Lewis, J. M., S. Lakshmivarahan, J. Hu, and R. Rabin, 2020: Placement of observations to correct return-flow forecasts. *Electronic J. Severe Storms Meteor.*, **15** (4), 1–20.

Electronic Journal of
SEVERE STORMS METEOROLOGY

Placement of Observations to Correct Return-Flow Forecasts

JOHN M. LEWIS
National Severe Storms Laboratory, Norman, Oklahoma, and University of Nevada-Reno, Reno, Nevada

S. LAKSHMIVARAHAN
School of Computer Science, University of Oklahoma, Norman, Oklahoma

JUNJUN HU
Cooperative Institute of Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma, and National Severe Storms Laboratory, Norman, Oklahoma

ROBERT RABIN
National Severe Storms Laboratory, Norman, Oklahoma, and Space Science and Engineering Center, Madison, Wisconsin

(Submitted 30 May 2020; in final form 15 December 2020)

ABSTRACT

The continued presence of systematic errors in operational forecasts of return flow over the Gulf of Mexico has motivated an investigation into this problem. The theme of the work is use of a low-order mixed-layer model that is faithful to the phenomenon in the context of dynamic data assimilation. Data assimilation experiments in the identical-twin mode determine the best place to make observations that minimize the forecast error through adjustment of model controls. The emphasized controls are those associated with the fluxes of heat and moisture from sea to atmosphere. Results indicate that the best observations are at that time and place when the outflowing continental air passes over the warmest sea surface temperatures. In the case studied, this warmest zone is directly over the Loop Current. Observations at times long after the modified air leaves these warmest waters lead to relatively poor control adjustments and little improvement in the forecast. If input to data assimilation is restricted to observations of a single model variable over short intervals of time (the order of several hours), results are relatively poor. Yet, a significant improvement is forthcoming if one of the observations is replaced by an observation from another model variable. This result is understood through arguments based on forecast sensitivity to model control. The paper ends with discussion of steps to be taken that hold promise for correcting systematic error in return-flow forecasts.

1. Introduction

Cold-air outflows over the Gulf of Mexico (GoM) were first studied by Palmén and Newton (1951). They were intent on understanding the exchange of momentum between mid-latitudes and the tropics in support of a mechanism that could maintain the subtropical jet stream—a general circulation viewpoint. Dallavalle and Bosart (1975) then studied cold-air outbreaks from a different perspective, viz., evolution of the synoptic flow and subtleties of interaction between the cold airmass and the GoM.

In the mid- to late 1970s, Texas A & M University Professor Walter Henry became fascinated with the anticyclonic turning of the cold airmass and its return to the Gulf's coastal plain in response to the eastward movement of the attending large-scale anticyclone (Henry, personal communication, 1987). He coined the term "return flow" to describe this process, and two of his master's degree students used the term in titles of their theses (Johnson 1976; Karnavas 1978). He later published a paper on the climatology of frontal penetrations into the GoM

Corresponding Author: John M. Lewis, Email: *John.Lewis@dri.edu*

during 1967–1977, with special emphasis on the return-flow events (Henry 1979, Table 2). Over the 11-y period of study, 88 return-flow events occurred during the cool season (November–April), and 22 return-flow events occurred during warm season (May–October, with none in July and only one in August)—an average of 8.0 events during the cool season and 2.0 events during the warm season.

In response to concerns about systematic errors in National Centers for Environmental Prediction's (NCEP's) operational forecasts of return flow over the GoM in the mid-1980s, a consortium of research organizations proposed investigation of this weather prediction problem to the National Science Foundation (NSF). NSSL took the lead in this proposal, but strong support came from several organizations, especially the Southern Region of the National Weather Service (NWS), the National Severe Storms Forecast Center (NSSFC, now the Storm Prediction Center, SPC), and the Space Science and Engineering Center (SSEC) at University of Wisconsin-Madison. The NSF award was granted in late 1987. The organizational structure of the project, called GUFMEX (Gulf of Mexico Experiment), and a summary of results from the field program in 1988 can be found in the paper by Lewis et al. (1989). During this program, six return flow events were observed from land, sea, and air in a coordinated fashion. These data were central to a series of contributions that appeared in the *Journal of Applied Meteorology*'s August 1992 issue.

Despite improvements in operational products at NCEP over the past 20–30 y that directly bear on GoM weather forecasts [1], improvements such as better sea surface temperatures (SSTs) (instituted in 2001) and convective schemes such as the Betts-Miller-Janjic scheme (Janjic 1994,2001), the systematic errors remain as discussed in the Supplement to Lewis et al. (2016; R. Maddox, Supplement). For a March 2015 return-flow event, Maddox (2016) compared forecasts from the large-scale spectral Global Forecast System (GFS) and the smaller-scale North American Mesoscale Forecast (NAM with a 12-km grid). In conclusion, he said:

> …while the current generation of operational numerical forecasts have improved, relative to the older models mentioned in the body of the paper [Lewis et al. 2016][2], there are still problems with many aspects of the model forecasts for air mass modification during transits over the warm waters of the GoM.

Difficulty in making accurate return-flow forecasts stem from the absence of upper-air observations over the GoM. The initial structure of the lower tropospheric air just prior to entry into the GoM is determined with fidelity, but after convective heating and moistening commences, it is often difficult to determine the boundary-layer structure. The following quotes from NOAA/NCEP meteorologist Geoffrey Manikin[3] discussed this continuing problem with forecasting the boundary layer structure. We include two quotes, one from 2007 and the other from 2020, respectively, separated by the slash ("/"):

> Moisture return is a key issue in convective forecasting, as subtle low-level changes can be the difference between a stable sounding or one able to support deep convection…the problem is most pronounced at the leading edge of the deeper moisture. It is often the case that there is a well-mixed temperature profile (constant θ), but a poor shape of the dewpoint curve. / My statement was definitely true in 2007, and I think it's still true for the most part today. …SST data is much better now, and model physics have improved greatly in the past 13 years, but we still lack vertical profile data over the Gulf. That said, I think we are much better now at simulating return flow in the models, but challenges remain, and the concept of determining where new obs would be best placed seems like a worthwhile pursuit.

Regarding the challenges that still remain, the important recent work of Cohen et al. (2017) has noted the extreme sensitivity to cold-season forecasted low-level thermodynamic structure (specifically low-CAPE) based on the choice of planetary boundary-layer scheme in high-resolution models like NAM and WRF [Weather

[1] Improvements began with efforts at the National Meteorological Center (NMC) which changed its name to National Centers for Environmental Prediction on 1 October 1995.

[2] Bracketed information inserted by the authors.

[3] Manikin is a member of the Environmental Prediction Center (EMC), a subunit within NCEP.

Research and Forecasting Model]. Cohen et al. (2017; section 3) showed that slight errors in the WRF forecast of these low-level structures in the southeast USA were central to poor severe-storm forecasts/outlooks.

Given that routine upper-air observations over the GoM are unavailable, it is prudent nevertheless to consider the hypothetical situation where upper-air observations will become available. Then the question becomes: In the context of data assimilation where forecasts are blended with observations, where is the ideal place to make observations? That is, placement that is most likely to correct forecast-model controls and thereby improve the forecast. This is the theme of the current investigation.

The paper is structured as follows: section 2 discusses the problems facing return-flow prediction. In section 3, the model and observations from GUFMEX are described. The data-assimilation methodology is given in section 4 and section 5 provides the results from numerical experiments. The discussions and conclusions are detailed in section 6. An appendix discusses the assimilation method and presents the equations governing the scheme.

2. Uncertainty in return-flow prediction: Ensemble viewpoint

Lewis et al. (2016) designed ensemble forecasts of a particular return flow event to identify forecast sensitivity to the controls that governed a mixed layer model—forecast uncertainty in response to changes to the control vector. Control elements were categorically divided into initial conditions (IC), boundary conditions (BC), and physical/empirical parameters (P). The controls in the ensemble forecast are described and numerically specified in tables found in Lewis et al. (2016; section 4). With the limited number of controls in the mixed layer model, roughly 50, the ensemble statistics of return flow could be determined very accurately based on 200 000 model executions. The controls used in the ensemble forecast are described and numerically specified in Lewis et al. (2016; section 4).

