Submesoscale processes and dynamics

Leif N. Thomas

Woods Hole Oceanographic Institution, Woods Hole, Massachusetts

Amit Tandon

Physics Department and SMAST, University of Massachusetts, Dartmouth, North Dartmouth, Massachusetts

Amala Mahadevan

Department of Earth Sciences, Boston University, Boston, Massachusetts

Abstract. Increased spatial resolution in recent observations and modeling has revealed a richness of structure and processes on lateral scales of a kilometer in the upper ocean. Processes at this scale, termed submesoscale, are distinguished by order one Rossby and Richardson numbers; their dynamics are distinct from those of the largely quasi-geostrophic mesoscale, as well as fully three-dimensional, small-scale, processes. Submesoscale processes make an important contribution to the vertical flux of mass, buoyancy, and tracers in the upper ocean. They flux potential vorticity through the mixed layer, enhance communication between the pycnocline and surface, and play a crucial role in changing the upper-ocean stratification and mixed-layer structure on a time scale of days.

In this review, we present a synthesis of upper-ocean submesoscale processes, arising in the presence of lateral buoyancy gradients. We describe their generation through frontogenesis, unforced instabilities, and forced motions due to buoyancy loss or down-front winds. Using the semi-geostrophic (SG) framework, we present physical arguments to help interpret several key aspects of submesoscale flows. These include the development of narrow elongated regions with $O(1)$ Rossby and Richardson numbers through frontogenesis, intense vertical velocities with a downward bias at these sites, and secondary circulations that redistribute buoyancy to stratify the mixed layer. We review some of the first parameterizations for submesoscale processes that attempt to capture their contribution to, firstly, vertical buoyancy fluxes and restratification by mixed layer instabilities and, secondly, the exchange of potential vorticity between the wind- and buoyancy-forced surface, mixed layer, and pycnocline. Submesoscale processes are emerging as vital for the transport of biogeochemical properties, for generating spatial heterogeneity that is critical for biogeochemical processes and mixing, and for the transfer of energy from the meso to small scales. Several studies are in progress to model, measure, analyze, understand, and parameterize these motions.

1. Introduction

The dynamics of $O(1)$ Rossby number, submesoscale processes at $O(1\text{km})$ in the upper ocean plays an important role in the vertical flux of mass, buoyancy, and tracers in the upper ocean. In addition, they are thought to be instrumental in transferring energy and properties from the largely adiabatic mesoscale ($\sim$10–100km) flow field, to a scale where mixing can occur. The breakdown of geostrophic balance in this regime leads to the development of secondary ageostrophic circulations with relatively large vertical velocities compared to those associated with the mesoscale. Hence they are vital to the exchange of properties between the surface mixed layer and thermocline.

The oceanic mesoscale flow field, characterized by a horizontal length scale of 10–100 km, has been studied extensively for its dynamics and its contribution to the lateral transport of heat, momentum and tracers via eddies. Similarly, three-dimensional processes at *small* length scales less than a kilometer (0.1–100 m) have been investigated for their contribution to mixing and energy dissipation. However,

submeso ($\sim$1 km) scales, that lie intermediate to meso- and small-scale three-dimensional motions, are less understood and only more recently brought to light through observational, modeling and analytical studies. The submesoscale, characterized by $\mathcal{O}(1)$ Rossby number dynamics, is not described appropriately by the traditional quasi-geostrophic theory that applies to mesoscales. It is not fully three-dimensional and nonhydrostatic, either, but is inevitably crucial to bridging the, meso and smaller, scales through processes and dynamics that we are just beginning to understand. The objective of this article is to review and synthesize the understanding of submesoscales put forth through recent diverse studies.

Our discussion will focus on the upper ocean, where submesoscale processes are particularly dominant due to the presence of lateral density gradients, vertical shear, weak stratification, a surface boundary that is conducive to frontogenesis. and a relatively small Rossby radius based on the mixed layer depth. This is not to say that submesoscales are solely an upper ocean phenomenon. In the ocean interior and abyss there exists a commonly observed class of vortical motions with length scales less than the mesoscale, termed submesoscale coherent vortices (SCVs). A thorough review on the observations and dynamics of SCVs can be found in *McWilliams* [1985].

The motivation to study submesoscales comes from several factors. Since the geometrical aspect (depth to length)

ratio, and Rossby number Ro, associated with meso and larger scale flow are $\ll 1$, and the Richardson number $Ri \gg 1$, the associated vertical velocities are $10^{-3} - 10^{-4}$ orders of magnitude smaller than the horizontal velocities, which are typically 0.1m-s^{-1}. However, localized submesoscale regions develop in which Ro and Ri are $O(1)$. At these sites, submesoscale dynamics generate vertical velocities of $O(10^{-3} ms^{-1})$ or $\sim$100m-day^{-1} that are typically an order of magnitude larger than those associated with the mesoscale, and play an important role in the vertical transport and mixing of properties in the upper ocean. The intense vertical velocities in the surface and mixed layer may communicate with the mesoscale up- and down-welling associated with the geostrophic frontal meander scale. Thus, submesoscale processes are instrumental in transferring properties and tracers, vertically, between the surface ocean and the interior. They enhance, for example, nutrient supply and the exchange of dissolved gases with the atmosphere.

A large part of the ocean's kinetic energy resides at meso and larger scales. At these scales, oceanic flow is largely two-dimensional and in a state of hydrostatic and geostrophic balance from which it is difficult to extract energy. A major conundrum [*McWilliams et al.*, 2001; *McWilliams*, 2003] therefore, is how energy is transferred from the meso scale to the small scale at which it can be dissipated through three-dimensional processes. The strong ageostrophic flow at submesoscales can extract energy from the balanced state and transfer it to smaller scales. The quasi two-dimensional mesoscale flow field is characterized by kinetic energy spectra with a slope of -3. Three-dimensional numerical simulations at progressively finer resolutions show that resolving submesoscale processes leads to flattening the kinetic energy spectra slope to -2 [*Capet et al.*, 2007a] and a transfer of energy to larger, as well as smaller scales [*Boccaletti et al.*, 2007].

Yet another factor associated with submesoscale instabilities, is the flux of potential vorticity to and from the surface to interior ocean and the change in stratification of the mixed layer. Submesoscale instabilities in the mixed layer are shown to hasten restratification and buoyancy transport several fold, as compared to what can be achieved through mesoscale baroclinic instability [*Fox-Kemper et al.*, 2007]. Hence their contribution to eddy transport can be significant. Present day global circulation models do not resolve submesocales; conceivably, this is the reason for the dearth of restratifying processes and mixed layers that are far too deep in the models. Hence parameterizing these processes is of interest to climate modeling. Similarly, the cumulative vertical flux of potential vorticity through submesoscale processes, can alter the potential vorticity budget of the thermocline and mixed layer [*Thomas*, 2005, 2007]. Submesoscales provide a pathway between the surface boundary layer where properties are changed by friction and diabatic processes, and the interior, which is largely adiabatic and conserves properties.

Resolving submesoscales within the mesoscale field has been a challenge for models and observations, but one that is being currently met through improvements in technology. Hydrographic surveys using towed vehicles (such as a SeaSoar) equipped with CTDs have revealed submesoscale features in the upper ocean around ocean fronts [e.g. *Pollard and Regier*, 1992; *Rudnick and Luyten*, 1996; *Lee et al.*, 2006b]. Recent high resolution observations centered around isopycnal front-tracking floats have captured rapid (occurring over a day) changes in the mixed layer stratification that cannot be ascribed to heating or cooling and, hence, are thought to result from processes involving the submesoscale [*Hosegood et al.*, 2006; *Lee et al.*, 2006a]. Further examples of submesoscale variability is seen in high resolution velocity fields from radar [*Shay et al.*, 2003] and sea-surface temperature fields from satellites [*Flament et al.*, 1985; *Capet et al.*, 2007b], as well as from biogeochemical sampling along ship

transects. We are at an exciting juncture because we are now able to achieve the required resolution in models and observations to capture this scale. The results from high resolution numerical modeling and analytical studies, several of which are discussed in this review, suggest that both forced and unforced instabilities drive submesoscale processes.

We begin Section 2 by defining submesoscales and describing phenomena with which they are associated. Further, we examine mechanisms that generate submesoscales in the upper ocean. In Section 3, we present a mathematical framework for understanding the secondary circulation associated with fronts, where submesoscales are found to occur. This is used to provide a dynamical explanation for several key features of submesoscale phenomena. In Section 4, we discuss the implications of submesoscales, which include mixed layer restratification, vertical transport and biogeochemical fluxes, and potential vorticity fluxes. Finally, we provide a discussion of outstanding questions and possible connections to other areas.

2. Phenomenology

2.1. What are submesoscales?

An active flow field in the upper ocean generates localized regions, typically along filaments or outcropping isopycnals, within which the relative vertical vorticity ζ equals or exceeds the planetary vorticity f. The dynamics within these regions differs from mesoscale dynamics characterized by small Rossby number ($Ro \ll 1$). We thus define submesoscale flows based on dynamics, as those where the Rossby number, defined by $Ro = \zeta/f$, is $O(1)$.

Since $\zeta \sim U/L$ and $Ro = U/fL$, the length scale for submesoscales, $L = U/f$. The horizontal velocity scale U can be defined further assuming thermal wind balance, $U \sim b_y H/f$, where buoyancy $b \equiv -g\rho/\rho_0$, y is the direction of lateral buoyancy gradient, ρ is the density anomaly from the mean density ρ_0, and H, the mixed layer depth, is the vertical extent over which velocity and lateral density gradients dominate. Thus, the submesoscale length scale can be expressed in terms of the lateral buoyancy gradient $M^2 \equiv b_y$, as $L = M^2 H/f^2$. Equivalently, since the vertical buoyancy gradient, or buoyancy frequency, $b_z \equiv N^2$ at the adjusted front scales with lateral buoyancy gradient as $N^2 = M^4/f^2$ [*Tandon and Garrett*, 1994], the submesoscale length scale can also be expressed in terms of the vertical buoyancy gradient as $L = NH/f$. Here, N and H are specific to the mixed layer and to the vicinity of the front; hence L is considerably smaller than the Rossby radius of deformation. For example, a mixed-layer depth $H = 50$m, $N = 10^{-3}$s^{-1} and $f = 0.5 \times 10^{-4}$s^{-1} gives a submesoscale $L = 1$km and $U = 0.1$ms^{-1}. Consequently, the bulk Richardson number Ri for the mixed layer, defined as $Ri = N^2 H^2/U^2$, is equal to $Ro^{-1/2}$ and is also $O(1)$ for submesocale flows.

The conditions on Ro and Ri imply that the aspect ratio of submesoscale flows $\Gamma = H/L$ scales as f/N. For typical oceanic conditions, $f/N \ll 1$, so that $\Gamma \ll 1$. Scaling the vertical momentum equation shows that the hydrostatic balance is accurate to $O(Ro^2\Gamma)$. Hence, even though $Ro = O(1)$ for submesoscale flows, $\Gamma \ll 1$ in the upper ocean, and these processes can, to a good approximation, be considered hydrostatic. Indeed, non-hydrostatic effects are difficult to detect in submesoscale model simulations at horizontal grid resolutions of 500 m [*Mahadevan*, 2006].

Next we describe mechanisms that are known to be active in generating submesoscales in the surface layer of the ocean. These include (1) frontogenesis, (2) unforced instabilities, such as the ageostrophic baroclinic instability [*Molemaker et al.*, 2005; *Boccaletti et al.*, 2007], and (3) forced
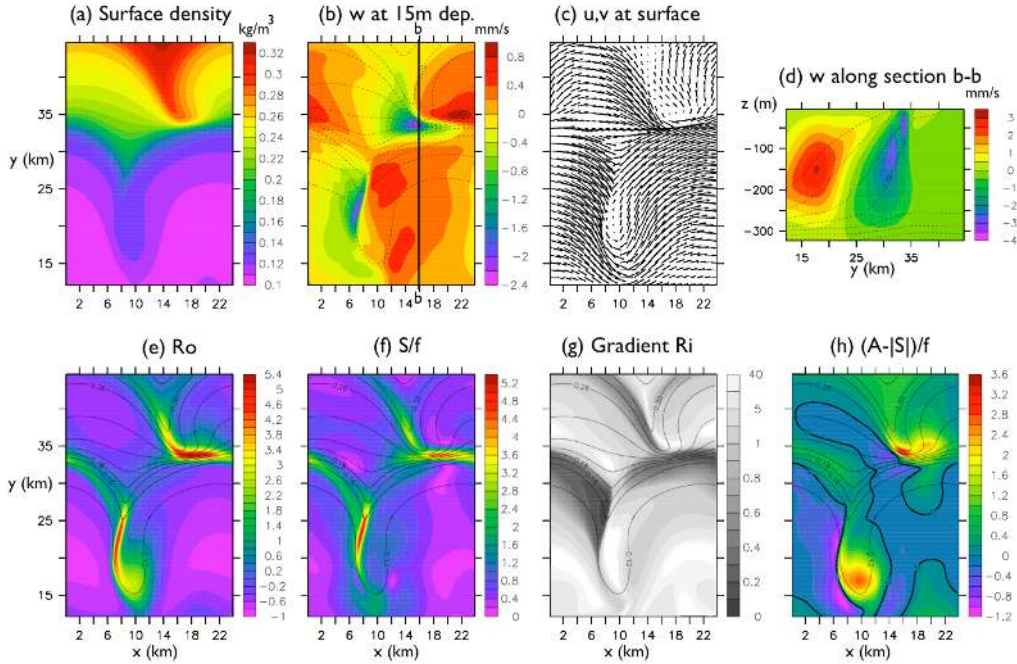
Figure 1. A region in the model domain where spontaneous frontogenesis has set up large shear and relative vorticity, strain rates and a strong ageostrophic secondary circulation. (a) Surface density (kg m^{-3}), (b) vertical velocity at 15 m depth (mm s^{-1}), (c) surface u, v velocities, (d) vertical section through the front at $x = 16$ km showing vertical velocity (red indicates upward, blue downward) and isopycnals (black), (e) $Ro = \zeta/f$, (f) S/f, (g) gradient Ri, and (h) $(A - |S|)/f$ with the zero contour shown as a dark black line. Light black contours indicate surface density.

motion, such as flows affected by buoyancy fluxes or friction at boundaries. We will use the results from a numerical model to individually demonstrate the above submesoscale mechanisms. Even though the submesoscale conditions are localized in space and time, the mesoscale flow field is crucial in generating them. In the ocean, it is likely that more than one submesoscale mechanism, and mesoscale dynamics, act in tandem to produce a complex submesoscale structure within the fabric of the mesoscale flow field.

