Developing a Convective-Scale EnKF Data Assimilation System for the Canadian MEOPAR Project

DOMINIK JACQUES, WEIGUANG CHANG, SEUNG-JONG BAEK, THOMAS MILEWSKI, AND LUC FILLION

Data Assimilation and Satellite Meteorology Research Section, Environment and Climate Change Canada, Dorval, Québec, Canada

KAO-SHEN CHUNG
National Central University, Jhongli, Taiwan

HAROLD RITCHIE
Environmental Numerical Prediction Research Section, Environment and Climate Change Canada, Dartmouth, Nova Scotia, Canada

(Manuscript received 8 April 2016, in final form 25 November 2016)

ABSTRACT

This study discusses the construction of a high-resolution ensemble Kalman filter system (the HREnKF) developed for the Marine Environmental Observation Prediction and Response (MEOPAR) network. The HREnKF runs at a horizontal resolution of 2.5 km and is intended to provide forecasts at lead times up to 12 h. This system was adapted from the global EnKF system in operation at Environment and Climate Change Canada. As a first development step, only the most necessary modifications have been implemented. The changes include an hourly cycling frequency, smaller localization radii, and the explicit representation of microphysical processes. To assess its performance and orient future developments, the HREnKF was continuously cycled for a period of 12 days. Verification against surface observations reveals that the skill of the forecasts initialized from the HREnKF is comparable to that of control forecasts also integrated at a resolution of 2.5 km. At 2.5 km, correlation lengths are smaller than those found at 15 and 50 km. These short correlation lengths demand a high observational density, which is not available over the west coast domain where the HREnKF was tested. The spatial and temporal variability of the correlations is also assessed for the HREnKF system. It is found that correlation patterns are complex and do not generally decrease monotonically away from the reference point around which they are estimated. This result is important as it indicates that separation distance may not be the ideal parameter to use as a basis for localization strategies.

1. Introduction

This study reports on the construction of an hourly cycled high-resolution EnKF system (the HREnKF) developed for the Marine Environmental Observation Prediction and Response (MEOPAR) network. MEOPAR’s goal is to improve Canada’s ability to observe, predict, and respond to marine hazards.1 The HREnKF has the potential to contribute toward the achievement of this goal by improving the quality of short-term forecasts over maritime regions. In the long term, it is envisioned that the HREnKF would be deployed in cases of environmental disasters, rescue operations, and other emergencies.

Rapidly cycled assimilation systems require sufficient data coverage, adequate covariance models, and “balanced” 1-h forecasts without significant mass–momentum adjustments. The two latter requirements are at least partially addressed by performing the assimilation within the EnKF framework (Evensen 1994; Houtekamer and Mitchell 1998). For this reason, EnKF systems are currently the object of active research. At the convective scales, recent EnKF studies...
have covered topics ranging from convective initiation (Lange and Craig 2014; Yussouf et al. 2015), precipitation forecasts and/or radar data assimilation (Schwartz et al. 2014; Sobash and Stensrud 2015; Snook et al. 2015), the assimilation of surface observations (Ancell et al. 2011, 2014; Ha and Snyder 2014; Sobash and Wicker 2015), model biases (Romine et al. 2013), and the representation of model errors (Romine et al. 2014; Harnisch and Keil 2015). All these studies demonstrate the potential of an EnKF system to yield improvements in short-term forecasts.

Based on these results, a real-time, hourly cycled EnKF system with convective-scale resolution has been implemented at NCAR (Schwartz et al. 2015). Similar systems are currently being investigated at the Met Office (Flowerdew 2015) and at the German meteorological service (Bick et al. 2016).

At Environment and Climate Change Canada (ECCC), a continuously cycled global EnKF system (the GEnKF) has been in operational use since 2005 (Houtekamer and Mitchell 2005). In November 2014, the GEnKF was upgraded to run at a horizontal resolution of approximately 50 km. A limited-area version of the GEnKF has also been used for various research projects since 2012 (Baek et al. 2012). This “regional” EnKF system (the REnKF) operates over a domain encompassing most of North America at a resolution of 15 km. The present study discusses the construction of a third EnKF system, the HREnKF, which is adapted from the GEnKF and is intended for use at the convective scales. As a first development step, only the most necessary modifications to the GEnKF have been implemented. Hourly cycling, explicit microphysics, and smaller localization radii are among the changes.

An early version of the HREnKF has been used in pilot experiments involving the assimilation of radar data (Chung et al. 2013; Chang et al. 2014). The assimilation of radar data is not attempted in the present study. Instead, the HREnKF is considered as an addition to the two other EnKF systems currently in use. The HREnKF, REnKF, and GEnKF are run concurrently and nested within one another. The principal goal pursued here is to assess whether the forecasts initialized from the HREnKF could improve over those initialized from the other systems. This comparison will allow us to assess if the modifications brought to the GEnKF were sufficient to obtain better forecast skill at the convective scales.

A priori, some improvements can reasonably be expected from the HREnKF. Running at higher resolution, this system is expected to provide better covariance estimates, which should result in better analyses. Also, the hourly update frequency of the HREnKF could help capture the rapidly evolving, small-scale atmospheric features that cannot be resolved at lower spatial and temporal resolutions. In the past, increasing the resolution of the GEnKF has resulted in better forecast performance (Houtekamer et al. 2014).

Notwithstanding the potential advantages of the HREnKF, demonstrating the improvements attributable to this system is expected to be difficult. The HREnKF is tested on a domain that encompasses the Strait of Georgia, which is located north of Vancouver Island. Approximately one-third of this domain is over the Coast Mountains; the remaining two-thirds are over the Pacific Ocean. Outside of the coastline and a few valleys, the observation coverage is sparse. This could lead to situations where the small-scale features that the HREnKF is able to resolve would be poorly sampled by the observation network.

For some quantities, such as the ensemble spread, it is difficult to predict whether the HREnKF will fare better than systems running at lower resolutions. In EnKF studies, the lack of variance, or “spread,” in trial ensembles is the issue that is most often mentioned (see all of the above EnKF references). Brousseau et al. (2011) show some evidence suggesting that greater spread can be achieved by increasing the resolution at which a trial ensemble is generated. The extent at which this phenomenon can be beneficial to the HREnKF is unknown.

The significant topography of the HREnKF domain is another factor for which the impact of increasing the model resolution is unclear. The verification of atmospheric forecasts over mountainous regions is notoriously difficult because of large representativity errors. However, Ancell et al. (2011) suggest that increasing the resolution within an EnKF system should be most beneficial over complex terrain because of improved estimates of the flow-dependent covariances. Similarly, a review by Mass et al. (2002) mentions many studies showing the benefits of increased resolution in regions of complex topography.

In some EnKF studies, resolution increases did not yield the expected improvements. For example, Lange and Craig (2014) found that the faster error growth associated with high-resolution forecasts rapidly nullified the improvements that could be obtained in analyses. Perhaps more worrying are the results of Ancell et al. (2014), which showed that refining the horizontal resolution from 12 to 4 km caused the analyses from an EnKF system to become inferior to those of a 4DVAR system using predefined error covariances.

