Can Observation Targeting Be a Wild Goose Chase? An Adjoint-Sensitivity Study of a U.S. East Coast Cyclone Forecast Bust

Daniel J. Lloveras$^{a,b}$, James D. Doyle$^c$, Dale R. Durran$^a$

$^a$Department of Atmospheric Sciences, University of Washington, Seattle, Washington, $^b$National Research Council, Monterey, California, $^c$Naval Research Laboratory, Monterey, California

Corresponding author: Daniel J. Lloveras, daniel.j.lloveras.ctr@us.navy.mil
ABSTRACT: We use a moist adjoint model to compute the initial-condition perturbations that minimize the significant 48–72-h synoptic-scale forecast errors associated with the 15 November 2018 East Coast cyclone. These adjoint-optimal perturbations, which have a maximum amplitude of about 2 K in potential temperature, are widespread, extending throughout the troposphere and across much of North America. We investigate the most impactful components of the perturbations by truncating them in physical and spectral space and rescaling them to be equal in an energy norm to the full, unmodified perturbations. When the perturbations are confined to localized target regions, including those with the highest adjoint sensitivity or the strongest convective instability, they have weaker impacts on the forecast than when the perturbations within the target regions are removed and the rest of the perturbations are retained. Additionally, when the perturbations are filtered to retain only wavelengths longer than 1000 km, they have greater impacts on the forecast than when the perturbations are filtered to retain only wavelengths shorter than 1000 km. These results suggest that the 15 November 2018 forecast bust was strongly sensitive to widespread, large-scale uncertainties in the initial conditions, rather than those in localized, small-scale regions that could more feasibly be reduced by targeted observations.
SIGNIFICANCE STATEMENT: Poor forecasts of midlatitude cyclones can cause tremendous socio-economic disruption via unexpected heavy precipitation and damaging winds. One approach to improving these forecasts involves targeting observations in localized regions where errors in the initial atmospheric state are expected to be most harmful. However, we find that a poor forecast of a midlatitude cyclone on 15 November 2018 was highly sensitive to widespread uncertainties in the initial state rather than those in a localized region. This suggests that targeted observations may not have been a fruitful strategy to improve the poor forecasts of this midlatitude cyclone.

1. Introduction

Accurate forecasts of midlatitude cyclones and their associated hazards are essential to protecting lives and property. Midlatitude-cyclone predictability has steadily increased over the last half-century, with present-day numerical weather prediction (NWP) models routinely producing accurate forecasts about one week in advance (Hoskins 2013; Bauer et al. 2015). Nevertheless, there are occasional midlatitude-cyclone “forecast busts” in which NWP models exhibit unusually large errors, even at much shorter 48–72-h lead times. In this study we focus on a forecast bust for a midlatitude cyclone that impacted the East Coast of the United States (U.S.) on 15 November 2018. Operational NWP models struggled to produce accurate forecasts of the location and intensity of the cyclone at 48–72-h lead times, and snow forecasts for some areas were quite poor at even shorter lead times. In particular, 12-h forecasts from the National Weather Service (NWS) predicted just 1–2 inches of snowfall for New York City (NYC), but more than 6 inches fell, breaking numerous daily records and shutting down travel in NYC (Novak et al. 2023).

One approach to mitigating forecast busts is observation targeting, which involves identifying regions in which initial-condition errors are expected to rapidly grow and assimilating additional observations in those regions to improve the analysis and reduce error growth. Many adjoint studies have suggested that midlatitude-cyclone forecasts are most sensitive to the initial conditions in regions that are localized enough to be adequately sampled by targeted radiosonde or aircraft observations. The sensitivity patterns in such studies are typically mesoscale, lower-tropospheric structures that tilt strongly upshear with height along baroclinic zones (Langland et al. 1995; Rabier et al. 1996; Langland et al. 1996; Ancell and Hakim 2007; Doyle et al. 2014, 2019). Although there were encouraging results from observation targeting campaigns for midlatitude cyclones during the
Fronts and Atlantic Storm Track Experiment (FASTEX) in 1997 and the North Pacific Experiment (NORPEX) in 1998 (Langland 2005), more recent results from The Observing Research and Predictability Experiment (THORPEX) from 2005–2014 indicated that improvements were generally positive, but marginal (Majumdar 2017). Explaining these underwhelming results is challenging because the degree of improvement depends on the availability of conventional observations, the data assimilation system, and the flow situation. For example, targeted aircraft observations have been shown to improve precipitation forecasts for atmospheric rivers impacting the U.S. West Coast (Lord et al. 2023). However, our hypothesis is that, for some midlatitude-cyclone forecast busts, the initial-condition sensitivity is so widespread and large in scale that targeting observations in a localized region would not lead to significant improvements.

Langland et al. (2002) considered whether the synoptic-scale forecast errors associated with the 25 January 2000 “surprise” snowstorm were sensitive to localized or widespread initial-condition uncertainties. They found that the adjoint-derived optimal perturbations for minimizing 72-h forecast errors for this storm were widespread, extending over large swaths of North America and the eastern Pacific. Langland et al. (2002) thus speculated that targeting a localized region for additional observations would have been an impractical method to mitigate this forecast bust. In contrast, Zhang et al. (2002) studied the same storm, instead focusing on shorter 24–36-h forecasts, and suggested that localized initial-condition errors played an important role in the bust by growing rapidly upscale from moist convection. Rodwell et al. (2013) focused on even longer 6-day forecast busts over Europe, and they found in their composite analysis of 584 cases that the initial conditions were associated with anomalously high convective available potential energy (CAPE) values in a relatively localized region over eastern North America. A potential interpretation of this result is that the forecasts suffered from rapid upscale error growth from convective regions, so targeted observations to reduce initial-condition errors in these regions might have mitigated the busts.

