from *SIAM News*, Volume 32, Number 7

Parallel Computation of Gravity Wave Turbulence in the Earth's Atmosphere

By Joseph M. Prusa, Piotr K. Smolarkiewicz, and Andrzej A. Wyszogrodzki

Gravity waves (GWs) are ubiquitous internal waves in which the restoring force is buoyancy. In the Earth's atmosphere, GWs are generated whenever a parcel of air is disturbed from its equilibrium position. Disturbing sources include topographic features (e.g., mountains), convective instabilities (e.g., thunderstorms), frontal motions, and oscillations of jet streams. When the amplitude of a GW approaches its inverse vertical wavenumber, the wave "breaks" and generates a localized region of intense turbulence.

APPLICATIONS ON ADVANCED ARCHITECTURE COMPUTERS

Greg Astfalk, Editor

Strong GWs near the surface in the lee of mountains are responsible for spectacular down-slope windstorms, such as the Boulder chinook, the Santa Ana winds, and the Yugoslavian bora. Such storms often generate hurricane-force winds and are a serious threat to personal safety, buildings, and surface and air transportation. At higher altitudes, GWs can be of sufficient magnitude to alter the mean state of the atmosphere.

The wave drag induced by the breaking of small-scale GWs (~10 km) is now thought to be an important issue in improving subgrid parameterizations for global atmospheric flow models. More generally, the propagation and evolution of GWs may be a key component in determining the spectral energy of the atmosphere at mesoscales (~100 km).

Physical Problem

We modeled the evolution of an internal gravity wave packet generated by a disturbance in the lower atmosphere and its subsequent breaking at higher altitudes. For direct resolution of as many of the physically important length scales as possible, a grid spacing of 380 meters was used throughout a $200 \times 30 \times 110$ km (zonal, meridional, vertical, respectively) domain. The result was a $544 \times 80 \times 291$ computational grid of 12.7 million points. The solution, generated on a CRAY T3E, showed spectacular detail (revealing the evolution of multiple subranges of turbulence) and clearly demonstrated the advantages of highly parallel technologies.

The disturbance was modeled as Gaussian in its spatial–temporal evolution, with a maximum forcing amplitude of 200 meters at $t = 2$ hours and peak half-widths of 2 km and 1 hour. The basic state of the atmosphere was one of uniform zonal wind ($u_o = -32$ meter-second^{-1}), stability ($N = 0.020$ second^{-1}), and density scale height ($H = 6.63$ km). N represents the Brunt–Väisällä, or buoyancy, frequency; it is the frequency of oscillation of a parcel of air undergoing adiabatic, stable vertical displacements about an equilibrium state. This basic state and forcing favored the development of a monochromatic, 2D primary (convective) instability at 130 minutes [4]. This was followed within a Brunt–Väisällä period (~5 minutes) by a secondary (spanwise) instability that induced 3D flow structures. The ensuing wavebreaking generated a highly inhomogeneous, anisotropic turbulence that developed in accord with the physics of a wave packet propagating into a very deep model atmosphere—it was not initialized according to any a priori turbulence model, nor was it constrained by domain size.

Some idea of the inhomogeneity of the wavefield can be gleaned from Figure 1, which shows a contour density plot of the natural log of the potential temperature (θ) field ($\ln(\theta) \sim$ entropy). The vertical plane of this plot is perpendicular to the zonal flow. The plot shows the wavefield to be homogeneous in the spanwise direction (left to right is south to north) but inhomogeneous in the vertical. (Notice that the complete altitude range is $15 \le z \le 125$ km. The regions above and below those shown in Figure 1 are very smooth and characterized by constant stratification.) Similar inhomogeneity occurs in the zonal direction (west to east). The anisotropy of the turbulence can be seen in Figure 2, which shows the velocity derivative skewness S_i:

$$S_i = -\left\langle \left(\frac{\partial u_i}{\partial x_i} \right)^3 \right\rangle \Big/ \left\langle \left(\frac{\partial u_i}{\partial x_i} \right)^2 \right\rangle^{3/2}$$