There is definitely a need for a better understanding of the conditions under which high-resolution assimilation systems are expected to fail or succeed. This study

2 A busy waterway that is a priority for the MEOPAR project.
attempts to answer this question in two different ways. First, the performance of the HREnKF's forecasts is compared to downscaled forecasts initialized with analyses obtained at lower resolutions. Second, the factors that are detrimental to the performance of the HREnKF system are identified. Some of the usual problems with EnKF systems are discussed along with two notable additions: the long spinup times for the generation of precipitation and the short correlation lengths in high-resolution ensembles.

This article is structured as follows. The experimental setup is presented in section 2, followed by the results of forecast verifications in section 3. Section 4 discusses the issues that were found to be important in the HREnKF. Finally, suggestions for improving the HREnKF and concluding remarks are included in section 5.

2. Experiment setup

a. Assimilation systems

Figure 1 shows the domains of the three EnKF systems that were run for this study. In light brown is the GEnKF, which has 256 members and covers the entire globe at a horizontal resolution of approximately 50 km. In green is the REnKF, which is also composed of 256 members and runs at a resolution of 15 km. The HREnKF system is shown in red and runs with 96 members at a resolution of 2.5 km. The HREnKF receives its boundary conditions from the REnKF, which, in turn, receives its boundary conditions from the GEnKF. All three systems are based on the GEnKF (described in Houtekamer and Mitchell 1998, 2005; Houtekamer et al. 2014) and use the ISBA land surface scheme (Noilhan and Planton 1989; Bélair et al. 2003).

For the three systems, the control variables are the $U$ and $V$ components of the wind, temperature, water vapor and surface pressure. The variance of model errors (as estimated by the ensemble spread) is not inflated in any way. In the GEnKF and the REnKF, all the variables that are not included in the control vector (dynamical and physical tendencies, cloud water, etc.) are lost during the analysis step of the assimilation cycle. All these variables have to be regenerated by the model after each analysis step. This approach is unsuitable to the HREnKF, which has explicit microphysics and assimilates observations every hour. For more continuity in this system, all the microphysical variables (cloud water, ice, rain, etc.) are “conserved” during the analysis step. This is an intermediate solution between losing all information on microphysics (as in the GEnKF and the REnKF) and including these variables in the control vector.

With the exception of horizontal resolution, the REnKF is nearly identical to the GEnKF. One notable difference between the two systems is the use of different parameterizations (and parameterization parameters) in the GEnKF [as per Table 5 in Houtekamer et al. (2014)] but not in the REnKF. Both systems share the following characteristics:

- Analyses are performed every 6 h.
- Deep convection is parameterized. In the GEnKF, convection is represented using either the Kuo (Geleyn 1985) or the Kain–Fritsch (Kain and Fritsch 1993) parameterizations. In the REnKF, only the Kain–Fritsch scheme is used.
All variables not being analyzed have to be regenerated by the model after each analysis period. This includes microphysical variables such as clouds and rain as well as “internal” model variables such as the dynamical and physical tendencies.

The localization of covariances is performed using the fifth-order piecewise rational function given by Eq. 4.10 in Gaspari and Cohn (1999). The cutoff distances at which covariances are forced to zero are on the order of 2000 km, as specified in Table 3 of Houtekamer et al. (2014). In the vertical direction, covariances are forced to zero in two units of ln p, as explained in Houtekamer et al. (2005).

Model errors are represented by the addition of correlated noise to the analyzed variables for each member of the analysis ensemble (Mitchell et al. 2002).

A digital filter (Fillion et al. 1995) with a 3-h window is applied at the moment of initializing the forecasts after the analysis period. This filter is applied in the cycled analysis system and the forecast ensemble.

In addition to being run at higher resolution, the HREnKF differs from the other systems in the following ways:

- Analyses are performed every hour.
- No deep convection parameterization is used. Microphysical processes are represented using the Milbrandt and Yau double-moment parameterization (Milbrandt and Yau 2005).
- All microphysical species are conserved during the analysis process such that the model does not have to regenerate precipitation every hour.
- The cutoff distance at which covariances are forced to zero is 60 km below 100 hPa and 100 km above. In a number of tuning experiments, these cutoff distances have been determined as the largest that could accommodate the short correlation lengths that are found in the HREnKF over topography (see section 4b). In these regions, the use of a larger localization distance resulted in analysis increments that were significantly affected by the sampling noise of covariance estimates.
- Model errors are not represented and the same physical parameterizations are used for all members of the ensemble.
- Forecasts are integrated directly from the analyses. No use is made of the digital filter.

The GEnKF, REnKF, and HREnKF were run continuously for a test period of 13 days during the month of February 2011.³ Four times a day, 12-h forecasts were generated from the HREnKF analyses. To reduce computational costs, forecasts were integrated from 40 (randomly chosen) members out of the 96 available in the analysis ensemble. When such subsampling is performed in the GEnKF, the average of the smaller forecast ensemble is adjusted to the average of the larger analysis ensemble. Experiments with such “recentering,” detailed in the appendix, showed that this process could trigger spurious convection in forecasts at a resolution of 2.5 km. Consequently, recentering was not applied to the HREnKF’s forecast ensemble.

Previous studies have demonstrated that higher resolution is especially beneficial over mountainous terrain (see a review by Mass et al. 2002). To evaluate the performance of the HREnKF, we wished to separate the impact of assimilating observations hourly from that of “simply” increasing the model resolution. To this end, the forecasts from the HREnKF were compared to downscaled forecasts from the REnKF. The term downscaled here refers to the fact that these forecasts were initialized with analyses from the REnKF (at a resolution of 15 km) but were integrated at a resolution of 2.5 km on the HREnKF’s domain. Because both forecast ensembles were integrated at the same resolution and used the same boundary conditions, the difference between them can only be attributed to the hourly assimilation of observations in the HREnKF.

b. Observations

With the exception of satellite radiances, the HREnKF was provided with the same observations as the REnKF and the GEnKF. We wished to assess the general behavior of the new system before more observations were added. Satellite radiances were not assimilated in the HREnKF because the 2.5-km grid spacing of this system is smaller than the footprint associated with such measurements (typically ≥ 15 km). The assimilation of satellite radiance would have required modifications to the forward operator. Such undertaking was left as future work.

All the observations assimilated by the HREnKF every hour of a typical day are depicted in Fig. 2. Surface observations come from the surface synoptic observation (SYNOP) network (in blue), ships (in purple), and drifting buoys (also in purple). Aircraft measurements (in green) provide most of the upper-air observations in the domain. During the course of one day, between one and five GPS radio occultations (red star at 1800 UTC) are assimilated. Over the ocean, satellite winds estimated from scatterometer (in pink) or cloud-tracking algorithms (in light orange) provide the observations available at the highest density. The information from

³ This month was chosen because it is part of a standard test period for model evaluation at ECCC.
- **SYNOP**
- Ships + Drifting buoys
- Airplanes

- **GPS**
- Radiosondes
- Scatterometer
- Satellite winds

FIG. 2. Observations assimilated on 10 Feb 2011.
two radiosondes (orange stars) is assimilated at 0000 and 1200 UTC.