2.2. Frontogenesis

Consider the flow field generated by a geostrophically balanced front in the upper mixed layer of the ocean, overlying a pycnocline. As the front becomes unstable and meanders, the nonlinear interaction of the the lateral velocity shear and buoyancy gradient, locally intensify the across-front buoyancy gradient. Strong frontogenetic action pinches outcropping isopycnals together, generating narrow regions in which the lateral shear and relative vorticity become very large, and the Ro and Ri become $\mathcal{O}(1)$. At these sites, the lateral strain rate $S \equiv ((u_x - v_y)^2 + (v_x + u_y)^2)^{1/2}$ is also large, and strong ageostrophic overturning circulation generates intense vertical velocities. In Fig. 1, we plot the density, horizontal and vertical velocities, strain rate, Ro, and Ri from a frontal region in a model simulation. The model was initialized with an across-front density variation of 0.27kg-m^{-3} across 20 km, (i.e., $|b_y| \approx 10^{-7}$s^{-2}) over a

deep mixed layer extending to 250m and allowed to evolve in an east-west periodic channel with solid southern and northern boundaries. Submesoscale frontogenesis is more easily seen when the mixed layer is deep, as the horizontal scale, which is dependent on H, is larger and more readily resolved in the numerical model. Here, the mixed layer is taken to be 250m deep in order to exaggerate frontogenesis. As the baroclinically unstable front meanders, the lateral buoyancy gradient is spontaneously, locally intensified in certain regions, as in Fig. 1, generating submesoscale conditions at sites approximately 5 km in width. This mechanism is ubiquitous to the upper ocean due to the presence of lateral buoyancy gradients and generates submesoscales when intensification can proceed without excessive frictional damping, or in a model with sufficient numerical resolution and minimal viscosity.

2.3. Unforced instabilities

The instabilities in the mixed layer regime where $Ro = \mathcal{O}(1)$ and and $Ri = Ro^{-1/2} = \mathcal{O}(1)$ are different from the geostrophic baroclinic mode in several respects. The ageostrophic baroclinic instability problem of a sheared rotating stratified flow in thermal wind balance with a constant horizontal buoyancy gradient was investigated using hydrostatic [*Stone*, 1966, 1970] and non-hydrostatic equations [*Stone*, 1971] for finite values of Ro. *Molemaker et al.*
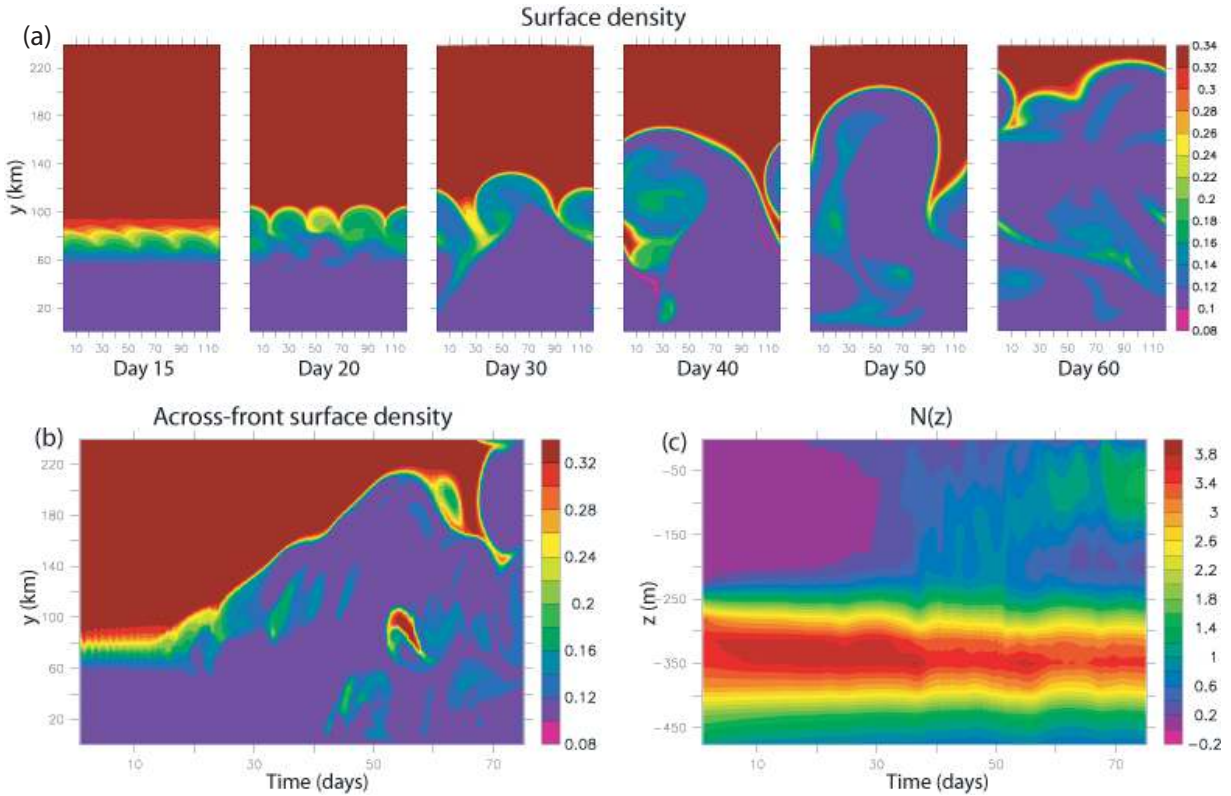
Figure 2. (a) Sequential figures of the surface density showing the evolution of the mixed layer instability over 45 days. (b) Hovmöller plot of the across-front surface density (c) Hovmöller plot showing the evolution of the horizontally averaged buoyancy frequency over time. The density is in units of kg m−3, while the buoyancy frequency is multiplied by 10^5 and is in s^{-1}.

[2005] extend these analyses by examining them in the context of loss of balance that leads to a forward energy cascade. Their instability analysis is applicable to the mixed layer regime, and shows two distinct instabilities. The largest growth rates arise for a geostrophic mode (large Ro Eady mode) which is mostly balanced and well captured by hydrostatic equations [*Stone*, 1966, 1970]. The non-geostrophic mode (called the anticyclonic ageostrophic instability or AAI in other flow situations) on the other hand, has a large non-hydrostatic unbalanced component, and generally has somewhat smaller growth rates than the geostrophic mode. (The growth rate is about one-quarter of the geostrophic mode for $Ro = 1$ and $N/f = 10$.) The non-geostrophic mode has the spatial structure of a critical layer, which suggests the generation of inertia-gravity waves. *Boccaletti et al.* [2007] investigate an instability problem similar to *Molemaker et al.* [2005] and *Stone* [1970], although with a reduced gravity bottom boundary condition. They show that baroclinic instability arises at the scale $L = NH/f$, where H is the mixed layer depth and N is the buoyancy frequency in the mixed layer, resulting in $L = 1$–10 km in the mixed layer. While *Boccaletti et al.* [2007] call this instability ageostrophic (since the instability arises in a high Ro and small Ri regime) this mixed layer instability (MLI) is the geostrophic mode of *Molemaker et al.* [2005] and is akin to the Eady mode for large Ro, small Ri. MLI, or the geostrophic mode, is mostly in balance, and its time scale,

given by $Ri^{1/2}/f$, is $O(1/f)$ since $Ri \sim 1$. This timescale is much shorter than that of mesoscale baroclinic instability. MLI hastens the slumping of fronts and hence leads to rapid restratification of the mixed layer, while the non-geostrophic mode or the ageostrophic anticyclonic instability (AAI) is more relevant to loss of balance and forward cascade of energy.

In Fig. 2, we show the evolution of the mixed layer (balanced) mode or MLI in a model simulation as it rapidly restratifies the mixed layer. The model is initialized with a north-south buoyancy gradient in the mixed layer, which is annihilated rapidly with depth in the upper thermocline. The east-west oriented front is seen to develop submesoscale meanders that merge and expand in length scale to the internal Rossby radius. As this occurs, lighter water overrides the denser side of front, and as the frontal outcrop is pushed back, the front slumps and the mixed layer restratifies. The simulation replicates the phenomena in *Boccaletti et al.* [2007]. Unless forced (by a loss of buoyancy at the surface, for example), the submesoscale wiggles at the front, which grow at $\mathcal{O}(f^{-1})$ time scales between days 5 and 10 of the simulation, transcend to mesoscale meanders within a couple of days by the up-scale transfer of energy. The evolution of meanders then progresses at mesoscale time scales, but submesoscale processes continue to dominate the secondary circulation and the vertical fluxes at the front, leading to restratification. The values of the local Ro and

gradient Ri during the evolution of MLI suggest that the instability is largely driven by the ageostrophic component of the flow, even though it is largely in gradient wind balance.

A departure from balanced dynamics, in which submesoscales most likely play an important role, facilitates a forward cascade of energy [*McWilliams et al.*, 2001]. The nongeostrophic, unbalanced mode in *Molemaker et al.* [2005], requires that the difference between the absolute vorticity, $A = f + v_x - u_y$, and the magnitude of the strain rate, i.e. $A - |S|$, changes sign. The frontal simulation in Figure 1 shows that this criteria is well satisfied in regions where active frontogenesis and MLI occur. However, diagnosis of the unbalanced mode remains a challenge because a method is required to decompose the flow into its balanced and unbalanced components. The imbalance itself has to be characterized as either a departure from a hydrostatic and cyclostrophically adjusted state (pressure imbalance), or as a departure from a state that is fully diagnosable from the Ertel Potential Vorticity (PV imbalance). These different imbalance characterizations leads to different decompositions. PV based decompositions for non-hydrostatic equations have been suggested by *McKiver and Dritschel* [2006a]; *Viúdez and Dritschel* [2003], wherein the PV and the two horizontal components of the ageostrophic vorticity are used as prognostic variables to iteratively find the state with minimum imbalance. Since the basic state of *Molemaker et al.* [2005] has constant PV, PV or extended PV based decompositions are impractical for their instability problem. Instead, they attempt an orthogonal decomposition into balanced and unbalanced modes by defining an energy-based norm, which minimizes a cost function to extract the balanced flow. In submesoscale resolving numerical simulations, *Mahadevan and Tandon* [2006] used the quasi-geostrophic omega equation to diagnose the balanced vertical velocities and compared these to the vertical velocities from the three-dimensional model. Likewise, one may diagnose the three-dimensional ageostrophic stream function (described in section 3.1.3) to represent the balanced flow, and compare it with the ASC circulation from the model. *Capet et al.* [2007a] calculate normalized departures from gradient wind balance in simulations of the California Current system. By successively increasing the horizontal model resolution, they see a greater departure from gradient wind balance, indicative of unbalanced instabilities.

2.4. Forced motions

Applying destabilizing atmospheric forcing to baroclinic mesoscale features can lead to the generation of submesoscale motions. Destabilizing atmospheric forcing is defined as that which tends to reduce the stratification, Richardson number, and, more generally, the PV of the upper ocean. Such forcing can arise from heat or salt fluxes that extract buoyancy from the ocean, or by down-front wind-stress.

