Improving Cloud Simulation for Air Quality Studies through Assimilation of Geostationary Satellite Observations in Retrospective Meteorological Modeling

ANDREW T. WHITE

Department of Atmospheric Science, University of Alabama in Huntsville, Huntsville, Alabama

ARASTOO POUR-BIAZAR AND KEVIN DOTY

Earth System Science Center, University of Alabama in Huntsville, Huntsville, Alabama

BRIGHT DORNBLASER

Texas Commission on Environmental Quality, Austin, Texas

RICHARD T. MCNIDER

Department of Atmospheric Science, University of Alabama in Huntsville, Huntsville, Alabama

(Manuscript received 16 May 2017, in final form 5 September 2017)

ABSTRACT

Development of clouds in space and time within numerical meteorological models as observed in nature is essential for producing an accurate representation of the physical atmosphere for input into air quality models. In this study, a new technique was developed to assimilate Geostationary Operational Environmental Satellite (GOES)-derived cloud fields into the Weather Research and Forecasting (WRF) meteorological model to improve the placement of clouds in space and time within the model. The simulations were performed on 36-, 12-, and 4-km grid-size domains covering the contiguous United States, the south-southeastern United States, and eastern Texas, respectively. The technique was tested over the month of August 2006. The results indicate that the assimilation technique significantly improves the agreement between the model-predicted and GOES-derived cloud fields. The daily average percentage increase in the cloud agreement was determined to be 14.02%, 11.29%, and 4.96% for the 36-, 12-, and 4-km domains, respectively. This was accomplished without degrading the model performance with respect to surface wind speed, temperature, and mixing ratio, which are important parameters for air quality applications; in some cases these variables were even slightly improved. The assimilation technique also produced improvements in the model-predicted precipitation and predicted downwelling shortwave radiation reaching the surface.

1. Introduction

Development of clouds in space and time within numerical weather prediction (NWP) models as observed in nature is essential for producing an accurate representation of the physical atmosphere for input into air quality models. This is because it is well known that air quality is directly coupled to meteorological conditions, and thus, errors in meteorology lead to errors in the predicted chemical composition of the air (e.g.,

Seaman 2000). The accurate prediction of clouds has been identified as a major source of error within NWP models (e.g., Pour-Biazar et al. 2007). This is because clouds directly modulate the energy balance over their area of influence, which is a driving mechanism for many processes within the atmosphere.

There are many specific mechanisms by which clouds alter atmospheric chemistry. First, clouds change the photodissociation reaction rates (photolysis rates) through an atmospheric column by scattering solar radiation, which alters the actinic flux. This increases the photolysis rates above the cloud and lowers them below the cloud (Tie et al. 2003). Clouds also affect the biogenic hydrocarbon emission

Corresponding author: Andrew T. White, awhite0427@gmail.com

rates by altering the air temperature and the photosynthetically active radiation (PAR) that reaches the biota (Kesselmeier and Staudt 1999). Additionally, clouds lead to the occurrence of aqueous reactions as chemical compounds interact with cloud droplets. An example of this is the production of secondary organic aerosol (SOA) in the presence of cloud droplets, as hypothesized by Blando and Turpin (2000) and further studied in multiple cases (Lim et al. 2010; Ervens et al. 2011). Clouds also modify the amount of vertical mixing. Clouds have the ability to mix pollutants from the boundary layer into the free troposphere, which has been shown to provide an important source of hydrogen oxide radicals to the upper troposphere (Tie et al. 2003). Additionally, clouds have the ability to alter the development of the boundary layer as they suppress the surface heat flux, which drives boundary layer development (Stull 1988). This then affects the vertical mixing. In the case of precipitating clouds, wet deposition of water-soluble species within the troposphere can be a significant removal mechanism of pollutants (Seaman 2000). Finally, convective clouds, which generate lightning, can also provide a significant source of nitrogen oxide (NO_x) into the free troposphere (Tie et al. 2003).

The simulation of clouds in NWP models has been improving over time with advancements in microphysical and planetary boundary layer parameterizations, combined with improvements in computational power to run more sophisticated parameterizations (Cintineo et al. 2014). Concurrent with improvements in the model physics, many research efforts have focused on the assimilation of observations to improve model performance, specifically cloud fields (Chen et al. 2015; Spero et al. 2014; Jones et al. 2014; Zhang et al. 2013; Jones et al. 2013; van der Veen 2013; Otkin 2010; Yucel et al. 2003). However, the improvement in the model forecast, in time, was limited. Chen et al. (2015) used a three-dimensional variational technique to assimilate satellite-derived cloud liquid water path and cloud ice water path from the Global Geostationary Gridded Cloud products at NASA and found improvements in temperature, humidity, and wind forecasts between 300 and 150 hPa. Jones et al. (2013) assimilated GOES cloud water path (CWP) using an ensemble Kalman filter (EnKF) approach and determined that the NWP model was able to consistently better represent the CWP field, but the improvement decreased during the forecast period. Analysis showed that the improvement in downwelling shortwave flux lasted 30 min, while the temperature improvements remained through a 90-min forecast. Van der Veen (2013) adjusted model-specific humidity profiles based on the Satellite Application Facility (SAF) cloud mask used on the Meteorological Satellite (Meteosat) Second Generation (MSG) satellite and improved cloud-cover forecasts. However, positive

impacts only lasted over 24 h on 59% of the forecasts completed. Jones et al. (2014) assimilated radar and satellite data simultaneously and found that the model analysis was improved, but the effectiveness of the assimilation generally lasted about 1 h. Yucel et al. (2003) assimilated GOES visible and infrared data, but also found that enhancements to the forecast lasted a maximum of 3 h. From this study, it was also concluded that the short-term impact of cloud assimilation in NWP models is caused by the inconsistency between the model dynamic field and the thermodynamic field. The problem is that the production and sustainability of cloud water is dependent on the water vapor and temperature environment that provides the needed relative humidity. Therefore, the added cloud water in the model, where the model has a dry environment, may not sustain for longer periods. Conversely, when liquid water is removed from the model where observations show no clouds, the model will continue to produce new water. Direct insertion of liquid water can even deteriorate model performance. As an example, attempting to insert clouds (based on observation) at a position where the model is clear means that the cloud is likely being inserted where the model has subsidence as opposed to lifting. Inserting water where the model has subsidence will cause evaporation and further increase subsidence, exactly the opposite of supporting the observed clouds.

Lien et al. (2016) also note that improving the model forecast beyond a few hours for precipitation has been difficult. Improving cloud forecasts, including non-precipitating clouds that are important for atmospheric chemistry, becomes even more challenging because of the reduction in the number of available observations. Standard weather service observations are not dense enough to be used for cloud specification, and the NWS WSR-88D network is not designed to be sensitive enough to retrieve cloud droplet information. Therefore, satellites remain the only platform that provides sufficient temporal and spatial resolution to quantify cloud fields. The *GOES-12* Imager has a spatial resolution of 1 km over the visible channel at $0.65 \mu\text{m}$ and 4-km resolution over the infrared channel at $10.7 \mu\text{m}$ for time scales down to 1 h or less. From this data, the cloud albedo can be retrieved from the visible channel, while the infrared channel is used to estimate the cloud-top heights. The GOES satellite retrieval of cloud albedo is described in Haines et al. (2004) and is based on an implementation of the Gautier et al. (1980) method with the improvements from Diak and Gautier (1983). Refer to Pour-Biazar et al. (2007) for further information on the satellite retrieval method.

While most studies focus on improving NWP models for forecasting applications, the focus of this study is to improve the simulation of clouds in time and space in

order to better represent the physical atmosphere for air quality studies. Where the use of an NWP model for forecasting and air quality diverges is that air quality studies are usually conducted as retrospective simulations. Therefore, observational and model analysis data are available throughout the time period of interest. This allows the four-dimensional data assimilation (FDDA) technique, based on Newtonian relaxation, or “nudging,” to be used throughout the simulation time period. Nudging works by adding artificial forcing terms to the governing equations to force the model toward an observed state (Stauffer and Seaman 1990). Nudging was originally implemented to create a dynamic initialization for a forecast model (Anthes 1974; Hoke and Anthes 1976) but is now routinely used for retrospective modeling simulations. There are typically two types of nudging that are routinely performed. The first is analysis nudging, which entails nudging to a gridded analysis field of meteorological variables at each model grid point. The second, observation nudging, consists of nudging toward point observations as they occur in space and time (e.g., Otte 2008). The use of nudging has also extended the time over which the model can be run before reinitializing while maintaining sufficient accuracy. Otte (2008) showed that there is no apparent degradation in model skill over 5.5-day simulation segments.

