Can monthly precipitation interpolation error be reduced by adding periphery climate stations? A case study in China’s land border areas
Jianyong Ma, Xiaodong Yan, Shanshan Hu and Yan Guo

ABSTRACT
Minor interpolation error of spatially continuous precipitation is increasingly in demand to support many climate studies. In this paper, based on the thin-plate smoothing splines (ANUSPLIN), we studied the effects of adding periphery stations on monthly precipitation interpolation errors in China with 184 stations from neighboring countries during 1971–2000. Here, we show that with the exception of the northern piedmont of the Himalayas, the interpolation accuracy of monthly precipitation was improved greatly in China’s border areas. Mean absolute error was reduced by an average of 2.8 mm month$^{-1}$ across 21 withheld stations. By incorporating 184 foreign stations into interpolation, the overestimated precipitation in the northern piedmont of the Himalayas can be primarily attributed to the drawback that ANUSPLIN had difficulty estimating sharply varying rain shadows in the Qinghai–Tibetan Plateau. Overall, these results mentioned above emphasized the importance of periphery stations to generate gridded precipitation datasets and the limitation of ANUSPLIN to simulate terrain-induced climate transitions.

Jianyong Ma
College of Plant Sciences and Technology, Huazhong Agricultural University, Wuhan 430070, China
and
State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing Normal University, Beijing 100875, China

Xiaodong Yan (corresponding author)
Shanshan Hu
Yan Guo
State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing Normal University, Beijing 100875, China
E-mail: yxd@bnu.edu.cn

Key words | ANUSPLIN software, interpolation error, monthly precipitation, periphery stations, thin-plate smoothing splines

INTRODUCTION
Precipitation, as the most important component of water resources on Earth, is one of the basic meteorological elements and affects human life directly and indirectly (Yatagai et al. 2012). In recent decades, a gridded precipitation dataset has become necessary for the investigation of climate change, the validation of climate models, and satellite precipitation products (Zhao & Yatagai 2014), and the detection of how climate impacts terrestrial ecosystems (Scholze et al. 2006), water resources (Thomas 2008), and hydrological processes (Haberlandt 2007).

Several interpolation methods have been developed to describe the spatial distribution of precipitation. Regions without major terrain features usually present the simplest climate pattern, for this, they could be probably handled by inverse distance weighted (Tabios & Salas 1985) and ordinary kriging (Isaaks & Srivastava 1989) if sufficient station information exists to reflect the major climate patterns (Daly 2006; McKenney et al. 2006). Situations characterized by significant terrain features can be reasonably handled by methods that explicitly account for elevation effects; these include co-kriging (Seo et al. 1999a, 1999b), ANUSPLIN (Hutchinson 1995), Daymet (Thornton et al. 1997), PRISM (Daly et al. 1994), and regional regression models (Goodale et al. 1998; Brown & Comrie 2002). Recently, wind information and terrain orientation have also been successfully used for precipitation interpolation in mountainous areas (Johansson & Chen 2005; Castro et al. 2014). Among these approaches, the outstanding advantages of ANUSPLIN are mainly shown in the following two aspects: on the one hand, a smoothing term is

doi: 10.2166/wcc.2016.126
tuned to minimize the cross-validation error automatically; on the other hand, the relationship between precipitation and elevation can vary spatially in ANUSPLIN. Therefore, this method has always been applied to generate gridded data in large domains on daily or monthly time scale (Price et al. 2000; Hijmans et al. 2005; Hong et al. 2005; Hofstra et al. 2008; Hopkinson et al. 2012).

