RESEARCH ARTICLE

Key Points:
- The ability of temperature, precipitation, and topography to explain spatial variability of δ^{18}O over Himalaya-Tibet is quantified
- Discrepancies between model- and observation-based investigation of the relationship of δ^{18}O with temperature, precipitation, and topography

Precipitation δ^{18}O over the Himalaya-Tibet orogen from ECHAM5-wiso simulations: Statistical analysis of temperature, topography and precipitation

Sebastian G. Mutz[1], Todd A. Ehlers[1], Jingmin Li[2], Christian Steger[2], Heiko Paeth[2], Martin Werner[3], and Christopher J. Poulsen[4]

[1]Department of Geosciences, University of Tuebingen, Tuebingen, Germany, [2]Institute of Geography, University of Wuerzburg, Wuerzburg, Germany, [3]Alfred Wegener Institute, Bremerhaven, Germany, [4]Earth and Environmental Sciences, Univeristy of Michigan, Ann Arbor, Michigan, USA

Abstract Variations in oxygen isotope compositions (δ^{18}O) provide insight into modern climate and past changes in climate and topography. In addition, in regions such as Tibet, geologic archives of isotope ratios record climate change driven by plateau uplift and therefore also provide information about the surface uplift history. A good understanding of modern-day controls on δ^{18}O is crucial for interpreting geologic δ^{18}O in this context. We use the ECHAM5-wiso global atmospheric general circulation model to calculate δ^{18}O in precipitation (δ^{18}O$_p$) for the present-day climate. In the region of the Tibetan Plateau, spatial variations of monthly means of δ^{18}O$_p$ are statistically related to spatial variations of 2 m air temperature and precipitation rate, as well as to topography. The size and location of investigated regions are varied in our study to capture regional differences in these relationships and the processes governing the modern δ^{18}O$_p$. In addition to correlation analyses, a cross-validated stepwise multiple regression is carried out using δ^{18}O$_p$ as the predictand, and topography and atmospheric variables (temperature and precipitation amount) as predictors. The 2 m air temperature and topography yield the highest spatial correlation coefficients of >0.9 and <-0.9, respectively, throughout most of the year. Particularly high correlation coefficients are calculated for the region along the Himalayan orogeny and parts of western China. The predictors explain $>90\%$ of the δ^{18}O$_p$ spatial variance in the same regions. The 2 m air temperature is the dominant predictor and contributes 93.6% to the total explained spatial variance on average. The results demonstrate that most of the δ^{18}O$_p$ pattern on and around the Tibetan Plateau can be explained by variation in 2 m air temperature and altitude. Correlation of the dependent predictors indicate that in low-altitude regions where topography does not determine temperature variability, the high correlation of temperature and δ^{18}O$_p$ may partially be explained by variations in precipitation rates.

1. Introduction

The Tibetan Plateau is the largest tectonically active orogen and plays a crucial role in regional and global climate. Its deformation history, the evolution of central Asian climate, and the interaction between orogen structures and climate in general remain a focus of research. This interaction takes place on different timescales. On geological timescales, the influence of climate on tectonics can largely be attributed to erosional processes [e.g., *Montgomery et al.*, 2001; *Willett et al.*, 2006; *Reiners et al.*, 2003], whereas on shorter timescales, mountain building can affect regional climate through the direct physical influence of topography [e.g., *Takahashi and Battisti*, 2007a, 2007b; *Insel et al.*, 2009; *Kutzbach et al.*, 1989; *Kutzbach et al.*, 1993; *Ruddiman and Kutzbach*, 1989].

The Tibetan Plateau in particular heavily influences regional and global climate due to its effect as a physical obstacle [*Raymo and Ruddiman*, 1992; *Kutzbach et al.*, 1993; *Thomas*, 1997; *Bohner*, 2006; *Molnar et al.*, 2010; *Boos and Kuang*, 2010]. The impact of tectonic uplift on climate has previously been assessed from observational, geological data sets [e.g., *Zhisheng et al.*, 2001; *Dettman et al.*, 2003] and climate models [e.g., *Kutzbach et al.*, 1993; *Bohner*, 2006; *Takahashi and Battisti*, 2007a, 2007b; *Ehlers and Poulsen*, 2009; *Insel et al.*, 2009]. However, only a few modeling studies investigated the effect of the Tibetan Plateau on climate [e.g., *Kutzbach et al.*, 1989; *Kutzbach et al.*, 1993; *Ruddiman and Kutzbach*, 1989; *Boos and Kuang*, 2010]. These point out the strong sensitivity of regional and global climate to changes in topography.

The modification of temperature and precipitation patterns resulting from Tibetan Plateau uplift ought to be reflected in records of precipitation δ^{18}O values as a consequence of changing temperatures and the amount

effect [*Dansgaard*, 1964; *Rozanski et al.*, 1993; *Gat*, 1996]. Studies of modern-day controls of the distribution of precipitation $\delta^{18}O$ across the Tibetan Plateau highlight the importance of precipitation rates associated with monsoon circulation [*Araguas-Araguas et al.*, 1998; *Tian et al.*, 2003]. Depending on the season and region, these can overshadow the dependence of precipitation $\delta^{18}O$ values on temperature [*Araguas-Araguas et al.*, 1998; *Tian et al.*, 2001a, 2001b, 2003; *Liu et al.*, 2008; *Yang et al.*, 2011].

Geological archives of $\delta^{18}O$ values can therefore give insight into Central Asian paleoclimate and the uplift history of the Tibetan Plateau [*Dettman et al.*, 2003]. However, interpreting these records requires a good understanding of the relationship between modern-day topography and climate with precipitation $\delta^{18}O$ and their geographic distribution. Several studies contributed to this understanding both on a more local scale based on station measurements [e.g., *Yang et al.*, 2011] and on a larger regional scale with the help of atmospheric models [e.g., *Yao et al.*, 2013]. The inverse relationship between $\delta^{18}O$ values and elevation is addressed in various studies [e.g., *Bershaw et al.*, 2012, *Liu et al.*, 2008; *Yao et al.*, 2013] as well as the latitudinal control on $\delta^{18}O$ values [e.g., *Liu et al.*, 2008]. *Tian et al.* [2001a] and *Yao et al.* [2013] also attempt to describe and identify the spatial distribution of $\delta^{18}O$ values and zones characterized by different climatic controls on it.

Observational data from the GNIP (Global Network of Isotopes in Precipitation) can provide the observational basis for studies in these regions [e.g., *Araguas-Araguas et al.*, 1998; *Tian et al.*, 2001a, 2001b; *Tian et al.*, 2003]. However, more than half of the stations in the GNIP database do not have oxygen isotope records covering more than 1 year [*International Atomic Energy Agency Water Resource Programme*, 1998; *Rozanski et al.*, 1993; *Bowen and Wilkinson*, 2002] and the spatial coverage of observed precipitation $\delta^{18}O$ in the region of interest is sparse.

Atmospheric general circulation models (AGCMs) equipped with an explicit diagnostic module for stable oxygen and hydrogen isotopes in water allow the modeling of the isotopic composition for the precipitation calculated by the AGCM [e.g., *Joussaume et al.*, 1984; *Jouzel et al.*, 1987; *Hoffmann et al.*, 1998] and produce a homogeneous set of climate and isotope values with the resolution used for the model simulation. The limited availability of suitable records, and the success of previous modeled results of precipitation $\delta^{18}O$ values [e.g., *Sturm et al.*, 2010, and references herein], encourage a modeling approach to investigate the influence of topography and climate on precipitation $\delta^{18}O$ value distribution.

This study complements previous work by contributing to a better quantitative understanding of the local climatic and topographic controls on the geographical distribution of precipitation $\delta^{18}O$ values with the aid of a model simulation of climate and oxygen isotopes in precipitation and a topography data set. Furthermore, it aims to highlight possible limitations for the use of model- and observation-based analyses investigating these controls. In particular, the following points are addressed: (1) the spatial correlation between topography, 2 m air temperature, and precipitation amount with precipitation $\delta^{18}O$; (2) the relationship of 2 m air temperature with topography and precipitation in the context of point (1); (3) how much of the spatial precipitation $\delta^{18}O$ variability is explained by the variability in topography and a chosen set of atmospheric variables; (4) the regional variations and causes in the $\delta^{18}O$ variance; and (5) the differences between results (1–4) using model output and observational and reanalysis data sets.

2. Data and Methods

2.1. Observation-Based Data Sets

The GTOPO30 data set, a product available from the U.S. Geological Survey, was processed and interpolated to T63 resolution. The resulting interpolated data set was used in the correlation and multiple regression analyses (section 2.4) and as one of the input variables for the ECHAM5-wiso simulation (section 2.2). Results from the GTOPO30 data set were cross-checked with the T63 topography from ECHAM5 boundary fields to ensure that they are identical. All topography-related variables that served as additional boundary conditions for the implemented ECHAM5-wiso simulation, such as orographic slope, orographic standard deviation, and orographic angle, were also derived from this data set.

Data sets from the European Centre for Medium-Range Weather Forecasts (ECMWF) reanalyses (ERA40) [*Uppala et al.*, 2005] and from the National Centers for Environmental Prediction and National Center for Atmospheric Research (NCEP/NCAR) reanalyses [*Kalnay et al.*, 1996; *Kistler et al.*, 2001] serve as observation-based climate data in this study. The reanalyses include data recovered by various measurement systems

and are dynamically interpolated by models to produce more homogeneous data sets. While these are not exclusively observation-based, the analyses conducted for this study requires this homogeneity of data. Additionally, the latest version (TS3.21) of the CRU (climate research unit) gridded climate data set [*Harris et al.*, 2013] is used for further validation of model output. CRU TS3.21 is derived from interpolation of values of land-based meteorological stations [*Harris et al.*, 2013]. For the greatest temporal model-data set overlap, long-term seasonal means of 2 m air temperature and precipitation are constructed from these data sets and compared to long-term seasonal means of modeled temperature and precipitation

2.2. ECHAM5-wiso Simulation

ECHAM5 is an atmospheric general circulation model [*Roeckner et al.*, 2003], developed at the Max Planck Institute for Meteorology (MPIM) and is based on the spectral weather forecast model of the ECMWF [*Simmons et al.*, 1989]. The isotope module in ECHAM5-wiso implements the calculation of oxygen and hydrogen isotopes in water as part of the hydrological computations in ECHAM and adds isotopic counterparts to the water variables in ECHAM [*Hoffmann et al.*, 1998; *Werner et al.*, 2011]. Since the third major version of ECHAM, the physics of fractionation of oxygen and hydrogen isotopes in water was built into the model. The implementation of isotopes in ECHAM5-wiso is done in the following way: For each phase of the water variables modeled by ECHAM5 (solid, liquid, and vapor), an isotopic counterpart is implemented. The water isotopologues, $H_2^{18}O$ and $HD^{16}O$, are treated as separate forms of water that are transported and transformed in parallel to the water calculated by ECHAM5. The transport scheme for both is semi-Lagrangian advection according to *Lin and Rood* [1996]. The isotope fractionation processes considered are divided into equilibrium processes, in which phase changes are slow enough to allow isotopic equilibrium, and nonequilibrium processes, which are a function of phase change velocity as well as kinetic fractionation processes [*Hoffmann et al.*, 1998].

The ECHAM5-wiso model simulation was run at a T63 spatial resolution (horizontal grid size $1.8° \times 1.8°$) and with 19 vertical levels (between surface and 10 hPa) for present-day boundary conditions. It was run for 43 model years in order to ensure a sufficiently large number of years for analysis, i.e., the standard climatological reference period of 30 years. This leaves a more than sufficient model spin-up time before analysis of the model output. The experiments are Atmospheric Model Intercomparison Project (AMIP)-style experiments with sea surface temperatures and sea ice concentrations annually varying from 1957 to 2000 and greenhouse gas concentrations for the same period.

2.3. Preparation of Variables

Monthly long-term averages were used in the analyses. These were obtained by calculating the means over the last 30 years of the model output for each month separately, thereby capturing the mean climate while keeping information on intermonthly variability. The $\delta^{18}O$ values were calculated from model convective- and large-scale precipitation and represents the total precipitation $\delta^{18}O$. The known factors controlling $\delta^{18}O$ in precipitation [*Dansgaard*, 1964; *Gat*, 1996] determined the choice of climate variables to be included in the analyses. Precipitation, 2 m air temperature, and model topography were chosen to represent controls of the amount, temperature, and altitude effect on precipitation $\delta^{18}O$. The same means were constructed from temporally corresponding NCEP and ERA40 data. Since model output and reanalyses are spatially homogeneous and the investigated region does not cover higher latitudes, unweighted means are used.

2.4. Correlation and Multiple Regression

All correlation and multiple regression procedures in this study analyze spatial variability rather than temporal variability. Thus, it was possible to incorporate topography in the same way that climate variables were used in the methods described in this section. Although the first-order control of 2 m air temperature is topography, especially in high-altitude regions, we include topography as an independent variable to identify discrepancies between the relationships of temperature and topography with oxygen isotope values. This allows us to identify where topographic variability correlated with isotope ratios is not reflected in temperature variability and where the correlation of temperature with $\delta^{18}O$ values may not be attributed to topography but to another variable that covaries with temperature where topography does not. It therefore helps better understand the underlying causes for the relationships between $\delta^{18}O$ and the predictors. The multiple regression procedure used in this study implements measures to avoid overfitting due to the resulting multicollinearity (see description of method).

The region was subdivided into model grid box areas, and Pearson correlation coefficients and multiple regression analyses were computed for each of those and each month separately. To ensure a sufficiently large sample size and assess the sensitivity of sample size, the analyses were carried out for model grid box areas of size 5×5, 7×7, and 9×9. The calculated values were assigned to the grid box at the center of each area. Each variable is correlated against the precipitation $\delta^{18}O$ separately. The significance level of the resulting Pearson correlation coefficient is determined by t tests.

The multiple correlation analysis allows the quantitative assessment of the regression relationship between a number of model variables and precipitation $\delta^{18}O$ values, referred to as the predictors and the predictands, respectively. It is also a means to get a quantitative estimate of how much of the variability in the precipitation $\delta^{18}O$ can be explained by the variability of each of the selected predictors [*von Storch and Zwiers*, 1999].

