

Ocean Dynamics

**Theoretical
Computational and
Observational
Oceanography**

**Special Issue
Multi-Scale Modelling:
Nested Grid and Unstructured Mesh Approaches**



 Springer

Chief Editor

Jörg-Olaf Wolff
 ICBM
 Physical Oceanography (Theory)
 Universität Oldenburg
 Postfach 2503
 26111 Oldenburg, Germany
 Telephone: +49-441-798-5343
 Fax: +49-441-798-3404
 e-mail: wolff@icbm.de

Special Issue Editors

Eric Deleersnijder
 Université Catholique de Louvain
 Louvain-la-Neuve, Belgium
 ericd@uclouvain.be

Pierre Lermusiaux
 Massachusetts Institute of Technology
 Cambridge, MA, USA
 pierrel@mit.edu

Editors

Jean-Marie Beckers
 University Liège, Liège, Belgium
 jm.beckers@ulg.ac.be

Franciscus Colijn
 GKSS, Geesthacht, Germany
 franciscus.colijn@gkss.de

Eric Deleersnijder
 Université Catholique de Louvain
 Louvain-la-Neuve, Belgium
 ericd@uclouvain.be

Joachim Dippner
 Institut für Ostseeforschung
 Warnemünde
 Rostock, Germany
 joachim.dippner@io-warnemuende.de

Tal Ezer
 Ocean, Earth & Atmospheric Sciences,
 Old Dominion University,
 Norfolk, USA
 tezer@odu.edu

Richard J. Greatbatch
 IFM-GEOMAR
 Kiel, Germany
 rgreatbatch@ifm-geomar.de

Tony Hirst
 CSIRO Marine and Atmospheric
 Research,
 Aspendale, Australia
 tony.hirst@csiro.au

Birgit Klein
 Bundesamt für Seeschifffahrt
 und Hydrographie,
 Hamburg, Germany
 birgit.klein@bsh.de

Rosemary Morrow
 LEGOS/GRGS,
 Toulouse, France
 rosemary.morrow@cnes.fr

Dirk Olbers
 Alfred-Wegener-Institut für Polar-
 und Meeresforschung,
 Bremerhaven, Germany
 dirk.olbers@awi.de

Andreas Oschlies
 IFM-GEOMAR
 Kiel, Germany
 aoschlies@ifm-geomar.de

Roger Proctor
 University of Tasmania,
 Hobart, Australia
 roger.proctor@utas.edu.au

Stephen R. Rintoul
 CSIRO Marine Atmospheric Research,
 Hobart, Australia
 steve.rintoul@csiro.au

Richard Signell
 U.S. Geological Survey,
 Woods Hole, USA
 rsignell@usgs.gov

Alejandro J. Souza
 Proudman Oceanographic
 Laboratory,
 Liverpool, United Kingdom
 ajso@pol.ac.uk

Jin-Song von Storch
 Max-Planck-Institut für Meteorologie,
 Hamburg, Germany
 jin-song.von.storch@zmaw.de

John Wilkin
 Institute of Marine and Coastal Sciences,
 Rutgers University,
 New Brunswick, USA
 wilkin@imcs.rutgers.edu

Han Winterwerp
 WL/Delft Hydraulics
 Delft, The Netherlands
 han.winterwerp@wldelft.nl

Instructions for authors:

Please go to the journal's website:
www.springer.com/10236

Table of Contents

- Deleersnijder E. and P.F.J. Lermusiaux, 2008, *Editorial - Multi-Scale Modeling: Nested-Grid and Unstructured-Mesh Approaches*, *Ocean Dynamics*, 58, 335-336
- Jones J.E. and A.M. Davies, 2008, Storm surge computations for the west coast of Britain using a finite element model (TELEMAC), *Ocean Dynamics*, 58, 3378-363
- Danilov S., Q. Wang, M. Losch, D. Sidorenko and J. Schröter, 2008, Modeling ocean circulation on unstructured meshes: comparison of two horizontal discretizations, *Ocean Dynamics*, 58, 365-374
- Yang B. and J. Sheng, 2008, Process study of a coastal circulation over the inner Scotian Shelf using a nested-grid ocean circulation model, with a special emphasis on the storm-induced circulation during tropical storm Alberto in 2006, *Ocean Dynamics*, 58, 375-396
- Bellafiore D., G. Umgiesser and A. Gucco, 2008, Modeling the water exchanges between the Venice Lagoon and the Adriatic Sea, *Ocean Dynamics*, 58, 397-413
- Debreu L. and E. Blayo, 2008, Two-way embedding algorithms: a review, *Ocean Dynamics*, 58, 415-428
- Harig S., Chaeroni, W.S. Pranowo and J. Behrens, 2008, Tsunami simulations over several scales - Comparison of approaches with unstructured meshes and nested grids, *Ocean Dynamics*, 58, 429-440
- Logutov O., 2008, A multigrid methodology for assimilation of measurements into regional tidal models, *Ocean Dynamics*, 58, 441-460
- Lambrechts J., R. Comblen, V. Legat, C. Geuzaine and J.-F. Remacle, 2008, Multiscale mesh generation on the sphere, *Ocean Dynamics*, 58, 461-473
- Ringler T., L. Ju and M. Gunzburger, 2008, A multiresolution method for climate system modeling: application of spherical centroidal Voronoi tessellations, *Ocean Dynamics*, 58, 475-498

Multi-scale modeling: nested-grid and unstructured-mesh approaches

Eric Deleersnijder · Pierre F. J. Lermusiaux

Received: 11 November 2008 / Accepted: 11 November 2008 / Published online: 27 November 2008
© Springer-Verlag 2008

In 1969, the *Journal of Computational Physics* published a seminal article by K. Bryan presenting the first ocean general circulation model. Since then, many numerical studies of the World Ocean, as well as regional or coastal flows, used models directly or indirectly inspired by the work of Bryan and his colleagues. A number of these models have evolved into highly modular and versatile computational systems, including multiple physical modules and options as well as varied biogeochemical, ecosystem and acoustics modeling capabilities. Several modeling systems are now well-documented tools, which are widely used in research institutions and various organizations around the world. The list of such modeling systems is large and too long to be summarized in this editorial.

Over the last three decades, significant progress has been made in the parameterization of subgrid-scale processes, in data assimilation methodologies and in boundary condition schemes, as well as in the efficient implementation of algorithms on fast vector and subsequently parallel computers, allowing higher and higher resolution in space and time. However, many of today's popular modeling systems

can still be regarded as members of the first generation of ocean models: at their core, rather similar geophysical fluid dynamics equations are solved numerically using a conservative finite-difference method on a structured grid.

Today, several aspects of structured-grid models could benefit from significant upgrades, learning from major advances in computational fluid dynamics. In particular, the use of a structured grid limits the flexibility in the spatial resolution and does not allow one to take full advantage of numerical algorithms such as finite volumes and finite elements, which can achieve their best performance when implemented on unstructured meshes.

Even though many of today's complex marine modeling and data assimilation systems have evolved significantly since Bryan's prototype, it would be challenging to modify them step-by-step from a structured-grid approach to an unstructured-grid one. Therefore, novel marine model design research is underway, paving the way for the second generation of ocean modeling systems. It is difficult to predict today if this new generation of ocean models will achieve its chief objective: widening the range of resolved scales of motion with increased efficiencies and accuracies, possibly allowing multi-resolution, multi-scale, and multi-dynamics numerical simulations of marine flows, all occurring seamlessly within distributed computing environments. In fact, hybrid approaches merging the advantages of structured and unstructured-grid modeling may be the way forward.

Whether or not unstructured mesh approaches will prevail is all the more difficult to predict now that structured mesh modelers have developed powerful solutions for increasing the resolution when and where needed. For instance, grid embedding is still a popular and useful method for enhancing model resolution. It can involve multiply nested domains and allows the relatively

Responsible editor: Jörg-Olaf Wolff

E. Deleersnijder (✉)
Université catholique de Louvain,
Centre for Systems Engineering and Applied Mechanics,
4 Avenue G. Lemaître,
B-1348 Louvain-la-Neuve, Belgium
e-mail: ericd@uclouvain.be

P. F. J. Lermusiaux
Massachusetts Institute of Technology,
Department of Mechanical Engineering,
Room 5-207, 77 Massachusetts Avenue,
Cambridge, MA 02139-4307, USA
e-mail: pierrel@mit.edu

straightforward use of different dynamics or models in each domain. Research is also underway for developing multi-grid, wavelet, and other multi-scale decompositions for the numerical solution of dynamical equations but also for the study of results, model evaluation or data assimilation.

This special issue presents a number of examples of the abovementioned developments. Ringler et al. examine the potential of spherical centroidal Voronoi tessellations for performing multi-resolution simulations; they apply this method to the Greenland ice sheet and the North Atlantic Ocean. Lambrechts et al. present a triangular mesh generation system and its applications to the World Ocean and various shelf seas, including the Great Barrier Reef, Australia. Finite element models on unstructured grids are described and utilized in several manuscripts. Bellafiore et al. study the Adriatic Sea and the Lagoon of Venice, while Jones and Davies simulate tides and storm surges along the western coast of Britain. Danilov et al. assess two finite element discretizations, i.e., a continuous element and a non-conforming one, and compare the results of these discretizations with those of a finite-difference model. In Harig et al., the tsunami generated by the great Sumatra–Andaman earthquake of 26 December 2004 is simulated by means of a finite element model. Comparisons are carried out with various types of data as well as with the results of a structured mesh model using a nested structured-grid system. A nested-grid ocean circulation model is also employed by Yang and Sheng to carry out a process study

on the Inner Scotian Shelf, Canada, focusing on the circulation induced by a tropical storm. Debreu and Blayo present a detailed review of two-way embedding algorithms for structured-grid models. Finally, Logutov develops a multi-scale assimilation scheme for tidal data within the framework of a multiply nested structured-grid barotropic tidal modeling approach.

As illustrated by these manuscripts, the next generation of ocean modelers is motivated by a wide range of research opportunities over a rich spectrum of needs. Future progress will involve fundamental and applied numerical and computational research as well as new multi-scale geophysical fluid modeling. Domains of ongoing interest range from estuaries to the global ocean, including coastal regions and shelf seas. New multi-scale modeling of physical as well as biological, chemical or interdisciplinary processes will flourish in the coming decades.

We are grateful to the authors for their contributions and to the chief-editor for his support in this endeavor. We are thankful to the reviewers for their time and help in assessing the manuscripts submitted to this special issue. Eric Deleersnijder is a Research associate with the Belgian National Fund for Scientific Research (FNRS); he is indebted to the Communauté Française de Belgique for its support through contract ARC 04/09-316. Pierre Lermusiaux is grateful to the Office of Naval Research for support under grant N00014-08-1-1097 to the Massachusetts Institute of Technology.

Storm surge computations for the west coast of Britain using a finite element model (TELEMAC)

John Eric Jones · Alan M. Davies

Received: 25 April 2008 / Accepted: 18 June 2008 / Published online: 17 July 2008
© Springer-Verlag 2008

Abstract An unstructured mesh finite element model of the sea region off the west coast of Britain is used to examine the storm surge event of November 1977. This period is chosen because accurate meteorological data to drive the model and coastal observations for validation purposes are available. In addition, previous published results from a coarse-grid (resolution 7 km) finite difference model of the region and high-resolution (1 km) limited area (namely eastern Irish Sea) model are available for comparison purposes. To enable a “like with like” comparison to be made, the finite element model covers the same domain and has the same meteorological forcing as these earlier finite difference models. In addition, the mesh is based on an identical set of water depths. Calculations show that the finite element model can reproduce both the “external” and “internal” components of the surge in the region. This shows that the “far field” (external) component of the surge can accurately propagate through the irregular mesh, and the model responds accurately, without over- or under-damping, to local wind forcing. Calculations show significant temporal and spatial variability in the surge in close agreement with that found in earlier finite difference calculations. In addition, root mean square errors between computed and observed surge are comparable to those found in previous finite difference calculations. The ability to vary the mesh in nearshore regions reveals appreciable small-scale variability that was not found in the previous finite difference solutions. However, the requirement to

perform a “like with like” comparison using the same water depths means that the full potential of the unstructured grid model to improve resolution in the nearshore region is inhibited. This is clearly evident in the Mersey estuary region where a higher resolution unstructured mesh model, forced with uniform winds, had shown high topographic variability due to small-scale variations in topography that are not resolved here. Despite the lack of high resolution in the nearshore region, the model showed results that were consistent with the previous storm surge models of the region. Calculations suggest that to improve on these earlier results, a finer nearshore mesh is required based upon accurate nearshore topography.

Keywords Storm surge · Finite element model · UK West coast

1 Introduction

Following on from the major flooding caused by the 1953 UK East coast storm surge event, the main focus in storm surge modelling has been the prediction of storm surge elevations on the European shelf. Of particular importance has been their accurate prediction in shallow coastal regions where flooding can occur during major storms. In regions such as the Irish Sea and North Sea, early research showed that limited area finite difference models failed to reproduce the observed surge due to their neglect of shelf wide wind events (the external surge). Consequently, early storm surge computations were performed with shelf-wide finite difference models (e.g. Davies and Flather 1977) that necessarily used coarse grids due to computational limitations; hence, their resolution was poor in coastal regions. Since the main objective of these models was the computation of surge

Responsible Editor: Eric Deleersnijder

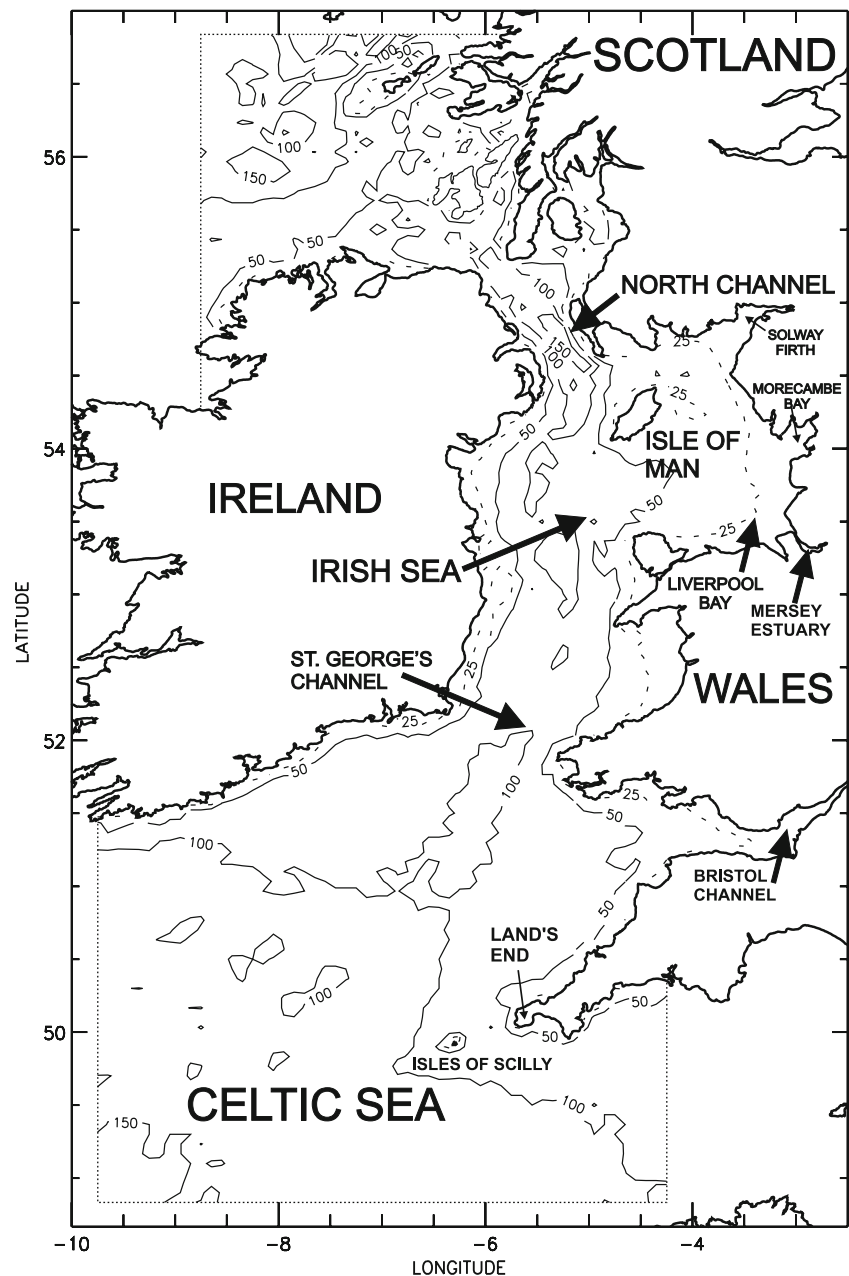
J. E. Jones (✉) · A. M. Davies
Proudman Oceanographic Laboratory,
6 Brownlow Street,
Liverpool L3 5DA Merseyside, UK
e-mail: jej@pol.ac.uk

elevations, the two-dimensional hydrodynamic equations were used.

With enhancements in computing power, finite difference grids in shelf-wide models were refined to the order of 12 km (Davies et al. 1998, 2000) and local (e.g. west coast of Britain) models of grid resolution 7 km (Davies and Jones 1992a, hereafter DJ92) or less (Jones and Davies 1996, 1998, hereafter JD98) were developed. However, such local models required open boundary input from a coarser shelf-wide model or a model which could account for changes produced by shelf-wide winds. For example, DJ92 used a 7-km resolution west coast of Britain model to simulate the November 1977 major surge event. For this

simulation, far field effects were taken into account using observations along the open boundaries. However, results showed that the 7-km grid of this model was not sufficiently fine to accurately resolve the local increase in storm surge elevation in the eastern Irish Sea. Consequently, in subsequent work, this model was used to provide boundary conditions for a limited area eastern Irish Sea model of 1-km resolution, resulting in improved surge accuracy in the region (JD98). A similar approach of using a high-resolution (of order 1 km) limited area model of the North Channel of the Irish Sea, forced by a coarse-grid large area model, was used by Davies et al. (2001) for a detailed study of tidal and wind forced currents in that area.

Fig. 1 Water depths in the region covered by the model and places named in the text

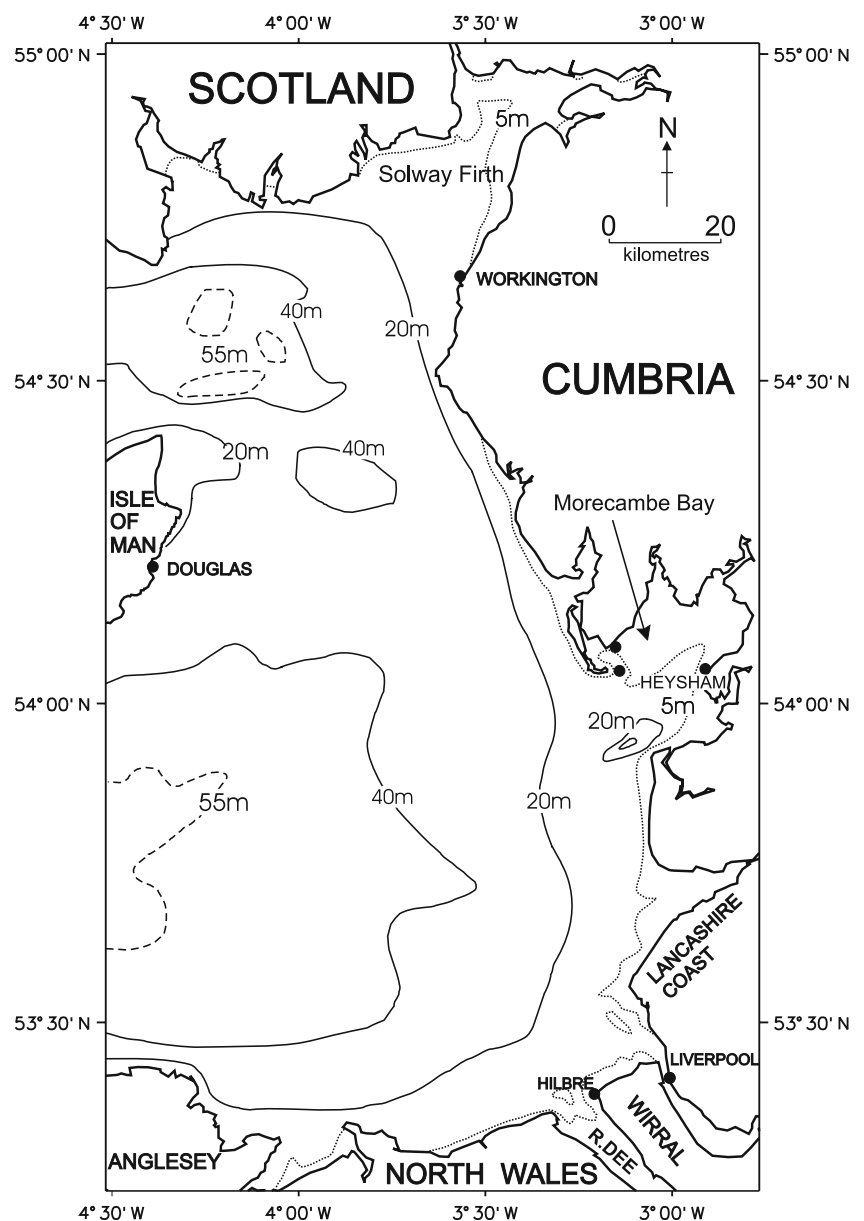


Although this approach of nesting a high-resolution limited area finite difference model within a coarser grid model enables a local improvement in resolution, if the nesting is two-way, then there may be problems at the interface between the two grids. In addition, as shown by Davies and Hall (2002) in regions of rapidly changing current magnitude and direction produced by local changes in topography, a nesting approach could give rise to significant errors in the currents. An alternative approach, the method examined here, is to use a finite element technique in which the grid resolution varies gradually in space. By this means, a coarse mesh can be used in offshore regions where the water is deep and surge elevations and currents show little spatial variability. As the water shallows and surge intensity and spatial variability increase in the

nearshore regions, then the grid is refined in these regions. In addition, surge propagation into estuaries can be readily accomplished without nesting.

The finite element model with its ability to refine the mesh in nearshore regions has been very successful in a number of problems (e.g. Werner 1995; Ip et al. 1998; Jones 2002; Fernandes et al. 2002, 2004; Walters 2005; Levasseur et al. 2007; Nicolle and Karputchev 2007). Although in theory the grading of the mesh is arbitrary, in practice, the computation of an optimal mesh is complex (e.g. Greenberg et al. 2007; Legrand et al. 2006, 2007; Hagen et al. 2001, 2002) and in the calculations presented here, in order to make rigorous comparisons with earlier finite difference and finite element models, no attempt was made to produce an optimal mesh.

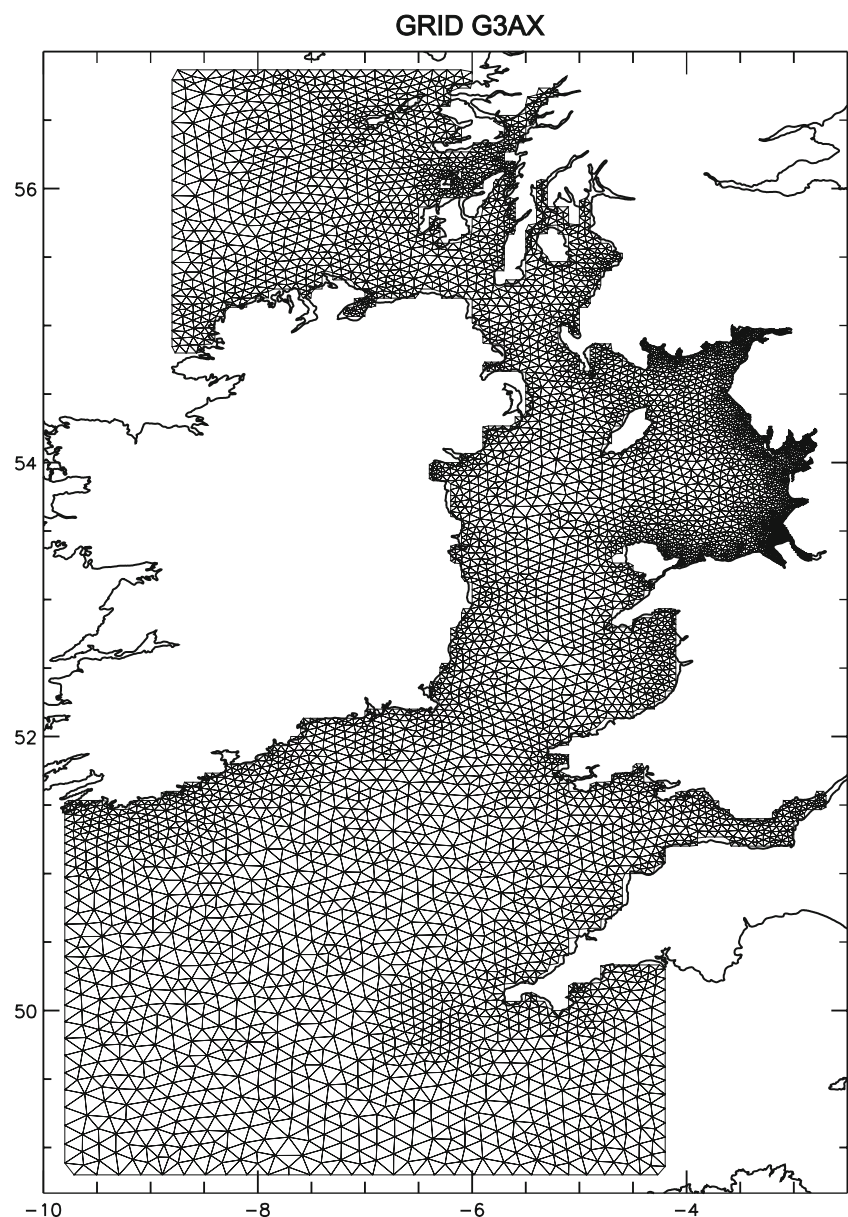
Fig. 2 Detailed topography of the eastern Irish Sea and location of tide gauges



Previous calculations using a finite element (TELEMAC) model of the west coast of Britain showed that it could reproduce the dominant (M_2 component) tide in the region (Jones and Davies 2005) to an accuracy comparable to an existing finite difference model (Davies and Jones 1992b). Furthermore, recent calculations showed (Jones and Davies 2007a) that the model was comparable in accuracy to both west coast (Davies and Jones 1992b) and high-resolution eastern Irish Sea models (Jones and Davies 1996) at reproducing the M_2 , S_2 , N_2 , K_1 and O_1 components of the tide. In addition, by refining the element size in the eastern Irish Sea region, the higher harmonics of the tide could be accurately reproduced (Jones and Davies 2007b). Furthermore, tidal residual currents in the region could be accurately simulated (Jones and Davies 2007c), and the

artificial flow in the nearshore region due to the “stair case” representation of the coast in finite difference models (Davies and Jones 1996) was not present. Recent calculations (Jones and Davies 2006) using this finite element model of the west coast of Britain showed that its response to steady orthogonal wind forcing was consistent to that found with well-established and proven finite difference models (Jones and Davies 2003a, b). In addition, the finer Eastern Irish Sea resolution in the finite element model showed small-scale wind-induced circulation features that were not present in the finite difference model (Jones and Davies 2003a, b). These calculations suggest that the finite element model developed previously should be able to reproduce storm surges in the eastern Irish Sea to an accuracy comparable to the nested high-resolution (1 km)

Fig. 3 Finite element grid used in the calculations



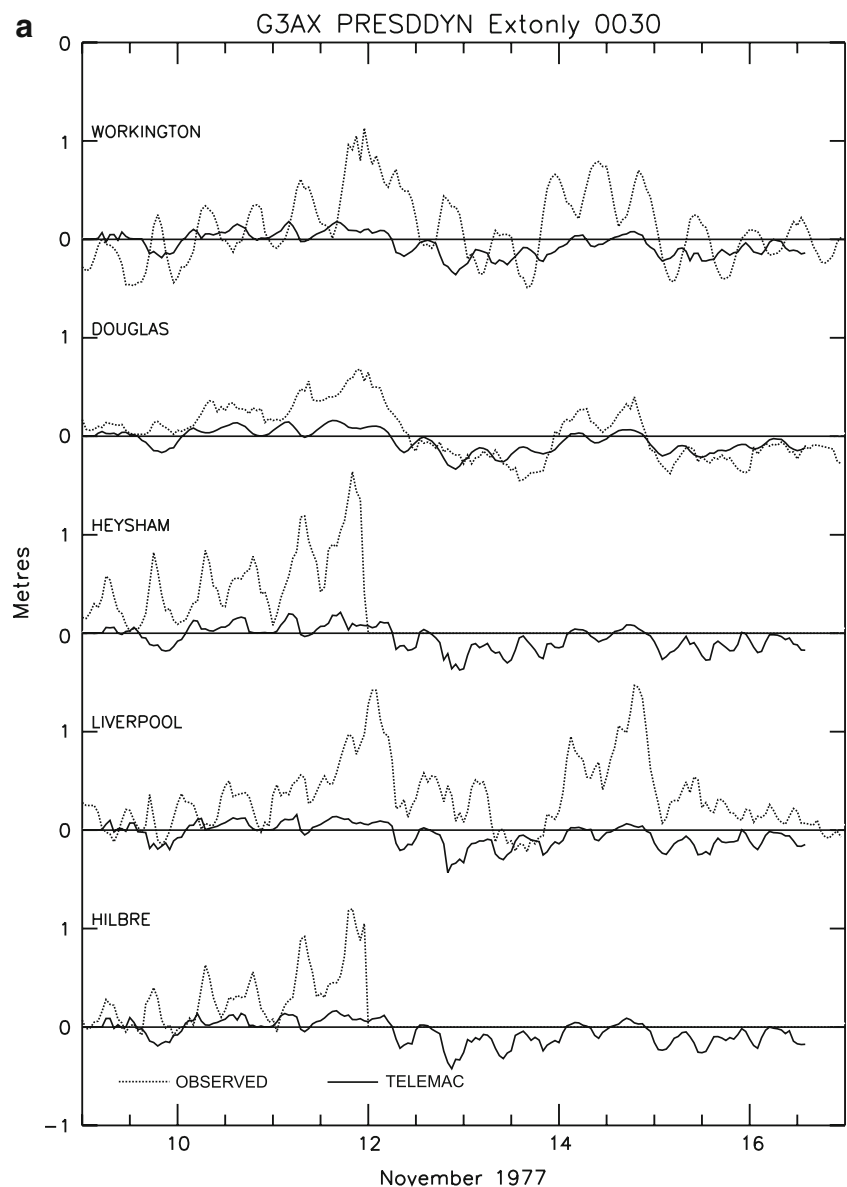
eastern Irish Sea and coarser (7 km) finite difference west coast models used previously.

Besides investigating the response of the west coast of Britain to wind forcing, the objective of these earlier calculations, namely Jones and Davies (2003a, 2006, 2008), was to examine the processes influencing surges in the region and the role of tide–surge interaction. However, the objective of this paper was to use the previous west coast finite element model, which has been validated against a range of tidal constituents, to examine the mechanisms (namely external and internal surge generation) producing the storm surge of November 1977. In addition, because this surge event has been computed with a range of finite difference models, the relative accuracy of the finite element model can be examined by comparing

with observations and finite difference solutions. To ensure that this is a meaningful comparison, the same regional extent, open boundary forcing and meteorological forcing, to that used previously (DJ92, JD98) was applied. In addition, the topography used in the model was identical over the Irish and Celtic Seas to that used in DJ92 and in the eastern Irish Sea to that in JD98. By this means, a rigorous comparison with these earlier finite difference models could be performed. The surge of November 1977 was chosen because an accurate meteorological data set was available with which to force the model.

The finite element model is discussed in the next section, with following sections describing the meteorological forcing and detailed model/data comparisons. A final section summarises the main results.

Fig. 4 Time series of computed surge at a number of eastern Irish Sea locations computed with **a** “external” surge introduced through the open boundary, **b** “internal” surge due to local meteorological forcing and **c** the “total” surge



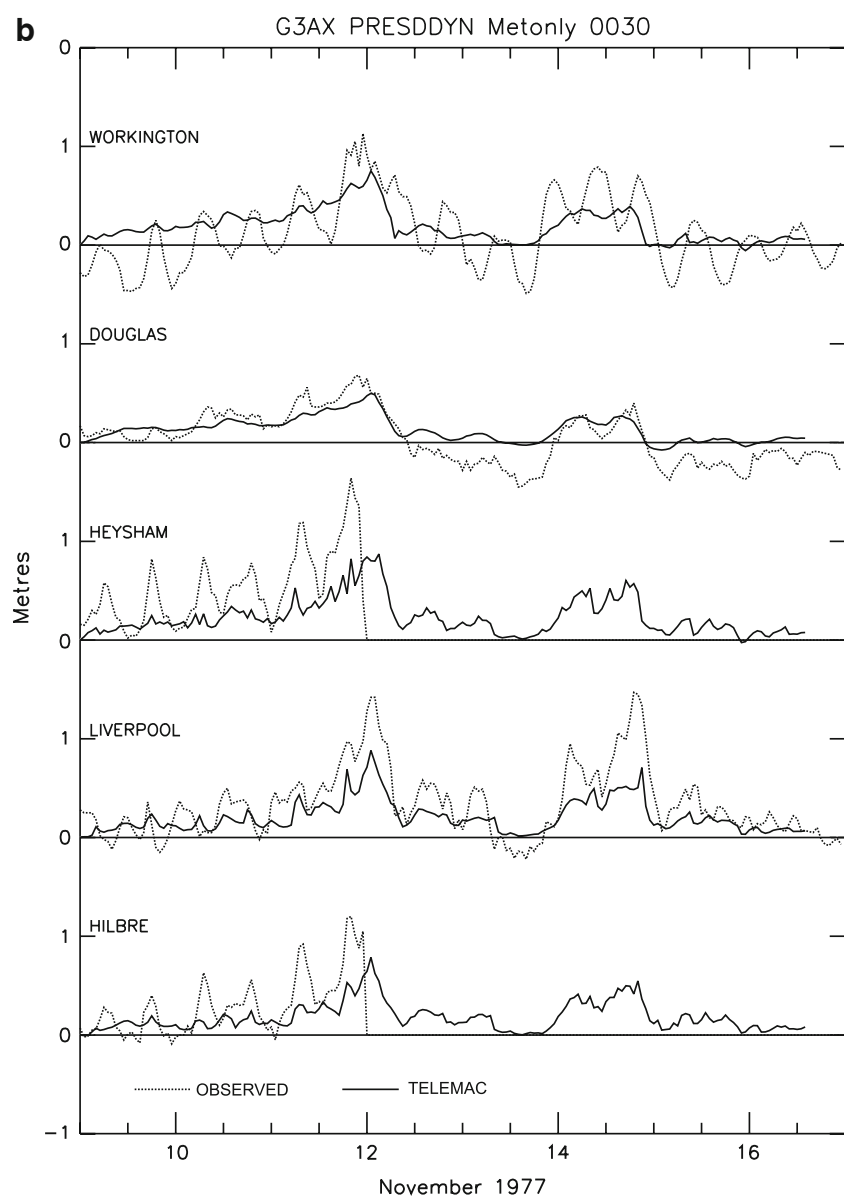
2 The finite element model and forcing

Since the focus of the paper is the application of a finite element model to the prediction of surge elevations in the Irish Sea for the major storm events in November 1977, it is sufficient to solve the two-dimensional vertically integrated hydrodynamic equations. However, since the region (Fig. 1) spans a range of latitudes, spherical coordinates were used as in earlier finite difference models. As the form of the nonlinear hydrodynamic equations using these coordinates is given elsewhere (DJ92, JD98), they will not be repeated here. As details of the numerical methods used in TELEMAC to solve the hydrodynamic equations have been reviewed in Jones and Davies (2006, 2007a, b, c) and references therein, they will not be repeated here. In

order to compare results with previous finite difference solutions, the region covered by the model was identical to that used in DJ92. In addition, the water depth distribution was based on DJ92, with the addition of more accurate water depths and coastal resolution in the eastern Irish Sea taken from JD98.

The water depth distribution in the region is characterised by depths of the order of 100 m in the Celtic Sea, deepening to 150 m at the south west of this region (Fig. 1). Within the Irish Sea on its western side, there is a deep channel, with water depths up to 100 m, although on its eastern side, the water is much shallower (less than 50 m) with extensive nearshore regions where the water depth is below 25 m (Fig. 2). In these nearshore regions, “wetting and drying” occurs over the tidal cycle. Water depths in the

Fig. 4 (continued)



North Channel can exceed 150 m (Fig. 1), with average depths to the north of this of order 100 m.

As tidal friction and tide–surge interaction are important in the region, the five dominant tidal constituents, namely M_2 , S_2 , N_2 , K_1 and O_1 , were included within all the calculations as input along the open boundary. This is consistent with the finite difference solution given in DJ92. It is important to note that this tidal forcing was used in all calculations, namely those including the computation of the external, internal and total surge (see later). Previous tidal calculations (Jones and Davies 2007a) showed that the finite element grid used in the present calculations (Fig. 3) could accurately reproduce these constituents and the associated higher harmonics produced by nonlinear interaction in the region. Although in the eastern Irish Sea storm surge calculations of JD98, using a limited area high-

resolution (1 km) model of the region, the input storm surge on the open boundary was taken without adjustment from the coarser (7 km) west coast model, the tidal input along the boundary was adjusted to give an accurate representation of the tide in the region. In the present finite element calculation, no tidal or storm surge adjustments were made to try and improve model accuracy, nor was the topography modified from that used in JD98.

To be consistent with DJ92 and JD98, identical meteorological forcing for the period 7–17 November 1977 was used, and the surge was computed by subtracting a tide-only solution from one involving tidal and meteorological forcing. As a detailed discussion of the meteorological forcing is presented in DJ92, it will not be given here. However, as shown in DJ92, the surge within the region covered by the west coast model (Fig. 3) is influenced by

Fig. 4 (continued)

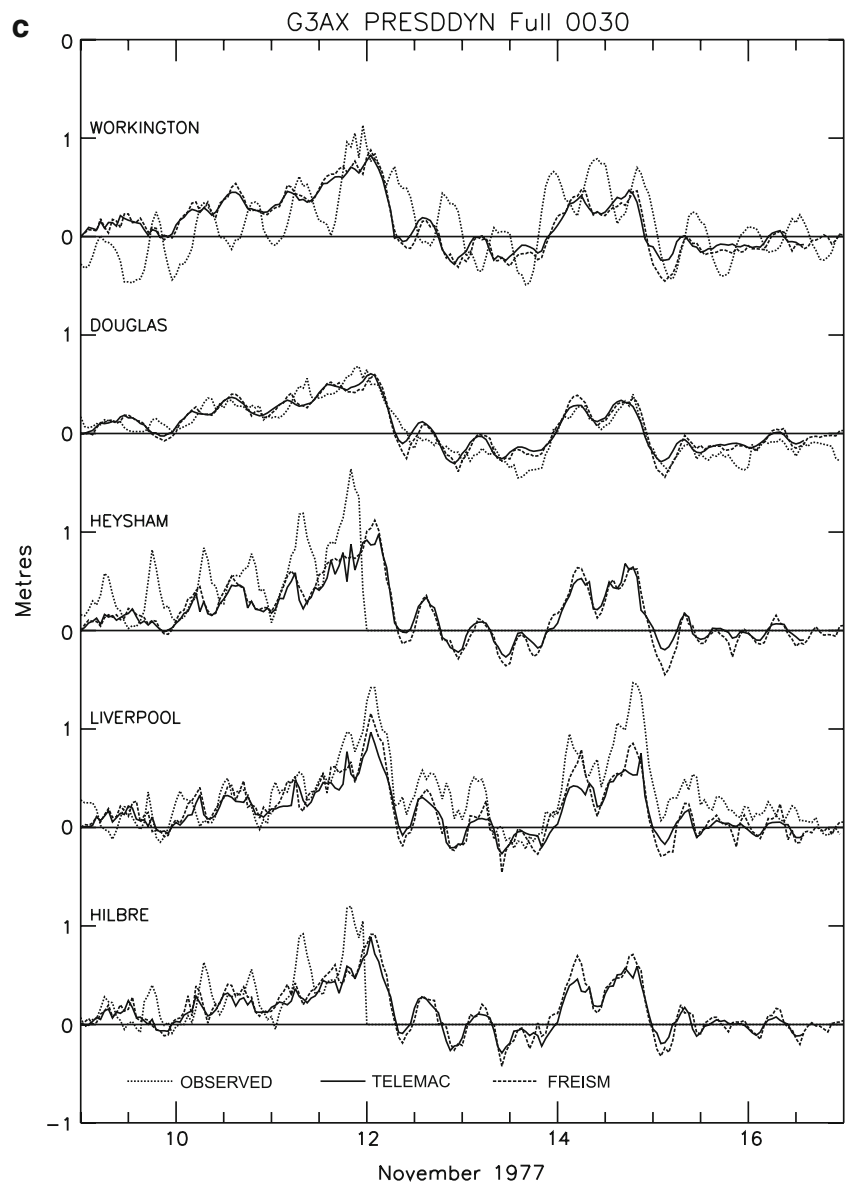


Table 1 Root mean square errors (cm) from the finite difference model (FD calc; Jones and Davies 2001) and finite element model (FE calc) at Liverpool and Douglas

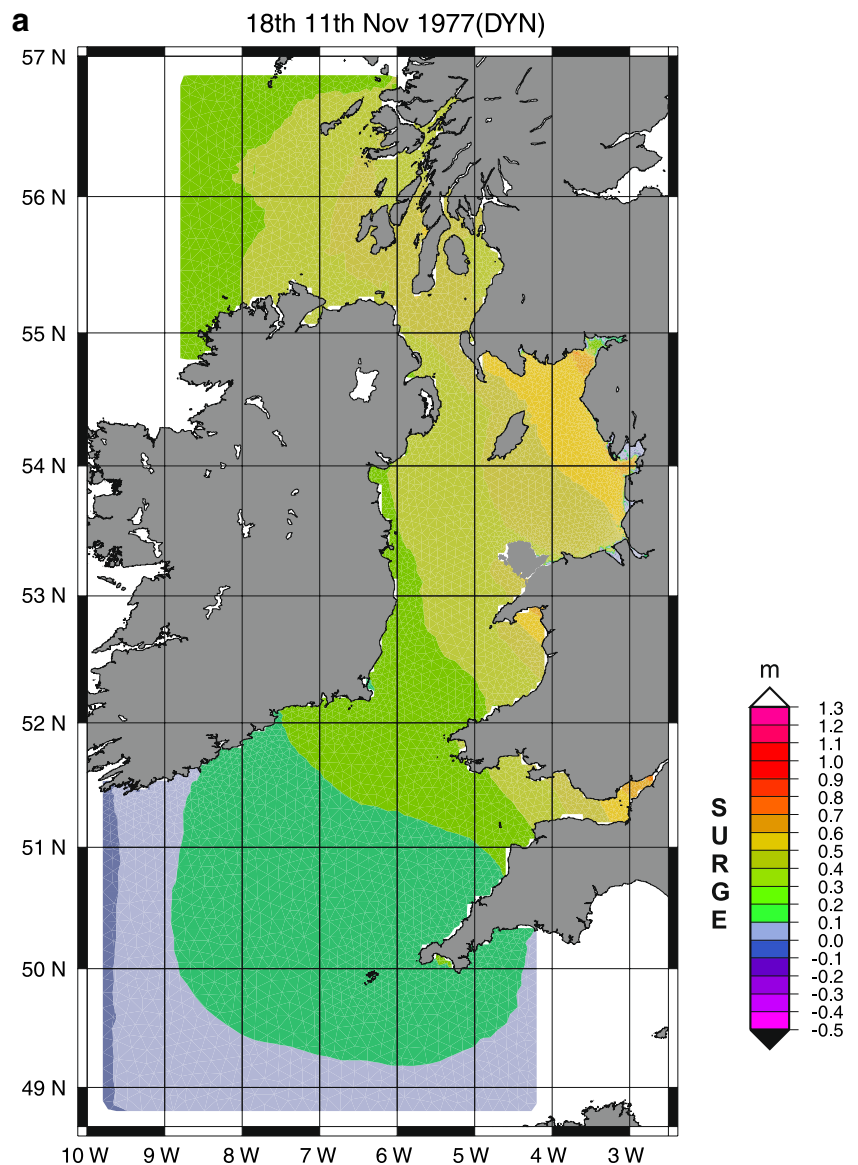
Calc	Port	Period			Total
		1	2	3	
FD	Liverpool	17.0	19.1	22.7	22.9
	Douglas	10.6	11.3	10.6	10.3
FE	Liverpool	21.5	23.2	28.9	27.7
	Douglas	10.0	11.8	11.2	11.0

both local winds and “far field” effects produced by wind forcing outside the model and the resulting flow into the region. Consequently, as in DJ92, it is necessary to take account of this external effect along the open boundary of

the model. To be consistent with DJ92, this was accomplished by linearly interpolating observed surge elevations from Castletownsend and Newlyn along the southern boundary of the model. In addition, observations from Malin were imposed along the northern boundary of the model.

To understand the influence of the external surge, and local wind-forced surge, upon the total surge, three calculations were performed with the finite element model using the grid given in Fig. 3. Initially (calc 1), the model was run with only boundary forcing to determine far field effects (external surge). Subsequently (calc 2), only local meteorological forcing was applied (internal surge), and finally (calc 3), the full surge was determined.

Fig. 5 Elevation contours over the whole region at **a** 1800 hours 11 Nov, **b** 0000 hours 12 Nov, **c** 0600 hours 12 Nov, **d** 1200 hours 14 Nov and **e** 1800 hour 14 Nov



3 Storm surge calculations

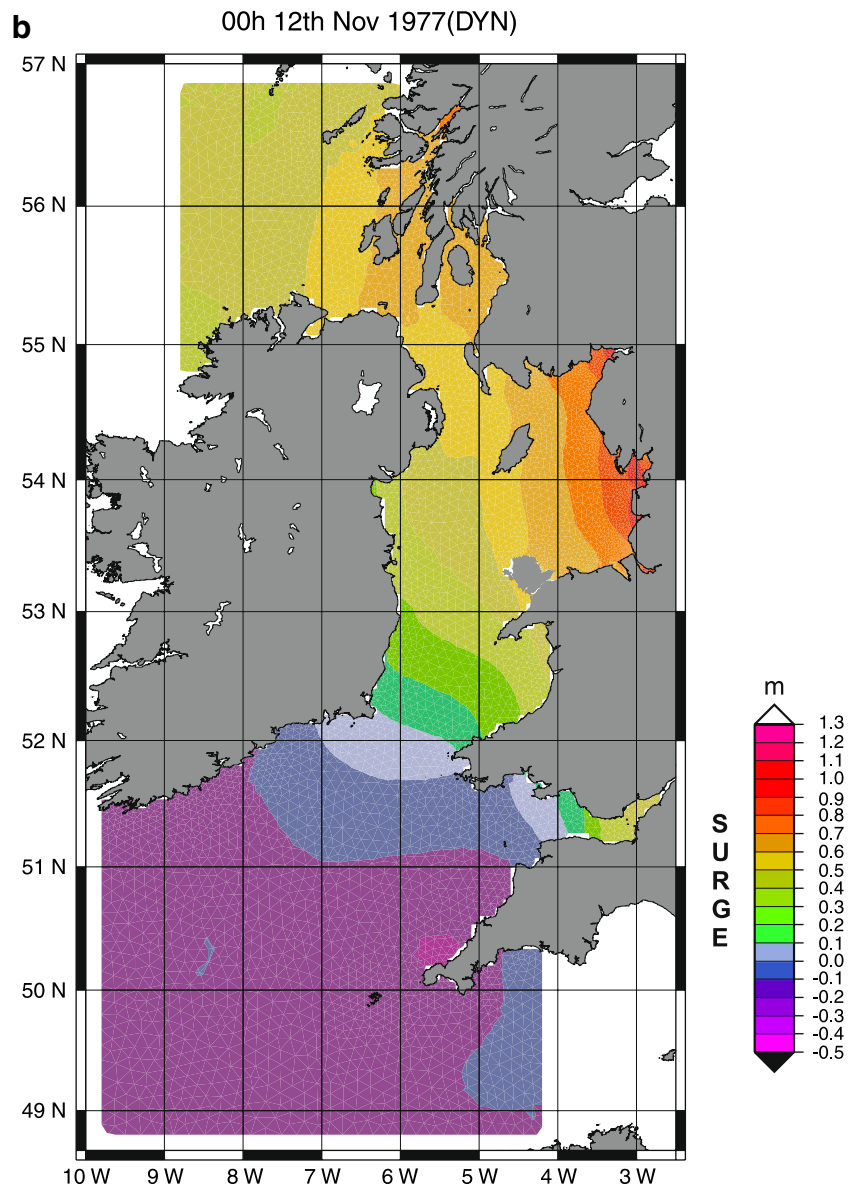
3.1 External surge

In an initial calculation (calc 1), the model was forced using only the external surge taken from observations (see DJ92 for details) applied along the open boundary, although as stated previously, tidal forcing through the open boundary was included. Consequently, no wind or atmospheric pressure gradient forcing was applied over the model domain. In this calculation, to be consistent with the coarse-grid (7-km resolution) finite difference model of the whole domain (Fig. 1), observations from Castletownsend and Newlyn were interpolated along the southern

boundary of the model. Observations from Malin were imposed along the northern boundary. This forcing was identical to that used in the coarse-grid model of DJ92.

Time series (Fig. 4a) of the storm surge elevation at a number of ports in the eastern Irish Sea (for locations, see Fig. 2) showed that although some of the main features of the surge at Douglas could be reproduced through open boundary forcing (the external surge), the model significantly underestimates surge peaks, which occurred at times of maximum wind forcing. However, at other times, surge elevations were reproduced, suggesting that these arose from “far field forcing” that propagated into the region through the open boundaries. In addition, the underestimation of the surge increased rapidly at shallow water

Fig. 5 (continued)



locations such as Liverpool, Workington and Hilbre, suggesting that local wind forcing in the regions was a major contributor to the surge.

3.2 Internal surge

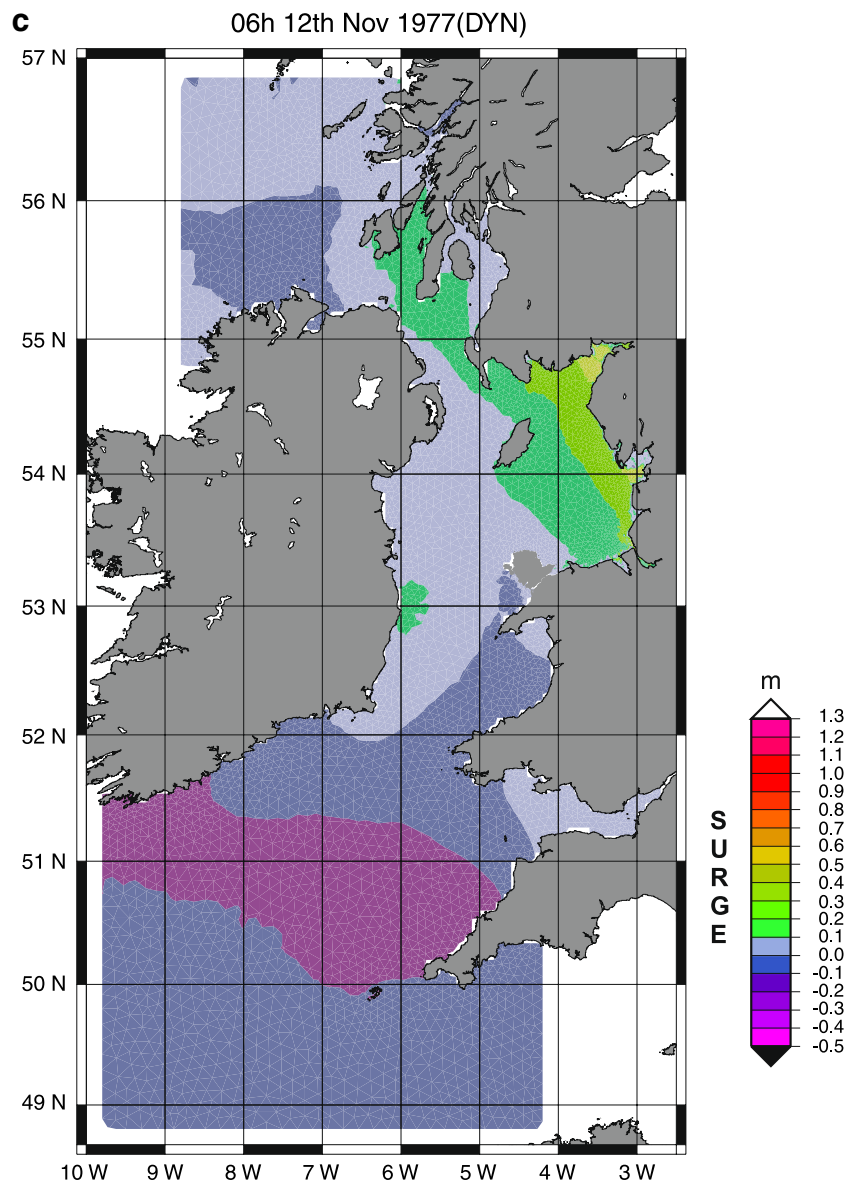
In a subsequent calculation (calc 2), no external surge was applied along the open boundary, although as previously, tidal forcing was included; however, now, meteorological forcing was provided by wind stresses and pressure gradients. Time series at all the eastern Irish Sea ports (Fig. 4b) show that the model could reproduce the major features of the observed surge. In particular, the rapid increase in surge elevations to give peak values at 0000 hours 12 Nov and during 14 Nov at times of strong

wind forcing was reproduced. However, at other times of weak wind forcing, the model failed to reproduce the small negative external surge. As discussed above, this was due to external forcing. Although the observed time variation of the surge is reproduced by the computed internal surge, its magnitude is below the observed. However, the magnitude of the computed surge increases as the water shallows, suggesting that the contribution of the internal surge to the total surge elevation will be more important in shallow than deep water regions.

3.3 Total surge

In a final calculation (calc 3), the model was forced with both the open boundary surge elevation as in calc 1 and

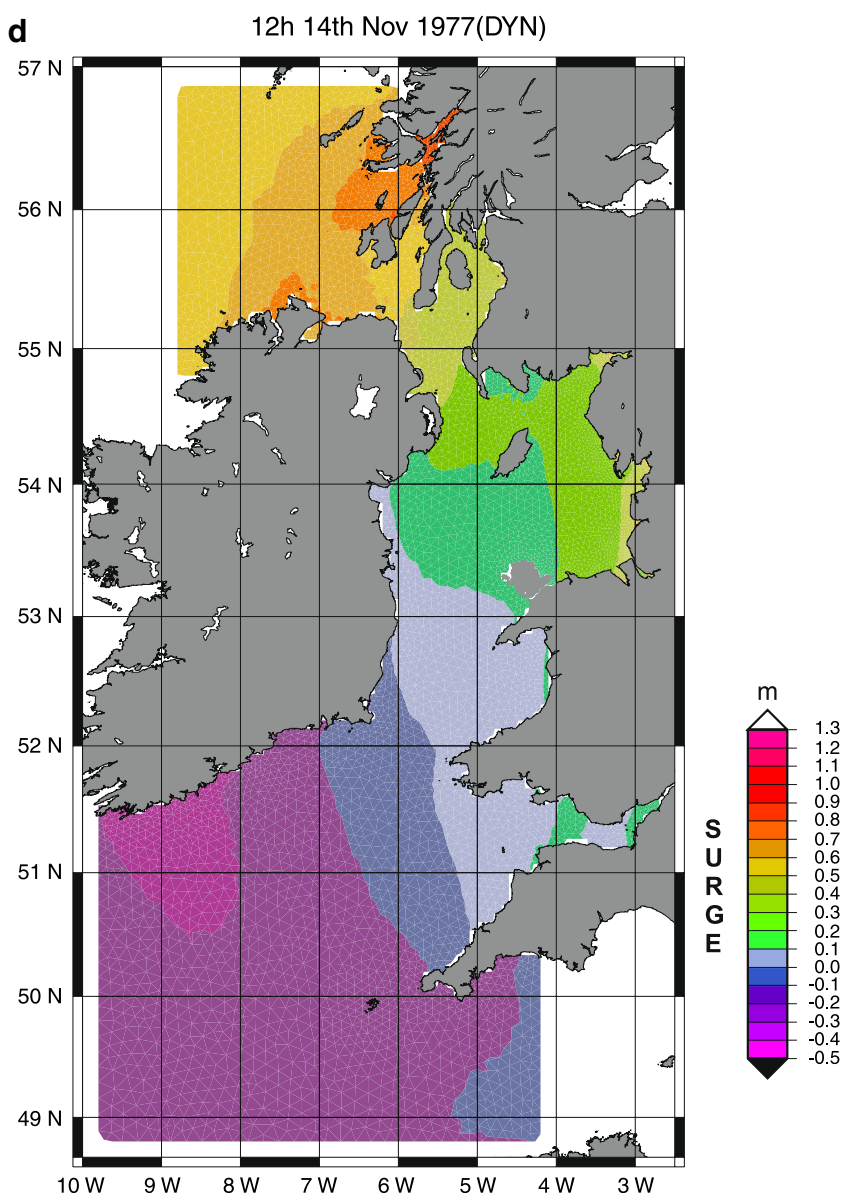
Fig. 5 (continued)



with the meteorological forcing as in calc 2. As previously (calcs 1 and 2), tidal forcing was included in the model. Time series at Douglas (Fig. 4c), located in deeper water on the western side of the eastern Irish Sea, shows that the finite element model can reproduce the observed features of the surge at this location, in particular the surge peaks that occur at 0000 hours 12 Nov and during 14 Nov. In addition, the time series is in close agreement with that computed by JD98. It is interesting that both the time series computed with the present finite element model and the finite difference model (FREISM) of JD98 are in such good agreement, considering that JD98 used a large area coarse-grid model (7 km) of the region shown in Fig. 1 to provide boundary conditions for a 1-km model of the eastern Irish Sea. In this 1-km model, the N_2 and S_2 tides, together with

the higher harmonics of the tide along the open boundary, had been adjusted to give the best possible solution in the interior. In the present finite element model, no such adjustment was made, and tide and surge have been propagated through the whole domain and into the eastern Irish Sea. This suggests that both the tide (a freely propagating wave) and the storm surge (a meteorologically forced event) can accurately propagate through the unstructured finite element mesh (Fig. 3). At shallow water locations, e.g. Liverpool, both the present and previous (JD98) solutions exhibit similar features, with both models failing to reproduce the full magnitude of the surges that occurred at 0000 hours 12 Nov and during 14 Nov. This could be due to a lack of detailed local meteorological forcing in the region, which calc 2 has shown to be

Fig. 5 (continued)

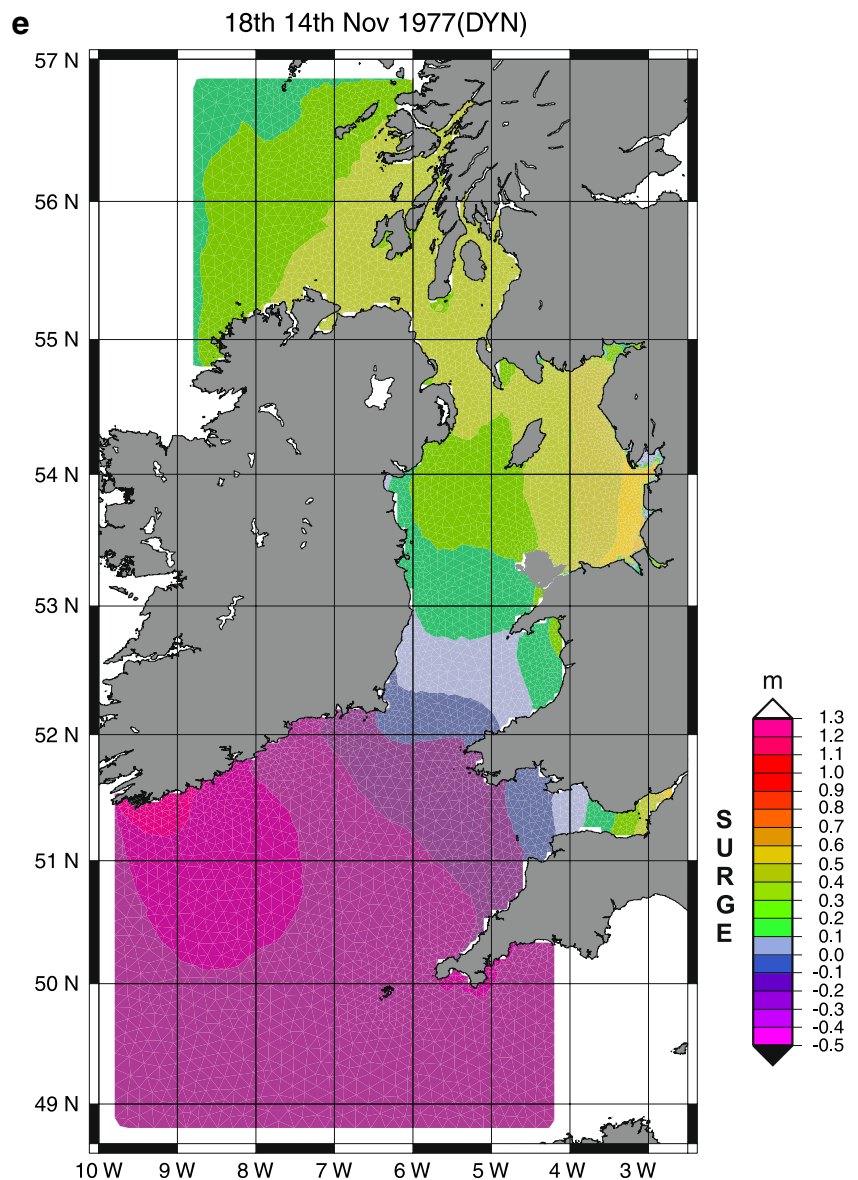


important in shallow water, or a lack of local resolution. In the case of the local eastern Irish Sea finite difference model, this was limited to 1 km. Although the finite element model uses finer elements in this region, in order to perform a “like with like” comparison, identical topography based on a 1-km grid was used. This suggests that in order to take full advantage of the finite element model’s ability to refine the grid in nearshore regions, with a possible improvement in accuracy, higher nearshore bathymetry is required. Although differences in the two solutions are evident at Liverpool, the two solutions are in closer agreement at Heysham and Workington. Interestingly, at Heysham, the finite element model shows some higher frequency “spikes” between 1800 hours 11 Nov and 0000 hours 12 Nov that were not present in the finite difference solution. The reason for this will be discussed later when

the dynamics of the surge in the Morecambe Bay area and its distribution over the whole west coast are examined.

