Ocean Dynamics manuscript No.
(will be inserted by the editor)

O. B. Andersen · G. D. Egbert · S. Y. Erofeeva · R. D. Ray

Mapping nonlinear shallow-water tides: A look at the past and future

In memory of Christian Le Provost

Received: 00 November 2005 / Accepted: date

Abstract Overtides and compound tides are generated by nonlinear mechanisms operative primarily in shallow waters. Their presence complicates tidal analysis owing to the multitude of new constituents and their possible frequency overlap with astronomical tides. The science of nonlinear tides has been greatly advanced by the pioneering researches of Christian Le Provost, who employed analytical theory, physical modeling, and numerical modeling in many extensive studies, especially of the tides of the English Channel. Le Provost's complementary work with satellite altimetry motivates our attempts to merge these two interests. After a brief review, we describe initial steps towards the assimilation of altimetry into models of nonlinear tides via generalized inverse methods. A series of barotropic inverse solutions is computed for the M_4 tide over the northwest European Shelf. Future applications of altimetry to regions with fewer *in situ* measurements will require improved understanding of error covariance models, since these control the tradeoffs between fitting hydrodynamics and data, a delicate issue in coastal regions. While M_4 can now be robustly determined along the Topex/Poseidon satellite ground-tracks, many other compound tides face serious aliasing problems.

Keywords Tides · Nonlinear tides · Overtides · Satellite altimetry

O. B. Andersen
Danish National Space Center
Juliane Maries vej 30
2100 Copenhagen, Denmark
Tel.: +45-3532-5754
Fax: +45-3536-2475
E-mail: oa@spacecenter.dk

G. D. Egbert & S. Y. Erofeeva
COAS, Oregon State University, Corvallis, OR, USA
E-mail: egbert@coas.oregonstate.edu
E-mail: serofeev@coas.oregonstate.edu

R. D. Ray
NASA/GSFC, Code 697, Greenbelt, MD, USA
E-mail: richard.ray@nasa.gov

1 Introduction

In the open ocean the tidal spectrum can normally be represented by a limited number of well-defined astronomical frequencies, primarily in the diurnal and semidiurnal bands. In coastal regions, however, the spectrum can become much more complex, as the semidiurnal and diurnal frequencies are mixed with a large number of "shallow water tides," having frequencies within the long-period, diurnal, semidiurnal, terdiurnal, and higher bands. These new frequencies are a consequence of the nonlinear interactions between the tidal waves as they propagate in shallow water.

Over the past few decades no one has done more to shed light on the nature of shallow-water tides than Christian Le Provost. Concentrating on the complicated, nonlinear tides of the English Channel, Le Provost attacked the problem with an unusually wide variety of methods:

- analytical analysis based on perturbation methods applied to the shallow water equations (Le Provost 1974, 1976, 1991; Kabbaj and Le Provost 1980).
- physical modeling with a large (14 meter) rotating platform, with an astonishing vertical precision of 0.02 mm, corresponding to cm-level precision in the real world (Chabert d'Hières and Le Provost 1976, 1979).
- time-stepping numerical modeling, with both finite-difference and finite-element formulations (Le Provost and Fornerino 1985).
- spectral numerical modeling—i.e., the frequency domain analogue to the time-stepping codes (Le Provost and Poncet 1977, 1978; Le Provost et al. 1981)
- data analysis (Chabert d'Hières and Le Provost 1979; Le Provost 1991; Ponchaut et al. 2001).

Owing to the difficult challenges posed by the nonlinearities in the equations of motion, Le Provost's analytical analyses proved to be of crucial importance to his subsequent successful physical and numerical modeling.

In this paper we address shallow-water tides in the context of another of Le Provost's longstanding interests—the tidal analysis of satellite altimeter data and the assimilation of altimetry into numerical tidal models (Le Provost et al.

1998; Le Provost 2001). Tidal data assimilation is now routinely applied at both global and regional scales, but it has not previously been used for nonlinear tides. Considering that numerical modeling is intrinsically more challenging for nonlinear tides than for linear tides—for example, spatial scales are shorter and the forcing is not known *a priori*—the use of data assimilation appears at first sight quite attractive. Unfortunately, nonlinear tides present their own special difficulties for assimilation. The work here represents a first step in this direction. The reader will discover that the problem is by no means solved.

After a brief review of nonlinear shallow-water tides, this paper focuses on mapping the M_4 barotropic tide over the northwest European Shelf, the region of Le Provost's keen interest for so many years. M_4 is the primary overtide of M_2, and it reaches an amplitude exceeding 30 cm in several parts of the English Channel. Our focus is primarily on tidal elevations rather than currents, and the models we employ are exclusively 2D barotropic. The use of generalized inverse methods with 3D tidal models is a promising approach for more detailed investigations of tidal currents and internal tides, but such work is in its infancy.

2 Shallow water tides

In coastal regions the tidal range is generally larger than in the open ocean, and the tidal waves are considerably more complex. The patterns of the tidal waves shorten as the wave speed reduces. Since long waves propagate as $\sqrt{(gH)}$, where H is the water depth and g is normal gravity, tidal wavelengths shorten dramatically in shallows—for example, 10 times shorter at 40 meters water depth than at 4000 meters water depth. Similarly, resonance or near-resonance responses add to the complexity of the tidal pattern and tend to intensify nonlinear effects.

Shallow-water constituents are caused by the nonlinear interaction of the astronomical semidiurnal and/or diurnal constituents when the tidal wave propagates over the shelf. Nonlinearity induces motion at multiples of the fundamental frequencies and at sums and differences of frequencies of interacting waves. Shallow-water tidal nomenclature typically refers to these as overtides and compound tides, respectively. That is, overtides are constituents that appear at multiples of one base frequency like M_4 or M_6, which have double and triple the frequency of M_2, while compound tides are constituents having frequencies at linear combinations of interacting tides—for example, MS_4 which is the interaction between M_2 and S_2, or MNS_2 which is a triple interaction between M_2, N_2, and S_2. See Table 1 below for further examples.

In the following we employ the depth-integrated 2D shallow water tidal equations, consisting of a momentum equation

$$\frac{\partial \mathbf{U}}{\partial t} + \mathbf{f} \times \mathbf{U} + \mathbf{F} + \mathbf{U} \cdot \nabla \mathbf{u} + A_H \nabla^2 \mathbf{U} = -gD\nabla(\eta - \bar{\eta}) \qquad (1)$$

and a continuity equation

$$-\frac{\partial \eta}{\partial t} = \nabla \cdot \mathbf{U}, \qquad (2)$$

where η is the elevation fluctuation of the sea surface, $\mathbf{U}$ is the tidal volume transport vector, equal to the depth-averaged tidal velocity $\mathbf{u}$ times total water depth $D = (H + \eta)$, and $\mathbf{f}$ is the Coriolis parameter, oriented vertically. Dissipation terms include a parameterization of bottom friction $\mathbf{F}$ and horizontal viscosity with eddy coefficient A_H. The equilibrium tide $\bar{\eta}$ includes two parts, one derived from the astronomical tide-generating potential, with allowance for the earth's body tide, and a second accounting for tidal loading and self-attraction.

Nonlinearities in Eqns (1–2) are responsible for the generation of overtides and compound tides. Nonlinearities are seen to enter the equations through the continuity equation—from the $\nabla \cdot \eta \mathbf{u}$ term—and through the advection and frictional terms in the momentum equation.

Dissipation due to bottom friction is normally described by a term quadratic in velocity: $\mathbf{F} = c\,|\mathbf{u}|\,\mathbf{U}/(H+\eta)$, where c is the nondimensional dissipation parameter, generally taken about 0.0025. This term is the major source of difficulty, with two nonlinear aspects: the quadratic part $|\mathbf{u}|\,\mathbf{U}$, and the elevation η in the denominator (Parker 1991). Work to linearize these terms in various ways (e.g., Le Provost 1973; Le Provost and Poncet 1977; Heaps 1978; Hunter 1979) has proven important for subsequent modeling efforts.

