Direct Evidence of an Oceanic Inverse Kinetic Energy Cascade from Satellite Altimetry

ROBERT B. SCOTT AND FAMING WANG

Institute for Geophysics, Jackson School of Geosciences, The University of Texas at Austin, Austin, Texas

(Manuscript received 17 June 2004, in final form 22 February 2005)

ABSTRACT

Sea surface height measurements from satellites reveal the turbulent properties of the South Pacific Ocean surface geostrophic circulation, both supporting and challenging different aspects of geostrophic turbulence theory. A near-universal shape of the spectral kinetic energy flux is found and provides direct evidence of a source of kinetic energy near to or smaller than the deformation radius, consistent with linear instability theory. The spectral kinetic energy flux also reveals a net inverse cascade (i.e., a cascade to larger spatial scale), consistent with two-dimensional turbulence phenomenology. However, stratified geostrophic turbulence theory predicts an inverse cascade for the barotropic mode only; energy in the large-scale baroclinic modes undergoes a direct cascade toward the first-mode deformation scale. Thus if the surface geostrophic flow is predominately the first baroclinic mode, as expected for oceanic stratification profiles, then the observed inverse cascade contradicts geostrophic turbulence theory. The latter interpretation is argued for. Furthermore, and consistent with this interpretation, the inverse cascade arrest scale does not follow the Rhines arrest scale, as one would expect for the barotropic mode. A tentative revision of theory is proposed that would resolve the conflicts; however, further observations and idealized modeling experiments are needed to confirm, or refute, the revision. It is noted that no inertial range was found for the inverse cascade range of the spectrum, implying inertial range scaling, such as the established $K^{-5/3}$ slope in the spectral kinetic energy density plot, is not applicable to the surface geostrophic flow.

1. Introduction

High-quality measurements of sea surface height (SSH) from satellites are revolutionizing the study of ocean dynamics and the fields affected by ocean circulation such as climate research. This article presents analysis of those measurements from the perspective of geostrophic turbulence, or quasi-two-dimensional turbulence phenomenology. Because of vertical density stratification and the earth's rotation, the large-scale atmospheric winds and oceanic currents are quasi two-dimensional (Gill 1982; Pedlosky 1987). This forms the primary motivation for studying quasi-two-dimensional turbulence, though motion of plasmas in a magnetic field can also display approximately two-dimensional turbulence behavior.

Corresponding author address: Robert Scott, Institute for Geophysics, Jackson School of Geosciences, The University of Texas at Austin, 4412 Spicewood Springs Rd., Bldg. 600, Austin, TX 78759.
E-mail: rscott@ig.utexas.edu

Since the pioneering, theoretical work of the late 1960s (Kraichnan 1967; Batchelor 1969), much progress has been made to extend the phenomenology to include further real-world complications such as stratification (Charney 1971), the change in Coriolis parameter with latitude (the β effect) (Rhines 1975), formation of coherent structures (McWilliams 1984), horizontal compressibility (Larichev and McWilliams 1991), bottom topography (Treguier and Hua 1988), bottom drag, and lateral viscosity (Scott 2001); see the review by Danilov and Gurarie (2000) for many further examples. Numerical simulations of simplified equations in idealized domains have remained the primary tool of investigation. A few laboratory experiments using soap films, electrolytes in a magnetic field, or rotating water tanks have revealed quasi-two-dimensional turbulence effects (Read 2001; Danilov and Gurarie 2000), helping to ground the work in reality, but questions of the applicability to real geophysical flows inevitably remain. The requirement of long time series of near-simultaneous measurements of ocean currents covering large regions has, until very recently, severely limited the possible

theory–data comparisons. With the advent of satellite altimetry, the World Ocean is now accessible as a laboratory to study quasi-two-dimensional turbulence, with all its real-world complications. Here we compare the observations of the kinetic energy cascade near the ocean surface with the predictions of turbulence phenomenological theory.

Since the pioneering work of Kolmogorov (1941a,b), the cornerstone of turbulence research has been the kinetic energy cascade (Frisch 1995). In the inertial range where direct viscous effects are negligible, energy is transferred between different length scales via triad interactions (also called "internal nonlinear transfer"). A distinguishing feature of two-dimensional (2D) turbulence is the counterintuitive cascade of energy from smaller to larger length scales. This is called the "inverse cascade" because it flows in the opposite direction to the "direct cascade" or "forward cascade" of its counterpart in 3D turbulence. For a rotating and stratified fluid such as the ocean or atmosphere, the situation is more complicated. The theory is reviewed in more detail in the following section. Here we note that at scales larger than or equal to the Rossby radius of deformation scale, L_D, the inverse cascade is confined to the barotropic (i.e., depth independent) flow, while motions with vertical structure (the higher baroclinic modes defined below) undergo a direct cascade. Upon reaching L_D baroclinic energy is largely converted to the barotropic mode.

While the quasi-2D kinetic energy cascade has been observed in numerical simulations, laboratory experiments, and observational atmospheric data, it has never been observed directly in the ocean. (Note that observing the $K^{-5/3}$ scaling of spectral kinetic energy density with wavenumber K does not indicate the direction of the cascade, since the same result holds for both the inverse cascade of 2D turbulence and the direct cascade of 3D turbulence. Direct evidence requires measurement of the spectral kinetic energy flux Π_K, defined as the flux of kinetic energy through total wavenumber K.) Indirect evidence is provided by apparent eddy mergers in altimeter data (B. Qiu 2004, personal communication) and with the observation that the peak in the sea surface height power spectrum shifts to lower wavenumber as one moves downstream of the eddy-producing East Pacific Ridge (Stewart et al. 1996). Stewart et al. suggest that the inverse cascade may be the mechanism driving the observed growth in eddy size. A similar result was found in the North Pacific Subtropical Countercurrent and Hawaiian Lee Countercurrent, where eddy generation by baroclinic instability peaks in winter, followed by a period of apparently free evolution (Kobashi and Kawamura 2002). Again the authors attribute the growth in scale to an inverse energy cascade. Using subsurface floats in the North Atlantic Ocean, LaCasce and Bower (2000) found their relative dispersion was consistent with an inverse energy cascade in the western (but not the eastern) part of the basin.

However, under the standard interpretation of altimeter measurements reflecting the first baroclinic mode, these observations of apparent growth in length scales are *not* consistent with the geostrophic turbulence theory of the inverse cascade. To see this, we must first relate the surface signal, what the altimeter sees, to the baroclinic modes, in which the theory is framed. Wunsch (1997) argues the altimeter reflects largely the first baroclinic mode since the extratropical, eddy kinetic energy is roughly equally partitioned between the barotropic and first baroclinic modes, and the first baroclinic mode is more strongly weighted at the surface because of surface-intensified density stratification. This can be made clearer using the two-layer model formulation. Here the surface reflects the upper-layer streamfunction ψ_1, which is related to the barotropic streamfunction $\psi_{\rm bt}$ and the baroclinic streamfunction $\psi_{\rm bc}$ by

$$\psi_1 = \psi_{\rm bt} + \frac{\psi_{\rm bc}}{\sqrt{\delta}}, \tag{1}$$

where $\delta = H_1/H_2$ is the ratio of layer depths (e.g., Arbic and Flierl 2003). Surface-intensified stratification corresponds to $\delta \ll 1$. If we also have similar kinetic energy levels in the barotropic and baroclinic modes, then assuming similar horizontal length scales apply for both modes then $\psi_{\rm bt} \sim \psi_{\rm bc}$, which leads to the approximation

$$\psi_1 \approx \frac{\psi_{\rm bc}}{\sqrt{\delta}}. \tag{2}$$

So for the extratropical ocean we expect the altimeter to reflect primarily, though of course not exclusively, the first baroclinic mode.

Thus based upon stratified, geostrophic turbulence theory, described in more detail below (Salmon 1980; Fu and Flierl 1980; Hua and Haidvogel 1986; Smith and Vallis 2001; Vallis 2005), we expect to find a cascade (of the first baroclinic mode) toward the deformation scale in the geostrophic surface currents observed with satellite altimetry. But Stewart et al. (1996) found the peak in the spectrum increasing with distance downstream of the ridge from 213- to 313-km wavelength. Because these scales are larger than 126 km $\approx 2\pi L_D$ at these high latitudes (Chelton et al. 1998), theory predicts a baroclinic cascade toward *smaller* scales. So if the altimeter is truly seeing the first baroclinic mode, then the cascade should have been in the opposite direction.

One might speculate that because the Southern Ocean is much more barotropic perhaps this overwhelms the greater surface weighting of the baroclinic mode, and the altimeter is actually seeing more barotropic than first baroclinic mode there. (We will examine this possibility more closely below.) However, the study of Kobashi and Kawamura (2002) was in the subtropics, and similar conclusions were found. For instance, they found a seasonal growth in the centroid of the zonal (meridional) surface velocity spectrum from about 500–700 km (300–400 km) in the North Pacific Subtropical Countercurrent region, defined as 21°–26°N, 140°–170°E. The area average deformation wavelength there is $2\pi \times 57.22 < 360$ km.[1] Slightly larger scales were found in the Hawaiian Lee Countercurrent region defined as 18°–20°N, 170°E–160°W, where the average wavelength of the Rossby radius of deformation is about 400 km. Because the growth in scale is extending well beyond $2\pi L_D$, especially for the zonal velocity, even in these subtropical regions, we claim there is an unresolved, and indeed unappreciated, contradiction between stratified geostrophic turbulence theory and observations. Furthermore, Kobashi and Kawamura also claim to observe the eddies evolving into zonally elongated, wavelike structures, as predicted by Rhines (1975). But again, theory holds this to be a barotropic mode phenomenon that does not apply to the baroclinic modes.

To address this discrepancy between theory and observations, we approached the problem directly by estimating the spectral kinetic energy flux Π_K with altimeter data. As described in the results section below, section 4, we actually found an *inverse* cascade at the surface for all locations in the extratropical South Pacific Ocean, consistent with previous observations. We tentatively propose an extension to geostrophic turbulence theory in section 5 to explain this discrepancy. First, we clarify the origin of the discrepancy by reviewing the relevant stratified, geostrophic turbulence theory in section 2. Then in section 3, we derive the equation containing the spectral energy flux, and present the data used to estimate it. The conclusions are summarized in section 6.