At the moment of assimilation, a quality check ensured that radiosonde and aircraft observations were at least 50 hPa above the surface. Also, observations were rejected if their value was found to be farther away than \(2\left(\sigma_b^2 + \sigma_o^2\right)^{1/2}\) from the trial ensemble mean at the observation’s location. The symbols \(\sigma_b\) and \(\sigma_o\) represent the standard deviation of the background and observation errors, respectively. This quality check is unique to the HREnKF and resulted in the rejection of approximately 10% of the observations.

The prescribed standard deviations of observation errors used in this study are the same as those used in the GEnKF and are given in Tables 1 and 2. Possible correlations between observation errors were neglected.

Radar measurements and surface observations in aviation routine weather report (METAR) format are only used for verifications purposes (see sections 3c,d). The assimilation of METAR observations is currently the object of a separate study (Chang et al. 2016, manuscript submitted to Atmos.–Ocean, hereafter CJFSJ) by our research group at ECCC in Dorval, Québec, Canada.

3. Results

a. Surface pressure tendencies

The domain-averaged absolute value of surface pressure tendencies (SPTs) is commonly examined to assess the behavior of continuously cycled assimilation systems (Huang and Lynch 1993; Benjamin et al. 2004; Bick et al. 2016, among others). Some amount of shock, associated with the generation of spurious gravity waves and a spike in SPT, is unavoidable after the analysis step. As the model is integrated, the noise is dissipated and the SPT goes down. It is desirable for the SPT to stabilize before the next analysis step is taken.

Figure 3 shows the SPT during the first 20 analysis cycles of the HREnKF. The SPT was estimated every minute (the model’s temporal resolution) from each one of the 96 members of the analysis ensemble. The thick red line indicates the ensemble average of the SPT. Color shadings extend from the 10th to the 90th percentile and are indicative of the variability of SPT within the analysis ensemble.

During the first hour, the model is continuously shocked as the topography’s resolution is continuously adjusted from 15 km (the resolution of the initial conditions) to 2.5 km. This period should be disregarded.

### Table 1. Prescribed standard deviation of upper-air observation errors as a function of altitude.

<table>
<thead>
<tr>
<th>Alt (hPa)</th>
<th>Radiosonde</th>
<th>AIREP</th>
<th>AMDAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alt (hPa)</td>
<td>T (°C)</td>
<td>T - T_d (°C)</td>
<td>U/V (m s^{-1})</td>
</tr>
<tr>
<td>3</td>
<td>5.0</td>
<td>4.6</td>
<td>4.5</td>
</tr>
<tr>
<td>5</td>
<td>3.6</td>
<td>4.6</td>
<td>5.5</td>
</tr>
<tr>
<td>7</td>
<td>2.5</td>
<td>4.6</td>
<td>4.5</td>
</tr>
<tr>
<td>10</td>
<td>2.1</td>
<td>4.6</td>
<td>3.5</td>
</tr>
<tr>
<td>20</td>
<td>1.7</td>
<td>4.6</td>
<td>3.0</td>
</tr>
<tr>
<td>30</td>
<td>1.9</td>
<td>4.6</td>
<td>2.8</td>
</tr>
<tr>
<td>50</td>
<td>1.9</td>
<td>4.6</td>
<td>2.9</td>
</tr>
<tr>
<td>100</td>
<td>1.4</td>
<td>4.6</td>
<td>3.0</td>
</tr>
<tr>
<td>150</td>
<td>1.3</td>
<td>4.6</td>
<td>3.0</td>
</tr>
<tr>
<td>200</td>
<td>1.1</td>
<td>4.6</td>
<td>3.0</td>
</tr>
<tr>
<td>250</td>
<td>0.9</td>
<td>3.4</td>
<td>2.6</td>
</tr>
<tr>
<td>300</td>
<td>0.7</td>
<td>3.4</td>
<td>2.5</td>
</tr>
<tr>
<td>400</td>
<td>0.7</td>
<td>3.6</td>
<td>2.4</td>
</tr>
<tr>
<td>500</td>
<td>0.8</td>
<td>3.9</td>
<td>2.3</td>
</tr>
<tr>
<td>700</td>
<td>0.9</td>
<td>3.6</td>
<td>2.2</td>
</tr>
<tr>
<td>850</td>
<td>1.0</td>
<td>3.1</td>
<td>2.1</td>
</tr>
<tr>
<td>925</td>
<td>1.3</td>
<td>2.3</td>
<td>2.1</td>
</tr>
<tr>
<td>1000</td>
<td>1.5</td>
<td>1.8</td>
<td>2.0</td>
</tr>
</tbody>
</table>

### Table 2. Prescribed standard deviation of surface observation errors.

<table>
<thead>
<tr>
<th>P (Pa)</th>
<th>T (°C)</th>
<th>T - T_d (°C)</th>
<th>U/V (m s^{-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scatterometer</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>SYNOP stations (manual and automatic)</td>
<td>75.0</td>
<td>2.0</td>
<td>3.0</td>
</tr>
<tr>
<td>SHIP manual</td>
<td>175.0</td>
<td>1.8</td>
<td>3.0</td>
</tr>
<tr>
<td>SHIP automatic</td>
<td>175.0</td>
<td>1.8</td>
<td>3.0</td>
</tr>
<tr>
<td>Drifting buoy</td>
<td>143.75</td>
<td>1.8</td>
<td>3.0</td>
</tr>
</tbody>
</table>
In general, the magnitudes of the SPT spike are larger at 0, 6, 12, and 18 h, when the REnKF (piloting the HREnKF) is updated. During this update, observations are assimilated and correlated noise (representing model errors) is added to the analyses of the REnKF. These modifications disrupt the continuity of the boundary conditions provided to the HREnKF and lead to the observed spike in SPT. This can be clearly observed in 2D maps of SPT (not shown here), where the imbalances originating at the boundaries have the largest magnitude. This problem could be alleviated by using shorter assimilation cycles in the REnKF. A possibly better solution would be to use an incremental analysis update (IAU) approach (Bloom et al. 1996; Lee et al. 2006; Buehner et al. 2015), which would distribute the discrepancy in time.

The shocks observed at intermediate times (when the REnKF is not updated) are caused by the hourly updates of the HREnKF. At these moments, animations of surface pressure tendencies (not shown) reveal fast-moving deep gravity waves radiating away from the locations where observations are assimilated. This indicates that some form of imbalance is introduced by the analysis process. Two plausible causes for these imbalances are identified. The first is the small localization radius used in the HREnKF. At the synoptic scales, localization has been shown (Mitchell et al. 2002; Greybush et al. 2011) to create imbalances with respect to geostrophy. It seems reasonable to suppose that localization would also be introducing imbalances at the convective scales. Second, imbalances may be produced when the high-amplitude small-scale features found in the different members of the trial ensemble become entirely decorrelated from one another. When this happens (as depicted in Chung et al. 2013; Jacques and Zawadzki 2015), correlation lengths become short, which makes it difficult to obtain spatially continuous analysis increments. In these cases, scale-dependent localization (Zhang et al. 2009; Buehner 2012; Miyoshi and Kondo 2013) could possibly mitigate the imbalances introduced at the analysis time.