A number of important studies now exist that have established which of the nonlinear terms in (1–2) are most significant in giving rise to which tidal effects. The work by Le Provost and Fornerino (1985), Pingree and Maddock (1987), and Walters and Werner (1991) demonstrated how the nonlinearities in the continuity equation dominate the generation of the M_4, MS_4, and other tides in the English Channel. They also quantified the contribution from advection around capes and bottom friction to this constituent. These others terms appear especially important around the Cherbourg Peninsula (Walters and Werner 1991). See Le Provost (1991) and Parker (1991) for comprehensive reviews.

3 Northwest European Shelf

The northwest European continental shelf (henceforth European Shelf), and particularly the English Channel, have sea surface variations ranging more than 13 meters during spring tides. Similarly the region is densely populated, and consequently much effort has been expended in the study of tides for safety, navigation, and other societal reasons. The European Shelf is probably the most intensively studied region in the world from a tidal perspective.

The most important nonlinear tides on the European Shelf are tabulated in Table 1, where constituents are included whose amplitudes at Dover (U.K.) exceed 1 cm. In many respects this is a much abbreviated table, since some shallow-water analyses can include well over a hundred constituents

Table 1 Principal nonlinear tides on the Northwest European Shelf

Tide	Origin	Doodson Number	Frequency ω (°/hr)	Alias period (days) T/P	ERS	GFO	Amplitude (cm) at Dover, UK
Long period							
MSf	M_2-S_2	073.555	1.0159	30.2	94.5	110.2	2.2
Semidiurnal							
MNS_2	$M_2+N_2-S_2$	227.655	27.4238	77.3	3166.1	98.7	2.6
$2MS_2$	$M_2+M_2-S_2$	237.555	27.9682	20.3	135.1	81.8	5.8
SNM_2	$S_2+N_2-M_2$	263.655	29.4556	21.0	129.5	35.4	4.0
$2MN_2$	$M_2+M_2-N_2$	265.455	29.5285	20.6	349.2	39.2	7.0
MSN_2	$M_2+S_2-N_2$	283.455	30.5444	51.9	129.5	60.8	3.6
$2SM_2$	$S_2+S_2-M_2$	291.555	31.0159	19.9	94.5	66.7	4.2
Terdiurnal							
MK_3	M_2+K_1	365.555	44.0252	96.8	127.5	392.7	1.5
Fourth diurnal							
MN_4	M_2+N_2	445.655	57.4238	244.5	3166.1	62.3	9.4
M_4	M_2+M_2	455.555	57.9682	31.1	135.1	158.6	25.5
ML_4	M_2+L_2	465.455	58.5126	30.9	74.4	34.9	3.0
MS_4	M_2+S_2	473.555	58.9841	1083.9	94.5	361.0	16.4
MK_4	M_2+K_2	475.555	59.0662	219.8	195.8	121.3	5.0
S_4	S_2+S_2	491.555	60.0000	29.4	∞	84.4	1.6
Sixth diurnal							
$2MN_6$	$M_2+M_2+N_2$	645.655	86.4079	83.3	91.7	77.5	3.8
M_6	$M_2+M_2+M_2$	655.555	86.9523	20.7	314.5	105.7	6.5
MSN_6	$M_2+S_2+N_2$	663.655	87.4238	47.4	3166.1	45.5	1.7
$2MS_6$	$M_2+M_2+S_2$	673.555	87.9682	65.9	135.1	2608.1	6.5
$2MK_6$	$M_2+M_2+K_2$	675.555	88.0503	48.4	77.6	196.4	1.8
$2SM_6$	$S_2+S_2+M_2$	691.555	88.9841	55.7	94.5	115.0	1.4
Eighth diurnal							
M_8	$M_2+M_2+M_2+M_2$	855.555	115.9364	27.4	72.6	79.3	2.0
$3MS_8$	$M_2+M_2+M_2+S_2$	873.555	116.9523	32.0	314.5	282.7	2.9

(e.g., Zetler and Cummings 1967), depending on the tidal regime. The relative lack of diurnal and terdiurnal tides in Table 1 stems from the general weakness of all diurnal tides in the Atlantic Ocean, since most of the terdiurnal constituents as well as the nonlinear diurnal constituents (e.g., MP_1) arise as double interactions between M_2 and diurnal tides.

By far the most important of the nonlinear tides is M_4. Aside from simple interpolations between coastal stations, the first cotidal chart of M_4 in this region, covering specifically the English Channel, was produced by Chabert d'Hières and Le Provost (1970). Their chart is reproduced here as Fig. 1. While the primary M_2 tide in the English Channel is essentially an eastward propagating Kelvin wave, its first overtide is seen to be far more complex, with two amphidromes existing within the channel about 250 km apart. Amplitudes of M_4 approach 40 cm at several locations along the coast of France.

In addition to Le Provost's modeling work in this region, cited above, we should also mention important work by Flather (1976), Pingree and Maddock (1978), Howarth and Pugh (1983), Davies (1986), Walters (1987), Werner and Lynch (1989), Walters and Werner (1991), Davies et al. (1997), Kwong et al. (1997), and Sinha and Pingree (1997). Some of these studies employed 3D models, which are essential for realistic simulations of tidal currents on the shelf.

4 Empirical estimates of nonlinear tides from satellite altimetry

With the launch of the Topex/Poseidon (T/P) satellite in 1992 and subsequent development of high-resolution global hydrodynamic models, the mapping of most semidiurnal and diurnal constituents improved dramatically over the deep oceans. In shallow water, however, discrepancies between these data-constrained global models are still plainly evident and unacceptably large (Andersen et al. 1995; Shum et al. 1997). Part of the reason for this (although certainly not the exclusive reason) is that none of these models thus far contains nonlinear constituents.

Although satellite altimetry can provide unique insights to the spatial distribution of nonlinear tides over continental shelves, the routine use of altimetry for this is impeded by several difficulties. Firstly, nonlinear tides have generally small amplitudes in most regions, relative to present altimetric noise levels, so signal-to-noise levels are often unfavorable. Secondly, the spatial scales of nonlinear tides are so small that the ground-track distances between satellite tracks are usually too large to allow straightforward empirical maps to be drawn. This necessitates some type of cross-track interpolation, either statistical (Andersen 1999) or hydrodynamical (see below). Thirdly, the alias periods of nonlinear tides as sampled by satellite are often unfavorable. Finally,

