Postprint: Applicability Assessment of TIGGE Precipitation Forecasts in Arid and Semi-Arid Regions of China

Authors: He Chaolu, Lü Haishen

Date: 2022-03-28T12:22:22+00:00

Abstract

As an important component of the currently most authoritative dataset for medium-range ensemble forecasting, the applicability of TIGGE precipitation products in the arid and semi-arid regions of China requires further investigation. Based on observed precipitation data from 2015–2017 in the arid and semi-arid regions of China, and employing metrics such as mean absolute deviation, root mean square error, and TS score, this study comprehensively evaluated the forecast performance of four models from the TIGGE data center—ECMWF, JMA, KMA, and UKMO—in the study region from multiple perspectives including precipitation amount forecast, categorical precipitation forecast, precipitation detection capability, and spatial forecast accuracy. The results indicate: (1) All four models demonstrate relatively good performance for light rain forecasts; when forecasting different precipitation categories, JMA exhibits the best performance for light rain, while no significant differences are observed among the four models for other precipitation categories; (2) For daily precipitation forecasting, KMA performs the worst, whereas ECMWF is the most accurate; (3) Evaluation results of precipitation detection capability under different precipitation thresholds reveal that ECMWF holds a comparative advantage, particularly evident when using 25 mm · d-1 as the threshold; (4) Spatial forecast accuracy verification results show that all models perform optimally within the range of 80°–100°E, 35°–45°N, primarily covering central Xinjiang and the tri-provincial junction of Xinjiang, Gansu, and Qinghai; among the models, ECMWF demonstrates more stable performance, while KMA performs poorly.

Full Text

Assessment of TIGGE Precipitation Forecast Models in Arid and Semi-Arid Regions of China

HE Chaolu[12], LYU Haishen[12], ZHU Yonghua[12], LI Wentao[12], XIE Bingqi[3], XU Kaili[12], LIU Mingwen[12]

[1]State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210098, Jiangsu, China
[2]College of Hydrology and Water Resources, Hohai University, Nanjing 210098, Jiangsu, China
[3]Pearl River Water Resources Research Institute, Pearl River Water Resources Commission, Guangzhou 510630, Guangdong, China

Abstract

Short- and medium-term precipitation forecast products are crucial for improving the prediction period and accuracy of flood forecasts. With global climate change, precipitation pattern prediction has become increasingly complex and important. However, the applicability of TIGGE precipitation products—a key component of the most authoritative dataset for short- and medium-term ensemble forecasts—in China' s arid and semi-arid regions remains underexplored. Based on measured precipitation data from 2015–2017 in these regions, this study comprehensively evaluated the forecast performance of four TIGGE models (ECMWF, JMA, KMA, and UKMO) using multiple indicators including mean absolute deviation, root mean square error, and TS score. The evaluation examined four aspects: precipitation amount forecasting, precipitation level classification, precipitation detection capability, and spatial forecast accuracy. Results demonstrate that all four models effectively forecast light rain, with the JMA model showing the best performance for light rain at different precipitation levels, while no significant differences were observed among models for other precipitation levels. The KMA model performed worst for daily precipitation forecasting, whereas the ECMWF model achieved the highest accuracy. Evaluation of precipitation detection capabilities under different thresholds revealed that ECMWF holds a distinct advantage, particularly when the threshold is 25 mm · d^{-1}. Spatial accuracy assessment indicated that all models performed optimally within the 80°–100°E and 35°–45°N range, primarily covering central Xinjiang and the tri-provincial junction of Xinjiang, Gansu, and Qinghai. Among all models, ECMWF demonstrated the most robust performance, while KMA performed the poorest.