Given these contradictory predictability studies and the underwhelming results from observation-targeting campaigns for midlatitude cyclones, the goal of this study is to investigate the predictability of a recent forecast bust, considering whether it was sensitive to localized initial-condition uncertainties that could feasibly have been targeted for observation. Our focus is on the 15 November 2018 East Coast cyclone, and our strategy is to use an adjoint model to compute the initial-condition perturbations that minimize 48–72-h errors associated with this forecast bust. We provide a syn-
optic overview of the storm in section 2. Section 3 outlines the methods, including the model configuration and computation of the adjoint perturbations. In sections 4 and 5 we discuss the results of our simulations with parameterized and explicit convection, respectively. The conclusions are in section 6.

2. Synoptic overview

Figure 1 shows the Global Forecast System (GFS) analysis for 500-hPa geopotential height and sea level pressure over the 72-h period from 1200 UTC 13 November 2018 to 1200 UTC 16 November 2018. At the beginning of this period, a positively tilted upper-level trough was oriented over the central U.S., and surface high pressure was present over much of the contiguous U.S. (Fig. 1a). Over the next 24 h, a closed upper-level low detached from the longwave trough, and this trough propagated eastward in association with a surface cyclone exiting to the northeast (Fig. 1b). Meanwhile, the surface high pressure extended northeastward, resulting in cool northerly flow throughout much of the northeastern U.S. The upper-level cutoff low then moved northeastward slowly and deepened slightly, and a closed surface low formed along the Kentucky-Tennessee border by 1200 UTC 15 November 2018 (Fig. 1c). This surface low dissipated over the Appalachian Mountains as the upper-level cutoff low moved toward the Mid-Atlantic Coast, a region with strong low-level baroclinity due to the warm Gulf Stream waters and cold air dammed east of the Appalachian Mountains after the previous day’s northerly flow. Consequently, the inverted trough off the coasts of North and South Carolina on 1200 UTC 15 November 2018 developed into the midlatitude cyclone that brought snowfall to the northeastern U.S. in the evening of 15 November 2018 and early morning of 16 November 2018. The cyclone deepened to about 995 hPa by 1200 UTC 16 November 2018, and the upper-level cutoff low was ingested back into the longwave pattern (Fig. 1d). This case was an example of a “Miller Type-B” storm in which a surface cyclone dissipates over the Appalachian Mountains and a more intense cyclone develops along the East Coast due to the strong low-level baroclinity in that region (Miller 1946).

The deterministic GFS exhibited significant errors in forecasting the evolution of the upper-level cutoff low and the ensuing surface cyclone. Figure 2a shows that the GFS forecast initialized at 1200 UTC 13 November 2018 had the cutoff low too far south compared to the analysis, with errors greater than 100 m in 500-hPa geopotential height at just a 48-h lead time. These height
Fig. 1. GFS analysis for sea level pressure (black contours every 4 hPa) and 500-hPa geopotential height (color fill every 10 dam) at (a) 1200 UTC 13 Nov 2018, (b) 1200 UTC 14 Nov 2018, (c) 1200 UTC 15 Nov 2018, and (d) 1200 UTC 16 Nov 2018.

errors were even greater than those at similar lead times in the medium-range European bust cases analyzed by Magnusson (2017). Figure 2b shows that this GFS forecast also had the surface cyclone too far south (and too weak) compared to the analysis at a 72-h lead time. At 1200 UTC 16 November 2018, the analysis had the low just east of Long Island with a minimum central pressure of about 995 hPa, whereas the forecast had the low further south near New Jersey with a minimum central pressure of about 1001 hPa, amounting to errors in sea level pressure of about 12 hPa. Given these large synoptic-scale errors, our focus is on the sensitivity of the 48-h forecast for the upper-level cutoff low to the initial conditions at 1200 UTC 13 November 2018. We also consider the implications of this sensitivity on the 72-h forecast for the surface cyclone.
3. Methods

a. Nonlinear model

We use the Advanced Research version of the Weather Research and Forecasting Model (WRF-ARW version 4.4, Skamarock et al. 2021) to run nonlinear numerical simulations of this storm over the 72-h period from 1200 UTC 13 November 2018 to 1200 UTC 16 November 2018. We run two sets of simulations: one with parameterized convection and another with explicit convection.

The simulations with parameterized convection have a Lambert conformal conic projection for the horizontal domain with 30-km grid spacing, 400×250 grid points, a center grid point at 45°N and 100°W, and true latitudes at 20°N and 70°N. Figure 3 shows the full extent of this domain, which covers the entirety of North America and significant portions of the northeast Pacific and northwest Atlantic. There are 45 staggered vertical levels with a model top at 50 hPa and the Rayleigh damping scheme from Klemp et al. (2008) in the top 5 km to prevent gravity-wave reflections. Other model physics include the Thompson microphysics scheme (Thompson et al. 2008), the Tiedtke cumulus
Fig. 3. 48-h forecast errors in 500-hPa geopotential height for the 30-km WRF simulation. The GFS analysis (black) and WRF forecast (magenta) are contoured every 15 dam. The forecast minus the analysis (color fill) is contoured every 30 m. The full extent of the WRF domain is plotted. The dotted black line shows the extent of the COAMPS domain. The green contour bounds the response-function region used to compute the adjoint perturbations.

scheme (Tiedtke 1989; Zhang et al. 2011), the Noah land surface model (Tewari et al. 2004), the Mellor–Yamada–Janjić (MYJ) planetary-boundary-layer and surface-layer schemes (Janjić 1994), and the Rapid Radiative Transfer Model for Global climate models (RRTMG) for longwave and shortwave radiation. The initial conditions come from the GFS analysis at 1200 UTC 13 November 2018, and the GFS forecast initialized at this time is used to update the lateral boundary conditions every 3 h. Figure 3 shows that the 30-km WRF forecast for 500-hPa geopotential height is similar to the GFS forecast at a 48-h lead time, with the cutoff low well south of the analysis (cf. Fig. 2a).