where $i = 1$, 2, or 3, denoting the zonal (Figure 2(a)), meridional (Figure 2(b)), or vertical (Figure 2(c)) wind and coordinate, respectively. Here $\langle \cdot \rangle$ denotes an averaging operator over a volume of space. For isotropic turbulence, $S_u = S_v = S_w \sim 0.5$ [3]. The zonal velocity derivative skewness S_u shown in Figure 2(a) matches this value in the region of vigorous breaking ($-70 \le x \le 10$ km). In Figure 2(b,c), however, smaller mean values of S_v and S_w are seen in this region, with S_v being very close to zero. This indicates that the vertical and, especially, the meridional velocity derivatives are more symmetrical functions about their mean values (i.e., they are Gaussian-like). These differences in S_i may result from preferential vortex stretching induced by the stratification of the basic state [3] as well as by the 2D nature of the primary instability.

We assessed the evolution of the turbulence by examining 1D energy spectra. At the onset of the primary instability ($t \sim 130$ minutes), the spectra showed energy strongly concentrated at a fundamental mode, with zonal wavelength $\lambda_x = 15.5$ km. A buoyancy subrange [11] rapidly developed upscale of the fundamental mode. With the onset of secondary instability ($t \sim 140$ minutes), an inertial subrange appeared upscale of the buoyancy subrange. The spectra also showed a second "inertial-like" regime at the longest wavelengths ($\lambda_x \geq 30$ km) for $120 \leq t \leq 150$ min. This time interval closely matches the period in which significant accelerations to the zonal mean wind occur and is thus consistent with a 2D inverse energy cascade [2]. For $t \geq 150$ min, breaking has extensively modified the background state and obliterated the primary wavefield.

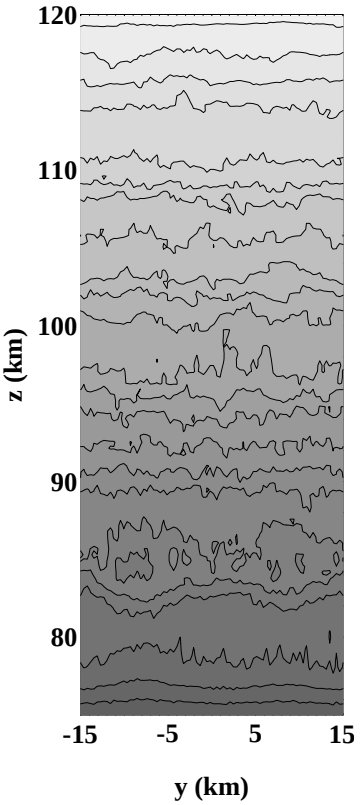

Figure 1. *Contour density plot of the potential temperature field ln(θ) in the vertical yz (meridional × vertical) plane at zonal location x = –35 km and time t = 155 minutes showing a region of vigorously breaking waves. Lighter grays denote larger values.*

Model Description

The underlying solver is an incompressible-type fluid model cast in a curvilinear rotating framework, with a subgrid-scale turbulence parameterization and water substance phase-change processes included. The distinctive aspect of our model is its numerical design, which incorporates a two-time-level, either semi-Lagrangian [7] or Eulerian [8], nonoscillatory forward-in-time (NFT) algorithm. The finite-difference approximations to the resulting trajectory-wise or point-wise integrals of the governing fluid equations have at least second-order accuracy. In the work described here, the semi-Lagrangian option was used.

The conservation laws for the dependent variables of the model can all be written in the compact Eulerian form:

$$\frac{\partial \rho^* \psi}{\partial t} + \nabla \cdot (\mathbf{v}^* \psi) = \rho^* F(\mathbf{\Psi}) \tag{1}$$

or, equivalently, in the compact Lagrangian form:

$$\frac{D\psi}{Dt} = F(\mathbf{\Psi}) \tag{2}$$

Here, ψ denotes any of the three Cartesian velocity components (u, v, w), the potential temperature, and water substance mixing ratios (e.g., vapor, cloud water, rain), as well as turbulent kinetic energy. $\rho^* \equiv \bar{\rho} G$ is the reference (Boussinesq-type) density profile premultiplied by the Jacobian of the coordinate transformation (from the Cartesian to the terrain-following, time-dependent, curvilinear framework). The advective velocity $\mathbf{v}^* \equiv \rho^* \mathbf{V} \equiv \rho^* (u, v, \omega)$, with ω denoting the "vertical" component of transformed (contravariant) velocity, satisfies the anelastic (filtered to remove sound waves) mass conservation law:

$$\frac{\partial \rho^*}{\partial t} + \nabla \cdot \mathbf{v}^* = 0 \tag{3}$$

and D/Dt ($\equiv \partial / \partial t + \mathbf{V} \cdot \nabla$) in (1) denotes the material derivative. The time derivative must be retained in (3) because of the time variation in the coordinate transformation [4]. The associated $F(\mathbf{\Psi})$ terms on the right-hand side of (1) and (2) are generally functionals of the vector $\mathbf{\Psi}$ of all dependent variables ψ. They represent the sum of the resolved and subgrid-scale parts of the total forcings. In the momentum equations the resolved terms include pressure-gradient forces, Coriolis accelerations, and buoyancy force, as well as wave-absorbing devices in the vicinity of open boundaries. In the thermodynamic equations the resolved terms include heat and moisture sink/sources resulting from the phase changes of water and the wave absorbers near the boundaries. The subgrid-scale forcing terms are fairly complex but standard. We employ a turbulence model based on the prognostic turbulent kinetic energy (TKE) equation or, optionally, its abbreviated version—the celebrated Smagorinsky model.

The integration of the discrete equations over a timestep uses a regular unstaggered mesh. We write the finite-difference approximations to (1) and (2) in the compact form:

$$\psi_{\mathbf{i}}^{n+1} = LE(\tilde{\psi}) + 0.5 \Delta t F_{\mathbf{i}}^{n+1} \tag{4}$$

Here, LE denotes either the advective semi-Lagrangian or the flux-form Eulerian NFT transport operator; $\tilde{\psi} \equiv \psi^n + 0.5 \Delta t F^n$ and the indices $\mathbf{i}$ and n have the usual meaning of spatial and temporal location on a (logically) rectangular Cartesian mesh. Completion of the model timestep requires the F^{n+1} values of forcings in (4). GW absorbers, Coriolis accelerations, condensation, and pressure-gradient forces are treated implicitly, whereas subgrid-scale terms and slow phase-change tendencies (such as rain formation or evaporation) are treated explicitly. The implicitness of the pressure-gradient forces is essential in that it enables projection of the preliminary values $LE(\tilde{\psi})$ onto solutions of the continuity equation (3). Here, this requires a straightforward algebraic inversion of the linear system (4) and the formulation of the boundary value problem for pressure implied by the continuity constraint (3).

The resulting elliptic equation is solved (subject to appropriate boundary conditions) via the generalized conjugate-residual approach—a preconditioned nonsymmetric Krylov solver [9]. The numerical stability of the computations is controlled by proper limiting of Courant or Lipschitz numbers $C = \|\Delta t \mathbf{V} / \Delta \mathbf{X}\|$ and $L = \left\|\Delta t\left(\partial \mathbf{V} / \partial \mathbf{x}\right)\right\|$, respectively, for the Eulerian or semi-Lagrangian variants of the model.

The problem was executed in two dimensions on a 544 × 1 × 291 grid for 120 minutes of the physical time. At this point, several Brunt–Väisällä periods before the onset of primary instability, we created the 3D domain by repeating the solution in the spanwise direction *y* and seeding the buoyancy field with low-amplitude (1% of the basic state) white noise. Further computations continued in 5-minute increments of physical time. The time chosen for 3D seeding was carefully selected on the basis of data generated in earlier 2D (cases LR, G1, and HR of [4]) and 3D [5] experiments. The run was terminated at 180 minutes, the point at which breaking had consumed the zonal extent of the domain.

