4286

VERTICAL STRUCTURE OF MESOSCALE OCEAN CURRENTS IN THE INDIAN OCEAN
– OBSERVATION, NUMERICAL MODELING AND THEORY

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE
UNIVERSITY OF HAWAII IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

OCEANOGRAPHY

MAY 2003

By

Shuiming Chen

Dissertation Committee:

Eric Firing, Chairperson
Doug Luther
Humio Mitsudera
Bo Qiu
Fei-Fei Jin

To the memory of my brother **Shuigen CHEN**

纪念敬爱的哥哥 陈水根

ACKNOWLEDGEMENTS

I would like to thank my dissertation committee Drs. Eric Firing, Doug Luther, Bo Qiu, Humio Mitsudera and Fei-Fei Jin for their helpful suggestions and comments. Particular gratitude goes to my advisor, Dr. Eric Firing, for his enthusiastic encouragement, guiding advice, insight, and patience. His idea to explore the model output of a high resolution GCM in light of observations was very important to this study.

The LADCP profiles were collected at sea by Dr. Eric Firing, Dr. Peter Hacker, Dr. Jules Hummon, and Ms. Elodie Kestenare on the R/V Knorr during the 1994-1995 Indian Ocean survey of the WOCE Hydrographic Program.

The JAMSTEC model output is made available by the kindness of Dr. Akio Ishida of JAMSTEC, Japan, and Dr. Humio Mitsudera.

The help and interest from Dr. Peter Hacker, Dr. Jules Hummon, Dr. Kathleen Donohue and Mr. Ted Durland during the course of this study are appreciated.

This study started with the Solaris machine "moli", then moved to the two-cpu Linux PC "ipu" with a raid array. Most of the text editing was done on the Linux laptop "manini". The hardware and software support from Dr. Eric Firing, Dr. Jules Hummon and Mr. Derek Young are appreciated.

ABSTRACT

The classical linear Rossby wave theory suggests that the barotropic and baroclinic modes of the mesoscale eddy field would disperse and become uncorrelated with each other in space and time. In contrast, the correlation between the barotropic and first baroclinic modes was noted from moored current meter records by Davis (1976). The sparse vertical sampling and inadequate record length left room for doubt and no dynamical explanation has been offered since then. In this study, hundreds of full-depth Lowered Acoustic Doppler Current Profile (LADCP) velocity profiles and the three years output from a General Circulation Model (GCM) are used. Analyses of the observation and model output confirm that outside the equatorial region the barotropic and first baroclinic modes are indeed correlated so that the velocity decreases throughout the water column from its maximum at the surface. We may call this the dominant vertical structure, which is quantified by the first Empirical Orthogonal Function (EOF) from the LADCP profiles and model output.

The phase speed of the dominant vertical structure in the model is approximately that of non-dispersive first mode long Rossby waves, even at frequencies above the Rossby wave cutoff. The model also shows that, even though the dominant vertical structure contains 50-80% of the variance, it alone provides an incomplete picture of the vertical structure of mesoscale eddies. There is a phase shift of almost 90° from top to bottom; in a zonal section, lines of constant phase slope down to the west. Or, the lower layer leads the upper layer, given westward propagation.

We could not find a satisfactory theory which could fully explain the following features: 1) correlation of the barotropic and first baroclinic mode; 2) enhanced westward propagation; and 3) phase shift in the vertical. The non-linear model with finite interface perturbation succeeds partially. The linear Rossby waves with bottom dissipation could explain the above features if dissipation is strong enough.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Mesoscale eddies dominate the open ocean velocity in mid-latitudes. This can be illustrated from the sea surface height (SSH) anomaly field either from a high resolution global circulation model (GCM) or from altimetric observation (figure 1.1). The motion around SSH elevations and depressions in mid-latitudes can be estimated by geostrophy: 10 cm SSH difference between two locations 100 km apart corresponds to $\sim$10 cm s^{-1} mean current across the section between the two locations. With more than 10 years of altimetric SSH anomaly, we have extensive information about surface mesoscale currents in mid-latitudes, including their horizontal (x,y) and temporal (t) variations (e.g., Stammer 1997), but our knowledge of their vertical structure is scattered and partial. We deduce the vertical structures of mesoscale motion mostly from moored current meter records, which have sparse space-time coverage and vertical resolution.

Though energetic, the mesoscale motion is not well understood. For example, there is still a controversy over the horizontal propagation of surface eddies inferred from altimetric SSH anomaly (Chelton and Schlax 1998). A solid theory should account for the horizontal propagation as well as the vertical structure, since they are different aspects of the same problem.

Review: Theories and Observations

The simplest theory about the vertical structure of mesoscale currents involves dynamical normal modes, which result from the linearized dynamics of a flat bottom ocean without mean flow and dissipation (Kundu 1990). These modes are orthonormal and complete, so that normal mode decomposition has long been the framework for describing the vertical structure of ocean currents, even though their dynamical relevance to any specific phenomenon, for example the mesoscale motions in the present study, is not clear. Similar modes occur in reduced gravity models such as the 2.5-layer model used by Liu (1999), but they are not orthogonal. Some other relevant theories are:

- **Shear Modes and Baroclinic Instability** This involves looking at the interaction between the mean flow and Rossby waves. There are two consequences due to the mean flow: one is the Doppler shift; another is to modify the background potential vorticity gradient, which is simply the planetary vorticity gradient (the β effect) without a mean flow, because of the isopycnal tilt corresponding to the vertical shear in the mean geostrophic flow. The resulting vertical structures—shear modes—will be different from the normal modes. If the mean flow shear is strong enough, the effective β changes sign in different layers, and the Rossby waves can become baroclinically unstable (Pedlosky 1987). The results from a two-layer model will be briefly discussed in the theoretical part of this study.

- **Equivalent Barotropic Mode** Killworth (1992) found that the six-year mean flow in the Fine Resolution Antarctic model is self-similar vertically (significant positive spatial correlation between different depths) and the velocity decays with depth. He called this structure the equivalent barotropic mode. He suggested that the steering of the mean flow by topography is important in the dynamics, but his argument seems incomplete (Chen 2000). In this study, we found the vertical structure of the mid-latitude mesoscale motion is similar to that of the mean flow in the Antarctic Circumpolar Circulation (ACC) region, except that the velocity decays more rapidly towards the bottom in the former. Because the so-called equivalent barotropic mode is probably not a dynamical mode (i.e., not an eigensolution of governing equations), we will avoid using the terminology. We will instead call it the dominant vertical structure as inferred from an Empirical Orthogonal Function (EOF) analysis.

- **Barotropization in Turbulence** Two-layer turbulence with weak dissipation in idealized numerical models (e.g., Rhines 1977 and Arbic and Flierl 2002) displays barotropization: the turbulence tends to be dominated by the barotropic mode with horizontal scales larger than the Rossby radius of deformation. With strong bottom friction (so the growth rate of eddies due to baroclinic instability and the decay rate due to bottom friction are of the same order), Arbic and Flierl (2002) demonstrate that an equilibrium near the Rossby radius of deformation is reached in which the barotropic and baroclinic modes are phase-locked with comparable amplitudes so that the flow in the lower layer is much weaker.

- **Isolated Structures (Modon)** One may use an isolated Gaussian eddy in either stationary or moving coordinates as "a building block" for the mesoscale eddy field, and the "no net angular momentum" theorem is then relevant (e.g., Flierl 1987):

  $$\beta \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \int_{-H}^{0} dz \psi = 0,$$

  where ψ is the geo-potential and a rigid lid and flat bottom are assumed. Notice that β plays a crucial role in the theorem. If we further assume ψ is separable so that $\psi \equiv f(x,y;t)g(z)$, the theorem implies

$$\beta \int_{-H}^{0} g(z)dz = 0,$$

i.e., the isolated eddy is baroclinic. The physics of the theorem is clearly stated in Flierl (1987): Due to the increase in the Coriolis parameter northward, there is a net north/south force on a recirculating flow. The net westward transport on the northern side of a cyclonic eddy, for example, must balance the net eastward transport on the southern side; yet the southward Coriolis force on the latter is not as strong as the northward Coriolis force on the former, then there is a net force, called "Rossby force" (Flierl 1987). In the case assuming a rigid lid and flat bottom, there is no other overall force that can balance the "Rossby force"; isolated features therefore cannot exist and remain isolated unless they have no net angular momentum then no "Rossby force". For an eddy in a one and half layer reduced gravity model, the eddy could remain isolated and move steadily westward at the speed approximately of the long Rossby wave (e.g., Nof 1982 and Cushman-Roisin *et al.* 1990).

Most of our observational evidence about the vertical structures of mesoscale motion has come from vertical arrays of moored current meters. The first attempt of this kind is by Davis (1976), using observations from the mid-ocean dynamics experiment - MODE I. He performed an Empirical Orthogonal Function (EOF) analysis (see chapter 2 for an introduction to the EOF analysis) and found considerable difference between the shapes of the empirical modes and the shapes of dynamical normal modes. Namely, the first EOF is unidirectional and surface-intensified; alternatively, in a decomposition into dynamical modes, the barotropic and first baroclinic modes are dominant and temporally correlated. The dominance and correlation were also shown in the survey by Wunsch (1997), who used a large number of globally distributed long current meter records. However, Wunsch expressed reservations about the EOF method due to its statistical uncertainty and did not intend to infer the dominant vertical structure (the first EOF) as Davis (1976) did. Notice that the time scale of the mesoscale motions is about 100 days, so that a few-year time series only gives us a few degrees of freedom. Another potential problem is the sparse vertical sampling.

We can also learn about the vertical structure of currents from hydrographic observations, but with a crucial limitation: only the baroclinic components can be calculated.

In term of vertical resolution, velocity profiles are ideal for the study of the vertical structures of currents. Such measurements date back to 1970s (Rossby 1974). With eight profiles at a site 70 km south of Bermuda, Rossby demonstrated the dominance of the barotropic and first baroclinic modes (48% and 39% of total variance respectively, a rough equipartition). The major problem with velocity profiles is that they contain energy *from the entire frequency spectrum*. Until recently, no long-term time series of velocity profiles has existed that allows us to filter out the high frequency internal waves and tides. Secondly, even though we have time series of a few years, the statistical uncertainty for the mesoscale motions remain.

Overall, sampling has been inadequate to clarify the vertical structure of the mesoscale motion.

Two Main Questions, Approaches and Results

Our main questions relate to the early speculation by Davis (1976):

their correlation coefficient [between the barotropic and first baroclinic modes] *is .55 here. Why this is so is uncertain: either the coupling may be brought about fortuitously or through systematic non-linear coupling between the modes (non-wave-like behavior).*

1. **What are the vertical structures of mesoscale motion in mid-latitudes and what are their propagation characteristics?**

   To address the first part of this question, we use the full-depth Lowered Acoustic Doppler Current Profiler (LADCP) profiles collected at all CTD rosette stations during the Indian Ocean survey of the WOCE Hydrographic Program (see chapter 2). Considering that the horizontal scale of the mesoscale motion is about 100 km and that the WOCE stations were typically separated by 55 km, we expect to have independent samples every other profile. With more than 300 profiles in our analysis, we have more than 150 degrees of freedom. Therefore, the space-time sampling of the present dataset is unprecedented for obtaining large number of degrees of freedom and rendering statistical confidence.

   Linear mode decomposition and EOF analysis are applied. We confirm that outside the equatorial region of the Indian Ocean, the barotropic and first baroclinic modes are correlated so as to reinforce each other in the upper ocean. The first EOF is labeled as the dominant vertical structure, and it has about 40% in middle latitudes and 80% in high latitudes.

   After finding a favorable comparison between the first EOFs from the LADCP profiles and from a numerical oceanic General Circulation Model (GCM) with high horizontal and vertical resolutions (the Japan Marine Science Technology Center (JAMSTEC) model, see appendix B), we used the model output to study eddy propagation and energy sources. In mid-latitudes, the propagation speeds of the dominant vertical structure in the model are close to those of the long Rossby waves of the first baroclinic mode, even at frequencies above the Rossby wave cutoff. In high latitudes (south of 30°S) the inferred speeds are faster than those of the long Rossby waves, while in lower latitudes the inferred speeds are slower. From the model output, we also learned that the lower layer is leading the upper layer, given westward propagation. Therefore, the dominant vertical structure, as quantified by the first EOF, is not a dynamical mode: $\psi(x, y, z; t)$ is not separable as $\psi(x, y, z; t) \approx f(x, y; t)g(z)$.

2. **What are the dynamics?**

   The dynamics should explain

   - The correlation of the barotropic and first baroclinic modes;
   - Different dispersion relations from those of linear Rossby waves;
   - The tile in lines of constant phase, down to the west.

   If the LADCP profiles and model output had shown that the dominant vertical structure is one of normal modes, e.g., the first baroclinic mode with one mid-depth zero-crossing, and that each mode propagates according to the dispersion relation of the linear Rossby wave of that mode, then we would claim that we understand the dynamics – we could write down the linear potential vorticity equation and solve it, and match the modal solution of that mode with what was observed. The reality is that there is no linear mode matching the dominant vertical structure in the observation and model output, and no linear mode dispersion relation matches the dispersion relation inferred from the model output.

The approach used in this study is to check factors not included in the linear Rossby wave theory: mean flow, non-linearity and bottom dissipation (Chapter 10). In other words, we add one of those terms to the potential vorticity equations for the linear Rossby waves, and see whether there is a modal solution which could explain all or part of what was observed and modeled. To be analytically tractable, a two-layer model or a constant N^2 model is used. We found that:

- Shear modes could be important in high latitudes;
- The two-layer non-linear model with a finite interface perturbation can marginally explain the phase shift and will work better if the flow is in transition but has not reached quasi-steady state;
- Bottom dissipation could explain the behavior in mid-latitudes, but we do not know if bottom friction in the ocean is strong enough to cause the observed vertical structure.

Because of the simplicity of the model and the uncertainty about the parameters in the models, the results are not decisive but could serve as hypotheses of further studies.

Layout

We will present the results in three parts:

1. Observation: LADCP
2. Numerical modeling: JAMSTEC GCM

3. Theories: two-layer models

Most of chapters will start with a brief summary of the chapter, and the conclusion chapter will detail the results for each part. As seen, the first two parts address our first question and the third part addresses our second question.

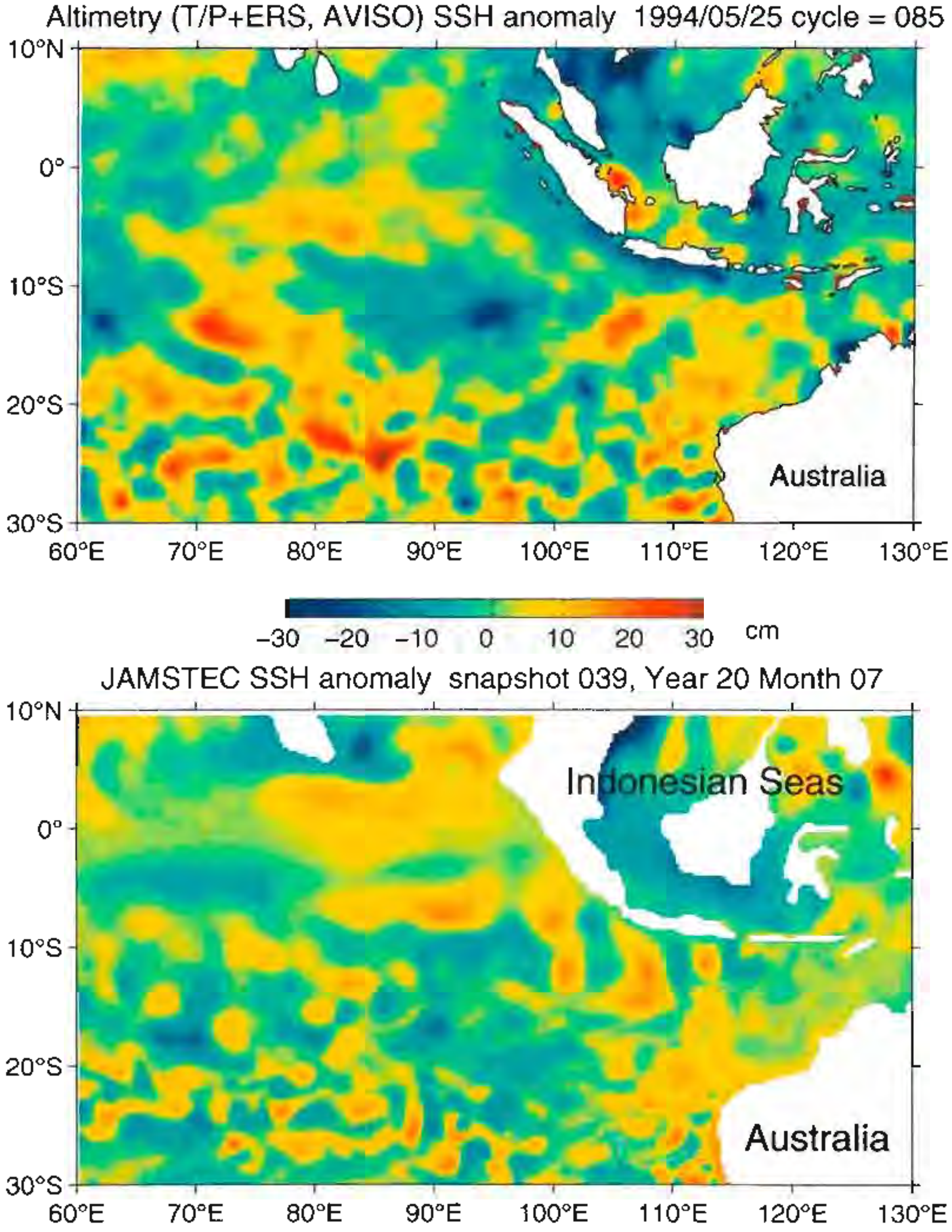


Figure 1.1: Snapshots of sea surface height anomaly in the eastern Indian Ocean, measured by altimetry (Archiving, Validation and Interpretation of Satellite Oceanographic Data (AVISO) group in CNES, French, http://www.oceanobs.com) and modeled by the JAMSTEC model (Ishida *et al* 1998). The anomaly in the upper panel is relative to the 7-year mean from 1993 to 1999) and the anomaly in the lower panel is relative to the 3-year mean from model 20 to 22). The same color scale is used in both panels. A brief description of the model is in appendix B.

Part I

OBSERVATION: LADCP

CHAPTER 2

DATA AND METHODS

2.1 Data

The Lowered Acoustic Doppler Current Profiler (LADCP) is a self-contained ADCP that is usually operated with a Conductivity-Temperature-Depth (CTD) rosette (Firing and Gorden 1990). The full-depth LADCP velocity profiles are typically gridded at 20 m intervals with ~40 m effective vertical resolution. A detailed description of the LADCP instrumentation and data processing algorithms as used in the data set presented here are given in Fischer and Visbeck (1993) and King *et al.* (2001).

As stated in King *et al.* (2001), the error analysis of LADCP profiles is complicated. Error could be related to instrument performance (e.g. compass error), as well as to environmental factors such as backscattering strength. Furthermore, error characteristics differ among different vertical scales. It is therefore not sensible to give a simple error estimate for a large dataset. Nonetheless, past experience suggests that final LADCP velocities have *rms* errors of a few cm s^{-1} (King *et al.* 2001).

The barotropic and first baroclinic mode projections of the velocity profiles are the most relevant in this study. The barotropic component is usually least worrisome, with an error 1-2 cm s^{-1} or less (Hacker *et al.* 1996). Estimated barotropic tides (Egbert *et al.* 1994) are removed from the LADCP profiles prior to analysis. The error in the first baroclinic component is not clear. The favorable comparison, however, between the first baroclinic mode of the LADCP velocity shear and that of the geostrophic shear in section 3.3 suggests that the component is not overwhelmed by noise.

The CTD profiles, used to calculate normal modes and geostrophic shears, are gridded at 2 decibar intervals, with an accuracy in temperature of ±0.001°C, salinity of ±0.002 *psu* and pressure of ±1 decibar (the WOCE standard for CTD measurements, King *et al.* 2001).

During the Indian Ocean survey of the WOCE Hydrographic Program, more than one thousand LADCP profiles were collected (figure 2.1) from December 1994 to December 1995. In order to analyze the dominant vertical structure of the mesoscale ocean currents in the mid-latitudes, we use the subset consisting of 366 profiles that are deeper than 4000 m and fall in the latitude range of 5° to 35°. There are two advantages for doing so: the subset is more homogeneous in terms of vertical range and latitudinal locations; the profiles within western boundary currents and over shallow water are mostly excluded.

2.2 Normal Mode Decomposition and EOF Analysis

The normal modes are the eigenfunctions of the following equation, which assumes linear ocean dynamics, with a flat bottom and without mean flows (e.g. Kundu 1990):

$$\frac{d}{dz}\frac{1}{N^2}\frac{dP_n}{dz}+\frac{1}{c_n^2}P_n=0, \qquad (\mathrm{n}=0,1,2,\cdots).$$

Top and bottom conditions are $\frac{dP_n}{dz}+\frac{N^2}{g}P_n=0 \quad (z=0)$, and $\frac{dP_n}{dz}|_{z=-H}=0$, respectively. N is the Brunt-Vaisala frequency, depending solely on the density profile. If N^2 is constant and the top condition is approximately $\frac{dP_n}{dz}|_{z=0}=0$ (rigid lid approximation), the above equation has solutions $P_n=\cos(n\pi z/H)$ $(n=0,1,2,\cdots$ and $z\in(-H,0))$. The eigenvalues are $c_n=\frac{NH}{n\pi}$ $(n=0,1,2,\cdots)$. $P_n(z)$ is called the n*th* mode of pressure and horizontal velocity. The barotropic mode, denoted by $n=0$, is simply a constant. The first baroclinic mode is denoted by $n=1$, the second by

$n = 2$, and so on.

A realistic N^2 profile will lift the zero-crossings of the above baroclinic modes upward but the overall vertical structures remain similar. Notice that the lower the mode (starting from zero for the barotropic mode and one for the first baroclinic mode, and so on) the larger the vertical scale. Conventionally, each mode is normalized so that its depth-averaged variance is unity and it is positive at the surface. In this study, the N^2 profiles are calculated from the one-time CTD profiles with which the LADCP profiles were collected together. The mode vertical structures, particularly of the first few modes, will not differ significantly from using either synoptic or long-term mean CTD profiles.

The n*th* mode amplitude (u_n) of the E-W component ($u(z)$) of an LADCP profile is defined as

$$u_n = \frac{1}{H}\int_{-H}^{0} u(z)P_n(z)dz.$$

Since the modes are orthonormal, the mode amplitudes are independent each other so that the distribution of the squared mode amplitudes over mode number can be viewed as one kind of variance-preserving energy spectrum of the LADCP profiles. As will be shown in the following, the EOFs are also orthonormal so that their associated variance $\lambda_i (i = 1, 2, \ldots, M)$ can be viewed as another kind of variance-preserving energy spectrum.

EOFs capture the inherent structures in the data. We view our LADCP dataset as two-dimensional: water depths (0 m, 20 m, ...) versus station numbers (1,2,..., which are arbitrarily assigned as the labels). The data matrix is defined as

$$\mathbf{A} = \begin{pmatrix} \cdots & u_{1i} & \cdots \\ \cdots & u_{2i} & \cdots \\ \cdots & \vdots & \cdots \\ \cdots & u_{Mi} & \cdots \end{pmatrix},$$

where the ith column is the LADCP profile at the ith station with M vertical grids. All LADCP profiles are truncated to a common depth of 4000 m, so $M = 201$. The EOFs $F_i(i = 1, 2, \ldots, M)$ are the eigenvectors of $\mathbf{A}\mathbf{A}^T$, i.e., the covariance matrix, satisfying:

$$\mathbf{A}\mathbf{A}^T F_i = \lambda_i F_i,$$

where λ_i is the corresponding eigenvalue of F_i. When $\lambda_i(i = 1, 2, \ldots, M)$ are sorted in descending order, F_1 is called the first EOF, F_2 the second EOF and so on.

Following Fukumori and Wunsch (1991), F_1 is the most "common" column vector, in the sense that it has the largest projection on all of the columns in $\mathbf{A}$:

$$||F_1^T\mathbf{A}|| = \sqrt{F_1^T\mathbf{A}\mathbf{A}^T F_1} = \sqrt{F_1^T\lambda_1 F_1} = \sqrt{\lambda_1}.$$

Notice that λ_1 is the maximum of all λ_i. The above equation also points out that $\sqrt{\lambda_i}$ is the rms of the projection amplitude of all of the columns in $\mathbf{A}$ on F_i. The total variance of $\mathbf{A}$ is the sum of all λ_i, i.e., $||\mathbf{A}||^2 = \sum_1^M \lambda_i$. The sorted $\lambda_i(i = 1, 2, \ldots, M)$ usually decrease very rapidly, so that the first few EOFs contain most of the variance of $\mathbf{A}$.

In this study, the velocity vectors are represented as $u + \mathbf{i}v$, so bfA is complex. Thus the covariance matrix is complex, as well as the EOFs, but the eigenvalues are real since the covariance matrix remains symmetric. We conventionally formulate the EOFs so that they are real at the surface.

Strictly speaking, the EOF analysis applied here is a complex EOF analysis (Preisendorfer 1988). However, as will be seen, the first complex EOF, which is our main focus, is dominated by the real component. If we apply the EOF analysis to east-west or north-south components separately, the resulting first EOFs are very similar .

In EOF analysis, the primary concern is the so-called effective degeneracy (North *et al* 1982), which stems from the degeneracy of the eigenvectors of a matrix. If two eigenvalues of the matrix are identical, their corresponding eigenvectors are not unique, i.e., any linear combination of these two eigenvectors will also be an eigenvector with the same eigenvalue. In EOF analysis, even if two eigenvalues are separate from each other but only by a small amount, the separation could be due to noise. In such a situation, the corresponding EOFs are not physically distinct and thus are misleading. As could be expected, the greater the separation of the estimated λ_i, the greater the significance of the EOFs (the less chance of effective degeneracy).

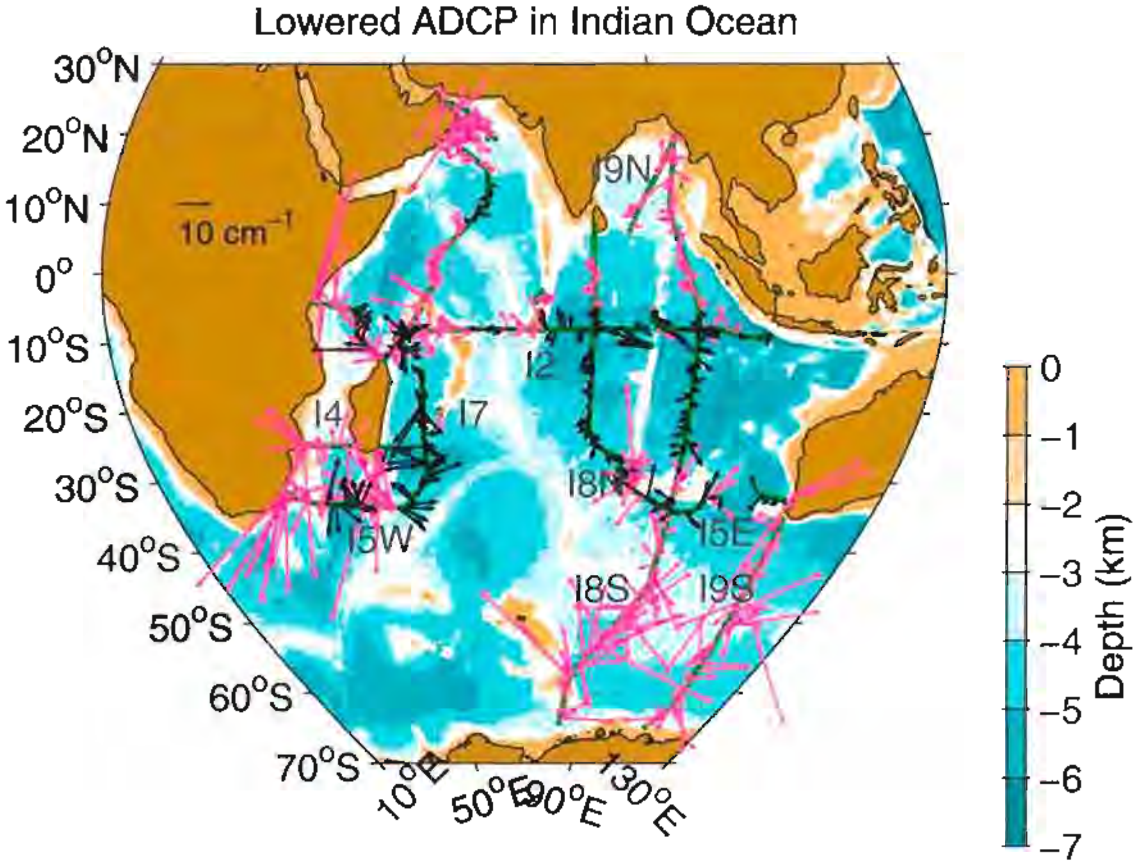


Figure 2.1: Indian Ocean topography and locations of about 1000 LADCP profiles (both black and purple vectors) from the WOCE Indian Ocean Hydrographic Program. Green dots are CTD stations. Labels are WOCE designations. One non-WOCE cruise by Doug Wilson, mostly a repeat of I7 and I5W, is also used in this study, but not shown here. The black vectors represent the mid-latitude subset, with 366 profiles deeper than 4000 m and in the 5° to 35° latitude range. The plot is a modification of a similar plot by Kathy Donohue.

CHAPTER 3

LADCP PROFILES AND LINKS TO MESOSCALE MOTION

The LADCP profiles are generally in the lowest modes (having large vertical scale). This is demonstrated by looking at the depth-average vectors, the profiles themselves and sectional contours of them, and formally by normal mode decomposition.

3.1 Glimpses of LADCP Profiles

Depth-average (barotropic component)

The depth-averages of the LADCP profiles, i.e., the barotropic components, are shown in figure 2.1. The largest barotropic velocities (~30 cm s^{-1}) mostly fall within the western boundary current: the Somali Current near the equator, the Agulhas Current around the southern tip of the African continent, the East Madagascar Current along the east coast of Madagascar and the Antarctic Circumpolar Current in the Southern Ocean. Large barotropic velocities were also located over shallow waters, such as the mouth of the Persian Gulf and the west coast of Australia. These large velocities are presumably dominated by tides or coastal currents.

In the mid-latitudes (black vectors in figure 2.1), the barotropic components are mostly 5-10 cm s^{-1}. They are a substantial part of the total velocities. Another feature, which will be tested by the lag-correlation along cruise tracks in the next section, concerns the horizontal scale. In figure 2.1, it is easy to spot the two adjacent vectors with similar magnitudes and directions. This is not true, however, for three or more vectors. Considering that the LADCP stations are typically 55 km apart, we can infer that the horizontal correlation scale in the LADCP barotropic velocities is

~50 km.

Velocity profiles

Figure 3.1 shows 40 individual LADCP profiles (E-W component only), which are randomly drawn from 366 profiles (their depth-averages are the black vectors in figure 2.1). The general impression is that small-scale wiggles and large-scale wobbles exist simultaneously. Looking at them more closely, one finds that the wiggles are not around the zero-velocity axis or any other fixed value, but around the wobbles, and that the amplitudes of the wobbles are significantly larger than those of the wiggles.

We shall call the wobbles low-mode motion, and the wiggles high-mode motion. The "low-mode" refers to the barotropic and first few baroclinic modes, while the "high-mode" refers to higher baroclinic modes. We were unable to match any of the wobble to any single low normal mode. Wobbles can be approximated by a combination of a few low normal modes, with the combination varying from one profile to another. When we put all combinations together, however, a certain pattern exists, which will be shown in the context of kinetic energy partition within normal modes.

Cross-basin sections

We present two sections, one along about 30°S (figure 3.2) and another along about 95°E (figure 3.3). As the contours show, the strong velocity cores usually extend more than one kilometer vertically. This is consistent with the individual velocity profiles (figure 3.1). Horizontally, the velocity cores extend about a degree along the cruise tracks.

Another observation, which is less straightforward but has important implication dynamically, is that we do not see any zero-crossings near 1000 m; we would expect to see them if the first baroclinic mode were dominating the synoptic velocity field. We see many instances where the velocity is in the same direction from surface to bottom, with surface intensification. The vertical structure will be quantified by EOF analysis in Chapter 4.

3.2 Normal Mode Decomposition

Normal modes are derived from simplified dynamics, but whether the simplified dynamics are applicable here has not been established. Normal model decomposition is therefore performed merely as a way of describing the observations statistically. For each station, we first project each component of the LADCP profile (u or v) onto normal modes ($P_n(z), n = 0, 1, 2, \cdots$). The resulting u_n and v_n then define the nth mode amplitude as $\sqrt{u_n^2 + v_n^2}$ ($n = 0, 1, 2, \cdots$).

cross-basin sections

The mode amplitudes for six latitudinal sections are shown in Figure 3.4 in the form of contours. Some features are as expected:

- The kinetic energy is mostly in the low modes;
- Near the equator, the kinetic energy usually spreads into modes higher than the third mode;
- In high latitudes (Southern Ocean, the first two panels), the barotropic and first baroclinic modes are dominant.

Other features, which may not be obvious, but are worth mentioning:

- The contours are not smooth. The amplitudes, however, are not jumping from one station to another, but rather are varying on a scale of about one degree latitude. This was also suggested by the sectional contours (figures 3.2 and 3.3).
- In the Southern Ocean, the barotropic mode significantly exceeds other modes. Away from the Southern Ocean, there are many instances where higher modes exceed lower modes.
- In the last two panels, we see a feature near 10°S that seems directly related to a major topographic feature, the Mascarene Plateau (see figure 2.1).

The mode amplitudes of one zonal section are shown in figure 3.5, along with the Smith-Sandwell topography. High modes are enhanced near major topographic features. Another noticeable feature is that the western basin is more energetic than the eastern basin. we might expect this kind of asymmetry due to the boundary current (the East Madagascar current) and complicated topography in the western basin.

Kinetic Energy Partition (KEP)

We focus on the barotropic mode and first ten baroclinic modes, since they account for most of the kinetic energy. The kinetic energy partition is defined as the percentage of total energy in each mode, i.e.

$$KEP_n = \frac{(u_n^2 + v_n^2)/2}{\sum_{m=0}^{\infty}(u_m^2 + v_m^2)/2} \times 100\% \quad (n = 0, 1, 2, \cdots).$$

From latitudinal sections (figure 3.4), we see that there is more energy in higher modes near the equator and that the barotropic mode is only dominant at high latitudes. To emphasize the latitudinal dependency, we binned the LADCP profiles for each 5 degrees latitude. Figure 3.6 (top panel) shows the mean KEP of all 11 modes for each 5-degree bin. The bottom panel shows the mean and standard deviation of the barotropic and first baroclinic modes along with a listing of the number of LADCP profiles per bin.

Figure 3.6 clearly demonstrates that

- The low modes are dominant. Even near the equator, the fifth mode has less than 10% of the kinetic energy;

- The barotropic mode dominates in high latitudes, while the kinetic energy is marginally dominated by the second baroclinic mode near the equator.

- In the mid-latitudes, we see the approximate equipartition between the barotropic and first baroclinic modes; each contains about 25% of the total variance, and the second baroclinic mode contains about 12%. This is similar to Wunsch (1997).

3.3 Comparison with CTD Geostrophy

Sectional comparison

We start with sectional contours of the LADCP profiles and CTD geostrophy. The geostrophy is referenced to the barotropic component of the LADCP profiles: the depth-average of the geostrophic velocity is equal to the depth-average of the average of the two LADCP profiles at corresponding stations. Recall that the barotropic component tends to be less noisy than the baroclinic components (King *et al* 2001). In figures 3.7 and 3.8, the cross-track LADCP velocities are compared with the corresponding geostrophy. The east-west and north-south LADCP velocities along these two sections were shown in figures 3.2 and 3.3. Since the cruise may not have gone along the lines of latitude or longitude, the cross-track velocity could differ from u or v in figures 3.2 and 3.3.

Generally speaking, the LADCP velocity and geostrophy are comparable. The LADCP profiles, however, have more small-scale structures. The geostrophic calculation inherently averages between stations, so that the high mode motions, usually with small horizontal scales, are filtered out. In figure 3.8, the geostrophy near the equator (e.g. less than 7°) is too noisy to be meaningful.

In the east part of I5E (figure 3.7), near the bottom ~1000 m, the geostrophy and LADCP velocity are different. This also seems to happen along 95°E (figure 3.8) in several locations. The reason is not clear. The noise in the LADCP profiles down there could have been enhanced because of weak scattering.

Comparison between LADCP velocity and CTD geostrophy

The criteria for choosing station pairs for comparison are as follows:

- stations must be located at least 10 degrees away from the equator. A small Coriolis parameter (f) near the equator will greatly exaggerate the noise in the dynamic height, so that the geostrophic calculation from one snapshot is unreliable.

- Distance between station pairs is less than 80 km but larger than 40 km. If the distance is too large, the LADCP profile pairs may miss peaks between stations while the geostrophic velocity does not. If the distance too small, the noise in the dynamic height will be exaggerated, similar to the situation with a small f.

- The common water depth of a station pair is larger than 2 km and the difference of the water depths between two stations is less than 1 km. This criterion excludes station pairs in shallow water and near steep topographic slopes, where internal waves, including internal tides, are usually significant.

The difference between the LADCP velocity and geostrophy could be due to: 1) measurement errors in both; 2) high frequency components (e.g. internal waves) and other non-geostrophic terms; and 3) a mismatch of the point-average of the LADCP profiles and the integral-average in the CTD geostrophy.

Figure 3.9 shows the first mode amplitudes of the LADCP profiles and geostrophy, station-pair by station-pair. The two basically follow each other. The favorable comparison indicates that the first baroclinic mode of the LADCP velocity is approximately in geostrophic balance.

There are, however, differences. In quite a few instances, we see that the geostrophy is larger than the LADCP at one station-pair, but smaller at an adjacent station-pair (e.g., see station pairs 123-124 and 147-148). This feature could be related to noise in the dynamic height: the noise in the dynamic height will bias higher for one station-pair and lower for the adjacent pair. The difference between the LADCP and geostrophy is generally 1-2 cm s^{-1}, but can be as large as $\sim$5 cm s^{-1}.

Figure 3.10 shows the correlation between the LADCP profiles and CTD geostrophy as a function of normal modes. We did not find a latitudinal dependence, but we see that the correlation coefficient curves are noisy (e.g. the green line) when the sample size is not large enough. The first and second modes have strong correlations, while the next two modes show weak correlations. After the fourth mode, there are no correlations. One explanation could be that the geostrophic estimates filter out high modes. This is supported by the mode amplitude variance in figure 3.11.

Figure 3.11 shows that the variance in the high modes of the geostrophy is weak, likely due to the integral-average inherent in the geostrophy. The drop-off in variance at higher modes is consistent with the smoothness of contours in the lower panels of figures 3.7 and 3.8. The first baroclinic mode variances for the LADCP and geostrophy profiles are 18.7 and 16.1 (cm s^{-1})^2 respectively. Thus, though it is not significant statistically, the variance in LADCP is ~16% greater than in geostrophy.

Breaking the data latitudinally (figure 3.12), we find a tendency that the higher latitude the more difference between the variances in the lower modes of the LADCP and geostrophy. Toward the equator, uncertainty in the geostrophic estimate due to noise in the dynamical height is enhanced by a smaller Coriolis parameter. On the other hand, the scale of the motion gets smaller toward high latitudes, thus the integral-average in the geostrophy may reduce the variance there.

In summary, the low modes of the cross-track LADCP velocity and CTD geostrophy are comparable. It follows that the low modes of the LADCP velocity are dominantly governed by geostrophic dynamics.

3.4 Spatial and Temporal Information in LADCP Profiles

We extract spatial and temporal information from the subset of 366 profiles, since most of the statistical analysis in the next section is performed on this subset.

Two types of lag correlation coefficients are calculated: one is from the profiles of the same cruises (upper panel of figure 3.13) and another from those of different cruises (lower panel of figure 3.13). The former should consist mainly of spatial information, if the LADCP velocity is dominantly the mesoscale motion, while the latter may consist of a mixture of spatial and temporal information. The detailed steps to calculate lag correlation coefficients are in Appendix A.

As shown in figure 2.1, the LADCP profiles were mainly sampled along longitudinal/latitudinal lines across the Indian Ocean basin within tens of days and with station intervals of $\sim$55 km. If we presume that the LADCP profiles are dominated by mesoscale motions with spatial scale 50 km, then the LADCP profiles just barely resolve the horizontal structure of the velocity field. This is the case suggested by figure 3.13a. The zero-crossing is at about 100 km, consistent with the scale of the mesoscale motion. Therefore, since most of the stations were 55 km apart, we conclude that we likely have 183 degrees of freedom for the 366 LADCP profiles used in the EOF analysis. This is critical. Existing moored current meter records have much fewer degrees of freedom for mesoscale motions, given their time scales are about 100 days.

From figure 3.13a, we see that the barotropic and first baroclinic modes have almost identical horizontal correlation scales. As will be discussed in the next section, this characteristic directly relates to the fact that the barotropic and first baroclinic modes are correlated.

Adjacent profiles from the same cruise should be viewed as a snapshot of the mesoscale motion, since they were sampled within a day. There are, however, profiles that are close to each other geographically, but from different cruises. The cruises were separated by 4 months to 11 months. The autocorrelation function for these profiles is shown in figure 3.13b. There is no significant correlation, even for the shortest distance. Physically, it shows that the current's time scale is shorter than 11 months – the longest separation for all involved cruises.

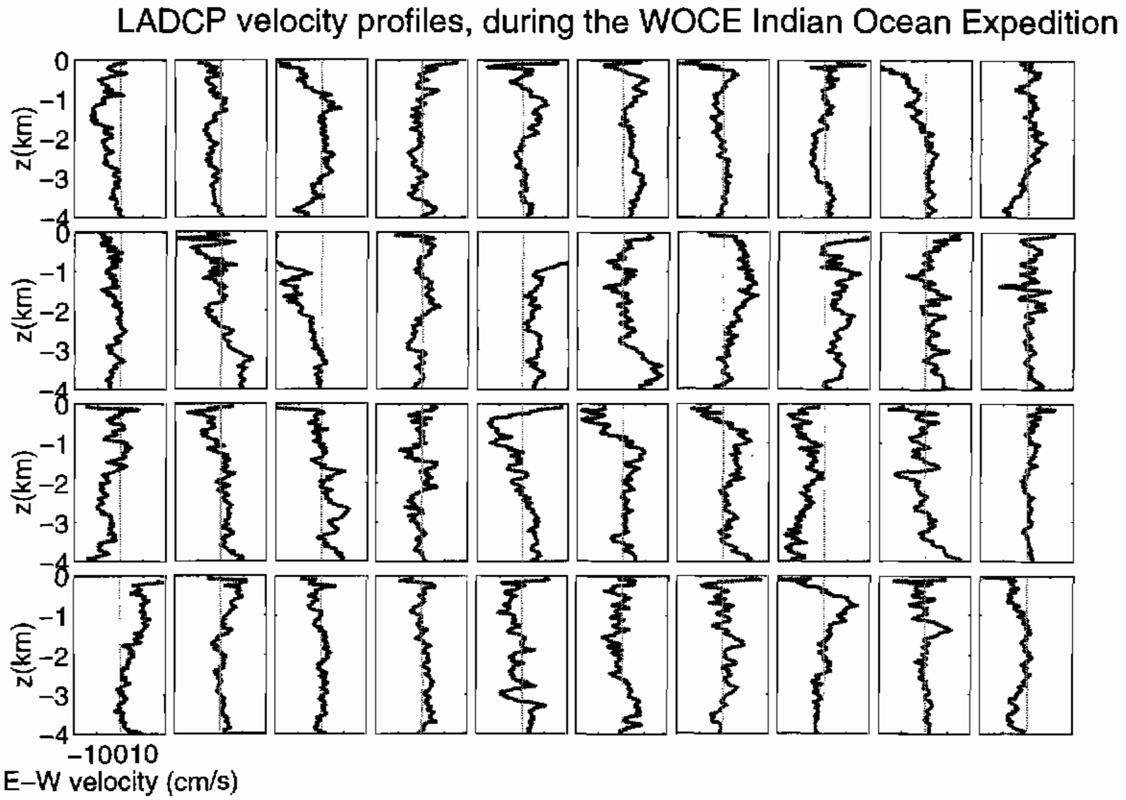


Figure 3.1: LADCP velocity profiles randomly chosen from the 366 subset indicated by the black vectors in figure 2.1. Only the E-W component is plotted. The profiles are truncated at 4000 m. In each panel, the thin straight line represents zero velocity. Each tick on the x-axis is 10 cm s^{-1} and the total range is from -20 cm s^{-1} to 20 cm s^{-1}. The profiles are truncated if the velocities are out of this range.

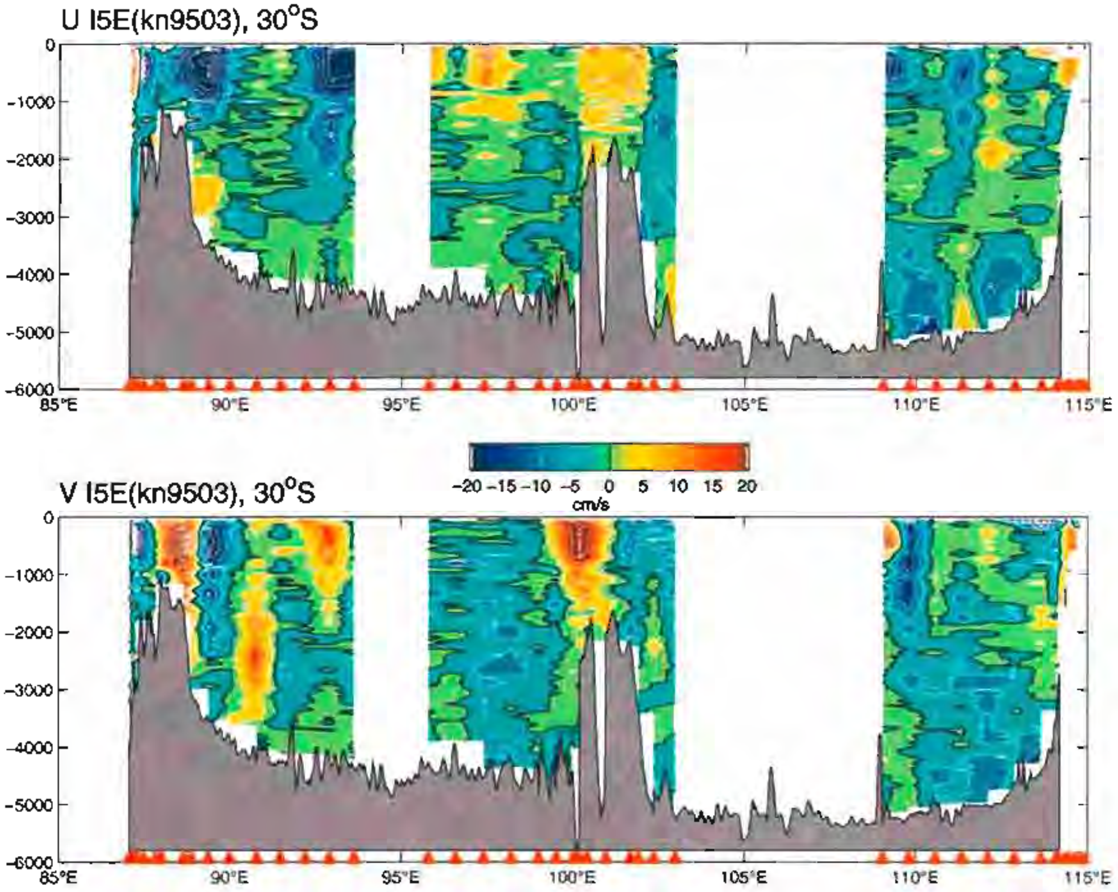


Figure 3.2: East-west (u, upper panel) and north-south (v, lower panel) currents from the LADCP profiles. This section is along 30°S in the eastern Indian Ocean (see figure 2.1). The two panels use the same color and contour schemes. As indicated in the color bar in the middle, the black contour is zero velocity and the white ones are ±5cm s^{-1}, ±10cm s^{-1}, $\cdots$. The red triangles indicate locations of the profiles. The topography is extracted from Smith-Sandwell topography. In the panel title, "I5E" is the WOCE designation of the cruise, and "kn9503" is the dataset name, which means the profiles were collected during march, 1995 on R/V Knorr.

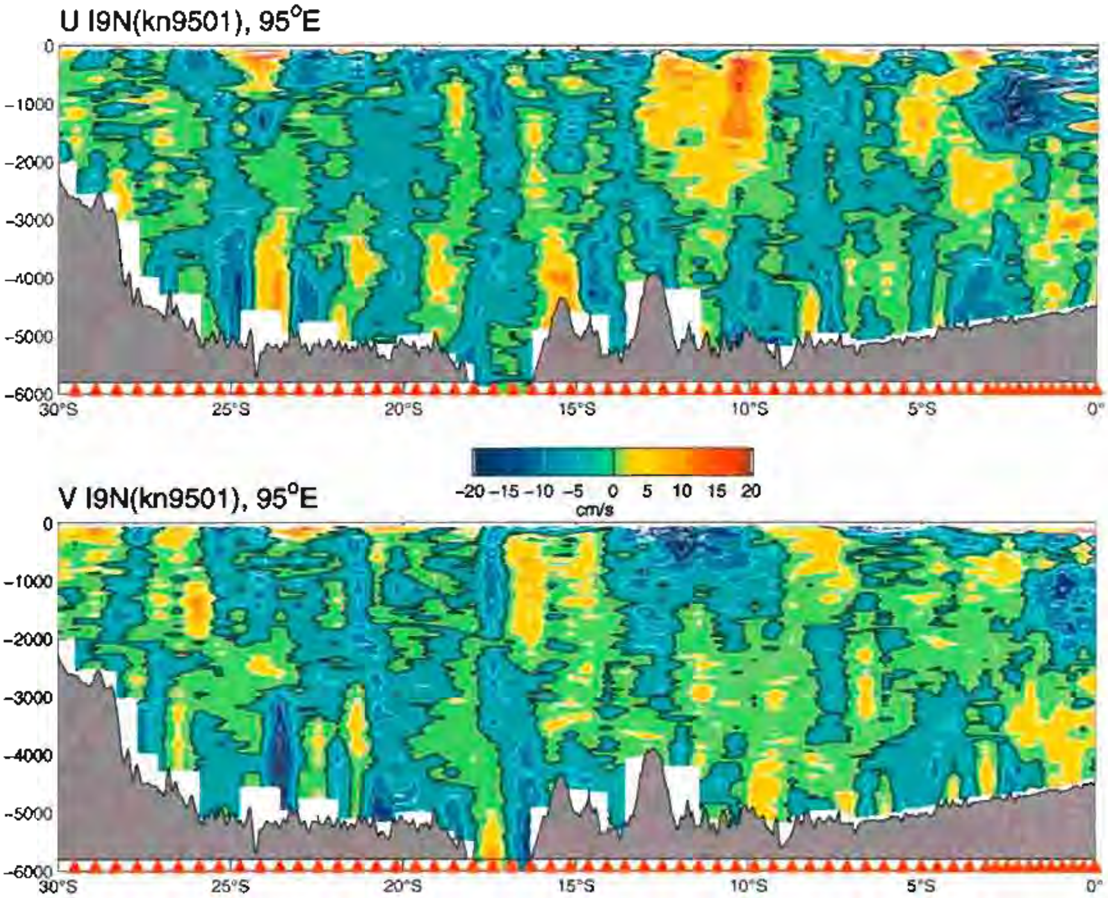


Figure 3.3: Same as figure 3.2, but this section is along 95°E in the eastern Indian Ocean (see figure 2.1).

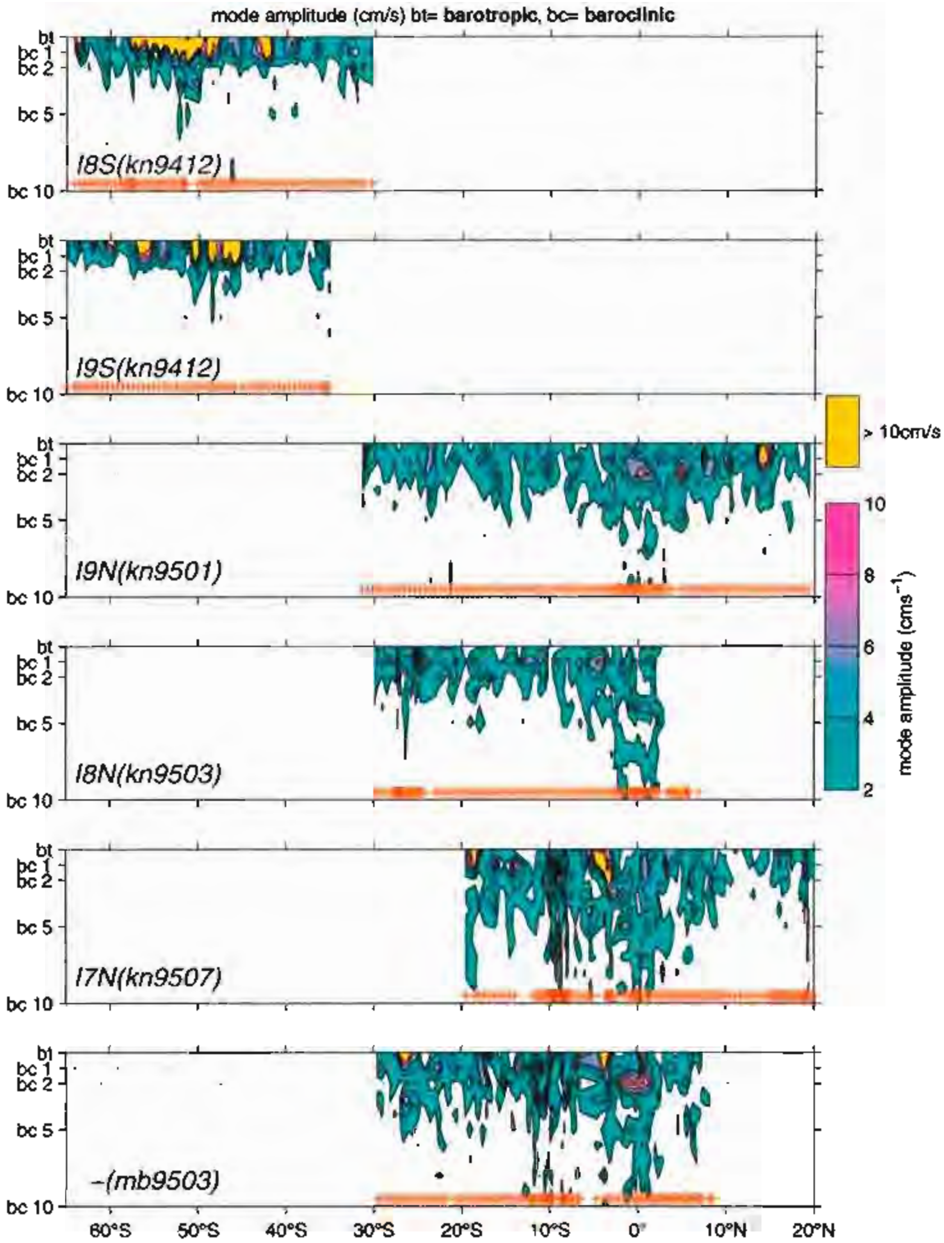


Figure 3.4: Mode amplitude variations with latitude. One panel represents one cruise. For the meaning of the labeling at the lower-left corner, refer to the caption in figure 3.2. The data for the last panel is non-WOCE, thus there is no WOCE designation; "mb" stands for R/V Malcolm Baldrige. The vertical coordinate is unusual: the top grid is for the barotropic mode, the second grid is for the first baroclinic mode, the third one is for the second baroclinic mode, and so on. The red crosses in the panels indicate locations of the LADCP profiles. All contours use the same color scale as shown at the right. The first contour is of 2 cm s^{-1}, and increases 2 cm s^{-1} for each contour until reaching 10 cm s^{-1} (the highest contour). Amplitudes above 10 cm s^{-1} are highlighted in yellow.

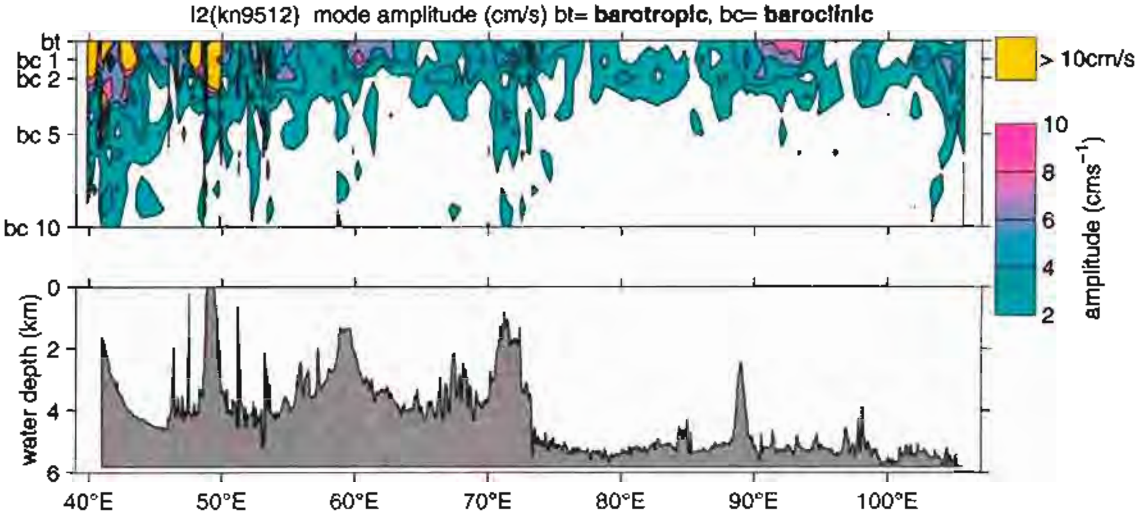


Figure 3.5: Same as figure 3.4, but for a zonal section. The cruise information is listed in the title. Also shown is the Smith-Sandwell topography along the cruise track.

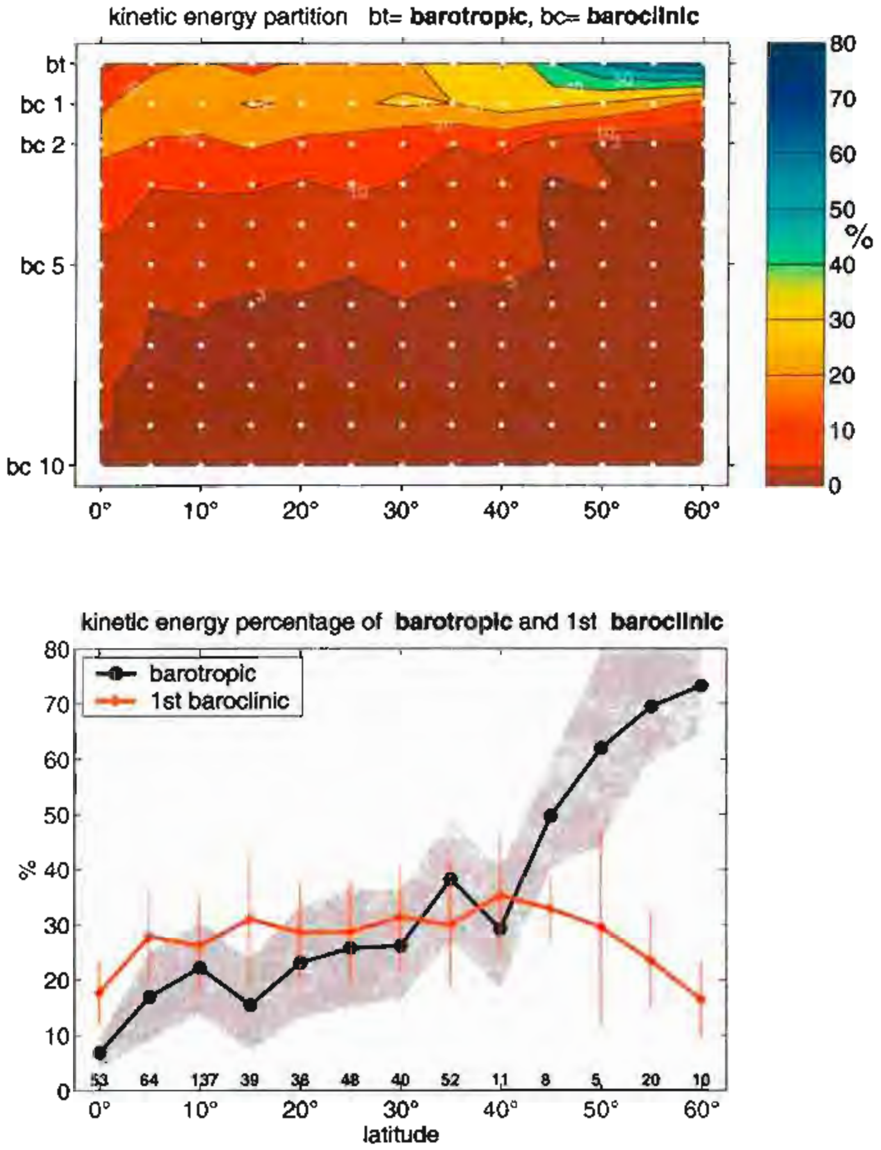


Figure 3.6: Kinetic energy partition (top). The method used to calculate the percentage is detailed in the text. The vertical coordinate is the same as in figure 3.4. The white dots represent discrete data points. With respect to the latitude, data is binned around the white dots ±2.5°. The kinetic energy percentage for the barotropic and first baroclinic modes (bottom). The values are the same as those in the top panel. The shaded area and error bars represent one standard deviation. The numbers near the abscissa are the number of the LADCP profiles within the corresponding bin.

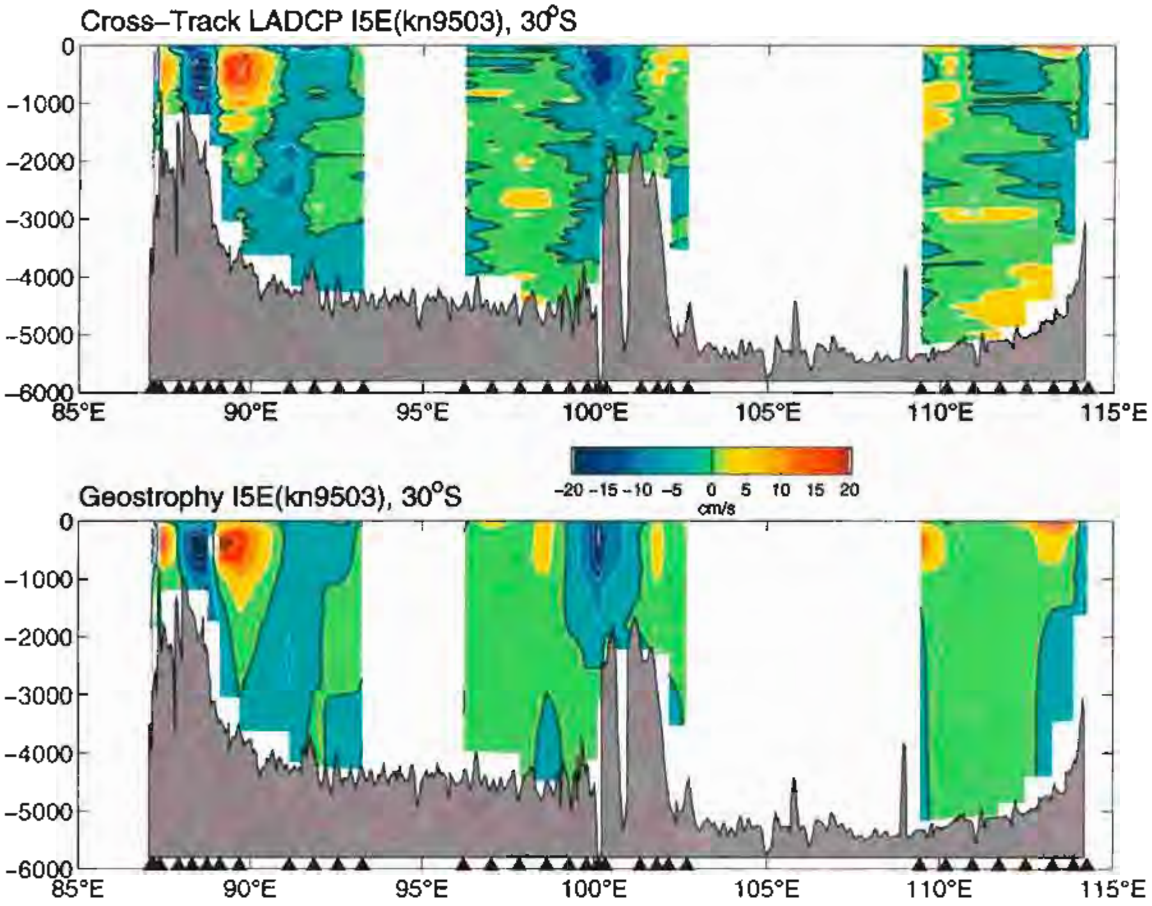


Figure 3.7: The cross-track LADCP velocities along 30°S (upper panel). The cruise is the same as the one in figure 3.2. Positive cross-track velocity is approximately southward. The LADCP profiles have been averaged between stations in accordance with the geostrophy. The topography in both panels is Smith-Sandwell topography. Both panels use the same color and contour scheme. The geostrophy calculated from simultaneous temperature and salinity profiles (lower panel). It is referenced to the depth-averaged LADCP velocity (see the text).