Through direct examination, we could confirm that the fast-moving waves generated at the analysis time had no perceivable influence on the different atmospheric fields being traversed. When first examining Fig. 3, we were satisfied with the SPT stabilization that occurred between the analyses. This conclusion will be reexamined in section 4a.

b. Ensemble spread at different locations

The impact of model resolution on the ensemble spread is now investigated. To get some idea about the relative importance of this factor, the impact of model resolution will be compared to the spatial variability of the spread found in the HREnKF’s domain. To this end, Fig. 4, shows the ensemble spread for three variables at six locations in the HREnKF’s domain. The three variables are the water vapor mixing ratio ($Q$), the temperature ($T$), and the magnitude of the wind velocity ($W$). The six locations are labeled from a to f and are illustrated in Fig. 5. The spread found in the GENKF, REnKF, and HREnKF are shown in light brown, green, and red, respectively. The ensemble of trial fields valid at 0600 UTC 13 February 2013 is used for these computations. Data at a single time (as opposed to a temporal average) are considered here to obtain an exact representation of the ensemble spread that would be used for a given assimilation window.

Mainly, Fig. 4 shows that, for a given assimilation window, the location in the domain has a much greater influence on the ensemble spread than the model resolution. For water vapor, the three assimilation systems behave similarly over the ocean (Figs. 4Qa–c), but not over land (Figs. 4Qd–f). The same can be said about wind velocity (Figs. 4Wa–f) with larger differences between the spread of the three assimilation systems. For wind, the ensemble spread can be sharply reduced near the surface over the topography. In Fig. 4Wd for example, the spread in all three systems is on the order of 0.1 m s$^{-1}$ at the lowest prognostic model level (at an altitude of 40 m). Such a small spread would effectively prevent the assimilation of a surface observation at this location.

The small spread near the surface is mostly explained by the “roughness length” by which the wind velocity is parameterized in the surface layer. This quantity determines the height (in meters) at which the log wind profile goes to zero. For example, point d is on steep terrain (as shown in Fig. 11HRd) that is characterized
by a large roughness length (37 m). This greatly constrains the magnitude of the wind velocity and reduces the ensemble spread at the lowest dynamical level. Conversely, point e is in a valley that is characterized by a smaller roughness length (1.5 m), which allows for more spread.

For temperature (Figs. 4Ta–f), the ensemble spread never exceeds 1°C. Neither the location within the domain or the resolution of the trial fields has a significant influence on this value. Because the prescribed observation errors for temperature are mostly between 1.0°C and 2.0°C (Tables 1 and 2), the influence of temperature observations on analyses can only be small.

Brousseau et al. (2011) showed that, on average, higher model resolution was associated with larger background errors in the lower troposphere. Figure 6 is provided for comparison with this result. It shows the spread (by which the background errors are estimated in EnKF systems) for Q, T, and W averaged over the HREnKF’s domain at 0600 UTC 13 February 2011. For velocities below 600 hPa and water vapor between 600 and 900 hPa the largest spread is indeed obtained with the HREnKF. However, for water vapor below 900 hPa and temperature below 700 hPa the largest spread is obtained with the GEnKF.

The interpretation of Fig. 6 is complicated by the fact that the GEnKF uses perturbed physics while the other two systems do not. On the HREnKF’s domain, this feature is highly effective at increasing spread below 700 hPa. This result is in agreement with the globally averaged ensemble spread shown in Fig. 10 of Houtekamer et al. (2009).

The color shadings shown in Fig. 6 extend ± two standard deviations about the GEnKF ensemble average and indicate the spatial variability observed for this quantity. A similar amount of variability is observed for the two other systems and is not shown for clarity. This variability dwarfs any differences in spread caused by the different model resolutions. As in Fig. 4, the geographical location is identified as the most influential factor determining the ensemble spread for a given assimilation window.

c. Surface verification

Figure 7 shows the results of surface verifications for the HREnKF and the downscaled forecasts in red and
blue, respectively. Each of the two ensembles is composed of 40 members. For clarity, only the forecasts initialized at 0000 and 1200 UTC are shown in Fig. 7. Surface verification was performed every hour against approximately 50 METAR observations that were not assimilated. The spatial distribution of these METARs is similar to that of the SYNOP stations shown in Fig. 2.

The standard deviation and bias of errors are plotted in Fig. 7. These two metrics were computed independently for each member of each forecast ensemble. For clarity, the 40 score values obtained for each forecast ensemble at each lead time are represented by a solid line (indicating the average score) and color shadings (extending ± two standard deviations from the average).

When considering the variability observed within the different forecast ensembles, most verification scores are not significantly different for the HREnKF and the downscaled forecasts. The exceptions to this are the biases of the dewpoint temperature and temperature (Figs. 7a,b). These biases are generally smaller for the downscaled forecasts. The repeated shocks caused by the hourly analyses in the HREnKF system could be responsible for the poorer performance observed.

The clear diurnal cycle of the biases indicates that important processes influencing surface observations are misrepresented in the model. Significant gains in forecast skill could be obtained by addressing the issues causing these biases. Because of the large temporal variability of the biases observed here, simple bias removal methods [such as the one used by Ancell et al. (2011)] are not expected to yield major improvements to the quality of forecasts. However, the use of an improved surface parameterization could prove helpful and is currently being investigated by CJFSJ.

Irrespective of the variable examined, approximately 70% of the observations are outside the range spanned by the ensemble. Because of this, a more traditional ensemble scoring metric such as the continuous rank probability score (CRPS: Hersbach 2000 and herein references) did not bring additional information to the scores plotted in Fig. 7. In the computation of the CRPS, these “outliers” are heavily weighted. When the proportion of these outliers is significant (as in the present case), the CRPS behaves like the mean absolute error.

In forecast verifications, one generally expects to observe larger errors at longer lead times. In Fig. 7, the errors fluctuate with the passing weather but they are
only minimally influenced by the forecast lead time. This indicates that even at the analysis time, the HREnKF is not capable of bringing measurable improvements to the surface conditions.

To rule out the possibility that these results may be caused by filter divergence, the root-mean-square (RMS) error with respect to aircraft observations has been compared to the ensemble spread. This verification (not shown here) shows that the RMS errors with respect to aircraft observations behave comparably to those shown in Fig. 7. The errors fluctuate with the weather and show no sign of improvements as the HREnKF is continuously cycled. The ensemble spread also fluctuates with the weather, but we do not observe the simultaneous increase in analysis error and decrease in ensemble spread that would be indicative of filter divergence (see, e.g., Fitzgerald 1971; Mitchell et al. 2002).

Other issues affecting the performance of the HREnKF are described in section 4.

d. Qualitative comparison with precipitation

In this section, we briefly investigate the skill of the HREnKF’s analyses to capture instantaneous precipitation fields. Figures 8a–f provide a qualitative verification of precipitation in forecasts from the HREnKF.
system. The probability of precipitation estimated from the 96 members of the analysis ensemble (Figs. 8a–c) can be compared with the precipitation observed by radars (Figs. 8d–f) at 1100, 1700, and 2300 UTC 3 February 2011. Ten-minute accumulations were considered for estimating the probability of precipitation while radar measurements are instantaneous reflectivity composites. It is worth reminding the reader that precipitation is not a variable currently analyzed in the HREnKF.