For the simulations with explicit convection, all of the settings are the same except the grid spacing is 4 km, there are $3000 \times 1875$ grid points, there are 60 staggered vertical levels, and there is no cumulus parameterization. Additionally, we initialize the model at 0900 UTC 13 November 2018 to allow for a 3-h spinup period due to the downscaling of the coarser-resolution initial
conditions. The solution does not diverge significantly from the analysis over this 3-h period (not shown), so for the initial-condition perturbation experiments discussed in section 5 we add the perturbations at the same time as in the experiments with parameterized convection (i.e., at 1200 UTC 13 November 2018).

b. Tangent-linear and adjoint models

Adjoint models (Errico 1997) are powerful tools for quantifying the influence of each component of the initial model state $x_0$ on some scalar aspect $J$ of the model state at a later time $x_t$,

$$J(x_t) = J[M(x_0)],$$

where $M$ is the nonlinear model and $J$ is the response function. The sensitivity of the response function to changes in the initial state can be estimated by the gradient

$$\frac{\partial J}{\partial x_0} = M^T_{0,t} \frac{\partial J}{\partial x_t},$$

where the tangent-linear model $M_{0,t}$ is obtained by linearizing $M$ at each time step along a reference nonlinear trajectory from the initial time to the final time. The transpose of the tangent-linear model $M^T_{0,t}$ is the adjoint model, and it maps the gradient of $J$ with respect to the final state to the gradient of $J$ with respect to the initial state.

The adjoint-sensitivity gradient $\partial J/\partial x_0$ expresses the impact of a unit change to each initial-state variable on the response function. However, it is often more physically meaningful to express the sensitivity fields as “optimal” perturbations to the initial conditions that either minimize or maximize $J$ subject to some scaling. Such scaling is necessary to satisfy the tangent-linear approximation that perturbations evolved by $M_{0,t}$ are similar in magnitude and pattern to those evolved by $M$. The perturbations can also be scaled so that they are comparable in magnitude to analysis errors in real-world data assimilation systems.

In this paper we compute adjoint-derived optimal perturbations following Doyle et al. (2014, 2019). Changes to $J$ are expressed as

$$J' = \sum_{m,j} \frac{\partial J}{\partial x_{m,j}} x'_{m,j},$$
where \( \partial J / \partial x_{m,j} \) is the adjoint sensitivity gradient of the response function with respect to the initial value of variable \( m \) at grid point \( j \), and \( x'_{m,j} \) are the adjoint perturbations. We define \( x'_{m,j} \) for each prognostic variable as

\[
x'_{m,j} = s \frac{\partial J}{w_m \partial x_{m,j}}.
\]

The variables have different relative magnitudes (e.g., 1 K is a small change in potential temperature, but 1 kg kg\(^{-1}\) is a huge change in water vapor mixing ratio), so the perturbations are weighted by the domain-maximum forecast differences for each variable \( m \) over the integration period,

\[
w_m = \left[ \max_j (|x'^m_j - x^m_0|) \right]^{-2}.
\]

Finally, the scaling parameter \( s \) is defined such that the maximum perturbation for wind, potential temperature, or water vapor does not exceed 1 m s\(^{-1}\), 1 K, or 1 g kg\(^{-1}\), respectively. The perturbations are optimal in the sense that they produce the greatest (positive or negative) change to the response function as expressed by (3) for the smallest-magnitude perturbations subject to the constraint imposed by the scaling parameter \( s \) in (4).

To compute the adjoint-derived optimal perturbations for the 15 November 2018 forecast bust, ideally we would use the adjoint model for WRF (Zhang et al. 2013), but the code is built within WRF’s four-dimensional variational data assimilation module, so there is no user interface for stand-alone adjoint runs. Instead, we use the adjoint of the atmospheric module of the Coupled Ocean-Atmosphere Mesoscale Prediction System (COAMPS; Amerault et al. 2008).

Our COAMPS domain is nearly identical to that of the WRF runs with parameterized convection, with the same Lambert conformal conic projection and 30-km horizontal grid spacing. However, we use a smaller domain size with 301 \times 201 grid points due to computational constraints and potential instabilities in the adjoint model associated with large variations in the map scale factor. This smaller domain is indicated by the dotted black line in Fig 3. As in the WRF runs with parameterized convection, we use the GFS analysis at 1200 UTC 13 November 2018 for the initial conditions and the GFS forecast initialized at 1200 UTC 13 November 2018 for the lateral boundary conditions.

