Article ID: 1006-8775(2018) 03-0334-12

EXPERIMENTAL APPLICATION OF A THINNING SCHEME FOR THE ASSIMILATION OF SURFACE OBSERVATION DATA IN GRAPES–3DVAR

ZHAO Hong (赵 虹)[1], LIU Yin (刘 寅)[2, 3]
(1. Luhe Meteorological Bureau, Nanjing 211599 China; 2. Jiangsu Meteorological Observation Center, Nanjing 210009 China; 3. Jiangsu Institute of Meteorological Sciences, Nanjing 210009 China)

Abstract: To reduce the spatial correlation of representation error in observations and computational complexity, we propose a thinning scheme that can extract typical observations within a certain range. This scheme is applied to the Global/Regional Assimilation and Prediction System (GRAPES) with three-dimensional variation (3DVAR) to study the effect of the thinning radius on the assimilation results. The assimilation experiments indicate that when the ratio of the model resolution to the observational resolution is 1:3, the simulated results for precipitation are relatively good and have a relatively high equitable threat score (ETS). Moreover, the analysis errors in the temperature and the specific humidity are the smallest, the dependence of the norm gradient vector of the objective function on the number of iterations is slow, gentle, and close to 0, and the minimization results in improved conditions.

Key words: thinning; typical observation; data assimilation
CLC number: P456 **Document code:** A
doi: 10.16555/j.1006-8775.2018.03.007

1 INTRODUCTION

The variation method has become the most common assimilation method (Lewis and Derber [1]; Dimet and Talagrand[2]; Talagrand and Courtier[3]). The basic idea of three-dimensional variation (3DVAR) is to provide an optimal estimate of atmospheric data based on observed forecast fields and to express the problem as one that minimizes the objective function (Lorenc[4]), namely, minimization of the quadratic functional between the analytical field and the observation field and between the analytical field and the background field. The objective function can be written as

$$J(x)=J_b+J_0=\frac{1}{2}[x-x_b]^T B^{-1}[x-x_b]+\frac{1}{2}[H(x)-y_0]^T(O+F)^{-1}[H(x)-y_0] \quad (1)$$

Here, J_b is called the background term, and J_0 is the observation term. x represents the atmospheric state, x_b represents the background field, y_0 contains the known observation data, B is the error covariance matrix of the background field, O is the error covariance matrix of the observational data, H is the observation operator, and F is the representative error covariance matrix of the observation operator H.

During the process of assimilation, an optimal analysis is conducted based on the statistical knowledge of the background error and observation error. The observation error can be divided into two components: the instrumental error of a measuring device and the representativeness error of an observation operator (Daley[5]; Schwartz and Benjamin[6]). The former is often considered to be white Gaussian noise and usually derived from instruments of the same specifications or measurements of the same variable (Richner and Phillips[7]). The latter is generated when the observation operator is converted from the model space to the observation space. It is thought to be responsible for spatial correlations in the observational error; it is difficult to estimate and expensive to specify in the minimization of the cost function (Lorenc[4]; Daley[5]). Liu and Rabier demonstrated in a simple one-dimensional (1D) framework that the representativeness error depends on a resolution function of the measuring instrument, observation density, model grid resolution, and specification of an observation operator[8]. For observations with spatially uncorrelated errors, increasing the observation density generally improves the analysis accuracy. However, for observations with spatially correlated errors, increasing the observation density beyond a threshold value will yield little or no improvement in the analysis accuracy. Bondarenko et al. expanded these studies to two-dimensional space and found that the observation error correlation coefficients should not exceed 0.3, which can be ignored in the assimilation without

Received 2017-10-11; **Revised** 2018-06-15; **Accepted** 2018-08-15
Foundation item: Young Meteorological Research of Jiangsu Provincial Meteorological Bureau (Q201611); Key Scientific Research Projects of Jiangsu Provincial Meteorological Bureau (KZ201605); Natural Science Foundation of Jiangsu Province (BK20161074); Beijige Fund of Jiangsu Institute of Meteorological Sciences (BJG201512)
Biography: ZHAO Hong, Ph. D., primarily undertaking research on surface data assimilation.
Corresponding author: LIU Yin, e-mail: liuyin200421@163.com

deteriorating the analysis quality[9]. However, the above studies are purely statistical and can only provide general guidelines because they do not target at real atmospheric situations. An observation may be representative of the atmospheric state over a wide area under stable weather conditions; however, this is not the case during stormy conditions. The situation in a real practical 3-D NWP (Numerical Weather Prediction) model is more sophisticated and requires further study.

Observation error correlations are difficult to estimate and expensive to specify in the assimilation process in operational practice, so the optimal analysis is usually not achievable. As a consequence, a suboptimal assimilation scheme is always used, in which the observation errors are assumed to be uncorrelated, thus, the observations with strongly correlated errors must be filtered out before assimilation. The error-decorrelation operation is called observation thinning, which is an efficient way to reduce effective error-correlation (Järvinen and Undén[10]).

On the other hand, high spatial density and temporal density of observations are potentially valuable for estimating an initial state in an operational data assimilation system. However, assimilating such a large number of observational datasets increases the computational costs, takes up more disk space, and leads to more time-consuming data transmission. Moreover, the high spatial and temporal data density can violate the assumption of independent observation errors in the operational and experimental assimilation schemes (Ochotta et al.[11]). In variational assimilation systems, the error correlations are unknown a priori, accounting for estimations of these correlations would require more complex observation error statistics, leading to an additional increase in computational costs. Therefore, we need to reduce the amount of data and extract the essential information.

In operational NWP, some scholars have proposed other thinning methods and applied them to intensive data. The simplest method may be the subsampling thinning approach, which systematically retains a few observations (Bradley[12]). Ramachandran proposed intelligent data thinning (IDT) (Ramachandran et al.[13]; Ramachandran et al.[14]), and Zavodsky et al.[15] found that the IDT algorithm resulted in the lowest RMS error over the coastal region where meteorological gradients are common, although it performed poorly over homogeneous regions. Ochotta et al.[11] discussed two thinning algorithms, which were devised and implemented for the thinning of Advanced TIROS Operational Vertical Sounder (ATOVS) satellite data. The first was called top-down clustering, which groups observations with similar spatial positions and measurement values into clusters that can be approximated by one representative measurement, and the second thins data through an estimation polynomial function and removes redundant measurements. The results showed that the two methods have their own advantages and disadvantages.

The above methods have been widely applied for the thinning of satellite data (Qin et al.[16]), but for surface observations, thinning methods have rarely been studied. This paper explores a new thinning scheme for real situations in heavy rainfall cases using a practical 3-D NWP model.

This paper is organized as follows. In the next section, we discuss the data and experiment design. Then, we introduce the thinning method and its application in detail. In sections 3 and 4, we present experimental results and investigations. Finally, we summarize our findings and address future work.

2 DATA AND METHODS

2.1 *Data*

In January 2008, frequent blizzards occurred in most regions of China, causing serious low-temperature hazards, such as rain, snow, and ice. In this study, because we only assimilate surface pressure in our experiments, the variable in the state vector x and the observation vector y_o in equation (1) is the surface pressure with one level. The pressure (y_o) data derive from surface synoptic observations (SYNOP) between 0000 UTC on January 1, 2008, and 1800 UTC on January 30, 2008, and the temporal resolution is three hours. There are 3,340 observation stations in the research region 0–70°N, 55–145°E.