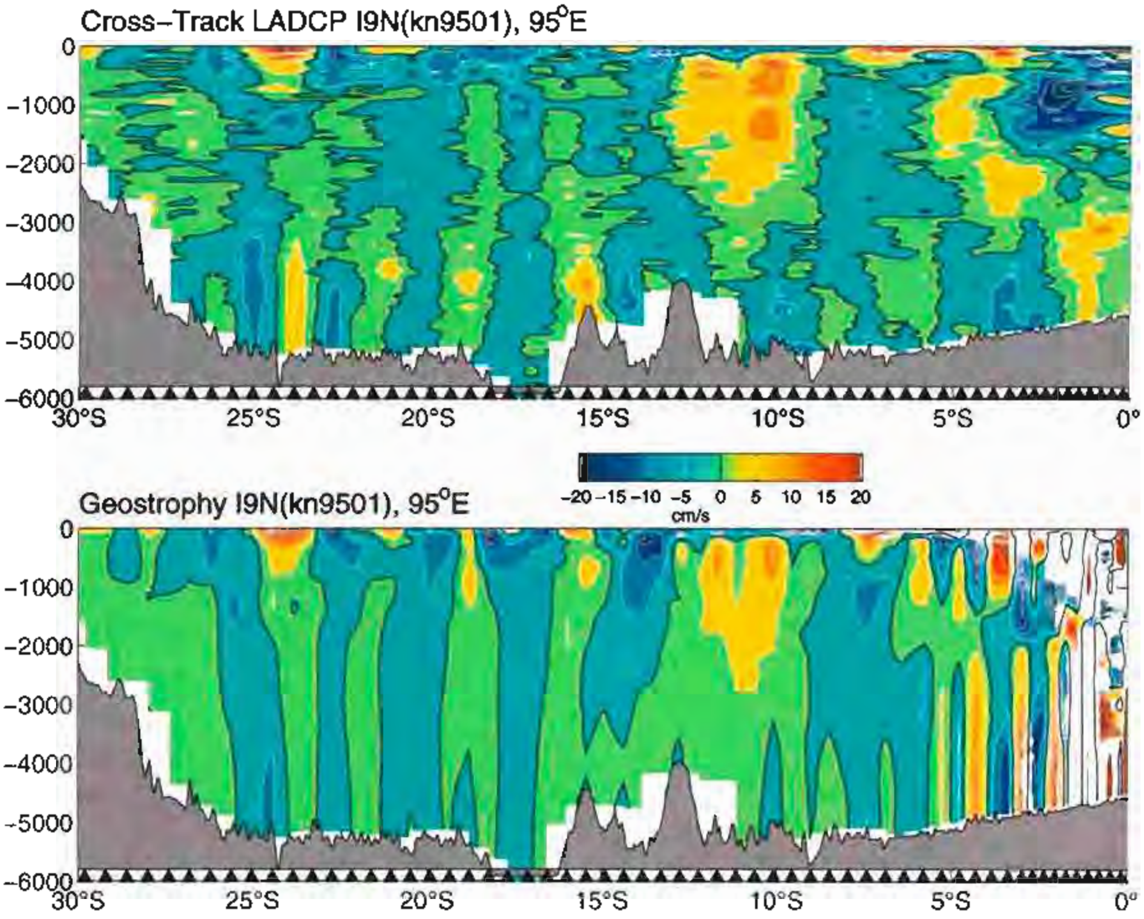


Figure 3.8: Same as figure 3.7 but for cruise I9N. Positive cross-track velocity is approximately eastward, similar to the *u* component in figure 3.3.

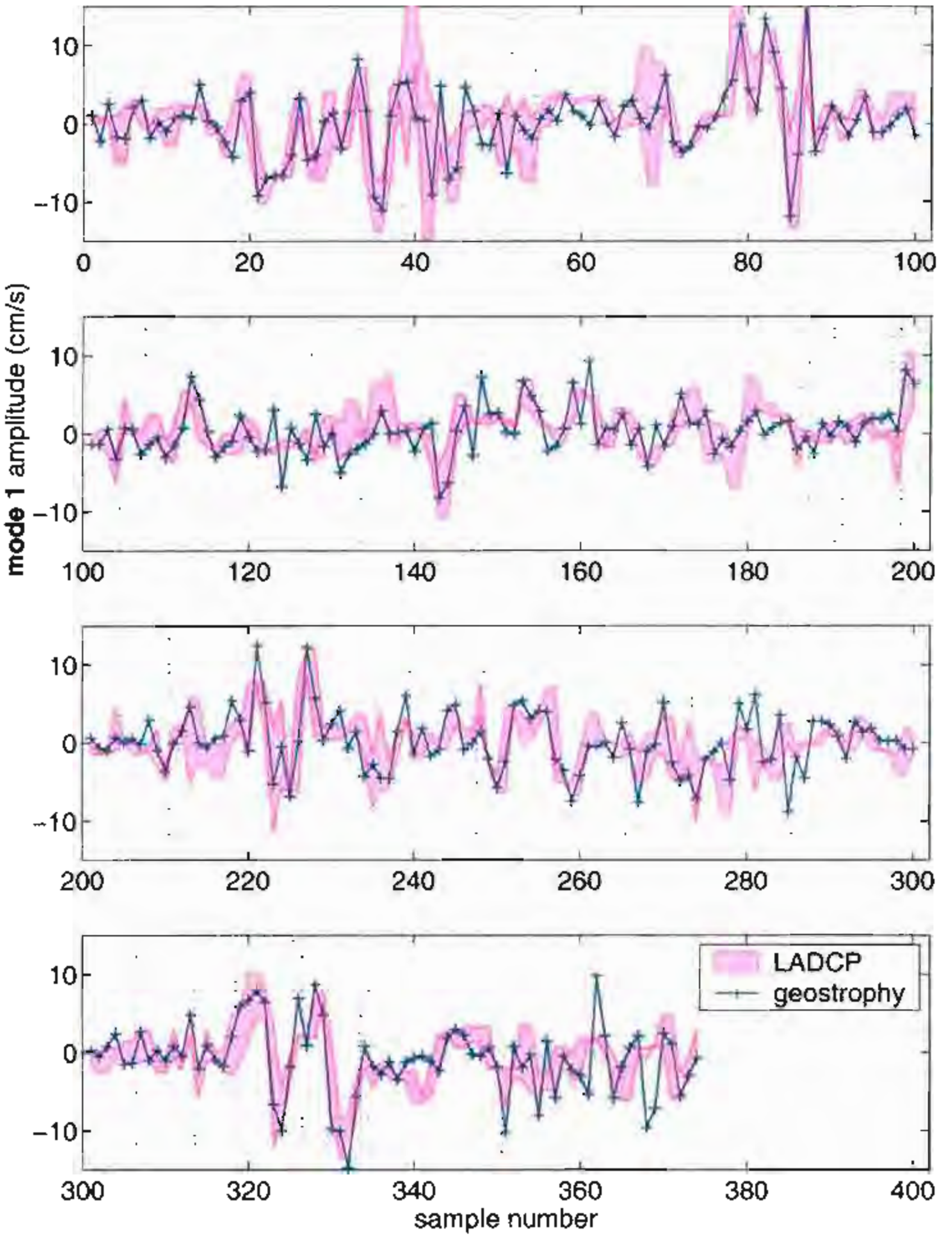


Figure 3.9: The amplitudes of the first baroclinic modes of the cross-track LADCP velocity (purple shaded area) and geostrophy (cross lines). The sample numbers (abscissa) are given arbitrarily, but the adjacent station-pairs are plotted adjacently. The lower (upper) bound of the purple shaded area is the smaller (larger) one of the first baroclinic mode amplitudes of the cross-track LADCP velocity at the two relevant CTD stations. Grey shaded or unshaded areas indicates that the station pairs are of the same cruise. The dotted lines indicate where the two adjacent station pairs are more then 80 km apart, so that the two station pairs do not share a common station. The criteria for choosing the station pairs are listed in the text.

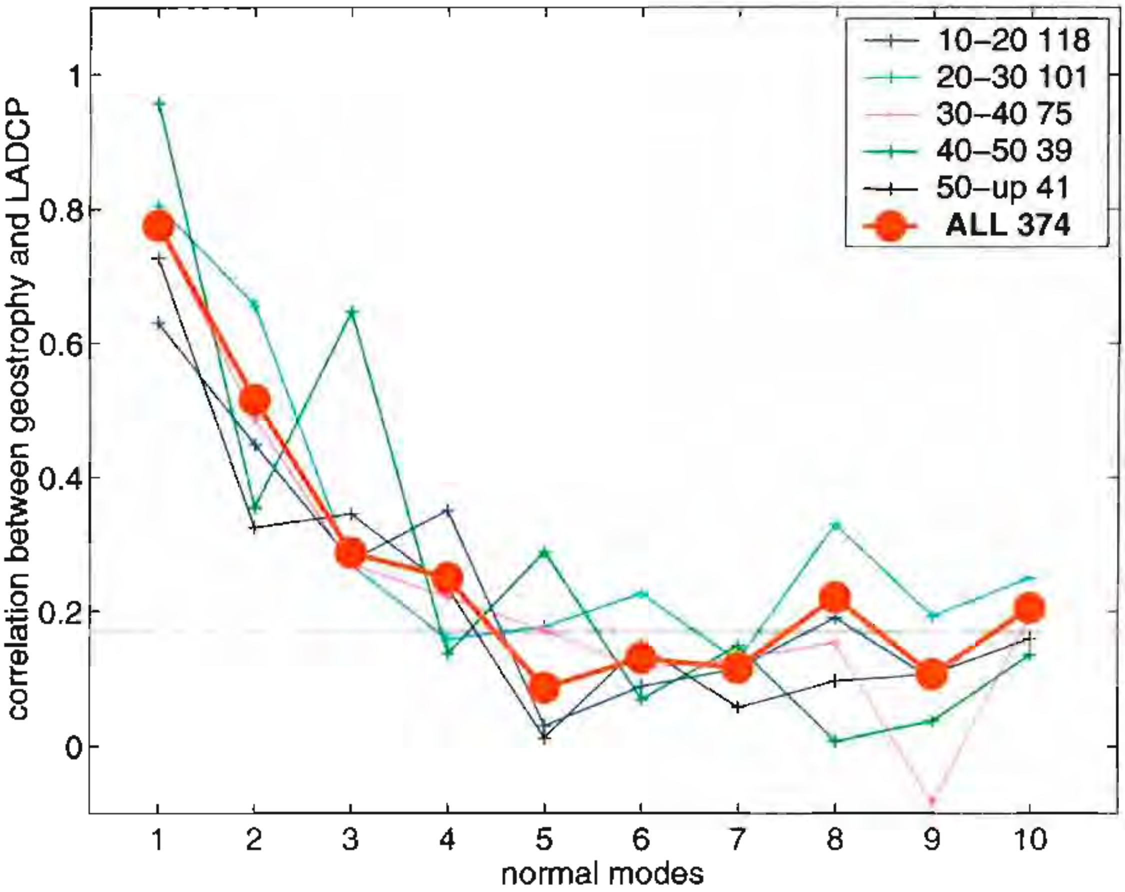


Figure 3.10: Correlation coefficients between the first ten baroclinic modes of the cross-track LADCP velocity and geostrophy. The station pairs are the same as those in figure 3.9. The thick red line is calculated from all available data. The gray straight line indicates no correlation level at 99%, by assuming the degrees of freedom equals 1/2 of the samples (187). We also break samples into five latitudinal sub-regions; the latitude range is shown in the first column in the legend. The second column in the legend shows the sample numbers within that particular latitude range.

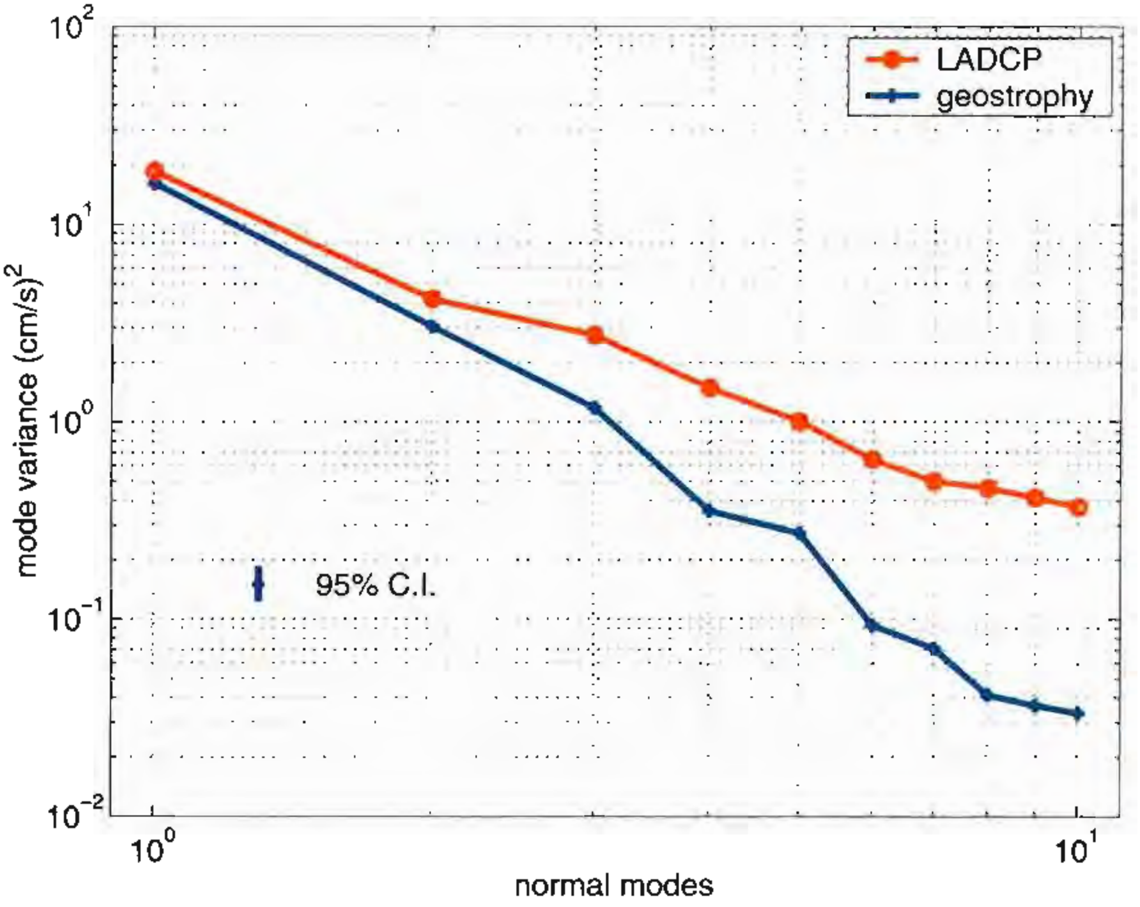


Figure 3.11: Mode amplitude variance of the first ten baroclinic modes of the cross-track LADCP velocity and geostrophy. The station pairs are the same as those in figure 3.9. The log scale is to mimic a variance-preserving wave number spectrum. The 95% confidence interval is based on a χ^2 distribution with 187 degrees of freedom (half of the station pairs in figure 3.9).

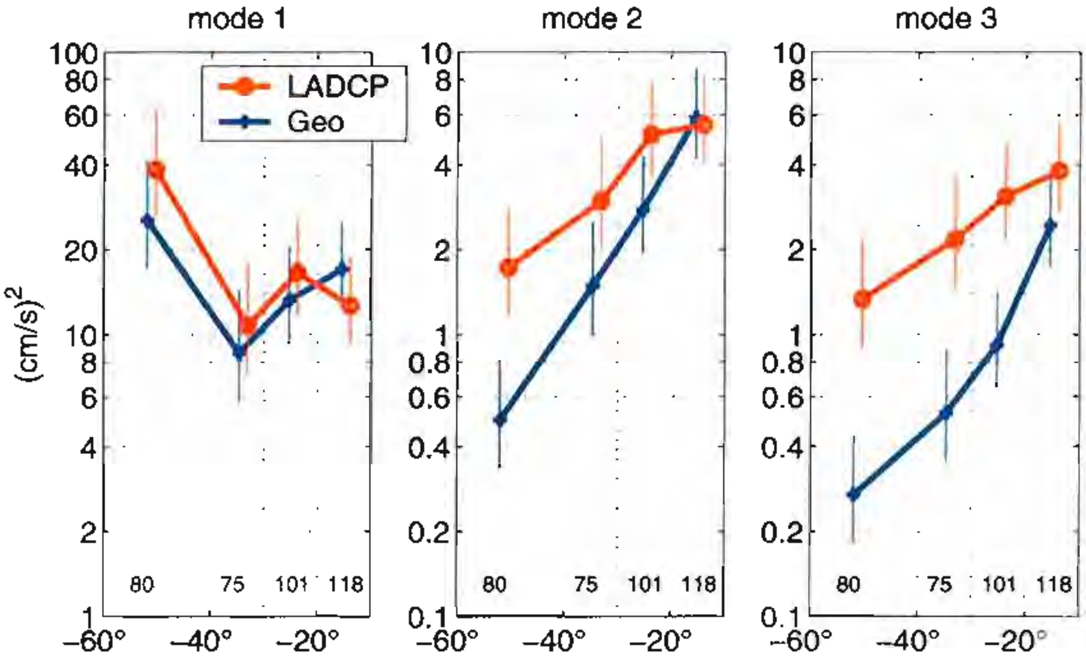


Figure 3.12: The latitudinal dependence of the mode amplitude variance of the first three baroclinic modes of the cross-track LADCP velocity and geostrophy. Four latitudinal groups are indicated by the grid lines: 10-20, 20-30, 30-40, and south of 40. The numbers of the station pairs of each group are listed in each panel. Part of the samples in the latitude range (-20° to -10°) is actually from the north hemisphere (see figure 2.1). Notice that different scales in the vertical axes are used in different panels. The vertical lines indicate the 95% confidence intervals based on a χ^2 distribution, whose degrees of freedom are half of the station pairs listed for each latitudinal group.

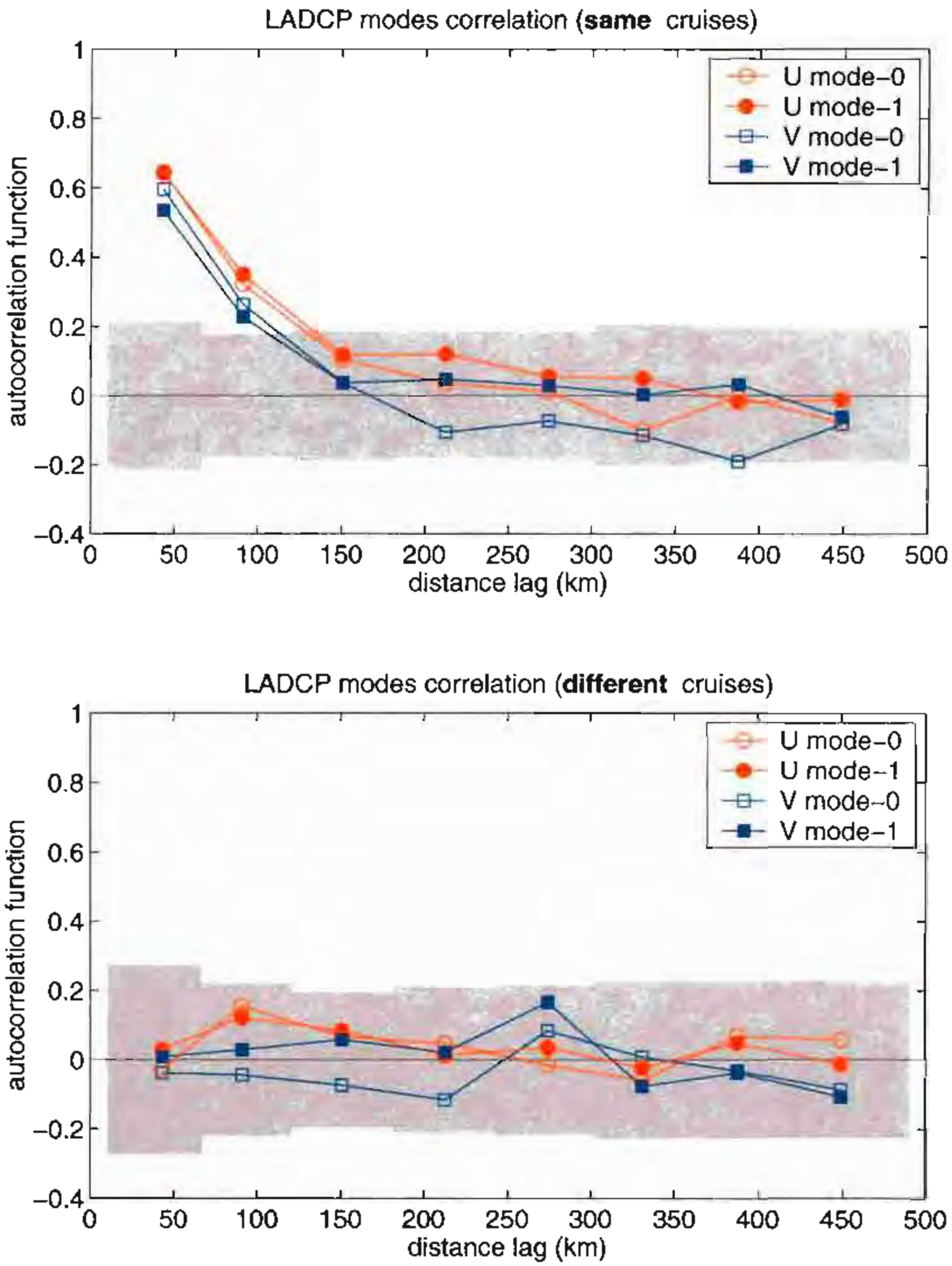


Figure 3.13: LADCP modes auto-correlations. The method used to calculate the correlation coefficients is detailed in Appendix A. The shading indicates the 99% no correlation level; the degrees of freedom are one half of the LADCP profiles used in calculating the correlation coefficients (usually more than 100).

CHAPTER 4

DOMINANT VERTICAL STRUCTURE IN THE LADCP PROFILES

The vertical structures of the mesoscale currents are different near the equator and in mid-latitudes and high latitudes. This is clear in figures 3.4 and 3.6. Figures 3.4 and 3.6 suggest that variations are gradual between 5 and 35 degrees latitude. Thus to produce statistics that are significant, we group 366 profiles in this region (the black vectors in figure 2.1) under "mid-latitude".

In previous sections, normal modes have been extensively employed to describe the vertical structure of the profiles. This does not imply, however, that the corresponding dynamics, i.e., the linear, flat-bottom and no mean flow dynamics, is applicable in this study. The barotropic and first baroclinic modes are the two most energetic modes in the LADCP profiles (figure 3.6). From contours in figures 3.2 and 3.3, however, we do not see much of the barotropic mode (constant over depth) or the first baroclinic mode (zero-crossing around 1000 m). Instead, there are many instances where the velocity profiles are unidirectional and surface-intensified. This is quantified by studying the relationship between the barotropic and first baroclinic modes and conducting the EOF analysis.

4.1 Correlation between Barotropic and the First Baroclinic Modes

Looking at the amplitudes of the barotropic and first baroclinic modes, station by station (component u in figure 4.1 and v in figure 4.2), the two curves do not match as well as those, for example, in figure 3.9, but they do show some correspondence. The correlation coefficients are 0.43 for u and 0.25 for v. We do not know why there is such a difference in u and v. For reference, the 99% no correlation level is 0.17, assuming one half of the 366 LADCP profiles are independent.

Figure 4.3 shows the barotropic and first baroclinic modes of the LADCP profiles as vectors. Notice the scales for the two panels are the same. First we see that the two modes have similar magnitudes, then we see that vectors from different modes are more or less parallel to one another. This relates to the fact that the two modes are correlated and is further quantified by the simple statistics in figure 4.4.

Figure 4.4 shows the distribution of the angles between the barotropic and first baroclinic modes. The lower panel shows that the two modes most likely align with each other within $\pm 45^o$. The upper panel shows an amplitude-dependency of such alignment. When $\sqrt{a_0 a_1}$ is large, the two modes more likely to align with each other. Among the 205 profiles with $\sqrt{a_0 a_1} > 3$, 93 profiles (about $\sim 45\%$) align within $\pm 45^o$.

4.2 EOF Analysis

Since the barotropic and first baroclinic modes in the mid-latitudes are most energetic and positively correlated, the actual vertical structure seen in the observation will tend to be unidirectional and surface-intensified. The EOF analysis confirms this speculation (figure 4.5b). Since we have enough degrees of freedom (more than 100) in the mid-latitudes (figure 4.5b), the first EOF is believed to be the dominant inherent vertical structure in the LADCP profiles, and the so-called effective degeneracy is unlikely to occur (North *et al.* 1982).

To put the mid-latitude results in the context of a larger latitudinal range, figure 4.5 shows the results from the profiles near the equator and in the Southern Ocean. Despite the concern about the statistical significance with fewer LADCP profiles, we see that the first EOF near the equator (figure 4.5a) has a lot of vertical structure that varies both in amplitude and in direction. The first EOF in the Southern Ocean (figure 4.5c) is unidirectional, similar to that in the mid-latitudes, but more barotropic; it does not tend to zero at the bottom.

The variation from the re-sampling is noticeably large for the Equator subset (figure 4.5a). The 94 LADCP profiles have a wide variation geographically from one cruise to another cruise (*ref*. figure 3.4). In figure 4.5c, we do not see much variation from the re-sampling, although it has the least number of profiles.

We tried to break the mid-latitude 366 profiles into two sets: south and north of 15°S. Both first EOFs are similar to that in Figure 4.5b, while the south one (in the higher latitude) is slightly more barotropic. The difference is much less dramatic than those between (a,b,c) of Figure 4.5. Such latitudinal dependency has been similarly seen in the kinetic energy partition (figure 3.6).

Figure 4.5d shows the variances of the EOFs. In the mid-latitudes, the first EOF accounts for about 42% of the variance, and the variance of the first EOF is almost double that of the second EOF mode. This is another indication that the effective degeneracy is unlikely to occur. As already mentioned, in the Southern Ocean, the LADCP profiles are overwhelmingly dominated by the first EOF.

Figure 4.6 shows how representative the first EOF is in each LADCP profile in the mid-latitudes. Neither a topographic enhancement nor a regional feature could be clearly identified.

Figure 4.7 shows that the stronger the currents the more the profiles could be represented by the first EOF. The strength of a LADCP profiles is defined by its depth-averaged velocity squared.

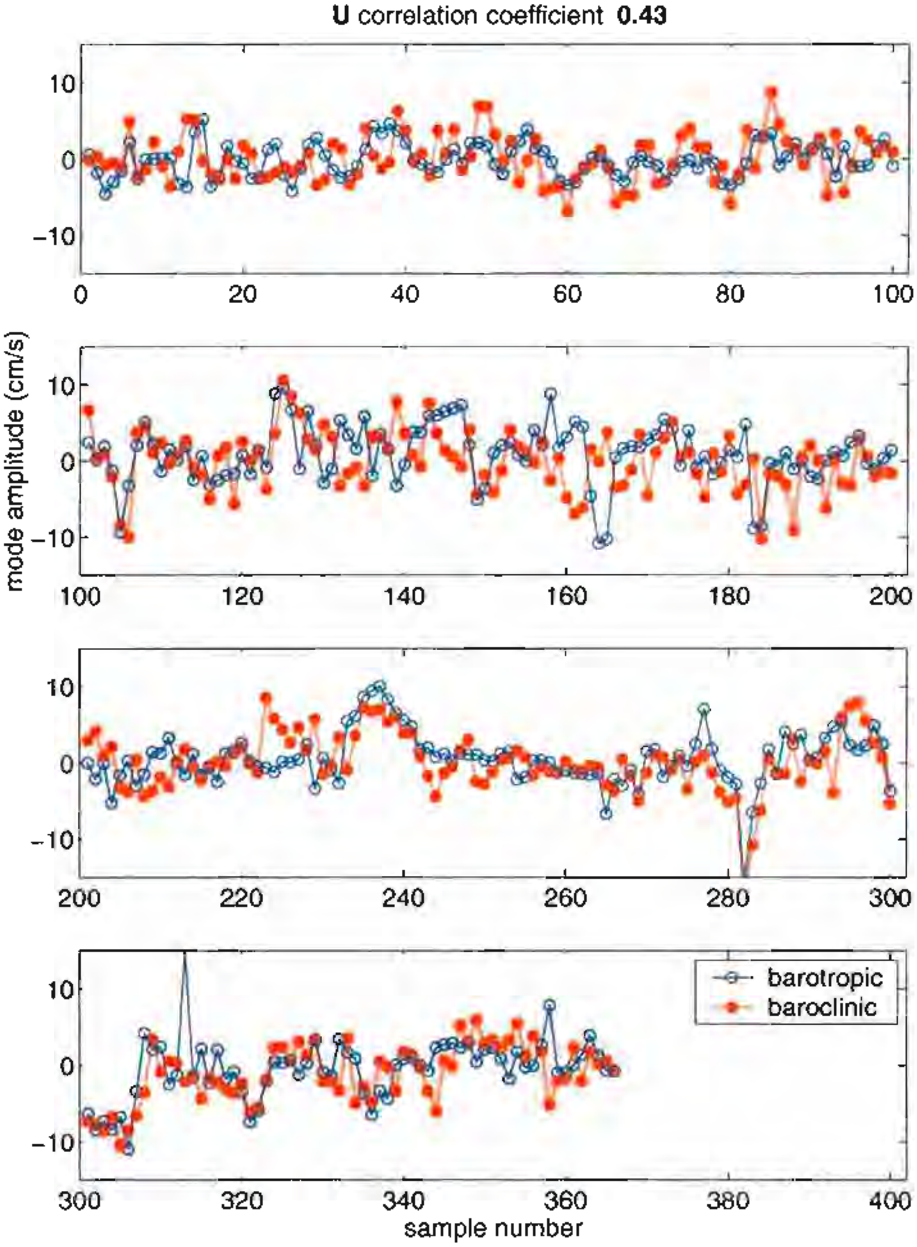


Figure 4.1: The amplitudes of barotropic and first baroclinic modes from 366 LADCP U profiles. The LADCP profiles are those indicated by the black vectors in figure 2.1. The profiles are plotted together if they are physically adjacent. Cruises are indicated by the shaded and unshaded areas. The title indicates the correlation coefficient between the two curves.

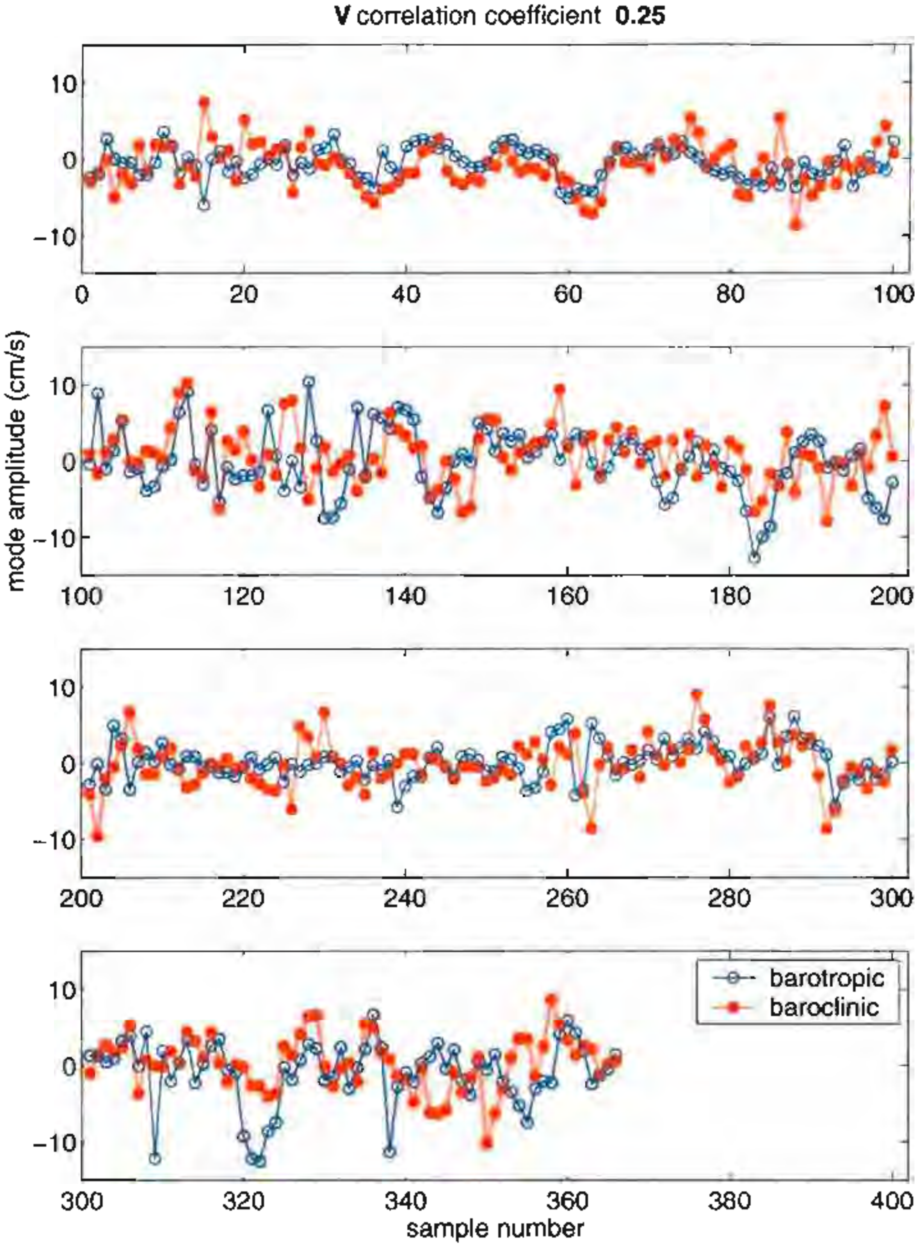


Figure 4.2: Same as figure 4.1, but for LADCP V component.

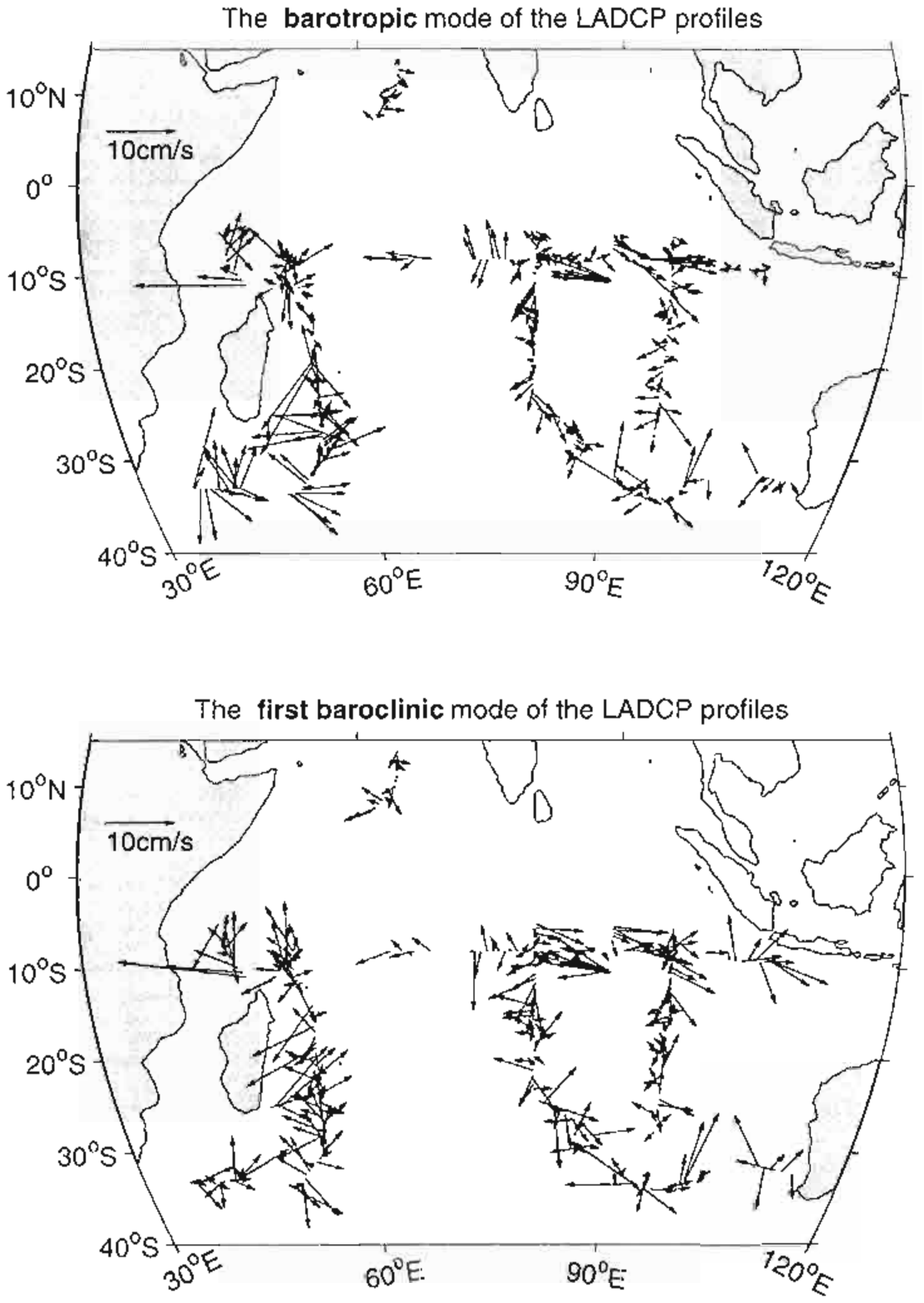


Figure 4.3: The u's and v's of the LADCP profiles are separately projected to the same barotropic and first baroclinic modes. Their mode amplitudes form the vectors in the figure. For clarity, the profiles from the repeated cruise by Doug Wilson is not included. The barotropic mode of the LADCP profiles (upper), which is just the depth-average of the LADCP velocity. The vectors are equivalent to the black ones in figure 2.1. The first baroclinic mode of the LADCP profiles (lower).

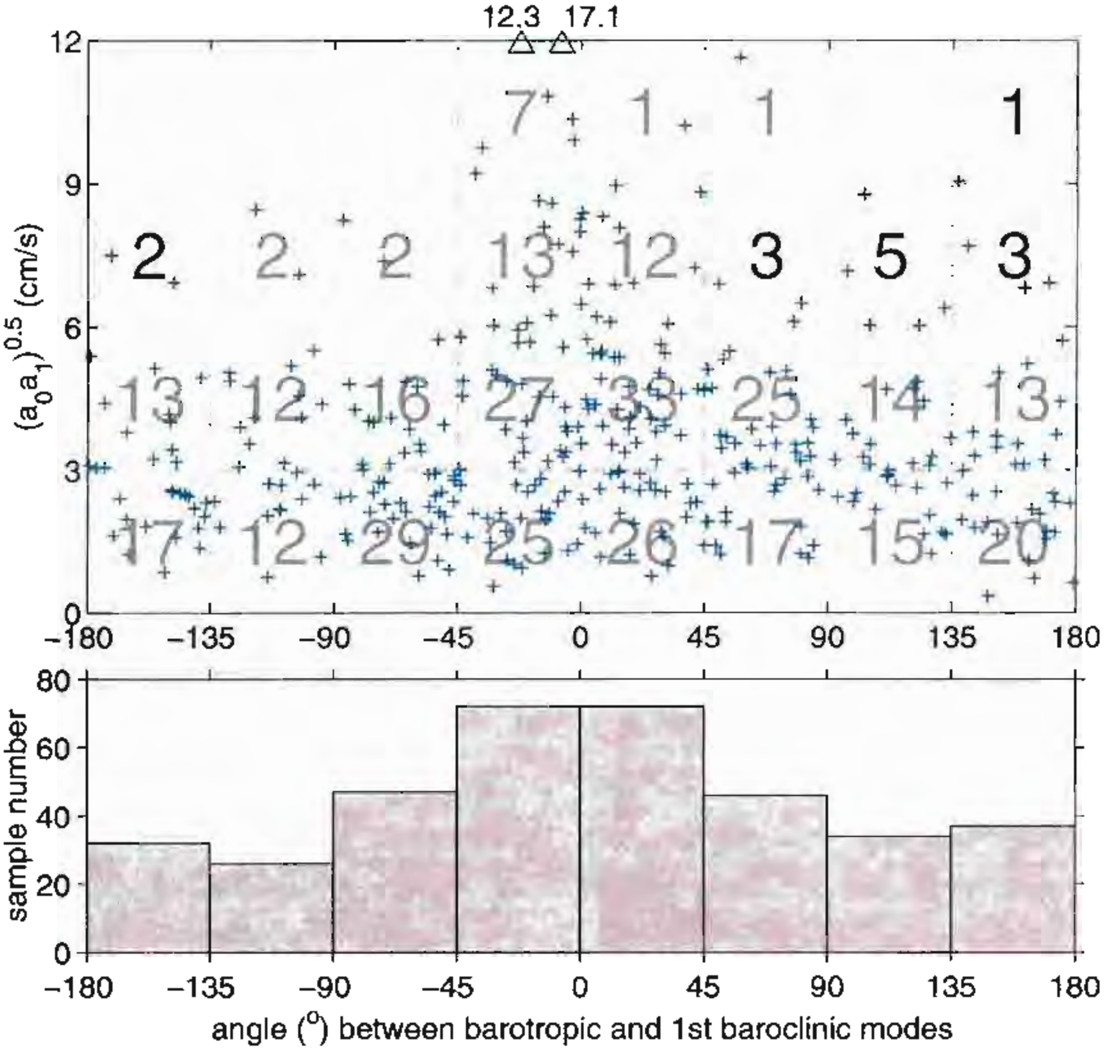


Figure 4.4: Angle between barotropic and 1st baroclinic velocities (upper) The abscissa is the angle in degrees between the two corresponding vectors in the upper and lower panels of figure 4.3 from the same LADCP profile, i.e., between the barotropic and first baroclinic modes of the LADCP profile. Positive angle is anti-clockwise from the barotropic mode to the first baroclinic mode. In the ordinate, a_0 is the amplitude of the barotropic mode, and a_1 is that of the first baroclinic mode; they are in fact the lengths of the corresponding vectors in figure 4.3. The numbers represent the number of profiles whose angle and $\sqrt{a_0 a_1}$ fall into the appropriate boxes, indicated by the dotted lines, except the box with two triangles. The triangles indicate two outliers; the numerics are their $\sqrt{a_0 a_1}$. These statistics are estimated using all 366 profiles in the mid-latitudes. The histogram of the angles between the barotropic and first baroclinic modes (lower). This is actually the summation of the upper panel along the ordinate.

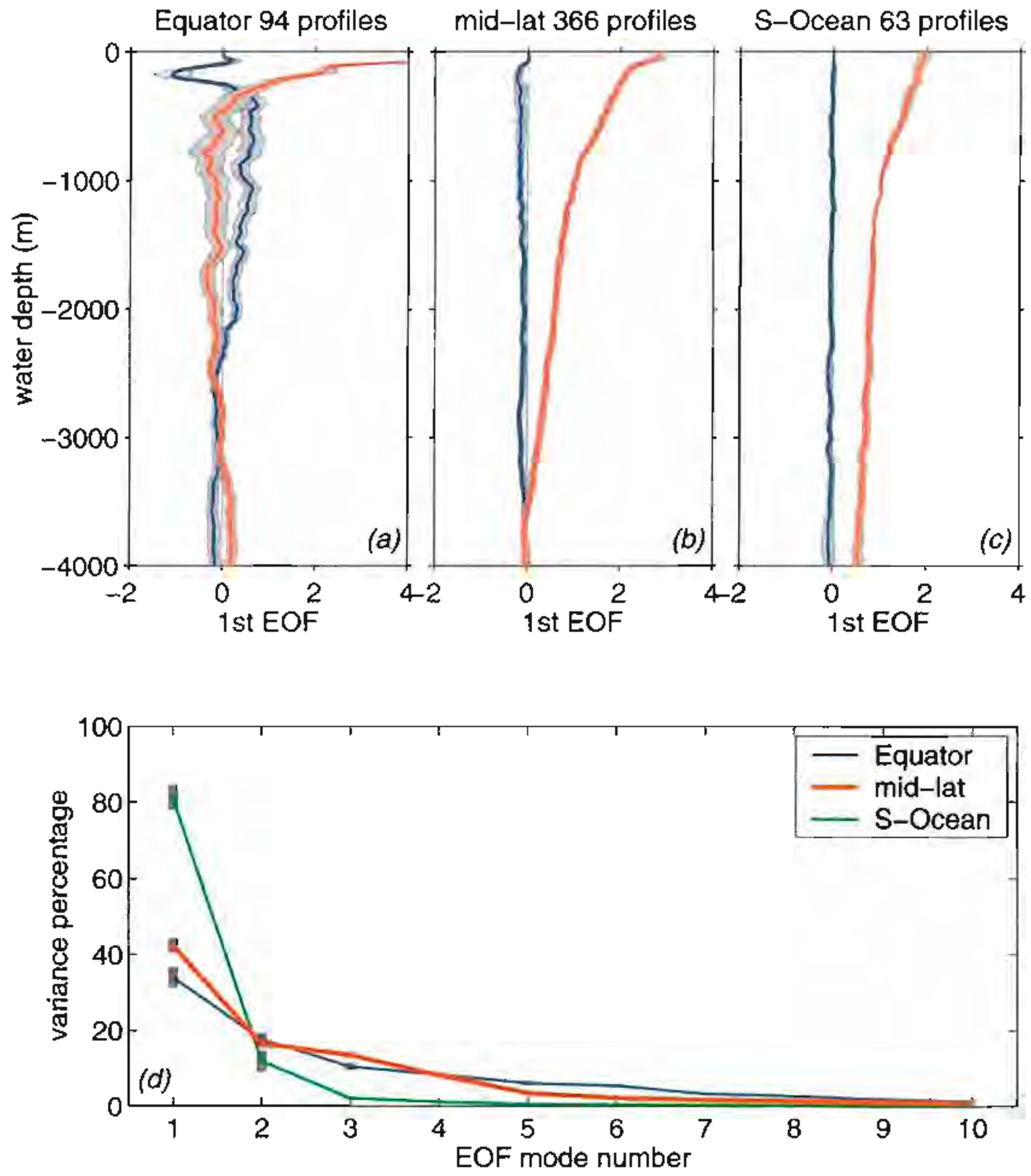


Figure 4.5: (a,b,c) The first EOF from LADCP profiles in three regions. (a) is 5 degree latitude near the equator; (b) is the mid-latitudes from 5 to 35 degrees latitude; (c) is south to 35 degree latitude (*ref.* figure 2.1). All profiles deeper than 4000 m are truncated before the EOF analysis. The shaded area indicates one standard deviation from 100 re-samplings of the LADCP profiles, assuming each profile has equal probability. (d) The variance percentage explained by the first ten EOF modes for each region. Its error bar is one standard deviation of the 100 re-samplings.

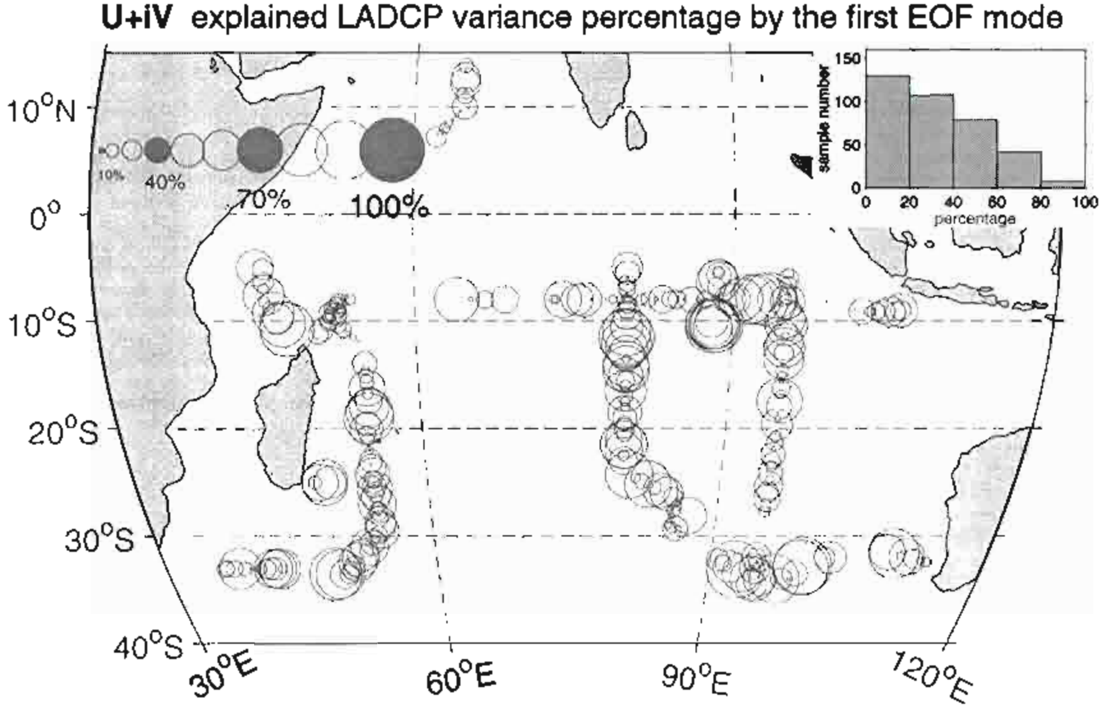


Figure 4.6: The percentage of LADCP variance in the first EOF. The legend is at the upper-left corner of the plot. For clarity, the data for the repeated cruise is not included. The inset at the upper-right is a histogram of the percentage of the explained variance over all 366 LADCP profiles.

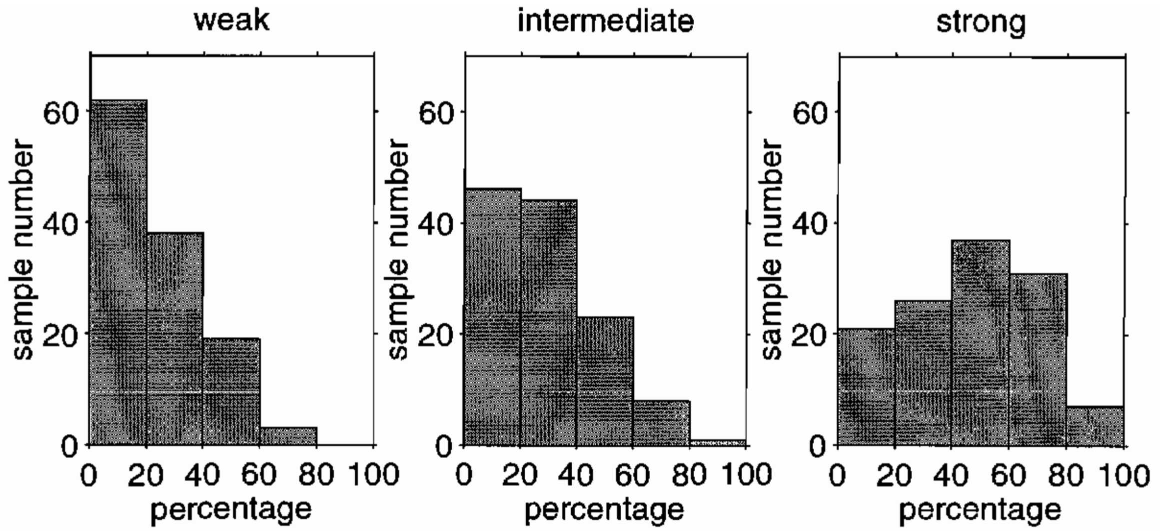


Figure 4.7: The percentage of variance in the first EOF. The 366 profiles are evenly divided into weak, intermediate and strong groups. The strength of a LADCP profiles is defined by its depth-averaged velocity squared.

Part II

NUMERICAL MODELING: JAMSTEC GCM

CHAPTER 5

MODEL CURRENTS AND COMPARISON WITH LADCP PROFILES

One snapshot (figure 1.1) already indicates that surface mesoscale motions are energetic and ubiquitous. Much more can be learned from three years of model output (219 5-day snapshots). We will present three-year means and their standard deviations (STDs) at different depths, which indicate the mesoscale motions' deep penetration and regional differences.

We do not compare the model output and LADCP velocity profiles at a given space and time. Instead, we will compare the EOFs from model output and LADCP profiles, with attention to the vertical structures and energy partition in EOFs. A favorable comparison, though in a statistical sense only, gives us the confidence to analyze model output in detail. The comprehensive picture from such a three year model output can not be obtained by observations alone.

5.1 Mean Currents and Variability

The study region (figure 5.1) mainly consists of two basins (Wharton Basin in the east and Indian Basin in the west) and three ridges (Ninety East Ridge in the middle and Central Indian Ridge and Chagos-Laccadive Ridge in the west). The basins are mostly deeper than 4000 m. Another major topographic feature is the Broken Plateau, west of Australia.

Three strong surface currents exist (upper-left, figure 5.2). Near the southern boundary of the study region, the strong eastward currents are part of the ACC system. The westward flow around 10-20°S is the South Equatorial Current. The eastward current near the Equator results from 1) the Equatorial Counter Current from January to March, 2) the Southwest Monsoon Current from July to September and 3) the eastward Equatorial Jets in between. Those features match the known general circulation in this region quite well (Tomczak and Godfrey 1994)

At both 5 m and 99 m (upper-left and upper right, figure 5.2), one can identify the strong southward Leeuwin Current along the west coast of Australia. Strong currents flow through the model Lambok Strait at about 115°E and 9°S. The model strait is 1 degree longitude wide and about 400 m deep.

The features in STDs closely relate to those in the mean currents (figures 5.3-5.6). The ACC and equatorial regions have high STD. From the maps at 5 m and 99 m, two bands of high STDs in the mid-latitudes originate in the Leeuwin Current and Lambok Strait. A few degrees away from the west coast of Australia and from the Lambok Strait, STDs decrease. This indicates that eddy energy is generated at those locations. The Broken Plateau, west of Australia, seems to steer eddies to the north. The maps at 1007 m and 2943 m show enhanced STDs near the Chagos-Laccadive Ridge.

The orientation of the ellipses generally follows the nearby topography, particularly in deeper layers. However, the orientation near the equator does not follow this rule.

5.2 Comparison with the LADCP Profiles

The locations of the model and LADCP velocity profiles, used to calculate the corresponding EOFs, are shown in figure 5.7. All the model and LADCP profiles are deeper than 4000 m and fall between 5° and 35°. Only one snapshot of model profiles is used.

As seen in the upper panel of figure 5.8, the vertical structures of the first EOFs from the LADCP and model profiles are unidirectional and surface-intensified. The lower panel of the figure shows EOF spectra or the variance in each EOF. The variance in the model first EOF is about half of that of the LADCP. For higher EOFs, the discrepancy is larger.

The lack of mesoscale energy in numerical models is common. The comparison between a one-degree resolution GCM and current meter record (Wunsch 1997) showed a factor of 10 or more difference over all periods. As suggested for the Parallel Ocean Climate Model (POCM) (quarter degree horizontal resolution and 20 vertical levels, Stammer *et al.* (1996)), possible reasons include:

- Deficiency in the simulated hydrographic structures, which prevent baroclinic instability processes from serving as an adequate energy source.
- The dissipation coefficients have to be large enough to keep the numerical integration stable. To decrease the dissipation coefficients, one has to increase the model resolution.
- Due to near thermal wind balance, any erosion of temperature and salinity structures, through the surface boundary conditions (relaxation of temperature and salinity with time scale 30 days) or by their own diffusion or the relaxation of temperature and salinity to the climatology below 2000 m, will in fact act as a friction term for the velocity field.
- The wind-stress forcing may be accurate only to a factor of 2.

Despite the limitations of numerical models listed above, the JAMSTEC model reproduced reasonably well the observed dominant vertical structure, i.e., the first EOF.

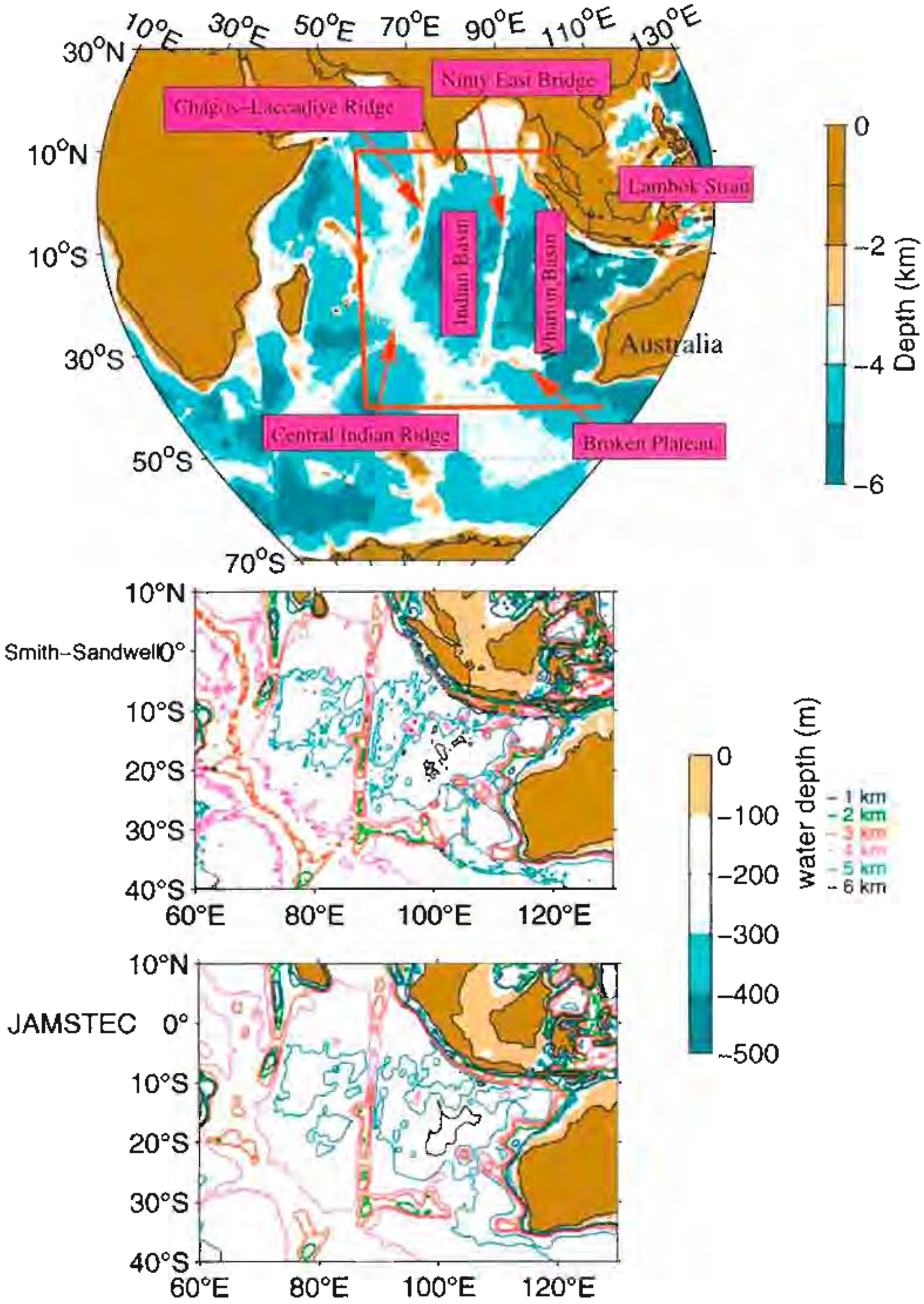


Figure 5.1: Indian Ocean and the study region. In the middle and lower panels, the shaded area is shallow than 500 m, and the color contours are 1, 2, $\cdots$, 6 km (the color scheme is at their right). "Smith-Sandwell" in the middle panel means the predicted topography by Smith and Sandwell (1994) and "JAMSTEC" in the lower panel means the depth used in the JAMSTEC model.

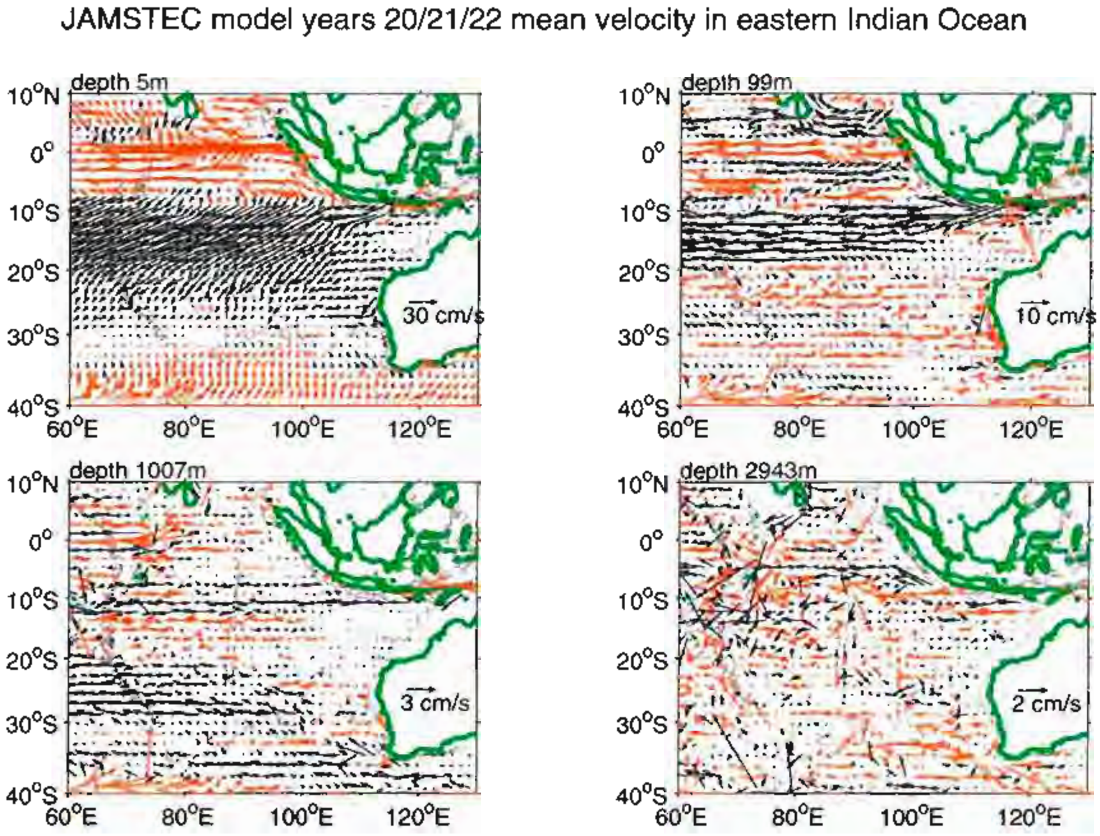


Figure 5.2: The three year mean currents at four different depths: 5 m, 99 m, 1007 m and 2943 m. Note that the velocity scale is different at each depth. Eastward (westward) currents are plots as red (black) arrows. Every sixth vector is plotted (i.e. the distance between vectors are $1\frac{1}{2}$ degrees along either latitude or longitude). The green contour indicates land and the thin gray contour is at 1000 m.

