A Model Instability Issue in the NCEP Global Forecast System Version 16 and Potential Solutions

Xiaqiong Zhou[1, 2] and Hann-Ming Henry Juang[1, 3]

1. NOAA/NCEP/EMC, College Park, MD, USA
2. NOAA/OAR/Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA
3. Institute of Atmospheric Physics, National Central University, Jongli, Taiwan, ROC

Correspondence to: Xiaqiong Zhou (xiaqiong.zhou@noaa.gov)

Abstract: The National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS) version 16 encountered a few model instability failures during the pre-operational real-time parallel runs. The model forecasts failed when an extremely small thickness depth appeared at the model's lowest layer during the landfall of strong tropical cyclones. A quick solution was to increase the value of minimum thickness depth, an arbitrary parameter introduced to prevent ~~the occurrence of extremely thin model layers, thus~~ numerical instability. This modification solved ~~the issue of the~~the ~~-~~model's numerical instability with a small impact on forecast skills. It was adopted in GFSv16 to implement this version of the operational system as planned.

Upon further investigation, it was determined that the extremely thin depth.~~Further investigation showed that~~ ~~the extremely small thickness depth~~ was a result of ~~occurred after~~ the advection of geopotential heights at the interfaces of model layers. In the FV3 dynamic core, the horizontal winds at interfaces for advection are calculated from the layer-mean values by solving a tridiagonal system of equations in the entire vertical column based on the Parabolic Spline Method (PSM) with high-order boundary conditions (BCs). We replaced the high-order BCs with zero-gradient BCs for the interface-wind reconstruction. The impact of the zero-gradient BCs was investigated by performing sensitivity experiments with GFSv16, idealized mountain ridge tests, and the Rapid Refresh Forecast System (RRFS). The results showed that zero-gradient BCs can fundamentally solve the instability and have little impact on the forecast performances and the numerical solution of idealized mountain tests. This option has been added to FV3 and will be utilized in the GFS (GFSv17/GEFSv13) and RRFS for operations in 2024.

1. Introduction

The Next Generation Global Prediction System (NGGPS) of the National Centers for Environmental Prediction (NCEP) is evolving into the Unified Forecast System (UFS). It is

designed to be the source system for NOAA's operational numerical weather prediction applications and acts as the foundation to better align collaboration with the U.S. modeling community (Ji and Toepfer 2016). The Geophysical Fluid Dynamics Laboratory (GFDL) Finite-Volume Cubed-Sphere (FV3) was chosen as the dynamical core for NGGPS in 2016 (Putman and Lin, 2007; Harris and Lin, 2013). The first major NGGPS model package was successfully implemented within the Global Forecast System (GFS). It became operational on 12 June 2019 as the GFS version 15 (referred to as GFDv15) to replace a legacy spectral model. It was further updated from version 15 to 16 (referred to as GFSv16) on 22 March 2021 with an increased number of vertical layers and model physics upgrades.

To fully assess GFSv16's forecasting ~~The~~performance, both the retrospective and real-time experiments were conducted, covering part of the 2018 hurricane season and the period from May 10, 2019, to real-time until~~before~~ the official implementation ~~was carried out to comprehensively evaluate the forecast performance of GFSv16~~. GFSv16 ~~has improved forecast skills compared with~~outperformed GFSv15 in many aspects, including ~~such as~~ better 500-hPa height anomaly correlation scores and synoptic patterns in the medium range, a better position of relevant frontal boundaries, reduced low-level cold bias during the cool season, and improved Quantitative Precipitation Forecast (QPF) Equitable Threat Scores (ETS) and biases in the medium range.

GFS is the most important operational global weather forecast system at NCEP/Environmental Modeling Center (EMC). It is not only widely used around the world, but also most of NCEP's forecast systems depend on GFS products. The stability of this operational system is critical to delivering reliable real-time products to its users and downstream forecast systems. GFSv16 encountered model instability issues as several cases during the real-time parallel runs crashed before reaching a 16-day forecast length. The diagnosis of the problematic cases in GFSv16 and corresponding proposed fixes are summarized in this study. The numerical model used in GFSv16 is introduced in Section 2. The diagnostic results are summarized in Section 3. Two potential solutions to fix model instability issues are discussed in Section 4. Section 5 introduces the impact of proposed fixes on forecast performances with sensitivity experiments. Summary and discussion are provided in Section 6.

2. Model configuration upgrades

GFSv16 uses a GFDL FV3-based model as its previous version GFSv15. A detailed description of the FV3 dynamic code can be found in the published papers of the GFDL FV3 team (Lin and Rood, 1997; Lin 2004; Harris and Lin 2013; Putman and Lin, 2007; Harris et al. 2020ab, Harris et al. 2021). Only a short summary is given here.

The cubed-sphere grid of the GFDL FV3 uses the equidistant gnomonic projection (Putman and Lin 2007) which divides each cube edge into N segments of equal length and generates a regular mesh on a sphere by connecting non-orthogonal coordinate lines along great circles between two opposite cubic edges. This projection has the advantage of being both equal-area and conformal, which allows for accurate representation of physical processes in both the horizontal and vertical dimensions.

There are two levels of time-stepping inside FV3. The inner time step (also referred to as the acoustic time step) is the integration of the dynamics along the Lagrangian surfaces, which includes computing the forward in-time horizontal flux terms along the Lagrangian surface, and the pressure-gradient force and elastic terms evaluated backwards-in-time. The outer time step is the vertical remapping process to re-grid the deformed Lagrangian surface to a reference coordinate.

The governing equations in FV3 in each horizontal layer are fully-compressible flux-form vector-invariant Euler equations (Harris and Lin, 2013). The momentum flux transportation is represented as vorticity flux and the gradient of the kinetic energy without gradients of vectors. The horizontal discretization of FV3 is derived using a two-grid system with the prognostic winds staggered on a D-grid and C-grid winds used to calculate the face-normal and time-mean fluxes across the cell interfaces (Lin and Rood, 1997). The C-grid winds are interpolated from D-grid winds and then advanced a half time step as the D-grid, except with lower-order fluxes, for efficiency.

The scalar advection scheme is based on the piecewise-parabolic method (PPM; Colella and Woodward, 1984) with a two-dimensional combination of one-dimensional flux methods (Lin and Rood, 1996). The same subgrid reconstruction unlimited scheme is used for mass, potential temperature, vorticity and momentum. The transport of tracers uses a simplified monotonicity constraint (Lin and Rood, 1997) and Huynh's second-order constraint (Putman and Lin, 2007).

