For How Long Should What Data Be Assimilated for the Mesoscale Forecasting of Convection and Why? Part I: On the Propagation of Initial Condition Errors and Their Implications for Data Assimilation

FRÉDÉRIC FABRY

McGill University, Montreal, Quebec, Canada

JUANZHEN SUN

National Center for Atmospheric Research, Boulder, Colorado*

(Manuscript received 24 November 2008, in final form 22 June 2009)

ABSTRACT

Data assimilation is used among other things to constrain the initial conditions of weather forecasting models by fitting the model fields to observations made over a certain time interval. In particular, it tries to tie incomplete data with model constraints to detect and correct for initial condition errors. This is possible only if initial condition errors leave their signature on the data assimilated and if the model is capable of faithfully reproducing such signatures. Using simulations of the evolution of convective storms in the Great Plains over an active 6-day period, the propagation of initial condition errors to other variables as well as their effect on the accuracy of the forecasts were investigated. Increasing the assimilation time window boosts the ability of assimilation systems to detect a variety of initial condition errors; however, limits to the predictability of convective events impose a maximum assimilation period that is a function of the type of measurements assimilated as well as of the type of errors one tries to correct for. These findings are then used to suggest changes in assimilation approaches to take into account the different predictability times of the model fields constrained by assimilation.

1. Perspective, paradigm, and plan

"There are too many types of data, too many new ones appearing every year, and there are too few people to work on their assimilation!" This outcry has been heard many times from data assimilation researchers, in particular from those working at the mesoscale. And it provided one of the initial impetuses for this work. Could there be a way to prioritize data assimilation efforts? On what basis? To approach this problem, one is required to think about data assimilation in a more conceptual way than usual. In particular, in this work, a greater emphasis will be put on the nature and characteristics of the data to be assimilated or of the model fields to be constrained.

Data assimilation is explicitly designed to constrain model variables with noisy measurements. But for data assimilation to succeed, three additional conditions must be met well enough. First, the difference between the assumed atmospheric state $\mathbf{x}'$ and the true atmospheric state $\mathbf{x}$ must result in a measurable difference between what the true observations $\mathbf{y}'$ would be with an atmospheric state $\mathbf{x}'$ and the true observations $\mathbf{y}$. In the absence of an observational signal, data assimilation will fail. Second, given an atmospheric state $\mathbf{x}$, a model can reproduce the true observations $\mathbf{y}$ and their time evolution. If the model-generated observations $H(\mathbf{x})$ cannot faithfully replicate the measured fields $\mathbf{y}$, providing correct observations to the assimilation system will only lead to an inaccurate initialization of the model. Failure to reproduce observations may come from both an inability to predict the atmospheric state and from an inaccurate observation operator. Third, the data assimilation system can use the difference between

* The National Center for Atmospheric Research is sponsored by the National Science Foundation.

Corresponding author address: Frédéric Fabry, Department of Atmospheric and Oceanic Sciences, McGill University, 805 Sherbrooke St. West, Montreal, QC H3A 2K6, Canada.
E-mail: frederic.fabry@mcgill.ca

DOI: 10.1175/2009MWR2883.1

the simulated observations $H(\mathbf{x}')$ and the true observations $\mathbf{y}$ to adjust the model state from $\mathbf{x}'$ to $\mathbf{x}$. This is a challenging undertaking given that the problem has a very large number of dimensions, and the process relies heavily on the ability of the data assimilation system to adjust the model state $\mathbf{x}'$ in the right direction in an attempt to get to the true atmospheric state $\mathbf{x}$. The presence of a signal in data, an accurate model, and an optimally functioning data assimilation system are all required to make data assimilation succeed.

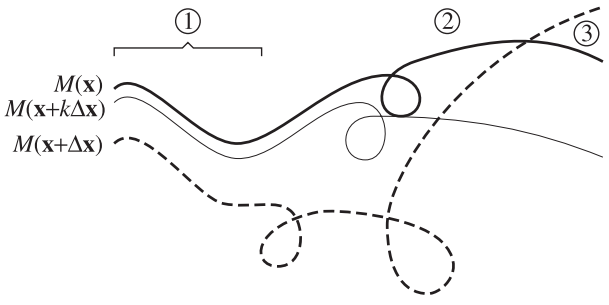
In response to the need to constrain the initial conditions of forecasting models, several data assimilation approaches were developed. Two somewhat different situations can be distinguished. In one case, enough data are available to constrain all variables more or less directly without having to rely on the time evolution of the weather events to close the system. Assimilation methods such as the three-dimensional variational data assimilation (3D-Var; Kalnay 2003, and references therein), or when one uses four-dimensional variational data assimilation (4D-Var) mainly to constrain the model at the precise time the observation was made, are based on this paradigm. In the other case, data are available to constrain only some of the variables, and the remaining ones are constrained using the model's ability to simulate them to be compatible with the time evolution of available observations. Methods such as 4D-Var (e.g., Talagrand 1997) are designed to handle such situations. For mesoscale forecasting, the second scenario generally applies: balloon soundings and other measurements of all model fields throughout the atmosphere are sparse in space and time; one must hence rely on the remote sensing of the time evolution of some of the fields to fill the gaps. The discussion to follow assumes the latter, with the emphasis being put on the forecasting of summertime convective weather.

If observations in a few variables are to gradually constrain the initial conditions of all variables, there must be ways to link the unobserved variables to the observed ones. Specifically, a change or an error in any unobserved variable must result in a mismatch between observed variables and their simulated value; this must occur within the duration of the assimilation interval during which observations at two or more times are used to constrain the model before a forecast is made. And that mismatch must be large enough to exceed measurement noise. This imposes a minimum assimilation interval, one that may depend on the field to be constrained through model dynamics. For example, some errors in unobserved variables may rapidly transfer to observed variables, while others may not. If a perturbation in, say, humidity, propagates quickly to other variables such as cloud and precipitation content, then its effect can be detected by assimilating the data from a variety of instruments over a short period. However, if the perturbation remains essentially in moisture form for a long period, then unless very long data assimilation intervals are used, it can only be detected by instruments that actually measure a quantity directly related to moisture.

To complicate matters, the pace at which these errors move from one field to another is also a function of scale. Small-scale features evolve rapidly, while large-scale patterns evolve slowly. For a forecast to be as accurate as possible for a long time, large-scale patterns must be well determined since they are the ones with the longest time of predictability. Consequently, the assimilation interval must be long enough to permit the detection of the slowest-evolving errors of the largest modeled patterns. If one optimistically assumes that patterns larger than 500 km are well constrained by upper-air synoptic observations, and that 500-km patterns have a lifetime of about 12 h (e.g., Lorenz 1969; Zhang et al. 2007), then one must assimilate data for a significant fraction of that time to observe an evolution in 500-km-scale patterns sufficient to constrain unobserved fields at that scale.

But for small-scale features, a long interval may be inadequate. Much like the speed of evolution, the time of predictability is also a function of scale as well as of model accuracy. If the model is incapable of forecasting the evolution of intense small-scale patterns over the duration of the assimilation interval, trying to assimilate such data over the assimilation interval is a counterproductive exercise. To illustrate this, let us assume that $M(\mathbf{x})$ is the current model state, and that $M(\mathbf{x} + \Delta\mathbf{x})$ is the truth one seeks (Fig. 1). To minimize the mismatch between the model state at the beginning of the assimilation interval and the information provided by observations, one generally first tries to find in what direction should $\mathbf{x}$ be pushed by computing $\partial M/\partial\mathbf{x}$. In the example of Fig. 1, moving the modeled initial conditions from $\mathbf{x}$ toward $\mathbf{x} + k\Delta\mathbf{x}$, $0 < k < 1$, improves the match with the true initial conditions $\mathbf{x} + \Delta\mathbf{x}$. But beyond the time of predictability, there will be many circumstances where such a move will worsen the fit with observations; in Fig. 1, these are illustrated as the contradictory regions. In these regions, the assimilation system will actually find it advantageous to move away from $\mathbf{x} + \Delta\mathbf{x}$ in order to improve the fit with observations. As time progresses and $k[M(\mathbf{x} + \Delta\mathbf{x}) - M(\mathbf{x})]$ and $k[\Delta\mathbf{x}\,\partial M/\partial\mathbf{x}]$, illustrated here as $[M(\mathbf{x} + k\Delta\mathbf{x}) - M(\mathbf{x})]$, become completely decorrelated, there will be a 50% chance that the assimilation of new data moves the initial conditions away from the correct solution.

