Hydrol. Earth Syst. Sci. Discuss., 7, C1594–C1604, 2010
www.hydrol-earth-syst-sci-discuss.net/7/C1594/2010/

Hydrology and Earth System Sciences Discussions

Interactive comment on "WRF simulation of a precipitation event over the Tibetan Plateau, China – an assessment using remote sensing and ground observations" *by* F. Maussion et al.

F. Maussion et al.

fabien.maussion@tu-berlin.de

Received and published: 28 July 2010

Authors' response to Anonymous Referee 1 (C1517, 22 Jul 2010)

General reply

We take note of the thorough review from the Referee, which is very critical to the approach and methods used in this study. This points out several weaknesses in the presentation of our results, and we appreciate being aware of them prior to publication. However, we also think that there may be a general misunderstanding about the frame and objectives of this study, and we wish to clarify this point here.

The main critic coming from Referee 1 is based on the modeling strategy that is not the one commonly used in forecasting systems. However, we did not intent assessing the WRF model in a forecasting configuration in this study. Moreover, the purpose of our study is optimizing the set-up of a model strategy including a suitable model configuration for producing a regional meteorological reanalysis for the study region, i.e., the region covered by the WRF domain of 10 km grid spacing, extending over the last decade. Such data do not exist at sufficient spatial and temporal resolution for the Tibetan Plateau but are essential for hydrological and glaciological studies depending on meteorological data sets as input. This holds particularly true in this region of complex terrain, where weather stations are sparsely distributed and completely missing at high elevations, such that spatial interpolation methods will not be able deriving reliable meteorological data.

Our strategy is using the WRF model for physically based, dynamical downscaling of the low-resolution gridded data set delivered by the Final Analysis product from the Global Forecasting System (GFS). Thus, this study was designed to examine the downscaling capabilities of the WRF model for liquid and solid precipitation for different model configurations. Observational data of suitable accuracy are required for this purpose.

We used three-hourly gridded data from the TRMM for assessing amount and timing of precipitation, gridded data from MODIS for delineation of areas of snowfall, and all publicly available daily data from Chinese weather stations for amount and type of precipitation. Our study was more comprehensive than described in the manuscript, so we will be able including more material, as proposed by the reviewer.

One of the important findings of our analyses is the fact that observational data are partly less reliable than the model results. This has been also reported in other studies, e.g. for problems in TRMM data showing inaccurate precipitation values over snow (Yin et al., 2008). This implies that data from different sources can only be compared to each other to validate the model, but no data set is suitable serving as accurate reference required for model verification.

The reviewer's comment signalizes that the purpose of the study is not sufficiently explained in the manuscript, and we would like to modify it accordingly in a revised version. In addition, the reviewer correctly states that we should put more effort on the method's particularities. We would be able and willing including more results as proposed by the reviewer in the specific comments.

Reply to specific comments:

RC: Usually, the model is set up to run continuously for one case study. For example, if an event lasted for 3 days, one would carry out a 72-h simulation. This way the model performance can be fairly evaluated, and the evolution and the development of the rainfall system can be well examined. However, the simulation in this study is separated for 7 runs and only the precipitation between 12-36 h of each run is used for verification. I don't understand the purpose of this kind of modeling setup.

AR: This statement is only true for model validation studies carried out for evaluating the forecasting skills of a NWP model. In modeling studies targeting on climate projections long-time integrations are carried out where the model is initialized once since no data exist for reinitialization. Reinitialization of a NWP model at a certain interval is, however, the common strategy in reanalysis studies preventing the model deviating too much from an initially observed state, and to deliver daily values that have been produced by the NWP model at a homogeneous skill level. The influence of different reinitialization intervals on model output has been subject of scientific studies (e.g. Lo et al., 2008 compared yearly, monthly and weekly runs with each other).

