Evaluation of the Forecast Performance of the CMA-MESO Model for Warm-Season Precipitation in the Tarim Basin

2026-07-23

ChinaRxiv

Abstract

The topography of the Tarim Basin is complex, and the uncertainty in warm-season precipitation forecasts is significant. To gain an in-depth understanding of the precipitation forecast performance of the China Meteorological Administration regional high-resolution numerical model (CMA-MESO) in the Tarim Basin, this study systematically evaluates the forecast biases of the model for warm-season precipitation in the Tarim Basin using multiple precipitation verification metrics, based on CMA-MESO forecast data and ground observations during the warm seasons of 2023—2024. The results show that CMA-MESO can generally simulate the spatial distribution characteristics of warm-season precipitation in the basin well. As elevation increases, the TS score rises, while the PC score decreases. The magnitude of precipitation frequency forecast bias is most pronounced in the sub-high-elevation zone, with the positive bias center mainly concentrated on the windward slope of the western Kunlun Mountains; the maximum forecast biases of 24 h and 1 h precipitation frequency are 44.1% and 11.6%, respectively. The diurnal variation of the TS score is significant, characterized by higher values during the daytime and lower values at night, and this day-night bias is most evident in the low-elevation zone. In terms of changes in precipitation intensity, the TS scores of the CMA-MESO model for 24 h and 1 h effective precipitation are 34.4% and 8.8%, respectively, but forecast performance declines rapidly as precipitation intensity increases. Notably, the probability density bias of 1 h precipitation forecasts is mainly concentrated in the range of 0.1~2 mm; targeted correction of forecasts within this range can more effectively reduce the overall precipitation frequency bias. The findings can provide a scientific basis for improving CMA-MESO precipitation forecasts over complex terrain in arid regions and for operational disaster prevention and mitigation.

Full Text

DOI: 10.13866/j.azr.2026.07.02
CSTR: 32277.14.AZR.20260702

Evaluation of the Forecast Performance of the CMA-MESO Model for Warm-Season Precipitation in the Tarim Basin

Nian Huan1, Zhao Li2, Yang Xia1, Chen Chunyan1, Li Ruoqi1, Sun Mingjing1

(1. Xinjiang Uygur Autonomous Region Meteorological Observatory, Urumqi, Xinjiang 830002; 2. Xinjiang Information Engineering School, Urumqi, Xinjiang 830013)

Abstract: The terrain of the Tarim Basin is complex, and the uncertainty in warm-season precipitation forecasting is considerable. To gain an in-depth understanding of the precipitation forecast performance of the China Meteorological Administration regional high-resolution numerical model (CMA-MESO) in the Tarim Basin, this study systematically evaluated the forecast biases of the model for warm-season precipitation in the basin by using CMA-MESO forecast data and surface observation data for the warm seasons of 2023–2024, together with multiple precipitation verification metrics. The results show that CMA-MESO can, overall, reasonably simulate the spatial distribution characteristics of warm-season precipitation in the basin. As elevation increases, the TS score rises, whereas the PC score decreases. The magnitude of the forecast bias in precipitation frequency is most pronounced in the sub-high-altitude zone, and the centers of positive bias are concentrated on the windward slope of the western Kunlun Mountains; the maximum forecast biases in 24 h and 1 h precipitation frequency are 44.1% and 11.6%, respectively. The diurnal variation of the TS score is significant and is characterized by higher values during the day and lower values at night; this day–night bias is most evident in low-elevation areas. In terms of changes in precipitation intensity, the TS scores of the CMA-MESO model for 24 h and 1 h effective precipitation are 34.4% and 8.8%, respectively, but forecast performance declines rapidly as precipitation intensity increases. Notably, the probability-density bias of the 1 h precipitation forecast is mainly concentrated in the 0.1–2 mm range; targeted correction of forecasts within this interval can more effectively reduce the overall bias in precipitation frequency. The findings provide a scientific basis for improving CMA-MESO precipitation forecasts over complex terrain in arid regions and for operational disaster prevention and mitigation.

Keywords: Tarim Basin; warm season; CMA-MESO; precipitation; forecast evaluation

As the core region of the arid area of Northwest China[1], the Tarim Basin is strongly affected by variations in warm-season (May–September) precipitation, which have important impacts on regional water-resource balance, the ecological environment, and agricultural production[2]. In recent years, against the background of global warming, summer in the arid region of Northwest China has shown a significant climatic trend toward "wetting" [3]. Summer precipitation in the western Tarim Basin has increased significantly since 1987, which may be related to midlatitude circulation anomalies associated with the Silk Road atmospheric teleconnection pattern[4]. However, the region has complex terrain and strong underlying-surface heterogeneity, and meso- and small-scale convective systems are active during the warm season, leading to frequent extreme precipitation events and posing great challenges to operational precipitation forecasting[5].

Numerical weather prediction models are the core tools of modern precipitation forecasting[6]. Owing to factors such as differences in model parameterization schemes and initial conditions, the characteristics of model forecast biases vary across different spatial and temporal scales, and verification and evaluation are important tasks for advancing the development and application of numerical model systems[7]. Under complex terrain conditions, the accuracy of precipitation forecasts is generally significantly lower than in plain areas[8–9], whereas high-resolution numerical models can better depict regional climatic characteristics[10]. Such models have marked advantages in forecasting short-term local convective weather with strong convection and play an extremely important role in the prediction of some high-impact weather processes.

The mesoscale numerical model developed by the China Meteorological Administration (China Meteorological Administration Mesoscale Numerical Weather Prediction Model, CMA-MESO) has significantly improved the temporally and spatially refined forecasting capability for short-duration heavy precipitation through technologies such as a high-resolution dynamical framework, multi-source data assimilation, and rapid cycling[11]. Studies have shown that the CMA-MESO model has superior precipitation forecast performance compared with mainstream numerical forecast models in China and abroad[12]. In flood-season precipitation forecasting, the TS scores of CMA-MESO for 3 h accumulated precipitation and heavy-precipitation thresholds are better than those of the global weather forecast model of the European Centre for Medium-Range Weather Forecasts (ECMWF)[13]. Comparisons of model performance in heavy-precipitation forecasting at different resolutions indicate that high-resolution models can more accurately simulate water-vapor transport and vertical motion over multiscale terrain, making the forecast precipitation area and intensity closer to observations[14]. Cai Yi et al.[15], He Mu

