Journal of Global Change Data & Discovery. 2019, 3(4): 343-348
DOI:10.3974/geodp.2019.04.05
www.geodoi.ac.cn

Global Change Research Data Publishing & Repository

Spatial Distribution Dataset of Annual Precipitation in Midwestern China (2010)

Jiang, Y. H.[1,2] Li, B. L.[1,2*] Yuan, Y. C.[1] Gao, X. Z.[1] Zhang, T.[1,2] Liu, Y.[1,2]
Li, Y.[1,2] Li, H.[3] Luo, Z. Y.[3] Ma, Q.[3] Wang, X. M.[3] Ciren, D. J.[4]

1. State Key Laboratory of Resources and Environmental Information Systems, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China;
2. University of Chinese Academy of Sciences, Beijing 100049, China;
3. Meteorological Bureau of Mongolian Autonomous County of Henan, Tibetan Autonomous Prefecture of Huangnan, Henan 811599, China;
4. Lhasa Meteorological Bureau, Lhasa 850000, China

Abstract: The dataset on the spatial distribution of annual precipitation in midwestern China (2010) is a simulation of regional precipitation with High Accuracy Surface Modeling (HASM). The method includes three steps: First, use the TRMM data (or station observations spatial interpolation results) to represent the stable change of precipitation in space, that is, the trend surface; second, calculate residue combined with station observations after removing the trend surface, and then incorporate residue in HASM to form unstable changed residual field in space; finally, add the trend surface and residual field to complete incorporation. The results show that HASM model with TRMM as the background field is significantly more precise and adaptive than traditional interpolation methods. Mean absolute error (MAE) and root mean square error (RMSE) are the accuracy evaluation indicator. The MAE and RMSE of TRMM based HASM simulation were 125.15 mm and 155.80 mm on a global scale and 167.53 mm and 228.81 mm on a local scale, respectively. Related research results were published in *Journal of Geo-Information Science* (Vol. 17, No. 8, 2015).

Keywords: precipitation; TRMM satellite; high accuracy surface modeling; midwestern China; Journal of Geo-Information Science

1 Introduction

Precipitation is an important environmental variable and plays an essential role in surface runoff, atmospheric movement, and agricultural resources. Large-scale precipitation data are usually based on the results of discrete site observations and are obtained using spatial interpolation. The traditional spatial interpolation method usually neglects the characteristics of the

Received: 07-11-2019; **Accepted:** 09-12-2019; **Published:** 24-12-2019
Foundations: Ministry of Science and Technology of P. R. China (2016YFC0500205, 2015CB954103)
***Corresponding Author:** Li, B. L. N-8884-2019, State Key Laboratory of Resources and Environmental Information Systems, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, libl@lreis.ac.cn
Data Citation: [1] Jiang, Y. H., Li, B. L., Yuan, Y. C., *et al*. Spatial distribution dataset of annual precipitation in midwestern China (2010) [J]. *Journal of Global Change Data* & *Discovery*, 2019, 3(4): 343–348. DOI: 10.3974/geodp.2019.04.05.
[2] Jiang, Y. H., Li B. L., Yuan, Y. C., *et al*. 0.25° Grid Precipitation Dataset covering midwestern China (2010) [DB/OL]. Global Change Research Data Publishing and Repository, 2019. DOI: 10.3974/geodb.2019.05.18.V1.

spatial surface itself, and ignores the constraint effect of the intrinsic factors on the surface reconstruction when modeling. High Accuracy Surface Modeling (HASM) can effectively resolve the issue of peak flattening and boundary turbulence that are inevitable among traditional interpolation methods[1]. However, in areas with fewer stations, the feasibility of the HASM model is greatly restricted. Satellite remote sensing data allows for large-scale and synchronous observation, which can make up for the limited observation range of ground stations. Therefore, the accuracy of spatial simulation of regional precipitation could be improved by integrating satellite precipitation product information while simulating precipitation with HASM[2].