In mountainous regions, topography has a strong effect on the magnitude and spatial distribution of precipitation, particularly in complex terrains where the depth of the moist boundary layer, rain shadow, wind speed, and direction have the most significant influence (Barros & Lettenmaier 1994; Ninyerola et al. 2000; Guan et al. 2005; Buytaert et al. 2006; Allamano et al. 2009; Castro et al. 2014). China, as one of the largest land territory countries, has mountains and hills in boundary regions with broad topographical gradients and complexity. The typical station spacing of more than 100 km in border areas is likely to be insufficient to represent actual precipitation patterns and, therefore, may directly increase the interpolation uncertainties in China’s boundary regions (Yuan et al. 2015). As noted by Daly (2006), a single station, and its particular topography regime, may impact the interpolated results for many tens of kilometers around the station; for this reason, the best approach to significantly reduce the interpolation uncertainties is to increase the number of meteorological stations (Buytaert et al. 2006; Daly 2006; Hofstra et al. 2008; Cuervo-Robayo et al. 2014; Ma et al. 2015). For example, Xie et al. (2007) used the PRISM model to construct a long-term daily gridded precipitation dataset for East Asia based on a dense observation network over different countries. Furthermore, Yatagai et al. (2012) discussed the impacts of periphery stations on daily precipitation interpolation in the Himalayas. Although these studies emphasized the importance of adding periphery climate sites during the interpolation process, to our current knowledge, it is not yet clear to what extent adding periphery climate stations could reduce monthly precipitation interpolation error, and no related studies selected China as a case study to discuss further.

The current paper initially evaluated the effects of periphery stations on monthly precipitation interpolation error in China from 1971 to 2000, and then discussed several potential limitations caused by data records and interpolation methods in our research.

DATA AND METHODS

Data

Monthly precipitation data of 610 climate stations over mainland China during 1971–2000 and 184 sites from neighboring countries over the same period was derived from the National Meteorological Information Center (NMIC) and the Global Historical Climate Network:Monthly Version 2.0 (GHCNM), respectively. In the NMIC dataset, several quality control procedures, including examining extreme values and internal consistency check, removing of questionable data, were used to guarantee the reliability of the observations by China Meteorological Administration. In the GHCNM dataset, since trace precipitation (<0.1 mm) without values was occasionally observed in some months, we filled in these gaps with a moderate value of 0.05 mm. In order to select appropriate numbers of climate stations outside China, the spatial range between periphery stations and China’s border was set to less than 600 km. Cuervo-Robayo et al. (2014) reported any reliable information from neighboring areas, to a certain extent, could improve the monthly precipitation interpolation accuracy of corresponding study areas. Therefore, the foreign stations with incomplete records were not excluded during the interpolation process in this study, such as the stations located in Nepal, where precipitation data were only available from 1971 to 1990. Overall, the 184 foreign climate sites selected from the GHCNM dataset were unevenly distributed among 12 countries: Nepal (58), Russia (40), Mongolia (33), Kazakhstan (17), Kyrgyzstan (11), Tajikistan (7), Pakistan (6), India (6), Laos (2), North Korea (2), Myanmar (1), and Vietnam (1). Corresponding spatial distribution can be found by referring to Figure 1(a). To quantitatively investigate the interpolation error in different areas, mainland China is divided into eight geographic regions that are shown in Figure 1(b).

Figure 2 reports the long-term variation of station numbers for NMIC and GHCNM datasets. Obviously, the number of GHCNM stations rapidly declined from 179 in
1975 down to 33 in 1995, while the NMIC presented a steady increase during 1971–1977 from which time onwards the number has stabilized, with a maximum being 610 from 1978 to 2000.

The ANUSPLIN model

ANUSPLIN is a suite of FORTRAN programs developed by Australian National University to generate spatial gridded climate data using thin-plate smoothing splines (Price et al. 2000; McKenney et al. 2006, 2008, 2011). According to Hutchinson & Xu (2013), the basic partial spline model for $N$ observed data values $Z_i$ is given by:

$$Z_i = f(x_i) + b^T y_i + e_i (i = 1, \ldots, N)$$  \hspace{1cm} (1)

where each $x_i$ is a $d$-dimensional vector of independent variables, $f$ is an unknown smooth function of $x_i$, each $y_i$ is a $p$-dimensional vector of independent covariates, $b$ is an unknown $p$-dimensional vector of coefficients of $y_i$ and each $e_i$ is an independent, zero mean error (ME) term with variance $w_i \sigma^2$, where $w_i$ is termed the relative error variance (known) and $\sigma^2$ is the error variance which is constant across all data points, but normally unknown. The model is reduced to an ordinary thin-plate spline model when there are no covariates ($p = 0$). In application here, ordinary spline is considered, with $x_i$ representing the coordinates for longitude, latitude, and appropriately scaled elevation (Hutchinson & Xu 2013). The function $f$ and the coefficient vector $b$ are determined by minimizing:

