Instabilities in the Shallow-Water System with a Semi-Lagrangian, Time-Centered Discretization

CHRISTOPHER SUBICH[a]

[a] *Environment and Climate Change Canada, Recherce en Prévision Numérique, Dorval, Quebec, Canada*

(Manuscript received 18 March 2021, accepted 26 November 2021)

ABSTRACT: Conventional wisdom suggests that the combination of semi-Lagrangian advection and an implicit treatment of gravity wave terms should result in a combined scheme for the shallow-water equations stable for high Courant numbers. This wisdom is well justified by linear analysis of the system about a uniform reference state with constant fluid depth and velocity, but it is only assumed to hold true in more complex scenarios. This work finds that this conventional wisdom no longer holds in more complicated flow regimes, in particular when the background state is given by steady-state flow past topography. Instead, this background state admits a wide range of instabilities that can lead to noise in atmospheric forecasts.

SIGNIFICANCE STATEMENT: This work shows that solutions to the shallow-water equations with a semi-Lagrangian treatment of advection and an implicit, time-centered treatment of gravity wave terms can be unstable when there is a background state of flow over topography. This basic algorithm is used by many operational weather-forecasting models to simulate the meteorological equations, and showing an instability in the simplified, shallow-water system suggests that a similar mechanism may be responsible for "noise" in operational weather forecasts under some circumstances. If this problem can be addressed, it could allow numerical weather models to operate with less dissipation, improving forecast quality.

KEYWORDS: Model errors; Numerical analysis/modeling; Numerical weather prediction/forecasting

1. Introduction

Developers of atmospheric dynamical cores regularly struggle against time step restrictions. The full Euler system of equations, which describe the terrestrial atmosphere below the rarefied exosphere (Izakov 1971), admits motions due to advection, gravity waves, and sound waves. Dynamical cores with local, explicit time-stepping find their maximum stable time steps limited by a Courant number restriction, whereby the fastest-propagating waves must travel no more than $O(1)$ grid lengths per time step. For nearly five decades (Robert et al. 1972), researchers have sought to extend the usable range of time steps, both through approximations to the Euler equation [Robert et al. (1972) using a hydrostatic approximation, others such as Arakawa and Konor (2009) and Dubos and Voitus (2014) removing sound waves through alternative formulations of the Euler equations including anelastic and pseudoincompressible forms] and through implicit treatment of some or all terms of the equation.

The semi-Lagrangian method is a popular treatment for the advective terms in the Euler equations. This method treats the material derivative nonlocally by integrating it in a Lagrangian frame, deriving "departure-point" values for fluid parcels that arrive on the designated computational grid at the end of each time step. First applied to weather modeling by Robert (1981) [with a historical review given by Staniforth and Côté (1991)], semi-Lagrangian advection further extended the admissible, stable time step size over the next-best algorithms by (often) allowing simulations with advective Courant numbers much larger than one. The approach quickly became popular, and it remains in use today at the heart of many operational atmospheric forecasting models (Girard et al. 2014; Wood et al. 2014; ECMWF 2020) in use today.

Despite the method's popularity, it can also be the source of computational "noise" in long simulations. The most well-known source of such noise is spurious orographic resonance, first described by Rivest et al. (1994). Here, in the case of subcritical flow over topography, the advection terms are treated accurately via a semi-Lagrangian approach, but a time-centered implicit treatment of the gravity wave terms slows down their phase speed, resulting in $c \to u_0$ in the hypothetical limit of an arbitrarily large time step. As a result, gravity waves that would ordinarily propagate upstream ($u_0 - c < 0$) can become trapped over topography, leading to resonance and inaccuracies. McDonald (1998) discusses additional sources of noise, including noise from some formulations of trajectory calculations, model-top boundary conditions, and physical parameterizations.

Despite these descriptions, there is no exhaustive list of "noise" sources in implicit, semi-Lagrangian simulations. Husain et al. (2020) has recently reported anomalous "valley warm pools" arising from flow over topography, and the anomaly is not adequately explained by the above phenomena.

Denotes content that is immediately available upon publication as open access.

Corresponding author: Christopher Subich, Christopher.Subich@ec.gc.ca

Publisher's Note: This article was revised on 16 September 2024 to remove the AMS copyright statement, which was mistakenly applied when originally published.

DOI: 10.1175/MWR-D-21-0054.1

Topography-mediated effects such as this may be more common in high-resolution simulations over mountainous regions, where the topographic signal can have both a large amplitude and short wavelength.

Although some sources of noise have been addressed individually, the operational GEM model (Girard et al. 2014; Husain et al. 2019) still sees a stubborn need for relatively large damping (Husain and Girard 2017) to address forecast noise. This is particularly concerning for emerging higher-order model discretizations, where overly diffusive or dissipative "fixes" can eat away at any performance advantage from an improved numerical foundation.

In this work, we further expand on the literature of potential instabilities and "noise" in semi-Lagrangian, time-centered simulations of the Euler equations by focusing on a type of secondary instability in the context of the shallow-water equations. In this family of instabilities, flow over topography first induces a steady-state response, but linearizing the discretized equations of motion about this background state reveal a rich field of instabilities. The largest of these instabilities can extract energy from the background flow into spurious modes over a wide range of topography scales and Courant numbers.

Overview

In section 2, we first discuss the shallow-water formulation used in this work, beginning with the continuous equations and ending with the linearized, discretized equations used to explore the system's instabilities. Section 3 discusses the instabilities in more detail, deriving necessary conditions based on the dispersion relation of the linearized shallow-water equations. Section 4 discusses potential mitigation strategies, and we conclude in section 5 with a discussion of the implications for future analysis and model development.

The numerical development in this work is intended to closely follow the discretizations and approximations made in the GEM model.

2. Shallow-water equations

a. Time-continuous equations

The one-dimensional shallow-water equations, which are as follows:

$$\phi_t + (u\phi)_x = 0 \quad \text{and} \tag{1a}$$

$$u_t + \left[\frac{1}{2}u^2 + g(\phi + H)\right]_x = 0, \tag{1b}$$

[where $(\cdot)_t$ and $(\cdot)_x$ denote partial derivatives in time and space, respectively] for a thin fluid with velocity u and layer depth ϕ above an orography of $H(x)$, serve as the model problem for this work. The linearized shallow-water equations (about a uniform background flow) arise from the vertical mode decomposition of the atmospheric Euler equations (Gill 1982, p. 107). The nonlinear form presented here is often used as a testing ground for discretizations intended for dynamical cores (e.g., Williamson et al. 1992; Ullrich et al. 2014; Thuburn and Cotter 2012), including explanations of instabilities in these discretizations (Peixoto et al. 2018; Bell et al. 2017).

While (1) is relatively simple, it captures the interaction between advection, topographic forcing, and gravity waves in a manner similar to atmospheric flow over mountainous regions. As presented, these equations neglect the effects of planetary rotation in order to simplify the system as much as possible, and the subsequent numerical experiments will focus on regimes that would have a Rossby number greater than one if rotation terms were present.

1) Steady-state solutions

In this work, we investigate instabilities in discretized versions of (1) under steady flow over oscillatory topography, which we use as a stylized reduction of flow over mountainous regions. The topography height $|H|$ is small relative to the characteristic layer thickness ϕ_0, and the background flow consists of a mean flow u_0 with small variations, where u_0 is significantly smaller than the characteristic phase speed $\sqrt{g\phi_0}$.

In this environment, the steady-state $(\overline{\phi}, \overline{u})$ is found by solving:

$$\overline{u}\,\overline{\phi} = u_0\phi_0 \quad \text{and} \tag{2a}$$

$$\frac{1}{2}\overline{u}^2 + g\left[\overline{\phi} + H(x)\right] = \frac{1}{2}u_0^2 + g\phi_0, \tag{2b}$$

which arise from setting the time-derivative terms in (1) to zero. This equation can be solved as a cubic equation in either $\overline{u}$ or $\overline{\phi}$, but if $\phi_0 \gg |H|$, a first-order expansion is also convenient. Taking H_0 to be a characteristic topographic height and H_0/ϕ_0 as a small parameter:

$$\overline{\phi} = \phi_0\left[1 - \frac{gH(x)}{\phi_0 g - u_0^2} + O\left(\frac{H_0^2}{\phi_0^2}\right)\right] \quad \text{and} \tag{3a}$$

$$\overline{u} = u_0\left[1 + \frac{gH(x)}{\phi_0 g - u_0^2} + O\left(\frac{H_0^2}{\phi_0^2}\right)\right]. \tag{3b}$$

When (3) is used as an approximate solution to (2), the residual unsteady oscillation remains small and bounded.

2) Linearization

Now, we wish to linearize (1) about $(\overline{\phi}, \overline{u})$, rather than the (ϕ_0, u_0) state more commonly used in the literature. To do this, we expand:

$$\phi \to \overline{\phi} + \epsilon\phi' \quad \text{and} \tag{4a}$$

$$u \to \overline{u} + \epsilon u', \tag{4b}$$

where ε is a small parameter used for perturbation analysis and $(\cdot)'$ is the respective perturbation.

Substituting (4) into (1), dropping terms of $O(\varepsilon^2)$, and canceling the steady-state terms gives the linearized equations:

$$\phi_t' + (\overline{u}\phi' + \overline{\phi}u')_x = 0 \tag{5a}$$

$$u_t' + (\overline{u}u' + g\phi')_x = 0. \tag{5b}$$

For the steady-state background flows considered in this paper, (5) is linearly stable, without exponentially growing solutions.[1]

In this work, we investigate instabilities in the discretized version of (1), rather than continuous instabilities that would reflect physical phenomena such as extraction of energy from a background current. The consequential approach is to discretize (1) in its full nonlinear form, then to look at the linear response of an infinitesimal perturbation about a steady background state.

b. Time discretization

To discretize (1) in time in a way similar to GEM, we first write it in an advective form using the material derivative $D/Dt = \partial/\partial t + u\partial/\partial x$, which tracks the evolution of a quantity following advective motion. The continuity and momentum equations become

$$\frac{D\phi}{Dt} + \phi u_x = 0 \quad \text{and} \tag{6a}$$

$$\frac{Du}{Dt} + g(\phi + H)_x = 0. \tag{6b}$$

To this we also add an equation expressing the identity $x_t = 0$:

$$\frac{Dx}{Dt} - u = 0. \tag{6c}$$

In the Lagrangian interpretation, Eq. (6c) expresses the simple idea that fluid parcels are advected by the local velocity field, and in other notational conventions this may be written as $\dot{x} = u$ or $(d/dt)x = u$.

