A Regional Hybrid Gain Data Assimilation System and Preliminary Evaluation Based on Radio Occultation Reflectivity Assimilation

Chih-Chien Chang[1], Shu-Chih Yang[1, 2], and Stephen G. Penny[3, 4]

[1]*Department of Atmospheric Sciences*, *National Central University*, *Taoyuan*, *Taiwan*
[2]*RIKEN Center for Computational Science*, *Kobe*, *Japan*
[3]*Cooperative Institute for Research in Environmental Sciences*, *University of Colorado Boulder*, *Colorado*, *USA*
[4]*NOAA Physical Sciences Laboratory*, *Boulder*, *Colorado*, *USA*

Abstract

A regional hybrid gain data assimilation (HGDA) system is newly developed using Weather Research and Forecasting model (WRF). The WRF-HGDA augments an ensemble-based Kalman filter (WRF-LETKF) with information from the variational analysis system (WRF-3DVAR) by combining their gain matrices. The performance of WRF-HGDA is evaluated by assimilating the GNSS radio occultation (RO) observations from the FORMOSAT-3/COSMIC (FS3/C) and the FORMOSAT-7/COSMIC2 (FS7/C2) under an Observing System Simulation Experiment (OSSE) framework. The results demonstrate that the variational correction improves the WRF-LETKF, with the equal-weighted WRF-HGDA outperforming its component DA systems in the moisture and wind fields when only conventional observations are assimilated. Assimilating additional RO data from FS7/C2 further improves the WRF-LETKF and WRF-HGDA systems. Although the variational correction for the mid-level temperature causes degradation in the WRF-HGDA analysis, this can be alleviated by adjusting the combination weight to include more flow-dependent information in WRF-HGDA at these levels. Further tuning of the static background error covariance used in WRF-3DVAR also brings some improvement in the WRF-HGDA wind analysis. Our results imply that a well-tuned variational system is critical for the accuracy of the regional HGDA analysis.

(**Citation**: Chang, C.-C., S.-C. Yang, and S. G. Penny, 2022: A regional hybrid gain data assimilation system and preliminary evaluation based on radio occultation reflectivity assimilation. *SOLA*, **18**, 33−40, doi:10.2151/sola.2022-006.)

1. Introduction

Hybrid data assimilation (HDA) has become the state-of-the-art approach for operational numerical weather prediction (NWP). Many HDA frameworks have been proposed to introduce flow-dependent background error covariance (**B**) estimated by a set of ensemble forecasts into a static climatological **B** as is used in the canonical form of the variational (VAR) methods (Lorenc 2003; Buehner 2005; Wang et al. 2007). The blending of different **B** matrices is typically implemented as a weighted sum of the two covariance terms (Hamill and Snyder 2000) and named as hybrid covariance data assimilation (HCDA). An alternative hybrid concept was proposed by Penny (2014) that applied the blending to the gain matrices. The hybrid gain data assimilation (HGDA) combines the gain matrix from an ensemble-based Kalman filter (EnKF) and a variational method. Because the gain matrix includes as components the **B** and observation error covariance (**R**) matrices, HGDA can further take the **R** into consideration in the hybridization. HGDA has been tested with various numerical models of different complexity, ranging from simple dynamics (Penny 2014; Chang et al. 2020; Azevedo et al. 2020) to large-scale models used for operational forecasts of atmosphere and ocean dynamics (Penny et al. 2015; Bonavita et al. 2015; Houtekamer et al. 2019, 2021). These studies have shown that HGDA can improve the analysis accuracy compared to either of the two original analysis methods used alone.

For regional data assimilation, studies have mainly focused on HCDA (Zhang and Zhang 2012; Poterjoy and Zhang, 2015; 2016). In this study, a regional HGDA system is newly developed to combine the WRF-LETKF and the WRF-3DVAR systems with the purpose of improving regional NWP by leveraging advantages from both systems.

One of the great challenges for improving local severe weather prediction is the accuracy of the initial moisture analysis (Yang et al. 2014). Among the space-borne observation platforms measuring moisture information, the Global Navigation Satellite System (GNSS) radio occultation (RO) observation has become an important observation type. Studies assimilating RO observations have shown positive impacts on tropical cyclone prediction (Huang et al. 2005, 2010; Chen et al. 2018, 2020) and the prediction of heavy rainfall (Yang et al. 2014; Huang et al. 2016; Chang and Yang 2022). The recently launched FORMOSAT-7/COSMIC2 (hereafter FS7/C2) provides about 5,000 RO profiles per day (Schreiner et al. 2020). In addition, FS7/C2 has a better penetration rate than the FORMOSAT-3/COSMIC (hereafter FS3/C; Chen et al. 2021). Lien et al. (2021) confirm that assimilating FS7/C2 observations has significant impacts on the operational global NWP system at the Central Weather Bureau.

Given the importance of the moisture analysis, this study aims to understand the characteristics of the newly developed WRF-HGDA in the context of RO assimilation and the inheritance of contributions from the component DA systems. We investigate how the additional adjustment derived from the WRF-HGDA system can modify the impact of assimilating GNSS-RO observations through the hybridization of the gain matrices, in particular with more RO profiles available from FS7/C2. The results of this study aim to improve understanding of the characteristics of HGDA implemented in a regional model and the impacts of assimilating observations produced by the latest RO mission. The sensitivity experiments may also be useful as a guide for adjusting the variational DA system used in HGDA. Given that WRF-HGDA is a newly developed system, the investigation is first conducted under the Observing System Simulation Experiment (OSSE) framework.

The remainder of this article is organized as follows. The methodology and the experimental setup are described in Section 2. Results are presented and discussed in Section 3, and a summary is given in Section 4.

2. Methodology

2.1 Hybrid Gain Data Assimilation (HGDA)

HGDA forms a hybrid combination of gain matrices derived from arbitrary sources. It was first proposed by Penny (2014) using EnKF and VAR components. Penny (2014) provided two implementations for conducting the HGDA and it was termed as

Corresponding author: Shu-Chih Yang, Department of Atmospheric Sciences, National Central University, 300 Jhong-Da Rd., JhongLi 32001, Taiwan. E-mail: shuchih.yang@atm.ncu.edu.tw.

scenarios a and b in Chang et al. (2020). Scenario a is a two-step *sequential* update that conceptually uses the VAR to adjust the analysis mean state of the EnKF, while scenario b is a *parallel* update that mirrors the traditional HCDA approaches to use the same background in VAR and EnKF and is more computationally economical. This study adopts scenario a as the HGDA and focuses on examining how the variational optimization improves the analysis mean state that has been updated by EnKF.

