

Comparison of Spatial Interpolation Methods Based on Rain Gauges for Annual Precipitation on the Tibetan Plateau

Xiaoke Zhang^{1,2}, Xuyang Lu^{2,3*}, Xiaodan Wang^{2**}

¹School of Public Administration, Hohai University, Nanjing 210098, China

²Key Laboratory of Mountain Surface Processes and Ecological Regulation, Institute of Mountain Hazards and Environment, CAS, Chengdu 610041, China

³Xainza Alpine Steppe and Wetland Ecosystem Observation and Experiment Station, CAS, Xainza 853100, China

Received: 8 September 2015

Accepted: 16 February 2016

Abstract

Accurate precipitation data are of great importance for environmental applications. Interpolation methods are usually applied to afford spatially distributed precipitation data. However, due to the scarcity of rain gauges, different spatial interpolation methods may result in deviations from the real spatial distribution of precipitation. In this study, three different interpolation methods were investigated with regard to their suitability for producing a spatial precipitation distribution on China's Tibetan Plateau. Precipitation data from 39 rain gauges were spatially interpolated using ordinary kriging, cokriging with covariates as elevation (Cok-elevation), and cokriging with covariates as tropical rainfall measuring mission (Cok-TRMM). The results showed that the mean absolute error (MAE), mean relative error (MRE), and root mean square error (RMSE) for Cok-TRMM amounted to 103.85 mm, 0.32, and 134.50 mm, respectively. These numbers were lower than the figures for ordinary kriging (MAE 111.01 mm, MRE 0.34, RMSE 144.86 mm) and Cok-elevation (MAE 111.43 mm, MRE 0.34, RMSE 144.35 mm). In addition, the correlation coefficient between observed and predicted values of Cok-TRMM ($r^2 = 0.53$) was higher than that for ordinary kriging ($r^2 = 0.46$) and Cok-elevation ($r^2 = 0.46$). Our results demonstrate that Cok-TRMM is more effective at producing a spatial precipitation distribution on the Tibetan Plateau and can serve as a new spatial interpolation method for precipitation in data-scarce regions.

Keywords: spatial interpolation, annual precipitation, Cokriging, TRMM, Tibetan Plateau

Introduction

Accurate precipitation data are important for many environmental studies and are the most direct and sensitive factors of climate change. At large scales, precipitation

spatial distribution patterns, which are usually spatially interpolated from point observations, have a significant impact on water resources and ecological environment management [1]. Precipitation interpolation results are influenced by a combination of interpolation methods, and by the number and spatial distribution of meteorological stations, among other things.

A wide range of interpolation schemes is available for producing spatial precipitation distributions, ranging

*e-mail: xylu@imde.ac.cn

**e-mail: wxd@imde.ac.cn

from such simple schemes as Thiessen polygons or inverse distance weighting [2, 3] to such complex statistical methods as geostatistical ordinary kriging and cokriging [4-6]. In geostatistical methods, geostatistical multivariate approaches are widely used to predict the precipitation values over a given area by considering secondary information sampled over the same area to enhance the interpolation quality [7]. Many studies have used a cheaper, widely available data source, elevation, or other parameters extracted from digital elevation models (DEMs) in utilizing the relationship between precipitation and elevation [8, 9]. However, the reliability of data from meteorological stations is limited by their lack of ample spatial coverage. Satellite remote sensing products are increasingly used as covariates because they afford spatially detailed information about precipitation distribution [10-12]. The spatially detailed information provided by satellite-based precipitation is even more significant in the context of data-scarce regions.

The Tibetan Plateau is located in the southern region of the Qinghai-Tibet Plateau (26.30°N~ 36.20°N, 78.15°E~ 99.04°E) and covers an area of over 1.2 million km². It makes up the main part of the Qinghai-Tibet Plateau [13]. The Tibetan Plateau has an average elevation of over 4,000 m, and the terrain slopes from the northwest to the southeast. The present study employed three different spatial interpolation methods – ordinary kriging, cokriging with covariates as elevation (Cok-elevation), and cokriging with covariates as Tropical Rainfall Measuring Mission (Cok-TRMM) – to produce a spatial precipitation distribution by utilizing precipitation data from 39 rain gauges on the Tibetan Plateau. Over the course of this process, the three methods were evaluated as to their suitability for this observed data-scarce region.