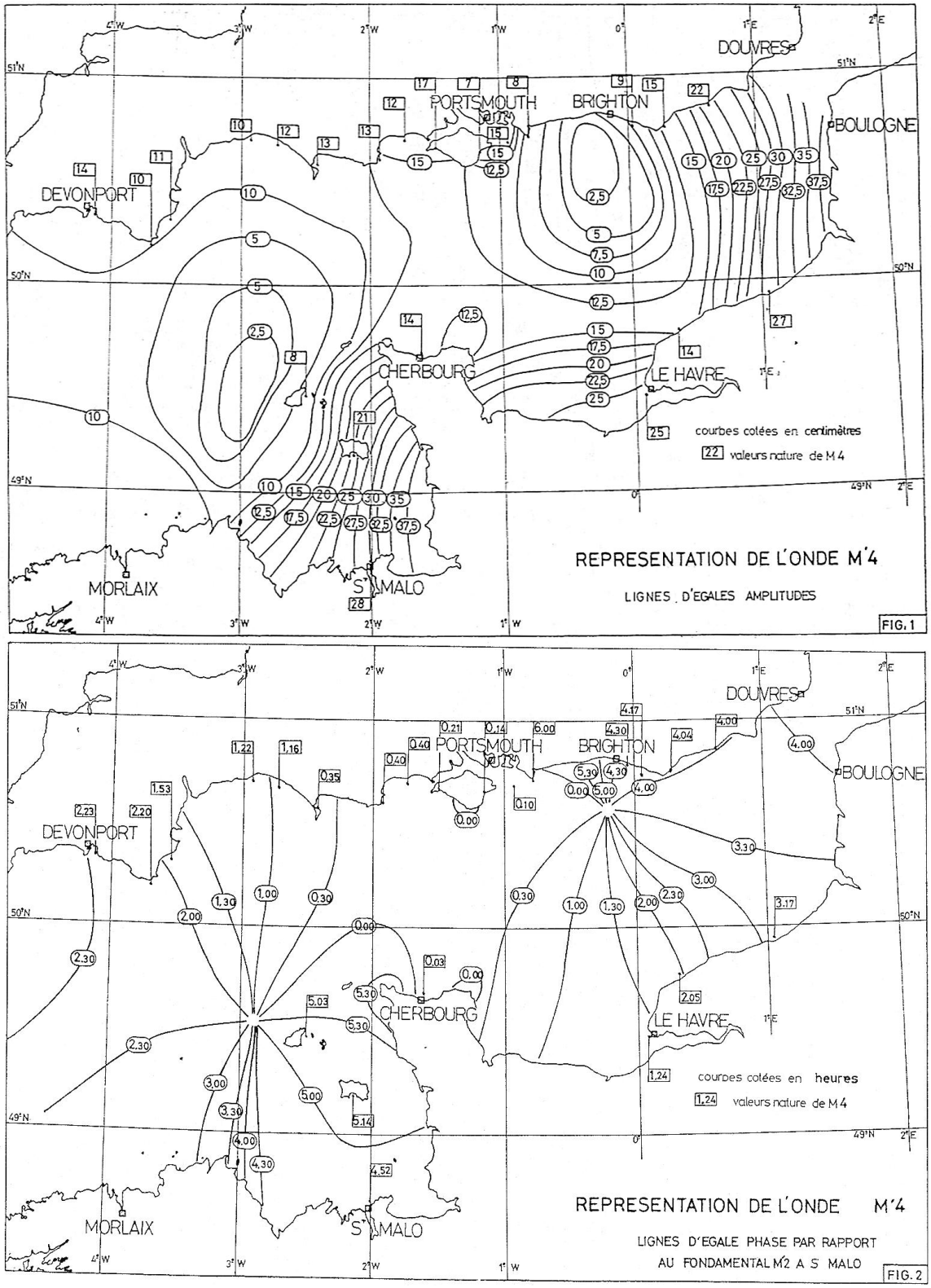

Fig. 1 Amplitude (top) and phase (bottom) of the M_4 tide in the English Channel, according to Chabert d'Hières and Le Provost (1970). Amplitudes in cm; phases in lunar hours.

Table 2 Primary linear/nonlinear coincident frequencies

Linear Tide	Doodson Number	Nonlinear Interaction(s)
Long period		
M_0	055.555	M_2-M_2, etc.
Mm	065.455	K_1-O_1
MSf	073.555	M_2-S_2
Mf	075.555	M_2-N_2
Diurnal		
Q_1	135.655	N_2-K_1
O_1	145.555	M_2-K_1
τ_1	147.555	M_2-P_1
M_1	155.655	N_2-O_1
P_1	163.555	S_2-K_1
K_1	165.555	M_2-O_1, S_2-P_1, K_2-K_1
OO_1	185.555	$K_1+K_1-O_1$
Semidiurnal		
ε_2	227.655	$M_2+N_2-S_2$
$2N_2$	235.755	$N_2+N_2-M_2$, $M_2+M_2-K_2$, O_1+O_1
μ_2	237.555	$M_2+M_2-S_2$
ν_2	247.455	$M_2+L_2-S_2$
λ_2	263.655	$S_2+N_2-M_2$
L_2	265.455	$M_2+M_2-N_2$
S_2	273.555	K_1+P_1
K_2	275.555	K_1+K_1
η_2	285.455	$M_2+K_2-N_2$

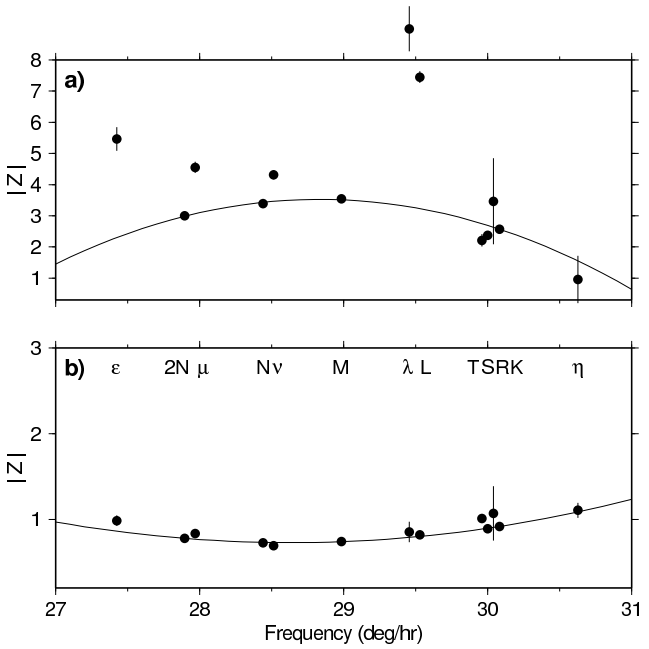

Fig. 2 Admittances of semidiurnal tides determined at (a) Dover, UK, and (b) Kwajalein Island, Pacific Ocean. Smooth curves represent a least-squares quadratic fit to the (2N, N, M, K) tides. The Dover tidal regime is highly nonlinear, and several tides are perturbed by nonlinear interactions (see Table 2). The Kwajalein tidal regime is fairly linear, and the admittance across the band is thus smooth. (Note differences in vertical scales.) Small jitter near S_2 is induced primarily by radiational effects (Munk and Cartwright 1966; Arbic 2005).

the number of satellite observations at any location, even after 12 years of the T/P mission, is still no more than 500, and sometimes much fewer. This is far less data than a typical tide gauge provides, and it limits the sort of analysis that can be done.

This final point—the limited amount of data—is important in nonlinear tidal regimes because of the greater number of independent constituents that must be estimated. Even in linear regimes most tidal analyses of altimetry overcome limited data by exploiting, either explicitly or implicitly, the smoothness of tidal admittances across all tidal bands, so that only a few tidal parameters need be estimated (e.g., Cartwright and Ray 1990; harmonic analyses can employ similar approximations using constituent modulations based on admittances). In nonlinear regimes, however, the admittances are no longer smooth—see Fig. 2—because of nonlinear tides whose frequencies coincide with those of linear astronomical tides; these lines nominally require independent estimation. Table 2 tabulates the most important of these coinciding frequencies.

One approach that warrants investigation is to employ the formalism of Munk and Cartwright (1966), and especially Cartwright (1968), to handle nonlinear tides through bilinear and trilinear admittances. In fact, Cartwright (1968) found that his method gave superior results with fewer independent free parameters than the standard approach of employing multiple compound tides. The application of such techniques to satellite altimetry over shallow seas may prove similarly beneficial.

The aliasing issue in altimetry arises because the sampling intervals of altimetric satellites are much longer than the tidal periods. The alias periods for the major astronomical tides have been discussed by many authors (e.g., Parke et al. 1987; Andersen and Knudsen 1997; Ray 1997). Table 1 lists the alias periods of the major nonlinear tides as they are sampled by three altimetric satellite missions: Topex/Poseidon and Jason-1 (repeat period 9.9156 days), Geosat and GFO (17.0505 d), and ERS–1/2 and Envisat (35.0000 d).