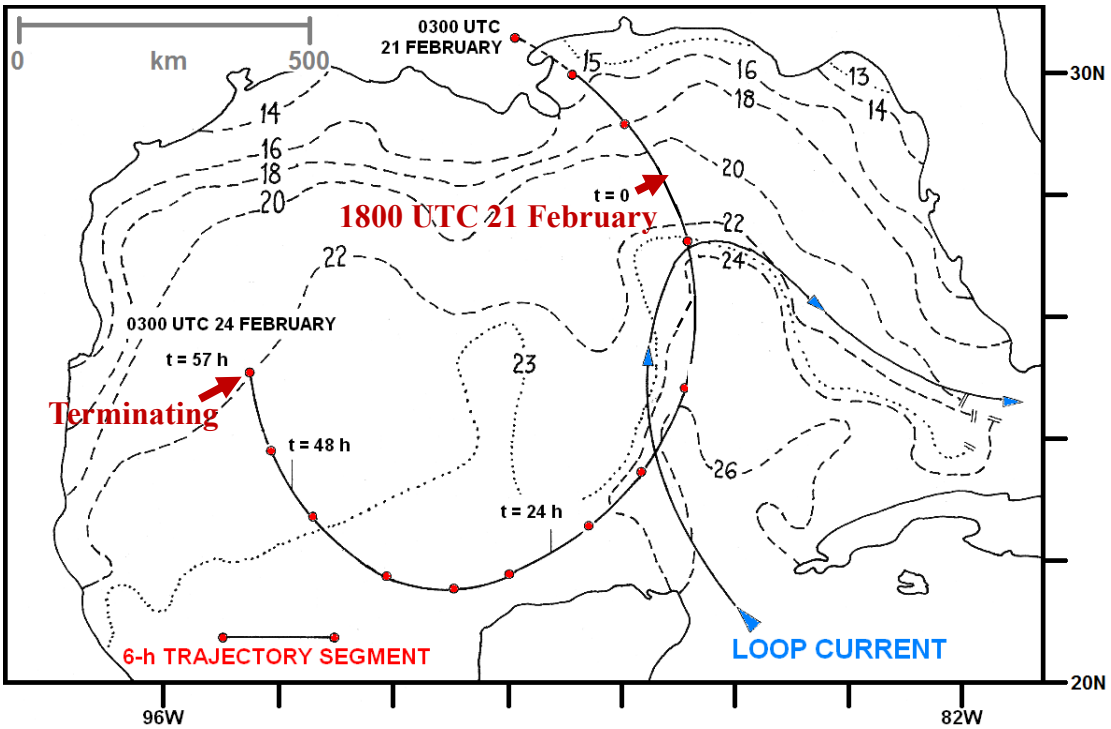

Figure 1: Lagrangian trajectory of modified continental air that left the coastal plain near New Orleans at 0300UTC 21 February and executed an elongated 3-day anticyclonic path over the GoM, where travel times along the trajectory are noted by small filled circles at 6-h intervals. Isotherms of sea surface temperature (in °C) are represented by dashed curves. The model integration begins at 1800UTC 21 February, as shown beside the symbol t = 0 (adapted from Fig. 3 in Lewis et al. 2016).

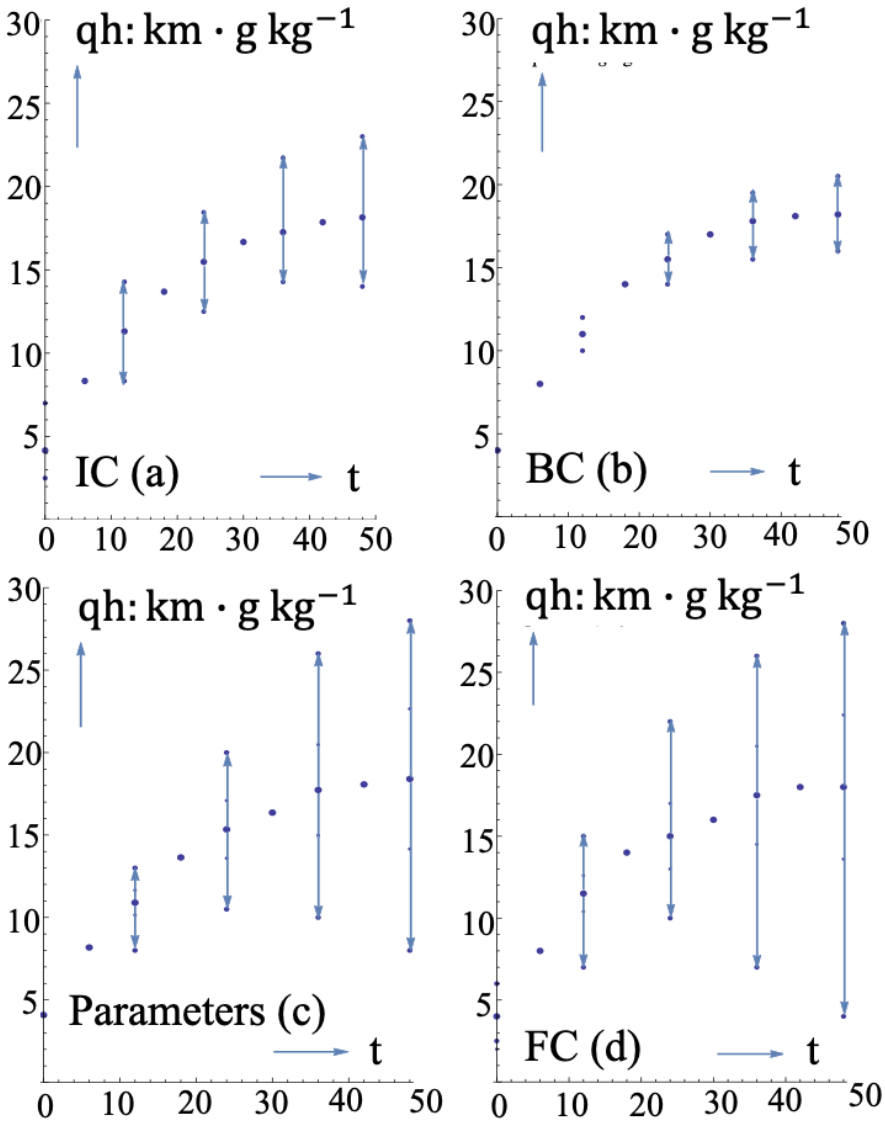

Figure 2: Uncertainty in ensemble prediction of qh, water vapor mass in the mixed-layer column, as a function of model controls: a) IC; initial conditions alone, b) BC; boundary conditions alone, c) P; empirical/physical parameters alone, and d) the full complement (FC) of controls.

The ensemble forecasts were made along a Lagrangian trajectory of low-level air associated with the return flow event of 21–25 February 1988. This trajectory is shown in Fig. 1. The boundaries of the warm Loop Current, a persistent feature in the eastern GoM are determined from both sea surface temperatues (SSTs) and deeper sea temperatures (at ~150-m depth for example) as discussed by Molinari (1987). In the year of our study, 1988, the Loop Current extended farther north than the climatological average and was therefore labeled a "deep intrusion". The boundaries of this Loop Current are quite precisely defined by the 24°C sea surface temperature (SST) contour shown in Fig. 1[4].

Figure 2 presents ensemble prediction results along the trajectory from t = 0 to t = 48 h for qh, the mass of water vapor in a mixed layer

[4] These boundaries can be verified by examining SSTs based on infrared radiance measurements during the week of 20–26 February 1988 as shown in Lewis (2007; Fig. 4).

of height h and water vapor mixing ratio q, a surrogate for precipitable water (PW in mm). The forecast panels present uncertainty due to IC alone (a), BC alone (b), P alone (c), and the full complement FC (d), uncertainty due to the combination of the three components.

The large dots in the panels display the mean values of water-vapor mass at 6-h intervals along the trajectory. The arrows emanating from mean values at 12-h intervals depict the range of ensemble forecasts. The standard deviations of the forecasts for P alone and FC are noted along the arrows by the small dots. The standard deviations of qh in units of $km \cdot g\ kg^{-1}$ at t = 48 h for each of the components follow: 1.12 (IC), 0.48 (BC), 4.20 (P), and 4.45 (FC). In units of mm for PW, P exhibits a relatively large range from 8–28 mm compared to the other components, and this range is slightly less than the range exhibited for FC, 4–29 mm. The range for BC is especially small, 17–20 mm, and IC is significant, 14–20 mm, but considerably smaller than the range for P.

In the case of severe-storm forecasting, the degree of uncertainty shown in Fig. 2 makes it difficult for forecasters to decide on the type of weather associated with a particular return-flow event—a wide spectrum that includes fog and stratus, shallow convection, and severe convective storms. The recent paper by Molina and Allen (2019) makes it clear that forecasting tornadic events is made especially challenging in return-flow situations.

3. The model and observations from GUFMEX

a. Observations

Our study relies on observations collected for a particular return-flow event during GUFMEX I in February–March 1988)[5]. The event occurred over the period 21–24 February and was associated with movement of a maritime Polar (mP) airmass from the Pacific Northwest to the GoM. The satellite imagery of this event over the period 21–23 February is shown in Fig. 1 of Lewis et al. 2016. The erosion of the low-level stably stratified airmass as it moved over the GoM's warm water was captured by a series of rawinsonde observations from Salvia, a U. S. Coast Guard ship that generally tended buoys in the GoM's shelf waters (Figs. 3 and 4). The salient features of potential temperature (θ) and water-vapor mixing ratio profiles follow:

θ[6] profile:

i) Superadiabtic layer, $\frac{d\theta}{dz} < 0, lowest\ 50 - 100\ m$

(ii) A layer of constant θ

(iii) A positive jump in θ atop the constant-θ layer

(iv) A layer of stable air, $\frac{d\theta}{dz} > 0$, above the jump

Water vapor profile:

(i) In the lowest 50–100 m, $\frac{dq}{dz} < 0$

(ii) Layer of (nearly constant) q

(iii) A negative jump in q atop the constant-q layer

(iv) A lapse of water vapor above the jump

The thermodynamic profiles conform to mixed-layer model theory as discussed in the review paper by Tennekes and Driedonks (1981). These standard profiles are diagrammed in Fig. 7 of Lewis et al. 2016.

b. Governing equations

The dynamical model used in this study follows from classic mixed-layer model theory. The first use of this mixed-layer model to forecast return flow is found in Liu et al. (1992). The trajectory of the low-level air shown in Fig. 1 defines the Lagrangian framework where speed and direction along the trajectory came from operational surface analyses at National Hurricane Center (NHC). From copies of these operational analyses at 6-h intervals, streamlines and isotachs were drawn by hand, and then following methodology found in Saucier (1955; section 10.06), the trajectory shown in Fig. 1 was created.

[5] Two field programs were conducted during Project GUFMEX (GUFMEX I in February/ March 1988 and GUFMEX II in March 1991).