At 1100 UTC, the HREnKF ensemble predicts an 80%–100% chance of precipitation over the topography in the northern half of the HREnKF’s domain. Radar observations are unavailable in this area, so it is impossible to verify whether this prediction was accurate. In the southern half of the domain, the ensemble predicts a 0% chance of precipitation but precipitation is observed by the radars. This situation means that the “truth” cannot be a realistic draw out of the probability density that is at the origin of the trial ensemble. In other words, the analysis ensemble does not “capture” the truth.

At 1700 and 2300 UTC, there is some overlap between the areas where precipitation is predicted by the ensemble and the observed precipitation. Nevertheless, it is difficult to attribute much skill to the HREnKF’s analyses in capturing the probability of occurrence of individual precipitation patterns. Several reasons can explain these poor results.

The first reason is the difficulty in predicting the short-lived small-scale precipitation features. In this respect, it is interesting to consider the results of Cookson-Hills et al. (2016) and work in preparation by P. Cookson-Hills et al. (2016, hereafter CH), who performed a formal verification of the precipitation forecasted by the REnKF, the downscaled REnKF, and the HREnKF. To match measurements from surface stations, 24-h accumulations were considered instead of instantaneous fields. At this lower temporal resolution, it is found that the forecasts from the HREnKF generally show significant skill [as measured by the fraction skill score (FSS)] for spatial scales greater than \(50–100\) km. If the negative precipitation bias depicted in Figs. 8a–f is also found in the 24-h accumulation for this day, the biases from different days compensate one another such that the average bias for all precipitation cases is small. This indicates that the physics parameterization of the 2.5-km model is not faulty in systematic ways.
A second reason explaining the difficulty in forecasting precipitation is that the cumulative improvement that can be obtained by cycling every hour is limited. Assuming a $50 \text{ km h}^{-1}$ displacement velocity for weather patterns (approximately correct for the period examined), an air parcel will remain inside the HREnKF’s domain for approximately 20 h. This means that this parcel can benefit from 20 assimilation cycles at most. Since few observations are found over the ocean, the parcel will benefit from assimilation for only a fraction of these 20 h. For this reason, analyses from the HREnKF are not better toward the end of the assimilation experiment than after 1 day of cycling.

A third reason for the poor precipitation forecasts is a problem with the development of precipitation in the analysis cycle. This problem is described in section 4a below.

### 4. Issues associated with high resolution

More development will be required for the HREnKF to become an effective forecasting tool. To orient future work on this system, we now identify its most problematic components.

#### a. Spinup time after analyses

In terms of the surface pressure tendencies (see section 3a), it appeared that the model shock could be mostly dissipated during the hour of model integration separating the analyses. This notion was further reinforced when we first examined the evolution of precipitation in the different members of the continuously cycled analysis ensemble. Figures 9a–d depict such an evolution for member 1 of the analysis ensemble at hourly intervals. When these figures (and the intermediate time steps) are animated, we can observe a number of relatively small patches of precipitation evolving and being advected. Careful examination does reveal hourly oscillations of the extent and intensity precipitation. However, these variations are very subtle and were initially thought to reflect the normal adjustment of the modeled fields to the newly assimilated data.

The rapid dissipation of gravity waves and the artifact-free evolution of precipitation has initially lead us to believe that the hourly assimilation periods (and the ensuing model shocks) had a negligible impact on the cycled analysis ensemble from the HREnKF. It is only after examining precipitation in the free forecasts that we realized that these analyses were, in fact, significantly affected by the hourly cycling. Consider Figs. 9e–h, which show accumulations of precipitation in a free forecast initialized from member 1 of the analysis ensemble at 0600 UTC. At 0700 UTC, the prevailing precipitation in the free forecast (Fig. 9e) and in the trial (Fig. 9a) are, by construction, identical. At longer lead times, however, the extent of the precipitation increases significantly in the free forecast (Figs. 9f–h) but not in the cycled system (Figs. 9b–d). This behavior is observed at all initialization times and for all members of the forecast ensemble.
In the literature, some authors (Romine et al. 2013, 2014; Schwartz et al. 2014) have diagnosed issues related to the spinup of precipitation by plotting time series of the areal coverage of precipitation exceeding various thresholds. Here, this metric is used to quantify the difference in precipitation extent found in the cycled ensemble and in the free forecasts. Figure 10 shows the areal coverage of precipitation for each member of the continuously cycled analysis ensemble (dark red) and the 12-h forecasts (pink) launched every 6 h. The areal coverage of precipitation was estimated here by counting the number of grid points exceeding a small threshold (0.04 mm in 10 min) in accumulations of the surface precipitation.

During the analysis cycle, the areal coverage of precipitation oscillates by a few percent every hour. To show these oscillations, one member of the analysis ensemble is highlighted in yellow. From this member, we can see that precipitation increases during the first 10 min after the analysis time (marked with thin vertical gray lines), diminishes for the next 20–30 min, and then increases again until the following analysis time. As mentioned above, a variation of a few percent in precipitation coverage was not thought to be indicative of critical issues with the HREnKF.

Contrary to this first assessment, serious problems are identified when we compare the areal coverage of precipitation in the cycled ensemble to that found in the free forecasts. In these forecasts, the areal coverage of precipitation always increases during the first 4–5 h of model integration. The areal coverage of precipitation for the HREnKF’s forecasts lags behind the bulk of the HREnKF forecasts by approximately 1 h. This means that the conservation of microphysical variables did, at best, speed up the spinup time by approximately 1 h.

When considering only Fig. 10, it is not possible to know which of the cycled or the free forecasts are the closest to the “truth.” However, the formal verification performed by (Cookson-Hills et al. 2016; CH) demonstrates that precipitation accumulations from the free forecasts have very small biases, on average, with respect to radars and surface station measurements. It is thus the cycled system that is in error. Experiments are currently being performed to determine the reason(s) and possible solution(s) to this problem.

In the literature, some authors (Romine et al. 2013, 2014; Schwartz et al. 2014) have diagnosed issues related to the spinup of precipitation by plotting time series of the areal coverage of precipitation exceeding various thresholds. Here, this metric is used to quantify the difference in precipitation extent found in the cycled ensemble and in the free forecasts. Figure 10 shows the areal coverage of precipitation for each member of the continuously cycled analysis ensemble (dark red) and the 12-h forecasts (pink) launched every 6 h. The areal coverage of precipitation was estimated here by counting the number of grid points exceeding a small threshold (0.04 mm in 10 min) in accumulations of the surface precipitation.

During the analysis cycle, the areal coverage of precipitation oscillates by a few percent every hour. To show these oscillations, one member of the analysis ensemble is highlighted in yellow. From this member, we can see that precipitation increases during the first 10 min after the analysis time (marked with thin vertical gray lines), diminishes for the next 20–30 min, and then increases again until the following analysis time. As mentioned above, a variation of a few percent in precipitation coverage was not thought to be indicative of critical issues with the HREnKF.

