# CAP Phase II Simulations for the Air Force HEL-JTO Project: Atmospheric Turbulence Simulations on NAVO's 3000-Processor IBM P4+ and ARL's 2000-Processor Intel Xeon EM64T Cluster

J. Werne[1], T. Lund[1], B.A. Pettersson-Reif[2], P. Sullivan[3], and D.C. Fritts[1]

[1]Colorado Research Associates, NorthWest Research Associates, Inc., Boulder, CO, USA
[2]Norwegian Defence Research Establishment, Kjeller, Norway
[3]Meso and Microscale Meteorology, National Center for Atmospheric Research
Boulder, CO, USA

**Abstract**

Shear turbulence induced by the Kelvin-Helmholtz (KH) instability in a stratified fluid is simulated in support of the Air Force High-Energy-Laser Joint Tactical Office (HEL-JTO) project using the new 3000-processor NAVO IBM Power 4+ system (Kraken). The results are used to 1) compare with and improve a dynamic LES method we have developed, and 2) provide the high-resolution simulation component of a new subgrid-scale (SGS) model we have developed for optical turbulence forecasting. We also discuss the relative performance of Kraken and the new 2000-processor Intel XEON EM64T cluster (Jvn) at ARL, pointing out differences which amount primarily to inter-node network performance. We suggest "% of wall time spent communicating" as a standard DoD benchmark criterion for very large distributed systems, as other more conventional criterion do not necessarily represent this important system and algorithm metric.

# 1 Introduction

The main objective of our Air Force HEL-JTO project is to help develop a reliable Atmospheric Decision Aid (ADA) to support high-altitude platforms, such as the Airborne Laser (ABL), the High Altitude Airship (HAA), and Unmanned Aerial Vehicles (UALs). The ADA will be integrated with the Air Force Weather Agency (AFWA) mesoscale forecasting model, *i.e.*, the WRF (Weather Research and Forecasting) model (`www.wrf-model.org`). The major challenge for this effort is description of the small vertical length scales associated with atmospheric clear air turbulence (CAT). In the upper troposphere and lower stratosphere, CAT results from internal gravity-wave breaking and shear instability with vertical scales predominantly between 1 m and 100 m, with larger layers reaching O(km). These are the length scales of the initiating wave and shear phenomena. Turbulence occurs on smaller scales and can reach larger horizontal scales via backtransfer. Since weather forecast models typically operate with approximately 300 m vertical resolution aloft, they cannot explicitly resolve the relevant dynamics, and such dynamics must therefore be modeled with an SGS model. Furthermore, because the mesoscale model's 300 m vertical resolution is too coarse to even describe the outer scales of motion for wind-shear- and wave-instability processes, the modeling challenge is much more severe than in conventional turbulence parameterization algorithms designed to filter resolved-scale motions in the inertial range of turbulence.

Our approach is to include the effects of wave breaking and shear instability via a probabilistic SGS model. We feel this is the most appropriate framework given the coarseness of the mesoscale-model resolution. An example of such a model for unresolved Reynolds stress $\tau_{ij}$ under stable atmospheric conditions is

$$[\,\tau_{ij}, Ri\,|\,Ri_M\,] \propto [\,Ri_M\,|\,Ri\,,\,\tau_{ij}\,]\,[\,\tau_{ij}\,|\,Ri\,]\,[\,Ri_M\,] \tag{1}$$

where quantities in brackets represent probability density functions (PDFs) and vertical bars indicate conditional probabilities. Equation (1) is derived from Bayes Theorem (see *e.g.*, Bernardo and Smith, 1994). The *joint posterior distribution* $[\tau_{ij}, Ri|Ri_M]$ is our SGS model for the unresolved Reynolds stress $\tau_{ij}$ and stability profile $Ri(z)$ given the model-predicted stability profile $Ri_M$. The *likelihood distribution* $[Ri_M|Ri, \tau_{ij}]$ is the stability profile $Ri_M(z)$ predicted by the model when the actual stability profile and Reynolds stress are $Ri(z)$ and $\tau_{ij}$. We can estimate $[Ri_M|Ri, \tau_{ij}] \approx [Ri_M|Ri]$ from observational data (*e.g.*, by filtering raw-insonde data). The remaining distributions $[\tau_{ij}|Ri]$ and $[Ri]$ are *prior distributions* which we can estimate from high-resolution DNS/LES results and atmospheric measurements, respectively. The method is flexible. For example, other SGS models can also be constructed, *e.g.*, $[u_i\theta, Ri|Ri_M]$, $[C_n^2, Ri|Ri_M]$, $[a_i, Ri|Ri_M]$, $[\sigma^2, Ri|Ri_M]$, *etc.*, where $u_i\theta$ is the heat flux, $C_n^2$ is the index-of-refraction 2nd-order structure constant, $a_i$ is an HAA acceleration component, and $\sigma^2$ is the Rytov variance along a path from a particular location, to name just a few.

In order to develop the needed physics *prior distributions* using a sufficient parameter survey (*i.e.*, appropriate ranges of $Ri$ and possibly Reynolds number $Re$ as well as background flow profiles and coupled stratified phenomena), an efficient and flexible large-eddy simulation (LES) code is needed. But before computing LES solutions to populate the required frequency distributions (*e.g.*, $[\tau_{ij}|Ri]$, *etc.*), we must first validate the LES algorithm used for atmospheric instability and turbulence dynamics under stably stratified conditions.

We have begun this process for a dynamic LES code we have developed via comparisons with very-high-resolution direct numerical simulations, DNS (Werne and Fritts 1999a,b, 2000), which have been validated previously (*e.g.*, Werne and Fritts 2000, Gibson-Wilde *et al.* 2000, Kelley *et al.* 2005). With this procedure, we have significantly improved the LES model presented in §3 from its initial formulation, but our results so far indicate that further improvement is possible and likely necessary. Unfortunately, further progress with existing DNS solutions is virtually impossible given their relatively small size, and hence, the relatively small spatial statistical sample the solutions represent.

The DoD CAP program has allowed us to overcome this difficulty by facilitating significantly larger DNS solutions (*i.e.*, a factor of 24 times larger than we achieved previously). The solutions obtained under the CAP program represent the largest and highest-resolution stratified wind-shear turbulence simulations ever conducted. They are an invaluable resource for our Air Force project, sufficient to further guide our LES developments, but also capable of directly evaluating physics *prior* distributions for the few cases we were able to obtain under the CAP program. This introduces a significant advantage over our original research plan, because it allows us to implement an end-to-end validation path for critical components of our modeling program.

# 2    Direct Numerical Simulations

## 2.1    Problem formulation

The problem studied is shear instability initiated with a hyperbolic tangent velocity profile $U = U_o \tanh(z/h)$ and linear temperature profile $T = \beta h$. $U_o$, $h$, and $\beta$ are the wind amplitude, initial half-shear-layer depth, and mean background temperature gradient, respectively. $z$ denotes the vertical direction; we use the convention that $x$ and $y$ are the streamwise and spanwise horizontal components. We also use $(u, v, w)$ for the $(x, y, z)$ velocity components. Below we employ units for space, time, velocity and temperature of $h$, $h/U_o$, $U_o$ and $\beta h$, respectively. Vorticity has units of $U_o/h$.

We apply the Boussinesq approximation to the Navier-Stokes equations, and solve the resultant equations of motion in a Cartesian geometry. Side boundaries are periodic, while top and bottom boundaries are impenetrable and stress-free. Three non-dimensional parameters specify the flow. They are the Reynolds number $Re = U_o\, h/\nu$, which quantifies the relative magnitude of the non-linear and diffusion terms; the Richardson number $Ri = g\alpha\beta h^2/U_o^2$, which is a measure of the fluid layer's stability; and the Prandtl number $Pr = \nu/\kappa$, which is a ratio of diffusion coefficients for momentum and heat. Here $g$, $\alpha$, $\nu$, and $\kappa$ are acceleration due to gravity, thermal expansion coefficient, molecular viscosity, and molecular diffusivity, respectively. We choose the values $Re = 2500$, $Pr = 1$, and the following three values for $Ri$: 0.05, 0.15, and 0.2. We note that when $Ri > 0.25$, the flow is stable and Kelvin-Helmholtz roll-up will not occur.

We employ domains of size $4\lambda \times 2\lambda \times 2\lambda$, where $\lambda$ is the wavelength for the most unstable asymptotic eigen mode for the KH instability. The streamwise length of the domain is chosen to admit four KH vortex tubes (*i.e.*, four KH billows). The spanwise width is selected to be sufficiently large to describe the breakdown of the KH billows at late times. The domain depth is selected so as to displace the numerical boundaries sufficiently from the turbulent mid-layer. Motion is initiated with the most unstable eigen mode with a vorticity amplitude of 0.07 (*i.e.*, 7% of the maximum initial mean vorticity) and a Kolmogorov noise field with vorticity amplitude 0.014.

The algorithm we use to solve the equations of motion is a pseudo-spectral method in which field variables are represented by their Fourier series, and time integration is handled by the 3rd-order Runga-Kutta (RK3) method of Spalart *et al.* (1991). Incompressibility is enforced via a two-streamfunction formalism in which the streamfunctions used are related to the vertical velocity and vertical vorticity fields. Nonlinear terms are treated explicitly by multiplication in physical space, while linear terms are handled implicitly. Real FFTs are conducted to move between physical and Fourier space. In order to take full advantage of the speed associated with cache memory, FFTs are always performed on contiguous data. This requires data transposes to successively rotate the $x$, $y$, and $z$ directions into the first array index. For the last step, a global transpose requiring all-to-all communication among processors is needed. For more details of the numerical algorithm, please refer to Julien *et al.* (1996).

## 2.2 Validation of the DNS Solutions

Results obtained previously with the DNS in a much smaller domain of size $\lambda \times \lambda/3 \times 2\lambda$ duplicated atmospheric turbulence measurements for the 2nd-order structure constants $C_U^2$ and $C_T^2$ associated with the streamwise velocity $U$ and the temperature $T$. The 2nd-order structure functions for $V$ and $W$ (spanwise and vertical velocities) were also obtained, though no atmospheric measurements were available at the time for validation. The slope of the 2nd-order structure functions and the inner scale were also reported, and they reproduced available atmospheric measurements, though they also reported a tendency to vary between the expected value of 2/3 and a smaller value of 2/5. See Werne and Fritts (2000) for details and references. Subsequent to our first publication, new aircraft measurements have been reported (Wroblewski *et al.* 2003) which verify our computed $C_V^2$ and $C_W^2$ values, as well as the occurrence of structure-function exponents of both 2/5 and 2/3 (compare results in Wroblewski *et al.* 2003 with Werne and Fritts 2000).

