

Dynamical Tests of a Deep Learning Weather Prediction Model[©]

GREGORY J. HAKIM^a AND SANJIT MASANAM^b

^a Department of Atmospheric Sciences, University of Washington, Seattle, Washington

^b Department of Physics, University of California at Santa Barbara, Santa Barbara, California

(Manuscript received 23 October 2023, in final form 21 April 2024, accepted 15 May 2024)

ABSTRACT: Global deep learning weather prediction models have recently been shown to produce forecasts that rival those from physics-based models run at operational centers. It is unclear whether these models have encoded atmospheric dynamics or simply pattern matching that produces the smallest forecast error. Answering this question is crucial to establishing the utility of these models as tools for basic science. Here, we subject one such model, Pangu-Weather, to a set of four classical dynamical experiments that do not resemble the model training data. Localized perturbations to the model output and the initial conditions are added to steady time-averaged conditions, to assess the propagation speed and structural evolution of signals away from the local source. Perturbing the model physics by adding a steady tropical heat source results in a classical Matsuno–Gill response near the heating and planetary waves that radiate into the extratropics. A localized disturbance on the winter-averaged North Pacific jet stream produces realistic extratropical cyclones and fronts, including the spontaneous emergence of polar lows. Perturbing the 500-hPa height field alone yields adjustment from a state of rest to one of wind–pressure balance over ~ 6 h. Localized subtropical low pressure systems produce Atlantic hurricanes, provided the initial amplitude exceeds about 4 hPa, and setting the initial humidity to zero eliminates hurricane development. We conclude that the model encodes realistic physics in all experiments and suggest that it can be used as a tool for rapidly testing a wide range of hypotheses.

KEYWORDS: Dynamics; Idealized models; Machine learning

1. Introduction

In the past few years, deep learning (DL) weather prediction models demonstrate forecast skill comparable to those from government operational centers (Bi et al. 2023; Kurth et al. 2023; Lam et al. 2023; Ben Bouallègue et al. 2024). These models are trained on ERA5 analyses and have forecast skill on initial conditions not contained in their training data. In contrast to DL approaches that explicitly enforce physical constraints (e.g., Beucler et al. 2021), it is unclear whether these models have encoded atmospheric physics, such as the dynamics of air motion and propagation of disturbances, or simply patterns that minimize the squared error of the next pattern in a sequence. If these models can be shown to produce physically realistic solutions, they offer an enormous opportunity for testing hypotheses much faster than is currently possible. More importantly, these models may offer a new path to discovery for multiscale problems, where solutions from physics models reflect uncertain parameterizations of poorly resolved processes such as moist convection, small-scale mixing, and surface fluxes.

Physical tests that examine the evolution of spatially localized disturbances are particularly effective in analyzing model physics, since the propagation of signals away from these disturbances is constrained by dynamics. For example, in the

small-amplitude limit, the group velocity in linear wave theory sets the speed of energy dispersion away from a local disturbance. Here, we apply the localized-disturbance approach to the Pangu-Weather model of Bi et al. (2023) using four canonical experiments, one involving perturbations to the model output and the other three involving perturbed initial conditions. The perturbations are applied to climatological time-mean steady states, which are smoother than any individual state that the model was trained on. These experiments are subjectively chosen, and while solutions are not compared directly to identical experiments in a physics-based model, they provide an important plausibility study to motivate such additional experiments. Our hypothesis at the start of this research was that localized features will immediately produce a global response, because no constraint was imposed to prevent this during model training.

In addition to running orders of magnitude faster than physics-based models, these experiments with the Pangu-Weather model are comparatively easy to configure. Performing any one of the experiments described here with a modern physics-based weather model is a significant undertaking, primarily due to complexities associated with model initialization (e.g., Daley 1993; Kalnay 2003). For example, one case we consider here involves idealized extratropical cyclone development over the Pacific Ocean, for which there have been various approaches using idealized basic-state jet streams and initial perturbations. Rotunno et al. (1994) provide an analytical approximation of an idealized jet stream originally shown in Simmons and Hoskins (1975), which is also used in later studies (e.g., Menchaca and Durran 2017). Rotunno et al. (1994) numerically condition the initial perturbation on the linearly most unstable normal mode of the jet stream, whereas

[©] Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/AIES-D-23-0090.s1>.

Corresponding author: Gregory J. Hakim, ghakim@uw.edu

[Menchaca and Durran \(2017\)](#) perturb the jet with a localized disturbance defined by potential vorticity inversion, followed by a time-filtering procedure to remove gravity waves. Recognizing the need for a standardized baroclinic wave test case, [Polvani et al. \(2004\)](#) and [Jablonowski and Williamson \(2006\)](#) developed protocols for such experiments in general circulation models. In particular, [Jablonowski and Williamson \(2006\)](#) provide analytical expressions for the jet stream and initial disturbance and compare solutions for four different models and varying grid resolutions to assess solution convergence. To apply this test to deep learning models, uniform surface boundary conditions are needed for these models, which the current-generation models do not have since they are trained on reanalysis grids. However, similarly configured deep learning models could be directly compared, and one motivation of this study is to provide a starting point for such “unit tests” of the model.

We proceed in [section 2](#) with a description of the experiments and the data used to conduct them. Results for the four experiments described above are presented in [section 3](#). Conclusions are drawn in [section 4](#).

2. Method and experiment design

The Pangu-Weather model uses a vision transformer architecture trained on ERA5 reanalysis data from 1979 to 2017 ([Bi et al. 2023](#)); the trained model weights are publicly available. Model variables consist of global gridded fields of geopotential height, specific humidity of water vapor, temperature, and vector wind components on 13 isobaric levels (1000, 925, 850, 700, 600, 500, 400, 300, 250, 200, 150, 100, and 50 hPa), and surface fields (mean sea level pressure, 2 m air temperature, and 10 m vector wind components). Data reside on the native 0.25° latitude-longitude grid of ERA5. There are four models, which are trained separately for different forecast lead times: 1, 3, 6, and 24 h. [Bi et al. \(2023\)](#) indicate that solutions are most accurate when using the sequence of models with the smallest number of steps to reach a desired lead time (e.g., a 32 h forecast uses the 24 h model, followed by the 6 h model and then two steps of the 1 h model).

