3B-01 WHY ARE RADAR DATA SO DIFFICULT TO ASSIMILATE SKILLFULLY?

Frédéric Fabry* and Véronique Meunier
McGill University, Montreal, Canada

1. MOTIVATION: UNSATISFACTORY BENEFITS

Radar is our best instrument to monitor and study storms. Data assimilation is our best approach to combine information from different sources. Numerical weather prediction is our best forecasting tool. Given these, one would expect that the assimilation of radar data into convective-scale models should provide the best storm forecasts. Yet, for example for precipitation forecasts, simple extrapolation of rainfall patterns often beat numerical forecasting with radar data assimilation for a couple of hours (Fig. 1).

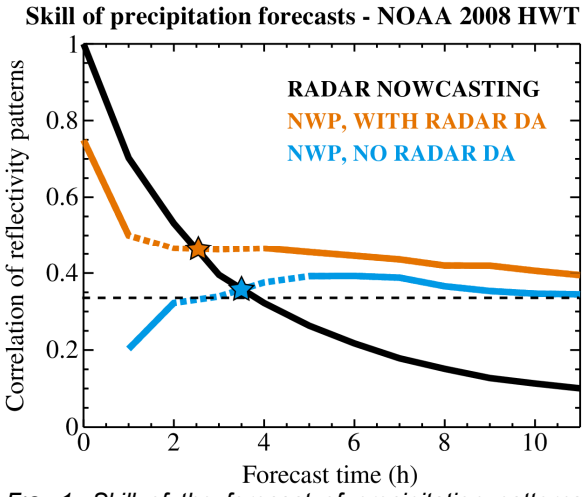

FIG. 1. Skill of the forecast of precipitation patterns made by three systems: the storm-scale ensemble forecasting system run by the Center for the Analysis and Prediction of Storms during NOAA's 2008 Hazardous Weather Testbed (HWT) Spring Experiment (Xue et al. 2008) without radar data assimilation (in cyan), the same system with radar data assimilation (in orange), and the McGill Algorithm of Precipitation Forecasting by Lagrangian Extrapolation (MAPLE; Germann and Zawadzki 2002; in black). Note the rapid initial decrease in skill of the NWP forecast with radar data assimilation. Image courtesy of Madalina Surcel based on the work of Surcel et al. (2015).

While there is no denying that radar data assimilation helps improve NWP forecasts (e.g., Dowell et al. 2011, Jones et al. 2015, and references therein), if these forecasts do not perform significantly better than extrapolation of echo movement, they are not as good as one should expect. If we examine forecast performance, we note that a) the skill of NWP aided by radar data assimilation often drops very rapidly in the first forecast hour (e.g., Fig. 1, or Fig. 10 of Aksoy et al. 2010) and/or does not have great skill immediately after assimilation (e.g., Fig. 15 of Mandapaka et al. 2014 or Fig. 8 of Supinie et al. 2017), and b) the assimilation of radar data yields significant gains for much shorter periods (e.g., about two hours) than that of other instruments (e.g., six hours for surface observations; L. Fillion, 2017, private communication).

This very rapid drop in skill suggests that radar data assimilation may not be as successful as that of other datasets. This is problematic given that we intend to largely rely on radar data assimilation to warn for storms on forecasts using numerical weather forecasting approaches (e.g., Stensrud et al. 2009, 2013). Understanding what makes radar data assimilation difficult is hence critical: What is fundamentally and/or practically different between the data from radar and that of other sources that may affect their assimilation, particularly in the context of convective-scale forecasting? Note that this question is different and complementary to the more traditional "*what makes convective-scale assimilation difficult*" for which we already have many answers (e.g., Yano et al. 2018). Since no simple experiment can be devised to answer the question at hand, a reasoning-based approach is adopted. The subject of our reflection will hence follow this axis: What could go wrong and why?

We shall accept as a starting hypothesis that data assimilation generally works given the existing evidence from its everyday use in operational forecast centers (e.g., Kwon et al. 2018). Given this hypothesis, the challenges of radar data assimilation must come from radar-related specificities, namely the context and conditions under which radar data assimilation is generally performed, peculiarities of the radar data themselves, and how radar data interface with the assimilation process. To understand what could go wrong, we must first collect the many puzzle pieces before we try putting them together; this involves documenting the characteristics of the data being assimilated and highlighting peculiarities that may cause difficulties. Issues arising from the interaction between these peculiarities and the assimilation machine will be subsequently investigated, sometimes numerically, sometimes using more qualitative reasoning, to expose the process by which the quality of either the analysis or the forecast resulting from the assimilation may be affected.

2. CONCEPTUAL OVERVIEW OF CORRECTION-BASED ASSIMILATION

At the risk of insulting the intelligence of data assimilation specialists, the first step in this investigation requires revisiting some of the key basic

* *Corresponding author address:* Frédéric Fabry, McGill University, Dept. of Atmospheric and Oceanic Sciences, Montreal QC, H3A0B9 Canada; frederic.fabry@mcgill.ca.

principles behind most of the current data assimilation approaches. The following narrative can only be incomplete but is merely meant to help understand the rest of this contribution. Readers interested in a more thorough introduction to the subject of data assimilation should consult, among others, the lecture notes by Bocquet (2019) and the book by Asch et al. (2016).

2.1 *Goal, Available Information, Process*

Data assimilation refers to techniques designed to combine optimally the information available about the state of the atmosphere at a specific time. The result of that optimal combination, referred to as the analysis, generally serves as the initial conditions for numerical weather forecasting. The information available to compute this analysis generally originates from two sources (Fig. 2):

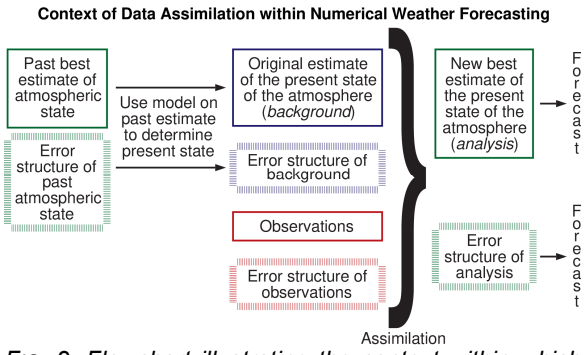

FIG. 2. Flowchart illustrating the context within which assimilation is performed. Adapted from Fabry (2015).

1) Background, or a-priori, information on the state of the atmosphere. This generally comes from a forecast made in the past but valid at the time of interest. It has the advantage of having information about all the atmospheric fields needed for a forecast, but this information has errors that need to be corrected;
2) Observations of the atmosphere made at or near the time of interest. These observations will be too sparse both in coverage and in atmospheric state variables to generate an analysis on their own, but they can be used to correct the errors of the background. One should not forget that these observations also have errors, either of measurement or of representativeness (e.g., the mismatch between the sampling volume of the observation and the volume represented by the analysis grid-point).

Because the background and observations are imperfect, the analysis cannot perfectly respect both sets of constraints. It is hence the task of the data assimilation system to generate the analysis that best combines the information provided by both sources of data. Most common assimilation approaches used nowadays such as those based on Kalman filters or variational approaches are "correction-based": The assimilation system computes the new analysis by adding corrections to the background to better respect the constraints provided by observations (the change to the background made by the assimilation system, or the difference between the analysis and the background, is referred to as the innovation). It attains this goal by simultaneously:

1) Minimizing the mismatch between the analysis and the background, accounting for background errors;
2) Minimizing the mismatch between actual observations and expected observations simulated using the analysis, accounting for the errors of observation, measurement simulation, and observation representativeness;
3) Spreading the innovation induced by observations from the locations and variables where the mismatch minimization occurred to other locations and variables.

2.2 *Error Dependencies to the Rescue*

The real power and magic of correction-based data assimilation results from the dependencies between errors, as it is what allows the assimilation system to spread the information coming from new measurements. Let us illustrate this process by analyzing the assimilation of a single measurement.

Consider a forecast for temperature such as Fig. 3 that we shall use as the background in an assimilation thought experiment. The expected temperature for Ocean City is 32°C, with a background error of ±1.4°C (not shown). We have a measurement at 29°C, with an observation error of ±1°C. Given the uncertainties on each, the optimum blend of these two constraints (32±1.4 and 29±1) is 30°C. The innovation on temperature at Ocean City resulting from the assimilated measurement is hence 30−32=−2°C. The analysis temperature at Ocean City is now computed (and its new uncertainty can also be estimated).

Since we lowered the temperature at Ocean City, should we change temperatures elsewhere? Consider Dover. Common sense suggests that it is very likely that the reason we made a mistake forecasting the temperature in Ocean City also applies in Dover; hence, it is reasonable to assume that we should also lower the temperature in Dover. By how much? Given the short distance between the two cities, the errors in their forecast are probably highly correlated and similar in magnitude; on that basis, we should probably change Dover's temperature by, say, −1.5°C, the exact value depending on the covariance between the two errors. How about New York? It is also possible that the reason we made a mistake forecasting the temperature in Ocean City also applies in New York, but not as likely as in Dover. Since there is a greater than 50% chance that the temperature in New York is also too high, we should slightly lower the temperature in New York too. For Chicago however, the processes that led to a temperature error in Ocean City are probably irrelevant there, and we have no reason to change temperatures that far from Ocean City.

The next question that arises is: What were the causes of this forecast error in Ocean City, and what have been their consequences? Causes may include increased cloudiness, or a stronger sea breeze from

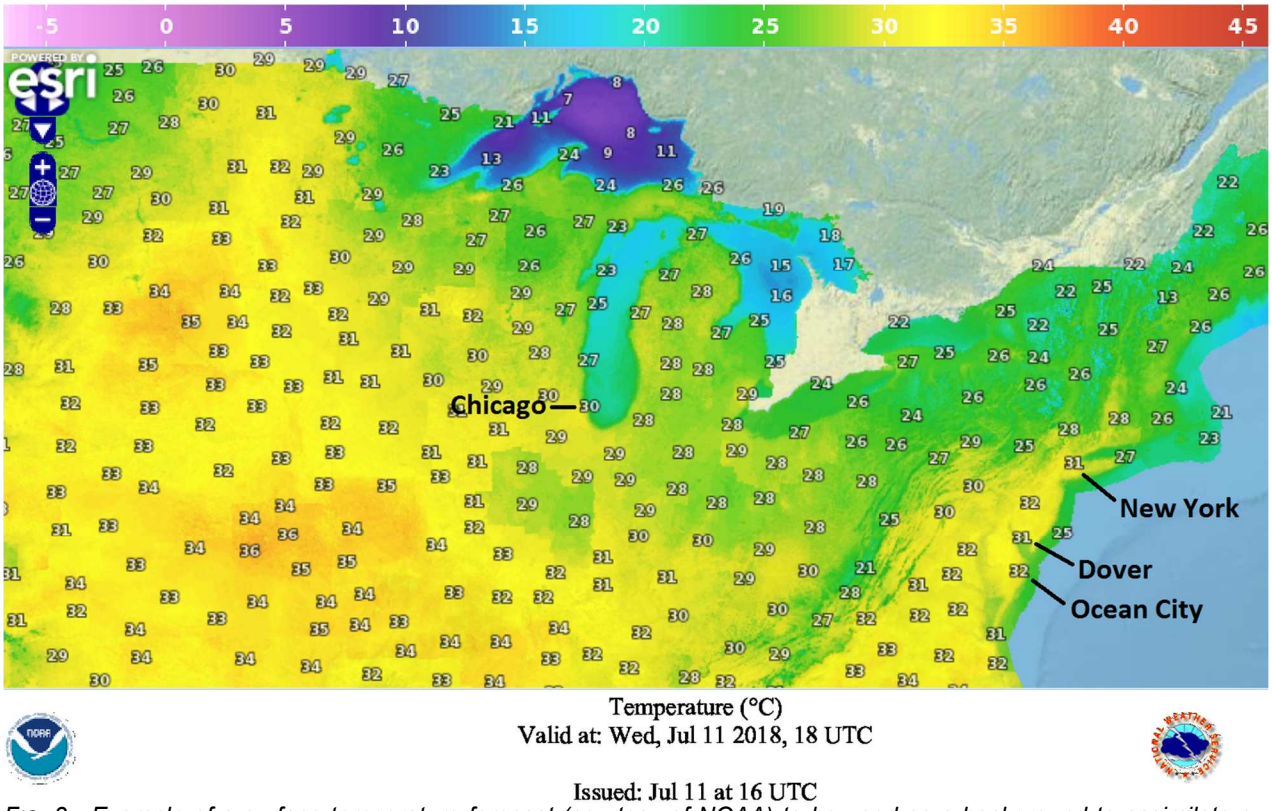

FIG. 3. Example of a surface temperature forecast (courtesy of NOAA) to be used as a background to assimilate a measurement at Ocean City.

the east; consequences may be a thinner planetary boundary layer, and increased surface humidity in it as a result. We should hence probably adjust those too. Suddenly, because we know that the surface temperature in Ocean City is wrong, we can modify many fields, and not only in Ocean City. This is only possible because errors in one quantity at one place, here the temperature at Ocean City, are physically or statistically linked with errors in other quantities or in other locations. Note that depending on the expected cause of forecast error, the sphere of influence of a correction will vary in size: If the likely cause of error is due to synoptic-scale processes, a single measurement can correct fields over large areas; if that cause is due to a localized phenomenon (e.g., convection), that sphere of influence will be more limited.