Data and Methods

Rain Gauge Data

Precipitation has an uneven spatial and temporal distribution on the Tibetan Plateau. Average annual precipitation is below 400 mm in most areas and gradually decreases from the southeast to the northwest. The precipitation is mostly concentrated between May and September, a period that accounts for 90% of the annual precipitation, and the precipitation has significant inter-annual variation [14, 15].

In this study, observed precipitation data stem from the China Meteorological Data Sharing Service System (www.cdc.nmic.). The system includes data from 39 rain gauge stations from 1998-2010, a period of more than 10 years (Fig. 1). Daily data from the rain gauge stations are available and can be used to create aggregate annual precipitation totals. The annual precipitation data include traces of precipitation from snow grains, fog, and dew, among other sources.

The 39 rain gauges on the Tibetan Plateau are poorly distributed; they are relatively dense in the southeast and very sparse in the northwest (Fig. 1). The precipitation data from these 39 rain gauges are used as ground truth for evaluating the quality of the precipitation estimates obtained from the interpolation methods.

Interpolation Methods

Spatial interpolation is the generation of evaluation values or attributes for unsampled or missing locations within the area covered by existing measurements. It is most often applied as a precursor to creating contour or

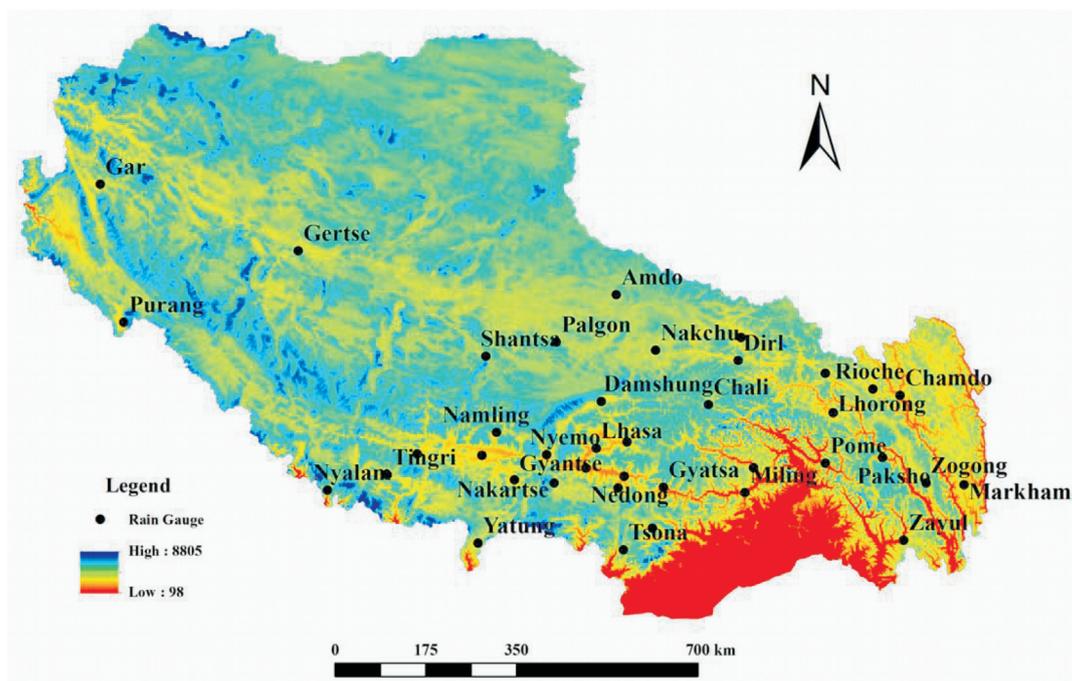


Fig. 1. Topography and location of rain gauges on the Tibetan Plateau, China.

isoline plots, the drawing of equal value lines to produce a realistic surface between measurement points [16, 17].

Interpolation methods can be grouped into deterministic and geostatistical methods. Deterministic interpolation methods create surfaces from observed points based on either the extent of similarity (e.g., IDW-Inverse distance weighted) or the degree of smoothing (e.g., CRS-Completely regularized spline). Geostatistical interpolation methods (e.g., kriging, cokriging) utilize statistical properties of the observed points. Geostatistical methods quantify spatial autocorrelation among observed points and account for the spatial configuration of sample points around the predicted location [18]. Ordinary kriging and cokriging were the geostatistical methods used in the study.