2.4.1. Buoyancy loss

Cooling a mesoscale flow uniformly or differentially cooling a laterally homogeneous layer can trigger submesoscale instabilities. The case of differential cooling was investigated by *Haine and Marshall* [1998] using nonhydrostatic numerical simulations run in two- and three-dimensional configurations. In both configurations, the differential cooling creates a lateral buoyancy gradient that forms a baroclinic jet through geostrophic adjustment. Due to the destabilizing forcing, the PV of the baroclinic jet is drawn to negative values, which in the 2D simulations triggers symmetric instability of submesoscale width ($\mathcal{O}(1 \text{ km})$) that mixes the PV to create a mixed layer with nearly zero PV, nonzero stratification, and order one Richardson number. In the 3D configuration, the jet goes unstable to baroclinic waves that have similar properties to the geostrophic mode described

in *Molemaker et al.* [2005] and the MLI of *Boccaletti et al.* [2007], i.e. with horizontal scales ~ 5 km and timescales of development of $\mathcal{O}(1 \text{ day})$. The baroclinic instability of *Haine and Marshall* [1998] rapidly restratifies the mixed layer and grows in horizontal scale, as seen in the numerical experiments of *Boccaletti et al.* [2007] and the solution presented in Fig. 2.

The case of uniform cooling of a mesoscale flow was investigated using nonhydrostatic numerical simulations by *Legg et al.* [1998] and *Legg and McWilliams* [2001] for a flow consisting of an initial mesoscale eddy field, and by *Yoshikawa et al.* [2001] for an ocean front. When a single cyclonic eddy is cooled in *Legg et al.* [1998], erosion of stratification reduces the Rossby radius of deformation to a value smaller than the eddy's radius and allows for the formation of submesoscale baroclinic waves on the edges of the eddy. Numerical experiments in which a mixture of baroclinic cyclonic and anticyclonic eddies are cooled were analyzed in *Legg and McWilliams* [2001]. In these experiments, the barotropic eddy kinetic energy is observed to increase at a faster rate as compared to the case without cooling, suggesting that cooling, by triggering submesoscale baroclinic instability, provides an additional means of converting available potential energy at mesoscales to barotropic eddy kinetic energy.

Yoshikawa et al. [2001] detailed the way in which destabilizing buoyancy fluxes at an ocean front can enhance the frontal vertical circulation. Comparing forced and unforced model runs, they showed that lateral strain associated with baroclinic instabilities was stronger for fronts forced by cooling and resulted in more intense frontogenesis. Using the semi-geostrophic omega equation, *Yoshikawa et al.* [2001] demonstrated that the amplified frontogenetic strain combined with the cooling-induced low PV in the mixed layer, causes an enhancement in the strength of the frontal vertical circulation. We will elaborate on the dynamics of this important result in section 3.1.1.

2.4.2. Down-front wind-stress

Down-front winds, i.e. winds blowing in the direction of the surface frontal jet, reduce the stratification and Richardson number in the surface boundary layer, and thus provide favorable conditions for submesoscale phenomena [*Thomas and Lee*, 2005; *Thomas*, 2005]. For this wind orientation, Ekman flow advects denser water over light, destabilizing the water column and triggering convective mixing. The strength of this mixing is set by the magnitude of the surface lateral buoyancy gradient and the Ekman transport, the latter being a function of the vorticity of the flow. This is because flows with $\mathcal{O}(1)$ Rossby number exhibit nonlinear Ekman dynamics in which the advection of momentum plays a role in the Ekman balance. This causes the Ekman transport to vary inversely with the absolute rather than planetary vorticity [*Stern*, 1965; *Niiler*, 1969; *Thomas and Rhines*, 2002]. One of the more striking results of nonlinear Ekman dynamics is that a spatially uniform wind-stress will induce Ekman pumping/suction if it is blowing over a current with lateral variations in its vorticity field. For example, a down-front wind blowing over a frontal jet will drive surface Ekman transport from the dense to light side of the front. The transport is weaker (stronger) on the cyclonic (anticyclonic) side of the jet and hence induces a secondary circulation with convergence and downwelling on the dense side of the front and divergence and upwelling at the center of the jet. Localized mixing caused by Ekman driven convection at the front drives an additional secondary circulation. The interplay of the two secondary circulations, vorticity, and lateral buoyancy gradient can result in frontal intensification and the formation of multiple fronts of submesoscale separation at a baroclinic zone forced by down-front winds [*Thomas and Lee*, 2005].

Convective mixing induced by down-front winds leads to a reduction in the PV. This modification of the PV is due

to an upward frictional PV flux that extracts PV from the ocean. The numerical simulations of *Thomas* [2005] demonstrate how this frictional PV flux is transmitted through the oceanic surface layer by submesoscale secondary circulations that downwell low PV from the surface while upwelling high PV from the pycnocline, yielding a vertical eddy PV flux that scales with the surface frictional PV flux. Although these numerical experiments were two-dimensional, fully three-dimensional simulations of wind-forced fronts have also shown such a vertical exchange of PV by submesoscale frontal circulations [*Capet et al.*, 2007b; *Mahadevan and Tandon*, 2006; *Thomas*, 2007]. In *Thomas* [2007], the subduction of the low PV surface water resulted in the formation of a submesoscale coherent vortex or intrathermocline eddy. The vortex was shown to exert an along-isopycal eddy PV flux that scales with the wind-driven frictional PV flux at the surface. The eddy PV flux drove an eddy-induced transport or bolus velocity down the outcropping isopycal which had an effect on the large-scale mean flow. The correspondence of the eddy and frictional PV fluxes in both the two and three-dimesional simulations suggests that a possible approach for parameterizing wind-forced submesoscale phenomena is to use a parameterization scheme based on PV fluxes, a subject that will be discussed further in section 4.2.

3. Submesoscale Dynamics

The time scale of variability for submesoscale flows is often not very distinct from the period of near-inertial internal gravity waves, but unlike internal gravity waves, submesoscale flows are to a large extent balanced. This implies that submesoscale dynamics are determined by a single scalar field from which all other variables (density, horizontal and vertical velocity, etc.) can be determined using an invertibility principle [*McKiver and Dritschel*, 2006b]. In the most familiar approximate forms of the primitive equations governing balanced flows, there are two choices for the controlling scalar field: the horizontal streamfunction (e.g. the balanced equations of *Gent and McWilliams* [1983]) and the potential vorticity (e.g. the quasi- and semi-geostrophic models of *Charney* [1948] and *Eliassen* [1948], respectively). Of all the models listed above, the quasi-geostrophic (QG) model places the most severe restriction on the Rossby and Richardson numbers. QG theory assumes that $Ro \ll 1$ and $1/Ri \ll 1$; consequently, it is not designed to accurately describe the dynamics of submesoscale phenomena. A thorough study of the advantages and disadvantages of each model listed above can be found in *McWilliams and Gent* [1980]. In general, for flows with significant curvature, such as submesoscale coherent vortices, the balance equations (BE) are the most accurate [*Gent et al.*, 1994]. while for relatively straight flows, the semi-geostrophic (SG) equations are accurate and provide insights into the dynamics of the intense fronts and vertical circulations typical of submesoscale features in a relatively simple manner. The BE have been described in the previous chapter by McWilliams, therefore, in the interest of brevity and to avoid redundancy, we will describe some of the key features of submesoscale phenomena, namely frontogenesis, strong vertical circulation, restratification, forward cascade through frontogenesis, and nonlinear Ekman effects using the dynamical framework of the semi-geostrophic equations.

3.1. Semi-geostrophic dynamics

In SG theory, the flow is decomposed into geostrophic and ageostrophic components $\mathbf{u} = \mathbf{u}_g + \mathbf{u}_{ag}$, where the geostrophic velocity is $\mathbf{u}_g \equiv \hat{\mathbf{k}} \times \nabla_h p / f$. and p is the pres-

sure. The SG equations are

$$\frac{D\mathbf{u}_g}{Dt} = -f\hat{\mathbf{k}} \times \mathbf{u}_{ag} \tag{1}$$

$$0 = -\frac{1}{\rho_o}\frac{\partial p}{\partial z} + b \tag{2}$$

$$\frac{Db}{Dt} = 0 \tag{3}$$

$$\nabla \cdot \mathbf{u}_{ag} = 0, \tag{4}$$

with $D/Dt = \partial/\partial t + (\mathbf{u}_g + \mathbf{u}_{ag}) \cdot \nabla$, and are valid if $D^2 u/Dt^2 \ll f^2 u$ and $D^2 v/Dt^2 \ll f^2 v$, i.e. if the Lagrangian timescale of variability of the flow is much longer than an inertial period [*Hoskins*, 1975].

3.1.1. Two-dimensional vertical circulation

Consider a front in the y-z plane, i.e. $b_y \neq 0$ and $b_x = 0$, where the along-front velocity u is purely geostrophic, i.e., $u = u_g$, and in thermal wind balance; $fu_{gz} = -b_y$. The two dimensional ageostrophic circulation can be described by an across-front overturning stream function ψ, where $(v_{ag}, w) = (\psi_z, -\psi_y)$. As first derived by *Eliassen* [1948]; *Sawyer* [1956], a single equation for ψ can be constructed by combining the y derivative of the buoyancy equation (3), with the z derivative of the zonal component of (1), yielding

$$F_2^2\frac{\partial^2\psi}{\partial z^2} + 2S_2^2\frac{\partial^2\psi}{\partial z\partial y} + N^2\frac{\partial^2\psi}{\partial y^2} = -2Q_2^g, \tag{5}$$

where $N^2 = b_z$, $S_2^2 = -b_y = fu_{gz}$, $F_2^2 = f(f - u_{gy})$, and Q_2^g is the y-component of the Q-vector

$$\mathbf{Q}^g = (Q_1^g, Q_2^g) = \left(-\frac{\partial\mathbf{u}_g}{\partial x}\cdot\nabla b, -\frac{\partial\mathbf{u}_g}{\partial y}\cdot\nabla b\right) \tag{6}$$

introduced by *Hoskins et al.* [1978]. A geostrophic flow with a nonzero Q-vector will modify the magnitude of the horizontal buoyancy gradient following the equation

$$\frac{D}{Dt}|\nabla_h b|^2 = \mathbf{Q}^g \cdot \nabla_h b \tag{7}$$

and will consequently disrupt the thermal wind balance of the flow. To restore geostrophy an ageostrophic secondary circulation is required, and its solution is governed by (5). Although (5) is formally valid only for two-dimensional ageostrophic motions, it is useful for diagnosing the importance of various mechanisms in generating an ASC and large vertical velocities in the presence of lateral buoyancy gradients. A three-dimensional version of this equation, with diabatic and frictional effects, is thus presented and discussed in Section 3.1.3.

The solution to (5) can be found using the method of Green's functions. The Green's function for ψ satisfies the following equation

$$F_2^2\frac{\partial^2 G}{\partial z^2} + 2S_2^2\frac{\partial^2 G}{\partial z\partial y} + N^2\frac{\partial^2 G}{\partial y^2} = \delta(y - \mathcal{Y}, z - \mathcal{Z}), \tag{8}$$

which, for any Q−vector distribution yields the ageostrophic circulation: $\psi = -2\iint G(y - \mathcal{Y}, z - \mathcal{Z})Q_2^g(\mathcal{Y}, \mathcal{Z})\mathrm{d}\mathcal{Y}\mathrm{d}\mathcal{Z} + \psi_h$, where ψ_h is a homogeneous solution to (4) that ensures that ψ satisfies the boundary conditions. The solution to (8) for constant coefficients is

$$G = \frac{1}{4\pi\sqrt{fq_{2D}}}\log|\mathrm{Arg}| \tag{9}$$

where

$$\mathrm{Arg} = \frac{[(y - \mathcal{Y}) - (z - \mathcal{Z})S_2^2/F_2^2]^2}{L_{SG}^2} + \frac{(z - \mathcal{Z})^2}{H^2}$$
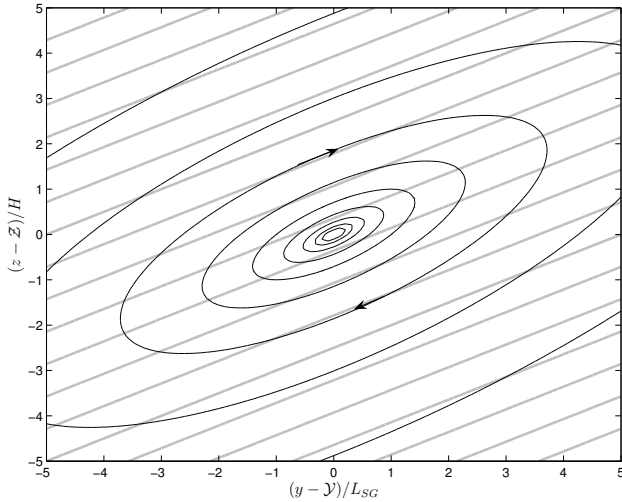
Figure 3. Ageostrophic secondary circulation G driven by a negative point source Q-vector, $Q_2^g < 0$, at $y = \mathcal{Y}, z = \mathcal{Z}$. Isopycnals (gray contours) slant upward to the north due to a southward buoyancy gradient. For this frontogenetic forcing, $\mathbf{Q}^g \cdot \nabla_h b > 0$, the circulation is thermally direct and tends to restratify the fluid.