In this paper, a novel approach for assimilating satellite-observed clouds is taken. The technique attempts to estimate a target grid-scale vertical velocity based on satellite observations. In effect, this technique provides the dynamical support, and due to latent heating from condensation, the thermodynamical support, needed to sustain or clear clouds. Nudging is used to assimilate the estimated dynamic field into the Advanced Research core of the Weather Research and Forecasting (WRF) Model (WRF-ARW; Skamarock et al. 2008) to improve the simulation of clouds in place and time for the use in retrospective air quality modeling studies. Section 2 describes the model configuration. Section 3 presents the assimilation technique methodology. Section 4 includes the analysis of the model performance, showing the impact of the assimilation technique. The final section provides the discussion and conclusions.

2. Model configuration

The WRF-ARW version 3.6.1 was used in this study to conduct two different simulations. The first is a control simulation (CNTRL), in which no satellite data are assimilated, and the second simulation includes satellite cloud assimilation (Cloud Assimilation). The simulations are performed over the August 2006 period, which

FIG. 1. The 36- (D01), 12- (D02), and 4-km (D03) modeling domains.

coincides with part of the Texas Air Quality Study (TexAQs) II field study, on a parent 36-km domain covering the contiguous United States (D01), a nested 12-km domain covering the south-southeastern United States (D02), and a nested 4-km domain covering eastern Texas (D03) (Fig. 1). A one-way nesting configuration was used to allow cloud assimilation to be conducted on each domain independently. The model's initial and lateral boundary conditions were provided by the National Centers for Environmental Prediction (NCEP) Eta Model 3-h analysis (<http://rda.ucar.edu/datasets/ds609.2/>). The physics and analysis nudging options used in this study are shown in Table 1. The selection of this configuration was based on our prior sensitivity studies in which this configuration yielded the best performance for control simulation during August 2006. The physics options include the Dudhia shortwave radiation scheme (Dudhia 1989); the Rapid Radiative Transfer Model (RRTM; Mlawer et al. 1997) for long-wave radiation; the Lin microphysics scheme (Lin et al. 1983); the Kain–Fritsch convective parameterization (Kain 2004) with the moisture-advection-based trigger (Ma and Tan 2009); the Yonsei University (YSU) planetary boundary layer (PBL) scheme (Hong et al. 2006); the unified Noah land surface model (Tewari et al. 2004) with four soil layers; and the revised fifth-generation Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model (MM5) Monin–Obukhov scheme (Jiménez et al. 2012). While the moisture-advection-based cumulus trigger function was originally developed for tropical cyclones, conducted sensitivity studies revealed small-to-negligible differences in surface wind speed, temperature, mixing ratio, and precipitation when compared to

TABLE 1. WRF physics and analysis nudging options.

	Domain 1 (36 km)	Domain 2 (12 km)	Domain 3 (4 km)
Shortwave radiation		Dudhia (Dudhia 1989)	
Longwave radiation		RRTM (Mlawer et al. 1997)	
Surface layer		MM5 Monin–Obukhov (Jiménez et al. 2012)	
Land surface		Unified Noah land surface model (Tewari et al. 2004) with four soil layers	
Planetary boundary layer		YSU PBL (Hong et al. 2006)	
Microphysics		Lin (Lin et al. 1983)	
Cumulus parameterization	Kain–Fritsch (Kain 2004) with the moisture-advection-based trigger (Ma and Tan 2009)		None
Analysis nudging	Yes (for both the CNTRL and Cloud Assimilation)		No for CNTRL; yes for Cloud Assimilation
U, V nudging coefficient		$3.0 \times 10^{-4} \text{ s}^{-1}$	
T nudging coefficient	$3.0 \times 10^{-4} \text{ s}^{-1}$		0.0 s^{-1}
Q nudging coefficient	$1.0 \times 10^{-5} \text{ s}^{-1}$		0.0 s^{-1}
Nudging within PBL		Yes for U and V ; no for Q and T	

the original Kain–Fritsch trigger function. However, the use of the moisture-advection trigger did lead to improved cloud agreement with GOES and improved model-predicted shortwave irradiance when compared to pyranometers. Thus, because these are the two fields that our technique is expected to have the most impact on, further improving the best possible control simulation shows greater overall impact. Analysis nudging of u and v wind components, temperature, and mixing ratio was used in the simulations with nudging coefficients of 3.0×10^{-4} , 3.0×10^{-4} , and $1.0 \times 10^{-5} \text{ s}^{-1}$, respectively. There was no nudging of temperature or mixing ratio within the PBL. For D03, analysis nudging was not performed for the CNTRL simulation. However, the u and v wind components were nudged into the Cloud Assimilation simulation as a way to assimilate the GOES cloud information.

3. Methodology

The assimilation technique is based on creating a dynamic environment that is supportive of grid-resolved cloud formation or removal through the use of GOES cloud information. The basic approach is to create positive vertical motion within the model to produce clouds and negative vertical motion to dissipate clouds, based on GOES cloud fields. The use of FDDA allows for the assimilation of horizontal components of the wind into the model continuously in space and time. Therefore, the developed method provides a path to convert GOES cloud fields into vertical velocity estimates that are used to derive horizontal wind fields to be assimilated into the model through WRF FDDA. With the use of continuous assimilation with FDDA, it was determined that modification to the initial condition was unnecessary. However, for the nested domains, lateral boundary conditions are supplied by the corresponding parent domain in which cloud assimilation was also applied.

The first step in the assimilation process is to determine the locations of disagreement between the model cloud fields and the GOES cloud fields. This is achieved by comparing satellite-derived cloud albedo with that of the model. To compare GOES-derived cloud albedo with the model cloud field directly, a model cloud albedo field, compatible with the satellite observation, needed to be derived. This was done by completing a model simulation without microphysics and cumulus parameterization schemes enabled to determine the maximum insolation predicted by the model's radiation scheme at every grid point over the model domain. The difference in the insolation fields from the model simulation with/without cloud fields allows a compatible model cloud albedo to be determined. While the use of cumulus radiation feedback does modify the model insolation fields, it was determined that it produced a negligible change in the cloud agreement and therefore was not used in this study. With the model cloud albedo known, the locations of disagreement with GOES are determined. The disagreement is split into two categories: under- and overprediction of clouds by the model. Underprediction areas are locations where the model has clear conditions and the satellite shows cloudiness. Conversely, overprediction areas are locations where the model is cloudy and the satellite indicates clear sky. A 10% cloud albedo threshold is used to identify if a model or GOES grid is flagged as cloudy. The 10% threshold is used to account for uncertainties in both the satellite retrieval (i.e., impact of aerosols and water vapor) and the model-derived cloud albedo. Because GOES-derived cloud albedos retrieved from the visible GOES channel are being used, the assimilation window is limited to the daytime. With the modeling domain covering the continental United States (CONUS), the assimilation window was chosen as 1400–2300 UTC to make sure that the solar zenith angle was as small as possible over the model domain. This reduces the error in the

satellite-derived variables and, to a larger degree, reduces the error in the model albedo calculation. The uncertainty in the model cloud albedo calculation increases when the maximum possible solar insolation received at the surface at a given solar hour is low (high zenith angle). This results in a cloud albedo calculation that is overly sensitive to any change in received insolation.

