https://doi.org/10.5194/egusphere-2024-1554
Preprint. Discussion started: 28 June 2024

Quantifying the analysis uncertainty for nowcasting application

Yanwei Zhu[1,2], Aitor Atencia[3], Markus Dabernig[3], Yong Wang[1,4]

[1]School of Atmospheric Science, Nanjing University of Information Science and Technology, Nanjing, China
[2]HuaFeng Research Lab for Weather Science and Applications, Nanjing University of Information Science and Technology, Nanjing, China
[3]GeoSphere Austria, Vienna, Austria
[4]CMA Earth System Modelling and Prediction Centre, China Meteorological Administration, Beijing, China

Correspondence to: Yong Wang (yong.wang@nuist.edu.cn)

Abstract. This study proposes a method to quantify the uncertainty of the error in the very high–resolution analysis in near–surface level for nowcasting application. We perturbed the first guess field and observation with Gaussian–distributed perturbations, which have variance equal to that of first guess error. The first guess error not only reflects the spatial characteristic of the difference between the first guess field and observation but also dominates the major uncertainty information of analysis errors. It is important to consider the attenuation of uncertainty dispersion caused by interpolation. Gaussian perturbations are combined with an inflation factor to estimate the attenuation of perturbation dispersion. To assess this method, it was applied to high–resolution analysis and nowcasting in the Beijing–Tianjin–Hebei region for hourly temperature, humidity and wind components. To evaluate the transmission of perturbation information in the nowcasting extrapolation, the ensemble analysis is used to compute ensemble nowcasting. The verifications show that the ensemble analysis has reasonable spread and high reliability, demonstrating effective and accurate quantification of the analysis error uncertainty. Verifications of ensemble nowcasting illustrate that the ensemble spread has effective growth within the nowcast extrapolation up to a lead time of 2 hours, but highly depends on the trend of NWP. The results prove that the propagation of analysis uncertainty representation in nowcasting extrapolation can match to the error increment, beneficial for estimating the near–surface nowcasting uncertainty.

1 Introduction

Nowcasting is essential for severe weather warnings and protecting life and property, as it accurately predicts high–impact weather events in near–real time (Wang et al., 2017a, b; Wastl et al., 2018; Schmid et al., 2019). A very high–resolution weather analysis is the base for a skilful nowcasting since the analysis describes the real–time atmospheric conditions accurately at the initial time (Wastl et al., 2021). However, due to the chaotic nature of the atmosphere, errors in data and the imperfect nowcasting models, nowcasting is with uncertainties (Lorenz, 1965; Leith, 1974; Kann et al., 2012; Glahn and Im, 2013; Wastl et al., 2019). To deal with these uncertainties in nowcasting, generating ensemble by appropriate perturbation is an effective method (Leutbecher et al., 2007; Leutbecher and Palmer, 2008).

https://doi.org/10.5194/egusphere-2024-1554
Preprint. Discussion started: 28 June 2024

In recent years, the use of ensemble nowcasting has become increasingly widespread (Wang et al., 2017b, 2021; Yang et al., 2023). Numerous studies have demonstrated that addressing all sources of the uncertainty represented by errors is a key aspect in generating ensemble nowcasting (Sun et al., 2014; Thiruvengadam, 2020). Due to the significant time dependency, nowcasting is sensitive to the uncertainty at the initial time. However, due to the limitation of the instrument accuracy and errors in computation, analyses are with uncertainty, which causes a major uncertainty in the initial time of nowcasting (Eibl and Steinacker, 2017; Keresturi er al., 2019). Hence, how to quantify the uncertainties in the analysis is one of the major challenges for ensemble nowcasting (Wang et al., 2017a; Taylor et al., 2022).

A widely recognized and applied approach to quantify the analysis uncertainty is introducing proper perturbations to generate ensemble, based on the characteristics of errors (Buizza et al., 2005; Zhu, 2005; Bouttier et al., 2016; Chen et al., 2016; Wang et al., 2017a; Lin et al., 2022). Most studies dedicated to tackling the uncertainty of nowcasting and only a few works explored the impact of overall uncertainty of analysis (Bouttier, 2019; Wang et al., 2021). Wang et al. (2014) and Suklitsch et al. (2015) presented that introducing additional perturbation to estimate the analysis error uncertainty could impact positively on simulating nowcasting uncertainty. Aire Limitée Adaptation dynamique Développement InterNational–Limited Area Ensemble Forecasting (ALADIN–LAEF) and Convection–permitting Limited–Area Ensemble Forecasting system (C–LAEF) are two skilful systems, which have ensemble analysis with 16 members to represent the analysis uncertainty. They consider the uncertainty caused by observation error through perturbing observation while the 16 first guess fields providing important uncertainty information within the three–dimensional background (more details in Wang et al., 2011; Bellus et al., 2016; Wastl et al., 2021). But neither ALADIN–LAEF nor C–LAEF discussed the impact of error uncertainty in analysis computation. Though Saetra et al. (2004) interpreted the impact of observation error on ensemble under–dispersion, Horányi et al. (2011) and Bellus et al. (2016, 2019) proved that perturbation can simulate the observation error in analysis, they did not delve into other errors in analysis, for instance, error made in interpolation (Wastl et al., 2021). Hence, consider the various errors of analysis is beneficial for quantifying the analysis uncertainty more accurately (Suklitsch et al., 2015).

The accurate analysis, which describes current atmosphere details, is typically derived by assimilating the first guess field (typically provided by numerical weather prediction models, NWP) and observation data (Randriamampianina and Storto, 2008; Kann et al., 2009; Haiden et al., 2010, 2011; Lin et al., 2022). The observations provide an approximate estimation of the true atmospheric values, while the three–dimensional first guess field provides a complete spatial structure within the interested region (Sun et al., 2013; Hoteit et al., 2015; Casellas et al., 2021). However, when combining observations with terrain–corrected first guess fields, interpolation errors caused by the algorithm are also present (Leutbecher and Palmer, 2008; Feng et al., 2020). The current works have not clearly indicated the impact of interpolation-induced error uncertainty on the analysis and nowcasting. Hence, it is of great significance to study how to accurately estimate the interpolation error of the analysis for a comprehensive understanding of analysis uncertainty and nowcasting uncertainty.

Integrated Nowcasting through Comprehensive Analysis (INCA, Haiden et al., 2010, 2011) system calibrates the first guess field (NWP) by automatic weather station observation. Within the NWP calibration, topographic characteristic factor

https://doi.org/10.5194/egusphere-2024-1554
Preprint. Discussion started: 28 June 2024

is used to correct the surface layer to match the actual terrain (Kann et al., 2009; Haiden et al., 2010, 2011). Seamless Integrated Weather Prediction and Applications (SIVA) is a nowcasting system based on INCA framework and is applicated in Beijing–Tianjin–Hebei region (BTH). There are dense stations and high resolution (1km×1km) first guess NWP in BTH, which are used to compute near–real–time analysis in SIVA. Since there is only one deterministic NWP model to as first guess, the uncertainty of SIVA analysis has not been taken into account at present version. Hence, it is necessary to consider the error within computation to quantify the uncertainty of SIVA analysis.

This work proposes a perturbation method to accurately quantify the analysis uncertainty of near–surface temperature, humidity and wind components in SIVA by introducing perturbation in analysis. An inflation factor is calculated to rescale the perturbation, avoiding the attenuation caused by interpolation. A set of ensemble analyses is generated to represent the analysis uncertainty in BTH. With cross–validation, the capability of ensemble analyses quantifying the uncertainty in no–station area is evaluated. To evaluate the impact of the perturbation in nowcasting extrapolation, ensemble analyses are applied in SIVA nowcasting to obtain ensemble nowcasting. Probabilistic verifications of ensemble nowcasting for months are aimed to evaluate the effectiveness of uncertainty representation of analysis in SIVA nowcasting.

This article is organized as follows. In section 2 the algorithm of SIVA analysis is introduced. Section 3 elaborates the characteristics of errors and perturbation methods. Section 4 is dedicated to the verification results of ensemble analysis and ensemble nowcasting. A summary and conclusions are given in section 5.

2 Method and data

SIVA is a nowcasting system built in BTH region, which creates analysis hourly at the horizontal resolution 1km for near–surface temperature, humidity, wind speed and precipitation. It calculates analysis in such a way that uses observation data and topographic parameters to calibrate the three–dimensional NWP first guess field. The automatic weather stations provide hourly observed value while the NWP output is available twice daily at 00:00 and 12:00 (UTC+8hour), up to a lead time 48 hour. These data can be obtained by submitting an application through the data website of the China Meteorological Administration. The algorithm of the 2m temperature, humidity and 10m wind speed analysis module of SIVA is mainly divided into the following:

Step 1: Three times linear interpolation of NWP data to SIVA grid as the first guess field.

