

# Applicability and Correction of ERA - Interim Reanalysis Precipitation Dataset in the Upper Reaches of the Yangtze River

Xiaofan Zeng, Qifang He\*, Na Zhao, Yiran Bai and Ruoyu Zhang

School of Hydropower & Information Engineering, Huazhong University of Science and Technology, Wuhan 430074, China

\*Corresponding author

**Abstract**—Based on the reanalysis data of ERA-interim, combined with the precipitation data of the measured sites in the basin, the reanalysis data in the upper reaches of the Yangtze River is studied. Calculating watersheds as well as the station annual precipitation, multi-year average precipitation. Using the indicators of precipitation deviation, percentage deviation, and qualified rate to figure out the applicability of ERA data in the basin in order to supplement the measured precipitation data. The results show that the trend of ERA precipitation in the year, spring, summer, autumn and winter is almost the same as that of the measured data, and the annual precipitation data of ERA can be corrected by using the conversion coefficient directly. In addition, the grid precipitation interpolation calculation of 65 measured sites in the basin is carried out to analyze the difference between ERA data and measured data. Calculating multi-year average precipitation of two datasets for each station, there is a significant difference between the estimated ERA precipitation and the measured precipitation at the high altitude in the basin, and the difference between the estimated ERA precipitation and the measured precipitation at the low altitude in the basin is smaller, and the precipitation deviation in the southeastern of the basin is smaller than that in the northwest of the basin. Using the conversion coefficient to convert annual, spring, summer, autumn and the winter scales of ERA interpolation precipitation data for each station, within the 30% deviation error threshold (compared with measured precipitation), the qualified rates for stations of the annual, spring, summer, autumn and winter scales after the fitting are 98.46%, 72.31%, 92.31%, 70.77% and 9.23% respectively.

**Keywords**—ERA-interim grid precipitation; the upper reaches of the Yangtze River; accuracy correction

## I. INTRODUCTION

In recent years, with the development of related technologies in the field of specialization, a large amount of precipitation products have emerged with satellite inversion, data assimilation and pattern operation. The reanalysis data set combines the observation data of ground, satellites and so on, can provide the hydrological meteorological elements with spatial and temporal continuity through the pattern assimilation, which can make up for the deficiency of traditional precipitation observation to a certain extent.

Since the 20th century, the experts have been devoting themselves to improving long-term precipitation data. By the middle of the 21st century, several representative precipitation

series precipitation data have been established around the world[1]. High-precision grid point precipitation data is an important input parameter for various hydrological models[2], and also it's necessary for the detection of forecasting models in scarce areas as well as important data sources for hydrological simulation calculations. The analysis and reanalysis precipitation data can make up for station observations, which shows important application values in the field of meteorology and hydrology[3].Jia compared the measured data with the earlier analysis of precipitation data (Humble data), the results show that the grid data to some extent can solve the problems of uneven distribution of stations and station accidental absence of records[4]. High-resolution monthly mean climatological data from the CRU in the UK indicate that the annual precipitation in China is consistent with the measured data between 1950 and 2000[5], the reconstruction data keeps highly consistency. Zhao et al[6,7]. using the TRMM multi-satellite precipitation data for the research of regional applicability, Adler et al[8],using monthly average precipitation results of 31 kinds of methods for comparative analysis, Sun et al[9], using different reanalysis datasets to evaluate the accuracy of the data in the Dongting Lake basin, the spatial dispersion of ERA-interim reanalysis data was significant, but the annual and seasonal precipitation deviations were relatively stable. Therefore, in order to further understand the applicability of reanalysis precipitation data released by ECMWF in China, in this paper, we selected the upper reaches of the Yangtze River as the research area, and analyzed the annual, seasonal variation trends and spatial distribution characteristics of ERA-interim precipitation data from the perspective of time and space. Based on the measured precipitation data of 65 sites in the basin, striving for being applied in practical work.