Our experiments involve adding perturbations to a steady climatological mean atmosphere. We perform the simulations by solving

$$\mathbf{x}(t+1) = \mathbf{N}[\mathbf{x}(t)] - \mathbf{d}\bar{\mathbf{x}} + \mathbf{f}. \quad (1)$$

Here, \mathbf{x} represents the model state vector, \mathbf{N} represents the Pangu-Weather model, and t represents time indexed according to the version of the model (i.e., $t+1$ means a 1-day forecast when using the 24-h version of the model and a 3-h forecast for the 3-h version). The term \mathbf{f} is a modification to the model output, taken here to be zero for all experiments except steady tropical heating, when it is fixed at a specified value. The term $\mathbf{d}\bar{\mathbf{x}}$ represents the one-step solution of the model that renders the climatological mean atmospheric steady state:

$$\mathbf{d}\bar{\mathbf{x}} = \mathbf{N}(\bar{\mathbf{x}}) - \bar{\mathbf{x}}. \quad (2)$$

We may then take $\bar{\mathbf{x}}$ independent of time. The full state vector, which we send to the Pangu-Weather model, is defined by

$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{x}'$, where \mathbf{x}' are anomalies from the climatological mean state. For $\mathbf{x}' = \mathbf{f} = 0$, (1) with (2) gives $\bar{\mathbf{x}}(t+1) = \mathbf{N}(\bar{\mathbf{x}}) - \mathbf{N}(\bar{\mathbf{x}}) + \bar{\mathbf{x}} = \bar{\mathbf{x}}$; i.e., $\bar{\mathbf{x}}$ is time independent.

Since we are interested in spatially localized perturbations, we define \mathbf{f} and the initial perturbations $\mathbf{x}'(t=0)$ using a function that decays to zero from a local maximum at a specified distance. For this purpose, we use the function defined by Eq. (4.10) in [Gaspari and Cohn \(1999\)](#) and define the distance at which the disturbance reaches zero by L :

$$G(r; L) = \begin{cases} -\frac{1}{4}r^5 + \frac{1}{2}r^4 + \frac{5}{8}r^3 - \frac{5}{3}r^2 + 1 & 0 \leq d \leq L/2 \\ \frac{1}{12}r^5 - \frac{1}{2}r^4 + \frac{5}{8}r^3 + \frac{5}{3}r^2 - 5r + 4 - \frac{21}{3}r & L/2 \leq d \leq L \\ 0 & L \leq d \end{cases} \quad (3)$$

Here, $r = 2d/L$ and d is the distance on the sphere from a central reference point.

For the steady heating experiment, we set \mathbf{f} to be a constant vector with zeros everywhere except for the temperature field within a horizontal region at all levels between 1000 and 200 hPa, where it is set to 0.1 K (day)⁻¹. The region is defined in longitude by (3) with $L = 10\,000$ km and centered at 120°E and in latitude ϕ by $\cos(6\phi)$ within 15° of the equator. The initial condition is given by $\bar{\mathbf{x}}$, which is set to the 0000 UTC 1979–2019 December–February (DJF) ERA5 time average.

For the extratropical cyclone experiment, we define the anomaly field $\mathbf{x}'(t=0)$ by simple linear regression of all fields in the state vector against a standardized time series of DJF 500-hPa geopotential height at the point 40°N, 150°E. The regressed field is then multiplied by (3) with $L = 2000$ km to ensure a spatially localized disturbance and added to the DJF time-mean field. We use the same perturbation initial condition for the geostrophic and hydrostatic adjustment experiments, except we set to zero all variables at all levels except the 500-hPa geopotential height.

For the hurricane experiments, we take the same approach as for the extratropical cyclone experiment, except in this case we use the July–September (JAS) mean state. The disturbance is defined by simple linear regression of all fields in the state vector against a standardized time series of JAS mean sea level pressure at the point 15°N, 40°W. The regressed field is then multiplied by (3) with $L = 1000$ km and added to the JAS time-mean field. For the results in this case, we perform simulations by scaling the perturbation field by a multiplicative constant to vary the strength of the initial low pressure system.

3. Results

a. Steady tropical heating

Tropical heating anomalies associated with long-lived phenomena such as El Niño and the Madden–Julian oscillation produce global changes in circulation. As a result, there is a long history of research to understand the mechanism relating tropical heating to circulation changes. The most well-known

experiments examine the response of a climatological atmospheric state to a steady heating pattern (e.g., Hoskins and Karoly 1981; Sardeshmukh and Hoskins 1988; Jin and Hoskins 1995), which reveal planetary waves radiating into the extratropics along great circle paths. Details of the response depend on the mean state, through the potential vorticity gradients that affect planetary wave propagation, and the structure of the heating field (Ting and Sardeshmukh 1993). A representative example showing the effect of the location of the heating on the wave response, and sensitivity to details in the mean state, is found in Fig. 7 of Ting and Sardeshmukh (1993).

The Pangu-Weather response to weak DJF tropical heating (0.1 K day^{-1}) shows a small 500-hPa height increase over the heating region after 5 days, with a negative anomaly to the north (Fig. 1a). The extratropical wave train extends downstream and increases in amplitude during days 5–20, with maximum anomalies over 100 m at day 20 (Figs. 1b,c). A wave train appears in both hemispheres, with larger amplitude in the Northern (wintertime) Hemisphere, which has the stronger waveguide for stationary waves. This response is qualitatively similar to classical results (e.g., Hoskins and Karoly 1981; Sardeshmukh and Hoskins 1988), with differences in details dependent on the location, shape, and temporal structure of the heating, seasonality, and other factors.

A closer examination of the response in the lower troposphere near the heating reveals a pattern similar to the classical Matsuno–Gill (Matsuno 1966; Gill 1980) response to steady tropical heating (Fig. 2). Along the equator, wind anomalies are convergent toward the western end of the heating region. This signature is associated with a Kelvin-wave response to the heating. Off the equator, the western end of the heating is flanked by cyclonic gyres in both hemispheres, which are associated with a mixed-Rossby–gravity wave response. Unlike idealized experiments, typically using shallow-water equations, these solutions are influenced by surface boundary conditions, so that there are flow distortions over the Maritime Continent in particular and myriad multiscale moist processes involving clouds and convection.

This experiment suggests that the Pangu-Weather model responds qualitatively, if not quantitatively, consistent with idealized experiments for tropical heating. Anomalies emerge smoothly and locally from the heat source and increase in amplitude with time as a nearly stationary wave response. Idealizing the problem further to the zonal-mean DJF basic state produces a similar response, with a wave train extending across the North Pacific to North America (Fig. S1 in the online supplemental material), but with differences in phase and amplitude related to the basic state on which the waves propagate. The Southern Hemisphere response is also notably weaker for the zonal-mean state, which may be related to a weaker extratropical waveguide (smaller potential vorticity gradients) in summer.

b. Extratropical cyclone development

The next experiment considers the time evolution of a localized 500-hPa trough at the western end of the North Pacific storm track (Fig. 3a), which is the canonical initial condition

preceding surface cyclogenesis (e.g., Gyakum and Danielson 2000; Hakim 2003; Yoshida and Asuma 2004). Observations and idealized modeling results show the development of a localized extratropical cyclone, with subsequent cyclones appearing downstream as the disturbance evolves into a spreading wave packet (e.g., Hakim 2003; Jablonowski and Williamson 2006).

