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

FRÉDÉRIC FABRY

McGill University, Montreal, Canada

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

ABSTRACT

The ability of data assimilation to correct for initial conditions depends on the presence of a usable signal in the variables observed as well as on the capability of instruments to detect that signal. In Part I, the nature, properties, and limits in the usability of signals in model variables were investigated. Here, the focus is on studying the skill of measurements to pull out a useful signal for data assimilation systems to use. Using model runs of the evolution of convective storms in the Great Plains over an active 6-day period, simulated measurements from a variety of instruments are evaluated in terms of their ability to detect various initial condition errors and to provide a signal above and beyond measurement errors. The usability of the signal for data assimilation is also investigated. Imaging remote sensing systems targeting cloud and precipitation properties such as radars and thermal IR imagers provided both the strongest signals and the hardest ones to assimilate to recover fields other than clouds and precipitation because of the nonlinear behavior of the sensors combined with the limited predictability of the signal observed. The performance of other sensors was also evaluated, leading to several unexpected results. If used with caution, these findings can help determine assimilation priorities for improving mesoscale forecasting.

1. Detecting initial condition errors

For data assimilation to succeed, the presence of a usable signal in data, an accurate model, and a properly functioning data assimilation system are all required. Much effort is being put on improving the mathematics and mechanics of data assimilation systems (e.g., Talagrand 1997; Kalnay 2003). Improving numerical models and their ability to simulate observations is the subject of even more work. The triad data–model–assimilation is evaluated either in specific case studies (e.g., Gao et al. 1999; Montmerle et al. 2002; Sun 2005) or in simulations of the usefulness of observing systems (e.g., Sokolovskiy et al. 2005; Tong and Xue 2005). But to the author's knowledge, nothing has been done to systematically study the occurrence of a measurable and usable signal in the data itself, at least in the context of the mesoscale forecasting of convection.

Both measurability and usability are important. Measurability refers both to the presence of a signal in the fields observed as well as the possibility of pulling a signal out of measurement noise. Usability deals with the ability of a data assimilation system to take that signal and use it to retrieve the error in initial conditions that caused the signal observed. This, in turn, depends on both on the mathematical ability of the assimilation system to go from the effect to the cause as well as whether the signal observed may have multiple causes or not.

To constrain all model variables everywhere, each possible type of initial condition error must cause an unexpected signal in the observations to be assimilated. In Fabry and Sun (2010, hereafter Part I) it was established that at the mesoscale, most initial condition errors in one variable started to contaminate other variables within 15 min, and that by 3 h, there had been enough interactions that one could not determine the origin of the initial condition errors simply by looking at the relative magnitude of errors in each variable. This is good news because it means that many types of sensors can get a signal from any particular initial condition error; however, it raises the question of our ability to determine what the origin of any discrepancy between expected measurements and observations is. A second important finding of Part I was that the response of a modeled atmosphere was not always linearly proportional to the magnitude of the

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: 10.1175/2009MWR2884.1

initial condition error. Though this result is not a surprise per se, the relatively short time over which the model behaved linearly given typical initial condition errors may limit our ability to assimilate data over long enough periods to constrain larger-scale and slower-evolving patterns. Because while on one hand, one would prefer to assimilate data for as long as possible to constrain patterns at all scales, the complex evolution of convective weather limited the time period over which different model variables could be usefully constrained. With a simulation made at 4-km resolution, moderate nonlinearities were seen after 90 min in winds, temperature, and humidity fields, while similar nonlinearities could be seen in clouds and precipitation fields in 15–30 min.

The conclusions of Part I were based on the assumption that one is perfectly measuring the model variables. When measurements from actual instruments are considered, new complications arise. Measurements have errors. The link between measurements and model variables is sometimes complex. The signal from instruments may have properties that make their assimilation more difficult. And some instruments measure more than one field, each of which has their own peculiarities.

It takes a considerable time to add a new type of measurements in a data assimilation system. The choice of where to focus data assimilation efforts is often a difficult one, as even observational system simulation experiments require much effort. But a great deal can be learned at much lower costs by limiting oneself to evaluating the strength and optimum usability of a dataset by a simulation experiment. For example, there would be no point in trying to assimilate a dataset that fails to deliver a usable signal. Furthermore, an analysis of the strength of the signal provided by different instruments is useful to determine the potential that their data has to correct for initial condition errors. In this work, the framework established by Part I for idealized model fields is used to study the properties and the detectability of the signal caused by initial condition errors in observations from a variety of sensors.