Once the disagreement areas between the model and satellite are known, a target vertical velocity and its location necessary to produce or dissipate clouds within the model is estimated. To create the derived vertical motion in the model, a one-dimensional variational technique based on O'Brien (1970) is used. The O'Brien (1970) technique provides a method to determine the adjusted divergence field needed to achieve the vertical motion prescribed at a certain height within a column based on an original first-guess field. While WRF-ARW's vertical coordinate is a sigma- p formulation, the divergence adjustment is calculated on a sigma- h grid. Therefore, the model is interpolated to the sigma- h coordinate for the calculations and then back to the sigma- p coordinate to assimilate into the model. During the interpolation, the vertical resolution is not maintained. Instead, the sigma- h grid is defined at a higher resolution, varying from 3 m near the surface to a maximum layer thickness of 300 m. Because the divergence adjustment is completed at a higher vertical resolution, vertical interpolation is performed regardless of the selected vertical coordinate. This, combined with the simple relationship between vertical velocity in sigma- h coordinates and a Cartesian vertical velocity w , makes the use of the sigma- h vertical coordinate advantageous. The final form used to calculate the adjusted divergence D_N^* at each level N in this work is given by

$$D_N^* = D_N + \frac{1}{\Delta h_N} \left[\frac{2N^*}{L^*(L^* + 1)} \right] (\rho_L \dot{h}_L - \rho_L \dot{h}_{TL}), \quad (1)$$

where D_N is the original divergence, Δh_N is the layer thickness, N^* is the current level index relative to the adjustment boundaries, L^* is the target level index relative to the adjustment boundaries, ρ_L is the density at the target level, $\dot{h}_L$ is the original sigma- h vertical velocity at the target level, and $\dot{h}_{TL}$ is the target sigma- h vertical velocity (see appendix for derivation). The adjusted divergence fields are then used to calculate the divergent component of the wind or the velocity potential. To determine the velocity potential from divergence, the Poisson equation shown in (2) is solved using a simultaneous overrelaxation (SOR) scheme,

$$MD_N = \nabla^2 \chi_N, \quad (2)$$

where M is the sigma- h height scale, D_N is divergence, and χ_N is the velocity potential at any level N . Dirichlet boundary conditions are used, prescribing zero velocity potential around the boundaries. As noted in Lynch (1988), the use of zero boundary conditions minimizes the kinetic energy of the divergent component of the wind. Once the velocity potential is known at each level, the new divergent component of the wind is added to the original rotational component of the wind field to fully construct a new horizontal wind field. Note that (2) is applied everywhere within the model domain, while the divergence adjustment is only applied to disagreement areas between the model and GOES. The result is a new divergent component of the wind at every grid point, but it also acts to balance the mass throughout the model domain. Therefore, the technique implicitly introduces cloud-induced subsidence across the model domain in response to the created clouds.

To calculate an adjusted divergence, initial and boundary conditions must be supplied to (1). While the initial condition is the original CNTRL simulation, the developed assimilation method is used to derive the necessary boundary conditions for the variational approach. These inputs are the target vertical velocity, the target height at which the target vertical velocity is reached, the bottom adjustment height, and the top adjustment height. To develop the necessary vertical motion within the column, (1) is applied twice. In the first instance, it is applied from the surface to the vertical velocity target height. In the second instance, it is applied from the level above the target height to the model top. At the target height boundary, the prescribed vertical velocity is set to the target vertical velocity. For the adjustment from the target level to the model top, the prescribed vertical velocity at the model top is set to zero. In both cases, the divergence adjustment is weighted so that the magnitude of the adjustment goes to zero at the prescribed bottom and top adjustment heights. Therefore, there is no adjustment of the divergence near the surface or above the tropopause [further details on the application of (1) for each case are provided in the appendix]. The approach for estimating the input parameters to the variational technique for the over- and underprediction cases is described below.

a. Underprediction

In the underprediction case, the model fails to produce clouds at the locations where GOES is cloudy. In these locations, the goal is to produce a cloud within the model that is comparable to the satellite field. This is basically done by identifying a parcel of air that can be

lifted to saturation in order to produce a cloud comparable to GOES observation. The distance that such a parcel will travel vertically constitutes the required vertical velocity. The steps to achieve this goal start with examining the observation using GOES cloud-top temperature and cloud albedo estimates. The observed cloud-top height is determined by calculating the height at which the model temperature is equal to the GOES IR temperature indicating cloud-top temperature. Then, the cloud thickness is estimated using the determined cloud-top height and the GOES cloud albedo. An empirical formula is used to estimate the vertical pressure change across the cloud, assuming both that an increase in cloud albedo indicates an increase in cloud thickness and that there is only one cloud layer. The empirical formula is given by

$$\Delta p = (P_{\text{Max}} - P_{\text{CTOP}}) \alpha_{\text{CLD}}^{(P_{\text{ICE}}/P_{\text{CTOP}})}, \quad (3)$$

where ΔP is the vertical pressure change, P_{Max} is the maximum between the pressure at the model column PBL height and the pressure at the first model level 1.5 km above the surface, P_{CTOP} is the model pressure at the determined satellite cloud-top height, α_{CLD} is the GOES cloud albedo, and P_{ICE} is a constant at 400 hPa. The constant P_{ICE} is used to account for the increased reflectance due to ice over liquid. The pressure difference is then converted into a height difference to obtain the estimated cloud thickness. The estimation of cloud thickness is only used as a first guess, with the intent to produce a cloud that roughly has the same radiative characteristics as the satellite-observed cloud. Additionally, the model is not constrained by the estimated cloud thickness. Thus, the uncertainty in cloud thickness estimation does not adversely impact the effectiveness of this technique. With the estimation of cloud thickness, it is then possible to determine the maximum height at which saturation must occur to maintain the ability of the model to produce an equivalently thick cloud. This height is calculated as the model column tropopause height, estimated by using potential vorticity (Hoskins et al. 1985), minus the estimated GOES cloud thickness. This height will be referred to as MAX_SAT. The MAX_SAT height is calculated to ensure that the divergence adjustment is completed below the tropopause height and that the distance below the tropopause is based on the first-guess GOES cloud thickness. Next, the lifted condensation level for each model vertical layer below MAX_SAT is determined. The layer that has to be lifted the smallest distance to become saturated and can achieve saturation before the MAX_SAT height is taken as the displacement height and is used to calculate the target vertical velocity necessary to produce a cloud within the model column. The target

vertical velocity is calculated by taking the distance the layer has to rise to reach saturation over 45 min. A 45-min time scale is selected because the GOES assimilation occurs on an hourly time scale, and thus, it is desirable to create the cloud within 15 min of observation. However, it is recognized that this time scale should be revisited based on the temporal frequency of observation.

Once the target model layer is defined, the other three components of the input to variational code are also determined. The target height is defined as the height at which the target model layer reaches its lifted condensation level. The top adjustment height is then determined to be the target height plus the estimated GOES cloud thickness. The bottom adjustment height is determined to be the target model layer minus 1 km. However, in determining the bottom adjustment height, thresholds are set to ensure that this level does not reach the surface. A threshold minimum height is set to two-thirds of the model column's PBL height or 500 m, whichever is greatest. This is done to ensure that the assimilated nudged winds minimally affect the surface layer of the model. The top and bottom adjustment heights and target height define the limits between which the divergence is adjusted through (1); see appendix for implementation. To ensure that the determined vertical velocities are not too large, a maximum vertical velocity limit is also imposed, such that no prescribed vertical velocity in one column is greater than the average calculated underprediction target vertical velocity. This is beneficial in the cases where the model column is too dry, making large vertical lifting necessary to reach saturation. Allowing these large vertical velocities results in model degradation rather than improvement. In subsequent studies, an adjustment to atmospheric moisture will be considered to alleviate this issue.

b. Overprediction

In the overprediction case, the model has created a cloud in a grid location that GOES indicated is not cloudy. The method used to remove these erroneous clouds is to create subsidence through the model cloud field to evaporate the cloud. This is completed by first determining the location of the cloud/cloud layers within the model column. The presence of clouds is determined by the total cloud liquid water (CLW) in each vertical layer. The CLW is represented by the sum of the liquid and ice water in each layer. The layer with the maximum amount of CLW is used to calculate the downward displacement necessary to sufficiently warm the air parcels in order to evaporate the cloud. The distance is calculated as the vertical distance that a parcel must sink and result in a relative humidity that is less than 60%. The

60% threshold is used to ensure sufficient evaporation has occurred. The calculated distance divided by 45 min then yields the target vertical velocity. The target height for adjustment is set to the height of the layer with the maximum CLW. The column cloud top and cloud base are then determined by setting a $1.0 \times 10^{-6} \text{ kg kg}^{-1}$ CLW threshold to indicate a cloud. The top adjustment height is then calculated as the model cloud-top height plus 1 km or the column tropopause height, whichever height is less. The bottom adjustment height is calculated as the height at which the layer containing the maximum CLW will evaporate minus 1 km, or the determined model column cloud base minus one-fourth of the distance between this point and the surface, whichever height is less. Similar to the underprediction case, a threshold minimum height for the bottom adjustment is set as the maximum between 500 m and two-thirds of the model PBL height to ensure the adjustment is minimally affecting the surface layer. Like the underprediction case, the top and bottom adjustment heights and target height define the limits between which the divergence is adjusted through (1); see appendix for implementation. Further, the maximum vertical velocity limit is also imposed for overprediction. In this case, this is beneficial in model columns that are too moist, making large vertical subsidence necessary to evaporate all the cloud water. Similar to the underprediction case, allowing these large vertical velocities results in model degradation rather than improvement.