$$\sum_{i=1}^{N} \left( \frac{Z_i - f(x_i) - b^T y_i}{w_i} \right)^2 + \rho J_m(f)$$  \hspace{1cm} (2)

where $J_m(f)$ is a measure of the complexity of $f$, the ‘roughness penalty’ is defined in terms of an integral of $m$th
order partial derivatives of \( f \) and \( \rho \) is a positive number called the smoothing parameter. As \( \rho \) approaches zero, the fitted function approaches an exact interpolant. As \( \rho \) approaches infinity, the function \( f \) approaches a least squares polynomial, with order depending on the order \( m \) of the roughness penalty. The value of the smoothing parameter is normally determined by minimizing a measure of predictive error of the fitted surface given by the generalized cross validation (Hutchinson & Xu 2013).

Since ANUSPLIN takes the elevation as a predictor for precipitation, we preferred the digital elevation model (ASTER GDEM) with 90 m resolution to drive ANUSPLIN for generating gridded precipitation data in China (http://gdem.ersdac.jspacesystems.or.jp/) (Figure 1(b)).

### Interpolation error evaluation

We designed two interpolated experiments to evaluate the interpolation error. Experiment I was a ANUSPLIN interpolation process using only the NMIC dataset, named as ‘NMIC’, while Experiment II combined NMIC and GHCNM dataset for ANUSPLIN processing, named as ‘NMIC + GHCNM’. According to Hijmans et al. (2005), we fitted a second-order spline for precipitation interpolation, with longitude, latitude, and elevation being independent variables. Meanwhile, square root transformation, as recommended by Hutchinson & Xu (2013), was applied to reduce positive skewed values and ignore all negative values of precipitation data in the interpolation process.

We also took two quantitative indices to assess the interpolation error. First, we used all of the available climate stations (i.e., none was withheld) to interpolate monthly precipitation at a spatial resolution of 10 km by 10 km, with model standard errors (MSEs) being automatically generated by ANUSPLIN at the same resolution. MSEs related to the error in the interpolation process could be useful to reflect the uncertainty in a spatially explicit form to a certain extent (Hutchinson & Xu 2013; Cuervo-Robayo et al. 2014). It is evaluated by the derived covariance structure of the surface coefficients in the following equation:

\[
MSE = \sqrt{a_x^T \times V a_x}
\]

where \( a_x \) is a vector depending on an arbitrary position \( x \), and \( V \) represents the error covariance matrix of the surface coefficients calculated by SPLINE package of the ANUSPLIN.

Second, we withheld 21 boundary climate stations from 610 NMIC sites to compare the difference between interpolated estimations and recorded observations. Note that in this interpolation process, the estimated values at the 21 withheld sites were calculated by directly interpolating observations at the remaining stations rather than by interpolating analysis values at nearby regular grid points. The ME and mean absolute error (MAE) between estimated monthly precipitation and corresponding observations at each withheld station were calculated to assess the interpolation accuracy, and the corresponding equations are listed as follows:

\[
ME = \frac{1}{N} \sum_{i=1}^{N} (e_i - o_i)
\]

\[
MAE = \frac{1}{N} \sum_{i=1}^{N} |e_i - o_i|
\]

where \( e_i \) and \( o_i \) is the estimated and observed monthly precipitation at each withheld station, respectively; \( N \) represents the number of months.

### RESULTS

#### Evaluation by MSEs

MSEs calculated by LAPPNT package of the ANUSPLIN could help to evaluate the interpolation error for each surface significantly. However, with few exceptions (McVicar et al. 2005; Cuervo-Robayo et al. 2014), this method has not been paid enough attention to in previous studies because of the high demand for computer resources; it is particularly so for long-term interpolated operation at a high spatial and temporal resolution. In this study, we developed MSEs of monthly precipitation with climatic records from 1971 to 2000 at the 10 km by 10 km grids; the spatial multi-year averages of MSEs in unit of mm/month can be summarized in Figure 3.

