Interactive comment on “Large scale atmospheric forcing and topographic modification of precipitation rates over High Asia – a neural network based approach” by L. Gerlitz et al.

Authors reply to
F. Maussion (Referee)

1 General Comments
This manuscript presents a new estimation of monthly precipitation rates generated for a large target region of High Asia (1989-2011). The very high resolution (1km) gridded precipitation data are obtained by training an artificial neural network, using large-scale dynamical fields from reanalysis data and local geographical features as input. The downscaling method is described, validated, and a few examples of its output are shown.

The study is timely and is a welcome application of downscaling approaches in this data-scarce region. However, I think that the manuscript could be improved substantially by shortening some less relevant text parts and by adding new figures and analyses. Two major issues have to be addressed before publication:

1.1 Validation strategy

The authors use 157 meteorological stations for calibration during the period 1989 to 2000, and further 18 “independent” stations for validation during the period 2000 to 2011. Little is told about what motivated the selection of these 18 stations, and a couple of them have long/frequent spurious data gaps, making them unusable. Furthermore, I would like the calibration/validation strategy to be better explained:

> *The validation stations were selected as to represent the major precipitation regimes in the target area – ranging from the humid Himalayas slopes to the arid steppe landscapes and deserts of the TP and the Tarim Basin. We will point out this information more clearly in a revised manuscript version.*
>
> *The stations with gaps are both situated in the Indian Plains and thus mark the border of our target area. Due to the fact that only a few stations in the lowlands were used for the model implementation, the manuscript shows, that the results are not reliable (1295 L6ff). The revised manuscript will focus on the core areas of the target area more clearly.*

• the calibration period starts in 1989, but ERA-Interim starts in 1979. Why?

• the considerable database of 157 stations is not used at it’s full potential if the calibration is done for 11 years only. Statistical downscaling approaches can take full advantage of all available data by using cross-validation techniques, in which data samples of data are used in turn for calibration and for validation (i.e. k-fold cross validation). This approach would have the advantage that a real estimation of the skill of the ANN can be estimated at all stations’ locations, not only on a fixed sample of validation stations for which the true representativeness cannot be assessed.

Statistical downscaling procedures and Climate regionalization techniques are often advantageous in applied science, particularly due to lower computational demands. All calculations were carried out on a personal computer, however, the training of the neural network took several weeks to converge.
A cross-validation strategy would certainly better ensure the quality of the downscaling approach, but would multiply the calculating time.

The separation of the data set seems to be the only possibility to conduct the validation spatially and temporally independent. The utilization of the whole time series might lead to overfitting of the meteorological predictor variables.

• the 18 stations used for validation are not truly independent, since they are also used to tune the degrees of freedom of the ANN (Fig. 5). As discussed by e.g. Elsner and Schmertmann (1994), it is crucial to define a truly independent data sample for validation, that have not been used for either training or tuning of the model.

The 18 validation records were used to validate the models, but not for the tuning of any model internal parameter. We implemented several neural network models and evaluated their predictive performance for the independent data set. Afterwards we utilized the "best" model for further investigations.

This is done eg. In Schönwiese, C.-D., Walter, A., and Brinckmann, S.: Statistical assessments of anthropogenic and natural global climate forcing. An update, Meteorol. Z., 19, 3–10,
-
I guess that your argument fruitful, if the stopping criteria of a neural network is based on the validation data set. This is not the case in our study. We will state that more clearly in our revised version.

To summarize, I suggest the authors to better explain the rationale of their validation strategy as well as their reasons for the choice of the calibration/validation/application time periods.

1.2 Objectives of the study

I believe that the study could be improved with more clearly stated objectives. The authors correctly discuss the advantages / drawbacks of dynamical and statistical downscaling methods, as well as the reasons for the need of new methods to estimate precipitation in High Asia. However, little is done to really discuss the added value of their new precipitation estimates.
Since they do not seem to wish to make their dataset freely available, more emphasis could be given on the question raised in the title of the study: "atmospheric forcing and topographic modification of precipitation rates". The analyses proposed in Fig. 7 and 8 are not really conclusive with respect to the title. With selected examples (for example "zoomed" target regions), they could discuss their results with regard to orographic precipitation more effectively. Currently available gridded datasets (TRMM, HAR) are too coarse for comparison, but the 5km TRMM rainfall climatologies from http://www.geog.ucsb.edu/~bodo/TRMM/index.php could be a starting point.

