Reducing a Tropical Cyclone Weak-Intensity Bias in a Global Numerical Weather Prediction System

Ron McTaggart-Cowan,[a] David S. Nolan,[b] Rabah Aider,[a] Martin Charron,[a] Jan-Huey Chen,[c] Jean-François Cossette,[a] Stéphane Gaudreault,[a] Syed Husain,[a] Linus Magnusson,[d] Abdessamad Qaddouri,[a] Leo Separovic,[a] Christopher Subich,[a] Jing Yang[a]

[a] *Atmospheric Numerical Weather Prediction Research Section, Environment and Climate Change Canada*

[b] *Rosenstiel School of Marine, Atmospheric, and Earth Science, University of Miami*

[c] *Geophysical Fluid Dynamics Laboratory, National Oceanographic and Atmospheric Administration*

[d] *European Centre for Medium-Range Weather Forecasts*

Corresponding author: Ron McTaggart-Cowan, ron.mctaggart-cowan@ec.gc.ca

ABSTRACT: The operational Canadian Global Deterministic Prediction System suffers from a weak-intensity bias for simulated tropical cyclones. The presence of this bias is confirmed in progressively simplified experiments using a hierarchical system development technique. Within a semi-idealized, simplified-physics framework, an unexpected insensitivity to the representation of relevant physical processes leads to investigation of the model's semi-Lagrangian dynamical core. The root cause of the weak-intensity bias is identified as excessive numerical dissipation caused by substantial off-centering in the two time-level time integration scheme used to solve the governing equations. Any (semi-)implicit semi-Lagrangian model that employs such off-centering to enhance numerical stability will be afflicted by a misalignment of the pressure gradient force in strong vortices. Although the associated drag is maximized in the tropical cyclone eyewall, the impact on storm intensity can be mitigated through an intercomparison-constrained adjustment of the model's temporal discretization. The revised configuration is more sensitive to changes in physical parameterizations and simulated tropical cyclone intensities are improved at each step of increasing experimental complexity. Although some rebalancing of the operational system may be required to adapt to the increased effective resolution, significant reduction of the weak-intensity bias will improve the quality of Canadian guidance for global tropical cyclone forecasting.

SIGNIFICANCE STATEMENT: Global numerical weather prediction systems provide important guidance to forecasters about tropical cyclone development, motion and intensity. Despite recent improvements in the Canadian operational model's ability to predict tropical cyclone formation, the system systematically under-predicts the intensity of these storms. In this study, we use a set of increasingly simplified experiments to identify the source of this error, which lies in the numerical time-stepping scheme used to solve the model equations. By decreasing numerical drag on the tropical cyclone circulation, intensity predictions that resemble those of other global modeling systems are achieved. This will improve the quality of Canadian tropical cyclone guidance for forecasters around the world.

1. Introduction

Accurate tropical cyclone predictions are essential for reducing the impacts of the hazards associated with these extreme events (Sharma and Berg 2022). Ongoing improvements in storm track prediction (Landsea and Cangialosi 2018; Heming et al. 2019) have allowed the focus of research efforts to shift towards the problem of forecasting storm intensity (Gall et al. 2013). However, accurately predicting the winds, rains and storm surges that accompany tropical cyclones remains a significant challenge despite recent progress in NWP and operational forecasting techniques (Cangialosi et al. 2020). This is particularly true in basins where storms are not well sampled by instrumented aircraft and for which little high-resolution NWP guidance is available. Meteorologists in such regions depend heavily on global model predictions for tropical cyclone forecasting (DeMaria et al. 2014; Courtney et al. 2019).

Limited spatial resolution in global NWP systems has historically meant that the tropical cyclone vortex is subject to significant spatial under-sampling and a systematic weak-intensity bias (Davis 2018). Even more problematic is the fact that such models are unable to resolve the internal structures and processes that control rapid intensity changes (Rogers et al. 2015). However, improvements in subgrid-scale parameterizations (hereafter referred to as "model physics") and the steady progress of global model resolution into the deep convective gray zone (Stevens et al. 2019) has led to the expectation that these systems should accurately represent most of the tropical cyclone life cycle (Judt et al. 2021).

The Canadian Global Deterministic Prediction System [GDPS; (Caron and Buehner 2022)] is run with a grid spacing of ∼15 km, placing it outside the gray zone but within the typical range for current operational systems. Using the 17 km configuration of the UKMO (UK Met Office) global model, Hodges and Klingaman (2019) identify a weak-intensity bias of 15 $\mathrm{m\,s^{-1}}$ (10 hPa) that they attribute primarily to insufficient resolution of the vortex. Majumdar et al. (2023) show that systematic errors in the wind-pressure relationship can also affect a model's ability to represent maximum wind speeds in the 9 km ECMWF system, a problem that persists even in a 5 km configuration. These expected limitations notwithstanding, the GDPS systematically under-predicts the intensity of mature storms (Yamaguchi et al. 2017). This conditional bias has significant forecasting implications because it hampers the system's ability to provide guidance for associated high-impact weather. This study therefore focuses on reducing the intensity bias in predictions for tropical cyclones of at least tropical storm strength (Simpson 1974).

Identifying the root cause of a systematic error in a complex NWP system is one research challenge; correcting it in a way that minimizes the risk of introducing additional error compensation is another. Frissoni et al. (2023) recommend the use of a hierarchical system development approach for attacking such problems (Jacob 2010), which is implemented using a "hierarchy of complexity" in the current study. This strategy, combined with standard model intercomparison, provides a powerful set of tools with which to identify error sources and to constrain individual components of the system. Here we pursue the hierarchical approach into the dynamical core of the numerical model to connect the tropical cyclone intensity bias to temporal discretization. This study therefore builds on the work of Walters et al. (2017), who found that changing time-integration parameters can affect storm intensity; however, the inclusion of numerous changes to the model made it impossible for the authors to identify the precise origin or extent of the observed sensitivity.

In this study we identify the numerical source of the tropical cyclone weak-intensity bias and design an experimental framework that allows us to develop an optimal dynamical core configuration. Documenting this investigation supports the WMO recommendation that "evaluations and specifics of upgrades to intensity guidance should be communicated to operational [tropical cyclone forecasting] centers." (Courtney et al. 2019). The data, models and methods used in this study are introduced in section 2. Steps down the hierarchy of complexity are taken in section 3, arriving finally at the semi-idealized, simplified-physics configuration used for the bulk of the study

TABLE 1. Gridded atmospheric analyses used in this study. The 1.5° grid spacing for ERA5 refers to a coarse-grained dataset for model evaluation derived from the original 0.28° source.

Product Name	Type	Grid Spacing (°)	Levels	Top (hPa)	Coordinate	Usage	Reference	Sections
CMC Analysis	Operational	0.135	84	0.1	Hybrid pressure	Initialization	Buehner et al. (2015)	2c1, 3a, 5c
ECMWF Analysis	Operational	0.075	137	0.01	Hybrid pressure	Initialization	ECMWF (2018a)	2c2, 3b, 5b
ERA5	Reanalysis	1.5	37	1	Pressure	Evaluation	Hersbach et al. (2020)	5c

(section 4). Once a solution is identified, expected behavior is confirmed as experiments step back up the hierarchy in section 5. The study concludes with a discussion of the findings in section 6.

2. Data, Model and Methods

The hierarchical system development and model intercomparison techniques employed in this study use a wide range of datasets, models, experimental protocols and diagnostic tools, each of which is described in this section.

a. Dataset Descriptions

Three different gridded analyses are used in different contexts as shown in Table 1. The CMC operational analysis is native to the GDPS (Buehner et al. 2015) and therefore provides the most direct estimate of sensitivities within the system. Operational ECMWF analyses are used as initializations for model intercomparisons (ECMWF 2018b). Finally, the ERA5 reanalysis is used as an independent reference for model evaluation.

Tropical cyclone guidance skill is assessed through comparisons with best track information issued by the Regional Specialized Meteorological Centre for each basin. Storm track, maximum wind and minimum central pressure estimates are obtained through the International Best Track Archive for Climate Stewardship [IBTrACS; Knapp et al. (2010)]. Only storms that reach a 35 kt wind-speed threshold are considered in this study (Hersbach et al. 2020).

b. Numerical Models

The Global Environmental Multiscale (GEM) model is used for all operational NWP applications at the Canadian Meteorological Centre. Girard et al. (2014) and Husain et al. (2019) describe the

TABLE 2. Description of model configurations used in this study unless otherwise noted.

Configuration	GEM	WRF
	General Parameters	
Grid Spacing	0.135°	15 km
Time Step	450 s	60 s
	Dynamical Core	
Advection	Semi-Lagrangian with cubic Lagrange interpolation	Third-order Eulerian (Skamarock and Gassmann 2011)
Grid Geometry	Latitude-longitude Yin-Yang (Qaddouri and Lee 2011) or limited-area	Latitude-longitude limited-area (Skamarock et al. 2019)
Horizontal Staggering	Arakawa C-grid (Arakawa and Moorthi 1988)	Arakawa C-grid (Arakawa and Moorthi 1988)
Time Integration	Two time-level iterative implicit	Third-order Runge-Kutta (Wicker and Skamarock 2002)
Vertical Coordinate	Hybrid terrain-following log-hydrostatic pressure (Girard et al. 2014)	Hybrid terrain-following dry mass (Park et al. 2013)
Vertical Staggering	Thermodynamic and dynamic variables (Girard et al. 2014)	Geopotential and vertical motion (Skamarock et al. 2019)
	Physical Parameterization Suite	
Boundary Layer	1.5-order closure (Bélair et al. 1999; McTaggart-Cowan and Zadra 2015)	First-order YSU closure (Hong et al. 2006)
Deep Convection	Mass-flux based on Kain and Fritsch (1990, 1992)	Mass-flux based on Kain (2004)
Microphysics	Grid-scale condensation (Sundqvist et al. 1989)	Five-category single-moment WSM5 (Hong et al. 2004)
Radiation	Correlated-k (Li and Barker 2005)	None
Shallow Convection	Mass-flux based on Bechtold et al. (2008)	None

GEM dynamical core, while McTaggart-Cowan et al. (2019a) document the available suite of physical parameterizations. The configuration adopted for this study follows that of the GDPS unless otherwise noted (Table 2).

The WRF-ARW model version 4.2.1 (Skamarock et al. 2019) is used to provide an independent reference solution in a semi-idealized framework (section 2c4). The WRF configuration (Table 2) is shown in the "real-shear" integrations of Nolan (2011) to be capable of generating reliable simulations[1] of tropical cyclone evolution in a range of tropical environments.

c. Testing Protocols and Intercomparison Projects

Each experimental protocol and intercomparison described in this section serves a specific purpose within the hierarchy of complexity as shown schematically in Fig. 1. The protocols are associated graphically with complexity through the width of the colored background to create an

[1]The term "simulation" is used generically throughout this study. Whether specific simulations could be further sub-categorized as "forecasts" or "hindcasts" depends on the context of the relevant experimental protocol. These distinctions do not impact interpretations of the results or the conclusions and have therefore been avoided in favor of consistency.

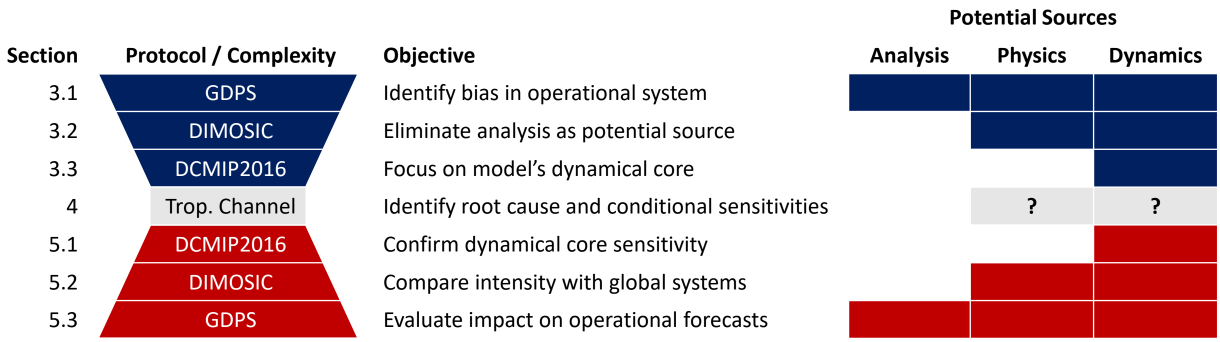


FIG. 1. Schematic of the hierarchy of modeling complexity used in this study. Blue shading indicate steps that occur prior to the correction of the weak-intensity bias, while red backgrounds represent post-correction steps. The gray background and question marks for the tropical channel ("Trop. Channel") protocol represents the execution of multiple experiments as error sources and sensitivities are assessed. The "Analysis" heading in the description of potential sources includes both initial and lower boundary conditions, while the "Physics" heading refers to the model's suite of physical parameterizations (Table 2).

hourglass shape that represents the hierarchy. Additional detail is provided by the colored panels at the right-hand side of the plot, which identifies the potential systematic error sources present at each step.

1) GDPS Forecast Sequences

The primary testing protocol for GDPS development consists of 10-day forecasts initialized from operational analyses for 2.5-month periods covering the boreal winter and summer seasons. Because of this study's focus on Northern Hemisphere tropical cyclones, the mid-June through August 2019 period is employed, with initializations at 36 h intervals for a total of 54 integrations. The model configuration follows that of the operational system, using a 0.135° Yin-Yang global grid with 84 levels that extend to 0.1 hPa. The first thermodynamic level is positioned at ~10 m above the surface, with 13 levels below 850 hPa in a standard atmosphere.

Although the operational GDPS forecast integration is coupled to ocean and sea-ice models (Smith et al. 2018), the atmosphere-only configuration used for GEM development and within the data assimilation system is employed throughout this study. Full coupling reduces mean tropical cyclone intensities by 1-2 $\mathrm{m\,s^{-1}}$, compounding the weak-intensity bias already present in atmospheric predictions. Although there is no reason to believe that the sensitivities documented in this study will be significantly altered by ocean coupling, coupled forecast sequences will be needed to confirm this assertion.

2) The DIMOSIC Intercomparison Project

Standard comparisons of operational model predictive skill are complicated by the fact that initial-state differences have significant impacts on short- and medium-range guidance. The DIfferent MOdel Same Initial Conditions (DIMOSIC) project was designed to remove this source of uncertainty (Magnusson et al. 2022).

All participants use operational ECMWF analyses (section 2a) to initialize their models at 3-day intervals over a 1-year period from 6 June 2018. The result is a set of 122 10-day simulations that diverge solely because of model differences. These data are regridded onto a common 0.5° global grid and made available to the community for further study. The GEM configuration used in the DIMOSIC project follows that of the atmosphere-only GDPS described above.

3) The DCMIP2016 Intercomparison Project

A stronger constraint on potential sources of differences across models is found in the 2016 Dynamical Core Model Intercomparison Project [DCMIP2016; Ullrich et al. (2017)], a protocol that includes simulation of a semi-idealized tropical cyclone using a highly simplified set of physical parameterizations (Reed and Jablonowski 2011, 2012). The configuration considered here employs a Kessler (1969) warm-rain scheme and a first-order turbulence closure (Reed and Jablonowski 2012).

All DCMIP2016 simulations use a 25 km variant of the GDPS configuration that is more consistent with the original protocol specifications (~0.5° grid) than the operational 15 km grid spacing. This permits direct comparison with Reed and Jablonowski (2012) and Willson et al. (2023), while avoiding the structural sensitivities noted in higher resolution runs initialized with the broad gyre-like circulation defined by the protocol. Use of an updated model version and the GDPS-like configuration relevant to this work means that the DCMIP2016 results shown here differ from the original project contribution.

4) The Tropical Channel Framework

A second semi-idealized framework is used to assess model sensitivities in an f-plane tropical aqua-channel configuration ($f = 5 \times 10^{-5}$ s^{-1}; ~20°N). Initial conditions are based on the Jordan (1958) thermodynamic profile over 28°C waters at 20°N. Shear is weak, with 5 m s^{-1} easterly winds between the surface and 850 hPa relaxing via a cosine function to 0 m s^{-1} at 200 hPa. Meridional temperature and pressure gradients are adjusted to thermal wind balance using the iterative procedure described in the appendix of Nolan (2011). This scheme also supports the insertion of a balanced tropical cyclone-like protovortex with maximum winds of 15 m s^{-1} at 1500 m altitude. This weak initial circulation is expected to strengthen given the 75 m s^{-1} (900 hPa) potential intensity of the prescribed environment (Emanuel 1988). This semi-idealized configuration precludes the investigation of complexities associated with landfalling tropical cyclones by design. The robustness of the study's results will instead be assessed by subsequent steps back up the hierarchy of complexity (Fig. 1).

Although this protocol is a useful way to connect simplified frameworks to the GDPS configuration, it has no analytic solution and it is not part of a broader intercomparison project. This means that the quality of GEM simulations cannot readily be evaluated in either an absolute or a relative sense. To fill this interpretation gap, results from a WRF simulation are used as a reference solution. The WRF domain is zonally periodic with free-slip boundary conditions at the north and south walls. It consists of 480x320 points with 15-km grid spacing and 60 levels extending to 20 km (Nolan et al. 2013). The same grid spacing is used in the 450x290 GEM configuration; however, 84 vertical levels extend to 0.1 hPa [~65 km; McTaggart-Cowan et al. (2019b)] and the domain is nested within the prescribed environmental conditions at the lateral boundaries. To ensure consistency between the simulations, the WRF initialization fields are interpolated directly onto the GEM grid, with a constant Brunt-Väisälä frequency (0.02 s^{-1}) and no vertical shear above 20 km altitude.

The tropical channel protocol is simplified by activating only those parameterizations that represent physical processes essential for tropical cyclone intensification. In the WRF reference, this means that only the planetary boundary layer, deep convection, and microphysical schemes are active (Table 2). Tight connections between the deep convection scheme and two other forms of moist convection in GEM [shallow and low-CAPE; McTaggart-Cowan et al. (2019b)] mean that they are also retained in GEM simulations unless otherwise noted. Although radiative heating is known to impact the structure (Trabing et al. 2019) and intensity (Wu et al. 2020) of simulated storms, its effects typically remain second-order compared to those of convective heating and turbulence. More importantly, the complexity of cloud-radiation interactions (Fovell et al. 2016; Ruppert Jr. et al. 2020) introduces additional indirect sensitivities that complicate interpretations of the results. For these reasons, no radiation scheme is used in this protocol.

To ensure the robustness of conclusions drawn from the semi-idealized framework, an ensemble comprising 10 perturbed members augments the unperturbed control for all GEM simulations. Inspired by Van Sang et al. (2008), random grid point meridional wind perturbations drawn from a uniform distribution over $[-0.01\ \mathrm{m\,s^{-1}}, 0.01\ \mathrm{m\,s^{-1}}]$ are added to the lowest prognostic level. This perturbation strategy is not intended to represent typical analysis uncertainty; it simply promotes the decorrelation of convective-scale elements across the ensemble without directly affecting mean-state evolution, thereby decreasing the sensitivity of the results to stochastic processes (Trabing et al. 2019).

d. Tropical Cyclone Tracking

Two different tropical cyclone tracking algorithms are used in this study. Each is used in its respective context for comparison with previous results and to avoid conflating model and tracker sensitivities. The adopted criteria ensure that tracking results focus on well-defined tropical cyclones rather than open waves or nascent vortices.