Results for Pangu-Weather show that after 2 days the trough has progressed to the central Pacific and begun to disperse, with the appearance of anticyclonic circulations both upstream and downstream (Fig. 3b). A surface cyclone develops to the east of the upper trough, with a smaller-scale secondary cyclone appearing upstream (Fig. 4b). By day 4, the upper trough has amplified and spread into a wave packet, with the leading edge along western North America (Fig. 3d), and a surface cyclone nearly coincident with the upper trough (Fig. 4d). Vertical alignment of extratropical cyclones is the hallmark of a developing cyclone that has reached the occluded phase of the life cycle. In contrast, the upstream surface cyclone remains downstream of the 500-hPa trough and continues to deepen past day 4. A second upstream cyclone appears at day 4 west of the date line. These cyclones are accompanied by temperature anomalies having the largest horizontal gradients near the surface cold front (Fig. S2).

All aspects of this idealized baroclinic development are consistent with observations and modeling (e.g., Jablonowski and Williamson 2006) of localized extratropical cyclone development. In particular, disturbances at the upstream end of the storm track produce a baroclinic wave packet (Simmons and Hoskins 1979) that disperses and moves downstream at the group velocity (faster than the phase of individual troughs). As we find here, these solutions also show both upstream surface development and downstream upper-level development (Simmons and Hoskins 1979; Chang 1993; Hakim 2003). Moreover, the upstream surface development we observe here has a relatively smaller spatial scale, resembling a “polar low,” which is frequently observed in winter over the North Pacific (e.g., Mullen 1983; Rasmusson 2003). Curiously, these polar lows appear first at the surface and have a warm core, suggestive of the importance of surface fluxes due to cold air moving over relatively warmer water (Emanuel and Rotunno 1989).

Idealizing the problem further to the zonal-mean DJF atmosphere produces a similar response, with a wave packet that spreads downstream toward Europe by day 10 (Fig. S3). Furthermore, repeating the experiment, but for summer conditions (JAS time mean), shows much weaker cyclone development and an absence of polar lows (not shown). We conclude that Pangu-Weather appears to have implicitly encoded the seasonally varying physical processes of oceanic extratropical cyclone development in the neural network weights that govern the dynamical evolution of its prognostic variables.

c. Geostrophic and hydrostatic adjustment

Here, we test an initial perturbation similar to the extratropical cyclone case, except that it is localized completely to

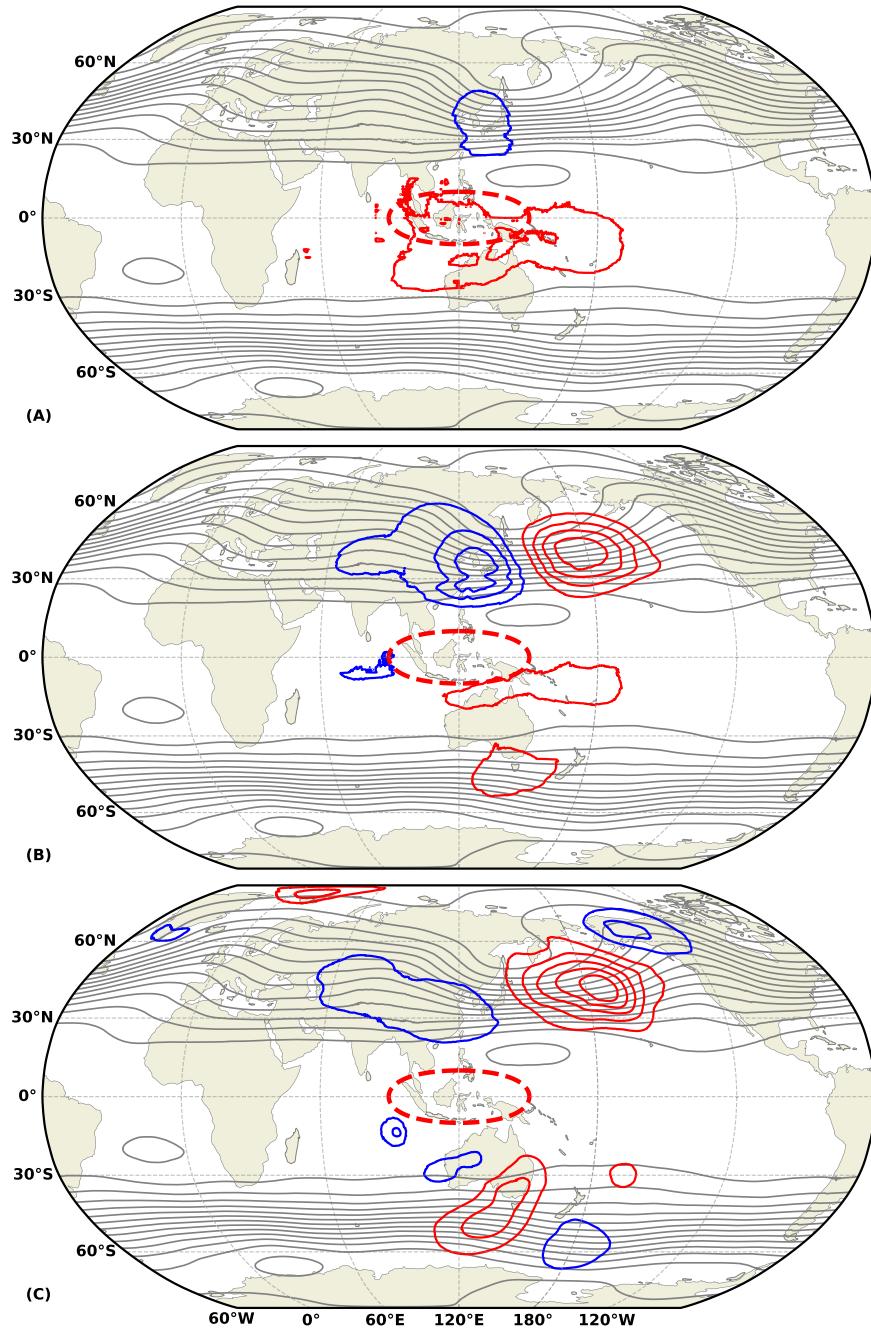


FIG. 1. Response in DJF 500-hPa geopotential height to steady tropical heating of 0.1 K day^{-1} within the region outlined by the dashed red line. The DJF-averaged geopotential height is shown by gray lines every 60 m, and anomalies are shown by red (positive) and blue (negative) lines; the zero contour is suppressed. Solutions are shown for (a) 5 days (contours every 0.3 m); (b) 10 days (contours every 2 m); and (c) 20 days (contours every 20 m).

the 500-hPa field; it does not extend in the vertical, and every other field has zero anomaly. We note that this initial condition lies outside all aspects of the ERA5 training data for the Pangu-Weather model, since ERA5 is produced with a hydrostatic model. This type of initial condition is unbalanced since there are no wind or temperature anomalies, whereas outside

the deep tropics one commonly finds the wind flowing along the height contours (as evident in Fig. 3a). The classic textbook example is the Rossby adjustment problem, consisting of an initially stationary shallow layer of water with a jump in the free surface, which evolves to a state of geostrophic balance (e.g., Holton and Hakim 2013, section 5.6). This is a