Contrary to this first assessment, serious problems are identified when we compare the areal coverage of precipitation in the cycled ensemble to that found in the free forecasts. In these forecasts, the areal coverage of precipitation always increases during the first 4–5 h of model integration. Ultimately, the precipitation coverage stabilizes at approximately 3 times the area found in the analysis cycle. This indicates that the spinup time necessary for the generation of precipitation is at least 4 h, a much longer time than the 1-h cycling frequency of the HREnKF.

The black lines in Fig. 10 indicate the areal coverage of precipitation in forecasts initialized every 6 h from the average of the analysis ensemble. These forecasts were started with no microphysical variables and also developed precipitation during the first 4–5 h of model integration. The areal coverage of precipitation for these forecasts lags behind the bulk of the HREnKF’s forecasts by approximately 1 h. This means that the conservation of microphysical variables did, at best, speed up the spinup time by approximately 1 h.

When considering only Fig. 10, it is not possible to know which of the cycled or the free forecasts are the closest to the “truth.” However, the formal verification performed by (Cookson-Hills et al. 2016; CH) demonstrates that precipitation accumulations from the free forecasts have very small biases, on average, with respect to radars and surface station measurements. It is thus the cycled system that is in error. Experiments are currently being performed to determine the reason(s) and possible solution(s) to this problem.

### b. The short extent of correlation lengths in high-resolution trials

The information from observations is propagated in space by use of the covariances between the different variables describing the atmosphere on the model grid. One advantage of the EnKF approach is the estimation of covariances that vary in space and time. Because the simulated atmospheric flow depends on the resolution at which the model is integrated, so do the covariances. In this section, we discuss the impact of model resolution on the estimated covariances and correlations.

Figure 11 shows vertical cross sections of autocorrelation for the \(U\) component of the wind as a function of resolution and location in the HREnKF’s domain. Each panel is designated by a combination of uppercase and lowercase letters. The uppercase letters HR, R, and G identify correlations estimated using the trial ensembles from the HREnKF, \(4\) REnKF, and GEnKF, respectively. The lowercase letters range from a to f and represent the geographical locations of the reference points at which

---

4 The correlations shown in this section were estimated from the HREnKF trial ensembles, which suffered from the spinup issue described in section 4a. To ensure that the correlations shown here were not affected by this problem, they were compared to correlations estimated from ensembles of 6- and 12-h forecasts valid at the same verification time. This comparison revealed only minor differences, confirming that the correlations being observed are not an artifact of the cycling procedure.
correlations are estimated. These locations, and the horizontal extent of the cross sections, are illustrated in Fig. 5.

No localization is applied to the correlations that are shown in Fig. 11. However, the short correlations, which are not significantly different from 0, appear in white. The threshold for the smallest significant correlations was determined using Eq. (12) of Houtekamer and Mitchell (1998):

\[
\frac{(\bar{\rho} - \rho)^2}{1 - \rho^2} \approx \frac{1}{N} (1 - \bar{\rho}^2)^2, \quad (1)
\]

which results from statistical theory and gives an approximation for the expected difference between the true correlation \( \rho \) of two normal distributions and the correlation \( \bar{\rho} \) estimated from \( N \) sample pairs. The overbar here denotes the expectation operator. If we let \( \sigma^2(\bar{\rho}) = (\rho - \bar{\rho})^2 \), then \( \bar{\rho} \pm 2\sigma(\bar{\rho}) \) gives a 95% confidence interval for the true correlation \( \rho \). With 256 sample pairs (\( N = 256 \)), correlations in the interval \([-0.12, 0.12]\) cannot preclude the existence of \( \rho = 0 \) at a 95% confidence level. In Fig. 11, correlations in this interval appear in white. With \( N = 96 \), correlations from the HREnKF are not significantly different from 0 in the interval \([-0.2, 0.2]\). Nevertheless, the same color palette is used for consistency.

The most striking feature in Fig. 11 is the short extent of the correlations observed for the HREnKF (Figs. 11HRa–f) as compared to the other systems. Similar reductions of correlation lengths with increasing resolutions have been shown by Brousseau et al. (2011) and Ancell et al. (2011). The short extent of the correlation reflects the more complex flow being resolved at higher resolutions. This argument lead Ancell et al. (2011) to suggest that increasing the resolution of an EnKF system may be most beneficial near complex terrain. On the other hand, as pointed out by Brousseau et al. (2011), sharp correlations demand that observations be available at a high density to ensure some continuity with the analysis increments.

In Fig. 11, the correlations estimated over land are significantly different from those estimated over the ocean. Over the ocean, the strongest correlations (>0.7) do not extend above the boundary layer at approximately 850 hPa. This result is comparable to the correlations shown by Hacker and Snyder (2005). It is also interesting to note how the strong correlations of the REnKF (Figs. 11Ra–c) extend farther away than those of the GEnKF (Figs. 11Ga–c). This is probably a consequence of the physics perturbations used in the GEnKF but not in the REnKF.

Over land, the relation between the resolution and correlation lengths is more straightforward. For the GEnKF (Figs. 11Gd–f) strong correlations are found at distances greater than 200 km and up to 500 hPa. For the REnKF (Figs. 11Rd–f), strong correlations (> 0.7 or ≤–0.7) do not extend as far as for the GEnKF. In the case of the HREnKF (Figs. 11HRd–f) strong correlations are only found near the reference points. By comparing the correlations estimated from these three systems, we learn that the assimilation of a given surface observation will have a significant influence over much greater distances at a resolution of 50 km than at a resolution of 2.5 km.

In Fig. 12 it is the correlation between the \( U \) component of the wind and the temperature that is shown. Over the ocean, the absolute magnitude of the
correlations does not generally exceed 0.4, indicating a relatively weak statistical relation between these two variables. The exception to this is point c in the GEnKF. Again, this result is comparable to those obtained by Hacker and Snyder (2005). Over land, the patterns of the $U$–$T$ correlations are substantially different for the three assimilation systems examined. In the REnKF and the GEnKF, significant correlations exist between the surface wind and temperature in the free atmosphere. Such interactions are lost in the HREnKF where cross correlations are weak everywhere.

When correlations are estimated from reference points at 700 hPa, in Fig. 13, the differences between the three assimilation systems become smaller. Over the ocean, the decrease in correlation lengths with increasing resolution is much more subtle. At locations a and b the correlations lengths estimated from the HREnKF are only slightly shorter than those estimated from the REnKF and the GEnKF. At location c, the correlation patterns are very similar for the three systems examined. As before, the impact of resolution is more pronounced over land, where the extent of correlation patterns is large in the GEnKF (Figs. 13Gd–f), small in the HREnKF (Figs. 13HRd–f), and somewhere in between in the REnKF (Figs. 13Rd–f).

If the correlations are conceptually easy to understand, it is the covariances that are used in producing the analysis increments. Since “raw” covariances are difficult to interpret, it is customary to examine the
spatial distribution of the increment obtained by using these covariances. In Fig. 14, the covariances have been normalized to represent the propagation of a 1 m s$^{-1}$ $U$-velocity increment (at the location of the reference point) on $U$ velocity in the rest of the domain. Figure 14 can be interpreted as the outcome of single-observation experiments where only the $U$ component of the wind was assimilated. Figure 14 differs from Fig. 13 in that the maximum increments are generally not found at the reference points. This is most notable for the GEnKF over land (Figs. 14Gd–f), where the maximum increments are observed as far as 200 km away from the reference points.