2. Determining signal strength

Ideally, one would like to know what is the skill or the ability of using data from an instrument in a data assimilation system to improve forecasts. In this work, however, the focus will be strictly on determining the presence of a signal and study some its properties, specifically how linear is the signal's response to changes in initial condition errors.

How much signal of the initial condition errors is there in observations? The answer depends on the nature and strength of initial condition errors, on the variable(s) measured by a given instrument, on the duration over which the observation is made, and on the extent with which the initial condition errors make their way on the variable(s) observed. The signal strength also depends on the expected magnitude of the changes in the observations $\mathbf{y}$ compared with the uncertainty $\sigma(\mathbf{y})$ in these observations. Finally, it is also a function of the ability of the model and data assimilation to use the variables observed.

Given an atmospheric state $\mathbf{x}$, and a perfect model whose initial conditions are in error by $\Delta\mathbf{x}$ at an initial time, one can compute what would be the time evolution of the error-free true observations $\mathbf{y}(\mathbf{x})$ coming from the atmosphere as well as the error-free observations $\mathbf{y}(\mathbf{x} + \Delta\mathbf{x})$ expected if the perturbed initial conditions were true. The strength S of the observational signal from a dataset y_i at a time T after the initial time is hence a result of the variance of the difference between the expected and true observations, normalized by the expected uncertainties in the observations:

$$S_{y_i}(\Delta\mathbf{x}, T) = \sum_{i=1}^{N} \frac{[y_i(\mathbf{x} + \Delta\mathbf{x}, T) - y_i(\mathbf{x}, T)]^2}{\sigma(y_i)^2}, \qquad (1)$$

where N is the number of observations coming from that sensor per unit time. As in Part I, the nonlinearity of the signal from each sensor can be quantified using

$$\mathrm{NLI}(y_i, T, \Delta\mathbf{x}) = \frac{\sum_{i=1}^{N} \left|[y_i(\mathbf{x}_o + k\Delta\mathbf{x}) - y_i(\mathbf{x}_o)] - k[y_i(\mathbf{x}_o + \Delta\mathbf{x}) - y_i(\mathbf{x}_o)]\right|}{\sum_{i=1}^{N} \left|[y_i(\mathbf{x}_o + k\Delta\mathbf{x}) - y_i(\mathbf{x}_o)]\right|}, \qquad (2)$$

where runs with reduced perturbations ($k = 1/8$) are also used in the calculations.

Part I used 16 summertime runs centered in the southern Great Plains to study the effect of perturbing initial conditions in 10 different variables and levels: low-level winds, midlevel winds, high-level winds, low-level temperatures, midlevel temperatures, high-level temperatures, low-level moisture, midlevel moisture, whole-atmosphere condensates, and soil moisture at all depths. The fields generated by these runs will now be

TABLE 1. Number of independent measurements N per hour per instrument and their uncertainty σ for each type of measurement simulated.

Instrument or platform	Variable measured	N (h^{-1})	Error σ	Notes and references
GPS receiver	Zenith wet delay (ZWD)	2	1 kg m^{-2} ILW equivalent	Braun et al. (2003) Martin et al. (2006)
	Surface pressure (P)	2	0.5 hPa	See surface station
Microwave radiometer	Integrated vapor (IWV)	24	1 kg m^{-2}	Martin et al. (2006)
	Integrated liquid (ILW)	24	0.1 kg m^{-2}	σ: Educated guess
Radar, long range (r = 230 km)	Doppler velocity (v_{DOP})	See text and Fig. 1	1 m s^{-1}	Standard σ used
	Reflectivity (Z)	See text and Fig. 1	1 dB	
Radar, short range (r = 50 km)	Refractivity (N)	2120; 1 point every 4 × 4 km^2 up to 30-km range, 12 times	2	σ: Educated guess
Radiosonde	Humidity (RH)	60	5%	Uncertainties based on those of the Vaisala RS92 radiosonde
	Temperature (T)	60	0.5°C	
	Winds (**u**)	60	1 m s^{-1}	
Rain gauge	15-min rain (R)	4	0.2 mm	Tipping-bucket uncertainty used
Satellite-borne IR imager	Brightness temperature (T_{BB})	4 every 4 × 4 km^2	1°C	Menzel and Purdom (1994)
Surface station	Humidity (RH)	12	3%	Uncertainties based on those of the Vaisala WXT510 weather transmitter (see online at www.vaisala.com)
	Pressure (P)	12	0.5 hPa	
	Temperature (T)	12	0.3°C	
	Winds (**u**)	12	0.3 m s^{-1}	
Wind profiler	Winds (**u**)	2 every 0.25 km up to 15-km height	1 m s^{-1}	For noninterferometric profilers
	Virtual temperature (T_v)	2 every 0.25 km up to 3-km height	1°C	May et al. (1989)

used to simulate what different sensors would have observed.