[6] θ is expressed in °C relative to Kelvin temperature 273.15 K (= 0°C) throughout the paper. Thus, a potential temperature of 283.15 K is expressed as 10°C.

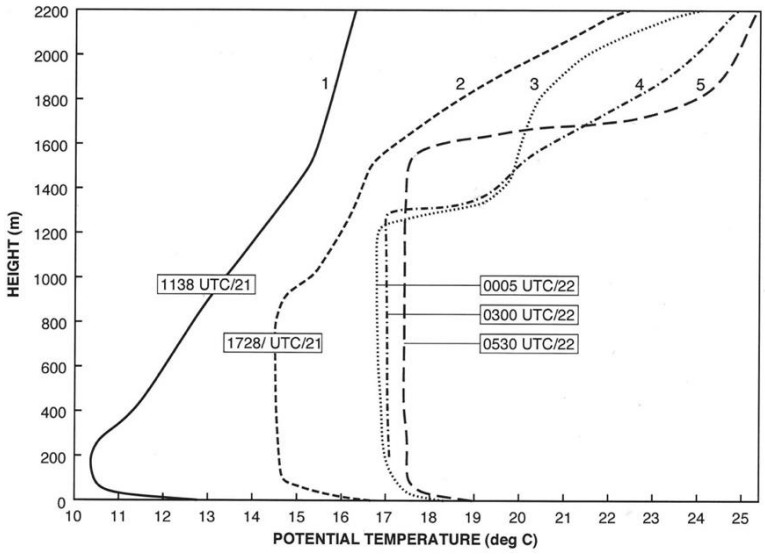

Figure 3: Vertical profiles of θ over the GoM (in °C relative to 273.5 K) derived from rawinsonde launches aboard U. S. Coast Guard ship Salvia at roughly 6-h intervals, following outflowing continental air along a near-Lagrangian trajectory in February 1988 (same as Fig. 5 in Lewis 2007).

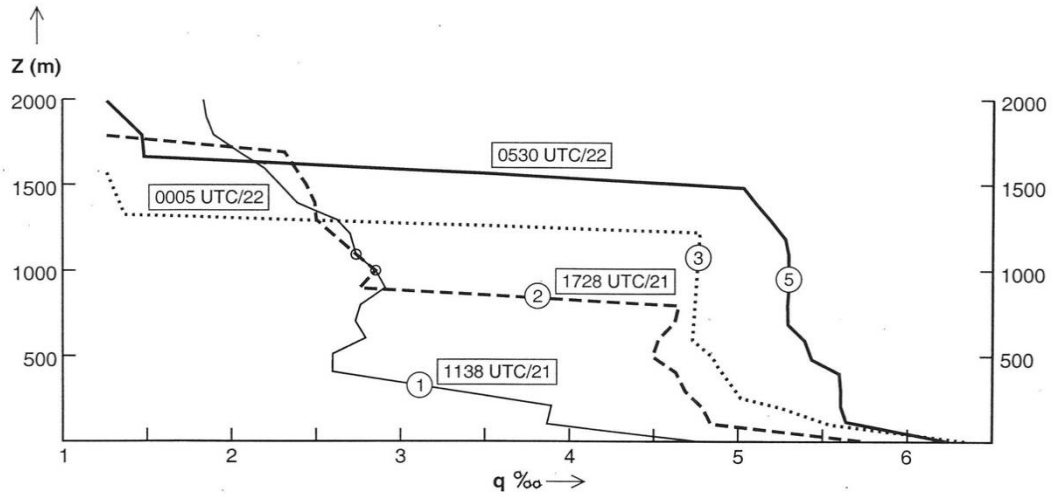

Figure 4: Vertical profiles of water vapor mixing ratio (expressed in g kg^{-1}) derived from the same data mentioned in the Fig. 3 caption (same as Fig. 6 in Lewis 2007).

There are five variables in the mixed-layer model: the constant potential temperature and constant water vapor mixing ratios (θ, q), respectively, jumps in potential temperature and mixing ratio atop the mixed layer (σ, μ), respectively, and the height of the mixed layer (H). Turbulence at the interface between the mixed layer and the stratified layer is accounted through an entrainment velocity w_e. The fluxes of heat and water vapor from the ocean to atmosphere, the driving forces within the model, are represented by $\overline{w'\theta'}]_s$ and $\overline{w'q'}]_s$, respectively, where subscript s refers to the sea surface. Thermodynamic structure as measured by rawinsondes generally agree with idealizations except for observed jumps. These jumps are smooth functions of height as opposed to strict discontinuities.

Equations that define time variation of variables along the Lagrangian trajectory follow:

$$\frac{d\theta}{dt} = (\overline{w'\theta'}]_s + w_e\sigma)H^{-1} \quad (1)$$

$$\frac{dH}{dt} = w_e + \bar{w} \quad (2)$$

$$\frac{d\sigma}{dt} = \gamma_\theta\left(\frac{dH}{dt} - \bar{w}\right) - \frac{d\theta}{dt} \quad (3)$$

$$\frac{dq}{dt} = (\overline{w'q'}]_s + w_e\mu\ )H^{-1} \quad (4)$$

$$\frac{d\mu}{dt} = \gamma_q\left(\frac{dH}{dt} - \bar{w}\right) - \frac{dq}{dt} \quad (5)$$

with entrainment velocity

$$w_e = \kappa\overline{w'\theta'}]_s\sigma^{-1}, \quad (6)$$

where κ is the empirically determined entrainment coefficient—typically a small fraction such as 0.2 or 0.3. The vertical gradients of potential temperature and water-vapor mixing ratio in the stratified layer, $\gamma_\theta > 0$ and $\gamma_q < 0$, respectively, are determined from rawinsonde observations (Figs. 3 and 4), and $\bar{w}$ is the large-scale subsidence determined from wind observations over the Gulf as shown in Fig. 10 of Lewis et al. 2016.

To achieve closure, fluxes are parameterized in terms of a first-order turbulent transfer processes where transfer coefficient C_T is used to represent heat and water vapor transfer as found in Panofsky and Dutton (1984):

$$\overline{w'\theta'}]_s = C_T V_s\,(\theta_s - \theta) \quad (7)$$

$$\overline{w'q'}]_s = C_T V_s\,(q_s - q) \quad (8)$$

where V_s is the translation speed along the trajectory and $(\theta_s - \theta)$ and $(q_s - q)$ are differences between saturated values of potential temperature and water vapor mixing ratio at the air/sea interface and the corresponding constant values of θ and q within the mixed layer.

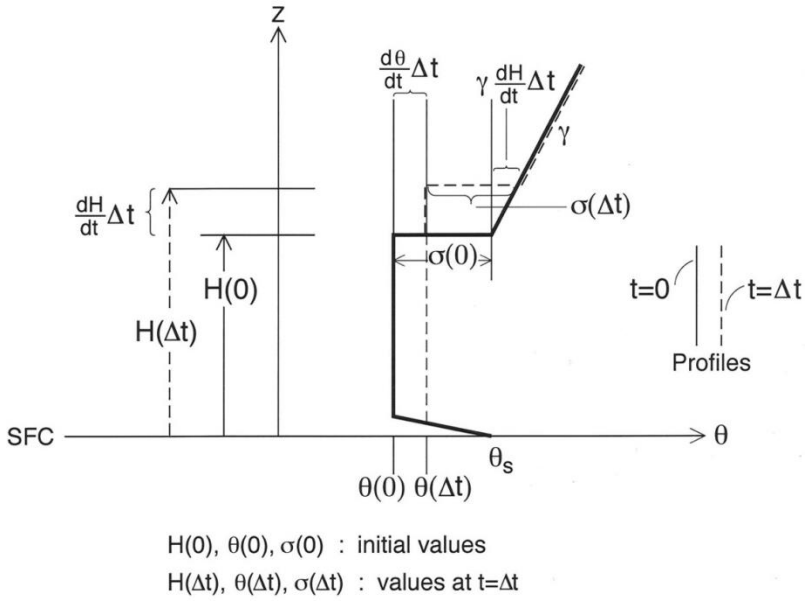

Figure 5: Geometry of changes in θ jump.

The physical interpretation of these mixed-layer equations is based on fluxes of heat and water vapor from sea to air, and entrainment of potential temperature and water vapor mixing ratio from the stratified layer into the mixed layer. Equations governing the change of potential temperature and water vapor mixing ratio, Eqs. (1) and (4), are quite similar—increases due to fluxes from sea to air, and increases (in the case of potential temperature) and decreases (in the case of water vapor mixing ratio) in response to entrainment from the stratified layer. The changes in these two variables are inversely proportional to the height of the layer. The temporal change in height of the mixed layer (Eq. 2) is governed by the interaction between large-scale subsidence (negative vertical velocity) and the entrainment velocity (positive).

The jumps in temperature and water vapor are best understood via Fig. 5. If height increases during the time interval $t = 0$ to Δt, then $w_e > -\overline{w} > 0$ and σ increases due to positive γ_θ. If θ increases during this time interval, then σ decreases. The same line of argument holds for the change in water-vapor mixing ratio jump. The main difference is that an increase of height leads to a decrease of water vapor mixing ratio due to the entrainment process.

