ACCSCIENCE PUBLISHING

Asian Journal of Water, Environment and Pollution

ORIGINAL RESEARCH ARTICLE

Adaptive bias correction for subseasonal temperature forecasting in East Asia

Hao Wu[1]* and Haonan Ji[2]

[1]School of Automation, Nanjing University of Information Science and Technology, Nanjing, Jiangsu, China

[2]Key Laboratory of Ecosystem Carbon Source and Sink, China Meteorological Administration, Wuxi University, Wuxi, Jiangsu, China

Abstract

Subseasonal (2–6-week) temperature prediction is important for climate-risk management, energy dispatch, and early warning of high-impact weather over East Asia. However, operational dynamical forecast systems often exhibit substantial lead-dependent biases in this region. In this study, we evaluate an adaptive bias correction (ABC) framework for East Asian subseasonal 2-m temperature forecasts based on Climate Forecast System Version 2 and European Centre for Medium-Range Weather Forecasts data, and compare its performance with raw forecasts and conventional debiasing. Results show that ABC consistently reduced forecast errors and improved spatial consistency across lead times, with particularly clear benefits during winter cold-surge conditions and summer high-temperature periods. The improvements arise from the complementary integration of climatological, dynamical, and persistence-based information, which helps address the spatiotemporal heterogeneity of forecast errors. These results suggest that adaptive post-processing offers a practical approach to improving operational subseasonal temperature prediction over East Asia.

Keywords: Machine learning; Subseasonal-to-seasonal prediction; Temperature forecasting; East Asia; Root mean square error; Bias correction

***Corresponding author:**
Hao Wu
(202312490014@nuist.edu.cn)

Citation: Wu H, Ji H. Adaptive bias correction for subseasonal temperature forecasting in East Asia. *Asian J Water Environ Pollut.* 2026;23(4):026100057. doi: 10.36922/AJWEP026100057

Received: March 3, 2026

Revised: April 28, 2026

Accepted: April 30, 2026

Published online: May 21, 2026

1. Introduction

Subseasonal-to-seasonal (S2S; 2–6-week) temperature prediction bridges the gap between medium-range weather forecasting and seasonal climate outlooks and is important for climate-risk management in agriculture, energy dispatch, and heat–cold hazard preparedness. The development of coordinated international efforts, particularly the S2S Prediction Project, has enabled systematic intercomparison of forecast systems and accelerated methodological advances through standardized datasets and near-real-time products.[1,2] Recent assessments further summarize the operational achievements of the S2S project and demonstrate that artificial intelligence-based postprocessing can systematically enhance extended-range skill in community challenges.[3,4] Despite these advances, actionable temperature prediction at 3–6-week leads remains challenging because forecast information gradually shifts from strong initial-condition control toward weaker boundary forcing, and predictability at this range is often flow-dependent and intermittent.[5,6]

At subseasonal leads, forecast skill is strongly modulated by slowly varying large-scale modes and their teleconnections, among which the Madden–Julian Oscillation (MJO) is particularly important.[7–9] The MJO can influence extratropical circulation weeks in advance, thereby providing a pathway for extended-range temperature predictability. However, this predictability is often conditional: skill is enhanced when tropical–extratropical coupling is well represented, but can decline rapidly when internal variability dominates or when model biases distort teleconnection pathways.[10,11] These characteristics make the extraction and calibration of low-frequency predictive signals especially important for operational temperature forecasting. Recent studies further show that preceding MJO conditions can enhance the subseasonal predictability of East Asian cold events through distinct dynamical pathways, while MJO prediction skill itself varies substantially across events depending on the large-scale background state, reinforcing the flow-dependent nature of subseasonal predictability.[12,13]

East Asia is a particularly challenging region for S2S temperature forecasting because strong land–sea thermal contrasts, complex topography, and monsoon-related intraseasonal variability amplify the spatiotemporal heterogeneity of forecast errors. In addition, high-impact events such as winter cold surges and summer heat episodes place further demands on forecast reliability.[14] Operational coupled models, such as the Climate Forecast System Version 2 (CFSv2), provide valuable ensemble forecasts at S2S leads[15,16], but systematic errors in circulation, land–atmosphere coupling, and parameterized physics can lead to pronounced regional and lead-dependent temperature biases over East Asia.[17] These challenges motivate a shift from purely "raw-model" interpretation toward systematic postprocessing designed to reduce error growth and improve anomaly-level consistency.[18] Postprocessing has therefore become an important means of improving forecast usability. Traditional approaches, including climatology-based debiasing, model output statistics, and distribution-mapping methods, can reduce part of the mean bias.[19,20] For East Asia, methods that incorporate regional circulation-dependent error structures have shown additional value beyond simple mean-bias correction.[21,22] Nevertheless, numerous conventional statistical methods rely on quasi-stationary forecast–observation relationships and fixed functional forms, which may limit their effectiveness when forecast errors vary strongly across seasons and circulation regimes.[23]

Recent advances in machine learning (ML) have provided new opportunities for subseasonal postprocessing by enabling nonlinear error representation and flexible calibration.[24] ML methods have been successfully applied to ensemble forecast postprocessing and model-error correction, and have also shown promise in improving key subseasonal predictability carriers such as MJO evolution.[25–28] More broadly, recent studies suggest that data-driven methods can learn flow-dependent forecast errors and enhance extended-range prediction skill in selected applications.[29–31] At the same time, purely data-driven prediction does not consistently outperform state-of-the-art numerical weather prediction across all regimes, highlighting the value of hybrid approaches that retain dynamical forecast information while adaptively correcting systematic errors.[32–34]

Within this hybrid framework, adaptive bias correction (ABC) is a promising postprocessing approach because it integrates climatological, dynamical, and persistence-based information in a unified and interpretable manner.[35] Rather than generating new predictability, ABC is designed to better exploit existing low-frequency forecast signals by reducing systematic bias and sampling noise. This makes it particularly relevant for subseasonal forecasting, where forecast errors are often nonstationary, flow-dependent, and regionally heterogeneous.[36]

In this study, we evaluate the performance of ABC for subseasonal (2–6-week) 2-m temperature prediction over East Asia using ensemble forecasts from CFSv2 and European Centre for Medium-Range Weather Forecasts (ECMWF)-based systems. We compare raw forecasts, conventional debiasing, and ABC-corrected forecasts in terms of root mean square error (RMSE) and anomaly correlation coefficient (ACC), and further assess their behavior in representative cold-surge and heat-event cases. By focusing on strict train–test separation and observability-consistent validation, we aim to clarify the practical value of ABC for operational subseasonal temperature forecasting over East Asia.

2. Data and methods

2.1. Data and experimental design

Historical temperature observations were obtained from the National Oceanic and Atmospheric Administration–Climate Prediction Center (CPC) global temperature dataset.[37] As surface-based truth, we use daily gridded 2-m air temperature (°C) at 1.5° × 1.5° resolution. A 14-day moving average was applied to both observations and forecasts to match the subseasonal prediction targets considered in this study. This processing suppressed day-to-day synoptic variability and emphasized lower-frequency temperature anomalies, which were more relevant to the weeks 1–2, 3–4, and 5–6 forecast evaluation.

In this study, the subseasonal targets for weeks 1–2, 3–4, and 5–6 were defined as three consecutive, non-overlapping 14-day windows following each forecast initialization, corresponding to days 1–14, 15–28, and 29–42, respectively. For each initialization, both forecast and observed temperatures were averaged over the same 14-day target window for verification. Although adjacent windows may overlap in an intermediate daily running-mean time series, the three lead windows used for evaluation in this study were non-overlapping. The study domain was East Asia (0–60° N, 70–140° E) (Figure 1).