and

$$L_{SG} = H \frac{\sqrt{f q_{2D}}}{F_2^2} \quad (10)$$

is the semi-geostrophic Rossby radius of deformation, H is a characteristic vertical lengthscale of the flow, and

$$q_{2D} = \frac{1}{f}(F_2^2 N^2 - S_2^4) = fN^2 \left[1 + Ro_{2D} - \frac{1}{Ri_{2D}}\right] \quad (11)$$

is the PV of the geostrophic flow if it were purely zonal and two-dimensional, i.e. $Ro_{2D} = -u_{gy}/f$ and $Ri_{2D} = N^2/(u_{gz})^2$ [*Eliassen*, 1951; *Hakim and Keyser*, 2001] ¿From the dependence of (9) on the PV, it can be seen that a solution does not exist for $fq_{2D} < 0$, indicating that the conditions under which a unique solution for (5) can be found is $fq_{2D} > 0$. G is plotted in figure 3 for a buoyancy field that decreases in the $y-$direction ($S_2^2 > 0$). Streamlines take the shape of tilted ellipses oriented at an angle $\theta = 0.5 \tan^{-1}[2S_2^2/(N^2 - F_2^2)]$. For typical conditions, $F_2^2 \ll |S_2^2| \ll N^2$, the ellipses are oriented parallel to isopycnals, $\tan\theta \approx S_2^2/N^2$, and the secondary circulations form slantwise motions along isopycnals. The sense of the secondary circulation of G is thermally direct, tending to flatten isopycnals. The Green's function is driven by a right hand side forcing of (5) corresponding to a negative point-source $Q-$vector, so that $\mathbf{Q}^g \cdot \nabla_h b > 0$, indicating that this secondary circulation is associated with frontogenetic forcing, i.e. $D|\nabla_h b|^2/Dt > 0$. If either the buoyancy gradient or the direction of the $Q-$vector were reversed, a thermally indirect circulation would be induced that tends to steepen isopycnals.

The vertical velocity associated with (9) is $-\partial G/\partial y$, or

$$w = -\frac{F_2^4}{2\pi(fq_{2D})^{3/2}H^2} \frac{(y - \mathcal{Y})}{\text{Arg}} \quad (12)$$

and has a magnitude that varies greatly with the PV and vertical vorticity of the geostrophic flow. This solution suggests that for a given geostrophic forcing, in regions of low

PV (where the fluid's Richardson number, stratification, or absolute vorticity is small), the vertical circulation is strong. This amplification of the vertical circulation in regions of low PV was also noted by *Yoshikawa et al.* [2001], as mentioned in section 2.4.2. The solution (12) also suggests w will be stronger in regions of cyclonic versus anticyclonic vorticity when both regions have the same PV. The enhancement of the vertical velocity in areas of low PV and high vertical vorticity is a consequence of the reduction in the characteristic horizontal lengthscale of the circulation (10), which is an effect that is more pronounced for flows with high Rossby numbers and low Richardson numbers such as those associated with submesoscale phenomena.

3.1.2. Horizontal deformation and frontal collapse

A simple yet dynamically insightful model for the formation of intense fronts in the submesoscale regime ($Ro \sim Ri \sim 1$) is the horizontal deformation model of *Hoskins and Bretherton* [1972]. In this model, the geostrophic flow and pressure field are decomposed into two parts

$$\begin{aligned} u_g &= \alpha x + u_g'(y, z, t) & v_g &= -\alpha y \\ p &= -\rho_o \alpha(fxy + \alpha x^2/2) + p'(y, z, t) \end{aligned} \quad (13)$$

associated with a deformation field with confluence $\alpha =$ constant > 0, and a time evolving, two-dimensional zonal flow u_g', respectively. In the context of submesoscale phenomena, the deformation field could be associated with a mesoscale eddy field, while u_g' could be interpreted as an evolving submesoscale frontal feature. The equations are solved by introducing a geostrophic coordinate $Y \equiv y - u_g'/f$. Substituting (13) into (1) yields $DY/Dt = -\alpha Y$, showing how following a fluid parcel, the geostrophic coordinate is compressed in time: $Y = Y_o \exp(-\alpha t)$ (Y_o is the geostrophic coordinate of the fluid parcel at $t = 0$). Transforming the geostrophic and hydrostatic relations from Cartesian (y, z) to geostrophic coordinates (Y, Z) (where $Z = z$), the geostrophic flow and buoyancy field can be
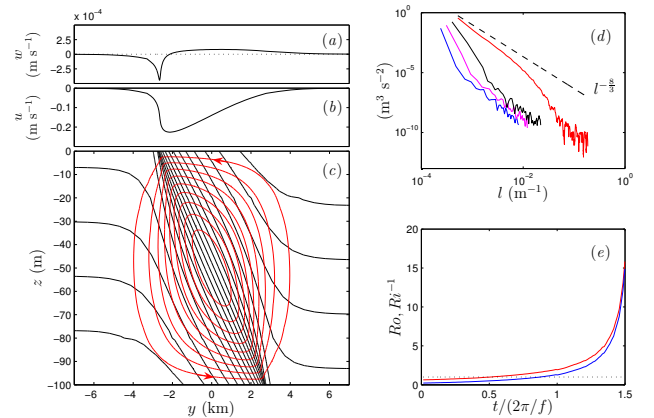

Figure 4. Solution of the horizontal deformation model for an initial buoyancy field $\mathcal{B} = \Delta b/2 \tanh(y/L)$: (a) w and (b) u at $z = -2.5$m, and (c) the overturning streamfunction ψ (red contours) and buoyancy b (black contours) at $t = 1.5$ inertial periods when frontal collapse at the boundaries is about to occur. (d) The horizontal spectrum of u at $z = 0$ for $t = 0.26$ (blue), 0.67 (magenta), 1.09 (black), and 1.50 (red) inertial periods. The dashed line denotes a -8/3 spectral slope. (e) Evolution of the maximum in the Rossby (blue) and inverse Richardson (red) number of the flow at $z = 0$. A dotted line at $Ro = Ri^{-1} = 1$ is plotted for reference.

expressed in terms of a horizontal streamfunction: $u'_g = -(1/f)\partial\Phi/\partial Y$, $b = \partial\Phi/\partial Z$. The streamfunction Φ can be related to the PV $q_g = (f - \partial u'_g/\partial y)N^2 - (\partial b/\partial y)^2/f$ through the following equation

$$f\frac{\partial^2\Phi}{\partial Z^2} + \frac{q_g}{f^2}\frac{\partial^2\Phi}{\partial Y^2} = q_g. \qquad (14)$$

When $fq_g > 0$, (14) is elliptic and given appropriate boundary conditions can be inverted to yield a unique solution for u'_g and b. If $fq_g < 0$, (14) is hyperbolic and a unique inversion cannot be performed, which is the indication of a break down in balance. Notice that this is the same solvability condition for (5), indicating that when $fq_g < 0$, a unique solution for the overturning streamfunction cannot be determined as well.

When the PV is constant, solutions to (14) are easy to obtain. Consider the configuration of a layer of fluid bounded at $z = 0, -H$ by rigid walls and initialized with a uniform stratification $N_o^2 = q_g/f$ and a weak laterally varying buoyancy field $b(t = 0, y, z) = \mathcal{B}(y)$. On the horizontal boundaries, the geostrophic coordinate of a fluid parcel that is initially at $y = Y_o$, is $Y = Y_o\exp(-\alpha t)$. The buoyancy of the particle is conserved and is equal to $\mathcal{B}(Y_o) = \mathcal{B}(Y\exp(\alpha t))$. Therefore, at $Z = 0, -H$, the buoyancy, and hence $\partial\Phi/\partial z$, is a known function of Y and (14) can be solved given lateral boundary conditions. The solution for an initially weak front that varies as $\mathcal{B} = \Delta b/2\tanh(y/L)$, with $\Delta b = 0.0016$ ms^{-2} and $L = 4$ km, in a background stratification of $N_o^2 = 4.2 \times 10^{-6}$ s^{-2} and forced by a confluence $\alpha = 0.1f$ ($f = 1 \times 10^{-4}$ s^{-1}) is shown in Figure 4. Since the geostrophic forcing is frontogenetic, a thermally direct secondary circulation with an upward vertical buoyancy flux $\overline{wb} > 0$ (the overline denotes a lateral average) is induced. The ageostrophic flow is itself frontogenetic and leads to the formation of a frontal discontinuity in the zonal velocity and buoyancy field that first appears at the horizontal boundaries in a finite amount of time. During frontogenesis, asymmetric vertical vorticity and vertical velocity distributions are generated. Near the upper (lower) boundaries, regions of intense cyclonic vorticity and downward (upward) vertical velocities coincide, consistent with the argument presented in section 3.1.3 that the characteristic lateral lengthscale of the secondary circulation (10) is compressed in areas of high absolute vorticity. The formation of the discontinuity in the zonal velocity field causes the horizontal wavenumber spectrum of u to flatten in time, asymptoting to a $l^{-8/3}$ spectrum as the time of frontal collapse approaches [*Andrews and Hoskins*, 1978]. As the frontal discontinuity forms, the maximum in the Rossby and inverse Richardson numbers grow rapidly, taking on values greater than one. The evolution of Ro and Ri^{-1} closely track one another. This is a consequence of conservation of PV. For large Rossby numbers, $Ro = -\partial u'_g/f \gg 1$, the PV is approximately $q_g \approx fN^2(Ro - Ri^{-1})$, so that for the PV to be conserved, $Ro \propto Ri^{-1}$. Many of these features of the deformation model: asymmetric vorticity and w distributions, restratifying secondary circulations, flat kinetic energy spectra, and coincident regions of high Rossby and low Richardson numbers are typical of submesoscale flows, and suggests that the simple physics encompassed in this model is quite relevant to the dynamics of submesoscale phenomena.

3.1.3. Three-dimensional overturning circulation with frictional and diabatic effects

The semi-geostrophic equations can be used to study fully three-dimensional flows as well, provided that the cross-stream length scale is much less than the radius of curvature of the flow [*Gent et al.*, 1994]. In particular, semi-geostrophic theory is useful for understanding the three-dimensional ageostrophic secondary circulation, which can concisely be described in terms of a vector streamfunction

(ϕ, ψ) such that

$$u_{ag} = \frac{\partial\phi}{\partial z}, \quad v_{ag} = \frac{\partial\psi}{\partial z}, \quad w = -\left(\frac{\partial\phi}{\partial x} + \frac{\partial\psi}{\partial y}\right) \qquad (15)$$

[*Hoskins and Draghici*, 1977]. If friction or diabatic effects are included, then the governing equation for the ageostrophic circulation is

$$\begin{pmatrix} \mathcal{L}_{11} & \mathcal{L}_{12} \\ \mathcal{L}_{21} & \mathcal{L}_{22} \end{pmatrix}\begin{pmatrix} \phi \\ \psi \end{pmatrix} = \underbrace{-2\begin{pmatrix} Q_1^g \\ Q_2^g \end{pmatrix}}_{I} + \underbrace{f\frac{\partial}{\partial z}\begin{pmatrix} Y \\ -X \end{pmatrix}}_{II} - \underbrace{\begin{pmatrix} \frac{\partial\mathcal{D}}{\partial x} \\ \frac{\partial\mathcal{D}}{\partial y} \end{pmatrix}}_{III},$$
$$(16)$$

where (X, Y) are the components of the horizontal frictional force and $\mathcal{D}$ represents diabatic processes,

$$\mathbf{F} = (X, Y); \quad \mathcal{D} \equiv \frac{Db}{Dt}, \qquad (17)$$

e.g., if lateral mixing of momentum and buoyancy are negligible, $(X, Y) = \partial_z(\tau^x, \tau^y)/\rho_o$ and $\mathcal{D} = -\partial_z F^B$, where (τ^x, τ^y) and F^B are the turbulent stress and vertical buoyancy flux, respectively. The differential operator on the left hand side of (4) is

$$\mathcal{L}_{11} = F_1^2\partial_{zz} - 2S_1^2\partial_{xz} + N^2\partial_{xx}$$
$$\mathcal{L}_{12} = -C^2\partial_{zz} + 2S_2^2\partial_{xz} + N^2\partial_{xy}$$
$$\mathcal{L}_{21} = -C^2\partial_{zz} - 2S_1^2\partial_{yz} + N^2\partial_{xy}$$
$$\mathcal{L}_{22} = F_2^2\partial_{zz} + 2S_2^2\partial_{yz} + N^2\partial_{yy},$$

where, N^2, F_2^2, and S_2^2 are the same as in (5), and the coefficients F_1^2, S_1^2, and C^2 are functions of the stratification, shear, and confluence/diffluence of the geostrophic flow as follows

$$F_1^2 = f\left(f + \frac{\partial v_g}{\partial x}\right), \ S_1^2 = f\frac{\partial v_g}{\partial z}, \ C^2 = -f\frac{\partial v_g}{\partial y} = f\frac{\partial u_g}{\partial x}.$$

As shown by *Hoskins and Draghici* [1977], equation (4) forced solely by the "geostrophic forcing" term I can be converted to the semi-geostrophic omega-equation by taking $-\partial/\partial x$ of the top row of (4) and adding it to $-\partial/\partial y$ of the second row of (4).