Keywords: TIGGE; precipitation forecast; precipitation level; spatial forecast accuracy; arid and semi-arid regions; China

1. Study Area and Data

1.1 Study Area

According to China' s annual precipitation distribution map, regions with annual precipitation below 400 mm are classified as arid and semi-arid areas—zones where precipitation is less than evaporation. The study area spans 27.27°-49.85°N and 73.43°-121.91°E, covering approximately 4.56×10^{6} km^2. Most of the region experiences a temperate continental climate characterized by scarce precipitation, low vegetation coverage, and extensive deserts, bare land, and saline-alkali soils. The boundary delineation follows the eco-geographical regionalization data from the Resource and Environment Science and Data Center of the Chinese Academy of Sciences (https://www.resdc.cn/data.aspx?DATAID=125).

1.2 Data Sources

1.2.1 Forecast Data This study utilized TIGGE forecast data from four operational centers: ECMWF (European Centre for Medium-Range Weather Forecasts), JMA (Japan Meteorological Agency), KMA (Korea Meteorological Administration), and UKMO (UK Met Office). Due to varying spatiotemporal resolutions across institutions, we standardized the forecast origin time to 00:00 UTC (corresponding to 08:00 Beijing time, UTC+8:00), set the spatial resolution to $0.5°\times0.5°$, and unified the forecast duration to 168 hours. Considering data completeness, we selected the period 2015-2017, which exhibited relatively complete forecast data. The forecast data can be freely obtained from https://apps.ecmwf.int/datasets/data/tigge/ in GRIB2 format. Using Python libraries xarray and cfgrib, we extracted the control forecast products from each model. Table 1 summarizes the basic information of the selected models.

1.2.2 Observed Data Observed precipitation data were obtained from the China Ground Precipitation Daily Gridded Dataset (version 2.0) established by the National Meteorological Information Center. This dataset is generated through spatial interpolation using local thin-plate smoothing splines, based on daily precipitation records from over 2,400 national meteorological stations and GTOPO30 elevation data (30 m $\times$ 30 m) resampled to $0.5°\times0.5°$ resolution. Quality assessment reveals that interpolation absolute errors are generally higher in southeastern China, with summer errors significantly larger than other seasons. The dataset tends to weaken precipitation intensity during heavy and moderate rain events while remaining closer to observed values during light rain events. It accurately captures precipitation spatial characteristics near major topographical features such as the Tibetan Plateau, Tianshan Mountains, and Tarim Basin. Consequently, this dataset serves as a reliable representation of observed precipitation in China' s arid and semi-arid regions. Table 1 presents the basic information of the observed data used in this study.

1.2.3 Data Processing To ensure comparability between forecast and observed data, we standardized their spatiotemporal scales through the following procedures: (1) Spatial scale unification: We applied bilinear interpolation to align forecast and observed data grids with 0.25° center differences, then extracted the study area using a mask. (2) Temporal scale unification: While forecast data represent 24-hour cumulative precipitation at 24-hour intervals, observed data are daily precipitation values. Given occasional missing forecast data, we first extracted daily precipitation data at 24-hour forecast steps, then removed corresponding observed data days for each model' s missing periods, ensuring temporal consistency between both datasets.

For precipitation amount evaluation, we employed the grid-averaging method to convert multi-model precipitation and observed data into area-averaged precipitation for the entire study region, enabling assessment of overall daily precipitation forecast effectiveness. Subsequently, we classified observed precipitation data according to the classification standards in Table 2, extracted forecast data at corresponding grid positions (same date and coordinates) using index matching, and generated separate datasets for different precipitation levels to evaluate classification forecast capability. Finally, we calculated regional forecast errors to reveal geographic variations among models.

Regarding precipitation detection capability, we selected threshold values of 0.1 $\mathrm{mm \cdot d^{-1}}$, 10 $\mathrm{mm \cdot d^{-1}}$, 25 $\mathrm{mm \cdot d^{-1}}$, and 50 $\mathrm{mm \cdot d^{-1}}$ based on the contingency table to compute statistical metrics, revealing detection capabilities across different models and lead times. This study exclusively examined precipitation events, excluding non-precipitation days from consideration.