The reviewer is right that we did not show why we have chosen daily intervals for reinitialization, and why we discarded the first twelve hours. After two days of simulation, we expect the model to be less accurate than during the first two days. In addition, the WRF model (as any NWP model) requires some spin-up time to reach a balanced state with the boundary conditions given by the GFS data. We have therefore chosen to discard the first twelve hours, and then to store one day of model simulation, i.e., the WRF model output until 36 hours. We have carried out various analyses comparing the results from daily reinitialization with those from single runs that have been only initialized once. Some examples are shown in the supplemental document attached to this reply, demonstrating that daily reinitialization is superior to single initialization. We would like including such analyses in a revised manuscript since they demonstrate that daily reinitialization significantly improves the consistency of the model results with the observations.

We are targeting on using the regional meteorological reanalysis for hydrological and glaciological applications. A similar concept has been successfully used e.g. by Box et al., 2006 over Greenland to drive an underlying mass balance model, also demonstrating the suitability of our daily reinitialization strategy.

RC: If the WRF simulations are to provide rainfall or snowfall information over the mountains, it will be more desirable to do verification of 24-h accumulated precipitation. For example, calculating the scores by comparing the data in Fig. 2.

AR: Such analyses have been carried out but we did not include them in the manuscript to avoid a lengthy presentation. The manuscript presents time series of hourly precipitation values from the WRF model runs in comparison with three-hourly values from TRMM and with daily values from the Chinese weather stations (see Fig. 6). We would include also daily analyses on a spatial basis if desired.

RC: The WRF is run with 3 nested domains of 30, 10, and 2 km resolutions. This is quite a high resolution simulation. The model performance should be verified in somehow a mesoscale way. But, instead, the precipitation is verified against TRMM and MODIS data using a 7-d accumulated rain or snow. This is really disappointing. Why don't the authors verify 24-h accumulated precipitation forecast against observations?

AR: The high spatial resolution (2 km) of the innermost domain is one of the major goals of our modeling approach, since we target on downscaling precipitation data for a region of highly complex terrain. Mountain-valley structures are often showing elevation differences of 1 to 2 km within short distances of less than 10 km. As shown e.g. by Zaengl, 2007, increasing the spatial resolution of a model domain in areas of complex terrain can be highly beneficial for simulating precipitation fields. Downscaling to such high resolutions cannot be done in a single step, because synoptic-scale weather patterns have to be considered in the model, too. Thus, the set-up of three two-way nested domains has been chosen.

However, studies like the one from Zaengl, 2007 have also shown that increasing spatial resolution does not improve the model quality for precipitation caused by embedded convection. Thus, we decided to examine the WRF model output at all spatial resolutions. Since the spatial extent of the outer domain is only used for optimizing simulation of precipitation in the nested medium (MD) and innermost small (SD) domains, we only selected observational data from TRMM, MODIS and meteorological data from Chinese stations in these domains for model validation.

We presented results from analyses involving observations in the manuscript (cf. Figure 6) exactly as proposed by the reviewer. We did even more detailed analyses since we compared the observed daily accumulated precipitation measurements from the stations with hourly values obtained from the WRF model at 10 km resolution (WRF10) for the whole seven-day period. In addition, we also included the three-hourly TRMM time series for comparison. The WRF10 was chosen since not all stations are situated within the SD.

Further, we used TRMM data of the MD for spatial validation. All spatial analyses have been carried out using the output from the WRF30 runs, since the spatial resolution of TRMM is 0.25 deg, i.e., almost matching this grid, to provide spatially homogeneous and consistent comparison results. We agree with the reviewer that inclusion of additional spatial analyses on a daily basis would deliver more details. We excluded these analyses, since we recognized that the aggregated results for the seven-day period are in accordance with the results from the daily analyses.

MODIS data on snow extent used in our model validation are not available on a daily basis, since clouds generally prevent acquisition of snow data at the land's surface by optical systems. Many clouds are present during the precipitation event, while only few clouds have been present before and after the event. Thus, we designed the validation study involving MODIS data for the whole seven-day period only.

