ADVANCES IN ATMOSPHERIC SCIENCES, VOL. 33, OCTOBER 2016, 1199–1208

# The Effects of Land Surface Process Perturbations in a Global Ensemble Forecast System

Guo DENG*[1], Yuejian ZHU[2], Jiandong GONG[1], Dehui CHEN[1], Richard WOBUS[2], and Zhe ZHANG[3]

[1]*National Meteorological Center, China Meteorological Administration, Beijing* 100081*, China*

[2]*Environmental Modeling Center, National Centers for Environmental Prediction,*
*5830 University Research Court, College Park, MD* 20740*, USA*

[3]*Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing* 100029*, China*

(Received 25 May 2016; revised 20 July 2016; accepted 25 July 2016)

ABSTRACT

Atmospheric variability is driven not only by internal dynamics, but also by external forcing, such as soil states, SST, snow, sea-ice cover, and so on. To investigate the forecast uncertainties and effects of land surface processes on numerical weather prediction, we added modules to perturb soil moisture and soil temperature into NCEP's Global Ensemble Forecast System (GEFS), and compared the results of a set of experiments involving different configurations of land surface and atmospheric perturbation. It was found that uncertainties in different soil layers varied due to the multiple timescales of interactions between land surface and atmospheric processes. Perturbations of the soil moisture and soil temperature at the land surface changed sensible and latent heat flux obviously, as compared to the less or indirect land surface perturbation experiment from the day-to-day forecasts. Soil state perturbations led to greater variation in surface heat fluxes that transferred to the upper troposphere, thus reflecting interactions and the response to atmospheric external forcing. Various verification scores were calculated in this study. The results indicated that taking the uncertainties of land surface processes into account in GEFS could contribute a slight improvement in forecast skill in terms of resolution and reliability, a noticeable reduction in forecast error, as well as an increase in ensemble spread in an under-dispersive system. This paper provides a preliminary evaluation of the effects of land surface processes on predictability. Further research using more complex and suitable methods is needed to fully explore our understanding in this area.

**Key words**: perturbation, land surface processes, GEFS, ensemble transform with rescaling

**Citation**: Deng, G., Y. J. Zhu, J. D. Gong, D. H. Chen, R. Wobus, and Z. Zhang, 2016: The effects of land surface process perturbations in a Global Ensemble Forecast System. *Adv. Atmos. Sci.*, **33**(10), 1199–1208, doi: 10.1007/s00376-016-6036-8.

## 1. Introduction

The importance of land surface processes to numerical weather prediction (NWP) has been recognized in recent years. The first few meters of ground below Earth's surface has a thermal capacity comparable to 1/10 of the entire atmospheric column, which could mean the change in atmospheric temperature through this layer is considerable (Lewis, 2007). It is generally agreed that land surface processes have a substantial influence on both large-scale and mesoscale circulation (Chen and Dudhia, 2001). Large-scale weather patterns are influenced by land surface processes as a consequence of change in moisture influx, static stability, convergence and divergence of flow patterns, vertical motions, and latent heating (Nicholson, 1988; Betts et al., 1996; Li and Zou, 2009). An improved understanding of atmosphere–land interaction, along with accurate measurements of land-surface properties, especially soil moisture, would constitute a major intellectual advantage. And potentially, such progress could lead to dramatic improvements in tackling a number of forecasting problems, including the location and timing of deep convection over land, quantitative precipitation forecasting, and seasonal climate prediction (National Research Council, 1998).

Among all currently available numerical prediction methods, ensemble forecasting has developed at a particularly fast pace during the last decade, and is expected to continue to play an increasingly important role in weather forecasting compared with other approaches. In ensemble forecasts, a set of different states is discretely sampled from a probability density function to account for uncertainty in the initial conditions. To achieve a reliable probabilistic weather forecast system, a series of schemes have been tested and applied by various NWP centers and researchers. For instance: the time-lagged method, which is a very simple but effective method (Yuan et al., 2008, 2009); the combined application of ensemble data assimilation and the singular vector based perturbations method at the European Centre for Medium-Range

* Corresponding author: Guo DENG
Email: deng719@cma.gov.cn

Weather Forecasts (Buizza et al., 2008, 2010); the ensemble Kalman filter (EnKF) method plus stochastic perturbation, which is operated at NCEP (Hou et al., 2016); the EnKF with a four-dimensional data method plus a kinetic energy backscatter algorithm, used at the Meteorological Service of Canada (Charron et al., 2010); and bred vectors, employed at the National Meteorological Center, China Meteorological Administration (Deng et al., 2010). Among these methods, at the present moment in time, the EnKF is particularly widely studied and applied as an initial condition perturbation or data assimilation method (Xue et al., 2006; Gao and Xue, 2008; Weng and Zhang, 2012).

Although ensemble products are playing an increasingly important role in daily probabilistic forecasts, the issue of unreliability and under-dispersion remains a known problem in the field of ensemble forecasting (Hamill and Colucci, 1997). Sutton et al. (2006) attributed the problems to the inadequate resolution of ensemble members (Mullen and Buizza, 2002), suboptimal methods for generating initial conditions (Hamill et al., 2000; Wang and Bishop, 2003), model biases related to problems in the parameterization of surface and boundary layer effects and the diurnal cycle (Davis et al., 2003), or a lack of perturbation in the characteristics of the land surface state. On the other hand, since atmospheric variability is driven not only by internal dynamics, but also by external forcing factors, such as soil states, SST, snow and sea-ice cover etc., consideration of the uncertainties and effects of land surface processes on the performance of an ensemble prediction system (EPS) is of great importance for improving its forecasting skill. Most methods dealing with the uncertainties are related to the initial state of the atmosphere, but only a small amount of work to perturb the initial state of the land surface in ensemble systems has been carried out thus far. Therefore, in most EPSs, the initial state of soil moisture and soil temperature is the same for each member in most currently available operational ensemble prediction systems (Wang et al., 2010). Sutton et al. (2006) tried to perturb the soil moisture to test its effect on temperature forecasts and precipitation forecasts; Wang et al. (2010) generated perturbations of surface variables, such as soil moisture content and surface temperature etc., to represent uncertainties in the surface initial conditions; McLay et al. (2012) introduced SST variation in the U.S. Navy's GEFS; while at the UK Met Office, its operational EPS contains SST (stochastic process) and soil-moisture perturbations (Tennant and Beare, 2014). Until now, most research related to land surface perturbation has been carried out in regional ensemble forecast systems. But what is the effect at the global scale (i.e. in a GEFS)? Furthermore, most studies have focused on the effects on near-surface variables, but what are the effects on forecast variables in the middle or upper levels? And what is the effect if we consider only the soil uncertainties in the ensemble forecast system? In the present work, using the addition of a module into NCEP's GEFS (Wei et al., 2005, 2008) to perturb the soil moisture and soil temperature, and comparing the results of a set of parallel experiments involving different configurations of land surface and atmospheric perturbation, we investigated whether or not the perturbation of soil states only could improve the system's forecasting skill. The aim in carrying out this study was to expand upon the relatively limited knowledge regarding land surface process perturbations in EPSs.

