

## Article

# Evaluation of the Four-dimensional Ensemble-Variational Hybrid Data Assimilation with Self-consistent Regional Background Error Covariance for Improved Hurricane Intensity Forecasts

Shixuan Zhang <sup>1,†</sup> , Zhaoxia Pu <sup>2,†</sup> 

<sup>1</sup> Pacific Northwest National Laboratory, Richland, WA 99354, USA

<sup>2</sup> Department of Atmospheric Sciences, University of Utah, Salt Lake City, UT 84112, USA

\* Correspondence: shixuan.zhang@pnnl.gov; zhaoxia.pu@utah.edu

† These authors contributed equally to this work.

Version August 13, 2020 submitted to Atmosphere

**Abstract:** The feasibility of a hurricane initialization framework based on the GSI-based four-dimensional ensemble-variational (GSI-4DEnVar) hybrid data assimilation system for the Hurricane Weather Research and Forecasting model (HWRF) model is evaluated in this study. The system considers the temporal evolution of error covariances via the use of four-dimensional ensemble perturbations that are provided by high-resolution, self-consistent HWRF ensemble forecasts. It is different from the configuration of the GSI-based three-dimensional ensemble-variational (GSI-3DEnVar) hybrid data assimilation system, similar to that used in the operational HWRF, which employs background error covariances provided by coarser-resolution global ensembles from the NCEP GFS ensemble Kalman filtering data assimilation system. In addition, our proposed initialization framework discards the empirical intensity correction in the vortex initialization package that is employed by the GSI-3DEnVar initialization framework in operational HWRF. Data assimilation and numerical simulation experiments for Hurricanes Joaquin (2015), Patricia (2015), and Matthew (2016) are conducted during their intensity changes. The impacts of two initialization frameworks on the HWRF analyses and forecasts are compared. It is found that GSI-4DEnVar leads to a reduction in track, minimum sea level pressure (MSLP), and maximum surface wind (MSW) forecast errors in all of the HWRF simulations, compared with the GSI-3DEnVar initialization framework. With assimilating high-resolution observations within the hurricane inner-core region, GSI-4DEnVar can produce reasonably well hurricane initial intensity without the empirical vortex intensity correction. Further diagnoses with Hurricane Joaquin indicate that GSI-4DEnVar can significantly alleviate the imbalances in the initial conditions and enhance the performance of the data assimilation and subsequent hurricane intensity and precipitation forecasts.

**Keywords:** GSI-3DEnVar; GSI-4DEnVar; HWRF; Hurricane intensity changes; Background error covariance

## 1. Introduction

The intensity forecasts for tropical cyclones (TCs) are of great importance for providing earlier warnings of TC-related damages [1,2]. However, accurate intensity forecasts, especially during TC intensity changes, remain a great challenge in the current operational and research community [3–7]. For instance, the official 48-h intensity forecast error only decreased by about 5% – 10% between 2000 and 2016 (see, e.g., [www.nhc.noaa.gov/verification/verify5.shtml](http://www.nhc.noaa.gov/verification/verify5.shtml)). Part of the difficulty in

improving intensity forecasts of TCs originates from uncertainties in the initial conditions provided for TC forecasting models [4,10]. At the U. S. National Centers for Environmental Prediction (NCEP), the Hurricane Weather Research and Forecasting (HWRF) model [11] has operationally provided real-time TC forecasts to the National Hurricane Center (NHC) for the Atlantic and eastern North Pacific basins since 2007 and was extended to all oceanic basins over the globe in 2014. The operational HWRF initialization adopts a vortex initialization (VI) package [12] and a Gridpoint Statistical Interpolation (GSI) based three-dimensional ensemble-variational (3DEnVar) hybrid data assimilation (DA) system [13,14]. The VI package is based on a dynamical initialization scheme that first relocates the vortex to its observed location, followed by a size and intensity correction process with adjustments to the wind, moisture, and thermodynamics fields [15]. After that, DA is performed to incorporate the available conventional, radar, and satellite data into the first guess from VI with GSI-3DEnVar, where the flow-dependent background error covariance is derived from a set of GFS EnKF ensemble forecasts [15]. This further improves the vortex structure and environment conditions for HWRF TC forecasts. The combination of VI and the GSI-3DEnVar DA system has substantially contributed to improvements in the track and intensity forecasts of HWRF, including those during changes in TC intensity [11].

Despite the significant progress in improving NCEP operational hurricane forecasts with the HWRF model and its data assimilation system, problems associated with the VI and DA system in HWRF have also been reported in recent studies. First, the VI in HWRF can lead to an unrealistic initial vortex inner-core structure [4] and vortex spin-down problems (e.g., a drop of 10 ~20 knots in the maximum surface wind in the first few hours of a forecast) in HWRF [10] in some cases. Second, the VI and DA in HWRF have been shown to counteract each other in some cases [16]. Third, the current GSI-3DEnVar system does not account for the temporal evolution of error covariance within the assimilation window [17], and the resolution of the GFS ensemble covariance is too coarse to properly resolve the error covariance of the TC inner core [10,15]. Due to these deficiencies, in some cases the initial conditions for HWRF are suboptimal, leading to the degradation of track and intensity forecasts, especially during changes in TC intensity [29]. Thus, resolving these problems is essential in order to further improve the forecasts of TC intensity changes in HWRF. Recent studies have explored ways to overcome these problems. [10] used high-resolution, self-consistent regional HWRF ensembles instead of a global ensemble to derive the flow-dependent background error covariances in GSI-3DEnVar for HWRF, resulting in a more realistic hurricane initial inner-core structure and leading to improved hurricane forecasting. In addition, the GSI-based four-dimensional ensemble-variational hybrid DA system (GSI-4DEnVar) has been developed for operational TC forecast models [15,17,18]. In GSI-4DEnVar, ensemble perturbations at multiple time periods spanning the assimilation window are used to incorporate the temporal evolution of the background error covariances into the DA system. Moreover, GSI-4DEnVar is computationally cheap relative to the traditional 4DVar DA system because it avoids using the tangent linear model and adjoint model [17,18]. Studies have shown that GSI-4DEnVar outperforms GSI-3DEnVar in producing more realistic TC analyses and can further improve TC intensity forecasts [15,17,18]. More importantly, recent studies have found that the assimilation of high-resolution inner-core observations (e.g., High-Definition Sounding System dropsonde data) based on an GSI-4DEnVar with a high-resolution background error covariance can outperform the VI scheme currently used in HWRF [30].

In light of these recent developments, this study assesses a revised initialization framework (referred to as "4DEnVar initialization framework" hereafter) with the following key changes in the initialization framework that is smiler to that used in the operational HWRF (referred to as "3DEnVar initialization framework" hereafter):

- the GSI-4DEnVar instead of GSI-3DEnVar is used to incorporate the real observations and generate the the initial conditions for HWRF forecast to consider the temporal evolution of error covariance within the assimilation window following the studies of [15], [17] and [18] ;
- the self-consistent high-resolution regional HWRF ensembles instead of the coarser-resolution GFS ensembles are used to derive the flow-dependent error covariance for the GSI hybrid data

assimilation system to properly resolve the error covariance of the TC inner core following the study in [10];

- the size and intensity correction process with adjustments to the wind, moisture, and thermodynamics fields are removed from the VI package to avoid the unrealistic initial vortex inner-core structure and vortex spin-down problems as pointed by [4] and [10].

The performance of the 4DEnVar initialization framework in terms of forecasting changes in TC intensity is compared with the 3DEnVar initialization framework. The answers to two major questions are investigated: first, whether the 4DEnVar initialization framework can at least partially resolve the problems associated with VI and DA in the 3DEnVar initialization framework, as mentioned above, and second, whether and why the 4DEnVar initialization framework outperforms the 3DEnVar initialization framework in terms of forecasts of TC intensity change in HWRF. These questions are answered by evaluating the track and intensity forecasts for three hurricane cases during their intensity changes.

The paper is organized as follows: Section 2 illustrates the methodology for the 3DEnVar and 4DEnVar initialization frameworks for HWRF. The study cases and experiment design are introduced in Section 3. Section 4 compares HWRF analyses and forecasts of Hurricane Joaquin (2015) during its rapid weakening phases with the two different initialization frameworks. Section 5 presents additional case studies for the track and intensity forecasts of Hurricanes Patricia (2015) and Matthew (2016) during their rapid intensification periods. Concluding remarks and a discussion are given in Section 6.

## 2. Methodology

### 2.1. HWRF model

HWRF Version 3.7a (V3.7a), which is functionally equivalent to the 2016 operational HWRF, is employed in this study [11,19]. The dynamical core used in HWRF V3.7a is the same as that in NCEP's WRF-Nonhydrostatic Mesoscale Model (NMM) [20]. The NMM core in HWRF V3.7a adopts a two-way interactive, movable, triply nested grid procedure. The atmospheric component of the NMM core in HWRF V3.7a is formulated with 61 vertical levels and a model top at 2 hPa. In addition, a suite of advanced physical parameterizations developed for TC applications is also employed. As indicated by [11], these schemes include the Geophysical Fluid Dynamics Laboratory (GFDL) surface-layer parameterization, the Noah Land Surface Model, the modified GFDL shortwave and longwave radiation scheme, the Ferrier-Aligo microphysical parameterization, the Global Forecast System (GFS) Planetary Boundary Layer (PBL) scheme, and the GFS simplified Arakawa-Schubert (SAS) cumulus scheme. The cumulus parameterization is used in the 18 km and 6 km resolution domains, but not in the 2 km horizontal grid spacing domain. A more detailed description of HWRF V3.7a can be found in [11].

### 2.2. DA schemes

In the 3DEnVar initialization framework, GSI-3DEnVar is used to incorporate the observations and improve HWRF initial conditions. In this system, the analysis increment is obtained by minimizing a cost function:

$$J(\delta\mathbf{x}) = \frac{1}{2}(\delta\mathbf{x})^T \left( \beta_1^{-1}\mathbf{B}_1 + \beta_2^{-1}(\mathbf{B}_2 \circ \mathbf{S}) \right)^{-1} (\delta\mathbf{x}) + \frac{1}{2}(\mathbf{y}' - \mathbf{H}[\delta\mathbf{x}])^T \mathbf{R}^{-1} (\mathbf{y}' - \mathbf{H}[\delta\mathbf{x}]) \quad (1)$$

Where  $\delta\mathbf{x} = \mathbf{x} - \mathbf{x}_b$  is the analysis increment,  $\mathbf{y}' = \mathbf{y} - \mathbf{H}[\delta\mathbf{x}]$  is the observation innovation.  $\mathbf{H}$  is a forward model or transferring operator, and  $\mathbf{R}$  is an observational error covariance matrix. The final analysis is  $\mathbf{x}_a = \mathbf{x}_b + \delta\mathbf{x}$ . The first term on the right-hand side is the background term with hybrid covariance  $\mathbf{B} = \beta_1^{-1}\mathbf{B}_1 + \beta_2^{-1}(\mathbf{B}_2 \circ \mathbf{S})$ . Namely, the background error covariance information is a combination of two sources: a static, pregenerated matrix  $\mathbf{B}_1$  and a flow-dependent matrix  $\mathbf{B}_2$  that is derived from the ensemble forecast. The Schur product with correlation matrix  $\mathbf{S}$  implies localization.

The second term is the observational term, and the total analysis increment  $\delta\mathbf{x}$  is associated with both static and flow-dependent background error covariances.  $\beta_1$  and  $\beta_2$  are weighting factors, but their inverse defines the weights placed on  $B_1$  and  $B_2$ , respectively. In practice,  $\beta_1^{-1} + \beta_2^{-1} = 1$  is satisfied. Detailed information and implications of GSI-3DEnVar can be found in Wang [22].