In order to quantify the accuracy of the storm surge computed with the finite element model and compare with earlier finite difference solutions (Jones and Davies 2001), root mean square (rms) errors based on differences between observed and computed surge were determined at Douglas and Liverpool. These ports were chosen to represent deep and shallow locations and because a continuous observed time series was available. Besides computing errors for the whole period (Table 1), sub-periods as in Jones and Davies (2001) were used. The first period is from 0000 hours 9 Nov to 1100 hours 12 Nov covering the first peak. The second is an extended version of the first, namely from 0000 hours 9 Nov until 2300 hours 13 Nov, in essence the period up to the start of the second peak. The third period

Fig. 5 (continued)



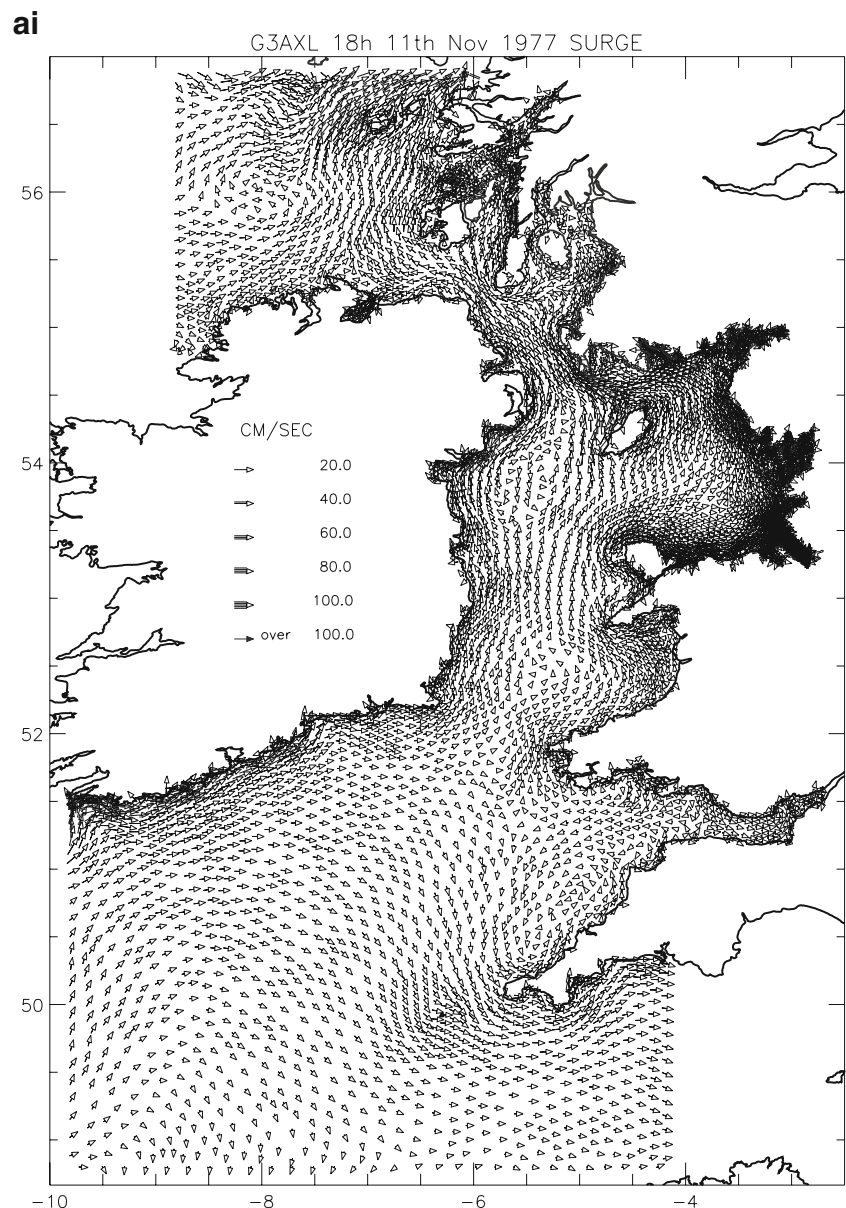
again starts at 0000 hours 9 Nov until 2300 hours 14 Nov and hence covers both surge peaks.

From rms errors in Table 1, it is evident that at Liverpool for all the periods, and the total, the rms error from the finite difference calculation is slightly less than that from the finite element model. There was, however, no significant difference in rms errors at Douglas which is located in deeper water. The fact that both models use the same topography and no attempt was made to improve the accuracy of the finite element model by enhancing local topography suggest that the difference in rms errors at Liverpool is not due to resolution. As discussed previously in the 1-km model of the eastern Irish Sea used by Jones and Davies (2001), the tidal input had been adjusted along the open boundary to give an optimal solution in the

interior. Since in the present finite element model no such adjustment was made and tide–surge interaction in the eastern Irish Sea has an important influence on the surge at Liverpool but not Douglas, it suggests that this may account for the very small difference in rms errors between the models at Liverpool.

Although calculations (Jones and Davies 2008) have shown that tide–surge interaction is important in this region, and hence a realistic storm surge could not be produced without including the tide, calculations showed that when the external and internal storm surges were added together (in essence time series in Fig. 4a,b), then the total surge was not significantly different (differences of less than 5 cm in the time series) from that found in the full surge calculation. This arises mainly from the fact that a

Fig. 6 **a** Current vectors at 1800 hours 11 Nov over (i) whole region, (ii) expanded plot in eastern Irish Sea, (iii) expanded plot in Morecambe Bay, (iv) expanded plot in Liverpool Bay, (v) expanded plot in Mersey estuary. **b** As in **a**, but at 0000 hours 12 Nov. **c** As in **a**, but omitting expanded plots (iii), (iv) and (v) at 0600 hours 12 Nov. **d** As in **a**, but omitting expanded plots (iii), (iv) and (v) at 1200 hours 14 Nov. **e** As in **a**, but omitting expanded plots (iii), (iv) and (v) at 1800 hours 14 Nov



large domain model was used, and hence, the external surge was small and only significant at times when the internal surge was negligible.

4 Storm surge dynamics

In order to understand the time variation of the surge within the eastern Irish Sea and compare its offshore distribution in detail with that from the 7-km west coast model (DJ92) and 1-km eastern Irish Sea model (JD98), it is essential to consider the time varying response of the whole region to storm forcing.

Consider initially the first major wind period (namely 1200 hours 11 Nov to 0600 hours 12 Nov, see DJ92 for

detailed meteorological charts). During this period, there were winds from the south–west over the region that forced water from the Celtic Sea into the Irish Sea, giving rise to an increase in sea level in the eastern Irish Sea particularly in the Solway estuary and Morecambe Bay regions. It is evident from Fig. 5a that although sea level rises at the entrance to the Solway estuary, Morecambe Bay and the Mersey estuary, the storm surge has not yet fully propagated into these shallow water regions. In addition, there is significant spatial variability in the storm surge elevation in these regions, reflecting local changes in bottom topography. In the Heysham region situated at the southern end of Morecambe Bay, there is significant small-scale variability in the surge produced by local topography. As the surge enters the region, “wetting and drying” can occur, giving

Fig. 6 (continued)

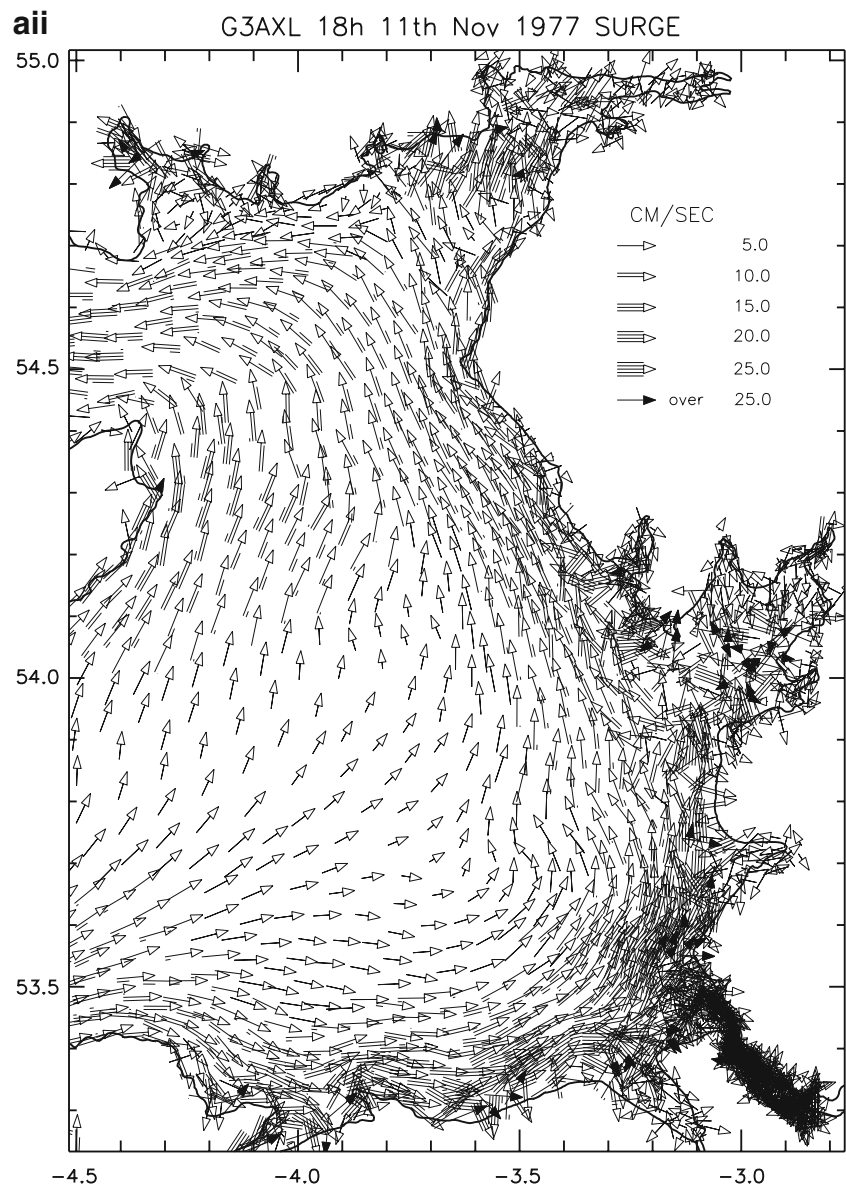
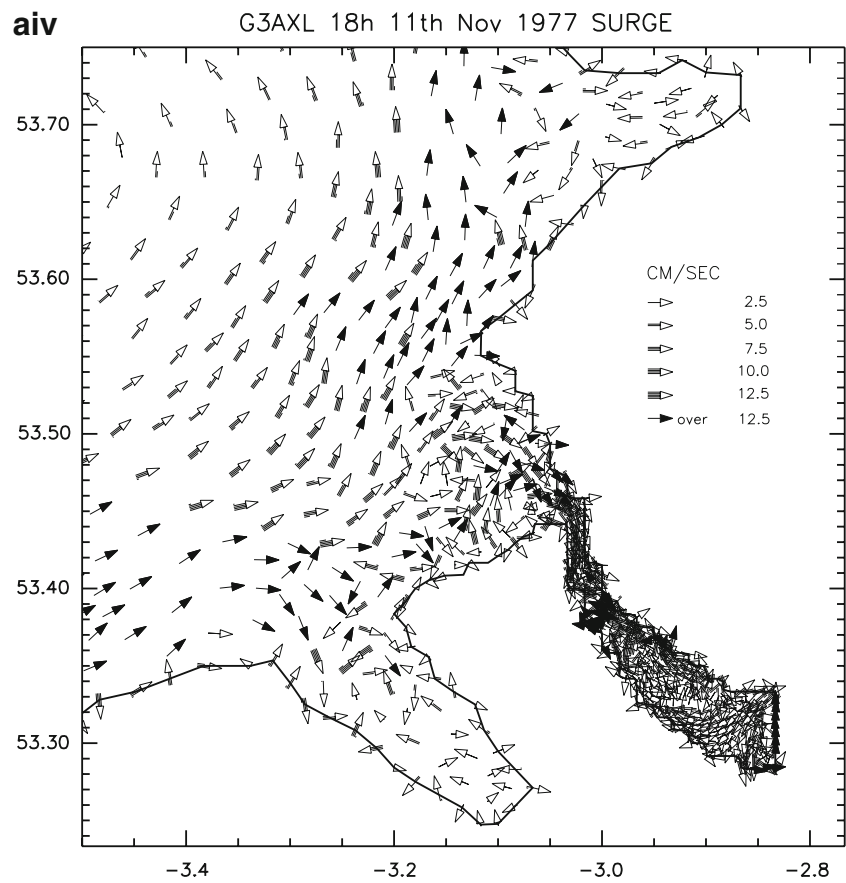
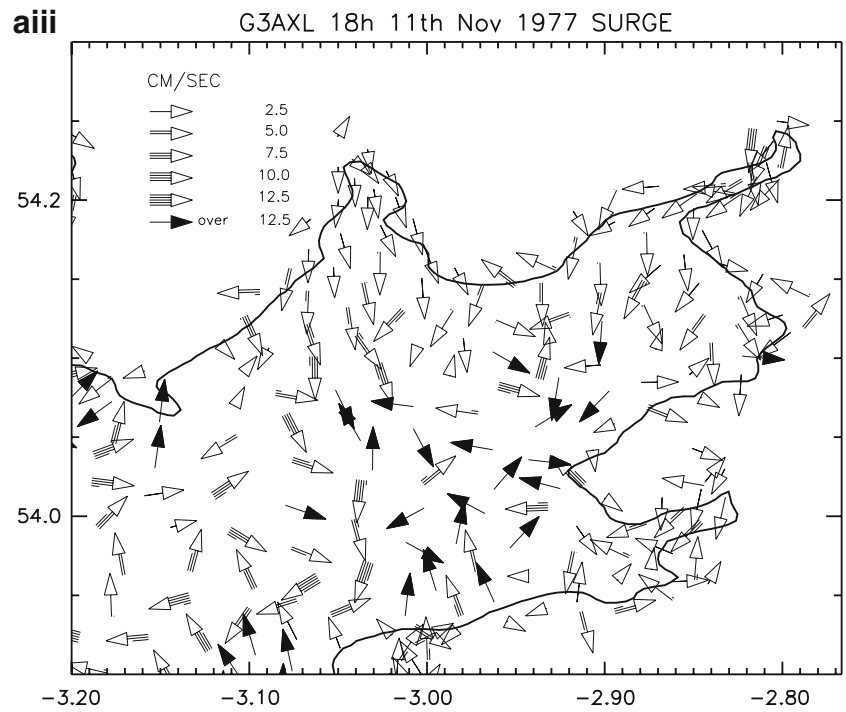


Fig. 6 (continued)



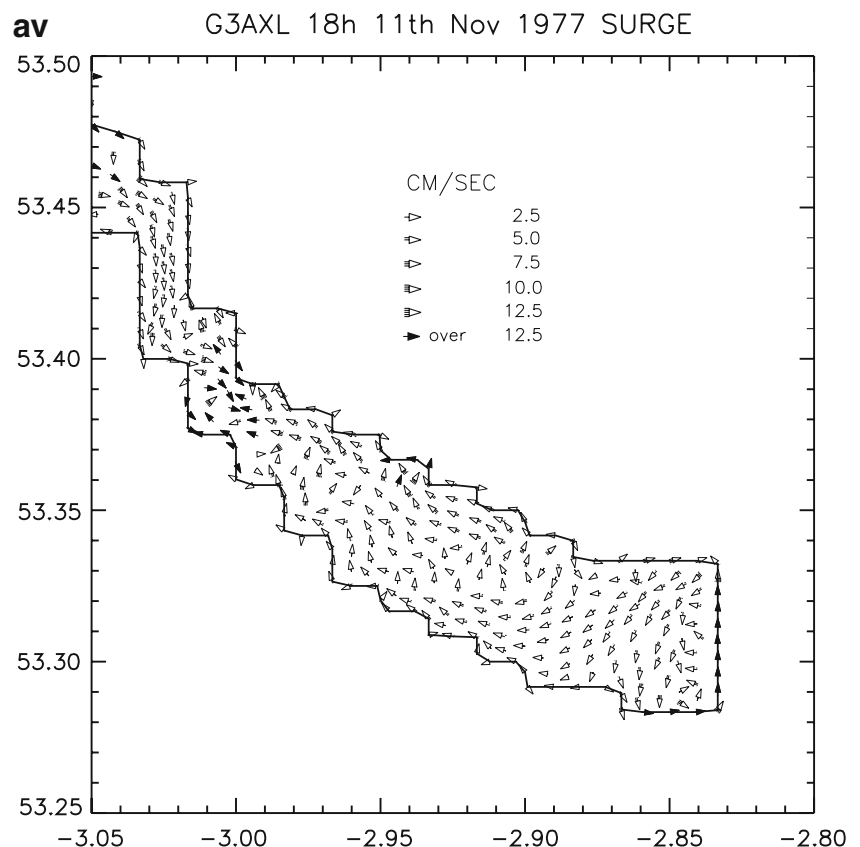
rise to short period oscillations in the surge elevation as shown at this time in Fig. 4c. The magnitude and spatial variability of the storm surge elevation in the eastern Irish Sea is consistent with that found in the high-resolution 1-km local area model of JD98 (compare Fig. 5a with Fig. 14a of JD98).

Current vectors over the whole region (Fig. 6a, *i*) show water flowing from the Celtic Sea into the eastern Irish Sea in the region to the south of the Isle of Man. In addition, some of the water that flows into the Irish Sea continues north along the east coast of Ireland, leaving the region through the North Channel. To the north of the Isle of Man, water leaves the eastern Irish Sea and flows out through the North Channel. The ability of the finite element model to refine the mesh in regions such as the North Channel, where previous work (Davies et al. 2001) required the nesting of a local high-resolution model, is a significant advantage over earlier work using regular finite difference grids. In addition, refining the mesh within the eastern Irish Sea allows more detail of the flow fields to be observed in the nearshore region (Fig. 6a, *ii*). It is evident from Fig. 6a, *ii* that there is a significant increase in current intensity in the nearshore region, with currents flowing parallel to the coastal boundary. As shown previously for the case of tidal residuals (Jones and Davies 2007c), coastal irregularities produced by finite difference “stair case” effects can

introduce spurious eddies in near coastal regions. Complex spatial variations are evident in the currents close to the entrances to shallow water estuaries, as can be seen in expanded plots of the Morecambe Bay (Fig. 6a, *iii*) and Liverpool Bay (Fig. 6a, *iv*) regions. At the entrance to Morecambe Bay (Fig. 6a, *iii*), namely at about 54.05N, $-3.1W$ (close to Heysham), currents change from 2.5 cm s^{-1} to more than 12.5 cm s^{-1} over one element. This suggests that to resolve exchange between these estuaries and the outside region and within the estuaries, fine mesh resolutions of the order of 50 m or less are required. A similar complex distribution of currents is evident in the Liverpool Bay area at the entrance to the Mersey estuary. This clearly shows that although storm surge elevations may vary smoothly in space, the currents exhibit significant small-scale variability that must be taken into account in any measurement programme.

Although the present model cannot take account of detailed topographic variations in the Mersey estuary, since to be consistent with earlier finite difference work (JD98) the topography in this region is taken from the 1-km grid of JD98, it is evident (Fig. 6a, *v*) that there is some spatial variability in the currents within the Mersey. This suggests that to reproduce storm surge events at Liverpool, a more accurate description of the Mersey such as that used by Jones and Davies (2006) is required.

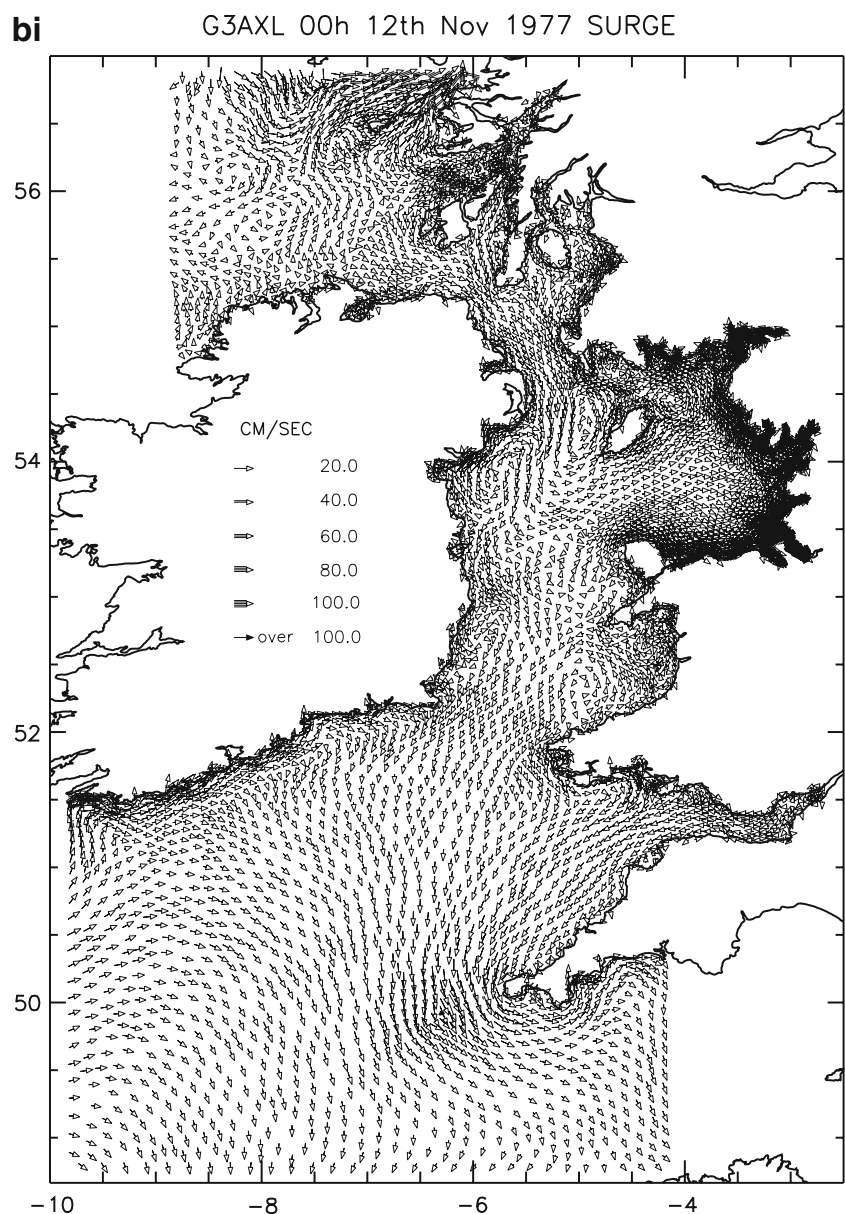
Fig. 6 (continued)



In essence, the surface elevation and current distributions shown in Figs. 5a and 6a were produced by south westerly winds. However, by 0000 hours 12 Nov, the wind direction changed to one from the northwest. This gave rise to a rapid increase in elevations in the eastern Irish Sea (Fig. 5b), although elevations decreased in the Celtic Sea. The alignment and distribution of elevation contours in the deep water regions (Fig. 5b) corresponds very closely with those found in the coarse-grid (7 km) model of DJ92 (compare Fig. 5b with Fig. 6b in DJ92). In the eastern Irish Sea where the mesh is much finer, elevation contours are in good agreement with the high-resolution (1 km) model results of JD98 (compare Fig. 5b with Fig. 14b in JD98).

Current vectors at 0000 hours 12 Nov reveal (Fig. 6b, i) that unlike previously (Fig. 6a, i) at this time, there is an inflow of water into the Irish Sea through the North Channel driven by the winds from the northwest. Some of this water flows due south in the deep region to the west of the Isle of Man, whilst some water flows into the eastern Irish Sea in the region to the south of the Isle of Man. An outflow from the eastern Irish Sea is evident to the north of the Isle of Man. Associated with this inflow and outflow a current gyre develops in the eastern Irish Sea to the northwest of the Isle of Man (Fig. 6b, ii). As previously in coastal regions, the flows are parallel to the coast, with intensity increasing very rapidly as the water shallows.

Fig. 6 (continued)



In the Morecambe Bay region (Fig. 6b, *iii*), there is less spatial variability in the currents than previously (Fig. 6a, *iii*), with a flow to the north in the shallow regions and an elevation gradient forced flow out of the bay in the deeper central channel. Similarly, in Liverpool Bay (Fig. 6b, *iv*), there is a more spatially coherent directly wind forced flow (Fig. 6b, *iv*), although at the entrance to the Mersey and within it, the current (Fig. 6b, *v*) has similar spatial variability to that found previously (Fig. 6a, *v*).

As illustrated at 1800 hours 11 Nov, although surge elevation contours show a uniform distribution, there is significant variability in the currents. This current variability tends to decrease at times of strong wind forcing due to the large-scale wind which is spatially coherent, setting the space scale rather than local elevation gradients in

nearshore regions which have a small space scale due to variation in topography. However, as will be shown later, in terms of flow away from coastal effects, where elevation gradients are more uniform, the large-scale pressure gradient forced flow is often spatially uniform.

By 0600 hours 12 Nov, the wind's magnitude has decreased, although surge elevations in the eastern Irish Sea remain significant (of order 30 cm; Fig. 5c). The presence of an appreciable west–east elevation gradient across the Irish Sea, comparable to that found in JD98 (compare Fig. 5c and Fig. 14c in JD98) and between the Irish Sea and the region beyond (compare Fig. 5c with Fig. 6c in DJ92), drives water out of the eastern Irish Sea both to the north and south of the Isle of Man (Fig. 6c, *i*). In addition, in the north, there is a net outflow from the Irish Sea both through

Fig. 6 (continued)

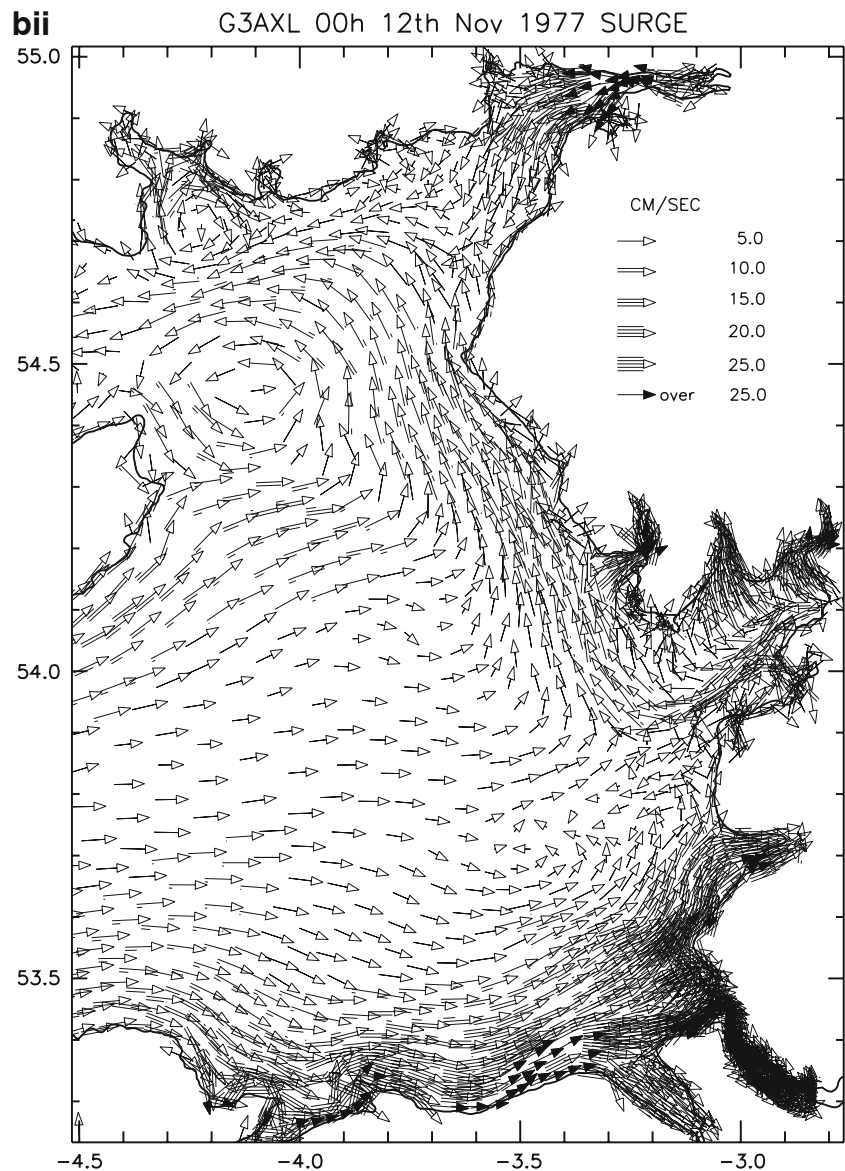
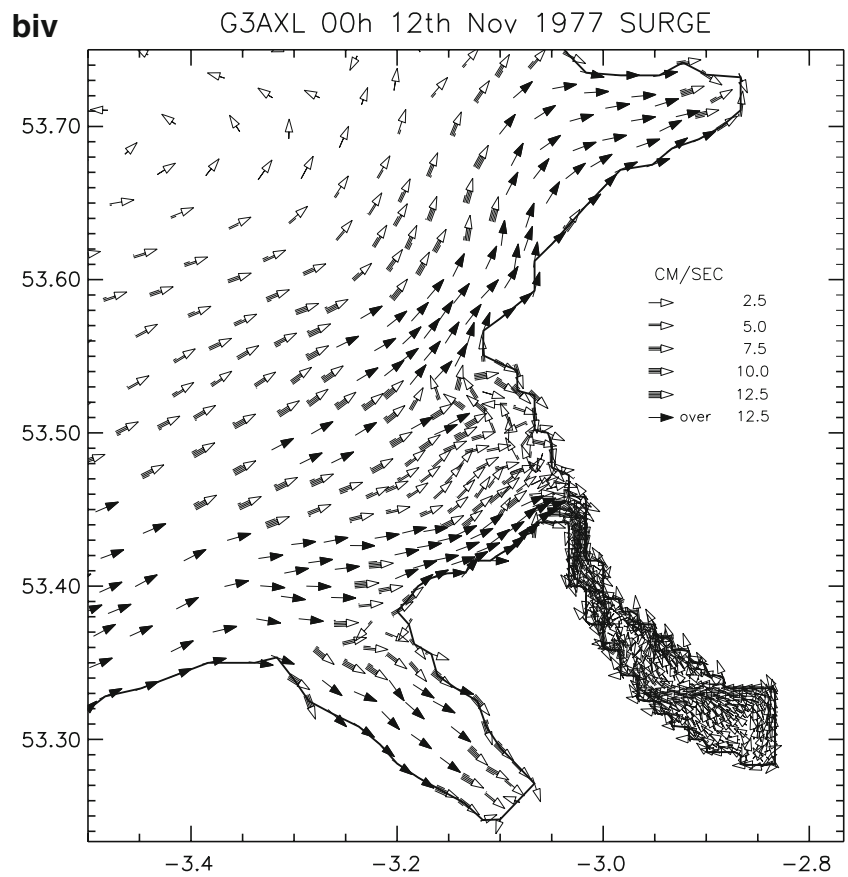
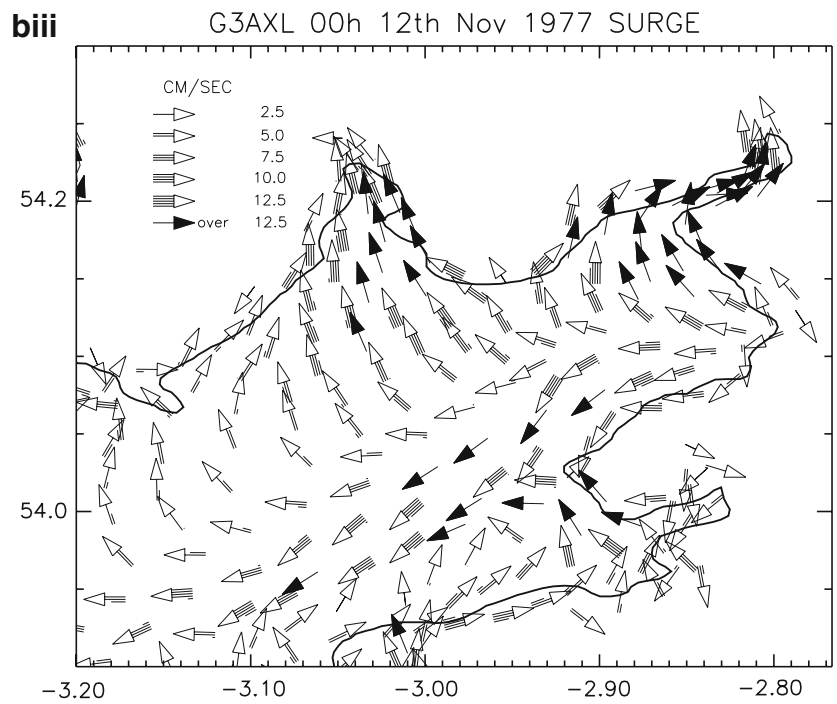


Fig. 6 (continued)



the North Channel and along the west coast of Scotland. A similar net outflow is evident in the south where water flows from the Irish Sea to the Celtic Sea. At this time, the wind field over the region is negligible, and hence, the storm surge elevation gradient that has developed over the region cannot be supported and forces flow out of the eastern Irish Sea. It is evident from Fig. 6c, *i* and *ii* that away from coastal boundary regions, the spatial variability of this elevation-pressure-driven flow is appreciably less than that found at times of wind forcing. However, in some regions, notably off the northeast corner of the Isle of Man, there is a bifurcation in the flow with some flow going to the north and another flow to the south, which leads to local small-scale variability. Similarly, in shallow coastal regions such as Morecambe Bay (expanded plot not shown) and the entrance to the Mersey in Liverpool Bay (expanded plot not shown), there is some local small-scale variability. In particular, in both these regions, current vectors show a strong offshore and out of estuary flow in the deeper water. Within the Mersey (expanded plot not shown), there is a near uniform outflow as sea surface elevations within the estuary decrease.

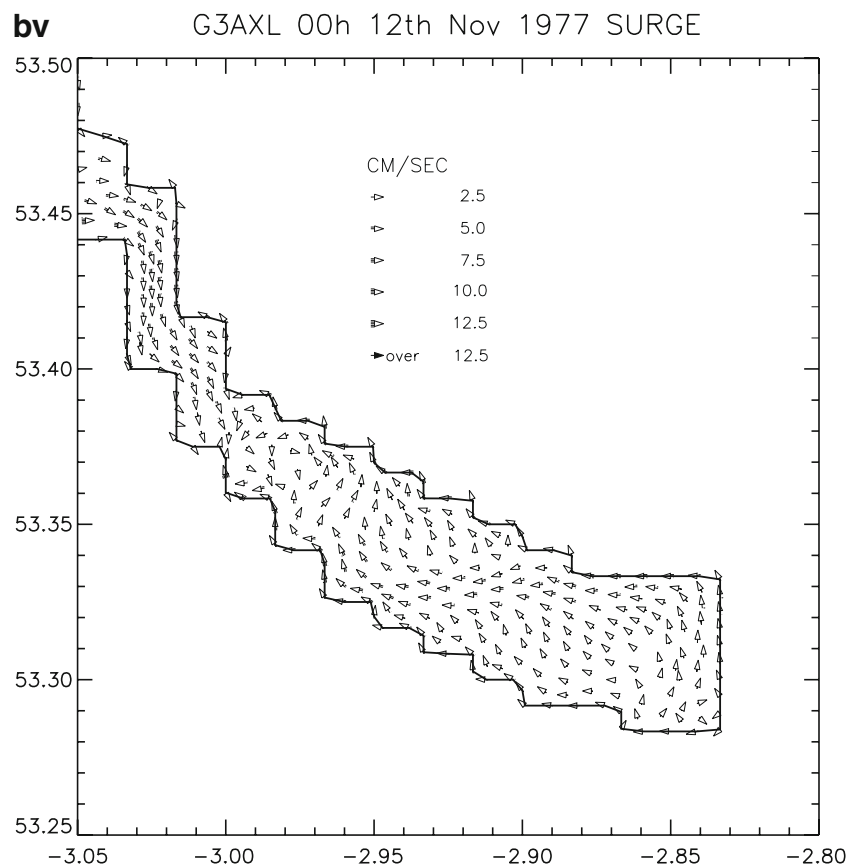
Calculations showed that as the depression that produced the surge of 0000 hours 12 Nov moved out of the region, surge elevations decreased to near zero. However, at

1200 hours 13 Nov, winds from the northwest moved over the region, and their intensity increased and their direction changed to winds from the north over the following 18 hours.

By 1200 hours 14 Nov, these winds had produced a decrease in elevation to the south of Ireland (Fig. 5d), with an increase in elevation to the north of Ireland and along the west coast of Scotland (Fig. 5d). The magnitude and spatial distribution of elevation contours associated with this wind event computed with the finite element model over the west coast region (Fig. 5d) correspond very closely with those computed with the coarser mesh 7-km finite difference grid of DJ92 (compare Fig. 5d with Fig. 6d of DJ92). Similarly, in the near coastal region of the eastern Irish Sea, the computed surge elevation is in good agreement with that computed by JD98 using a limited area high-resolution (1 km) model (compare Fig. 5d with Fig. 14d of JD98).

As surface wind stresses over the region decreased and changed to a wind stress from the north, the elevation gradient to the north of the North Channel could not be maintained and water flowed from the west coast of Scotland region through the North Channel and into the Irish Sea (Fig. 6d, *i*). A significant proportion of this water flows south in the deep channel to the west of the Isle of Man, with some water entering the eastern Irish Sea

Fig. 6 (continued)

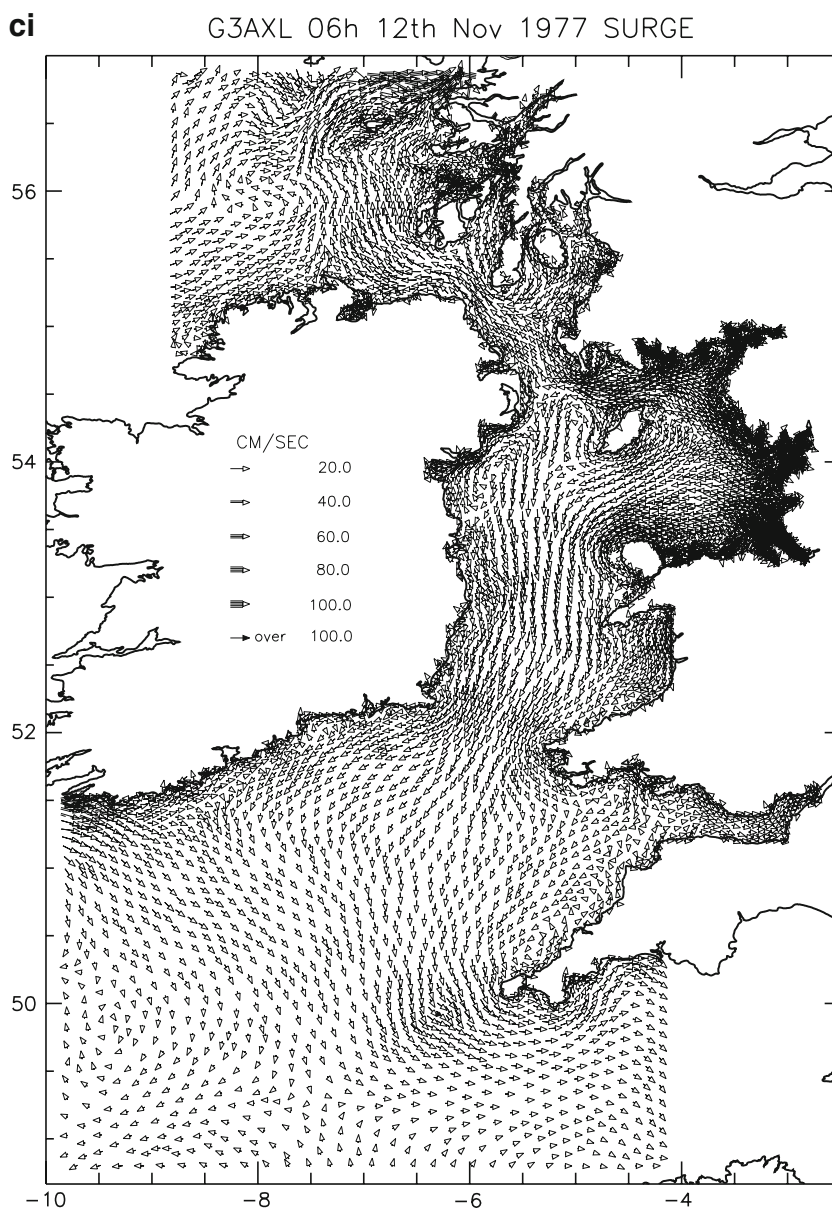


(Fig. 6d, *i*). Within the eastern Irish Sea, as previously, away from the coastal boundary layer, the flow field is fairly spatially uniform, although it changes rapidly in the region of estuaries (Fig. 6d, *ii*). Within estuaries, see for example Morecambe Bay, there is appreciable small-scale variability in the currents (expanded plot not shown) associated with changes in topography. Similarly, at the entrance to the Mersey (expanded plot not shown) and within the Mersey (expanded plot not shown), there is some spatial variability. However, since the wind field at 1200 hours 14 Nov has declined from its previous maximum value of 1 Pa (see JD98 for details of wind stress) and the flow is mainly surface elevation gradient forced into the Mersey from Liverpool Bay, the distribution

of current vectors in the Mersey is substantially smoother than that found previously.

The change in wind direction and increase in magnitude between 1200 hours and 1800 hours 14 Nov gives rise to a positive surge in the eastern Irish Sea at 1800 hours 14 Nov (Fig. 5e). However, the winds from the north, namely offshore winds in the region to the south of Ireland, produce a negative surge of about 50 cm to the south of Ireland (Fig. 5e). The location and magnitude of this negative surge and the distribution of elevation contours in the Celtic and Irish Sea (Fig. 5e) is in close agreement with that computed in DJ92 with the 7-km finite difference grid (compare Fig. 5e with Fig. 6e in DJ92). Similarly, in the eastern Irish Sea, the rapid increase in surge elevation

Fig. 6 (continued)



as the coast is approached and the subsequent decrease within the estuaries corresponds to that computed by JD98 using the fine grid model of the region (compare Fig. 5e with Fig. 14e in JD98).

The influence of the strong wind from the north at this time (1800 hours 14 Nov) is to force water from the west of Scotland through the North Channel and into the Irish Sea (Fig. 6e, *i*). Within the Irish Sea, there is a flow to the south in the deep channel (Fig. 1) to the west of the Isle of Man that enters the Celtic Sea (Fig. 6e, *i*). Some of the water flowing through the Irish Sea enters the eastern Irish Sea along the northern and southern coastal regions of the Isle of Man (Fig. 6e, *i*). Within the eastern Irish Sea, there is significant spatial variability in the currents (Fig. 6e, *ii*)

produced by a combination of local sea level rise along the coast, which, in Liverpool Bay, produces an offshore flow and wind forced flow from the Irish Sea entering the region. In addition, directly wind forced currents within the eastern Irish Sea contributed to this spatial variability. The net effect of this combined forcing is to produce a large current gyre to the east of the Isle of Man and a number of near shore gyres (Fig. 6e, *ii*). Within Morecambe Bay (expanded plot not shown), there is significant spatial variability in both current magnitude and direction due to variations in topography and the interplay between local elevation gradient forced flow and that due to direct wind forcing.

Similarly, in Liverpool Bay (expanded plot not shown) away from the nearshore region, there is a uniform offshore

Fig. 6 (continued)

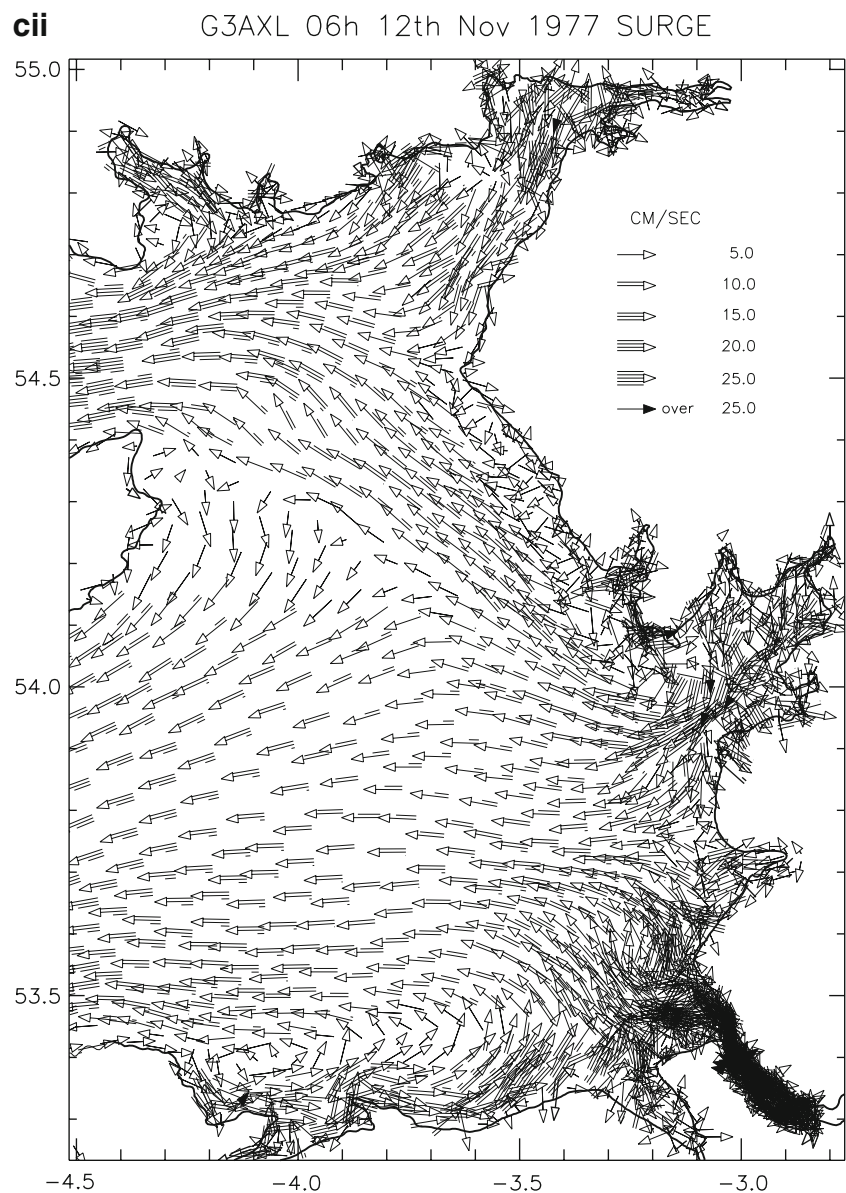
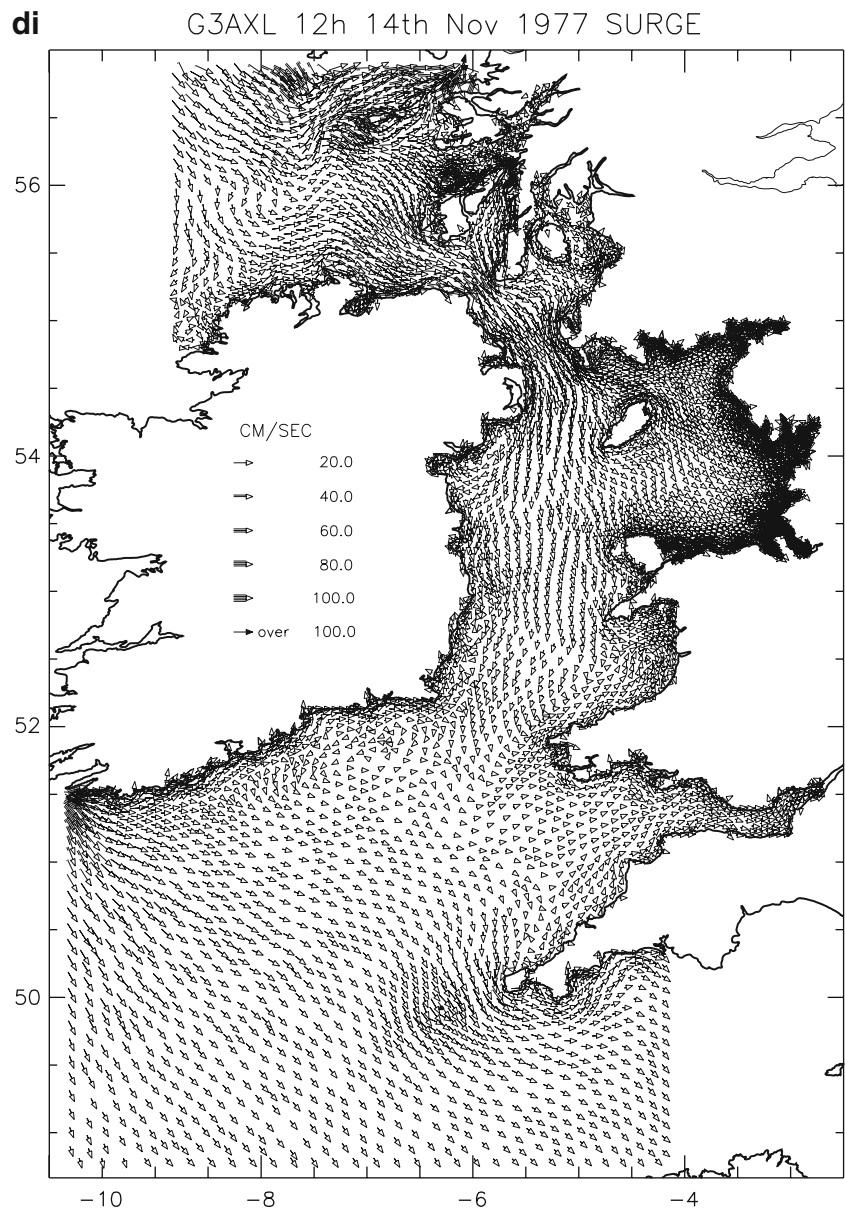


Fig. 6 (continued)

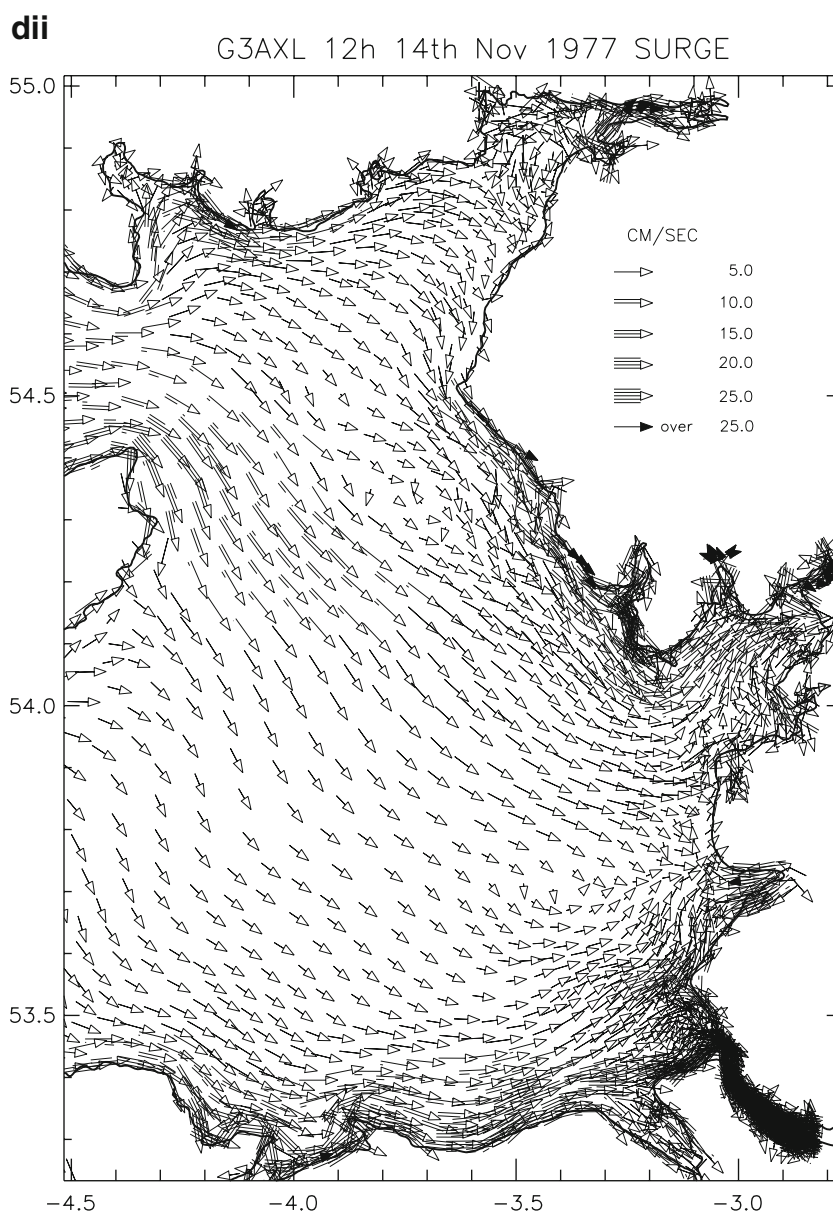


flow; however, in the nearshore region, a detailed examination shows that the currents exhibit significant spatial variability. Nevertheless, within the Mersey estuary, a more uniform distribution of currents is evident. The uniform flow in the Mersey found in the present calculations primarily arises from a lack of resolution in this region. This is due to the fact that in order to compare surge solutions derived with the finite element model with those computed by JD98, identical topography was used within the Mersey. As shown by Jones and Davies (2006) for the case of the response of the Mersey to uniform wind forcing, when more detailed and accurate topography together with element sizes of order 50 m are used in the Mersey, then a more complex spatial pattern arises. However, the whole

region surge plots clearly show that there was significant spatial and temporal variability in both the surge elevations and currents during the November 1977 storm surge event. This variability arises from time and space variations in the meteorology and the significant changes in water depth over the region. The existence of an accurate meteorological data set for model forcing and coastal gauges for comparison makes it an ideal period for testing models and inter-comparing their performance.