Kriging has become a primary tool in geostatistics for point interpolation. Kriging produces weights from surrounding observed values to predict values at unobserved locations. The closest observed values generally have the most influence on kriging. Kriging weights are based on a semivariogram developed from the spatial structure of the data. To create a realistic surface or map, predicted values are created for locations in the research region based on the semivariogram and the spatial distribution of nearby observed values. There are four different kriging types: ordinary kriging, simple kriging, universal kriging, and disjunctive kriging. This study relied on ordinary kriging, which is the most common and widely used kriging method. It estimates the value of climatic variables at a given point from values at surrounding stations and a variogram model for that variable [18-20].

Cokriging is similar to kriging, except that it makes use of additional covariates, such as climatic variables and elevation. Thus, cokriging takes advantage of inter-variable correlation. Cokriging is most effective when the covariates are highly correlated [20, 21]. For the Tibetan Plateau, two covariates were taken into account: i) elevation and ii) tropical rainfall measuring mission (TRMM) data for Cokriging.

Elevation is commonly used as a covariate for precipitation [22, 23]. Because of the scarce rain gauges, in some years no significant correlation was found between precipitation and elevation data for the Tibetan Plateau (Table 1). The use of Cok-elevation interpolation methods could result in differences and, hence, in deviations from the actual spatial distribution of precipitation.

We therefore used TRMM data as covariates to improve interpolant precision. The TRMM data was taken from www.trmm.gsfc.nasa.gov. TRMM data collection began in 1997 in a partnership between the U.S. National Aeronautics and Space Administration (NASA) and the Japan Aerospace Exploration Agency (JAXA). TRMM is a satellite-based program for observing tropical precipitation and helping to quantify the associated distribution and transport of latent heat, which drives the global atmospheric system. TRMM 3B42 v6 precipitation products are available at a spatial resolution of 0.25° × 0.25° and a 3-h time step [24-26]. TRMM 3B42 data are summed in order to determine the annual precipitation [27]. TRMM data yield a high correlation

Table 1. Correlation coefficient between different covariates and annual precipitation on the Tibetan Plateau (**Correlation is significant at the 0.01 level, *Correlation is significant at the 0.05 level).

Year	Elevation	TRMM
1998	-0.610**	0.837**
1999	-0.437**	0.731**
2000	-0.530**	0.733**
2001	-0.325*	0.673**
2002	-0.547**	0.750**
2003	-0.269	0.823**
2004	-0.581**	0.679**
2005	-0.262	0.812**
2006	-0.481**	0.820**
2007	-0.522**	0.786**
2008	-0.251	0.687**
2009	-0.116	0.858**
2010	-0.574**	0.761**

coefficient when compared with measured annual precipitation (Table 1).

Validation

Interpolation results must be assessed by statistics that indicate the degree of concordance between the models and reality. Through various statistical calculations, the differences between the predictions obtained by means of the different interpolation methods and the real data recorded at the rain gauges can be observed. In this paper, the interpolation results were validated via cross-validation. The cross-validation was conducted by estimating the yearly time series from 1998 to 2010 at one gauge using all other rain gauges. The cross-validation was used to verify the accuracy of the different interpolation methods. To evaluate the goodness-of-fit, we calculated the mean absolute error (MAE), the mean relative error (MRE), and the root mean square error (RMSE) [3, 16, 28]. The MRE represents the percentage of error between the observed and predicted values, while the RMSE and MAE summarize the mean difference in the units of observed and predicted values [3]. The MAE, MRE, and RMSE are defined as follows:

$$MAE = \frac{1}{n} \sum_{i=1}^n |Z_{oi} - Z_{ei}|$$

$$MRE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Z_{oi} - Z_{ei}}{Z_{oi}} \right|$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Z_{oi} - Z_{ei})^2}$$