The applied multiple regression procedure is a cross-validated procedure, involving a stepwise multiple regression with random bootstrapping [*Paeth and Hense*, 2003; *Paeth et al.*, 2006; *Mutz et al.*, 2015]. The multiple regression was adjusted to the problems addressed in this study and carried out as follows: The included predictors are sorted by their importance, which is determined by linear regression analyses between the n-dimensional predictand, where n is the number of total grid boxes in the investigated area (25, 49, and 81 for grid box areas 5×5, 7×7, and 9×9, respectively), and the predictor series $\vec{x}_1, \ldots, \vec{x}_k$, where k is the number of predictors used in the multiple regression, in this case 3 (temperature, precipitation, and topography). The most important predictor is selected for the regression analysis, and the residual $\vec{\varepsilon}$ is calculated. The residual is then used in the regression analysis with the subsequent predictor, and a new residual is determined. This method is continued with all predictors included in the multiple regression model. For cross validation [*Michaelsen*, 1987], a reasonable number of random values are retained before each regression analysis to leave sufficient values for training the algorithm. The regression model is trained with the dependent data, i.e., the data left after the removal of the randomly selected bootstrap values. The model is then applied to the retained data, and the RMSE (root-mean-square error) between the independent predictand values and regression model estimates is calculated. Once the RMSE can no longer be reduced by the addition of another predictor, the currently investigated predictor is not included in the final model and the multiple regression is terminated. This results in an exclusion of predictors that do not provide any additional information about the independent data. The whole process is repeated 1000 times. Each time, a different set of bootstrap values is retained and the resulting regression equation is saved. The coefficients are averaged over all iterations for the final model. To ensure the robustness of the model, a filter for predictors is introduced before the computation of the final multiple regression. Only predictors that pass through the filter in at least 50% of all iterations are included.

The correlation and multiple regression analyses are conducted first using the prepared model output and interpolated GTOPO30 topography data set. The procedures are then repeated with the prepared NCEP variables, an interpolated observation-based $\delta^{18}O$ data set [*Bowen and Revenaugh*, 2003], and the same topography data.

2.5. Cluster Analysis

The same analyses as described in the previous section were carried out for different isotope zones on the Tibetan Plateau. The Tibetan Plateau has previously been subdivided into three zones based on spatiotemporal patterns of precipitation $\delta^{18}O$ and their relationship to climatic controls [*Yao et al.*, 2013; *Li et al.* 2016]. In this study, we evaluate these findings and attempt to quantitatively constrain the geographical extent of the domains by subdividing the plateau using a cluster analysis and on the basis of precipitation $\delta^{18}O$ variability. Three variables describe $\delta^{18}O$ variability in the analysis: annual mean $\delta^{18}O$, the amplitude of annual $\delta^{18}O$ values, and $\delta^{18}O$ values during the monsoon season (July–September). They are used in a centroid clustering method with randomized regroupment [*Paeth*, 2004; *Bahrenberg et al.*, 1992], where the Mahalanobis distance is used as a measure for unit-independent statistical distance in multidimensional variable space. *Tian et al.* [2001a] and *Yao et al.* [2013] divide the Tibetan plateau into three distinct domains based on $\delta^{18}O$ variability and its relationship with temperature and precipitation. These domains represent regions influenced by different climate system components, namely, the Westerlies in the north and the Indian monsoon in the south. The central domain is a transitional zone. The cluster number of three is therefore explicitly prescribed based on these findings.
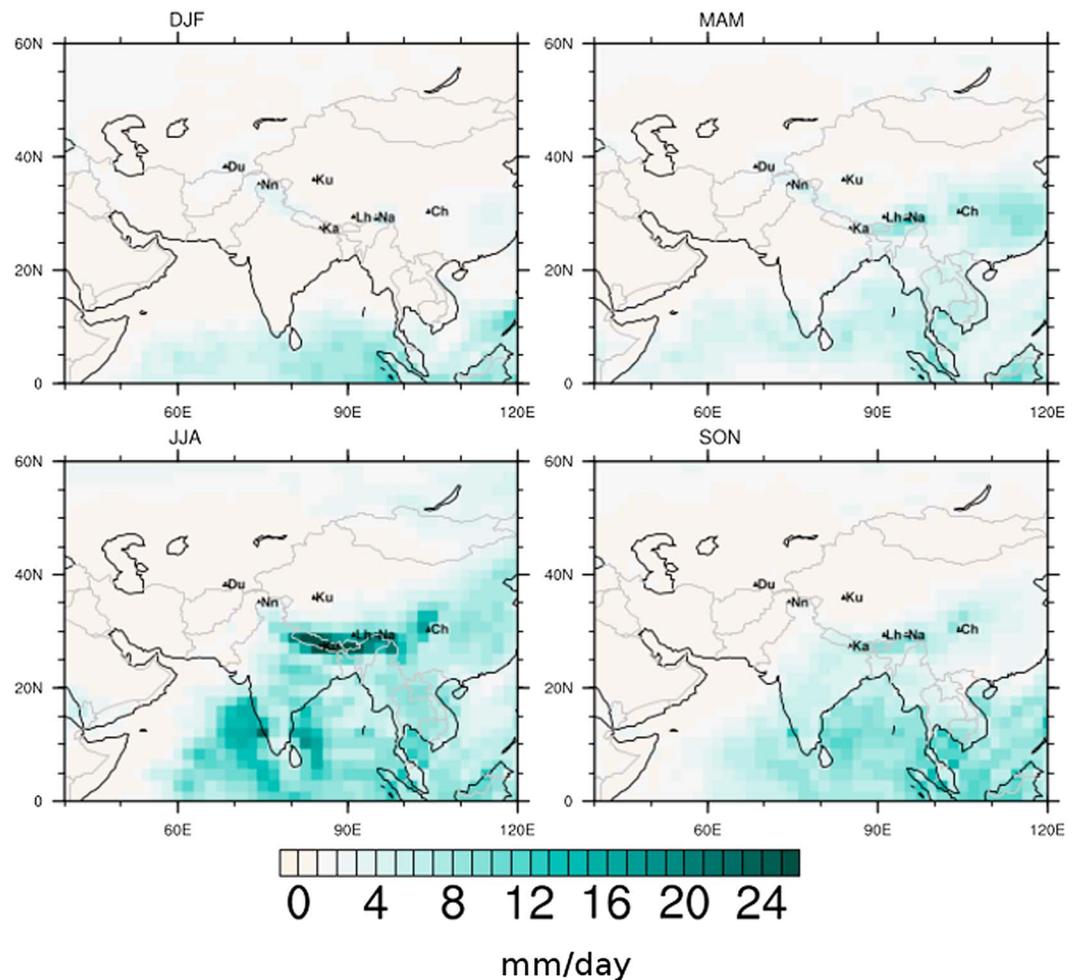
Figure 1. Long-term (30 year) means of ECHAM5-wiso total precipitation for seasons DJF (top left), MAM (top right), JJA (bottom left), and SON (bottom right).

The cluster analyses are first carried out using the prepared model output and then repeated with the observation-based δ^{18}O data set. The same method is later applied on the basis of the three sets of Pearson's R values calculated by correlating precipitation, 2 m temperature, and topography with δ^{18}O. This provides a method of clustering based on investigated relationships between δ^{18}O and its controls rather than δ^{18}O variability. In order to ensure that only values on the plateau are used, a topographic high pass filter (HPF) was applied to all values in the investigated region prior to the analysis. An HPF cutoff height of 2 km was used.

2.6. Model-Observation Comparisons

While this study focuses more on the comparison of the results obtained from model output and observation-based data (using the methods described in previous sections), direct comparisons of modeled and observed values are made in order to assess the feasibility of the long-term means of relevant model variables. The CRU TS3.21 data set [*Harris et al.*, 2013] and NCEP/NCAR [*Kalnay et al.*, 1996; *Kistler et al.*, 2001] and ERA40 [*Uppala et al.*, 2005] reanalyses products were used to validate simulated climate. The *Bowen and Revenaugh* [2003] data set was used to validate the simulated precipitation δ^{18}O. For these comparisons, values from the temporal overlaps of the reanalyses time series and analyzed model time period (1970–2000) were used. For brevity, only comparisons of ECHAM5-wiso and ERA40 precipitation and 2 m air temperature are shown here. The remaining direct comparisons of values are included as difference plots in the supporting information (Figures S1–S4). To allow the reader to put these differences in context of interobservational data set differences, plots of ERA40 and NCEP precipitation and 2 m air temperature are included. Comparison of the reanalysis products with CRU are included in the supporting information (Figures S5–S8).
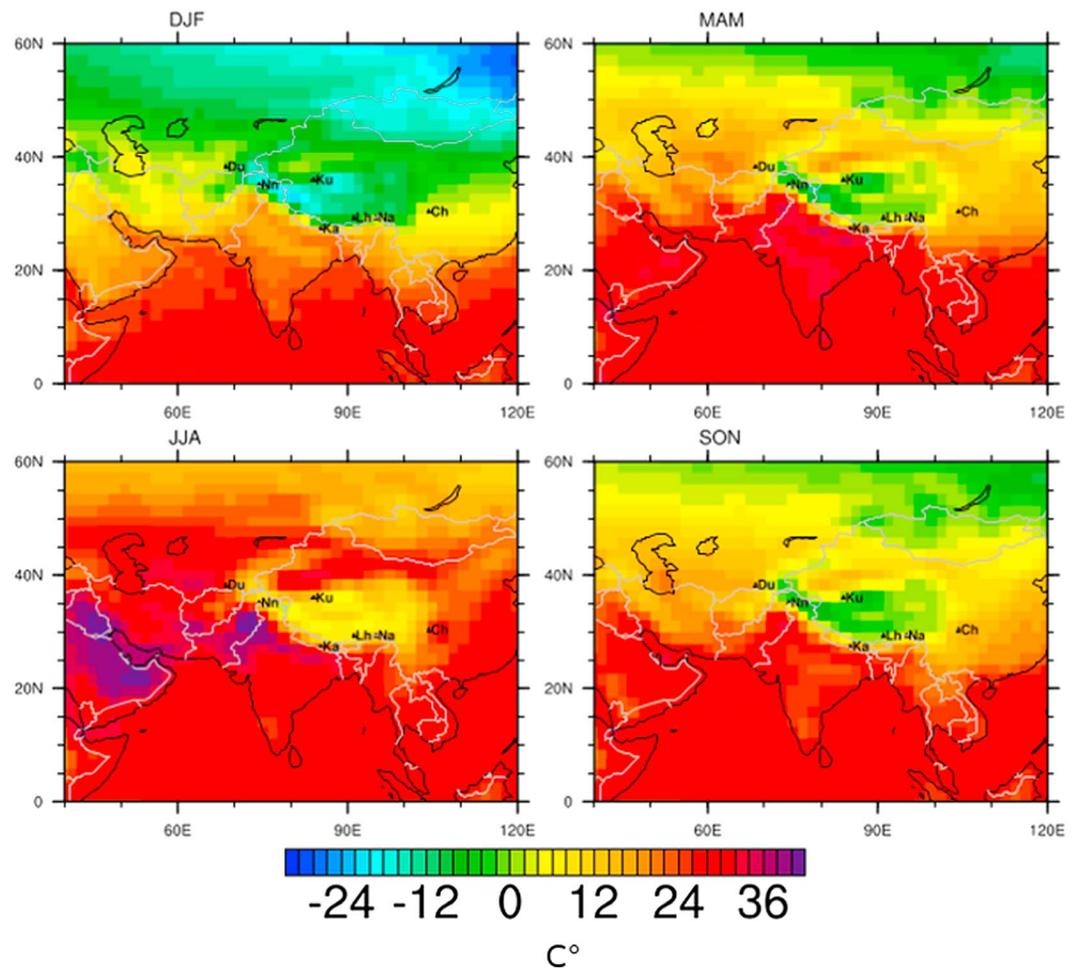
Figure 2. Long-term (30 year) means of ECHAM5-wiso 2 m air temperature for seasons DJF (top left), MAM (top right), JJA (bottom left), and SON (bottom right).

3. Results

In this section, 30 year long-term means of model output are shown. Furthermore, plots of the Pearson correlation coefficients from the correlation analyses between the chosen predictor variables and $\delta^{18}O$ values are shown. Results from the multiple regression analysis are also presented. In this study, the term "monsoon season" is used to describe the major summer rain period on the Indian subcontinent (aka Indian summer monsoon). Results are presented by seasons and divided into DJF (December, January, and February), MAM (March, April, and May), JJA (June, July, and August), and SON (September, October, and November). For better spatial orientation, the geographical locations of several towns and mountains are marked on the panel plots: Dushanbe (Du), Nanga Parbat (Nn), Kunlun mountains (Ku), Kathmandu (Ka), Lhasa (Lh), Namche Barwa (Na), and Chengdu (Ch). These locations were chosen to cover particular regions of interest. Dushanbe and Chengu lie to the west and east of the Tibetan Plateau, the Kunlun Mountains and Lhasa are locations on the northern and southern plateau, respectively, Nanga Parbat and Namche Barwa lie in the orogen syntaxes, and Kathmandu is located just south of the high plateau.