The hybridization is performed by blending the solutions of the VAR analysis ($\mathbf{x}_a^{\mathrm{VAR}}$) and the EnKF analysis mean state ($\bar{\mathbf{x}}_a^{\mathrm{EnKF}}$) with a given combination weight, α,

$$\bar{\mathbf{x}}_a^{\mathrm{HGDA}} = (1-\alpha)\bar{\mathbf{x}}_a^{\mathrm{EnKF}} + \alpha\mathbf{x}_a^{\mathrm{VAR}}. \quad (1)$$

The VAR analysis is obtained by taking the EnKF analysis mean state as the background field in HGDA scenario a. Substituting the background state and analysis increment form into the analysis solution (Eq. 1), we have

$$\bar{\mathbf{x}}_a^{\mathrm{HGDA}} = \bar{\mathbf{x}}_b + [\mathbf{K}^{\mathrm{EnKF}} + \alpha\mathbf{K}^{\mathrm{VAR}}(\mathbf{I} - \mathbf{H}\mathbf{K}^{\mathrm{EnKF}})](\mathbf{y}^{\mathrm{o}} - \mathbf{H}\bar{\mathbf{x}}_{\mathbf{b}}), \quad (2)$$

where $\bar{\mathbf{x}}_b$ is the background ensemble mean state, $\mathbf{y}^{\mathrm{o}}$ is the observation vector, $\mathbf{H}$ is the observation operator, and $\mathbf{K}^{\mathrm{EnKF}}$ and $\mathbf{K}^{\mathrm{VAR}}$ are the gain matrices obtained from the EnKF, and the VAR, respectively. By defining the hybrid gain matrix ($\widehat{\mathbf{K}}$) as:

$$\widehat{\mathbf{K}} = [\mathbf{K}^{\mathrm{EnKF}} + \alpha\mathbf{K}^{\mathrm{VAR}}(\mathbf{I} - \mathbf{H}\mathbf{K}^{\mathrm{EnKF}})], \quad (3)$$

the hybrid analysis state is simply expressed as:

$$\bar{\mathbf{x}}_a^{\mathrm{HGDA}} = \bar{\mathbf{x}}_b + \widehat{\mathbf{K}}(\mathbf{y}^{\mathrm{o}} - \mathbf{H}\bar{\mathbf{x}}_{\mathbf{b}}). \quad (4)$$

Equation (3) demonstrates how the HGDA combines the gain matrices. More detailed derivations are given in Penny (2014).

At each cycle, the HGDA re-centers the analysis ensemble perturbation ($\mathbf{X}_a'^{\mathrm{EnKF}}$) at the updated hybrid analysis state:

$$\mathbf{X}^{\mathrm{HGDA}} = \bar{\mathbf{x}}_a^{\mathrm{HGDA}}\mathbf{v}^T + \mathbf{X}_a'^{\mathrm{EnKF}}, \quad (5)$$

where $\mathbf{v}$ is a column vector of ones. Since the HGDA is developed from the EnKF perspective, the re-centering process treats the hybrid analysis state as the EnKF analysis mean state to be centered upon, which forms a hybridized analysis ensemble ($\mathbf{X}^{\mathrm{HGDA}}$) for conducting the next DA cycle. Such a re-centering of an EnKF analysis mean toward a more accurate mean state improves the representation of the dynamical uncertainties that are used to form the $\mathbf{B}$ matrix (Yang et al. 2012; Chang et al. 2014). Thus, the HGDA not only adjusts the EnKF analysis mean state, but also indirectly improves the estimated $\mathbf{B}$ matrix through the ensemble evolution of the EnKF components as the DA process is cycled.

2.2 Experimental configuration

a. DA systems

Three DA systems are implemented here with the regional WRF (Weather Research and Forecasting) model (Skamarock et al. 2008): WRF-3DVAR (Barker et al. 2012), WRF-LETKF (Yang et al. 2014), and WRF-HGDA. The fundamental parameters used in WRF-3DVAR are tuned for this experiment. For instance, the static climatological $\mathbf{B}$ matrix is generated using the NMC (National Meteorological Center) method (Parrish and Derber 1992) with the CV5 control variables (i.e. the streamfunction, unbalanced velocity potential, unbalanced surface pressure, unbalanced temperature, and pseudo relative humidity). To adequately reflect the characterization of RO observations, the error variance and the characteristic length-scale related to the temperature and moisture fields are tuned to optimize the performance of WRF-3DVAR, based on the assimilation of conventional and FS3/C observations. The WRF-LETKF employs 36 ensemble members with a horizontal localization radius of 250 km and a multiplicative inflation of 8%. An equal combination weight (i.e. α = 0.5) is chosen for this preliminary WRF-HGDA study.

b. OSSE configuration

An OSSE is conducted to evaluate the impacts of assimilating the RO observations with the WRF-HGDA. During the experiment period, there were three hurricanes (Helene, Gordon, and Florence) that occurred over the Atlantic Ocean. To generate the nature run, a 5-day ensemble forecast with 50 members is initialized at 1200 UTC 10 September 2006 in order to cover the early lifetime of Hurricane Helene. The ensemble member that has minimum accumulated track error for Hurricane Helene is selected as the nature run. To account for model errors, the nature run is simulated with the WRF V3.7.1 configuration while the experimental simulations use the WRF V3.2.1 configuration. These imposed model differences act as a surrogate for the true model errors, and increase the uncertainties of the representativeness of the numerical simulations. Both the nature run and experimental simulation have a grid spacing of 27 km and 31 vertical layers.

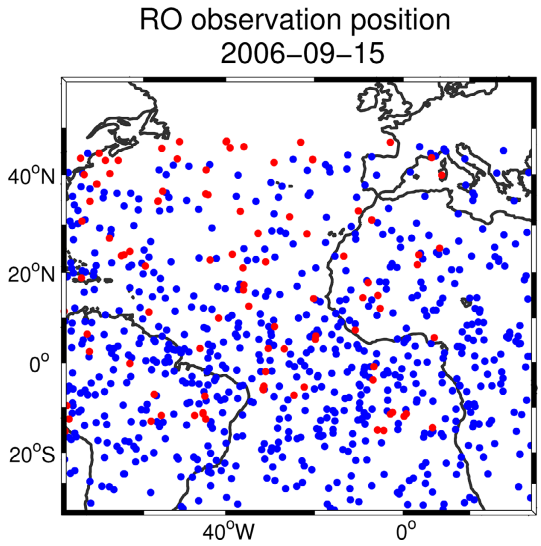

Fig. 1. The distribution of FS3/C (red points) and FS7/C2 (blue points) RO observations in the experimental domain.