The Yangtze River is the largest river in China. The economic belts, resource belts and industrial belt formed by many large, medium and small cities in the basin, which plays an important role in the sustainable development of our society and economy. However, due to the combination of natural factors such as climate, geographical location and topography and human activities, there are many problems such as floods and waterlogging in the Yangtze River basin. In particular, the Yangtze River Basin has suffered several floods since the 1990s. The flooding in the upper reaches of the Yangtze River is mainly caused by heavy rain, and the flood time is basically consistent with the heavy rainy season[10]. In order to

reduce the loss caused by floods, this study focuses on the ERA-interim data released by ECMWF. Comparing with the precipitation data from the observed sites in the study area, and analyzing the applicability of ERA-interim data in the upper reaches of the Yangtze River from the time and space scales, and calibrating them to supplement the rainfall data in the basin, which helps to solve the problem of inaccurate precipitation studies due to the spatial distribution of the stations in the upper reaches of the Yangtze River.

## II. RESEARCH AREA AND DATA BASE

### A. Research Area

The research area of this paper is the upper reaches of the Yangtze River, located between  $90^{\circ} \sim 115^{\circ} \text{ E}$  and  $25^{\circ} \sim 36^{\circ} \text{ N}$ , starting from the east of the Lhasa in the Qinghai-Tibet Plateau to the east of Yichang, Hubei. It's about 4,500 km long, the controlling area of basin is about 1 million square kilometers[11], the regional climate belongs to the Qinghai-Tibet alpine zone and the subtropical monsoon region, it's rainfall is mainly from the plateau snowmelt. The Branch of the Yangtze River are Jinsha River, Yalong River, Minjiang River, Jialing River, Wujiang etc. The terrain of the area is complex, the precipitation space and time distribution vary greatly. A total of 65 sites surveyed in the basin were selected in this study. The distribution of weather stations is shown in Figure 1.

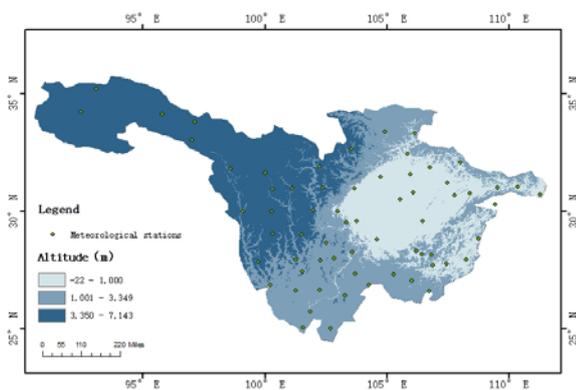


FIGURE 1. SKETCH OF STUDY AREA

### B. Data Base

The daily precipitation data of 65 ground meteorological stations in the upper reaches of the Yangtze River are provided by the meteorological data center of the National Meteorological Center of China Meteorological Administration. The precipitation data are the daily precipitation from 1979 to 2012.

The ERA-interim reanalysis dataset is the most recent global reanalysis data released by the European Center for Medium Range Weather Forecasts (ECMWF). Compared with the previously released ERA-40, the data is improved in terms of resolution, model physical parameters, four-dimensional variational method, error correction, etc. In order to maintain consistency with the time scale of measured meteorological data, daily grid precipitation data from 1979 to 2012 are

selected with a spatial resolution of  $0.25^{\circ} * 0.25^{\circ}$ , hereinafter referred to as ERA data.

The DEM data uses SRTM data of  $90\text{m} * 90\text{m}$  resolution.