To discretize (6), we take a fully time-centered, two-level approach. The material derivative is replaced by a single difference, and terms appearing outside the material derivative are replaced by the average of old and new-time level values, following the flow. Using $(\cdot)^-$ for values at the old time level and $(\cdot)^+$ for values at the new time level, (6) becomes

$$\phi^+ - \mathrm{SL}(\phi^-) + \frac{\Delta t}{2}\left[\phi^+ u_x^+ + \mathrm{SL}(\phi^- u_x^-)\right] = 0, \tag{7a}$$

$$u^+ - \mathrm{SL}(u^-) + g\frac{\Delta t}{2}\left\{(\phi^+ + H)_x + \mathrm{SL}\left[(\phi^- + H)_x\right]\right\} = 0, \quad \text{and} \tag{7b}$$

$$x - \mathrm{SL}(x) - \frac{\Delta t}{2}\left[u^+ + \mathrm{SL}(u^-)\right] = 0, \tag{7c}$$

where SL is introduced as the semi-Lagrangian interpolation operator, which is made explicit here because the perturbation expansion of (7) [in (10)] will cause perturbations to this operator. The $(\cdot)^-$ and $(\cdot)^+$ fields are both functions of x, and $\mathrm{SL}[f(x)]$ performs the interpolation to give $f[x - r(x)]$ for some displacement function r depending on both u and Δt. As a consequence, $x - r(x)$ is also $\mathrm{SL}(x)$.

Equation (7c) acts to define fluid parcel trajectories over the finite interval. Rearranged, it gives $\mathrm{SL}(x) = x - (\Delta t/2)[u^+ + \mathrm{SL}(u^-)]$, which is a restatement of the implicit trapezoidal rule for calculating trajectories.

STEADY-STATE SOLUTIONS AND LINEARIZATION

As with the fully continuous equations, we are interested in linearizing about steady-state solutions to (7a). Unlike (1), however, in this Eulerian frame there is no simple time (partial) derivative to set to zero. Instead, at steady state $\phi^+ = \phi^- = \overline{\phi}$ and $u^+ = u^- = \overline{u}$, giving

$$\overline{\phi} - \overline{\mathrm{SL}}(\overline{\phi}) + \frac{\Delta t}{2}\left\{\overline{\phi}\overline{u}_x + \overline{\mathrm{SL}}\left[(\overline{\phi}\overline{u}_x)\right]\right\} = 0, \tag{8a}$$

$$\overline{u} - \overline{\mathrm{SL}}(\overline{u}) + g\frac{\Delta t}{2}\left\{(\overline{\phi} + H)_x + \overline{\mathrm{SL}}\left[(\overline{\phi} + H)_x\right]\right\} = 0, \quad \text{and} \tag{8b}$$

$$x - \overline{\mathrm{SL}}(x) - \frac{\Delta t}{2}\left[\overline{u} + \overline{\mathrm{SL}}(\overline{u})\right] = 0, \tag{8c}$$

where the semi-Lagrangian operator is denoted $\overline{\mathrm{SL}}$ to emphasize that fluid parcel trajectories are based on the steady-state velocities.

Unlike the fully continuous equations, (8) is not guaranteed to have well-behaved solutions. For $\Delta t > 1/\max(\overline{u}_x)$, the Lipschitz criterion (Diamantakis 2014; Staniforth and Côté 1991) fails, leading to crossing fluid parcel trajectories and non-monotonic $x - \overline{\mathrm{SL}}(x)$. Additionally, periodic H imposes an orographic Courant number constraint, where $\overline{u}\Delta t$ should remain smaller than the characteristic wavelength of the topography, lest a physical resonance arise. Finally, these equations can be subject to the spurious orographic resonance described by Rivest et al. (1994).

For realistic scenarios where H is small compared to the mean fluid depth, however, these restrictions are not very binding. In the small-topography limit, $\overline{u}_x$ is itself small, and in operational models Δt is generally taken small enough to satisfy an orographic Courant number restriction. Spurious orographic resonance is more stubborn, but it only affects a

[1] This can also be shown from an energy perspective, following Eldred and Randall (2017). The total energy $\int (1/2)[\phi u^2 + g\phi(2H + \phi)]dx$ is conserved, and expanding this with (4) leads to bounded variations in ϕ' and u' provided $\overline{u}^2 - g(\epsilon\phi' + \overline{\phi}) < 0$. For the cases discussed here, $\overline{u}^2$ is significantly smaller than $g\phi_0$, and this relationship is safely true throughout the domain.

small portion of the parameter space considered in this work.

As a result, steady-state solutions to (8) are generally available. As with section 2a, we are interested in solutions that are close to a "topography-free" state of (ϕ_0, u_0). The steady-state $\overline{\phi}$ and $\overline{u}$ can be computed from any spatially discretized version of (8) through Newton's method, and the approximation of (3) provides a fine initial guess.

Given a steady-state solution, we are now interested in determining its stability. We again expand $\phi(x, t)$ and $u(x, t)$ as the sum of $\overline{\phi}$ and $\overline{u}$, respectively, plus a small perturbation, and because we are interested in exponential growth we assume that this time-varying perturbation has a characteristic eigenvalue λ. This gives

$$\overline{\phi} \to \overline{\phi} + \epsilon\phi', \quad \phi^+ \to \overline{\phi} + \epsilon\lambda\phi', \tag{9a}$$

$$\overline{u} \to \overline{u} + \epsilon u', \quad u^+ \to \overline{u} + \epsilon\lambda u'. \tag{9b}$$

Substituting (9) into (7) first begins with (7c) in order to define SL up to $O(\epsilon)$. This substitution gives

$$\mathrm{SL}(x) = x - \frac{\Delta t}{2}\left[(1 + \mathrm{SL})(\overline{u}) + \epsilon(\lambda + \mathrm{SL})u'\right]. \tag{10}$$

Now, we can expand SL itself in ε, as $\mathrm{SL} \to \overline{\mathrm{SL}} + \epsilon\mathrm{SL}'$. Applying this to (10) and using (8c) gives

$$\epsilon\mathrm{SL}'(x) = -\frac{\Delta t}{2}\,\epsilon(\lambda + \overline{\mathrm{SL}})(u') + O(\epsilon^2). \tag{11}$$

Remembering the definition of SL as an interpolation operator with $\mathrm{SL}[f(x)] = f[\mathrm{SL}(x)]$, we can now express SL in terms of $\overline{\mathrm{SL}}$ and a spatial derivative using a Taylor series expansion:

$$\begin{aligned}
\mathrm{SL}[f(x)] &= f\left[\overline{\mathrm{SL}}(x) + \epsilon\mathrm{SL}'(x)\right] \\
&= f\left[\overline{\mathrm{SL}}(x)\right] + \epsilon\mathrm{SL}'(x)\partial_x\left\{f\left[\overline{\mathrm{SL}}(x)\right]\right\} + O(\epsilon^2) \\
&= f\left[\overline{\mathrm{SL}}(x)\right] - \epsilon\frac{\Delta t}{2}(\lambda + \overline{\mathrm{SL}})(u')\partial_x\left\{f\left[\overline{\mathrm{SL}}(x)\right]\right\} \\
&\quad + O(\epsilon^2) \\
&= \overline{\mathrm{SL}}[f(x)] - \epsilon\,\frac{\Delta t}{2}(\lambda + \overline{\mathrm{SL}})(u')\partial_x\left\{\overline{\mathrm{SL}}[f(x)]\right\} \\
&\quad + O(\epsilon^2).
\end{aligned} \tag{12}$$

It is important to stress here that the spatial derivative ∂_x is that induced by an infinitesimal increment to the fluid parcel trajectory. In the case of the fully discrete system, this implied derivative operator is not necessarily the same as the derivative operator used for the gravity wave terms of (7a) and (7b).

With (12) in hand, (9) can finally be applied to (7a) and (7b) (dropping $O(\varepsilon^2)$ and higher-order terms) to give the generalized eigenvalue problem for small perturbations. Doing so gives

$$\begin{aligned}
&\overline{\mathrm{SL}}\left[\left(-1 + \frac{\Delta t}{2}\overline{u}_x\right)\phi' + \frac{\Delta t}{2}\overline{\phi}u'_x\right] + \frac{\Delta t}{2}\overline{\mathrm{SL}}(u')\partial_x\overline{\mathrm{SL}}\left(\overline{\phi} - \frac{\Delta t}{2}\overline{\phi}\overline{u}_x\right) \\
&= -\lambda\left[\left(1 + \frac{\Delta t}{2}\overline{u}_x\right)\phi' + \frac{\Delta t}{2}\overline{\phi}u'_x + \frac{\Delta t}{2}u'\partial_x\overline{\mathrm{SL}}\left(\overline{\phi} - \frac{\Delta t}{2}\overline{\phi}\overline{u}_x\right)\right] \quad \text{and}
\end{aligned} \tag{13a}$$

$$\begin{aligned}
&\overline{\mathrm{SL}}\left\{\left(-u' + g\frac{\Delta t}{2}\phi'_x\right) + \frac{\Delta t}{2}\overline{\mathrm{SL}}(u')\partial_x\overline{\mathrm{SL}}\left[\overline{u} - g\frac{\Delta t}{2}(\overline{\phi} + H)_x\right]\right\} \\
&= -\lambda\left\{u' + g\frac{\Delta t}{2}\phi'_x + \frac{\Delta t}{2}u'\partial_x\overline{\mathrm{SL}}\left[\overline{u} - g\frac{\Delta t}{2}(\overline{\phi} + H)_x\right]\right\},
\end{aligned} \tag{13b}$$

where (8) has been used to cancel the steady-state solution.

The time-discretized system (7), with a steady-state background flow given by (8), is unstable to infinitesimal perturbations if any of the eigenvalues in the linearized system (13) have a magnitude greater than one. Since (13) is not analytically tractable, we instead proceed by discretizing in space and solving the discrete eigenvalue problem.

c. Spatial discretization and numerical implementation

To investigate the stability of (13), we implement a spatial discretization in MATLAB, alongside an equivalent discretization of (7) (to allow for time integration) and (8) (to compute steady-state solutions). Overall, this discretization follows the basic framework of GEM, using an Arakawa C-grid. In one dimension, this becomes a simple grid staggering, where u and ϕ points are offset by $(1/2)\Delta x$, and for the sake of this analysis we take the underlying grid to be periodic in x with an overall domain size of L_x. The topographic height H is also assigned to the ϕ grid.