3.1. ECHAM5-wiso Simulation Results

The ECHAM5-wiso simulated 30 year seasonal means of 2 m temperature, total precipitation rate, and total precipitation $\delta^{18}O$ are presented in Figures 1–3. Results show that the total precipitation rate over central Asia typically stays below 5 mm/day outside the summer season (Figure 1). During the summer monsoon,
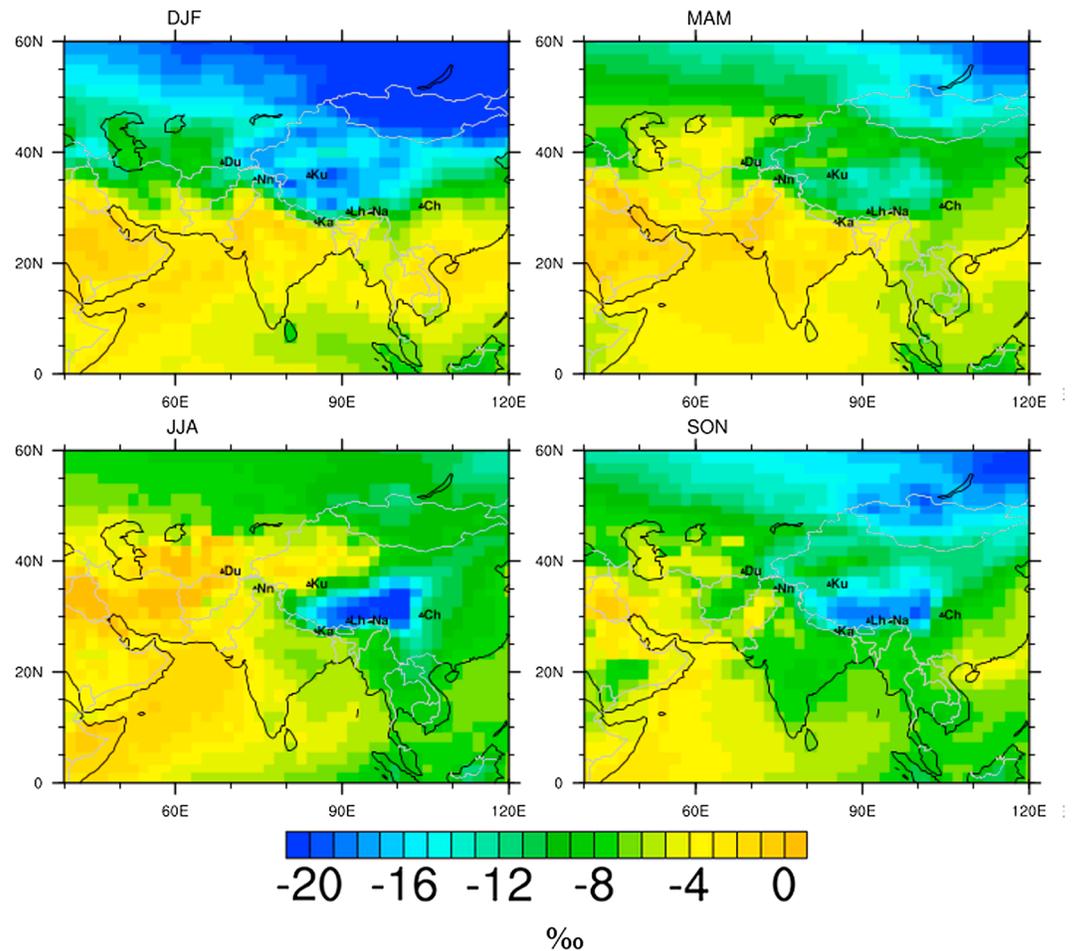
Figure 3. Long-term (30 year) means of ECHAM5-wiso total precipitation δ^{18}O for seasons DJF (top left), MAM (top right), JJA (bottom left), and SON (bottom right).

the southeastern Himalaya experiences anomalously high precipitation ≥ 25 mm/day. Modeled temperatures reproduce the latitude-related differences in temperatures of central Asia and the Indian subcontinent (Figure 2). The model output highlights constant low temperatures across the Tibetan Plateau relative to the temperatures of the surrounding regions that lie at approximately the same latitude but lower altitude. The temperature on the plateau itself varies between −20°C outside the monsoon seasons and up to 10°C during the monsoon season. Simulated precipitation δ^{18}O values show a similar altitude-specific geographic distribution (Figure 3). The typical range of −25 to 5‰ is covered in all seasons, although the average δ^{18}O values outside the monsoon season are lower. The values on the plateau are usually lower than the values of surrounding regions that lie on approximately the same latitude but at a lower altitude.

A comparison of long-term seasonal means of modeled and ERA40 precipitation show the largest discrepancies in the summer season JJA (Figure 4). ECHAM5-wiso overestimates precipitation along the eastern Himalayan orogen and to a lesser degree in low-altitude China. This overestimation persists when model results are compared to NCEP reanalyses and CRU (Figures S1 and S3 in the supporting information). A similar overprediction of precipitation in the eastern Himalaya, especially in summer, was also noticed in comparisons to the HAR30 (High Asia Refined analysis) data set [*Maussion et al.*, 2014] for a model simulation with a different temporal coverage that overlaps with that of the HAR time series. However, these comparisons are not shown here, because a different time period is covered by the simulations presented in this study and HAR is not yet suitable for calculations of long-term trends due to the short time period covered by the data set.
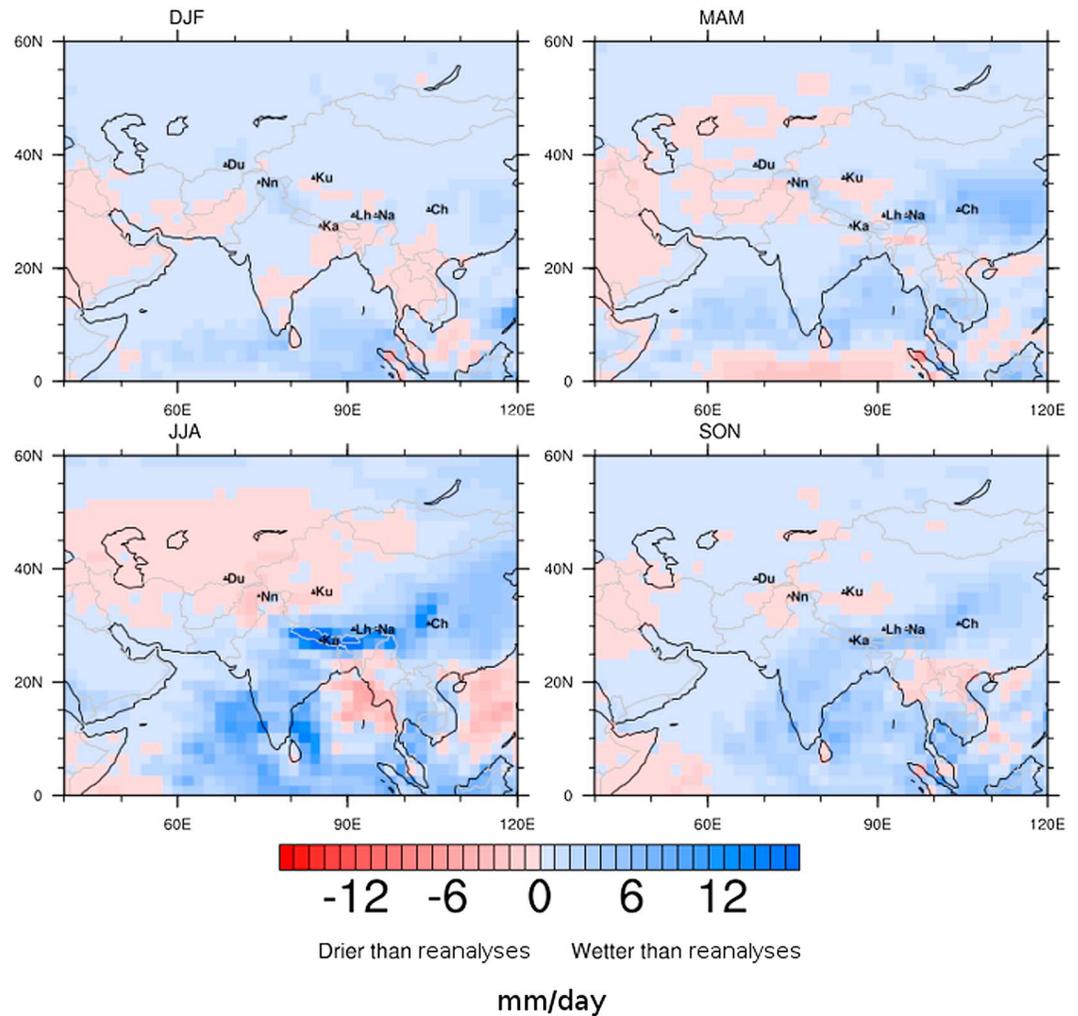
Figure 4. Difference between ECHAM5-wiso and ERA40 precipitation for seasons DJF (top left), MAM (top right), JJA (bottom left), and SON (bottom right). Long-term means of ERA40 precipitation were subtracted from ECHAM5-wiso total precipitation means for their temporal overlap (1970–2000).

A comparison of long-term seasonal means of modeled and ERA40 2 m air temperature reveals that the model overestimates temperatures in northwest Pakistan, especially in summer (Figure 5). This model bias also persists for comparisons to NCEP reanalysis and CRU data. Precipitation differences between the two reanalysis products (Figure 6) and NCEP and CRU data (supporting information) show discrepancies of a similar magnitude in low-altitude China and Southeast Asia during the summer season but no significant disagreement along the Himalayan orogen. On the Tibetan Plateau, 2 m air temperature differences between the reanalysis products (Figure 7) and NCEP and CRU data (see supporting information) are larger than differences between modeled and ERA40 temperature. The largest discrepancies between modeled and interpolated observation-based $\delta^{18}O$ values of the *Bowen and Revenaugh* [2003] data set exist on parts of the Tibetan Plateau covering the Kunlun Mountains (≤10‰), north of Lake Balkhash in Kazakhstan, and between Lhasa and Chengdu in summer (Figure 8).

3.2. Correlation Analysis

Pearson's R values were calculated by correlation analyses between each predictor variable and $\delta^{18}O$ values. Statistically significant Pearson's R values with $\alpha \leq 0.05$ are shown in Figures 9–11. Statistically insignificant R values are shown in white. Results are presented for the analyses using 7×7 grid box windows.

The statistically significant Pearson's R values attained in the correlation analysis between total precipitation and $\delta^{18}O$ values typically range from 0.4 to 0.6 and -0.2 to -0.8 (Figure 9). Anomalously high positive values
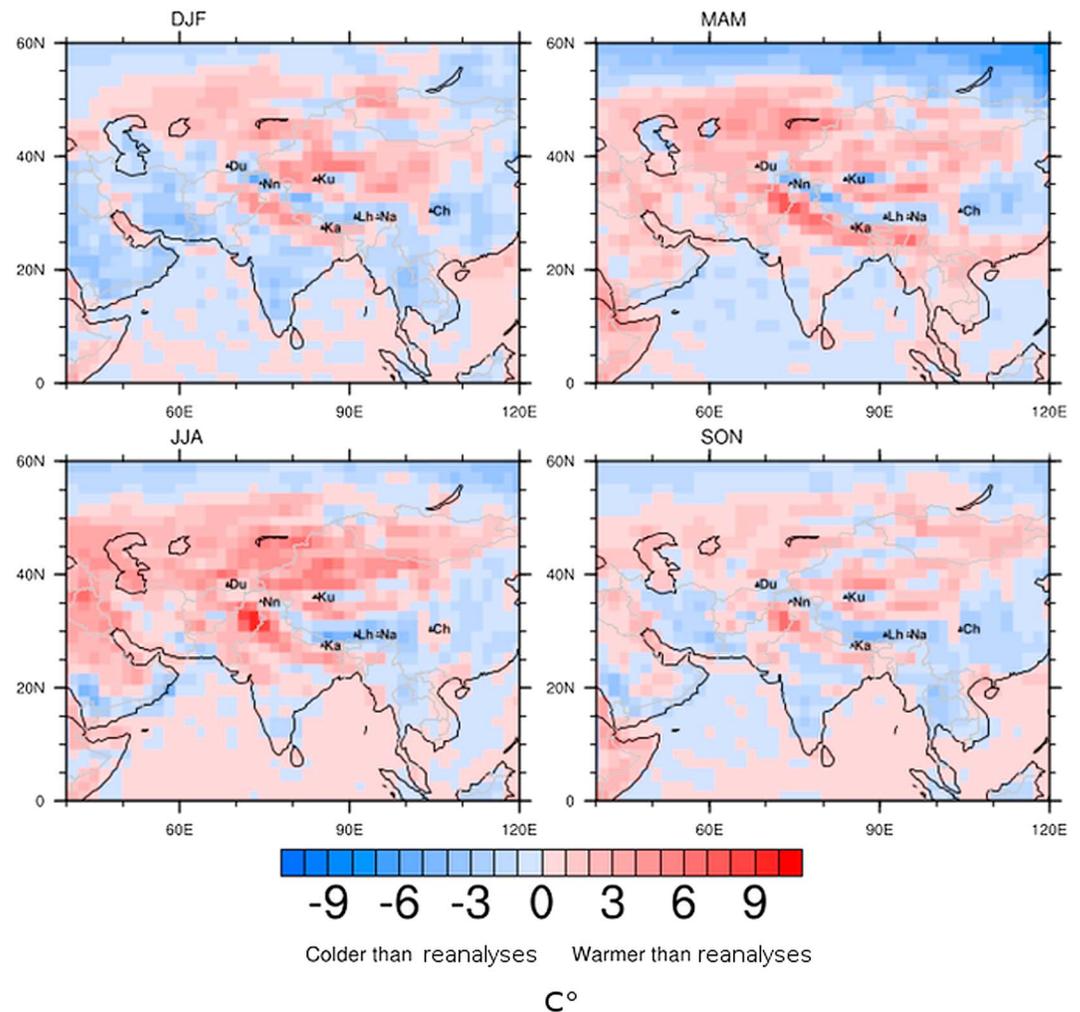
Figure 5. Difference between ECHAM5-wiso and ERA40 2 m air temperature for seasons DJF (top left), MAM (top right), JJA (bottom left), and SON (bottom right). Long-term means of ERA40 2 m air temperature were subtracted from ECHAM5-wiso 2 m air temperature means for their temporal overlap (1970–2000).

of over 0.8 are calculated for MAM in the north of the investigated region and northeast of Chengdu particularly outside the summer season. Anomalously high negative values of under −0.8 can be seen in the north of the investigated region from MAM to SON, around Dushanbe and north of it during MAM and SON, and in parts of northern India from DJF to JJA. During DJF and MAM, the most negative values extend to the northwest toward Nanga Parbat and Dushanbe. During the seasons MAM, JJA, and SON, the negative Pearson's R values are predominant. The area covering the easternmost Tibetan Plateau and the region just off the plateau, approximately between Lhasa and Chengdu, is characterized by positive Pearson's R values throughout the year. A second notable feature in the geographical distribution of Pearson's R values is the high values along parts of the Himalayan orogen, particularly from DJF to MAM.