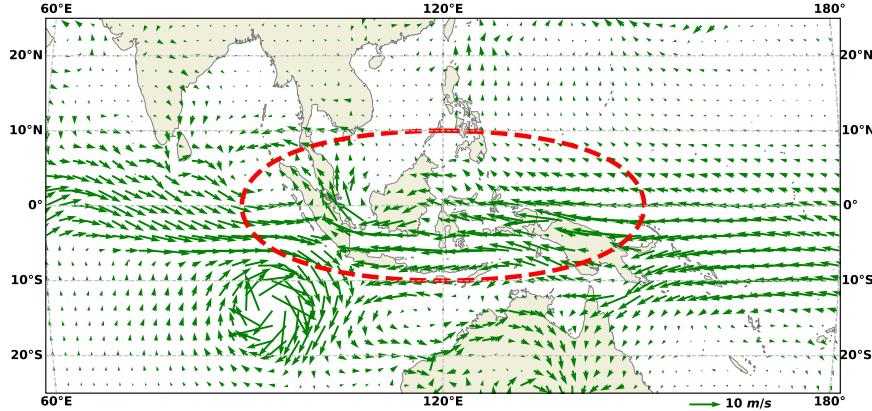


FIG. 2. The 850-hPa anomaly wind vectors for the steady heating experiment after 20 days. The red dashed line outlines the region of steady heating.

particularly hard test and one that likely cannot be performed without additional modification using a physics-based model, since unbalanced initial conditions produce rapid oscillations that are difficult to resolve. An example showing the initial response of a localized anomaly analogous to the one considered here (but of opposite sign) is shown in Lelong and Sundermeyer (2005, their Fig. 5). The initial response is divergent flow, which under the action of Coriolis turning rotates over an inertial period to a state of wind–pressure balance.

Since we are interested in the initial response at short time scales, we use the 1-, 3-, and 6-h versions of the Pangu-Weather model here. Results show that, after 1 h, the wind accelerates from rest in the initial conditions to about 5 m s^{-1} and is convergent on the area of low geopotential height (Fig. 5a). The center of convergence is to the west of the lowest height, which increases to -89 m from -100 m in the initial condition. At 3 h, the wind accelerates to a maximum of about 10 m s^{-1} and remains convergent on the area of low height, for which the minimum has increased to -74 m (Fig. 5b). The wind direction has turned clockwise at all locations compared to the 1-h solution, as one expects from the Coriolis turning of the accelerating wind in the direction of the pressure gradient force. At 6 h, the wind direction has continued to rotate clockwise such that it is nearly parallel to the geopotential height contours everywhere, reflecting a closer balance between the wind and geopotential height fields (Fig. 5c). The height minimum has increased to -58 m , reflecting a conversion of available potential energy to kinetic energy.

A quarter turn of a Foucault pendulum at 40°N takes $\sim 9 \text{ h}$, so the adjustment in the wind field indicated by the Pangu-Weather solution is consistent with physical expectations. Once again, we conclude that the solution for this idealized initial-value problem is qualitatively, if not quantitatively, consistent with the expected dynamics.

Repeating the experiment, except for an initial disturbance on the equator, produces a notably different response (Fig. S4). The velocity field is again convergent on the area of low geopotential height, except in this case convergence is directed on the center of the low. The difference may be due to the basic-state jet stream in the previous case, with fast westerly winds and a

strong meridional potential vorticity gradient that promotes westward Rossby-wave propagation. Another notable aspect of the equatorial case is the slower Coriolis turning of the wind and the fact that the model has learned about the asymmetry in this turning about the equator. An analysis of the time difference in the anomalous zonal wind on the equator reveals signals that propagate in both directions at around 20 m s^{-1} (Fig. S5), typical of tropical gravity waves. A weaker signal is also evident at the speed of sound (dashed black lines). Finally, we note that this figure shows the incompatibility between the different versions of the Pangu-Weather model, with abrupt differences in the time tendency at intervals of 3, 6, and 24 h. These single-step “shocks” do not appear to adversely affect the solution at subsequent times but will affect temporal diagnostic calculations that span several time steps of the model.

d. Atlantic hurricane development

The last example concerns the evolution of a localized disturbance in the subtropics for the JAS averaged conditions. Seeds of Atlantic hurricanes take the form of weak low pressure systems, which may develop into mature storms given the right environmental conditions. Finite-amplitude disturbances are thought to be needed to reduce the time to development while the storm is in a favorable environment (e.g., McBride and Zehr 1981; Rotunno and Emanuel 1987; Nolan et al. 2007). Here, we perform experiments for a localized area of low pressure at a reference location ($15^\circ\text{N}, 40^\circ\text{W}$) and vary the initial amplitude. The three-dimensional perturbation is constructed similarly to the initial condition for the extratropical cyclone case, by regressing all variables and locations onto the mean sea level pressure at the reference location.

Results show that the low pressure systems take a familiar track toward the northwest around the climatological subtropical area of high pressure (Fig. 6). Stronger initial conditions take a progressively northward track, which is consistent with the known physical basis due to the increasing amplitude of azimuthal wavenumber-1 asymmetries (“ β gyres”). Although Pangu-Weather may at best poorly resolve these features, the