It is not immediately clear if all model fields and datasets are affected identically by these considerations.



① Linear regime: $M(\mathbf{x}+\Delta\mathbf{x}) \approx M(\mathbf{x}) + \Delta\mathbf{x}\,(\partial M/\partial \mathbf{x})$
$\rightarrow M(\mathbf{x}+k\Delta\mathbf{x}) - M(\mathbf{x}) \approx k\,[M(\mathbf{x}+\Delta\mathbf{x}) - M(\mathbf{x})]$

② Nonlin. regime: $M(\mathbf{x}+k\Delta\mathbf{x}) - M(\mathbf{x}) \neq k\,[M(\mathbf{x}+\Delta\mathbf{x}) - M(\mathbf{x})]$

③ Contradictory region: Adjusting $\mathbf{x}$ towards $\mathbf{x}+k\Delta\mathbf{x}$ at the initial time worsens the fit with the data from $M(\mathbf{x}+\Delta\mathbf{x})$

FIG. 1. Schematic illustration of the trajectory in phase space of three model runs M starting with three different initial conditions: $\mathbf{x}$ the original initial conditions, $\mathbf{x} + \Delta\mathbf{x}$ the initial conditions that would minimize the errors with respect to observations, and $\mathbf{x} + k\Delta\mathbf{x}$. The latter represents initial conditions that an algorithm trying to minimize the difference between $M(\mathbf{x})$ and $M(\mathbf{x} + \Delta\mathbf{x})$ might try in an attempt to evaluate whether moving the initial conditions from $\mathbf{x}$ toward $\mathbf{x} + \Delta\mathbf{x}$ improves the fit with observations. Three regimes can be observed: 1) a linear regime, where $M(\mathbf{x} + k\Delta\mathbf{x})$ is $k/(1 - k)$ closer to $M(\mathbf{x})$ than to $M(\mathbf{x} + \Delta\mathbf{x})$, such a regime makes it easy for error minimization algorithms to adjust initial conditions from $\mathbf{x}$ to $\mathbf{x} + \Delta\mathbf{x}$; 2) a nonlinear regime, where $M(\mathbf{x} + k\Delta\mathbf{x})$ is not $k/(1 - k)$ closer to $M(\mathbf{x})$ than to $M(\mathbf{x} + \Delta\mathbf{x})$, such a regime makes it harder for error minimization algorithms to adjust initial conditions from $\mathbf{x}$ to $\mathbf{x} + \Delta\mathbf{x}$; 3) a contradictory regime, a special case of the nonlinear regime, where changing initial conditions from $\mathbf{x}$ to $\mathbf{x} + k\Delta\mathbf{x}$, i.e., closer to the $\mathbf{x} + \Delta\mathbf{x}$ that minimizes the difference with observations, actually worsens the fit between M and $M(\mathbf{x} + \Delta\mathbf{x})$, possibly misleading the minimization algorithm to go away from the right answer.

Some fields such as cloud water evolve in more complex ways, especially via the release of convective instabilities, as a result of which nonlinearities build up faster at convective scales than at synoptic scales (Hohenegger and Schär 2007). More fundamentally, it has not been quantified how easily and clearly one gets a signal of an error in one field by observing other fields. For example, the evolution of some fields such as winds and temperature are dynamically coupled, but no such coupling exists for water vapor amounts. An additional confounding effect is the sometimes complex or indirect link between observations and model fields. Finally, errors in some fields may have greater impacts on forecast quality than others. All these factors influence the performance of data assimilation systems and should affect the choices of approaches and observed variables used in assimilation. Yet at present, computing resource and data availability considerations seem to be the main factors taken into account.

Additional factors to bear in mind should include: Based on physical considerations, what data should be assimilated, and for how long? Are there datasets that provide more or better information or do they target fields or regions in the atmosphere whose uncertainty has a greater impact on the final forecasts? To answer these questions, two broad topics must be investigated. One is related with the propagation of initial condition errors from one variable to the next: How well can one detect errors in one field by observing the time evolution of another? How does predictability affect the assimilation period to be used for different fields? Is the resulting assimilation period sufficient? And in any case, which initial condition error(s) cause the worst forecasts and should be detected in priority? This paper will provide some answers to these questions. The second set of issues deals with the ability of instruments to extract the wanted signal. What are different instruments measuring and how well? How are their data affected by the issues discussed above? Is the wanted signal detectable amidst measurement noise? Part II (Fabry 2010) will consider these questions.

The data assimilation community has certainly recognized the problems of model error and atmospheric predictability in general and dealt with them by designing a data assimilation system suitable for specific applications. For example, a 6–12-h window was used for large-scale data assimilation (Rabier et al. 2000; Zou and Kuo 1996) while a window as small as 12 min was used for the assimilation of convective-scale data assimilation using radar observations (Sun 2005). The incremental data assimilation approach was introduced by Courtier et al. (1994) to handle the issue of the nonlinear growth of model errors. Various other approaches were proposed to deal with model errors in data assimilation systems (e.g., Derber 1989; Zupanski 1997; Bennett et al. 1996). Finally, Trémolet (2006) proposed a weak-constraint 4D-Var in which subwindows were implemented and the model was used as a strong constraint only in each subwindow. In this paper, we do not intend to develop a methodology to treat the model errors, but rather to take a step back to examine how the forecast errors propagate with time and from one variable to another, with the motivation that the findings can provide guidance for the design of future data assimilation systems and the choice of observations to be assimilated.

The approach used in this work is as follows. First, the impact of different initial condition errors is investigated. Some types of errors may result in large forecast errors, and are therefore more important to detect. Others may cause limited forecasts errors, at least over some time scales, and if they may not be ignored, they should be deemphasized. Next, one must look at the properties of the transfer of the errors from one variable to another. These properties include the strength and speed of the interactions between variables, as well as how

predictability issues affect our ability to use that information in the context of data assimilation. This exercise will allow us to determine which field(s) should be constrained in priority as well as provide some hints as how to best assimilate data. Finally, we will reflect on the implications of the results on data assimilation, especially at the mesoscale. In this work, we chose to restrict the focus on data assimilation considerations and not on storm dynamics questions even though the questions asked and approach used are conducive to both types of work.

2. Data generation

First, an appropriate dataset must be obtained to help answer our questions. For this project, knowledge as complete as possible of the entire system is needed. This alone rules out using true but incomplete atmospheric measurements and calls for a modeling-based experiment. We hence resort to an identical twin experiment (Daley 1991). First, a series of plausible convective events is simulated. These control runs will constitute "the truth." Then, initial conditions are perturbed by plausible errors, and these are used to generate "the forecasts" that will be compared with the truth. These forecasts will have errors whose magnitude can be evaluated as a function of the type of errors in the initial conditions. These runs can then also be used to study the transfer of initial condition errors, as well as to simulate measurements by different sensors.

a. Model, domain, and period used

To perform this numerical experiment, a significant number of plausible convective events needed to be simulated. The tool used was the Advanced Research Weather Research and Forecasting (WRF) model, version 2.2 (Skamarock et al. 2005), with Thompson et al. (2004) microphysics, the Yonsei University PBL, the radiation scheme following Mlawer et al. (1997) for the longwave and the approach used by Dudhia (1989) for the shortwave radiation, and the Noah land surface model (Chen and Dudhia 2001). The domain covered a 1600 km × 1600 km area centered on the Great Plains similar to the one used by Xiao and Sun (2007). The model runs have a 4-km resolution in the horizontal and 28 vertical levels. The choice of these parameters was dictated by the need of having enough physically plausible simulations to have statistical significance without making the project too large to be completed.

To add to the realism, the runs were initialized using analyses from actual events that occurred in the Great Plains between 10 and 16 June 2002. This period was particularly rich in convective events: in the center of the domain, where the International H_2O Project (IHOP_2002) field experiment was taking place (Weckwerth et al. 2004), a large variety of squall lines and multicell storms raged, having a variety of initiation mechanisms (Wilson and Roberts 2006); in addition, the region was swept by a large mesoscale convective system and more storms were triggered outside of the IHOP_2002 domain next to the Gulf Coast and near the foothills of Colorado and New Mexico. The net result is that in a short period of time, a microcosm of convective events was observed and large amounts of precipitation were recorded in the domain and simulated by the control runs. Note that the rainfall amounts shown in Fig. 2 were accumulated over the sixteen 12-h control runs described below.