## III. METHODOLOGY

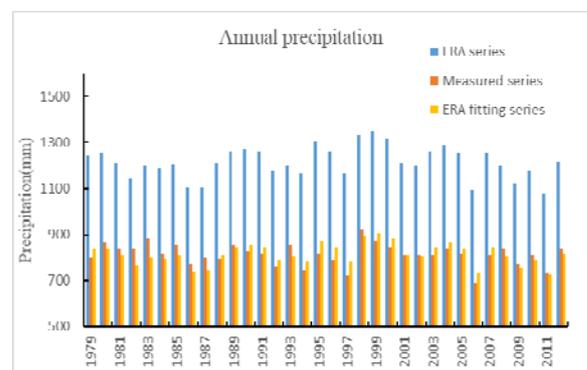
In this study, annual, seasonal and multi-year average precipitation of two datasets at each site were calculated to analyze the trend of annual and seasonal changes of two datasets.

Two indicators of deviation absolute value and deviation percentage were used to analyze the difference of precipitation between the two sets of data. In addition, the accuracy of ERA data was corrected by using the conversion coefficient method according to the characteristics of precipitation difference between the two datasets. Absolute deviation and deviation percentage reflect the difference between the ERA data and the measured data as well as the proportion. The conversion coefficient is used for ERA data correction. It refers to the correction of the ERA data multiplied by some conversion coefficient according to the difference between precipitation in the basin and measured precipitation.

## IV. RESULTS

### A. Time Scale

Calculating annual and seasonal precipitation of two datasets, and the trend of precipitation in the basin is shown in Figure 2 below. It can be seen from Figure 2 (a-e) that annual and seasonal ERA reanalysis precipitation data are both higher than the measured data. As can be seen from the curves in the figure, the ERA precipitation data are numerically larger than the measured precipitation, and annual, spring, summer, autumn, winter precipitation increases as high as 50%, 81%, 27%, 55% and 187% respectively. Although there is a large difference in two datasets on the magnitude, the trend of datasets over the years in time scales is similar. Therefore, the annual and seasonal timescales grid precipitation can be directly converted using the conversion coefficient in order to make up for the lack of the measured data in practical study.



(a)

