Geophys. J. Int. (1997) **128**, 708–722

Combination of TOPEX/POSEIDON data with a hydrographic inversion for determination of the oceanic general circulation and its relation to geoid accuracy

Alexandre Ganachaud,[1] Carl Wunsch,[1] Myung-Chan Kim[2] and Byron Tapley[2]

[1] *Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, MA* 02139, *USA. E-mail: ganacho@gulf.mit.edu; cwunsch@pond.mit.edu*

[2] *Center for Space Research, University of Texas, Austin, TX, USA. E-mail: byaa672@charon.cc.utexas.edu; tapley@utcsr.ae.utexas.edu*

Accepted 1996 November 12. Received 1996 October 28; in original form 1996 May 28

SUMMARY

A global estimate of the absolute oceanic general circulation from a geostrophic inversion of *in situ* hydrographic data is tested against and then combined with an estimate obtained from TOPEX/POSEIDON altimetric data and a geoid model computed using the JGM-3 gravity-field solution. Within the quantitative uncertainties of both the hydrographic inversion and the geoid estimate, the two estimates derived by very different methods are consistent. When the *in situ* inversion is combined with the altimetry/geoid scheme using a recursive inverse procedure, a new solution, fully consistent with both hydrography and altimetry, is found. There is, however, little reduction in the uncertainties of the calculated ocean circulation and its mass and heat fluxes because the best available geoid estimate remains noisy relative to the purely oceanographic inferences. The conclusion drawn from this is that the comparatively large errors present in the existing geoid models now limit the ability of satellite altimeter data to improve directly the general ocean circulation models derived from *in situ* measurements. Because improvements in the geoid could be realized through a dedicated spaceborne gravity recovery mission, the impact of hypothetical much better, future geoid estimates on the circulation uncertainty is also quantified, showing significant hypothetical reductions in the uncertainties of oceanic transport calculations. Full ocean general circulation models could better exploit both existing oceanographic data and future gravity-mission data, but their present use is severely limited by the inability to quantify their error budgets.

Key words: gravity, oceanic models, satellite altimetry.

1 INTRODUCTION

In this paper, we explore the problem of estimating the general circulation of the ocean through the combination of *in situ* hydrographic data with absolute satellite altimetric observations. Wunsch & Gaposchkin (1980) discussed the problem and showed schematically how it could be carried out. At that time, the existing altimetric and geoid data were far too crude to attempt a real calculation. Recently, the advent of data of extraordinary accuracy and precision from the TOPEX/POSEIDON mission, and the ancillary great improvements made in estimates of the Earth's gravity field (Tapley *et al.* 1997) lead us to re-open the question of whether improvements can be made in the knowledge of the large-scale steady ocean circulation through the combination of hydrography and altimetry. The reader is reminded, however, that this goal was regarded as perhaps the most difficult of all those set for the mission (TOPEX Science Working Group 1981), and a major focus of this paper is to make a diagnosis of the factors that may today still limit its achievement.

Our general approach is based upon using prior knowledge of the ocean circulation through quantitative use of the classical oceanographic fields of temperature, salinity, etc. That is, one builds on existing information. The mathematical/physical framework that we will use is the same as that outlined by Wunsch & Gaposchkin (1980), being based on hydrostatic quasi-geostrophic ocean dynamics, long transoceanic hydrographic sections, and altimeter-measured sea-surface height relative to an estimated marine geoid. The inverse formalism has been described extensively in the literature; see for example Wunsch (1978) or Wunsch (1996).

In Section 2 we present the inverse model, Section 3 describes the altimetric measurement and the error due to the

geoid, and Section 4 outlines the method for combining the results from hydrography with the altimetric data. It is found in Section 5 that the large errors in the present geoid limit the impact of the altimetric data, our figure of merit being the integrated heat flux across the hydrographic sections. Finally, Section 6 describes the impact of a hypothetical more accurate degree 70 geoid resulting from a possible future gravity mission. If all other errors are ignored, the reduction in the heat flux uncertainties is as high as 60 per cent.

2 PURE HYDROGRAPHIC ESTIMATE

Consider, as an example, a so-called hydrographic section of measured temperature and salinity, running in an east–west direction, and let x be the distance coordinate along the section, with z the corresponding vertical. If $\rho(x, z)$ is the density field as determined from the shipboard measurements, then the 'thermal-wind' equations produce the absolute velocity

$$v_{\mathrm{abs}}(x,z)=v_{\mathrm{rel}}(x,z)+b(x)=-\frac{g}{\rho f}\int_{z_0}^{z}\frac{\partial\rho}{\partial x}\,dz+b(x), \tag{1}$$

where z_0 is an arbitrary depth, called the reference-level, and $b(x)$ is the reference-level velocity. v_{rel} is the relative velocity, based solely upon the observed density field, g is local gravity, and f is the local Coriolis parameter. Both v and b are directed normal to the section. Classically, $b(x)$ was regarded as impossible to determine from shipboard measurements alone, and thus circulation schemes were based upon attempts to guess depths z_0 where v_{abs}, and hence $b(x)$, might plausibly be assumed to vanish. The 'box-inverse' method discussed in the references cited produces estimates of $b(x)$ by demanding that v_{abs} (and the corresponding velocities in the zonal direction) should satisfy physically plausible sets of constraints such as conservation of mass, salt, nutrients, or potential vorticity. The result is an estimate, $\tilde{b}_i=\tilde{b}(x_i)$, where x_i are a set of discrete positions lying between successive pairs of hydrographic stations. (In general, estimates of specific quantities will be denoted by a tilde.) By definition, an inverse method is one for which a quantitative uncertainty estimate for the solution, e.g.

$$\mathbf{P}(1)=\langle(\tilde{\mathbf{b}}-\mathbf{b})(\tilde{\mathbf{b}}-\mathbf{b})^{\mathrm{T}}\rangle, \tag{2}$$

is known. The possibility of determining $\mathbf{P}(1)$ is what dictates the use of the present linear inverse model, rather than any of several more sophisticated, but non-linear, models that are available.

As outlined by Wunsch & Gaposchkin (1980), and also discussed by the TOPEX Science Working Group (1981), if altimetry is to provide new information about the ocean circulation, the accuracy of the measurements must produce significant improvements in the norm of $\mathbf{P}(1)$. For many purposes we care less about the uncertainties of the individual $\tilde{b}_i$ than we do about various weighted linear combinations, $\mathbf{a}^{\mathrm{T}}\tilde{\mathbf{b}}$, where the elements of $\mathbf{a}$ are constants,with the product representing integrated mass, heat, or other property transports. Thus one requires improvement either in the diagonal elements of $\mathbf{P}(1)$ (the variances of the $\tilde{b}_i$) or in the variances of the various transports, given by

$$\langle(\mathbf{a}^{\mathrm{T}}\tilde{\mathbf{b}}-\mathbf{a}^{\mathrm{T}}\mathbf{b})^2\rangle=\mathbf{a}^{\mathrm{T}}\mathbf{P}(1)\mathbf{a}. \tag{3}$$

A full discussion of the general estimation methods can be found in the references (see Bennett 1992; Wunsch 1996). But one point must be emphasized: the constraints employed generally reflect flow fields integrated over large distances and areas (see Fig. 1). That is, although the ocean circulation is known to be highly time-variable, the estimates based on long-distance integrations such as transoceanic sections make the plausible, but unproven, assumption that time-variable elements tend to become very small. This assumption is reasonably well-founded, in view of the observed stability over decades of the large-scale density structure, but its validity must be a constant focus of vigilance. The same assumption is applied to integrals across intense boundary currents such as

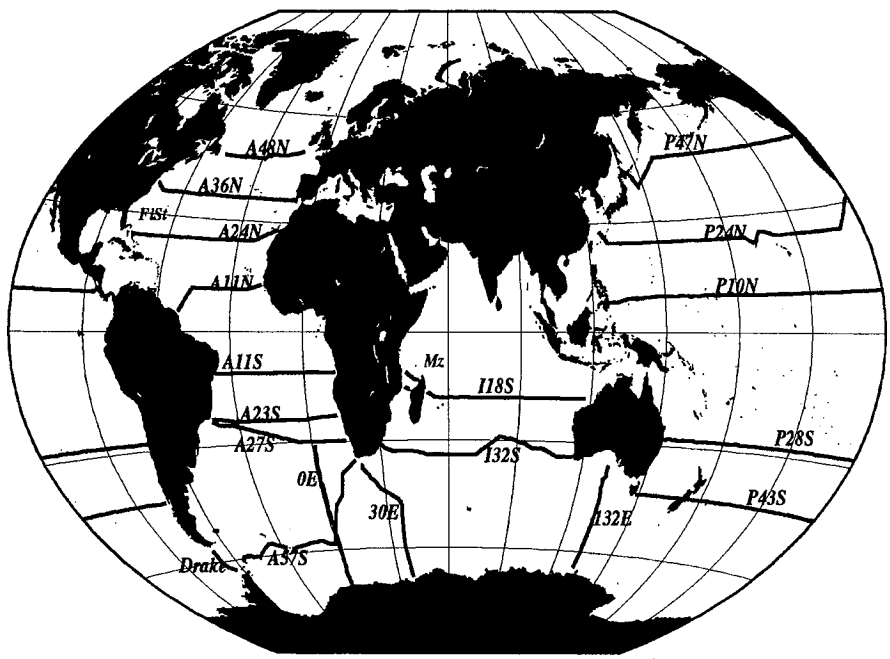

Figure 1. The station positions of the hydrographic/CTD sections from Macdonald (1995) and Macdonald & Wunsch (1996).

the Gulf Stream where it is bordered by Florida and the Bahamas, and where intensive observation programmes over several decades have shown no convincing evidence for secular shifts. We will, however, return to the time-variation problem.

Here, the prior estimate of the ocean circulation is taken from the recent computation of Macdonald (1995) and Macdonald & Wunsch (1996), which is based on the box inverse method of Wunsch (1978). The data consist of hydrography (temperature, salinity, oxygen, nutrients) as measured along the transoceanic lines shown in Fig. 1. These data—about 1600 pairs of stations—are used to calculate the thermal wind relative to deep reference levels, which vary with location (typically at about 1500 m depth) and produce a first estimate $v_{\rm rel}(\mathbf{r}_i, z)$ of the oceanic geostrophic flow perpendicular to the lines between each station pair located at $\mathbf{r}_i$. (Horizontal positions, $\mathbf{r}$, are in spherical coordinates.) Integration constants $b(\mathbf{r}_i)$ are estimated by requiring that the $\mathbf{v}_{\rm abs}$ satisfy near-conservation of mass, salt, oxygen, silicate and phosphate over the large areas shown in Fig. 1. A prior covariance $\mathbf{P}(0)$ is assigned to $\mathbf{b}$, resulting in, as described in the references, estimated values $\tilde{b}_i(1)$ and hence $\tilde{v}_{\rm abs}(\mathbf{r}_i, z, 1)$ with a posterior uncertainty in $\tilde{b}_i(1)$, given by $\mathbf{P}(1)$. The argument, 1, is used to indicate that this estimate is the first of two that we will make of the reference-level velocity and to distinguish it from the prior uncertainty. The model noise is assigned a prior covariance $\mathbf{R}_{nn}$. The approximately 1600×1600 matrix $\mathbf{P}(0)$ was taken to be diagonal, with values of 1 cm^2 s^{-2} over the ocean interior, and larger values, approaching 100 cm^2 s^{-2}, in the western boundary currents. Owing to the limited information available in the inversion, $\tilde{\mathbf{b}}(1)$ is generally 'smooth'—that is, only its long-wavelength elements are estimated, with short wavelengths remaining in the solution nullspace. The specific spatial scales in $\tilde{\mathbf{b}}$ that are constrained are, however, a strong function of geography. For example, inversions that take the Drake Passage mass transport as accurately known produce values for $\tilde{\mathbf{b}}$ that are well determined on the horizontal scale of the Passage, O(800 km). In contrast, the requirement for mass conservation across the width of the Pacific Ocean determines well the corresponding elements, b_i, on a 10 000 km scale, leaving the 500 km scale poorly constrained. This regional variation in the scales of uncertainty of prior knowledge is an essential feature of oceanography (and of fluid flows more generally) and is an important complication in the combined use of altimetry and hydrography.