Another axis of research could be to continue further the discussion started in the sensitivity analysis (Fig. 8). Would it be possible to select the most important predictors based on Fig. 8 and then realize a distributed evaluation of their relative importance? (for example, a map of the 0.1, 0.5 and 0.9 percentiles of the influence of elevation or the wind index)

> *Our study is funded by a BMBF bundle project on regional and local scale climate variations and their influence on various climate sensitive environments. The choice of the target area as well as the required parameters were made in cooperation with our project partners.*
>
> *The suggestion of zoomed target areas (or precipitation profiles) is certainly good, particularly for the Himalayan slopes.*
> *Additionally we will improve the sensitivity analysis by assessing the effect of the topographic predictor variables for different atmospheric situations.*

As a whole, I find the study interesting and hope to see it published in a more elaborated form (see also specific comments below).

Best regards,
Fabien Maussion
Institute of Meteorology and Geophysics
University of Innsbruck

2 Specific comments

Throughout the manuscript:

The ratio text/figures is not really balanced. The text can be sometimes repetitive. As an example: P1277 L20 ! P1278 L5 contains several sentences repeating more or less the same information and could be shortened. Chapter 2 is a comprehensive review of knowledge about precipitation in High-Asia, but it could be skipped almost entirely without much lost for the rest of the paper. I am not asking for an entire suppression but I suggest the authors to consider shortening their manuscript as a whole.

> *Section 2 reviews the major precipitation regimes of the target area. We included that information in our manuscript in order to clarify the choice of relevant predictor variables afterwards. However a shortening of that introductory chapter without losing the most important information is certainly constructive,*

Presentation of the results. In their current form, the figures are difficult to interpret for the reader:

• the validation stations are marked in red in Fig.1 but there is no simple way for the reader to know which validation station is where

> *in fig 6/8 the geographical coordinates of the stations are specified. This allows to find their location in the map. If necessary we could certainly name the stations in fig 1.*

• while Fig. 6 allows a coarse evaluation of the model skill (which looks quite good), it is not really possible to assess an important aspect of the precipitation estimates: their capacity to reproduce observed inter-annual variability. More analyses/figures are needed here

> *Figures showing the observed and modeled annual precipitation for the validation stations will be included in the revised version.*

• the choice of a continuous diverging color table and of topographical colorshading makes it really difficult to associate a pixel on the plot with a precipitation value (e.g. Stauffer et al., 2014). Fig. 7 (top-right) is a good example of this problem: I find it difficult to distinguish more than three zones of precipitation, and since the dataset is of very high resolution it would be good to be able to distinguish between topographical (artificial) shading and orographic (real) precipitation.

> *It is truly difficult to show the results of high resolution climatic fields for such a large target area. A suggested above, we will definitively show zoomed maps for selected areas.*
> *Further we will try to better present the precipitation fields by using discrete color schemes.*

Since I expect the text to be modified substantially I will not make too many specific comments here:

L24: “precipitation-genetic” does not seem to be a very commonly used expression.

P1284
Fig 1: please explain the choice of the target area. Is there a specific reason for omitting the Karakoram/Hindu-kush? Doesn’t the presence of very different settings (e.g. Indian lowlands VS central TP) make the job of the ANN considerably more difficult? The stations that are referred to in the text should be named in Fig. 1

> *As stated above, the target area is chosen by in order to support our project partners with high resolution climate estimates. The presence of different precipitation regimes makes the implementation of adequate statistical transfer functions more complicated on the one hand, on the other it enables the investigation of specific topo-climatic processes under varying atmospheric conditions.*

L5: how do ANN handle missing data? How many gaps were found in the time series?

> *Most of the time series were complete. Missing data were eliminated. We will clearly point that out in our revised manuscript.*

L7: I am not sure if ERA-Interim is assimilating precipitation directly. In all cases, it is not relevant since ERA precipitation is not used as predictor.

> *Precipitation is not assimilated in the ERA-Interim reanalysis. However, the large large scale atmospheric predictors are. This is an advantage for our study, since we can assume, that the predictor variables are perfectly simulated by the reanalysis product.*

L12: these other predictands are mentioned here, but they are not validated and almost not used afterwards (in particular number of precip days). Are they also influenced by orography?

> *The wet days, which were used as additional predictors did not reveal important additional information, since the variable is highly correlated with the monthly mean and the maximum daily precipitation. We will give some more information about the predictant.*

P1287
L2: 4.4 percent: in Fig. 2, the number indicated is 0.044. I guess this is an error

> *0.044 indicates 4,4% explained variance. Could you please specify, why you think that this could be an error?*

L10-L12: EOFs 4 to 6. Indeed, they explain less variance as the other ones but, as stated by the authors, a very large part of the variance is due to the annual cycle (EOFs 1 and 2). The remaining EOFS could become particularly important for inter-annual / intra-seasonnal variability.

> *This is true and for that reason we used all 6 EOFs as potential predictor variables. However the sensitivity analysisindicates, that the model is not very sensitive to EOF 4-6.*

P1288
L17-20: your domain also includes lowlands. Discuss the choice of the 500hPa level in this case.