The spatial derivatives used in the gravity wave terms $[(\cdot)_x]$ are approximated as divided differences, where the approximate derivative "lives" on the other grid.

For the initial analysis, the semi-Lagrangian operator will be implemented through a Fourier expansion, expanding the interpolated function in its Fourier series and evaluating the terms at the departure point (Boyd 2001). This preserves the amplitude of wave-like functions, and it will allow us to more directly examine the stability implications of the temporal discretization. The use of polynomial interpolation, which is more conventional and used in the operational GEM model, is discussed in section 4. The ∂_x terms which appear in (13) are evaluated in the semi-Lagrangian context. For the Fourier interpolation this is computed by evaluating the derivative of the Fourier series off-grid, and for polynomial interpolation this is computed by evaluating the derivative of the respective Lagrange interpolating polynomials.

For time-varying and steady-state computations, the semi-Lagrangian trajectories (those implied in the SL and $\overline{\mathrm{SL}}$ operators) are found by iterating (7c) and (8c) to convergence.

The linear system corresponding to (13) is full because of the Fourier interpolation in the semi-Lagrangian terms, but computation of its eigenvalues and eigenvectors is still quite tractable because the one-dimensional system is relatively small. To explore this system, we consider a parameter space

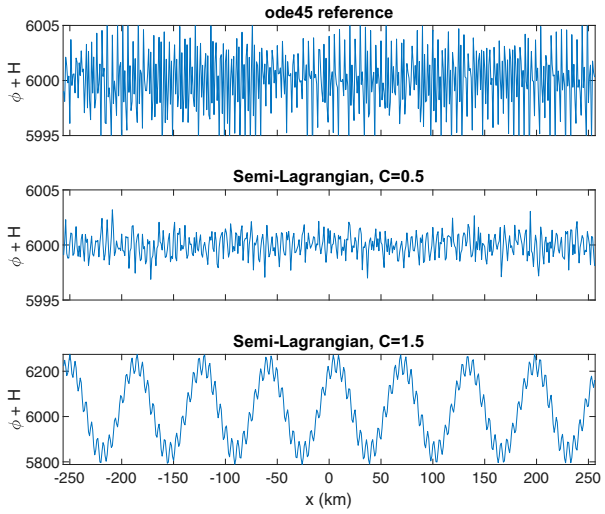

FIG. 1. Layer height ($\phi + H$) after 4 simulated hours for the shallow-water system as described in section 2c, with a topography wavelength of $5\Delta x$, initial conditions given by (3), and an initial random perturbation in ϕ of white noise with a standard deviation of 1 m. (top) Reference solution computed with a Fourier spectral method and Matlab's *ode45*; (middle) solution with the semi-Lagrangian, implicit method described in this paper and a Courant number of 0.5; and (bottom) the semi-Lagrangian, implicit method and a Courant number of 1.5. Note the larger vertical axis range in the bottom panel.

that is stylistically representative of flow past topography in high-resolution, local-area simulations of the atmosphere:

- $\Delta x = 1$ km, with $N_x = 513$ equally spaced points[2] on both the ϕ and u grids giving an overall domain size (L_x) of 513 km
- $\phi_0 = 6$ km, approximately the atmospheric scale height, and $u_0 = 20$ m s^{-1}, borrowed from Williamson et al. (1992)'s first test case.
- $H(x) = H_0 \sin(k_{\text{hill}} x)$ to present topography with a single wavelength. $H_0 = 100$ m, and k_{hill} will vary below.

This specification ensures that the magnitude of the topographic perturbation (H_0/ϕ_0) is small (about 1.7×10^{-2}), so the steady-state solution to (8) is a small perturbation to the uniform state (ϕ_0, u_0). The characteristic Froude number $(u_0/c = u_0/\sqrt{g\phi_0})$ is much smaller than one, so the gravity waves produced in this system should be subcritical and propagate both up and downstream. The Rossby number ($c/L_x f$) is also comfortably greater than one for rotation rates corresponding to the midlatitudes, suggesting that the rotation effects neglected in (1) should not significantly affect these results.

A typical example of this parameter regime is shown in Fig. 1, which depicts the layer height ($\phi + H$) after 4 h of simulated time under a reference solution (a Fourier-transform spectral method for the spatial discretization and Matlab's *ode45* routine

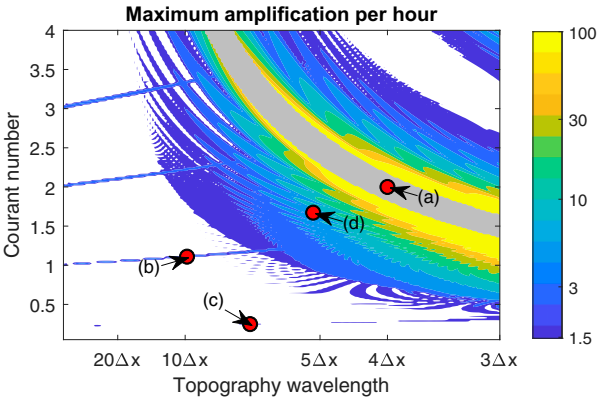

FIG. 2. Normalized maximum growth rates for the system of (13) under the conditions described in section 2c, showing the growth rate per hour of the most unstable mode for each (C, k_{hill}) pairing. Four different unstable branches appear: (a) is spurious orographic resonance, (b) is advective aliasing, (c) is wave spatial aliasing, and (d) is wave temporal aliasing. The instability growth rate is shown via the colored regions, with a lower bound of 1.5 (50% amplification per hour) chosen to highlight the (b) and (c) branches.

for time integration) and this discretized system at Courant numbers ($C = \Delta t u_0/\Delta x$) of 0.5 and 1.5, with $k_{\text{hill}} \approx 2\pi/(5\Delta x)$, and initial conditions given by the approximation (3).[3] While the reference solution and small-time-step semi-Lagrangian solution are well behaved, generating a small gravity wave field of amplitude comparable to the initial perturbation, the large-time-step semi-Lagrangian solution develops an extremely large gravity wave, with an amplitude several times larger than the topography's sinusoidal amplitude..

This is one example of the instabilities present in the time-discretized system (7), and these instabilities are present over a large fraction of the reasonable parameter space. Varying Δt and k_{hill} changes the effective Courant number and topographic forcing wavenumber. Varying C between 0.05 and 4 (Δt between 2.5 and 200 s) over 200 equally spaced steps and k_{hill} between $2\pi/2.5L_x$ and $2\pi/3\Delta x$ over 167 equally spaced steps and finding the most unstable eigenmode of the linearized system (13) gives the graph of Fig. 2.

Throughout this parameter space, the Lipschitz trajectory crossing criterion ($\Delta_t|u_x| < 1$) is well respected, and the obstacle Courant number ($\Delta t|u|$ divided by the topographic wavelength) remains below one everywhere except the upper-right corner of Fig. 2.

3. Analysis

Most of the instabilities depicted in Fig. 2 take on a generic form: a wavelike eigenmode that oscillates at a

[2] Using an odd number of points eliminates the exact, $2\Delta x$ "sawtooth" mode on the grid, which slightly simplifies the subsequent numerical analysis by allowing every remaining mode to be identified as upstream or downstream propagating.

[3] This example does not use the steady state computed after the time discretization because (8) depends on the time step and time discretization. Since each panel of the figure would then have slightly different initial conditions, it would not be a like-for-like comparison.

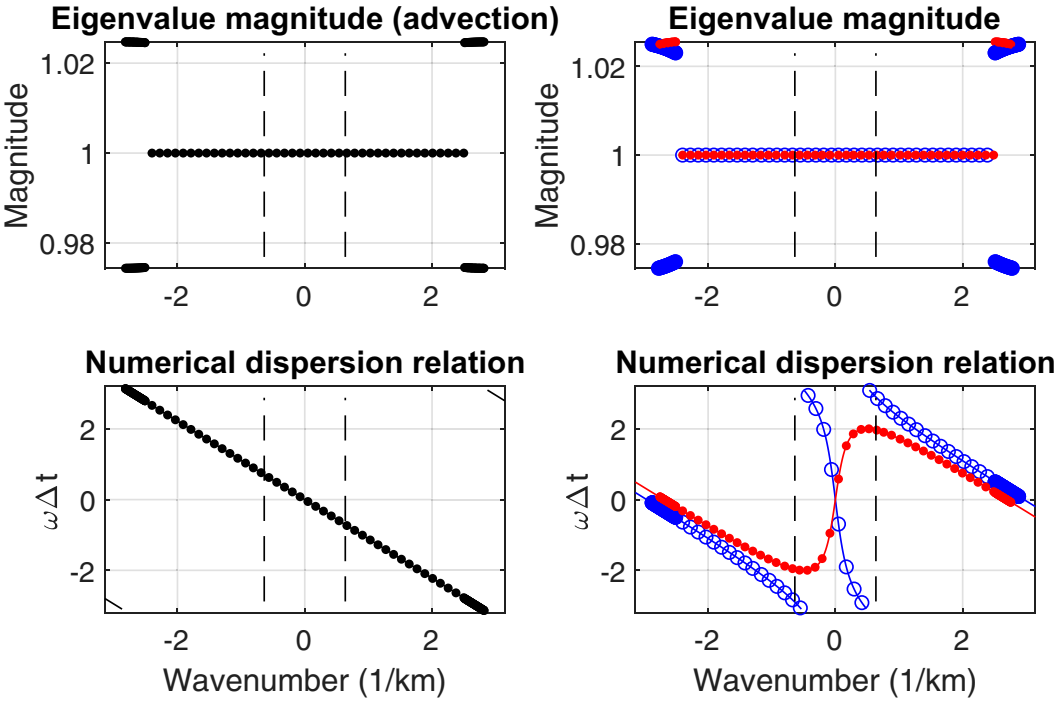

FIG. 3. (top) Eigenvalue magnitudes and (bottom) dispersion relation for the (left) advective operator and (right) shallow-water operators (open circles in blue for downstream-propagating waves and solid circles in red for upstream-propagating waves) at the point marked (b) in Fig. 2. The dashed black vertical lines mark $\pm k_{\text{hill}}$. In the bottom panels, symbols show the calculated dispersion relation while the line shows the linear estimate. To improve readability, only a selection of modes with near-unit-amplitude eigenvalues are shown.

single, dominant wavenumber k, with smaller-scale oscillations at shorter wavelengths, with corresponding wavenumbers $k \pm jk_{\text{hill}}$ for integer j. A typical example is shown in the bottom panel of Fig. 1.