Model physics closely follow those of Doyle et al. (2014, 2019) and include a modified version of the Rutledge and Hobbs (1983) microphysics scheme, a modified Kuo convective parameterization
(Molinari 1985), a turbulent kinetic energy budget to parameterize turbulent mixing and diffusion processes in the boundary layer and free atmosphere (Hodur 1997), and a modified version of the Louis (1979) surface-layer parameterization. The tangent-linear and adjoint models use identical physics parameterizations as the nonlinear COAMPS model, but gradients and perturbations associated with vertical diffusion are neglected following Mahfouf (1999). There are 45 vertical levels with a model top at 30 km and a sponge upper-boundary condition over the top 10 km to prevent gravity-wave reflections.

c. Diagnostic forecast correction via iterative adjoint computations

Our goal in using the adjoint model is to compute the initial-condition perturbations that minimize the 48-h forecast errors for the upper-level cutoff low. Thus, we define the response function $J$ as the squared difference between the forecast potential vorticity (PV) and the analysis PV at 1200 UTC 15 November 2018 averaged over a region containing the cutoff low. The horizontal extent of this region is marked by the green contour in Fig. 3, and it extends vertically between the model levels closest to 500 and 300 hPa.

Starting from the nonlinear COAMPS simulation with GFS-analysis initial conditions $x_0$, we use the adjoint model to compute the perturbations $\delta x_0^{(1)}$ that produce the greatest tangent-linear decrease in $J$ following the procedure described in section 3b. These perturbations are then added to obtain the initial conditions for a new nonlinear run, 

$$x_0^{(1)} = x_0 + \delta x_0^{(1)}.$$  (6)

The nonlinear forecast starting from $x_0^{(1)}$ produces a cutoff low that is closer to the analysis than the original forecast, but $J$ can be reduced even further by using a new adjoint run to compute the optimal perturbations $\delta x_0^{(2)}$ corresponding to the improved nonlinear trajectory. This procedure is then iterated to minimize $J$ in a manner broadly analogous to the iterative cost-function minimization in variational data assimilation. Previous studies have used similar iterative adjoint methods to obtain the diagnostic initial-condition perturbations that correct forecast busts (Rabier et al. 1996; Langland et al. 2002; Kleist and Morgan 2005a). Rather than employing a convergence criterion, we iterate $n$ times until we qualitatively determine that the solution is sufficiently close
to the analysis. This results in the initial conditions $x_0^{(n)}$ that yield the corrected forecast,

$$x_0^{(n)} = x_0 + \delta x_0^{(1)} + \delta x_0^{(2)} + \ldots + \delta x_0^{(n)},$$

and the optimal perturbations that produce this correction are $\delta x_0 = x_0^{(n)} - x_0$. We use 9 iterations to produce a nonlinear COAMPS forecast that is sufficiently close to the analysis, and then the accumulated adjoint perturbations $\delta x_0$ are added to the WRF initial conditions.

Figure 4 demonstrates that the adjoint-derived perturbations $\delta x_0$ move the 30-km WRF forecast for the upper-level cutoff low substantially closer to the GFS analysis. In the 48-h control forecast with unperturbed initial conditions, the cutoff low is well southwest of the analysis and the average squared PV error in the response-function region is about 3.5 PVU$^2$ (Fig. 4a). When the adjoint-derived perturbations are added to the initial conditions, the 48-h forecast for the cutoff low is in nearly the same location as the analysis, and the average squared PV error in the response-function region is reduced to 0.8 PVU$^2$ (Fig. 4b). Notably, these perturbations are effective at improving the 30-km WRF forecast despite being computed using the COAMPS adjoint, suggesting that the sensitivity results are robust and more dependent on the initial conditions (which are identical between the simulations) than the model.

Figure 5 shows the structure of the adjoint-derived potential-temperature perturbations. These perturbations, which have a maximum amplitude of about 2 K, suggest that the 48-h forecast of the upper-level cutoff low, despite being a rather localized feature, is strongly sensitive to changes in the initial conditions throughout the troposphere and across much of North America. At low levels, the perturbations are confined primarily to the Gulf of Mexico (Fig. 5d), but further up in the troposphere the perturbations extend across most of western North America, with high-amplitude structures over Texas and Louisiana at 700 hPa (Fig. 5c), southwestern Mexico at 500 hPa (Fig. 5b), and the Pacific Northwest at 300 hPa (Fig. 5a). Overall, the forecast for the cutoff low appears to be sensitive to the initial specification of the entire longwave pattern, including both the trough that the low eventually detached from as well as the upstream ridge.

An important limitation of this adjoint method is that it does not take into account the likelihood that the sensitivity patterns represent actual errors in the analysis (Isaksen et al. 2005), so it cannot be used to definitively demonstrate that targeting observations in a localized region would not have
Fig. 4. 48-h errors (valid at 1200 UTC 15 Nov 2018) in the forecast of the upper-level cutoff low for the (a) control and (b) adjoint-perturbed 30-km WRF simulations. 400-hPa geopotential height is contoured every 10 dam for the GFS analysis (black) and the WRF forecast (magenta). The WRF forecast minus the GFS analysis for PV averaged between 300 and 500 hPa (color fill) is contoured every 1.5 PV units (1 PVU = \(10^{-6} \text{ K m}^2 \text{ s}^{-1} \text{ kg}^{-1}\)). The green contour bounds the response-function region used to compute the adjoint perturbations.

helped to mitigate this forecast bust. Nevertheless, the widespread sensitivity patterns shown in Fig. 5 do suggest, in a diagnostic sense, that the biggest potential for improvement in the 48-h forecast of the upper-level cutoff low comes from changes to the initial conditions that would be difficult to achieve by targeting a localized area with radiosonde or aircraft observations. Another important limitation of our method is that it does not consider the influence of model error. However, the fact that the GFS, WRF, and COAMPS models exhibit similar errors and the adjoint-derived perturbations from COAMPS evolve similarly in WRF suggests that initial-condition errors were more important to this forecast bust than model errors.
Fig. 5. Adjoint-derived potential-temperature perturbations (color fill every 0.4 K) and GFS analysis for geopotential height (black contours in dam) at 1200 UTC 13 Nov 2018 and at (a) 300, (b) 500, (c) 700, and (d) 850 hPa.