The remainder of the paper is organized as follows: Section 2 provides a brief description of the parallel experimental design in the GEFS. Section 3 reports the uncertainties and changes in variables as a result of land surface perturbation. A probabilistic verification of the results regarding the predictability of the GEFS under the different configurations is presented in section 4. Finally, discussion and conclusions are provided in section 5.

## 2. Configuration of the GEFS

The NCEP's GEFS (http://www.emc.ncep.noaa.gov/gmb/yzhu/html/ENS_IMP.html) was developed based on the earlier Global Forecast System (GFS) (Version 8.0.0, T126L28, NCEP Office Note 442) (Global Climate and Weather Modeling Branch, 2003). The horizontal resolution is approximately 110 km in both the analysis and forecast model for the four GFS cycles at 0000, 0600, 1200 and 1800 UTC. The vertical resolution is 64 hybrid layers for the entire 16-day forecast. The GFS land-surface model component is the Noah Land Surface Model (Noah LSM; Chen et al., 1996). Its land-surface parameterization has four subsurface layers (10, 40, 100 and 200 cm). The model also contains an improved algorithm of frozen soil, ground heat flux, and energy/water balance at the surface, along with reformulated infiltration and runoff functions and an upgraded vegetation fraction. The heat capacity, thermal and hydraulic diffusivity, and hydraulic conductivity coefficients are a function of the soil moisture content (Pan and Mahrt, 1987). To obtain initial values of soil moisture and soil temperature, Noah LSM cycles continuously on itself in the Global Data Assimilation System cycles. Values are updated at each model forecast integration time step in response to land-surface forcing (precipitation, surface solar radiation, and near-surface parameters: temperature, humidity, and wind speed) (Campana and Caplan, 2005).

The initial perturbations of the GEFS are generated by an ensemble transform (ET) with rescaling technique, and the methods are the same as employed in Bishop and Toth (1999), Wei et al. (2008) and Deng et al. (2012). To test the effect of land surface process perturbations on the forecasting skill of the GEFS, parallel experiments were devised (Fig. 1). On one side, perturbations were included only in the atmospheric component, named the "control run". Its characteristics included: initial perturbation (ET technique) in the atmospheric component, with four ensemble members; and the tropical cyclone relocation technique. On the other side, in the sensitivity run, perturbations were included in both the atmospheric and land surface processes, besides all the characteristics in the control run, with two methods: the "replacement run" began with the control surface analysis (i.e. cold start), and then, after one cycle (6 h), each member used its

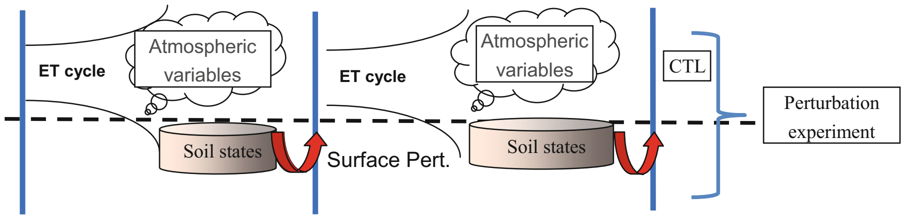


**Fig. 1.** Configuration of the atmosphere perturbation (control) run and the atmosphere/surface perturbation (replacement and rescaled) runs.

forecasted soil temperature and soil moisture from the previous cycle for its initial surface condition, and so on until the end of the experiment; in the "rescaled run", the soil moisture and temperature differences between each forecast member and the deterministic GFS model forecast were added. To maintain the values of the perturbations within a reasonable range, the maximum amplitude of the perturbations was scaled to the climate reference values. For all the perturbation tests, the variables perturbed included soil temperature (four layers: 0–10 cm; 10–40 cm; 40–100 cm; 100–200 cm) and soil moisture (four layers, similar to soil temperature, including soil volumetric water content in the fraction and liquid soil moisture). To avoid the model drifting after long-term integration (several months later), an exponential function of soil moisture (as well as soil temperature) perturbation and soil climate was devised in the experiment [after the land surface process devised in the Global Spectral Forecast Model (T213/T639) at the National Meteorological Center, China Meteorological Administration]. That is, at the beginning of the model integration period, the perturbation part was maintained as a comparatively larger component, and then the climate states gradually dominated; after three months of integration, the soil states would finally convert to the model climate value. As for the rescaled run, because the sum of soil states perturbations was near zero, it would also prevent model from drifting (Tennant and Beare, 2014). The test period was from 1200 UTC 22 August 2006 to 1200 UTC 24 September 2006.

## 3. Uncertainties and variation resulting from land surface perturbation

As described in section 2, the four members in the control experiments used the same initial land surface conditions, whereas they were different in the sensitivity run. Therefore, the differences in the results of the three experiments could only result from the uncertainties in the soil temperature and soil moisture. To investigate whether these uncertainties impose any impacts on the GEFS, the changes in the land surface processes and free atmosphere were explored through comparison with the control experiment.

### 3.1. *Soil perturbation variation*

To analyze the effects of perturbing land surface variables on the predictability of the GEFS, we began by examining the variation in soil properties due to land surface perturbation. Firstly, the average volumetric soil moisture difference between the four perturbation members (replacement or rescaled run) and the control experiment (four members) at the start and at a later time (e.g., one week later) was examined (not shown). It was found that, at the very beginning of model integration (second integration cycle after a cold start), the differences between the two experiments were apparent. This indicated that land surface process uncertainties had been introduced into the ensemble system and the interaction between land surface processes and the atmosphere subsequently took place. Although the soil moisture difference was small at the beginning, the difference grew rapidly as the model integrated, indicating strong soil moisture exchange among land surface processes and the atmosphere compared with the control experiment. It is interesting to note that large soil moisture differences did not necessarily correspond to large soil temperature differences, and vice-versa. This indicated that, although the soil moisture and temperature interacted with the atmosphere above, the uncertainties varied temporally and spatially for different elements. Similar characteristics were found in all the other soil levels. However, in the deeper soil layers, the change in soil temperature or moisture decreased rapidly compared with the levels above (Fig. 2 and 3); that is, the deeper down in the soil, the less of a difference was obtained between the perturbed and control runs. An explanation for this might be the fact that deeper soil layers possess more stable thermodynamic and humid characteristics. It was noticed, for instance, that uncertainties in both soil moisture and temperature were large in high-altitude mountain areas, such as the Tibetan Plateau and Iranian Plateau, which may have resulted from fewer high quality surface observations, but nevertheless affected the model's integration and forecasting ability. To investigate the soil variation due to land surface perturbation more thoroughly, the time series of soil spread across the Northern Hemisphere during the experimental period were examined. The ensemble spread was used to measure forecast uncertainties, which was calculated by the deviation of ensemble forecasts from their mean.