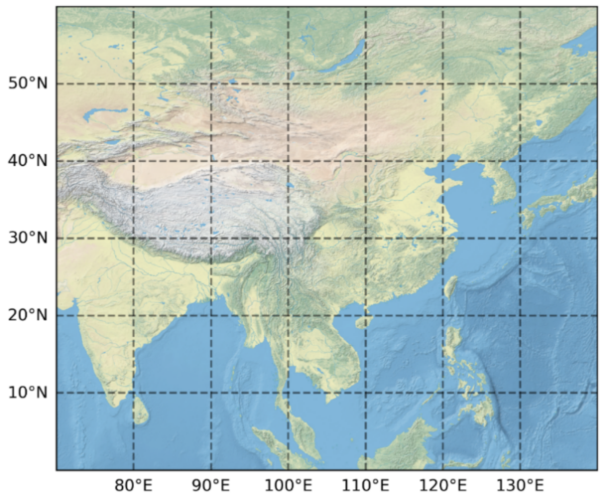

Figure 1. Study domain in East Asia (0–60° N, 70–140° E)

We analyzed ensemble forecasts from CFSv2 and ECMWF. CFSv2 forecasts were obtained from the SubX project.[38] CFSv2 is a coupled dynamical model developed by the United States National Centers for Environmental Prediction and provides 2-m temperature forecasts with a 6-hour temporal resolution and 32 ensemble members. In this study, the available real-time CFSv2 forecasts covered 2017–2021, and the corresponding hindcast archive used for bias correction covered 1999–2017. For each grid point and valid date, the four 6-hourly forecasts were first averaged to obtain daily values, and a 14-day moving average was then applied. After this averaging, an additional temporal smoothing step was introduced by averaging the forecast valid for target date t at lead l with the forecast initialized one day earlier and verified for the same target date (i.e., (t, l) and $(t - 1, l + 1)$). This step was not intended to further remove synoptic-scale variability after the 14-day averaging, but rather to reduce small discontinuities between adjacent initialization cycles and to stabilize lead-dependent forecast estimates. Because the two forecasts corresponded to the same target date (and therefore the same 14-day verification window) but originated from adjacent initialization cycles, averaging them yielded a more stable estimate for subseasonal verification without changing the forecast's temporal definition.

The ECMWF forecasts were from the S2S database, which included one control member and 50 perturbed members. In our dataset, the real-time ECMWF forecasts used for evaluation covered 2015–2021 and included 51 members (one control member and 50 perturbed members), whereas the corresponding reforecast archive used for debiasing covered 1995–2014 and included 11 members (one control member and 10 perturbed members). The same preprocessing procedure was applied to the ECMWF members, and the ensemble mean was used for deterministic evaluation after preprocessing and correction. The two forecast systems were evaluated separately throughout this study rather than merged into a single multi-model ensemble. After preprocessing, both forecast datasets were interpolated to a common 1.5° × 1.5° grid for comparison with CPC observations. The analysis period covered 2018–2021. Although this period was sufficient to evaluate representative subseasonal forecast behavior, the relatively short sample length may limit the robustness of the results, especially for event-based statistics.

2.2. Conventional bias-correction procedures

For ECMWF's 51-member ensemble, we applied a dynamic climatological reference for correction.[39] For each target forecast date and grid point, we extracted historical data from the 1995–2014 reforecast archive within a ±6-day window around the same month–day and computed the mean of the 11-member ECMWF reforecast ensemble over that window as the climatological bias baseline. This bias baseline was estimated from the reforecast ensemble statistics and then applied member-wise to each member of the real-time 51-member ECMWF ensemble. We subtracted this baseline from each member of the real-time forecast ensemble, and add the long-term climatological mean of observations over the same window to obtain the bias-corrected probabilistic forecasts. The observational climatology used in this debiased (DEB) procedure was defined using the same calendar-window framework as the anomaly reference used for ACC calculation.

For CFSv2, we follow the correction workflow in[40]. We construct a 32-member forecast ensemble by aggregating the most recent eight initialization times at 6-hour intervals. For each target date and grid point, we extract historical dates from 1999–2017 that exactly match the month–day and compute the mean climatological bias from the corresponding 32-member CFSv2 hindcast ensemble. We then subtract this hindcast climatological mean from each member of the real-time ensemble and add the observational climatological mean for the same calendar date, forming a two-way compensation that removes systematic bias. As with ECMWF, the correction is applied to each member of the real-time CFSv2 ensemble, while the bias reference is estimated from the corresponding hindcast ensemble statistics.

2.3. Adaptive bias correction algorithm

We implement the ABC algorithm as described by Mouatadid *et al.*[35] (Figure 2). The procedure is summarized below:

(i) Inputs included dynamical ensemble forecasts (51 members for ECMWF: one control + 50 perturbed; 32 members for CFSv2), historical gridded observations of temperature, a climatological baseline, the target date, lead time, and the number of training years.
(ii) Climatology++ (climatological baseline optimization): based on the seasonal characteristics of the target date, Climatology++ selected an optimal climatological reference from historical observations around the same season (e.g., the past 12 years within ±35 days of the same calendar date). Using a robust error criterion, it dynamically chose the optimal climatological median as a static reference, improving seasonal consistency.
(iii) Dynamical++ (dynamical bias correction): Dynamical++ filtered historical data within a seasonally similar window, dynamically fused raw forecasts from multiple initialization dates and lead times to form an ensemble mean, computed the mean bias between historical observations and the historical ensemble mean, and added this bias as a correction term to the current dynamical forecast, reducing systematic warm/cold biases.
(iv) Persistence++ (trend modeling): To capture recent weather evolution, Persistence++ built grid-point regression models that integrated lagged observations, the climatological baseline, and the dynamical ensemble mean. The correction term was estimated via least squares to represent local temperature tendencies and to compensate for dynamical-model lag in responding to rapidly evolving events.
(v) The final ABC forecast was produced by uniformly averaging the outputs of Climatology++, Dynamical++, and Persistence++, ensuring balanced fusion of multiple information sources.
(vi) For each target date and each member of the dynamical

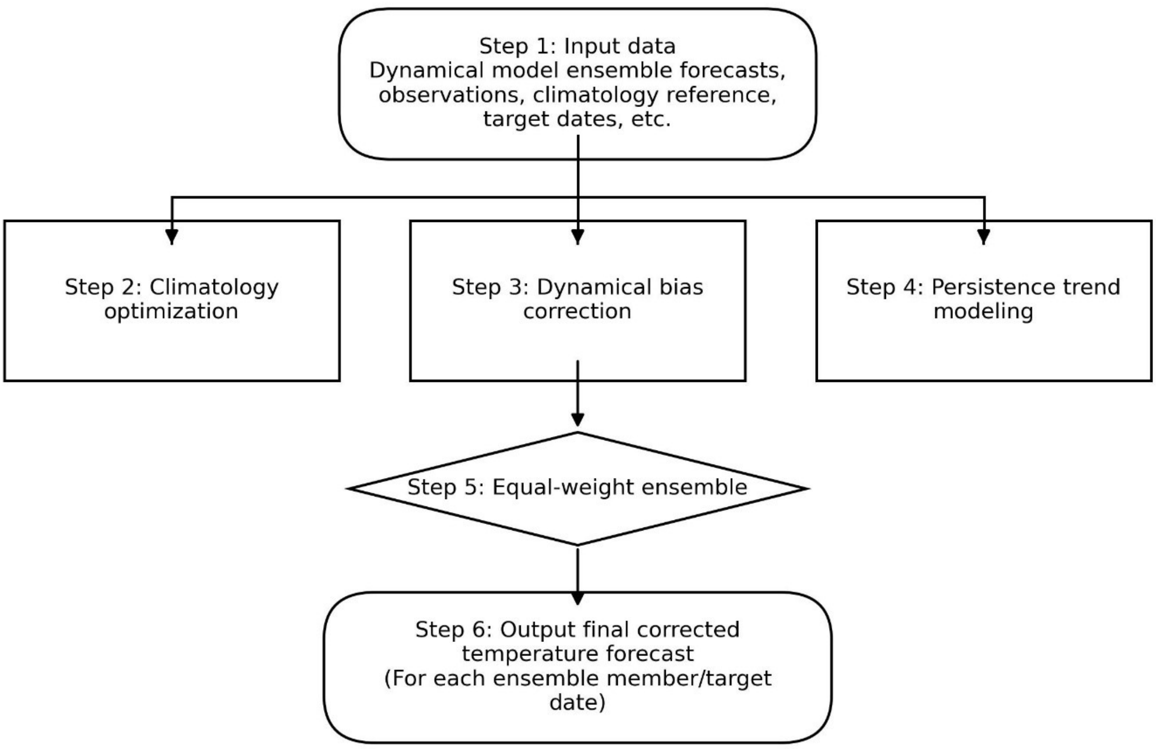

Figure 2. Schematic framework of the adaptive bias correction algorithm

ensemble, we took an equal-weight average of the three sub-model predictions to obtain the final bias-corrected daily temperature forecasts for 2018–2021.