For T/P the alias period of the largest shallow-water constituent M_4 is 31.05 days. It is well separated from all important shallow-water and astronomical constituents in T/P altimetry. Some other shallow-water tides are less fortunate. MS_4 has an alias period of 1084 days and will therefore be prone to contamination by interannual variations in sea level with periods near 3 years. M_6 has an alias of 20.70 d, fairly short, but unfortunately this is close to the alias of the L_2 (and coincident $2MN_2$) tide; they are formally separable only after 18 years.

For the sun-synchronous ERS and Envisat satellites the situation is far worse. Table 1 suggests that these satellites are not suited at all for recovering shallow-water constituents. Sun-synchronous orbits sample solar tides S_n, for all species n, at a single phase, giving an alias speed of zero. Consequently, interaction constituents involving S_2 end up with identical aliases; for example, $2SM_2$ and $2SM_6$ have speeds identical to M_2, while $2MS_2$ and $2MS_6$ have speeds of twice the M_2 speed, just like M_4, All these constituents are thus inseparable. In general, the ERS satellites are of

dubious value in measuring all solar tides (Andersen and Knudsen 1997).

Andersen (1999) addressed nonlinear shallow-water tidal problems using satellite altimetry and was able to derive estimates of M_4 and M_6 with favorable comparison to tide gauges on the Northwest European shelf. Attempts for other constituents like MN_4, $2SM_2$, and M_8 were unsuccessful, partly owing to their small amplitudes. Initial studies of this sort, based on purely empirical tidal estimates, are useful and necessary, and they pave the way for more sophisticated high-resolution hydrodynamic models that can assimilate altimeter data.

With more than a decade of measurements along the original T/P ground tracks, we find that M_4 can now be estimated at each location along-track with high precision. We estimate M_4 standard errors in most locations to be of order 5–15 mm. At a few locations, especially in intense boundary currents or near land, the uncertainties grow substantially. On the European Shelf this unfortunately occurs most often for some high-tide points in the English Channel. For the newer interlaced T/P tracks, flown since September 2002, there is far less data and standard errors are correspondingly higher—of order 35 mm or larger, which is barely acceptable for the study of shallow-water nonlinear tides. We use these along-track M_4 estimates in the next section as input to the assimilation studies.

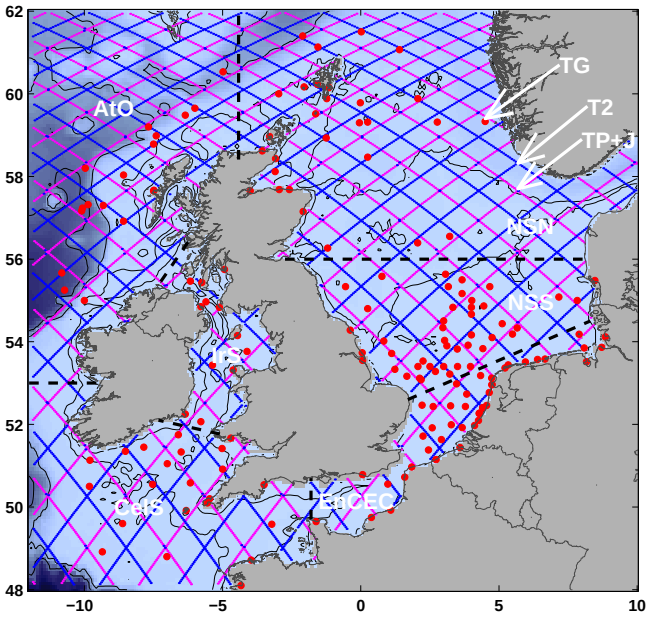

Fig. 3 Data sites on bathymetry background. Topex/Poseidon + Jason tracks are shown with magenta, Topex2 tracks with blue. Tide gauge locations are shown with red dots. Bathymetry is shown with color and contours for depth 10, 50,100, 200, 300 and 1000 m.

5 Satellite data assimilation

Variational data assimilation provides a rigorous way to combine dynamical models with data, improving the accuracy of maps of tidal elevations and currents from what can be obtained from hydrodynamic modeling alone (Egbert and Bennett 1996). As noted above, such methods are especially attractive for shallow water where the spatial scales are undersampled by altimeters. In this final section we explore application of data assimilation methods to mapping of M_4 on the European shelf.

M_4 tidal solutions assimilating the altimeter data were obtained with OTIS (OSU Tidal Inversion Software; Egbert and Erofeeva 2002). OTIS implements all stages of data assimilation, including computing a prior solution to the forward modeling problem, fitting data, and posterior error analysis of the inverse solution. OTIS allows assimilation of a wide range of data types using a representer approach (Bennett 1992; Egbert et al. 1994). This is the first use of OTIS for inverse modeling of nonlinear tides, requiring some modifications, especially in the treatment of error covariances. These are discussed more fully below.

5.1 Model domain

The model domain extends from 48°N–62°N, 12°W–10°E, including the European shelf, the North Sea and adjacent Atlantic Ocean (Figure 3), with a resolution of 1/12 degree. Bathymetry was obtained by merging gridded data from several sources. The bathymetry in this area is complicated and its accuracy critical for accurate modeling of shallow water tides such as M_4. Shortcomings in our present compilation, including its relatively coarse resolution (cf. Jones and Davies 1996), certainly limit our present ability to model these tides.

5.2 Data

Data sets used for assimilation and solution validation include: (a) 364 orbit cycles of Topex/Poseidon data (1992–2002), augmented by 72 orbit cycles of Jason-1 data following the same ground track (2002–2004); (b) 69 orbit cycles of Topex2 data on tracks interleaved between the original T/P tracks (2002–2004); and (c) 158 tide gauges on the European Shelf and coast. Locations of these data are shown along with the bathymetry in Fig. 3

The tide gauge data are a compilation from several sources, including bottom-pressure data from Cartwright and Zetler (1985; later supplemented by Smithson 1992), coastal gauges from various national authorities, and pelagic gauges mounted on oil rigs in the southern North Sea.

To analyze and discuss model performance the tide gauges were divided into six groups, as indicated by the dashed lines in Fig. 3. In the following we generally refer to these areas with the abbreviations:

EnCEC	:	English Channel and European Coast
NSN	:	North Sea, north of 56°N
NSS	:	North Sea, south of 56°N
AtO	:	Atlantic Ocean
CelS	:	Celtic Sea
IrS	:	Irish Sea

5.3 Dynamical model

The prior solution was obtained by time stepping the nonlinear shallow-water equations (1–2). Boundary conditions are no-flow across (and partial-slip along) the coast, and specification of elevations on (and free-slip along) any open boundaries. Note that nonlinear terms in (1–2) are relatively unimportant for global scale modeling of the major tidal constituents, and thus have been completely omitted in many past applications of OTIS. The nonlinear terms are of course critical to generation of the overtide M_4 considered here, especially the term in the continuity equation (Parker 1991).

To model nonlinear constituents on the European shelf domain of Fig. 3, the model was forced by the modified tide generating potential $\bar{\eta}$ for the four largest semidiurnal constituents (M_2, S_2, N_2, K_2), and the two largest diurnal constituents (K_1, O_1). The explicit time-stepping shallow-water equation solver was run for 46 days, with harmonic analysis for all constituents of interest, including nonlinear constituents M_4, MS_4 and MN_4 over the final 23 days of the run. Here we consider only results for M_4. For the six linear constituents elevations on the open boundary were taken from a global data-assimilating tidal model (TPXO7.0, a more recent version of the global solution described in Egbert and Erofeeva 2002). For the nonlinear tides boundary elevations were taken from a 0.5° global time-stepping solution, essentially the prior model for TPXO.7. Note that self-attraction and loading forces for the nonlinear tides are essentially negligible, since these forces decay rapidly at high wavenumbers; they are therefore neglected in the dynamical model

Prior solution elevations for M_4 are shown in Fig. 4b. For comparison the FES2004 global M_4 solution (Lyard et al., this issue) is shown in Fig. 4a for the same area. The two models have many similarities. Both show the two English Channel amphidromes, as originally deduced by Chabert d'Hières and Le Provost (1970) and seen in Fig. 1. But there are also significant differences—for example, in the Irish Sea and around Scotland. (Jones and Davies (1996) give amplitudes slightly exceeding 15 cm along the eastern shore of the Irish Sea, which agrees somewhat more closely with our prior.) Root mean square (RMS) misfits to tide gauges, broken down into the geographic groupings defined in Fig. 3, are given for both our prior model and for FES2004 in Table 3.

