Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Impacts of Multi–Observation Interests on Rainstorm Forecasts in East China: An OSSE–Based Study

Yakai Guo

guoykhmb@126.com

Henan Meteorological Bureau, China Meteorological Administration

Changliang Shao

China Meteorological Administration

Guanjun Niu

China Meteorological Administration

Yuting Dong

Alfred Wegener Institute

Article

Keywords: Observation impact, forecast sensitivity, Observation System Simulation Experiments, rainstorm, data assimilation

Posted Date: July 16th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-7091557/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Impacts of Multi–Observation Interests on Rainstorm Forecasts in East China: An OSSE–Based Study

Yakai Guo [1, 5, *], Changliang Shao [2, 5, +], Guanjun Niu [3, 5], and Yuting Dong [4, 5]

[1]Henan Meteorological Bureau, China Meteorological Administration, Zhengzhou 450003, China
[2]Meteorological Observation Centre, China Meteorological Administration, Beijing 100081, China
[3]Meteorological Development and Planning Institute, China Meteorological Administration, Beijing 100081, China
[4]Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven 27570, Germany
[5]Observation System Evaluation Innovation Group, China Meteorological Administration, Beijing 100081, China
*Corresponding: guoykhmb@126.com, shaocl@cma.gov.cn

Abstract:

Existing Observation System Simulation Experiment (OSSE) studies have limitations in evaluating regional rainstorms in East China due to the lack of known atmospheric truth and varied interests. This study introduces a comprehensive framework that combines multiple forecast metrics (synoptic, objective, energy, economic) with various observation interests (element, region, density, layer) using identical twin designs. The framework assesses their impacts on a central East China rainstorm event (19 July 2016) through flexible assimilation-forecast systems. Key findings include: (1) significant disparities in mutual sensitivity between metrics and observation interests, highlighting the limitations of single-interest approaches; (2) coordinated optimization strategies that identify specific observation interests (e.g., surface temperature, mid-east region, southern density, 500 hPa layers) that enhance forecast accuracy, while others (e.g., dew element, central-west region, northwestern density, 925 hPa layers) require cautious reduction; (3) universal improvements in forecasts through assimilation across different reference truths, but with notable benefit-cost paradoxes that underscore the need for enhanced validation of synoptic metrics and reference authenticity. These results provide actionable insights for improving the harmonization of observation and forecast systems in regional numerical models, ultimately enhancing extreme weather prediction and disaster mitigation.

Key words: Observation impact, forecast sensitivity, Observation System Simulation Experiments, rainstorm, data assimilation

1. Introduction

The frequency and intensity of regional extreme weather events have been on the rise under the backdrop of climate change [43,9], posing significant challenges to weather forecasting and disaster prevention. East China, characterized by complex synoptic systems and limited observational data, is particularly vulnerable to these extreme events. Therefore, optimizing regional observations to enhance the accuracy of rainstorm forecasts has become an urgent priority [48,38].

Observation System Simulation Experiments (OSSEs) have long been recognized as a valuable tool for evaluating the impact of planned observations on numerical weather prediction (NWP). OSSEs measure the potential impact of new or modified observation systems on NWP forecasts by simulating the assimilation of hypothetical observations into a numerical model [3,16,41]. However, existing OSSE studies often lack a comprehensive evaluation of multiple forecast metrics for regional rainstorms in East China, primarily due to the unknown atmospheric truth and the diverse interests in both forecast and observation. This gap in research hinders the effective optimization of regional observation networks and the improvement of local weather forecasting capabilities. Early OSSE studies primarily focused on global-scale applications, utilizing identical twin designs based on data denial to assess the impact of existing observations on historical forecasts [6,5]. These studies, known as Observation System Experiments (OSEs), have been widely applied in global NWP centers with customized configurations tailored to different forecast interests [36,11,23]. Subsequently, to address the challenges of new observation plans and their NWP applications, OSSEs were developed by incorporating simulated observations derived from a tailored atmosphere and their corresponding assimilation schemes [49,24]. These advancements have enabled the assessment of simulated observations' impact on historical NWP forecasts and provided guidance for observation design and application.

Despite these advancements, traditional OSSEs face limitations when applied to regional studies due to the high cost of obtaining reference truth and the substantial differences in interests between forecast and observation. For instance, forecast verification often focuses on specific weather phenomena, while observation characteristics may vary widely [3,16,41]. To overcome these challenges, recent studies have explored the use of flexible verification metrics

and varied observation interests within the OSSE framework [1,25,51,52,8,22,27]. These efforts aim to provide a more comprehensive evaluation of observation impacts on regional weather forecasts. However, the validity of these approaches is often limited by the short lead forecast times of their reference datasets and the inherent uncertainties in model and observation parameters [32,29,10,28,34,30].

In this study, we focus on a typical warm-sector rainstorm event that occurred in central East China on 19 July 2016. This event has been well-documented in previous studies, highlighting the large-scale predictability and significant regional uncertainties [13]. Recent advancements in regional observation intensification and assimilation schemes, such as three-dimensional variation (3DVar) and ensemble root mean square filter (EnSRF), have significantly improved the monitoring and forecasting of such events [45]. However, the complex characteristics of these intensified observations, including spatiotemporal discontinuities and large terrain gaps, have exacerbated the disharmony between regional NWP's observation and forecast systems [26,47]. Consequently, the impacts of multiple forecast and observation interests on regional rainstorm forecasts in central East China remain largely unexplored.

To address these gaps, this study introduces a comprehensive framework that combines multiple forecast metrics and observation interests. The framework employs identical twin designs to assess the mutual sensitivities between forecast metrics and observation interests, the relative impacts of different observation interests, and the effects of reference truth on forecast accuracy. By leveraging flexible assimilation–forecast systems, this study aims to provide actionable insights for optimizing local observations and improving regional NWP models, ultimately enhancing local weather forecasting and disaster mitigation capabilities.

The initial section sets the stage by introducing the problem of interest, namely the impacts of multi–observation interests on rainstorm forecasts in East China. Section 2 details the methodologies employed in this study. Section 3 presents the experimental setup, including the specifics of the rainstorm event selected for study, the design of the assimilation experiments, and the reference datasets used for validation. Section 4 delves into the results of the study, discussing the impacts of data assimilation on forecast accuracy, the mutual sensitivities between forecast metrics and observation interests, and the relative impacts of different

observation interests on forecast performance. Section 5 provides a discussion of the findings, interpreting the results in the context of existing literature and theory, and exploring the implications for the design of future observation networks and data assimilation strategies. The paper concludes in Section 6 with a summary of the key findings and their significance for improving weather forecasting and disaster prevention efforts in East China. It also suggests directions for future research to build upon the insights gained from this study. Finally, the Appendices contain detailed descriptions of the calculation formulas for the various forecast metrics used in the study, providing transparency and reproducibility of the research methods.

2. Method

To comprehensively evaluate the impacts of various observation interests on regional rainstorm forecasts in East China, this study employs a combination of multiple forecast metrics and observation interests within an identical twin OSSE framework. The methodology involves the following key components: (1) Forecast Metrics; (2) Observation Interests; (3) Mutual Sensitivities; (4) Relative Impacts; (5) Reference Effects.

2.1 Forecast Metrics

This study utilizes a diverse set of forecast metrics to evaluate the impacts of different observation interests on the rainstorm forecasts. These metrics are categorized into four types: synoptic, objective, energy, and economic. The specific metrics and their definitions are summarized in **Table 1**. The distance between experimental data and reference truth (RT) in the horizontal dimensions is defined as dis, while the distance between control data and RT is defined as dis_c. A smaller dis than dis_c indicates a benefit, or a cost otherwise. The detailed formulas for calculating these distances are provided in the appendix.

Table 1 Description of the forecast metrics used in this study

Type	Metric	Formula*	Distance*
Synoptic	*THE*	$dis^t - dis_c^t$	$dis = abs(\Delta LCL)$
	VOR	$\sum_{t=1}^{t=nt}(dis^t - dis_c^t)$	$dis = VCD(\Delta lat, \Delta lon)$
	SUR	$\sum_{t=1}^{t=nt}(dis^t - dis_c^t)$	$for\ vector\ (v_1, v_2)$:
Objective	*UPP*	$\sum_{t=1}^{t=nt}\sum_{k=1}^{k=nk}(dis^{k,t} - dis_c^{k,t})$	$dis = RMSVE(\Delta v_1, \Delta v_2)$
	PRO	$\sum_{t=1}^{t=nt}(dis^{k,t} - dis_c^{k,t})$	$else$: $dis = RMSE(\Delta u)$
Energy	*MSE*	$\sum_{t=1}^{t=nt}\sum_{k=1}^{k=nk}(dis^{k,t} - dis_c^{k,t})$	$dis = RMSE(\Delta MSE)$

	TKE	$\sum_{t=1}^{t=nt}\sum_{k=1}^{k=nk}(dis^{k,t}-dis_c^{k,t})$	$dis = RMSE(\Delta TKE)$
Economic	*ECO*	$dis - dis_c$	$dis = 1 - ECLV$

*The superscripts t and nt represent t^{th} step and total step number respectively in the study time period. The superscripts k and nk represent k^{th} layer and total layer number respectively. The symbol Δ represents the difference between the experiment data and RT. (lat, lon) represents the latitude and longitude of the vortex center. (v_1, v_2) and u represent the vector and unique variables respectively. The details about distance calculation should refer to the appendix section.

2.1.1 Synoptic

Thermodynamic (*THE*): Based on the T–lnp framework, this metric quantifies the benefit/cost of the lifting condensation level (*LCL*) distance at Zhengzhou (ZZ) station at 12:00 prior to the core rainstorm onset [12].

Vortex (*VOR*): This metric quantifies the benefit/cost of the path distance of the low-level vortex center during the rainstorm event [14,39,50].

2.1.2 Objective

Surface Element (*SUR*): This metric quantifies the benefit/cost of errors in different near-surface atmospheric states (T2, T2-D2, and W10) over the entire weather period.

Upper Element (*UPP*): This metric corresponds to different atmospheric states (T, RH, and W) for the entire layer.