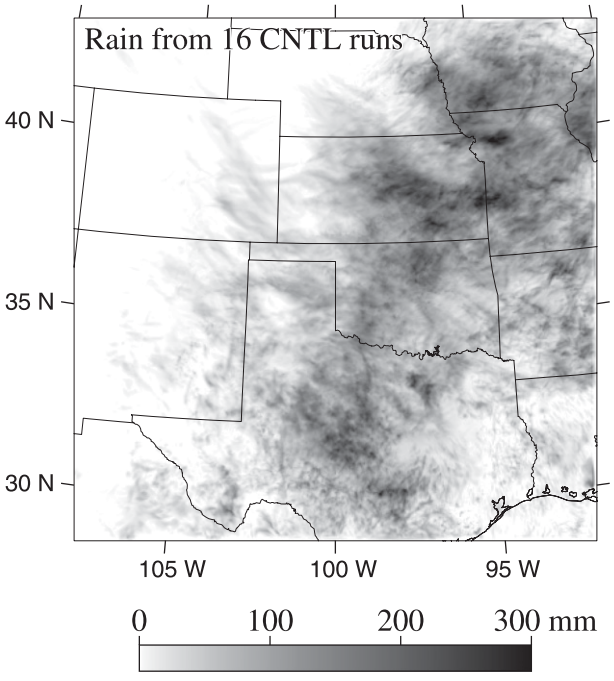


FIG. 2. Simulation domain and model rainfall accumulation of the 16 control runs used in this study.

b. The model runs

In total, 16 sets of runs were done in this work covering the period between 10 and 16 June 2002, one every 9 h. Each set of runs starts with a 3-h run initialized by the analysis valid at $T - 3$ h made by the National Centers for Environmental Prediction (NCEP) for its Eta model (Fig. 3). The goal of this run is to spin up the dynamics and generate believable fields down to the model resolution. The resulting fields produced at time T are then used as initial conditions to start our 12-h control run that will act as our truth. These initial conditions were then perturbed to simulate 10 different types of initialization errors, and each set of these modified initial conditions was used as a starting point to what will be referred to as a perturbed run. Details on the nature of these

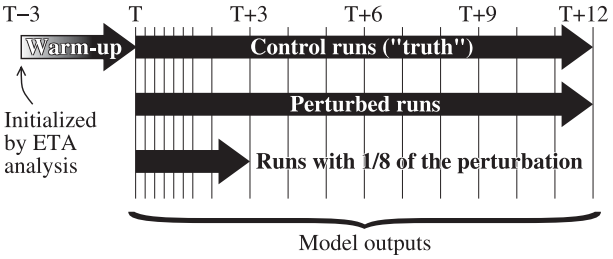


FIG. 3. Timeline of the initialization and outputs of the 16 control runs and the associated perturbed runs. Vertical lines represent times for which a model output was obtained.

perturbations will follow in section 2c. Outputs from the control and the perturbed runs are available every 15 min up to $T + 90$ min and every hour between $T + 2$ h and $T + 12$ h. In addition, for each 12-h perturbed run, a 3-h run was done with 1/8th of the perturbation size used in the 12-h run. These reduced perturbation runs were used to test the linearity and predictability of the forecast errors from the perturbed runs.

c. *Initial condition perturbations*

There are many possible ways one can perturb initial fields. In this work, we have chosen to perturb individual fields to an extent that is comparable with the uncertainty with which these fields are generally known. Because our goal is to estimate the ability of different sensors to detect reasonable initialization errors, perturbing initial fields in a way that is comparable to our lack of knowledge of these fields seemed like a natural choice. The choice of perturbing one field at a time was one of convenience, but it also allowed us to evaluate the impact of the uncertainty in each field on the quality of the forecast. However, we could not investigate the effect of possible interactions between fields, which can be important especially if the errors between fields are correlated.

Five sets of fields were perturbed: winds, temperature, humidity, condensates (i.e., clouds and precipitation), and soil moisture as a proxy for surface properties. But both the processes by which forecast uncertainties grow and the ability of different instruments to measure atmospheric fields also depend on altitude. We then settled on 10 different types of perturbations: low-level winds, midlevel winds, high-level winds, low-level temperatures, midlevel temperatures, high-level temperature, low-level moisture, midlevel moisture, whole-atmosphere condensates, and soil moisture at all depths.

A first challenge was to find estimates of the uncertainties in our knowledge of these fields. This implied not only determining the magnitude of the uncertainties, but also the correlation structure of these fields in the horizontal and the vertical. For many fields, at least one reference was found on which to base our perturbations: the spatial structure of mid- and high-level temperatures and winds were well characterized by Nastrom and Gage (1985); low-level temperature and humidity structure were measured by Fabry (2006) and Lenschow and Sun (2007), the latter also describing the scale-dependence of low-level winds. But for several others, information is extremely limited. For condensates, we used the spatial structure data of drop size distribution uncertainty from G.W. Lee (2006, personal communication). We however failed to find useful information on the magnitude of the uncertainty in low-level winds and on the spatial properties of midlevel humidity and soil moisture. For those, we had to make reasonable guesses. Each set of perturbations was generated by filtering a Gaussian-distributed noise to make its power spectra follow the curves in Fig. 4; the result was then scaled in such a way that the standard deviation σ of the resulting perturbation matched the values plotted on the same figure.

While information was found on the horizontal structure of different fields, we were less lucky with their vertical structure. We expect the vertical correlation of the small-scale structures in these fields to be high, but found no quantitative information. The solution adopted was to assume perfect correlation of the errors in the vertical, and the perturbations obtained above were multiplied by a shaping function $f(\eta)$ such that

$$
\begin{aligned}
f(\eta) &= \cos^2\left[\frac{\pi}{.35}(1-\eta)\right], && \text{for} \quad .825 \le \eta \le 1, \quad \text{for low-level perturbations;}\\
f(\eta) &= \cos^2\left[\frac{\pi}{.35}(.675-\eta)\right], && \text{for} \quad .5 \le \eta \le .85, \quad \text{for mid-level perturbations;}\\
f(\eta) &= \cos^2\left[\frac{\pi}{.35}(.325-\eta)\right], && \text{for} \quad .15 \le \eta \le .5, \quad \text{for high-level perturbations; and}\\
f(\eta) &= 1, && \text{for whole-atmosphere or whole-soil perturbations,}
\end{aligned}
\tag{1}
$$

where η is the Eta level in the WRF model defined by $\eta = (P - P_{\text{top}})/(P_o - P_{\text{top}})$, with P being the pressure at a given level, P_o is the surface pressure in the same column, and P_{top} is the pressure at the top of the model, set at 5 kPa. These shaping functions confine the low-level perturbations to the bottom 1.5 km of the atmosphere, the midlevel perturbations to the bottom half of the free troposphere away from the boundary layer, and the high-level perturbations to the top half of the free troposphere.

Hence, for each of the perturbed runs, the field to be perturbed was modified by adding the correlated noise specified above and in Fig. 4, and weighted by the shaping function in (1). This perturbation was made directly on the model field itself without the use of the WRF data assimilation system. This was done to eliminate any constraints and any possible modification that a data assimilation system might set. The WRF model reacted reasonably well to these forced perturbations, except for the temperature ones. For these, we adjusted the spacing between geopotential levels to account for the fractional change in temperature as well as modified the surface pressure to minimize the acoustic waves and Brunt–Väisälä oscillations that were triggered by these perturbations.

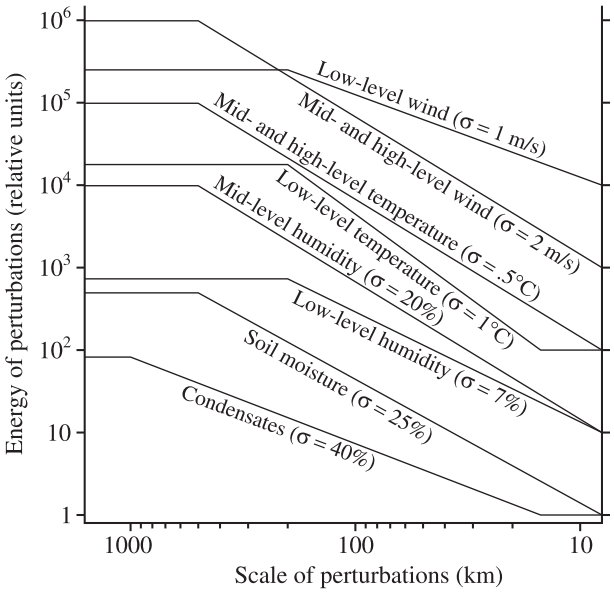


FIG. 4. Power spectra and magnitude of the different perturbations introduced in the initial conditions of the perturbed runs.