Tropical cyclone tracking at the CMC employs a variant of the Sinclair (1997) vorticity algorithm. A Cressman (1959) filter with a radius of 300 km is first applied to sea level pressure to remove subsynoptic-scale structures, followed by identification of local minima. To be classified as a tropical cyclone, the candidate low must have a maximum in cyclonic 850 hPa relative vorticity that exceeds $5\times10^{-5}\ \mathrm{s^{-1}}$ within a radius of 150 km, a 250-850 hPa thickness maximum > 9350 m within 150 km, peak 10-m winds that exceed $11\ \mathrm{m\,s^{-1}}$ within 225 km, and 900-600 hPa thickness asymmetry < 25 m averaged over a 500 km radius (Sinclair 2004). A track is generated only if the cyclone persists for 24 h or more in the forecast. Tests with the CMC algorithm confirm that the average number of tracked cyclones present during the summer-2019 testing period (section 2c1) closely matches best track data (2.3 and 2.4, respectively) and that tracking is relatively insensitive to reasonable changes to the criteria listed above.

Tropical cyclone evaluation in the DIMOSIC project is based on the Harris et al. (2016) tracking algorithm (Chen et al. 2023). This technique also uses the smoothed sea level pressure field to identify candidate centers. The 850 hPa cyclonic relative vorticity threshold used in this algorithm is a more permissive 15×10^{-5} s^{-1}, with the additional condition of a mean 500-300 hPa temperature anomaly > 2 K within 500 km of the center applied to identify warm-core cyclones. To be tracked as a tropical cyclone, the candidate center needs to persist for at least 72 h and must maintain a warm core for at least 36 consecutive hours and 48 h in total over 10-day DIMOSIC forecasts.

e. Diagnostic and Evaluation Techniques

Calculations of azimuthal mean quantities begin with a reprojection of model fields into storm-centered cylindrical coordinates using bicubic interpolation. The cylindrical grid is defined with 11 km radial and 3° azimuthal grid spacing. This configuration yields approximately isotropic grid cells at a radius of 2° and avoids sampling-induced aliasing within ~3° of the center.

Uncertainty is assessed whenever possible using 1000-member bootstrapping with replacement to compute 95% confidence intervals for the mean values shown in plots. When the mean of one set of results lies outside the confidence interval of another, the null hypothesis of equal means can be rejected at the 95% level.

3. Prevalence of the Weak-Intensity Bias in GEM

Differences between the intensity of simulated tropical cyclones and best-track estimates are expected in relatively low-resolution global NWP models (Davis 2018). Although the GDPS employs a 0.135° (15 km) grid, its effective resolution approaches ~120 km based on free-tropospheric kinetic energy spectra [Skamarock (2004); not shown]. In addition to under-resolving relevant features, (Rogers et al. 2015), the ~225 km^2 footprint of GDPS grid cells means that modeled winds suffer from representativeness errors when compared to maximum wind estimates (Knaff et al. 2021). Despite these limitations, the results described in this section show that the GDPS suffers from more severe weak-intensity biases than do equivalent NWP systems.

a. A Weak Bias in Operational GDPS Predictions

The expectation of underprediction has meant that weak storms in the GDPS have not historically been considered a major problem. Recent changes to physical parameterizations have improved tropical cyclone predictions in general (Zadra et al. 2014; McTaggart-Cowan et al. 2019b); however, mean 72 h intensity errors remain approximately -7 $\mathrm{m\,s^{-1}}$ (6 hPa) for the limited sample (34) of tropical cyclones in the summer-2019 period (section 2c1).

Annual WGNE tropical cyclone assessments performed by the JMA have indicated that these biases are larger than those of other global modeling systems (Yamaguchi et al. 2017). An updated 2021 assessment (Fig. 2) confirms that there has been no notable improvement in GDPS biases despite model upgrades and the reduction of grid spacing from 0.35° to 0.135° over the intervening period. The model continues to suffer from a conditional intensity bias: tropical cyclones with best-track central pressures above 980 hPa are associated with a limited weak-intensity bias, while stronger storms suffer from a large intensity deficit (Fig. 2a). Other global modeling systems included in the assessment appear to be more capable of representing the full range of storm intensities (Fig. 2b-d), with the UKMO model predicting particularly strong storms (Fig. 2c).

An important caveat is that GDPS data continue to be retrieved on a 1° grid for the WGNE evaluation, while datasets for the other systems follow the native model grid more closely (annotations in Fig. 2). The impact that this inconsistency has on the results is difficult to quantify; however, this assessment suggests that the GDPS remains an outlier in terms of tropical cyclone intensity biases.

b. Constraining Analysis Uncertainty with DIMOSIC

The influence of differing initial and lower boundary conditions on simulated tropical cyclone intensity is impossible to determine based on the evaluation of operational guidance alone. However, the DIMOSIC project eliminates this uncertainty to permit a more direct evaluation of model behavior (section 2c2). Chen et al. (2023) show that the GDPS-configured GEM model (labeled as “CMC” in their Fig. 5) lies on the weak-cyclone end of the predicted intensity distribution, with global mean biases of approximately -15 $\mathrm{m\,s^{-1}}$ (15 hPa).

Using a reference model with 13 km grid spacing, Chen et al. (2023) show that the impact of aggregation onto the 0.5° DIMOSIC exchange grid is roughly -4 $\mathrm{m\,s^{-1}}$ (5 hPa). Although this

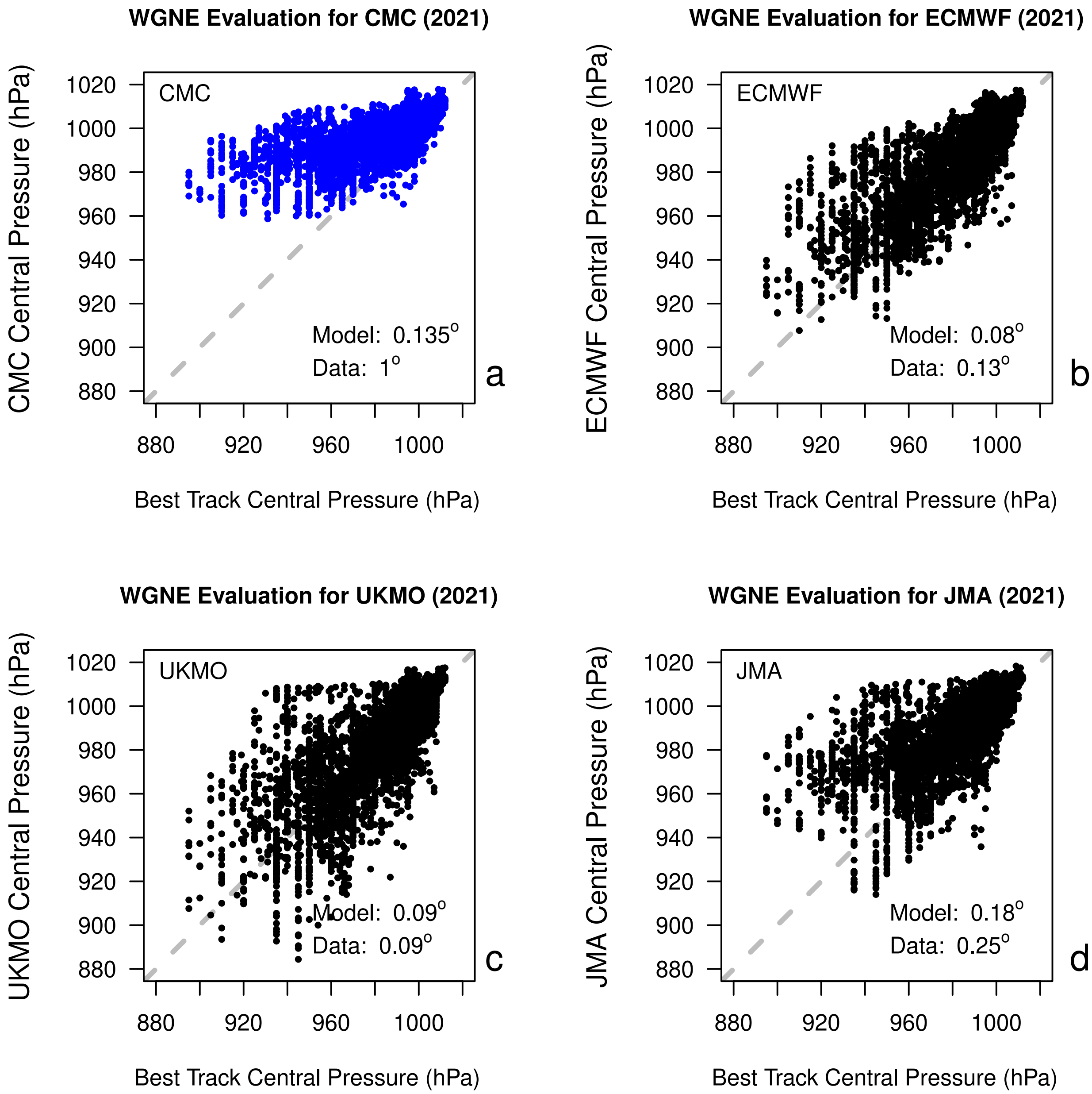


FIG. 2. Scatter plot of 72 h model-predicted (ordinate) versus best-track estimated (abscissa) central pressures of tropical cyclones across the global domain in 2021, assessed as described by Yamaguchi et al. (2017). Results are shown for the operational global guidance generated by the CMC (the GDPS; a), ECMWF (b), UKMO (c) and JMA (d). The diagonal is indicated with a dashed gray line on each panel for reference. The native grid spacing for each model is labeled as "Model", while the spacing of the latitude-longitude grid used to retrieve forecasts is labeled as "Data" on each panel.

operation explains much of the difference in bias estimates between DIMOSIC and the GDPS evaluation described above, the underlying systematic error remains evident. Its reproduction under DIMOSIC constraints and in the presence of significant changes in analyzed tropical moisture (Magnusson et al. 2022) suggest that the GEM model itself is a leading source of the GDPS bias: contributions from atmospheric and SST analyses appear to be limited.

c. Focusing on the Dynamical Core with DCMIP2016

Despite the constrains applied in the DIMOSIC project, the complexity of full-model intercomparison makes it difficult to identify candidate sources of the weak-intensity bias within GEM. The DCMIP2016 tropical cyclone test represents a step down in the hierarchy of model complexity that eliminates initial condition, lower boundary and model physics differences simultaneously (section 2c3).

Despite developing in an environment that is highly favorable to tropical cyclone intensification, wind speeds in the 25-km GEM-simulated storm reach only 20 $\mathrm{m\,s^{-1}}$ (970 hPa; Fig. 3a and b). These results resemble those of the T340 spectral semi-Lagrangian dynamical core employed by Reed and Jablonowski (2012), standing in stark contrast to the intense storms depicted by other formulations (their Fig. 6). Differences are not restricted to the lower-level structure of the storm (Fig. 3c): the GEM weak-intensity bias extends throughout the troposphere.

These results suggest that the GEM dynamical core contributes to the weak-intensity bias. However, the relatively coarse resolution prescribed by the protocol complicates quantitative interpretation of DCMIP2016 sensitivities in the GDPS context.

4. Root Cause Analysis using the Tropical Channel Framework

The persistence of a weak-intensity bias in increasingly simplified contexts motivates another step down the hierarchy of complexity. This will allow us to identify the root cause of the error in a GDPS-like configuration within a framework that is sufficiently constrained to limit the potential for error compensation as possible solutions are explored. The relevant characteristics of the simulations described in this section are summarized in Table 3 for reference.

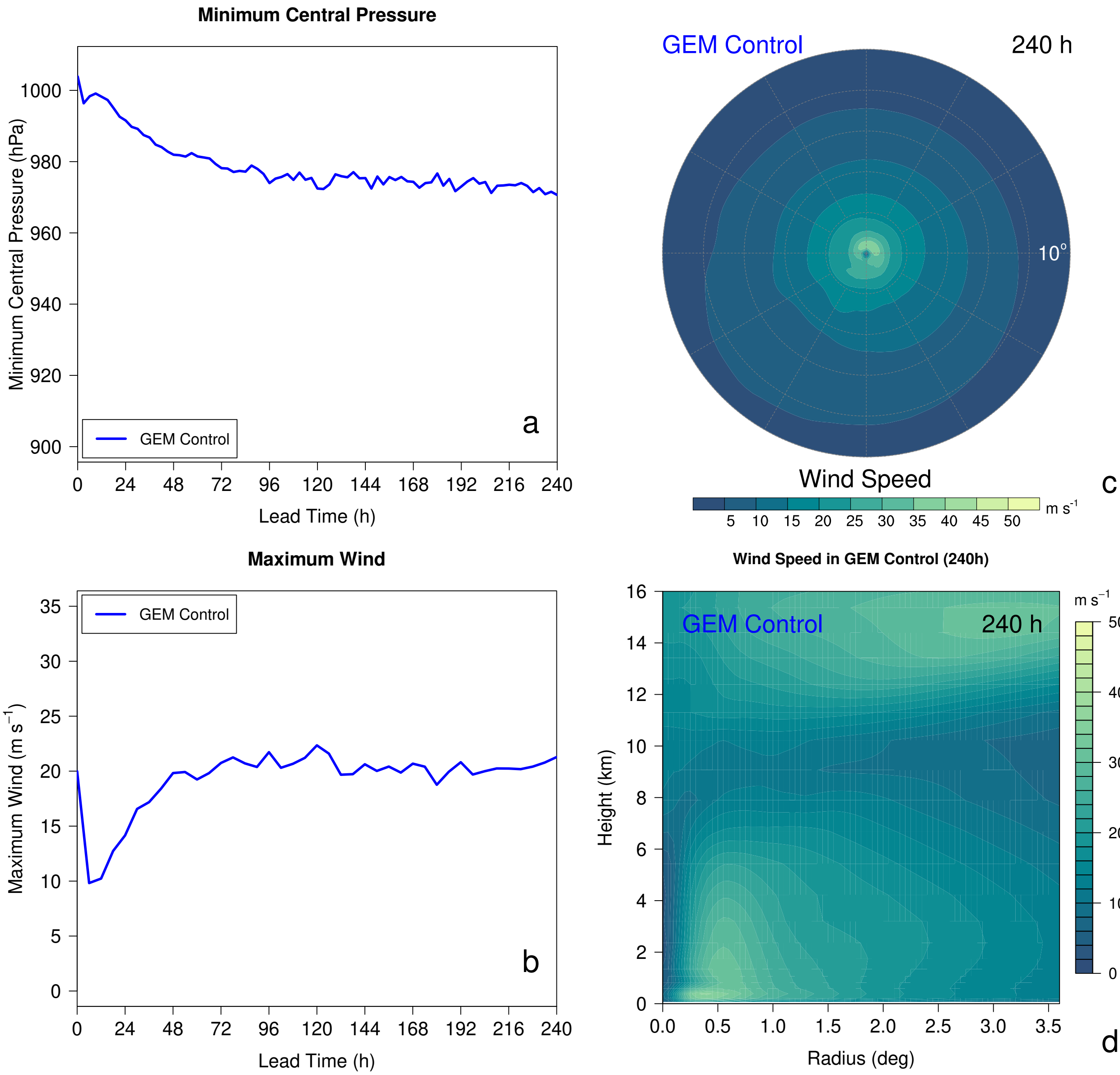


FIG. 3. Time series of tropical cyclone minimum central pressure (a; in hPa) and maximum first-level wind speed (b; in $\mathrm{m\,s^{-1}}$) in the DCMIP2016 tropical cyclone test case using the project-specified simplified physical parameterization package with 25-km grid spacing. The storm-centered tropical cyclone wind field at the second model level (c; approximately 200 m above the surface) and radius-height section of azimuthally averaged wind speed (d) are shown after 240 h of integration in $\mathrm{m\,s^{-1}}$ as indicated on the color bars. Although an updated color palette is used here for accessibility, readers interested in making a direct comparison to Fig. 5 of Reed and Jablonowski (2012) may refer to section S2 of Supplemental Material.

TABLE 3. Reference for simulations using the tropical channel framework discussed in section 4. Additional details about specific configurations and terminology are provided in the text.

Name	Model	Physical Parameterizations	Off-Centering	Type	Sections	Plotting Color
GEM control	GEM	GDPS physics	0.6	Ensemble	4a, 4b, 4c	Blue
OFFB5	GEM	GDPS physics	0.5	Ensemble	4c	Magenta
OFFB51	GEM	GDPS physics	0.51	Ensemble	4c, 4d	Red
PHYWRF	GEM	Unified WRF-type physics	0.51	Ensemble	4d	Green
PHYWRFo	GEM	Unified WRF-type physics	0.6	Ensemble	4b	Orange
WRF reference	WRF	Nolan (2011) WRF physics	-	Deterministic	4a, 4b, 4c, 4d	Black

a. Reference and Control Integrations

Both the WRF reference and GEM control simulations predict the development of the initial vortex into a tropical cyclone over the eight days of integration (Fig. 4). However, the storm characteristics are dramatically different in the two models.

The strength of the circulation in the WRF reference simulation remains steady over the first 48 h of integration (Fig. 4), at which point a convective outbreak initiates rapid intensification (Kaplan et al. 2010). The wind field in the WRF reference simulation contracts throughout this phase in response to sustained latent heating and precipitation within the radius of maximum wind [Fig. 5a; Stern et al. (2015); Smith and Montgomery (2016); Rogers (2021)]. The tropical cyclone's structure becomes very compact (Fig. 5b), consistent with the neglect of radiative transfer (Fovell et al. 2016). There is little evidence of outer rainbands (Fig. 5a) because subsidence in the secondary circulation effectively suppresses convection beyond the eyewall (Fig. 5c). The storm remains in a quasi-steady mature state for nearly 24 h (from 120 h to 144 h; Fig. 4), with a central pressure near 920 hPa and wind speeds nearing 60 $\mathrm{m\,s^{-1}}$. The inner core expands progressively thereafter (Fig. 5a), leading to weakening over the final 48 h of integration (Fig. 4).