The typical values of parameters in the mixed-layer equations are based on thermodynamic structures from rawinsondes for the 21–24 February 1988 case study and values found from the literature. These values follow:

$$C_T = 10^{-3}$$

$$C_T V_s = 1.5 \times 10^{-2}\ m\ s^{-1}$$

$$\gamma_\theta = 6.6 \cdot 10^{-3}\ ℃\ m^{-1}$$

$$\gamma_q = -10^{-3}\ g\ kg^{-1}\ m^{-1}$$

$$\overline{w} = -(0.5 — 1.0)\ cm\ s^{-1}$$

$$\kappa = 0.2 – 0.3 \qquad (9)$$

The literature used to estimate these parameters follows: (1) C_T, Stage and Businger (1981) and Kondo (1975), and (2) κ , Tennekes and Driedonks (1981). The model integration begins at 1800 UTC 21 February, identified by t = 0 along the trajectory shown in Fig. 1. The initial conditions for variables in the mixed layer model are taken from the 1728 UTC 21 February 1988 rawinsondes (Figs. 3 and 4), and given by

$$\theta = 14.5\ ℃$$

$$H = 0.90\ \text{km}$$

$$\sigma = 0.5\ ℃$$

$$q = 4.5\ \text{g kg}^{-1}$$

$$\mu = -1.50\ \text{g kg}^{-1} \qquad (10)$$

The boundary conditions along the trajectory are θ_s and q_s at the air/sea interface. These fields depend on sea surface temperature and surface pressure. The sea surface temperatures come from SSEC's weekly averaged temperatures shown in Fig. 1. The surface pressure comes from NHC's operational sea-level analysis available at 6-h intervals.

4. Data assimilation

a. Identical twin strategy: True and forecasted states

We assume systematic error is present in the dynamical model due to uncertainty in control parameters that govern turbulent transfer of heat and water vapor from sea to air—in particular, the turbulent exchange coefficient C_T and the entrainment parameter κ. To allow a variable exchange coefficient, we represent the exchange coefficient as $C_T = c \cdot 10^{-3}$ where c is a nondimensional factor the order of 1. We also assume systematic error in the forecast is due to uncertainty in the control element $\overline{w}$, the large-scale subsidence. However, we assume that this element of control will not be corrected. This simulates the case where a particular element of control is not suspected of error, and therefore the data assimilation process does not attempt to correct it. In this case, the data assimilation process will make corrections to κ and c, but not to $\overline{w}$. Even if κ and c are perfectly corrected, the forecast will contain error due to the uncorrected $\overline{w}$.

Our approach follows "identical twin" methodology where two differing states are defined. The controls for the two states differ only in the specification of c, κ, and $\overline{w}$. That is, the other elements of control (initial conditions, boundary conditions, and other physical/ empirical parameters) are identical for the two states—one labeled "truth" and the other labeled "forecast". Observations are created from "truth". The true state assumes $(\kappa, c, \overline{w}) = (0.3, 1.2, -0.6)$ and the forecast state assumes $(\kappa, c, \overline{w}) = (0.2, 1.0, -0.5)$. The subsidence parameter $\overline{w}$ is nondimensionalized with a scale equal to $1\ cm\ s^{-1}$. Thus, $\overline{w} = -0.6$ and -0.5 represent a subsidence of 0.6 cm s^{-1} and 0.5 cm s^{-1}, respectively, for the true /observed state and the forecast state.

b. Observations created from true state

At 6-h intervals along the Lagrangian trajectory (starting at t = 0), observations are created by integrating the governing equation with true control. The observation vector $\mathbf{Z}(t)$ at each time is given by $\mathbf{Z}(t) = (Z_\theta(t), Z_H(t), Z_q(t))^T$, observations of the potential temperature, height of the mixed

layer, and water vapor mixing ratio, respectively. The jumps in temperature and mixing ratio are not included in the observation vector since large uncertainty exists for these variables. The forecast vector $\mathbf{F}(t) = (\theta(t), H(t), q(t))^T$, forecasts of potential temperature, height of the mixed layer, and water vapor mixing ratio, respectively, is determined by integrating the governing equations with forecast-state controls. From vectors $\mathbf{F}(t)$ and $\boldsymbol{Z}(t)$, the forecast error or innovation $\mathbf{I}(t) = \mathbf{Z}(t) - \mathbf{F}(t)$ is known. The sum of the squares of the three components of $\mathbf{I}(t)$ constitutes the cost function at the particular time. Data assimilation seeks to minimize this cost function by optimally adjusting the two forecast controls $(\kappa, c) = (0.2, 1.0)$.

c. Forecast Sensitivity Method (FSM) of data assimilation

The variational data assimilation process used in this study is labeled Forecast Sensitivity Method (FSM) (Lakshmivarahan and Lewis, 2010; Lakshmivarahan, et al., 2017, Ch. 2). The development of this method as applied to return-flow dynamical constraints is found in the appendix. Since the cost function associated with FSM and other optimal methods of assimilation typically involve physical variables with differing units, it is important to put variables in nondimensional form such that they have the same order of magnitude. In this case, scales 10 ℃, 10 g kg^{-1}, and 1 km are used for θ, q and H, respectively. This scaling delivers values the order of 1 for each of the variables. Graphical display of results will be presented in nondimensional form unless otherwise specified.

5. Results from the numerical experiments

a. Forecast errors and forecast sensitivities

The observed and forecased states of $\theta, H,$ and q are displayed in Fig. 6. The difference between observed and forecasted states for θ and H are largest around t ≈20 h, while the largest differences between observed and forecasted states for q are beyond 20 h. Forecast errors for θ and H are relatively small near t = 0 and beyond t = 40 h whereas the forecast errors for q are small up through t = 10 h.

The turbulent heat flux is proportional to the difference between forecasted θ and the boundary value θ_s while the turbulent moisture flux is proportional to the difference between forecasted q and the boundary value q_s. These difference fields are shown in Lewis et al. (2016)—Fig.8 (dimensional form). From this figure, convective forcing clearly vanishes beyond t ≈50 h while vapor flux forcing is still active beyond t ≈50 h although relatively small compared to maximum forcing near t ≈10 h. In the absence of significant thermal forcing beyond t ≈36 h, near the southmost point along the trajectory, the experiment will be restricted to the time interval between t = 0 to t = 36 h—the active penetrative convection period.

It is also important to examine the forecast sensitivity fields generated through solution to the augmented dynamical equations. These sensitivities are displayed in Figs. 7 and 8—Fig. 7 exhibiting forecast sensitivity to κ and Fig. 8 exhibiting forecast sensitivity to c. All forecast sensitivities with respect to the two parameters are positive except for the forecasted q with respect to κ. This negative sensitivity can be understood through examination of Eq. (4), the equation governing time evolution of q. In this equation we note that an increase in κ leads to an increase in moisture flux from the surface, but the same increase in κ increases the negativity of the second term in this equation, $w_e \mu H^{-1}$ where this second term outweighs the first term. However, in the case of forecasted q sensitivity to c, the first term (positive) outweighs the second term (negative) and leads to a positive value of forecast sensitivity to c.

The fields of sensitivity generally reach a "saturation point", that point where sensitivity is relatively constant. In effect, the sensitivity functions accumulate the action of the fluxes up through t = 36 h (period of active convection), but beyond that point forcing is small to nonexistent. For the height sensitivity fields, there comes a point where subsidence dominates surface forcing and height sensitivity begins to decrease.

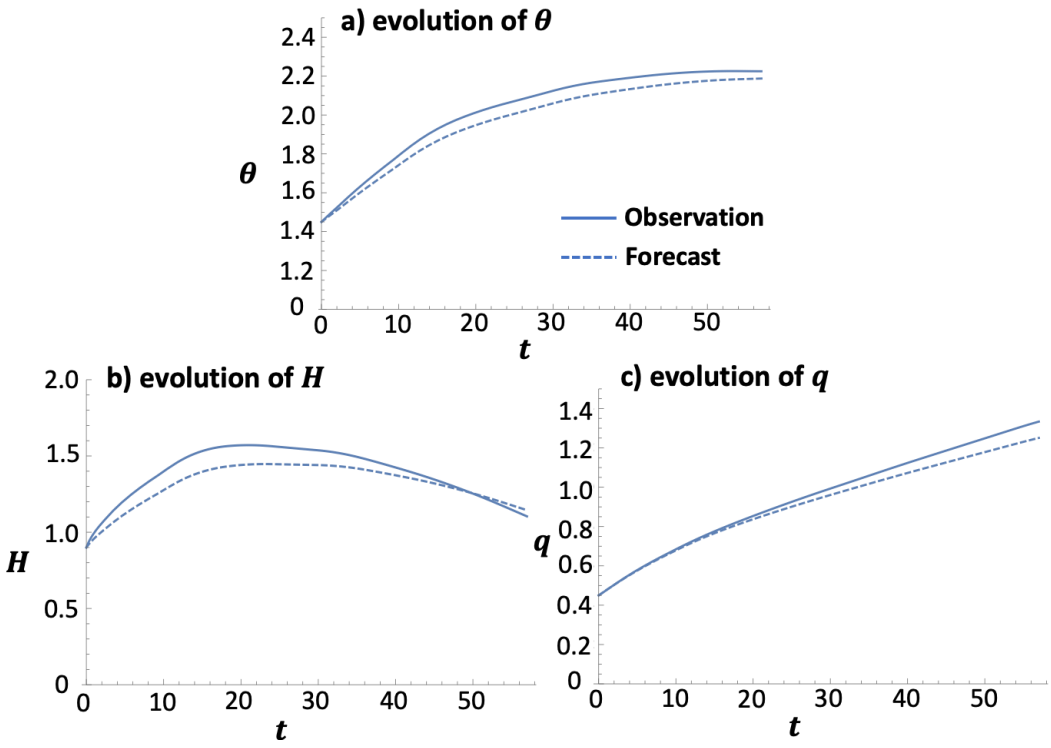

Figure 6: The evolution of the observation field (solid lines) and the forecasted field (dashed), over the 57-h period defined by the trajectory.