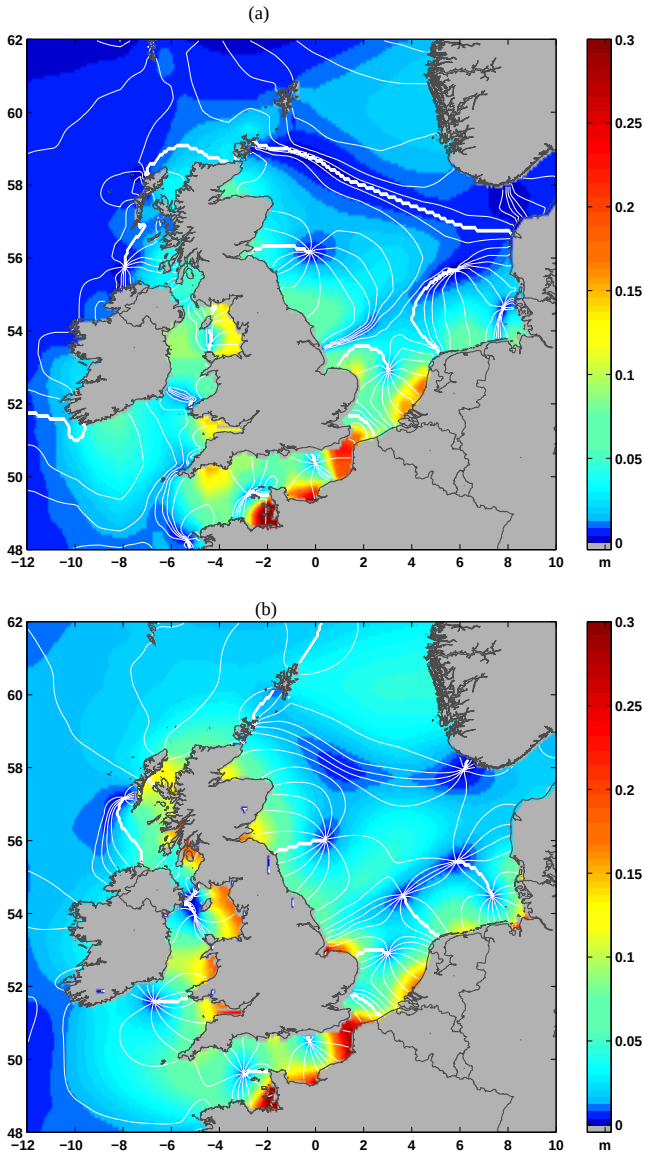

Fig. 4 Cotidal charts of the M_4 tide from (a) the FES2004 solution and (b) our prior time-stepping solution. Amplitudes are shown with color; phase contours are shown every 30°. Data for (a) were kindly provided courtesy of Florent Lyard and colleagues.

5.4 Inverse solutions

For major tidal constituents the only significant nonlinearity in Eq. (1) is the quadratic bottom drag. This is a comparatively crude parameterization of friction, at best only an approximation to the effects of dissipative processes in shallow seas. For example, the usual quadratic model does not include the effective dissipation of the barotropic tides due to baroclinic conversion over topography. Compared with the errors in the simple friction parameterization, other nonlinear terms in (1) are negligible for major constituents. As a consequence, linear dynamics are used for tidal data assimilation in OTIS, with the quadratic friction linearized using currents from the prior model to compute a spatially varying drag coefficient (Egbert and Erofeeva 2002) and with other nonlinear terms omitted. These approximations allow us to adopt a frequency domain approach for computing inverse

solutions with OTIS. We follow the same general strategy here, both for the linear and nonlinear tides. For the latter the nonlinearity in the continuity and momentum equations obviously cannot be neglected; as Parker (1991) and others show, nonlinear tides are generated primarily in the very shallow areas where η becomes comparable to H. Over most of the domain, however, propagation of nonlinear tides can be modeled adequately using purely linear dynamics. We thus use the nonlinear model to generate the prior solution, but the linearized shallow-water equations to define the data assimilation penalty functional. With this approximation errors should be allowed for in both the linearized continuity equation and the momentum equation when assimilating data.

Note that other approaches to data assimilation for nonlinear tides might be worth considering. For example, one could develop the tangent-linear equations for the time domain problem and retain the coupling between constitutents (e.g., between M_2 and M_4) throughout the assimilation. Such an approach would be considerably more involved than the frequency domain scheme we pursue here.

5.5 Dynamical error covariances

In the variational data assimilation approach, error covariances define *a priori* assumptions about the magnitude and spatial structure of errors in dynamical equations, forcing, and boundary conditions. The error covariance models used in OTIS are a compromise between efficient and simple implementation, and scientifically justifiable error hypotheses. For linear constituents such as M_2 the continuity equation, which in the flux form used here is just a statement of mass conservation, is taken as exact. Errors are allowed in the momentum balance equations, primarily to account for inaccuracies in bathymetry and in the parameterization of frictional effects. Covariances for the momentum equations are taken to have a constant spatial correlation length scale for simplicity. Variances, which are spatially variable, are estimated using elevations and currents from the prior model, together with assumptions about the relative errors in bathymetry and in the dissipation parameterization. Egbert et al. (1994) and Egbert and Erofeeva (2002) provide further details on the standard OTIS covariance model. For the assimilation of the linear constituents, relative errors in the bathymetry are assumed to be 5%, and errors in the bottom drag parameterization are taken to be 100%. The correlation scale for errors in the momentum equations is assumed to be 50 km, and errors in open boundary data are assumed to have a standard deviation that is 5% of the average boundary forcing amplitude, with constant decorrelation length scale of 500 km. For an overtide such as M_4 it is no longer clear that the continuity equation should be assumed exact. We briefly consider this possibility here. Consider the elevations and currents expanded to include terms of M_2 and M_4 frequencies (ω and 2ω, respectively):

$$\mathbf{u} = \mathbf{u}_2 e^{i\omega t} + \mathbf{u}_4 e^{2i\omega t} \tag{3}$$

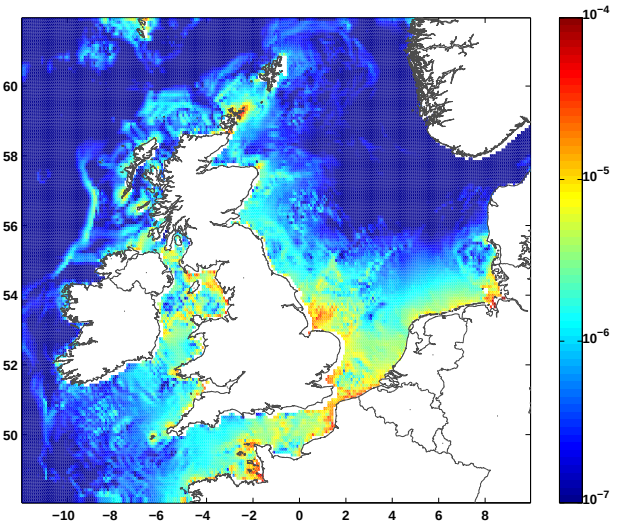

Fig. 5 Error in continuity equation due to M_2 forcing ($m\,s^{-1}$). Note logarithmic color scale.

$$\eta = \eta_2 e^{i\omega t} + \eta_4 e^{2i\omega t}$$

Inserting (3) into (2) and collecting terms of like frequency, we have for M_4

$$\frac{\partial \eta}{\partial t} = \nabla \cdot H\mathbf{u}_4 + \nabla \cdot \eta_2 \mathbf{u}_2. \tag{4}$$

The term $\nabla \cdot \eta_2 \mathbf{u}_2$ is essentially a forcing which drives the M_4 tide. To the extent that the M_2 tidal elevations and currents which define this term in the nonlinear prior solution are inexact, there will be an error in this forcing, and hence in the linearized continuity equation for the M_4 constituent (but not the continuity equation for the M_2 frequency).