Figure 2 presents the time series of soil temperature spread for the four-member rescaled perturbation, replacement perturbation, and control experiments, at different soil depths. Because there was no soil perturbation in the control test, the spread for the control was close to zero. It is clear that the spread at the near-surface soil level reached a

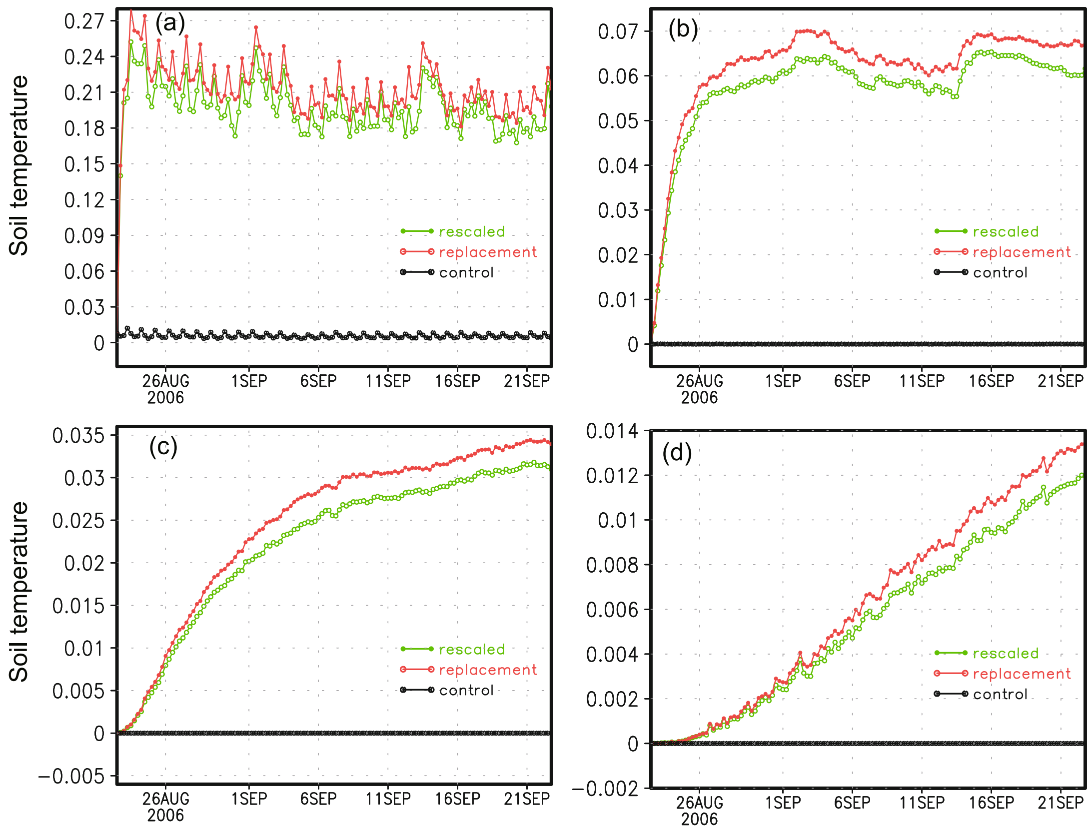


**Fig. 2.** Temporal evolution of soil temperature spread in the perturbation experiments at different soil depths: (a) 0–10 cm; (b) 10–40 cm; (c) 40–100 cm; (d) 100–200 cm (units: K).