2. Background geostrophic turbulence theory

The cascade properties of stratified, geostrophic turbulence in the atmosphere and the ocean share similarities, of course, but also differ qualitatively because of the difference in density stratification profiles. For the atmosphere the stratification is approximately depth independent throughout most of the troposphere. In the two-layer model, this is represented with a light layer overlying a denser layer of equal mean thickness. Using such a model, Salmon (1980, 1998) presents a simple, powerful argument for a direct cascade for the baroclinic mode (two layers moving in opposite directions), and a cascade diverging away from the deformation scale, with a net inverse cascade for the barotropic mode (two layers moving together). This is shown schematically in terms of the spectral energy flux in Fig. 1 (solid lines). The sun acts as a power source at large scales, driving baroclinic potential and kinetic energy. This baroclinic, total energy cascades downscale in a direct cascade (green line), being conserved over the interial range. Near the deformation radius, most of this energy is transferred to the barotropic mode via barotropization, while a small fraction continues to small-scale dissipation. The barotropic energy flux (blue line), which is almost completely kinetic energy, diverges away from the deformation scale. Most of the energy cascades to larger scales, through a net inverse cascade, though a small fraction cascades to small dissipation scales.

For the ocean the stratification is surface intensified, which leads to qualitative differences. Fu and Flierl (1980) analyzed the rates of triad interactions for density stratification typical of the midlatitude ocean. The main difference from the uniform stratification case is that energy concentrates in the first baroclinic mode, *converges* to the deformation scale, then transfers to the barotropic mode, as suggested by Rhines (1977) based upon numerical simulations of decaying turbulence. All the barotropic energy then cascades upscale in an inverse cascade, also consistent with Rhines (1977). The spectral flux for the first baroclinic and barotropic modes would look the same as in Fig. 1 at scales larger than the deformation scale. However, for $K > 1/L_D$, both fluxes would be negative (fluxing to larger scale), and are represented as dashed lines in Fig. 1. The origin for the discrepancy between the cascade directions at scales smaller than L_D in these two pictures is unclear. In fact, we are unaware of any reference to this discrepancy. At least regarding the baroclinic spectral fluxes, the discrepancy may arise because of Fu and Flierl considering several baroclinic modes, while Salmon only considers the first baroclinic mode (G. Vallis 2005, personal communication). The higher baroclinic modes cascade energy to the first mode at scales smaller than L_D, and from there it cascades upscale. However, Salmon's two-layer model does not have energy input to the

[1] This estimate was made using Chelton et al. (1998) values, which were made available, with permission from Chelton, in netCDF format [http://www.ig.utexas.edu/people/staff/rscott/index.htm].

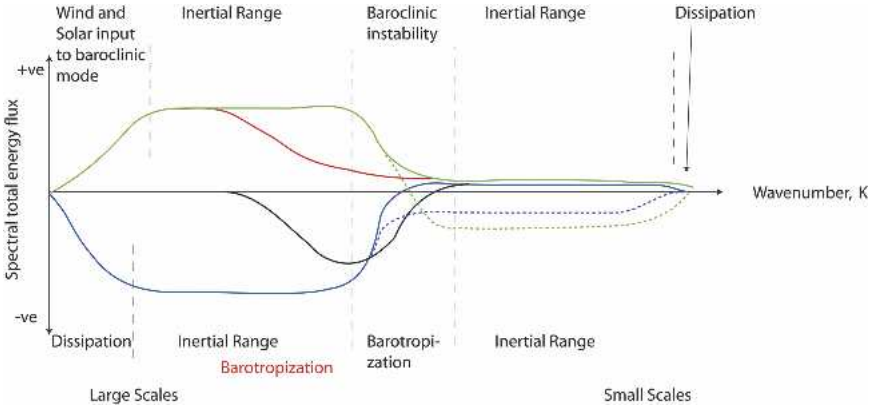

FIG. 1. Schematic of spectral total energy flux Π_K vs total wavenumber K for the baroclinic mode (green line) and barotropic mode (blue line). For the baroclinic mode, the total energy includes a potential energy component. The solid lines are interpreted in terms of spectral flux from Salmon's (1980) Fig. 1 (reproduced in Salmon 1998's Fig. 6.6 and Vallis 2005, his chapter 10). The dashed lines are the modifications to the spectral flux implicit in Rhines's (1977) Fig. 19 and Fu and Flierl's (1980) Fig. 12. The red line is our proposed modification to the baroclinic total energy cascade, which is reduced in amplitude by an upscale flux of kinetic energy shown in black. The black line is assumed to be of similar shape to the universal kinetic energy flux found in the altimeter data, e.g., Fig. 4. Moving upscale, this baroclinic flux peters out, we suspect, because of barotropization.

first baroclinic mode at scales smaller than L_D, so there is no energy to cascade upscale. The discrepancy for the barotropic mode remains obscure to us.

Despite Fu and Flierl's method being "crude," in the authors' own words, these results were indeed consistent with the high-resolution, numerical, quasigeostrophic (QG) simulations of decaying turbulence by Smith and Vallis (2001), although spectral fluxes were not presented in that study. Their focus was on the concentration of energy in the first baroclinic mode as a theoretical explanation for the correspondance between observed eddy length scales in altimeter data and L_D (Stammer 1997). Spectral flux divergences were presented by Hua and Haidvogel (1986), who analyzed numerical solutions of stratified geostrophic turbulence with six vertical modes. They found that, while the isomorphism between 2D and stratified 3D turbulence only strictly applies for the case of constant density stratification, in practice the inverse cascade of the barotropic mode still occurred for an exponential density profile that is surface intensified, and therefore representative of the ocean. Their Figs. 11 and 12 plot $T(K) \equiv -\partial\Pi_K/\partial K$. Integrating "by eye," or observing that there is only one zero crossing one can deduce[2] that $\Pi_K < 0$ for all K for the barotropic mode, consistent with Fu and Flierl (1980) and contrary to Salmon (1980). However, the first baroclinic mode is consistent with Salmon's prediction, since $\Pi_K > 0$ for all K. (Note that the single, small positive point at $K = 3$ in Hua and Haidvogel's Fig. 12b is unlikely to be statistically significant. Unfortunately there are no standard error bars.)

In summary, theory agrees on the cascade direction for scales larger than L_D: there is a direct (inverse) cascade in the first baroclinic (barotropic) mode. Apparently there is an unresolved discrepancy between Salmon (1980) and Fu and Flierl (1980) regarding the cascade directions for $K > 1/L_D$. However, the numerical simulations of Hua and Haidvogel (1986) imply an inverse cascade for all scales in the barotropic mode, and a direct cascade for all scales in the first baroclinic mode. This discrepancy will not be crucial here, since mostly we will be comparing our results with theory at scales larger than L_D, simply because the altimeter is better able to resolve these scales. More fundamental, it seems to us, is the discrepancy between the suggestions of an inverse cascade in altimeter data (all at scales comparable to or larger than L_D) and the consistent theoretical and numerical prediction of a direct cascade in the baroclinic mode at these scales. To address this discrepancy, we measured the spectral kinetic energy flux in the altimeter data, as described in the next section.

[2] This follows from conservation of energy, $\int_0^\infty T(K)\, dK = 0$. With a single zero crossing in $T(K)$, the contribution to Π_K from the positive lobe at low K must be balanced by the negative lobe contribution as $K \to \infty$.

3. Methodology and data sources

a. Derivation of the spectral energy equation

The spectral kinetic energy flux Π_K is the flux of kinetic energy through total wavenumber K. Physically, this flux arises through the advective terms in the kinetic energy equation. They are neither sources nor sinks of energy, since they average to zero on a periodic or enclosed domain, yet they can redistribute energy between different length scales via the "triad interactions" discussed in turbulence literature. To reveal the meaning of Π_K and assumptions behind our estimate, we derive the expression for Π_K following the general method advocated by Frisch (1995, p. 25). For the special case of isotropic or at least homogeneous flow, the spectral energy flux is also related to the longitudinal component of the third-order structure function tensor (Lindborg 1999; Frisch 1995). However, because of the large spatial inhomogeneities in the ocean over the length scales of interest, we require the more general formulation. Frisch starts from the incompressible, forced-dissipative Navier–Stokes equations on a 3D, periodic domain. Here we follow the same development, but applied to the equations of motion applicable to mesoscale, oceanic flow on a doubly periodic, rectangular domain. While the ocean is clearly not a doubly periodic rectangle, this assumption is applied to formally isolate the region. This artificially eliminates flow interactions at the boundaries, but this effect will be small if the region is large enough that interior terms are more significant than boundary terms. Furthermore, isolating a rectangular region artificially imposes a largest length scale (gravest mode of the rectangle) that can be studied. However, we find that an interesting range of length scales can be studied.

We seek an evolution equation for the kinetic energy of the large-scale flow. First, to define the large-scale flow, represent the horizontal current $\mathbf{u} = (u, v)$ by its Fourier series,

$$\mathbf{u}(x, y) = \sum_{k_x,k_y} \hat{\mathbf{u}} \exp i(k_x x + k_y y),$$

where $(k_x, k_y) = (m, n)\, 2\pi/L$, with $m, n \in \mathrm{Z}$. Then we can define the low-pass-filtered and high-pass-filtered fields,

$$\mathbf{u}_K^<(x, y) \equiv \sum_{K'<K} \hat{\mathbf{u}} \exp i(k_x x + k_y y) \quad \text{and}$$

$$\mathbf{u}_K^>(x, y) \equiv \sum_{K'>K} \hat{\mathbf{u}} \exp i(k_x x + k_y y),$$

where the total wavenumber K is the magnitude of the vector wavenumber, that is, $K^2 = (k_x^2 + k_y^2)$. Thus $\mathbf{u}_K^<(x, y) = (u_K^<, v_K^<)$ is the flow field with wavelengths larger than $2\pi/K$. Clearly the associated area averaged kinetic energy density, $\mathrm{KE}_K^<$, is just

$$\mathrm{KE}_K^< = \langle \mathbf{u}_K^< \cdot \mathbf{u}_K^< \rangle / 2,$$

where $\langle\,\rangle$ denotes area average.