Given in Figure 3(a) and 3(b), the spatial distributions of MSEs were highly consistent between the two experiments, with a steady decrease from South China (Region VI) to...
Northwest China (Region I). Note, meanwhile, that the error was moderately reduced after incorporating the GHCNM dataset into interpolation for most areas of China (Figure 3(c)). This is especially true in land border areas over Northeast China (Region III) and Northwest China (Region I), where MSEs decreased by 2–4 mm. In non-border areas, only in a small part of Central China (Region V), South China (Region VI), and Southwest China (Region VII) was a slight increment, rather than decrement, in MSEs found with the difference being 0–1 mm. However, there was a small region below the Qinghai–Tibetan Plateau (Region VIII) that we need to pay attention to, where the MSEs were significantly increased after the GHCNM dataset was used.

For further detail of the MSEs of both experiments in a seasonal pattern, corresponding results regarding monthly variation for multi-year average are shown in Figure 4. Here, we knew that higher MSEs of both experiments were found in summer (June–August) and the lowest errors occurred in winter (December–February). With respect to the difference between the two experiments, the effects varied by seasons.
More specifically, the greatest reductions in monthly MSEs were observed in August, July, May, and June with average reductions being \(-0.57\), \(-0.44\), \(-0.41\), and \(-0.39\) mm, respectively. Nevertheless, from December to March, the MSEs generated by NMIC + GHCNM were slightly higher, rather than lower, than that of NMIC. These various effects on months can be mainly attributed to the uneven seasonal distribution and high spatial variation of precipitation over China.

**Evaluation by withheld data**

Analysis of data from 21 withheld sites with a sample size of 7,560 site-months verified the equivalence relation between estimated and observed values for NMIC and NMIC + GHCNM datasets in Figure 5(a) and 5(b), respectively. Collectively, by applying the GHCNM dataset, the interpolation accuracy of monthly precipitation in boundary stations over China was greatly improved when compared with the results only from the NMIC dataset. For instance, the slopes of the regression lines became closer to 1.0, changing from 0.92 to 1.01, and Pearson correlation coefficient \((r)\) became better than that of NMIC, increasing from 0.83 up to 0.88 \((p < 0.001)\). In addition, root mean square error fell from 27.8 mm down to 25.4 mm, and the reduction almost reached 2.5 mm month\(^{-1}\). Nevertheless, of note, numerous overestimated outliers marked by the rectangle in Figure 5(b) were unexpectedly observed when the GHCNM dataset was utilized. To investigate the causes of this phenomenon more specifically, we further mapped the corresponding interpolation error at each withheld station in Figure 6.

As shown in Figure 6, the ME and MAE across 21 withheld stations were derived from the multi-year average from 1971 to 2000. In summary, parts of the results were consistent with that of the MSEs method to a certain extent. Compared with the results from NMIC, the interpolation biases from NMIC + GHCNM were reduced at most withheld stations. Note, however, that the observed values of two stations located in the northern piedmont of the Himalayas (Region VIII) were markedly overestimated after incorporating foreign stations into interpolation, with ME varying from \(-34\) to 17.4 mm month\(^{-1}\) and \(-7.1\) to

---

**Figure 4** | Climatology of monthly MSEs between NMIC and NMIC + GHCNM in China. Grey and white histograms signify the results from NMIC and NMIC + GHCNM, respectively. Error bars refer to one standard deviation and are derived from the spread of spatial grid numbers in China.