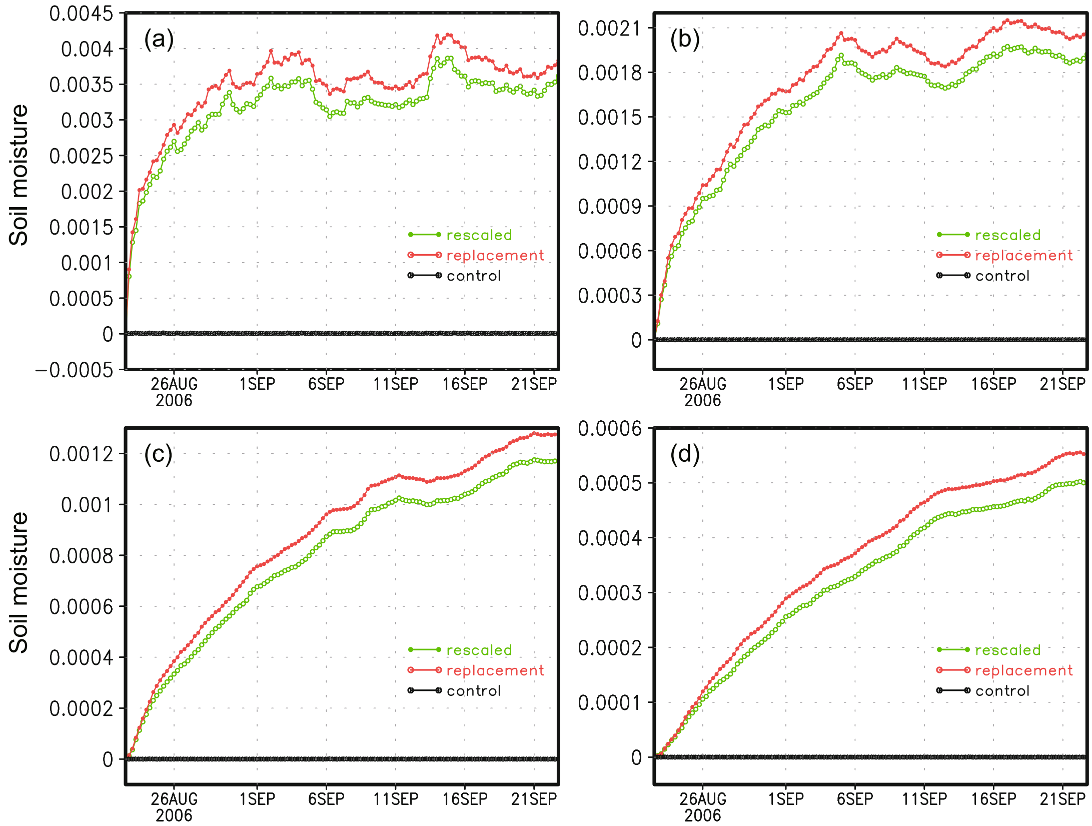


**Fig. 3.** As in Fig. 2 but for soil moisture (units: %).

steady state immediately, despite significant fluctuation [Fig. 2a, 0–10 cm, which reflects the range of probable soil temperature uncertainties at this level; the calculation area was the Eurasian continent (20°–80°N, 0°–150°E)]. By contrast, the soil temperature spread at deeper layers (Figs. 2b–d) presented a rapid increase with time, meaning the uncertainties of soil temperature at these levels did not even reach a saturation value within the experimental period. This phenomenon could be explained by the fact that the timescales at which the land surface interacts and responds to atmospheric forcing differ greatly with soil depth (Viterbo and Beljaars, 1995; Beljaars et al., 2007). Studies indicate that the timescales at which the atmosphere and land surface processes interact range from instantaneous to seasonal (Beljaars et al., 2007). Furthermore, Viterbo and Beljaars (1995) tried to deduce the timescales associated with each soil layer by using the soil heat budget function, and concluded that the timescales of interaction among soil layers depend on the soil depth, the heat capacity, and soil moisture; for any given layer, the interactions with lower layers operate at longer timescales than interactions with upper layers, and the timescales range from fractions of a day to around 150 days. Therefore, the timescales of interactions between the atmosphere and land surface processes in our experiments were expected to differ from the diurnal to seasonal scale, since there are four soil layers in the land surface processes of the GEFS. At the top level, the timescale of interactions between land surface processes and the atmosphere was very short (not longer than one day), so their interactions reached a balanced state very quickly. In the latter stages of the experiment, there was a tendency for the spread to decrease slightly compared with the earlier period, and this phenomenon probably resulted from the imperfections of our experimental design: if the perturbations had been devised more strategically, such as the non-cycling surface breeding in Wang et al. (2010), the effect would probably have been more obvious. At the second soil level (10–40 cm), the spread grew steadily in the later stages of the experiment, and for the third and last layer, the slopes were larger, implying a longer timescale of interactions. The evolution of spread for soil moisture was similar to that of soil temperature, albeit there were some differences in the variation range and slope (Fig. 3). Given the finding that the spread of soil moisture and temperature continued to increase with soil depth, to determine the overall effect of land surface process perturbations on the predictability of the GEFS should require a longer model integration time.

### 3.2. *Effect of land surface perturbations on atmospheric variables*

It is known that land surface processes play an important role in NWP: as the surface heats up during the day, sensible energy is transferred to the atmosphere, moisture evaporates from the soil or transpires from plants (latent heating), and soil in the lower levels is heated. Changes in land-surface properties have been shown to influence the heat and moisture fluxes within the PBL, which influences convective available potential energy and other measures of deep cumulus cloud activity (Pan and Mahrt, 1987; Pielke, 2001; Sutton et al., 2006). The effect of land surface processes in a numerical prediction system is reflected explicitly and inexplicitly in the boundary layer dynamic and thermodynamic equations; for example, the friction term in the momentum equation, the sensible and latent heating in the energy equation, and the local water vapor budget in the moisture conservation equation. The soil moisture and temperature interact with the atmosphere above in the form of sensible heat flux and evapotranspiration (latent heat flux heat flux). The latent and sensible heat flux within the PBL affect the development of convection and precipitation—a mechanism that operates globally (Pielke, 2001). Therefore, discussing the distribution of sensible and latent flux is key to understanding how land surface processes affect the forecasting skill of tools such as the GEFS.

Figure 4 shows the difference between the two perturbed tests and control experiments (ensemble mean), and the spread for each test within the Eurasian continent. A positive value in Fig. 4a means that the overall forecasted surface latent heat flux in the perturbed run was larger than in the

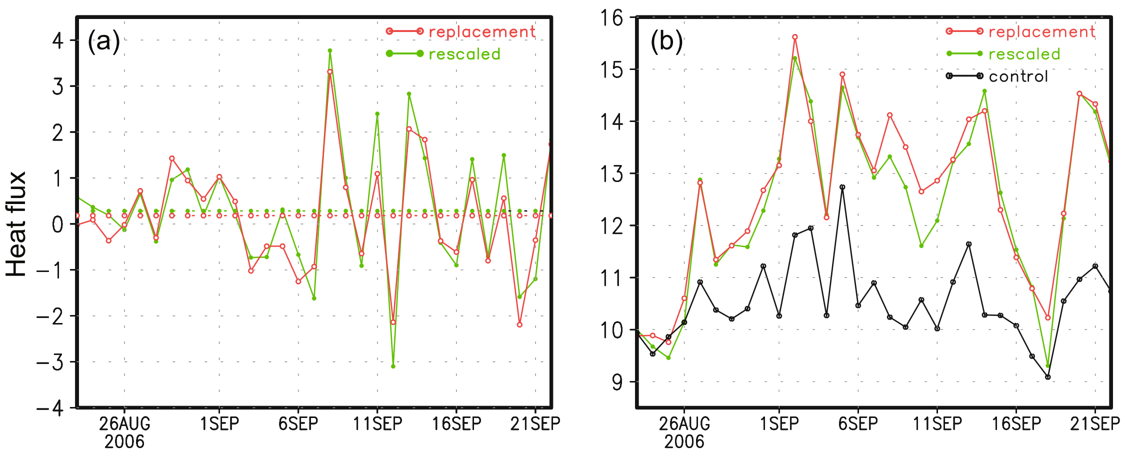


**Fig. 4.** Time series of average surface latent heat flux (units: W m$^{-2}$) for the (a) difference between the perturbed and control experiments, and (b) spread for each test, within the Eurasian continent.

control run, while a negative value indicates a lower overall latent heat exchange. It is clear that uncertainties in the GEFS resulted in a comparatively larger surface latent heat flux during the one-month experiment over the area; and in view of the spread, the two perturbation tests were obviously larger than the test without soil perturbation. Meanwhile, for surface sensible heat flux, we found a reduction between the control test and the two perturbations (Fig. 5). The positive or negative value between the perturbation tests and the control run was not the point of our focus, but by combining with the spread we were able to find that the perturbation of land surface processes did indeed contribute to obvious variation in forecasted heat fluxes. Therefore, we can confidently conclude that uncertainties in land surface processes tend to change the exchanges of surface sensible and latent heat flux in systems such as the GEFS, therefore affecting the development of atmospheric processes.

To investigate the effects of land surface process perturbations on day-to-day forecasting, the 5-day lead time forecast with the initial date of 1 September was arbitrarily selected (Fig. 6). It can be seen that all three of these groups of ensemble means differed obviously from one another; for instance, geopotential height and 2-m temperature. Furthermore, the time series of average relative humidity in the replacement perturbation, rescaled perturbation and control tests over the area of focus were calculated (not shown), and the results also indicated that uncertainties in land surface processes contributed to quite different forecast results from day to day, therefore affecting the performance of the ensemble forecasts. Because land surface uncertainties within the GEFS result in a change in the surface energy budget, it follows that partitioning of thermal energy between latent and sensible heat flux (Dr. Jun DU, NCEP, 2006, personal communication), and further alteration of PBL processes, convection, radiation, and other processes in the free atmosphere, will also take place. At the same time, these variations in free-atmospheric processes will feed back to the land surface in perturbation experiments, and thus the interactions between land surface processes and the atmosphere cycles and induces forecast differences between the perturbation and control experiments.