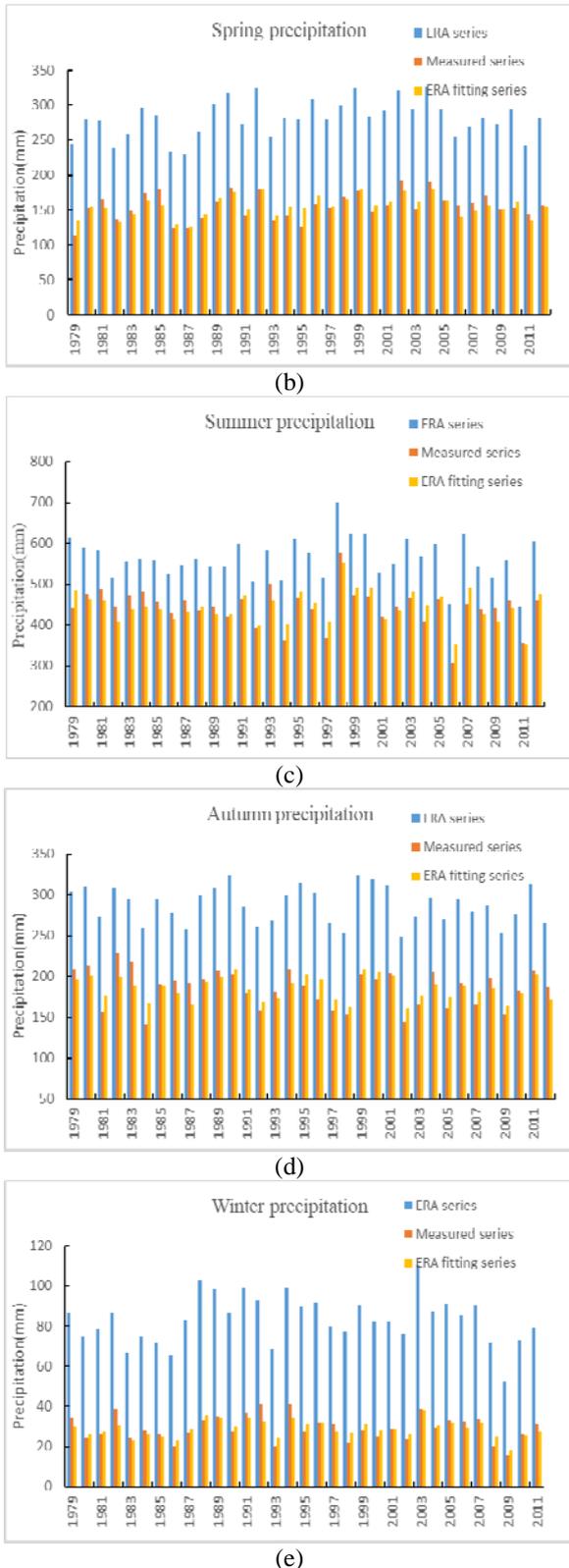


FIGURE II. PRECIPITATION OF DIFFERENT TIME SCALES IN THE UPPER REACHES

B. Space Scale

Using the inverse distance weight interpolation module in ArcGIS, the grid precipitation data are interpolated to 65 observation stations to analyze the applicability of ERA reanalysis data in the upper reaches of the Yangtze River. According to two datasets of sites for many years the average precipitation, calculate the deviation of precipitation and the percentage deviation, the spatial distribution of the results shown in Figure 3. Compared with the figure 1, the estimated ERA precipitation at the observation site with high altitude in the basin significantly increases with the measured precipitation. The estimated ERA precipitation at the observation site with relatively low elevation is relatively small compared with the measured precipitation, and the precipitation deviation in the southeastern part of the basin is smaller than that in the northwest part of the basin.

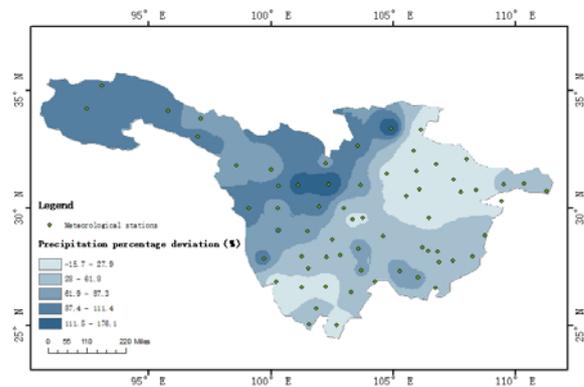


FIGURE III. MULTI- YEAR AVERAGE PRECIPITATION DEVIATION PERCENTAGE OF EACH STATION

C. Grid Precipitation Correction

The advantage of ERA data is that it can reflect the spatial distribution of precipitation. Its deficiency is the poor accuracy of the data. The advantage of the measured data is the high rainfall accuracy, but there are some limitations in reflecting the characteristics of rainfall-time-space variation. Comparing the spatial distribution of grid precipitation with the measured precipitation, we modify the grid rainfall data. According to the observed precipitation of each ground observation station, the conversion coefficient method is used to correct the grid precipitation. The probability of rainfall relative error exceeding 30% for the measured precipitation and grid interpolation precipitation in 65 stations on the annual and seasonal time scales is counted.

The annual and summer series of ERA data at each site showed good fitting effects. The average annual precipitation deviation was 8.14% and the average percentage deviation of summer precipitation was 19.1%. The results are smaller than that of spring, autumn and winter scales. In addition, the qualified rates of the site was calculated, within the precipitation deviation percentage of 30%, and the annual fitting series was as high as 98.46%. The qualified rates of the fitting series in each season (spring, summer, autumn and winter) was 72.31%, 92.31%, 70.77% and 9.23%. The result is that the best is for summer, followed by spring and autumn,

and the winter precipitation with large uncertainties. Overall, the corrected grid precipitation can be used to supplement the measured precipitation and improve the accuracy of historical precipitation datas.