The evaluation of the pressure-gradient force in FV3 remains a 4^{th}-order accuracy and is consistent with Newton's 3^{rd} law of motion and achieved by finite-volume integration about a grid

cell (Lin 1997). The "Vertically Lagrangian" dynamics of Lin (2004) were extended with the non-hydrostatic pressure gradient computation of Lin (1997) and included a traditional semi-implicit solver for fast vertically propagating sound waves and gravity waves with efficient computation and great accuracy.

The Lagrangian vertical coordinate (Lin, 2004) is one unique aspect of the FV3, in which each vertical layer resembles that of a shallow water system and is allowed to deform freely during the horizontal integration. It is periodically remapped by vertically redistributing mass, momentum, and energy to a predefined Eulerian coordinate to prevent severe distortion of the Lagrangian surfaces. Vertical transport occurs implicitly from horizontal transport along Lagrangian surfaces.

GFSv16 is built on 13-km quasi-uniform grids having six tiles globally with each tile having 768×768 grid cells. The physics time step is 150 seconds, ~~,~~ ~~i~~In GFSv16 with~~,~~ the "remapping" time step ~~is~~ 75 seconds and the shortest acoustic timestep ~~is~~ 12.5 seconds ~~In GFSv16,~~. The model uses the sigma pressure hybrid coordinate with near surface sigma levels, blended sigma/constant pressure levels in mid-atmosphere, constant pressure levels above.

The major upgrade of GFSv16 from GFSv15 includes an increased number of vertical layers from 64 and 127 with the extended model top from 54 km to 80 km and physics upgrades. The upgraded physics parameterization includes a new scheme to parameterize both stationary and non-stationary gravity waves (Alpert et al. 2019; Yudin et al. 2016;~~,~~ Yudin et al. 2018), a new scale-aware turbulent kinetic energy-based moist eddy-diffusivity mass-flux vertical turbulence mixing scheme to better represent the planetary boundary layer processes (Han and Bretherton, 2019), the improved solar radiation absorption by water clouds and the cloud-overlapping algorithm for the Rapid Radiative Transfer Model for GCMs (RRTMG) (Iacono et al., 2008), and improved GFDL cloud microphysics for computing ice cloud effective radius (Harris et al., 2020ab, Zhou et al. 2019).

3. The study of failed cases

There were eight failed cases during the GFSv16 retrospective and real-time parallel run. They are the cases with the forecast starting times at 00Z 22 Sep. 2018, 18Z 22 July 2020, 06Z 2 Sep. 2020, 06Z and 18Z 3 Sep. 2020, 12Z 4 Sep. 2020, 06Z 5 Sep. 2020, 00Z 6 Sep. 2020 respectively.

A series of sensitivity tests were performed to increase the model stability. Several methods available in FV3 for numerical diffusion to maintain model stability and control energy cascading

were tested. For example, a Rayleigh damping method can be used to dampen the winds to zero with the shortest timescale (tau) at the top increasing with pressure until reaching a defined cutoff pressure level. The minimum timescale and the cutoff pressure level were tuned to apply a stronger Rayleigh damping. Other parameters such as the non-dimensional divergence damping coefficient, the Smagorinsky-type damping coefficient, and the parameters that control the sponge-layer damping to the top three layers of the model were also tuned. However, the instability issues could not be completely solved with these modifications. Although some of the cases were able to finish 16-day forecasts, not all of them became stable, indicating that further improvements were needed to address the model stability issue.

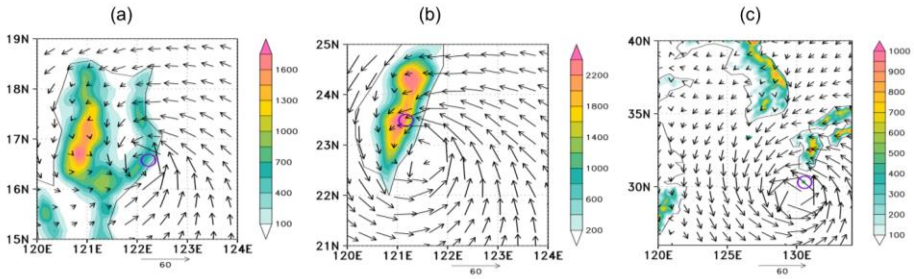

Fig. 1 Horizontal ~~W~~wind fields (m/s) at the lowest level before model crash from the cases with the initial starting time at a) 18Z 22 July 2020, b) 00Z 22 Sep. 2018 and c) 06Z 2 Sep. 2020. The shading is terrain height (unit: m). The open circles mark the location of the crash.

The diagnosis of these cases indicated that all model failures were related to the landfall of strong tropical cyclones. Negative layer thickness in the pressure between lower and upper interfaces or not-a-number (NaN) layer thickness in geopotential height was observed. All failures occurred at grid points located over land when the eyewall of a strong tropical cyclone made landfall from the east. For example, the forecast starting from 18Z July 22 2020 failed when a strong tropical cyclone reached the Philippine east coast with strong onshore winds of about 40-50 m/s (Fig. 1a). In another case, the forecast was interrupted at a grid over the Taiwan Central Mountain area when a strong tropical cyclone started to make landfall. The other six cases were related to Tropical Cyclone Haiseng (2020) when it approached Yakushima Island south of Japan (Fig. 1c).

By examining the model prognostic variables in each acoustic time step (12.5s), we found that unrealistic downdrafts occurred before the failure of the model integration. Figure 2 shows that the vertical motion at the specific grid point increases with time in the case with the initial time at 18Z 18 July 2020. The updraft greater than 5 m/s abruptly changes to an unrealistically large downdraft with an amplitude greater than 200 m/s in one acoustic time step, which directly results in the model failure in the next time step. Figure 3 shows a similar variation of the vertical motion in the case with the initial forecast time at 0000 UTC on Sep. 22, 2018. Similar phenomena were observed in other six cases related to the landfall of Haiseng (2020) (not shown).

