Responses to Referee comments on manuscript ESDD-5-1-2014

General author response/commentary:

We thank the reviewers for their constructive and insightful commentary which will help us to improve our final manuscript.

i) Response to Referee M. Disse [5, C580-C581, 2014]

[i.C1] (p. 1120 / line 6-11) *"…What is the benefit compared to classical bias correction?"*

[i.R1] The assertion that there will be a benefit to bias correction, or more precisely, bias characterisation relates to the aggregation of biases calculated for individual grid cells or point locations into more general assessments of gridded dataset performance. Our contention is that the defined climate zones are rational, object geographic zones (spatial units) within which one ought to be able to expect consistent gridded dataset performance. The converse assertion is that aggregating bias statistics which lump climatically dissimilar zones could lead to masking of biases, e.g. averaging cool biases over deserts with warm biases over mountains to conclude near neutral mean bias.

[i.C2] (p. 1122, line 9-12) *"Please explain how the objective climate classification will improve the inter-comparison of gridded dataset performance in the sub-regions. Which new approach will be provided?"*

[i.R2] This answer is linked to the previous response (i.R1). We assert that our findings demonstrate that an objective basis is required to group individual observation sites into subsets for coherent, consistent characterisation of the biases of gridded datasets. As an illustration, if stations from across the South Asian or Central Asian domains were lumped as a single set, important differences in performance between climate zones, e.g. the Central Asian deserts and the Ganges-Brahmaputra Delta.

[i.C3] (figures and tables) *"In order to get a quick understanding of the figures and tables, I would suggest to explain all abbreviations of the respective figures / tables."*

[i.R3] We will make the following changes to tables:

Table 2. Variables used for Himalayan region climate classification.

Variable	Season	Physical importance
Precipitation	Annual Total	Humid vs arid climates
	ONDJFM ("rabi")	Westerly (extra-tropical) weather system climate influence
	AMJJAS ("kharif")	Monsoonal weather system climate influence
T_{avg} daily mean near surface air temperature	DJF	Indicator of precipitation state (solid versus liquid) and available energy to drive hydrological processes (meltwater generation) and crop growth (transpiration); as such indicator of hydrological regime (pluvial, nival or glacial)
	MAM	
	JJA	
DTR diurnal temperature range	DJF	(inverse) Indicator of moisture conditions, i.e. relative humidity and cloud cover, as both suppress DTR; as such proxy for cloud cover further informs regarding circulation influences
	MAM	
	JJA	
SW_{net} net downward shortwave radiation at the surface	DJF	Indicator of land surface state (snow covered or bare) and available energy to drive hydrological processes (meltwater generation) and crop growth (transpiration) ; as such indicator of hydrological regime (pluvial, nival or glacial)
	MAM	
	JJA	

Table 5. Variability of primary Himalayan region climate zones (8 clusters) area in the Hadley Centre downscaled perturbed physics ensemble, Regionally Quantify Uncertainty in Model Predictions (RQUMP), for South Asia.

We intend to modify the figure captions to read as follows:
Figure 1. Geographic context of the – Himalayan arc and adjacent plains – study area including elevation and areas with > 33% under irrigation (hashed). Data sources include the United Nations Food and Agriculture Organisation (FAO) and the United States Geological Survey Global 30 Arc-Second Digital Elevation Model (GTOPO30).

Figure 2. Ensemble precipitation climatology and normalised comparison of individual contributions from reanalyses used in this study. ONDJFM is the abbreviation for the period from October to March, referred to regionally as "Rabi." AMMJJAS is the abbreviation for the period from April to September, referred to regionally as "Kharif."

Figure 3. Ensemble energy input (temperature and radiation) climatology and normalised comparison of individual contributions from reanalyses used in this study. SW_{net} is net downward shortwave radiation at the surface. T_{avg} is daily mean near surface air temperature. DTR is diurnal temperature range. DJF is the (Winter) period December through February. MAM is the (Spring) period March through May. JJA is the (Summer) period June through August.

Figure 4. Comparison of the first three principal components (PCs) from each of the reanalyses used in this study. PCs are calculated from the Principal Component Analysis (PCA) input standardised variables using the PCA output weighting factors. PCs are thus dimensionless and values are expressed in standard deviations.

Figure 5. Comparison of climate classifications resulting from the use of 8, 12 and 16 clusters (k) on principal components from the individual reanalyses. Large units in the legend refer to zones for the k=8 case.