An energy evolution equation can be obtained from the geostrophic relative vorticity equation. We use the dimensional version of Eq. (6.3.17) of Pedlosky (1987),

$$\frac{\partial}{\partial t} \nabla^2 \psi = -J(\psi, \nabla^2 \psi + \beta y) + \frac{U}{L} \frac{1}{\rho_s} \frac{\partial}{\partial_z} (\rho_s w_1) - d, \quad (3)$$

where ψ is the streamfunction for the geostrophic current; $(u, v) = (-\psi_y, \psi_x)$; ∇^2 is the 2D horizontal Lapacian operator; J is the Jacobian operator; $J(A, B) = A_x B_y - A_y B_x$; ρ_s is the density of the background state (or the long-term mean); U and L are scales of the velocity and horizontal length, respectively; and w_1 is the dimensionless, order-Rossby-number vertical velocity, and we have added an unspecified dissipation, d. This is the governing equation for motions with horizontal scales large enough for the geostrophic approximation to hold, and with topographic height scales much less than the depth. The evolution of $\mathrm{KE}_K^<$ is obtained by multiplying the relative vorticity Eq. (3) by low-pass-filtered ψ and integrating over the domain. Repeatedly using integration by parts, one finds

$$\frac{\partial}{\partial t} \mathrm{KE}_K^< = -\Pi_K + \mathcal{F} - \mathcal{D}, \quad (4)$$

where

$$\Pi_K = \langle \mathbf{u}_K^< \cdot (\mathbf{u}_K^< \cdot \nabla \mathbf{u}_K^>) \rangle + \langle \mathbf{u}_K^< \cdot (\mathbf{u}_K^> \cdot \nabla \mathbf{u}_K^>) \rangle. \quad (5)$$

The forcing term $\mathcal{F}$ arises from the vortex-stretching term only,

$$\mathcal{F} = \frac{U}{L} \left\langle \psi_K^< \frac{1}{\rho_s} \frac{\partial}{\partial z} (\rho_s w_1) \right\rangle.$$

Note the β term dropped out, which is not surprising because the Coriolis force is perpendicular to velocity, and so does no work. The formal dissipation term $\mathcal{D}$ arises from the d in the vorticity equation.

The forcing term $\mathcal{F}$, we argue below, largely arises from baroclinic instability. But baroclinic instability can also be understood as a triad interaction (Salmon 1980; Vallis 2005), with one wavenumber close to zero (representing the mean flow) losing energy to two other wavenumbers of similar size close to the deformation scale. Indeed this energy flux is the blue branch of Fig. 1. So the question arises as to why we see this process as a source of kinetic energy $\mathcal{F}$, while it can also be represented as a spectral flux. The answer is that it is a spectral flux of total energy. This is not contained in Π_K

because it is a nonlocal triad that allows energy to jump from very large scales to near the deformation scale. The large-scale energy is almost exclusively potential energy, which we cannot measure. However, we see the arrival of kinetic energy near L_D.

While our focus is estimating Π_K from oceanic observations, we discuss the spectral kinetic energy density equation to aid in its interpretation. Differentiating (4) with respect to K, we obtain an evolution equation for the spectral kinetic energy density (or "energy spectrum"),

$$\frac{\partial}{\partial t} E(K) = T(K) + F(K) - D(K). \tag{6}$$

The term $T(K)$ is the convergence of the spectral energy flux, and acts as a source of kinetic energy at total wavenumber K. A formal expression for $T(K)$ involves an integral over all triad interactions [Lesieur 1997, his Eq. VI(3–2)]. Because $T(K)$ involves a further derivative, it is noisier, so we favor estimating Π_K. Furthermore, it is easier to visually estimate the slope of Π_K to obtain $T(K)$ than it is to visually estimate the integral of $T(K)$ to obtain Π_K, as we did to interpret Hua and Haidvogel (1986) in section 2. The long-term mean of (6) requires, for a statistical steady state, a balance between the time mean of the terms on the rhs; the flux convergence $T(K)$ and forcing $F(K)$ must be balanced by the dissipation $D(K)$ at all wavenumbers K. For future reference, in terms of the spectral energy flux, a statistical steady state [or constant $\overline{E}(K)$] requires,

$$\frac{\partial \overline{\Pi}_K}{\partial K} = \overline{F}(K) - \overline{D}(K), \tag{7}$$

where $\bar{x}$ is the long-term mean of x. The forcing term here arises only from vortex stretching, though if we had included other forcing such as surface wind stress in the vorticity Eq. (3), then the forcing term would also include the direct wind work. The form of the dissipation is beyond the scope of this work, though typical parameterizations used in turbulence studies would be

$$D(K) = 2\nu K^2 E(K) + 2rE(K),$$

where ν is the eddy viscosity, and r is the linear bottom drag parameter. See Scott (2001) for their effect on shape of $E(K)$ in freely decaying, 2D turbulence.

b. Data sources and processing

Inspection of (5) reveals why this has remained an elusive quantity; the data requirements are very demanding! We require snapshots of 2D velocity fields, over regions large enough to contain the gravest modes of interest (at least several hundred kilometers) at spatial resolution sufficient to capture the smallest scales of interest (Rossby radius of deformation and smaller). Furthermore, Π_K in (5) is at an instant in time. However, the time series were found to be very erratic, and only through time averaging over *many years* could a statistically significant signal be uncovered.

Fortunately satellite altimeter data are reaching the point where we can evaluate the time mean of (5) with some statistical significance over an interesting range of length scales. The geostrophic velocity anomaly was calculated from the weekly sea surface height anomaly fields optimally interpolated to a ⅓° Mercator grid by Aviso from Ocean Topography Experiment (TOPEX)/Poseidon and the first and second European Remote Sensing Satellites (*ERS-1* and *ERS-2*) data (LeTraon et al. 1998). The time-mean geostrophic velocity was estimated from the long-term mean dynamic topography relative to 2000-m depth (Gill 1982, p. 215) obtained from the temperature and salinity data of the *World Ocean Atlas 2001* (Conkright et al. 2002).

The Fourier decomposition required for calculating (5) was performed with a fast Fourier transform (FFT) algorithm, and was efficiently implemented on regions with 32×32, 64×64, and 128×128 grid points on the ⅓° Mercator grid (i.e., with zonal width either 10⅔°, 21⅓°, or 42⅔°, longitude by a similar meridional width). Smaller regions were found to be too noisy, most likely indicating the importance of boundary terms relative to interior terms. The spatial mean and linear trend in x and y were removed from the SSH, and a Hamming window was applied prior to the FFT. This reduces Gibbs phenomenon by ensuring the SSH smoothly approaches zero at the boundaries. The results were found to be not sensitive to the choice of window. The dependence on the size of region is addressed in Figs. 2 and 3.

c. Summary of assumptions and limitations of the estimate

The equation in (5) for the spectral kinetic energy flux is of identical form as the very general expression obtained by Frisch (1995, section 2.4) from the incompressible Navier–Stokes equation for the full 3D velocity field (though of course then the velocity is the 3D vector field, and the gradient operator is also 3D). Two important simplifications arose in deriving (5) for mesoscale oceanic flow. 1) Because the vertical velocity w is generally much smaller than the horizontal currents u, v, the kinetic energy is dominated by the horizontal currents: $\mathrm{KE} \approx (u^2 + v^2)/2$. 2) The vertical advection of momentum (and hence KE) is much smaller than the horizontal advection. While this may be violated in regions where the bottom topography scale to depth scale

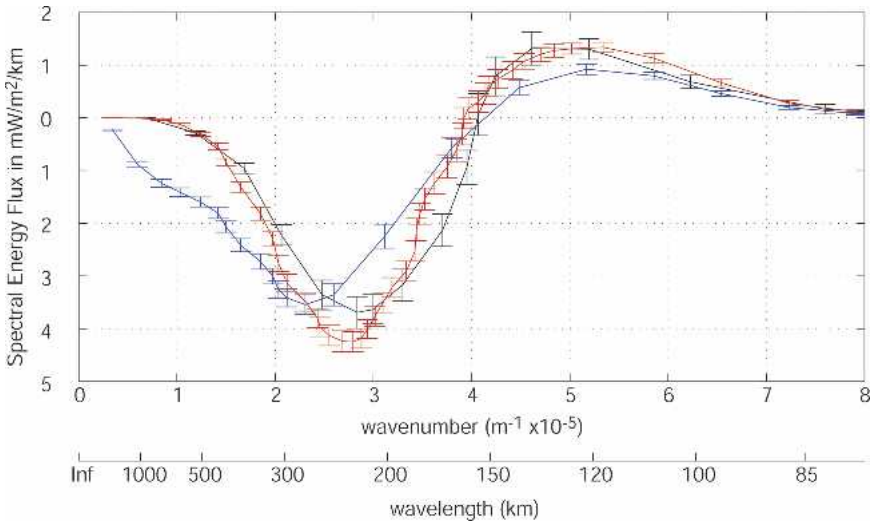

FIG. 2. Time mean, spectral kinetic energy flux $\overline{\Pi}(K)$ vs total wavenumber K in a homogeneous ACC region (rectangles centered at 57°S, 120°W): black curve using SSH on a 32 × 32 grid, red curve using SSH on a 64 × 64 grid, blue curve using velocity on a 64 × 64 grid. Positive slope reveals a source of energy. The larger negative lobe reveals a net inverse cascade to lower wavenumber. Error bars represent standard error.

ratio is greater than order Rossby number, see Pedlosky (1987, p. 348), or perhaps also in frontal regions where the thermal wind shear is large, it generally holds for mesoscale motion in the oceanic interior.

No further assumptions are contained in (5). In fact, (5) is easily generalized to compressible flow, but the results were found to be not sensitive to this. Note that nothing has been assumed about barotropic versus baroclinic modes.

The data sources impose further limitations, however, beyond just measurement error. Using altimeter data restricted us to considering geostrophic currents near the oceanic surface. For the geostrophic balance to hold, we must avoid regions too near the equator (where the Coriolis parameter approaches zero). Away from the equator, and at mesoscale length scales, Ekman currents are likely the strongest ageostrophic currents. At depths greater than a few hundred meters, the geostrophic current may be significantly different than the surface values, because of thermal wind shear (proportional to the horizontal density gradient). Thus the analysis is best interpreted as applying just below the Ekman layer (a few tens of meters). The 10-day sampling time of the TOPEX/Poseidon satellite is too long to resolve barotropic Rossby waves, and other high-frequency motions, that create an aliasing error in the altimeter data (Gille and Hughes 2001). The Nyquist wavelength ranged from about 74 km at the equator to 37 km at 60°S, representing the smallest scales resolved by the data. However, the altimeter data is strongly damped at these smaller scales during the gridding process. Although not strictly necessary, the calculations are much faster with the f-plane approximation, and ignoring the small variations in grid spacing within a given rectangle. Last, as in Frisch's development, the periodic domain was introduced. As discussed above, this imposes the largest length scale interactions that can be studied, and introduces some error through iso-

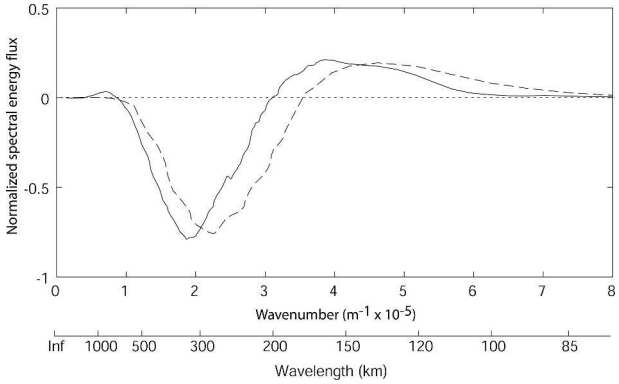

FIG. 3. Time mean, spectral kinetic energy flux normalized by peak-to-peak amplitude $\overline{\Pi}'(K)$ vs total wavenumber K, for region centered at 45°S, 140°W. The solid curve represents an average over 128 × 128 grid points, while the dashed curve is an average over four curves corresponding to surrounding regions with 64 × 64 grid points each. Curves were smoothed with a five-point binomial filter prior to normalization. The reasonable agreement between the two curves lends support to the adequacy of the smaller regions to minimize the boundary effects.

lating the region from interactions at the boundary. Despite these limitations, we found a reasonably clean, statistically significant signal that was easily interpreted and revealed a universal shape, as discussed below.

