

Large scale atmospheric forcing and topographic modification of precipitation rates over High Asia – a neural network based approach

L. Gerlitz, O. Conrad, and J. Böhner

–

Final Response to the reviews of the Discussion Paper ESDD 2014-50

Dear Ladies and Gentlemen,

Thank you very much for your fruitful comments concerning our submitted manuscript. The revised manuscript has been recently uploaded. Please find below our detailed response to your general and specific recommendations.

Reviewer 1

1.1 Validation strategy

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

We have to admit, that at the start of our project work the ERA-Interim data were only available from 1989 onwards and we didn't consider to extend the data set. However the investigation of available observations revealed that high quality observational data (without to many data gaps) were only available recent decades. We point out the observational constraints in our revised manuscript in the data section.

- **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.**

- **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.**

As already stated, a cross-fold-evaluation of the modeling approach would lead to an inappropriate increase of computational demands. Due to contradicting reviews (reviewer 2 in general found our evaluation adequate) we decided to keep our evaluation strategy. Particularly the fact, that the results coincide with other studies mentioned in our manuscript supports the suitability of our approach. In our revised manuscript, we mentioned your concerns in the evaluation section. In general we tried to highlight the feasibility character of our study in the introduction and focused on the idea of the modeling approach as a promising alternative to traditional statistical and dynamical downscaling applications.

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)

As stated we tried to revise our main objectives and highlighted the feasibility character of our approach. In this regard we highlighted the disadvantages of site specific statistical downscaling and extremely computational demanding dynamical downscaling applications in our introduction and suggest our approach as a promising alternative.

With the objective of better analyzing the large scale and local scale mechanisms, resulting a spatial and temporal precipitation differentiation, we improved our sensitivity analysis. Particularly the sensitivity of the model to topographic predictor variables was assessed separately for moist and dry atmospheric conditions.

Additionally we provide a map with annual precipitation sums for the entire target area as well as for few enlargements in our revised manuscript.

2. Specific comments

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.

We tried to avoid repetitive text parts in our revised manuscript, particularly in the introduction and the general section on pluviometric regimes. In the latter section we tried to abstain from rather unimportant information. However, we believe that the major part of the section is important for the understanding of the choice of adequate predictor variables.

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.

We revised our maps of monthly precipitation sums. For the new Figure of annual sums we likewise used discrete color schemes.

P1227

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

The term was substituted.

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

The choice of the target area was limited by available observations. Additionally the publication is part of an interdisciplinary project, which mainly deals with the Central and Eastern part of the Plateau. However, we think it is rather unusual to mention that in the manuscript. Fig. 1 shows the names of stations which were used for the validation procedure in our revised version.

P1285

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

Missing data were removed in a data preprocessing step. We point that out in our revised manuscript in the data section.

The sentence concerning the data assimilation of precipitation was misleading and has been removed.

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 maximum daily precipitation was used for the calculation of the precipitation sensitivity which is illustrated in Fig. 7 and is addressed by the sensitivity analysis. The number of rainy days is actually a function of precipitation sum and intensity and didn't reveal any new information. We clearly state that in the revised version (see results).

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-seasonal variability.

As already stated in our first response these EOFs were certainly used for the downscaling approach. However the sensitivity analysis did not show a reliable response of the model to EOF 4-6.

P1288

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

This is done in one sentence in our revised version. In general we point out more clearly (particularly in the results section), that the results for the Indian lowlands are not reliable due to an insufficient number of observations in that particular subregion.

P1289

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

The speed is not included, the equations for the derivation are given in section 2.

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?

As already stated in our first response, the normalization of the predictands should not lead to any problems, since the ANN model approximates any skewed non-linear function. The occurrence of negative values was negligible.

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 model with only few neurons has a large wet-bias for the central Asian deserts. This is now clearly stated in our manuscript. That indicates that models with too few neurons are not capable to capture the varying climatic characteristics of the target domain.

As stated in our first response, the assessment of spatially distributed precipitation rates requires spatially and temporally independent data for evaluation. The error of the neural network model for the 157 station decreases with model complexity. If we used the same stations for another period for the identification of the best architecture, this would certainly lead to an overfit of the topographic predictor variables, since in that case, these are not free.

As already stated we highlighted your concerns in our revised version.

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.

The sharp precipitation gradient in the Himalayas above 4000 m elevation is actually evident in our data set. However this gradient coincides with TRMM-based studies such as

Shrestha, D., Singh, P. and Nakamura, K.: Spatiotemporal variation of rainfall over the central Himalayan region revealed by TRMM Precipitation Radar, *J. Geophys. Res.*, 117(D22), D22106, doi:10.1029/2012JD018140, 2012.

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?

Figure 6 has been revised. Annual Precipitation sums were included.

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 artefacts, particularly at the Eastern margins of the Plateau, were probably incorrectly interpreted as orographic precipitation. In our revised manuscript, we highlight that problem of our data set (particularly with the enlargement of annual precipitation amounts).

The gradient at the Himalayan slopes can be observed. As already mentioned, the gradient in the lowlands is not reproduced, which most likely a result of the insufficient number of observational records.

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 amounts has been 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 choice of the locations for the exemplary sensitivity analysis was based on the idea to show results under different pluviometric regimes. We tried to state that more clearly in our revised manuscript. The Figure in general has been revised.