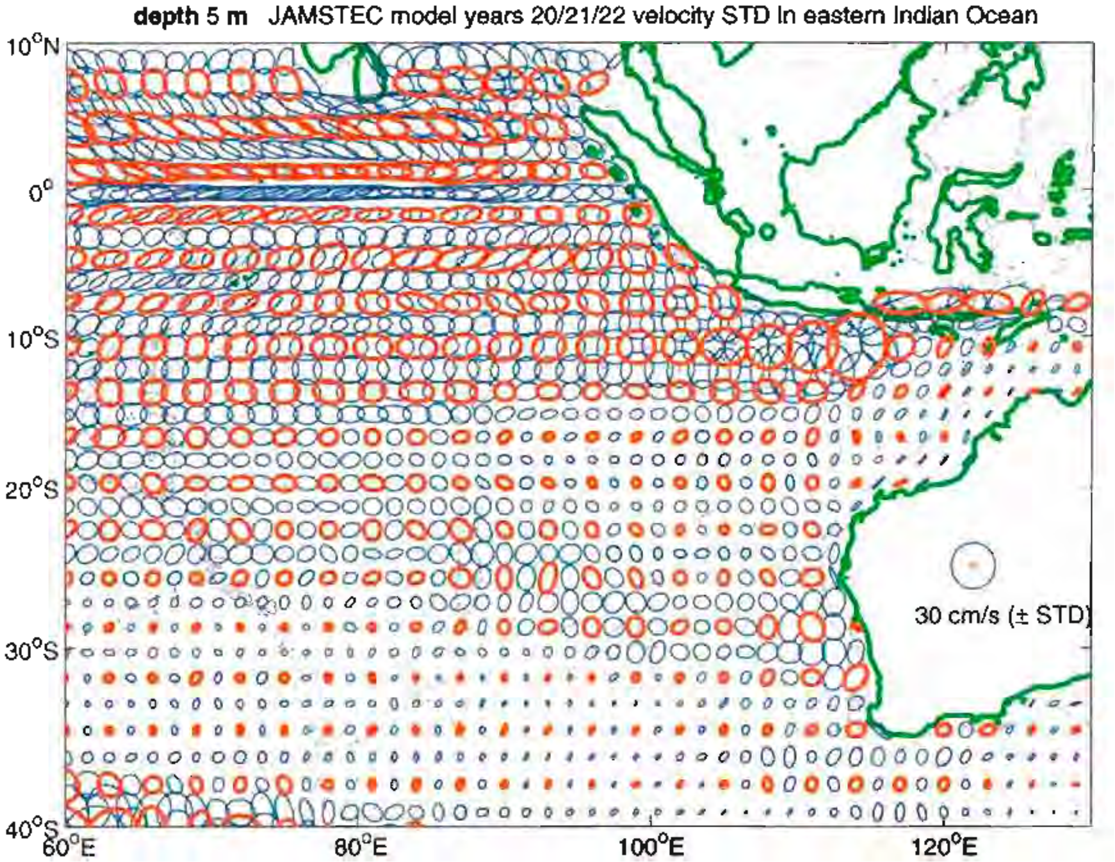


Figure 5.3: The standard deviation of the velocity at depth 5 m, sampled at the same locations as in figure 5.2. The variance ellipse is first calculated from the model output. The STD ellipse is defined as such that its major and minor axes are square root of those of the variance ellipse, and the orientation is same as that of the variance ellipse. The scale corresponds that for the vectors in figure 5.2. For clarity, 1 out of 4 ellipses is plotted in red.

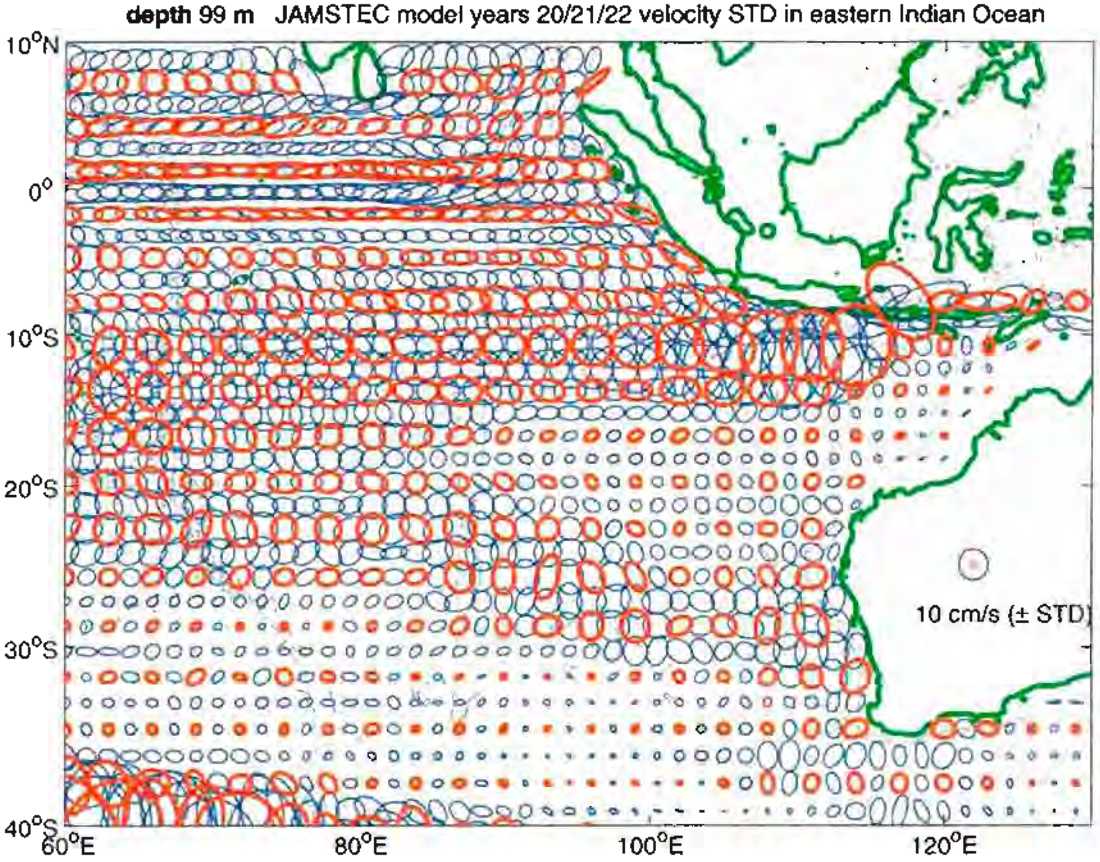


Figure 5.4: Same as figure 5.3 but for depth 99 m.

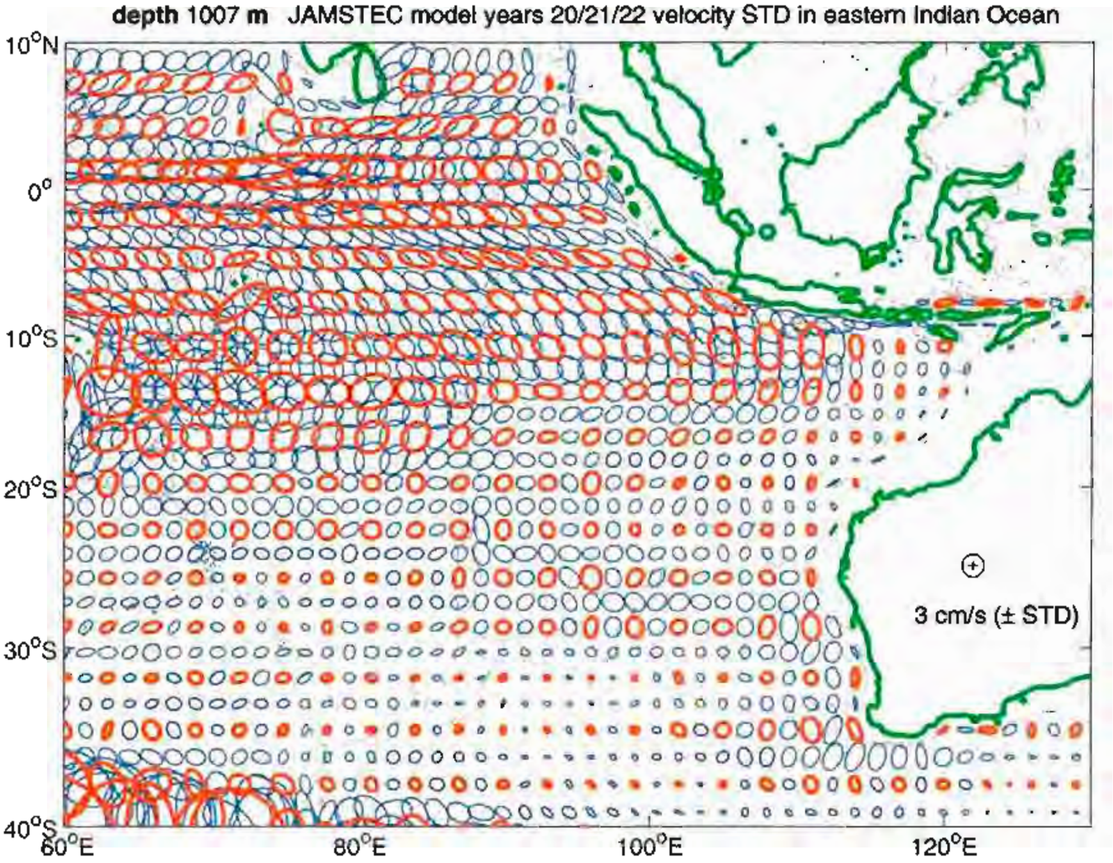


Figure 5.5: Same as figure 5.3 but for depth 1007 m.

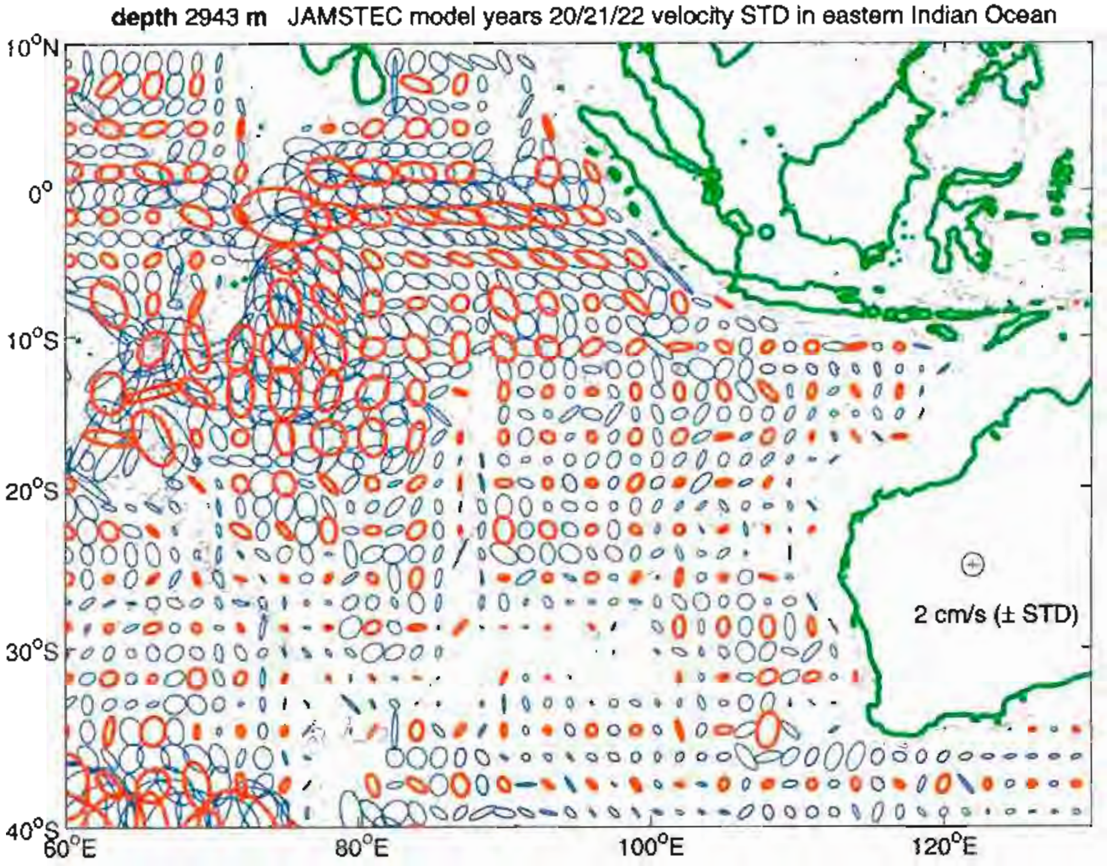


Figure 5.6: Same as figure 5.3 but for depth 2943 m.

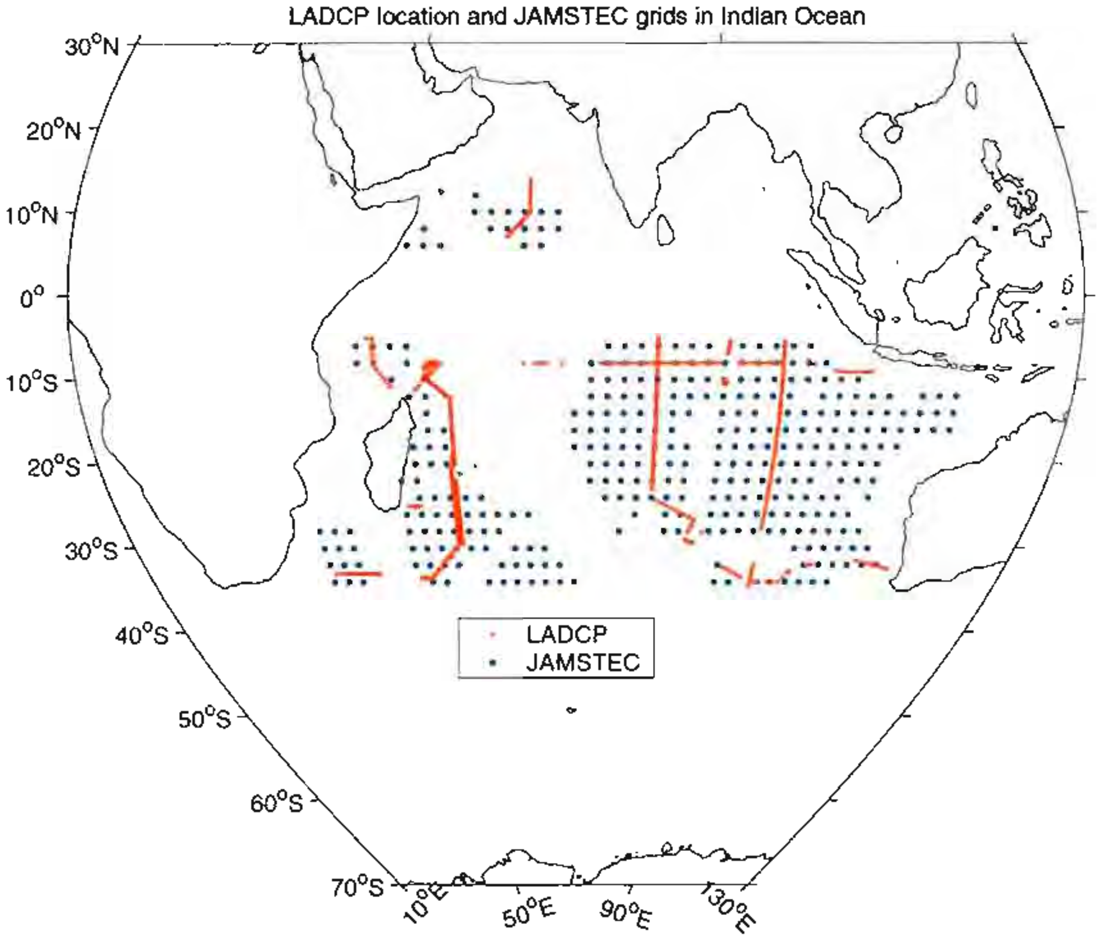
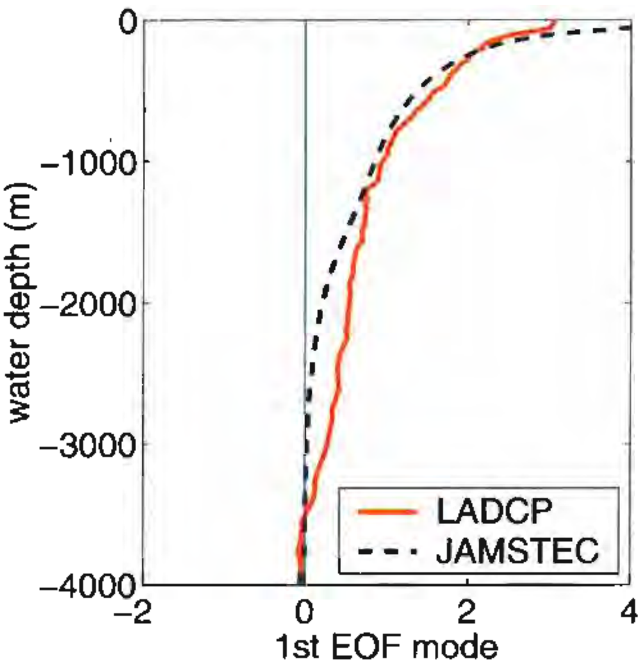


Figure 5.7: The red dots are the locations of the LADCP profiles used to calculate the EOFs in figure 5.8. The blue dots are the 2x2-degree grids at which the model velocities are used to calculate the EOFs in figure 5.8. All LADCP and model velocity profiles are deeper than 4000 m.



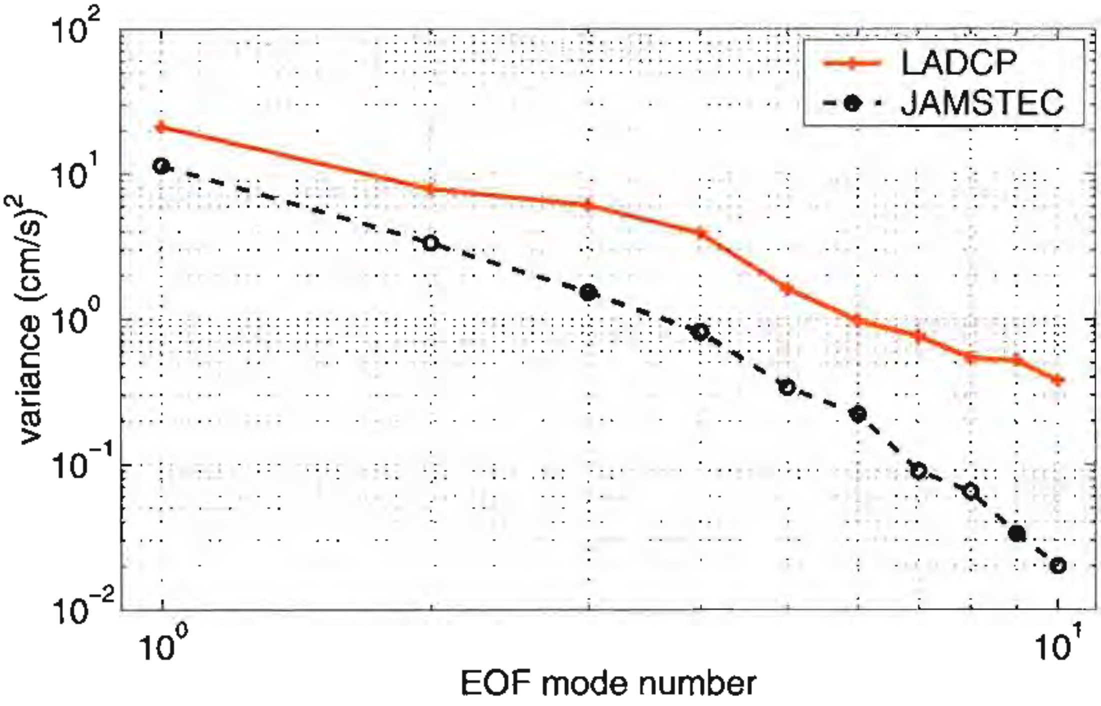


Figure 5.8: (upper) The first EOFs from 366 LADCP profiles and 328 model profiles. The locations of these profiles are shown in figure 5.7. All profiles are truncated to 4000 m before EOF analysis. (lower) The variance (the eigenvalue in EOF analysis) of the first ten EOFs from LADCP and model.

CHAPTER 6

THE FIRST EOF OF THE MODEL

The first EOF of model velocities is unidirectional and surface intensified (upper panel, figure 5.8). In this chapter, we will demonstrate how typical of the first EOF is in the snapshots of velocity components along longitudinal sections and time series of horizontal velocity profiles at given locations.

For comparison to the LADCP observation, the model EOFs (figure 5.8) are calculated using simultaneous velocity profiles from a large area. Hereafter, we will calculate the EOFs from the so-called geostrophic profiles over three years at each single location. This allows us to see how the vertical structure varies with position. The geostrophic profile is the velocity profile obtained from the geostrophic calculation referenced to the average of the model velocities at the deepest five depths. The main difference between the original model velocity profile and geostrophic profile is the near-surface wind-driven Ekman flow.

At each horizontal point, we obtain one first EOF. The sectional contours of the first EOFs will demonstrate the vertical structure in the upper panel of figure 5.8 is quite similar in the mid-latitudes.

6.1 Unidirectional Currents

Three snapshots in time of the model velocities along 30°S (figure 6.1) and the time series of the model velocities at given locations (figure 6.2) show the upper layer velocities extend to the bottom without switching direction and with decreasing magnitude, as in the first EOF in figure 5.8.

The percentages of unidirectional velocity profiles (figure 6.3) display three expected features: 1) one has little chance of finding a unidirectional profile near the equator, where the currents change directions (eastward or westward) over depth, 2) high percentages are found within the ACC region – both the ACC itself and associated eddies have significant barotropic components (Donohue *et al.* 2001), and 3) hot spots are associated with major topographic features (e.g. the Ninety East Ridge and the Chagos-Laccadive Ridge).

The obvious patchiness in figure 6.3 can not be explained fully in such a simple way. A close look in the small region, denoted by the white rectangle in figure 6.3, shows the possible role of topography. In figure 6.4, we see a connection between the topographic steps and local minima, but the topography does not affect much either downstream or upstream. Looking back at figure 6.1, we can see such local effects, especially at the step near 105°E, where a propagating current feature passes the topographic step without significant distortion. From the conservation of mass, we expect adjustments to the flow structure, as it runs across the topographic steps. The vertical resolution in the deep ocean is about 400 m (see appendix B) and it is possible that the connection between the topographic steps and local minima is a numerical artifact.

Topography can not account for all the patchiness. Over the flat topography, variability is still apparent. This is probably because three years is not long enough to yield stable statistics. Assuming a time scale of about 100 days, we have about ten degrees of freedom. As seen in figure 6.5, in which the percentages for each single year are shown, the large variations over flat bottom from one year to another could possibly result in the patchiness.

6.2 Spatial Distribution of the First EOF

Figure 6.6 shows the percentage of variance in the first EOF. Except for the equatorial region, the percentages are mostly above 50% and often as high as 80%. Patchiness exists as in figure 6.3. We identify three regions with high percentages: one in the south associated with the ACC system, one west of Australia (20-30°S), and another near the model Lambok Strait (10-15°S). The percentages are relatively low between those bands.

Figures 6.7, 6.8 and 6.9 show the vertical structures of the EOF along three latitudes: two of them cross the strong bands (27.5°S and 12.5°S) and the third falls between (17.5°S). In the strong bands, the vertical structure at neighboring grids varies gradually. At few locations along the weak band (17.5°S) the vertical structure varies abruptly. Notice that at those locations the variances are relatively low in the first EOF and relatively high in the second EOF. Therefore, the variances in the two EOFs are close (see the lower panel of figure 6.8), then so-called effective degeneracy could occur (North *et al.* 1982) and extra cautions are needed for interpreting those EOFs.

Surface geostrophic streamfunctions of the first EOF

Hereafter, the analyses will be carried out in boxes of 10 degrees longitude and 5 degrees latitude (figure 6.10).

Since the EOFs are calculated from the geostrophic profiles, their surface currents should be largely geostrophic. The streamfunction associated with these surface currents will be called the surface geostrophic streamfunction. The surface geostrophic streamfunction is directly comparable to the SSH, whose propagation characteristics have been extensively studied using altimetric observations (e.g., Chelton and Schlax 1996 and Zang and Wunsch 1999). In the next chapter, the surface geostrophic streamfunction of the first EOF will be used to infer propagation speeds of the mesoscale features.

The algorithm for calculating this streamfunction is outlined below (see figure 6.11): for the box with 25 velocity points (red circles), we apply one of the following equations 24 times in either north-south or east-west directions,:

$$\begin{cases} -f_0 v &= -g\frac{\partial h'}{\partial x} \\ f_0 u &= -g\frac{\partial h'}{\partial y}, \end{cases}$$

where f_0 is chosen as the Coriolis parameter averaged over the box and g is the gravitational acceleration. h' has the same units as SSH and the above equations are just the geostrophic relation on a f-plane. Therefore, we call h' the surface geostrophic streamfunction and denote it as ψ_{EOF}. We will obtain 24 equations for 16 unknown geostrophic streamfunctions for the situation in figure 6.11. The velocity used in the geostrophic relation is the average of the two adjacent velocities. The rule of the least squares is applied. The integral constant is constrained by assuming a zero-average spatially. The calculation is made for each snapshot; a matrix of 861×861 for each box in figure 6.10 is inverted. One example is in figure 6.12. The matching between the vectors and streamfunctions confirms that the surface currents of the first EOF are largely non-divergent and presumably geostrophic.

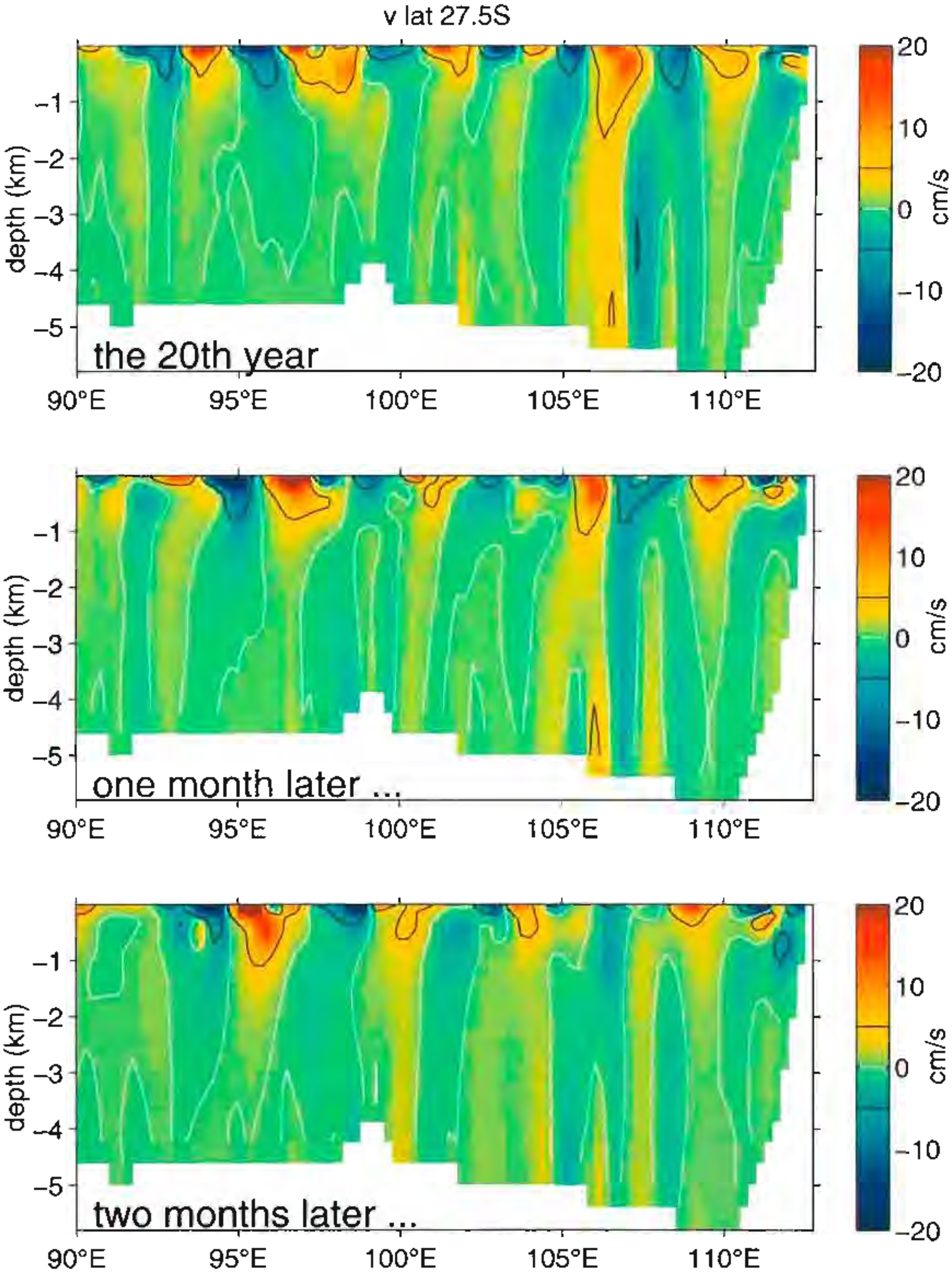


Figure 6.1: The north-south component of the model velocities along latitude 27.5°S. Panels are three snapshots one month apart. The section is west of Australia where instabilities of the Leeuwin Current generate eddies. The white contours are zero velocities. The black contours are ±5 cm s^{-1}.

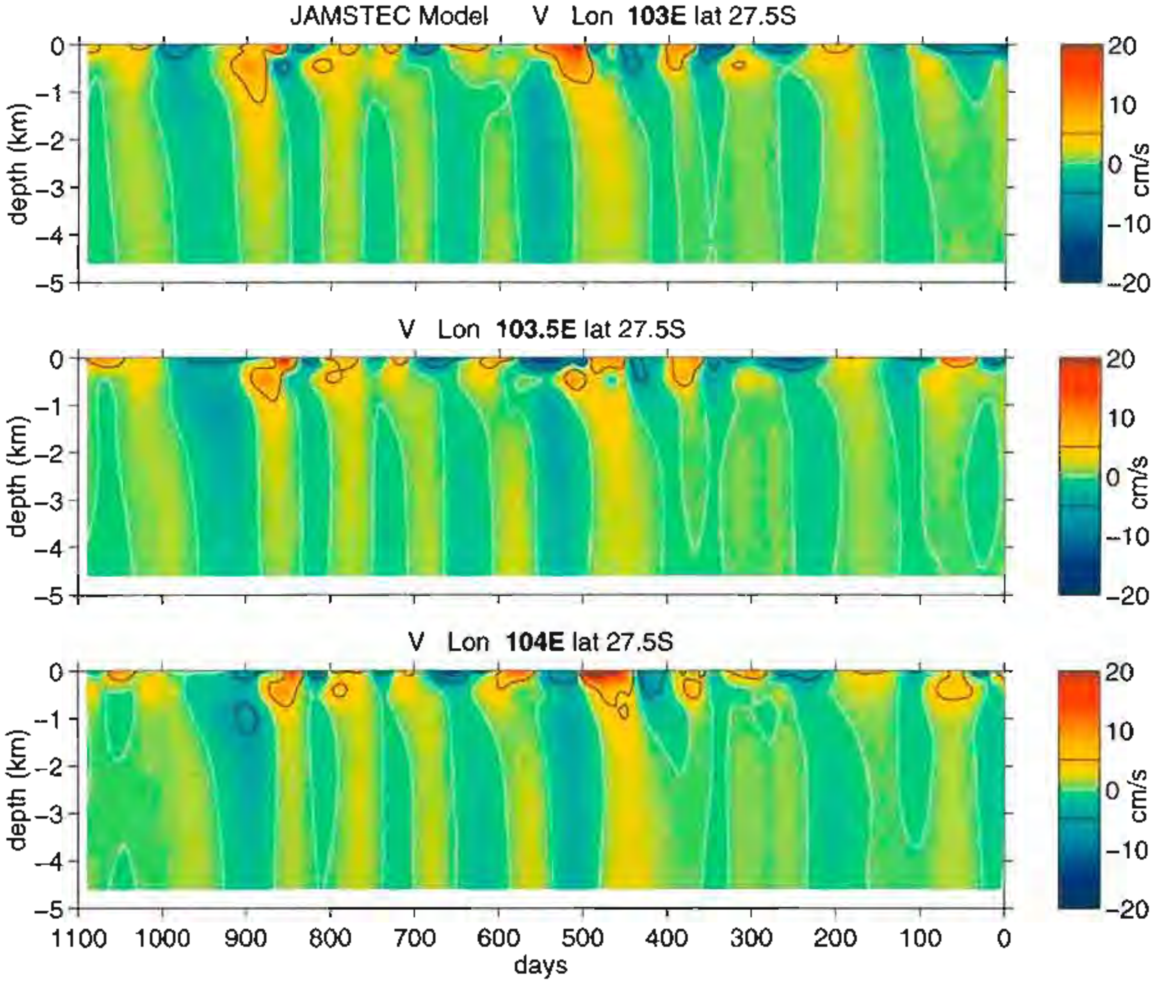


Figure 6.2: The north-south component of the model velocities at three grids, close to the west coast of Australia. The latitude is the same as in figure 6.1. The grids are a half degree longitude apart. Note the time axis runs from the right to the left, to show the westward propagation. The white contours are zero velocities. The black contours are ±5 cm s^{-1}.

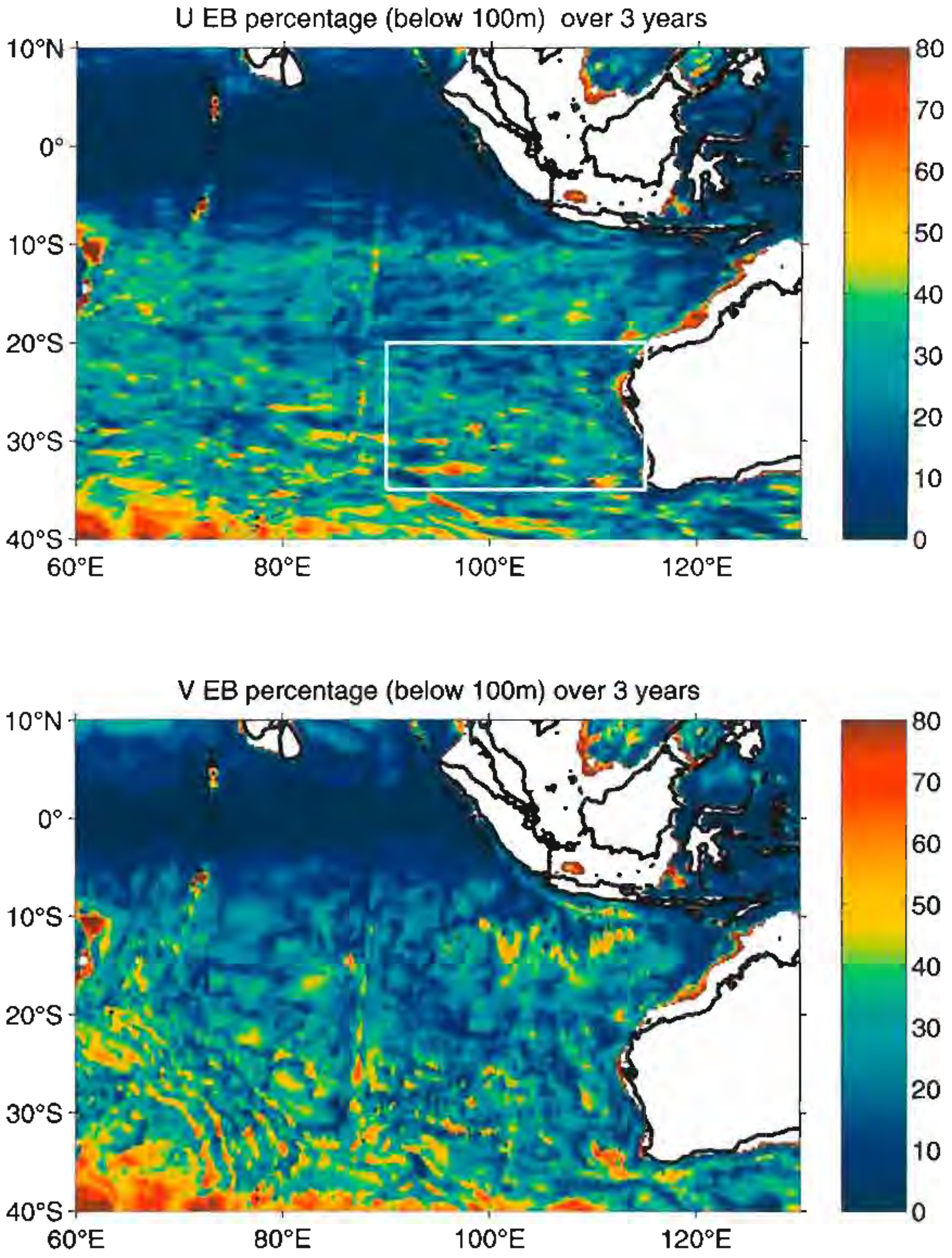


Figure 6.3: The percentage of unidirectional U and V profiles (below 100 m) over three years: in how many of the 219 snapshots do the U or V from 100 m to the bottom have the same sign at the given point?. The white rectangle in the upper panel is examined further in figure 6.4 with topographic contours.

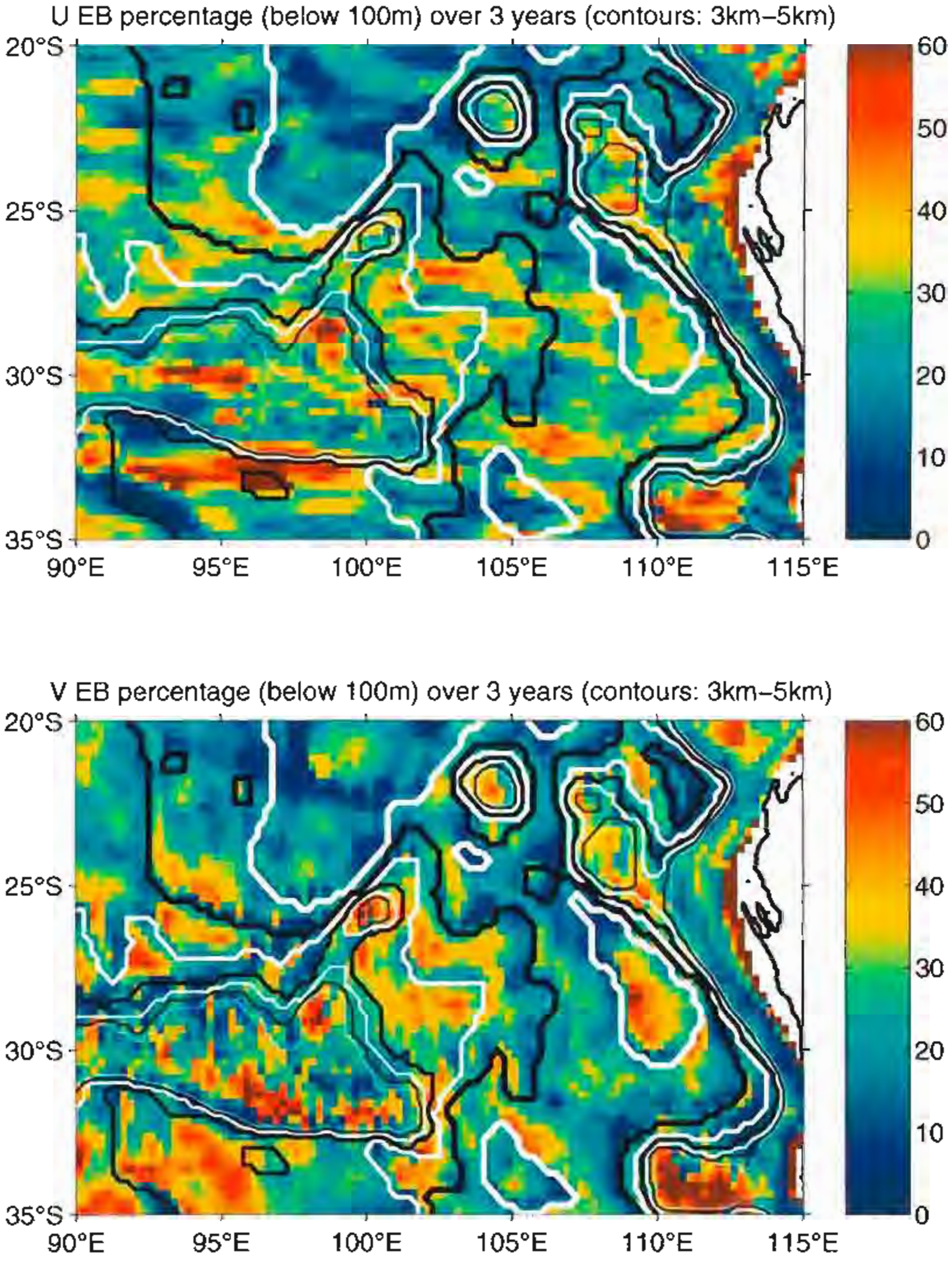


Figure 6.4: The percentage of unidirectional U and V profiles (below 100 m) over three years with topographic contours. The topographic contours were chosen so that the topography is exactly flat between the contours. The depths of contours are those (red lines) in figure B.1. Specifically, the four white contours, from thin to thick, are 3082 m, 3709 m, 4421 m and 5200 m; the black contours, from thin to thick, are 3383 m, 4055 m and 4801 m. The coast at the right is the west coast of Australia. The shallow part near 30°S is the Broken Plateau.

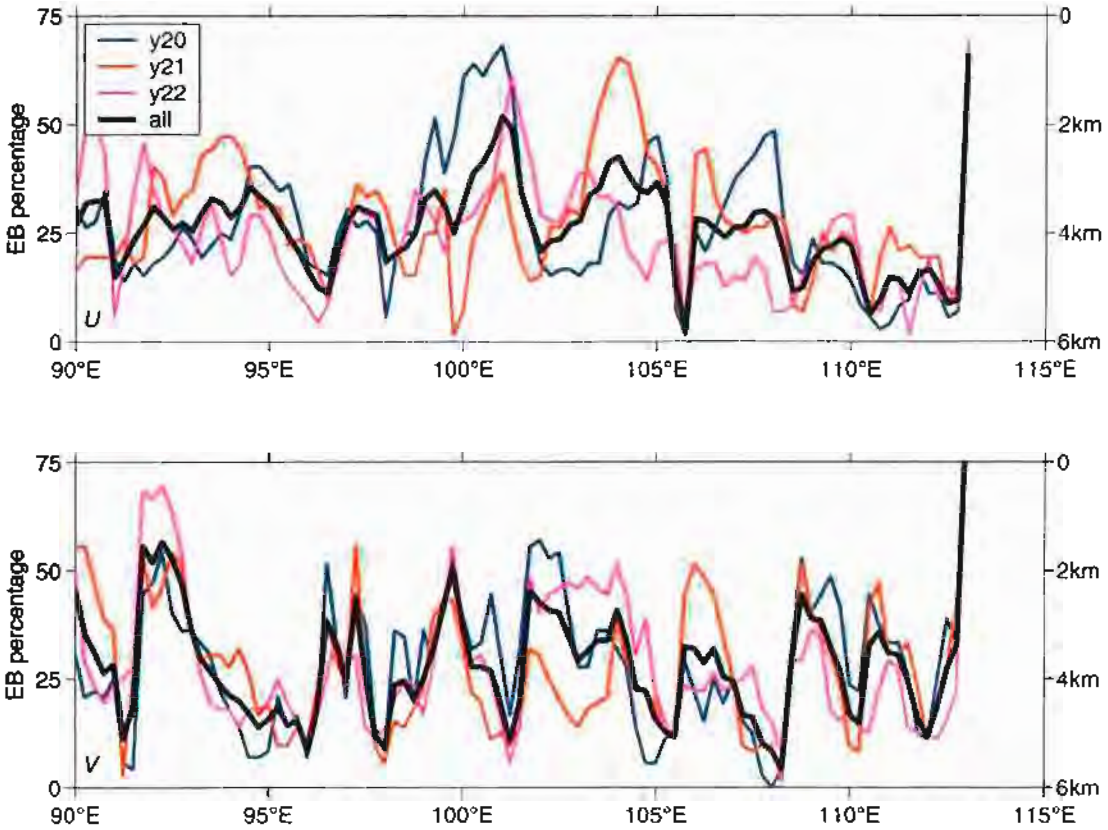


Figure 6.5: The percentage of unidirectional U and V profiles (below 100 m) over three years (thick black) or over one of three years along 27.5°S ('y20' refers to model year 20). The upper (lower) panel is for the east-west (north-south) velocity component. The shading shows the model topography, with a scale on the right.

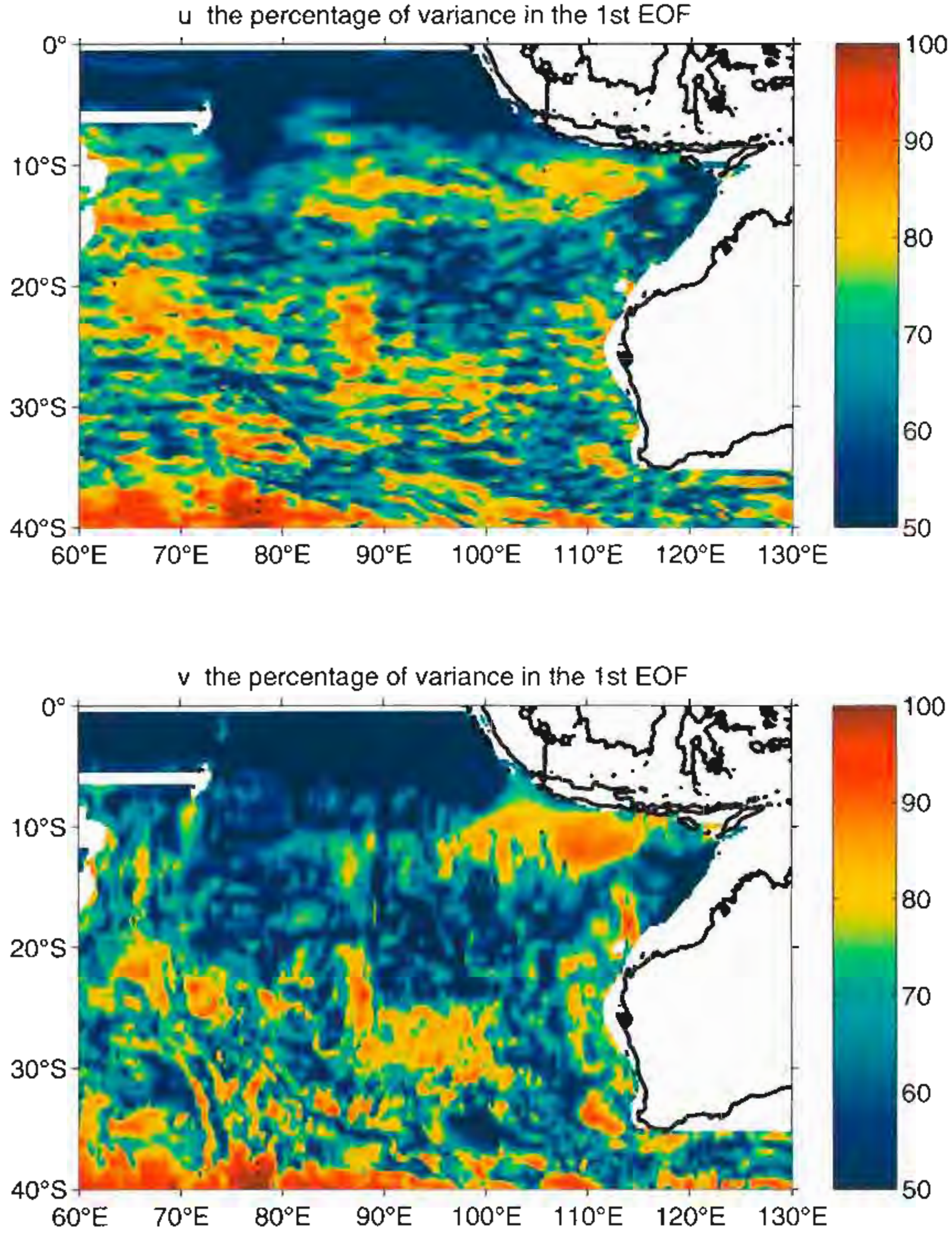


Figure 6.6: The percentage of variance in the first EOF. Upper panel is for east-west component and the lower panel for north-south one.

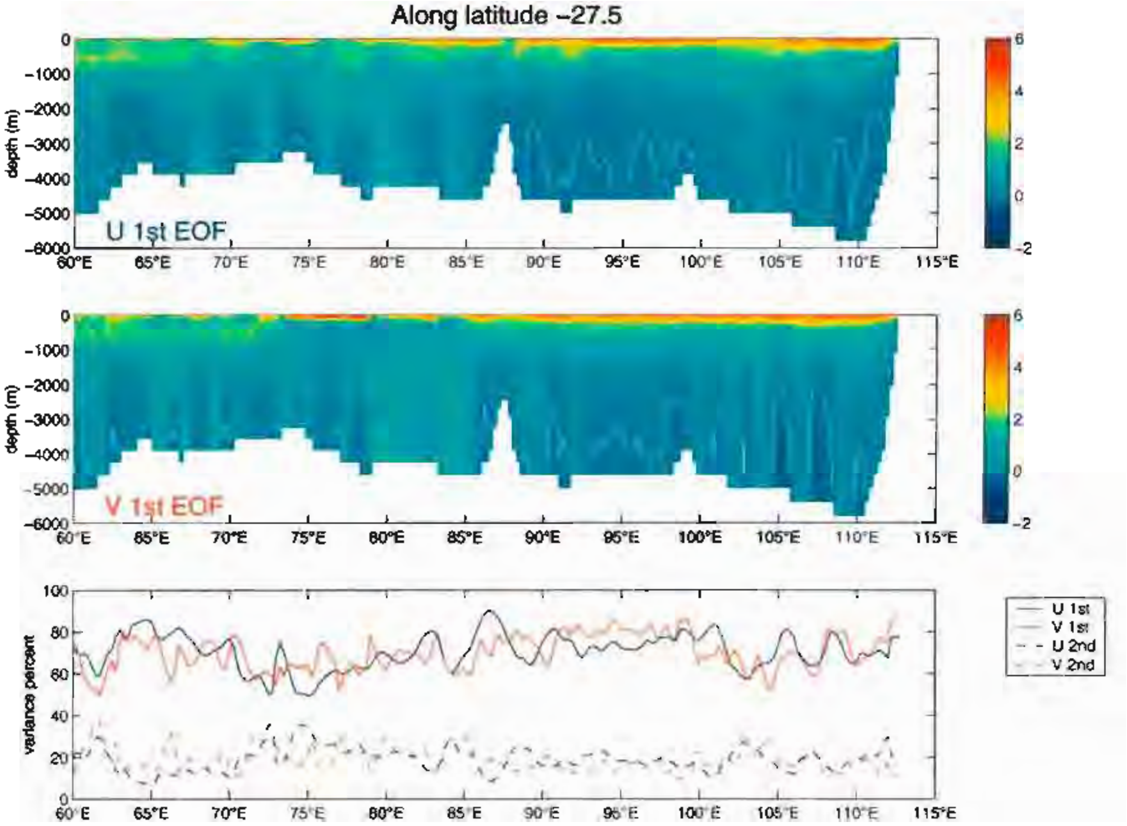


Figure 6.7: The first EOF along latitude 27.5°S. The first EOF is normalized such that its depth-average squared is unity and its surface value is positive. The white contours in the upper and middle panels are zeros. The lower panel is the variance percentages in the first and second modes.

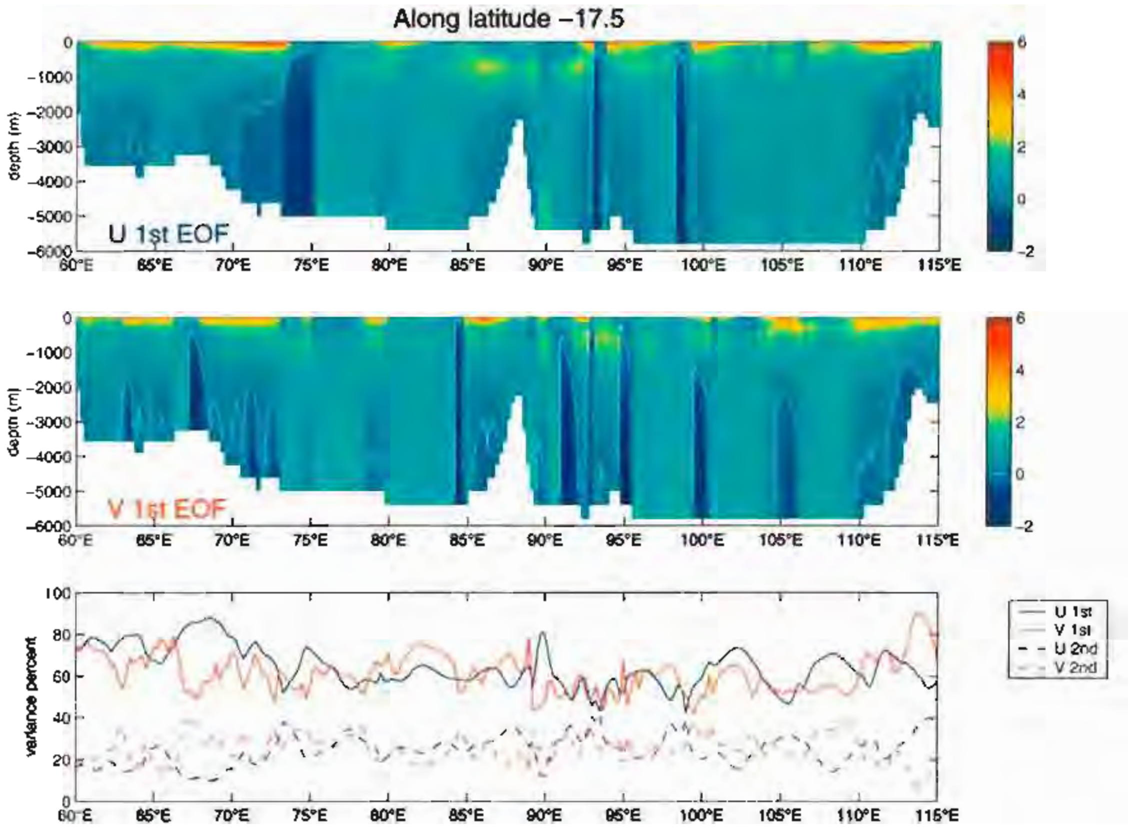


Figure 6.8: As figure 6.7 but along latitude 17.5°S.

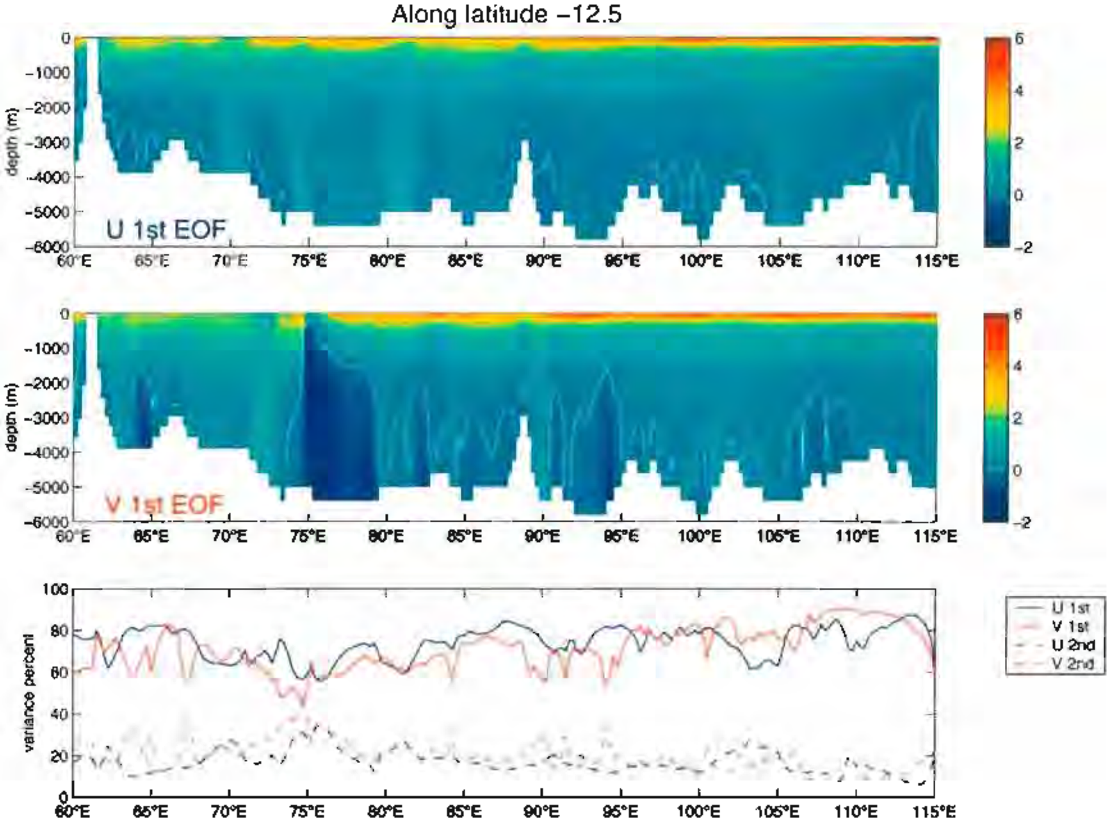


Figure 6.9: As figure 6.7 but along latitude 12.5°S.

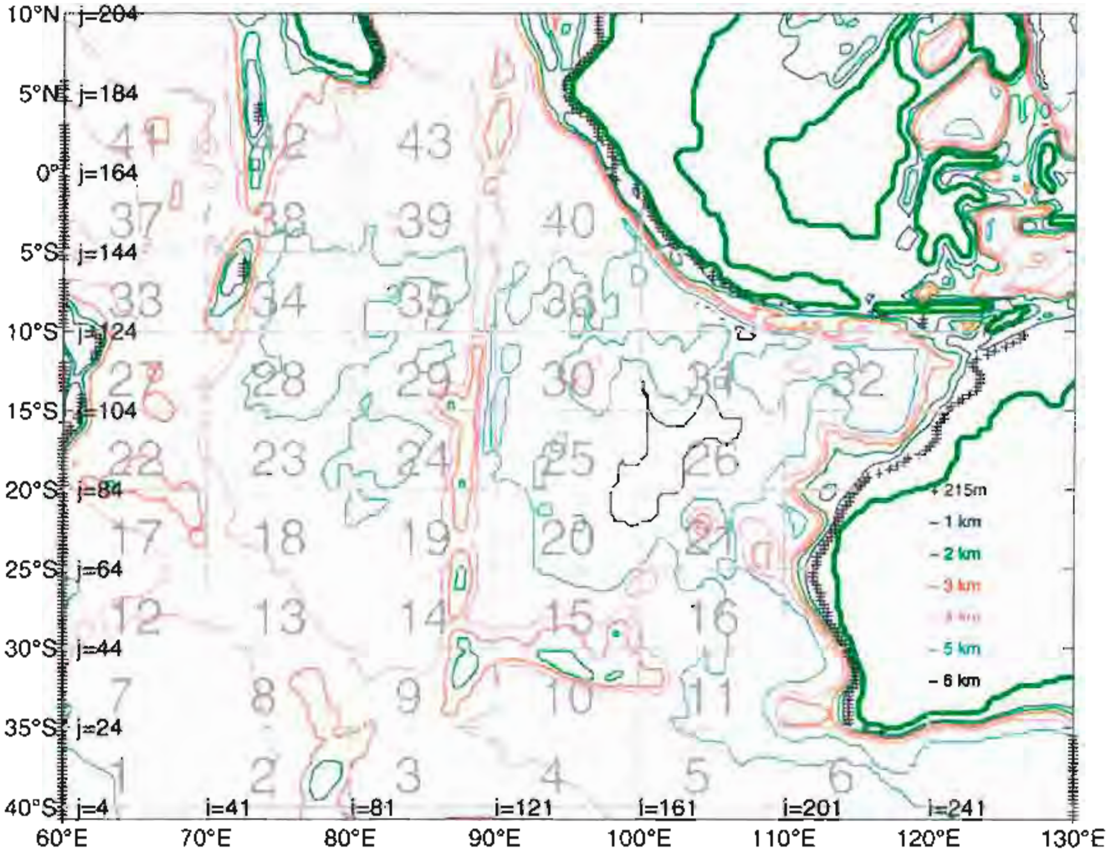


Figure 6.10: The 10 degrees longitude by 5 degrees latitude boxes. (i,j) are the grid cell indices. The black crosses indicate where the water depth is 215 m or the land boundaries of the model.

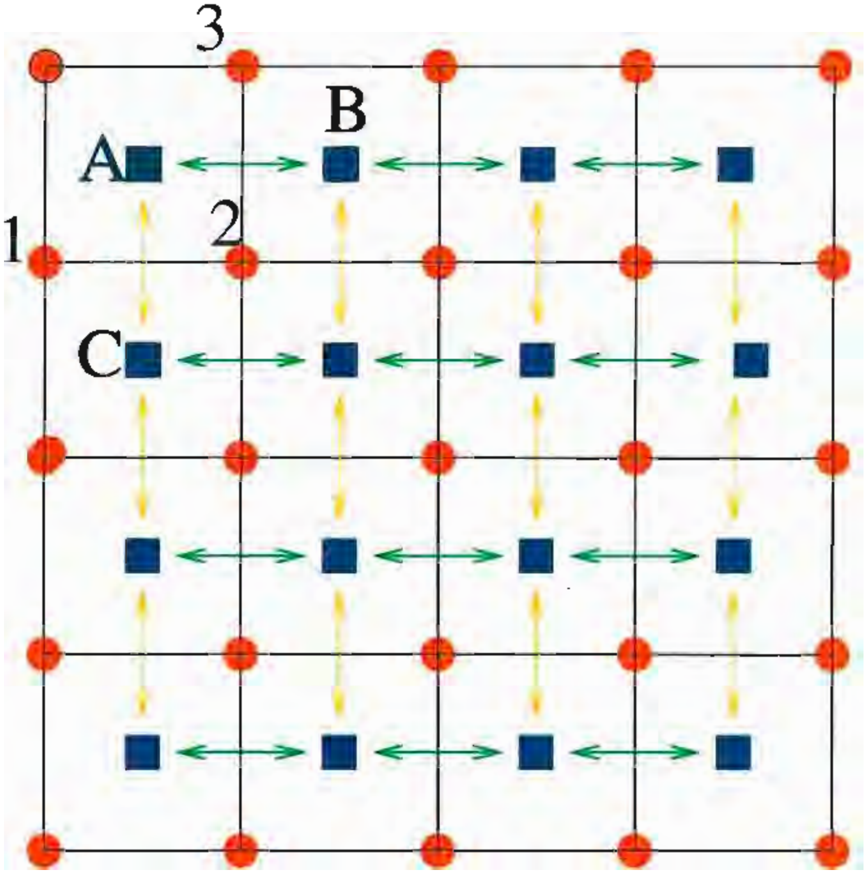


Figure 6.11: The scheme to calculate the geostrophic streamfunction. Red circles are locations of the surface currents of the first EOF. Blue squares are locations of the geostrophic streamfunctions. The geostrophic relation in east-west direction is applied at green arrows (e.g. between A and B) and the geostrophic relationship in north-south direction is applied at yellow arrows (e.g. between A and C).

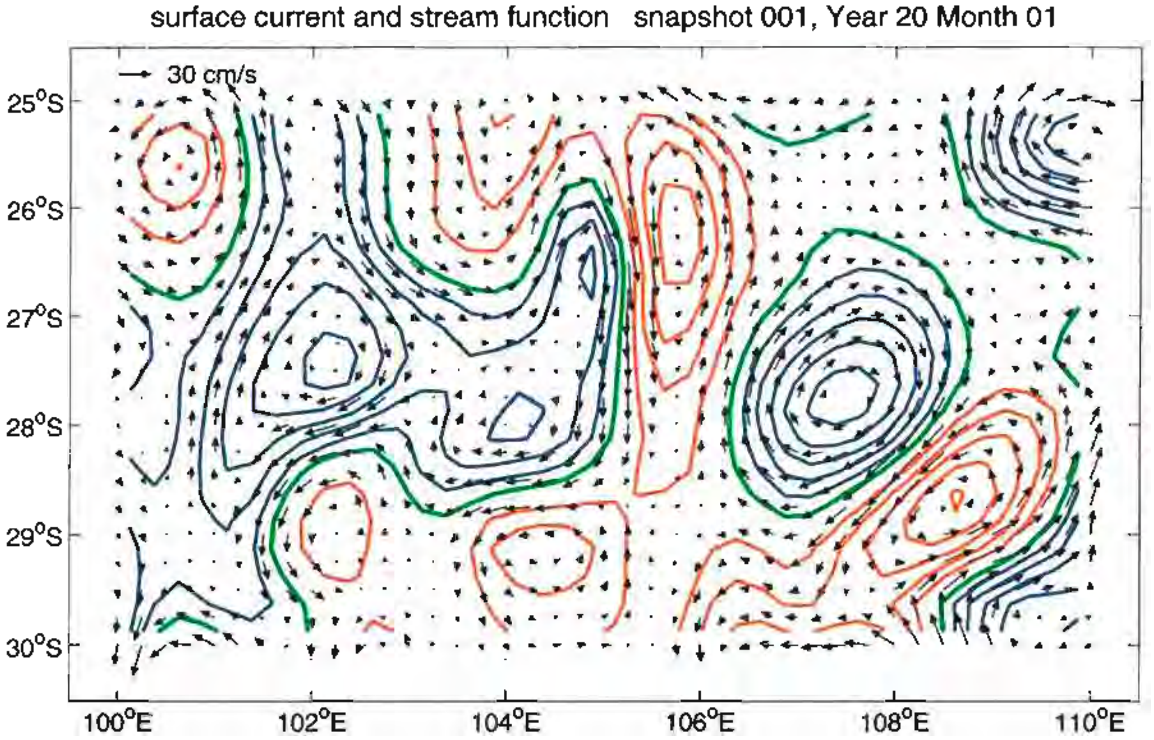


Figure 6.12: One snapshot of the surface currents and associated geostrophic streamfunction of the first EOF in box 16 (see figure 6.10). Zero is given by the green contours. The contour interval is 3 cm. Red (blue) contours are positive (negative). All the data are plotted.