Because of these and other validation results comparing our DNS solutions with atmospheric measurements, we are confident that the DNS can be used as an effective guide for our LES modeling efforts, but the smaller domain originally used proved to be too small to adequately constrain the statistics needed for our LES work. This is not the case with the new DNS solutions reported here.

## 2.3 Time Evolution for Ri=0.05 and 0.2

Figure 1a shows the time evolution of the volume-integrated kinetic ($KE$, solid curve) and an estimate of the potential energy ($PE$, dashed curve) per unit volume for our most turbulent solution, *i.e.*, the case with $Ri = 0.05$. The background mean velocity and temperature have been removed, hence,

$$KE = (L_x L_y h)^{-1} \int \left[ \left( U - \overline{U}(z) \right)^2 + V^2 + W^2 \right] d^3x \qquad (2)$$

$$PE = (L_x L_y h)^{-1} Ri \int \left( T - \overline{T}(z) \right)^2 d^3x . \qquad (3)$$

The normalization factor $L_x L_y h$ includes the horizontal extent of the layer $L_x L_y$ in order to correct for the effect domain size has on the integrated energy. We use $h$ for the vertical length instead of $L_z$ because the turbulent motions do not fill the full depth of the simulated domain. Figure 1b shows the time evolution of the maximum values of the three vorticity components $\omega_i$, which indicate when turbulent fluctuations and mixing are active.

The initial peaks in $KE$ and $PE$ occur at time $t \approx 37$ when the primary KH billows reach their maximum amplitude. For $Ri = 0.05$, this amplitude, as measured by the depth of the billows, is approximately $6h$. By this time the billows have completed roughly two revolutions in near-solid-body rotation and effectively wrapped four alternating layers of vorticity and temperature around their outer edges. This occurs when the flow is still laminar and predominantly two dimensional (2D), though weak 3D variation is evident inside each of the billow cores; see Figure 2a. Strong three dimensionality develops for $Ri = 0.05$ during $45 < t < 55$ when the layers of alternating vorticity and temperature enveloping the billows
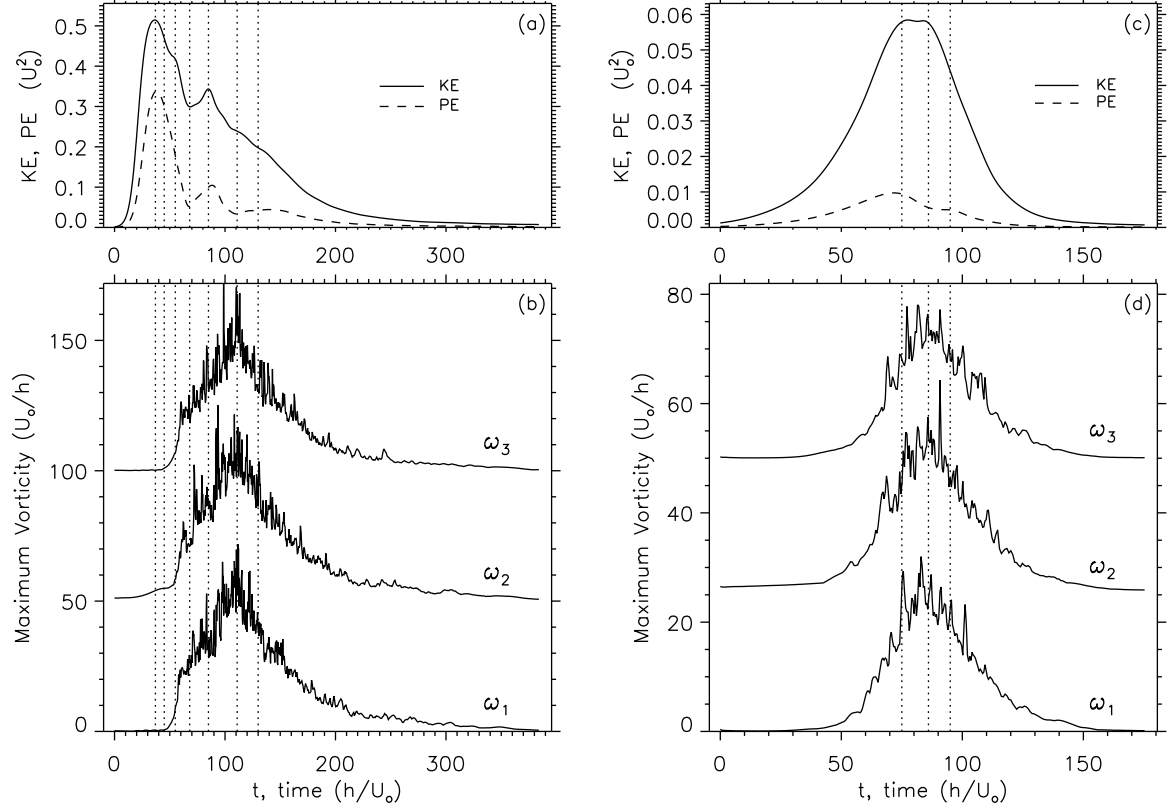
Figure 1: Time evolution of fluctuation kinetic energy KE and an estimate for the potential energy PE (panels a and c) and the maximum of each vorticity component (panels b and d). Panels a and b are for Ri=0.05. Panels c and d are for Ri=0.2.

become unstable to pairs of counter-rotating vortex tubes, which are oriented orthogonally relative to the billow cores (see Figure 2b). This secondary instability triggers the onset of turbulence in the outer edges of the KH billows (see Figure 2c), which expands laterally until the full horizontal extent of the layer becomes turbulent.

The second peak in $KE$ and in $PE$ exhibited in Figure 1a at $t \approx 85$ coincides with the moment turbulence reaches the interior of the billow cores; see Figure 2d. It indicates when the combined intensity of turbulence and the KH billows is maximal. At this point the billow energy rapidly decreases, while the energy of the fluctuations grows more slowly. The turbulent fluctuations peak at $t \approx 111$, precisely when the maximum vorticity is realized; see Figure 2e. This time also coincides with the local minimum in $PE$ because the vigorous turbulent mixing homogenizes the temperature field and eradicates temperature variance for the billow and fluctuations (see Figure 1a), and for the mean as well.