2. Methods

2.1 Statistical Metrics

We employed commonly used indicators to evaluate precipitation forecast performance:

- **Mean Absolute Deviation (MAD)**: Measures forecast deviation direction and magnitude
- **Root Mean Square Error (RMSE)**: Reflects deviation between forecast and observed sequences without considering direction, with greater sensitivity to larger errors
- **Underestimation Rate (Sl) and Underestimation Error (Xl)**: Reflect average degree of forecast values being lower than observed values
- **Overestimation Rate (Sg) and Overestimation Error (Xg)**: Reflect average degree of forecast values being higher than observed values

All metrics approach zero for perfect forecasts, with values closer to zero indicating better model performance.

The formulas are:

$$\text{Bias} = \frac{1}{N}\sum_{t=1}^{N}(F_t - O_t)$$

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(F_t - O_t)^2}$$

$$S_g = \frac{N_g}{N}, \quad X_g = \frac{\sum_t X_{gt}}{N_g}$$

$$S_l = \frac{N_l}{N}, \quad X_l = \frac{\sum_t X_{lt}}{N_l}$$

where F_t and O_t represent forecast and observed precipitation on day t, respectively; N is the total number of forecast days; X_{gt} is the overestimation error on day t; X_{lt} is the underestimation error on day t; N_g is the total number of days with overestimation; and N_l is the total number of days with underestimation.

2.2 Classification Metrics

Beyond continuous variable evaluation, assessing precipitation event detection capability is essential. Using the contingency table approach proposed by Wilks, we calculated classification metrics for different precipitation thresholds to comprehensively evaluate detection capabilities of the four TIGGE models.

The contingency table structure is:

Forecast	Observed	Yes	No
Yes		H	F
No		M	

where: - **H** (Hits): Events detected by both observed and forecast data - **F** (False alarms): Events detected by forecast but not observed - **M** (Misses): Events detected by observed but not forecast data

We computed three key indicators: - **TS Score (Threat Score)**: Considers both successful detection and false alarms, reflecting actual detection capability - **POD (Probability of Detection)**: Proportion of observed precipitation events successfully forecasted - **FAR (False Alarm Ratio)**: Proportion of forecast precipitation events that did not occur

All three metrics range from 0 to 1, where higher TS and POD values and lower FAR values indicate better forecast performance. The formulas are:

$$\text{TS} = \frac{H}{H+F+M}, \quad \text{POD} = \frac{H}{H+M}, \quad \text{FAR} = \frac{F}{H+F}$$

3. Results and Analysis

3.1 Daily Precipitation Forecast Evaluation

To assess overall applicability of TIGGE models across the study area, we processed multi-model precipitation and observed data into area-averaged precipitation using the grid-averaging method. Five evaluation metrics (MAD, RMSE, Sl, Xl, Sg, and Xg) were applied to evaluate 1–7 day lead time forecasts.

Results show that all four models exhibit negative bias, indicating systematic underestimation of precipitation amounts. The underestimation rate (Sl) ranges from 0.65 to 0.85, significantly higher than the overestimation rate (Sg) which ranges from 0.15 to 0.35. Although Sl decreases with longer lead times while Sg increases, the predominant underestimation remains the primary source of precipitation forecast bias.

Forecast errors increase mainly during the first 1-3 days, subsequently stabilizing without significant changes. The ECMWF model demonstrates the best performance, particularly for 1–3 day lead times, while KMA performs worst. Overall, ECMWF shows superior daily precipitation forecasting capability with better forecast timeliness.

3.2 Precipitation Classification Forecast Evaluation

Beyond overall precipitation assessment, evaluating forecasts by precipitation level is crucial. We classified observed precipitation data according to Table 2, extracted corresponding forecast data at matching grid positions, and computed five evaluation metrics (MAD, RMSE, Sl, Xl, Sg, Xg) for different intensity levels across 1–7 day lead times.