Vertical Profile (*PRO*): This metric corresponds to different atmospheric states (T, RH, and W) for stratified layers (i.e., vertical profiles).

2.1.3 Energy

Moist Static Energy (*MSE*): This metric quantifies the benefit/cost of errors in the moist static energy of the lower atmosphere (1000–700 hPa) over the entire weather period [31].

Turbulent Kinetic Energy (*TKE*): This metric quantifies the benefit/cost of errors in the turbulent kinetic energy of the lower atmosphere [46].

2.1.4 Economic

Economic Cost-Loss Value (*ECO*): This metric quantifies the benefit/cost of the ECLV distance for forecasted disaster-causing precipitation (i.e., rainstorm and heavy rainstorm) during the study period [35,44].

2.2 Observation Interests

The observation interests are categorized into four types: Element, Region, Density, and Layer. These interests are designed to reflect the practical needs for observing typical weather processes and to evaluate the influence of actual observational tasks on weather forecasting.

The geographical distribution of these observation interests is illustrated in **Figure 1**, which provides an overview of the main concerned interests of our study region, including the early observing stations and the four defined regions at 06:00 UTC on 19 July 2016. The figure also shows the subsequent rainstorm event (total rainfall exceeding 50 mm: contoured; time-averaged winds of 700 hPa: vector) covering the period from 06:00 to 18:00, along with the elevation data from the USGS' 3-second SRTM (shaded; units: m). The early reported observations include automatic weather stations (AWS; blue scattered dots), sounding (red circles), Doppler radars (red boxes), Global Navigation Satellite System/Meteorology water vapor (GNSS/MET; red crosses), wind profile radar (WP; red triangles), and microwave radiometer (MR; red inverted triangles). The main rivers and province boundaries are indicated by ticked cyan and gray lines, respectively.

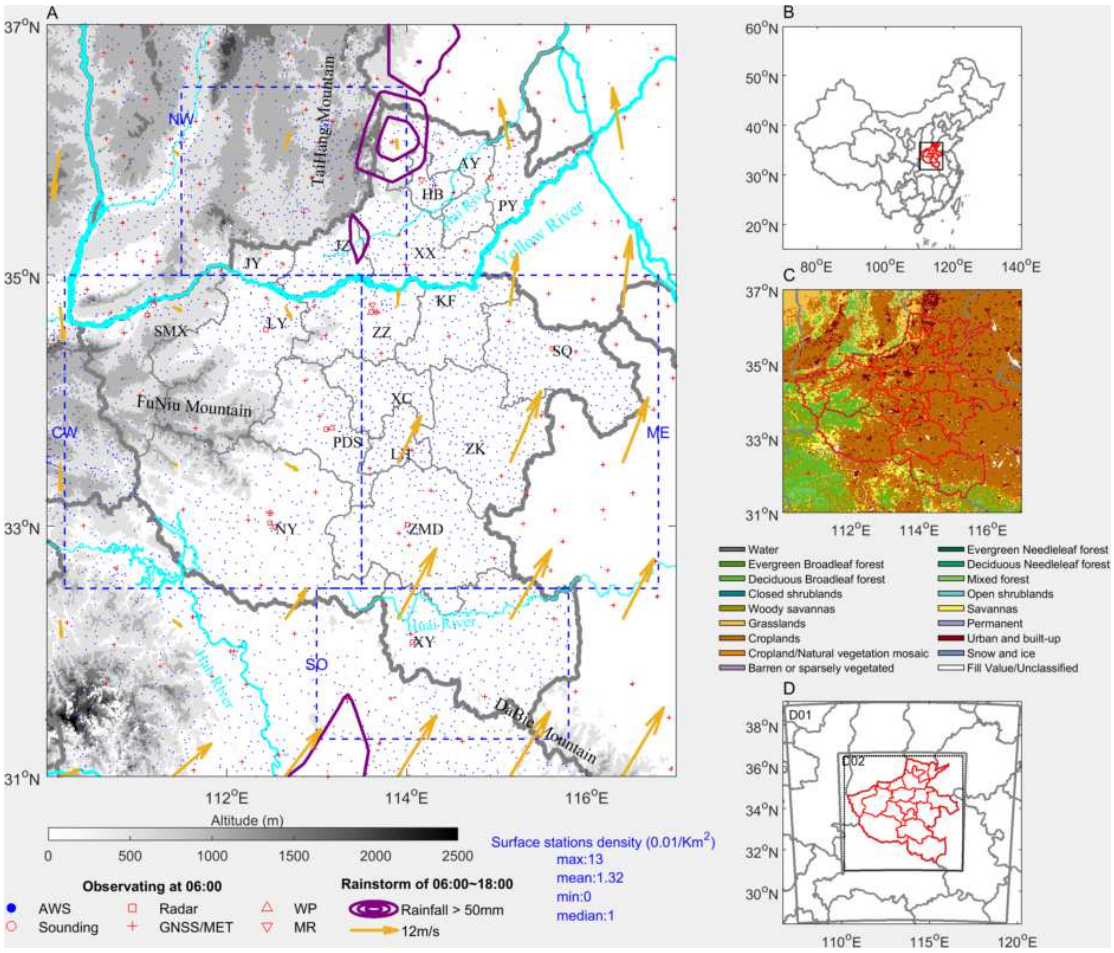


Figure 1 A. Overview of main concerned interests of our study region such as the early observing (scattered) and four regions (blue dashed boxes) at UTC 06:00 19 July 2016, the following rainstorm (total rainfall exceeding 50 mm: contoured; time averaged winds of 700hPa: vector) covering 06:00~18:00, and the elevation (shaded; units: m) of USGs' 3s SRTM. The early reported observations includes the automatic weather stations (AWS; blue scattered dots), sounding (read circles), Doppler radars (red boxes), Global Navigation Satellite System/Meteorology water vapor (GNSS/MET; red crosses), wind profile radar (WP; red triangles), microwave radiometer (MR; red inverted triangles), and the main rivers and province boundaries are in ticked cyan and ticked gray lines respectively. B. Location (black box) of Henan province (red lines) in China. C. The MODIS land cover (shaded). D. Model domains (solid lines) and our study region (dotted lines). During A, ZZ = Zhengzhou, KF = Kaifeng, LY = Luoyang, PDS = Pingdingshan, AY = Anyang, HB = Hebi, XX = Xinxiang, JZ = JiaoZuo, PY = Puyang, XC = Xuchang, LH = Luohe, SMX = Sanmenxia, NY = Nanyang, SQ = Shangqiu, XY = Xinyang, ZK = Zhoukou, ZMD = Zhumadian, JY = Jiyuan.

2.2.1 Element

Temperature (T2): Surface temperature observations.

Dew Point (D2): Surface dew point observations.

Wind Vector (W10): Surface wind vector observations.

2.2.2 Region

Northwest (NW): Covering the southern part of the TaiHang mountains (111.5°E to 114.0°E, 35.0°N to 36.5°N).

Central-West (CW): Covering the FuNiu mountain (110.2°E to 113.5°E, 32.5°N to 35.0°N).

South (SO): Covering the northern side of the DaBie mountains (111.3°E to 115.8°E, 31.3°N to 32.5°N).

Middle-East (ME): Covering the Yellow-Huai grain-producing area (113.5°E to 116.8°E, 32.5°N to 35.0°N).

2.2.3 Density

6 km: High-density observations.

12 km: Medium-density observations.

30 km: Low-density observations.

2.2.4 Layer

200 hPa: Upper atmospheric layer.

500 hPa: Mid-atmospheric layer.

700 hPa: Lower mid-atmospheric layer.

850 hPa: Lower atmospheric layer.

925 hPa: Near-surface atmospheric layer.

2.3 Mutual Sensitivities

To identify and analyze the sensitivity characteristics of forecast metrics (*M*) to observation interests (OIG), we define the correlation of benefit rankings (i.e., Spearman's rank correlation coefficient, CC) between different sub-metrics of the same *M* type and different *M* types as self-CC and non-self-CC, respectively. Larger values of self-CC and non-self-CC indicate closer benefit rankings between sub-metrics of the same *M* type and between different *M* types within a certain OIG type, suggesting insensitivity of *M* to that OIG type. The significance of these correlations is further analyzed using the distribution of significance levels (p-values) to reveal the overall confidence characteristics of the sensitivity of *M* type to OIG.

Similarly, the sensitivity of a given OIG type to *M* is analyzed by swapping *M* and OIG in the above analysis. This provides qualitative insights into the overall sensitivity of *M* or OIG types, offering objective suggestions for their practical applications.

2.4 Relative Impacts

To objectively evaluate the primary influence of homogeneous OIG types on the entire *M* category, we introduce linear transformations across all *M* within the same OIG type as $\widetilde{M_i} = \frac{M_i}{max(abs(\boldsymbol{M}))}$, where i represents the i^{th} metric, ne denotes the length of *M*, and $\widetilde{M}$ varies in [−1, 1]. Noted that, due to the unsolved scaling of extra upper layer dimensionalities, except *PRO*, other metrics are all scaled during this work (Table 1). Moreover, the significance of OIG's benefit/cost associated with the overall *M* are further analyzed by examining the

extreme values of M and their corresponding OIG pairs, i.e., C_{max}/ B_{max} represents the actual extreme benefit/cost corresponding to *M*. This analysis helps identify the beneficial generality and cautious extremes of each OIG type.

2.5 Reference Effects

Given the inherent selectability of RT in M calculations, this study examines the impact of RT selection on the sensitivity of OIG and/or *M* by comparing self-CC across varying RT. Additionally, the distribution differences of dis under different RT scenarios are analyzed to clarify how RT choices affect the relative impacts of OIG. This comprehensive methodology allows for a detailed assessment of the impacts of various observation interests on regional rainstorm forecasts, providing valuable insights for optimizing local observation networks and improving regional NWP models.