Such dependencies between errors are what allow correction-based data assimilation systems to correct much of the background with a relatively small number of incomplete observations and obtain a new analysis. Often, these error dependencies are assumed to be linear in variance, i.e.,

$$\sigma^2_{var} = k^2 \sigma^2_{H(x)} + \sigma^2_{unexplained}, \quad (1)$$

$$\text{with } k = correl[var, \mathrm{H}(\boldsymbol{x})] \frac{\sigma^2_{var}}{\sigma^2_{\mathrm{H}(\mathrm{x})}}, \quad (2)$$

where *var* is a state variable of model state **x** at one location, H(**x**) is the simulated measurement given a model in state **x**, σ_{var} and $\sigma_{\mathrm{H}(\mathrm{x})}$ are the expected uncertainties on *var* and on H(**x**), and correl[*var*,H(**x**)] is the linear correlation between *var* and H(**x**) across possible atmospheric states, a quantity either assumed or estimated, for example from an ensemble of backgrounds. If one knows σ_{var}, $\sigma_{\mathrm{H}(\mathrm{x})}$ and correl(*var*,H(**x**)), one can propagate an innovation computed for each observation *y* into *var* using $\Delta_{var} = k\left[y - H\left(\mathbf{x}_{background}\right)\right]$, and have *var* benefit from the reduction in uncertainty on H(**x**) via

$$\sigma^2_{var_{analysis}} = \sigma^2_{var_{background}} - k^2\left[\sigma^2_{H\left(\mathbf{x}_{background}\right)} - \sigma^2_{H\left(\mathbf{x}_{analysis}\right)}\right], \quad (3)$$

$$\text{with } \sigma^{-2}_{H\left(\mathbf{x}_{analysis}\right)} = \sigma^{-2}_{H\left(\mathbf{x}_{background}\right)} + \sigma^{-2}_{y}. \quad (4)$$

Simplifying to the extreme, existing correction-based assimilation approaches use the same basis for spreading the innovation; the primary difference between them lies in the assumptions, approaches and algorithms used to estimate σ_{var}, $\sigma_{\mathrm{H}(\mathrm{x})}$ and correl(*var*,H(**x**)) for every combination of state variables and observations.

2.3 *The Complex Nature of Error Dependencies*

Let us continue our assimilation experiment of the temperature in Ocean City. We have just received shocking news: The actual temperature is 21°C, 11°C cooler than expected. Once we confirm that this is not an erroneous measurement, how do we use it to correct other fields? Our previous assumption that the temperature error was probably due to increased cloudiness or a stronger sea breeze becomes doubtful. The most likely process that can lead to such a temperature error is an unpredicted storm outflow. Because the phenomena that cause such large errors are different from those that cause small ones, our ability to project a correction on temperatures elsewhere and on other fields will also change: How far does the storm outflow affecting

Ocean City go? Probably not to Dover, and certainly not to New York, unless they are themselves affected by different storm outflows. However, the atmosphere overall appears to be more unstable than expected as it can support unanticipated storms; but at any one place, the surface temperature could be either higher or lower depending whether that location is under the influence of a storm outflow or not. We hence have learned something important, that the atmosphere is more unstable, but how to best use that information is not clear. In parallel, up to where and in what direction do we extend the storm outflow beyond Ocean City? That could depend on the wind simultaneously measured in Ocean City: If it has a much stronger component than expected from a specific direction, say the west, the center of the outflow pattern and the associated storm are probably in that direction. Hence, where the corrections get applied could also depend on the magnitude of some measured property.

What this illustrates is that the magnitude of the mismatch between expected and real observations can also shape how any correction should be applied for optimal results. Since small perturbations in one atmospheric property generally cause a small linear response on other properties, the assumption of linearly-related errors works best under such situations. Large mismatches imply a vastly different atmospheric state, one where higher-order error terms become significant and where error linearity breaks down. These generally occur when the evolution of either the real atmosphere or the model state is determined by an atmospheric instability but not the other. Paradoxically, because major instabilities develop in specific ways that have been well documented, if one can extract the proper information by combining available data, one might make more appropriate choices on where and how to correct atmospheric properties.

3. CONTEXT OF ASSIMILATION AND NATURE OF DATA

The assimilation algorithm operates on data that have specific characteristics and error properties. Let us first focus on the characteristics of the background information and of the data at the meso- and convective scales where radar data assimilation is generally being performed.

3.1 *Quality of A-Priori Information*

Many mesoscale and convective-scale assimilation efforts start with background information originating from a global model analysis. These analyses are already based on the assimilation of data from a wide range of instruments. Smaller-scale data assimilation systems then take this background and add unused denser data such as from radars, satellite imagers, and in-situ measurements from aircrafts and surface stations (e.g., Benjamin et al. 2016).

What is the uncertainty of the global analysis used as background? This is not well known, especially at the smaller scales relevant to convective storms. To get a first order estimate on that uncertainty, we chose to look at the mismatch between the analyses of the control member of ensemble forecast systems from many global prediction centers: If all analyses from control members are similar, there is a good chance they are close to accurate; if they disagree, most must be wrong. On that basis, the average difference between these analyses likely represents a lower-bound estimate on their expected uncertainty.

Using data from The International Grand Global Ensemble (TIGGE; Bougeault et al. 2010), the analyses of many global centers were compared for June and July 2018 in the midlatitudes of the Northern Hemisphere. The results of that comparison for key atmospheric properties shaping convective storms (Fig. 4) suggest that while their structure is well captured at scales larger than 1000 km, it is poorly known at scales smaller than 500 km in summer, largely irrespective of the atmospheric property considered. Hence, correct meso-β and meso-γ structures must be rebuilt from either the terrain forcing, the assimilation of additional information, or arising processes at smaller scales. Because meso-β atmospheric structures evolve more slowly than smaller entities such as individual convective cells, they provide skill to longer-lead time forecasts. If data assimilation cannot rebuild cells and larger-scale structures, forecasts will suffer.

3.2 *Data Density Considerations*

What data are available to rebuild those structures, particularly at scales relevant to convection? Table 1 lists the number of measurements available over the typical scale (10×10 km) and lifetime (1 h) of a storm cell. Supposing that independent data are available at roughly 1 km resolution over 20 elevation angles and are measured every 5 minutes, radar provides approximately 25000 measurements of reflectivity and a somewhat smaller number of valid Doppler velocity measurements. Except for satellite data, nothing comes close to the density of radar information: as illustrated in Table 1, rare are the constraints from other data sources in the context of the scale and lifetime of a convective cell. By default, it will hence largely be the responsibility of radar data to correct for the missing or erroneous information of the background at meso-β and meso-γ scales.

But are 40000 constraints over the lifetime of a convective cell a little or a lot? For reference, in the 1980s, we primarily relied on radiosonde measurements to forecast synoptic-scale systems. For a typical 2000-km storm lasting five days, the information used from about 40 stations (one every 315 km in the continental US) launching radiosondes twice a day consisted of four measurements (temperature, humidity, and two wind components) at eight mandatory levels up to 20 kPa, or 12800 direct constraints. But because at these scales one can benefit from hydrostatic and quasi-geostrophic balance to some extent, and if we add surface stations, the actual number of constraints on atmospheric fields probably approached 20000. Hence, the number of constraints obtained by a radar

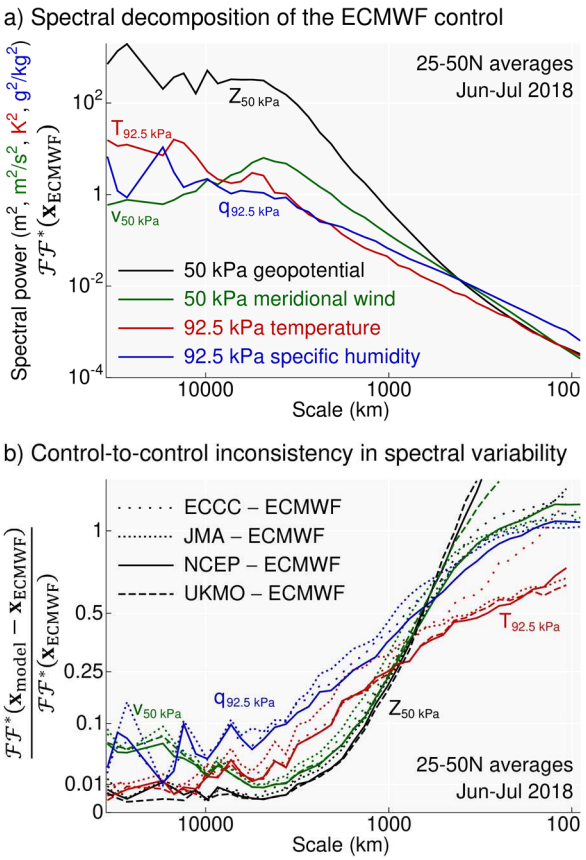

FIG. 4. a) Spectral decomposition of 50-kPa geopotential height (black), 50-kPa meridional wind (green), 92.5-kPa temperature (red) and 92.5-kPa specific humidity (blue) along lines of constant latitude and averaged over 25°-50°N in June and July 2018 for the analysis of the control member of the ECMWF ensemble system. Except at small scales, curves for control members of other global ensemble systems are very similar. b) Ratio of 1) the average spectral decomposition of differences between control members, and of 2) the average spectral decomposition of the ECMWF control, for the same atmospheric properties as in a). Ratios of 0 imply that the two control members considered are identical at that scale while ratios of 1 arise when the magnitude of control-to-control differences are comparable to that of the patterns in the ECMWF control. For each of the four properties listed above, four curves illustrate the relative difference between ECMWF controls and those of four other centers: The Environment and Climate Change Canada (ECCC, spaced dotted lines), the Japanese Meteorological Agency (JMA, tightly dotted lines), the National Center for Environmental Prediction (NCEP, solid line), and the UK Meteorological Office (UKMO, dashed lines). For all considered properties and analysis pairs, considerable inconsistencies exist at scales below 500 km, much larger than the resolution of all the models considered.

TABLE 1. Measurements of any atmospheric variable per hour over a 10×10 km area comparable to the size of a convective cell. For this exercise, a typical surface station reports information for five variables (pressure, temperature, dew point, winds, and precipitation).