CHAPTER 7

WESTWARD PROPAGATION OF THE FIRST EOF

Generally, the first EOF in the eastern Indian Ocean is neither the barotropic nor the first baroclinic mode but some combination. The propagation characteristics of the first EOF also fall between the barotropic and baroclinic modes at mid-latitudes and high latitudes, but not at low latitudes. These kinds of behavior have been identified in SSH anomaly from altimetry (e.g. Chelton and Schlax 1996) and Zang and Wunsch 1999) but the dynamics are still the subject of debate. The present model has good horizontal resolution, and we are not confined to long Rossby waves when interpreting model results.

We will also try to see whether 1) there is north-south propagation of the mesoscale motions and 2) the east-west propagation speed is amplitude-dependent. Both of them, however, are not obvious from the JAMSTEC model output.

The lag-correlation, counting and the coherence methods are used. The first two methods are in physical space and the third is in wavenumber space. Since mesoscale motion is broadband in nature, the description in physical space is more appealing. The extreme situation is a δ function field. The best description of such a field is the movement of the single peak, although we could certainly describe the field in wavenumber space.

The lag-correlation and counting methods estimate one "bulk" propagation speed for one box (the boxes are shown in figure 6.10), without any frequency or wavenumber information. However, they are also different. The lag-correlation method is variance-weighted at each length scale. For a red spectrum, the lag-correlation method emphasizes long length scales. The counting method counts the peaks (local extrema) in the time series without regard for their magnitudes and so trends will not directly affect the results. When we view a 2-D animation of SSH, our eyes usually catch the peaks; the counting method is designed to do the same thing. For a slightly red spectrum with a sudden cutoff near the mesoscale, this method emphasizes the cutoff frequency.

Before going into details, we may look back again at figures 6.1 and 6.2 where westward propagation was visible. In figure 6.1, we easily see the perturbations move westward as time advances. In figure 6.2, the perturbations in the easternmost grid (lower panel) are found in grids to the west (middle and upper panels) at later times. Both figures show the vertical structure does not change substantially over two months (figure 6.1) or over one degree longitude (figure 6.2).

7.1 Lag Correlation Method

ψ_{EOF} is a functions of latitude, longitude, and time. The correlation is lagged in latitude and longitude but not in time (e.g. upper panel of figure 7.1) and then in time as well (e.g. lower panel of figure 7.1). The westward offset of the maximum in the lower panel indicates westward propagation and is seen in plots for other boxes. The maximum is also offset in north-south direction, and thus we infer the north-south propagation of the perturbation. The propagation speeds for all snapshots are used to get a mean and its STD for that time-lag and box.

There is usually a secondary maximum east of the origin in the lag-correlation maps like the lower panel of figure 7.1. If the field consists of only one westward propagating Fourier component, then the lag-correlation is periodic and there is another maximum one wavelength to the east. On the other hand, if the field consists of only one westward propagating solitary-type perturbation, there would be no secondary maximum. The actual mesoscale field is not periodic, so we find a maximum in the west and a secondary maximum in the east, just like the lower panel of figure 7.1. Therefore, a maximum in the east and a western secondary maximum is an indication of enhanced noise. Under such logic, the propagation speeds in which all the maxima over three years are found in the west are highlighted as closed circles and thick lines in figures 7.2 and 7.3.

Though messy, the results for several time-lags are presented together. The methodologically limiting factors are different for different time-lags and depend on the perturbations' propagation speed. Note that propagation speeds of Rossby waves vary substantially from high to low latitudes. If propagation is slow, the perturbation may not be able to pass through a single grid cell in a short lag times. For each time lag, the minimum detectable speed is one eighth degree longitude divided by the lag time, which is about 3 cm s^{-1} if the lag time is five days. The minimum detectable speed squared is added to the variance which affects STD in figures 7.2 and 7.3. When the lag time is larger than 5 days, this limiting factor is minor. If propagation is fast, the perturbation may have moved out of the box at large lag times. The shaded areas in figures 7.2 indicate the speeds at which the perturbations at the center of the box will move out of the box within the lag time. The estimated propagation speeds will be biased to smaller values, if the propagation speeds are close to the shading. The shading is not exactly same for each panel, since an east-west grid cell (one quarter degree longitude) has different length (e.g. in kilometers) at different latitudes.

There are quite large variations in the westward propagation speeds, as indication by the STDs (figure 7.2). Nonetheless, the maxima in lag correlation maps are usually obvious. Figure 7.2 shows that south of 15°S, the propagation speeds are close to those of long Rossby waves. At lower latitudes, the speeds are less than those of long Rossby waves. From 5-10°S, the speeds are only about one quarter of the long Rossby wave speed.

Figure 7.3 shows the north-south propagation speeds. There is no significant north-south propagation, but its variability is large. Using the next two methods, we will only look at the east-west propagation. Figure 7.4 shows the magnitude of the surface geostrophic streamfunctions within binned propagation speeds, in attempt to identify any amplitude-dependency of the propagations. Most boxes show a weak tendency that strong eddy fields seem to propagate faster. However, some boxes show an opposite tendency.

7.2 Counting Method

The counting method is illustrated in figure 7.5. In the upper panel, the peaks (local extremes) of two time series, one degree longitude apart, are evolving, but there is no ambiguity in matching most of them. The peaks propagate westward at various speeds, but a statistically stable mean exists as seen in the middle panel. The red line in the middle panel is the lag-correlation for the two time series. The histogram is appealingly narrow so the results from the histogram are preferred, while the results from the lag correlation methods are shown as a reference.

The lower panel is an attempt to see whether the propagation speed is amplitude-dependent. The variation is so large that we will not draw conclusions, similar to the situation in figure 7.4. As in figure 7.4, there is a weak indication the stronger peaks propagate faster: the maximum of the amplitude curve has a larger W-shift than that of the histogram.

The statistics in the middle panel are based on the peak differences and not on the peaks themselves. Therefore, pre-filtering the time series with a 35-day running mean may not directly limit the statistics here. In other words, the peak differences can be less than 35 days.

Figure 7.6 shows all the estimates for different boxes and different space-lags, which are almost the same for some boxes, such as 16, 15, 21 and 30. However, there are large discrepancies for others. The speeds inferred from medians are most robust with respect to space lags. The speeds from the maxima are usually faster than those from the medians. For Gaussian distributions, the estimates from the medians should be same as those from the maxima.

The overall results are basically similar to those from the lag correlation method. Looking more closely, however, the propagation speeds from the counting method are generally larger than those of the lag-coefficient method, especially south of 30°S. The counting method emphasizes local extrema and is not directly affected by the long-term trends. As seen in the upper panel of figure 7.5, the trend with time scale 200-300 days does not affect the results directly. As pointed out earlier, the lag correlation method may emphasize larger spatial scales and longer time scales, while the counting methods may emphasize mesoscale motions. The coherence method in the next section could shed some light on the faster propagation of mesoscale motions (near the first Rossby radius of deformation) compared to longer time scales.

7.3 *Coherence Method*

The previous methods estimate propagation speeds more in the bulk sense: one speed for one box and one time. Such methods work best if the fields are coherent. The motions at different frequencies and wavenumbers propagate at same speed. In other words, the dispersion curve is a straight line or

$$\psi(x, y, t) \equiv \psi(x - ct, y)$$

which certainly oversimplifies reality. The coherence method is designed to look at the propagation at different frequencies and wavenumbers.

From two three-year time series, one degree longitude apart, we could estimate co-spectra (lower-left panel of Figure 7.7) and phase spectra (lower-right panel of Figure 7.7) in the frequency domain. We assume the phase difference at each frequency between two time series is solely due to the westward propagation. Then we infer negative east-west wavelengths from the phase difference at that frequency. For example, if the phase difference is 90° at 0.01 cpd, the wavelength would be $\frac{360}{90} = 4 \times 1^\circ = 4^\circ$ longitude at that frequency. Thus, each frequency is paired by wavelength or wavenumber. Finally, we plot the wavenumber-frequency pairs together with the linear Rossby wave dispersion relationship (the upper panel of Figure 7.7). The lower-left panel shows that the waves at frequencies less than 0.02 cpd are energetic; this is generally true for other boxes. For this reason, we concentrate on the waves with frequencies lower than 0.02 cpd in figure 7.8.

The dispersion relationship of the first mode linear baroclinic Rossby wave is

$$\omega = -\frac{\beta k}{k^2 + l^2 + (1/2\pi R)^2},$$

where R is the first mode Rossby radius of deformation and ω and (k, l) are in units of cpd and cpkm, respectively, as in Figure 7.7. The curve in the figure is that of $l = 0$, and is the upper bound for all l. If waves from the model output fall above the curve of $l = 0$, there is a discrepancy between the model output and linear first baroclinic Rossby wave theory. From figure 7.8, such discrepancy exists for all boxes excepts those from 5-10°S.

At lower frequencies, waves tend to propagate at the long Rossby wave speed. The tendency is clear except in the boxes closest to the equator. At high (lower) latitudes, the waves at frequencies other than the lowest frequencies propagate faster (slower) than long Rossby waves. This pattern in figure 7.8 is consistent with the estimates from the previous two methods. In figure 7.8, within 15-35°S the phase speed of the first EOF is approximately that of non-dispersive Rossby waves of the first mode even at frequencies above the Rossby wave cutoff. The dispersion relations within 15-35°S in figure 7.8 are astonishingly similar to those from the altimetric analyses by Zhang and Wunsch (1999, their figures 4 and 5).

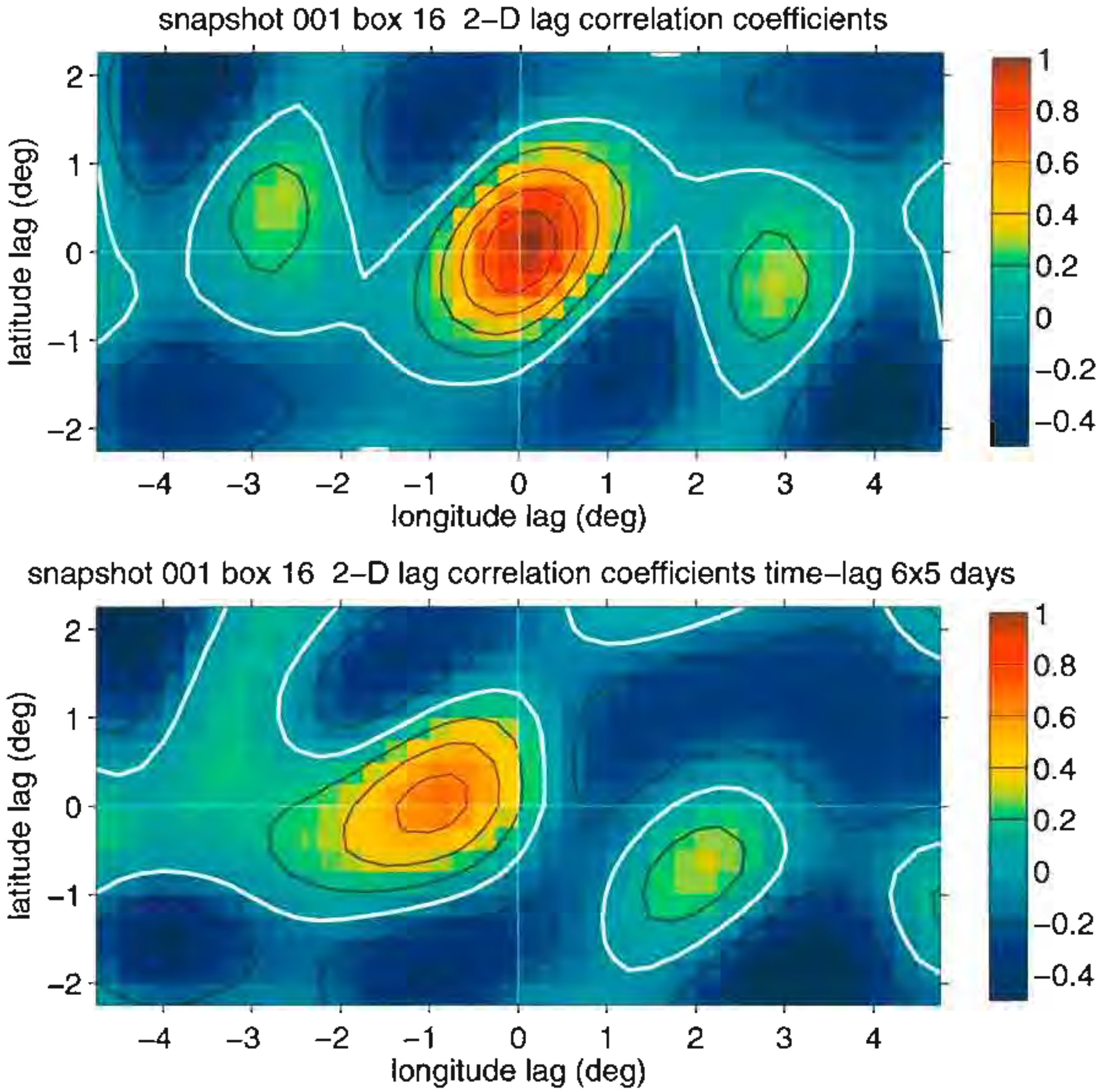


Figure 7.1: Upper panel: auto lag-correlation of ψ_{EOF} in box 16. Lower panel: lag-correlation between two ψ_{EOF} snapshots 30 days apart.

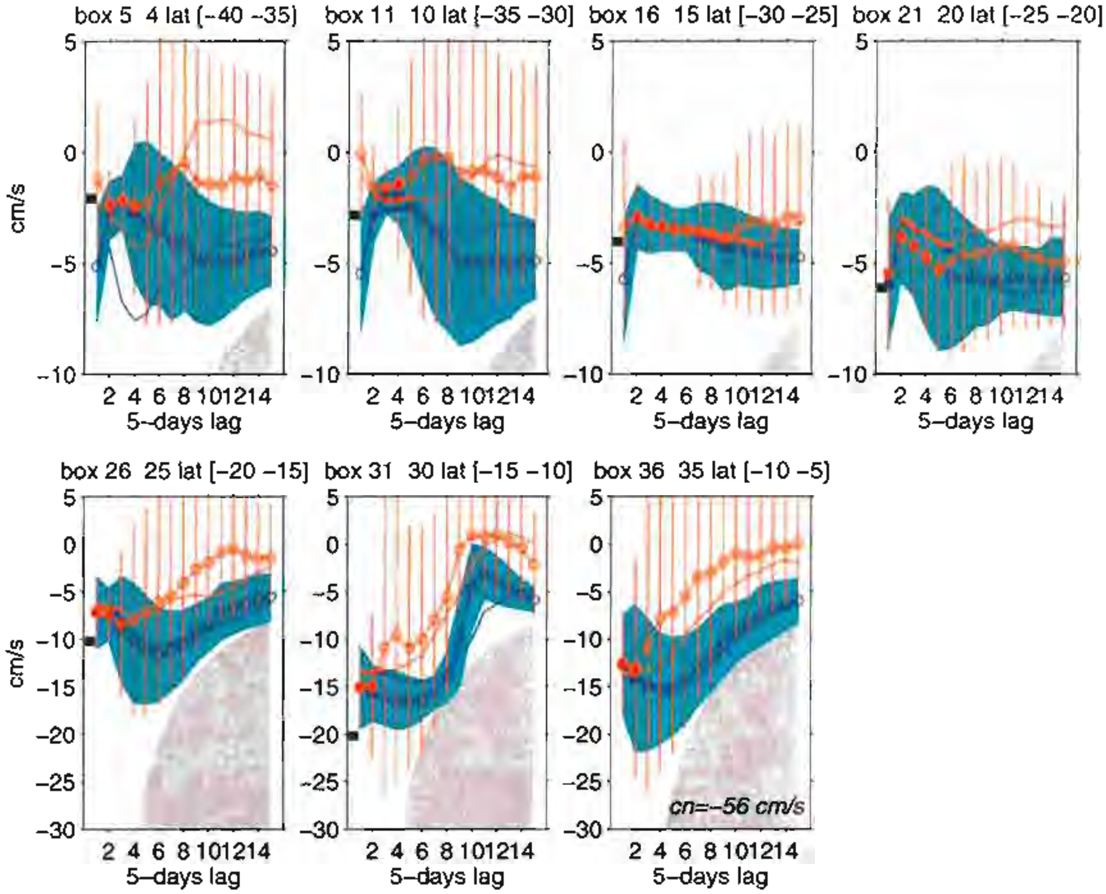


Figure 7.2: Each panel shows the east-west propagation speeds for two neighboring boxes in the east-west direction (see figure 6.10). The abscissa indicates the time-lags used to calculated the lag-correlation. The red/blue circles with red error bars or blue shading are mean and STD of propagating speeds over three years for the first box. The red/blue lines are the means for the second box. The means associated with blue circles or blue lines are calculated from the westward propagation speeds only. Therefore, coincidence of two means for the same box is a strong indication of the robustness of the calculation. Those means are highlighted by the closed circle and thick lines. Note there is no such coincidence for box 4. The close squares near the ordinate indicate the speed of the long Rossby waves. In the last panel, the speed of the long Rossby wave is out of the axis range, so we show the speeds in the plot. The gray shading in each panel is explained in the text.

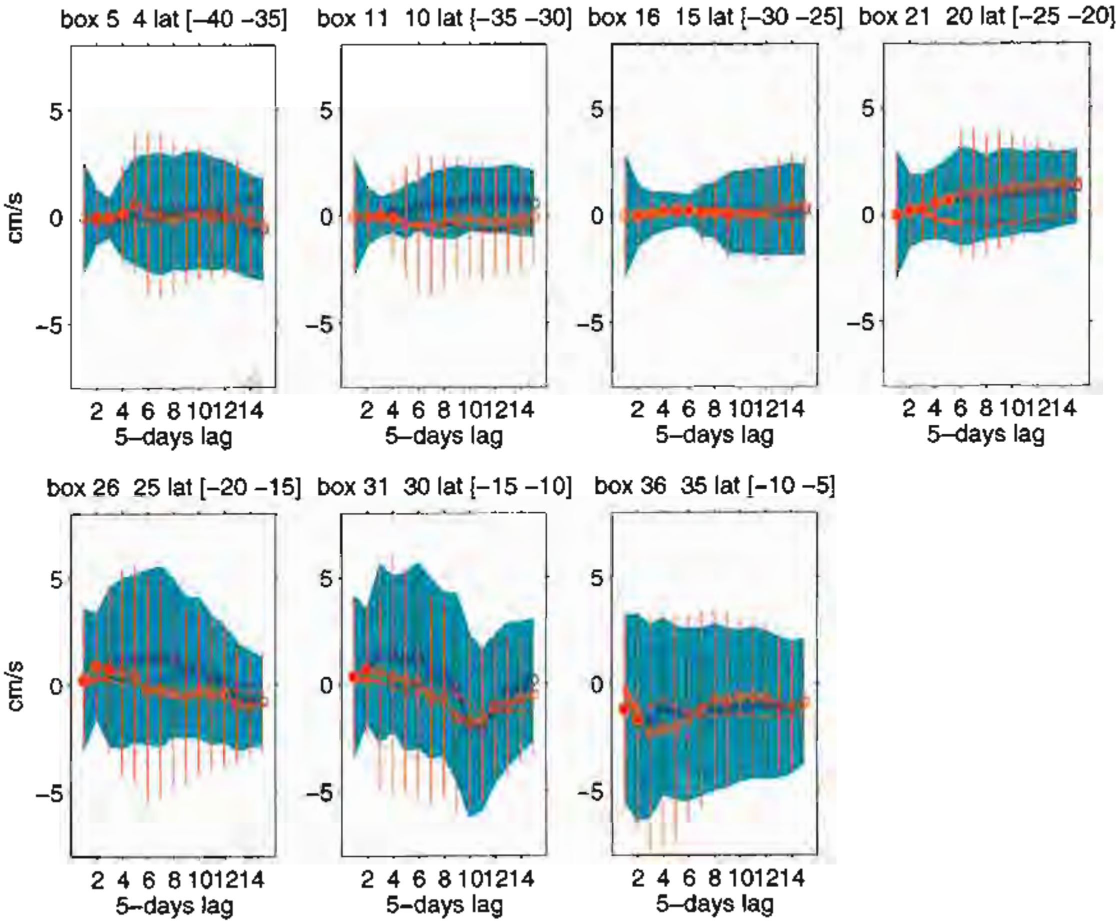


Figure 7.3: Same as figure 7.3 but for the north-south propagation speed. The close circles and thick lines are based on figure 7.2.

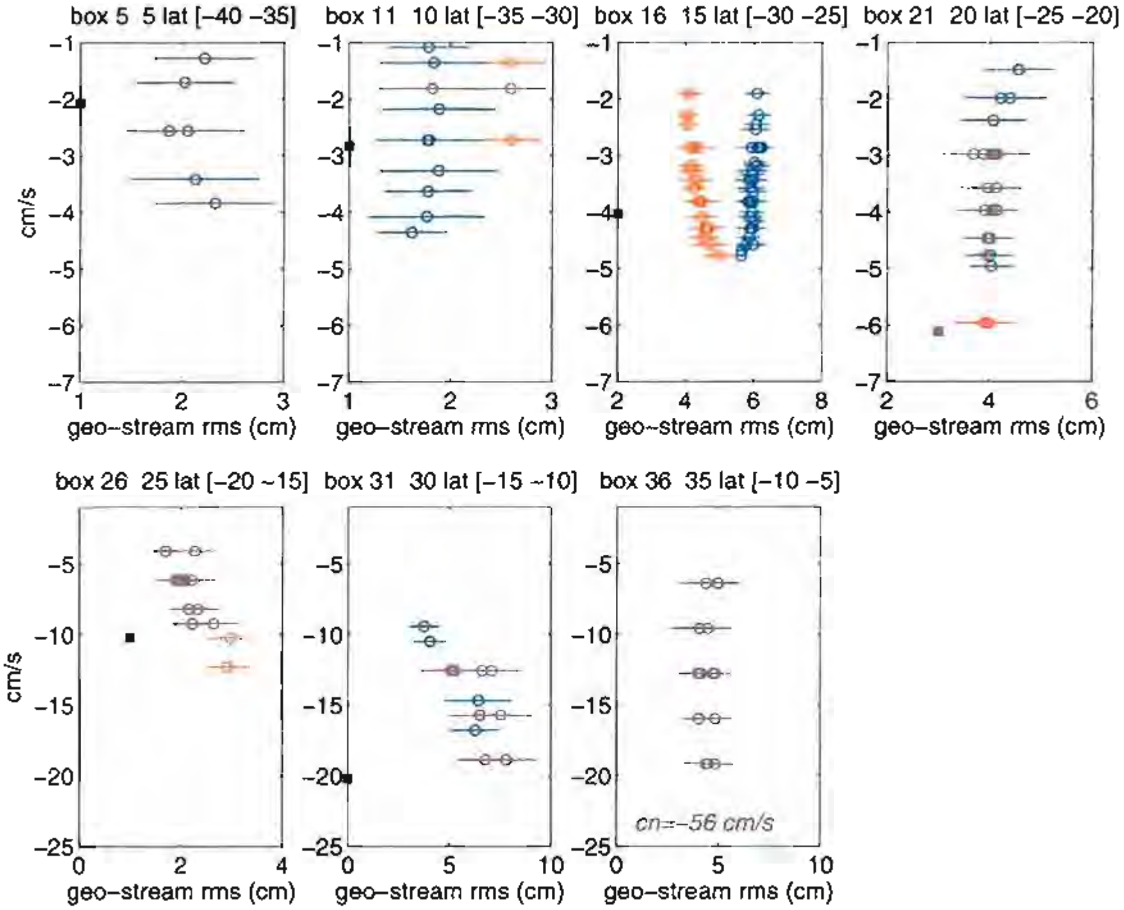


Figure 7.4: Only those results with closed circles and thick lines in figure 7.2 are further analyzed here. The red is for the first box and the blue for the second. The *rms* amplitude is defined as the STD of one snapshot ψ_{EOF} over space. For each time lag with closed circle or thick line, we have a time series of inferred propagation speed and a time series of the corresponding *rms* amplitude. Then we bin the propagation speeds around 20 discrete speeds (20 discrete grids west to the origin divided by the lag time) and group the *rms* amplitude correspondingly. Next the average speed of each bin and average amplitude of the corresponding group is plotted. Some propagation speed bins have more than one amplitude because they are from different time-lags.

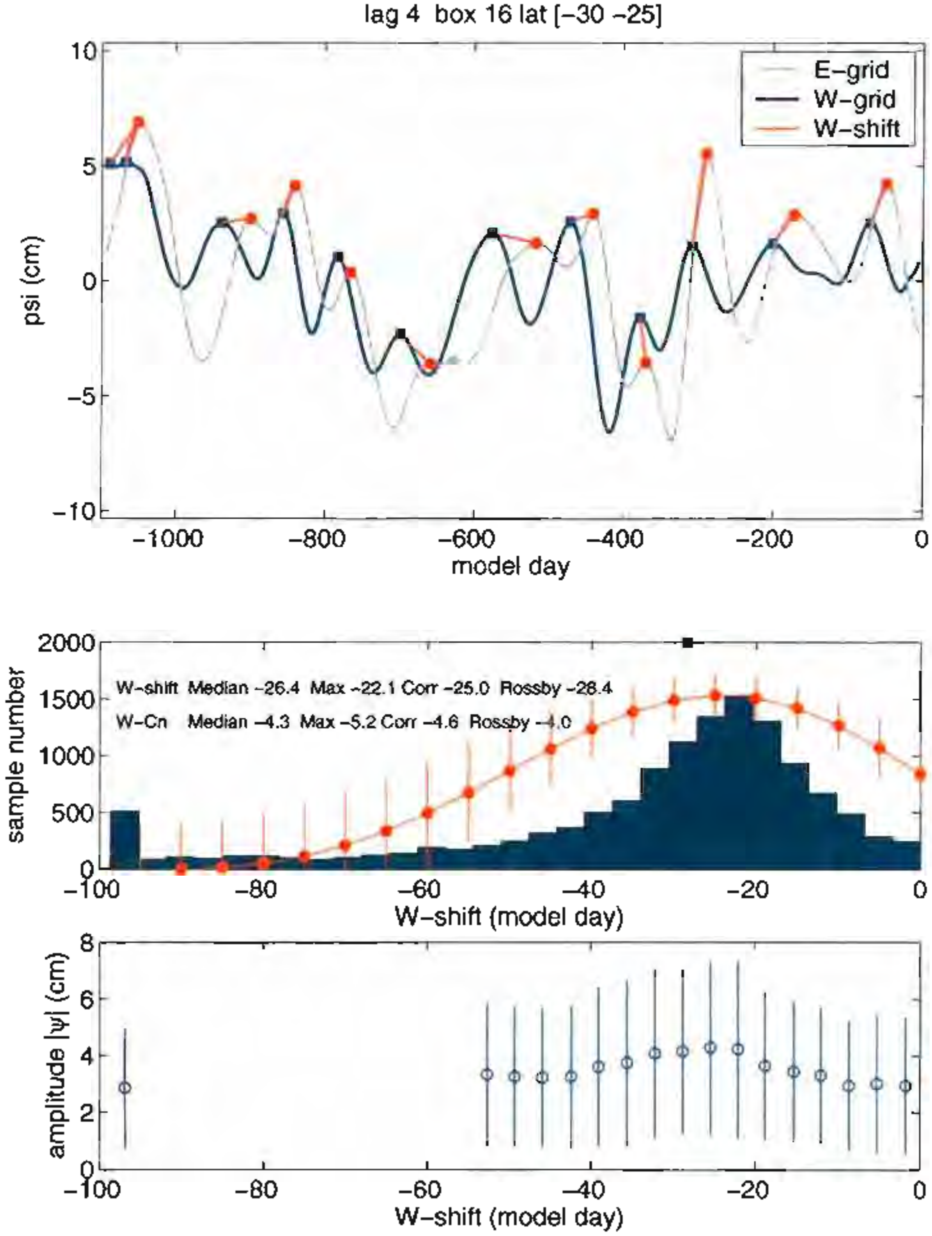


Figure 7.5: The counting method for ψ_{EOF}. For two three-year series (blue and gray lines in upper panel, one degree longitude apart): 1) perform 35 days running-mean; 2) pair each peak in blue (west) time series with the peak in gray (east) line at the nearest earlier time; 3) record the time difference of those two peaks, and the amplitude of the blue peak. For each box there are 10×21 such three-year series pairs. The statistics for all pairs are shown in the middle and lower panels. In the middle panel, the maximum samples in one bin of the histogram is required to be less or equal to 1500. The red line is the mean and STD of the lag-correlations of all available pairs. In the lower panel, the peak amplitudes are averaged over the same bins as in the middle panel; the error bars are STDs.

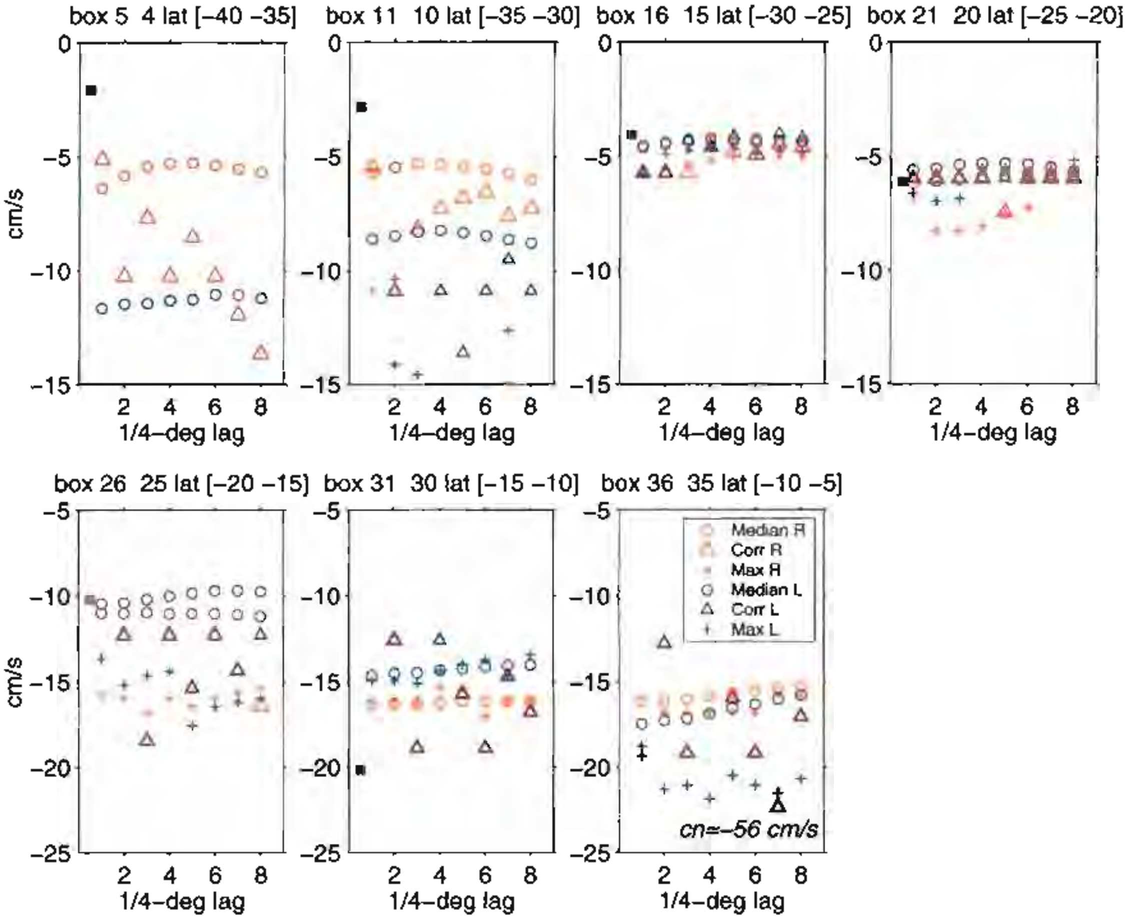


Figure 7.6: The statistics based on those in the middle panel of figure 7.5 for different boxes and different time lags. In each panel, the boxes and corresponding latitude range are in the subtitle. The red color is for the first box and the blue for the second box. The legend is in the last panel.

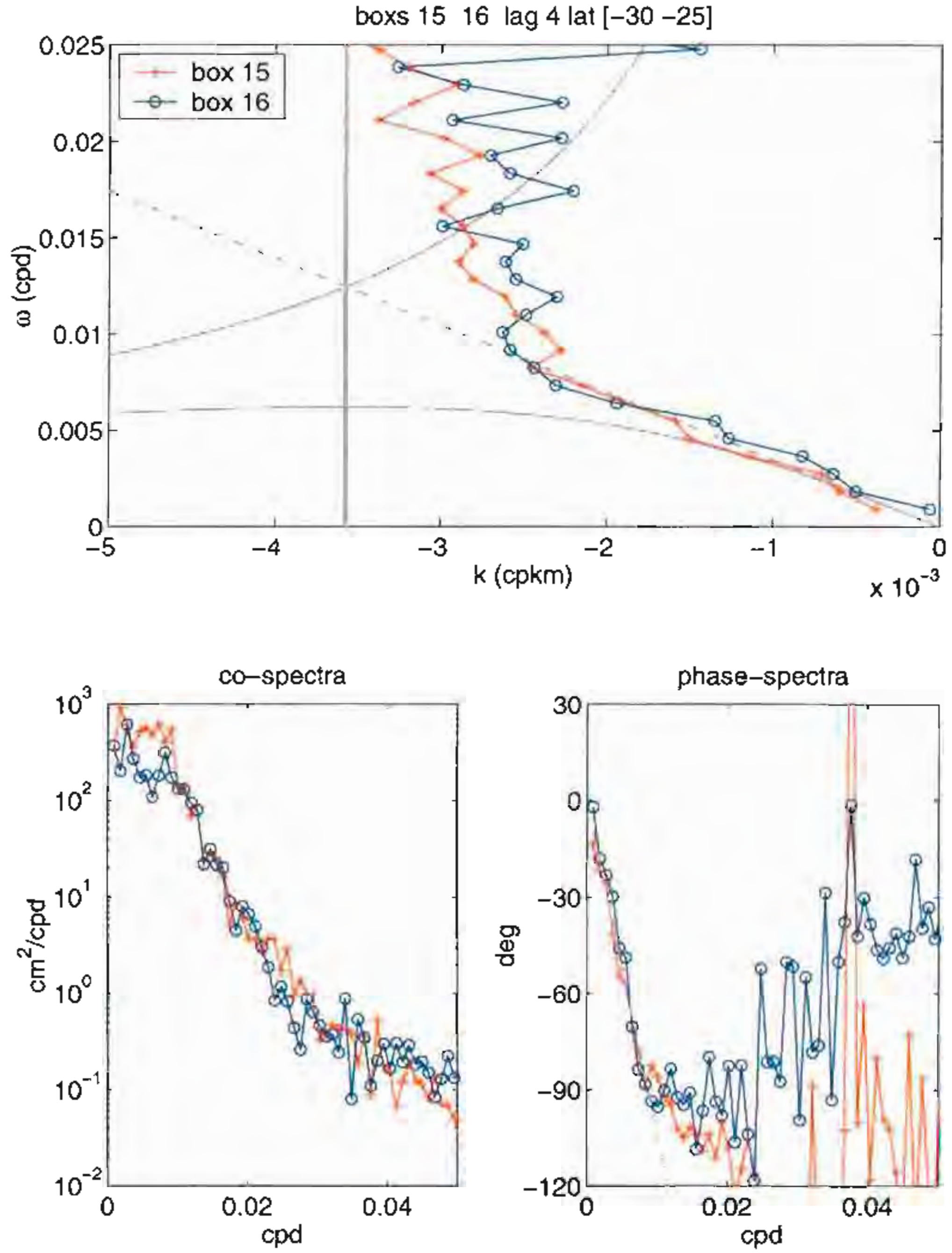


Figure 7.7: One cross spectrum is obtained from any three-year time series pair, one degree longitude apart. The final cross-spectra is averaged over all possible pairs. The lower-left panel is the co-spectra and lower-tight panel is the phase spectra. The red is for the first box (box 15) and blue for the second (box 16). The phase in the lower-right panel is converted into wavelength as explained in the text. The wavelength-frequency pairs are then plotted in the upper panel. The vertical line in the upper panel corresponds to $\frac{1}{2\pi R}$, where R is the first Rossby radius of deformation. The dash line in the lower-right panel show the phase difference if the waves are all propagating at the speed of long Rossby waves.

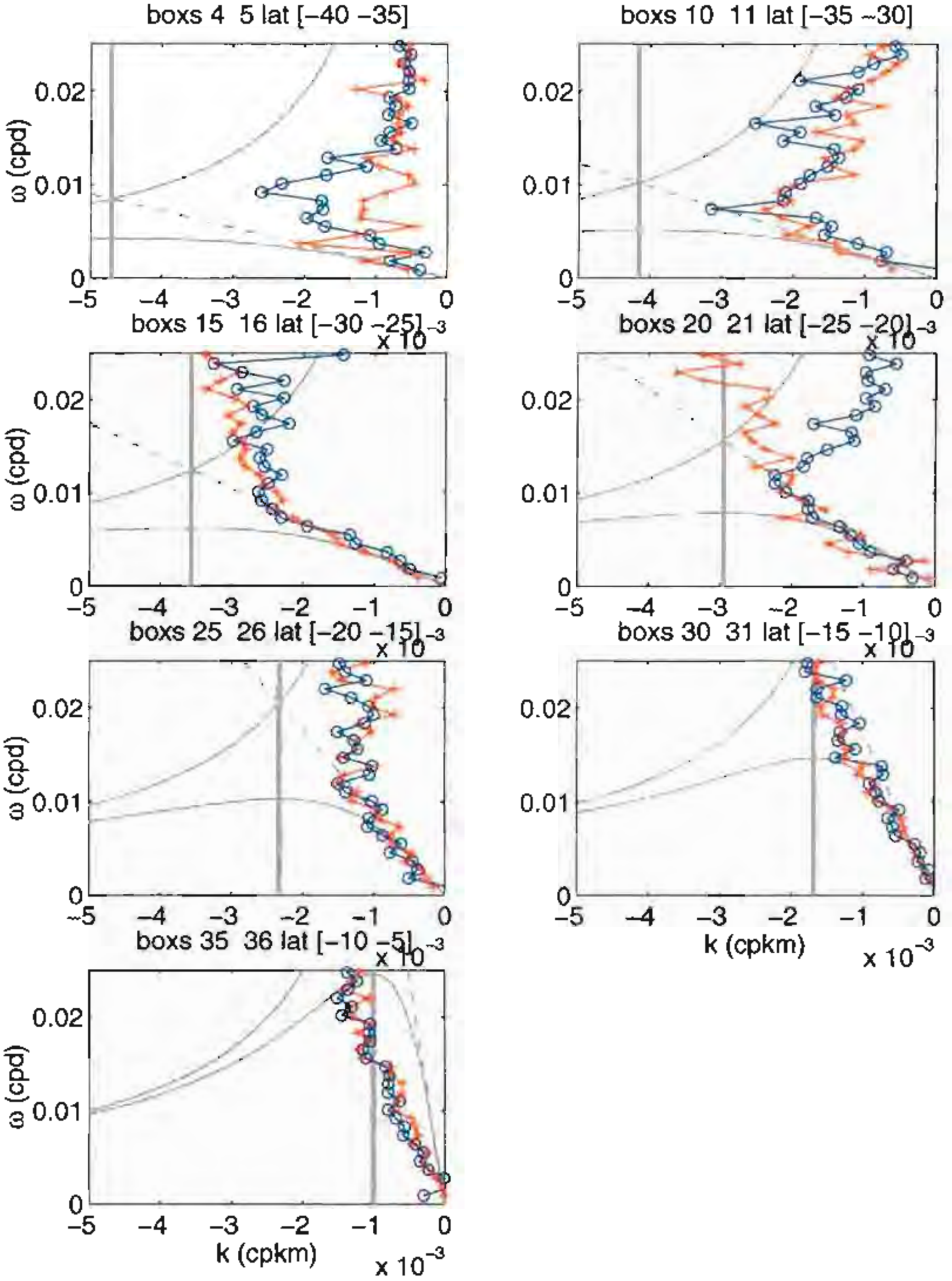


Figure 7.8: Same as the upper panel of figure 7.7 but with more boxes.

CHAPTER 8

BEYOND THE FIRST EOF

The analysis of the LADCP profiles convinced us that the first EOF could represent the vertical structure better than either the barotropic or the first baroclinic modes alone. In the JAMSTEC model, the first EOF contains 50-80% of the variance The unidirectional and surface-intensified vertical structure of the first EOF is similar across most of the eastern Indian Ocean.

Then, are the second and higher EOFs are totally negligible? The four dimensional field $\psi(x, y, z; t)$ is exactly expressed as

$$\psi(x, y, z; t) = \sum_{j=1}^{M} F_j(z) a_j(x, y; t),$$

where F_j represents the vertical structure of the jth EOF, and M is the total number of EOFs which is 55 for 55 levels of the model output. The questions are how important the EOFs greater than 2? Is it satisfactory to simply use the first EOF to represent the original field? For example,

$$\psi(x, y, z; t) \approx F_1(z) a_1(x, y; t). \tag{8.1}$$

Dynamically, can we seek a solution, which is separable with respect to z as in equation 8.1 or is $F_1(z)$ a dynamical mode (i.e. is it an eigensolution of governing equations)? We do not discuss dynamics here. That is the subject of the theoretical part of this study. Here we will test whether equation 8.1 is consistent with model output.

Equation 8.1, as all other separable expressions, implies the currents at different depths are either in phase or 180° out of phase, depending on the sign of $F_1(z)$. Since the first EOF is almost unidirectional, e.g., it has same sign over depth, we may expect model in phase over the entire water column if equation 8.1 holds. Clearly this is not true; the time series of geostrophic profiles (figure 6.2) or the sectional contours (figure 6.1) show the lower layer phase leads the upper layer for west propagation.

To quantify this statement, the wavenumber cross-spectra between different depths are calculated (figure 8.1). Figure 8.1 shows the lower layer (3882 m) leads the upper layer (190 m) by 90°, given westward propagation. After computing spectra over more depths, we find the phase lead is gradual (figure 8.2).

Figure 8.3 shows the vertical correlations of all boxes. One general characteristic is that there are higher correlations at wavenumbers less than the first Rossby radius of deformation. We see three different regions. One is near the equator, which exhibits almost no vertical correlation. One is in the ACC, where high vertical correlations extend towards the bottom. In most regions, the correlation decays towards the bottom, but is significant for wavenumbers less than the first Rossby radius of deformation.

Figure 8.4 shows the phase differences between different levels. It is a robust feature that the lower layer leads the upper layer. In the ACC region, the phases leads are small, but in most regions the phase leads are approximately 90°.

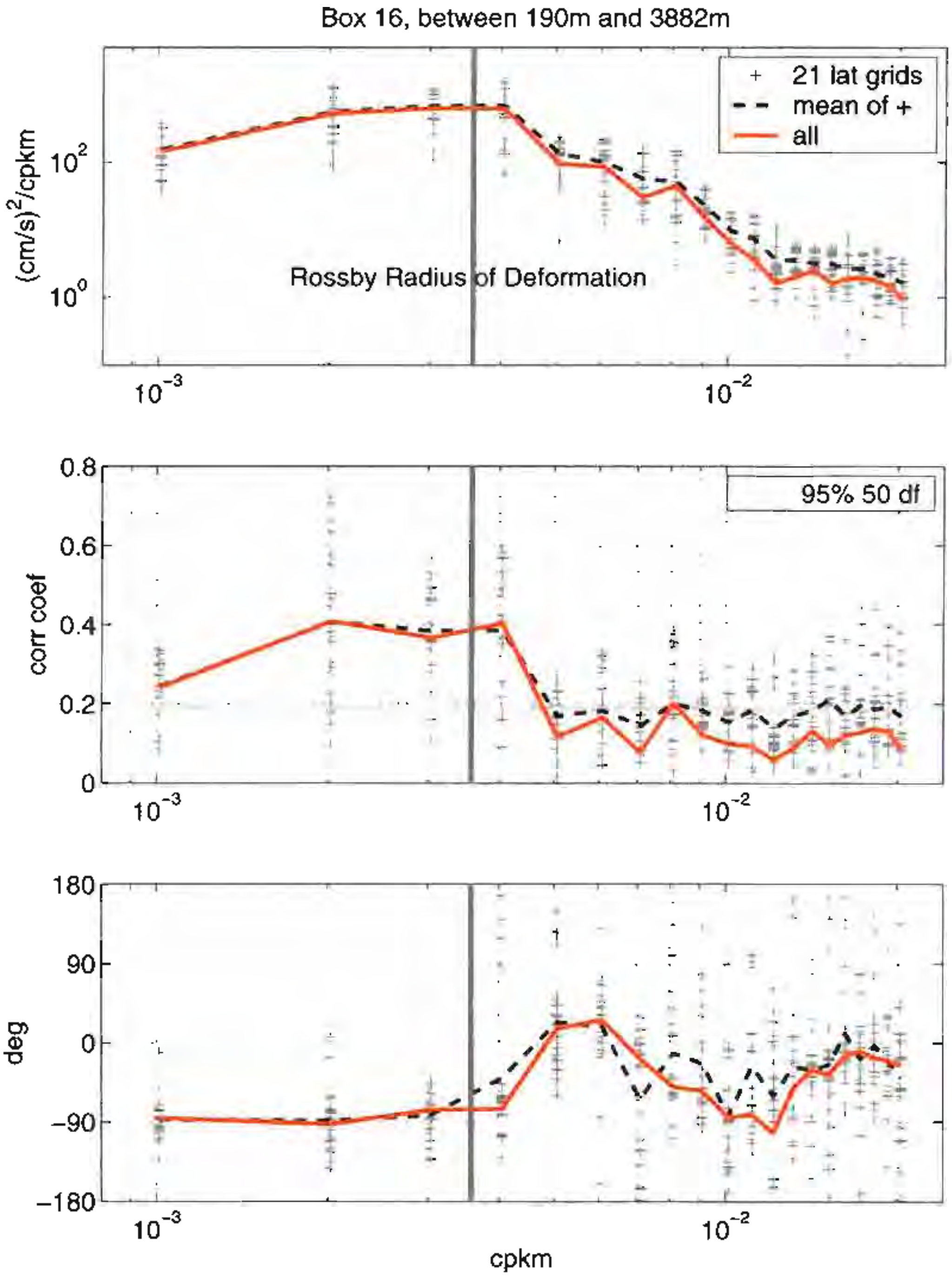


Figure 8.1: Upper: co-spectra between depths 190 m and 3882 m. Middle : correlation coefficients at each wavenumber. Lower: phase-spectra. The red lines are the averages over time and latitude together, and the crosses are the average over time for each latitude line. The dashed lines are the average of those crosses. The gray line in the middle panel indicates the 95% confidence level assuming 50 degrees of freedom. The vertical lines mark the wavenumber corresponding to the first Rossby radius of deformation.

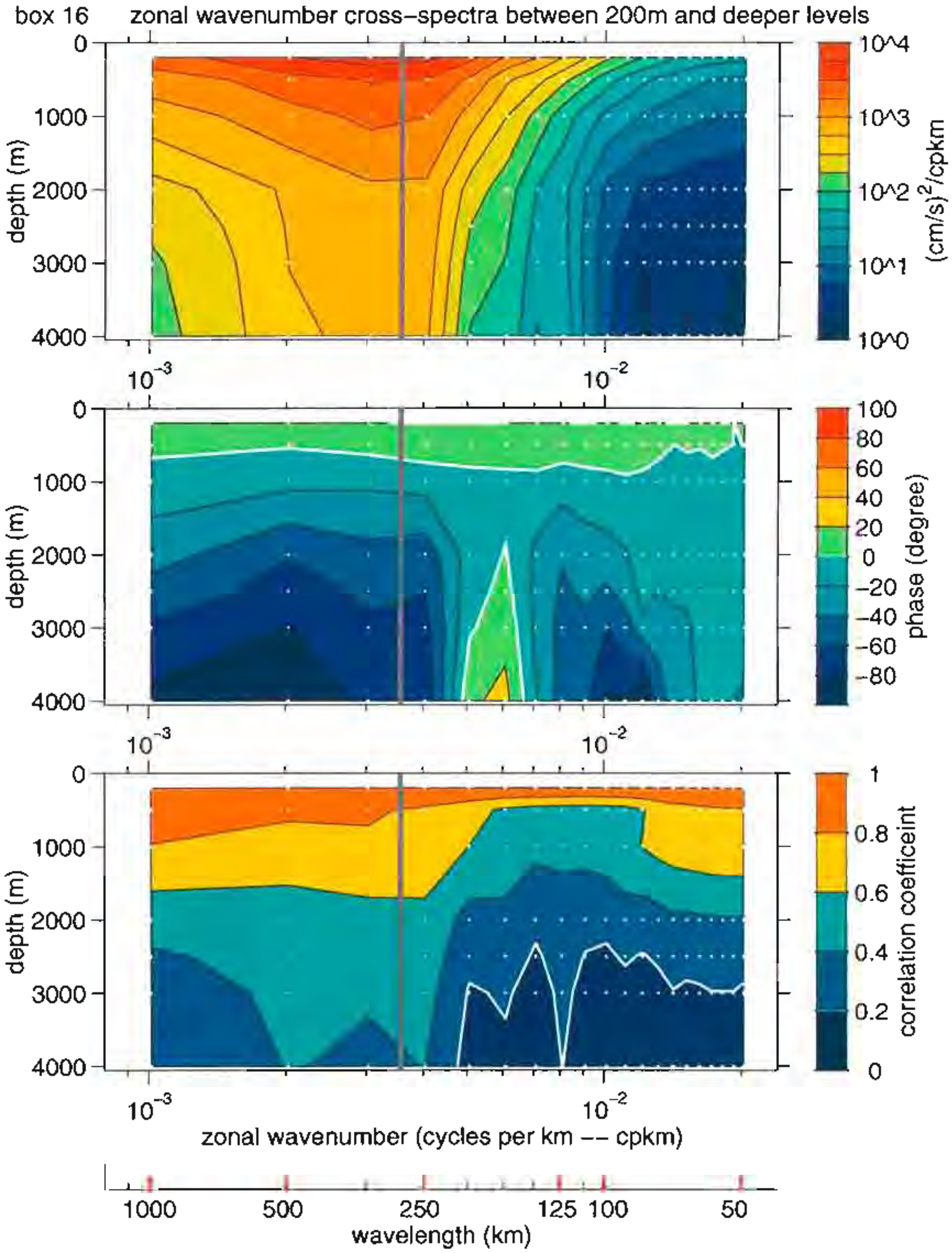


Figure 8.2: Same as the red lines in figure 8.1 but including other depths.

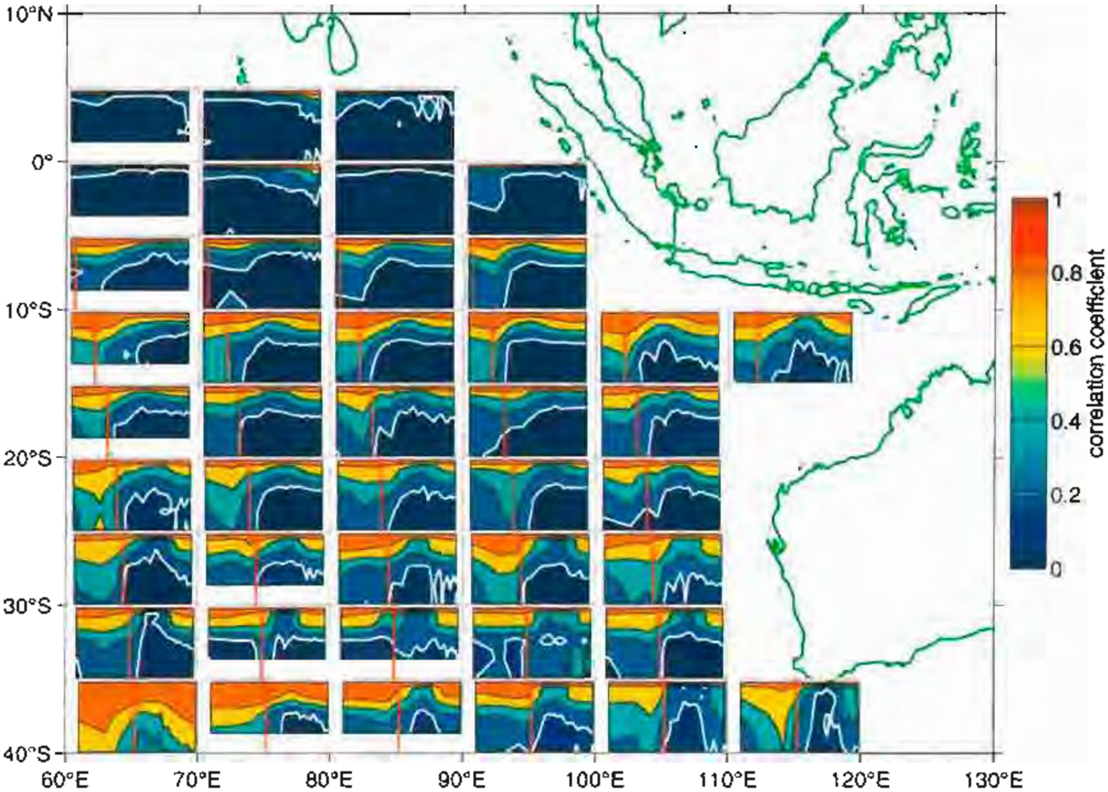


Figure 8.3: Same as the lower panel of figure 8.2 but including all the boxes. The white contours are 0.2, which indicates the correlation is under the 95% confidence level. The contours change color every 0.2.

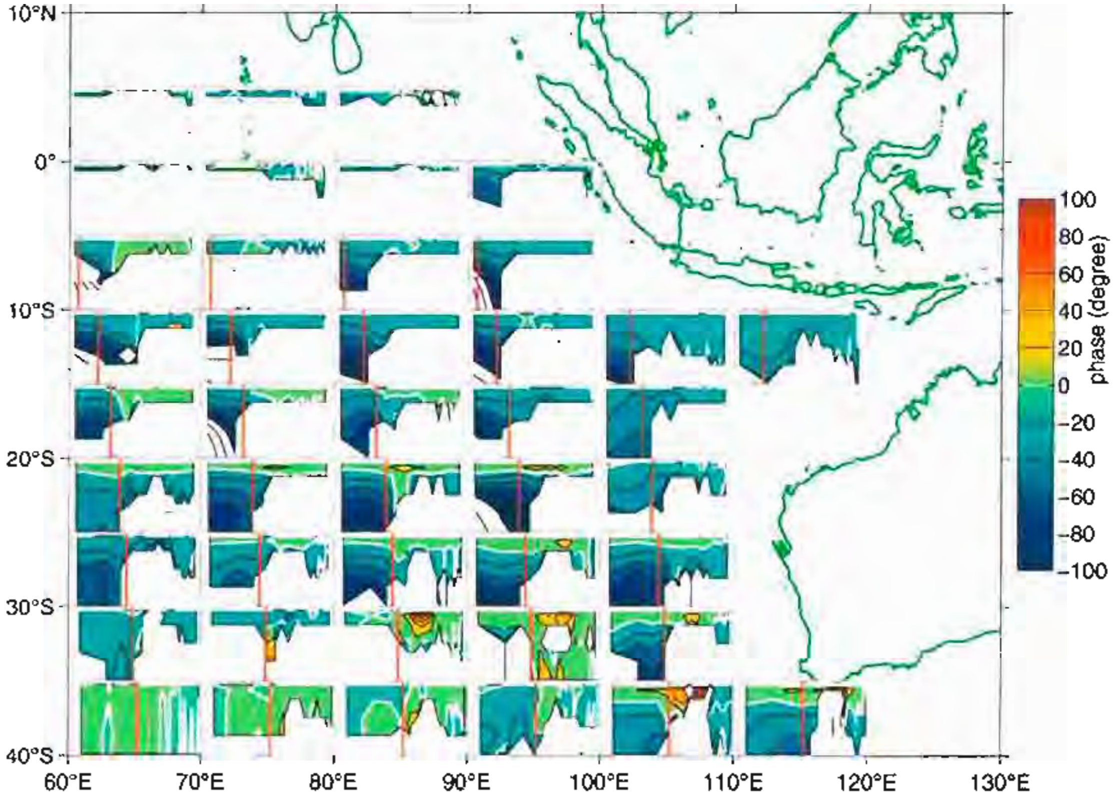


Figure 8.4: Same as the middle panel of figure 8.2 but including all the boxes. Only those contours with correlation coefficients higher than 0.2 are shown. The color changes every 20°. If high correlation coefficients are out of the range [-100° 100°], the contours shown have interval of 20°.

Part III

THEORY: TWO-LAYER MODELS

CHAPTER 9

LINEAR ROSSBY WAVES: DISPERSION OF A GAUSSIAN EDDY

Starting with a Gaussian perturbation (eddy) in the upper layer and a resting lower layer, the two layer model evolves according to the linear Rossby wave dynamics without any forcing and dissipation. The dynamics of the model are obvious; our exercise will focus on one aspect of the kinematics: dispersion. One usually thinks about the dispersion of waves of one particular vertical mode with a spectrum of horizontal wavelengths: its different Fourier components will propagate at different speeds. Here we will emphasize the dispersion among different vertical modes, namely the barotropic and first baroclinic modes in a two-layer model. To quantify the dispersion, the correlations between the two modes in horizontal space are calculated as functions of time. This correlation has been used an important index to demonstrate that the mesoscale motion measured by LADCP profiles deviates from the linear Rossby wave theory.

9.1 The Two-layer Model

Besides f and β, the Coriolis parameter and its latitudinal derivative, the parameters in a two-layer model are layer mean depths H_1 (upper) and H_2 (lower), and the reduced gravity g'. Based on these parameters, one can define length and time scales for the Rossby waves:

$$R = \frac{\sqrt{g'H_1}}{f}, \qquad T = \frac{1}{\beta R},$$

and a depth ratio δ:

$$\delta = \frac{H_2}{H_1}.$$

When we nondimensionalize by these length and time scales, the resulting linear, quasi-geostrophic and inviscid two-layer model is:

$$\begin{cases} \frac{\partial}{\partial t}[\nabla^2\psi_1 + (\psi_2 - \psi_1)] + \frac{\partial \psi_1}{\partial x} & = & 0 \\ \frac{\partial}{\partial t}[\nabla^2\psi_2 + \frac{1}{\delta}(\psi_1 - \psi_2)] + \frac{\partial \psi_2}{\partial x} & = & 0, \end{cases} \tag{9.1}$$

where ψ_1 and ψ_2 are the geo-potentials of the upper and lower layers respectively. Notice that a rigid-lid and flat bottom are assumed in the above equations, so that the stretching effect relates only to the interface perturbation, which is proportional to $\psi_1 - \psi_2$. The initial condition is

$$\begin{cases} \psi_1 & = & \exp(-\frac{x^2+y^2}{2L^2}) \\ \psi_2 & = & 0 \end{cases} \quad \text{when} \quad t = 0$$

This axisymmetric eddy has been used previously (e.g., Chassignet and Cushman-Roisin 1991). The azimuthal velocity v_θ for the axisymmetric eddy is:

$$v_\theta = \frac{d\psi_1}{dr} = -\frac{r}{L^2}e^{-\frac{r^2}{2L^2}}$$

and its radial gradient is

$$\frac{dv_\theta}{dr} = \frac{1}{L^2}(\frac{r^2}{L^2} - 1)e^{-\frac{r^2}{2L^2}}$$

so that v_θ reaches its maximum at r=L, where $r = \sqrt{x^2+y^2}$. The factor 2 in the exponential is usually included so that the azimuthal velocity has its maximum at $r = L$. As we will see, some results depend sensitively on the magnitude of L. For Gulf Stream Rings, as the examples of large eddies, the measurements suggest that L is about 3 (Dewar and Gailliard 1994). We will often choose L=1, 2 or 3. Also notice that the length scale is the first Rossby radius of deformation, which varies with the latitude, so that for the same size eddies, L will be larger in high latitudes where R is smaller.

As seen in equation 9.1, the upper and lower layers are coupled. However, the barotropic and baroclinic modes of the motion are not. The dynamical normal modes for the two layer model are

$$\text{barotropic} \quad \begin{pmatrix} 1 \\ 1 \end{pmatrix}; \qquad \text{baroclinic} \quad \begin{pmatrix} \sqrt{\delta} \\ -\frac{1}{\sqrt{\delta}} \end{pmatrix}$$

whose amplitudes (ψ_T and ψ_c) are governed by

$$\begin{cases} \frac{\partial}{\partial t}[\nabla^2\psi_T] + \frac{\partial \psi_T}{\partial x} & = & 0 \\ \frac{\partial}{\partial t}[\nabla^2\psi_c - (1+\frac{1}{\delta}))\psi_c] + \frac{\partial \psi_c}{\partial x} & = & 0 \end{cases}$$

The relationships between the mode amplitudes and layer amplitudes are

$$\begin{cases} \psi_T & = & \frac{\psi_1+\delta\psi_2}{1+\delta} \\ \psi_c & = & \frac{\sqrt{\delta}}{1+\delta}(\psi_1 - \psi_2) \end{cases} \quad \text{and} \quad \begin{cases} \psi_1 & = & \psi_T + \sqrt{\delta}\psi_c \\ \psi_2 & = & \psi_T - \frac{\psi_c}{\sqrt{\delta}}. \end{cases}$$

The initial condition becomes

$$\begin{cases} \psi_T & = & \frac{1}{1+\delta}\exp(-\frac{x^2+y^2}{2L^2}) \\ \psi_c & = & \frac{\sqrt{\delta}}{1+\delta}\exp(-\frac{x^2+y^2}{2L^2}) \end{cases} \quad \text{when} \quad \text{t} = 0$$

The solutions are obtained in appendix C:

$$\begin{aligned} \psi_T &= \frac{1}{1+\delta}\frac{L^2}{2}\int_0^\infty d\mathcal{Z}e^{-\frac{\mathcal{Z}L^2}{2}} J_0\left(\sqrt{\mathcal{Z}}\sqrt{(\frac{t}{\mathcal{Z}}+x)^2+y^2}\right) \\ \psi_c &= \frac{\sqrt{\delta}}{1+\delta}\frac{L^2}{2}\int_0^\infty d\mathcal{Z}e^{-\frac{\mathcal{Z}L^2}{2}} J_0\left(\sqrt{\mathcal{Z}}\sqrt{(\frac{t}{\mathcal{Z}+1+1/\delta}+x)^2+y^2}\right) \end{aligned}$$

where J_0 is the zero-order Bessel function of the first kind.

9.2 Dispersion

The linear Rossby waves are dispersive. We will show that the dispersion between barotropic and baroclinic modes is different from that between different horizontal wavenumbers of the same mode.

We first show the x-t Hovmüller diagram, in terms of baroclinic and barotropic modes and in terms of the upper and lower layer streamfunctions, for one depth ratio, $\delta = 5$ but three different values of L (figures 9.1-9.3). Initially, the amplitude of the baroclinic mode is $\sqrt{\delta} = \sqrt{5}$ times as large as that of the barotropic mode. Thus the upper layer is mostly dominated by the baroclinic mode, while in the lower layer the two modes cancel each other. After the barotropic mode has dispersed, both layers are dominated by the baroclinic mode.

The barotropic waves usually disperse faster than the baroclinic waves. The dependency of the dispersion on the size of the initial Gaussian eddy is different for the barotropic and baroclinic waves. The larger the initial eddy (L increases from 1 to 3), the slower the dispersion of the baroclinic waves but the faster the dispersion of the barotropic waves (see figures 9.1-9.3).

Next, we show the evolution in a 2-dimensional view for L=1 and $\delta = 5$ (figures 9.4 and 9.5). The salient features are

- Long waves propagate to the west, so that there is no sign change west of the original location (x=0). Shorter waves propagate to the east. This is true for both barotropic and baroclinic modes;
- Symmetry with respect to the north-south coordinate (y);
- As shown in x-t Hovmüller diagram, the barotropic mode disperses faster than the baroclinic mode. The main core of the barotropic mode is hardly identifiable after t=5, while that of the baroclinic mode can be identified in the last panel;
- For the barotropic mode, we see that there are local extrema away from the x-axis for $t \leq 4.8$.

The correlation between ψ_T and ψ_c at any given time (figure 9.6) is defined

$$\mathcal{R}(t) = \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\psi_T\psi_c dxdy}{\sqrt{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\psi_T^2 dxdy\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\psi_c^2 dxdy}}$$

Obviously $\mathcal{R}(0) = 1$, since at t=0 ψ_c is proportional to ψ_T: $\psi_c = \sqrt{\delta}\psi_T$. Figure 9.6 shows that the larger the initial eddy, the faster the loss of correlation between the barotropic and baroclinic modes.