Pearson's R values from analyses based on observed precipitation rates (NCEP) and $\delta^{18}O$ values exhibit a similar pattern, including the high positive values between Lhasa and Chengdu. However, they often reveal a weaker statistical relationship outside the monsoon season than R values based on model output. Only during JJA, R values based on observations show a more widespread, strong negative correlation in the northwest of the investigated region, north of the Caspian and Aral Seas.

The statistically significant Pearson's R values attained in the correlation analysis between 2 m temperatures and $\delta^{18}O$ values are predominantly positive and typically range from 0.2 to over 0.8 (Figure 10). In the region between the Aral Sea and the Arabian Sea, calculated Pearson's R values are low in the seasons SON and MAM
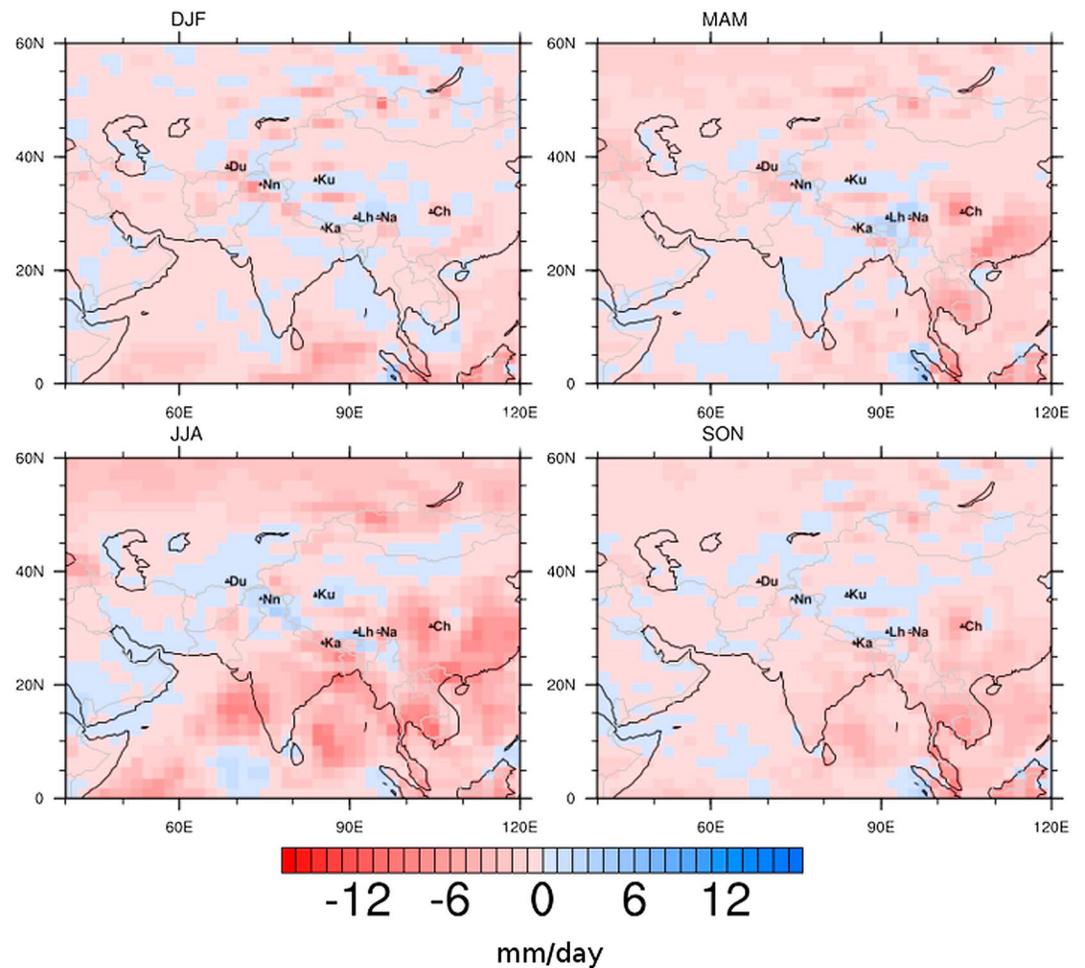
Figure 6. Difference between NCEP and ERA40 precipitation for seasons DJF (top left), MAM (top right), JJA (bottom left), and SON (bottom right). Long-term means of ERA40 precipitation were subtracted from NCEP precipitation means for their temporal overlap.

and insignificant in summer. There are a few significant R values for the Indian subcontinent during any season of the year. Along the eastern Himalayan orogen, south of Kathmandu, Pearson's R values lie between 0.8 and 1.0 from JJA to DJF. For most seasons, few significant correlation coefficients were obtained for the region west of Lake Baikal. The correlation analysis for the region east of the Kunlun mountains also yielded few significant Pearon's R values during the summer season. Overall, there are fewer significant Pearson's R values during the summer monsoon than during the remaining year.

Pearson's R values from analyses based on observed 2 m air temperature (NCEP) and $\delta^{18}O$ values exhibit a similar pattern. In a few regions, including the Tibetan plateau and north of the plateau, R values based on observations are lower (by approximately 0.2–0.4) but mostly still significant. For most seasons, observation-based analyses yield high positive (>0.8) R values on the Indian subcontinent, where model output-based analyses yield no or few statistically significant R values.

The statistically significant Pearson's R values attained in the correlation analysis between the interpolated GTOPO30 topography and $\delta^{18}O$ values are predominantly negative and typically range from -0.4 to over -0.8 (Figure 11). Most of the highest negative values can be seen on the Tibetan Plateau, the Himalayan orogen, and in eastern China. Correlation analyses for the region northeast of the Aral Sea yield only a few significant Pearson's R values throughout the year, possibly due to low-relief terrain. The region north of Lake Balkhash is characterized by a positive correlation between $\delta^{18}O$ and topography throughout the year. The statistical relationship generally weakens north of China, east of the plateau, and on the Indian subcontinent.
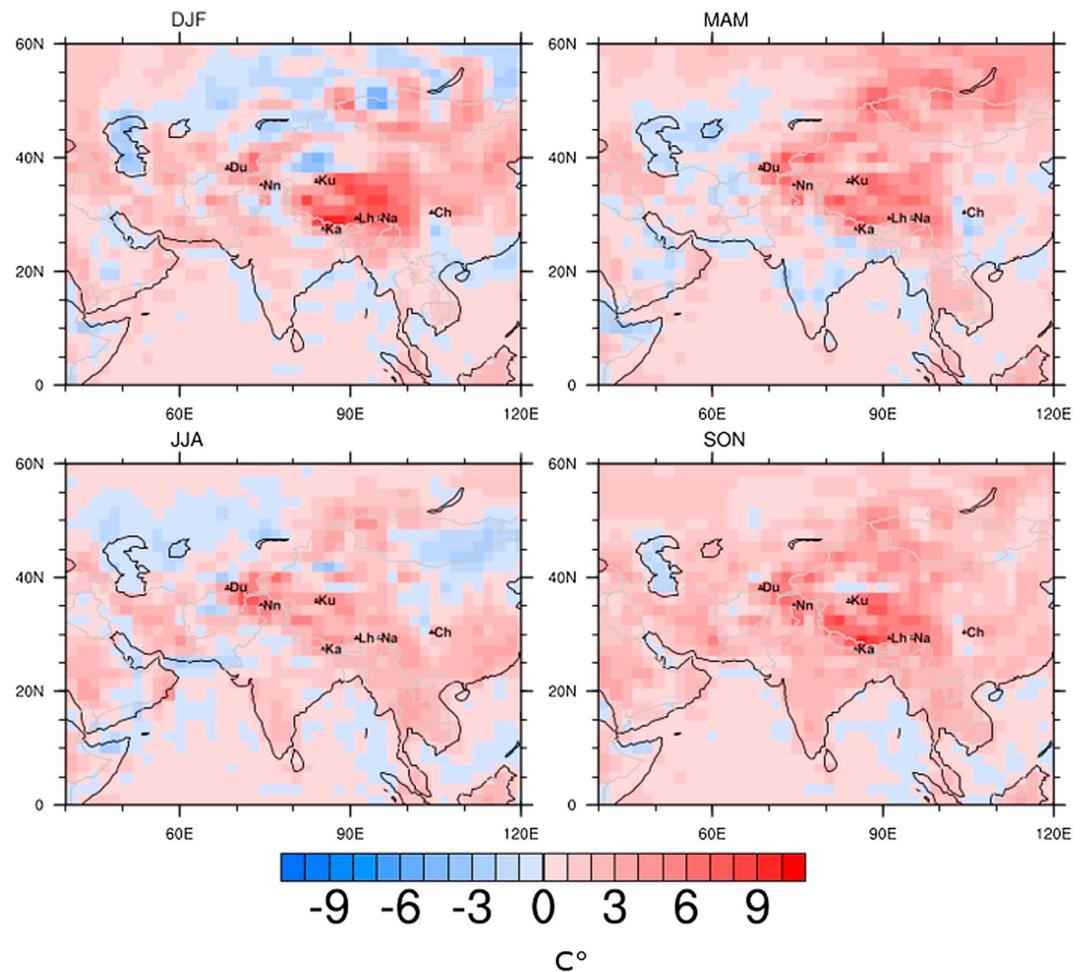
Figure 7. Difference between NCEP and ERA40 2 m air temperature for seasons DJF (top left), MAM (top right), JJA (bottom left), and SON (bottom right). Long-term means of ERA40 2 m air temperature were subtracted from NCEP 2 m air temperature means for their temporal overlap.

Correlation of GTOPO30 topography and observation-based δ^{18}O values is of very limited use and would result in overfitting, since δ^{18}O values in the *Bowen and Revenaugh* [2003] data set are interpolated with topography and actual observations are rare.

The results from the correlation analysis carried out with 5×5 and 9×9 grid box sized geographical areas (not shown) reveal that most of the patterns of statistically significant values described in this section are persistent but slightly weakened for 9×9 and sparser for 5×5. Due to a masking procedure that ignores R values that are calculated using grid boxes over bodies of water, the missing value margin at the continent edges increases in area with an increase in the size of the geographical areas used in the analysis. The otherwise conspicuous pattern in high values at the Himalayan orogen are therefore lost to some extent in the analysis using 9×9 grid box sized areas.

Overall, the geographical coverage of high R values is greatest for correlation coefficients calculated from temperature and δ^{18}O. Since temperature is controlled by altitude, especially at higher elevations, high widespread R values across the plateau are expected. However, temperature also yields high R values north of the investigated regions where topography often yields no significant R values. An analysis of the difference in the strength of correlation, i.e., difference of absolute correlation coefficients, revealed that large discrepancies between R values from topography-δ^{18}O and temperature-δ^{18}O analyses spatially often coincide with small discrepancies between R values from precipitation-δ^{18}O and temperature-δ^{18}O analyses. To investigate further, temperature was correlated with topography and precipitation for each month. For brevity, the results are presented as cumulative correlation coefficients (Figure 12) by adding up the values calculated
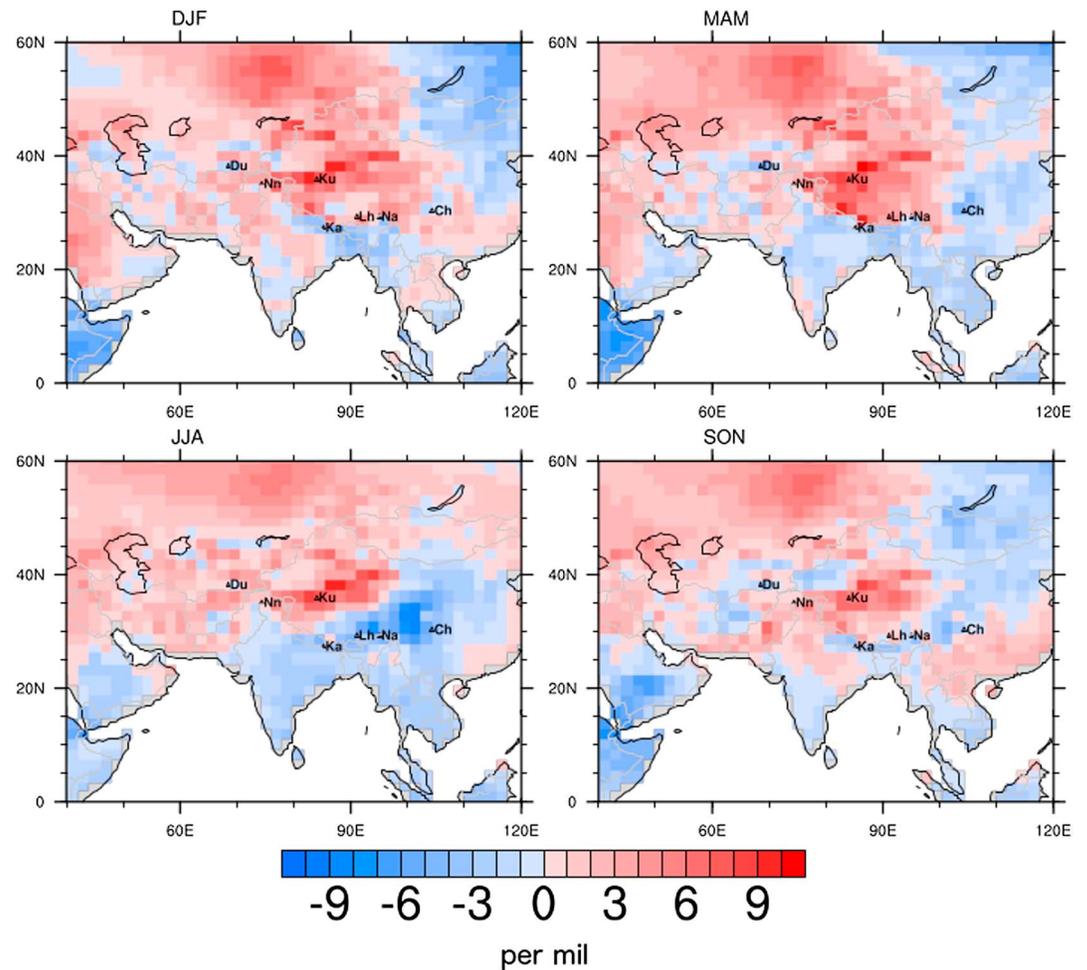
Figure 8. Difference in precipitation δ^{18}O between ECHAM5-wiso and the *Bowen and Revenaugh* [2003] data set for seasons DJF (top left), MAM (top right), JJA (bottom left), and SON (bottom right). Precipitation δ^{18}O values from the *Bowen and Revenaugh* [2003] data set were subtracted from ECHAM5-wiso precipitation δ^{18}O values for each season.

for each month, thus preserving information that may be lost in an analysis of annual values. Results reveal a high statistical relationship of temperature and topography predominantly in the south of the investigated region, north of Lake Balkhash, and north of the Caspian and Aral Seas (Figure 12a). The statistical relationship between temperature and precipitation is highest in the northwest of the investigated region, Mongolia, northwestern India, and southeast of Chengdu (Figure 12b).