The GEM-simulated tropical cyclone intensifies slowly over the first 48 h of the simulation, temporarily achieving a lower central pressure than the WRF reference (Fig. 4). Although development accelerates after this time, the deepening rate never meets the rapid intensification threshold (Kaplan et al. 2010). The circulation in the GEM control integration remains much broader and more diffuse than the WRF reference, even as it nears peak intensity (Fig. 6b). Active outer rainbands (Figs. 6a and b) limit tropical cyclone strength (Wang 2009) despite environmental subsaturation (Cornforth and Hoskins 2009), resulting in a poorly developed secondary circulation

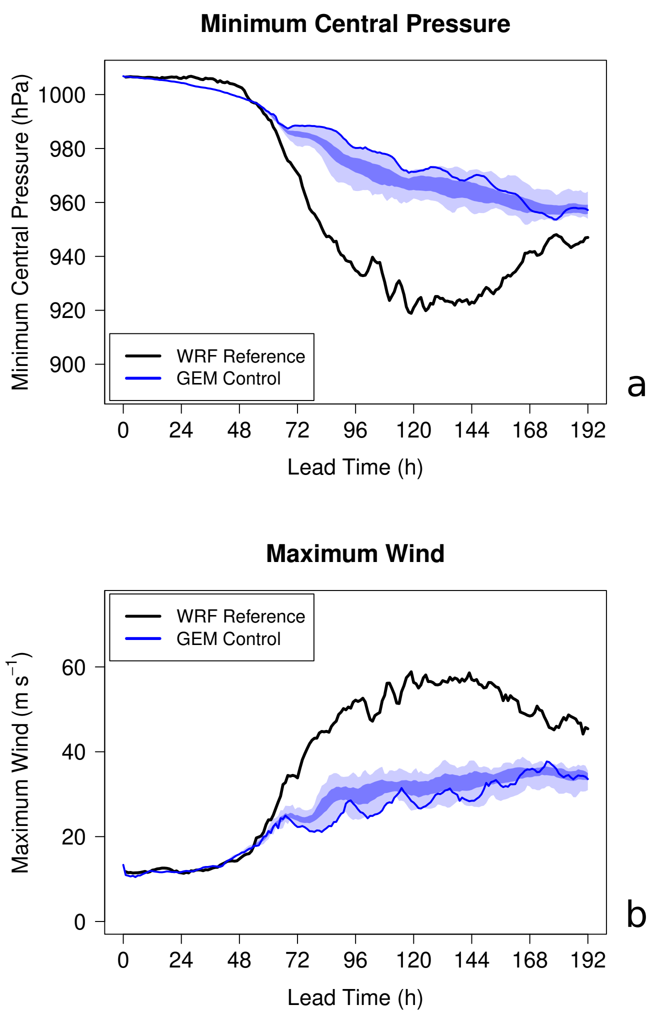


FIG. 4. Time series of tropical cyclone minimum central pressure (a; in hPa) and maximum 10-m wind speed (b; in m s^{-1}) under the tropical channel framework. The WRF reference simulation results are shown in a black solid line, while the results for the GEM control ensemble are shown in blue. Light shading covers the range of values spanned by the ensemble, while dark shading indicates the 95% confidence interval for the ensemble mean. The results for the unperturbed control member are shown with a thin solid line for reference.

(Fig. 6c). These features promote secondary eyewall formation (Wang and Tan 2020; Rozoff et al. 2012) despite the fact convective rings are not typically observed in such weak storms (Willoughby et al. 1982). The associated eyewall replacement cycles (Sitkowski et al. 2011) are responsible for periodic intensity fluctuations in the GEM control (Fig. 4b).

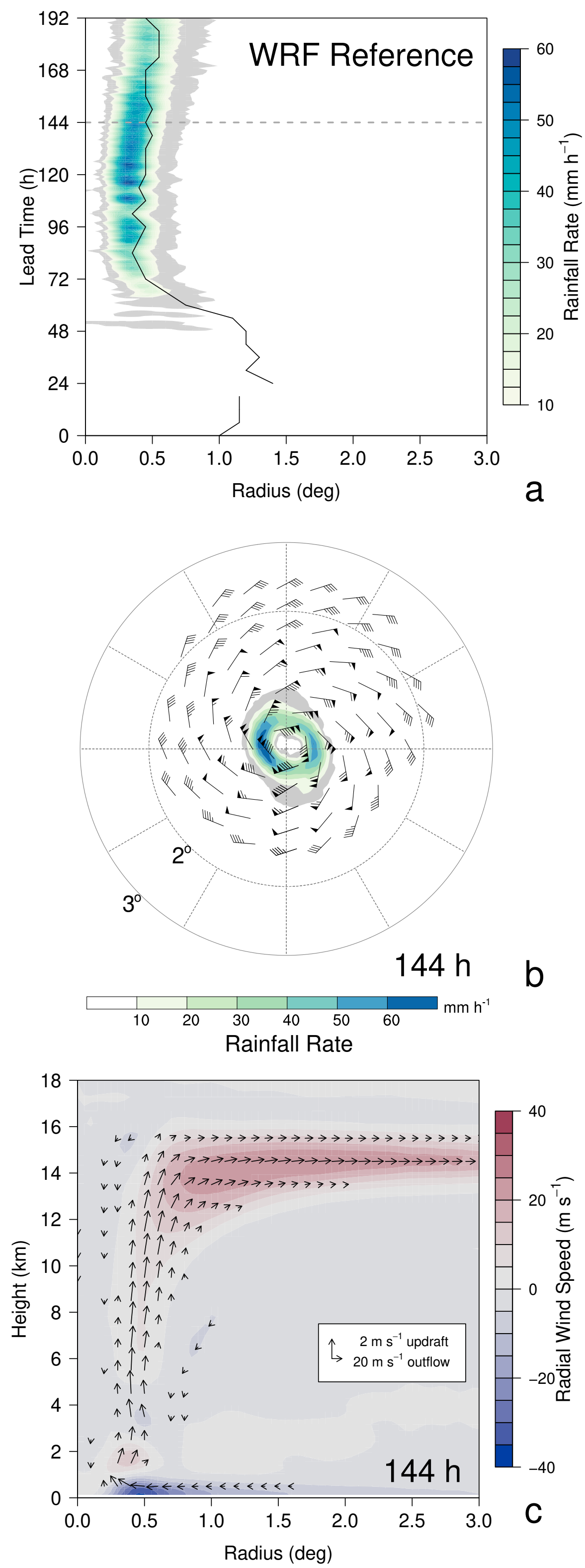


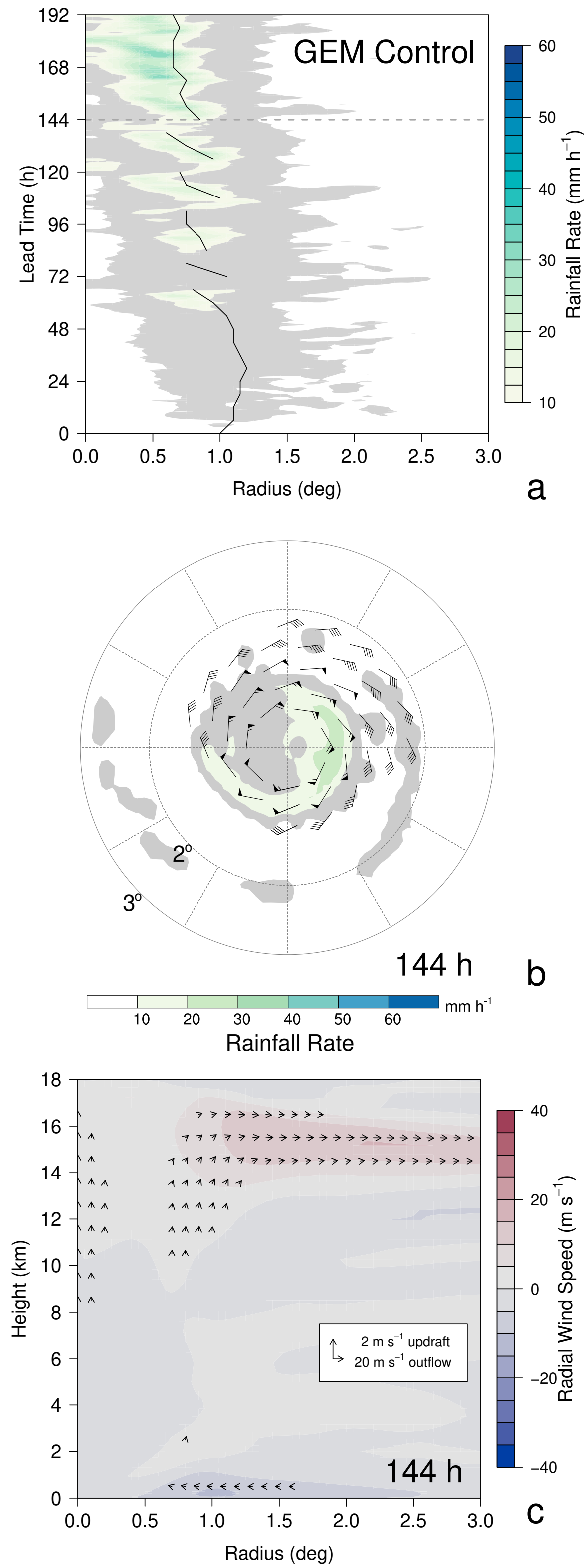
FIG. 5. See caption on next page.

FIG. 5. Summary of the tropical cyclone in the WRF reference simulation. Panel (a) shows the time evolution of the azimuthally averaged rainfall rate (color-shaded in $\mathrm{mm\,h^{-1}}$ as shown on the color bar), with light gray shading for rain rates >1 $\mathrm{mm\,h^{-1}}$. The radius of maximum wind at 2 km (Rogers 2021) is plotted with a solid black line, discontinuous to indicate the development of secondary wind maxima. A dashed gray line indicates the 144 h lead time. Panel (b) shows the 144 h precipitation rate [plotted as in (a)] and 10 m winds with short, long and pennant barbs indicating 2.5 $\mathrm{m\,s^{-1}}$, 5 $\mathrm{m\,s^{-1}}$ and 25 $\mathrm{m\,s^{-1}}$ winds, respectively. Barbs are only plotted for values >17.5 $\mathrm{m\,s^{-1}}$, indicative of tropical storm-force winds. The 2° and 3° storm-centered range rings are plotted using dashed lines in (b) for reference. Panel (c) shows the radius-height section of the 144 h azimuthally averaged secondary circulation (vectors as shown in the reference inset, with small magnitudes masked), and radial wind speeds color-shaded in $\mathrm{m\,s^{-1}}$ as shown on the color bar.

Differences in intensification rate between the WRF reference and GEM control may indicate that the models favor different forms of deepening (Holliday and Thompson 1979; Ryglicki et al. 2018; Judt et al. 2023); however, there is no independent way to evaluate the relative accuracy of the depictions. What is more certain is that the weak-shear environment is ideal for the development of a vortex whose strength approaches its potential intensity (900 hPa and 75 $\mathrm{m\,s^{-1}}$; section 2c4). Even qualitatively accounting for the under-resolution of the tropical cyclone core in these model configurations, it is clear that the WRF reference better represents expected storm strength than the GEM control. This reproduction of the weak-intensity bias makes the tropical channel framework an ideal testbed for identifying the leading factors that contribute to this systematic error.

b. Sensitivity to Physical Parameterizations

A logical place to begin the search for specific factors contributing to a tropical cyclone weak-intensity bias is the model's suite of physical parameterizations. Underestimation of surface enthalpy fluxes or deficits in condensation heating would directly contribute to insufficient vortex strength by depriving the system of its primary energy source. Minimizing parameterization differences between the GEM control and WRF reference configurations is an efficient way to determine the potential impact of physical process representation on GEM's weak-intensity bias. Although each parameterization change was tested individually, for brevity only their combined effects on the simulation are discussed.



364 FIG. 6. Summary of the tropical cyclone in the unperturbed member of the GEM control ensemble. Plotting
365 follows the conventions adopted for Fig. 5.

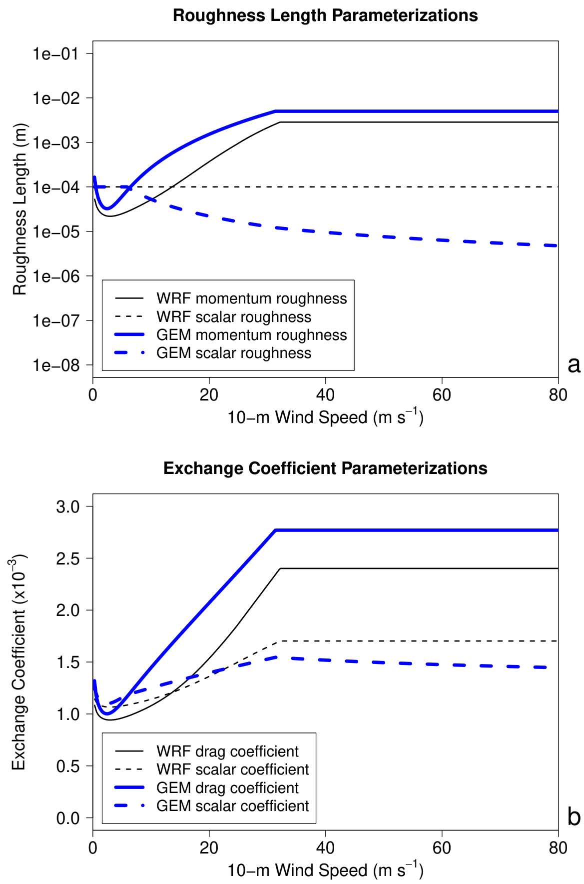


FIG. 7. Dependence of roughness lengths for momentum (solid) and scalars (dashed) on 10-m wind speed in the WRF reference (thin black; "isftcflux=1") and the GEM control (thick blue). A logarithmic ordinate is used in (a) because of the large range of roughness values. The equivalent relationship between winds and momentum (drag) and scalar exchange coefficients is shown with a linear ordinate in (b).

Surface exchange coefficients in the WRF reference are computed with the "isftcflux=1" configuration [Eq. 10 of Green and Zhang (2013)]. This formulation limits the momentum roughness length at high wind speeds (Powell et al. 2003) and holds the scalar roughness length constant (Fig. 7a). Replacing GEM estimates with these values is expected to increase storm intensity by enhancing moist enthalpy fluxes as the circulation accelerates in a reduced-drag environment (Fig. 7b).

The turbulent fluxes serve as the lower boundary condition for the boundary layer parameterization, which represents vertical eddy transports. The TKE-based closure used in GEM (Bélair et al. 1999; McTaggart-Cowan and Zadra 2015) differs significantly from the parameterized K-profile closure of WRF's YSU scheme (Hong et al. 2006). Unification was therefore only achievable through the implementation of the latter in the GEM physics suite. With this addition, the two models have similar representations of unresolved turbulence and boundary layer depth.

Although deep moist convection is parameterized using variants of Kain and Fritsch (1993) in both models, important differences have evolved over time. The GEM implementation has thus been modified to resemble its WRF counterpart more closely. Convective momentum transport has been removed and the convective velocity scale-based trigger function (McTaggart-Cowan et al. 2019b) has been replaced with the Kain (2004) LCL-based trigger. Although these modifications are known to produce inferior guidance in general, they harmonize key components of the parameterization. Similarly, the shallow convection and low-CAPE schemes used in GEM are deactivated to unify the model configurations.

The WSM5 microphysics scheme employed in the WRF reference is more advanced than the Sundqvist et al. (1989)-based condensation scheme used in the GEM control. However, GEM tests using the Predicted Particle Properties (P3) scheme (Morrison and Milbrandt 2015) reveal little sensitivity in this case, consistent with equivalent WRF integrations that use alternative microphysical options (not shown) and full-complexity simulations of strong storms in the tropics (Park et al. 2020). In light of these results and the lack of radiative feedback in this protocol, no change was made to GEM's representation of gridscale clouds and precipitation.

The GEM configuration resulting from this unification of surface flux, turbulence and moist convective processes is identified as PHYWRFo (the reason for the appended "o" will become apparent in section 4c). The model appears to be unphysically insensitive to these fundamental changes to key parameterizations (Figs. 8a and b). The weak-intensity bias persists despite increased rainfall within the radius of maximum wind (Fig. 8c). Although inward-propagating bands no longer perturb the circulation, the simulated storm is unable to sustain a cloud-free eye. This allows us to conclude that the weak-intensity bias likely lies outside GEM's suite of physical parameterizations.

c. Sensitivity to Dynamical Core Configurations

The search for potential error sources in the dynamical core is guided by preliminary DCMIP2016 results (section 3c). Reed and Jablonowski (2012) hypothesize that the weak-intensity bias that they observe in their spectral semi-Lagrangian dynamical core is related to excessive numerical

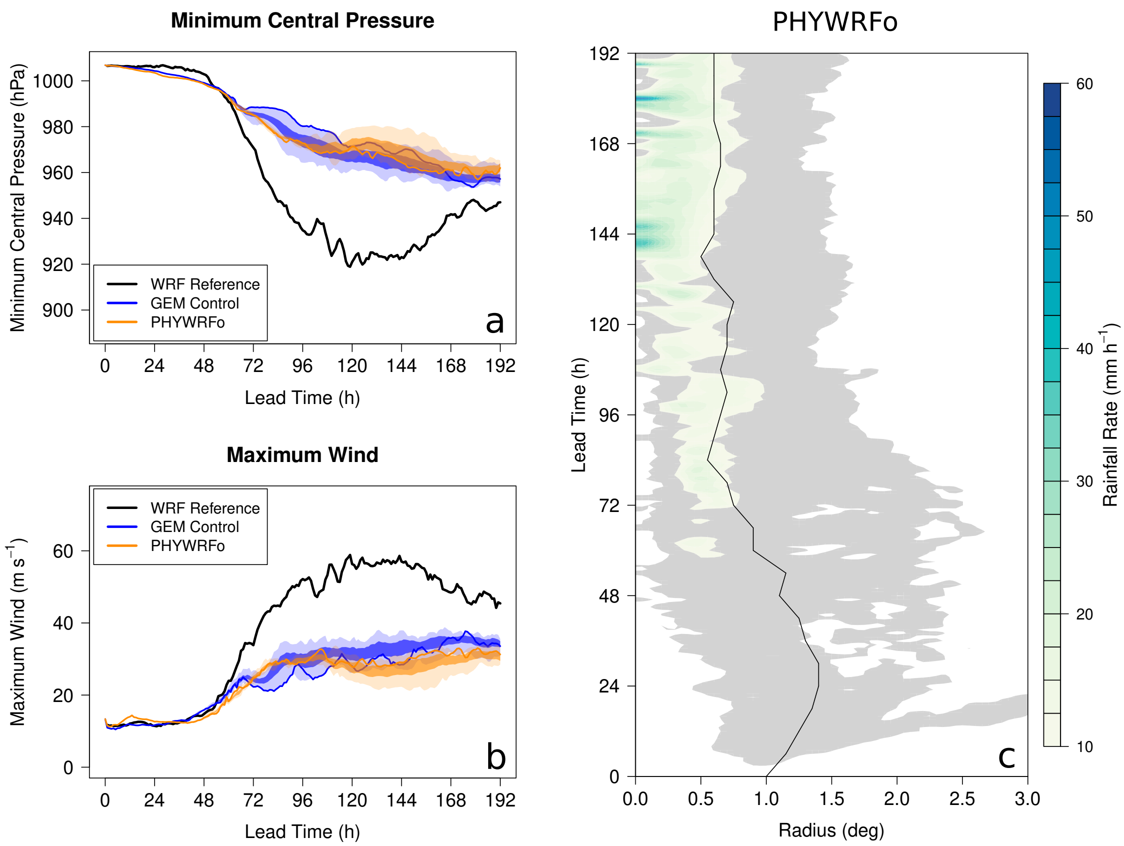


FIG. 8. Summary of the tropical cyclone in the PHYWRFo simulations. Central pressure (a) and maximum near-surface wind (b) time series are plotted as in Fig. 4. The evolution of azimuthal-mean rainfall rate and radius of maximum wind (c) follows the conventions adopted for Fig. 5a.

dissipation. Despite significant formulation differences, this conjecture is valuable guidance for the root cause analysis in GEM.