Received: 2025-08-08; **Revised:** 2025-12-04
Funding: Key Research and Development Program of Xinjiang Uygur Autonomous Region (2023B03019-2); Source Risk Identification and Dynamic Monitoring and Forecasting of Rain–Snow–Ice Events Causing Floods in Alpine Mountain Areas (2022B03021-1); Natural Science Foundation of Xinjiang Uygur Autonomous Region (2022D01A292)
Author profile: Nian Huan (1982–), male, senior engineer, mainly engaged in research on objective correction techniques for short- and medium-range numerical models. E-mail: mouh198209@foxmail.com
Corresponding author: Chen Chunyan. E-mail: chency357@sina.com

et al. [16] studied precipitation forecasts by the CMA-MESO model during the rainy season and found that it could accurately capture daily precipitation amounts and precipitation frequency, but that its bias was significant when forecasting the diurnal variation of precipitation over complex terrain. Some studies have also argued, however, that kilometer-scale numerical models still cannot accurately simulate turbulence and shallow-convection processes at the boundary of the troposphere over land [17–18]. Therefore, the forecasting performance of high-resolution numerical models still requires further evaluation. Existing studies, however, have mostly focused on monsoon regions such as eastern and southwestern China. These regions are dominated by low-elevation terrain; in summer, under the influence of water vapor transported by the monsoon, low-level specific humidity can exceed 15 $g \cdot kg^{-1}$, and the boundary-layer top is high and the mixed layer thick, conditions that are highly favorable for the occurrence of widespread stratiform clouds and persistent convective precipitation [19–20]. By contrast, the high mountains above 4000 m around the Tarim Basin strongly block and divert airflow, reducing the entry of water vapor into the basin. Low-level specific humidity in the Kashgar area can reach 4.3 $g \cdot kg^{-1}$, the highest specific humidity in the basin [21]. The basin' s water vapor is supplied mainly by relay transport from the midlatitude westerlies; the mean column-integrated atmospheric water vapor content is only 8 $kg \cdot m^{-2}$, and the net summer water-vapor budget of the basin as a whole is negative [22]. Because the water-vapor content is extremely low, the triggering of severe convection depends largely on orographic lifting, resulting in a short convective life cycle and small spatial scale; precipitation thus exhibits the extreme characteristics of "high intensity, short duration, and strong locality." The precipitation-forming conditions in the Tarim Basin differ markedly from those in monsoon regions in terms of terrain, water vapor, and convective mechanisms, meaning that evaluation conclusions for high-resolution models in monsoon regions should not be directly applied to arid regions. At present, studies evaluating the precipitation-forecasting performance of high-resolution numerical models in arid regions are scarce. It is therefore necessary to evaluate the precipitation-forecasting performance of high-resolution numerical models for the Tarim Basin.

Using CMA-MESO model forecasts and automatic-station precipitation observations for the warm seasons of 2023–2024, this study comprehensively applies multiple evaluation methods to analyze the CMA-MESO model' s ability to capture precipitation of different intensities in the Tarim Basin and the spatiotemporal characteristics of forecast biases. It also explores the causes of model biases, with the aim of providing a theoretical basis for improving high-resolution numerical-model precipitation forecasts and their operational application in arid regions.

1 Data and Methods

1.1 Overview of the Study Area

The Tarim Basin is the largest inland arid basin in China (Fig. 1). It is surrounded on its northern, western, and southern sides by high mountain ranges, with mean elevations in the mountainous areas exceeding 4000 m. External water-vapor transport is blocked by the high mountains, and the annual mean precipitation is extremely low, forming a typical arid-climate region. The spatiotemporal distribution of precipitation in this region exhibits three typical features: (1) a pronounced diurnal variation, with precipitation peaks concentrated from evening to the early morning of the following day; (2) few rainstorm days, but strong extremity and high disaster risk—when short-duration heavy precipitation occurs, the daily precipitation amount can reach 60% of the annual precipitation [23]; and (3) significant orographic forcing. These features make precipitation forecasting in the basin more challenging than in monsoon regions, making it an important region for testing the performance of numerical models. Based on China' s terrestrial geomorphological indicators [24], this paper divides the basin into low-elevation areas (0–1000 m, accounting for 19% of the observation stations), middle-elevation areas (1000–2000 m, accounting for 72% of the observation stations), and sub-high-elevation areas (2000–4000 m, accounting for 9% of the observation stations), in order to describe the vertical distribution characteristics of model forecast biases.

1.2 Brief Introduction to the CMA-MESO Model

The version of the CMA-MESO model used in this paper is v5.1. Its main technical framework and physical-parameterization configuration [25] are as follows: the model adopts a semi-

Note: The base map was produced using the standard map of the Department of Natural Resources of the Xinjiang Uygur Autonomous Region; the map approval number is Xin S (2021) 047. The boundaries of the base map were not modified. The same applies below.

Fig. 1 Spatial distribution of the meteorological stations in the Tarim Basin

an implicit semi-Lagrangian scheme and a fully compressible nonhydrostatic dynamic model framework; no cumulus-convection parameterization scheme was activated. The forecast temporal resolution was 1 h. In the vertical direction, a height-based terrain-following coordinate and Charney–Philips vertical staggering were adopted, with a total of 50 layers. In the horizontal direction, spherical coordinates and an Arakawa-C staggered latitude-longitude grid were used, with a resolution of $0.03^\circ \times 0.03^\circ$. The physical parameterization schemes included the Noah land-surface process scheme, the Monin–Obukhov near-surface-layer scheme, the MRF boundary-layer scheme, the RRTM longwave radiation scheme, the Dudhia shortwave radiation scheme, and the WSM6 cloud microphysics scheme.

1.3 Data Sources

The station precipitation data for May–September 2023–2024 and the CMA-MESO model precipitation forecasts (hereinafter referred to as NWP) provided by the Information Center of the Meteorological Bureau of Xinjiang Uygur Autonomous Region were selected. They include: (1) hourly precipitation observations from 674 automatic weather observation stations in the Tarim Basin (hereinafter referred to as SPO), after quality control; and (2) precipitation forecasts initialized at 20:00 (Beijing time, the same below). Owing to operational requirements and delays in model-forecast transmission, forecast lead times of 13–36 h were selected.