Step 2: Revise the NWP data by inverse distance weight interpolation (IDW) using ground observation data to obtain a three–dimensional revised field (3DRF). The interpolation process in this step is denoted as 3D interpolation.

Step 3: Combining the topographic features and observation data, the 3DRF is revised again and then vertical interpolated to the lowest model level to obtain the near–surface revised field (NRF). The interpolation process in this step is denoted as 2D interpolation.

https://doi.org/10.5194/egusphere-2024-1554
Preprint. Discussion started: 28 June 2024

Since SIVA is a new version of INCA built in BTH, readers can find more details of the algorithm in Haiden et al. (2010, 2011). To avoid the confusion of the word error, we denote the difference between NWP and observation in Step 2 as error1 and the difference in near–surface level between calibrated NWP and observation in Step 3 as error2.

The experimental periods are in August 2020 (hereafter called summer) and in February 2021 (winter). With the high horizontal resolution, SIVA can describe the geographic features in detail. In addition, 21 vertical levels corresponding to different altitude such as 0 m, 200 m … 4000 m above the ground (wind speed has 32 levels in vertical with intervals 125 m) are to match the stations at different altitudes ensure that NWP can be corrected in combination with the topographic parameters at station location. Three meteorological elements are discussed in this study: temperature (T2m) and humidity (relative humidity RH2m and specific humidity QQ2m) 2 meters above the surface and the wind speed component (U10m and V10m) of 10 meters above the ground. Relative humidity is converted from specific humidity in SIVA and uncertainty is expressed by the perturbation of specific humidity.

BTH has a unique and complex topographic structure, with the northwestern part between two mountain ranges and the southeastern part belonging to the North China Plain (Fig. 1). Such topographic information and station altitude is used to imposes terrain constraints on the SIVA grid points. There are 1670 automatic weather stations in BTH that can pass quality control to be used in T2m and humidity analysis while 2351 stations are available for wind components analysis. Selected 151 stations randomly as test set to verify the performance of the obtained ensemble, while the remaining 1519 stations to be training set participating SIVA computation to generate ensemble (Fig. 1). Since wind components have different quality control framework with temperature, the total stations for wind are 2350 and the test stations are 192.

https://doi.org/10.5194/egusphere-2024-1554
Preprint. Discussion started: 28 June 2024

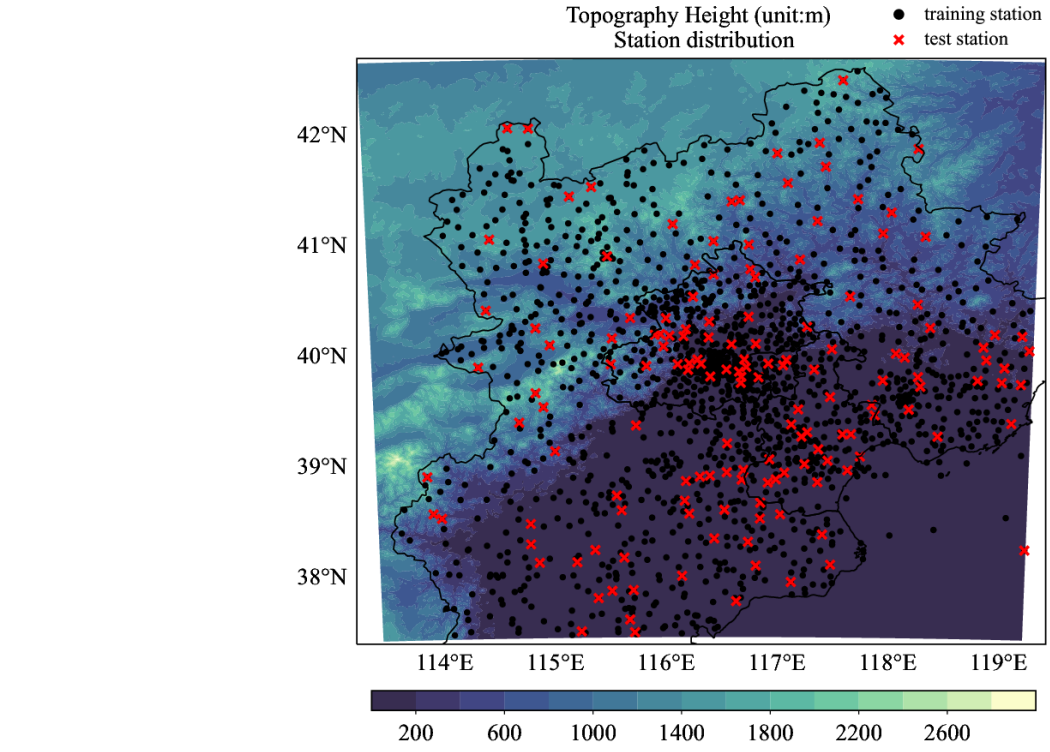


Fig. 1. Topography height (unit: m) and station distribution (used in T2m and QQ2m, black dots: training stations, red dots: test stations) over the Beijing–Tianjin–Hebei region.

To evaluate the effectiveness of ensemble analyses in nowcasting extrapolation, ensemble nowcasting is computed by the ensemble analysis. The ensemble nowcasting runs started at each hour and extrapolates up to 2h. Verification of ensemble nowcasting is covered the test stations in figure 1.

3 Perturbation method

For T2m, error1 is based on the principle of minimal required correction to filter out the excessive forecast errors, leading to the existence of a partial 0 value for the error though the true error1 value is nonzero. Therefore, the forecast error of NWP is one of the important sources of the uncertainty of error1. Daily variations of T2m are more significant in the mountains than in the plains. And due to terrain constrain, the spatial distribution of error1 and error2 exhibits distinct topographic features and the error of SIVA analysis against to observation will reflect these geographically characteristics (Fig. S1).

To estimate the uncertainty of NWP in the entire field, the standard deviation (STD) of error1 is interpolated into the grids. In every grid point, the perturbation is a Gaussian noise with 0 mean and STD equal to that of error1. Since error1 is the distance of NWP to the observation in the lowest model level, it implicates the downward (upward) shift that constrains

https://doi.org/10.5194/egusphere-2024-1554
Preprint. Discussion started: 28 June 2024

model height to valley–floor–surface (mountain peak or ridge). This means in vertical interpolation the forecast error will combine with geographic feature. In addition, interpolating 3DRF to surface level depends on the level–to–level error within NWP. Hence, the NWP perturbation is sampled from Gaussian noise with shape (M, Z) where M is number of members and Z is the numbers of levels. Such perturbation can represent the NWP systemic uncertainty in vertical dimension and the horizontal distribution of interpolated error1 provides the spatial characteristic. The difference between the results of ensembles with different numbers of members is only around 1%. To save computational costs and ensure a sufficient number of members, the number of ensemble members is set 20.

Though perturbed NWP already provide some noise information, the uncertainty range represented by error1 at each station will still be underestimated due to the filtering effect of the principle of minimal required correction. Thus, it is important to perturb the observation. Similarly, the perturbation is sampled randomly from Gaussian noise with 0 mean and standard deviation equal to that of error1 in each station.

SIVA interpolates error2 horizontally and combines it with inversion factors at each grid point to obtain 2D surface layer correction. The 2D interpolation entails scaling the error by the terrain difference between SIVA grid point and altitude of the corresponding stations. Therefore, the no–station areas receive perturbation from 3DRF but lack the information of 2D surface layer correction in corresponding points, i.e., these areas exist a loss of perturbation information. As a result, the corresponding ensemble spread will be underestimated. This can be considered as perturbation attenuation caused by 2D interpolation. To tackle the effects of 2D interpolation, an inflation factor is calculated. The inflation factor is the ratio of root–mean–square error (RMSE) and spread in test stations of the ensemble obtained by perturbing both NWP and observation. It serves as an indicator of the attenuation of perturbation information after interpolation. The figure S2 shows the spatial distribution of inflation factor, it relates to the test stations distribution, indicates the degree of the under–dispersion. Combining with the inflation factor, the perturbations are generated as follows:

$$\mathbf{NWP}'_{(i,j)} = \mathbf{NWP}_{(i,j)} + \mathbf{N}(0, STD_{2D-error1(i,j)} * infla\_factor_{(i,j)}), \quad (1)$$

$$\mathbf{OBS}'_{k} = \mathbf{OBS}_{k} + \mathbf{N}(0, STD_{error1(k)}). \quad (2)$$

Here, 2D–error1 represents the interpolated STD of error1. The subscript (i, j) denotes the SIVA grid point and k indicates the kth station. It should be emphasized that the observation of wind has not been perturbed. Since in SIVA, the true values of wind components are obtained through the sine and cosine conversion, perturbing the wind components by error–based noise cannot reflect the uncertainty causing by the wind direction.