1) Use of Off-Centering in GEM

The GEM dynamical core employs iteratively implicit time discretization in conjunction with semi-Lagrangian advection (Girard et al. 2014). The model equations are represented in the form,

$$\frac{dF_i}{dt} + G_i = 0 \quad , \tag{1}$$

where F_i is a prognostic variable with dynamical forcings G_i for the i^{th} equation. Adopting a two time-level scheme, Eq. 1 is discretized using the trapezoidal rule as,

$$\frac{F_i^A - F_i^D}{\delta t} + bG_i^A + (1-b)\,G_i^D = 0 \quad , \tag{2}$$

for time step δt, where superscript "A" refers to the trajectory arrival location at time t, while superscript "D" refers to the computed trajectory departure point at time $t - \delta t$. Most important for the current discussion is b, an off-centering parameter introduced to control the resonant growth of spurious structures generated by sharp gradients in flows whose Courant number approaches or exceeds unity (Rivest et al. 1994). This parameter is also known as the "decentering" or "time weighting" parameter and is related to the $\epsilon = 2b - 1$ used by Jablonowski and Williamson (2011).

2) The Impact of Numerical Damping

A value of $b = 0.5$ implies no off-centering, such that time integration scheme reduces to the Crank-Nicholson method. As b is increased, the damping effects of the technique intensify and the second-order accuracy of the scheme drops to first-order (Jablonowski and Williamson 2011). Although b could in principle contain spatiotemporal variability and be independent for each equation, a single value of $b = 0.6$ is currently used in all GEM configurations.

The results of the OFFB5 experiment, identical to the GEM control but with $b = 0.5$, suggest that Reed and Jablonowski (2012) were correct to posit that off-centering could limit simulated tropical cyclone intensity (Fig. 9). The storm undergoes rapid intensification between 48 h and 96 h, with an intensification rate approaching that of the WRF reference. A quasi-equilibrium is established for the subsequent 48 h, with a central pressure of ~935 hPa and maximum winds approaching 50 $\mathrm{m\,s^{-1}}$ (Figs. 9a and b). A second phase of intensification takes place thereafter as eyewall precipitation intensifies near the radius of maximum wind (Fig. 9c). Central pressures in some members fall below 890 hPa and maximum winds exceed 60 $\mathrm{m\,s^{-1}}$, double the peak values seen in the GEM control.

The OFFB5 simulation shows that GEM is highly sensitive to off-centering in this semi-idealized experiment, an indication that this may be an important contributor to the model's weak-intensity bias in more complete configurations. However, $b = 0.5$ is not an admissible value for GEM simulations that include orography (Subich 2022) and leads to numerical instability even in this

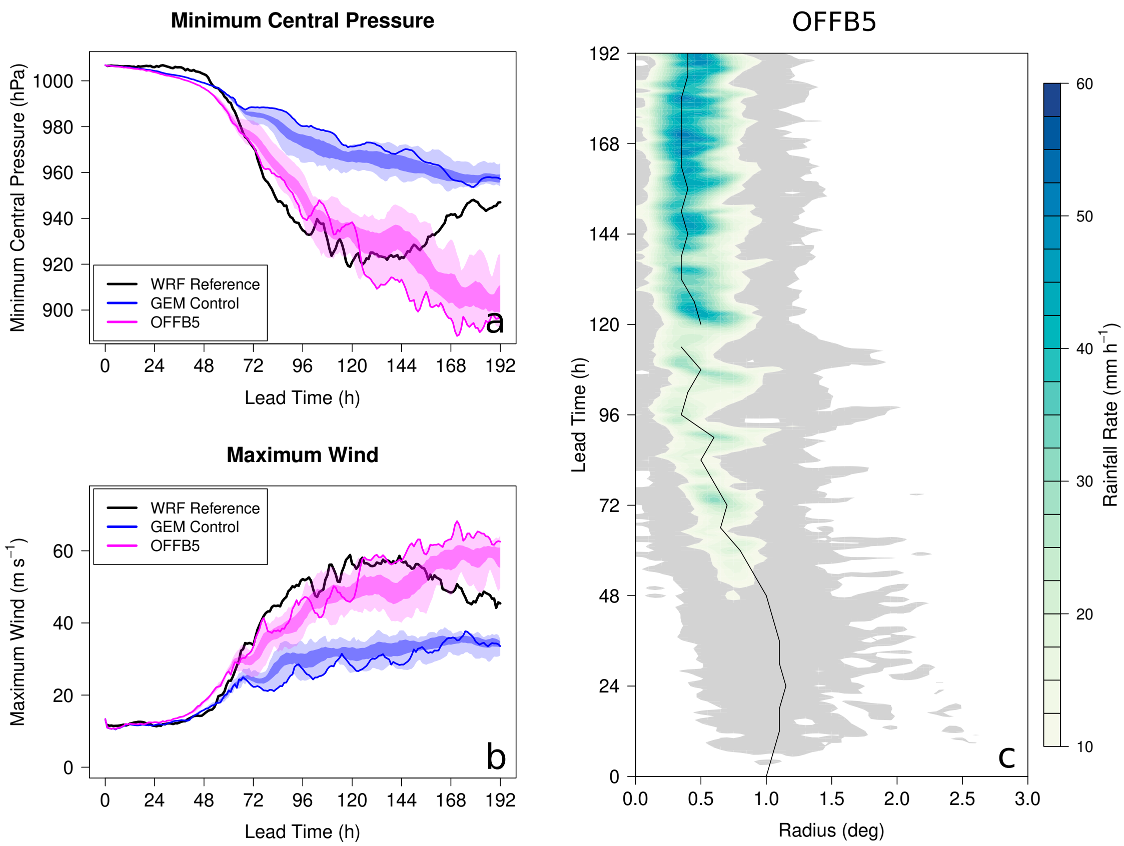


FIG. 9. Summary of the tropical cyclone in the OFFB5 simulations, plotted as in Fig. 8.

simplified framework when run in non-hydrostatic mode (not shown). An additional experimental setup is needed to pinpoint the source of the sensitivity and to establish a value for $b > 0.5$ without introducing compensating errors. For example, insufficient suppression of unstable modes in the dynamical core could be misdiagnosed as improved storm intensity if they are controlled by excessively diffusive physical parameterizations.

3) Error Description and Mitigation using Vortex Spin-Down

A spin-down experiment is designed to evaluate the inherent numerical dissipation of GEM dynamics through comparison to an equivalent WRF simulation in the context of a strong tropical cyclone. The WRF reference simulation is modified to turn off all physical parameterizations after 144 h of integration, when the storm is in its mature phase. This state is also used to initialize dynamics-only GEM simulations. The models' atmospheres become adiabatic and inviscid, de-

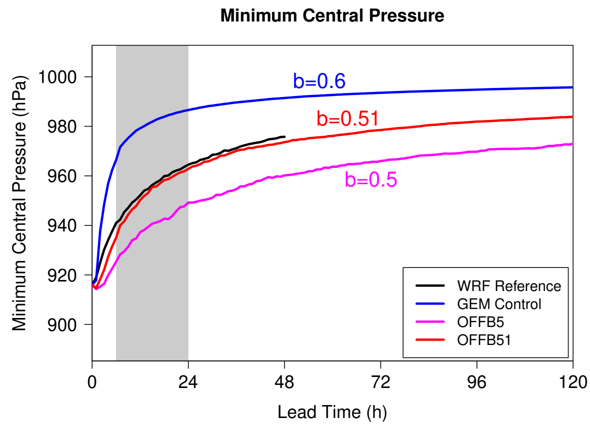


FIG. 10. Time series of tropical cyclone central pressure evolution in the spin-down experiments. Although all simulations are run without physical parameterizations, dynamical core configuration names match those used throughout this section (Table 3). The WRF reference simulation is shown for the 48 h of dynamics-only integration that follows the 144-h spin-down initialization (a total 196 h total run length as used throughout this study), while GEM simulations are extended by an additional 72 h to confirm sensitivities. Gray shading indicates the time period used for piggyback averaging.

priving the storm of the boundary layer convergence and eyewall heating required to maintain its secondary circulation. The vortex undergoes an equivalent barotropic form of spin-down through internal dynamics (e.g., radiation of waves during balance adjustments) and the inherent dissipation of the dynamical cores themselves.

The circulation decays quickly in WRF, with the central pressure of the storm rising from 915 hPa to 975 hPa in just 48 h (Fig. 10). Weakening rates in GEM depend strongly on the value of b, with the control integration ($b = 0.6$) virtually eliminating the vortex in just 24 h. The circulation persists for much longer in the OFFB5 ($b = 0.5$) configuration; however, increased temporal variability is indicative of potential noise problems when off-centering is completely eliminated. The OFFB51 experiment ($b = 0.51$) yields vortex evolution that closely resembles that of the WRF spin-down integration.

The first step in diagnosing the source of this sensitivity involves the inviscid tangential wind budget (Hendricks et al. 2004),

$$\overbrace{\frac{\partial \overline{v}}{\partial t}}^{\text{Model Tendency}} = \overbrace{-\overline{u}\left(\overline{\zeta}+f\right) - \overline{w}\frac{\partial \overline{v}}{\partial z} - \overline{u'\zeta'} - \overline{w'\frac{\partial v'}{\partial z}}}^{\text{Dynamic Forcing}} + \overbrace{D_T}^{\text{Residual Acceleration}} , \qquad (3)$$

where v and u are the tangential and radial wind components, w is vertical motion, ζ is relative vorticity and D_T is a residual acceleration to be discussed shortly. Overbars indicate azimuthal means and primes denote departures therefrom. While mathematically well-posed, direct application of Eq. 3 to the spin-down simulations (Fig. 10) leads to the trivial conclusion that the vortex in the GEM control integration is "weaker because it is weaker". Instead, we adopt a piggybacking approach (Grabowski 2014) in which the vortex evolution follows that of the OFFB5 simulation and the model predicts one-step changes away from this state using $b = 0.6$. Averaging these steps allows us to diagnose the direct impact of off-centering while remaining fixed to the $b = 0.5$ slow-decay solution.

The tangential wind budget for the OFFB5 simulation reveals slow vortex spin-down (Fig. 11a) despite weak inflow-driven acceleration from the dynamic forcings (Fig. 11b). Most relevant here, however, is the residual acceleration (Fig. 11c). The D_T term incorporates all changes to the primary circulation that are not captured by the inviscid momentum equation, including the effects of numerical dissipation in the dynamical core. Residual acceleration in the OFFB5 simulation does not exceed $-5\ \mathrm{m\,s^{-1}\,h^{-1}}$, consistent with the inherent damping of iteratively implicit time stepping and semi-Lagrangian advection.

The magnitude of D_T increases dramatically when off-centering is applied in the GEM control configuration (Fig. 11f). Diagnosed now as the departure from OFFB5 accelerations to be consistent with the piggybacking approach, the numerical deceleration approaches $20\ \mathrm{m\,s^{-1}\,h^{-1}}$ within the radius of maximum wind. This drag-like forcing induces radial inflow throughout the lower troposphere (Fig. 12) as numerically slowed tangential winds are deflected by the pressure gradient force to reestablish gradient balance [Fig. 11d and Smith et al. (2009)]. The implied deep-layer convergence at the vortex core leads to rapid filling through a process that is analogous to tropical

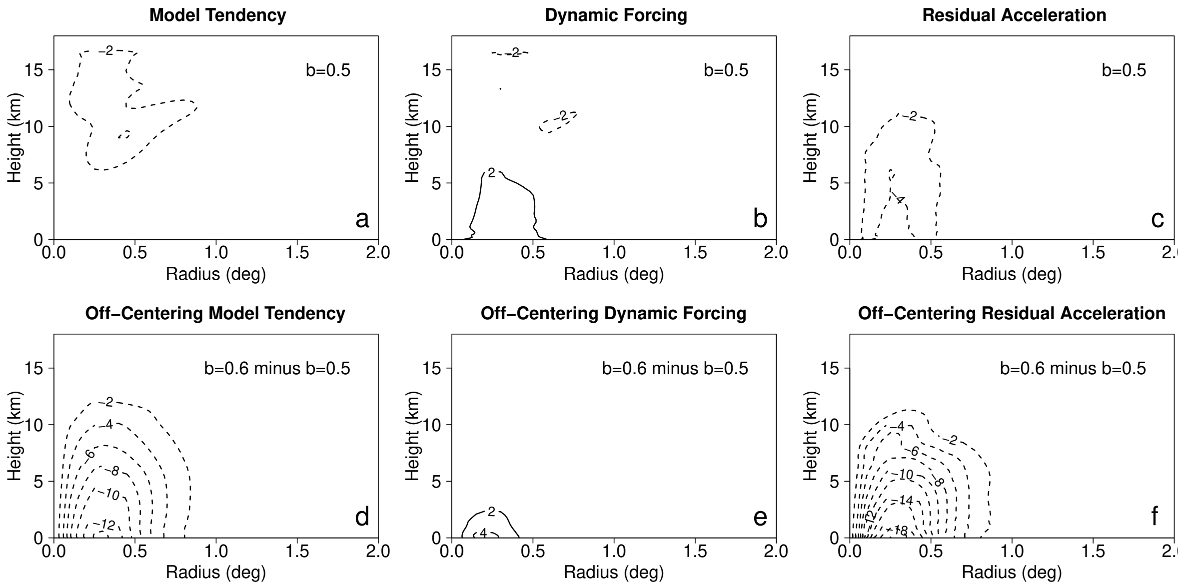


FIG. 11. Radius-height sections of tangential momentum budget terms (Eq. 4) for the model tendency (a), dynamic forcing (b) and residual acceleration (c) in the OFFB5 spin-down experiment (top row). The additional accelerations induced by $b = 0.6$ in the piggybacked GEM control simulation for the same terms are shown in the bottom row. Contours appear at 2 $\mathrm{m\,s^{-1}\,h^{-1}}$ intervals with dashed negatives and no plotting of the zero contour. All values are averaged between 6 h and 24 h integration times (gray shading in Fig. 10).

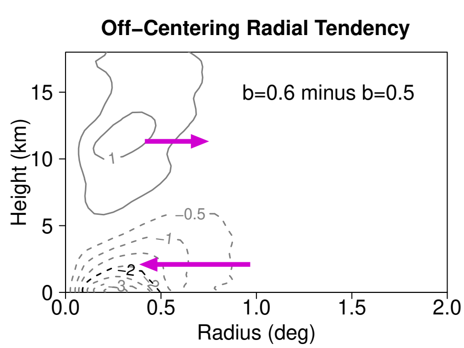


FIG. 12. As in Fig. 11d, but for radial accelerations and using a contour interval of 0.5 $\mathrm{m\,s^{-1}\,h^{-1}}$. Negative values denote inflow, with schematic wind vectors plotted in magenta for clarity.

cyclone landfall, but with friction acting throughout the circulation instead of only at the surface (Chen and Chavas 2020; Hlywiak and Nolan 2021).

The reason that off-centering decelerates the primary circulation can be understood through analysis of the pressure gradient terms on the right-hand side of GEM's discretized momentum

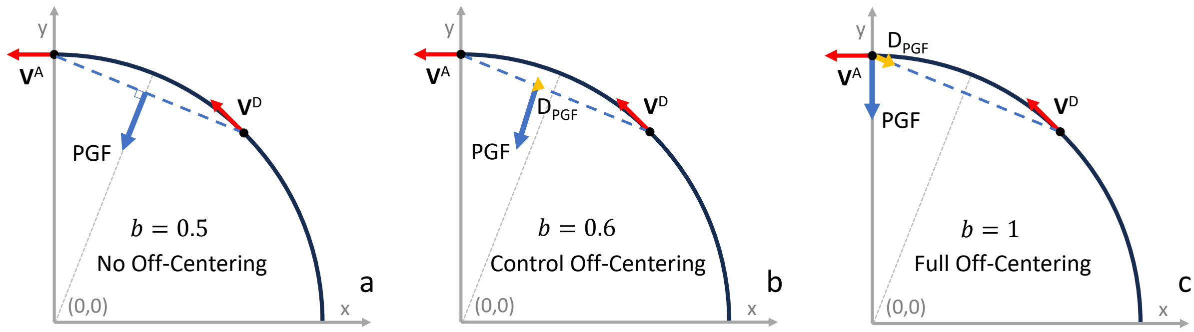


FIG. 13. Schematic of the impact of off-centering values of 0.5 (a), 0.6 (b) and 1.0 (c) for a vortex centered at the origin. Wind vectors are shown as red arrows, the computed pressure gradient force (PGF) in blue and the component of the pressure gradient force that opposes the flow along the trajectory (labeled D_{PGF}) in yellow (b and c). A linear back-trajectory connects arrival ("A") and departure ("D") points (black dots) with a dashed blue line.

equation,

$$\frac{\mathbf{V}^A - \mathbf{V}^D}{\delta t} = -bR_d\,(T_v \nabla \ln p)^A - (1-b)R_d\,(T_v \nabla \ln p)^D + bS^A + (1-b)S^D \quad , \tag{4}$$

where T_v the virtual temperature, p pressure, R_d is the gas constant for dry air and the S terms represent additional forcings. A schematic representation of this expression shows that when $b = 0.5$ (Fig. 13a) the pressure gradient force is valid at the time-centered linear trajectory midpoint and is therefore perpendicular to the tangential wind on the vortex segment as expected for gradient-balanced flow. The transported wind vector is thus rotated to follow the circular path without a change in speed.

When off-centering is introduced, the pressure gradient force is valid closer to the arrival point and time and is no longer orthogonal to the trajectory (Fig. 13b). Its orientation becomes increasingly perpendicular to the arrival wind vector for larger b (Fig. 13c), with a projected component tangent to the arc midpoint that opposes the flow along the full trajectory,

$$D_{PGF} = -\frac{\chi\left(\frac{v^2}{r} + fv\right)}{\sqrt{1+\chi^2}} \quad \text{where} \quad \chi = 2\,(b - {}^1\!/\!_2)\tan\left(\frac{v\delta t}{2r}\right) \quad , \tag{5}$$

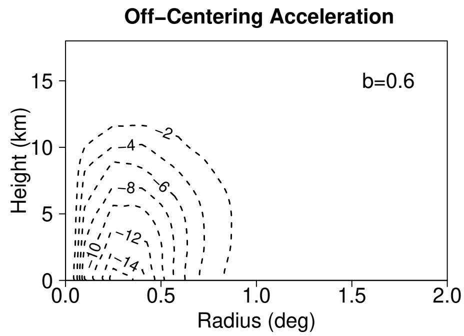


FIG. 14. Tangential wind acceleration expected for $b = 0.6$ via Eq. 5, plotted as in Fig. 11f for direct comparison.

as derived in appendix A (yellow arrows in Fig. 13). This means that the pressure gradient force actively slows the tangential wind rather than simply rotating the vector to maintain the steady-state circulation, a numerical error that disappears for $b = 0.5$ and in the small-step limit. Comparison of Figs. 11f and 14 shows that D_{PGF} explains the full structure of the residual acceleration (D_T). This misalignment of the pressure gradient force therefore drives the spin-down of the vortex in the simulation.