4. Results

To validate the assimilation technique, both the CNTRL and Cloud Assimilation model simulations were compared to surface observations and GOES retrievals. Surface observations were used to determine the model bias and root-mean-square error (RMSE) of each simulation with respect to surface wind speed, temperature, mixing ratio, and precipitation. An evaluation with respect to upper-air radiosonde data was completed for wind speed, temperature, and mixing ratio as well. Both model simulations were also compared to observed surface shortwave radiation to determine the change in the predicted downwelling shortwave radiation by the model due to the assimilation. GOES retrievals of cloud albedo and insolation were used to show the spatial effects of the assimilation technique.

a. Comparison with satellite observation

To determine how well each model simulation performed with respect to GOES, a cloud agreement index (AI) based on a contingency table, shown in Table 2, was

TABLE 2. Contingency table used to derive the AI.

		WRF	
		Cloudy	Clear
GOES	Cloudy	<i>A</i>	<i>B</i>
	Clear	<i>C</i>	<i>D</i>

devised to quantify the model cloud performance. The AI is calculated as the percentage of the total grids that agree with the satellite and is expressed as

$$\text{AI} = \frac{A + D}{A + B + C + D}. \quad (4)$$

In (4), as in Table 2, *A* represents the number of grids where both the model and GOES indicate clouds, *B* represents the number of grids where GOES indicates clouds and the model is clear (underprediction), *C* represents the number of grids where GOES is clear and the model is cloudy (overprediction), and *D* represents the number of grids where both GOES and the model are clear. The grid cells were characterized as cloudy or clear based on the derived cloud albedos for the model and satellite. As discussed previously, a 10% cloud albedo threshold was used to signify if a model or GOES grid is flagged as cloudy. The value of AI varies from 0 to 1, with closer to 1 indicating better model agreement with GOES. The AI was calculated hourly from 1500 to 2300 UTC for each day for all three modeling domains. The initial assimilation hour, 1400 UTC, was omitted from the AI evaluation to allow time for the assimilated clouds to spin up. The hourly AI was then averaged to produce a daily average AI for the simulation period. The daily average AI for both model simulations can be seen in Figs. 2–4 for the 36-, 12-, and 4-km domains, respectively. The average percentage change in the AI over the simulation period was determined to be 14.02%, 11.29%, and 4.96% for the 36-, 12-, and 4-km domains, respectively. For both the 36- and 12-km domains, the assimilation technique was found to consistently improve the placement of clouds in space and time when compared to GOES. At the higher grid resolution of 4 km, the assimilation technique did not perform quite as well. However, improvement in the AI overall was still achieved. The variability in the AI at this grid resolution can be attributed to the more transient nature of clouds within the smaller grid cell over a time period of 1 h. In other words, when the cloud fields throughout the domain are characterized as fair weather cumulus clouds or isolated convective cells, the time scale for cloud development and advection to neighboring grid cells becomes smaller than the input of the hourly nudging field. However, when a large area of the domain is covered by clouds, the assimilation technique still performs adequately. The

FIG. 2. Average daily AI for the CNTRL (blue) and Cloud Assimilation (red) for the 36-km simulation.

assimilation technique also improves the cloud agreement by a nearly uniform amount across the 1500–2300 UTC time period, as shown in Fig. 5. The percentage increase in the AI due to the assimilation peaks around the 1600–1800 UTC time period and then slightly decreases by a few percent as 2300 UTC is approached. While the magnitudes of the hourly averaged AI are different across the three domains, the overall pattern is the same.

Evaluating the spatial fields of model/GOES for cloud agreement, cloud albedo, and insolation reveals the overall effect of the assimilation technique. Figure 6 shows the cloud agreement for 1700 UTC 12 August 2006 for the 36-km domain. For this particular hour, the gross increase in the AI was determined to be 14.9% (a 22% change in the cloud agreement). It can be clearly seen that the cloud assimilation is able to remove nearly all of the cloud overprediction in the CNTRL simulation over the Gulf of Mexico. Another overprediction area that the assimilation was able to correct was over southern Arizona into northwestern Mexico. The cloud assimilation was also able to correct underprediction areas by creating clouds in an

area stretching from Missouri to Minnesota that was completely missed by the CNTRL simulation. The cloud assimilation further expanded cloud coverage in the Four Corners region, the Southeast, the north-central United States, the Montana area, and north Texas. The AI plots are further validated by the spatial plots of derived cloud albedo (Fig. 7) and surface insolation (Fig. 8). From these two figures, it is clear that the spatial pattern of clouds in the Cloud Assimilation simulation is much closer to that of GOES observation than the CNTRL simulation. The cloud agreement spatial plots for the 12-km domains at the same time are shown in Fig. 9. For the 12-km domain, the amount of cloud underprediction by the model was reduced as the amount of cloud coverage increased over western Tennessee, Arkansas, Missouri, and north Texas. It can also be seen that the amount of cloud overprediction by the model is reduced over the Gulf of Mexico. At the highest grid resolution of 4 km, the improvement is not as significant. However, as evident in Fig. 10, the overall cloud pattern is in better agreement with the GOES observation. This particular hour also shows some of the

FIG. 3. As in Fig. 2, but for the 12-km simulation.

FIG. 4. As in Fig. 2, but for the 4-km simulation.

challenges of using hourly data for higher-resolution grids. In trying to correct some of the underprediction areas over central Louisiana, there is the new creation of overprediction areas in neighboring grid cells. Similarly, this occurs along the coast of Texas. The assimilation technique also had trouble creating the clouds, as observed by GOES, immediately inland of the coast throughout Texas. These clouds are optically thin and are described by low reflectance, as observed by GOES and shown in Fig. 10. Our technique distributes the prescribed vertical motion in the underprediction areas throughout a given column in a way that is proportional to the observed cloud thickness. Thus, the lower the GOES cloud albedo, the more dependent on the original CNTRL flow field the end result becomes. In higher spatial resolution simulations, such as 4 km, the vertical motion within each grid cell can become larger than in grids with coarser resolution, and thus, small changes are less effective.

b. Evaluation with respect to surface observations

The model simulations were then evaluated with respect to surface observations. The hourly observation data for wind speed, temperature, and mixing ratio from National Weather Service (NWS) stations were retrieved from the NCAR Research Data Archive (<http://rda.ucar.edu/datasets/ds472.0/>). The METSTAT package developed by Ramboll Environ (<http://www.camx.com/download/support-software.aspx>) was used to calculate the mean bias and RMSE for wind speed, temperature, and mixing ratio. METSTAT interpolates all model data to the observation points before calculating the difference statistics. For the 36-km domain, the surface statistics for wind speed, temperature, and mixing ratio were found to vary, depending on the day and variable. The cloud assimilation was found to effectively decrease the temperature across the domain as shown in Fig. 11 by the increase in the negative bias. However, the temperature RMSE was decreased for the cloud assimilation as shown in Fig. 12.

FIG. 5. Average hourly AI for the CNTRL (blue) and Cloud Assimilation (red) for the 36-km simulation.

FIG. 6. The spatial agreement between the model and GOES cloud fields for (a) CNTRL and (b) Cloud Assimilation at 1700 UTC 12 Aug 2006 for the 36-km simulation. Green represents locations where the model and GOES were clear, red is locations where the model is cloudy and GOES is clear, orange is areas where the model is clear when GOES is cloudy, and gray is locations where both the model and GOES were cloudy. The AI is the percentage of grids within the domain that are in agreement with GOES.

This indicates that the creation of more clouds in the correct places led to a reduction in temperature due to shading, which then decreased the error and further pushed the overall inherent model bias more negative. For wind speed and mixing ratio, the cloud assimilation overall slightly increased the RMSE. While these errors do increase, both wind speed and mixing ratio are highly variable in space, and thus, a coarse resolution is not expected to be representative of point observations for these variables. This is further supported by the fact that on the 12- and 4-km domains, the consistent daily increase in the RMSE is not present. For the 12-km simulation (Fig. 13), the cloud assimilation consistently improves the mixing ratio RMSE, while at 4 km, there is more variability but still a reduction in the error on half of the days through the month. The variation at the 4-km

resolution is likely due to the technique contaminating neighboring grid cells, as discussed above. For wind speed, the results are similar. For the 12-km simulation, the wind speed RMSE of the CNTRL and Cloud Assimilation simulations were similar, with no consistent increase or reduction through the time period. Further reducing the grid spacing to 4 km, the wind speed RMSE is significantly improved in the Cloud Assimilation simulation, as shown in Fig. 14. As recommended for regulatory modeling, the CNTRL simulation is run without analysis nudging at 4 km, as the input analysis data become too coarse to justifiably use at this grid resolution. However, in regulatory modeling, the usual practice is observation nudging, which was not performed in this study. In not doing analysis nudging in the CNTRL simulation, the simulation tended to produce higher

FIG. 7. The derived cloud albedo for (a) CNTRL, (b) Cloud Assimilation, and (c) GOES at 1700 UTC 12 Aug 2006 for the 36-km simulation.