3. Model and Data

3.1 Model

3.1.1 WRF Model

The WRF model (version 3.9.1) is employed in this study to simulate the regional rainstorm event. The model uses a two-way nested scheme with one-way feedback, meaning that only the feedback from the outer domain to the inner domain is considered. The model domains are centered at 113.45°E, 33.85°N, with the outer domain consisting of 100×100 grids at a 12 km resolution and the inner domain consisting of 166×160 grids at a 4 km resolution. The model integration time step is set to 30 seconds. The model includes 51 vertical layers, with 21 layers below 1 km.

For the outer domain, the Kain-Fritsch (KF) convection scheme [21] is used, while convection is turned off in the inner domain. Other physical processes are modeled using the same schemes for both domains: the Thompson scheme for cloud microphysics [42], the Yonsei University (YSU) scheme for boundary layer physics [17], the Rapid Radiative Transfer Model (RRTMG) scheme for both long-wave and short-wave radiation physics [17], the Monin-Obukhov similarity scheme for surface physics [20], the Noah Land Surface Model (LSM) scheme for land surface physics [40], and the Unified Canopy Model (UCM) scheme for canopy physics [7].

3.1.2 Assimilation

The processes of the two data assimilation–forecast schemes adopted in this study are illustrated in **Figure 2**. The cold start firstly uses forcing datasets to drive WRF to get boundary and initial values (x^0). For 3DVar [4], a following 6 hour (6h) warm up is conducted to get the initial first guess (x^f), then the static background covariance using the global–scale forecast error variance (or CV3) [33], is assembled to assimilate the surface simulated observations. However, for EnSRF [29], the following 6h warm ups that initialized with a ensemble of 30 members, that adding x^0 with random perturbation of the standard deviation (δ^f) of various model sates, i.e., temperature (δ_T^f, 1.5 K), geopotential altitude (δ_{GH}^f, 10 gpm), zonal winds (δ_U^f, 2 m/s), and meridional winds (δ_V^f, 2 m/s), to achieve the initial ensemble forecasts (x_i^f), then the real–time ensemble background covariance using a fixed–coefficient inflation scheme (i.e., $(1+\alpha)$) [2] with α being 0.5 and a distance–weighted piece–wise localization [15] with the horizontal and vertical radii being 20 and 10 Km respectively, is conducted to assimilate the upper simulated observations.

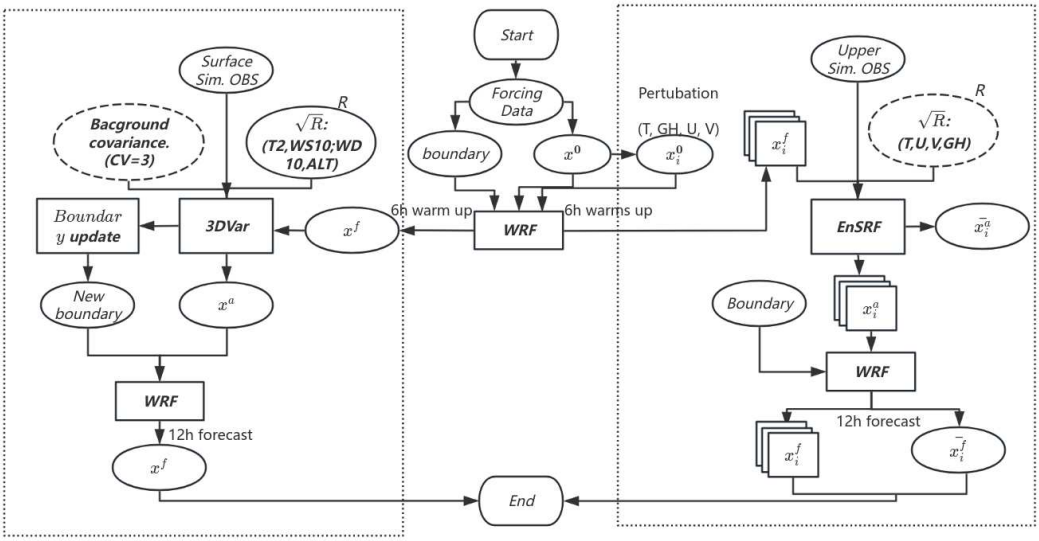


Figure 2 The flowchart of local WRF–3DVar and –EnSRF systems.

In addition, 3DVar adopts the observation error ($\sqrt{R}$) with a standard deviation (δ^o) of various SYNOP assimilation entries, i.e., 10m wind speed (δ_{WS10}^o, 1.1 m/s), 10m wind direction (δ_{WD10}^o, 1.1°), 2m temperature (δ_{T2}^o, 2 K) and observation screen height (δ_{ALT}^o, 6.0 m), while EnSRF adopts $\sqrt{R}$ with a standard deviation (δ^o, who equals to δ^f) [19] of various SOUNDING assimilation entries, i.e., temperature (δ_T^o, 1.5 K), geopotential altitude (δ_{GH}^o, 10 gpm), zonal winds (δ_U^o, 2 m/s), and meridional winds (δ_V^o, 2 m/s). Especially, the thresholds of the various

corresponding observation innovation (i.e., observation minus background or OMB) for both 3DVar and EnSRF all equal to $5\delta^o$.

3.2 Data

3.2.1 Simulated Observation

Simulated observations are derived from the European Centre for Medium-Range Weather Forecasts (ECMWF) real-time operational forecast datasets, which are composited with layer-different datasets of varied horizontal resolutions. The upper-level ECMWF data have a resolution of around 25 km, while the surface-level data have a resolution of around 12 km. The real surface Automatic Weather Stations (AWS) resolution can approach nearly 10 km, indicating that their surface-level resolutions are fortuitously close. The ECMWF data are bi-linearly interpolated to resolutions of 12 km and 6 km to obtain the simulated observations at upper-level and surface-level, respectively, named as EC12 and EC6.

For the simulated surface-level observation (SURF), the screen heights for temperature and dew point are 2 meters above the ground level (T2 and D2, respectively), while the screen height for wind is 10 meters above the ground level (W10). These conform to the EC6 surface-level layers. For the simulated vertical-level observation (HIGH), the vertical layers for suited elements, including temperature, winds, and altitude (T, W, and H), conform to the EC12 pressure layers. SURF and HIGH jointly form the tailored observations of this study.

3.2.2 Forcing Datasets

The final reanalysis (FNL) dataset with a 0.25° resolution, derived from the National Centers for Environmental Prediction (NCEP), is used to drive the WRF model. The FNL dataset contains 27 pressure layers, 1 surface layer, and 4 soil layers. The USGS' SRTM DEM with 3-second resolution and the NASA's MODIS land cover datasets with 5-second resolution are used as the WRF's underlying terrain and land cover, respectively (**Figure 1**A and C).

3.2.3 Validation Dimensions

In addition to using the ECMWF data as the RT, this study also employs the full-observation assimilation of each OIG as the tailored reference to investigate the effects of RT differences. Given that the spatiotemporal resolution of ECMWF data is significantly lower than that of the WRF–model–based tailored reference data, the ECMWF data's spatiotemporal

resolution (i.e., 3–hour interval, horizontal spacing of 6/12 km, and 10 pressure levels) is adopted as the baseline dimension for verification.

4. Experiment

4.1 Experimental Design

The rainstorm event that occurred around the local afternoon (06:00 UTC) on July 19, 2016, and lasted for 12 hours, crossing the ZZ–XX–AY area northeasterly, is used as the case study. The WRF model is initiated with FNL at 00:00 UTC. After a 6–hour warm–up, assimilation with four observation interest groups (OIG), designed as allow and deny twins, is performed to achieve varied initial conditions at 06:00 UTC. These initial conditions drive the WRF model to produce 12–hour forecasts.

The experiment design is detailed in **Table 2**. Except for Density, all other OIGs are conducted as allow and deny types with opposite parameter behaviors to quantify the impacts of individual observation interests. Density is conducted as density enhancement with different spatial parameters of the ongoing mentioned four interest regions to qualify the impacts of regional density increase or decrease. The experiment without assimilation is taken as the control run (CTR), while the varied tailored assimilation runs and the EC6/EC12 datasets are taken as the two types of desirable reference truth, i.e., tailored and ECMWF.

Table 2 Experiment description

OIG	Type	Parameter Design*	Notes
Element	Allow	EA1, EA2, and EA3: Single assimilation of 2m temperature (T2), 2m dew (D2), and 10m winds (W10) respectively compared to CTR.	Quantify impacts of individual surface element of SURF.
	Deny	ED1, ED2, and ED3: Denial assimilation of T2, D2, and W10 respectively compared to SR	Quantify impacts of excluding individual surface element from SURF.
Region	Allow	RA1, RA2, RA3, and RA4: Single assimilation of the data in NW (111.5–114°E, 35–36.5°N), CW (110.2–113.5°E, 32.5–35°N), SO (111.3–115.8°E, 31.3–32.5°N), and ME (113.5–116.8°E, 32.5–35°N) regions respectively compared to CTR.	Quantify impacts of individual specific region of SURF.

Region	Deny	RD1, RD2, RD3, and RD4: Denial assimilation of the data in NW, CW, SO, and ME regions respectively compared to SR.	Quantify impacts of excluding individual specific region from SURF.
Density	Enhancement	RA1, RA1$_{.5}$, and RA1$_{.2}$: Single assimilation of the data with density of 6km (100%), 12km (50%), 30km (20%) respectively in NW region. (RA2, RA2$_{.5}$, RA2$_{.2}$), (RA3, RA3$_{.5}$, RA3$_{.2}$), and (RA4, RA4$_{.5}$, RA4$_{.2}$) are the same to (RA1, RA1$_{.5}$, RA1$_{.2}$) but for those in CW, SO, and ME regions respectively.	Quantify impacts of density increasing or decreasing of individual specific region of SURF.
Layer	Allow	LA1, LA2, LA3, LA4, and LA5: Single assimilation of the data of 200hPa, 500hPa, 700hPa, 850hPa, and 925hPa respectively compared to CTL.	Quantify impacts of individual layer of HIGH.
	Deny	LD1, LD2, LD3, LD4, and LD5: Denial assimilation of the data of 200hPa, 500hPa, 700hPa, 850hPa, and 925hPa respectively compared to UR.	Quantify impacts of excluding individual layer from HIGH.