Figure 6. Ensemble spatial statistics for annual cycles of precipitation (left) and DTR (right) by climate zone (8 clusters). DTR is diurnal temperature range.

Figure 7. Ensemble spatial statistics for annual cycles of Tavg and SWnet by climate zone (8 clusters). SW_{net} is net downward shortwave radiation at the surface. T_{avg} is daily mean near surface air temperature.

Figure 8. Comparison of climate classifications resulting from the use of 8 clusters on principal components of the control period (1970 to 1999) from the individual members of the Hadley Centre RQUMP perturbed physics ensemble downscaled over South Asia.

[i.C4] (Figure 4) *"The unit of the legend is not clear (-5 to +5)."*
[i.R4] As also stated in the previous response, we are adding the following text to the figure caption: PCs are calculated from the Principal Component Analysis (PCA) input standardised variables using the PCA output weighting factors. PCs are thus dimensionless and values are expressed in standard deviations.

[i.C5] (Figure 6 and 7) *"The quality of the figures should be improved. Is the printing resolution sufficient to distinguish the differences? Can you add some (general) statistics in order to evaluate the differences of the four reanalysis datasets (e.g. test statistics for common mean value)"*
[i.R5] With regards to the figure quality, we have review the files we supplied for article submission, and the problem appears to have occurred during generation of the pdf file as the original figure was of far greater

resolution. In any case we will coordinate with editor and journal production staff to ensure that the graphics are both of the highest possible quality and meet the standards of the publisher.
With regards to evaluation of the differences between the reanalyses, we consider the evaluation of differences between individual reanalyses to be beyond the scope of the present work. We are of this opinion because central to the interpretation reanalysis data in general is understanding that both data assimilation and forecasting models used to generate the datasets and the fixing of mean surface elevations at coarse spatial resolution introduce biases. We consider that it would be of greater interest to compare statistics between derived climate zones from a single reanalysis although in our opinion step would also be out of scope. If the editor, however, is of the opinion that this would be an essential addition to the paper, we are willing to generate the additional table(s) and add accompanying text to the table. We will await specific instruction from the editor before make these changes.

ii) Response to Referee J. Böhner [5, C587-C590, 2014]

[ii.C0] (General comments)

[ii.R0] We appreciate the Referee's comprehensive assessment of our work. We recognise that definitive substantiation of our assertions regarding the importance of the climate classification scheme would require a much greater volume of work (e.g. crop yield impact assessments based climate change projections) to deliver the *"novel scientific insights"* required as underpinning.

[ii.C1] (page 1107, lines 1-3) *"… What I miss is the method …a major alternative would have been to stay closer to the original horizontal discretization (e.g 0.5 x 0.5 Degree Lat./Long.). A more sophisticated altitude-adjustment is likewise an increasingly established method …"*

[ii.R1a] To clarify, the "method" was to simply subdivide the "native resolution" grid cells of each reanalysis into 0.25 degree resolution cells while preserving the numerical values of the "parent cell". This approach was chosen, with the aim of minimising interpolation, because 0.25 decimal degrees is literally the common denominator of the differing spatial resolutions of the reanalyses (0.5, 0.75, 1.25 decimal degrees). In terms of software tools, subdivision was performed with the "gdalwarp" command-line tool (gdal.org). This was the case with the exception of the longitudinal spacing (0.667 decimal degrees) of NASA MERRA where gdalwarp applied bilinear sampling to perform the interpolation. To restate, the selection of 0.25 decimal degree resolution for intercomparison was selected to minimise interpolation and thus preserve the reanalysis values as distributed by their producers. The utilisation of a 0.5x0.5 degree grid would have imposed substantial interpolation for all reanalyses with the exception of NCEP CFSR.
[ii.R1b]The approach of "altitude-adjustment" could have been applied based on a much substantially finer spatial grid, e.g using the GTOPO30 DEM, but this would have introduced a step change in complexity and entailed assumptions about whether additional parameters beyond elevation were required to "downscale" the reanalysis at the sub-grid level. Selection of these additional parameters might vary in validity between individual reanalyses. We agree that within relatively limited spatial domains temperature can be translated based "lapse rates" although other factors, such as "terrain aspect" (direction of exposition), can also play a role. The downscaling of precipitation and radiative fluxes would be more complex with orgraphic barriers (topographic wind-shadows) and elevation-influenced cloud cover playing a role. Further exploration of these issues would be potentially interesting work, but we consider it beyond the scope of the present study. At (pending) the editor's instruction we are willing develop these themes further in (a new subsection) of the "Discussion" portion of the manuscript.
[ii.R1c] With regards to the suggestion of adding a "comparative orography figure" analogous to the comparative climatologies in Figures 2 and 3, we had considered this and could develop it partially with confidence. We are satisfied with the clarity with which "invariant orography" is distributed/published by ECMWF and NASA for ERA-Interim and NASA MERRA respectively. For JRA-55 and NCEP CFSR the orographic fields/variables we have been able to identify appear, in contrast, to be time-varying and thus we