Synthetic observations are generated by mapping the nature run states to the desired observed variables and then adding random noise which is sampled from a standard Gaussian distribution multiplied by the observation error. The observations assimilated in this study include conventional GTS (Global Telecommunication System) observations (includes the synoptic stations, radiosondes, ships, buoys, and aircraft reports) and RO refractivity observations from FS3/C and FS7/C2. The locations of the synthetic GTS and FS3/C RO observations are taken from real data, while the FS7/C2 observation locations are obtained from an orbit simulator.

c. Experimental setup

Three experiments were conducted with different sets of observations. The first experiment only assimilates the synthetic conventional GTS observations via different DA systems. The purpose of this experiment is to understand the general performance of each DA system. The second and the third experiments assimilate the conventional GTS observations and additionally with the FS3/C observations and the FS7/C2 observations, respectively. Figure 1a shows the distribution of the FS3/C and FS7/C2 observations. The initial ensemble was generated by adding perturbations that are randomly drawn from the static background error covariance used in the WRF-3DVAR (Torn et al. 2006) to the NCEP (National Centers for Environmental Prediction) FNL reanalysis. After a 3-day ensemble forecast to spin-up the ensemble perturbations, a 2-day assimilation cycle from 1200 UTC 13 September to 1200 UTC 15 September was performed with a 6-h interval. Rather than analyzing any particular hurricane case, this study examines the overall characteristics of the DA systems with the focus on RO assimilation.

Although the 2-day assimilation seems to be short for the DA system to reach its asymptotic level of performance, the fundamental features of WRF-HGDA are well exhibited. The flow-

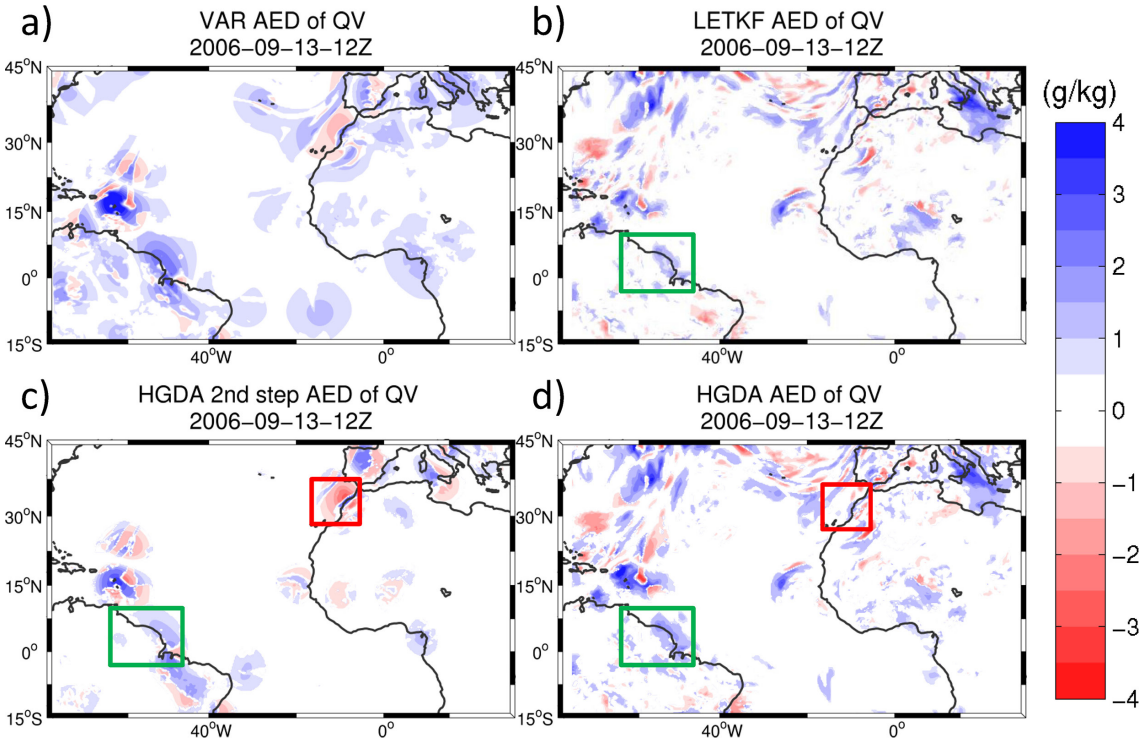

Fig. 2. The AED of horizontal water vapor (QV) analysis field at 700 hPa of (a) the stand-alone WRF-3DVAR, (b) the stand-alone WRF-LETKF (i.e. the first step update of WRF-HGDA), (c) the second step update of WRF-HGDA, and (d) the WRF-HGDA.

dependency has been established in the ensemble-based **B** matrix within the 2-day period, such as the dynamic uncertainties of the hurricane development. Despite the fact that the resolution is not high enough to represent the detailed structure of TC, the model state during this period is able to capture the evolution of TCs over the Atlantic well enough to investigate how to use the information from the VAR with a static **B** and the EnKF with the dynamic-evolved **B** for the hybrid data assimilation.

3. Results

3.1 General performance of DA systems with conventional observations

The performance of a DA system can be evaluated using a statistical quantity, the absolute error difference (AED) between the analysis (Ana) and the background states (Bgd):

$$AED = abs(Bgd - Truth) - abs(Ana - Truth), \quad (6)$$

where 'Truth' is provided by the nature run. A positive AED indicates improvement. Figure 2 shows the AED of moisture field at 700 hPa with conventional observation assimilation. The stand-alone WRF-3DVAR provides an efficient moisture correction in the first DA cycle (Fig. 2a) while the WRF-LETKF produces flow-dependent corrections with a shorter horizontal length-scale due to the localization. Given that the same background state for the first DA cycle, the WRF-LETKF produces an equivalent result to the first step update of the WRF-HGDA. We note that the variational correction at the second step shows a positive AED around the coast of South America (the green box in Fig. 2c), and thus the positive AED of WRF-HGDA (Fig. 2d) extends over a bigger region with a larger magnitude than the WRF-LETKF. However, the WRF-3DVAR also introduces inaccuracies (the red box in Fig. 2c) into the WRF-HGDA. We note that the moisture is tricky to handle in a variational system with the static **B** and the moisture analysis increment may contaminate the WRF-3DVAR analysis (Bannister et al. 2019). As mentioned by Chang et al. (2020), the combination weight used in the WRF-HGDA helps mitigate the invalid variational correction, resulting in a milder degradation. Hence, with more DA cycles, the WRF-HGDA has the smallest domain-averaged root mean square error (RMSE) with respect to the nature run. For instance, the domain-averaged RMSE of moisture at 950 hPa of WRF-3DVAR, WRF-LETKF, and WRF-HGDA are 1.471 g Kg^{-1}, 1.340 g Kg^{-1}, and 1.313 g Kg^{-1} in the last DA cycle, respectively. The performance of WRF-LETKF is close to that of WRF-HGDA, while the WRF-3DVAR is the worst performer.