The dataset[3] was developed on the basis of the readily available TRMM 3B43 V7 data product with a resolution 0.25°×0.25°[4–5]. To resolve the issue of poor HASM accuracy in regions with fewer stations, the TRMM data were set as the background field (trend surface) by using hybrid interpolation, and the residual field was revised with HASM (after removing the trend) to improve HASM's depicting capacity of spatial details for regional precipitation. The revised product was compared with the results generated from ground station interpolation to validate its effectiveness.

2 Metadata of Dataset

The metadata for "0.25° grid precipitation dataset covering midwestern China (2010)"[3] is summarized in Table 1. It includes the name, authors, geographical region, year of the data, temporal resolution, spatial resolution, data files, data publisher, and data sharing policy, etc.

Table 1 Metadata summary for "0.25° grid precipitation dataset covering midwestern China (2010)"

Items	Description
Dataset full name	0.25° grid precipitation dataset covering midwestern China (2010)
Dataset short name	PrecipMidwesternChina2010
Authors	Jiang, Y. H. N-8765-2019, State Key Laboratory of Resources and Environmental Information Systems, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, jiangyh@lreis.ac.cn Li, B. L. N-8884-2019, State Key Laboratory of Resources and Environmental Information Systems, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, libl@lreis.ac.cn Yuan, Y. C. N-9047-2019, State Key Laboratory of Resources and Environmental Information Systems, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, yuanyc@lreis.ac.cn Gao, X. Z. N-1655-2019, State Key Laboratory of Resources and Environmental Information Systems, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, gaoxz@lreis.ac.cn Liu, Y. N-8844-2019, State Key Laboratory of Resources and Environmental Information Systems, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, liuy.18b@igsnrr.an.c.cn Li, Y. Y-4384-2019, State Key Laboratory of Resources and Environmental Information Systems, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, liying9391@126.com Luo, Z. Y., Meteorological Bureau of Mongolian Autonomous County of Henan, Tibetan Autonomous Prefecture of Huangnan, 393352158@qq.com Ma, Q., Meteorological Bureau of Mongolian Autonomous County of Henan, Tibetan Autonomous Prefecture of Huangnan, 107082968@qq.com Wang, X. M., Meteorological Bureau of Mongolian Autonomous County of Henan, Tibetan Autonomous Prefecture of Huangnan, 2444869807@qq.com Ciren D. J., Lhasa Meteorological Bureau, LSNSE111@126.com
Geographical region	25°N–35°N, 105°E–115°E, including Chongqing, Guizhou, Hunan, Hubei, the majority of Henan, eastern Sichuan, southern Shaanxi and southeastern Gansu, with total area of 1.06×10^6 km^2

(*To be continued on the next page*)

(*Continued*)

Item	Description
Year	2010 Temporal resolution 1 month
Spatial resolution	0.25°×0.25°
Data format	.tif
Data size	48.8 KB (after compression)
Data files	The dataset is composed of 20 files
Foundations	Ministry of Science and Technology of P. R. China (2016YFC0500205, 2015CB954103)
Computing envi-roment	Matlab 2011b; ArcGIS campus license of Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences
Data publisher	Global Change Research Data Publishing & Repository, http://www.geodoi.ac.cn
Address	No. 11A, Datun Road, Chaoyang District, Beijing 100101, China
Data sharing policy	***Data*** from the Global Change Research Data Publishing & Repository includes metadata, datasets (data products), and publications (in this case, in the *Journal of Global Change Data & Discovery*). ***Data*** sharing policy includes: (1) ***Data*** are openly available and can be free downloaded via the Internet; (2) End users are encouraged to use ***Data*** subject to citation; (3) Users, who are by definition also value-added service providers, are welcome to redistribute ***Data*** subject to written permission from the GCdataPR Editorial Office and the issuance of a ***Data*** redistribution license; and (4) If ***Data*** are used to compile new datasets, the 'ten percent principal' should be followed such that ***Data*** records utilized should not surpass 10% of the new dataset contents, while sources should be clearly noted in suitable places in the new dataset[6]
Communication and searchable system	DOI, DCI, CSCD, WDS/ISC, GEOSS, China GEOSS