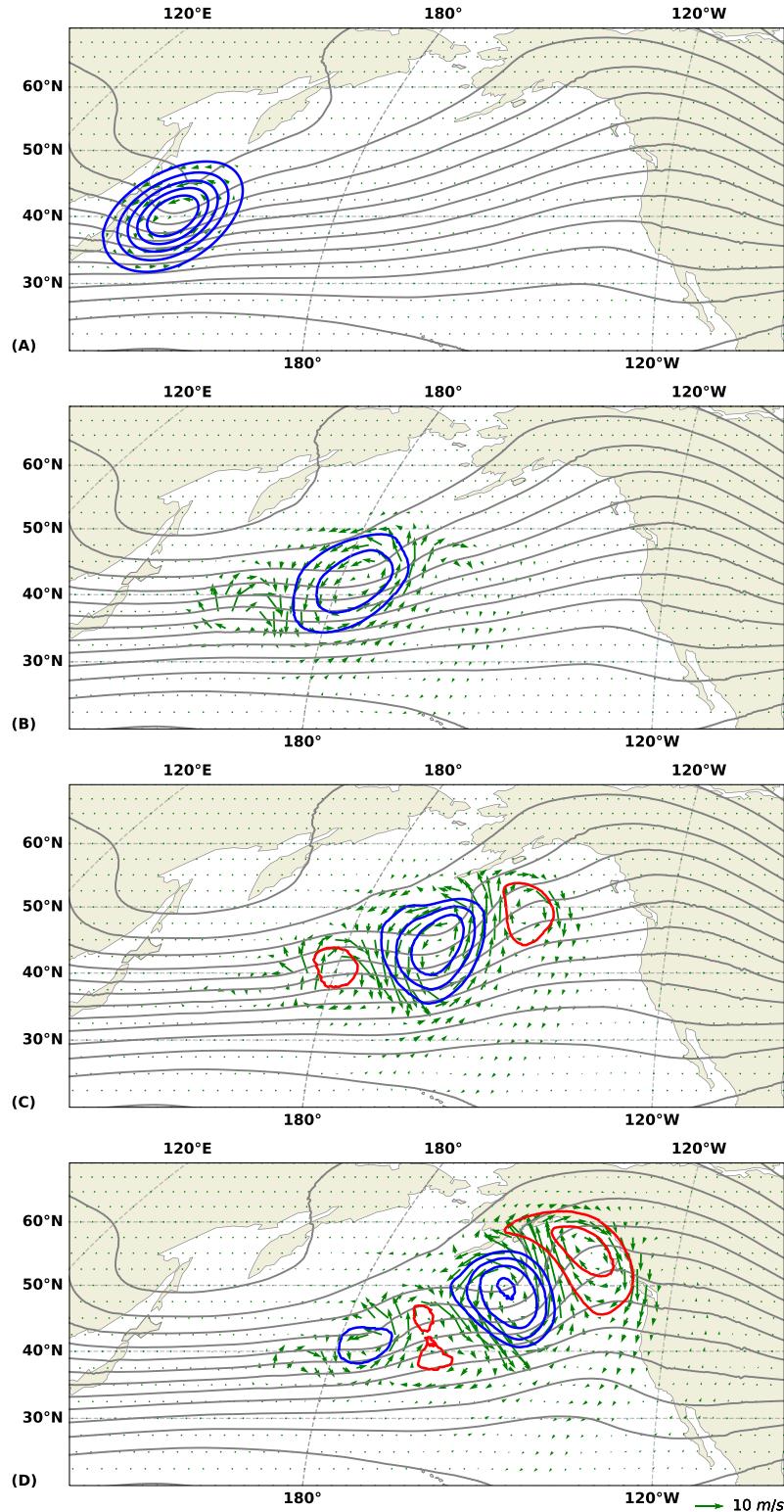


FIG. 3. Solution at 500 hPa for a localized disturbance on the DJF atmosphere. The full geopotential height is shown by gray lines every 60 m, and anomalies from the DJF average are shown by red (positive) and blue (negative) lines every 20 m; the zero contour is suppressed. Green arrows show the anomalous vector wind. Solutions are shown at (a) 0 (the specified initial condition); (b) 2; (c) 3; and (d) 4 days.

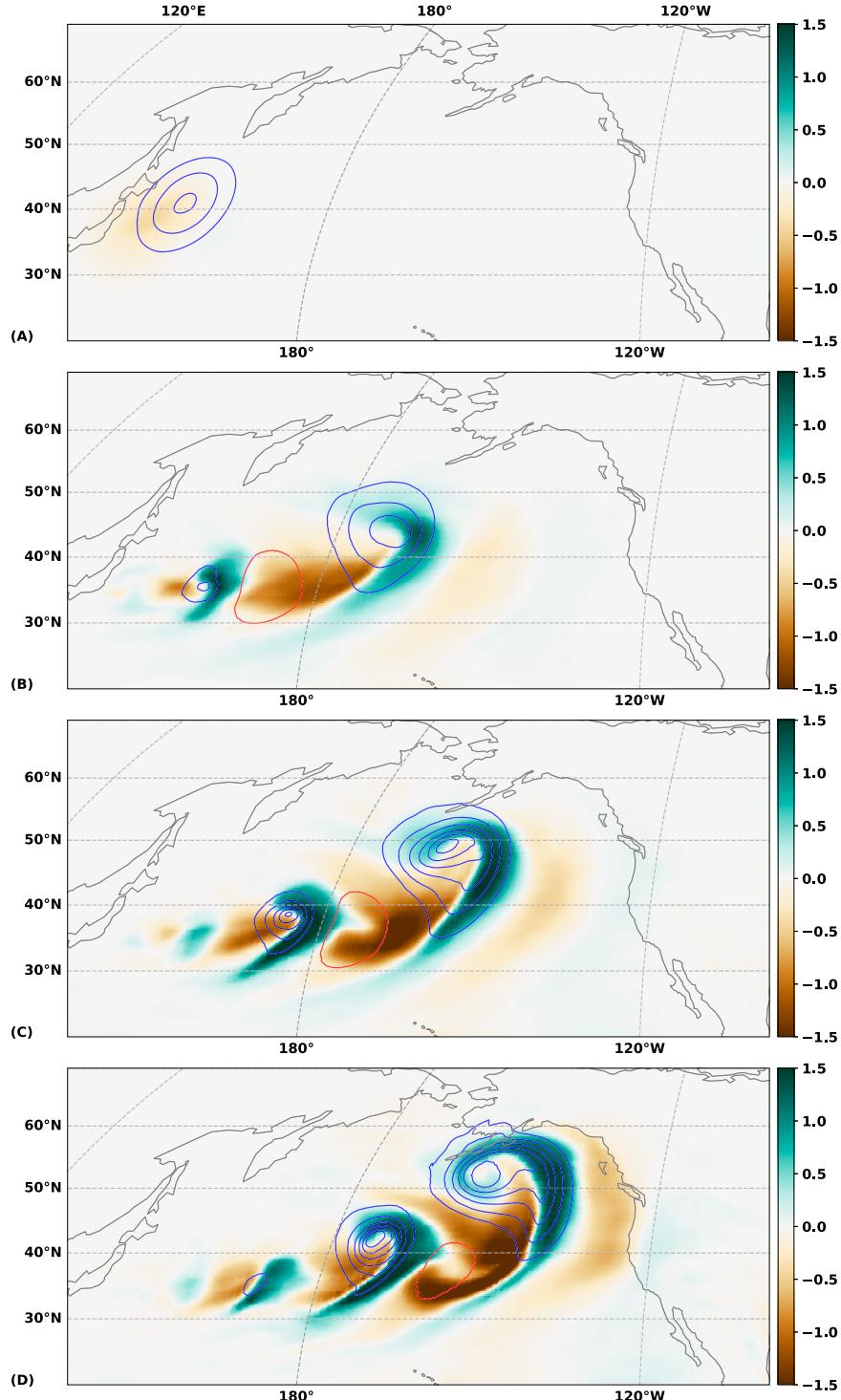


FIG. 4. Surface cyclones associated with the solution in Fig. 3. Anomalies in mean sea level pressure are shown every 2 hPa, with red (blue) lines for positive (negative) values; the zero contour is suppressed. Water vapor specific humidity anomalies (g kg^{-1}) at 850 hPa are shaded. Solutions are shown at (a) 0 (the specified initial condition); (b) 2; (c) 3; and (d) 4 days.