2.4. Evaluation metrics

We evaluate forecast performance using the RMSE and ACC, representing error magnitude and spatial consistency, respectively.

The RMSE is defined as in **Equation 1**:

$$RMSE = \frac{1}{N_{forecasts}} \sum_{i}^{N_{forecasts}} \sqrt{N_{lat} N_{lon} \sum_{j}^{N_{lat}} \sum_{k}^{N_{lon}} L(j) \left(f_{i,j,k} - t_{i,j,k} \right)^2} \quad (1)$$

where f denotes the forecast, t denotes the observed temperature, and $L(j)$ is the latitude weight at grid j, defined as **Equation 2**:

$$L(j) = \frac{\cos\left(lat(j)\right)}{\frac{1}{N_{lat}} \sum_{j}^{N_{lat}} \cos\left(lat(j)\right)} \quad (2)$$

The latitude-weighted ACC is defined as **Equation 3**:

$$ACC = \frac{\sum_{i,j,k} L(j) f'_{i,j,k} t'_{i,j,k}}{\sqrt{\sum_{i,j,k} L(j) f'^{2}_{i,j,k} \sum_{i,j,k} L(j) t'^{2}_{i,j,k}}} \quad (3)$$

where primes (') denote anomalies relative to climatology. ACC quantifies the spatial similarity between the forecast and observed fields.

3. Results

3.1. Overall performance of adaptive bias correction in subseasonal temperature prediction

We systematically assessed the impact of ABC on subseasonal temperature forecasts from CFSv2 and ECMWF over East Asia. Unless otherwise stated, all percentage changes reported in this section were calculated relative to the corresponding raw forecast (RAW) for the same model, lead window, and season. Figure 3 summarizes the RMSE of 2-m temperature forecasts at lead windows of 1–2, 3–4, and 5–6 weeks during 2018–2021 for raw CFSv2 and ECMWF forecasts and their corresponding ABC-corrected results.

Overall, ABC markedly reduced RMSE relative to the RAW for both CFSv2 and ECMWF, demonstrating robust improvements across all lead windows. Figure 3A–C shows the multi-year mean RMSE at lead times of 1–2, 3–4, and 5–6 weeks, respectively. For both forecast systems, raw-model errors generally increased with lead time, with RMSE values exceeding 4 °C at 3–6-week leads. After ABC correction, RMSE was substantially reduced across all three lead windows, by approximately 60–75% relative to RAW.

Figure 3D–F further illustrates the seasonal mean RMSE averaged over 2018–2021. Raw forecasts exhibited clear seasonal dependence, with relatively larger errors in winter and summer, whereas ABC substantially weakened this seasonal contrast. In particular, ABC-corrected ECMWF forecasts consistently achieved the lowest RMSE in all seasons and lead times, remaining below 1.5 °C even at 5–6-week leads. These results indicated that ABC not only reduced the overall error magnitude but also improved the stability of subseasonal temperature forecasts across both seasons and lead times. The seasonal comparison further showed that the largest raw-forecast errors occur in summer and winter, while the ABC-corrected forecasts remained comparatively stable in all seasons. This behavior suggested that ABC was effective in reducing lead-dependent error accumulation and enhancing cross-season robustness. Taken together, Figure 3 demonstrates three main features of ABC performance: (i) consistent RMSE reduction relative to RAW, (ii) effective suppression of error growth at longer leads, and (iii) robust improvements across both forecast systems and all seasons.

3.2. Contributions of the adaptive bias correction sub-modules

To assess the contributions of the Dynamical++ and Persistence++ components to the ABC framework, Figure 4 presents seasonal comparisons of RMSE over East Asia for CFSv2 and ECMWF at lead windows of 1–2, 3–4, and 5–6 weeks during 2018–2021. Specifically, Figure 4A compares raw CFSv2 forecasts with their dynamically corrected counterparts (raw_cfs vs. cfsv2pp), and conventional persistence forecasts with the enhanced persistence scheme for CFSv2 (persistence vs. perpp_cfs), Figure 4B compares raw ECMWF forecasts with their dynamically corrected counterparts (raw_ec vs. ecmwfpp), and conventional persistence forecasts with the enhanced persistence scheme for ECMWF (persistence vs. perpp_ec).

Across both forecast systems, the dynamically corrected forecasts consistently yield lower RMSE than the corresponding raw model forecasts in all seasons. This improvement was evident at all three lead windows and was particularly clear at 3–4 and 5–6 weeks, indicating that Dynamical++ effectively reduced lead-dependent forecast errors. For both CFSv2 and ECMWF, the seasonal comparisons showed that the error reduction was robust across winter, spring, summer, and autumn, although the

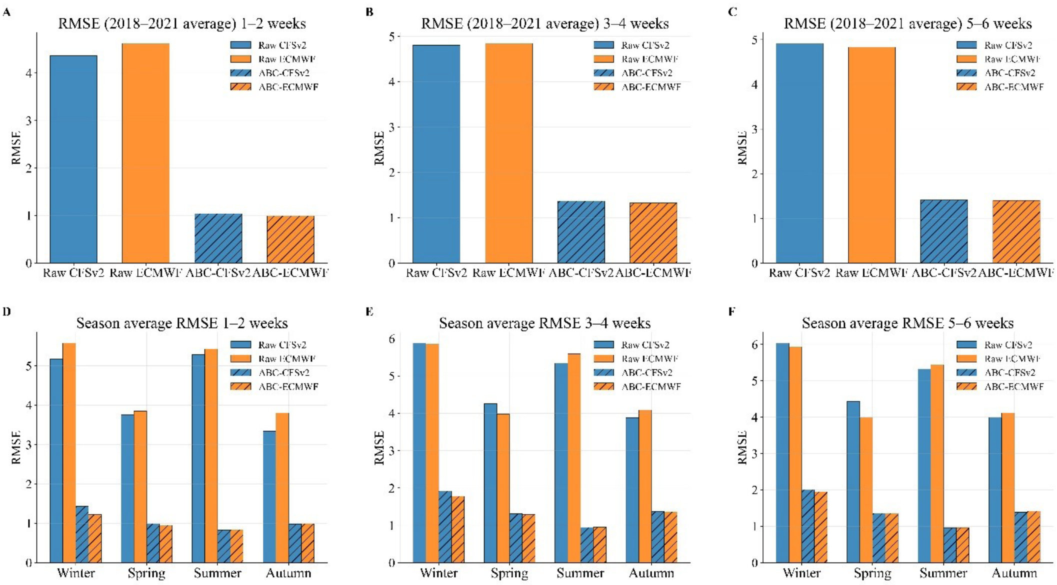

Figure 3. Comparison of root mean square error (RMSE) for subseasonal (2–6-week) 2-m temperature forecasts over East Asia during 2018–2021. (A–C) Multi-year mean RMSE at lead times of 1–2, 3–4, and 5–6 weeks for raw CFSv2, raw ECMWF, and their corresponding adaptive bias correction (ABC) forecasts. (D–F) Seasonal mean RMSE averaged over 2018–2021 for the same lead-time windows.

Abbreviations: ABC: Adaptive bias correction; CFSv2: Climate Forecast System Version 2; ECMWF: European Centre for Medium-Range Weather Forecasts; RMSE: Root mean square error.