After time $t \approx 111$, the maximum vorticity fluctuations begin to decay. After $t \approx 130$ the billows break up (see Figure 2f) and the turbulence becomes more horizontally homogeneous. During this transition to turbulent decay, the fluctuation $PE$ grows weakly, from $t \approx 111$ to $t \approx 150$. After $t \approx 150$, all fluctuations decay, turbulence length scales expand, and the formation of high- and low-speed streaks occurs, organizing larger horizontal scales of motion.
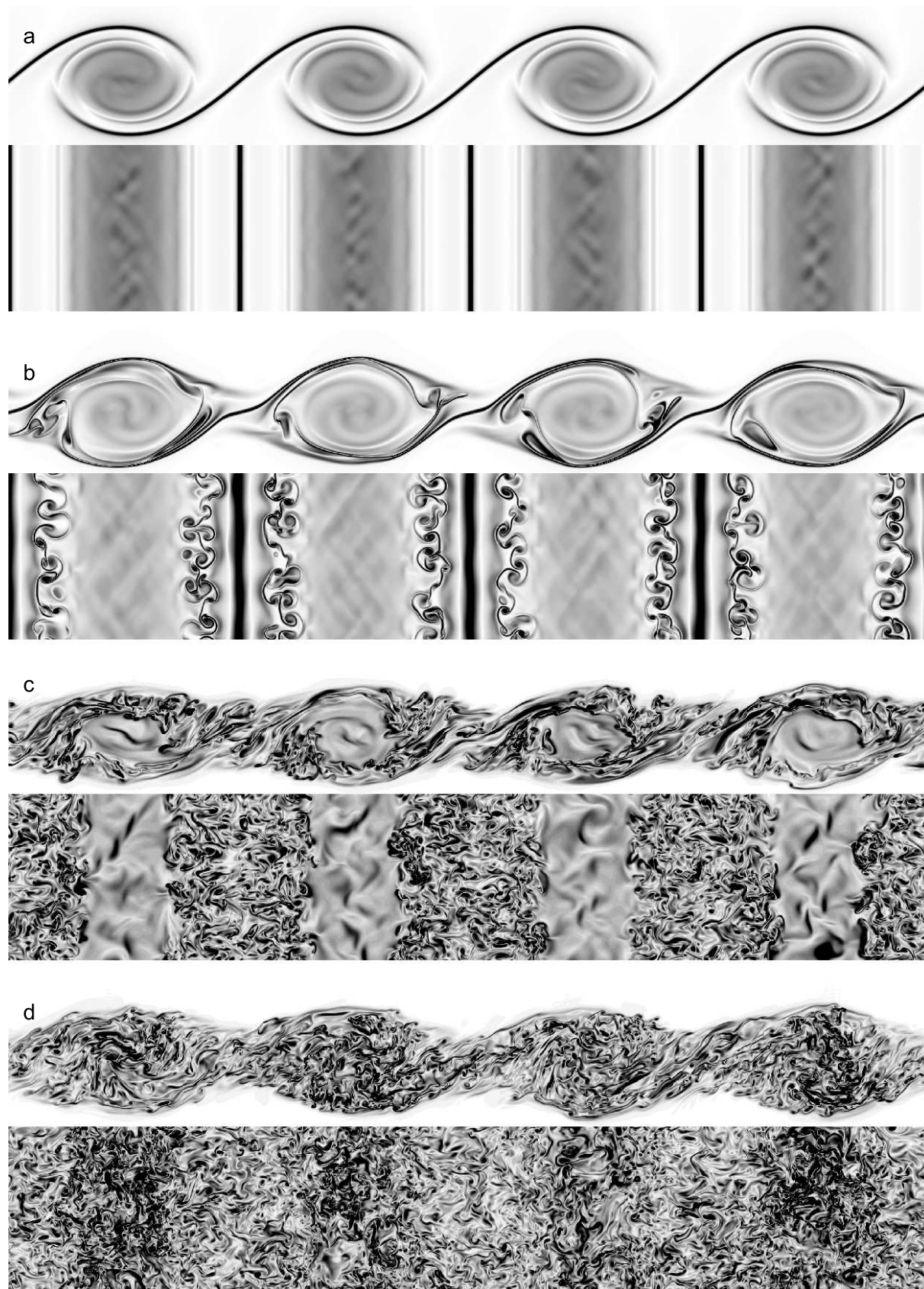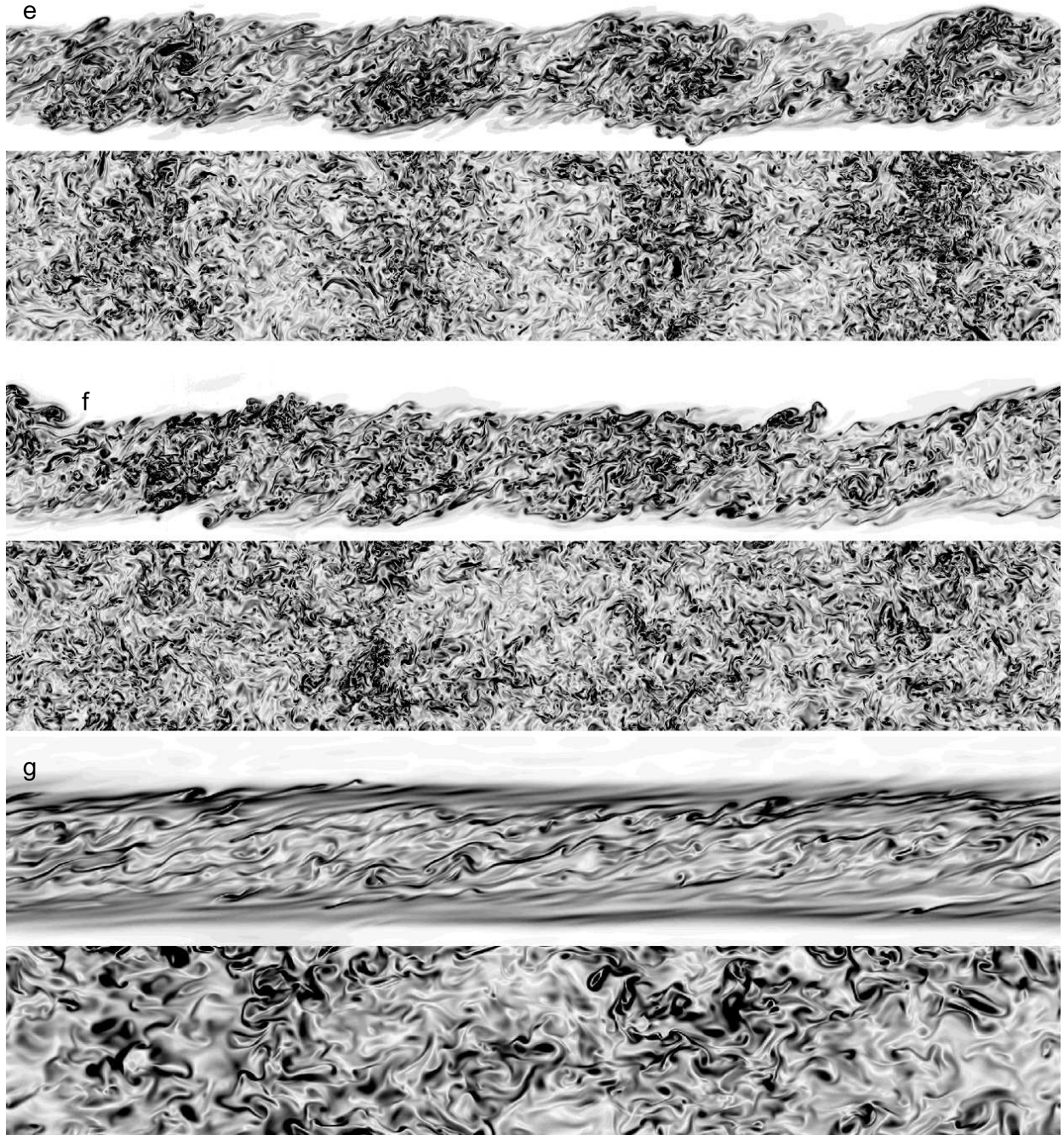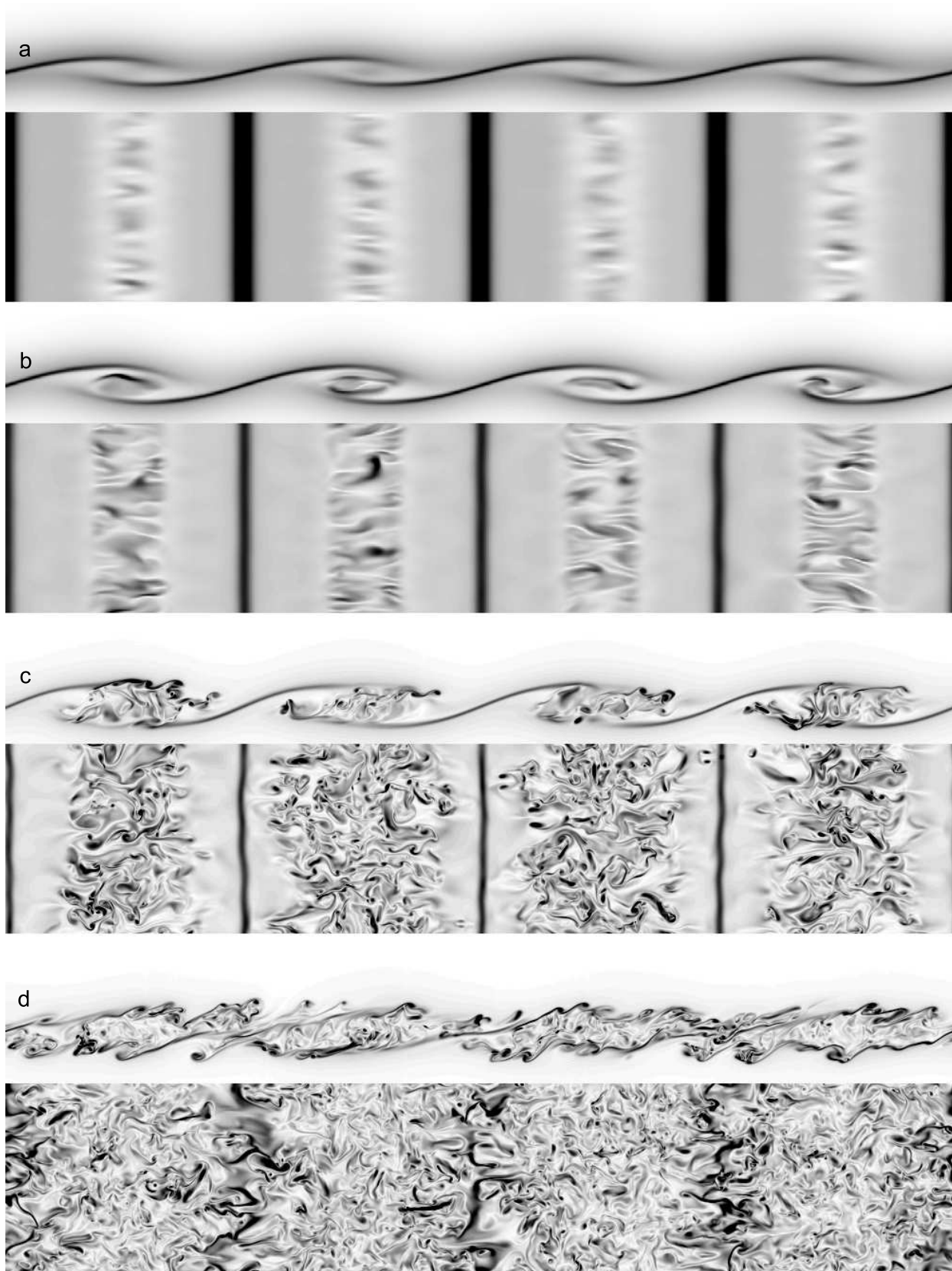
Figure 2: Vorticity slices for the Kelvin-Helmholtz instability with Ri=0.05 at seven distinct times during the flow evolution. For each time, two panels are shown: 1) a side view of the four billows and 2) a top view of 40% of the mid-plane. The times chosen correspond to when (a) billows reach maximum amplitude, $t = 37$; (b) secondary instability triggers the transition to turbulence, $t = 54$; (c) KE and PE dip to local minima before secondary peaks in KE and PE occur, $t = 68$; (d) KE and PE exhibit secondary maxima, $t = 85$; (e) turbulence intensity and vorticity magnitude attain maximum values, coincident with a local minimum in PE, $t = 111$, (f) turbulence first becomes horizontally homogeneous, $t = 130$; and (g) turbulence decay exhibits a change in character, $t = 251$. Panels a-d appear on the preceding page.
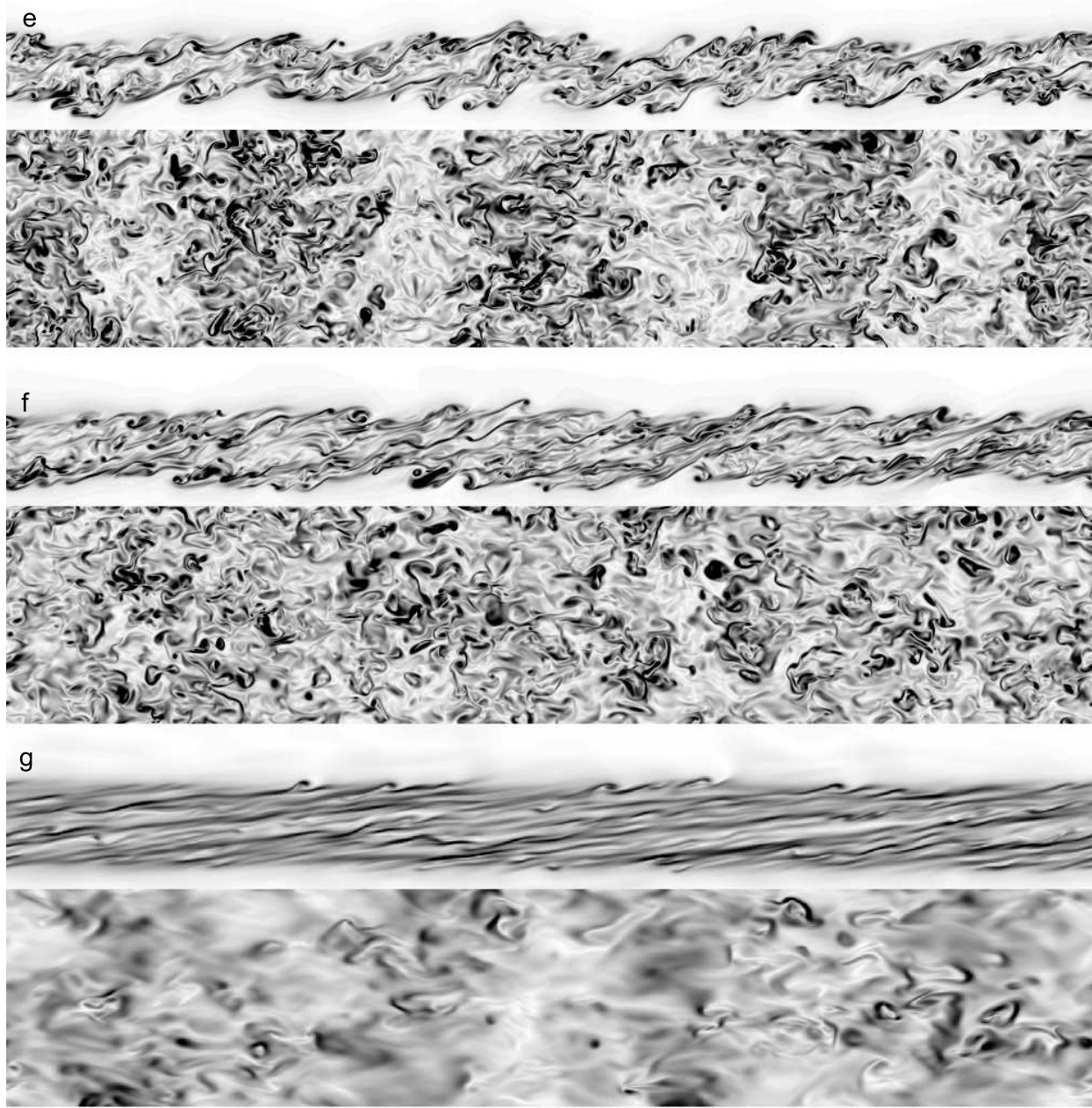
Figure 3: Vorticity slices for the Kelvin-Helmholtz instability with Ri=0.2. Seven times are presented, as in Figure 2, but because the two Ri cases differ, some of the times correspond to different stages in flow evolution and morphology. The times associated with (a) and (b) are the same as in Figure 2, *i.e.*, $t = 37$ and 54; they serve to demonstrate the different flow evolutions at early time. In panel (a) the flow is laminar, but with significant fluctuations in the core regions of the billows. In panel (b) we see turbulence developing in the billow cores. Panel (c) shows $t = 75$ when KE and PE are maximal. Turbulence is beginning to migrate from the cores at this time. Panel (d) at $t = 86$ shows the time when turbulence intensity and vorticity fluctuations are maximal, coincident with a local minimum in PE. Panel (e) depicts the time when KE and PE both begin to decay exponentially, $t = 95$. Panel (f) shows the moment when the flow first becomes horizontally homogeneous, $t = 105$, and Panel (g) depicts the evolution at $t = 130$ late in the decay phase.