Comparison with a previous coarse-grid large area model (DJ92) and a high-resolution limited area Eastern Irish Sea model (JD98) showed that the finite element model with its graded mesh could reproduce the large-scale variability of the surge over the region. In addition, its finer

Fig. 6 (continued)



mesh in the eastern Irish Sea could resolve the small-scale variability of the surge in this region.

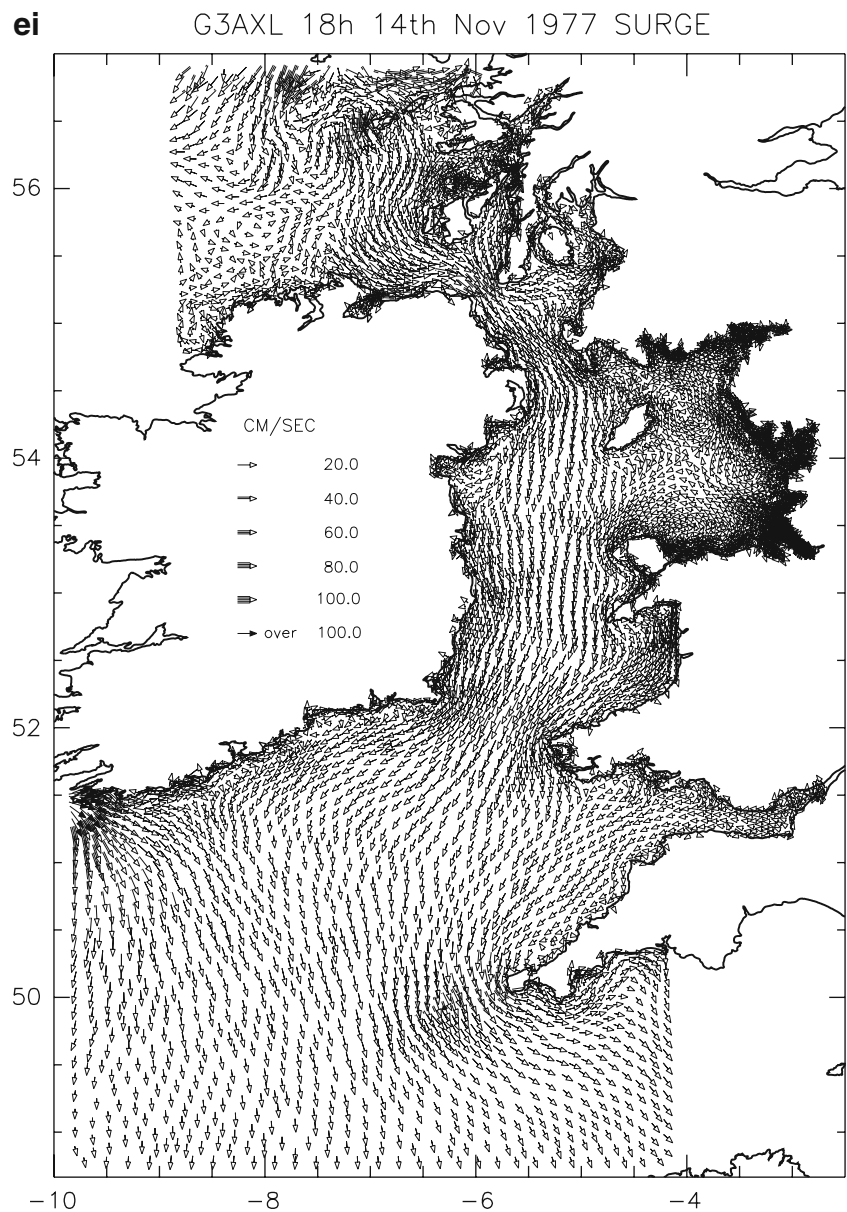
5 Concluding remarks

An unstructured mesh finite element model of the sea region off the west coast of Britain has been used to model the storm surge event of November 1977. This period was chosen because an accurate meteorological data set was available to drive the model and coastal gauge data could be used to validate the model. In addition, the solution from a large area coarse-grid (7 km) finite difference model covering an identical region and a limited area (eastern Irish

Sea) high-resolution (1 km) model was available for comparison. By covering the same area as the coarse-grid model and using identical water depth distributions to those used in the coarse- and high-resolution limited area model, a valid comparison with earlier finite difference solutions could be made, since all models used the same forcing.

Calculations showed that the external component of the surge was significant at all locations, although in the eastern Irish Sea, wind forcing, namely the internal surge, was a major contributor. Consequently, in any storm surge simulation, the model had to accurately propagate the surge into the eastern Irish Sea and account for wind forcing in the region and local changes in topography. The finite element model with its ability to vary the mesh could

Fig. 6 (continued)



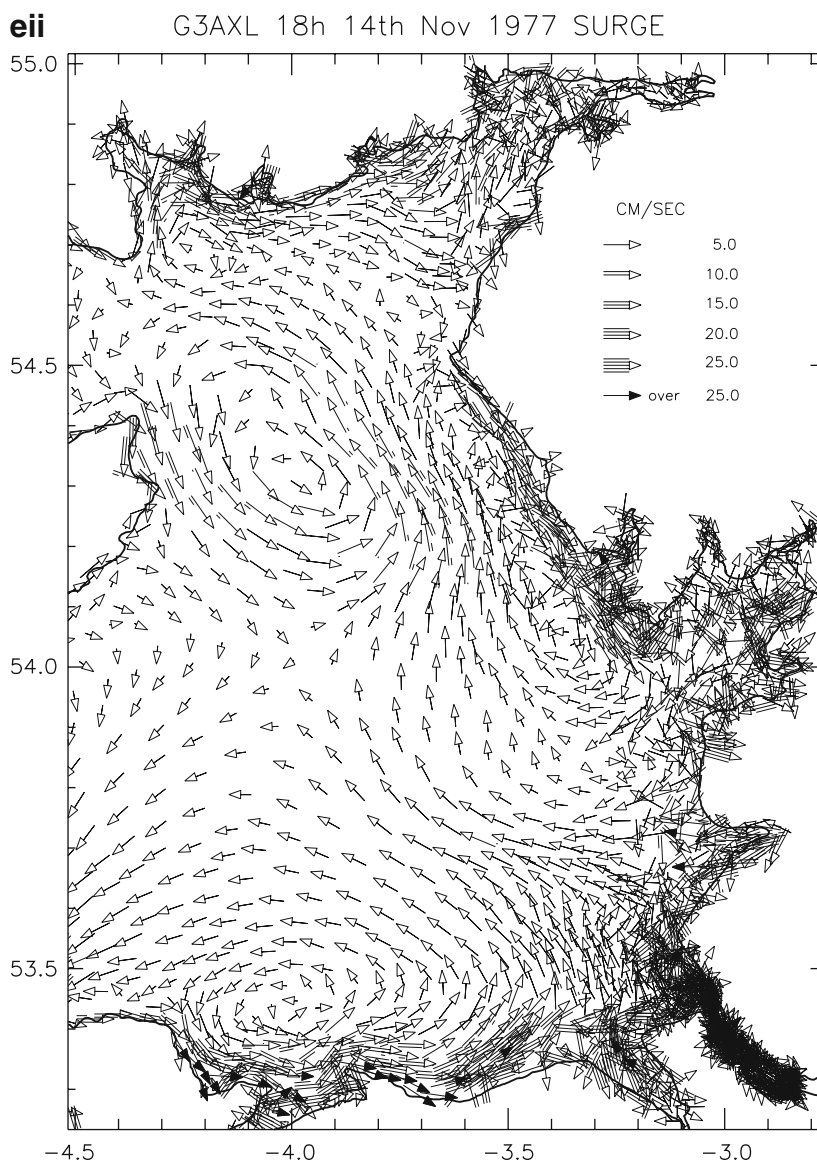
accurately resolve narrow channels such as the North Channel (Figs. 1 and 3) where previously nested high-resolution models had been used (Davies et al. 2001) and the nearshore region of the eastern Irish Sea. This area had previously been modelled with a limited area high-resolution (1 km) model (JD98).

Time series of the storm surge elevations at ports in the eastern Irish Sea showed that the finite element model with its fine mesh in this region could reproduce surge elevations with comparable accuracy to the 1-km model of JD98. Similarly, surface elevation contours both within the eastern Irish Sea and the region beyond where the mesh was much coarser could be reproduced by the model and were in close agreement with those computed in DJ92 and JD98. By

using the same region, open boundary and meteorological forcing together with identical water depths showed that the solution was independent of the numerical method used to derive it (namely finite difference or finite element). In addition, the close agreement in the solutions showed that the finite element model with its irregular mesh could accurately reproduce the storm surge propagation through the region and account for local changes due to wind forcing and nearshore fine-scale topography.

Although the use of identical topography enabled a valid model inter-comparison to be made, the use of a coarse representation of the nearshore region of Liverpool Bay and the Mersey as in JD98 meant that the storm surge was not accurately resolved in this region. The importance of high

Fig. 6 (continued)



resolution in the Mersey in resolving details of the wind-induced flow in this area has recently been reported by Jones and Davies (2006) for the case of uniform steady wind forcing. This suggests that in any future calculations of storm forced flow in the Liverpool Bay and Mersey estuary, an enhanced mesh refinement in this region as in Jones and Davies (2006) is required. This effect of such a mesh enhancement on storm surges is currently being investigated.

Certainly, the ability of the finite element model to refine the mesh in regions of rapidly changing topography where high resolution and accuracy is required (e.g. the nearshore region or within an estuary) is a major advantage over the nested finite difference approach used previously (JD98, Davies et al. 2001). As shown here, at the entrance to many estuaries, there is significant spatial variability in the flow,

suggesting that nesting a fine mesh finite difference model of an estuary within a larger model would be very difficult. As shown by Davies and Hall (2002), nesting a local area model within a coarser model in regions of rapidly varying flow can significantly influence and sometimes reverse the flow in the limited area model.

These calculations suggest that the use of an unstructured mesh model with a progressively finer grid in the nearshore region should improve storm surge prediction, provided detailed and accurate bottom topography is available in these areas. However, as the mesh is refined, other processes such as wave–current interaction (Davies and Lawrence 1995; Jones and Davies 2001) and the three-dimensional nature of the flow become important (JD98) and may need to be considered in order to significantly improve storm surge prediction in an irregular mesh model.

Acknowledgements The origin of the TELEMAC system is EDF-LNHE and is therefore ©EDF-LNHE. The authors are indebted to R.A. Smith for help in preparing diagrams and L. Parry for typing the paper.

References

- Davies AM, Flather RA (1977) Computation of the storm surge of 1–6 April 1973 using numerical models of the northwest European continental shelf and North Sea. *Dtsch Hydrogr Z* 30:139–162
- Davies AM, Jones JE (1992a) A three-dimensional wind driven circulation model of the Celtic and Irish Seas. *Cont Shelf Res* 12:159–188
- Davies AM, Jones JE (1992b) A three dimensional model of the M_2 , S_2 , N_2 , K_1 and O_1 tides in the Celtic and Irish Sea. *Prog Oceanogr* 29:197–234
- Davies AM, Lawrence J (1995) Modelling the effect of wave–current interaction on the three-dimensional wind driven circulation of the eastern Irish Sea. *J Phys Oceanogr* 25:29–45
- Davies AM, Jones JE (1996) Sensitivity of tidal bed stress distributions, near bed currents, overtides and tidal residuals to frictional effects in the eastern Irish Sea. *J Phys Oceanogr* 26:2553–2575
- Davies AM, Hall P (2002) Numerical problems associated with coupling models in shelf edge regions. *Appl Math Model* 26:807–831
- Davies AM, Kwong SCM, Flather RA (1998) A three-dimensional model of wind-driven circulation on the shelf: application to the storm of January 1993. *Cont Shelf Res* 18:289–340
- Davies AM, Kwong SCM, Flather RA (2000) On determining the role of wind wave turbulence and grid resolution upon computed storm driven currents. *Cont Shelf Res* 20:1825–1888
- Davies AM, Hall P, Howarth MJ, Knight P, Player R (2001) A detailed comparison of measured and modeled wind driven currents in the North Channel of the Irish Sea. *J Geophys Res* 106:19683–19713
- Fernandes EHL, Dyer KR, Moller OO, Niencheski LFH (2002) The Patos lagoon hydrodynamics during an El Nino event (1998). *Cont Shelf Res* 22:1699–1713
- Fernandes EHL, Marino-Tapia I, Dyer KR, Moller OO (2004) The attenuation of tidal and subtidal oscillations in the Patos Lagoon estuary. *Ocean Dynamics* 54:348–359
- Greenberg DA, Dupont F, Lyard FH, Lynch DR, Werner FE (2007) Resolution issues in numerical models of oceanic and coastal circulation. *Cont Shelf Res* 27:1317–1343
- Hagen S, Westerink J, Kolar R, Horstmann O (2001) Two dimensional unstructured mesh generation for tidal models. *Int J Numer Methods Fluids* 35:669–686
- Hagen SC, Horstmann O, Bennett RJ (2002) An unstructured mesh generation algorithm for shallow water modeling. *Int J Comput Fluid Dyn* 16(2):83–91
- Ip JTC, Lynch DR, Friedrichs CT (1998) Simulation of estuarine flooding and dewatering with application to Great Bay, New Hampshire. *Estuar Coast Shelf Sci* 47:119–141
- Jones JE (2002) Coastal and shelf-sea modelling in the European context. *Oceanogr Mar Biol: an Annual Review* 40:37–141
- 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 eastern Irish Sea. *Estuar Coast Shelf Sci* 42:311–346
- Jones JE, Davies AM (1998) Storm surge computations for the Irish Sea using a three-dimensional numerical model including wave–current interaction. *Cont Shelf Res* 18:201–251
- Jones JE, Davies AM (2001) Influence of wave–current interaction and high frequency forcing upon storm induced currents and elevations. *Estuar Coast Shelf Sci* 53:397–414
- Jones JE, Davies AM (2003a) Processes influencing storm-induced currents in the Irish Sea. *J Phys Oceanogr* 33:88–104
- Jones JE, Davies AM (2003b) On combining current observations and models to investigate the wind induced circulation of the eastern Irish Sea. *Cont Shelf Res* 23:415–434
- Jones JE, Davies AM (2005) An intercomparison between finite difference and finite element (TELEMAC) approaches to modelling west coast of Britain tides. *Ocean Dynamics* 55:178–198
- Jones JE, Davies AM (2006) Application of a finite element model (TELEMAC) to computing the wind induced response of the Irish Sea. *Cont Shelf Res* 26:1519–1541
- Jones JE, Davies AM (2007a) A high resolution finite element model of the M_2 , M_4 , M_6 , S_2 , N_2 , K_1 and O_1 tides off the west coast of Britain. *Ocean Model* 19:70–100
- Jones JE, Davies AM (2007b) On the sensitivity of computed higher tidal harmonics to mesh size in a finite element model. *Cont Shelf Res* 27:1908–1927
- Jones JE, Davies AM (2007c) On the sensitivity of tidal residuals off the west coast of Britain to mesh resolution. *Cont Shelf Res* 27:64–81
- Jones JE, Davies AM (2008) On the modification of tides in shallow water regions by wind effects. *J Geophys Res* 113:C05014
- Legrand S, Deleersnijder E, Hanert E, Legat V, Wolanski E (2006) High-resolution, unstructured meshes for hydrodynamic models of the Great Barrier Reef, Australia. *Estuar Coast Shelf Sci* 68:36–46
- Legrand S, Deleersnijder E, Delhez E, Legat V (2007) Unstructured, anisotropic mesh generation for the Northwestern European continental shelf, the continental slope and the neighbouring ocean. *Cont Shelf Res* 27:1344–1356
- Levasseur A, Shi L, Wells NC, Purdie DA, Kelly-Gerreyn BA (2007) A three-dimensional hydrodynamic model of estuarine circulation with an application to Southampton Water, U.K. *Estuar Coast Shelf Sci* 73:753–767
- Nicolle A, Karputchev M (2007) Evidence for spatially variable friction from tidal amplification and asymmetry in a shallow semi-diurnal embayment: the Pertuis Breton Bay of Biscay, France. *Cont Shelf Res* 27:2346–2356
- Walters RA (2005) Coastal ocean models: two useful finite element methods. *Cont Shelf Res* 25:775–793
- Werner FE (1995) A field test case for tidally forced flows: a review of the tidal flow forum. In: Lynch DR, Davies AM (eds) Quantitative skill assessment for coastal ocean models. American Geophysical Union, USA, pp 269–284

Modeling ocean circulation on unstructured meshes: comparison of two horizontal discretizations

Sergey Danilov · Qiang Wang · Martin Losch ·
Dmitry Sidorenko · Jens Schröter

Received: 19 February 2008 / Accepted: 18 June 2008 / Published online: 5 August 2008
© Springer-Verlag 2008

Abstract Finite-element models on unstructured meshes are frequently formulated in terms of continuous linear elements, which suffer from pressure modes and require stabilization. Alternatively, horizontal velocities may be represented with linear nonconforming elements. While the latter formulation uses three times more degrees of freedom for the velocity, it does not support pressure modes. The effects of stabilization are estimated by comparing the performance of continuous linear and nonconforming versions of the finite-element ocean circulation model (FEOM) in two simple configurations: a Munk gyre and baroclinic turbulence in a zonally reentrant channel. It is shown that, outside the free slip boundary layers, the presence of stabilization does not lead to noticeable effects if its strength is kept within certain limits. In order to evaluate the performance of FEOM, the baroclinic turbulence test is repeated with the MIT general circulation model (MITgcm), which serves as a benchmark, and reasonable agreement between different model codes is found. The two versions of FEOM have a similar computational cost, but both are significantly slower (per node) than the regular-mesh MITgcm. The paper also provides a brief description of the implementation of the nonconforming version of FEOM.

Keywords Unstructured mesh ocean modeling · Finite-element method

Responsible Editor: Eric Deleersnijder

S. Danilov (✉) · Q. Wang · M. Losch ·
D. Sidorenko · J. Schröter
Alfred Wegener Institute for Polar and Marine Research,
Bremerhaven, Germany
e-mail: Sergey.Danilov@awi.de

1 Introduction

Unstructured meshes suggest a number of conceptual advantages for ocean modeling such as variable resolution and continuous representation of coastlines (or a recent review, see Pain et al. 2005). The former makes nesting unnecessary and is of potential interest for designing configurations with regional focus imbedded into a coarse resolution global ocean. However, current practical usage of unstructured meshes in oceanography is mostly limited to tidal and coastal applications where the propagation speed dependence of surface gravity and barotropic Kelvin waves on fluid depth is effectively taken into account by using variable spatial resolution. Applications of unstructured meshes to simulating large-scale ocean circulation are not common, which is partly associated with their relatively low computational efficiency as compared to their structured grid counterparts. The continuing search for computationally efficient and robust algorithms explains the recent emergence of several models working on unstructured meshes (see, e.g., Casulli and Walters 2000; Chen et al. 2003; Danilov et al. 2004; Ford et al. 2004; Zhang et al. 2004; Walters 2005; Stuhne and Peltier 2006; Fringer et al. 2006; Wang et al. 2008a; Zhang and Baptista 2008; White et al. 2008a). These models differ in the discretization method (finite elements or finite volumes), discretization type (representation of variables on meshes), solution algorithm, and area of application (coastal or large-scale). Because of strong vertical stratification of the real ocean, most models are formulated for vertically aligned meshes, implying that only the surface mesh is unstructured.

The numerical efficiency and accuracy of the models based on the finite-element (FE) method are defined

by functional spaces (polynomial order of functions) selected to represent variables. Although high-order elements allow achieving high spatial accuracy, most practical approaches use low-order polynomials. In this way, the degrees of freedom are used to represent the complex geometry of the computational domain. QUODDY (Lynch et al. 1996), ADCIRC (Westerink et al. 1992), and MOG2D (Carrère and Lyard 2003), as well as models formulated in Danilov et al. (2004) and Wang et al. (2008a) [finite-element ocean circulation model (FEOM)], use a linear continuous representation for velocity and elevation, while the nonhydrostatic ICOM (Ford et al. 2004) works on quadrilateral surface meshes and uses continuous linear (CL) velocities and elementwise constant pressure. The approach by Walters (2005) is formulated with the so-called RT0 element and is close to the finite-volume approach of Casulli and Walters (2000). It uses elementwise constant pressure (elevation) and associates the normal components of velocity with the edges of elements.

The CL representation for velocity and elevation (P_1 - P_1 discretization) is an obvious choice because it requires minimum memory storage and results in a reasonable number of operations for assembling the right-hand sides (RHS) of model equations. Since this representation uses the full vector of horizontal velocities handling the Coriolis operator, which requires special care with finite-volume, C-grid-type discretization, poses no particular problem here. Its notorious difficulty is, however, the spurious pressure (elevation) modes. These are eliminated by different stabilization techniques, such as the generalized wave continuity equation method (used by QUODDY, ADCIRC, MOG2D), Galerkin least squares method (Codina and Soto 1997; Danilov et al. 2004), or pressure (elevation) split method (Zienkiewicz and Taylor 2000; Codina and Zienkiewicz 2002; Wang et al. 2008a). All methods share the drawback that the horizontal velocity field satisfies a modified, as opposed to exact, vertically integrated continuity equation. The bias of numerical solutions caused by stabilization is difficult to assess, except for simple cases.

Recently, it was shown that using the so-called linear nonconforming (NC) representation of velocity (to be explained in Section 2.1) suggests a number of advantages both in numerical accuracy and computational efficiency (see Hanert et al. 2005; Le Roux et al. 2005). In particular, such discretization is free of pressure modes and does not require stabilization. It serves as a basis to a model by White et al. (2008a) that further demonstrates the usability of this approach.

The aim of this paper is to assess the effects of stabilization by comparing the performance of two dis-

cretizations of horizontal velocity, the CL and NC, as implemented in FEOM. We show that the influence of stabilization on flow dynamics is felt as numerical viscosity, leading to a noticeable difference in dynamics of boundary flows subject to free slip boundary conditions. However, the effects of stabilization can be made reasonably small in other situations by controlling the strength of stabilization. We also show that, despite the fact that NC elements suggest a number of simplifications to the numerical algorithm, they do not make this method cheaper in terms of CPU time. Thus, both methods can be recommended, yet one is advised to check that the CL discretization is not overstabilized. Our implementation of the NC code is different from that suggested by White et al. (2008a) in that we apply a pressure correction algorithm instead of introducing a barotropic mode.

An additional task is to carry out an elementary comparison of the performance of FEOM with the structured-grid finite-volume MIT general circulation model (MITgcm) (Marshall et al. 1997) in terms of both CPU time and characteristics of the simulated fields. The mesh used for this comparison is regular. It is triangulated for FEOM and used as it is by the MITgcm. We show that, in the current implementation, FEOM is approximately a factor of 10 slower than the MITgcm in the particular configuration we have used. This comparison is computer-architecture dependent, but the slowness factor already appears to be close to the performance limit of FE codes using CL or NC elements, as these codes simply need many more operations to assemble RHSs of equations than regular grid codes, leaving alone mass matrix inversions.

The CL implementation of FEOM is given by Wang et al. (2008a), and here, we only mention the details that are different from the NC implementation, which is briefly described in Section 2. Section 3 presents experiments with the Munk gyre aimed at evaluating the dependence of numerical viscosity on the stabilization parameter. In Section 4, we present experiments with baroclinic turbulence in a zonally re-entrant channel and compare levels of eddy kinetic energy achieved in both P_1 and P_1^{NC} cases with those of MITgcm. Conclusions and discussion are presented in the last section.

2 Model details

2.1 NC setup

Both NC and CL implementations use CL functions to represent elevation and tracers. The difference between them is in the discretization of horizontal ve-

locity. In the NC case, the horizontal velocities are expanded in basis functions that are products of NC linear functions $P_1^{NC}(x, y)$ with CL functions $P_1(z)$. The NC linear functions are associated with the edges of triangles. They are equal to 1 at their edge and go linearly to -1 at the opposing node. CL functions are associated with nodes. They are equal to 1 at their node and go linearly to 0 at neighboring nodes. The essential advantage of NC linear functions is their orthogonality on triangles. Since the number of edges is approximately three times that of nodes, there is no nullspace of the discrete gradient operator—no spurious pressure modes, and no stabilization is needed. These two properties, (1) the diagonal mass matrix and (2) the absence of stabilization, suggest that the NC representation of velocity is a promising choice for formulating a model. On the downside is the increased number of degrees of freedom and extra computational load in computing momentum advection due to the discontinuity of these functions.

The property of orthogonality is strictly maintained on z -level meshes for which the integration over vertical and horizontal coordinates on elements are independent. In the case of generalized vertical coordinates, elements are no longer rectangular prisms and Jacobians of transform from parent to physical domains are functions of horizontal coordinates. This circumstance destroys the orthogonality property. In order to maintain the efficiency of the code, horizontal lumping or special quadrature rules (see White et al. 2008a) have to be applied in such cases, which partly contaminates the mathematical elegance of the approach. If elements are only gently deformed, the deviations from orthogonality are small and horizontal lumping remains a good compromise. Although the functionality of generalized vertical coordinates is supported in the code, here, we will only discuss the case of z -level grids for brevity.

The momentum equation is discretized using the second-order Adams–Bashforth method for the Coriolis term and the momentum advection. Applying the Adams–Bashforth method to horizontal viscosity terms is not recommended and can even increase instability, but it is convenient and seldom leads to problems. The Coriolis term can be treated semi-implicitly, but this can be beneficial only on coarse meshes and is not discussed here. The contribution from vertical viscosity can be optionally treated implicitly (when viscosity is large and CFL limiting may occur). The contribution from the sea surface height (ssh) is implicit to suppress inertia-gravity waves. For this reason, the method remains only first-order accurate in time. One can easily make it second-order if required at almost no cost (by

taking ssh semiimplicitly), but this is not recommended for large-scale ocean applications. The time-discretized momentum and vertically integrated continuity equations are

$$\delta(\mathbf{u}^{n+1} - \mathbf{u}^n) + g\nabla\eta^{n+1} - \partial_z A_v \partial_z \mathbf{u}^{n+1} = \mathbf{R}^{n+1/2}, \tag{1}$$

$$\delta(\eta^{n+1} - \eta^n) + \nabla \cdot \int_{-H(x,y)}^0 \mathbf{u}^{n+1} dz = 0. \tag{2}$$

Here, $\delta = 1/\Delta t$, where Δt is the time step, A_v is the vertical viscosity coefficient, g the acceleration due to gravity, \mathbf{u} is the horizontal velocity, η is the elevation, and n marks time steps. The RHS of the vertically integrated continuity equation is set to zero for simplicity.

The RHS of Eq. 1 contains all terms of the momentum equation other than time derivative, surface pressure, and vertical viscosity. The Coriolis, pressure, and viscous terms are computed as

$$\mathbf{R}^{n+1/2} = -\nabla p_h^{n+1/2}/\rho_0 + \nabla A_h \nabla \mathbf{u}^n - (3/2 + \varepsilon)(\mathbf{f} \times \mathbf{u} + (\mathbf{v}\nabla)\mathbf{u})^n + (1/2 + \varepsilon)(\mathbf{f} \times \mathbf{u} + (\mathbf{v}\nabla)\mathbf{u})^{n-1}$$

Here, ε is a small constant chosen to stabilize the second-order Adams–Bashforth scheme, p_h is the pressure due to the weight of the fluid counted from $z = 0$, ρ_0 is the reference density, $\mathbf{v} = (\mathbf{u}, w)$ is the full velocity, A_h the horizontal viscosity coefficient, and \mathbf{f} the Coriolis parameter.

The ideology of solving the pair of Eqs. 1 and 2 is standard (pressure correction method) and is similar to that used in the CL setup of FEOM (Wang et al. 2008a) and in other models working with implicit free surface (e.g., MITgcm).

First, during the prediction step, one solves for \mathbf{u}^*

$$\delta(\mathbf{u}^* - \mathbf{u}^n) + \partial_z A_v \partial_z \mathbf{u}^* + g\nabla\eta^n = \mathbf{R}^{n+1/2}. \tag{3}$$

The predicted velocity is then corrected by solving

$$\delta(\mathbf{u}^{n+1} - \mathbf{u}^*) + g\nabla(\eta^{n+1} - \eta^n) = 0. \tag{4}$$

Formally, combining Eqs. 3 and 4, one does not recover the original Eq. 1. There is a small difference due to the omission of the viscous contribution from Eq. 4. This difference, however, does not destroy the time accuracy of the method (cf. with discussion in Ford et al. 2004).

Discretizing Eqs. 3 and 4, one gets the following matrix equations:

$$(\delta M + D)\mathbf{u}^* = \delta M\mathbf{u}^n - g\mathbf{G}\eta^n + \mathbf{R}^{n+1/2}, \tag{5}$$

and

$$(M\mathbf{u}^{n+1} - M\mathbf{u}^*) + g\Delta t\mathbf{G}(\eta^{n+1} - \eta^n) = 0. \tag{6}$$

Here, the notation used for continuous fields is preserved for their discrete counterparts because the meaning is clear from the context. The matrices introduced above are the mass matrix

$$M_{ij} = \int N_i N_j d\Omega,$$

the matrix of vertical viscosity

$$D_{ij} = \int A_v \partial_z N_i \partial_z N_j d\Omega,$$

and of the gradient operator

$$\mathbf{G}_{ij} = \int \mathbf{N}_i \nabla \bar{N}_j d\Omega.$$

Here, N_i and \bar{N}_i are the basis functions used to represent velocity and elevation. Let us look at the structure of M and D matrices. Due to the orthogonality of the horizontal basis functions, these contain only links between vertically aligned nodes. This implies that the problem of matrix inversion is split into E_{2D} (the number of 2D edges) subproblems, each of which can be inverted effectively by the sweep algorithm. Since the number of edges is three times that of the nodes, this inversion is relatively expensive, yet not as expensive as applying iterative solvers to invert the stiffness matrices in the CL case when implicit vertical viscosity is used.

In situations when the vertical viscosity does not introduce CFL limitations for the selected time step Δt , it can be included into the \mathbf{R} term. If, additionally, one lumps the mass matrix in the vertical direction, there is no longer any need for a matrix inversion, and a very effective numerical algorithm follows.

Expressing velocity from Eq. 6, one gets

$$\mathbf{u}^{n+1} = \mathbf{u}^* - g\Delta t M^{-1} \mathbf{G}(\eta^{n+1} - \eta^n). \quad (7)$$

Now, we *first* discretize the vertically integrated continuity Eq. 2 and then substitute Eq. 7 to obtain

$$\delta M_\eta \Delta \eta + g\Delta t \mathbf{G}^T M^{-1} \mathbf{G} \Delta \eta = \mathbf{G}^T \mathbf{u}^*. \quad (8)$$

Here, $\Delta \eta = \eta^{n+1} - \eta^n$, and M_η is the mass matrix of the elevation problem. This step is essentially different from its analog in the CL case where we first rearrange the continuous equations and then apply the FE discretization (see Section 2.2). In the NC case, the discretized vertically integrated continuity equation will be satisfied by \mathbf{u}^{n+1} on completing the time step (solving Eq. 8 and updating the horizontal velocity via Eq. 6).

Assembling matrix $\mathbf{G}^T M^{-1} \mathbf{G}$ is tricky but feasible and has to be done only once during the initialization phase of the model run. If vertical lumping is applied to the velocity mass matrix, the assembly task is substantially simplified.

In summary, the solution algorithm goes through the following steps:

- Compute \mathbf{u}^* from Eq. 5 by inverting $\delta M + D$. This requires solving E_{2D} (the number of surface edges) subsystems of equations for vertically aligned edges. If the mass matrix is vertically lumped and vertical viscosity is taken explicitly, this step becomes elementary.
- Compute elevation from Eq. 8 by inverting the matrix $\delta M_\eta + g\Delta t \mathbf{G}^T M^{-1} \mathbf{G}$. This is the matrix of size N_{2D} (the number of surface nodes), which is assembled outside the time stepping loop.
- Update the velocity according to Eq. 7. This step is elementary if the mass matrix is vertically lumped, and it is associated with the sweep algorithm otherwise.

The computation of vertical velocity and solving the tracer advection-diffusions equation follow the same ideology as in the CL approach and are not discussed here. In particular, the Taylor–Galerkin and FCT advection schemes are available. All physical parameterization options of the CL code are also supported. The code is MPI parallelized and uses PETSc to solve for elevation.

2.2 Brief summary of the CL approach

In order to remove the spurious pressure modes, an analog of Eqs. 5 and 6 is written as

$$(\delta M + D)\mathbf{u}^* = \delta M\mathbf{u}^n - g\gamma \mathbf{G}\eta^n + \mathbf{R}^{n+1/2}, \quad (9)$$

$$(M\mathbf{u}^{n+1} - M\mathbf{u}^*) + g\Delta t \mathbf{G}(\eta^{n+1} - \gamma\eta^n) = 0. \quad (10)$$

Here, the difference to the NC case is in adding a multiplier γ to the η^n term. The strength of stabilization turns out to be proportional to $(1 - \gamma)$. In practical applications, $\gamma = 0.95 - 0.97$ works well, and in some cases, even values closer to 1 lead to a stable algorithm. The major difficulty in solving Eq. 9 is the matrix inversion (horizontal basis functions are not orthogonal on elements), and for numerical efficiency, the vertical diffusion is delegated to the RHS whenever possible. Inverting the mass matrix is then done iteratively, as explained in Wang et al. (2008a), without calling PETSc.

The other essential distinction from the NC case is that, in order to solve for the elevation, one does not use Eq. 10. Instead, one first expresses \mathbf{u}^{n+1} from Eq. 4 modified by including γ and substitutes it into Eq. 2, and *then* discretizes the emerging equation. The difference to the NC algorithm is the replacement of operator $-\mathbf{G}^T M^{-1} \mathbf{G}$ by the Laplacian operator, which

does not support the pressure modes of operator \mathbf{G} . The price for this (necessary) modification is that \mathbf{u}^{n+1} as found from Eq. 10 does not satisfy the vertically integrated continuity equation exactly. It is its unprojected counterpart from Eq. 4 that does. The latter is used to solve for vertical velocity and to advect tracers, while the former is used in the momentum equation. The difference between them is small (it is only due to reprojection on linear functions) but important for the consistency of the code. The need to keep two types of the horizontal velocity is the major conceptual disadvantage of the CL code. One of these velocities (\mathbf{u}^{n+1}) satisfies boundary conditions but does not exactly satisfy the discrete vertically integrated continuity equation. The other one, $\mathbf{u}^* - g\Delta t \nabla(\eta^{n+1} - \gamma\eta^n)$, on the contrary, does it, but it satisfies only weak impermeability boundary conditions. For this reason, one should expect main manifestations of effects from stabilization in boundary layers.

It should be remembered that all methods of stabilizations used in models employing $P_1 - P_1$ discretization have similar problems because the essence of stabilization is the regularization of the vertically integrated continuity equation. They all introduce the Laplacian operator (instead of or in addition to the true operator $-\mathbf{G}^T M^{-1} \mathbf{G}$) into the vertically integrated continuity equation, and in this way, have the same mathematical basis. The approach used by us has the advantage of explicitly providing the expression for the horizontal velocity, which ensures consistency with the vertical velocity equation.

The solution of the dynamical part follows the same three basic steps as in the NC case, but the prediction and correction steps require the inversion of the mass matrix or (mass + vertical viscosity) matrix. The inversion of the mass matrices can be done effectively and does not slow down the algorithm as much as the inversion of the (mass + vertical viscosity) matrix.

Applying the stabilization may, however, affect the dynamics. We address the question of the consequences of using stabilization by comparing the solutions from the code with stabilization (CL) and without it (NC).

2.3 Momentum advection in the NC code

Before describing the results of our numerical experiments, some remarks are necessary on the implementation of the momentum advection in the NC case. There are two points. The first one is on which form of the momentum advection term is preferable, and the second one is on how to properly treat discontinuities.

In order to guarantee momentum conservation, it seems reasonable to use the conservative form of the

momentum advection $\nabla_3 \cdot (\mathbf{v}\mathbf{u})$ where $\mathbf{v} = (\mathbf{u}, w)$ is full velocity and $\nabla_3 = (\nabla, \partial_z)$. Indeed, if it is projected on an appropriate (and for a while) differentiable function $\tilde{\mathbf{u}}$, and integrated by parts, the result is

$$\mathbf{A} = - \int \mathbf{v} \cdot (\mathbf{u} \cdot \nabla_3 \tilde{\mathbf{u}}) d\Omega,$$

with surface integrals over impermeable lateral walls set to zero. This form (with added flux penalties in the NC case because of discontinuity of basis functions) guarantees that the discretized momentum advection sums to zero. Indeed, the sum of test functions on an element is 1, which reduces the elemental part of the integral to zero. Also, the additional penalties always sum to zero in this case too.

The hidden inconsistency here is that, in writing the momentum advection as $\nabla_3(\mathbf{v}\mathbf{u})$ instead of $(\mathbf{v} \cdot \nabla_3)\mathbf{u}$, one exploits the fact that the divergence of the full velocity is zero. This is true in the continuous case, but for the FE discretization, it holds only in the projection on a particular set of functions. This sense is not respected by \mathbf{A} , which implies that enforcing momentum conservation introduces noisy sinks and sources, and our experience is that the form \mathbf{A} (with additional penalties) works only provided that the explicit viscosity is relatively high, and it remains prone to instabilities in flows with “rich” dynamics. It is worth mentioning that a similar difficulty with the tracer equations is avoided by respecting the consistency requirement on functional spaces used to represent tracers and vertical velocity (see White et al. 2008b; Wang et al. 2008a).

The nonconserving form $(\mathbf{v} \cdot \nabla_3)\mathbf{u}$ works stably and allows for an order of magnitude smaller explicit viscosity. We use this form in numerical experiments. One more stably working variant is to project the horizontal velocity on P_1 functional space to get \mathbf{u}_c , where the subscript “c” stands for “continuous,” and write the momentum advection term as $(\mathbf{v} \cdot \nabla)\mathbf{u}_c$. The advantage of this form is that it is globally conservative ($\int (\mathbf{v} \cdot \nabla)\mathbf{u}_c d\Omega = 0$ because w is solved by projecting it on functions from the same space as \mathbf{u}_c) and that it does not require taking into account continuity penalties across vertical faces of elements. Its potential disadvantage is that a numerically efficient way of solving for \mathbf{u}_c is achieved by lumping the mass matrix, which smoothes the momentum advection and thus introduces a mild scheme dissipation. We compare the performance of both advection schemes in Section 4.

We will now address the other question on how to correctly write additional penalties arising due to the discontinuous nature of NC elements. Writing the penalties is straightforward if equations are written in

the conservative form because fluxes are to be continuous across the boundaries of elements. This ideology cannot be applied to the nonconservative form of the momentum advection. A suitable method is to represent velocity as $\mathbf{u} = \sum_e \mathbf{u}_e \theta_e$ and introduce the same representation for test functions. Here, summation is over elements; \mathbf{u}_e denotes velocity on element e and θ_e equals one inside element e and is zero outside it. With this representation, velocities and test functions are defined everywhere in the computational domain. On manipulating products of θ_e and delta-functions arising from the differentiation, one splits the form like \mathbf{A} into a sum of integrals over interiors of elements and singular contributions from the vertical faces. For the nonconservative momentum advection form, the result is

$$\int \tilde{\mathbf{u}}(\mathbf{v} \cdot \nabla_3) \mathbf{u} d\Omega = \sum_e \int_e \tilde{\mathbf{u}}(\mathbf{v} \cdot \nabla_3) \mathbf{u} dS - \sum_{\text{v.f.}} \int_{\text{v.f.}} \langle \tilde{\mathbf{u}} \rangle \langle \mathbf{u} \mathbf{n} \rangle [\mathbf{u}] dS.$$

Here, subscript “v.f.” stands for vertical faces between elements, \mathbf{n} is the normal to the respective face, and the notations $[\]$ and $\langle \rangle$ of Hanert et al. (2005) for the difference and half sum are used. Such a form of singular terms is only valid for NC elements, and additional terms appear in the general case. It is easy to see that the momentum advection remains nonconservative because, on setting the test function to unity, all terms do not sum to zero. Yet, it produces a consistent approximation.

3 Influence of stabilization on representation of boundary layers

CL and NC setups are applied to simulate Munk gyre flow driven by the wind stress

$$\tau_x = -\tau_0 \cos(\pi y/L)$$

with $\tau_0 = 0.1 \text{ N/m}^2$ in a rectangular box of 1,500 by 1,500 km on a β -plane $f = f_0 + \beta y$, where $f_0 = 10^{-4} \text{ s}^{-1}$, and $\beta = 2 \times 10^{-11} \text{ m}^{-1} \text{ s}^{-1}$ (L is the length of the box in the meridional direction and y is counted from the southern wall). The horizontal viscosity is set to $A_h = 540 \text{ m}^2 \text{ s}^{-1}$. The thickness of the western boundary layer scales as $\delta_M = (A_h/\beta)^{1/3} = 30 \text{ km}$. The real width of this layer (defined as the distance to zero crossing of the meridional velocity) is larger and depends on the boundary conditions. It is approximately $3\delta_M$ for the no-slip case. The Munk problem with its

narrow boundary layers provides a strict test for the effect of (over) stabilization.

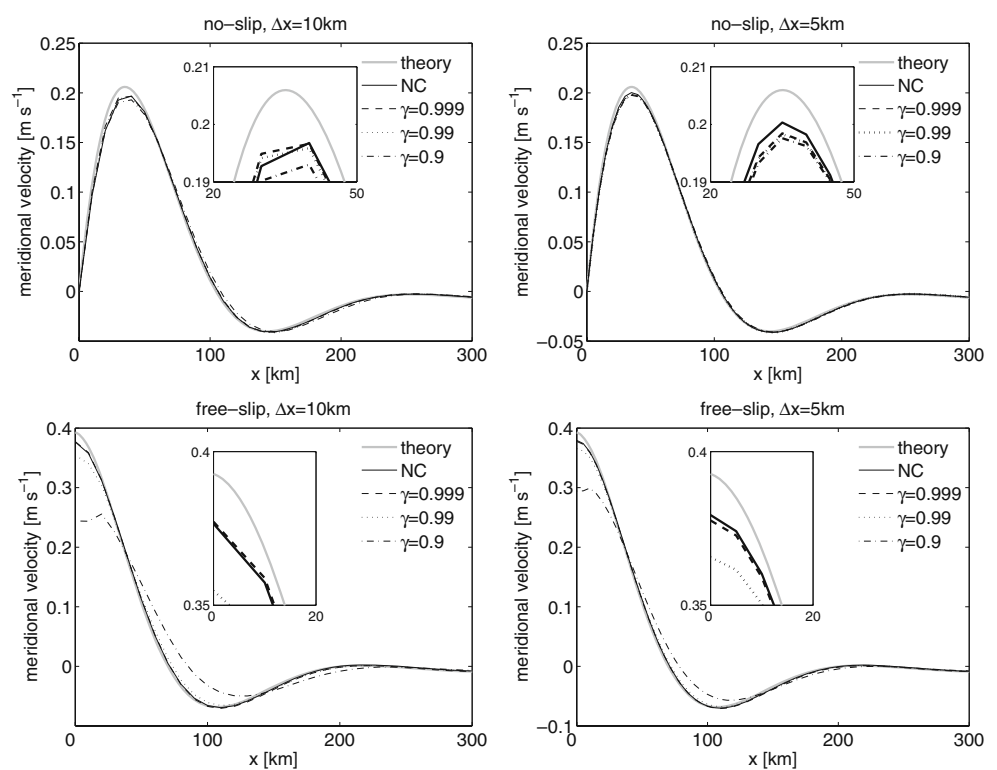
Two different horizontal resolutions of 10 and 5 km are used in the region of the western boundary. The resolution is reduced smoothly to 50 km in the interior of the domain. In the vertical, 10 unevenly spaced levels are used. There is no baroclinic forcing (no stratification), and momentum advection is off in order to compare model solutions with existing analytical solutions (see, for example, Pedlosky 1996). The experiments were run with time step of 1 h for 3 years. This time is still insufficient to reach a fully steady state, but deviations from equilibrium are already very small.

Two sets of experiments were carried out, one with no slip and the other with free slip boundary conditions. In both sets, CL runs were conducted with stabilization parameter $\gamma = 0.9, 0.99, \text{ and } 0.999$. Upper panels of Fig. 1 show the comparison of model and analytical solutions for the no-slip case. It displays the meridional velocity profile drawn across the central latitude of the domain. There is no large difference between the CL and NC solutions in this case, and both discretizations lead to results that are very close to the analytical solution (the existing difference between model solutions and the analytical solution is due to the finite resolution). The influence of stabilization remains on a moderate level beginning from $\gamma = 0.9$, in agreement with our practice. The effects of stabilization are becoming noticeable below this value, and deviation of the velocity amplitude from the analytical solution can be in excess of 15% for $\gamma = 0.5$ (not shown).

However, the influence of stabilization can be much stronger in free-slip cases, as shown in the lower panels of Fig. 1. While the NC solution provides a very reasonable approach to the analytical solution, the CL solutions approach the NC results only at very low levels of stabilization. By comparing profiles of meridional velocity, it can easily be concluded that, except for the case of very low stabilization, there is extra friction against the western wall introduced by the presence of stabilization. Indeed, the part of the jet adjacent to the wall is retarded. This friction is small and does not distort the no-slip boundary layers (which are frictional by themselves). It broadens free-slip western boundary layers to accommodate for the gyre transport with reduced velocity amplitude. In free-slip cases, using $\gamma = 0.9$ produces results that are hardly acceptable.

The effect of stabilization becomes less pronounced as the mesh is refined but the horizontal viscosity coefficient is kept fixed. In such cases, boundary layers become better resolved and the same values of the stabilization parameter lead to smaller deviations from the analytical solutions. We observe this by comparing

Fig. 1 The meridional velocity profile in the Munk gyre. The resolution is 10 km (left panels) and 5 km (right panels) in the western boundary layer. Upper panels correspond to the no-slip case and the lower panels correspond to the free-slip case. The influence of stabilization is small even for $\gamma = 0.9$ in the no-slip case but very strong in the free-slip case



results obtained on 10- and 5-km grids (left and right panels in Fig. 1). In the no-slip case, there is practically no difference between cases with three different γ , yet the NC solution is slightly more accurate and closer to the analytical solution. The agreement clearly improves in the free-slip case too, but $\gamma = 0.9$ still leads to significant errors. In the limit of very fine resolution, the effect of stabilization will be small. This can hardly lead to practical consequences, as in ocean modeling, one tries to reduce viscosity together with increasing resolution, thus making boundary layers thinner. So the practical recipe is rather to keep the stabilization as small as possible in order to minimize its effects.

4 Baroclinic turbulence in a channel

The question that still remains is what happens outside the boundary layers. In order to qualitatively answer, it we carry out a set of numerical experiments on baroclinic turbulence in a zonally reentrant channel. The turbulence statistics are rather sensitive to the explicit and scheme-implicit dissipation, and in this way, the effects of stabilization can also be diagnosed. This comparison is not so clean as in the Munk gyre case, as the properties of numerical baroclinic turbulence are affected by the momentum advection, which

is different in CL and NC cases (tracer advection is implemented in the same way but can also be affected by different representations of the horizontal velocity). The experiments are conducted in a channel of 15 degrees in latitude and 40 degrees in zonal direction centered at 37.5° N at a resolution of about 16 km. In the vertical direction, the mesh contains 16 equally spaced levels going to the depth of 1,500 m. A linear equation of state is used with stratification being due to temperature only. The initial temperature distribution has horizontal and vertical gradients of 0.5×10^{-5} and 8.2×10^{-3} K/m, respectively. The stratification is maintained by relaxation to the initial profile in 1.5° northern and southern zones with relaxation coefficient dropping from $1/3 \text{ day}^{-1}$ at the walls to zero outside the 1.5° zones.

Both (CL and NC) cases are run with the FCT advection scheme with explicit horizontal diffusivity of $50 \text{ m}^2/\text{s}$. For the NC case, runs with true NC and reprojected implementations of the momentum advection are carried out. The biharmonic viscosity of $2.5 \times 10^{10} \text{ m}^4/\text{s}$ is used, both versions are using (the same) explicit vertical viscosity and diffusion, and the mass matrix of the momentum equation is vertically lumped in the NC case for numerical efficiency.

In addition to comparing the performance of two versions of FEOM, similar computations were performed

with MITgcm (Marshall et al. 1997). The same mesh, initial temperature stratification, relaxation in southern and northern buffer zones, and physical parameters were used in these simulations. This serves two goals. The first one is to confront FEOM results with those of a more traditional model. The second one is the CPU time comparison. It is frequently mentioned that unstructured grid FE codes are slower (for the same amount of nodes) than structured grid codes, and here, we took the opportunity to measure the extent to which they are slower.

Each case is run for 3 years and eddy kinetic and eddy available potential energies are compared in the upper (kinetic) and lower (available potential) panels of Fig. 2. The eddy part of velocity and density fields are defined here as deviations from zonal mean values, and computation of the available potential energy follows the quasigeostrophic rules. As an illustration that the flow is in a turbulent regime, Fig. 3 shows a snapshot of the temperature field displaying a variety of eddy features characteristic of the well developed baroclinic turbulence. Since the channel is relatively wide, the characteristics of eddy flow are sensitive to the details of dissipation and, in particular, of the bottom drag. The latter is parameterized through the quadratic law with

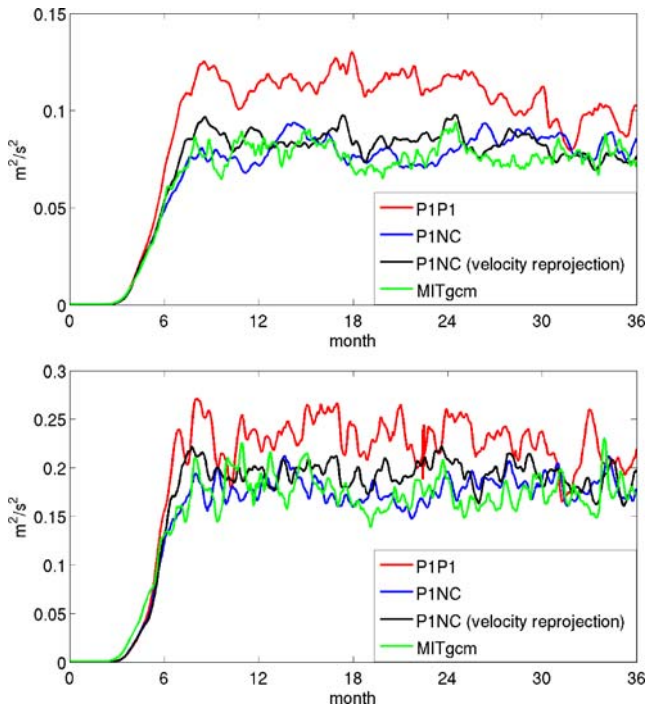


Fig. 2 Doubled eddy turbulent kinetic energy (*upper panel*) and eddy available potential energy (*lower panel*) in CL FEOM, NC FEOM (projected and true momentum advection), and MITgcm. The CL version simulates higher levels of energy compared to the NC version

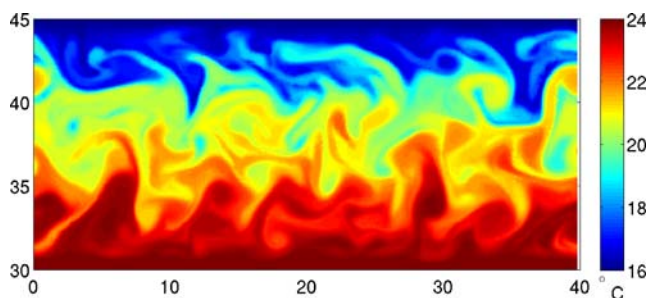


Fig. 3 A snapshot of temperature field at 100-m depth at the end of integration in the CL FEOM

the drag coefficient $C_d = 0.0025$. Removing the bottom drag leads to the appearance of large eddies occupying the entire channel in width.

The levels of eddy kinetic and eddy available potential energies are similar in CL and NC setups. In the NC case, both momentum advection schemes show similar results, which is counterintuitive, because the reprojected variant of momentum advection is expected to introduce some dissipation. The other counterintuitive fact is that CL levels of the eddy kinetic and available potential energies are slightly higher than the levels of the NC setup. They, however, show a tendency toward the NC results. The conclusion is that the stabilization present in the CL version does not damage the performance and does not lead to additional dissipation inside the domain.

The MITgcm results match very closely the results of both NC versions. Given the sensitivity of the eddy energy levels to the total dissipation (explicit and scheme-implicit) and recognizing the differences in discretization and implementation, the agreement between all simulations is surprisingly good.

5 Numerical efficiency

Despite the fact that no true matrix inversion is needed in the predictor and corrector steps of the NC setup in our experiments (explicit vertical viscosity and vertical lumping are used), the NC code is only marginally faster (less than 10%) than the CL code, which always does mass matrix inversions to solve for \mathbf{u}^* and full velocity. One model year of the channel experiment (mesh of 0.4 million nodes) takes approximately 8 h on a single node (8 Power4 1.7-GHz processors) of IBM p655 in both implementations with $\Delta t = 15$ min.

Although this performance is already reasonably fast, it is still much slower than the performance of structured grid codes. The performance of the MITgcm on the same mesh (but without triangulation) and on

the same architecture is considerably faster. On a single processor, CL FEOM CPU time per time step compares to that of MITgcm as $6.5/0.35 \text{ s} \approx 18$, while on the full node (eight processors), the ratio is $0.85/0.14 \text{ s} \approx 6$. Note that the MITgcm code does not scale optimally because of the details of the architecture and the size of the mesh. This gives a slowness factor of about 10, which can perhaps be only slightly improved for FE $P_1 - P_1$ or $P_1^{nc} - P_1$ codes on unstructured triangular grids. The reason is, as mentioned in Section 1, a much higher number of operations needed to assemble the RHSs of equations on unstructured meshes, the need to invert mass matrices, and the relatively high cost of the FCT algorithm on unstructured meshes.

The essential part of the CPU load in our implementation of NC code is the assembly of the RHS of the momentum equation. The edge penalty part of the momentum advection term is not so expensive by itself, but it becomes more so in parallel implementation where the cycle over edges is to be repeated twice, first to collect combinations of velocities on edges (which are to be communicated between neighboring processors) and then to compute the contributions to the RHS. This splitting also requires additional storage. It turns out that the predictor/corrector steps together require comparable time in NC (the NC version is slightly faster) and CL setups, although in the latter case, mass matrix inversions are needed.

The analysis of CPU time per model time step shows that the horizontal velocity part takes about 40% of the total time, and 10% more are required to solve for elevation. This share depends on the tracer advection scheme and on how many tracers are advected. The FCT scheme is relatively expensive, and in the channel experiments, only temperature is advected (it takes 33%). Adding salinity would increase the weight of the tracer part to 50%, and it will dominate if more tracers are used.

The relationship between the CPU time of CL and NC cases can change in some special cases. For example, the CL code slows down in the case of free-slip boundary conditions (rotations are performed to local normal and tangential directions). In this case, one deals with matrices of doubled size (because velocity components become coupled) and uses a more sophisticated algorithm for matrix inversion. This leads to about 30% overhead. In the same vein, using local coordinate frames on elements (instead of spherical coordinates) in future setups will lead to more overhead in the CL case than in the NC case. Finally, the CL code is more expensive in the implicit vertical viscosity option (unless horizontal lumping is applied). These are the reasons why the NC setup is of further

potential interest. It should, however, be mentioned that the CL setup is more robust with respect to the choice of viscosity and time step, so it remains a good choice too.

6 Conclusions

In this paper, we tried to estimate effects of stabilization of CL of FEOM by comparing model results to analytical solutions and to results of the NC setup, which does not require stabilization. The artifacts introduced by stabilization depend on the problem under consideration and can be strong close to free-slip boundaries. Except for this, there is no strong influence provided the stabilization strength is sufficiently small ($\gamma \geq 0.95$). The broadening of the western boundary current in the Munk gyre experiment remains small, and both eddy TKE and APE in the turbulent flow in a zonally reentrant channel are not affected by stabilization and reach approximately the same levels in both CL and NC cases. The overall conclusion from this comparison is that stabilization does not damage solutions in many cases but should be kept as small as possible in order to minimize its influence in boundary layers. Surely, the simple cases considered here do not exclude other possible effects that may emerge in longer simulations and under other circumstances, but they simply give an idea of consequences of using the stabilization. The potential of the CL code is illustrated by a study of overflows in Wang et al. (2008b), and other examples with both setups will be presented elsewhere.

The intercomparison of CL and NC discretizations as implemented in FEOM supports the currently forming opinion (Hanert et al. 2005; Le Roux et al. 2005) that NC discretization of horizontal velocities is a good alternative to CL horizontal velocities mostly used thus far in FE codes. The obvious advantage of the NC representation of the horizontal velocity is the absence of pressure modes, which leads to a cleaner pressure correction algorithm and consistent vertically integrated continuity equation. Orthogonality of NC basis functions in horizontal directions is another advantage, but taking it rigorously into account requires z -levels in the vertical direction. On generalized vertical grids with deformed elements, the orthogonality is destroyed and horizontal lumping (or special quadrature rules) is needed to get a numerically efficient algorithm. The robustness of different implementations in such situations remains to be explored.

In practice, the advantages of the NC elements do not lead to a faster code, as the number of operations needed to assemble RHSs is as high as or greater

than with the CL discretization. In perspective, working toward codes that do not need a longitude–latitude basis may change the situation in favor of the NC implementation.

The progress in the acceptance of unstructured grids for modeling large-scale ocean circulation is heavily dependent on the numerical efficiency of unstructured grid models, and the important task for future research is to establish the optimal approach. The current work is a step in this direction.