1.4 Evaluation Metrics

According to the operational verification standards of the National Meteorological Center, the nearest-neighbor interpolation method was used to interpolate the gridded values of the numerical model to nearby observation stations, and multiple metrics were used for evaluation. The evaluation metrics included the threat score (TS), rain/no-rain forecast accuracy (PC), probability of detection (POD), false alarm ratio (FAR), missed alarm ratio (MAR), frequency bias (Bias), and frequency-forecast bias rate (r). The calculation formulas are as follows:

$$\mathrm{TS} = \frac{H}{H+M+F} \times 100\% \tag{1}$$

$$\mathrm{PC} = \frac{H}{H+M+F+CN} \times 100\% \tag{2}$$

$$\mathrm{Bias} = \frac{H+F}{H+M} \times 100\% \tag{3}$$

$$\mathrm{POD} = \frac{H}{H+M} \times 100\% \tag{4}$$

$$\mathrm{FAR} = \frac{M}{H+M} \times 100\% \tag{5}$$

$$\mathrm{MAR} = \frac{F}{H+F} \times 100\% \tag{6}$$

$$r_i = \frac{g_i - l_i}{l_i} \times 100\% \tag{7}$$

where H is the number of correctly forecast precipitation events; F is the number of false-alarm precipitation events; M is the number of missed precipitation events; and CN is the number of correctly forecast non-precipitation events. When Bias $> 100\%$, the number of false alarms is greater than the number of missed forecasts; when Bias $< 100\%$, the number of false alarms is less than the number of missed forecasts. r_i is the precipitation-frequency forecast bias rate; when $r_i > 0$, it is a positive bias, indicating that the forecast precipitation frequency is higher than the observed frequency; when $r_i < 0$, it is a negative bias, indicating that the forecast precipitation frequency is lower than the observed frequency. i denotes the precipitation-intensity class; g is the forecast precipitation frequency, and l is the observed precipitation frequency.

According to the China Meteorological Administration precipitation grid-evaluation rules[26] and the Xinjiang precipitation operational specifications[27], this study classified 24 h precipitation R_{24} into light rain (0.1 mm $\leq R_{24} < 6.1$ mm), moderate rain (6.1 mm $\leq R_{24} < 12.1$ mm), heavy rain (12.1 mm $\leq R_{24} < 24.1$ mm), and rainstorm ($R_{24} \geq 24.1$ mm). The 1 h precipitation amount R_1 was divided into ordinary precipitation (0.1 mm $\leq R_1 < 10$ mm) and intense precipitation ($R_1 \geq 10.0$ mm). In addition, based on precipitation frequency, 1 h ordinary precipitation[28] was divided into weak precipitation (0.1 mm $\leq R_1 < 2.0$ mm), moderate precipitation (2.0 mm $\leq R_1 < 5.0$ mm), and moderately strong precipitation (5.0 mm $\leq R_1 < 10.0$ mm). When $R_{24} \geq 0.1$ mm ($R_1 \geq 0.1$ mm), effective precipitation occurred within 24 h (1 h), and it was counted as one precipitation event.

2 Results and Analysis

2.1 Spatial Distribution of Warm-Season Precipitation and Biases

2.1.1 Spatial Distribution of Precipitation

As can be seen from Fig. 2, both the accumulated precipitation amount and precipitation hours in the warm season increased with rising elevation. Specifically, at observation stations in low-elevation areas, the mean warm-season precipitation amount was approximately 46.4 mm, and the mean precipitation hours were approximately 55.5 h. At observation stations in middle-elevation areas, the mean warm-season precipitation amount was approximately 63.9 mm, and the mean precipitation hours were approximately 70.8 h. At observation stations in sub-high-elevation areas, the mean warm-season precipitation amount was approximately

Figure panels: (a) accumulated precipitation amount; (b) accumulated precipitation hours. Legend: low elevation, middle elevation, sub-high elevation. Color bars: precipitation amount/mm; precipitation hours/h; elevation/m.

Fig. 2 Spatial distribution of accumulated precipitation amount and precipitation hours at observational stations over the Tarim Basin during the warm season.

140.2 mm, and the mean precipitation duration is about 156.9 h. Among them, at some observing stations the precipitation amount exceeds 300 mm and the precipitation duration exceeds 260 h; these are mainly concentrated on the windward slope of the western Kunlun Mountains above 2400 m, where terrain uplift has a significant enhancement effect on precipitation.

2.1.2 Spatial distribution of forecast bias

As shown in Table 1, the proportions of stations with positive bias in the 24 h and 1 h precipitation forecasts both increase rapidly with elevation. In terms of bias magnitude, the mean bias, mean absolute bias, and maximum and minimum bias amplitudes in the sub-high-elevation region are all significantly higher than those in the low- and mid-elevation regions. Among them, the maximum biases of the 24 h and 1 h precipitation-frequency forecasts are 44.1% and 11.6%, respectively, and the minimum biases are −21.1% and −6.1%, respectively.

As shown in Fig. 3, driven by the lifting of westerly airflow, the positive-bias center is concentrated on the windward slope of the western Kunlun Mountains (37°-39°N, 74°-76°E). Influenced by the foehn effect of southwesterly airflow, the negative-bias center is dispersed over the leeward slope of the middle Kunlun Mountains (35.5°-37°N, 80°-85°E). The bias amplitude in the sub-high-elevation region is significantly larger than that in the lower-elevation regions, and the positive-bias center is markedly higher than the negative-bias center. This indicates a sensitive coupling relationship between terrain slope direction and the physical processes in the model.

2.2 Precipitation forecast performance at different elevations

2.2.1 Evaluation of precipitation forecast performance at different elevations

As shown in Table 2, forecast performance differs significantly among elevations. Judging from the TS and PC scores, the 24 h forecasts are 34.4% and 81.3%, respectively; the 1 h forecasts are 11% and 96.2%, respectively. The trends of the PC and TS scores with increasing elevation are opposite. This is because the PC score emphasizes the accuracy of rain/no-rain classification, whereas the TS score assesses the proportion of correctly forecast precipitation. This shows that the model has good indicative value for warm-season precipitation forecasting, especially with better spatiotemporal agreement and discrimination capability for 24 h accumulated precipitation; this is consistent with the conclusion that "the predictability of long-duration accumulated precipitation is higher than that of short-duration precipitation" for the CMA-MESO model over the complex terrain of the Ili Valley[1]. The Bias score can assess the balance between the number of missed precipitation forecasts and false alarms. Overall, the Bias scores of the 24 h and 1 h forecasts are 107.9% and 100.7%, respectively, which are very close to the ideal value (100%). Comparison among elevations shows that precipitation tends to be underforecast in the low-elevation region and overforecast in the sub-high-elevation region, while the precipitation forecasts in the mid-elevation region are the most balanced; the Bias scores of the 24 h and 1 h forecasts are 103.2% and 89%, respectively.