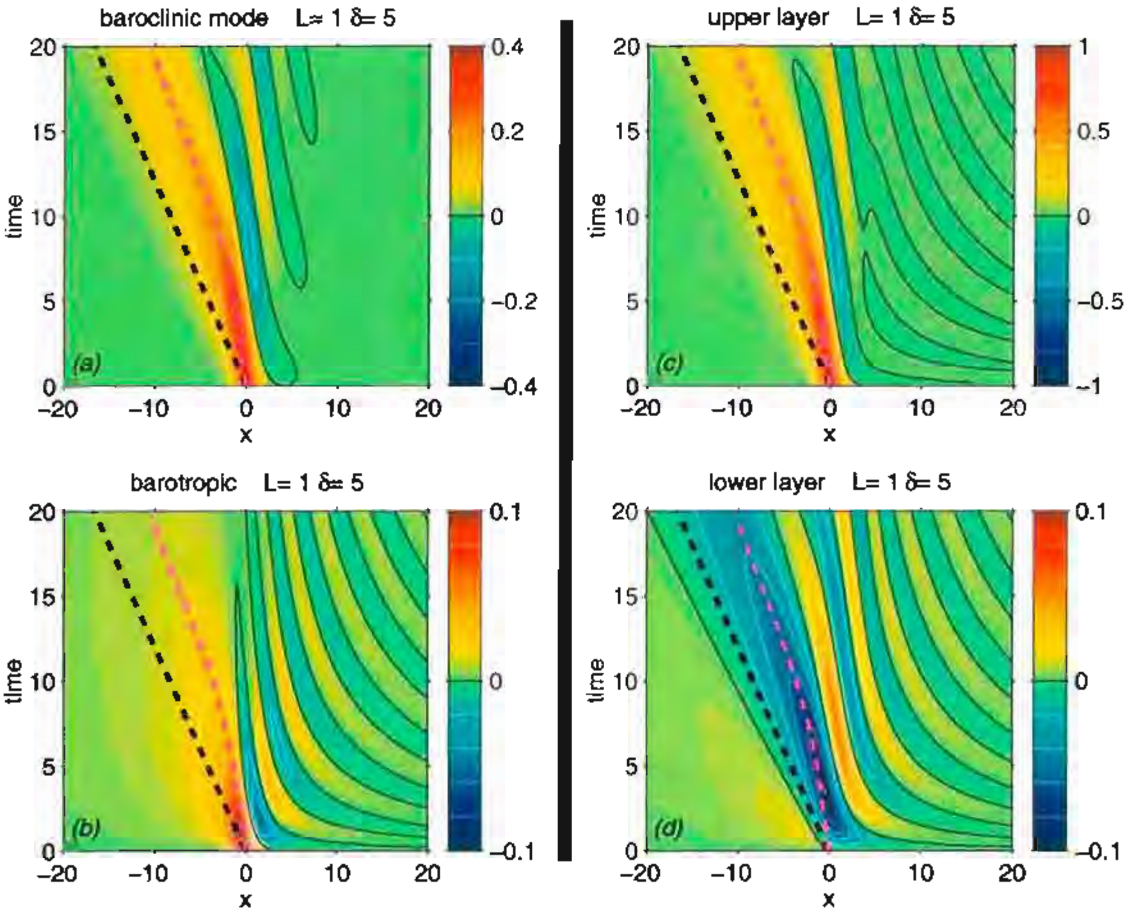


Figure 9.1: Dispersion of the Gaussian eddy in the two-layer model (L=1 and δ =5). The left two panels are the baroclinic (a) and barotropic (b) modes, and the right two panels for the upper (c) and lower (d) layers. Notice the color scales change from one panel to another, but the thin black lines always represent the zero contour. The dashed black lines (same for all panels) show the westward propagations of the first mode long Rossby waves, which is actually $-\frac{1}{1+1/\delta}$, and the dashed pink lines (same for all panels) indicate the maximum of the baroclinic mode at each time in panel (a). The "x" is non-dimensionalized by $R = \sqrt{g'H_1}/f$ and the time by $T = 1/\beta R$.

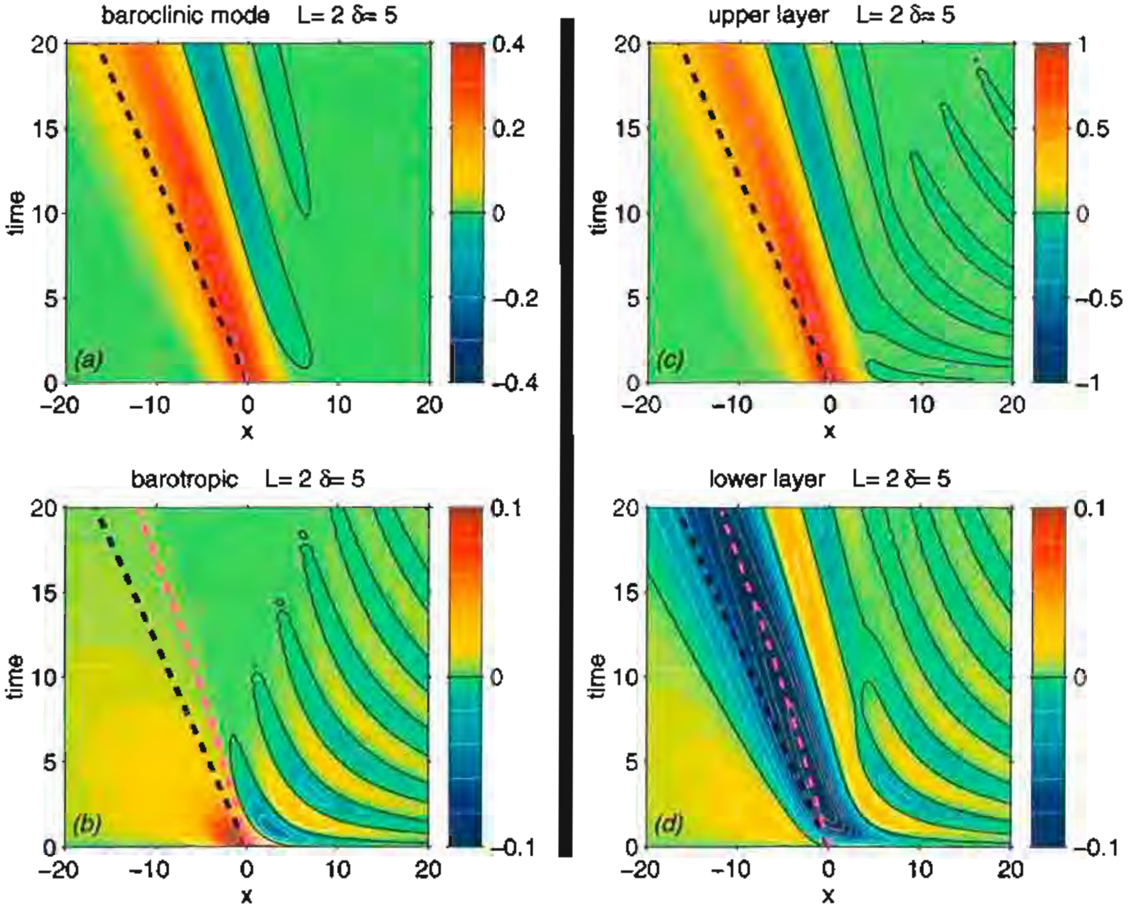


Figure 9.2: Dispersion of the Gaussian eddy in the two-layer model (L=2 and δ =5). Other notations are the same as in figure 9.1.

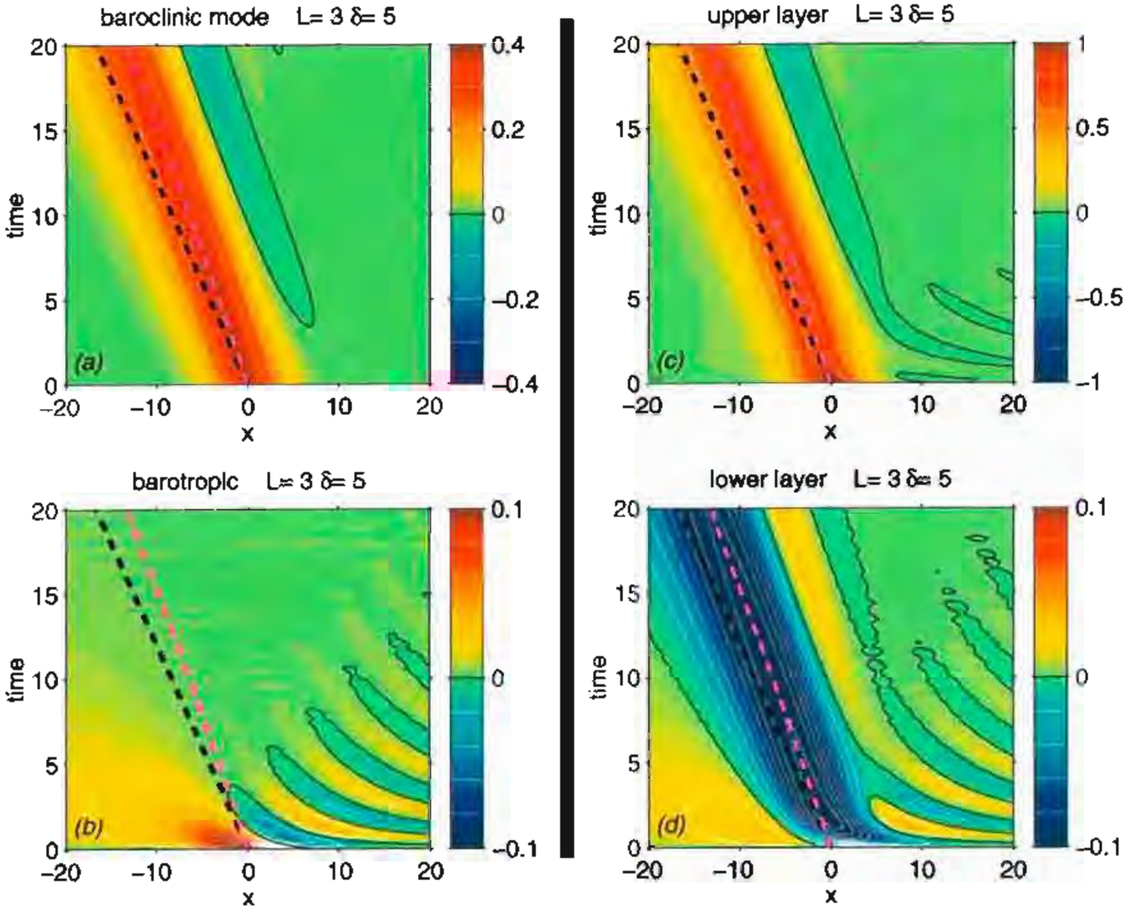


Figure 9.3: Dispersion of the Gaussian eddy in the two-layer model (L=3 and δ =5). Other notations are the same as in figure 9.1.

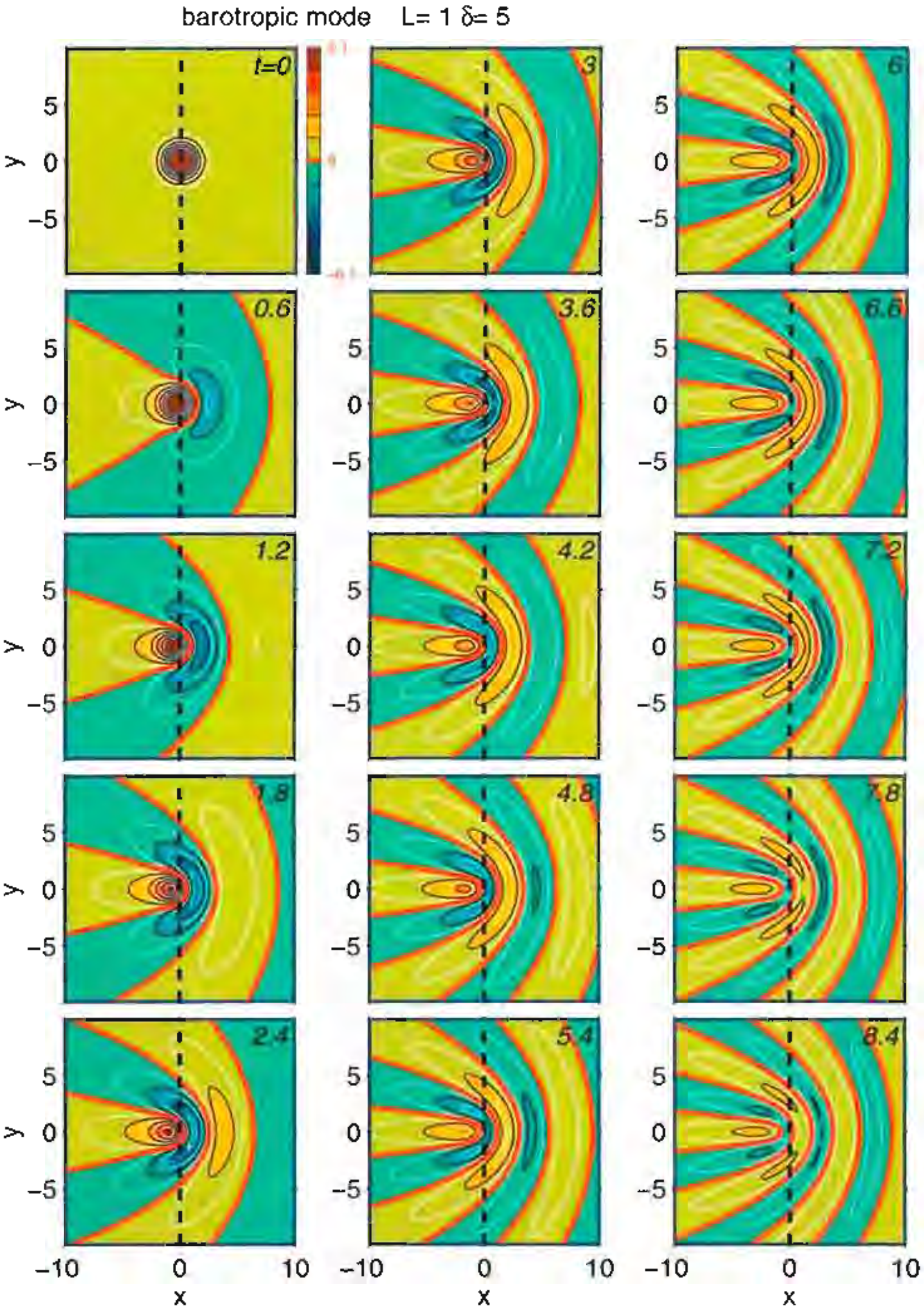


Figure 9.4: 2-D temporal evolutions of initial barotropic waves (L=1 and δ =5). The corresponding x-t plot is in figure 9.1 (b). The time is at the upper-right corner of each panel. The space coordinates x and y are non-dimensionalized by $R = \sqrt{g'H_1}/f$ and the time by $T = 1/\beta R$. All panels use the same color scale as in the first panel: the red lines are zeros, with blue negative and yellow positive; the black lines are $\pm 0.02, \pm 0.04, \cdots$ and the white lines are $\pm 0.01, \pm 0.03, \cdots$.

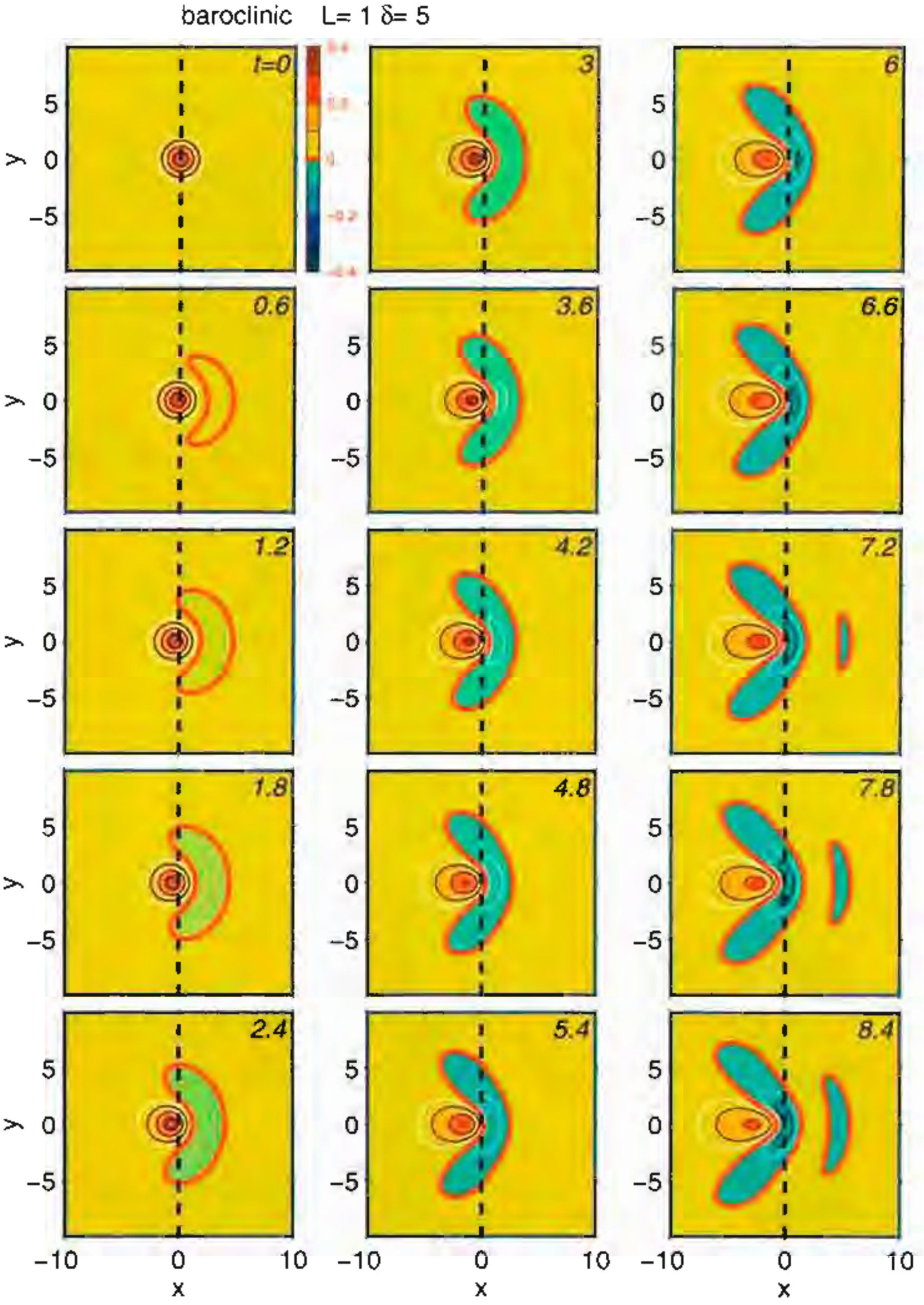


Figure 9.5: 2-D temporal evolutions of initial baroclinic waves (L=1 and δ =5). The corresponding x-t plot is in figure 9.1 (a). The time is at the upper-right corner of each panel. The space coordinates x and y are non-dimensionalized by $R = \sqrt{g'H_1}/f$ and the time by $T = 1/\beta R$. All panels use the same color scale as in the first panel: the red lines are zeros, with blue negative and yellow positive; the black lines are $\pm 0.1, \pm 0.2, \cdots$ and the white lines are $\pm 0.05, \pm 0.15, \cdots$.

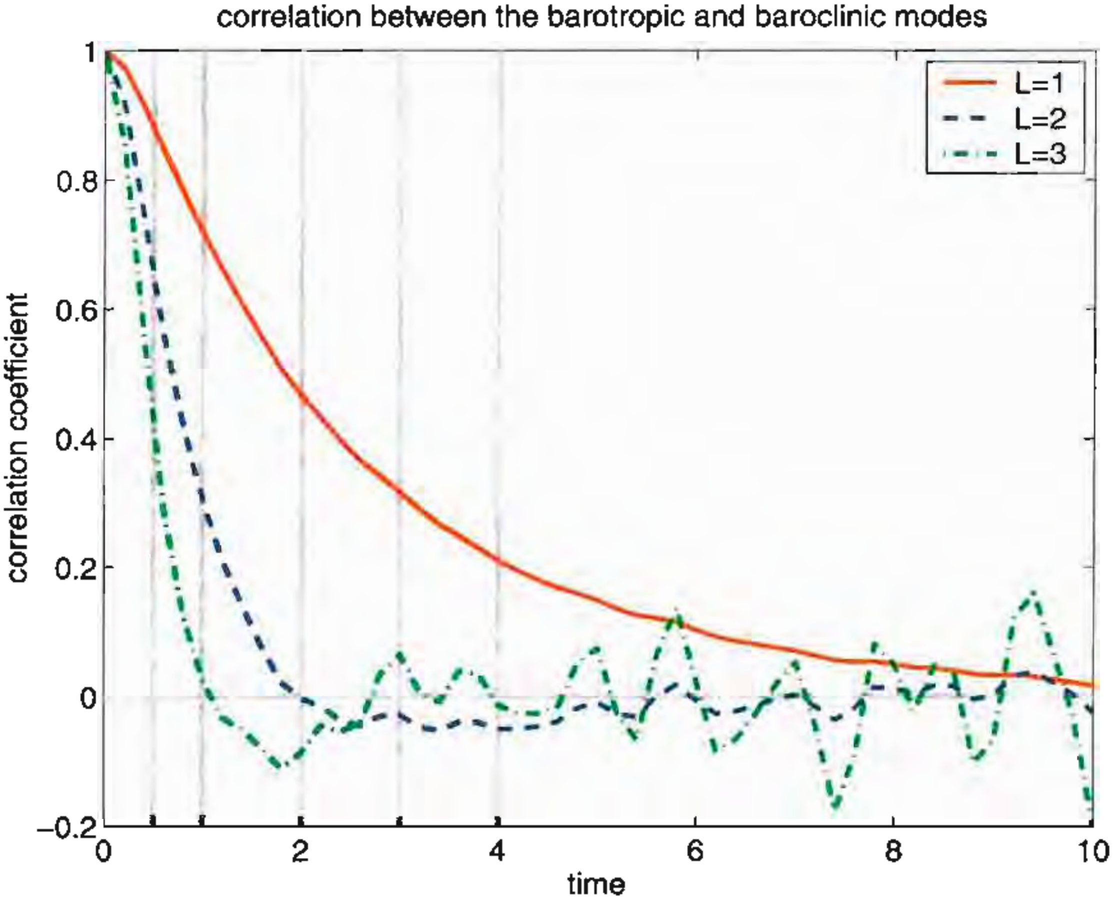


Figure 9.6: Spatial correlation coefficients between the barotropic and baroclinic modes as a function of time. For L=1, the spatial distribution of the mode streamfunctions are shown in figures 9.4 and 9.5.

CHAPTER 10

WESTWARD PROPAGATION

The LADCP profiles in the Indian Ocean show that the vertical structure of the mesoscale ocean currents in mid-latitudes is dominantly unidirectional and surface-intensified. The JAMSTEC model reveals the same feature. The JAMSTEC model further shows that the mesoscale motions associated with the vertical structures are mainly generated near the eastern boundary and then maintain this structure while propagating westward across the basin with a speed approximately equal to that of the long Rossby wave. The model also demonstrates a phase shift in the vertical: the lower layer leads the upper layer for westward-propagating eddies. The question here is why the vertical structure is sustained during its westward propagation. The propagation speed and phase shift in the vertical are two other aspects of the same problem to be addressed at same time. The generation mechanism near the eastern boundary will not be discussed here.

Another manifestation of the dominant vertical structure is the correlation of the barotropic and first baroclinic modes; such correlation could not be explained by the linear Rossby wave theory. Therefore, we need to have a close look at the factors which are not included in the linear theory: 1) non-linearity, 2) mean flow, 3) dissipation and 4) bottom topography. The model shows that the vertical structures propagate westward over the basin without much distortion by local topography. Also the dominant vertical structure tails off toward the bottom. These two features may indicate that the topography does not play a first-order role in the dynamics of the vertical structure. Therefore we will not explore the role of bottom topography in this study. We will investigate factors 1) through 3) in the sections 10.1 through 10.3, after briefly previewing the main issues involved in each.

Vertically sheared mean flow will modify the normal modes into so-called shear modes, although not in the long wave limit ($K \rightarrow 0$). The vertical structures of the shear modes could be similar to the dominant vertical structures, provided that the mean flow shear (defined as the mean flow difference between upper and lower layers) is large enough, compared to the speed of the long Rossby waves. Given the latitudinal dependence of the speeds of long Rossby waves, we conclude that the shear modes would not be significant in middle and low latitudes. Also notice that the mean of the flow is ill-defined if the flow is dominantly mesoscale. The shear modes could be significant in the ACC region, where the mean flow shear is strong and well defined and the long Rossby waves are relatively slow.

We next turn to non-linearity: to seek a solution such as $\psi(x - ct, y, z)$ in the non-linear model. To be analytically tractable, we assume that 1) non-linearity is important in the upper layer of the two-layer model but not in the lower layer; 2) the flow in the upper layer is radially symmetric so that the Jacobian non-linear terms do not appear; 3) the depth ratio is large so that the flow in the lower layer is weak. The non-linearity of the model comes from the interface perturbation whose amplitude is finite with respect to the upper layer thickness. The upper layer solution is a soliton which translates steadily westward, and the lower layer solution is a quasi-steady wave field forced by the upper layer soliton. The quasi-steady wave field is obtained by solving an initial value problem. The vertical structure of the solution is quite sensitive to the size of the upper layer eddy. When the upper layer eddy is small, the vertical structure is somewhat comparable to the dominant vertical structure seen in the observation and model with a phase shift. Before the lower layer reaches a quasi-steady state (we will call it an intermediate solution), the comparison becomes more favorable for all sizes of upper layer eddies. The intermediate solution solution, not sought originally, does not invalidate the assumptions in the non-linear model.

Motivated by two facts: one is the vertical phase shift in the model and another is the existence of bottom Ekman layer in the model, we next investigate the effect of bottom friction on the propagation of Rossby waves. The phase shift in the model, i.e., the lower layer leads the upper layer, implies an upward phase propagation, which in turn implies a downward energy flow in the vertical. The energy consumed by bottom friction could be its sink. The existence of a bottom Ekman layer in the model is demonstrated by the correlation between the vertical velocity and relative vorticity near the bottom, as suggested by the bottom Ekman layer theory.

The bottom friction simply sets up a constant exponential decay rate for all wavenumbers of the barotropic motion. For the baroclinic motion, it is far more complicated. There is always potential energy accompanying the kinetic energy, so the bottom friction is less efficient in decaying the baroclinic mode. The exponential decay rate is wavenumber-dependent: the larger the horizontal scale of the baroclinic motion, the slower the decay rate. Another important perspective is the vertical structures of the modal solutions for Rossby waves with bottom friction. We find that the vertical structure evolving from the baroclinic mode is surface-intensified and tails off towards the bottom, and that the one evolving from the barotropic mode is bottom-intensified. The later is dissipated much faster than the former, so that the bottom friction performs a natural selection of the vertical structure evolving from the baroclinic mode: surface-intensified and unidirectional. The above results certainly depend on the strength of the bottom friction. The question about the importance of bottom friction in the real oceans remains.

10.1 Shear Modes: Interaction with Mean Flows

Interaction with mean flows will modify the behavior of Rossby waves. The effect of the mean flow could be so drastic that the Rossby waves are unstable, which barotropic instability resulting from the horizontal shear of the mean flow, and baroclinic instability resulting from the vertical shear. However, our attention is directed to the vertical structure of stable Rossby waves with a mean flow. The calculation by Killworth *et al.* (1997) demonstrates that, based on a global hydrographic dataset, Rossby waves interacting with the east-west geostrophic shears are mostly stable in the world oceans, with only a few exception.

Mainly motivated by altimetric observations, the effect of a mean flow on Rossby waves has been investigated extensively in continuous models (e.g., Killworth *et al.* 1997) and in layered models (e.g., Dewar 1998). However, all of these studies have used the long wave approximation. As will be shown, under the long wave approximation the mean flow will has no effect on the vertical structure of Rossby waves in a two-layer model.

The vertical structures of stable Rossby waves with a mean flow are called shear modes. The shear modes could certainly be traced back to the normal modes when the mean flow gets small, and are generally not orthogonal. The present formalism is the same as that in Pedlosky (1987), but his emphasis is on baroclinic instability. The main conclusions from the simplest model, a two-layer quasi-geostrophic model, are:

- In the long wave limit (wavenumber $K = \sqrt{k^2 + l^2} \to 0$), a mean flow has no effect on the vertical structure, that is, the vertical structures remain those of the normal modes. The effect of the mean flow is a Doppler shift only, which is equal to the barotropic component of the mean flow—There is no Doppler shift due to the baroclinic component (so-called non-Doppler effect). Away from the long wave limit, there will be a Doppler shift even though the mean flow is baroclinic (provided the two layers differs in thickness).

- The following parameter naturally emerges in the expression of the vertical structure:
$$U_s^* = 2K^2R^2\frac{U_s}{|c_L|},$$
where U_s is the mean flow shear (it is actually the difference between the upper layer and lower layer velocities, $U_1 - U_2$, but we refer it as the mean flow shear in the present two-layer model), R the Rossby radius of deformation and c_L the speed of the long Rossby wave. Therefore the significance of the mean flow shear depends on its magnitude relative to the speed of the long Rossby waves.

10.1.1 The Two-layer Model

Assuming the mean flow has east-west components only : U_1 (upper layer) and U_2 (lower layer), then the Rossby waves with a rigid lid are governed by the following potential vorticity equations:

$$\left\{\begin{array}{rcl} (\frac{\partial}{\partial t} + U_1\frac{\partial}{\partial x})[\nabla^2\psi_1 + F_1(\psi_2 - \psi_1)] + [\beta + F_1(U_1 - U_2)]\frac{\partial\psi_1}{\partial x} & = & 0, \\ (\frac{\partial}{\partial t} + U_2\frac{\partial}{\partial x})[\nabla^2\psi_2 + F_2(\psi_1 - \psi_2)] + [\beta + F_2(U_2 - U_1)]\frac{\partial\psi_2}{\partial x} & = & 0 \end{array}\right. \qquad (10.1)$$

where ψ_1, ψ_2 are the geostrophic stream functions in layer 1 (upper) and layer 2 (lower) respectively, and $F_i = f^2/g'H_i(\mathrm{i} = 1, 2)$. The model is similar to the two layer model in Pedlosky (1987), where one can find a thorough discussion of baroclinic instability. Here we will concentrate on neutral Rossby waves and their vertical structures.

Substituting the modal wave solution

$$\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} e^{i(kx+ly-\omega t)}$$

into the governing equations 10.1, we have the following eigenvalue equations:

$$\left\{ \begin{array}{rcl} (U_1 - c)[-K^2 A_1 + F_1(A_2 - A_1)] + [\beta + F_1(U_1 - U_2)]A_1 & = & 0, \\ (U_2 - c)[-K^2 A_2 + F_2(A_1 - A_2)] + [\beta + F_2(U_2 - U_1)]A_2 & = & 0 \end{array} \right. \tag{10.2}$$

where $c = \omega/k$, and $K^2 = k^2 + l^2$.

<u>*Dispersion relation*</u>

For non-trivial A_1 and A_2, we have

$$\begin{array}{rcl} c & = & U_m - \frac{\beta}{2}(\frac{1}{K^2} + \frac{1}{K^2+F_1+F_2}) \\ & & \pm\frac{\sqrt{(\beta+F_1U_s')(\beta-F_2U_s')+\frac{K^4}{4}U_s'^2}}{2}(\frac{1}{K^2} - \frac{1}{K^2+F_1+F_2}), \end{array} \tag{10.3}$$

where

$$U_m = \frac{(K^2 + 2F_2)U_1 + (K^2 + 2F_1)U_2}{2(K^2 + F_1 + F_2)}, \text{ and } U_s' = U_s\frac{2K^2}{F_1 + F_2} = (U_1 - U_2)\frac{2K^2}{F_1 + F_2}.$$

In general, U_m is a function of wavenumber (K) and of the ratios: $\frac{F_1}{F_1+F_2}$ and $\frac{F_2}{F_1+F_2}$. There are two situations in which U_m is simply the depth average of the mean flow:

- In the long-wave limit ($K = 0$): $U_m = \frac{F_2U_1+F_1U_2}{F_1+F_2} = \frac{H_1U_1+H_2U_2}{H_1+H_2}$.
- If the two layers have the same water depth (i.e., $F_1 = F_2$), then for any wavenumber U_m is simply the arithmetic average (also the depth average) of the mean flow of the two layers: $U_m = \frac{U_1+U_2}{2}$.

In equation 10.3, it is U_s', instead of U_s, that appears in the dispersion relation. U_s' is a wavenumber-modified shear, which equals U_s only when $|K| = \frac{1}{\sqrt{2}R}$.

In the long-wave limit ($K \to 0$), U_m is the depth average of the mean flows and U_s' is zero. Thus the only effect of the mean flows is a constant Doppler shift; the shear of the mean flow has no effect—the waves are always stable. As will shown later, the vertical structures do not change either, i.e, the vertical structures are those of the normal modes without mean flow.

The Rossby waves will be unstable when the radicand in equation 10.3 becomes negative. A necessary condition for instability is $(\beta + F_1 U_s')(\beta - F_2 U_s') < 0$. From equations 10.1, we see that $\beta + F_1 U_s$ and $\beta - F_2 U_s$ are the background potential vorticity gradients in the upper and lower layers. When $|K| = \frac{1}{\sqrt{2}R}$, then the necessary condition is a change in the sign of the background potential vorticity gradients from upper to lower layer.

10.1.2 Vertical Structure: Shear Modes

With mean flow shears, the vertical structure of a Rossby wave varies not only with shear (U_s) but also with wavenumber (K); with $U_s = 0$, the vertical structure is independent of wavenumber. From 10.2, we have the vertical structure

$$\frac{A_2}{A_1} = \frac{\frac{F_1 - F_2}{F_1 + F_2}\beta + \frac{K^2}{2}U_s' \mp \sqrt{(\beta + F_1 U_s')(\beta - F_2 U_s') + \frac{K^4}{4}U_s'^2}}{\frac{2F_1}{F_1 + F_2}\beta + F_1 U_s'}. \tag{10.4}$$

First we will show that, in the long wave limit, the vertical structures are those of the normal modes without mean flow. As shown in last section, $U_s' = 0$ in the long wave limit, so we have

$$\frac{A_2}{A_1} = -\frac{F_2}{F_1} = -\frac{H_1}{H_2} \qquad \text{and} \qquad \frac{A_1}{A_2} = 1.$$

They are the vertical structures of the normal modes without mean flows (baroclinic and barotropic modes respectively).

Defining

$$U_s^* = \frac{U_s'}{\beta/(F_1+F_2)} = \frac{2K^2}{\beta}U_s = \frac{2K^2}{F_1+F_2}\frac{U_s}{|c_L|},$$

then the vertical structures become

$$\frac{A_2}{A_1} = \frac{F_1 - F_2 + \frac{K^2}{2}U_s^* \mp (F_1+F_2)\sqrt{(1+\frac{F_1}{F_1+F_2}U_s^*)(1-\frac{F_2}{F_1+F_2}U_s^*) + \frac{K^4}{4(F_1+F_2)^2}U_s^{*2}}}{2F_1 + F_1 U_s^*},$$

where $c_L = \beta/(F_1+F_2)$. Notice that the β effect has been absorbed in U_s^*. As U_s' is the natural parameter for c, U_s^* is the natural parameter for the vertical structure.

In a two-layer model without mean flows, the two dynamical modes are orthogonal. If we denote them as (c_1, A_1, A_2) and (c_2, A_1', A_2'), we have

$$H_1 A_1 A_1' + H_2 A_2 A_2' = 0. \quad \text{or} \quad F_2 A_1 A_1' + F_1 A_2 A_2' = 0. \tag{10.5}$$

With mean flows, they are not orthogonal any more. It is possible to get a relationship similar to equation 10.5 with scaling coefficients before the terms $A_1 A_1'$ and $A_2 A_2'$. The scaling coefficients may give us some insights into why the vertical structures differ from those of normal modes. In order to get the coefficients, we have four equations from equations 10.2:

$$\begin{cases} \{-K^2 - F_1 + \frac{\beta + F_1 U_s}{U_1 - c_1}\}A_1 & +F_1 A_2 & = & 0 \\ \{-K^2 - F_2 + \frac{\beta - F_2 U_s}{U_2 - c_1}\}A_2 & +F_2 A_1 & = & 0 \\ \{-K^2 - F_1 + \frac{\beta + F_1 U_s}{U_1 - c_2}\}A_1' & +F_1 A_2' & = & 0 \\ \{-K^2 - F_2 + \frac{\beta - F_2 U_s}{U_2 - c_2}\}A_2' & +F_2 A_1' & = & 0. \end{cases}$$

Multiplying the equations with $F_2 A_1'$, $F_1 A_2'$, $-F_2 A_1$ and $-F_1 A_2$ respectively and adding them together yield:

$$(\beta + F_1 U_s)(\frac{1}{U_1 - c_1} - \frac{1}{U_1 - c_2})F_2 A_1 A_1' + (\beta - F_2 U_s)(\frac{1}{U_2 - c_1} - \frac{1}{U_2 - c_2})F_1 A_2 A_2' = 0,$$

which can be rewritten as

$$(\beta+F_1U_s)(\frac{1}{(c_1-U_1)(c_2-U_1)})F_2A_1A_1'+(\beta-F_2U_s)(\frac{1}{(c_1-U_2)(c_2-U_2)})F_1A_2A_2'=0, \tag{10.6}$$

When the mean flow has no shear ($U_s=0$, i.e., $U_1=U_2$), the above equation reduces to equation 10.5. From the above equation, we can identify the physical factors which change the vertical structures from those of normal modes:

- The background potential vorticity gradients ($\beta+F_1U_s$ and $\beta-F_2U_s$) are different in the two layers;
- The wave speeds seen by a person following the mean flows ($c_i-U_j, i=1,2, j=1,2$) are different in the upper and lower layers.

These are actually the two familiar consequences of a mean flow: modification of the background potential vorticity gradient and advection.

Latitudinal dependence

The mean flow shear naturally combines with the β effect as follows

$$U_s^*=\frac{U_s'}{|c_L|}=\frac{2K^2}{F_1+F_2}\frac{U_s}{|c_L|}=2K^2R^2\frac{U_s}{|c_L|},$$

where $c_L=-\beta/(F_1+F_2)$ and $R^{-2}=F_1+F_2$. c_L can be expressed as follows

$$c_L=-\frac{\beta c_1^2}{f^2},$$

where f is the Coriolis parameter and c_1 is the phase speed of the first mode gravity wave without a mean flow. As calculated from a global climatological dataset, c_1 ranges approximately from 2 m s^{-1} to 3 m s^{-1} (Chelton *et al.* 1998). Thus we have a good estimate of c_L. β and f are latitudinally dependent, and so is c_L (figure 10.1).

As seen in the figure, c_L varies from ~2 cm s^{-1} at 50 degrees latitude to ~20 cm s^{-1} at 10 degrees latitude. The variation within 10-20 degrees latitude is the greatest: from 7 cm s^{-1} at 20 degrees to triple that value at 10 degrees.

Assuming U_s is a few cm s^{-1} and the horizontal scale is close to the Rossby radius of deformation, then the only region where the shear modes will be significant is in the high latitudes. In the Antarctic Circumpolar Current region, the U_s could be ~50 cm s^{-1}, and the shear modes would be significant even for the horizontal scales somewhat larger than the first Rossby radius of deformation. In middle and low latitudes, the shear modes are not likely to be significant.

There are some caveats here:

- We considered zonal mean flow only. As we know, all meridional mean flows with vertical shears are baroclinically unstable. It sounds dramatic. Then one may wonder what kind of effect a meridional mean flow will have on the vertical structure. For the case with meridional mean flow V_1 and V_2 only, after we define (see Pedlosky (1987))

  $$\tilde{\beta} = \beta \tfrac{k}{K}$$
  $$\tilde{U}_1 = V_1 \tfrac{l}{K} \quad \text{and} \quad \tilde{U}_2 = V_2 \tfrac{l}{K}$$
  $$\tilde{c} = \tfrac{\omega}{K},$$

  we will get the same equations as equations 10.2 except for the tilde terms. Therefore, the situation will certainly be more complicated with more dependence on wavenumbers. However, if $\frac{k}{K}$ and $\frac{l}{K}$ are $\sim \bigcirc(1)$, the tilde terms are close to the untilded and the present results would be indicative.

- In the present model, the mean flow is assumed to be steady with large horizontal scales. To define such flow will be problematic.

10.2 Non-linearity

Nof (1982), Cushman-Roisin and Tang (1990) and others derived an integral constraint for one and one half layer models, which states that an isolated eddy will propagate approximately at the speed of long Rossby waves. This seems like the right direction since the JAMSTEC model output does show that the eddies are basically propagating at such a speed. However, when the same idea is applied to a model with a rigid lid and a flat bottom, the integral constraint becomes the so-called "no net angular momentum" theorem (Flierl 1987) which states that the translating isolated structure must be baroclinic. This constraint applies to primitive equation, as well as quasi-geostrophic models. Because of the theorem, we will restrict the non-linearity to the upper layer where a translating isolated structure is possible, and allow linear Rossby waves in the lower layer to radiate away so as not to be isolated. In the real ocean the upper layer is usually energetic and the lower layer flow is weak, so that it is reasonable to assume that non-linearity is important in the upper layer but not in the lower layer.

In quasi-geostrophic dynamics, either the continuous stratification model:

$$\frac{\partial}{\partial t}(\nabla^2\psi + \frac{\partial}{\partial z}\frac{f^2}{N^2}\frac{\partial\psi}{\partial z}) + J(\psi, \nabla^2\psi + \frac{\partial}{\partial z}\frac{f^2}{N^2}\frac{\partial\psi}{\partial z}) + \beta\frac{\partial\psi}{\partial x} = 0,$$

with a rigid lid and a flat bottom

$$\frac{\partial}{\partial t}\frac{\partial\psi}{\partial z} + J(\psi, \frac{\partial\psi}{\partial z}) = 0, \quad \text{at} \quad z = -H \quad \text{and} \quad 0$$

or the two-layer model

$$\left.\begin{aligned} \tfrac{\partial}{\partial t}[\nabla^2\psi_1 + F_1(\psi_2 - \psi_1)] + J(\psi_1, \nabla^2\psi_1 + +F_1(\psi_2 - \psi_1)) + \beta\tfrac{\partial\psi_1}{\partial x} &= 0 \\ \tfrac{\partial}{\partial t}[\nabla^2\psi_2 + F_2(\psi_1 - \psi_2)] + J(\psi_2, \nabla^2\psi_2 + +F_2(\psi_1 - \psi_2)) + \beta\tfrac{\partial\psi_2}{\partial x} &= 0 \end{aligned}\right\}$$

also with a rigid lid and a flat bottom, the non-linearity comes from the Jacobian terms. The enigma is that the Jacobian terms will disappear if we use the radially symmetric eddy as the prototype ($\partial/\partial\theta = 0$), because

$$\begin{aligned} J(A,B) &= \frac{\partial A}{\partial x}\frac{\partial B}{\partial y} - \frac{\partial A}{\partial y}\frac{\partial B}{\partial x} \qquad \text{(Cartesian coordinates)} \\ &= \frac{\partial A}{\partial r}\frac{\partial B}{r\partial\theta} - \frac{\partial A}{r\partial\theta}\frac{\partial B}{\partial r} \qquad \text{(polar coordinates)}. \end{aligned}$$

Therefore, we will focus on the non-linearity associated with the finite interface perturbation with respect to the upper layer depth. This non-linearity arises from the continuity equation and has been studied in the context of so-called intermediate geostrophy (Charney and Flierl 1982) and frontal geostrophy (Cushman-Roisin 1986).

To keep the problem analytically tractable, we further assume that the depth ratio is large so that the flow in the lower layer is weak. The upper layer non-linear equation is solved without any influence from the lower layer, and an upper layer positive soliton which translates steadily westward is obtained. A quasi-steady wave field in the lower layer forced by the upper layer soliton is obtained by solving an initial-value problem. Overall, we find an approximate translating solution of the non-linear model, i.e., $\psi = \psi(x - ct, y, z)$.

With respect to an upper layer positive soliton, the vertical structure refers to the relative locations of the lower layer negative maximum (east of the soliton and always well defined) and the positive extreme (west of the soliton but sometimes poorly defined). We would consider the non-linear theory successful in modeling the dominant vertical structure if 1) the westward positive extreme is well defined and 2) the upper layer positive soliton is centered approximately between the lower layer positive extreme and negative maximum, so that one could identify a barotropic component (connection of positive extremes in the two layers) with a slope downward and to the west. For smaller upper layer solitons (e.g., $L = 1$), the model works fairly well, but for larger solitons the westward positive extreme is poorly defined and the upper layer soliton is too close to the lower layer negative maximum so that the vertical structure is more baroclinic. Overall, the non-linear model has a limited performance. However, at intermediate time steps (t=3 instead of t=10), the positive extreme west of the negative maximum is better defined, so that we could identify a barotropic structure with almost 90 degree phase shift—the upper layer eddy is located between the positive extreme and negative maximum in the lower layer. In a realistic situation, the eddy field is intermittent, so that the quasi-steady solution is rarely reached and the intermediate solution could be more relevant to reality. Notice that the intermediate solution does not invalidate the assumptions of the non-linear model.

The non-linear theory does not satisfactorily model the vertical structure in the observation and in the JAMSTEC model. Restricted by the needs of analytical tractability, the two-layer model without interaction between layers could be too simple. Furthermore, we used a single soliton as the prototype of the eddy field; this could be another limitation of the model. The altimetric observations and JAMSTEC model output show much horizontal interaction among eddies; an individual eddy sometimes does not persist very long.

It is emphasized that our study does not imply that the non-linear Jacobian terms are unimportant; the limited performance of our model may instead indicate that the neglected Jacobian terms are important. Restricted by the analytical tractability, we simply do not address those terms. How to analytically address the effect of the Jacobian terms on the vertical structure remains a challenge, while numerical modeling studies, pioneered by Rhines (1977), suggested so-called barotropization at certain stages of geostrophic turbulence.

10.2.1 The Two-layer Model

Our model configuration (figure 10.2b) is similar to that used by Flierl (1984) to study Rossby wave radiation from a lens-like Gulf Stream warm ring (figure 10.2a). However, we consider finite interface perturbations without interface out-cropping. We will assume a geostrophic balance at the leading order, so the upper layer governing equation is the potential vorticity equation, while Flierl (1984) used the primitive equations in the upper pool. For the lower layer we will use the same dynamics as Flierl (1984).

The potential vorticity equations for each layer (figure 10.2b) are:

$$\frac{\partial}{\partial t}\nabla^2\psi_i + J(\psi_i, \nabla^2\psi_i) + \beta\frac{\partial\psi_i}{\partial x} - f\frac{\partial w}{\partial z} = 0 \qquad (\mathrm{i} = 1, 2), \tag{10.7}$$

where ψ_i (i=1,2) are the geo-potentials for the upper and lower layers respectively. The vertical velocities at the surface and bottom are zero. At the interface

$$w = \frac{d\zeta}{dt} = \frac{f}{g'}\frac{d(\psi_2 - \psi_1)}{dt},$$

in which $\frac{d}{dt}$ is $\frac{\partial}{\partial t} + J(\psi_1, \ )$, or $\frac{\partial}{\partial t} + J(\psi_2, \ )$ depending on where (either the upper layer or the lower layer) it is evaluated.

ψ_i are constant with z within each layer. Vertically integrating the potential vorticity equations 10.7 from $(-H_1 + \zeta)$ to zero (upper layer) and from $(-H_1 - H_2)$ to $(-H_1 + \zeta)$ (lower layer) yields

$$(H_1 - \zeta)[\frac{\partial}{\partial t}\nabla^2\psi_1 + J(\psi_1, \nabla^2\psi_1) + \beta\frac{\partial\psi_1}{\partial x}] + \frac{f^2}{g'}\frac{d(\psi_2 - \psi_1)}{dt} = 0, \tag{10.8}$$

and

$$(H_2+\zeta)[\frac{\partial}{\partial t}\nabla^2\psi_2 + J(\psi_2,\nabla^2\psi_2) + \beta\frac{\partial\psi_2}{\partial x}] + \frac{f^2}{g'}\frac{d(\psi_1-\psi_2)}{dt} = 0. \tag{10.9}$$

After a few rearrangements,

$$\begin{aligned}\frac{\partial}{\partial t}[\nabla^2\psi_1 + F_1(\psi_2-\psi_1)] + J(\psi_1,\nabla^2\psi_1 + F_1(\psi_2-\psi_1)) + \beta\frac{\partial\psi_1}{\partial x}& \\ -\{\frac{\zeta}{H_1}[\frac{\partial}{\partial t}\nabla^2\psi_1 + J(\psi_1,\nabla^2\psi_1) + \beta\frac{\partial\psi_1}{\partial x}]\} &= 0, \quad (10.10)\end{aligned}$$

and

$$\begin{aligned}\frac{\partial}{\partial t}[\nabla^2\psi_2 + F_2(\psi_1-\psi_2)] + J(\psi_2,\nabla^2\psi_2 + F_1(\psi_1-\psi_2)) + \beta\frac{\partial\psi_2}{\partial x}& \\ +\{\frac{\zeta}{H_2}[\frac{\partial}{\partial t}\nabla^2\psi_2 + J(\psi_2,\nabla^2\psi_2) + \beta\frac{\partial\psi_2}{\partial x}]\} &= 0 \quad (10.11)\end{aligned}$$

where F_i (i=1,2) are defined as $F_i = \frac{f^2}{g'H_i}$ $(\mathrm{i}=1,2)$. Since $\zeta = \frac{f}{g'}(\psi_2-\psi_1)$, equations 10.10 and 10.11 are for two unknowns ψ_1 and ψ_2. If $|\zeta| \ll H_1$ and $|\zeta| \ll H_2$, the parts associated with the curly brackets can be neglected, and equations 10.10 and 10.11 reduce to the familiar equations of a two-layer QG model (e.g. Pedlosky 1987). The extra non-linear terms are part of the potential vorticity as demonstrated as follows for the upper layer

$$\begin{aligned}\tfrac{d}{dt}(\tfrac{f+\xi}{H_1-\zeta}) &= \tfrac{d}{dt}\{\tfrac{f_0+\beta y+\xi}{H_1}(1+\tfrac{\zeta}{H_1}-\cdots)\} \\ &= \tfrac{1}{H_1}\tfrac{d}{dt}\{f_0\tfrac{\zeta}{H_1} + \beta y + \xi + (\beta y+\xi)\tfrac{\zeta}{H_1} - \cdots\} \\ &\approx \tfrac{1}{H_1}[\tfrac{d}{dt}\{f_0\tfrac{\zeta}{H_1} + \beta y + \xi\} + \underline{\tfrac{\zeta}{H_1}\tfrac{d}{dt}\{(\beta y+\xi)\}} - \cdots]\end{aligned}$$

ζ is finite comparing with H_1 but will be much less than H_2. As a result, we will neglect the curly bracket term in equation 10.11. Furthermore, the flow in the lower layer is assumed weak, so we will also neglect the advection term (the second term) in equation 10.11. Therefore, the lower layer is governed by the following linear equation

$$\frac{\partial}{\partial t}[\nabla^2\psi_2 + F_2(\psi_1-\psi_2)] + \beta\frac{\partial\psi_2}{\partial x} = 0. \tag{10.12}$$

We will not neglect the advection terms for the upper layer (equation 10.10), since the currents there are strong. However, since we assume that the isolated structure in the upper layer is radially symmetric ($\partial/\partial\theta = 0$) , the Jacobian term is identically zero. In the real world, the circular characteristic of most eddies, though not perfectly radially symmetric, will greatly reduce the non-linear effect of the advection terms. We will also neglect the influence of ψ_2, i.e., we will not consider the feedback from the lower layer to the upper layer. This is justified by equation 10.15, the non-dimensionalized form of equation 10.12. Since the lower layer is assumed forced by the upper layer, equation 10.15 suggests that ψ_2 in the same order of $\frac{1}{\delta}\psi_1$. In equation 10.12 the influence of ψ_2 is in the form of $\psi_1 - \psi_2$, so that we could neglect the feedback from the lower layer as long as $\frac{1}{\delta} = \frac{H_1}{H_2} \ll 1$.

Finally, the upper layer is governed by

$$\frac{\partial}{\partial t}[\nabla^2\psi_1 - F_1\psi_1] + \beta\frac{\partial\psi_1}{\partial x} + \frac{f\psi_1}{g'H_1}[\frac{\partial}{\partial t}\nabla^2\psi_1 + \beta\frac{\partial\psi_1}{\partial x}] = 0. \tag{10.13}$$

The following characteristic scales and parameters:

length $R = \frac{\sqrt{g'H_1}}{f} \sim 30$ km

time $T = \frac{1}{\beta R} \sim 20$ days

velocity $U \sim 20$ cm s^{-1}

Rossby number R_0 $\frac{U}{fR} = \frac{U}{\sqrt{g'H_1}} \sim 0.1$

depth ratio δ $\frac{H_2}{H_1} \sim 5$

β $\sim 2 \times 10^{-11}$ m^{-1}s^{-1}

$\sqrt{g'H_1}$ ~ 3 m s^{-1}

f $\sim 1 \times 10^{-4}$ s^{-1}

H_1 ~ 800 m

H_2 ~ 4000 m,

are used to non-dimensionalize equations 10.12 and 10.13. U is the characteristic velocity in the upper layer. Since the lower layer is linear, we can use the same characteristic scales. Also notice that the scale of ψ_i is uR (i=1,2). The non-dimensioned governing equations for the upper and lower layers of our model are

$$\frac{\partial}{\partial t}[\nabla^2\psi_1 - \psi_1] + \frac{\partial \psi_1}{\partial x} + R_o\psi_1[\frac{\partial}{\partial t}\nabla^2\psi_1 + \frac{\partial \psi_1}{\partial x}] = 0 \quad \text{(upper layer)}, \quad (10.14)$$

$$\frac{\partial}{\partial t}[\nabla^2\psi_2 + \frac{1}{\delta}(\psi_1 - \psi_2)] + \frac{\partial \psi_2}{\partial x} = 0 \quad \text{(lower layer)}. \quad (10.15)$$

For clarity, the same notations as their dimensional counterparts are used in above equations. The appearance of $1/\delta$ in equation 10.15 is because we utilized the same time and horizontal scales for the upper and lower layers. It will disappear, if we choose different scales for the lower layer as follows

$$R_2 = R\sqrt{\delta}, \quad \text{and} \quad T_2 = T/\sqrt{\delta}.$$

If we neglect the ψ_1 related term in equation 10.15, which is actually a forcing term, then the homogeneous solution of equation 10.15 will be a function of the new independent variables non-dimensionalized by R_2 and T_2. In such a sense, the lower layer will have a larger horizontal scale and a shorter time scale than the upper layer when $\delta > 1$.

comments

The assumptions leading to equations 10.14 and 10.15 are

1. Geostrophy allows us to use streamfunctions, express the advective non-linear terms in Jacobian form, and relate the Rossby number R_o to the interface perturbation scale (δ_H) as follows: assuming the lower layer flow is small, $U \sim \frac{g'\delta_H}{fR}$ and $R_o = \frac{U}{fR} = \frac{g'\delta_H}{f^2R^2} = \frac{\delta_H}{H_1}$.

2. The depth ratio is large, so that we can assume that the motion in the lower layer is weak, and that the interface perturbation is infinitesimal with respect to the lower layer depth.

3. The non-linearity due to the interface perturbation is important in the upper layer but not in the lower layer.

4. The flow in the upper layer is radially symmetric so that the Jacobian non-linear terms are not considered. The flow in the lower layer is weak so that the Jacobian terms are also not considered.

It is clear that our model is very simplified in order to illustrate the effect of the non-linearity analytically. The model is intended to represent mid-latitude mesoscale motion. At high latitudes, to assume a weak flow in the lower layer may not be suitable; near the equator, equatorial dynamics should be applied.

In our model, we retain two horizontal scales: R and $R\sqrt{\delta}$, or equivalently, $\frac{\sqrt{g'H_1}}{f}$ and $\frac{\sqrt{g'H_2}}{f}$. In this respect, it is different from the modal solution which has only one horizontal scale $\frac{1}{f}\sqrt{\frac{g'H_1H_2}{H_1+H_2}}$. It is further emphasized that the horizontal structure of the lower layer will be different from that in the upper layer: one is an isolated structure and another is a wave field extending far away. That means, in such situation, that the solutions in the two layers are not similar, i.e., $\psi_2(x,y;t) \neq C \cdot \psi_1(x,y;t)$ (C is a constant) and not separable $\psi(x,y,z;t) \neq f(z)g(x,y;t)$ for continuous stratification. This is totally different from the dynamics of the linear normal modes. For the linear baroclinic mode of the two-layer model, $\psi_2 = C \cdot \psi_1$ where $C = -\frac{H_2}{H_1}$.

10.2.2 Upper Layer: Solitons

The upper layer governing equation (equation 10.14) is

$$\frac{\partial}{\partial t}[\nabla^2\psi_1 - \psi_1] + \frac{\partial \psi_1}{\partial x} + R_o\psi_1[\frac{\partial}{\partial t}\nabla^2\psi_1 + \frac{\partial \psi_1}{\partial x}] = 0,$$

where we have linear wave terms $-\frac{\partial \psi_1}{\partial t} + \frac{\partial \psi_1}{\partial x}$, a dispersion term $\frac{\partial \nabla^2 \psi_1}{\partial t}$, and non-linear terms multiplied by R_o. A soliton requires that 1) the first order of the equation is a linear wave equation and 2) the dispersion and non-linear terms appear together in the second order term. To satisfy these requirements, we stretch the coordinates as follows

$$(x', y') = \sqrt{\hat{S}}(x, y) \text{ and } t' = \sqrt{\hat{S}}(t)$$

Assume $\hat{S} \sim R_o$, then the equation becomes

$$\frac{\partial \psi_1}{\partial t} - \frac{\partial \psi_1}{\partial x} - R_o\psi_1\frac{\partial \psi_1}{\partial x} - \hat{S}\frac{\partial \nabla^2 \psi_1}{\partial t} = 0. \qquad (10.16)$$

Notice we have neglected the term in order of $R_o\hat{S}$ and dropped the primes. Assuming $\frac{\partial}{\partial y} = 0$, we could obtain one-dimensional solitons analytically (appendix D). Here we present the two-dimensional solitons or radial solitons, which propagate westward steadily. However, no analytical solution could be obtained, and numerical techniques are needed.

Using the coordinate transformation $s = x - ct$, the governing equation 10.16 becomes

$$\frac{\partial}{\partial s}[(1 + c)\psi_1 + \frac{R_o}{2}\psi_1^2 - \hat{S}c(\frac{\partial^2 \psi_1}{\partial s^2} + \frac{\partial^2 \psi_1}{\partial y^2})] = 0.$$

For a soliton, $\psi_1 = 0$ and $\frac{\partial^2 \psi_1}{\partial s^2} + \frac{\partial^2 \psi_1}{\partial y^2} = 0$ when$|s| \to \infty$, so that

$$(1 + c)\psi_1 + \frac{R_o}{2}\psi_1^2 - \hat{S}c(\frac{\partial^2 \psi_1}{\partial s^2} + \frac{\partial^2 \psi_1}{\partial y^2}) = 0$$

with a zero integration constant. For a radial soliton, $\frac{\partial^2\psi_1}{\partial s^2} + \frac{\partial^2\psi_1}{\partial y^2} = \frac{1}{r}\frac{d}{dr}r\frac{d\psi_1}{dr} = \frac{d^2\psi_1}{dr^2} + \frac{1}{r}\frac{d\psi_1}{dr}$, where r=$\sqrt{s^2+y^2}$. Finally, the governing equation for the radial stationary soliton is

$$(1+c)\psi_1 + \frac{R_o}{2}\psi_1^2 - \hat{S}c(\frac{d^2\psi_1}{dr^2} + \frac{1}{r}\frac{d\psi_1}{dr}) = 0. \tag{10.17}$$

The boundary conditions are

$$\begin{cases} \text{non singular} \quad \text{at} \quad r = 0 \\ \psi_1 = \frac{d\psi_1}{dr} = 0 \quad \text{at} \quad r = \infty \end{cases}$$

$\psi_1 = 0$ is its trivial solution. Our purpose is to find a particular c and its corresponding non-zero ψ_1 decaying towards zero as $r \to \infty$, given R_o and $\hat{S}$. The numerical scheme is the Runge-Kutta integration of the following first-order differential equations:

$$\begin{cases} \frac{d\psi_1}{dr} &= G \\ \frac{dG}{dr} &= \frac{1}{\hat{S}c}[(1+c)\psi_1 + \frac{R_o}{2}\psi_1^2] - \frac{1}{r}G \end{cases}$$

with conditions

$$\psi_1 = A,\ G = 0 \quad \text{when} \quad r = 0.$$

where A is the peak value of the 2-D soliton at the center. The integration is from $r = 0$ to $r = \infty$.

When $R_o = 0$, i.e, no non-linear term, equation 10.17 becomes a zeroth-order Bessel equation:

$$\frac{d^2\psi_1}{dr^2} + \frac{1}{r}\frac{d\psi_1}{dr} + (-\frac{1+c}{\hat{S}c})\psi_1 = 0.$$

When $-1 < c < 0$, its two independent solutions are Bessel functions of the zeroth-order:

$$J_0(\sqrt{\frac{1+c}{\hat{S}|c|}}r) \quad \text{and} \quad Y_0(\sqrt{\frac{1+c}{\hat{S}|c|}}r).$$

When $c < -1$ or $c > 0$, they are modified Bessel function of the zeroth-order:

$$I_0(\sqrt{\frac{|1+c|}{\hat{S}|c|}}r) \quad \text{and} \quad K_0(\sqrt{\frac{|1+c|}{\hat{S}|c|}}r).$$

To be non-singular at r=0 , only are J_0 and I_0 permissible.

Figure 10.3 shows the numerical solution of equation 10.17 when A=1. For most values of c, the numerical solution behaves indeed like J_0 or I_0. Near $c = 0^-$, the solution decreases from $\psi_1 = 1$ very fast but keeps an oscillation, so that the contours are indiscernible. Near $c = 0^+$, the solution increases dramatically. When $c = 0$, $\psi_1 = 0$ or $-1/R_o$ (see equation 10.17), so that $\psi_1 = 1$ at $r = 0$ is a bad value. Since these situations are not of interest to us, we grayed out the solution near $c = 0$. Recall that we are interested in the solution which is finite near the center and decays towards zero as $r \to \infty$. Figure 10.3 suggests that the only possible c for such a solution is slightly less than -1.

A close look at ψ_1 near $c = -1.02$ (figure 10.4) does show such a soliton-type solution between -1.021 and -1.020. We tracked down the value c to a high precision as shown in figure 10.5. In figure 10.5, we also plotted the 1-D soliton discussed in appendix D with the same parameters ($A = 1$, $R_o = 0.1$ and $\hat{S} = 0.1$), and the Gaussian curve as in figure D.1. All three curves are visually similar to each other. Figure 10.6 shows similar results for a stronger soliton, $A = 10$. From these two cases, we can conclude that the Gaussian curves are good representations of both 1-D and 2-D solitons.

10.2.3 Lower Layer: Forced Barotropic Waves

The governing equation of the lower layer (equation 10.15) is:

$$\frac{\partial}{\partial t}[\nabla^2\psi_2 + \frac{1}{\delta}(\psi_1 - \psi_2)] + \frac{\partial \psi_2}{\partial x} = 0$$

which becomes in the moving coordinates (s=x-ct,y;t)

$$\frac{\partial}{\partial t}[\nabla^2\psi_2 - \tfrac{1}{\delta}\psi_2] + \underset{\text{I}}{\frac{\partial\psi_2}{\partial s}} - c\frac{\partial}{\partial s}[\ \underset{\text{II}}{\nabla^2\psi_2} - \underset{\text{III}}{\tfrac{1}{\delta}\psi_2}\ ] = \underset{\text{RT}}{\frac{c}{\delta}\frac{\partial\psi_1}{\partial s}} \tag{10.18}$$

The forcing term at the right side (RT) is the upper layer soliton discussed in the previous section. c is chosen as the translating speed of the soliton, so that the forcing term is stationary in the moving coordinates, i.e., $\psi_1 \equiv \psi_1(s,y)$. We will use the Gaussian eddy $\psi_1 = \exp(-\frac{s^2+y^2}{2L^2})$ to approximate the upper layer soliton, because of its simple Fourier transform $L^2\exp(-\frac{L^2(k^2+l^2)}{2})$. The amplitude of ψ_1 does not matter, because equation 10.18 is linear, but the size of the Gaussian eddy, characterized by L, matters, since it determines the wavenumber content of the forcing.

