

Velocity Probability Density Functions from Altimetry

SARAH T. GILLE*

School of Environmental Sciences, University of East Anglia, Norwich, United Kingdom

STEFAN G. LLEWELLYN SMITH[†]

Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, United Kingdom

(Manuscript received 2 July 1998, in final form 23 March 1999)

ABSTRACT

Probability density functions (pdfs) are employed to evaluate the distribution of velocities in the global ocean. This study computes pdfs of ocean surface velocity using altimetric data from the TOPEX/Poseidon satellite. Results show that the shape of the observed pdfs changes with the size of the domain over which they are calculated: if data are drawn from a small region of the ocean, the pdfs are Gaussian. As the area of the ocean considered increases, the pdfs take on more exponential shapes. The appearance of exponential pdfs is particularly pronounced when data are drawn from a large range of latitudes, while data drawn from constant latitude tend to have a more Gaussian pdf. The authors show that this distinction between zonal and meridional regions is also observed in acoustic Doppler current profiler measurements.

The authors propose a simple statistical model to explain the observed velocity pdfs. This explanation depends on the fact that root-mean-squared velocity (or the width of velocity pdf) varies throughout the ocean. The velocity pdf is predicted from the distribution of the mean-squared velocity. The model matches the observations in predicting a pdf that is parabolic for small velocities with generalized exponential decay for large velocities.

1. Introduction

Turbulent motions in the ocean are responsible for mixing chemical tracers, potential vorticity, and other quantities. Probability density functions (pdfs) are one of the basic statistical tools that can be used to describe these turbulent motions (e.g., Frisch 1995). In this paper, we look at pdfs of surface velocity measurements from the TOPEX/Poseidon satellite altimeter.

Because motions in the ocean are subject to the constraints of rapid rotation and strong stratification, ocean mesoscale variability resembles two-dimensional turbulence or its generalizations such as beta-plane turbulence and geostrophic turbulence (Rhines 1979; Lesieur 1997; Salmon 1998). These flows are characterized by coherent vortices, separated by near-irrotational re-

gions dominated by strong straining motions (McWilliams 1984). In the ocean, as in the atmosphere, these structures are associated with substantial mixing along two-dimensional isentropic or isopycnal surfaces (Samelson 1996; Haynes 1999).

In an early numerical study of two-dimensional turbulence, Lilly (1969) remarked that the components of velocity appeared to be non-Gaussian. More recently Jiménez (1996) and Min et al. (1996) independently calculated the theoretical velocity pdfs for ensembles of point vortices and of nonsingular vortices with identical strength. Both predicted Gaussian pdfs for infinite domains. Jiménez also found Gaussian distributions when he analyzed simulations of two-dimensional turbulence. One reason the two-dimensional turbulence results may resemble the point vortex results is because the simulations of two-dimensional turbulence exhibited coherent vortices, all of roughly the same extent.

From the work of Jiménez (1996) and Min et al. (1996), we might expect the ocean itself to have Gaussian velocity pdfs. However, there are a number of ways in which the ocean differs from the simulations of two-dimensional turbulence considered by Jiménez. In contrast with the ocean, the numerical simulations took place in a flat-bottomed, doubly periodic domain, with no stratification and no topography. (Jiménez noted that finite-size effects due to the doubly periodic domain can

* Current affiliation: Department of Earth System Science, University of California, Irvine, Irvine, California.

[†] Current affiliation: Department of Mechanical and Aerospace Engineering, University of California, San Diego, La Jolla, California.

Corresponding author address: Dr. Sarah Gille, Department of Earth System Science, University of California, Irvine, Irvine, CA 92697-3100.
E-mail: sgille@uci.edu

have a significant impact on predicted pdfs.) In addition, in two-dimensional turbulence, the Rossby radius is infinite and the eddies have a limited range of sizes, while the ocean contains a wide distribution of eddy sizes. Moreover, two-dimensional turbulence calculations are usually spindown simulations, and do not have time-dependent and spatially inhomogeneous forcing. The differences between two-dimensional turbulence simulations and the ocean imply that oceanic velocity pdfs may differ from the predictions of point vortex theory.

Recent investigations with Lagrangian data from floats (Bracco et al. 2000) and from numerical models (Weiss et al. 1998) have found evidence for exponential tails on velocity pdfs. In this paper, we will show that in regions of uniform eddy kinetic energy the pdfs of geostrophic surface velocity are roughly Gaussian. However, in regions with varying levels of eddy kinetic energy, pdfs tend to be more exponential, in agreement with the float observations of Bracco et al. (2000). We will suggest that oceanic velocity pdfs are non-Gaussian as a result of the inhomogeneity of eddy kinetic energy, and we will provide a statistical framework for predicting velocity pdfs.

Investigating velocity pdfs on a global scale would be difficult with measurements from floats or current meters because of the paucity of in situ data compared with the enormous number of realizations required to compute reliable pdfs. In contrast, the TOPEX satellite altimeter provides on the order of 150 000 usable estimates of geostrophic velocity over the global ocean every 10 days. Previous global analyses of the TOPEX measurements have focused on the spectral characteristics of the measurements, eddy kinetic energy, length scales of variability, and steric height changes due to seasonal heating and cooling (Stammer and Wunsch 1994; Wunsch and Stammer 1995; Stammer 1997; McClean et al. 1997). Llewellyn Smith and Gille (1998) describe altimetric velocity pdfs and velocity gradient pdfs in the context of turbulence theory, but to our knowledge, this study represents the first detailed discussion of altimetric pdfs in an oceanographic framework.

We begin in section 2 with a brief description of pdfs and their relation to eddy kinetic energy and spectra. Section 3 discusses the results from TOPEX altimeter data, which show that pdfs calculated over small regions of the ocean may often be Gaussian (or nearly Gaussian) but tend to be more exponential when data from larger regions of the ocean are considered. In section 4 we explain how the exponential velocity distributions that we find in altimeter data can be predicted statistically, based on the variations in background eddy kinetic energy. The results are summarized in section 5. The appendix contains asymptotic calculations relevant to section 4.

2. Probability density functions of velocity

Consider a time series of a component of velocity, v , at a point in space. In a turbulent flow, this signal will

be random. Formally for a pdf $p(v)$, the integral $\int_u^{u+\delta u} p(v) dv$ represents the probability that a sampled velocity should take a value between u and $u + \delta u$. To estimate the pdf of the observed velocity samples, we compute a histogram, normalize it by the number of samples, and divide by the velocity bin width.

From the pdf $p(v)$ for a set of velocity observations, we can calculate the moments μ_n about the mean μ . These are

$$\mu_n = \int_{-\infty}^{\infty} (v - \mu)^n p(v) dv, \quad (1)$$

where μ_1 is zero since we have subtracted out the sample mean μ . The second moment, μ_2 , represents the variance or spread of the distribution. The next higher moments give the skewness ($\mu_3/\mu_2^{3/2}$) and flatness (μ_4/μ_2^2) of the distribution (Priestley 1981).

The variance of the velocity field, μ_2 , is of particular interest. The velocity variance is related to eddy kinetic energy. We shall define the eddy kinetic energy (EKE) over a region with area A to be

$$\text{EKE} = \frac{1}{2A} \int |\mathbf{v}(\mathbf{x})|^2 d^2x. \quad (2)$$

If the velocity field is isotropic, then we may consider only one component, v , and (2) may be recast in probabilistic language as

$$\text{EKE} = \int_{-\infty}^{\infty} v^2 p(v) dv = \mu_2, \quad (3)$$

where $p(v)$ is the pdf of velocity in the region of interest. Since the variance is related to the spread of the pdf, flows with low variances have low eddy kinetic energy and narrow pdfs.