4. Experiments with parameterized convection

The sensitivity patterns for this forecast bust are generally widespread, but the highest-amplitude structures appear to be more localized (e.g., the feature over southeastern Texas at 700 hPa in Fig. 5c). Thus, it is possible that adding only those perturbations confined to a localized region of high sensitivity might still lead to substantial improvement in the forecast for the upper-level cutoff low. In this section we present 30-km WRF simulations in which the adjoint-derived perturbations are modified to examine the components of the perturbations that are most effective at improving the forecast.

To introduce the experiments we focus on 600 hPa, which is the level at which the potential-temperature perturbations reach their maximum. The WRF simulation with unperturbed initial conditions is the “control” experiment (Fig. 6a). The “full” experiment (Fig. 6d) corresponds to
Fig. 6. 600-hPa potential-temperature perturbations (color fill every 0.5 K) for the (b) target, (c) short-\( \lambda \), (d) full, (e) target\(^c \), and (f) long-\( \lambda \) experiments. The GFS analysis for 600-hPa geopotential height at 1200 UTC 13 Nov 2018 is contoured in black every 15 dam. The green contours in (b) and (e) bound the target region.

the WRF simulation in which the full, unmodified adjoint perturbations are added to the initial conditions. For the “target” experiment (Fig. 6b), the perturbations are confined to a 15° × 15° region containing the maximum potential-temperature perturbation (which is along the Kansas-Oklahoma border) and a high-amplitude wave pattern over Texas. In set theory, the complement of a set \( A \), denoted \( A^c \), is the set of elements not in \( A \). Thus, the “target\(^c \)” experiment (Fig. 6e) involves setting the perturbations within the target region to zero and retaining the rest. To address the issue of scale dependence, we truncate the perturbations in spectral space to filter the perturbations by wavelength \( \lambda \). For the “short-\( \lambda \)” experiment (Fig. 6c) only wavelengths shorter than 1000 km are retained, and for the “long-\( \lambda \)” experiment (Fig. 6f) only wavelengths longer than 1000 km are retained.
Although these experiments are configured based on the structure of the potential-temperature perturbations, we similarly modify the perturbations to the other prognostic variables. Also, to ensure that the perturbations are comparable in initial magnitude, we rescale the modified perturbations so that the domain-integrated difference total energy (DTE; Zhang et al. 2007) equals that of the full adjoint perturbations. The DTE is given by

\[
\text{DTE} = \frac{1}{2} \left[ (\delta u)^2 + (\delta v)^2 + \kappa (\delta T)^2 \right],
\]

where \(\delta u\), \(\delta v\) and \(\delta T\) are the differences between the perturbed and control fields of zonal wind, meridional wind, and temperature, respectively, \(\kappa = c_p / T_r\), \(c_p = 1004 \text{ J K}^{-1} \text{ kg}^{-1}\) is the specific heat at constant pressure, and \(T_r = 270 \text{ K}\) is the reference temperature.

Figure 7 shows the impacts of these perturbations on the 48-h, 30-km WRF forecast for the cutoff low. In the target experiment (Fig. 7b), the localized perturbations do produce a cutoff low that is closer to the analysis in both location and intensity than the unperturbed control simulation (cf. Fig. 7a). Nevertheless, these improvements pale in comparison to those of the full adjoint perturbations (cf. Fig. 7d), as the cutoff low in the target experiment is still well southwest of the analysis and there remain substantial PV errors. The target \(\mathcal{C}\) perturbations (Fig. 7e) improve the forecast for the cutoff low significantly more than the target perturbations, especially in terms of location. In fact, the target \(\mathcal{C}\) experiment has only slightly greater PV errors than the full experiment. Interestingly, both the short-\(\lambda\) (Fig. 7c) and long-\(\lambda\) (Fig. 7f) perturbations produce similar significant improvements to the 48-h forecast, although neither match the improvement attained by the full or target \(\mathcal{C}\) perturbations.

Figure 8 shows that the forecast accuracy for the upper-level cutoff low at 1200 UTC 15 November 2018 is strongly associated with the forecast accuracy for the surface cyclone 24 h later. The control WRF forecast at 72 h (Fig. 8a) has the low over southern New Jersey with a minimum central pressure of about 1001 hPa, whereas the analysis has the low near Long Island with a minimum central pressure of about 995 hPa, amounting to errors in sea level pressure of about 15 hPa. These characteristics are similar to the 72-h GFS forecast for sea level pressure (cf. Fig. 2b). The full adjoint perturbations (Fig. 8d) significantly improve the forecast for both the location and intensity of the surface cyclone. This is noteworthy because the perturbations are only optimized to reduce the forecast errors in the upper-level PV field at 48 h, yet they also substantially improve the
representation of the surface cyclone 24 h later. This suggests that the upper-level cutoff low was a key synoptic-scale forecast challenge associated with this bust.