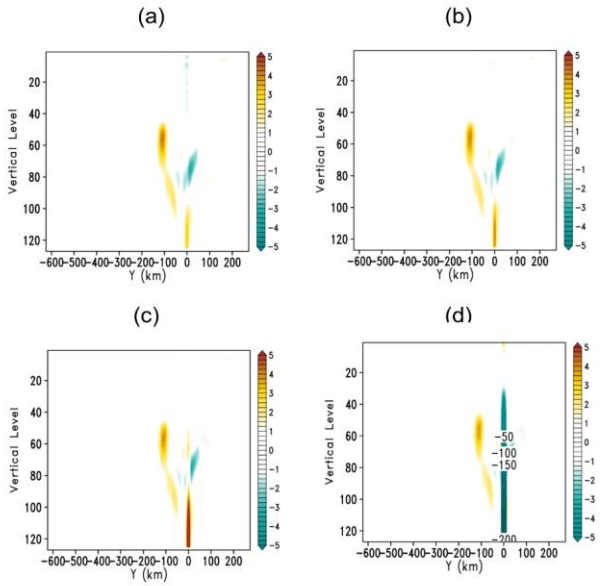

Fig. 2 Vertical section of vertical velocity (m/s) through the location of the crash along with the model y-directional grids before the crash for the case with the initial ~~starting~~ forecasting time at 18Z 18 July 2020. A), b), c), and d) represent 4, 3, 2, and 1 acoustic time step before crash respectively. The y-axis represents the number of vertical levels, with level 1 being at the top of the model.

The hydrostatic and nonhydrostatic solvers in FV3 are "switchable" at runtime through a name list option (Harris et al. 2020ab). The nonhydrostatic solver augments the hydrostatic solver by introducing the prognostic variables, namely vertical velocity *w*, non-hydrostatic pressure and

height thickness δz. The nonhydrostatic pressure is diagnosed as a deviation with $p' = p - p^*$ where p is full pressure calculated from the ideal gas law:

$$p = (R_d \theta_v \frac{\delta m}{\delta z})^{\gamma} \quad (1)$$

δm and δz are the mass and height thickness. $\theta v = Tv\,(\frac{p_0}{p})^{\kappa}$ is virtual potential temperature, where p_0 =1 Pa in FV3, $\kappa = \left(1 + \frac{C_{vm}}{R_d(1+\epsilon q_v)}\right)^{-1}$, and Tv is the "condensate modified" virtual temperature. R_d is the gas constant for dry air. The parameter $\gamma = (1 + \kappa)^{-1}$. C_{vm} is the "moist" specific heat capacities under constant volume and q_v is specific humidity.

Non-hydrostatic pressure perturbation p' and w in the Lagrangian vertical coordinates are solved using a semi-implicit solver, in which the fully implicit time-difference scheme yields a tridiagonal matrix system of equations for vertical velocity w. This system requires coefficients and weights related to p' and layer thickness δz to solve w with the Thomas algorithm (Thomas, 1949). ~~In the FV3 model, all the algorithms are formulated in a finite-volume manner. This means that the above variables of interest are represented as cell- or layer-means.~~

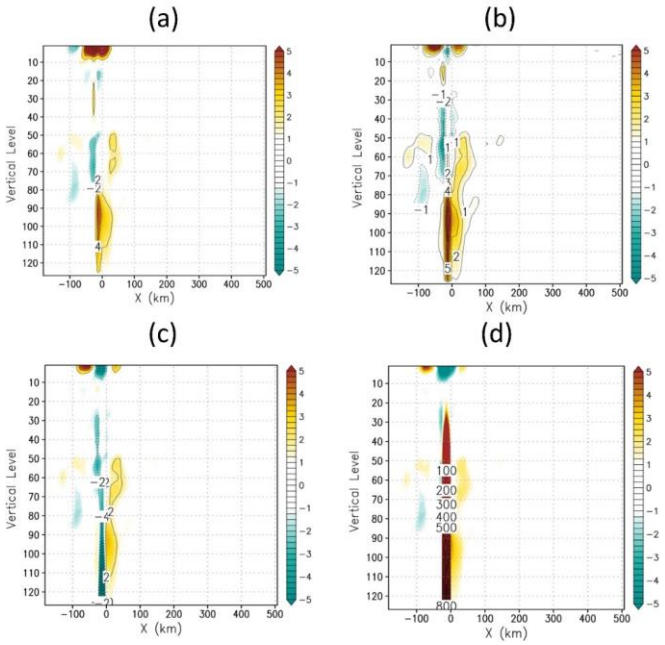

Fig. 3 As Fig. 2 except for the case with an initial starting time of 00Z 22 Sep. 2018 at each

Formatted: Indent: Left: 0", Hanging: 0.44"

acoustic time step before the crash.

All relevant variables before the crash time including p^*, p, θ_v, δz, and the mass δm to calculate w were investigated. In the FV3 model, all the algorithms are formulated in a finite-volume manner. This means that the above variables of interest are represented as cell- or layer-means. Figure 4 shows that θ_v and δm remain reasonable and consistent before the crash (Figs. 4c and 4d). The unrealistic value of the full pressure (larger than 5000 hPa) appears at the model's lowest level at about 200 seconds before the model crash (Fig. 4a), while the hydrostatic pressure remains reasonable (about 900 hPa) with time (Fig. 4b). The slight discontinuity of these variables every six acoustic time steps is a result of the vertical remapping process. GFSv16 has 127 vertical layers with the lowest layer about 20 m thick on average. The value of δz close to zero 200s prior to the crash is quite unusual (Fig. 4e). The unrealistically increased p' and the full pressure at the lowest level before the crash come from the occurrence of extremely small δz while computed from the ideal gas law formula. Extremely large downdrafts are generated through the non-hydrostatic semi-implicit solver from p', which eventually leads to the model failure. The model instability is a result of the presence of extremely small δz at the lowest model level.

The calculation of δz, the vertical difference of geopotential height z between the Lagrangian surfaces before the non-hydrostatic adjustment was investigated. The forward-in-time advective