3. Forecast errors with different perturbations

a. Defining a unified model error

Our next task was to define an error equation to characterize model errors in winds, temperature, humidity, and precipitation. Ideally, one would like to find some reasonable equivalence between the magnitudes of the errors from each field. A great starting point is provided by Talagrand (1981) and Ehrendorfer and Errico (1995) who studied the energy difference between two similar model states, one at rest and one perturbed, and obtained:

$$E = \frac{1}{2}\int_D \left[\Delta u^2 + \Delta v^2 + \frac{c_p}{T_r}\Delta T^2 + RT_r\left(\frac{\Delta p_s}{p_r}\right)^2\right] dD, \quad (2)$$

where the differences in zonal wind Δu and meridional wind Δv and the differences in temperature ΔT and in surface pressure Δp_s are combined with the specific heat at constant pressure c_p, the gas constant R for dry air, a reference temperature T_r, and a reference pressure p_r, integrated over the whole domain D, to yield the difference in the energy E between two model states. Equation (2) is useful for two reasons. First, it offers a common "currency," energy, to gauge the equivalence between errors. Second, it proposes a framework and a method to compare the effect of differences in wind and of differences in heat. With it, two error terms can be defined, the kinetic energy difference (KED) and the thermal energy difference (TED) between states:

$$\text{KED} = \frac{1}{2}\int_D (\Delta u^2 + \Delta v^2)\, dD \quad (3)$$

and

$$\text{TED} = \frac{c_p}{2T_r}\int_D \Delta T^2\, dD. \quad (4)$$

Following Ehrendorfer and Errico (1995), $T_r = 270$ K was used. In subsequent work, Ehrendorfer et al. (1999) also derived an error term for the latent energy difference (LED) that we also used:

$$\text{LED} = \frac{L^2}{2c_p T_r}\int_D \Delta r_v^2\, dD, \quad (5)$$

with L being the latent heat of vaporization and Δr_v being the change in the vapor mixing ratio.

It proved difficult to find an energy equivalence for differences in condensates that was compatible with the framework established by Talagrand (1981). In the end, we chose to use the potential energy of hydrometeors as our metric, on the grounds that the drag caused by falling raindrops would result in the transfer of the potential energy of falling hydrometeors into kinetic energy of air. Hence, the energy differences due to condensates (CED) was set to

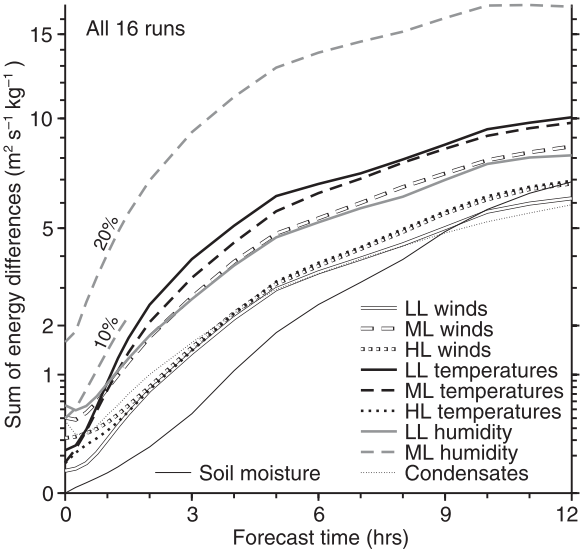


FIG. 5. SED averaged over all 16 runs between the perturbed and the control runs as a function of the forecast time for the different types of initial condition errors simulated. Note the unusual parabolic *y* axis. In the legend, LL stands for low-level, ML stands for midlevel, and HL stands for high-level.

$$\mathrm{CED} = g \int_D h|\Delta r_r + \Delta r_s + \Delta r_g| \, dD, \qquad (6)$$

with g being the acceleration due to gravity, h being height, and Δr_r, Δr_s, and Δr_g being the changes in the mixing ratios of rain, snow, and graupel, respectively. Note that since liquid and ice clouds do not fall significantly, they were excluded from the calculation. The resulting energy is probably overestimated compared to the others, but no convincing rationale was found to reduce it.

Finally, most of the analysis of model errors will be done using the sum of energy differences (SED), defined as

$$\mathrm{SED} = \mathrm{KED} + \mathrm{TED} + \mathrm{LED} + \mathrm{CED}. \qquad (7)$$

The integration of model energy differences was limited to the troposphere, and all results are presented in terms of energy differences per unit mass of dry air.

b. Average model errors versus forecast time

The control and perturbed runs were first used to quantify the forecast errors arising from the different types of perturbations. Originally, we simply wanted to use these errors to weigh the ability of different sensors to detect a wrong forecast with the magnitude of the forecast errors resulting from each given scenario. But a quick look at forecast errors proved to be interesting per se.

Figure 5 shows the SED per unit mass of dry air as a function of time for different types of perturbations.

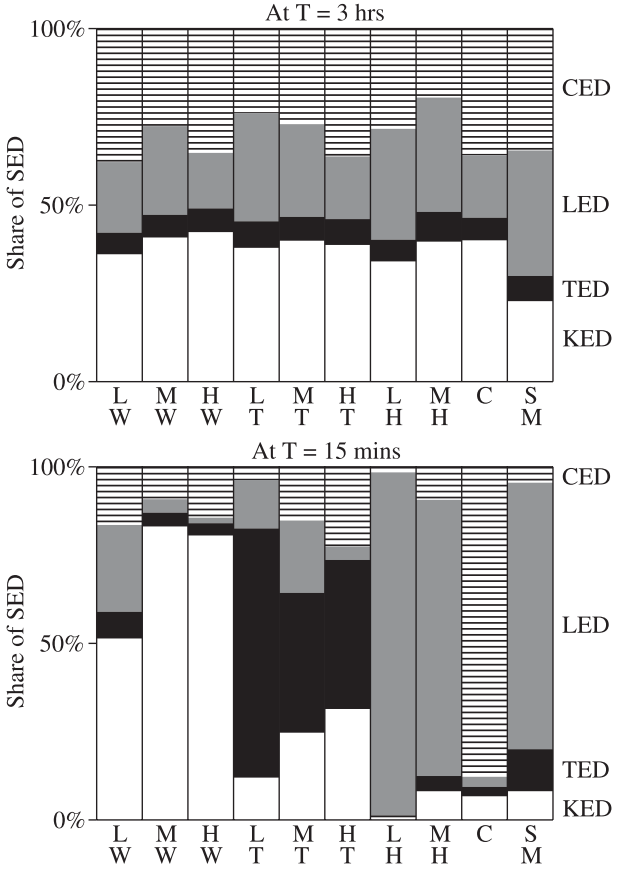


FIG. 6. Share of the SED among its different components [KED (white), TED (black), LED (gray), and CED (stripes)] (top) 3 h and (bottom) 15 min after the initial perturbations. Each bar graph is for a different type of perturbations. (from left to right) Low-level winds (LW), midlevel winds (MW), high-level winds (HW), low-level temperatures (LT), midlevel temperatures (MT), high-level temperatures (HT), low-level humidity (LH), midlevel humidity (MH), condensates (C), and soil moisture (SM) perturbations.

There is much information in Fig. 5, but two points are particularly relevant in the context of this work. First, the uncertainty in midlevel humidity appears to be dominating all other uncertainties, particularly at short forecast times. Initially, the possibility that the humidity perturbations used in our simulations were too large was considered. To verify this hypothesis, we complemented the long runs where a ±20% uncertainty in humidity (or a dewpoint uncertainty of nearly 3°C over 500-km scales) was assumed with an additional set of 1.5-h runs where a ±10% uncertainty in humidity (or 1.5°C in dewpoints) was assumed. Even in this scenario, midlevel humidity again proved to be the main source of error in mesoscale forecasts.