4. Results

a. Robustness of the spectral kinetic energy flux estimates

In Fig. 2 we plot the spectral kinetic energy flux $\overline{\Pi}_K$, averaged over 559 weeks and over rectangular regions in the Antarctic Circumpolar Current (ACC) region of the South Pacific Ocean, versus total wavenumber K. The black (red) curve shows $\overline{\Pi}_K$ averaged over a 32×32 (64×64) contiguous point rectangle on the ⅓° Mercator grid. Flux Π_K was found to fluctuate wildly in time, because of real variations and because of errors in the altimeter data. The error bars represent the standard error estimated from the 559 week-long time series, assuming independent weeks. Typical SSH decorrelation timescales are 10–20 days (LeTraon et al. 1998), so these are probably roughly $\sqrt{2}$ too optimistic. The red and black curves generally agree within statistical uncertainty, confirming the small influence of boundary terms, neglected in (5) because of periodicity. The remaining discrepancy may be partly due to inhomogeneity over the region producing real differences in $\overline{\Pi}_K$. For the blue curve, the geostrophic velocity was first estimated with central differences on a 64×64 grid; for the other curves the SSH was first transformed to Fourier modes, providing a much higher order approximation to the derivatives in the geostrophic relation. Influences of this methodology (comparing the blue and black curves) reveal more significant discrepancy at lower wavenumber. However, all three results reveal the same general shape, allowing for a qualitative interpretation. Note for instance the robust zero crossing point around 160 km.

To further test the influence of artificially isolating a finite region of the ocean, we performed an analysis for a 128×128 grid point region in the central South Pacific. Because the region is highly inhomogeneous, we normalized $\overline{\Pi}_K$ by its peak-to-peak amplitude (solid line in Fig. 3) and compared this with the average of the normalized $\overline{\Pi}_K$ in the surrounding region obtained with 64×64 grid points (dashed line in Fig. 3). From here on, Π_K normalized in this way will be denoted Π'_K. Since the larger region was over 3300 km across, the calculation accuracy is compromised by the variations in grid spacing and Coriolis parameter across the region. However it provides an opportunity to further assess the influence of boundary terms and finite gravest mode. While there does appear to be a slight shift toward longer wavelengths in Fig. 3 for the larger domain size, the shift is quite small considering the gravest mode of the domain has been doubled. We interpret this as further evidence that the 64×64 gridpoint region is of adequate size to capture the major part of the $\overline{\Pi}_K$ signal. This was not the case, however, for the spectral kinetic energy density estimates, discussed below. Note the small positive peak at very low wavenumber that is only visible for the larger region (solid line). This may be a hint of the direct cascade of baroclinic energy driven via stirring by the barotropic mode (Salmon 1980). To analyze this further would require analysis of even larger regions. But even for the largest ocean basins, it is not possible to increase the meridional extend of the regions much further (certainly not by a factor of 2).

b. The inverse cascade

For length scales where $\overline{\Pi}_K$ has positive slope (i.e., the flux is divergent), the nonlinear terms tend to remove energy in the long-term mean, which implies a power source to maintain a constant long-term mean spectral energy density, see (7). Denoting K_f as the point of the zero crossing, the negative (positive) region with $K < K_f$ ($K > K_f$) has energy flowing to larger (smaller) scales; moreover, $\overline{\Pi}_K$ has positive slope at $K = K_f$, indicating a source near that wavenumber. That more energy flows to lower wavenumber implies a net inverse cascade.

If indeed the altimeter data predominately reveals the first baroclinic mode, as discussed in the introduction, then the observations actually contradict the theory of barotropic straining of baroclinic potential vorticity yielding a forward tracer-like cascade of baroclinic energy toward smaller length scales (Salmon 1980). Recall that this theory is grounded in the numerical simulations of Hua and Haidvogel (1986). To resolve this discrepancy, the choices seem to be to accept an inverse cascade in the first baroclinic mode, at least at scales slightly larger than the deformation scale, or to interpret the altimeter data as reflecting the barotropic mode, at least for the triads leading to the inverse cascade. The latter possibility is addressed in section 4d. The former possibility requires a modification to turbulence theory.

Convergent $\overline{\Pi}_K$ (i.e., negative slope) implies a kinetic energy sink for those length scales. Figure 2 reveals two such regions, one at large and one at small length scales. The decay of energy for wavelengths smaller than about 150 km is likely due more to the finite SSH grid resolution than to real energy dissipation; small scales are artificially removed from the altimeter data during the gridding process. The energy sink at large scales

represents an arrest of the inverse cascade, and is discussed in section 4d.

The qualitative results in Figs. 2 and 3 were found for almost all regions of the World Ocean that we studied, revealing a near-universal character of the spectral kinetic energy flux. The results of this section are only a subset of a more extensive analysis that included almost all of the North Pacific, and some regions in the South Atlantic. We found the main ideas were completely presented by only showing the South Pacific results.

Interestingly, the shape of the spectral flux looks intriguingly similar to the corresponding result in the atmosphere (Boer and Shepherd 1983, their Fig. 10a). They estimated the kinetic energy flux through spherical harmonic spectral space, averaged over either the globe or the Northern Hemisphere, and over depth.[3]

Since the amplitude of the flux typically varied by an order of magnitude across the South Pacific, it was advantageous to normalize by its peak-to-peak amplitude. Each subplot of Fig. 4 shows the normalized spectral kinetic energy flux $\overline{\Pi}'_K$ for various longitudes between 170°E and 80°W in the South Pacific along a given latitude band between 55° and 15°S. Each line plots $\overline{\Pi}'_K$ averaged over a 64 × 64 point rectangle on the ⅓° Mercator grid. Regions with more than five missing data points, due primarily to land or ice, were removed. As in Fig. 2, we see for all regions in the 55°–25°S latitude bands a clear, net inverse cascade with about a 3:1 ratio between the upscale (negative lobe) and downscale (positive lobe) fluxes. The data are noisier at 15°S but also reveal a net inverse cascade.

The magnitude of the spectral kinetic energy flux in Fig. 2 is presented as an energy flux per kilometer of ocean depth. This was necessary because the depth of the currents is not known. However, if the altimeter is predominately reflecting the first baroclinic mode, then an order of magnitude estimate is that of the depth of the main thermocline, which is about a kilometer. This implies fluxes of a few milliwatts per meter squared. We argue below that this power comes from baroclinic instability, which is a conversion of mean flow energy at very large scale (and hence largely potential energy) to eddy energy. Ultimately we require an equivalent energy source at the large scales to maintain a steady state. Scott (1999b) estimated the mean wind power input to the surface geostrophic flow in the North Pacific using altimeter data, JGM-3 geoid model, and National Centers for Environmental Prediction (NCEP) winds and found an average of 0.9 ± 0.4 mW m^{-2}, where the error estimate is one standard deviation of the mean, estimated using both the NCEP winds and the gravity model error covariance matrix; see paper for methodology. For the Southern Ocean, Scott (1999a) found 3.73 ± 0.28 mW m^{-2}. A global average of 2.68 ± 0.17 mW m^{-2} is similar to that found independently by Wunsch (1998), though he did not estimate the error bars. Oort et al. (1994) found that the buoyancy forcing provided similar power input as the wind stress, though others contend that this must be zero (Wang and Huang 2002; Wunsch and Ferrari 2004). We conclude that the spectral kinetic energy flux found here is of the correct order of magnitude, if it reflects the motions as deep as the main thermocline.

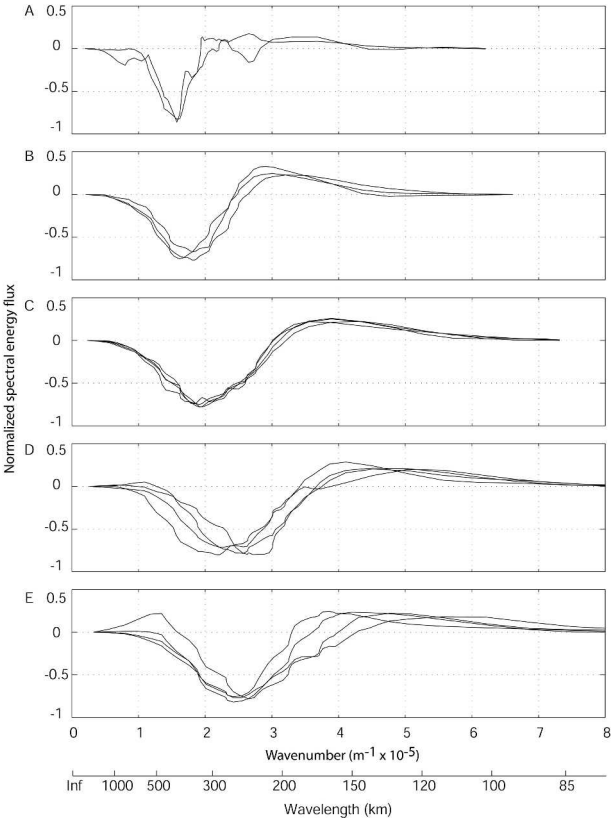

FIG. 4. Time-mean, spectral kinetic energy flux normalized by peak-to-peak amplitude $\overline{\Pi}'(K)$ vs total wavenumber K for various longitudes in the South Pacific. Latitude bands centered at (a) 15°, (b) 25°, (c) 35°, (d) 45°, and (e) 55°S. Curves were smoothed with a five-point binomial filter prior to normalization. Each curve represents an average over 64 × 64 grid points.

c. The length scale of the kinetic energy source

The plots in Fig. 4 reveal a shift to lower wavenumber approaching the equator. The shift in the source can be understood by linear instability theory, as we now discuss. The change in the arrest scale with latitude is taken up thereafter, but remains less clear.