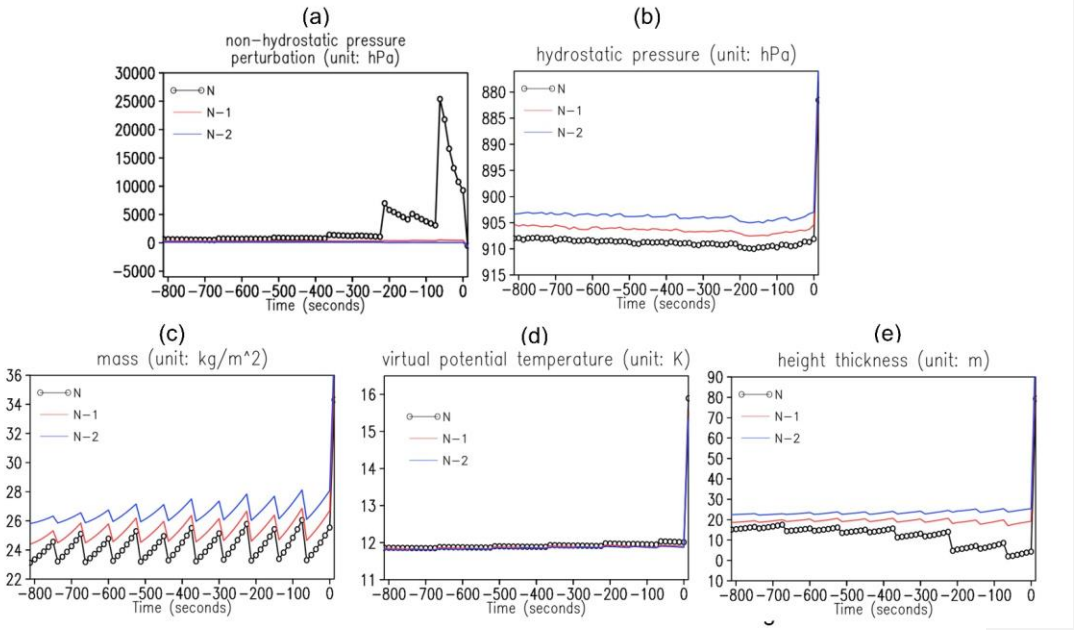

Fig. 4 The time series of a) non-hydrostatic pressure, b) hydrostatic pressure, c) mass, d) virtual potential temperature and e) thickness depth in height at the crash grid for the case with the initial starting time 18Z 18 July 2020. The black curves represent the model's lowest level (marked by ~~N~~N), while the blue and red represent the second and third lowest levels (marked with ~~km~~N-1 and N~~km~~-2). The open circle marks each acoustic time step in the time series.

processes are performed to generate the partially-updated z before the non-hydrostatic adjustment in the FV3 dynamics. Note that the update of z through advection processes does not directly solve an equation for the volume of a grid cell (δz) and it is forward-in-time as the sum of the advective height flux along with the Lagrangian interfaces and the vertical distortion of the surfaces by the gradient of z. The previous study found that the advection of δz created excessive noise near steep topography (Harris et al. 2021) and it was more difficult to guarantee the kinematic surface condition without perpendicular flow to the surface with the advection of δz.

To advance z on the interfaces, the advection winds are interpolated from layer means onto the layer interfaces. Figure 5 shows the time series of z at the crash location before and after the advection process. The value of z at the lowest level before advection remains constant as it is the height of the topography (Fig. 5a). After advection, there is a significant change ~~It has a great change after advection and becomes very close to~~as and it becomes close to the value of z at the second lowest level (Fig 5b), starting around~~beginning about~~ 200 seconds before the crash. There are no ~~significant~~ noticeable changes of z ~~before and after the advection~~ at the second and third lowest levels both before and after the advection. As δz at the lowest level represents the difference between the height at the lowest two interfaces, t~~T~~he advection process at the lowest level is responsible for the decreased thickness depth seen in Fig. 4e.

4. Potential solutions

The forward-in-time advection of geopotential height is a part of the acoustic time step in which the Lagrangian surface is allowed to freely deform. An artificial limiter is defined as the minimum thickness depth after the geopotential height advection to enhance its monotonicity in the vertical. This limiter is defined as *dz_min* in FV3 with 2 meters as the default in FV3. It only takes in effect when the thickness of geopotential occurred in the model is smaller than the default value. We found that increasing this limiter value from 2 to 6 meters can effectively avoid model crashes. All eight cases can finish 16-day forecasts with this modification.

To examine whether increasing this artificial limiter violates general model states in GFSv16, the possibility for δz to reach the minimum thickness depth of 6 was investigated in both crash cases and successful cases. The successful cases were randomly selected from the retrospective runs among the cases that can complete 16-day forecasts successfully. There were no extremely small δz values in any grids from randomly selected successful cases. δz less than 6 meters likely only occurred at the breakpoint in crash cases and~~. This examination indicates that~~ this artificial limiter is only used in very rare situations. Forecast-only experiments also showed that this fix had very little impact on the forecast skill. Since any changes in the forecast performance were not desirable at the final retrospective test stage for the implementation of GFSv16, this method was considered a suitable temporary fix for GFSv16. It was adopted for the GFSv16 implementation ~~and this fix allowed GFSv16 to~~so that GFSv16 can be implemented at the time.

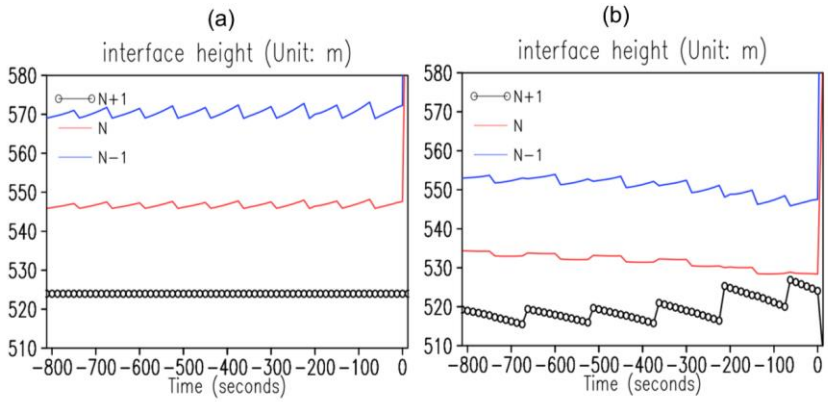

Fig.5 As Fig. 4 except for the geopotential height at the model's three lowest interface layers (marked as N+1, N, N-1 respectively) a) before and b) after the advection procedure for the case with the initial starting time 18Z 18 July 2020.

A sensitivity experiment was performed by restarting the model about 1 hour before the crash with increased minimum thickness depth. The geopotential height after the advection was forced to be greater than the artificial limiter. The abrupt change of the geopotential height was observed at the original crash location and time, then it backed to a normal range after several acoustic time steps (Fig. 6a). The model can successfully finish 16-day forecasts. The increased minimum thickness depth can prevent the model from crashing, but it does not fundamentally solve the model instability issue. In addition, this arbitrary limiter should be used with caution and the height of the model's lowest level should be considered to select a reasonable value for the limiter.

The advection process to update *z* in FV3 was examined since the model instability issue likely originated from the advection of *z* at the model's lowest level. To update z, the advection winds and vertical velocity are reconstructed from layer means onto the layer interfaces by solving a tridiagonal system of equations based on the Parabolic Spline Method (PSM, Zerroukat, et al., 2006).