The numerical values of $\mathbf{P}(1)$ are very important in the following because the impact of the altimetry is measured by its ability to reduce significantly the solution uncertainty. Portions of two rows of $\mathbf{P}(1)$ are displayed in Fig. 2. Note that the variance at these interior ocean locations remains near the prior value of 1 cm^2 s^{-2} as a result of there being comparatively few (compared to the number of system unknowns) available constraints. However, as will be seen, the remaining errors near 1 cm^2 s^{-2} represent a benchmark against which to measure geoid accuracy.

From the hydrographic solution, the estimated absolute surface geostrophic velocity is

$$\tilde{v}_{\rm s,h}(\mathbf{r}_i, z=0, 1) \equiv v_{\rm rel}(\mathbf{r}_i, 0) + \tilde{b}_i(1) = \frac{g}{f} \left. \frac{\partial \tilde{\eta}_{\rm h}(\mathbf{r}_i, 1)}{\partial x} \right|_i , \tag{4}$$

where $\eta_{\rm h}$ is the surface elevation. Thus eq. (4) produces an estimate from the *in situ* data of the surface slope, $\partial \tilde{\eta}_{\rm h}(\mathbf{r}_i, 1)/\partial x$, that should be measured by the altimeter at position i. $\mathbf{P}(1)$ represents the contribution to the uncertainty of that slope owing to the reference-level velocity, but not any uncertainty arising from $v_{\rm rel}$, and this latter uncertainty is a central issue in the following.

The use of recursive inverse methods makes it possible to constrain further the solution of the original problem. Martel & Wunsch (1993a) built a finite-difference model and included constraints not just from hydrography, but also from current meters, floats and other observations. The model yielded a solution for the North Atlantic circulation that was

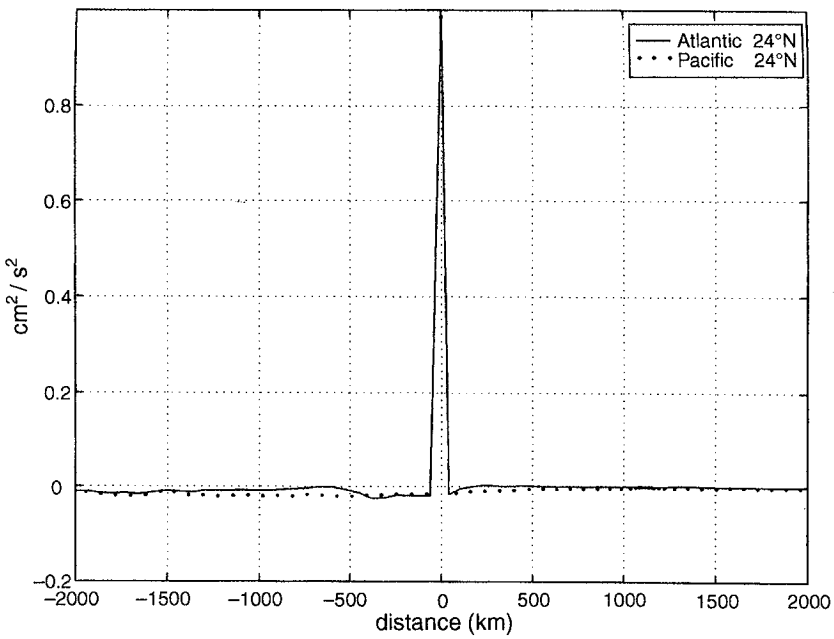

Figure 2. Portions of two rows of the error covariance from the inverse model $\mathbf{P}(1)$, one centred in the middle of the section at 24°N in the Atlantic, the other (dotted) from the middle of the section at the same latitude in the Pacific. Most of the structure lies near the central peak, whose magnitude is the error variance of a typical open ocean solution element.

compared in Martel & Wunsch (1993b) to the velocities from the GEOSAT altimeter. Within the very large uncertainties arising from the altimeter measurements and the crude geoid estimate being used, there was marginal agreement with the hydrographic estimate, but no further progress was possible.

The TOPEX/POSEIDON altimetric data and the estimated long-wavelength marine geoid associated with the JGM-3 gravity-field model (Tapley *et al.* 1996) are of much higher quality than have ever been available before. We therefore re-examine the idea of learning something of the absolute circulation by obtaining an improved estimate of the absolute ocean circulation by combining these data with the prior solution hydrography as discussed in the next section.

Before proceeding to the altimetry, it may be helpful to clarify the interpretation of the hydrographic estimate. At least three depictions of the large-scale ocean circulation can be discussed: (1) the instantaneous, absolute circulation; (2) the time-average circulation; and (3) the instantaneous-circulation anomaly measured relative to the time-averaged flow. (3) is not of direct concern here because time fluctuations can be determined independently of the geoid with altimetry. If the hydrography were a true instantaneous measure of the ocean temperature and salinity fields, then up to spatial coverage issues, the present prior estimate would represent (1). However with the available hydrography obtained from multidecadal samples, $\tilde{v}_{\mathrm{abs}}(\mathbf{r}_i, z, 1)$ cannot represent the true instantaneous absolute circulation unless the circulation is time-independent, which it manifestly is not. Similarly, the estimate cannot be considered a true time-averaged flow. How, then, should the flow be interpreted?

First, it must be recognized that understanding the meaning of flows inferred from non-instantaneous hydrography is a central issue in modern observational oceanography and is not peculiar to the use of altimetry. Altimetry does require a direct confrontation with the question of the meaning of the flow, which is otherwise often simply not discussed. There is, however, considerable evidence that flow fields integrated across ocean basins or straits such as that of Florida or the Drake Passage are stable over decades (or more precisely, that there are no statistically significant fluctuations visible with the existing data). We thus postulate that the *large-scale* components of $\tilde{v}_{\mathrm{abs}}(\mathbf{r}_i, z, 1)$ are, within some comparatively small error, a representation of both the multidecadal time average and the instantaneous absolute flow field. On shorter scales, $\tilde{v}_{\mathrm{abs}}(\mathbf{r}_i, z, 1)$ contains elements that are required to satisfy the constraints of the inversion, but whose detailed structure is ephemeral. Later in this paper, we will quantify the meaning of 'long' and 'short' scales.

One can usefully contrast the approach we are using here with what can be regarded as the preferred or ultimate method: existing oceanic general circulation models (GCMs) are capable of calculating complete 3-D temperature, salinity, pressure and flow fields from boundary conditions of estimates of atmospheric stress and buoyancy exchanges with the ocean. If further constrained to agreement, within error estimates, of *in situ* temperatures and salinities, one would have prior estimates of the complete oceanic fields at arbitrary locations, not restricted to the hydrographic lines. The reason we do not use this type of model is that there is no known practical way of computing quantitative estimates of the uncertainties of the resulting oceanic state estimates, which would include the sea-surface elevation. These models have state vectors of immense dimensions (exceeding 10^8 elements) and the governing equations are highly non-linear. We have opted instead for a linearized model for which an explicit state vector uncertainty covariance can be computed, but which restricts us to the hydrographic line positions.

There are also steady non-linear models (e.g. Mercier 1986; Wunsch 1994). These require that one partitions the surface elevation between changes in the deep reference-level velocity and those in the density field. On short scales (e.g. 500–1000 km and less), it is clear that changes in the density field are required, and probably dominate (see, for example, the solution in Wunsch 1994). However, on the transoceanic scales that predominate here, there is no evidence of significant temporal changes in the density field, and it is the deep pressure field that appears to require modification. In any event, to the extent that our model ocean is too simple, we will produce *a priori* oceanic state estimates with greater than necessary uncertainties, and will therefore tend to *overestimate* the impact of improved geoid estimates.

3 THE ALTIMETRIC ESTIMATE

A satellite-borne altimeter measures the shape of the sea surface relative to the spacecraft. Knowledge of the position of the spacecraft relative to the centre of the Earth produces an estimate of the absolute sea-surface shape $S(\phi, \lambda, t)$, where ϕ and λ are latitude and longitude respectively. The oceanographically important quantity is the deviation,

$$\eta(\phi, \lambda, t) = S(\phi, \lambda, t) - \Phi(\phi, \lambda), \tag{5}$$

where $\Phi(\phi, \lambda)$ is the gravitational/centrifugal equipotential called the geoid. Let the altimetrically determined value be denoted η_{alt} to distinguish it from the value implied by the hydrography in eq. (4). With η_{alt} known along a zonal line, in analogy to (4), altimetry provides an independent estimate of the absolute surface velocity:

$$v_{\mathrm{s,a}} = \frac{g}{f}\frac{\partial \eta_{\mathrm{alt}}}{\partial x}, \tag{6}$$

or, in vector form,

$$\tilde{\mathbf{v}}_{\mathrm{s,a}} = \begin{bmatrix} \tilde{u}_{\mathrm{s,a}} \\ \tilde{v}_{\mathrm{s,a}} \end{bmatrix} = \begin{bmatrix} (g/f)\partial \eta_{\mathrm{alt}}/\partial y \\ -(g/f)\partial \eta_{\mathrm{alt}}/\partial x \end{bmatrix}, \tag{7}$$

for the east and north components respectively (x and y are the local eastward and northward Cartesian coordinates, but in practice spherical coordinates are used). The interpolation to, and projection of components at, the station pair positions was carried out as described below for the computation of the geoid slope error.

Let $\tilde{v}_{\mathrm{s,a}}(\phi_i, \lambda_i, A_i)$ denote the altimetric estimate of the absolute surface geostrophic velocity normal to the the ith hydrographic station pair, where A_i is the local azimuth angle (see Fig. 3). Are the 1600 estimates $\tilde{\mathbf{v}}_{\mathrm{s,a}} = \{\tilde{v}_{\mathrm{s,a}}(\phi_i, \lambda_i, A_i)\}$ consistent with $\tilde{\mathbf{v}}_{\mathrm{s,h}}$ from the hydrographic estimate, and do they carry information capable of improving those estimates of $\mathbf{b}$? To answer both these questions, we must obtain an estimate, $\mathbf{R}_{\mathrm{aa}}$, of the covariance matrix of the errors in $\tilde{\mathbf{v}}_{\mathrm{s,a}}$.

The TOPEX/POSEIDON data set used is the 10 day gridded 2°×2° form described by King, Stammer & Wunsch (1994) and Stammer & Wunsch (1994). The resulting surface height, S, was further smoothed with a 4°×4° Shapiro (1970) filter. An estimate, $\tilde{\eta}_{\mathrm{alt}}$, of the mean absolute sea-surface height (relative to the JGM-3 geoid; see Tapley *et al.* 1997) is formed

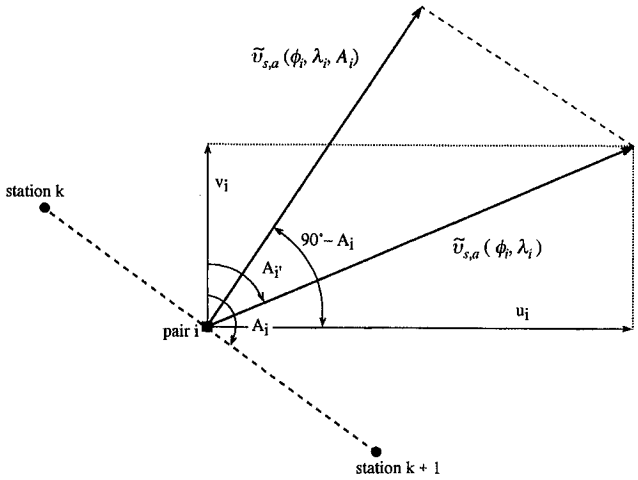

Figure 3. Normal projection of the velocity from altimetry, $\tilde{v}_{sa}(\phi_i, \lambda_i)$, onto the station pairs.

by averaging the time-series over a 2 yr (cycles 9–81) period. $\tilde{\eta}_{alt}$ is an estimate of the time-mean ocean dynamic topography, η, independent of the *in situ* data inversion.

3.1 The JGM-3 geoid slope error covariance

$\mathbf{R}_{aa}$ describes all of the errors in the determination of η_{alt} and hence of $\mathbf{v}_{s,a}$, and it is here assumed to be completely dominated by the uncertainty of the JGM-3 implied geoid. All other errors are represented by a small white-noise contribution. This simplifying assumption is justified by the dominance in the altimetric measurement error of the gravity field error (e.g. Tapley *et al.* 1994a; Tsaoussi & Koblinsky 1994). Although the accuracy of the TOPEX/POSEIDON determination of S is better than about 4 cm (Tapley *et al.* 1994b), the JGM-3 geoid has estimated errors over the ocean of about 36 cm at wavelengths represented by a spherical harmonic expansion to degree 70 (errors owing to omitted shorter space scales are accounted for separately).