Figures 1c,d and 3 present similar information for $Ri = 0.2$, *i.e.*, our most stable case. The evolution is notably different. First, and most importantly, because stratification is more effective in impeding vertical motions, the primary KH billows are much more shallow and less circular in cross section. Second, the $Ri = 0.2$ billows do not sustain laminar motion; instead they exhibit turbulent fluctuations immediately upon formation. We believe the enhanced stability of the billow cores for $Ri = 0.05$ results from their near-solid-body rotation. For discussion of the stabilizing influence of flow rotation, see Bradshaw (1969).

Figure 1c shows that the $Ri = 0.2$ case evolves in a somewhat simpler manner than the $Ri = 0.05$ case, due primarily to the rapid appearance of turbulence for $Ri = 0.2$. Because the billows are unstable immediately, they are not sufficiently coherent to experience secondary instability. This explains the lack of a second peak in $KE$ and $PE$ after the initial peak forms at $t \approx 73$. Also note that the maximum values attained for $KE$ and $PE$ are significantly lower than for the more turbulent $Ri = 0.05$ case. This occurs for two reasons. First, the higher stratification impedes vertical motions. Second, the enhanced stability of the billow cores for the $Ri = 0.05$ case allows the billows to continue developing coherently long after they would have succumbed to turbulence had the stabilizing influence of billow rotation not been present. This allows the billow energy to initially surge to very high values when $Ri$ is small. Both of these effects are nonlinear because they operate on the KH billows after they have attained finite and large amplitudes. The combination of the two effects results in $KE$ which is nearly 9 times larger for $Ri = 0.05$ than for $Ri = 0.2$.

An interesting aspect of the dynamics for $Ri = 0.2$ is the appearance of very coherent, horizontally propagating vortex rings from $t \approx 65$ to $t \approx 80$, *i.e.*, during the development of the dominant peak in $KE$. These coherent vortices form when the billow cores are vigorously turbulent, but the braid regions between billow cores are still laminar. A braid region contains a strong vortex sheet which connects the top of one billow to the bottom of its upstream neighbor. Figure 3c, lower panel, exhibits a horizontal slice through three of these horizontally propagating vortex rings, two of which can be seen at the upper left and right edges of the third billow from the left. Because the higher stratification present for $Ri = 0.2$ impedes vertical mixing, turbulent motion is effectively channeled along the horizontal, and well-formed vortex rings are the most efficient means of horizontal migration of kinetic energy.[1]

Vorticity reaches its maximum level, then begins to decay, at time $t \approx 86$; see Figure 3d. As with the $Ri = 0.05$ case, $PE$ dips when mixing is most vigorous; this occurs when vorticity is maximal (*i.e.*, $t \approx 86$). Also as in the $Ri = 0.05$ case, $PE$ grows very weakly (or merely plateaus) immediately as the turbulence begins to subside. After $t \approx 95$ (Figure 3e), all turbulence velocity and temperature fluctuations decay. At $t \approx 105$ (Figure 3f), the turbulence becomes horizontally homogeneous.

---

[1]We note that the number of individual vortex rings observed is less than the size increase from our previous smaller simulations and the new CAP simulations; *i.e.*, we observe fewer than 24 vortex rings. Hence, the previous simulations were too small to observe even one of these vortex rings on average.
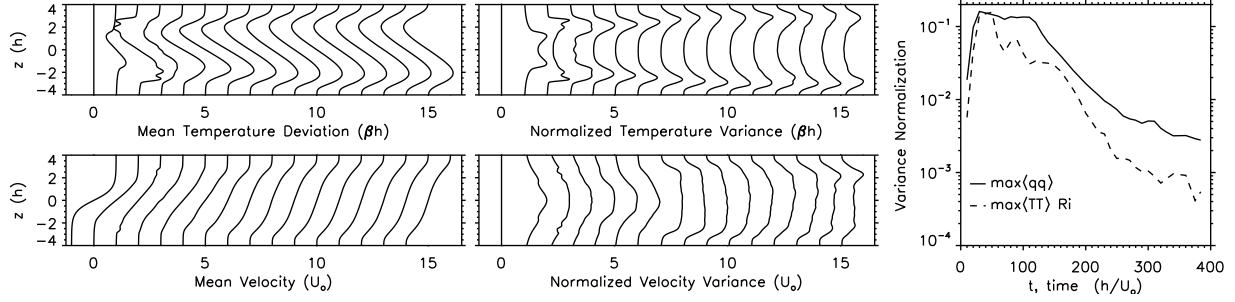
Figure 4: Ri=0.05 – Horizontally averaged profiles for the mean temperature (upper left) and mean streamwise velocity (lower left) and for the normalized temperature variance (upper middle) and normalized velocity variance (lower middle). The variance normalization factors used in the middle panels are shown in the rightmost panel. The solid (dashed) curve depicts the velocity variance $\langle qq \rangle = \langle u^2 + v^2 + w^2 \rangle$ (temperature variance $\langle TT \rangle$).

## 2.4  Flow Profiles for Ri=0.05 and Ri=0.2

The left panels of Figure 4 show the time evolution of the mean temperature deviation $T - z$ and the streamwise velocity as functions of height for $Ri = 0.05$. $z$ is the initial profile for $T$, therefore $T - z$ is the deviation of $T$ from its initial profile. The curves are separated by 20 time units. The middle panels show the normalized temperature and velocity variance, *i.e.*, the variance divided by its maximum. The right panel shows the maximum used to normalize the variance curves in the middle panels. In the figure, $\langle qq \rangle = \langle (U - \overline{U}(z))^2 + V^2 + W^2 \rangle$ is the sum over the three velocity components. Also, max $\langle TT \rangle$ has been multiplied by $Ri$ in the right panel so that each of the two curves represents a contribution to the total energy, which is approximated by $E = (q^2 + Ri\,T^2)/2$. As is evident from the decay with time shown in the right panel, normalizing the variance curves is necessary to see the variation with height of the profiles at late times.

The figure shows that the mean velocity and temperature are modified significantly during the course of the flow evolution. The deviation in the mean temperature exhibits a characteristic S-shape, with the layer rapidly growing from $2h$ to roughly $6h$ in depth. The mean velocity approaches a nearly constant gradient, though deviations from a strictly linear profile at late times are easily discernible. This occurs because of the vigorous mixing at midlayer (see the lower middle panel), which produces positive curvature in $U(z)$ at $z = 0$. The variance in temperature (upper middle panel) exhibits peaks near the top and bottom of the turbulent layer. This occurs because the mean temperature gradient is expelled from the middle of the layer by turbulent mixing, and it accumulates at the edges as a result. Where the mean temperature gradient is large, so too is the potential for producing temperature variance. Note the inverse relationship between the temperature and velocity variances: whereas the velocity variance peaks near midlayer and is understandably weaker near the layer edges, the temperature variance has the opposite behavior, *i.e.*, it attains its maximum near the layer edges, and is minimal near the middle of the layer.