Due to the impact of the principle of minimal required correction in T2m analysis, a significant decrease in the dispersion of perturbed error1 can be observed compare to unperturbed error1 (Fig. 2a). Nevertheless, the impact of the inflation factor guarantees that the dispersion of ensemble errors will not attenuate excessively (Fig. 2a, b). The fact that both training and test stations exhibit error dispersion consistent with perturbed error1 substantiates the effective dissemination of the perturbation information in no–station area through the topographic-based interpolation. The generated ensembles have uncertainty representation with equal reliability in whole areas (Fig. 2).

https://doi.org/10.5194/egusphere-2024-1554
Preprint. Discussion started: 28 June 2024

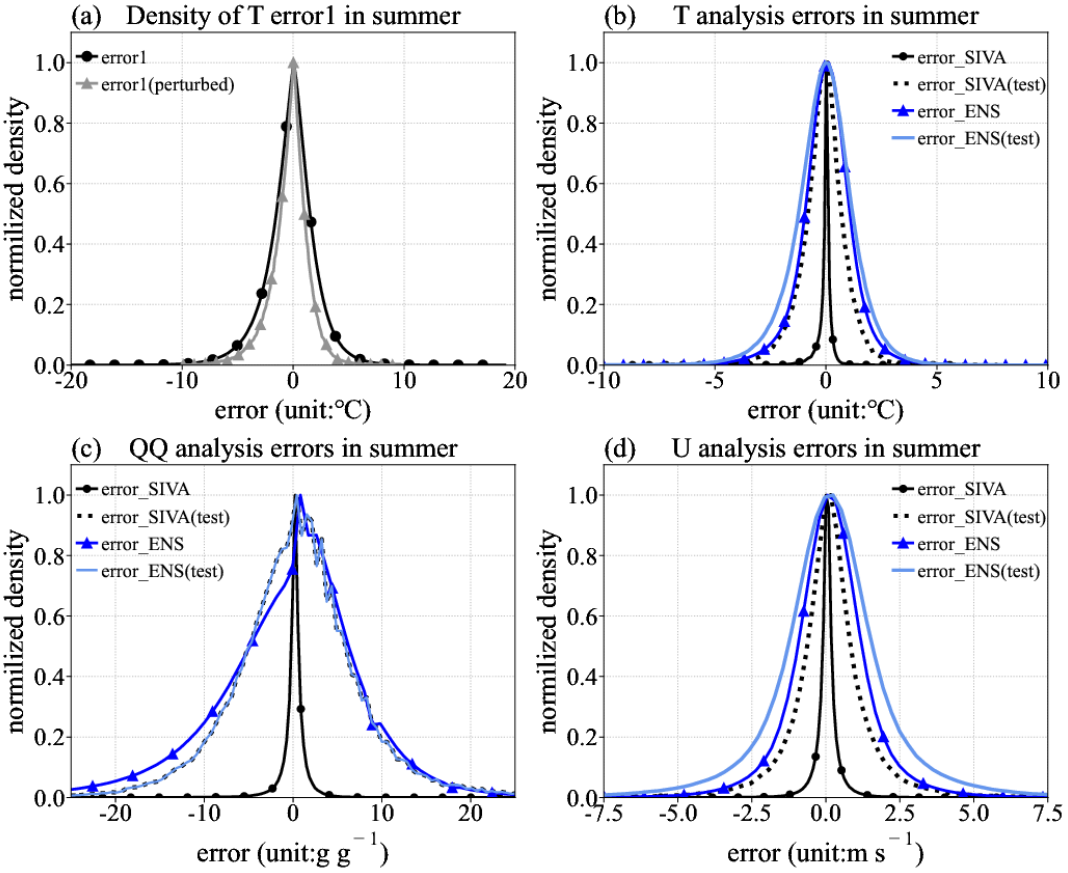


Fig. 2. Probability density function (PDF) of errors. The error1 without perturbation is denoted by the red line and with perturbation by the green line (a). Black (blue) solid line is the distribution of the error between SIVA deterministic analysis (ensemble analysis) and observations, over all training stations. Dashed line and light blue line are the distribution of the error for test stations.

4 Verification

4.1 Verification of ensemble Analysis

The deterministic SIVA analysis (hereafter call CANA) is utilized as a reference to assess the performance of the generated ensemble estimating the error uncertainty in test stations. In cross–validation, the commonly used probability validation scores such as RMSE, ensemble spread, Talagrand diagram and reliability diagram are used to evaluate the effectiveness of uncertainty representation at the test stations.

https://doi.org/10.5194/egusphere-2024-1554
Preprint. Discussion started: 28 June 2024

4.1.1 Ensemble RMSE and Spread

For the verification of T2m, QQ2m and RH2m ensembles, a total of 151 stations are used (one station in southeast is only available in summer and there are 150 test stations in winter). The wind components have different quality control procedure with T2m and QQ2m, and there are 191 test stations to evaluate the performance. The value of RMSE, and ensemble spread are given in table1.

Table 1. RMSE and spread of the ensemble analyses cross validation for a summer and a winter month. The subscript C (E) denotes the RMSE of CANA (ensemble analysis). Rrs is the ratio of ensemble RMSE and spread. Averaged over all the test stations and whole period.

	Summer				Winter			
	$RMSE_C$	$RMSE_E$	Spread	R_{rs}	$RMSE_C$	$RMSE_E$	Spread	R_{rs}
T2m (℃)	0.96	0.96	0.89	1.06	1.23	1.23	1.18	1.03
RH2m (%)	6.92	6.73	4.58	1.37	6.57	6.57	5.67	1.17
QQ2m (g g^{-1})	1.09	1.08	0.95	1.12	0.24	0.23	0.23	0.99
U10m (m s^{-1})	1.25	1.28	1.29	0.99	1.90	1.96	1.96	0.99
V10m (m s^{-1})	1.37	1.41	1.40	0.99	1.85	1.91	1.90	1.00

The ensembles for T2m have RMSE lower than that of the CANA not only for the average over all stations but also for individual stations. The average RMSE in summer has been reduced by around 0.01℃ (Table 1) and the mean RMSE reduction is approximately 0.0076 ℃. RMSE reduction is little in winter month, due to the more significant diurnal cycle of temperature in cold season. In mountainous, the RMSE receives a negative effect due to the lack of precise information about inversion heights. There is a marine station in summer, located far away in the southeast, which receives perturbation information mainly from 3DRF and only obtains NRF information with little weighted. With the effect of inflation factor, the perturbation in the corresponding grids of that station has been amplified to align the ensemble spread with RMSE. In addition, the perturbed NWP reflects the spatial uncertainty on the grid points. Hence, the distance weight in 3D interpolation of 3DRF can take into account the uncertainty derived by the geographic characteristic of NWP model. Most of the test stations have spread consistent with RMSE, indicating the accuracy of the ensemble in estimating the error uncertainty (Fig. S3).

https://doi.org/10.5194/egusphere-2024-1554
Preprint. Discussion started: 28 June 2024

The QQ2m has similar characteristic as T2m while the value limitation in variables conversions impacts the convey of perturbation information, which derives an under–estimated dispersion for RH2m. Nevertheless, the ensemble of humidity maintains the errors consistency with that of CANA and provides a reliable uncertainty estimation of errors.

For wind components, error uncertainties are totally represented by the perturbation of first guess field. The increased RMSE of ensemble mean might be due to the lack of observation perturbation or the divergence constraint of wind components. But the consistency of ensemble spread and RMSE indicates the effective representation of the error uncertainty. There is a probably way that use multiple first guess input to take into account the error uncertainty caused by the wind components conversion in stations

4.1.2 Probability scores

Not only the summer ensemble, but also the winter ensemble exhibits a flat distribution, albeit there is a slightly cold bias in summer. Since the over–estimated temperature in the single first guess NWP derived a systematic negative error for T2m. Due to the impact of Gaussian–distributed noise, this cold bias retains in the perturbed error1 and conveys to the ensemble. This can also be observed in the reliability curve (Fig. 3).