Error covariances of gravity-field models have been studied frequently in terms of their propagation into geoid height estimates (Rapp 1993; Haagmans & Gelderen 1991). Because $\tilde{\eta}$ is of dynamical significance only through its derivatives, $\mathbf{R}_{aa}$ is proportional to the covariance of the geoid *slope*, not of the geoid itself.

Let δs_i be the geoid slope error between hydrographic station pair i, then $\langle \delta s_i \delta s_j \rangle$ is the error covariance in pairs i,j. A general procedure for computing this slope error covariance is summarized in Appendix A. Then,

$$\mathbf{R}_{aa} = \left\{ \frac{g_i}{f_i} \langle \delta s_i \delta s_j \rangle \frac{g_j}{f_j} \right\} \tag{8}$$

is the required covariance.

Because $\mathbf{R}_{aa}$ is the crucial element in determining the degree to which altimetry can improve knowledge of the ocean circulation, it is helpful to have some understanding of its structure. Fig. 4 shows that it is dominated by its diagonal and near-diagonal elements, but there are non-negligible covariances between altimetric velocity errors between different sections and even between different oceans. The diagonal elements are large (see Fig. 5 for JGM-3 at $N=70$), corresponding to root-mean-square (rms) geostrophic velocity errors between 25 and 60 cm s^{-1}, and rising to near 200 cm s^{-1} near the western boundaries. Much of the latitudinal structure arises from the increase in $1/f$ towards the equator.

The geoid slope uncertainty is a function of wavelength as well as geography. For a simple quantification of the JGM-3 geoid slope error, define the rms variance of the slope error over a specific geographic domain Λ:

$$\overline{\delta s}(\Lambda) = \sqrt{\frac{1}{\Lambda} \int_\Lambda \frac{1}{2} (\delta s_x^2 + \delta s_y^2)\, d\Lambda}, \tag{9}$$

where δs_x, δs_y denote the zonal and meridional components of the geoid slope error (not to be confused with the values δs_i defined between station pairs). Fig. 6 shows $\overline{\delta s}$ globally and over the ocean from 66°S to 66°N computed by varying the cut-off degree, N, of the spherical harmonic expansion between 0 and 70. Note the rapid increase in $\overline{\delta s}$ when $N > 25$.

The rows (or columns) of $\mathbf{R}_{aa}$ show a general pattern close to that of a truncated white-noise field (see Appendix B). Fig. 7(dots) displays a row centred at a station pair in the 24°N Pacific section and shows a quasi-periodicity close to 5°. This periodicity is mainly determined by the cut-off degree, N. With $N=70$ the shortest wavelength present in JGM-3 is approximately 40 000 km/70 ≈ 570 km (one might define the 'scale' as being one-quarter of the wavelength or about 150 km), but the implied error in the altimetric velocities is very large at the shortest wavelengths. In principle, one can use the $N=70$ representation of the geoid, because the information that the long wavelengths are more reliably determined than the short ones is contained in $\mathbf{R}_{aa}$. Numerically, however, difficulties arise in working with matrices whose elements range over many orders of magnitude; the errors accumulate as N increases. For the calculations described below, we have found it preferable to reduce the representation to $N=25$, and to recompute the corresponding $\mathbf{R}_{aa}$. The bold line in Fig. 7 displays the analogue for this reduced representation; now the oscillations occur over about 14° (1600 km wavelengths). The implied slope uncertainty is equivalent to geostrophic velocity errors of about 2–3 cm s^{-1} rms. Truncation at even lower degrees is possible, but then even fewer constraints on the ocean circulation become possible, and $N=25$ is a reasonable compromise at the present time.

4 COMBINING THE TWO DATA SETS

Relative to the JGM-3 geoid, then, the two-year average altimetric surface described above can be used to produce an estimate of $\mathbf{v}_{s,a}$ averaged over the two years December 1992–December 1994. Not only are there differences in the methodology used in obtaining $\mathbf{v}_a$ from that used in the hydrographic inversion, but one is comparing a two-year average in the early 1990s with the inhomogeneous result of the hydrography. That is, the hydrography and altimetry represent different oceans, unless the circulation was steady over the approximately 25 years spanned by all the data.

4.1 Comparison and errors

The first question to be asked is whether the altimeter and *in situ* inversion results for the surface geostrophic velocity are consistent with each other within the uncertainties of

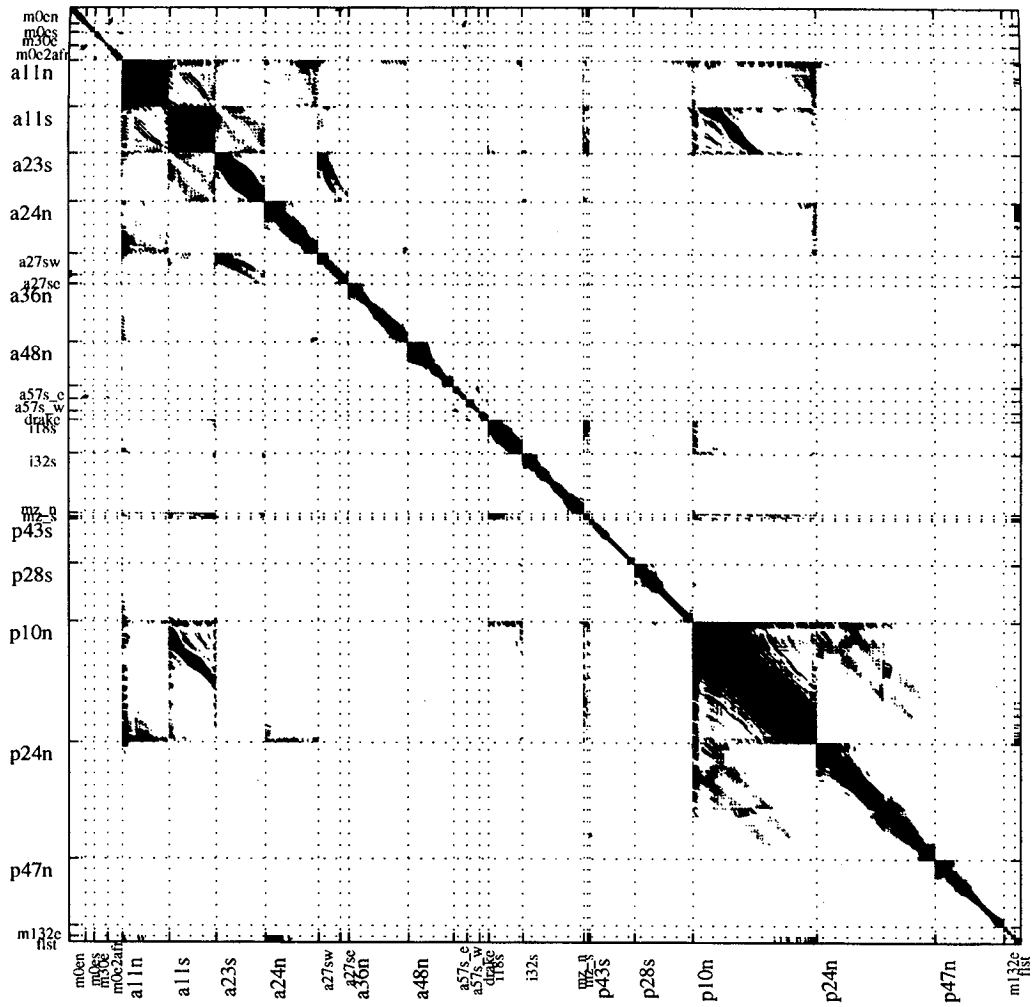

Figure 4. Structure of the error covariance matrix of the surface geostrophic velocity from altimetry under the assumption that the error is wholly due to the JGM-3 geoid slope error. Black dots indicate numerical values with a magnitude exceeding 50 (cm s^{-1})^2.

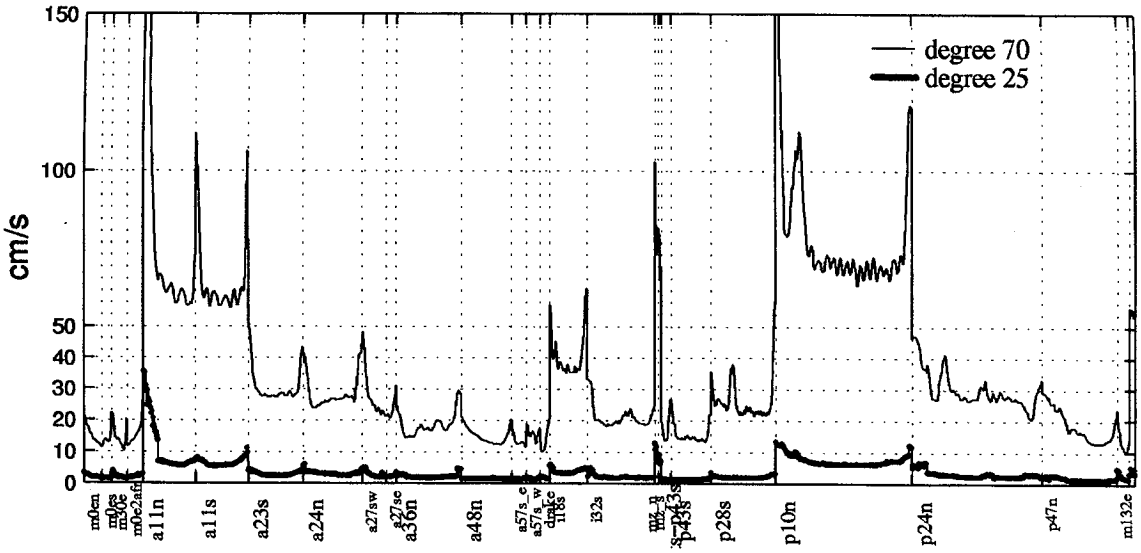

Figure 5. Square root of the variance of the surface velocity errors (1 μ radian ≈10 cm s^{-1} with $f \approx 10^{-4}$) obtained using the diagonal elements of the JGM-3 geoid slope error covariance matrix, for wavelengths longer than 500 km (degree 70, upper curve) and longer than 1600 km (degree 25, lower curve). Section names are listed on the abscissa. Lower-latitude sections are much more uncertain because of the decrease there of f.

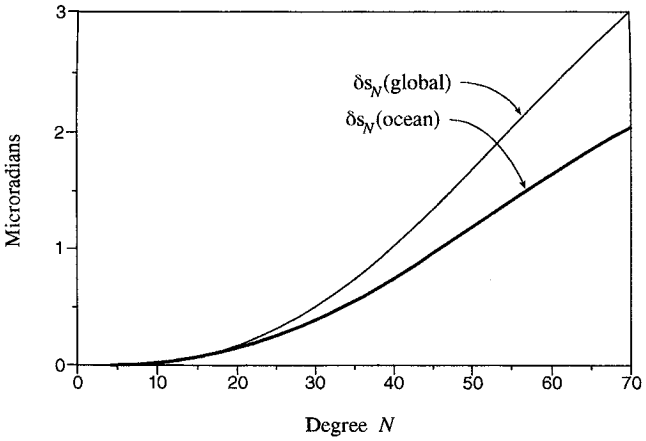

Figure 6. Square root of the variances of the JGM-3 geoid slope error calculated both globally and over the ocean only as a function of the cut-off degree in the spherical harmonic expansion.

each. A full answer is very complicated because, in addition to the temporal averaging question, each is represented by a 1600×1600 covariance matrix and the question should be answered scale-by-scale and position-by-position. We will proceed in two stages: first, we will compare point-by-point values and their uncertainty estimates, and then we will attempt to force the *in situ* model to agreement with the altimetry. The second step, if successful, simultaneously demonstrates consistency, and produces a new solution using the information from both *in situ* and altimetric data.