Figure 5 shows similar data for $Ri = 0.2$. Here the buoyancy time, which is proportional to $\sqrt{Ri}$, is half that for $Ri = 0.05$, so the profiles are separated by only 10 time units. The profiles are in stark contrast to the lower $Ri$ case. First, the evolving turbulent layer is
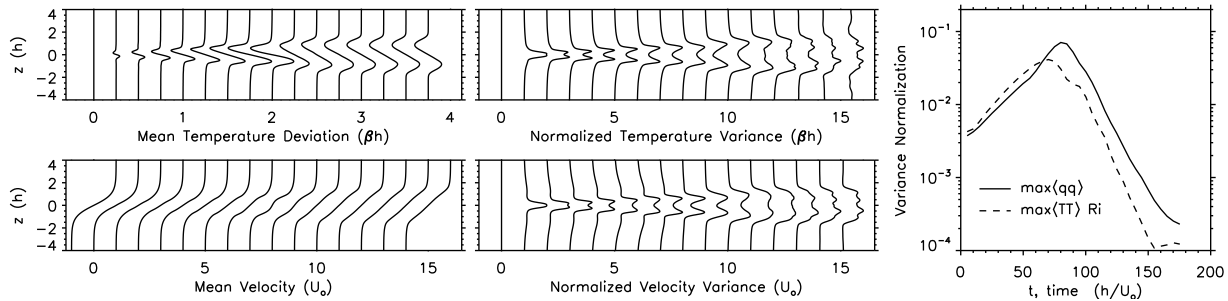
Figure 5: Ri=0.2 – Horizontally averaged profiles for the mean temperature (upper left) and mean streamwise velocity (lower left) and for the normalized temperature variance (upper middle) and normalized velocity variance (lower middle). The variance normalization factors used in the middle panels are shown in the rightmost panel. The solid (dashed) curve depicts the velocity variance $\langle qq \rangle = \langle u^2 + v^2 + w^2 \rangle$ (temperature variance $\langle TT \rangle$).

significantly shallower. It reaches only approximately half of the depth of the $Ri = 0.05$ layer. Second, the deviation $T - z$ is significantly smaller: note that the scale of the upper left panel of Figure 5 is four times smaller than that used for Figure 4. Also notice that the mean velocity gradient is more uniform at midlayer. And, perhaps most significantly, note that there is a fundamental difference between the shapes of the variance profiles in Figures 4 and 5. The $Ri = 0.2$ case does not exhibit the "inverse relationship" between max $\langle qq \rangle$ and max $\langle TT \rangle$ that we see for the lower Richardson number. On the contrary, the two profiles for $Ri = 0.2$ appear nearly identical in shape. The reason for the stark difference is 1) the relative efficiency of mixing for the two cases (mixing is initially much more efficient for the lower value of $Ri$) and 2) the relative depths of the two layers (the lower $Ri$ results in a deeper layer). The second effect insures that the higher $Ri$ case will retain strong velocity gradients, especially in the edge regions of the flow (see lower left panel of Figure 5 at the middle time shown, which is $t = 80$). As a result, the local stability profile (*i.e.*, the diagnosed Richardson number versus height) for this case dips back below the stability limit of $Ri(z) = 0.25$ near the edges, reinvigorating the dynamics in the edge regions and giving rise to renewed turbulence $KE$ production there. This is remarkable because at a slightly earlier time ($t = 70$) the entire stability profile had already been elevated above 0.25, indicating that the time of active dynamics had already passed.

The right panels of Figures 4 and 5 posses the same scale for the abscissa; hence, the magnitudes of $\langle qq \rangle$ and $Ri \langle TT \rangle$ can be compared for the two cases. The scale for the ordinate axes differs by a factor of two because the buoyancy time scale is proportional to $\sqrt{Ri}$. The most evident difference between the $Ri = 0.05$ and 0.2 cases is the extremely rapid instability growth and higher energies attained in the $Ri = 0.05$ case. In contrast, the $Ri = 0.2$ case develops more gradually. We also note that stratification has a much more evident impact on the developing dynamics for the higher $Ri$ case because $PE$ is larger than $KE$ during the initial growth phase for $Ri = 0.2$; this is not the case for $Ri = 0.05$. Finally, the decay in energy is exponential and rapid for $Ri = 0.2$, while it is much more gradual for $Ri = 0.05$ (even when the $\sqrt{Ri}$ factor is accounted for), and it appears to experience a shift in the controlling dynamics at $t \approx 250$ (see the change in the decay behavior in the upper right panel of Figure 4).
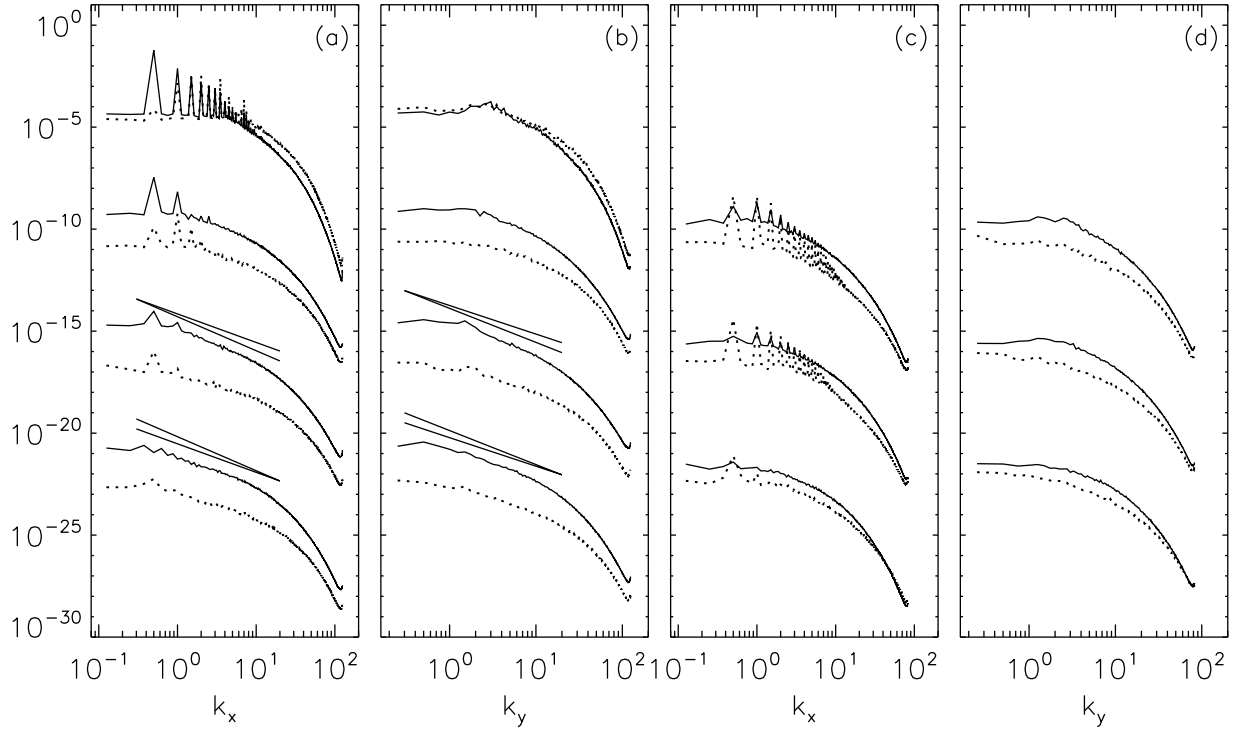
**Figure 6:** Horizontal wave-number spectra at mid-layer for velocity magnitude (solid) and temperature (dotted). The temperature spectra have been multiplied by $Ri$ to adjust their relative magnitude in proportion to their approximate contribution to the total energy. (a) streamwise spectra for $Ri = 0.05$ (b) spanwise spectra for $Ri = 0.05$, (c) streamwise spectra for $Ri = 0.2$, (d) spanwise spectra for $Ri = 0.2$. Panels (a) and (b) show spectra at $t = 54$ (the time of secondary instability), $t = 85$ (secondary peak in KE and PE), $t = 111$ (peak vorticity and minimum PE), and $t = 130$ (first time of horizontal homoegneity). Panels (c) and (d) show spectra at $t = 75$ (primary peak in KE and PE), $t = 86$ (peak vorticity and minimum PE), and $t = 105$ (first time of horizontal homogeneity). Earlier times are shown higher on the plot. Later times are shifted down successively by $10^{-6}$.

## 2.5 Turbulence Spectra and Resolution Requirements

Figure 6 presents horizontal spectra at mid-layer. Panels (a) and (b) show results for $Ri = 0.05$, while panels (c) and (d) show results for $Ri = 0.2$. Four times are shown for $Ri = 0.05$, including the time of secondary instability of the primary billows (top curve) occurring at $t = 54$. An equivalent spectrum is not shown for $Ri = 0.2$, since this case does not exhibit secondary instability and instead transitions directly from the primary billows to a turbulent state.

The time of secondary instability for $Ri = 0.05$ is characterized by strong peaks at $k_x = 2\pi/\lambda_x \approx 0.5$ and higher harmonics, where $\lambda_x = 12.566$ is the most unstable asymptotic eigenmode computed from linear stability theory. A broad and much weaker spanwise peak also occurs from $k_y = 2\pi/\lambda_y \approx 3$ to 8. The initial mode appeared near $k_y \approx 8$ at $t = 45$, and by time $t = 54$ this initial peak has spread to lower wavenumbers. We note that at $t = 54$, the temperature spectrum (multiplied by $Ri$) is larger than the velocity spectrum

at high wavenumber, indicating the importance of stratification at this very early time. At later times (lower) curves, the temperature assumes a less significant role.

Note that as time progresses, the primary mode is reduced in magnitude as turbulence overcomes the billows (Panel a). By $t = 130$, the flow is horizontally homogeneous, and the primary mode is barely visible.

Near the lower two sets of spectra in Panels (a) and (b), *i.e.*, times corresponding to peak vorticity (or maximum mixing) and horizontal homogeneity (or billow break up) we have added lines with slopes of $-5/3$ and $-7/5$. It appears that $t = 111$ has a kinetic energy spectrum with a near $k^{-5/3}$ form and a temperature-variance spectrum that is somewhat shallower, and $t = 130$ has a weaker slope for both kinetic energy and temperature that is consistent with $k^{-7/5}$. These slopes have been reported previously for simulations of stratified shear flow (Werne and Fritts 2000) and in aircraft spectra at the edge of the jet stream over Australia (Wroblewski *et al.* 2003). Using the slope indicators in Figure 6a,b as guides, it is clear that when turbulence peaks, the $Ri = 0.05$ simulations contain as much as 1.0 to 1.5 decades of inertial range turbulence; *e.g.*, see Panels (a) and (b).