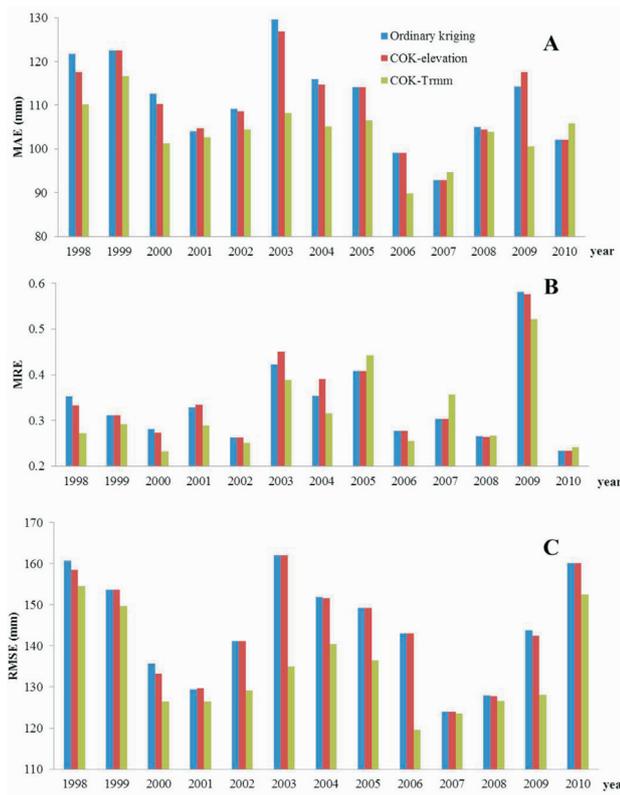


Fig. 2. Mean absolute error (MAE: A), mean relative error (MRE: B), and root mean square error (RMSE: C) for three interpolation methods.

...where Z_{oi} and Z_{ei} represent the i th observation value and predicted value, and n denotes the number of rain gauges. Smaller values of these indicators indicate that the predicted value is closer to the observed one [16, 28].

Results

The Suitability of Three Different Interpolation Methods

In general, interpolation methods that used covariates had an advantage over univariate methods. The results, based on cross-validation (which were indicated by MAE, MRE, and RMSE for the different interpolation methods), are presented in Fig. 2. Among the interpolation methods, ordinary kriging and cokriging with covariate as elevation (Cok-elevation) presented similar performances. Comparing covariates, the use of TRMM led to significantly better results than did the use of elevation pattern as a covariate.

The MAE, MRE, and RMSE values ranged from 89.87 mm to 129.63 mm, from 0.23 to 0.58, and from 119.54 mm to 162.07 mm, respectively, across the three different spatial interpolation methods used for data collected from 1998-2010 (Fig. 2). The MAE, MRE, and RMSE of ordinary kriging were close to the values for Cok-elevation, while the average values amounted