For light rain, MAD ranges from 0.4 to 4.3 mm and RMSE from 0.4 to 0.9 mm, indicating high forecast capability that gradually declines with increasing precipitation levels. For heavy rain, MAD increases to 3.3–4.3 mm and RMSE to 0.7–0.6 mm. A notable case occurred on July 19, 2016, when a heavy rain event in the eastern study area (109°–111°E, 38°–39.5°N) with maximum single-grid daily precipitation of 44–60 mm was detected only by ECMWF, while other models missed it entirely and showed significant overestimation.

Across all precipitation levels, overestimation rates exceed underestimation rates, with overestimation errors (except for light rain) greater than underestimation errors. This results in slight negative bias for light rain but substantial positive bias for other levels, indicating that forecast overestimation constitutes the main error source beyond light rain.

Inter-model comparison reveals minimal differences during light rain events, where JMA performs best. For moderate, heavy, and torrential rain, ECMWF demonstrates clear superiority, particularly during the first 1–3 days when error reduction is most significant. The ranking of model performance from best to worst is: ECMWF, UKMO, JMA, and KMA.

3.3 Precipitation Detection Capability Evaluation

Classification metrics were calculated for thresholds of 0.1 $\mathrm{mm \cdot d^{-1}}$ (light rain), 10 $\mathrm{mm \cdot d^{-1}}$ (moderate rain), 25 $\mathrm{mm \cdot d^{-1}}$ (heavy rain), and 50 $\mathrm{mm \cdot d^{-1}}$ (torrential rain) across 1-7 day lead times using TS, POD, and FAR indicators.

For the 0.1 $\mathrm{mm \cdot d^{-1}}$ threshold, TS scores range from 0.32 to 0.48, POD from 0.5 to 0.8, and FAR from 0.35 to 0.55, indicating relatively good detection capability for light rain that deteriorates with increasing thresholds. POD decreases while FAR increases consistently across all models, demonstrating reduced detection capability at higher precipitation thresholds. The most significant degradation occurs during the first 1-3 days.

Model comparison shows greater differentiation for heavy rain, minimal differences for moderate rain, and poor performance for torrential rain (making comparisons meaningless). At the 0.1 $\mathrm{mm \cdot d^{-1}}$ threshold, although KMA shows higher POD, it also exhibits the highest FAR, suggesting a tendency toward over-forecasting light rain. At the 25 $\mathrm{mm \cdot d^{-1}}$ threshold, ECMWF demonstrates clear advantages with more robust performance.

3.4 Spatial Forecast Accuracy Evaluation

To analyze spatial patterns, we calculated RMSE at each grid point by comparing TIGGE model precipitation data with corresponding observed data. Grid results were interpolated to contour lines at 2 $\mathrm{mm \cdot d^{-1}}$ resolution, with areas exceeding 6 $\mathrm{mm \cdot d^{-1}}$ uniformly colored to visualize spatial accuracy patterns against a provincial administrative base map.

Spatial distribution reveals that most regions show RMSE below 3 $\mathrm{mm \cdot d^{-1}}$, with over half the area below 1 $\mathrm{mm \cdot d^{-1}}$, indicating generally good forecast accuracy across the study area. However, distinct spatial patterns emerge: RMSE increases from west to east, peaking in the mid-eastern region before decreasing. High-error areas coincide with high-precipitation zones, suggesting that abundant precipitation reduces forecast accuracy—more frequent precipitation corresponds to lower forecast precision.

All models perform best in the 80°-100°E and 35°-45°N range, primarily covering central Xinjiang and the Xinjiang-Gansu-Qinghai tri-provincial junction. Performance is poorest in the 100°-120°E and 35°-40°N range. ECMWF consistently outperforms other models across different lead times, demonstrating more stable performance, particularly at the 25 $\mathrm{mm \cdot d^{-1}}$ threshold.