Data source	Number of observations
Upper air observation of any variable using planes, radiosondes, GNSS receivers, etc.	≈ 0–10 (0–8 about thermodynamics)
Surface observations of any variable	≈ 0–10 (0–4 about thermodynamics)
Geostationary satellite (per thermal IR channel)	≈ 1,200 (largely from cloud top up)
Radar reflectivity	≈ 25,000
Radial velocity (assuming 60% echo coverage in storms)	≈ 15,000 in storms, 0 in no echo areas

on a convective cell is of the same order of magnitude as what was available in the 1980s to forecast synoptic-scale systems. There are however two key differences: First, while the data from radiosonde covered most key dynamic and thermodynamic fields, the data from radar are largely limited to precipitation and one wind component. Second, outside of stormy areas, the information provided by radar largely collapses to "no measurable echo".

An interesting sidebar is that in the 1980s, with the data provided by radiosondes, one could usefully forecast synoptic-scale storms over a period of five days, which roughly corresponds to the lifetime of the event itself. If the same logic applies for thunderstorms, one can at best reasonably hope to forecast the outcome of 10-km scale structures like convective cells for about their one-hour lifetime (see Fig. 9 of Sun et al. 2014) with the data that are currently available. This sets a sobering limit to what could ultimately be achieved with storm forecasting without additional data constraints.

3.3 *Operationally-Available Radar Data*

3.3.1 *THE INFORMATION FROM REFLECTIVITY*

Let us now focus on the nature of the information provided by radar, starting with reflectivity. The intensity of the radar echo at traditional surveillance radar wavelengths is mainly dictated by the number and size of the largest hydrometeors, as reflectivity largely depends on the sixth moment of the hydrometeor size distribution (e.g., $Z = \int N(D)D^6$ in rain). Radars have a range-dependent minimum sensitivity below which "*no measurable echo*" is reported. In the absence of significant beam blockage or attenuation, that threshold is such that most ice clouds and precipitation are observed, but liquid clouds are not. Reflectivity, or the sixth moment of the drop size distribution, is generally not a modeled quantity; to simulate their measurements from model state variables such as the mixing ratio of different hydrometeors, some assumptions must be made by

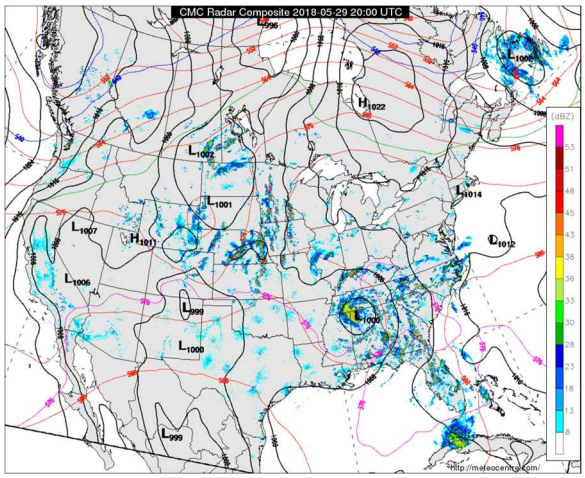

FIG. 5. Composite of reflectivity from Canadian and US radars on a late spring afternoon on which are overlaid black contours of sea-level pressure, red contours of constant 50-kPa heights and purple contours of constant 100-50 kPa thicknesses. Contrast the spottiness of reflectivity compared to the smoothness of pressure, reflecting the strong variability at small scales and the weak variability at large scales of precipitation fields. Image courtesy of the Canadian Meteorological Centre (CMC) and of meteocentre.com.

the assimilation system. Such assumptions being imperfect, the resulting error is generally correlated in space and time (e.g., Fig. 1c of Lee et al. 2007) and must be included in the error of the observational term. These correlated errors will generally dwarf accepted measurement errors (Keeler and Ellis 2000) unless the radar is not properly calibrated. In addition, only reflectivity measurements arising from hydrometeors should be assimilated, and contamination caused by ground, biological, or other echoes should be removed prior to assimilation.

More important is the spatial and intensity structures of the precipitation field sampled by radar reflectivity measurements. Meteorologically, precipitation is an intermittent double-threshold process: first, clouds occur primarily in updrafts when saturation is reached; then, precipitation grows only when enough clouds are present. As a result, it is sparse, covering only a few percent of the United States on average in summer, and also much more structured at small scales than other atmospheric fields (Fig. 5). The consequences of the unusual topology of precipitation fields on assimilation will be examined in Section 3. The relatively high structure of precipitation at small scales combined with the lower quality of background information at scales below 500 km (Fig. 4) cause precipitation background errors to be generally larger in magnitude and have shorter correlation distances than those of other fields.

In addition, radar cannot accurately measure reflectivity under some circumstances, particularly when echoes from hydrometeors are too weak or contaminated by other echoes. In such situations, "*no measurable weather echo*" is generally reported. We often take a shortcut and assume that this means "*zero reflectivity*". Complications arise when we assimilate logarithmic reflectivity (dBZ), as a value of minus infinity ensues. Since $-\infty$ cannot be used numerically, this problem is solved by setting all reflectivity values below a certain threshold to an arbitrary value such as 5 or 10 dBZ (e.g., Tong and Xue 2005; Gastaldo et al. 2018). This choice also conveniently eliminates many weak non-meteorological echoes. Ignoring for now the consequences of this choice, one should realize that fundamentally, the information being reported by radar is not "*zero reflectivity*", but instead "*reflectivity is below a threshold*". In the absence of attenuation or beam blockage, that threshold Z_{min} simply depends on radar characteristics and range. For example, on WSR-88Ds in precipitation mode, it is approximately $20\log_{10}(r) - 40$ dBZ, where r is the range to the target in kilometers. Sometimes, echoes were present but then suppressed because they were deemed to be primarily from non-weather targets such as insects or the surface. In that case, all we know is that any possible weather echo is much less intense than that of the contaminating echo. These two examples illustrate the lack of specificity of the "*reflectivity is below a threshold*" information and how it does not equate "*zero reflectivity*".

3.3.2 *REFLECTIVITY UNITS AND VALUES*

Radars are generally able to detect equivalent reflectivity factors Z_e ranging from less than 0.1 to more than 10,000,000 mm^6/m^3. For a variety of largely historical technical and practical reasons, logarithmic or decibel units (dBZ, or $10\times\log_{10}(Z_e)$) are generally used to archive and display reflectivity values. In addition, the distribution of reflectivity values from precipitation is unusual (top of Fig. 6): while, at any given height, most atmospheric fields generally have a quasi-symmetric distribution around an average, precipitation intensity and its radar reflectivity have distributions that follow more closely exponentials with a delta-function added at zero precipitation. In that context, using dBZ values for reflectivity has a potentially interesting property: the distribution of $10\times\log_{10}(Z_e)$ values is closer to a normal distribution (bottom of Fig. 6), and so, it is hoped, would the distribution of their errors. For all these reasons, reflectivity values in dBZ units are generally assimilated. The wisdom behind this choice will be examined in Section 6.a.1.

3.3.3 *THE "MEAN" DOPPLER VELOCITY INFORMATION*

The mean Doppler velocity estimates the radial component of target velocity toward or away from the radar. Its measurement is only available in the presence of echoes. In many ways, it is easier to assimilate than reflectivity as it does not suffer from calibration or attenuation bias and it is more directly related to model state variables. It is also the only direct constraint obtained concerning storm dynamics. Consequently, most researchers find that radial velocity assimilation help improve forecasts, and Doppler data are assimilated by many operational centers (Gustafsson et al. 2018).

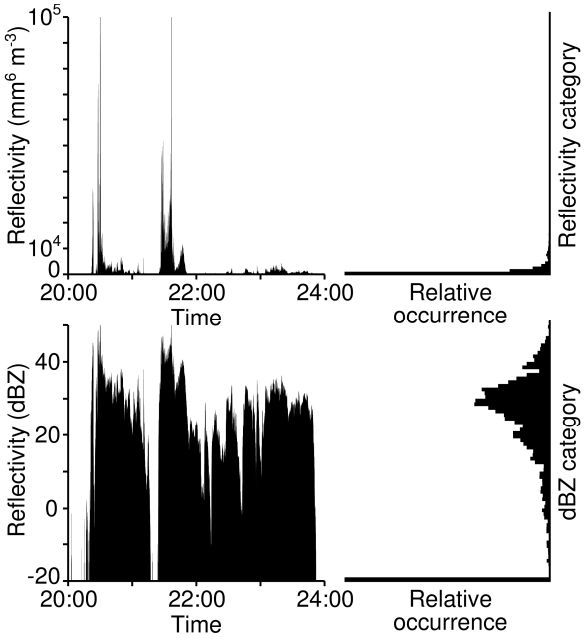

FIG. 6. Time series of reflectivity at one location from a rain event plotted using (top) a linear reflectivity scale and (bottom) a logarithmic reflectivity scale. On the right, the associated histograms of occurrence of different reflectivity are plotted sideways to match the reflectivity values on the left of the plot. Histograms of reflectivity generally resemble exponentials, while those of logarithm of reflectivity look more like normal distributions to which are added a second peak for reflectivity values corresponding to "no echo" (adapted from Fabry 2015).

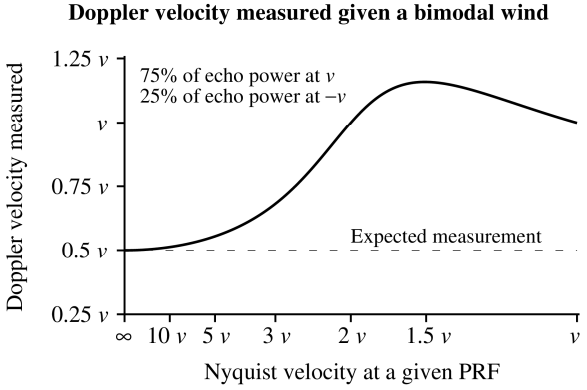

FIG. 7. Mean Doppler velocity measured by a radar as a function of Nyquist velocity when 75% of the received echo power comes with a radial velocity v and 25% with a radial velocity –v. While, using common sense and all currently-used observation operators, one would expect a measurement of v/2, this is not what standard radar signal processing algorithms such as pulse pair or Fourier transform "on a circle" (Keeler and Passarelli 1990) return as result. Note that the Nyquist velocity considered here is the raw one from a single pulse repetition frequency (PRF), not the combined value arising from the use of multiple PRFs.

Doppler velocity assimilation is however not free of problems. Target velocity is largely wind velocity plus fall velocity which can be significant for rain. It can be contaminated by surface clutter, birds, and insects, though this contamination is diminishing with improved radar signal processing. Under some circumstances, it can be aliased. Lesser known is that the signal processing algorithms estimating mean Doppler velocity are biased when the distribution of radial velocities is non-symmetric (e.g., Fig. 7, (4.5) of Zrnic 1979). And while perfect operators such as Eq. (9) of Fabry (2010) can handle these situations, they are too complicated to be usable. This leads to errors correlated in space and results in artificial convergence or divergence patterns as the slowly elevating beam gradually traverses regions of vertical wind shear.

3.3.4 *DUAL-POLARIZATION INFORMATION*

The increasing availability of dual-polarization radars is stimulating research on the assimilation of their data (Li et al. 2017; Augros et al. 2018; and references therein). Dual-polarization parameters such as specific differential phase K_{DP}, differential reflectivity Z_{DR}, and co-polar correlation coefficient ρ_{HV} provide constraints on precipitation mass, mean target shape and its variability. Measurements or retrieved properties can then be assimilated to improve precipitation simulation. While some benefits are observed in rain, the unknown and complex relationship between model-simulated properties such as hydrometeor type, number, and mixing ratio, and radar observations in ice remains an obstacle (e.g., Posselt et al. 2015).