Recently, the NCEP operational GSI-3DEnVar system has been extended to include four-dimensional (4D) ensemble perturbations (GSI-4DEnVar). The analysis increment in GSI-4DEnVar is obtained by minimizing a cost function:

$$J(\delta\mathbf{x}) = \frac{1}{2}(\delta\mathbf{x})^T \left( \beta_1^{-1}\mathbf{B}_1 + \beta_2^{-1}(\mathbf{B}_2 \circ \mathbf{S}) \right)^{-1} (\delta\mathbf{x}) + \frac{1}{2} \sum_{t=1}^m (\mathbf{y}'_t - \mathbf{H}_t[\delta\mathbf{x}_t])^T \mathbf{R}_t^{-1} (\mathbf{y}'_t - \mathbf{H}_t[\delta\mathbf{x}_t]) \quad (2)$$

where the third term is extended to use the asynchronous observations up to  $m$  time levels compared with equation (1), and the other terms are the same as those in equation (1). The increment  $\delta\mathbf{x}_t$  is defined as:

$$\delta\mathbf{x}_t = (\delta\mathbf{x}) + \mathbf{T} \sum_{t=1}^m (\mathbf{a}_k \circ (\mathbf{x}_k^e)_t) \quad (3)$$

where  $\mathbf{a}$  is the extended control variable [22], which defines the localization applied to the ensemble perturbation  $(\mathbf{x}_k^e)_t$  (only in the diagonal direction).  $\mathbf{T}$  is a matrix that transfers the ensemble perturbations from the ensemble grid space into the model grid space. Equation (3) indicates that the increments in GSI-4DEnVar are prescribed through linear combinations of the 4D ensemble perturbations plus the contribution from static background error covariance. Thus, the tangential linear (TL) model and adjoint (AD) model in the traditional 4DVar are avoided. Detailed explanations of the GS-4DEnVar algorithm can be found in Wang and Lei [17], Kleist and Ide [18], Buehner *et al.* [23] and Lorenc *et al.* [24]. In addition, the GSI-4DEnVar system has been applied in the NCEP operational GFS system since 2015, while it has not been used in the operational HWRF.

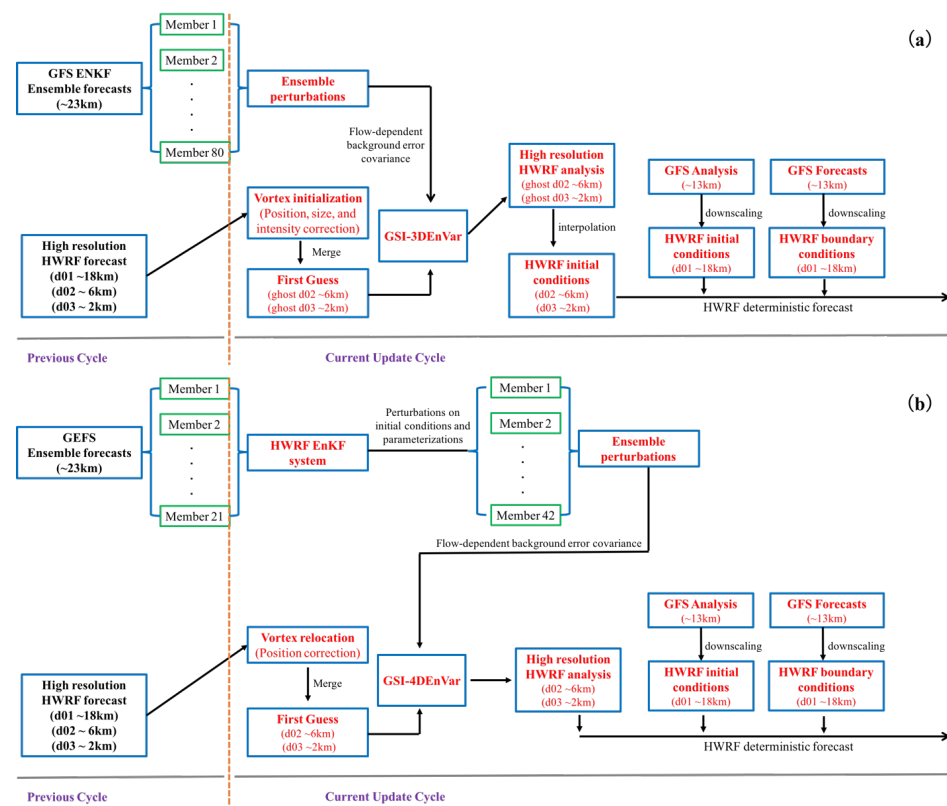
### 2.3. HWRF initialization frameworks

In the GSI-3DEnVar settings, the HWRF initialization framework is a combination of a VI package and the GSI-3DEnVar DA system. This setting is similar to the DA used in operational HWRF. As shown in Figure 1a, the key procedures involved in this framework include:

- VI is applied on the previous cycle's HWRF forecasts in the d01 (~18 km), d02 (~6 km), and d03 (~2 km) domains (see Figure 2a) to provide the first guess for DA. The VI enables the relocation, resizing, and intensity correction of the vortex using the NHC tropical cyclone vital statistics (TCVitals) database in order to correct the storm position and intensity approach to the real-time estimation.
- After VI, the improved initial conditions in d01, d02, and d03 are merged into two large domains, ghost d02 (~6 km) and ghost d03 (~2 km) (see Fig. 2a). The ghost d02 and ghost d03 domains are an extension of forecast domains d02 and d03, respectively. The size of ghost d02 (ghost d03) is about two times larger than that of d02 (d03). The goal of this approach is to include more observations in the GSI-3DEnVar DA system for HWRF and to obtain a more realistic analysis near the domain boundaries of d02 and d03.
- Observations available for the operational HWRF are then assimilated into ghost d02 and ghost d03 by the GSI-3DEnVar system to further improve the initial conditions for the HWRF forecast. The 6-h NCEP operational GFS 80-member ensemble forecast at a resolution of T574 (~23 km) from the previous analysis cycle is used to provide the flow-dependent background error covariance for GSI-3DEnVar within a 6-h DA window.

We note that GSI-3DVar does not consider the temporal evolution of error covariances.

- After DA, the model fields in ghost d02 and ghost d03 are then interpolated back into the d02 and d03 domains to form the initial conditions of these two domains. The initial conditions in d01 are downscaled from GFS real-time analysis in the current HWRf analysis cycle. In addition, a “blending” scheme is activated after the VI and DA processes. This scheme blends the vortex from the VI process with the vortex after DA and use the blended vortex in the final analysis for the HWRf forecast. The effect of this scheme will eliminate the DA increments within 150 km of the center below 400 hPa. The details for this procedure can be referred to [11]. Moreover, the GFS forecasts are used to provide the boundary conditions for the HWRf model. With all of above information, the deterministic HWRf forecast during a specific period (e.g. 126 h) is made.



**Figure 1.** Flow charts for the HWRf simulations based on (a) the 3DVar initialization framework and (b) 4DVar initialization framework.

In this study, we evaluate the 4DVar initialization framework for HWRf. The differences between the 4DVar and 3DVar initialization frameworks are listed in Table 1. In addition, as shown in Figure 1b, the HWRf simulation with this initialization framework contains the following steps:

- A parallel run of the HWRf regional ensemble system [21] initialized by the forecasts from the GEFS is performed first to generate a 42-member self-consistent HWRf ensemble at 6-km resolution. Here, the self-consistent means that the ensemble forecasts use the same resolution and model configurations as the HWRf model. The detailed treatments and parameters used for HWRf ensemble forecasts follow those in Zhang *et al.* [21]. These 42-member self-consistent HWRf ensembles (6 km) are then used to provide the flow-dependent background error covariances for the GSI-4DVar DA

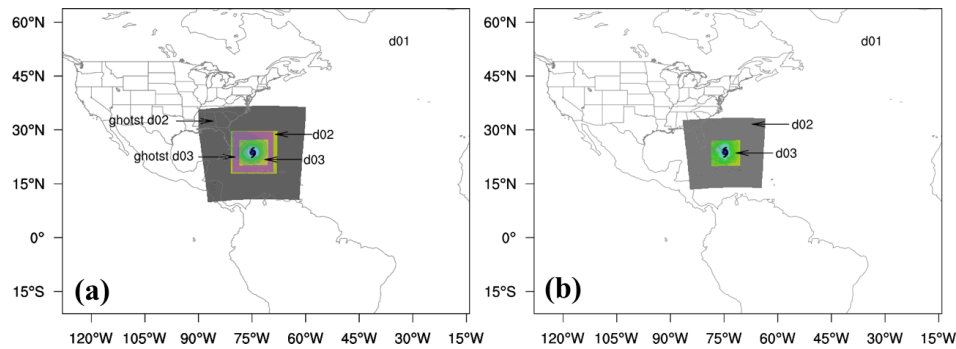


**Table 1.** Differences between the 3DEnVar and the 4DEnVar initialization frameworks for HWRF.

Configuration	HWRF 3DEnVar initialization framework	HWRF 4DEnVar initialization framework
DA scheme	GSI-3D-EnVar	GSI-4D-EnVar
Flow-dependent error covariance	GFS 80-member ensembles (~23 km)	HWRF 42-member ensembles (~6km)
Vortex initialization	Vortex relocation & Intensity correction	Vortex relocation
DA domains	ghost d02 (30° × 30°) ghost d03 (12° × 12°)	d02 (25° × 25°) d03 (7.5° × 9°)
Forecast domains	d01 (80° × 80°) d02 (13° × 13°) d03 (7.5° × 9°)	d01 (80 × 80°) d02 (25 × 25°) d03 (7.5 × 9°)

system in 4DEnVar initialization framework. In this way, the DA system considers the temporal evolution of error covariances via the use of four-dimensional ensemble perturbations that are provided by high-resolution, self-consistent HWRF ensemble forecasts.

- Vortex relocation is performed on the previous cycle’s HWRF forecasts in the d02 (~6 km) and d03 (~2 km) domains (see Figure 2b) before DA. Vortex relocation corrects the storm position based on the NHC tropical cyclone vital statistics (TCVitals) database without changes in storm size and intensity. In addition, the size of d02 in Figure 2b for 4DEnVar initialization framework is about two times larger than that in Figure 2a for the 3DEnVar initialization framework, which is used to include more observations in the GSI-4DEnVar system and to obtain a more realistic analysis in the storm environment. The sizes of d01 and d03 in Figure 2b are the same as those in Figure 2a.
- After vortex relocation, the observations available for operational HWRF are then assimilated directly into d02 and d03 by the GSI-4DEnVar system to further improve the initial conditions in these domains. In GSI-4DEnVar, the DA window is divided into 6 observational bins; the 3-h, 4-h, 5-h, 6-h, 7-h, 8-h, and 9-h self-consistent HWRF 42-member ensemble forecasts from the first step are used to provide the flow-dependent background error covariance, realizing the time evolution of background error covariance within a 6-h DA window. We note that the 1-h observational bins are used in this study for GSI-4DEnVar as the previous study has demonstrated that the performance of GSI-4DEnVar can be enhanced with a denser observational bin [30].
- As in the 3DEnVar initialization framework, the initial conditions in d01 (Figure 2b) for 4DEnVar initialization framework are also downscaled from the GFS real-time analysis in the current HWRF analysis cycle. This is combined with the analyses from the above step for d02 and d03 to provide the initial conditions for the HWRF forecast. The GFS forecasts are still used to provide the boundary conditions. Finally, a deterministic HWRF forecast for a specific time period (e.g., 126 h) can be made.