References

- Carrère L, Lyard F (2003) Modeling the barotropic response of the global ocean to atmospheric wind and pressure forcing—comparisons with observations. *Geophys Res Lett* 30:1275
- Cassuli V, Walters RA (2000) An unstructured grid, three-dimensional model based on the shallow water equations. *Int J Numer Methods Fluids* 32:331–348
- Chen C, Liu H, Beardsley RC (2003) An unstructured grid, finite-volume, three-dimensional, primitive equations ocean model: applications to coastal ocean and estuaries. *J Atmos Ocean Technol* 20:159–186
- Codina R, Soto O (1997) Finite element solution of the Stokes problem with dominating Coriolis force. *Comput Methods Appl Mech Eng* 142:215–234
- Codina R, Zienkiewicz OC (2002) CBS versus GLS stabilization of the incompressible Navier–Stokes equations and the role of the time step as the stabilization parameter. *Commun Numer Methods Eng* 18:99–112
- Danilov S, Kivman G, Schröter J (2004) A finite element ocean model: principles and evaluation. *Ocean Model* 6:125–150
- Ford R, Pain CC, Piggott MD, Goddard AJH, de Oliveira CRE, Uplemby AP (2004) A non-hydrostatic finite-element model for three-dimensional stratified flows. Part I: model formulation. *Mon Weather Rev* 132:2832–2844
- Fringer OB, Gerritsen M, Street RL (2006) An unstructured-grid, finite-volume, nonhydrostatic, parallel coastal ocean simulator. *Ocean Model* 14:139–173
- Hanert E, Le Roux DY, Legat V, Deleersnijder E (2005) An efficient Eulerian finite element method for the shallow water equations. *Ocean Model* 10:115–136
- Le Roux DY, Sène A, Rostand V, Hanert E (2005) On some spurious mode issues in shallow-water models using a linear algebra approach. *Ocean Model* 10:83–94
- Lynch DR, Ip JTC, Naimie, CE, Werner FE (1996) Comprehensive coastal circulation model with application to the Gulf of Maine. *Cont Shelf Res* 16:875–906
- Marshall J, Adcroft A, Hill C, Perelman L, Heisey C (1997) A finite-volume, incompressible Navier–Stokes model for studies of the ocean on parallel computers. *J Geophys Res* 102:5753–5766
- Pain CC, Piggott MD, Goddard AJH, Fang F, Gorman GJ, Marshall DP, Eaton MD, Power PW, de Oliveira CRE (2005) Three-dimensional unstructured mesh ocean modelling. *Ocean Model* 10:5–33
- Pedlosky J (1996) *Ocean circulation theory*. Springer, Berlin Heidelberg New York
- Stuhne GR, Peltier WR (2006) A robust unstructured grid discretization for 3-dimensional hydrostatic flows in spherical geometry: a new numerical structure for ocean general circulation modeling. *J Comput Phys* 213:704–729
- Walters RA (2005) Coastal ocean models: two useful finite-element methods. *Cont Shelf Res* 25:775–793
- Wang Q, Danilov S, Schröter J (2008a) Finite element ocean circulation model based on triangular prismatic elements, with application in studying the effect of topography representation. *J Geophys Res* 113:C05015
- Wang Q, Danilov S, Schröter J (2008b) Comparison of overflow simulations on different vertical grids using the finite element ocean circulation model. *Ocean Model* 20:313–335. doi:10.1016/j.ocemod.2007.10.005
- Westerink JJ, Luettich RA, Blain CA, Scheffner NW (1992) ADCIRC: an advanced three-dimensional circulation model for shelves, coasts and estuaries; report 2: users manual for ADCIRC-2DDI. Contractors report to the US Army Corps of Engineers. US Army Corps of Engineers, Washington D.C.
- White L, Deleersnijder E, Legat V (2008a) A three-dimensional unstructured mesh shallow-water model, with application to the flows around an island and in a wind driven, elongated basin. *Ocean Model* 22:26–47
- White L, Legat V, Deleersnijder E (2008b) Tracer conservation for three-dimensional, finite element, free-surface, ocean modeling on moving prismatic meshes. *Mon Weather Rev* 136:420–442
- Zhang Y, Baptista AM (2008) SELFE: A semi-implicit Eulerian–Lagrangian finite-element model for cross-scale ocean circulation. *Ocean Model* 21:71–96
- Zhang YL, Baptista AM, Myers EP (2004) A cross-scale model for 3D baroclinic circulation in estuary–plume–shelf systems: I. Formulation and skill assessment. *Cont Shelf Res* 24:2187–2214
- Zienkiewicz OC, Taylor RL (2000) *The finite element method*. V. 3. Fluid dynamics. Butterworth–Heinemann, Oxford

Process study of coastal circulation over the inner Scotian Shelf using a nested-grid ocean circulation model, with a special emphasis on the storm-induced circulation during tropical storm Alberto in 2006

Bo Yang · Jinyu Sheng

Received: 2 March 2008 / Accepted: 8 September 2008 / Published online: 7 October 2008
© Springer-Verlag 2008

Abstract This study examines main physical processes affecting the three-dimensional (3D) circulation and hydrographic distributions over the inner Scotian Shelf (ISS) in June and July 2006 using a nested-grid coastal ocean circulation modeling system known as the NCOPS-LB. The nested-grid system has five relocatable downscaling sub-models, with the outermost submodel of a coarse horizontal resolution of $(1/12)^\circ$ for simulating storm surges and barotropic shelf waves over the Eastern Canadian shelf and the innermost submodel of a fine resolution of ~ 180 m for simulating the 3D coastal circulation and hydrography over Lunenburg Bay of Nova Scotia in the default setup. The NCOPS-LB is driven by meteorological and astronomical forcing and used to study the storm-induced circulation over the ISS during tropical storm Alberto. Model results demonstrate that the coastal circulation and hydrographic distributions over the ISS are affected significantly by tides, local wind forcing, and remotely generated coastal waves during the study period.

Keywords Physical process · Nested-grid · Numerical modeling · Scotian shelf · Storm-induced · Tropical storm Alberto

1 Introduction

The Eastern Canadian Shelf (ECS) referred to in this study consists of the Labrador Shelf (LS), Newfoundland Shelf (NS), the Grand Banks (GB), the Gulf of St. Lawrence (GSL), the Scotian Shelf (SS), and the Gulf of Maine (GOM). The three-dimensional (3D) circulation and hydrographic distributions in this region are affected significantly by wind, tides, buoyancy forcing associated with sea surface heat and freshwater fluxes and river runoff, and occasionally by tropical storms and hurricanes (Smith and Schwing 1991). On September 26, 2003, for an example, Hurricane Juan moved northward from Bermuda and made landfall near Halifax with a maximum sustained wind speed of about 160 km h^{-1} and caused significant intense inertial currents in the upper waters of the SS and localized coastal flooding along coastlines of Nova Scotia (Fogarty et al. 2006; Sheng et al. 2006). On June 15, 2006, tropical storm Alberto moved northeastward over the Canadian Atlantic Provinces with a maximum peak wind of about 100 km h^{-1} , resulting in large sea states over the ECS. Significant progress has been made in recent years in understanding the physical processes of storm-induced circulation over the ECS (Petrie et al. 1987; Davidson et al. 2001; Zhai et al. 2004; Sheng et al. 2006; Wang et al. 2007; Zhai et al. 2008a). Many important scientific issues, however, remain to be addressed, including the interaction of the storm-induced circulation with coastal topography, dynamics of coastal upwelling and eddy development, and characteristics of coastal cool water plumes.

A coastal ocean observatory was established in summer 2000 in Lunenburg Bay (LB), Nova Scotia as part of a multi-agency research program of marine environmental observation and prediction for Canadian Atlantic coastal waters known as Center for Marine Environmental Prediction (CMEP)-Bay (Sheng and Wang 2004, www.cmpe.ca). The

Responsible editor: Eric Deleersnijder

B. Yang · J. Sheng (✉)
Department of Oceanography, Dalhousie University,
Halifax, Nova Scotia B3H 4J1, Canada
e-mail: Jinyu.Sheng@Dal.Ca

B. Yang
Physical Oceanography Laboratory, Ocean University of China,
Qindao 266100, People's Republic of China

main purpose of CMEP-Bay was to develop a marine environmental observation and prediction system for Canadian Atlantic coastal waters, using data assimilative and coupled models guided directly by real-time observations (Safer 2002; Wang et al. 2007). As part of the CMEP-LB, a nested-grid coastal ocean circulation prediction system known as the NCOPS-LB was developed (Sheng and Yang 2008; 2008, unpublished manuscript) based on the integration of a prototype operational shelf circulation forecast system known as Dalcoast3 (Thompson et al. 2007; Ohashi et al. 2008, unpublished manuscript) and a two-level nested-grid coastal circulation model for LB (Zhai et al. 2008a; Sheng et al. 2008). The main objective of this study is to use the NCOPS-LB to examine the main physical processes influencing the three-dimensional circulation and hydrographic distributions over the inner Scotian Shelf (ISS) in June and July 2006, with a special emphasis on the storm-induced coastal circulation during tropical storm Alberto.

The remainder of the paper is arranged as follows: Section 2 discusses the physical conditions over the SS during Alberto based on the in situ and satellite remote sensing observations. Section 3 describes the model setup and external forcing. Section 3 presents the simulated circulation and hydrographic distributions over the study region during Alberto. Section 4 examines the main physical processes influencing the 3D circulation over the ISS based on the model results in three numerical experiments. The final section is the summary and conclusion.

2 Meteorological and oceanographic conditions during tropical storm Alberto

Tropical storm Alberto was the first tropical storm of the 2006 Atlantic hurricane season (www.nhc.noaa.gov). It formed on

June 10, 2006 as a tropical depression in the northwestern Caribbean Sea and moved northwestward to the southeastern Gulf of Mexico. It became a tropical storm in the morning of June 11 with the peak intensity of about 110 km h^{-1} . The storm moved northeastward through Northern Florida and weakened to a tropical depression in the early morning of June 14 over Georgia. The storm emerged off the mid-Atlantic coast of the United States at the late night of June 14 and accelerated northeastward to the Canadian Atlantic provinces (Fig. 1). The storm deepened into a powerful post-tropical storm as it swept the Scotian Shelf on June 15 with the maximum peak wind of about 100 km h^{-1} .

Figure 2 presents the sea level atmospheric pressures (SLPs) and wind stress over the SS-GSL region calculated or extracted from numerical weather forecasts produced by the Meteorological Service of Canada (MSC). At day 166.75 (18:00 UTC June 15), the storm center was located at the southern SS with a maximum wind stress of $\sim 1.5 \text{ Pa}$, and the wind stress was southwestward over the ISS (Fig. 2a). At day 167.0 (00:00 June 16), the storm center approached the central SS with a maximum wind stress of $\sim 2 \text{ Pa}$ on the right hand side of the track (Fig. 2b). The wind stress was roughly onshore over the eastern SS and offshore over the western SS at this time. The storm center moved to the northeastern SS at day 167.25 (06:00 June 16) with a maximum wind stress of $\sim 1 \text{ Pa}$, and the wind stress in LB at this time was relatively weak and roughly eastward (Fig. 2c). The storm reached the western GB and made landfall over Southern Newfoundland at the noon of June 16, with a maximum wind stress of $\sim 0.75 \text{ Pa}$ (Fig. 2d).

The six hourly positions of “best track” of Alberto (Fig. 2) were constructed by the Canadian Hurricane Center (CHC) from multisource measurements and can be used to validate the performance of the MSC numerical weather model in predicting the center position of the storm. The storm track of

Fig. 1 Storm track of tropical storm Alberto in June 2006. Abbreviations are used for the Newfoundland (NF), Labrador Sea (LS), Grand Banks (GB), Flemish Cap (FC), Gulf of St. Lawrence (GSL), Gulf of Maine (GOM), Scotian Shelf (SS), and Lunenburg Bay (LB)



Alberto was very close and parallel to the south coast of Nova Scotia as shown in Figs. 1 and 2, with a typical translational speed of 13 m s^{-1} over the SS. In comparison with the CHC best storm track, the MSC numerical weather forecast model reproduces reasonably well the movement of the storm from days 166.75 to 167.50 (Figs. 2a,d). There are some noticeable differences between the storm centers produced by the MSC numerical forecast model and the CHC best storm track from days 167.00 to 167.25 (Figs. 2b,c).

The 3D circulation and hydrographic distributions on the SS were affected significantly by the storm. Figure 3 presents the satellite-measured sea surface temperatures (SSTs) over the ISS on June 17, 2006, 1 day after Alberto swept the study region. The satellite-measured SSTs were constructed by Chris Jones (personal comm. 2007) at Dalhousie University from remote sensing measurements of the moderate resolution imaging spectroradiometer (MODIS) using a two-dimensional (2D) optimal interpolation scheme. The MODIS SSTs on June 17 were characterized by relatively cool surface waters over areas between isobaths of 100 and 170 m and relatively warmer surface waters near the coast and further offshore. There are two plausible

reasons for the appearance of narrow bands of cool surface waters in Fig. 3: one is due to the SST cooling associated with the passage of Alberto, and the other is due to the advection of relatively cold surface waters from the GSL by the inshore Scotian Current. We will discuss in Section 4 that the passage of Alberto is the primary reason for narrow bands of cool surface waters over the ISS on June 16, 2006.

Temperature and salinity observations were made at station 2 of Halifax Line on days 159 (June 8), 187 (July 16), 201 (July 20), and 214 (August 2) in 2006 (Fig. 4), respectively, by the Bedford Institute of Oceanography (Blair Greenan, personal comm., 2007). The vertical stratification at station 2 in June and July 2006 was characterized by relatively uniform temperature and salinity in the surface mixed layer (SML) and a sharp thermocline at the bottom of the SML with a significant decrease in temperature (with a gradual increase in salinity). Beneath the thermocline, there were an upper intermediate layer between 40 and 80 m with relatively uniform temperature and a gradual increase in salinity, a lower intermediate layer between 80 and 150 m with gradual increases in both temperature and salinity, and a deep layer of deeper than 150 m with relatively uniform

Fig. 2 SLPs and wind stress over the Scotian Shelf and Gulf of St. Lawrence during tropical storm Alberto extracted or calculated from the three hourly numerical weather forecasts produced by the MSC at: **a** day 166.75 (18:00 UTC June 15), **b** day 167.00 (00:00 June 16), **c** day 167.25 (06:00 June 16), and **d** day 167.50 (12:00 June 16) of 2006. Wind stress vectors are plotted at every second model grid point of the MSC numerical weather model. The contour interval for SLPs is 2 hPa. The *solid (open) circles* represent the current (non-current) position of the six hourly “best track” of produced by the Canadian Hurricane Center (www.atl.ec.gc.ca/wether/hurricane/images)

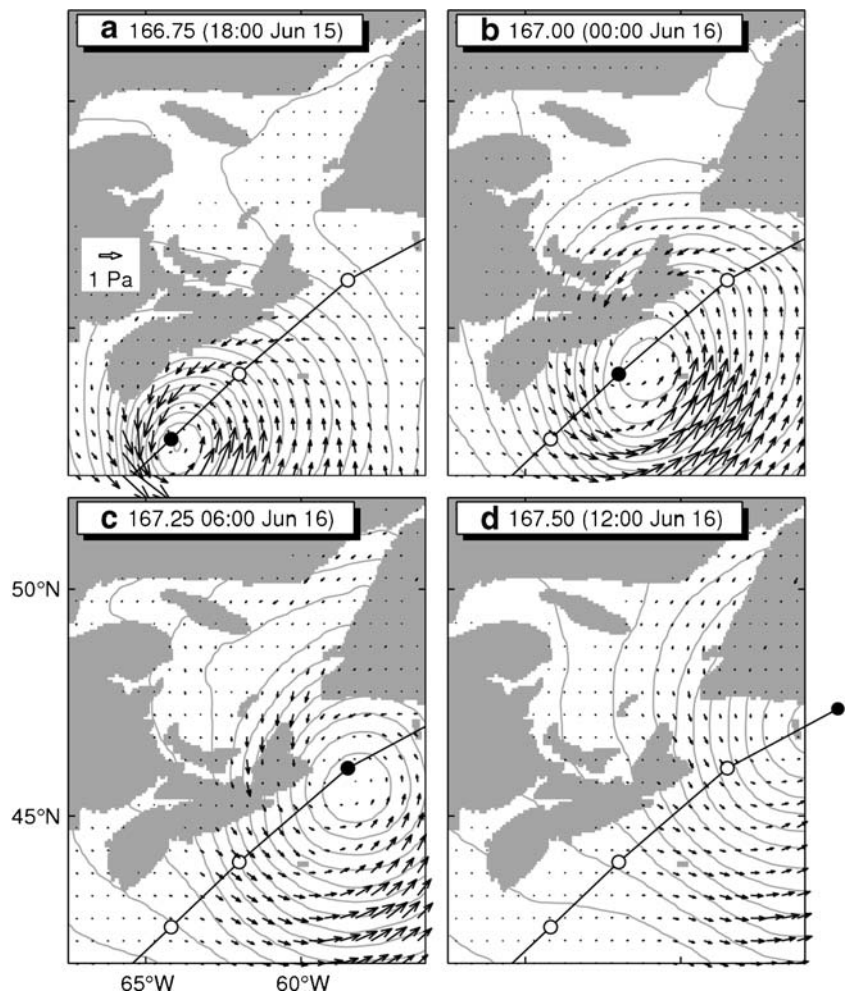
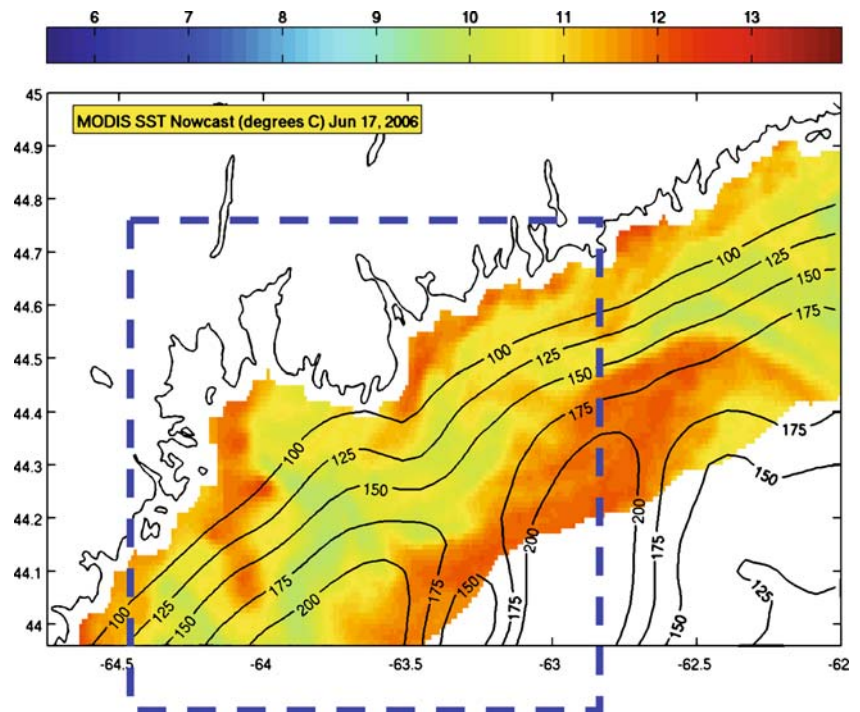


Fig. 3 SSTs over the inner Scotian Shelf on June 17, 2008, which is 1 day after the passage of Alberto, constructed from MODIS measurements (courtesy of Chris Jones, Dalhousie University). The area marked by the dashed box is the domain of submodel L3 shown in Fig. 17



temperature and salinity at station 2 in June and July 2006 (Fig. 4). The thickness of the SML at station 2 was ~20 m on day 159 and ~10 m on other three observation days. The observed temperature and salinity in the SML at the site were about 9°C and 30.2 ppt on day 159; about 12.5°C and 31 ppt on day 187; and about 18.5°C and 31 ppt on day 201, respectively. Significant vertical temperature gradients occurred in the thermohaline on days 187 and 201, and the thickness of the upper intermediate layer was relatively thick on day 159 and relatively thin on days 187 and 201.

The near-surface (1 m) temperatures observed at the two mooring sites in LB (Fig. 5) exhibit a similar warming trend as those at station 2 from days 155 to 166 but with significant high-frequency variability. Vertical stratifications were present during this study period, except for two storm events around days 159 and 167, respectively. The vertical hydrographic stratification was reduced significantly during

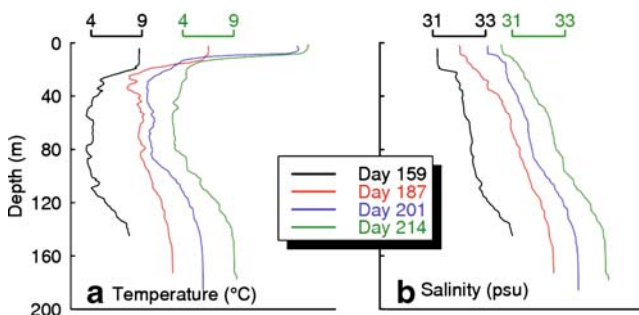


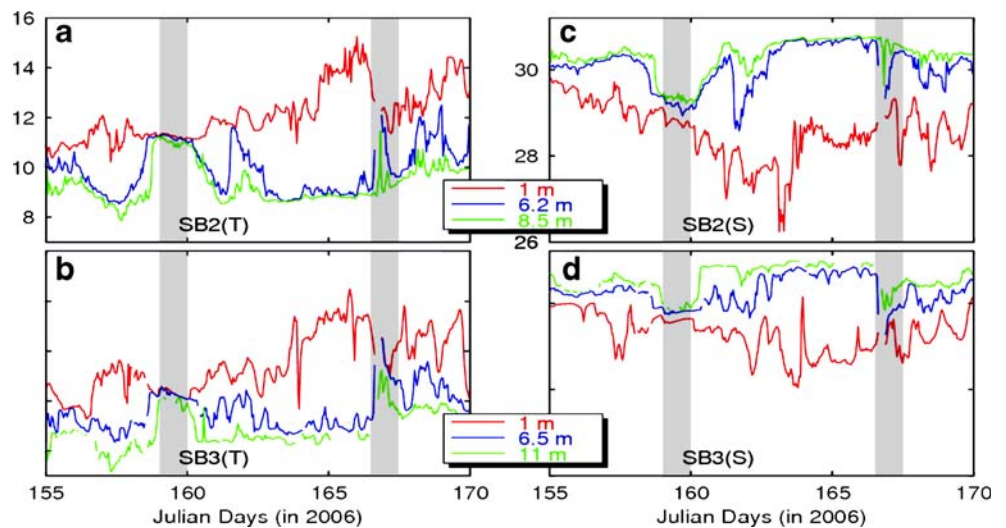
Fig. 4 Vertical profiles of observed temperatures (left) and salinities (right) at station 2 of Halifax Line at days 159 (June 8, black), 187 (June 16, red), 201 (June 20, blue), and 214 (August 2, green) in 2006

the two storm events and reestablished gradually after the storms. The observed near-surface salinities in LB decreased from 30 to 28 psu with significant high-frequency variability. The power spectral analysis of observed near-surface salinities at SB2 and SB3 demonstrates that temporal variability of salinity in the upper water column in LB was mainly dominated by the energy in the semidiurnal frequency band and in the subtidal frequency band of ~1 cpd (not shown). The latter frequency band was associated most likely with the propagation of baroclinic waves excited by the storms propagating from the Rose Bay to Lunenburg Bay as discussed in Zhai et al. (2008b).

Vertical profiles of horizontal currents observed by the ADCPs at mooring sites SB2 and SB3 in LB shown in Fig. 6 demonstrate that the horizontal currents in LB had significant variability in the tidal and synoptic weather frequency band (ADCP data in the top 3 m were not shown in Fig. 6 due to the acoustic scattering from air bubbles). During the first storm around day 159, relatively large north-eastward currents occurred at SB2, with a maximum speed of ~20 cm s⁻¹ at 8 m depth and in the roughly opposite direction to the wind forcing, indicating that large subsurface currents were associated mainly with the return flow forced by the pressure gradient generated by wind stress at earlier times. The eastward components of the observed currents at SB3 were much stronger than the northward components, in consistence with the nearly east-west orientation of the local bottom topography.

The in situ and satellite remote sensing observations presented above provide valuable information of the 3D

Fig. 5 Time series of observed temperatures (*left*) and salinities (*right*) at three different depths at mooring sites SB2 (*upper*) and SB3 (*lower*) in Lunenburg Bay of Nova Scotia from days 155 (June 4) to 170 (July 19) 2006. *Shading areas* represent periods of two storm events



circulation and hydrographic distributions over the ISS. Further understanding of physical processes for circulation in June and July 2006 over the study region, particularly during tropical storm Alberto, will be gained from model results discussed in Sections 4 and 5.

3 The ocean circulation model setup and external forcing

A nested-grid coastal circulation prediction system known as the NCOPS-LB used in this study is the integration of the shelf circulation forecast system known as Dalcoast3 (Thompson et al. 2007; Ohashi et al. 2008, unpublished manuscript) and a two-level nested coastal circulation model for Lunenburg Bay (Zhai et al. 2008a; Sheng et al. 2008). The nested-grid system has five relocatable submodels (Fig. 7), which consist of, in the descending order of the model domain sizes, (1) a storm surge submodel for the Northwest Atlantic Ocean between 72° and 42° W and between 38° and 60° N (submodel L1); (2) a shelf circulation submodel for the GSL-SS and the adjacent slope between 67.5° and 56° W and between 41.75° and 52° N (submodel L2); (3) a limited-area shelf circulation submodel for the ISS between 64.4° and 62.9° W and between 43.8° and 44.8° N (submodel L3); (4) a coastal circulation submodel for the three-bay region of Mahone Bay (MB), LB, and Rose Bay (RB) of Nova Scotia (submodel L4); and (5) a coastal circulation submodel for LB (submodel L5) in the default setup. The horizontal resolutions of the five submodels are (1/12)°, (1/16)°, ~2 km, ~500 m, and ~180 m, respectively.

It should be noted that submodel L1 is 2D and barotropic (in which model temperature and salinity are set to constant in time and space) and simulates storm surges and generations or propagations of barotropic shelf waves excited by meteorological forcing over the ECS. The other four submodels are 3D and baroclinic. Submodels L1 and L2 are the default setup of Dalcoast3 constructed from the σ -level Princeton Ocean Model (POM; Mellor 2004). Submodel L2 uses 30 σ levels in the vertical, with various spacing between adjacent σ levels doubling between the sea surface and the tenth σ level at $\sigma = -0.0909$ and constant spacing of 0.0455 below the tenth σ level. The sub-grid

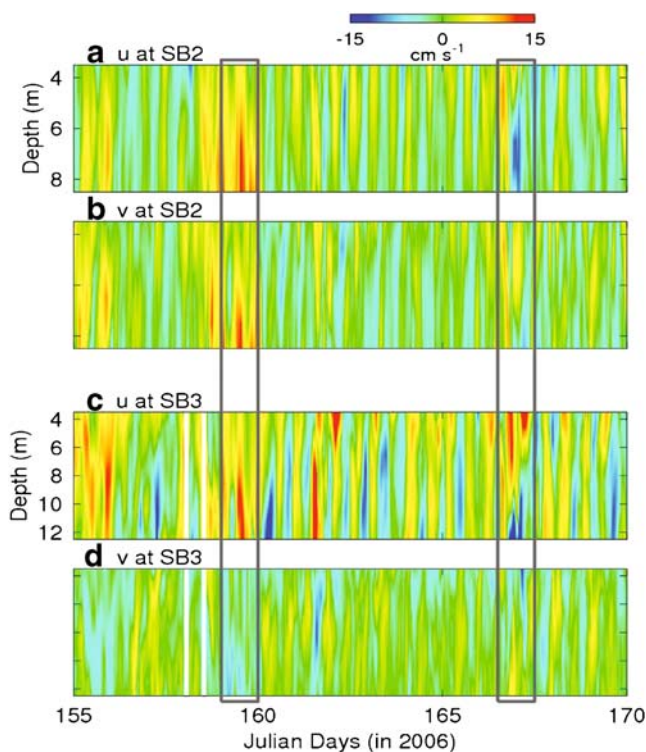


Fig. 6 Time-depth distributions of eastward and northward components of observed currents at SB2 and SB3 in Lunenburg Bay of Nova Scotia from days 155 to 170 in 2006. *Two gray boxes* indicate two wind events in June, 2006

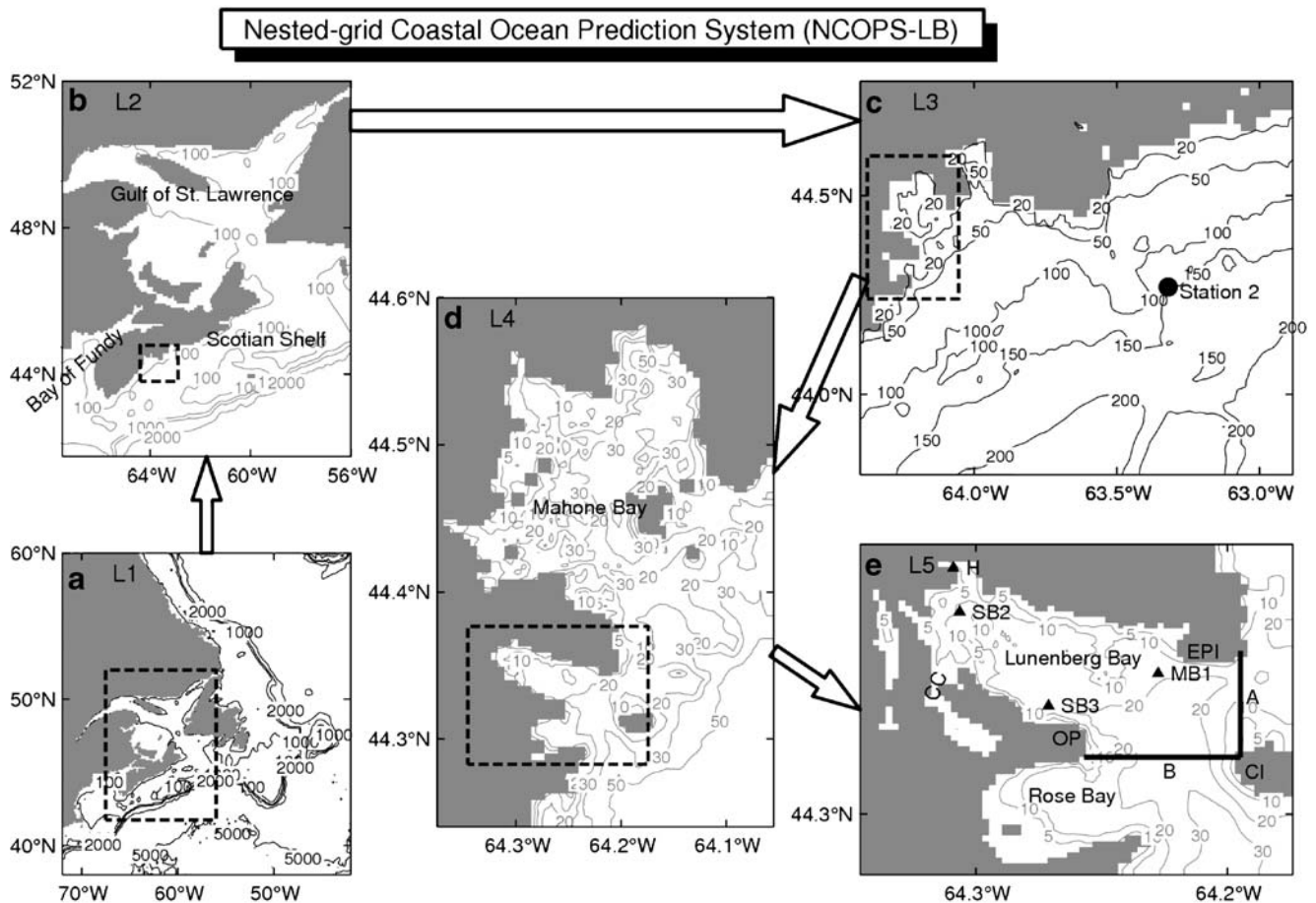


Fig. 7 Major bathymetric features of five submodels of a nested-grid coastal ocean circulation prediction system for Canadian Atlantic coastal waters (NCOPS-LB). **a** Submodel L1 and **b** submodel L2 are Dalcoast3 based on POM developed by Thompson et al. (2007) and **d** submodel L4 and **e** submodel L5 are the nested-grid coastal circulation model based on CANDIE developed by Zhai et al.

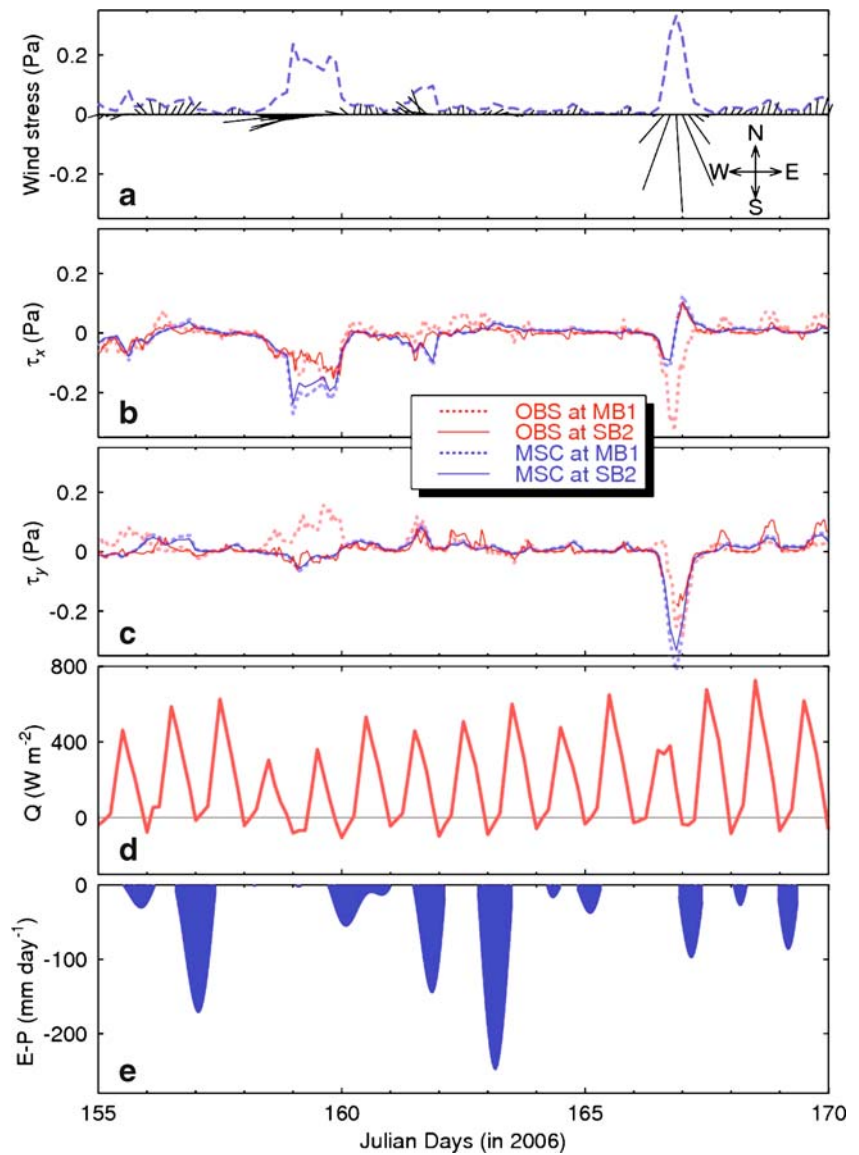
(2008a, b). **c** Submodel L3 is an intermediate resolution circulation model for the inner Scotian Shelf also based on CANDIE. Abbreviations are used for Corkum's Channel (CC), Ovens Point (OP), East Point Island (EPI), and Cross Island (CI). The horizontal heat fluxes through transect A and B are presented in Section 6

scale mixing parameterizations used in Dalcoast3 are the horizontal mixing scheme suggested by Smagorinsky (1963) and the turbulent closure scheme developed by Mellor and Yamada (1982). Submodels 3–5 are constructed from the z -level ocean circulation model known as CANDIE with free surface (Sheng et al. 1998; Zhai et al. 2008a) and use 24 z levels in the vertical with a vertical resolution of 1 m except for 3 m for the top z level and 4.5 m below 20 m depth. In submodels L3–5, the vertical mixing scheme of Large et al. (1994) is used for vertical eddy viscosity and diffusivity coefficients K_m and K_h , and the horizontal mixing scheme of Smagorinsky (1963) is used for the horizontal eddy viscosity coefficient A_m , with the horizontal turbulent Prandtl Number A_{hm} set to 0.1, where A_h is the horizontal eddy diffusivity coefficient.

The nested-grid system is forced by meteorological and astronomic forcing to simulate the 3D circulation and hydrographic distributions over the ECS from days 155 (June 4) to 170 (June 19) 2006. The meteorological forcing includes

SLP and surface wind stress extracted or converted from three hourly numerical weather forecast fields produced by the MSC. The surface wind stress is converted from the MSC wind speeds (Fig. 2) using the bulk formula of Large and Pond (1981). To demonstrate the temporal variability of the meteorological forcing, we examine time series of the three hourly MSC wind stress in LB (Fig. 8). The MSC wind stress was relatively weak (~ 0.05 Pa) and roughly northward during the study period, except for two storm events. The first storm event occurred from days 158 to 160 with a maximum wind stress of ~ 0.2 Pa associated with a low pressure system moving northward over the eastern GOM (not shown). The second event occurred on days 166 and 167 with a maximum wind stress of ~ 0.3 Pa associated with tropical storm Alberto moving northeastward over the central SS (Fig. 2). In comparison of wind stress converted from observed wind at SB2 and MB1 in LB, the MSC wind is in good agreement with the observations, with some noticeable differences during the two storms (Figs. 8b,c).

Fig. 8 Time series of **a** the stick plot of wind stress, **b** the eastward components, and **c** northward components of observed (red) and MSC (blue) wind stress, **d** the net sea surface heat flux at SB2 and **e** the diagnosed net sea surface freshwater flux from days 155 to 170 in 2006. The dashed line in (a) represents the amplitude of wind stress



The net heat flux (Q_{net}) at the sea surface is calculated using the following bulk formulae (Gill 1982):

$$Q_{net} = Q_I + Q_B + Q_L + Q_S \tag{1}$$

where Q_I is the flux of sunlight into the ocean, Q_B is the net upward flux of long-wave radiation from the ocean, Q_L is the latent heat flux carried by evaporated water, and Q_S is the sensible heat flux due to conduction. Each term on the right hand side of Eq. 1 is positive if the flux is from the atmosphere to the ocean. In this study, Q_I is taken from six hourly NCEP fields, and Q_B , Q_L , and Q_S are calculated using empirical formulas based on the NCEP air temperature, relative humidity, MSC wind speed, a constant cloud cover of 0.5, and model-calculated SST. The estimated net heat flux during the study period (Fig. 8d) has significant day-to-day variability with large positive values (heat gain by the ocean waters) of more than 500 W m^{-2} near

noontimes and near zero or small negative values (heat loss by the ocean waters) during nights in June 2006, except for the two storm events (Fig. 8d). During these two storm events, the daily maximum sea surface fluxes near noontimes in LB decrease to $\sim 400 \text{ W m}^{-2}$. Furthermore, to reduce the model bias in simulating temperature and salinity distributions over the three-bay region, the external forcing in submodels L4 and L5 also includes an additional net surface heat (and freshwater) flux diagnosed from model results using the following procedure. We first run the NCOPS-LB with the net sea surface heat flux calculated by Eq. 1. We calculate differences between the modeled sea surface temperature (SST_m) and the observed sea surface temperature (SST_o) at SB3 in LB. The additional net heat flux in submodels L4 and L5 is then estimated based on $\hat{Q} = \frac{\Delta \tau}{\rho_0 C_p} \frac{\partial}{\partial t} (SST_o - SST_m)$. Similarly, the net freshwater flux at the sea surface in submodels L4 and L5 (Fig. 8e) is

estimated based on $E - P = \frac{\Delta z}{SSS_o} \frac{\partial}{\partial t} (SSS_o - SSS_m)$, where SSS_o is the observed sea surface salinity at SB3, and SSS_m is the corresponding model sea surface salinity produced by the NCOPS-LB using all the external forcing except for zero sea surface freshwater flux.

It should be noted that additional correction terms expressed in terms of restoring boundary conditions to seasonal mean hydrographic climatology are added to the net sea surface heat and freshwater fluxes in submodel L2 using the spectral nudging approach in order to reduce the model drift in simulating the seasonal cycle of the sea surface temperature and salinity. Readers are referred to Thompson et al. (2007) and Ohashi et al. (2008, unpublished manuscript) for more discussion on the correction terms and the spectral nudging approach.

The following boundary conditions are used in the NCOPS-LB (Sheng and Yang 2008). At lateral closed boundaries of all the five submodels of the nested-grid prediction system, the normal flow, tangential stress of the currents, and horizontal fluxes of temperature and salinity are set to zero. The Sommerfeld radiation condition is used for the sea surface elevations and depth-mean currents along the open boundaries of submodel L1. Along the open boundaries of submodel L2, the model surface elevation and depth-mean currents are restored to the combination of the sea surface elevations and depth-mean currents produced by submodel L1 and those produced by WebTide. WebTide is a graphic user interface which includes a simple tidal prediction program over the ECS based upon tidal harmonic constants of major tidal constituents precalculated from previous tidal ocean model results (Dupont et al. 2002). Model temperature and salinity along the open boundaries of submodel L2 are advected outward if the normal flow at the open boundaries is outward and restored to the seasonal mean climatology if the model flow is inward (Ohashi et al. 2008, unpublished manuscript). For the other three submodels of the NCOPS-LB based on CANDIE (i.e. submodels L3–L5), the following one-way nesting technique is used for specifying open boundary conditions. The Orlanski (1976) radiation condition is first used to determine whether the open boundary is passive or active. If the open boundary is passive, the model variables (including temperature, salinity, and 3D currents) at the open boundaries are restored to the model results produced by the upper-level submodels, with a restoring time scale of 1.2 h. If the open boundary is active, the model variables at the open boundaries are advected outward. An additional adjustment is made to the depth-mean currents along the open boundaries based on the open boundary condition of Davies and Flather (1978) in order to have proper propagations of (barotropic) tidal forcing and meteorologically forced surface elevations and depth-mean flows from open boundaries of submodel L2 to the lower-level submodels of the NCOPS-LB.

4 Storm-induced circulation during tropical storm Alberto

The NCOPS-LB is initialized from seasonal hydrographic climatology on day 121 (May 1) 2006 and integrated for 3 months using all the external forcing presented in Section 3, which is referred to as the control run (Exp-CR) in this study. The model results in Exp-CR for a 15-day period from days 155 (June 4) to 170 (June 19) are presented in this paper, except where otherwise stated.

Figure 9 shows the simulated sea surface elevations, surface (or near-surface) temperatures and currents produced by the NCOPS-LB at day 166.75 (18:00 June 15) in 2006. The center of the storm is located at the southwestern SS at this time (Fig. 2a). The total sea surface elevations (i.e., the sum of the adjusted sea level and isostatic sea level due to the inverse barometer effect, Gill 1982) at day 166.75 produced by submodel L1 (Fig. 9a) are positive over the ECS (in the barotropic response to wind forcing and sea level atmospheric pressure perturbations), with a maximum value of ~50 cm near the storm center on the southern SS due mainly to the large inverse barometer effect at the center of the storm. Submodel L1 also generates relatively large depth-mean currents over shallow water regions of the ECS and very strong southwestward currents over the coastal region of the western SS at this time.

The surface currents over the GSL-SS region produced by submodel L2 at day 166.75 are characterized by a large-scale cyclonic recirculation over the SS and southwestern GSL and relatively weak surface currents over the eastern SS and the central GSL (Fig. 9b). The simulated SSTs at this time are relatively cool and ~8°C on the SS and the northern and eastern GSL and relatively warm and ~12°C over the southwestern GSL and deep waters of the SS.

The near-surface (1.5 m) currents over coastal waters of the central SS produced by submodel L3 at day 166.75 (Fig. 9c) are relatively uniform and nearly southwestward with a typical speed of ~50 cm s⁻¹. The simulated near-surface temperatures are about 8°C over the ISS except for relatively warmer and about 12°C over bays and inlets.

Over the three-bay region (i.e., MB, LB, and RB), the near-surface currents produced by submodel L4 (Fig. 9d) are nearly southwestward, which are relatively stronger over outer MB and outer LB, and adjacent offshore regions and relatively weaker over inner LB and inner MB. The simulated near-surface temperatures produced by submodel L4 are about 12°C over inner MB and inner LB and relatively cooler and about 8°C over the offshore region of the three bays.

The near-surface currents in LB produced by submodel L5 at day 166.75 (Fig. 9e) are characterized by an intense southwestward throughflow of ~50 cm s⁻¹ in the offshore

Nested-grid Coastal Ocean Prediction System (NCOPS-LB)

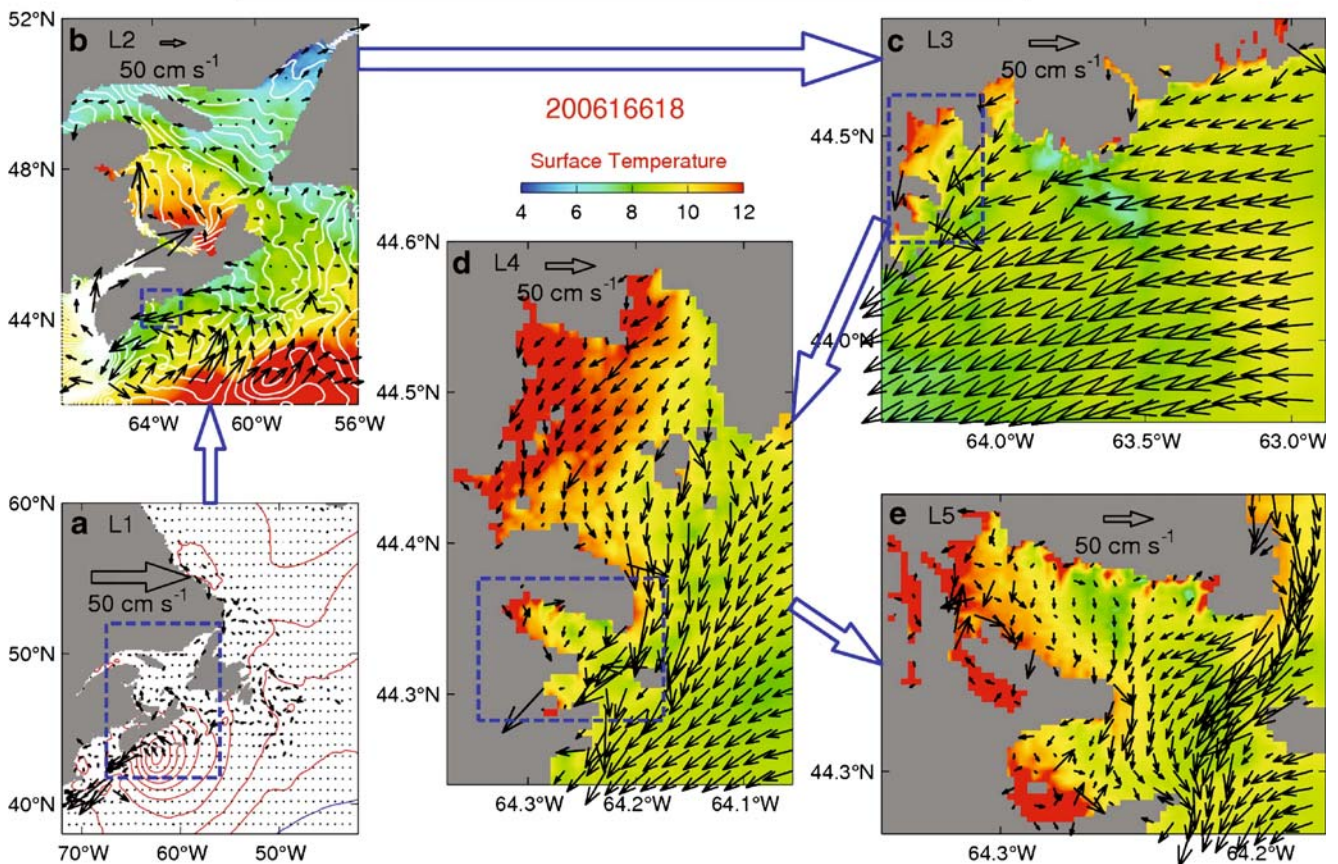


Fig. 9 **a** Depth-mean currents (*black arrows*) and total surface elevations (*red or blue contours* for positive or negative values with the contour interval of 5 cm) produced by submodel L1; **b** surface currents (*black arrows*) and sea surface temperatures (images) produced by submodel L2; and **c–e** near-surface (1.5 m) currents

and temperature produced by submodels L3–L5 of the NCOPS-LB in the control run at day 166.75 (18:00 June 15) in 2006. The center of the storm at this time is located over the central Scotian Shelf. For clarity, velocity vectors are plotted at every **a** ninth, **b** tenth, **c** sixth, **d** third, and **e** third model grid point

deep waters off LB and relatively weak and southwestward currents inside LB. There is a strong northeastward jet through Corkum’s Channel due mainly to the tidal circulation through the channel. The simulated near-surface (1.5 m) temperatures produced by submodel L5 at this time are relatively warmer and about 12°C over western and northern LB and relatively cooler in the deep waters region off LB (Fig. 9e). There are several small-size pools of cool surface waters over northeastern LB due to localized wind-driven upwelling.

At day 167.5 (12:00 June 16, 2006), the center of the storm moves onto the NF. The simulated total surface elevations at this time in the barotropic response to wind and atmospheric pressure perturbations produced by submodel L1 (Fig. 10a) have a maximum value of about 30 cm over the inshore region of the NS; about 10 to 20 cm over the GB and the eastern SS and the southern GSL and relatively small over other regions of the ECS. There are relatively large depth-mean currents at this time over the

GB, the southern GSL and the ISS, associated mainly with the propagation of the barotropic shelf waves excited by Alberto.

Submodel L2 generates strong and southward surface flows over the western GB, the southern GSL and the eastern SS, and strong southwestern tidal currents in the Bay of Fundy at day 167.5 (Fig. 10b). The simulated SSTs produced by submodel L2 at this time are relatively cool and ~7°C over the eastern SS and the northern and eastern GSL and relatively warmer and ~11°C over the southwestern GSL and deep waters off the SS.

The simulated near-surface (1.5 m) currents produced by submodel L3 at day 167.5 (Fig. 10c) are relatively weaker over the ISS with more small-scale features except for some large currents over isolated coastal areas, in comparison with those on the previous day (Fig. 9c). The simulated near-surface temperatures are relatively warmer and about 9°C over the eastern part and about 7°C to 8°C over the western part of the ISS (Fig. 10c). Over the three-bay

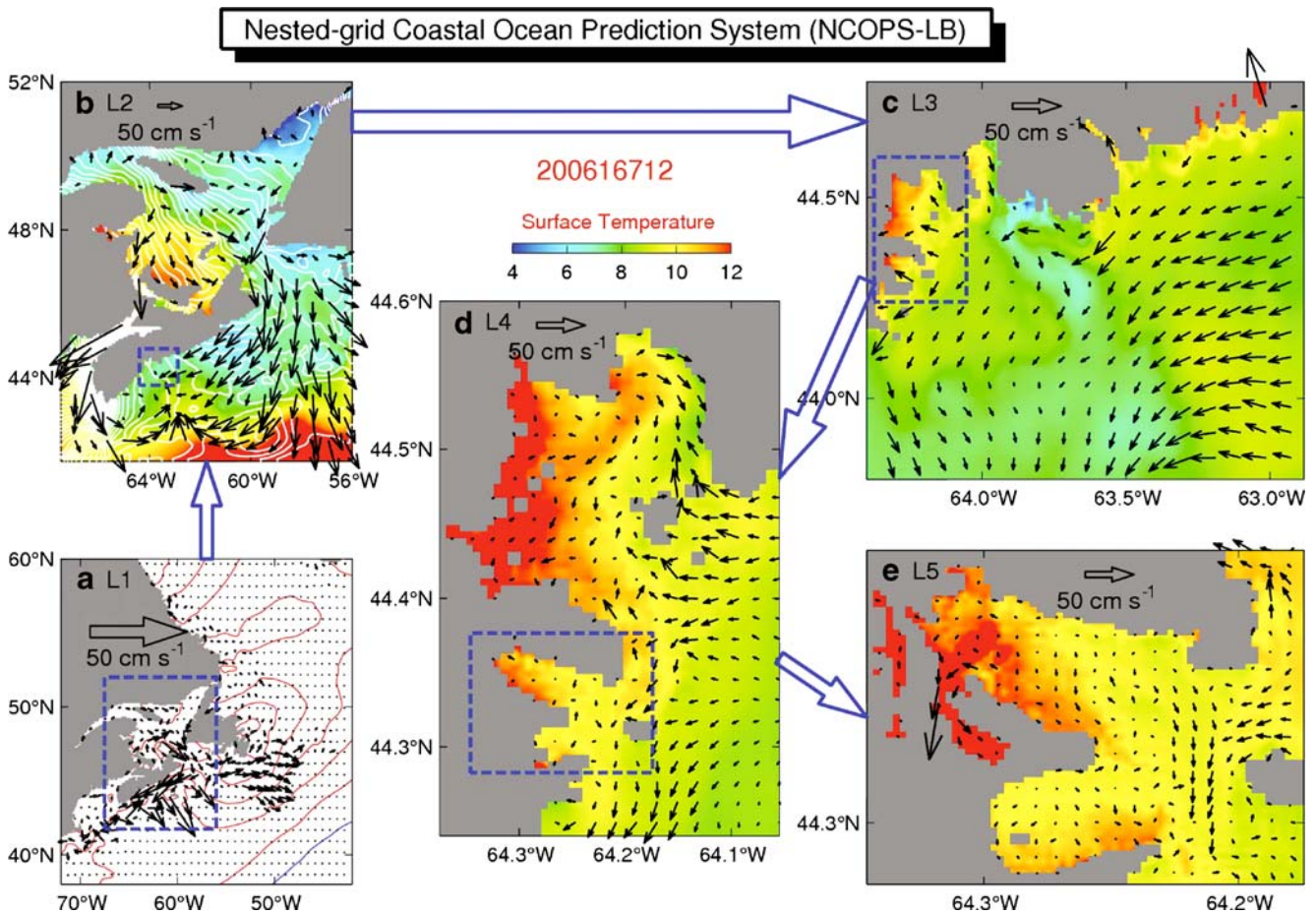


Fig. 10 **a** Depth-mean currents (*black arrows*) and total surface elevations (*red or blue contours* for positive or negative values with the contour interval of 5 cm) produced by submodel L1; **b** surface currents (*black arrows*) and sea surface temperatures (*images*) produced by submodel L2; and (**c–e**) near-surface (1.5 m) currents

and temperature produced by submodels L3–L5 of the NCOPS-LB in control run at day 167.5 (12:00 June 16) in 2006. The center of the storm at this time was located over the inshore region of the Newfoundland Shelf. For clarity, velocity vectors are plotted at every **a** ninth, **b** tenth, **c** sixth, **d** third, and **e** third model grid point

region (Fig. 10d), the simulated near-surface currents are roughly northwestward over outer MB and relatively weak over inner MB and inside LB and RB. The simulated near-surface temperatures are about 10°C in the three-bay region at this time.

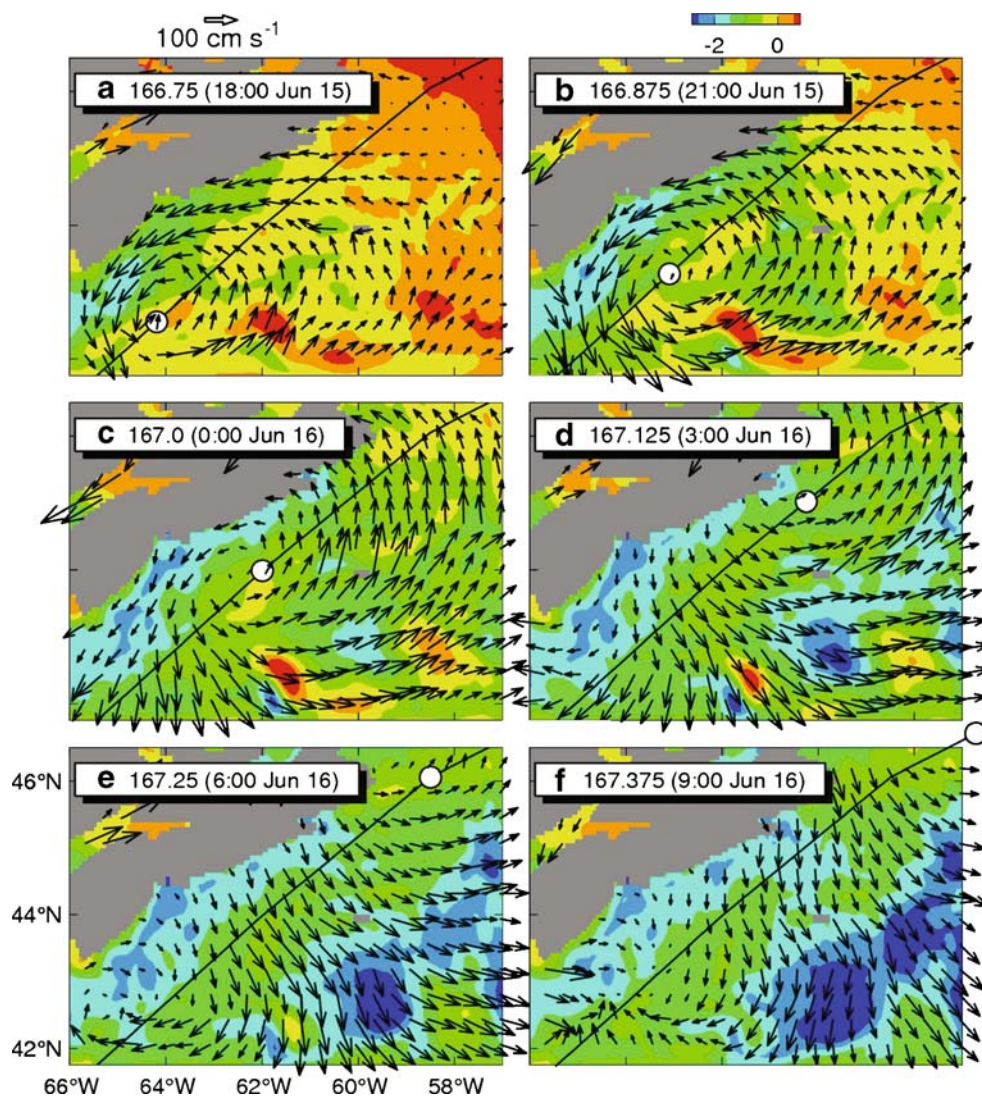
The near-surface currents produced by submodel L5 at day 167.5 (12:00 June 16) are characterized by relatively weak currents over central and inner LB and RB and strong southwestern jet through Corkum's Channel and northwestward jet through the narrow mouth between Upper and Lower South Coves due mainly to the tides (Fig. 10e). The simulated near-surface temperatures are relatively warmer over Rose Bay and western and northern LB and relatively cooler over the southeastern LB and deep waters the south of Cross Island (CI).