Because the observation elements are continuous in space, the observations used in this study are highly correlated within a certain range. As shown in Fig.1, by calculating the distance between two observation stations and the corresponding surface pressure correlation coefficient, the surface pressure is very similar between the two stations, indicating strong correlations. We typically assume the observation error to be uncorrelated and filter out the observations with strongly correlated errors to meet the assumption as much as possible. However, in operational practice, the observation error correlations are difficult to estimate and expensive to specify in the assimilation process. We believe that thinning the observations can reduce the observations with strongly correlated errors and meet the assumption to some extent. In this paper, we apply the regional model of the Global/Regional Assimilation and Prediction System (GRAPES) to examine the thinning and assimilation of observation data; the assimilated variable is the surface pressure.

2.2 *Experiment design*

"GRAPES" is a type of new generation numerical forecasting system, developed independently by Chinese scientists (Wang et al.[17]). The GRAPES regional mesoscale numerical forecasting system (GRAPES_Meso) is an important part of the basic GRAPES system. It has been continuously improved according to the weather characteristics of China since

its implementation for operational application in the National Meteorological Center of China in 2006. The GRAPES_Meso model is currently used by many researchers (Wang et al.[18]; Yang and Shen[19]; Liu et al.[20]). The GRAPES_Meso model (along with its assimilation system) used in this paper is V3.1.0.1.

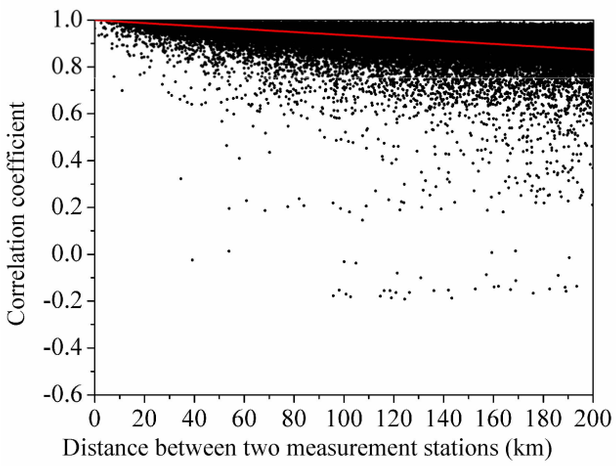


Figure 1. Scatter plot of the distance between two observation stations and their correlation coefficient (the red straight line represents the fitted line).

In the GRAPES model, the observation operator H can be written as follows:

$$H=H_{qc}H_pH_s$$

Here, H_s is the spatial interpolation operator, which interpolates variables from grid points to observation locations; bilinear interpolation and cubic spline interpolation are used in the horizontal and vertical dimensions, respectively. H_p represents a physical transformation operation, which converts the model variables to observation variables. When the observation variable is pressure, the physical transformation in the observation operator converts the dimensionless pressure to pressure. H_{qc} indicates the quality control based on observations and simulated observations.

Because we assimilate only surface pressure in our experiments, the variable in the state vector x is the surface pressure with one level. The pressure data are obtained from SYNOP between 0000 UTC on January 1, 2008, and 1800 UTC on January 30, 2008, and the temporal resolution is three hours. There are 3,340 observation stations in the research region 0–70°N, 55–145°E. The variable in the observation vector y_o is the only surface pressure at the stations.

There are 301 × 241 grid points in the numerical simulation. The coordinate center is (111° E, 29° N), which is located in Changde, Hunan, China. There are 49 levels in the vertical dimension and the top pressure is 35 hPa. The model uses the WDM-6 microphysical scheme, the Grell-3 cumulus parameterization scheme, the MYJ boundary layer scheme, the Dudhia shortwave radiation scheme, the RRTM longwave radiation scheme and the Noah land surface process scheme.

The assimilation experiments are designed as follows:

(1) Using Global Forecast System (GFS) data as the initial field and boundary conditions, the experiments are first integrated from 1800 UTC on January 26 to 0000 UTC on January 27. Next, the results of 0000 UTC on January 27 from the GRAPES model are used as the background data, and we assimilate surface pressure at this time; by using this analogy, we assimilate surface data five times at six-hour intervals. Lastly, the experiments are integrated for 24 hours to obtain the final numerical results. Fig.2 shows the assimilation process:

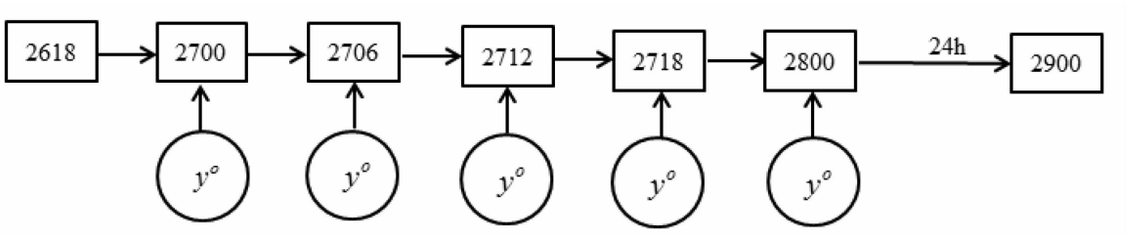


Figure 2. Flowchart for the assimilation process. y^o indicates surface observations.

(2) To study the effect of the thinning radius on assimilation, six different radii are used. y^o in each experiment indicates the observations with different thinning radii.

(3) The background error covariance B is obtained using the NMC (National Meteorological Center) method (Wang et al.[21]), and the observed error covariance is the default value from the model.

2.3 *Thinning scheme*

The subsampling thinning approach chooses data every few observations; thus, the data retained by this method is random and may not represent the characteristics of the removed data. As mentioned above, the IDT method performed well over the coastal region and poorly over homogeneous regions. Surface stations in the eastern part of China are densely

distributed and relatively uniform, so the IDT method is not suitable for surface observations used in this paper. While top-down clustering thinning and polynomial function thinning retain data that are different from surrounding observations, those data cannot effectively represent the variable characteristics of the study area. Thus, it is necessary to design a new thinning scheme for surface observations. The thinning scheme should be suitable for any area, including coastal regions and homogeneous regions, and more importantly, the data retained by the thinning scheme should represent the variable characteristics of the study area.

Due to the high correlation between two adjacent measurement stations, the scheme adopted for thinning the observation data is as follows.

(1) With the model grid as the center and r as the radius, calculate the temporal correlation coefficient between each observation station and the other stations within this radius.

(2) Calculate the average correlation coefficient between each observation station within this radius and all nearby stations.

(3) Select the observation station with the largest average correlation coefficient within this radius.

The circles in the thinning scheme overlap. If an observation is located in the overlap of the circles, we retain this observation for only the overlapping circles. The purpose of this operation is to select typical representative observations within the radius; a schematic of the grid points and the thinning radius is shown in Fig.3.