3.3. Multiple Regression and Cluster Analysis

Results of the multiple regression analysis are presented in Figure 13. The selection of results is again presented as seasonal means as in previous sections. The percentage of the variance in precipitation δ^{18}O explained by the variables included in the multiple regression analysis, namely, total precipitation, 2 m temperature, and topography, ranges from 0% to over 90% (Figure 12). During the summer season, the average explained variance in the region of interest is lower than in all of the other seasons.

In the region between the Aral Sea and the Arabian Sea, the southern Indian subcontinent, east of the Kunlun Mountains and along the Russian-Mongolian border west of Lake Baikal, the average percentage of explained variance is lowest and often within 0%–10%. The region between the Aral Sea and the Arabian Sea holds higher values in winter. In Kazakhstan, southern Russia, eastern China, and along the Himalayan orogen, the percentage of explained variance is often highest and sometimes 90–100%. The explained variance along the Himalayan orogen at and just south of Kathmandu is particularly high in winter, but more restricted to the eastern Himalayan orogen during JJA and SON, and lowest in the MAM season.
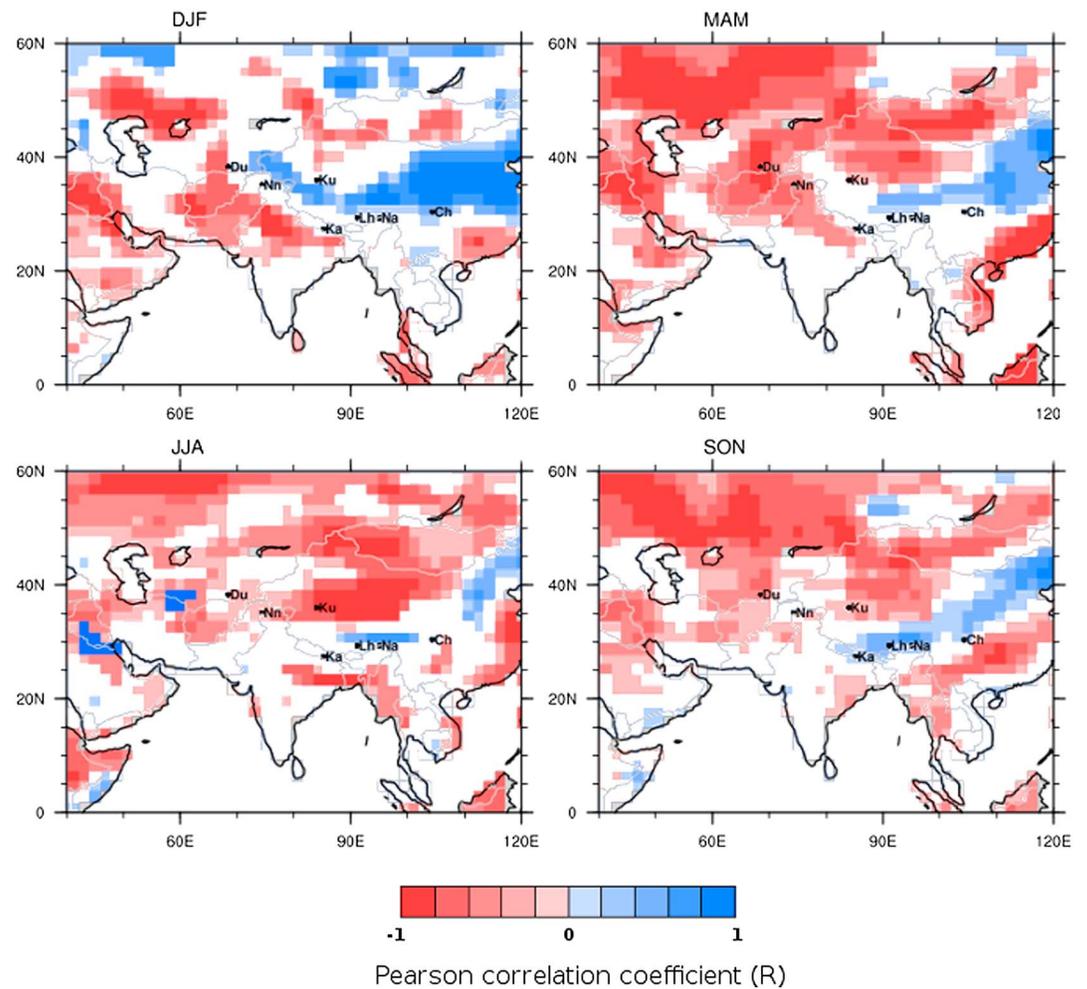
Figure 9. Correlation coefficients calculated from long-term (30 year) means of total precipitation and precipitation δ^{18}O values for seasons DJF (top left), MAM (top right), JJA (bottom left), and SON (bottom right).

Similar to the results from correlation analyses, the pattern of explained variance calculated in the multiple regression procedure persists but weakens with a grid box area increase.

Results from the multiple regression analysis carried out over the complete study area, from 40°E to 120°E and 0°N to 60°N, shows similarly high total explained variance, ranging from ~50% to ~90% over space and months analyzed. As in the results from the analyses using smaller geographic areas, a decrease in the explained variance in the summer season can also be observed here. This temporally coincides with a weakened statistical δ^{18}O-temperature relationship and slightly stronger δ^{18}O-precipitation and δ^{18}O-topography relationship. Averaged over all seasons, the total explained spatial variance (73.7) can be broken down into that coming from 2 m air temperature (83.8% of the total), from topography (16.2%), and from total precipitation (1.2%).

The multiple regression analysis conducted on the basis of observation-based data does not yield the same high values of explained variance. The highest values are calculated for August and the JJA season on the plateau, reaching up to approximately 80%. However, most values are below 50%. Other regions of high explained variance include the eastern Himalayan orogen syntax near Kathmandu and Lhasa and the region north of the Caspian and Aral Seas. Even though the explained variance calculated on the basis of observational data sets is much lower than that calculated on the basis of model output, the regions characterized by highest explained variance are the same. This result underscores the utility of using an isotope-tracking AGCM to understand spatial variability in δ^{18}O-precipitation for regions where limited data are available.
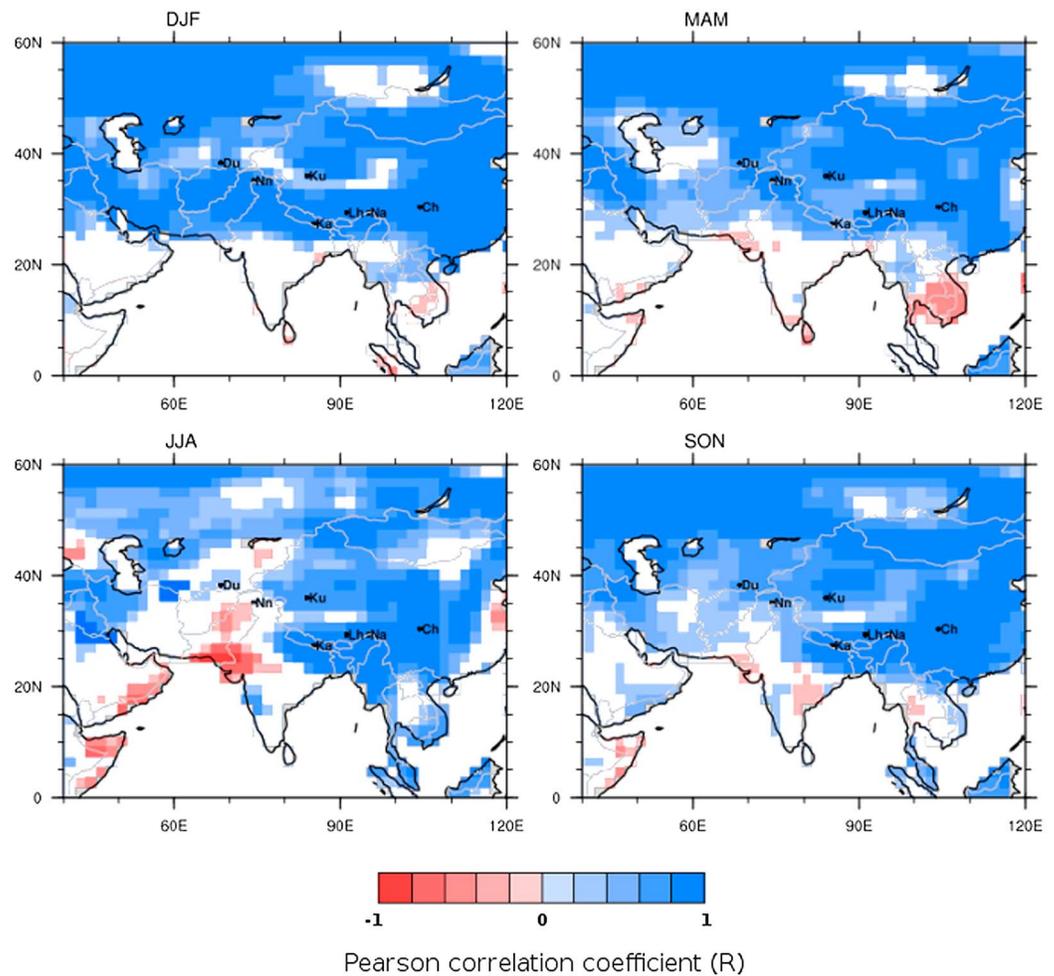
Figure 10. Correlation coefficients calculated from long-term (30 year) means of 2 m air temperature and precipitation $\delta^{18}O$ values for seasons DJF (top left), MAM (top right), JJA (bottom left), and SON (bottom right).

Results from the cluster analysis carried out on the basis of modeled precipitation $\delta^{18}O$ variability show three zones along a southeast-northwest axis on the Tibetan Plateau. Explained $\delta^{18}O$ variance is similarly high in all clusters (92%, 93%, and 97% in clusters 1, 2, and 3, respectively). Contributions by topography, temperature, and precipitation vary from 64%, 35%, and 1% in cluster 1 to 87%, 13%, and 0% in cluster 2 and 40%, 60%, and 0% in cluster 3 (Figure 14). Topography and temperature are highly correlated in the clusters ($R = 0.94$, 0.94, and 0.96 in clusters 1, 2, and 3, respectively). Cluster 1 is characterized by high annual (−9.7) and monsoon season (−5.4) means and large amplitude of values (12.8). Cluster 2 is characterized by moderate annual (−11.6) and monsoon mean (−11.4) and moderate amplitude of values (11). Cluster 3 is characterized by low annual (−15.2) and monsoon mean (−18) and low amplitude of values (9.5). The geographical coverage of the third cluster corresponds to the distribution of extremely low values in August and other monsoon months. Respectively, 40, 32, and 40 grid boxes are assigned to clusters 1, 2, and 3.

The cluster analysis carried out on the basis of observed precipitation $\delta^{18}O$ yields similar results (Figure 15), but cluster 3 is more central than cluster 3 calculated on the basis of model output. Cluster 1 is characterized by high annual (−10.2) and monsoon season (−7.1) means and large amplitude of values (11.5). Cluster 2 is characterized by moderate annual (−14.6) and monsoon mean (−12.3) and moderate amplitude of values (10.1). Cluster 3 is characterized by low annual (−17.2) and monsoon mean (−16.2) and low amplitude of values (8.6). Respectively, 37, 35, and 40 grid boxes are assigned to clusters 1, 2, and 3.