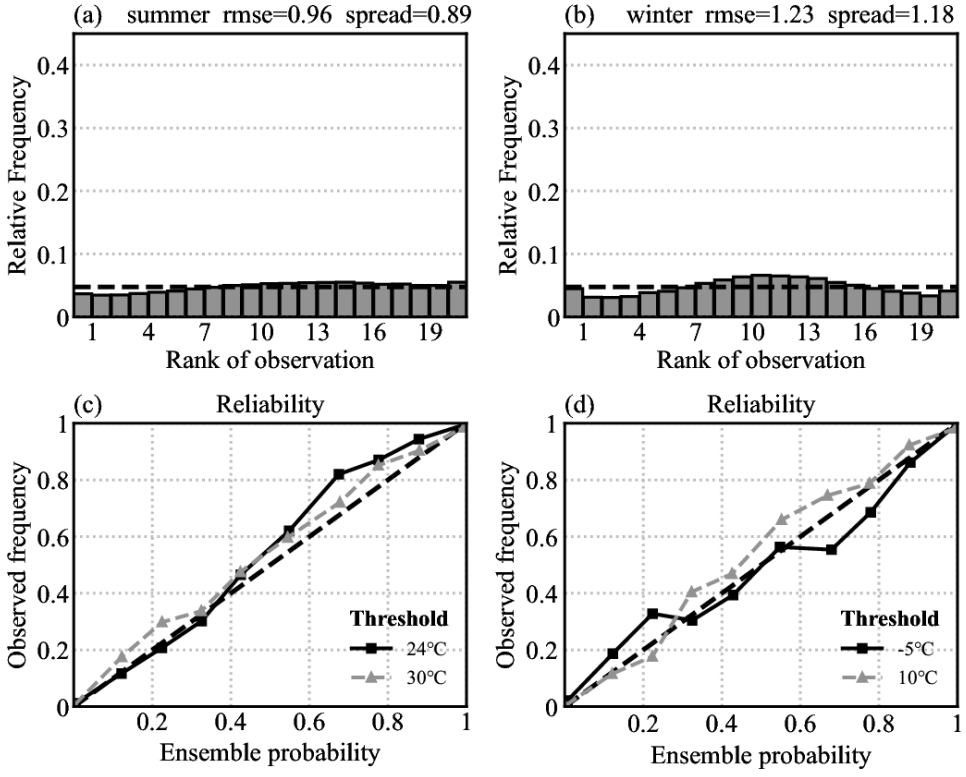


Fig. 3. Talagrand (a, b) and reliability diagrams (c, d) for T2m at analysis time, averaged over all test stations.

https://doi.org/10.5194/egusphere-2024-1554
Preprint. Discussion started: 28 June 2024

In summer the QQ2m ensemble has batter verification scores than that in winter, which is consistent with T2m (Fig. 4). Since in winter the more pronounced negative bias in T2m indicates a greater variability in 3D interpolation weights. The 3DRF of specific humidity, on the other hand, relies on the topographic characteristics of temperature NWP errors. Consequently, there is a slight increase in RMSE among some stations for the specific humidity ensemble in winter (Fig. S4).

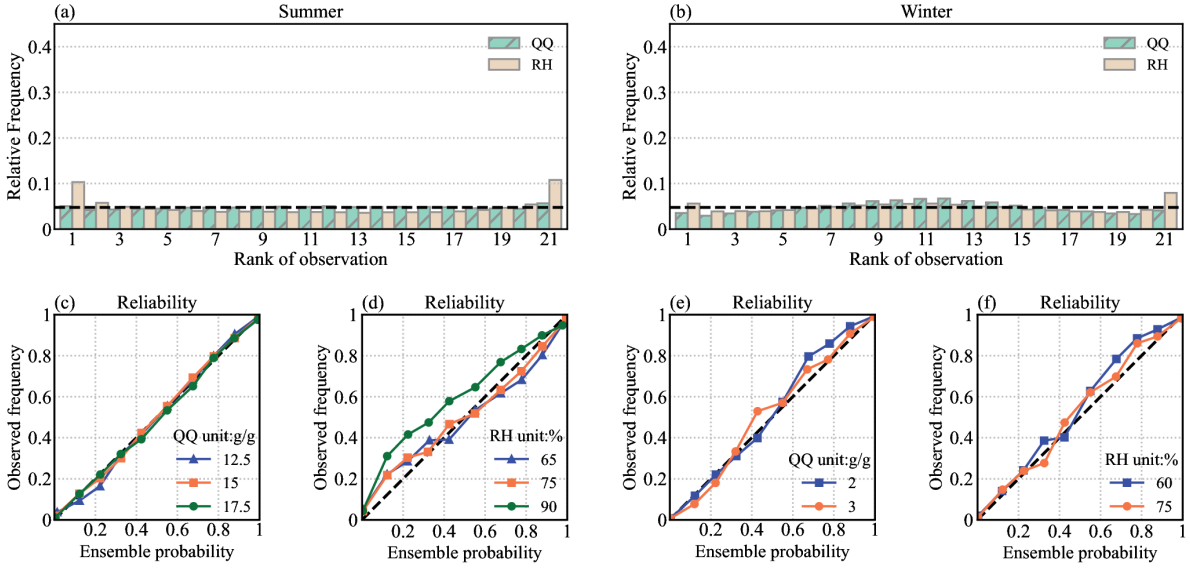
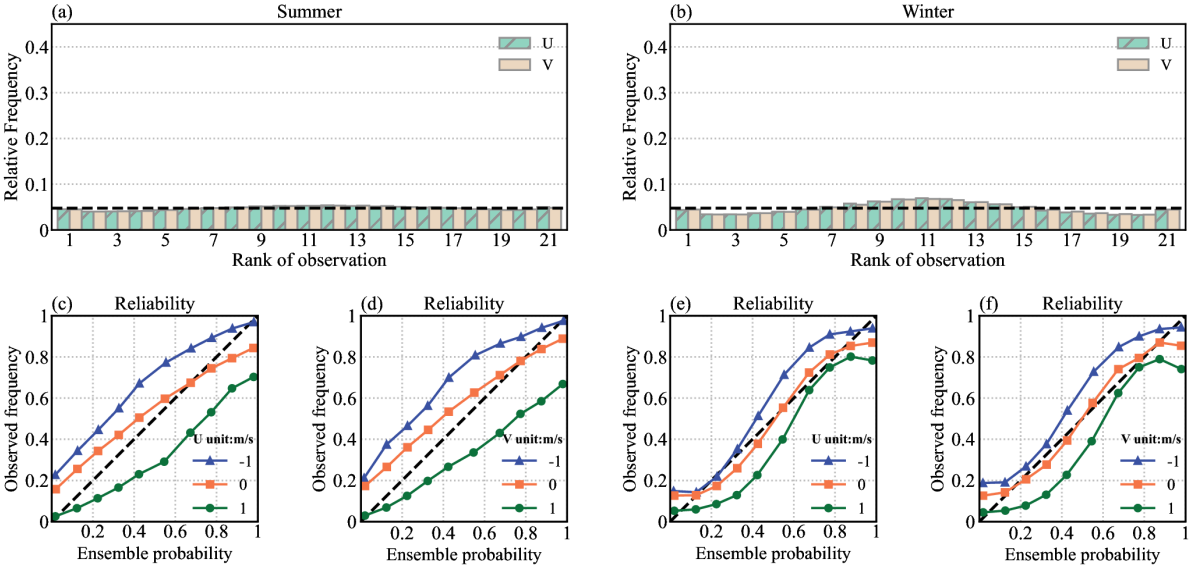


Fig. 4. Talagrand (a, b) and reliability diagrams (c ~ f) for QQ (unit: g g^{-1}) and RH (unit: %) at analysis time, averaged over all test stations.

The ensemble for QQ2m in summer has high resolution and reliability at different thresholds, as well as a flat Talagrand distribution, illustrating that the uncertainty has been estimated accurately (Fig. 4). The conversion of RH2m involves temperature, and in order to avoid the influence of temperature uncertainty on the variable conversion, only deterministic temperature is used in the humidity module. Therefore, in this study the interplay effects among different variables have not been taken into account. As a result, systematic biases in T2m within SIVA will propagate into RH2m, causing certain dry biases and moist biases (Fig. 4).