In the extratropics, a likely source of kinetic energy

[3] We thank Greg Flato for raising this question.

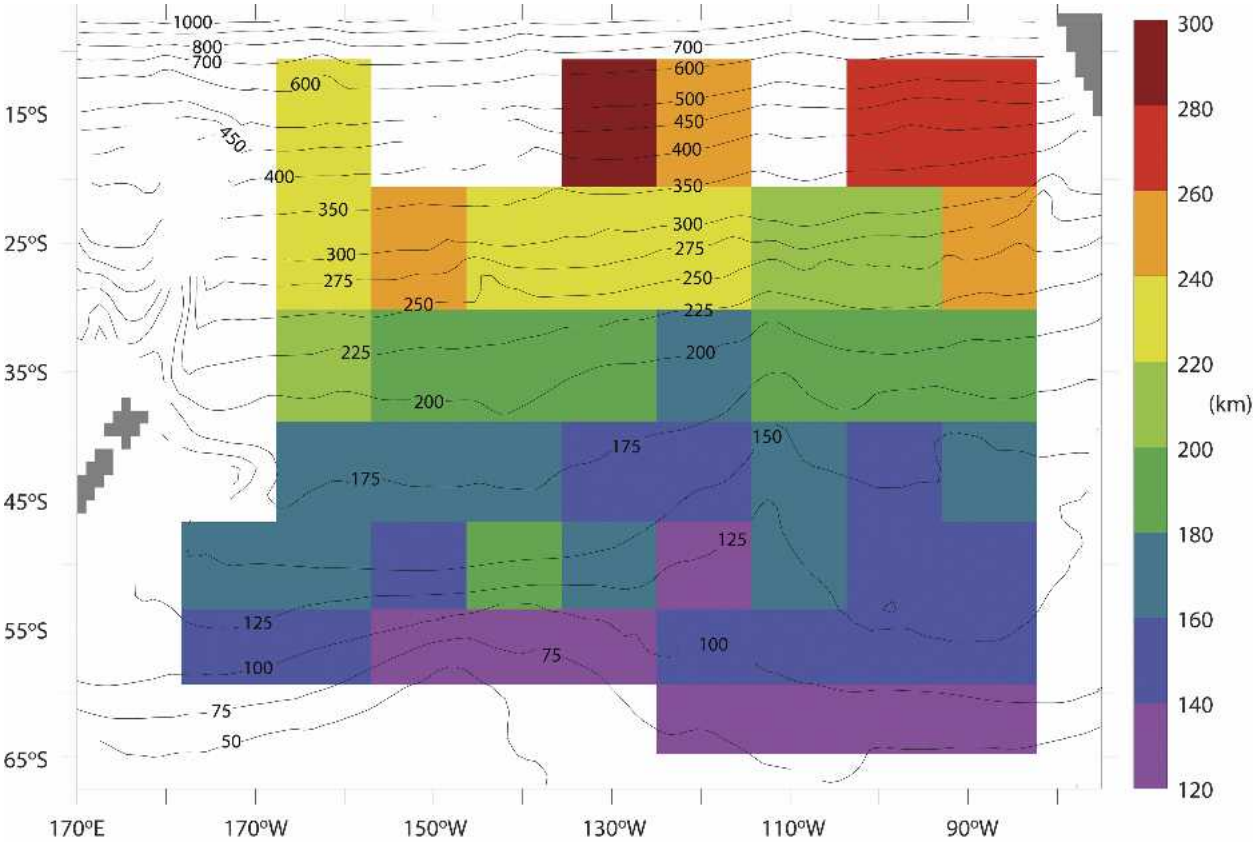

FIG. 5. Energy source wavelength (estimated from the zero of $\overline{\Pi}_k$) in color (km); rough estimate from linear instability theory, $2\pi L_D$, shown in contours (km).

near K_f is baroclinic instability (Stammer 1997, and references therein). This is the process whereby potential energy is converted to kinetic energy in a rotating, stratified fluid. Indeed, this likelihood is supported by linear stability theory since the most unstable wavelength for first mode baroclinic instability is around $2\pi L_D$ (where L_D is the first Rossby radius of deformation) and this matches closely the zero crossing wavenumber K_f. (Actually, a better comparison with the most unstable wavelength would be the wavelength of the steepest slope of $\overline{\Pi}_K$. However, given the limitations of the data, the zero crossing formed a more robust approximation to this than actual estimates of the steepest slope; see Fig. 2.) In Fig. 5 we compare $2\pi L_D$ (contours) with $2\pi/K_f$ (shaded colors), where L_D was taken from Chelton et al. (1998). Note that both length scales tend to increase substantially approaching the equator. The quantitative agreement is reasonable south of about 30°S. But note that it gets systematically worse approaching the equator, with L_D greater than the length scale of the energy source. This is consistent with two predictions from Phillips's model of baroclinic instability (Pedlosky 1987, p. 561): (i) analysis predicts length scales less than L_D being excited in regions of negative shear, which applies to most of the tropical Pacific; (ii) also the thermocline is shallower approaching the equator, which further decreases the length scales of instability relative to L_D.

Bo Qiu (2004, personal communication) pointed out that closer to the equator, baroclinic instability is inhibited while barotropic instability becomes more important (see Qiu and Chen 2004), so the northernmost row of rectangles in Fig. 5 (centered on 15°S) may overlap regions with both barotropic and baroclinic instability. However, the mechanism of barotropic instability is the Reynolds stresses, which are explicitly resolved by the spectral kinetic energy flux. Hence, this process would provide a positive contribution to $\overline{\Pi}_K$ between the length scale of the mean flow and the eddy generation. This should be contrasted with baroclinic instability, which converts potential to kinetic energy, and hence appears as a source for kinetic energy. In short, barotropic instability appears in $\overline{\Pi}_K$ in (7), while baroclinic instability appears in $F(K)$ in (7).

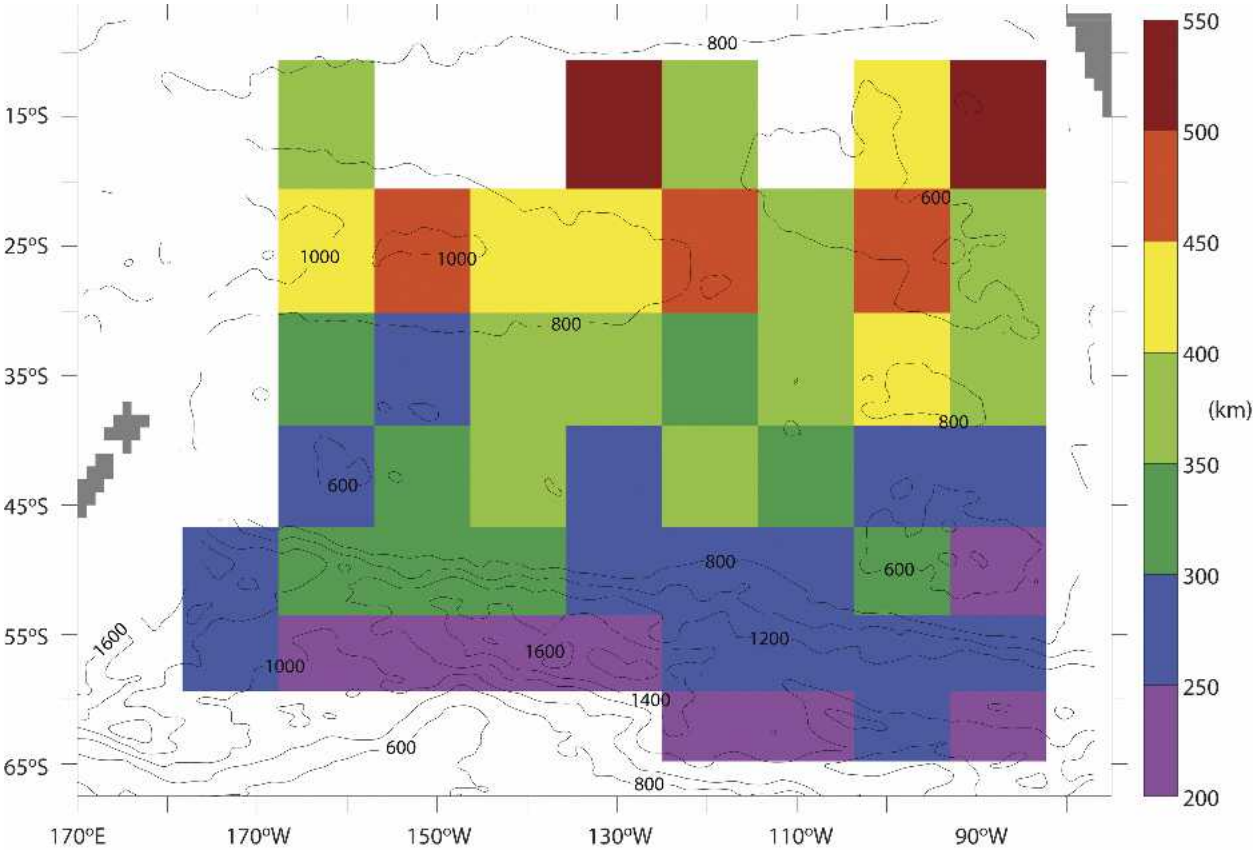

FIG. 6. Arrest wavelength (estimated from the minimum wavelength at which $\overline{\Pi}'_k = -1/2$), in color (km); estimate from β effect, $2\pi L_R$, shown in contours (contour interval = 200 km).

d. The length scale of the cascade arrest

As stated above in section 4b, one possible solution to the conflict between the observed inverse cascade and geostrophic turbulence theory is to assume that the upscale flux of kinetic energy is in the barotropic mode. We emphasize again that this is unlikely to be the case, given the argument leading up to (2). We now show that, even with this assumption of the predominance of the barotropic mode, the results are also not consistently in agreement with theory.

The theory of Rhines (1975, 1977) predicts that the upscale flux of kinetic energy in the barotropic mode is arrested on a rotating sphere because of the β effect. Various estimates of the arrest scale have evolved from this theory (Vallis and Maltrud 1993; Danilov and Gurarie 2000). We used the original Rhines scale ($L_R = \sqrt{2U_{\text{rms}}/\beta}$, see contours in Fig. 6); the scale of Holloway and Hendershott (1977) had similar spatial structure but with greater variation in amplitude over the basin. For comparison, the arrest scale was estimated as the first point at which $\overline{\Pi}'_K = -1/2$ (see colors in Fig. 6). This is an admittedly arbitrary definition of arrest scale. Recall from Figs. 2 and 3 that the arrest scale appears to be the least robust feature, depending upon the methodology used to estimate $\overline{\Pi}_K$. These results in particular then need to be interpreted with some caution, and the specific values are not to be trusted. (Furthermore, the factor of 2 in the definition of L_R is arbitrary; Danilov and Gurarie 2000). However, the *regional variations* in arrest scale were found to depend much less on the computation methodology, and so we focus our attention on this aspect of the theory to data comparison.