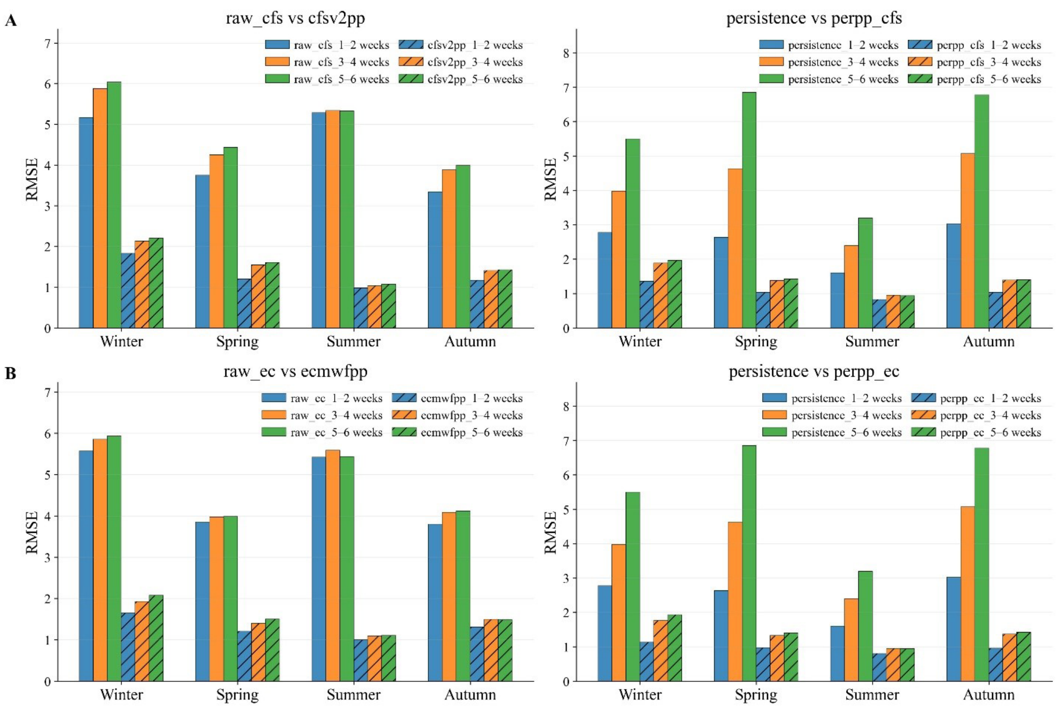

Figure 4. Seasonal comparison of RMSE for the Dynamical++ and Persistence++ components in the ABC framework over East Asia within 2018–2021. (A) Raw CFSv2 versus cfsv2pp, persistence versus perpp_cfs. (B) raw ECMWF versus ecmwfpp and persistence versus perpp_ec. Each panel shows results for winter, spring, summer, and autumn at lead windows of 1–2, 3–4, and 5–6 weeks.

Abbreviations: ABC: Adaptive bias correction; CFSv2: Climate Forecast System Version 2; ECMWF: European Centre for Medium-Range Weather Forecasts; RMSE: Root mean square error.

magnitude of improvement varies somewhat by season and forecast system. A similar pattern was found for the persistence-based pathway. Compared to conventional persistence, the enhanced persistence schemes (perpp_cfs and perpp_ec) produced substantially lower RMSE in all seasons and across all lead windows. The improvements were especially notable at longer leads, where conventional persistence degraded more rapidly. This behavior suggested that Persistence++ provided useful additional information for representing evolving temperature tendencies beyond what could be captured by simple persistence alone.

We did not include a direct comparison between Climatology and Climatology++ in these figures because the purpose of these two panels was to isolate lead-time-dependent error-growth control for forecast-like components. Climatology-based baselines were largely lead-invariant and were mainly used as a reference term in ABC rather than a competitive subseasonal forecast on their own. Instead, the effect of Climatology++ was reflected indirectly through its interaction with the dynamical and persistence corrections in the full ABC framework, where it stabilized the seasonal baseline and constrained drift under varying background climate states.

Therefore, Figure 4 focuses on the two components that were most directly associated with lead-dependent error growth and its reduction. Taken together, these results indicate that the overall improvement of ABC arises from the complementary effects of its sub-modules. Dynamical++ contributed by correcting systematic model-dependent errors, while Persistence++ improved the representation of recent temperature evolution. Climatology++, although not shown separately here, provided a stable seasonal baseline that helps constrain drift in the full ABC framework. Their combined use supported the robust improvements in forecast performance shown in Section 3.1.

3.3. Spatial characteristics of forecast improvement

We further compared spatial distributions of RMSE and ACC for CFSv2 and ECMWF under RAW, standard debiasing, and ABC at 1–2, 3–4, and 5–6-week leads (Figure 5). The number shown at the top of each panel denotes the corresponding domain-averaged RMSE or ACC over East Asia for that method and lead window. Overall, ABC consistently reduced errors and improved spatial coherence across regions and lead times, indicating strong adaptability and generalization over East Asia.

The RMSE maps showed that raw CFSv2 forecasts exhibited relatively large errors over the northern mid-latitudes, including the Mongolian Plateau and the North China Plain, and that these errors generally increased with lead time. Conventional debiasing reduced RMSE over part of the domain, but notable errors remained over northern East Asia. In comparison, ABC further reduced RMSE across most regions, with particularly clear improvements in North and Northeast China. Raw ECMWF forecasts showed lower RMSE at short leads, but higher errors at longer leads in several regions. DEB improved part of the short-lead error structure, but the spatial improvement remained uneven. ABC further reduced RMSE across much of the domain, with marked improvements along the eastern margin of the Tibetan Plateau and in the Yangtze–Huaihe River basin.

The ACC maps provided a consistent picture. At lead windows of 3–4 and 5–6 weeks, raw forecasts showed substantial spatial mismatch, with low or even negative ACC over northern Asia and the Tibetan Plateau. Conventional debiasing yielded only limited ACC improvement, and its benefit weakened noticeably at longer leads. In contrast, ABC substantially improved spatial correlation for both forecast systems. For CFSv2, ABC maintained relatively higher ACC values across short to long leads, while for ECMWF, ABC also produced broader regions of positive, relatively large ACC than either RAW or DEB. High-ACC regions under ABC are mainly located in Southeast Asia, the Yangtze–Huaihe basin, and South China.

Taken together, these results show that ABC not only reduced the magnitude of temperature forecast errors but also improved the spatial consistency of subseasonal temperature anomalies. Relative to the RAW, the RMSE reductions and ACC increases were generally more pronounced under ABC than under DEB, indicating that the adaptive combination of climatological, dynamical, and persistence-based information provided more effective spatial correction than static debiasing alone.

To further illustrate the improvement in spatial bias structure, Figure 6 compares the temperature-bias patterns of RAW and ABC at different lead windows. For both forecast systems, raw forecasts exhibited widespread cold bias across much of East Asia, particularly over the Mongolian and Tibetan Plateaus and Northeast Asia. Although cold bias remained evident at longer leads, the signed bias did not necessarily increase monotonically with lead time. This is because Figure 6 shows mean signed bias rather than total error magnitude, and positive and negative regional errors may partially offset each other after temporal and spatial averaging. Therefore, the main purpose of Figure 6 is to show the spatial structure of the systematic warm/cold bias and its correction by ABC, rather than to show the growth of total forecast error with lead time. From this perspective, ABC substantially reduced the magnitude of these biases and improved their