The relative performances of the simulations with modified adjoint perturbations are similar in their 72-h surface forecasts as in their 48-h upper-level forecasts. The targetC perturbations (Fig. 8e) produce substantial reductions in both location and intensity errors in the surface cyclone that are comparable to that of the full adjoint perturbations, whereas the target experiment (Fig. 8b) performs nearly as poorly as the unperturbed control simulation. Both the short-λ (Fig. 8c) and

Fig. 7. 48-h errors (valid at 1200 UTC 15 Nov 2018) in the forecast of the upper-level cutoff low for the (a) control, (b) target, (c) short-λ, (d) full, (e) targetC, and (f) long-λ experiments. 400-hPa geopotential height is contoured every 10 dam for the GFS analysis (black) and the WRF forecast (magenta). The WRF forecast minus the GFS analysis for PV averaged between 300 and 500 hPa (color fill) is contoured every 1.5 PVU.
Fig. 8. 72-h errors (valid at 1200 UTC 16 Nov 2018) in the forecast of the surface cyclone for the (a) control, (b) target, (c) short-λ, (d) full, (e) target^C, and (f) long-λ experiments. Sea level pressure is contoured every 6 hPa for the GFS analysis (black) and the WRF forecast (magenta). The WRF forecast minus the GFS analysis for sea level pressure (color fill) is contoured every 3.5 hPa.

Long-λ (Fig. 8f) perturbations yield substantial improvements in the 72-h surface-cyclone forecast, but neither match the improvements attained by the full or target^C perturbations.

Overall, the results of these WRF simulations with parameterized convection indicate that the widespread contributions of the adjoint perturbations are essential to yielding substantial reductions in forecast error: confining the perturbations to a localized target region significantly impedes their ability to improve the forecast, whereas eliminating the perturbations in that target region or over a range of wavelengths does not.
5. Experiments with explicit convection

a. Comparison with the 30-km results

The moist adjoint model is run with parameterized convection and 30-km grid spacing in order to satisfy the tangent-linear approximation, mitigate nonlinear instabilities, and avoid spurious sensitivities. Nevertheless, there has been considerable debate over the importance of moist convection to midlatitude-cyclone predictability, with some studies suggesting that upscale error growth from convection plays a significant role in forecast busts (Zhang et al. 2002; Rodwell et al. 2013) and others arguing that such upscale growth is overwhelmed by relatively small-amplitude initial errors on much larger scales (Langland et al. 2002; Durran et al. 2013; Lloveras et al. 2023). To address this issue, we present a set of initial-condition-perturbation experiments in our convection-permitting WRF runs with 4-km grid spacing. These simulations use the same adjoint perturbations as in section 4, but interpolated onto the 4-km mesh.

Figure 9 shows that the full adjoint perturbations do not correct the 48-h forecast for the upper-level cutoff low in the 4-km WRF simulation, in stark contrast to the substantial correction produced by the perturbations in the 30-km WRF run (cf. Fig. 4). The cutoff low in the 4-km control WRF forecast (Fig. 9a) is less far south of the analysis than the corresponding 30-km run, and the low is east of the analysis rather than west. As a result, when the adjoint-derived perturbations move the cutoff low to the northeast in the 4-km simulation (the same effect as in the 30-km run), they push the low further away from the analysis, increasing the PV error rather than reducing it (Fig. 9b). This indicates that the adjoint-derived perturbations are not optimal for improving the 4-km WRF forecast because of the nonlinearities associated with representing smaller-scale processes. This also suggests that the convective parameterization is an important aspect of model error that is not addressed with our adjoint method. Nevertheless, the adjoint perturbations still have a significant impact on the forecast of the cutoff low in the 4-km WRF run. Thus, in the remainder of this section we focus on how different initial-condition perturbations produce forecast changes relative to the unperturbed control simulation, rather than on their performance relative to the analysis. In other words, we use the perturbations to investigate the initial-condition sensitivity in the 4-km simulation without expecting the perturbations to improve the forecast accuracy.
To investigate the components of the adjoint perturbations that produce the greatest changes relative to the control simulation, we run 4-km WRF simulations with the same set of perturbations as in the 30-km experiments (Fig. 6), but interpolated onto the 4-km mesh. Figure 10 shows the impacts of these perturbations on the 48-h forecast of the upper-level cutoff low. Note that in these plots, the more impactful perturbations exhibit more intense colors because they produce greater differences relative to the control simulation. This is in contrast to our presentation of the 30-km results (e.g., Fig. 7) in which the more impactful perturbations exhibit less intense colors because they produce solutions that are closer to the analysis.

Figures 10d and 10e show that the full and target perturbations have the greatest impacts on the 48-h forecast of the upper-level cutoff low, consistent with the 30-km experiments. Both of these perturbations deepen the cutoff low and move it more than 100 km to the northeast of the control. In contrast, the target perturbations (Fig. 10b) only deepen the low slightly and do not introduce any significant differences in location. The short- and long- $\lambda$ perturbations (Figs. 10c,f) both produce greater impacts than the target perturbations, also consistent with the 30-km experiments. However, the long- $\lambda$ perturbations cause significantly greater changes than the short- $\lambda$ perturbations to both

![Fig. 9. As in Fig. 4, but for the 4-km WRF forecasts.](image-url)
Fig. 10. 48-h changes (valid at 1200 UTC 15 Nov 2018) to the 4-km WRF forecast of the upper-level cutoff low produced by the (b) target, (c) short-\(\lambda\), (d) full, (e) target\(^c\), and (f) long-\(\lambda\) perturbations. 400-hPa geopotential height is contoured every 10 dam for the control (black) and perturbed (magenta) simulations. The perturbed run minus the control run for PV averaged between 300 and 500 hPa (color fill) is contoured every 1.5 PVU.

The intensity and location of the cutoff low in the 4-km simulations, whereas they are about equally impactful in the 30-km runs (cf. Figs. 7c,f).