4. Conclusions

This study evaluated TIGGE precipitation forecasts from ECMWF, JMA, KMA, and UKMO models in China' s arid and semi-arid regions using 2015-2017 daily precipitation data. A comprehensive applicability assessment was conducted from four perspectives: precipitation amount forecasting, precipitation level classification, precipitation detection capability, and spatial forecast accuracy. Key findings include:

1. **Precipitation Amount Forecasting**: All four models demonstrate strong forecasting capability for area-averaged daily precipitation across 1–7 day lead times, with good timeliness. ECMWF performs optimally, while KMA performs poorest. The models show slight underestimation for light rain but overestimation for other precipitation levels, with daily precipitation forecasts generally biased low.

2. **Precipitation Classification Forecasting**: The models exhibit superior performance for light rain, with MAD of 4.30 mm and RMSE of 0.44 mm, but poorer performance for torrential rain and above, with MAD reaching 44.00 mm. ECMWF shows the best performance, particularly for precipitation below 10 mm, while KMA performs worst. Inter-model differences are minimal for light rain but become pronounced for moderate to heavy precipitation.

3. **Precipitation Detection Capability**: All models demonstrate precipitation detection ability across different thresholds. ECMWF holds a distinct advantage, particularly at the 25 $mm \cdot d^{-1}$ threshold. Spatial accuracy tests reveal optimal performance in the 80°–100°E and 35°–45°N range (central Xinjiang and tri-provincial junction), with poorest performance in the 100°–120°E and 35°–40°N range.

4. **Common Characteristics**: All models show consistent patterns: forecast accuracy degrades with increasing lead time, with the most significant degradation during days 1–2; accuracy metrics deteriorate with increasing precipitation levels; and spatial patterns show decreasing accuracy from west to east, correlating with precipitation frequency.

These findings provide a foundation for precipitation correction, model integration, and improved medium-range forecasting in the study area. Specific recommendations include: applying different correction coefficients below and above certain thresholds to address systematic biases; assigning higher weights to ECMWF in multi-model ensembles; and conducting seasonal evaluations for applications beyond summer.

References

[1] Wang Haixia, Zhi Xiefei. Statistical downscaling research of precipitation forecast based on TIGGE multimodel ensemble[J]. Journal of the Meteorological Sciences, 2015, 35(4): 430-437.

[2] Du Kailian, Ge Yi, Zhang Jiali, et al. Study on the evaluation of TIGGE precipitation forecast precision in Chishan Lake Catchment[J]. Jiangsu Water Resources, 2020(7): 48-51.

[3] Amit B, Vinay K, Anjali S, et al. Application of multimodel superensemble

technique on the TIGGE suite of operational models[J]. Geomatics, 2021, 1(1): 81-91.

[4] Amini S, Azizian A, Arasteh P D. How reliable are TIGGE daily deterministic precipitation forecasts over different climate and topographic conditions of Iran?[J]. Meteorological Applications, 2021, 28(4): e2013, doi: 10.1002/met.2013.

[5] Zhang Chengjun, Ji Xiaoling, Ma Jingren, et al. Verification of numerical forecast and its application products in weather forecast in Ningxia[J]. Journal of Arid Meteorology, 2017, 35(1): 148-156.

[6] Jin Junliang, Shu Zhangkang, Chen Min, et al. Meteo-hydrological coupled runoff forecasting based on numerical weather prediction products[J]. Advances in Water Science, 2019, 30(3): 316-325.

[7] Bao Hongjun, Wang Lili, Shen Xueshun, et al. A review: Advances of flood forecasting of hydro-meteorological forecast technology[J]. Meteorological Monthly, 2016, 42(9): 1045-1057.

[8] Yong Yanlan, Yang Yunchuan, Zhao Tongtiegang, et al. Verification metrics of hydrological ensemble forecasts and their applications to seasonal precipitation forecasts of the Xijiang River Basin[J]. Journal of China Hydrology, 2020, 40(6): 31-39.