A second point of interest from Fig. 5 is that it shows that the relative magnitude of errors varies significantly

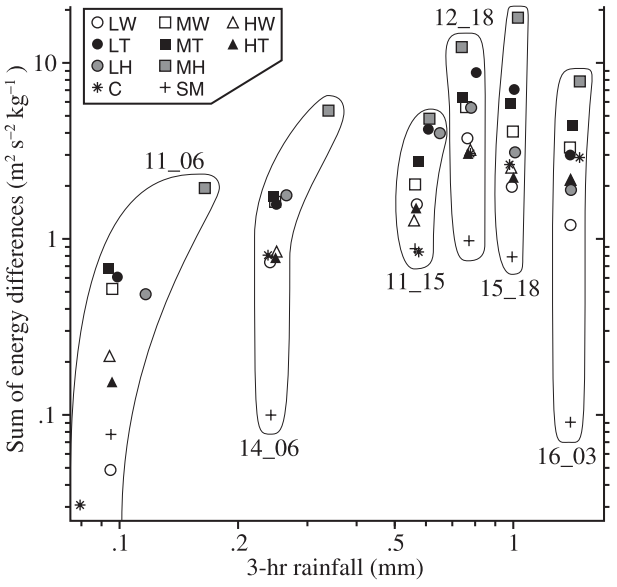


FIG. 7. Sum of energy differences at $T = 3$ h as a function of the domain-averaged rainfall accumulation that occurred in the first 3 h of the run. Six of the 16 times were chosen for this plot to make the figure less busy; runs were sorted by average rainfall, and every third case is plotted beginning with the set of runs producing the greatest average precipitation. The day (DD) in June 2002 and the hour (HH) in local daylight time (LDT) when the runs started ($T = 0$) are shown next to each set of points using a DD_HH format.

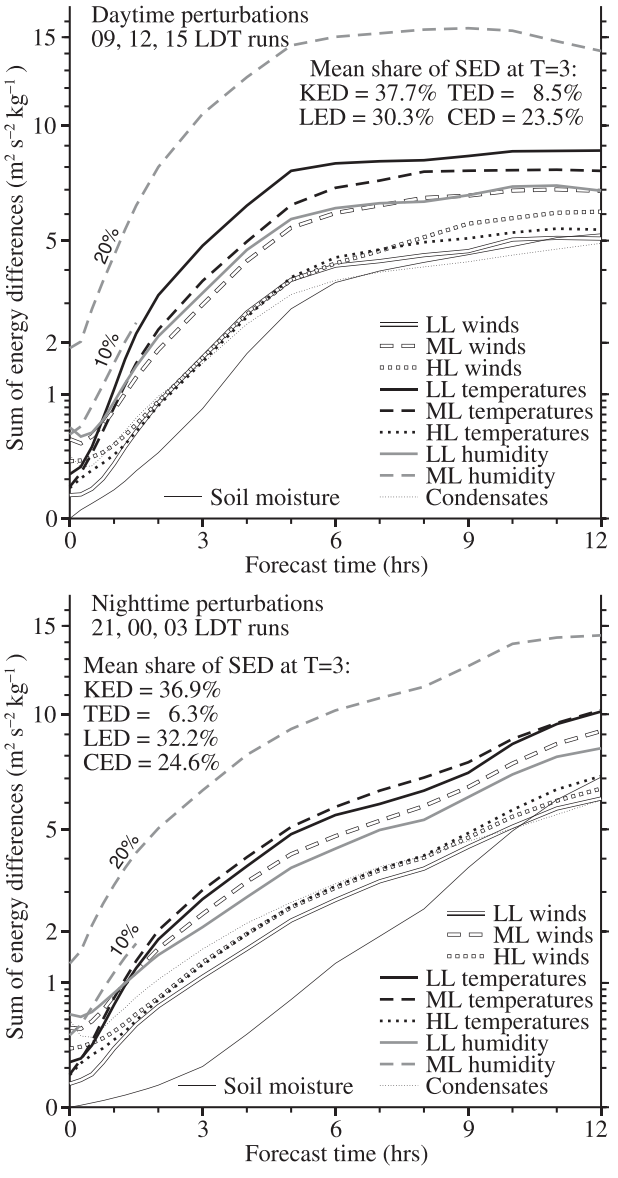


FIG. 8. SED between the perturbed and the control runs as a function of forecast time for runs starting during the (top) day and (bottom) night. Although there are significant differences in the magnitude of errors between daytime and nighttime runs, the relative magnitude of errors between the different types of initial condition perturbations as well as the distribution of the forecast errors among the four components considered (KED, TED, LED, and CED) remain mostly unchanged.

with the forecast duration. At $T = 12$ h, all but the runs with the midlevel humidity perturbations yield comparable SED. But for forecasts of a few hours, several groups can be identified: at the top thrones the midlevel humidity perturbation runs; a factor of 2–3 below, one can find the low-level temperature, midlevel temperature, low-level humidity, and midlevel wind runs; another factor of 2 below sit the condensates, low- and high-level winds, and high-level temperature runs; another factor of 5 below, the soil moisture runs closes the march. Consequently, depending on the forecast horizon of interest, the attention one should pay to different fields ought to be different: while the uncertainty on all model variables should be reduced nearly equally to improve long mesoscale forecasts, a more targeted approach focusing on humidity and temperature at low and midlevels as well as on midlevel winds should yield the best dividends for shorter forecasts.

Figure 6 shows the share of energy differences among the different terms listed in (7). By $T = 3$ h, almost all memory of the field perturbed is lost, and all runs show similar proportions of errors in temperature, winds, humidity, and precipitation. Since these proportions are similar, selecting another measure of forecast error than our SED would yield similar results as those shown in Fig. 5 beyond 3 h.

As can be expected, there is considerable variability in the magnitude of the perturbations (Fig. 7). Much of the observed variability is correlated with precipitation, in line with ideas suggesting that moist convective instability is the main growth mechanism of errors initially (e.g., Zhang et al. 2007). Some of the variability can also be explained by contrasting daytime and nighttime runs (Fig. 8). But

somewhat to our surprise, the pecking order of perturbations was not changed significantly between day and night, and only small relative enhancements were observed during the day for boundary layer–based perturbations. This can also be observed on a run-by-run basis (Fig. 7). Such a result gives us confidence that the findings presented here have robustness and should probably apply to other times and regions provided that the atmospheric conditions are conducive to deep convection.

Perturbations, especially the rapidly growing ones, quickly propagate to other variables. For most perturbations, already after 15 min (bottom of Fig. 6), a significant share of the original perturbation has been redistributed. This implies that if the assimilation period is long enough, most types of instruments have an opportunity to observe a signal resulting from the perturbations. But a few perturbations, like those in low-level humidity, have slow starts, and they will be nearly impossible to detect using short assimilation periods by instruments incapable of sensing the quantity perturbed. In the case of low-level humidity perturbations, the slow start in their transformation is associated with the time required to advect them to higher levels where they will then affect cloud and subsequently precipitation formation. It is possible that this delayed transfer of the low-level humidity perturbations was exacerbated by the slow, fat updrafts generated by 4-km resolution model runs; the extent to which the slow transformation is physical and not dependent on model resolution should be further studied.

In all cases, after a relatively short time, all perturbations in one variable will provide measurable signals on other variables, but how useful a signal?

4. Predictability of the response to perturbations

Many data assimilation systems work best if a perturbation $\Delta\mathbf{x}$ to the model state $\mathbf{x}_o$ results in a linear response of the model:

$$M(\mathbf{x}_o + \Delta\mathbf{x}) \approx M(\mathbf{x}_o) + \Delta\mathbf{x}\left(\frac{\partial M}{\partial \mathbf{x}}\right)_{\mathbf{x}_o}, \tag{8}$$

where $M(\mathbf{x})$ represents the operations executed by the weather prediction model on the initial model state $\mathbf{x}$. If (8) is true, then a perturbation that is a fraction k of the original one should trigger a response k times as intense from the model:

$$M(\mathbf{x}_o + k\Delta\mathbf{x}) - M(\mathbf{x}_o) \approx k[M(\mathbf{x}_o + \Delta\mathbf{x}) - M(\mathbf{x}_o)]. \tag{9}$$

This kind of behavior is expected in models for small enough perturbations and is generally valid for a short enough time much smaller than the time of predictability. This was illustrated in Fig. 1 as the ''linear regime.'' Hence, for each variable v that is a subset of the model state $\mathbf{x}$, one can define a nonlinearity index (NLI) such that

$$\mathrm{NLI}(v, T, \Delta\mathbf{x}) = \frac{\int_D \left|[v(\mathbf{x}_o + k\Delta\mathbf{x}) - v(\mathbf{x}_o)] - k[v(\mathbf{x}_o + \Delta\mathbf{x}) - v(\mathbf{x}_o)]\right| dD}{\int_D \left|v(\mathbf{x}_o + k\Delta\mathbf{x}) - v(\mathbf{x}_o)\right| dD}. \tag{10}$$

The NLI integrated over the domain D will depend on the variable v considered, the time of integration T, and the perturbations $\Delta\mathbf{x}$ introduced in the model initial conditions. As long as the model is in its linear stage, the two terms in the numerator should cancel each other because of (9), and the NLI should be near 0. As the integration progresses, model evolution will become increasingly nonlinear, and the two terms in the numerator will fail to cancel each other, resulting in an increase in the NLI. It also turns out that after a long enough time T, the smaller perturbation has generally grown significantly more than the larger one, and the second term of the numerator becomes much smaller than the first as well as decorrelated from it. As a result, the NLI tends to 1.