**Figure 5** | The relationship between estimated and observed precipitation for NMIC (a) and NMIC + GHCNM (b) across 360 × 21 month-stations. The solid lines represent 1:1 correspondence line and short dashes are the linear regressions fitted to the data. Grey and open circles, respectively, denote the results generated by NMIC and NMIC + GHCNM.
26.4 mm month$^{-1}$, respectively (Figure 6(a)). These extraordinary overestimations corresponded to the outliers mentioned in Figure 5(b). To better understand the actual interpolation bias, the MAE, computed station-to-station, is displayed in Figure 6(b), again referring to the results of NMIC and NMIC$^+$GHCNM datasets. With the exception of the two stations mentioned above, the technique ‘adding periphery climate stations’ allowed substantial removal of interpolation error across the remaining withheld sites. In the present study, this effect was particularly so in Northwest China (Region I) and Northeast China (Region III) with MAE reduction of 1.2–13.6 mm month$^{-1}$ and 2.8–5.7 mm month$^{-1}$, respectively (Figure 6(b)).

Similar to the MSEs method, to further detect the seasonal pattern of interpolation error over all withheld sites, the boxplots shown in Figure 7 report the monthly variations of ME and MAE for two experiments. Owing to the fact that winter exhibits the least precipitation of all seasons over the entire regions (Zhang et al. 2009; Sui et al. 2012), the box ranges of ME and MAE were the narrowest in both experiments in winter, relative to other seasons (Figure 7). In contrast, more than 50% of the total annual precipitation falls in summer in China (Sui et al. 2012), and convective processes which produce localized rainfall events could result in high spatial variability in this season; therefore, interpolation errors in summer reached the greatest with averaged MAE being 27.7 and 22.7 mm month$^{-1}$ in NMIC and NMIC$^+$GHCNM datasets, respectively (Table 1). Furthermore, compared with the former dataset, the box ranges of ME and MAE generated by the latter dataset were significantly reduced, again reflecting that numerous errors had been removed by utilizing the GHCNM dataset. Given the monthly variation of averaged ME in Table 1, although the observed precipitation from the NMIC dataset had been underestimated in most months due to the sparse observation networks near border areas, these negative biases were markedly removed in the interpolation process of NMIC$^+$GHCNM datasets. Meanwhile, an overall reduction in MAE for each month was observed after using periphery stations (Table 1), with an average reduction of 2.8 mm month$^{-1}$.

**Spatial distribution of interpolated precipitation**

Figure 8(a) and 8(b) show the multi-year average of annual precipitation for the NMIC and NMIC$^+$GHCNM datasets, respectively. The border areas of Region VII and VIII, shown in Figure 8, are moderately enlarged to be convenient for comparing the interpolation difference between two experiments.
In summary, the spatial distributions of interpolated precipitation are similar between the two experiments. The monsoon system coupled with topography generates a remarkable change for annual total precipitation (Zhai et al. 2008), from less than 100 mm in the northwestern desert interior to more than 2,000 mm on the southeastern coast (Figure 8). Furthermore, precipitation in the mountainous regions is more abundant than the plains (Sun et al. 2015) due to forced uplift and cooling of moisture-bearing winds by terrain barriers; for example, in Qaidam Basin, the annual precipitation is less than 50 mm while the value is up to 400 mm in the neighboring Qilian Mountain. However, there exists obvious interpolation differences between two experiments in border areas, such as the north of Northwest China (Region I), the east of Northeast China (Region III), and the south of Tibet Autonomous Region (Region VII and VIII). Among these border areas, using the GHCNM dataset removed substantial negative effects introduced by sparse observation coverage, and corresponding spatial distribution was more accurate than that produced only using the NMIC dataset. It was especially true in the

![Figure 7](image)