Equation 10.18 is solved with following conditions:

$$\psi_2 = 0, \quad \text{when t} = 0$$

$$\psi_2 \to 0 \quad \text{in the far field, i.e., r} = \sqrt{\mathrm{s}^2 + \mathrm{y}^2} \to \infty$$

The solution is obtained in appendix E:

$$\psi_2 = \frac{L^2}{2\delta}\int_0^\infty \frac{e^{-\frac{\mathcal{Z}L^2}{2}}}{\mathcal{Z} + \frac{1}{\delta} + \frac{1}{c}} \left(J_0(\sqrt{\mathcal{Z}}\sqrt{s^2+y^2}) - J_0(\sqrt{\mathcal{Z}}\sqrt{[s + (\frac{1}{\mathcal{Z}+\frac{1}{\delta}} + c)t]^2 + y^2}\,) \right) d\mathcal{Z}$$

where J_0 is the zero-order of Bessel function of the first kind.

Our goal here is to obtain a quasi-steady solution of equations 10.14 and10.15 in the form $\psi_1(x-ct,y)$ and $\psi_2(x-ct,y)$, where c is the translation speed. This is shown in figure 10.7 (the intermediate solutions of the lower layer between an initial resting state and those in figure 10.7 are shown in appendix E). The lower layer is filled with barotropic Rossby waves. Short Rossby waves are east of the forcing (marked by the white vertical lines in figure 10.7) and long Rossby waves are west of the forcing. The sausage shapes are also typical of the barotropic Rossby waves. These features were seen in the laboratory (Firing and Beardsley 1976) and in theoretical and numerical models (e.g. Longuet Higgins 1965, McWilliams *et al.* 1986 and Chassignet and Cushman-Roisin 1991).

The main response of the lower layer is the negative maximum, which is always east of the forcing. This is clearer for a smaller upper layer eddy (the left column of figure 10.7), and can be understood by looking at the balances of equation 10.18, where c is close to -1 ($c = -1$ is used here). In the quasi-steady state, the left side of the equation consists of terms I, II and III. For a large depth ratio ($\delta = 4$ is used here), term III is always smaller than term I. For a larger upper layer eddy, the horizontal scale is large so that term II is small. The dominant balance is then between term RT and term I. i.e., forcing and β effect. The outcome of this balance is $\psi_2 \propto -\psi_1/\delta$, i.e., a pure baroclinic vertical structure and ψ_2 and ψ_1 will align with each other vertically. The physics is the following: when the upper layer warm eddy moves west, it will shrink the lower layer west of the upper layer eddy and stretch the lower layer east of it. Since the response is dominated by the β effect, the water parcels west of the upper layer eddy have to move southward to conserve potential vorticity and the water parcels east of it have to move northward. Thus the motion of the lower layer is cyclonic, opposite to that of the upper layer. In this balance, the rate of change of the water depth of the lower layer is important. Under the center of the upper layer eddy, the lower layer is depressed the most but the change rate is zero so that the north-south velocity is zero. In other words, the lower layer cyclonic motion is centered under the upper layer eddy.

The dominant balance is used to explain the existing of the negative maximum. To understand why it is east of the forcing, we need to consider the time-independent homogeneous equation including terms I, II and III, which is

$$\frac{\partial}{\partial s}\left[\nabla^2\psi_2 + (-\frac{1}{c} - \frac{1}{\delta})\psi_2\right] = 0.$$

Its general solution ψ_{2g} in polar coordinate ($r = \sqrt{s^2+y^2}, \theta$) is a combination of Bessel functions:

$$\psi_{2g} = \sum_{n=0}^{\infty} A_n J_n(\sqrt{-1/c - 1/\delta}\, r)\cos(n\theta). \tag{10.19}$$

Notice that c is about -1. A_n are dependent on what happened before reaching the quasi-steady state.

In the general solution 10.19, the $n = 0$ component is radially symmetric so that it does not change the picture from the dominant balance between the forcing and β terms. The $n = 1$ component would be the most relevant; it is anti-symmetric in east-west direction. The origin of the anti-symmetric motion can be seen by checking the temporal and forcing terms: west of the upper layer eddy, negative (anticyclonic) relative vorticity is generated and east of it positive relative (cyclonic) vorticity is generated. Recall that the dominant balance results in a cyclonic motion centered under the upper layer eddy. Therefore, the $n = 1$ component will offset the center of the cyclonic motion (the negative maximum) eastward. Notice that this offset will be more significant if the horizontal scale is smaller so that the relative vorticity (the Laplacian term) gets larger.

Because the negative maximum is east of the forcing, we should connect the upper layer positive to the lower layer positive west of the forcing when illustrating the vertical structure. The existence of such a positive extreme is illustrated in figures E.1 to E.3 in appendix E. For a small upper layer eddy, the westward positive extreme is stronger relative to the negative maximum than in the case of a large eddy. At intermediate time steps when the lower layer is less steady (t=3 instead of t=10), the westward positive extrema are also stronger. In those two scenarios we can identify the vertical structure with an almost 90 degrees phase shift—the upper layer eddy is located between the positive and negative extrema in the lower layer. The phase shifts are consistent with what we found based on the JAMSTEC model output (figure 8.4). Notice that the intermediate solution does not invalidate the assumptions in the non-linear model. If the upper layer eddy evolves in time, probably the intermediate solutions (figure 10.8) are more applicable than the quasi-steady solution (figure 10.7), but this is just a speculation.

Although we can identify a phase shift consistent with the JAMSTEC model output, the positive extreme west of the upper layer eddy is weak and the main response of the lower layer is instead the baroclinic flow; we conclude that the non-linear model does not fully explain the dominant vertical structures seen in the observations and JAMSTEC model. However, we may be able to do a better job by having more layers and emphasizing the transitional nature of the flows. These issues are very challenging, especially if one wishes to tackle them analytically.

It is interesting to compare our theoretical results with two numerical modeling studies: one is a QG model (figure 10.9, from McWilliams *et al.* 1986) and another uses the primitive equations (figure 10.10, from Chassignet and Cushman-Roisin 1991). The parameters in McWilliams *et al* (1986) are $L = 1$ and $\delta = 4$. Figure 10.9 was averaged over time approximately from t=2 to t=5, so that the figure should be comparable with the left column of figure 10.8. The parameters in Chassignet and Cushman-Roisin (1991) are $L = 1.3$ and $\delta = 4$. The times for those rows are approximately t=1, 4 and 7. The similarity suggests that the forced linear Rossby wave theory captured the dynamics of the lower layer quite well.

10.3 Bottom Dissipation

Dissipation (or friction) is not considered in the linear Rossby wave theory. Since the empirical forms for dissipation are usually linear, it is not mathematically difficult to show its effect on Rossby waves. The effect of Rayleigh friction on free internal waves in the equatorial waveguide was studied by Mofjeld (1981). Since the Rayleigh friction was introduced after the vertical mode decomposition, the focus of Mofjeld (1981) was on the horizontal structures of the equatorial waves. In this study, we will introduce bottom friction on the total flow. The effect of the bottom dissipation on the vertical structure of the Rossby waves is one of our concerns.

Not much attention has been paid to the effect of bottom friction on Rossby waves so far. One possible reason is that observational evidence on the dissipation of the Rossby waves is rare. Luther (1982) estimated that the decay scale of the 4-6 day barotropic Rossby wave in the Pacific Ocean is about equal to its period, based on the widening of the peak of the sea-level frequency spectrum at 4-6 days. Fu (2003) inferred that the damping timescale for intraseasonal (30 days to one year) barotropic variability is generally longer than 20 days by fitting the altimetric observations into the barotropic vorticity equation with Rayleigh friction. Other knowledge about the dissipation of barotropic and baroclinic motion (not particularly of the barotropic and baroclinic Rossby waves though) in the ocean comes indirectly from the energy budget and from the spin-down of individual eddies such as the Gulf Stream rings. A survey by Wunsch (1998) indicated that, on a global scale, barotropic motion has a decay scale of ~100 days and baroclinic motion has a decay scale of ~10 years.

Because of the limited observational evidence about the dissipation of Rossby waves, we have little knowledge of the appropriate dissipation coefficients, and the theoretical results are therefore merely suggestive.

The dissipation under consideration is bottom dissipation. The horizontal dissipation in the biharmonic form does not affect the vertical structure of Rossby waves in a two-layer model (append G), i.e., the vertical structures of barotropic and baroclinic modes remain the same in the decaying (in time) or localized (in space) Rossby waves. As shown in appendix F, the bottom dissipation seems much more important than interior dissipation. In the real oceans, the bottom Ekman spirals and log-layers, obviously due to the bottom friction, were observed, for example, within the dense overflow from the Nordic Seas into the North Atlantic by Johnson and Sandford (1992) using an expendable current profiler (XCP) with ~1 m resolution—see Saunders (2001) for a recent review. We also know that the bottom drag is one of two dominant along-stream forces in the so-called streamtube models for dense outflow: the potential energy released from the descent of the outflow down the continental slope was dissipated mainly by bottom stress—the outflow was not accelerated due to the release of the potential energy. This was tested by the intensive measurements within the Mediterranean outflow (Baringer and Price 1997). Baringer and Price (1997) further showed that the inferred bottom stress can be parameterized by the familiar quadratic drag law with the coefficient c_D between 2-12$\times 10^{-3}$. Recall that the quadratic drag law with $c_D = 1.2 \times 10^{-3}$ was used in the JAMSTEC model as well as most other GCMs.

To compare with the situation of Rossby waves in stratified oceans, which will be discussed in the following sections, we first show how the bottom friction affects the barotropic waves in a homogeneous ocean. The Rossby waves will certainly be always depth-independent—the vertical structure is not an issue here. The potential vorticity equation is

$$\frac{\partial}{\partial t}(\nabla^2\psi - F\psi) + \beta\frac{\partial\psi}{\partial x} = -\gamma\nabla^2\psi \tag{10.20}$$

where ψ is the geostrophic streamfunction, and γ is the bottom friction coefficient with the dimension of inverse time. Other notations are conventional. Substituting the wave solution $\psi = Ae^{i(kx+ly-\omega t)}$ into equation 10.20, one obtains the dispersion relation $\omega = -\frac{\beta k}{k^2+l^2+F} - i\gamma$. Therefore, the propagation (real part of the dispersion relation) of the barotropic Rossby waves in a homogeneous ocean is not affected, but they all decay as $\psi \sim e^{-\gamma t}$, regardless of wavenumber. Here we see that γ acts exactly and simply as an exponential decay scale. As shown in the following sections, this is not the case for the Rossby waves in a stratified ocean; the waves will decay much slower than $e^{-\gamma t}$ and their propagation will also be altered.

Two forms of bottom friction are used in the layered and continuously stratified models respectively. One is the Rayleigh form which can be traced back to the Rayleigh friction in the momentum equations and will give rise to an extra term identical to $\gamma\nabla^2\psi$ in equation 10.20. Another form relates to the bottom Ekman layer theory, which states

$$w = sign(f)\sqrt{\frac{A_v}{2|f|}}\xi = sign(f)\mu\xi,$$

where w and ξ are the vertical velocity and relative vorticity at the top of the bottom Ekman layer, and the parameter μ has the dimension of length. For a continuously stratified QG model, the bottom Ekman boundary layer theory presents a new boundary condition

$$-\frac{f}{N^2}\frac{\partial^2\psi}{\partial z\partial t} = w_E = \mu\xi, \qquad \text{at} \quad \text{z} = -\text{H}$$

to replace the usual boundary condition $w = 0$. The two forms are equivalent under some circumstances. If we integrate the continuously stratified QG model into a layered model and include the bottom Ekman layer, the equation of the deepest layer will have an extra term $-\mu\xi$, which is equivalent to the Rayleigh friction.

Our focuses are on the propagation (dispersion relationship) and vertical structure. Two simple models are used: a two-layer model and a constant N^2 model. The main conclusions are:

1. The bottom dissipation, if strong enough, is a dynamically possible cause for the discrepancy between the observation and linear classical theory.

2. In the two-layer model, the barotropic branch, which reverts to the usual dynamical barotropic mode when the bottom dissipation gets small, is bottom-intensified and decays quickly, while the baroclinic branch is surface-intensified and decays slowly. The decay scale for the baroclinic branch in the two-layer model is more than 10 times longer than the Rossby wave time scale $(\frac{1}{\beta R})$. The situation for the constant N^2 model is similar but more complicated.

3. The first baroclinic Rossby wave without dissipation will be 180 degrees out of phase between the upper and lower layers. In the baroclinic branch of the present two-layer model, the phase shifts are reduced towards 90 degrees, depending on the wavenumber and strength of bottom dissipation. In the continuously stratified model, the phase shift can be anywhere between 0 and 180 degrees.

4. The westward propagation of the baroclinic branch is accelerated, as seen in the JAMSTEC model. The $\psi|_{z=-H} = 0$ model is the limiting case of the dissipative Rossby wave. For long Rossby waves, the westward speed of Rossby waves can be enhanced no more than 4 times.

10.3.1 The Two-layer Model

The two-layer model is same as that in chapter 9 with addition of the bottom dissipation $-\gamma'\nabla^2\psi_2$:

$$\left\{\begin{array}{lcl} \frac{\partial}{\partial t}[\nabla^2\psi_1 + (\psi_2 - \psi_1)] + \frac{\partial \psi_1}{\partial x} & = & 0 \\ \frac{\partial}{\partial t}[\nabla^2\psi_2 + \frac{1}{\delta}(\psi_1 - \psi_2)] + \frac{\partial \psi_2}{\partial x} & = & -\gamma'\nabla^2\psi_2. \end{array}\right. \tag{10.21}$$

The length and time scales are still $R = \frac{\sqrt{g'H_1}}{f}, T = \frac{1}{\beta R}$, and the bottom friction coefficient γ with the dimension of inverse time is non-dimensionalized by the time scale as $\gamma' = \gamma T$.

The wave solution is assumed:

$$\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \begin{pmatrix} \psi_1' \\ \psi_2' \end{pmatrix} e^{\mathrm{i}(kx+ly-\omega t)}.$$

Notice that if both ψ_1' and ψ_2' are real, then the phase relationship between ψ_1' or ψ_2' is either in phase (ψ_1' and ψ_2' have same sign) or 180° out of phase (ψ_1' and ψ_2' have opposite signs). However, if either ψ_1' or ψ_2' has an imaginary part, then the phase difference will be between 0 and 180°. This is the situation for dissipative Rossby waves. Substituting the wave solution into equation 10.21, we have

$$\begin{bmatrix} \omega(k^2+l^2+1)+k & -\omega \\ -\frac{\omega}{\delta} & \omega(k^2+l^2+\frac{1}{\delta})+k+\gamma(k^2+l^2)\mathbf{i} \end{bmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = 0 \tag{10.22}$$

where we have dropped the primes for ψ_1', ψ_2' and γ'. To have non-trivial solutions, the determinant of the coefficients matrix has to be zero. After some algebraic manipulations, we have

$$\omega_{1,2} = \frac{-[2k(k^2+l^2)+k(1+\frac{1}{\delta})+(k^2+l^2)(k^2+l^2+1)\gamma\mathbf{i}] \pm \sqrt{I}}{2(k^2+l^2)(k^2+l^2+1+\frac{1}{\delta})} \tag{10.23}$$

where

$$I = [k(1 + \frac{1}{\delta}) + (k^2 + l^2)(k^2 + l^2 + 1)\gamma \mathbf{i}]^2 - 4k(k^2 + l^2)^2 \frac{\gamma}{\delta}\mathbf{i}.$$

When $\gamma = 0$, then $I = [k(1 + \frac{1}{\delta})]^2$ and

$$\omega_1 = -\frac{k}{k^2 + l^2}, \qquad \omega_2 = -\frac{k}{k^2 + l^2 + 1 + \frac{1}{\delta}}.$$

These are the dispersion relationships for the barotropic and baroclinic Rossby waves without bottom dissipation.

Small γ and $k, l \sim O(1)$

Neglecting $O(\gamma^2)$ terms, we have

$$I \approx [k(1 + \frac{1}{\delta})]^2 + 2k(k^2 + l^2)[(1 + \frac{1}{\delta})(k^2 + l^2 + 1) - 2\frac{k^2 + l^2}{\delta}]\gamma \mathbf{i}$$

and

$$\sqrt{I} \approx k(1 + \frac{1}{\delta} + (k^2 + l^2)(k^2 + l^2 + 1)\gamma \mathbf{i} - \frac{2(k^2 + l^2)^2}{1 + \delta}\gamma \mathbf{i}.$$

Then

$$\begin{array}{rcl} \omega_1 & \approx & -\frac{k}{k^2+l^2} - \frac{\delta}{1+\delta}\gamma \mathbf{i} \\ \omega_2 & \approx & -\frac{k}{k^2+l^2+1+\frac{1}{\delta}} - \frac{k^2+l^2}{k^2+l^2+1+\frac{1}{\delta}}\frac{1}{1+\delta}\gamma \mathbf{i} \end{array} \qquad (10.24)$$

Obviously, ω_1 relates to the barotropic Rossby wave and ω_2 to the baroclinic Rossby waves, so that we call the former the barotropic branch and the latter the baroclinic branch. The imaginary parts are the decay rates of the waves. The decay rate of the barotropic branch is wavenumber-independent for small γ and $k, l \sim O(1)$. For other situations as will be seen, the dependency of the decay rate on wavenumbers is weak. The decay rate converges to γ when $\delta \gg 1$. The factor $\frac{\delta}{1+\delta}$ is the ratio of the lower layer depth to the total depth.

The decay rate for the baroclinic branch is very different from that of the barotropic branch. There are two factors involved. One is $\frac{1}{1+\delta}$, which is the ratio of the upper layer depth to the total depth. The other is $\frac{k^2+l^2}{k^2+l^2+1+\frac{1}{\delta}}$, which is actually the ratio of the kinetic energy to the total energy (kinetic and potential energy) for the (k,l) Fourier components of the baroclinic mode. Because of the bottom dissipation, the baroclinic mode is not a modal solution anymore, but we assume the baroclinic mode is still a good approximation for illustrative purposes due to the smallness of γ. The baroclinic mode is

$$\begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix} = \psi_c \begin{bmatrix} \sqrt{\frac{H_2}{H_1}} \\ -\sqrt{\frac{H_1}{H_2}} \end{bmatrix}.$$

The kinetic energy per unit mass is then

$$KE = \frac{H_1}{2}(\nabla\psi_1)^2 + \frac{H_2}{2}(\nabla\psi_2)^2 = \frac{H_1+H_2}{2}(\nabla\psi_c)^2.$$

The dimensional potential energy per unit mass is $PE = \frac{1}{2}g'\zeta^2$ with $\zeta = \frac{f(\psi_2-\psi_1)}{g'}$. This is

$$PE = \frac{1}{2}\frac{f^2(\psi_2-\psi_1)^2}{g'} = \frac{f^2}{2g'}\psi_c^2(\sqrt{\frac{H_2}{H_1}}+\sqrt{\frac{H_1}{H_2}})^2 = \frac{1}{2}\frac{f^2(H_1+H_2)^2}{g'H_1H_2}\psi_c^2 = \frac{H_1+H_2}{2R^2}(1+\frac{1}{\delta})\psi_c^2,$$

Where R is the length scale of the model: $R = \sqrt{g'H_1}/f$. For a single (k,l) Fourier component of ψ_c, $(\nabla\psi_c)^2 = (k^2+l^2)\psi_c^2$, and we have

$$\frac{KE}{KE+PE} = \frac{k^2+l^2}{k^2+l^2+1+\frac{1}{\delta}},$$

where the wavenumber (k,l) has been non-dimensionalized by R. For small wavenumbers, this factor is small, and so is the decay rate; this is understandable because of the large potential energy reservoir. For the barotropic mode, there is no such potential energy reservoir, so the decay rate remains the same for all wavenumbers.

The vertical structure from the first row of the matrix in equation 10.22 is

$$\Gamma = \frac{\psi_2}{\psi_1} = k^2 + l^2 + 1 + \frac{k}{\omega} = k^2 + l^2 + 1 + \frac{k}{\omega_r + \omega_i \mathbf{i}},$$

where ω_r and ω_i are the real and imaginary parts of ω (equation 10.24), and ω_r is actually the frequency without bottom dissipation. With the same assumption: a small γ and $k, l \sim O(1)$, we have $\omega_i \ll \omega_r$, so

$$\Gamma \approx k^2 + l^2 + 1 + \frac{k}{\omega_r}(1 - \frac{\omega_i}{\omega_r}\mathbf{i}) = k^2 + l^2 + 1 + \frac{k}{\omega_r} - \frac{k\omega_i}{\omega_r^2}\mathbf{i},$$

whose residual is $O(\gamma^2)$. For the barotropic and baroclinic branches respectively, Γ becomes,

$$\Gamma_1 = 1 - \frac{k\omega_i}{\omega_r^2}\mathbf{i} \quad \text{and} \quad \Gamma_2 = -\frac{1}{\delta} - \frac{k\omega_i}{\omega_r^2}\mathbf{i}.$$

If there is no dissipation ($\gamma = 0$), we have

$$\Gamma_1 = 1 \quad \text{and} \quad \Gamma_2 = -\frac{1}{\delta} = -\frac{H_1}{H_2}$$

which correspond to the usual dynamical modes. Notice that $angle(\Gamma_1) = 0°$ and $angle(\Gamma_2) = 180°$, i,e., in phase for the barotropic mode and 180° out of phase for the baroclinic mode. For decaying ($\omega_i < 0$) and westward ($k < 0$) Rossby waves, the imaginary parts of $\Gamma_{1,2}$ are negative. This implies that the phase for the baroclinic branch is in the third quadrant $[-180°, -90°)$, and that for the barotropic branch it is in the fourth quadrant $(-90°, 0°]$ (figure 10.11). It is emphasized that, in order to get the pattern in figure 10.11, one needs to satisfy two conditions: one is $\omega_i < 0$ (decaying) and another is $k < 0$ (westward propagation). If $k > 0$ (then $\omega_r < 0$ and $\omega_i < 0$), the phase will be in the second and first quadrants for the baroclinic and barotropic branches respectively.

General cases

It is difficult to pin down the magnitude of γ. The smallness imposed in the previous section was mainly for a convenient discussion. Notice that γ is non-dimensionalized by $\frac{1}{\beta R}$, which is about 100 days. Assuming an eddy decays in 100 days or so—this seems not far from our present knowledge—γ would be about 1. Assuming an eddy decays in 1000 days or so—this seems not far from our present knowledge either—γ would be 0.1. Here we are assuming the eddy is barotropic, so γ directly relates to the eddy decay rate. As emphasized in the previous section, this is not the case for baroclinic eddies. For the same decay rate, γ for the baroclinic eddy needs to be much larger, although we do not know how large it should be. We will allow γ to vary from 0.1 to 5 or so. For the continuously stratified case we may have a better situation, for the JAMSTEC model output may shed some light on this issue.

As seen in the top row of figure 10.12, the propagation characteristics of the barotropic branch do not change much, while there are significant changes in the dispersion relations of the baroclinic branches for l=0 and 0.5. As a general trend, the waves of the baroclinic branch propagate faster than the baroclinic Rossby waves without dissipation (same wavenumber k but larger ω_r), and the waves of the barotropic branch are slower than the barotropic Rossby waves without dissipation (same wavenumber k but smaller ω_r). As in the case of small γ (equation 10.24), the decay rate (-ω_i) of the barotropic branch converges to γ, and the decay rate of the baroclinic branch is generally much less than γ (the top row of figure 10.12).

We also notice that the decay rate of the baroclinic branch decreases towards zero in the limit of large or small wavenumber k. The decay rate decreases when k is small, because there is more and more accompanying potential energy; the decay rate decreases when k is large, because the lower layer velocity goes to zero—see the second row in figure 10.12. We could interpret the latter case in an alternative way. For large wavenumbers, the stretching terms are negligible, so that ψ_1 and ψ_2 have their own QG potential vorticity equations. ψ_2 decays as in the barotropic situation with a decay rate of approximately γ, and ψ_1 does not decay at all because there is no dissipation in the upper layer at all. The two modes are approximately

$$\begin{bmatrix} \psi_1 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 \\ \psi_2 \end{bmatrix}.$$

All of the above inferences are consistent with figure 10.12.

In figure 10.13, when γ gets large the decreasing of the decay rate can be attributed to the surface intensification. In the lower layer potential vorticity equation (see equation 10.21), the main balance of the baroclinic branch would be between $\frac{\partial}{\partial t}\frac{\psi_1}{\delta}$ and $-\gamma'\nabla^2\psi_2$. In such a balance, the larger the γ', the smaller the ψ_2 relative to ψ_1, thus more surface-intensified. The most interesting feature in the figure is the maximum decay rate for the baroclinic branch. The maximum (-0.0451) is obtained when k=-0.65, l=0 and γ=1.4. Therefore, the decay scale of any wave in the baroclinic branch (i.e., surface-intensified ones) is 20 times more than the time scale, which is defined here as $1/\beta R$.

Surface intensification is the reason for the small decay rates in the case of large γ. Because of the strong bottom friction (defined by a large γ but not by a fast decay, though we have no solid foundation for the range of γ) and no dissipation in the interior, the surviving waves should have minimum velocities near the bottom. Then the strong bottom friction does not dissipate much energy anymore, and a small decay rate will be seen. It is emphasized that a weak bottom flow and small decay rate do not mean there is no strong bottom friction. As argued, the strong bottom friction mandates a weak flow near the bottom, not the other way around.

10.3.2 Constant N^2 Model

In the two-layer model the bottom dissipation was in the Rayleigh form. In a continuously stratified model, we will represent the bottom dissipation by the bottom Ekman boundary layer theory. To keep the problem simple and tractable, we will study a constant N^2 model.

The model

The dimensional linear QG potential vorticity equation with bottom dissipation is

$$\begin{cases} \frac{\partial}{\partial t}(\nabla^2\psi + \frac{\partial}{\partial z}\frac{f^2}{N^2}\frac{\partial\psi}{\partial z}) + \beta\frac{\partial\psi}{\partial x} = 0 \\ z = 0, \quad \frac{\partial\psi}{\partial z} = 0 \\ z = -H, \quad -\frac{|f|}{N^2}\frac{\partial^2\psi}{\partial z\partial t} = \mu\nabla^2\psi \end{cases} \tag{10.25}$$

where we have used the bottom Ekman vertical velocity

$$w_E = sign(f)\sqrt{\frac{A_v}{2|f|}}\xi = sign(f)\mu\xi = sign(f)\mu\nabla^2\psi.$$

A_v, the vertical eddy viscosity, is assumed constant in the bottom Ekman boundary layer. μ is positive, with the dimension of length.

Choosing the following vertical, horizontal and time scales

$$H, \quad R = \frac{NH}{\pi f}, \quad \text{and} \quad T = \frac{1}{\beta R},$$

to non-dimensionalize the potential vorticity equation and boundary conditions, and assuming $\psi = \psi'(z)e^{\mathbf{i}(kx+ly-\omega t)}$ we have

$$\begin{cases} \frac{\partial^2\psi'}{\partial z^2} + \pi^2(-\frac{k}{\omega} - (k^2+l^2))\psi' = 0 \\ z=0, \quad \frac{\partial\psi'}{\partial z} = 0 \\ z=-1, \quad \frac{\partial\psi'}{\partial z} = \mathbf{i}\mu'\frac{k^2+l^2}{\omega}\psi' \end{cases} \tag{10.26}$$

where

$$\mu' = \mu\frac{HN^2}{|f|L^2}\frac{1}{\beta R} = \pi^2\frac{\mu}{H}\frac{|f|}{\beta R}.$$

μ' is the sole explicit parameter of the model. Its possible range is estimated as follows. The analysis from the JAMSTEC model (appendix F) shows that μ is about 10 m in the mid-latitude Indian Ocean, thus for a 4000 m depth ocean $\frac{\mu}{H} \sim 0.003$. The gravity wave speeds for the constant N^2 model are $c_n = \frac{NH}{n\pi}$ $(n = 1, 2, ...)$, then we have $R = c_1/f$ and $\frac{f}{\beta R} = \frac{f^2}{\beta c_1} \approx \frac{10^{-8}}{2\times10^{-11}\times2.5} = 200$, where we have assumed $c_1 \sim 2.5$m s^{-1}. Therefore

$$\mu' \sim \pi^2 \sim 10.$$

This is certainly a rough estimate; we will consider any value between 0.1 and 100 is possible. We also notice that N^2 in the bottom boundary condition should take the value near the bottom (N_b^2), which is much smaller than the depth-average $\overline{N}^2$ used to define the horizontal scale. Therefore there may be a factor $\frac{N_b^2}{\overline{N}^2}$ in μ' at first glance. According to the WKB scaling, however, the vertical scale in $\frac{\partial}{\partial z}$ is large near the bottom and proportional to $\frac{\overline{N}}{N_b}$. Thus the overall factor should be $\frac{N_b}{\overline{N}}$ in μ'. Assuming that the buoyancy frequency squared can be approximated by an exponential profile with a decay scale 1000 m: $N^2 = N_0^2e^{z/1000}$, then the factor $\frac{N_b}{\overline{N}} \approx 0.3$ for a 4000 m depth ocean. If the decay scale is 3000 m (likely in the Southern Ocean), then the factor is approximately 0.7.

Dispersion relation

To solve equation 10.26, we define

$$A_n^2 = -\frac{k}{\omega} - (k^2 + l^2).$$

If $\mu' = 0$, then the solution of equation 10.26 will be a function of A_n^2 only, with no additional dependence on (k, l, ω). In this sense, the vertical structure (of the usual dynamical modes without bottom dissipation) is independent of wavenumber and frequency. This is not the case for the dissipative Rossby waves.

The possible solutions of equation 10.26, satisfying the upper boundary condition $\partial\psi'/\partial z|_{z=0} = 0$, are $\psi' = \cos(A_n \pi z)$. Substituting this into the lower boundary condition $\frac{\partial\psi'}{\partial z}|_{z=-1} = \mathrm{i}\mu' \frac{k^2+l^2}{\omega}\psi'$, we have

$$\tan(A_n \pi) = \mathrm{i}\frac{\mu'(k^2 + l^2)}{A_n \omega \pi}. \qquad (10.27)$$

For A_n is a function of (k, l, ω), equation 10.27 is the dispersion relation between k, l and ω with parameter μ'.

Due to the periodicity of $\tan(A_n\pi)$, we have multiple modes as with the usual dynamical modes. For given μ' and l, the dispersion curves (relationship between k and real component of ω, i.e., ω_r) for the first two modes are displayed along with the imaginary component (ω_i) and vertical structures (amplitude and phase) in figures 10.14 to 10.17. The two dispersion curves cross in some plots, such as figure 10.16 for $\mu' = 10$ and l=0. This does not mean that there are multiple roots in equation 10.27, since the corresponding ω_i are different.

At first, let's look at the results for l=0. When μ' =10 and 30, the results are comparable to those of the two-layer model with bottom dissipation. We see that the barotropic branch is bottom-intensified and has a large decay rate; and the baroclinic branches are surface-intensified and have small decay rates. For the baroclinic branches, the decay rate decrease for both high and low wavenumbers. We also notice the decay rate is smaller for $\mu' = 30$ than for $\mu' = 10$. This indicates that there is a maximum decay rate, similar to the two-layer model, although the parameter there is γ instead of μ'. The behavior for μ' =1 and 3 is strange. The baroclinic branch looks like the usual dynamical baroclinic mode at the low wavenumber end, but it becomes bottom-intensified at the high wavenumber end. The decay rate does decrease towards the low wavenumber end but not towards the high wavenumber end. These results would be more comparable to the two-layer model and to the results for larger μ', if we combined the low wavenumber part of the baroclinic branch with the high wavenumber part of the barotropic branch.

Results for l=0.5 and 1 are similar to those for l=0. Only the dispersion curves are shown in figure 10.19. For l=1, the frequency of the baroclinic branch is higher than that of the barotropic branch; this is contrary to the usual dynamical modes.

In the two-layer model, the phase shift of the baroclinic branch lies in the third quadrant (-180, -90). Here the phase of the baroclinic branch could be in the third and fourth quadrants (-180, 0). Similar to the two-layer model, the stronger the dissipation, the smaller the phase shift. Comparing the panels at the lower-left corner of figures 10.14 to 10.17 (l=0), the contours are more squeezed towards the bottom (z=-1) as μ' is increased.

Dissipative Rossby wave with $\psi|_{z=-H} = 0$

The two-layer model in the previous section clearly shows that bottom dissipation modifies the usual dynamical modes into two non-orthogonal modes: one surface-intensified and decaying slowly and another bottom-intensified and decaying quickly. The continuously stratified model in this section is more complicated but the general trend is the same. The stronger the bottom dissipation (in terms of μ, not of the decay rate) the clearer this trend. Since the barotropic branch decays quickly, we concentrate on the baroclinic branch. When the bottom dissipation is 10 time larger than the estimate from the JAMSTEC model, the baroclinic branch becomes unidirectional and surface-intensified and ψ approaches zero near the bottom (left column, figure 10.18) except at the smallest wavenumbers. This prompts us to look at the Rossby waves with $\psi|_{z=-H} = 0$:

$$\begin{cases} \frac{\partial}{\partial t}(\nabla^2\psi + \frac{\partial}{\partial z}\frac{f^2}{N^2}\frac{\partial\psi}{\partial z}) + \beta\frac{\partial\psi}{\partial x} = 0 \\ z = 0, \quad \frac{\partial\psi}{\partial z} = 0 \\ z = -H, \quad \psi = 0 \end{cases} \tag{10.28}$$

Compared with the usual Rossby wave equation, equation 10.28 abandons the bottom condition on vertical velocity $w = 0$. $\psi = 0$ at the flat bottom is equivalent to $u = v = 0$ under geostrophy.

It is interesting to compare the results from equation 10.28 with those of equation 10.25. Using the same scales and notations as in equation 10.25, we have following dispersion relationship

$$\cos(A_n\pi) = 0, \quad \text{i.e.} \quad A_n\pi = \frac{2n-1}{2}\pi \quad (\text{n} = 1, 2, \cdots)$$

That is

$$\omega = -\frac{k}{k^2 + l^2 + (\frac{2n-1}{2})^2} \quad (\text{n} = 1, 2, \cdots). \tag{10.29}$$

The vertical structure is $\cos(A_n\pi z) = 0$ with $\text{z} \in [-1, 0]$, which has a zero vertical phase shift for $n = 1$. In terms of the dispersion relation (figure 10.19), vertical phase shift (lower-left of the figure 10.18) and vertical structure (figure 10.20),we would conclude the $\psi|_{z=-H} = 0$ Rossby waves are the limiting cases of the dissipative Rossby waves. For small (k,l) waves, i.e., the long Rossby waves, the $\psi|_{z=-H} = 0$ model shows that the westward phase speed (equation 10.29, n=1 and ignoring $k^2 + l^2$ in the denominator) and group speed are four times faster than those of the usual Rossby wave of the first mode. Recall, with present characteristic scales, the speed of the long Rossby waves of the usual first baroclinic mode is -1.

The $\psi|_{z=-H} = 0$ model does not have any decay mechanism. For strong bottom dissipation, the dissipation happens within a thin layer near the bottom in which phases change rapidly (see figures 10.18, for example). Outside the thin layer, the two models (the $\psi|_{z=-H} = 0$ model and the $\frac{\partial \psi'}{\partial z}|_{z=-1} = \mathbf{i}\mu' \frac{k^2+l^2}{\omega}\psi'$ model with a large μ') are similar. The $\psi|_{z=-H} = 0$ model excludes the barotropic bottom-intensified branch, which will decay quickly anyway. In a word, the $\psi|_{z=-H} = 0$ model is physically meaningful, and represents the main aspects of the Rossby waves with strong bottom dissipation. Mathematically, the boundary condition:

$$z = -1, \quad \frac{\partial \psi'}{\partial z} = \mathbf{i}\mu' \frac{k^2 + l^2}{\omega} \psi'$$

in equation 10.26 requires that $\psi \to 0$ if $\mu' \to \infty$.

As a serious drawback of present work, evidence is lacking regarding the rate of the bottom dissipation of the Rossby waves in particular. In the ocean, but not in the model, the dissipation near the bottom could be enhanced by small-scale bottom topography (Vanneste 2000).

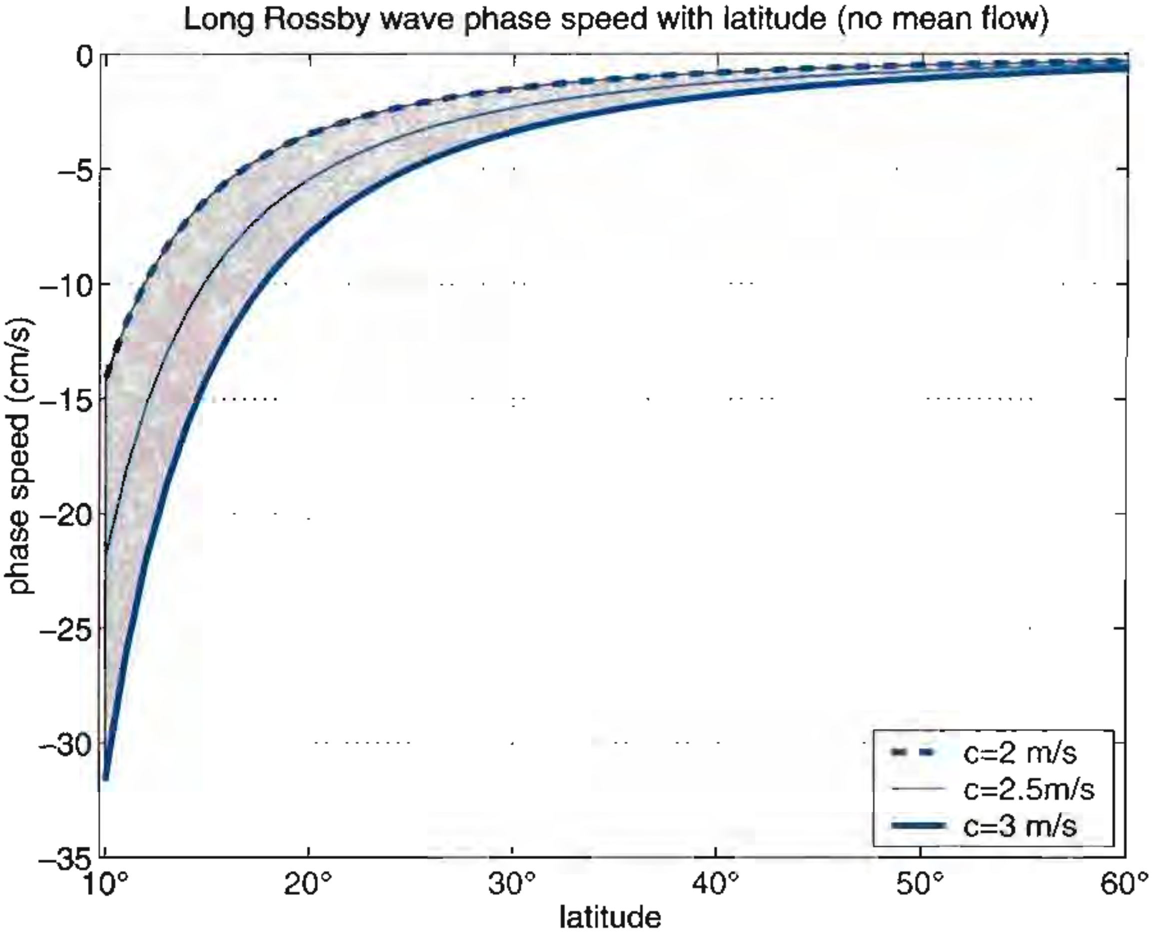


Figure 10.1: Latitudinal dependence of the long Rossby wave speed (c_L).

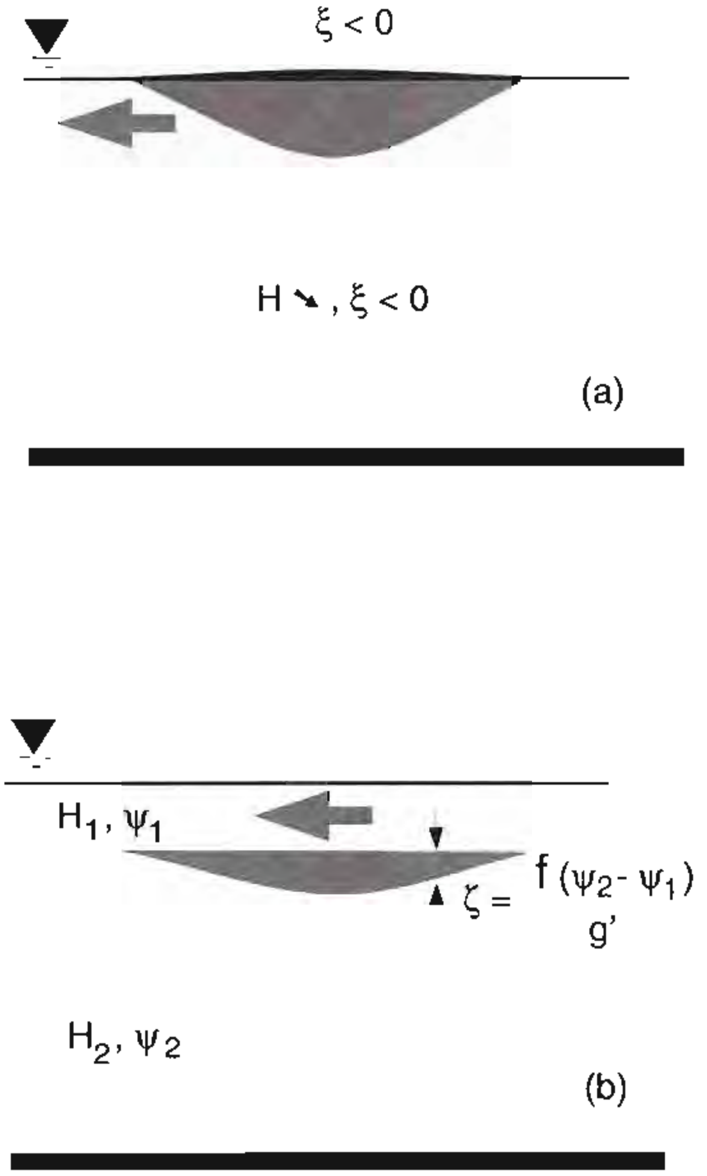


Figure 10.2: (a) The two-layer model for Gulf Stream rings (Flierl 1984). The warm upper layer has a finite volume, thus the interface outcrops. When the warm water moves westward, the lower layer is compressed, thus it will have a negative relative vorticity to conserve potential vorticity. Therefore, the upper and lower layers will rotate in the same direction. (b) The perturbation is isolated, but there is no interface outcropping as in the upper panel. The interface perturbation is not infinitesimal but of the order of the Rossby number, compared to the thickness of the upper layer.

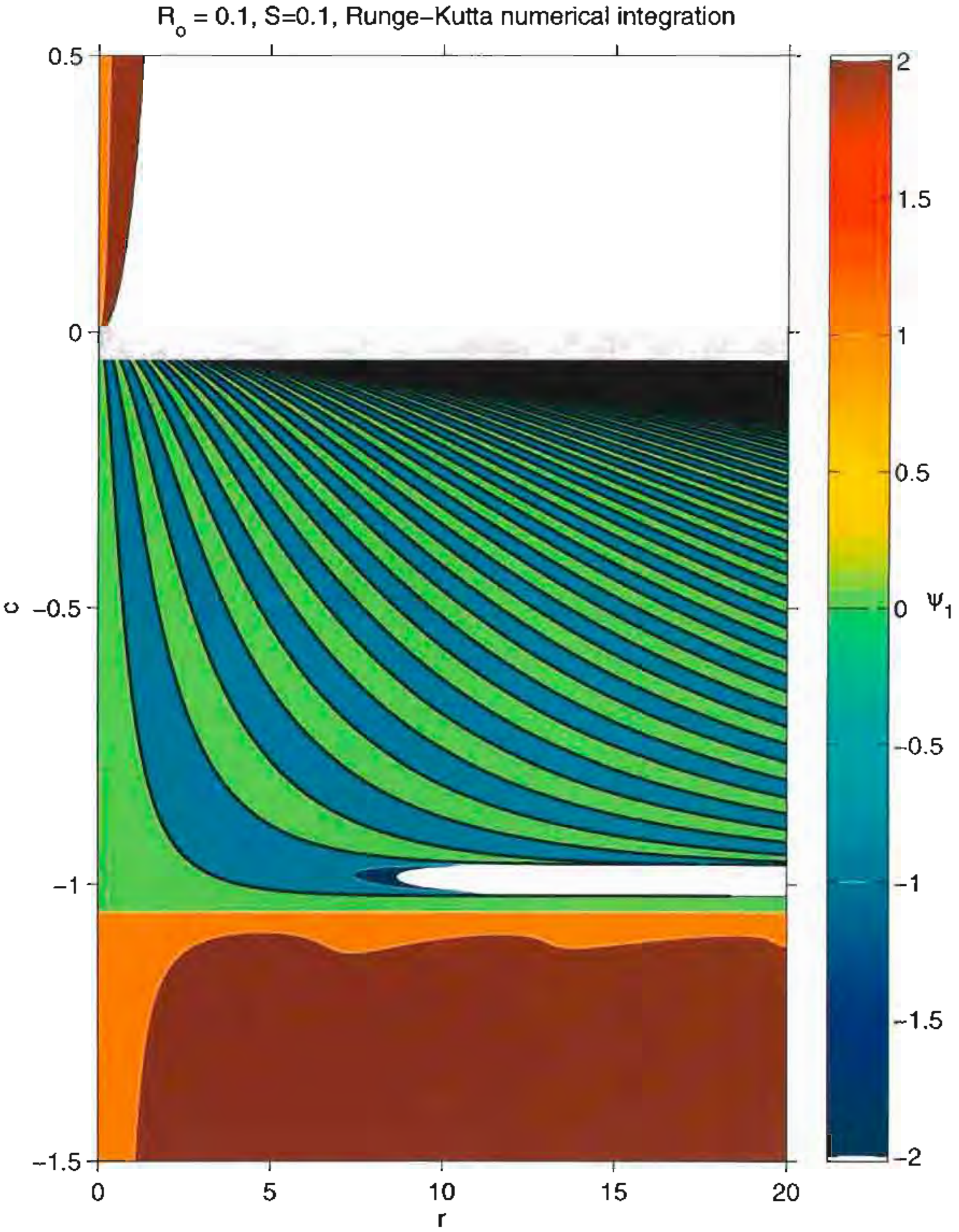


Figure 10.3: The numerical solution of equation 10.17 with $\psi_1 = 1$ at $r = 0$. The numerical integration is from $r = 0$ outward. The gray region is where the integration overflows. $R_o = 0.1$ and $\hat{S} = 0.1$. The color is the color of the lowest value within the contour range. For example, for the range [0 1], the color is that of the color at 0 in the color bar.

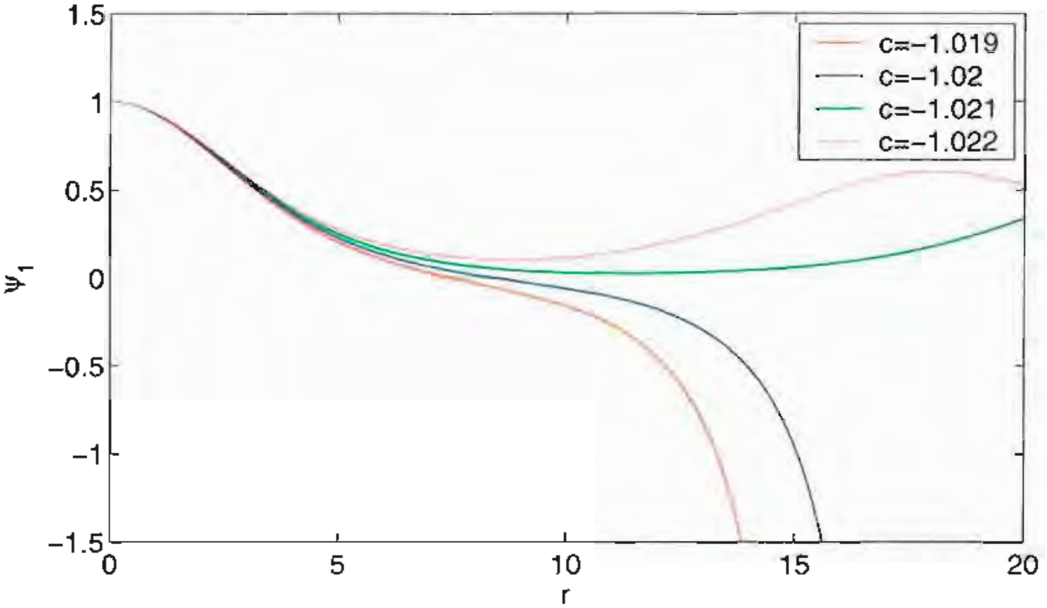


Figure 10.4: A close look at the solution of equation 10.17 near $c = -1.02$.

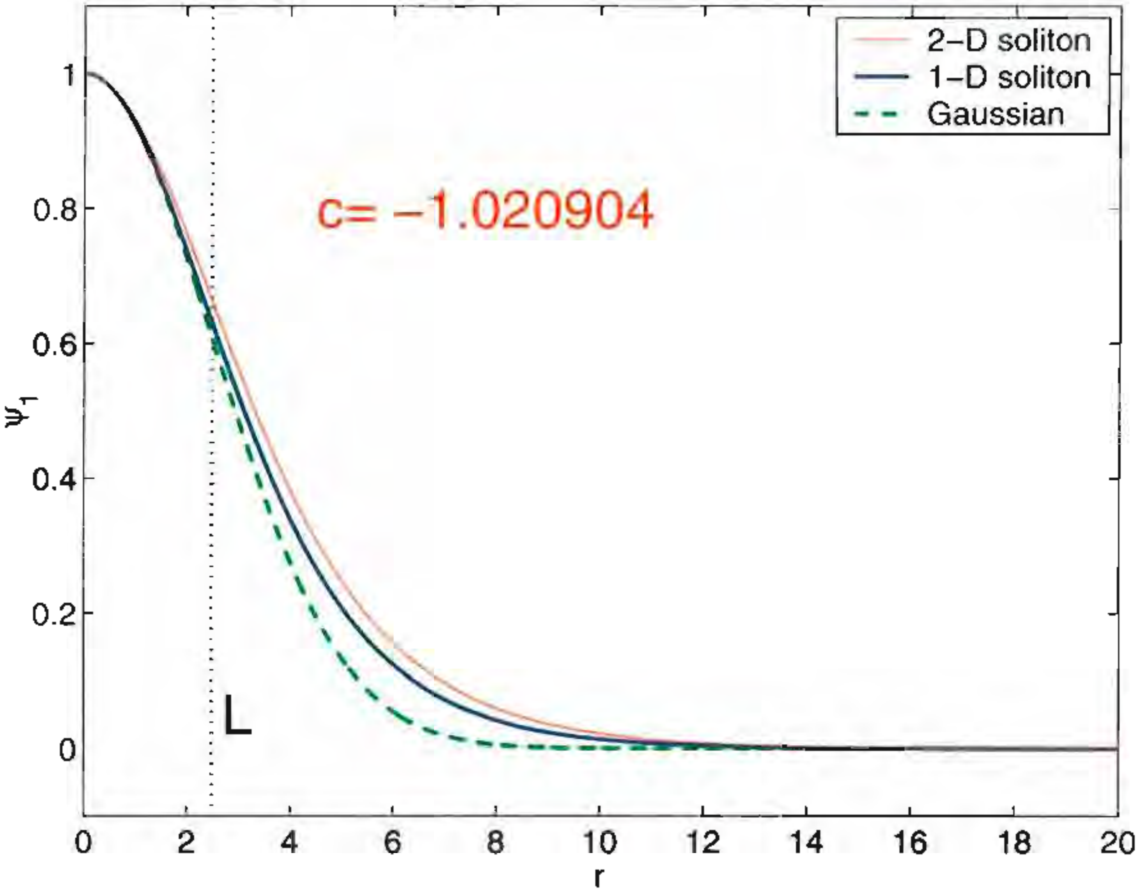


Figure 10.5: The shape of the 2-D soliton for $A = 1$ with parameters $R_o = 0.1$ and $\hat{S} = 0.1$. Also shown is the 1-D soliton with the same values for A, R_o and $\hat{S}$ (see appendix D). The Gaussian curve has the same decay scale as the 1-D soliton—refer to figure D.1.

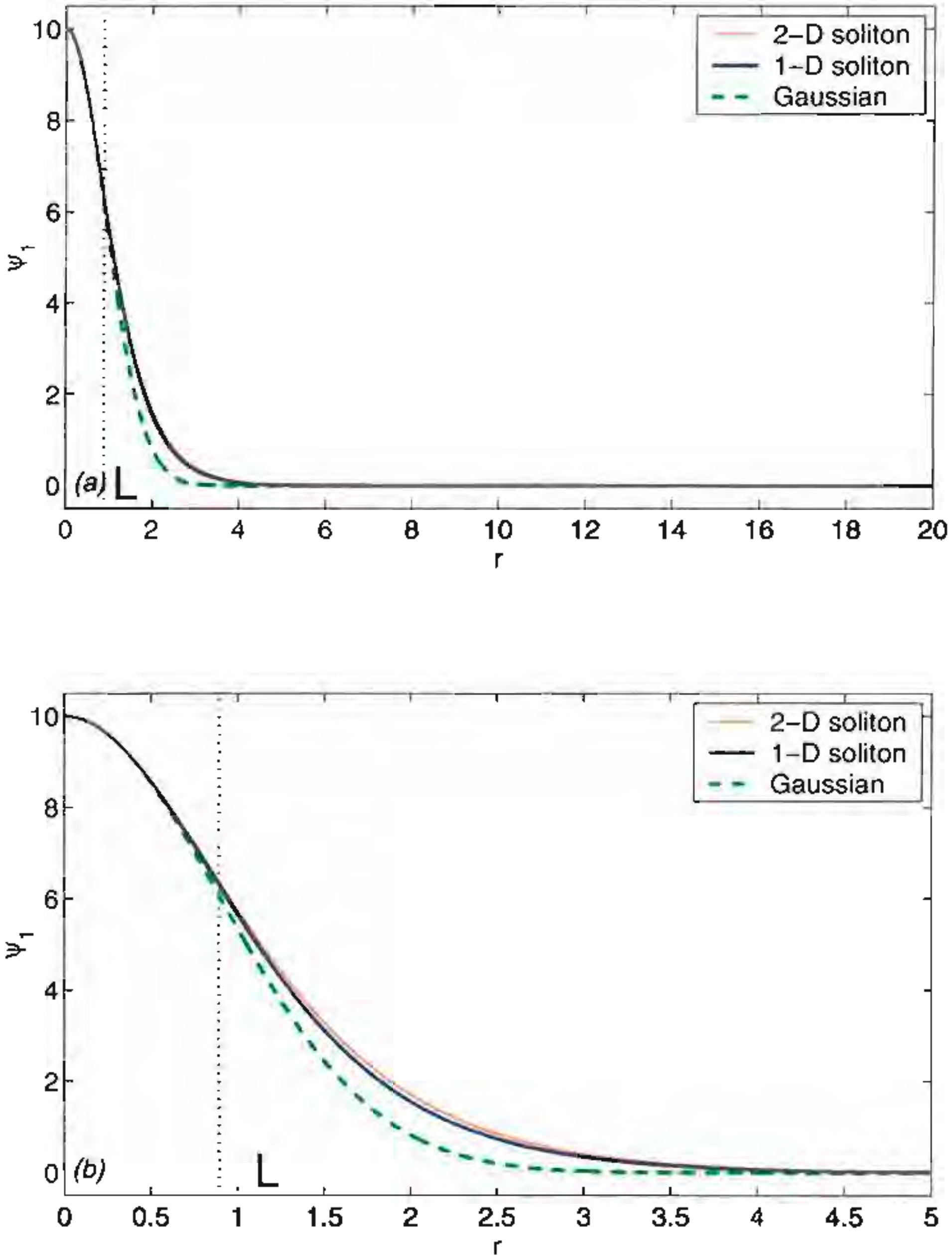


Figure 10.6: The shape of the 2-D soliton for $A = 10$ with parameters $R_o = 0.1$ and $\hat{S} = 0.1$. Also shown is the 1-D soliton with same values for A, R_o and $\hat{S}$. The Gaussian curve has same decay scale as the 1-D soliton. The lower panel is just a zoom-in of the upper panel

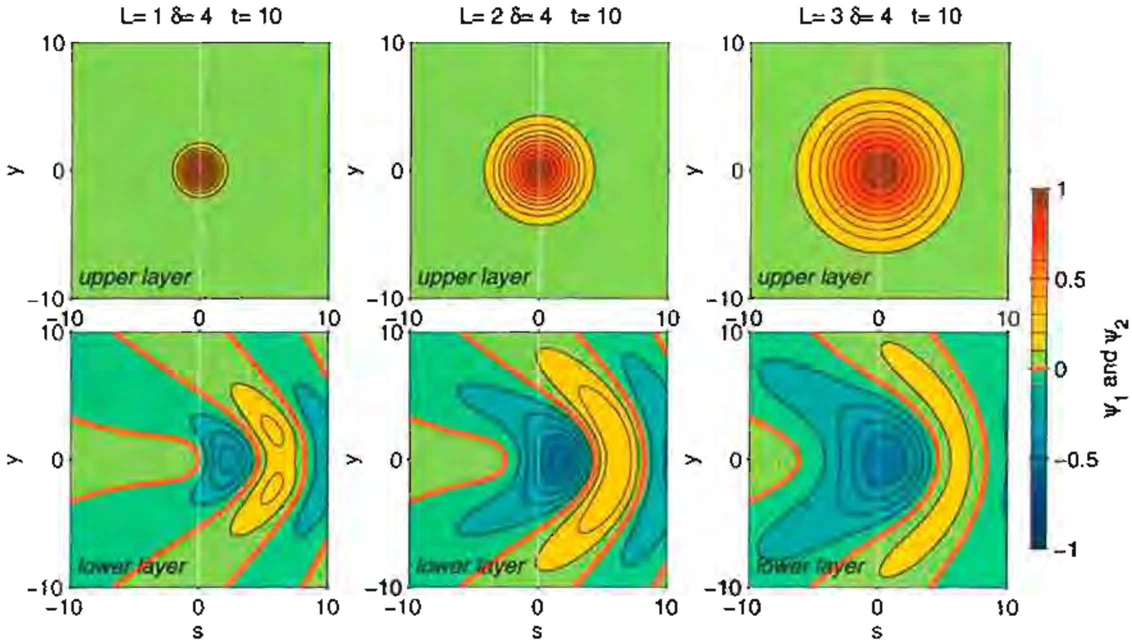


Figure 10.7: Lower layer streamfunction ψ_2 (lower row) at t=10. The lower layer, initially in the resting state, is forced by the Gaussian eddy in the upper layer (upper row). Left column for L=1, middle column for L=2 and the right column for L=3. All panels use the same color scale. The depth ratio is $\delta = 4$ for all columns.

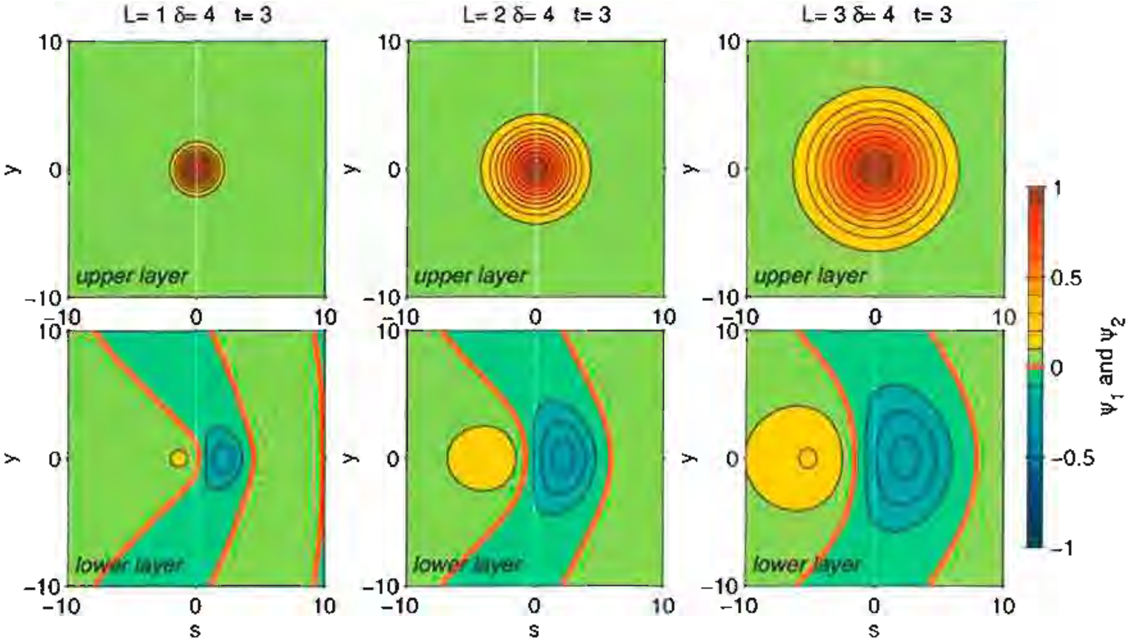


Figure 10.8: Same as figure 10.7 but for t=3; these are intermediate states before the quasi-steady states shown in figure 10.7.

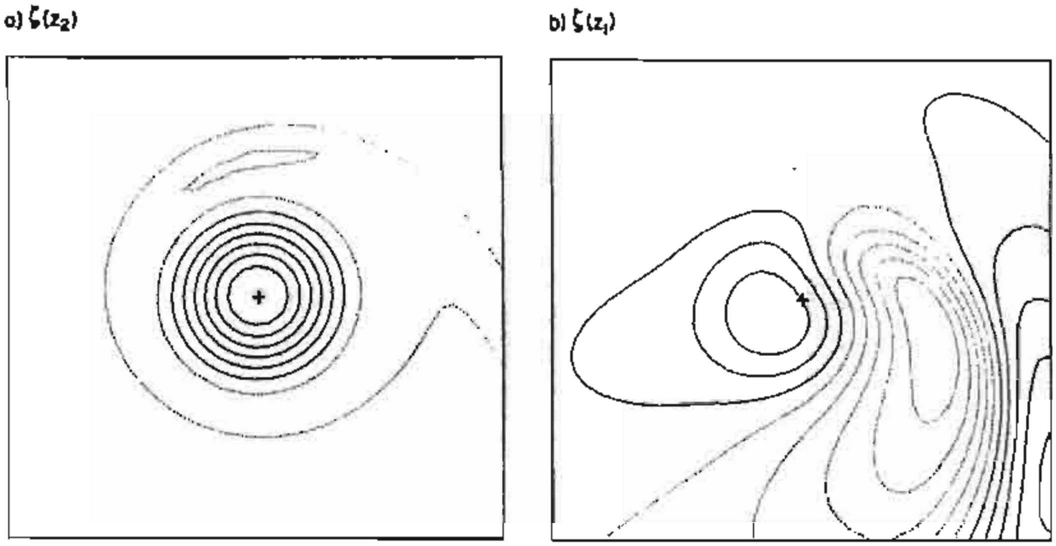


Figure 10.9: Relative vorticities of the upper (left) and lower (right) layers in the QG model of McWilliams *et al.* (1986). The contour interval is three times larger in the upper layer than in the lower layer. Using the notation of the present study, $L = 1$ and $\delta = 4$. The relative vorticities shown are the averages over time from t=2 to t=5 in our scaling.