The necessary conditions for these instabilities are the existence of two wavelike solutions to the linearized equations with distinct wavenumbers k_1 and k_2, where

$$k_1 \pm k_{\text{hill}} \approx k_2 \mod 2\pi/\Delta x \quad \text{and} \tag{14a}$$

$$\omega(k_1) \approx \omega(k_2) \mod 2\pi/\Delta t. \tag{14b}$$

When the instability exists, forcing induced by the background flow causes waves of mode k_1 to generate waves of (approximately) mode k_2 and vice versa. When the condition of (14b) holds, these waves remain resonant with each other, so this wave generation process adds energy in an additive, coherent way. For the purposes studied here, it also suffices for (14b) to use the linearized dispersion relation $\omega(k)$ of (7), with the linearization taken about the topography-free reference state (ϕ_0, u_0):

$$\omega(k) = u_0 k \pm \frac{2}{\Delta t} \arctan\left[\frac{\Delta t}{\Delta x} c_0 \sin\left(\frac{1}{2} k \Delta x\right)\right], \tag{15}$$

where $c_0 = \sqrt{g\phi_0}$ is the (nondispersive) phase speed of waves in the continuous shallow-water system and of long waves ($k \to 0$) in the discretized system.

The exact mechanics of the coupling change depending on the branch of the instability marked in Fig. 2.

a. Spurious orographic resonance

The most striking feature of Fig. 2 is the large central region excluded from the contour plot. Here, the steady-state solution process described in section 2b fails through the well-known "spurious orographic resonance" process, with the excluded region matching the first unstable region in Fig. 1 of Rivest et al. (1994). In the region of Fig. 2 with orographic resonance, $\omega(k_{\text{hill}}) \approx 0$, and perturbation of the initially undisturbed system by topographic forcing generates standing waves with zero phase speed. In the language of this section, the $k = 0$ "wave"—the steady-state solution—shares $\omega = 0$ with gravity waves that would have $k = \pm k_{\text{hill}}$, and the topography at $k = k_{\text{hill}}$ provides the coupling mechanism.

In a time-dependent simulation initialized from a uniform state, the topographic forcing accumulates with no linear decay or other damping, and this leads to no physically relevant steady-state solution.

In practice, weather models such as GEM use off-centering to partially mitigate this issue, which will be discussed further in section 4.

b. Advective aliasing

Figure 3 depicts the situation with $k_{\text{hill}} \approx 0.633\ \text{km}^{-1}$ (wavelength = $9.92\Delta x$) and $\Delta t = 55.5$ s, corresponding to point (b) in

Fig. 2, giving an advective Courant number of 1.11—comfortably within the expected range of validity for a semi-Lagrangian, implicit method. In this figure, the eigenmodes of the advective operator $[\overline{\mathrm{SL}}(f) = \lambda f]$ and shallow-water time-stepping Eqs. (13) are identified as single-wavenumber modes based on their strongest single components [i.e., a mode with form $\exp(\mathbf{i}k_0x) + \varepsilon \exp(\mathbf{i}k_1x)$ would be identified as having wavenumber k_0], and their magnitudes and phase angles are shown in the top and bottom panels, respectively. Modes with a principal component near the Nyquist limit ($k_0 \approx \pm\pi\ \mathrm{km}^{-1}$) are no longer of unit amplitude or less; instead these modes have paired into growing and decaying components. Except for these growing/decaying modes, the phase angles remain very close to those predicted by a simple linear dispersion relation (linearizing the discrete system about ϕ_0 and u_0, rather than $\overline{\phi}$ and $\overline{u}$).

This instability arises because of the behavior of the semi-Lagrangian operator $\overline{\mathrm{SL}}$. Taking the approximate steady-state solution from (3b) and substituting the sinusoidal profile of H gives

$$\overline{u} = u_0\left[1 + \frac{gH_0\sin(k_{\mathrm{hill}}x)}{\phi_0 g - u_0^2} + O\left(\frac{H_0^2}{\phi_0^2}\right)\right]. \tag{16}$$

This sinusoidal perturbation is replicated in $\overline{\mathrm{SL}}$, where $\overline{\mathrm{SL}}(x) = x - \Delta t u_0 - D\Delta t\sin(k_{\mathrm{hill}}x - \overline{\theta}) + O\left(H_0^2/\phi_0^2\right)$, where $\overline{\theta}$ is a phase shift arising from applying the trapezoidal rule to evaluated trajectories in (8c) and D is proportional to u_0H_0/ϕ_0.

Applying this operator to an arbitrary f gives

$$\begin{aligned}\overline{\mathrm{SL}}[f(x)] &= f(x - \overline{\mathrm{SL}}\,x) = f(x - \Delta t u_0)\\ &\quad - D\Delta t\sin(k_{\mathrm{hill}}x + \overline{\theta})f_x(x - \Delta t u_0) + O\left(\frac{H_0^2}{\phi_0^2}\right).\end{aligned} \tag{17}$$

If f is the complex exponential $\exp(\mathbf{i}k_0x)$, then dropping higher-order terms simplifies this to

$$\begin{aligned}&\overline{\mathrm{SL}}[\exp(\mathbf{i}kx)] = \exp[\mathbf{i}k_0(x - u_0\Delta t)][1 - D\mathbf{i}k\sin(k_{\mathrm{hill}}x + \overline{\theta})] \quad \text{or}\\ &\overline{\mathrm{SL}}[\exp(\mathbf{i}kx)] = \exp(-\mathbf{i}k_0u_0\Delta t)[\exp(\mathbf{i}k_0x)] + E\mathbf{i}k_0\exp[\mathbf{i}(k_0 + k_{\mathrm{hill}}x)]\\ &\qquad + F\mathbf{i}k_0\exp[\mathbf{i}(k_0 - k_{\mathrm{hill}}x)],\end{aligned} \tag{18}$$

where the $\sin(k_{\mathrm{hill}}x + \overline{\theta})$ term is expanded to give (18), with constants E and F that absorb the phase shift $\overline{\theta}$. Note the net effect of (18): $\overline{\mathrm{SL}}$ takes a single-mode signal, rotates its phase by $\exp(-\mathbf{i}k_0\Delta t u_0)$ in the complex plane, and adds new components corresponding to the sum and difference frequencies with an amplitude proportional to H_0/ϕ_0.

This gives rise to aliasing when k_0 is near the Nyquist limit $k_{\mathrm{nyq}} = \pi/\Delta x$. For such a positive k_0, $k_0 + k_{\mathrm{hill}} > k_{\mathrm{nyq}}$, and on the discrete grid this high-wavenumber wave appears identical to a wave with $k_1 = k_0 + k_{\mathrm{hill}} - 2k_{\mathrm{nyq}}$. This aliasing in turn leads to an instability when the corresponding temporal frequencies are aliased, with $\exp(-\mathbf{i}k_0\Delta t u_0) \approx \exp(-\mathbf{i}k_1\Delta t u_0)$.

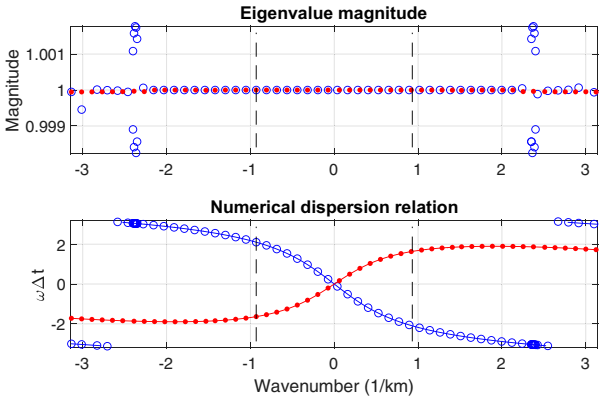

FIG. 4. (top) Eigenvalue magnitudes and (bottom) dispersion relation for the shallow-water operator at the point marked (c) in Fig. 2. The dashed vertical lines denote $\pm k_{\mathrm{hill}}$. As in Fig. 3, the blue open circles correspond to downstream-propagating modes, while the red filled circles correspond to upstream-propagating modes. For clarity, a reduced set of near-unit-amplitude modes are shown, and in the bottom panel the solid lines show the dispersion relation calculated from (15).

When this condition is satisfied, waves with wavenumbers k_0 and k_1 oscillate together in time, and when the topographic forcing produces waves of mode k_1 out of the k_0 mode (or vice versa) this energy accumulates.

The temporal resonance occurs when the two phase shifts are equivalent modulo 2π, or

$$\Delta t u_0 \approx \frac{2n\pi}{2k_{\mathrm{nyq}} - k_{\mathrm{hill}}} \tag{19}$$

for integer n. For the case depicted in Fig. 3, $\Delta t u_0 = 1.11\Delta x$, $k_{\mathrm{nyq}} = \pi/\Delta x$, and $k_{\mathrm{hill}} = 2\pi/(9.92\Delta x)$, giving $n = 0.998$. The unstable regions extending rightwards in Fig. 2 from Courant numbers of 2 and 3 correspond to $n = 2$ and 3, respectively.

In brief, topographic forcing via the advection operator causes a propagating, single-wavenumber mode to produce daughter waves separated in wavenumber by k_{hill}, which may spatially alias across the Nyquist frequency. When the leading mode and its daughter waves share a time-aliased temporal frequency, this wave production remains coherent, leading to a growing response with amplitude proportional to H_0/ϕ_0.

This instability of this section arises from the semi-Lagrangian operator itself, and it can potentially act as a noise-generating process for passively advected tracers such as moisture or chemical constituents. This instability is inherent to a finite-time semi-Lagrangian advection, and it does not depend on any particular trajectory calculation methodology: a function that is barely resolved on a discrete grid becomes unresolved, with aliasing error, after its locally compressed by semi-Lagrangian trajectories that converge.

c. *Wave spatial aliasing*

An analogous effect can also occur for gravity wave modes, even if the instability of section 3b is not present. Figure 4

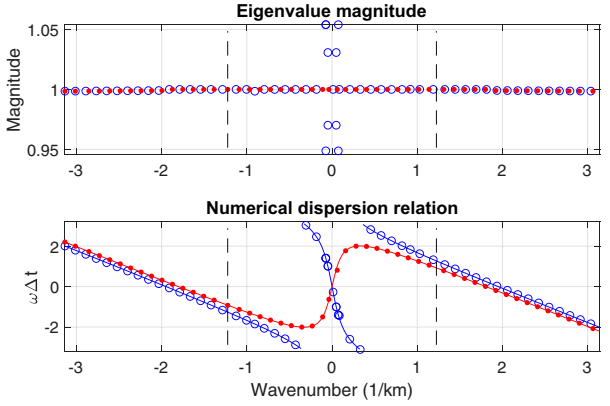

FIG. 5. (top) Eigenvalue magnitudes and (bottom) dispersion relation for the shallow-water operator at the point marked (d) in Fig. 2. The dashed vertical lines mark $\pm k_{\text{hill}}$. The symbols and solid lines have the same meaning as in Fig. 4.