**Figure 2.** (a) Forecast domains, as indicated by d01, d02, and d03, and HWRF data assimilation domains, as indicated by ghost d02 (black shaded area), and ghost d03 (pink shaded area) for the 3DEnVar initialization framework; and (b) forecast domains and data assimilation domains as indicated by d01, d02 and d03 for 4DEnVar initialization framework. Data assimilation in the 4DEnVar initialization framework is directly performed on d02 and d03. The black symbol indicates the storm center from NHC best track (black storm sign).

### 3. Performances on the simulation of the rapid weakening of Hurricane Joaquin

In this section, the performances of HWRF analyses and forecasts with two different initialization frameworks are compared. Hurricane Joaquin, which occurred over the Atlantic Ocean in the 2015 hurricane season, is selected as a case study. A detailed description of this case can be seen in Berg [25]. Since this study focuses on forecasts of TC intensity changes, the initialization and HWRF simulation are performed on the rapid weakening (RW) phase of Hurricane Joaquin from 1800 UTC 02 to 1800 UTC 06 October 2015. The HWRF model is initialized at 1200 UTC 02 October 2015 and allowed to spin-up until 1800 UTC 02 October 2015, and then a DA process is performed and cycled from 1800 UTC 02 to 1800 UTC 03 October 2015 in 6-h windows, for a total of 24 hours. All datasets available for the operational HWRF (see the list of types at [http://www.emc.ncep.noaa.gov/mmb/data\\_processing/prepbufr.doc/table\\_2.htm](http://www.emc.ncep.noaa.gov/mmb/data_processing/prepbufr.doc/table_2.htm) and [table18.htm](http://www.emc.ncep.noaa.gov/mmb/data_processing/prepbufr.doc/table18.htm)), along with available high-resolution Tail Doppler Radar (TDR) and HDSS dropsonde observations, are assimilated. After DA, a 72-h forecast from 1800 UTC 03 to 1800 UTC 06 October 2015 is then performed to predict the RW of Hurricane Joaquin. We note that the 72-h forecast period is selected as it has already covered the whole intensity changes period for Hurricane Joaquin, thus, we did not run a 5-day simulation following the HWRF routines to save the computational time. Two experiments are conducted: HWRF-OPR1 and HWRF-RVS1. HWRF-OPR1 adopts the 3DEnVar initialization framework (Fig. 1a), while HWRF-RVS1 uses the 4DEnVar initialization framework (Fig. 1b). Table 2 summarizes all experiments and their acronyms.

#### 3.1. Balances in HWRF initial conditions

A previous study has proved that gradient imbalances exist in the initial conditions with the operational HWRF initialization framework in some cases due to the use of VI before DA, leading to the vortex spin-down problems and the degradation of the intensity forecast [10]. Thus, we first investigate whether the 4DEnVar initialization framework can alleviate this problem. Following Pu *et al.* [10], the azimuthally averaged net radial force  $F$ , which represents the difference between the local radial pressure gradient and the sum of the centrifugal and Coriolis forces, is calculated to examine the dynamical balances in the HWRF initial vortices generated by the two different initialization frameworks.

Figures 3a and b show the radius-height cross-sections of  $F$  isopleths before and after VI in HWRF-OPR1 at 1800 UTC 03 October 2015 (analysis time). It clearly reveals the strong supergradient wind fields in the inner-core region of the vortex after the VI procedure (Fig. 3a, b), indicating that VI induces these supergradient wind fields. DA in HWRF-OPR1 does not eliminate the supergradient wind fields, as the distribution and magnitude of  $F$  in Fig. 3c are nearly the same as in Fig. 3b. However,

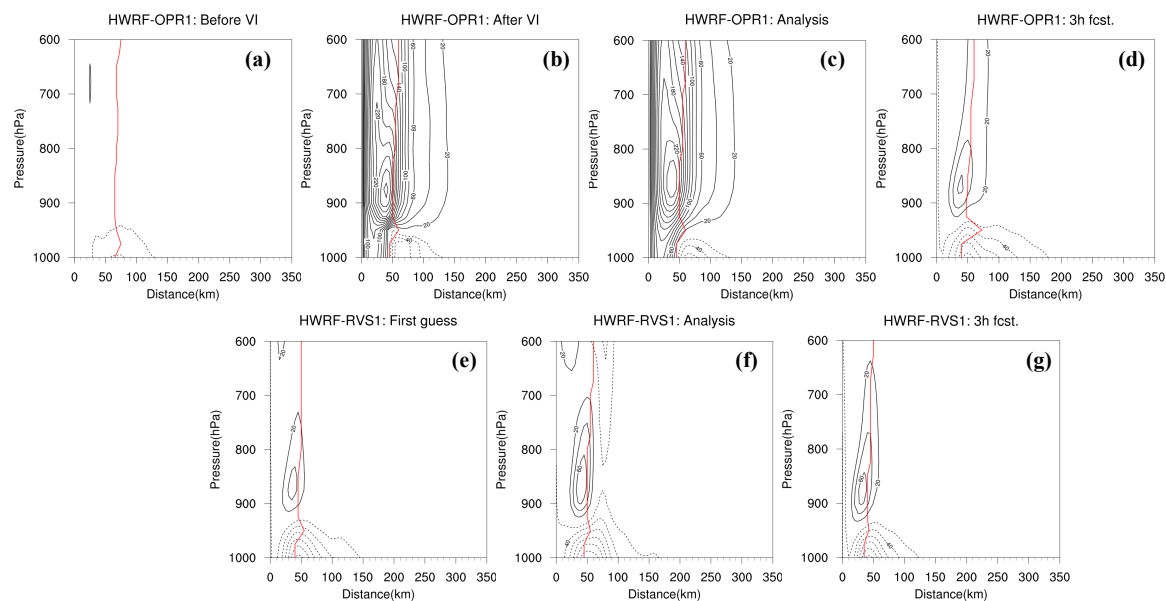
Table 2. List of experiments.

Exps.	HWRF initialization	Hurricane Case	DA period <sup>a</sup>	Simulation period
HWRF-OPR1	3DEnVar (Fig.1a)	Joaquin (2015)	1800UTC 02 – 1800 UTC 03 Oct.	1800UTC 03 – 1800 UTC 06 Oct.
HWRF-RVS1	4DEnVar (Fig.1b)	Joaquin (2015)	1800UTC 02 – 1800 UTC 03 Oct.	1800UTC 03 – 1800 UTC 06 Oct.
HWRF-OPR2	3DEnVar (Fig.1a)	Patricia (2015)	1800UTC 20 – 1800 UTC 21 Oct.	1800UTC 21 – 1800 UTC 24 Oct.
HWRF-RVS2	4DEnVar (Fig.1b)	Patricia (2015)	1800UTC 20 – 1800 UTC 21 Oct.	1800UTC 21 – 1800 UTC 24 Oct.
HWRF-OPR3	3DEnVar (Fig.1a)	Matthew (2016)	1800UTC 28 – 1800 UTC 29 Sep.	1800UTC 29 Sep. – 1800 UTC 02 Oct.
HWRF-RVS3	4DEnVar (Fig.1b)	Matthew (2016)	1800UTC 28 – 1800 UTC 29 Sep.	1800UTC 29 Sep.– 1800 UTC 02 Oct.

<sup>a</sup> The observations that are assimilated by each simulation include conventional and satellite observations as well as HDSS dropsonde and Tail Doppler Radar (TDR) radial wind data.



the distribution and magnitude of  $F$  in HWRF-OPR1 are adjusted significantly after the 3-h forecast (Fig. 3d). As a result, the gradient wind balance is established in the inner core region with radius  $< 50$  km, and the supergradient and subgradient wind fields are concentrated around the radius of maximum wind (Fig. 3d). This indicates that the supergradient wind fields induced by VI in Fig. 3b and c are not realistic, and the model has to adjust these imbalanced structures in the first few hours' forecasts. In contrast to HWRF-OPR1, the gradient wind fields are much more balanced in HWRF-RVS1 for the first guess and the analysis fields (Fig. 3e), and no significant adjustment is revealed during the first 3-h forecast period (Fig. 3f). This means that 4DVar initialization framework improves the balances in the initial conditions. In addition, a comparison of Fig. 3b with Fig. 3e shows that the improvements in HWRF-RVS1 derive mainly from not using the size and intensity correction in VI for HWRF-OPR1.

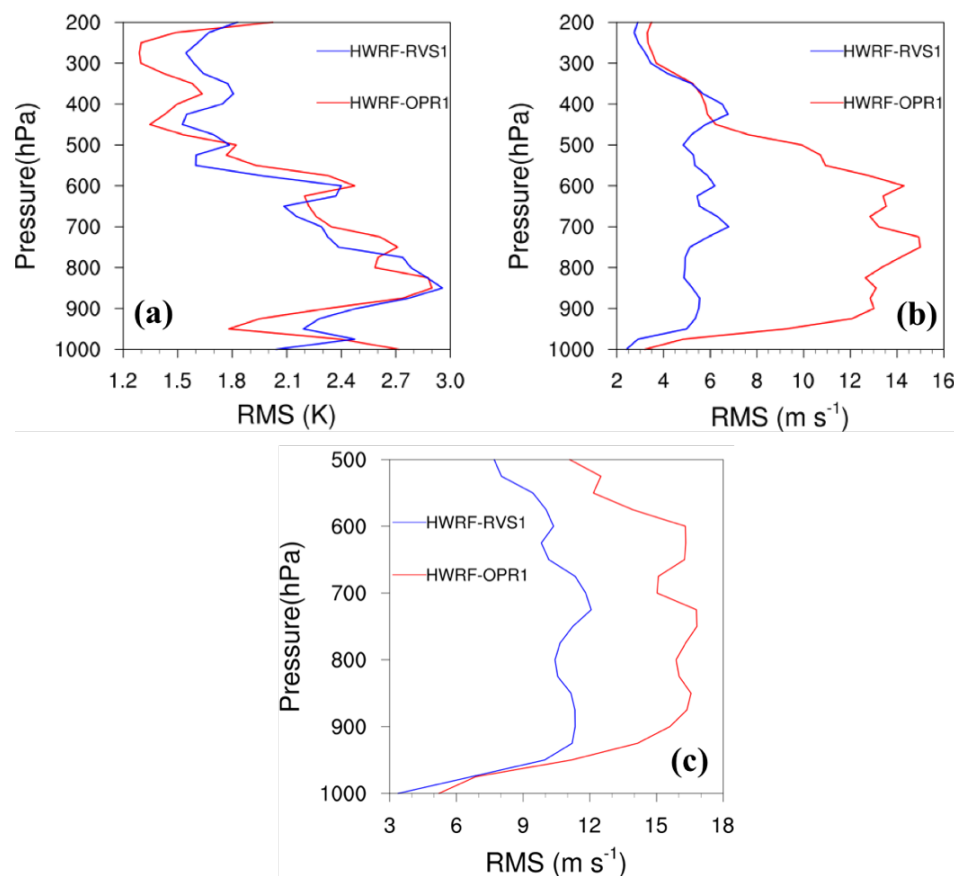


**Figure 3.** Radius-height cross sections of the isopleths of net radial force per unit mass (unit:  $\text{m s}^{-1} \text{h}^{-1}$ ). The top four panels show the results from HWRF-OPR1 that are diagnosed (a) before VI, (b) after VI, (c) after data assimilation (i.e., analysis) and (d) 3-h forecast at the analysis cycle of 1800 UTC 03 October 2015. The bottom three panels represent the results from HWRF-RVS1 that are diagnosed (e) after the vortex relocation (i.e., first guess), (f) after data assimilation (i.e., analysis) and (g) 3-h forecast at the analysis cycle of 1800 UTC 03 October 2015. The contour interval is  $20 \text{ m s}^{-1} \text{h}^{-1}$ , with dashed lines indicating negative values. The zero contour is not plotted. The red lines indicate the radius of maximum wind.