Similar to the "geostrophic forcing", vertically varying frictional forces or laterally varying buoyancy sources/sinks (terms II and III respectively), will disrupt the thermal wind balance and hence drive ageostrophic circulations [*Eliassen*, 1951] Also, it is worth noticing that the equations governing nonlinear Ekman dynamics are encompassed in (4) when term II is balanced on the left hand side by terms with coefficients F_1^2 and F_2^2. For example, if the geostrophic flow were purely zonal and forced by a zonal wind stress τ_w^x, then the meridional Ekman transport $M_e^y = \int_{-\infty}^0 vdz$ can be found by vertically integrating term II twice and dividing by F_2^2 (assuming that u_g is approximately constant through the thickness of the Ekman layer) yielding

$$M_e^y = -\frac{\tau_w^x}{\rho_o(f - \partial u_g/\partial y|_{z=0})}, \qquad (18)$$

which is the solution of *Stern* [1965] and *Niiler* [1969] alluded to in section 2.4.2. Term III is important at wind-forced fronts where Ekman flow can advect dense water over light, generating convection and turbulent mixing of buoyancy at the front. The strength of this mixing is set by a wind-driven buoyancy flux

$$F_{wind}^B = \mathbf{M}_e \cdot \nabla_h b|_{z=0}, \qquad (19)$$

which is a function of both the lateral buoyancy gradient and, as a result of (18), the vertical vorticity of the front. This mixing drives (through term III) frontogenetic ageostrophic secondary circulations [*Thomas and Lee*, 2005].

For submesoscale flows in the upper ocean, all three forcing terms of (4) are likely to play a role in driving an overturning circulation. In inferring the vertical velocities from high resolution hydrographic and velocity observations in the upper ocean, is is often assumed that term I dominates the dynamics (i.e. the vertical velocity follows the omega-equation) and that the ASC is well described by the quasi-geostrophic version of (4), i.e. with $F_1^2 \rightarrow f^2$, $F_2^2 \rightarrow f^2$, $C^2 \rightarrow 0$, $N^2 \rightarrow N^2$, $S_1^2 \rightarrow 0$, and $S_2^2 \rightarrow 0$ [e.g. *Tintore et al.*, 1991; *Pollard and Regier*, 1992; *Rudnick*, 1996; *Pinot et al.*, 1996; *Shearman et al.*, 1999; *Viúdez and Dritschel*, 2004; *Vélez-Belchí et al.*, 2005]. The validity of these assumptions comes into question for submesoscale flows, especially if the flows are exposed to atmospheric forcing. The QG omega-equation does not work well for diagnosing vertical velocities when tested with a high-resolution numerical simulation that generates $\mathcal{O}(1)Ro$, as it misdiagnoses the position and sign of the most intense vertical velocities near the surface [*Mahadevan and Tandon*, 2006]. To account for effects arising from high Rossby and low Richardson numbers, friction, and diabatic processes, methods for inferring the overturning circulation that are more appropriate for submesoscale phenomena have been derived using variants of (4) [e.g. *Nagai et al.*, 2006; *Thomas and Lee*, 2007].

To yield a unique solution for the ASC for given boundary conditions, (4) must be elliptic. Equation (4) switches from being elliptic to hyperbolic, and the semi-geostrophic limit of balance is crossed when the three-dimensional semi-geostrophic PV

$$ q_{3D} = \left[\left(F_1^2 F_2^2 - C^4 \right) N^2 + 2C^2 S_1^2 S_2^2 - F_2^2 S_1^4 - F_1^2 S_2^4 \right] / f^3 \tag{20} $$

multiplied by the Coriolis parameter becomes negative, i.e. $fq_{3D} < 0$. In three-dimensional numerical simulations [e.g. *Mahadevan and Tandon*, 2006], these regions where the PV changes sign tend to coincide with the sites of most intense submesoscale activity. *Hoskins* [1975] demonstrated that the semi-geostrophic equations without friction and diabatic terms, i.e. (1)-(4), can be manipulated to form a conservation law for q_{3D}, $Dq_{3D}/Dt = 0$. Consequently, in an adiabatic, inviscid geostrophic flow that initially has a PV field with $fq_{3D} > 0$ everywhere, the semi-geostrophic equations can be integrated forward for all times since the semi-geostrophic limit of balance can never be crossed. When friction or diabatic effects are present, this is not always the case, since the PV can be driven to negative values under certain conditions, as described in the next section.

3.2. Frictional or diabatic modification of the potential vorticity

Changes in the PV arise from convergences/divergences of PV fluxes

$$ \frac{\partial q}{\partial t} = -\nabla \cdot \left(\mathbf{u}q + \mathbf{J}^{na} \right), \tag{21} $$

where $q = (f\hat{k} + \nabla \times \mathbf{u}) \cdot \nabla b$ is the full, unapproximated Ertel PV and

$$ \mathbf{J}^{na} = \nabla b \times \mathbf{F} - \mathcal{D}(f\hat{k} + \nabla \times \mathbf{u}) \tag{22} $$

is the non-advective PV flux [*Marshall and Nurser*, 1992]. *Thomas* [2005] shows that friction or diabatic processes, i.e. (17), acting at the sea surface will result in a reduction of the PV when

$$ fJ_z^{na} = f \left[\nabla_h b \times \mathbf{F} \cdot \hat{k} - \mathcal{D}(f + \zeta) \right] \Big|_{z=0} > 0, \tag{23} $$

where $\zeta = v_x - u_y$. Destabilizing atmospheric buoyancy fluxes reduce the buoyancy in the upper ocean $\mathcal{D} = Db/Dt < 0$, which for inertially stable flows, $f(f + \zeta) > 0$, results in a diabatic PV flux that satisfies condition (23) and reduces the PV. Friction can either input or extract PV from the fluid, depending on the orientation of the frictional force and the lateral buoyancy gradient. Down-front winds drive PV fluxes that meet condition (23) and, as described in section 2.4.2, Ekman driven convection ensues to mix the stratification and reduce the PV. Friction injects PV into the fluid when a baroclinic current is forced by upfront winds or during frictional spin-down by vertical mixing of momentum [*Boccaletti et al.*, 2007; *Thomas and Ferrari*, 2007]. Regardless of whether friction increases or decreases the PV, frictional modification of PV at the sea surface is largest in regions with strong lateral buoyancy gradients. Consequently, submesoscale phenomena with their enhanced baroclinicity are especially prone to frictional PV change.

4. IMPLICATIONS

As highlighted in the previous section, in the submesoscale regime $Ro \sim Ri \sim 1$, vertical motions are enhanced. The presence of upper ocean submesoscale motions consequently results in an enhancement vertical fluxes of buoyancy, PV, and biogeochemical tracers in the mixed layer. In this section, the implications of the submesoscale vertical tracer fluxes on budgets of the buoyancy, PV, and biological productivity will be described, and possible parameterization schemes will be discussed.

4.1. Effect on mixed layer stratification

The mixed layer instabilities (MLI) described in Section 2.3, release available potential energy from upper ocean fronts by inducing an upward buoyancy flux $\overline{w'b'}$ that tends to restratify the mixed layer, e.g. figure 2(c). Submesoscale vertical buoyancy fluxes play an important role in the buoyancy budget of the mixed layer by competing with, or augmenting, buoyancy fluxes associated with small-scale turbulent motions. This may be expressed as

$$ \partial_t \bar{b} + \nabla_h \cdot \left(\overline{\mathbf{u}_h' b'} + \overline{\mathbf{u}}_h \bar{b} \right) + \partial_z \left(\overline{w'b'} + \overline{w}\bar{b} \right) = \overline{\mathcal{D}}, \tag{24} $$

where, the overline denotes a lateral average, primes denote the deviation from that average, and the turbulent effects are encompassed in the diabatic term $\overline{\mathcal{D}}$. The horizontal and vertical eddy fluxes of buoyancy on the left hand side of (24) have contributions from both mesoscale and submesoscale flows. Submesoscale resolving numerical experiments show that the the vertical buoyancy flux $\overline{w'b'}$ is dominated by submesoscale eddies, while the lateral eddy buoyancy flux $\overline{\mathbf{u}_h' b'}$ is largely mesoscale [*Capet et al.*, 2007a; *Fox-Kemper et al.*, 2007]. While $\overline{w'b'}$ acts to redistribute buoyancy rather than generate a net input of buoyancy (as would be induced by an actual surface heat flux), expressing the vertical buoyancy flux in units of a heat flux gives the reader an appreciation for the strong restratifying capacity of submesoscale flows and motivates a parameterization. The submesoscale vertical buoyancy fluxes simulated in the numerical experiments of *Capet et al.* [2007a] had values equivalent to heat fluxes of $\mathcal{O}(100 \text{ W-m}^{-2})$. Having said this, *Capet et al.* [2007a] find that in their forced model simulations, the sea surface temperature is not very sensitive to increasing grid resolution, because the increase in restratifying submesoscale vertical buoyancy fluxes, though significant, is offset by an equivalent change in $\overline{\mathcal{D}}$ that tends to destratify the mixed layer.

Fox-Kemper et al. [2007] propose parameterizing the submesoscale vertical buoyancy flux using an overturning streamfunction, i.e. $\overline{w'b'} = \Psi \overline{b}_y$ (y indicates the cross-front direction). Using both dynamical and scaling arguments, they suggest that Ψ takes the form

$$\Psi = C_e \mu(z) \frac{H^2 \overline{\overline{b}_y}^z}{|f|}, \qquad (25)$$

where H is the mixed layer depth and $\overline{\overline{b}_y}^z$ is the highest-resolution cross-front buoyancy gradient in a non-submesoscale permitting simulation vertically averaged over the mixed layer. In performing a suite of numerical experiments with varying H, f, and lateral buoyancy gradients, some forced by a diurnally varying heat flux, but none by wind-stress, *Fox-Kemper et al.* [2007] showed that the buoyancy fluxes predicted by (25) compare well with those of the numerical solutions when the constant C_e is in the range 0.04-0.06 and the vertical structure function is $\mu(z) = -4z(z + H)/H^2$. Owing to its strong dependence on H, implementation of (25) in a global circulation model has the greatest impact in regions with deep mixed layers and is found to alleviate the problem of overly deep mixed layers at high latitudes [*Fox-Kemper, pers. comm.*].

4.2. Frictionally driven eddy potential vorticity fluxes

As described in section 2.4.2, down-front, wind-forced baroclinic flows form submesoscale circulations that advect PV and induce eddy PV fluxes. In two-dimensional simulations, submesoscale ASCs drive vertical PV fluxes that scale with the surface frictional PV flux. i.e.,

$$\overline{w'q'} \propto \overline{\nabla_h b \times \mathbf{F}} \cdot \hat{k}\big|_{z=0}, \qquad (26)$$

where q is the full Ertel PV, the overline denotes a lateral average, and primes denote a deviation from that average [*Thomas*, 2005]. At a fully three-dimensional front, an analogous result holds: along-isopycnal eddy PV fluxes on isopycnal surfaces that outcrop at a down-front, wind-forced front, scale with the frictional PV flux averaged over the outcrop window of the isopycnal (see Figure 5). Thus,

$$\int \int_{z=z_b}^{z=z_t} q''v'' \mathrm{d}z\mathrm{d}x \sim \iint_{\mathcal{A}} \nabla_h b \times \mathbf{F} \cdot \hat{k}\big|_{z=0} \mathrm{d}y\mathrm{d}x, \qquad (27)$$

where z_t, z_b are the top and bottom depths of the isopycnal layer, $\mathcal{A}$ is the area of the outcrop, x is the along-front direction, and the double primes denote the deviation from the thickness weighted isopycnal mean, i.e. $v'' = v - \overline{vh}^x/\overline{h}^x$, $h = z_t - z_b$ (the overline denotes a x average) [*Thomas*, 2007]. Results from a high resolution numerical simulation of a down-front, wind-forced front are shown in Figure 5 to illustrate the correspondence between eddy and frictional PV fluxes. The details of the initial configuration and numerics of the simulation can be found in *Thomas* [2007]. The key elements of the experiment are: the front is initially aligned in the x-direction and is forced by a spatially uniform down-front wind that persists through the 7.3 day long experiment. (This differs from the experiment described in *Thomas* [2007] where the wind-stress is shut off after 3.6 days.) As can be seen in panel (b) of Figure 5, the destruction of PV by the winds results in the formation of low, even negative, PV fluid where the frontal isopycnals outcrop into the surface Ekman layer ($z > -40$ m and 100 km $> y > 90$ km). Streamers of low PV surface fluid are subducted into the interior by the submesoscale instabilities, while high PV interior fluid is drawn to the surface, so that a positive eddy PV flux is induced at the y locations of the eddies. Once the eddies grow to finite amplitude, relation (27) holds, Figure

5 (d). In an experiment where the winds are turned off, (27) also holds, but the eddy PV flux, which persists after the winds, scales with the frictional PV flux at the time of the wind forcing, suggesting that there is a temporally nonlocal relation between the two fluxes [*Thomas*, 2007].