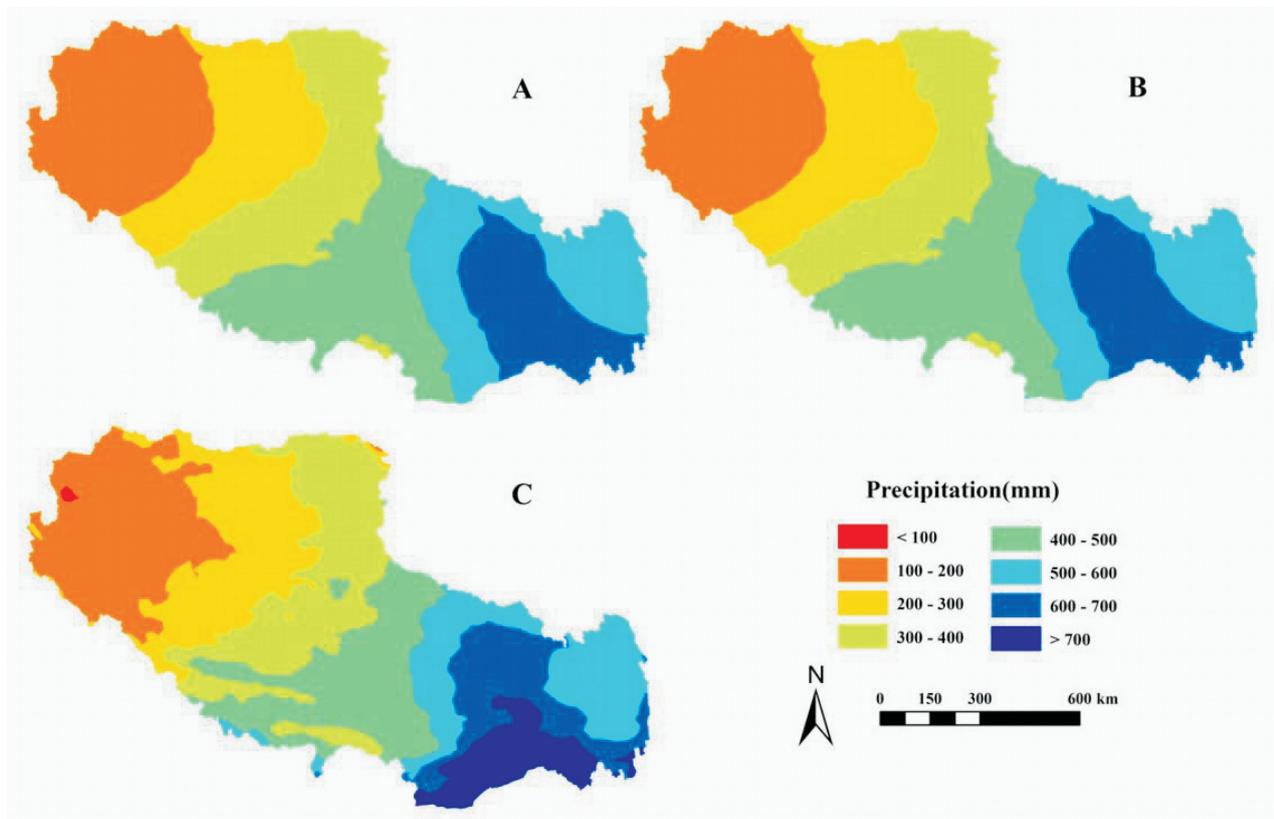


Fig. 3. Average annual precipitation distribution on the Tibetan Plateau from 1998-2010 using three interpolation methods (A: Ordinary kriging, B: Cok-elevation, C: Cok-TRMM).

to 111.01 mm, 0.34, 144.86 mm, 111.43 mm, 0.34, and 144.35 mm. The average values of MAE (103.85 mm), MRE (0.32), and RMSE (134.50 mm) were clearly lower for Cok-TRMM than the values obtained from the ordinary kriging and Cok-elevation methods. In most years, Cok-TRMM produced better results since this method yielded the lowest values of MAE, MRE, and RMSE among the three interpolation methods.

The Relationship between Observed Data and Interpolated Data

The precipitation interpolation maps of the Tibetan Plateau showed different spatial distribution patterns depending on which of the different interpolation methods was used (Fig. 3). A clear northwest-southeast precipitation gradient exists in all the maps, with high precipitation in the southeast (more than 700 mm) and low precipitation in the northwest (less than 100 mm) of the Tibetan Plateau. Ordinary kriging and Cok-elevation were generally consistent with one another, producing similar precipitation spatial distribution patterns. However, when TRMM data were used as the covariate, the precipitation distribution reflected the spatial heterogeneity of the Tibetan Plateau.

Correlation relationships between observed precipitation data (from 39 rain gauge stations) and predicted values from spatial interpolation methods are presented in Fig. 4. The straight 45° line indicates absolute coincidence between observed and predicted data. The correlation coefficient values have all passed the significance test ($p < 0.01$). Among them, the correlation coefficient value of Cok-TRMM ($r^2 = 0.53$) was higher than the correlation coefficient values for ordinary kriging ($r^2 = 0.46$) and Cok-elevation ($r^2 = 0.46$). Therefore, cokriging could improve the accuracy of precipitation interpolation for the Tibetan Plateau when TRMM data are used as a covariate.

Discussion

Improvement of interpolation results by using additive information from a covariate (Cok-TRMM) is obviously a recommendation of this study, as interpolation methods using covariates had an advantage over univariate methods (ordinary kriging). This finding is in agreement with those of other studies, in which cokriging (distance from the coastline as a covariate) was found to be the best interpolation method for the Tibetan Plateau [29]. Kriging requires a sufficient number of data points to generate a reliable semivariogram, and autocorrelation with univariate (ordinary kriging) increases almost linearly with decreasing distance. Therefore, ordinary kriging is not the method of choice for scarce rain gauges.

Cok-elevation for precipitation interpolation can also provide better results in many studies [18]. However, its superiority was not obvious when compared to ordinary kriging in this study. If high elevation precipitation