This "balance of forces" description of off-centering-induced spin-down does not depend on 3D vortex structure and can be similarly diagnosed in the shallow water system (appendix B). Application to a tropical cyclone-like circulation shows that three separate regimes of tangential accelerations exist, all of which suffer from numerical drag on the tangential wind that is first-order in δt. The friction-like forcing is strongest where the outer boundary of the vortex core meets the inner edge of the eyewall, exactly where maximum D_{PGF}-induced deceleration is observed (Fig. 14). Even small off-centering in the shallow water context therefore yields rapid vortex decay as in the full 3D case (c.f. Figs. 10 and B2).

4) THE IMPACT OF REDUCED OFF-CENTERING

Returning to the original semi-idealized configuration, the tropical cyclone in OFFB51 undergoes a period of rapid intensification to reach a mature-state intensity that is similar to that of the WRF reference (Fig. 15). Although the storm still possesses inwards-propagating rainbands (Fig. 16a), they are less pronounced than those noted in the GEM control (Fig. 6a). The tropical cyclone's primary eyewall contracts to a scale similar to that of the WRF reference in association with a strong secondary circulation despite reduced updraft speeds (cf. Figs. 5 and 16).

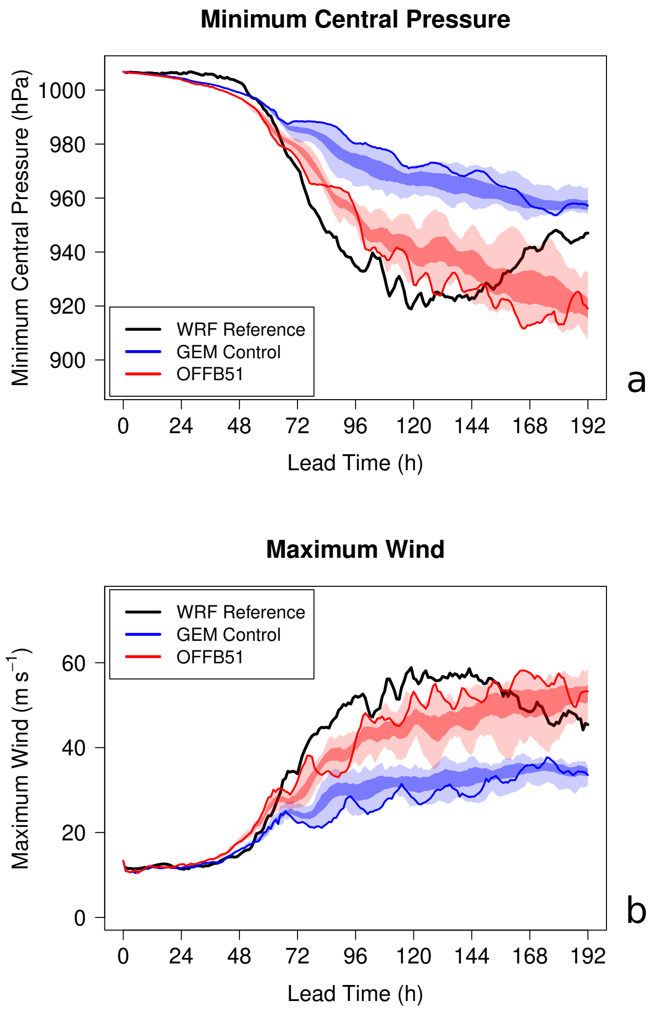


FIG. 15. Time series of tropical cyclone intensity evolution in the OFFB51 simulations, plotted as in Fig. 4.

The robustness of the simulated storm's response to off-centering is assessed through additional sensitivity tests described in sections S3 and S4 of Supplemental Material. Time step reductions (Figs. S3 and S4) lead to progressively stronger tropical cyclones in $b = 0.6$ integrations because the associated drag scales with δt (Eq. 5). The OFFB51 configuration shows much-reduced δt sensitivity until other dissipative sources in the dynamical core prevent convergence in the small-step limit (Fig. S4). Results from the tropical aqua-channel simulations also appear to be robust to changes in the prescribed thermodynamic environment, with $\pm 10\%$ changes in relative humidity having no significant impact on storm strength (section S4 of Supplemental Material). These results augur well for OFFB51-based intensity bias reductions in more complex experimental frameworks.

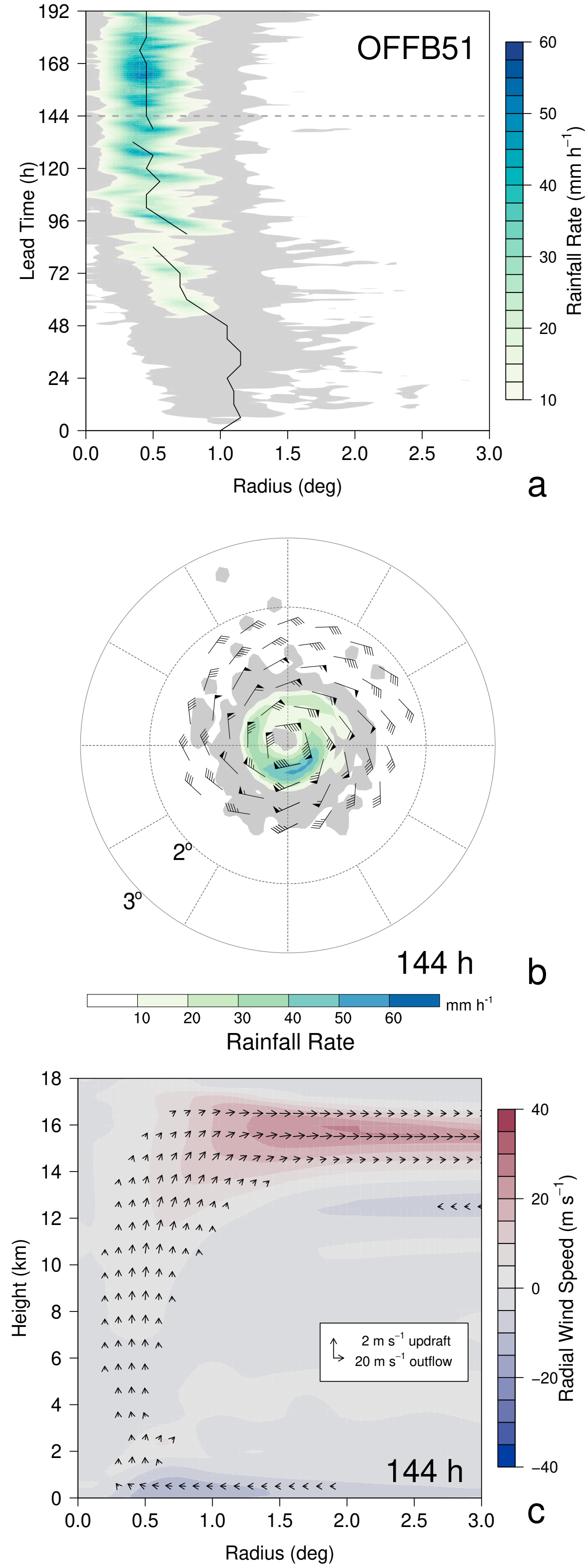


FIG. 16. Summary of the tropical cyclone in the unperturbed member of the OFFB51 ensemble. Plotting follows the conventions adopted for Fig. 5.

d. Conditional Physical Parameterization Sensitivity

The development of a strong tropical cyclone in OFFB51 presents an opportunity to revisit the sensitivities to model physics diagnosed in section 4b. The question to be answered here is whether the weak-intensity bias induced by aggressive off-centering ($b = 0.6$) dampened the response to changes in the surface flux, boundary layer and deep convective parameterizations. The PHYWRF configuration considered here is therefore identical to PHYWRFo except that $b = 0.51$ such that the final "o" (off-centered) is removed from the experiment name. The results of the PHYWRF simulation are compared to those of OFFB51 to isolate sensitivities to physical parameterizations in the reduced-dissipation context.

The simulated tropical cyclone intensity in PHYWRF slightly exceeds that of OFFB51 (Figs. 17a and b). Although this appears to imply that the results are once again unphysically insensitive to fundamental parameterization changes, the structure of the simulated storm tells a different story. The remaining inward-propagating rainbands in OFFB51 that limit intensification by repeatedly depriving the inner eyewall of moist enthalpy and momentum fluxes (Houze et al. 2007; Zhou and Wang 2011) are absent from the PHYWRF integrations. This reduces intensity fluctuations and ensemble spread (Figs. 17a and b) as the simulated eyewall maintains a strong, coherent structure throughout the storm's mature phase (Fig. 17c). This important storm-scale process distinction yields a tropical cyclone in PHYWRF whose structural evolution resembles that of the WRF reference (cf. Figs. 5a and 17c).

One aspect of the storm life cycle that remains distinct between the models is the gradual intensification over the first 48 h of all GEM integrations (e.g., Figs. 17a and b). Although potentially related to increased heating and precipitation in the near-storm environment, the source of this model-specific behavior has not been identified.

The overall similarity between PHYWRF and the WRF reference shows that expected physical responses emerge once the excessive dissipation in GEM is controlled. This highlights the importance of considering conditional sensitivities even in reduced-complexity protocols, particularly when experiments (e.g., the spin-down test described in section 4c3) can be used to constrain key components of the system.

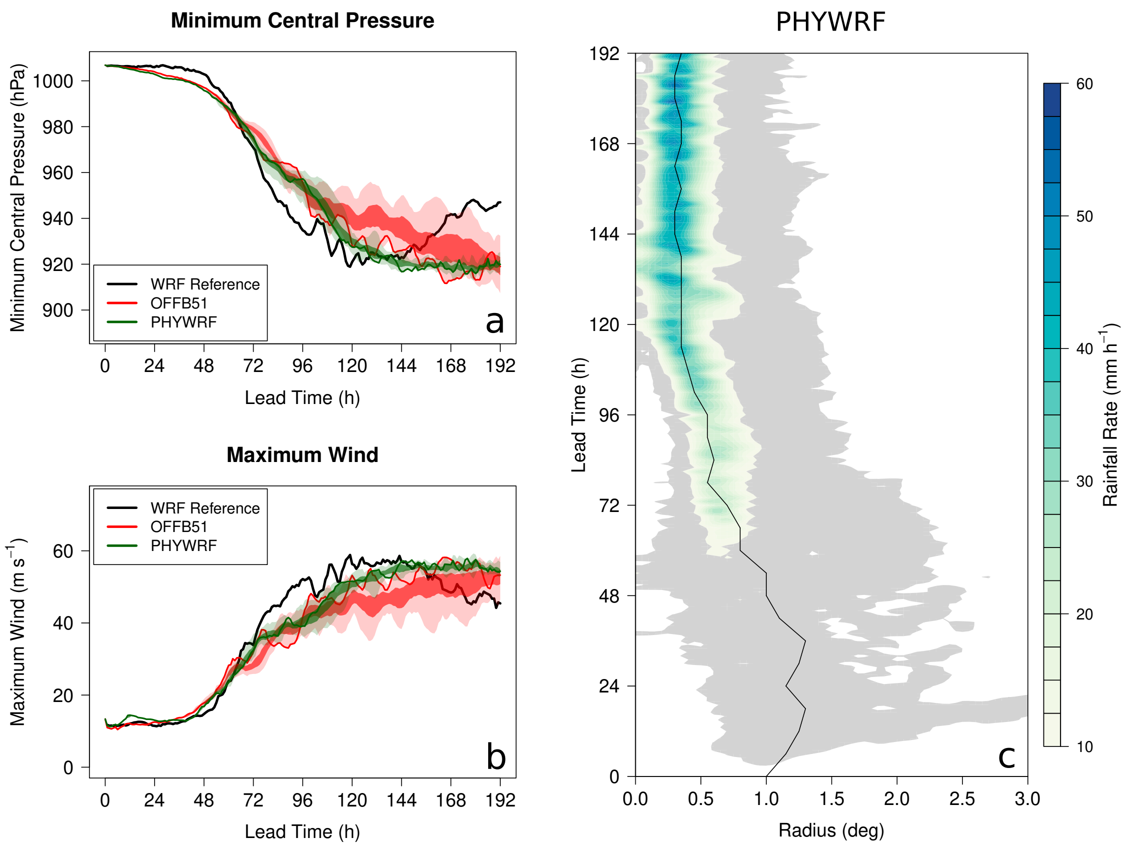


FIG. 17. Summary of the tropical cyclone in the PHYWRF simulations, plotted as in Fig. 8.

5. Assessing the Impact of Reduced Off-Centering

Simulations using the tropical channel protocol have allowed us to identify and mitigate the root cause of the weak-intensity bias. However, the next steps back up the hierarchy of complexity (Fig. 1) require the selection of a more complete GEM configuration based on one of two potential candidates: OFFB51 or PHYWRF. Both yield storms whose strengths approach the potential intensity (Fig. 17), making it impossible to dismiss either of them on theoretical grounds.

A practical consideration is that the success of subsequent steps will be evaluated against results from the current operational model. Because this configuration has been optimized for skill across a broad range of metrics, minimizing changes to it will reduce the risk of disrupting the well-balanced system (Hourdin et al. 2017; Tuppi et al. 2023). The OFFB51 configuration has therefore been selected to serve as the basis for further assessment, a choice that amounts to adopting $b = 0.51$ in the GDPS-like configurations discussed in section 3. The impact of this change in isolation can therefore be documented as complexity is reintroduced.

a. Impact on the DCMIP2016 Simulation

Reduced off-centering yields a substantial increase in tropical cyclone intensity in the DCMIP2016 simulation (Figs. 18a and b). The compact cyclonic circulation (Fig. 18c) also extends to a greater altitude, with 15 $\mathrm{m\,s^{-1}}$ winds extending throughout the depth of the troposphere (cf. Figs. 3d and 18d). These changes bring GEM results more in line with those of other participating models [e.g., Fig. 8 of Willson et al. (2023)].

The increase in tropical cyclone intensity with $b = 0.51$ is also evident in the wind-pressure relationship (Fig. 19), with OFFB51 results shifted to higher intensity along model-derived wind-pressure curves. Although there is no observational reference in the DCMIP2016 protocol, this change in gradient-balanced intensity is consistent with increased model resolution (Magnusson et al. 2019). The implied increase in GEM's effective resolution directly increases model efficiency by enhancing the accuracy of the solution without additional computational cost (Skamarock 2004).

b. Impact on DIMOSIC Intercomparison

The impact of reduced off-centering on tropical cyclone intensity in the DIMOSIC simulations shows that the sensitivity documented in more simplified contexts is robust in full GEM configurations (Fig. 20). A 2.5 $\mathrm{m\,s^{-1}}$ (5 hPa) mean intensity increase (Figs. 20a and b) yields similar reductions in root mean square errors (Figs. 20c and d) to bring GEM results into line with those of equivalent participating models.

Although this investigation focuses on tropical cyclone intensity, the changes in storm depth noted above (Fig. 18d) have the potential to affect track predictions (DeMaria et al. 2022). The year-long design of the DIMOSIC protocol provides sufficient sampling of events to reveal an improvement in track guidance through 60 h (Fig. 21a). Although relatively modest in absolute terms, errors are reduced by nearly 50% with respect to the operational ECMWF benchmark (Fig. 21b).

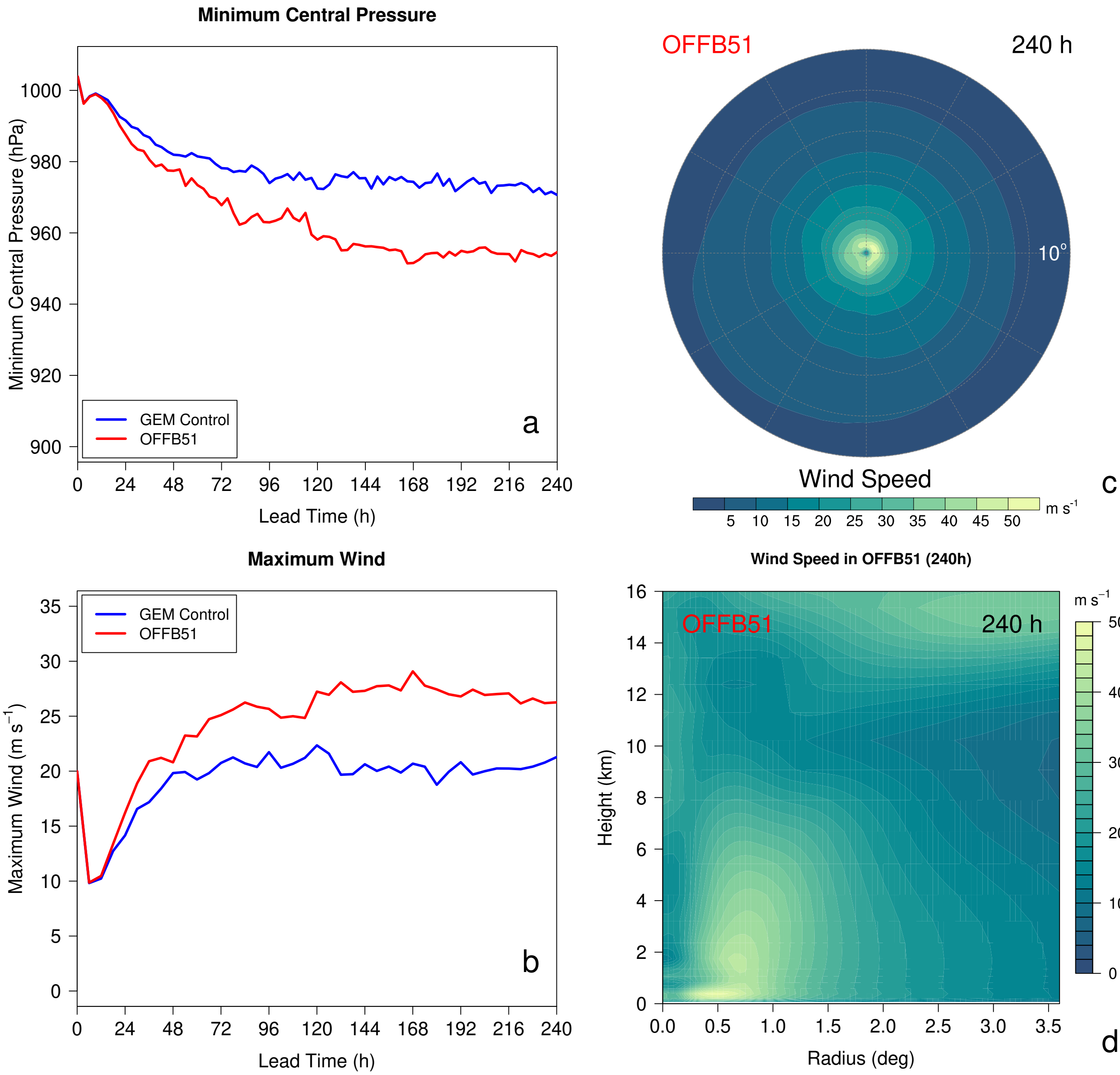


FIG. 18. Summary of DCMIP2016 results for the 25 km OFFB51 configuration, plotted as in Fig. 3.

c. Implementation in the GDPS

The promising results obtained in simplified contexts provide motivation for testing the OFFB51 configuration in a full GDPS forecast sequence (section 2c1). As an incremental step made without system rebalancing, the results discussed in this section should be considered a checkpoint in ongoing model development rather than an end point in themselves. In addition to the evaluation of tropical cyclone predictions in the operational system presented here, an analysis of the impact of adopting $b = 0.51$ on global guidance is provided in Supplemental Material (section S5). Changes

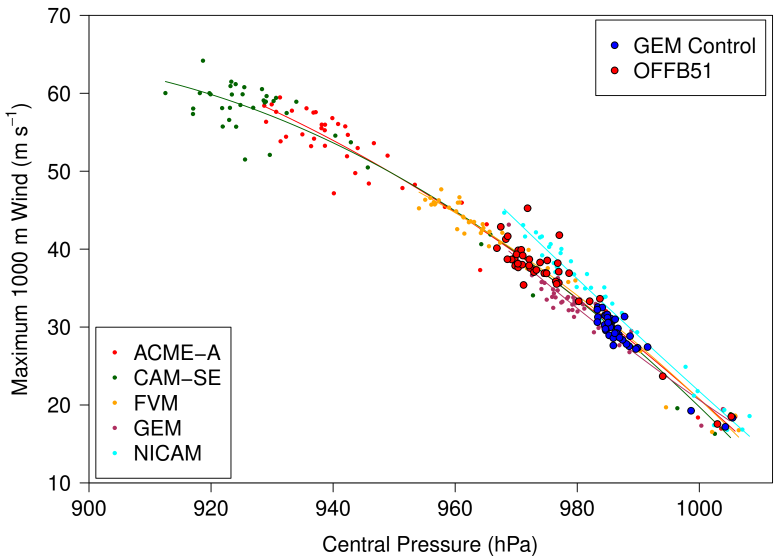


FIG. 19. Wind-pressure relationships for 25 km simulations using the DCMIP2016 tropical cyclone protocol. Minimum central pressure (abscissa) and maximum azimuthally averaged 1000 m winds (ordinate) are plotted at 6 h intervals for participating models as in Fig. 6 of Willson et al. (2023). Results of the 25 km GEM control and OFFB51 configurations are shown with large, black-outlined blue and red symbols, respectively. Empirical quadratic wind-pressure relationships for each participating model are shown in thin solid lines whose color matches that of the corresponding points. Model acronyms in the legend follow the definitions in Table 4 of Willson et al. (2023), including: the Energy Exascale Earth System Model (ACME-A), the Community Atmosphere Spectral Element Model (CAM-SE), the Finite Volume Module of the Integrated Forecasting System (FVM), and the Non-hydrostatic Icosahedral Atmospheric Model (NICAM).

to headline scores are modest; however, the model's kinetic energy spectrum and depiction of the strong winds in the stratospheric polar vortex appear to be improved.