No explicit subgrid-scale viscosity was employed in this simulation. Instead,the removal of energy at the grid scale was effected with the monotonicity option in the semi-Lagrangian interpolator [7]. The foundation for this option is a topological constraint whereby no two streamtubes are allowed to intersect. Essentially, energy is removed at the grid scale to the extent needed to avoid local negative entropy production [4]. This corresponds well to the Kolmogorov micro-scale, which is the same order of magnitude as the grid size at the initial altitude of breaking.

We conducted additional 3D simulations at a lower resolution (625 m) on the CRAY T3D at the National Center for Atmospheric Research (NCAR), using both the TKE and Smagorinsky subgrid-scale models. Compared with the monotonicity option alone, the subgrid-scale results appeared to be overly damped with respect to fine-scale features, and did not show inertial subranges in spectra at the highest wavenumbers. Although the subgrid models can be tuned to produce an inertial subrange, the nonstationarity of the turbulence makes such tuning a tedious task. Our NFT method with either monotonic interpolation (semi-Lagrangian model) or flux-form advection (Eulerian model) bypasses this difficulty.

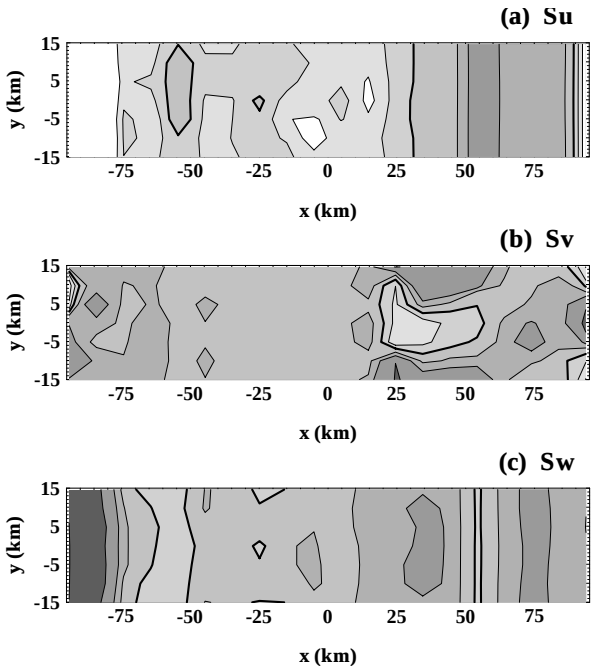

Figure 2. *Contour density plot of velocity derivative skewnesses S_u, S_v, and S_w in the horizontal xy (zonal × meridional) plane at altitude z = 100 km and time t = 155 minutes. The averaging volumes for the S_i computations were (10 × 5 × 7.5 km) in (zonal × meridional × vertical) extent. The region of vigorous wavebreaking/broken waves is –70 ≤ x ≤ 10 km. Contours are shown at S_i = –1, –0.35, 0, 0.35 (*bold*), 0.65, and 1. Lighter grays denote larger values.*

Parallelization Versus Portability Strategy

In recent years a number of new machines based on massively parallel processing (MPP) technology have become available for large-scale computations in science and engineering. While these machines make possible levels of computation that were previously unthinkable, a number of important problems can still be addressed successfully with conventional vector supercomputers, single-processor scalar workstations, and even modern PCs. For the best utilization of the wide range of computing resources now available for science and engineering, application codes require a high degree of portability between different systems.

The Fortran code adopted for the purpose of this study has been described in the literature [1,6]. The traditional CFL stability condition for the Eulerian algorithm limits local communications to nearest neighboring points on the mesh; the semi-Lagrangian algorithm admits Courant numbers exceeding unity and results in irregular communication patterns extending over a number of grid points. To take full advantage of MPP systems, we used a single-program, multiple-data (SPMD) message-passing approach to implement the solver. This parallelization approach was found to be superior to the use of High-Performance Fortran (HPF).

For the GW experiment, we used a two-dimensional horizontal decomposition of the grid (the default for our applications) and explicitly dimensioned each array to contain a subgrid plus extra space for a copy of the neighboring processors' boundary cells. The extra boundary points, often referred to as "halo cells," are used to minimize communication. The number of halo cells depends on the local stencils used in the model algorithm and on the maximum Courant number. In the simulations reported here, $C \leq 3$.