The model resolution selected in this experiment is 0.3°, namely, dx=dy ≈30 km, and the values of the radius, r, are 30, 60, 90, 120, and 150 km. Fig.4 shows the distribution of observation stations for different thinning radii, and Fig.4a shows the actual spatial distribution of the surface observation stations. The observation stations are densely distributed in the eastern part of China, and the horizontal spacing is approximately 20–30 km. According to Figs.4b–4f, as the thinning radius becomes larger, the distribution of stations becomes sparser. The numbers of observations are 2,804, 2,096, 1,552, 1,271, and 1,062 for r=30, 60, 90, 120, and 150, respectively. Moreover, in the eastern part of China, the stations are essentially uniformly distributed.

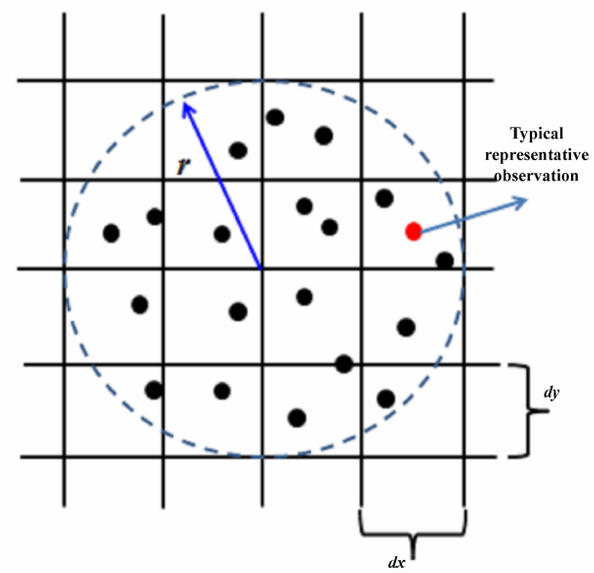
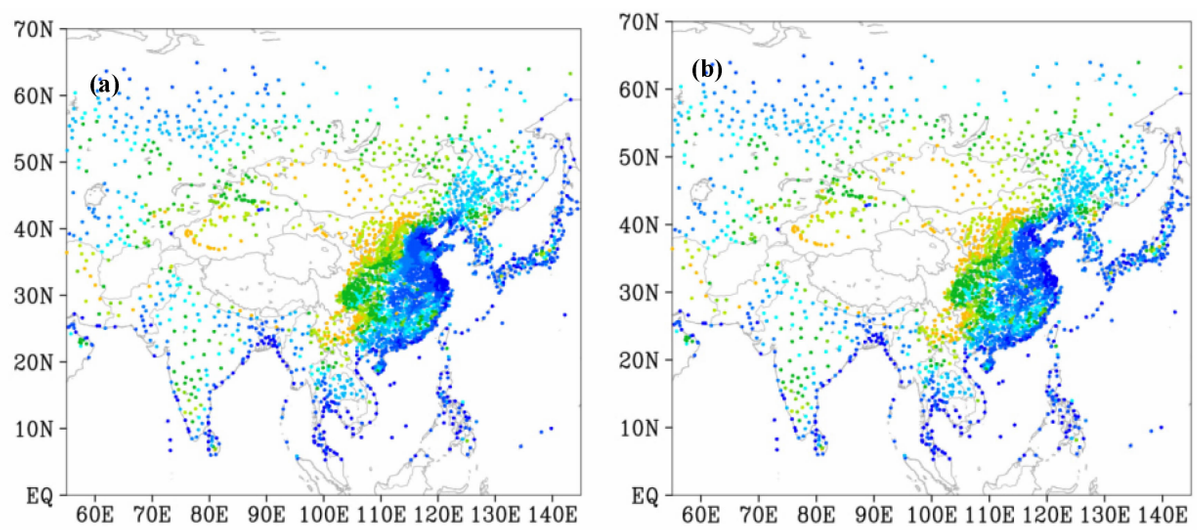


Figure 3. Schematic of the thinning scheme, where dx and dy are the model resolution in the longitudinal and latitudinal directions, respectively, and r represents the thinning radius. Each point represents an observation; the red point represents the typical observation within the radius.



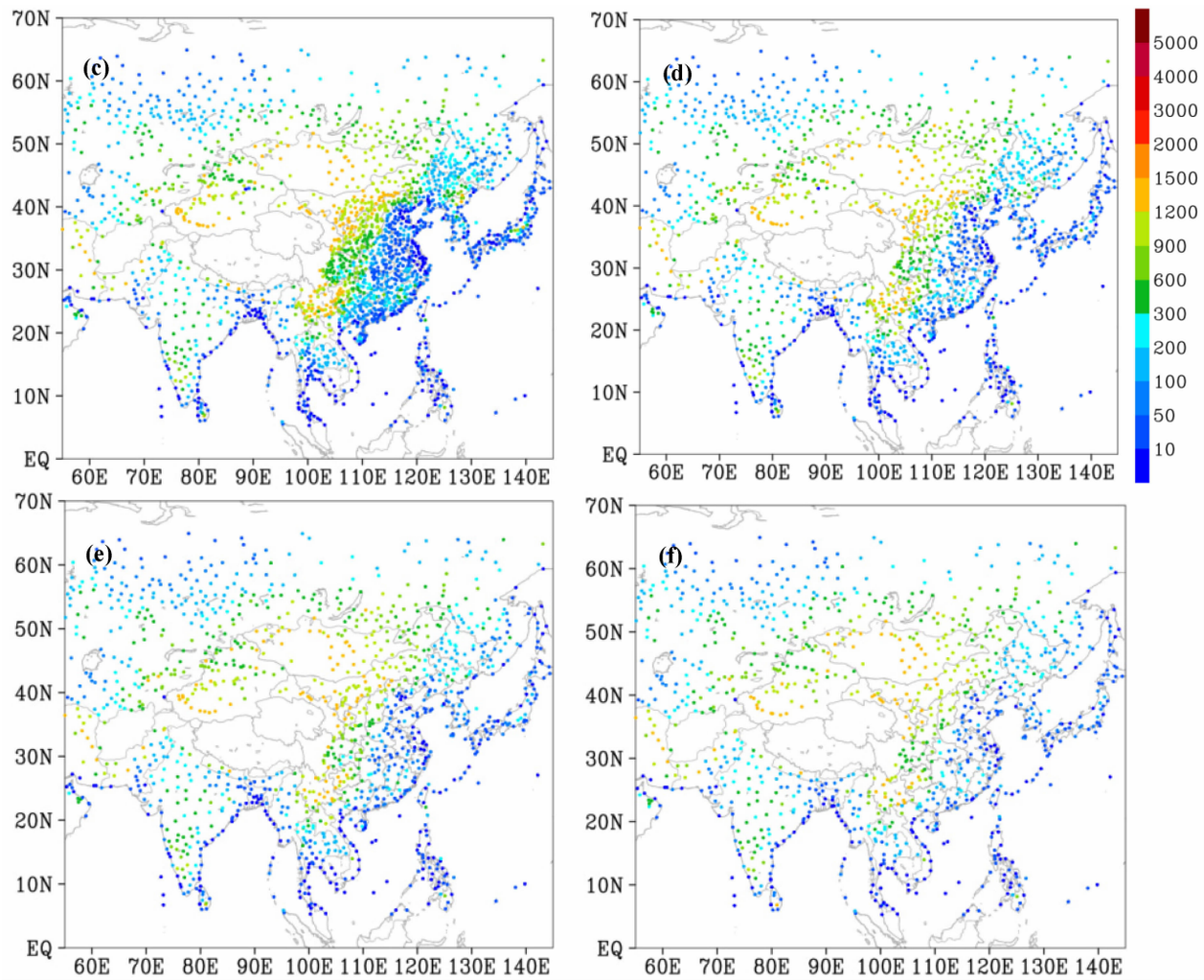


Figure 4. Distribution of observation stations for different thinning radii: (a) r=0 km (no thinning), (b) r=30 km, (c) r=60 km, (d) r=90 km, (e) r=120 km, and (f) r=150 km. The different colors represent the elevation of the observation stations (unit: m). The number in the lower-right corner represents the number of observations retained for different thinning radii.