In this paper we will consider both the Gaussian distribution, $p_1(v)$:

$$p_1(v) = \frac{1}{\sqrt{2\pi}s} \exp\left(-\frac{v^2}{2s^2}\right), \quad (4)$$

and the symmetric exponential distribution $p_2(v)$:

$$p_2(v) = \frac{1}{\sqrt{2}s} \exp\left(-\frac{|v|\sqrt{2}}{s}\right). \quad (5)$$

Both (4) and (5) are written so that they have variance s^2 . Thus, the EKE of measurements drawn from either distribution will be identical, and the distributions will be distinguishable only when we examine higher-order moments.

More than just having the same EKE, data drawn from different pdfs can also have nearly identical spectra. Figure 1 shows spectra and pdfs for data selected randomly from a uniform distribution, a Gaussian (or normal) distribution, and an exponential distribution, all with the same mean and variance. The exponentially and normally distributed random data are generated by

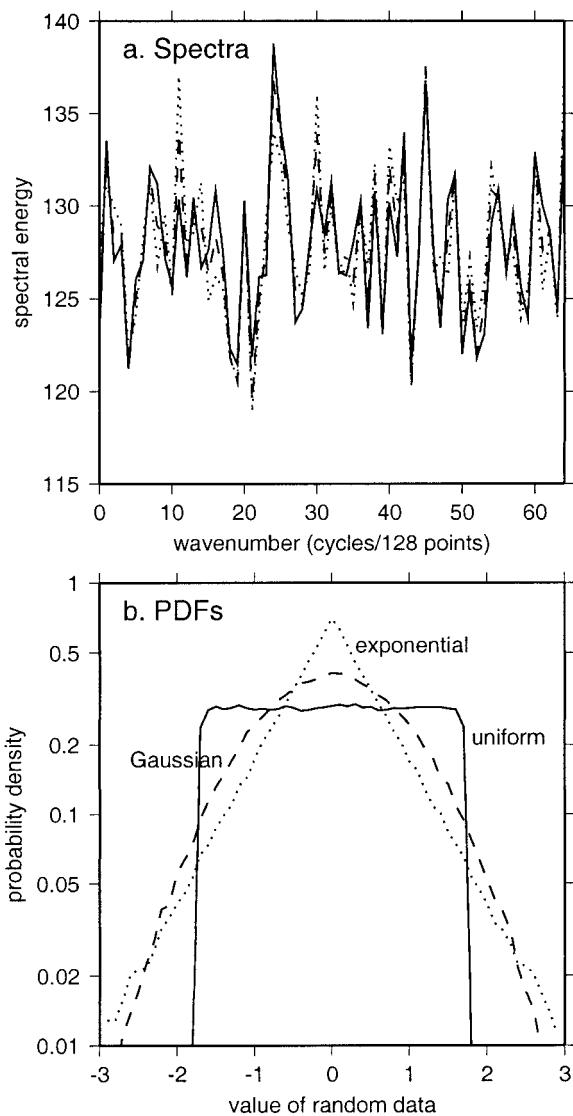


FIG. 1. (a) Mean spectra for 1000 realizations of 128-point random data series. The solid line is the spectrum for random data that have a uniform distribution between $\pm\sqrt{3}$. Data from the uniform distribution are then transformed to have Gaussian (dashed line) and exponential (dotted line) distributions with the same standard deviation. Even though these lines indicate means of 1000 spectra, they match closely because the peaks and troughs in the initial uniform distribution occur at the same points in the Gaussian and exponential distributions. (b) Pdfs for the 128 000 points used to make the spectra show that the data distributions are distinctly different, although the standard deviations are the same.

transforming the uniformly distributed noise. Therefore the highs and lows in the data series coincide, and the datasets are more than 90% correlated. As Fig. 1a indicates, the resulting spectra are nearly indistinguishable, although their pdfs in Fig. 1b are distinctly different. In semilogarithmic coordinates, the Gaussian pdf is parabolic, while the exponential pdf is triangular. Pdfs of the form $\exp(-|v|^\alpha)$ with $0 < \alpha < 2$ are non-Gaussian and will be termed “generalized exponential” distri-

butions. Thus compared with a Gaussian pdf, an exponential pdf has longer “tails,” containing more extreme values. From the example of Fig. 1, we can see that the information in pdfs is complementary to the spatial and temporal structures derived from spectral analysis. Pdfs do not provide information about time-scales or lengthscales, but they do indicate the probability of extreme events.

Based on spectral analysis of altimeter data from the global ocean, Stammer (1997) noted that velocity wavenumber spectra have nearly the same slope everywhere and that differences in eddy kinetic energy are the major factor distinguishing different parts of the ocean. In the next section, we show that pdfs, like wavenumber spectra, have a nearly universal shape: within small regions of the ocean, velocity pdfs are almost always Gaussian, but the width of these pdfs varies with EKE.

3. Velocity distributions from the TOPEX altimeter

The TOPEX/Poseidon satellite has provided measurements of ocean surface height since September 1992. The satellite carries two radar altimeters. Since the Poseidon altimeter has higher noise levels, here we employ only the measurements from the TOPEX altimeter, which operates about 90% of the time. One TOPEX/Poseidon cycle of 9.9 days has 127 ascending and 127 descending ground tracks, giving a between-track spacing of about 300 km at the equator. The alongtrack footprints are about 6.5 km in diameter. For details about TOPEX/Poseidon, see for example Fu et al. (1994).

In this analysis we consider only the time-varying component of sea surface height. Surface velocity anomalies with respect to the mean circulation can be computed based on the geostrophic relationship $v = -(g/f)\partial\eta/\partial l$, where v is the cross-track velocity anomaly, g is gravity, f is the Coriolis parameter, η is the sea surface height anomaly relative to time-mean sea surface height, and l is the distance along a satellite ground track.

To compute sea surface slopes, the first 171 cycles of TOPEX data are used with the JGM-3 orbits. We correct the data for solid earth tides, oceanic tides (using the CSR 3.0 tidal model), the pole tide, wet and dry tropospheric delays, and EM bias (Benada 1997). We do not apply the ionospheric correction because, even spatially smoothed, it increases the variance of the geostrophic velocities on the length scales we are examining. Following the techniques developed by Yale et al. (1995) and Sandwell and Smith (1997) for geophysical analysis, point-to-point sea surface slopes are computed for alongtrack altimeter measurements. A 101-point Parks–McClellan filter is applied to the slopes to remove noise with wavelengths shorter than about 40 km. Then the filtered slopes are demeaned and converted to geostrophic velocities. Ascending and descending ground tracks are oriented differently, and the altimeter

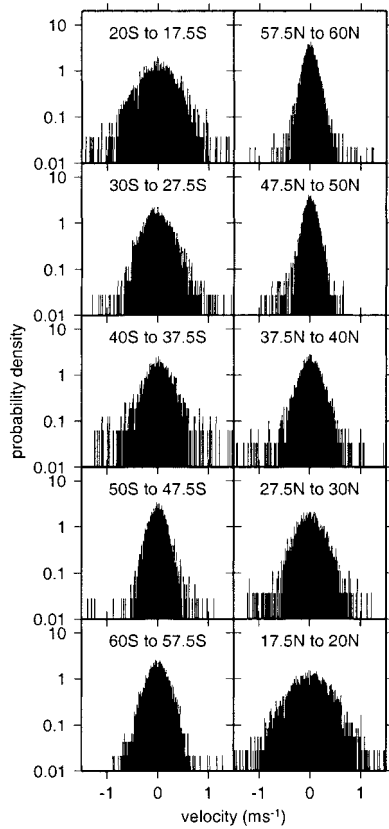


FIG. 2. Probability density distributions of TOPEX descending track velocities in selected 2.5° by 2.5° boxes along the date line, between 177.5°E and 180° (along the P14 WOCE cruise track).

measures only cross-track velocities. Therefore, to allow for possible discrepancies between different velocity directions, ascending and descending groundtracks are treated separately.