3.2 Impacts of RO observation

RO refractivity is a function of temperature, moisture and pressure and is not a direct observation of a single prognostic model variable. A single observation experiment was conducted to understand the fundamental characteristics of assimilating an RO refractivity observation for each DA system. The refractivity innovation is one unit of refractivity larger than the background at 900 hPa and the location is chosen to be within the circulation of the Hurricane Helene. Compared with the moisture correction of WRF-LETKF (Fig. 3d), the WRF-3DVAR (Fig. 3a) produces a broader and stronger correction that dominates the moisture correction of the WRF-HGDA (Fig. 3g). The amplitude of the temperature correction resulting from WRF-LETKF (Fig. 3e) is larger than that of WRF-3DVAR (Fig. 3b), thus the WRF-HGDA (Fig. 3h) gains a smaller temperature adjustment from the WRF-3DVAR. In WRF-3DVAR, the background error correlation between zonal wind and refractivity is weak (Fig. 3c), so the WRF-3DVAR obtains limited corrections from the RO assimilation. In contrast, the WRF-LETKF has a strong cross-variable error correlation between these variables owing to the high dynamical uncertainties around TC. Figure 3f suggests that assimilating the refractivity can enhance the cyclonic structure of Hurricane Helene in the WRF-LETKF, and the WRF-HGDA (Fig. 3i) has also inherited this capability. Since the characteristics of assimilating the refractivity observation with DA systems are similar in the zonal and meridional winds, only the results for zonal winds are

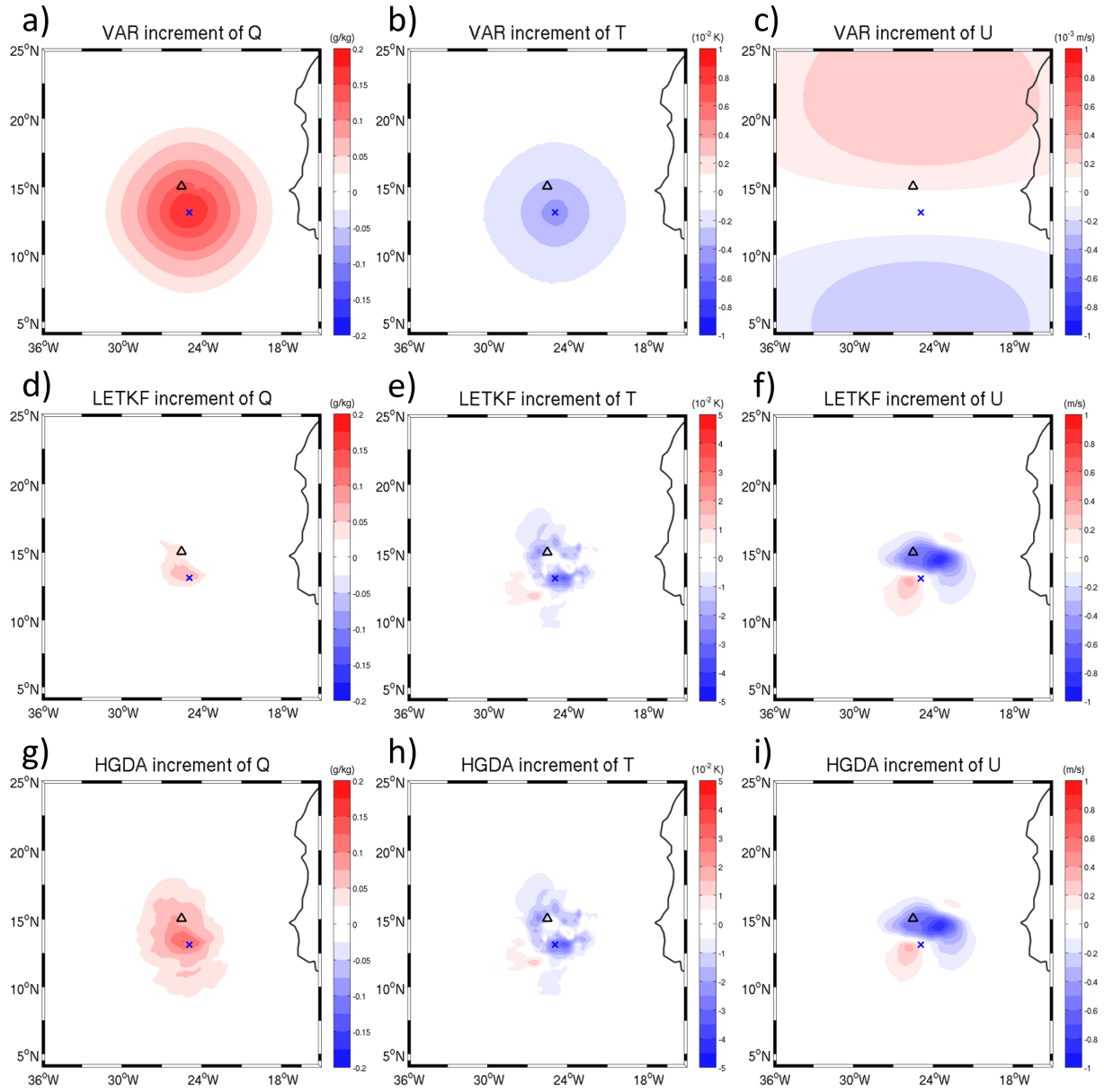

Fig. 3. The analysis increment produced by assimilating the pseudo refractivity observation at 900 hPa. The panels from top to bottom display results using WRF-3DVAR, WRF-LETKF, and WRF-HGDA. The panels from left to right display water vapor (Q), temperature (T), and zonal wind (U). The blue cross and the black triangle indicate the location of pseudo observation and the eye of Hurricane Helene in the background state, respectively. Note that while the analysis increments for Q and T are in general agreement in sign between the WRF-3DVAR and WRF-LETKF, the analysis increments are quite different in sign and structure for the zonal wind.

shown.

Next, cycled DA experiments are conducted. In addition to the conventional observations, the FS3/C and the FS7/C2 experiments assimilate the synthetic RO observations from the FS3/C and the FS7/C2, respectively. Figure 4 shows the domain-averaged RMSE of the FS3/C and the FS7/C2 experiment at the first and fifth DA cycles. The performance of the stand-alone WRF-3DVAR assimilating either FS3/C observations or FS7/C2 observations is comparable in general. This may be partially attributed to the static **B** matrix which is well-tuned for the stand-alone WRF-3DVAR with only GTS and FS3/C observations assimilated. Considering that the moisture distribution is inhomogeneous, the use of a static **B** matrix can introduce degradation in the moisture analysis (e.g. Fig. 2a) due to the homogeneity assumption applied in modeling a static **B**. Given this, assimilating more RO profiles available from FS7/C2 may not be as beneficial for the moisture analysis. Also, the static **B** matrix may need to be tuned for the assimilation of FS7/C2 observations. Both the WRF-LETKF and the WRF-HGDA show positive impacts on the moisture and wind fields when FS7/C2 observations is assimilated and this advantage becomes greater when assimilating more RO observations (FS3/C vs. FS7/C2) at later DA cycles. However, the temperature field has different behavior. At the lowest level, the WRF-LETKF has the worst performance. We suspect this could be due to the limited ensemble spread, which is constrained by the fixed lower boundary condition such as the unperturbed sea surface temperature. The variational correction of WRF-HGDA provides useful information for mitigating this deficiency in the WRF-LETKF. At the levels between 900−700 hPa and 500−300 hPa, the WRF-LETKF analysis better fits to the nature run, with a smaller RMSE than the WRF-3DVAR. This advantage also becomes clearer when assimilating more RO observations and when conducting more DA