3 DESCRIPTION OF THE WEATHER PROCESS

There was a notable precipitation process in southern China on January 27 –29, 2008. The precipitation during this period was mainly due to a strong blocking high that was steadily maintained over western Siberia. The cold, dry air converged with warm, moist air brought by the southern branch of a trough and the subtropical high over the middle and lower reaches of the Yangtze River and southern China, which caused a large-scale weather pattern conducive to rain, snow, and ice.

Figure 5 shows the actual 24-h precipitation from 0000 UTC on January 28 to 0000 UTC on January 29. The cumulative precipitation field is obtained from station observations. Precipitation values are interpolated from sites to model grids (30 km× 30 km) based on the Cressman interpolation method.

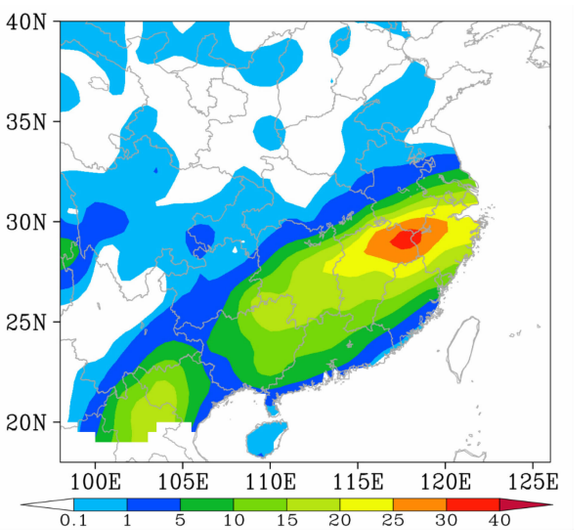


Figure 5. The 24-h cumulative precipitation from 0000 UTC on January 28 to 0000 UTC on January 29 from surface stations.

According to Fig.5, the precipitation was mainly concentrated in the southeastern coastal region of China with a southwest-northeast orientation. There were two relatively large precipitation centers in the precipitation zone. One was located in the southern part of Yunnan province, where the rainfall exceeded 20 mm, and the other was located at the border of the four provinces. In it, the amount of precipitation was greater than 30 mm.

4 ANALYSIS OF THE ASSIMILATION EXPERIMENT RESULTS

4.1 *The 24–h cumulative precipitation*

We used GRAPES-3DVAR to assimilate the surface pressure under different thinning radii and control the quality of these observations. Through bulk assimilation experiments, we studied the influence of different thinning radii on the simulation performance of the model. Fig.6 shows the 24-h cumulative precipitation from 0000 UT on January 28 to 0000 UTC on January 29 for different thinning radii. The 24-h cumulative precipitation derive from numerical model outputs, which is integrated for 24 hours from the analyzed field at 0000 UTC on January 28 using GRAPES model. The resolution of the modeled precipitation is 30 km×30 km.

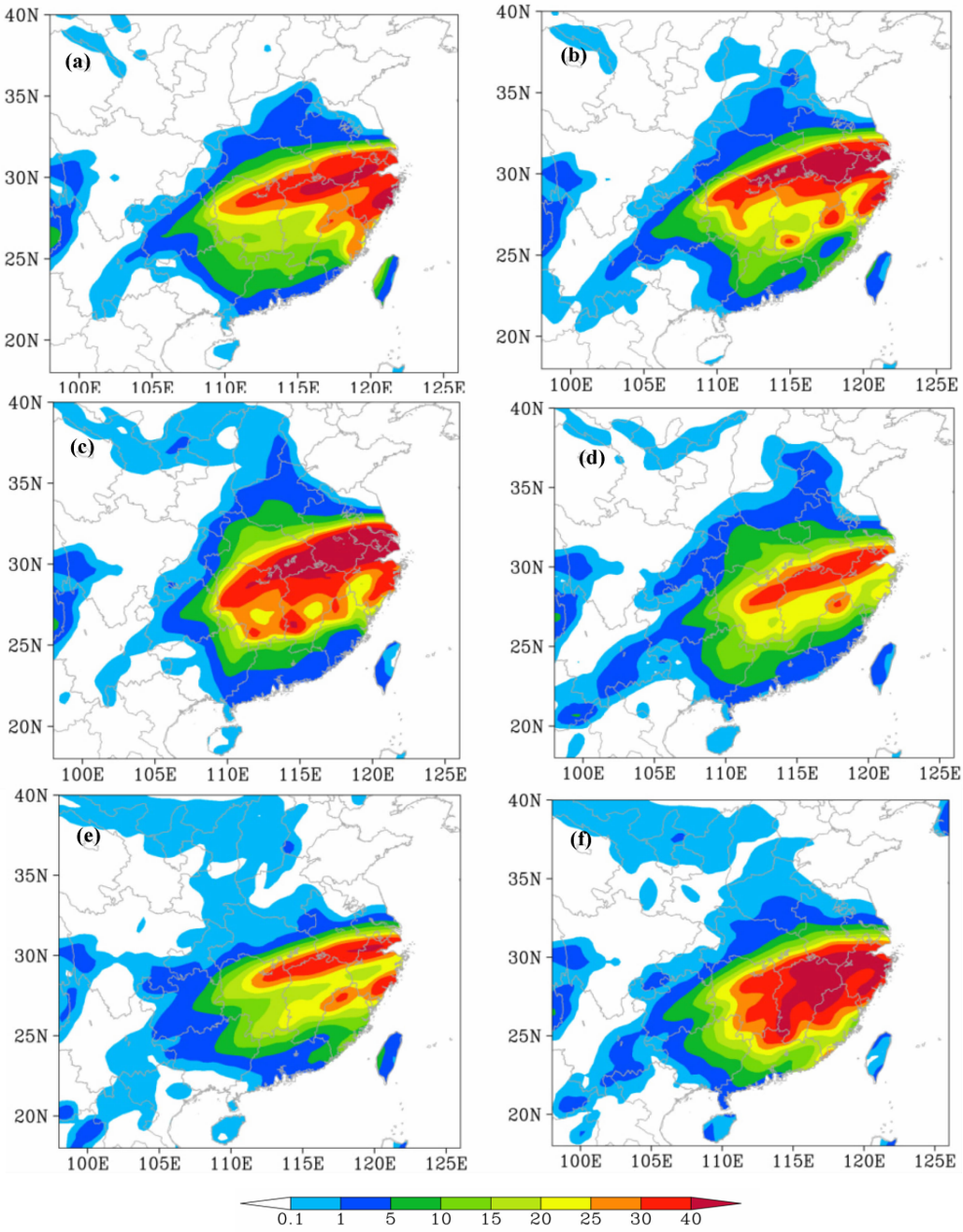


Figure 6. The 24-h cumulative precipitation from 0000 UTC on January 28 to 0000 UTC on January 29: (a) r=0 km (non-thining); (b) r=30 km; (c) r=60 km; (d) r=90 km; (e) r=120 km; and (f) r=150 km (unit: mm).