The spectra for $Ri = 0.2$ look very different from those for $Ri = 0.05$. First, no secondary instability in span is observed (neither in physical space, see Figures 2 and 3, nor via peaks in the spanwise spectra). Second, the primary mode at $k_x = 2\pi/9.5$ and its harmonics exhibit relatively much stronger peaks in temperature variance than in kinetic energy when compared to the $Ri = 0.05$ results. Finally, no clear power-law behavior is observed, and therefore no inertial range spectrum can be clearly identified. The lack of secondary instability for $Ri = 0.2$ results because the primary billows succumb to small-scale turbulent motions as soon as the billows form. The relatively stronger peaks in the temperature spectrum result due to the relatively greater importance of the buoyancy field for the higher stratification level. And finally, the lack of an inertial range results because buoyancy is much more effective at inhibiting turbulent-kinetic-energy growth for the $Ri = 0.2$ case.

Resolution requirements for the simulations are determined by the largest energetic wavenumbers that develop. Resolutions satisfying $k_{max}\eta = \pi\eta/\delta x = 1.45 - 1.75$ (or $\delta x/\eta = 1.8 - 2.2$) produce acceptable results. Here $\eta = (\nu^3/\epsilon)^{1/4}$ is the Kolmogorov length scale, and $\delta x$ is the grid spacing. Figure 7 presents $\delta x/\eta$ for the $Ri = 0.05$ and $Ri = 0.2$ solutions. The finest grid resolution used for $Ri = 0.05$ was $\delta x = 0.0168$, with the number of $x$ grid-points given by Nx = 3000, while $Ri = 0.2$ required only $\delta x = 0.019$ and Nx = 2000. Note that the domain size for $Ri = 0.05$ is $L_x = 50.26$, while $L_x = 38$ is all that is required for $Ri = 0.2$. This is because the most unstable asymptotic eigen mode is 32% smaller for the higher-stratification case.

Significantly less resolution (a factor of 1.5 less in each spatial direction) is required for the high stratification case because the domain is smaller and the level of turbulence attained is less. As a result, the $Ri = 0.2$ case is $1.5^4 \approx 5$ times less expensive to compute than the $Ri = 0.05$ case. Twenty days of computing on as many as 1500 processors were required for the $Ri = 0.05$ case, while only 10 days of computing on at most 500 processors were required for the $Ri = 0.2$ case.
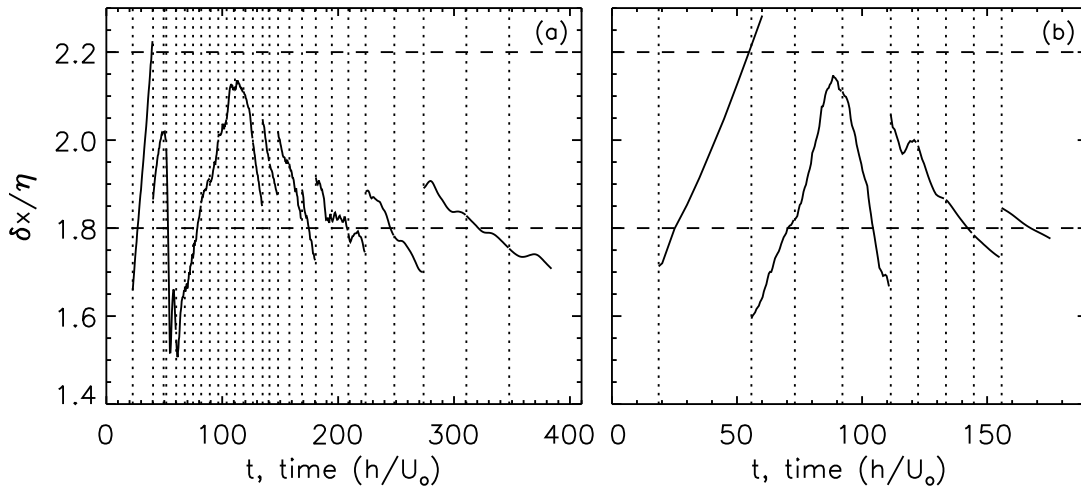
Figure 7: Resolution evaluation for the DNS simulations using $\delta x/\eta$ for (a) $Ri = 0.05$ and (b) $Ri = 0.2$. $\delta x$ is the grid spacing and $\eta$ is the minimum Kolmogorov scale evaluated by averaging on horizontal planes. Values between 1.8 and 2.2 are acceptable; however, care must be taken at early times when the flow is not horizontally homogeneous. Vertical dotted lines indicate the initial times for separate runs required to complete the simulation sequence.

# 3 Large Eddy Simulations

We shift now to a brief discussion of preliminary results obtained with the LES code for the $Ri = 0.05$ case. We also discuss some results for solutions obtained for the smaller domains computed previoiusly. The problem addressed is challenging for LES methods in general because it encompasses two transitions: 1) from an initial high-Reynolds-number laminar state to the development of instability and turbulence, and 2) from the resulting turbulent state to turbulence decay, restratification, and relaminarization. We will examine the successes and challenges of our current LES model and describe steps we plan to take to further improve the algorithm.

## 3.1 Numerical Procedure

The LES algorithm has evolved since conducting our initial comparisons with the DNS on the smaller domains used previously. Originally we employed a turbulent kinetic energy (TKE) model for the eddy viscosity $\nu_t$, and we modeled the eddy diffusivity $\kappa_t$ by specifying a turbulent Prandtl number $Pr_t = \nu_t/\kappa_t = 1/(1 + 2\ell/\Delta)$. Here $\ell$ is a model length scale and $\Delta$ is the spectral cut-off filter width. For details, see Moeng (1984). Based on our initial comparisons with the DNS, we have migrated to a dynamic procedure (Germano *et al.* 1991) employing two filter widths $\Delta$ and $2\Delta$ in order to determine the constants in the Smagorinsky model for the turbulent stress, $\tau_{ij} = -2C_\tau \Delta^2 |S| S_{ij}$, and in a similar gradient transport model

for the eddy diffusivity, $q_i = -C_q \Delta^2 |S| \partial T / \partial x_i$. Here $C_\tau$ and $C_q$ are the model constants, $\Delta$ is the primary filter size (taken to be the grid spacing), and $|S|$ is the magnitude of the resolved-scale strain rate. By using Germano's identity (Germano *et al.* 1991) and similar expressions for $tau_{ij}$ and $q_i$, but written for the second filter width $2\Delta$, we can solve for $C_\tau$ and $C_q$ as a function of space and time. Our solution procedure involves a least-squares fit to the filtered fields averaged on horizontal planes. Such horizontal averaging is appropriate at late times when turbulence becomes horizontally homogeneous, but the impact of this choice should be evaluated at early time when the flow is organized coherently into well-defined KH billows which exhibit strong streamwise variation.

The numerical algorithm we employ is a hybrid pseudo-spectral/finite-difference method in which the periodic horizontal directions are treated with Fourier-series representations, and the vertical direction employs a staggered 2nd-order finite-difference algorithm which conserves resolved-scale mass, momentum, kinetic and thermal energy exactly. Time stepping is as with our DNS code, *i.e.*, we use the RK3 scheme of Spalart *et al.* (1991). Part of the time-stepping procedure unique to the LES code is a global Poisson solver for the pressure field which insures incompressibility at each RK3 sub-step. Hence, the LES code confronts the same data-flow issue as the DNS code, namely, FFTs are done most efficiently on contiguous data, but so too is the Poisson inversion the LES performs for the pressure; therefore, an all-to-all global communication pattern at each sub-step is required to move between the two desired data layouts for the LES model, just as it is for the DNS code. This has implications for the optimal parallel architecture on which both codes run efficiently. See §4 for more discussion on this point.

## 3.2   LES-DNS Comparison

In this section we compare solutions obtained with the DNS and LES codes. Figure 8 shows time evolutions for $KE$ and $PE$ for both models. The LES is computed with 6 times fewer grid points in each spatial direction. It therefore requires roughly $6^3 \approx 200$ times less memory and $6^4 \approx 1300$ times fewer CPU cycles than the DNS to complete.

The figure demonstrates that the dynamic LES procedure qualitatively tracks the volume-integrated $KE$ and $PE$ time traces, managing to even describe the secondary maxima in $KE$ and $PE$ at $t \approx 85$. This is impressive given that the flow undergoes two transitions. Nevertheless, at late times the LES systematically over-predicts both $KE$ and $PE$ by more than 40%, and at early times (*i.e.*, just before the secondary maxima in $KE$ and $PE$) the LES under-predicts $KE$ by 14% and $PE$ by 50%. We note that given the sensitive dependence on initial conditions exhibited by turbulent flows, once the LES solution diverges significantly from the DNS, it is unlikely that the LES can re-establish a time-accurate solution. Therefore it is critically important that the LES tracks the initial evolution for as long as possible.