[1] Reference [28].

2.2.2 Diurnal variation in precipitation forecast scores at different elevations

As shown in Fig. 4, with increasing elevation, the PC score decreases slightly, but the amplitude of diurnal variation is not obvious. As shown in Table 2, the frequency of hourly precipitation is only 2.4%, and the high PC score is mainly contributed by the correct forecasting of "no precipitation" events; its indicative significance for actual precipitation forecasting capability is limited. From the diurnal variation of the TS score for the model's effective precipitation forecasts, the TS score is characterized by being high during the day and low at night. Overall, the daytime (09:00-20:00, same below) mean is 12.3%, and the nighttime (21:00-08:00 the next day, same below) mean is 10%. Comparing dif-

Table 1 Forecast bias of 24 h and 1 h precipitation frequency for the CMA-MESO model

Tab. 1 Forecast biases of 24 h and 1 h precipitation frequency for the CMA-MESO model

Forecast period	Elevation/m	Proportion of stations with positive bias/%	Mean bias/%	Mean absolute bias/%	Maximum bias/%	Minimum bias/%
24 h	$ $1000	16.7	−2.7	3.3	5.9	−14.5
24 h	1000–2000	48.2	0.5	3.5	32.6	−12.5
24 h	2000–4000	70.2	6.8	10.2	44.1	−21.1
1 h	$ $1000	11.7	−0.6	0.6	0.3	−3.0
1 h	1000–2000	34.0	−0.2	0.57	8.2	−2.4
1 h	2000–4000	63.4	1.1	2.1	11.6	−6.1

Figure panels

- **(a) 24 h precipitation**
  Legend: low elevation; mid-elevation; sub-high elevation. Elevation/m. Precipitation-frequency forecast bias/%. Scale: −50, −40, −30, −20, −10, 0, 10, 20, 30, 40, 50.
- **(b) 1 h precipitation**
  Legend: low elevation; mid-elevation; sub-high elevation. Elevation/m. Precipitation-frequency forecast bias/%. Scale: −12, −9, −6, −3, 0, 3, 6, 9, 12.

Fig. 3 Spatial distribution of precipitation-frequency forecast bias in CMA-MESO

Tab. 2 Verification scores of precipitation forecasts by the CMA-MESO model

Accumulated time/h	Elevation/m	Precipitation frequency/%	TS/%	PC/%	Bias/%
24	Overall	18.4	34.4	81.3	107.9
24	$ $1000	12.3	26.2	87.3	77.7
24	1000–2000	16.3	31.6	82.8	103.2
24	2000–4000	29.0	41.3	73.1	123.6
1	Overall	2.4	11.0	96.2	100.7
1	$ $1000	1.5	10.7	98.0	61.5
1	1000–2000	1.9	9.6	97.0	89.0
1	2000–4000	4.3	13.1	92.5	126.5

By elevation, it can also be seen that the highest mean daytime and nighttime TS scores both occurred in the sub-high-elevation zone, at 14.2% and 12.6%, respectively; the lowest daytime and nighttime mean values occurred in the mid-elevation and low-elevation zones, at 10.8% and 6.4%, respectively. In addition, the day–night difference in TS was smallest in the sub-high-elevation zone (1.6%) and most pronounced in the low-elevation zone (7.6%).

To explain the day–night difference in TS, the POD, FAR, and MAR scores were further diagnosed. Figure 5 shows that the daytime POD reached 22.3%, markedly higher than the nighttime value of 17.8%; meanwhile, FAR and MAR increased from daytime to nighttime by 2.1% and 4.5%, respectively. Thus, the higher daytime POD, together with lower FAR and MAR, contributed to the day–night difference in TS. In terms of elevation differences, missed forecasts dominated in the low- and mid-elevation zones, whereas false alarms were more prominent in the sub-high-elevation zone. This differs from results obtained for the complex terrain of Southwest China: the CMA-MESO model generally overestimates precipitation frequency in monsoon regions[29], whereas in the basin it mainly underestimates nighttime precipitation frequency. This bias structure reflects an asymmetric model response to precipitation produced by different forcing mechanisms under complex terrain. In the low-elevation zone, precipitation is mostly dominated by shallow convection or local thermally driven convection triggered by daytime surface heating. The model may miss weak precipitation because the boundary-layer scheme produces excessively strong cooling of the stable nocturnal boundary layer and insufficient mixing of low-level water vapor; this is especially evident at night when thermal forcing is weak. In the sub-high-elevation zone, precipitation is more strongly controlled by large-scale westerly water-vapor flow around the terrain and by orographic lifting. If the terrain height is smoothed, the upslope lifting rate on windward slopes is too high, and the conversion efficiency from cloud water to raindrops in the microphysics scheme is biased high, systematic false alarms may result[30].

2.3 Forecast performance for precipitation of different intensities

2.3.1 Frequencies of precipitation of different intensities The proportion of no precipitation in the basin is far greater than that of effective precipitation, and the proportions of the various precipitation intensities decrease rapidly as intensity increases. Figure 6a shows that the observed and forecast proportions of 24 h precipitation (including no precipitation; the same below) were 81.61%, 15.97%, 1.6%, 0.66%, 0.16% and 80.15%, 17.39%, 1.6%, 0.75%, 0.11%, respectively; Fig. 6b

Visible labels in Fig. 4: (a) Overall; (b) Low elevation; (c) Mid-elevation; (d) Sub-high elevation. Axes: PC/%; TS/%; Time. Legend: PC; TS; daytime mean TS; nighttime mean TS.

Fig. 4 Diurnal variations of precipitation forecast TS and PC scores for the CMA-MESO model

Fig. 5 Diurnal variations of warm-season dichotomous verification metrics (POD, FAR, MAR) and precipitation frequency for the CMA-MESO model

In-figure labels:
(a) Overall; (b) Low elevation; (c) Middle elevation; (d) Sub-high elevation.
Left y-axis: Precipitation frequency/%; right y-axis: precipitation score/%; x-axis: Time.
Legends: SPO, NWP, POD, FAR, MAR.