The bias for wind ensembles that can be observed in the reliability curve is primarily caused by the uncertainty representation in wind direction (Fig. 5c~f). Since in this study, we consider the zonal and meridional direction of U and V component. The positive value of U10m (V10m) means the westerly (southerly) wind, while negative value means easterly (northerly) wind. For example, when the U10m is +1 m s^{-1}, the ensemble values may show -1 m s^{-1}, leading to significant bias. The flat Talagrand distribution is due to the symmetry of wind speed about the value of 0 m s^{-1}. A probably resolution is to take into account the impact of wind direction in observation.

https://doi.org/10.5194/egusphere-2024-1554
Preprint. Discussion started: 28 June 2024



235 **Fig. 5. Talagrand (a, b) and reliability diagrams (c ~ f) for U10m and V10m (unit: m s^{-1}) at analysis time, averaged over all test stations.**

4.2 Verification of nowcasting

The verifications of ensemble analysis indicate the reliability of the error uncertainty estimation. The ensemble provides a
240 spread in analysis time and exhibit a reasonable error that is consistent with the deterministic analysis. To evaluate the transmission effect of ensemble spread in extrapolation, ensemble analysis is utilized to compute ensemble nowcasting. Because of the impact of value range limitation in variable conversion, direct computation of ensemble nowcasting for RH2m by using ensemble analysis is not feasible. For wind, SIVA assigns a weight to NWP based on an extrapolation step function. As the effectiveness of extrapolation method gradually diminishes with the prolongation of time steps, the weight
245 of uncertainty information carried by ensemble analysis will not be 1 when extrapolating up to +3 hours. Hence, the ensemble nowcasting is with maximum lead time 2 hours.

4.2.1 Ensemble BIAS and RMSE

Figure 6 shows that at the initial time the bias of ensemble mean is almost equal to that of deterministic nowcasting (hereafter called CNOW). Since the perturbation is Gaussian–distributed and the inflation of NWP perturbation is symmetric
250 about the mean. Meanwhile, the ensemble nowcasting shows a bias shift compared to the bias of CNOW. The slightly bias shift is due to the systematic bias of ensemble analysis.

https://doi.org/10.5194/egusphere-2024-1554
Preprint. Discussion started: 28 June 2024

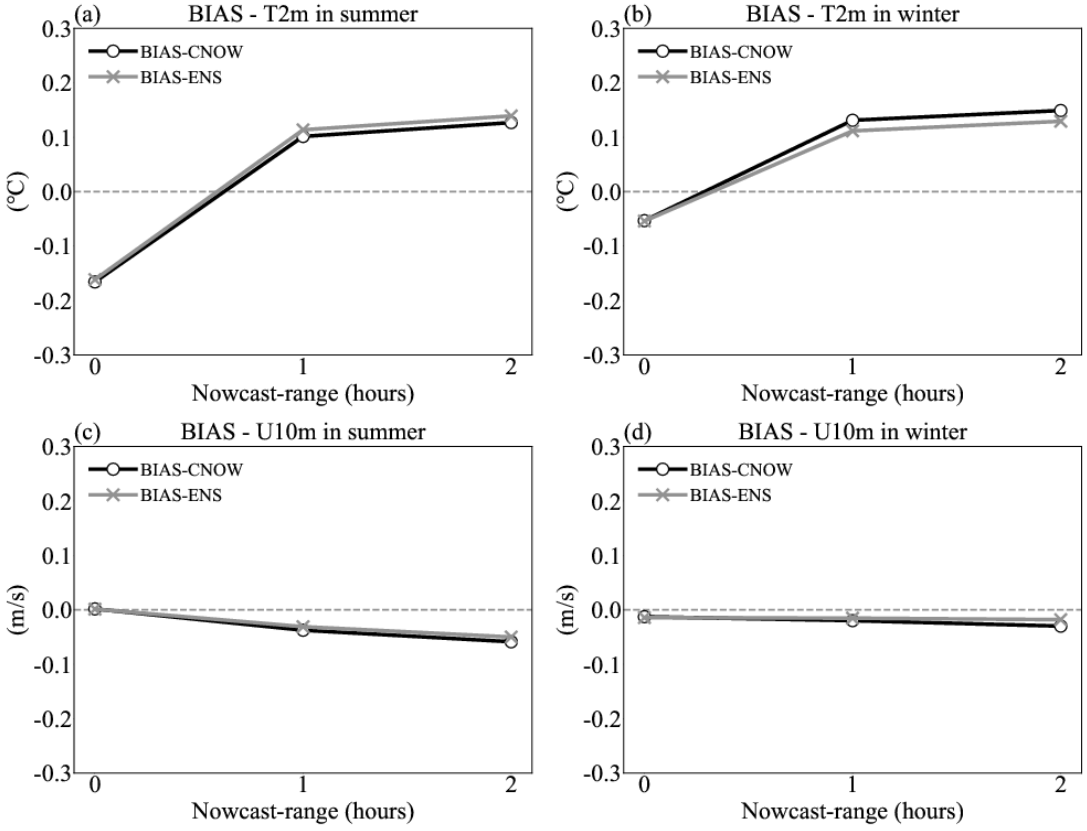


Fig. 6. Bias of deterministic nowcasting (black, CNOW, ○) and ensemble mean (grey, ENS, ×) for T2m (a, b) and U10m (c, d), averaged over all stations and whole period.

255

The RMSE and spread in figure 7 demonstrate that the analysis uncertainty can convey effectively through extrapolation. At the initial time, the variance of ensemble error, for test stations, is consistent with the variance of deterministic analysis error (Fig 2, 7). In the nowcasting range, spread can match to the stable growth trend of RMSE. We also calculate the ensemble nowcasting for specific humidity and due to the similar calculation framework of specific humidity and
260 temperature, the results are generally consistent. For wind components, the spread is consistent with RMSE both in analysis time and in nowcast time. This might be attributed to the estimation of the lowest model layer through divergence constraint after wind correction (3DRF), thereby avoiding dispersion attenuation caused by 2D interpolation. The error scores indicate that ensemble nowcasting has a reasonable spread that can be effectively transmitted without causing abnormal increment in error.

https://doi.org/10.5194/egusphere-2024-1554
Preprint. Discussion started: 28 June 2024

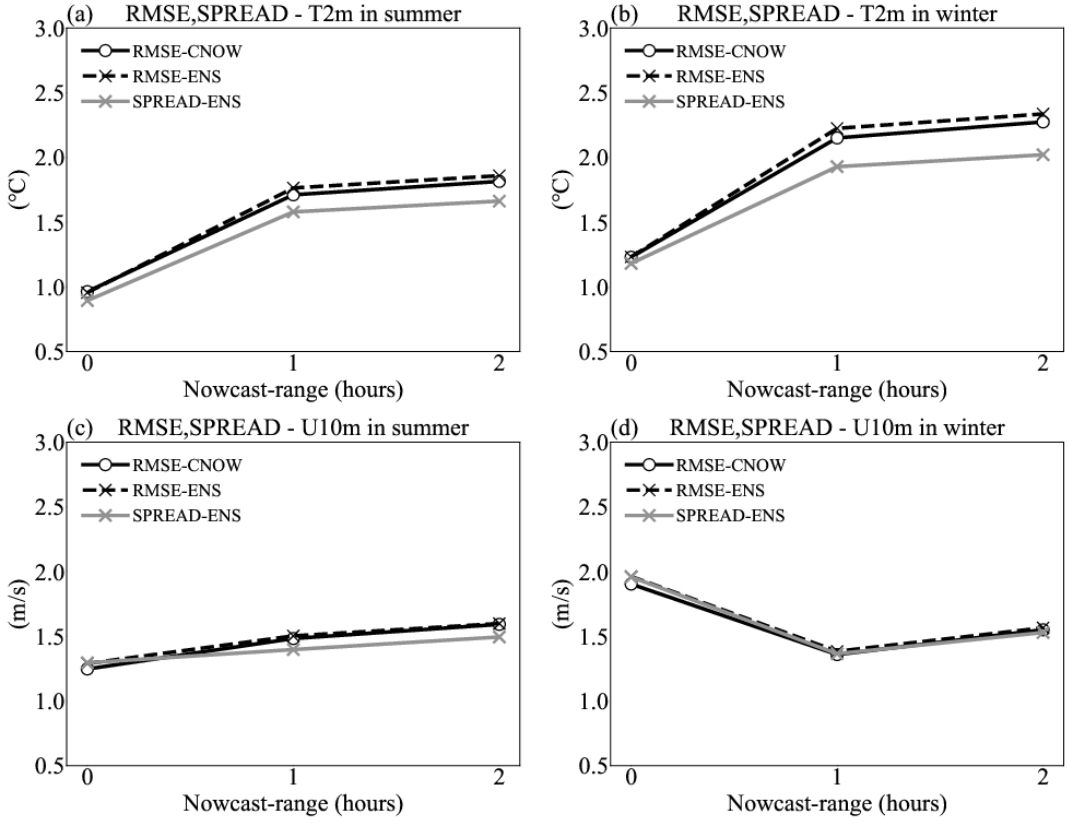


Fig. 7. RMSE of deterministic nowcasting (black, ○) and ensemble mean (black, ×), spread of ensemble nowcasting (grey, ×) for T2m (a, b) and U10m (c, d), averaged over all stations and whole period in summer and in winter.