To examine the dynamic response of upper waters over the ISS to tropical storm Alberto, we calculate the SST differences (ΔSST) between the model SST during Alberto and the one without the storm (chosen as the daily mean

SST on day 165 as the first order of approximation) to represent the storm-induced ocean response produced by the model. There are several interesting features in the model-calculated ΔSST shown in Fig. 11. During the passage of Alberto over the middle Scotian Shelf, the storm generates significant surface water cooling to the left of the storm between the storm track and the south coast of Nova Scotia (Fig. 11a–c), due mainly to the interaction of coastal topography with intense southwestward coastal currents as part of the storm-induced cyclonic jet under Alberto. After passage of the storm, the SST cooling near the coast reaches a maximum of about -2°C (Fig. 11c), with a cool water tongue expanding gradually to the slope water region off the southwestern SS advected by the storm-induced south or southeastward currents (Fig. 11b,c).

There are several localized maxima of SST cooling of about -2.5°C behind the storm over the slope water region off the SS to the right of the storm track (Fig. 11d–f), which differ from the continuous strip of SST cooling near the

Fig. 11 Snapshots of surface currents (*arrows*) and changes of SSTs during tropical storm Alberto produced by submodel L2 at days **a** 166.75, **b** 166.875, **c** 167.0, **d** 167.125, **e** 167.25, and **f** 167.375. For clarity, velocity vectors are plotted at every sixth model grid point



coast during the storm as shown in Fig. 11a–c. The SST cooling over the slope water region to the right of the storm track intensifies and reaches its maximum at about 200 km to the right of the storm track. The rightward bias of SST cooling behind the storm is due mainly to the rightward bias of strong entrainment and mixing associated with the storm-induced inertial surface currents, which is one of important physical processes for the storm-induced circulation (Price 1981; Sheng et al. 2006).

The SST warming over the eastern SS and Laurentian Channel during the passage of the storm shown in Fig. 11a–c is mainly due to horizontal convergence of relatively warm surface waters ahead of the storm. In addition, there are several small-size pools of warm surface waters (with ΔS_{ST} warming of about 0.5°C) on the right side of the track, due mainly to the advection of warm surface water from the vicinity of the Gulf Stream. These warm surface waters are gradually dissipated with time (Fig. 11e).

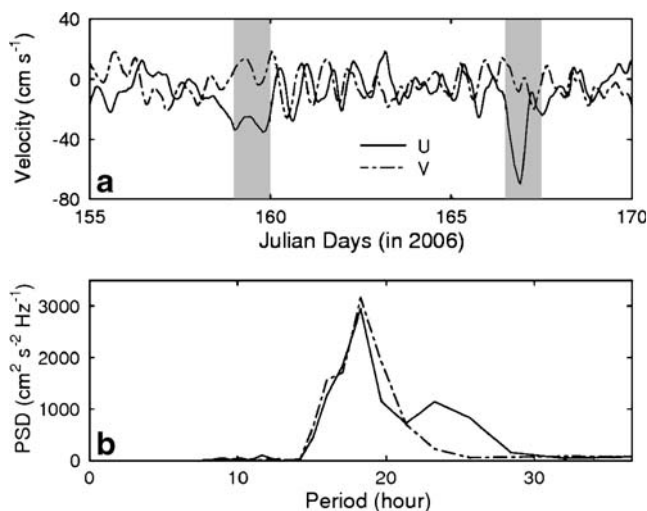


Fig. 12 **a** Time series of model-calculated subtidal horizontal currents at station 2 of Halifax line, **b** spectrum of model-calculated subtidal horizontal currents at station 2. *Shading areas* in **(a)** represent periods of two storm events in June 2006

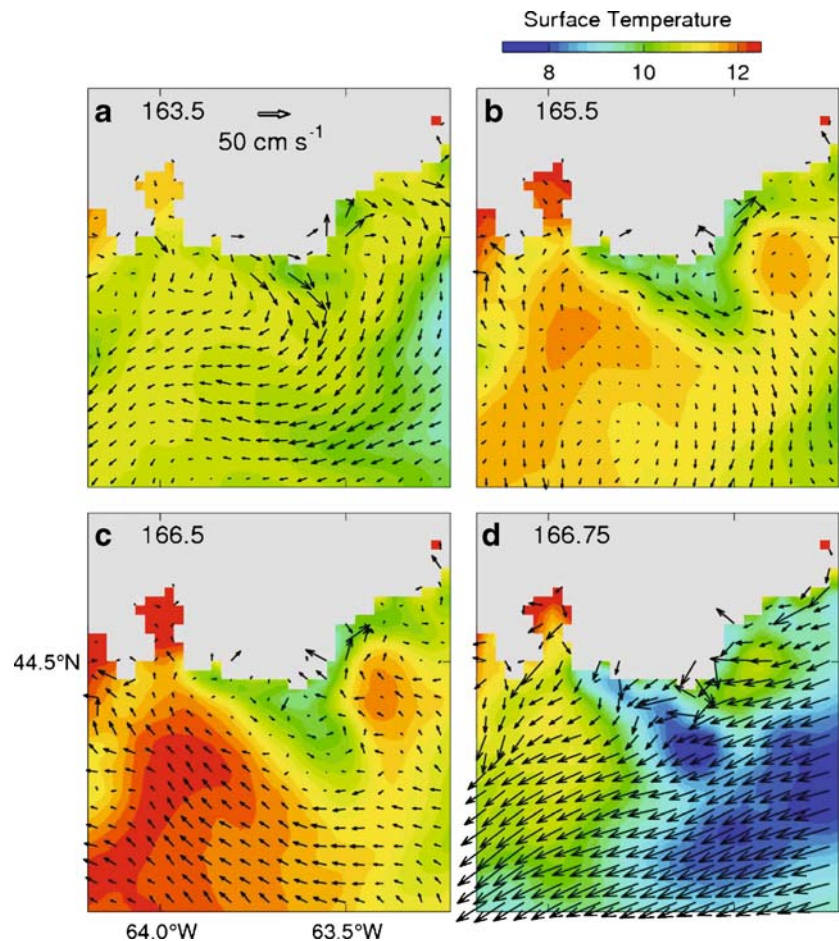
The model-calculated surface currents on the SS shown in Fig. 11 are characterized by intense cyclonic currents under the storm (Fig. 11a–d) and near-inertial currents behind the storm (Fig. 11d–f), which are also consistent with previous findings of Sheng et al. (2006) for the storm-induced circulation over the SS and adjacent deep waters during Hurricane Juan in 2003. At day 166.75, the center of Alberto is located over the Southwestern SS and the storm-induced surface circulation is characterized by a cyclonic gyre over Southwestern SS. At a given point within this gyre, the surface currents rotate clockwise associated with inertial oscillations. The rightward bias of the intense inertial currents behind the storm is also evident, particularly in Fig. 11c,d.

To further examine the storm-induced inertial circulation on the ISS, we conduct the power spectral analysis of nontidal surface currents extracted from model results at Station 2 of Halifax line (Fig. 12a). After the strong wind events around days 159 and 167, the nontidal currents have significant high-frequency oscillations with a typical period of about 18 h (Fig. 12b), which is comparable to, and slightly longer than the periods of inertial oscillations defined as $2\pi/f$ (where f is the Coriolis parameter) to be ~ 17.2 h at this location.

The other important feature of model results over the ISS is the generation of a surface cool water plume extending from the coast near Halifax to offshore waters associated with local upwelling during a 3-day period from day 163 to 166 before Alberto (Fig. 13). The simulated SST at day 163.5 (12:00 June 12) produced by submodel L3 (Fig. 13a) demonstrates the occurrence of coastal upwelling near Halifax. This upwelling event is mainly associated with the outward (divergent) surface currents forced by local wind forcing. The predominant wind direction at Halifax is nearly southeastward with a typical speed of 5 m s^{-1} on that day. The cool water plume continues to expand and forms a cool water tongue reaching the isobaths of 100 m within 3 days (Fig. 13b,c). There is a southeastward baroclinic jet-like flow along the southwestern edge of the cool water plume. This cool water plume disappears by day 166.75 due mainly to strong currents and mixing induced by Alberto (Fig. 13d).

The time-depth distribution of water temperatures at SB2 and SB3 produced by submodel L5 in the Exp-CR (Fig. 14) has the similar temporal and vertical variability to the observed hydrography shown in Fig. 5 but with a much higher resolution in the vertical. The water column at the two sites is weakly stratified from day 155 to the first storm

Fig. 13 Snapshots of near-surface (1.5 m) temperatures and currents produced by submodel L3 at days **a** 163.5, **b** 165.5, **c** 166.5, and **d** 166.75. For clarity, velocity vectors are plotted at every second model grid point



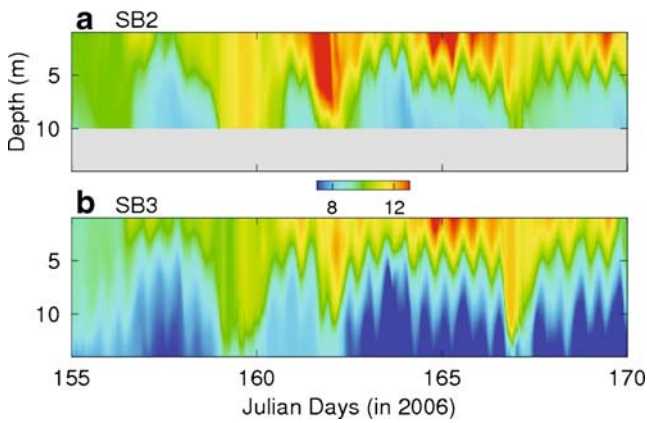


Fig. 14 Time-depth distributions of temperatures at SB2 and SB3 produced by submodel L5 in the control run

event around day 159. The vertical stratification is diminished during three wind events respectively around days 159, 162, and 167, with vertical displacements of isotherms associated with storm-induced upwelling or downwelling at the two sites. The water column is restratified gradually about 2 days after each storm event. Figure 14 also demonstrates that the model temperatures (and salinities) at the two sites in LB also have significant temporal variability at the M_2 tidal frequency.

5 Comparison of model results with observations

The simulated surface elevations and currents produced by submodel L5 of the NCOPS-LB in the Exp-CR are compared with the observations made in LB during the study period to assess the performance of the nested-grid prediction system. The ocean observing system in LB was established in summer 2002 and had been operational from spring to fall during the last 7 years. The core of the observing system consists of three solar-powered buoy nodes in central LB, of which the locations are marked by SB2, SB3, and MB1 in Fig. 7e. Each buoy node communicated with the shore station at Battery Point via a wireless Ethernet link. The data were then transferred to Dalhousie University in Halifax over the internet in near-real time. Each buoy node had a suite of oceanographic and meteorological sensors. The sensors on each buoy node relevant to this study included a pressure gauge and an upward-looking 1.5-MHz Sontek Acoustic Doppler profiler, both deployed on a bottom-mounted pod connected to the buoy by a power and communications cable. An anemometer was mounted on each buoy at 3-m height. In addition to the buoy nodes, a bottom-mounted pressure–temperature sensor was also deployed at site H inside Lunenburg Harbor during the study period. Readers are referred to Wang et al. (2007) for more detailed descriptions of the ocean observing system in LB.

We quantify the hindcast skill of the nested-grid system in simulating the observed temporal variability in terms of the γ^2 value defined as the variance of model errors normalized by the observed variance (Thompson and Sheng 1997). The smaller γ^2 is, the better the agreement between the model and observations. In this paper, $\gamma_c^2 = 1$ is used as a threshold for assessing the performance of the prediction system.

5.1 Sea surface elevations

Figure 15 presents time series of observed and simulated (adjusted) sea surface elevations at three sites MB1, SB2, and H in LB during a 15-day period from days 155 to 170 in 2006. The observed sea surface elevations in LB are dominated by the semidiurnal tide M_2 , which is consistent with previous studies (Thompson et al. 1998; Sheng and Wang 2004; Wang et al. 2007). As mentioned earlier, the coastal sea levels and circulations are also strongly affected by the meteorological forcing during the extreme weather conditions. The γ^2 values between the observed and simulated (adjusted) sea surface elevations at these three sites are very small and about 0.06, indicating that the NCOPS-LB performs very well in reconstructing the observed surface elevations in LB.

5.2 Currents

Figure 16 presents the observed and simulated currents at three depths of 4.5, 6.5, and 8.5 m at two buoy nodes (SB2 and SB3) in LB during the 15-day study period. The observed currents are extracted from the ADCP measurements, and the simulated currents are produced by submodel L5. The observed currents in LB have significant

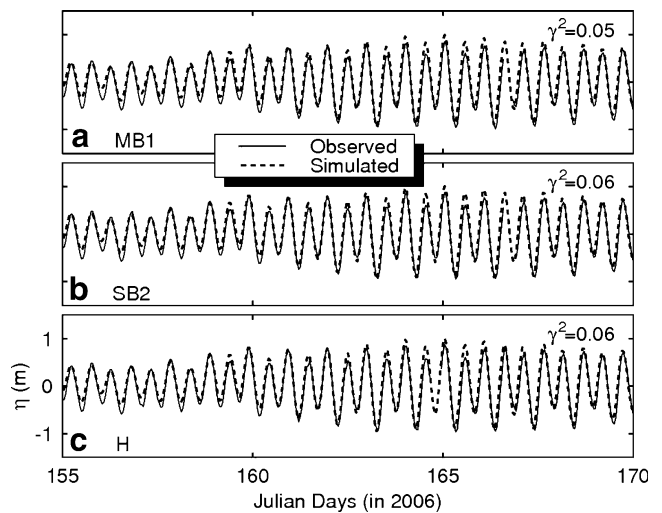
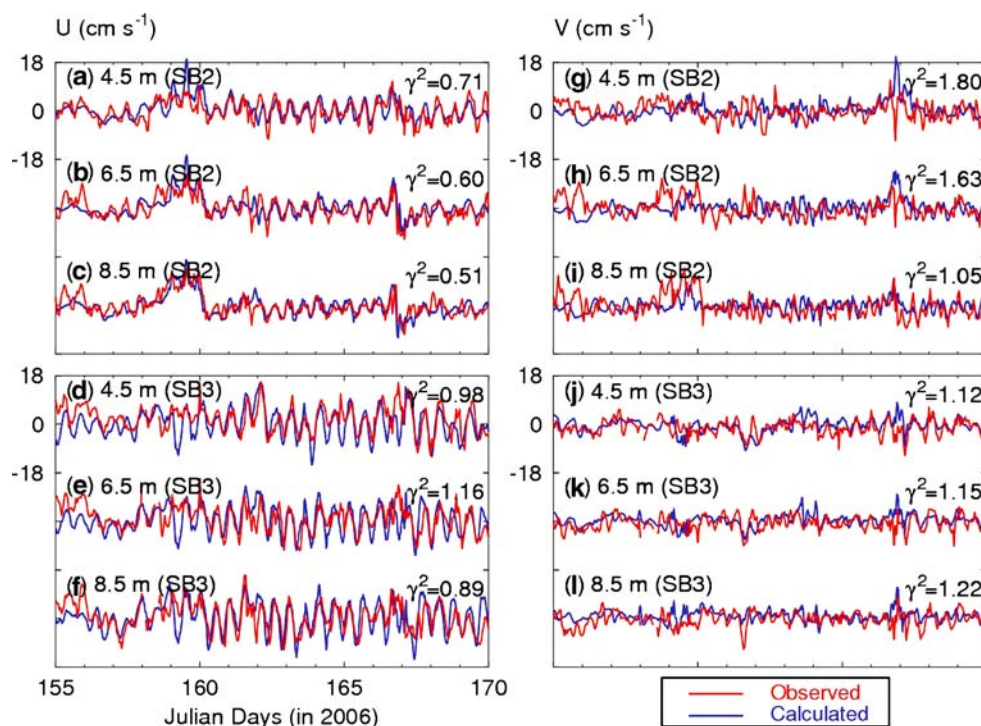


Fig. 15 Comparison of observed (solid) and simulated (dashed) adjusted sea surface elevations at sites **a** MB1, **b** SB2, and **c** H in Lunenburg Bay of Nova Scotia. The simulated results are produced by submodel L5 in the control run

Fig. 16 Time series of (left) eastward and (right) northward components of observed (red line) and simulated (blue line) currents at depths of 4.5, 6.5, and 8.5 m for site (top) SB2, site (middle) SB3 and site (bottom) MB1 in Lunenburg Bay of Nova Scotia. The simulated results are produced by submodel L5 in the control run



temporal and spatial variability with relatively stronger flows at SB3 and relatively weaker flows at SB2. Submodel L5 reproduces moderately well the temporal variability and vertical structures of the eastward components of the observed currents at SB2, with the γ^2 values between 0.5 and 0.7 (Fig. 16a–c). In comparison, submodel L5 performs less well in simulating the northward components of the observed currents at the three depths at site SB2 particularly during the two storms, with the γ^2 values ranging between 1.0 and 1.8 (Fig. 16g–i). Although exact reasons are not known for the model deficiency in simulating the northward components of the currents at SB2, the currents at this site are affected significantly by an intense tidal jet running northeastward from Corkum’s Channel to inner LB (Sheng and Wang 2004), which is highly nonlinear and could not easily be simulated. The observed currents at SB3 have stronger tidal currents in the eastward direction (Fig. 16d–f) than in the northward direction (Fig. 16j–l). Figure 16 demonstrates that submodel L5 simulates the eastward components (with γ^2 values to be 0.9–1.2) slightly better than the northward components (with γ^2 values to be 1.1–1.2) of the observed currents at site SB3.

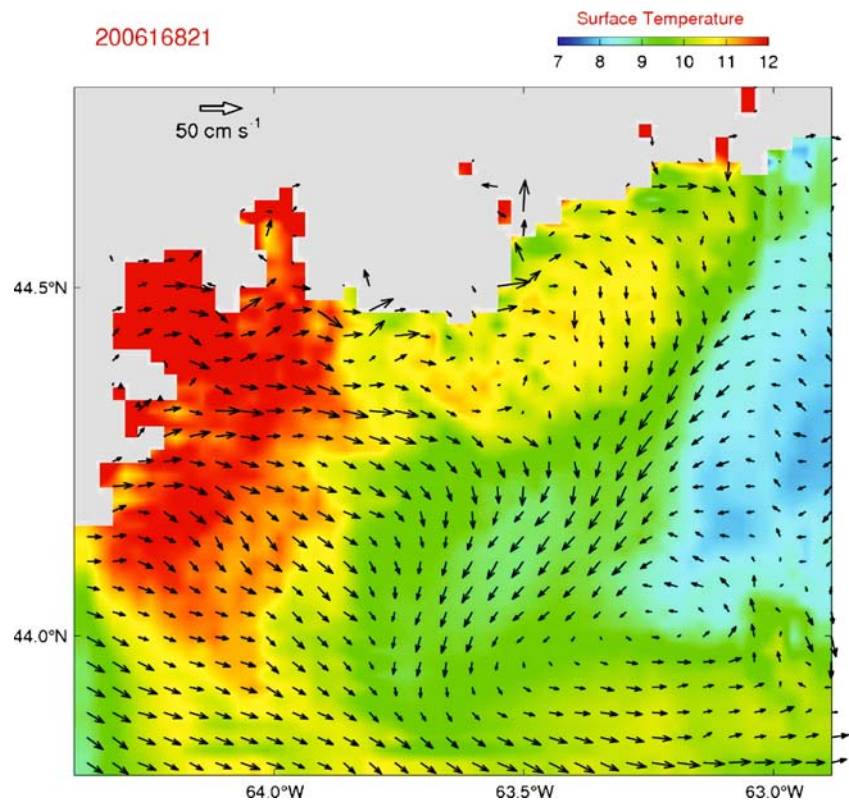
5.3 Temperature

Figure 17 shows the simulated SST over the ISS produced by submodel L3 in the control run at day 168.875 (18:00 June 17), which is 1 day after the Alberto’s passage. Compared to the SST measured by the MODIS (Chris Jones, personal comm., 2007) in Fig. 3, submodel L3 captures the general features of

the observed SST over the ISS characterized by relatively warm surface waters near the coast and offshore, separated by a narrow band of relatively cool SST and a relatively warm water pool over the three-bay region and surrounding area. The cool surface waters over the ISS are mainly due to the advection from the eastern SS. In comparison with the satellite measurements (Fig. 3), the simulated cold water band is overdiffused and the offshore SST is underestimated due mainly to the model deficiency associated with a relatively coarse horizontal resolution of ~ 2 km and less realistic sub-grid scale mixing parameters used in submodel L3. It should be noted that the MODIS measures SST from a thin “skin” surface layer, while the model temperatures shown in Fig. 17 are at 1.5 m, which should be lower than the satellite measured SSTs as expected.

We next compare the model temperatures and salinities with hydrographic observations at mooring sites SB2 and SB3 in LB during the study period. Figure 18 demonstrates that submodel L5 of the NCOPS-LB reproduces moderately well the general temporal and vertical variations of water temperatures and salinities in LB. Submodel L5 also captures approximately the observed upwelling or downwelling events with periods of 1–2 days during large wind events around days 159 and 167, respectively. However, the simulated subsurface temperature on days 161 and 162 is about 4°C and 2°C higher than the observations at SB2 (Fig. 18b,c) and SB3 (Fig. 18e), respectively, indicating the model overestimates the downwelling during this period, particularly at SB2. As mentioned earlier, the MSC wind

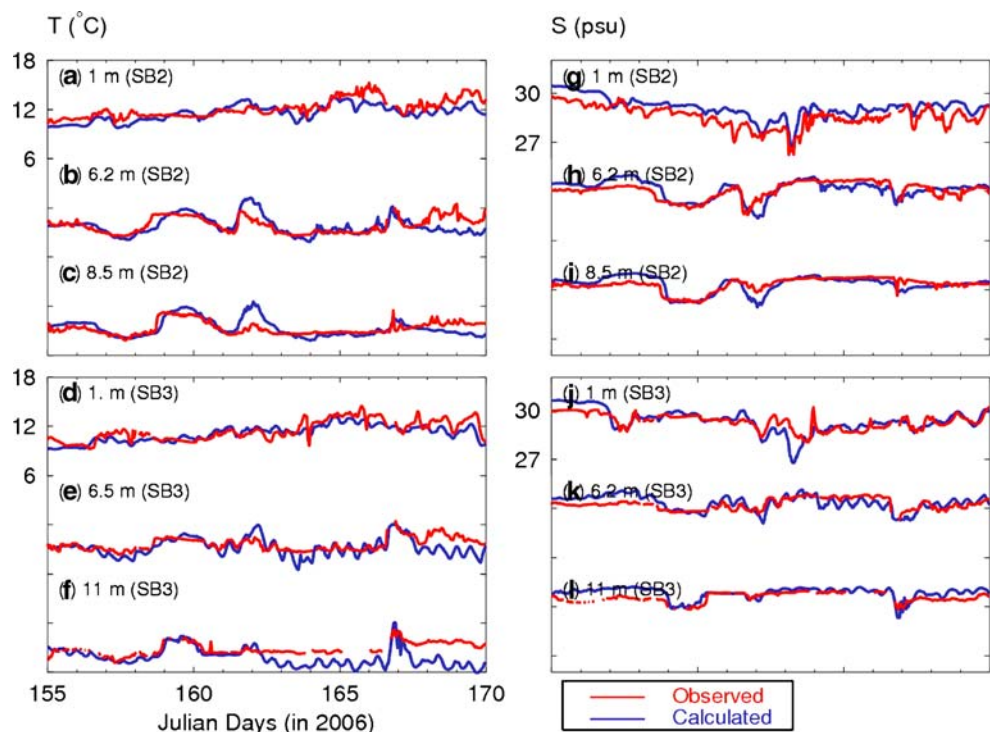
Fig. 17 The near-surface (1.5 m) temperature and currents produced by submodel L3 in the control run at day 168.875 (21:00 June 17) in 2006. For clarity, velocity vectors are plotted at every second model grid point



stress overestimates the observed westward wind stress in the bay on day 161 (Fig. 8b), suggesting that the bias in the MSC wind forcing used in the simulation could be a factor for the poor model performance in reproducing subsurface temperature during this 2-day period. A simple model

sensitivity study demonstrates that model results using the observed wind stress at site MB1 in submodel L5 (not shown) are in slightly better agreement with the observed subsurface temperature during the period of days 161–162 than those in the control run (it should be noted that the

Fig. 18 Comparison of observed and simulated temperature and salinity at two depths at mooring sites SB2 and SB3 in Lunenburg Bay. The simulated results are produced by submodel L5 in the control run



model currents with the use of observed wind stress at MBI are not always better than those with the use of MSC wind stress during the whole study period). After the Alberto's passage, there are cold biases of about 2°C at SB2 and about 3°C in subsurface layer at SB3, due mostly likely to unrealistically strong advection from the outside of LB (Zhai et al. 2008a). In the next section, the role of the heat exchange between LB and offshore waters associated with the throughflow will be examined.

There are also salinity biases of about 0.5 psu at 1 m at site SB2 during the whole study period (Fig. 18g). One plausible explanation is that the low surface salinity observed at SB2 during the study period could be partially influenced by the local freshwater runoff, which is not considered in this study (due to the lack of observations). During the period from days 161 to 162, the simulated subsurface salinity at SB2 is about ~0.5 psu lower than the observations, due mainly to overestimated downwelling in submodel L5 during this period.

6 Processes influencing coastal circulations over the inner Scotian Shelf

Two additional numerical experiments are conducted to investigate the role of local wind forcing and the remotely generated waves in affecting the coastal circulation and hydrographic distributions in LB and adjacent areas during the study period. The nested-grid system is forced by all the external forcing described in Section 2 except for zero wind stress in submodel L5 in the first additional experiment (Exp-RGW), and by all the external forcing except for the zero wind forcing in submodel L1-L4 in the second additional experiment (Exp-LWF). All other model parameters in the two experiments are the same as in the Exp-CR. Dynamically, it means that the circulation produced by submodel L5 in Exp-RGW is not affected by the local wind but affected by the coastal waves generated remotely by the wind forcing outside LB. In Exp-LWF, the circulation produced by submodel L5 is affected by the local wind

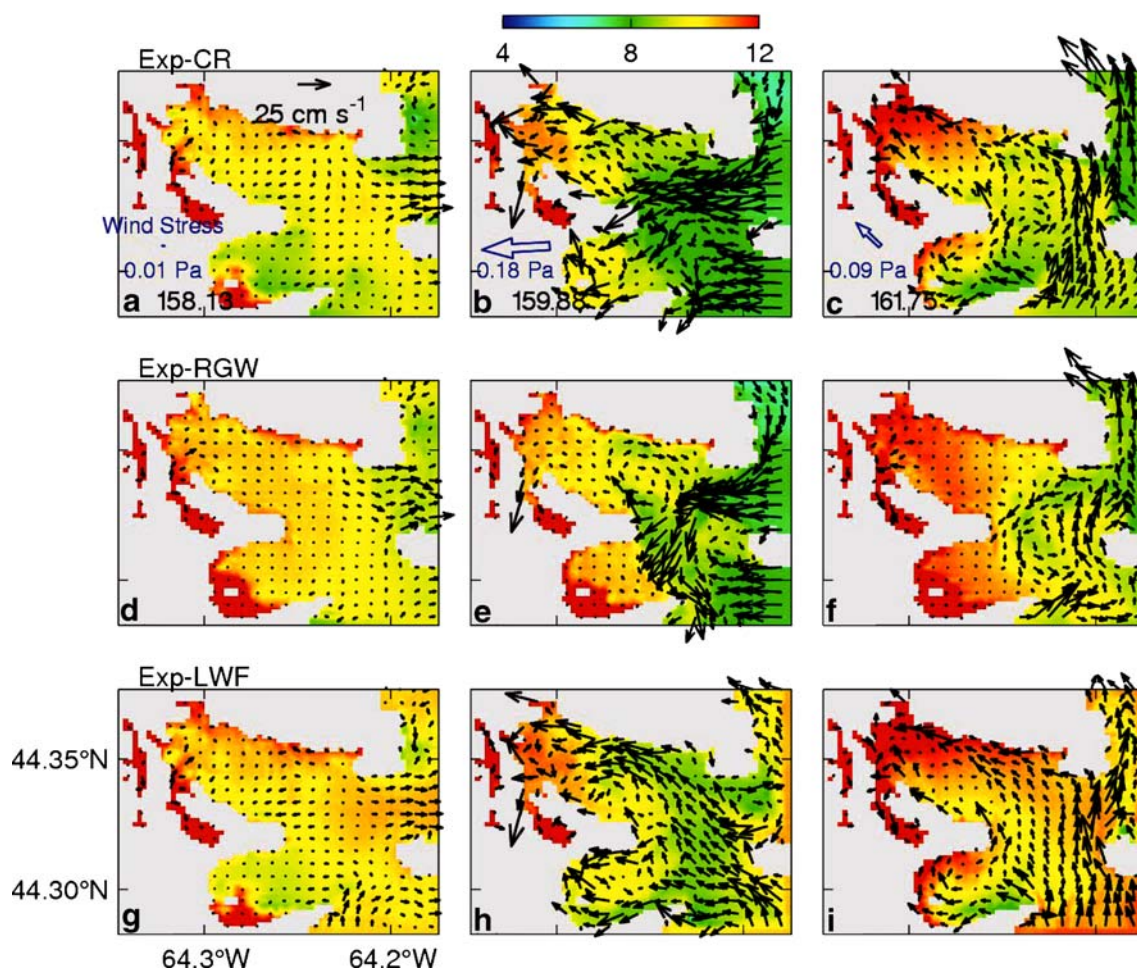


Fig. 19 Near-surface (1.5 m) currents and temperatures at days **a** 158.125, **b** 159.875, and **c** 161.75 produced by submodel L5 in the Exp-CR; **d–f** in experiment Exp-RGW forced by all the external forcing except for the local wind in submodel level 5; **g–i** in

experiment Exp-LWF forced by all the external forcing except for the wind forcing in submodels L1–L4. *Open arrows* denote wind stress at SB2 in control run and Exp-LWF. Velocity vectors are plotted at every third model grid point

forcing but not by the remotely generated coastal waves since the wind forcing in submodels L1–L4 is set to zero. The remotely generated coastal waves discussed here could be either barotropic or baroclinic. As demonstrated recently by Zhai et al. (2008b), baroclinic Kelvin waves can be excited by wind forcing in neighboring MB and then propagate into LB.

Figures 19 and 20 present the near-surface (1.5 m) and subsurface (8.5 m) currents, near-surface temperatures, and differences between the near-surface and subsurface temperatures at three different times produced by submodel L5 of the NCOPS-LB in three experiments: Exp-CR, Exp-RGW, and Exp-LWF, respectively. At day 158.125 (03:00 June 7), the local wind stress in LB is very weak and ~ 0.01 Pa. The model-calculated near-surface and subsurface currents in the three experiments at this time are highly comparable and characterized by relatively weak currents in LB and a northeastward tidal jet over Corkum’s Channel and relatively strong and eastward currents over the channel between East Point Island

(EPI) and CI (Figs. 19a,d,g and 20a,d,g). Some noticeable differences occur in the model results at day 158.125 between the three experiments. In comparison with the model results in the control run, the near-surface currents at day 158.125 are relatively weaker inside LB in Exp-RGW due to the lack of local wind forcing (Fig. 19a,d) and slightly stronger with relatively warmer near-surface temperature over the open waters near the southern open boundary in Exp-LWF due mainly to the lack of the influence of remotely generated coastal waves (Fig. 19a,g).

By day 159.875 (21:00 June 8), the wind stress has changed to westward and increased to ~ 0.18 Pa. The model results in Exp-CR are characterized by a strong and nearly southwestward throughflow to the south of EPI and a strong northwestward flow along the northeastern coast of LB in the near-surface layer (Fig. 19b) and a cyclonic gyre in the deep waters in the subsurface layer (Fig. 20b). There are some similarity in the model calculated near-surface and subsurface currents and temperatures at this time between

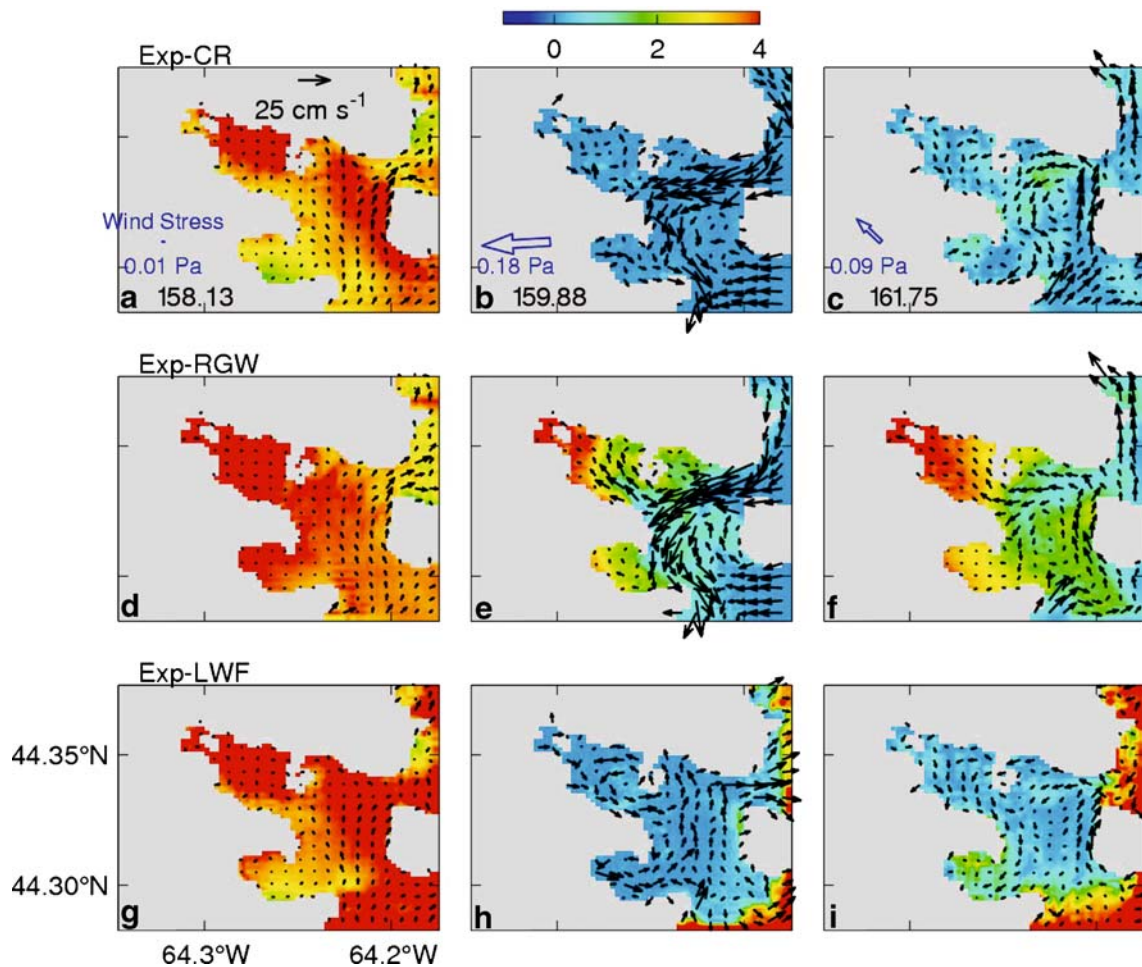


Fig. 20 Sub-surface (8.5 m) currents and temperature differences between those at 1.5 and 8.5 m at days **a** 158.125, **b** 159.875, and **c** 161.75 produced by submodel L5 in the Exp-CR; **d–f** in experiment Exp-RGW forced by all the external forcing except for the local wind

in submodel L5; **g–i** in experiment Exp-LWF forced by all the external forcing except for the wind forcing in submodels L1–L4. *Open arrows* denote wind stress at SB2 in control run and Exp-LWF. Velocity vectors are plotted at every third model grid point

Exp-CR and Exp-RGW over the deep waters off LB and between Exp-CR and Exp-LWF inside LB and RB, indicating that the throughflow in deep waters off LB in the control run is strongly affected by combination of remotely generated waves and local wind forcing, and the circulation over inner LB in Exp-CR are mainly affected by the local wind forcing.

The wind stress has changed to northwestward and decreased to ~ 0.1 Pa by day 161.75 (18:00 June 10). At this time, the throughflow in the deep waters off LB in Exp-CR runs almost northward and separates into two branches before reaching EPI, with one branch veering northwestward and entering the inner LB along the northeast coast of LB and the other branch turning around EPI and flowing northward along the eastern side of EPI (Figs. 19c and 20c). The model near-surface currents and temperature in Exp-RGW are similar to those in the control run over the deep waters off LB but significantly different from the control run inside LB (Fig. 19c,f). By comparison, the near-surface currents and temperature in Exp-LWF are similar to the control run inside LB (Fig. 19c, i). The subsurface currents and temperatures in Exp-RGW and Exp-LWF differ significantly from the model results in the control run at this time (Fig. 20c,f,i). Therefore, the combination of the local wind forcing and remotely generated waves affects the 3D circulation and hydrographic distributions in the region at this time.

Figures 21 and 22 presents the time-depth distributions of water temperatures at mooring sites SB2 and SB3 in LB produced by submodel L5 in Exp-RGW and Exp-LWF, respectively. In Exp-RGW, the model temperatures at the two sites are characterized by vertical periodic movements of isotherms at the M_2 tidal frequency and a gradual warm trend throughout the water column due mainly to the horizontal transport of surface heat into LB from neighboring coastal waters (the surface heat flux is small since the wind forcing is set to zero in this experiment) during the

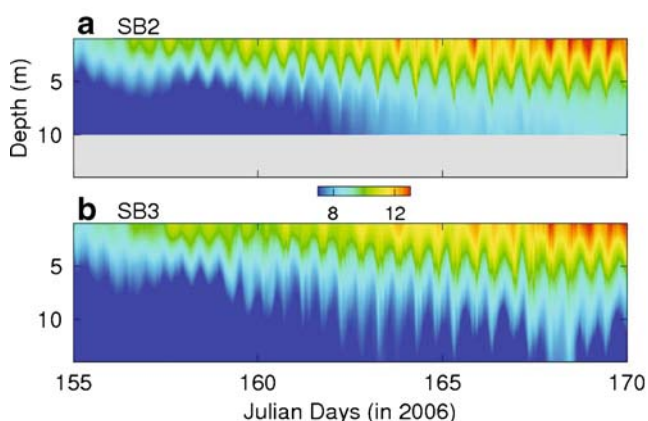


Fig. 21 Time-depth distributions of temperatures at mooring sites SB2 and SB3 in Lunenburg Bay of Nova Scotia produced by submodel L5 in Exp-RGW

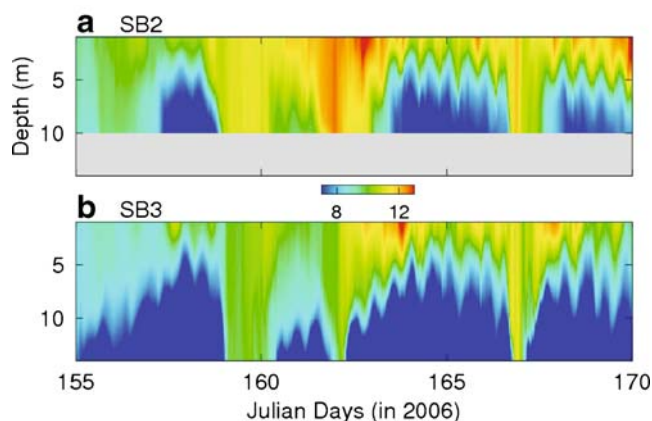


Fig. 22 Time-depth distributions of temperatures at mooring sites SB2 and SB3 in Lunenburg Bay of Nova Scotia produced by submodel L5 in Exp-LWF

study period (Fig. 21). In comparison, model temperatures at the two sites in Exp-LWF have significant low-frequency variability (Fig. 22). The water column in Exp-LWF is almost well mixed during three relatively strong wind events around days 159, 162, and 167 respectively, and restratified about 2 days after each wind event (Fig. 22). Some cool waters appear below 10-m depth at SB3 after passage of Alberto in Exp-LWF due mainly to the advection of offshore waters by the wind-induced throughflow (Fig. 22). A comparison of model results in the three experiments (Fig. 14) indicates that water temperatures in LB are affected by tides, local wind forcing and sea surface heat (and freshwater) fluxes at the sea surface, and heat (and salinity) exchange between the bay and neighboring coastal and shelf waters.

An empirical orthogonal function (EOF) analysis is conducted to determine the dominant temporal and spatial patterns of the nontidal circulation at 3.5 and 8.5 m from model results from days 155 to 170 in three experiments. The nontidal currents are extracted from the total currents produced by submodel L5 using a tidal analysis package known as T_TIDE (Pawlowicz et al. 2002) and a low-pass filter with a cutoff period of 25 h to eliminate any additional tidal components. The near-surface (3.5 m) and subsurface (8.5 m) circulation patterns (or eigenvectors) and time series of mode coefficients of the first two EOFs in the three experiments (Exp-CR, Exp-RGW, and Exp-LWF) are presented in Fig. 23. The first EOF explains about 51%, 53%, and 41% of the total variance of the model nontidal currents at the two depths over the domain of submodel L5 in Exp-CR, Exp-RGW, and Exp-LWF, respectively (Fig. 23a–f). The second EOF explains, respectively, $\sim 17\%$, 23%, and 20% of the total variance in the three experiments respectively (Fig. 23g–i).

The near-surface (1.5 m) eigenvectors of the first EOF (EOF1) in the control run (Exp-CR) are characterized by a

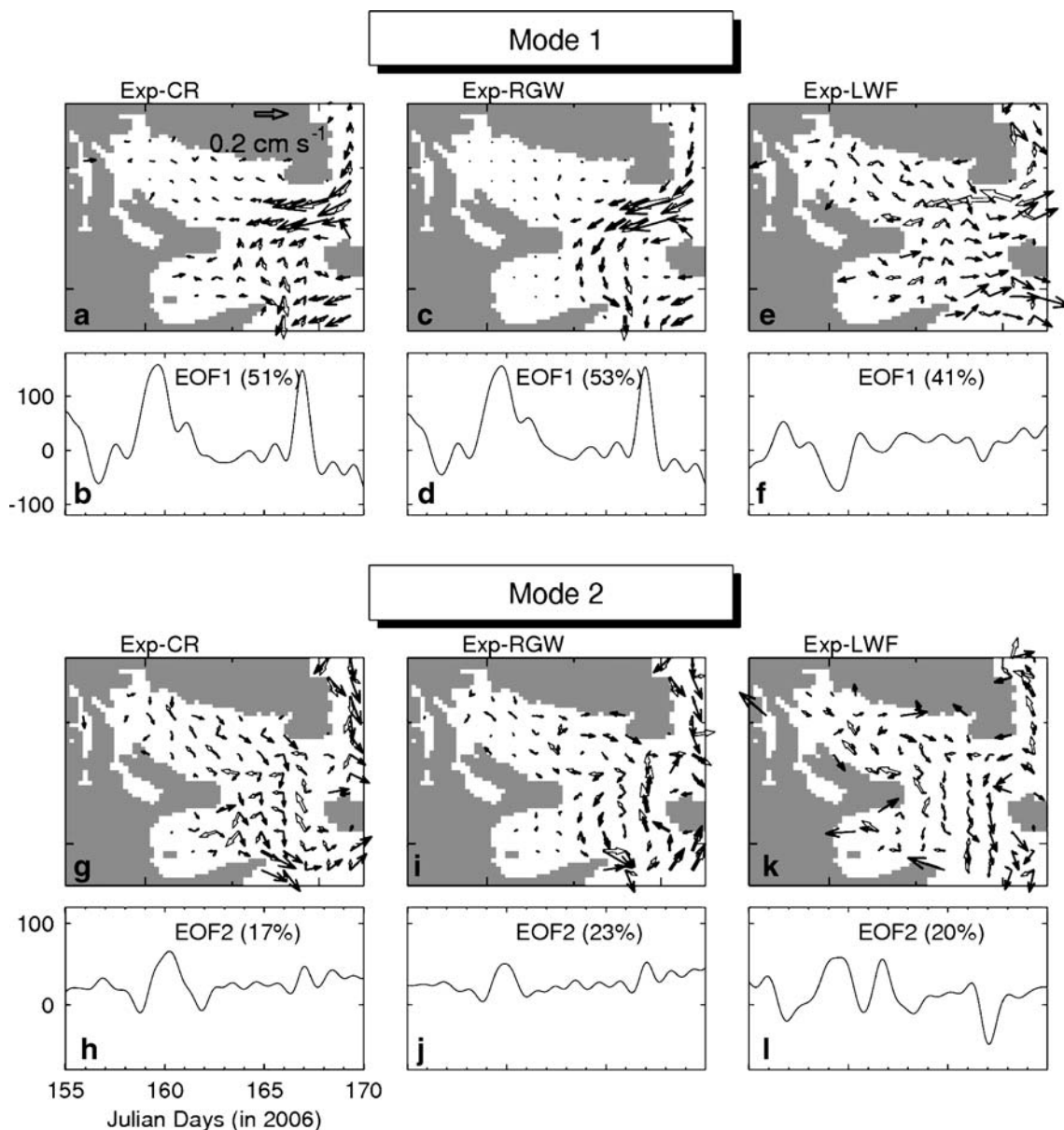


Fig. 23 EOF analysis of nontidal currents produced by submodel L5 (a, b, g, h) in the Exp-CR; c, d, i, j in experiment Exp-RGW forced by all the external forcing except for local wind in submodel L5; e, f, k, l in experiment Exp-LWF forced by all the external forcing except for wind forcing in submodels L1–L4. Eigenvectors of (a, c, e) the first-

mode (*EOF1*) and (g, i, k) the second-mode (*EOF2*) at 3.5 m (*solid arrow*) and 8.5 m (*open arrow*) depths; the time-varying mode coefficients of (b, d, f) *EOF1* and (h, j, l) *EOF2* during the period from days 155 to 170, 2006. For clarity, velocity vectors are plotted at every fifth model grid point

relatively strong throughflow in the deep waters to the southeast of LB around CI and weaker currents inside LB (Fig. 23a). The subsurface (8.5 m) eigenvectors of *EOF1* in Exp-CR have a similar spatial pattern with smaller amplitudes as the near-surface eigenvectors. There are some noticeable differences in direction between the near-surface and subsurface eigenvectors of the first EOF in Exp-CR, particularly inside LB and RB. The mode coefficient of *EOF1* has two large positive peaks around days 159.5 and 167, respectively, which are correlated strongly with wind forcing during the two storm events

(Fig. 8a). A multiregression analysis demonstrates that the mode coefficient of *EOF1* in Exp-CR has maximum correlation ($r=0.87$) with the MSC wind stress at a time lag of about 4 h, with the wind forcing leading.

The near-surface eigenvectors of the second EOF (*EOF2*) in Exp-CR are characterized by relatively strong and broad currents in the deep waters to the east of RB and outer LB; a narrow coastal jet along the northern coast of LB; and relatively weak currents along the southern coast of LB (Fig. 23g). The subsurface eigenvectors of *EOF2* have a very different circulation pattern inside LB with

relatively strong currents along the southern coast of LB and strong currents in the deep waters to the east of RB. The mode coefficients of EOF2 are relatively small except for around day 160 associated with the first storm during which the vertical stratifications are weak. The general pattern of EOF2 represents a two-layer circulation associated with wind-induced coastal upwelling or downwelling inside LB, which can easily be explained in terms of vertical stratifications in the bay.

The eigenvectors and mode coefficients of EOF1 in Exp-RGW are very similar to those in the control run (Fig. 23a–d), with major differences in vertical structures of the circulation patterns. The near-surface and subsurface eigenvectors of EOF1 in Exp-RGW are almost in the same direction, which differs from those in Exp-CR. The mode coefficient of EOF1 in Exp-RGW is also highly correlated with the wind stress with the maximum correlation coefficient of 0.86 at a time lag of about 5 h, with the wind forcing leading, which are very similar to the values in the control run, indicating that the temporal variability of the large-scale circulation in LB and adjacent areas in Exp-CR is affected significantly by coastal waves that are generated remotely by wind forcing outside LB which enter the model domain of L5 through the open waters between LB and CI and between RB and CI.

It is interesting to note that the correlation coefficients for the first EOFs in Exp-LWF with the wind stress are high and about 0.88 and 0.84 for EOF1 and EOF2, respectively, both at zero time lag, indicating that the circulation patterns shown in Fig. 23e,k, are forced directly by local wind. The large-scale spatial distributions of eigenvectors in Exp-LWF (Fig. 23e,k) differ significantly from those in Exp-CR (and Exp-RGW). Nevertheless, there are some similarities between eigenvectors in Exp-LWF and Exp-CR over LB and RB, particularly in terms of vertical structures of the eigenvectors, further indicating the important role of local wind forcing in generating the 3D circulation over the coastal waters of the ISS.

To examine the relative role of the local and nonlocal processes in determining the thermal structure in LB, the heat budget analysis is conducted for waters over LB marked the coastline of LB and two transects A and B (see Fig. 7e for positions of the two transects). The time change of the total heat content in LB can be written as

$$\frac{\partial H}{\partial t} = \Gamma_S + \Gamma_A + \Gamma_B \quad (2)$$

where $H = \iiint TdV$, $\Gamma_S = \iint Q/\rho_0 C_p dS$, $\Gamma_A = -\iint uTdS_A$, and $\Gamma_B = \iint vTdS_B$, which represent respectively the volume-integrated heat content, the surface-integrated net heat flux, and the area-integrated heat transports through transects A and B. The contribution from horizontal mixing at transects A and B to the time-change of the total heat content is assumed to be small and ignored in Eq. 2. We calculate each

term in Eq. 2 using model results in the control run. Figure 24 demonstrates that the time-integrated heat budget in LB is balanced dominantly between the inward heat transport through transect A and the outward heat transport through transect associated mainly with the throughflow in the deep waters off LB.

During the two strong wind events around days 159 and 167, respectively, the heat transports through transect A and B increase significantly due mainly to the strong northwestward throughflow (Fig. 23a,b), which explains the early findings by Zhai et al. (2008a) that the temperature distribution at SB3 is very sensitive to changes in open boundary conditions. The underestimated subsurface temperature in the control run after the passage of Alberto (Fig. 19) is mainly due to the model deficiency in reproducing either the throughflow associated with wind forcing in large-scale and less realistic vertical mixing in the model.

7 Summary and conclusion

A multiple nested-grid coastal ocean circulation prediction system known as the NCOPS-LB was used to examine the main physical processes affecting the three-dimensional circulation and hydrographic distributions over the ISS, with a special emphasis on the storm-induced circulation during tropical storm Alberto in June 2006. The nested-grid system has five downscaling, relocatable submodels with the outmost submodel covering the Eastern Canadian continental shelf from Labrador Shelf to the Gulf of Maine and the innermost submodel covering LB of Nova Scotia in the default setup. The NCOPS-LB is forced by the three hourly atmospheric forcing based on the numerical weather forecasts provided by the Meteorological Service of Canada and astronomical forcing based on WebTide.

The NCOPS-LB was used to simulate the 3D circulation and water mass distributions during tropical storm Alberto

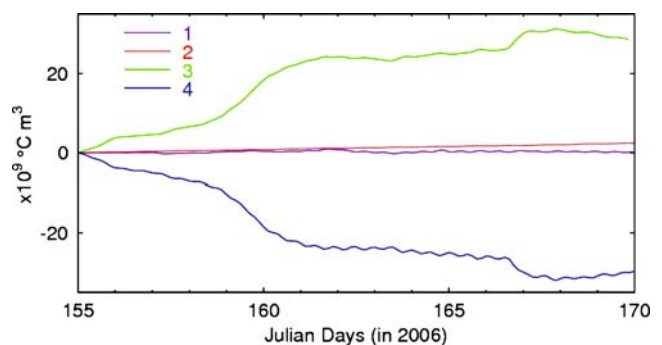


Fig. 24 Time series of 1 the volume-integrated heat content reference to the value at day 155.0 ($H(t)-H(t_0)$), 2 time-integrated surface heat flux contribution ($\int_{t_0}^t \Gamma_S dt$); 3 time-integrated heat transports through transect A ($\int_{t_0}^t \Gamma_A dt$); and 4 time-integrated heat transports through transect B ($\int_{t_0}^t \Gamma_B dt$)

in June 2006 and July in this study. The nested-grid system produces intense cyclonic currents under the storm and strong near-inertial currents behind the storm, with rightward bias of SST cooling behind the storm. Due to the interaction of storm-induced currents and coastal topography, there is also strong SST cooling over the coastal waters of the ISS to the left of the storm track during tropical storm Alberto. Model results suggest that the throughflow over outer LB and in the adjacent deep waters is affected strongly by tides, local wind forcing, and the remotely generated waves (RGWs), and the 3D circulation and upwelling or downwelling inside LB are affected strongly by the local wind and weakly by RGWs.

The model results produced by submodel L5 of the NCOPS-LB, which has the finest horizontal resolution of ~180 m among the five submodels, were compared with the observations made by a multidisciplinary ocean observing system in Lunenburg Bay to assess the performance of the NCOPS-LB. The nested-grid system has reasonable hindcast skills in simulating sea levels and tidal currents during tropical storm Alberto. The NCOPS-LB performs less well in simulating nontidal currents and hydrography, due partially to the model deficiency in sub-grid scale mixing and nonlinear dynamics over coastal waters, less realistic representations of wind forcing, and lack of realistic freshwater fluxes at the sea surface and at the coast. In addition, a simple one-way nesting technique was used in this study. To reduce the model drift between different submodels of the NCOPS-LB, the two-way nesting technique based on the semi-prognostic method will be used (Sheng et al. 2001; Greatbatch et al. 2004; Sheng et al. 2005).

Acknowledgments The authors wish to thank Li Zhai, Kyoko Ohashi, Keith Thompson, Jun Zhao, Liang Wang, Hal Ritchie, Mike Dowd, Hal Ritchie, Serge Desjardins, Doug Schillinger, Alex Hay, John Cullen for their contributions in developing the NCOPS-LB. Comments from two anonymous reviewers were very useful when revising the manuscript. This study was supported by the Canadian Foundation for Climate and Atmospheric Studies (CFCAS), and the NSERC/MARTEC/MSC Industrial Research Chair.

References

- Davidson FJM, Greatbatch RJ, de Young B (2001) Asymmetry in the response of a stratified coastal embayment to wind forcing. *J Geophys Res* 106:7001–7015
- Davies AM, Flather RA (1978) Computing extreme meteorologically induced currents, with application to the northwest European continental shelf. *Cont Shelf Res* 7:643–683
- Dupont F, Hannah CG, Greenburg DA, Cherniawsky JY, Naimie CE (2002) Modelling system for tides. *Can Tech Rep Hydrogr Ocean Sci* 221. Fisheries and Oceans Canada, Bedford Institute of Oceanography, Dartmouth, Canada. 72 pp
- Fogarty CT, Greatbatch RJ, Ritchie H (2006) The role of anomalously warm sea surface temperatures on the intensity of Hurricane Juan

- (2003) during its approach to Nova Scotia. *Mon Weather Rev* 134:1484–1504
- Gill AE (1982) *Atmosphere-ocean dynamics*. Academic, San Diego, p 662
- Greatbatch RJ, Sheng J, Eden C, Tang L, Zhai X, Zhao J (2004) The semi-prognostic method. *Cont Shelf Res* 24:2149–2165
- Large WG, Pond S (1981) Open ocean momentum flux measurements in moderate to strong winds. *J Phys Oceanogr* 11:324–336
- Large WG, McWilliams JC, Doney SC (1994) Oceanic vertical mixing: a review and a model with a nonlocal boundary layer parameterization. *Rev Geophys* 32:363–403
- Mellor GL (2004) *Users guide for a three-dimensional, primitive equation, numerical ocean model*. Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, p 56
- Mellor GL, Yamada T (1982) Development of a turbulence closure model for geophysical fluid problems. *Rev Geophys Space Phys* 20:851–875
- Orlanski I (1976) A simple boundary condition for unbounded hyperbolic flows. *J Comput Phys* 21:251–269
- Pawlowicz R, Beardsley B, Lentz S (2002) Classical tidal harmonic analysis including error estimates in MATLAB using T_TIDE. *Comput Geosci* 28:929–937
- Petrie B, Topliss BJ, Wright DG (1987) Coastal upwelling and eddy development off Nova Scotia. *J Geophys Res* 29:12979–12991
- Price JF (1981) Upper ocean response to a hurricane. *J Phys Oceanogr* 11:153–175
- Safer A (2002) MEPS: a prototype for the study of coastal dynamics. *Sea Technology* 40:10–14
- Sheng J, Wang L (2004) Numerical study of tidal circulation and nonlinear dynamics in Lunenburg Bay, Nova Scotia. *J Geophys Res* 109:C10018. doi:10.1029/2004JC002404
- Sheng J, Yang B (2008) A five-level nested-grid coastal circulation prediction system for Canadian Atlantic coastal waters, Proceedings of 10th international conference on estuarine and coastal modeling, in press
- Sheng J, Wright DG, Greatbatch RJ, Dietrich DE (1998) CANDIE: a new version of the DieCAST ocean circulation model. *J Atmos and Ocean Tech* 15:1414–1432
- Sheng J, Greatbatch RJ, Wright DG (2001) Improving the utility of ocean circulation models through adjustment of the momentum balance. *J Geophys Res* 106:16711–16728
- Sheng J, Greatbatch RJ, Zhai X, Tang L (2005) A new two-way nesting technique based on the smoothed semi-prognostic method. *Ocean Dynamics* 55:162–177. doi:10.1007/s10236-005-00005-6
- Sheng J, Zhai X, Greatbatch RJ (2006) Numerical study of the storm-induced circulation on the Scotia shelf during Hurricane Juan using a nested-grid ocean model. *Prog Oceanogr* 70:233–254
- Sheng J, Zhao J, Zhai L (2008) Examination of circulation, dispersion and connectivity in Lunenburg Bay of Nova Scotia using a nested-grid circulation model. *J Marine Systems*, in press
- Smagorinsky J (1963) General circulation experiments with the primitive equation. I. The basic experiment. *Mon Wea Rev* 21: 99–165
- Smith PC, Schwing FB (1991) Mean circulation and variability on the Eastern Canadian Continental Shelf. *Cont Shelf Res* 11:977–1012
- Thompson KR, Sheng J (1997) Subtidal circulation on the Scotian Shelf: assessing the hindcast skill of a linear, barotropic model. *J Geophys Res* 102:24987–25003
- Thompson KR, Kelley DE, Sturley D, Topliss B, Leal R (1998) Nearshore circulation and synthetic aperture radar: an exploratory study. *Int J Remote Sensing* 19:1161–1178
- Thompson KR, Ohashi K, Sheng J, Bobanovic J, Ou J (2007) Suppressing bias and drift of coastal circulation models through the assimilation of seasonal climatologies of temperature and salinity. *Cont Shelf Res* 27:1303–1316

- Wang L, Sheng J, Hay AE, Schillinger DJ (2007) Storm induced circulation in Lunenburg Bay of Nova Scotia. *J Phys Oceanogr* 37:873–895
- Zhai X, Greatbatch RJ, Sheng J (2004) Advective spreading of storm-induced inertial oscillations in a model of the northwest Atlantic Ocean. *Geophys Res Lett* 31:L14315. doi:[10.101029/2004/GL020084](https://doi.org/10.101029/2004/GL020084)
- Zhai L, Sheng J, Greatbatch RJ (2008a) Application of a nested-grid ocean circulation model to Lunenburg Bay of Nova Scotia: verification against observations. *J Geophys Res* 113:C02024. doi:[10.1029/2007/JC004230](https://doi.org/10.1029/2007/JC004230)
- Zhai L, Sheng J, Greatbatch RJ (2008b) Baroclinic dynamics of wind-driven circulation in a stratified bay: a numerical study using models of varying complexity. *Cont Shelf Res*. doi:[10.1016/j.csr.2008.05.005](https://doi.org/10.1016/j.csr.2008.05.005)

Modeling the water exchanges between the Venice Lagoon and the Adriatic Sea

Debora Bellafiore · Georg Umgiesser · Andrea Cucco

Received: 3 March 2008 / Accepted: 20 September 2008 / Published online: 23 October 2008
© Springer-Verlag 2008

Abstract A hydrodynamic model of the Venice Lagoon and the Adriatic Sea has been developed in order to study the exchanges at the inlets of the Venice Lagoon, a complex morphological area connecting the sea and the lagoon. The model solves the shallow water equations on a spatial domain discretized by a staggered finite element grid. The grid represents the Adriatic Sea and the Venice Lagoon with different spatial resolutions varying from 30 m for the smallest channels of the lagoon to 30 km for the inner areas of the central Adriatic Sea. Data from more than ten tide gauges displaced in the Adriatic Sea have been used in the calibration of the simulated water levels. After the calibration, the tidal wave propagation in the North Adriatic and in the Venice Lagoon is well reproduced by the model. To validate the model results, empirical flux data measured by acoustic Doppler current profiler probes installed inside the inlets of Lido and Malamocco have been used and the exchanges through the three inlets of the Venice Lagoon have been analyzed. The comparison between modeled and

measured fluxes at the inlets outlines the efficiency of the model to reproduce both tide- and wind-induced water exchanges between the sea and the lagoon. Even in complex areas, where highly varying resolution is needed, the model is suitable for the simulation of the dominating physical processes.