3.4 *Summarizing the Assimilation Context*

Compared with many other measurements, radar data are more challenging to interpret: Reflectivity is a measure of higher moments of the hydrometeor size distribution than are generally modeled numerically; its estimate for weak and contaminated values is poor and ambiguous to interpret; Doppler velocity can be biased; polarization signals are weak and difficult to simulate; last but not least, the errors on measurements and their simulation are not well quantified and their spatiotemporal correlation not well known. By themselves, these issues would make radar data assimilation difficult. But they are probably minor compared to what follows.

4. ARISING ISSUES WITH ASSIMILATION

4.1 *Relevance of the Measured Information*

Radar has become a valued operational instrument for weather surveillance thanks to its ability to characterize storms and monitor their impacts. Therefore, we naturally think that its data should be equally skillful at helping us forecast these storms numerically. But radar is much better at detecting the threats of storms than at inferring their causes. In fact, radars detect precipitating cells relatively late in their life cycle. Consider the textbook evolution of

convective cells: Storms begin thanks to the combination of environment conditions that provide its fuel (temperature and humidity profiles), and a forcing mechanism that breaks the capping inversion preventing the storm to form earlier (e.g., via low-level convergence). Except sometimes for the forcing mechanism itself, radar is largely blind to these environmental conditions. As the unstable air parcel accelerates upwards, clouds form rapidly. But again, radar is largely blind to water clouds. Only when coalescence or ice crystal growth is significant enough do echoes appear aloft as illustrated by first echo studies (e.g., Knight et al. 1983) and numerical modeling of storms cells (e.g., Fig. 2 of Murakami 1990). By that time, the storm is well underway, and precipitation starts very quickly. It is at the latter stages of storm evolution that radar provides its best information. Unfortunately, this does not give us much time to use radar information to make a valuable numerical forecast for that cell. Note that the late appearance of precipitation during the release of an atmospheric instability is not limited to convective processes but is also the case for baroclinic systems.

We can complement the time-evolution view above with information obtained by considering atmospheric evolution in a process-oriented context. The box-and-arrow diagram in Fig. 8 illustrates the many interactions between different atmospheric properties relevant at weather timescales. Radars measure precipitation properties, and one component of the 3D wind where echoes are available. But the main driving force of storm dynamics is the pressure gradient arising from temperature and/or density contrasts, while the main control on storm intensity is low-level temperature and humidity. And though the dynamical and thermodynamic components of atmospheric motion are tightly coupled (top of Fig. 8), the water cycle largely follows one large loop whose last atmospheric component is precipitation (bottom of Fig. 8). In fact, the property best constrained by radar measurements, precipitation, 1) is the furthest from storm drivers, and 2) is short-lived without support, reducing its interactions with other fields. Indeed, if precipitation is modified in a model analysis and nothing else, that precipitation will fall out quickly without having had much time to influence the evolution of other atmospheric properties except through drag and evaporation. Only if assimilation properly changes the fields that cause precipitation may it have a lasting positive effect.

This discussion illustrates two additional challenges of radar data assimilation at convective scales: Direct information on storm properties arrive late in the time evolution of storms, and the atmospheric fields well measured by radar are remotely connected with those that shape the evolution of present and future events. We must hence largely rely on indirect information, namely what can current storm characteristics tell us about storm drivers in the past, and what can we infer from them concerning storm drivers in the future.

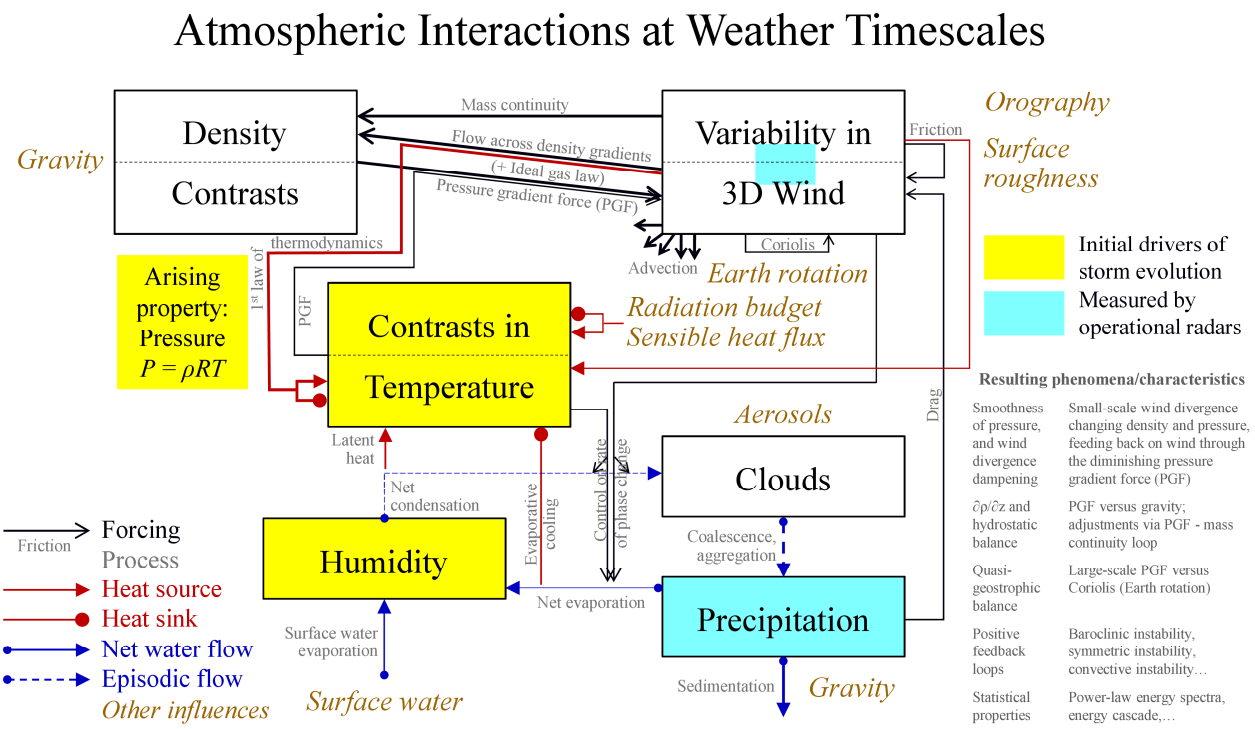

FIG. 8. System diagram illustrating critical interactions between different properties of an evolving atmosphere for time scales relevant to weather forecasting. Two interacting subsystems can be identified: The dynamics subsystems (largely in the top of the diagram) where interactions occur continuously largely through differential equations, and the water cycle (largely in the bottom of the diagram) where interactions tend to be more episodic. Highlighted elements contrast the initial drivers of storm evolution (in yellow) with properties constrained by radar (in cyan). Some properties arising from atmospheric interactions are also listed on the bottom right.

4.2 *Radar Innovation Sparseness*

This is where the nature of the precipitation field causes challenges. Over most of the troposphere, precipitation is zero (e.g., Fig. 5). This fact is generally well reproduced by models. In fact, for much of the model grid-space, the background and the measurements agree perfectly that echoes are absent. Radar data hence provide no innovation in these areas, and there is no reason to modify the background. Observations will induce an innovation only when they disagree with the background, and this occurs only in clustered areas where either the background or observations have precipitation. This limits the spatial extent of innovations introduced by the assimilation of radar data. More importantly, it leaves large regions unconstrained by radar measurements, including many that will play important roles in the future evolution of weather events.

Consider the tornadic supercell storm in the rounded rectangle in Fig. 9a that struck Tuscaloosa AL in 2011. It is a perfect example of the type of events we want to forecast well in advance, say with a lead-time of an hour. To forecast such a storm, we need to know the inflow that will provide its heat and moisture (identified as the colored letters A and B in Fig. 9b). We also need to know the properties of the air that will make up the surrounding environment and feed its downdrafts (colored letters C to E). An hour before, where is that air? It is in the areas bounding the white letters in Fig. 9b. Most of those are devoid of radar echoes; therefore, radar cannot obtain much information about them, at least directly. Note that in this case these areas are also covered by high-level clouds, limiting any help that could come from spaceborne imagers.

Therefore, to improve forecasts, it is critical for the available information to be projected from innovative precipitating areas out to other regions devoid of new constraints. While the need to project information outward is always present in data assimilation, it is surprisingly strong for radar data, given that data density should not be an issue: While data density is indeed high, because of the clustered nature of precipitation areas, there are huge regions where radar measurements do not directly innovate the background state. We have also seen that there are significant errors at scales up to hundreds of kilometers (Fig. 4) and gaps between precipitation areas comparable or exceeding such distances (Fig. 5). It is therefore imperative for assimilation systems to project information up to hundreds of kilometers outward to help improve forecasts of a few hours. In Fig. 9 for example, if information obtained from the supercells cannot reach the areas labelled with a white A and B, radar data assimilation cannot be used to improve our knowledge of the properties of the air that will feed the supercell updraft in an hour, which severely limits possible forecast improvement. If we want to improve forecasts beyond one hour, information must be projected even farther. And, at this time (Table 1), radar data is largely the information that must be used for that task.

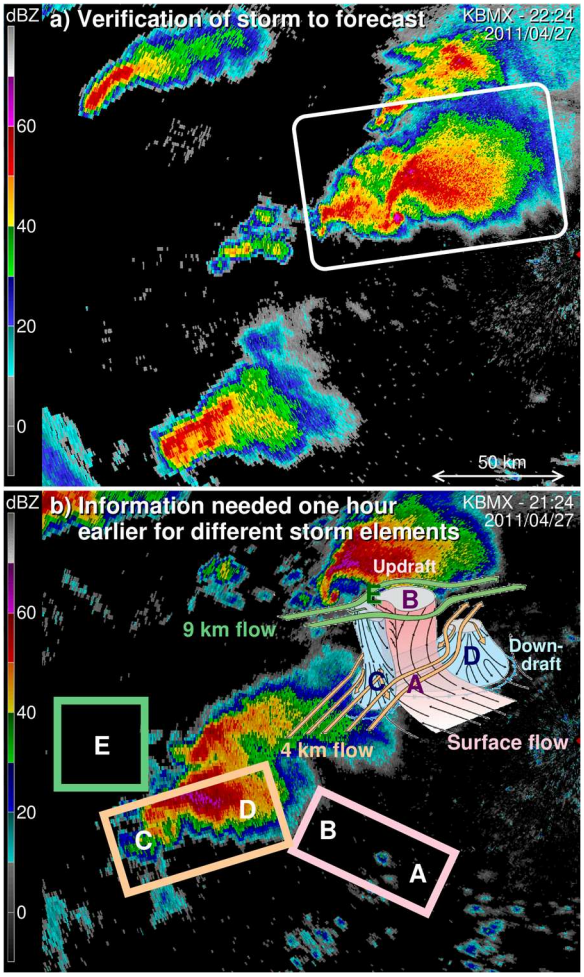

FIG. 9. Example illustrating the limited amount of direct constraints from radar data on areas that will shape the evolution of severe events. a) Radar reflectivity image of a tornadic supercell storm at time T that we wanted to forecast at T−1 hr. b) Composite image of i) a conceptual diagram of a supercell storm at the position of the storm in a), and of ii) the reflectivity image at T−1 hr. On the diagram courtesy of Markowski and Richardson (2010), key areas are labelled by colored letters. The air in these areas at time T originates at T−1 hr from the colored rectangle areas with corresponding white labels. Most of these are in echo-free regions; as a result, the properties of the air in these source areas are unlikely to be directly constrained by radar measurements.

