameters. At least everything is guessing in here. It is probably that the inclusion of the canopy stands in ESCIMO improves the whole model as it was shown by Warscher (2013) but they have used a consistent model package and consistent validation data for showing the effect of different model components on the quality of the results. Here we have a highly calibrated model and an improvement can be due to the inclusion of the named effects or we have a typical case for being right for the wrong reasons."*

We do not agree with the reviewer's statement "everything is guessing in here". All snowmelt models were quantitatively validated including RMSE, correlation and a comparison of standard deviations. In contrast to exclusively using a single measure, this enables gaining insight into the question why models do not perfectly match the observed time series, as is always the case. The correlation gives some information about the mismatch in phase, whereas the RMSE provides information about the bias of the model (even though the central pattern RMSE is somewhat limited with respect to that). In our opinion, both Taylor plots adequately represent the model performance for single models and among all models between both scales. However, we could provide a table with all performance measures including an additional column that includes Nash-Sutcliffe model efficiencies.

The comparison of water vapour mass flux simulations (Fig. 11) is not intended to be a validation of models because adequate measurements are not available. We do not state that any of the models perform better with respect to runoff as a result of these processes representations. We merely pointed out that the inclusion of canopy processes is generally more reliable and may be relevant for other sites.

The concept behind this comparison is that processes may change due to scaling, which may also result in different representations of specific processes. We think that the different sign of water vapour mass flux is worth noting because the different representations of processes may lead to opposing results with respect to this component of the water balance. In other regions, this difference may also affect the water balance simulation to a higher degree. In conclusion, Fig. 11 is neither meant for model validation nor for showing the effect of different model components on the quality of results (as described by Warscher et al., 2013). We will clarify that this comparison is intended to emphasize the different behaviour of the tested models for plausibility only.

References:

Andréassian, V., Perrin, C., Parent, E., and Bárdossy, A.: The Court of Miracles of Hydrology: can failure stories contribute to hydrological science?, Hydrolog. Sci. J., 55, 6, 849-856, http://dx.doi.org/10.1080/02626667.2010.506050, 2010.

Warscher, M., Strasser, U., Kraller, G., Marke, T., Franz, H., and Kunstmann, H.: Performance of complex snow cover descriptions in a distributed hydrological model system: a case study for the high Alpine terrain of the Berchtesgaden Alps, Water Resour. Res.,

HESSD

11, C1521–C1531, 2014

Interactive Comment

49, 2619–2637, 2013.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 11, 4063, 2014.

HESSD

11, C1521–C1531, 2014

Interactive Comment