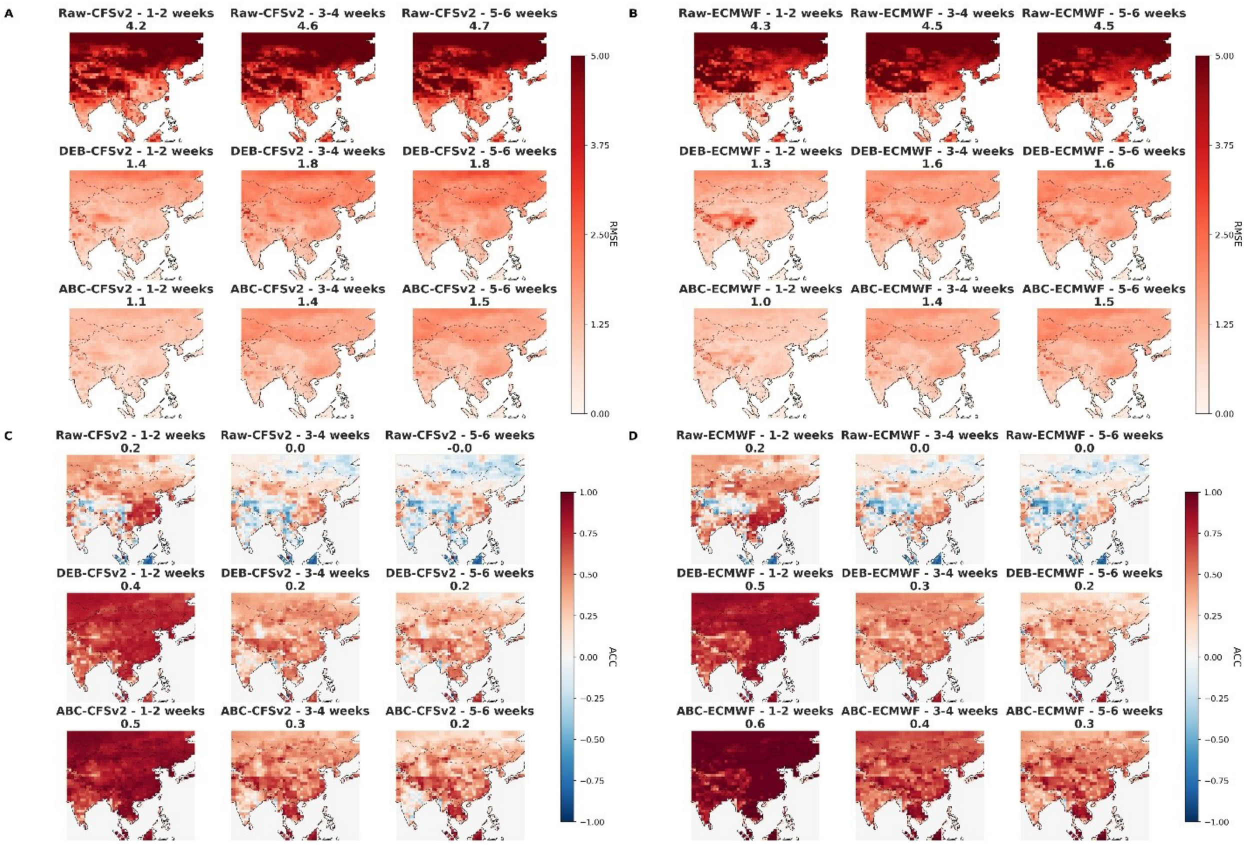

Figure 5. Spatial distributions of RMSE and ACC for CFSv2 and ECMWF under RAW, DEB, and ABC at lead windows of 1–2, 3–4, and 5–6 weeks over East Asia during 2018–2021. (A, B) RMSE. (C, D) ACC. The number shown at the top of each panel denotes the corresponding domain-averaged RMSE (for panels A and B) or domain-averaged ACC (for panels C and D).
Abbreviations: ACC: Anomaly correlation coefficient; CFSv2: Climate Forecast System Version 2; ECMWF: European Centre for Medium-Range Weather Forecasts; RMSE: Root mean square error.

spatial distribution, particularly over mid–high latitudes and regions with complex terrain. This improvement was also reflected in the domain-averaged mean absolute bias shown in each panel, which provided a complementary measure of the overall bias magnitude without cancellation between positive and negative regional errors.

For ECMWF, the ABC-corrected bias was compressed to a much narrower range over most of the domain, while for CFSv2, the bias reduction was especially evident at 3–4 and 5–6 weeks. Therefore, Figure 6 should be interpreted primarily as evidence of improved bias structure, whereas the lead-time growth of total forecast error was more directly reflected by the RMSE results in Figure 5. These improvements were consistent with ABC's ability to better control large-scale systematic errors in extended-range forecasting, although our analysis did not explicitly diagnose the underlying circulation processes.

3.4. Case-study evaluation of cold-surge and heat-event forecasts

To further examine the practical behavior of the different correction methods during high-impact events, we included two illustrative cases: a winter cold-surge event in January 2021 and a summer heat event on 7 July 2021. These two cases were selected as illustrative examples of winter and summer high-impact temperature events during 2018–2021 because both exhibited clear regional extent and pronounced temperature anomalies, making them suitable for visually comparing the forecast behavior of RAW, DEB, and ABC under contrasting cold and warm extreme conditions. They were not selected as the strongest events in the evaluation period, and they were intended to provide event-level illustrations rather than a statistical representation of all cold-surge and heat events during 2018–2021.

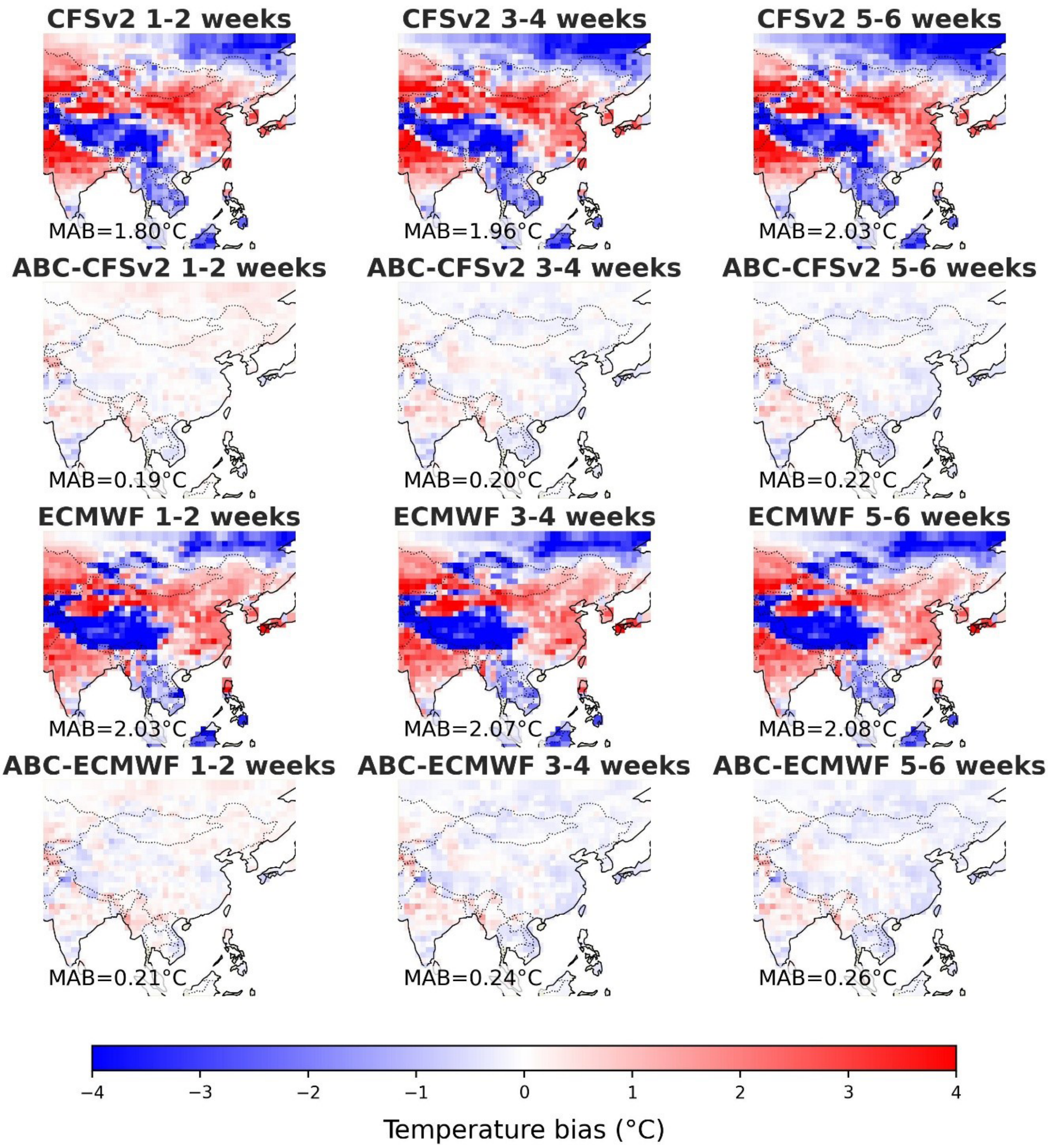

Figure 6. Spatial distribution of temperature bias for CFSv2 and ECMWF under raw forecasts and ABC at different lead times.
Abbreviations: ABC: Adaptive bias correction; CFSv2: Climate Forecast System Version 2; ECMWF: European Centre for Medium-Range Weather Forecasts; MAB: Mean absolute bias.

Figure 7 shows the cold-surge case, in which cold air invaded North and Northeast China, with temperatures dropping below −10 °C. Raw ECMWF generally overestimated temperature, with 1-2 week and 3-4 week RMSE of 5.59 °C and 6.40 °C, respectively. Conventional debiasing reduced part of the mean error but still left notable warm bias over key cold-surge regions. In comparison, ABC achieved the lowest RMSE across all lead windows; bias maps showed that ABC suppressed warm-bias spread and kept biases mostly within ±1.5 °C, indicating improved correction performance in this cold-surge case.