Clustering based on R values calculated in the correlation analyses using model output results in one cluster occupying much of the southern Tibetan Plateau, one occupying the northern and western plateau, and a smaller
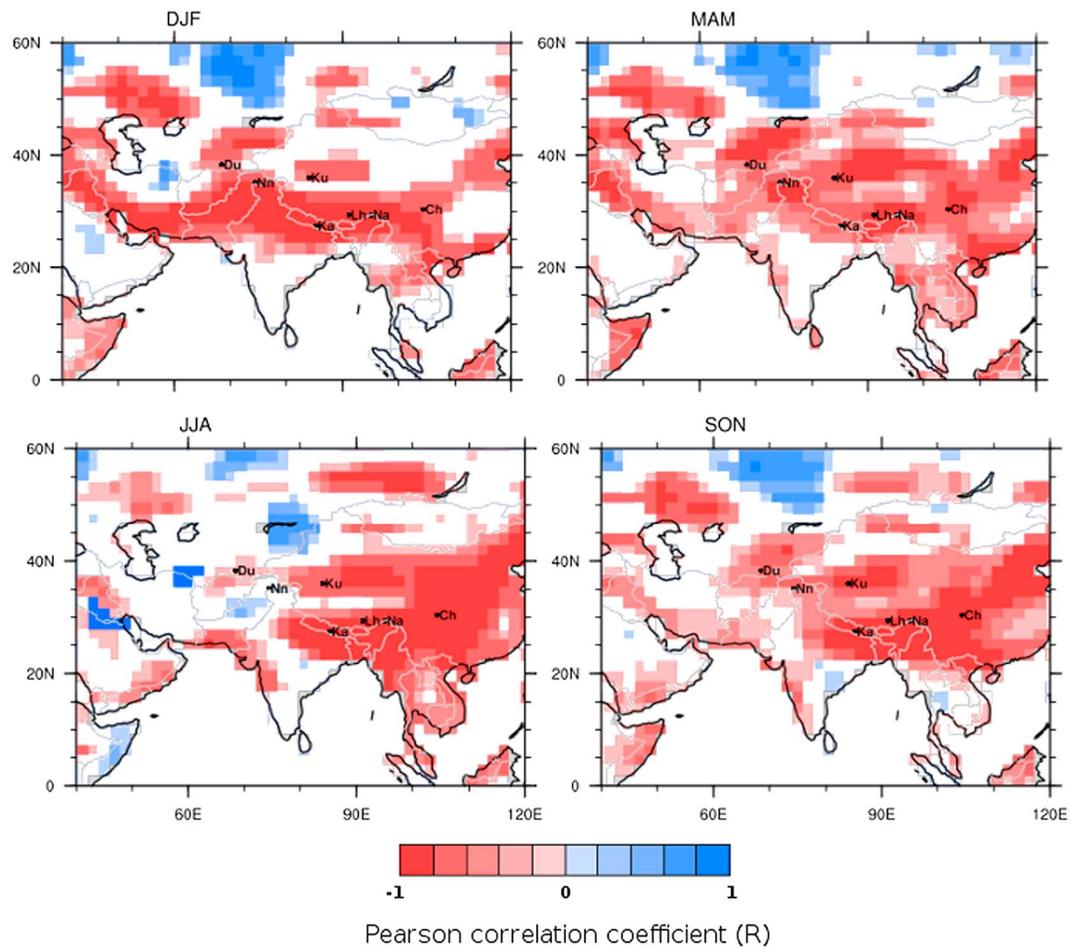
Figure 11. Correlation coefficients calculated from topography and long-term (30 year) means of precipitation $\delta^{18}O$ values for seasons DJF (top left), MAM (top right), JJA (bottom left), and SON (bottom right).
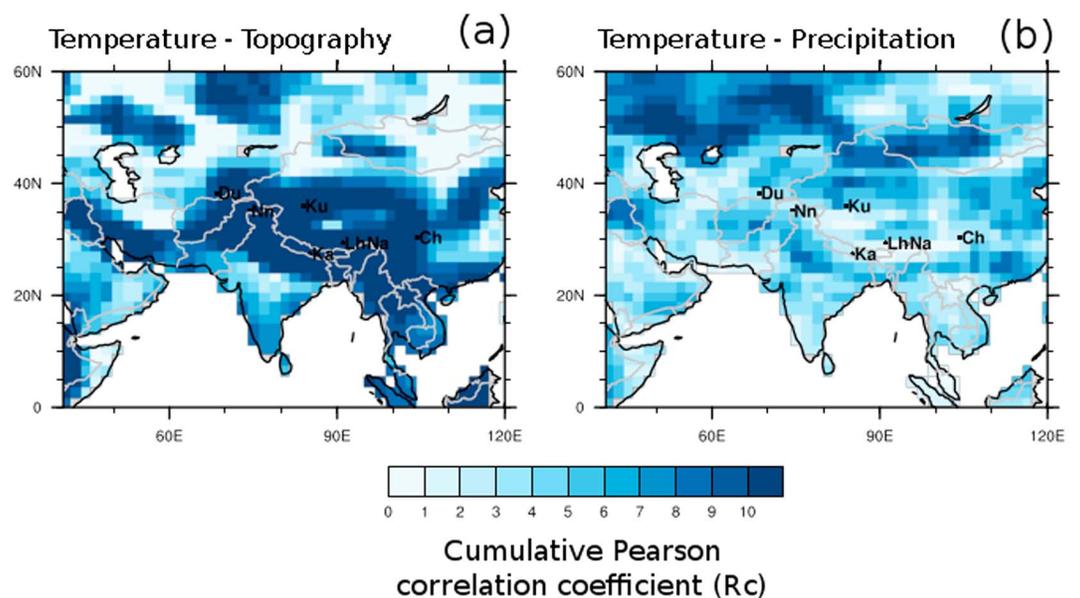

Figure 12. Cumulative absolute monthly correlation coefficients for analyses of (a) temperature and topography, and (b) temperature and precipitation.
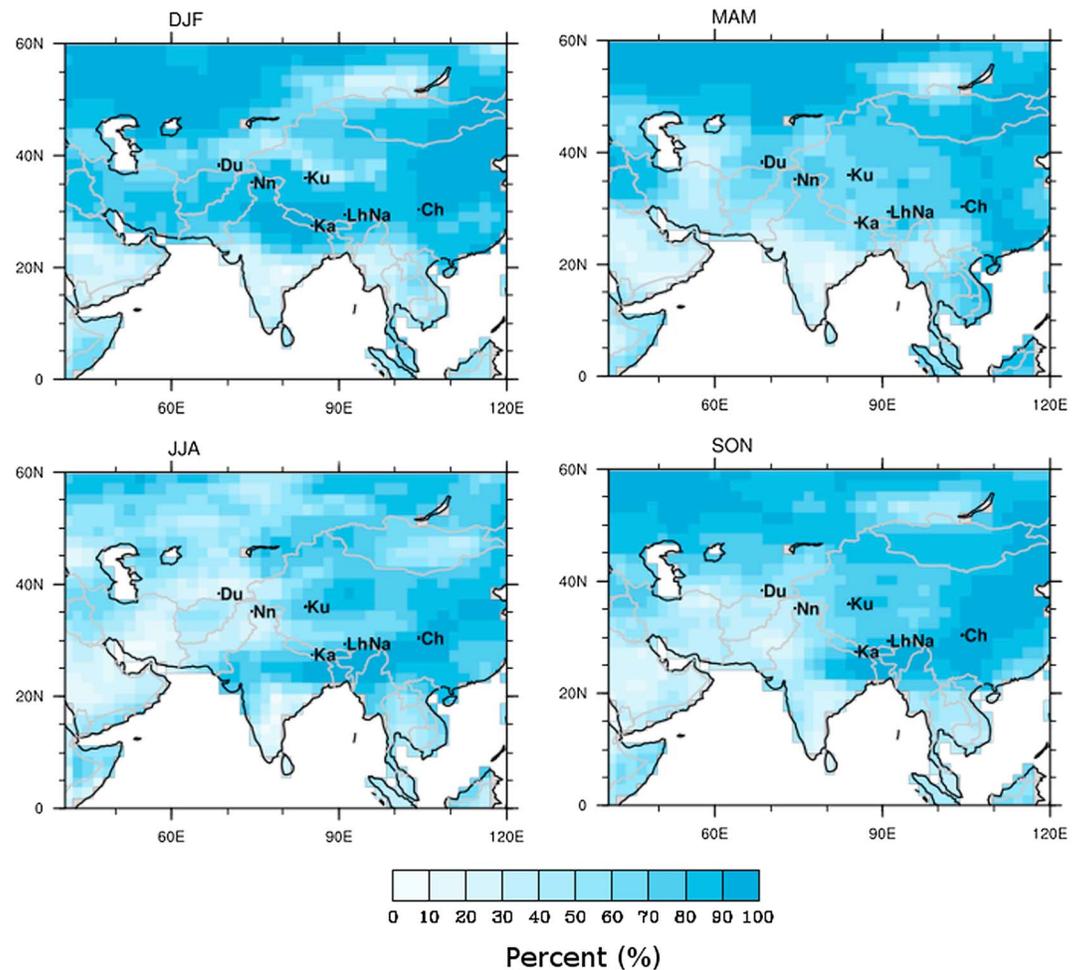
Figure 13. Total $\delta^{18}O$ variance explained by all predictors (temperature, topography, and precipitation) for seasons DJF (top left), MAM (top right), JJA (bottom left), and SON (bottom right).

cluster between the northern and southern clusters on the eastern plateau (Figure 16). The larger clusters are characterized by high correlation of temperature and topography with $\delta^{18}O$, while the middle zone is characterized by a weaker relationship with temperature and topography and a stronger relationship with precipitation.

4. Discussion

4.1. Synthesis of Results and Processes Influencing Precipitation $\delta^{18}O$

Generally, ECHAM5-wiso accurately reproduces temperature and precipitation patterns in the investigated region. In most cases, discrepancies between model and observation-based (ERA40, NCEP, and CRU) precipitation and 2 m air temperature do not exceed discrepancies between the observation-based data products. A notable exception is the model bias in precipitation along the Himalayan orogen. ECHAM5-wiso significantly overestimates precipitation in this region, and this bias persists in comparison to all observation-based data products. ECHAM5-wiso also overestimates near-surface temperatures on the plateau. However, there seem to be similarly large discrepancies in temperature between observation-based data products (Figure 7 and S8 in supporting information). These model biases in precipitation and temperature have already been identified for an older version of the model (ECHAM4-iso) and the GCM of the Laboratoire de Météorologie Dynamique (LMDZ-iso) by *Gao et al.* [2011], and resulting limitations for the interpretation of model output ought to be taken into consideration. Modeled precipitation $\delta^{18}O$ values show large differences to the observation-based *Bowen and Revenaugh* [2003] data set on parts of the Tibetan Plateau, eastern Kazakhstan, and between Lhasa and Chengdu in summer. However, since the density of observations in the
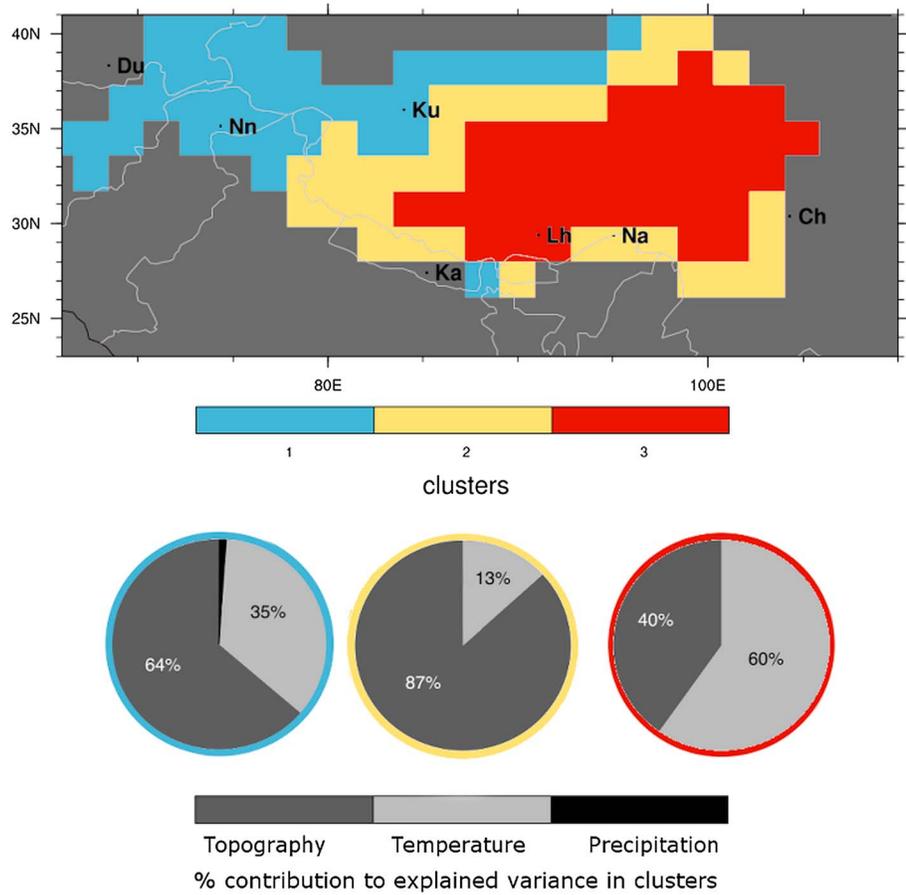
Figure 14. (top) Clusters on the Tibetan Plateau determined by the cluster analysis on the basis of modeled $\delta^{18}O$ variability and (bottom) annual mean percentage contributions of predictors (temperature, topography, and precipitation) to the total explained $\delta^{18}O$ variance as calculated in the multiple regression.

investigated region is low, these differences may be due to model biases as well as the spatial interpolation method applied by *Bowen and Revenaugh* [2003].

The correlation coefficients calculated from precipitation means and precipitation $\delta^{18}O$ values are predominantly negative. One notable observation in the spatial distribution of these values is a pattern of strong
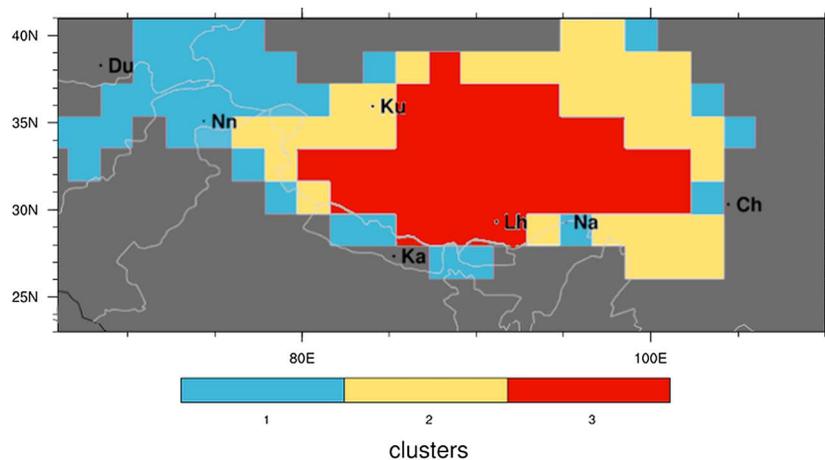

Figure 15. Clusters on the Tibetan Plateau determined by the cluster analysis on the basis of observed $\delta^{18}O$ variability.