In this work, we have chosen to focus on in situ instruments and ground-based remote sensors: surface stations, rain gauges, radiosondes, radars, wind profilers, radiometers, and global positioning satellite (GPS) receivers at the surface. The only satellite-borne sensor considered was the thermal IR imager. To properly compute the skill of an instrument, one must simulate its measurement accurately, determine the number of independent measurements N per unit time, and agree on the uncertainty σ of each measurement.

3. The quest for Ns and σs

The mismatch between model-simulated and actual measurements depends on the physics of the measurement process, on the number of measurements available to a real-time or a postfacto user, and on the ability of the model and of the data assimilation system to properly simulate the measurements. For example, let us consider radar reflectivity Z, one of the variables simulated in this work that is the least directly related to a model variable. If the model microphysics is truly capable of simulating reflectivity such as with Szyrmer et al. (2005), then the data can be used fully and the skill calculation should only depend on true measurement accuracy. But if the model microphysics only simulates mass and/or number concentration, and then derives reflectivity, then the uncertainty should also include the uncertainty in the conversion between mass and Z. In that case, what would be a 1-dB (25%) measurement error suddenly becomes a 3-dB equivalent (100%) interpretation and measurement error based on years of experience measuring drop-size distributions. This diminishes the skill of Z measurements by a factor of 10. Furthermore, the spatial and temporal correlation structure of that uncertainty should also enter into the picture. Since the focus of this work is on data, we chose to exclude considerations about model limitations from our analysis, but these should be kept in mind in any real situation as the effective ability of a measurement to be used to detect initialization errors also depends on them.

Table 1 summarizes the values used for the number of independent measurements and measurement accuracy for most of the instruments studied. In all cases, when

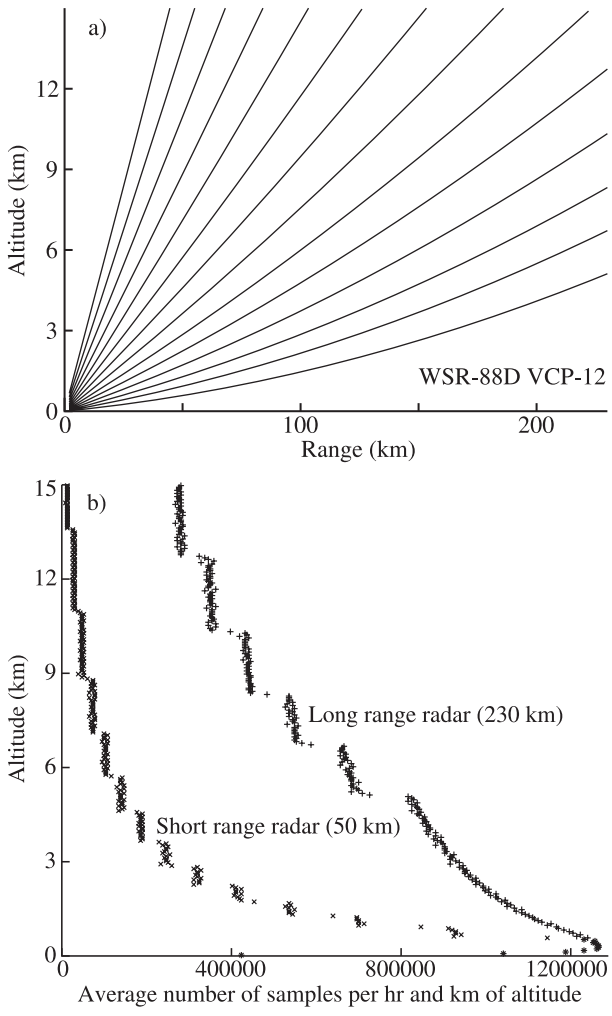

FIG. 1. (a) Altitude sampled by the WSR-88D VCP-12 with its 14 elevation angles as a function of range. (b) Resulting average number of reflectivity samples per hour and per kilometer of altitude as a function of height for a long- and short-range radar.

a choice was available between values estimated in a real-time context and a research one, the real-time numbers were preferred. For instruments with different uncertainty and data availability characteristics than those simulated, it is an easy matter to scale the results that will follow using (1) as a guide. Because of the complex sampling strategy of radars, the values of N chosen for reflectivity and Doppler velocity deserve some explanations.