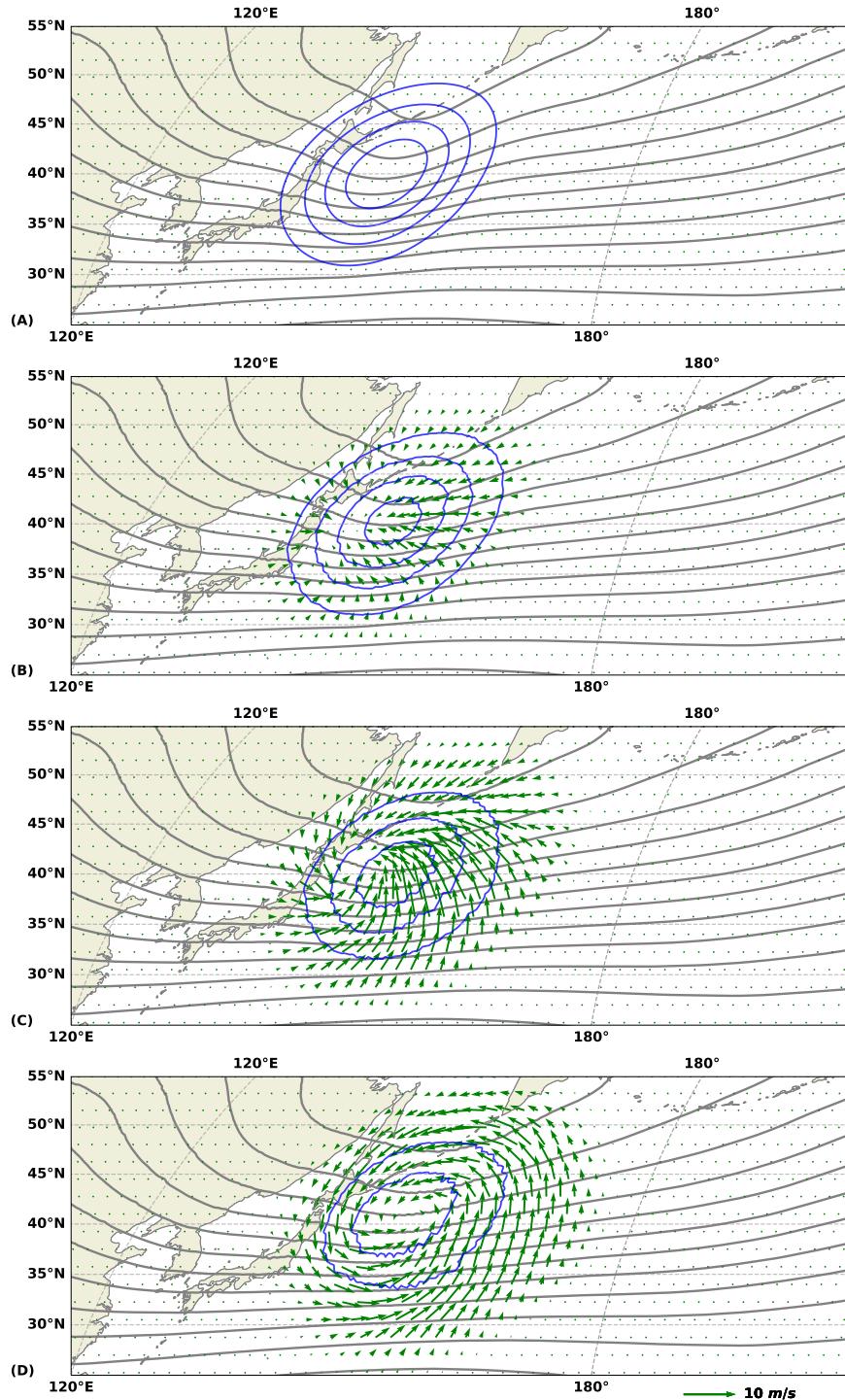


FIG. 5. Solution at 500 hPa for the geostrophic adjustment problem consisting of a localized geopotential height disturbance on the DJF-averaged atmosphere. The full geopotential height is shown by gray lines every 60 m, and negative anomalies are shown by blue lines every 20 m; the zero contour is suppressed. Green arrows show the anomalous vector wind. Solutions are shown at (a) $t = 0$ (the specified initial condition); (b) 1; (c) 3; and (d) 6 h.

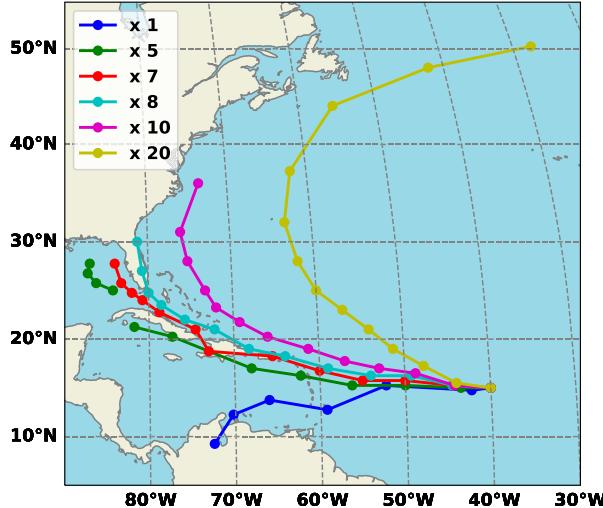


FIG. 6. Tracks of mean sea level pressure minima for experiments seeding Atlantic hurricanes on the July–September-averaged atmosphere. All experiments are initialized with a surface-based low pressure system at 15°N, 40°W, and initial amplitude by a scaling factor on the climatological JAS standard deviation at that location (indicated in the legend).

neural network has identified this physical relationship between the strength of tropical cyclones and a northward track.

For initial disturbances with anomalous mean sea level pressure less than about ~ 4 hPa, the storms do not intensify, whereas initial disturbances stronger than this rapidly intensify (Fig. 7). An additional experiment for the $10\times$ disturbance was performed by setting the water vapor specific humidity to zero, and unlike the original case that rapidly develops, the dry system rapidly decays. Pangu-Weather does not explicitly model condensational heating, but the model has the conditional association between water vapor content and the development of tropical cyclones.