Figure 11 shows that the perturbations that cause the greatest changes to the upper-level cutoff low at 48 h cause the greatest changes to the surface cyclone at 72 h, consistent with the 30-km experiments. The full and target\(^c\) perturbations (Fig. 11d,e) produce nearly identical changes to the surface cyclone, moving the low pressure center several hundred km to the northeast and resulting in perturbations in sea level pressure of nearly 15 hPa. In contrast, the target perturbations produce only a very minor northeastward shift in the cyclone’s location (Fig. 11b). The short- and long-\(\lambda\)
Fig. 11. 72-h changes (valid at 1200 UTC 16 Nov 2018) to the 4-km WRF forecast of the surface cyclone produced by the (b) target, (c) short-λ, (d) full, (e) target$^C$, and (f) long-λ perturbations. Sea level pressure is contoured every 6 hPa for the control (black) and perturbed (magenta) simulations. The perturbed run minus the control run for sea level pressure (color fill) is contoured every 3.5 hPa.

perturbations (Fig. 11c,f) produce similar shifts to those in the full and target$^C$ experiments, but as for the 48-h forecast of the cutoff low, the long-λ perturbations have a greater impact on the surface cyclone at 72 h than the short-λ perturbations.

In summary, the initial-condition perturbations that are most effective at modifying the 4-km forecasts with explicit convection are the same as those for the 30-km forecasts with parameterized convection (i.e., widespread and large in scale) except that, unlike the 30-km forecasts, the long-λ perturbations have greater impacts than the short-λ perturbations.
Fig. 12. Convective available potential energy (CAPE, color fill every 400 J kg\(^{-1}\)) from the GFS analysis at
(a) 1200 UTC 13 Nov 2018, (b) 1200 UTC 14 Nov 2018, (c) 1200 UTC 15 Nov 2018, and (d) 1200 UTC 16
Nov 2018. The GFS analysis for 500-hPa geopotential height is contoured in black every 15 dam. The magenta
contours bound the CAPE-target region.

b. Growth of perturbations in high-CAPE regions

Rodwell et al. (2013) determined that a localized high-CAPE region over eastern North America
was important to 6-day forecast busts over Europe. Our focus is on shorter 48–72-h lead times,
but we also are interested in whether initial-condition sensitivity in high-CAPE regions played an
important role in the 15 November 2018 forecast bust. Figure 12 shows the surface-based CAPE
from GFS analysis over the 72-h period of our WRF simulations. At the initial time, there were very
high CAPE values greater than 1600 J K\(^{-1}\) kg\(^{-1}\) off the coast of southwestern Mexico and over the
Caribbean Sea. There was also substantial CAPE (with values greater than 1200 J K\(^{-1}\) kg\(^{-1}\)) over
the Gulf of Mexico and extending over the southeastern U.S. and the northwest Atlantic. These
high-CAPE regions remained mostly stationary as the system developed, with the CAPE values
gradually decreasing over the 72-h period.
Recalling that the forecast for the upper-level cutoff low is sensitive to low-level perturbations over the Gulf of Mexico and the southeastern U.S. (Fig. 5d), we conduct a second target experiment in which the adjoint-derived perturbations are confined to a 15° × 15° high-CAPE region over the Gulf of Mexico (denoted by the magenta contour in Fig. 12a). We call this the “CAPE-target” experiment, and we conduct a corresponding “CAPE-targetC” experiment in which the perturbations within this region are removed and those outside the region are retained. As in the previous experiments, we rescale these perturbations to have the same domain-integrated DTE as the full adjoint perturbations.

Fig. 13. (a),(b) As in Fig. 10, but for the CAPE-target and CAPE-targetC experiments. (c),(d) As in Fig. 11, but for the CAPE-target and CAPE-targetC experiments.
Figure 13 shows the impacts of the CAPE-target\(^C\) and CAPE-target perturbations on the 48-h upper-level and 72-h surface forecasts. The CAPE-target perturbations cause noticeable eastward shifts in the upper-level cutoff low at 48 h (Fig. 13a) and the surface cyclone at 72 h (Fig. 13c). This is in contrast to the other target perturbations, which do not significantly impact either feature (cf. Figs. 10b, 11b). The changes produced by the CAPE-target perturbations are comparable with those in the CAPE-target\(^C\) experiment, although the CAPE-target\(^C\) perturbations yield slightly greater shifts, especially in the northward direction, for both the upper-level cutoff at 48 h (Fig. 13b) and the surface cyclone at 72 h (Fig. 13d). Nevertheless, neither the CAPE-target\(^C\) nor CAPE-target perturbations have as much of an impact as the full or target\(^C\) perturbations (cf. Figs. 10d,e, 11d,e). These experiments indicate that the high-CAPE region of strong low-level sensitivity is indeed an impactful component of the adjoint perturbations, but that the forecast is even more sensitive to widespread changes in the initial conditions.

c. DTE growth

To provide a quantitative, overall assessment of the impact of each set of perturbations on the simulations with explicit convection, Fig. 14 depicts the 72-h evolution of domain-integrated DTE, where the differences are computed between the perturbed and control simulations. Recall that we configure all perturbations to be equal in initial DTE (Fig. 14a). The DTE time series show that the greatest differences at 48–72-h lead times are produced by the perturbations that are widespread or large in scale (solid curves) rather than by the perturbations that are localized or small in scale (dashed curves). This result is best illustrated by the plot with the linear ordinate axis (Fig. 14b). In the first 30 h the CAPE-target experiment has greater DTE than the other experiments, and the short-\(\lambda\) experiment has similar DTE to the long-\(\lambda\) and CAPE-target\(^C\) experiments. However, by 48 h the full, target\(^C\), long-\(\lambda\), and CAPE-target\(^C\) experiments have greater DTE than the target, short-\(\lambda\), and CAPE-target experiments, and this separation becomes even more significant by 72 h.