In this study, nearly five years of geostrophic velocities are sorted into 2.5° longitude by 2.5° latitude boxes, and pdfs are computed for each of these boxes. On average there are 2800 ascending and 2800 descending velocities in each 2.5° box. By averaging both in space and time, we implicitly assume that the turbulence is statistically steady and homogeneous inside each box. Data within 10° of the equator are discarded since the geostrophic relationship breaks down as f approaches

zero, and data more than 60° from the equator are omitted because of frequent ice cover at high latitudes.

Figure 2 shows the pdfs for selected 2.5° boxes located along the date line (between 177.5°E and 180°). All of the pdfs are centered around zero velocity, because the mean velocities are removed when the altimeter data are processed. Compared with the sample pdfs in Fig. 1, we see that the pdfs in Fig. 2 are more Gaussian than exponential for the most part since the distributions appear more or less parabolic. However, the tails of the pdfs show evidence for numerous data outliers, which is not surprising given the noise inherent in altimeter measurements. Like EKE (Stammer 1997), the widths of these pdfs are substantially greater at low latitudes than at high latitudes, indicating that the pdfs are not identical throughout the global ocean. In contrast, Fig. 3 shows pdfs spaced at 20-degree intervals across the South Pacific. All of these pdfs have nearly the same width and a shape characteristic of Gaussian pdfs. The differences between Figs. 2 and 3 suggest that in the middle of the ocean basins, velocity pdfs vary more in the meridional direction than in the zonal direction.

If we merge data from boxes along either a zonal line or a meridional line, we find that the resulting pdfs are significantly different. Figure 4 shows the pdfs obtained by combining all data from descending tracks in boxes located along the date line or all data distributed zonally between 32.5° and 30°S . The pdf for the meridional box drops more slowly to zero than does the pdf for the zonal box, indicating that the meridional distribution of velocities is more exponential, while the zonal distribution is more Gaussian.

For comparison, Fig. 4 also shows the pdfs for zonal velocities measured by acoustic Doppler current profiler (ADCP) during hydrographic cruises carried out as part of the World Ocean Circulation Experiment along the same zonal and meridional lines as the satellite data shown (WOCE Data Products Committee 1998). Velocities shown here are from 140-m depth, chosen to be well below the upper ocean Ekman layer, but the results would be the same at any other depth. The ADCP results are noisy because many fewer data points are available: 575 points on the meridional line and 869 points on the zonal line, compared with 130 000 and 78 000 for the meridional and zonal TOPEX pdfs respectively. In ad-

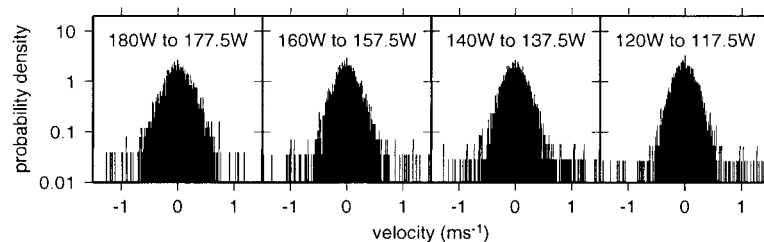


FIG. 3. Probability density distributions of TOPEX descending track velocities in selected 2.5° by 2.5° boxes between 32.5°S and 30°S (along the P6 WOCE cruise track).

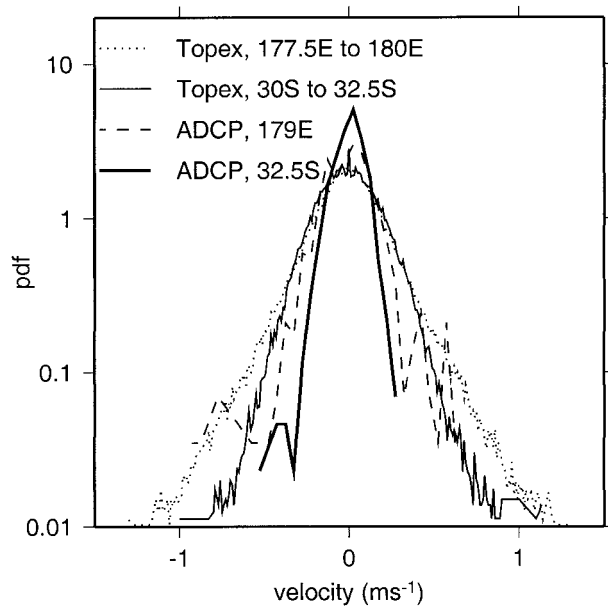


FIG. 4. Probability density distribution of velocities for elongated meridional and zonal boxes in the Pacific Ocean. Pdfs for descending TOPEX tracks represent (dotted line) the meridional region from 60°S to 60°N between 177.5°E and 180° (with the exception of measurements within 10° of the equator), and (thin black line) the zonal region between 32.5°S and 30°S from 180° to 115°W. Pdfs for ADCP zonal velocities from 140-m depth from (gray line) the meridional P14 cruise (179°E between 10° and 51°N) and (heavy black line) the zonal P6 cruise (32.5°S between 175°E and 110°W) are also shown for comparison.

dition, the ADCP velocity range is smaller both because the measurements represent hourly averages and because velocities attenuate with depth. Nonetheless, the ADCP pdfs show the same trends as the TOPEX pdfs. Velocities from the meridional P14 section have more outliers than velocities from the zonal P6 section, again indicating that a meridional box of velocities has a more exponential pdf, while a zonal box has a more Gaussian pdf.

Statistical methods such as the chi-squared or Kolmogorov–Smirnov tests can be used to determine whether data distributions are Gaussian or exponential (D’Agostino 1982). These tests are sensitive measures that determine the likelihood that randomly selected variables drawn from a known probability distribution would more closely duplicate the pdf than the observations. TOPEX altimeter measurements contain significant high-wavenumber noise due to a variety of instrumental and geophysical effects including satellite tracking errors, sea ice, and the influence of strong winds (Stammer 1997). These are likely to result in numerous outliers in the velocity field so that the velocity distributions are expected to have long tails. Within the stringent criteria imposed by statistical tests, these outliers dominate the velocity distributions, and the data appear to be drawn neither from Gaussian nor from exponential distributions.

TABLE 1. Percentage of boxes for which a Gaussian distribution is better than an exponential distribution at representing the observed data. The first column indicates the sizes of the boxes over which data are averaged, measured in degrees latitude by degrees longitude. The middle three columns indicate the fraction of pdfs best represented by a Gaussian distribution for descending, ascending, and combined analyses, where velocities are unnormalized. In the fifth column, velocities from each 2.5° box have been normalized by the box variance before histograms are computed and boxes are merged. Data within 10° of the equator and data more than 60° from the equator are omitted.

Box size (lat × long)	Unnormalized			Normalized Both
	Descending	Ascending	Both	
2.5° × 2.5°	77	78	81	76
5° × 5°	76	77	80	89
10° × 10°	70	71	71	91
20° × 20°	61	65	65	93
30° × 30°	53	50	54	93
40° × 40°	41	44	44	96
60° × 60°	0	8	8	100
Global	0	0	0	100
2.5° × 40°	71	70	74	93
2.5° × 60°	67	68	67	92
40° × 2.5°	51	50	53	94
60° × 2.5°	19	16	17	93

For this study, instead of using strict statistical tests, we least squares fitted Gaussian and exponential distributions to the observed velocity pdfs. Since we assumed that outliers might corrupt the results, only the core of the pdf was least squares fitted. Our pdfs had velocity bins that were 0.025 m s⁻¹ wide, and bins containing fewer than five data points or with a value of $p(v)$ less than 2% of the pdf maximum were discarded from the fit. We then assessed whether the misfit between the observed pdf and the fitted distribution was larger for the Gaussian or for the exponential distribution.