In making the comparisons we can directly compare the surface velocities $\tilde{\mathbf{v}}_{s,h}(1)=\mathbf{v}_R+\tilde{\mathbf{b}}(1)$ (from the hydrographic inversion) with $\tilde{\mathbf{v}}_{s,a}$, or we can compare $\tilde{\mathbf{b}}$ with $\tilde{\mathbf{v}}_{s,a}-\mathbf{v}_R$, i.e. the altimetric estimate of the reference-level velocity under the assumption that $\mathbf{v}_R$ was unchanged in time. If this assumption were correct, all elements were perfect, and the spatial resolution of the three velocities were identical, then in a totally consistent, complete (in the sense of adequate information) result,

$$\tilde{\mathbf{v}}_{s,a}-\mathbf{v}_R=\tilde{\mathbf{b}}(1). \tag{10}$$

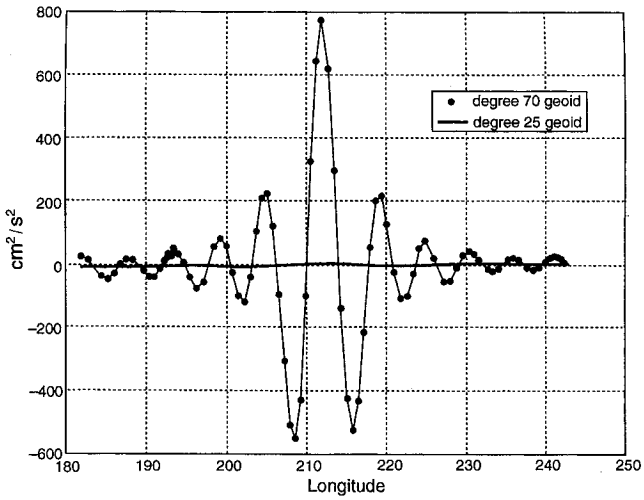

Figure 7. Portion of a typical row (or column) vector of $\mathbf{R}_{aa}$ for a zonal oceanographic section at 24°N in the Pacific based on the JGM-3 error covariance for degree 70 (dots) and degree 25 (bold line). Each dot represents a station pair location. The central point is located at the maximum. Compare to Fig. 2.

This equality fails for at least four reasons: (1) there is an expected error in $\tilde{\mathbf{b}}(1)$, with covariance $\mathbf{P}(1)$; (2) there is an error in $\tilde{\mathbf{v}}_{s,a}$, with covariance $\mathbf{R}_{aa}$; (3) $\mathbf{v}_R$ is not time-independent; and (4) the spatial resolution of $\mathbf{v}_R$ is that of the hydrographic station pairs, the resolution of $\tilde{\mathbf{v}}_{s,a}$ is given by the cutoff degree, N, and the resolution of $\tilde{\mathbf{b}}$ is that of the *in situ* inversion, which as noted above is spatially variable.

To accommodate point (3), let the covariance of the time-dependent portion of $\mathbf{v}_R$ be $\mathbf{R}_{tt}$. The covariance of the error in the left-hand side of (10) is then $\mathbf{R}_{aa}+\mathbf{R}_{tt}$, assuming no correlation between the errors in altimetry and in oceanic variability. Because the $\mathbf{R}_{aa}$, $\mathbf{R}_{tt}$ are additive, we could regard the temporal variability as an error in the altimetry. In this view, an accurate depiction of the 'hydrographic ocean' is the goal. An equally valid goal is the estimation of the ocean during the 2 yr altimetric period: then, $\mathbf{R}_{tt}$ would be an error in the hydrography as it represents the 'altimetric ocean'. Either point of view is correct, but we will adhere to the former.

Ideally, $\mathbf{v}_R$ should be the 2 yr mean relative velocity sampled in the same way as the satellite measurement $\tilde{\mathbf{v}}_{s,a}$. Because it is impossible to obtain such measurements, we took $\mathbf{v}_R$ from the hydrographic section data, thus introducing time-dependent eddy and other noise into the altimetric observations and $\mathbf{R}_{tt}$. The estimate (10) of the reference-level velocity is very noisy and some small-scale fluctuations from $\mathbf{v}_R$ can be as large as 100 cm s^{-1} (see Appendix C).

Specification of $\mathbf{R}_{tt}$ is a large, complex subject involving representation of oceanic variability on all space- and time-scales. Stammer (1997) used the along-track altimeter data over the interior of the ocean to compute an estimate of the global sea-surface slope spectrum, assuming isotropy. This slope spectrum is shown in Fig. 8(a); a frequency/wavenumber version will be displayed elsewhere. (The slope spectrum is employed because it can be readily computed globally, whereas the geostrophic velocity involves a division of the slope by f, producing an equatorial singularity.) The total slope variance is equivalent to a time-dependent geostrophic velocity variance of 150 cm^2 s^{-2} at 45° latitude. An important feature of this slope spectrum is the decay at large scales. In our case, it is an error estimate, meaning that the patterns in $\mathbf{v}_R$ are more accurate on large scales. The total variance for scales larger than 1600 km is less than 4 cm^2 s^{-2} at 45° latitude.

The interpretation of this value for comparison with altimetric noise involves at least two distinct problems. The spectrum (Fig. 8a) is computed from two years of data, and the hydrography spans 25 years. It is dangerous to extrapolate from the 2 yr frequency/wavenumber spectrum to that representing a non-existent 25 yr record. The frequency spectrum has a characteristic 'red' structure (Wunsch, in preparation) and at these scales, a 25 yr record would exhibit more energy than a 2 yr record, but how much more is not clear. From this point of view, we must interpret the 4 cm^2 s^{-2} as a lower bound.

On the other hand, it is incorrect to ascribe the entirety of this variability to the reference-level velocities, b_i, because doing so would imply that all oceanic variability was barotropic (in the rigorous sense of representing a velocity with no vertical structure, and completely unrelated to the density field). On the contrary, it appears (Wunsch, in preparation) that at long periods—longer than about 100 days—the bulk of oceanic velocity variability is dominated by baroclinic

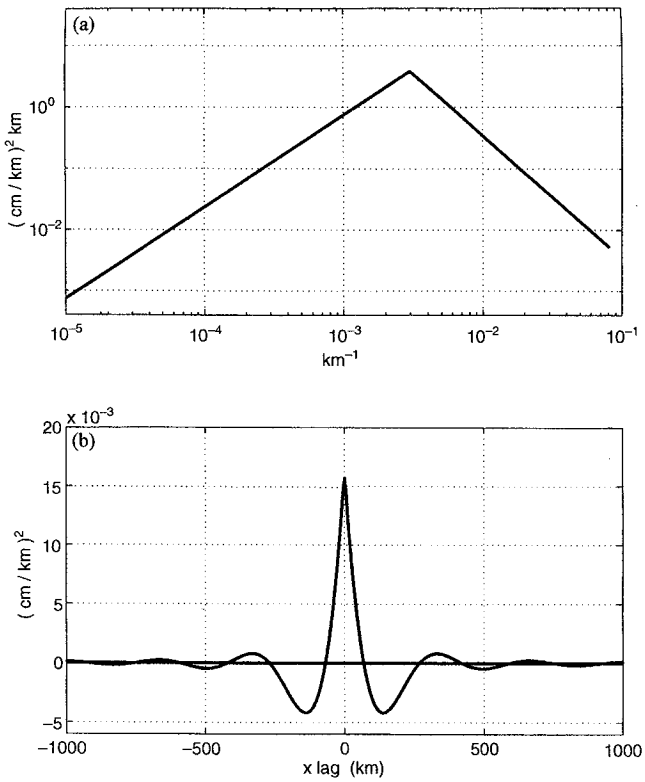

Figure 8. (a) Slope spectrum from Stammer (1997) of the surface variability. The total variance corresponds to 1.5×10^{-2} cm^2 km^{-2} or 150 cm^2 s^{-2} at 45° latitude. Energy decreases markedly at both short and long scales. (High wavenumber slope is not definitive; see Wunsch & Stammer 1995). (b) Covariance function, $\Gamma_s(x)$, deduced from the spectrum in (a).

motions, those directly associated with changes in the interior density structure. Because Macdonald's (1995) inversion is a linear one, there is no provision in it for modifying the initial density-field estimates (for a discussion of this problem, which is highly non-linear, see Wunsch 1994). From this viewpoint, the estimated variance of 4 cm^2 s^{-2} is an upper bound on the actual time variability of b_i over two years. Accepting this spectrum as having rough validity, we calculated the covariance of $\mathbf{v}_R$ as a function of the distance between points (Fig. 8b). The resulting covariance function, $\Gamma_s(\psi_{ij})$, was then interpolated between each pair to form the $\mathbf{R}_{tt}$ matrix,

$$\mathbf{R}_{tt}(i,j)=\left\{\frac{g_i}{f_i}\Gamma_s(\psi_{ij})\frac{g_j}{f_j}\right\}, \tag{11}$$

where $\psi(i,j)$ is the distance between station pair i and pair j.

Concerning point (4) above, there are two ways of representing the differences due to the differing spatial resolutions. We can either add to the various uncertainty matrices additional matrices representing the structures of the unresolved components, or we can represent the different fields as filtered versions of the true ones. We will employ a combination of these two approaches. First note that $\mathbf{P}(1)$ as computed by Macdonald (1995) includes the so-called null-space contribution, that is, those structures that were not resolved by the inversion, in contrast to the error covariances for the geoids which omit these contributions. Re-writing (10) as

$$\mathbf{y}(2)=\mathbf{F}(\tilde{\mathbf{v}}_{s,a}-\mathbf{v}_R)+\mathbf{n}=\mathbf{F}\mathbf{b}+\mathbf{n}, \tag{12}$$

where $\mathbf{F}$ is a matrix representing the filtering operation that the altimetric measurement applies to the surface velocity, one obtains

$$\langle\mathbf{n}\mathbf{n}^T\rangle=\mathbf{F}(\mathbf{R}_{aa}+\mathbf{R}_{tt})\mathbf{F}^T\equiv\mathbf{R}(2).$$

The filtering operation in (12) is slightly complicated. The point altimetric values represent a *spatial* average over about 5° for $N=70$, and over about 14° for $N=25$, defining the averaging distance in terms of the wavelength. These scales are much larger than the station spacings in the hydrographic lines, and a comparison of altimetric and hydrographic-inversion estimates can only be made if the hydrographic velocities are averaged equivalently. At best we can average the $\tilde{\mathbf{v}}_{sh}$, etc. only linearly along the hydrographic lines. Another minor issue is that because hydrographic stations are not uniformly spaced, any filtering carried out of $\tilde{\mathbf{v}}_{sh}$, etc. will be done with non-uniformly spaced velocity elements (Appendix B).

Let $S(\theta,\lambda)$ be the correct absolute sea-surface elevation at colatitude θ, and longitude λ. A truncated spherical harmonic expansion (e.g. Wunsch 1991) is then the weighted average

$$\bar{S}(\theta,\lambda)=\int_0^{2\pi}\int_0^{\pi}W(\theta,\lambda,\theta',\lambda')S(\theta',\lambda')\sin\theta'd\theta'd\lambda', \tag{13}$$

where the averaging kernel W is approximately

$$W(\theta,\lambda,\theta',\lambda')=\sum_{n=0}^{n=N}\sum_{m=-n}^{m=n}Y_n^m(\theta',\lambda')Y_n^m(\theta,\lambda)^*. \tag{14}$$

Y_n^m are the fully normalized spherical harmonics effectively removing wavelengths shorter than $\gamma_c=360°/N$.

To what extent can the spatial average (13) be expected to agree with the linear average implicit in $\tilde{\mathbf{b}}(1)$? If the ocean were homogeneous and isotropic, one could answer this question in the mean square with some elementary statistical calculations. In practice, homogeneity and isotropy demonstrably fail. To obtain at least a sense of the deviations expected, we took the TOPEX/POSEIDON data from a 2°×2° set of gridded values [formed from the along-track data as described by Stammer & Wunsch (1994)]. From these data, a set of gridded velocities, $\mathbf{v}_{s,a}-\mathbf{v}_R$, were computed along the hydrographic lines and taken to be representative of the type of structure present in $\tilde{\mathbf{b}}$, which as noted above is primarily long wavelength in nature. These surface-velocity values were then interpolated at each station pair and filtered along the sections with the filter $\mathbf{F}$ set to remove wavelengths shorter than 360/25° (1600 km). We assigned this signal to a one-dimensionally filtered long-wavelength velocity. On the other hand, the *same* velocities from this 2°×2° grid were spatially filtered in two dimensions by using the truncated spherical harmonic expansion for $N=25$. These latter velocities were identically interpolated at the station pairs and compared to the former one-dimensionally averaged ones. Fig. 9 shows the result of this comparison for the zonal section Pacific 28°S and for the meridional section 132°E. Within the limits of present data accuracy, it appears that the linear and spatial averages are indistinguishable.