Fig. 6 Cumulative frequency and proportion of precipitation events with different intensities

In-figure labels:
(a) 24 h precipitation; (b) 1 h precipitation.
Left y-axis: cumulative count/events; right y-axis: cumulative-count proportion/%; x-axis: precipitation intensity.
Legends: SPO count, NWP count, SPO proportion, NWP proportion.
Precipitation-intensity categories:
(a) No precipitation, light rain, moderate rain, heavy rain, rainstorm.
(b) No precipitation, weak precipitation, moderate precipitation, moderately strong precipitation, strong precipitation.

indicate that the observed and forecast proportions of 1 h precipitation are 97.64%, 2.09%, 0.22%, 0.04%, 0.01% and 97.61%, 2.12%, 0.22%, 0.04%, 0.005%, respectively. Comparing the forecast bias rate r of precipitation frequency for each precipitation intensity shows that, for 24 h precipitation, r is -1.8%, 8.9%, -0.1%, 13.2%, and -31.1%, respectively, whereas for 1 h precipitation, r is -0.02%, 1.5%, -0.4%, -13.6%, and -63.2%, respectively. The above results show that, when the CMA-MESO model forecasts precipitation of relatively high intensity, r is significantly biased high; this is consistent with the widespread tendency of the ECMWF model to underestimate frequency in rainstorm forecasts for Xinjiang [31].

2.3.2 Forecast performance for precipitation of different intensities

It can be seen from Fig. 7 that, as precipitation intensity increases, the TS score decreases. Among them, for 24 h precipitation of different intensities ($R_{24} \geq 0.1$ mm, $R_{24} \geq 6.1$ mm, $R_{24} \geq 12.1$ mm and $R_{24} \geq$

Fig. 7. Forecast scores (TS, FAR, MAR) of the CMA-MESO model for different precipitation intensities.
Panels: (a) 24 h precipitation; (b) 1 h precipitation. The y-axis is precipitation score/% and the x-axis is precipitation amount/mm.

For precipitation of different 24 h intensities ($R_{24} \geq 0.1$ mm, $R_{24} \geq 6.1$ mm, $R_{24} \geq 12.1$ mm, and $R_{24} \geq 24.1$ mm), the TS scores were 34.4%, 13.3%, 6.8%, and 4%, respectively. For precipitation of different 1 h intensities ($R_1 \geq 0.1$ mm, $R_1 \geq 2$ mm, $R_1 \geq 5$ mm, and $R_1 \geq 10$ mm), the TS scores were 8.8%, 2.4%, 0.8%, and 0.1%, respectively. Among them, the TS scores for $R_{24} \geq 12.1$ mm and $R_1 \geq 5$ mm were relatively low. In addition, FAR and MAR increased with precipitation intensity. The difference between the 24 h and 1 h forecasts was that, when $R_{24} < 12.1$ mm, FAR was greater than MAR; when $R_{24} \geq 24.1$ mm, FAR was smaller than MAR. By contrast, FAR and MAR were similar for the different precipitation intensities in the 1 h forecasts. Compared with the 24 h precipitation forecasts of the ECMWF model[31], the CMA-MESO model had slightly better TS scores when $R_{24} < 12.1$ mm, whereas the ECMWF model had better TS scores when $R_{24} \geq 12.1$ mm. Therefore, relying on perturbations in the initial fields and physical processes introduced by the ensemble prediction system, ECMWF can better characterize the probability distribution of rainstorm events. CMA-MESO has better forecasting skill for light and moderate rain, but its forecasting performance is insufficient for heavy rain and above. This indicates that, in regions where moisture is deficient and the threshold of convective available potential energy is sensitive, if low-level moisture transport in the initial field is too weak or the phase of westerly disturbances is biased, convection may fail to be triggered; conversely, if orographic lifting or boundary-layer heating is overestimated, spurious heavy precipitation may be produced. This nonlinear response amplifies the uncertainty in heavy-precipitation forecasting[32].

2.3.3 Precipitation probability density distribution Figure 8 shows the probability density distribution of warm-season precipitation in the basin and the forecast deviation in precipitation density. Both the observations and forecasts of 24 h and 1 h precipitation exhibited a “sharp-peak, long-tail”distribution, with small-magnitude precipitation highly concentrated and extreme events very sparse. Specifically, 93.7% of the 24 h precipitation samples were distributed within 0.1-10 mm, and 97.6% of the 1 h precipitation samples were distributed within 0.1-5 mm. Moreover, the probability-density peaks of both 24 h and 1 h precipitation were concentrated at 0.1-1 mm. The density deviation reflects the strength of systematic model bias; the density deviation of the 24 h forecast was mainly dis-

Fig. 8. Probability density distribution of precipitation and forecast deviation in probability density.
Panels: (a) 24 h precipitation; (b) 1 h precipitation. The x-axis is precipitation amount/mm. The left y-axis is probability density, and the right y-axis is the density difference between NWP and SPO. Legends: SPO, NWP, and density difference.

distributed over 0.1–10 mm and exhibited a “bimodal” structure. Among them, the bias amplitude was largest at 0.1–1 mm, with a peak of 0.02; the bias amplitude at 4.1–5 mm was secondary, at approximately −0.006. By contrast, the density bias of the 1 h forecast was concentrated at 1.1–2 mm, showing a “unimodal”structure, with a peak of 0.01. The systematic overestimation of trace precipitation is usually related to an excessively low triggering threshold for shallow convection, overmixing in the boundary layer, and insufficient collision-coalescence efficiency in the cloud microphysics scheme; the underestimation in the moderate-to-heavy rainfall range is associated with insufficient vertical transport by deep convection and an imperfect representation of orographic-lifting triggering mechanisms.

3 Discussion

The CMA-MESO model can, to a certain extent, simulate the spatial distribution characteristics of warm-season precipitation. However, along the western margin of the basin, the positive-bias center of precipitation frequency is 2.1 times the absolute value of the negative-bias center, indicating that the model’s ability to simulate the sub-high-elevation topographic transition zone is relatively insufficient. Precipitation amount and precipitation hours increase by more than 200% from the middle-elevation zone to the sub-high-elevation zone, far exceeding the increase from the low-elevation zone to the middle-elevation zone (< 50%). This feature bears some similarity to the “precipitation overestimation” phenomenon of the WRF model over the steep Sichuan Plateau-Sichuan Basin transition zone[33], confirming that high-resolution models commonly over-simulate orographic precipitation in regions with large-scale terrain gradients. Nevertheless, the vertical structures of precipitation in the two regions differ. In the Sichuan Basin, the precipitation peak occurs at elevations of 200–1200 m, and precipitation rapidly decreases if elevation continues to rise[34]; in the Tarim Basin, however, precipitation continues to increase with elevation, reflecting the distinctive efficiency of orographic lifting under the water-vapor-limited background of arid regions.