3 Methods

3.1 Algorithm Principles

Precipitation interpolation results are more uncertain in sparse stations and in areas with greater ground surface change, and remote sensing precipitation products could well manifest the spatial distribution patterns of precipitation. Therefore, in this study, we combine these two data sets together based on HASM model to estimate regional rainfall. First, using TRMM data (or station observations spatial interpolation results) to represent the stable change of precipitation in space, that is, the trend surface; second, calculate the residue combined with station observations after removing the trend surface, and then incorporate the residue in HASM to form unstable changed residual field in space; finally, add the trend surface and residual field to complete incorporation.

3.2 Technical Route

In this interpolation method, residual field was simulated in HASM. IDW, Spline, Kriging interpolation results and TRMM data were applied to calculate the trend surface, which were expressed as HASM_I, HASM_S, HASM_K and HASM_T, respectively. HASM model equation sets were resolved by iteration with Preconditioned Conjugate Gradient[7] (Figure 1).

Randomly-selected weather stations were used as modeling points, and the rest stations were global validation point. Both of them were used to test the model simulation accuracy in large scale. Without changing the modelling points, all local weather stations were used to test the accuracy in small scale. The background field was simulated by using HASM with TRMM data, IDW, Spline and Kriging interpolation results separately, and the simulation accuracy was compared.

4 Results and Validation

4.1 Data Products

The dataset is composed of 20 files. Among them, hasm_idw2010.tif, hasm_kriging-2010.tif, hasm_spline2010.tif, and hasm_tr-mm2010.tif represent the 2010 precipitation data with IDW interpolation, Kriging interpolation, Spline interpolation, and TRMM data results as the HASM drive field and using hybrid interpolation.

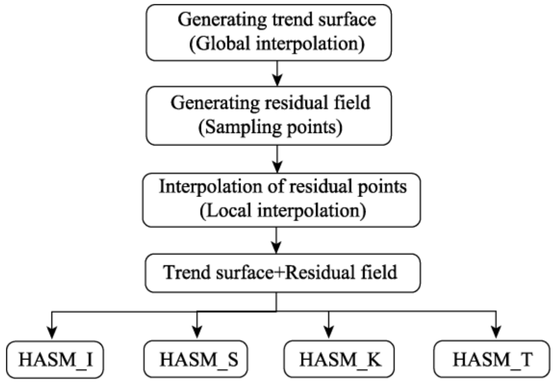

Figure 1 Technical route for research and development of the dataset

4.2 Data Results

Similar spatial distribution patterns of precipitation existed among the four methods (Figure 2). Differences in spatial distribution of HASM simulation with different interpolations were insignificant, and precipitation showed a decrease trend from southeast to northwest overall.

HASM simulation results with TRMM as the drive filed had more information on spatial changes in precipitation in regions without station observation, which were superior to other HASM simulation results with traditional interpolation methods. Place A in Figure 2 is located in the rapid change area from humid and semi-humid region to semiarid region, and precipitation should have a decrease trend from southeast to northwest. Among the four methods, only HASM-T simulated this change. Place B in Figure 2 is in Weihe Plain, and precipitation distribution simulated in HASM_T was consistent with the previous studies. Place C is in the upstream of Lishui River on the northern branch of Wuling Mountains, and is one of the heaviest precipitation of Hunan province. Place D is in Hengyang-Shaoyang hugelland region, a "arid area" in Hunan province[9]. HASM_T distinctly reflected the precipitation of these regions.