To estimate the magnitude of this error we calculate the difference between the M_4 forcing term calculated using the prior and inverse solutions:

$$|\nabla \cdot (\mathbf{u}_{\text{prior}} \eta_{\text{prior}} - \mathbf{u}_{\text{inv}} \eta_{\text{inv}})| \tag{5}$$

The result is plotted in Fig. 5 on a logarithmic scale. Note that errors in bathymetry make an additional contribution to errors in the forcing of (4), since computation of $\mathbf{u}_2$ from the transports requires dividing by H. However, for plausible bathymetry error levels (e.g., 5–10%) this additional term is not significant. Because the estimated error varies by many orders of magnitude, the peaks in Fig. 5 were truncated, and then the result was smoothed to define spatially varying variances for errors in the M_4 continuity equation. This error covariance was then used, with a Monte Carlo approach, to compute prior error variances for elevations, under the assumption that the momentum equations were exact. These were compared to the elevation prior error variances obtained under the assumption of an exact continuity equation, with the standard OTIS model for errors in the momentum equations. For the European Shelf we find that errors in the continuity equation contribute only about 10% to the elevation error, and are thus probably negligible. Given

Table 3 M_4 RMS tide-gauge misfit (cm)

	All	EnCEC	NSN	NSS	AtO	CelS	IrS
FES'04	4.43	6.70	2.10	3.91	0.61	4.34	3.73
Prior	3.96	5.20	3.55	4.09	1.81	2.68	3.58
Inv-1	7.36	14.52	0.77	1.43	0.52	0.86	1.54
Inv-2	7.33	14.46	0.76	1.43	0.52	0.87	1.54
Inv-3	6.72	13.22	0.78	1.51	0.52	0.82	1.57
Inv-4	1.99	2.62	0.79	2.49	0.52	0.99	1.57

Table 4 M_4 RMS misfit to T/P+Jason data (cm)

	All	EnCEC	NSN	NSS	AtO	CelS	IrS
FES'04	2.45	7.06	1.68	3.35	0.99	3.47	4.88
Prior	2.73	5.29	2.68	3.80	2.26	2.28	3.86
Inv-1	0.73	1.74	0.72	0.89	0.59	0.81	1.16
Inv-2	0.74	1.97	0.72	0.89	0.59	0.81	1.17
Inv-3	0.75	1.98	0.74	0.90	0.59	0.80	1.18
Inv-4	1.06	4.96	0.75	1.85	0.59	0.91	1.18

Table 5 M_4 RMS misfit to Topex2 data (cm)

	All	EnCEC	NSN	NSS	AtO	CelS	IrS
FES'04	4.45	12.62	4.59	6.51	2.69	4.17	12.34
Prior	4.58	10.60	5.10	6.65	3.38	3.62	9.19
Inv-1	3.67	7.77	4.17	5.07	2.64	2.82	8.96
Inv-2	3.68	7.86	4.17	5.07	2.70	2.79	8.98
Inv-3	3.57	7.59	4.11	4.81	2.60	2.78	8.95
Inv-4	3.71	11.37	4.11	5.21	2.60	2.78	8.95

the very approximate nature of our dynamical error covariances, using the standard OTIS covariance (with an exact continuity equation) seems well justified. These results also provide support for our comparatively simple frequency domain assimilation approach: errors associated with the nonlinear terms that couple constituents are likely less important than uncertainties in modeling dissipative processes. We experimented with combining the continuity error covariance with the standard OTIS momentum error covariance, using a range of relative scalings. Although details of the inverse solutions depend upon the covariance, solution quality (as assessed by comparison to validation data) was not clearly improved by allowing for errors in the continuity equation. For all results discussed subsequently we consider results obtained with the standard OTIS covariance.

5.6 Inversion results

A number of data assimilation experiments were carried out for the M_4 constituent on the European shelf, varying model error covariance assumptions, bathymetry, and subsets of data that were fit. Four of these will be described in some detail below, and they differ primarily in the adopted input data, as follows:

Inv-1 Uses all T/P-Jason data on original T/P tracks.
Inv-2 As in Inv-1, but deleting some data in EnCEC area.
Inv-3 As in Inv-2, plus Topex2 interleaved data.
Inv-4 As in Inv-3, plus a subset of gauge data.

In all cases we use not actual altimetric sea-surface height records but rather the previously estimated along-track M_4 harmonic constants, as discussed above in Section 4.

The four inverse solutions, the prior model obtained from time-stepping of the nonlinear equations, and the global FES-2004 M_4 solution are compared in Tables 3–5, which give RMS misfits to harmonic constants for the tide gauges, and for the two altimeter datasets.

For inverse solution Inv-1 all of the T/P-Jason along-track harmonic constants are fit. Relative to the prior solution, RMS tide gauge misfits are significantly reduced in 5 of the 6 areas by this solution (Table 3). However, misfits for the tide gauges in the English Channel and along the European coast increase dramatically, from 5.20 cm to 14.52 cm. Since this is the area with the largest M_4 amplitudes, and because increases in misfit for this area are so large, the net effect of the inversion for the full set of 158 tide gauges is an increase in overall RMS misfit. Of course, Inv-1 significantly reduces RMS misfits for the altimeter data which are fitted in the inversion (Table 4). More interestingly, Inv-1 also noticeably reduces RMS misfits for the Topex2 data set (not included in Inv-1) in all of the six areas (Table 5). Note that RMS misfits are much larger for the Topex2 dataset than for T/P + Jason, as would be expected from the relatively short duration of time series along the new ground-track.

In area EnCEC the inversion increases RMS misfits for tide gauges, but reduces misfits for independent Topex2 altimetry. This suggests some inconsistency for M_4 harmonic constants for tide gauges and altimetry in this area. To investigate this possibility we compared harmonic-constant residuals (with the prior solution subtracted) for both the tide gauges and the along-track altimetry. As might be inferred from Tables 3–5, over most of the area the tide gauge and harmonic-constant residuals are consistent; fitting one dataset can be expected to also reduce misfit in the other. An exception is in the English Channel, where we found that even the signs of the residuals disagree for the two datasets. Given these discrepancies, and considering the likelihood that altimetry data so close to the coasts might be of poor quality, we recomputed the inversion after omitting altimeter data in the EnCEC area that were obviously inconsistent with nearby tide gauges, or otherwise appeared noisy. This inverse solution, Inv-2, reduces the RMS misfit for EnCEC only slightly, and is otherwise very similar to the first case considered.

Inverting Topex2 data along with T/P and Jason data (case Inv-3 in Tables 3–5; for this and the next case we still omitted the questionable altimetry data from the English Channel) also has only a small effect on the solution. EnCEC and overall RMS misfits are slightly reduced, but remain significantly larger than those obtained for the prior solution. Several factors may contribute to the poor comparison of these inverse solutions to the EnCEC tide gauges. Obviously errors in the tide-gauge harmonic constants might

be a contributing factor, but we have little useful information regarding the quality of these data. A related possibility is that some of the coastal tide gauges are likely located in areas (e.g., in rivers or small estuaries) which are poorly represented in our relatively coarse resolution ($\frac{1}{12}^\circ$) numerical model. This difficulty might be expected to be especially severe for nonlinear tides, which can be generated very locally, and depend strongly on fine-scale bathymetric details. Obviously, any tide gauges affected by such local features are not really appropriate for validating a regional scale model of the entire European shelf.