> *As already stated the core of our target area ist the TP and the adjacent mountains. The results for the lowlands are not very reliable for the lowlands (1295 L6ff).*

P1289
L11-12: does the strength of the wind field also influences the wind effect parameter? Would it be useful to also include wind speed as a predictor?

> *The wind speed has not been used as a predictor. However the wind shear was utilized but the sensitivity analysis didn't reveal any significant effect.*

P1290
L7: are the predictands also normalized? This does not make much sense for the considered variables. How do ANN handle skewed distributions? Do they predict negative precipitation values as linear models would?

> *Due to the non-linear and non-parametric structure, skewed distributions shouldn't be a problem for ANNs.*
> *Some negative values were were detected for some January situations, but were not very pronounced (>-5mm). This makes complex ANN superior compared with linear models.*

Fig. 5. It seems that the choice of the number of neurons is strongly influenced by one or two stations showing the largest errors (the other stations do not vary much between 2 to 8 neurons). This calls for the use of a much larger sample of data for this tuning procedure. Why not using the 157 stations during 2001-2011 to select the degrees of freedom and then use independent stations for validation?

The two stations showing the largest errors for ANNs with only few hidden neurons are located in the arid landscapes of the northern target area (P1294 L4). This indicates a large wet bias of those models for the arid regions. Thus, the deviation of only one more two specific stations indicates the importance of a non-linear model architecture for such a huge target area.

The use of the training stations for the evaluation of different neural network architectures is risky, since the topographic predictor variables wouldn't be independent in that case. This would certainly lead to overfitting and a rather unrealistic spatial distribution of precipitation as shown in fig 5.

This color scale has an abrupt change of color at 260 mm . This creates a completely arbitrary and non-physical "border" in the Himalayas, which makes any interpretation difficult.

Fig5? The "border" in the map shows the abrupt precipitation decrease at higher elevations and is actually not a result of the color scheme. The maps by Bookhagen and Burbank (2006) indicate similar results.

Fig. 6. I suggest to ignore the stations with too many data gaps. The Y-axis range should be adapted to the precipitation sums, the first four plots are barely readable. Units are mm.month-1? Would it be possible to add annual precipitation as dots for example, to assess inter annual variability?

The lowland stations will be ignored due to the low reliability and the large number of gaps.
Annual precipitation will be either included in that figure or analyzed in an additional one.

Fig. 7. I am concerned about the high amount of precipitation in in January and the spurious artefacts in July at the mountain ranges. But these could be related to the colorscales. Can you detect an east to west decreasing gradient at the Himalaya range as documented by e.g. Bookhagen and Burbank (2010) or Maussion et al. (2014)?

The East west gradient can be detected during July (although not as strong as for example shown by Bookhagen and Burbank). During January, the western Himalaya shows a maximum of precipitation.
Color schemes will be revised.

P1297 L16 and P1298 L25: Why not show the figures? 8 figures is not too many for a paper and especially the annual amounts would be very interesting to show (annual amounts are a variable people can evaluate more easily than monthly sums).

A map of annual precipitation sums will be included!

P1299, Fig 8. What motivated the choice of the four example stations? Without a location on the map it is difficult to assess these results. The difference between left and right panels is not explained in the legend. It could be better to scale the Y-axis in percentage of total amounts instead of mm.

The stations utilized for the sensitivity analysis are exemplary for different climatic regimes. Thus they show a very clear picture of the major predictant variables in the subregions of our target area.

P1303 L5: The HAR recently changed its name to "High Asia Refined analysis" (http://www.klima.tu-berlin.de/har). Maussion et al (2014) is the correct reference.

References

Bookhagen, B. and Burbank, D. W.: Toward a complete Himalayan hydrological budget: Spatiotemporal distribution of snowmelt and rainfall and their impact on river discharge, J. Geophys. Res., 115(F3), F03019, doi:10.1029/2009JF001426, 2010.

Elsner, J. B. and Schmertmann, C. P.: Assessing Forecast Skill through Cross Validation, Weather Forecast., 9(4), 619–624, doi:10.1175/1520-0434(1994)009<0619:AFSTCV>2.0.CO;2, 1994.

Maussion, F., Scherer, D., Mölg, T., Collier, E., Curio, J. and Finkelnburg, R.: Precipitation Seasonality and Variability over the Tibetan Plateau as Resolved by the High Asia Reanalysis*, J. Clim., 27(5), 1910–1927, doi:10.1175/JCLI-D-13-00282.1, 2014.

Stauffer, R., Mayr, G. J., Dabernig, M. and Zeileis, A.: Somewhere over the rainbow: How to make effective use of colors in meteorological visualizations, Bull. Am. Meteorol. Soc., 70(306), 140710055335002, doi:10.1175/BAMS-D-13-00155.1, 2014.