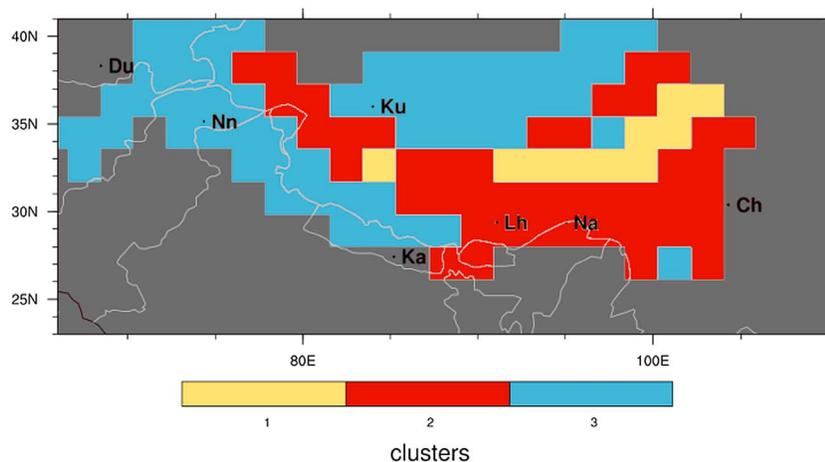**Figure 16.** Clusters on the Tibetan Plateau determined by the cluster analysis on the basis of correlation of $\delta^{18}O$ values with precipitation, temperature, and topography.

negative correlation coefficients south of the Himalaya that becomes more pronounced and stretches farther to the northwest toward Nanga Parbat and Dushanbe in seasons DJF and MAM. This temporally coincides with higher average precipitation rates around Nanga Parbat and Dushanbe. This negative correlation may be attributed to rainout. However, the record high precipitation rates at Kathmandu and along the Himalayan orogen in summer are not reflected in the Pearson's R values presented here, possibly due to the fact that only variance in space is considered. R values calculated from observation-based data sets are lower except in summer, when strong negative correlation is more widespread north of the Caspian and Aral Seas.

The correlation coefficients calculated from 2 m air temperature means and precipitation $\delta^{18}O$ values are predominantly positive and particularly high (>0.9) in the region around Chengdu and along the eastern Himalayan orogen in all seasons. The high Pearson's R values along the Himalayan orogen spatially coincide with a steep temperature and topography gradient and may be attributed to the rainout effect, i.e., the depletion of isotopic values through the loss of moisture from the air as a result of cooling [*Dansgaard*, 1964; *Gat*, 1996]. Results from observation-based analyses yield similar results. The main differences are high positive R values on the Indian subcontinent and 0.2–0.4 lower R values on the plateau, north of the plateau, where stations providing observations are sparse. Overall, correlation of observation-based data yield lower R values. This may reflect an overestimation of local climatic controls on the $\delta^{18}O$ distribution by the model. In an investigation of the influence of the Asian monsoon on $\delta^{18}O$, *Vuille et al.* [2005] have already identified an overestimation of the local climatic controls, in particular temperature, by the older ECHAM4-iso.

The correlation coefficients calculated from GTOPO30 topography and precipitation $\delta^{18}O$ values are predominantly negative and particularly high (< -0.9) along the Himalayan orogen and on parts of the plateau and China. The high correlation coefficients along the Himalayan orogen coincide with a steep topographic gradient and may be attributed to the effect of altitude on precipitation $\delta^{18}O$ values. An increase in altitude favors a decrease in temperature, cooling of air masses, and depletion of isotopic values due to the loss of moisture from the air [*Dansgaard*, 1964; *Gat*, 1996].

Overall, the distribution of Pearson's R values calculated from topography and precipitation $\delta^{18}O$ is similar to that of the values calculated from 2 m air temperature and precipitation $\delta^{18}O$ but spatially more restricted to the plateau and surrounding regions due to the high covariance and strong physical relationship of temperature and topography in regions of great topographic relief. The correlation analysis using 2 m air temperature and isotopic values more frequently yielded high correlation coefficients (>0.9) in Kazakhstan and southern Russia.

Correlation analyses using topography and temperature yield high R values especially in the southern high-altitude regions and lower values in the lower altitude north. Temperature and precipitation are highly correlated mostly in the northwest of the investigated region. This spatially coincides with areas where temperature is highly correlated with $\delta^{18}O$ values but topography is not. Taking this into account, the high

correlation of temperature with δ^{18}O values in regions with notable topographic features can be explained by the inverse physical relationship between altitude and temperature. The correlation of δ^{18}O with temperature in lower altitude northwest characterized by no significant correlation between temperature and topography may ultimately not be attributed to a topographic control. The high correlation of temperature with δ^{18}O values in this region may instead be related to precipitation-related processes.

The amount of precipitation δ^{18}O variance that can be explained by the three factors, precipitation, temperature, and topography, in the multiple regression analysis is highest along the Himalaya and China. The 2 m air temperature is the dominant predictor in almost all regions. In the regression analysis using the entire study area, it contributes 93.6% of the total explained spatial variance, while topography and total precipitation only contribute 4.5% and 1.9%, respectively. The average contribution of topography to explained δ^{18}O variance based on moving geographical windows is 16.2%, and the average δ^{18}O variance explained by topography in the three precipitation δ^{18}O clusters on the plateau are 64%, 87%, and 40%. These results are significant as they show that in some regions, topography shows a more statistically robust relationship with modeled δ^{18}O variability than temperature and may therefore be a better predictor for spatial δ^{18}O variability, especially on the plateau. During the summer season (JJA), less variance can be explained than during other seasons. This can be observed for results of all regression analyses and may be due to the failure to include a variable in the spatial analyses that adequately represents the effects of vapor source on precipitation δ^{18}O. Upstream circulation has been demonstrated to be one of the important controls on precipitation δ^{18}O over the Tibetan Plateau [*Vuille et al.*, 2005; *Gao et al.*, 2011, 2013; *Bershaw et al.*, 2012; *Yao et al.*, 2013; *He et al.*, 2015]. *Gao et al.* [2011] note that the older ECHAM4-iso simulations underestimate the amount effect in summer at the southern margin of the plateau. Part of the unexplained variance, especially during the monsoon season, may be due to the inability of the model, and the restrictions of the spatial correlation and multiple regression applied here, to capture this effect. This explanation is also supported by *Li et al.* [2016] highlighting the importance of vapor source as a control on δ^{18}O during the monsoon season.

While the explained variances calculated in the multiple regression procedure using observation-based data is much lower, the regions of highest explained variance are the same.

4.2. Comparison to Other Studies

Previous work by *Tian et al.* [2003] and *Araguas-Araguas* [1998] were based on observational data and describe different climatic controls on δ^{18}O value distribution in Southeast Asia. Based on a <9 year record of meteorological station data, *Tian et al.* [2003] describe a strong relationship between enhanced monsoon activity, increased precipitation rates, and a depletion in heavy isotopes on a 1500 km southwest-northeast transect of the Tibetan Plateau. This is accompanied by a weak δ^{18}O-temperature relationship on seasonal and annual scales [*Tian et al.*, 2003]. *Araguas-Araguas* [1998] also describes the amount effect overshadowing the temperature effect in certain regions such as Lhasa. While the results of this study also show a weaker δ^{18}O-temperature relationship during the summer monsoon, precipitation rates continue to explain very little of the spatial δ^{18}O variance. This could be due to the lower spatial resolution and timescales used in this study or inadequately represented precipitation rates. However, the strong relationship between δ^{18}O values and temperature on a monthly scale, described by *Tian et al.* [2003], is consistent with the findings of this study.

Liu et al. [2008] used a model to predict δ^{18}O distribution over China by using geographical parameters, namely, latitude and altitude, as proxies for the more direct controls of the isotopic composition of precipitation (i.e., temperature driven rainout effect and vapor source control). In this study, Pearson correlation coefficients of altitude and δ^{18}O values are high over the plateau and China throughout most of the year and show, not surprisingly, a similar yet inverse relationship to that of temperature with δ^{18}O values. However, the relationship of temperature with precipitation δ^{18}O remains strong in regions where the relationship of altitude and δ^{18}O weakens, thus highlighting the regional limits of using altitude rather than temperature as a predictor of δ^{18}O in statistical models.

Studies by *Vuille et al.* [2005] and *Gao et al.* [2011] use ECHAM4-iso to understand isotopic variability. Both studies highlight ECHAM4-iso's deficiencies in the model's ability to simulate different aspects of monsoon precipitation. *Gao et al.* [2011] also list an overestimation of precipitation, underestimation of temperature on the Tibetan Plateau, a stronger than observed correlation of precipitation and temperature with δ^{18}O, an underestimation of the amount effect in summer, and overestimation of the effect of temperature and

precipitation in winter at the southern Tibetan Plateau as biases for ECHAM4-iso (and LMDZ). ECHAM5-wiso output generated for this study shows a stronger (spatial) correlation between simulated temperature and precipitation with $\delta^{18}O$ than between observation-based data sets. However, the sparse $\delta^{18}O$ observation record in the investigated region limits the conclusions that can be drawn from the comparison of model correlations with observation-based correlations. ECHAM5-wiso also overestimates precipitation at the southern Tibetan Plateau during the monsoon season. However, it does not show an underestimation of temperatures on the Tibetan Plateau. Findings of a high correlation between temperature and $\delta^{18}O$ north of the region influenced by the Indian Monsoon empirically complement findings of *Tian et al*. [2001a, 2001b]. The cluster analysis based on variables describing only model precipitation $\delta^{18}O$ variability split up the domain of the Tibetan Plateau into roughly southeast to northwest trending zones and provides empirical support for the more qualitative division based on $\delta^{18}O$ values [*Li et al*. 2016]. The analysis based on the relationships between precipitation $\delta^{18}O$ and temperature, precipitation, and topography (established in this study) yields two major zones in the north and south of the eastern plateau and a smaller zone between them. The latter supports the division into three north-south trending zones as discussed in, e.g., *Tian et al*. [2001a] and *Yao et al*. [2013] based on controls of $\delta^{18}O$ values rather than $\delta^{18}O$ variability only. The weaker correlation of topography and temperature with $\delta^{18}O$ in the smaller middle zone between the Himalayas and Tanggula Mountains south of the Kunlun Mountains is consistent with reported loss of altitude effect [*Yao et al*., 2013]. It is accompanied by stronger correlation with precipitation and spatially overlaps with the region identified by *Tian et al*. [2001a] in which summer moisture originates from the Bay of Bengal.

4.3. Methodological Limitations

In addition to general model uncertainties of ECHAM5, there are a number of limitations of the statistical methods used in this study. Due to the incremental procedure of the multiple regression that prevents an overestimation of explained variance (section 2.4), similar and strongly dependent predictors, such as 2 m air temperature and topography in this study, may "compete" to be chosen as the first predictor. This may result in little explained $\delta^{18}O$ variance by one of the competing predictors despite a high correlation between them or even in a complete exclusion of the predictor from the statistical model. The multicollinearity of topography and temperature results in such a competition and may be the reason why temperature explains significantly more isotopic variance than topography despite the high correlation of topography with $\delta^{18}O$ values.

The residual or unexplained $\delta^{18}O$ variance may be the result of omissions of relevant factors from our analysis. One such example of this includes the effect of the distance of the investigated geographical window to the coast along the vapor source trajectory. Such an isotopic depletion factor was identified as important by *Dansgaard* [1964]. For part of the investigated region, i.e., the southern Tibet, Nepal, and northern India, upstream convective activity during the monsoon season has been listed as a major control of spatial and temporal $\delta^{18}O$ variability [*Vuille et al*., 2005; *Gao et al*., 2011, 2013; *Bershaw et al*., 2012; *Yao et al*., 2013; *He et al*., 2015]. While the importance of local climate in explaining the spatiotemporal variability of precipitation $\delta^{18}O$ at the southern Tibetan Plateau has also been highlighted [*Gao et al*., 2011], they may account only for a small fraction of the $\delta^{18}O$ variability during the monsoon season [*Gao et al*., 2013]. This suggests that caution is needed when interpreting the results of this study for that particular region during the monsoon season.

Unfortunately, the analysis conducted here with different size windows for analysis around a point is not amendable to include predictors that adequately capture upstream convective activity and, hence, is the focus of our efforts using a stream trace analysis in the companion paper to this one by *Li et al*. [2016]. As part of this study, several other efforts have been made to quantify the effect of the rainout along the path traveled by an air mass on precipitation $\delta^{18}O$. These include the calculation of the distance to the nearest body of water, the calculation of the distance to the nearest body of water weighted by the frequency and degree of deviation of wind directions from the direct path, and the calculation of the distance between an investigated geographical window and a region over the sea identified as relevant to the local precipitation by correlation with latent heat flux over the sea. When used in the regression procedure, the weighted distances contributed most ($> \sim 5\%$) to explaining the $\delta^{18}O$ variance that had not been explained by the other predictors used in this study. This predictor may be refined by making a more informed choice of a

possible water source region and developing a simple dynamic model that is able to accommodate geographical shifts in the water source region. None of the attempts we made produced satisfactory results for quantifying the effects of upstream circulation over Tibet on precipitation $\delta^{18}O$ with the correlation and regression analyses. This may explain the decrease in explained $\delta^{18}O$ variance in summer and remains one of the study's shortcomings.

Finally, model performance may introduce a bias in the results of the model-based multiple regression analysis. While ECHAM5 and other GCMs are able to accurately simulate near-surface air temperatures, mesoscale circulation is not well resolved and regional orography is poorly represented [e.g., *Cohen*, 1990]. Since the main results of this study are based on relatively small geographical domains, varying model performance for temperature and precipitation may result in an underestimation of the role of precipitation in explaining the spatial variability of precipitation $\delta^{18}O$.

The caveats associated with our analysis are mentioned to highlight that some degree of (unquantifiable) uncertainty exists in our results. The approach presented here does, however, provide a robust statistical analysis of the common regional factors attributed to spatial precipitation $\delta^{18}O$ variations.