4.3 *Obstacles to Information Propagation*

The uncertainty on state variables far away from storms can only be reduced if we can devise relationships between values or errors of characteristics of radar echoes (reflectivity, radial velocity, etc.) and values or errors of atmospheric properties well outside echo regions. Usually, these relationships are based on the covariance between errors in observational constraints and errors in state variables. To be useful for assimilation, these relationships must first exist, either physically or statistically, and they must also either be known a

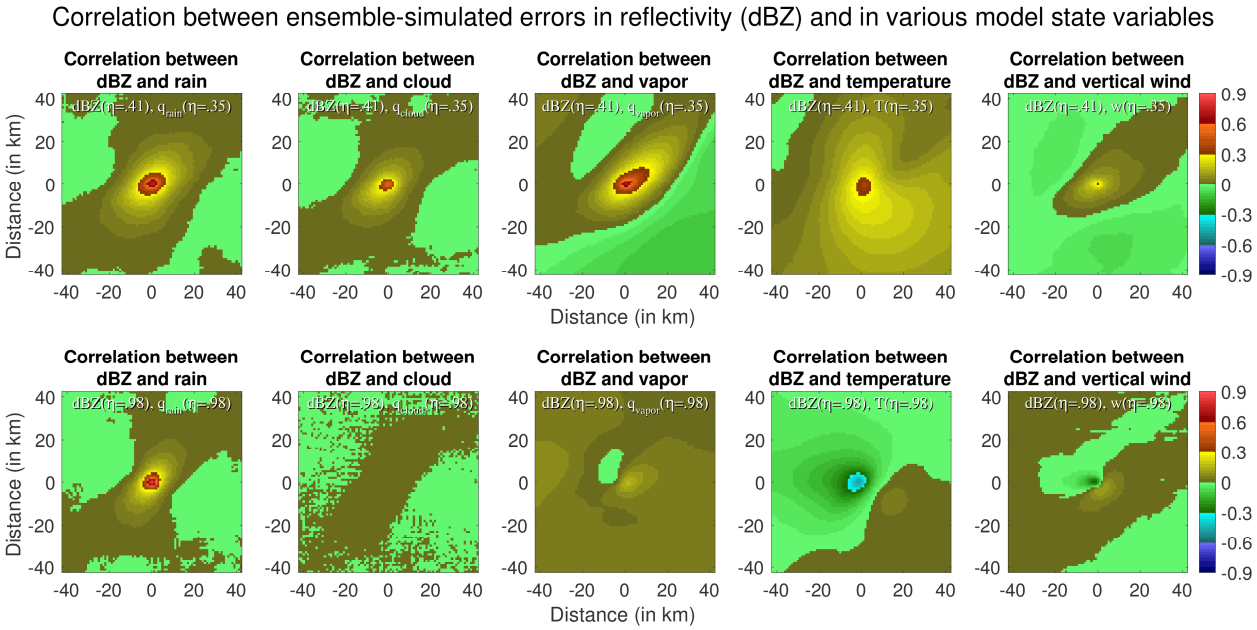

FIG. 10. Average correlation between simulated reflectivity measurements from the ensemble generated by Jacques and Zawadzki (2015) and nearby background state variables of rain mixing ratio q_{rain}, cloud mixing ratio q_{cloud}, water vapor mixing ratio q_{vapor}, temperature T, and vertical velocity w in the upper troposphere (top row) and the lower boundary layer (bottom row). In all cases, limited correlation is observed beyond a few kilometers. And while reflectivity measurements in the upper troposphere (here at eta-level η of 0.41) can provide some information on upper troposphere state variables (η=0.35), those in the boundary layer (η=0.98) are not as skillful at constraining boundary-layer state variables (η=0.98). For this calculation, all reflectivity measurements below 10 dBZ were set to 5 dBZ, and at least one member had to have a reflectivity greater than 10 dBZ for the resulting correlation matrix to be included in our calculation.

priori (like in 3DVar) or determined based on available information (like in ensemble-based approaches).

Unfortunately, precipitation 1) is a delayed outcome of its root causes (Section 4.1) and 2) has smaller-scale errors than other fields (Section 3.3.1), both complicating the relationship between precipitation intensity and other atmospheric properties locally and at larger distances. As a result, the correlation between errors on point precipitation values and on other properties is on average weak (Fig. 10). This is particularly the case at low levels where, by the time the rainfall arrives, the dynamical processes that led to its creation occurred a long time ago. Nonetheless, the correlation (ρ) between observed quantities and state variables must be strongly positive or strongly negative for effective error reduction: Given a perfect observation, the error on the correlated state variable will be reduced by a fraction $\left(1-\sqrt{1-\rho^2}\right)$ ((A5) of Jacques et al. 2018). As such, correlations of 0.2, 0.4, 0.6, and 0.8 will then respectively lead to an error reduction of only 2%, 8%, 20%, and 40%. If observations have errors, the resulting innovations and error reduction will be even smaller. Since high correlations between precipitation and other state variables are rare, especially at longer distances (Fig. 10), greatly reducing errors on other state variables by assimilating precipitation is difficult.

While the theory of using error covariances to reduce the uncertainty in initial conditions is sound, it is only truly effective if error correlations are high. In the atmosphere, this occurs primarily when errors are small enough to grow linearly and when higher-order error terms or unexpected instabilities have minor consequences. However, in precipitation in general, and particularly in convection, instabilities play key roles. In fact, precipitation largely arises as a delayed response from atmospheric instabilities, whether baroclinic, convective, or otherwise. Since radar largely provides information on precipitating areas, that information will hence generally be concentrated in areas where instabilities are being or have just been released. A resulting challenge is that while background errors might have been small and Gaussian-distributed before the triggering of the instability, this is no longer the case by the time one detects a significant difference between simulated and real radar measurements. Outside of these areas, radar generally only reports the absence of precipitation.

This has a few consequences that affect the potential performance of radar data assimilation positively and negatively: 1) Radar provides most of its information in regions where more constraints are badly needed to document the rapid evolution of the atmospheric state arising from an unfolding instability; 2) The very nature of the rapidly-evolving atmosphere in unstable regimes makes any information more difficult to use, especially if that information is limited to a small subset of atmospheric properties as is the case for radar; 3) Contrary to most other data sources, radar provides little information in more stable atmospheric regimes where error reduction could be easier to achieve with incomplete constraints; 4) The exception to the above is when a

radar observation of no significant echo invalidates an instability release in a model. Indeed, many researchers find that suppressing model precipitation where it does not occur contributes significantly to the skill of radar data assimilation (section 2c of Wattrelot et al. 2014 and references therein).

4.4 *Resulting Forecast*

What will then happen when radar data are assimilated? While observed fields such as precipitation and one wind component will be significantly corrected in the analysis, key unobserved fields such as temperature and humidity will only be marginally updated from the background, and only in or very close to precipitating areas. The storm dynamics and thermodynamics will hence be largely incompatible with the updated precipitation field. When the forecasting model will then be run, after some time, the modified precipitation will have fallen and be replaced by one compatible with other fields; the precipitation will then largely resemble that of the background. The model will then appear to have forgotten the assimilated data and returned to a trajectory somewhat resembling the one prior to assimilation.

This problem goes beyond the ability to forecast properly a storm in the "distant" future. For example, let us consider what happens when assimilating observations of a burgeoning cell in a background that lacks such convection but that is otherwise ripe for it. The intimidating Fig. 11 contrasts the evolution of a model cell in a WRF nature run with that of an identical twin model run from one member of a fifty-member ensemble driven by assimilating reflectivity data in an Ensemble Adjustment Kalman filter (EAKF) framework using the Data Assimilation Research Testbed (DART, Anderson et al. 2009). Focusing first on the nature run, we observe that the growing cell is driven by boundary-layer convergence that feeds a growing updraft, saturates air, and generates increasing amounts of precipitation increasingly higher as time progresses. The modeled cell behaves differently: While, at every time step, precipitation is reasonably well corrected by reflectivity data assimilation, changes in dynamical and thermodynamic properties are insufficient in the bottom half of the troposphere. One reason for winds is that, especially at low levels, heavier precipitation can be associated with stronger updrafts early in the cell's lifecycle, or stronger downdrafts late, as a result

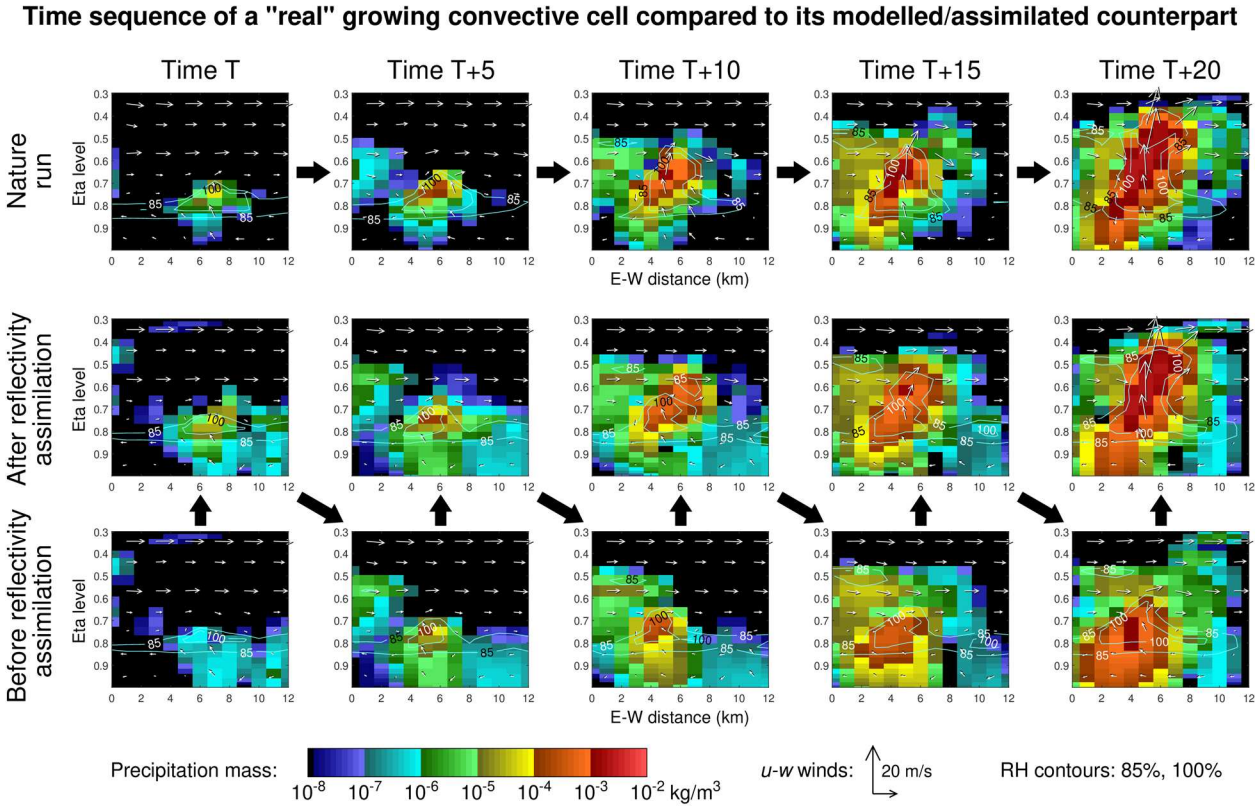

FIG. 11. Comparison of the time sequence of a "real" growing cell (the nature run of an identical twin experiment) with that of the same cell whose evolution is driven by assimilating reflectivity "data" from the nature run. Each frame represents an east-west cross-section of precipitation mass in color and of zonal (u) and vertical (w) winds as vectors, on which are superposed contours of 85% and 100% relative humidity in cyan. Block arrows illustrate the progression of time on the nature run row and that of the assimilation and modeling cycles in the two bottom rows. While precipitation mass is relatively well assimilated at every step, changes in winds and humidity, among others, are generally insufficient in the bottom half of the troposphere. As a result, the modeled cell is anemic and does not evolve like the real one. Model runs are from Sodhi and Fabry (2020) and the assimilation is based on EAKF under DART and done with reflectivity data at nine levels every 1 km horizontally with a 3-km localization window.

of which limited net covariance exists between precipitation and updrafts if members have both growing and mature cells. At higher levels, precipitation and updrafts are generally collocated, hence assimilation performs better there. Another reason for humidity is that many precipitation intensities are associated with relative humidity near 100% (e.g., $RH \approx RH_{environment} + (1 - RH_{environment})[1 - exp(-AR)]$, A being an arbitrary positive "constant"). When a linear relationship between errors in *R* and in *RH* is derived by the assimilation system, it leads to under-correction in *RH* in weak precipitation and over-correction in strong precipitation. As a result of these two non-linear relationships, 1) the lower-level updrafts (up to T+15 min) and downdrafts (T+20 min) in the precipitation core are not well captured, and 2) the air is generally not saturated with humidity where precipitation is added at the top of the core except at T+20 min. Hence, when the model runs, evaporation suppresses the already-weak updraft and precipitation fails to grow sufficiently. In the end, the modeled cell does not flow or generate precipitation, sensible and latent heat like the real cell, which is why it cannot evolve in the same way. It will also be unable to correctly interact with neighboring storms.