4.2.2 Probability scores

To evaluate the probabilistic skill of the ensemble nowcasting, Talagrand and reliability diagrams for T2m and U10m were computed. Figure 8 shows the probabilistic skill for T2m. The Talagrand diagram shows a warm bias for T2m both in summer and in winter. As observed in the verification of T2m ensemble analysis, there exists systematic cold bias (Fig 3). The NWP at each time step will subtract the previous one to represent the predictive trend of variables, thereby converting the cold bias into a warm bias in the opposite site.

For reliability verification, the events chosen are the quartile temperature both in summer and winter in BTH region. The reliability diagram indicates that ensemble nowcasting is more reliable in summer than in winter. Since nowcasting in summer can better match to verification probabilities compare to nowcasting in winter. But there are under (over) nowcast of probabilities for warm (cold) events in winter. This is consistent with the reliability score for T2m ensemble analysis.

https://doi.org/10.5194/egusphere-2024-1554
Preprint. Discussion started: 28 June 2024

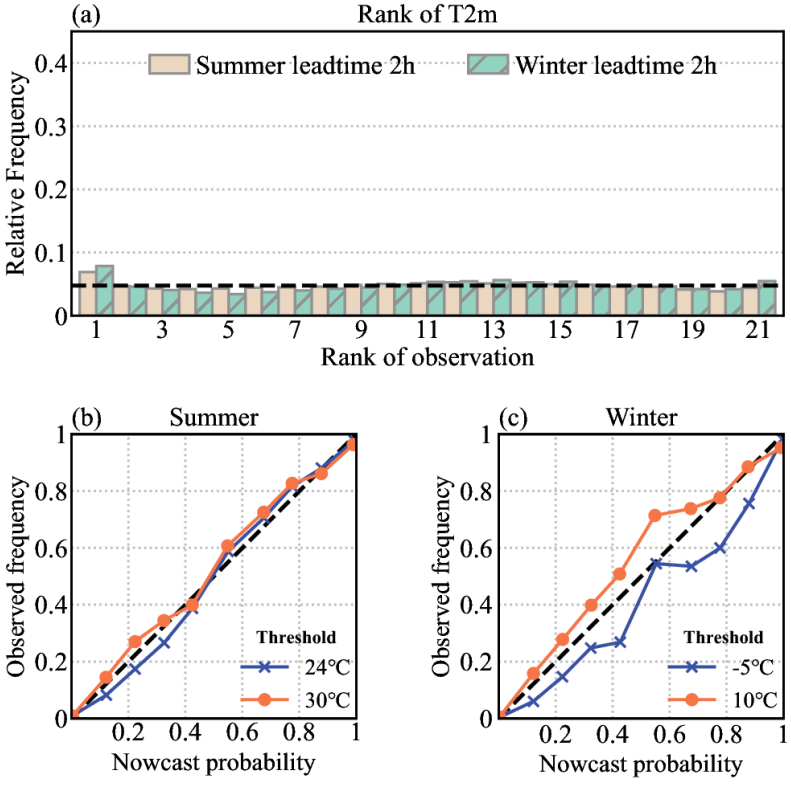


Fig. 8. Talagrand (a) and reliability diagrams (b, c) for the ensemble nowcasting at lead time of +2 hour of T2m (unit: ℃).

U10m and V10m have similar verification characteristic. The verification of U10m shows that ensemble nowcasting in summer has better reliability and better resolution than that in winter (Fig 9). The evident bias for summer U10m, as shown in the reliability curve (Fig 9b), indicates that ensemble lacks skill to represent the wind direction uncertainty during periods of significant wind variability. This is due to the lack of appropriate observational perturbations.

https://doi.org/10.5194/egusphere-2024-1554
Preprint. Discussion started: 28 June 2024

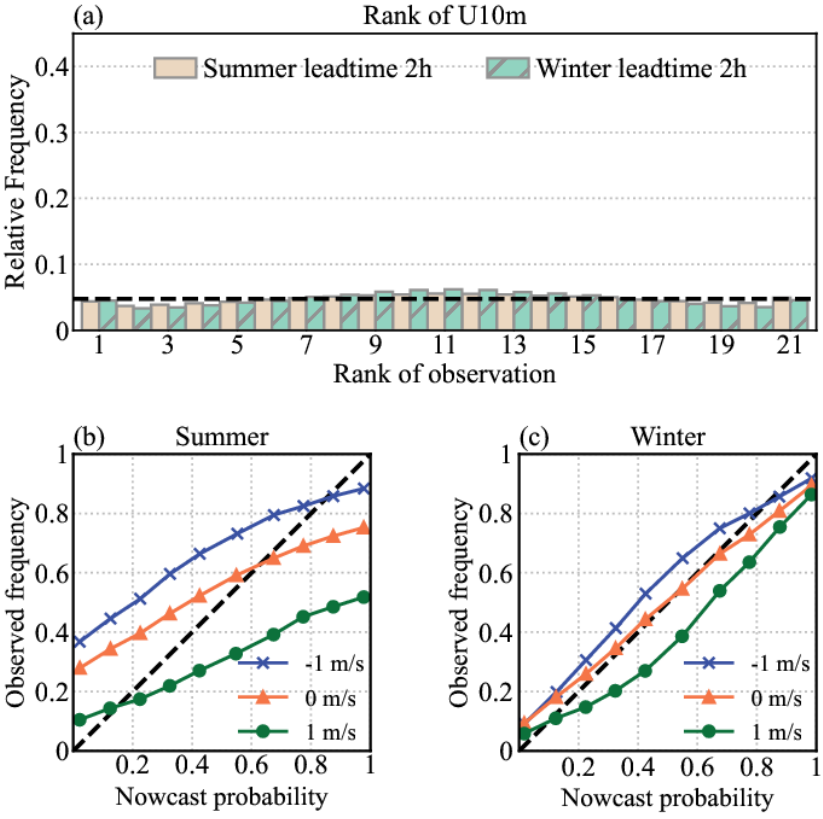


Fig. 9. Talagrand (a) and reliability diagrams (b, c) for the ensemble nowcasting at lead time of +2 hour of U10m (unit: m s^{-1}).

5 Conclusion and discussion

This study proposes an efficient perturbation method to quantify the uncertainty of near–surface analysis accurately. Generating Gaussian–distributed noises associated with the characteristic of the error between observation and first guess field can simulate the propagation of the error uncertainty within the analysis computation. An inflation factor is computed to simulate the attenuation of perturbation dispersion as it conveys through the interpolation process. The generated ensembles demonstrate reliability and capability for estimating the uncertainty of analysis error. The ensemble analysis provides a reasonable estimation of surface uncertainty for nowcasting at initial time, which can match to the error increment in nowcasting extrapolation.

Perturbing both observation and first guess field by noise reflects the dispersion of the first guess error, which is the main source of the analysis error uncertainty, especially in no–station areas. For temperature, the spatial uncertainty caused by geographic characteristics can be taken into account by combining the perturbed field with terrain information in the

https://doi.org/10.5194/egusphere-2024-1554
Preprint. Discussion started: 28 June 2024

interpolation. For humidity, the intrinsic relationship with temperature affects the estimation of the error uncertainty. For wind components, the variation of perturbation information in the divergence constraint is not considered, leading to a certain negative effect in RMSE, which could be a task to improve in future work.

Verifications of ensemble analysis for all variables illustrate the representation of analysis error uncertainty is reliable. But the wind components suffer the influence of divergence constraint. The perturbation simulates the error variation but lack contribute to correct the velocity value, resulting in negative effect in RMSE. Relative humidity is affected by the conversion of variables, leading an under–dispersive spread.

The flat Talagrand diagrams demonstrate that ensembles can estimate the probable range of true value in no–station area. The high resolution and reliability indicate that the ensembles can classify samples probability correctly. Perturbing the errors in analysis computation can accurately estimate the uncertainty for near–surface variables in both numerical value and spatial distribution.

The ensemble analysis provides reliable spread at initial time for nowcasting. Not only the errors of ensemble nowcasting can match to the errors of deterministic nowcasting, but also the growth of ensemble spread is consistent with the growth trend of error in nowcasting extrapolation. Since this method does not take into account the uncertainty derived by NWP systematic error and the relativity of different variables, the ensemble nowcasting are slightly under-dispersive.