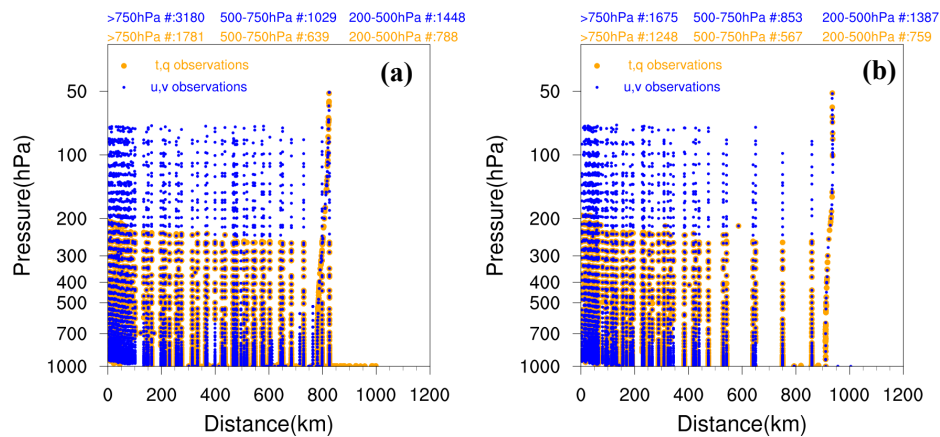
### 3.1.1. Fit to observations

The 4DVar initialization framework can enhance the performance of DA. Figures 4a and b show the root-mean-square (RMS) fit of the analysis to in-situ observations within the d02 region (Table 1), averaged over the DA period from 1800 UTC 02 to 1800 UTC 03 October 2015. The analyzed temperature from HWRF-RVS1 and HWRF-OPR1 fit the observational temperature similarly, except that HWRF-RVS1 leads to a slightly better fit in the midlevel (750–500 hPa) than HWRF-OPR1 does (Figure 4a). In addition, the root-mean-square fit of the analyzed wind to observations is significantly improved by HWRF-RVS1. As shown in Figure 4b, the RMS of wind speed within 900–500 hPa in HWRF-RVS1 is about 2–3 times smaller than that in HWRF-OPR1. Overall, the analyses from HWRF-RVS1 lead to a better fit of the analysis to the observations than HWRF-OPR1 does, but the improvements are mainly seen in the dynamical fields ( $u$ ,  $v$  wind) rather than thermal fields (i.e. temperature and humidity). Figure 5 further shows the distribution of observations that are

assimilated by GSI at 1800 UTC 02 and 1800 UTC 03 October 2015. We can see that two times more wind observations are assimilated by the HWRF DA system relative to the temperature and humidity data (see the numbers at the top of Figures 5a and b). This partially explains the results from HWRF-RVS1 show significant improvement in the wind but much less significant improvement in temperature in Figure 4. In addition, Figure 5 reveals that the temperature and humidity observations within the inner-core region are mostly concentrated in the boundary layer regions (below 700hPa), while the observations are much sparse in the middle (500–750 hPa) and upper levels (200–500 hPa). In particular, there are no temperature and humidity observations within the inner-core region (radius < 800 km) above 200 hPa (Figure 5). In this way, we would expect the worse root-mean-square fit of the analyzed temperature to observations in the upper levels (above 500 hPa) as shown in Figure 4a.

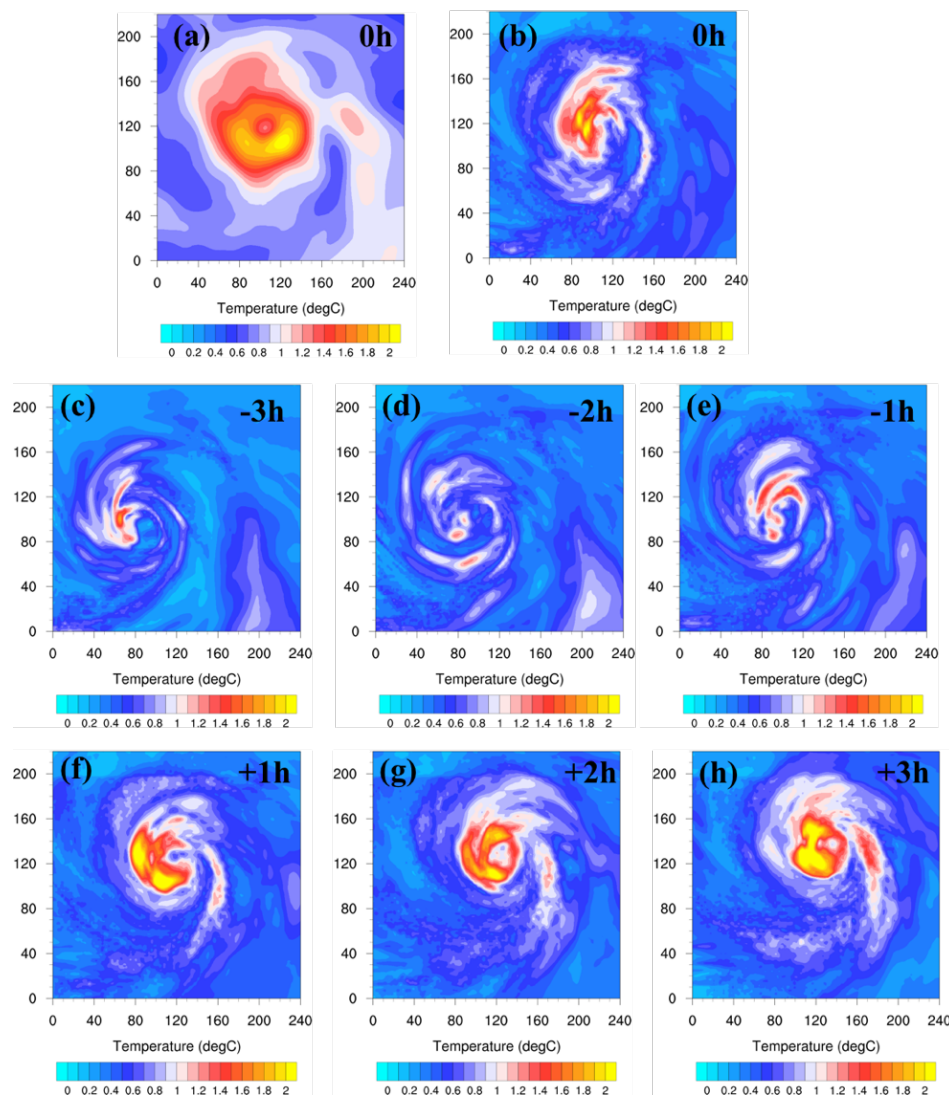


**Figure 4.** The RMS fit of the analysis to in situ observations within d02 region (Table 1) averaged over the data assimilation period from 1800UTC 02 to 1800 UTC 03 October 2015 for (a) temperature and (b) wind as a function of pressure. (c) the RMS fit of the analyzed wind to in situ observations within d03 from 1000 hPa to 500 hPa during the data assimilation period from 1800UTC 02 to 1800 UTC 03 October 2015. Red lines indicate the results from HWRF-OPR1, and blue lines indicate the results from HWRF-RVS1.



**Figure 5.** the vertical distribution of observations that are assimilated by GSI for Hurricane Joaquin at (a) 1800 UTC 2 Oct 2015 and (b) 1800 UTC 3 Oct 2015. The large orange dots denote the temperature (t) and humidity (q) data; while the small blue dots represent wind data (u and v). The numbers of observations that are assimilated in three different levels (below 750 hPa, 500 – 750 hPa, and 200 – 500 hPa) are also shown on the top of (a) and (b). The x-axis represents the distance to the storm center.

Figure 4c further illustrates the root-mean-square (RMS) fit of the analyzed wind speed to in-situ observations within the d03 region (Table 1), averaged over the DA period from 1800 UTC 02 to 1800 UTC 03 October 2015. We can see that HWRF-RVS1 produces a much smaller RMS of wind speeds extended from 900 hPa to 500 hPa than HWRF-OPR1 does. Two factors in the 4DEnVar initialization framework can contribute to the improvements in the HWRF analysis in HWRF-RVS1: first, the use of high-resolution flow-dependent error covariances in HWRF-RVS1. Figures 6a and b show the ensemble spreads of temperature at 850 hPa from the 80-member GFS EnKF ensembles for HWRF-OPR1 and the 42-member HWRF regional ensembles for HWRF-RVS1 at 1800 UTC 03 October 2015. Apparently, the 42-member HWRF ensemble members provide many detailed storm-scale structural features for temperature (Figure 6b) compared with the 80-member global ensembles (Figure 6a). The same features are also observed in the wind and relative humidity fields (Figures not shown). Compared with the coarser-resolution global ensemble forecasts, high-resolution regional ensemble forecasts in the HWRF native model domains can produce much more realistic (storm-scale) correlation structures among these variables and enable DA to incorporate the convective-scale information from observations into the HWRF analysis. Second, the flow-dependent background error covariances in HWRF-RVS1 evolve with time within a 6-h DA window. As shown in Figure 6c-h, the ensemble spreads of temperature in HWRF-RVS1 reveal different structures at different times within the DA window. The variation of the ensemble perturbations within the DA window will give the DA system a greater ability to capture the high-frequency changes in atmospheric variables in the observational field (e.g., wind field). With these advantages, 4DEnVar initialization framework can better incorporate the observational information related to TCs into the HWRF initial conditions, relative to the operational initialization framework. As a result, the analysis field in HWRF-RVS1 is more realistic than that in HWRF-OPR1, and the RMS errors of wind in HWRF-RVS1 are smaller than those in HWRF-OPR1.

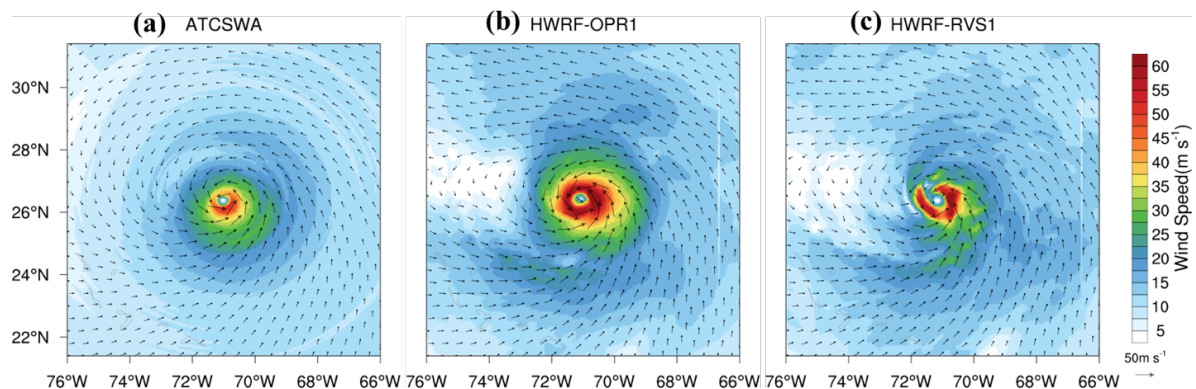


**Figure 6.** Ensemble spreads of temperature (unit:  $^{\circ}\text{C}$ ) at 850 hPa within 6-h analysis window (-3h–3h) around 1800 UTC 03 October 2015 from (a) GFS EnKF 80-member ensemble forecasts at  $\sim 23\text{km}$  horizontal resolution and (b–h) HWRF 42-member ensemble forecasts at  $6\text{km}$  horizontal resolution. The numbers in each plot indicate the differences between the ensemble forecast time and the analysis time.