Figure 15 is provided for completeness and shows the correlation between the $U$ component of the wind at 700 hPa with temperature in the rest of the domain. Over the ocean, correlations are generally low with maximum absolute values of approximately 0.4. Interestingly, the strength of the correlations increases with model resolution at point c. Over land, stronger correlations are observed. At location f, the correlation patterns bear close connections to the respective topographies of each of the assimilation systems.

The magnitudes of the cross correlations shown in Fig. 15 are relatively small. Figure 16 shows the temperature adjustments resulting from a 1 m s$^{-1}$ increment for the $U$ component of the wind at the reference points. At the locations where the response is strong, temperature adjustments on the order of 0.2°C are observed. Given that wind increments of several meters per second are

**Fig. 14.** Influence on $U$ velocity in the rest of the domain of a 1 m s$^{-1}$ increment of $U$ velocity at the positions indicated by the black dots.

**Fig. 15.** As in Fig. 12, but for the correlation of the $U$ component of the wind (at 700 hPa) with temperature in the rest of the domain.
common, the temperature increments induced through cross covariances can be important.

As before, the maximum increments are often found some distance away from the reference points. This is especially evident in Fig. 16HRb, where the maximum increments are found between 100 and 200 km away from the reference point. Such “offset” correlations are in agreement with geostrophic balance. See, for example, Fig. 5.7 in Daley (1991), which shows the correlation between the $U$ component of the wind and geopotential estimated for geostrophic conditions.

With the current localization cutoff distance of 60 km in the HREnKF, such analysis increments would have been entirely filtered out. This figure thus demonstrates the need for spatially variable localization radii in the HREnKF system.

Many results have been presented in this section. The following conclusions were found to be generally applicable.

- In the boundary layer, increasing the model resolution from 50 to 2.5 km always resulted in a reduction of the observed correlation lengths. This proved true whether same-variable (Fig. 11) or cross-variable (Fig. 15) correlations were examined. Over the ocean, increasing the resolution from 50 to 15 km resulted in increased correlation lengths. We believe that this is caused by the perturbed physics in the GEnKF but not in the REnKF.
- In the free atmosphere, the relation between correlation lengths and model resolution is more tenuous, especially over the ocean (Figs. 13 and 15).
- Topography has a strong influence on the magnitudes and patterns of correlations (Figs. 11–16).
- Large-scale correlation patterns are often found in the low-resolution trials from the GEnKF. They are less evident in trials from the REnKF and altogether absent in trials from the HREnKF (Figs. 11 and 12, at locations d, e, and f).
- The locations of the maximum and minimum increments rarely coincide with the locations at which an observation is assimilated (Figs. 14 and 16).
- The correlation of the $V$ component of the wind, temperature, and humidity were also examined. Even though the correlation patterns varied, the general conclusions stated above remain applicable to these other variables.

c. The temporal variability of correlation patterns

The previous section showed the variability of the correlations lengths estimated at different locations in the domain. We now examine the temporal variability of correlations in the HREnKF system.

Figure 17 illustrates the autocorrelation of the $U$-velocity field estimated at 1200 UTC every day between 2 and 13 February 2014. Correlations are estimated with respect to the reference points a and e (see Fig. 5) at the lowest model level. At both locations, the horizontal and vertical extents of the correlation vary significantly from day to day. For example, consider the difference in the correlations estimated on 4–6 February at location a.

Interestingly, areas of negative correlations are sometimes seen far above the reference points. This is especially evident at location e on 3 and 6 February. The low values of correlations estimated ($\leq -0.4$) and the extent of the features suggest that these zones of negative correlations have a physical cause and are not simply a consequence of sampling errors. Topography
appears to play a role in their occurrence as they are observed more frequently at location e (on 3–6, 9–11, and 13 February) than at location a (weak signal is observed on 2 February). This observation is also true for the other reference points, which have been examined but are not shown here. The areas of negative correlations only weakly depend on the separation distance from the reference point and will be mostly filtered out by the current localization strategy. This result again demonstrates the need for an adaptive localization strategy that would not depend only on the separation distance.

The temporal variability of correlations has also been examined at other locations, for other atmospheric fields and different altitudes. This examination revealed that, the horizontal and vertical extent of the correlations patterns may vary by a factor of 10 from day to day.

5. Discussion and conclusions

A high-resolution EnKF system (the HREnKF) is currently being developed within the context of the Marine Environmental Observation Prediction and Response (MEOPAR) network. The objective of this network is to improve the accuracy of short-term atmospheric forecasts over the maritime regions of Canada. This study reports on the first experiments with the HREnKF on a domain located over the west coast of North America. From an assimilation perspective, this is a challenging area extending from the Coast Mountains out over the Pacific Ocean. Almost everywhere, the density of observations is low.

The basis for the construction of the HREnKF is the global ensemble Kalman filter system (the GEnKF) that has been run operationally since 2005 (Houtekamer and Mitchell 2005). The HREnKF differs from the GEnKF in that it runs at a horizontal resolution of 2.5 km, assimilates observations every hour, employs much shorter localization radii and uses explicit microphysics. The microphysical species are not part of the control vector and are not directly assimilated. However, their state is conserved during the analysis cycle.

The first goal of this study was to determine whether the above changes would be sufficient for the HREnKF to yield better short-term forecasts than those initialized from low-resolution analyses. Second, we wished to identify the components of the HREnKF that were most in need of improvement. To this end, an experiment was designed where three EnKF systems were run concurrently and nested within one another. The HREnKF received its boundary conditions from the regional EnKF system (the REnKF), which itself received its boundary conditions from the GEnKF. All systems were continuously cycled during a test period of 13 days during February of 2011. Four times a day, 12-h forecasts were

![Fig. 17. Autocorrelation of the $U$ velocity field at 1200 UTC every day between 2 and 13 Feb. The correlations are estimated against a reference point at the lowest model level at the locations a and e (see Fig. 5). The data source for this figure consists of the trial fields from the 96 members of the HREnKF's analysis ensemble. The color palette is the same as for the preceding correlation figures with the exception that the values in the interval $[-0.2, 0.2]$ (which, per Eq. (1), are not statistically different from 0) appear in white.](http://journals.ametsoc.org/mwr/article-pdf/145/4/1473/4786580/mwr-d-16-0135_1.pdf)
integrated from the HREnKF’s analyses. Both the hourly analysis ensemble and the 6-hourly forecast ensemble have been examined.

The HREnKF’s forecast ensembles were composed of 40 members whose initial conditions were chosen randomly from the 96 members available in the analysis cycle. An experiment was performed to determine if it was desirable to “recenter” the average of the small forecast ensemble to that of the larger analysis ensemble. Such recentering was found to trigger spurious convection in the forecasts and was not used thereafter.