The above bias characteristics are closely related to the model' s own physical framework. Therefore, forecast biases may arise from the microphysics scheme' s overestimated efficiency in converting cloud water into ice-phase hydrometeors. Water vapor in the basin is extremely scarce, and near the altitude where ice-phase processes are initiated, the microphysics scheme may overestimate the efficiency with which cloud water is converted into ice particles[35]. The model may also have biases in estimating saturation vapor pressure. Saturation vapor pressure decreases exponentially with altitude, causing water-vapor condensation during orographic lifting to exhibit nonlinear characteristics[36]. On the windward slopes of the western basin, ascending airflow cools due to adiabatic expansion, leading to a rapid decrease in saturation vapor pressure. If the model overestimates the cooling rate or underestimates the ambient temperature, the air reaches saturation too early, thereby amplifying the condensation amount and producing systematic false alarms of windward-slope precipitation. The model' s representation of terrain is distorted. Constrained by horizontal resolution and terrain-smoothing treatment, if the truly steep terrain is excessively simplified, the lifting structure on the windward slope cannot be realistically reproduced. On this basis, the terrain gradient is still represented in the model as discrete steps; together with the absence of realistic flow-around pathways, this instead induces overly strong concentrated lifting near the top of the windward slope, causing vertical velocity and condensation rate to be systematically overestimated[37]. The model may also have biases in simulating turbulence and the wind field. Low-level turbulence and wind-field disturbances induced by the basin' s complex terrain can alter the falling trajectories and residence times of ice-phase hydrometeors, while the model' s description of such subgrid-scale dynamical and microphysical interactions remains inadequate; this may instead exert a misleading influence on precipitation forecasts[38].

This study reveals that the bias centers in CMA-MESO precipitation-frequency forecasts exhibit significant coupling between elevation gradients and terrain orientation, providing a new target for the correction and optimization of precipitation forecasts over complex terrain in arid regions. The complex basin topography and sparse station network not only constrain the model initial fields but also increase the uncertainty of evaluation. In the future, it will be necessary to enhance multisource observational assimilation capability. By improving data assimilation techniques and cloud microphysics schemes, the model' s representation of water-vapor transport, condensation efficiency, and terrain response can be jointly optimized, thereby improving precipitation-forecast capability over complex terrain in arid regions.

4 Conclusions

Based on a multi-index verification framework, this study comprehensively evaluated the performance of the CMA-MESO model in forecasting 24 h and 1 h precipitation over the Tarim Basin during the warm seasons of 2023–2024. The main conclusions are as follows:

(1) Biases in precipitation-frequency forecasts show obvious elevation dependence. Forecast biases are mainly negative in the low- and middle-elevation zones, whereas they are mainly positive in the sub-high-elevation zone, where the biases are the largest. In terms of spatial distribution, the positive-bias centers are concentrated on the windward slopes of the sub-high-elevation zone along the western margin of the basin, forming a north–south belt along the mountains; the negative-bias centers are scattered on the leeward slopes along the southern margin of the basin, extending east–west. These characteristics indicate that the model has an evident systematic tendency to overforecast precipitation frequency in the sub-high-elevation topographic transition zone, and that this bias is closely related to terrain forcing.

(2) The TS scores of model precipitation forecasts exhibit a diurnal variation characterized by "higher in the daytime and lower at night." The day-night difference in TS scores is largest in the low-elevation zone, mainly because missed forecasts are more frequent; the difference is smallest in the sub-high-elevation zone because false alarms are prominent. These features reflect an asymmetry in the model' s response to precipitation-forcing mechanisms at different elevations. This elevation-dependent bias structure provides a scientific basis for subsequent bias correction by region and forecast lead time.

(3) TS scores decrease rapidly as precipitation intensity increases. The model performs poorly in forecasting 24 h rainstorms (TS = 4%) and 1 h heavy precipitation (TS = 0.1%), highlighting its limitations in forecasting stronger precipitation. This arises from initial-field biases caused by sparse observations in the basin and the insufficient adaptability of physical schemes to water-vapor-deficient environments. In high-resolution models, even slight biases in water vapor and boundary-layer disturbances can trigger false alarms or missed forecasts.

(4) Judging from the probability-density distribution of precipitation, the bias of the 24 h forecast exhibits a bimodal structure, whereas the bias of the 1 h forecast exhibits a unimodal structure and its bias distribution is more concentrated. This feature indicates that, compared with directly correcting 24 h precipitation forecasts, starting from 1 h precipitation forecasts to construct a correction scheme and then iterating it to the 24 h scale can achieve higher correction efficiency and stability in operations, providing a better technical pathway for improving forecast accuracy.

References

[1] Jin Chen, He Qing, Huang Qian. A review of rainstorm research in southern Xinjiang[J]. Transactions of Atmospheric Sciences, 2023, 46(1): 82–96. [Jin

Chen, He Qing, Huang Qian. A review on rainstorm research in southern Xinjiang[J]. Transactions Atmospheric Sciences, 2023, 46(1): 82–96.]

[2] Zhao Yuzhi, Yang Jianjun. Spatio-temporal pattern of water-resource carrying capacity and coupling coordination of subsystems in southern Xinjiang[J]. Arid Zone Research, 2023, 40(2): 213-223. [Zhao Yuzhi, Yang Jianjun. Spatio-temporal pattern of water resource carrying capacity, coupling and coordination of subsystems in southern Xinjiang[J]. Arid Zone Research, 2023, 40(2): 213–223.]

[3] IPCC. Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change[R]. Cambridge, UK and New York, NY, USA: Cambridge University Press, 2022.

[4] Qu L L, Zhao Y, Zhou Y M, et al. Why has the summer rainfall increased prominently in the west Tarim Basin of Northwest China since 2010?[J]. Atmospheric Research, 2023, 284: 106602.