[9] Zhi Xiefei, Huang Wen. Multimodel ensemble forecasts of surface air temperature and precipitation over China by using Kalman filter[J]. Transactions of Atmospheric Sciences, 2019, 42(2): 197-206.

[10] Wang Jianqun, Duan Rong, Cai Chenkai. Application of TIGGE precipitation forecast in Shihe catchment of Huaihe River Basin[J]. Journal of Hohai University (Natural Sciences Edition), 2020, 48(1): 14-21.

[11] Patrice D, Jean Luc M, Richard A. Hydrological ensemble forecasting using a multi-model framework[J]. Journal of Hydrology, 2021, 600: 126537, doi.org/10.1016/j.jhydrol.2021.126537.

[12] Liu Zhangjun, Guo Shenglian, Xu Xinfa, et al. Bayesian probabilistic hydrological forecasting: Progress and prospects[J]. Journal of Hydraulic Engineering, 2019, 50(12): 1467-1478.

[13] Lei Xiaohui, Wang Hao, Liao Weihong, et al. Advances in hydrometeorological forecast under changing environment[J]. Journal of Hydraulic Engineering, 2018, 49(1): 9-18.

[14] Shu Zhangkang, Wang Lin, Jin Junliang, et al. Verification and integration of TIGGE multimode precipitation forecast products[J]. Hydro-Science and Engineering, 2021(2): 10-19.

[15] Zhao Linna, Wu Hao, Tian Fuyou, et al. Assessment of probabilistic precipitation forecasts for the Huaihe Basin using TIGGE data[J]. Meteorological Monthly, 2010, 36(7): 133-142.

[16] Zhao Yufei, Zhu Jiang. Assessing quality of grid daily precipitation datasets in China in recent 50 years[J]. Plateau Meteorology, 2015, 34(1): 50-58.

[17] Sun Mingyue, Lyu Haishen, Zhu Yonghua, et al. Applicability assessment of two meteorological datasets in areas lacking data with the Hutubi River Basin as an example[J]. Arid Zone Research, 2022, 39(1): 94-103.

[18] Gao Yanhong, Xu Jianwei, Zhang Meng, et al. Advances in the study of the 400 mm isohyet migrations and wetness and dryness changes on the Chinese mainland[J]. Advances in Earth Science, 2020, 35(11): 1101-1112.

[19] Wang Xiaofeng, Zhang Mingming, Yin Lichang, et al. Study on the driving factors in desertification process in arid and semi-arid region of China from 2000 to 2015[J]. Ecology and Environmental Sciences, 2019, 28(5): 948-957.

[20] Ma Jiehua, Sun Jianqi, Wang Jun, et al. Real-time prediction for 2018 JJA extreme precipitation and landslides[J]. Transactions of Atmospheric Sciences, 2019, 42(1): 93-99.

[21] Chen Tingxing, Lyu Haishen, Zhu Yonghua. Analysis of flood characteristics in Xiying River Basin based on GEV distribution[J]. Arid Zone Research, 2021, 38(6): 1563-1569.

[22] Wang Zaiwen, Liang Xudong, Fan Shuiyong, et al. Impact of spatial resolution on precipitation forecast score in numerical weather prediction models[J]. Torrential Rain and Disasters, 2016, 35(1): 10-16.

[23] Wilks D S. Statistical Methods in the Atmospheric Sciences[M]. USA: Academic Press, 2011: 301-394.

[24] Du Yaling, Lu Guihua, Wu Zhiyong, et al. Assessment of TIGGE medium-term precipitation forecast in Jiangsu Province[J]. South-North Water Transfers and Water Science & Technology, 2017, 15(6): 17-25, 44.

[25] Yin Zhiyuan, Wang Zhibin, Li Jun, et al. An experimental study on the prediction of flood using coupled WRF-Topmodel model[J]. Acta Meteorologica Sinica, 2017, 75(4): 672-684.

Note: Figure translations are in progress. See original paper for figures.