The percentages shown in the middle three columns of Table 1 indicate the fraction of boxes throughout the global ocean for which a Gaussian distribution provides a better fit than an exponential distribution. Within individual 2.5° by 2.5° boxes, the Gaussian distribution is better about 80% of the time. However, as the box size increases, we find that the data are progressively more exponential and less Gaussian. As shown in the lower section of Table 1, this effect is particularly pronounced if the boxes are long in the meridional direction. In contrast, boxes that are meridionally thin but zonally broad have pdfs that are almost as likely to be Gaussian as 2.5° × 2.5° boxes. The statistics for ascending and descending tracks are nearly the same, suggesting no strong anisotropy between the two-track orientations and allowing us to combine ascending and descending data for further analysis.

As noted in section 2, the fitted Gaussian or exponential width should be related to the eddy kinetic energy. Figure 5a shows the mean widths s for the Gaussian distribution $p_1(v)$ from (4), and Fig. 5b indicates the mean rms velocity (the square root of EKE) for each

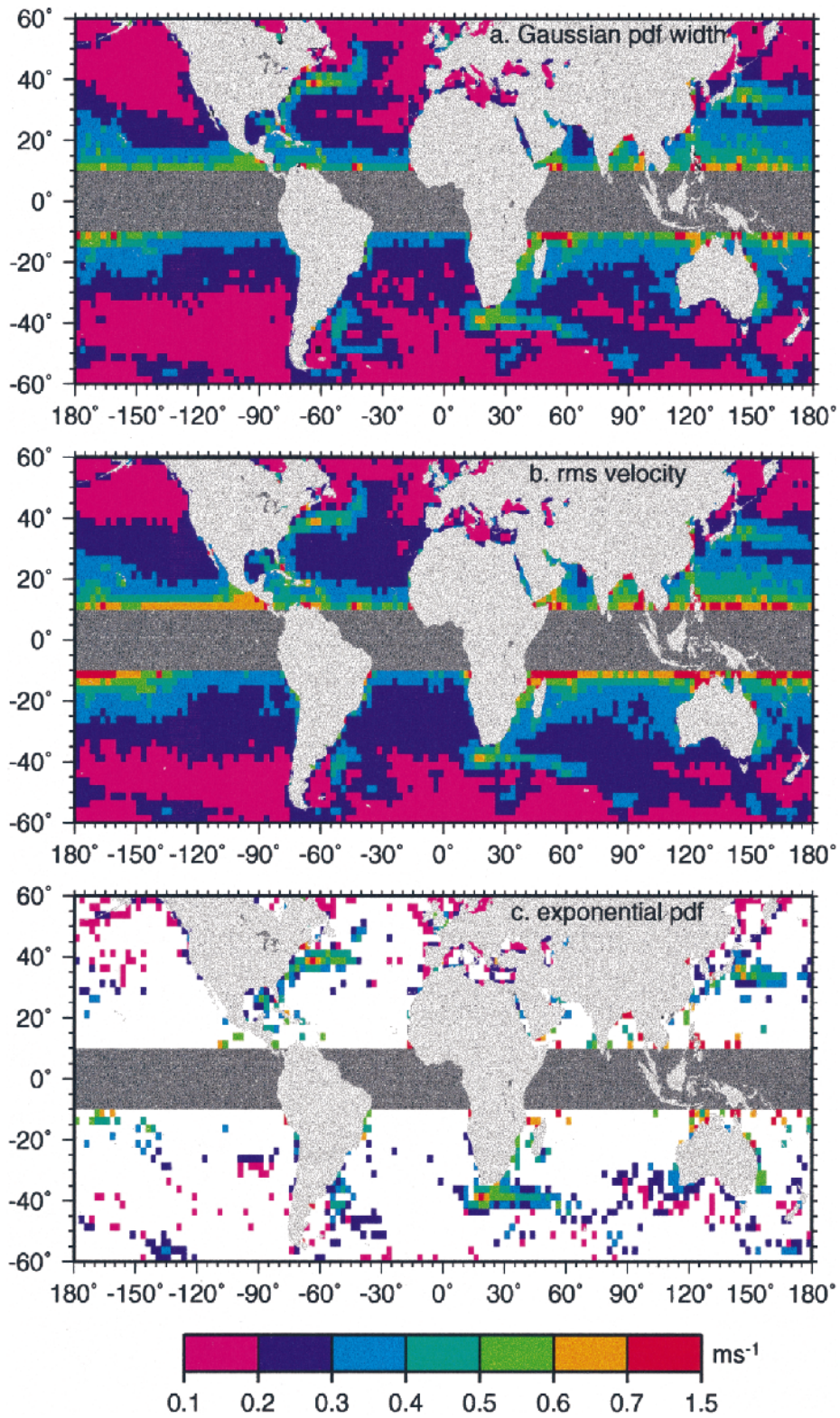


FIG. 5. (a) Width of Gaussian distribution, (b) estimated rms velocity for 2.5° by 2.5° boxes, and (c) width of the exponential distribution in boxes in which it provides a better fit than a Gaussian distribution.

2.5° by 2.5° box in the global ocean for all of the ascending and descending satellite tracks. The spatial patterns match closely, though the widths in Fig. 5a are systematically smaller than the rms velocities, because the least squares fitting procedure omitted outliers that can strongly influence the rms velocity estimates. The implied zonally averaged energies in these 2.5° boxes are roughly double those estimated by Stammer (1997) because we include the local variability within 2.5° boxes in our energy calculation.

Energy varies with latitude primarily because of the inverse relation between geostrophic velocity and the Coriolis parameter f . In addition, EKE is high in the regions where western boundary currents leave the continental slopes. Together these two effects mean that EKE varies more dramatically in the meridional direction than in the zonal direction.

Figure 5c indicates which data boxes are better represented by an exponential pdf rather than a Gaussian pdf. These points appear most frequently in regions such as the Gulf of Mexico, the Gulf Stream, the Kuroshio Extension, the Malvinas confluence, and the Agulhas retroflection that are marked not only by large rms velocities but also by strong variations in rms velocity over short length scales associated with baroclinic instability. Thus a single 2.5° box might encompass subregions with a broad range of pdf widths. Coastal regions including the coastal upwelling region on the west coast of Africa and the shallow region of strong tides near Britain and Ireland also tend to have non-Gaussian pdfs. However, tide-model and orbit errors in shallow water may make these pdfs suspect. In contrast, the North Pacific subtropical countercurrent, where Qiu (1999) identified a strong seasonal cycle, does not have an exponential pdf. Likewise the Antarctic Circumpolar Current is more Gaussian than exponential.

When all boxes of ascending and descending velocities are combined, we obtain the global velocity pdf shown in Fig. 6 based on 24.8 million velocity measurements. Like the meridional pdfs in Fig. 4, this pdf is strongly exponential, but shows a parabolic flattening near $v = 0$. In the range $-1.5 < v < 1.5 \text{ m s}^{-1}$, the mean rms misfit between the log of the observed pdf and the log of the fitted Gaussian distribution is 8.6. An exponential distribution provides a better fit, with a mean rms misfit of only 0.27. The observed pdf is most accurately described by a function of the form

$$p_j(v) = A \exp\left(-\frac{av^2}{1 + b|v|^\alpha}\right), \quad (6)$$

which is Gaussian near $v = 0$ and has a generalized exponential structure for large $|v|$. From an empirical fit, we obtain $a = 25.7$, $b = 4.4$, $\alpha = 1.5$. The constant $A = 1.99$ is determined so that $p_j(v)$ is correctly normalized. In log units, the resulting mean rms misfit is 0.06.