The perturbed first guess errors, as well as the spread and error of resulting ensemble, are associated with the analysis error observed in test stations. The perturbation method in this study provides reference for the lack of means for accurately representing the uncertainty of near–surface deterministic analysis. This method can contribute to the estimation of near–surface analysis uncertainty in both nowcasting applications and ensemble nowcasting development. Further improvements can be made by considering the uncertainty of estimating the first guess error of multi–source NWP, in order to obtain a more comprehensive spatial uncertainty factor as well as the inflation factor. Taking the uncertainty of observation error uncertainty into account can be another priority in future development.

Code and data availability. The information data, example data and the corresponding codes for generating perturbation are archived on Zenodo: https://doi.org/10.5281/zenodo.11243716. (Zhu, 2024). Due to the confidentiality policy, the code and datasets of SIVA that utilized in this study are not in the public domain and cannot be distributed.

https://doi.org/10.5194/egusphere-2024-1554
Preprint. Discussion started: 28 June 2024

Appendix A: Case of generating temperature perturbation

Listing A1 demonstrates how to generate perturbation for both observation and NWP. This case is for 2m temperature while specific humidity and wind components have similar process. The input file sta_inf and grid_inf are the example temperature data and the corresponding standard deviation of first guess error. Due to the confidentiality agreement, these files only include the temperature value. The variable inflation in grid_inf is the inflation factor which is used to rescale the perturbed NWP. Since the standard deviation of perturbation that generated by the function numpy.random.normal exists offsets, the factor sc_sta and sc_grid are used to ensure the scale of perturbation be consistent with the scale of first guess error.

https://doi.org/10.5194/egusphere-2024-1554
Preprint. Discussion started: 28 June 2024

```python
import xarray as xr
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

member_number = 10                                  # Number of ensemble members
nz = 21                                             # Number of vertical levels
nj = 581                                            # Number of grids in Latitudinal direction
ni = 511                                            # Number of grids in Meridian

# Read the information for generating random noise
sta_inf = pd.read_csv('./observation_data.csv')     # Information in station

grid_inf = xr.open_dataset('./grid_information.nc')  # Information in grid
grid_std = grid_inf.error_std_grid.values           # Standard deviation(STD) of first guess error in grid
nwp_data = grid_inf.nwp.values                      # Example of NWP data
inflation = grid_inf.inflation.values               # Inflation factor

########################################################
# Generating random noise at each station that use to calculate analysis and nowcasting
sta_noise = np.empty([member_number, len(sta_inf)])
for s_num in range(len(sta_inf)):
    gen_noise_grid = np.random.normal(loc=0, scale=sta_inf['error_std'][s_num], size=member_number)
    # Rescale the standard deviation (STD) of noise
    # to be consistent with the STD of first guess error in each station
    sc_sta = np.std(gen_noise_grid) / sta_inf['error_std'][s_num]
    sta_noise[:, s_num] = (gen_noise_grid - np.nanmean(gen_noise_grid)) / sc_sta

# Generating random noise at each grid point
nwp_noise = np.empty([member_number, nz, nj, ni])
for j in range(nj):
    for i in range(ni):
        gen_noise_grid = np.random.normal(loc=0, scale=grid_std[j, i], size=[member_number, nz])
        sc_grid = np.std(gen_noise_grid) / grid_std[j, i]
        nwp_noise[:, :, j, i] = (gen_noise_grid - np.nanmean(gen_noise_grid)) / sc_grid

# Perturbing NWP
per_nwp = np.empty([member_number, nz, nj, ni])
for x in range(member_number):
    per_nwp[x, :, :, :] = nwp_data + nwp_noise[x, :, :, :]
per_mean = np.nanmean(per_nwp, axis=0)

for x in range(1, member_number + 1):
    # Perturbing observation
    sta_inf['tt'] += sta_noise[x - 1]

    # Inflating the perturbed NWP
    new_nwp = (np.einsum('kij,ij->kij', (per_nwp[x - 1] - per_mean), inflation)) + per_mean

    plt.imshow(new_nwp[0], cmap='jet', origin='lower')
    plt.show()
```

Listing A1. Process of generating random perturbation and rescaling the perturbed NWP by inflation factor.

https://doi.org/10.5194/egusphere-2024-1554
Preprint. Discussion started: 28 June 2024

Author contributions. Yanwei Zhu and Yong Wang equally contributed to this work. Yong Wang and Aitor Atencia proposed the method; Yanwei Zhu applied the method and perform the experiments; Yanwei Zhu wrote the manuscript draft and all authors reviewed and edited the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We are grateful to Huafeng Meteorological Media Group and the HuaFeng Research Lab for Weather Science and Applications, Nanjing University of Information Science & Technology for their support and technical assistance in this research. The authors express gratitude for funding the three projects on "HUAFENG Forecast Applications" with Grant No.CY-J2020007, CY-J2021002 and CY-2022ZDA01.

References

Bellus, M., Wang, Y., and Meier, F.: Perturbing Surface Initial Conditions in a Regional Ensemble Prediction System, Mon. Wea. Rev., 144, 3377–3390, https://doi.org/10.1175/MWR-D-16-0038.1, 2016.

Bellus, M., Weidle, F., Wittmann, C., Wang, Y., Tasku, S., and Tudor, M.: Aire Limitée Adaptation dynamique Développement InterNational-Limited Area Ensemble Forecasting (ALADIN-LAEF), Adv. Sci. Res., 16, 63–68, https://doi.org/10.5194/asr-16-63-2019, 2019.

Bouttier, F.: The ensemble forecasting. Encyclopedia of the Environment, Accessed 13 March 2024, https://www.encyclopedie-environnement.org/en/air-en/overall-forecast/, 2019.

Bouttier, F., Raynaud, L., Nuissier, O., and Ménétrier, B.: Sensitivity of the AROME ensemble to initial and surface perturbations during HyMeX, Q. J. R. Meteorol. Soc., 142, 390–403, https://doi.org/10.1002/qj.2622, 2016.

Buizza, R., Houtekamer, P. L., Pellerin, G., Toth, Z., Zhu, Y.-J., and Wei, M.-Z.: A Comparison of the ECMWF, MSC, and NCEP Global Ensemble Prediction Systems, Mon. Wea. Rev., 133, 1076–1097, https://doi.org/10.1175/MWR2905.1, 2005.

Casellas, E., Bech, J., Veciana, R., Pineda, N., Miró, J., Moré, J., Rigo, T., and Sairoui, A.: Nowcasting the precipitation phase combining weather radar data, surface observations and NWP model forecasts, Q. J. R. Meteorol. Soc., 147, 3135–3153, https://doi.org/10.1002/qj.4121, 2021.

Chen, X.-C., Zhao, K., Sun, J.-Z., Zhou, B.-W., and Lee, W.-C.: Assimilating Surface Observations in a Four-Dimensional Variational Doppler Radar Data Assimilation System to Improve the Analysis and Forecast of a Squall Line Case, Adv. Atmos. Sci., 33, 1106–1119, https://doi.org/10.1007/s00376-016-5290-0, 2016.

Eibl, B. and Steinacker, R.: Treatment of deterministic perturbations and stochastic processes within a quality control scheme, Geosci. Instrum. Method. Data Syst. Discuss. [Perprint], https://doi.org/10.5194/gi-2017-42, 2017.

https://doi.org/10.5194/egusphere-2024-1554
Preprint. Discussion started: 28 June 2024

Keresturi, E., Wang, Y., Meier, F., Weidle, F., Wittmann, C., and Atencia, A.: Improving initial condition perturbations in a convection-permitting ensemble prediction system, Q. J. R. Meteorol. Soc., 145, 993–1012, https://doi.org/10.1002/qj.3473, 2019.

Feng, J., Toth, Z., Peña, M., and Zhang, J.: Partition of Analysis and Forecast Error Variance into Growing and Decaying Components, Q. J. R. Meteorol. Soc., 146, 1302–1321, https://doi.org/10.1002/qj.3738, 2020.

Glahn, B. and Im, J.-S.: Error estimation of objective analysis of surface observations, J. Operational Meteorol., 1, 114–127, https://doi.org/10.15191/nwajom.2013.0111, 2013.

Haiden, T., Kann, A., Pistotnik, G., Stadlbacher, K., and Wittmann, C.: Integrated Nowcasting through Comprehensive Analysis (INCA) System description, ZAMG report, 61 pp, http://www.zamg.ac.at/fix/INCA_system.pdf, 2010

Haiden, T., Kann, A., Wittmann, C., Pistotnik, G., Bica, B., and Gruber, C.: The Integrated Nowcasting through Comprehensive Analysis (INCA) System and Its Validation over the Eastern Alpine Region, Wea. Forecasting, 26, 166–183, https://doi.org/10.1175/2010WAF2222451.1, 2011.