Keywords Hydrodynamic model · Finite elements · Venice Lagoon · Adriatic Sea · Sea–lagoon exchange

1 Introduction

The lagoon of Venice is a complex and unique environment both because of the internal hydrodynamics and the high variety in its morphological characteristics. Because of the presence of channels, shallow flats and multiple connections with the open sea the lagoon is continuously changing. For these aspects, the Venice Lagoon and the coastal areas in front of it can be seen as a big laboratory for the study of hydrodynamic processes.

On the other hand, the area deserves its main importance from the presence of the city of Venice inside the lagoon. Many studies are driven by the need to preserve this natural environment. In the last decades, due to the increased frequency of the flooding events and to the deterioration of the water quality, the ongoing research has been focused to control and preserve the hydrological, morphological and, bio-geo-chemical characteristics of the lagoon. Its connections with the open sea play a central role, and the study of the mass balance through the inlets is fundamental in order to monitor the conservation of the lagoon itself.

Responsible editor: Pierre Lermusiaux

D. Bellafiore · G. Umgiesser (✉)
Institute of Marine Science–National Research Council
(ISMAR-CNR), Riva dei Sette Martiri 1364/a,
30122 Venice, Italy
e-mail: georg.umgiesser@ismar.cnr.it

D. Bellafiore
EuroMediterranean Centre for Climate Change (CMCC),
c/o Consorzio Venezia Ricerche, Venice, Italy

A. Cucco
Institute for the Coastal Marine Environment–National
Research Council (IAMC-CNR), Oristano, Italy

The city of Venice is an island situated approximately in the center of the lagoon with other important islands in the southern and in the northern part. The lagoon is connected to the Adriatic Sea through three inlets that guarantee the water exchange with the open sea. The southern and the central inlets (Chioggia and Malamocco, respectively) are about 500 m wide, whereas the north-most inlet (Lido) is nearly 1,000 m wide. The maximum depth is around 8 m for Chioggia and 14 m for Malamocco and Lido.

The inlets have undergone major changes in the second part of the nineteenth and the first part of the twentieth century. Due to silting up of the entrances, only small boats could pass in this period. Therefore, jetties have been constructed that reach 2–3 km into the Adriatic Sea and that gave the inlets the shape and morphology that can nowadays be observed.

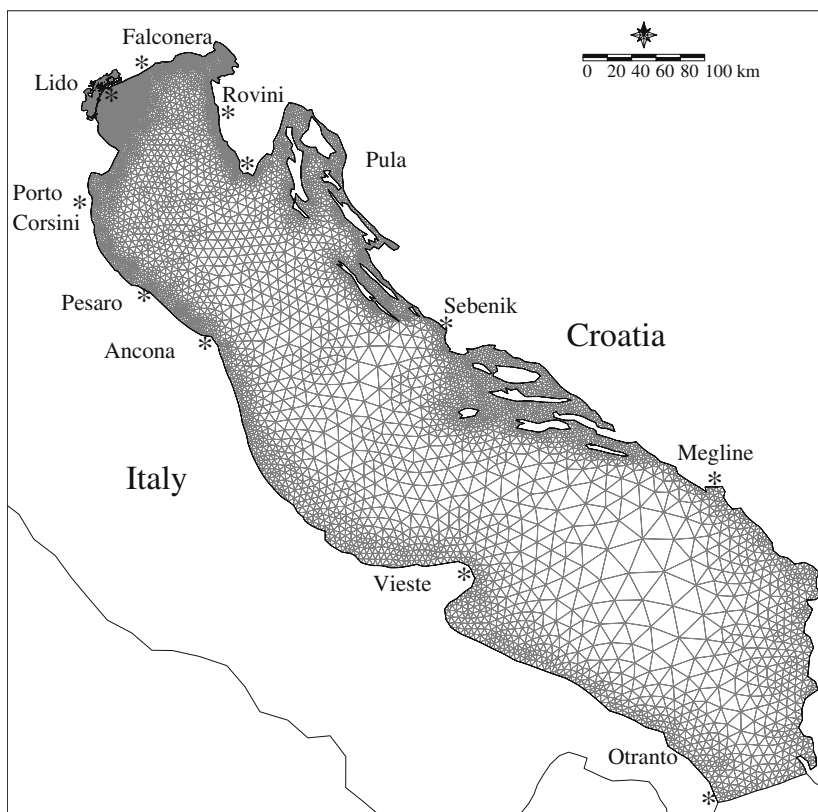
To our knowledge, in the past no major studies have been carried out that try to measure or describe the exchange and its mechanisms through the inlets. This might be surprising, if the importance of the inlets for the maintenance of the lagoon is considered. Only recently a major effort in this direction has been undertaken, investigating the exchange mechanism between the lagoon and the Adriatic Sea, both by measurements of fluxes and biochemical parameters at the

inlets (Bianchi et al. 2004). Bottom mounted acoustic Doppler current profilers (ADCPs) have been installed at the three inlets and data have been analyzed (Gačić et al. 2002). With the data of these field campaigns a hydrodynamic model could be validated. One part of the modeling effort is presented in this work.

In the recent past, few modeling studies have been carried out to investigate the Sea–Lagoon water exchange mechanism. In particular in Umgiesser (2000), a first attempt to understand the residual currents due to the most prominent wind regimes present in the Northern Adriatic can be found. A finite element model (Umgiesser and Bergamasco 1993, 1995) to simulate the residual currents of a complete year (1987) is used there. The other study (Bergamasco et al. 1998) applied the Princeton Ocean Model (POM; Blumberg and Mellor 1987; Mellor 1991) to the lagoon and the coastal area of the Adriatic Sea to simulate both hydrodynamics and primary production.

Both approaches suffered from main deficiencies. Both models could not be validated with data because flux measurements were not available at the inlets. Moreover, the second study applied the POM model with a 1,200-m grid size, too coarse to resolve the important hydrodynamic features of the Venice Lagoon. On the other hand, the first study used a calibra-

Fig. 1 The numerical domain of the Adriatic Sea and the location of the tide gauges considered during the calibration process



ted model for the water levels inside the lagoon with good resolution of the channel system (due to its finite element method), but failed to describe well the interface dynamics, since the model domain ended exactly at the inlets.

This article is finally concerned with the interaction process modeling, making a further step in understanding and describing how the Venice Lagoon interacts with the open sea. The suitability of finite element approach allows the reproduction of the hydrodynamics in such a complex modeling area with the necessary high resolution. The finite element model, even if 3D in structure, is applied here in its 2D version. The model domain comprises both the Venice Lagoon and the whole Adriatic Sea, in order to move the open boundary as far away as possible from the area of interest. In this study, the model is calibrated with harmonic constants of tidal data available at the inlets of the Lagoon and validated in some stations around the Adriatic Sea (Fig. 1). The fluxes computed at the inlets are then validated with the astronomical contribution of the measured fluxes. Additionally, the model capability to reproduce residual signals due to the meteorological forcings at the inlets is investigated.

2 Materials and methods

In this section, a description of the numerical model and a summary of the adopted experimental data (sea surface elevation, wind, and ADCP data) are given. Thereafter, the general simulation set-up is presented.

2.1 The hydrodynamic model

The hydrodynamic model applied here has been developed at ISMAR-CNR (Istituto di Scienze Marine; Umgiesser and Bergamasco 1993, 1995). It has already been applied successfully in its 2D version to the Venice Lagoon (Umgiesser 2000; Melaku Canu et al. 2001; Umgiesser et al. 2004) and in its 3D version to the Adriatic Sea (Umgiesser and Bergamasco 1998).

The model uses finite elements for horizontal spatial integration and a semi-implicit algorithm for integration in time. The finite element method allows high flexibility with its subdivision of the numerical domain in triangles varying in form and size. It is especially suited to reproduce the geometry and the hydrodynamics of complex shallow water basins such as the Venice Lagoon with its narrow channels and small islands.

The numerical computation has been carried out on a spatial domain that represents the Venice Lagoon and the Adriatic Sea by means of a grid that consists

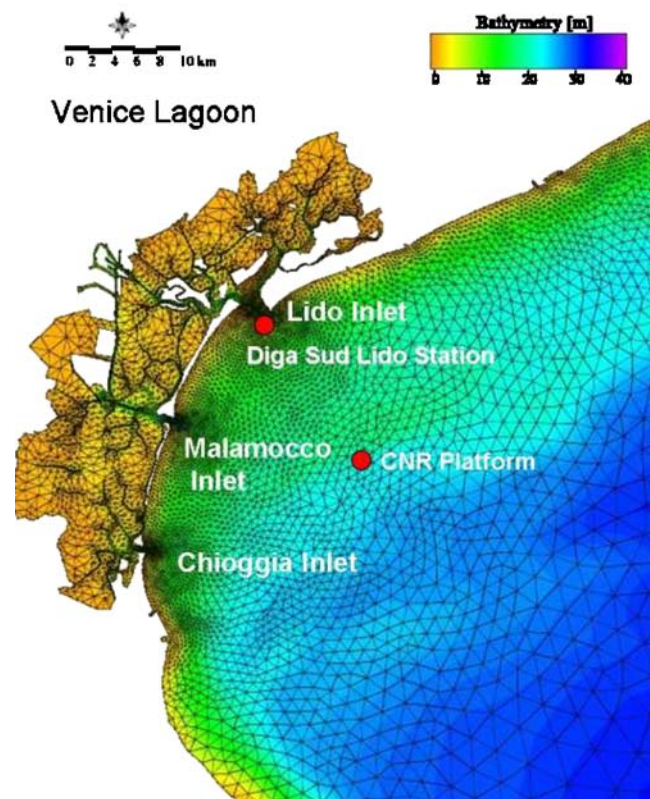


Fig. 2 Bathymetry and finite element grid of the Venice Lagoon and the Gulf of Venice, a subset of the numerical domain of the model

of 15,619 nodes and 28,827 triangular elements (Figs. 1 and 2) with a resolution varying between 30 m close to the lagoon inlets and inside the lagoon and 30 km in the central areas of the Adriatic Sea.

In the lagoon, due to its shallow areas and the relatively big amplitude of tides (± 50 cm during spring tide), stratification of water masses can develop only far from the inlets where the tidal energy is low. On the other hand, inside the inlets, where the water velocities are high (over 1 m/s), the velocity shear creates enough turbulence to mix the water column (Gačić et al. 2002). Consequently, the water exchanges between the lagoon and the sea that are driven mainly by the tide and the wind action are essentially barotropic in nature (Gačić et al. 2002). Therefore, the model has been applied in its 2D version.

The model resolves the shallow water equations in their formulations with levels and transports. The horizontal diffusion and the advective terms in the momentum equation are fully explicitly treated. The Coriolis force and the barotropic pressure gradient terms in the momentum equation and the divergence term in the continuity equation are semi-implicitly treated. The friction term is treated fully implicitly for stability

reasons due to the very shallow nature of the lagoon. This discretization provides unconditional stability for what concerns the effects of the fast gravity waves, the bottom friction and the Coriolis acceleration (Umgiesser and Bergamasco 1995).

The 2D shallow water equations, here applied, are

$$\frac{\partial U}{\partial t} - fV + gH \frac{\partial \zeta}{\partial x} + RU + X = 0 \quad (1)$$

$$\frac{\partial V}{\partial t} + fU + gH \frac{\partial \zeta}{\partial y} + RV + Y = 0 \quad (2)$$

$$\frac{\partial \zeta}{\partial t} + \frac{\partial U}{\partial x} + \frac{\partial V}{\partial y} = 0 \quad (3)$$

with U, V the barotropic transports in the x, y direction, g the gravitational acceleration, f the Coriolis parameter, ζ the sea level, with $H = h + \zeta$ the total water depth, and h the undisturbed water depth.

The friction term R is given by $R = c_B \sqrt{U^2 + V^2} / H^2$ which corresponds to a quadratic bottom friction formulation. In this expression, the bottom drag coefficient c_B can be assumed either constant (in the Adriatic Sea) or dependent on the depth through the Strickler formula

$$c_B = \frac{g}{C^2} \quad C = k_s H^{1/6} \quad (4)$$

with C the Chezy coefficient and k_s the Strickler coefficient (Venice Lagoon).

The terms X, Y contain all terms that are treated explicitly in the time integration. They read

$$X = \frac{U}{H} \frac{\partial U}{\partial x} + \frac{V}{H} \frac{\partial U}{\partial y} + \frac{H}{\rho_0} \frac{\partial p_a}{\partial x} - \frac{\tau_x^{\text{wind}}}{\rho_0} - A_H \left(\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} \right) \quad (5)$$

$$Y = \frac{U}{H} \frac{\partial V}{\partial x} + \frac{V}{H} \frac{\partial V}{\partial y} + \frac{H}{\rho_0} \frac{\partial p_a}{\partial y} - \frac{\tau_y^{\text{wind}}}{\rho_0} - A_H \left(\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} \right) \quad (6)$$

where ρ_0 is the constant water density, p_a the atmospheric pressure, A_H the horizontal eddy viscosity and $(\tau_x^{\text{wind}}, \tau_y^{\text{wind}})$ are the wind stress terms expressed as

$$\tau_x^{\text{wind}} = c_D \rho_a w_x \sqrt{w_x^2 + w_y^2} \quad \tau_y^{\text{wind}} = c_D \rho_a w_y \sqrt{w_x^2 + w_y^2} \quad (7)$$

where c_D is the wind drag coefficient, ρ_a the air density and (w_x, w_y) the x, y components of the wind velocity, respectively.

At the lateral open boundaries of the domain, the water levels are prescribed while at the closed boundaries the normal velocity is set to zero and the tangential velocity is a free parameter. This corresponds to a full slip condition. The model allows also for flooding and drying of the shallow water flats. This is especially important for the Venice Lagoon, since about 15% of the area is partially wet and dry during a spring tidal cycle. The flooding and drying mechanism has been implemented in a mass consistent way, and spurious oscillations that are generated by the algorithm are damped out fast.

An element is considered dry if at least one of its nodes has a total water depth of less than 5 cm. In this case, the element is taken out of the dynamic system, its water mass is conserved, and an extrapolation of the water level is done from the adjacent wet nodes to its dry nodes. After all of its nodes show a water depth of more than 10 cm, the element is again integrated into the system with the same water mass it had when it was taken out. The numerical oscillations that would appear when switching on and off the element can be controlled by adjusting the time weighting parameter (0.5 in the semi-implicit case). In practice, a value of 0.55 (a little more implicit) suffices to completely damp out these oscillations in a few time steps. The wet and dry mechanism and its implementation in the model are fully described in Umgiesser and Bergamasco (1993) and Umgiesser et al. (2004).

2.2 The available data

In this work two different sets of measured data have been used to calibrate and validate model results. The first dataset consists of the harmonic constants (amplitude and phase) of the main tides that characterizes the sea surface elevation (SSE) of the Adriatic Sea. For the lagoon inlets and for the Otranto station, the harmonic constants are computed from available time series, ISMAR data for Lido and APAT (Environmental Protection Agency) data for Otranto. The second dataset contains the discharge data through two of the three lagoon inlets measured by ADCP probes.

2.2.1 Tidal data

The Adriatic Sea is a basin characterized by moderate tides. The highest amplitudes are found in the northern sub-basin, extending from the line connecting Pesaro to Kamenjak to the Northern coast (Fig. 1), where the amplitude of the M2 reaches 0.266 m. Tidal observations in the Northern sub-basin have been collected from the middle of the eighteenth century. Three relevant

Table 1 Comparison between model results (m) and observations (o) of the amplitude H and phase g of the most energetic tidal constituents (M2, S2, and K1) at the tidal stations in the northern part of the Adriatic Sea

Site		H_m [cm]	H_o [cm]	$H_m - H_o$ [cm]	$(H_m - H_o)/H_o$ [%]	g_m [°]	g_o [°]	$g_m - g_o$ [°]
Lido	M2	23.30	23.38	-0.08	-0.3	286.2	285.1	0.6
	S2	14.67	14.71	-0.04	-0.3	298.1	296.7	1.4
	K1	20.53	20.00	0.53	2.6	78.5	77.8	0.7
Porto Corsini	M2	16.54	15.50	1.04	6.7	301.2	305.10	-3.9
	S2	10.31	9.2	1.11	12.1	306.9	310.0	-3.1
	K1	18.89	16.5	2.39	14.5	92.5	89.3	3.2
Falconera	M2	23.01	23.8	-0.79	-3.3	285.0	291.1	-6.1
	S2	14.80	14.10	0.70	5.0	290.3	297	-6.7
	K1	20.49	19.0	1.49	7.84	85.4	87.3	-1.9
Trieste	M2	24.47	26.4	-1.93	-7.31	280.9	278.1	2.8
	S2	15.84	16.00	-0.16	-1.0	286.1	284.0	2.1
	K1	20.81	19.3	1.51	7.82	83.4	78.3	5.1
Rovinj	M2	17.94	19.10	-1.16	-6.1	272.7	272.1	0.6
	S2	11.38	11.2	0.18	1.6	277.2	277.0	0.2
	K1	19.13	16.7	2.43	14.55	80.2	79.3	0.9
Pula	M2	14.82	15.0	-0.18	-1.2	266.1	267.1	-1.0
	S2	9.28	8.7	0.58	6.7	269.9	273.0	-3.1
	K1	18.19	16.1	2.09	12.98	78.4	77.3	1.1

stations are present in the Northern part of the basin, Trieste, Venezia-Lido, and Rovinj (Fig. 1), where the monitoring of tides covers more than a century.

Modeling the tidal currents in the Adriatic Sea should be able to reproduce the behavior of the seven main tidal constituents in the basin, the four semi-diurnal ones M2, S2, N2, K2, and the three diurnal ones, K1, O1, and P1 (Polli 1960). The knowledge about tides in the Adriatic Sea derives from previous works (Polli 1960, 1961; Mosetti 1987) where the main harmonic constituents were derived analyzing the tide gauge measurements in a number of stations along both the western and the eastern coast (Fig. 1). Ten of these stations, covering the whole basin, have been chosen to compare modeled and measured harmonic constituents

after the calibration procedure. For these points, the harmonic constants obtained from Polli (1960) for the most energetic semi-diurnal and diurnal tides (M2, S2, and K1) are reported in Tables 1 and 2.

As a convention, all phases in Tables 1 and 2 are given with respect to the meridian of central Europe, Greenwich plus 1 h. As explained in Tomasin (2005), the harmonic constants are subject to the periodical changes in the orbits of the sun and moon and for this reason they slightly change in time. The algorithm described in Tomasin (2005) gives the opportunity to update the date shown in Polli (1960), annualizing them for the year 2002.

For what concerns the tides in the Venice Lagoon, measured data of amplification and delay of the main

Table 2 Comparison between model results (m) and observations (o) of the amplitude H and phase g of the most energetic tidal constituents (M2, S2, and K1) at the tidal stations in the southern part of the Adriatic Sea

Site		H_m [cm]	H_o [cm]	$H_m - H_o$ [cm]	$(H_m - H_o)/H_o$ [%]	g_m [°]	g_o [°]	$g_m - g_o$ [°]
Pesaro	M2	11.23	12.7	-1.47	-11.57	310.5	313.1	-2.6
	S2	6.77	6.8	-0.03	-0.4	316.9	313.0	3.9
	K1	17.34	16.0	1.34	8.38	95.3	92.3	3.0
Ancona	M2	6.15	6.5	-0.35	-5.38	324.0	334.1	-10.1
	S2	3.46	3.5	-0.04	-1.1	333.9	347.0	-13.1
	K1	15.59	13.7	1.89	13.80	96.3	96.3	0.0
Sebenik	M2	6.16	6.2	-0.04	-0.65	138.9	137.1	1.8
	S2	4.46	4.4	0.06	1.4	138.7	132.0	6.7
	K1	11.48	9.7	1.78	18.35	70.9	65.3	5.6
Vieste	M2	8.71	9.3	-0.59	-6.34	103.8	107.1	-3.3
	S2	5.81	6.0	-0.19	-3.2	110.0	115.0	-5.0
	K1	6.92	5.3	1.62	30.57	102.8	99.3	3.5
Megline	M2	9.440	9.0	0.44	4.88	108.3	101.1	7.2
	S2	6.16	5.9	0.26	4.4	112	103.0	9.0
	K1	6.85	5.2	1.65	31.73	73.8	60.3	13.5

diurnal and semi-diurnal constituents with respect to the Diga Sud Lido tide gauge are reported for many stations inside the lagoon by Goldmann et al. (1975). This dataset has been considered for the calibration of the model to reproduce the tidal propagation in the Venice Lagoon (Umgiesser et al. 2004). Recently, new harmonic constants, calculated from tidal gauges time series, have been provided for the Venice Lagoon by the Environment Protection Agency APAT (Ferla et al. 2007). In the same paper, even an accurate analysis on long-term variations in the tidal regime in the Venice Lagoon has been presented. From this dataset, the harmonic constants for the Lido station have been derived. From a comparison between the cited papers, the two sets of harmonic constants computed for the lagoon are significantly different, which might be ascribed mainly to the influence of the deep morphological changes occurred in the last decades (Ferla et al. 2007).

2.2.2 The flow rate ADCP data

Recent campaigns in the area of interest provide ADCP data for the years 2001–2002. The aim of these measurements is to study the time-dependent variability of the inlet currents as well as of water exchange rates. The current data have been collected using bottom-mounted ADCPs installed in each inlet. Current speed and direction along the water column are recorded every 10 min with a vertical resolution of 1 m in selected locations inside the inlets. During a preliminary phase, measurement campaigns have been carried out to estimate the relationship between the vertically averaged water velocity collected by the fixed ADCP and the inlet flow rate. About 100 ship-borne ADCP surveys were conducted to estimate the water inflow and outflow through each inlet both during spring and neap tide. Comparing the discharge results with the average velocity collected by the bottom mounted ADCP for the same period, the parameters of a linear correlation function have been calculated (Gačić et al. 2004). Therefore, the flow rate is available every 10 min applying the calculated linear regression formula to the vertically integrated measured current values (Gačić et al. 2002).

2.2.3 Wind data

The meteorological contribution can affect strongly the dynamics of the water exchanges at the lagoon inlets (Gačić et al. 2002). The residual water level in the Adriatic Sea and the related residual flow through the lagoon inlets are generated by the intense meteoro-

logical phenomena that occur in this area. Therefore, in this work, wind and pressure fields over the whole Adriatic Sea have been considered as surface boundary forcings. The adopted wind fields are the $0.5^\circ \times 0.5^\circ$ T511 European Centre for Medium Weather Forecast (ECMWF) 6-h data.

As previous studies stressed (Cavaleri and Bertotti 1996), these data underestimate the real meteorological status, especially in the Venice Gulf, where the meteorological phenomena are more intense at the synoptic scale. In the last years, various correction factors have been applied for both wind direction and wind intensity of ECMWF atmospheric fields computed for the Mediterranean. In this work, a spatially variable correction factor provided by Cavaleri and Bertotti (1996) has been applied to the wind dataset used to force the model. In Cavaleri and Bertotti (1996), the corrected wind field has been compared with measured data in the CNR Oceanographic Platform, a station 15 km offshore in front of the Venice Lagoon (Fig. 2).

The same correction mask has been adopted to set up a storm surge finite element model made operative at the Venice Municipality (Bajo et al. 2007) with the aim of forecasting the water level elevation generated by the meteorological action (the residual water level). During this application, it has been seen that it was not necessary to apply the correction mask to the pressure because the pressure field is reproduced correctly by the ECMWF model over the whole basin (Zampato et al. 2005).

2.3 The general simulation set-up

All simulations presented in this work are for the year 2002 and have been carried out using a time step of 300 s. This time step could be achieved due to the unconditionally stable time integration scheme of the finite element model.

The bathymetric data interpolated on the Adriatic Sea finite element grid are taken from the NOAA 1-min resolution dataset.

All simulations have been run with the same wind forcing of the year 2002 and the lateral boundary conditions at the Otranto strait for the water levels derived from the harmonic constants and the storm surge model. The open boundary has been chosen as a straight line across the Strait of Otranto (Fig. 1). This parameterization was suggested by the results obtained in a previous work by Cushman-Roisin and Naimie (2002) who applied with success a finite element tidal model of the Adriatic Sea. This seems a logical choice for the Adriatic Sea being the Strait of Otranto the

narrowest part of the whole basin. Furthermore, being the Otranto line far away from the study area (Venice Lagoon and northern Adriatic coast), disturbances to the numerical solution for the region of interest, induced by the artificial imposition of boundary conditions, can be avoided. At each node of the Otranto open boundary, water level time series are imposed. They are obtained by adding a surge signal coming from an operational storm surge model (Bajo et al. 2007) to the astronomical tide computed from harmonic constants.

The contribution of the fresh water input released inside the basin by the main Italian rivers is not taken into account. Flow discharge measurements at the Malamocco inlet register a flow rate of 10,000 m³/s, and all three inlets together may show flow rates as high as 24,000 m³/s. These numbers can be compared with the main river in the Adriatic Sea, the Po river, that shows an average discharge rate of 1,500 m³/s (Gačić et al. 2002).

Even if the river runoff strongly influences the baroclinic circulation in the Adriatic Sea, it is far from giving appreciable contributions to the barotropic water circulation in the coastal area in front of the Venice Lagoon. In particular, considering the Po river runoff, the advective fluxes and the water level gradient induced by the river discharge induce a momentum flux over the water column which influences the water circulation only locally. Simulations in Umgiesser and Bergamasco (1998) show how the jet created by the Po first deviates to North East but then deflects southward due to the conservation of potential vorticity. Maps show how it does not influence the study area near the inlets of the Venice Lagoon. Therefore, the influence of the barotropic contribution of the Po river runoff on the sea–lagoon water exchanges and local water circulation can be neglected.

The bottom boundary condition enforces quadratic friction based on barotropic (depth-averaged) velocity according to the classical quadratic drag law. In this application, two formulations for the bottom friction coefficient c_B have been considered. For the deeper areas of the domain (Adriatic Sea), the coefficient is imposed to be constant with a value of $2.5 \cdot 10^{-3}$. However, in the shallow parts (Venice Lagoon), the Strickler formula has been used where the friction parameter depends on the water depth. This approach was suggested by the application of the model to the Venice Lagoon alone (Umgiesser et al. 2004). Inside the lagoon, different Strickler coefficients have been used to distinguish the behavior of channels, tidal flats, and the rest, and at the three inlets slightly different values of this coefficient are used in inflow and in outflow conditions to parameterize the unresolved physical processes such as

the local head loss due to sudden expansion of the outflowing water (Umgiesser et al. 2004).

What concerns the wind drag coefficient c_D , a constant value of $2.5 \cdot 10^{-3}$ has been adopted. The lateral diffusion parameter A_H has been set to 0. The initial conditions for the water levels and the velocities are zero in the whole basin and a spin-up time of 30 days for all the simulations has been imposed (Cucco and Umgiesser 2005). This time was enough to damp out all the noise that was introduced through the initial conditions.

3 Results

In this section, the model results are presented. First, the model calibration through harmonic constants is described and then the water levels are validated. In the third section, the fluxes through the lagoon inlets are computed and compared with the ADCP data. The fourth part analyzes the reproduction of the total levels and fluxes, considering both the astronomical and the meteorological signals. Finally, the focus is the reproduction of strong wind events, considering the two major winds blowing in the Northern Adriatic. The first one, called Bora from NE, is a cold wind that characterizes extreme events generally occurring during winter time (Orlić et al. 1992). The second one, Scirocco, is a wet, warm wind, coming from SE largely present in autumn and mostly responsible for the flooding events of the city of Venice (Orlić et al. 1992).

3.1 Calibration of harmonic constants

The model has been calibrated to reproduce the harmonic constants of the tidal constituents M2, S2, K1, N2, K2, O1, and P1 in the Adriatic Sea and, more specifically, in the Gulf of Venice. These tidal constituents are known for Otranto from a 9-year-long hourly water level time series recorded by APAT at the Otranto station. They are representative for the tides of the western Italian border of the strait. In addition to the astronomical signal, surge values, produced by a numerical model applied to the whole Mediterranean Sea, are imposed at the Otranto open boundary (Bajo et al. 2007). This surge model is adopted by the Venice Municipality, running operationally to predict the flooding in the city.

To determine the tidal constituents for the eastern Albanian border, an iterative procedure has been applied which can be schematically explained as follows. First, the harmonic constants of Albania are set to the same values as the ones for Otranto. A 1-year-long time

series is produced from these constituents and the surge signal is added to these water levels. For every node on the border, a linear interpolation is carried out between the water levels at Otranto and Albania.

With these water levels for the open boundary, a 1-year-long (8,760 h) simulation is carried out. This period is needed because it is the necessary and sufficient time to separate the K2 and S2 frequencies in the harmonic analysis of results (Foreman 1996).

At the end of each calibration run, the harmonic constants are extracted from the model for the station closest to the study area, Diga Sud Lido (Fig. 1), and compared with the empirical ones. The differences between modeled and observed harmonic constants are then used in changing the harmonic constants imposed for the Albanian side of the open boundary. In particular, the proportion of observed and modeled amplitudes and the time lag for the phases is used to compute new tidal constituents for Albania. With these new values, a new set of boundary conditions is computed for each node on the open boundary as explained above and a new run is performed.

The calibration is ended when the difference between the amplitude of the three main measured and modeled harmonic constants (M2, S2, K1) falls below 3%, and the phase shift is less than 2° at Diga Sud Lido. To reach this result, around 20 iterations were needed. At this point, a comparison for the other ten stations in the Adriatic Sea is performed to assure a realistic tidal behavior in the whole Adriatic Sea and not only an ad hoc solution for the Venice Lagoon.

3.2 Water level validation

First, a comparison between modeled and measured harmonic constants at Diga Sud Lido station has been carried out. In Table 3, the set of harmonic constants obtained at Diga Sud Lido from the model calibration and from measurements taken by ISMAR-CNR in the year 2002 is shown. In Table 4, the final set of harmonic

Table 3 Comparison between model results (m) and observations (o) of the amplitude H and phase g of the whole set of tidal data at Diga Sud Lido station

Constituent	H_m [cm]	H_o [cm]	g_m [°]	g_o [°]
M2	23.00	23.38	292.5	291.4
S2	14.67	14.71	298.2	296.7
N2	4.13	4.16	290.4	290.1
K2	5.67	5.83	312.8	311.3
K1	20.53	20.00	88.6	87.9
O1	6.32	6.55	59.0	53.4
P1	5.56	5.39	70.9	72.0

Table 4 Open boundary condition

Constituent	H_O [cm]	g_O [°]	H_A [cm]	g_A [°]
M2	6.8	133.1	8.58	59.33
S2	4.0	141.0	4.65	71.89
N2	1.2	131.2	1.61	1.84
K2	1.1	136.9	1.52	221.33
K1	2.3	91.1	4.04	19.33
O1	0.9	75.7	1.71	33.77
P1	0.8	85.1	1.37	29.04

Amplitude H and phase g of the main tidal constituents for the western (Otranto, O) and eastern station (Albania, A) of the SSE imposed as forcing along the open boundary of Otranto

constants of the open boundary (Otranto and Albania) are listed. In Tables 1 and 2, the amplitudes and the phase lags of the main semi-diurnal and diurnal tides (M2, S2, and K1) are compared with the observations along the coasts of the Adriatic sea (see Fig. 1 for reference). The phase in the table is expressed with respect to the meridian of central Europe, as conventionally applied in Polli (1960) and annualized for the year 2002 as previously explained.

As we could expect, the calibrated model reproduces the SSE at Diga Sud Lido station with high accuracy. We do not show the comparison graph for Lido station because the results are quite obvious even on the basis of the harmonic constants shown in Table 3: small differences with observed data for the harmonic constants in this station are registered, with differences of millimeters in amplitude and some degrees for the phases; the highest difference at Diga Sud Lido is less than 5° for the O1 component which is one of the less important components of the ones taken into account.

The area of interest for this application is the Northern Adriatic Sea, which is limited in the Southern part by the line connecting Pesaro to Pula (Fig. 1). The stations located in this area are Falconera, Porto Corsini, Trieste, Rovinj, Pesaro, and Pula (Table 1). In the northern half of the basin, the two semi-diurnal components are well reproduced with an absolute difference in amplitude never higher than 2 cm (Table 1). Following the North-East coast of the basin, amplitude differences are oscillating slightly around zero. These two harmonic components are quite in phase with measurements with differences never higher than -6.7° . This value is reached in the Falconera station. There is a 20-min delay in the modeled data for these components. In the same stations, the major diurnal component is well reproduced, with -1.93 cm (-7%) as the highest difference in Trieste. When considering what is the main aim of this work, the modeling of the interaction between the Venice Lagoon

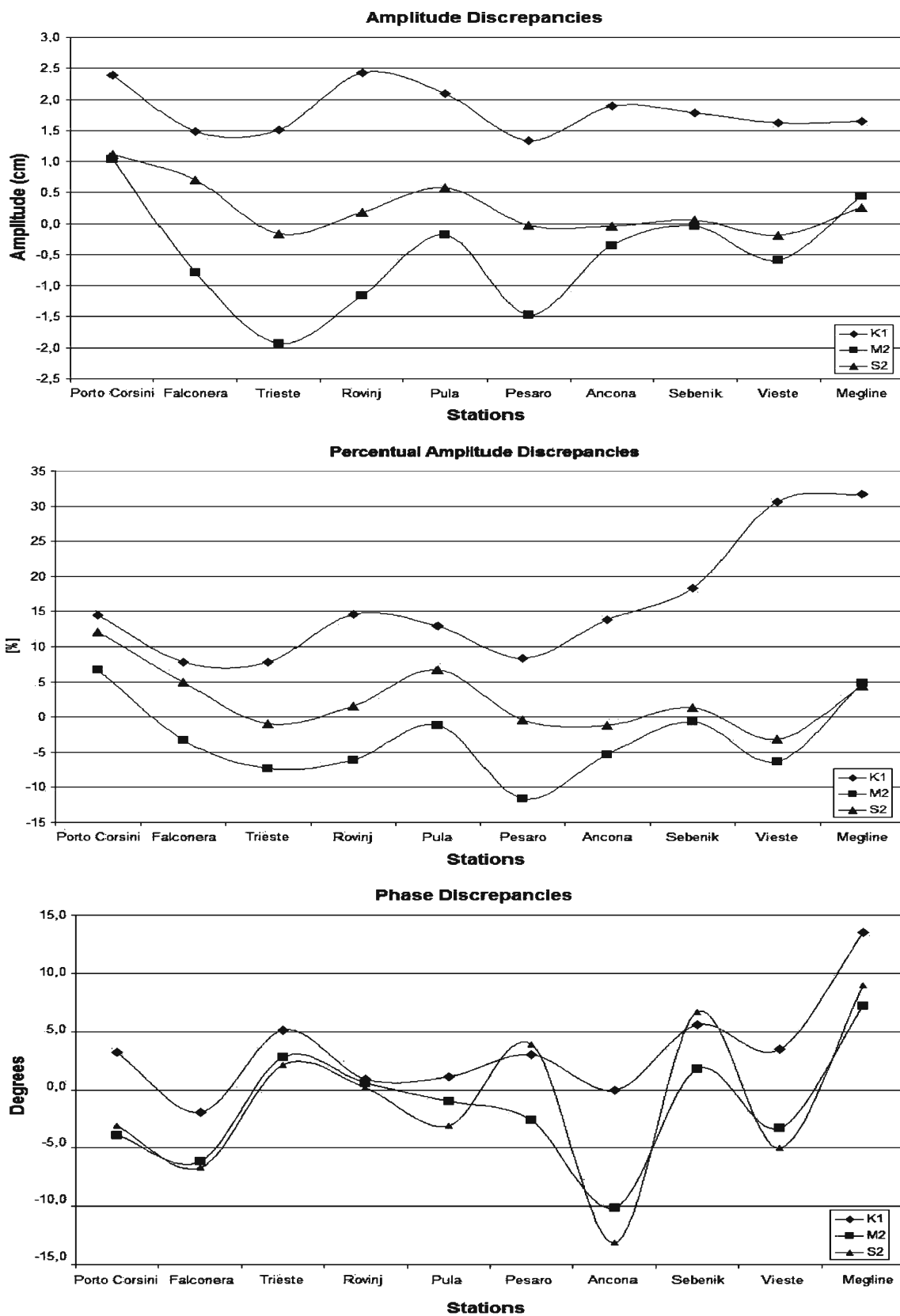


Fig. 3 Graph of amplitude, relative amplitude and phase differences of the three main modeled and measured harmonic constants (M2, S2, K1) for the ten stations in the Adriatic Sea

and the open sea, these results can be considered satisfactory.

For the other stations in the Adriatic Sea, the result for Ancona station is striking. Near Ancona, the M2 amphidromic point is located. This means that tidal variations are small and water level measurements can be affected by higher errors because of the low signal to noise level. This could explain the high difference in phase (around -12°). Going South, a general overestimation of the diurnal component K1 can be observed but it has to be mentioned that the difference never exceeds 1.5 cm, even if this translates into a high error in terms of percentage (Fig. 3, Table 2).

The relatively high amplitude errors for K1 in the stations of the Dalmatian coast should be stressed. This might be due to the grid resolution along the coast. Because of the interest in simulating the North Adriatic coastal dynamics, the applied finite element grid has high resolution in proximity of the Venice Lagoon. The Dalmatian coast is less resolved than the Northern part of the Adriatic Sea because of the attempt to limit the number of elements of the grid and to lower the time of computation.

For the other two components, M2 and S2, no general trends in amplitude difference can be observed. Phase differences do oscillate around zero for all the stations along the Adriatic Sea, increasing in module approaching the southern open boundary at the Otranto Strait (Fig. 3). A smooth anticipation of the real signal along the low resolved eastern coast of the southern Adriatic and a specular delay along the western littoral, near Vieste station (Fig. 1), is observed for these harmonic constituents. The presence of the open boundary, where water levels are imposed, is certainly influencing the results for the stations located close to it.

In conclusion, the choice of moving the open boundary to the Otranto strait was correct. In the southern part of the basin, errors are higher, mostly due to the influence of the open boundary. On the other hand, the results in the Northern part of the basin do not evidence an influence of these modeling choices. What concerns the model results of the propagation of the tide inside the Venice Lagoon we always refer to Umgiesser et al. (2004) for a more accurate description.

3.3 Flux data validation

Once the model had been calibrated and validated with the water levels, it has been used to estimate the water exchange through the three inlets. The presence of both tidal and meteorological data forcings has to be stressed because the simulation of barotropic currents in the Adriatic Sea needs to consider the whole set of significant forcings. From the evidence of measurements at the inlets, it can be deduced that the flow acceleration is generated by the pressure gradient due to the sea level slope. Velocities are of the order of 1 m/s and the bottom friction term is balanced by the local acceleration and the horizontal pressure gradient (Gačić et al. 2004). The wind action is mainly barotropic at the inlets (Gačić et al. 2002). The conjunction of tidal cycles and atmospheric pressure and wind produces extreme events propagating into the lagoon from the inlets, in particular when SE wind (Scirocco) is blowing in a low pressure configuration (generally in winter time; Gačić et al. 2004).

To validate the model results, only the tidal signal of the measured discharge data has been considered. Harmonic analysis is therefore applied to the ADCP data collected at Lido and Malamocco inlets for the

Fig. 4 Time series of total, residual and tidal components of fluxes from ADCP data in the Malamocco inlet for November 2002

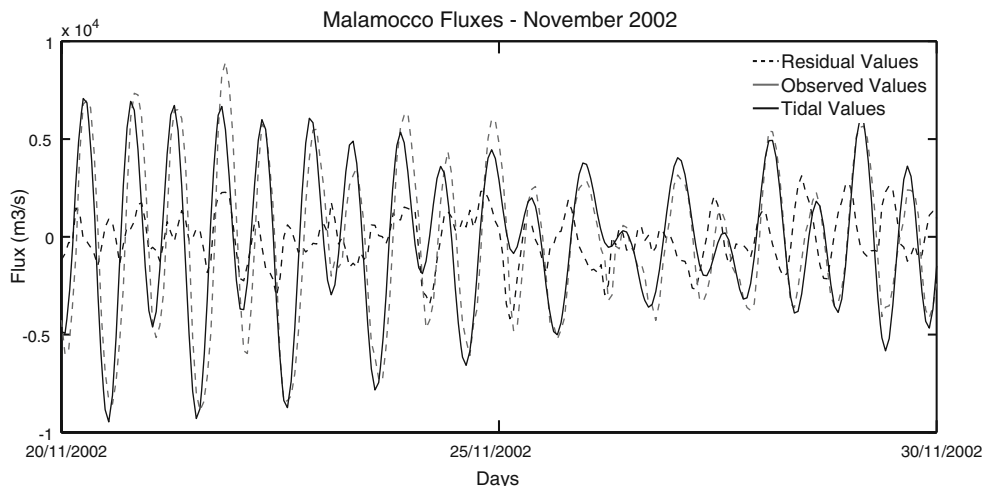
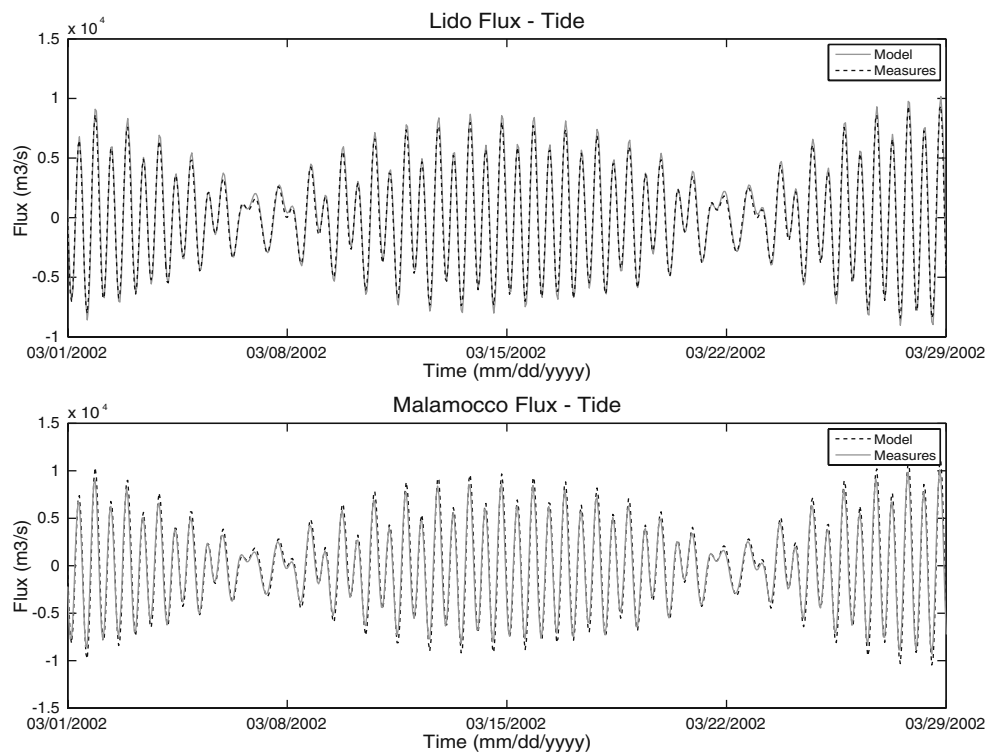


Fig. 5 Tidal discharge time series comparison between modeled and measured data in the Lido and Malamocco inlets



year 2002, as has been done before for the water levels (Fig. 4). The contribution of the three major diurnal and semi-diurnal tidal components to the total flow rate—M2, S2, and K1—is obtained and the parameters related to these constituents are presented in Table 5. As it can be seen from the graphs in Fig. 5, the two signals match nicely. To quantify the error, a comparison between the computed flux amplitudes and phases has been done in Table 5.

Analyzing the tidal reproduction of water fluxes, different considerations can be done for diurnal and semi-diurnal components. Satisfactory results are obtained in simulating the M2 and S2 components: the amplitude error is always less than 3.65% for both inlets (Fig. 5, Table 5). The M2 and S2 modeled fluxes anticipate the measured ones by around 20 min at the

Lido inlet and 30 min at the Malamocco inlet. The worst reproduction, as in the case of water levels, is again for the K1 constituent. Following the results of the harmonic analysis, from ADCP data it can be stated that M2, S2, and K1 components represent about 50%, 30%, and 20% of tidal signal (Table 6). Weighting each component with these values, the comprehensive tidal difference in amplitude is 2.7% for Lido the inlet and 2.9% for the Malamocco inlet. The phase anticipation of modeled tidal fluxes is 23 min for Lido and 30 min for Malamocco. Moving from harmonic constituents to the modeled and observed tidal time series, the computed correlation coefficient is 0.99 both for Lido and Malamocco inlets.

Moreover, the model reproduces an additional aspect, already seen in measurements (Gačić et al. 2002).

Table 5 Water discharge of the Lido and Malamocco inlet

Inlet		A_m [m ³ /s]	A_o [m ³ /s]	$A_m - A_o$ [m ³ /s]	$(A_m - A_o)/A_o$ [%]	g_m [°]	g_o [°]	$g_m - g_o$ [°]
Lido	M2	4400.1	4366.8	33.28	0.76	254.9	244.8	10.1
	S2	2596.3	2504.8	91.5	3.65	267.2	255.4	11.8
	K1	1895.8	1790.4	105.4	5.9	21.3	14.4	6.8
Malamocco	M2	4725.2	4682.0	43.2	0.92	245.5	232.1	13.4
	S2	2825.0	2790.7	34.3	1.23	258.4	243.1	15.3
	K1	1892.1	1717.4	174.7	10.2	13.2	4.0	9.2

Comparison between empirical data (o) (derived from ADCP measurements) and model results (m) of the amplitude A and phase g of the most energetic diurnal and semi-diurnal frequencies (M2, S2, and K1)

Table 6 Weighted differences of discharges in amplitude (A) and phase (g) (in terms of minutes) for the Lido and Malamocco inlets

Inlet		Perc of modeled tidal fluxes [%]	A [%]	g [min]
Lido	M2	49.5	2.7	23
	S2	29.2		
	K1	21.3		
Malamocco	M2	50	2.9	30
	S2	30		
	K1	20		

The observed semi-diurnal signal at the Malamocco inlet anticipates the Lido one by about half an hour (Gačić et al. 2002). The simulated phase lag is of about 20 min, nicely matched, even if slightly underestimated.

The flux modeling is strictly connected with two factors: the capability in reproducing the dynamics in the lagoon and the parameterization of the bottom stress inside the three connection channels. The bottom friction coefficients chosen during the calibration phase are specific for each inlet with different values applied to simulate the inflow and outflow dynamics. The final values chosen are the ones that more uniformly distribute the error on all three inlets and reproduce the interaction processes.

To explain the differences with the empirical data, we can distinguish two different groups of errors: one related to the measured data and one related to the numerical model. The empirical discharge data, as explained in Section 2.3, have been obtained from the vertically averaged water velocity value measured by ADCP probe for one fixed location inside each inlet. From Gačić et al. (2002), we know (Table 1) that the ADCP velocity data are affected by an error of around 5% for the M2 and 10% for the K1. Since the discharge data are computed from the velocity data by

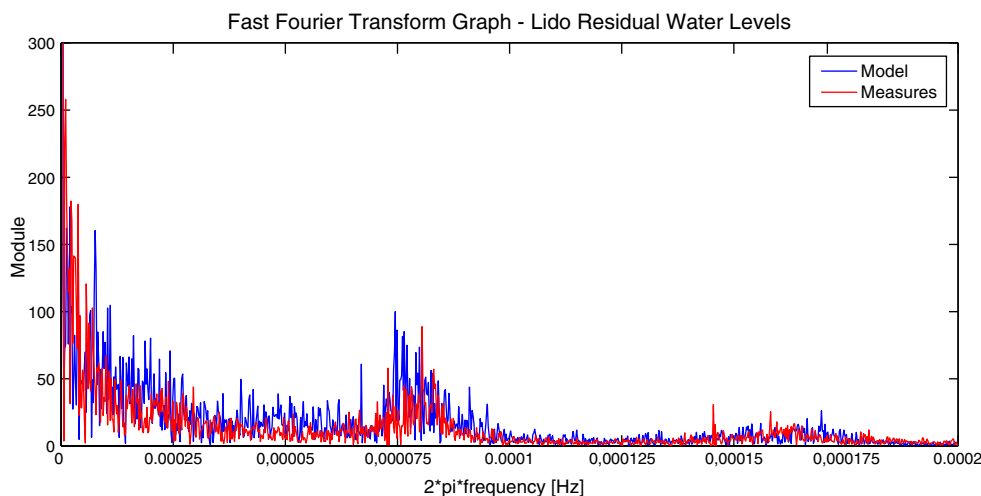
applying a regression formula, the errors given above are certainly a lower bound for the errors of the flux data.

A further source of error for the empirical data is related to the harmonic analysis. In fact, due to the high residual signal generated by intense meteorological events, the amplitude and the phase values obtained for each main tidal constituents from the harmonic analysis are subject to high relative errors.

What concerns the model, a source of error can be related to the reproduction of the SSE outside and inside the inlets. The discharge rate through the inlet is dependent on the SSE gradient along the inlet channel and therefore the accuracy of the SSE inside and outside the lagoon strongly affects the model results. The increase in resolution for what concerns the lagoon, describing in a more realistic way shallow areas and channels, in terms of internal dynamics and bottom friction, can improve the internal reproduction of SSE. Finally, as mentioned in Ferla et al. (2007), the harmonic constants measured in the lagoon have changed significantly during the last decades because of the morphological changes. It is realistic to suppose that the use of new bathymetric data to reproduce the morphology could reduce the error of the model for the SSE and consequently also the discharge rate through the inlets.

3.4 Residual levels and fluxes

An analysis on the reproduced fluxes with respect to the meteorological impact on hydrodynamics has also been carried out. In Fig. 6 the fast Fourier transform (FFT) graph for the residual signal, calculated subtracting the tides from water level at the Lido inlet, is presented. For this final part, the Foreman harmonic constant analysis

Fig. 6 Fast Fourier transform of residual components of levels in the Lido inlet

has been applied to produce the residual water level time series (Foreman 1978).

An interesting aspect is the presence of two peaks, related to the 11 and 22 h periodical signals. These are the two main seiches of the Adriatic Sea (Tomasin and Pirazzoli 1999) and the model is able to reproduce them, better the 22-h seiche than the 11-h one. Comparing model and real data signals, it can be noted that the model, even if it is able to reproduce the 11-h seiche, presents a displacement in frequency, with a shift to higher frequency values (Fig. 6).

Applying the same kind of analysis to the residual fluxes for Lido and Malamocco once more, the two seiches are registered by the model (Fig. 7) which are even more evident. It seems that some semi-diurnal components still remain in the residual signals, perhaps the overtide portion of sea breeze. The model does not reproduce them but this can be explained by the meteorological forcing that was applied coming from ECMWF model data. These data have a frequency of 6 h and it is certainly not possible, with this coarse

temporal and spatial resolution of 0.5°, to produce a signal of limited space and time extension like the sea breeze. Additionally, it has to be said that the meteorological data time step, 6 h, could introduce some artificial frequencies in the model results and that could explain the peak around 3.2E-4 Hz in the model data. These results could be interfaced with other studies done on residual currents at the Venice Lagoon inlets from ADCP data (Gačić et al. 2004). In Gačić et al. (2004), the seiches signals can be isolated and the obtained information is comparable with the one from our study, even if the analysis has been done on axial velocities and not on fluxes as done here.

Considering the whole modeled fluxes for the Lido and Malamocco inlets, it can be noted that the tidal signal contained is 95% of the total variance for the Lido inlet and 97% for the Malamocco inlet, values comparable with the ones obtained by measurements (Gačić et al. 2004). From Fig. 7, it can be seen that the residual variability is due to the meteorological forcings.

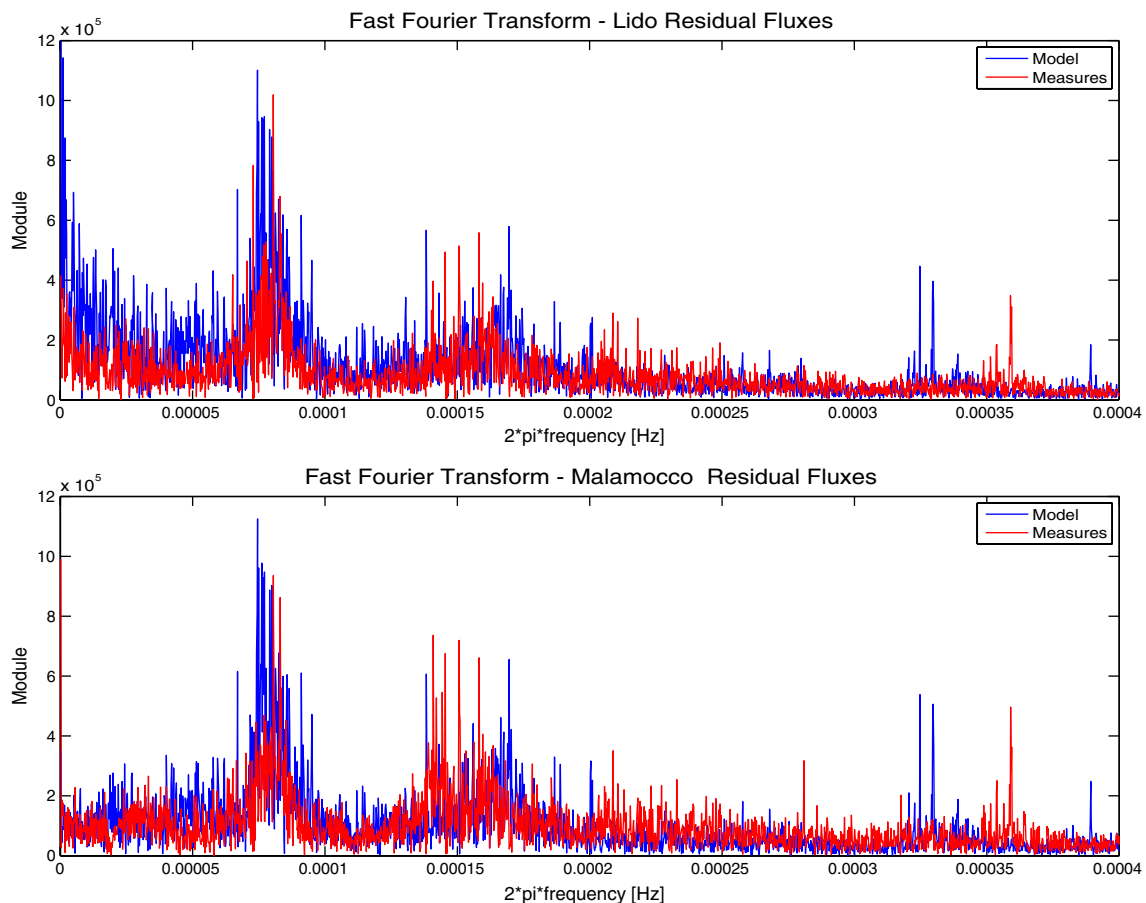


Fig. 7 Fast Fourier transform of residual component of fluxes in the Lido and Malamocco inlets

3.5 Extreme wind events

The whole simulated period has been analyzed and a particular portion, from the 20th of November to the end of December, has been considered. This period shows the two main extreme events of the year, with a strong Scirocco wind blowing around the 24th of November and a long Bora event lasting several days before the 8th of December.

In Fig. 8, the wind data measured by the Oceanographic Platform CNR for the chosen period are shown, both speed and direction.