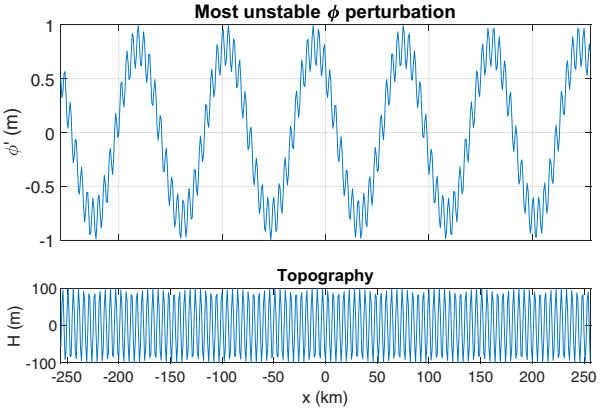

FIG. 6. (top) Perturbation in the ϕ field corresponding to the most unstable mode in Fig. 5. (bottom) The unstable mode consists of a large, long-wavelength oscillation modified by an oscillation with a wavelength similar to that of the underlying topography.

depicts the situation with $k_{\text{hill}} \approx 0.953\ \text{km}^{-1}$ (wavelength = $6.59\Delta x$) and $\Delta t = 12.43$ s, giving an advective Courant number of approximately 0.25. This set of parameters is marked as (c) in Fig. 2, along the thin line of instability at low Courant numbers.

Finding any instability at such a low Courant number is surprising, but downstream-propagating gravity waves suffer from the same sort of aliasing as in Fig. 3. In this case, the shallow-water operators themselves contribute the necessary secondary modes to a perturbation that initially oscillates with a single wavenumber, but the full dispersion relation possesses the same sort of temporal aliasing.

In the linearized dispersion relation of (15), we seek two wave modes k_0 and $k_1 = k_0 + k_{\text{hill}} - 2k_{\text{nyq}}$ such that $\exp[-\mathbf{i}\omega(k_0)\Delta t] \approx \exp[-\mathbf{i}\omega(k_1)\Delta t]$, or

$$\begin{aligned} &u_0\Delta t(2k_{\text{nyq}} - k_{\text{hill}}) \\ &+ 2\arctan\left[\frac{\Delta t}{\Delta x}\, c_0 \sin\left(\frac{1}{2}k_0\Delta x\right)\right] \\ &- 2\arctan\left[\frac{\Delta t}{\Delta x}\, c_0 \sin\left(\frac{1}{2}k_1\Delta x\right)\right] = 2n\pi, \end{aligned} \tag{20}$$

where again n is approximately an integer. Equation (20) has no simple closed-form solution for k_0 given n, but it can be evaluated numerically. The most unstable (positive wavenumber) mode in Fig. 4 has wavenumber $k_0 \approx 2.44\ \text{km}^{-1}$, and that wavenumber and the corresponding $k_1 \approx 2.88\ \text{km}^{-1}$ satisfy (20) with $n = 1$ and a residual of 2.7×10^{-4}.

Unlike the conditions for advective aliasing in (19), the more complicated functional form of (20) results in at most a few near-resonant spatial frequencies for any given configuration.

d. Wave temporal aliasing

With larger Courant numbers, the temporal aliasing can occur independently of spatial aliasing. Figure 5 depicts the eigenspectrum with $k_{\text{hill}} \approx 1.22\ \text{km}^{-1}$ (wavelength = $5.14\Delta x$) and $\Delta t = 83.60$ s, giving an advective Courant number of approximately 1.672. This set of parameters is marked as (d) in Fig. 2, and this sort of behavior (for various unstable perturbation wavenumbers) is typical in the large region of instability near the limit of spurious orographic resonance.

The form of the most unstable mode is shown in Fig. 6. For this set of parameters, the unstable mode takes the form of a long wave (with wavelength of about 85.7 km, or $k_0 \approx 7.33 \times 10^{-2}\ \text{km}^{-1}$), which has a shorter wave superimposed on it at lower amplitude. That smaller wave has a wavelength comparable to the topography, with $k_1 \approx 1.15\ \text{km}^{-1}$. Both waves are well resolved on the grid.

Here, the linearized dispersion relation of (15) is still applicable, and the temporal aliasing relationship is

$$\begin{aligned} &u_0\Delta t k_{\text{hill}} + 2\arctan\left[\frac{\Delta t}{\Delta x}\, c_0 \sin\left(\frac{1}{2}k_0\Delta x\right)\right] \\ &- 2\arctan\left[\frac{\Delta t}{\Delta x}\, c_0 \sin\left(\frac{1}{2}k_1\Delta x\right)\right] = 2n\pi, \end{aligned} \tag{21}$$

developed analogously to (20). $k_1 \approx k_0 - k_{\text{hill}}$, and (21) holds with $n = 1$ and a residual of 1.44×10^{-4}. Taken independently, the long/slow wave and the short/fast wave have nearly identical phase shifts over each time step, and the topography-induced background flow causes each wave to generate its counterpart. The overall effect is that the perturbation grows by about 5% per time step, or by a factor of 9.5 per simulated hour.

This instability requires a higher Courant number than that of section 3c, since the frequency wraparound necessary for (15) to alias must occur over two waves separated by k_{hill} rather than the much larger k_{nyq}. This occurs more readily for downstream-propagating waves, where the background velocity and gravity wave propagation effects combine to give a much higher effective gravity wave Courant number.

Overall, the semi-Lagrangian treatment of advection is "doing its job" quite well: from one time step to the next it

advects the ϕ and u fields with the background velocity. However, the net phase speed of a linear mode is given not only by the background velocity, but also by its propagation as a gravity wave. When the conditions of (21) are met, a pair of modes separated by k_{hill} oscillate at the same frequency, while topographic forcing causes each mode of the pair to feed energy into its partner. This instability is a "moving spurious orographic resonance," and it does not depend on the precise details of semi-Lagrangian trajectory calculations.

4. Mitigations

Section 3 demonstrates that the shallow-water system, adapted to steady flow over topography, is unstable to certain infinitesimal perturbations over a wide parameter range. However, operational dynamical cores already contain diffusive and dissipative elements that are not present in the idealized, time-centered discretization discussed here, and some of these stabilizing elements may also affect these instabilities. This section discusses the extent to which realistic choices made in dynamical cores might mitigate the presence or strength of the instabilities.

For this purpose, the instability described in section 3d will be used as a prototypical example, with modifications as noted.

a. Need for mitigations

The first approach to mitigate these instabilities may be to do nothing at all. Over most of the parameter space in Fig. 2, these instabilities grow relatively slowly (by a factor of 3–10 per hour). Additionally, a typical forecast domain does not look like a sinusoidal mountain range, so an instability will generally have only a finite "contact time" to extract energy from the background flow. If left totally unmitigated, how badly can these instabilities corrupt an otherwise smooth, steady solution?

To observe this, we take the configuration of section 3d, perturb ϕ pointwise by white noise with a standard deviation of 1 m, or 10^{-2} times the amplitude of the hill, and simulate the resulting system via (7) with a Newton solver used for the nonlinear implicit terms. The simulation is run for a period of 8 h, and typical results are shown in Fig. 7. Despite the random perturbation, the system effectively selects the most quickly growing modes, and those begin to dominate over the 2–4-h interval.

After reaching an initial root-mean-square amplitude of 100 m, the growth rate of the perturbation begins to slow down. By the end of the simulation, the perturbation has not become large enough to cause the model to "crash," but the perturbation amplitude has reached about 600 m, several times the amplitude of the underlying topography. While the perturbation growth has effectively saturated because of nonlinear effects, it has effectively converted all of the energy from the background flow variation into the unstable mode. The overall flow structure no longer reflects reality.

b. Adjustments to the nonlinear solver

Operational weather models typically do not solve the nonlinear implicit system of (7) to convergence. As an example, in most operational configurations the GEM dynamical core

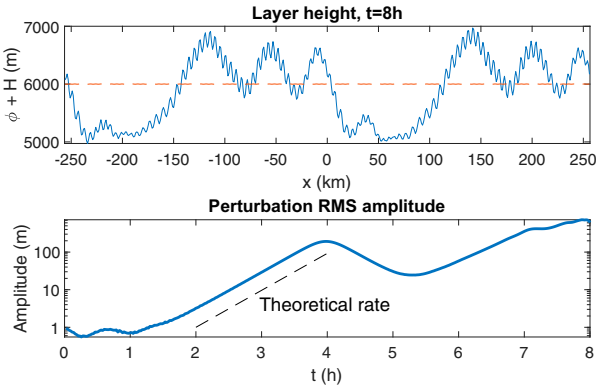

FIG. 7. (top) Layer height ($\phi + H$, solid) corresponding to the case of section 3d after 8 h of simulated time, beginning with a random ϕ perturbation from the steady state (dashed). (bottom) Root-mean-square amplitude of the perturbation $\phi - \overline{\phi}$ over the simulated period, with a dashed line denoting the theoretically expected growth rate. On the axis scales used here, the "exact" solution is visually indistinguishable from the steady state.

truncates the equivalent solution process (for the Euler or hydrostatic equations) at four iterations, consisting of two outer trajectory iterations and two inner, nonlinear iterations (with frozen semi-Lagrangian trajectories). The inner solution step involves a linear solve using the common quasi-linear approximation, where the full Jacobian of the Euler equation system is replaced with a linearization about a prescribed atmospheric state.