In terms of tropical cyclones, this final step of the investigation confirms that the OFFB51 configuration of the GDPS yields the expected reduction in the system's weak-intensity bias (Fig. 22). Maximum wind speeds increase by up to 3 m s^{-1} as central pressures drop by nearly 5 hPa. The increase of mean storm strength with lead time is consistent with the imprint of the model's weak-intensity bias on the initializing analysis.

Other standard tropical cyclone statistics do not show significant sensitivity to off-centering (not shown). For track forecasts, this result remains consistent with the DIMOSIC assessment given that day 1-3 improvements are unlikely in GDPS forecast sequences initialized with excessively

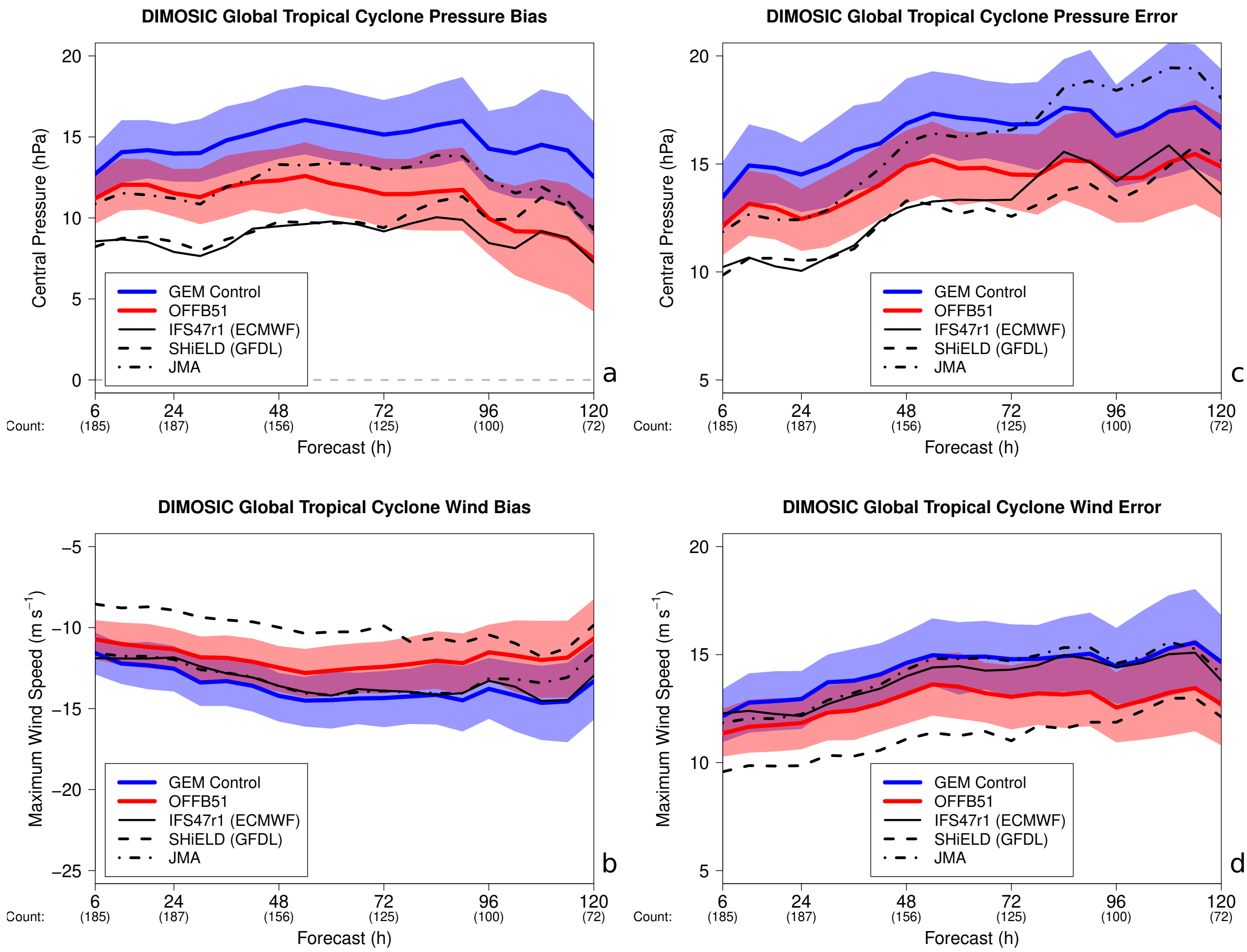


FIG. 20. Tropical cyclone intensity evaluation from the DIMOSIC project. Minimum central pressure bias (a) and root mean square error (c) are shown in hPa for a subset of participating DIMOSIC models (black lines with styles as shown in the plot legends), including the GEM control (blue) and OFFB51 (red) configurations. The 95% confidence interval for the plotted mean values are semi-transparently color-shaded for the GEM configurations. Equivalent plots of maximum 10 m wind speed bias (b) and root mean square error (d) are shown in m s^{-1} along the bottom row. The number of best track fixes that contribute to the plotted scores is shown in parentheses below the lead times along the abscissa. The zero line is plotted with gray dashing in (a) for reference.

weak storms. Both intensity and track results highlight the need for a full data assimilation cycle based on the OFFB51 configuration.

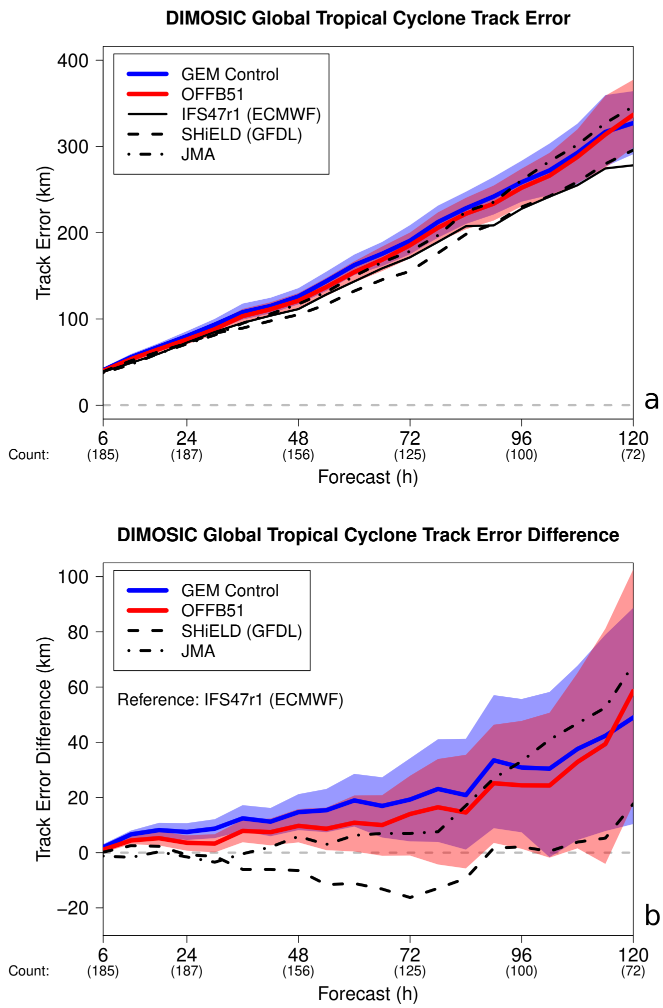


FIG. 21. Tropical cyclone track evaluation from the DIMOSIC project. The root mean square track error (a) is shown for the same subset of participating models as in Fig. 20, following the same plotting conventions. The difference between track errors in selected models and the ECMWF reference (IFS47r1: the forward model for generation of the operational ECMWF analyses used in the project) provides additional information about relative track forecast skill (b). The zero line is plotted with grey dashing for reference.

6. Discussion

A hierarchy of modeling complexity was used in this study to identify the source of a tropical cyclone weak-intensity bias in the Global Deterministic Prediction System (GDPS). The presence of the bias was confirmed at each step towards a semi-idealized framework based on a simplified model configuration. The resilience of the bias to fundamental changes to the physical parameterization

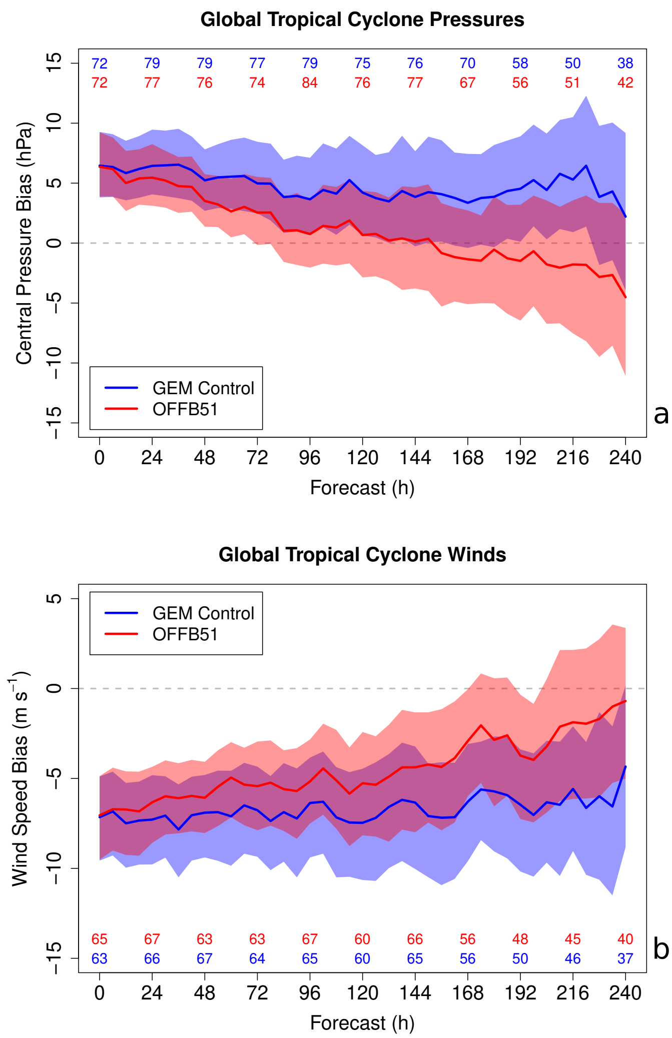


FIG. 22. Forecast time series of global tropical cyclone central pressure (a; in hPa) and maximum 10-m wind speed (b; in m s^{-1}) biases in the GEM control (blue) and OFFB51 (red) GDPS configurations. Mean biases are shown in solid lines, with the 95% confidence interval for the mean shown with semi-transparent color shading. A dashed gray line represents zero bias in each panel. The number of individual forecasts with a tracked tropical cyclone that was matched to an observed storm for the bias calculations is shown at 24-h intervals color-coded for each experiment across the top (a) or bottom (b) of the plot.

suite led to a closer examination of GEM's dynamical core that identified off-centering in the time-stepping scheme as the primary factor limiting simulated tropical cyclone intensity.

A dry vortex spin-down test designed to assess numerical dissipation showed the need for dramatic off-centering reduction (from $b = 0.6$ to $b = 0.51$). Subsequent re-evaluation of physical parameterization changes revealed important conditional sensitivities in the model; however, the leading contributor to the weak-intensity bias remained off-centering itself. This assertion stayed true as complexity was added back into the system, ultimately leading to the conclusion that tropical cyclone intensities in the reduced-dissipation GEM configuration resemble those of other global models with similar nominal resolutions. Although some rebalancing of physical parameterizations to account for increased effective resolution may be needed, reduced off-centering will serve as an important departure point for continued system development.

This study highlights the power of hierarchical development techniques, applied here as the progressive simplification of experimental protocols. As envisioned by Frissoni et al. (2023), this framework facilitated both the identification of the error source and its mitigation. The intercomparisons used throughout the investigation further increased the likelihood that the intensity bias reduction was achieved through physically relevant improvements to the simulations, rather than by error compensations within the system.

The proposed reduction in off-centering is consistent with progress made by other operational centers that employ implicit or semi-implicit time discretization. Both ECMWF and Météo France use alternative techniques to control spurious wave amplification without increasing dissipation or reducing accuracy (Ritchie and Tanguay 1996). Although this strategy has been found to decrease forecast skill in GEM, the fact that ECMWF guidance exhibits a relatively small tropical cyclone central pressure bias is consistent with the conclusions drawn here (Chen et al. 2023). In a more analogous system, the UKMO was able to reduce off-centering to b=0.55 with the introduction of a new dynamical core (Wood et al. 2014). Walters et al. (2017) attribute the significant intensification of tropical cyclones in ENDGame [e.g. Fig. 5 of Chen et al. (2023)] in part to this reduction in "implicit damping". Model intercomparison in the vortex spin-down framework developed here would help to determine whether there is a generally optimal value for the off-centering parameter, or whether implementation differences make it truly system-specific.

Although the GDPS is the main source of medium-range guidance for operational forecasters, it is not the only NWP system run at the CMC. The sensitivities of the global ensemble [39 km grid spacing; McTaggart-Cowan et al. (2021)] are typically found to be similar to those of the GDPS, such that simulated tropical cyclone intensities are expected to benefit from off-centering reduction. However, preliminary tests in the high-resolution system [2.5 km grid spacing; Milbrandt et al.

(2016)] suggest that $b = 0.51$ is a necessary but not sufficient condition for intensity bias reductions in the convection-permitting context. The conditional sensitivities identified in this study will serve as the basis for future efforts to improve intensity predictions in high-resolution configurations.

Reduction of the tropical cyclone weak-intensity bias is important for both high-impact weather forecasts and longer-range predictions involving tropical-extratropical interactions (Keller et al. 2019). Tropical cyclones also represent a stress-test for model formulations, with improved predictions an indication that the model better reproduces atmospheric extremes. In combination, these factors suggest that the proposed reduction of numerical dissipation in GEM will yield important benefits for the quality of guidance generated by Canadian NWP systems.

Acknowledgments. The authors thank Drs. Jean-François Caron and Ayrton Zadra for their work on the CMC tropical cyclone tracker and Vivian Lee for her support on DCMIP2016 configurations. The 2021 data for the WGNE evaluation shown in Fig. 2 were provided by Dr. Masashi Ujiie, while those for the DCMIP2016 tropical cyclone results shown in Fig. 19 were provided by Justin Willson and Dr. Kevin Reed. We sincerely appreciate these investigators' willingness to share their datasets and analysis techniques. The authors also appreciate Dr. Timothy Marchock's review of an earlier draft of this study. Suggestions made by editor Dr. Tommaso Benacchio and three anonymous reviewers were essential to deepening the analysis and preparing this work for publication. D. Nolan was supported by the Tropical Cyclone Rapid Intensification (TCRI) program of the Office of Naval Research, grant N00014-20-1-2075.

Data availability statement. The source code for the GEM model used in this study is available at https://github.com/ECCC-ASTD-MRD/gem/releases/tag/5.2.0-a24. Version 4.2.1 of the WRF model can be retrieved from https://github.com/wrf-model/WRF/archive/v4.2.1.tar.gz. Diagnostics are computed using the MetCal library (tag 2.0.0), available from https://sourceforge.net/projects/metcal. Model configurations, ensemble perturbation, diagnostic and plotting software is available at https://doi.org/10.5281/zenodo.8187835 (McTaggart-Cowan et al. 2023). Raw model outputs are too large to archive externally, but will be stored at ECCC for at least five years and will be made freely available upon request.

APPENDIX A

An Analytic Expression for Pressure Gradient Force Misalignment

An expression for the drag induced by off-centering via the misalignment of the pressure gradient force can be derived geometrically based on the schematic shown in Fig. A1. An underlying assumption adopted here is that the magnitude of the pressure gradient force (PGF) vector is insensitive to the small radial displacements implied by movement along the linear trajectory (blue dashed line in Fig. A1). The quality of results described in section 4c3 shows that this assumption does not lead to significant error in the final expression.