*SR = Surface Reference, and UR = Upper Reference.

The flowchart of this study can be divided into three parts (Fig. 3). Firstly, data preparation primarily involves the observation extraction of various twin pairs (such as allowed and denied) within four interest groups (i.e., Element, Region, Density, and Layer), followed by data assimilation to obtain forecasts incorporating reference observations and forecasts incorporating observation interests. Secondly, benefit/cost calculation mainly consists of extracting surface and upper–level state data (e.g., temperature, humidity, and wind) from various forecasts, and calculating the benefits/costs of different observation–interest informed forecasts using various evaluation metrics against different reference (i.e., EC and tailored data) and control (CTL) forecasts. Lastly, benefit/cost evaluation primarily entails ranking and/or scaling the benefits/costs under various metrics and references, as well as conducting quantitative analysis of their differences and/or characteristics, thereby deriving the coordinated observation and forecast interests.

4.2 Assimilation and Forecast

The data assimilation process involves the following steps:

Data Preparation: Extract observations from various twin pairs (allow and deny) within the four interest groups (Element, Region, Density, and Layer).

Data Assimilation: Use the 3DVar and EnSRF schemes to assimilate the observations and generate initial conditions for the WRF model.

Forecast Generation: Drive the WRF model with the assimilated initial conditions to produce 12-hour forecasts.

The flowchart of this study is shown in **Figure 3**, which outlines the data preparation, benefit/cost calculation, and benefit/cost evaluation processes. The data preparation involves extracting observations from various twin pairs within the four interest groups. The benefit/cost calculation involves extracting state data from various forecasts and calculating the benefits/costs using the defined metrics against different references (EC and tailored data) and control forecasts. The benefit/cost evaluation involves ranking and scaling the benefits/costs under various metrics and references, and conducting quantitative analysis of their differences and characteristics.

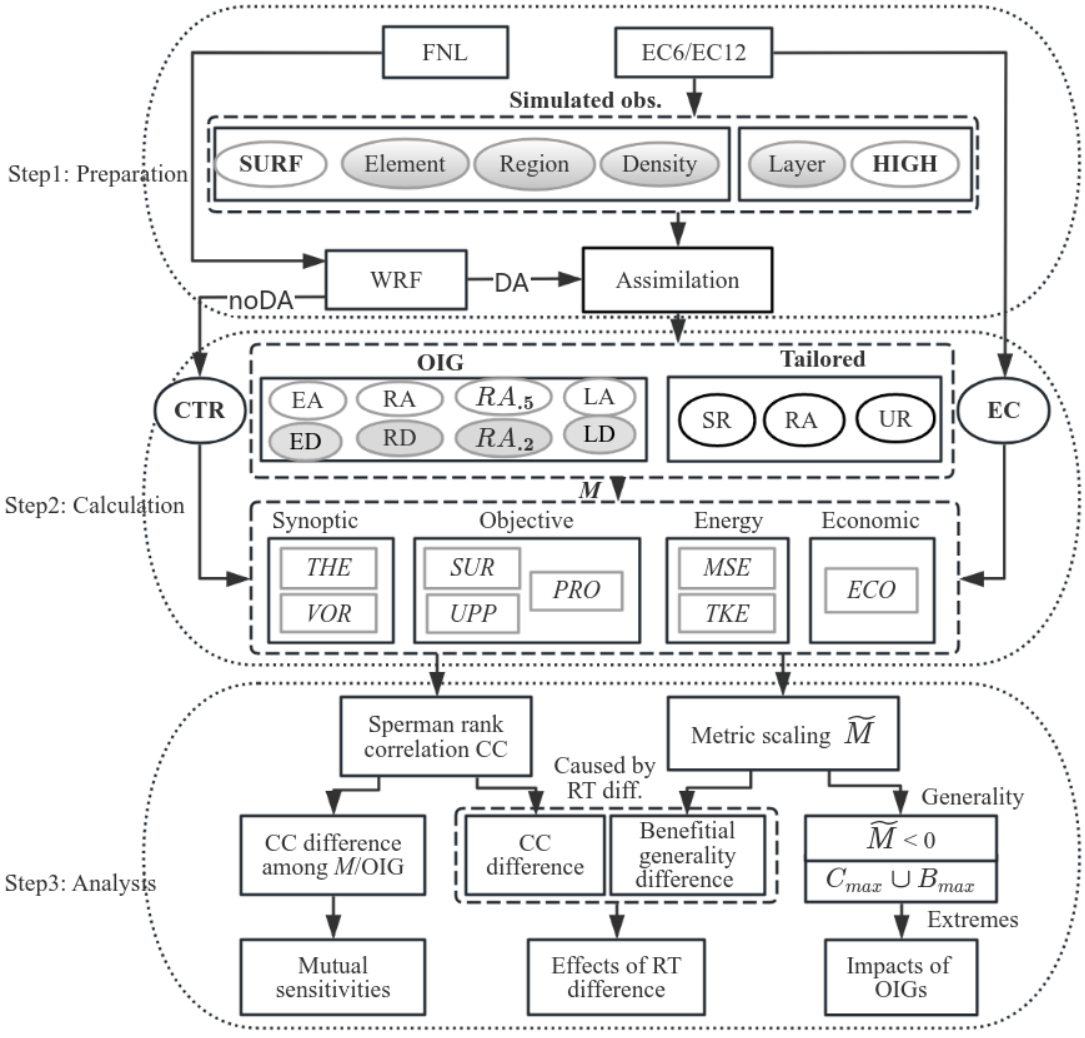


Figure 3 Flowchart of this study. RT = Reference Truth.

4.3 Validation and Analysis

The validation and analysis of the forecasts involve the following steps:

Comparison with Reference Truth: Compare the forecasts generated from different assimilation schemes with the reference truth (EC and tailored data) to assess the accuracy of the forecasts.

Benefit/Cost Analysis: Calculate the benefits/costs of different observation–interest–informed forecasts using the defined metrics.

Sensitivity Analysis: Analyze the sensitivity of forecast metrics to different observation interests, the sensitivity of observation interests to different forecast metrics, and the impact of reference truth selection on the results.

The results of these analyses provide insights into the mutual sensitivities between forecast metrics and observation interests, the relative impacts of different observation interests, and the effects of reference truth on forecast accuracy. These insights are crucial for optimizing local observation networks and improving regional NWP models.

5. Result

5.1 Effects of Data Assimilation

The assimilation of different observation interests significantly impacts the initial conditions and subsequent forecasts. For 3DVar assimilation, the surface reference (SR) absorbed the most T2, U10, and V10 observations, while the region-specific assimilation (RA) absorbed the least. For EnSRF assimilation, the upper reference (UR) absorbed the most T, U, and V observations, while the layer-specific assimilation (LA) absorbed the least. The observation minus background (OMB) quantities showed consistent U-shaped distributions for surface observations and approximately linear distributions for upper-level observations.

Figure 4A illustrates the OMB quantity in assimilation of each OIG, showing the number of observations assimilated for each variable. The results indicate that the surface reference (SR) absorbed the most observations, while the region-specific assimilation (RA) absorbed the least.

Figure 4B shows the T–lnP diagram of THE at ZZ station at 12:00 on July 19, 2016. The temperature profiles among various datasets are close, but significant differences are observed in the dew profiles, with the ECMWF data indicating a notably dry upper atmosphere compared to other datasets.

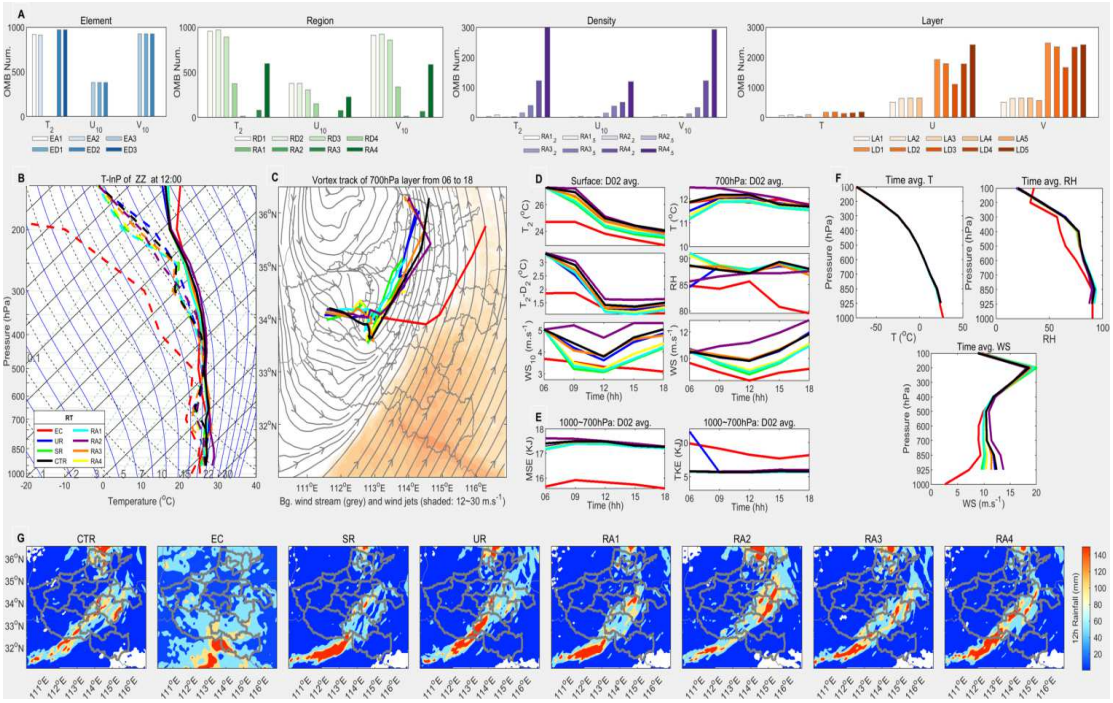


Figure 4 A. The observation minus forecast (OMB) quantity in assimilation of each OIG, x– and y–axis represent variable and OMB quantity respectively. B. The T–lnP diagram of *THE* at ZZ station at 12:00 on July 19, 2016, where the colored thick solid and dashed lines represent the temperature and dew of different reference truths, while the fine solid black line, fine solid blue line, fine dashed black line, and fine dashed green line represent isotherms, moist adiabats, dry adiabats, and saturation vapor mixing ratio, respectively. C. The 700hPa vortex tracking from 06:00 to 18:00 on July 19, 2016. D. The variation of domain (D02) averaged states of the surface layer and the 700hPa layer against time. E is the same to D, but for the energy states of 1000–700hPa layer. F. The vertical profiles of spatiotemporally averaged upper states. G. The 12 hour total rainfall at 18:00 on July 19, 2016.