The equivalent to this in the shallow-water context is to treat (7a) and (7b) as a fixed point iteration. Rather than compute ϕ^+ and u^+ directly, instead treat $\phi^+ = \phi^* + \tilde{\phi}$ and $u^+ = u^* + \tilde{u}$, where ϕ^* and u^* are values from the previous iteration (or an initial guess) and $\tilde{\phi}$ and $\tilde{u}$ are increments to be computed. Using this expansion in (7a) and (7b) and applying the quasi-linear approximation to the increment terms gives

$$\tilde{\phi} + \frac{\Delta t}{2}\phi_{\text{ref}}\tilde{u}_x = \text{SL}\left[\phi^-\left(1 - \frac{\Delta t}{2}u_x^-\right)\right] - \phi^*\left(1 + \frac{\Delta t}{2}u_x^*\right), \tag{22a}$$

$$\tilde{u} + \frac{\Delta t}{2}g\tilde{\phi}_x = \text{SL}\left[u^- - \frac{\Delta t}{2}g(\phi^- + H)_x\right] - u^* - \frac{\Delta t}{2}g(\phi^* + H)_x, \tag{22b}$$

where ϕ_{ref} is a reference layer thickness.

Under a wide range of circumstances, iterating (22) converges to the solution of (7), giving stability results identical to those of section 4a. Using even two iterations leads to instability growth at a rate very close to that predicted for the full nonlinear system; this is shown in the top panel of Fig. 8.

However, taking just one iteration is special. With the initial guess $\phi^* = \phi^-$ and $u^* = u^-$, this constitutes a *semi*-implicit time-

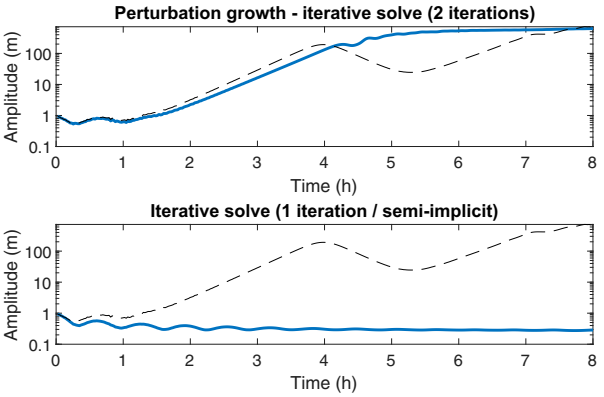

FIG. 8. Root-mean-square amplitude (solid) of the random perturbation whose evolution was shown in Fig. 7 over an 8-h simulated period, with an iterative solver truncated after (top) two and (bottom) one iterations, using a basic state of $\phi_{\text{ref}} = 1.1\phi_0$. The bottom panel corresponds to semi-implicit time-stepping, and in each panel the growth rate curve of Fig. 7 is shown (dashed) for reference.

stepping algorithm[4] equivalent to that described for the Euler equations by Simmons and Temperton (1997), where the time centering of the gravity wave terms is applied only to the reference state. This algorithm is generally stable when $\phi_{\text{ref}} > \max(\phi)$, and here we take $\phi_{\text{ref}} = 1.1\phi_0$. Under these circumstances, the random perturbation decays with time, leading to the amplitude graph shown in the bottom panel of Fig. 8.

Unfortunately, semi-implicit time stepping is also in general less accurate than the fully implicit trapezoidal rule. Because the time-averaging is not applied to $\phi - \phi_{\text{ref}}$ or u in the gravity wave terms, only small perturbations from the reference state are propagated with second order in time. Large "nonlinear" motions are evaluated on a forward-in-time basis at first order, and the $\phi_{\text{ref}} > \max(\phi)$ requirement also contributes to the loss of accuracy by pushing the reference state further from a true linearization of the problem.

The SETTLS approach (Hortal 2002), incorporated into the IFS model (ECMWF 2020), regains a second-order discretization by extrapolating the nonlinear terms along the semi-Lagrangian trajectory, but a straightforward formulation has a reduced stability region compared to the fully implicit formulation (Durran and Reinecke 2004), such that dissipation included in the linear term may still be required for overall, unconditional stability. It is not clear whether the secondary instability described in this work has an equivalent with SETTLS extrapolation.

[4] The nomenclature in the literature is not quite consistent here: Girard et al. (2014) and Wood et al. (2014) use similar time-stepping algorithms that treat the nonlinear terms implicitly with a restricted number of iterations, but the former describes the method as "implicit" and the latter "semi-implicit," while Benacchio and Wood (2016) suggests the phrase "iterative implicit" for a nonlinear, implicit system solved with a fixed number of iterations. Here, we use the phrase "semi-implicit" to mean that the time averaging is not applied to some portion of (1).

c. Semi-Lagrangian dissipation

Another potential mitigation for these instabilities is the dissipation inherent in practical semi-Lagrangian interpolation. Thus far, the semi-Lagrangian interpolation used has been a Fourier interpolation that exactly reproduces single frequencies up to the grid resolution (Nyquist) limit. This requires $O(N^{2d})$ operations to interpolate N^d points over d dimensions, so it is impractical in any realistic three-dimensional model. Instead, these models generally rely on local polynomial interpolation, and cubic interpolation is a popular choice (e.g., ECMWF 2020; Girard et al. 2014).

This polynomial-based interpolation no longer preserves wave amplitudes, and instead it dissipates waves in proportion to their wavenumber (relative to the grid scale) and the distance between the interpolated, off-grid point and the nearest on-grid location. For example, bilinear interpolation to a point $x' = x_j + \delta$ (relative to on-grid points x_0 and $x_0 + \Delta x$, with $0 \le \delta \le \Delta x$) is

$$f(x_0 + \delta) \approx \left(1 - \frac{\delta}{\Delta x}\right) f(x_0) + \frac{\delta}{\Delta x} f(x_0 + \Delta x). \tag{23}$$

For a single complex exponential mode $f(x) = \exp(\mathbf{i}kx)$, and (23) becomes

$$f(x_0 + \delta) \approx \exp(\mathbf{i}kx_0)\left\{1 + \frac{\delta}{\Delta x}\left[-1 + \exp(\mathbf{i}k\Delta x)\right]\right\}. \tag{24}$$

This expression has a magnitude of $1 + 2(\delta/\Delta x - \delta^2/\Delta x^2) \times [\cos(k\Delta x) - 1]$, which reaches its minimum of 0 at $\delta = \Delta x/2$ (halfway between grid points) and $k = \pi/\Delta x$ (the Nyquist limit). Cubic interpolation via Lagrange interpolating polynomials follows a similar relationship, cubic in $\cos(k\Delta x)$.

The effect of both linear and cubic interpolation is to selectively damp higher frequencies, and in the context of the shallow-water equations this damping becomes preferentially stronger when the advective Courant number is close to $n + 1/2$ for integer n. For linear interpolation, this effect is quite severe, and in the low-k limit it acts as a strong diffusive term. For cubic interpolation, the effect is one of a higher order (in k) hyperdiffusion, and as a result cubic interpolation is preferred for horizontal interpolation in modern dynamical cores.

Either form of polynomial interpolation eliminates the high-wavenumber instabilities shown in section 3b and 3c because they both strongly dissipate waves near the Nyquist limit. However, neither the strong diffusion of linear interpolation nor the weaker hyperdiffusion of cubic interpolation eliminates the instabilities described in section 3d. Figures 9 and 10 show growth rates over the (C, k_{hill}) parameter space analogously to Fig. 2 for linear and cubic semi-Lagrangian interpolation respectively, and the same basic pattern is evident.

The growth rates are overall lower for linear interpolation, but this comes at the cost of unacceptable levels of diffusion—especially for tracer-like constituents such as atmospheric moisture content. Additionally, the case of linear semi-Lagrangian interpolation remains unstable at near-integer Courant numbers, with this residual instability caused by

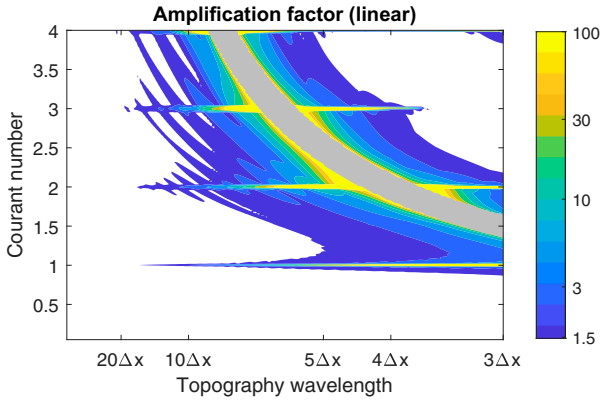

FIG. 9. As in Fig. 2, but computed using linear interpolation for the semi-Lagrangian operators in the computation of the steady-state and linearized operators. A similar pattern of instability exists, but it is somewhat diminished in amplitude because of the dissipation of the semi-Lagrangian operator. Thin bands of instability remain at integer Courant numbers.

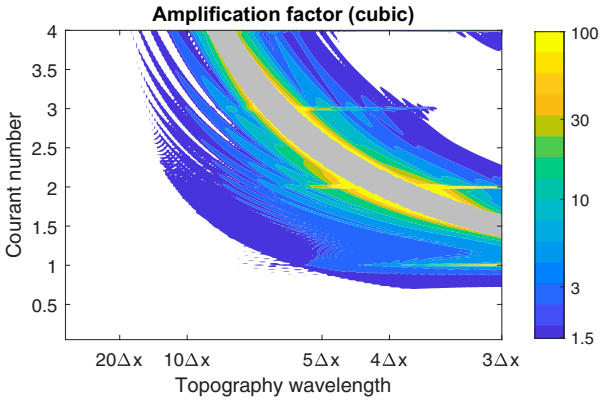

FIG. 10. As in Figs. 2 and 9, but computed using cubic interpolation for the semi-Lagrangian operators in the computation of the steady-state and linearized operators. The resulting region of instability is larger and more intense than that arising from linear interpolation, and the instability bands at integral Courant numbers are less apparent.

abrupt changes of the interpolating stencil when departure points of adjacent cells lie alternately just to the left and just to the right of grid points. However, these bands of instability are very thin, and even if they were observed in a realistic simulation it would take only a small adjustment to the time step to stabilize the simulation.

Cubic interpolation gives instability growth rates that are intermediate between the linear and Fourier interpolation, as would be expected from its higher order.

d. Off-centering

Dynamical cores commonly implement an off-centering parameter to Eqs. (7) for stability purposes. Off-centering adds dissipation to the gravity wave modes of the system, and this dissipation acts to limit the effects of spurious orographic resonance. However, the impact on the resonant growth discussed in this work is less clear.