FIG. 8. Surface insolation for (a) CNTRL, (b) Cloud Assimilation, and (c) GOES at 1700 UTC 12 Aug 2006 for the 36-km simulation.

wind speeds throughout the day. In the Cloud Assimilation simulation, however, analysis nudging is still used. But in this case, the assimilated wind field that is based off the 4-km CNTRL simulation's predicted wind field is assimilated into the model during the 1400–2300 UTC time period. For the remaining hours of the day (0000–1300 UTC), the interpolated analysis field is still nudged into the model. The introduction of an assimilated nudging field tends to regulate the wind speeds, which then reduced the error throughout the day. The Cloud Assimilation simulation was found to consistently reduce the daily temperature RMSE at 12 km but produced more variable results at 4 km, as shown in Figs. 15 and 16, respectively. The variability in the 4-km results is, once again, attributed to neighboring gridcell contamination. While not shown here, the model simulations were also evaluated against radiosonde data to ensure that the upper air and surface statistics were similar. It was found that there was not a consistent increase or decrease in the model bias or

RMSE when compared to radiosonde data, but the Cloud Assimilation simulation performed better as the grid spacing decreased. Overall, the surface statistics for wind speed, temperature, and mixing ratio show that the cloud assimilation technique can produce results that are modestly better or comparable to a control model simulation. Most of the improvement occurs with the temperature field due to the impact clouds have on the radiative fields.

c. Precipitation evaluation

Each model simulation's predicted surface precipitation was also validated against NCEP's Stage IV product. The Stage IV product is a multisensor (i.e., radar, gauge, and satellite) precipitation analysis produced from a mosaic of the 12 CONUS NWS River Forecast Centers' multisensor hourly and 6-hourly analysis (Lin and Mitchell 2005). The 4-km Stage IV product was interpolated to the model grids by taking an area average of the grids that fell within each model grid. Once the Stage IV analysis was regridded to the model

FIG. 9. As in Fig. 6, but for the 12-km simulation.

FIG. 10. As in Fig. 7, but for the 4-km simulation.

domains, the domainwide mean bias and RMSE of each model simulation were computed. For the 36-km CONUS domain, the RMSE improvement is minimal. In fact, the error increases slightly on the majority of the days throughout the simulation time period. However, because Stage IV data are mostly available over land, precipitation improvement over the ocean is likely missed, which is where large cloud improvements on the 36-km domain occurred (where Fig. 6 shows a good representation of the amount of improvement). Precipitation is also generally more localized than a 36-km grid can accurately represent, especially in the summer-time environment. At the 12-km domain, which covers a smaller ocean area, the RMSE for the Cloud Assimilation simulation shows more improvements. Further reducing the grid spacing to 4 km, there is a continued reduction in the RMSE, shown in Fig. 17, provided by the Cloud Assimilation simulation. These results are in agreement with our physical model that improving the placement of clouds in space and time should improve the predicted precipitation by the model.

d. Surface radiation evaluation

Next, both model simulations were validated against surface pyranometer measurements in order to quantify the change in predicted solar insolation at the surface due to the cloud assimilation technique. The U.S. Climate Reference Network (USCRN) insolation data were used to validate the model output (Diamond et al. 2013). USCRN stations are equipped with a Kipp and Zonen SP Lite pyranometer, which creates a voltage output that is proportional to the incoming solar radiation. During the August 2006 time period, the USCRN network provided a maximum temporal resolution of solar insolation of 1 h. However, the data were hourly averaged solar insolation measurements while the model-predicted values were instantaneous. Therefore, similar to Otkin et al. (2005), the USCRN time stamps were shifted to represent the half hour of the averaging interval, and the WRF-predicted solar insolation was linearly interpolated to the same time. The model-predicted value chosen to represent the observation was determined by taking the closest

FIG. 11. Daily temperature mean bias for the CNTRL and Cloud Assimilation simulations for the 36-km simulation.

FIG. 12. Daily temperature RMSE for the CNTRL and Cloud Assimilation simulations for the 36-km simulation.

model value to the observation from the grid cell containing the USCRN station and the eight adjacent grids. This was done to better capture the high variability in solar radiation at a point observation over the hour averaging period. Figure 18 shows the difference in mean bias and RMSE over the 36-km domain for the CNTRL and Cloud Assimilation simulations. The Cloud Assimilation simulation shows a nearly uniform reduction in bias and RMSE across the entire domain. However, over one station near Mobile, Alabama, there was an increase in bias but little change in RMSE. Table 3 shows the average mean bias and RMSE for each simulation over the August 2006 time period. For all three domains, the use of the cloud assimilation technique reduces the bias and RMSE in the model insolation field when compared to surface observations. For the 12-km domain, the improvement in the model insolation was determined to be less than the other two domains. In particular, the 12-km domain increased the bias over many of the stations located over the Southeast states

and, for a few stations, slightly increased the RMSE. This result can be mainly attributed to the transient nature of convection in this region and the relatively few USCRN stations over the area. Further model comparisons to the Soil Climate Analysis Network (SCAN) pyranometer observations, which had significantly more stations over states like Alabama and Mississippi, showed that Cloud Assimilation simulation reduced the bias by $\sim 14 \text{ W m}^{-2}$ and RMSE by $\sim 20 \text{ W m}^{-2}$ over the 12-km domain. SCAN data were not used for the entire evaluation because the coverage over the entire CONUS region was not as uniform as USCRN during the August 2006 time period. Additionally, there were multiple SCAN stations in the northern and western part of the CONUS domain that were reporting erroneous clear sky values that were too low for summertime. This highlights the fact that observations have sources of error and that great care must be taken when comparing model values to point observations. Overall, the results indicate that the cloud assimilation technique does improve the predicted solar insolation. This

FIG. 13. Daily mixing ratio RMSE for the CNTRL and Cloud Assimilation simulations for the 12-km simulation.

FIG. 14. Daily wind speed RMSE for the CNTRL and Cloud Assimilation simulations for the 4-km simulation.

is due to the improvement in cloud prediction in place and time producing a more accurate atmospheric optical depth.

5. Conclusions

In this study, GOES satellite-derived cloud albedo and cloud-top temperature information was assimilated into the WRF Model to improve the placement of clouds in space and time. A technique for estimating vertical velocities necessary to produce and sustain the observed clouds within the model was developed. By employing a variational technique, the estimated velocities were used to generate horizontal wind components necessary to support such vertical velocities. Horizontal winds were assimilated in the model through WRF FDDA. The results from a CNTRL WRF simulation without cloud assimilation (but employing analysis nudging) and a Cloud Assimilation WRF simulation were compared over the August 2006 time period. The simulations were performed on three different domains: 36-, 12-, and 4-km grid spacings. The results of each simulation were compared to

surface observations of wind speed, temperature, mixing ratio, precipitation, and shortwave radiation. The results were also compared to satellite-derived cloud fields and upper-air radiosonde data.

The results indicated that the assimilation technique greatly improved the placement of clouds in space and time when compared to satellite-derived cloud fields. The average percent increase in the daily average agreement between the model and GOES cloud fields was found to be 14.02%, 11.29%, and 4.96% for the 36-, 12-, and 4-km domains, respectively. For certain time periods, the improvements exceed 20%. As expected with improved representation of clouds across the modeling domain, there was also a reduction in the model surface temperature error when compared to surface observations. The improvement in cloud fields also tended to reduce the model surface mixing ratio error when compared to surface observations. However, more variability was evident, especially on the 36- and 4-km domains. The result at 36 km is to be expected for mixing ratio at this coarse grid resolution due to the high

FIG. 15. As in Fig. 12, but for the 12-km simulation.