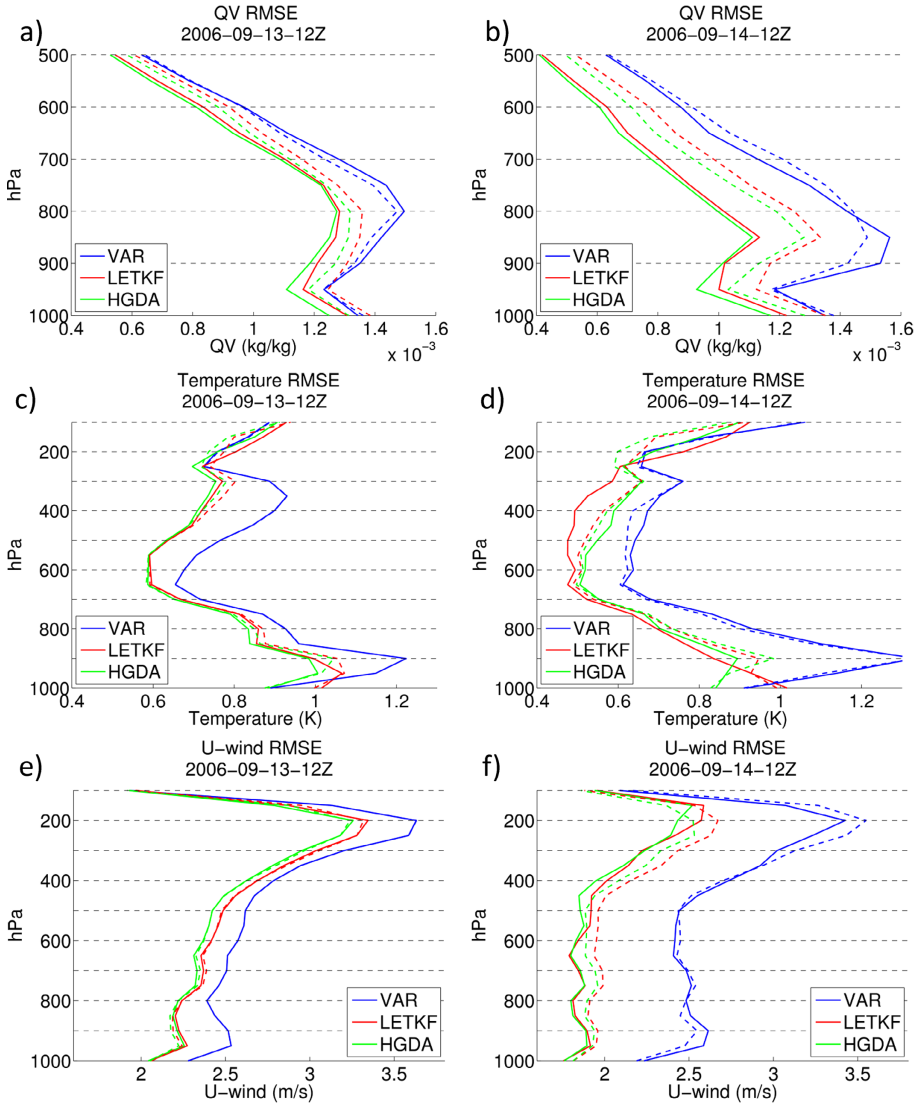

Fig. 4. The domain averaged RMSE of (a) water vapor (QV) at the first DA cycle, and (b) the fifth DA cycle. (c) and (d), (e) and (f) are the same as (a) and (b) but for the temperature and zonal wind fields, respectively. The blue, red, and green lines show the result of the WRF-3DVAR, the WRF-LETKF, and the WRF-HGDA, respectively, assimilating the FS3/C (dashed lines) and FS7/C2 (solid lines) observations. It is noted that the difference between the blue dashed and dotted lines in (c) and (e) are indistinguishable because the RMSE difference is very small when assimilating FS3/C or FS7/C2 observations for the first DA cycle.

cycles (Fig. 4d).

Relative to the WRF-LETKF, the larger temperature RMSE between 900 to 300hPa presented in the WRF-HGDA is further explained by Fig. 5. Figure 5 shows the AED of temperature field in the fifth DA cycle of FS7/C2 assimilation at 700 hPa. At this time, both the stand-alone WRF-3DVAR and the WRF-LETKF generally provide valid correction to the background error, improving the accuracy in temperature. However, for the WRF-HGDA, the inaccurate variational correction (Fig. 5c) that is potentially made by the improper static **B** used in WRF-3DVAR shows not only in areas where the stand-alone WRF-3DVAR has degradation, but also in areas where WRF-LETKF provides a weak improvement, such as the African continent. In other words, although the variational correction indeed provides corrections in areas that WRF-LETKF barely corrects, most of the corrections lead to degradation. Even though the combination weight plays the role of relaxing the invalid variational adjustment back toward the WRF-LETKF analysis (red box in Figs. 5c and 5d), the invalid variational adjustment still degrades the WRF-HGDA relative to WRF-LETKF between 900 to 300 hPa.

3.3 Results of the sensitivity experiments with WRF-HGDA

As mentioned in the previous section, the WRF-HGDA gains clear improvement in the moisture, the wind, and the low-level temperature fields, while there is a degradation of temperature analysis between 900 to 300 hPa. To further understand the potential of the WRF-HGDA, two sensitivity experiments are conducted to modify the static **B** matrix and adjust the combination weight. The FS7/C2 experiment using the default setup is named as the control run.