Isopycnal eddy PV fluxes are important because they drive an eddy-induced transport velocity (also known as a bolus velocity which can be related to the residual circulation of the transform Eulerian mean equations [*Plumb and Ferrari*, 2005]) that can play an important role in the advection of tracers [*Greatbatch*, 1998]. When the tracer is the buoyancy, advection by the eddy transport velocity can result in restratification. Indeed, the process of mixed layer restratification by MLI described by *Fox-Kemper et al.* [2007] and summarized in section 4.1, is expressed in terms of the flattening of isopycnals by an overturning stream function associated with the residual circulation. In the wind-forced problem, the eddy transport velocity $v^* = \overline{vh}^x/\overline{h}^x - \overline{v}^x$ was found to scale with the eddy PV flux, i.e. $v^* \approx -\iint q''v'' \mathrm{d}z\mathrm{d}x/(\Delta b L_x f)$, where Δb is the buoyancy difference across the isopycnal layer and L_x is the zonal width of the domain [*Thomas*, 2007]. Since the eddy PV flux also scaled with the surface frictional PV flux, i.e. (27), and the magnitude of frictional force is set by the strength of the wind-stress τ and Ekman depth δ_e, i.e. $|\mathbf{F}| \sim \tau/(\rho_o \delta_e)$, it was shown that the eddy transport velocity scaled with the Ekman flow $v_e \sim \tau/(\rho f \delta_e)$ [*Thomas*, 2007]. This result suggests that a closure scheme for submesoscale eddy PV fluxes and residual circulation at wind-forced baroclinic currents would have a dependence on the wind-stress, which would contrast to parameterizations based on down-gradient PV fluxes that make no explicit reference to atmospheric forcing.

4.3. Vertical transport and its effect on biogeochemistry

Vertical, diapycnal transport between the sea surface and the upper thermocline is instrumental in supplying nutrients for phytoplankton production, exchanging gases between the atmosphere and the ocean, and transferring heat, salt and momentum from the surface to the interior ocean. A number of oceanic properties are modified in the surface layer because the ocean experiences surface wind stress, heat and freshwater fluxes, exchanges gases with the atmosphere, and has a biologically active euphotic layer in which a large amount of organic matter is produced. The rates of property modification in the surface layer are inhibited by the rate of vertical transport across its base, as is evidenced by the strong gradient in the concentration profiles of properties in the upper 500–1000 m of the ocean.

The mechanisms by which the physical exchange of water and properties is achieved between the actively forced surface layer of the ocean and the relatively quiescent thermocline have posed a long-standing question in our understanding of the upper ocean. For example, estimates of new production (phytoplankton production relying on a fresh, rather than recycled, supply of nutrients) based on oxygen utilization and cycling rates [*Platt and Harrison*, 1985; *Jenkins and Goldman*, 1985; *Emerson et al.*, 1997] and helium fluxes [*Jenkins*, 1988], are much higher in the subtropical gyres than can be accounted for through the physical circulation in global carbon cycle models [*Najjar et al.*, 1992; *Maier-Reimer*, 1993]. Several studies, such as *McGillicuddy and Robinson* [1997] and *McGillicuddy et al.* [1998], suggest that eddies, which proliferate the ocean, act to pump nutrients to the euphotic zone. However, a basin-wide estimate for the eddy pumping fluxes [*Oschlies*, 2002; *Martin and Pondaven*, 2003] turns out to be inadequate in supplying the nutrient flux required to sustain the observed levels of productivity in the subtropical gyres. Further, the excursion
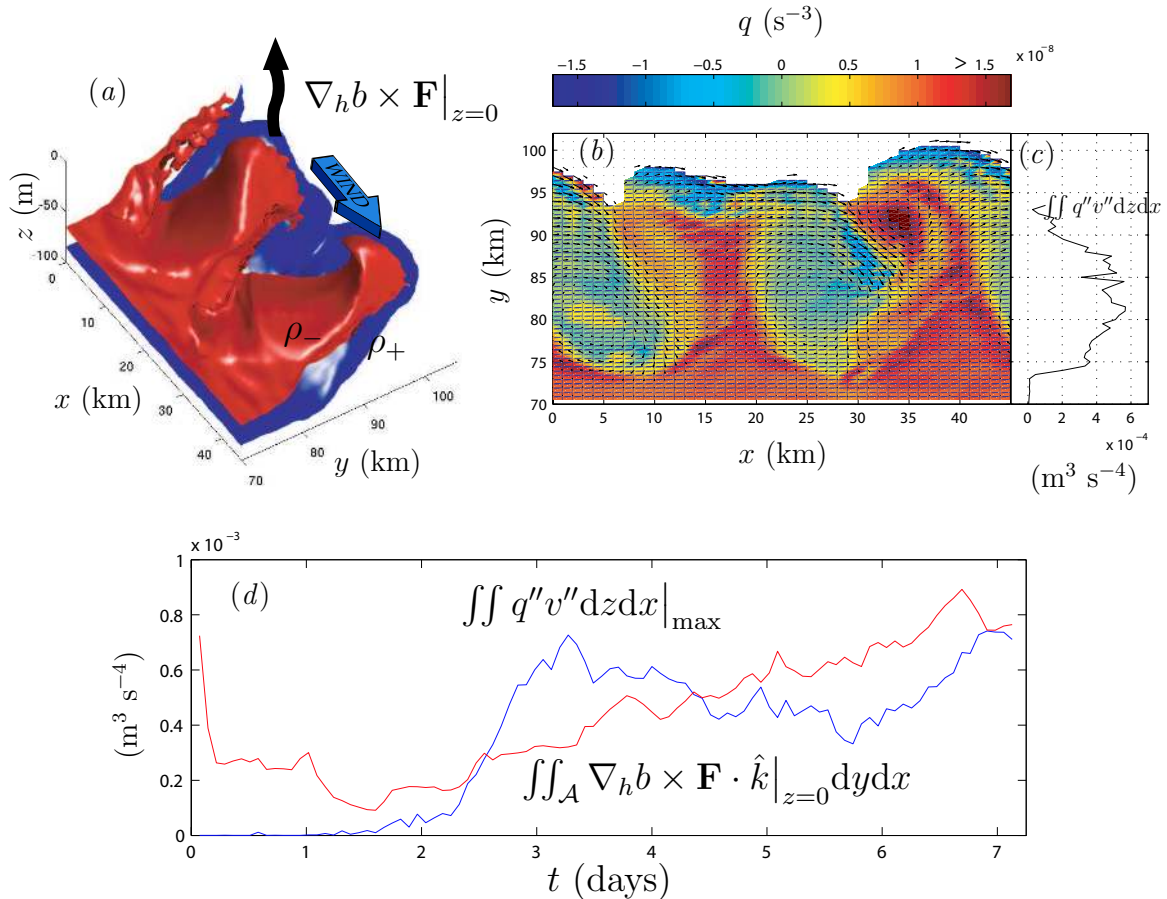
Figure 5. An example of submesoscale eddy PV fluxes driven by winds. (*a*) Down-front winds of strength 0.2 N m^{-2} forcing a front induce an upward frictional PV flux, triggering frontal instabilities that distort the bounding frontal isopycnal surfaces ρ_- (red) and ρ_+ (blue) ($\Delta\rho = \rho_+ - \rho_- = 0.2$ kg m^{-3}), shown here $t = 4.1$ days after the onset of the winds. (*b*) Isopycnal map of the PV (shades) and velocity (vectors) averaged in the vertical over the isopycnal layer shown in (*a*), illustrates the manner in which the instabilities subduct low PV from the surface while upwelling high PV from the pycnocline. (*c*) The correlation of the velocity and PV fields results in a net positive meridional eddy PV flux along the isopycnal layer $\iint q''v'' \mathrm{d}z\mathrm{d}x > 0$. (*d*) A timeseries of the maximum value of the eddy PV flux with respect to y (blue) and the frictional PV flux integrated over the outcrop area (red) reveal that the two fluxes scale with one another after the initial growth of the instabilities, i.e. $t > 3$ days.

of isopycnals at eddy centers does not support a diapycnal or advective vertical flux.

Fronts, on the other hand, are known to be sites for intense vertical advection and exchange [e.g. *Pollard and Regier*, 1990]. They are ubiquitous to the oceans and can intensify spontaneously through frontogenesis. Modeling studies of fronts show the generation of strong vertical velocities in frontal regions. At grid resolutions coarser than 2 km in the horizontal, vertical velocities are found to intensify with an increase in model resolution (from 40 to 10 km in *Mahadevan and Archer* [2000] for example, and from 6 to 2 km in *Lévy et al.* [2001]), suggesting finer- or submeso-scale processes at play. With submesoscale resolving grid spacing (≤ 1 km in the horizontal), the vertical velocity patterns begin to show convergence in simulations at various resolutions in [*Mahadevan*, 2006]. Though these simulations are largely in hydrostatic balance at this scale, the strongest ASCs generate $\mathcal{O}(1$ mm-s$^{-1})$ vertical velocities close to the surface, almost an order of magnitude larger than those found in earlier, coarser resolution observational and modeling stud-

ies of fronts, suggesting that vertical motion is dominated by submesoscale processes.

The submesoscale resolving model simulations in *Mahadevan and Tandon* [2006], show a preponderance of positive vorticity at submesoscales. Downward velocities are more intense and narrowly confined in elongated regions as compared to upward velocities. Even a weak, sustained downfront wind stress (of magnitude 0.025Nm^{-2}) generates a profusion of frontal structures which exhibit very large strain rates, relative vorticity and vertical velocities in the surface mixed layer. These characteristics can be explained by SG dynamics, as shown in Section 3.1.1. In contrast to the upper ocean, the velocity structure in the stratified pycnocline remains largely mesoscale in character. The communication of vertical submesoscale fluxes from the upper ocean to the deeper mesoscale circulation remains to be investigated.

Submesoscales processes have a significant effect on biological productivity in multiple ways. Firstly, the vertical fluxes supported by submesoscale processes have substantial effects on phytoplankton productivity. This is limited by the availability of nutrients that are at depth, and light, which
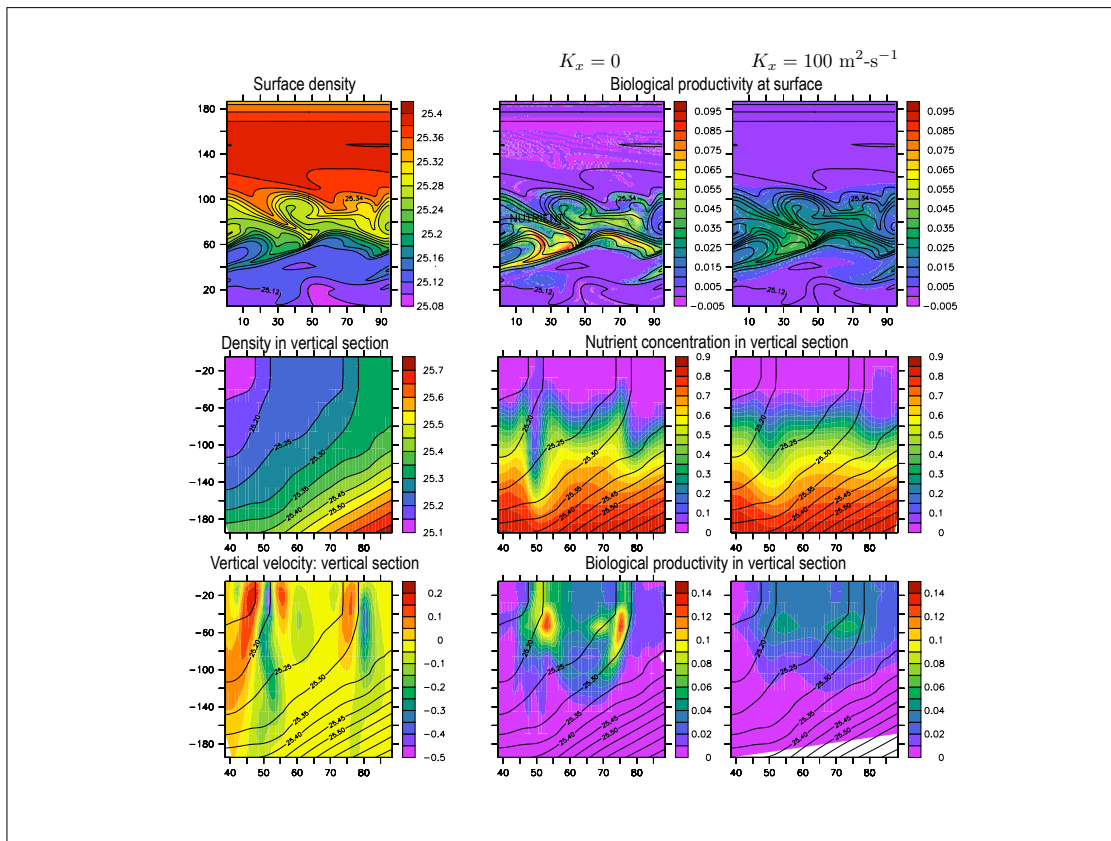
Figure 6. Modeling the nutrient and biological productivity due to submesoscale vertical nutrient supply shows a reduction in productivity when the lateral diffusivity of the biological tracers is increased 100 fold. The leftmost column shows the physical variables, density anomaly (in kg-m^3) and vertical velocity (in mm- s^{-1}). These are the same for two simulations compared. The middle column shows how the vertical submesoscale velocities supply nutrient and fuel biological productivity in the euphotic zone. The nutrient concentration is normalized between 0 and 1 and biological productivity is the rate of uptake of nutrient. The right column shows simulations with increased lateral diffusivity of the biogeochemical tracers. In this case, vertical nutrient transport is suppressed as pathways for transport are obliterated by diffusivity.

is at the surface. Since vertical motion is strongly linked to horizontal strain, regions of high strain in which tracer filaments are stretched and stirred by lateral motions at the surface, are also regions of intense up- or down-welling. The vertical velocities can, in turn, affect the lateral stirring of a phytoplankton bloom and enhance phytoplankton productivity through nutrient supply in filaments, as well as carbon export through the subduction of organic matter, and the air-sea transfer of carbon dioxide. A further effect of submesoscale processes on ocean biogeochemistry is that they generate lateral gradients at scales of $O(1km)$ that enhance lateral mixing, as well as create spatial heterogeneity in property distributions. A number of biological processes are affected by distribution patchiness, since they are limited by access to nutrients, oxygen, or properties that are involved in a reaction.