Figure 8 shows the July 2021 heat case over northern East Asia (30°–47° N, 110°–135° E). Raw CFSv2 captured the broad warm anomaly pattern but exhibited increasing warm bias over North and Northeast China as lead time increased, with the RMSE reaching 3.87 °C at the longest shown lead. Conventional debiasing reduced warm bias at short leads (e.g., the 1–2-week RMSE was reduced to 0.84 °C), but at longer leads, it introduced notable cold bias in parts of northern East Asia. ABC consistently produced the lowest RMSE among the compared methods, reducing the 1–2-week RMSE to 0.73 °C and limiting the longest-lead RMSE to 2.99 °C.

Taken together, these two illustrative cases show that ABC improved forecast performance for both cold and heat extremes relative to RAW and DEB. In both cases, the ABC-corrected fields exhibited lower RMSE and more realistic spatial temperature patterns than the RAW fields, suggesting that the adaptive framework could improve forecast accuracy in these two high-impact temperature cases. However, these case studies were intended only as illustrative event-level examples, and the robustness of the improvement across a larger set of cold-surge and heat events remained to be assessed in future work.

4. Discussion

This study demonstrated that ABC substantially improved subseasonal (2–6-week) temperature forecasts over East Asia relative to both raw dynamical outputs and conventional bias-correction methods. The improvements were evident across multiple dimensions, including reduced error magnitude, enhanced spatial coherence, and increased robustness during extreme temperature

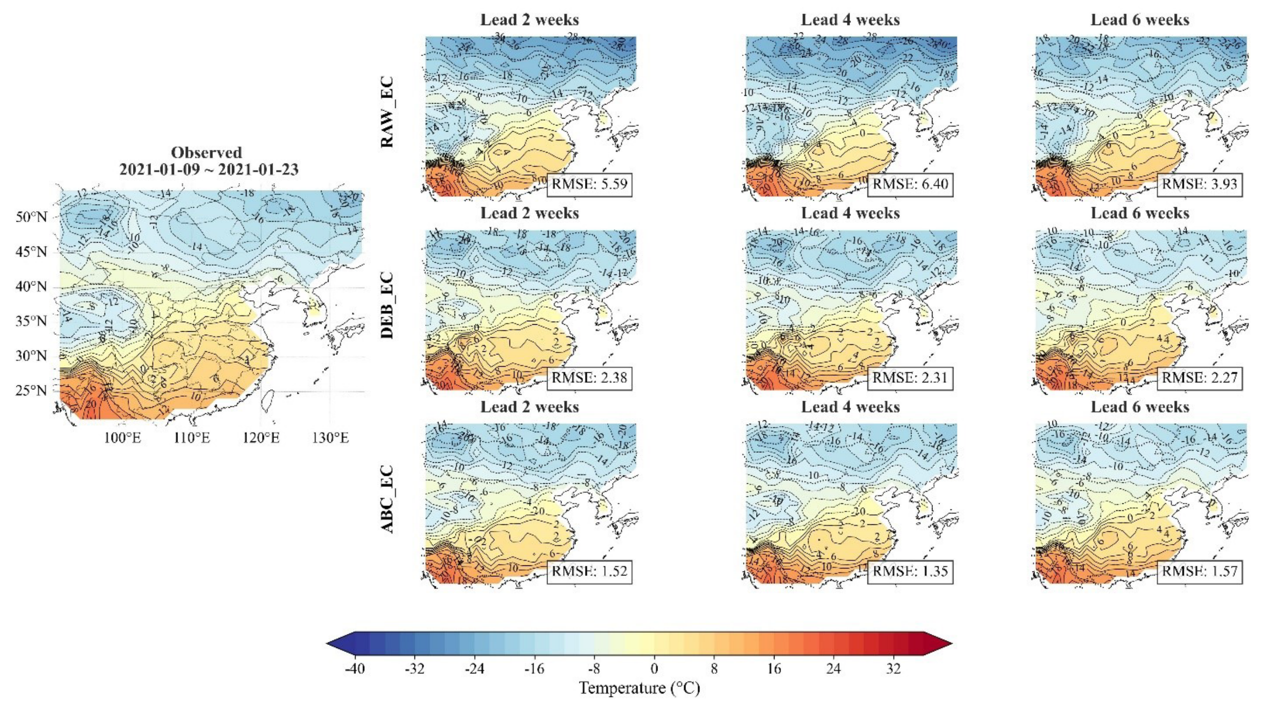

Figure 7. Observed and forecast 2-m temperature fields for the January 2021 East Asian cold-surge event under different methods at different lead windows. Shading denotes 2-m temperature (°C), and the black contours indicate 2-m temperature isotherms at 2 °C intervals. The boxed value in each forecast panel denotes the corresponding RMSE.

Abbreviations: ABC: Adaptive bias correction; DEB: Debiased; EC: European Centre for Medium-Range Weather Forecasts; RAW: Corresponding raw forecast; RMSE: Root mean square error.

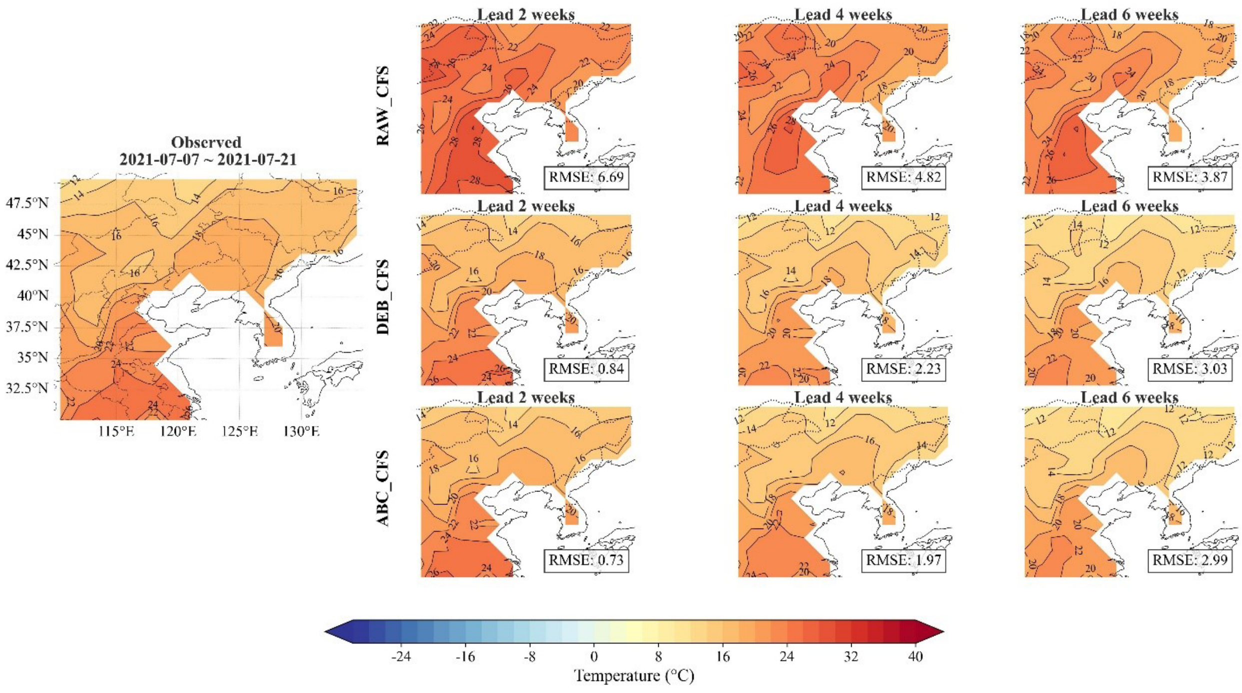

Figure 8. Observed and forecast 2-m temperature fields for the 7 July 2021 heat event over northern East Asia under different methods at different lead windows. Shading denotes 2-m temperature (°C), and the black contours indicate 2-m temperature isotherms at 2 °C intervals. The boxed value in each forecast panel denotes the corresponding RMSE.

Abbreviations: ABC: Adaptive bias correction; CFS: Climate Forecast System Version 2; DEB: Debiased; RAW: Corresponding raw forecast; RMSE: Root mean square error.

events. These results provide strong support for our working hypothesis that static, climatology-based bias correction is insufficient at the subseasonal range, and that adaptive integration of climatology, dynamical-model error characteristics, and recent persistence information is essential for suppressing nonlinear error growth.