Chang et al. (2020) pointed out that the performance of HGDA is sensitive to the static **B** matrix. Although we have tuned both the background error variance and horizontal length-scale of the temperature and moisture variables, which are the variables directly related to the RO refractivity, we noticed that WRF-3DVAR has a much larger RMSE in the wind field (Fig. 4). Since the moisture transport depends on the dynamic processes, the accuracy of the wind field can affect the impact of RO assimilation. For improving the performance of WRF-HGDA, it is not sufficient to tune the static **B** matrix only for the variables related to the refractivity. Hence, a sensitivity test is performed with the modifications on the characteristic length-scale and variances of static **B** matrix to

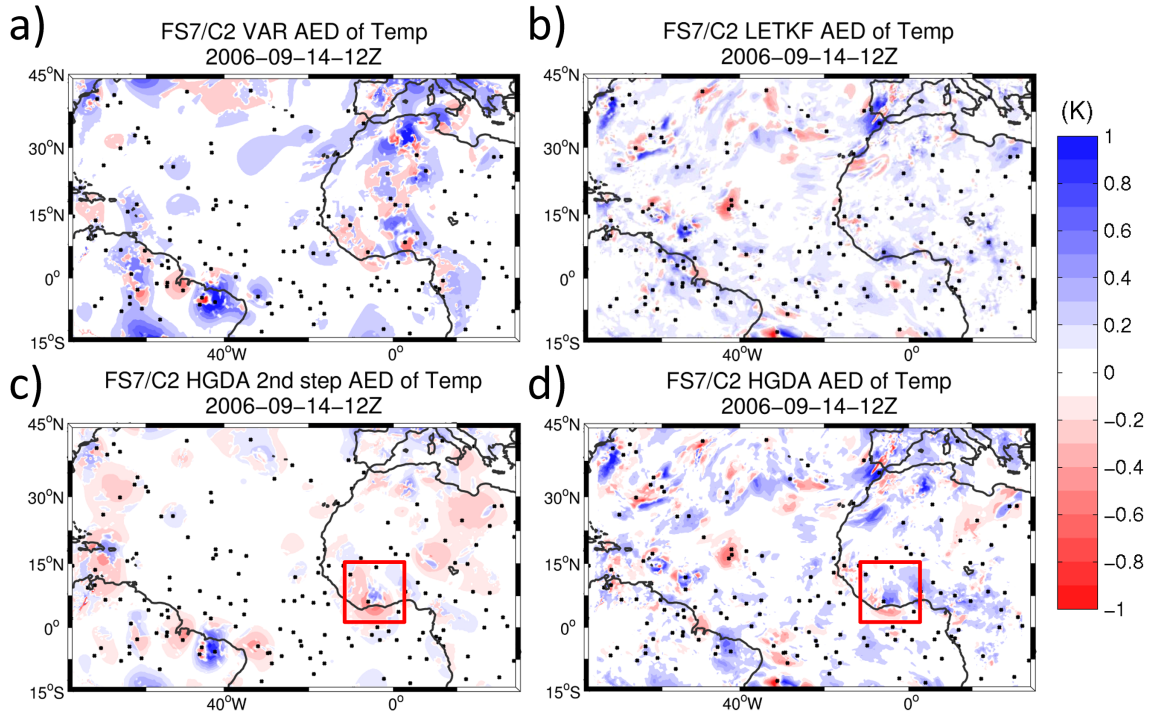

Fig. 5. The AED of temperature analysis field at 700 hPa at the fifth DA cycle when additionally assimilating FS7/C2 observations for: (a) the stand-alone WRF-3DVAR, (b) the stand-alone WRF-LETKF, (c) the variational correction at the second step update of WRF-HGDA, and (d) the total adjustment of WRF-HGDA.

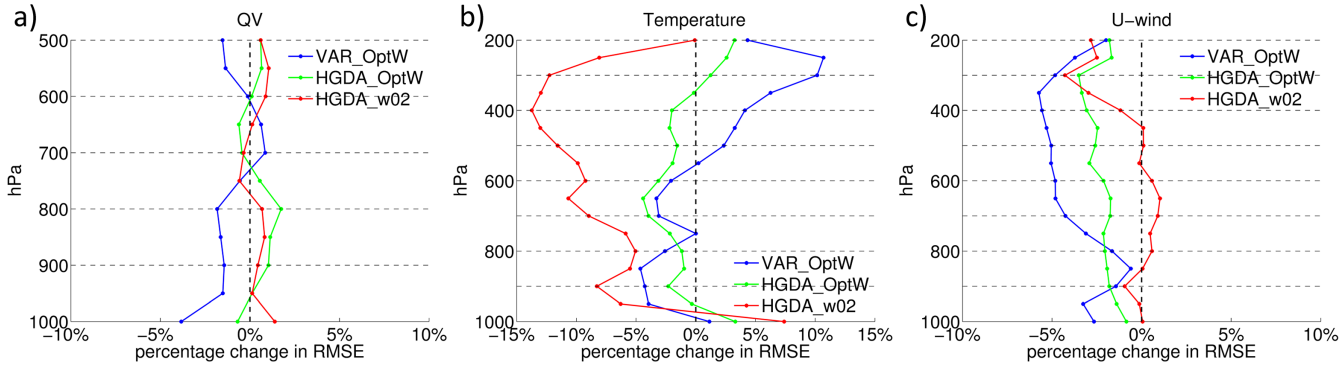

Fig. 6. The percentage change in averaged RMSE of sensitivity experiment and control run in the last four DA cycles for a) moisture, b) temperature, and c) zonal wind fields. The negative value indicates a positive impact on the changes of sensitivity experiment. The blue (green) line represents the WRF-3DVAR (WRF-HGDA) using the static **B** matrix optimized with considering the wind component. The red line represents the sensitivity experiment of the WRF-HGDA with $\alpha = 0.2$.

consider not only the temperature and moisture related control variables but also the control variables related to the horizontal wind fields. The tuning is also based on the FS3/C assimilation with stand-alone WRF-3DVAR. Experiments using this new tuned static **B** matrix are named with 'OptW' (optimization with wind component).

Figure 6 shows the percentage change in averaged RMSE of the sensitivity experiments in the last four DA analysis cycles relative to the control run. A negative value indicates a positive impact obtained from the sensitivity experiment. Given the significant improvement of VAR_OptW, the HGDA_OptW inherits the benefit of an improved variational analysis in the wind field (Fig. 6c). However, the new tuned static **B** matrix cannot provide a further improvement on temperature above 500 hPa and slightly degrades the performance on the moisture field. This result demonstrates the importance of the static **B** matrix, which is non-trivial to optimize for the purpose of improving the WRF-HGDA. It is worth pointing out that such impact should not be attributed to the use of RO observations directly since the correlation between refractivity and zonal wind background errors is weak (Fig. 3c).

Another drawback for the HDA is that the combination weight, which affects the performance of the hybrid system severely, is determined empirically. To illustrate the sensitivity to the combination weight, the HGDA_w02 experiment decreases the combination weight from the default value ($\alpha = 0.5$) to $\alpha = 0.2$ and thus a stronger weightage is placed on the flow-dependent component of WRF-HGDA. Another purpose is to understand how to mitigate the invalid variational correction in the temperature field. Despite that the equal-weighted WRF-HGDA outperforms the HGDA_w02 at the first analysis cycle (not shown), HGDA_w02 leads to an improvement in temperature as the DA cycling progresses. This reflects the fact that the correction from EnKF becomes more useful when the flow dependency of the ensemble-based B is well built (Caya et al. 2005; Kalnay and Yang 2010). Nevertheless,

HGDA_w02 leads to an improvement in temperature as the DA cycling progresses. The degradation of WRF-HGDA caused by the variational adjustment at the mid-levels is reduced in the HGDA_w02. Nevertheless, the issue of insufficient ensemble spread at the lowest level unavoidably becomes severe with more information from the WRF-LETKF. Results of adjusting the combination weight suggest the WRF-HGDA can be improved if the weight could be varied with the altitude, especially in the temperature correction, for example potentially following Azevedo et al. (2020).