6. Conclusions

In this paper we explore the predictability of the 15 November 2018 East Coast cyclone, focusing on the initial-condition uncertainties leading to significant synoptic-scale forecast errors at relatively short 48–72-h lead times. We use iterative runs of a moist adjoint model to obtain the
Fig. 14. (a) Time series of domain-integrated difference total energy (DTE; m$^2$ s$^{-2}$) through 72 h of lead time for all initial-condition-perturbation experiments with explicit convection. (b) As in (a), but with a linear ordinate axis instead of logarithmic.

We consider whether the most impactful components of the perturbations are widespread or localized by truncating them in physical space and rescaling them to be equal in an energy norm to the full, unmodified perturbations. When the perturbations are confined to a $15^\circ \times 15^\circ$ target region containing the highest-amplitude feature, they do not produce significant forecast improvements. In contrast, when the perturbations within this target region are removed and the rest of the perturbations are retained, they yield substantial forecast improvements comparable to those produced by the full perturbations. We also consider whether the most impactful components of the perturbations are small or large in scale by truncating them in spectral space, retaining only wavelengths shorter or longer than 1000 km, and rescaling them to be equal in an energy norm to the
full perturbations. Both the short- and long-wavelength perturbations produce substantial forecast improvements, albeit less than those yielded by the full perturbations. These results suggest that this forecast bust was sensitive to widespread uncertainties in the initial conditions, rather than those in a localized region or over a specific range of wavelengths.

The adjoint perturbations for the 30-km parameterized-convection forecasts are not optimal for improving the 4-km convection-permitting forecasts, but they do have a substantial impact when compared to the unperturbed control simulation. Thus, with our convection-permitting simulations we use the adjoint perturbations to explore the initial-condition sensitivity of this bust without focusing on improvements to forecast accuracy. When the perturbations are confined to $15^\circ \times 15^\circ$ target regions, including one that encompasses a region with high CAPE and strong low-level sensitivity over the Gulf of Mexico, they are less impactful than when the perturbations inside the target regions are removed and the rest of the perturbations are retained. Also, the perturbations with only wavelengths longer than 1000 km retained are more impactful than those with only wavelengths shorter than 1000 km retained.

Given that the most impactful initial-condition perturbations for this forecast bust are widespread and large in scale, it seems unlikely that assimilating additional observations in a localized target region could have mitigated the bust. This conclusion is similar to that of Langland et al. (2002) in their study of the 25 January 2000 “surprise” snowstorm. Nevertheless, as we note in section 3c, an important limitation of our diagnostic method is that it does not take into account the likelihood that the analysis errors project onto the sensitivity pattern. Ultimately, the degree of improvement from any targeted observing system depends on the availability of conventional observations and the data assimilation system. Even so, our sensitivity study suggests that focusing on improving the fidelity of the analysis over widespread, large-scale regions (as opposed to localized, small-scale regions that can more feasibly be sampled by targeted radiosonde or aircraft observations) may be the best strategy for mitigating similar 48–72-h midlatitude-cyclone forecast busts.

Another important limitation of our study is that we examine only one midlatitude cyclone with only one response function. The response-function choice does impact adjoint-sensitivity patterns (Kleist and Morgan 2005b), but the localized versus widespread nature of such patterns (and thus the opportunity for improvement from targeted observations) can vary significantly between cases, even for similar response functions. For example, Jung and Kim (2009) investigated the sensitivity
of a heavy snowfall event over the Korean Peninsula using a forecast-error response function and an iterative adjoint method similar to ours, but their optimal perturbations were much more localized. Additionally, Doyle et al. (2014) used a kinetic-energy response function to compute the adjoint sensitivity for Cyclone Xynthia and found highly localized patterns, whereas Doyle et al. (2019) used the same type of response function to study Cyclone Desmond and found more widespread patterns. In contrast, other flow situations like atmospheric rivers more consistently exhibit localized sensitivity patterns across a variety of cases and response functions (Reynolds et al. 2019), which may explain why they more reliably experience forecast improvements from targeted observations (Stone et al. 2020; Lord et al. 2023; DeHaan et al. 2023). Future research on initial-condition sensitivity in midlatitude cyclones could help identify which bust situations, if any, are likely to see improvement from targeted observations.
**Acknowledgments.** Daniel J. Lloveras and Dale R. Durran acknowledge the support of Grant N00014-20-1-2387 from the Office of Naval Research. James D. Doyle acknowledges the support of the Naval Research Laboratory Base Program through Program Element 0601153N. We acknowledge computational resources from the Navy Department of Defense Supercomputing Resource Center in Stennis, Mississippi.

**Data availability statement.** The model output files used in this study are too large to publicly archive or to transfer. However, the control simulations can be easily reproduced because the GFS datasets used for the initial and boundary conditions are publicly available (https://rda.ucar.edu/datasets/ds084.1) and the WRF model is publicly available (https://github.com/wrf-model/WRF/releases/tag/v4.4). The perturbed simulations can be reproduced by adding the adjoint perturbations to the initial conditions, which are archived with the University of Washington Libraries ResearchWorks repository (https://hdl.handle.net/1773/51016).

**References**