If we normalize the velocities in each 2.5° box by the

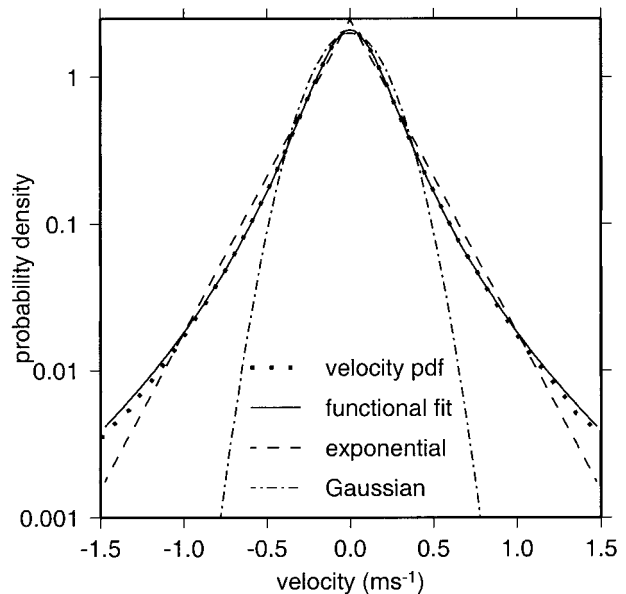


FIG. 6. Probability density distributions of all TOPEX descending and ascending track velocities for the global ocean (black dotted line), along with (gray line) the empirical functional fit represented by (6), (dashed line) an exponential fit to the pdf, and (gray dashed line) a Gaussian fit.

local variance before combining them to compute regional pdfs, we find that nearly all of the boxes have Gaussian pdfs, as shown in the fifth column of Table 1. With the exception of boundary current regions, normalized pdfs display a universal Gaussian shape, with no variations due to latitude. In contrast, Stammer (1997) showed that, in addition to varying with EKE, spectra also “cut off” at lower wavenumbers near the equator. Variance-normalized pdfs contain no wavenumber information and no signature analogous to low wavenumber cutoff at low latitudes.

To summarize, pdfs are Gaussian for regions with constant energy such as small boxes or zonal boxes, but they are exponential in regions with large variations in mesoscale energy such as the 60° by 60° and 60° by 2.5° boxes in Table 1 or the meridional boxes of Fig. 4. That EKE [or equivalently the width of $p(v)$] is more variable in regions with exponential pdfs will provide the key element to explaining why global ocean velocity pdfs are non-Gaussian.

4. Statistics of velocity pdfs

Since the two-dimensional turbulence results of Jiménez suggest that Eulerian velocities should have Gaussian distributions, the non-Gaussian pdfs in the previous section might seem unexpected. In this section we show that a simple statistical model, coupled with knowledge of the distribution of eddy kinetic energy in the ocean, is sufficient to explain why the distribution of velocities can be exponential.

Assume that at any given location, velocities have a Gaussian probability density function $p_1(v)$ as in (4). Now suppose that the width s of the distribution varies throughout the ocean. In the lexicon of turbulence theory, this is equivalent to stating that energy dissipation is intermittent (Castaing et al. 1990). The term “intermittent” is used with a variety of definitions to describe signals that appear bursty or are patchy in space. Physically, intermittency occurs when different subsets of a signal do not have the same distribution or, in other words, when a field appears to have different pdfs for different regions in space or time. Our observation that pdf widths vary in the meridional direction is one example of this phenomenon.

Kolmogorov (1962) first explored this possibility by taking s to have a lognormal distribution, and Castaing et al. (1990) examined the consequences of this for the distribution of velocity increments in three-dimensional turbulence. In this study, we will draw s from a more general set of pdfs $q(s)$. The pdf that we expect to be determined from observations is then given by

$$p(v) = \frac{1}{\sqrt{2\pi}} \int_0^\infty \exp\left(-\frac{v^2}{2s^2}\right) \frac{q(s)}{s} ds. \quad (7)$$

As a simple example, if we take $q(s) = \delta(s - s_0)$, the velocity pdf reduces to a Gaussian distribution with variance s_0^2 . This corresponds to the observation that most small boxes have Gaussian velocity pdfs: in limited

regions, the velocity variance s^2 is distributed over a narrow range. This is also the case that we construct in the fifth column of Table 1 when we normalize the velocities in each of the 2.5° boxes by the local variance, and hence produce strictly Gaussian pdfs.

In general, the pdf $p(v)$ will depend on the form of the function $q(s)$. The appendix discusses how we can obtain information about the asymptotic behavior of $p(v)$ for large and small v . Straight exponential tails will be obtained for distributions $q(s)$ that decay like $\exp(-s^2)$.

For the altimeter observations of the global ocean, Fig. 7 shows probability distributions of pdf width and rms velocity. Although the most frequent width is 0.2 m s^{-1} , much larger widths are observed in some of the 2.5° by 2.5° boxes. The distribution is noisy, but may be fitted approximately by a function that increases parabolically near the origin, and decreases exponentially for large s .

The behavior near $s = 0$ of the curves of Fig. 7 is smooth, and the results of the appendix show that the pdf $p(v)$ will be parabolic near the origin. This feature does not depend on the precise form of the function used to fit the data in Fig. 7.

For large s , we pick pdfs $q(s)$ that decay exponentially. We know from the appendix that such distributions lead to velocity pdfs that have generalized exponential decay. We shall consider two possibilities. The first is a Rayleigh–Maxwell distribution with the origin shifted to s_0 :

$$q(s) = \begin{cases} \frac{1}{2^{(m-1)/2} a^{m+1} \Gamma\left(\frac{m+1}{2}\right)} (s - s_0)^m \exp\left(-\frac{(s - s_0)^2}{2a^2}\right), & s > s_0 \\ 0, & s < s_0. \end{cases} \quad (8)$$

The special case $m = 1$ corresponds to the classical Rayleigh distribution. For the case $s_0 = 0$, the pdf $p(v)$ can be calculated analytically (Holdom 1997) and takes the form

$$p(v) = \frac{1}{\sqrt{\pi} 2^{m/2} a^{m/2+1} \Gamma\left(\frac{m+1}{2}\right)} |v|^{m/2} K_{m/2}(|v|/a), \quad (9)$$

where K is a modified Bessel function. The case $m = 1$ can be written in the simpler form

$$p(v) = \frac{1}{2a} e^{-|v|/a}, \quad (10)$$

a symmetric exponential distribution. Fitting the distribution (8) to the width of the Gaussian pdfs gives the

following values when m is taken to be unknown: $a = 0.16$, $s_0 = 0.12$, and $m = 0.04$. If we fix $m = 1$, the corresponding values are $a = 0.18$ and $s_0 = 0.04$.

We shall also consider the Gamma probability distribution, again with a shift of origin:

$$q(s) = \begin{cases} \frac{1}{a^{m+1} \Gamma(m+1)} (s - s_0)^m \exp\left(-\frac{|s - s_0|}{a}\right), & s > s_0 \\ 0, & s < s_0. \end{cases} \quad (11)$$

Fitting the Gamma distribution (11) gives $a = 0.10$, $s_0 = 0.12$, and $m = 0.46$. The results of these fits, shown in Fig. 8, show that all three functions duplicate the leading features of the observed $q(s)$ but that the Gamma distribution most closely matches the observations,

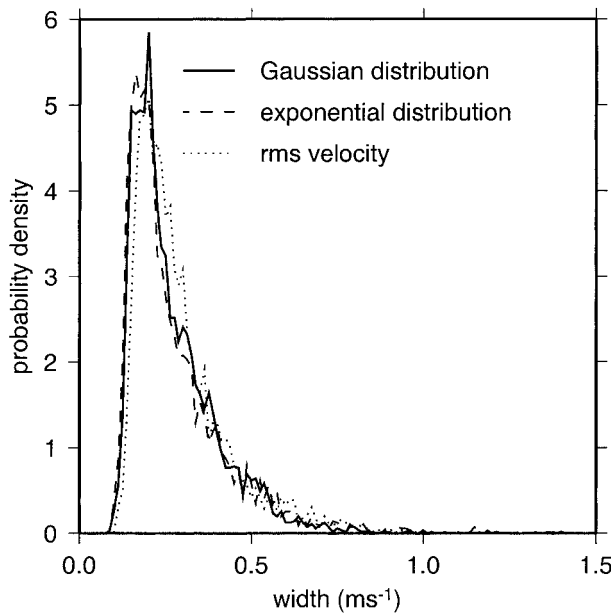


FIG. 7. Distribution of pdf widths (in m s^{-1}) for the Gaussian pdf, and the exponential pdf determined by least squares fitting the observed data for each 2.5° by 2.5° box between 10° and 60° north and south of the equator. Also shown is the distribution of rms velocity for the same boxes.