FIG. 16. As in Fig. 12, but for the 4-km simulation.

variability of mixing ratio in space. At 4 km, the grid resolution starts to become too fine for hourly satellite data in some instances. This leads to the increased variability. The improvement in the model-predicted surface wind speed was found to be variable across the three domains. On the 36-km grid, the assimilation technique increased the wind speed error slightly, but on the 12-km grid, the error between the two simulations was comparable. Because of the high variability of wind speed in space, it is expected that the 12-km domain is more representative of surface wind speeds. At the 4-km grid spacing, however, there was a consistent reduction in the surface wind speed error. The improvement in the cloud placement also improved the model precipitation error. This was especially evident at the finer grid resolutions of 12 and 4 km. This is expected in a summertime environment for domains covering the southern portion of the United States because rainfall is more convective in nature, associated with single-cell convective storms that are highly variable in space. Finally, the results also indicated that the model-predicted downwelling

shortwave radiation in assimilation run better represented the observed USCRN data. This is expected, as more accurate atmospheric optical depth profiles should lead to more accurate surface insolation estimates. Overall, the cloud assimilation technique greatly improves the placement of clouds in space and time. In doing so, model-predicted values of wind speed, temperature, mixing ratio, precipitation, and solar insolation are improved or comparable to a control model simulation.

This study shows that while the agreement between the model-predicted clouds and GOES-derived cloud fields is improved across all three domains of different grid resolutions, there are instances in which the technique performs better than others. At the 36-km grid resolution, the agreement between model and GOES cloud fields is at a maximum, which, in turn, reduced the shortwave irradiance bias and RMSE by greater than 17 W m^{-2} . However, it was determined that there is minimal, if any, improvement in other model fields as a result. At the 4-km grid resolution, the cloud agreement was still increased, but it

FIG. 17. Daily precipitation RMSE for the CNTRL and Cloud Assimilation simulations for the 4-km simulation.

FIG. 18. Model performance with respect to USCRN pyranometers observations for domain 1 (36 km). (left) Mean bias absolute difference and (right) the RMSE difference between the CNTRL and Cloud Assimilation simulations (negative indicates a reduction in bias and error when cloud assimilation is performed).

was by a smaller amount, while the predicted shortwave irradiance, precipitation, temperature, moisture, and wind speed were improved significantly. This is likely because of the assimilation time scale. At an assimilation time scale of 1 h, the clouds that are being developed likely move out of the grid box before the next comparison hour. Thus, if the surrounding grids where the cloud was developed should be clear, the result is a flagged model overprediction as the cloud is advected to the neighboring grids. In this case, the correct cloud was created, but because of the assimilation time scale being longer than the advection time scale, the result is disagreement rather than agreement with GOES. This type of scenario is typical of a summertime environment. These single-cell thunderstorms are characterized as meso- γ scale, which occur on time scales of 1 h, as indicated by Orlanski (1975). Thus, it would be necessary to have data with higher temporal resolution (more frequently than 1 h) to accurately represent them. For phenomena that are larger than the meso- γ scale, such as mesoscale convective systems that persist longer than 1 h, the assimilation technique performs adequately. The large improvement in shortwave irradiance at 4 km further supports this notion because the model-predicted value closest to the observation from the immediate surrounding grid points is used in the statistics calculation. Therefore,

with the use of a 1-h assimilation time scale, the results indicate that the assimilation has the greatest overall improvement at the grid resolution near 12 km.

The dependence on the assimilation time scale is a problem that requires further research to improve the portability of the assimilation technique to finer grid resolutions. The recent launch of the *GOES-16* satellite, equipped with the Advanced Baseline Imager (ABI), will also aid in our ability to have higher temporal and spatial resolution data. This new dataset is expected to improve the ability of the assimilation technique for higher-resolution model simulations, which are expected to become the norm for both meteorological and air-quality simulations. It is also expected that the improvement in cloud prediction can be achieved in any meteorological model that also has an FDDA-type option. The assimilation technique was implemented to not have any model physics dependencies. But it is noted that different model physics combinations will produce different initial cloud solutions and therefore lead to different initial disagreement between model and GOES cloud fields. Thus, the amount of improvement attained depends on the simulated environment, and the exact magnitude of the improvements will be case dependent.

Finally, the results presented here should be viewed as an initial attempt in applying a new technique for

TABLE 3. Model statistics for shortwave irradiance compared to USCRN data.

	Domain 1 (36 km)		Domain 2 (12 km)		Domain 3 (4 km)	
	Mean bias (W m^{-2})	RMSE (W m^{-2})	Mean bias (W m^{-2})	RMSE (W m^{-2})	Mean bias (W m^{-2})	RMSE (W m^{-2})
CNTRL	34.6	110.1	54.8	144.6	73.0	156.3
Cloud Assimilation	17.2	91.4	53.8	136.3	54.9	112.6

assimilating satellite-observed clouds in the numerical weather forecast models. While the results from this study are promising and show a significant improvement in model performance, many of the underlying assumptions used in this study can be revisited and improved. Furthermore, application and evaluation of this technique for other episodes and over multiple summers can lead to better refinement of such assumptions.

Acknowledgments. This study was accomplished by partial support from NASA Science Mission Directorate Applied Sciences Program (NASA Grant NNX15AN63G) and the Texas Commission on Environmental Quality (TCEQ). Note the results in this study do not necessarily reflect policy or science positions by the funding agencies. Additionally, we would like to thank the two anonymous reviewers and our editor Dr. Juazhen Sun, whose time and comments improved the quality of this paper.

APPENDIX

Divergence Adjustment

a. Divergence adjustment derivation

The divergence adjustment based on O'Brien (1970) in the sigma- h coordinate system is derived in the following manner. Starting with the continuity equation in sigma- h coordinates,

$$M \frac{\partial \rho}{\partial t} + \frac{\partial \rho u M}{\partial x} + \frac{\partial \rho v M}{\partial y} + \frac{\partial \rho \dot{h} M}{\partial h} = 0, \quad (\text{A1})$$

where $M = H - E$, H is the constant height above the mean sea level (MSL) of the flat model top, E is the terrain height, and $\dot{h}$ is the sigma- h vertical velocity. Solving for the vertical component and integrating between two levels yields

$$\begin{aligned} (\rho_k \dot{h}_k) - (\rho_{k-1} \dot{h}_{k-1}) &= - \int_{k-1}^k \frac{\partial \rho}{\partial t} dh \\ &\quad - \frac{1}{M} \int_{k-1}^k \left(\frac{\partial \rho u M}{\partial x} + \frac{\partial \rho v M}{\partial y} \right) dh. \end{aligned} \quad (\text{A2})$$

The divergence and density tendency are defined as

$$D_k \Delta h_k = \frac{1}{M} \int_{k-1}^k \left(\frac{\partial \rho u M}{\partial x} + \frac{\partial \rho v M}{\partial y} \right) dh, \quad (\text{A3})$$

$$T_k \Delta h_k = - \int_{k-1}^k \frac{\partial \rho}{\partial t} dh. \quad (\text{A4})$$

Substituting (A3) and (A4) into (A2) and applying the result to a series of layers from the ground to height L yields

$$\rho_L \dot{h}_L = \rho_1 \dot{h}_1 + \sum_{k=1}^{k=L} T_k \Delta h_k - \sum_{k=1}^{k=L} D_k \Delta h_k. \quad (\text{A5})$$

To achieve a target vertical velocity $\dot{h}_{TL}$ at height L , we use adjusted divergences D_k^* :

$$\rho_L \dot{h}_{TL} = \rho_1 \dot{h}_1 + \sum_{k=1}^{k=L} T_k \Delta h_k - \sum_{k=1}^{k=L} D_k^* \Delta h_k. \quad (\text{A6})$$

To find the adjusted divergence, as in O'Brien (1970), the one-dimensional variational function is given by

$$\begin{aligned} J &= \left[\sum_{k=1}^{k=L} \kappa_k (D_k^* - D_k)^2 \Delta h_k^2 \right] \\ &\quad + 2\lambda \left[\rho_1 \dot{h}_1 - \rho_L \dot{h}_{TL} + \sum_{k=1}^{k=L} T_k \Delta h_k - \sum_{k=1}^{k=L} D_k^* \Delta h_k \right], \end{aligned} \quad (\text{A7})$$

where κ_k are the Gauss precision moduli and λ is a Lagrangian multiplier. Equation (A7) is the functional of the difference between the adjusted divergence and the first-guess divergence that satisfies the constraint of the continuity equation. The precision moduli are related to the error variance σ_k by

$$\kappa_k = \frac{1}{2\sigma_k^2}. \quad (\text{A8})$$

The adjusted divergence values are then determined using variational calculus. Therefore, the minimum adjustment occurs where the first derivative goes to zero. Therefore, substituting (A8) into (A7) and setting the derivative of (A7) with respect to adjusted divergence at an individual level N equal to zero yields

$$\frac{1}{\sigma_N^2} (D_N^* - D_N) \Delta h_N - 2\lambda = 0. \quad (\text{A9})$$