In numerical experiments, the rate of vertical exchange in the upper ocean can be quantified by tracking the cumulative vertical flux of a tracer with a vertical gradient in concentration. The flux is dependent not only on the vertical velocities, but also on the tracer's vertical concentration gradient and its evolution over time. In numerical

experiments described below, the vertical tracer profile is restored in time. Faster restoring maintains sharper gradients and results in higher fluxes. The ratio of the vertical advection time scale to the restoring time scale (analogous to a reaction time scale) is an important parameter (termed the Damköhler number) that affects not only the tracer flux and the time-averaged vertical tracer profile, but also the spatial heterogeneity of the tracer distribution in the upper ocean. Submesoscale processes increase spatial heterogeneity or variance of the spatial distribution of biogeochemical properties, either by introducing submesoscale concentration anomalies due to vertical advection across strong vertical gradients, or by the lateral straining and drawing out of filaments. In the former case, an increase in restoring (or reaction) time, shifts the spatial variance towards larger scales [*Mahadevan and Campbell*, 2002]. In the latter, slower reaction time scales offer greater scope for filamentation and increased submesoscale spatial variance [*Abraham*, 1998].

The enhancement of vertical tracer flux by wind forced submesoscales can be seen in Fig. 7, where we compare the vertical tracer flux from a numerical experiment in which submesoscales are generated by surface wind, to a case without wind forcing and submesoscale structure. In these simulations, the mixed layer is 50 m deep. Two tracers, one that

is plentiful at depth and depleted in the surface, and the other with the inverse characteristic, are restored to their mean initial profiles on a time scale of 3 days. The (positive) vertical flux of tracer is the amount of tracer that is added per time step to perform the described restoration of the tracer profile. By treating the tracers identically in both the wind-forced and unforced cases, the difference in the estimated vertical flux can be attributed entirely to the differences in their associated vertical velocities, which turns out to be substantial.

In another experiment (Fig. 6) the lateral variability in tracer distribution introduced by submesoscale processes is suppressed by increasing the lateral diffusivity of the tracer. The result is that narrow pathways for vertical transport are obliterated and vertical tracer flux is substantially diminished. When the tracer is considered to be nitrate, and its vertical transport results in phytoplankton productivity, the lateral diffusion of tracer diminishes phytoplankton productivity in the model.

5. Discussion

As described in section 4.1 restratification of the mixed-layer by submesoscale flows is accomplished at the expense of the available potential energy (APE) stored in the larger-scale baroclinic currents in which they form. This characteristic of submesoscale flows, which is a consequence of their strong secondary circulations, hints at their potentially important role in the energy budget of the ocean, a subject that is explored in this section.

The power spectrum of kinetic energy is often used to characterize the energy distribution in terms of length scales. While three-dimensional turbulence spectra show a characteristic -5/3 power law (or -5/3 slope in log-log space)

Figure 7. Comparison between the vertical flux of two tracers: C1, which is prescribed to be abundant at depth and surface depleted, and C2: abundant in the surface and depleted at depth. To the right, we show the time evolution of the vertical flux integrated over the domain of the model, as a function of time (ranging from 0 to 80 days). The black line shows the fluxes from the model run which develops submesoscale structure due to a constant down-front wind. The red line shows the flux from the model run without wind, and a paucity of submesoscale structure. The tracer concentrations are normalized between 0 and 1; the fluxes represent the rate at which the tracers are restored in both cases.

and the cascade of energy to smaller scales, stratified quasi-geostrophic mesoscale flow (geostrophic turbulence) spectra have a -3 slope [*Vallis*, 2006, e.g.] and a reverse energy cascade. With increasing grid resolution in a primitive equation model, *Capet et al.* [2007a] find the spectral slope transitions from -3 at mesoscale resolutions, to -2 at submesoscales in the upper ocean (also seen in *Legg and McWilliams* [2001]). *Boccaletti et al.* [2007] and *Fox-Kemper et al.* [2007] show that the energy peak triggered by mixed layer instability is transferred to both larger and smaller scales. The reverse cascade is physically manifest through the evolution of mixed layer eddies into larger ones, while the forward cascade is intimately tied to frontogenesis and frontal instabilities and, as revealed by the detailed energetics analysis of *Capet et al.* [2007c], is almost entirely associated with advection by the horizontally divergent, ageostrophic component of the flow.

Tulloch and Smith [2006] explain the transition from -3 to -5/3 slope in atmospheric spectra derived from aircraft observations using a finite depth surface quasi-geostrophic model (henceforth, SQG). SQG has also been advanced by *Hakim et al.* [2002] and *Lapeyre et al.* [2006] as appropriate for submesoscales as the model is capable of a forward energy cascade. As shown in Section 3.1.2, the frontogenetic motion described by semi-geostrophic equations also produces a forward energy cascade, since smaller length scale structures arise as frontogenesis occurs. In contrast to the QG equations, the SG equations do not neglect advection by ageostrophic flow, suggesting that the forward cascade described by SG models is more representative of the down-scale energy transfer simulated in *Capet et al.* [2007c]. In addition, there is a difference in the frontogenetic time scale between the SQG and SG models; an SQG model would need a much longer time than an SG model to develop frontal discontinuity [*Held et al.*, 1995; *Hoskins*, 1982].

As discussed in Section 3.1.2 and *Andrews and Hoskins* [1978], the semi-geostrophic frontal equation leads to flattening the slope of the KE spectrum to -8/3 at the time of frontal collapse. *Boyd* [1992] shows that this slope is ephemeral, and as the front evolves into a discontinuity, the slope of the energy spectra would flatten further to -2, which is characteristic of jump discontinuities. Numerical experiments that are actively being forced, either by buoyancy loss [*Legg and McWilliams*, 2001] or by wind-stress [*Capet et al.*, 2007a], also show a -2 slope for EKE, raising questions about the relative contributions of atmospheric forcing, frontogenesis, and APE release by instability, in the energetics of submesoscale flows.

Many questions remain open in regards with submesoscales. The numerical experiments described in section 4.2 emphasized the role of wind-forcing in driving submesoscale eddy PV fluxes. It remains to be seen if a similar result holds true for submesoscale motions driven by atmospheric buoyancy fluxes. Another question is whether there a general relation between atmospheric forcing induced non-advective PV fluxes (frictional or diabatic) and eddy PV fluxes, or if it is specific to the case of down-front winds.

The cumulative effects of the submesoscale on larger scales are beginning to be assessed in terms of mixed layer restratification (Fox-Kemper and Hallberg, 2007, in preparation), though it is not well understood how submesoscales affect the PV budget at the large scale. A review of the general ocean circulation energetics by *Wunsch and Ferrari* [2004] notes that the energy pathways from mesoscale eddies to internal waves are not well understood and have not been quantified. Since it turns out that the limits of balance coincide with the growth of unbalanced instabilities at the submesoscale, it is plausible that the energy pathway from mesoscale to internal waves is via unbalanced submesoscale instabilities [*Molemaker et al.*, 2005]. However, it remains to
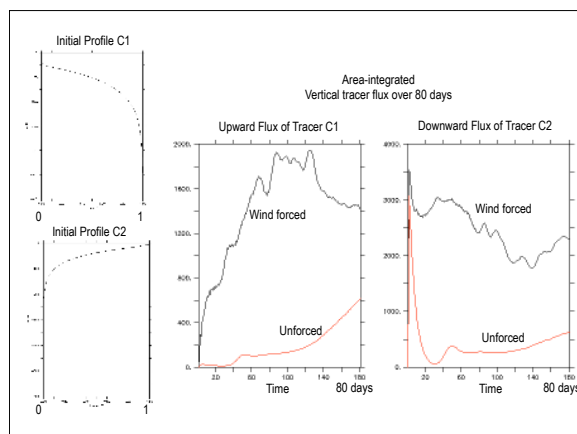
be shown whether this breakdown in balance leads to triggering energy into internal gravity waves, or whether direct interactions between the balanced submesoscale flows and a prexisting internal wave field affects energy transfer from large to small scales [*Bühler and McIntyre*, 2005; *Polzin*, 2006]. Whatever the mechanism, the downscale cascade of energy from submesoscales could have far-reaching implications on mixing and dissipation. This, plus the implications for vertical transport of buoyancy, momentum and biogeochemical properties, makes it important to improve our understanding of submesoscales.

References

Abraham, E. (1998), The generation of plankton patchiness by turbulent stirring, *Nature*, *391*, 577–580.

Andrews, D. G., and B. J. Hoskins (1978), Energy spectra predicted by semi-geostrophic theories of frontogenesis, *J. Atmos. Sci.*, *35*, 509–512.

Boccaletti, G., R. Ferrari, and B. Fox-Kemper (2007), Mixed layer instabilities and restratification, *J. Phys. Oceanogr.*, in press.

Boyd, J. P. (1992), The energy spectrum of fronts: Time evolution of shocks in burgers' equation, *J. Atmos. Sci.*, *49*, 128–139.

Bühler, O., and M. E. McIntyre (2005), Wave capture and wave-vortex duality, *J. Fluid Mech.*, *534*, 67–95.

Capet, X., J. McWilliams, M. Molemaker, and A. Shchepetkin (2007a), Mesoscale to submesoscale transition in the California Current system: Flow structure, eddy flux, and observational tests, *J. Phys. Oceanogr.*, p. submitted.

Capet, X., J. McWilliams, M. Molemaker, and A. Shchepetkin (2007b), Mesoscale to submesoscale transition in the California Current system: Frontal processes, *J. Phys. Oceanogr.*, p. submitted.

Capet, X., J. McWilliams, M. Molemaker, and A. Shchepetkin (2007c), Mesoscale to submesoscale transition in the California Current system: Energy balance and flux, *J. Phys. Oceanogr.*, p. submitted.

Charney, J. G. (1948), On the scale of atmospheric motions, *Geophys. Publ. Oslo*, *17*(2), 1–17.

Eliassen, A. (1948), The quasi-static equations of motion, *Geofys. Publikasjoner*, *17*(3).

Eliassen, A. (1951), Slow thermally or frictionally controlled meridional circulation in a circular vortex, *Astrophys. Norv.*, *5*, 19–60.

Emerson, S., P. Quay, D. Karl, C. Winn, L. Tupas, and M. Landry (1997), Experimental determination of the organic carbon flux from open-ocean surface waters, *Nature*, *389*, 951–954.

Flament, P., L. Armi, and L. Washburn (1985), The evolving structure of an upwelling filament, *J. Geophys. Res.*, *90*, 11,765–11,778.

Fox-Kemper, B., R. Ferrari, and R. Hallberg (2007), Parameterization of mixed layer eddies: I: Theory and diagnosis, *Journal of Physical Oceanography (submitted)*.

Gent, P. R., and J. C. McWilliams (1983), Consistent balanced models in bounded and periodic domains, *Dyn. Atmos. Oceans*, *7*, 67–93.

Gent, P. R., J. C. McWilliams, and C. Snyder (1994), Scaling analysis of curved fronts: validity of the balance equations and semigeostrophy, *J. Atmos. Sci.*, *51*, 160–163.