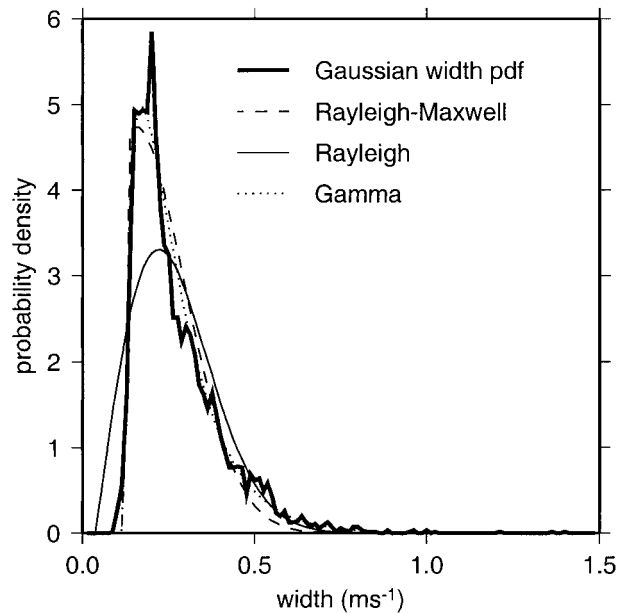


FIG. 8. Functional fits of width pdfs $q(s)$ to the distribution of Gaussian widths shown in Fig. 7.

while the Rayleigh distribution is least successful. We also examined a lognormal distribution for $q(s)$ but found poor agreement.

Using the $q(s)$ functions shown in Fig. 8 along with the assumption that velocities are normally distributed within small regions, we can now predict the global form of $p(v)$, by integrating (7). We already know from the asymptotic analysis of the appendix that $p(v)$ will decay like $\exp(-|v|)$ for the Rayleigh–Maxwell distribution and like $\exp(-|v|^{2/3})$ for the Gamma distribution. In Fig. 9 we show the integrated $p(v)$ in the three cases. Predictions based on all three distributions closely match the basic structure of the observed $q(s)$, by capturing both the parabolic behavior of the pdf near $v = 0$ and the generalized exponential decay of the pdf tails. The three distributions each have different velocity ranges in which they provide the best match to $p(v)$. However all three distributions underpredict the peak value of the pdf by about 20%, overpredict the observations around $\pm 0.5 \text{ m s}^{-1}$, and underpredict the tails. In the range $-1.5 < v < 1.5 \text{ m s}^{-1}$ the rms difference between the log of the observed distribution and the log of the theoretical distribution is 0.63 for the Rayleigh–Maxwell distribution, 0.34 for the Rayleigh distribution, and 0.20 for the Gamma distribution.

The imperfections in the predicted values of $p(v)$ can be attributed to a number of causes. First, ocean velocities may not always be locally Gaussian as we assumed in (7). As an alternative, we considered the possibility that ocean velocities have the exponential distribution $p_2(v)$ from (5). In this case the corresponding model,

analogous to (5), predicts global pdfs that are 50% larger than the observed velocity pdf, do not flatten out near $v = 0$, and in general do not fit the observations as well as the results depicted in Fig. 9. Since exponential velocity pdfs lead us to overpredict $p(0)$ and Gaussian pdfs lead us to underpredict $p(0)$, we might infer that, at the 2.5° resolution of this study, the ocean itself includes a few regions with exponential velocity pdfs and a majority of regions with Gaussian pdfs, as suggested in Fig. 5c.

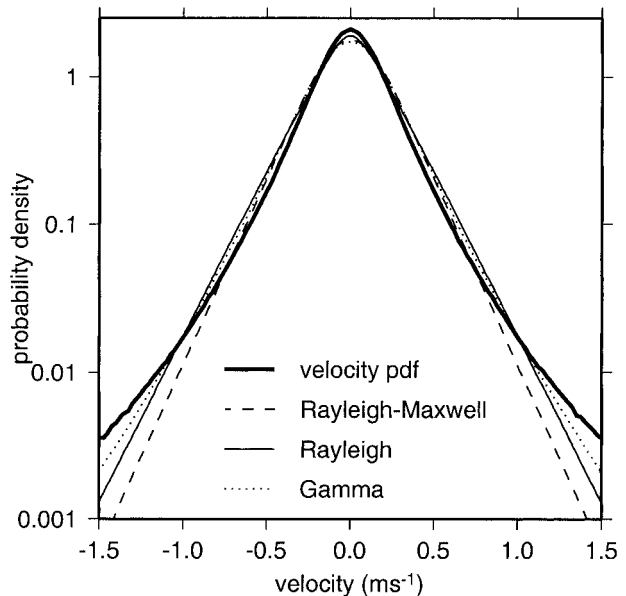


FIG. 9. The observed global pdf (heavy lines) and theoretical predictions of $p(v)$ based on the $q(s)$ functions shown in Fig. 8.

A second reason why our predicted values of $p(v)$ might not match the observations perfectly in Fig. 9 may stem from the fact that $q(s)$ is determined from the pdf widths of the 2.5° boxes and is not weighted by the number of measurements available in each box. We considered correcting for this effect by weighting the widths s by the number of samples in each box when we estimated $q(s)$. However, this strategy did not work well because the resulting $q(s)$ had most of its samples near $s = 0.2$ so that it behaved more like a delta function. Therefore our predicted $p(v)$ had a significantly more Gaussian structure than the data would indicate.

Overall, the good match between the observed and predicted velocity pdfs in Fig. 9 suggests that the hypothesis invoked at the beginning of this section is consistent so that, within small regions of the ocean, velocities usually have Gaussian pdfs. This agrees with the results of two-dimensional turbulence (Jiménez 1996). Over large regions of the ocean, the hypotheses underlying point vortex models and two-dimensional turbulence are not satisfied: bottom depth changes, continents bound the circulation, spatially varying winds force the flow, and eddy sizes and strengths depend on latitude, stratification, and baroclinic instability of the local mean flow. As a result, when larger regions of the ocean are considered, EKE (or pdf width) varies substantially, and velocity pdfs are nearly exponential, with smooth parabolic behavior near $v = 0$, in accord with the theory outlined in this section.

Although this analysis is based on Eulerian velocities, the pdfs that we find in Fig. 9 resemble the results obtained by Bracco et al. (2000) from Lagrangian float data taken from a broad range of latitudes in the Atlantic Ocean. There are no doubt differences in the statistics of Eulerian and Lagrangian velocities (Davis 1983), but the basic characteristics of the statistical model presented here for Eulerian velocities should also apply to Lagrangian measurements. Thus floats that are confined to small and energetically homogeneous regions of the ocean might produce Gaussian pdfs, while floats that travel through regions with a wide range of background eddy kinetic energy are likely to yield exponential pdfs as seen by Bracco et al. (2000).