4. Conclusions

We have tested the Pangu-Weather deep learning weather prediction model on a set of four canonical experiments aimed at probing its dynamical response to local perturbations. These perturbations are helpful for determining whether disturbances evolve and propagate in a physically meaningful manner. Our hypothesis at the outset of this work was that these localized features would immediately produce a global response because there is no constraint to prevent this during model training. The fact that each experiment produced signal propagation and structural evolution qualitatively in accord with previous research in meteorology suggests that the model has encoded realistic physics. While we do not make a direct comparison to solutions from a physics-based model, the results here provide a proof of concept motivating such experiments. We note that, due to differences in numerics, boundary conditions, and parameterizations for unresolved scales and processes, solutions from physics-based models for these experiments will differ in details, and it

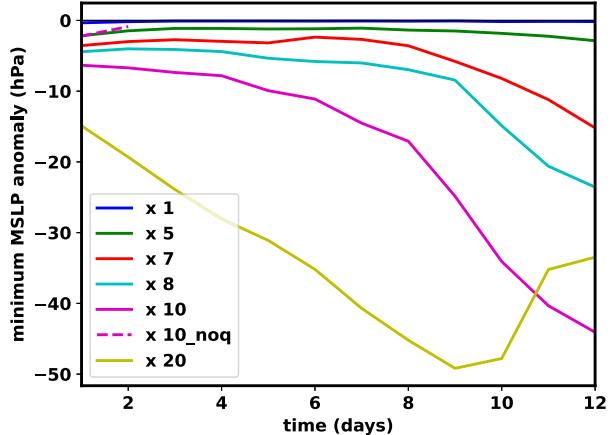


FIG. 7. Intensity of the low pressure systems tracked in Fig. 6 in terms of anomalous mean sea level pressure (hPa) as a function of time (days).

would be interesting to see whether the Pangu-Weather solutions fall within the uncertainty of the physics-based models. Perhaps equally useful is an application of the experiments presented here for comparing deep learning models. Since many models are trained on the same reanalysis data, issues related to surface boundary conditions are less important, facilitating model comparisons for what may be regarded as unit tests.

Results from the canonical experiments show qualitative, if not quantitative, agreement with studies of similar phenomena in observations and numerical simulations. This agreement ranges from hourly time scales for the geostrophic adjustment process to approximately steady features beyond 10 days associated with stationary tropical heating. Highlights from these experiments, summarized in Table 1, include a Matsuno-Gill response and extratropical planetary wave response to steady tropical heating; baroclinic wave packet emergence and polar low development in the cold air mass associated with a North Pacific extratropical cyclone; divergent flow yielding to rotational flow for an unbalanced initial condition; and the importance of initial-vortex amplitude and water vapor in the development and track of Atlantic hurricanes.

We conclude that the Pangu-Weather model encodes realistic physics for the experiments considered here, motivating future basic research using this tool. Several attributes make this model particularly powerful for atmospheric dynamics and scientific hypothesis testing. First, the simulations are computationally inexpensive compared to traditional global weather models. This enables large ensembles, including iterations over varying parameters, initial conditions, and perturbations to model output. Second, experiments are extremely easy to configure, and the model is very forgiving in aspects that physics models are not. For example, initial imbalances in physics-based models can produce spurious oscillations at the model time step that are difficult to remove or filter without affecting the resolved scales of interest. Therefore, we speculate that models like Pangu-Weather might be particularly useful for the rapid evaluation of hypotheses, allowing

TABLE 1. Summary of subjective evaluation of the physics experiments. Qualitative summary evaluation labels are physically “consistent,” “inconsistent,” or “mixed.” The remaining columns summarize notable strengths and weaknesses.

Experiment	Evaluation	Strength	Weakness
Tropical heating	Consistent	Teleconnections; Matsuno–Gill response	Weak remote response
Baroclinic development	Consistent	Cyclogenesis; downstream development	None
Geostrophic adjustment	Mixed	Divergent wind; Coriolis turning	Shocks from different models
Tropical cyclone	Consistent	Development sensitive to moisture	Poleward turning too strong?

tests over a wide range of ideas to quickly narrow the scope of investigation for experiments using expensive physics-based models. Among many possibilities, one particularly interesting path of research employs deep learning models to examine multiscale phenomena involving convective clouds, such as the Madden–Julian oscillation, where physics-based models and theory have not yet approximated the essential physical processes.

Acknowledgments. We thank Steve Penny for conversations related to deep learning models in the geosciences and Mike Pritchard for comments on an earlier draft of the manuscript. Comments from three anonymous reviewers and editor David John Gagne were helpful in improving the clarity of the manuscript. GJH acknowledges support for this research from NSF Award 2202526 and Heising-Simons Foundation Award 2023-4715.

Data availability statement. Code and data to reproduce results in this paper can be found at this GitHub repository: <https://github.com/modons/DL-weather-dynamics>.