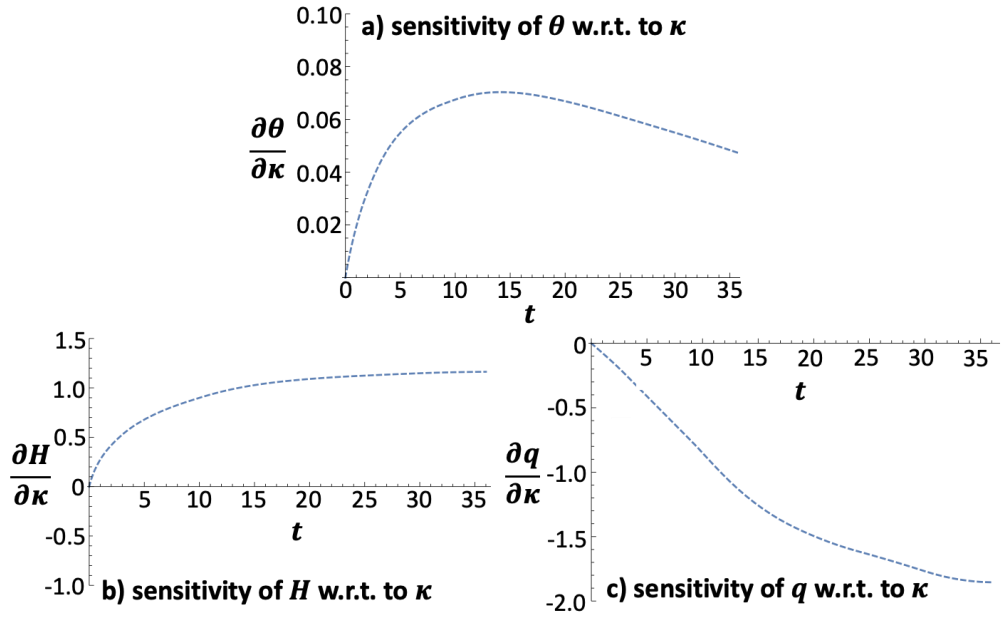

Figure 7: The sensitivity of forecasts $\theta, H, and\ q$ with respect to κ.

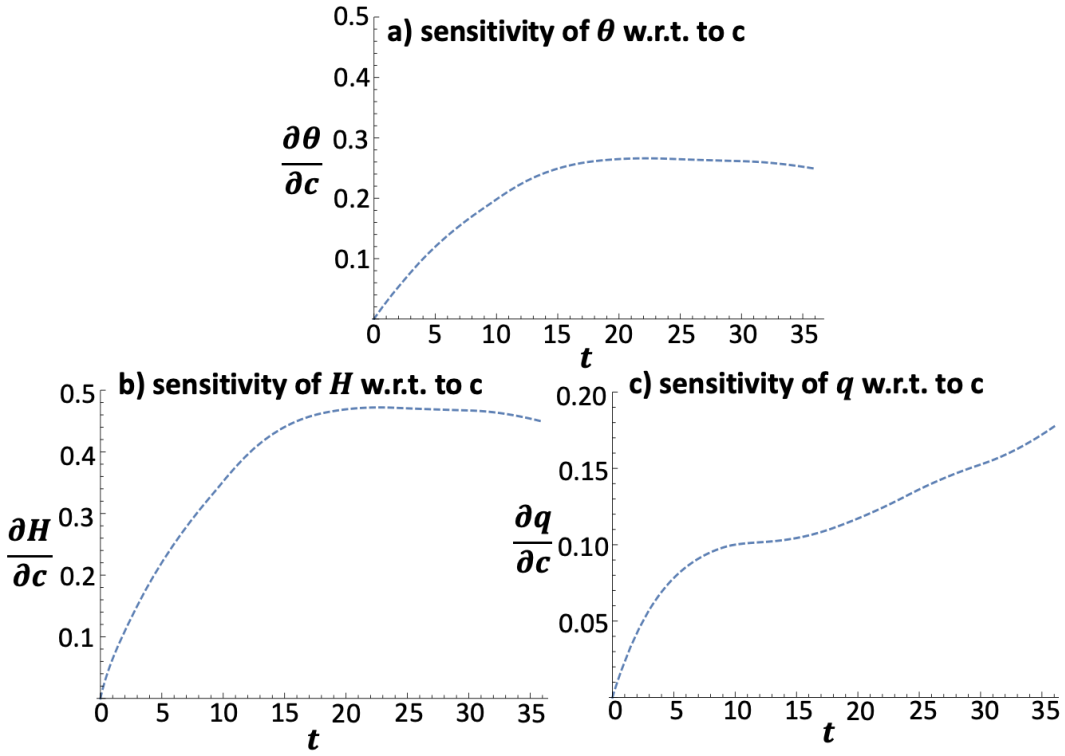

Figure 8: The sensitivity of forecasts $\theta, H, and\ q$ with respect to c.

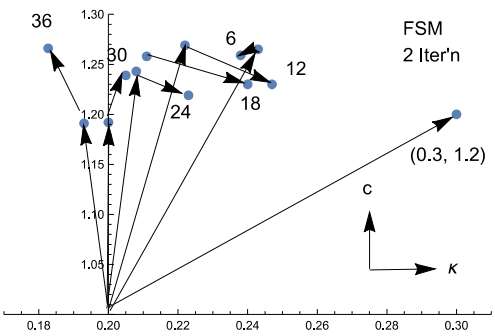

Figure 9: Adjustment of controls (κ, c) after two iterations of the FSM data assimilation scheme using only those observations at separate times t= 6, 12, …, 36.

b. Observation placement

Adjustments to control are made at 6-h increments along the trajectory at t = 6, 12, 18, 24, 30, and 36 h—that portion of the trajectory where convective forcing is active. The process adjusts control at each 6-h increment separately. Once the adjustment is made, these values of control are used to integrate the governing equations over the time period t = 0 to t = 57 h. Two iterations are made at each time. In all cases, the initial operating point in the space of control is $(\kappa, c) = (0.2, 1.0)$. Adjustments to control at each time are shown in Fig. 9 (two iterations). The adjustments are viewed relative to the initial operating point. In this same figure,

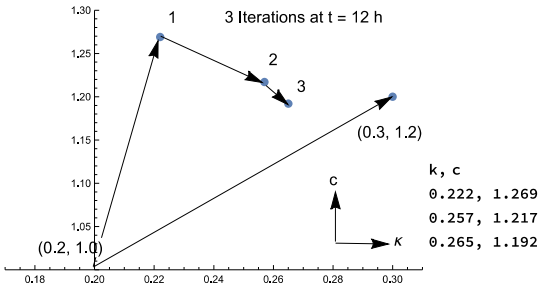

Figure 10: Adjustment of controls (κ, c) after 3 iterations using observations at t = 12 h.

we identify the true control $(\kappa, c) = (0.3, 1.2)$ and draw a vector from the initial operating point to true control. This vector serves to graphically indicate the best path from the initial operating point to true control. Adjustments associated with t = 12 h deliver the best results, an adjusted control closest to the true control. One further iteration for the t = 12 h adjustment brings it slightly closer to true control as shown in Fig. 10. The iterative process is stopped when further

iterations produce only small changes in the value of the cost function. The termination point is generally associated with flatness of the cost function in the space of control.

The quantitative aspects of the iterative data assimilation process are shown for adjustment at two times (Table 1), t = 12 h and t = 36 h—the best result at t = 12 h and the poorest result at t = 36 h. The relative smallness of the condition number for t = 12 h is also an indication that adjustment at this time will give better results. The condition number as a function of time between t = 0 and t = 36 is displayed in Fig. 11.

Having determined that the best results come from data assimilation at t = 12, we plot the evolution of the mixed-layer variables using the adjusted controls based on 4 iterations at this time. These controls are inserted into the dynamical equations to make the forecast. This forecast is compared to the observations in Fig. 12. Note that the fit of height to observations at later times is not as good as the fit at earlier times. This is because height is strongly sensitive to subsidence, and the observed structure is associated with larger subsidence than forecasted.

As time approaches t = 0, the condition number becomes exponentially large and essentially warns against making observations close to t = 0. At this point in time, the forecast errors become extremely small as do the sensitivity fields. If initial conditions were among the controls to be adjusted, the condition number would not become extremely large near t = 0 since the initial analysis error (incorrect initial condition) and forecast sensitivity to initial conditions would be nonzero at t = 0.

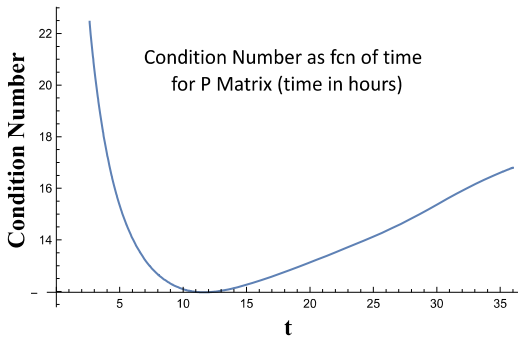

Figure 11: Time variation of condition number from the P matrix that was used in FSM assimilation.