The Weather Surveillance Radar-1988 Doppler (WSR-88D) radars were designed to have accuracies of 1 dB in reflectivity and 1 m s^{-1} in Doppler velocities at resolutions of 1 km in range and 1° in azimuth. These values were used for σ. The total number of samples available and its height distribution depend on the volume coverage pattern (VCP) used to scan the atmosphere. For this test, the VCP-12 pattern was assumed (Fig. 1a). Given a maximum range, the height distribution of the number of samples per hour for reflectivity can be computed (Fig. 1b). Two maximum ranges were considered: 230 km for long-range WSR-88D radars, and 50 km for short-range radars such as those developed by the Center for Collaborative Adaptive Sensing of the Atmosphere (CASA; McLaughlin et al. 2005). The number of samples for Doppler velocity depends on the presence of targets to scatter the radar waves. Targets were assumed present in precipitation (at least 10 dBZ of equivalent reflectivity due to snow, graupel, and rain), and in conditions where insect echoes would be likely, chosen to be everywhere in the troposphere where the temperature exceeded 10°C. A final Doppler velocity scenario was considered for short-range radars lacking the sensitivity to detect insects. In the case of short-range radars, no additional skill was attributed to the potential gains adaptive scanning could provide.

4. Results from each sensor

We took from Part I the control runs and the runs perturbed by different plausible initial condition errors. We then computed the signal observed in (1) by all the sensors simulated. The strength of the signal for each sensor depends on the nature of the perturbation or initial condition error, the time since the original perturbation, and the measurement capabilities of the sensor. Furthermore, the nonlinearity issues that were discussed in Part I apply to measurements in the same way as they did with model variables; we therefore also computed the nonlinearity index of each variable measured using (2).

Figure 2 shows the sum of all the signal strengths resulting from all the initial condition errors. The signal strength plotted is the strength per sensor and per hour of data collected. Because of the large amount of information in Fig. 2, the results are discussed instrument by instrument. But first, a few words on the interpretation of signal strength. Instruments or datasets providing signal strengths that largely exceed 1 when integrated over the assimilation period can be used to detect the presence of initial condition errors. That being said, these initial condition errors are often spread among many variables and several tens or hundreds of thousands of model grid points. As a result, the greater the signal strength, the more useful a dataset will be to quantify the effect of such errors. This does not necessarily imply improved ability to determine where such errors originated from, but a stronger signal should help.

a. Wind profilers and RASS

For the wind profiler and the radio acoustic sounding system (RASS), a good system capable of providing