Adding an off-centering parameter α extends (7) to

$$\phi^+ - \mathrm{SL}(\phi^-) + \Delta t\left[\alpha\phi^+ u_x^+ + (1-\alpha)\mathrm{SL}(\phi^- u_x^-)\right] = 0, \tag{25a}$$

$$u^+ - \mathrm{SL}(u^-) + g\Delta t\left\{\alpha(\phi^+ + H)_x + (1-\alpha)\mathrm{SL}\left[(\phi^- + H)_x\right]\right\} = 0, \text{ and} \tag{25b}$$

$$x - \mathrm{SL}(x) - \Delta t\left[\alpha u^+ - (1-\alpha)\mathrm{SL}(u^-)\right] = 0, \tag{25c}$$

where $\alpha = 0.5$ corresponds to a fully centered time discretization [equal to the unmodified (7)] and $\alpha = 1.0$ corresponds to a fully implicit Euler method for the gravity wave terms. Varying α changes the implicit trapezoidal rule used for integrating the time-continuous shallow-water equations, and the overall method is only second order in time when $\alpha = 0.5$. Nevertheless, $O(\Delta t^2) + (\alpha - 0.5)O(\Delta t)$ may still be small in practice, so small increases in α are often considered negligible.

The off-centering parameter affects the gravity wave modes of (7) by introducing a dissipation proportional to the temporal frequency. Linearizing (7) about no topography with the fluid at rest[5] ($H = 0$, $u = 0 + u'$, $\phi = \phi_0 + \phi'$) gives

$$\phi'^+ + \Delta t\alpha\phi_0 u_x'^+ = \phi'^- - \Delta t(1-\alpha)\phi_0 u_x'^-, \tag{26a}$$

$$u'^+ + g\Delta t\alpha\phi_x'^+ = u'^- - g\Delta t(1-\alpha)\phi_x'^-. \tag{26b}$$

Decomposing ϕ and u in terms of single-mode waves in space and time $\{\exp[\mathbf{i}(kx - \omega t)]\}$ gives the dispersion relation:

$$e^{-\mathbf{i}\omega\Delta t} = \frac{1 \pm \dfrac{2\mathbf{i}}{\Delta x} c\Delta t \sin\left(k\dfrac{\Delta x}{2}\right)(1-\alpha)}{1 \mp \dfrac{2\mathbf{i}}{\Delta x} c\Delta t \sin\left(k\dfrac{\Delta x}{2}\right)\alpha}. \tag{27}$$

In the limit $\alpha = 0.5 + \varepsilon$, this dispersion relation becomes

$$e^{-\mathbf{i}\omega\Delta t} = \frac{\Delta x \pm \mathbf{i}c\Delta t \sin\left(k\dfrac{\Delta x}{2}\right)}{\Delta x \pm \mathbf{i}c\Delta t \sin\left(k\dfrac{\Delta x}{2}\right)} - \epsilon\frac{4c^2\Delta t^2 \sin^2\left(k\dfrac{\Delta x}{2}\right)}{\left[\Delta x - \mathbf{i}c\Delta t \sin\left(k\dfrac{\Delta x}{2}\right)\right]^2} + O(\epsilon^2), \tag{28}$$

[5] Linearizing about a constant $u_0 \neq 0$ gives a dispersion relation of essentially the same functional form, just shifted by $\exp(\mathbf{i}ku_0\Delta t)$, because the semi-Lagrangian advection of u and ϕ fully acts before the off-centering modification.

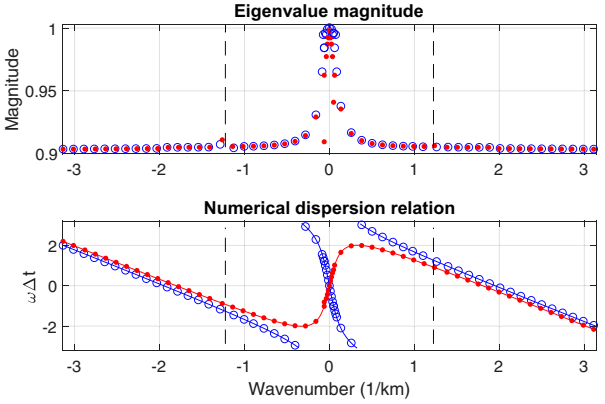

FIG. 11. As in Fig. 5, with an off-centering parameter of α = 0.525. The maximum eigenvalue magnitude is now 1.000 394, leading to a negligible growth rate for the unstable mode, but higher frequency gravity waves have an eigenvalue magnitude of about 0.9, representing a 10% dissipation per time step.

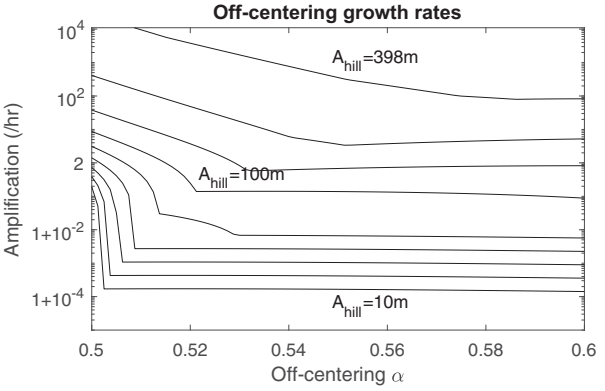

FIG. 12. Growth rates for the most unstable mode of the case discussed in section 3d, for hill heights $10 \leq A_{\text{hill}} \leq 398$ m, spaced evenly on a logarithmic scale, and off-centering parameter $0.5 \leq \alpha \leq 0.6$. The $A_{\text{hill}} = 100$ m line corresponds exactly to that shown in section 3d, and its growth rate is trivial for $\alpha \geq 0.52$. Larger hill heights quickly overcome the damping effect of off-centering.

where the second term gives a dissipation factor proportional to ε and k^2 (in the long wave limit).

Although not directly applicable to the more complicated basic state of flow past topography, this functional form gives us reason to worry: dissipation is relatively weak for longer waves, and the instabilities generated by the mechanism of section 3d have a longwave component.

Testing the interaction of off-centering with the case shown in Fig. 5 initially shows promise—an off-centering parameter as small as $\alpha = 0.525$ effectively eliminates the instability at the cost of dissipating higher-frequency gravity waves, giving the gravity wave spectrum shown in Fig. 11. The maximum eigenvalue in this case remains slightly greater than one (about 1.003), but the resulting growth rate is about 15% h^{-1}, small enough that it would likely be unnoticed in a realistic simulation.

However, one other degree of freedom remains: the topography amplitude.

From the analysis in section 3d, the overall magnitude of the instability is closely related to the ratio of A_{hill} to ϕ_0. Thus far we have worked with $A_{\text{hill}} = 100$ m in dimensional units to ensure that the background flow is still a small perturbation from uniform motion, but mountains tend to be significantly taller than 100 m. Figure 12 depicts the growth rate when A_{hill} and α are modified simultaneously, and the results are not encouraging—a hill amplitude of 398 m recovers the instability with α as large as 0.6.

5. Discussion and conclusions

An instability, driven by the interaction between the steady response of flow over topography and shortwave perturbations, is present in the semi-Lagrangian, time-centered implicit discretization of the shallow-water equations over a significant portion of the Courant number and topographic wavelength parameter space. The most prevalent instabilities are similar to the well-known phenomenon of spurious orographic resonance, but they occur under more general circumstances. These instabilities take their energy from the variations in the background state, transferring energy into their respective unstable modes. The instability strength is proportional to the amplitude of the topography relative to the background layer height ϕ_0.

In time-dependent simulations, the growth of the instability tends to saturate at a level proportional to the difference between the steady-state background flow and a hypothetical no-topography, constant background flow. In a larger-scale model, this sort of saturation would be more likely to appear as "noise" or implausible results rather than model crashes or infinite results, but interactions with other aspects of the model such as physical parameterizations may cause further corruption of the simulation. Without additional feedback, there does not appear to be a direct mechanism to fuel unlimited, exponential growth.

The more specialized instabilities of advective resonance (section 3b) and spatial aliasing (section 3c) involve modes with high spatial frequencies that are not well resolved on the grid. These instabilities are effectively suppressed by modest damping from off-centering or through the use of polynomial interpolation for the semi-Lagrangian advection. In contrast, the temporal resonance described in section 3d consists of an instability with a long-wavelength component and a short-wavelength component comparable to the length scale of the topography itself, and it is much less affected by these sorts of damping mechanisms. Since its oscillations are at least as well resolved as the topography itself, it is also unlikely to benefit from changes exclusively in the semi-Lagrangian operator, such as positivity-preserving interpolation.

The instabilities shown here do not appear to depend strongly on the exact form of the semi-Lagrangian operator, and in particular there seems to be no strong dependency on how the trajectories are computed. The numerical examples presented here all used iteratively computed trajectory calculations, and the form of the instability remains the same with even small numbers of iterations that do not result in fully converged trajectories. Even more extreme adjustments such as time-explicit

trajectory computations (basing the trajectories only on u^-, not shown for brevity) fail to dislodge the instability.

a. Appearance in realistic simulations

If this orographic instability is present in the shallow-water system, then why do operational systems based on this algorithm nonetheless produce successful forecasts and analyses?

The answer may lie in the slow growth rates of the instability and the nature of Earth itself. The problem domain established in section 2c is specifically designed to highlight the instability: it consists of periodic topography with a single wavelength over a relatively large domain. The calculated growth rates over much of the parameter space in Fig. 2 are of the order of a factor of 10 per hour, but the phase speed of gravity waves in this system is about 250 m s^{-1}. In 1 h, a gravity wave will have propagated about 900 km.

As a general rule, Earth's surface does not look like a perfectly periodic, sinusoidal undulation of a single wavelength. A mode that is unstable in one region will likely find itself in a different regime 900 km downstream, having had only a small "contact time" to grow from the instability. Since the unstable modes (e.g., Fig. 6) depend on k_{hill}, a mode that grows in one region is not likely to grow quickly in another.

Higher-resolution, local area simulations, however, are more likely to have topographies that resemble those studied here. The relevant time scale for the instability is proportional to the Courant number, so higher resolution simulations that use a shorter time step would see this instability grow proportionally more quickly in time. Very high-resolution simulations with resolutions finer than 1 km over mountainous regions (with correspondingly large "hill amplitudes") may be particularly susceptible to this instability.