[5] Wu Xiaodan, Luo Min, Meng Fanhao, et al. Analysis of the spatio-temporal evolution characteristics of extreme climate events in Xinjiang under the background of climate warming and humidification[J]. Arid Zone Research, 2022, 39(6): 1695–1705. [Wu Xiaodan, Luo Min, Meng Fanhao, et al. New characteristics of spatio-temporal evolution of extreme climate events in Xinjiang under the background of warm and humid climate[J]. Arid Zone Research, 2022, 39(6): 1695–1705.]

[6] Zhang Jin, Sun Jian, Shen Xueshun, et al. Development of key technologies and operational application of the CMA-GFS V4.0 model[J]. Journal of Applied Meteorological Science, 2023, 34(5): 513–526. [Zhang Jin, Sun Jian, Shen Xueshun, et al. Key model technologies of CMA-GFS V4.0 and application to operational forecast[J]. Journal of Applied Meteorological Science, 2023, 34(5): 513–526.]

[7] Cao Yong, Bao Hongjun, Zhang Hengde, et al. A seamless quantitative precipitation forecasting model based on rapid rolling update technology[J]. Journal of Hohai University (Natural Sciences), 2021, 49(4): 303–308. [Cao Yong, Bao Hongjun, Zhang Hengde, et al. Seamless quantitative precipitation forecasting model based on rapid rolling update technique[J]. Journal of Hohai University (Natural Sciences), 2021, 49(4): 303–308.]

[8] Mehran A, Aghakouchak A, Phillips T J. Evaluation of CMIP5 continental precipitation simulations relative to satellite based gauge-adjusted observations[J]. Journal of Geophysical Research: Atmospheres, 2014, 119: 1695–1707.

[9] Guilizia C, Camilloni I. Comparative analysis of the ability of a set of CMIP3 and CMIP5 global climate models to represent precipitation in South American[J]. International Journal of Climatology, 2015, 35: 583–595.

[10] Collier E, Immerzeel W W. High-resolution modeling of atmospheric dynam-

ics in the Nepalese Himalaya[J]. Journal of Geophysical Research: Atmospheres, 2015, 120(19): 9882–9896.

[11] Zhuang Zhaorong, Jiang Yuan, Tian Weihong, et al. CMA-MESO hourly rapid-update assimilation forecast system and preliminary analysis of its short-term forecasting effect[J]. Chinese Journal of Atmospheric Sciences, 2023, 47(4): 925–942. [Zhuang Zhaorong, Jiang Yuan, Tian Weihong, et al. Hourly rapid updating assimilation forecast system of CMA-MESO and preliminary analysis of short-term forecasting effect[J]. Chinese Journal of Atmospheric Sciences, 2023, 47(4): 925–942.]

[12] Liu Jing, Ren Chuan, Zhao Ziqi, et al. Comparative verification analysis of heavy-precipitation forecasts from high-resolution models in multiple regions[J]. Meteorological Monthly, 2022, 48(10): 1292–1302. [Liu Jing, Ren Chuan, Zhao Ziqi, et al. Comparative analysis on verification of heavy rainfall forecasts in different regional models[J]. Meteorological Monthly, 2022, 48(10): 1292–1302.]

[13] Huang Liping, Deng Liantang, Wang Ruichun, et al. Integration and application of key technologies of CMA-MESO[J]. Journal of Applied Meteorological Science, 2022, 33(6): 641–654. [Huang Liping, Deng Liantang, Wang Ruichun, et al. Key technologies of CMA-MESO and application to operational forecast[J]. Journal of Applied Meteorological Science, 2022, 33(6): 641–654.]

[14] Ye Mao, Chen Liangliu, Wang Jingzhuo, et al. Impact of horizontal resolution on the ensemble forecast skill of CMA-REPS[J]. Meteorological Monthly, 2025, 51(2): 153–166. [Ye Mao, Chen Liangliu, Wang Jingzhuo, et al. Impact of horizontal resolution on ensemble forecast skill of CMA-REPS[J]. Meteorological Monthly, 2025, 51(2): 153–166.]

[15] Cai Yi, Xu Zhifang, Gong Xi, et al. Evaluation of precipitation forecasts of the CMA-MESO model in summer 2021[J]. Journal of Arid Meteorology, 2023, 41(3): 503–515. [Cai Yi, Xu Zhifang, Gong Xi, et al. Evaluation of precipitation forecast of CMA-MESO model in summer of 2021[J]. Journal of Arid Meteorology, 2023, 41(3): 503–515.]

[16] He Mu, Yang Fuyan, Tang Weiyao, et al. Performance evaluation of high-resolution numerical models for precipitation forecasts during the 2023 flood season in Guizhou[J]. Journal of Mountain Meteorology, 2024, 48(6): 28–35. [He Mu, Yang Fuyan, Tang Weiyao, et al. Evaluation of three high-resolution operational weather forecast models for precipitation in Guizhou Province during the rainy season 2023[J]. Journal of Mountain Meteorology, 2024, 48(6): 28–35.]

[17] Noda A T, Oouchi K, Satoh M, et al. Importance of the subgrid-scale turbulent moist process: Cloud distribution in global cloud-resolving simulations[J]. Atmospheric Research, 2010, 96(2–3): 208–217.

[18] Bogenschutz P A, Krueger S K. A simplified PDF parameterization of subgrid-scale clouds and turbulence for cloud-resolving models[J]. Journal of Advances in Modeling Earth Systems, 2013, 5(2): 195–211.

[19] Qiao Liang, Zhang Qiang, Yue Ping, et al. Analysis of changes in the atmospheric boundary-layer structure during the transition from a non-monsoon region to a monsoon region[J]. Chinese Journal of Atmospheric Sciences, 2019, 43(2): 251–265. [Qiao Liang, Zhang Qiang, Yue Ping, et al. Analysis of changes in

the structure of atmospheric boundary layer from non-monsoon zone to monsoon zone[J]. Chinese Journal of Atmospheric Sciences, 2019, 43(2): 251-265.]

[20] Li Qingchun, Cheng Conglan, Quan Jiping, et al. Comparison of the characteristics and effects of low-level jets in two summer rainstorm processes of different intensities[J]. Meteorological Monthly, 2022, 48(11): 1384–1401.

[21] Yang Tao, Yang Lianmei. Climatic characteristics of atmospheric specific humidity in Xinjiang from 1961 to 2018 based on radiosonde observation data[J]. Desert and Oasis Meteorology, 2021, 15(3): 77–84.