5. Summary

We have calculated pdfs from nearly five years of global surface velocity anomalies measured by the TOPEX altimeter. We divided the ocean into 2.5° by 2.5° boxes, omitting data within 10° of the equator where geostrophic velocities are most sensitive to noise. We least squares fitted Gaussian and exponential distributions to the pdfs in each of these boxes. In more than 80% of boxes, the Gaussian distribution fitted the observed pdfs more closely than the exponential, as might have been predicted from calculations based on two-dimensional turbulence theory (Jiménez 1996). The boxes in which pdfs were better fitted by an exponential

distribution corresponded primarily to regions such as western boundary currents, where baroclinic instability of jets and fronts, for example, is expected to be strongest and correspondingly where eddy kinetic energy is not only high but also varies significantly in space. These locations may also be marked by strong temporal variability of eddy kinetic energy, which would have similar consequences in our analysis.

When 2.5° boxes were combined to make larger boxes, we found that as the size of the box increased, the pdfs were increasingly likely to be more exponential than Gaussian. For boxes larger than 30° by 30° , exponential pdfs were a better fit in the majority of cases. This effect was also observed in cases where we extended the box size in the meridional direction, while extending the box in the zonal direction yielded a Gaussian distribution. Although more limited in quantity, ADCP data from the meridional WOCE line along 179°E and from the zonal section at 32.5°S showed similar results. For the global ocean the velocity pdf, based on nearly 25 million observations, showed a parabolic shape near $v = 0$ with generalized exponential tails for large $|v|$. Recent Lagrangian studies with floats and numerical models have also shown exponential pdfs (Bracco et al. 2000; Weiss et al. 1998). If we normalize the velocity data by their variance, then the resulting pdfs are Gaussian for large boxes.

We proposed an explanation for why ocean pdfs should be non-Gaussian based on the observation that the rms velocity within small boxes, which is roughly equivalent to the width of the velocity pdf, varies substantially over large regions of the ocean, particularly in the meridional direction. Our theory also explains why pdfs computed from normalized velocities are Gaussian as box size increases. Based on the geographic variation in the width s of velocity pdfs, we developed a pdf for the width itself, $q(s)$. This probabilistic framework allowed us to superpose Gaussian distributions of varying widths, and we showed that the resulting global velocity pdf should be non-Gaussian.

We considered several simple analytical models for the variance distribution $q(s)$ and fitted these to the observations of s from the altimeter data. The best fit was given by a Gamma distribution. We then used these analytic fits to calculate the velocity pdfs numerically. The calculated pdfs are roughly exponential, but flatten near $v = 0$, closely matching the observations. As shown in the appendix, the parabolic behavior near $v = 0$ occurs, because $q(0) = 0$ and $q(s)$ is smooth near the origin.

Based on these results, we conclude that ocean velocities have a Gaussian distribution within small regions of the ocean, and only within these small energetically homogeneous regions is simple two-dimensional turbulence likely to provide plausible predictions of ocean variability. On a global scale, the superposition of velocity pdfs from regions with differing eddy kinetic energy will result in non-Gaussian pdfs.

Acknowledgments. Conversations with Peter Haynes, Joe LaCasce, and Antonello Provenzale were invaluable. Mara Yale and David Sandwell provided invaluable assistance in processing the altimeter measurements. The anonymous reviewers provided helpful comments. This research was supported by a Research Fellowship from Queens' College, Cambridge (SGLS), by a fellowship from the North Atlantic Treaty Organization awarded in 1997 (STG), and by JPL Contract 960875 with Scripps Institution of Oceanography. Financial support from the Transport processes in the Atmosphere and the Oceans (TAO) programme of the European Science Foundation is also acknowledged.

APPENDIX

Asymptotic Behavior of the pdf $p(v)$

We consider the behavior of the function $p(v)$, defined by the integral (7), for large and small v . This integral is well-defined for all values of v . The constraints on the behavior of $q(s)$ will become clear during the course of the analysis. (For more background on some of the results that we will use, see, e.g., Hinch 1991.)

a. The tails of the pdf: The case of large v

We first consider the case of large v . The pdf $q(s)$ decays for large s and the nature of its decay is relevant

to these asymptotics. We shall consider three possibilities: algebraic decay, generalized exponential decay, and lognormal decay. These three possibilities correspond to

$$q(s) = \begin{cases} A_1 \hat{q}(s) s^{-a} \\ A_2 s^b \exp(-(s/s_0)^\beta) \\ A_3 s^{-1} \exp(-\ln^2 s/\lambda^2), \end{cases} \quad (A1)$$

where the A are normalization constants. The first case contains a function $\hat{q}(s)$, which tends to 1 for large s .

In the first case, the change of variable $y = 1/(2s^2)$ leads to the integral

$$p(v) = \frac{A_1}{\sqrt{2\pi}} \int_0^\infty e^{-v^2 y} \hat{q}([2y]^{-1/2}) (2y)^{-1+a/2} dy. \quad (A2)$$

Watson's lemma then gives

$$p(v) \sim A_1 2^{(a-3)/2} \pi^{-1/2} \frac{\Gamma(a/2)}{v^a}. \quad (A3)$$

The pdf $p(v)$ thus decays algebraically, which is too slow to match the data.

In the second case, the change of variable $s = v^{2/(2+\beta)} y$ leads to

$$p(v) = A_2 \frac{v^{2b/(2+\beta)}}{\sqrt{2\pi}} \int_0^\infty \exp\left\{-v^{2\beta/(2+\beta)} \left[\frac{1}{2y^2} + \left(\frac{y}{s_0}\right)^\beta\right]\right\} y^{b-1} dy. \quad (A4)$$

Laplace's method applied to this integral gives

$$p(v) \sim A_2 \frac{\beta^{-(b+1)/(2+\beta)}}{\sqrt{2+\beta}} s_0^{\beta(3-b)/(2+\beta)} v^{(2b-\beta)/(2+\beta)} \times \exp\left[-\left(\frac{v}{s_0}\right)^{2\beta/(2+\beta)} \left(1 + \frac{\beta}{2}\right) \beta^{-\beta/(2+\beta)}\right]. \quad (A5)$$

This implies generalized exponential decay for large v , which could represent the observed oceanic pdfs.

The third case is less transparent. The resulting integral cannot be cast into a form directly suitable for Laplace's method. However, the dominant behavior still comes from the point where the argument of the integrand is maximal. The change of variable $y = 1/s$ gives

$$p(v) = \frac{A_3}{\sqrt{2\pi}} \int_0^\infty \exp\left[-\frac{v^2 y^2}{2} - \frac{\ln^2 y}{2\lambda^2}\right] dy. \quad (A6)$$

For large v , the maximum in the square bracket comes from a point y_m where $y_m \ll 1$. The value y_m is determined by the transcendental equation

$$y_m^2 = \frac{\ln y_m^{-1}}{v^2 \lambda^2}. \quad (A7)$$

An approximation to y_m may be found by iteration and takes the form

$$y_m = \frac{(\ln \lambda v)^{1/2}}{\lambda v} + \dots \quad (A8)$$

The large- v behavior for $p(v)$ is then given by

$$p(v) \sim A_3 \left(\frac{2\pi}{v^2 + (1 - \ln y_m)/y_m^2 \lambda^2}\right)^{1/2} \times \exp\left[-\frac{v^2 y_m^2}{2} - \frac{\ln^2 y_m}{2\lambda^2}\right]. \quad (A9)$$

An approximation to this, which is not asymptotic but takes a simple form, is

$$p(v) \sim A_3 \left(\frac{\pi}{2v^2}\right)^{1/2} \exp\left[-\frac{\ln^2 v}{2\lambda^2}\right]. \quad (A10)$$

The decay for $p(v)$ is between exponential and algebraic.