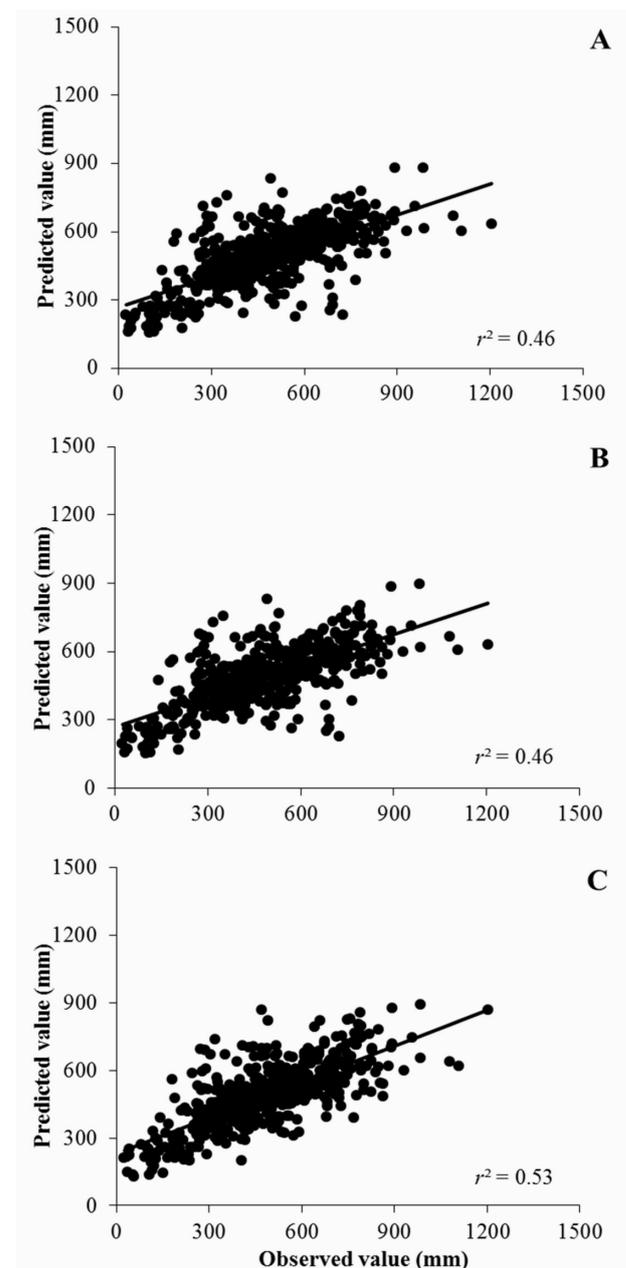


Fig. 4. Relationship between observed precipitation values (39 rain gauge stations) and predicted values using three interpolation methods (A: Ordinary kriging, B: Cok-elevation, C: Cok-TRMM) on the Tibetan Plateau from 1998-2010.

measurements are missing or large-scale processes regulate the small-scale orographic precipitation effects, a significant correlation between elevation and precipitation may not be acquired from the data. This is often the case, especially in data-scarce regions. Consequently, precipitation dependence on elevation cannot be explored in the study area; Cok-elevation thus delivered a weak performance in predicting annual precipitation of the Tibetan Plateau.

The reason for Cok-elevation's weak performance may be the low correlation between annual precipitation and elevation on the Tibetan Plateau. The highest correlation coefficient ($r = -0.61^{**}$) was obtained in 1998; thus, the

Cok-elevation method performed better than ordinary kriging that year (Fig. 2). No significant correlations were found between annual precipitation and elevation in 2003, 2005, 2008, and 2009 (Table 1). Therefore, the precision of the interpolation method that relied on elevation as a covariate demonstrated little improvement compared with ordinary kriging. This finding is consistent with that of previous research, in which DEM as a covariate was found to serve as a poor interpolation method on the Tibetan Plateau [29].

Satellite data as a covariate is increasingly used for precipitation interpolation [1]. The results demonstrate that the annual precipitation pattern obtained from TRMM data serves as a valuable covariate. Within the geographic range from 38°S to 38°N, TRMM data as an alternative to empirical covariates provides a spatial precipitation pattern. Cok-TRMM displayed the best performance across the three interpolation methods because its spatial resolution (0.25° grid) makes it superior to empirical covariates.

The reason behind Cok-TRMM's superior performance may be the high correlation between annual precipitation and TRMM data. The correlation coefficients were all over 0.65, particularly in 1998, 2003, 2005, 2006, and 2009; they even exceeded 0.80 (Table 1). In these years, Cok-TRMM performed significantly better than the other two interpolation methods. Specifically, Cok-TRMM was more effective when the covariates were highly correlated. For the MAE in 2007 and 2010 and the MRE in 2005, 2007, and 2010, the results for Cok-TRMM were relatively poor. One possible reason is that in Gar, TRMM data overestimate annual precipitation when the observed data is less than 50 mm. Therefore, the contribution of secondary data to the cokriging assessment depends not only on the correlation between principal and secondary variables but also on their patterns of spatial continuity. The TRMM pattern shows great spatial variability, resulting in certain areas (the northwest) having more and other areas (the southeast) having less precipitation. When comparing the covariates, the use of TRMM yielded significantly better results than did the use of elevation patterns as a covariate.