The NLI is hence a useful bulk quantity to measure the nonlinearity in the response of the model to a perturbation after a certain time. It is tied with predictability, but somewhat loosely. Of greater relevance to this work is how good an indicator it is of the potential performance of data assimilation. If NLI = 0, the minimization of the cost function that measures the mismatch between true observations and model-generated observations can be done easily: adjoints and minimization routines can work efficiently, and observation errors are the only difficulty that a minimization algorithm encounters. As the NLI increases, even with perfect observations, the ''topography'' of the cost function becomes more complicated: complex valleys and multiple

minima start to appear, and minimizations can be done less efficiently.

In terms of its effect on data assimilation, an easier quantity to grasp than the NLI is the relative frequency of the contradictory regions discussed in Fig. 1. Under these circumstances, by attempting to minimize the difference between simulated and measured observations, data assimilation systems will be lured away from the correct answer instead of toward it. We can define a contradictory information index (CII) that measures the ratio of contradictory regions to that of noncontradictory ones:

$$\mathrm{CII}(\upsilon, T, \Delta\mathbf{x}) = \frac{\int_D \left|\mathrm{sgn}[\upsilon(\mathbf{x}_o + k\Delta\mathbf{x}) - \upsilon(\mathbf{x}_o)] - \mathrm{sgn}[\upsilon(\mathbf{x}_o + \Delta\mathbf{x}) - \upsilon(\mathbf{x}_o)]\right| dD}{\int_D \left|\mathrm{sgn}[\upsilon(\mathbf{x}_o + k\Delta\mathbf{x}) - \upsilon(\mathbf{x}_o)] + \mathrm{sgn}[\upsilon(\mathbf{x}_o + \Delta\mathbf{x}) - \upsilon(\mathbf{x}_o)]\right| dD}, \tag{11}$$

where

$$\mathrm{sgn}(x) = \begin{cases} 1: & x > 0 \\ 0: & x = 0 \\ -1: & x < 0 \end{cases}. \tag{12}$$

If CII is 0, all data for which there is a signal pushes toward the correct answer. When there are as many points pulling away from the correct answer than there are pushing toward it, CII is 1.

As illustrated in Fig. 3, two sets of perturbed runs were done: 12-h-long ones with the full perturbations that we analyzed in section 3, and 3-h-long ones with perturbations a fraction $k = 1/8$ of the original ones. By using these two sets of perturbed runs in conjunction with (10) and (11), one can study the linearity, or lack of, of the perturbations from each model variable for each type of perturbation as well as the importance of contradictory information. Here, we will focus on the NLI and the CII of each variable as a function of time, but averaged for all perturbation scenarios and with a weight proportional to the amplitude of the change in υ:

$$\overline{\mathrm{NLI}}(\mathbf{v}, T) = \frac{\sum_{\mathrm{All}\Delta\mathbf{x}} \left|\upsilon(\mathbf{x}_o + \Delta\mathbf{x}) - \upsilon(\mathbf{x}_o)\right| \mathrm{NLI}(\mathbf{v}, T, \Delta\mathbf{x})}{\sum_{\mathrm{All}\Delta\mathbf{x}} \left|\upsilon(\mathbf{x}_o + \Delta\mathbf{x}) - \upsilon(\mathbf{x}_o)\right|}, \tag{13}$$

for the NLI and a similar equation for the CII. Since all perturbations were defined based on current uncertainties in initialization, they should be equiprobable, and (13) can be used as a representative measure of the nonlinearity of the response of specific variables in the model to typical initialization errors. Only a systematic study will tell how much nonlinear behavior the data assimilation systems can tolerate for each variable. Here, we chose to use two thresholds as signposts, NLI = 0.5 as a ''best before'' threshold, and NLI = 0.9 as a ''challenging beyond'' threshold.

The top of Fig. 9 shows the nonlinearity and contradictory indices of different model variables. Two families of variables can be identified. A first family whose perturbations evolve less nonlinearly includes horizontal winds, temperature, and humidity. The perturbations in these variables cross the NLI = 0.5 threshold in 90 min. Condensates form the second family. Their perturbations become nonlinear more rapidly, and the NLI = 0.5 threshold is crossed in between 15 and 40 min while the NLI = 0.9 threshold is passed after 1–2 h. This is a result of both the fact that clouds and precipitation often form as a result of convection, a highly nonlinear smaller-scale process, as well as because cloud and precipitation amounts vary nonlinearly with changes in other fields. For example, as total water amount increases, cloud amounts will remain zero until saturation is reached; beyond that point, it will grow at a steady rate until precipitation starts to remove cloud by coalescence. The overall outcome of this test is that condensates should be much more difficult to assimilate over long periods than other variables, and their assimilation over periods beyond an hour may be very difficult. CII results show similar patterns, though the CII have much smaller values than the NLI. Still, a modest CII of 0.11 implies that for 10% of the model grid points where a signal of an initial condition error can be detected, most minimization algorithms will tend to degrade the solution instead of improving it, even with perfect data and a well-functioning assimilation system.

Larger-scale patterns evolve more slowly and are more predictable. Furthermore, linearization approximations become inaccurate first at smaller scales (e.g., Tanguay et al. 1995). Consequently, smoothed patterns are often more predictable than unsmoothed ones (e.g., Bellon and Zawadzki 1994; Germann et al. 2006). By the same token, smoothing model perturbations should decrease their nonlinear behavior, making smoothed variables easier to assimilate over longer periods. We chose to test this idea by smoothing the control and perturbed model outputs