The following equation represents the relationship between the interface value $\hat{q}_{i-\frac{1}{2}}$ and layer-mean value $\bar{q}_i$ (~~i~~i=1, 2, …N) from the model top to the lowest level in a computational one-dimensional discretized domain with PSM (Zerroukat et al., 2006):

$$\frac{1}{h_i}\hat{q}_{i-\frac{1}{2}} + 2\left(\frac{1}{h_i} + \frac{1}{h_{i+1}}\right)\hat{q}_{i+\frac{1}{2}} + \frac{1}{h_{i+1}}\hat{q}_{i+\frac{3}{2}} = 3\left(\frac{1}{h_i}\bar{q}_i + \frac{1}{h_{i+1}}\bar{q}_{i+1}\right) \quad (2)$$

where h_i is the spatial interval between two interfaces $h_i = z_{i+\frac{1}{2}} - z_{i-\frac{1}{2}}$(i=1, 2...N) and q represents horizontal wind components u and v here. Equations (2) define a linear system of equations for the unknown interface values $\hat{q}_{i-\frac{1}{2}}$ in terms of the layer-mean values $\bar{q}_i$. Boundary

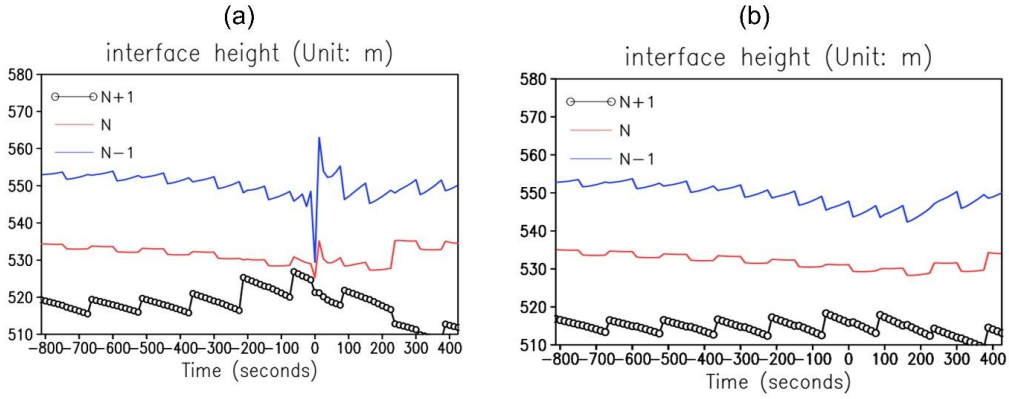

Fig. 6 The time series of the geopotential height at the model's three lowest interface layers (marked as N+1, N, N-1 respectively) after the advection at the original crash location in the sensitivity experiments with proposed fixes: a) increased *dz_min* from 2 to 6 and b) zero-gradient BCs for the case with the initial starting time 18Z 18 July 2020.

conditions are required to close the problem. FV3 uses the following equations for the upper and lower boundary to solve the horizontal winds at the model interfaces as a problem of a tridiagonal system:

$$g_1(g_1 + 0.5)\hat{q}_{\frac{1}{2}} + [1 + g_1(g_1 + 1.5)]\hat{q}_{\frac{3}{2}} = 2g_1(1 + g_1)\bar{q}_1 + \bar{q}_2 \quad (3)$$

$$[1 + g_N(g_N + 1.5)]\hat{q}_{N-\frac{1}{2}} + g_N(g_N + 0.5)\hat{q}_{N+\frac{1}{2}} = \bar{q}_{N-1} + 2g_N(1 + g_N)\bar{q}_N \quad (4)$$

where $g_1 = h_2/h_1$ and $g_N = h_N/h_{N-1}$ with level 1 at the model top and level N at the lowest level.

We proposed to use zero-gradient BCs, that is $\frac{dq}{dz}$=0 at the endpoints $z_{\frac{1}{2}}$ and $z_{N+\frac{1}{2}}$ corresponding to an assumption of zero slope there. Applying these zero-gradient BCs leads to

$$2\hat{q}_{\frac{1}{2}} + \hat{q}_{\frac{3}{2}} = 3\bar{q}_1 \quad (5)$$

$$\hat{q}_{N-\frac{1}{2}} + 2\hat{q}_{N+\frac{1}{2}} = 3\bar{q}_N \quad (6)$$

The original BCs used in FV3 as shown in Eq. (3) and (4) are named as high-order BCs thereafter in contrast with the zero-gradient BCs we proposed.

The comparison of the vertical profiles with two different BCs shows that the reconstructed winds are similar in these two types BCs when the vertical shear of the layer-mean winds in the lower levels is relatively small (Fig. 7a). With larger vertical shear, the overshooting and undershooting of the reconstructed winds at the lowest two layers are more evident by using higher-order BCs than zero-gradient BCs while interior winds remain similar (Fig. 7b). The vertical shear of interface winds at the lowest several layers are smaller with zero-gradient BCs than with high-order BCs.

With the application of the zero-gradient BCs, all originally crashed cases can finish 16-day forecasts successfully. A sensitivity experiment was performed similarly for zero-gradient BCs. Fig. 6b shows that applying the zero-gradient BCs avoids unrealistic δz values. No extremely small δz was found during the model integration. This method is better than increasing the artificial thickness depth limiter as it fundamentally solves the occurrence of unrealistic δz values at the model's lowest level.

PSM is third-order accurate in space for a non-uniform grid and fourth-order accurate for a uniform grid (Zerroukat et al. 2006). The reconstructed winds at the BCs with high-order BCs may retain a relatively higher accuracy. However, it can be worse in the case of sharp/under-resolved gradients with significant overshoots/undershoots due to a larger degree of freedom. Constraints are usually required for the reconstruction to prevent overshoots/undershoots with respect to the layer-mean values (Shchepetkin and McWilliams, 1998; Zerroukat et al. 2006). Our method is reducing the order of the reconstruction polynomial for BCs. It is worth noting that the zero-gradient condition is only used at the model's upper and lower edge levels. The parabolic spline as the reconstructed function remains valid for the inner layers. In addition, the reconstructed horizontal winds are only used for the advection of geopotential height. The revised BCs do not impact the layer-mean prognostic wind fields directly.

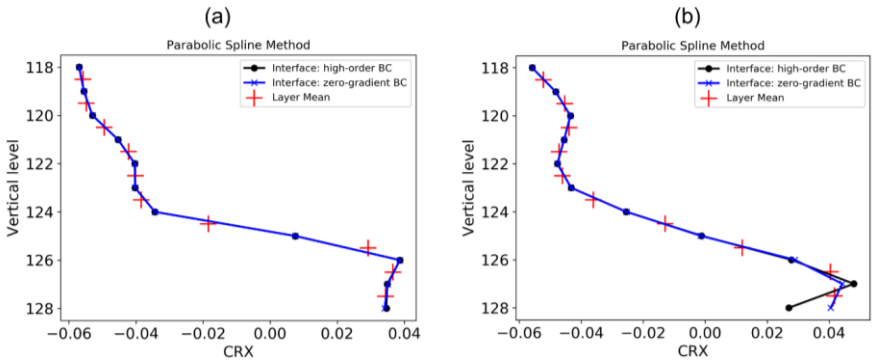

Fig. 7 The vertical profile of Courant numbers in x-axis ($c_x = \triangle t \cdot u/\Delta x$) at two grids with a) smaller and b) larger vertical gradients. The red cross represents the layer-mean value while the black and blue represent the interface values reconstructed with high-order BCs and zero-gradient BCs.