4.1. Interpretation in the context of previous studies

Previous studies have consistently shown that subseasonal temperature prediction skill deteriorates rapidly with lead time due to the combined effects of initial-condition uncertainty, model-physics deficiencies, and climate drift, particularly over regions influenced by complex circulation systems, such as the East Asian monsoon. Traditional bias-correction approaches—such as climatological anomaly adjustment, model output statistics, and quantile mapping—have been shown to remove part of the mean bias but often fail to maintain skill at longer leads because they assume temporal stationarity in forecast errors.[41] Our results are fully consistent with these findings: while conventional debiasing reduces short-lead errors, its effectiveness decays markedly beyond 3–4 weeks and offers limited improvement in spatial anomaly correlation.

In contrast, the ABC framework aligns with recent advances that emphasize adaptive and hybrid correction strategies. Similar to prior work across North America and Europe, we found that ABC significantly improved both RMSE and ACC at extended lead times. Furthermore, this study extended earlier findings in three important ways. First, it demonstrated that ABC remained effective over the East Asian monsoon region, where strong intraseasonal variability, land–sea thermal contrasts, and complex topography posed additional challenges. Second, the method showed consistent skill gains across two structurally different forecast systems (CFSv2 and ECMWF), indicating that its effectiveness was not model-specific. Moreover, the event-based analyses in Section 3.4 showed that ABC also yielded lower RMSE and more realistic spatial temperature patterns than RAW and DEB during the January 2021 cold-surge case and the July 2021 heat event.

4.2. Possible interpretation of adaptive bias correction performance

The improved performance of ABC likely reflects the complementary roles of its three sub-components.

Climatology++ provides a dynamically optimized seasonal baseline that mitigates climate drift and suppresses seasonally dependent biases, both of which are common in CFSv2 and ECMWF. Dynamical++ reduces systematic model-dependent errors by leveraging historical forecast–observation relationships under similar seasonal conditions, thereby improving the correction of seasonally varying temperature biases. Persistence++ captures recent temperature tendencies and compensates for the lagged response of dynamical models to rapidly evolving atmospheric processes.

Importantly, the equal-weight fusion of these components allows ABC to balance stability and adaptability. During relatively stable periods, climatological information constrains overfitting and noise amplification, whereas during transition seasons or extreme events, persistence and dynamical adjustments become more influential. This multi-source combination is consistent with the observed reductions in RMSE and improvements in spatial anomaly correlation, particularly at 5–6-week leads where conventional approaches show limited effectiveness.

4.3. Implications for subseasonal prediction and applications

From an operational perspective, the improvement shown here is relevant for early warning and decision-making in East Asia. More stable temperature forecasts at subseasonal leads can support agriculture, energy management, and risk preparedness for cold surges and heat events. In this sense, ABC can be viewed as a practical adaptive post-processing framework that helps convert existing dynamical forecast information into more usable temperature guidance.

More broadly, the results highlight a shift in subseasonal forecasting philosophy: rather than treating bias correction as a purely statistical post-processing step, it should be viewed as a dynamic extension of the forecast system, capable of interacting with both model physics and observed atmospheric evolution. This perspective is consistent with emerging hybrid dynamical–statistical frameworks and supports the growing integration of adaptive algorithms into operational S2S prediction systems.

4.4. Limitations and future research directions

Despite its advantages, several limitations warrant further investigation. First, the present study was conducted at a relatively coarse spatial resolution (1.5° × 1.5°). While this is appropriate for large-scale assessment, finer-resolution forecasts may exhibit different bias structures, particularly over complex terrain such as the Tibetan Plateau and its surrounding regions. Future work should evaluate ABC using higher-resolution forecast systems to assess its scalability and effectiveness for local extremes.

Second, the computational cost of ABC is higher than that of static debiasing because it requires rolling training, historical-analog-based sampling, and dynamic optimization of multiple correction pathways. Although this remains feasible for retrospective evaluation, further simplification may be required for real-time operational use, implying a trade-off between forecast accuracy and computational efficiency.

Finally, although this study focuses on temperature, extending ABC to other subseasonal variables may be challenging. For example, precipitation exhibits stronger intermittency, greater non-Gaussianity, and higher sensitivity of extremes to small forecast shifts, all of which may complicate both calibration and validation. These issues should be examined carefully in future work.

In summary, this study provides strong evidence that ABC represents a robust and transferable pathway for improving subseasonal temperature prediction over East Asia. By explicitly accounting for spatiotemporal heterogeneity in model errors and leveraging multiple sources of predictive information, ABC offers a valuable bridge between dynamical forecasting and practical climate services.

5. Conclusion

In conclusion, ABC substantially improves subseasonal (2–6-week) 2-m temperature forecasts over East Asia compared with raw forecasts and conventional debiasing. By integrating climatological, dynamical, and persistence information, ABC reduces forecast errors, improves spatial consistency, and remains effective across seasons, lead times, forecast systems, and representative cold-surge and heat-event cases. These findings suggest that ABC is a robust and transferable postprocessing framework with strong potential for operational subseasonal temperature prediction and climate-risk decision support in East Asia.

Acknowledgments

None.

Funding

This study was jointly funded by China State Railway Group Science and Technology Research and Development Program (N2024T008), Guangdong Basic and Applied Basic Research Foundation.

Conflict of interest

The authors declare they have no competing interests.

Author contributions

Conceptualization: Hao Wu
Formal analysis: Hao Wu
Investigation: All authors
Methodology: Hao Wu
Writing–original draft: Hao Wu
Writing–review & editing: Hao Wu

Availability of data

All data used in this study are publicly available. The temperature data from the NOAA Climate Prediction Center (CPC) are available at ftp://ftp.cdc.noaa.gov/Datasets/cpc_global_temp/. The CFSv2 forecast data produced by the Subseasonal Experiment (SubX) can be retrieved from http://iridl.ldeo.columbia.edu/SOURCES/.Models/.SubX/.NCEP/. ECMWF S2S real-time forecast data are available at https://apps.ecmwf.int/datasets/data/s2s-realtime-instantaneous-accum-ecmf/levtype=sfc/type=cf/, and the corresponding reforecast data are available at https://apps.ecmwf.int/datasets/data/s2s-reforecasts-instantaneous-accum-ecmf/levtype=sfc/type=cf/.

References

1. Vitart F, Ardilouze C, Bonet A, *et al*. The Subseasonal to Seasonal (S2S) Prediction Project Database. *Bull Am Meteorol Soc.* 2017;98(1):163-173.

   doi: 10.1175/BAMS-D-16-0017.1

2. Vitart F, Robertson AW, Anderson DLT. Subseasonal to Seasonal Prediction Project: Bridging the Gap between Weather and Climate. *WMO Bull.* 2012;61(2):23-28. Available from: https://wmo.int/media/magazine-article/subseasonal-seasonal-prediction-project-bridging-gap-between-weather-and-climate [Last accessed on March 3, 2026].

3. Vitart F, Robertson AW, Brookshaw A, *et al*. The WWRP/WCRP S2S project and its achievements. *Bull Am Meteorol Soc.* 2025;106(5):E791-E808.

   doi: 10.1175/BAMS-D-24-0047.1

4. Vitart F, Robertson AW, Spring A, *et al*. Outcomes of the WMO prize challenge to improve subseasonal to seasonal predictions using artificial intelligence. *Bull Am Meteorol Soc.* 2022;103(12):E2878-E2886.

   doi: 10.1175/BAMS-D-22-0046.1

5. Pegion K, Sardeshmukh PD. Prospects for improving subseasonal predictions. *Mon Weather Rev.* 2011;139(11):3648-3666.

   doi: 10.1175/MWR-D-11-00004.1

6. von Neumann J. Some remarks on the problem of forecasting climatic fluctuations. In: Pfeffer RL, ed. *Dynamics of Climate*. Oxford, England: Pergamon Press; 1960:9-11.

   doi: 10.1016/B978-1-4831-9890-3.50009-8

7. Xiang B, Harris L, Delworth TL, *et al*. S2S prediction in GFDL SPEAR: MJO diversity and teleconnections. *Bull Am Meteorol Soc.* 2022;103(2):E463-E484.

   doi: 10.1175/BAMS-D-21-0124.1

8. Waliser DE, Jones C, Schemm JKE, Graham NE. A statistical extended-range tropical forecast model based on the slow evolution of the Madden-Julian oscillation. *J Clim.* 1999;12(7):1918-1939.