Reflecting on the analyses in this section, we surmise that, at its root, the rapid skill loss observed in precipitation after radar data assimilation (Fig. 1) stems from the inability of assimilation systems to correct unobserved fields, including those away from storms. This arises from the fact that radar primarily takes measurements in precipitation, a clustered field with large areas of zero value and that arises from atmospheric instabilities where errors are large, have limited spatial correlation, and have nonlinear relationships with other state variables that will ultimately dictate its future evolution.

5. COMPOUNDING PROBLEMS

Furthermore, we should not forget additional difficulties linked with the field of precipitation, its sampling with radar, and its modeling.

5.1 *The Lack of Large-Scale Variability of Rainfall*

As illustrated in Fig. 5, compared to other atmospheric fields, rainfall has considerably more small-scale variability. As a result, the errors in forecasted rainfall will also be smaller-scale errors. Because such forecasts are used as background fields by assimilation systems, an additional obstacle arise: The successful modeling of rainfall relies on the ability of models to simulate detailed smaller-scale processes, something they often do poorly because of limited resolution and microphysics. The lack of effective model resolution limits the model's ability to simulate what can be observed and hence the benefit of innovation from observations. A case could be made that since models cannot simulate the finest-scale processes of precipitation shaping what radars measure, perhaps we should not assimilate measurements at the highest available resolution. Figuring out what better observational constraint to provide should increase the effectiveness of radar data assimilation.

5.2 *Radar-Measured vs. Modelled Properties*

Radar measures beam- and range-weighted reflectivity, a property linked with the sixth moment of precipitation. In echo areas, it also measures other properties with beam, range, and reflectivity weighting. Models simulate a few moments of precipitation on a grid, but generally not the sixth. In the process of simulating radar observations, the assimilation system must convert the moments simulated by models into reflectivity, a process not unlike what radar meteorologists do when converting reflectivity to rainfall rate using Z-R relationships. It must also do so accounting for the geometry of radar measurements. Since neither calculations are done perfectly, errors that are correlated in space will result. For example, the same microphysical processes that cause drops to be unusually large or small at one location also act in nearby areas. In the 70 years we have had Z-R relationships, we have not definitely characterized the statistical properties of their error, especially their spatial error covariance. This must now be achieved to help data assimilation (DA).

Additionally, assimilation approaches are based on the hypothesis that data are not biased. Calibrated data everywhere are hence a necessity, and so are unbiased observation simulations.

5.3 *The National Nature of Radar and DA Work*

Lastly, each radar system is a unique instrument with unique quirks. Some homogeneity exists within each country, but despite standardization efforts, it is difficult to keep track of peculiarities in the hardware, operation, and data processing of radars from each nation. These also evolve: At the time of this writing, in the US, version 18.2 of the radar data processing system is being used, the first dating circa 1990. And each time a new processing system is being used, the accuracy, biases, and usability of radar data may change. And all radar data assimilation efforts must be duplicated in every country and adapted to each assimilation system.

6. HOW TO CONFRONT THESE DIFFICULTIES?

Radar data assimilation is difficult and prone to failure for a multitude of reasons summarized in Table 2. But, for the foreseeable future, it will largely remain all that we have to enable the use of NWP for storm warnings. We must therefore try to make the best of a difficult situation. Doing so requires confronting the challenges specific to radar data assimilation, recognizing that unique problems may require unique solutions.

While the picture we painted until now may look bleak, we do not believe that the situation is hopeless. For example, many numerical experiments such as Crook (1996) show how small changes in environmental humidity or temperature lead to large

TABLE 2. Conditions of success and possible causes of failure of radar data assimilation.

Conditions of success of data assimilation	Some radar-specific challenges
***Observability condition*:** The analysis or its error will differ from the background or its error if measurements add value to the simulated observations from the background	1) Operational radars measure a limited number of properties; 2) over most of the model domain, observations provide no new information given the absence of both real and simulated echoes
Reproducibility condition: The model must have the variables, physics, and resolution needed to 1) generate the real current atmospheric state and 2) simulate accurately the physics of the measurement process	1) The small-scale processes that shape rainfall and its reflectivity are often not resolvable; 2) accurate observation simulation is too complex or impossible given the available information (e.g., most model microphysics do not have the information needed to compute reflectivity correctly)
Error characterization condition: Observations and their simulation must be statistically unbiased, and their error covariance must be known	Radar measurements contain artifacts and have poorly-characterized correlated errors; simplistic observations simulation also introduce undetermined correlated errors
Usability condition: The assimilation system must be able to use the observation-background mismatch to efficiently adjust the model state at the grid-point where the observation occurred	Imperfect relationship between errors in radar reflectivity (especially in dBZ values) and in related state variables such as precipitation mixing ratio
***Propagability condition**:* Innovations in observations can be propagated to state variables at other locations if usable relationships between them exist and are known to or determinable by the assimilation system	Convective patterns have primarily small-scale variability and their error covariance with other state variables is limited, especially at larger distances and in the context of large background errors
***Relevance condition**:* Changed fields must play a significant role in the future evolution of the weather patterns of interest for forecasts to be improved	Correcting precipitation has the least impact on the future outcome of storms; key fields such as temperature and humidity are harder to modify using radar data

changes in precipitation and its patterns. Hence, one should be able to take advantage of that fact to improve our knowledge of the said temperature or humidity. We simply seem to fail to achieve this with current practices. In parallel, instabilities often evolve in expectable ways, and this could potentially be exploited. The challenge becomes finding new approaches that account for the strengths and limitations of radar data as well as the nature of the problems we face. Two key goals should be sought in priority: 1) Correcting fields far away from precipitation, and 2) Devising error-reduction methods that work well in the context of large and rapidly-growing background uncertainty in atmospherically unstable regions.

We do not know how to successfully face these challenges yet. What follows are a set of possible and complementary avenues of inquiry arising from the issues raised in the previous sections. Several ideas are explored, but the list is far from exhaustive, its purpose being primarily to stimulate reflection.

6.1 *Simpler Adjustments to Current Approaches*

6.1.1 *To dB or not to dB? More Complex Relationships*

We mentioned in Section 3.3.2 why reflectivity measurements Z are generally assimilated in dBZ or log(Z) units. Because reflectivity and the precipitation rate R from any individual type of precipitation are closely linked through a power-law relationship, the previous statement is akin to saying that we are generally assimilating $\log(R)_{radar}$ measurements from radar. Traditional assimilation systems then estimate or use a linear relationship between errors in $\log(R)_{bkgd}$ in the background and those in state variables. Once the assimilation system has combined the $\log(R)_{bkgd}$ and $\log(R)_{radar}$ to determine a new $\log(R)_{anal}$ for the analysis, the innovation $[\log(R)_{anal} - \log(R)_{bkgd}]$ is then used to update state variables by $[\log(R)_{anal} - \log(R)_{bkgd}]\mathbf{\Delta}$. Assimilating reflectivity in dBZ units hence leads to the following proportionality relationship: If an innovation a in log(R) is associated with a state change $a\mathbf{\Delta}$, then an innovation b in log(R) should be associated with a state change $b\mathbf{\Delta}$.

In that case, the following example is assumed to be true: If a change in rainfall R from 1 mm/hr to 10 mm/hr ($\log_{10}(R)$ changing by 1) is associated with an increase in updraft velocity w of 0.3 m/s, then a change in R from 1 mm/hr to 100 mm/hr ($\log_{10}(R)$ changing by 2) leads to an increase in w of 0.6 m/s. At face value, this is ridiculous. We know that, physically, the rate of water vapor condensation is proportional to w, and if clouds do not accumulate mass, rainfall must be linearly linked to the rate of condensation and hence to w. Changes in w are not linked to changes in log(R): While updrafts of 0-0.1 m/s and 0.3-0.4 m/s would sustain rainfalls of 1 mm/hr and 10 mm/hr respectively at steady state, an updraft of 0.6-0.7 m/s will not be sufficient for a 100 mm/hr rainfall. Continuing this example, a rainfall of 0.1 mm/hr ($\log_{10}(R)$ changing by −1) would then lead to a downdraft of 0.2-0.3 m/s, also an unlikely scenario.

These inconsistencies arise because of the assumed linearity between increments in simulated

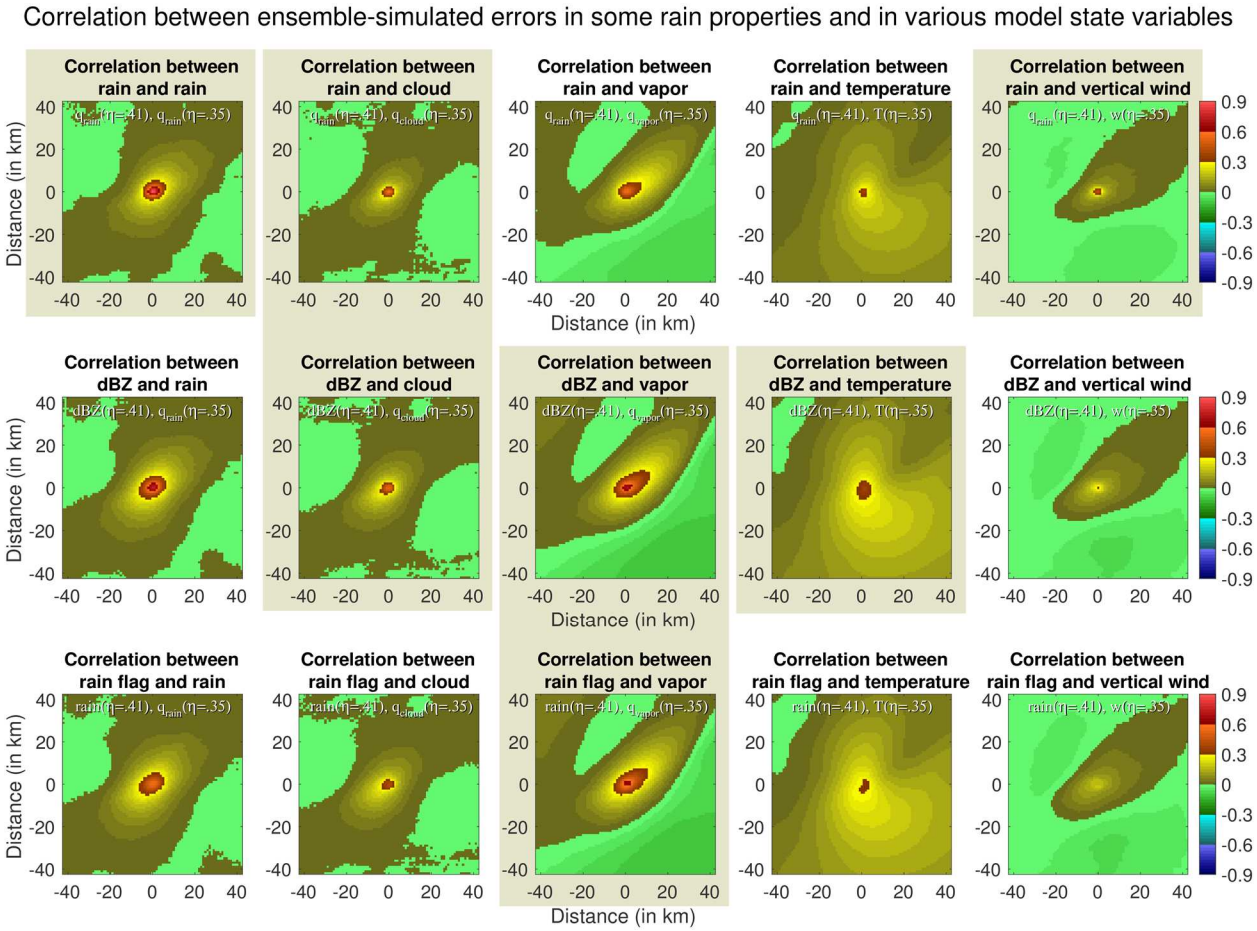

FIG. 12. Average correlation between either rain mixing ratio (top row), simulated reflectivity measurements (middle row), and a flag of presence of echoes greater than 10 dBZ (bottom row) from the ensemble generated by Jacques and Zawadzki (2015) and nearby background state variables of rain mixing ratio q_{rain}, cloud mixing ratio q_{cloud}, water vapor mixing ratio q_{vapor}, temperature T, and vertical velocity w in the upper troposphere. Shaded figure backgrounds highlight for which of the three radar-derived quantities we observe stronger correlation with each state variable.

observations and increments in state variables. Here, because assimilated observations are proportional to log(R), this assumption leads to improbable results. Note that if errors were small, the non-linearity of the logarithm function would not matter. But since errors are often large, it becomes important.