REFERENCES

- Ben Bouallègue, Z., and Coauthors, 2024: The rise of data-driven weather forecasting: A first statistical assessment of machine learning-based weather forecasts in an operational-like context. *Bull. Amer. Meteor. Soc.*, **105**, E864–E883, <https://doi.org/10.1175/BAMS-D-23-0162.1>.
- Beucler, T., M. Pritchard, S. Rasp, J. Ott, P. Baldi, and P. Gentine, 2021: Enforcing analytic constraints in neural networks emulating physical systems. *Phys. Rev. Lett.*, **126**, 098302, <https://doi.org/10.1103/PhysRevLett.126.098302>.
- Bi, K., L. Xie, H. Zhang, X. Chen, X. Gu, and Q. Tian, 2023: Accurate medium-range global weather forecasting with 3D neural networks. *Nature*, **619**, 533–538, <https://doi.org/10.1038/s41586-023-06185-3>.
- Chang, E. K. M., 1993: Downstream development of baroclinic waves as inferred from regression analysis. *J. Atmos. Sci.*, **50**, 2038–2053, [https://doi.org/10.1175/1520-0469\(1993\)050<2038:DDOBWA>2.0.CO;2](https://doi.org/10.1175/1520-0469(1993)050<2038:DDOBWA>2.0.CO;2).
- Daley, R., 1993: *Atmospheric Data Analysis*. Cambridge University Press, 472 pp.
- Emanuel, K. A., and R. Rotunno, 1989: Polar lows as Arctic hurricanes. *Tellus*, **41A** (1), 1–17, <https://doi.org/10.3402/tellusa.v41i1.11817>.
- Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723–757, <https://doi.org/10.1002/qj.49712555417>.
- Gill, A. E., 1980: Some simple solutions for heat-induced tropical circulation. *Quart. J. Roy. Meteor. Soc.*, **106**, 447–462, <https://doi.org/10.1002/qj.49710644905>.
- Gyakum, J. R., and R. E. Danielson, 2000: Analysis of meteorological precursors to ordinary and explosive cyclogenesis in the western North Pacific. *Mon. Wea. Rev.*, **128**, 851–863, [https://doi.org/10.1175/1520-0493\(2000\)128<851:AOMPTO>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128<851:AOMPTO>2.0.CO;2).
- Hakim, G. J., 2003: Developing wave packets in the North Pacific storm track. *Mon. Wea. Rev.*, **131**, 2824–2837, [https://doi.org/10.1175/1520-0493\(2003\)131<2824:DWPITN>2.0.CO;2](https://doi.org/10.1175/1520-0493(2003)131<2824:DWPITN>2.0.CO;2).
- Holton, J. R., and G. J. Hakim, 2013: *An Introduction to Dynamic Meteorology*. Academic Press, 552 pp.
- Hoskins, B. J., and D. J. Karoly, 1981: The steady linear response of a spherical atmosphere to thermal and orographic forcing. *J. Atmos. Sci.*, **38**, 1179–1196, [https://doi.org/10.1175/1520-0469\(1981\)038<1179:TSLROA>2.0.CO;2](https://doi.org/10.1175/1520-0469(1981)038<1179:TSLROA>2.0.CO;2).
- Jablonowski, C., and D. L. Williamson, 2006: A baroclinic instability test case for atmospheric model dynamical cores. *Quart. J. Roy. Meteor. Soc.*, **132**, 2943–2975, <https://doi.org/10.1256/qj.06.12>.
- Jin, F., and B. J. Hoskins, 1995: The direct response to tropical heating in a baroclinic atmosphere. *J. Atmos. Sci.*, **52**, 307–319, [https://doi.org/10.1175/1520-0469\(1995\)052<0307:TDRTTH>2.0.CO;2](https://doi.org/10.1175/1520-0469(1995)052<0307:TDRTTH>2.0.CO;2).
- Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, 341 pp.
- Kurth, T., and Coauthors, 2023: FourCastNet: Accelerating global high-resolution weather forecasting using adaptive Fourier neural operators. *Proc. of the Platform for Advanced Scientific Computing Conf.*, Davos, Switzerland, Association for Computing Machinery, 1–11, <https://doi.org/10.1145/3592979.3593412>.
- Lam, R., and Coauthors, 2023: Learning skillful medium-range global weather forecasting. *Science*, **382**, 1416–1421, <https://doi.org/10.1126/science.adl2336>.
- Lelong, M.-P., and M. A. Sundermeyer, 2005: Geostrophic adjustment of an isolated diapycnal mixing event and its implications for small-scale lateral dispersion. *J. Phys. Oceanogr.*, **35**, 2352–2367, <https://doi.org/10.1175/JPO2835.1>.
- Matsuno, T., 1966: Quasi-geostrophic motions in the equatorial area. *J. Meteor. Soc. Japan*, **44**, 25–43, https://doi.org/10.2151/jmsj1965.44.1_25.
- McBride, J. L., and R. Zehr, 1981: Observational analysis of tropical cyclone formation. Part II: Comparison of non-developing versus developing systems. *J. Atmos. Sci.*, **38**, 1132–1151, [https://doi.org/10.1175/1520-0469\(1981\)038<1132:OAOTCF>2.0.CO;2](https://doi.org/10.1175/1520-0469(1981)038<1132:OAOTCF>2.0.CO;2).
- Menchaca, M. Q., and D. R. Durran, 2017: Mountain waves, downslope winds, and low-level blocking forced by a midlatitude cyclone encountering an isolated ridge. *J. Atmos. Sci.*, **74**, 617–639, <https://doi.org/10.1175/JAS-D-16-0092.1>.
- Mullen, S. L., 1983: Explosive cyclogenesis associated with cyclones in polar air streams. *Mon. Wea. Rev.*, **111**, 1537–1543,

- [https://doi.org/10.1175/1520-0493\(1983\)111<1537:ECAWCI>2.0.CO;2](https://doi.org/10.1175/1520-0493(1983)111<1537:ECAWCI>2.0.CO;2).
- Nolan, D. S., E. D. Rappin, and K. A. Emanuel, 2007: Tropical cyclogenesis sensitivity to environmental parameters in radiative–convective equilibrium. *Quart. J. Roy. Meteor. Soc.*, **133**, 2085–2107, <https://doi.org/10.1002/qj.170>.
- Polvani, L. M., R. K. Scott, and S. J. Thomas, 2004: Numerically converged solutions of the global primitive equations for testing the dynamical core of atmospheric GCMs. *Mon. Wea. Rev.*, **132**, 2539–2552, <https://doi.org/10.1175/MWR2788.1>.
- Rasmussen, E. A., 2003: Polar lows. *A Half Century of Progress in Meteorology: A Tribute to Richard Reed*, Springer, 61–78.
- Rotunno, R., and K. A. Emanuel, 1987: An air-sea interaction theory for tropical cyclones. Part II: Evolutionary study using a nonhydrostatic axisymmetric numerical model. *J. Atmos. Sci.*, **44**, 542–561, [https://doi.org/10.1175/1520-0469\(1987\)044<0542:AAITFT>2.0.CO;2](https://doi.org/10.1175/1520-0469(1987)044<0542:AAITFT>2.0.CO;2).
- , W. C. Skamarock, and C. Snyder, 1994: An analysis of frontogenesis in numerical simulations of baroclinic waves. *J. Atmos. Sci.*, **51**, 3373–3398, [https://doi.org/10.1175/1520-0469\(1994\)051<3373:AAOFIN>2.0.CO;2](https://doi.org/10.1175/1520-0469(1994)051<3373:AAOFIN>2.0.CO;2).
- Sardeshmukh, P. D., and B. J. Hoskins, 1988: The generation of global rotational flow by steady idealized tropical divergence. *J. Atmos. Sci.*, **45**, 1228–1251, [https://doi.org/10.1175/1520-0469\(1988\)045<1228:TGOGRF>2.0.CO;2](https://doi.org/10.1175/1520-0469(1988)045<1228:TGOGRF>2.0.CO;2).
- Simmons, A. J., and B. J. Hoskins, 1975: A comparison of spectral and finite-difference simulations of a growing baroclinic wave. *Quart. J. Roy. Meteor. Soc.*, **101**, 551–565, <https://doi.org/10.1002/qj.49710142912>.
- , and —, 1979: The downstream and upstream development of unstable baroclinic waves. *J. Atmos. Sci.*, **36**, 1239–1254, [https://doi.org/10.1175/1520-0469\(1979\)036<1239:TDAUDO>2.0.CO;2](https://doi.org/10.1175/1520-0469(1979)036<1239:TDAUDO>2.0.CO;2).
- Ting, M., and P. D. Sardeshmukh, 1993: Factors determining the extratropical response to equatorial diabatic heating anomalies. *J. Atmos. Sci.*, **50**, 907–918, [https://doi.org/10.1175/1520-0469\(1993\)050<0907:FDETERT>2.0.CO;2](https://doi.org/10.1175/1520-0469(1993)050<0907:FDETERT>2.0.CO;2).
- Yoshida, A., and Y. Asuma, 2004: Structures and environment of explosively developing extratropical cyclones in the northwestern Pacific region. *Mon. Wea. Rev.*, **132**, 1121–1142, [https://doi.org/10.1175/1520-0493\(2004\)132<1121:SAEOED>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<1121:SAEOED>2.0.CO;2).