The Pearson correlation coefficient has been computed, first on the complete chosen time window, and then on the two singular extreme wind events. The significance level chosen is 0.05. High correlation is registered for fluxes, 0.89 for the total temporal range, 0.96 for the Bora event and 0.89 for the Scirocco event. In the three cases the p value is nearly 0.

On the other hand the correlation of water levels is higher for the two single wind events, 0.9 for Scirocco

and 0.94 for Bora, compared with the correlation of the whole time series, 0.82. This could be ascribed to the choice of imposing modeled storm surge forcings at the open boundary instead of the measured ones. In fact, it has to be mentioned that storm surge modeling is a difficult task because of the strong dependence on the choice of meteorological forcings. Anyway, it can be stressed that the correlation computed here is comparable with the one in Bajo et al. (2007).

Because the sub-sample for the Scirocco wind event consists only of around 30 values, it has been decided to compute also the Kendall's Tau, better suited for small samples. The Kendall's Tau gives a value of 0.74 for the water levels (p value 2.2E-9), and 0.84 for the fluxes (p value 1.266E-14). These values prove a good correlation even on this short wind event.

It is interesting to note that total fluxes seem to be reproduced better than total water levels (Fig. 9). This can be explained because fluxes at the inlets are mainly tidally driven (Gačić et al. 2004) and, as shown, the tidal

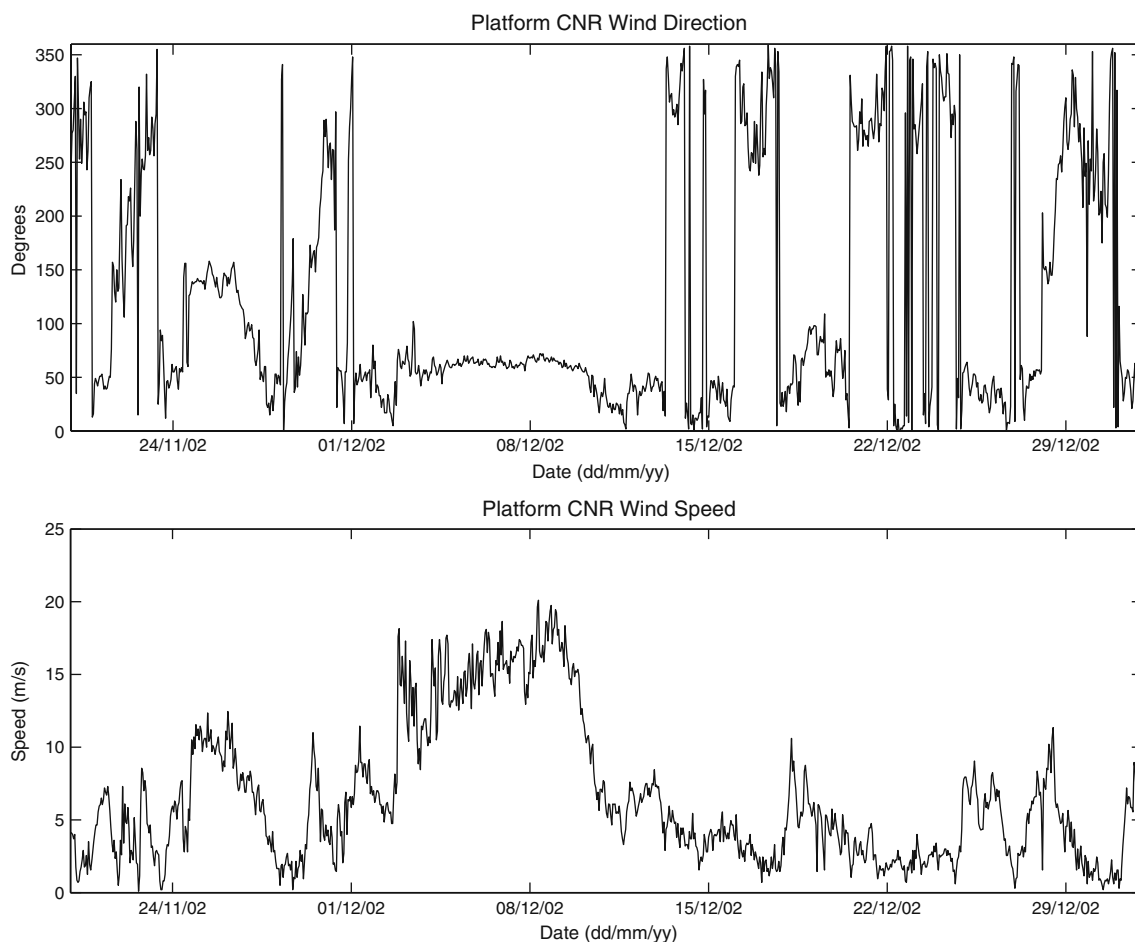


Fig. 8 Wind data (direction and speed) for a significant extreme event period of December 2002

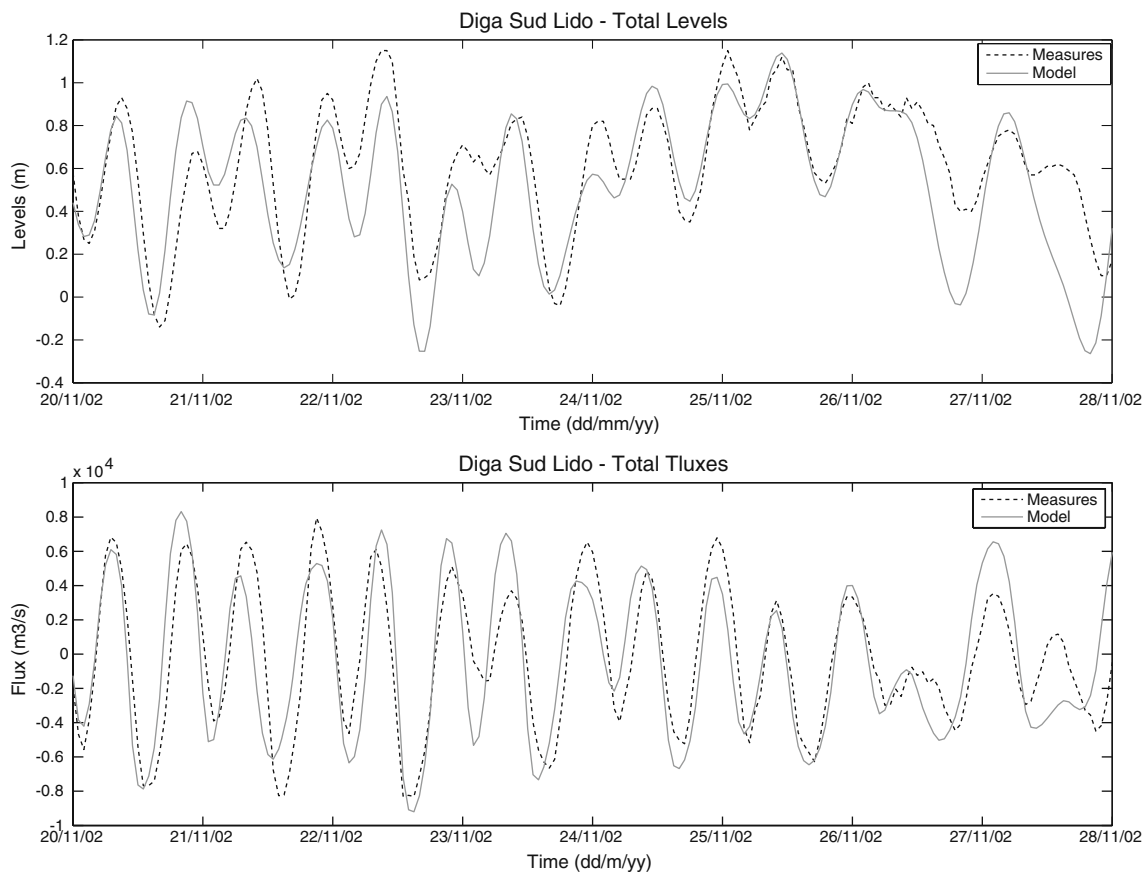


Fig. 9 Comparison between residual levels and fluxes modeled and measured in Lido inlet

signal is better reproduced than the residual one. As an example, Fig. 9 shows the comparison between the measured and the modeled total water levels and fluxes time series, stressing clearly the good match.

Some differences in extreme event reproduction can be also connected with the rough resolution of ECMWF data, 1/2 of a degree, which does not reproduce adequately the wind variability in the study area. A previous work compared results from the SHYFEM model, in its operational version, forced with different wind fields, the ECMWF ones described above and the Limited Area Model (LAMI) wind fields (Zampato et al. 2007). The latter one produced more intense wind fields and was capable of resolving the meso-scale features of the wind in the study area with a higher correlation with measured data (Zampato et al. 2007). Unfortunately, LAMI wind fields were not available for the considered period.

4 Summary and conclusions

The objective of the present work was to reproduce by numerical modeling the water exchange dynamics

through the three lagoon inlets. Before this work, no major studies have been carried out that try to model the exchange mechanisms through the Venice Lagoon inlets. It is also the first work in this area that tries to model the interaction dynamics, running together a model for the lagoon and the Adriatic Sea. The increased spatial resolution achieved in the inlet area permits a realistic study of the interaction processes. We limited the study to barotropic processes and, for this reason, the model has been applied in its 2D version to reproduce the discharge rate through the lagoon inlets driven by the action of the tide and wind (Gačić et al. 2002).

The spatial resolution of the model varies from 30 m inside the Venice Lagoon inlets to 30 km in the middle of the Adriatic Sea, with a transition zone close to the lagoon inlets characterized by a resolution of 300 m. The open boundary has been chosen as a straight line across the Strait of Otranto on which the tidal SSE is imposed. The model has been calibrated to reproduce the tidal propagation in the Adriatic Sea and in the Venice Lagoon. The SSE along the eastern end of the open boundary line of Otranto and the bottom friction coefficient inside the Venice Lagoon

basin were changed during the calibration. For validation, the model results have been compared to measured data of harmonic constants collected in stations located along both the West and the East coast of the Adriatic Sea.

In this article, only the results referred to the tidal propagation in the Adriatic Sea are reported because the tidal dynamics inside the Venice Lagoon and the results of the related calibration process is treated in Umgiesser et al. (2004). For what concerns the tide, the model reproduces the set of harmonic constants for the Diga Sud Lido station for the most energetic tidal constituents (M2, S2, and K1). The results obtained for the other stations close to the Venice Gulf, Falconera, Trieste, and Rovinj, are in good agreement with the empirical data. The SSE offshore the Venice Lagoon is in fact reproduced with an error always lower than 2 cm for the amplitude of the main tides and with a phase difference not higher than 10°. These results are in accordance with the task of this work, which is the reproduction of the tidal flow through the three lagoon inlets.

Finally, the water fluxes through the inlets have been computed by the model when tide and wind are forcing the basin. The results have been compared with empirical data of water discharge derived from ADCP measurements collected inside each inlet. The model reproduces the fluxes through the inlets with a good agreement for both of them and an average overestimation less than 2.9% with a time anticipation of about 20 min. The frequency analysis conducted using the FFT both on residual signals of levels and fluxes stresses the capability of the model to reproduce seiches and its effects on circulation in the interaction areas. Free oscillations are correctly reproduced.

The SHYFEM model has been already applied at the whole Mediterranean basin to forecast, focusing on the Venetian area, water levels (Bajo et al. 2007). The comparison between modeled and measured residual levels time series showed how the coupling of tidal and meteorological forcings, including the surge impact, leads to acceptable results.

The success of these simulations on the reproduction both of the SSE inside and outside the Venice Lagoon and of the tidal flow through the lagoon inlets indicates that the finite element model is performing adequately on the barotropic mode. For the first time, the physical processes that govern the interaction between the two basins are reproduced, as the flux modeling shows. Clearly, the good results reproducing discharges through the inlets are due to the fact that the lagoon is modeled together with the Adriatic Sea. Without the

lagoon, the water levels could be certainly reproduced, but not the fluxes between the sea and the lagoon.

Therefore, this model can be realistically considered as a fundamental support for other studies that aim to investigate the nutrient, salinity, pollutant, and sediment budgets of the lagoon and their exchange with the open sea.

Acknowledgements The authors want to thank the CMCC Project for the financial support. This work has been also partially funded by the CORILA projects 3.2 *Hydrodynamics and morphology of the Venice Lagoon* and 3.5 *Quantity and quality of exchanges between lagoon and sea* which also provided the necessary flux data. The wind dataset of the CNR Platform has been provided by the Venice Municipality. Special thanks to Professor Tomasin for the precious help in dealing with the subject of tides and harmonic analysis.

References

- Bajo M, Zampato L, Umgiesser G, Cucco A, Canestrelli P (2007) A finite element operational model for the storm surge prediction in Venice. *Estuar Coast Shelf Sci* 75:236–249
- Bergamasco A, Carniel S, Pastres R, Pecelik G (1998) A unified approach to the modelling of the Venice Lagoon–Adriatic Sea ecosystem. *Estuar Coast Shelf Sci* 46:483–492
- Bianchi F, Ravagnan E, Acri F, Bernardi-Aubry F, Boldrin A, Camatti E, Cassin D, Turchetto M (2004) Variability and fluxes of hydrology, nutrients and particulate matter between the Venice Lagoon and the Adriatic Sea. Preliminary results (years 2001–2002). *J Mar Syst* 51(v–vi):49–64
- Blumberg AF, Mellor GL (1987) A description of a coastal ocean circulation model. In: Heaps NS (ed) *Three dimensional ocean model*. American Geophysical Union, Washington D.C., pp 1–16
- Cavaleri L, Bertotti L (1996) In search for the correct wind and wave fields in a minor basin. *Mon Weather Rev* 125(8):1964–1975
- Cucco A, Umgiesser G (2005) Modeling the tide induced water exchanges between the Venice Lagoon and the Adriatic Sea. In: *Scientific research and safeguarding of Venice. Proceedings of the annual meeting of the CORILA research program, 2003 results, vol III*. Venice, Italy, pp 385–402
- Cushman-Roisin B, Naimie CE (2002) A 3D-finite element model for the Adriatic tides. *J Mar Syst* 37(4):279–297
- Ferla M, Cordella M, Michielli L, Rusconi A (2007) Long-term variation on the sea level and the tidal regime in the lagoon of Venice. *Estuar Coast Shelf Sci* 75:214–222
- Foreman MGG (1978) Manual for tidal currents analysis and prediction. In: *Pacific marine science report, vol 78, no 6*. Institute of Ocean Sciences, Patricia Bay, p 57
- Foreman MGG (1996) Manual for tidal heights analysis and prediction. *Pac Mar Sci Rep* 77(10):58
- Gačić M, Kovačević V, Mazzoldi A, Paduan JD, Arena F, Mosquera IM, Gelsi G, Arcari G (2002) Measuring water exchange between the Venetian Lagoon and the open sea. *EOS, Trans Am Geophys Union* 83(20):217–222
- Gačić M, Mosquera IM, Kovačević V, Mazzoldi A, Cardin V, Arena F, Gelsi G (2004) Temporal variations of water flow between the Venetian Lagoon and the open sea. *J Mar Syst* 51:33–47

- Goldmann A, Rabagliati R, Sguazzero P (1975) Characteristics of the tidal wave in the lagoon of Venice. Technical Report 47, IBM, Venice Scientific Center
- Melaku Canu D, Umgiesser G, Solidoro C (2001) Short term simulations under winter conditions in the lagoon of Venice: a contribution to the environmental impact assessment of a temporary closure of the inlets. *Ecol Model* 138(1–3):215–230
- Mellor GL (1991) An equation of state for numerical models of ocean and estuaries. *J Atmos Ocean Technol* 8:609–611
- Mosetti F (1987) Distribuzione delle maree nei mari italiani. *Boll Oceanol Teor Appl* 5(1):65–72
- Orlić M, Gačić M, Violette PEL (1992) The currents and circulation of the Adriatic Sea. *Oceanol Acta* 15:109–124
- Polli S (1960) La propagazione delle maree nell'Alto adriatico. Publications of the Istituto Sperimentale Talassografico—Trieste 370
- Polli S (1961) La propagazione delle maree nel golfo di Venezia. Publications of the Istituto Sperimentale Talassografico—Trieste 385
- Tomasin A (2005) The software Polifemo for tidal analysis. Technical Note 202, ISMAR-CNR Institute of Marine Science Venice
- Tomasin A, Pirazzoli PA (1999) The seiches in the Adriatic Sea. *Atti dell'Istituto Veneto di Scienze, Lettere ed Arti CLVII*, pp 299–316
- Umgiesser G (2000) Modeling residual currents in the Venice Lagoon. In: Yanagi T (ed) *Interaction between estuaries, coastal seas and shelf seas*. Terra Scientific Publishing Company (TERRAPUB), Tokyo, pp 107–124
- Umgiesser G, Bergamasco A (1993) A staggered grid finite element model of the Venice Lagoon. In: Morgan K, Offiate E, Periaux J, Peraire J, Zienkiewicz OC (eds) *Finite elements in fluids, proceedings of the VIII international conference, 20–23 September 1993, Barcelona*. Pineridge Press, Barcelona, pp 659–668
- Umgiesser G, Bergamasco A (1995) Outline of a primitive equation finite element model. In: *Rapporto e Studi*, vol XII. Istituto Veneto di Scienze, Lettere ed Arti, Venice, Italy, pp 291–320
- Umgiesser G, Bergamasco A (1998) The spreading of the Po plume and the Italian coastal current. In: Dronkers J, Scheffers M (eds) *Physics of estuaries and coastal seas*. Rotterdam, pp 267–275
- Umgiesser G, Canu DM, Cucco A, Solidoro C (2004) A finite element model for the Venice Lagoon. *Development, set-up, calibration and validation*. *J Mar Syst* 51: 123–145
- Zampato L, Canestrelli P, Tomasin A, Zecchetto S (2005) A comparison between modelled and observed atmospheric fields over the Adriatic and Tyrrhenian seas. Technical Report 263, ISMAR-CNR, Venezia, Italy
- Zampato L, Umgiesser G, Zecchetto S (2007) Sea level forecasting in Venice through high resolution meteorological fields. *Estuar Coast Shelf Sci* 75:223–235

Two-way embedding algorithms: a review

Submitted to *Ocean Dynamics: Special Issue on Multi-Scale Modelling: Nested Grid and Unstructured Mesh Approaches*

Laurent Debreu · Eric Blayo

Received: 21 April 2008 / Accepted: 11 September 2008 / Published online: 17 October 2008
© Springer-Verlag 2008

Abstract Local mesh refinement features have now been added to a number of numerical ocean models. In its crudest form, a high-resolution grid is embedded (or nested) in a coarse-resolution grid, which covers the entire domain, and the two grids interact. The aim of this paper is to review existing two-way grid embedding algorithms. The basic algorithms and specificities related to ocean modelling are first described. Then, we address several important issues: conservation properties, design of interpolation/restriction operators, and noise control techniques.

Keywords Two-way embedding · Mesh refinement · Structured grids

1 Introduction

An increase of the horizontal resolution of a numerical ocean model still remains a key point in the improvement of the realism of its solutions, mainly through a better representation of small scales and domain geometry. This increase of resolution is generally not performed everywhere in the domain, both because it is not necessary from a physical point of view and be-

cause of limited computational resources. That is why a number of ocean models include local mesh refinement features. The idea is to refine the mesh where (and potentially when) necessary according to the objective of the simulation.

For models based on a structured grid, a possible way to locally increase the resolution is to use a grid with variable resolution. Starting from a uniform grid, a mapping is introduced to produce a grid with increased resolution in areas of interest (Zhuo and Qingcun 1995). This approach has the advantage that the model is still written on a single grid and, thus, there is almost no additional coding complexity, and properties like conservation are easily handled. The disadvantages lie mainly in the difficulty of generating grids with good numerical properties and in the fact that the time step constraint relies on the smallest grid cell (in the case of explicit time stepping schemes).

The second approach, which is the subject of the present paper, is to embed a high-resolution (HR) version of the same model at a specific location. The information is exchanged in two ways between the coarse-resolution (CR) solution and the HR solution: the coarse grid provides the boundary conditions of the child grid, while the HR solution is used to update the coarse grid solution in the common domain. The major interest of this approach is that the coarse grid provides accurate boundary conditions for incoming information along the HR grid boundaries, while the update (or feedback) step enables outgoing information to leave the fine grid much better than in one-way interaction (i.e., without feedback from the HR to the CR solution). In this approach, even if the same model is applied, numerical schemes and/or physical parameterizations can be adapted to the grid resolution.

Responsible Editor: Pierre Lermusiaux

L. Debreu (✉)
INRIA, Laboratoire Jean Kuntzmann, 51 rue des
Mathématiques, 38400 Saint Martin d'Hères, France
e-mail: Laurent.Debreu@inria.fr

E. Blayo
Joseph Fourier University, Laboratoire Jean Kuntzmann,
Grenoble, France

We will discuss here only refinement of the mesh along the horizontal directions. A few applications of vertical mesh refinement have also been reported. In Fox and Maskell (1995), the vertical refinement is applied to the whole water column. For similar reasons to the nondecomposition of the vertical direction in parallel applications (use of implicit schemes and particular treatment of the barotropic mode), local mesh refinement on the vertical introduces additional difficulties, some of them very close to the ones reported in this paper for horizontal refinement.

Concerning horizontal refinement, we will focus here on several important aspects, including conservation, intergrid transfer operators, and noise control techniques. We begin in Section 2 with a description of the basic algorithms underlying the two-way mesh refinement method and then give an overview of difficulties that can arise in practical numerical simulations. Conservation issues are addressed in Section 3. In Section 4, intergrid transfer operators (interpolations and updates) are examined and important details of implementations are given. Noise control techniques, which aim at making the algorithm more robust, are presented in Section 5.

2 Two-way nesting: algorithms

In this paper, we will consider only one HR grid embedded in the CR grid. Extension to more than one grid and/or more than one level of refinement does not present additional complexities.

So, let us consider Ω as the domain covered by the coarse grid and a subdomain ω , covered by the fine grid. γ is the boundary of ω . At the discrete level, the numerical grids are denoted by Ω_H and ω_h (cf. Fig. 1). ω_H is the part of the grid Ω_H corresponding to the domain ω .

2.1 Basic algorithm

The ratio between the coarse and fine horizontal mesh sizes is an integer ρ , the mesh refinement factor ($\rho = 2$ in Fig. 1). Typical values of ρ in actual applications range from 2 to 5. Using larger values introduces too strong a change in the resolved scales on the different grids. In that case, it may not be appropriate to try to enforce a strong coupling between the two grids, and methods with weaker interactions than the ones we will discuss in this paper could be of interest (Sheng et al. 2005).

Choosing an odd mesh refinement factor simplifies grid interactions since, in that case, a coarse grid point

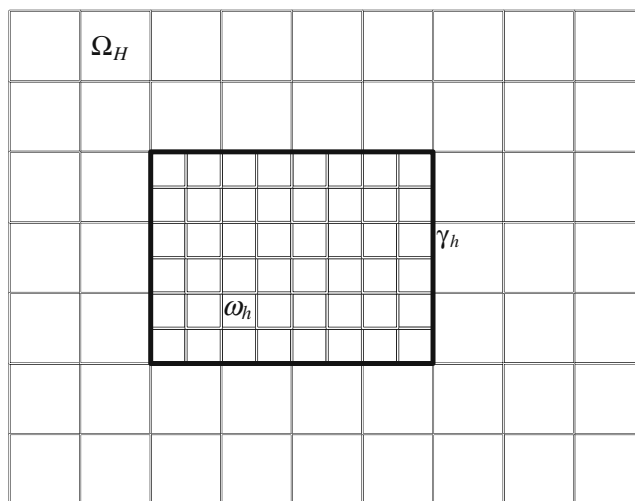


Fig. 1 Local mesh refinement of a structured grid with a mesh refinement factor of 2. The HR grid ω_h is embedded in the CR grid Ω_H

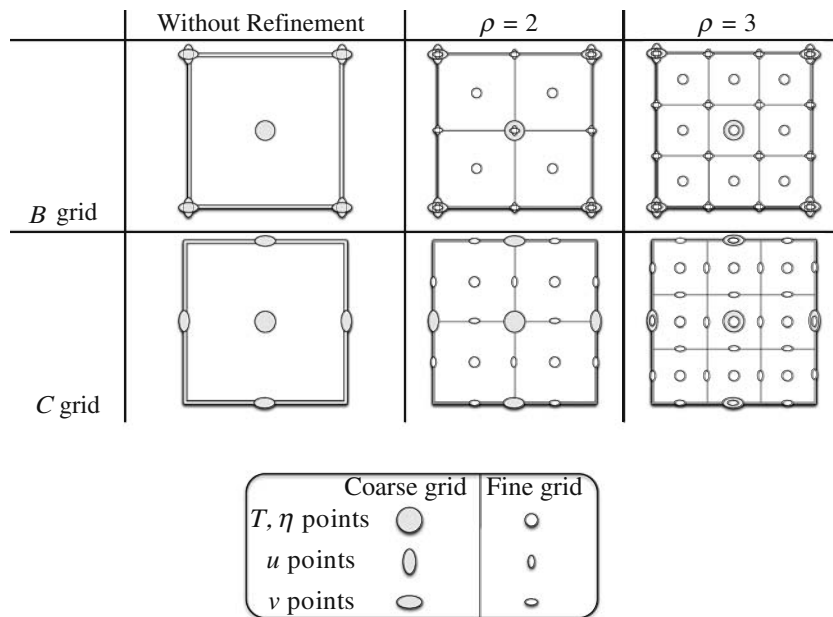
always has one underlying HR point. Figure 2 shows, for B and C grids (using the Arakawa classification), that if a variable is staggered (cell- or face-centered) and if the mesh refinement factor is even, then a coarse grid point does not have a corresponding point on the fine grid.

Associated with this spatial mesh refinement, time refinement may also be applied. Most of the ocean models now use explicit time integration algorithms both for computational reasons (efficiency on distributed memory parallel computers) and for accuracy reasons (smaller dispersion errors than implicit schemes). The model is then subject to a given Courant–Friedrichs–Lewy stability condition, and the ratio $\Delta t/\Delta x$ must be kept smaller than a given value on the whole grid hierarchy. As we shall see, refining in time induces several difficulties but is required for some applications, especially when only a small portion of the coarse grid domain is refined or when there are several levels of refinement. In these cases, keeping the same time step over the whole grid hierarchy greatly increases the computational cost. The integration algorithm for a time refinement factor ρ_t of 3 is depicted in Fig. 3.

The model is first integrated on the coarse grid with a time step equal to Δt_c , and then, the fine grid is integrated ρ_t times with $\Delta t_f = \Delta t_c/\rho_t$. Interpolations of coarse grid boundary data occur at the end of each fine grid time step and updates occur at the end of the coarse time step, when the two solutions have been advanced to the same physical time.

Let \mathcal{L} represent the model integration from time t^n to time t^{n+1} and $\mathcal{L}_c, \mathcal{L}_f$ its discretizations on the

Fig. 2 Refinement of a cell with even ($\rho = 2$) and odd ($\rho = 3$) refinement factors for a *B* and a *C* grid. Only one coarse cell is shown. When $\rho = 3$, a coarse grid point always has a corresponding fine grid point



CR and HR grids. Let P denote the interpolation (or prolongation) operator from Ω_H to γ_h , the boundary of ω_h , and R the update (or restriction) operator from ω_h to ω_H . Then, assuming that the model is fully explicit, the algorithm can be written in the following simplified form:

1. $u_c^{n+1} = \mathcal{L}_c(u_c^n)$
2. For $m = 1 \dots \rho_t$ do
 - $u_f^{n+\frac{m}{\rho_t}} = \mathcal{L}_f\left(u_f^{n+\frac{(m-1)}{\rho_t}}\right)$
 - $u_f^{n+\frac{m}{\rho_t}}|_{\gamma_h} = P(u_c^n, u_c^{n+1})$
3. $u_c^{n+1}|_{\omega_H} = R(u_f^{n+1})$

P , the interpolation operator, includes a time interpolation. This time interpolation is done linearly in most of

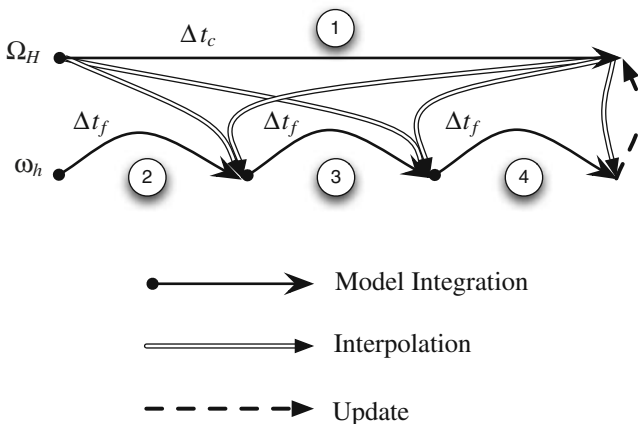


Fig. 3 Integration algorithm for a time refinement factor of 3

the applications and, thus, requires only the knowledge of u_c at time t^n and t^{n+1} . To our knowledge, the effect of increasing the interpolation order in time has not been studied in the literature. This is because, as for numerical models on a uniform grid, the error relative to time is expected to be lower than the one relative to spatial interpolation. However, as for numerical models, it may not be the case anymore due to the increase of horizontal resolution and to the use of higher-order spatial schemes. New experiments are probably required to test the validity of the linear interpolation. Note, finally, that interpolation and restrictions can also be applied to time tendencies instead of instantaneous values (Jones 1977; Spall and Holland 1991; Oey and Chen 1992).

2.2 Complexities associated with time stepping

The numerical schemes are seldom as simple as previously described, and this can lead to several complications. In a numerical ocean model, some of them are linked to the particular treatment of the gravity waves associated with the evolution of the barotropic mode. This treatment leads either to the use of an implicit time stepping, thus leading to the solution of an elliptic system, or to the use of a time splitting method.

2.2.1 Elliptic systems

Elliptic systems can come from a 2D implicit solve (e.g., implicit treatment of free surface (Dukowicz and Smith 1994) or rigid lid approximation). A 3D system also has

to be solved if the hydrostatic assumption is removed. If we denote the associated linear system by:

$$\mathcal{A}v = \mathcal{B},$$

then, after discretization, on the nested grid, the simplest choice is to force the HR solution with an interpolation of the coarse grid solution:

$$a) \mathcal{A}_c v_c = \mathcal{B}_c \text{ on } \Omega_H, \quad b) \begin{cases} \mathcal{A}_f v_f = \mathcal{B}_f \text{ on } \omega_h \\ v_{f|\gamma_h} = P v_c \end{cases} \quad (1)$$

This implies continuity of the solution at the interface γ_h , but the resulting gradient across the boundary is discontinuous. Additionally, a loss of accuracy is introduced because errors produced by the coarse grid resolution (a) propagate inside the HR domain through the boundary forcing in step b).

A first step toward better accuracy is to update the right-hand side of the coarse grid equation so that Eq. 1 is replaced by Eq. 2:

$$a) \mathcal{A}_c v_c = \begin{cases} R\mathcal{B}_f \text{ in } \omega_H \\ \mathcal{B}_c \text{ in } \Omega_H \setminus \omega_H \end{cases}, \quad b) \begin{cases} \mathcal{A}_f v_f = \mathcal{B}_f \text{ on } \omega_h \\ v_{f|\gamma_h} = P v_c \end{cases} \quad (2)$$

In Spall and Holland (1991), this technique is applied to a rigid lid ocean model. The linear system corresponds to the computation of the streamfunction tendency and the right-hand side is the barotropic vorticity tendency. The barotropic vorticity tendency on the coarse grid is updated by the one computed on the fine grid.

However, despite this improvement, the gradient of the solution across the interface is still discontinuous. To be properly solved, the two-grid system should be handled as a truly local multilevel system. This can be done using local defect correction methods (see Laugier et al. 1996 for an application to a rigid lid model and a comparison with results obtained in Spall and Holland 1991). The system can also be seen as discretized on a single composite grid. In this approach, the operator \mathcal{A} is modified at the boundary and involves both coarse- and fine-grid variables. The resulting system is:

$$\begin{pmatrix} \mathcal{A}_c & 0 \\ 0 & \mathcal{A}_f \\ \mathcal{A}_{c\gamma} & \mathcal{A}_{f\gamma} \end{pmatrix} \begin{pmatrix} v_{c|\Omega_H \setminus \omega_H} \\ v_f \end{pmatrix} = \begin{cases} \mathcal{B}_c \text{ in } \Omega_H \setminus \omega_H \\ \mathcal{B}_f \text{ on } \omega_h \\ \mathcal{B}_\gamma \text{ in } \gamma_h \end{cases} \quad (3)$$

$\mathcal{A}_{c\gamma} v_{c|\Omega_H \setminus \omega_H} + \mathcal{A}_{f\gamma} v_f = \mathcal{B}_\gamma$ is an approximation of the original system on the interface, which leads at convergence to continuity of both the solution and its gradient across the boundary. See Martin and Cartwright (1996) for applications of the local multilevel method to the solution of a Poisson equation.

Remark

- Algorithms 2 and 3 are more difficult to apply when time refinement is used. Indeed, the right-hand side \mathcal{B}_c and \mathcal{B}_f should be computed at the same time, and the way to write the system at intermediate time steps on the fine grids is not obvious.
- Since we are using two-way techniques, the right-hand side \mathcal{B}_c , which depends on coarse grid variables that have been updated, is never significantly different from $R\mathcal{B}_f$. In some cases, the update schemes used for the primary variables are precisely chosen in such a way that $\mathcal{B}_c = R\mathcal{B}_f$ in ω_H . However, this leads most of the time to the use of very simple restriction operators (e.g., average, see Clark and Farley 1984) that do not have good filtering properties and, thus, can affect the quality of the solution (see Section 4.1).

2.2.2 Free surface and time splitting

Similar problems are also inherent when a time-splitting is used for the treatment of the external mode. A number of models now use the time-splitting approach of Blumberg and Mellor (1987) and Killworth et al. (1991). The barotropic time step is set to a fraction of the baroclinic time step. The barotropic mode is then integrated as a solution of a shallow water model forced by the vertical integral of the right-hand side of the 3D equations. At the end of this integration, the newly computed barotropic quantities are used to replace the barotropic part of the 3D fields. In order to prevent aliasing errors, two approaches can be taken. The first one, as in Killworth et al. (1991), is to use in the barotropic time steps a time integration scheme that sufficiently damps high temporal frequencies (e.g., Euler Backward). In this case, the fine and coarse grids can interact during the small barotropic time steps throughout the barotropic quantities (see Oey and Chen 1992). This is illustrated in Fig. 4 with a time refinement factor of 2.

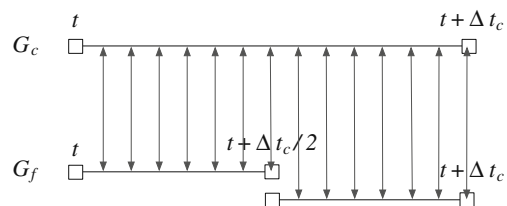


Fig. 4 Time splitting algorithm with a time refinement ratio equal to 2. Intermediate times (due to time refinement) on the fine grid are not shown. G_c is the coarse grid and G_f is the fine grid

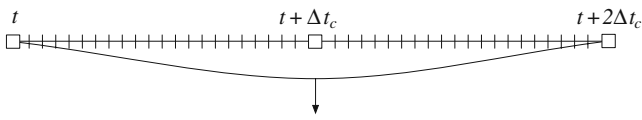


Fig. 5 Time splitting algorithm with filtering using flat weights over $[t, t + 2\Delta t_c]$. Intermediate instantaneous values are used to compute a filtered value at time $t + \Delta t_c$

The second approach consists in applying a filter to the barotropic quantities to filter out the high temporal frequencies. The integration extends to a time larger than $t + \Delta t_c$, and a filtering formula, the result of which is centered at time $t + \Delta t_c$, is applied. For example, as shown in Fig. 5, one may extend the integration period to $t + 2\Delta t_c$ and use flat weights over $[t, t + 2\Delta t_c]$ (constant weights equal to $1/N$, where N is the number of barotropic time steps) to compute the average at time $t + \Delta t_c$. When filtering is applied, it is relatively easy to interact between grids when no time refinement is done, as in Fig. 6 (Barth et al. 2005).

However, when both filtering and time refinement are used (cf. Fig. 7), a new problem arises, due to the fact that some coarse grid points do not have fine grid points located at the same time. For this reason, the coupling between coarse and fine grids cannot be done anymore through the instantaneous barotropic quantities.

In that case, a simple choice is to perform the coupling at the baroclinic level (Penven 2006), i.e., the exchange occurs only between filtered quantities. However, this is like having the system in a one-way mode for the external mode and, as for implicit solvers, errors produced on the coarse grid can then propagate inside the HR grid. One way to perform the coupling at the barotropic level is to exchange information between another set of variables corresponding to intermediate quantities (Debreu et al. 2008a). Using intermediate weights, quantities centered at times between t and $t + \Delta t_c$ are computed so that coupling as without filtering (Fig. 4) is permitted.

Partial conclusion As we have seen, besides the simple algorithm, several additional complexities arise

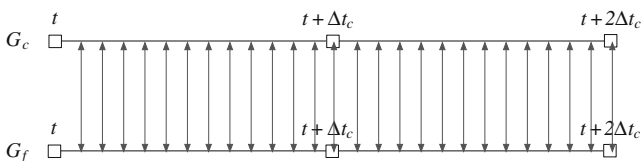


Fig. 6 Time splitting algorithm without time refinement and with filtering

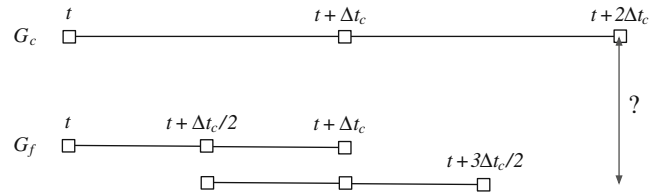


Fig. 7 Time splitting algorithm with time refinement equal to 2 and with filtering

when implicit schemes or time splitting schemes are used. Moreover, it is clear that, in both cases, use of time refinement leads to additional difficulties. Looking at the problem as a truly multilevel system in the case of elliptic systems and doing the coupling at the barotropic level in the case of a time splitting scheme are probably the best choices. Additionally, if more simple techniques are used, then it brings strong constraints on the update operators in order to maintain the coarse-grid right-hand side in agreement with the fine-grid right-hand side. Most of the time, these constraints lead to the use of very simple update operators that do not have good numerical properties (see Section 4.1).

3 Conservation

On a uniform grid, conservation is guaranteed when internal numerical schemes are written in flux form. Concerning two-way embedded models, it is not so common that the resulting system is conservative. Ensuring conservation leads to several computational issues and imposes strong requirements on intergrid transfer operators that can lead to a loss of accuracy. However, it is recommended for long-term integration. After reviewing the basic requirements for conservation on an embedded grid, two familiar approaches for the preservation of conservation properties are described.

3.1 Definition and discretization

Let us consider a two-dimensional domain and q , a solution of the following equation written in conservative form:

$$\frac{\partial q}{\partial t} + \frac{\partial f}{\partial x} + \frac{\partial g}{\partial y} = 0,$$

where f and g may contain both advective and diffusive fluxes.

Then, assuming that fluxes f and g cancel on the boundaries of Ω , Q_Ω , integral of q over the domain Ω , is constant in time (at the continuous level).

$$Q_\Omega(t) = \int_\Omega q(x, y, t) \, dx \, dy \Rightarrow \frac{dQ_\Omega(t)}{dt} = \int_{\partial\Omega} f \, ds + \int_{\partial\Omega} g \, ds = 0$$

In the nested grid system, the quantity Q_Ω is defined by the summation over the HR domain ω and its complement in Ω :

$$Q_\Omega = Q_\omega + Q_{\Omega \setminus \omega}$$

Now, let us make the following assumptions for the sake of simplicity:

- As shown in Fig. 8, we consider a two-dimensional domain infinite in both x and y directions. The left (respectively, right) part of the domain is at coarse (respectively, high) resolution.
- The variable $q_{i,j}^n$ is cell-centered.
- The time stepping scheme is an explicit Euler scheme.

$$q_{i,j}^{n+1} = q_{i,j}^n - \frac{\Delta t}{\Delta x \Delta y} (F_{i,j} - F_{i-1,j}) - \frac{\Delta t}{\Delta x \Delta y} (G_{i,j} - G_{i,j-1}) \tag{4}$$

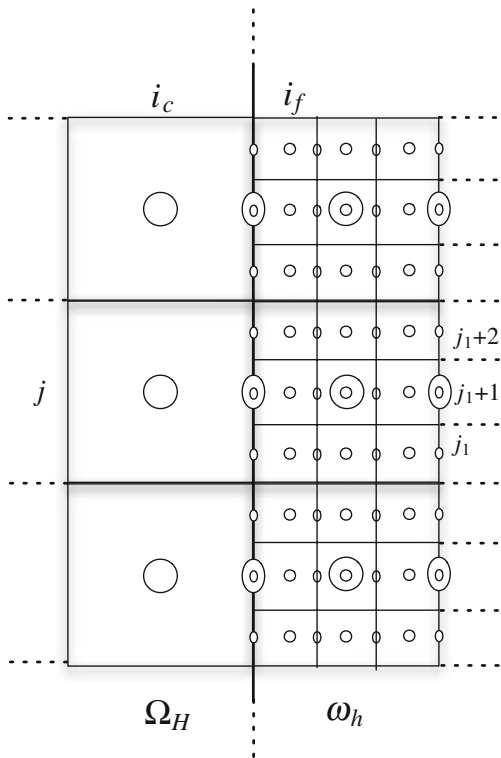


Fig. 8 The CR domain Ω_H on the left and the HR domain ω_h on the right for a C-grid with a mesh refinement factor of 3

where $F_{i,j}, G_{i,j}$ are volumetric fluxes: $F_{i,j} = f_{i,j} \Delta y$, $G_{i,j} = g_{i,j} \Delta x$.

At discrete level, Q_Ω at time $t = t^n$ is given by

$$Q_\Omega^n = \sum_{i \leq i_c, j} \Delta x_c \Delta y_c q_{i,j}^{c,n} + \sum_{i \geq i_f, j} \Delta x_f \Delta y_f q_{i,j}^{f,n}$$

where, as shown in Fig. 8, i_c and i_f denote the first coarse and fine indices close to the interface.

Then after one time step, according to Eq. 4, we obtain

$$Q_\Omega^{n+1} = Q_\Omega^n - \Delta t_c \sum_j F_{i_c, j}^{c,n} + \Delta t_f \sum_{p=0}^{\rho_t-1} \sum_{j_f=j_1}^{j_1+\rho-1} F_{i_f-1, j_f}^{n+p/\rho_t} = Q_\Omega^n - \Delta t_c \sum_j \left(F_{i_c, j}^n - \frac{1}{\rho_t} \sum_{p=0}^{\rho_t-1} \sum_{j_f=j_1}^{j_1+\rho-1} F_{i_f-1, j_f}^{n+p/\rho_t} \right) \tag{5}$$

In general, there will be a misfit between the coarse and fine grid fluxes so that the flux differences of the right-hand side do not cancel and conservation is artificially lost: $Q_\Omega^{n+1} \neq Q_\Omega^n$.

Conservation can be achieved if the fine grid fluxes derive from a conservative interpolation (in time and space) of the coarse grid flux so that the following relation holds:

$$\frac{1}{\rho_t} \sum_{p=0}^{\rho_t-1} \sum_{j_f=j_1-\rho/2}^{j_1+\rho/2} F_{i_f-1, j_f}^{n+p/\rho_t} = F_{i_c, j}^n \tag{6}$$

This approach is applied in several ocean models for volume conservation (q being the free surface and F, G being the vertically integrated transport in x and y directions). It is indeed sufficient to produce a conservative interpolation of the transport for the global volume to be conserved. A number of classical conservative interpolation schemes can then be used (see Section 4.2).

Let us consider more generally the computation of a boundary flux on a C-grid as indicated in Fig. 9 and let us suppose that we are trying to compute the flux of a tracer on the boundary so that:

$$f = uT$$

The computation of fine grid flux on the boundary involves the result of the interpolation of the velocity u on the boundary (grey ellipses in Fig. 9). This interpolation has to be conservative in order not to create an artificial loss of conservation. The fine grid flux also

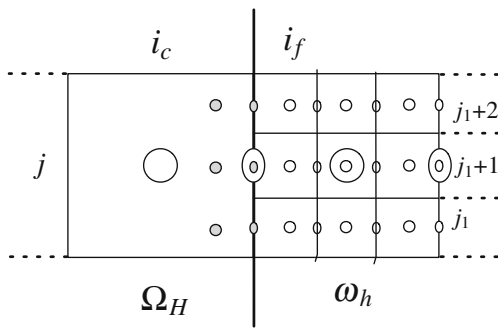


Fig. 9 Computation of a boundary flux on a C-grid

involves the computation of the tracer’s value on the interface that makes use of both interpolated values (grey circles in Fig. 9) and internal values in the fine-grid domain. Thus, we may suppose that the resulting fine-grid fluxes are actually more accurate than the coarse-grid ones, so that trying to enforce Eq. 6 is not the best choice. Two other approaches to conservation, that share several similarities, are now introduced: the flux correction algorithm and the Kurihara method.

3.2 Flux correction algorithm

The flux correction algorithm comes from the adaptive mesh refinement community. It follows from the algorithm of Berger and Olinger (1984) and Berger and Colella (1989). The idea is to apply a modification of the coarse grid variables that take into account the misfit between coarse and fine grid fluxes.

Starting from Eq. 5, a correction is applied to the coarse grid variable at time $n + 1$ near the boundary as follows:

$$q_{i_c,j}^{n+1,*} = q_{i_c,j}^{n+1} + \frac{\Delta t_c}{\Delta x_c \Delta y_c} \left(F_{i_c,j}^n - \frac{1}{\rho_t} \sum_{p=0}^{\rho_t-1} \sum_{j_f=j-\rho/2}^{j+\rho/2} F_{i_f-1,j_f}^{n+p/\rho_t} \right) \tag{7}$$

Basically, it means that the coarse grid variable has been integrated using, on the right interface, fluxes computed by the fine grid:

$$q_{i_c,j}^{n+1,*} = q_{i_c,j}^n - \frac{\Delta t_c}{\Delta x_c \Delta y_c} \left(\frac{1}{\rho_t} \sum_{p=0}^{\rho_t-1} \sum_{j_f=j-\rho/2}^{j+\rho/2} F_{i_f-1,j_f}^{n+p/\rho_t} - F_{i_c-1,j} \right) \tag{8}$$

The algorithm is easy to implement assuming that it is possible to write the time evolution of q in terms of flux divergences, as it was indeed the case for the Euler time scheme. The program stores the fine-grid fluxes at the

boundary and makes a summation in time and in space over the fine-grid cells.

With other time-stepping schemes, e.g., leap-frog associated to an Asselin filter, this is not achievable (even without the Asselin filter, it is more complex since two arrays of fluxes should be stored, see Herrnstein et al. 2005).

When using this approach, care must be taken concerning potential stability issues. Indeed, if the original fluxes are computed in a centered way, the resulting scheme is always biased and instabilities can occur (Olsson and Petersson 1996; Debreu et al. 2008a). One remedy is to compute the interfacial fluxes using information along characteristics of the flow (see Part-Enander and Sjogreen 1994; Blayo and Debreu 2005 for an interpretation of usual open-boundary conditions in terms of characteristic variables). This is actually the case in the adaptive mesh refinement community where, due to the hyperbolic nature of the problems, the internal numerical schemes already make use of decompositions into characteristic variables.

Moreover, there is an associated loss of accuracy at the boundary. Indeed, it can be shown that if the fluxes are computed with a second-order accurate approximation, use of Eq. 8 leads to a first-order only approximation of the original equation (Debreu et al. 2008a). These two remarks also apply to the Kurihara method that we now describe.

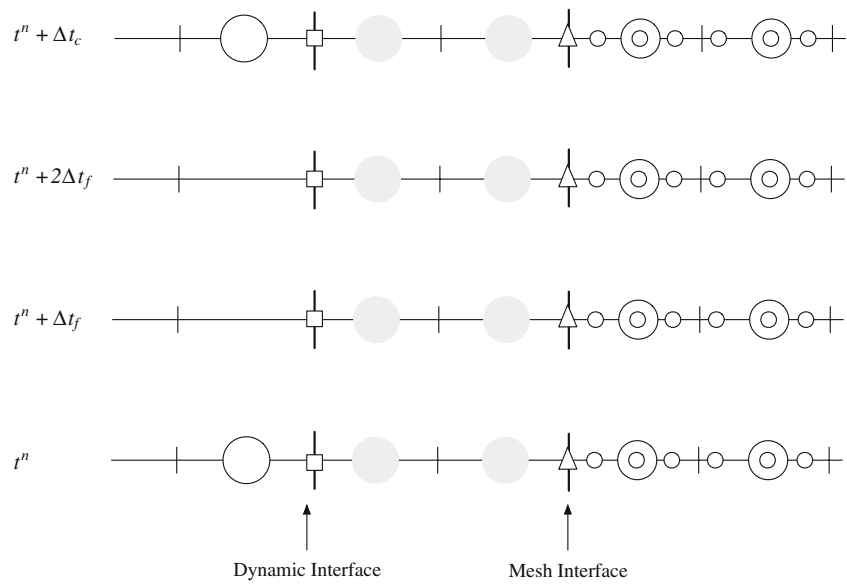
3.3 Kurihara method

The Kurihara method (Kurihara et al. 1979) has been used by several authors to enforce conservation (Kurihara and Bender 1980; Ginis et al. 1998). It is important to note that it applies to a model with all variables cell-centered. To understand the method, it is first necessary to explain that the scheme makes use of an intermediate area, composed of two coarse-grid cells (grey circles in Fig. 10), where the space resolution is the one of the coarse grid and the time resolution is the one of the fine grid.

In this method, conservation is achieved thanks to two principles:

- The coarse grid computes its fluxes, which are then interpolated in time in a conservative way at \square points to force the separation area.
- A spatial linear interpolation of variables is performed at Δ points and the computed flux is used both for integrating the coarse-grid cell on the left and the fine-grid cell on the right. In two dimensions, a summation is applied along the y axis just as for the flux correction approach.

Fig. 10 The three zones of the Kurihara method: on the left, the coarse grid, on the right, the fine grid. In between, an area composed of two coarse-grid cells with the space resolution of the coarse grid and the time resolution of the fine grid



Thus, the resulting scheme is conservative. The problems of loss of conservation due to time refinement and to space refinement have been treated separately. It is clear that, if there is no time refinement, then the flux correction approach is equivalent to the Kurihara method if the boundary variables used in the flux computation have also been linearly interpolated.

As already said, the original method is designed for a nonstaggered grid with all variables located at the cell's center. Similar to the flux-correction approach, the Kurihara method is more difficult to implement for non-cell-centered variables (for example, in order to preserve conservation of momentum on a C-grid, cf. Sobel 1976).

Partial conclusion We have presented two different approaches for enforcing conservation. Due to the presence of the transition area, the Kurihara method is more difficult to implement than the flux-correction approach. Additionally, although possible, it is more difficult to implement such methods for quantities not defined at the centers of the cells. The flux correction approach is easier to implement; however, care about potential stability issues must be taken, especially when the original model computes its fluxes using second-order centered differences.

When conservation properties are not preserved in a two-way simulation, it is important, especially for long-term integration, to have an idea of the artificial loss of conservation. That is why the misfit between coarse and fine grid fluxes on the interface should always be diagnosed.

4 Intergrid transfer operators

For very simple equations (e.g., 1D advection-diffusion), the properties of the two-grid methods can be analyzed with simple techniques related to the study of numerical schemes on nonuniform grids (Harrison and Elsberry 1972; Berger 1985; Olsson and Petersson 1996). In the linear case, it can even be extended to more complex systems of equations using matrix stability analysis (Heggelund and Bernsten 2002). However, in the more general case, the design of intergrid transfer operators must rely on more crude considerations.

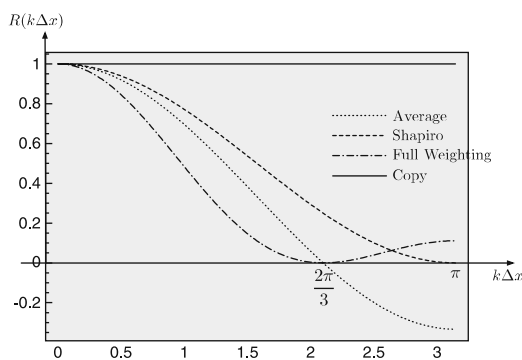
4.1 Update schemes

Obviously, the restriction operator has a crucial role in two-way nesting algorithms. If we just think in terms of resolved scales, the two following properties should hold:

1. The transfer of information should be maximum for scales well resolved on the coarse grid. Otherwise, the order of the approximation will be lowered.
2. Small scales should be strongly filtered. Otherwise, by aliasing, noise will be produced on the coarse grid.

In several applications, authors have made the choice to use the average operator for the feedback step: the coarse grid value in a cell is replaced by the area-weighted sum of the fine grid values in the same cell:

$$u^c = \frac{1}{\Delta x^c \Delta y^c} \sum_{i,j} \Delta x^f \Delta y^f u^f$$



Copy	$u^c = u^f$
Average	$u^c = \frac{1}{3} (u_{j-1}^f + u_j^f + u_{j+1}^f)$
Shapiro	$u^c = \frac{1}{4} (u_{j-1}^f + 2u_j^f + u_{j+1}^f)$
Full Weighting	$u^c = \frac{1}{9} (u_{j-2}^f + 2u_{j-1}^f + 3u_j^f + 2u_{j+1}^f + u_{j+2}^f)$

Fig. 11 Transfer functions for usual one-dimensional restriction operators with a mesh refinement factor $\rho = 3$ for typical filters (k is the wavenumber). The corresponding formulas are also given

Note that this is often seen as a requirement for conservation but, as has been previously seen, it is not.

In an adaptive mesh refinement context, the restriction is still made with this average formula because the grid can move from one time step to another, so that an average restriction is necessary to globally maintain the conservation before and after the regriding step (periodic modification of the grid hierarchy following a refinement criterion). Note that this was also the main reason for using the average restriction operator in the Kurihara paper (Kurihara et al. 1979), where the methodology is applied to moving mesh methods. Other familiar restriction schemes are written in Fig. 11, along with their corresponding transfer functions.

Recently, in Debreu et al. (2008a), it was shown that the use of higher-order restriction operators, using a larger stencil (full weighted schemes), can lead to strong improvements. Even if the full-weighted operator damps the well resolved scales a little more, it also produces a strong damping of the small scales, which is required to inhibit aliasing and noise on the coarse grid.

4.2 Interpolation

Numerous interpolation operators have been used in two-way nesting methods. See Koch and McQueen (1987) and Zhang et al. (1986) for the details in several models. The basic idea is that the order of the

interpolation order should be high enough not to lower the global order of approximation and, as usual when using high-order interpolation schemes, care must be taken about the potential oscillations produced by the schemes (this problem is well known, for example, when using Lagrange interpolation operators). As previously seen, it is also advantageous to have conservative interpolations (in particular, for the interpolation of normal velocities on the boundaries). To fulfill these two conditions, several methods can be used: parabolic conservative interpolation or approximation Clark and Farley (1984), parabolic interpolation with minimization of second-order derivative Barth et al. (2005), piecewise parabolic method (PPM), weighted essentially nonoscillatory (WENO) method, advection equivalent interpolation schemes Alapaty et al. (1998).

Several authors also mention the fact that interpolation and restriction operators must be devised in agreement with one another. This is mainly justified when implicit solvers are used. The main reason is that, after feedback, the right-hand side of the elliptic equation, computed on the coarse grid, must be in agreement with the one of the fine grid (Clark and Farley 1984). In that case, assuming a conservative interpolation has been used, the average restriction operator is the most natural. Let us take the example of the free surface equation written under the following form (Eq. 9):

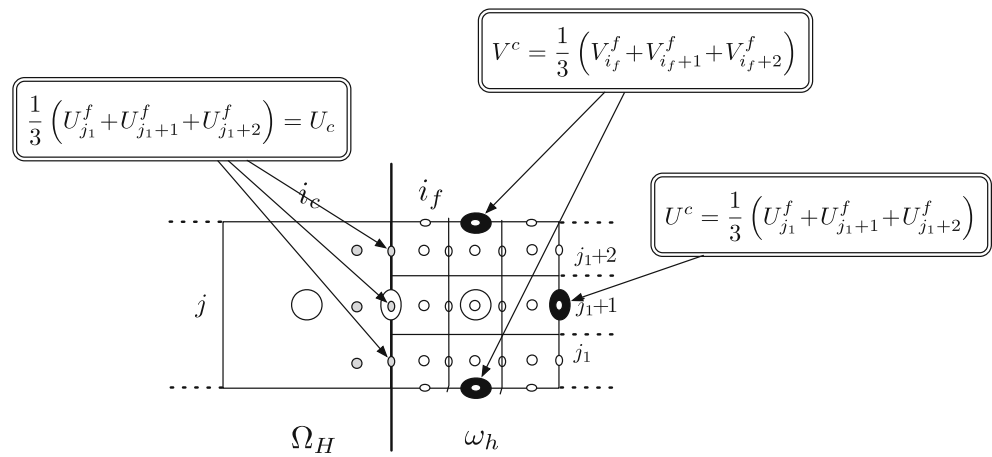
$$\frac{\partial \eta}{\partial t} + \nabla \cdot (U, V) = 0 \tag{9}$$

where η is the free surface and (U, V) is the barotropic transports in x and y directions. If the transport U has been conservatively interpolated on the fine grid boundary, then a simple update (average in one direction and simple copy in the other direction, as shown in Fig. 12) of the transports leads to a barotropic divergence in the coarse-grid cell equal to the average of the barotropic divergence on the fine-grid cell:

$$\nabla \cdot (U^c, V^c) = \frac{1}{9} \sum_{i,j} \nabla \cdot (U_i^f, V_j^f) \tag{10}$$

However, as we already saw (Section 2.2.1), the constraint given by Eq. 10, which actually corresponds to an equality between tendencies of the free surface equation instead of instantaneous values, can be removed if either the right-hand side of the coarse grid is directly updated by the right-hand side of the fine grid or if a truly multilevel solver is used. Removing this constraint allows the use of more scale-selective update operators than those used here.