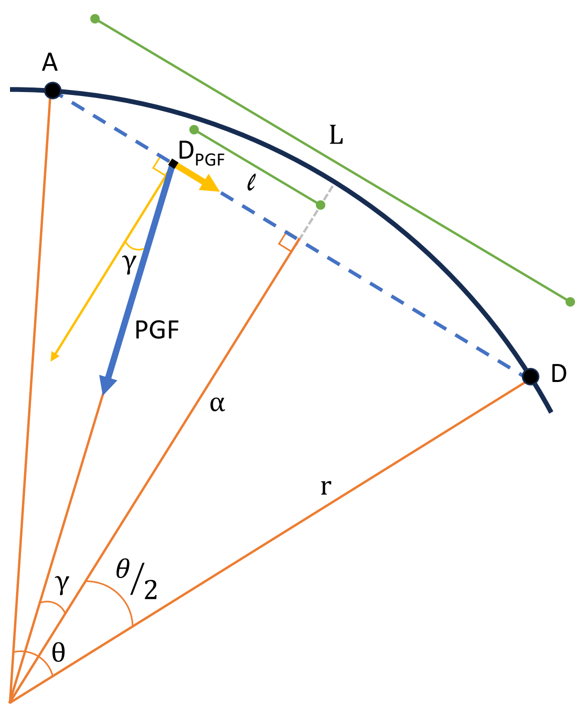


FIG. A1. Geometry of pressure gradient force (PGF) misalignment for $b > 0.5$, plotted following the conventions used for Fig. 13. Symbols are defined in the text.

The component of pressure gradient force oriented along the D⟶A linear trajectory is,

$$D_{PGF} = -(PGF)\sin\gamma \quad \text{where} \quad \gamma = \tan^{-1}\frac{\ell}{\alpha} \quad \text{but,} \tag{A1}$$

$$\alpha = r\cos\left(\frac{\theta}{2}\right) \quad \text{and} \quad \ell = (b - {}^{1}\!/_{2})\,L \quad , \tag{A2}$$

$$\tag{A3}$$

is the distance between the linear trajectory midpoint (at $L/2$) and point at which the PGF is valid as per Eq. 4. Using,

$$L = 2r\sin\left(\frac{\theta}{2}\right) \quad , \tag{A4}$$

yields the expression,

$$D_{PGF} = -(PGF)\sin\left\{\tan^{-1}\left[\frac{2\left(b-{}^{1}\!/\!{}_{2}\right)\sin\left(\frac{\theta}{2}\right)}{\cos\left(\frac{\theta}{2}\right)}\right]\right\} \tag{A5}$$

$$= -(PGF)\sin\left\{\tan^{-1}\left[2\left(b-{}^{1}\!/\!{}_{2}\right)\tan\left(\frac{\theta}{2}\right)\right]\right\} \quad . \tag{A6}$$

The angle (θ) swept over a time step by a parcel travelling at tangential speed v is simply $\frac{v\delta t}{r}$, so

$$D_{PGF} = -(PGF)\sin\left\{\tan^{-1}\left[2\left(b-{}^{1}\!/\!{}_{2}\right)\tan\left(\frac{v\delta t}{2r}\right)\right]\right\} \quad . \tag{A7}$$

Using the trigonometric identity for inverse functions,

$$D_{PGF} = \frac{-(PGF)\chi}{\sqrt{1+\chi^2}} \quad \text{where} \quad \chi = 2\left(b-{}^{1}\!/\!{}_{2}\right)\tan\left(\frac{v\delta t}{2}\right) \quad . \tag{A8}$$

To complete this analysis we employ gradient balance,

$$\frac{v^2}{r} + fv = (PGF) \quad \text{for} \quad (PGF) = \frac{1}{\rho}\frac{\partial p}{\partial r} \quad , \tag{A9}$$

to obtain the final estimate of acceleration related to misalignment of the pressure gradient force,

$$D_{PGF}(v,r,f,\delta t,b) = -\frac{\chi\left(\frac{v^2}{r}+fv\right)}{\sqrt{1+\chi^2}} \quad . \tag{A10}$$

APPENDIX B

Numerical Spin-Down in a Shallow-Water Vortex

The description of the tropical cyclone spin-down process (section 4c3) does not rely on 3D storm structure. The shallow-water system is therefore used here to quantify the vortex decay induced by off-centering in a minimum-complexity framework.

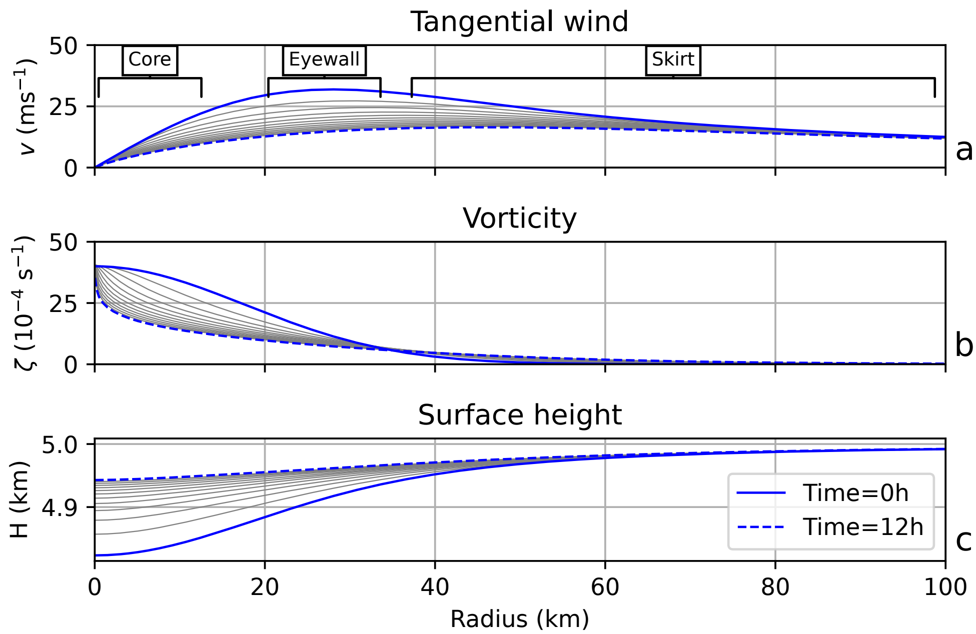


FIG. B1. Structure of tangential wind (a), vorticity (b) and height (c) for $b = 0.6$ from the initialization (solid blue line) to the 12 h state (dashed blue line) at hourly increments (thin black lines). The approximate radial bands that represent the three vortex regions treated separately in section c of this appendix are annotated in (a).

a. Model Description

The model is initialized with a Gaussian vortex that approximates a solid-body core and an irrotational "skirt". Similar to a Rankine vortex, the radius of maximum wind (eyewall) is located between these two components (Fig B1a). An important advantage of the Gaussian vortex is the finite width of this approximate eyewall region, where the maximum deceleration is found to occur.

In this axisymmetric (1D) framework, the initial relative vorticity (ζ) is a function of radius from the center (r) and is given by,

$$\zeta = \zeta_\circ e^{-r^2/L^2} \quad , \tag{B1}$$

which implies a tangential wind field,

$$v = \frac{L^2 \zeta_\circ}{2r}\left(1 - e^{r^2/L^2}\right) \quad . \tag{B2}$$

Setting parameters L and $\zeta_\circ$ to 25 km and 30×10^{-4} s^{-1} yields maximum winds of 32 m s^{-1} at $r = 28$ km (blue contours in Fig. B1). The initial layer height (H) is in centripetal balance,

$$g\frac{\partial H}{\partial r} = -v^2/r \quad , \tag{B3}$$

where $g = 9.81$ m s^{-2} is gravitational acceleration and $H = 5$ km is used as the far-field boundary condition for numerical solution.

The solution is propagated forwards in time using a spectral method to essentially eliminate discretization error as a complicating factor in this analysis. A total of $N = 128$ solution points are placed at Gauss-Legendre quadrature points on the interval $x \in (-1,1)$, which is scaled to the interval $r \in (0, \inf)$ by the relationship $r = L\sqrt{N}\tan\left[\frac{\pi}{4}(1+x)\right]$.

The time-discretized shallow water equations are,

$$\frac{U^A - U^D}{\delta t} = g\left[b\left(\frac{\partial H}{\partial x}\right)^A + (1-b)\left(\frac{\partial H}{\partial x}\right)^A\right] \tag{B4}$$

$$\frac{V^A - V^D}{\delta t} = g\left[b\left(\frac{\partial H}{\partial y}\right)^A + (1-b)\left(\frac{\partial H}{\partial y}\right)^A\right] \tag{B5}$$

$$\frac{H^A - H^D}{\delta t} = b\left[H\left(\frac{\partial U}{\partial x} + \frac{\partial V}{\partial y}\right)\right]^A + (1-b)\left[H\left(\frac{\partial U}{\partial x} + \frac{\partial V}{\partial y}\right)\right]^D \quad , \tag{B6}$$

where U and V are winds in the x and y directions on the model's Cartesian grid. The system is solved using four iterations for each time step ($\delta t = 450$ s). Mapping model winds into tangential (v) and radial (u) components is accomplished via the transforms,

$$U = v\frac{-y}{\sqrt{x^2+y^2}} + u\frac{x}{\sqrt{x^2+y^2}} \tag{B7}$$

$$V = v\frac{x}{\sqrt{x^2+y^2}} + u\frac{y}{\sqrt{x^2+y^2}} \quad . \tag{B8}$$

To close the discretized system, the implicit solve of (U,V,H) employs a boundary condition of $V(y=0) = 0$ m s^{-1} and $\frac{\partial H}{\partial y} = 0$ instead of disretizing the system at the smallest radius, a formulation that avoids the formation of a cusp at the origin. No boundary conditon at infinity is necessary.

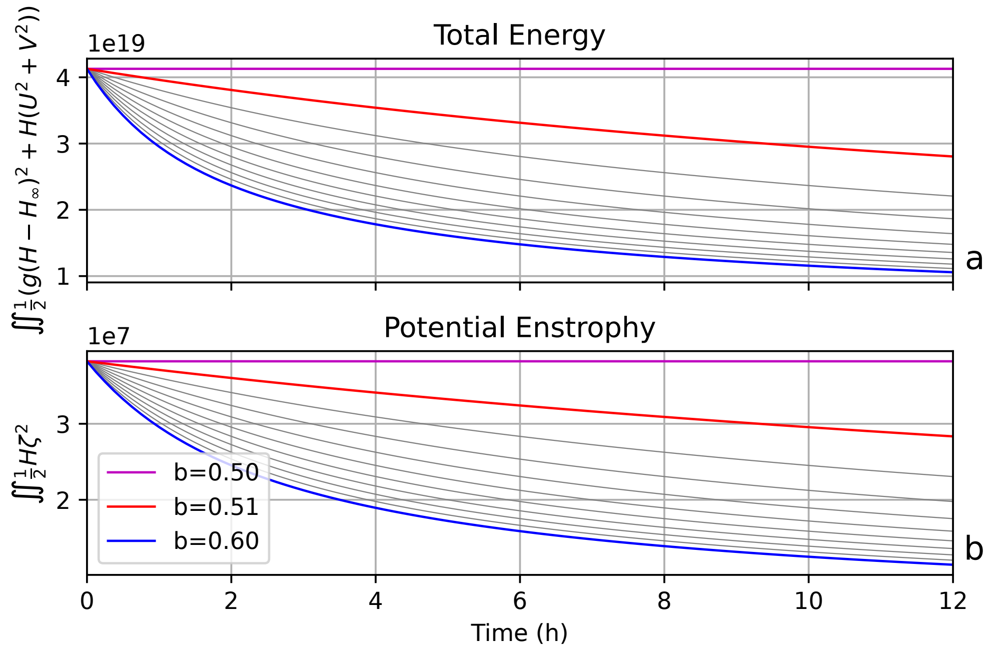


FIG. B2. Time evolution of domain-integrated total energy (a) and potential enstrophy (b) for values of b from 0.5 (magenta) to 0.6 (blue) at 0.01 increments (thin black lines). Evolution of the vortex with $b = 0.51$ is highlighted in red for consistency with Fig. 10.

b. Vortex Spin-Down

The model's discretization of the shallow water system exposes it to the misalignment of the pressure gradient force described in section 4c3 of the main text (Fig. 13). Adopting $b = 0.6$ as in the GEM control configuration leads to a rapid spin-down of the vortex (Fig. B1). The similiarity between the speed of initial decay in this low-order system and the rapid filling in the spin-down test (Fig. 10) suggests strongly that the numerical error sources represented here dominate 3D vortex evolution.

The justification for characterizing the effects of these numerical errors as a "3D friction" in section 4c3 is found in Fig. B2. The rapid decay of total energy and potential enstrophy indicates that the vortex spin-down is directly related to non-conservation rather than energy cascades or radial expansion. The large sensitivity of these otherwise-conserved quantities to even small values of off-centering is again consistent with the results of the GEM spin-down simulations that lead to the proposed $b = 0.51$ dynamical core configuration.

c. *Radius-Dependent Impacts of Off-Centering*

The Gaussian vortex is split into three conceptual sub-regions (shown schematically in Fig. B1a) in which impact of off-centering induced numerical drag is assessed: a core in solid-body rotation, an irrotational skirt and an eyewall with approximately constant wind speeds. For simplicity only the point $(x,y) = (0,1)$ is considered (the arrival point in Fig. 13) and the vortex parameters are scaled such that $U(0,1) = -1$, $H(0,1) = 1$ and $g = 1$. Drag is evaluated in the limit $\delta t \to 0$ such that higher orders represent smaller contributions to vortex deceleration. Consistent with the piggybacking approach employed in section 4c3, the leading-order effects of the drag are isolated here by holding the flow constant for the purposes of trajectory, divergence and pressure gradient calculations.

1) Numerical Drag in the Solid-Body Core

The normalized solid-body core is described by,

$$U = -y \tag{B9}$$

$$V = x \tag{B10}$$

$$H = \frac{1}{2} + \frac{1}{2}\left(x^2 + y^2\right) \quad , \tag{B11}$$

for which Eulerian changes in state variables from current to future times (superscripts "-" and "+", respectively) are,

$$\frac{U^+ - U^-}{\delta t} = \frac{y\delta t}{2}(2b-1) + O\left(\delta t^2\right) \tag{B12}$$

$$\frac{V^+ - V^-}{\delta t} = -\frac{y\delta t^2}{4}(2b-1) + O\left(\delta t^4\right) \tag{B13}$$

$$\frac{H^+ - H^-}{\delta t} = 0 + O\left(\delta t^4\right) \quad , \tag{B14}$$

showing that off-centering causes a direct spin-down of the vortex in this region. The deceleration is proportional to radius for the simple $y \to r$ mapping at $(x,y) = (0,1)$ and thus to v through Eq. B9. This means that the leading term in Eq. B12 represents first-order friction for any $b > 0.5$. The error in radial acceleration is one order higher in δt and layer height in the solid-body core is conserved.

2) Numerical Drag in the Irrotational Skirt

In the normalized irrotational skirt,

$$U = \frac{-y}{x^2+y^2} \tag{B15}$$
$$V = \frac{x}{x^2+y^2} \tag{B16}$$
$$H = \frac{3}{2} - \frac{1}{2\left(x^2+y^2\right)} \quad , \tag{B17}$$

which evolve following,

$$\frac{U^+ - U^-}{\delta t} = \frac{\delta t}{2y^5}(2b-1) + O\left(\delta t^3\right) \tag{B18}$$
$$\frac{V^+ - V^-}{\delta t} = -\frac{\delta t^2}{4y^7} b(2b-1) + O\left(\delta t^4\right) \tag{B19}$$
$$\frac{H^+ - H^-}{\delta t} = 0 + O\left(\delta t^4\right) \quad . \tag{B20}$$

The leading error induced by off-centering affects the tangential wind and scales as $\delta t(2b-1)$ in this region. However, the effects of this numerical drag are concentrated in the inner portion of the irrotational skirt ($y^{-5} \implies r^{-5}$) closest to the eyewall.

3) Numerical Drag in the Eyewall

The eyewall is defined as the region close to the radius of maximum wind where the flow is well approximated by,

$$U = \frac{-y}{\sqrt{x^2+y^2}} \tag{B21}$$
$$V = \frac{x}{\sqrt{x^2+y^2}} \tag{B22}$$
$$H = 1 + \frac{1}{2}\log\left(x^2+y^2\right) \quad . \tag{B23}$$

Solution of this system requires power series expansion followed by matching of terms by order of δt. These steps yield tendencies that have the form,

$$\frac{U^+ - U^-}{\delta t} = \frac{\delta t}{2y}(2b-1) + O\left(\delta t^3\right) \tag{B24}$$

$$\frac{V^+ - V^-}{\delta t} = -\frac{\delta t^2}{4y^3} b(2b-1) + O\left(\delta t^4\right) \tag{B25}$$

$$\frac{H^+ - H^-}{\delta t} = 0 + O\left(\delta t^4\right) \quad . \tag{B26}$$

The r^{-1} scaling in the $O(\delta t)$ tangential deceleration term at $x = 0$ (Eq. B24) implies that the effects of numerical drag are maximized in the eyewall. This is consistent with both the rapid decay in this region observed in the shallow water system (Fig. B1a) and the tangential wind decelerations diagnosed in the GEM spin-down simulations (Fig. 10).

References

Arakawa, A., and S. Moorthi, 1988: Baroclinic instability in vertically discrete systems. *J. Atmos. Sci.*, **45**, 1688–1707.

Bechtold, P., M. Köhler, T. Jung, F. Doblas-Reyes, M. Leutbecher, M. J. Rodwell, F. Vitart, and G. Balsamo, 2008: Advances in simulating atmospheric variability with the ECMWF model: from synoptic to decadal time scales. *Quart. J. Roy. Meteor. Soc.*, **134**, 1337–1351.

Bélair, S., J. Mailhot, J. W. Strapp, and J. I. MacPherson, 1999: An examination of local versus non-local aspects of a tke-based boundary layer scheme in clear convective conditions. *J. Appl. Meteor.*, **38**, 1499–1518.

Buehner, M., and Coauthors, 2015: Implementation of deterministic weather forecasting systems based on ensemble–variational data assimilation at Environment Canada. *Mon. Wea. Rev.*, **143**, 2532–2559.

Cangialosi, J. P., E. Blake, M. DeMaria, A. Penny, A. Latto, E. Rappaport, and V. Tallapragada, 2020: Recent progress in tropical cyclone intensity forecasting at the National Hurricane Center. *Wea. Forecasting*, **35**, 1913–1922.

Caron, J.-F., and M. Buehner, 2022: Implementation of scale-dependent background-error covariance localization in the canadian global deterministic prediction system. *Wea. Forecasting*, **37**, 1567–1580.

Chen, J., and D. R. Chavas, 2020: The transient responses of an axisymmetric tropical cyclone to instantaneous surface roughening and drying. *J. Atmos. Sci.*, **77**, 2807–2834.

Chen, J.-H., L. Zhou, L. Magnusson, R. McTaggart-Cowan, and M. Köhler, 2023: Tropical cyclone forecasts in the dimosic project - medium range forecast models with common initial conditions. *Earth and Space Sci.*

Cornforth, R. J., and B. J. Hoskins, 2009: Understanding African easterly waves: a moist singular vector approach. *Atm. Sci. Lett.*, **10**, 185–191.

Courtney, J., and Coauthors, 2019: Operational perspectives on tropical cyclone intensity change part 1: recent advances in intensity guidance. *Trop. Cyclone Res. and Rev.*, **8**, 123–133.

Cressman, G. P., 1959: An operational objective analysis system. *Mon. Wea. Rev.*, **87**, 367–374.