4.3 Validation

It can bee seen from Figure 2 that accuracy of HASM simulation with TRMM as the drive filed is superior to that of HASM simulation with traditional interpolation methods as the drive field. In the global validation results, the MAE and RMSE of HASM_T were 125 and 156 mm, respectively (Table 2). The MAEs of HASM_I, HASM_S and HASM_K were 212, 234 and 192 mm, 70%, 87%, and 54% higher than that of HASM_T, respectively; the RMSEs were 260, 328 and 241 mm, 67%, 110%, and 54% higher than that of HASM_T, respectively. In the local validation results, the MAE and RMSE of HASM_T was 168 and 229 mm, respectively. The MAEs of HASM_I, HASM_S and HASM_K were 196, 198, and 197 mm, 17%, 18% and 17% higher than that of HASM_T, respectively; the RMSEs were 263, 260, and 256 mm, 15%, 14% and 12% higher than that of HASM_T, respectively.

Table 2 HASM calculation accuracy in different background fields (mm)

Validation point	HSAM_I		HSAM_S		HSAM_K		HSAM_T	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
Global	212	260	234	328	192	241	125	156
Local	196	263	198	260	197	256	168	229

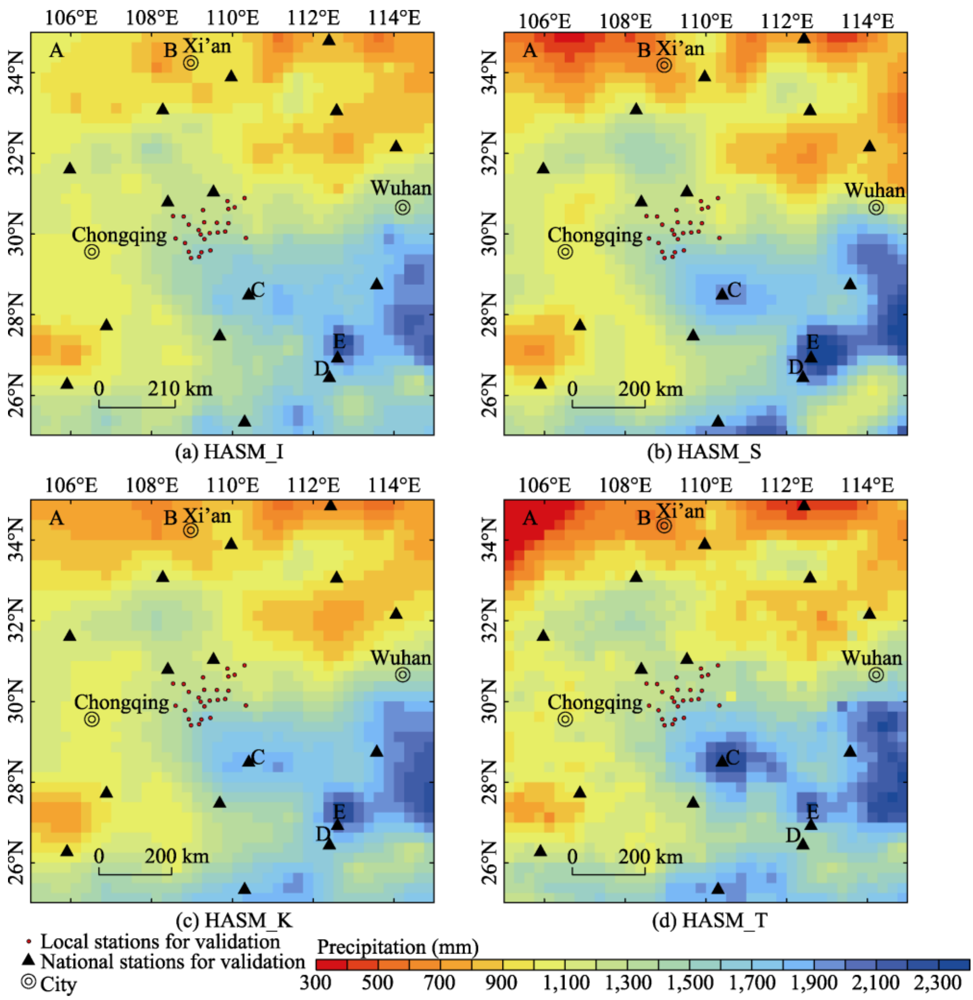

Figure 2 Precipitation distribution simulated by HASM with different background fields