Table 1: Stepwise correction of control. Symbols and terms as described in the text.

t = 12 h

Iteration	Control		$-\nabla J$		J	Condition Number
	κ	c	$-\nabla_\kappa J$	$-\nabla_c J$		
Start	0.200	1.000	0.1225	0.0648	0.0101	12.0
1	0.222	1.269	0.0445	0.0003	0.0006	17.6
2	0.248	1.229	0.0171	–0.0011	0.0002	15.3
3	0.259	1.202	0.0065	–0.0004	0.0000(8)	14.1
4	0.265	1.191	0.0024	–0.0002	0.0000(8)	13.7

t = 36 h

Iteration	Control		$-\nabla J$		J	Condition Number
	κ	c	$-\nabla_\kappa J$	$-\nabla_c J$		
Start	0.2000	1.000	0.0054	0.0550	0.0054	16.8
1	0.1930	1.191	0.0476	0.0006	0.0004	25.9
2	0.2021	1.1891	0.0257	–0.0024	0.0004	24.2
3	0.2077	1.1739	0.0111	0.0010	0.0004	25.2

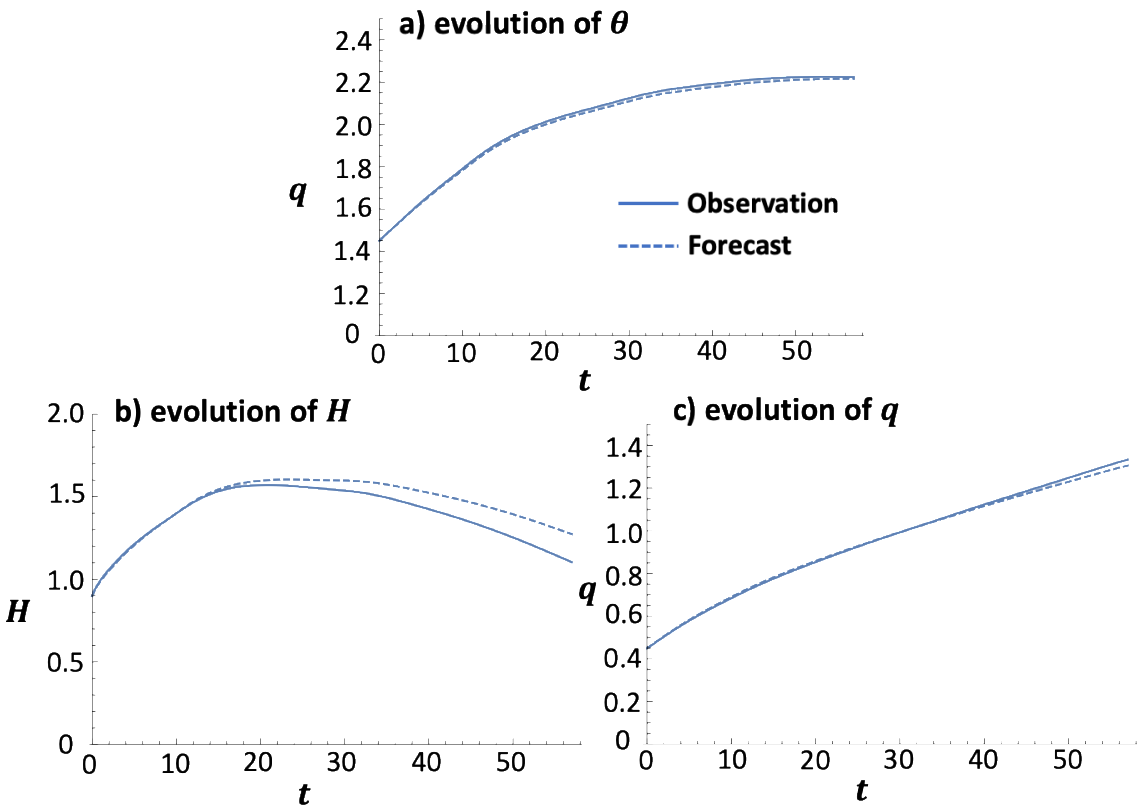

Figure 12: Comparison of observation (the true state) and forecasted state based on the best adjustments at t = 12 h.

From discussion in David Blackwell's classic undergraduate textbook, *Basic Statistics* (Blackwell, 1969, Ch. 6), we ask the question: "What is the *worth of the predictor(s)*?". Or alternately stated, what is the value of adjusting controls to remove the systematic bias in the forecast? The worth W is given by

$$W(v) = \frac{SE(v) - SE_{adj}(v)}{SE(v)} \qquad (11)$$

where v is one of the variables $(\theta, H, or\ q)$, $SE(v)$ is the squared error of the initially prescribed forecast, and $SE_{adj}(v)$ is the squared error in the adjusted forecast. If $SE_{adj}(v) = SE(v)$, then worth of the predictor $W = 0$, but if $SE_{adj}(v) = 0$ (forecast and observations coincide), then worth $W = 1$. Using $SE(v)$ and $SE_{adj}(v)$ from the unadjusted forecasts and observations (shown in Fig. 6) and the adjusted forecast and observations (shown in Fig. 12), we find $W(\theta) = 0.96,\ W(H) = 0.82, and\ W(q) = 0.92$. The good-fit adjusted forecasts (visual comparison of Figs. 12 with Fig. 6) are confirmed with these quantitative measures.

6. Discussions and conclusions

A low-order model that is faithful to the evolution of boundary layers in return-flow episodes has been used to determine ideal observation placement over the GoM. Through data assimilation, the ideal observation placement leads to model-control correction that in turn serves to improve the forecast plagued by systematic error.

Results from the numerical experiment indicate that optimal observation placement is over the area of greatest flux of heat and moisture from sea to air: continental air passing over the Loop Current. Observations at times long after the modified air leaves these warmest waters lead to relatively poor adjustments to control. A combination of observations (i.e., observations of more than a single boundary-layer variable), also leads to improved results. This fact was uncovered by examination of forecast sensitivity to elements of control, a feature of the FSM data assimilation scheme.

Examination of ocean-based rawinsonde observations as shown in Figs. 3 and 4 during the outflow phase, and the corresponding rawinsonde observations during the return-flow phase of this same event (Fig. 7 of Lewis and Crisp 1992), makes it clear that observations showing this detail in thermodynamic-structure would go far to improve operational return-flow forecasts. However, in this age of emphasis on remote sensing, especially over the world sea, there is little chance that this will happen through the use of rawinsondes.

Following valued discussions with the NCEP/EMC data assimilation group (D. Kleist and C. Thomas 2020, personal communication) and satellite meteorologists (S. Kidder and W. Smith 2020, personal communication), the best chance for getting boundary layer thermodynamic structures comparable to those from rawinsondes will eventually come from satellite measurements of blended total precipitable water (bTPW) and from retrieved temperature/vapor profiles from hyperspectral radiance instruments (now on polar orbiters like Suomi NPP and NOAA-20). Theoretical limits for vertical resolution, temperature, and humidity from the hyperspectral radiance instruments—1–3 km, 1–2 K, and 10–20%—have yet to be achieved, but the active research projects designed to do so have been detailed in Weisz et al. (2020).

The online bTPW at https://www.star.nesdis.noaa.gov/portfolio/detail_bTPW.php, produced daily at the NOAA STAR Center for Satellite Applications and Research and within NESDIS, , also holds promise for delivering quality estimates of water vapor mixing ratio. The strengths (and weaknesses) of TPW measurements from various satellites have been accounted in this blended product. In the return-flow situations, water in the mixed layer is water vapor with only a slight amount to spray near the surface, so the bTPW would be a measure of the water-vapor mixing ratio. Further, since the air is typically dry above the mixed layer, higher levels contribute little to the vertically integrated TPW.

This research has paid strict attention to a typical yet singular return flow event. We know of significant variance in return flows, as clearly presented by Henry (1979) and in the climatological study by Crisp and Lewis (1992). Thus, some caution is necessary in assuming that the results in this study apply to all return flow episodes. Yet, it's safe to assume that satellite measurements of the mixed-layer temperature, water vapor mixing ratio, and height of the mixed layer, or some combination of these variables, in the vicinity of those places where the cold and dry continental airmass traverses the warmest SSTs will go far to improve operational forecasts of return flow. Further, as found from Maddox's study of recent operational forecasts of return flow (Maddox 2016), the trajectories appear to be forecasted well, and that would allow identification of locations where the continental air will pass over those warmest waters. Thus, maybe 12–24 h before the cold/dry air enters the Gulf, efforts can be made to make satellite observations at these optimal locations.

ACKNOWLEDGMENTS

Conversations with Rich Thompson and Roger Edwards at an early stage of thinking about the research went far to help design the project. The following NCEP/EMC meteorologists supplied valued information about data assimilation and problems with operational forecasts of return flow: Zavisa Janjic (posthumously), Geoff Manikin, Daryl Kleist, and Cristine Thomas. Input on satellite data from Stan Kidder and Bill Smith helped us understand the bTPW and hyperspectral products, respectively.

Three formal reviewers appointed by EJSSM editorial staff put great effort into their reviews that went far to strengthen arguments and identify weaknesses in the presentation. Informal reviews by Pam Heinselman and Jack Kain helped iron out confusing statements in an early version of the manuscript, and NSSL draftsperson Joan O'Bannon expertly drafted several of the figures.

Support was provided by NOAA/Office of Oceanic and Atmospheric Research under the NOAA-University of Oklahoma Agreement NA111OAR430072, U. S. Department of Commerce.