## 4. Evaluation of predictability due to land surface perturbation

For probabilistic forecasts, there are many existing verification methods to help with judging the quality of a forecast system. Some measures assess the reliability or resolution, while others provide a combined measure of both. No single verification measure provides complete information on the quality of a product (Stanski et al., 1989). The resolution is defined as a forecast system's ability to distinguish, ahead of time, different outcomes of the real atmosphere. Resolution, as the inherent predictive value of a forecast system, is one of two important forecast attributes most sought after by developers of forecast systems, could be only enhanced through improving forecast system. Reliability, however, is equally important in real-world applications (Toth et al., 2006). It refers to the ability to provide unbiased probability estimates for forecasts. To assess the effect of land surface processes on the GEFS, various scores that evaluate the performance of probability forecasts were calculated for the three experiments.

### 4.1. *Relative Operating Characteristic area*

The Relative Operating Characteristic (ROC) curve is a plot of the hit rate as a function of the false alarm rate of a series of deterministic forecasts, obtained from the probability distribution by considering several probability thresholds, from $p = 0\%$ (event systematically forecasted) to $p = 100\%$ (event never forecasted) (Atger, 1999). It measures the ability of the forecast to discriminate between events and non-events, and indicates the characteristic attribute of resolution. The area under the curve (the "ROC area") is a useful summary measure of forecast skill, and for a perfect ensemble prediction system, ROC area = 1 (Richardson, 2000). Figures 7a and b show the ROC area scores for the 1000 hPa and 500 hPa geopotential heights, respectively. It can be seen that, at the short forecast lead times (1–5 days), there was no obvious difference among the three experiments; as the

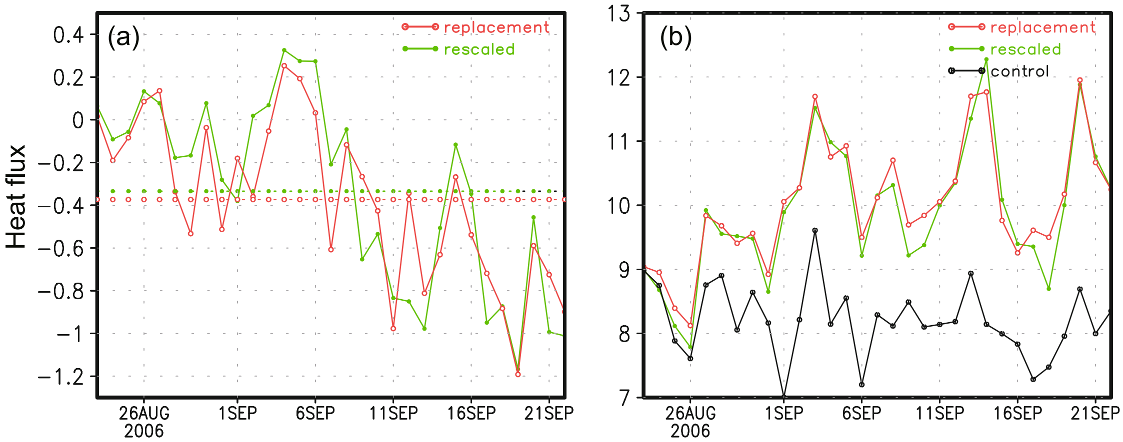


**Fig. 5.** As in Fig. 4 but for sensible heat flux (units: W m$^{-2}$).

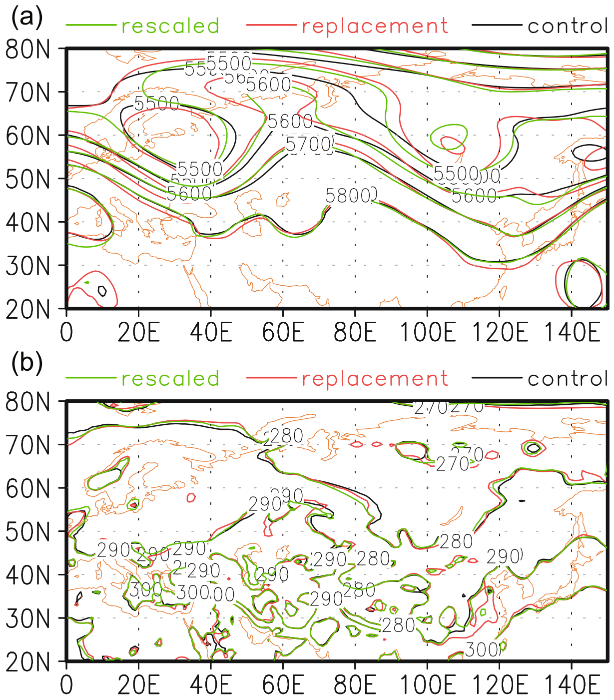


**Fig. 6.** A 5-day lead time forecast ensemble mean for the three ensemble tests at 500 hPa (initial time is 1 September 2006): (a) geopotential height (units: gpm); (b) 2-m temperature (units: K).

forecast lead time increased, from day 4 to day 9, the ROC area in the two perturbation experiments was slightly better than in the control. As the lead time increased beyond 10 days, however, the scores of the rescaled and control experiments were almost the same. From the low level (1000 hPa) to the mid-level (500 hPa), the effects seemed to grow larger. This can be explained by the fact that soil moisture and temperature uncertainties affect PBL and radiation processes, among others, directly. As height increases, more complex physics is involved, and thus the effects are enlarged. A lower level of improvement in forecast skill between the perturbation and control tests resides in the fact that there were too few members (four members for each test, due to limitations in computing resources); it is known that the skill of an ensemble forecast generally increases with an increase in the number of members (Langford and Hendon, 2011), Moreover, from the ROC area score, it seems that considering the variation in land surface processes could slightly increase the resolution of global prediction systems.

### 4.2. *Continuous ranked probability score*

The continuous ranked probability score (CRPS) measures the difference between the forecasted and observed cumulative density functions of scalar variables (Candille et al., 2007). It evaluates both the reliability and resolution; furthermore, the CRP skill score (CRPSS) has an advantage of being sensitive to a whole range of values of the parameter of interest, that does not depend on predefined classes at the same time. It evaluates the characteristics of both the resolution and reliability. Similar to the ROC area, the CRPSSs for the perturbation experiments were slightly larger than for the control run at most forecast lead times (beyond day 5, Fig. 8). Likewise, a tendency was found for a higher CRPSS at higher levels (vertically) in the model, illustrating that perturbation of soil moisture and soil temperature contribute to an overall slight improvement of forecast skill in resolution and reliability. However, the replacement test produced a lower score than the other two at the lead times of 15 and 16 days, indicating that perturbations were too large, in comparison with the rescaled perturbation test.