5 Discussion and Conclusion

In this study, TRMM satellite precipitation data was used as the background field and the residual field was revised by HASM model to estimate regional rainfall. Midwestern China, which are featured by extensive high mountains and plains, is chosen as the study area to model the spatial distribution of its total precipitation in 2010. HASM simulation with TRMM as the background field is significantly more accurate than traditional interpolation methods. On a global scale, MAE and RMSE were 125.15 and 155.80 mm, respectively; on a local scale, MAE and RMSE were 167.53 and 228.81 mm, respectively. HASM simulation results with TRMM as the background field at local region reflected the spatial patterns of precipitation better, and the results not only showed the rapid changes in precipitation, reflecting peak precipitation areas, but also effectively avoided extreme value caused by observations in traditional spatial interpolation algorithms. Besides, its relative error in each sub-region was also smaller than the other methods. At present, in the relevant applications of the spatial interpolation of meteorological elements, the shortage of data caused by sparse stations is inevitable, which leads to uncertainties in the simulation results of meteorological elements. Therefore, combining the HASM model with other auxiliary data that can reflect the results of the spatial distribution of the meteorological elements to make up the shortage of the number of available sampling points is the following research frontier of HASM.

Author Contributions

Li, B. L. and Jiang, Y. H. designed the dataset development; Liu, Y. and Li, Y. processed the TRMM precipitation data; Zhang, T. and Yuan, Y. C. designed the model and algorithm; Gao, X. Z. performed data validation; Li, H., Luo, Z. Y., Ma, Q., Wang, X. M., and Ciren, D. J. provided the weather station data. Jiang, Y. H. and Li, B. L. wrote the data paper.

References

[1] Zhao, N., Jiao, Y. M. Downscaling of TRMM satellite precipitation data [J]. *Journal of Geo-Information Science*, 2018, 20(10): 20–27.

[2] Zhang, T., Li, B. L., Zhao, N., *et al*. Analysis on high accuracy surface modeling in regional rainfall estimation combined with TRMM data [J]. *Journal of Geo-Information Science*, 2015, 17(8): 895–901.

[3] Jiang, Y. H., Li, B. L., Yuan, Y. C., *et al*. 0.25° grid precipitation dataset covering midwestern China (2010) [DB/OL]. Global Change Research Data Publishing and Repository, 2019. DOI: 10.3974/geodb.2019.05.18.V1.

[4] Kummerow, C., Barnes, W. The tropical rainfall measuring mission (TRMM) sensor package [J]. *Journal of Atmospheric and Oceanic Technology*, 1998, 15: 809–817.

[5] Huffman, G. J., Bolvin, D. T., Nelkin, E. J., *et al*. The TRMM multisatellite precipitation analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales [J]. *Journal of Hydrometeorology*, 2007, 8(1): 38–55.

[6] GCdataPR Editorial Office. GCdataPR data sharing policy [OL]. DOI: 10.3974/dp.policy.2014.05 (Updated 2017).

[7] Golub, G. H., van Loan, C. F. Matrix Computations [M]. Beijing: Posts & Telecom Press, 2009.

[8] Zhao, X., Ma, J. Z., Nie, Z. Q., *et al*. Spatiotemporal pattern of precipitation in Tianshui region and its response to the Indian Ocean dipole [J]. *Arid Zone Research*, 2009, 26(6): 917–922

[9] You, X., Li, C., Yu, W. H. Analysis of precipitation characteristics in Hunan province [J]. *Hunan Hydro & Power*, 2004(6): 41–43.