Figure 4C depicts the 700hPa vortex tracking from 06:00 to 18:00 on July 19, 2016. The vortex paths show significant differences due to large variations in vortex development among the OSSE experiments. The ECMWF data show a faster eastward acceleration of the vortex compared to the tailored reference truth (RT).

Figure 4D shows the variation of domain-averaged states of the surface layer and the 700hPa layer against time. The results indicate that the tailored RT and control (CTL) forecasts are very close, while the ECMWF data show a relatively cold and saturated surface with weak winds.

Figure 4E shows the variation of domain-averaged energy states of the 1000-700hPa layer against time. The ECMWF data show a lower moist static energy (MSE) but a higher turbulent kinetic energy (TKE) compared to the tailored RT and CTL forecasts.

Figure 4F shows the vertical profiles of spatiotemporally averaged upper states. The results indicate that the ECMWF data show a relative drier troposphere but a sharper low-level vertical wind shear compared to the tailored RT and CTL forecasts.

Figure 4G shows the 12-hour total rainfall at 18:00 on July 19, 2016. The results indicate consistent southwest-northeastern oriented belts but with significant regional differences. The CTL and tailored RT forecasts produced heavier rainfall than the ECMWF data over the northern region.

5.2 Mutual Sensitivities of Interests

The mutual sensitivities between forecast metrics and observation interests were analyzed using Spearman's rank correlation coefficients (CC). The results showed that the synoptic-type metric exhibited the highest sensitivity to observation interests, while the energy-type metric showed the second highest sensitivity. The economic metric exhibited the least sensitivity. The density-type observation interest showed the least sensitivity to forecast metrics, while the region-type interest showed the highest sensitivity.

Figure 5A shows the violin plot of Spearman correlation coefficients (CC) of benefit/cost ranking against different forecast metrics and observation interests. The results indicate that the synoptic-type metric has the highest sensitivity, followed by the energy-type metric, with the economic metric showing the least sensitivity.

Figure 5B shows the violin plot of CC for different observation interest types. The results indicate that the density-type observation interest has the least sensitivity, followed by the layer-type interest, with the region-type interest showing the highest sensitivity.

Figure 5C and **Figure 5**D show the significance levels (p-values) of the CC for forecast metrics and observation interests, respectively. The results indicate low statistical confidence in the sensitivity measures across forecast metrics, reflecting diminished sensitivity credibility. However, the density-type observation interest shows high reliability in benefit correlations.

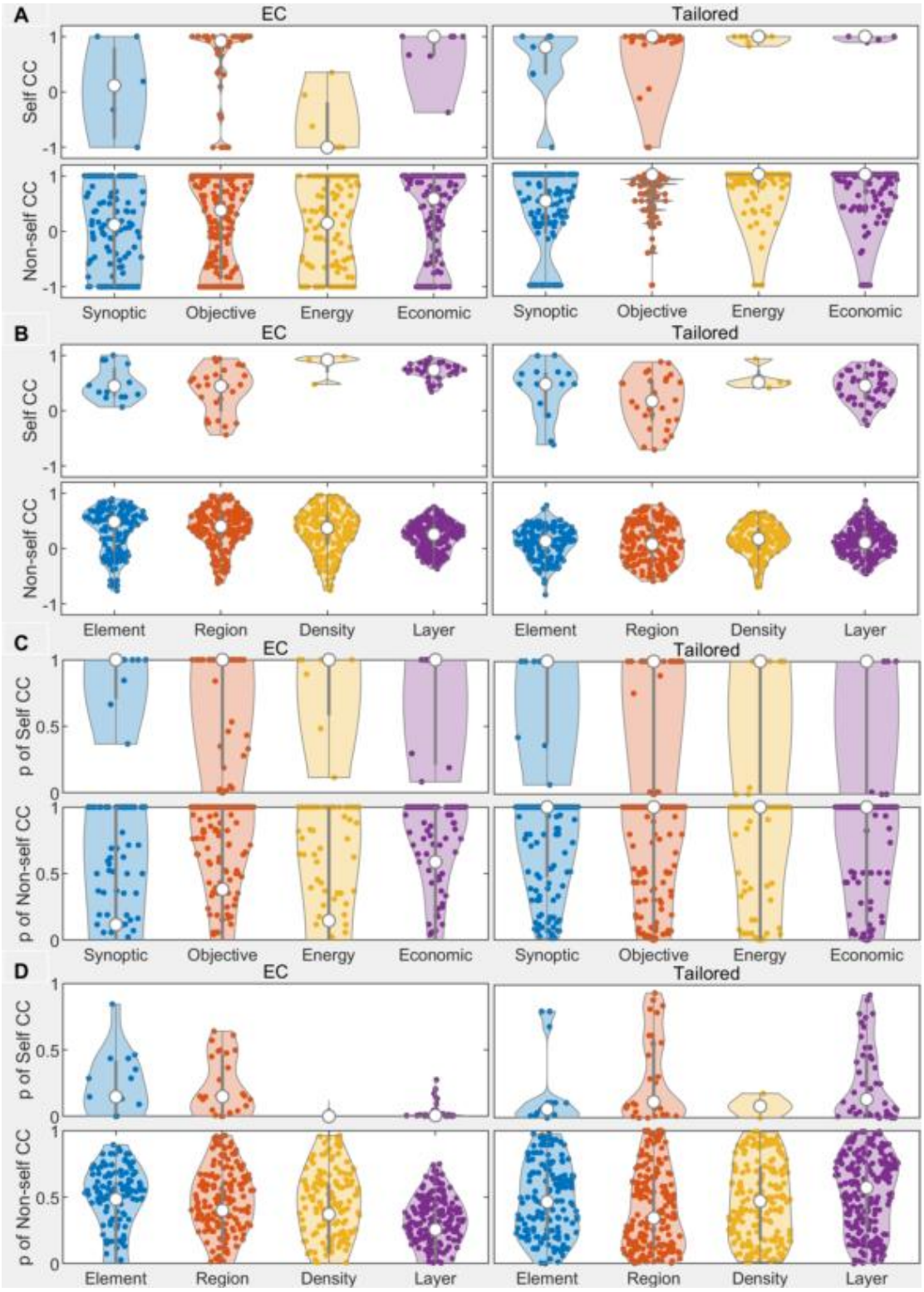


Figure 5 Violin plot of Sperman correlation coefficients (CC; units: 1) of benefit/cost ranking against different metric (A) and OIG types (B). C and D are same to A and B respectively, but for their significance levels p. The violin edges represent the kernel density distributions. The boxes, thick and slim whiskers represent the ranges of the four percentiles, median outliers and extreme outliers respectively, while the circles represent the median.

5.3 Relative Impacts of Observations

The relative impacts of different observation interests on forecast metrics were evaluated by comparing the benefits/costs across various metrics and references. The results showed that the dew point (D2) element and the central–west (CW) region had the most significant impacts on forecasts, while the temperature (T2) element and the middle–east (ME) region contributed the most to improving overall forecasts. The density enhancement from 30 km to 12 km demonstrated optimal robustness across all interests.

Figure 6A and **Figure 6**B show the bubble charts of the benefit/cost for various observation interests across different forecast metrics, referenced to ECMWF and tailored truth, respectively. The results indicate that the dew point (D2) element and the central–west (CW) region have the most significant impacts on forecasts.

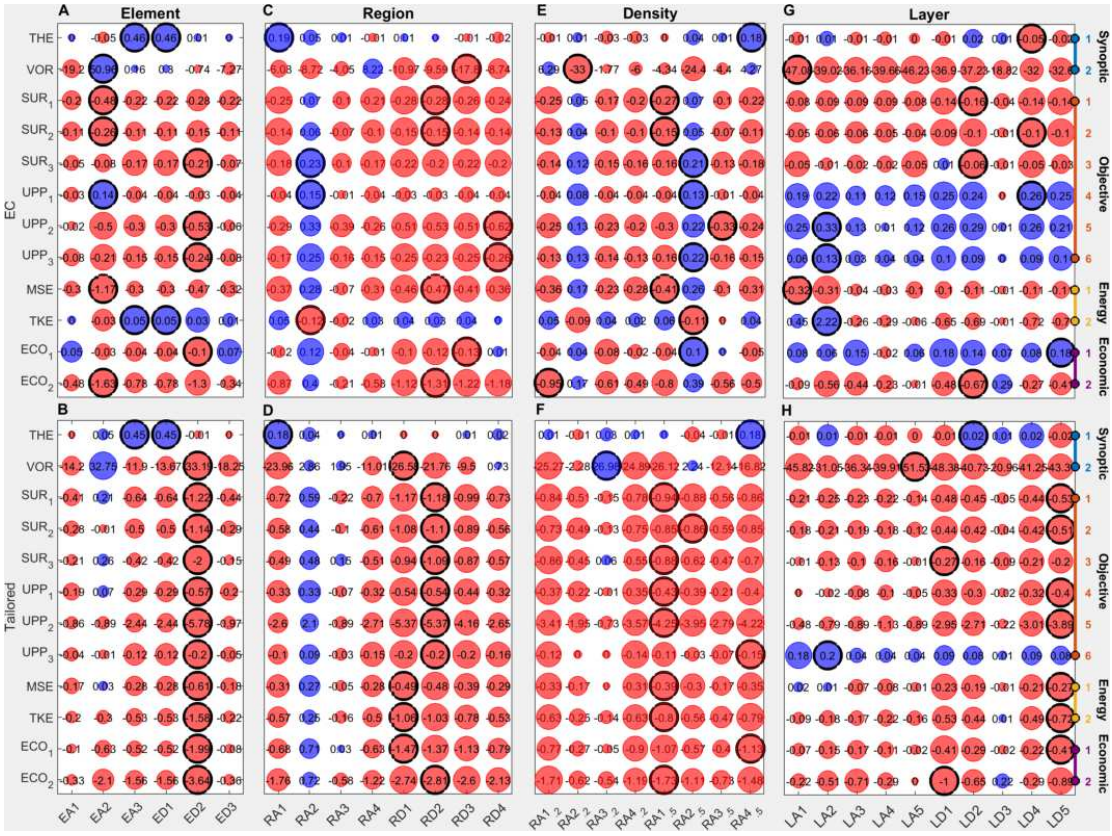


Figure 6 Bubble chart of the benefit/cost for various OIG across different *M*, where red and blue represent benefits and costs, respectively, and the bubble sizes, texts and thick circles indicate the scaled, actual, and C_{max}/B_{max} of *M* respectively, across single OIG type. A shows the benefit/cost of Element, referenced to EC, with x–axis representing members within one OIG and y–axis representing *M* names. B is the same to A, but referenced to tailored truth. C(D),