When they were first considered, time series of domain-averaged surface pressure tendencies seemed to indicate that the GEM model was able to dissipate most of the “shock” caused by the analysis step of the assimilation cycle. Also, visual inspection of the precipitation patterns in the cycled analysis members showed no signs of pathological behavior. It is only by comparing the areal coverage of precipitation from the cycled analysis members and from the free forecasts that we could diagnose a serious problem with the development of precipitation in the cycled system.

We then examined the ensemble spread found in trial ensembles at different altitudes and geographical locations. By far, geographical location was found to be the most influential parameter affecting the ensemble spread. In the free atmosphere, the spread in wind velocity approximately matched the prescribed observation errors. However, the ensemble spread was found to be too small in the boundary layer. Over land, the ensemble spread was so small that it prevented the information from surface observations from influencing analyses in any significant way. Solutions for improving the near-surface ensemble spread in the HREnKF are explored in CJFSJ. The approach tested includes varying the surface parameterizations and using ensembles of surface analyses.

Surface verifications revealed that the forecasts initialized from the HREnKF’s analyses were not better than downscaled forecasts initialized from the REnKF analyses. For precipitation, the mismatch between the forecasted and observed quantities was such that quantitative verification was not attempted. These verification results established that high resolution and minimal changes to the GEnKF were not sufficient to improve short-term forecasts over the west coast of Canada.

Two issues specific to the HREnKF have then been identified: 1) the long spinup time of the precipitation mentioned above and 2) the short correlation lengths observed at higher resolutions. We believe that the long spinup times can be addressed by conserving more variables during the analysis step. Finding which variables are to be conserved to avoid perturbing the development of precipitation will be the subject of future work.

The “issue” of the short correlation lengths at higher model resolutions is more fundamental. To be clear, these correlations are not “wrong” in the sense that they do not indicate problems with the HREnKF’s ensemble. On the contrary, one explanation for the short correlation lengths at higher resolution is that the small-scale atmospheric flow is better resolved at higher resolutions. Within this context, the shorter correlation lengths are “good,” as they are likely a better approximation to the true propagation of information in the atmosphere. This interpretation is put forward by Brousseau et al. (2011) and Ancell et al. (2011), who discussed the reduction of correlation lengths at higher resolution. In the present study, this situation is particularly evident over topography.

Another explanation for the reduction of correlation lengths at higher resolution is that the high-energy/small-scale features have entirely diverged in the different members of a given trial ensemble. Chung et al. (2013) and Jacques and Zawadzki (2015) demonstrated this phenomenon for cases of summer convection. In the present study, we believe that decorrelated small scales could be responsible for the short correlation lengths observed in the boundary layer over the ocean. Because short correlation lengths are associated with convection (where the dissociation of small-scale features will occur at shorter forecast lead times), one expects problems with displacement errors (see, e.g., Fig. 8). One risks applying short correlations to observations that, in reality, are not influenced by convection and, conversely, long correlations to observations that are affected by convection.

Irrespective of the explanation for the short correlation distances found at high resolution, they ended up having a negative impact on the HREnKF. Most everywhere, the correlation lengths are shorter than the average separation distance between the available observations. Because of this, spatially continuous analysis increments could not be obtained. Discussing the temporal reduction of the assimilation window, Macpherson (1991) describes a worst-case scenario in which “the assimilation would consist of a sequence of multiple-iteration analyses with only a few data at most time steps.” To some extent, this situation was encountered here.

The comparison of correlations estimated at different model resolutions is a key contribution of this study since it provides a framework for estimating the maximum resolution of an assimilation system based on the available observational networks. In most locations, the correlation lengths of the HREnKF do not extend beyond a few tens of kilometers. For spatially continuous analysis increments, observations have to be available at this resolution. In this respect, observations from satellites and radars offer the best prospects for improving the performances of the HREnKF. Before this becomes a reality,
better quality controls will be required. The specification of observation error correlations may also become necessary for using these high-density observations. Over land, the assimilation of aviation routine weather reports (METARs) would add approximately 50 surface observations to each assimilation cycle. The assimilation of such observations is currently under investigation (CJFSJ).

The covariances shown in sections 4b, c are also useful for orienting the choice of future localization strategies. The examination of covariances has demonstrated that the distance over which meaningful signal could be obtained varied significantly depending on the location in the domain and the time at which correlations are estimated. This suggests that flow-dependent localization strategies could be beneficial to the HREnKF. It was also shown that the position of the maximum increment generally did not coincide with the position of the observations. This is especially true for cross-variable increments. This result is important, as it indicates that separation distance may not be the ideal parameter to use as a basis for localization strategies.

The present study provides a baseline against which future versions of the HREnKF will be compared. Increasing the ensemble spread near the surface, reducing model shocks after the analysis step, and implementing a better localization strategy were all identified as priorities. The need for more fundamental research on the estimation of correlations and localizations was also shown.

Acknowledgments. The authors are grateful to Dr. Peter Houtekamer and Dr. Herschel Mitchell, who both provided detailed internal reviews of an early version of this manuscript.

APPENDIX

The Negative Impact of Recentering at the Convective Scale

Hourly cycling and explicit microphysics make the HREnKF more expensive than either the GEnKF or the REnKF. Consequently, it was decided that 96 members would be integrated in the continuous analysis cycle rather than the 256 members of the GEnKF and the REnKF. For further cost reductions, the number of members in forecast ensembles was reduced to 41.

One member of the forecast ensemble was initialized from the ensemble mean. For the remaining 40 members, we chose randomly among the 96 members available in the analysis ensemble. In the GEnKF, such a reduction of ensemble size is accompanied by “recentering” the ensemble subset such that its mean becomes identical to that of the larger source ensemble. At the convective scales, it is not obvious that one should proceed to such recentering. Because of the nonlinear nature of high-resolution atmospheric flows, recentering will be a source of imbalances and model shock. Also,
microphysical variables have nonnegative distributions that do not readily lend themselves to recentering.

Unsure of whether recentering should be applied or not, forecasts were generated both with and without recentering. Figures A1a–h show the impact of recentering for one forecast initialized at 1200 UTC 2 February 2011. The top row (Figs. A1a–d) shows 1-min accumulations of precipitation with no recentering. The bottom row (Figs. A1e–h) shows the same metric with recentering applied. Recentering was applied to the $U$ and $V$ components of the wind, temperature, dewpoint temperature departures, and surface pressure. With the exception of the water vapor, all microphysical variables are identical in the initial conditions of the two forecasts. For this reason, Figs. A1a and A1e are nearly identical.

However, the two forecasts show some significant differences 1 h later. An organized line of precipitation (indicated by black arrows) is observed in the recentered forecasts (Figs. A1f,g) but not in its non-recentered counterpart (Figs. A1b,c). After close inspection, it was found that this line of convection was caused by the recentering process, which caused an inconsistency at the model boundary.

The impact of recentering can also be found within the domain. After 1 h of model integration, there is considerably more convective activity in the recentered member (circled area in Figs. A1f–h) than in the non-recentered member (Figs. A1b–d). This increased activity persists throughout the forecast.

These results indicate that recentering can be damaging at resolutions on the order of 1 km. Recentered forecasts were not used elsewhere in this study.

REFERENCES


Ha, S.-Y., and C. Snyder, 2014: Influence of surface observations in mesoscale data assimilation using an ensemble Kalman