RC: The comparison with TRMM data is only to tell the performance of WRF30. If the results are not bad, it would be more interesting to discuss the high-resolution simulations of the event in WRF10 and WRF2. But, the authors chose not to examine any other fields except for precipitation. If the rainfall is comparable with the TRMM observation, what will be the purpose of the simulation if the authors did not look further? We could just use TRMM data to provide the precipitation information over the Tibetan.

AR: Generally, our study focuses on precipitation!

Since all analyses for model validation have been carried out within the MD, the comparison of the WRF30 output with TRMM data is interesting since the WRF30 runs have benefited from the higher-resolution runs by the two-way nesting method. We did not show this point in the manuscript, but would be able to include an analysis demonstrating the difference between our nested runs and runs using only one-way nesting, where the WRF30 results are not influenced by the higher-resolution WRF10 and WRF2 runs.

Table 3 presents a comparison of station measurements and TRMM observations together with the results from the WRF30 and WRF10 runs. Where possible, the results from the WRF2 runs are also shown in Table 3. The interesting finding is that the r.m.s. deviation of the TRMM data to the station data is much higher (31.9 mm/7 days) than the WRF30 (23.4 mm/7 days), and the WRF10 results where even better (13.6 mm/7 days). These values imply that TRMM data are not only less accurate than the WRF model results but also do not provide sufficient detail in regions of complex terrain.

RC: Of course it would be another story if the model setup is designed in a real-time *forecast mode. In that case, one can say that WRF can be used for forecasting purpose. However, this is not the intention of the study.*

AR: That is correct.

RC: The same situation aforementioned applies to the comparison between WRF10/WRF2 and MODIS snow data. I cannot see much value from the comparison. If I have MODIS data, I will trust the observation more than the model simulation. Furthermore, the verification is only made for the 2mm/7d and 7mm/7d thresholds. The authors did not show how much snow was estimated in the MODIS data. But, I guess 2 or 7mm of snow is very small for a 7-d accumulation. The scores are usually higher for smaller thresholds.

AR: MODIS data on snow extent at high spatial resolution of 500 m are a suitable way for analyzing WRF2 results by observations on a spatially distributed basis (comparison of WRF2 results for individual grid points with station data have been presented in Table 3). As shown for the TRMM data, observational data can be less accurate than model results. Thus, trusting in data is not always a suitable approach.

MODIS does not deliver data on snow water equivalent, and thus no direct information on precipitation amount is available. A further argument against using MODIS data for regional meteorological reanalysis is the limited availability of snow data over time due to the already mentioned problems with clouds. This is the reason that the analyses involving MODIS have been carried out for the whole seven-day period only. Instead, WRF provides both precipitation type (snow/rain) and amounts (mean intensities) at arbitrary, regular intervals (e.g. days). Thus, model validation by MODIS is a useful approach to assess the model's capabilities in producing daily data on snowfall. Since we are targeting on using the regional meteorological reanalysis for hydrological and glaciological applications, snow data are of great importance.

Figure 4 demonstrates that we did not only use two distinct threshold values (2 and 7 mm/7 days) but the full range of thresholds. The results shown in Figure 4 have been used for choosing the threshold values for which additional spatial analyses have been presented in the manuscript. We would be able including spatially distributed results for any of the thresholds used in Figure 4 if desired.

The reviewer is correct that a value of 7 mm/7 days or less is a small value. Large areas of the MD have been snow-free as indicated by MODIS demonstrating that small snowfall rates are common in this region, even during this exceptional event. High skill scores for low thresholds indicate that the model is able to sensitively distinguish between snow and rain, respectively between snow and no precipitation. Figure 4 shows that the WRF10 results in the SD are superior to the WRF2 results as measured by the HSS, but differences are generally small, mainly because of the two-way nesting. Interestingly, the HSS of the WRF10 in the MD reaches its maximum not at lowest thresholds, in contrast to the WRF2.

We would anyway take the reviewer's comments as motivation for including more details on this kind of analysis. Results from WRF30 runs for the MD and SD could also be shown in Figures 4 and 5, because the reader would then be able to assess the influence of the spatial resolution on the model's skill with respect to precipitation.