When the resolution of the model grid was set to 0.3°, the area with more than 25 mm of precipitation increased and expanded southward to cover all of Zhejiang. When the thinning radius was set to 30 km, the area with more than 25 mm of precipitation remained relatively large, and a precipitation center with more than 50 mm of precipitation appeared north of Hangzhou Bay. When r was 60 km, the area with more than 25 mm of precipitation increased and extended from northern Hunan and Jiangxi southward to their central and southern regions. Moreover, the area with more than 50 mm of precipitation increased, and the center moved into the southern part of Anhui. When r was 90 km, the area with more than 25 mm of precipitation shrank to a long, narrow zone stretching across the four provinces of Hunan, Jiangxi, Anhui, and Zhejiang with a southwest-northeast trend; in addition, there was a small precipitation center on the border of Jiangxi and Fujian. In this experiment, the center with excessive precipitation (more than 50 mm) disappeared. When r=120 km, the spatial distribution of precipitation was similar to the distribution for r=90 km; the area of the southern precipitation center with more than 25 mm of precipitation increased and extended eastward to the southern part of Zhejiang. When r=150 km, the area with more than 25 mm of precipitation was very large, and the precipitation zone expanded southward to include Hunan, Jiangxi, and most regions of Zhejiang. Moreover, a precipitation center with more than 50 mm of precipitation appeared at the border of Anhui, Jiangxi, and Zhejiang. Compared with the precipitation results for other thinning radii, the area and amount of precipitation of the precipitation center for a thinning radius of 90 km was closest to the actual situation.

4.2 *ETS Score*

The Equitable Threat Score (ETS) can effectively evaluate the simulation ability of each experiment. ETS≤0 indicates an invalid forecast, ETS>0 indicates a valid forecast, and ETS=0 denotes the best forecast. The formula for the ETS is

$$ETS=\frac{N_A-R(a)}{(N_A+N_B+N_C-R(a))} \quad (1)$$

and

$$R(a)=\frac{(N_A+N_B)(N_C+N_D)}{(N_A+N_B+N_C+N_D)} \quad (2)$$

N_A, N_B , N_C and N_D are defined in Table 1:

Table 1. Classification of the precipitation test.

observe\forecast	Yes	No
Yes	N_A	N_C
No	N_B	N_D

Figures 7 and 8 show the 24-h cumulative precipitation ETS score and the 6-h cumulative precipitation ETS score, respectively, for different experiments. The precipitation scores of the different experiments indicate that when the thinning radius was set to 90 km, the model performed well at forecasting the amount of precipitation. For the 6-h cumulative precipitation, when the thinning radius was set to 90 km, the improvement in the estimate of the low level of precipitation was relatively obvious (8a and 8b).

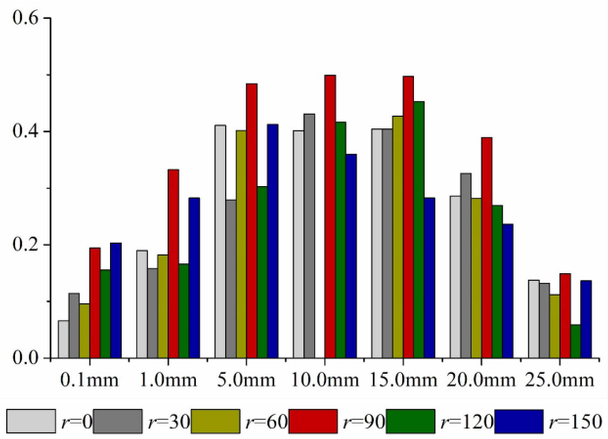


Figure 7. ETS of the 24-h cumulative precipitation from 0000 UTC on January 28 to 0000 UTC on January 29, 2008, in the different experiments.

4.3 *Investigation of the reason for the improvement in the precipitation estimates*

In the numerical model, we applied five intermittent assimilation processes to the pressure. Fig.9 shows the variation of the analysis error in the surface specific humidity and the surface temperature with the thinning radius for the five intermittent assimilation processes. The analysis error on the surface specific humidity and temperature in Fig.9 is the error between analysis field and surface observations. We interpolate the values at grids to stations using the bilinear method and then calculate the average error between them.

According to Fig.9, the analysis error of the specific humidity increased, decreased, and then increased again as the thinning radius increased. In particular, when the thinning radius was set to 90 km, the analysis error in the specific humidity reached a minimum in most of these five assimilation processes. Similar to the specific humidity, the analysis error in the surface temperature first increased, decreased and then slowly increased again as the thinning radius increased; it was smallest when the thinning radius was set to 90 km.

As shown in Fig.9, the analysis error in the specific humidity and surface temperature reached a minimum in most of the five assimilation processes when the thinning radius was set to 90 km. The analysis error was larger for both r>90 km and r<90 km, which means that the thinning scheme proposed in this paper has an optimal thinning radius. The results obtained using the

operational 3-D variational assimilation system were similar to those from the ideal 1-D variational assimilation system[8]. Therefore, to understand the large values for small values of r, more experiments are needed.