5. Sensitivity experiments with zero-gradient BCs

The impact of zero-gradient BCs on forecast performance was investigated with different model configurations. The experiment design and results are discussed including idealized mountain ridge tests and real-case tests with the same configuration as GFSv16 and the high-resolution regional application in EMC.

The mountain waves could be sensitive to the model's lower boundary conditions (Smith 2007). The impact of the BC change on the geopotential height advection on the mountain waves was investigated. An idealized mountain ridge test with an adiabatic condition, a uniform flow of 8 m/s over a ridge mountain was performed. This is a modified version of the Dynamical Core Model Intercomparison Project (DCMIP) case 2.1 with a quasi-2D mountain ridge with a ridge height of 250m (Ullrich et al., 2016). Instead of assuming a small Earth, the idealized mountain-ridge experiment was tested on a doubly-periodic domain. The model top is 50 hPa with a horizontal resolution of 500 m. Zero-gradient BCs are utilized in the upper and lower boundaries in the sensitivity experiment. The mountain wave patterns are similar in these two experiments. Although slightly larger at the lowest levels, the difference ~~between the two zonal wind fields~~ is ~~still~~ negligible throughout the entire domain. (Fig.8). At both the upper and lower boundaries of the model, identical boundary condition methods are implemented. However, the damping applied

to absorb vertically propagating waves may have a lesser effect at the upper boundary. This can be attributed to two forms of damping utilized, namely Rayleigh damping which reduces wind speed to zero within a specific timescale, and a sponge damping layer that employs second-order damping to divergence, vorticity, mass, and w-flux in the top three layers of the model.

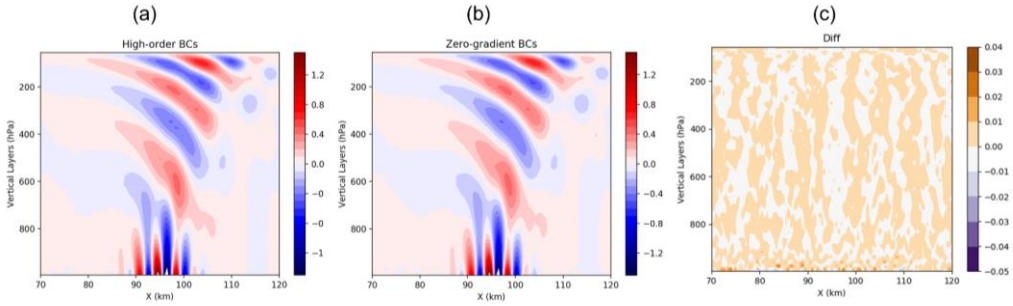

Fig. 8 Cross sections of vertical velocity (m/s) along the equator for orographic mountain ridge on the earth (quasi-2D ridge in a barotropic zonal flow). The x-axis is measured in kilometers (km) to represent distance and the y-axis is the vertical coordinate in pressure (hPa) in (a) the control and (b) the sensitivity run with zero-gradient BCs. c) is the difference between these two runs.

A group of sensitivity experiments was performed by using the GFSv16 as the control. A sensitivity experiment was performed by replacing high-order BCs in the control with zero-gradient BCs at both the lower and upper boundaries of the model to reconstruct horizontal winds with PSM were replaced. 10-day forecasts were compared with initial times from June to October 2020 every five days with 00Z only. The EMC Global NWP Model Verification Package was used for the verification (Yang et al. 2006). This verification package is a standard evaluation tool for the GFS upgrade and implementation with verification scores comparing gridded model data to both point-based rawinsonde and surface station observations and GFS gridded analysis. The model forecast statistics in terms of the Root Mean Square Error (RMSE), bias, and anomaly correlation for conventional variables, as well as tropical cyclone intensity and track forecasts over the Atlantic, Eastern Pacific, and West Northern Pacific and precipitation threat skill scores over CONUS. The comparison of these forecast verification metrics shows that the sensitivity experiments with zero-gradient BCs have similar forecast performance without significant differences to those of GFSv16 with high-order BCs (not shown).

The Rapid Refresh Forecast System (RRFS) is another important FV3-based UFS application in EMC. It is the NOAA next-generation convection-allowing, rapidly-updated ensemble prediction system of a limited area (The continental U.S, CONUS) (Black et al., 2021), currently scheduled for operational implementation in late 2023. The operational configuration features a 3 km grid spacing covering North America and include forecasts every hour out to 18 hours, with extensions to 60 hours four times per day at 00Z, 06Z, 12Z, and 18Z. Each forecast is planned to be composed of 9-10 members.

The impact of the zero-gradient BCs on the high-resolution forecasts was also investigated based on the RRFS configuration. The ensemble members with the Mellor–Yamada–Nakanishi–Niino (MYNN) (Nakanishi 2001; Olson et al., 2019) Planetary Boundary Layer (PBL) and Thompson MP scheme were used as the control to investigate the impact of zero-gradient BCs. Figure 9 shows that the precipitation distribution from 12-36 hours in the experiment with zero-gradient BCs resembles that in the control. The use of zero-gradient BCs does not significantly change the forecast results for high-resolution forecasts.

6. Summary and discussion

GFS is one of the most important operational global weather forecast systems at NCEP/EMC. The stability of GFS on model integration is as important as its forecast skills to deliver dependable real-time products to its users and downstream forecast systems. The model instability issue of GFSv16 caught our attention when several cases in its real-time parallel runs failed to finish 16-day forecasts. The analysis of these cases showed that the model integration was interrupted after the presence of a very thin layer depth corresponding to a largely deformed layer surface at the model's lowest level in tropical cyclones during the landfall after the advection of geopotential height.

An artificial limiter is defined in FV3 to ensure that the minimum layer depth in FV3 after the advection is not less than a default value to maintain the monotonicity of geopotential height in the vertical. Sensitivity tests showed that increasing the value of this artificial parameter from the default value of 2 meters to 6 meters can fix the model instability issue. An abrupt change of geopotential height at the model's lowest interface was still observed with an increased value of the limiter, but all previously crashed cases can finish 16-day forecasts. This method was effective to solve the model instability issue and was adopted to GFSv16 so that the GFSv16 can be implemented as scheduled. Nevertheless, this method lacks a scientific foundation and the root reason corresponding to the model instability remains unknown.