Davis, C. A., 2018: Resolving tropical cyclone intensity in models. *Geophys. Res. Lett.*, **45**, 2082–2087.

DeMaria, M., C. R. Sampson, J. A. Knaff, and K. D. Musgrave, 2014: Is tropical cyclone intensity guidance improving? *Bull. Amer. Meteor. Soc.*, **95**, 387–398.

DeMaria, M., and Coauthors, 2022: The National Hurricane Center tropical cyclone model guidance suite. *Wea. Forecasting*, 2141–2159.

ECMWF, 2018a: IFS documentation - Cy45r1. Part IV: physical processes. Tech. rep., ECMWF, Reading, United Kingdom.

ECMWF, 2018b: Ifs documentation CY45r1. Tech. rep., ECMWF Tech. Rep. Https://www.ecmwf.int/en/publications/ifs-documentation.

Emanuel, K. A., 1988: Observational evidence of slantwise convective adjustment. *Mon. Wea. Rev.*, **116**, 1805–1816.

Fovell, R. G., Y. P. Bu, K. L. Corbosiero, W. Tung, Y. Cao, H. Kuo, L. Hsu, and H. Su, 2016: Influence of cloud microphysics and radiation on tropical cyclone structure and motion. *Meteor. Monogr.*, **56**, 11.2–11.27.

Frissoni, A., and Coauthors, 2023: Systematic errors in weather and climate models: Challenges and opportunities in complex coupled modeling systems. *Bull. Amer. Meteor. Soc.*

Gall, R., J. Franklin, F. Marks, E. N. Rappaport, and F. Toepfer, 2013: The Hurricane Forecast Improvement Project. *Bull. Amer. Meteor. Soc.*, **94**, 329–343.

Girard, C., and Coauthors, 2014: Staggered vertical discretization of the Canadian Global Environmental Multiscale (GEM) model using a coordinate of the log-hydrostatic-pressure type. *Mon. Wea. Rev.*, **142**, 1183–1196.

Grabowski, W. W., 2014: Extracting microphysical impacts in large-eddy simulations of shallow convection. *J. Atmos. Sci.*, **71**, 4493–4499.

Green, B. W., and F. Zhang, 2013: Impacts of air–sea flux parameterizations on the intensity and structure of tropical cyclones. *Mon. Wea. Rev.*, **141**.

Harris, L. M., S. Lin, and C. Tu, 2016: High-resolution climate simulations using GFDL HiRAM with a stretched global grid. *J. Climate*, **29**, 4293–4314.

Heming, J. T., and Coauthors, 2019: Review of recent progress in tropical cyclone track forecasting and expression of uncertainties. *Trop. Cyclone Res. and Rev.*, **8**, 181–218.

Hendricks, E. A., M. T. Montgomery, and C. A. Davis, 2004: The role of ”vortical” hot towers in the formation of tropical cyclone Diana (1984). *J. Atmos. Sci.*, **61**, 1209–1232.

Hersbach, H., and Coauthors, 2020: The era5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146**, 1999–2049.

Hlywiak, J., and D. S. Nolan, 2021: The response of the near-surface tropical cyclone wind field to inland surface roughness length and soil moisture content during and after landfall. *J. Atmos. Sci.*, **78**, 983–1000.

Hodges, K. I., and N. P. Klingaman, 2019: Prediction errors of tropical cyclones in the western north pacific in the Met Office global forecast model. *Wea. Forecasting*, **34**, 1189–1209.

Holliday, C. R., and A. H. Thompson, 1979: Climatological characteristics of rapidly intensifying typhoons. *Mon. Wea. Rev.*, **107**, 1022–1034.

Hong, S.-Y., J. Dudhia, and S.-H. Chen, 2004: A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation. *Mon. Wea. Rev.*, **132**, 103–120.

Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318–2341.

Hourdin, F., and Coauthors, 2017: The art and science of climate model tuning. *Bull. Amer. Meteor. Soc.*, **98**, 589–602.

Houze, R. A., S. S. Chen, B. F. Smull, W.-C. Lee, and M. M. Bell, 2007: Hurricane intensity and eyewall replacement. *Science*, **315**, 1235–1239.

Husain, S. Z., C. Girard, A. Qaddouri, and A. Plante, 2019: A new dynamical core of the Global Environmental Multiscale (GEM) model with a height-based terrain-following vertical coordinate. *Mon. Wea. Rev.*, **147**, 2555–2578.

Jablonowski, C., and D. L. Williamson, 2011: *Numerical Techniques for Global Atmospheric Models*, chap. Chapter 13: The Pros and Cons of Diffusion, Filters and Fixers in Atmospheric General Circulation Models, 381–493. Springer.

Jacob, C., 2010: Accelerating progress in global atmospheric model development through improved parameterizations. *Bull. Amer. Meteor. Soc.*, **91**, 1189–1209.

Jordan, C. L., 1958: Mean soundings for the West Indies area. *J. Meteor.*, **15**, 91–97.

Judt, F., R. Rios-Berrios, and G. H. Bryan, 2023: Marathon vs. sprint: two modes of tropical cyclone rapid intensification in a global convection-permitting simulation. *Mon. Wea. Rev.*

Judt, F., and Coauthors, 2021: Tropical cyclones in global storm-resolving models. *J. Meteor. Soc. Japan*, **99**, 579–602.

Kain, J., and J. M. Fritsch, 1993: Convective parameterization for mesoscale models: The Kain-Fritsch scheme, chapter 16. *The representation of cumulus convection in numerical models of the atmosphere*, K. A. Emanuel, and D. J. Raymond, Eds., Meteo. Mon., American Meteorological Society, 165–170.

Kain, J. S., 2004: The Kain-Firtsch convective parameterization: an update. *J. Appl. Meteor.*, **43**, 170–181.

Kain, J. S., and J. M. Fritsch, 1990: A one-dimensional entraining/detraining plume model and its application in convective parameterization. *J. Atmos. Sci.*, **47**, 2784–2802.

Kain, J. S., and J. M. Fritsch, 1992: The role of the convective 'trigger' function in numerical forecasts of mesoscale convective systems. *Meteor. Atmos. Phys.*, **49**, 93–106.

Kaplan, J., M. DeMaria, and J. A. Knaff, 2010: A revised tropical cyclone rapid intensification index for the Atlantic and Eastern North Pacific basins. *Wea. Forecasting*, **25**, 220–241.

Keller, C. M., J. H. Grams, and Coauthors, 2019: The extratropical transition of tropical cyclones. Part II: Interaction with the midlatitude flow, downstream impacts, and implications for predictability. *Mon. Wea. Rev.*, **147**, 1077–1106.

Kessler, E., 1969: On the distribution and continuity of water substance in atmosphere circulations. *Meteor. Monogr.*, Vol. 32, Amer. Meteor. Soc.

Knaff, J. A., and Coauthors, 2021: Estimating tropical cyclone surface winds: Current status, emerging technologies, historical evolution, and a look to the future. *Trop. Cyclone Res. and Rev.*, **10**, 125–150.

Knapp, K. R., M. C. Kruk, D. H. Levinson, H. J. Diamond, and C. J. Neumann, 2010: The International Best Track Archive for Climate Stewardship (IBTrACS): Unifying tropical cyclone best track data. *Bull. Amer. Meteor. Soc.*, **91**, 363–376.

Landsea, C., and J. P. Cangialosi, 2018: Have we reached the limits of predictability for tropical cyclone track forecasting? *Bull. Amer. Meteor. Soc.*, **99**, 2237–2243.

Li, J., and H. W. Barker, 2005: A radiation algorithm with correlated-k distribution. Part I: local thermal equilibrium. *J. Atmos. Sci.*, **62**, 286–309.

Magnusson, L., and Coauthors, 2019: ECMWF activities for improved hurricane forecasts. *Bull. Amer. Meteor. Soc.*, **100**, 445–458.

Magnusson, L., and Coauthors, 2022: Skill of medium-range forecast models using the same initial conditions. *Bull. Amer. Meteor. Soc.*, **103**, E2050–E2068.

Majumdar, S. J., L. Magnusson, P. Bechtold, J.-R. Bidlot, and J. D. Doyle, 2023: Advanced tropical cyclone prediction using the experimental global ECMWF and operational regional COAMPS-TC systems. *Mon. Wea. Rev.*, **151**, 2029–2048.

McTaggart-Cowan, R., L. Separovic, M. Charron, X. Deng, N. Gagnon, P. Houtekamer, and A. Patoine, 2021: Using stochastically perturbed parameterizations to represent model uncertainty, part II: comparison with existing techniques in an operational ensemble. *Mon. Wea. Rev.*

McTaggart-Cowan, R., P. A. Vaillancourt, A. Zadra, L. Separovic, S. Covec, and D. Kirshbaum, 2019a: A Lagrangian perspective on parameterizing deep convection. *Mon. Wea. Rev.*, **147**, 4127–4149.

McTaggart-Cowan, R., and A. Zadra, 2015: Representing Richardson number hysteresis in the NWP boundary layer. *Mon. Wea. Rev.*, **143**, 1232–1258.

McTaggart-Cowan, R., and Coauthors, 2019b: Modernization of atmospheric physics in Canadian NWP. *J. Adv. Model. Earth Syst.*, **11**, 3593–3635.

McTaggart-Cowan, R., and Coauthors, 2023: Supporting dataset for ”Reducing a tropical cyclone weak-intensity bias in a global numerical weather prediction system” (version 1). Zenodo, access date 26 July 2023, https://doi.org/10.5281/zenodo.8187835.

Milbrandt, J. A., S. Bélair, M. Faucher, M. Vallée, M. L. Carrera, and A. Glazer, 2016: The Pan-Canadian High Resolution (2.5 km) Deterministic Prediction System. *Wea. Forecasting*, **31**, 1791–1816.

Morrison, H., and J. A. Milbrandt, 2015: Parameterization of cloud microphysics based on the prediction of bulk ice particle properties. Part I: scheme description and idealized tests. *J. Atmos. Sci.*, **72**, 287–311.

Nolan, D. ., R. Atlas, K. T. Bhatia, and L. R. Bucci, 2013: Development and validation of a hurricane nature run using the joint OSSE nature run and the WRF model. *J. Adv. Model. Earth Syst.*, **5**, 382–405.

Nolan, D. S., 2011: Evaluating environmental favorableness for tropical cyclone development with the method of point-downscaling. *J. Adv. Model. Earth Syst.*, **3**, M08 001, 28 pp.

Park, J., D.-H. Cha, M. K. Lee, J. Moon, S.-J. Hahm, K. Noh, J. C. L. Chan, and M. Bell, 2020: Impact of cloud microphysics schemes on tropical cyclone forecast over the western North Pacific. *J. Geophys. Res.-Atmospheres*, **125**, e2019JD032 288.

Park, S.-H., W. Skamarock, J. Klemp, L. Fowler, and M. Duda, 2013: Evaluation of global atmospheric solvers using extension of the Jablonowski and Williamson baroclinic wave test case. *Mon. Wea. Rev.*, **141**, 3116–3129.

Powell, M. D., P. J. Vickery, and T. A. Reinhold, 2003: Reduced drag coefficients for high wind speeds in tropical cyclones. *Nature*, **422**, 279–283.

Qaddouri, A., and V. Lee, 2011: The Canadian Global Environmental Multiscale model on the Yin-Yang grid system. *Quart. J. Roy. Meteor. Soc.*, **137**, 1913–1926.

Reed, K. A., and C. Jablonowski, 2011: An analytic vortex initialization technique for idealized tropical cyclone studies in AGCMs. *Mon. Wea. Rev.*, **139**, 689–710.

Reed, K. A., and C. Jablonowski, 2012: Idealized tropical cyclone simulations of intermediate complexity: a test case for AGCMs. *J. Adv. Model. Earth Syst.*, **4**, M04 001.

Ritchie, H., and M. Tanguay, 1996: A comparison of spatially averaged Eulerian and semi-Lagrangian treatements of mountains. *Mon. Wea. Rev.*, **124**, 167–181.

Rivest, C., A. Staniforth, and A. Robert, 1994: Spurious resonant response of semi-lagrangian discretizations to orographic forcing: Diagnosis and solution. *Mon. Wea. Rev.*, **122**, 366–376.

Rogers, R., 2021: Recent advances in our understanding of tropical cyclone intensity change processes from airborne observations. *Atmosphere*, **12**, 36 pp.

Rogers, R. F., P. D. Reasor, and J. A. Zhang, 2015: Multiscale structure and evolution of Hurricane Earl (2010) during rapid intensification. *Mon. Wea. Rev.*, **143**, 536–562.

Rozoff, C. M., D. S. Nolan, J. P. Kossin, F. Zhang, and J. Fang, 2012: The roles of an expanding wind field and inertial stability in tropical cyclone secondary eyewall formation. *J. Atmos. Sci.*, **69**, 2621–2643.

Ruppert Jr., J. H., W. A. A., X. Tang, and E. L. Duran, 2020: The critical role of cloud-infrared radiation feedback in tropical cyclone development. *Proc. Natl. Acad. Sci. USA*, **117**, 27 884–27 892.

Ryglicki, D. R., J. H. Cossunth, D. Hodyss, and J. D. Doyle, 2018: The unexpected rapid intensification of tropical cyclones in moderate vertical wind shear. Part I: overview and observations. *Mon. Wea. Rev.*, **146**, 3773–3800.

Sharma, M., and R. Berg, 2022: Topic 5: Forecasting tropical cyclone hazards and impacts. Tech. rep., Tenth International Workshop on Tropical Cyclones (IWTC-10), World Meteorological Organization, 16 pp pp.

Simpson, R. H., 1974: The hurricane disaster potential scale. *Weatherwise*, **27**, 169–186.

Sinclair, M., 1997: Objective identification of cyclones and their circulation intensity, and climatology. *Wea. Forecasting*, **12**, 595–612.

Sinclair, M. R., 2004: Extratropical transition of Southwest Pacific tropical cyclones. Part II: midlatitude circulation characteristics. *Mon. Wea. Rev.*, **132**, 2145–2168.

Sitkowski, M., J. P. Kossin, and C. M. Rozoff, 2011: Intensity and structure changes during hurricane eyewall replacement cycles. *Mon. Wea. Rev.*, **139**, 3829–3847.

Skamarock, W. C., 2004: Evaluating mesoscale NWP models using kinetic energy spectra. *Mon. Wea. Rev.*, **132**, 3019–3032.

Skamarock, W. C., and A. Gassmann, 2011: Conservative transport schemes for spherical geodesic grids: High-order flux operators for ODE-based time integration. *Mon. Wea. Rev.*, **139**, 2962–2975.

Skamarock, W. C., and Coauthors, 2019: A description of the Advanced Research WRF model version 4. Tech. Rep. NCAR/TN-556+STR, NCAR Technical Note.

Smith, G. C., and Coauthors, 2018: Impact of coupling with an ice-ocean model on global medium-range NWP forecast skill,. *Mon. Wea. Rev.*, **146**, 1157–1180.

Smith, R. K., and M. T. Montgomery, 2016: The efficiency of diabatic heating and tropical cyclone intensification. *Quart. J. Roy. Meteor. Soc.*, **142**, 2081–2086.

Smith, R. K., M. T. Montgomery, and N. Van Sang, 2009: Tropical cyclone spin-up revisited. *Quart. J. Roy. Meteor. Soc.*, **135**, 1321–1335.

Stern, D. P., J. L. VIgh, D. S. Nolan, and F. Zhang, 2015: Revisiting the relationship between eyewall contraction and intensification. *J. Atmos. Sci.*, **72**, 1283–1306.

Stevens, B., and Coauthors, 2019: DYAMOND: the DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains. *Prog. Earth Planet. Sci.*, **6 (61)**, 17 pp.

Subich, C., 2022: Instabilities in the shallow-water system with semi-Lagrangian, time-centered discretization. *Mon. Wea. Rev.*, **150**, 467–480.

Sundqvist, H., E. Berge, and J. E. Kristjánsson, 1989: Condensation and cloud parameterization studies with a mesoscale numerical weather prediction model. *Mon. Wea. Rev.*, **117**, 1641–1657.

Trabing, B. C., M. M. Bell, and B. R. Brown, 2019: Impacts of radiation and upper-tropospheric temperatures on tropical cyclone structure and intensity. *J. Atmos. Sci.*, **76**, 135–153.

Tuppi, L., M. Ekblom, P. Ollinaho, and H. Järvinen, 2023: Simultaneous optimization of 20 key parameters of the Integrated Forecasting System of ECMWF Using OpenIFS: Part I (effect on deterministic forecasts). *Mon. Wea. Rev.*

Ullrich, P. A., and Coauthors, 2017: DCMIP2016: a review of non-hydrostatic dynamical core design and intercomparison of participating models. *Geosci. Mod. Dev.*, **10**, 4477–4509.

Van Sang, N., R. K. Smith, and M. T. Montgomery, 2008: Tropical cyclone intensification and predictability in three dimensions. *Quart. J. Roy. Meteor. Soc.*, **134**, 563–582.

Walters, D., and Coauthors, 2017: the Met Office Model Global Atmosphere 6.0/6.1 and JULES Global Land 6.0/6.1 configurations. *Geosci. Mod. Dev.*, **10**, 1487–1520.

Wang, Y., 2009: How do outer spiral rainbands affect tropical cyclone structure and intensity? *J. Atmos. Sci.*, **66**, 1250–1273.

Wang, Y., and Z. Tan, 2020: Outer rainbands-driven secondary eyewall formation of tropical cyclones. *J. Atmos. Sci.*, **77**, 2217–2236.

Wicker, L. J., and W. C. Skamarock, 2002: Time splitting methods for elastic models using forward time schemes. *Mon. Wea. Rev.*, **130**, 2088–2097.

Willoughby, H. E., J. A. Clos, and M. G. Shoreibah, 1982: Concentric eyewalls, secondary wind maxima, and the evolution of the hurricane vortex. *J. Atmos. Sci.*, **39**, 395–411.

Willson, J. L., and Coauthors, 2023: DCMIP2016: the tropical cyclone test case. *Geosci. Mod. Dev.*

Wood, N., and Coauthors, 2014: An inherently mass-conserving semi-implicit semi-Lagrangian discretization of the deep-atmosphere global non-hydrostatic equations. *Quart. J. Roy. Meteor. Soc.*, **140**, 1505–1520.

Wu, S.-N., B. J. Soden, and G. J. Alaka, 2020: The influence of radiation on the prediction of tropical cyclone intensification in a forecast model. *Geophys. Res. Lett.*, **50**, e2022GL099 442.

Yamaguchi, M., J. Ishida, H. Sato, and M. Nakagawa, 2017: WGNE intercomparison of tropical cyclone forecasts by operational NWP models: A quarter century and beyond. *Bull. Amer. Meteor. Soc.*, **98**, 2337–2349.

Zadra, A., R. McTaggart-Cowan, P. A. Vaillancourt, M. Roch, S. Bélair, and A.-M. Leduc, 2014: Evaluation of tropical cyclones in the Canadian global modeling system: sensitivity to moist process parameterization. *Mon. Wea. Rev.*, **142**, 1197–1220.

Zhou, X., and B. Wang, 2011: Mechanism of concentric eyewall replacement cycles and associated intensity change. *J. Atmos. Sci.*, **68**, 972–988.