Although it is probable that some tide gauges should be excluded from the validation set, it is more than a little disturbing that misfit levels increase so dramatically in the inverse solutions, relative to the prior. Indeed, large misfits due to poor quality or unrepresentative tide gauges should be as likely for the prior as for the inverse solutions. In fact, the substantial increase occurs largely because amplitudes in the inverse solutions become anomalously large in some local areas along the coast. In these areas prior variances are unusually large, due primarily to local dynamical resonances in the numerical model. Because prior variances are large, large changes from the prior are allowed. In these areas fitting nearby offshore data can result in large amplitudes at the coast if there are no local data available to constrain the inversion.

One obvious solution in our case is to invert some of the tide gauges, and thus better constrain the solution in the area of large misfit. For the final inversion, Inv-4, we fitted all of the EnCEC tide gauges along with all of the altimetry. This solution, not surprisingly, provides by far the best overall fit to the tide gauges, and arguably provides the best overall model for M_4 on the European shelf. However, the situation is far from completely satisfying. Fitting the subset of tide gauges leads to some degradation in the overall fit to the altimetry. Increases in misfit for both tide gauges and altimetry in the southern North Sea (NSS) are particularly disturbing; obviously Inv-4 is not the best M_4 tidal solution in this area, even if it is the best compromise of the solutions considered here. Our inability to fit all data sets together strongly suggests that our hypothesis concerning dynamical error covariances is flawed. As noted above, a number of variants on the standard OTIS covariance were tested. Although details in the misfit statistics were sensitive to the covariance used, overall results were comparable to those presented here.

Amplitude and phase for the Inv-4 inverse solution is shown in Fig. 6a, and the amplitude of the correction to the prior is shown in Fig. 6b. Significant changes with amplitude 3–5 cm are seen to occur over almost the entire domain. Larger changes of 5–8 cm occur in the North Sea and in the Celtic Sea near the western end of the English Channel. Even larger changes (over 15 cm) occur in small areas along the coast of Scotland and in the English Channel. Note that in some areas where the prior and FES2004 exhibit significant differences (Fig. 4) agreement of the inverse solution with FES2004 is much improved (e.g., around Scotland and in the open Atlantic Ocean). These are areas where tide-

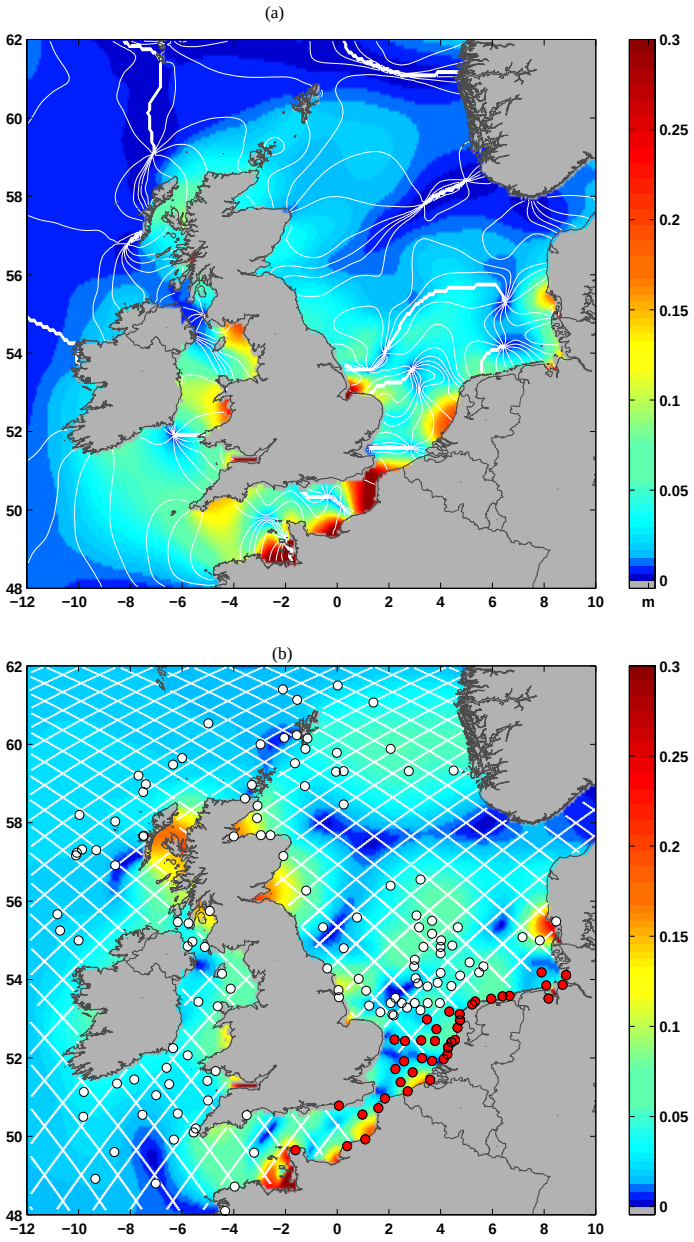

Fig. 6 (a) Inverse model Inv-4 amplitude and phase, and (b) change in complex amplitude relative to the prior. Satellite tracks are shown with white. Validation tide gauges are shown with circles; assimilated tide gauges are colored red.

gauge misfits were larger for the prior than for FES2004 (Table 3). These deficiencies in the OTIS prior were clearly corrected by assimilation of the satellite data.

6 Concluding Remarks

In this paper we have briefly reviewed some recent progress in the science of shallow-water nonlinear tides. From the pioneering work of Chabert d'Hières and Le Provost in physical modeling of the English Channel and simple limited

time-stepping models, the science has advanced to a state where even a global model such as FES2004 appears to be remarkably accurate over the European Shelf. There is, however, a strong need for powerful data-assimilation methods to supplement hydrodynamical models. In many shallow-water regions, bathymetric errors limit the development of accurate hydrodynamical models, and in regions where tide-gauge data are lacking or suspect, remote sensing via satellite altimetry provides one of the few means of constraining such models. Future progress in mapping nonlinear tides throughout the globe surely depends on further development of these methods.

One possible refinement of the variational data assimilation approach used here would be to use a more complete development of the tangent linear to the shallow water equations. To maintain the frequency domain approach required to map specific tidal constituents, this would entail inversion for all relevant linear and nonlinear constituents simultaneously. A possible advantage of a scheme such as this is that, for example, corrections to M_2 obtained by data assimilation might result in improvements to the generation terms for M_4.

Future progress also depends on extending such methods to 3D models. We have focused primarily on tidal elevations, partly because of their important geodetic applications. Yet determination of tidal currents is necessary for studying a whole range of important coastal processes, from turbulence and mixing rates to sediment and nutrient transport (e.g., Prandle 1997). Realistic simulations of shelf currents require 3D models such as those of, for example, Davies et al. (1997) for the European Shelf.

The results presented here for M_4 represent a first step toward applying generalized inverse methods (Egbert and Erofeeva 2002) to overtides or compound tides in relatively localized, shallow-water areas. Such methods have proven highly successful and adaptable for linear tides on a global or basin scale, but as we have seen, nonlinear tides in regions such as the English Channel present their own special difficulties, including aliasing and other sampling problems, as well as uncertainties about error-covariance models.

As a more general point, this study raises an important issue concerning variational tidal data assimilation in and near shallow coastal areas: absent sufficient local constraints there may be areas where assimilation of offshore altimetry or mooring data may result in inordinately large amplitudes being extrapolated to the coast. This is, in fact, an issue for all tidal constituents. It was a problem Christian Le Provost noticed, and was troubled by, when he began applying variational methods to the global FES solutions. As our results here suggest, this problem remains unresolved and deserving of further effort.

Acknowledgements We are indebted to Florent Lyard and Fabian Lefèvre for use of the FES2004 tidal solution prior to publication. We thank Philip Woodworth and several reviewers for useful comments.