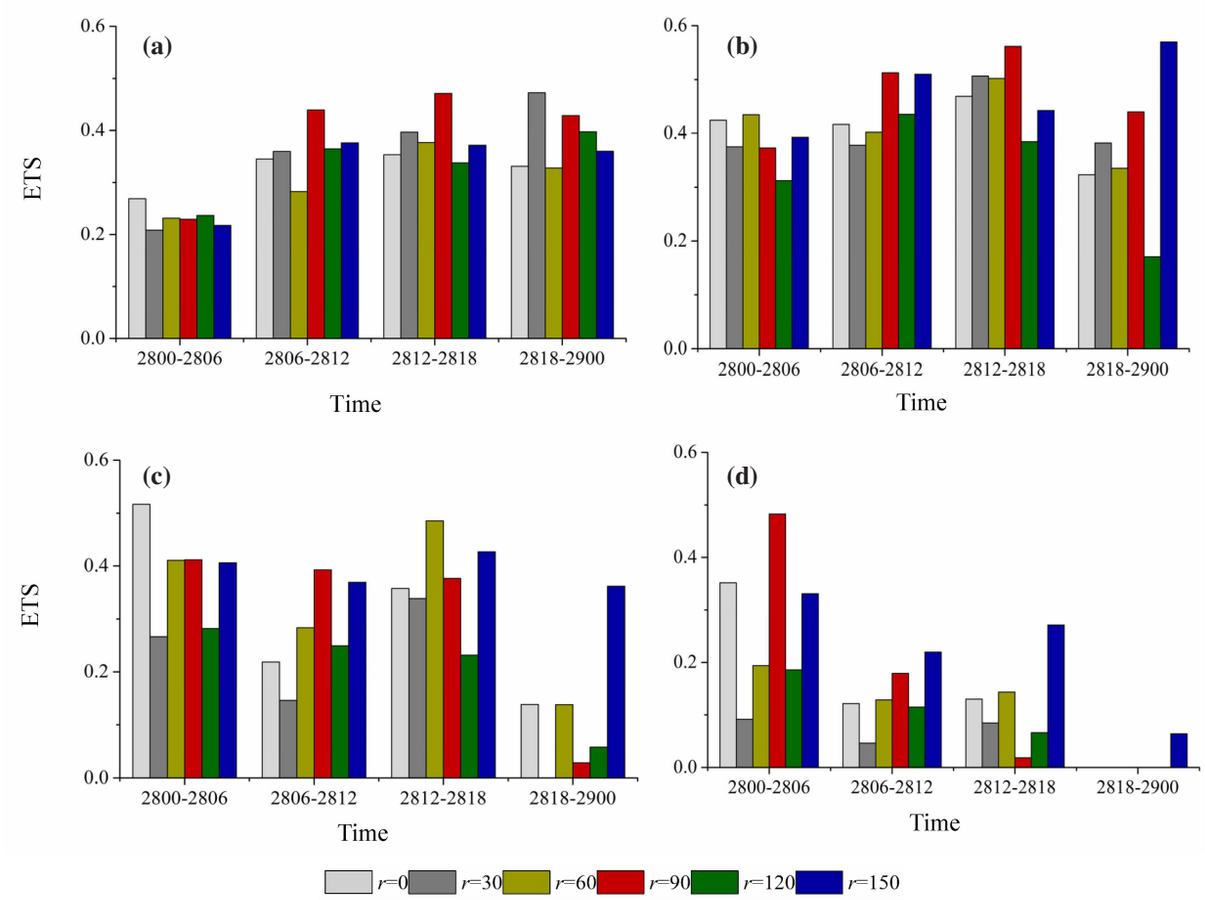


Figure 8. ETS of the 6-h cumulative precipitation when it was greater than (a) 0.1 mm, (b) 1 mm, (c) 5 mm, and (d) 10 mm in different experiments.

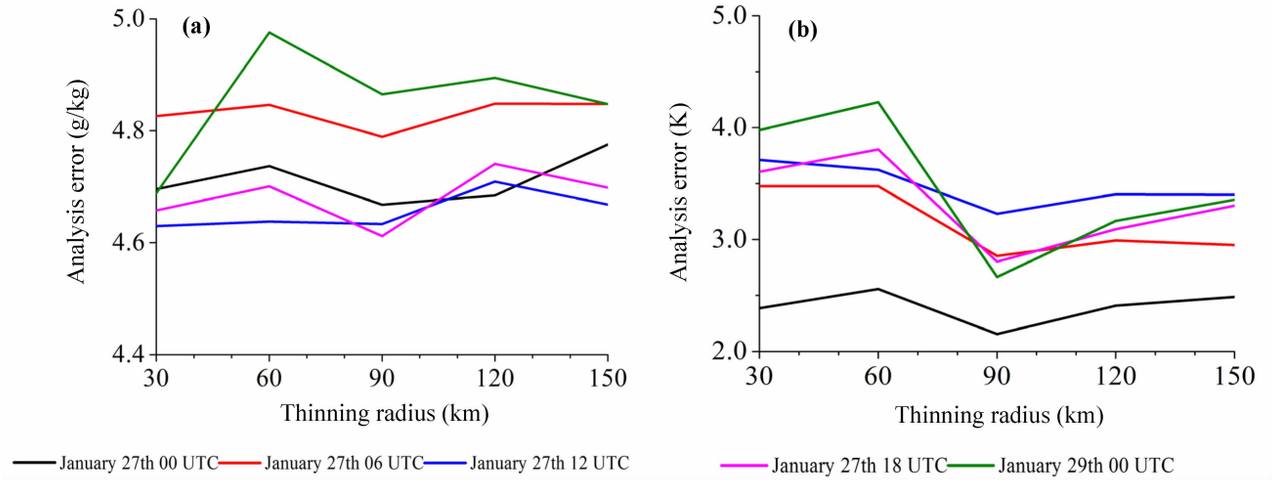


Figure 9. Variation in the analysis error in the surface specific humidity (a, unit: g/kg) and the surface temperature (b, unit: K) with thinning radius (unit: km).

In GRAPES-3DVAR, the minimization algorithm adopts the limited-memory Broyden-Fletcher-Goldfarb-Shanno (LBFGS) algorithm, i.e., the limited-memory BFGS method, to solve the problem, and the number of iterations is set to 100. Because the cost function in the five experiments has different orders of magnitude, we process the cost function with normalization. Fig.10 shows the variation of the normalized cost function J, (10a) and the norm gradient vector G, (10b) with the number of iterations for different thinning radii. The cost function decreased as the number of iterations increased, and the minimization process converged in all of the experiments. Saturation was reached after 20 iterations of the minimization algorithm. Regardless, the decrease in the cost function in the experiment with r=90 km performed the best for all cases, reaching saturation after only 10 iterations, which means most of the data misfit between the model and observations was reduced earlier than in other experiments (Zou et al.[22]).

During the minimization process, G decreased constantly, and, eventually, the normalized J approached a stable value, while G continuously approached 0. For different thinning radii, the numbers of iterations used for the objective function were 50, 42, 37, 49, and 44. When the thinning radius was set to r=90 km, the number of iterations was smallest.