hesitate to attempt to portray the orographic heights utilised these two reanalyses. This is a further reason we did not attempt an “altitude adjustment.”
[ii.R2d] With regards to the assertion of “advantages of higher resolution (of NCEP-CFSR)”, while our initial assumption too was that this would be an “edge” for this reanalysis dataset over its counterparts, our work both in this study and in others (manuscripts in preparation for other journals) in fact indicate at best equivalent and more often inferior performance by CFSR over this geographic domain particularly when compared to ERA-Interim.

[ii.C2] (page 1104, lines 24-26 and page 1105, lines 1-7) *“…it would have been more appropriate to use consistently 6h values, available for all reanalyses, and to depict the advantages of a higher temporal resolution in one sentence …”*
[ii.R2] We found it more appropriate to present the each dataset as it is in order to illustrate its strengths and limitations. More particularly, since we were performing principal components analysis (PCA) and clustering within individual reanalyses, identifying differences in climate classification which could arise from varying time-steps was an element of interest. We assert that the findings of similar DTR (magnitude) loadings between ERA-Interim (6h timestep) and NASA-MERRA (hourly time-step) -- in composition of the third principle component (PC3) shown in Table 3 of the discussion paper – demonstrate the “information content” of DTR as a variable is preserved in the coarser time-resolution datasets.

[ii.C3] (page 1121, lines 4-25) *" ... A comparison of classification results for different time-slices would be more appropriate to illustrate its added values but is still in progress. Hence, I suggest to shorten and move this aspect into the conclusion section as outlook."*
[ii.R3] We agree that the primary value of utilisation of the ensemble RCM outputs is an eventual time-slice comparison. We maintain that the present manuscript configuration is preferable because: a) an essential step in establishing the validity of RCM outputs is comparison of “control climate” results to climate conditions described by datasets constrained by observations (in this case reanalyses) and thus the comparison of the ensemble RCM climate classification to reanalysis-derived zonings is crucial and worthy of detailed examination; and b) for practical reasons of length, inclusion of the future RCM time-slides would lead to an “unwieldly” manuscript. We are willing to revisit manuscript structure at the editor’s instruction.

[ii.C4] (page 1137, Figure 4 and page 1130, Table 3) *“I’m a bit surprised that PC2 was dominated by precipitation inputs (Table 3) whilst in Figure 4 the shape of the Tibetan Plateau is quiet clearly represented in all four reanalyses ... however, please check.”*
[ii.R4] We have confirmed, the reason the Himalaya Arc/Tibetan plateau shape is evident in the geographic distribution of PC2 is because of orographically-forced precipitation not due to air temperature. As can be seen in Figure 4, there is a “doughnut hole” (of varying size) in PC2 in each of the reanalyses over the arid central plateau area. This relative aridity, along with that over the Central Asian deserts and the Indus Valley, can be seen in the ensemble mean climatologies in Figure 2.

[ii.TC1] (page 1104, line 26 – page 1105, line 1) *“In all cases daily means were calculated as the mean of the available sub-daily time-steps.”* ***(page 1106, line 6-7)*** *“Hence, Tavg (mean temperature) and DTR – both calculated from tmax (maximum temperature) and Tmin (minimum temperature) [....].” The averaging methods are contradicting, please check.”*
[ii.TR1] We recognise that this is unclear both statements are correct. The first refers to the methodology of the study. The second refers to Tavg and DTR reported from observing stations. We will amend the latter statement (p.1106) as follows: “Hence Tavg and DTR, which together describe the diurnal temperature cycle and can be calculated at stations recording solely Tmax and Tmin -- …”

[ii.TC2] (page 1137, Figure 4) *“Please add an information about the units in the legend.”*

[ii.TR2] We will amend the figure caption as follows: “Figure 4. Comparison of the first three principal components (PCs) from each of the reanalyses used in this study. PCs are calculated from the Principal Component Analysis (PCA) input standardised variables using the PCA output weighting factors. PCs are thus dimensionless and values are expressed in standard deviations.