### 4.3. *Measurements of the ensemble mean*

Due to the deficiency in the ensemble technique and the limited number of ensemble members, almost all current EPSs are under-dispersive, which remains as a great chal-

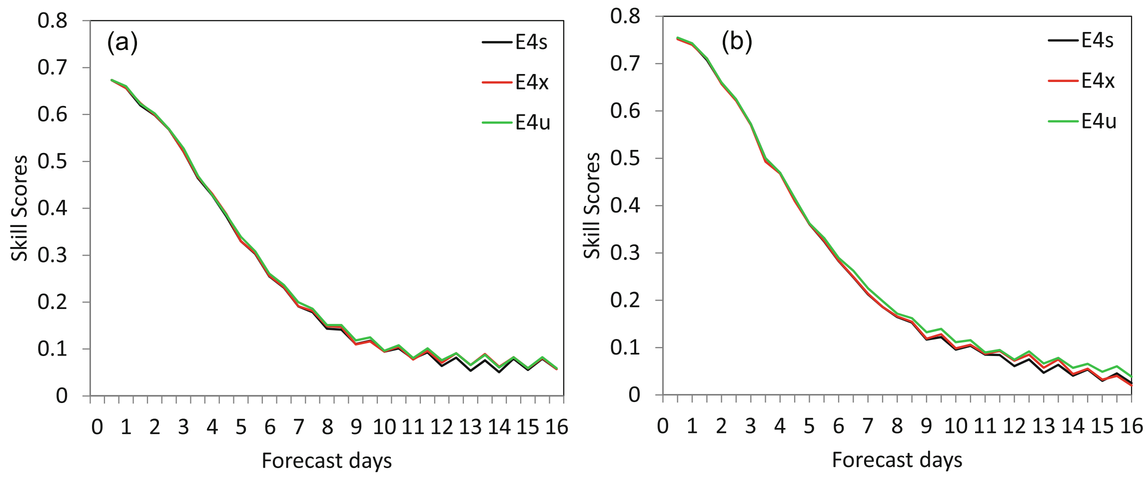


**Fig. 7.** ROC area for the rescaled perturbation (green), replacement perturbation (red) and control (black) experiments for (a) 1000 hPa geopotential height and (b) 500 hPa geopotential height, averaged over the verification domain (Eurasian continent) and over the verification period from 23 August to 24 September 2006 (E4s: control run; E4x: replacement run; E4u: rescaled run).

lenge in ensemble forecasts. For the ensemble verification score, it shows that the ensemble spread (distance between the ensemble mean and ensemble members) is less than the ensemble RMSE (distance between the ensemble mean and the analysis). Figure 9 compares the ensemble mean and spread of the 500 hPa and 1000 hPa geopotential heights. The results indicated that, unlike the slight improvement in forecast skill in terms of resolution and reliability, the perturbation of surface variables in the GEFS contributes to a noticeable reduction in forecast error, as well as an increase in ensemble spread in an under-dispersive system.

Besides the scores mentioned above, other verification methods were employed to evaluate the performance of the probability forecasts, and the results were similar. All in all, mostly positive results were obtained for the GEFS when considering the uncertainties of land surface processes. This is due to the fact that, unlike the direct perturbation of atmospheric variables, it takes time for land surface process uncertainties to play a role, through interactions between the atmosphere and the land surface, i.e., the characteristics of soil are much more stable than those of air, and so there is a clear time delay in the saturation of soil spread. Finally, due to the limitations of computing resources, too few ensemble members were used in the experiments (four members for each group), which more than likely affected the results (Sutton et al., 2006). Furthermore, a more wisely devised perturbation scheme, more ensemble members, and a longer experiment period are expected to improve the forecast skill.

## 5. Discussion and conclusions

Land surface processes have a profound impact on the overlying atmosphere on all time scales, including the storm scale, meso-scale, weather, sub-seasonal to seasonal, and climate scales. This study took into account the uncertainties in land surface processes by adding a module into the NCEP's GEFS and testing the influence on its predictability. Three experiments were conducted, and the preliminary results can be summarized as follows:

(1) The variations of soil temperature and soil moisture in the GEFS were examined to illustrate the uncertainties in land surface processes. The spread of the soil states reflected the timescales of interactions between the atmosphere and land surface processes, ranging from fractions of a day to the seasonal scale. The ensemble spread reached a steady state immediately at the near-surface soil level; but with deeper soil underneath the surface, the time it took for the spread to saturate increased. Therefore, a successive integration period of

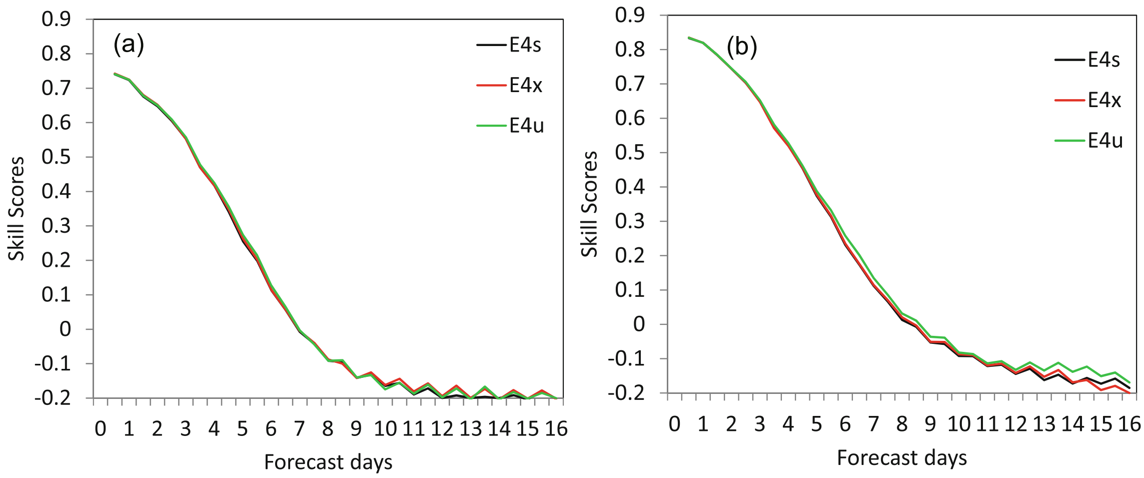


**Fig. 8.** As in Fig. 7 but for the CRPSS.

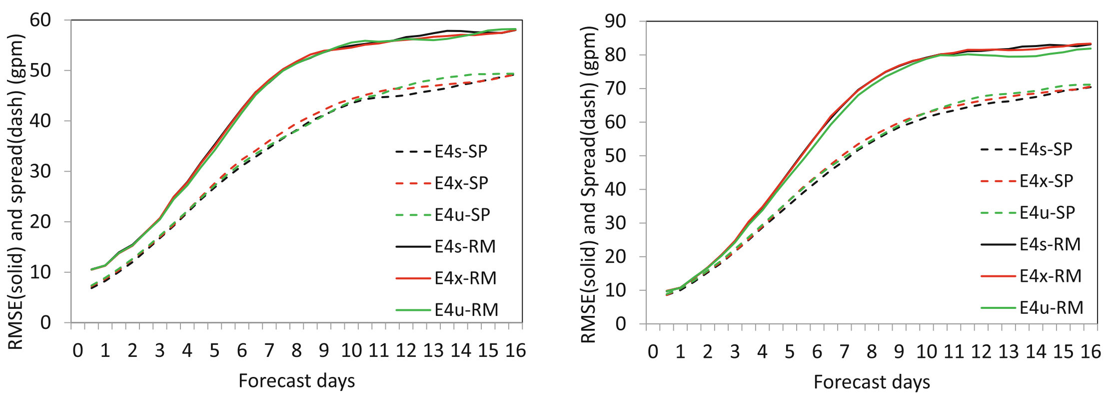


**Fig. 9.** As in Fig. 7 but for the evolution of spread (SP, dashed) and RMSE (RM, solid).

more than 6 months is required in the GEFS to fully represent the effects of land surface perturbation.