[22] Ma Chao, Liu Yan, Liu Jing, et al. Spatial and temporal distribution characteristics of cloud water resources in the Tarim Basin in summer[J]. Arid Zone Research, 2025, 42(2): 223–235.

[23] Yang Xia, Zhou Hongkui, Xu Tingting, et al. Comparative analysis of the fine-scale characteristics of different types of rainstorms in southern Xinjiang during summer[J]. Arid Zone Research, 2021, 38(3): 747–756.

[24] Li Bingyuan, Pan Baotian, Han Jiafu. Basic terrestrial geomorphological types in China and their delineation indices[J]. Quaternary Sciences, 2008, 28(4): 535–543.

[25] Qiu Peiyu, Shen Xueshun. Application of the zero lateral flux scheme in the CMA-MESO model[J]. Journal of Tropical Meteorology, 2024, 40(3): 468–476.

[26] Wei Qing, Dai Kan, Lin Jian, et al. Evaluation of the 2016–2018 fine-gridded precipitation and temperature forecasts[J]. Meteorological Monthly, 2020, 46(10): 1272–1285.

[27] Zhang Jiabao, Su Qiyuan, Sun Shenqing, et al. Guide Handbook on Xinjiang Short-Term Weather Forecast[M]. Urumqi: Xinjiang People Press, 1986.

[28] Mou Huan, Chen Chunyan, Yang Xia, et al. Evaluation of hourly and daily precipitation forecasting performance of the CMA-MESO model in the warm season: a case study of the Ili River Valley[J]. Arid Land Geography, 2025, 48(2): 179–189.

[29] Xie Yiyun, Wang Jianjie. Preliminary study on the deviations and causes of precipitation prediction by the GRAPES kilometer-scale model in complex terrain areas of Southwest China[J]. Acta Meteorologica Sinica, 2021, 79(5): 732–749.

[30] Chen Haoming, Li Puxi, Zhao Yan. A review and outlook of verification and evaluation of precipitation forecasts at convection-permitting resolution[J].

Advances in Meteorological Science and Technology, 2021, 11(3): 155-164.

[31] Yang Liu, Yang Xia, Diao Peng, et al. Analysis of the forecast performance of the ECMWF model for the diurnal variation characteristics of summer precipitation on the northern slope of the Kunlun Mountains[J]. Arid Zone Research, 2025, 42(1): 27-39.

[32] Wang Xiaofeng, Xu Xiaolin, Xu Tong, et al. Verification of extreme precipitation forecasts from SMS-WARMSV 2.0[J]. Advances in Meteorological Science and Technology, 2017, 7(6): 67-74.

[33] Huang Chuhui, Niu Jinlong, Li Guoping, et al. A correction method for forecast precipitation in steep-terrain transition zones based on the Southwest Center WRF ADAS real-time modeling system[J]. Arid Meteorology, 2022, 40(2): 317-326.

[34] Zhou Qiuxue, Kang Lan, Jiang Xingwen, et al. Relationship between heavy rainfall and altitude in mountainous areas along the margins of the Sichuan Basin[J]. Meteorological Monthly, 2019, 45(6): 811-819.

[35] Li Zhe, Chen Jiong, Ma Zhanshan, et al. Deviation distribution features of CMA-GFS cloud prediction[J]. Journal of Applied Meteorological Science, 2022, 33(5): 527-540.

[36] Houze R A Jr. Orographic effects on precipitating clouds[J]. Reviews of Geophysics, 2012, 50(1): RG1001.

[37] Liu Yi, Chen Dehui, Hu Jianglin, et al. An impact study of the orographic effective scales for the GRAPES_MESO model with idealized numerical simulations[J]. Journal of Tropical Meteorology, 2011, 27(1): 53-62.

[38] Davies L A, Brown A R. Assessment of which scales of orography can be credibly resolved in a numerical model[J]. Quarterly Journal of the Royal Meteorological Society, 2001, 127: 1225-1237.

Evaluation of the forecast performance of the CMA-MESO model for warm-season precipitation over the Tarim Basin

MOU Huan1, ZHAO Li2, YANG Xia1, CHEN Chunyan1, LI Ruqi1, SUN Mingjing1

(1. Xinjiang Uygur Autonomous Region Meteorological Observatory, Urumqi 830002, Xinjiang, China;
2. Xinjiang Information Engineering School, Urumqi 830013, Xinjiang, China)

Abstract: The complex terrain of the Tarim Basin and active local convective weather during the warm season significantly increase uncertainty in numerical precipitation forecasting, making it difficult to meet the demand for refined precipitation forecasts. To better understand the forecast bias characteristics of the China Meteorological Administration Mesoscale (CMA-MESO) numerical weather prediction model in this region and enhance its forecast performance, this study systematically evaluated the CMA-MESO model' s capability in forecasting 24 hour accumulated precipitation and 1 hour accumulated precipitation during the warm season (from May to September) over the Tarim Basin. The evaluation used forecast data from the CMA-MESO model and surface observational data from the 2023–2024 warm seasons, employing multiple operational precipitation verification metrics. The results showed that the CMA-MESO model could generally simulate the spatial distribution characteristics of precipitation in the basin during the warm season. As the altitude increased, the Threat Score (TS) rose, while the Probability of Correct detection (PC) decreased. The forecast bias in precipitation frequency was most significant in the moderate-elevation zones, with a positive bias center zone concentrated on the windward slope of the western Kunlun Mountains. The maximum forecast bias in 24 hour accumulated precipitation frequency and 1 hour accumulated precipitation frequency was 44.1% and 11.6%, respectively. The diurnal variation of the TS was significant, characterized by higher values during the day and lower values at night; this diurnal bias was most pronounced in the low-altitude areas. In terms of precipitation intensity, the TS for effective 24 hour accumulated precipitation and 1 hour accumulated precipitation in the CMA-MESO model were 34.4% and 8.8%, respectively. However, the forecast performance declined rapidly with increasing precipitation intensity. Notably, the overestimation in the frequency of 1 hour accumulated precipitation was primarily confined to the 0.1–2 mm range, and targeted corrections to forecasts in this interval could more effectively reduce the overall precipitation frequency Bias. The results of this study provide a scientific basis for improving the precipitation forecasts of the CMA-MESO model in complex terrain in arid areas and support disaster prevention and mitigation operations.

Keywords: Tarim Basin; warm season; CMA-MESO; rainfall; forecast evaluation

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