Formatted: Indent: First line: 0.31"

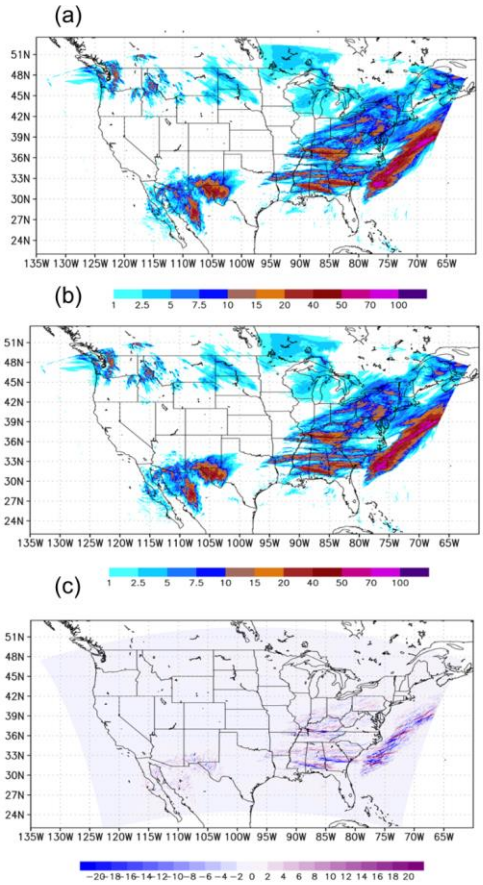

Fig. 9 24-hr precipitation (mm) from 12-36 hour in (a) the control and (b) the sensitivity run with zero-gradient BCs. c) is the precipitation difference between these two runs. The control run has the same configuration as the ensemble member one in RRFS with the initial starting time on 00Z March 2, 2020.

~~crashed cases can finish 16-day forecasts. This method was effective to solve the model instability issue and was adopted to GFSv16 so that the GFSv16 can be implemented as scheduled. Nevertheless,~~

~~this method lacks a scientific foundation and the root reason corresponding to the model instability remains unknown.~~

Further investigation suggested that the presence of an extremely thin layer at the model's lowest layer was related to the reconstruction of interface winds from layer mean winds for the advection of geopotential height along the Lagrangian surfaces. In FV3, the horizontal winds are calculated from layer means onto the layer interfaces by solving a tridiagonal system of equations based on PSM (Zerroukat, et al., 2006) with high-order BCs. It was found that the high-order BCs easily produce overshoots or undershoots in areas with large vertical wind shear. The lower boundary in a landfall tropical cyclone was a perfect condition for the occurrence of overshoots/undershoots with high-order BCs, which led to a heavily distorted Lagrangian surface and triggered unstable conditions. The change of BCs from high-order to zero-gradients at the lowest interface removed spurious under/overshoots near steep terrain with vertical wind shears, thus avoiding a distorted geopotential height interface so that the model remains in stable conditions.

The impact of the zero-gradient BCs for the tridiagonal system on the forecast results was very minor. The zero-gradient condition for BCs was only valid at the model's lowest/highest interface. The reconstructed horizontal wind profile at sub-grids remained a parabolic spline as defined in terms of the layer-mean values. In addition, the reconstructed interface horizontal winds were only used in the advection of geopotential height. The zero-gradient BCs did not impact the prognostic layer-mean wind fields directly. The zero-gradient BCs had been committed to the Unified Forecast System (UFS) as an alternative method for the forward-in-time advection of geopotential height.

Even though the model instability issue only was found during the landfall of tropical storms in GFSv16, it could be the case in any situations with strong vertical shear of winds at the lower and upper boundary. It was found that the zero-gradient BCs can effectively improve the model forecast stability for RRFS in non-tropical cyclone cases. This option has been included in the RRFS package for the real-time parallel runs. For the GFS, the artificial limiter used in GFSv16 will be replaced by the option of zero-gradient BCs to stabilize the model forecasts in the next-generation coupled GFSv17/Global Ensemble Forecast System (GEFS v13). Both the RRFS and GFSv17/GEFSv13 target operational implementation in 2024.

Code/Data availability: The numerical model simulations upon which this study is based are too large to archive or to transfer. Instead, we provide all the information needed to replicate the simulations; we used the model version GFSv16. The model code, the scripts to compile and run the model, and the scripts to reproduce the figures in this work are available at https://doi.org/10.5281/zenodo.7555839. The initial condition files used in this study are the GFS/GDAS analysis data but only the recent production is available for the public at https://www.ftp.ncep.noaa.gov/data/nccf/com/gfs/prod/. Two potential fixes we discussed in this paper can be tested by adding *dz_min* or *psm_bc* in the model input name list.~~The model code, compilation script, the scripts to run the model and the name list setting are available at https://github.com/NOAA-EMC/global-workflow/tree/gfs.v16.2.2. The initial condition files used in this study are the GFS/GDAS analysis data but only the recent production is available for the public at https://www.ftp.ncep.noaa.gov/data/nccf/com/gfs/prod/. Two potential fixes we discussed in this paper can be tested by adding *dz_min* or *psm_bc* in the model input name list.~~

Author contribution: Xiaqiong Zhou and Hann-Ming Henry Juan designed the experiments and Xiaqiong Zhou carried them out. Xiaqiong Zhou prepared the manuscript with contributions from the co-author.

Acknowledgment: We thank the GFDL modeling group for their support, especially Lucas Harris, Xi Chen and Linjiong Zhou, and EMC colleagues as well, such as Fanglin Yang, Sajal Kar, and others for their insightful suggestions and discussions. We also thank Miodrag Rancic and Kevin Viner for their careful EMC internal review.

References

Alpert, J. C, Yudin, V. A., and Strobach, E.: Atmospheric gravity wave sources correlated with resolved-scale GW activity and sub-grid scale parameterization in the FV3gfs model, ~~In~~ AGU Fall Meeting 2019, San Francisco, CA, USA, 9 -13 December 2019.

Black, T. L., Abeles, J. A., Blake, B. T., Jovic, D., Rogers, E., Zhang, X., Aligo, E. A., Dawson, L. C., Lin, Y., Strobach, E., Shafran, P. C., and Carley, J. R: A limited area modeling capability for the Finite-Volume Cubed-Sphere (FV3) dynamical core and comparison with a global two-way nest, J. Adv. Model Earth Sys., **13,** e2021MS002483. https://doi.org/10.1029/2021MS002483, 2021.