For each of the sections, we compared the altimetric estimate, $\mathbf{y}(2)$ to the *in situ* estimate $\mathbf{F}\tilde{\mathbf{b}}(1)=\mathbf{E}(2)\tilde{\mathbf{b}}(1)$, as shown in Fig. 10 for the zonal Pacific 28°S section and the

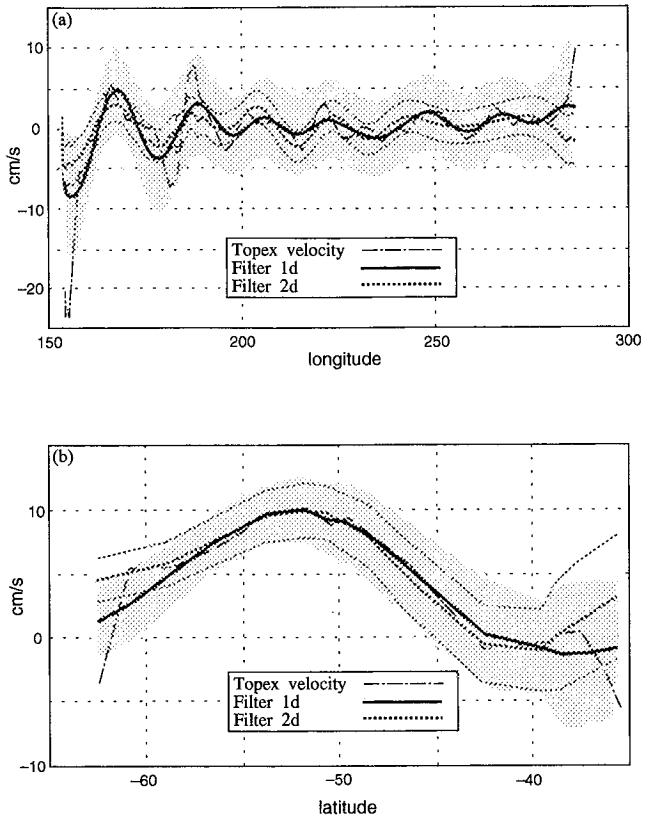

Figure 9. Comparison of TOPEX/POSEIDON-derived velocities between the along-section (1-D) filtered version (bold line, with uncertainties in the shaded region) with those from the spherical-harmonic (2-D) filter (dotted lines are the uncertainty). Results are for (a) the zonal section at 28°S in the Pacific and (b) the meridional 132°E section south of Australia.

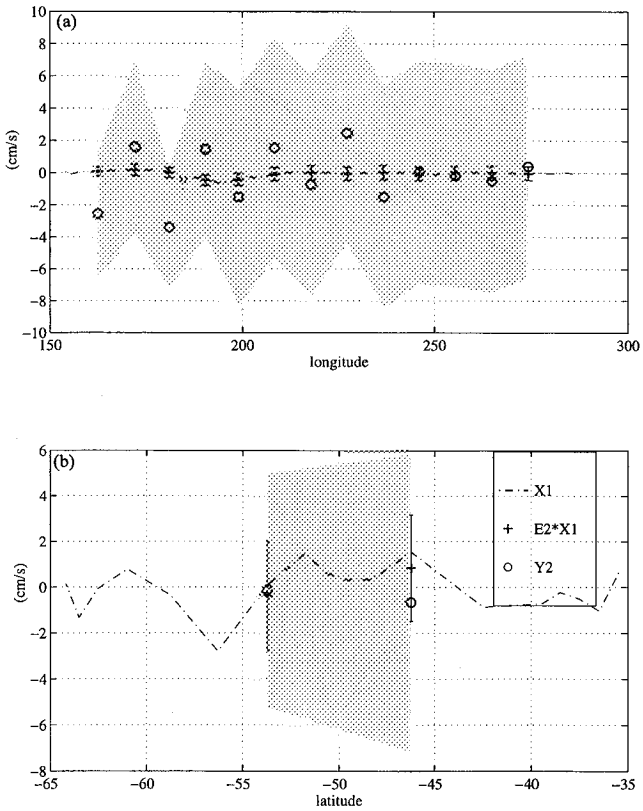

Figure 10. Reference-level velocities in (a) the zonal Pacific 28°S section and (b) in the Southern Ocean meridional 132°E section. Circles are the observations, $\mathbf{y}(2)$ from the TOPEX/POSEIDON data and crosses are the estimates from the hydrographic model, $\mathbf{E}(2)\tilde{\mathbf{b}}(1)=\mathbf{F}\tilde{\mathbf{b}}(1)$. Both data were filtered with a low-pass 14° (1600 km) filter along the section. The uncertainty in $\mathbf{y}(2)$ lies in the shaded region, which almost always includes the hydrographic estimates.

Southern Ocean meridional section at 132°E. Here, as with the remaining sections, the two estimates agree almost everywhere within one standard deviation of the two uncertainties. The satellite data tend to show larger poleward velocities near the western boundaries; it is not entirely clear whether the *in situ* model is underestimating the velocity field there, or whether there is some edge effect due to the different averaging types.

4.2 Recursive improvement

Having demonstrated at least rough statistical consistency of the two estimates of the reference-level velocity, we now turn to the problem of combining them. We use a form of recursive estimation (see e.g. Wunsch 1996) in which the new estimate of the reference level velocity is $\tilde{\mathbf{b}}(2)$:

$$\tilde{\mathbf{b}}(2)=\tilde{\mathbf{b}}(1)+\mathbf{K}(2)[\mathbf{y}(2)-\mathbf{E}(2)\tilde{\mathbf{b}}(1)], \tag{15}$$

$$\mathbf{P}(2)=\mathbf{P}(1)-\mathbf{K}(2)\mathbf{E}(2)\mathbf{P}(1), \tag{16}$$

$$\mathbf{K}(2)=\mathbf{P}(1)\mathbf{E}(2)^{\mathrm{T}}[\mathbf{E}(2)\mathbf{P}(1)\mathbf{E}(2)^{\mathrm{T}}+\mathbf{R}(2)]^{-1}. \tag{17}$$

As already noted, the uncertainty of $\tilde{\mathbf{b}}(1)$, that is the diagonal elements of $\mathbf{P}(1)$, ranges from less than 1 cm^2 s^{-2} for most of the ocean to 100 cm^2 s^{-2} in the western boundary current and circumpolar current regions. Furthermore, various integral constraints on $\tilde{\mathbf{b}}(1)$ are formally accurate to the equivalent of about 2 Sv. It is these generally small estimated error variances with which the added altimetric constraints must 'compete'.

Because for reasons of numerical accuracy we chose to work with the geoid carried only to degree and order 25, the $\mathbf{E}(2)$ matrix is thus the low-pass filter $\mathbf{F}$ with cut-off at $\gamma_c \approx 360/25 = 14.4° = 1600$ km. To reduce the computational load, we divided the sections (Fig. 1) into large regions (North Pacific, North Atlantic, etc.) and calculated separately $\tilde{\mathbf{b}}(2)$ and $\mathbf{P}(2)$ over those regions, losing the correlations in the errors between the different regions. Those correlations are, however, weak, and this truncation has no noticeable influence on the results. The end result is a new estimate of the flow fields and their transports which is found to be consistent with both the prior model and the altimetric constraints. (Strict mass-transport constraints in the Southern Ocean are, however, lost, as they depend upon the interregion correlations.)

The norm of $\mathbf{K}(2)$, which controls the magnitude of the change made to $\tilde{\mathbf{b}}(1)$, is dependent on the estimated noise covariance, $\mathbf{R}(2)$, as well as on the structure of $\mathbf{P}(1)$; that is, the large diagonal elements of $\mathbf{P}(1)$ do not by themselves imply that the norm of $\mathbf{K}(2)$ will be small; rather, the structure of the matrix is involved through the product $\mathbf{E}(2)\mathbf{P}(1)\mathbf{E}(2)^{\mathrm{T}}$. In practice, however, the changes implied in the diagonal elements of $\mathbf{P}(1)$ by eqs (15) and (17) are less than 3 per cent almost everywhere. Exceptions occur, for example in the

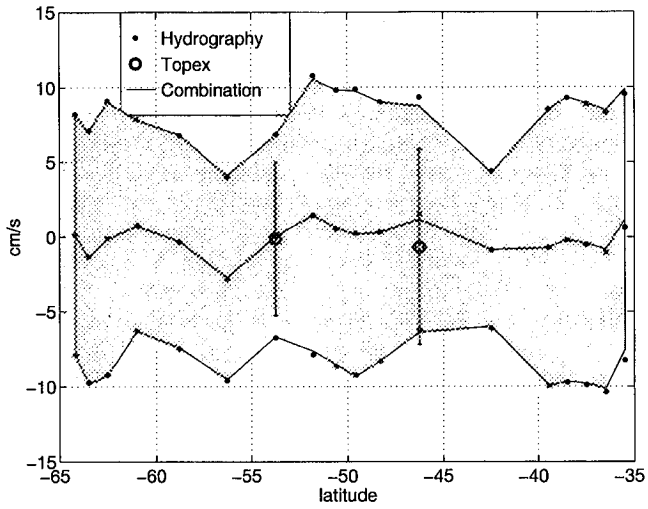

Figure 11. Influence of the TOPEX/POSEIDON altimetric data on the reference-level velocities. The 132°E section is more heavily influenced by the altimetry because it is initially one of the most uncertain sections in the hydrographic results; other sections are only weakly influenced. Dots are the solution $\tilde{\mathbf{b}}(1)$ and the standard error from the hydrographic model. Circles are the observations $\mathbf{y}(2)$ from TOPEX/POSEIDON with their standard error. The solid line shows the combined solution with the uncertainty given by the shaded area. The shaded area extends beyond the TOPEX/POSEIDON uncertainty because the total uncertainty includes short scales not represented by the altimetry.

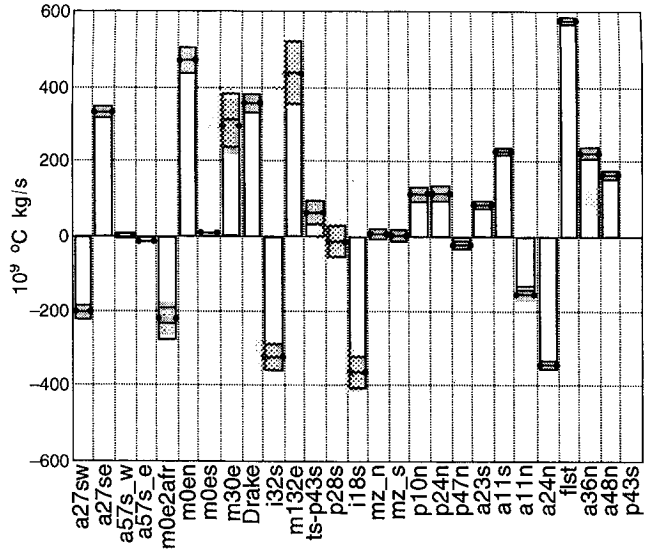

Figure 12. Temperature fluxes and uncertainties for each section. The results from the hydrographic inversion alone are denoted by a pair of dots, with an uncertainty given by the shaded area. The combined result is given by a horizontal line segment, which is often visually indistinguishable from the hydrographic result (an exception is e.g. along section m30e). The uncertainty of the combined solution is given by two horizontal line segments on either side of the solution; these again are often indistinguishable from those of the pure hydrographic solution. Note that on this panel the flux across section A11N does not account for the North Brazil current; the total value with the Brazil current included is close to $+330\times10^9$ °C kg s^{-1}.

circumpolar current (see Fig. 11), where 6 per cent reductions are made by the altimetry reflecting the generally poorer resolution there in the purely hydrographic inversion.

One must be careful, however, before concluding that the altimetry is adding little to the prior estimate: property transports, e.g. of mass and heat, are integral properties of the flow field, and thus sensitive to long-wavelength information in the reference-level velocity. The combined result can, in principle, produce changes in integrated transports, even though the changes in any given element b_i are quite small. To understand this possibility one can either examine the structure of $\mathbf{P}(2)$ relative to that of $\mathbf{P}(1)$, or, as we have chosen to do, compare the property transports before and after the addition of the altimetry.