   doi: 10.1175/1520-0442(1999)012<1918:ASERTF>2.0.CO;2

9. Stan C, Zheng C, Chang EKM, *et al*. Advances in the prediction of MJO teleconnections in the S2S forecast systems. *Bull Am Meteorol Soc.* 2022;103(6):E1426-E1447.

   doi: 10.1175/BAMS-D-21-0130.1

10. Tseng KC, Barnes EA, Maloney E. The importance of past MJO activity in determining the future state of the midlatitude circulation. *J Clim.* 2020;33(6):2131-2147.

    doi: 10.1175/JCLI-D-19-0512.1

11. Zhou Y, Thompson KR, Lu Y. Mapping the relationship between Northern Hemisphere winter surface air temperature and the Madden-Julian oscillation. *Mon Weather Rev.* 2011;139(8):2439-2454.

    doi: 10.1175/2011MWR3587.1

12. Wang L, Zhou X, Hsu,PC. Preceding MJO enhances the subseasonal predictability of East Asian cold events via two dynamic pathways. *Sci China Earth Sci.* 2026;69(4):1259-1272.

    doi: 10.1007/s11430-025-1832-2

13. Zhou X, Wang L, Hsu PC, Li T, Xiang B. Understanding the factors controlling MJO prediction skill across events. *J Clim*. 2024;37(20):5323-5336.

    doi: 10.1175/JCLI-D-23-0635.1

14. Weyn JA, Durran DR, Caruana R. Can machines learn to predict weather? Using deep learning to predict gridded 500-hPa geopotential height from historical weather data. *J Adv Model Earth Syst.* 2019;11(8):2680-2693.

    doi: 10.1029/2019MS001705

15. Saha S, Moorthi S, Wu X, *et al*. The NCEP Climate Forecast System Version 2. *J Clim.* 2014;27(6):2185-2208.

    doi: 10.1175/JCLI-D-12-00823.1

16. Pegion K, Kirtman BP, Becker E, *et al*. The Subseasonal Experiment (SubX): A multimodel subseasonal prediction experiment. *Bull Am Meteorol Soc.* 2019;100(10):2043-2060.

doi: 10.1175/BAMS-D-18-0270.1

17. Tang S, Qiao S, Feng T, *et al*. Predictability of the mid-summer surface air temperature over the Yangtze River valley in the NCEP Climate Forecast System. *Int J Climatol.* 2021;41(2):811-829.

 doi: 10.1002/joc.6670

18. Liu L, Bai H, Feng G, *et al*. Evaluation and correction of sub-seasonal dynamic model forecast of precipitation in eastern China. *Clim Dyn.* 2023;61(9-10):4643-4659.

 doi: 10.1007/s00382-023-06788-6

19. Kang IS, Shukla J. Dynamic seasonal prediction and predictability of the monsoon. In: Wang B, ed. *The Asian Monsoon*. Berlin, Germany: Springer; 2006:585-612.

 doi: 10.1007/3-540-37722-0_15

20. Cleveland WS, Devlin SJ. Locally weighted regression: an approach to regression analysis by local fitting. *J Am Stat Assoc.* 1988;83(403):596-610.

 doi: 10.1080/01621459.1988.10478639

21. Lyu Y, Zhi X, Zhu S, *et al*. Statistical calibrations of surface air temperature forecasts over East Asia using pattern projection methods. *Weather Forecast.* 2021;36(5):1661-1674.

 doi: 10.1175/WAF-D-21-0043.1

22. Zhu S, Zhang L, Jiang H, *et al*. Pattern projection calibrations on subseasonal forecasts of surface air temperature over East Asia. *Weather Forecast.* 2023;38(6):865-878.

 doi: 10.1175/WAF-D-22-0046.1

23. Huang Z, Zhao T, Xu W, *et al*. A seven-parameter Bernoulli-Gamma-Gaussian model to calibrate subseasonal to seasonal precipitation forecasts. *J Hydrol.* 2022;610(1):127896.

 doi: 10.1016/j.jhydrol.2022.127896

24. Scher S. Toward data-driven weather and climate forecasting: approximating a simple general circulation model with deep learning. *Geophys Res Lett.* 2018;45(22):12616-12622.

 doi: 10.1029/2018GL080704

25. Rasp S, Lerch S. Neural networks for postprocessing ensemble weather forecasts. *Mon Weather Rev.* 2018;146(11):3885-3900.

 doi: 10.1175/MWR-D-18-0187.1

26. Bonavita M, Laloyaux P. Machine learning for model error inference and correction. *J Adv Model Earth Syst.* 2020;12(12):e2020MS002232.

 doi: 10.1029/2020MS002232

27. Kim H, Ham YG, Joo YS, Son SW. Deep learning for bias correction of MJO prediction. *Nat Commun.* 2021;12(1):3087.

 doi: 10.1038/s41467-021-23406-3

28. Jin W, Zhang W, Hu J, *et al*. Using the residual network module to correct the sub-seasonal high temperature forecast. *Front Earth Sci.* 2022;9:760766.

 doi: 10.3389/feart.2021.760766

29. Ling F, Li Y, Luo JJ, *et al*. Two deep learning-based bias-correction pathways improve summer precipitation prediction over China. *Environ Res Lett.* 2022;17(12):124025.

 doi: 10.1088/1748-9326/aca68a

30. Xie J, Hsu PC, Hu Y, *et al*. Skillful extended-range forecast of rainfall and extreme events in East China based on deep learning. *Weather Forecast.* 2023;38(3):467-486.

 doi: 10.1175/WAF-D-22-0132.1

31. Lyu Y, Zhu SP, Zhi XF, *et al*. Improving subseasonal-to-seasonal prediction of summer extreme precipitation over southern China based on a deep learning method. *Geophys Res Lett.* 2023;50(24):e2023GL106245.

 doi: 10.1029/2023GL106245

32. Weyn JA, Durran DR, Caruana R, Cresswell-Clay N. Sub-seasonal forecasting with a large ensemble of deep-learning weather prediction models. *J Adv Model Earth Syst.* 2021;13(7):e2021MS002502.

 doi: 10.1029/2021MS002502

33. Schultz MG, Betancourt C, Gong B, *et al*. Can deep learning beat numerical weather prediction? *Philos Trans A Math Phys Eng Sci*. 2021;379(2194):20200097.

 doi: 10.1098/rsta.2020.0097

34. Kashinath K, Mustafa M, Albert A, *et al*. Physics-informed machine learning: case studies for weather and climate modelling. *Philos Trans A Math Phys Eng Sci.* 2021;379(2194):20200093.

 doi: 10.1098/rsta.2020.0093

35. Mouatadid S, Orenstein P, Flaspohler G, *et al*. Adaptive bias correction for improved subseasonal forecasting. *Nat Commun.* 2023;14(1):3482.

 doi: 10.1038/s41467-023-38874-y

36. Risbey JS, Squire DT, Black AS, *et al*. Standard assessments of climate forecast skill can be misleading. *Nat Commun.* 2021;12(1):4346.

 doi: 10.1038/s41467-021-23771-z

37. Fan Y, van den Dool H. A global monthly land surface air temperature analysis for 1948-present. *J Geophys Res Atmos.* 2008;113(D1):D01103.

 doi: 10.1029/2007JD008470

38. Kirtman BP, Pegion K, DelSole T, *et al*. The Subseasonal Experiment (SubX) dataset. IRI Data Library. Published 2017.

 doi: 10.7916/D8PG249H

39. European Centre for Medium-Range Weather Forecasts. Operational archive (ENS extended): re-forecast products. Available from: https://www.ecmwf.int/en/forecasts/dataset/operational-archive [Last accessed on March 3, 2026].

40. Raff D, Nowak K, Cifelli R, Brekke LD, Webb RS. Sub-seasonal climate forecast rodeo. Presented at: AGU Fall Meeting; December 11-15, 2017; New Orleans, LA. Available from: https://ui.adsabs.harvard.edu/abs/2018AGUFM.A41L3155R/abstract [Last accessed on March 3, 2026].

41. Cheng Y, Wang L, Zhou X, Wang L. Quasi-biweekly and quasi-monthly intraseasonal oscillations in North American winter atmospheric circulation. *Clim Dyn.* 2025;63(12):457.

doi: 10.1007/s00382-025-07942-y