APPENDIX:
Mechanics of data assimilation

A1. Sensitivity calculation

The Forecast Sensitivity Method (FSM) requires knowledge of the forecast sensitivities to κ and c. With five variables in the mixed layer model and two elements of control that will be adjusted, there are ten forecast sensitivities, five associated with κ and five associated with c. These sensitivities are represented as vectors

$$\mathbf{S_\kappa} = (\frac{\partial\theta}{\partial\kappa}, \frac{\partial H}{\partial\kappa}, \frac{\partial q}{\partial\kappa}, \frac{\partial\sigma}{\partial\kappa}, \frac{\partial\mu}{\partial\kappa})^T \text{ and}$$

$$\mathbf{S_c} = (\frac{\partial\theta}{\partial c}, \frac{\partial H}{\partial c}, \frac{\partial q}{\partial c}, \frac{\partial\sigma}{\partial c}, \frac{\partial\mu}{\partial c})^T.$$

The equations governing these sensitivity variables are found by differentiating the basic set of mixed layer equations with respect to κ, yielding five equations, and with respect to c, delivering another set of five equations. Thus, the coupled set of equations used in this study involves the basic set and the ten sensitivity equations, a coupled set of fifteen ordinary differential equations. To give some substance to the process, the derivation of the forecast sensitivity of θ with respect to κ follows:

$$\frac{\partial}{\partial\kappa}\left(\frac{d\theta}{dt}\right) = \frac{d}{dt}(\frac{\partial\theta}{\partial\kappa}) = \frac{\partial}{\partial\kappa}\left[cV_s(1+\kappa)\frac{(\theta_s-\theta)}{H}\right]$$

$$= cV_s\,\frac{(\theta_s-\theta)}{H} + cV_s\,(1+\kappa)\frac{H\frac{\partial(\theta_s-\theta)}{\partial\kappa} - (\theta_s-\theta)\frac{\partial H}{\partial\kappa}}{H^2} \tag{A.1}$$

where c accounts for variability in the transfer coefficient. The time evolution of $\frac{\partial\theta}{\partial k}$ depends on solutions to the basic mixed-layer equations through θ and H, and also dependent on the forecast height sensitivity to κ, $\frac{\partial H}{\partial\kappa}$, which is governed by an equation similar to the one just derived. The initial conditions for each of these ten "augmented" equations are equal to zero. That is, if you change these parameters, the values of the basic variables do not change at the initial time. The sensitivity equations depend on the solution to the basic equations, but not vice versa. And as will be noted below in this appendix, the FSM only requires knowledge of forecast sensitivities associated with the observed variables (θ, H, q), i. e., six of the ten forecast sensitivities — $\frac{\partial\theta}{\partial\kappa}, \frac{\partial H}{\partial\kappa}, \frac{\partial q}{\partial\kappa}, \frac{\partial\theta}{\partial c}, \frac{\partial H}{\partial c}, \frac{\partial q}{\partial c}$. But knowledge of these six sensitivities requires knowledge of all ten sensitivities.

A2. Adjustments to control

To simplify notation, the three forecast variables are represented as a column vector:

$$\mathbf{X}(t) = \begin{bmatrix}\theta(t)\\ H(t)\\ q(t)\end{bmatrix} = \begin{bmatrix}x_1(t)\\ x_2(t)\\ x_3(t)\end{bmatrix} \tag{A.2}$$

Further, we represent the observations of these variables, in the same order as the forecasts, as

$$\mathbf{Z}(t) = \begin{bmatrix}z_1(t)\\ z_2(t)\\ z_3(t)\end{bmatrix} \tag{A.3}$$

The cost function J is defined as the one-half the sum of squared differences between forecast and observations:

$$J = \tfrac{1}{2}[\mathbf{X}(t) - \mathbf{Z}(t)]^T[\mathbf{X}(t) - \mathbf{Z}(t)] \tag{A.4}$$

where $(\ldots)^T$ represents the transpose of a vector or matrix. J is an implicit function of the variable control parameters κ and c through the forecasts.

The unique characteristic of FSM is determination of adjusted control parameters through minimization of the cost function that depends on knowledge of forecast sensitivities to κ and c. The fundamental assumption is that adjusted forecasts, denoted by $\widehat{[\ldots]}$, take the following form based on a first-order Taylor expansion about the current estimate of control:

$$\hat{x}_1\,(t) = x_1\,(t) + \frac{\partial x_1\,(t)}{\partial\kappa}\Delta\kappa + \frac{\partial x_1(t)}{\partial c}\Delta c \tag{A.5}$$

$$\hat{x}_2\,(t) = x_2\,(t) + \frac{\partial x_2\,(t)}{\partial\kappa}\Delta\kappa + \frac{\partial x_2(t)}{\partial c}\Delta c \tag{A.6}$$

$$\hat{x}_3\,(t) = x_3\,(t) + \frac{\partial x_3\,(t)}{\partial\kappa}\Delta\kappa + \frac{\partial x_3(t)}{\partial c}\Delta c \tag{A.7}$$

Or in matrix form

$$\widehat{\mathbf{X}}(t) = \begin{bmatrix}\hat{x}_1\,(t)\\ \hat{x}_2\,(t)\\ \hat{x}_3\,(t)\end{bmatrix} =$$

$$\mathbf{X}(t) + \begin{bmatrix}\frac{\partial x_1(t)}{\partial\kappa} & \frac{\partial x_1(t)}{\partial c}\\ \frac{\partial x_2(t)}{\partial\kappa} & \frac{\partial x_2(t)}{\partial c}\\ \frac{\partial x_3(t)}{\partial\kappa} & \frac{\partial x_3(t)}{\partial c}\end{bmatrix}\begin{bmatrix}\Delta\kappa\\ \Delta c\end{bmatrix} = \mathbf{X}(t) + \mathbf{S}(t)\Delta\mathbb{C} \tag{A.8}$$

where $\mathbf{S}(t)$ is the sensitivity matrix and $\Delta\mathbb{C}$ is incremental-change control vector. The first-order approximation to the cost function takes the form

$$J \approx \frac{1}{2}[\mathbf{X}(t) + \mathbf{S}(t)\Delta\mathbb{C} - \mathbf{Z}(t)]^T[\mathbf{X}(t) + \mathbf{S}(t)\Delta\mathbb{C} - \mathbf{Z}(t)] \quad \text{(A.9)}$$

a quadratic function in the unknown incremental-change control vector. A necessary condition for minimization of the cost function is vanishing of its derivative with respect to the unknown incremental-change control vector. Taking the derivative and setting it equal to zero gives the result

$$\Delta\mathbb{C} = (\mathbf{S}^T\mathbf{S})^{-1}\mathbf{S}^T\mathbf{D} = \mathbf{P}^{-1}\,\mathbf{S}^T\mathbf{D}, \quad \text{(A.10)}$$

where $\mathbf{P} \equiv (\mathbf{S}^T\mathbf{S})$ and $\mathbf{D}(t) = \mathbf{Z}(t) - \mathbf{X}(t)$, difference between observation and forecast (i.e. Innovation)—new information brought out by observation $\mathbf{Z}(t)$ beyond that provided by the forecast $\mathbf{X}(t)$.

By adding the adjustment to the initial value of control, a new control is found. It is used to find a solution to the mixed-layer equations that has reduced the cost function. The process is used in an iterative fashion where the new forecast and new sensitivities and the original observations are used to produce successive estimates of the control. The process is terminated when sequential values of the cost function show little change.

A3. Gradient of the cost function J, $\nabla J = (\nabla_\kappa J, \nabla_c J)^T$, and condition number

The gradient of J in the space of variable control, the $\kappa - c$ space, is an important quantity in all variational data assimilation schemes—3D-Var and 4-D Var schemes as well as FSM as developed and discussed in Lewis et al. (2006) and Lakshmivarahan et al. (2017). At an operating point in this space (current estimate of the variable control), it is important to know the value of each component of this vector, a two-component vector in our case. In minimization of the cost function (a scalar quantity) by steepest descent or descent generally, a "flat" region of J around the operating point presents a problem. In this situation, the pathway to the cost function minimum is ill-defined. In FSM, this problem is identified by the condition number of the **P** matrix that was defined above. The condition number, the ratio of the largest eigenvalue of the matrix to the smaller eigenvalue of **P** matrix, indicates the presence of a flat region when this number is very large (orders of magnitude >1). Thus, in our study, we calculate the condition number for each adjustment to control, but we also calculate the 2-componnet gradient of the cost function. With knowledge of the sensitivities and Innovation, the gradient of the cost functional J in Eq. (A.4) can be found as follows:

$$\nabla J = (\nabla_\kappa J, \nabla_c J)^T = \mathbf{S}^T\big(\mathbf{X}(t) - \mathbf{Z}(t)\big), \quad \text{(A.11)}$$

the product of $\mathbf{S}^T$ (2×3 matrix) and the innovation (3×1 column vector) resulting in the 2×1 gradient vector.

REFERENCES

Blackwell, D., 1969: *Basic Statistics*. McGraw-Hill 143 pp.

Cohen, A. E., S. M. Cavallo, M. C. Coniglio, H. E. Brooks, and I. L. Jirak, 2017: Evaluation of multiple planetary boundary layer schemes in Southeast U. S. cold season thunderstorm environments. *Wea. Forecasting*, **32**, 1857–1884.

Crisp, C., and J. M. Lewis, 1992: Return flow in the Gulf of Mexico, Part I: A classifacatory approach with a global historical perspective. *J. Appl. Meteor*., **31**, 868–881.

Dallavalle, P., and L. Bosart, 1975: A synoptic investigation of anticyclogenesis accompanying North American polar air outbreaks. *Mon. Wea. Rev*., **103**, 941–956.

Henry, W., 1979: Some aspects of the fate of cold fronts in the Gulf of Mexico. *Mon. Wea. Rev.*, **107**, 1078–1082.