The results (A9) and (A10) also give the asymptotic behavior of the velocity difference pdf considered by Castaing et al. (1990) in the case of zero skewness.

b. The center of the pdf: Near $v = 0$

For small v , the behavior of $p(v)$ is determined by the behavior of $q(s)$ near the point $s = 0$. Since $q(s)$ is a pdf, it must be $o(s^{-1})$ for small and large s (to ensure that the normalization integral exists). However, we can immediately see that if $q(s) = o(1)$ for small s , then

$$p(v) \sim \frac{1}{\sqrt{2\pi}} \int_0^\infty \frac{q(s)}{s} ds + o(1) \quad (\text{A11})$$

for small v . We shall therefore insist on the condition $q(s) = o(1)$ for small s to ensure that $p(0)$ is finite. This corresponds to requiring that the probability that $s = 0$ be vanishingly small.

It is apparent that if $q(s) = o(s^2)$ near the origin, then

$$p(v) \sim \frac{1}{\sqrt{2\pi}} \int_0^\infty \frac{q(s)}{s} ds - \frac{v^2}{2\sqrt{2\pi}} \int_0^\infty \frac{q(s)}{s^3} ds + o(v^2). \quad (\text{A12})$$

Therefore $p(v)$ is parabolic near the origin. The case where $q(s)$ is larger than $o(s^2)$ near the origin leads to intermediate behavior where $p(v) \sim p_0 + O(v, v \ln v)$ depending on the precise form of $q(s)$.

We have shown that the asymptotic behavior of $p(v)$ for large and small v may be obtained from the corresponding behavior of $q(s)$. In particular, algebraic decay of $q(s)$ for large s leads to algebraic decay of $p(v)$. Generalized exponential decay of $q(s)$ leads to exponential decay of $p(v)$, but with a different power in the exponent. The special case of Gaussian decay in $q(s)$ however leads to Gaussian decay in $p(v)$. A lognormal form for $q(s)$ leads to almost lognormal decay in $p(v)$. If $q(s)$ is sufficiently smooth, namely $o(s^2)$, then $p(v)$ will be parabolic near the origin.

REFERENCES

- Benada, J. R., 1997: Merged GDR (TOPEX/POSEIDON) Generation B User's Handbook. PO.DAAC User Services Office, Jet Propulsion Laboratory, M/S 300–320. [Available from Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, CA 91109.]
- Bracco, A., J. H. LaCasce, and A. Provenzale, 2000: Velocity pdfs for oceanic floats. *J. Phys. Oceanogr.*, in press.
- Castaing, B., Y. Gagne, and E. J. Hopfinger, 1990: Velocity probability density functions of high Reynolds-number turbulence. *Physica D*, **46**, 177–200.
- D'Agostino, R. B., 1982: Tests for departures from normality. *Encyclopedia of Statistical Sciences*, S. Kotz and N. L. Johnson, Eds., Wiley, 315–324.
- Davis, R. E., 1983: Oceanic property transport, Lagrangian particle statistics, and their prediction. *J. Mar. Res.*, **41**, 163–194.
- Frisch, U., 1995: *Turbulence: The Legacy of A. N. Kolmogorov*. Cambridge University Press, 296 pp.
- Fu, L. L., E. J. Christensen, C. A. Yamarone Jr., M. Lefebvre, Y. Ménard, M. Dorrer, and P. Escudier, 1994: TOPEX/POSEIDON mission overview. *J. Geophys. Res.*, **99**, 24 369–24 381.
- Haynes, P. H., 1999: Transport, stirring and mixing in the atmosphere. *Mixing, Chaos and Turbulence*, H. Chaté and E. Villermanx, Eds., Plenum, 229–272.
- Hinch, E. J., 1991: *Perturbation Methods*. Cambridge University Press, 160 pp.
- Holdom, B., 1997: Intermittency and the slow approach to the Kolmogorov scaling. *Phys. Rev. E*, **55**, 7000–7004.
- Jiménez, J., 1996: Probability densities in two-dimensional turbulence. *J. Fluid Mech.*, **313**, 223–240.
- Kolmogorov, A. N., 1962: A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds number. *J. Fluid Mech.*, **13**, 82–85.
- Lesieur, M., 1997: *Turbulence in Fluids: Stochastic and Numerical Modelling*. 3d ed. Kluwer, 515 pp.
- Lilly, D. K., 1969: Numerical simulation of developing and decaying two-dimensional turbulence. *J. Fluid Mech.*, **45**, 395–415.
- Llewellyn Smith, S. G., and S. T. Gille, 1998: Probability density functions of large-scale turbulence in the ocean. *Phys. Rev. Lett.*, **81**, 5249–5252.
- McClean, J. L., A. J. Semtner, and V. Zlotnicki, 1997: Comparisons of mesoscale variability in the Semtner-Chervin $1/4^\circ$ model, the Los Alamos Parallel Ocean Program $1/6^\circ$ model, and TOPEX/POSEIDON data. *J. Geophys. Res.*, **102**, 25 203–25 226.
- McWilliams, J. C., 1984: The emergence of coherent isolated vortices in turbulent flow. *J. Fluid Mech.*, **146**, 21–43.
- Min, I. A., I. Mezić, and A. Leonard, 1996: Lévy stable-distributions for velocity and velocity difference in systems of vortex elements. *Phys. Fluids*, **8**, 1169–1180.
- Priestley, M. B., 1981: *Spectral Analysis and Time Series*. Academic Press, 890 pp.
- Qiu, B., 1999: Seasonal eddy field modulation of the North Pacific Subtropical Countercurrent: TOPEX/Poseidon observations and theory. *J. Phys. Oceanogr.*, **29**, 2471–2486.
- Rhines, P. B., 1979: Geostrophic turbulence. *Ann. Rev. Fluid Mech.*, **11**, 401–443.
- Salmon, R., 1998: *Lectures on Geophysical Fluid Dynamics*. Oxford University Press, 378 pp.
- Samelson, R. M., 1996: Chaotic transport by mesoscale motions. *Stochastic Modelling in Physical Oceanography*, R. J. Adler, P. Müller, and B. L. Rozovskii, Eds., Birkhäuser, 423–438.
- Sandwell, D. T., and W. H. F. Smith, 1997: Marine gravity anomaly from Geosat and ERS 1 satellite altimetry. *J. Geophys. Res.*, **102**, 10 039–10 054.
- Stammer, D., 1997: Global characteristics of ocean variability estimated from regional TOPEX/POSEIDON altimeter measurements. *J. Phys. Oceanogr.*, **27**, 1743–1769.
- , and C. Wunsch, 1994: Preliminary assessment of the accuracy and precision of TOPEX/POSEIDON altimeter data with respect to the large-scale ocean circulation. *J. Geophys. Res.*, **99**, 24 584–24 604.
- Weiss, J. B., A. Provenzale, and J. C. McWilliams, 1998: Lagrangian dynamics in high-dimensional point-vortex systems. *Phys. Fluids*, **10**, 1929–1941.
- WOCE Data Products Committee, 1998: WOCE Global Data Version 1.0. Tech. Rep. 158, WOCE International Project Office, CD-ROM, 13 disks. [Available from Southampton Oceanography Centre, European Way, Southampton SO14 3ZH, United Kingdom.]
- Wunsch, C., and D. Stammer, 1995: The global frequency–wavenumber spectra of oceanic variability estimated from TOPEX/POSEIDON altimetric measurements. *J. Geophys. Res.*, **100**, 24 895–24 910.
- Yale, M. M., D. T. Sandwell, and W. H. F. Smith, 1995: Comparison of along-track resolution of stacked Geosat, ERS 1 and TOPEX satellite altimeters. *J. Geophys. Res.*, **100**, 15 117–15 127.