**Figure 7** Boxplots for monthly variations of ME (a) and MAE (b) between NMIC and NMIC + GHCNM. Monthly statistical results are derived from the sample size of 30 × 21 year-stations. The grey and white boxes denote the NMIC and NMIC + GHCNM, respectively. The upper and lower whiskers of each box are the 90th and 10th percentile of samples, respectively.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ME</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NMIC</td>
<td>−1.5</td>
<td>−2.0</td>
<td>−2.7</td>
<td>−3.1</td>
<td>−2.7</td>
<td>−0.4</td>
<td>2.4</td>
<td>5.8</td>
<td>−1.5</td>
<td>−3.6</td>
<td>−0.9</td>
<td>−1.5</td>
</tr>
<tr>
<td>NMIC + GHCNM</td>
<td>−0.4</td>
<td>−0.9</td>
<td>−1.1</td>
<td>1.1</td>
<td>4.5</td>
<td>8.7</td>
<td>13.7</td>
<td>13.0</td>
<td>7.4</td>
<td>0.5</td>
<td>0.2</td>
<td>−0.6</td>
</tr>
<tr>
<td>MAE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NMIC</td>
<td>5.0</td>
<td>6.3</td>
<td>8.5</td>
<td>11.5</td>
<td>16.4</td>
<td>24.9</td>
<td>30.0</td>
<td>28.1</td>
<td>18.0</td>
<td>12.7</td>
<td>6.3</td>
<td>5.6</td>
</tr>
<tr>
<td>NMIC + GHCNM</td>
<td>4.3</td>
<td>5.3</td>
<td>6.8</td>
<td>8.5</td>
<td>13.6</td>
<td>19.6</td>
<td>24.8</td>
<td>23.6</td>
<td>14.5</td>
<td>9.9</td>
<td>4.8</td>
<td>4.3</td>
</tr>
</tbody>
</table>
region of Yarlung Zangbo Grand Canyon (Region VII), where the annual precipitation increased from approximately 500–1,300 mm for the NMIC up to 800–1,700 mm for the NMIC + GHCNM. The latter result was more consistent with previous studies, such as those of Xie et al. (2007) and Yatagai et al. (2012). Note, however, that the precipitation in the neighboring south of Region VIII unexpectedly increased to 700–1,300 mm after incorporating 58 Nepal stations into interpolation. This phenomenon conflicted with the actual precipitation regimes in the northern piedmont of the Himalayas due to the fact that the Qinghai-Tibetan Plateau always markedly blocks and uplifts the moisture-bearing airflow from the Indian Ocean and, therefore, the windward slope of the Himalayas (south) receives far more rainfall than the leeward slope (north). For these reasons, the technique ‘adding periphery climate stations’ was likely to be unsuitable to interpolate precipitation in the northern piedmont of the Himalayas using the ANUSPLIN model.

**DISCUSSION**

In this paper, on the basis of ANUSPLIN software, the technique ‘adding periphery climate stations’ was used to examine the interpolation accuracy of monthly precipitation in China. Although simple, this method has removed substantial interpolation error caused by sparse coverage in border areas. Here, we argue that the climate stations only across research areas are not the best preference when generating a gridded precipitation dataset. Instead, combined climate stations with more reliable information around the periphery areas are more appropriate if the terrain-induced precipitation transitions in the interpolation process are to be properly solved. For example, when taking precipitation as an input variable for watershed hydrology models (e.g., SWAT, SWIM), the gridded data generated by hydrological stations and periphery climate sites, to some extent, could effectively improve the accuracy of final assessment results through reducing the uncertainty of input rainfall data.

**Figure 8** A visual comparison in annual precipitation for 30-year (1971–2000) average at the resolution of 10 by 10 km grid between NMIC (a) and NMIC + GHCN (b) over China. The monthly precipitation is aggregated into annual value for spatial distribution.
Nevertheless, we found that the following three major issues should be paid sufficient attention when using the ANUSPLIN model to generate gridded precipitation data. First, although the ANUSPLIN appeared to have performed reasonably well, even with a much reduced number of climate stations in a period of several decades (McKenney et al. 2006), the inconsistent station density across the time series was still an error source in the interpolation process to a certain extent. In application here, the reduced climate stations during 1991–2000 in the GHCNM dataset indeed affected the interpolation error of the stations located in the northern piedmont of the Himalayas (Figure 9). As shown in Figure 9, presumably in response to the shrinking network of periphery stations in the GHCNM dataset (Figure 2), there was an obvious trend towards reduced MAE as the years proceeded in NMIC+-GHCNM interpolation; the difference between NMIC+GHCNM and NMIC datasets notably declined from 19.3 mm month$^{-1}$ for the 1971–1990 period down to 4.7 mm month$^{-1}$ for the 1991–2000 period, demonstrating that the impacts of inconsistent station density across time series on interpolation accuracy actually varied by region in ANUSPLIN.