E(F), G(H) are the same to A(B), but for Region, Density, and Layer respectively. The *M*' type (colored lines) and it index (colored numbers) are shown at the right y–axis.

Figure 6C and **Figure 6**D show the bubble charts for region-type observation interests, indicating that the central-west (CW) region has the most significant overall impact on forecasts, while the middle-east (ME) region contributes the most to improving overall forecasts.

Figure 6E and **Figure 6**F show the bubble charts for density-type observation interests, indicating that the density enhancement from 30 km to 12 km demonstrates optimal robustness across all interests.

Figure 6G and **Figure 6**H show the bubble charts for layer-type observation interests, indicating that the 925 hPa layer has the most significant overall impact on forecasts, while the 500 hPa layer contributes the most to improving overall forecasts.

Table 3 Benefit numbers of single interest and extreme numbers of interest pair. summarizes the benefit numbers of single interests and extreme numbers of interest pairs for each observation interest group (OIG). The table provides detailed information on the number of metrics showing benefits and the number of extreme values for each interest, highlighting the overall impact of each observation interest on the forecasts.

Table 3 Benefit numbers of single interest and extreme numbers of interest pair.

OIG	$(M < 0)$*; single interest	$(C_{max} \cup B_{max})$*; interest pair
Element	21; EA1	3; EA1/ED1
	16; EA2	
	20; EA3	**21**; EA2/ED2
	20; ED1	
	22; ED2	3; EA3/ED3
	21; ED3	
Region	21; RA1	6; RA1/RD1
	2; RA2	
	19; RA3	**14**; RA2/RD2
	21; RA4	
	22; RD1	2; RA3/RD3
	22; RD2	
	22; RD3	2; RA4/RD4
	20; RD4	

Density	21; RA1$_{.2}$	**11**; RA1$_{.2}$/RA1$_{.5}$
	14; RA2$_{.2}$	
	20; RA3$_{.2}$	7; RA2$_{.2}$/RA2$_{.5}$
	22; RA4$_{.2}$	
	22; RA1$_{.5}$	2; RA3$_{.2}$/RA3$_{.5}$
	13; RA2$_{.5}$	
	22; RA3$_{.5}$	4; RA4$_{.2}$/RA4$_{.5}$
	22; RA4$_{.5}$	
Layer	17; LA1	4; LA1/LD1
	16; LA2	
	19; LA3	8; LA2/LD2
	20; LA4	
	19; LA5	0; LA3/LD3
	18; LD1	
	17; LD2	3; LA4/LD4
	16; LD3	
	18; LD4	**9**; LA5/LD5
	19; LD5	

*The bold numbers represents the maximum during one OIG.

The table shows that for the Element–type OIG, the dew point (D2) element has the most significant overall impact on forecasts, with the most extreme values and the highest number of metrics showing benefits when denied (ED2). For the Region–type OIG, the central–west (CW) region has the most significant overall impact, with the most extreme values and the highest number of metrics showing benefits when denied (RD2). For the Density–type OIG, the density enhancement in the northwest (NW) region has the most significant overall impact, with the most extreme values and the highest number of metrics showing benefits when increased from 30 km to 12 km (RA1$_{.2}$ and RA1$_{.5}$). For the Layer–type OIG, the 925 hPa layer has the most significant overall impact, with the most extreme values and the highest number of metrics showing benefits when denied (LD5).

5.4 Effects of Reference Truth

The choice of reference truth significantly impacted the sensitivity of forecast metrics to observation interests. The tailored reference showed lower sensitivity compared to the ECMWF reference, indicating that the tailored reference may overestimate the benefits of assimilation due to self–validation. The ECMWF reference, on the other hand, showed a more pessimistic performance, suggesting that it may underestimate the benefits of assimilation.

Figure 7A shows the scatter distribution of the self–CC of forecast metrics against different reference truths. The results indicate that the tailored reference exhibits a greater proportion of larger CC values compared to the ECMWF reference, suggesting lower sensitivity of forecast metrics to observation interests in the tailored reference.

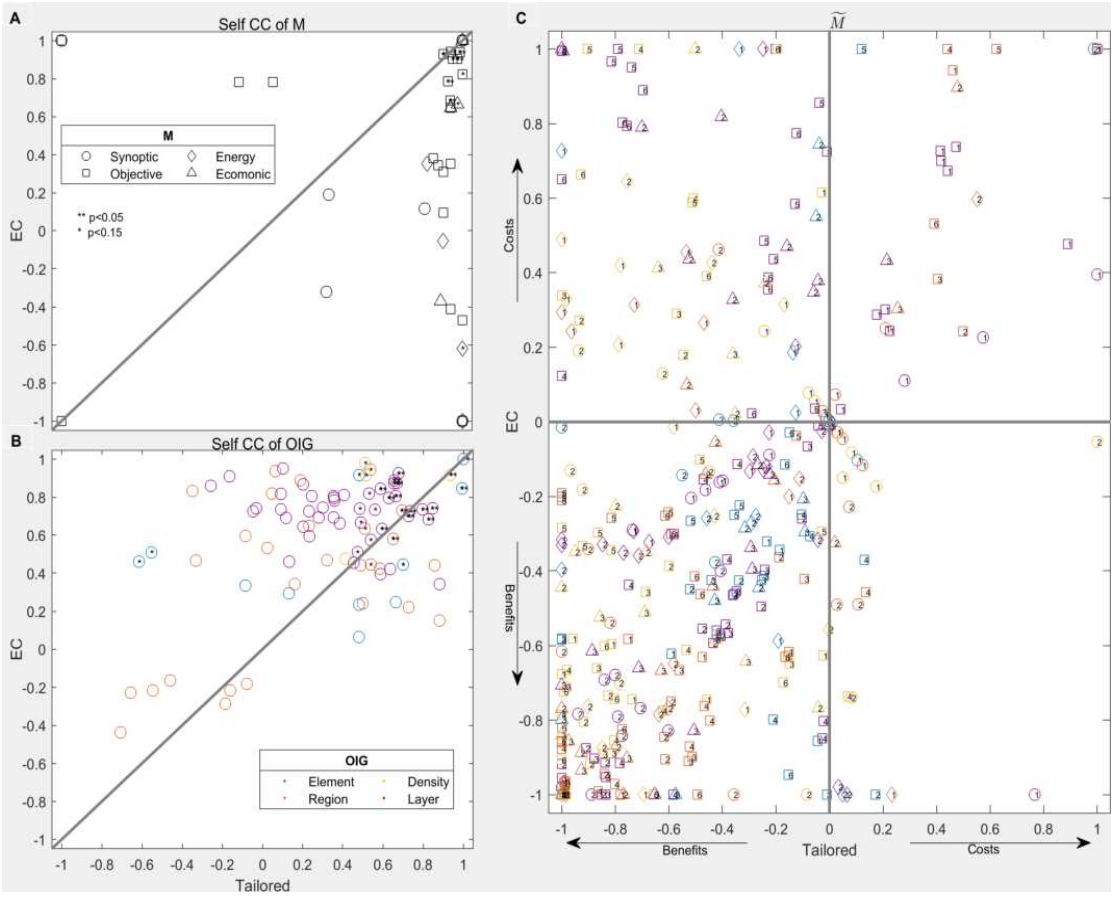


Figure 7 Scatter distribution of the self–CC of *M* (A), the self–CC of OIG (B), and $\widetilde{\boldsymbol{M}}$ (C) against different RT, that y– and x–axis represent EC and tailored respectively, and different metric and OIG types are shown in markers and colors respectively. The bold gray lines in A and B represent a equatable CC of different RT, while those in C represent the boundaries between benefit and cost within different RT. Also, the significance levels ($p<0.05$ and $p<0.15$) of CC within both EC and tailored are shown in A and B, and the metric indexes of Fig. 6 are shown in C.

Figure 7B shows the scatter distribution of the self–CC of observation interests against different reference truths. The results indicate that the tailored reference exhibits a higher frequency of smaller CC values compared to the ECMWF reference, suggesting higher sensitivity of observation interests to forecast metrics in the tailored reference.

Figure 7C shows the distribution of benefit/cost phases against different reference truths. The results indicate that joint benefits are the most prominent feature, highlighting the positive impact of assimilating different observational interests on forecasting. The significant contradictions within the ECMWF cost but tailored benefit zone highlight the need for careful trade–off of reference truth for layer– and density–type observation interests..

6. Discussion

6.1 Key Findings and Implications

The comprehensive evaluation of various observation interests on regional rainstorm forecasts in East China using multiple forecast metrics and an identical twin OSSE framework has yielded several key findings. These findings provide valuable insights into the mutual sensitivities and relative impacts of different observation interests, emphasizing the need for a balanced approach in managing complex observation interests within local NWP systems.