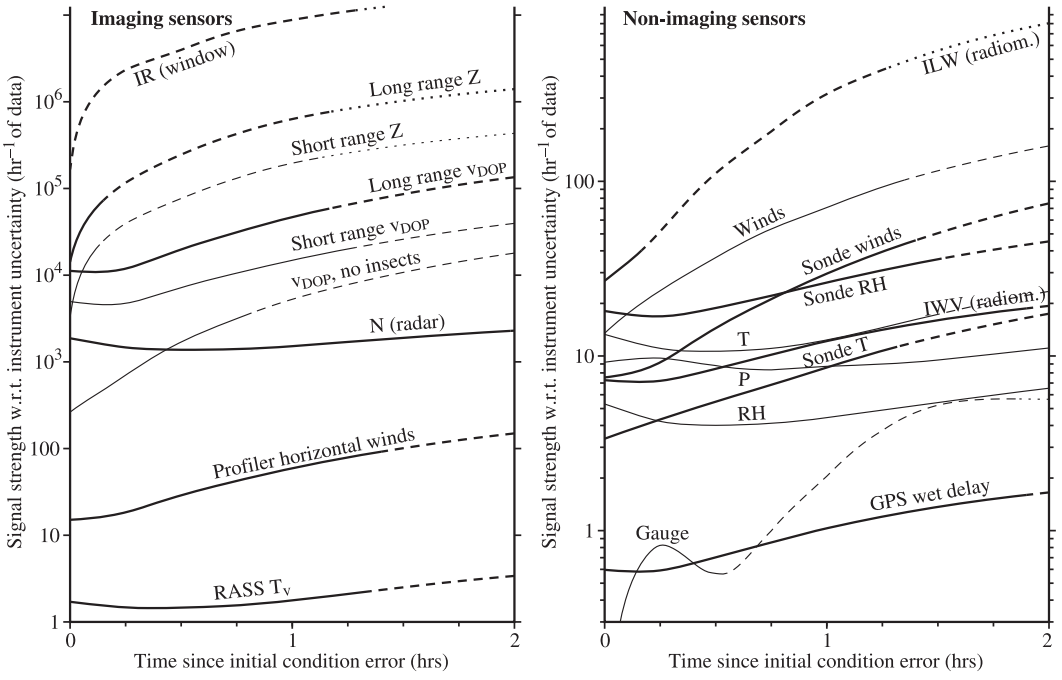

FIG. 2. Variance of the difference between observations in the perturbed runs and those in the control runs, normalized to the variance of measurement uncertainty, as a function of time for each sensor simulated. Lines are solid if the nonlinearity index (NLI) is less than 0.5, dashed if the NLI is less than 0.9, and dotted beyond. To help clarify the plot, sensors were split in two groups, imaging remote sensors to the left, and all other sensors to the right. Thin lines were used on the left plot for short-range radar results and on the right plot for the results for surface weather stations.

winds up to 15 km and virtual temperature up to 3 km at all times was assumed. At the initial time, the signal caused by expected errors in all variables was 20 times larger than measurement noise per hour of wind profiler data collected, and twice the measurement noise per hour of RASS data collected. Though the signal is weak compared to other imaging sensors, it is not strongly nonlinear up to 80 min. The RASS signal is dominated by low-level temperature errors at all times while the wind profiler signal is first due to wind errors at the initial time, and then to a combination of winds, temperature, and midlevel humidity errors after half an hour, the latter being responsible for the buildup in the nonlinearity of the signal. This pattern will repeat for other instruments: after a certain amount of time, differences due to midlevel humidity errors end up both dominating the signal and causing most of the nonlinearities, the pace at which it occurs depending on how quickly humidity errors make their way into the fields observed by the instrument considered. For instruments measuring winds and ironically humidity, this happens more slowly. For others that sample fields directly impacted by humidity changes such as instruments measuring clouds or precipitation like radars or satellite imagers, this will occur within the first 15 min.

b. Radars

As can be seen by the computed strength of the signals, radar data are extremely sensitive to changes in the initial conditions. Signal strength of reflectivity and velocity data are comparable at the initial time, at least if the velocity data from insects can be used. But because of the rapid divergence in time of precipitation patterns between the perturbed and the control runs, the reflectivity signal rapidly becomes 10 times stronger than the velocity one. Unfortunately that rapid divergence becomes strongly nonlinear beyond 10 min, severely limiting the assimilation of reflectivity over long time intervals. As a result, by the time errors in variables other than condensate amounts can propagate to radar reflectivity differences, mainly from all midlevel fields, the reflectivity signal becomes hard to use. Velocity data remain more linear, especially outside storm regions. But even in storm regions, velocity data can be assimilated for longer periods than reflectivity data, as illustrated by the 45-min linearity time obtained for Doppler velocity from storms only; here too, errors in all midlevel fields give the strongest signal. Refractivity data give the weakest signal of all, but it is linear up to 2 h; by the nature of the measurement, its signal at all times is dominated by low-level humidity and temperature errors.

The signal strength from short-range radars is obviously weaker than for longer-range ones because of their smaller spatial coverage, but the linearity characteristics are similar. If one contrasts the curves with and without insect echoes, one will note that 90% of the signal computed for the velocity data from short-range radars at the initial time comes from insect returns, with that number somewhat diminishing to 70% after an hour. Hence, for short-range radars too insensitive to detect Doppler returns from insects, velocity data loses a large part of its appeal.

A final word of caution: because of the massive amounts of data collected by radar per unit time, it does not take much of a difference in the observed fields to result into huge computed signals. A single storm that is a bit larger or smaller than expected will cause hundreds of radar pixels to be altered, and a mean change of 10 dB on the reflectivity of these pixels will result in signal strength in the tens of thousands. So even though large signals are detected, they do not necessarily translate into major changes in the fields observed by the naked eye. Also, in practice, inaccuracies in modeling radar reflectivity may well dwarf the signal caused by initial condition errors.