When necessary, the halo cell information is updated by having each processor exchange information with its neighbors. This communication process can be further economized by having halos updated only partially, with the selected information exchanged between processors depending on the finite-difference algorithms employed. Reduction operations like sums and extrema, unavoidable in fluid models, require global exchanges of information between all processors. In general, to exchange messages between processors we use the portable and widely supported MPI (Message-Passing Interface) standard. However, on the CRAY T3D and T3E we also have the option of using SHMEM (shared–distributed memory data-passing support) library routines.

To enhance the portability of the code, we use the same halo-update subprograms on the distributed- or shared-memory parallel architectures as on single-processor machines. On single-processor and shared-memory platforms, all updates are elementary. They employ one processor for the total domain dimension, with halos used to set appropriate conditions at the domain boundaries.

In this case, there is no need for an explicit message-passing protocol, and only selected parts of total arrays are rewritten to halo cells on the processor.

Input/output (I/O) operations also cause serious problems for code portability. In general, outputted fields should be independent of the machine size and the number of parallel processors used in simulations. This is especially important for postprocessing of large simulations. The output files must also be readable on different platforms with different binary file formats. In our code, one processor performs all I/O communication operations, collecting arrays from certain processors and distributing them among others. These I/O operations depend on the particular computer at hand and are important to the overall model performance. Keeping the total grid array on one processor does have the disadvantage of limiting the size of the application. This drawback is more than offset by the simplicity in coding possible with this approach, which seems to offer optimal performance.

Performance

We used the 512-processor CRAY T3E at the National Energy Research Scientific Computing Center (NERSC) in Berkeley to conduct our numerical experiments. Table 1 outlines the computational cost of the GW experiment performed versus the overall model performance (measured by the wall-clock time) as functions of the simulated physical time, timestep Δt, number of timesteps N_t, and average number of iterations N_{it} in the elliptic Krylov solver (per timestep) per 5-minute portion of the experiment.

In addition to summarizing elementary aspects of the model efficiency, this table illustrates the important point that overall model performance (in addition to the relative performance of various model components, such as advective transport and the pressure solver) is an elusive entity—it is a complicated function of the simulated flow. Consider, for instance, that at 145 minutes, at the onset of vigorous wavebreaking, accuracy arguments [4,7] dictate halving the timestep. As the flow becomes more quiescent following the onset of breaking, however, the elliptic solver converges (see [10] for a discussion of the convergence criteria) after only a third as many iterations. Table 2 gives a similar set of data for an earlier, lower-resolution (at 625 m) but otherwise identical, experiment run on a 24-processor CRAY J90 at NCAR in multitasking mode [5].

Although the J90 experiment was more than an order of magnitude smaller than the T3E experiment, the total J90 wall-clock time was 20 times greater than that for the T3E. A more objective assessment of the performance on the GW experiment can be obtained with an elementary analysis using Amdahl's law and a comparison based on the CPU times listed in Tables 1 and 2. Given a typical performance of 75 Mflops (well established for our code) on one processor of the J90, the comparison indicates that an average speed of 10 Gflops was achieved on the T3E with 99.0% parallelization of the code. This level of parallelization is especially noteworthy because of the elliptic Krylov solver that lies at the heart of the computations. Additional information, particularly with regard to scalability and cross-platform performance issues, can be found in [6].

Physical Time (minutes)	Δt (seconds)	N_t	N_{it}	Wall-clock (seconds)	CPU (seconds)	S_u
125–130	5	60	32	1156	583159	504
145–150	2.5	120	19	1500	757776	505
165–170	2.5	120	12	1278	645268	505

Table 1. *GW-breaking experiment on the CRAY T3E. Wall-clock time, CPU time, and speedup (S_u) history as functions of the simulated physical time, timestep Δt, and number of iterations in the pressure solver for selected 5-minute portions of the experiment.*

Physical Time (minutes)	Δt (seconds)	N_t	N_{it}	Wall-clock (seconds)	CPU (seconds)	S_u
125–130	10	30	22	4531	7584	1.7
145–150	5	60	25	36060	16002	0.4
165–170	4	75	15	9326	17984	1.9

Table 2. *Same as Table 1 but for lower-resolution experiment on the CRAY J90.*

Regardless of any objective model-performance measures, the single most important outcome of this study cannot be overstated: In our earlier GW experiments on the CRAY J90 at NCAR, simulation of 5 minutes of physical time required several days (including time spent waiting in economy queues). The present experiments on the 512-processor T3E, on a grid an order of magnitude larger, were executed essentially overnight for the same 5-minute period of simulated physical time!