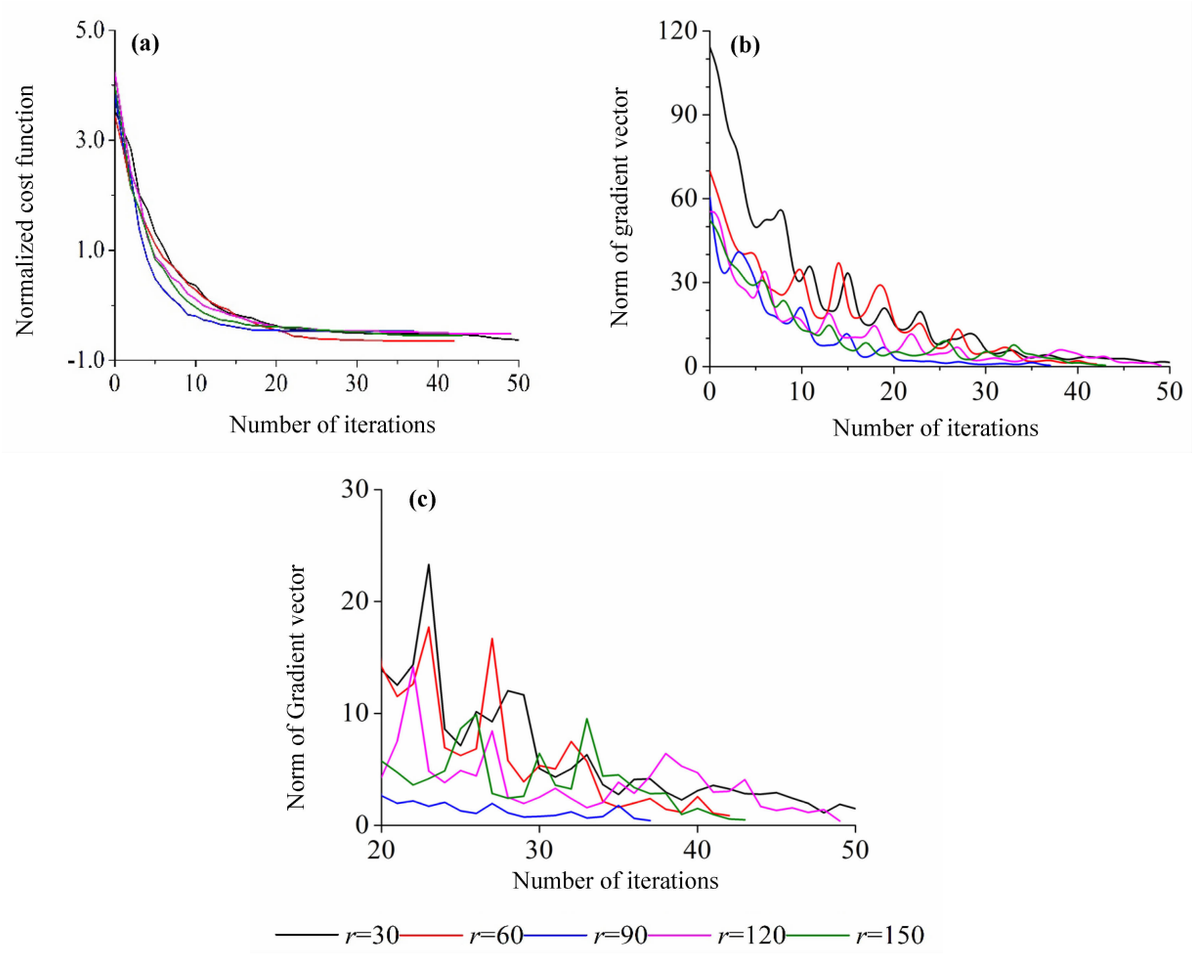


Figure 10. Variation of the normalized objective function J (a), the norm gradient vector G with the number of iterations for different thinning radii (unit: km) (b) and the norm gradient vector G for 20–50 iterations (c).

With three-dimensional variations, the ideal situation is that when the objective function reaches the optimum solution, the norm gradient vector G approaches 0. However, because the numerical model is complex and the assimilation data are uncertain, this ideal model is very difficult to obtain. This is mainly because the accuracy of the model does not reflect its agreement with the observation data (Županski[23]; Županski and Mesinger[24]; Županski[25]; Tsuyuki[26]). To further understand the variation in G with the number of iterations for different thinning radii, we expanded the region with 20–50 iterations in Fig.10b (Fig.10c). When the thinning radius was set to 30, 60, 120, or 150 km, the norm gradient vector G exhibited a relatively vigorous saw-tooth pattern as the number of iterations increased. When the thinning radius was set to 90 km, the norm gradient vector G exhibited relatively slow and gentle variations with the number of iterations, and the value of the norm gradient vector G was very close to 0. This result indicates that the minimization was

improved when the thinning radius was 90 km.

4.4 *Investigation of the reason for the improvement in the precipitation estimates*

We performed another set of experiments with a model resolution of 0.15°; the corresponding values of the radius (r) were 15, 30, 45, 60 and 75 km. Fig.11 shows the 24-h cumulative precipitation ETS scores of different experiments with a model resolution of 0.15°. The result suggests that when the thinning radius was set to 45 km or 60 km, the model performed better at predicting precipitation than the other experiments.

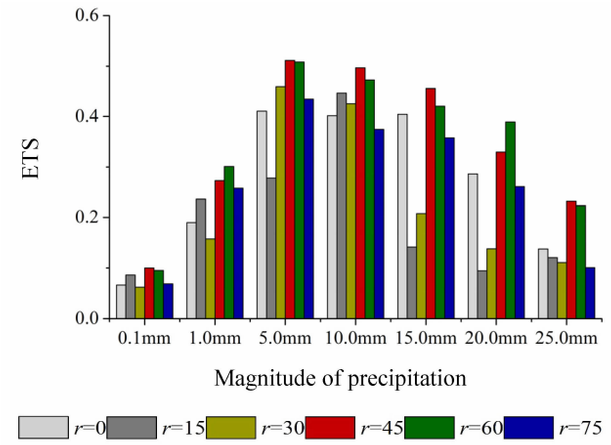


Figure 11. ETS of the 24-h cumulative precipitation from 0000 UTC on January 28 to 0000 UTC on January 29, 2008 in the different experiments with a model resolution of 0.15°.

Figure 12 shows the norm gradient vector G versus the number of iterations for different thinning radii. Additionally, we expanded that with 20–50 iterations. Using different thinning radii with a model resolution of 0.15°, the numbers of iterations applied in the objective function were 48, 45, 41, 41 and 46. The iteration was terminated first when the thinning radius was set to r=45 km or r=60 km, and the norm of gradient vector of the two experiments became steady after 10 iterations (Fig.12a). From the expanded G (Fig.12b), it could be found that the experiments with r=45 km and r=60 km performed better than other experiments. Although the G of the two experiments still exhibited a saw-tooth pattern as the number of iterations increased, slower and more moderate variations than other experiments were observed with the value of G closer to 0. Moreover, the G of the r=45 km experiment was slightly better than that of the r=60 km experiment.

The results of experiments with a model resolution of 0.15° show that when the ratio between the model resolution and observation radius is 1:3 or 1:4, the precipitation scores of different levels are relatively high, and the minimizations are improved. The ratio of 1:3 is slightly better than that of 1:4. However, because we performed only two sets of experiments for the precipitation case, the conclusions have some limitations. More cases of the thinning scheme are needed to be studied and definite conclusions are expected to be made.

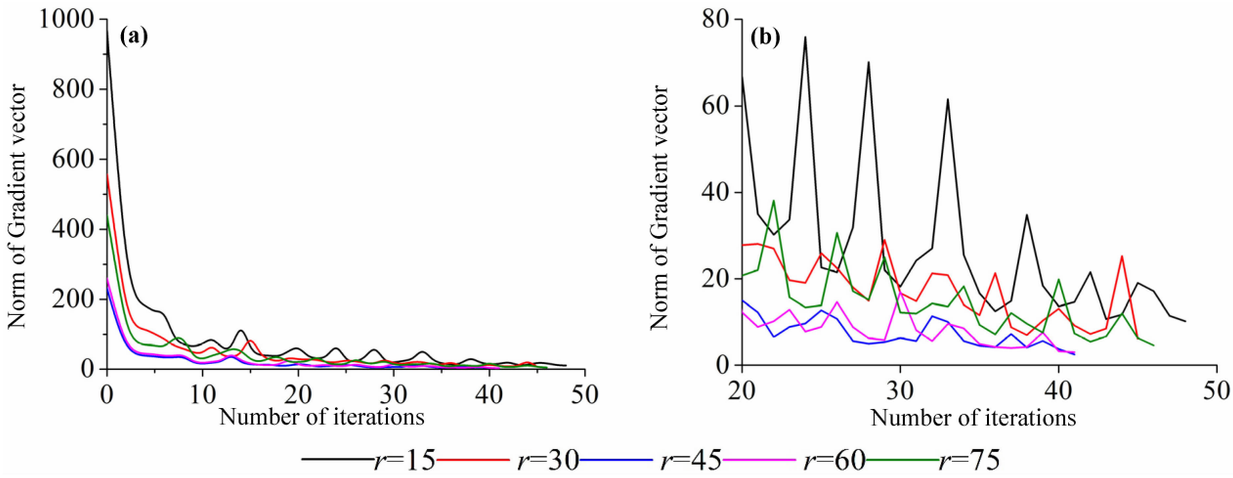


Figure 12. Variation in the norm gradient vector G with the number of iterations for different thinning radii (unit: km) (a) and the norm gradient vector G for 20–50 iterations (b) in the different experiments with a model resolution of 0.15°.