TRMM3B42 has a low (0.25° grid) spatial resolution and often does not meet requirements. Therefore, spatial interpolation methods thus need to be studied. The spatial resolution after interpolation can match other remote sensing satellite image data. The results can be applied to many models, including CASA and SWAT, among others. Improved interpolation accuracy can yield reliable model results that are particularly valuable in data-scarce regions. In the future, we should observe a greater degree of spatial variability, as evidenced by variable precipitation patterns.

Conclusions

Precipitation is a significant parameter in ecology, hydrology, and meteorology; it is an essential component of the global water cycle. In this study, three different

interpolation methods (ordinary kriging, cokriging with covariates as elevation, and cokriging with covariates as TRMM data) were applied to compare predicted precipitation spatial distribution on the Tibetan Plateau. An annual precipitation spatial distribution with a 250 km resolution from 1998 to 2010 was produced. Comparing and analyzing the MAE, MRE, and RMSE for three geostatistical interpolation methods indicated that the use of elevation patterns as a covariate (Cok-elevation) gives results similar to those obtained via ordinary kriging. The Cok-TRMM method, on the other hand, produces significantly better results than either of the other two methods. Thus, the use of TRMM data as a covariate could improve interpolation precision, providing a new spatial interpolation method for precipitation in data-scarce regions.

Acknowledgements

This study is funded by the Action-Plan of the Chinese Academy of Sciences (CAS) for West Development (KZCX2-XB3-08), the 135 Strategic Programs of the Institute of Mountain Hazards and Environment (sds-135-1203-03), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB03030505) and the Fundamental Research Funds for the Central Universities (2015B13714).

References

1. WAGNER P.D., FIENER P., WILKEN F., KUMAR S., SCHNEIDER K. Comparison and evaluation of spatial interpolation schemes for daily rainfall in data scarce regions. *J. Hydrol.* **464**, 388, **2012**.
2. HABERLANDT U. Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event. *J. Hydrol.* **332** (1-2), 144, **2007**.
3. VICENTE-SERRANO S.M., SAZ-SANCHEZ M.A., CUADRAT J.M. Comparative analysis of interpolation methods in the middle Ebro Valley (Spain): application to annual precipitation and temperature. *Clim. Res.* **24** (2), 161, **2003**.
4. BORGA M., VIZZACCARO A. On the interpolation of hydrologic variables: Formal equivalence of multiquadratic surface fitting and kriging. *J. Hydrol.* **195** (1-4), 160, **1997**.
5. ZHANG X.S., SRINIVASAN R. GIS-based spatial precipitation estimation: A comparison of geostatistical approaches (1). *J. Am. Water Resour. As.* **45** (4), 894, **2009**.
6. HWANG Y., CLARK M., RAJAGOPALAN B., LEAVESLEY G. Spatial interpolation schemes of daily precipitation for hydrologic modeling. *Stoch. Env. Res. Risk. A.* **26** (2), 295, **2012**.
7. FEKI H., SLIMANI M., CUDENNEC C. Incorporating elevation in rainfall interpolation in Tunisia using geostatistical methods. *Hydrolog. Sci. J.* **57** (7), 1294, **2012**.
8. GOOVAERTS P. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *J. Hydrol.* **228** (1-2), 113, **2000**.
9. LLOYD C. D. Assessing the effect of integrating elevation data into the estimation of monthly precipitation in Great Britain. *J. Hydrol.* **308** (1-4), 128, **2005**.