Careful examination of Figure 8 reveals that the LES departs from the DNS at $t = 45$, precisely when the secondary instability of the primary billows attains a non-linear amplitude (see Figure 2b). This is perhaps expected, as the SGS turbulence model in the LES is not designed to describe a high-$Re$ laminar solution with extremely small spatial scales. In addition, the horizontal averaging we currently use to evaluate model constants is
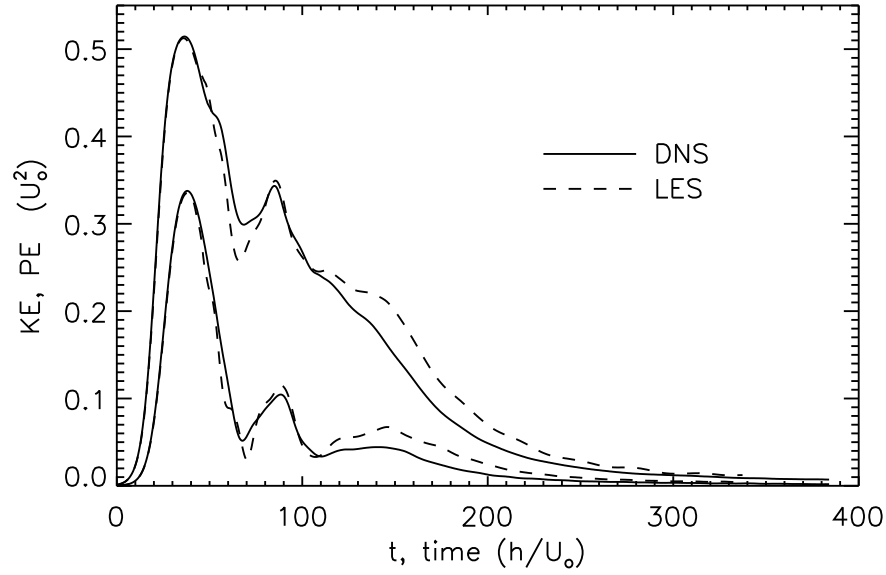
Figure 8: KE and PE for $Ri = 0.05$ for the DNS (solid) and LES (dashed) solutions. KE is represented by the upper two curves; PE is represented by the lower two curves.

inappropriate at this early time. As a result, our next planned improvement to the LES model is to replace horizontal averages with spanwise averages. If the spanwise-average statistics prove insufficient to provide a reasonably constrained description of the flow, we can combine the results for the four individual billows to improve the statistical signal-to-noise ratio.

In order to better understand the discrepancies between the DNS and LES solutions (especially at late times), Figure 9a,b presents $KE$ and $PE$ evaluated at mid-layer and in the edge regions of the flow. The edges are located by the local maxima in the temperature variance. Results for both edges are averaged together in Figure 9b. The figure shows that at mid-layer $PE$ is insignificant after $t \approx 68$ (Figure 2c). This occurs because temperature fluctuations are homogenized by vigorous turbulent motion early in the flow's evolution for $Ri = 0.05$. Temperature homogenization at mid-layer persists throughout the simulation, and the LES predicts this reasonably well, *i.e.*, it over-predicts the mid-layer $KE$ by only 10% to 15%. In contrast, the edge regions exhibit significant temperature variance at all times. This occurs because the mean temperature gradient in the edge regions is sharpened by mixing in the core of the layer, and therefore turbulent production of temperature variance $\langle w\theta \rangle \partial_z \overline{T}$ is enhanced near the edges. Enhanced edge-region mean thermal gradients result because the temperature contrast that is eradicated in the middle of the layer by turbulent mixing is forced to concentrate near the edges. The LES also predicts enhanced temperature variance near the layer edges, but it significantly over-predicts both $KE$ and $PE$ here; *e.g.*, the LES $KE$ ($PE$) excess is 45% (50%) in the edges; see Figure 9b after $t \approx 100$.

The lower two panels of Figure 9 compliment the upper two panels by presenting the flow's energy fractions, *i.e.*, the relative contributions the temperature and velocity components make to $KE + PE$. From panel (c) we can see that the correct relative proportions for
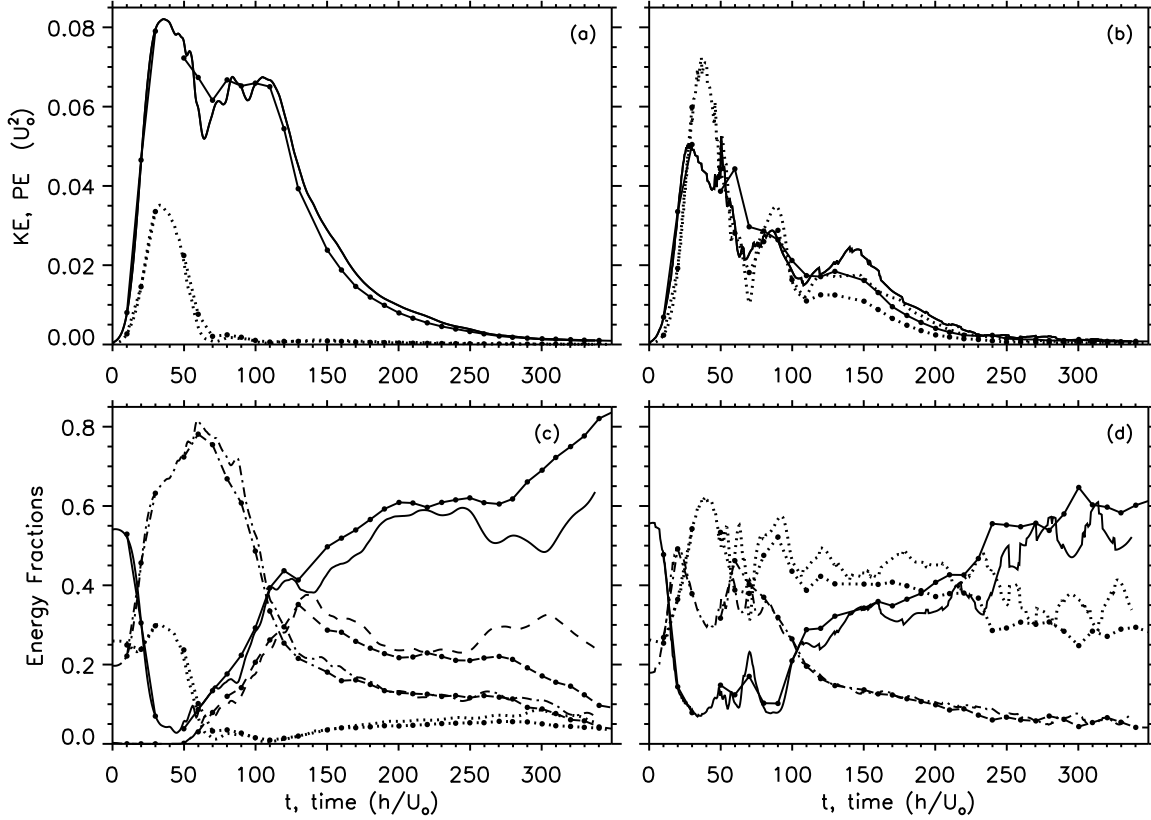
Figure 9: Comparison of LES and DNS results. (a) Mid-layer $KE$ (solid) and $PE$ (dotted). (b) Edge-region $KE$ and $PE$. (c) Mid-layer velocity and temperature variances normalized by $KE + PE$. (d) Edge-region variances normalized by $KE + PE$. For (c) and (d), line styles are solid $\langle u^2 \rangle$, dashed $\langle v^2 \rangle$, dot-dashed $\langle w^2 \rangle$ and dotted $\langle \theta^2 \rangle$. LES results are shown with continuous curves. DNS data are shown with solid symbols. All curves show averaged results, where averaging is conducted over the entire horizontal extent of the domain and vertically over a limited range in $z$ defined by $\Delta z = 1.0$. The edge regions are defined by the location of maximum temperature variance. Results for the top and bottom edges are averaged together in (b) and (d).

$\langle u^2 \rangle$, $\langle v^2 \rangle$, $\langle w^2 \rangle$, and $Ri\langle \theta^2 \rangle$ at mid-layer are all reasonably captured by the LES out to $t \approx 250$ (when the mid-layer dynamics become stratification dominated). Some discrepancy is apparent for $\langle u^2 \rangle$ and $\langle v^2 \rangle$ for $t > 130$, but this is roughly 10% or less, whereas the departure at $t \approx 250$ grows to 100% for $\langle v^2 \rangle$ and 25% for $\langle u^2 \rangle$. The edge regions, shown in panel (d), show near-perfect agreement for the $\langle v^2 \rangle$ and $\langle w^2 \rangle$ proportions, but $\langle \theta^2 \rangle$ ($\langle u^2 \rangle$) appears to be systematically over (under) predicted by the LES (relative to $KE + PE$) after $t \approx 100$. We note that nearly all times are stratification dominated in the edge regions of the mixing layer.

The results of Figure 9 indicate that the LES solutions falter when stratification dominates the dynamics. It is important to note that when this occurs, the spectral shape differs from that assumed by the SGS model. Hence, though the dynamic procedure we employ can adjust the model constants as a function of space and time, it is not set up to also adjust

the spectral slope. We suspect this may be an important cause of the remaining differences between the DNS and LES solutions at late times. We therefore plan to modify the dynamic procedure to adjust both the model constants and the spectral slope. This will require the introduction of yet a third filter width to the dynamic procedure.

## 3.3   The Optimal Cut-off Filter Width $\Delta$

By experimenting with the cut-off filter width $\Delta$ used by the LES code, we have deduced that the optimal value is approximately six times larger than the DNS Nyquist mode. This produces the best overall agreement with the DNS, and it significantly reduces the computational cost. However, a noteworthy problem is the dependence the LES solutions exhibit on $\Delta$. In particular, we notice that as $\Delta$ is increased, the predicted eddy-viscosity and eddy-diffusivity values systematically drift above the same quantities obtained from filtered DNS data. This might be expected if the larger values of $\Delta$ happened to be outside the inertial range of turbulence for the DNS. However, upon examining Figure 6a,b, it is quite apparent that $\Delta = 6\,\delta x$ is barely at the high wavenumber end of the turbulence inertial range. Hence, further increases in $\Delta$ should sweep through the inertial range, and other values of $\Delta$ should produce satisfactory results. We do not know what is causing this anomalous dependence on $\Delta$ in the LES results, but we are hopeful that the two improvements we currently plan (*i.e.*, 1. spanwise rather than full horizontal averaging when evaluating model constants, and 2. the introduction of a third filter width) will help alleviate the problem.