To a researcher or meteorologist looking at the outputs from such a simulation, this instability would likely appear as a stubborn but small amplification of "noise" otherwise present in any output from an operational system. This "noise" would be generated from the dynamical core itself, absent physical parameterizations of moist processes, but since it would occur over mountainous regions the vertical coordinate may receive undue blame.

b. Extensions to the Euler equations

Further tests involving extensions of this analysis to the three-dimensional Euler system will help characterize to what extent current and near-future forecasting models are impacted by this instability.

It seems very likely that these instabilities are still present in the more complicated system. The root cause of the instability is the interaction between gravity waves, which are slowed by the time-centered discretization, advection by the variable background flow, which is treated quite accurately via the semi-Lagrangian approach, and topographic forcing, which produces sum and difference modes. This same set of effects should be present in the three-dimensional system regardless of its choice of vertical coordinate, but adding the extra dimension will also extend the gravity wave spectrum in ways that may allow for more coupling and resonance.

c. Future work

In the shortest term, the approach to controlling these instabilities must be to control the "operating envelope" of the dynamical core by restricting the topographic scales presented to the model. Filtering topography with a characteristic wavelength shorter than 5–6Δx would eliminate the most unstable regions show in Fig. 2, particularly in combination with cubic interpolation in the semi-Lagrangian operator (Fig. 10). This restriction may be unsatisfying as a physical representation, however, since in the presence of a radiative forcing model topography is also responsible for driving anabatic and katabatic winds, which are not fundamentally propagating gravity waves. If instead the time step must be reduced for stability, the criticism of Bartello and Thomas (1996) to the claimed performance advantages of semi-Lagrangian advection will become even more applicable.

In the longer term, the best approach will likely be to revisit the time-stepping algorithm. The semi-implicit approach computes the nonlinear, time-implicit terms through a single linear iteration based on a reference background state, and it appears to be immune to the temporal aliasing instability. However, this comes at the cost of reducing the time accuracy of the system or introducing time extrapolation, which itself comes with stability constraints. There may be room for compromise through a split approach, where a full nonlinear iteration applies to the coarsest features of the solution go through several iterations of a nonlinear solution process and the finest scales are handled semi-implicitly.

Another promising approach may ultimately be to return to an Eulerian formulation of advection, avoiding the instability by treating advection consistently with the discretization of the gravity wave restoring forces. This approach is under study with spectral element codes (Thomas and Loft 2005). Alternatively, the consistency can be enforced at the time-discretization level through exponential propagation as described in Gaudreault and Pudykiewicz (2016). Peixoto and Schreiber (2019) also demonstrates the combination of a semi-Lagrangian treatment of advection with exponential propagation of the other terms in the shallow-water equations.

Ultimately, the potential for this sort of instability must be considered by those designing dynamical cores. These instabilities affect a wide section of the parameter space spanned by Courant numbers and the topography wavelength, but they are invisible to an analysis linearized about an atmosphere at rest.

Acknowledgments. The author would like to thank Drs. Syed Husain and Janusz Pudykiewicz for helpful comments on an early draft of this paper. Additionally, the author would like to thank Dr. Nigel Wood and two other anonymous reviewers for their helpful comments during peer review.

Data availability statement. The Matlab script files used to generate the figures in this paper are available on Github at https://github.com/csubich/subich_mwr_instabilities_2021, under an LGPL license.

REFERENCES

Arakawa, A., and C. S. Konor, 2009: Unification of the anelastic and quasi-hydrostatic systems of equations. *Mon. Wea. Rev.*, **137**, 710–726, https://doi.org/10.1175/2008MWR2520.1.

Bartello, P., and S. J. Thomas, 1996: The cost-effectiveness of semi-Lagrangian advection. *Mon. Wea. Rev.*, **124**, 2883–2897, https://doi.org/10.1175/1520-0493(1996)124<2883:TCEOSL>2.0.CO;2.

Bell, M. J., P. S. Peixoto, and J. Thuburn, 2017: Numerical instabilities of vector-invariant momentum equations on rectangular C-grids. *Quart. J. Roy. Meteor. Soc.*, **143**, 563–581, https://doi.org/10.1002/qj.2950.

Benacchio, T., and N. Wood, 2016: Semi-implicit semi-Lagrangian modelling of the atmosphere: A Met Office perspective. *Commun. Appl. Ind. Math.*, **7**, 4–25, https://doi.org/10.1515/caim-2016-0020.

Boyd, J. P., 2001: *Chebyshev & Fourier Spectral Methods.* 2nd ed. Courier Corporation, 698 pp.

Diamantakis, M., 2014: The semi-Lagrangian technique in atmospheric modelling: Current status and future challenges. *Seminar on Recent Developments in Numerical Methods for Atmosphere and Ocean Modelling, ECMWF*, Shinfield Park, Reading, 183–200, https://www.ecmwf.int/node/9054.

Dubos, T., and F. Voitus, 2014: A semihydrostatic theory of gravity-dominated compressible flow. *J. Atmos. Sci.*, **71**, 4621–4638, https://doi.org/10.1175/JAS-D-14-0080.1.

Durran, D. R., and P. A. Reinecke, 2004: Instability in a class of explicit two-time-level semi-Lagrangian schemes. *Quart. J. Roy. Meteor. Soc.*, **130**, 365–369, https://doi.org/10.1256/qj.03.14.

ECMWF, 2020: IFS Documentation CY47R1. Part IV: Physical processes. IFS documentation, ECMWF, https://doi.org/10.21957/cpmkqvhja.

Eldred, C., and D. Randall, 2017: Total energy and potential enstrophy conserving schemes for the shallow water equations using Hamiltonian methods—Part 1: Derivation and properties. *Geosci. Model Dev.*, **10**, 791–810, https://doi.org/10.5194/gmd-10-791-2017.

Gaudreault, S., and J. A. Pudykiewicz, 2016: An efficient exponential time integration method for the numerical solution of the shallow water equations on the sphere. *J. Comput. Phys.*, **322**, 827–848, https://doi.org/10.1016/j.jcp.2016.07.012.

Gill, A. E., 1982: *Atmosphere-Ocean Dynamics.* Academic Press, 662 pp.

Girard, C., and Coauthors, 2014: Staggered vertical discretization of the Canadian environmental multiscale (GEM) model using a coordinate of the log-hydrostatic-pressure type. *Mon. Wea. Rev.*, **142**, 1183–1196, https://doi.org/10.1175/MWR-D-13-00255.1.

Hortal, M., 2002: The development and testing of a new two-time-level semi-Lagrangian scheme (SETTLS) in the ECMWF forecast model. *Quart. J. Roy. Meteor. Soc.*, **128**, 1671–1687, https://doi.org/10.1002/qj.200212858314.

Husain, S. Z., and C. Girard, 2017: Impact of consistent semi-Lagrangian trajectory calculations on numerical weather prediction performance. *Mon. Wea. Rev.*, **145**, 4127–4150, https://doi.org/10.1175/MWR-D-17-0138.1.

——, ——, 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, https://doi.org/10.1175/MWR-D-18-0438.1.

——, L. Separovic, and C. Girard, 2020: On the need of orography filtering in a semi-Lagrangian atmospheric model with a terrain-following vertical coordinate. *19th Conf. on Mountain Meteorology*, virtual, Amer. Meteor. Soc., 7.2, https://ams.confex.com/ams/19Mountain/webprogram/Paper376266.html.

Izakov, M. N., 1971: On theoretical models of the structure and dynamics of the Earth's thermosphere. *Space Sci. Rev.*, **12**, 261–298, https://doi.org/10.1007/BF00165510.

McDonald, A., 1998: The origin of noise in semi-Lagrangian integrations. Tech. Rep. 55, Met Éireann, 28 pp.

Peixoto, P. S., and M. Schreiber, 2019: Semi-Lagrangian exponential integration with application to the rotating shallow water equations. *SIAM J. Sci. Comput.*, **41**, B903–B928, https://doi.org/10.1137/18M1206497.

——, J. Thuburn, and M. J. Bell, 2018: Numerical instabilities of spherical shallow-water models considering small equivalent depths. *Quart. J. Roy. Meteor. Soc.*, **144**, 156–171, https://doi.org/10.1002/qj.3191.

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, https://doi.org/10.1175/1520-0493(1994)122<0366:SRROSL>2.0.CO;2.

Robert, A., 1981: A stable numerical integration scheme for the primitive meteorological equations. *Atmos.–Ocean*, **19**, 35–46, https://doi.org/10.1080/07055900.1981.9649098.

——, J. Henderson, and C. Turnbull, 1972: An implicit time integration scheme for baroclinic models of the atmosphere. *Mon. Wea. Rev.*, **100**, 329–335, https://doi.org/10.1175/1520-0493(1972)100<0329:AITISF>2.3.CO;2.

Simmons, A. J., and C. Temperton, 1997: Stability of a two-time-level semi-implicit integration scheme for gravity wave motion. *Mon. Wea. Rev.*, **125**, 600–615, https://doi.org/10.1175/1520-0493(1997)125<0600:SOATTL>2.0.CO;2.

Staniforth, A., and J. Côté, 1991: Semi-Lagrangian integration schemes for atmospheric models—A review. *Mon. Wea. Rev.*, **119**, 2206–2223, https://doi.org/10.1175/1520-0493(1991)119<2206:SLISFA>2.0.CO;2.

Thomas, S. J., and R. D. Loft, 2005: The NCAR spectral element climate dynamical core: Semi-implicit Eulerian formulation. *J. Sci. Comput.*, **25**, 307–322, https://doi.org/10.1007/s10915-004-4646-2.

Thuburn, J., and C. Cotter, 2012: A framework for mimetic discretization of the rotating shallow-water equations on arbitrary polygonal grids. *SIAM J. Sci. Comput.*, **34**, B203–B225, https://doi.org/10.1137/110850293.

Ullrich, P. A., P. H. Lauritzen, and C. Jablonowski, 2014: A high-order fully explicit flux-form semi-Lagrangian shallow-water model. *Int. J. Numer. Methods Fluids*, **75**, 103–133, https://doi.org/10.1002/fld.3887.

Williamson, D. L., J. B. Drake, J. J. Hack, R. Jakob, and P. N. Swarztrauber, 1992: A standard test set for numerical approximations to the shallow water equations in spherical geometry. *J. Comput. Phys.*, **102**, 211–224, https://doi.org/10.1016/S0021-9991(05)80016-6.

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, https://doi.org/10.1002/qj.2235.