Janjic, Z., 1994: The step–mountain coordinate model: Further development of the convective, viscous sublayer, and turbulence closure schemes. *Mon. Wea. Rev*., **122**, 927–945.

——, 2001: Nonsingular implementation of the Mellor-Yamada level 2.5 scheme in the NCEP/EMC meso model. NOAA/NWS/NCEP/EMC Office Note 437, 61 pp.

Johnson, J. R., 1976: The origin, structure and modification of return flow over the Gulf of Mexico. M. S. thesis, Texas A & M University, 61 pp.

Karnavas, G. R., 1978: On polar airmass modification over the Gulf of Mexico during periods of return flow and development of low clouds. M. S. thesis, Texas A & M University, 69 pp.

Kleist, D., R. Mahajan, and C. Thomas, 2018: Data assimilation in the Next-Generation Global Prediction System (NGGPS) Era: Initial implementation of FV3-based Global Forecast System (GFS), *Joint Center for Satellite Data Assimilation*, Fall Quarterly, 1–10.

Kondo, J., 1975: Air-sea bulk transfer coefficients in diabatic conditions. *Bound.-Layer Meteor*., **9**, 91–112.

Lakshmivarahan, S., and J. Lewis, 2010: Forward sensitivity approach to dynamic data assimilation. *Adv. Meteor*., **375615**, 1–12, doi: 10.1155/2010/375615.

——, J. M. Lewis, and Jabrzemski, 2017: *Forecast Error Correction using Dynamic Data Assimilation*. Springer, 270 pp.

Lewis, J. M., 2007: Use of a mixed-layer model to investigate problems in operational predicition of return flow. *Mon. Wea. Rev*., **135**, 2610–2628.

——, and C. A. Crisp, 1992: Return flow in the Gulf of Mexico, Part II: Variability of return flow thermodynamics inferred from trajectories over the Gulf. *J. Appl. Meteor*., **31**, 882–898.

——, C. Hayden, R. Merrill, and J. Schneider, 1989: GUFMEX: A study of return flow in the Gulf of Mexico. *Bull. Amer. Meteor. Soc*., **70**, 24–29.

——, S. Lakshmivarahan, and S. Dhall, 2006: *Dynamic Data Assimilation: A Least Squares Approach.* Cambridge Univ. Press, 654 pp.

——, ——, J. Hu, R. A. Maddox, R. Edwards, R. Thompson, and S. F. Corfidi, 2016: Ensemble forecasting of return flow over the Gulf of Mexico. *Electronic J. Severe Storms Meteor.*, **11** (4), 1–26.

Liu, Q., J. Lewis, and J. M. Schneider, 1992: A study of cold air modification over the Gulf of Mexico using *in situ* data and mixed-layer modeling. *J. Appl. Meteor*., **31**, 909–924.

Maddox, R. A., 2016: Ensemble forecasting of return flow over the Gulf of Mexico: Supplemental material. *Electronic J. Severe Storms Meteor.*, **11** (4), A1–A10.

Molina, M. J., and J. T. Allen, 2019: On the moisture origins of tornadic thunderstorms. *J. Climate*, **32**, 4321–4346.

Molinari, R. L., 1987: Air mass modification over the eastern Gulf of Mexico as a function of surface wind fields and Loop Currrent position. *Mon. Wea. Rev*., **115**, 646–652.

Palmén, E., and C. Newton, 1951: On the three-dimensional motions in an outbreak of polar air. *J. Meteor*., **8**, 25–39.

Panofsky, H., and J. Dutton, 1984: *Atmospheric Turbulence: Models and Methods for Engineering Application*s. John Wiley and Sons, 397 pp.

Saucier, W. J., 1955: *Principles of Meteorological Analysis*. University of Chicago Press, 438 pp.

Stage, S., and J. Businger, 1981: A model for entrainment into a cloud-topped marine boundary layer. Part I: Development of a model and application to a cold air outbreak episode. *J. Atmos. Sci.*, **38**, 2213–2229.

Tennekes, H., and A. Driedonks, 1981: Basic entrainment equations for the atmospheric boundary layer. *Bound.-Layer Meteor*., **20**, 515–529.

Weisz, E., N. Smith, and W. L. Smith, 2015: The use of hyperspectral sounding information to monitor atmospheric tendencies leading to severe storms. *Earth Space Sci.*, **2**, 369–377, doi: 10.1002/2015EA000122.

REVIEWER COMMENTS

[Authors' responses in *blue italics.*]

REVIEWER A (Lewis D. Grasso):

Initial Review:

Recommendation: Accept with minor revisions.

Substantive Comments: End of section 4.2, the Innovation is defined as $\mathbf{Z}(t) - \boldsymbol{F}(t)$, that is obs minus forecast. Yet in the beginning of section 5, Innovations are advertised as being plotted in Fig. 7. However, the figure caption for Fig. 7 states "…forecast-minus-observations…"; that is, $\boldsymbol{F}(t) - \boldsymbol{Z}(t)$. Please clarify.

"Innovation" (difference between forecast and observation) is standard terminology for the data assimilation community, but it is generally unfamiliar to those outside this community. Indeed, it is appropriate to introduce this word alongside the more familiar term "forecast error". Further, the discussion of Fig. 7 has been changed to avoid this confusion.

Section 5a: Innovations are stated as being "small" and "very small". Perhaps you would be willing to quantify those two terms. One risk of using "small" and "very small" is that one might say, "Well, if the difference between a forecast and observations is small or very small, then the forecast does not need to be improved."

Point well taken. In first paragraph, discussion has been modified slightly to avoid the interpretation that "forecast does not need to be improved". The nondimensional form of the variables also contributes to the sense of very small differences.

End of paper: Results from one experiment that used one trajectory; that is, you have a sample size of one. From that one sample, you then provide guidance for the general population of all trajectories for all return flow situations. As a way to alert readers about pitfalls of making inferences from small sample sizes, would you be willing to add some text to inform readers of a cautionary note—you pick the word/phase if you don't like "cautionary note"—that the paper contains one experiment and one trajectory and the ability to make inferences about general return flows?

How true a statement! To caution the reader, [a cautionary note was added].

[Minor comments omitted…]

Second Review:

Recommendation: Accept.

General Comment: All of my comments/suggestions have been acceptably addressed; hence, I accept this manuscript.

REVIEWER B (William A. Gallus Jr.):

Initial Review:

Recommendation: Accept with minor revisions.

I believe more context is needed in your last paragraph. You do a good job in the paper making the case for where it would be most important to have observations, and the discussion seems to imply this is the area where upper air observations should be made. You then end up concluding that only satellites

currently provide what is needed. But I am confused because I thought satellite data is already being incorporated for DA, and since satellite data cover large sections of the planet "equally" (meaning the satellite does not have spotty coverage related to land or sea presence), wouldn't these data already be getting used over the regions where the continental air flows over the Loop Current? I get the impression maybe the satellite measurements that you talk about are not yet being used in DA? If this is the case, you should make that clearer in this discussion.

We have received a most heartening response from Daryl Kleist, chief of the data assimilation and quality control unit at NCEP, and one of the scientists working in that division, Cathy Thomas. We have included [a] statement to bring the readership up-to-date on work at NCEP related to inclusion of satellite data in GSI.

[Minor comments omitted…]

Second Review:

Recommendation: Accept with minor revisions.

General Comment: This article has been substantially improved by the changes made by the authors. I only have one remaining minor concern.

[Minor comments omitted…]

REVIEWER C (Ariel E. Cohen):

Initial Review:

Recommendation: Accept with minor revisions.

Specific Comments: The need to reference Manikin is outstanding and very well-articulated, and I'm really appreciative of the incorporation of this direct statement. I feel that the combination of this statement, and the Molina and Allen publication provides some substantiation and motivation for the operational need to evaluate and improve moisture-return forecasts. However, I think a bit more motivation would be helpful to further motivate just how important it is. I do recommend referencing Cohen et al. (2007) that specifically highlights the sensitivities of boundary-layer moisture profiles within the return-flow regimes and warm sectors of extratropical cyclones during the cold season (i. e., low-CAPE, high-shear regimes). The support of severe convection in this regime often benefit greatly from dynamical perturbation pressure processes resulting from updraft-shear interactions, in many cases where boundary-layer moisture is only partially modified, as opposed to reaching complete equilibrium. It is in these environments where the low-CAPE nature results in substantial difference in severe-thunderstorm-threat assessment for marginal differences in PBL thermodynamic structure as addressed by Cohen et al. (2017), further motivating the need to improve return-flow processes. I think addressing the present work in the context of this motivator may be useful from a big-picture perspective—not to mention other return-flow regimes supporting convective weather.

Mentioning the work of Cohen et al. (2017) is a must. It seems most appropriate to place this discussion right after Manikin's quote in the Introduction.

Is the use of "innovations" standard here? I understand where the authors are coming from, though I wonder if "forecast errors" might represent the description more.

"Innovation" (difference between forecast and observation) is standard terminology for the data assimilation community, but it is generally unfamiliar to those outside this community. Indeed, it is appropriate to introduce this word alongside the more familiar term "forecast error". Thus, in the first

paragraph of section 4.1, the following change is made: "…vectors $\mathbf{F}(t)$ *and* (t) *, the forecast error or Innovation* $\mathbf{I}\ (t)$ = $\mathbf{Z}(t) - \mathbf{F}(t)$ *is known. The sum…"*

[Minor comments omitted…]

Second Review:

Recommendation: Accept with minor revision.

General Comment: The Lewis et. al. paper is in great shape. Only things I found were very minor...

[Minor comments omitted…]