Colella, P., and Woodward, P. R.: The piecewise parabolic method (PPM) for gas-dynamical simulations, J. Comput. Phys., 54, 174-201, 1984.

Han, J. and Bretherton, C. S.: Scale-aware TKE-based moist eddy-diffusivity mass-flux (EDMF) parameterization for vertical turbulent mixing interacting with cumulus convection, Wea. And Forecasting, 34(4), 869–886, 2019.

Harris, L. M., and Lin, S. J.: A two-way nested global-regional dynamical core on the cubed-sphere grid, Mon. Wea. Rev., 141(1), 283– 306, 2013.

Harris, L., and Coauthors: GFDL SHiELD: A unified system for weather-to-seasonal: A two-way nested global-regional dynamical core on the cubed-sphere grid, Mon. Wea. Rev., 141(1), 283– 306, 2020a.

Harris, L., Chen, X., Zhou, L., and Chen, J.-H.: The Nonhydrostatic solver of the GFDL ~~Finite~~finite ~~Volume~~ volume ~~Cubed~~cubed ~~Sphere~~ sphere ~~Dynamical~~ dynamical ~~Core~~core, Tech. Rep. 2020-003, Geophysical Fluid Dynamics Laboratory. URL https://repository.library.noaa.gov/view/noaa/27489, 2020b.

Harris, L., Chen, X., Putman, W., Zhou, L., and Chen, J-H.: A scientific description of the GFDL finite-volume cubed-sphere dynamical core. NOAA Technical memorandum OAR GFDL, 2021-001, https://doi.org/10.25923/6nhs-5897, 2021.

Formatted: Font: Not Italic
Formatted: Font: (Default) Times New Roman
Formatted: Font: Not Italic
Formatted: Font: Not Italic
Formatted: Font: Not Italic
Formatted: Font: Not Italic
Formatted: Font: Not Italic
Formatted: Font: Not Italic
Formatted: Font: Not Bold
Formatted: Font: Not Bold

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing by long-lived greenhouse gasses: Calculations with the AER radiative transfer models, J. Geophys. Res., 113, 1-8, 2008.

Ji, M. and Toepfer F.: Dynamical core evaluation test report for NOAA's Next Generation Global Prediction System (NGGPS), NOAA IR ID#-18653, 2016.

Lin, S.-J.: A finite-volume integration method for computing pressure gradient force in general vertical coordinates, Q. J. R. Meteorol. Soc.,Quarterly Journal of the Royal Meteorological Society, 123 (542), 1749–1762, 1997.

Lin, S. J.: A "vertically Lagrangian" finite-volume dynamical core for global models, Mon. Wea. Rev., 132(10), 2293– 2307, 2004.

Lin S.-J. and Rood, R. B.: Multidimensional flux-form semi-Lagrangian transport schemes, Mon. Wea. Rev., 124 (9), 1996.

Lin S.-J. and Rood, R. B.: An explicit flux-form semi-Lagrangian shallow-water model on the sphere, *Q. J. R. Meteorol. Soc.*, 123, 2477-2498, 1997.

Nakanishi, M.: Improvement of the Mellor–Yamada turbulence closure model based on large-eddy simulation data, Bound. Layer Meteor., 99, 349–378, 2001.

Olson, J. B., Kenyon, J. S., Angevine, W. A. Brown, J. M., Pagowski, M. and Suselj, K.: A description of the MYNN-EDMF scheme and the coupling to other components in WRF-ARW. Technical memorandum, NOAA OAR GSD-61, March 2019.

Putman, M. and Lin, S.-J.: Finite-volume transport on various cubed-sphere grids, J. Comp. Phys., 227, 55-78, 2007.

Smith, R. B.: Interacting mountain waves and boundary layers, *J. Atmos. Sci.*, 64, 594-606, 2007.

Shchepetkin, A. F., and McWilliams, J. C.: Quasi-Monotone advection schemes based on explicit locally adaptive dissipation, Mon. Wea. Rev. 126(6), 1541-1580, 1998.

Thomas, L. H.: Elliptic problems in linear differential equations over a network, Watson Sci. Comput. Lab Report, Columbia University, New York, USA, Employing https://www.weather.gov/media/notification/pdf2/scn21-20_gfsv16.0_aac.pdf, 1949.

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Ullrich, P. A., Jablonowski, C., Kent, J., Lauritzen, P. H., Nair, R., Reed, K. A., Zarzycki, C. M., Hall, D. M., Dazlich, D., Heikes, R., Konor, C., Randall, D., Dubos, T., Meurdesoif, Y.,Chen, X., Harris, L., Kühnlein, C., Lee, V., Qaddouri, A., Girard, C., Giorgetta, M., Reinert, D., Klemp, J., Park, S.-H., Skamarock, W., Miura, H., Ohno, T., Yoshida, R., Walko, R., Reinecke, A., and Viner, K.: DCMIP2016: a review of non-hydrostatic dynamical core design and intercomparison of participating models, Geosci. Model Dev., 10, 4477–4509, https://doi.org/10.5194/gmd-10-4477-2017, 2017.

Yang, F. H-L Pan, S. K. Krueger, Moorthi, S. and Lord, S. J.: Evaluation of the NCEP Global Forecast System at the ARM SGP site, Mon. Wea. Rev, 134, 3668-3690, 2006.

Yudin, V. A., R. A. Akmaev, T. J. Fuller-Rowell, and J.C. Alpert, 2016: Gravity wave physics in the NOAA environmental modeling system. ~~In~~ International SPARC Gravity Wave Symposium, volume 48, page 012024, 2016.

Yudin V. A, Akmaev, R. A. Alpert, J. C., Fuller-Rowell T. J., and Karol S. I., 2018: Gravity wave physics and dynamics in the ~~fv3~~FV3-based atmosphere models extended into the mesosphere, ~~I~~In Am. Meteorol. Soc., ~~editor,~~ 25th Conference on Numerical Weather Prediction, Denver, CO, USA, 4-8 Jun, 2018.

Zerroukat, M., Wood, N., and Staniforth, A.: The Parabolic Spline Method (PSM) for conservative transport problems, Inter. J. Num. Meth. In Flu., 51, 1297-1318, 2006.

Zhou, L., Lin, S.-J., Chen, J.-H., Harris, L. M., Chen, X., and Rees, S. L.: Toward convective-scale prediction within the next generation global prediction system, Bull. Amer. Meteo. Soc., 100 (7), 1225–1243, 2019.

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Formatted: Font: Not Italic