### 3.1.2. Analysis of storm structure

The improvement of the fit of analyzed wind speed to the in-situ observations in HWRF-RVS1 implies that HWRF-RVS1 may produce better analyzed wind structure than HWRF-OPR1 does. Figure 7 illustrates the 10-m wind analysis field from HWRF-OPR1 and HWRF-RVS1, compared against the Aircraft-based Tropical Cyclone Surface Wind Analysis (ATCSWA) provided by the National Oceanic and Atmospheric Administration/National Environmental Satellite Data and Information Service (NOAA/NESDIS). Previous studies [8,9] have proven that the basic vortex structure can be well reproduced by ATCSWA data. Thus, the ATCSWA can be used to evaluate the vortex structure in the HWRF analysis and forecast. Compared with the ATCSWA wind field across the center of Hurricane Joaquin (Figure 7a), the HWRF-OPR1 analysis at 1800 UTC 03 October 2015 (Figure 7b) overestimates the maximum wind speed in the inner-core region and fails to capture the asymmetric vortex structure as shown in ATCSWA. In contrast, the HWRF-RVS1 analysis (Figure 7c) reproduces a more reasonable magnitude of maximum wind speed than HWRF-OPR1 does and captures the

asymmetric wind structure as shown in ATCSWA. Thus, 4DEnVar initialization framework produces more realistic surface wind structure than the operational initialization framework does.

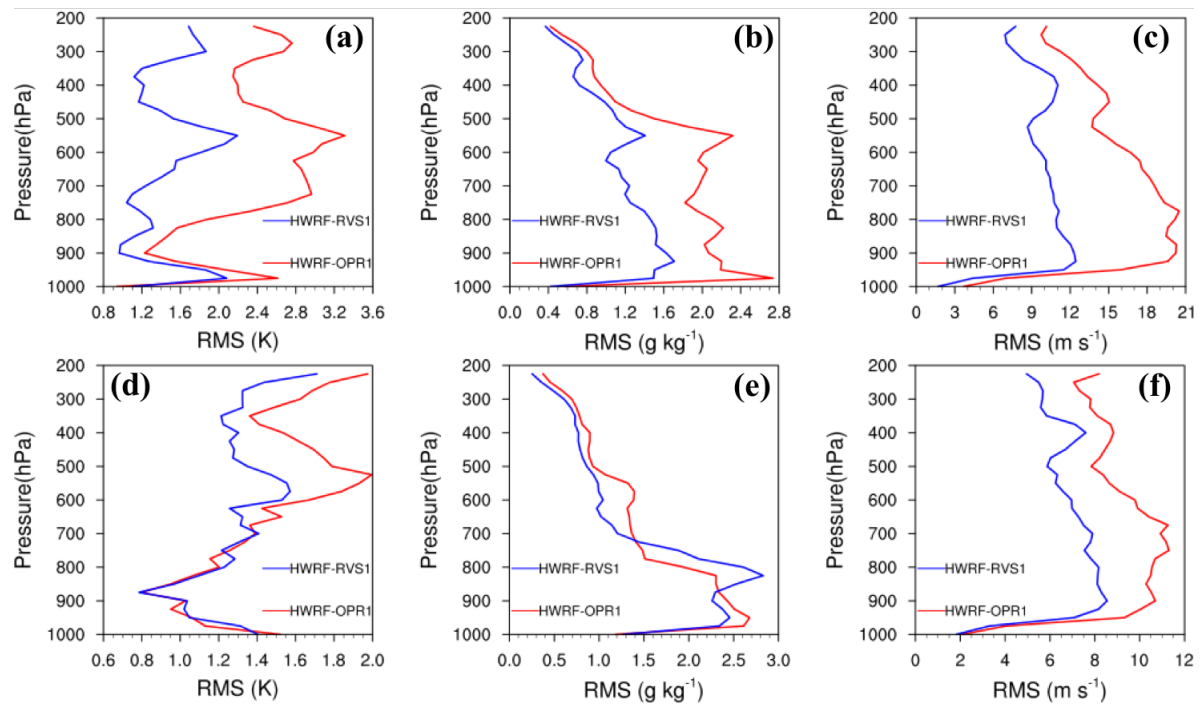


**Figure 7.** Vortex structure of Hurricane Joaquin at 1800 UTC 03 October 2015 as revealed by wind speeds (shaded; unit:  $\text{m s}^{-1}$ ) and vectors at 10-m height from (a) ATCSWA, (b) HWRF-OPR1 and (c) HWRF-RVS1.

### 3.1.3. HWRF forecast

The two initialization frameworks also lead to different HWRF forecasts during the rapid weakening of Hurricane Joaquin. Figure 8 illustrates the RMS errors for temperature, specific humidity, and wind forecasts with the HWRF-OPR1 and HWRF-RVS1 experiments verified against the HDSS dropsonde observations at 1800 UTC 04 and 1800 UTC 05 October 2015 over the innermost domain (d03, Table 1). The HDSS dropsonde observations are from the Tropical Cyclone Intensity (TCI) 2015 field campaign (available online at <https://data.eol.ucar.edu/dataset/488.004>). The data is quality controlled using a combined "subjective-objective" procedure which utilizes the Atmospheric Sounding Processing Environment (ASPEN) software (Bell et al. 2016). After the quality control, there are 74 (80) total profile observations at 1800 UTC 04 (05) October 2015. In this study, the grid point values of temperature, specific humidity and wind speed from HWRF simulations are first interpolated into a HDSS dropsonde observational point, then the average RMS errors between the HWRF forecasts and HDSS dropsondes are calculated and compared. As shown in Figure 8, the RMS errors of temperature, specific humidity, and wind speed at the 24-h and 48-h forecasts in HWRF-RVS1 (blue lines) are smaller than those in HWRF-OPR1 (red lines). This indicates that the forecasts produced by HWRF-RVS1 are more skillful than those of HWRF-OPR1. In addition, the reduction in the RMS errors of wind in HWRF-RVS1 is significant in both the 24-h and 48-h forecasts, whereas the reduction in the RMS errors of temperature and specific humidity are significant only in the 24-h forecast.

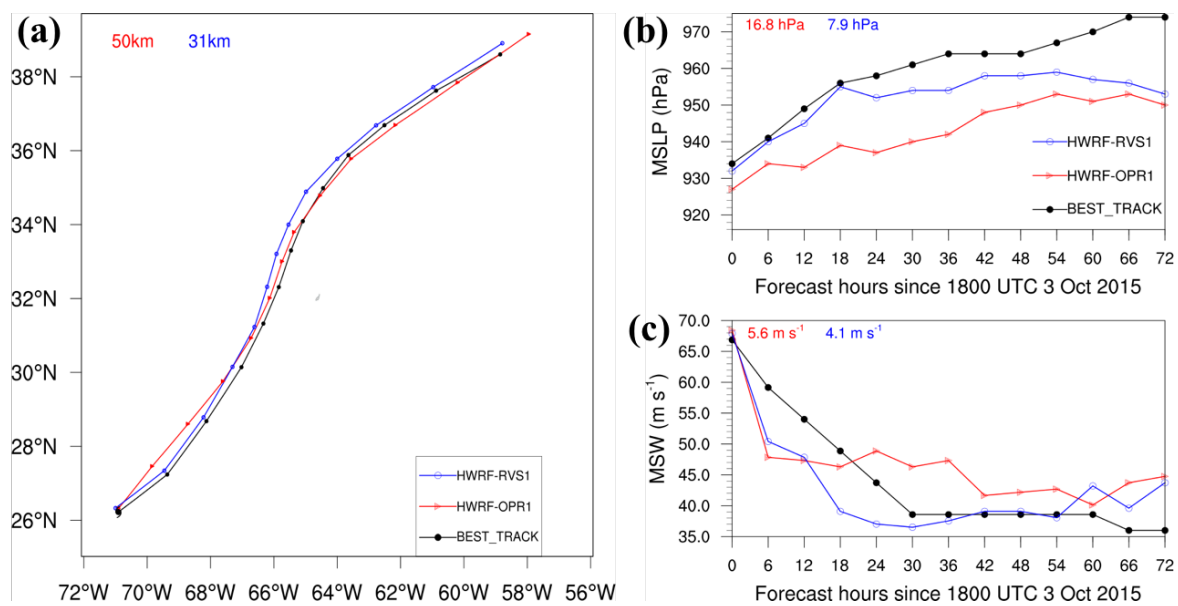




**Figure 8.** The RMS fit of the the forecasts to the dropsonde observations for (left) temperature, (middle) specific humidity and (right) wind speed as a function of pressure at (a-c) 24-h and (d-f) 48-h forecast lead times for the experiments of HWRf-OPR1 (red lines) and HWRf-RVS1 (blue lines)

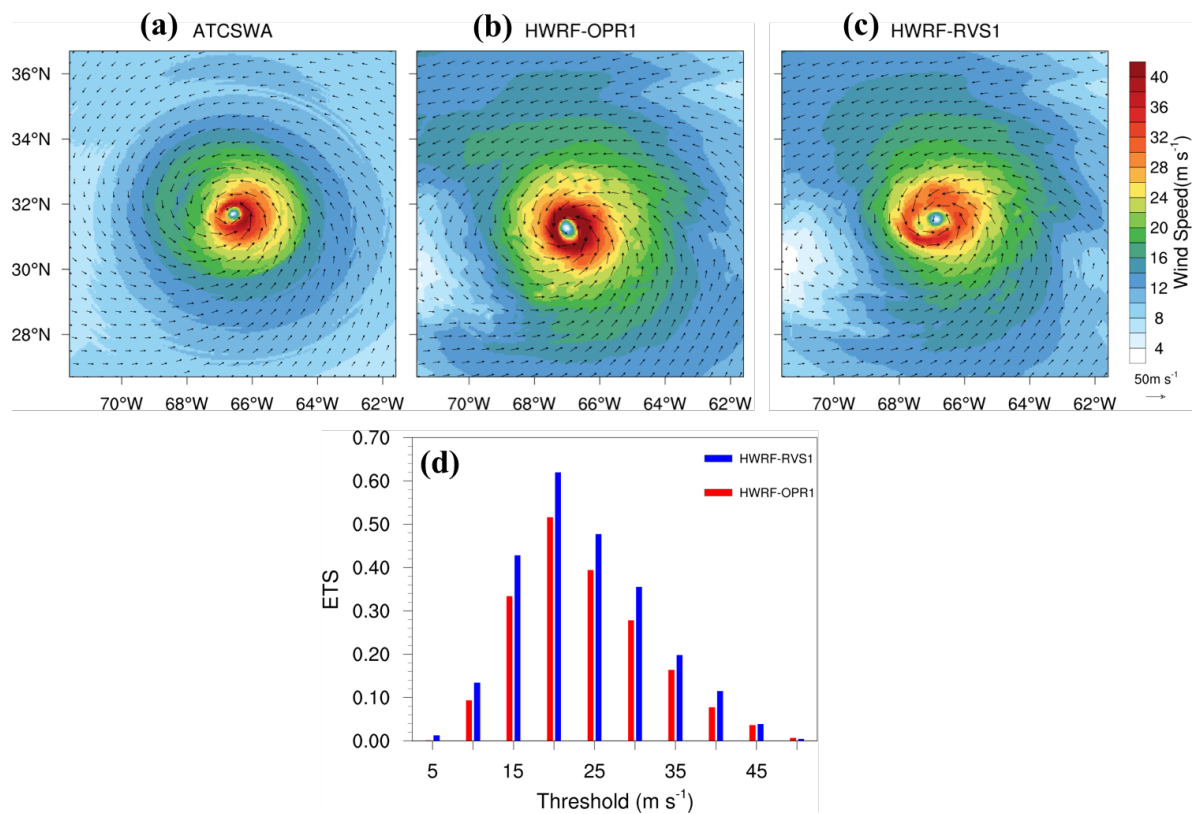
Figure 9 further shows the track and intensity in terms of the minimum sea level pressure (MSLP) and maximum surface wind (MSW) produced by the two different initialization frameworks from 1800 UTC 03 October to 1800 UTC 06 October 2015. It is verified that HWRf-RVS1 leads to better track and intensity forecasts than HWRf-OPR1 does. Specifically, both HWRf-OPR1 and HWRf-RVS1 produce realistic track forecasts, as the track errors over a 72-h forecast period are less than 60 km (Figure 9a). However, the cross-track errors in the first 12 h and the along-track errors after the 48-h forecast in HWRf-OPR1 are reduced in HWRf-RVS1. As a result, HWRf-RVS1 leads to a 37% average track error reduction over the 72-h forecast period compared with HWRf-OPR1 (Figure 9a). As for the intensity forecasts, HWRf-RVS1 reduces the 72-h averaged MSLP and MSW errors in HWRf-OPR1 by 53% and 26%, respectively (Figure 9b-c). More importantly, the unrealistic wind-pressure relationship in HWRf-OPR1 is eliminated by HWRf-RVS1, as the overestimation of MSLP in HWRf-OPR1 is significantly mitigated by HWRf-RVS1.