(2) Land surface process uncertainties resulted in large sensible and latent heat flux changes in the perturbation experiments compared to the control run, and the influences of land surface processes propagated to the upper troposphere via PBL processes, convection, and other activities. Locally, or in terms of day-to-day forecasting, there were great differences between the perturbed and control experiments.

(3) To assess the effects of land surface process perturbations on the GEFS, various scores, such as the ROC area, CRPSS, ensemble mean forecast error and spread, were calculated to evaluate the performance of the probability forecasts. The results indicated that the perturbation of surface variables in the GEFS contributes to slight improvement in forecast skill in terms of resolution and reliability, a noticeable reduction in forecast error, as well as an increase in ensemble spread in an under-dispersive system. The improvement is small at the surface, but the effect becomes increasingly obvious with depth due to interactions or feedback among surface processes and the free atmosphere. Considering the small number of ensemble members in the experiments, we expect the land surface perturbations to potentially have a greater impact in baroclinic zones, which is important for increasing ensemble spread in under-dispersive systems.

(4) Two different perturbation schemes were designed in this study. It seems that the rescaled experiment showed more skill than the replacement experiment, indicating that it is necessary to control the ranges of perturbation. Moreover, a state-of-the-art land surface perturbation might help to further improve the GEFS' forecast skills. For both schemes, the effects of interactions between land surface processes and the atmosphere differed with variables (soil moisture, soil temperature, geopotential height, temperature field, wind fields etc.), due to the timescales and mechanisms of interactions involved. Limited by computing resources, there were only four members for each ensemble group, which would have greatly affected the results. Therefore, this paper serves only as a preliminary exploration in this field. More complex and suitable methods need to be devised and applied to examine the effects of land surface process perturbations, such as the ET method for land surface processes, the perturbation of more variables (SST, sea-ice, near-surface temperatures, humidity etc.), a longer testing period, and more ensemble members.

***Acknowledgements.*** We are grateful to the computing resources and data provided by the NCEP. This research was supported by the National Fundamental (973) Research Program of China (Grant No. 2013CB430100),the Special Fund for Meteorological Scientific Research in the Public Interest (Grant No. GYHY201506005), and the National Natural Science Foundation of China (Grant Nos. 41475097, 41075079, 41275065 and 41475054) . Two anonymous reviewers provided constructive comments that helped improve the manuscript. Dr. Jun Du from NCEP provided much help with the structure of the manuscript, as well as helping to improve the English.

**REFERENCES**

Atger, F., 1999: The skill of ensemble prediction systems. *Mon. Wea. Rev.*, **127**, 1941–1953, doi: 10.1175/1520-0493(1999)127<1941:TSOEPS>2.0.CO;2.

Beljaars, A., G. Balsamo, A. Betts, and P. Viterbo, 2007: Atmosphere/surface interactions in the ECMWF model at high latitudes. *Proc. of ECMWF Seminar on Polar Meteorology*, 4–8 September 2006, Reading, ECMWF, 153–168.

Betts, A. K., J. H. Ball, A. C. M. Beljaars, M. J. Miller, and P. A. Viterbo, 1996: The land surface–atmosphere interaction: A review based on observational and global modeling perspectives. *J. Geophys. Res.*, **101**(D3), 7209–7225, doi: 10.1029/95JD02135.

Bishop, C. H., and Z. Toth, 1999: Ensemble transformation and adaptive observations. *J. Atmos. Sci.*, **56**, 1748–1765, doi: 10.1175/1520-0469(1999)056<1748:ETAAO>2.0.CO;2.

Buizza, R., M. Leutbecher, and L. Isaksen, 2008: Potential use of an ensemble of analyses in the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **134**(637), 2051–2066, doi: 10.1002/qj.346.

Buizza, R., M. Leutbecher, L. Isaksen, and J. Haseler, 2010: Combined use of EDA- and SV-based perturbations in the EPS. ECMWF Newsletter No. 123, 22–28. [Available from http://www.ecmwf.int/sites/default/files/elibrary/2010/14602-newsletter-no123-spring-2010.pdf].

Campana, K., and P. Caplan, 2005: Technical procedures bulletin for the T382 Global Forecast System. Environmental Modeling Center, National Centers for Environmental Prediction. [Available online at http://www.emc.ncep.noaa.gov/gc_wmb/Documentation/TPBoct05/T382.TPB.FINAL.htm].

Candille G., C. Côté, P. L. Houtekamer, and G. Pellerin, 2007: Verification of an ensemble prediction system against observations. *Mon. Wea. Rev.*, **135**, 2688–2699, doi: 10.1175/MWR3414.1.

Charron, M., G. Pellerin, L. Spacek, P. L. Houtekamer, N. Gagnon, H. L. Mitchell, and L. Michelin, 2010: Toward random sampling of model error in the Canadian ensemble prediction system. *Mon. Wea. Rev.*, **138**, 1877–1901, doi: 10.1175/2009MWR3187.1.

Chen, F., and Coauthors, 1996: Modeling of land surface evaporation by four schemes and comparison with FIFE observations. *J. Geophys. Res.*, **101**(D3), 7251–7268.

Chen, F. and J. Dudhia, 2001: Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Mon. Wea. Rev.*, **129**, 569–585, doi: 10.1029/95JD02165.

Davis, C. A., K. W. Manning, R. E. Carbone, S. B. Trier, and J. D. Tuttle, 2003: Coherence of warm-season continental rainfall in numerical weather prediction models. *Mon. Wea. Rev.*, **131**, 2667–2679, doi: 10.1175/1520-0493(2003)131<2667:COWCRI>2.0.CO;2.

Deng, G., and Coanthors, 2010: Development of mesoscale ensemble prediction system at National Meteorological Center. *Journal of Applied Meteorological Science*, **21**(5), 513—523, doi: 10.3969/j.issn.1001-7313.2010.05.001. (in Chinese)

Deng, G., H. Tian, X. L. Li, J. Chen, J. D. Gong, and M. Y. Jiao, 2012: A comparison of breeding and ensemble transform vectors for global ensemble generation. *Acta Meteorologica Sinica*, **26**(1), 52–61, doi: 10.1007/s13351-012-0105-4. (in Chinese)

Gao, J. D., and M. Xue, 2008: An efficient dual-resolution approach for ensemble data assimilation and tests with simulated Doppler radar data. *Mon. Wea. Rev.*, **136**, 945–963, doi: 10.1175/2007MWR2120.1.