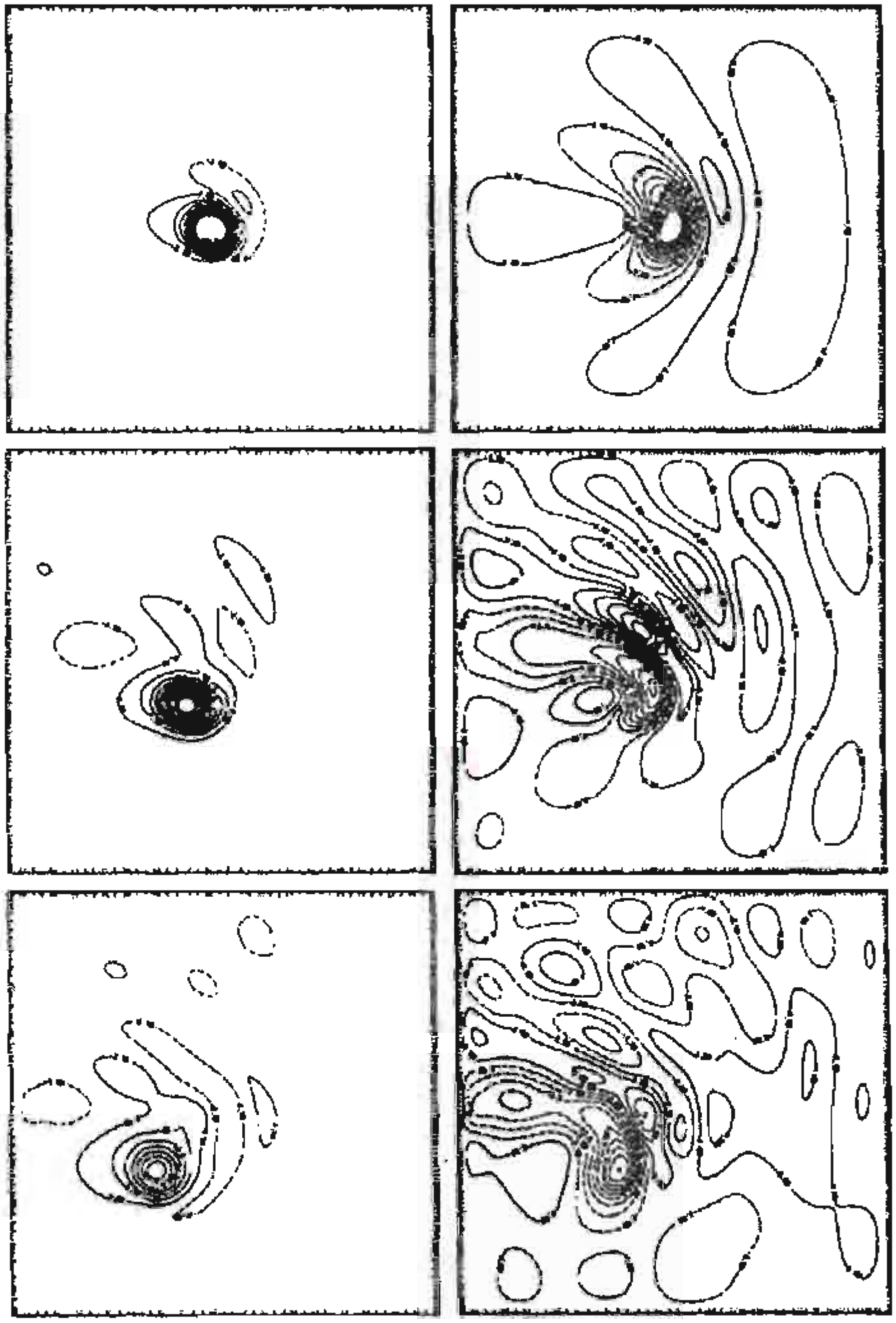


Figure 10.10: Interface elevation (left column) and transport streamfunction of the lower layer (right column) from Chassignet and Cushman-Roisin (1991). Using the notation of the present study, $L = 1.3$ and $\delta = 4$. The times for the rows are approximately t=1, 4 and 7 in the present scaling.

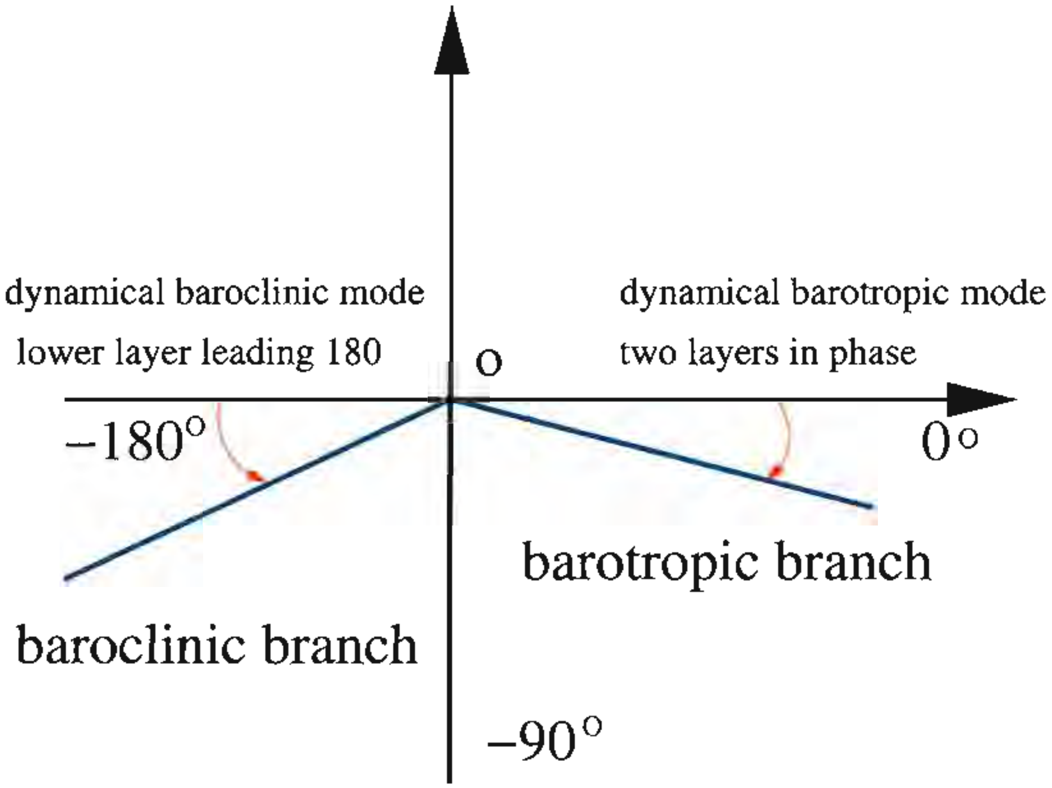


Figure 10.11: In the case of small γ and $k, l \sim O(1)$, the phase relation between the upper and lower layers. The blue lines are angle(ψ_2/ψ_1). If the blue line is -150°, then the lower layer is leading the upper layer 150°, given westward propagation. According to this definition, the lower layer of the usual dynamical baroclinic mode is leading the upper layer 180°, and the lower layer of the dynamical barotropic mode is leading the upper layer 0°, i.e, the two layers are in phase.

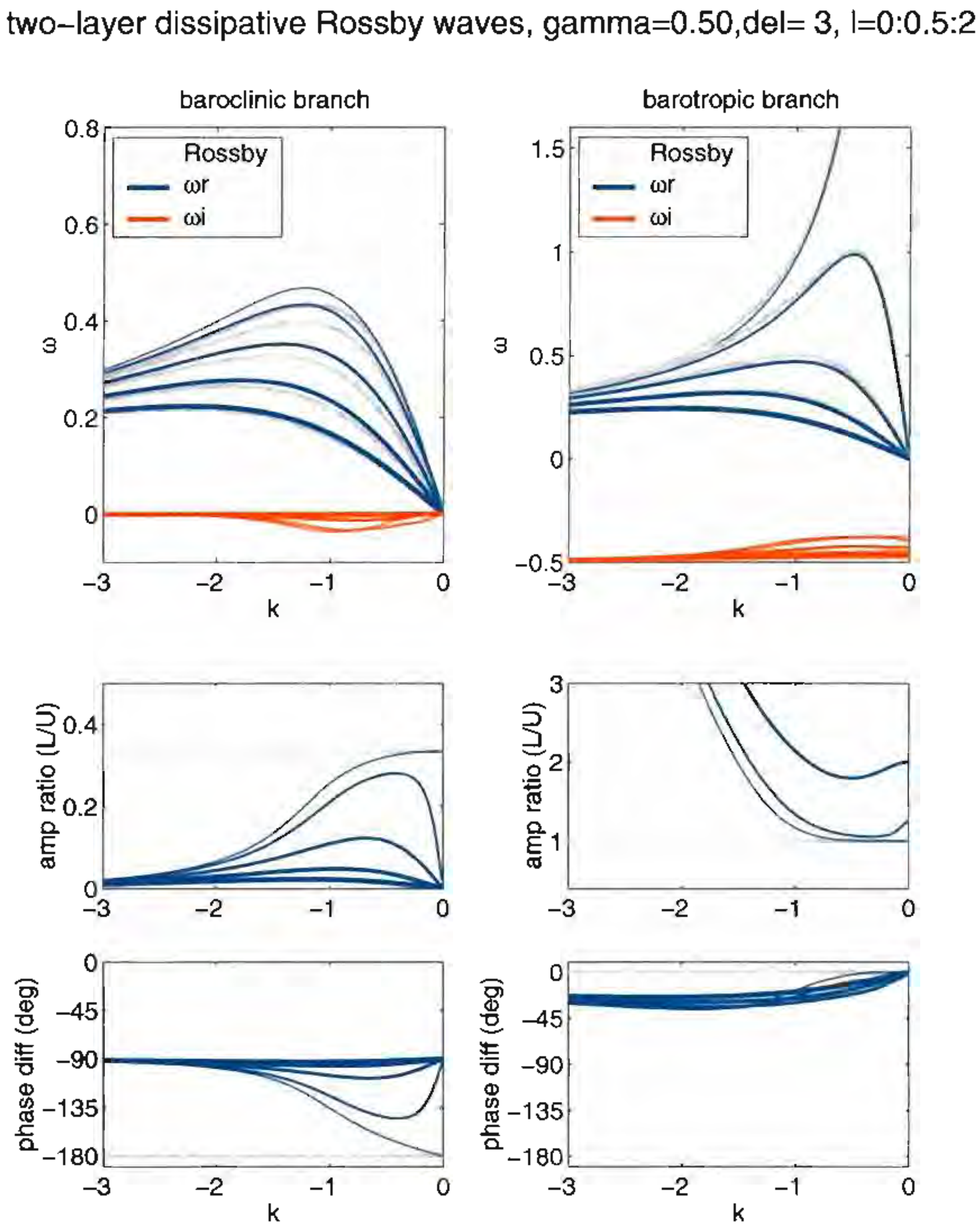


Figure 10.12: The dispersion relation and vertical structure of the dissipative Rossby waves in the two-layer model. Left column is the baroclinic branch and right column is the barotropic one. Four different lines are for different north-south wavenumbers: the thicker the line, the larger the wavenumber l. The "amp ratio (L/U)" of the middle row is $|\psi_2/\psi_1|$; the "phase diff (deg)" in the last row is angle(ψ_2/ψ_1), as in figure 10.11

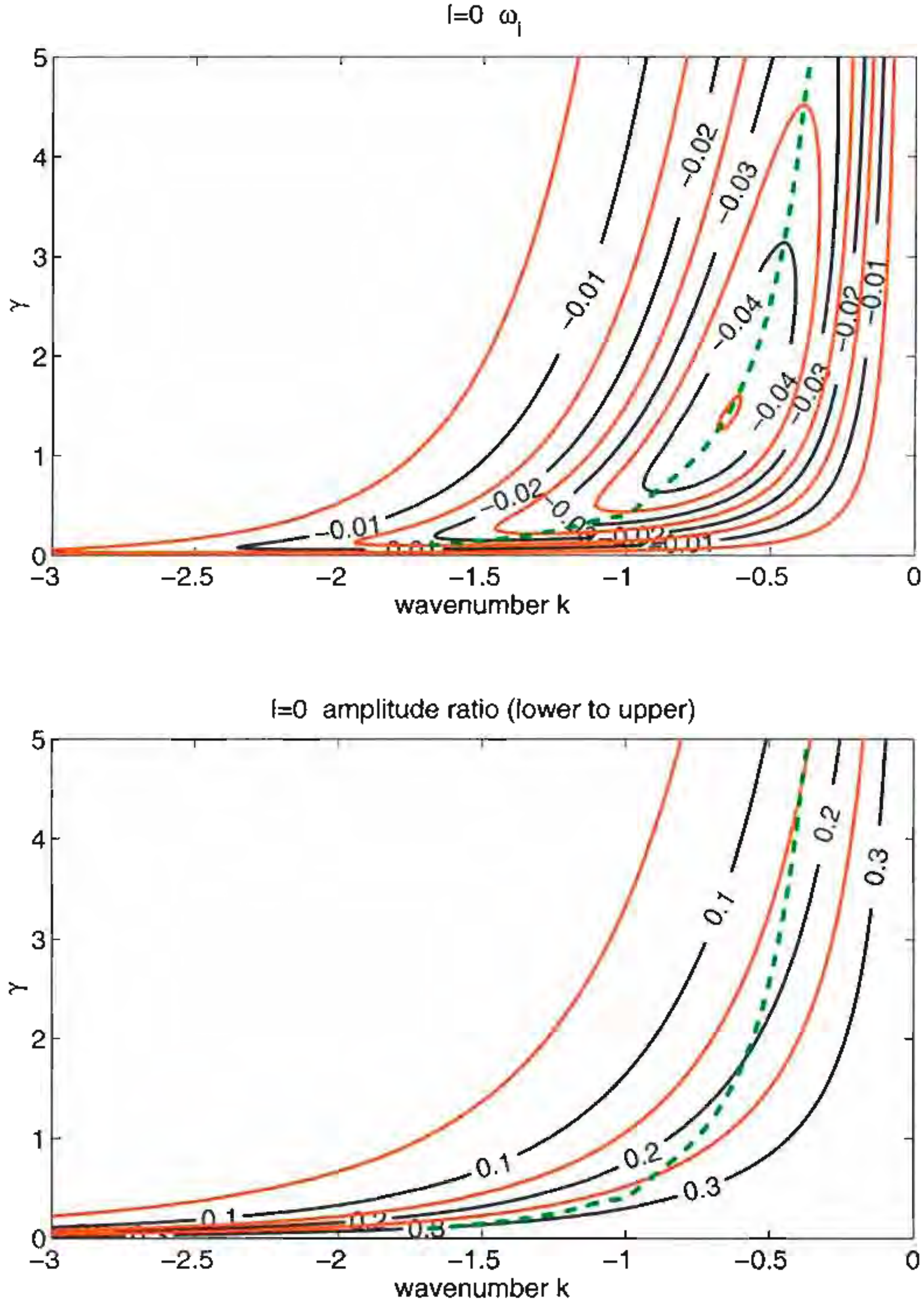


Figure 10.13: Both panels are for the baroclinic branch and north-south wavenumber l=0. Upper panel: the decay rates ω_i as γ changes. Lower panel: the vertical structure (ratio of the lower layer to the upper layer ($|\psi_2/\psi_1|$). The red contours are between the black labelled contours. The maximum in the upper panel is $\omega_i = -0.00451$. And the maximum in the lower panel is $\frac{1}{\delta} = 0.33$ without any dissipation (see middle row of figure 10.12). The green lines indicate where ω_i reaches its local extreme for a given γ.

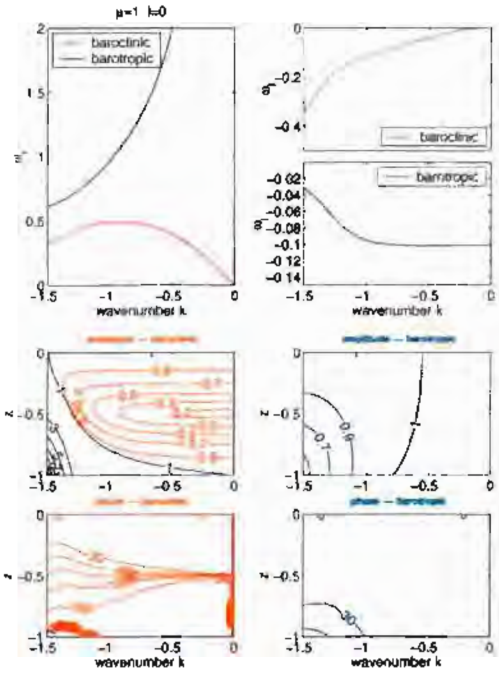


Figure 10.14: Dispersion relationship for $\mu' = 1$ and l=0.

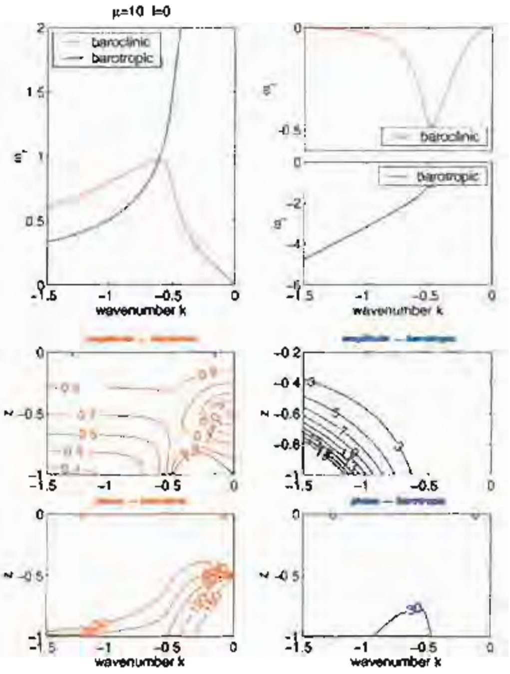


Figure 10.16: Dispersion relationship for $\mu' = 10$ and l=0.

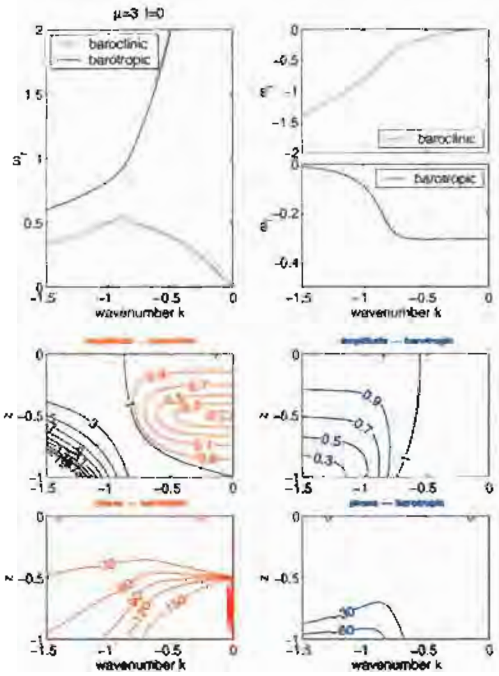


Figure 10.15: Dispersion relationship for $\mu' = 3$ and l=0.

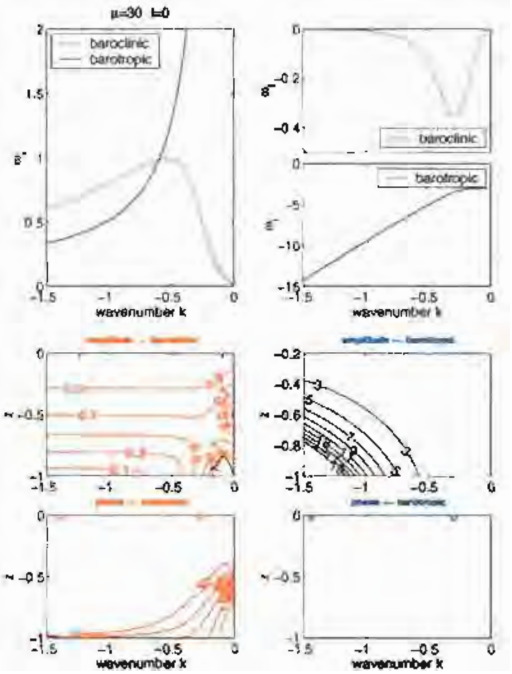


Figure 10.17: Dispersion relationship for $\mu' = 30$ and l=0.

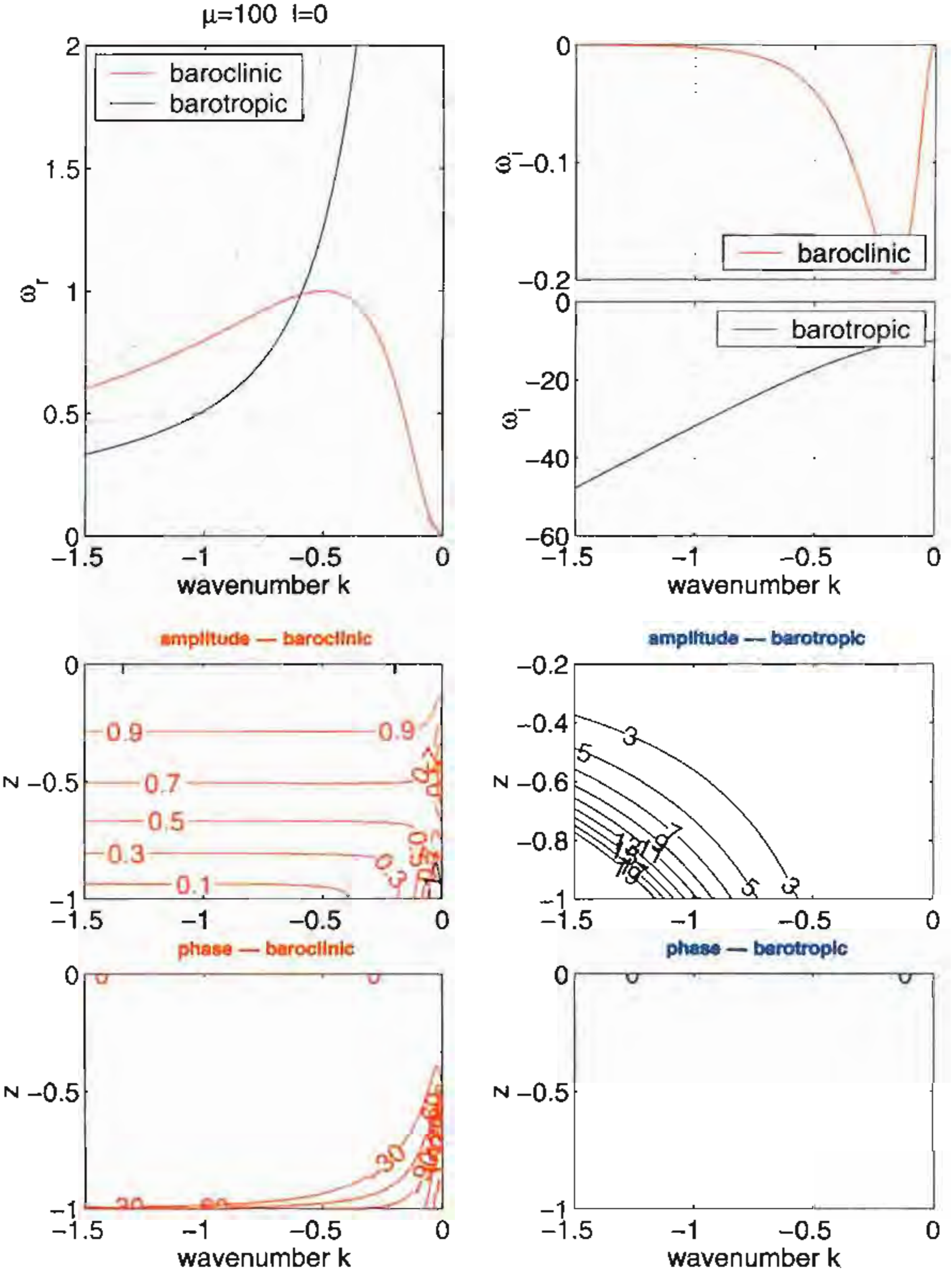


Figure 10.18: Dispersion relationship for $\mu' = 100$ and l=0.

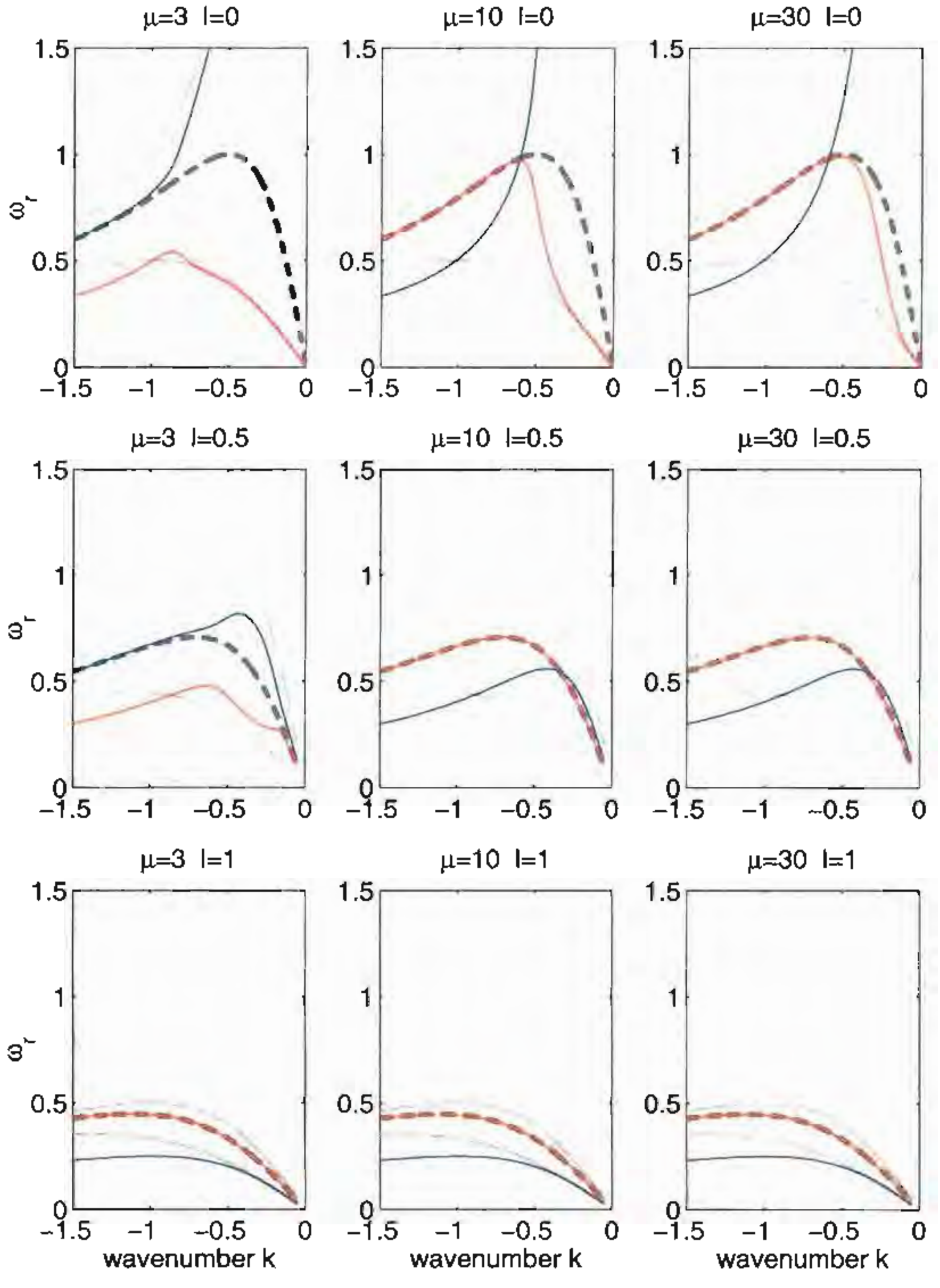


Figure 10.19: Dispersion relationships for the usual barotropic and first baroclinic Rossby waves (the upper and lower gray lines respectively), the dissipative Rossby waves (red and blue lines for the barotropic and baroclinic branched respectively) and the Rossby waves with bottom boundary condition $\psi|_{z=-H} = 0$ (dash line).

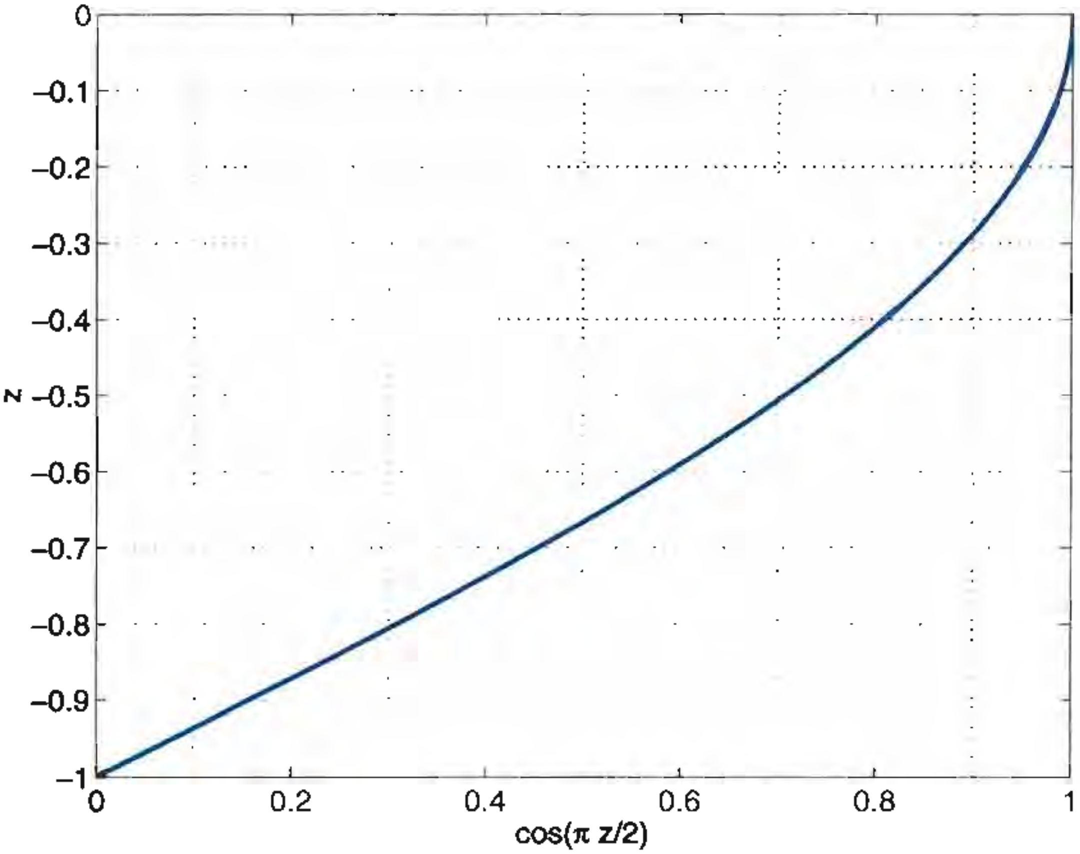


Figure 10.20: The vertical structure of the Rossby waves with bottom boundary condition $\psi|_{z=-H} = 0$. $n = 1$.

CHAPTER 11

CONCLUSIONS

The LADCP profiles show that the dominant vertical structures of mesoscale motions in mid-latitudes are unidirectional and surface-intensified. The JAMSTEC model output further shows that these structures propagate westward at approximately the speed of first baroclinic mode long Rossby waves, even at frequencies above the Rossby wave cutoff.

The study of the dynamics is not conclusive, because of the simplicity of the analytically tractable models and the uncertainty about the parameters therein. The following results, however, would serve as hypotheses in future studies:

- Shear modes could be important in high latitudes;
- The two-layer non-linear model with finite interface perturbation could marginally explain the vertical phase shift, particularly if considering the transitional nature of the flow;
- Bottom dissipation could explain the feature in mid-latitudes, but we do not know much about it.

11.1 Observation: LADCP

In order to understand the vertical structure of mesoscale ocean currents, a large number of LADCP profiles in the Indian Ocean, with extensive horizontal coverage, is used in this study. We presume that mesoscale motion dominates our LADCP profiles; this is supported by the favorable comparison between the LADCP velocity and CTD geostrophy.

The LADCP velocity is dominated by the low modes. The barotropic mode absolutely dominates in high latitudes. In mid-latitudes, we see the approximate equipartition between the barotropic and first baroclinic modes; each contains about 25% of the total variance, and the second baroclinic mode contains about 12%. Near the equator, the kinetic energy spreads over a few low modes but the fifth mode has less than 10% of the total kinetic energy.

The dominant vertical structure in middle and high latitudes, as quantified by the first EOFs with a variance of ~40% and ~80% of the total respectively, is that the velocity tends to be unidirectional and surface intensified. This is also suggested by the positive correlation between the barotropic and first baroclinic modes of the LADCP profiles, and by the vector alignment of the barotropic and first baroclinic modes of the LADCP profiles.

The correlation between the barotropic and first baroclinic modes conflicts with the classical linear Rossby wave theory. If the theory holds, the barotropic and baroclinic modes will disperse, so they should be uncorrelated in an ensemble of profiles such as ours.

An LADCP dataset in the Indian Ocean is used in this study, because it is the most complete, but preliminary calculations indicate that results are similar in other oceans.

11.2 Numericial Modeling: JAMSTEC Model

After finding a favorable comparison between the first EOFs from the LADCP profiles and from the JAMSTEC model, we used the latter to explore the horizontal and temporal characteristics of the dominant vertical structure. The major contributions are

- At each latitude, the first EOF propagates westward at the same speed as has been found from altimetric SSH anomalies.
- The westward propagating vertical structure has a phase shift in the vertical, with the lower layer leading the upper layer. Therefore, the first EOF alone is not a dynamical mode: $\psi(x, y, z; t)$ is not separable such as $\psi(x, y, z; t) \approx f(x, y; t)g(z)$.

In more detail, the conclusions from the analysis of the model output are:

1. We surveyed the eddy energy distribution horizontally and vertically. Beside the ACC region and the region near the major topography Chagos-Laccadive, the strong eddy bands correspond to two generating sites: the throughflow near the model Lambok Strait and the Leeuwin Current along the west coast of Australia. The rms is above 20cm s^{-1} near the surface and $\sim$3cm s^{-1} at 3000 m depth.
2. We calculated the first EOF at each location. The variance in the first EOF ranges from 50% of the total in weak eddy energy regions and up to 80% in strong eddy energy regions. The vertical structure of the first EOFs is quite similar throughout the study region: unidirectional and surface intensified. Because of the high percentage of variance and similarity of the shape of the first EOF, it is a good descriptor of the vertical structure of mesoscale currents.
3. Based on lag-coefficient and counting methods, the characteristics of the westward propagation of the first EOF is similar to those inferred from the altimetric observations, although the latter is mainly for longer length scales. In mid-latitudes, the speeds are close to those of the long Rossby waves. In high

latitudes (south of 30°S) the inferred speeds are faster than those of the long Rossby waves of the first baroclinic mode, while in the lower latitudes the inferred speeds are slower. The coherence method shows that, in mid-latitudes, the propagation characteristics remain the same even at frequencies above the Rossby wave cutoff.

4. Although the first EOF is a good descriptor, other EOFs are dynamically important. If the original fields could be approximated by the first EOF itself, we would expect no change in phase over the depth. However, we identified a lower layer phase advance of about 90° in mid-latitudes. This important feature will be used as one of the facts, against which the theories will be tested in the theoretical part of this study.

The following features are contrary to the linear Rossby wave theory:

- The correlation of the barotropic and first baroclinic modes;
- Different dispersion relations from those of linear Rossby waves;
- The phase leading of the lower layer relative to the upper layer, given westward propagation.

11.3 Theory: Two-layer Models

A few simple models (all but one are two-layer models) are utilized to explore the dynamics relevant to the dominant vertical structures in the observation and JAMSTEC model output. Two aspects of the dynamics, dispersion and westward propagation, are studied; the major portion of the study is focused on the westward propagation:

1. Using Gaussian eddies as the simple prototype of the eddy field, we show that, under the classic linear dynamics, the barotropic and baroclinic modes of the Gaussian eddy will soon be uncorrelated even though they initially are totally correlated.

2. The vertical structures of shear modes could be similar to the dominant vertical structure, provided that the mean flow shear is large compared to the speeds of the long Rossby waves (c_L). However, this is unlikely to be true in the mid to low latitudes. Also notice that $|c_L|$ varies by a factor of three from *20° to 10°*. However, the situation in the high latitudes may be different: the mean flow is well defined on the order of tens cm s^{-1} and c_L is 1-2 cm s^{-1}, so that the shear modes could be significant.

3. The non-linear model considers a finite interface perturbation. The model does not satisfactorily explain the dominant vertical structure. However, there are two features worthing noting. One is that in the quasi-steady solution the negative maximum of the lower layer is always east of the positive maximum of the upper layer. The positive extreme (but weak) being west of the positive maximum of the upper layer suggests that the lower layer is leading the upper layer by less than 180°. This is not inconsistent with the phase shift from the GCM output. No other known dynamics so far provides such a vertical structure. Another feature is that, during the intermediate steps, the positive extreme is better defined and the phase advance of the lower layer approaches 90°: the upper layer positive maximum is located approximately between the lower layer positive and negative extremes. This is more consistent with the vertical phase shift inferred from the GCM output.

4. Bottom dissipation is a dynamically possible cause for the discrepancy between the observation and classical theory if it is strong enough—but we do not know much about bottom dissipation. In the two-layer model, the barotropic branch is bottom-intensified and decays quickly. So the more interesting mode is the baroclinic branch, which is surface-intensified and decays slowly. The decay scale for the baroclinic branch in the two-layer model is more than 20 times larger than the Rossby wave time scale ($\frac{1}{\beta R}$). The inefficiency of its decay is because of the potential energy accompanying the kinetic energy and surface intensification (or weak flow near the bottom). Much different decay rates for the barotropic and baroclinic branches may suggest a natural selection – the baroclinic branch would be more observable.

   We usually argue that the bottom dissipation will not be important because of the surface intensification. Here we show that a strong bottom dissipation (in terms of γ or μ—the coefficients) without interior dissipation mandates the surface intensification. The phase shifts in the baroclinic branch of the present two layer model are close to 90°. This is consistent with the vertical phase shift in the GCM output. It is demonstrated that the $\psi|_{z=-H} = 0$ model is the limiting case of dissipative Rossby waves. For long Rossby waves, the westward speed of the waves could be enhanced no more than 4 times.

11.4 The Future

A global description of the vertical structures from observations and from model output is necessary. The mid-latitudes of the Indian Ocean are eddy energetic, as is the northern Atlantic, where the vertical structure seen in LADCP profiles seems similar. How about other parts of world oceans, which have less eddy energetic regions such as Southeast Pacific? Another development could be to analyze the model output forced by the ECMWF winds from 1982 till 2000 (ECMWF run), other than by the climatological winds (climatological run, used in this study). Both differences and agreements from these two runs would be interesting.

This study proceeded from observation to model output and then to theory. We would go though the procedure more than once, in order to test the results from simple models, specifically:

- **Shear Modes** We concluded the shear modes could play a role in the high-latitudes. Could we identify it using the JAMSTEC model output, considering the difficulty in defining mean flow? Is the non-dimensionalized mean flow shear $U_s^* = 2K^2R^2U_s/|c_L|$ actually very useful for understanding the role of the mean flow shear on the vertical structures?

- **Non-linear Model** The model is highly simplified for one special type of non-linearity—finite interface perturbation. The single Gaussian eddy configuration is also very idealized. Meanwhile, the model output shows that the Jacobian could be important and that adjacent eddies could interact. Therefore the challenge is to identify the dynamics in such a complicated evolving field.

- **Bottom Dissipation** The preliminary diagnoses showed the bottom dissipation could be important in the model (appendix F). The first question is whether the framework of the Ekman bottom layer is good for our purpose, since the Ekman bottom layer is not resolved by the model and the inferred vertical eddy viscosity is as large as that near the surface. Could we diagnose the energetics to see similar bottom dissipation? With respect to the spatial distribution of the correlation coefficients (figure F.2), if we assume that the significant correlation indicates an Ekman bottom layer, then there are some regions without clear Ekman bottom layers. Why should this be?

We did not address how the eddies are generated in the Indian Ocean. The JAMSTEC model clearly shows that the mesoscale features are mainly generated near the eastern boundary and then overwhelmingly propagate westward. Two eddy generation sites can be identified from the model output in the Eastern Indian Ocean:

1. the west coast of Australia, along which the Leeuwin Current flows. The current is seen in the model mean velocity. The baroclinic instability of the Leeuwin Current is assumed to be the generation mechanism. The model shows bursts of eddies during certain times of the year, presumably when the current is more unstable.
2. near the model Lambok Strait. The nearby eddies directly connect to the throughflow in the model Lambok Strait. The intrusion of shallow external water into a deep ocean will initiate the motion over whole water column by the geostrophic adjustment.

The model output would enable us to diagnose these hypothesized mechanism.

The following two topics are important and yet to be understood:

1. **Bottom Dissipation and Rossby waves** Because the bottom dissipation could be empirically expressed as a linear term, we have a quite good understanding about its role on the Rossby waves. It suggests that we need a strong bottom dissipation as in the JAMSTEC model to have significant impact on the Rossby waves. Converting the model bottom dissipation to vertical eddy

viscosity, it is as large as that near the surface. Is it realistic? In real oceans the small-scale topography could enhance the bottom dissipation (Vanneste 2000). Thus it is possible that the bottom dissipation in the JAMSTEC model actually reflects both bottom friction and form drag due to rough topography. Thus the first question is whether form drag could be as large as that in the JAMSTEC model.

The second question would be what the bottom dissipation does to the Rossby waves. In this study, the comparison between the theory and the model output is rather qualitative—could we simulate the results numerically?

To address these two equations, we could run a box Rossby wave model similar to that used in Qiu *et al.* (1997) but not reduced gravity, which is forced near the eastern boundary by the Kelvin waves originating from the equatorial region. We could test whether the Rossby waves would be modified by increasing bottom dissipation in the way predicted in this study, and whether it is true that the more dissipative the wave the faster the propagation.

2. **Non-linearity – the Jacobian Terms** As seen, we keep the theoretical study within the Rossby waves framework. We do not look at the geostrophic turbulence resulting from the advective non-linear terms (the Jacobian terms). The model output shows that the advective non-linear terms could be important, especially at time scales shorter than 100 days (figure 11.1).

   Another viewpoint considers nonlinear interactions between Rossby waves of different vertical modes. Its motivation is that the square of the first baroclinic mode is similar to that of the vertical structure seen in the observation and model output. Since the coupling of the different modes is in the first order, this interaction (Jacobian terms) has to be one of the dominant terms. This is the combination of the Rossby waves and geostrophic turbulence.

We would certainly like to learn more about the Jacobian terms using the JAMSTEC model. Then we could use the same Rossby wave model as a starting point but force it more vigorously.

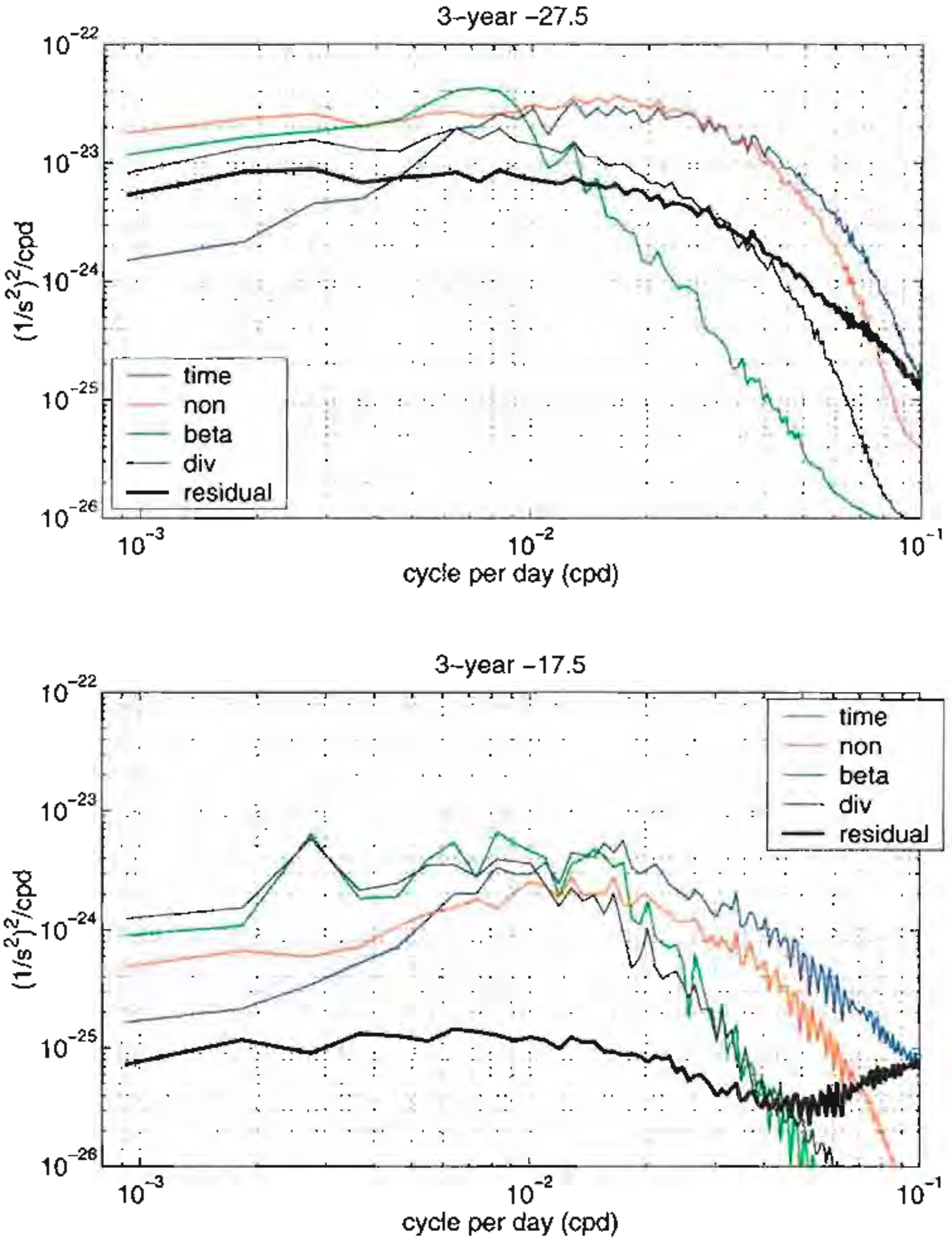


Figure 11.1: The spectra of each terms of the vorticity equation: time - temporal variation of the relative vorticity; non - 3-D advection of the relative vorticity; beta - β term; div - the stretching term, and the residual is the residual of the time after subtracting all others. The 3-year original time series at a 1/4 degree by 1/4 degree grid at 190m is separately used to calculated the Fourier coefficient, then the squared coefficients within 95-100E and 25-30S (upper panel) or 15-20S (lower panel) are averaged to produce the spectra shown. Totally, 441 time series' are used to calculate one spectrum, but the degree of freedom is unknown due to the correlation among the time series at adjacent grids.

APPENDIX A

CALCULATION OF THE LAG CORRELATION COEFFICIENTS

The steps used to calculate the lag correlation coefficients for one velocity component (either u or v) and one normal mode (either barotropic or baroclinic mode) are as follows:

1. For each of the 366 profiles, we define a element which includes the mode amplitude, geographic location (latitude and longitude) and cruise identification. We then arbitrarily order those 366 elements into one single series **P**.

2. For any distance range, e.g., from 0 km to 60 km, start two new series **A** and **B**. For each element (denoted as **a**) of the series **P** , test all elements behind it (denoted as **b**). Whenever the distance between the two stations of **a** and **b** is within that distance range (i.e. from 0 km to 60 km), we add **a** to series **A** and **b** to series **B**. Notice that series **A** and **B** have the same length and could be longer than 366.

3. Between series **A** and **B**, the station pair (one from **A** and another from **B**, and they have the same index in their corresponding series) could be from the same cruise or from different cruises—recall we have cruise identifications for each element. Two correlation coefficients are calculated: one for the same cruise (upper panel of figure 3.13) and another for different cruises (lower panel of figure 3.13).

4. repeat steps 2 and 3 for another distance range.

APPENDIX B

THE JAMSTEC MODEL

The model, a MOM2 ocean global circulation model (GCM), was implemented in the Japan Marine Science Technology Center (JAMSTEC) by Ishida *et al.* (1998). We will call the model as JAMSTEC model. The model has horizontal resolution 0.25° latitudinally and longitudinally and vertical resolution 55 levels. The 55 vertical levels are listed below, where 'mid-depth' is where temperature, salinity and horizontal velocities locate, and 'thickness' is the thickness of the z-coordinate layer, in the middle of which the temperature, salinity and horizontal velocities locate. The velocities (means and their variability over model years 20-22) at the highlighted depths in the Indian Ocean are shown in chapter 5.

level	**1**	2	3	4	5	6	7	8	9	**10**	11
mid-depth	**-5**	-15	-25	-35	-45	-55	-65	-76	-87	**-100**	-114
thickness	**10**	10	10	10	10	10	10	11	12	**13**	15

level	12	13	14	15	16	17	18	19	20	21	22
mid-depth	-130	-147	-167	-190	-215	-242	-273	-307	-344	-384	-427
thickness	17	19	21	24	26	29	32	35	38	42	45

level	23	24	25	26	27	28	29	30	31	**32**	33
mid-depth	-473	-522	-575	-630	-687	-747	-810	-874	-940	**-1007**	-1076
thickness	48	51	54	56	59	61	64	65	67	**68**	69

level	34	35	36	37	38	39	40	41	42	43	44
mid-depth	-1145	-1215	-1285	-1357	-1435	-1524	-1626	-1745	-1884	-2045	-2231
thickness	70	70	70	74	82	94	110	128	150	173	198

level	45	46	**47**	48	49	50	51	52	53	54	55
mid-depth	-2442	-2680	**-2943**	-3233	-3546	-3882	-4238	-4611	-4998	-5395	-5798
thickness	224	251	**277**	302	325	347	365	381	393	401	405

The model starts from a rest state with annually averaged temperature and salinity of the 1982 Levitus climatology. No heat and freshwater exchanges are considered but a linear restoring of the temperature and salinity in the first model level toward the 1982 Levitus climatology. The restoring time scale is such that the water properties near the surface within 50 m is restored within 30 days. It is impossible to integrate the model for a long time enough to spin up the deep ocean because a short time step is required by the fine horizontal grids. Restoring terms with a 2 years time scale toward the 1982 Levitus climatology are therefore included in the temperature and salinity equations below 2000 m. The Hellerman and Rosenstein wind stress is used to force the model. The temperature and salinity of the climatology are linearly interpolated into to the model grids. Meanwhile the wind stress is interpolated using hyperbolic cubic patches, in order to obtain smoothly varying derivatives,specially the the curl of the wind stress.

The model uses staged B grids.

The first 2 model years is the initial spin-up stage when the annually averaged climatologies (temperature, salinity and wind stress) and the harmonic horizontal dissipation are used. After that, monthly climatologies and the highly scale selective biharmonic horizontal dissipation, with coefficients -1×10^{19} $\mathrm{cm}^4\mathrm{s}^{-1}$ for momentum, temperature and salinity, are used. The vertical dissipation scheme is the Pacanowski and Philander formulation.

The model topography is derived from the NOAA National Geophysical Data Center dataset (ETOPO5) with 5' resolution latitudinally and longitudinally. The original topography from ETOPO5 is too steep for our purpose, a Gaussian spatial smoother with the radius varying continuously from 0.5° above 1000 m to 1° at 6000 m is applied. Additionally, single grid holes and spikes and small islands are objectively removed. In the MOM2 code, the topography is stepwise: the water depths are exactly those gray dots in figure B.1 and can not be between. Because the model topography is constructed in such a way, perfect flat bottom is resulted in some locations.

The output of the model years 20-22 are used in this study.

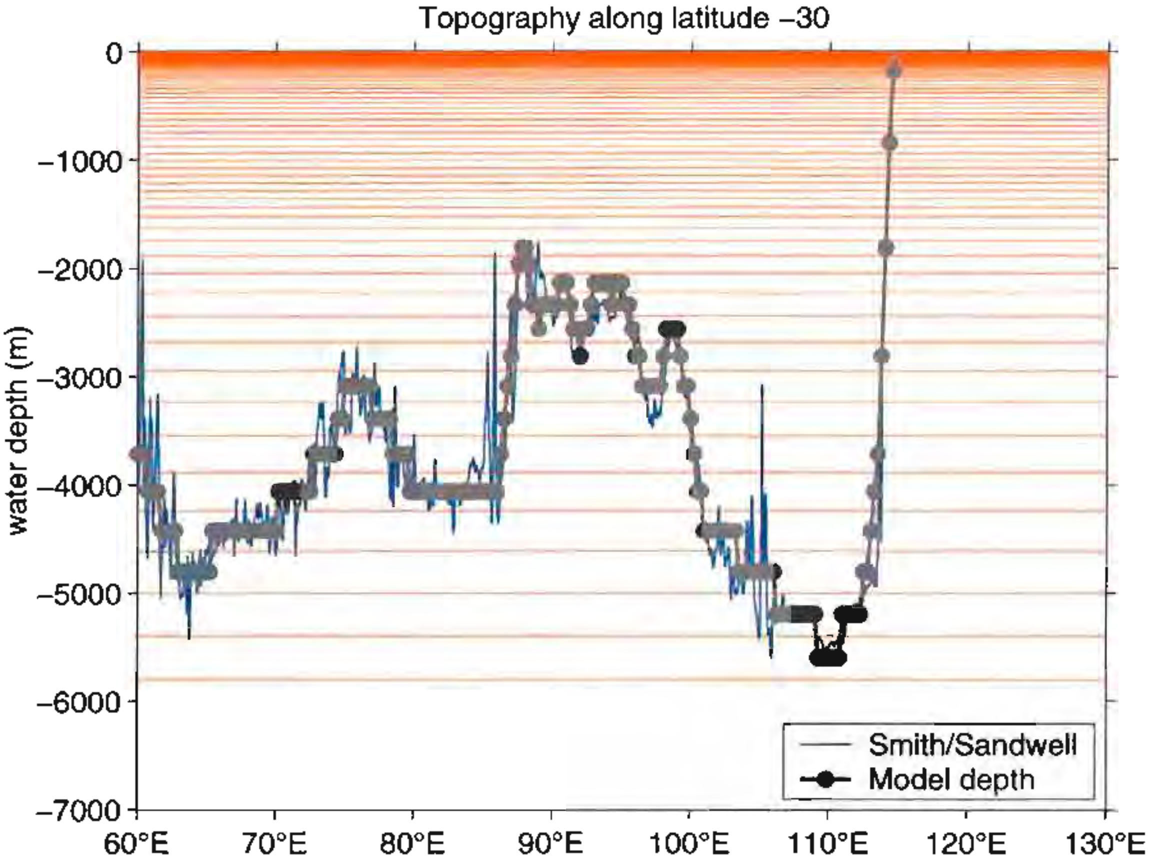


Figure B.1: Vertical grids and model topography in the Indian Ocean along 30°S. The red lines are where temperature, salinity and horizontal velocities locate. The gray dotted line is the model bottom where $w = 0$ and the bottom drag law are applied. The Smith-Sandwell topography (thin blue line) is shown for comparison.

APPENDIX C

DISPERSION OF A GAUSSIAN EDDY

We first solve for baroclinic mode ψ_c. The equation and initial condition are:

$$\begin{cases} \frac{\partial}{\partial t}[\nabla^2\psi_c - (1+\frac{1}{\delta}))\psi_c] + \frac{\partial \psi_c}{\partial x} = 0 \\ \psi_c|_{t=0} = \frac{\sqrt{\delta}}{1+\delta}\exp(-\frac{x^2+y^2}{2L^2}) \end{cases} \tag{C.1}$$

with boundary condition

$$\psi_c \to 0, \quad \text{when} \quad \sqrt{x^2+y^2} \to \infty$$

Following 2-D Fourier transforms* are introduced:

$$\begin{cases} \Psi_c(k,l;t) = \mathcal{F}(\psi_c) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\psi_c(x,y;t)e^{-\mathrm{i}(kx+ly)}dxdy \\ \psi_c(x,y;t) = \overline{\mathcal{F}}(\Psi_c) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Psi_c(k,l;t)e^{\mathrm{i}(kx+ly)}dkdl, \end{cases}$$

and applied to equation C.1:

$$\begin{cases} \frac{\partial \Psi_c}{\partial t} - \frac{ik}{k^2+l^2+1+\frac{1}{\delta}}\Psi_c = 0 \\ \Psi_c|_{t=0} = \Psi_{c0} = \frac{\sqrt{\delta}}{1+\delta}L^2\exp(-\frac{(k^2+l^2)L^2}{2}) \end{cases}$$

Its solution is

$$\begin{aligned} \Psi_c &= \Psi_c|_{t=0}e^{\frac{\mathrm{i}kt}{k^2+l^2+1+1/\delta}} \\ &= \frac{\sqrt{\delta}}{1+\delta}L^2 e^{-\frac{(k^2+l^2)L^2}{2}} e^{\frac{\mathrm{i}kt}{k^2+l^2+1+1/\delta}} \end{aligned}$$

Then $\psi_c(x,y;t)$ is

$$\begin{aligned} \psi_c &= \frac{\sqrt{\delta}}{1+\delta}L^2 \cdot \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{(k^2+l^2)L^2}{2}} e^{\frac{\mathrm{i}kt}{k^2+l^2+1+1/\delta}} e^{\mathrm{i}(kx+ly)}dkdl \\ &= \frac{\sqrt{\delta}}{1+\delta}L^2 \cdot \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{(k^2+l^2)L^2}{2}} e^{\mathrm{i}k(\frac{t}{k^2+l^2+1+1/\delta}+x)+\mathrm{i}ly}dkdl \end{aligned}$$

Defining

$$(k,l) = \kappa(\cos\theta, \sin\theta), \quad \text{and} \quad \mathcal{Z} \equiv \kappa^2 = k^2 + l^2$$

we have

$$\psi_c = \frac{\sqrt{\delta}}{1+\delta}L^2 \cdot \frac{1}{2\pi}\int_0^{\infty} d\kappa \int_0^{2\pi} \kappa d\theta e^{-\frac{\kappa^2 L^2}{2}} e^{\mathrm{i}\kappa\{(\frac{t}{\kappa^2+1+1/\delta}+x)\cos\theta + y\sin\theta\}}$$

Using the relation

$$\int_0^{2\pi} e^{\mathrm{i}(A\cos\theta + B\sin\theta)}d\theta = 2\pi J_0(\sqrt{A^2+B^2})$$

---

[*] The similar 1-D transforms were used in page 131, Pedlosky (1987)

where J_0 is the zero-order Bessel of the first kind, we have

$$\begin{aligned}\psi_c &= \tfrac{\sqrt{\delta}}{1+\delta} L^2 \int_0^\infty \kappa d\kappa e^{-\frac{\kappa^2 L^2}{2}} J_0\left(\kappa\sqrt{(\tfrac{t}{\kappa^2+1+1/\delta}+x)^2+y^2}\right) \\ &= \tfrac{\sqrt{\delta}}{1+\delta}\tfrac{L^2}{2} \int_0^\infty d\mathcal{Z} e^{-\frac{\mathcal{Z}L^2}{2}} J_0\left(\sqrt{\mathcal{Z}}\sqrt{(\tfrac{t}{\mathcal{Z}+1+1/\delta}+x)^2+y^2}\right)\end{aligned}$$

The solution for ψ_T could be obtained by substituting $\frac{1}{1+\delta}$ for $\frac{\sqrt{\delta}}{1+\delta}$ and omitting $1+1/\delta$ inside J_0:

$$\psi_T = \frac{1}{1+\delta}\frac{L^2}{2}\int_0^\infty d\mathcal{Z} e^{-\frac{\mathcal{Z}L^2}{2}} J_0\left(\sqrt{\mathcal{Z}}\sqrt{(\frac{t}{\mathcal{Z}}+x)^2+y^2}\right)$$

APPENDIX D

ONE-DIMENSIONAL SOLITARY WAVES

If the lower layer is quiet (as in a $1\frac{1}{2}$ layer model) or the influence of the lower layer is neglected (as in this study), then the downward interface η^* (a sign difference from the usual upward interface perturbation, * means dimensional quantities)

$$\eta^* = \frac{f\psi_1^*}{g'} \quad \text{and} \quad u^* \sim \frac{g'}{f}\frac{\partial \eta^*}{\partial y}$$

geostrophically. Choosing δH as the scale of η^* and R as the length scale, the velocity scale $U = g'\delta H/fR$ and the scale of ψ_1^* is $UR = g'\delta H/f$. Therefore in the non-dimensional form

$$\eta = \psi_1.$$

In this appendix, we substitute η for ψ_1. When the lower layer cropping, i.e., $\eta^* = -H_1$,

$$\eta = -\frac{H_1}{\delta H} = -\frac{\frac{g'H_1}{fR}}{\frac{g'\delta H}{fR}} = -\frac{\frac{g'H_1}{f^2R^2}fR}{U} = -\frac{fR}{U} = -\frac{1}{R_o}$$

if we chose the horizontal scale

$$R = \frac{\sqrt{g'H_1}}{f} \quad \text{and} \quad R_o = \frac{U}{fR},$$

as we have done in Section 10.2.1. R is the Rossby radius of deformation and R_o is the Rossby number. Notice all the characteristic scales are consistent with what in Section 10.2.1.

Considering one-dimensional solutions, we assume $\frac{\partial}{\partial y} = 0$. Equation 10.16 becomes:

$$\frac{\partial \eta}{\partial t} - \frac{\partial \eta}{\partial x} - R_o\eta\frac{\partial \eta}{\partial x} - \hat{S}\frac{\partial}{\partial t}\frac{\partial^2 \eta}{\partial x^2} = 0. \tag{D.1}$$

For solitary waves, we define

$$s = x - ct,$$

and

$$\frac{\partial}{\partial x} = \frac{d}{ds} \qquad \text{and} \qquad \frac{\partial}{\partial t} = -c\frac{d}{ds},$$

where c is the constant east-west propagation speed. Then equation D.1 becomes

$$(1+c)\frac{d\eta}{ds} + R_o\eta\frac{d\eta}{ds} - \hat{S}c\frac{d^3\eta}{ds^3} = 0.$$

which is, after integrated,

$$(1+c)\eta + \frac{R_o}{2}\eta^2 - \hat{S}c\frac{d^2\eta}{ds^2} = C.$$

Multiplying above equation $d\eta/ds$ and integrating once more yield

$$(1+c)\frac{\eta^2}{2} + \frac{R_o}{6}\eta^3 - \frac{\hat{S}c}{2}(\frac{d\eta}{ds})^2 = C\eta + D,$$

which is

$$(\frac{d\eta}{ds})^2 = \frac{R_o}{3\hat{S}c}[\eta^3 + \frac{3(1+c)}{R_o}\eta^2 - \frac{6C}{R_o}\eta - \frac{6D}{R_o}] \equiv \frac{R_o}{3\hat{S}c}F(\eta), \tag{D.2}$$

where $F(\eta)$ is

$$F(\eta) = \eta^3 + \frac{3(1+c)}{R_o}\eta^2 - \frac{6C}{R_o}\eta - \frac{6D}{R_o}.$$

C and D are two integration constants, relating to the initial states of perturbation, and constrain the solution through the roots of the three-order polynomial $F(\eta)$.

Solitons have one single peak and decay to zero toward the infinity. Therefore, we have

$$\eta = 0, \qquad \frac{d\eta}{ds} = 0, \qquad \frac{d^2\eta}{ds^2} = 0, \qquad \ldots \qquad (s \to \pm\infty)$$

Making use of above relations, we have

$$C = 0 \qquad \text{and} \qquad D = 0$$

in equation D.2, which becomes

$$(\frac{d\eta}{ds})^2 = \frac{R_o}{3\hat{S}c}\eta^2(\eta + \frac{3(1+c)}{R_o}). \tag{D.3}$$

The peak (local extrema) of the soliton happens where $\frac{d\eta}{ds} = 0$. When letting $\frac{d\eta}{ds} = 0$ in equation D.3, we obtains the peak itself—the non-zero root of the right side—which is

$$A = -\frac{3(1+c)}{R_o}. \tag{D.4}$$

We may call A as the amplitude of the soliton. Since R_o could be expressed as $\delta H/H_1$, where δH is the scale of the interface perturbation and H_1 is the upper layer depth, then AR_o is the amplitude of the soliton relative to the upper layer depth. Being not outcropping requires $AR_o > -1$; recall η is defined positive downward. Equation D.4 suggests that the propagation speed is

$$c = -1 - \frac{AR_o}{3} \tag{D.5}$$

Since AR_o has a minimum -1, c has a maximum $-\frac{2}{3}$. This means the solitons always propagate westward. We also notice that the propagation speed is linearly proportional to its amplitude.

Equation D.3 further constrains c so that $c < -1$, i.e., fast than all linear Rossby waves. If $-1 < c < -\frac{2}{3}$, $\eta < -\frac{3(1+c)}{R_o} < 0$ in order to make the right side of equation D.3 positive. This is contrary to $\eta = 0$ when $s \to \infty$. Moreover, equation D.5 indicate that A has to be positive since $c < -1$. This implies that the soliton has to have the interface downward and anti-cyclonic.

The soliton expression is

$$\eta = A\text{sech}^2(\sqrt{\frac{AR_o}{12\hat{S}(-c)}}\, s) = A\text{sech}^2\left(\sqrt{\frac{AR_o}{12\hat{S}(1+AR_o/3)}}\,(x-ct)\right)$$

The width of the soliton is defined as

$$L = \sqrt{\frac{6\hat{S}(-c)}{AR_o}} = \sqrt{\frac{6\hat{S}(1+AR_o/3)}{AR_o}},$$

where the soliton falls off to about $e^{-1/2} \approx 0.61$ of its peak (figure D.1). In figure D.1, we also plotted the Gaussian function with the same width. They are similar overall, but the soliton decays slightly slower than the Gaussian function.

As seen in figure D.2, the width L deceases when the soliton gets stronger. After the amplitude A beyond 3, however, the decrement gets less dramatically.