## 4   Parallel Performance on Kraken and Jvn

As part of the evaluation criteria for receiving CAP Phase II computer time, acceptable parallel scalability must be demonstrated for jobs as large as those being proposed. The results of our scalability tests demonstrate that our DNS code is extremely efficient. We measured the asymptotic parallelization efficiency in two ways. The first method, which employed Amdahl's Law (Amdahl 1967, Gustafson 1998, Lewis and El-Rewini 1992), showed the code to be 99.976% parallel when running on up to 2400 processors on the NAVO IBM Power 4+ (Kraken). The second method used the scaled grind time $g_n$ to evaluate the performance. The grind time is defined by $g_n = g_1/n^m$, where $g_1$ is the time required to complete a calculation on one processor, and $g_n$ is the time needed when $n$ processors are used. The parallelization efficiency reported by this method is $m = 95\%$. Both of these tests were conducted by fixing the problem size per processor, so that as larger numbers of processors were employed, a larger total problem size was treated. Relative to the Cray T3E which was previously at the ERDC Major Shared Resource Center (MSRC), the speed-up we experienced on Kraken was 5.2. This is an acceptable level of performance given that the theoretical speed-up for these two machines is 5.7. We achieve this high computational efficiency by minimizing the amount of serial operations in the code, conducting efficient parallel I/O (Werne *et al.* 2000), and using very fast one-sided communication calls available via David Klepacki's SHMEM library routines (Klepacki 2004).
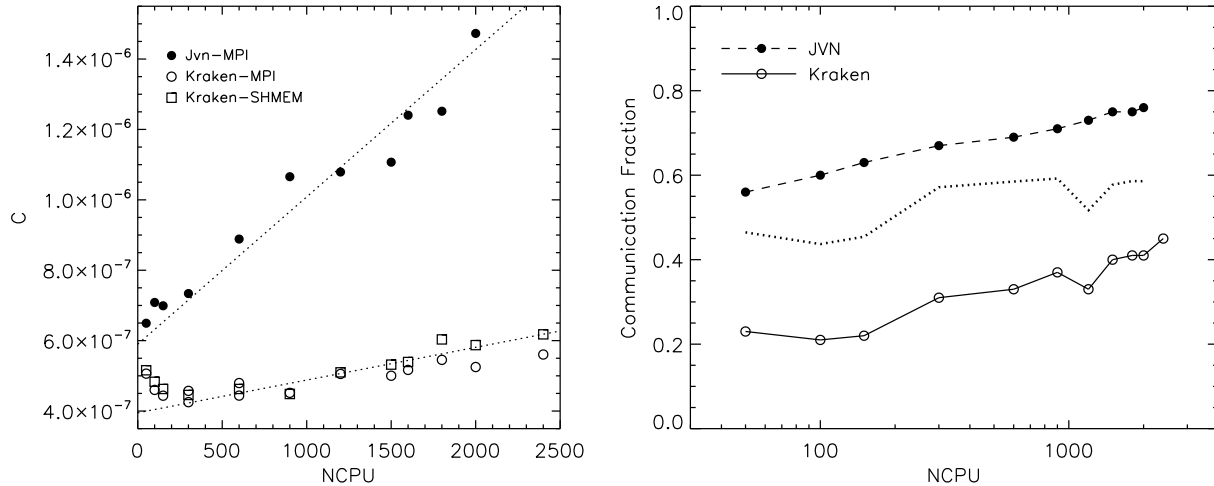
Figure 10: (a) Parallel scalability tests for the DNS code running on Kraken and Jvn. $C = (walltime) \times NCPU/(N_t N_g \log N_g)$ is proportional to the ratio of the CPU seconds to the total number of operations completed. $N_t$ is the number of time steps, and $N_g$ is the total grid size. The $\log N_g$ factor appears because FFTs dominate the computational cycles, and the cost of a 3D FFT is proportional to $N_g \log N_g$. (b) Fraction of time DNS code spends performing inter-processor communication on Kraken and Jvn. All tests are conducted with fixed problem-size per processor of 280 Mb.

Identical tests on the ARL Intel Xeon EM64T cluster (Jvn) produced similarly impressive parallelization numbers; *e.g.*, 99.89% when using Amdahl's Law and $m = 88\%$ according to the grind time. These results are presented in Figure 10a. We note that the per-processor peak flop rate on Jvn is slightly higher than that on Kraken (7.2 versus 6.8 Gflops/processor).

We caution that the quoted parallelization measures do not effectively communicate the entire performance assessment for a computer code and platform combination. For example, Figure 10a demonstrates that, despite similar peak flop rates and high parallelization efficiencies on both machines, our code requires nearly three times as much time to complete on Jvn as Kraken when running on 2000 processors. This occurs because the code consumes significantly more time conducting inter-processor communication on Jvn than on Kraken. And since communication operations can execute in parallel, this explains why high parallelization performance can be consistent with slower than expected execution times.

In an attempt to understand the performance deficit on Jvn compared to Kraken, we measured the percentage of time spent conducting all-to-all communications associated with our global transpose operation. We found that while Kraken can spend as much as 45% of its time communicating when running on 2400 processors, Jvn spends as much as 75% when running on 2000 processors. See Figure 10b. As a result, despite the slightly higher theoretical peak flop rate on Jvn, simulations take significantly longer to complete. We note that the upward drift in the communication fraction results because the all-to-all communication pattern uses decreasing message sizes as $NCPU$ is increased (*i.e.*, message sizes scale with $1/NCPU$).

It is understandable that Jvn spends more time in communication than Kraken, given

that Jvn's per-processor bandwidth (250 Mb/s) is nearly half that of Kraken (450 Mb/s). However, the ratio, which is only 1.8, does not suggest the relative performance indicated by Figure 10. In order to assess this performance difference in greater detail, we profiled the inter-processor and inter-node send/receive communication speeds on the two systems and compared the results with the theoretical limits suggested by the network hardware. We found that when Jvn conducts communication between two processors located on separate nodes, it nearly achieves its theoretical inter-node bandwidth of 500 Mb/s (it achieves 489 Mb/s), with a zero-message-size latency of 9.6 $\mu sec$. In contrast, Kraken manages 1.85 Gb/s, which is only about half the theoretical limit of 3.6 Gb/s possible with its two links per node. Kraken's latency was measured to be 7.6 $\mu sec$. When two processors per node send and receive messages with two processors on another node, Jvn's bandwidth performance remains high, but its latency jumps by a factor of two, while Kraken's latency remains low, but its bandwidth nearly doubles to 3.45 Gb/s. It appears that when the IBM performs send/receive communication between single processors on each of two nodes, it only uses one network link per node, while communication between more processors per node is required to trigger usage of both links.

We also conducted tests involving seven processors per node on Kraken (Jvn only has two processors per node), because this reflects the communication pattern of our DNS code on the IBM. The timing tests demonstrated that Kraken exceeds its quoted performance figures, delivering up to 4.75 Gb/s total over the two communication links, while its latency remains low at 8.2 $\mu sec$.

The dotted curve in Figure 10b incorporates information from these tests, including the specific times measured when passing messages of the size used during the parallel scalability tests. The magnitude of the curve indicates the relative performance difference between the two systems based on the send/receive tests. As is apparent in the figure, Jvn's performance remains somewhat worse than expected, indicating that improvements should be possible through communication-software improvements and/or an improved MPI implementation. We estimate a theoretical speed increase for communication of 25% should be possible for the current system.

The results of this analysis are a bit disappointing for the two new systems examined because the inter-node network performance seems no better than that possible on the Cray T3D circa 1994. This seems an unacceptable lag given the 45-fold increase in per-processor flop rates between the T3D and these new systems, and it explains why such large fractions of time are being spent communicating (see Figure 10b) as opposed to computing. For this reason we feel "percent of time communicating" should be a criterion used when benchmarking code performance on new parallel systems. Inter-processor network bandwidth will only become more important as system size continues to grow.

### References

1. Amdahl, G.M., "Validity Of Single-Processor Approach to Achieving Large-Scale Computing Capability, Proceedings of American Federation of Information Processing Societies

(AFIPS) Conference, Reston, VA, pp.483–485, 1967.

2. Bernardo, J.M., and A.F.M. Smith, Bayesian Theory, Wiley, UK, 1994

3. Bradshaw, P., "The analogy between streamline curvature and buoyancy in turbulent shear flows. J. Fluid Mech., **54**, 177-191, 1969.

4. Germano, M., U. Piomelli, P. Moin, and W.H. Cabot, "A Dynamic Subgrid-scale Eddy Viscosity Model", *Phys. Fluids A*, **3**, 1760–1765, 1991.

5. Gibson-Wilde, D., J. Werne, D.C. Fritts, and R. Hill, "Direct Numerical Simulation of VHF Radar Measurements of Turbulence in the Mesosphere," *Radio Science*, **35**, 783, 2000.

6. Gustafson, J.L., "Reevaluating Amdahls Law, Communications of the Association for Computing Machinery (CACM), **31**(5), 532–533, 1998.

7. Julien, K., S. Legg, J. McWilliams, J. Werne, "Rapidly Rotating Turbulent Rayleigh-Bénard Convection," J. Fluid Mech., **322**, 243, 1996.

8. Kelley, M.C., C.Y. Chen, R.R. Beland, R. Woodman, J.L. Chau, and J. Werne, "Persistence of a Kelvin-Helmholtz Instability Complex in the Upper Troposphere," *J. Geophys. Res.* (in press).

9. Klepacki, D., "The IBM High Performance Computing Toolkit," Trace Library Documentation, IBM, 2004.

10. Lewis, T.G. and H. El-Rewini, "Introduction to Parallel Computing, Prentice Hall, pp.32–33, 1992.

11. Moeng, C.-H., "A Large-eddy simulation model for the study of planetary boundary-layer turbulence. J. Atmos. Sci., **41**, 2052–2062, 1984.

12. Spalart, P.R., R.D. Moser and M.M. Rogers, "Spectral methods for the Navier-Stokes equations with one infinite and two periodic directions," J. Comput. Phys., **96**, 297–324, 1991.

13. Werne, J., P.Adams, and D.Sanders, "Linear scaling during production runs: conquering the I/O bottleneck," ARSC Cray T3E Users' Group Newsletter, http://www.arsc.edu/pubs/T3Enews/T3Enews193.shtml, March 31, 2000.

14. Werne, J. and D. C. Fritts, "Stratified shear turbulence: Evolution and Statistics" (cover article) Geophys. Res. Lett., **26**, 439–442, 1999a.

15. Werne, J. and D. C. Fritts, "Turbulence and Mixing in a Stratified Shear Layer: 3D K-H Simulations at $Re = 24,000$" European Geophysical Society, XXIV General Assembly. April 1999, The Hague, 1999b.

16. Werne, J. and D.C. Fritts, "Structure Functions in Stratified Shear Turbulence," DoD HPC User Group Conference, Albuquerque, NM. June 2000.

17. Wroblewski, D., O. Cote, J. Hacker, T.L. Crawford, and R.J. Dobosy, "Refractive turbulence in the upper troposhere and lower stratosphere: analysis of aircraft measurements using structure functions," 12th Symposium on Meteorological Observations and Instrumentation, Long Beach, February, 2003.