Second, the poor performance in the northern piedmont of the Himalayas after incorporating Nepal stations into interpolation may be attributed to drawbacks of the interpolation method. On one hand, as reported by McKenney et al. (2006), ANUSPLIN could fit thin-plate smoothing splines through all station data in three dimensions (i.e., latitude, longitude, and elevation in this study), and the relationship between precipitation and elevation can vary spatially. However, a spline is defined by smoothly varying in ANUSPLIN, which has difficulty in simulating sharply varying climate transition. In the application here, due to rain shadows in the Qinghai–Tibetan Plateau, sharp precipitation transition varies in the windward and leeward slopes of the Himalayas. On the other hand, the uncertainty from the elevation could also increase the difficulties in interpolation. In the current paper, although we used ASTER GDEM with 90 m resolution to drive ANUSPLIN, elevation estimates differ substantially among individual DEM products. Accordingly, further efforts should be focused on the interpolation of precipitation with an ensemble of DEM products (e.g., SRTM3, ACE2, and GMTED2010).

Third, following Daly et al. (1994, 2002), the interpolation method, PRISM, calculates station weight based on the extensive spatial climate knowledge pool which assesses a station’s physiographic similarity to the target gridded cell. The knowledge pool and station weighting functions could account for spatial variations in climate caused by elevation, terrain orientation, terrain barrier, and topographic position (Daly 2006). For these reasons, PRISM may be a better strategy to investigate the effects of adding Nepal’s climate stations on precipitation interpolation errors over the northern piedmont of the Himalayas. Furthermore, the spatial range of the periphery stations could affect the interpolation results. Further efforts need to focus on investigating the distance between the periphery station and border station, and by doing that, we would clearly know how the precipitation

Figure 9 | MAE temporal variation of a withheld station located in the northern piedmont of the Himalayas.
interpolation error is affected by adding periphery stations, especially for land border areas.

CONCLUSIONS

On the basis of NMIC and GHCNM datasets from 1971 to 2000, the effects of periphery stations on monthly precipitation interpolation errors in China were studied in the current paper. We summarized the main conclusions as follows:

1. MSEs of monthly precipitation, either from the NMIC or NMIC + GHCNM dataset, steadily decreased from South China to Northwest China at the resolution of 10 km by 10 km, and MSEs derived from the NMIC + GHCNM were much lower than that of NMIC in border areas, particularly in Northwest China and Northeast China.

2. With the exception of stations located in the northern piedmont of the Himalayas, the estimations generated by the NMIC + GHCNM were much closer to the observations than that of the NMIC across all withheld stations. Compared with NMIC, the MAE derived from NMIC + GHCNM was reduced by 2.8 mm month$^{-1}$, with the maximum and minimum reduction occurring in summer and winter, respectively, again demonstrating the interpolation accuracy of monthly precipitation was improved greatly in border areas through adding periphery stations.

3. The performances of interpolated precipitation are consistent between two experiments at the resolution of 10 km by 10 km, with the exception of the north of Northwest China, east of Northeast China, and south of Tibet Autonomous Region. Among these border areas, although the result generated by the NMIC + GHCNM was close to the appropriate distribution, the overestimated precipitation in the northern piedmont of the Himalayas was still substantially observed. This phenomenon can be primarily attributed to the fact that thin-plate smoothing splines in ANUSPLIN were defined by smoothly varying, and therefore had difficulty in estimating sharply varying rain shadows in the Qinghai–Tibetan Plateau.

The results mentioned above emphasized the importance of periphery stations to generate gridded precipitation datasets and the limitation of ANUSPLIN to simulate terrain-induced climate transitions.

ACKNOWLEDGEMENTS

This study was supported by the National Natural Science Foundation of China (41330527 and 31570473). Special thanks to two anonymous reviewers for their useful discussions and suggestions to our manuscript.

REFERENCES