Acknowledgments

The authors thank William Anderson for his advice on parallelization issues and comments on an earlier version of the manuscript. This work has been supported in part by the Department of Energy's Computer Hardware, Advanced Mathematics, Model Physics (CHAMMP) research program and by National Science Foundation grant #ATM9616811. The use of the 512-PE CRAY T3E at NERSC and of the 32-PE CRAY T3E at ICM is gratefully acknowledged.

References

[1] W.D. Anderson and P.K. Smolarkiewicz, *A comparison of high performance Fortran and message-passing parallelization of a geophysical fluid model*, in Parallel Computational Fluid Dynamics: Algorithms and Results Using Advanced Computers, Schiano et al. (eds.),

Elsevier Science, 1997, 384–391.

[2] R.H. Kraichnan, *Inertial ranges in two-dimensional turbulence*, Phys. Fluids, 10 (1967), 1417–1423.

[3] M. Lesieur, *Turbulence in Fluids*, in Fluid Mechanics and its Applications, 40, Kluwer Academic Publishers, Dordrecht, 1997, 515.

[4] J.M. Prusa, P.K. Smolarkiewicz, and R.R. Garcia, *Propagation and breaking at high altitudes of gravity waves excited by tropospheric forcing*, J. Atmos. Sci., 53 (1996), 2186–2216.

[5] J.M. Prusa, R.R. Garcia, and P.K. Smolarkiewicz, *Three-dimensional evolution of gravity wave breaking in the mesosphere*, Preprint Vol., 11th Conference on Atmospheric and Oceanic Fluid Dynamics, Tacoma, Washington, June 23–27, 1997, J3–J4.

[6] J.M. Prusa, P.K. Smolarkiewicz, and A.A. Wyszogrodzki, *Simulations of gravity wave induced turbulence using 512 PE Cray T3E*, Preprint Vol., 2nd International Workshop on Software Engineering and Code Design in Parallel Meteorological and Oceanographic Applications, Scottsdale, Arizona, June 15–18, NASA GSFC/CP–1998–206860, 1998, 139–159.

[7] P.K. Smolarkiewicz and J.A. Pudykiewicz, *A class of semi-Lagrangian approximations for fluids*, J. Atmos. Sci., 49 (1992), 2082–2096.

[8] P.K. Smolarkiewicz and L.G. Margolin, *On forward-in-time differencing for fluids: Extension to curvilinear coordinates*, Mon. Wea. Rev., 121 (1993), 1847–1859.

[9] P.K. Smolarkiewicz and L.G. Margolin, *Variational solver for elliptic problems in atmospheric flows*, Appl. Math. Comp. Sci., 4 (1994), 527–551.

[10] P.K. Smolarkiewicz, V. Grubisic, and L.G. Margolin, *On forward-in-time differencing for fluids: Stopping criteria for iterative solutions of anelastic pressure equations*, Mon. Wea. Rev., 125 (1997), 647–654.

[11] J. Weinstock, *Theoretical gravity wave spectrum in the atmosphere: Strong and weak wave interaction*s, Radio Sci., 20 (1985), 1295–1300.

Joseph M. Prusa (prusa@iastate.edu) is an associate professor in the Department of Mechanical Engineering at Iowa State University. Piotr K. Smolarkiewicz (smolar@ncar.ucar.edu) is a senior scientist in the Mesoscale and Microscale Meteorology Division at the National Center for Atmospheric Research. Andrzej A. Wyszogrodzki (andii@ncar.ucar.edu) is a graduate student in the Department of Physics at the University of Warsaw.