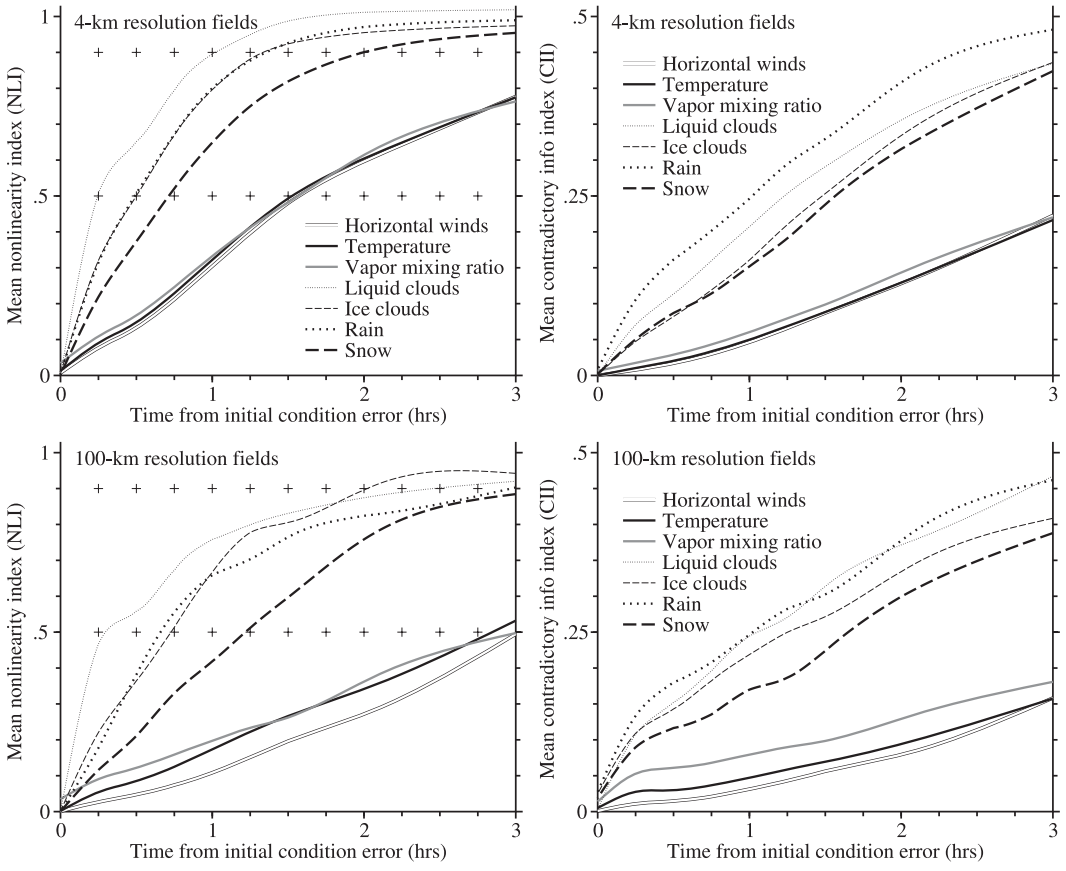


FIG. 9. (left) Mean nonlinearity and (right) contradictory information indices as defined in (13) as a function of the time (top) since the initial condition error for different model variables at the full model resolution and (bottom) when fields are smoothed over 100 km × 100 km regions. NLI and CII values are computed at times when model outputs are available (see Fig. 3), and interpolated in between. The crosses at the NLI = 0.5 and NLI = 0.9 threshold levels are simply markers to ease readability.

over 100 km × 100 km areas. The NLI and CII of smoothed variables are shown in the bottom of Fig. 9. As expected, smoothing decreased both indices. For winds, temperature, and humidity, the smoothing doubled the time taken until the NLI = 0.5 threshold is crossed to nearly 3 h. Unexpectedly, the gains for condensates proved to be much more limited, even in relative terms. Only 5 min was gained for liquid clouds, and 10 min for rain and ice clouds. Smoothing can only be done if fields of the variables to be smoothed are measured; ironically, there are many more sensors measuring fields of condensates such as radars and satellite imagers than there are sensors measuring fields of other variables. One may also want to consider the benefit of assimilating point measurements smoothed in time; the limited time resolution with which model outputs were archived prevented us from testing this idea. Another interesting result is that smoothing model variables had a much smaller effect on the CII than on the NLI. In fact, for short forecasts, smoothing slightly increased the fraction of contradictory information: as a result of smoothing, a larger fraction of data points were affected by isolated pockets of strong and highly nonlinear response associated with convection.

5. Consequences on data assimilation

Let us consider the results discussed previously and how their consequences impact data assimilation. First, an obvious observation: Fitting data beyond their time of predictability has a negative impact. If a significant fraction of the data pulls away from the correct answer as indicated by higher CII values, it prevents assimilation from having the desired effect. If not understood properly, the resulting reduced performance of the assimilation may then possibly be attributed to the poorness or worthlessness of data, and not to the fact that their information has become unusable beyond a certain assimilation interval. And after noticing that data do not properly correct the model state, one might mistakenly assign low weight to the information provided by

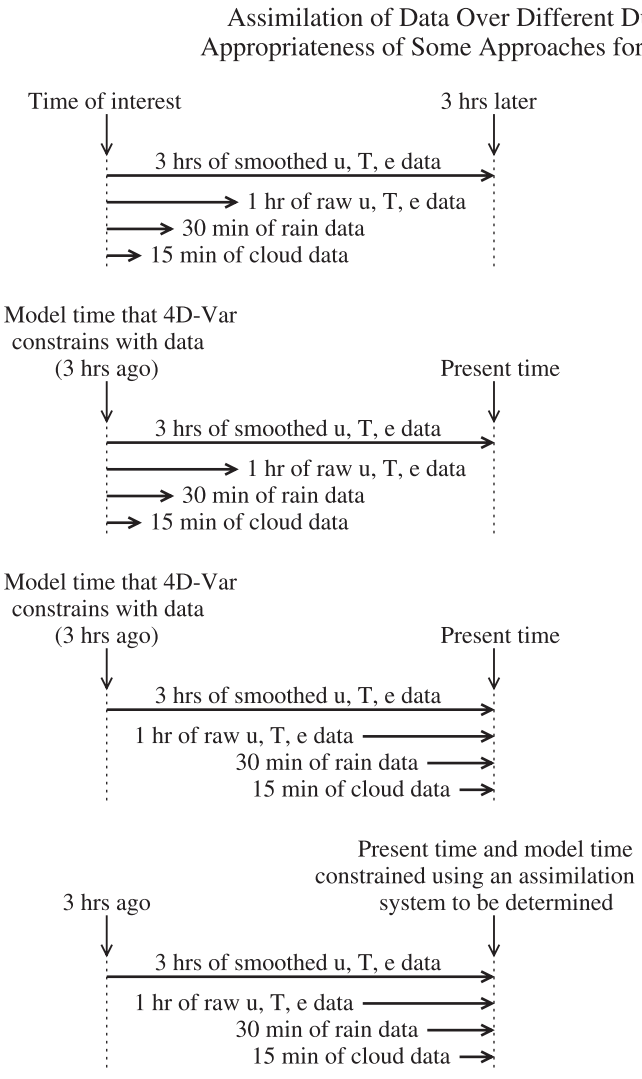


FIG. 10. (left) Illustration of four data assimilation approaches. Horizontal arrows illustrate the time period over which different datasets are being assimilated. (right) Comments on the strengths or weaknesses of these different approaches for case reanalysis or real-time uses.

measurements for the whole assimilation period. The solution to this issue is to limit the assimilation interval to a period short enough for the data to be predictable. For example, if one finds that in actual fact data with a NLI above 0.5 are more detrimental than beneficial to the assimilation process, one comes to the conclusion that the optimum assimilation interval must be dataset dependent. In this example, using Fig. 9 as a guide, one would assimilate up to 3 h of smoothed winds, temperature, and humidity; 90 min of high-resolution winds, temperature, and humidity; 30–45 min of precipitation data; and 15–30 min of cloud data. Note again that this assessment assumes a perfect model and error-free data. In practice, these times may well be shorter. And they may well be too short to allow for a proper monitoring of the evolution of intermediate and large-scale patterns, limiting our ability to initialize them properly in data assimilation systems.

How should the approach used to constrain models with data be modified to take into account limits imposed by predictability? A common data assimilation approach is to use a cost function that measures the mismatch between all observations within an assimilation interval and model-simulated observations. Until now, one used a single cost function that included measurement errors and sometimes model error (Trémolet 2006), but did not explicitly include any consideration on the predictability of different initial condition errors and how it affects different variables. An approach based on a single cost function can still be used, but the

minimization of the cost function ought to be designed in a way that dampens or nullifies all the information coming from unpredictable data. For example, data that gradually becomes unusable beyond a certain time could be given progressively lower weights on the basis of predictability limits, even though their accuracy remains unquestioned. This approach would take into account the issue raised by Fig. 9. Whether these weights could be chosen dynamically based on the weather situation or not is an interesting open question.

Other suggested changes to data assimilation are more fundamental. Consider our proposal to assimilate up to 3 h of smoothed winds, temperature, and humidity; 90 min of high-resolution winds, temperature, and humidity; 30–45 min of precipitation data; and 15–30 min of cloud data. Under current 4D-Var formalism, data assimilation adjusts the atmospheric conditions at the *beginning* of the assimilation interval based on the data collected during the whole period (Fig. 10). For reanalysis purposes, this is fine. It is more problematic for real-time use. If one wants to assimilate 3 h of smoothed wind data together with 15–30 min of cloud data, then that cloud data would have to be nearly 3 h old: current cloud data is more than 2.5 h after the initial-condition time whose variables we try to constrain, and hence well beyond predictability limits. A possible way out is to rethink 4D-Var data assimilation by constraining the data at the *end* time of the assimilation interval, though the irreversibility of some processes such as diffusion might make such a solution difficult to realize. But this way, the latest data could be used for all datasets.