5 TRANSPORT SHIFTS

Transports of mass, temperature, salinity, oxygen, silicate and phosphate and their respective uncertainties associated with solutions $\tilde{\mathbf{b}}(1)$ and $\tilde{\mathbf{b}}(2)$ were computed. The changes in the estimated transports are generally less than 1 per cent everywhere, except along the sections 30°E, 132°E and 0°E (that from the Cape of Good Hope), all crossing the Antarctic circumpolar current, which had the most uncertain $\tilde{\mathbf{b}}(1)$. In these places, the transports of some of the properties (including temperature) are modified by up to 7 per cent and their uncertainties decrease by up to 5 per cent, but the net mass transport is very little changed, probably because of the small errors in the mass conservation constraints that determine $\mathbf{b}(1)$. The box residuals remain relatively unchanged. In Fig. 12 the temperature fluxes for each section and the changes induced by the TOPEX/POSEIDON data are shown. Table 1 summarizes the heat-flux uncertainty change over large regions.

The conclusions from these calculation are then two-fold and straightforward: (1) there is no inconsistency of estimates of the surface velocities $\mathbf{v}_{s,h}(1)$ and $\mathbf{v}_{s,a}$ within their respective uncertainty estimates, and (2) the additional information provided to the hydrographic inversion by the absolute altimetry makes no qualitative, and only a small quantitative, change in the solution.

Although this result might be regarded as disappointing, it is in many ways a major achievement: altimetry and marine geodesy have progressed in the past 15 years from a situation in which the implied errors in absolute sea-surface estimates relative to the geoid had to be measured in metres, to one in which the errors are a few centimetres, an improvement of two orders of magnitude. Furthermore, two entirely different methods for estimating surface topography, hydrographic inversions and absolute altimetry, have been brought into agreement at an accuracy of a few centimetres per second. The issue now is whether one more order-of-magnitude

Table 1. Percentage reduction in the heat-flux uncertainty using various assumed geoid slope error covariances. Results are grouped by region. JGM-3 is the present geoid estimate used for TOPEX/POSEIDON analysis; the Sharma (1995) degree 25 result is more conservative than the Bettadpur (private communication, 1996) computation at degrees 25 and 70.

Region/Geoid	N. Atl.	N. Pac.	S. Pac./Ind.	ACC
JGM-3	0	0	0	1–6
Sharma 25	1–10	5–30	15–35	20–40
Bettadpur 25	1–20	8–35	20–45	20–40
Bettadpur 70	1–30	15–50	30–60	30–60

improvement in the altimetric accuracies can be achieved in the next 15 years.

6 IMPROVED GEOID ACCURACY

We turn, next, to the important question of the impact on estimates of the ocean circulation of potential improvements in the geoid estimates from any of a variety of hypothetical spaceborne gravity missions. Such missions have been the focus of discussion for more than 20 years, a discussion that has become more important with the current availability of the high-accuracy TOPEX/POSEIDON data. Recent discussions include NASA (1987), Bettadpur (1993), Elema (1993), Sharma (1995), Rummel, Sneeuw & Mueller (1995), and Balmino *et al.* (1996).

Some comment is also required concerning our particular measure of the impact of more accurate geoid data. We have chosen to study improvements in the uncertainty of integrated mass and temperature (heat) fluxes. Because the ocean circulation is important for a vast variety of phenomena, one can easily imagine focusing on a different goal (the flux of carbon in the East Greenland current, to pick an arbitrary example) and for which the resulting impact of an improved geoid might be very different. Such choices of goals are, however, inevitable in science and we make no apology for our particular figures of merit: the heat and mass fluxes are central elements in the important problem of understanding the global climate. Other investigators with other goals are urged to study similarly the impact of improved geoids. Their conclusions may well be different from ours.

In attempting to determine the degree of the improvement in knowledge of the ocean circulation that would result from an improvement in geoid accuracy, one must be careful to separate the impact of errors occuring in the *in situ* inversion that are intrinsic from those that are accidents of the particular database or method being used. A complete and clear answer to the question posed is impossible both because full error analyses of the potential gravity measurements are not available, and because one can only speculate as to the characteristics of future *in situ* oceanic observations and their analyses. As has been stated above, Macdonald's solution was necessarily based upon a hydrographic survey spanning 25 years. This property of the hydrography introduces the time-dependent error, with covariance $\mathbf{R}_{tt}$, and provides a limit on the utility of even a perfect geoid if the goal is to estimate the ocean over those 25 years. If one worked instead with a hydrographic survey that was essentially instantaneous (defined to be conducted over order 10 days, the time required for TOPEX/POSEIDON to sample the ocean once), then the elements of the long-wavelength time-dependent error $\mathbf{R}_{tt}(2)=\mathbf{F}\mathbf{R}_{tt}\mathbf{F}^{T}$ would be much smaller. Such rapid surveys of single transoceanic lines are today possible in principle with existing fast-profiling devices. Alternatively, and perhaps more realistically, ocean general circulation models are capable of producing global instantaneous values, and the ultimate goal of our exercise is to apply the altimetric constraints to such models. We will suppose, then, that $\mathbf{R}_{tt}$ (and $\mathbf{F}\mathbf{R}_{tt}\mathbf{F}^{T}$) will in the future have much reduced norms, becoming effectively vanishingly small. (We ignore the important practical obstacle, discussed above, that existing general circulation model results are not accompanied by uncertainty estimates, and to proceed we are postulating that they are comparable to those from the *in situ* inversion.)

Sharma (1995) produced an estimate of the geoid error that would result from a hypothetical low–low satellite-to-satellite tracking mission. Further details are given in Appendix D. The result from this simulation showed a significantly improved recovery of the geopotential compared to the JGM-3 field. This estimate was converted into the equivalent error covariance matrix $\mathbf{R}_{aa}$ at $N=25$ as used above, and the property transport fluxes recomputed. Now the rms geoid error to this degree corresponds to about 0.15 cm s^{-1} at 45° (see Fig. 13). $\mathbf{R}_{tt}$ was greatly reduced, to an equivalent rms error of about (0.02 cm s^{-1})^2 at 45° latitude, and is no longer limiting. In the South Pacific, Indian and Southern oceans, the mass-flux uncertainties are reduced by about 15 per cent, with heat-flux uncertainty reductions of 30–40 per cent. Elsewhere, there are only slight mass-flux error reductions (less than 4 per cent), while the North Pacific heat-flux uncertainty is reduced by about 30 per cent. In contrast, heat-flux uncertainties in the North Atlantic drop by between 1 and 8 per cent, presumably reflecting the much smaller scale of this basin.

Sharma's analysis is being repeated (S. V. Bettadpur, private communication, 1996) under the assumption that both satellites are tracked by the Global Positioning System (GPS) and with an increased mission lifetime. A preliminary quantification of Bettadpur's analysis is illustrated in Fig. 13 in terms of the geostrophic velocity error. Sharma's result was evidently conservative and it is plausible that a real mission outcome will be much more accurate than he found. Assuming Bettadpur's result is correct, $\mathbf{R}_{tt}$ as used above becomes the limiting factor. To explore the best possible outcome,we set $\mathbf{R}_{tt}$ to zero, along with all other errors, with outcome displayed in Table 1.

At degree 70, Bettadpur's analysis produces a reduction in the heat-flux uncertainty by a factor of two. The remaining errors arise from the short scales that are not resolved by the degree 70 geoid. To render this residual more concrete, one row of the error covariance matrix was Fourier transformed to produce an approximate wavenumber spectral estimate, as depicted in Fig. 14. (The row corresponds to the covariance of a station pair at 18°S in the centre of the Indian Ocean.) The spectral density is shown for: (1) the initial hydrographic uncertainty, $\mathbf{P}(1)$, (2) the uncertainty after improvement with

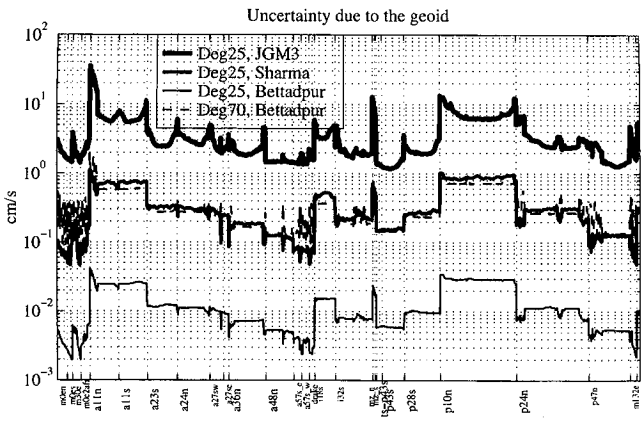

Figure 13. Root-mean-square geoid error from JGM-3 and a hypothetical gravity mission. Shown is the geostrophic velocity error implied at each station pair for the existing JGM-3 geoid (very heavy line), Sharma's simulation at degree 25 (bold line), Bettadpur's simulation at degree 25 (normal line) and at degree 70 (dashed line)

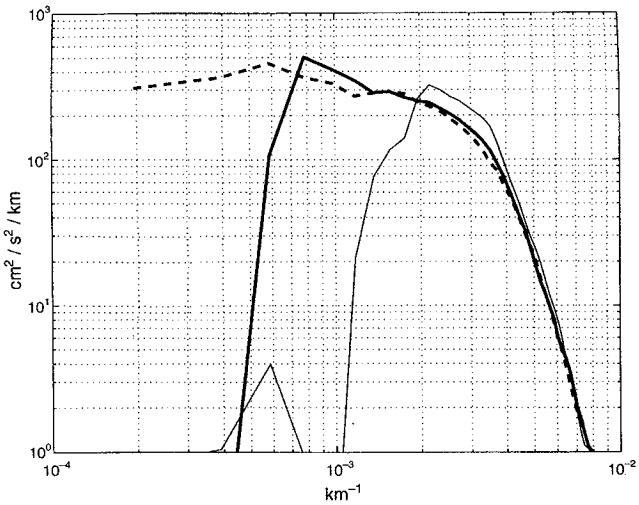

Figure 14. Approximate reference-level velocity error spectral estimates: from the hydrographic model (dashed line), after improvement with Bettadpur's geoid at degree 25 (heavy line) and at degree 70 (light solid line). The considerable remaining energy at high wavenumbers is the origin of the final uncertainty in the oceanic heat fluxes. Estimated by Fourier transforming one row of $\mathbf{P}(2)$ corresponding to a covariance along 18°S in the Indian Ocean.

Bettadpur's estimate at degree 25 and (3) at degree 70. The accurate geoids eliminate almost all the errors for scales longer than the cut-off wavelength, but leave the shorter scales untouched. Because oceanic heat fluxes remain sensitive to correlation between these short-scale flow fields and the temperature field, there is a considerable uncertainty remaining in the computations.

Other estimates exist for error budgets of different gravity-mission designs (Rummel *et al.* 1995; Balmino *et al.* 1996; B. Bills, private communication, 1996); however, no slope-error covariance has been made available to us for these other missions. A simple final experiment was therefore attempted in which $\mathbf{R}_{aa}$ was given a purely white-noise structure with total variance of 3 cm^2 s^{-2} at 45° spread uniformly across all scales down to the station spacing. Heat-flux errors decreased between about 20 per cent and 60 per cent in the circumpolar sections, but only slightly elsewhere. If, as has been suggested, realistic gravity missions could do far better than the equivalent of 3 cm^2 s^{-2}, then much larger error decreases are evidently possible. Until complete error budgets for such missions are available, there seems little point at present in pursuing these calculations further.

7 SUMMARY AND CONCLUSIONS

We draw three simple conclusions from this lengthy calculation:

(1) The error budgets for JGM-3 and for the *in situ* inversion of Macdonald (1995) demonstrate that the geoid estimate is consistent with what is known of the large-scale general circulation;

(2) the present accuracies of JGM-3 are inadequate to improve significantly the *in situ* result, even one using a simple, if reasonable level of no motion (Macdonald's starting point);

(3) proposed gravity missions, in so far as they quantitatively resemble the one for which we have had adequate information to study the impact, will improve existing estimates of mass and heat fluxes, and should reduce the largest existing uncertainties by about a factor of two.

With the remaining errors being attributed to smaller scales, the impact of higher-order geoids may be limited by the time variability as it is maximum on scales of a few hundred kilometres (Fig. 8). The most likely way that these errors can be removed is by the use of simultaneous hydrographic and altimetric measurements over large (but non-global) regions.