Greatbatch, R. J. (1998), Exploring the relationship between eddy-induced transport velocity, vertical momentum transfer, and the isopycnal flux of potential vorticity, *J. Phys. Oceanogr.*, *28*, 422–432.

Haine, T. W. N., and J. C. Marshall (1998), Gravitational, symmetric, and baroclinic instability of the ocean mixed layer, *J. Phys. Oceanogr.*, *28*, 634–658.

Hakim, G., and D. Keyser (2001), Canonical frontal circulation patterns in terms of Green's functions for the Sawyer-Eliassen equation, *Quart. J. Roy. Meteor. Soc.*, *127*, 1795–1814.

Hakim, G., C. Snyder, and D. J. Muraki (2002), A new surface model for cyclone-anticyclone asymmetry, *J. Atmos. Sci.*, pp. 2405–2420.

Held, I. M., R. T. Pierrehumbert, S. T. Garner, and K. L. Swanson (1995), Surface quasi-geostrophic dynamics, *J. Fluid Mech.*, *282*, 1–20.

Hosegood, P., M. Gregg, and M. Alford (2006), Submesoscale lateral density structure in the oceanic surface mixed layer, *Geophys. Res. Lett. (submitted)*.

Hoskins, B. J. (1975), The geostrophic momentum approximation and the semi-geostrophic equations., *J. Atmos. Sci.*, *32*, 233–242.

Hoskins, B. J. (1982), The mathematical theory of frontogenesis., *Annu. Rev. Fluid Mech.*, *14*, 131–151.

Hoskins, B. J., and F. P. Bretherton (1972), Atmospheric frontogenesis models: Mathematical formulation and solution, *J. Atmos. Sci.*, *29*, 11–37.

Hoskins, B. J., and I. Draghici (1977), The forcing of ageostrophic motion according to the semi-geostrophic equations and in an isentropic coordinate model., *J. Atmos. Sci.*, *34*, 1859–1867.

Hoskins, B. J., I. Draghici, and H. C. Davies (1978), A new look at the ω-equation., *Qt. J. R. Met. Soc.*, *104*, 31–38.

Jenkins, W. (1988), Nitrate flux into the euphotic zone near Bermuda, *Nature*, *331*, 521–523.

Jenkins, W., and J. Goldman (1985), Seasonal oxygen cycling and primary production in the Sargasso Sea, *Journal of Marine Research*, *43*, 465–491.

Lapeyre, G., P. Klein, and B. L. Hua (2006), Oceanic restratification forced by surface frontogenesis, *J. Phys. Oceanogr.*, *36*, 1577–1590.

Lee, C., E. D'Asaro, and R. Harcourt (2006a), Mixed Layer Restratification: Early Results from the AESOP Program, *Eos Trans. AGU, 87(52), Fall Meet. Suppl., Abstract OS51E-04*.

Lee, C. M., L. N. Thomas, and Y. Yoshikawa (2006b), Intermediate water formation at the Japan/East Sea subpolar front, *Oceanography*, *19*, 110–121.

Legg, S., and J. C. McWilliams (2001), Convective modifications of a geostrophic eddy field, *J. Phys. Oceanogr.*, *31*, 874–891.

Legg, S., J. C. McWilliams, and J. Gao (1998), Localization of deep ocean convection by a mesoscale eddy, *J. Phys. Oceanogr.*, *28*, 944–970.

Lévy, M., P. Klein, and A.-M. Treguier (2001), Impacts of submesoscale physics on production and subduction of phytoplankton in an oligotrophic regime, *Journal of Marine Research*, *59*, 535–565.

Mahadevan, A. (2006), Modeling vertical motion at ocean fronts: Are nonhydrostatic effects relevant at submesoscales?, *Ocean Modelling*, *14*, 222–240.

Mahadevan, A., and D. Archer (2000), Modeling the impact of fronts and mesoscale circulation on the nutrient supply and biogeochemistry of the upper ocean, *Journal of Geophysical Research*, *105*(C1), 1209–1225.

Mahadevan, A., and J. Campbell (2002), Biogeochemical patchiness at the sea surface, *Geophysical Research Letters*, *29*(19), 1926, doi:10.10292001GL014,116.

Mahadevan, A., and A. Tandon (2006), An analysis of mechanisms for submesoscale vertical motion at ocean fronts, *Ocean Modelling*, *14 (3-4)*, 241–256.

Maier-Reimer, E. (1993), Geochemical cycles in an ocean general circulation model, *Global Biogeochemical Cycles*, *7*, 645–678.

Marshall, J. C., and A. J. G. Nurser (1992), Fluid dynamics of oceanic thermocline ventilation., *J. Phys. Oceanogr.*, *22*, 583–595.

Martin, A., and P. Pondaven (2003), On estimates for the vertical nitrate flux due to eddy-pumping, *Journal of Geophysical Research*, *108*(C11), 3359, doi: 10.1029/2003JC001,841.

McGillicuddy, D., Jr., and A. Robinson (1997), Eddy induced nutrient supply and new production in the Sargasso Sea, *Deep Sea Research, Part I*, *44*(8), 1427–1450.

McGillicuddy, D., Jr., A. Robinson, D. Siegel, H. Jannasch, R. Johnson, T. Dickey, J. McNeil, A. Michaels, and A. Knap (1998), Influence of mesoscale eddies on new production in the Sargasso Sea, *Nature*, *394*, 263–266.

McKiver, and D. Dritschel (2006a), Balance in non-hydrostatic rotating stratified turbulence, *J. Fluid Mech.*, p. submitted.

McKiver, W. J., and D. G. Dritschel (2006b), Balance in non-hydrostatic rotating stratified turbulence, *J. Fluid Mech.*, submitted.

McWilliams, J. (2003), Diagnostic force balance and its liimits, in *Nonlinear Processes in Geophysical Fluid Dynamics*, pp. 287–304, Kluwer Academic Publishers.

McWilliams, J. C. (1985), Submesoscale, coherent vorticies in the ocean, *Rev. Geophys.*, *23*, 165–182.

McWilliams, J. C., and P. R. Gent (1980), Intermediate models of planetary circulations in the atmosphere and ocean, *J. Atmos. Sci.*, *37*, 1657–1678.

McWilliams, J. C., M. J. Molemaker, and I. Yavneh (2001), From stirring to mixing of momentum: Cascades from balanced flows to dissipation, in *Proceedings of the 12th 'Aha Huliko'a Hawaiian Winter Workshop*, University of Hawaii, pp. 59–66.

Molemaker, M. J., J. C. McWilliams, and I. Yavneh (2005), Baroclinic instability and loss of balance, *J. Phys. Oceanogr.*, *35*, 1505–1517.

Nagai, T., A. Tandon, and D. L. Rudnick (2006), Two-dimensional ageostrophic secondary circulation at ocean fronts due to vertical mixing and large-scale deformation, *J. Geophys. Res.*, *111*, C09,038, doi:10.1029/2005JC002964.

Najjar, R., J. Sarmiento, and J. Toggweiler (1992), Downward transport and fate of organic matter in the ocean: Simulations with a general circulation model, *Global Biogeochemical Cycles*, *6*, 45–76.

Niiler, P. (1969), On the Ekman divergence in an oceanic jet., *J. Geophys. Res.*, *74*, 7048–7052.

Oschlies, A. (2002), Can eddies make ocean deserts bloom?, *Global Biogeochemical Cycles*, *16*, 1106, doi:10.1029/2001GB001,830.

Pinot, J.-M., J. Tintore, and D.-P. Wang (1996), A study of the omega equation for diagnosing vertical motions at ocean fronts, *Journal of Marine Research*, *54*, 239259.

Platt, T., and W. G. Harrison (1985), Biogenic fluxes of carbon and oxygen in the ocean, *Nature*, *318*, 55–58.

Plumb, R. A., and R. Ferrari (2005), Transformed Eulerian-mean theory. Part I:nonquasigeostrophic theory for eddies on a zonal-mean flow, *J. Phys. Oceanogr.*, *35*, 165–174.

Pollard, R., and L. Regier (1990), Large variations in potential vorticity at small spatial scales in the upper ocean, *Nature*, *348*, 227–229.

Pollard, R., and L. Regier (1992), Vorticity and vertical circulation at an ocean front, *Journal of Physical Oceanography*, *22*, 609–625.

Polzin, K. L. (2006), How Rossby waves break. Results from POLYMODE and the end of the enstrophy cascade, *J. Phys. Oceanogr.*, submitted.

Rudnick, D. L. (1996), Intensive surveys of the Azores front. Part II: Inferring the geostrophic and vertical velocity fields, *Journal of Geophysical Research*, *101*(C7), 16,291–16,303.

Rudnick, D. L., and J. R. Luyten (1996), Intensive surveys of the Azores front: 1. Tracers and dynamics., *J. Geophys. Res.*, *101*, 923–939.

Sawyer, J. (1956), The vertical circulation at meteorological fronts and its relation to frontogenesis, *Proc. Royal. Soc. London, Ser. A*, *234*, 346–362.

Shay, L., T. Cook, and P. An (2003), Submesoscale coastal ocean flows detected by very high frequency radar and autonomous underwater vehicles, *J. Atmos. Oceanic Technol.*, pp. 1583–1599.

Shearman, R., J. Barth, and P. Kosro (1999), Diagnosis of three-dimensional circulation associated with mesoscale motion in the California current, *Journal of Physical Oceanography*, *29*, 651–670.

Stern, M. E. (1965), Interaction of a uniform wind stress with a geostrophic vortex., *Deep-Sea Res.*, *12*, 355–367.

Stone, P. (1966), On non-geostrophic baroclinic stability, *Journal of the Atmospheric Sciences*, *23*, 390–400.

Stone, P. (1970), On non-geostrophic baroclinic stability: Part II, *Journal of the Atmospheric Sciences*, *27*, 721–727.

Stone, P. (1971), Baroclinic instability under non-hydrostatic conditions, *Journal of Fluid Mechanics*, *45*, 659–671.

Tandon, A., and C. Garrett (1994), Mixed layer restratification due to a horizontal density gradient, *Journal of Physical Oceanography*, *24*, 1419–1424.

Thomas, L. N. (2005), Destruction of potential vorticity by winds, *J. Phys. Oceanogr.*, *35*, 2457–2466.

Thomas, L. N. (2007), Formation of intrathermocline eddies at ocean fronts by wind-driven destruction of potential vorticity, *Dyn. Atmos. Oceans*, accepted.

Thomas, L. N., and R. Ferrari (2007), Friction, frontogenesis, frontal instabilities and the stratification of the surface mixed layer, *J. Phys. Oceanogr.*, in preparation.

Thomas, L. N., and C. M. Lee (2005), Intensification of ocean fronts by down-front winds, *J. Phys. Oceanogr.*, *35*, 1086–1102.

Thomas, L. N., and C. M. Lee (2007), Inverse method for determining mesoscale and submesoscale vertical circulation with application to the subpolar front of the Japan/East Sea, *J. Phys. Oceanogr.*, in preparation.

Thomas, L. N., and P. B. Rhines (2002), Nonlinear stratified spin-up., *J. Fluid Mech.*, *473*, 211–244.

Tintore, J., D. Gomis, S. Alonso, and G. Parrilla (1991), Mesoscale dynamics and vertical motion in the alboran sea, *Journal of Physical Oceanography*, *21*, 811–823.

Tulloch, R., and K. Smith (2006), A theory for the atmospheric energy spectrum: Depth limited temperature anomalies at the tropopause, *Proc. Nat. Acad. Sci. USA*, *103*(40), 14,690–14,694.

Vallis, G. K. (2006), *Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-scale Circulation*, 745 pp., Cambridge University Press.

Vélez-Belchí, P., M. Vargas-Yáñez, and J. Tintoré (2005), Observation of a western Alborán gyre migration event, *Progress in Oceanography*, *66*, 190–210.

Viúdez, A., and D. Dritschel (2003), A balanced approach to modelling rotating stably stratified geophysical flows, *J. Fluid Mech.*, *488*, 123–150.

Viúdez, A., and D. Dritschel (2004), Potential vorticity and the quasigeostrophic and semigeostrophic mesoscale vertical velocity, *Journal of Physical Oceanography*, *34*, 865887.

Wunsch, C., and R. Ferrari (2004), Vertical mixing, energy and the general circulation of the oceans, *Annual Review of Fluid Mechanics*, *36*, 281–314.

Yoshikawa, Y., K. Akitomo, and T. Awaji (2001), Formation process of intermediate water in baroclinic current under cooling, *J. Geophys. Res.*, *106*, 1033–1051.

L. N. Thomas, Department of Physical Oceanography, Woods Hole Oceanographic Institution, MS 21, Woods Hole, MA 02543. (e-mail: lthomas@whoi.edu)

A. Tandon, Physics Department and SMAST, University of Massachusetts, Dartmouth, 285 Old Westport Road, North Dartmouth, MA 02747. (e-mail: atandon@umassd.edu)

A. Mahadevan, Department of Earth Sciences, Boston University, 685 Commenwealth Avenue, Boston, MA 02215. (e-mail: amala@bu.edu)