4.4. Conclusions

The primary conclusion of this study is that predictors considered in the statistical model, namely, 2 m air temperature, precipitation, and topography, are able to explain most of the precipitation $\delta^{18}O$ spatial variance over the Tibetan Plateau. More specifically, topography and topography-related variations in temperature show the strongest statistical relationship with precipitation $\delta^{18}O$ values over the Himalaya-Tibet region, indicating that most of the modeled spatial $\delta^{18}O$ variability can be explained by the rainout effect, i.e., the depletion through loss of moisture from the air as a result of cooling. While temperature yields most of the highest Pearson correlation coefficients for different geographical regions investigated, topography shows an even stronger correlation along the Himalaya Mountains, suggesting that topographic variability in these regions empirically better represents the controls of the rainout effect than 2 m air temperature. In regions where temperature is the dominant predictor but primarily is controlled by topography, like in the region of the Tibetan plateau, much of the precipitation $\delta^{18}O$ value distribution can ultimately also be attributed to topography, although the control of the rainout effect is better represented by temperature variability here. In regions where topography has no apparent control over temperature, results from temperature-precipitation correlation analyses suggest that the high correlation of temperature with $\delta^{18}O$ values may in part be attributed to variability in precipitation rates. The smallest amount of the precipitation $\delta^{18}O$ variance that can be explained in our analysis occurs between the Aral Sea and the Arabian Sea. Values of explained variance are particularly low in summer when the otherwise strong statistical relationship between temperature and $\delta^{18}O$ is weakened. The correlation of temperature with $\delta^{18}O$ is also weak in the region influenced by the Indian Monsoon. The inability of the model and methods to adequately capture the effect of upstream circulation on $\delta^{18}O$ variability in Tibet, which has previously been identified as a major control in the region during the monsoon season, is one of this study's limitations and may be the reason for the low explained $\delta^{18}O$ variance in summer in the regions affected by the monsoon.

It can be concluded that the temperature and altitude effect are the dominant regional controls of the spatial distribution of isotopic values at least outside the summer monsoon season and north of the region influenced by the Indian Monsoon.

References

Araguas-Araguas, L., K. Froehlich, and K. Rozanski (1998), Stable isotope composition of precipitation over Southeast Asia, *J. Geophys. Res.*, *103*(D22), 28,721–28,742, doi:10.1029/98JD02582.

Bahrenberg, G., E. Giese, and J. Nipper (1992), Multivariate Statistik, in *Statistische Methoden in der Geographie 2*, Gebrüder Borntraeger Verlagsbuchhandlung Berlin, Stuttgart.

Bershaw, J., S. M. Penny, and C. N. Garzione (2012), Stable isotopes of modern water across the Himalaya and eastern Tibetan Plateau: Implications for estimates of paleoelevation and paleoclimate, *J. Geophys. Res.*, *117*, D02110, doi:10.1029/2011JD016132.

Bohner, J. (2006), General climatic controls and topoclimatic variations in Central and High Asia, *Boreas*, *35*(2), 279–295.

Boos, W. R., and Z. Kuang (2010), Dominant control of the South Asian monsoon by orographic insulation versus plateau heating, *Nature*, *463*, 218–222.

Bowen, G. J., and B. Wilkinson (2002), Spatial distribution of delta O-18 in meteoric precipitation, *Geology*, *30*(4), 315–318.

Bowen, G. J., and J. Revenaugh (2003), Interpolating the isotopic composition of modern meteoric precipitation, *Water Resour. Res.*, *39*(10), 1299, doi:10.1029/2003WR002086.

Cohen, S. J. (1990), Bringing the global warming issue close to home: The challenge of regional impact studies, *Bull. Am. Meteorol. Soc.*, *71*, 520–526.

Dansgaard, W. (1964), Stable isotopes in precipitation, *Tellus*, *16*(4), 436–468.

Dettman, D. L., X. M. Fang, C. N. Garzione, and J. J. Li (2003), Uplift-driven climate change at 12 Ma: A long δ18O record from the NE margin of the Tibetan plateau, *Earth Planet. Sci. Lett.*, *214*(1-2), 267–277.

Ehlers, T. A., and C. J. Poulsen (2009), Influence of Andean uplift on climate and paleoaltimetry estimates, *Earth Planet. Sci. Lett.*, *281*(3–4), 238–248.

Gao, J., V. Masson-Delmotte, T. Yao, L. Tian, C. Risi, and G. Hoffmann (2011), Precipitation water stable isotopes in the south Tibetan Plateau: Observations and modeling, *J. Clim.*, *24*(13), 3161–3178.

Gao, J., V. Masson-Delmotte, C. Risi, Y. He, and T. Yao (2013), What controls precipitation $\delta^{18}O$ in the southern Tibetan Plateau at seasonal and intra-seasonal scales? A case study at Lhasa and Nyalam, *Tellus Ser. B*, *65*, 21043, doi:10.3402/tellusb.v65i0.21043.

Gat, J. R. (1996), Oxygen and hydrogen isotopes in the hydrologic cycle, *Annu. Rev. Earth Planet. Sci.*, *24*, 225–262.

Harris, I., P. D. Jones, T. J. Osborn, and D. H. Lister (2013), Updated high-resolution grids of monthly climatic observations—The CRU TS3.10 Dataset, *Int. J. Climatol.*, doi:10.1002/joc.3711.

He, Y., et al. (2015), Impact of atmospheric convection on south Tibet summer precipitation isotopologue composition using a combination of in situ measurements, satellite data, and atmospheric general circulation modeling, *J. Geophys. Res. Atmos.*, *120*, 3852–3871, doi:10.1002/2014JD022180.

Hoffmann, G., M. Werner, and M. Heimann (1998), Water isotope module of the ECHAM atmospheric general circulation model: A study on timescales from days to several years (vol 103, pg 16871, 1998), *J. Geophys. Res.*, *103*(D18), 23,323–23,323, doi:10.1029/98JD02857.

Insel, N., C. J. Poulsen, and T. A. Ehlers (2009), Influence of the Andes Mountains on South American moisture transport, convection, and precipitation, *Clim. Dyn.*, *35*(7–8), 1477–1492.

International Atomic Energy Agency Water Resource Programme (1998), Global network for isotopes in pre-cipitation. The GNIP database. Release 3, October 1999. [Avalable at http://www-naweb.iaea.org/napc/ih/IHS_resources_gnip.html(06/2012).]

Joussaume, S., R. Sadourn, and J. Jouzel (1984), A general-circulation model of water isotope cycles in the atmosphere, *Nature*, *311*(5981), 24–29.

Jouzel, J., G. L. Russell, R. J. Suozzo, R. D. Koster, J. W. C. White, and W. S. Broecker (1987), Simulations of the HDO and H$_2$ O-18 atmospheric cycles using the NASA GISS general-circulation model – the seasonal cycle for present-day conditions, *J. Geophys. Res.*, *92*(D12), 14,739–14,760, doi:10.1029/JD092iD12p14739.

Kalnay, E., et al. (1996), The NCEP/NCAR 40-year reanalysis project, *Bull. Am. Meteorol. Soc.*, *77*(3), 437–471.

Kistler, R., et al. (2001), The NCEP-NCAR 50-Year Reanalysis: Monthly means CD-ROM and documentation, *Bull. Am. Meteorol. Soc.*, *82*(2), 247–267.

Kutzbach, J. E., P. J. Guetter, W. F. Ruddiman, and W. L. Prell (1989), Sensitivity of climate to late Cenozoic uplift in southern Asia and the American West—Numerical experiments, *J. Geophys. Res.*, *94*(D15), 18,393–18,407, doi:10.1029/JD094iD15p18393.

Kutzbach, J. E., W. L. Prell, and W. F. Ruddiman (1993), Sensitivity of Eurasian climate to surface uplift of the Tibetan Plateau, *J. Geol.*, *101*(2), 177–190.

Li, J., T. A. Ehlers, S. G. Mutz, C. Steger, H. Paeth, M. Werner, C. J. Poulsen, and R. Reng (2016), Modern Precipitation δ^{18}O and Trajectory Analysis over the Himalaya-Tibet Orogen from ECHAM5-wiso Simulations, *J. Geophys. Res. Atmos.*, *121*, doi:10.1002/2016JD024818.

Lin, S.-J., and R. B. Rood (1996), Multidimensional flux-form semi-Lagrangian transport schemes, *Monthly Weather Rev.*, *124*, 2046–2070.

Liu, Z., L. Tian, X. Chai, and T. Yao (2008), A model-based determination of spatial variation of precipitation δ18O over China, *Chem. Geol.*, *249*, 203–212, doi:10.1016/j.chemgeo.2007.12.011.

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

Michaelsen, J. (1987), Cross-validation in statistical climate forecast models, *J. Climate Appl. Meteorol.*, *26*(11), 1589–1600.

Molnar, P., W. R. Boos, and D. S. Battisti (2010), Orographic controls on climate and paleoclimate of Asia: Thermal and mechanical roles for the Tibetan Plateau, *Annu. Rev. Earth Planet. Sci.*, *38*, 77–102.

Montgomery, D. R., G. Balco, and S. D. Willett (2001), Climate, tectonics, and the morphology of the Andes, *Geology*, *29*(7), 579–582.

Mutz, S., H. Paeth, and S. Winkler (2015), Modelling of future mass balance changes of Norwegian glaciers by application of a dynamical-statistical model, *Clim. Dyn.*, *46*, 1581–1597, doi:10.1007/s00382-015-2663-5.

Paeth, H. (2004), Key factors in African climate change evaluated by a regional climate model, *Erdkunde*, *58*, 290–315.

Paeth, H., and A. Hense (2003), Seasonal forecast of sub-Sahelian rainfall using cross validated model output statistics, *Meteorol. Z.*, *12*(3), 157–173.

Paeth, H., R. Girmes, G. Menz, and A. Hense (2006), Improving seasonal forecasting in the low latitudes, *Mon. Weather Rev.*, *134*(7), 1859–1879.

Raymo, M. E., and W. F. Ruddiman (1992), Tectonic forcing of late Cenozoic climate, *Nature*, *359*(6391), 117–122.

Reiners, P. W., T. A. Ehlers, S. G. Mitchell, and D. R. Montgomery (2003), Couples spatial variations in precipitation and long-term erosion rates across the Washington Cascades, *Nature*, *426*, 645–647.

Roeckner, E., et al. (2003), The atmospheric general circulation model ECHAM5. Part I: Model description, Rep. 349, *Rep.*, 127 pp., Max Planck Inst. for Meteorol., Hamburg.

Rozanski K., L. Araguas-Araguas, and R. Gonfiantini (1993), Isotopic patterns in modern global precipitation, in *Climate Change in Continental Isotopic Records*, edited by P. K. Swart et al., AGU, Washington, D. C., doi:10.1029/GM078p0001.

Ruddiman, W. F., and J. E. Kutzbach (1989), Forcing of late Cenozoic Northern Hemisphere climate by plateau uplift in southern Asia and the American West, *J. Geophys. Res.*, *94*(D15), 18,409–18,427, doi:10.1029/JD094iD15p18409.

Simmons, A. J., D. M. Burridge, M. Jarraud, C. Girard, and W. Wergen (1989), The ECMWF Medium-Range prediction models development of the numerical formulations and the impact of increased resolution, *Meteorol. Atmos. Phys.*, *40*(1–3), 28–60.

Sturm, C., Q. Zhang, and D. Noone (2010), An introduction to stable water isotopes in climate models: Benefits of forward proxy modelling for paleoclimatology, *Clim. Past*, *6*, 115–129.

Takahashi, K., and D. S. Battisti (2007a), Processes controlling the mean tropical pacific precipitation pattern. Part I: The Andes and the eastern Pacific ITCZ, *J. Clim.*, *20*(14), 3434–3451.

Takahashi, K., and D. S. Battisti (2007b), Processes controlling the mean tropical pacific precipitation pattern. Part II: The SPCZ and the southeast pacific dry zone, *J. Clim.*, *20*(23), 5696–5706.

Thomas, A. (1997), The climate of the Gongga Shan range, Sichuan Province, PR China, *Arct. Alp. Res.*, *29*(2), 226–232.

Tian, L., T. D. Yao, A. Numaguti, and W. Z. Sun (2001a), Stable isotope variations in monsoon precipitation on the Tibetan Plateau, *J. Meteorol. Soc. Jpn.*, *79*(5), 959–966.

Tian, L., V. Masson-Delmotte, M. Stiévenard, T. Yao, and J. Jouzel (2001b), Tibetan Plateau summer monsoon northward extent revealed by measurements of water stable isotopes, *J. Geophys. Res.*, *106*, 28,081–28,088, doi:10.1029/2001JD900186.

Tian, L., T. Yao, P. F. Schuster, J. W. C. White, K. Ichiyanagi, E. Pendall, J. Pu, and Y. Wu (2003), Oxygen-18 concentrations in recent precipitation and ice cores on the Tibetan Plateau, *J. Geophys. Res.*, *108*(D9), 4293, doi:10.1029/2002JD002173.

Uppala, S. M., et al. (2005), The ERA-40 re-analysis, *Q. J. R. Meteorol. Soc.*, *131*, 2961–3012.

von Storch, H., and F. Zwiers (1999), *Statistical Analysis in Climate Research*, 496 pp., Cambridge Univ. Press, Cambridge, U. K.

Vuille, M., M. Werner, R. S. Bradley, and F. Keimig (2005), Stable isotopes in precipitation in the Asian monsoon region, *J. Geophys. Res.*, *110*, D23108, doi:10.1029/2005JD006022.

Werner, M., P. M. Langebroek, T. Carlsen, M. Herold, and G. Lohmann (2011), Stable water isotopes in the ECHAM5 general circulation model: Toward high-resolution isotope modeling on a global scale, *J. Geophys. Res.*, *116*, D15109, doi:10.1029/2011JD015681.

Willett, S. D., F. Schlunegger, and V. Picotti (2006), Messinian climate change and erosional destruction of the central European Alps, *Geology*, *34*(8), 613–616.

Yang, X., T. Yao, W. Yang, W. Yu, and D. Qu (2011), Co-existence of temperature and amount effects on precipitation δ18O in the Asian monsoon region, *Geophys. Res. Lett.*, *38*, L21809, doi:10.1029/2011GL049353.

Yao, T., et al. (2013), A review of climatic controls on δ18O in precipitation over the Tibetan Plateau: Observations and simulations, *Rev. Geophys.*, *51*, 525–548, doi:10.1002/rog.20023.

Zhisheng, A., J. E. Kutzbach, W. L. Prell, and S. C. Porter (2001), Evolution of Asian monsoons and phased uplift of the Himalaya-Tibetan plateau since Late Miocene times, *Nature*, *411*(6833), 62–66.