The last point requires further exploration. The *in situ* inversion was carried out under the *a priori* assumption that average deep-ocean velocities do not exceed about 1 cm s^{-1} rms over all scales, down to the station spacing, (near 50 km on average). Furthermore, Macdonald (1995) estimated the formal mass transport errors of her result as approximately 2×10^9 kg s^{-1} (2 Sv), which places a strong constraint on the integral properties of this deep flow field. This uncertainty may be somewhat optimistic, but is surely accurate to within an order of magnitude. Choose any specific horizontal scale, L, over which the general circulation is believed known already to within 2 Sv. To reduce this uncertainty signficantly, the geoid slope error would have to be sufficiently small that the implied transport error is less than 2 Sv. To be specific, suppose the mean section water depth is $h = 5000$ m (a bit large), at latitude 30°. Two Sverdrups of transport corresponds to a surface elevation change of about 0.3 cm, and this sets a threshold for significant improvement in the uncertainties of the *in situ* inversion by any possible gravity mission. (It is a peculiarity of geostrophic balance that the lateral distance, Δx, over which the elevation changes is irrelevant: the velocity corresponding to an elevation change $\Delta\eta$ is proportional to $\Delta\eta/\Delta x$, and so the corresponding transport is proportional to the velocity multiplied by the cross-sectional area, or $\Delta\eta h$.) If the true circulation uncertainty were closer to 20 Sv, the corresponding requirement on useful surface-slope measurements would be reduced accordingly.

The alternative approach to synthesizing and using the gravity/altimetry is, as alluded to in the Introduction, to employ full general circulation models capable of computing instantaneous oceanic states from realistic surface forcing functions. Here the issue remains of the general difficulty of producing useful model error covariances; learning how to do so will be a major effort for the next several years.

It is not the purpose of this paper to draw conclusions about the benefits of any particular gravity measurement programme, but only to provide some perspective for those who must make decisions about such missions and to provide a methodology for exploiting future data streams. Evaluation of any particular measurement strategy has to occur within the context of existing knowledge.

However, this discussion reinforces an old idea—that estimates from *in situ* data and the dynamics of the ocean circulation imply a geoid estimate. Thus one strategy for *marine* geoid improvement is to construct better estimates of the ocean circulation through improvements in the dynamical models and the *in situ* database. Subtraction of the implied sea surface from the altimetric sea-surface (eq. 1) produces a geoid estimate, and indeed we have made such estimates using ocean general circulation numerical models. More generally, one exploits all existing knowledge, whatever its origin, of the

ocean circulation and the marine gravity field; the best geoid estimate thus requires a full symbiosis of marine geodesy and physical oceanography.

ACKNOWLEDGMENTS

Supported in part by the National Aeronautics and Space Administration (Grants NAGW-918) and through the Jet Propulsion Laboratory (Contracts 958125). We had discussions with and assistance from A. Macdonald concerning the inverse model, D. Stammer about the $\mathbf{R}_{tt}$ matrix, and useful comments from D. Stammer, R. Rummel, R. Rapp and J.-F. Minster.

REFERENCES

Abramowitz, M. & Stegun, I.A., 1968. *Handbook of Mathematical Functions,* Dover, New York, NY.

Balmino, G., Sabadini, R., Tscherning, C. & Woodworth, P., 1996. *Gravity-field and steady-state ocean circulation mission,* ESA SP-1196(1), European Space Agency Publications Division, Noordwijk, the Netherlands.

Bennett, A.F., 1992. *Inverse Methods in Physical Oceanography,* Cambridge University Press, Cambridge.

Bettadpur, S.V., 1993. A simulation study of high degree and order geopotential determination using satellite gravity gradiometry, *PhD thesis,* Department of Aerospace Engineering and Engineering Mechanics, University of Texas, Austin, TX.

Bracewell, R.N., 1978. *The Fourier Transform and its Applications,* McGraw-Hill, New-York, NY.

Elema, I.A., 1993. Influence of geoid model uncertainty on the determination of ocean circulation with satellite altimetry, *PhD thesis,* Technical University Delft, Faculteit der Geodesie, the Netherlands.

Haagmans, R.H.N. & van Gelderen, M., 1991. Error variances–covariances of GEM-T1: their characteristics and implications in geoid computation, *J. geophys. Res.,* **96,** 20 011–20 022

Heiskanen, W.A. & Moritz, H., 1967. *Physical Geodesy,* W.H. Freeman, San Francisco, CA.

King, C., Stammer, D. & Wunsch, C., 1994. *The CMPO/MIT TOPEX/POSEIDON Altimetric Data Set,* Report No. 30, Center for Global Climate Change Science, Massachusetts Institute of Technology, MA.

Macdonald, A., 1995. Oceanic fluxes of mass, heat and freshwater: a global estimate and perspective, *PhD thesis,* Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA.

Macdonald, A.R. & Wunsch, C., 1996. An estimate of the global ocean circulation and heat flux, *Nature,* **382,** 436–439.

Martel, F. & Wunsch, C., 1993a. The North Atlantic circulation, in the early 1980s—An estimate from inversion of a finite difference model, *J. Phys. Oceanogr.,* **23,** 898–924.

Martel, F. & Wunsch, C., 1993b. Combined inversion of hydrography, current meter data and altimetric elevations for the North Atlantic circulation, *Manuscripta Geodaetica,* **18,** 219–266.

Mercier, H., 1986. Determining the general circulation of the ocean: a non-linear inverse problem, *J. geophys. Res.,* **91,** 5103–5109.

National Aeronautics and Space Administration, 1987. *Geophysical and Geodetic Requirements for Global Gravity Field Measurements, 1987–2000,* Division of Earth Sciences and Applications, Geodynamics Branch, Report of the Gravity Workshop, February 1987, Goddard Space Flight Center, Maryland, USA.

Rapp, R.H., 1993. Geoid undulation accuracy, *IEEE Trans. Geoscience and Remote Sensing,* **31,** 365–370.

Rummel, R., Sneeuw, N. & Mueller, J., 1995. *Geodetic Requirements and Prospects. Study of Gravity Explorer Mission Requirements (A Simulation Study),* Study for Daimler–Benz Aerospace, Dornier, Stuttgart, Germany.

Shapiro, R., 1970. Smoothing, filtering, and boundary effects, *Rev. Geophys.,* **8,** 358–387.

Sharma, J., 1995. Precise determination of the geopotential with a low–low satellite-to-satellite tracking mission, *PhD thesis,* Department of Aerospace Engineering and Engineering Mechanics, University of Texas, Austin, TX.

Stammer, D., 1997. Global characteristics of ocean variability from regional TOPEX/POSEIDON altimeter measurements, *J. Phys. Oceanogr,* in press.

Stammer, D. & Wunsch, C., 1994. Preliminary assessment of the accuracy and precision of the TOPEX/POSEIDON altimeter data with respect to the large scale ocean circulation, *J. geophys. Res.,* **99,** 24 584–24 604.

Tapley, B.D., Chambers, D.P., Shum, C.K., Eanes, R.J. & Ries, J.C., 1994a. Accuracy assessment of the large-scale dynamic ocean topography from TOPEX/POSEIDON altimeter data with respect to the large scale ocean circulation, *J. geophys. Res.,* **99,** 24 605–24 617.

Tapley, B.D. *et al.,* 1994b. Precision orbit determination for TOPEX/POSEIDON, *J. geophys. Res.,* **99,** 24 383–24 404.

Tapley, B.D. *et al.,* 1997. The JGM-3 gravity model, *J. geophys. Res.,* in press.

TOPEX Science Working Group, 1981. *Satellite Altimetric Measurements of the Ocean,* Document 400-111, Jet Propulsion Laboratory, Pasadena, CA.

Tsaoussi, L. & Koblinsky, C., 1994. An error covariance model for sea-surface topography and velocity derived from TOPEX/POSEIDON altimetry, *J. geophys. Res.,* **99,** 24 669–24 683.

Wunsch, C., 1978. The North Atlantic general circulation west of 50°W determined by inverse methods, *Rev. Geophys.,* **16,** 583–620.

Wunsch, C., 1991. Global-scale sea surface variability from combined altimetric and tide gauge measurements, *J. geophys. Res.,* **96,** 15 053–15 082.

Wunsch, C., 1994. Dynamically consistent hydrography and absolute velocity in the eastern North Atlantic Ocean, *J. geophys. Res.,* **99,** 14 071–14 090.

Wunsch, C., 1996. *The Ocean Circulation Inverse Problem,* Cambridge University Press, Cambridge.

Wunsch, C. & Gaposchkin, E.M., 1980. On using satellite altimetry to determine the general circulation of the oceans with application to geoid improvement, *Rev. Geophys.,* **18,** 725-745.

Wunsch, C. & Stammer, D., 1995. The global frequency–wavenumber spectrum of oceanic variability estimated from TOPEX/POSEIDON altimetric measurements, *J. geophys. Res.,* **100,** 24 895–24 910.

APPENDIX A: GEOID SLOPE ERROR COVARIANCE

Given the error covariance for the coefficients of a spherical harmonic geopotential model, the corresponding error covariance of geoid slopes associated with a set of scattered geographic locations and directions can be calculated. Using Brun's formula (Heiskanen & Moritz 1967, p. 85), consider the geoid height error, δh, expressed as a function of the geodetic latitude, ϕ, and longitude, λ:

$$\delta h(\phi,\lambda)=\sum_{m=-N}^{N}\delta H_m(\phi)q_m(\lambda), \tag{A1}$$

with

$$\delta H_m(\phi)=\frac{\mu}{\gamma_0 r_0}\sum_{n=|m|}^{N}\left(\frac{a}{r_0}\right)^n \bar{P}_{n|m|}(\sin\phi_0)\delta\bar{C}_{nm} \tag{A2}$$

and

$$q_m(\lambda)=\begin{cases}\cos m\lambda, & m\geq 0\\ -\sin m\lambda, & m<0\end{cases}, \tag{A3}$$

where a, μ denote the semi-major axis and the gravitational constant of a mean earth ellipsoid to which the geoid height is referenced, and γ_0, r_0 and ϕ_0 are the normal gravity, the geocentric radius, and the geocentric latitude evaluated at the surface of the mean earth ellipsoid as functions of ϕ. $\bar{P}_{n|m|}$ denotes the fully normalized Legendre polynomial of degree n and order $|m|$, the extension to negative order using eq. (A3). $\delta\bar{C}_{nm}$ denotes the error in the corresponding spherical harmonic coefficient, and N is the cut-off degree.

Taking the surface gradient of (A1) in the direction of azimuth angle A, one obtains the geoid slope error,

$$\delta s(\phi, \lambda, A) = \delta s_x(\phi, \lambda) \sin A + \delta s_y(\phi, \lambda) \cos A\,, \tag{A4}$$

where δs_x, δs_y denote the zonal and meridional components,

$$\delta s_x(\phi, \lambda) = \frac{1}{r_0 \cos\phi} \frac{d\delta h}{d\lambda} = \sum_{m=-N}^{N} \delta U_m(\phi) p_m(\lambda)\,,$$

$$\delta s_y(\phi, \lambda) = \frac{1}{r_0} \frac{d\delta h}{d\phi} = \sum_{m=-N}^{N} \delta V_m(\phi) q_m(\lambda)\,, \tag{A5}$$

with

$$\delta U_m(\phi) = \sum_{n=|m|}^{N} X_{nm}(\phi)\delta\bar{C}_{nm}\,, \qquad \delta V_m(\phi) = \sum_{n=|m|}^{N} Y_{nm}(\phi)\delta\bar{C}_{nm}\,, \tag{A6}$$

$$X_{nm}(\phi) = \frac{\mu}{\gamma_0 r_0^2}\left(\frac{a}{r_0}\right)^n \frac{m\bar{P}_{n|m|}}{\cos\phi}\,, \qquad Y_{nm}(\phi) = \frac{\mu}{\gamma_0 r_0^2}\left(\frac{a}{r_0}\right)^n \frac{d\bar{P}_{n|m|}}{d\phi}\,, \tag{A7}$$

and

$$p_m(\lambda) = -q_m(-\lambda)\,. \tag{A8}$$

Consider now two geoid slopes, $s(\phi_1, \lambda_1, A_1)$ and $s(\phi_2, \lambda_2, A_2)$. Their error covariance can be expressed as

$$\mathbf{R}_{ss} = \langle \delta\mathbf{s}(\phi_1, \lambda_1, A_1)\delta\mathbf{s}(\phi_2, \lambda_2, A_2)\rangle = \mathbf{a}_1^{\mathrm{T}} \langle \delta\mathbf{s}(\phi_1, \lambda_1)\delta\mathbf{s}(\phi_2, \lambda_2)\rangle \mathbf{a}_2\,, \tag{A9}$$

with

$$\langle \delta\mathbf{s}(\phi_1,\lambda_1)\delta\mathbf{s}(\phi_2,\lambda_2)\rangle$$