References

Arbic BK (2005) Atmospheric forcing of the oceanic semidiurnal tide. Geophys Res Lett 32, L02610.

Andersen OB (1994) Ocean tides in the northern North Atlantic Ocean and adjacent seas from ERS-1 altimetry. J Geophys Res 99: 22557–22573.

Andersen OB (1999) Shallow water tides on the northwest European shelf from TOPEX/POSEIDON altimetry. J Geophys Res 104: 7729–7741.

Andersen OB, Woodworth PL, Flather RA (1995) Intercomparison of recent global ocean tide models. J Geophys Res 100(C12): 25261–25282.

Andersen OB, Knudsen P (1997) Multi-satellite ocean tide modelling—the K_1 constituent. Prog Oceanogr 40: 197–216.

Bennett AF (1992) Inverse Methods in Physical Oceanography. Cambridge Univ. Press, 346 pp.

Cartwright DE (1968) A unified analysis of tides and surges round north and east Britain. Philos Trans Royal Soc London A263: 1–55.

Cartwright DE, Ray RD (1990) Oceanic tides from Geosat altimetry. J Geophys Res 95: 3069–3090.

Cartwright DE, Zetler BD (1985) Pelagic Tidal Constants—2, Pub Sci 33, Int Assoc Phys Sci Oceans, Paris, 59 pp.

Chabert d'Hières G, Le Provost C (1970) Determination des characteristiques des ondes harmoniques M2 et M4 dans la manche sur modele reduit hydraulique. C R Acad Sci A270: 1703–1706.

Chabert d'Hières G, Le Provost C (1976) On the use of an hydraulic model to study non linear tidal deformations in shallow waters: Applications to the English Channel. Memoires de la Societe Royal de Liege 6: 113–124

Chabert d'Hières G, Le Provost C (1979) Atlas des composantes harmoniques de la marée dans la Manche, Ann Hydrographiques 6: 5–36.

Davies AM (1986) A three-dimensional model of the northwest European shelf with application to the M_4 tide. J Phys Oceanogr 16: 797–813.

Davies AM, Kwong SCM, Flather RA (1997) Formulation of a variable-function three-dimensional model, with applications to the M_2 and M_4 tide on the Northwest European continental shelf. Cont Shelf Res 17: 165–204.

Egbert GD, Bennett AF (1996) Data assimilation methods for ocean tides, in Modern Approaches to Data Assimilation in Ocean Modeling (Ed: P Malanotte-Rizzoli), pp 147–179, Elsevier, Amsterdam.

Egbert GD, Bennett AF, Foreman MGG (1994) Topex/Poseidon tides estimated using a global inverse model. J Geophys Res 99: 24821–24852.

Egbert GD, Erofeeva SY (2002) Efficient inverse modeling of barotropic ocean tides. J Atmos Oceanic Tech 19: 183–204.

Flather RA (1976) A tidal model of the north-west European continental shelf. Mem Soc R des Sci Liege (ser 6) 10: 141–164.

Heaps NS (1978) Linearized vertically integrated equations for residual circulation in coastal seas. Dt Hydrogr Z 31: 147–169.

Howrath MJ, Pugh DT (1983) Observations of tides over the continental shelf of northwest Europe, in Physical Oceanography of Coastal and Shelf Seas, (ed: D Johns), pp. 135-185, Elsevier, New York.

Hunter JR (1979) On the interaction of M_2 and M_4 tidal velocities in relation to quadratic and higher power laws. Dtsch Hydrogr Z 22: 146–153.

Jones JE, Davies AM (1996) A high-resolution, three-dimensional model of the M_2, M_4, M_6, S_2, N_2, K_1, and O_1 tides in the Irish Sea. Estuarine Coastal Shelf Sci 42: 311–346.

Kabbaj A, Le Provost C (1980) Non-linear tidal wavel in channels: a pertubation method adapted to the importance of quadratic bottom friction. Tellus 32: 143–163.

Kwong SCM, Davies AM, Flather RA (1997) A three-dimensional model of the principal tides on the European Shelf. Prog Oceanogr 39: 205–262.

Le Provost C (1973) Décomposition spectrale du terme quadratique de frottement dans les équations des marées littorales. C R Acad Sci 276A, 571–574 and 653–656.

Le Provost C (1974) Contribution à l'etude des marées dans les mers littorales. Application à la Manche. Thèse d'état, Grenoble, 228 pp.

Le Provost C (1976) Theoretical analysis of the tidal wave spectrum in shallow water areas. Mem S Roy Sci Liege 10: 97–111.

Le Provost C, Poncet A (1977) Sur une méthode numérique pour calculer les marées océaniques et littorales. C R Acad Sci B285, 349–352.

Le Provost C, Poncet A (1978) Finite-element method for spectral modeling of tides. Int J Num Meth Eng 12: 853–871.

Le Provost C, Rougier G, Poncet A (1981) Numerical modelling of the harmonic constituents of the tides with aplication to the English Channel. J Phys Oceanog 11: 123–138.

Le Provost C, Fornerino M (1985) Tidal spectroscopy of the English Channel with a numerical model. J Phys Oceanog 15: 1009–1031.

Le Provost C (1991) Generation of overtides and compound tides (review), in: Tidal Hydrodynamics (Ed: B. B. Parker), John Wiley, pp 263–295.

Le Provost C (2001) Chapter 6: Ocean Tides, in: Satellite Altimetry and Earth Sciences (Eds: L.-L. Fu and A. Cazenave), Academic Press, London, pp 267–304.

Munk WH Cartwright DE (1966) Tidal spectroscopy and prediction. Philos Trans R Soc London A259: 533–583.

Parke ME, Stewart RH, Farless DL, Cartwright DE (1987) On the choice of orbits for an altimetric satellite to study ocean circulation and tides. J Geophys Res 92(C11): 11693–11707.

Parker BB (1991) The relative importance of the various nonlinear mechanisms in a wide range of tidal interactions (review), in: Tidal Hydrodynamics (Ed: B B Parker), John Wiley, pp 263–295.

Pingree RD, Maddock L (1978) The M_4 tide in the English Channel derived from a nonlinear numerical model of the M_2 tide. Deep-Sea Res 25: 52–63.

Ponchaut F, Lyard F, Le Provost C (2001) An analysis of the tidal signal in the WOCE sea level dataset. J Atmos Oceanic Tech 18: 77–91.

Prandle D (1997) Tidal currents in shelf seas—their nature and impacts. Progr Oceanogr 40: 245–261.

Ray RD (1997) Spectral analysis of highly aliased sea-level signals. J Geophys Res 103: 24991–25003.

Shum CK, Woodworth PL, Andersen OB, Egbert GD, Francis O, King C, Klosko SM, Le Provost C, Li X, Molines JM, Parke M, Ray RD, Schlax M, Stammer D, Tierney C, P. Vincent P, Wunch C (1997) Accuracy assesment of recent ocean tide models. J Geophys Res 102(C11): 25173–25194.

Sinha B, Pingree RD (1997) The principal lunar semdiurnal tide and its harmonics: baseline solutions for M_2 and M_4 constituents on the northwest European continental shelf. Cont Shelf Res 17: 1321–1365.

Smithson MJ (1992) Pelagic Tidal Constants—3. IAPSO Pub Sci 35, 191 pp.

Walters RA (1987) A numerical model for tides and currents in the English Channel and the southern North Sea. Adv Water Res 10: 138–148.

Walters RA, Werner FE (1991) Nonlinear generation of overtides, compound tides, and residuals, in Tidal Hydrodynamics (Ed: BB Parker), pp 297–320, John Wiley, New York.

Werner FE, Lynch DR (1989) Harmonic structure of English Channel/Southern Bight tides from a wave equation simulation. Adv Water Res 12: 121–142.

Zelter BD, Cummings RA (1967) A harmonic method for predicting shallow-water tides. J Mar Res 25: 103–114.