Horányi, A., Mile, M., and Szucs, M.: Latest developments around the ALADIN operational short-range ensemble prediction system in Hungary, Tellus A, 63, 642–651, https://doi.org/10.1111/j.1600-0870.2011.00518.x, 2011.

Hoteit, I., Pham, D.-T., Gharamti, M. E., and Luo, X.: Mitigating Observation Perturbation Sampling Errors in the Stochastic EnKF, Mon. Wea. Rev., 143, 2918–2936, https://doi.org/10.1175/MWR-D-14-00088.1, 2015.

Kann, A., Wittmann, C., Wang, Y., and Ma, X.-L.: Calibrating 2-m Temperature of Limited-Area Ensemble Forecasts Using High-Resolution Analysis, Mon. Wea. Rev., 137, 3373–3387, https://doi.org/10.1175/2009MWR2793.1, 2009.

Kann, A., Pistotnik, G., and Bica, B.: INCA-CE: a Central European initiative in nowcasting severe weather and its applications, Adv. Sci. Res., 8, 67–75, https://doi.org/10.5194/asr-8-67-2012, 2012.

Leith, C. E.: Theoretical Skill of Monte Carlo Forecasts, Mon. Wea. Rev., 102, 409–418, https://doi.org/10.1175/1520-0493(1974)102<0409:TSOMCF>2.0.CO;2, 1974.

Leutbecher, M., Buizza, R., and Isaksen, L.: Ensemble forecasting and flow-dependent estimates of initial uncertainty. ECMWF Workshop on Flow-Dependent Aspects of Data Assimilation, Reading, United Kingdom, ECMWF, 185–201. https://www.ecmwf.int/sites/default/files/elibrary/2007/10731-ensemble-forecasting-and-flow-dependent-estimates-initial-uncertainty.pdf, 2007.

Leutbecher, M. and Palmer, T. N.: Ensemble forecasting, J. Comput. Phys., 227, 3515–3539, https://doi.org/10.1016/j.jcp.2007.02.014, 2008.

Lin, X.-X., Feng, Y.-R., Xu, D.-S., Jian, Y.-T., Huang, F., and Huang, J.-C.: Improving the Nowcasting of Strong Convection by Assimilating Both Wind and Reflectivity Observations of Phased Array Radar: A Case Study, J. Meteorol. Res., 36, 61–78, https://doi.org/10.1007/s13351-022-1034-5, 2022.

Lorenz, E. N.: A study of the predictability of a 28-variable atmospheric model, Tellus A, 17, 321–333, https://doi.org/10.3402/tellusa.v17i3.9076, 1965.

https://doi.org/10.5194/egusphere-2024-1554
Preprint. Discussion started: 28 June 2024

Randriamampianina, R. and Storto, A.: ALADIN-HARMONIE/Norway and its assimilation system – the implementation phase, HIRLAM Newsletter, 54, 20–30, https://api.semanticscholar.org/CorpusID:59364871, 2008.

Saetra, Ø., Hersbach, H., Bidlot, J., and Richardson, D. S.: Effects of Observation Errors on the Statistics for Ensemble Spread and Reliability, Mon. Wea. Rev., 132, 1487–1501, https://doi.org/10.1175/1520-0493(2004)132<1487:EOOEOT>2.0.CO;2, 2004.

Schmid, F., Bañon, L., Agersten, S., Atencia, A., Coning, E., Kann, A., Wang, Y., and Wapler, K.: Conference Report: Third European Nowcasting Conference, Meteorol. Z., 28, 447–450, https://doi.org/10.1127/metz/2019/0983, 2019.

Suklitsch, M., Kann, A., and Bica, B.: Towards an integrated probabilistic nowcasting system (En-INCA), Adv. Sci. Res., 12, 51–55, https://doi.org/10.5194/asr-12-51-2015, 2015.

Sun, J.-Z., Xue, M., Wilson, J., Zawadzki, I., Ballard, S., Onvlee-Hooimeyer, J., Joe, P., Barker, D., Li, P., Golding, B., et al.: Use of NWP for Nowcasting Convective Precipitation: Recent Progress and Challenges, Bull. Amer. Meteor. Soc., 95, 409–426, https://doi.org/10.1175/BAMS-D-11-00263.1, 2014.

Taylor, C., Klein, C., Dione, C., Parker, D., Marsham, J., Diop, C., Fletcher, J., Chaibou, A., Nafissa, D., and Semeena, V.: Nowcasting Tracks of Severe Convective Storms in West Africa from Observations of Land Surface State, Environ. Res. Lett., 17, 034016, https://doi.org/10.1088/1748-9326/ac536d, 2022.

Thiruvengadam, P., Indu, J., and Ghosh, S.: Significance of 4DVAR Radar Data Assimilation in Weather Research and Forecast Model-Based Nowcasting System, J. Geophys. Res.: Atmos., 125, e2019JD031369, https://doi.org/10.1029/2019JD031369, 2020.

Wang, J.-Z., Chen, J., Zhang, H.-B., Tian, H., and Shi, Y.-N.: Initial Perturbations Based on Ensemble Transform Kalman Filter with Rescaling Method for Ensemble Forecasting, Wea. Forecasting, 36, 823–842, https://doi.org/10.1175/WAF-D-20-0176.1, 2021.

Wang, Y., Bellus, M., Wittmann, C., Steinheimer, M., Weidle, F., Kann, A., Ivatek-Šahdan, S., Tian, W.-H., Ma, X.-L., Tascu, S., Bazile, E.: The Central European limited-area ensemble forecasting system: ALADIN–LAEF, Q. J. R. Meteorol. Soc., 137, 483–502, https://doi.org/10.1002/qj.751, 2011.

Wang, Y., Coning, E., Harou, A., Jacobs, W., Joe, P., Nikitina, L., Roberts, R., Wang, J.-J., Wilson, J., Atencia, A., Bica, B., Brown, B., Goodmann, S., Kann, A., Li, P., Monterio, I., Schmid, F., Seed, A., Sun, J.: Guidelines for Nowcasting Techniques, World Meteorological Organization, 82pp, https://library.wmo.int/doc_num.php?explnum_id=3795, 2017a.

Wang, Y., Meirold-Mautner, I., Kann, A., Šajn Slak, A., Simon, A., Vivoda, J., Bica, B., Böcskör, E., Brezková, L., and Dantinger, J.: Integrating nowcasting with crisis management and risk prevention in a transnational and interdisciplinary framework, Meteorol. Z., 26, 459–473, https://doi.org/10.1127/metz/2017/0843, 2017b.

Wang, Y., Bellus, M., Geleyn, J., Ma, X.-L., Tian, W.-H., and Weidle, F.: A New Method for Generating Initial Condition Perturbations in a Regional Ensemble Prediction System: Blending, Mon. Wea. Rev., 142, 2043–2059, https://doi.org/10.1175/MWR-D-12-00354.1, 2014.

https://doi.org/10.5194/egusphere-2024-1554
Preprint. Discussion started: 28 June 2024

Wastl, C., Simon, A., Wang, Y., Kulmer, M., Baár, P., Bölöni, G., Dantinger, J., Ehrlich, A., Fischer, A., and Heizler, Z.: A seamless probabilistic forecasting system for decision making in Civil Protection, Meteorol. Z., 27, 417–430, https://doi.org/10.1127/metz/2018/902, 2018.

Wastl, C., Wang, Y., Atencia, A., and Wittmann, C.: Independent perturbations for physics parametrization tendencies in a convection–permitting ensemble (pSPPT), Geosci. Model Dev., 12, 261–273, https://doi.org/10.5194/gmd-12-261-2019, 2019.

Wastl, C., Wang, Y., Atencia, A., Weidle, F., Wittmann, C., Zingerle, C., Keresturi, E.: C-LAEF: Convection-permitting Limited-Area Ensemble Forecasting system, Q. J. R. Meteorol. Soc., 147, 1431–1451, https://doi.org/10.1002/qj.3986, 2021.

Yang, L., Cheng, C.-L., Xia, Y., Chen, M., Chen, M.-X., Zhang, H.-B., and Huang, X.-Y.: Evaluation of the added value of probabilistic nowcasting ensemble forecasts on regional ensemble forecasts, Adv. Atmos. Sci., 40, 937−951, https://doi.org/10.1007/s00376-022-2056-8, 2023.

Zhu, Y.-J.: Ensemble forecast: A new approach to uncertainty and predictability, Adv. Atmos. Sci., 22, 781–788, https://doi.org/10.1007/BF02918678, 2005.

Zhu, Y.-W.: Information for generating perturbation for ensemble analysis and ensemble nowcasting, Zenodo [code and data set], https://doi.org/10.5281/zenodo.11243716. 2024.