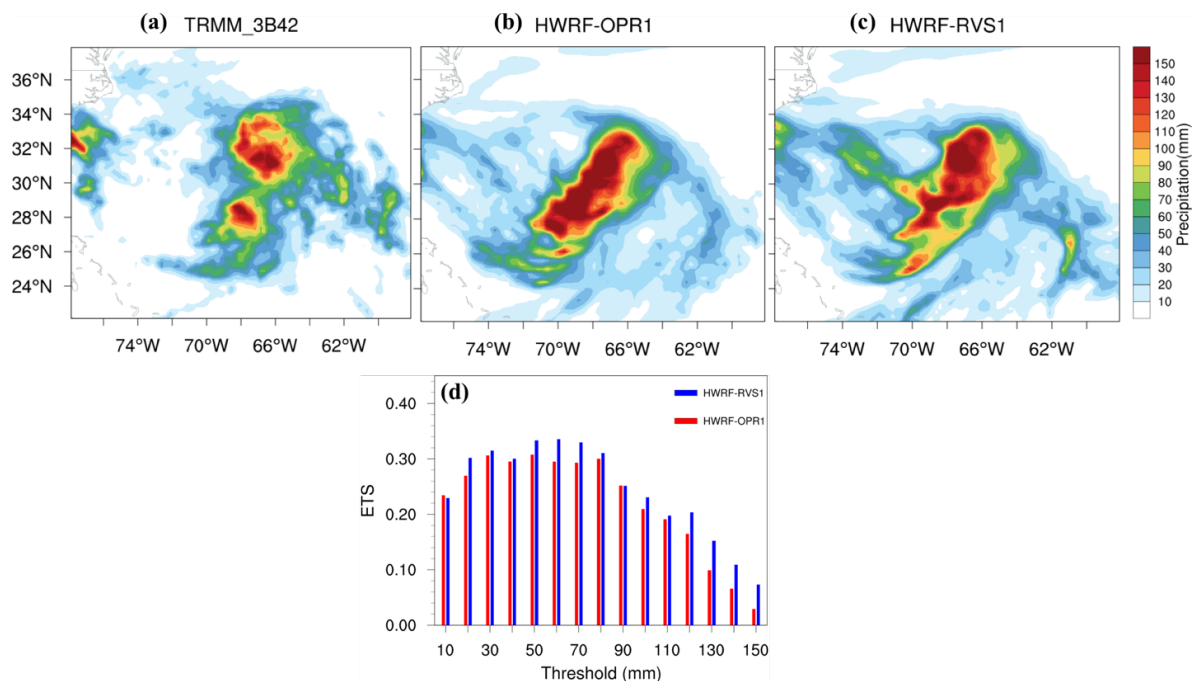
**Figure 9.** Time series (at 6-h intervals) of (a) track, (b) minimum sea level pressure (MSLP, hPa) and (c) maximum surface winds (MSW,  $\text{m s}^{-1}$ ) for Hurricane Joaquin during 1800 UTC 03–1800 UTC 06 October 2015. The colored numbers on the top of figures represent the average track (a), MSLP (b) and MSW (c) errors over the whole 72 h forecast period.

Quantitative wind and precipitation forecasts from the two different initialization frameworks are also diagnosed. Figure 10 shows the distribution of the 10-m wind field from the 24-h forecast of HWRF-OPR1 and HWRF-RVS1, compared against the 10-m wind analysis from ATCSWA at 1800 UTC 03 October 2015. Relative to ATCSWA (Figure 10a), the large overestimation of wind speed around the storm center in HWRF-OPR1 (Figure 10b) is not seen in HWRF-RVS1 (Figure 10c). In addition, HWRF-RVS1 (Figure 10c) partially captures the asymmetric wind structure shown in ATCSWA (Figure 10a), whereas HWRF-OPR1 (Figure 10b) leads to symmetric wind structure. Therefore, HWRF-RVS1 leads to a more realistic storm structure forecast than HWRF-OPR1 does. Following Wang (2014), mean equitable threat scores (ETSs) are further calculated for the forecasted 10-m wind fields from HWRF-OPR1 and HWRF-RVS1 compared against the ATCSWA wind analysis fields during the rapid weakening of Hurricane Joaquin from 1800 UCT 03 to 0000 UTC 05 October 2015. The ETS measures the quantitatively forecast accuracy relative to the observational or analysis field. As depicted in Figure 10d, the mean ETSs in HWRF-RVS1 are higher than those in HWRF-OPR1 for all the thresholds from  $10 \text{ m s}^{-1}$  to  $50 \text{ m s}^{-1}$ , indicating that HWRF-RVS1 produces more skillful wind forecasts during the whole rapid weakening of Hurricane Joaquin than HWRF-OPR1 does.



**Figure 10.** Vortex structure of Hurricane Joaquin at 1800 UTC 04 October 2015 as revealed by wind speeds (shaded; unit:  $\text{m s}^{-1}$ ) and vectors at 10m height from (a) ATCSWA, (b) HWRF-OPR and (c) HWRF-RVS. (d) the mean equitable threat score (ETS) of the 10-m forecasted wind field from HWRF-OPR1 (red bars) and HWRF-RVS1 (blue bars) compared against the 10-m wind analysis from ATCSWA with thresholds of 5, 10, 15, 20, 25, 30, 35, 40, 45 and 50  $\text{m s}^{-1}$ . The mean ETS is calculated during the rapid weakening of Hurricane Joaquin from 1800 UTC 03 to 0000 UTC 05 October 2015 with a 6-h interval.

Figure 11 also illustrates the 30-h accumulated precipitation distribution during the rapid weakening of Hurricane Joaquin from 1800 UTC 03 to 0000 UTC 05 October 2015, compared with precipitation observations from the Tropical Rainfall Measuring Mission (TRMM) 3B42 products. It is clear that HWRF-OPR1 (Fig. 11b) produces too much rainfall along the storm track relative to TRMM 3B42 (Fig. 11a), while HWRF-RVS1 captures the magnitude and distribution of precipitation below 30°N in TRMM 3B42 (Figure 11a, c), and the overall precipitation distribution in HWRF-RVS1 (Fig. 11c) is much more realistic than that in HWRF-OPR1 (Fig. 11b). In order to quantitatively verify the precipitation forecasts (QPFs) in HWRF-OPR1 and HWRF-RVS1, ETSs are also calculated for the 30-h accumulated precipitation from HWRF-OPR1 and HWRF-RVS1 compared with the precipitation observations from TRMM 3B42. As shown in Fig. 11d, the ETSs in HWRF-RVS1 are higher than those in HWRF-OPR1 for all thresholds from 10 mm to 150 mm, indicating that HWRF-RVS1 also produces more skillful precipitation forecasts during the rapid weakening of Hurricane Joaquin than HWRF-OPR1 does.

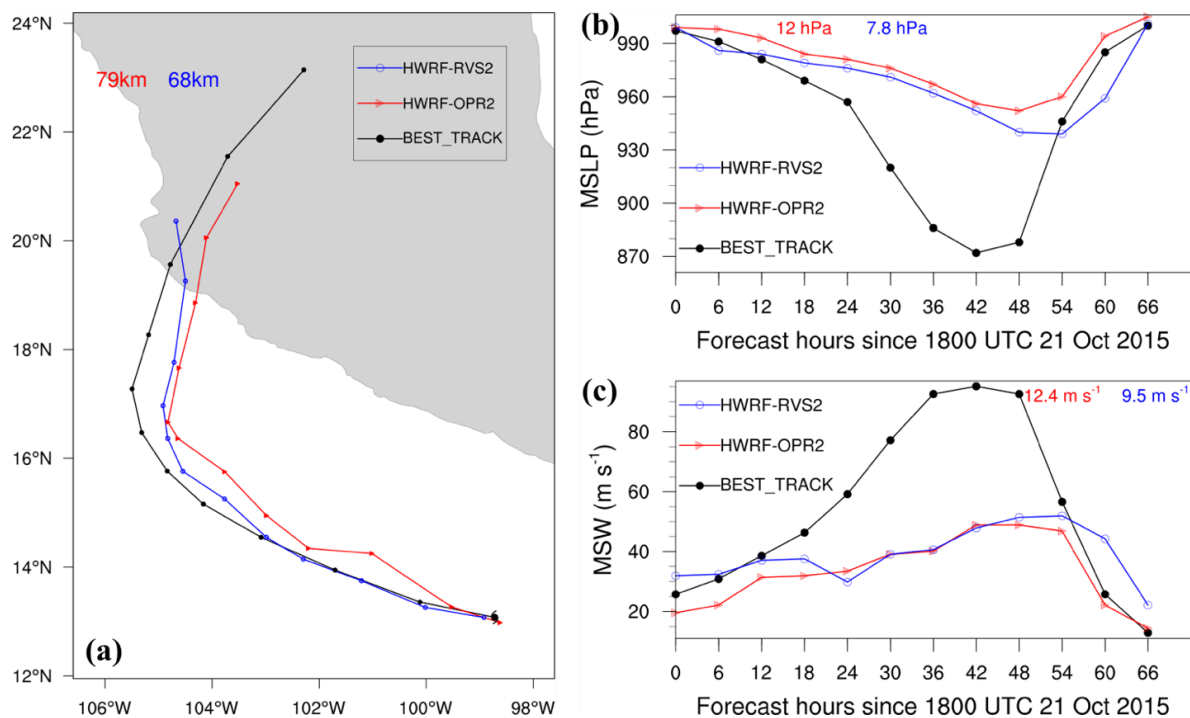


**Figure 11.** Accumulated precipitation (unit: mm) from (a) TRMM 3B42, (b) HWRf-OPR1 and (c) HWRf-RVS1 during the rapid weakening of Hurricane Joaquin from 1800 UTC 03 to 0000 UTC 05 October 2015. (d) the equitable threat score (ETS) of the 30-h accumulated precipitation from HWRf-OPR1 (red bars) and HWRf-RVS1 (blue bars) compared against TRMM 3B42 precipitation observations with thresholds of 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140 and 150 mm.