4. Conclusion

A newly developed regional DA system, the WRF-HGDA, is established and its potential is examined within an OSSE framework. The HGDA combines the gain matrices from component DA systems rather than forming a hybrid combination of only the **B** matrices. In this study, the WRF-HGDA is composed of the component systems WRF-3DVAR and WRF-LETKF. The purposes of this study are to (1) improve the understanding of the characteristics of the WRF-HGDA and the impacts by assimilating different RO datasets (the FS3/C and the FS7/C2), and (2) examine the sensitivity of WRF-HGDA via adjusting the static **B** matrix used in WRF-3DVAR.

Results show that the variational adjustment of the WRF-HGDA improves the analysis of moisture and wind fields compared with the WRF-LETKF. Assimilating more RO data in the FS7/C2 experiment further reduces the RMSE of those variables with the WRF-HGDA and WRF-LETKF systems. With the help from WRF-3DVAR, the WRF-HGDA outperforms the WRF-LETKF in the temperature fields at the low levels (below 900 hPa) owing to the limited ensemble spread constrained by the boundary conditions. However, for the temperature field at the mid-levels, the variational correction is improper, and thus the WRF-LETKF performs best. This degradation of WRF-HGDA can be attributed to the invalid variational adjustment, primarily due to inaccuracies in the static **B** matrix, and changing the combination weight alleviates such negative impacts. A further tuning of the static **B** matrix by considering the temperature, moisture, and wind components improves the impact of the stand-alone WRF-3DVAR and the WRF-HGDA on the wind field, but the impact on other variables is not significant.

This study examines the performance of HGDA in a regional model. The benefits of WRF-HGDA mainly depend on the effectiveness of the variational correction, which is closely associated with the static **B** matrix. It implies that a well-tuned WRF-3DVAR implemented with the WRF-HGDA is a key component for improving the regional HDA system. The importance of the DA components with a static **B** or a flow-dependent **B** may vary with the model resolution. As noted in Feng and Wang (2021), they only use the full flow-dependent **B** in the innermost domain of a nested setup without combining the static **B** in the hybrid EnVAR system for a TC study with a cloud-resolving model because the static **B** results in a negative impact as the resolution is increased. Therefore, how to optimize the use of component DA systems through the combination weight is crucial to not only the HCDA but also the HGDA. To further improve the WRF-HGDA, other methods should be examined, such as applying a spatially varied weight (Azevedo et al. 2020) or removing the need for a combination weight by using an orthogonal update (Chang et al. 2020).

Acknowledgement

This study is supported by the Ministry of Science and Technology of Taiwan under Grant MOST-110-2121-M-008 -007 and the National Space Organization (NSPO) of Taiwan under Grant NSPO-S-110316. C.-C. Chang is supported by the Taiwan Central Weather Bureau (CWB) Grant MOTC-CWB-110-M-08 and the Ministry of Science and Technology of Taiwan under Grant MOST 110-2811-M-008-561. S.G.P. acknowledges support from NOAA grants NA20OAR4600277, NA19NES4320002, and NA18NWS4680048, and the Office of Naval Research (ONR) grants N00014-19-1-2522 and N00014-20-1-2580.

Edited by: T.-Y. Koh

Reference

Azevedo, H. B., L. G. G. De Gonçalves, E. Kalnay, and M. Wespetal, 2020: Dynamically weighted hybrid gain data assimilation: Perfect model testing. *Tellus, Ser. A Dyn. Meteor. Oceanogr.*, **72**, 1−11, doi:10.1080/16000870.2020.1835310.

Barker, D., X.-Y. Huang, Z. Liu, T. Auligné, X. Zhang, S. Rugg, R. Ajjaji, A. Bourgeois, J. Bray, Y. Chen, M. Demirtas, Y.-R. Guo, T. Henderson, W. Huang, H.-C. Lin, J. Michalakes, S. Rizvi, and X. Zhang, 2012: The Weather Research and Forecasting model's community variational/ensemble data assimilation system: WRFDA. *Bull. Amer. Meteor. Soc.*, **93**, 831−843, doi:10.1175/BAMS-D-11-00167.1.

Bannister, R. N., H. G. Chipilski, and O. Martinez-Alvarado, 2019: Techniques and challenges in the assimilation of atmospheric water observations for numerical weather prediction towards convective scales. *Quart. J. Roy. Meteor. Soc.*, **146**, 1−48, doi:10.1002/qj.3652.

Bonavita, M., M. Hamrud, and L. Isaksen, 2015: EnKF and hybrid gain ensemble data assimilation. Part II: EnKF and hybrid gain results. *Mon. Wea. Rev.*, **143**, 4865−4882, doi:10.1175/MWR-D-15-0071.1.

Buehner, M., 2005: Ensemble-derived stationary and flow-dependent background-error covariances: Evaluation in a quasi-operational NWP setting. *Quart. J. Roy. Meteor. Soc.*, **131**, 1013−1043, doi:10.1256/qj.04.15.

Caya, A., J. Sun, and C. Snyder, 2005: A comparison between the 4DVAR and the ensemble Kalman Filter techniques for radar data assimilation. *Mon. Wea. Rev.*, **133**, 3081−3094, doi:10.1175/MWR3021.1.

Chang, C.-C., and S.-C. Yang, 2022: Impact of assimilating Formosat-7/COSMIC-II GNSS radio occultation data on heavy rainfall prediction in Taiwan. *Terr. Atmos. Ocean. Sci.* (accepted).

Chang, C.-C., S.-C. Yang, and C. Keppenne, 2014: Applications of the mean recentering scheme to improve typhoon track prediction: A case study of Typhoon Nanmadol (2011). J. *Meteor. Soc. Japan. Ser. II*, **92**, 559−584, doi:10.2151/jmsj.2014-604.

Chang, C.-C., S. G. Penny, and S.-C. Yang, 2020: Hybrid gain data assimilation using variational corrections in the subspace orthogonal to the ensemble. *Mon. Wea. Rev.*, **148**, 2331−2350, doi:10.1175/MWR-D-19-0128.1.

Chen, S.-Y., H. Zhao, and C.-Y. Huang, 2018: Impacts of GNSS radio occultation data on predictions of two super-intense typhoons with WRF hybrid variational-ensemble data assimilation. *J. Aeronaut. Astronaut. Aviat.*, **50**, 347−364, doi:10.6125/JoAAA.201812_50(4).02.