RC: The result presented here does not necessarily guarantee a good model performance. If the authors intent to evaluate the model performance over the Tibetan, more case studies are needed. It can not be over emphasized by just one case. If possible, a statistical evaluation of one-month simulations would be better. For example, how does the model perform if there is no precipitation?

AR: Generally, no model validation study is able to prove the applicability of a model under all possible conditions. Such studies are always using certain meteorological situations as statistical samples of a much larger population.

The reviewer is right that the performance of the model when there is no precipitation is an important question. For this reason, we analyzed the propagation of the precipitation front, e.g. by the timing of precipitation for the Chinese stations (Figure 6). But our manuscript does not address this aspect comprehensively. Inclusion of daily spatial analyses, as proposed by the reviewer, would deliver more information also for areas without precipitation.

If we would extent our study over one month, as proposed by the reviewer, the model period would include meteorological situations without precipitation. Then, a suggestion could be to perform the model validation for a different season or a different year. Finally we would have to perform the whole regional meteorological analysis, and would possibly see that we did not use a suitable model set-up. This is exactly what we wanted to avoid by our study design, and we think that our study could be very helpful to other researchers in setting up their own NWP model to be used for non-forecasting purposes like generating input data for hydrological modeling.

We could nevertheless offer including results from WRF runs carried out for the whole month (October 2008), validated by TRMM data for the MD. The interesting aspect of this analysis is that it reveals how much the individual event has contributed to monthly precipitation. This will demonstrate the importance of such exceptional short-term events on longer time scales.

RC: The conclusion about different settings of model physics is not convincing. It is made only based from a simple precipitation score. More examinations are needed in order to make a robust conclusion about the performance of different physics.

AR: We have presented all relevant scores commonly used in comparable studies not only one score as stated by the reviewer. But we could also include the contingency tables and samples sizes in Table 5 and Figures 3 and 5, such that the reader is able to compute all scores by himself/herself.

The results from the comparison of the WRF runs for different model physics are presented in Table 5. Results are validated using TRMM, MODIS and the station observations for comparison. Our main finding is that there are only small differences between the experiments. This is interesting, because it demonstrates that there is no high sen-

HESSD
7, C1594–C1604, 2010

sitivity of the model physics for reanalyzing even a strong precipitation event commonly regarded as the most critical meteorological situation.

Of course, we could perform more experiments. But as already mentioned, we would finally end up in producing the regional meteorological reanalysis for all model physics options, which would be a waste of computing time.

RC: Overall, there is not much scientific contribution, and not much we can learn for the forecast purpose, either, from this paper.

AR: We hope that our explanations will give the Reviewer and other persons following the discussion the impression that we would provide scientifically valuable information in a revised manuscript.

References:

Box, J. E., Bromwich, D. H., Veenhuis, B. A., Bai, L.-S., Stroeve, J. C., Rogers, J. C., Steffen, K., Haran, T., and Wang, S.-H.: Greenland ice sheet surface mass balance variability (1988–2004) from calibrated polar MM5 output, J. Climate, 19, 2783–2800, 2006.

Lo, J. C.-F., Yang, Z.-L. , and Pielke Sr., R. A.: Assessment of three dynamical climate downscaling methods using the Weather Research and Forecasting (WRF) model, J. Geophys. Res.-Atmos., 113, D09112, 2008.

Yin, Z.-Y., Zhang, X., Liu, X., Colella, M., and Chen, X.: An assessment of the biases of satellite rainfall estimates over the Tibetan Plateau and correction methods based on topographic analysis, J. Hydrometeorol., 9, 301–326, 2008.

Zaengl, G.: To what extent does increased model resolution improve simulated precipitation fields? A case study of two north-Alpine heavy-rainfall events, Meteorol. Z., 16, 571-580, 2007.

Please also note the supplement to this comment:

HESSD

7, C1594–C1604, 2010

Interactive Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

http://www.hydrol-earth-syst-sci-discuss.net/7/C1594/2010/hessd-7-C1594-2010-supplement.pdf

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 7, 3551, 2010.