$$= \sum_{m_1=-N}^{N} \sum_{m_2=-N}^{N} \mathbf{p}_{m_1}(\lambda_1) \langle \delta\mathbf{U}_{m_1}^{T}(\phi_1)\delta\mathbf{U}_{m_2}(\phi_2)\rangle \mathbf{p}_{m_2}^{\mathrm{T}}(\lambda_2) \tag{A10}$$

and

$$\langle \delta\mathbf{U}_{m_1}(\phi_1)\delta\mathbf{U}_{m_2}(\phi_2)\rangle$$

$$= \sum_{n_1=|m_1|}^{N} \sum_{n_2=|m_2|}^{N} \mathbf{X}_{n_1 m_1}^{T}(\phi_1) \langle \delta\bar{C}_{n_1 m_1}\delta\bar{C}_{n_2 m_2}\rangle \mathbf{X}_{n_2 m_2}^{\mathrm{T}}(\phi_2)\,, \tag{A11}$$

where vector/matrix notation is used such that $\delta\mathbf{s} = \{\delta s_x, \delta s_y\}^{\mathrm{T}}$, $\delta\mathbf{U} = \{\delta U, \delta V\}^{\mathrm{T}}$, $\mathbf{X} = \{X, Y\}^{\mathrm{T}}$, $\mathbf{p} = \{p, q\}$ and $\mathbf{a_i} = \{\sin A_i, \cos A_i\}^{\mathrm{T}}$. Each element of the two-by-two matrix equation (A10) is essentially a 2-D Fourier series whose coefficients are given in (A11). We will refer to (A11) as the 'latitude-lumped geoid slope error covariance' (LEC). When the size of the geopotential covariance, and/or the number of slopes is large, the covariance calculation becomes a computationally demanding task. Because the computational burden lies mainly in the LECs, we first generated them for all possible pairs of 1° intervals of latitudes, and stored them in a direct access device. For L latitudes, there exist $L(L+1)/2$ latitude pairs and each pair has $4(2N+1)^2$ LECs. To construct the 1600 by 1600 slope error covariance matrix, the stored LECs were interpolated to each hydrographic station pair in terms of bicubic Lagrange polynomials over four latitudes by four latitudes. Longitudinal evaluation of (A10) followed by (A9) then completes the task.

APPENDIX B: COVARIANCE OF A WHITE-NOISE FIELD

For a homogeneous isotropic field on a sphere, the spatial covariance between two arbitrary points (i and j) can be expressed as:

$$C_v(i,j) = C_v(\theta_i, \lambda_i, \theta_j, \lambda_j) = \frac{1}{4\pi} \sum_{n=0}^{N} a_n^2 (2n+1) P_n(\cos\psi_{ij})\,, \tag{B1}$$

$$\cos\psi_{ij} = \cos\theta_i \cos\theta_j + \sin\theta_i \sin\theta_j \cos(\lambda_i - \lambda_j) \tag{B2}$$

(e.g. Wunsch & Stammer 1995), where P_n is the nth Legendre polynomial, ψ_{ij} is the spherical angular distance between the two points, and α_n is called the degree variance. Let the field be white to the cut-off degree N, i.e. $\alpha_n = 1$, for $n \le N$. As shown in Fig. B1 (for $N = 70$), the covariance for such a truncated white-noise field has a local pattern similar to a sinc function with a pseudo period of about $360°/N$. This pseudo-period can be qualitatively explained by the expression (Abramowitz & Stegun 1968, eq. 8.9.1)

$$C_{ij} = \frac{N+1}{4\pi(1-\cos\psi_{ij})}[P_N(\cos\psi_{ij}) - P_{N+1}(\cos\psi_{ij})]\,,$$

which can also be approximated as (Abramowitz & Stegun 1968, eq. 8.7.3)

$$C_{ij} \approx G_N[2\cos(N+3/2)\psi_{ij} - \cos(N+7/2)\psi_{ij}]\sin(\psi_{ij}/2)\,, \tag{B3}$$

where G_N is a function of N and $(1-\cos\psi_{ij})$. The period of the leading term is $360°/(N+3/2)$. The apparent period observed in the JGM-3 geoid slope error covariance (developed to $N = 70$) is larger than that of the white-noise field by about 20

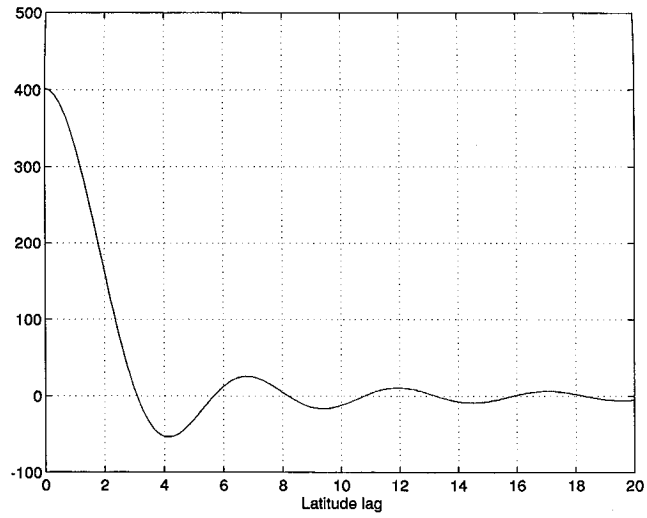

Figure B1. Covariance vector of a white-noise field expanded in spherical harmonics at degree and order 70.

per cent in some locations, and thus deviates significantly from white noise.

APPENDIX C: ALONG-SECTION FILTERING OF THE UNEVENLY SPACED VELOCITIES

An ideal 1-D filter that removes wavelengths shorter than ψ_c is

$$f = \frac{1}{\psi_c}\,\mathrm{sinc}\left(\frac{\psi}{\psi_c}\right)$$

(Bracewell 1978) for a linear distance ψ (in our case ψ is the spherical angular distance). For any 1-D field $v(\psi)$ (here the velocity normal to the section) the filtered field $\tilde{v}$ is obtained from the convolution

$$\tilde{v}(\psi) = \int_{-\infty}^{+\infty} f(\psi' - \psi) v(\psi')\, d\psi', \qquad \text{(C1)}$$

where the integration is carried out numerically and the bounds are finite. To avoid Gibbs' effects, one may use a tapering scheme such as a Hamming window. With a window size of $6\psi_c$ and a sampling interval of 0.5° (the average distance between station pairs), the response of such a tapered filter is shown in Fig. C1. For the velocity data, which are unevenly spaced, trapezoidal integration was used:

$$\bar{v}_i = \sum_{j=j_1(i)}^{j_2(i)-1} \frac{1}{2}(\psi_{i,j+1} - \psi_{i,j})(f_{ij} c_{ij} v_j + f_{i,j} c_{i,j+1} v_{j+1}), \qquad \text{(C2)}$$

where $\psi_{i,j}$ is the distance betwen pair i and pair j. $j_1(i), j_2(i)$ are defined as the first and last pairs of the section in the $6\psi_c$ window centred at pair i. $c_{ij} = \cos(A_i - A_j)$ is the direction cosine projecting velocity v_j (with azimuth angle A_j) onto the direction of velocity v_i (with azimuth angle A_i). The filtering scheme (C2) will be represented in matrix form as

$$\bar{\mathbf{v}} = \mathbf{F}_N \mathbf{v}, \qquad \text{(C3)}$$

where $\mathbf{F}_N$ is a banded matrix whose bandwidth corresponds to the window size $6\psi_c$. The effect of the filter $\mathbf{F}_{70}$ ($\psi_c = 5.14°$) is illustrated in Fig. C2.

APPENDIX D: A LOW–LOW SATELLITE-TO-SATELLITE GRAVIMETRIC MISSION

Sharma (1995) produced an estimate of the accuracy with which gravity could be recovered from a hypothetical low–low satellite-to-satellite tracking mission. In particular, the assumption was made that two coplanar satellites with 5° separation are in a near-circular near-polar orbit at 250 km altitude. The lead satellite is assumed to be tracked by GPS and the two satellites are assumed to measure mutually relative ranges and range rates with precisions of 10 μm and 1 μm respectively. The analysis of a single 10 day repeat cycle was carried out with a measurement sampling interval of 30 s in the presence of 1 nm s^{-2} white-noise acceleration errors. Further details can be found in Sharma (1995).

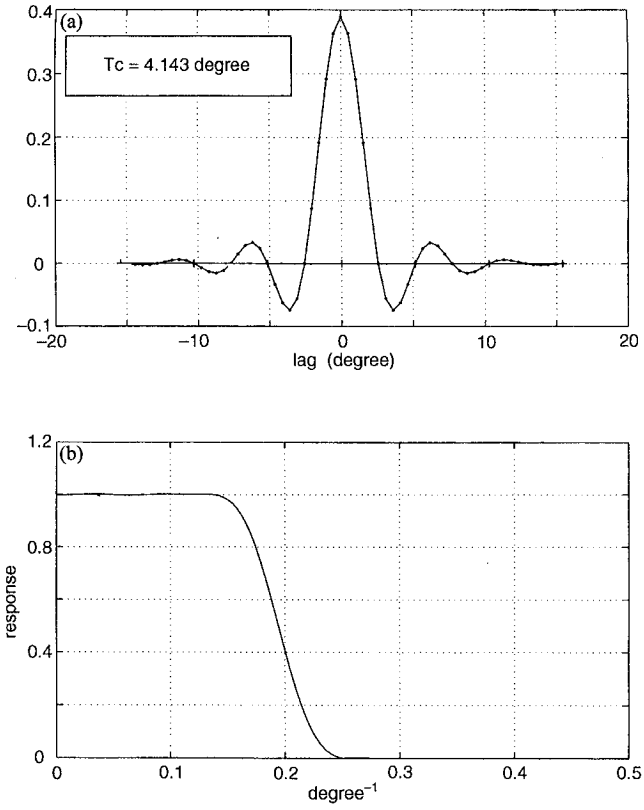

Figure C1. Wavenumber response of the ideal 1-D filter $f(\psi, \psi_c) = \mathrm{sinc}(\psi/\psi_c)$ with a Hamming window $h(\psi, \psi_c) = 0.54 + 0.46\cos(\psi/6\psi_c)$. (a) The tapered filter $f(\psi, \psi_c)h(\psi, \psi_c)$ with $\psi_c = 360/70 = 5.14°$. f is sampled at a 0.5° step, the mean distance between station pairs. (b) The amplitude drops at $\psi_c^{-1} \simeq 0.2$.

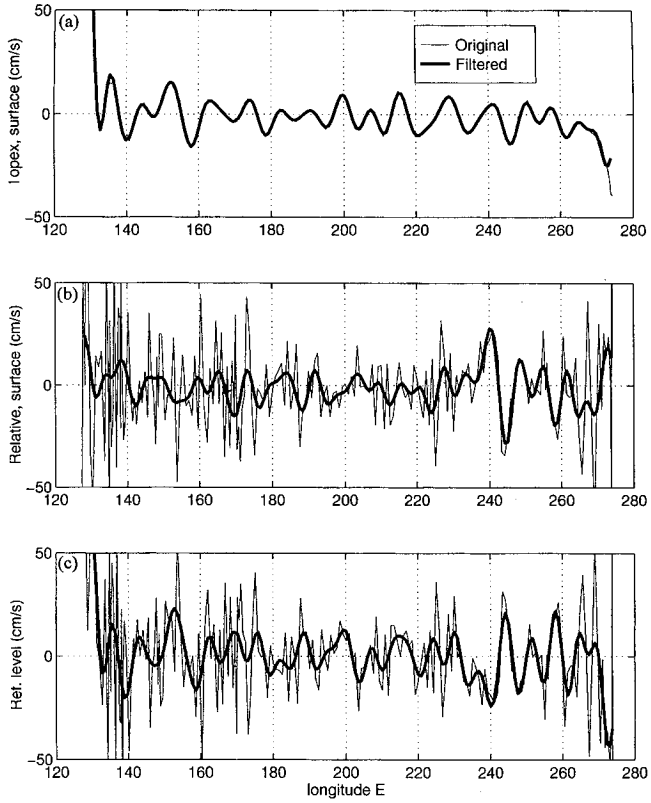

Figure C2. Effect of the 1-D filter $\mathbf{F}_{70}$ on the velocity data on the zonal section Pacific 10°N. (a) The original TOPEX/POSEIDON data remain quite unchanged as they do not contain short wavelengths. (b) The small scales are removed from the relative velocity, yielding a less noisy estimate of the reference level velocity shown in (c).