Chen, S.-Y., Y.-H. Kuo, and C.-Y. Huang, 2020: The impact of GPS ro data on the prediction of tropical cyclogenesis using a nonlocal observation operator: An initial assessment. *Mon. Wea. Rev.*, **148**, 2701−2717, doi:10.1175/MWR-D-19-0286.1.

Feng, J., and X. G. Wang, 2021: Impact of increasing horizontal and vertical resolution during the HWRF hybrid EnVar data assimilation on the analysis and prediction of Hurricane Patricia (2015). *Mon. Wea. Rev.*, **149**, 419−441, doi:10.1175/MWR-D-20-0144.1.

Hamill, T. M., and C. Snyder, 2000: A hybrid ensemble Kalman Filter–3D variational analysis scheme. *Mon. Wea. Rev.*, **128**, 2905−2919, doi:10.1175/1520-0493(2000)128<2905:AHEKFV>2.0.CO;2.

Houtekamer, P. L., M. Buehner, and M. De La Chevrotière, 2019: Using the hybrid gain algorithm to sample data assimilation uncertainty. *Quart. J. Roy. Meteor. Soc.*, **145**, 35−56, doi:

10.1002/qj.3426.

Houtekamer, P. L., B. He, D. Jacques,R. McTaggart-Cowan, L. Separovic, P. A. Vaillancourt, A. Zadra, and X. Deng, 2021: Use of a genetic algorithm to optimize a numerical weather prediction system. *Mon. Wea. Rev.*, **149**, 1089−1104, doi: 10.1175/MWR-D-20-0238.1.

Huang, C.-Y., Y.-H. Kuo, S.-H. Chen, and F. Vandenberghe, 2005: Improvements in typhoon forecasts with assimilated GPS occultation refractivity. *Wea. Forecasting*, **20**, 931−953, doi:10.1175/WAF874.1.

Huang, C.-Y., Y.-H. Kuo, S.-Y. Chen, C.-T. Terng, F.-C. Chien, P.-L. Lin, M.-T. Kueh, S.-H. Chen, M.-J. Yang, C.-J. Wang, and S. K. A. V. P. Rao Anisetty, 2010: Impact of GPS radio occultation data assimilation on regional weather predictions. *GPS Solut.*, **14**, 35−49, doi:10.1007/s10291-009-0144-1.

Huang, C.-Y., S.-Y. Chen, S. K. A. V. P. Rao Anisetty, S.-C. Yang, and L.-F. Hsiao, 2016: An impact study of GPS radio occultation observations on frontal rainfall prediction with a local bending angle operator. *Wea. Forecasting*, **31**, 129−150, doi:10.1175/WAF-D-15-0085.1.

Kalnay, E., and S. C. Yang, 2010: Accelerating the spin-up of ensemble Kalman filtering. *Quart. J. Roy. Meteor. Soc.*, **136**, 1644−1651, doi:10.1002/qj.652.

Lien, G.-Y., C.-H. Lin, Z.-M. Huang, W.-H. Teng, J.-H. Chen, C.-C. Lin, H.-H. Ho, J.-Y. Huang, J.-S. Hong, C.-P. Cheng, and C.-Y. Huang, 2021: Assimilation impact of early FORMOSAT-7/COSMIC-2 GNSS radio occultation data with Taiwan's CWB Global Forecast System. *Mon. Wea. Rev.*, **149**, 2171−2191, doi:10.1175/mwr-d-20-0267.1.

Lorenc, A. C., 2003: The potential of the ensemble Kalman filter for NWP - A comparison with 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **129**, 3183−3203, doi:10.1256/qj.02.132.

Penny, S. G., 2014: The hybrid local ensemble transform Kalman filter. *Mon. Wea. Rev.*, **142**, 2139−2149, doi:10.1175/MWR-D-13-00131.1.

Penny, S. G., D. W. Behringer, J. A. Carton, and E. Kalnay, 2015: A hybrid global ocean data assimilation system at NCEP. *Mon. Wea. Rev.*, **143**, 4660−4677, doi:10.1175/MWR-D-14-00376.1.

Poterjoy, J., and F. Zhang, 2015: Systematic comparison of four-dimensional data assimilation methods with and without a tangent linear model using hybrid background error covariance: E4DVar versus 4DEnVar. *Mon. Wea. Rev.*, **143**, 1601−1621, doi:10.1175/MWR-D-14-00224.1.

Poterjoy, J., and F. Zhang, 2016: Comparison of hybrid four-dimensional data assimilation methods with and without the tangent linear and adjoint models for predicting the life cycle of Hurricane Karl (2010). *Mon. Wea. Rev.*, **144**, 1449−1468, doi:10.1175/MWR-D-15-0116.1.

Schreiner, W. S., and co-authors, 2020: COSMIC-2 radio occultation constellation: First results. *Geophys. Res. Lett.*, **47**, 1−7, doi:10.1029/2019GL086841.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. Barker, M. G. Duda, X.-Y. Huang, W. Wang, and J. G. Powers, 2008: *A Description of the Advanced Research WRF Version 3* (No. NCAR/TN-475+STR). University Corporation for Atmospheric Research, doi:10.5065/D68S4MVH.

Torn, R. D., G. J. Hakim, and C. Snyder, 2006: Boundary conditions for limited-area ensemble Kalman filters. *Mon. Wea. Rev.*, **134**, 2490−2502, doi:10.1175/MWR3187.1.

Wang, X., C. Snyder, and T. M. Hamill, 2007: On the theoretical equivalence of differently proposed ensemble - 3DVAR hybrid analysis schemes. *Mon. Wea. Rev.*, **135**, 222−227, doi:10.1175/MWR3282.1.

Yang, S.-C., E. Kalnay, and B. Hunt, 2012: Handling nonlinearity in an ensemble Kalman filter: Experiments with the three-variable Lorenz model. *Mon. Wea. Rev.*, **140**, 2628−2646, doi:10.1175/MWR-D-11-00313.1.

Yang, S.-C., S.-H. Chen, S.-Y. Chen, C.-Y. Huang, and C.-S. Chen, 2014: Evaluating the impact of the COSMIC RO bending angle data on predicting the heavy precipitation episode on 16 June 2008 during SoWMEX-IOP8. *Mon. Wea. Rev.*, **142**, 4139−4163, doi:10.1175/MWR-D-13-00275.1.

Zhang, M., and F. Zhang, 2012: E4DVar: Coupling an ensemble Kalman filter with four-dimensional variational data assimilation in a limited-area weather prediction model. *Mon. Wea. Rev.*, **140**, 587−600, doi:10.1175/MWR-D-11-00023.1.

Manuscript received 20 October 2021, accepted 3 January 2022
SOLA: *https://www.jstage.jst.go.jp/browse/sola/*