Global Climate and Weather Modeling Branch, 2003: The GFS Atmospheric Model. NCEP Office Note 442, Environmental Modeling Center. [Available online at http://www.nws.noaa.gov/ost/climate/STIP/AGFS_DOC_1103.pdf]

Hamill, T. M., and S. J. Colucci, 1997: Verification of Eta-RSM short-range ensemble forecasts. *Mon. Wea. Rev.*, **125**, 1312–1327, doi: 10.1175/1520-0493(1997)125<1312:VOERSR>2.0.CO;2.

Hamill, T. M., C. Snyder, and R. E. Morss, 2000: A comparison of probabilistic forecasts from bred, singular-vector, and perturbed observation ensembles. *Mon. Wea. Rev.*, **128**, 1835–1851, doi: 10.1175/1520-0493(2000)128<1835:ACOPFF>2.0.CO;2.

Hou, D., and Coauthors,., 2016: Current State of NCEP Global Ensemble. 7th NCEP Ensemble Users Workshop, 13-15 June 2016, College Park, MD.

Langford, S., and H. H. Hendon, 2011: Assessment of international seasonal rainfall forecasts for Australia and the benefit of multi-model ensembles for improving reliability. CAWCR Technical Report No. 039.

Lewis, F. R., 2007: *Weather Prediction by Numerical Process*. 2nd ed. Cambridge University Press, 236pp.

Li, X. F., and C. Z. Zou, 2009: Effects of sea surface temperature, radiation, cloud microphysics, and diurnal variations on vertical structures of tropical tropospheric temperature: A two-dimensional equilibrium cloud-resolving modeling study. *Meteor. Atmos. Phys.*, **105**, 85–98, doi: 10.1007/s00703-009-0039-2.

McLay, J. G., M. K. Flatau, C. A. Reynolds, J. Cummings, T. Hogan, and P. J. Flatau, 2012: Inclusion of sea-surface temperature variation in the U. S. Navy ensemble-transform global ensemble prediction system. *J. Geophys. Res.: Atmos.*, **117**, D19120, doi: 10.1029/2011JD016937.

Mullen, S. L., and R. Buizza, 2002: The impact of horizontal resolution and ensemble size on probabilistic forecasts of precipitation by the ECMWF ensemble prediction system. *Wea. Forecasting*, **17**, 173–191, doi:10.1175/1520-0434(2002)017<0173:TIOHRA>2.0.CO;2.

National Research Council, 1998: *The Atmospheric Sciences Entering the Twenty-first Century*. National Academy Press, 364 pp.

Nicholson, S. E., 1988: Land Surface Atmosphere Interaction: Physical processes and Surface Changes and their Impact. *Progress in Physical Geography*, **12**, 36–65, doi: 10.1177/030913338801200102.

Pan, H.-L., and L. Mahrt, 1987: Interaction between soil hydrology and boundary layer development. *Bound.-Layer Meteor.*, **38**, 185–202, doi: 10.1007/BF00121563.

Pielke, R. A., 2001: Influence of the spatial distribution of vegetation and soils on the prediction of cumulus Convective rainfall. *Rev. Geophys.*, **39**, 151–177, doi: 10.1029/1999RG000072.

Richardson, D. S., 2000: Skill and relative economic value of the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **126**, 649–667.

Stanski, H. R., L. J. Wilson, and W. R. Burrows, 1989: Survey of common verification methods in meteorology. World Weather Watch Tech. Rept. No.8, WMO/TD No.358, WMO, Geneva, 114 pp.

Sutton, C., T. M. Hamill, and T. T. Warner, 2006: Will perturbing soil moisture improve warm-season ensemble forecasts? A proof of concept. *Mon. Wea. Rev.*, **134**, 3174–3189, doi: 10.1175/MWR3248.1.

Tennant, W., and S. Beare, 2014: New schemes to perturb sea-surface temperature and soil moisture content in MOGREPS. *Quart. J. Roy. Meteor. Soc.*, **140**, 1150–1160, doi: 10.1002/qj.2202.

Toth, Z., O. Talagrand, and Y. J. Zhu, 2006: The attributes of forecast systems: a general framework for the evaluation and calibration of weather forecasts. Tim Palmer and Renate Hagedorn, Predictability of Weather and Climate, Ed., Cambridge University Press, 584–595.

Viterbo, P., and A. C. M. Beljaars, 1995: An improved land surface parameterization scheme in the ECMWF model and its validation. *J. Climate*, **8**, 2716–2748, doi: 10.1175/1520-0442(1995)008<2716:AILSPS>2.0.CO;2.

Wang, X. G., and C. H. Bishop, 2003: A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes. *J. Atmos. Sci.*, **60**, 1140–1158, doi: 10.1175/1520-0469(2003)060<1140:ACOBAE>2.0.CO;2.

Wang, Y., A. Kann, M. Bellus, , J. Pailleux, and C. Wittmann, 2010: A strategy for perturbing surface initial conditions in LAMEPS. *Atmos. Sci. Lett.*, **11**(2), 108–113, doi: 10.1002/asl.260.

Wei, M. Z., Z. Toth, R. Wobus, Y. J. Zhu, C. H. Bishop, and X. G. Wang, 2005: Ensemble Transform Kalman Filter-based ensemble perturbations in an operational global prediction system at NCEP. *Tellus A*, **58**(1), 28–44, doi: 10.1111/j.1600-0870.2006.00159.x.

Wei, M. Z., Z. Toth, R. Wobus, and Y. J. Zhu, 2008: Initial perturbations based on the ensemble transform (ET) technique in the NCEP global operational forecast system. *Tellus A*, **60**, 62–79, doi: 10.1111/j.1600-0870.2007.00273.x.

Weng, Y. H., and F. Q. Zhang, 2012: Assimilating airborne Doppler radar observations with an ensemble Kalman filter for convection-permitting hurricane initialization and prediction: Katrina (2005). *Mon. Wea. Rev.*, **140**, 841–859, doi: 10.1175/2011MWR3602.1.

Xue, M., M. J. Tong, and K. K. Droegemeier, 2006: An OSSE framework based on the ensemble square root Kalman filter for evaluating impact of data from radar networks on thunderstorm analysis and forecasting. *J. Atmos. Oceanic Technol.*, **23**, 46–66, doi: 10.1175/JTECH1835.1.

Yuan, H. J., J. A. McGinley, P. J. Schultz, C. J. Anderson, and C. G. Lu, 2008: Short-range precipitation forecasts from time-lagged multimodel ensembles during the HMT-West-2006 campaign. *Journal of Hydrometeorology*, **9**, 477–491, doi: 10.1175/2007JHM879.1.

Yuan, H. L., C. G. Lu, J. A. McGinley, P. J. Schultz, B. D. Jamison, L. Wharton, and C. J. Anderson, 2009: Evaluation of short-range quantitative precipitation forecasts from a time-lagged multimodel ensemble. *Wea. Forecasting*, **24**, 18–38, doi: 10.1175/2008WAF2007053.1.