Summing (A9) up to level L and solving for the Lagrangian multiplier yields

$$2\lambda = \frac{\sum_{k=1}^{k=L} (D_k^* - D_k) \Delta h_k}{\sum_{k=1}^{k=L} \sigma_k^2}. \quad (\text{A10})$$

Substituting (A10) into (A9) and solving for the adjusted divergence:

$$D_N^* = D_N + \frac{1}{\Delta h_N} \frac{\sigma_N^2}{\sum_{k=1}^{k=L} \sigma_k^2} \left(\sum_{k=1}^{k=L} D_k^* \Delta h_k - \sum_{k=1}^{k=L} D_k \Delta h_k \right). \quad (\text{A11})$$

Substituting (A5) and (A6) into (A11) and simplifying yields

$$D_N^* = D_N + \frac{1}{\Delta h_N} \frac{\sigma_N^2}{\sum_{k=1}^{k=L} \sigma_k^2} (\rho_L \dot{h}_L - \rho_L \dot{h}_{TL}). \quad (\text{A12})$$

To achieve the maximum adjustment at the target level L , it is necessary to model the error variances as being approximately proportional to height or the height index:

$$\sigma_N^2 = \beta N. \quad (\text{A13})$$

Summing for all levels,

$$\sum_{k=1}^{k=L} \sigma_k^2 = \beta \sum_{k=1}^{k=L} k = \beta \frac{L(L+1)}{2}. \quad (\text{A14})$$

Substituting (A13) and (A14) into (A12) and simplifying gives the equation for adjusted divergence at level N necessary to meet a target vertical velocity:

$$D_N^* = D_N + \frac{1}{\Delta h_N} \left[\frac{2N}{L(L+1)} \right] (\rho_L \dot{h}_L - \rho_L \dot{h}_{TL}). \quad (\text{A15})$$

In the actual application of (A15), the quantity $2N/[L(L+1)]$ is actually applied as a scaling parameter

for the adjustment. Essentially, it models that the original error in the divergence field is maximized near the vertical velocity target height and reduces as distance from the target height increases. Because the adjustment is applied over a specified layer within any model column, levels relative to prescribed adjustment boundaries are introduced within the scaling parameter. This yields the final form for adjusted divergence:

$$D_N^* = D_N + \frac{1}{\Delta h_N} \left[\frac{2N^*}{L^*(L^*+1)} \right] (\rho_L \dot{h}_L - \rho_L \dot{h}_{TL}), \quad (\text{A16})$$

where N^* is the current level index relative to the adjustment boundaries, and L^* is the target level index relative to the adjustment boundaries.

b. Divergence adjustment boundary conditions

For every model column where there is disagreement between the model and GOES cloud fields, the divergence adjustment is applied in two parts. First, (A16) is applied from the surface to the prescribed target height. The described analytical method determines the target height location L and the target vertical velocity $\dot{h}_{TL}$. From the initial model output, it is possible to determine the original divergence at any level D_N , the layer thickness Δh_N , the density at the target height ρ_L , and the original vertical velocity at the target height $\dot{h}_L$. The scaling parameter coefficients for this case are given as

$$N^*(N=1 \rightarrow L) = \begin{cases} N - \text{ADJ}_{\text{BOT}} + 1, & N - \text{ADJ}_{\text{BOT}} > 0 \\ 0, & N - \text{ADJ}_{\text{BOT}} \leq 0 \end{cases}, \quad (\text{A17})$$

$$L^* = L - \text{ADJ}_{\text{BOT}} + 1, \quad (\text{A18})$$

where N is the current level index relative to the surface, L is the target height level index relative to the surface, and ADJ_{BOT} is the bottom adjustment level determined by the developed analytical method. This shows that the scaling parameter is maximized at $N = L$ and reduces to zero as $N \rightarrow \text{ADJ}_{\text{BOT}}$. Therefore, no

adjustment occurs below the prescribed bottom adjustment height, while the maximum adjustment occurs at the target level.

In the second case, the adjustment is applied from one level above the target height to the model top. Based on (A16), the target height L becomes the model top with a prescribed target vertical velocity of zero. All other components are still known from the initial model output, but the scaling parameter coefficients for this case become

$$N^*(N = \text{TL} + 1 \rightarrow L) = \begin{cases} \text{ADJ}_{\text{TOP}} - N + 1, & \text{ADJ}_{\text{TOP}} - N > 0 \\ 0, & \text{ADJ}_{\text{TOP}} - N \leq 0 \end{cases}, \quad (\text{A19})$$

$$L^* = \text{ADJ}_{\text{TOP}} - \text{TL} + 2, \quad (\text{A20})$$

where ADJ_{TOP} is the top adjustment level, and TL is the vertical velocity target level determined by the developed analytical method. The +2 in (A20) comes from starting at one model level above TL. This shows that the scaling parameter is maximized at $N = \text{TL} + 1$ and reduces to zero as $N \rightarrow \text{ADJ}_{\text{TOP}}$. Therefore, no adjustment occurs above the prescribed top adjustment height while the maximum adjustment occurs near the prescribed vertical velocity target level.