c. Thermal IR imagers

Because of their high sampling frequency (4 times an hour) and their complete coverage of the domain, satellite imagers provide data that are extremely sensitive to any change in the soil temperature or in the cloud-top temperature and coverage. In fact, it is the only sensor that provides any information on errors in soil properties at the initial time. But the greatest signal comes from changes in midlevel moisture in saturated or near-saturated air, resulting in the disappearance or appearance of clouds. Keep in mind that here and elsewhere, ''midlevel'' is defined as the lower troposphere above the boundary layer (e.g., from about 1 km AGL to halfway between the ground and the tropopause). The atmosphere is conditionally unstable in many regions where air is humid enough that a 20% change in relative humidity will result in the appearance or disappearance of clouds. In these regions, a change in moisture results in a step-function change in buoyancy; a parcel that was unstable may then become stable, or vice versa. This then results in a completely different forecast outcome. Furthermore, the observed brightness temperature changes suddenly with the appearance of clouds, and then much more slowly as the clouds become denser. The combination of the nonlinear change in IR brightness temperature change with cloud water content, coupled with the step-function trigger or disappearance of free convection, makes $\partial T_{BB}/\partial \mathbf{x}$, the change in IR signal with respect to atmospheric conditions, fluctuate wildly with small changes in $\mathbf{x}$. The thermal IR signal is hence extremely nonlinear at the initial time and even more so at latter times, much worse than the data of any other sensor considered here. As a result, the direct assimilation of time sequences of T_{BB} data may prove challenging if one uses minimization techniques relying on linear assumptions. Even though it is generally better to assimilate the rawest data possible, in this particular case, one should explore the possibility of assimilating a derived quantity q instead of T_{BB} such that $\partial q/\partial \mathbf{x}$ is a more continuous function than $\partial T_{BB}/\partial \mathbf{x}$. In the case of T_{BB} data from the IR window, it is not clear what such a quantity q could be: a T_{BB} only available in cloud-free regions is a possibility, but such a dataset would lose of a lot of its usefulness. Note that this issue also applies to radiometric sounding measurements made in more opaque regions of the IR, even though measurements at these wavelengths are less affected by cloud appearance and disappearance.

d. Microwave radiometers

It was assumed here that microwave radiometers measured without bias errors both integrated liquid water (ILW) and integrated water vapor (IWV) and only those fields. Good signal of the diverging runs was observed on both variables. The ILW signal is stronger and less linear because of the nonlinear behavior of clouds in the simulations, the linearity of the signal being comparable to that of radar reflectivity. The IWV measurement is good at detecting humidity errors at all times and can sense a weaker signal associated with initial condition perturbations in midlevel temperatures and winds after half an hour. The overall signal is also one of the most linear of the sensors considered.

e. Ground-based GPS receivers

Here, we chose to consider ground-based GPS receivers, not satellite-based ones. And although GPS receivers measure properties similar to those of the microwave radiometer IWV, the signal strength is much weaker because of the smaller frequency of measurement and the larger sampling errors. In fact, given the magnitude in the perturbations in initial conditions considered here, a GPS receiver cannot observe a signal comparable to measurement noise after averaging 1 h of data, the lowest signal of all the sensors in this study. That may change somewhat as the accuracy of GPS measurements of IWV increase.

f. Radiosondes

The interpretation of the radiosonde scores is somewhat more complicated because of the short-lived

nature of the instrument and the multiple fields measured. Here, scores are considered in terms of radiosonde hours. To obtain the scores mentioned, one would have to have one radiosonde in the air at all times collecting data, a new radiosonde replacing immediately a dead one typically every 30 min or so. Hence, if only one radiosonde was used, the integrated score over 30 min should be considered with the starting time being how long after the beginning of the assimilation window the radiosonde was launched.

Winds and humidity data give the strongest signal, followed by temperature about 3 times weaker. All three signals are linear enough to be easily assimilated for 90 min. And with three sensors, radiosondes can detect signals from all initial condition errors except soil moisture and, to a lesser extent, condensates.

g. Surface stations

Surface station scores were penalized compared to the radiosonde scores on the grounds that typically, data frequency is much smaller: except in research networks, most stations report in real time at best every 5 min, if not once an hour. Measurement accuracy is somewhat higher, but since our knowledge of the magnitude of fields at the surface is generally better than aloft except for temperatures, signals of discrepancies between expected and measured values are weaker. The result is then that the greatest signal observed from surface stations is by far from wind errors, followed by low-level temperature, pressure, relative humidity, and finally rain gauges. At the surface, the relative humidity signal was associated with both low-level humidity and to a lesser extent soil moisture errors, pressure perturbations with midlevel winds and low-level humidity errors, temperature signal with low-level temperature errors, and wind signal with low-level wind and temperature errors. More than for any other sensor, one should take the results from surface stations with caution: they are based on the assumption that models are perfectly capable of simulating reality 2 m above ground level, a feat that is far from assured in reality.