We are not obliged to assimilate reflectivity as $10\log_{10}(Z)$. We can take advantage of the fact that $Z \approx a_{ZR}R^{b_{ZR}}$, or also $Z \approx a_{Zq}q_{rain}^{b_{Zq}}$, to assimilate $Z^{1/b_{ZR}}$ or $Z^{1/b_{Zq}}$, quantities roughly proportional to the rainfall R and to the mixing ratio q_{rain} respectively. This transformation would not be unprecedented, as satellite measurements are often assimilated using the transformed quantity brightness temperature T_{BB}, and not as rawer measurements of radiances (e.g., Garand 2003). This is both for reasons of convenience, errors tending to be constant in T_{BB}, and because it leads to near-linear relationships between innovations in T_{BB} and innovations in wanted properties such as T. The main benefit of assimilating a quantity which is proportional to R or q_{rain} would be its stronger link to other state variables whose errors we are desperately trying to reduce.

However, we do not expected errors in all atmospheric properties to be better correlated with rainfall errors. Errors in humidity for example probably depend more on the presence of enough precipitation than on variations of its intensity. We can for example contrast the correlation between errors in state variables with errors in rain mass, in the logarithm of reflectivity, or in the presence of echoes greater than 10 dBZ (Fig. 12). It confirms that for some state variables such as precipitation and updrafts, assimilation rain amounts would lead to better error reduction, while for vapor, assimilating the mere presence of echoes does as well as any richer quantity. What Fig. 12 illustrates is that using some physically-based reasoning, we can determine which relationship should perform better, and use it to our advantage. Here we have limited ourselves to three simple scenarios, but more could be explored.

6.1.2 *MEMBER RELEVANCE AND CONDITIONAL SELECTION*

The use of more complex relationships between errors in radar-measured quantities and in state variables is actually a timid response to a much more complex problem. Initial background errors are large (Fig. 4). Especially in the context of an atmosphere conditionally unstable for convection, these then grow

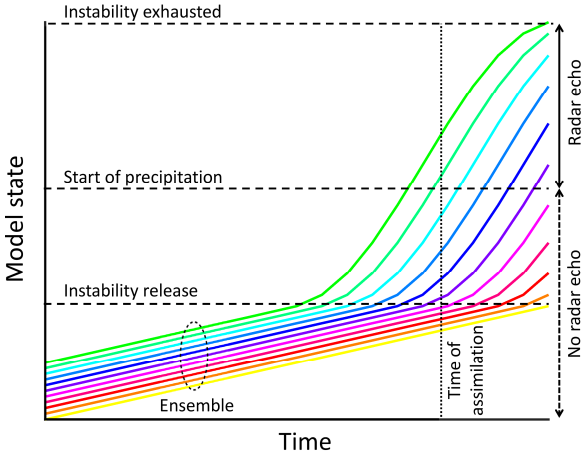

FIG. 13. Illustration of the evolution of the phase state of ensemble members as a function of time as they experience the release of a convective instability at different times. At the annotated time of assimilation, some members (yellow to red) have not experienced the instability, while only a subset of them (in green) have precipitation. The task of the assimilation system is to devise a way to correct the model state from all those members using radar observations given a truth that may lie in any of the interval areas in the vertical.

rapidly, destroying simple error relationships. Figure 13 conceptually illustrates this process: It is not uncommon to have, for any specific region, members with no convection, some with growing convection, and some with mature convection. In the absence of echoes, no relationship is expected between echo strength and any state variable; in the growing stages of a cell, updrafts, cloudiness, and temperature positively correlate with echo strength, or more accurately with echo intensification; in the mature stage, the opposite occurs, stronger precipitation cores being associated with stronger downdrafts and cold pools. Under such circumstances, any attempt to find a single relationship, however sophisticated, between errors in echo strength and state variables is doomed: If the assimilation system tries to derive unique relationships applicable to all situations, the result will be an odd mixture of all of them. Consequently, error reduction is unlikely to be very effective.

This discussion highlights the importance of meteorological context and how it affects the applicability of expected relationships between observed and unobserved quantities. Context can include meteorological relevance, such as whether the members are in the correct phase of the evolution of the cell. It can also be spatial, such as whether the pixel considered is at the center of the cell, ahead, or behind.

A better approach may be to use simulated and real observations to determine the pertinent context, and then derive a more appropriate relationship accordingly. As a simple example, let us assume that we know from observations that the state variables to be corrected are in the middle of the precipitation core; we can hence repeat the computations done in Fig. 12 but only selecting members where a local reflectivity maximum is within 1 km of the point of interest. The result, shown in Fig. 14, shows that much higher error covariance can be achieved. This implies that relevant members could be better corrected.

What this leaves out is what to do with the irrelevant members. For those, a very different type of error correction must be considered. In the past, these have included, among others, moving patterns around (e.g., Brewster 2003a,b; Stratman et al. 2018), or adding/suppressing surface heat and column moisture to create/suppress storms using ad-hoc processes (e.g., Wattrelot et al. 2014; de Lozar et al. 2018). These and new approaches should be explored more thoroughly.

6.1.3 *LARGER-SCALE CORRECTIONS*

As seen previously, precipitation varies considerably at small scales, and proportionally less at larger scales than other fields (Fig. 5). As a result, errors in distant atmospheric properties are unlikely to be correlated with errors in point rainfall (Fig. 10): Errors in rainfall at any particular point are instead generally caused by storm displacement errors or morphological differences in the precipitation pattern. But if smaller-scale errors or patterns are filtered, error correlation increases (e.g., Fig. 15). This is partly because rapidly-evolving small scale patterns with large errors are suppressed, and slowly-evolving

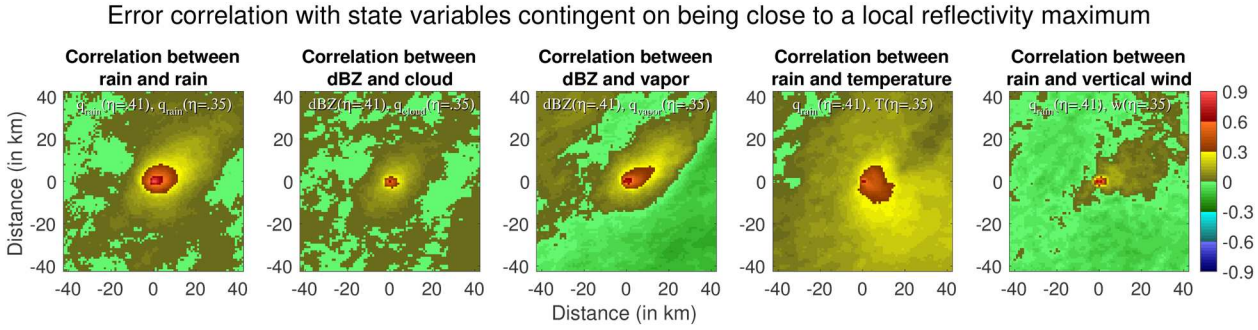

FIG. 14. Average correlation between either rain mixing ratio or simulated reflectivity measurements from the ensemble generated by Jacques and Zawadzki (2015) and nearby background state variables of rain mixing ratio q_{rain}, cloud mixing ratio q_{cloud}, water vapor mixing ratio q_{vapor}, temperature T, and vertical velocity w in the upper troposphere. As opposed to Fig. 10, average error correlations in excess of 0.6 can be observed for all these background variables.

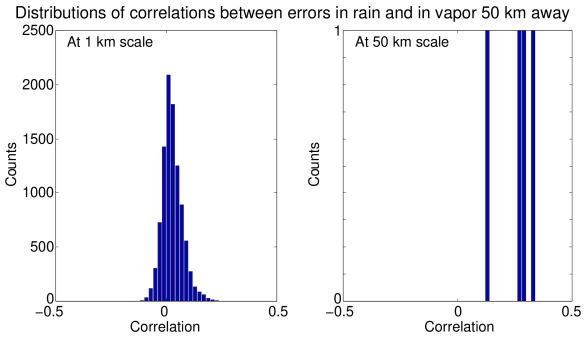

FIG. 15. Distribution of correlations at 1-km altitude between errors in rainfall and in vapor amounts 50 km to the east of the rain for the ensemble generated by Jacques and Zawadzki (2015). Correlations of 1×1-km scale values are shown on the left, while those at 50×50-km scale are shown on the right. As state variables and observations are smoothed, error correlation increases.

larger-scale patterns with more linear errors can be revealed. Multi-scale approaches vary from smoothing background error covariances (e.g., Miyoshi and Kondo 2013; Caron and Buehner 2018) to smoothing observations (Sodhi and Fabry 2020). By effectively smoothing error patterns, both approaches allow the use of much larger localization windows that increase the projection of information. Here too, other ideas would be welcome.

6.2 *Diversifying Approaches*

The ideas proposed above can be implemented within traditional adjustment-based assimilation systems. But to better tackle the challenges posed by both limited information propagation and the large background errors, radically different approaches may have to be considered. We believe for example that the time evolution and the spatial structure of echo patterns are not well exploited in convective-scale assimilation, yet they could provide valuable information. When one only has synoptic point data, all one can rely on is error covariance. But when data are available everywhere and all the time, derived quantities can also be exploited such as the position and strength of key features (reflectivity cores, low-level radial divergence, tropospheric-wide velocity couplets, etc.). In parallel, we should also better exploit additional radar and remote-sensed measurements to help better constrain storms as well as their environments. We believe that traditionally-assimilated radar data do not provide enough variety of constraints to effectively reduce background errors.

Perhaps should we also embrace the information provided by instability releases. The atmosphere is unpredictable because of instabilities, and precipitation generally occurs in the middle or final stages of the release of atmospheric instabilities. Paradoxically, we rarely have a better constraint on the state of the atmosphere than at the instant an instability is being triggered, as conditions for their triggering are often very specific. Therefore, the instability triggering event itself is a powerful piece of information if it could be intelligently taken advantage of.

Then, to deal with large errors at the assimilation level, more heuristic methods may have to be explored. For example, Pérez Hortal et al. (2019) constructs analyses by selecting, for each vertical column in the model, the ensemble member whose areal precipitation is locally closest to the observed values. Variations on particle filters may also be possible, among others, to eliminate members whose errors become so large that their information becomes misleading.