REFERENCES

- Anthes, R. A., 1974: Data assimilation and initialization of hurricane prediction models. *J. Atmos. Sci.*, **31**, 702–719, [https://doi.org/10.1175/1520-0469\(1974\)031<0702:DAAIOH>2.0.CO;2](https://doi.org/10.1175/1520-0469(1974)031<0702:DAAIOH>2.0.CO;2).
- Blando, J. D., and B. J. Turpin, 2000: Secondary organic aerosol formation in cloud and fog droplets: A literature evaluation of plausibility. *Atmos. Environ.*, **34**, 1623–1632, [https://doi.org/10.1016/S1352-2310\(99\)00392-1](https://doi.org/10.1016/S1352-2310(99)00392-1).
- Chen, Y., H. Wang, J. Min, X.-Y. Huang, P. Minnis, R. Zhang, J. Haggerty, and R. Palikonda, 2015: Variational assimilation of cloud liquid/ice water path and its impact on NWP. *J. Appl. Meteor. Climatol.*, **54**, 1809–1825, <https://doi.org/10.1175/JAMC-D-14-0243.1>.
- Cintineo, R., J. A. Otkin, M. Xue, and F. Kong, 2014: Evaluating the performance of planetary boundary layer and cloud microphysical parameterization schemes in convection-permitting ensemble forecasts using synthetic GOES-13 satellite observations. *Mon. Wea. Rev.*, **142**, 163–182, <https://doi.org/10.1175/MWR-D-13-00143.1>.
- Diak, G. R., and C. Gautier, 1983: Improvements to a simple physical model for estimating insolation from GOES data. *J. Climate Appl. Meteor.*, **22**, 505–508, [https://doi.org/10.1175/1520-0450\(1983\)022%3C0505:ITASPM%3E2.0.CO;2](https://doi.org/10.1175/1520-0450(1983)022%3C0505:ITASPM%3E2.0.CO;2).
- Diamond, H. J., and Coauthors, 2013: U.S. Climate Reference Network after one decade of operations: Status and assessment. *Bull. Amer. Meteor. Soc.*, **94**, 485–498, <https://doi.org/10.1175/BAMS-D-12-00170.1>.
- Dudhia, J., 1989: Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.*, **46**, 3077–3107, [https://doi.org/10.1175/1520-0469\(1989\)046<3077:NSOCOD>2.0.CO;2](https://doi.org/10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2).
- Ervens, B., B. J. Turpin, and R. J. Weber, 2011: Secondary organic aerosol formation in cloud droplets and aqueous particles (aqSOA): A review of laboratory, field and model studies. *Atmos. Chem. Phys.*, **11**, 11 069–11 102, <https://doi.org/10.5194/acp-11-11069-2011>.
- Gautier, C., G. Diak, and S. Masse, 1980: A simple physical model to estimate incident solar radiation at the surface from GOES satellite data. *J. Appl. Meteor.*, **19**, 1005–1012, [https://doi.org/10.1175/1520-0450\(1980\)019<1005:ASPMTE>2.0.CO;2](https://doi.org/10.1175/1520-0450(1980)019<1005:ASPMTE>2.0.CO;2).
- Haines, S. L., R. J. Suggs, and G. J. Jedlovec, 2004: The Geostationary Operational Environmental Satellite (GOES) Product Generation System. Tech. Memo NASA/TM-2004-213286, 62 pp., <https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20050019524.pdf>.
- Hoke, J. E., and R. A. Anthes, 1976: The initialization of numerical models by a dynamic-initialization technique. *Mon. Wea. Rev.*, **104**, 1551–1556, [https://doi.org/10.1175/1520-0493\(1976\)104<1551:TIONMB>2.0.CO;2](https://doi.org/10.1175/1520-0493(1976)104<1551:TIONMB>2.0.CO;2).
- Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318–2341, <https://doi.org/10.1175/MWR3199.1>.
- Hoskins, B. J., M. E. McIntyre, and A. W. Robertson, 1985: On the use and significance of isentropic potential vorticity maps. *Quart. J. Roy. Meteor. Soc.*, **111**, 877–946, <https://doi.org/10.1002/qj.49711147002>.
- Jiménez, P. A., J. Dudhia, J. F. González-Rouco, J. Navarro, J. P. Montávez, and E. García-Bustamante, 2012: A revised scheme for the WRF surface layer formulation. *Mon. Wea. Rev.*, **140**, 898–918, <https://doi.org/10.1175/MWR-D-11-00056.1>.
- Jones, T. A., D. J. Stensrud, P. Minnis, and R. Palikonda, 2013: Evaluation of a forward operator to assimilate cloud water path into WRF-DART. *Mon. Wea. Rev.*, **141**, 2272–2289, <https://doi.org/10.1175/MWR-D-12-00238.1>.
- , J. A. Otkin, D. J. Stensrud, and K. Knopfmeier, 2014: Forecast evaluation of an observing system simulation experiment assimilating both radar and satellite data. *Mon. Wea. Rev.*, **142**, 107–124, <https://doi.org/10.1175/MWR-D-13-00151.1>.
- Kain, J. S., 2004: The Kain–Fritsch convective parameterization: An update. *J. Appl. Meteor.*, **43**, 170–181, [https://doi.org/10.1175/1520-0450\(2004\)043<0170:TKCPAU>2.0.CO;2](https://doi.org/10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2).
- Kesselmeier, J., and M. Staudt, 1999: Biogenic volatile organic compounds (VOC): An overview on emission, physiology and ecology. *J. Atmos. Chem.*, **33**, 23–88, <https://doi.org/10.1023/A:1006127516791>.
- Lien, G.-Y., E. Kalnay, T. Miyoshi, and G. J. Huffman, 2016: Statistical properties of global precipitation in the NCEP GFS model and TMPA observations for data assimilation. *Mon. Wea. Rev.*, **144**, 663–679, <https://doi.org/10.1175/MWR-D-15-0150.1>.
- Lim, Y. B., Y. Tan, M. J. Perri, S. P. Seitzinger, and B. J. Turpin, 2010: Aqueous chemistry and its role in secondary organic aerosol (SOA) formation. *Atmos. Chem. Phys.*, **10**, 10 521–10 539, <https://doi.org/10.5194/acp-10-10521-2010>.
- Lin, Y., and K. E. Mitchell, 2005: The NCEP Stage II/IV hourly precipitation analyses: Development and applications. *19th Conf. on Hydrology*, San Diego, CA, Amer. Meteor. Soc., 1.2, https://ams.confex.com/ams/Annual2005/techprogram/paper_83847.htm.
- Lin, Y.-L., R. D. Farley, and H. D. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Climate Appl. Meteor.*, **22**, 1065–1092, [https://doi.org/10.1175/1520-0450\(1983\)022<1065:BPOTSF>2.0.CO;2](https://doi.org/10.1175/1520-0450(1983)022<1065:BPOTSF>2.0.CO;2).
- Lynch, P., 1988: Deducing the wind from vorticity and divergence. *Mon. Wea. Rev.*, **116**, 86–93, [https://doi.org/10.1175/1520-0493\(1988\)116<0086:DTWFVA>2.0.CO;2](https://doi.org/10.1175/1520-0493(1988)116<0086:DTWFVA>2.0.CO;2).
- Ma, L.-M., and Z.-M. Tan, 2009: Improving the behavior of the cumulus parameterization for tropical cyclone prediction: Convection trigger. *Atmos. Res.*, **92**, 190–211, <https://doi.org/10.1016/j.atmosres.2008.09.022>.
- Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.*, **102**, 16 663–16 682, <https://doi.org/10.1029/97JD00237>.
- O'Brien, J. J., 1970: Alternative solutions to the classical vertical velocity problem. *J. Appl. Meteor.*, **9**, 197–203, [https://doi.org/10.1175/1520-0450\(1970\)009<0197:ASTTCV>2.0.CO;2](https://doi.org/10.1175/1520-0450(1970)009<0197:ASTTCV>2.0.CO;2).

- Orlanski, I., 1975: A rational subdivision of scales for atmospheric processes. *Bull. Amer. Meteor. Soc.*, **56**, 527–530.
- Otkin, J. A., 2010: Clear and cloudy sky infrared brightness temperature assimilation using an ensemble Kalman filter. *J. Geophys. Res.*, **115**, D19207, <https://doi.org/10.1029/2009JD013759>.
- , M. C. Anderson, J. R. Mecikalski, and G. R. Diak, 2005: Validation of GOES-based insolation estimates using data from the U.S. Climate Reference Network. *J. Hydrometeorol.*, **6**, 460–475, <https://doi.org/10.1175/JHM440.1>.
- Otte, T. L., 2008: The impact of nudging in the meteorological model for retrospective air quality simulations. Part I: Evaluation against national observation networks. *J. Appl. Meteor. Climatol.*, **47**, 1853–1867, <https://doi.org/10.1175/2007JAMC1790.1>.
- Pour-Biazar, A., and Coauthors, 2007: Correcting photolysis rates on the basis of satellite observed clouds. *J. Geophys. Res.*, **112**, D10302, <https://doi.org/10.1029/2006JD007422>.
- Seaman, N. L., 2000: Meteorological modeling for air-quality assessments. *Atmos. Environ.*, **34**, 2231–2259, [https://doi.org/10.1016/S1352-2310\(99\)00466-5](https://doi.org/10.1016/S1352-2310(99)00466-5).
- Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., <https://doi.org/10.5065/D68S4MVH>.
- Spero, T. L., M. J. Otte, J. H. Bowden, and C. G. Nolte, 2014: Improving the representation of clouds, radiation, and precipitation using spectral nudging in the Weather Research and Forecasting model. *J. Geophys. Res. Atmos.*, **119**, 11 682–11 694, <https://doi.org/10.1002/2014JD022173>.
- Stauffer, D. R., and N. L. Seaman, 1990: Use of four-dimensional data assimilation in a limited-area mesoscale model. Part I: Experiments with synoptic-scale data. *Mon. Wea. Rev.*, **118**, 1250–1277, [https://doi.org/10.1175/1520-0493\(1990\)118<1250:UOFDDA>2.0.CO;2](https://doi.org/10.1175/1520-0493(1990)118<1250:UOFDDA>2.0.CO;2).
- Stull, R. B., 1988: *An Introduction to Boundary Layer Meteorology*. Kluwer Academic, 666 pp.
- Tewari, M., and Coauthors, 2004: Implementation and verification of the unified Noah land surface model in the WRF Model. *20th Conf. on Weather Analysis and Forecasting/16th Conf. on Numerical Weather Prediction*, Seattle, WA, Amer. Meteor. Soc., 14.2a, https://ams.confex.com/ams/84Annual/techprogram/paper_69061.htm.
- Tie, X., S. Madronich, S. Walters, R. Zhang, P. Rasch, and W. Collins, 2003: Effect of clouds on photolysis and oxidants in the troposphere. *J. Geophys. Res.*, **108**, 4642, <https://doi.org/10.1029/2003JD003659>.
- van der Veen, S. H., 2013: Improving NWP model cloud forecasts using meteosat second-generation imagery. *Mon. Wea. Rev.*, **141**, 1545–1557, <https://doi.org/10.1175/MWR-D-12-00021.1>.
- Yucel, I., W. J. Shuttleworth, X. Gao, and S. Sorooshian, 2003: Short-term performance of MM5 with cloud-cover assimilation from satellite observations. *Mon. Wea. Rev.*, **131**, 1797–1810, <https://doi.org/10.1175/2565.1>.
- Zhang, S. Q., M. Zupanski, A. Y. Hou, X. Lin, and S. H. Cheung, 2013: Assimilation of precipitation-affected radiances in a cloud-resolving WRF ensemble data assimilation system. *Mon. Wea. Rev.*, **141**, 754–772, <https://doi.org/10.1175/MWR-D-12-00055.1>.