h. Reducing initialization errors

To reduce initialization errors, one must hence focus on the fields whose errors affect forecast performance the most, and then chose instruments and datasets that give the strongest and easiest signal to assimilate on these errors. For example, in Part I, we identified inaccuracies in midlevel moisture as the dominant cause of forecast errors, especially for short lead time forecasts. In that case, assimilating measurements sensitive to humidity errors would bear the most fruits. Many sensors measured signals associated with midlevel humidity errors, but most of these may be difficult to assimilate because of the nonlinearity of their signal. Of the datasets considered in this study, the best to minimize these errors appear to be IWV from microwave radiometers, and RH from radiosondes for periods longer than an hour, and Doppler velocity from radar for periods of tens of minutes. The first two should be easier to assimilate because of the direct link between their data and midlevel humidity, while Doppler velocity differences are much less directly linked to humidity and a good model is required to take full advantage of that data.

5. How to use and not misuse these results

I shudder with terror at the thought that infrastructure administrators could exploit the results from Fig. 2 to make cost-effectiveness calculations in an attempt to determine which types of instruments should be preferred over others. While this study provides some unique information on how well different sensors are able to detect the presence of initial condition errors, it also has limits that are important to remember:

1) This work focused on determining the presence of an observable signal and somewhat on the linearity of that signal. It did not begin to touch the issue of how well assimilation systems are able to use the signal especially when the field measured is related indirectly or in a complex manner to model variables. For example, a signal in the surface pressure data was observed for almost all initialization errors, but it is not clear that one could use that signal to determine the nature and magnitude of the initialization error.
2) The issue of the complementarities of the information between different sensors, or the lack of it, was not looked at. Different sensors may be sensitive to a particular field, say humidity, but depending on the technology used or the heights sampled, the information from these sensors may be highly redundant or not.
3) This study focused on the forecast of convective weather of the central Great Plains variety at mesoscale time and space scales only.
4) Its results are based on our best estimates of what errors currently exist in both model fields and measurements. As instruments and model initialization improve, these results will change.
5) They are also based on the assumption that the 4-km resolution runs used in our identical-twin simulations are realistic enough not to miss crucial effects or phenomena that would significantly alter forecast outcomes.

In the end, the most valuable aspect of this study is the approach we used and what information it can provide, but the exercise needs to be repeated in the appropriate context to be most useful. The results presented here are more intended to be used as a guide for assimilation efforts: which model variables seem to be the most critical to constrain, over what duration, and which technologies or datasets show the most promises.

In this context, what have we learned? A first question we had was how well can one use time series data from a sensor whose data are sensitive to changes in one model field to constrain other model fields. This is only possible if the errors introduced in one model field propagate quickly enough to other fields that its signal can be observed. Another crucial aspect, the linearity assumption, is whether the signal strength is proportional enough to the magnitude of the initialization error to be assimilated effectively. While initialization errors in one field quickly propagate to all other atmospheric variables, they do not always result in an assimilable signal. In particular, uncertainties in midlevel humidity, the ones that caused the largest forecast errors, also give the least linear signals on other fields. As a result, measurements from imaging sensors targeting properties of condensates like those of brightness temperatures from satellite imagers and reflectivity from radars provide a signal that cannot be easily assimilated over long periods. These will be useful to assess the location of clouds and precipitation at the initial time and constrain some of the other fields to have values compatible with the presence of clouds and precipitation; but one will have difficulties to use a time sequence of data over long periods to retrieve the forcing processes of the observed clouds and precipitation. The net result of this exercise is that to improve forecasts, one has little choice but to constrain midlevel humidity as much as possible with data from instruments sensitive to midlevel humidity such as radiosondes and radiometers or possibly a very large number of GPS sensors. If our assumptions on typical errors are correct, we found that errors on other variables have smaller effects and have a more linear response than those from moisture. Furthermore, temperature and winds are dynamically related, easing the retrieval of one knowing the other. For those, a variety of sensors can be used with radar measurements of Doppler velocity heading the list of the instruments considered, followed by in situ measurements by radiosondes and surface stations. In the end, one can see that at the mesoscale, there is no magic bullet: to be successful, one must assimilate data from a variety of sources to properly initialize models.