However, we refused to change the scale into percentage, with the objective not to lose the information on the actual model response.

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.

Reference has been changed

Reviewer 2

1. Overall comments

- Evaluation/validation procedure: the validation procedure is based on the use of 18 independent in-situ stations, for which the data in the period 2000-2011 have been considered.

On the contrary, 157 stations in the period 1989-2000 have been used for the analysis of local-scale precipitation rates as an essential part of the downscaling/ANN framework (i.e., for calibration purposes). I would like to find in the paper more explanations

about the choice of those specific 18 stations used in the evaluation procedure.

In addition, I suppose the observed datasets include missing data and gaps. How are these gaps (if any) handled by the ANN?

The stations in the evaluations were subjectively chosen with the objective of representing the different climatic subregions of this vast target area. We tried to state that more clearly in our revised version.

Concerning the data gaps. The data set had few. Missing values however were rejected and thus do not lead to any problems for the model implementations. This is likewise stated in our revised manuscript version.

- Why the authors did not use a cross-validation strategy in order to better exploit information from all the available stations? Is the reason related to the computational time that is required for such an “exercise”?

As supposed, a cross validation strategy would increase the computational demands of the approach in an inappropriate manner. In our revised introduction we highlighted the low demands of our approach as an advantage, particularly in comparison with sophisticated dynamical downscaling applications.

However, in our Evaluation-section, we no pointed out, that cross validation techniques would certainly better assess the overalls performance of the model.

Additionally, for the evaluation of distributed statistical downscaling applications, spatially and temporally independent data sets are important. Simple cross validation techniques (omitting one station or one time step at each time) cannot be spatially and temporally independent, which might lead to a overoptimistic results.

- The proposed method is very interesting and effective for obtaining high spatial resolution

(1 km) precipitation fields. However, as the authors also highlight in the conclusions, the achieved monthly temporal resolution is really too coarse to be of use for impact and assessment studies requiring, at least, a daily resolution. Therefore, the method is interesting but it does not seem to be really (fully) effective for practical application

purposes, or unless another temporal downscaling is applied. Please suggest some “solutions” to this drawback or discuss it better.

In our original manuscript we already discussed the possibility of combining dynamical downscaling and statistical parameterization techniques. Dynamical downscaling applications are particularly advantageous for the representation of mesoscale atmospheric processes. Thus dynamically downscaled fields could improve the suitability of free atmospheric predictor variables for the presented approach. If for example frontal systems are well captured by the atmospheric model, an assessment of daily precipitation rates might become possible. In the revised manuscript we tried to better highlight that “solution”.

- The vertical gradient of precipitation is still a poorly understood phenomenon (e.g., Barry, 2012, Rangwala and Miller, 2012). Several studies in the literature reported increasing precipitation with altitude. However, some works found a precipitation increase up to the highest elevations, while others indicated a limiting elevation above which precipitation amounts saturate. Please at least mention this issue somewhere in the paper (e.g., Section 2), since the elevational precipitation gradient is discussed, for example at page 1294. At the same time, I found Section 2 a bit too long and maybe repetitive in some parts.

We included the discussion of Barry (2012) into our manuscript.

2. Specific comments

Page 1281, Line 22: Please add in this section the following reference (Filippi et al., 2014, see below), since it discusses another important aspect of the role of western weather patterns in the western stretches of the area and associated precipitation, i.e., the mechanisms that can explain the well-known relationship between the NAO teleconnection pattern and winter precipitation in that region.

The Reference and a short sentence concerning the importance of extratropical atmospheric modes has been included in the revised version.

- Page 1285, lines 7-11: I do not understand very well the two sentences reported at those lines and their causal relationship. To my knowledge, estimates of precipitation associated with the ERA-Interim reanalysis are produced by the forecast model, based on temperature and humidity information derived from the assimilated observations. That is, there is no direct assimilation of precipitation observations in the reanalysis system. Please explain better/rephrase the sentence in order to avoid confusion. The reanalysis system does assimilate the large scale fields indeed, which is good for the approach employed in this study. Again, specify better.

That particular sentence didn't make much sense and thus has been deleted. For sure precipitation is not assimilated, thus there are no problems arising. Thanks for clarifying.

- Page 1285, lines 14-21: The authors use the ERA-Interim monthly mean fields of humidity at two atmospheric levels (500 and 200 hPa) and of wind at 500 hPa. These fields are resampled (interpolated) at 1 km spatial resolution, much finer than the original resolution of the reanalysis. To what extent does the interpolation introduce error/uncertainties in the procedure? Has this issue been addressed/taken into account by the authors?

Since the approach only utilizes free atmospheric variables, which actually should be „smooth“ we do not see any problems with the interpolation procedure. Subgrid (mesoscale) processes are not represented by ERA-Interim, but the large scale characteristics are well captured by the resampled fields. We specified that in our revised manuscript.

- Page 1286, Line 7 (the same: Page 1290 line 27; Page 1302 line 17): While its meaning is clear, I don't like very much the expression “precipitation-genetic”

The term has been substituted

3. Figures

Figure 8: Could you please specify the difference between left and right panels in the figure caption?

The Figure has been completely revised. The figure caption has been changed in order to better explain the diagrams.