Firstly, the synoptic–type metric exhibited the highest sensitivity to observation interests. This indicates that the selection of observational interests significantly impacts the forecast accuracy for synoptic–scale features. This finding underscores the importance of carefully choosing observation interests when focusing on synoptic-scale forecasts. The energy–type metric showed the second highest sensitivity, followed by the objective metric, with the economic metric showing the least sensitivity. This suggests that different forecast metrics may require different strategies for optimizing observation interests.

Secondly, the relative impacts of different observation interests on forecast metrics were evaluated. The dew point (D2) element and the central–west (CW) region had the most significant impacts on forecasts, indicating that these observation interests should be prioritized in local NWP systems. The density enhancement from 30 km to 12 km demonstrated optimal robustness across all interests, highlighting the importance of increasing observation density in specific regions to improve forecast accuracy. The 925 hPa layer showed the most significant overall impact on forecasts, suggesting that lower atmospheric layers should be given more attention in observation and assimilation strategies.

Thirdly, the choice of reference truth significantly impacted the sensitivity of forecast metrics to observation interests. The tailored reference showed lower sensitivity compared to the ECMWF reference, indicating that the tailored reference may overestimate the benefits of

assimilation due to self–validation. The ECMWF reference showed a more pessimistic performance, suggesting that it may underestimate the benefits of assimilation. This highlights the need for careful consideration of reference truth selection in OSSEs to ensure reliable evaluation results.

6.2 Challenges and Future Work

Despite the valuable insights provided by this study, several challenges remain, and future work is needed to address these issues.

Firstly, the high sensitivity of synoptic metrics to observation interests indicates the need for further research to enhance their broader applicability. Future work should focus on developing more robust synoptic metrics that can better handle the complexities of regional weather systems. This will help in improving the accuracy and reliability of synoptic–scale forecasts. Secondly, the significant impact of reference truth on the evaluation results highlights the need for more reliable and higher–resolution reference datasets. Future work should explore the use of downscaled ECMWF data or other high-resolution datasets to improve the accuracy of reference truth in OSSEs. This will ensure that the evaluation results are more representative of real–world conditions. Thirdly, the results suggest that certain observation interests, such as dew point and central–west region, should be prioritized in local NWP systems. Future work should focus on optimizing the design of observation networks to better capture these critical interests and improve forecast accuracy. This may involve the deployment of additional observation stations in key regions and the use of advanced observational technologies. Lastly, the study highlights the importance of using appropriate data assimilation schemes to effectively integrate different observation interests into NWP models. Future work should explore the use of advanced assimilation techniques, such as hybrid schemes combining 3DVar and EnSRF, to further improve the assimilation of observational data. This will help in enhancing the overall performance of NWP models and improving the accuracy of weather forecasts.

7. Conclusion

In conclusion, this study provides a comprehensive evaluation of the impacts of various observation interests on regional rainstorm forecasts in East China using a combination of multiple forecast metrics and an identical twin OSSE framework. The results emphasize the

need for a balanced approach in managing complex observation interests within local NWP systems, considering the significant trade-offs and sensitivities involved. Future work should focus on enhancing the applicability of synoptic metrics, improving the accuracy of reference truth, optimizing observation network design, and advancing data assimilation techniques to further improve local weather forecasting and disaster mitigation strategies. These efforts will contribute to better preparedness and response to extreme weather events, ultimately enhancing public safety and reducing the socio-economic impacts of severe weather.

Funding: This research was jointly funded by the Henan Provincial Natural Science Foundation Project (grant number: 242300421367) and the Open Project of KLME CIC–FEMD NUIST (grant numbers: KLME202407).

Data Availability Statement:

The final reanalysis dataset (FNL) used in this study is openly available at https://rda.ucar.edu/datasets/d083002/dataaccess. The land cover dataset (MCD12Q1) is openly available at https://lpdaac.usgs.gov/products/mcd12q1v061/. The terrain dataset (SRTM) is openly available at https://srtm.csi.cgiar.org/download. The experimental data and code are openly available at 10.5281/zenodo.15644895.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix

All the calculation formulas in the definition of distance (***dis***).

A1. Lifting condensation level *(LCL)* calculation:

$$LCL = 123 * (T2 - D2) \quad (A1)$$

A2. Vortex center distance *(VCD)* calculation:

$$VCD(\Delta lat, \Delta lon) = R * 2 * atan2(\sqrt{\alpha}, \sqrt{1-\alpha}),$$
$$\alpha = sin^2(\Delta lat/2) + cos(lat) * cos(lat0) * sin^2(\Delta lon/2) \quad (A2–1)$$

$$\Delta lat = lat - lat^r; \Delta lon = lon - lon^r \quad (A2–2)$$

The (lat, lon) and (lat^r, lon^r) represent the vortexes' center location in (latitude, longitude) of the experiment and referenced data respectively, R represents the earth's radius and equals to 6743 Km. And α is a corresponding parameter.

A3. Root mean square error *(RMSE)* and root mean square vector error *(RMSVE)* calculation:

$$RMSE(\Delta a) = \sqrt{\frac{\sum_{i=1}^{i=nx}\sum_{j=1}^{j=ny}(\Delta a_i^j - \overline{\Delta a})^2}{2}}, \quad \Delta a = a - a^r \tag{A3–1}$$

$$RMSVE(\Delta u, \Delta v) = \sqrt{\frac{\sum_{i=1}^{i=nx}\sum_{j=1}^{j=ny}((\Delta u_i^j - \overline{\Delta u})^2 + (\Delta v_i^j - \overline{\Delta v})^2)}{2}}, \quad \Delta u = u - u^r, \Delta v = v - v^r \tag{A3–2}$$

The i and j represent the positions of horizontally orthogonal grid points respectively, while nx and ny represent their total numbers. The a and a^r represent the state variable of the experiment and reference data respectively, while (u, v) and (u^r, v^r) represent for the wind vector of the experiment and reference data respectively.

A4. *Moist static energy (MSE)* calculation:

$$MSE = C_p * T + g * H + L_v * Q \tag{A4}$$

The C_p, g, L_v represent the constants of heat capacity (1.005 KJ·kg^{-1}·K^{-1}), gravitational acceleration (9.8 m·s^{-2}) and latent heat capacity (2.5008 KJ·kg^{-1}·K^{-1}) respectively.

A5. Turbulent kinetic energy *(TKE)* calculation:

$$TKE = ((\frac{\partial u}{\partial h})^2 + (\frac{\partial v}{\partial h})^2) * C * K_m \tag{A4}$$

The K_m and C represent the momentum turbulence diffusion coefficient (1.0 m²/s) and the neutral stability empirical constant (0.02) respectively.

A6. Economic Cost-Loss Value *(ECLV)* calculation:

$$ECLV = \begin{cases} \frac{(cl*(h+f-1))+m}{cl*(br-1)}, when\ cl < br \\ \frac{(cl*(h+f))+m-br}{cl*(br-1)}, when\ cl \geq br \end{cases} \tag{A5}$$

The cl, br, h, f, and m represent the cost/loss rate, base rate, hit rate, false alarm rate, and miss rate respectively.

References

1.Abbasnezhadi, K., Rousseau, A. N., Koenig, K. A., Zahmatkesh, Z., and Wruth, A. M. (2019). Hydrological assessment of meteorological network density through data assimilation simulation. *J. Hydrol.*, 569, 844–858.

2.Anderson, J.L. (2007). An adaptive covariance inflation error correction algorithm for ensemble filters. *Tellus A*, 59(2), 210—224.

3.Arnold, J.C.P., Dey, C.H. Observing–Systems Simulation Experiments: Past, Present, and Future. *BAMS*, 1986, 67(6), 687–695.

4.Barker, D., Huang, X.Y., Liu, Z., Auligné, T., Zhang, X., Rugg, S., Ajjaji, R., Bourgeois, A., Bray, J., Chen, Y.S., et al. (2012). The weather research and forecasting model's community variational/ensemble data assimilation system: WRFDA. *BAMS*, 93(6), 831–843.

5.Bormann, N., Lawrence, H., and Farnan, J. (2019). Global observing system experiments in the ECMWF assimilation system (ECMWF Technical Memoranda, No. 839). *ECMWF,* pp. 1–18. (https://doi.org/10.21957/sr184iyz, accessed on 31 May 2025)

6.Bouttier, F., and Kelly, G. (2001). Observing–system experiments in the ECMWF 4DVAR assimilation system. *Q. J. R. Meteorol. Soc.*, 127, 1469–1488.

7.Chen, F., Kusaka, H., Bornstein, R., Ching, J., Grimmond, C.S.B., Grossman–Clarke, S., Loridan, T., Manning, K., Martilli, A., Miao, S., et al. (2011). The integrated WRF/urban modeling system: Development, evaluation, and applications to urban environmental problems. *Int. J. Climatol.*, 31(2), 273–288.

8.Chen, K., Hong, P.F., Han, W., Li, Z.Y., Wang, H., Wang, J.C., Chen, H., Zhang, Z.G., Xie, Z.C. (2021). Geostationary microwave observation system simulation experiments using the GRAPES 4DVar. *Acta Meteorologica Sinica*, 79(5), 769–785. (In Chinese)

9.Coumou, D., and Rahmstorf, S. (2012). A decade of weather extremes. *Nature Clim. Change*, 2, 491–496.

10.Curnel, Y., de Wit, A.J.W., Duveiller, G., and Defourny, P. (2011). Potential performances of remotely sensed LAI assimilation in WOFOST model based on an OSS Experiment. *Agric. Forest Meteorol.*, 151(12), 1843–1855.

11.Day, J.J., Sandu, I., Magnusson, L., Rodwell, M.J., Lawrence, H., Bormann, N., and Jung, T. (2019). Increased Arctic influence on the midlatitude flow during Scandinavian Blocking episodes. *Q. J. R. Meteorol. Soc.*, 145(725), 3846–3862.

12.Emanuel, K.A. (1994). Stability. In: online eds. Atmospheric Convection. *Oxford Academic*, pp. 165–188.