For much of the basin the arrest-scale estimate is roughly $L_R/2$ and there is even a hint of a local maximum in arrest scale corresponding to the tongue of large L_R along the high kinetic energy band centered at 25°S. (A similar and more convincing correspondence was observed along the Kuroshio Extension in the North Pacific.) However, the prominent maximum in Rhines scale associated with the high kinetic energy of the ACC has *no* correspondence with the arrest-scale estimate, which is in fact a minimum there.

The simple interpretation of this lack of correspon-

dence between the arrest and L_R is that, as we suggested in the introduction, the altimeter signal is largely the first baroclinic mode, and the arrest by β is for the barotropic mode. However, some more subtle factors may also be at play.

Caveats to the arrest at L_R have been put forward. For instance, LaCasce (2002) argued for arrest by basin modes, rather than free Rossby waves, with the implication that the arrest should be isotropic, but still scales with L_R. Recently Theiss (2004) has proposed that only in regions where $L_D > L_R$ are the Rossby waves fast enough to arrest the cascade. But $L_D < L_R$ almost everywhere except within a narrow band within a few degrees of the equator. For example, L_D is several times smaller than L_R along 25°S, and yet L_R appears to have some association with the arrest scale there. Furthermore, Theiss used the dispersion relation for baroclinic Rossby waves, as opposed to the barotropic dispersion relation that is usually used. The external Rossby radius (~2000 km) is almost everywhere larger than the Rhines scale, and so the barotropic Rossby waves are fast enough to arrest the cascade. Eric Kunze (2005, personal communication) has made the interesting suggestion that topography may play a role in arresting the cascade through, for instance, topographic β; J. LaCasce added that steep topography can inhibit barotropization, making the ocean appear more equivalent barotropic. If so, the results of Larichev and McWilliams (1991) apply; they found the inverse cascade is slowed down considerably above the deformation scale in a 1.5-layer model. While undoubtedly bottom topography is important (see, e.g., Rhines 1977; Treguier and Hua 1988), we did not find a systematic relationship between topography and cascade properties.[4] Thus we seek another explanation.

We mention two processes that possibly affect the upscale cascade. First, barring the effects of steep topography mentioned above, we expect a "barotropization"—that is, upscale cascade in vertical mode accompanying the cascade to larger horizontal scale, which produces an apparent arrest. That is, the barotropization transfers kinetic energy down the water column, reducing the surface expression and decreasing the signal seen by the altimeter data. Furthermore, barotropic motions interact with bottom topography, which may provide the ultimate sink of oceanic mesoscale eddies. At lower latitudes where the energy source is much closer to L_R, the cascade may be arrested by the β effect before the barotropization can substantially reduce the surface signal. The extensive subsurface data required to verify this possibility with oceanic observations are, unfortunately, not currently available. A second possibly important process is that kinetic energy in the upscale cascade is converted to potential energy (D. Straub 2004, personal communication). However, this is not entirely convincing because this implies a systematic correspondence between the apparent arrest and the deformation scale, which is clearly missing in Fig. 8, described later.

In short, it is *not clear* what is leading to the arrest, or apparent arrest, of the upscale kinetic energy cascade seen in the altimeter data. Thus we conclude that the lack of correspondence between the apparent arrest scale, and the Rhines scale *cannot* be used, at this time, to confirm that the altimeter is definitely seeing the first baroclinic mode. Barotropization is consistent with our observations, but remains speculative.

e. The spectral kinetic energy density and flux, and inertial ranges

Much of turbulence theory is based upon the existence of an inertial range, where the energy and enstrophy cascading away from the source scale is essentially conserved. An inertial range in the enstrophy cascade, where K is larger than the local maximum in $\overline{\Pi}_K$, may be affected by the resolution limitations of the altimeter data.[5] Henceforth, we focus on the upscale cascade range, roughly $K < K_f$, that is well resolved.

An inertial range requires $\overline{\Pi}_K \approx$ constant over an extended range in wavenumber K. Yet this is conspicuously absent in all curves in Figs. 2, 3, and 4. This lack of an inertial range explains why Stammer (1997) observed that altimeter derived eddy length scales were closer to L_D than L_R. For if an energy cascade inertial range were present in the surface geostrophic flow, then this kinetic energy would continue to cascade from the source near L_D up to the L_R, providing a maximum of spectral kinetic energy density near L_R.

Note that Smith and Vallis (2001) also attempt to explain Stammer's observation. Starting with the view that the altimeter reflects the first baroclinic mode, they use numerical simulations of decaying, QG turbulence to show that the surface intensified stratification leads to an accumulation of kinetic energy in the first baro-

[4] A member of the audience at a presentation of this work at the Canadian Centre for Climate Modelling and Analysis (CCCma) (seminar September 2004 in Victoria, British Columbia, Canada) suggested that this is due to the bottom topography not affecting the baroclinic modes as strongly as the barotropic mode.

[5] However, it was pointed out by the audience at CCCma (seminar September 2004 in Victoria, British Columbia, Canada), the most convincing inertial ranges appear for the enstrophy cascade region of Fig. 4e, suggesting that perhaps resolution is *not* a problem!

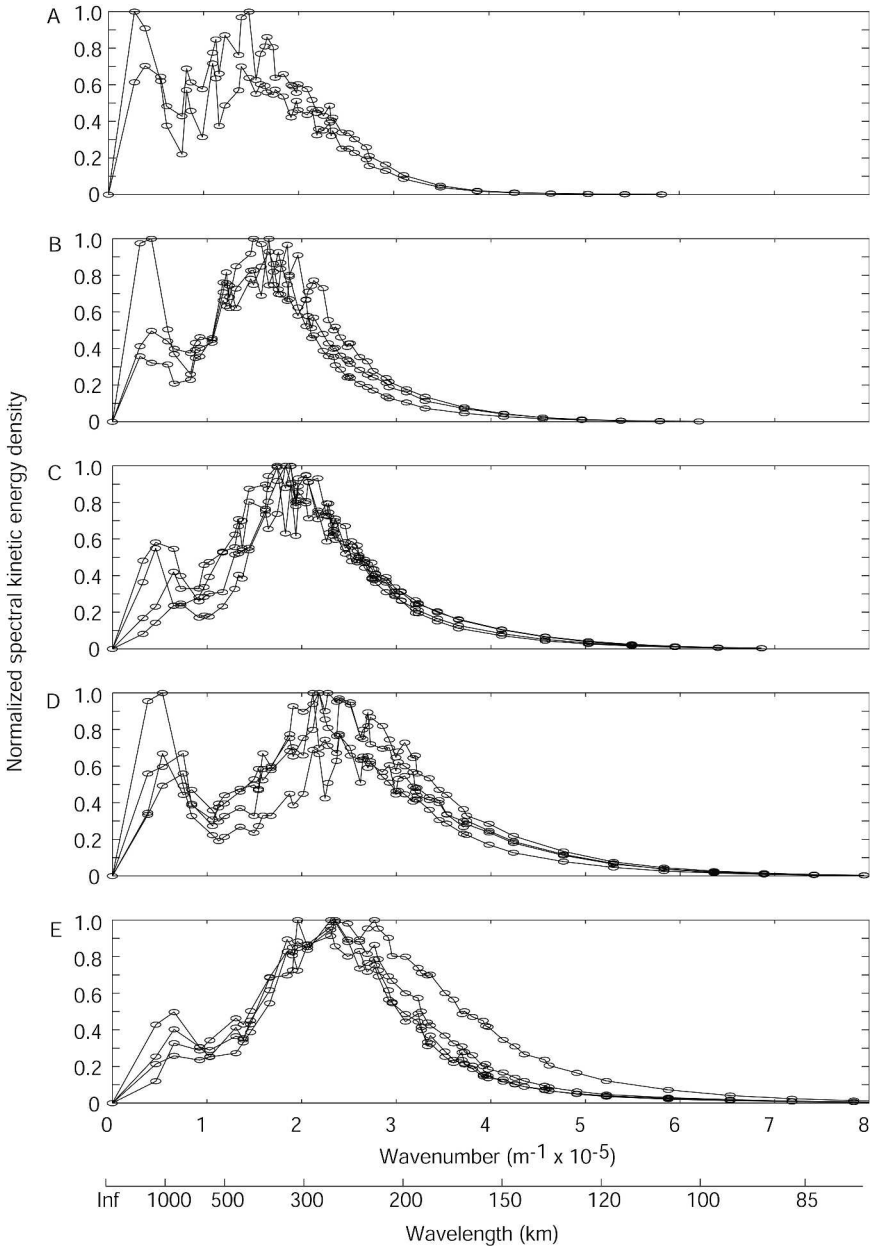

FIG. 7. As in Fig. 4, but for the spectral kinetic energy density normalized by peak amplitude $E'(K)$.

clinic mode. Because the inverse cascade is generally believed to only apply to the barotropic mode, they argue that the altimeter data should not reflect the Rhines scale. However, this argument is inconsistent with the observations, presented here, that reveal a ubiquitous inverse cascade in the surface geostrophic flow. In other words, the answer is not that there is no inverse cascade for the surface kinetic energy, but rather that it is arrested long before the Rhines scale.