Another conclusion one can draw from this exercise is that there exist considerable opportunities for other sensors. In Part I, we found that the greatest current source of forecast errors are due to 1) uncertainties in midlevel moisture and 2) uncertainties in low- and midlevel temperature, low-level humidity, and midlevel winds. Given this result, one should first consider the performance of other sensors or techniques such as satellite-borne retrieval of temperature and humidity by profiling in the IR (Hayden and Schmit 1994) or by GPS receivers in space (Ware et al. 1996). But given that the first does not work in cloudy areas and that mesoscale coverage of GPS-based soundings would only be feasible in a distant future, one should consider developing new instruments targeting the variables listed above, ideally in that order.

6. In brief

The work of Part I was extended by testing the ability of various common meteorological sensors to detect a signal of initial condition errors. Scanning (radar) and framing (imagers) sensors gave the strongest signals. But these sensors generally measure quantities sensitive to properties of condensates, such as radar reflectivity or brightness temperature; these data cannot be assimilated easily for long periods because of the nonlinear response of these quantities to perturbations in humidity, the variable whose uncertainty introduces the greatest forecast error. Datasets that lend themselves to be assimilated over a longer time window are Doppler wind measurements and data from humidity-measuring instruments. But given the looser dynamical relationships between thermodynamics and dynamics at the mesoscale than at larger scales, one should strive to measure all fields or risk missing important information.

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 made possible by a sabbatical leave financed by McGill University. Thanks to Eunha Lim, Qingnong Xiao, and Sherrie Fredrick from NCAR/MMM for their technical assistance with the WRF model. The scientific input from Jenny Sun, Peggy Lemone, and Don Lenschow from NCAR/MMM, and from Barry Turner from McGill University is also recognized.

REFERENCES

Braun, J., C. Rocken, and J. Liljegren, 2003: Comparisons of line-of-sight water vapor observations using the Global Positioning System and a pointing microwave radiometer. *J. Atmos. Oceanic Technol.,* **20,** 606–612.

Fabry, F., and J. Sun, 2010: 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 *Mon. Wea. Rev.,* **138,** 242–255.

Gao, J.-D., M. Xue, A. Shapiro, and K. K. Droegemeier, 1999: A variational method for the analysis of three-dimensional wind fields from two Doppler radars. *Mon. Wea. Rev.,* **127,** 2128–2142.

Hayden, C. M., and T. J. Schmit, 1994: GOES-I temperature and moisture retrievals and associated gradient wind estimates. Preprints, *Seventh Conf. on Satellite Meteorology and Oceanography,* Monterey, CA, Amer. Meteor. Soc., 477–480.

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

Martin, L., C. Mätzler, T. J. Hewison, and D. Ruffieux, 2006: Intercomparison of integrated water vapour measurements. *Meteor. Z.,* **15,** 57–64.

May, P. T., K. P. Moran, and R. G. Strauch, 1989: The accuracy of RASS temperature measurements. *J. Appl. Meteor.,* **28,** 1329–1335.

McLaughlin, D. J., and Coauthors, 2005: Distributed Collaborative Adaptive Sensing (DCAS) for improved detection, understanding, and prediction of atmospheric hazards. Preprints, *Ninth Symp. on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface (IOAS-AOLS),* San Diego, CA, Amer. Meteor. Soc., 11.3. [Available online at http://ams.confex.com/ams/Annual2005/techprogram/paper_87890.htm.]

Menzel, W. P., and J. P. W. Purdom, 1994: Introducing GOES-I: The first of a new generation of geostationary operational environmental satellites. *Bull. Amer. Meteor. Soc.,* **75,** 757–781.

Montmerle, T., A. Caya, and I. Zawadzki, 2002: Short-term numerical forecasting of a shallow storms complex using bistatic and single-Doppler radar data. *Wea. Forecasting,* **17,** 1211–1225.

Sokolovskiy, S., Y.-H. Kuo, and W. Wang, 2005: Assessing the accuracy of a linearized observation operator for assimilation of radio occultation data: Case simulations with a high-resolution weather model. *Mon. Wea. Rev.,* **133,** 2200–2212.

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

Szyrmer, W., S. Laroche, and I. Zawadzki, 2005: A microphysical bulk formulation based on scaling normalization of the particle size distribution. Part I: Description. *J. Atmos. Sci.,* **62,** 4206–4221.

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

Tong, M., and M. Xue, 2005: Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS experiments. *Mon. Wea. Rev.,* **133,** 1789–1807.

Ware, R., and Coauthors, 1996: GPS sounding of the atmosphere from low earth orbit: Preliminary results. *Bull. Amer. Meteor. Soc.,* **77,** 19–40.