## V. CONCLUSIONS

Based on the comparison of the ERA-interim reanalysis precipitation and 65 sites measured precipitation provided by the meteorological data center of the National Meteorological Center of China Meteorological Administration, analyzing the applicability of the ERA precipitation data in the upper reaches of the Yangtze River, the following conclusions are drawn.

For annual and seasonal precipitation, the ERA data are larger than the measured precipitation data, but the trend of the two datasets are almost the same. The fitted ERA series calculated by using the conversion coefficient has good fitting effect with the measured series.

The estimated ERA precipitation at the observation site with high altitude in the basin significantly increases with the measured precipitation. The estimated ERA precipitation at the observation site with relatively low elevation is relatively small compared with the measured precipitation, and the precipitation deviation in the southeastern part of the basin is smaller than that in the northwest part of the basin.

After converting the ERA data of annual and seasonal, the qualified rates (deviation error less than 30%) for stations of the annual, spring, summer, autumn and winter scales after the fitting are 98.46%, 72.31%, 92.31%, 70.77% and 9.23% respectively.

## ACKNOWLEDGEMENTS

This research has been financially supported by the national Key Research and Development Program (2016YFC0401005, 2016YFC0401005), and the Funds for the Central Universities, HUST (2016YXZD046, 2017KIFYXJ 191).

## REFERENCES

- [1] Qiang Zhang, Xin Ruan, Anyuan Xiong. Establishment and assessment of the grid air temperature data sets in China for the past 57 years, *Journal of Applied Meteorological Science*. 20(04), pp.385-393, in 2009
- [2] Yan Shen, Mingnong Feng, Hongzheng Zhang, Feng Gao. Interpolation methods of China daily precipitation data, *Journal of Applied Meteorological Science*. 21(03), pp.279-286, in 2010
- [3] Shaoning Lv, Jun Wei, Rong Liu. Applicability and potential of the different precipitation data in mainland China. *Plateau Meteorology*. 30(03), pp.628-640, in 2011
- [4] Pengqun Jia. Comparison between observational data and grid data of precipitation for the last one hundred years in China, *Journal of Applied Meteorological Science*, 02, pp.54-62, in 1999
- [5] Xinyu Wen, Shaowu Wang, Jinhong Zhu, David VINER. An overview of China climate change over the 20th century using UK UEA/CRU high resolution grid data, *Chinese Journal of Atmospheric Sciences*, 05, pp.894-904, in 2006
- [6] Yuan Zhao, Li Zhao, Tian Xiao, SanJun Zhao, Jiliang Xuan, Chaolun Li, et al. Spatial and temporal variation of picoplankton distribution in the Yellow Sea, China[J]. *Chinese Journal of Oceanology and Limnology*, 29(01), pp.150-161, in 2011
- [7] Hongwei Zeng, Lijuan Li. Accuracy validation of TRMM 3B43 data in Lancang River Basin, *Acta Geographica Sinica*, 66(07), pp.994-1004, in 2011
- [8] Adler R F, Kidd C, Petty G, et al. Inter-comparison of global precipitation products: The third precipitation inter-comparison project (PIP-3)[J]. *Bull Amer Meteor Soc*, 82 (7):, pp.1377-1396, in 2001
- [9] Jia Sun, Xinping Zhang, Yimin Huang. Evaluation of precipitation from ERA-INTERIM, CRU, GPCP and TRMM reanalysis data in the Dongting Lake Basin, *Resources and Environment in the Yangtze Basin*. 24(11), pp.1850-1859, in 2015
- [10] Lujun Zhang, Yongfu Qian. A study on the feature of precipitation concentration and its relation to flood-producing in the Yangtze River Valley of China, *Chinese Journal of Geophysics*, 04, pp.622-630, in 2004
- [11] Changchun Tang, Wei Sun. Comprehensive evaluation of land spatial development suitability of the Yangtze River Basin, *Acta Geographica Sinica*, 67(12), pp.1587-1598, in 2012