#### 4. Forecasts of the rapid intensification of Hurricanes Patricia and Matthew

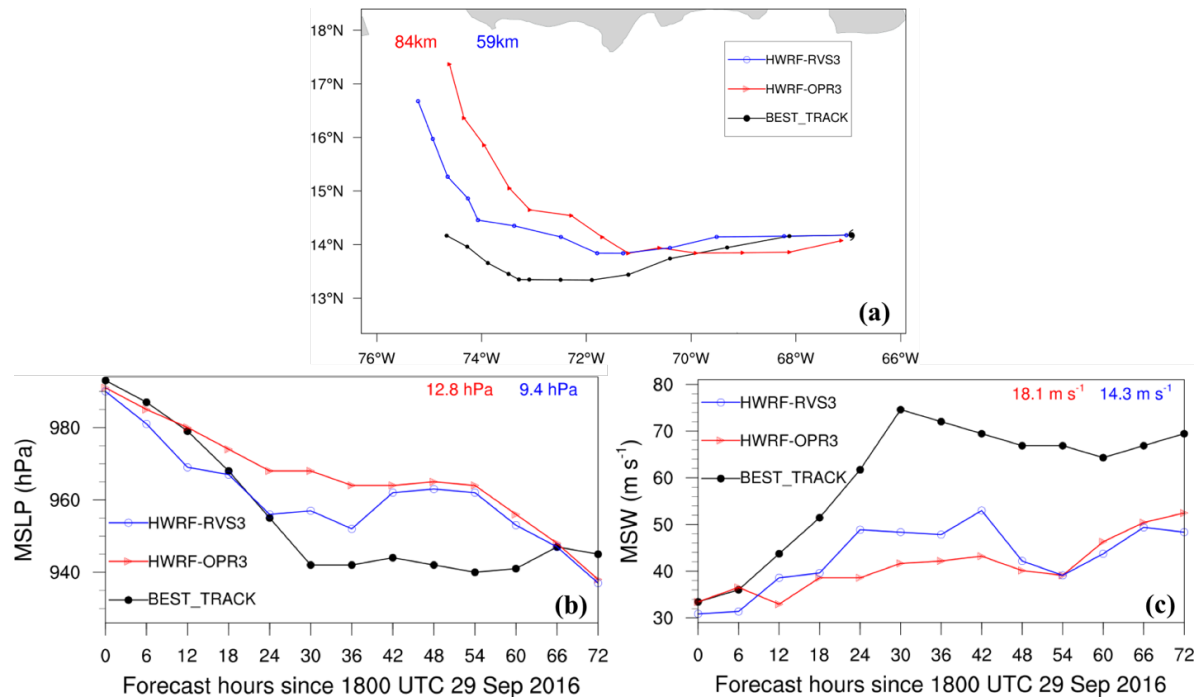
As we have shown in the previous section, the 4DnVar initialization framework leads to a more skillful forecast of TC intensity change than the 3DnVar initialization framework does. In order to further confirm these results, HWRf simulations with two different HWRf initialization frameworks are also conducted on Hurricanes Patricia and Matthew during their rapid intensification period. Detailed descriptions of these two cases can be found in Kimberlain *et al.* [26] and Stewart [27], respectively. For Hurricane Patricia, the HWRf model is initialized at 1200 UTC 20 October 2015, the model is allowed to spin-up until 1800 UTC 20 October 2015, and then a DA process is performed that is cycled from 1800 UTC 20 to 1800 UTC 21 October 2015 every 6 h, for total of 24 hours. A 66-h forecast from 1800 UTC 21 to 1200 UTC 24 October 2015 is then performed to predict intensity changes of Hurricane Patricia. For Hurricane Matthew, the HWRf model is initialized at 1200 UTC 28 September 2016, allowed to spin-up until 1800 UTC 28 September 2016, and the cycled DA is performed from 1800 UTC 28 to 1800 UTC 29 September 2016 in 6-h windows, for total of 24 hours. Finally, a 72-h forecast from 1800 UTC 29 to 1800 UTC 02 October 2016 is made to simulate the period of intensity changes of Hurricane Matthew. As with Hurricane Joaquin, two HWRf experiments are designed for Hurricanes Patricia (Matthew) to understand the effects of the two different initialization frameworks. Details of the experimental design are listed in Table 2.

Figure 12 shows the track and intensity in terms of minimum sea level pressure (MSLP) and maximum surface wind (MSW) produced by the two different initialization frameworks for Hurricane Patricia from 1800 UTC 21 to 1200 UTC 24 October 2015. Overall, the rapid intensification of Hurricane Patricia is not well captured by either HWRf-OPR2 or HWRf-RVS2 (Fig. 12b, c). However, HWRf-RVS2 outperforms HWRf-OPR2 in track, MSLP, and MSW forecasts (Fig. 12a-c). Compared with HWRf-OPR2, the mean track, MSLP, and MSW errors over the first 24-h forecast in HWRf-RVS2 are reduced by 14%, 35%, and 23%, respectively. In addition, HWRf-RVS2 leads to smaller cross-track errors than HWRf-OPR2 over the whole 66 h of simulations.



**Figure 12.** Time series (at 6-h intervals) of (a) track, (b) minimum sea level pressure (MSLP, hPa) and (c) maximum surface winds (MSW,  $\text{m s}^{-1}$ ) for Hurricane Patricia during 1800 UTC 21–1200 UTC 24 October. The colored numbers on the top of figures represent the average track (a), MSLP (b) and MSW (c) errors over the whole 24-h forecast period.

Furthermore, Figure 13 illustrates the 72-h track and intensity forecasts for Hurricane Matthew during its rapid intensification from 1800 UTC 29 September to 1800 UTC 02 October 2016. It is clear that HWRf-RVS1 outperforms HWRf-OPR1 in terms of track, MSLP, and MSW forecasts (Fig. 13a-c). Compared with HWRf-OPR1, the mean track, MSLP, and MSW errors over the first 48-h forecast in HWRf-RVS1 are reduced by 30%, 27%, and 21%, respectively. More importantly, the intensification of Hurricane Matthew in the first 24 hours is well captured by HWRf-RVS1, as the track, MSLP, and MSW forecasts are nearly the same as those in the best-track data. Overall, the results for Hurricanes Patricia and Matthew further suggest that 4DVar initialization framework indeed improves forecasts of TC intensity change in HWRf.



**Figure 13.** Time series (at 6-h intervals) of (a) track, (b) minimum sea level pressure (MSLP, hPa) and (c) maximum surface winds (MSW,  $\text{m s}^{-1}$ ) for Hurricane Matthew during 1800 UTC 29 September–1800 UTC 02 October 2016. The colored numbers on the top of figures represent the average track (a), MSLP (b) and MSW (c) errors over the whole 24-h forecast period.

## 5. Concluding remarks

This study examined the forecast skill of TC intensity changes in a research version of the HWRf model with two different initialization frameworks. The first initialization framework is similar to that used in the operational HWRf, which employs the GSI-3DEnVar DA system with background error covariances provided by the coarser-resolution GFS EnKF ensembles. The second initialization framework is different from the first one, it uses GSI-4DEnVar with background error covariances provided by high-resolution self-consistent HWRf ensembles. Various HWRf simulations with the two different initialization frameworks are performed on three major hurricanes, including Joaquin, Patricia, and Matthew, during their intensity changes. The results suggest that:

- 4DEnVar initialization framework has shown the potentials of leading to a better HWRf forecast of Hurricanes Joaquin, Patricia, and Matthew during their changes in intensity. Compared with the simulations with the 3DEnVar initialization framework, the 4DEnVar initialization framework leads to a reduction in track, MSLP, and MSW forecast errors, at least in the first 24-h forecast in all the HWRf simulations. However, the extent of the improvement is sensitive to the specific case. Specifically, the 4DEnVar initialization framework produces substantial improvements in the forecast of the rapid weakening of Hurricane Joaquin, while it produces weak improvements in the forecasts of the rapid intensification of Hurricanes Patricia and Matthew.
- Further diagnoses for Hurricane Joaquin indicate that 4DEnVar initialization framework, which does not use size and intensity correction in the vortex initialization, can significantly alleviate the gradient imbalances in the HWRf initial conditions produced by the operational initialization framework. In addition, 4DEnVar initialization framework can enhance the performance of DA, as it produces smaller HWRf analysis errors than the operational initialization framework does. This is because DA in 4DEnVar initialization framework not only involves the variation of the background error covariances within the DA window but also uses high-resolution self-consistent HWRf ensembles that can provide detailed storm-scale features, which enhances the incorporation of observational information into the model initial conditions.



- The evaluation of the HWRF-forecasted 10-m wind and accumulated precipitation for Hurricane Joaquin shows that 4DEnVar initialization framework produces higher ETS scores for wind and precipitation during the intensity changes of Hurricane Joaquin than the operational initialization framework does. This indicates that 4DEnVar initialization framework also has the potential to improve quantitative wind and precipitation forecasts in HWRF.

Overall, the proposed revised initialization framework in this study reveals the potential to improve forecasts of TC intensity changes in HWRF. It suggests that the use of high-resolution self-consistent HWRF ensembles and the more advanced GSI-4DEnVar data assimilation systems can lead to noticeable improvements in the vortex initializations and forecasts of hurricanes using HWRF. Fortunately, advancing DA systems is of the utmost importance to improve hurricane forecasting models, like HWRF and its eventually successors [28]. In particular, the HWRF-based ensemble prediction system (EPS) has been implemented in operational TC forecasts in NCEP/EMC since 2014 [21], which provides strong support for the practical use of 4DEnVar initialization framework in the operational HWRF system. In addition, implementation of 4DEnVar in the NCEP global forecast system will make it possible to eventually move 4DEnVar into the HWRF DA system. We believe that our investigations in this study can provide useful guidance for future work. Furthermore, the new satellite-derived products such as enhanced AMVs and CYGNSS sea surface wind developed in current observational projects and TC field campaigns could be incorporated into 4DEnVar initialization framework to further improve hurricane initialization[29–32]. Finally, our evaluation in this study mainly focuses on the track and intensity forecasts for three selected hurricane cases during their rapid intensity changes. A larger sample of hurricane cases will be necessary to evaluate the consistency of the impacts on the HWRF Model provided by the revised initialization framework, and the detailed reasoning and physical understanding on the data assimilation results also need to be further discussed.

**Author Contributions:** Conceptualization, Z. Pu and S. Zhang; methodology, S. Zhang; software, S. Zhang; validation, S. Zhang and Z. Pu; formal analysis, S. Zhang; investigation, S. Zhang; resources, Z. Pu; data curation, S. Zhang and Z. Pu; writing—original draft preparation, S. Zhang; writing—review and editing, S. Zhang and Z. Pu; visualization, S. Zhang; supervision, Z. Pu; project administration, Z. Pu; funding acquisition, Z. Pu. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by NOAA/NWS Awards #NA14NWS4680025 and #NA190AR4590239.