10. VELASCO-FORERO C.A., SEMPERE-TORRES D., CASSIRAGA E.F., GOMEZ-HERNANDEZ J.J. A non-parametric automatic blending methodology to estimate rainfall fields from rain gauge and radar data. *Adv. Water Resour.* **32** (7), 986, **2009**.
11. OUMA Y.O., OWITI T., KIPKORIR E., KIBIYI J., TATEISHI R. Multitemporal comparative analysis of TRMM-3B42 satellite-estimated rainfall with surface gauge data at basin scales: daily, decadal and monthly evaluations. *Int. J. Remote Sens.* **33** (24), 7662, **2012**.
12. KIZZA M., WESTERBERG I., RODHE, A., NTALE H. K. Estimating areal rainfall over Lake Victoria and its basin using ground-based and satellite data. *J. Hydrol.* **464**, 401, **2012**.
13. SHU S.J., YU Z.F., WANG Y., MA M. A statistic model for the spatial distribution of precipitation estimation over the Tibetan complex terrain. *Chinese J. Geophys.* **48** (3), 535, **2005**.
14. WANG M., ZHOU C.P., WU L., ZHANG G.L., OYANG H. Application of satellite-based rainfall estimate on the precipitation in Tibet Plateau. *Plateau Meteor.* **31** (5), 1215, **2012**.
15. ZHAO J.F., GUO J.P., FANG S.B., MAO F. Trends of Tibet's dry-wet condition under future climate scenario. *Chinese J. Agrometeorol.* **32** (1), 61, **2011**.
16. TEEGAVARAPU R.S.V., MESKELE T., PATHAK C.S. Geo-spatial grid-based transformations of precipitation estimates using spatial interpolation methods. *Comput. Geosci-UK.* **40**, 28, **2012**.
17. TAYLOR S.P., HAYWOOD A.M., VALDES P.J., SELLWOOD B.W. An evaluation of two spatial interpolation techniques in global sea-surface temperature reconstructions: last glacial maximum and pliocene case studies. *Quaternary Sci. Rev.* **23** (9-10), 1041, **2004**.
18. APAYDIN H., SONMEZ F.K., YILDIRIM Y.E. Spatial interpolation techniques for climate data in the GAP region in Turkey. *Clim. Res.* **28** (1), 31, **2004**.
19. CARUSO C., QUARTA F. Interpolation methods comparison. *Comput. Math. Appl.* **35** (12), 109, **1998**.
20. NALDER I.A., WEIN R.W. Spatial interpolation of climatic normals: test of a new method in the Canadian boreal forest. *Agr. Forest Meteorol.* **92** (4), 211, **1998**.
21. ASHRAF M., LOFTIS J.C., HUBBARD K.G. Application of geostatistics to evaluate partial weather station networks. *Agr. Forest Meteorol.* **84** (3-4), 255, **1997**.
22. BUYTAERT W., CELLERI R., WILLEMS P., DE BIEVRE B., WYSEURE G. Spatial and temporal rainfall variability in mountainous areas: A case study from the south Ecuadorian Andes. *J. Hydrol.* **329** (3-4), 413, **2006**.
23. VERWORN A., HABERLANDT U. Spatial interpolation of hourly rainfall-effect of additional information, variogram inference and storm properties. *Hydrol. Earth Syst. Sc.* **15** (2), 569, **2011**.
24. WOLFF D.B., MARKS D.A., AMITAI E., SILBERSTEIN D.S., FISHER B.L., TOKAY A., WANG J., PIPPITT J.L. Ground validation for the tropical rainfall measuring mission (TRMM). *J. Atmos. Oceanic Technol.* **22** (4), 365, **2005**.
25. PIERRE C., BERGAMETTI G., MARTICORENA B., MOUGIN E., LEBEL T., ALI A. Pluriannual comparisons of satellite-based rainfall products over the Sahelian belt for seasonal vegetation modeling. *J. Geophys. Res.* **116** (D18), **2011**.
26. HUFFMAN G.J., ADLER R.F., BOLVIN D.T., GU G.J., NELKIN E.J., BOWMAN K.P., HONG Y., STOCKER E.F., WOLFF D.B. The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.* **8** (1), 38, **2007**.
27. YANG Y.C., CHEN G.W., FAN J.H., SUN J., LI W.P. Accuracy validation of TRMM3B42 data in Sichuan basin and the surrounding areas. *Scientia Meteor. Sinica.* **33** (5), 526, **2013**.
28. WANG Z., WU Y.J., LIANG F.C., CHANG S.L., SHI Q.D. Study on spatial interpolation method of annual precipitation in Xinjiang. *Chinese J. Agrometeorol.* **32** (3), 331, **2011**.
29. LIU Z.J., YU X.X., WANG S.S., SHANG G.D. Comparative analysis of three covariates methods in thin-plate smoothing splines for interpolating precipitation. *Prog. Geogr.* **31** (1), 56, **2012**.