In Fig. 7 we plot the spectral kinetic energy density, $E(K)$, as a function of total wavenumber K. The subplots correspond to the same latitude bands as in Fig. 4, and of course used the same data. Again the spectra are roughly independent of longitude, but show a similar shift to larger scales approaching the equator. To facilitate comparison between Figs. 4 and 7, we average over longitudes and plot the corresponding curves on the same plots (Fig. 8). It is now clear that for all latitude bands, the length scale of the arrest of the spectral kinetic energy flux corresponds to the left shoulder of the

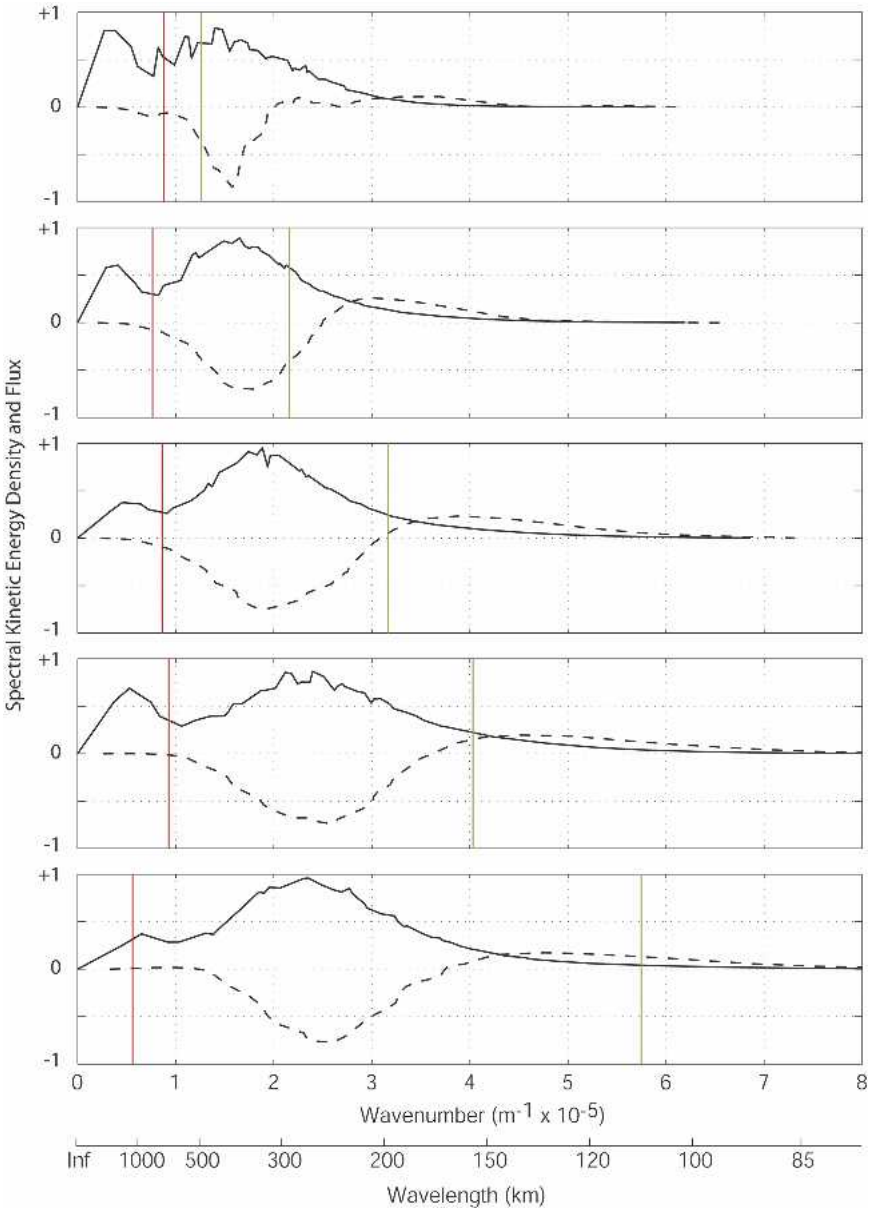

FIG. 8. The zonal average of $E'(K)$ from Fig. 7, solid line; zonal average of $\overline{\Pi}'_K$ from Fig. 4, dashed line.

main peak in the spectral kinetic energy density. Thus the arrest of the kinetic energy cascade explains the dominant scales in the mesoscale, surface geostrophic flow. What in turn explains the arrest is still unclear (see previous section).

In the lower latitudes we see evidence of a Rhines arrest, but in the ACC region another explanation is required, such as hypothesized in section 4d. This lack of an inertial range implies that the standard $K^{-5/3}$ scaling estimate of the kinetic energy spectral slope is not applicable to the surface geostrophic circulation. Indeed our calculations of such spectra in the extratropical Pacific based upon both gridded and along-track altimeter data did not reveal a constant spectral slope upscale of the energy source (where one should find the inverse cascade), consistent with previous estimates (Stammer 1997). Downscale of the source a near-constant spectral slope (enstrophy cascade) was often found, but the data are less believable in that range because of lack of spatial resolution.

It is interesting to note in Fig. 7 the second peak in $E(K)$ at lower wavenumber, which to our knowledge

has not been previously observed. We confirmed that this peak represents the large-scale time mean flow. (The peak disappears when we remove the time-mean current and replot the spectra, not shown.) Not too surprisingly, the size the low wavenumber peak was sensitive to the size of the rectangular region used. For instance, using the data from Fig. 3 with 128×128 points, the time-mean peak was much larger (not shown). Despite this limitation, Fig. 8 nicely displays the spectral gap between the mesoscale flow, driven by baroclinic instability at high wavenumber, and the time-mean flow, driven by surface wind stress at very low wavenumber. It is not surprising that the mean flow has such large scale since it is driven largely by time-mean wind stress and buoyancy flux that also have very large scales. However, it was not clear a priori that there should be a spectral gap. That there is a gap is good news in the sense that the eddy viscosity closure, generally considered only an ad hoc, pragmatic necessity of low-resolution climate models, is actually more mathematically justified when there is a gap (see Lesieur 1997 and reference therein). Perhaps this is relevant in explaining the surprising usefulness of the low-resolution, noneddy-revealing versions of these models.

5. Discussion

In this section we speculate that the discrepancy between the observed cascade direction and the geostrophic turbulence theory may reflect a limitation in the theory. We tentatively propose that there may be an inverse energy cascade in the first baroclinic mode. This would of course explain our observations presented herein, while maintaining the interpretation of the altimeter reflecting primarily the first baroclinic mode. While this requires modifying the cascade schematic of Fig. 1, the modification is rather minor. The proposed upscale kinetic energy cascade (black line) would only partially reduce the forward flux of total, baroclinic energy; recall this total energy forward flux is the *source* of the kinetic energy, arriving near the deformation scale from the large-scale mean flow via baroclinic instability. Some of this source (we estimate about 1/4) continues to cascade toward higher wavenumber. So only about 3/4 is available to cascade upscale. In Fig. 1 we have drawn our modification to the total, direct flux as the red line. That is, the red line is the sum of the blue line and the black line.[6] Near the baroclinic instability region the upscale baroclinic energy reduces the forward total energy flux by about 3/4. Moving upscale, this reduction tapers out to zero, leaving just the forward flux of total energy (green line). The reduction tapers out, we suspect, because of barotropization. Hence we have indicated this barotropization in red letters, below the proposed upscale flux of baroclinic energy. Where the black line becomes positive, there is no longer a correction, and so the red line meets the green line. The black line is related to the spectral kinetic energy flux found in the altimeter data, but includes a potential energy component that we are unable to reveal with the altimeter data.

To some readers, it may seem radical to propose a change in established turbulence phenomenology. But we note that the forward cascade at very low wavenumber remains unchanged, as it must since this is well grounded theoretically. At very large scales the evolution of the baroclinic mode streamfunction approaches that of the passive tracer being advected by the barotropic flow (Rhines 1977; Salmon 1980; Vallis 2005). Indeed, we see some evidence for this in Figs. 3 and 4, albeit rather unimpressive evidence. Furthermore, recall that the blue line in Fig. 1 is mostly invisible to our measurements of Π_K, as explained in section 3a, because it is mostly the potential energy flux associated with baroclinic instability. This flux proceeds not by a local cascade, but jumps from very large-scale potential energy to near the deformation scale, via a nonlocal triad.

Our proposed change only becomes significant as one approaches the deformation scale. Here the passive tracer argument becomes less applicable and one must turn to numerical simulations to sort out this case. The turbulence phenomenology is based upon numerical experiments that reflect the behavior of a highly barotropic ocean. For instance, in Hua and Haidvogel (1986), the ratio of barotropic to baroclinic kinetic energy was about 3, which is about 3 times what current-meter moorings suggest for the real ocean (Wunsch 1997). Thus the stirring of the baroclinic mode by the barotropic mode is exaggerated in these simulations. Furthermore, the decaying simulations of Smith and Vallis (2001) typically have very large barotropic to baroclinic energy ratios; see their Table 1. The one exception, though, is case D with ratio about 1, and it is the only case with final state baroclinic length scales *larger* than the first deformation scale. This suggests that energy in the baroclinic mode may have cascaded toward larger scales, away from the deformation scale, in this particular experiment.

To other readers, this may seem like a small detail in the turbulent cascade phenomenology, or even that our view is consistent with previous theories. Indeed, the direction of *total* energy flux in the baroclinic mode

[6] Keep in mind they were drawn by hand and therefore do not sum perfectly.

remains unchanged. However, insofar as previous theories were ambiguous about the direction of the *kinetic* energy flux near the deformation scale, we have simply pinned down that direction. We emphasize however, that our view has real implications for the surface observations of the ocean. Indeed it was necessary to reconcile the previous observations as well as our observations with the standard interpretation of the altimeter data as primarily reflecting the first baroclinic mode (Smith and Vallis 2001; Wunsch 1997).

6. Conclusions

The universal shape of the spectral kinetic energy flux presented here reveals a ubiquitous source of kinetic energy at mesoscale wavelengths. The length scale is consistent with the power source resulting from baroclinic instability. This supports the possibility that eddy kinetic energy in midocean gyres results largely from baroclinic instability. Spall (2000) explored this possibility using an idealized model of the wind-driven gyre, and showed that because the meridional flows are more unstable than the traditionally studied zonal flows, large-eddy kinetic energy can be extracted from the mean flow potential energy. A complementary result was found by Arbic and Flierl (2004) for a doubly periodic, homogeneous, QG model. Studies using hydrographic data have also revealed the importance of baroclinic instability (Qiu and Chen 2004; Kobashi and Kawamura 2002; Qiu 1999), but were focused on specific regions.

At the surface, about 3/4 of the kinetic energy cascades upscale, while only about 1/4 cascades downscale. This confirms the existence of an inverse cascade of kinetic energy suggested previously in altimeter data (Stewart et al. 1996; Kobashi and Kawamura 2002). However, this runs contrary to stratified, geostrophic turbulence theory. There is strong evidence that the altimeter signals predominantly reflect the first baroclinic mode and hence the motions of the main thermocline (Wunsch 1997). With this assumption we find the magnitude of the spectral kinetic energy flux is of the same order as the flux of mechanical energy from the surface wind stress to the surface geostrophic flow (Scott 1999a,b; Wunsch 1998). However, theory predicts a cascade of baroclinic energy toward the deformation scale, exactly opposite to what we found here. In the discussion section we outlined our arguments why the theory may have to be modified to accommodate these observations.