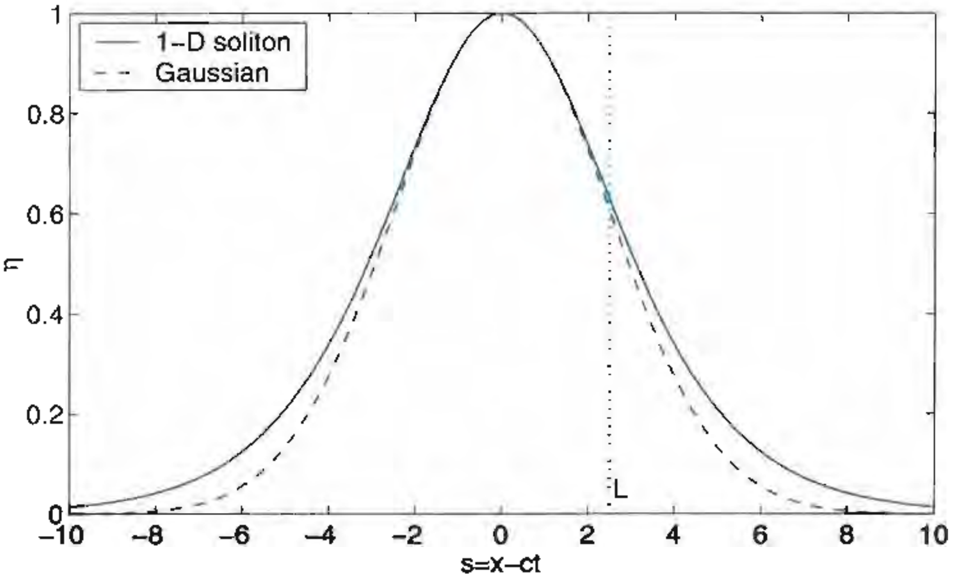


Figure D.1: The 1-D soliton with parameters: $\hat{S} = 0.1$, $R_o = 0.1$ and $A = 1$. $c = -1.03$ and $L \approx 2.5$. A 1-D Gaussian function $e^{-\frac{s^2}{2L^2}}$ is also shown.

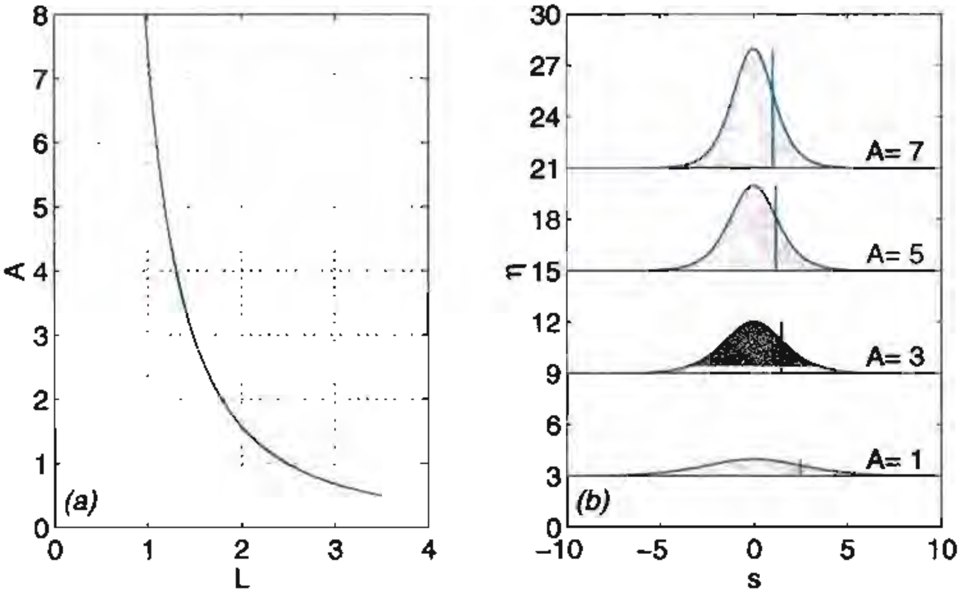


Figure D.2: (a) The width (L) of the solitons as function of their amplitudes (A). (b) The shapes of the solitons when the amplitudes A are 1, 3, 5 and 7. η is offsetted vertically. The vertical lines indicate the widths of the solitons.

APPENDIX E

FORCED ROSSBY WAVES

Applying the 2-D Fourier transform (from (s, y) to (k, l)), which has been used in appendix C, to following equation (which is in fact equation 10.18)

$$\frac{\partial}{\partial t}[\nabla^2\psi_2 - \frac{1}{\delta}\psi_2] + \frac{\partial\psi_2}{\partial s} - c\frac{\partial}{\partial s}[\nabla^2\psi_2 - \frac{1}{\delta}\psi_2] = \frac{c}{\delta}\frac{\partial\psi_1}{\partial s}$$

yields

$$\frac{\partial\Psi_2}{\partial t} - (\frac{\mathrm{i}k}{k^2+l^2+\frac{1}{\delta}} + \mathrm{i}kc)\Psi_2 = \frac{c}{\delta}\frac{-\mathrm{i}k}{k^2+l^2+\frac{1}{\delta}}\mathcal{F}(\psi_1)$$

with the initial condition $\Psi_2|_{t=0} = 0$. Its homogeneous solution is

$$\Psi_{2h} = Ae^{(\frac{\mathrm{i}k}{k^2+l^2+\frac{1}{\delta}}+\mathrm{i}kc)t}$$

where A is an integral constant to be determined, and the particular solution is

$$\Psi_{2p} = \frac{1}{\delta(k^2+l^2+\frac{1}{\delta}+\frac{1}{c})}\mathcal{F}(\psi_1)$$

Therefore the solution is

$$\begin{aligned}
\Psi_2 &= \Psi_{2p} + \Psi_{2h} \\
&= \Psi_{2p}(1 - e^{(\frac{\mathrm{i}k}{k^2+l^2+\frac{1}{\delta}}+\mathrm{i}kc)t}) \\
&= \mathcal{F}(\psi_1)\tfrac{1}{\delta(k^2+l^2+\frac{1}{\delta}+\frac{1}{c})}(1 - e^{(\frac{\mathrm{i}k}{k^2+l^2+\frac{1}{\delta}}+\mathrm{i}kc)t}) \\
&= \tfrac{L^2}{\delta(k^2+l^2+\frac{1}{\delta}+\frac{1}{c})}e^{-(k^2+l^2)L^2/2}(1 - e^{(\frac{\mathrm{i}k}{k^2+l^2+\frac{1}{\delta}}+\mathrm{i}kc)t})
\end{aligned}$$

where the initial condition $\Psi_2|_{t=0} = 0$ has been applied and

$$\mathcal{F}(\psi_1) = L^2e^{-(k^2+l^2)L^2/2},$$

since $\psi_1 = e^{-(x^2+y^2)/2L^2}$. The inverse Fourier transform of $\Psi_2(k, l; t)$ results

$$\psi_2(s, y; t) = \frac{L^2}{2\pi\delta}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{e^{-(k^2+l^2)L^2/2}}{k^2+l^2+\frac{1}{\delta}+\frac{1}{c}}(1 - e^{(\frac{\mathrm{i}k}{k^2+l^2+\frac{1}{\delta}}+\mathrm{i}kc)t})e^{\mathrm{i}(ks+ly)}dkdl$$

Denoting

$$(k, l) = \kappa(\cos\theta, \sin\theta)$$

then

$$\psi_2 = \frac{L^2}{2\pi\delta}\int_0^{\infty}\int_0^{2\pi}\frac{e^{-\kappa^2L^2/2}}{\kappa^2+\frac{1}{\delta}+\frac{1}{c}}(1 - e^{\mathrm{i}\kappa\cos\theta(\frac{1}{\kappa^2+\frac{1}{\delta}}+c)t})e^{\mathrm{i}\kappa(s\cos\theta+y\sin\theta)}\kappa d\kappa d\theta$$

Making use of

$$\int_0^{2\pi} e^{\mathrm{i}(A\cos\theta+B\sin\theta)}d\theta = 2\pi J_0(\sqrt{A^2+B^2}),$$

where J_0 is the zero-oder of Bessel function of the first kind, we have

$$\int_0^{2\pi} e^{\mathrm{i}\kappa(s\cos\theta+y\sin\theta)}d\theta = 2\pi J_0(\kappa\sqrt{s^2+y^2})$$

and

$$\begin{aligned}\int_0^{2\pi} e^{\mathrm{i}\kappa\cos\theta(\frac{1}{\kappa^2+\frac{1}{\delta}}+c)t}\cdot e^{\mathrm{i}\kappa(s\cos\theta+y\sin\theta)}d\theta &= \int_0^{2\pi} e^{\mathrm{i}\kappa([s+(\frac{1}{\kappa^2+\frac{1}{\delta}}+c)t]\cos\theta+y\sin\theta)}d\theta \\ &= 2\pi J_0(\kappa\sqrt{[s+(\tfrac{1}{\kappa^2+\frac{1}{\delta}}+c)t]^2+y^2})\end{aligned}$$

Therefore

$$\begin{aligned}\psi_2 &= \frac{d^2}{\delta}\int_0^\infty \frac{e^{-\kappa^2d^2/2}}{\kappa^2+\frac{1}{\delta}+\frac{1}{c}}\left(J_0(\kappa\sqrt{s^2+y^2}) - J_0(\kappa\sqrt{[s+(\tfrac{1}{\kappa^2+\frac{1}{\delta}}+c)t]^2+y^2})\right)\kappa d\kappa \\ &= \frac{d^2}{2\delta}\int_0^\infty \frac{e^{-\frac{\mathcal{Z}d^2}{2}}}{\mathcal{Z}+\frac{1}{\delta}+\frac{1}{c}}\left(J_0(\sqrt{\mathcal{Z}}\sqrt{s^2+y^2}) - J_0(\sqrt{\mathcal{Z}}\sqrt{[s+(\tfrac{1}{\mathcal{Z}+\frac{1}{\delta}}+c)t]^2+y^2})\right)d\mathcal{Z}\end{aligned} \tag{E.1}$$

where $\mathcal{Z} \equiv \kappa^2 = k^2 + l^2$. When $\mathcal{Z} + \frac{1}{\delta} + \frac{1}{c} \to 0$ (this apparent singularity will happen because c is about -1),

$$\frac{1}{\mathcal{Z}+\frac{1}{\delta}} + c = \frac{c}{\mathcal{Z}+\frac{1}{\delta}}(\mathcal{Z}+\frac{1}{\delta}+\frac{1}{c}) \to 0$$

so that

$$J_0(\sqrt{\mathcal{Z}}\sqrt{s^2+y^2}) - J_0(\sqrt{\mathcal{Z}}\sqrt{[s+(\frac{1}{\mathcal{Z}+\frac{1}{\delta}}+c)t]^2+y^2}) \to 0$$

This means that the singularity at $\mathcal{Z} + \frac{1}{\delta} + \frac{1}{c} = 0$ could be eliminated. When numerically integrating equation E.1, we may simply skip its small neighbor.

We first show ψ_2, as well as ψ_1, along x-axis in figure E.1, E.2 and E.3. We keep $\delta = 4$ but vary L from 1 to 2 and 3. From these figures we see ψ_2 under the upper layer soliton, which is moving westward at the speed -1, reaches the quasi-steady state in a short time. Most of the variation of ψ_2 associates the wake waves behind the upper layer soliton. Figures figure E.4, E.5 and E.6 show the two-dimensional views of ψ_2 at sequential times. Pay attention to the relative locations of the upper layer soliton center (the white crosses), and the ψ_2 minima and local positive extrema west of them.

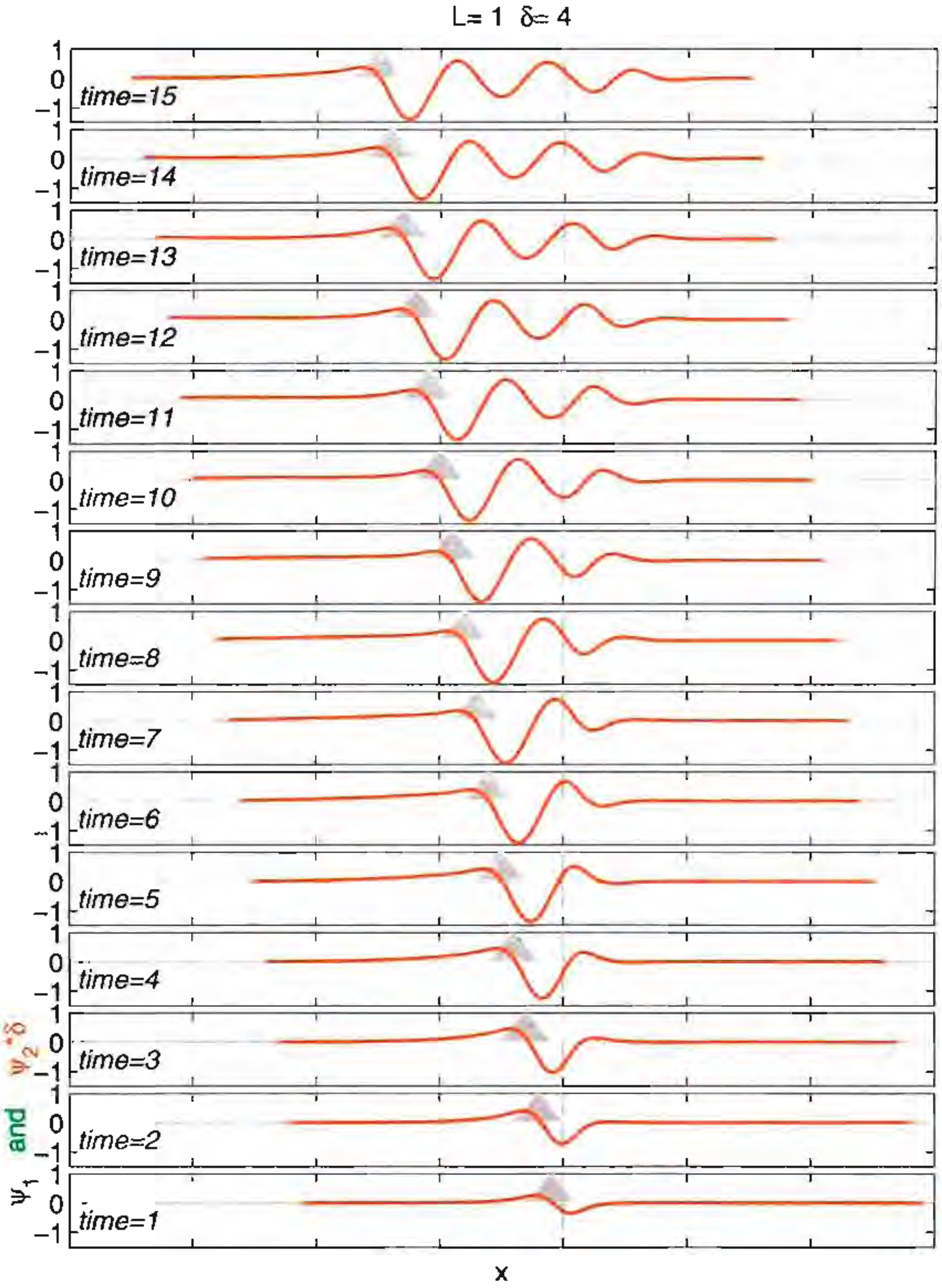


Figure E.1: The upper layer (ψ_1, gray shadings) and lower layer ($\psi_2\delta$, red lines) streamfunctions along x-axis (y=0) at different times. The upper layer Gaussian eddy propagates westward at speed c=-1. $L = 1$ and $\delta = 4$. Notice ψ_2 is scaled by δ.

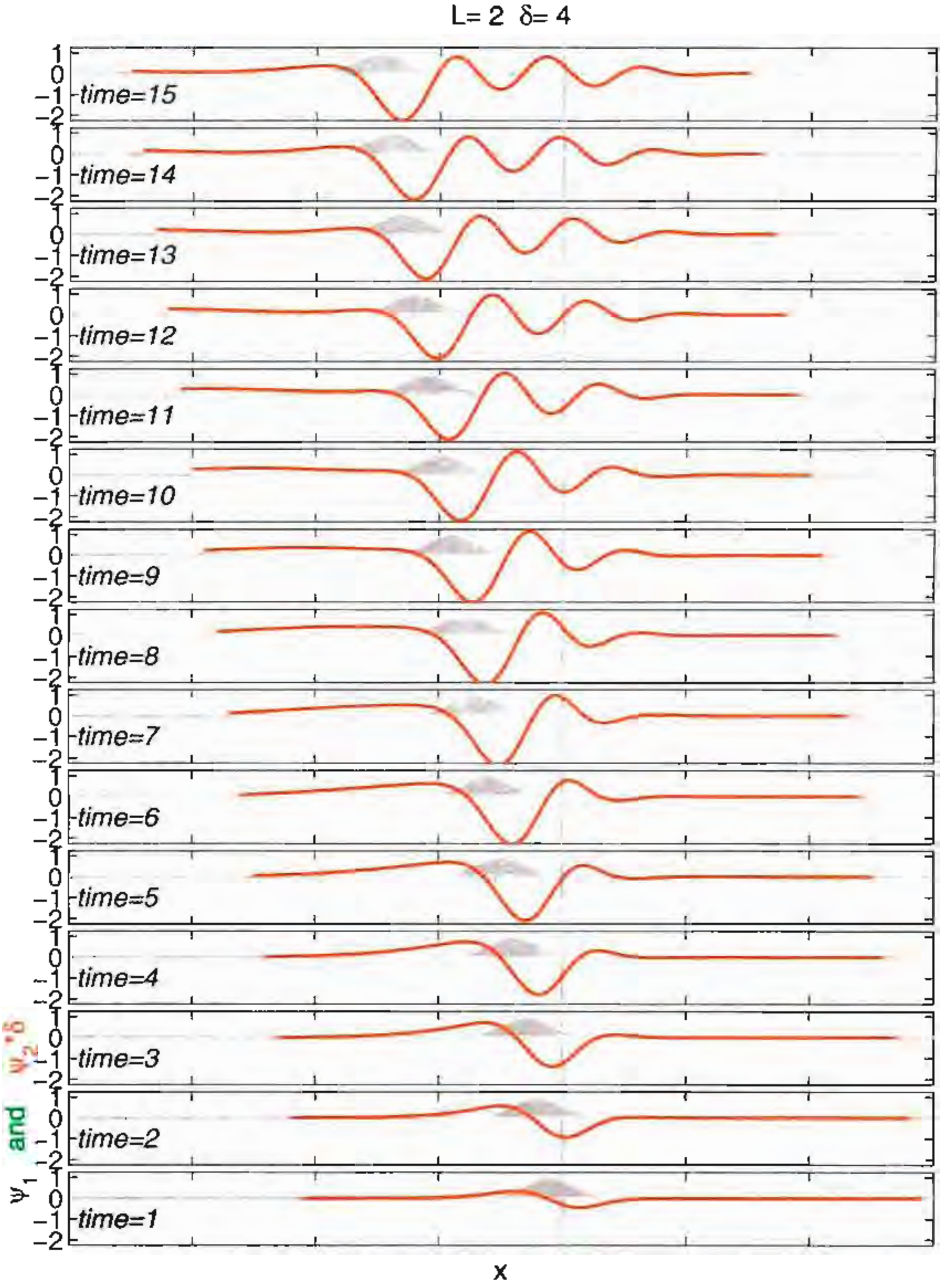


Figure E.2: Same as in figure E.1 but L=2

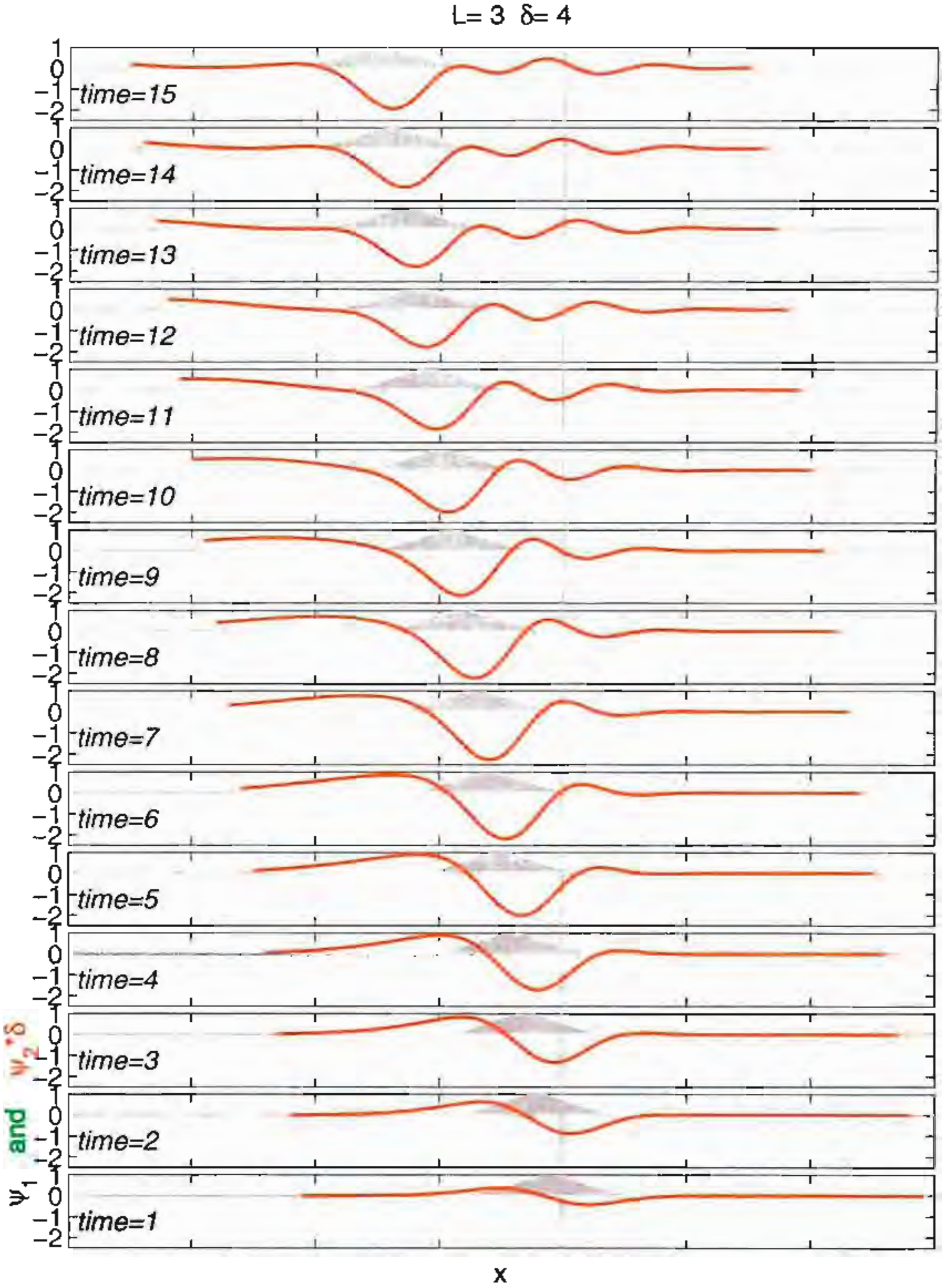


Figure E.3: Same as in figure E.1 but L=3

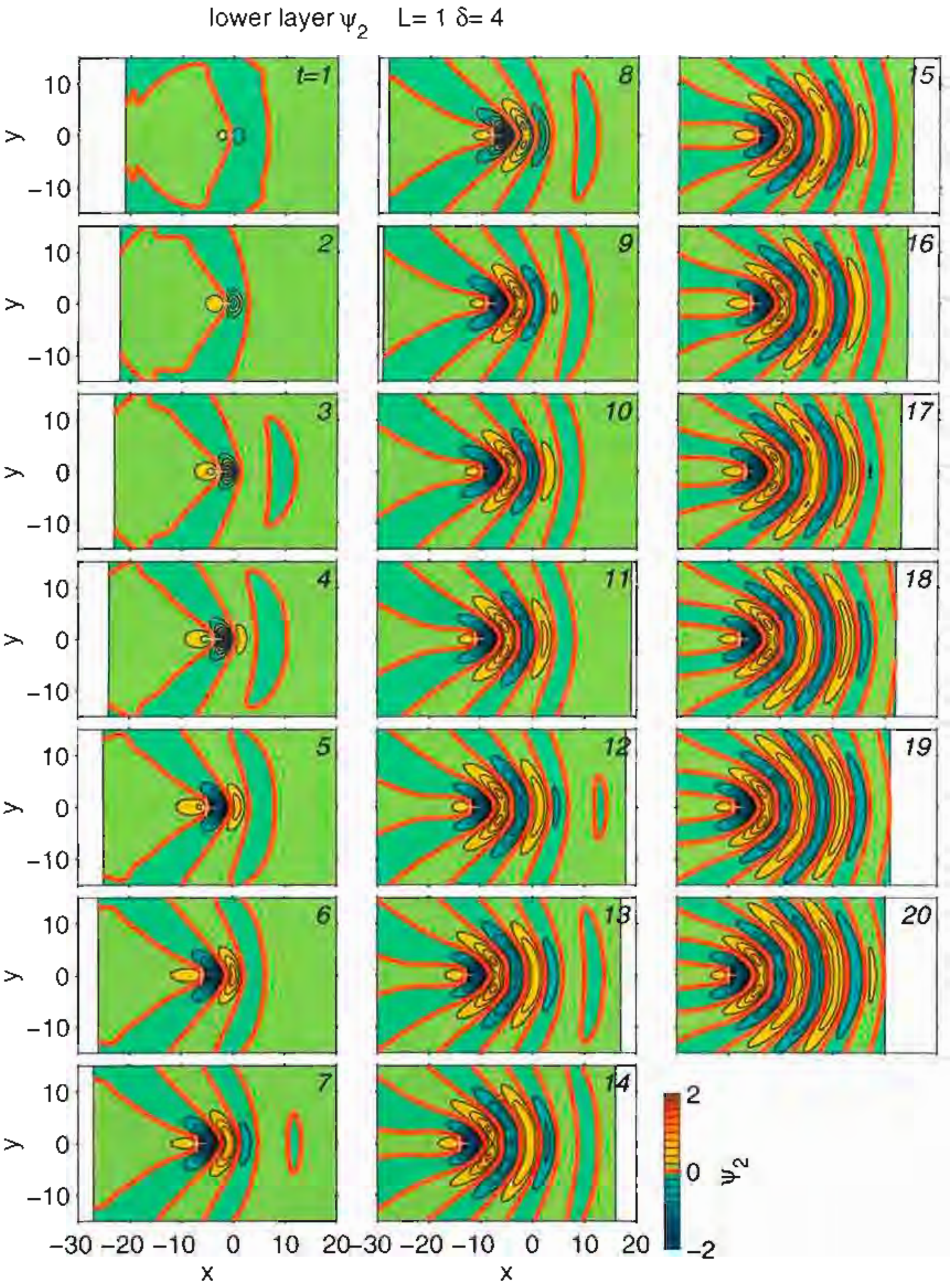


Figure E.4: Two-dimensional lower layer streamfunctions at sequential times. The time is indicated in the right corner of each panel. Location of the upper layer Gaussian eddy, which is assumed propagating westward at speed c=-1, is marked by the white cross. $L = 1$ and $\delta = 4$. Notice ψ_2 is scaled by δ. All panels use the same color scale.

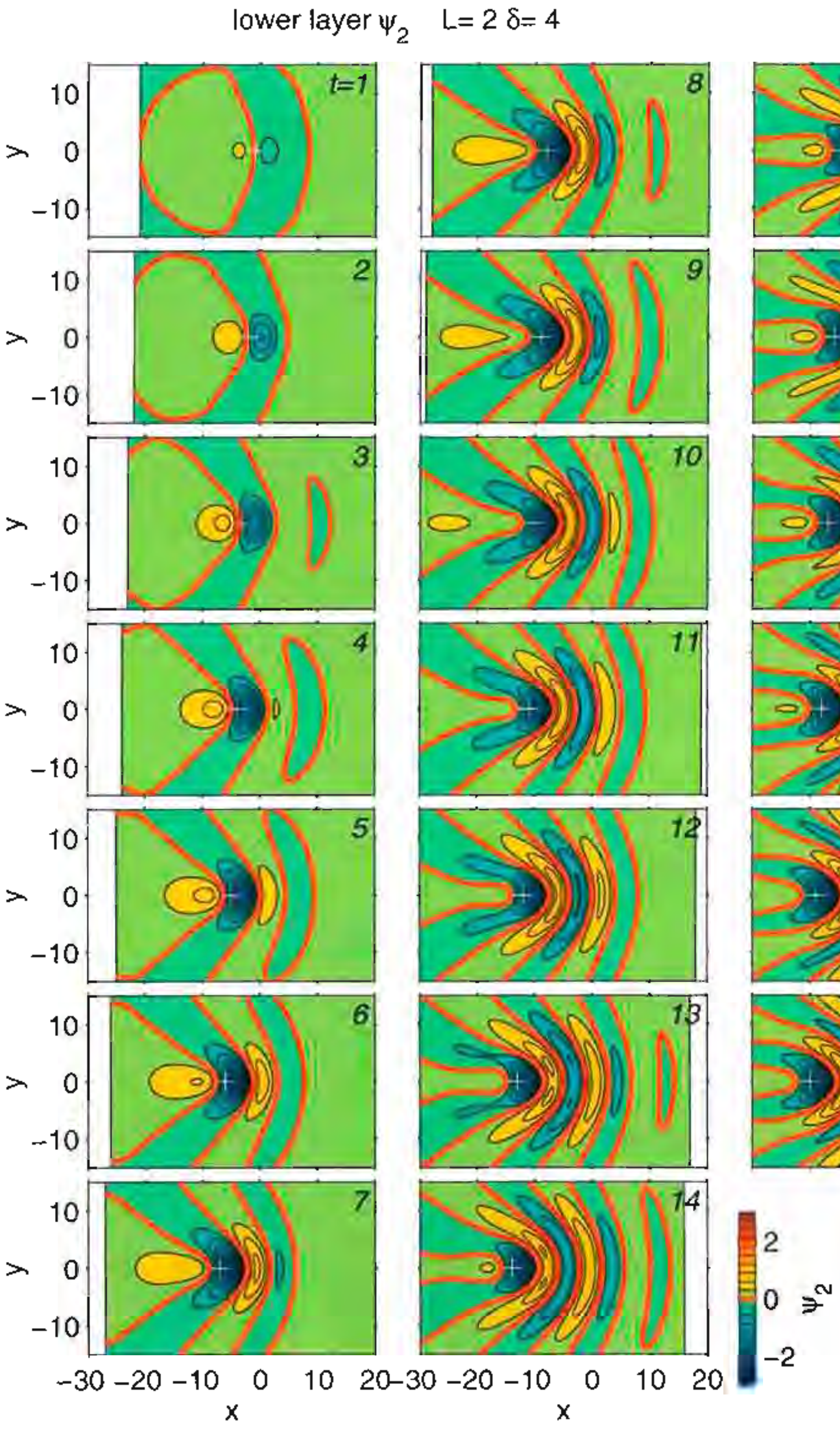


Figure E.5: Same as in figure E.4 but L=2

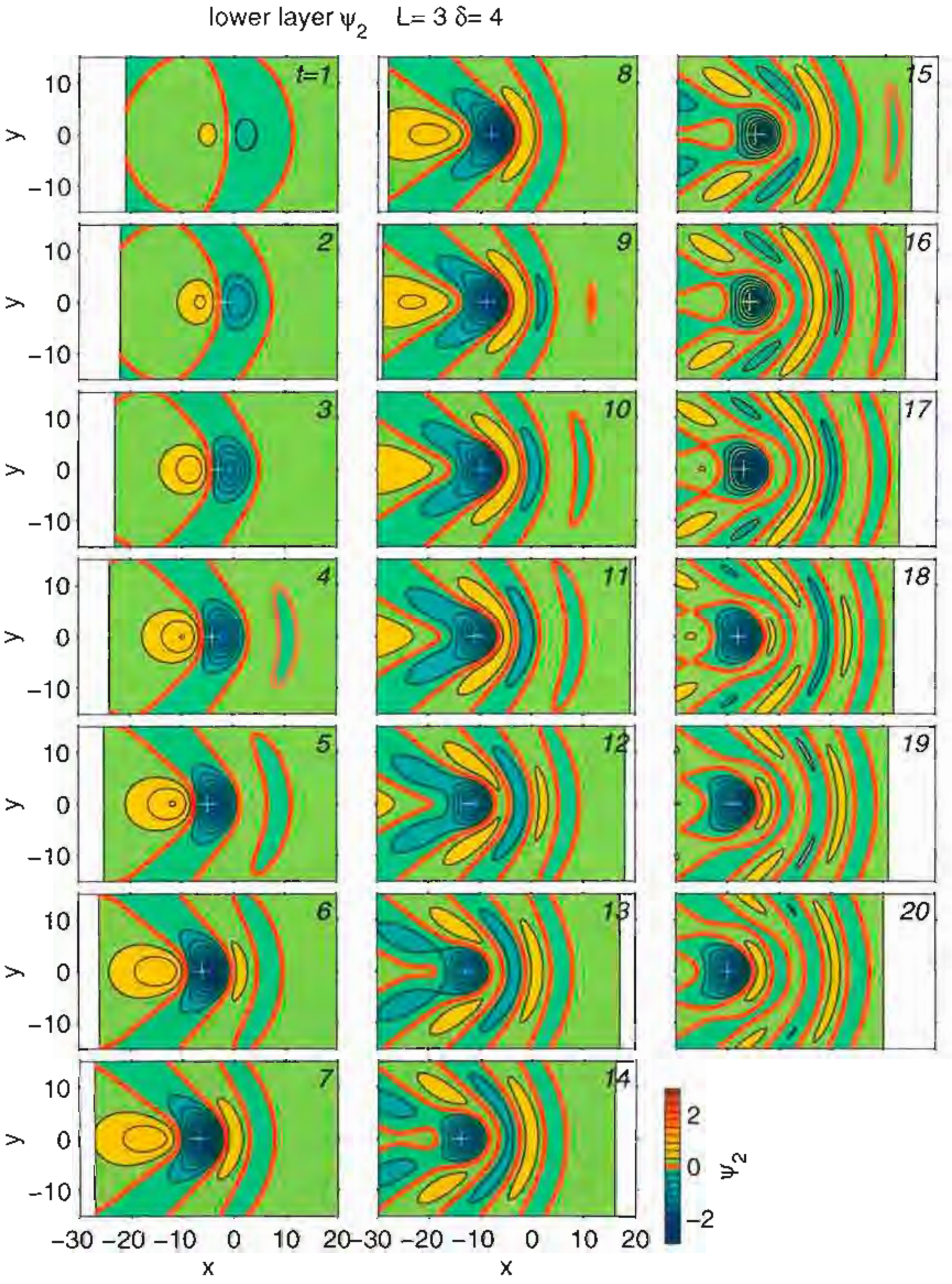


Figure E.6: Same as in figure E.4 but L=3

APPENDIX F

BOTTOM DISSIPATION IN THE JAMSTEC MODEL

The JAMSTEC model has 55 levels vertically, but its resolution near the bottom of a 4000-5000 m water column is 300-400 m (see appendix B). As a reference, thickness of surface Ekman layers is usually about 50 m, thus the bottom Ekman layer is unlikely resolved in the model. Meanwhile, the bottom friction is presented in the momentum equation of the deepest layer, in the same manner as the wind stress in the momentum equation of the most upper layer. The bottom friction is parameterized in the model as:

$$\tau_b = \rho_0 c_D \vec{u}|\vec{u}| \tag{F.1}$$

where c_D is the drag coefficient (0.0012) and $\vec{u}$ is the velocity 150-200 m above the bottom boundary (specifically in the middle of the last layer). There is not much attention to the bottom friction in GCMs so far. On may argue that the bottom friction can not be be important because the velocity near the bottom is usually small. However, the situation may be in the opposite way: because of the bottom friction, the velocity near the bottom is small.

We will not diagnose the role of the bottom friction in the momentum balance. The bottom Ekman layer is unlikely resolved, so the spiral-type velocity vertical shear will not appear in the model output. Instead, we will diagnose its role in the vorticity balance by analyzing the correlation between the vertical velocity and relative vorticity in the deepest layer.

If

- the lower layer of the ocean is homogeneous and the deepest layer is in or inside of the homogeneous layer;
- the total flow in the deepest layer is the supposition of depth-independent flow and a bottom Ekman flow;
- the vertical velocity in the middle of the deepest layer is the Ekman pumping of the bottom Ekman layer and the relative vorticity in the middle of the deepest layer represents the interior of the homogeneous layer,

then vertical velocity (w_E) and relative vorticity (ξ_N) in the middle of the deepest layer are possibly be correlated by the bottom Ekman layer theory (see Pedlosky 1987) as follows:

$$w_E = sign(f)\sqrt{\frac{A_v}{2|f|}}\xi_N \equiv sign(f)\mu\xi_N. \tag{F.2}$$

where A_v is the constant vertical eddy viscosity and f the Coriolis parameter. $sign(f)\mu$ is the regression coefficient if w_E and ξ_N are two random variables.

Equation F.2 can also be interpreted as the following. The linear vorticity equation with vertical stress is

$$\frac{\partial \xi}{\partial t} - f\frac{\partial w}{\partial z} = \frac{\partial \vec{k} \cdot (\nabla \times \vec{\tau})}{\partial z}, \tag{F.3}$$

where the temporal term is included for the completeness. Assuming

$$\begin{array}{ll} w = 0,\ \vec{k} \cdot (\nabla \times \vec{\tau}) = \Gamma \vec{k} \cdot (\nabla \times \vec{u_N}) = \Gamma \xi_N & \text{at the bottom} \\ w = w_E,\ \vec{k} \cdot (\nabla \times \vec{\tau}) = 0 & \text{outside the boundary layer} \end{array}$$

The Rayleigh bottom friction, instead of the quadratic drag law (equation F.1), is assumed. Integrating equation F.3 from the bottom until outside of the boundary layer, and neglecting the temporal term, we have

$$f w_E = \Gamma \xi_N$$

which is equivalent to equation F.2.

The Ekman layer theory with a constant eddy viscosity is used as a framework (equation F.2). Doing so has two advantages: 1) the estimate of μ can be readily applied to the QG dynamics (section 10.3); 2) after estimating μ, we could further estimate A_v using equation F.2, whose magnitude will indicate the strength of the bottom friction in the JAMSTEC model.

For each grid point with flat model topography (figure F.1), we will first estimate w_E and ξ_N for each time, then calculate correlation and regression coefficients between the three model years serieses of w_E and ξ_N.

Above flat model topography, the vertical velocities will be less affected by the horizontal velocities crossing the topographic slope. The vertical velocity is the depth change of a particular neutral surface (Jackett and McDougall 1997) at the T/S grid point. The detailed steps are: first to find out the densest neutral surface which always appears over the three model years; second to estimate the depth of that particular neutral surface at each time; then to estimate the vertical velocity using the local time variation—no advection effect is included. The relative vorticity is calculated from surrounding 4 u/v grid of the deepest layer.

The correlation coefficient between w_E and ξ_N is shown in figure F.2. We find a negative tendency. This is consistent with the bottom Ekman layer theory (equation F.2); notice f is negative in the south hemisphere. Meanwhile, we see non-negligible part of the region has positive correlation. Given that the positive correlation is generally small (see the histogram in figure F.2), we should conclude that the correlation in the region is insignificant. However, we do not have any explanation for this insignificance.

We then estimate μ for the grid with flat bottom topography, using equation F.2. They varies significantly from one grid point to another, but we have a stable distribution if they are averaged longitudinally (figure F.3, upper panel). It is quite interesting to see the magnitude of μ tends to follow $\frac{1}{\sqrt{|f|}}$, as suggested by the Ekman layer theory. Notice $\mu = \sqrt{\frac{A_v}{2|f|}}$ and the thickness of the Ekman boundary layer is $\delta_E = \sqrt{2A_v}/|f|$, so that

$$\delta_E = 2\mu.$$

Therefore, figure F.3 (upper panel) suggests that the thickness of the Ekman boundary layer is about 20 m. This is comparable to the upper wind-driven Ekman layer. Whether this is realistic needs further investigation. Further more, we infer A_v from the longitudinally averaged μ. The magnitude of A_v is 100 $\mathrm{cm^2s^{-1}}$ (figure F.3, lower panel). This is the upper limit of the Philander and Pacanowski mixing scheme used in the JAMSTEC model. In the interior away from the surface and bottom layers, A_v is usually only a few $\mathrm{cm^2s^{-1}}$.

As seen, we interpret the correlation between w_E and ξ_N under the framework of the bottom Ekman layer. This may not be the best, but the results strong suggest that the bottom friction is playing a role in the vorticity balance of presumably mesoscale motions. This prompts us to understand how the bottom friction affects Rossby waves—this has never been studied before.

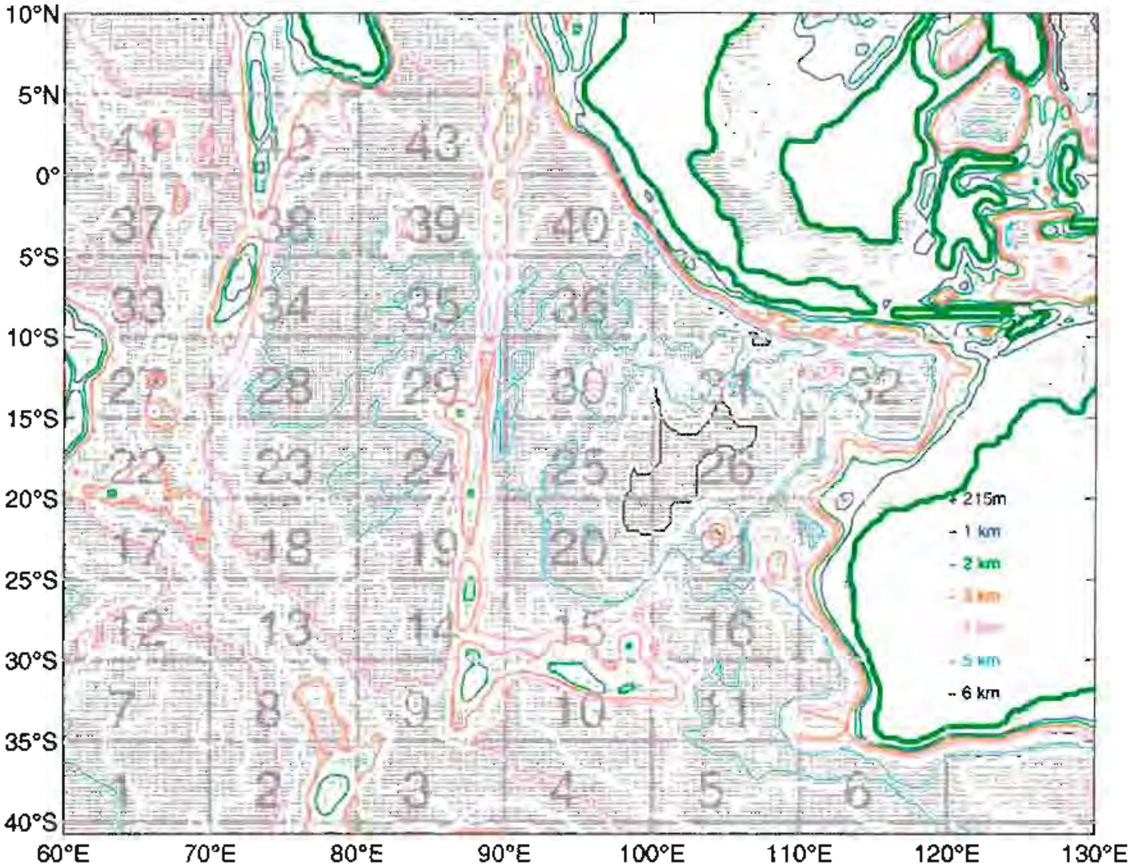


Figure F.1: The JAMSTEC model topography in the Indian Ocean. The flat bottom (dotted) is the T/S grid point where its 4 surrounding u/v grids have the same depths.

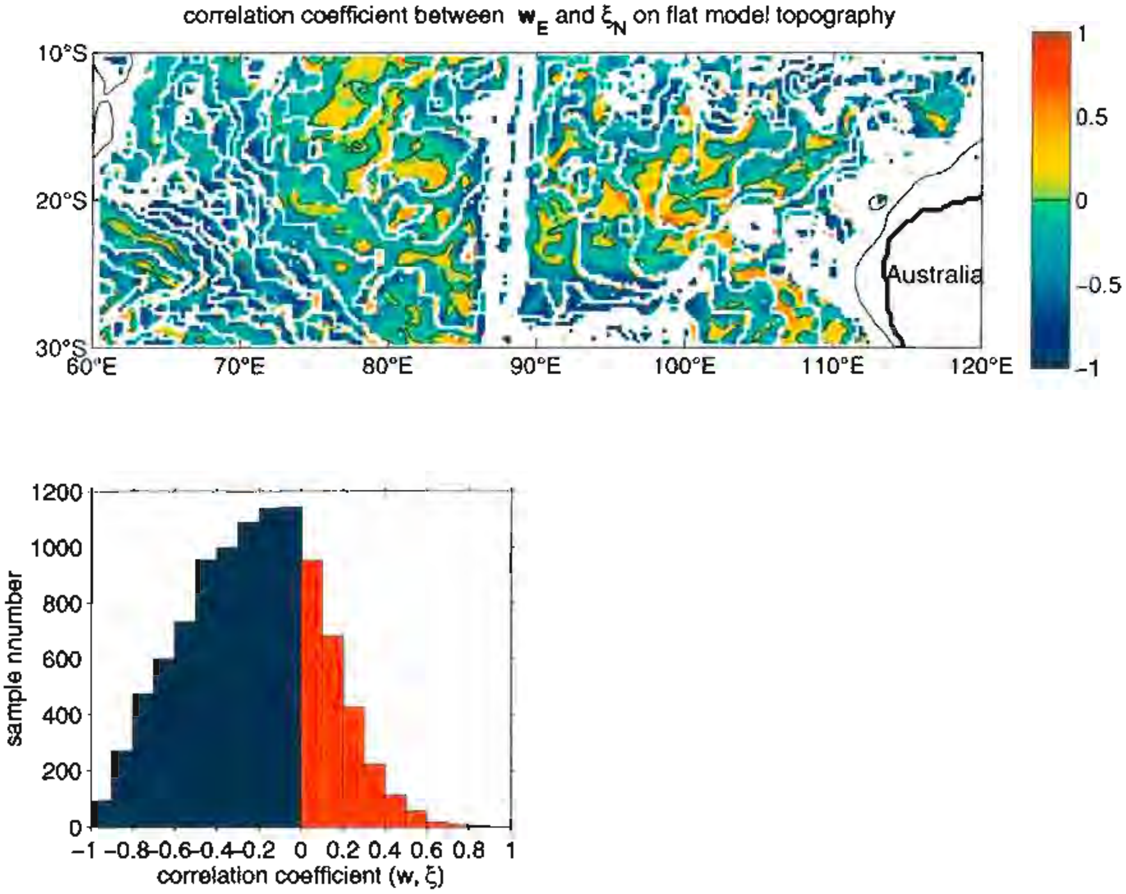


Figure F.2: Upper panel: correlation coefficient between vertical velocity and relative vorticity over three years at flat model topography. Lower panel: histogram of the correlation coefficients in the upper panel. The black contours are zeros. The blank in the upper panel indicates sloping topography.

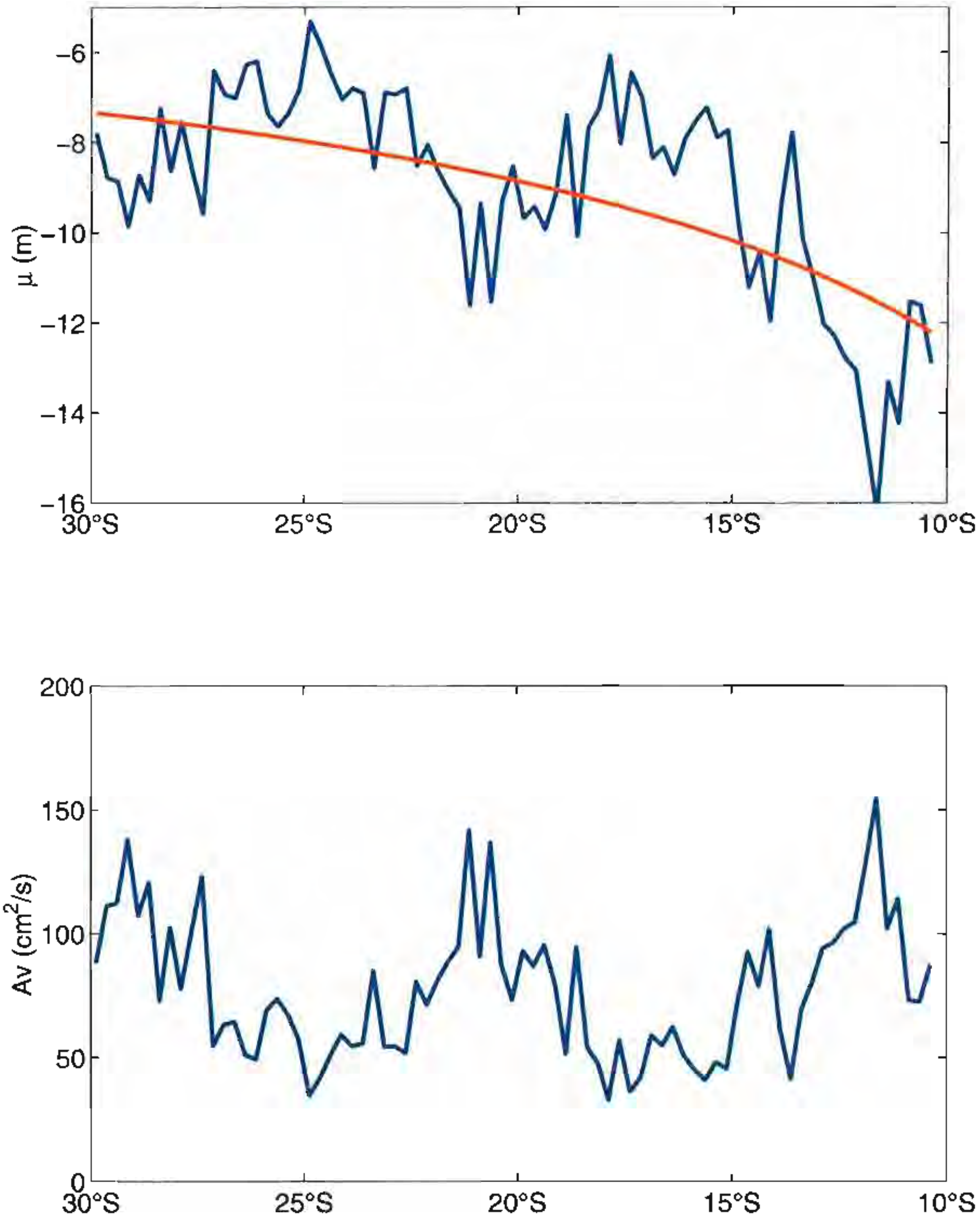


Figure F.3: Upper panel: μ is the regression coefficient between w and ξ. At each dot between 30°S and 10°S (figure F.1), there is one estimate of μ. This panel shows the longitudinally averaged μ. The red line is the best fit by $\frac{1}{\sqrt{f}}$. Lower panel: assuming the bottom Ekman layer theory with constant A_v applicable, we inverted A_v from μ in the upper panel using $\mu = \sqrt{\frac{A_v}{2|f|}}$.

APPENDIX G

HORIZONTAL DISSIPATION

The non-dimensional two-layer model is (see equation 9.1 in chapter 9)

$$\left\{\begin{array}{lcl} \frac{\partial}{\partial t}[\nabla^2\psi_1 + (\psi_2 - \psi_1)] + \frac{\partial \psi_1}{\partial x} & = & A_H\nabla^2(\nabla^2\psi_1) \\ \frac{\partial}{\partial t}[\nabla^2\psi_2 + \frac{1}{\delta}(\psi_1 - \psi_2)] + \frac{\partial \psi_2}{\partial x} & = & A_H\nabla^2(\nabla^2\psi_2) \end{array}\right. \tag{G.1}$$

with addition of the non-dimensional biharmonic terms:

$$A_H = A_H^* \frac{T}{R^2} = \frac{A_H^*}{\beta R^3}$$

Substituting following wave solutions

$$\left(\begin{array}{c} \psi_1 \\ \psi_2 \end{array}\right) = \left(\begin{array}{c} \psi_1' \\ \psi_2' \end{array}\right) e^{\mathrm{i}(kx+ly-\omega t)}$$

into equation G.1, we have

$$\left[\begin{array}{ll} \omega(k^2+l^2+1)+k+\mathrm{i}\mathcal{A} & -\omega \\ -\frac{\omega}{\delta} & \omega(k^2+l^2+\frac{1}{\delta})+k+\mathrm{i}\mathcal{A} \end{array}\right] \left(\begin{array}{c} \psi_1' \\ \psi_2' \end{array}\right) = 0 \tag{G.2}$$

where $\mathcal{A} = A_H(k^2+l^2)^2$.

To have nontrivial solutions, the determination of the coefficient matrix in equation G.2 is zero. This lead to

$$\begin{array}{lcl} \omega_1 & = & -\frac{k+\mathrm{i}\mathcal{A}}{k^2+l^2} \\ \omega_2 & = & -\frac{k+\mathrm{i}\mathcal{A}}{k^2+l^2+1+\frac{1}{\delta}}. \end{array} \tag{G.3}$$

$\omega_{1,2}$ certainly reduce to the dispersion relationships of the barotropic and baroclinic modes respectively if $\mathcal{A} = 0$.

The vertical structure (ψ_1'/ψ_2') can be obtained by the first equation equation G.2

$$(k^2+l^2+1+\frac{k+\mathrm{i}\mathcal{A}}{\omega_{1,2}})\psi_1' - \psi_2' = 0.$$

As seen in equation G.3, $\frac{k+\mathrm{i}\mathcal{A}}{\omega_{1,2}}$ is $-(k^2+l^2)$ or $-(k^2+l^2+1+1/\delta)$, thus

$$\frac{\psi_1'}{\psi_2'} = 1 \text{ and } \frac{\psi_1'}{\psi_2'} = -\delta$$

which are in fact the vertical structure of the barotropic and baroclinic modes without the horizontal dissipation. In other words, the horizontal dissipation will not modify the vertical structure of normal modes. On the other hand, we will expect that the waves decay because $\omega_{1,2}$ have imaginary parts, or that the waves are localized near source regions because the wavenumbers have imaginary parts. One example of the latter was given numerically and analytically by Qiu *et al.* (1997), who showed the Rossby waves decay away from the eastern boundary where the Rossby waves are initiated by the Kelvin waves from the equatorial region.

References

Arbic, Brian K. and G. R. Flierl, 2002: *Baroclinically unstable geostrophic turbulence in the limits strong and weak friction: application to mid-ocean eddies.* J. Phys. Oceanogr., submitted.

Baringer, M. O. and J. F. Price, 1997: Momentum and energy balance of the Mediterranean outflow. *J. Phys. Oceanogr.,* **27,** 1678-1692.

Charney J. G. and G. R. Flierl, 1982: Oceanic analogues of large-scale atmospheric motions. *Evolution of Physical Oceanography, Warren and Wunsch (eds.),* **Chapter 18,** 504-548.

Chassignet E. and B. Cushman-Roisin, 1991: On the influence of a lower layer on the propagation of nonlinear oceanic eddies. *J. Phys. Oceanogr.,* **21,** 939-957.

Chen, Shuiming, 2000: *Note on Killworth's model about equivalent barotropic mode.* Unpublished, 2pp.

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

Chelton, D. B. and M. G. Schlax, 1996: Global observations of oceanic Rossby waves. *Science,* **272,** 234-238.

Clarke, R. A., 1971: Solitary and cnoidal planetary waves. *Geophysical Fluid Dynamics,* **2,** 343-354.

Cushman-Roisin, B., 1986: Frontal geostrophic dynamics. *J. Phys. Oceanogr.,* **16,** 132-143.

Cushman-Roisin, B, E. P. Chassignet and B. Tang, 1990: Westward motion of mesoscale eddies. *J. Phys. Oceanogr.,* **20,** 758-768.

Cushman-Roisin, B. and B. Tang, 1990: Geostrophic turbulence and emergence of eddies beyond the radius of deformation. *J. Phys. Oceanogr.,* **20,** 97-113.

Cushman-Roisin, B. E. P. Chassignet and B. Tang, 1990: Westward motion of mesoscale eddies. *J. Phys. Oceanogr.,* **20,** 758-768.

Cushman-Roisin, B. and B. Tang, 1990: Geostrophic turbulence and emergence of eddies beyond the radius of deformation. *J. Phys. Oceanogr.,* **20,** 97-113.

Cushman-Roisin, B., G. G. Sutyrin and B. Tang, 1992: Two-layer geostrophic dynamics. Part I: governing equations. *J. Phys. Oceanogr.,* **22,** 117-127.

Davis, R.E., 1975: Statistical methods in *Dynamics and the analysis of MODE-I.* Rep. of the MODE-I dynamics group, MIT, Cambridge, 250pp.

Dewar, W. K. and C. Gailliard, 1994: The dynamics of barotropic dominated rings. *J. Phys. Oceanogr.,* **24,** 5-29.

Dewar, W. K., 1998: On "too fast" baroclinic planetary waves in the general circulation. *J. Phys. Oceanogr.,* **28,** 1739-2649.

Donohue, K.A., E. Firing, and S. Chen, 2001: Absolute Geostrophic Velocity within the Subantarctic Front in the Pacific Ocean. *J. Geophys. Res.,* **106,** 19,869-19,882.

Egbert, G. D., A. F. Bennett and M. G. G. Foreman, 1994: Topex/Poseidon tides estimated using a global inverse model. *J. Geophys. Res.,* **99,** 24,821-24,852.

Firing, E. and R. C. Beardsley, 1976: The behavior of a barotropic eddy on a β-plane. *J. Phys. Oceanogr.,* **6,** 57-65.

Firing, E., 1998: Lowered ADCP development and use in WOCE. *International WOCE Newsletter,* **30,** 10-14.

Firing, E. and R. L. Gordon, 1990: *Deep ocean Doppler current profiling.* Proc. fourth IEEE working conference on current measurements, Clinton, MD, Current Measurement Technology Committee of the Oceanic Engineering society, 192-201.

Fisher, J. and M. Visbeck, 1993: Deep velocity profiling with self-contained ADCPs. *J. Atmos. Oceanic Technol.,* **10,** 764-773.

Fleirl, G. R., 1979: Baroclinic solitary waves with radial symmetry. *Dyn. Atmos. Oceans,* **3,** 15-38.

Fleirl, G. R., 1984: Rossby wave radiation from a strongly nonlinear warm eddy. *J. Phys. Oceanogr.,* **14,** 47-58.

Fleirl, G. R., 1987: Isolated eddy models in geophysics. *Ann. Rev. Fluid Mech.,* **19,** 493-530.

Fu, L.-L., 2003: Wind-forced intraseasonal sea level variability of the extratropical oceans. *J. Phys. Oceanogr.,* **33,** 436-449.

Fukumori, I. and C. Wunsch, 1991: Efficient representation of the North Atlantic hydrographic and chemical distributions. *Prog. Oceanogr.,* **27,** 111-195.

Gill, A. E., J. S. A. Green and A. J. Simmons, 1974: Energy partition in the large-scale ocean circulation and the production of mid-ocean eddies. *Deep-Sea Res.,* **21,** 499-528.

Hacker, P., E. Firing, W. D. Wilson and R. Molinari, 1996: Direct observations of the current structure east of the Bahamas. *Geophys. Res. Ltrs.,* **23,** 1127-1130.

Ishida, A., Y. Kashino, H. Mitsudera, N. Yoshioka, and T. Kadokura, 1998: Preliminary results of a global high-resolution GCM experiment. *the Journal of the Faculty of Science, Hokkaido University Series VII (geoph ysics),* **11,** 441-460.

Jackett, D. R. and T. J. McDougall, 1997: A neutral density variable for the world's oceans. *J. Phys. Oceanogr.,* **27,** 237-263.

Johnson, G. C. and T. B. Sanford, 1992: Secondary circulation in the Faroe Bank Channel outflow. *J. Phys. Oceanogr.,* **22,** 927-933.

Killworth, P. D., 1992: An equivalent-barotropic mode in the Fine Resolution Antarctic Model. *J. Phys. Oceanogr.,* **22,** 1379-1387.

Killworth, P. D., D. B. Chelton and R. A. DeSzoeke, 1997: The speed of observed and theoretical long extratropical planetary waves. *J. Phys. Oceanogr.,* **27,** 1946-1966.

King, B. A., E. Firing and T. M. Joyce, 2001: Shipboard observations during WOCE. In: *Ocean Circulation and Climate – Observing and Modelling the Global Ocean. G. Siedler, J. Church and J. Gould (eds), Academic Press. The international geophysics series,* **77,** 715pp.

Kundu, Pijush K., 1990: *Fluid Mechanics.* Academic Press, 638pp.

LeBlond, P. H. and L. A. Mysak, 1978: *Waves in the Ocean.* Elsevier Scientific Publishing, 602pp.

Liu, Zhengyu, 1999: Planetary wave modes in the thermocline: non-Doppler-shift mode, advective mode and Green mode. *Q. J. R. Meteorol. Soc,* **125,** 1315-1339.

Longuet-Higgins, M. S., 1965: The response of a stratified ocean to stationary or moving wind-systems. *Deep-Sea Res.,* **12,** 923-973.

Luther, D., 1982: Evidence of a 4-6 day barotropic, planetary oscillation of the Pacific Ocean. *J. Phys. Oceanogr.,* **12,** 644-657.

Matsuura T. and T. Yamagata, 1982: On the evolution of nonlinear planetary eddies larger than the radius of deformation. *J. Phys. Oceanogr.,* **22,** 440-456.

McWilliam, J. C., P. R. Gent and N. C. Norton, 1986: The evolution of balanced, low-mode vortices on the β-plane. *J. Phys. Oceanogr.,* **16,** 838-855.

Mofjeld, H., 1981: An analytic theory on how friction affects free internal waves in the equatorial waveguide. *J. Phys. Oceanogr.,* **11,** 1585-1590.

Muller, T. and G. Siedler, 1992: Multi-year current time series in the eastern North Atlantic Ocean. *J. Marine Res.,* **50,** 63-98.

Nof, Doron, 1982: On the β-induced movement of isolated baroclinic eddies. *J. Phys. Oceanogr.,* **11,** 1662-1672.

North, G. R., T. L. Bell, R. F. Cahalan and F. J. Moeng, 1982: Sampling errors in the estimation of empirical orthogonal functions. *J. Atmos. Sci.,* **110,** 699-706.

Pacanowski, R. C., 1995: *MOM 2 documentation: user's guide and reference manual.* GFDL Ocean Technical Report #3, 232pp.

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

Preisendorfer, Rudolph W., 1988: *Principal component analysis in meteorology and oceanography.* Elsevier, 425pp.

Rossby, T., 1974: Studies of the vertical structure of horizontal currents near Bermuda. *J. Geophys. Res.,* **79,** 1781-1791.

Qiu, B., Weifeng Miao and P. Muller, 1997: Propagation and decay of forced and free baroclinic Rossby waves in off-equatorial oceans. *J. Phys. Oceanogr.,* **27,** 2405-2417.

Redekopp L. G., 1977: On the theory of solitary Rossby waves. *Journal of Fluid Mechanics,* **82,** 725-745.

Rhines P., 1977: The dynamics of unsteady currents. *The Sea,* **6,** 189-318.

Rossby, T., 1974: Studies of the vertical structure of horizontal currents near Bermuda. *J. Geophys. Res.,* **79,** 1781-1791.

Saunders, P. M., 2001: The dense northern overflows. In: *Ocean Circulation and Climate – Observing and Modelling the Global Ocean. G. Siedler, J. Church and J. Gould (eds), Academic Press. The international geophysics series,* **77,** 715pp.

Stammer, Detlef, R. Tokmakian, A. Semtner and C. Wunsch, 1996: How well does a 1/4^o global circulation model simulate large-scale oceanic observations?. *J. Geophys. Res.,* **101,** 25,779-25,811.

Stammer Detlef, 1997: Global characteristics of ocean variablity estimated from regional Topex/Poseidon altimeter measurements. *J. Phys. Oceanogr.,* **27,** 1743-1769.

Smith, W. H. F. and D. T. Sandwell, 1994: Bathymetric prediction from dense altimetry and sparse shipboard bathymetry. *J. Geophys. Res.,* **99,** 21,803-21,824.

Tomczak, M. and J. S. godfrey, 1994: *Regional oceanography: an introduction.* Pergamon, 422pp.

Vanneste Jacques, 2000: Enhanced dissipation for quasi-geostrophic motion over small-scale topography. *J. Fluid Mech.,* **407,** 105-122.

Wang, B., 1990: On the asymmetry of baroclinic instability between easterly and westerly shear. *Tellus,* **42A,** 463-468.

Wunsch, C., 1997: The Vertical Partition of Oceanic Horizontal Kinetic Energy. *J. Phys. Oceanogr.,* **27,** 1770-1794.

Wunsch, C., Uncertainly in oceanic general circulation models: *in Monte Carlo simulation in oceanography by Muller, Peter and Diane Henderson (eds).* Proceedings, 'Aha Huliko'a, Hawaiian Winter Workshop, Jan 1997, Honolulu, Hawaii, 206pp.

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

Zang X. and C. Wunsch, 1999: The observed dispersion relationship for North Pacific Rossby wave motions. *J. Phys. Oceanogr.,* **29,** 2303-2317.