13.Fu, J., Ma, X., Chen, T., Zhang, F., Zhang, X., Sun, J., Quan, W., Yang, S., and Shen, X. (2017). Characteristics and Synoptic Mechanism of the July 2016 Extreme Precipitation Event in North China. *Meteorological Monthly*, 43(5), 528–539. (In Chinese)

14.Fu, S.M., Zhang, J.P., Tang, H., Jiang, L.Z., and Sun, J.H. (2020). A new mesoscale–vortex identification metric: restricted vorticity and its application. *Environ. Res. Lett.*, 15(12), 124053.

15.Gaspari, G., and Cohn, S.E. (1999). Construction of correlation functions in two and three dimensions. *Q. J. R. Meteorol. Soc.*, 125(554), 723–757.
16.Hoffman, R.N., and Atlas, R. (2016). Future Observing System Simulation Experiments. *BAMS*, 97(9), 1601–1616.
17.Hong, S.Y., Noh, Y., Dudhia, J. (2006). A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, 134, 2318–2341.
18.Iacono, M.J., Delamere, J.S., Mlawer, E.J., Shephard, M.W., Clough, S.A., Collins, W.D. (2008). Radiative forcing by long–lived greenhouse gases: Calculations with the AER radiative transfer models. *J. Geophys. Res. Atmos.*, 113, D13103.
19.Jaynes, E.T. (2003). Elementary sampling theory. In G.L. Bretthorst, Ed., *Probability theory: The logic of science*, 51–85. Cambridge University Press, Cambridge.
20.Jimenez, P.A., Dudhia, J., Gonzalez–Rouco, J.F., Navarro, J., Montavez, J.P., Garcia–Bustamante, E. (2012). A revised scheme for the WRF surface layer formulation. *Mon. Wea. Rev.*, 140, 898–918.
21.Kain, J.S. (2004). The Kain–Fritsch convective parameterization: An update. *J. Appl. Meteor.*, 43, 170–181.
22.Kwon, Y., Yoon, Y., Forman, B. A., Kumar, S. V., and Wang, L. (2021). Quantifying the observational requirements of a space–borne LiDAR snow mission. *J. Hydrol.*, 601, 126709.
23.Lawrence, H., Bormann, N., Sandu, I., Day, J., Farnan, J., and Bauer, P. (2019). Use and impact of Arctic observations in the ECMWF Numerical Weather Prediction system. *Q. J. R. Meteorol. Soc.*, 145(725), 3432–3454.
24.Li, J., Zhang, P., Chen, L., Han, W., Lu Q.F., Gong, X.Y., Wang, P., and Li, Z.L. (2021). An Overview on the Development of the Observing System Simulation Experiments (OSSEs). *Adv. Meteor. Sci. Technol.*, 11(3), 27–34. (In Chinese)
25.Liu, L.L., Li, Y.H., Wang, X.L.,Wang, W., Wu, B.G., Qiu, X.B., and Chen, J. (2019). Study of Sea Surface Meteorological Observation Station Schemes Based on Observational System Simulation Experiments. *Clim. Environ. Res.*, 24(6): 711–722. (In Chinese)
26.Luo, Y.L., Sun, J.S., Li, Y., Xia, R.D., Du, Y., Yang, S., Zhang, Y.C., Chen, J., Dai, K., Shen, X.S., et al. (2020). Scientific research and forecasting of heavy rain in China: Achievements in the 40 years since the reform and opening up. *Acta Meteorologica Sinica*, 78(03), 419–450. (In Chinese)
27.Maejima, Y., Kawabata, T., Seko, H., and Miyoshi, T. (2022). Observing System Simulation Experiments of a Rich Phased Array Weather Radar Network Covering Kyushu for the July 2020 Heavy Rainfall Event. Sola, 18(0), 25–32.
28.Melet, A., Verron, J., and Brankart, J.M. (2012). Potential outcomes of glider data assimilation in the Solomon Sea: Control of the water mass properties and parameter estimation. *J. Marine Syst.*, 94, 232–246.
29.Mishra, A.K., and Krishnamurti, T.N. (2011). Observing System Simulation Experiment for Global Precipitation Mission. *Pure Appl. Geophys.*, 169(3), 353–365.

30.Nag, S., Gatebe, C. K., and Weck, O.D. (2015). Observing system simulations for small satellite formations estimating bidirectional reflectance. *Int. J. Appl. Earth Obs. Geoinform.*, 43, 102–118

31.Neelin, J.D., and Held, I.M. (1987). Modeling tropical convergence based on the moist static energy budget. *Mon. Wea. Rev.*, 115(1), 3–12.

32.Nicolsky, D.J., Romanovsky, V.E., and Panteleev, G.G. (2009). Estimation of soil thermal properties using in–situ temperature measurements in the active layer and permafrost. *Cold Reg. Sci. Technol.*, 55(1), 120–129.

33.Parrish, D.F. and Derber, J.C. (1992). The National Meteorological Center's Spectral Statistical Interpolation Analysis System. *Mon. Wea. Rev.*, 120, 1747–1763.

34.Qin, Y.Y., Gong, J.D., and Li, Z.C. (2012). Assimilation of Doppler Radar Observations with an Ensemble Kalman Filter: OSS Experiments. *Meteor. Monthly*, 38(5), 513–525. (In Chinese)

35.Richardson, D.S. (2000). Skill and relative economic value of the ECMWF ensemble prediction system. *Q. J. R. Meteorol. Soc.*, 126, 649–667.

36.Sato, K., Inoue, J., Yamazaki, A., Kim, J.H., Maturilli, M., Dethloff, K., Hudson, S. R., and Granskog, M. A. (2017). Improved forecasts of winter weather extremes over midlatitudes with extra Arctic observations. *J. Geophys. Res. Oceans*, 122(2), 775–787.

37.Shao C.L., and Min, J.Z. 2015. A Study of the Assimilation of Surface Automatic Weather Station Data Using the Ensemble Square Root Filter. *Chinese J. Atmos. Sci.*,39(1):1–11. (In Chinese)

38.Shou, S.W. (2019). Research progress in the synoptic meteorology of heavy rainstorms in China. *J. Nat. Disasters*, 38(5), 450–463. (In Chinese)

39.Tang, H., Fu, S.M., Sun, J.H., and Zhou, X.X. (2023). Three–Dimensional Objective Identification of the Tibetan Plateau Vortex Based on a Reanalysis Wind Field with a High Spatial and Temporal Resolution. *J. Atmos. Sci.*, 47(3): 698−712. (In Chinese)

40.Tewari, M., Chen, F., Wang, W., Dudhia, J., LeMone, M.A., Mitchell, K., Ek, M., Gayno, G., Wegiel, J., Cuenca, R.H. (2004). Implementation and verification of the unified NOAH land surface model in the WRF model. *In Proceedings of the 20th Conference on Weather Analysis and Forecasting/16th Conference on Numerical Weather Prediction*, Seattle, WA, USA, 12–16 January 2004, pp. 11–15.

41.Timmermans, R. M. A., Lahoz, W. A., Attié, J. L., Peuch, V. H., Curier, R. L., Edwards, D. P., Eskes, H. J., and Builtjes, P. J. H. (2015). Observing System Simulation Experiments for air quality. *Atmos. Environment*, 115, 199–213.

42.Thompson, G., Field, P., Rasmussen, R., Hall, W. (2008). Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization. *Mon. Wea. Rev.*, 136, 5095–5115.

43.Trenberth, K., Dai, A., Rasmussen, R. and Parsons, D. (2003). The Changing Character of Precipitation. *BAMS*, 84(9), 1205–1217.

44.Wilks, D.S. (2001). A skill score based on economic value for probability forecasts. *Met. Appl.*, 8, 209–219.

45.Xia, R. and Zhang, D. (2019). An Observational Analysis of Three Extreme Rainfall Episodes of 19–20 July 2016 along the Taihang Mountains in North China. *Mon. Wea. Rev.*, 147(11), 4199–4220.

46.Xian, J., Qiu, Z., Luo, H., Hu, Y., Lin, X., Lu, C., Yang, Y., Yang, H., and Zhang, N. (2025). Turbulent energy budget analysis based on coherent wind lidar observations, *Atmos. Chem. Phys.*, 25, 441–457.

47.Xu, Z.F., Gong, J.D., Wang, J.J., Li, Z.C. (2007). A study of assimilation of surface observational data in complex terrain Part Ⅱ: Representative error of the elevation difference between model surface and observation site. *Chinese J. Atmos. Sci.*, 31(3): 449–458. (In Chinese)

48.Yu, X.D., (2002). Strengthening the construction of observation network and promoting the development of numerical forecasting business. *Meteor. Sci. Technol.*, 30(1), 15–18. (In Chinese)

49.Zeng, X., Atlas, R., Birk, R. J., Carr, F. H., Carrier, M. J., Cucurull, L., Hooke, W. H., Kalnay, E., Murtugudde, R., Posselt, D. J., Russell, J. L., Tyndall, D. P., Weller, R. A., and Zhang, F. (2020). Use of Observing System Simulation Experiments in the United States. *BAMS*, 101(8), E1427–E1438.

50.Zhang, J.P., Jin, S.L., Feng, S.L., Han, H., Fu, S.M. (2023). A New Objective Identification Method for Mesoscale Vortices: Three–Dimensional Tracking and Quantitative Evaluation. *Chinese J. Atmos. Sci.*, 47(5): 1434−1450. (In Chinese)

51.Zhao, X., Zhu, J., Cheng, L., Liu, Y., and Liu, Y. (2020). An Observing System Simulation Experiment to Assess the Potential Impact of a Virtual Mobile Communication Tower–based Observation Network on Weather Forecasting Accuracy in China. Part 1: Weather Stations with a Typical Mobile Tower Height of 40 m. *AAS*, 37(6), 617–633.

52.Zhu, Y., Li, Y., and Peng, S. (2020). On evaluating the effect of assimilating glider–observed T/S profiles with different horizontal resolutions and assimilation frequencies. *Ocean Dynam.*, 70(6), 827–837.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Guoetal2025MvsOIGsuppletmentsSRV1.docx