5 SUMMARY AND DISCUSSION

In this paper, to reduce the strong correlation of the representation error in observations and agree with the assumption of data assimilation, we propose a scheme for the spatial thinning of observation data based on the correlation coefficient between two observation stations. The main conclusions are as follows:

(1) Due to the high correlation between two adjacent observation stations, the thinning scheme described in this paper can select the optimum observation site within the observation radius.

(2) The use of this scheme in two sets of experiments in GRAPES-3DVAR indicates that when the ratio between the model resolution and observation radius is 1:3, the simulated results for the precipitation are closer to the actual situation and the precipitation scores at different times and for different levels are also relatively high. If the thinning radius is too large or too small, the improvement in the precipitation estimates is less than ideal.

(3) The reason for the improvement in the precipitation estimates is that when the model resolution and observation radius was 1:3, the error in the analytical field for specific humidity and temperature was the smallest, and the norm gradient vector G of the decrease in the cost function was close to 0. Moreover, the change is slow and gentle, which indicates that it reaches a quasi-stationary state in which the accuracy of the model and its agreement with the observation data are optimal.

In this paper, we conduct a preliminary investigation and study the agreement between the thinning radius and the model resolution. In the future, we will need more cases to study the application of this scheme in detail and to derive systematic conclusions. The experiments with other thinning schemes and a comparison of the difference between them will be conducted in our future study.

REFERENCES:

[1] LEWIS J M, DERBER J C. The use of adjoint equations to solve a variational adjustment problem with advective constraints [J]. Tellus A: Dyn Meteor Oceanogr, 1985, 37(4): 309-322.

[2] LE DIMET F X, TALAGRAND O. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects [J]. Tellus A: Dyn Meteor Oceanogr, 1986, 38(2): 97-110.

[3] TALAGRAND O, COURTIER P. Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory [J]. Quart J Roy Meteor Soc, 1987, 113(478): 1311-1328.

[4] LORENC A C. Analysis methods for numerical weather prediction [J]. Quart J Roy Meteor Soc, 1986, 112(474): 1177-1194.

[5] DALEY R. Estimating observation error statistics for atmospheric data assimilation [J] Annales Geophysicae, 1993, 11(7): 634-647.

[6] SCHWARTZ B, BENJAMIN S G. A comparison of temperature and wind measurements from ACARS-equipped aircraft and rawinsondes [J]. Wea Forecasting, 1995, 10(3): 528-544.

[7] RICHNER H, PHILLIPS P D. Reproducibility of VIZ radiosonde data and some sources of error [J]. J Appl Meteor, 1981, 20(8): 954-962.

[8] LIU Z Q, RABIER F. The interaction between model resolution, observation resolution and observation density in data assimilation: a one-dimensional study [J]. Quart J Roy Meteor Soc, 2002, 128(582): 1367-1386.

[9] BONDARENKO V, OCHOTTA T, SAUPE D, et al. The interaction between model resolution, observation resolution and observations density in data assimilation: A two-dimensional study [C]// 11th Symposium on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface, San Antonio, Texas, 2007: 5.

[10] JÄRVINEN H, UNDÉN P. Observation screening and first guess quality control in the ECMWF 3D-Var data assimilation system [J]. ECMWF Res Dep Tech Memor, 1997, 19(236): 1-33.

[11] OCHOTTA T, GEBHARDT C, SAUPE D, et al. Adaptive thinning of atmospheric observations in data assimilation with vector quantization and filtering methods [J]. Quart J Roy Meteor Soc, 2005, 131(613): 3427-3437.

[12] BRADLEY T Z. Evaluation of an innovation variance methodology for real-time data reduction of satellite data streams [C]// 18th Conference on Probability and Statistics in the Atmospheric Sciences. Atlanta, Georgia, 2006: 1-6.

[13] RAMACHANDRAN R, LI X, MOVVA S, et al. Intelligent data thinning algorithm for earth system numerical model research and application [C]// 21st International Conference on Interactive Information Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology. San Diego, California, 2005: 1-6.

[14] RAMACHANDRAN R, LI X, MOVVA S, et al. An improved data reduction tool in support of the real-time assimilation of NASA satellite data streams [C]// NASA Science Technology Conference Proceedings 2007. Adelphi, Maryland, 2007: 1-7.

[15] ZAVODSKY B, LAZARUS S, LI X, et al. Intelligent data thinning algorithms for satellite imagery [C]// Geoscience and Remote Sensing Symposium. Boston, Massachusetts, 2008: 1-4.

[16] QIN Z, ZOU X, WENG F. Evaluating added benefits of assimilating GOES imager radiance data in GSI for coastal QPFs [J]. Mon Wea Rev, 2013, 141(1): 75-92.

[17] WANG Jin-cheng, LU Hui-juan, HAN Wei, et al. Improvements and performance of the Operational GRAPES_GFS 3DVar System [J]. J Appl Meteor Sci, 2017, 28(1): 11-24 (in Chinese).

[18] WANG Jin-cheng, GONG Jian-dong, WANG Rui-chun. Estimation of background error for brightness temperature in GRAPES 3DVar and its application in radiance data background quality control [J]. Acta Meteor Sinica, 2016, 74(3): 397-409 (in Chinese).

[19] YANG J, SHEN X. The construction of SCM in GRAPES and its applications in two field experiment simulations [J]. Adv Atmos Sci, 2011, 28(3): 534-550 (in Chinese).

[20] LIU H, XUE J, GU J, et al. Radar data assimilation of the GRAPES model and experimental results in a typhoon case [J]. Adv Atmos Sci, 2012, 29 (2): 344-358 (in Chinese).

[21] WANG Jin-cheng, ZHUANG Zhao-rong, HAN Wei, et al. An improvement of background error covariance in the global GRAPES varitional data assimilation and its impact on the analysis and prediction: statistics of the three-dimensional struture of background error covariance [J]. Acta Meteor Sinica, 2014, 72(1): 62-78 (in Chinese).

[22] ZOU X, KUO Y H, GUO Y R. Assimilation of

atmospheric radio refractivity using a nonhydrostatic adjoint model [J]. Mon Wea Rev, 1995, 123 (7): 2229-2250.
[23] ŽUPANSKI M. A preconditioning algorithm for large-scale minimization problems [J]. Tellus A, 1993, 45 (5): 478-492.
[24] ŽUPANSKI D, MESINGER F. Four-dimensional variational assimilation of precipitation data [J]. Mon Wea Rev, 1995, 123(4): 1112-1127.
[25] ŽUPANSKI M. A preconditioning algorithm for four-dimensional variational data assimilation [J]. Mon Wea Rev, 1996, 124(11): 2562-2573.
[26] TSUYUKI T. Variational data assimilation in the tropics using precipitation data. Part III: assimilation of SSM/I precipitation

Citation: ZHAO Hong and LIU Yin. Experimental application of a thinning scheme for the assimilation of surface observation data in GRAPES-3DVAR [J]. J Trop Meteor, 2018, 24(3): 334-345.