**Acknowledgments:** The NOAA Jet supercomputer, supported by NOAA/ESRL and the Center for High-Performance Computing (CHPC) at the University of Utah, are greatly appreciated.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

NCEP	National Centers for Environmental Prediction
GEFS	Global Ensemble Forecast System
GSI	Gridpoint Statistical Interpolation
GSI-3DEnVar	GSI-based three-dimensional ensemble-variational hybrid data assimilation system
GSI-4DEnVar	GSI-based four-dimensional ensemble-variational hybrid data assimilation system
HWRF	Hurricane Weather Research and Forecasting model
GFS	NCEP Global Forecast System

## References

1. Elsberry, R.L. Achievement of USWRP Hurricane Landfall Research Goal. *Bulletin of the American Meteorological Society* **2005**, *86*, 643–646. doi:10.1175/BAMS-86-5-643.
2. Liu, J.; Zhang, F.; Pu, Z. Numerical simulation of the rapid intensification of Hurricane Katrina (2005): Sensitivity to boundary layer parameterization schemes. *Advances in Atmospheric Sciences* **2017**, *34*, 482–496. doi:10.1007/s00376-016-6209-5.



3. Elsberry, R.L.; Lambert, T.D.B.; Boothe, M.A. Accuracy of Atlantic and Eastern North Pacific Tropical Cyclone Intensity Forecast Guidance. *Weather and Forecasting* **2007**, *22*, 747–762. doi:10.1175/WAF1015.1.
4. Zhang, F.; Weng, Y.; Gamache, J.F.; Marks, F.D. Performance of convection-permitting hurricane initialization and prediction during 2008?2010 with ensemble data assimilation of inner-core airborne Doppler radar observations. *Geophysical Research Letters* **2011**, *38*. doi:10.1029/2011GL048469.
5. Rogers, R.; Aberson, S.; Aksoy, A.; Annane, B.; Black, M.; Cione, J.; Dorst, N.; Dunion, J.; Gamache, J.; Goldenberg, S.; Gopalakrishnan, S.; Kaplan, J.; Klotz, B.; Lorsolo, S.; Marks, F.; Murillo, S.; Powell, M.; Reasor, P.; Sellwood, K.; Uhlhorn, E.; Vukicevic, T.; Zhang, J.; Zhang, X. NOAA'S Hurricane Intensity Forecasting Experiment: A Progress Report. *Bulletin of the American Meteorological Society* **2013**, *94*, 859–882. doi:10.1175/BAMS-D-12-00089.1.
6. DeMaria, M.; Mainelli, M.; Shay, L.K.; Knaff, J.A.; Kaplan, J. Further Improvements to the Statistical Hurricane Intensity Prediction Scheme (SHIPS). *Weather and Forecasting* **2005**, *20*, 531–543. doi:10.1175/WAF862.1.
7. DeMaria, M.; DeMaria, R.T.; Knaff, J.A.; Molenaar, D.A Tropical Cyclone Lightning and Rapid Intensity Change. *Monthly Weather Review* **2012**, *140*, 1828–1842. doi:10.1175/MWR-D-11-00236.1.
8. Knaff, J.A.; Longmore, S.P. ; DeMaria, R.T.; Molenaar, D.A Tropical Cyclone Lightning and Rapid Intensity Change. *J. Appl. Meteor. Climat.* **2015**, *54* (2), 463–478. doi:10.1175/JAMC-D-14-0112.1.
9. Knaff, J.A.; Zehr, R.M Reexamination of tropical cyclone wind?pressure relationships. *Weather and Forecasting* **2007**, *22*, 71–88. doi:10.1175/MWR-D-11-00236.1.
10. Pu, Z.; Zhang, S.; Tong, M.; Tallapragada, V. Influence of the Self-Consistent Regional Ensemble Background Error Covariance on Hurricane Inner-Core Data Assimilation with the GSI-Based Hybrid System for HWRF. *Journal of the Atmospheric Sciences* **2016**, *73*, 4911–4925. doi:10.1175/JAS-D-16-0017.1.
11. Tallapragada, V.; Bernardet, L.; Biswas, M.K.; Ginis, I.; Kwon, Y.; Liu, Q.; Marchok, T.; Sheinin, D.; Thomas, B.; Tong, M.; Trahan, S.; Wang, W.; Yablonsky, R.; Zhang, X. Hurricane Weather Research and Forecasting (HWRF) Model: 2015 Scientific Documentation (No. NCAR/TN-522+STR). *HWRF Development Testbed Center Tech. Rep* **2016**. doi:10.5065/D6ZP44B5.
12. Liu, Q.; Surgi, N.; Lord, S.; Wu, W.; Parrish, S.; Gopalakrishnan, S.; Waldrop, J.; Gamache, J. Hurricane Initialization in HWRF Model. Preprints. 27th Conference on Hurricanes and Tropical Meteorology, Monterey, CA, 2006.
13. Wu, W.S.; Purser, R.J.; Parrish, D.F. Three-Dimensional Variational Analysis with Spatially Inhomogeneous Covariances. *Monthly Weather Review* **2002**, *130*, 2905–2916. doi:10.1175/1520-0493(2002)130<2905:TDVAWS>2.0.CO;2.
14. Wang, X.; Parrish, D.; Kleist, D.; Whitaker, J. GSI 3DVar-Based Ensemble?Variational Hybrid Data Assimilation for NCEP Global Forecast System: Single-Resolution Experiments. *Monthly Weather Review* **2013**, *141*, 4098–4117. doi:10.1175/MWR-D-12-00141.1.
15. Lu, X.; Wang, X.; Li, Y.; Tong, M.; Ma, X. GSI-based ensemble-variational hybrid data assimilation for HWRF for hurricane initialization and prediction: impact of various error covariances for airborne radar observation assimilation. *Quarterly Journal of the Royal Meteorological Society* **2017**, *143*, 223–239. doi:10.1002/qj.2914.
16. Tallapragada, V.; Kieu, C.; Kwon, Y.; Trahan, S.; Liu, Q.; Zhang, Z.; Kwon, I.H. Evaluation of Storm Structure from the Operational HWRF during 2012 Implementation. *Monthly Weather Review* **2014**, *142*, 4308–4325. doi:10.1175/MWR-D-13-00010.1.
17. Wang, X.; Lei, T. GSI-Based Four-Dimensional Ensemble?Variational (4DEnsVar) Data Assimilation: Formulation and Single-Resolution Experiments with Real Data for NCEP Global Forecast System. *Monthly Weather Review* **2014**, *142*, 3303–3325. doi:10.1175/MWR-D-13-00303.1.
18. Kleist, D.T.; Ide, K. An OSSE-Based Evaluation of Hybrid Variational?Ensemble Data Assimilation for the NCEP GFS. Part II: 4DEnVar and Hybrid Variants. *Monthly Weather Review* **2015**, *143*, 452–470. doi:10.1175/MWR-D-13-00350.1.
19. Gopalakrishnan, S.G.; Marks, F.; Zhang, X.; Bao, J.W.; Yeh, K.S.; Atlas, R. The Experimental HWRF System: A Study on the Influence of Horizontal Resolution on the Structure and Intensity Changes in Tropical Cyclones Using an Idealized Framework. *Monthly Weather Review* **2011**, *139*, 1762–1784. doi:10.1175/2010MWR3535.1.

20. Tallapragada, V.; Bernardet, L.; Biswas, M.K.; Ginis, I.; Kwon, Y.; Liu, Q.; Marchok, T.; Sheinin, D.; Thomas, B.; Tong, M.; Trahan, S.; Wang, W.; Yablonsky, R.; Zhang, X. Scientific Documentation for the NMM Solver (No. NCAR/TN-477+STR). *National Center for Atmospheric Research University Corporation for Atmospheric Research* **2010**. doi:10.5065/D6MW2F3Z.
21. Zhang, Z.; Tallapragada, V.; Kieu, C.; Trahan, S.; Wang, W. HWRF Based Ensemble Prediction System Using Perturbations from GEFS and Stochastic Convective Trigger Function. *Tropical Cyclone Research and Review* **2014**, 3, 145 – 161. doi:https://doi.org/10.6057/2014TCRR03.02.
22. Wang, X. Incorporating Ensemble Covariance in the Gridpoint Statistical Interpolation Variational Minimization: A Mathematical Framework. *Monthly Weather Review* **2010**, 138, 2990–2995. doi:10.1175/2010MWR3245.1.
23. Buehner, M.; Morneau, J.; Charette, C. Four-dimensional ensemble-variational data assimilation for global deterministic weather prediction. *Nonlinear Processes in Geophysics* **2013**, 20, 669–682. doi:10.5194/npg-20-669-2013.
24. Lorenc, A.C.; Bowler, N.E.; Clayton, A.M.; Pring, S.R.; Fairbairn, D. Comparison of Hybrid-4DVar and Hybrid-4DVar Data Assimilation Methods for Global NWP. *Monthly Weather Review* **2015**, 143, 212–229. doi:10.1175/MWR-D-14-00195.1.
25. Berg, R. Tropical cyclone report: Hurricane Joaquin, 28 September–7 October 2015. NOAA. Technical report, NHC Tech. Rep. AL112015, 36 pp., <http://www.nhc.noaa.gov/data/tcr/>, 2016.
26. Kimberlain, T.; Blake, E.; Cangialosi, J. Tropical cyclone report: hurricane patricia 20–24 October 2015. *National Hurricane Center* **2016**.
27. Stewart, S. Hurricane Matthew (AL142016) 28 September–9 October 2016. National Hurricane Center Tropical Cyclone Report. *National Hurricane Center* **2017**.
28. Biswas, M.K.; Abarca, S.; Bernardet, L.; Ginis, I.; Grell, E.; Iacono, M.; Kalina, E.; Liu, B.; Liu, Q.; Marchok, T.; Mehra, A.; Newman, K.; Sippel, J.; Tallapragada, V.; Thomas, B.; Wang, W.; Winterbottom, H.; Zhang, Z. Hurricane Weather Research and Forecasting (HWRF) Model: 2018 Scientific Documentation. *HWRF Development Testbed Center Tech. Rep* **2018**.
29. Zhang, S.; Pu, Z.; Velden, C. Impact of Enhanced Atmospheric Motion Vectors on HWRF Hurricane Analyses and Forecasts with Different Data Assimilation Configurations. *Monthly Weather Review* **2018**, 146, 1549–1569. doi:10.1175/MWR-D-17-0136.1.
30. Zhang, S.; Pu, Z. Numerical Simulation of Rapid Weakening of Hurricane Joaquin with Assimilation of High-Definition Sounding System Dropsondes during the Tropical Cyclone Intensity Experiment: Comparison of Three- and Four-Dimensional Ensemble?Variational Data Assimilation. *Weather and Forecasting* **2019**, 34, 521–538. doi:10.1175/WAF-D-18-0151.1.
31. Cui, Z.; Pu, Z.; Tallapragada, V.; Atlas, R.; Ruf, C.S. A Preliminary Impact Study of CYGNSS Ocean Surface Wind Speeds on Numerical Simulations of Hurricanes. *Geophysical Research Letters* **2019**, 46, 2984–2992. doi:10.1029/2019GL082236.
32. Lu, X.; Wang, X. Improving Hurricane Analyses and Predictions with TCI, IFEX Field Campaign Observations, and CIMSS AMVs Using the Advanced Hybrid Data Assimilation System for HWRF. Part II: Observation Impacts on the Analysis and Prediction of Patricia (2015). *Monthly Weather Review* **2020**, 148, 1407–1430. doi:10.1175/MWR-D-19-0075.1.