Colleagues have asked: ''Why not assimilate the data over short windows as time progresses; is that not the same and simpler than a long window?'' Not necessarily. Consider a situation where winds aloft evolve with time because of a propagating low-level cold outflow, and that the cold air is not directly observed. If successive assimilation windows are used, assimilation systems have the choice of directly adjusting the wind or adjusting temperature, among others. If the time required for the atmosphere to adjust to a temperature change is longer than the assimilation window, changing temperatures will be less favored by the minimization algorithm than a direct change in the winds. In addition, the consistently changing wind may be interpreted as unreliable data and given lower weights as a result. In a one window situation, the assimilation system must contend with the wind acceleration and find a physically compatible cause for it; the model will then be better constrained. This situation is less likely at the synoptic scale where measurements are more plentiful and most fields are better constrained together via hydrostatic balance and geostrophic approximation. This is not the case at the mesoscale where measurements are very sparse for temperature and humidity in particular. In such situations, using long assimilation windows will yield different, and hopefully better, results.

6. Review

Except for errors in low-level moisture and soil properties, most initial condition errors in one variable propagate rapidly to other variables, allowing them to be detected by a variety of instruments and possibly to be corrected by data assimilation. Different types of errors will have different effects on forecast accuracy. In parallel, the usability of the information from various instruments for data assimilation purposes depends on the strength of the signal caused by the initial condition error and the linearity of the model's response to that error. Using initial condition errors compatible with today's accuracy in model initializations, it was found that uncertainties in midlevel moisture caused the greatest uncertainties in forecasts. For mesoscale forecasts exceeding 6 h, the uncertainty in all other variables had a comparable effect on forecast accuracy; for forecasts of up to a few hours, uncertainties in low-level temperatures, midlevel temperatures, low-level moisture, and midlevel winds stood out as the greatest cause of forecast uncertainty after midlevel moisture.

To constrain model variables with data, one would like to assimilate data for a long enough interval to constrain sluggish large-scale patterns while not reaching the predictability limits for small-scale patterns. At the 4-km resolution of our simulation, it was found that predictability imposed limits on the assimilation interval that are different for different model variables, condensates being harder to assimilate than other variables for long time periods. If one assumes a reasonable though debatable threshold based on the nonlinearity of the response to perturbations (NLI = 0.5), one finds that winds, temperature, and humidity can be assimilated for 90 min, but precipitation and cloud data are limited to 30–45 and 15–30 min, respectively. An optimal assimilation must hence be done over different assimilation intervals for different datasets, and that requires adjusting the approach used to fit model fields to observations. This includes diminishing the weight of unpredictable observations in the cost function J, as well as reformulating 4D-Var to constrain the model fields at the end of the assimilation interval rather than at its beginning.

All the conclusions above are based on a well-functioning data assimilation system, error-free information on model variables, and a perfect model. The numbers quoted in this paper are hence best-case estimates. Real data with their measurement errors and sometimes indirect link to model variables add another

level of complications that are dealt in Part II (Fabry 2010).

Acknowledgments. This work was funded by the Natural Science and Engineering Research Council of Canada and the Canadian Foundation for Climate and Atmospheric Sciences and was made possible by a sabbatical leave of the first author financed by McGill University. The authors thank Eunha Lim, Qingnong Xiao, and Sherrie Fredrick from NCAR/MMM for their technical assistance with the WRF model. We also recognize the scientific and technical input from Kay, Chris, Rich, Yongsheng, Wei, Michael, Morris, Joseph, Bill, and David from NCAR/MMM.

REFERENCES

Bellon, A., and I. Zawadzki, 1994: Forecasting of hourly accumulations by optimal extrapolation of radar maps. *J. Hydrol.,* **157,** 211–233.

Bennett, A., B. Chua, and L. Leslie, 1996: Generalized inversion of a global numerical weather prediction model. *Meteor. Atmos. Phys.,* **60,** 165–178.

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

Courtier, P., J.-N. Thepaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4d-Var using an incremental approach. *Quart. J. Roy. Meteor. Soc.,* **120,** 1367–1387.

Daley, R., 1991: *Atmospheric Data Analysis*. Cambridge University Press, 457 pp.

Derber, J., 1989: A variational continuous assimilation technique. *Mon. Wea. Rev.,* **117,** 2437–2446.

Dudhia, J., 1989: Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.,* **46,** 3077–3107.

Ehrendorfer, M., and R. M. Errico, 1995: Mesoscale predictability and the spectrum of optimal perturbations. *J. Atmos. Sci.,* **52,** 3475–3500.

——, ——, and K. D. Raeder, 1999: Singular-vector perturbation growth in a primitive equation model with moist physics. *J. Atmos. Sci.,* **56,** 1627–1648.

Fabry, F., 2006: The spatial variability of moisture in the boundary layer and its effect on convection initiation: Project-long characterization. *Mon. Wea. Rev.,* **134,** 79–91.

——, 2010: For how long should what data be assimilated for the mesoscale forecasting of convection and why? Part II: On the observation signal from different sensors. *Mon. Wea. Rev.,* **138,** 256–264.

Germann, U., I. Zawadzki, and B. Turner, 2006: Scale dependence of predictability of precipitation from continental radar images. Part IV: Limits to prediction. *J. Atmos. Sci.,* **63,** 2092–2108.

Hohenegger, C., and C. Schär, 2007: Atmospheric predictability at synoptic versus cloud-resolving scales. *Bull. Amer. Meteor. Soc.,* **88,** 1783–1793.

Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, 341 pp.

Lenschow, D. H., and J. Sun, 2007: The spectral composition of fluxes and variances over land and sea out to the mesoscale. *Bound.-Layer Meteor.,* **125,** 63–84.

Lorenz, E., 1969: The predictability of a flow which possesses many scales of motion. *Tellus,* **21,** 289–307.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radioative transfer for inhomogeneous atmosphere: RRTM, a validated correlated k-model for the longwave. *J. Geophys. Res.,* **102** (D14), 16 663–16 682.

Nastrom, G. D., and K. S. Gage, 1985: A climatology of atmospheric wavenumber spectra of wind and temperature observed by commercial aircraft. *J. Atmos. Sci.,* **42,** 950–960.

Rabier, F., H. Järvinen, E. Klinker, J.-F. Mahfouf, and A. Simmons, 2000: The ECMWF operational implementation of four dimensional variational assimilation. Part I: Experimental results with simplified physics. *Quart. J. Roy. Meteor. Soc.,* **126,** 1143–1170.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers, 2005: A description of the Advanced Research WRF version 2. NCAR Tech. Note NCAR/TN-468+STR, 88 pp.

Sun, J., 2005: Initialization and numerical forecasting of a supercell storm observed during STEPS. *Mon. Wea. Rev.,* **133,** 793–813.

Talagrand, O., 1981: A study of the dynamics of four-dimensional data assimilation. *Tellus,* **33,** 43–60.

——, 1997: Assimilation of observations, an introduction. *J. Meteor. Soc. Japan,* **75,** 191–209.

Tanguay, M., P. Bartello, and P. Gauthier, 1995: Four-dimensional data assimilation with a wide range of scales. *Tellus,* **47A,** 974–997.

Thompson, G., R. M. Rasmussen, and K. Manning, 2004: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis. *Mon. Wea. Rev.,* **132,** 519–542.

Trémolet, Y., 2006: Accounting for an imperfect model in 4D-Var. *Quart. J. Roy. Meteor. Soc.,* **132,** 2483–2504.

Weckwerth, T. M., and Coauthors, 2004: An overview of the International H_2O Project (IHOP_2002) and some preliminary highlights. *Bull. Amer. Meteor. Soc.,* **85,** 253–277.

Wilson, J. W., and R. D. Roberts, 2006: Summary of convective storm initiation and evolution during IHOP: Observational and modeling perspective. *Mon. Wea. Rev.,* **134,** 23–47.

Xiao, Q., and J. Sun, 2007: Multiple radar data assimilation and short-range quantitative precipitation forecasting of a squall line observed during IHOP_2002. *Mon. Wea. Rev.,* **135,** 3381–3404.

Zhang, F., N. Bei, R. Rotunno, C. Snyder, and C. C. Epifanio, 2007: Mesoscale predictability of moist baroclinic waves: Convection-permitting experiments and multistage error growth dynamics. *J. Atmos. Sci.,* **64,** 3579–3594.

Zou, X., and Y.-H. Kuo, 1996: Rainfall assimilation through an optimal control of initial and boundary conditions in a limited-area mesoscale model. *Mon. Wea. Rev.,* **124,** 2859–2882.

Zupanski, D., 1997: A general weak constraint applicable to operational 4DVAR data assimilation systems. *Mon. Wea. Rev.,* **125,** 2274–2292.