The arrest of the cascade is also mysterious. We found a vague suggestion at low latitudes that the Rhines scale may be relevant, but at high latitudes it clearly is not. This may indicate an arrest (or more properly an "apparent arrest") in kinetic energy flux arising from barotropization, which transfers the energy down the water column, reducing the surface signal seen by the altimeter.

The results here demonstrate that altimeter data have reached an accuracy, spatial resolution, and duration to address fundamental questions of ocean dynamics, such as the source of eddy kinetic energy and the existence of the inverse cascade, with real data. While far from resolving these questions, the observations can at least play a role in transforming the questions to a higher level.

Acknowledgments. The altimeter products have been produced by the CLS Space Oceanography Division as part of the European Union's Environment and Climate project AGORA (ENV4-CT9560113) and DUACS (ENV4-CT96-0357) with financial support from the CEO programme (Centre for Earth Observation) and Midi-Pyrenees regional council. Data were made available by the AVISO/Altimetry operations center. Hydrographic data have been provided by the U.S. National Oceanographic Data Center via the the *World Ocean Atlas 2001*. We thank Dudley B. Chelton, Roland A. deSzoeke and Michael G. Schlax for making the Rossby radius data available online at http://www.coas.oregonstate.edu/research/po/research/chelton/index.html. The interpretation and presentation herein benefited from discussion with Shafer Smith, David Straub, Brian Arbic, Bo Qiu, Eric Kunze, Joe LaCasce, Bach Lien Hua, Charles Jackson, and Christina Holland, which is greatly appreciated by the authors. Rick Salmon and an anonymous reviewer also provided useful feedback, for which we are very thankful. We also thank the Canadian Centre for Climate Modelling and Analysis and the Institute of Ocean Sciences, in British Columbia, Canada, for hosting seminars in which we received useful feedback on this work. Both authors were supported by the G. Unger Vetlesen Foundation for this project. This work is dedicated to Rupert Ford, whose boundless energy and enthusiasm were an inspiration to our field.

REFERENCES

Arbic, B. K., and G. R. Flierl, 2003: Coherent vortices and kinetic energy ribbons in asymptotic, quasi two-dimensional f-plane turbulence. *Phys. Fluids,* **15,** 2177–2189.

——, and ——, 2004: Effects of mean flow direction on energy, isotropy, and coherence of baroclinically unstable beta-plane geostrophic turbulence. *J. Phys. Oceanogr.,* **34,** 77–93.

Batchelor, G. K., 1969: Computation of the energy spectrum in homogeneous two-dimensional turbulence. *Phys. Fluids Suppl.,* **12B,** 233–239.

Boer, G. J., and T. G. Shepherd, 1983: Large scale two-dimensional turbulence in the atmosphere. *J. Atmos. Sci.,* **40,** 164–184.

Charney, J., 1971: Geostrophic turbulence. *J. Atmos. Sci.,* **28,** 1087–1095.

Chelton, D. B., R. A. deSzoeke, M. G. Schlax, K. E. Naggar, and N. Siwertz, 1998: Geographical variability of the first-baroclinic Rossby radius of deformation. *J. Phys. Oceanogr.,* **28,** 433–460.

Conkright, M., R. Locarnini, H. Garcia, T. O'Brien, T. Boyer, C. Stephens, and J. Antonov, 2002: *World Ocean Atlas 2001: Objective Analyses, Data Statistics, and Figures.* CD-ROM documentation, National Oceanographic Data Center, 17 pp.

Danilov, S. D., and D. Gurarie, 2000: Quasi-two-dimensional turbulence. *Physics-Uspeki,* **43,** 863–900.

Frisch, U., 1995: *Turbulence: The Legacy of A. N. Kolmogorov.* Cambridge University Press, 296 pp.

Fu, L.-L., and G. R. Flierl, 1980: Nonlinear energy and enstrophy transfers in a realistically stratified ocean. *Dyn. Atmos. Oceans,* **4,** 219–246.

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

Gille, S. T., and C. W. Hughes, 2001: Aliasing of high-frequency variability by altimetry: Evaluation from bottom pressure recorders. *Geophys. Res. Lett.,* **28,** 1755–1758.

Holloway, G., and M. Hendershott, 1977: Stochastic closure for nonlinear Rossby waves. *J. Fluid Mech.,* **82,** 747–765.

Hua, B. L., and D. B. Haidvogel, 1986: Numerical simulations of the vertical structure of quasi-geostrophic turbulence. *J. Atmos. Sci.,* **43,** 2923–2936.

Kobashi, F., and H. Kawamura, 2002: Seasonal variation and instability nature of the North Pacific Subtropical Countercurrent and the Hawaiian Lee Countercurrent. *J. Geophys. Res.,* **107,** 3185, doi:10.1029/2001JC001225.

Kolmogorov, A. N., 1941a: Dissipation of energy in the locally isotropic turbulence (English translation). *Dokl. Akad. Nauk SSSR.,* **32,** 19–21. [Reprinted in *Proc. Roy. Soc. London,* **434A,** 15–17.]

——, 1941b: The local structure of turbulence in incompressible viscour fluid for very large Reynolds number (English translation). *Dokl. Akad. Nauk SSSR.,* **30,** 299–303. [Reprinted in *Proc. Roy. Soc. London,* **434A,** 9–13.]

Kraichnan, R. H., 1967: Inertial ranges in two-dimensional turbulence. *Phys. Fluids,* **10,** 1417–1423.

LaCasce, J. H., 2002: On turbulence and normal modes in a basin. *J. Mar. Res.,* **60,** 431–460.

——, and A. Bower, 2000: Relative dispersion in the subsurface North Atlantic. *J. Mar. Res.,* **58,** 863–894.

Larichev, V. D., and J. C. McWilliams, 1991: Weakly decaying turbulence in an equivalent-barotropic fluid. *Phys. Fluids,* **3A,** 938–950.

Lesieur, M., 1997: *Turbulence in Fluids*. 3d ed., Fluid Mechanics and Its Applications, Vol. 40, Kluwer Academic, 515 pp.

LeTraon, P.-Y., F. Nadal, and N. Ducet, 1998: An improved mapping method of multisatellite altimeter data. *J. Atmos. Oceanic Technol.,* **15,** 522–534.

Lindborg, E., 1999: Can the atmospheric kinetic energy spectra be explained by two-dimensional turbulence? *J. Fluid Mech.,* **388,** 259–288.

McWilliams, J. C., 1984: The emergence of isolated coherent vortices in turbulent flow. *J. Fluid Mech.,* **146,** 21–43.

Oort, A., L. Anderson, and J. Peixoto, 1994: Estimates of the energy cycle of the oceans. *J. Geophys. Res.,* **99,** 7665–7688.

Pedlosky, J., 1987: *Geophysical Fluid Dynamics.* 2d ed. Springer-Verlag, 710 pp.

Qiu, B., 1999: Seasonal eddy field modulation of the North Pacific subtropical countercurrent: TOPEX/Poseidon observations and theory. *J. Phys. Oceanogr.,* **29,** 2471–2486.

——, and S. Chen, 2004: Seasonal modulations in the eddy field of the South Pacific Ocean. *J. Phys. Oceanogr.,* **34,** 1515–1527.

Read, P. L., 2001: Transition to geostrophic turbulence in the laboratory, and as a paradigm in atmospheres and oceans. *Surv. Geophys.,* **22,** 265–317.

Rhines, P. B., 1975: Waves and turbulence on a β-plane. *J. Fluid Mech.,* **69,** 417–443.

——, 1977: The dynamics of unsteady currents. *The Sea,* E. Goldberg et al., Eds., Marine Modeling, Vol. 6, Wiley and Sons, 189–318.

Salmon, R., 1980: Baroclinic instability and geostrophic turbulence. *Geophys. Astrophys. Fluid Dyn.,* **15,** 167–211.

——, 1998: *Lectures on Geophysical Fluid Dynamics.* Oxford University Press, 378 pp.

Scott, R. B., 1999a: Geostrophic energetics and the small viscosity behaviour of an idealized ocean circulation model. Ph.D. dissertation, McGill University, 124 pp. [Available from Schulich Library of Science and Engineering, Macdonald Stewart Library Building, McGill University, 809 Sherbrooke St. West, Montreal, QC H3A 2K6, Canada.]

——, 1999b: Mechanical energy flux to the surface geostrophic flow using TOPEX/Poseidon data. *23d EGS General Assembly,* Nice, France, European Geophysical Society, 399–402.

——, 2001: Evolution of energy and enstrophy containing scales in decaying, two-dimensional turbulence with friction. *Phys. Fluids,* **13,** 2739–2742.

Smith, K. S., and G. K. Vallis, 2001: The scales and equilibration of midocean eddies: Freely evolving flow. *J. Phys. Oceanogr.,* **31,** 554–571.

Spall, M. A., 2000: Generation of strong mesoscale eddies by weak ocean gyres. *J. Mar. Res.,* **58,** 97–116.

Stammer, D., 1997: Global characteristics of ocean variability estimated from regional TOPEX/Poseidon altimeter measurements. *J. Phys. Oceanogr.,* **27,** 1743–1769.

Stewart, R. H., C. Shum, B. Tapley, and L. Ji, 1996: Statistics of geostrophic turbulence in the Southern Ocean from satellite altimetry and numerical models. *Physica D,* **98,** 599–613.

Theiss, J., 2004: Equatorward energy cascade, critical latitude, and the predominance of cyclonic vorticies in geostrophic turbulence. *J. Phys. Oceanogr.,* **34,** 1663–1678.

Treguier, A.-M., and B. L. Hua, 1988: Influence of bottom topography on stratified quasi-geostrophic turbulence in the ocean. *Geophys. Astrophys. Fluid Dyn.,* **43,** 265–305.

Vallis, G. K., 2005: *Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-Scale Circulation*. Cambridge University Press, in press. [Available online at www.princeton.edu/~gkv/aofd/.]

——, and M. Maltrud, 1993: Generation of mean flows and jets on a beta plane and over topography. *J. Phys. Oceanogr.,* **23,** 1346–1362.

Wang, W., and R. X. Huang, 2002: Gravitational potential energy sinks/sources in the oceans. *AGU Fall Meeting Abstracts,* San Francisco, CA, Amer. Geophys. Union, D251.

Wunsch, C., 1997: The vertical partition of oceanic horizontal kinetic energy and the spectrum of global variability. *J. Phys. Oceanogr.,* **27,** 1770–1794.

——, 1998: The work done by the wind on the oceanic general circulation. *J. Phys. Oceanogr.,* **28,** 2332–2340.

——, and R. Ferrari, 2004: Vertical mixing, energy, and the general circulation of the oceans. *Annu. Rev. Fluid Mech.,* **36,** 281–314.