There is no magic bullet. Each success will only bring us a bit closer to victory. But a wise colleague of ours keeps repeating "I'll take a few percent of something over 100% of nothing any day" (G. McCourt, 2010-2019, personal communication). It is only via considerable efforts that radar data assimilation can considerably help improving convective-scale NWP.

6.3 *Acknowledgements*

Heartfelt thanks to Isztar Zawadzki who encouraged the first author to write this manuscript. The authors acknowledge the help of Andrés Pérez Hortal for the Introduction and of Dominik Jacques and Jagdeep Singh Sodhi for their ensemble forecast data. This project was undertaken with the financial support of the Government of Canada provided through the Department of the Environment through Grants and Contributions GCXE16S123. This work is also partly based on TIGGE data. TIGGE (The Interactive Grand Global Ensemble, Bougeault et al. 2010) is an initiative of the World Weather Research Programme (WWRP).

7. REFERENCES

Aksoy, A., D.C. Dowell, and C. Snyder, 2010: A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part II: Short-range ensemble forecasts. Mon. *Wea. Rev.*, **138**, 1273–1292, https://doi.org/10.1175/2009MWR3086.1.

Anderson, J., T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Avellano, 2009: The Data Assimilation Research Testbed: A community facility. *Bull. Amer. Meteor. Soc.,* **90**, 1283–1296, https://doi.org/10.1175/2009BAMS2618.1.

Asch, M., M. Bocquet, and M. Nodet, 2016: *Data Assimilation – Methods, Algorithms, and Applications*. Society for Industrial and Applied Mathematics, Philadelphia, 295 pp.

Augros, C., O. Caumont, V. Ducrocq, and N. Gaussiat, 2018: Assimilation of radar dual-polarization observations in the AROME model. *Quart. J. Royal Meteor. Soc.*, **144**, 1352–1368, https://doi.org/10.1002/qj.3269.

Benjamin, S. G., and Coauthors, 2016: A North American hourly assimilation and model forecast cycle: The Rapid Refresh. *Mon. Wea. Rev.*, **144**, 1669–1694, https://doi.org/10.1175/MWR-D-15-0242.1.

Bocquet, M., 2019: *Introduction to the Principles and Methods of Data Assimilation in the Geosciences*. Lectures notes of the data assimilation course, École des Ponts ParisTech, http://cerea.enpc.fr/HomePages/bocquet/teaching/assim-mb-en.pdf (accessed 15 August 2019).

Bougeault, P., and Coauthors, 2010: The THORPEX Interactive Grand Global Ensemble. *Bull. Amer. Meteor. Soc.*, **91**, 1059–1072, https://doi.org/10.1175/2010BAMS2853.1.

Crook, N. A., 1996: Sensitivity of moist convection forced by boundary layer processes to low-level thermodynamic fields. *Mon. Wea. Rev.*, **124**, 1767–1785, https://doi.org/10.1175/1520-0493(1996)124<1767:SOMCFB>2.0.CO;2.

De Lozar, A., A. Seifert and U. Blahak, 2018: Direct assimilation of 3D radar reflectivities with an ensemble-based data assimilation system. *Sixth International Symposium on Data Assimilation*, Munich, Germany, 5–9 March 2018, presentation 2.2, https://isda2018.wavestoweather.de/program/radar-d-a-and-nowcasting/o2_2_delozar.pdf.

Dowell, D.C., L.J. Wicker, and C. Snyder, 2011: Ensemble Kalman Filter assimilation of radar observations of the 8 May 2003 Oklahoma City supercell: Influences of reflectivity observations on storm-scale analyses. *Mon. Wea. Rev.*, **139**, 272–294, https://doi.org/10.1175/2010MWR3438.1.

Fabry, F., 2010: Radial velocity measurement simulations: Common errors, approximations, or omissions and their impact on estimation accuracy. *Proc. Sixth European Conf. on Radar in Meteorology and Hydrology*, Sibiu, Romania, ERAD, 17.2, http://www.erad2010.com/pdf/oral/thursday/nwp1/02_ERAD2010_0154.pdf.

Fabry, F., 2015: *Radar Meteorology – Principles and Practice*. Cambridge University Press, Cambridge, 272 pp.

Garand, L., 2003: Toward an integrated land–ocean surface skin temperature analysis from the variational assimilation of infrared radiances. *J. Appl. Meteor.*, **42**, 570–583, https://doi.org/10.1175/1520-0450(2003)042<0570:TAILSS>2.0.CO;2.

Gastaldo, T., V. Poli, C. Marsigli, P. P. Alberoni, and T. Paccagnella, 2018: Data assimilation of radar reflectivity volumes in a LETKF scheme. *Nonlinear Process Geophys.*, **25**, 747–764, https://doi.org/10.5194/npg-25-747-2018.

Gustafsson, N., and co-authors, 2018: Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres. *Quart. J. Roy. Meteorol. Soc.*, **144**, 1218–1256. https://doi.org/10.1002/qj.3179.

Jacques, D., and I. Zawadzki, 2015: The impacts of representing the correlation of errors in radar data assimilation. Part II: Model output as background estimates. *Mon. Wea. Rev.*, **143**, 2637–2656.

Jacques, D., D. Michelson, J.-F. Caron, and Luc Fillion, 2018: Latent heat nudging in the Canadian Regional Deterministic Prediction System. *Mon. Wea. Rev.*, **146**, 3995–4014, https://doi.org/10.1175/MWR-D-18-0118.1.

Jones, T.A., D. Stensrud, L. Wicker, P. Minnis, and R. Palikonda, 2015: Simultaneous radar and satellite data storm-scale assimilation using an Ensemble Kalman Filter approach for 24 May 2011. *Mon. Wea. Rev.*, **143**, 165–194, https://doi.org/10.1175/MWR-D-14-00180.1.

Keeler, R. J., and S. M. Ellis, 2000: Observational error covariance matrices for radar data assimilation. *Phys. Chem. Earth, Part B: Hydrol. Oceans Atmos.*, **25**, 1277–1280, https://doi.org/10.1016/S1464-1909(00)00193-3.

Keeler, R. J., and R. E. Passarelli, 1990: Signal processing for atmospheric radars. *Radar in Meteorology*, D. Atlas (Ed.), Amer. Meteor. Soc., Boston, 199–229.

Knight, C.A., W.D. Hall, and P.M. Roskowski, 1983: Visual Cloud Histories Related to First Radar Echo Formation in Northeast Colorado Cumulus. *J. Climate Appl. Meteor.*, **22**, 1022–1040. https://doi.org/10.1175/1520-0450%281983%29022%3C1022%3AVCHRTF%3E2.0.CO%3B2Kwon, I., S. English, W. Bell, R. Potthast, A. Collard, and B. Ruston, 2018: Assessment of progress and status of data assimilation in numerical weather prediction. *Bull. Amer. Meteor. Soc.*, **99**, ES75–ES79, https://doi.org/10.1175/BAMS-D-17-0266.1.

Lee, G.W., A.W. Seed, and I. Zawadzki, 2007: Modeling the variability of drop size distributions in space and time. *J. Appl. Meteor. Climatol.*, **46**, 742–756, https://doi.org/10.1175/JAM2505.1.

Li, X., J. R. Mecikalski, D. Posselt, 2017 : An ice-phase microphysics forward model and preliminary results of polarimetric radar data assimilation. Mon. Wea. Rev., **145**, 683–708, https://doi.org/10.1175/mwr-d-16-0035.1.

Mandapaka, P. V., U. Germann, L. Panziera, A. Hering, 2014: Can Lagrangian extrapolation of radar fields be used for precipitation nowcasting over complex alpine orography? *Wea. Forecasting*, **27**, 28–49, https://doi.org/10.1175/WAF-D-11-00050.1.

Miyoshi, T., and K. Kondo, 2013: A multi-scale localization approach to an Ensemble Kalman filter. *Scientific Online Letters on the Atmosphere*, **9**, 170−173, https://doi.org/10.2151/sola.2013-038.

Murakami, M., 1990: Numerical modeling of dynamical and microphysical evolution of an isolated convective cloud. *J. Meteorol. Soc. Japan*, **68**, 107–128, https://doi.org/10.2151/jmsj1965.68.2_107.

Pérez Hortal, A.A., I. Zawadzki, and M.K. Yau, 2019: A heuristic approach for precipitation data assimilation. Characterization using OSSEs. *Mon. Wea. Rev.*, in press, https://doi.org/10.1175/MWR-D-19-0034.1.

Posselt, D., X. Li, S. Tushaus, and J. Mecikalski, 2015: Assimilation of dual-polarization radar observations in mixed- and ice-phase regions of convective storms: Information content and forward model errors. *Mon. Wea. Rev.*, **143**, 2611–2636, https://doi.org/10.1175/MWR-D-14-00347.1.

Sodhi, J.S., and F. Fabry, 2020: Multiscale assimilation of radar reflectivity. 24th Conference on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface (IOAS-AOLS), Boston MA, Amer. Meteor. Soc.

Stensrud, D.J., M. Xue, L.J. Wicker, K.E. Kelleher, M.P. Foster, J.T. Schaefer, R.S. Schneider, S.G. Benjamin, S.S. Weygandt, J.T. Ferree, and J.P. Tuell, 2009: Convective-scale warn-on-forecast system. *Bull. Amer. Meteor. Soc.*, **90**, 1487–1500, https://doi.org/10.1175/2009BAMS2795.1.

Stensrud, D.J., and Coauthors, 2013: Progress and challenges with Warn-on-Forecast. *Atmos. Res.*, **123**, 2–16, https://doi.org/10.1016/j.atmosres.2012.04.004.

Stratman, D.R., C.K. Potvin, L.J. Wicker, 2018: Correcting storm displacement errors in ensembles using the feature alignment technique (FAT). *Mon. Weather Rev.*, **146**, 2125–2145, doi: 10.1175/MWR-D-17-0357.1.

Sun, J., and Coauthors, 2014: Use of NWP for nowcasting convective precipitation: Recent progress and challenges. *Bull. Amer. Meteor. Soc.*, **95**, 409–426, https://doi.org/10.1175/BAMS-D-11-00263.1.

Supinie, T.A., N. Yussouf, Y. Jung, M. Xue, J. Cheng, and S. Wang, 2017: Comparison of the analyses and forecasts of a tornadic supercell storm from assimilating phased-array radar and WSR-88D observations. *Wea. Forecasting*, **32**, 1379–1401, https://doi.org/10.1175/WAF-D-16-0159.1.

Surcel, M., I. Zawadzki, and M. K. Yau, 2015: A study on the scale dependence of the predictability of precipitation patterns. *J. Atmos. Sci.*, **72**, 216–235, https://doi.org/10.1175/JAS-D-14-0071.1.

Tong, M., and M. Xue, 2005: Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS experiments. *Mon. Weather Rev.*, **133**, 1789–1807, https://doi.org/10.1175/MWR2898.1.

Wattrelot, E., O. Caumont, and J. Mahfouf, 2014: Operational implementation of the 1D+3D-Var assimilation method of radar reflectivity data in the AROME model. *Mon. Wea. Rev.*, **142**, 1852–1873, https://doi.org/10.1175/MWR-D-13-00230.1.

Xue, M., and Coauthors, 2008: CAPS realtime storm-scale ensemble and high-resolution forecasts as part of the NOAA Hazardous Weather Testbed 2008 Spring Experiment. Preprints, 24th Conf. on Severe Local Storms, Savannah, GA, Amer. Meteor. Soc., 12.2. [Available online at https://ams.confex.com/ams/24SLS/techprogram/paper_142036.htm.]

Yano, J.-I., and Coauthors, 2018: Scientific challenges of convective-scale numerical weather prediction. *Bull. Amer. Meteor. Soc.*, **99**, 699–710, https://doi.org/10.1175/BAMS-D-17-0125.1.

Zrnic, D. S., 1979: Estimation of spectral moments for weather echoes. *IEEE Transactions on Geoscience Electronics*, **17**, 113–128, https://doi.org/10.1109/TGE.1979.294638.