Fig. 12 Update and interpolation schemes for the preservation of the barotropic divergence on a C-grid with a mesh refinement factor of 3. The transport (U) is conservatively interpolated on the boundary (grey ellipses) and then simple update operators are used at coarse grid points (black ellipses)



4.3 Inconsistency at boundary

We now look at a particular problem that also induces differences between implementation of intergrid transfer operators. Let us rewrite the explicit algorithm given in Section 2.1

1. $u_c^{n+1} = \mathcal{L}_c(u_c^n)$
2. FOR $m = 1 \dots \rho_t$ DO
 $u_f^{n+\frac{m}{\rho_t}} = \mathcal{L}_f \left(u_f^{n+\frac{(m-1)}{\rho_t}} \right)$
 $u_f^{n+\frac{m}{\rho_t}} \Big|_{\gamma_h} = P(u_c^n, u_c^{n+1})$
3. $u_c^{n+1} \Big|_{\omega_H} = R(u_f^{n+1}(u_c^{n+1}))$,

where we insist on the fact that u_f^{n+1} is indeed dependent on u_c^{n+1} through the boundary interpolation so that an inconsistency (sometimes called overspecification) is inherent to the third step. Note that this does not apply to variables entirely lying on the interface, like normal velocities on a C-grid, because the coarse grid variables used in the interpolation (along the interface) are not updated in the feedback step. On a C-grid, this inconsistency occurs for tracers and tangential velocities. One possible way to remove this inconsistency is to modify the interpolation step 2 by doing intermediate updates so that step 2 is modified as follows:

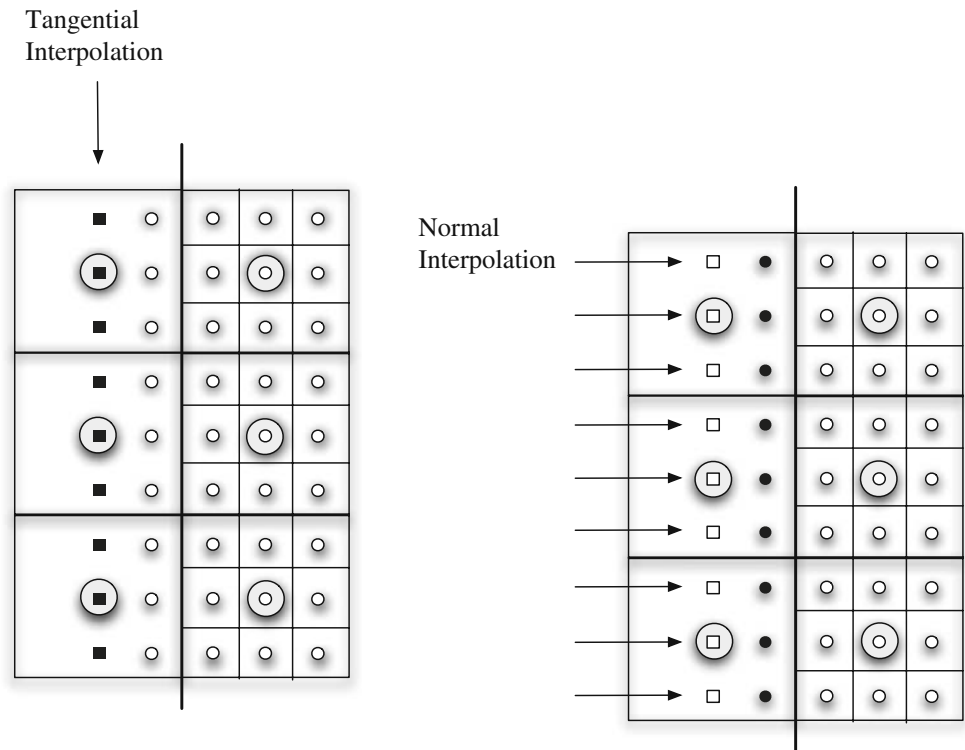
$$u_f^{n+\frac{m}{\rho_t}} \Big|_{\gamma_h} = P(u_c^n, u_c^{n+1}) \rightarrow u_f^{n+\frac{m}{\rho_t}} \Big|_{\gamma_h} = P \left(u_c^n, \left\{ \begin{array}{l} u_c^{n+1} \text{ in } \Omega_H \setminus \omega_H \\ R \left(u_f^{n+\frac{m}{\rho_t}} \Big|_{\gamma_h} \right) \text{ in } \omega_H, \end{array} \right\} \right)$$

eventually using a fixed point iteration if the operators P and/or R are nonlinear.

To our knowledge, this method has not been explored. In the literature, several other methods have been developed to try to remove this contradiction:

- Interpolation operators. It is possible to write the interpolations so that no coarse grid points that will be modified during the feedback step (i.e., inside the HR grid domain) are used in the interpolation process. In order to do this, a tangential interpolation is first applied, followed by a normal interpolation (Fig. 13). The idea is also to use the HR solution as soon as available, thus lowering interpolation errors (in both space and time). Note that, here too, care must be taken about stability issues. As an example, using a high-order interpolation, which involves several fine-grid points inside the fine-grid domain, biases the interpolation schemes and can produce unstable solutions if the flow is actually entering the fine grid. Therefore, in practice, the normal interpolation should be computed using the direction of the flow. Again, this is especially true when the model computes its fluxes using second-order centered differences.
- Update operators—separation of dynamic and feedback interface. Another remedy to remove this inconsistency is to modify the feedback step. Several authors have proposed to separate the feedback interface from the dynamic interface (sometimes also called the input interface) where the boundary values are interpolated (Phillips and Shukla 1973; Zhang et al. 1986; Oey and Chen 1992). Similarly, the Kurihara method can also be seen as introducing a separation between forcing and feedback interfaces. Figure 14 shows a separation of the dynamic and feedback interface by one coarse-grid cell. This separation can also

Fig. 13 Tangent interpolation followed by a normal interpolation. A set of fictitious points (*squared boxes*) are first obtained by interpolation along the tangential direction; this is followed by a normal interpolation that, in addition to these points, uses fine-grid points inside the fine-grid domain. Note that the coarse-grid points inside the fine-grid domain are not used in the interpolation process

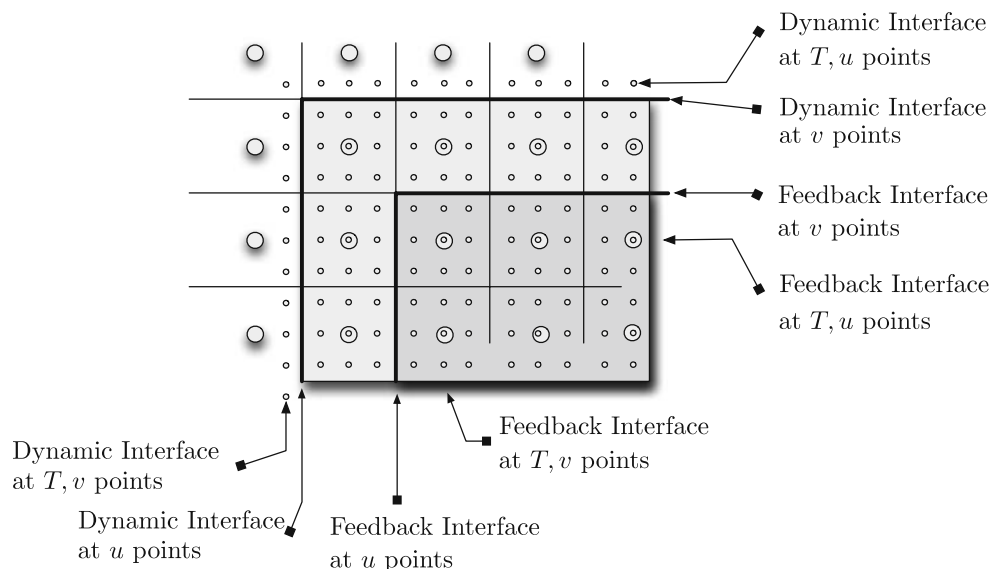


be composed of two coarse-grid cells. In these cases, the contradiction has been removed since the coarse-grid points used in the interpolations are not updated.

Another reason for using a mesh separation between dynamic and feedback interfaces is that, if noise is produced, it will be larger near the dynamic interface so that it is safer not to use the fine grid values near the dynamic interfaces.

However, as mentioned by Spall and Holland (1991), setting the feedback interface closer to the dynamic interface maximizes the information transferred to the coarse grid. In addition, since there is no feedback in this area, the consistency between coarse and fine solutions will be less than with feedback, and this inconsistency has to be lowered by the use of noise control techniques that will indeed recombine and smooth the HR and CR

Fig. 14 Separation of dynamic and feedback interface on a C-grid for a mesh refinement factor of 3. Without a separation of these interfaces, the tracers and tangential velocities would also be updated in the *light grey area*



solutions in the area between dynamic and feedback interfaces.

Partial conclusion From the authors' experience, update operators have a greater impact on the quality of the solution than interpolation operators (assuming that these last ones are properly built, e.g., conservative when they have to be). The usual update operators (average, Shapiro filter...) do not have sufficiently good numerical properties to lead to stable simulations. In particular, the use of update operators that produce a stronger damping of small scales is required.

The technique of separation of dynamic and feedback interfaces can be used associated with noise control techniques to make robust two-way embedding simulations. However, errors introduced by these methods may be found to be very large when quantitatively evaluated.

5 Noise control

Two principles can motivate the use of noise control techniques:

- Maintain a strong consistency between HR and CR solutions in the area where solutions interact (i.e., near the common interface).
- Prevent waves reflection: waves unresolved on the coarse grid reflect inside the fine grid domain and have to be filtered in some way.

It is generally accepted that one criterion for choosing the noise control technique is that the associated modifications should cancel if the refinement ratio is equal to 1, so that if the simulation is made with a mesh refinement factor of 1, the result is identical to the result computed with a uniform grid. Of course, the use of noise control techniques modifies the original differential operator near the interface and, thus, lowers the order of approximation.

Two-way embedding models can make use of noise control techniques that are traditionally used in open-boundary and/or one-way embedded simulations. We briefly recall here some of them:

- Use of a time-stepping algorithm or other internal numerical scheme that damps small wavelengths. This is the crudest form of noise control technique: the original model solution is sufficiently damped so that small scales are strongly removed.

- Relaxation methods
 - Blending (flow relaxation scheme)

$$q_f^* = (1 - \mu_{x,\partial\omega})q_f + \mu_{x,\partial\omega} Pq_c$$

$\mu_{x,\partial\omega}$ is a coefficient varying from 1 on the interface $\partial\omega$ to 0 away to the interface (Davies 1976). This is used in Oey and Chen (1992) in the separation area between feedback and dynamic interfaces.

- Nudging

$$\frac{\partial q_f}{\partial t} = \dots - \frac{(q_f - Pq_c)}{T_d}$$

The discrete form of nudging is equivalent to the flow relaxation scheme (Martinsen and Engedahl 1987).

- Increase of dissipation coefficient—sponge layer. When using a sponge layer, it is advantageous to apply it to the difference between external (CR) and internal (HR) fields.

$$\frac{\partial q_f}{\partial t} = \dots + (-1)^{n+1} (\Delta)^n [\mu_{x,\partial\omega}(q_f - Pq_c)] \quad (11)$$

Here, the objective is clearly to damp small scales as defined by the restriction operator. Since the coarse grid variable q_c has been updated ($q_c = \mathcal{R}q_f$), the diffusion term mostly acts on the scales lying in the kernel of the restriction operator ($\mathcal{R}q_f = 0$). Note that the preceding relaxation methods can be put under this category just by choosing $n = 0$ in Eq. 11 so that they produce a less selective damping than for $n \geq 1$.

The advantage of this formulation is also that the diffusion term naturally cancels if there is no refinement. Using differences between coarse- and fine-grid variables can also be at the basis of the derivation of open boundary schemes suitable for embedding methods (e.g., Perkins et al. 1997; Oddo and Pinaridi 2008). Additionally, in a flux correction approach of conservation, it is easy to maintain conservation just by adding the diffusive flux of Eq. 11.

Moreover, due to the inability to build ideal restriction operators, it may also be useful to smooth the coarse-grid solution inside the fine-grid domain. The filtering is only applied to the coarse-grid points inside the fine-grid domain in order not to violate conservation principles. As an example, the smoothing–desmoothing algorithm of Bender et al. (1993) can be applied on the first coarse grid points near the interfaces.

Partial conclusion Two-way grid embedding algorithms should be first evaluated without using noise control techniques in order not to hide fundamental problems of the algorithm. Because the primary requirement is that the associated modifications of the original model should cancel if the refinement factor is equal to 1, most of the methods are formulated in terms of differences between coarse and fine fields as in Eq. 11. Choosing $n = 1$ in Eq. 11 (laplacian diffusion) is a good compromise between a low computational cost and a selective damping of the difference. For $n \geq 1$, the correction can be easily written in terms of flux divergences, thus allowing the conservation properties to be maintained (e.g., using the flux correction approach).

6 Conclusion

This paper reviews existing two-way embedding techniques. The emphasis has been put on what the sources of potential approximations and errors are. In particular, we have shown how the use of implicit solver and/or time splitting algorithms and the use of refinement in time can introduce several problems in comparison with an idealized model. Conservation issues can be tackled in a simple way, using the flux correction approach (at least for quantities based on cell centered variables: mass, tracers...). There is also a large choice of intergrid transfer operators; the choice between one of them is a trade-off between conservation, accuracy, robustness, and coding complexities. Some guidelines have been given as partial conclusions of each section.

Several issues related to grid embedding have not been covered. Computational implementation of embedding can be done by hand or by using dedicated software (e.g., AGRIF Debreu et al. 2008b; RSL Michalakes 1998). In the context of operational oceanography and production of HR forecasts at particular areas of interest, HR grid initialization schemes that are in agreement with the two-way interactions have to be derived. As most of the operational systems are now using data assimilation methods, one possible way to look at the problem is to introduce the HR grid initialization in the assimilation process, and this is an active research subject.

Acknowledgements This work is a contribution to the MERSEA project. Partial support from the European Commission under contract SIP3-CT-2003-502885 is gratefully acknowledged. We would like to thank two anonymous reviewers for their comments and careful reading of the manuscript.

References

- Alapaty K, Mathur R, Odman T (1998) Intercomparison of spatial interpolation schemes for use in nested grid models. *Mon Weather Rev* 126(1):243–249
- Barth A, Alvera-Azcarate A, Rixen M, Beckers J-M (2005) Two-way nested model of mesoscale circulation features in the Ligurian Sea. *Prog Oceanogr* 66:171–189
- Bender MA, Ross RJ, Tuleya RE, Kurihara Y (1993) Improvements in tropical cyclone track and intensity forecasts using the GFDL initialization system. *Mon Weather Rev* 121:2046–2061
- Berger MJ (1985) Stability of interfaces with mesh refinement. *Math Comput* 45:301–318
- Berger MJ, Olinger J (1984) Adaptive mesh refinement for hyperbolic partial differential equations. *J Comput Phys* 53(3):484–512
- Berger MJ, Colella P (1989) Local adaptive mesh refinement for shock hydrodynamics. *J Comput Phys* 82(1):64–84
- Blayo E, Debreu L (2005) Revisiting open boundary conditions from the point of view of characteristic variables. *Ocean Model* 9:231–252
- Blumberg AF, Mellor GL (1987) A description of a three-dimensional coastal ocean circulation model. In: Heaps N (ed) *Three-dimensional coastal ocean models*, vol 4. American Geophysical Union, Washington, DC, p 208
- Clark TL, Farley D (1984) Severe downslope windstorm calculations in two and three spatial dimensions using anelastic interactive grid nesting: a possible mechanism for gustiness. *J Atmos Sci* 41:329–350
- Davies, HC (1976) A lateral boundary formulation for multi-level prediction models. *Q J R Meteorol Soc* 102:405–418
- Debreu L, Marchesiello P, Penven P (2008a) Two-Way embedding algorithms for a split-explicit free surface ocean model. *Ocean Model* (in preparation)
- Debreu L, Vouland C, Blayo E (2008b) AGRIF: Adaptive Grid Refinement In Fortran. *Comput Geosci* 34(1):8–13
- Dukowicz JK, Smith RD (1994) Implicit free-surface method for the Bryan-Cox-Semtner ocean model. *J Geophys Res* 99:7991:8014
- Fox AD, Maskell SJ (1995) Two-way interactive nesting of primitive equation ocean models with topography. *J Phys Oceanogr* 25:2977–2996
- Ginis I, Richardson RM, Rothstein LM (1998) Design of a multiply nested primitive equation ocean model. *Mon Weather Rev* 126:1054–1079
- Harrison EJ, Elsberry RL (1972) A method for incorporating nested finite grids in the solution of systems of geophysical equations. *J Atmos Sci* 29:1235–1245
- Heggelund Y, Bernsten J (2002) A method for analysing nesting techniques for the linearized shallow water equations. *Int J Numer Methods Fluids* 38:163–185
- Herrnstein A, Wickett M, Rodrigue G (2005) Structured adaptive mesh refinement using leapfrog time integration on a staggered grid for ocean models. *Ocean Model* 9(3): 283–304
- Jones RW (1977) A nested grid for a three-dimensional model of a tropical cyclone. *J Atmos Sci* 34:1528–1553
- Killworth PD, Stainforth D, Webb DJ, Paterson SM (1991) The development of a free-surface Bryan-Cox-Semtner ocean model. *J Phys Oceanogr* 21:1333–1348
- Koch SE, McQueen JT (1987) A survey of nested grid techniques and their potential for use within the MASS weather prediction model. NASA Technical Memorandum 87808

- Kurihara Y, Tripoli GJ, Bender MA (1979) Design of a movable nested-mesh primitive equation model. *Mon Weather Rev* 107:239–249
- Kurihara Y, Bender MA (1980) Use of a movable nested-mesh model for tracking a small vortex. *Mon Weather Rev* 108:1792:1809
- Laugier M, Angot P, Mortier L (1996) Nested grid methods for an ocean model: a comparative study. *Int J Numer Methods Fluids* 23(11):1163–1195
- Martin DF, Cartwright KL (1996) Solving Poisson's equation using adaptive mesh refinement. Technical Report UCB/ERL, M96/66 UC, Berkeley
- Martinsen EA, Engedahl HE (1987) Implementation and testing of a lateral boundary scheme as an open boundary condition in a barotropic ocean model. *Coast Eng* 11: 603–627
- Michalakes J (1998) Runtime System Library (RSL) for parallelization of gridded weather models with nesting. <http://www-unix.mcs.anl.gov/~michalak/rsl>
- Oddo P, Pinardi N (2008) Lateral open boundary conditions for nested limited area models: a scale selective approach. *Ocean Model* 20:134–156
- Oey LY, Chen P (1992) A nested-grid ocean model: with application to the simulation of meanders and eddies in the Norwegian Coastal Current. *J Geophys Res* 97(20): 063–086
- Olsson F, Petersson NA (1996) Stability of interpolation on overlapping grids. *Comput Fluids* 25(6):583–605
- Part-Enander E, Sjogreen B (1994) Conservative and non-conservative interpolation between overlapping grids for finite volume solutions of hyperbolic problems. *Comput Fluids* 23(3):551–574
- Penven P, Debreu L, Marchesiello P, McWilliams JC (2006) Evaluation and application of the ROMS 1-way embedding procedure to the central california upwelling system. *Ocean Model* 12:157–187
- Perkins AL, Smedstad LF, Blake DW, Heburn GW, Wallcraft AJ (1997) A new nested boundary condition for a primitive-equation ocean model. *J Geophys Res* 102(C2):3483–3500
- Phillips NA, Shukla J (1973) On the strategy of combining coarse and fine grids meshes in numerical weather prediction. *J Appl Meteorol* 12:763–770
- Sheng J, Greatbatch RJ, Zhai X, Tang L (2005) A new two-way nesting technique for ocean modeling based on the smoothed semi-prognostic method. *Ocean Dyn* 55:162–177
- Sobel JP (1976) Nested grids in numerical weather prediction and an application to a mesoscale streak. Ph.D. thesis, Pennsylvania State University, p 135
- Spall MA, Holland WR (1991) A nested primitive equation model for oceanic applications. *J Phys Oceanogr* 21:205–220
- Zhang DL, Chang HR, Seaman NL, Warner TT, Fritsch JM (1986) A two-way interactive nesting procedure with variable terrain resolution. *Mon Weather Rev* 114:1330–1339
- Zhuo L, Qingcun Z (1995) The application of time-dependent adaptive mesh model in typhoon track prediction. *Chin J Atmos Sci* 19:303–310

Tsunami simulations on several scales

Comparison of approaches with unstructured meshes and nested grids

Sven Harig · Chaeroni · Widodo S. Pranowo ·
Jörn Behrens

Received: 28 April 2008 / Accepted: 24 October 2008 / Published online: 14 November 2008
© Springer-Verlag 2008

Abstract The tsunami event generated by the great Sumatra–Andaman earthquake on 26 December 2004 was simulated with the recently developed model TsunAWI. The model is based on the finite element method, which allows for a very flexible discretization of the model domain. This is demonstrated by a triangulation of the whole Indian Ocean with a resolution of about 14 km in the deep ocean but a considerably higher resolution of about 500 m in the coastal area. A special focus is put on the Banda Aceh region in the Northern tip of Sumatra. This area was heavily hit by the tsunami and the highest resolution in this area is about 40 m in order to include inundation processes in the model simulation. We compare model results to tide gauge data from all around the Indian Ocean, to satellite altimetry, and field measurements of flow depth in selected locations of the Aceh region. Furthermore, we compare the model results of TsunAWI to the results of a nested grid model (TUNAMI-N3) with the same initial conditions and identical bathymetry and topography in the Aceh region. It turns out that TsunAWI gives accurate estimates of arrival times in distant locations and in the same mesh gives good

inundation results when compared to field measurements and nested grid results.

Keywords Tsunami · Numerical modeling · Finite element method · Unstructured meshes · Nested grids · Shallow water equations · Model comparison

1 Introduction

The devastating tsunami generated by the great Sumatra–Andaman earthquake on December 26 in 2004 has triggered many activities aimed at the establishment of a tsunami early warning system for the Indian Ocean. As part of the German aid contribution in the framework of the German–Indonesian Tsunami Early Warning System (GITEWS) (www.gitews.de), a tsunami wave propagation model (TsunAWI) is being developed at AWI (Behrens 2008). Numerical modeling relies on a discretization of the physical space. The two major strategies in this respect are the regular and the unstructured discretization of the domain under consideration. The largest part of models used for tsunami simulations at the moment are based on regular meshes. Many of these have been successfully applied to the December 2004 tsunami event (see Geist et al. 2007; Kowalik et al. 2007; Grilli et al. 2007; Titov and Gonzalez 1997). Even the global extent has been investigated in Kowalik et al. (2005) and Titov et al. (2005). On the other hand, ocean models based on unstructured grids are being developed by many institutions and wave propagation problems were among the first applications in the ocean successfully tackled with this technology (see, for example, Piatanesi et al.

Responsible Editor: Eric Deleersnijder

S. Harig (✉) · W. S. Pranowo · J. Behrens
Alfred Wegener Institute for Polar and
Marine Research (AWI), Bremerhaven, Germany
e-mail: Sven.Harig@awi.de

W. S. Pranowo
e-mail: Widodo.Pranowo@awi.de

Chaeroni
Coastal Dynamic Research Center BPPT,
Jogyakarta, Indonesia

1999; Tinti and Gavagni 1995; Walters 2006; Walters and Goff 2003).

TsunAWI is based on an unstructured grid approach employing finite elements to solve the governing equations. Development of this model has started as a spin-off from the 3D ocean model FEOM, which is being developed at AWI (see Danilov et al. 2004). Motivation for the choice of unstructured grids is the ability to simulate wave processes on different scales. In the deep ocean, coarse resolution (several kilometers) is sufficient to represent the wave adequately. However, while reaching the shoreline, the tsunami steepens, the wave may break, contributions of nonlinear terms become more important, and much higher resolution is needed to include even rough estimates of the inundation processes. In finite difference approaches, this problem is tackled by a series of nested grids ranging from about 1 km in the coarsest grid to less than 100 m in the finest (see Fig. 2). In an unstructured grid, all constraints on the local resolution mentioned above may be satisfied in one grid where, additionally, transitions of nodal density are smooth.

In the framework of an early warning system model, results are crucial as important quantities such as arrival time and estimated wave heights are derived from precalculated scenarios of future tsunamis [e.g., a description of the Japanese early warning system is given in Furumoto et al. (1999)]. TUNAMI is recommended within the Tsunami Inundation Modeling Experiment project and has been used in many countries for tsunami simulations already. A direct comparison of this model to TsunAWI is therefore especially helpful for gaining insights into possible advantages and disadvantages of the finite element approach in unstructured grids.

The subject of the present study is the tsunami generated by the Sumatra–Andaman earthquake on 26 December 2004. The initial conditions are taken from the reconstruction of the rupture as described in Tanioka et al. (2006). We compare model results to the following data:

- Arrival times and wave heights in tide gauge records from rim countries of the Indian Ocean
- Satellite altimetry data obtained from Jason-1, Envisat and Topex-Poseidon
- Field measurements in selected positions in Banda Aceh region

Additionally, we compare results of both models, TsunAWI and TUNAMI-N3, with respect to arrival times and wave height in virtual gauges, as well as the

inundation obtained in Aceh region. It turns out that arrival times and estimated amplitude of the first wave crest coincide well, given the uncertainties in approximation and data. The same is true for the inundation area, which coincides well in both models. Field data in selected positions were used to adjust the friction parameters of the models.

Section 2 describes the model setup for the numerical experiments. Section 3 contains some details on the models TUNAMI and TsunAWI. Section 4 describes the results of both models before conclusions are drawn in Section 5.

2 Tsunami source model

There are many studies on the rupture process of the 2004 Sumatra–Andaman earthquake using seismic data estimating the moment magnitude (M_w) about 9.0 to 9.3 (Ammon et al. 2005; Park et al. 2005). For water wave modeling purposes, a precise description of the fault mechanism is needed. In Tanioka et al. (2006), the rupture process of the 2004 Sumatra–Andaman earthquake has been estimated by tsunami wave forms at tide gauges, as well as coastal coseismic vertical deformations. It turns out that a rupture speed of 1.7 km/s gives the best match between tide gauge data in five stations and the synthetic tsunami propagation. These results are used to derive initial conditions for both models. The source area of the 2004 Sumatra–Andaman earthquake is divided into 12 subfaults. The event is described by consecutive activation of the corresponding 12 fault plates located as shown in Fig. 1. Assuming the rupture speed mentioned above, the whole earthquake takes 12 min. Each fault plate moves as described by a set of Okada parameters (see Okada 1985, for a theoretical background) specifying the location, bearing, and slip of each plate. Table 1 gives the timing and amount of slip for the plates depicted in Fig. 1. The large slip values of plates A and C correspond to the largest bottom deformations in these regions. It is assumed that the bottom and ocean surface move at the same rate. Whenever a subfault is to be moved during model integration (as given in Table 1), the model is reinitialized, i.e., the bottom and sea surface height are adjusted instantly at the same rate in order to conserve volume.

There are other approaches to the parameters of the rupture mechanism. Banerjee et al. (2007) presents coseismic slip distributions inverted from GPS data, whereas Hirata et al. (2006) optimizes the earthquake

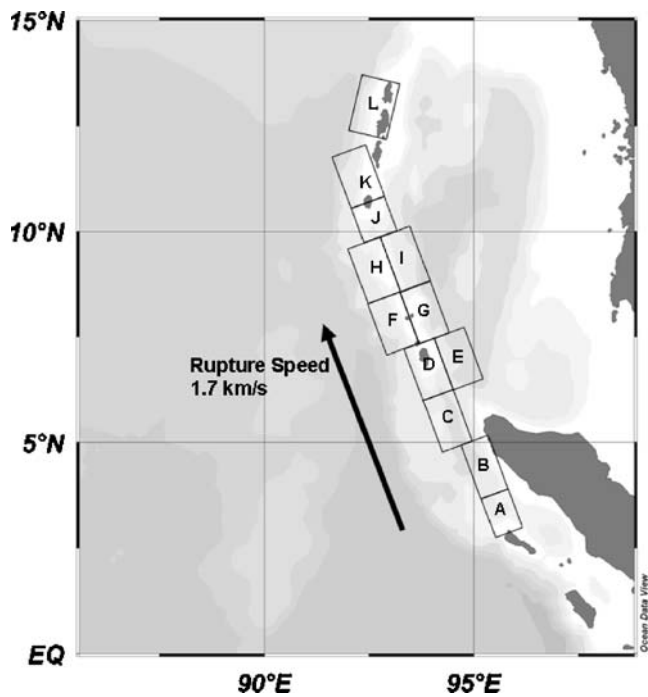


Fig. 1 Subfaults as proposed in Tanioka et al. (2006). The rupture area has been decomposed into 12 subregions. With the indicated rupture speed, the whole rupture process takes 12 min

parameters with respect to satellite altimetry. We did additional experiments with initializations derived from these sources. Initial conditions taken from Banerjee et al. (2007) improved matching with Indian tide gauge data, whereas matching to satellite altimetry was improved using the initialization from Hirata et al. (2006). However, none of the approaches resulted in consistent agreement of the modeling results to all of the available measurement data.

Table 1 Fault plates A–L specifying the initial conditions

Fault plate	Length [km]	Width [km]	Depth [km]	Strike angle [deg]	Dip angle [deg]	Time [min]	Slip [m]
A	100	100	10	340	10	0	26.1
B	160	100	10	340	10	1	0.0
C	150	90	10	340	10	1	29.6
D	150	100	10	340	10	3	7.3
E	150	100	27	340	10	3	10.9
F	150	100	5	340	10	5	7.8
G	150	100	22	340	10	5	0.0
H	150	100	5	340	10	7	12.1
I	150	100	22	340	10	7	16.5
J	100	110	10	340	10	8	16.6
K	150	110	10	340	3	10	7.7
L	100	110	5	10	17	12	1.4

The slip values correspond to the largest values proposed in Tanioka et al. (2006) for the case of a rupture speed of 1.7 km/s

3 Model description

Both TsunAWI and TUNAMI are based on the non-linear shallow water theory. The parameterization of bottom friction is based on Manning’s approach. The friction coefficient n ranges from 0.01 for smooth concrete to 0.06 for poor natural channels (see Arcement and Schneider 1984). This parameter varies between 0.015 and 0.04 in both models; however, it is assumed constant for all the coast. Tidal forcing is not included in the simulations.

3.1 TsunAWI

TsunAWI has been developed within the framework of GITEWS. It is based on an unstructured mesh approach applying finite elements for discretization. Model results depend heavily on the quality of the underlying grid. Therefore, numerics and mesh generation are described separately.

3.1.1 Governing equations and numerics

The velocity formulation of the shallow water equations is used, i.e., the governing equations are

$$\partial_t \mathbf{v} + (\mathbf{v} \cdot \nabla) \mathbf{v} + \mathbf{f} \times \mathbf{v} + g \nabla \zeta + \frac{gn^2 \mathbf{v} |\mathbf{v}|}{H^{4/3}} - \nabla \cdot (A_h \nabla \mathbf{v}) = 0 \tag{1}$$

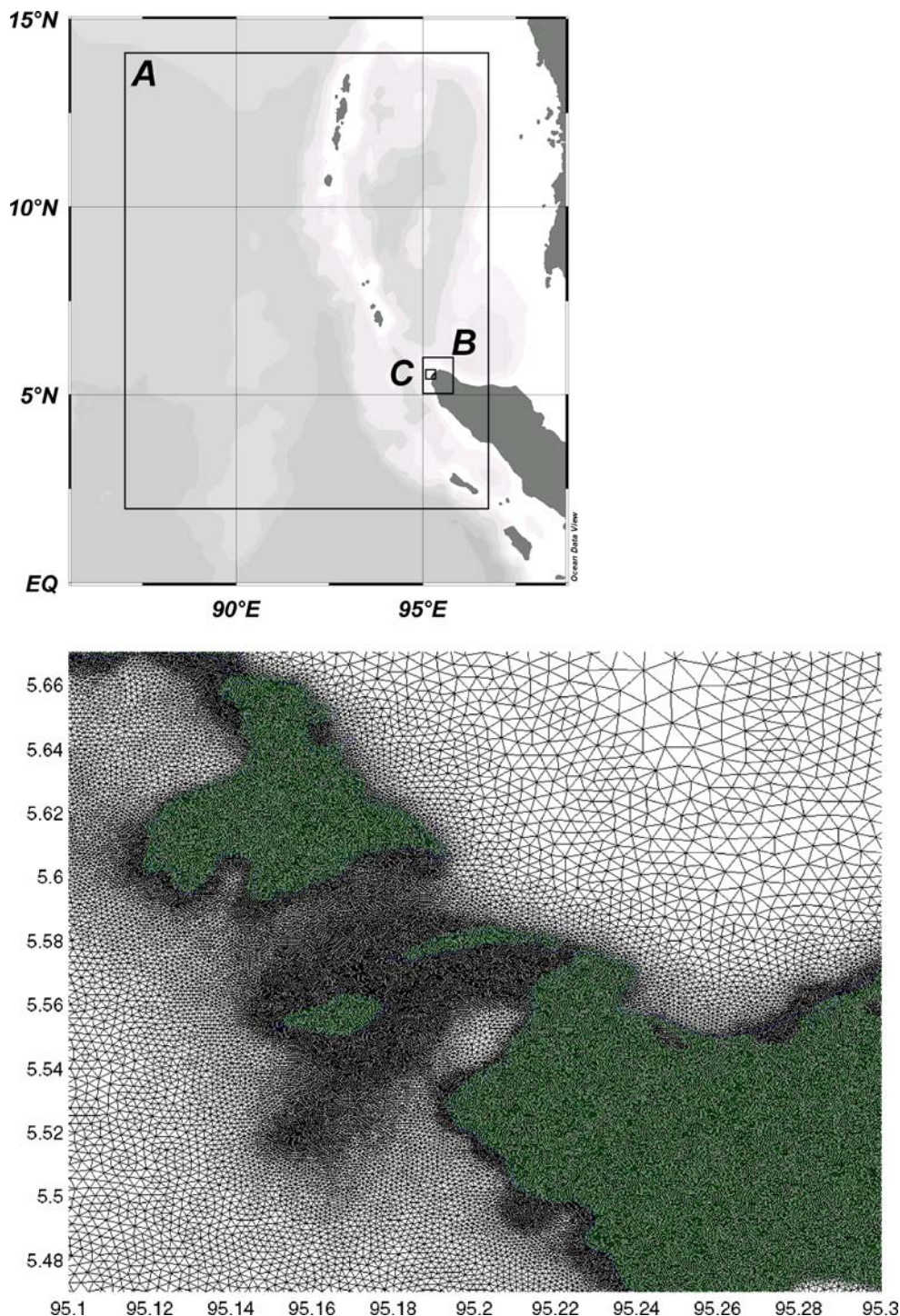
$$\partial_t \zeta + \nabla \cdot (\mathbf{v} H) = 0, \tag{2}$$

where $\mathbf{v} = (u(t, \mathbf{r}), v(t, \mathbf{r}))$ denotes the horizontal depth averaged velocity and $H = h(\mathbf{r}) + \zeta(t, \mathbf{r})$ denotes the total water depth, with ζ the time-dependent sea surface elevation relative to the mean water depth h .

The discretization of Eqs. 1 and 2 follows the ideas published in Hanert et al. (2005). Approximate solutions are determined in the spaces of piecewise linear basis functions for ζ and nonconforming linear basis functions for \mathbf{v} . This combination ($P^1 - P^1_{NC}$) is known to be well suited for wave propagation problems (see

Hanert et al. 2005; Le Roux and Lin 1998). Parameterizations of bottom friction are based on empirical studies of open channel flow. Most frequently formulas by Manning and Chézy are used. We implemented both variants; in the present study, a Manning formulation was employed. The coefficient n is assumed constant

Fig. 2 The upper panel shows the nested grids A, B, and C in TUNAMI (grid A: $1,280 \times 1,354$ nodes, $dx = 900$ m; grid B: 297×357 nodes, $dx = 300$ m; grid C: 289×274 nodes, $dx = 100$ m). The lower figure displays a close up of the TsunAWI mesh to Aceh region (approximately domain C in upper panel). Green nodes are land nodes (initially dry)



in all of the domain and takes values between 0.015 and 0.04. Comparisons of inundation results with field measurements are used to tune this factor. The horizontal diffusion is needed for numerical stability. It is of Smagorinsky type as described in Stacey and Nowak (1995). The factor A_h is constant on elements and the value in each triangle depends on the size and the velocity gradient. The diffusion coefficient in triangle τ is given by

$$A_h(\tau) = cA(\tau)\sqrt{(\partial_x u)^2 + \frac{1}{2}(\partial_x v + \partial_y u)^2 + (\partial_y v)^2}$$

where $A(\tau)$ denotes the area of τ , all gradients are determined in τ , and c is a constant. A benchmark experiment (Okushiri test case) was used to compare the damping in the model to measurements and the actual value of c was chosen in order to guarantee a stable model yet minimizing the damping. Time stepping is performed with a Leap-Frog scheme. A Roberts–

Table 2 Model characteristics

Parameter	TsunAWI	TUNAMI-N3
Number of nodes	5.066 Mio	1.918 Mio
Time stepping	Leap frog	Leap frog
Manning coefficient	0.035	0.025
Time step	0.5 s	1 s in all grids
Resolution [m]	40–14500	900 (A), 300 (B), 100 (C)

Asselin filter is used to eliminate the numerical mode. The time step in all experiments is 0.5 s.

The run up scheme implemented in TsunAWI follows the ideas of Lynett et al. (2002). Once the wave crest reaches the shore, the gradient of the sea surface height is extrapolated from wet nodes to dry nodes on land. This information is used to calculate the velocity and sea surface height in the next time step. In this approach the new shoreline after the wetting process is not limited to computational nodes; instead, it might

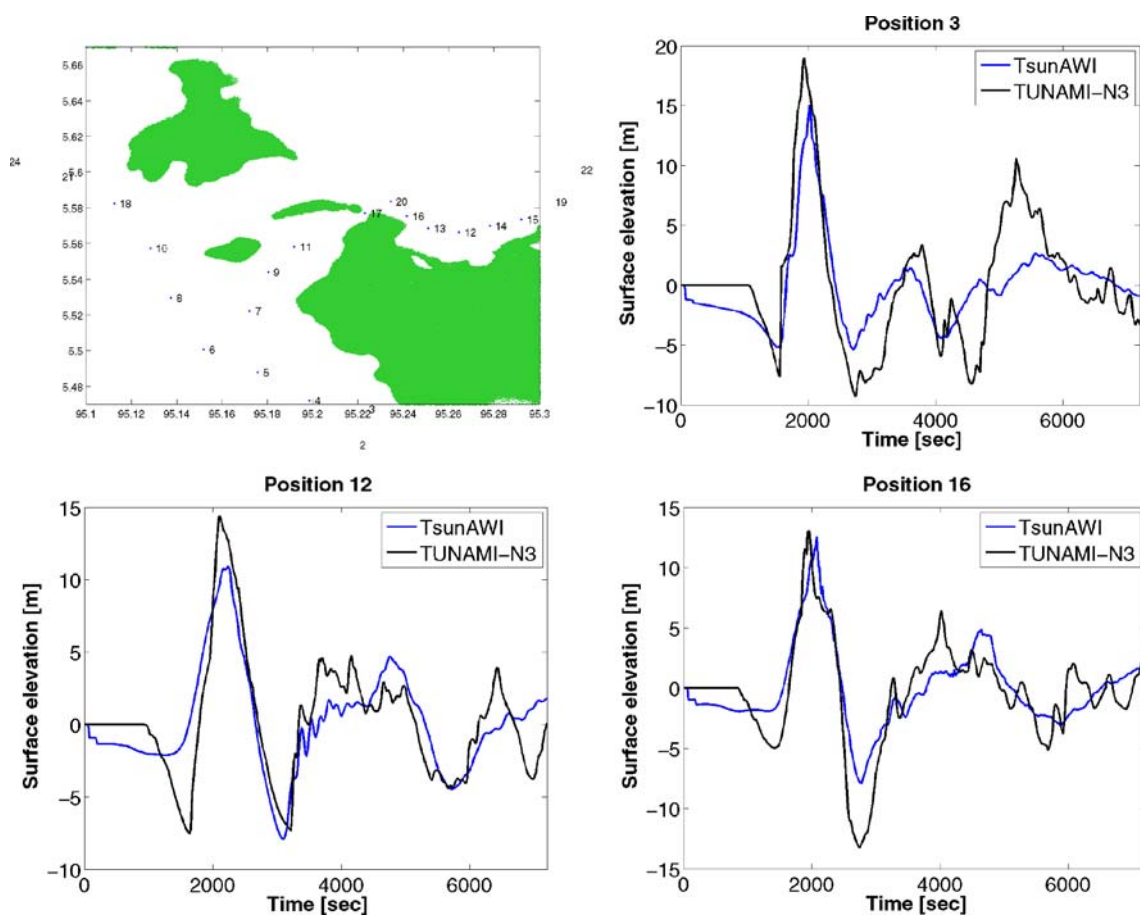


Fig. 3 Time series obtained with both TsunAWI and TUNAMI-N3 in selected locations. All virtual gauges are shown in the upper left panel. Green areas are land nodes contained in the model

Table 3 The upper part compares arrival times obtained with TUNAMI and TsunAWI at the nine field stations depicted in Fig. 8 and the lower part compares the arrival times at six tide gauges (taken from Rabinovich and Thomson 2007) to the values obtained with TsunAWI

	Station	Gauge data [sec]	TUNAMI-N3 [sec]	TsunAWI (ΔT) [sec]
	Station 1	–	2,520	2,550(+30)
	Station 2	–	2,300	2,260(–40)
	Station 3	–	2,260	2,160(–100)
	Station 4	–	2,160	2,100(–60)
	Station 5	–	2,740	2,780(+40)
	Station 6	–	2,560	2,660(+100)
	Station 7	–	2,680	2,760(+80)
	Station 8	–	1,880	2,020(+140)
	Station 9	–	1,780	1,800(+20)
	Colombo	10,200	–	9,900(–300)
	Male	11,700	–	11,540(–160)
	Diego Garcia	13,560	–	13,240(–320)
	Salalah	25,740	–	25,900(+160)
	Lamu	31,980	–	31,500(–480)
	Kochi	16,920	–	16,700(–220)

Numbers in brackets are time differences between TsunAWI and TUNAMI in the upper part and differences between model and data in the lower part

be located at arbitrary positions in between nodes. The same is true for the drying process. Different extrapolation schemes were implemented following both a linear interpolatory and a linear least squares strategy. The scheme was tested in a simple, quasi-one-dimensional run-up benchmark with linear slope as well as a complicated three dimensional domain (Okushiri test-case) and did perform well. More details of the model are given in Behrens et al. (2007).

3.1.2 Mesh generation

The quality of the mesh is crucial for the quality of the model results. The unstructured mesh in the present study was produced with the freely available mesh generator TRIANGLE by Jonathan Shewchuk, see Shewchuk (1996). The resulting meshes were smoothed to improve the overall mesh quality. The coarsest resolution of the mesh is 14 km in the deep ocean, whereas the coastal area is better resolved. The cell size on land is 500 m in the broad scale and about 100 m in Aceh region. Mesh generation as well as model topography and bathymetry are based on the GEBCO (GEB 2008) data set in a resolution of one arc minute. These data however were improved by several sources. Topography data were obtained from results of the Shuttle radar topography mission SRTM (SRT 2007), which were processed by DLR. Ship cruise data by the research vessels SONNE and SCOTT were included as well. In the domain of TUNAMI grid C, the same topography and bathymetry provided by BPPT Yogyakarta was used by both models.

Mesh generation is performed in three steps. After the boundary of the model domain has been specified as a polygon line, TRIANGLE is used to produce a

basic triangulation containing all boundary nodes and allowing edge lengths not larger than the coarsest acceptable resolution. In the second step, this triangulation is refined according to user-defined rules. These rules are linked to TRIANGLE by a C-routine. Since the phase velocity of the tsunami is given by \sqrt{gh} in linear approximation, the desired resolution is directly linked to the bathymetry. On the other hand, steep bathymetry leads to a steepening of the wave, and such areas must be well resolved. Therefore, the refinement algorithm employed in this study is as follows: Each triangle is refined until all the edges of the resulting triangle fulfill the criteria

$$\Delta x \leq \min \left\{ c_t \sqrt{gh}, c_v \frac{h}{\nabla h} \right\},$$

where Δx denotes edge length and c_t [s] and c_v (dimensionless) are constants. The CFL criterion connects c_t to the largest possible time step. Larger values of c_t result in a coarser overall mesh and a larger time step might be used.

After TRIANGLE has completed the mesh refinement relaxation methods as described in Frey and Field (1991) are applied in the third and final step. The whole model domain with corresponding nodal density is shown in Fig. 5. A small part of the smoothed mesh is displayed in Fig. 2.

3.2 TUNAMI-N3

TUNAMI-N3 (Tohoku University Numerical Analysis Model for Investigation of Near-field tsunamis Vers. 3) utilizes the consistent flux formulation for

the discharge $\mathbf{M} = (M, N) = (uH, vH)$. The governing equations are given by

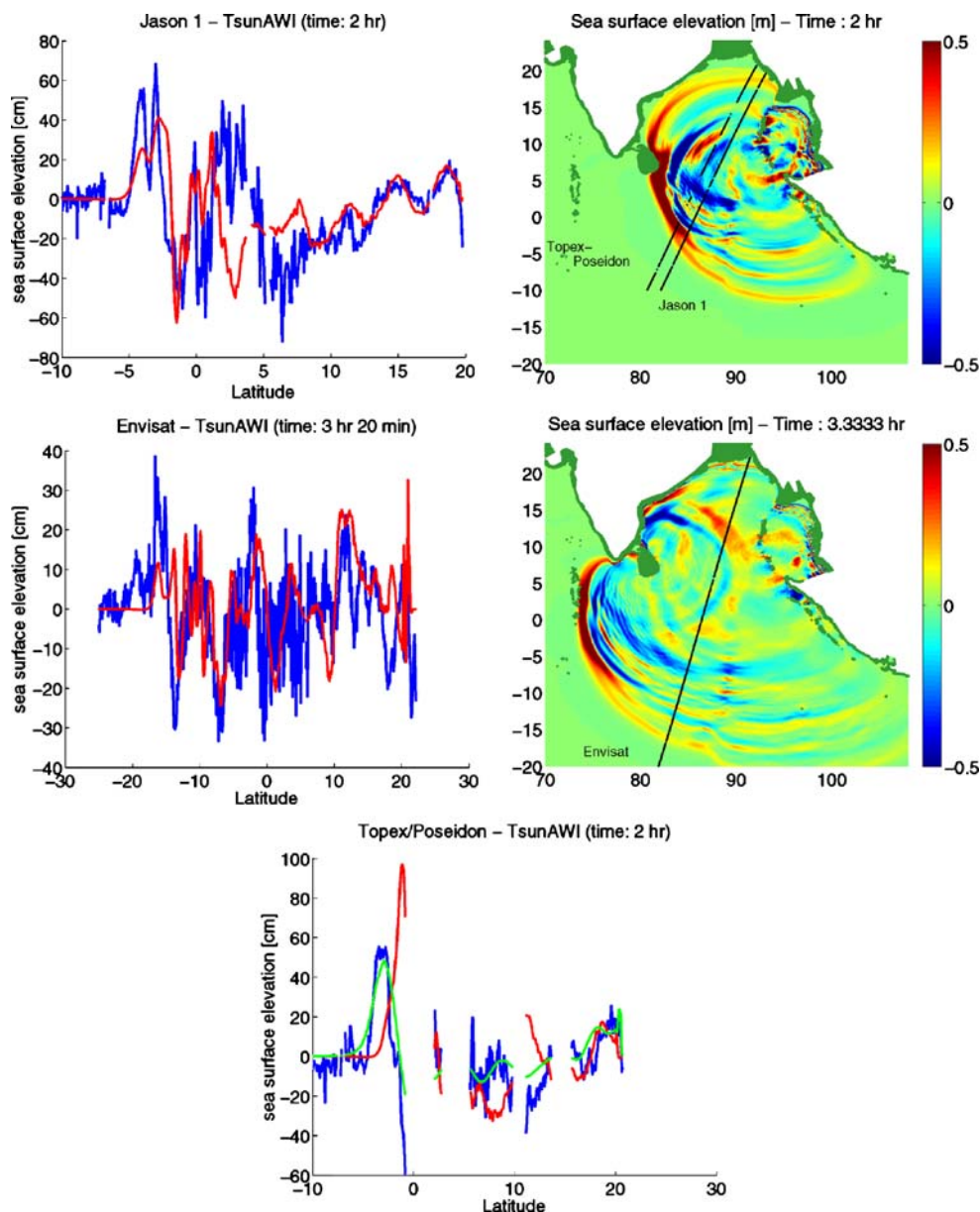
$$\partial_t M + \partial_x(M^2/H) + \partial_y(MN/H) + gH\partial_x\zeta + \frac{gn^2 M|\mathbf{M}|}{H^{7/3}} = 0,$$

$$\partial_t N + \partial_x(MN/H) + \partial_y(N^2/H) + gH\partial_y\zeta + \frac{gn^2 N|\mathbf{M}|}{H^{7/3}} = 0,$$

$$\partial_t\zeta + \nabla \cdot (\mathbf{M}) = 0.$$

The Coriolis force is neglected in this approach. A detailed description of the model is given in Imamura et al. (2006). A sequence of three nested grids is used in the comparison with highest resolution of 100 m in the Aceh region. The spacial extent of all three grids is shown in Fig. 2. A Leap-frog scheme is used for time integration with a time step of one second in all three grids. The time stepping is performed subsequently from coarse to fine grids. The coarser grid provides boundary conditions for the next finer grid and after the time stepping has been performed in the finest grid the coarser grids are updated by averaging the fine cells making up a coarse cell.

Fig. 4 Comparisons to satellite altimetry data (blue lines) by Jason-1 (upper panels) and Envisat (middle panels). The red lines mark model results after 2 h (upper and lowest panels) and 3 h 20 min. The lowest figure displays the comparison to Topex-Poseidon data. The green line marks the result from an alternative initial condition (Banerjee et al. 2007). The green areas along the coast in the left figures are land nodes contained in the model domain



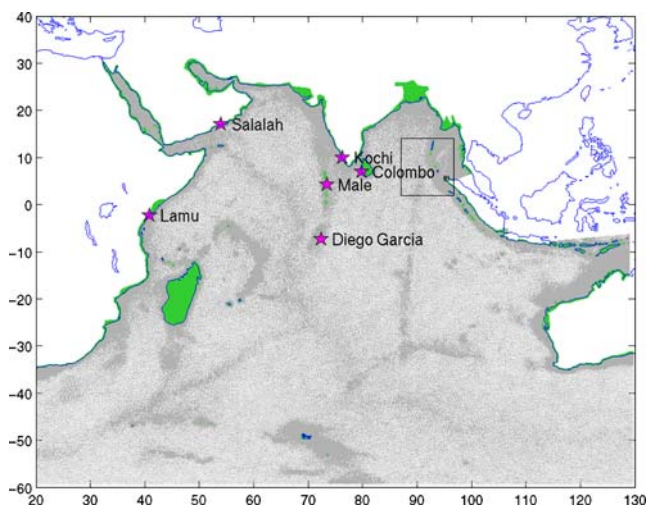


Fig. 5 Model domain of the unstructured grid and positions of the tide gauges that are used for far field model verification. The resolution of this grid ranges from 14 km in the deep ocean to 500 m at the coast and 100 m at the northern tip of Sumatra. The shades of grey correspond to the density of nodes. The green areas are land nodes contained in the model domain. The black squares correspond to the nested grids of TUNAMI

Run-up is only considered in the finest grid and is calculated cell wise. In each time step, the water level along the border separating dry and wet cells is checked with respect to a flooding criterion. If the flow depth in the wet cell exceeds the topography value in the neighboring dry cell the latter is submerged.

4 Discussion of the results

Subject of this study are in the first place model results, which are relevant in an early warning process. We investigate arrival times, wave heights near the coast, and inundation areas obtained in both models. Additionally, we compare the inundation results to field measurements in selected locations to find appropriate values for the Manning friction coefficient n in Eq. 1. Some characteristics of the models are summarized in Table 2.

4.1 Wave propagation

4.1.1 Early stages

The tsunami reaches the Northern tip of Sumatra in about 30 min. Since not many gauge data documenting arrival times in this area are available we compare arrival times and wave heights obtained by both models in virtual gauges close to the shore line in the area of highest resolution. Figure 3 displays both the po-

sitions for comparison, as well as the results in three exemplary locations South and North of Aceh region. In all panels, the position of the first wave crest and the magnitude coincide well. The upper part of Table 3 compares the arrival times determined by both models in the locations displayed in Fig. 8. The differences are acceptable given that different inundation schemes are implemented in TUNAMI and TsunAWI and the resolution in the coarse part of the unstructured grid is considerably lower.

The two models do not agree well in later stages of the time series. However, the behavior of the wave after the coast has already been reached depends strongly on reflections at the shoreline. In the coarse grids A and B of the TUNAMI experiments, no inundation is implemented, and therefore, less realistic results must be expected once reflections from those grids reach the finest grid C. In this context, TsunAWI should give more consistent results as the broad scale resolution is the same for all parts of the coast and inundation is implemented all along the coast.

4.1.2 Comparison to satellite altimetry

The tsunami was observed by several satellites. Jason-1 crossed the equator at 2:55 AM UTC around 2 h after the earthquake (Hayashi 2008) and the onboard altimeter recorded the tsunami in the Bay of Bengal. By comparison to the previous average heights, the signature of the wave was determined. These data are available from NOAA (NOA 2008). It is shown in Fig. 4 together with the model state of TsunAWI after 2 h of integration time. The double peak structure between 5°S and the equator arising from the superposition of the partial waves generated by subfaults A and C (compare Fig. 1) are visible, although the magnitudes are too low compared to the data.

Topex-Poseidon altimetry data contain several gaps and are displayed in the lowest panel. Additionally to the experiments with initial conditions from Tanioka et al. (2006) (red line), the green line displays the results from an alternative initial condition taken from Banerjee et al. (2007). This experiment does not show the double peak structure; however, the location of the leading wave crest is improved when compared to Topex-Poseidon data.

Envisat crossed the equator at around 4:14 AM UTC and thus registered the wave from approximately 3 h and 10 min until 3 h and 21 min after the earthquake. The middle panels display the comparison to a model snapshot after 3 h 20 min of integration. The reflected wave from Sri Lanka (10°N–15°N) is reproduced by the model although the leading crest to the south is too low.

Fig. 6 Time series (time in hours) of the sea surface elevation (in meters) in positions of the actual tide gauges displayed in Fig. 5. Starting point of the earthquake was assumed on 0:59 AM UTC on 26 December 2004

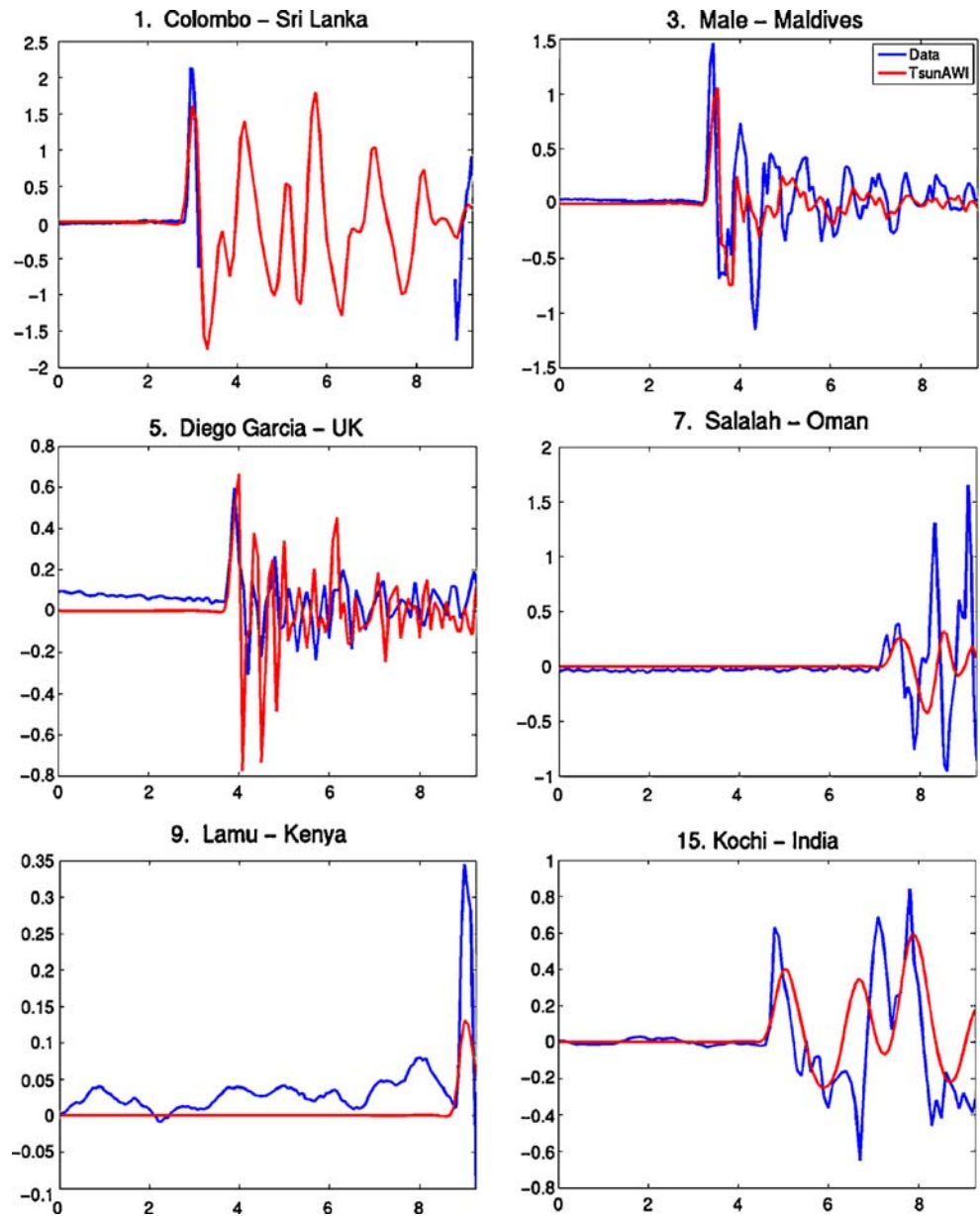
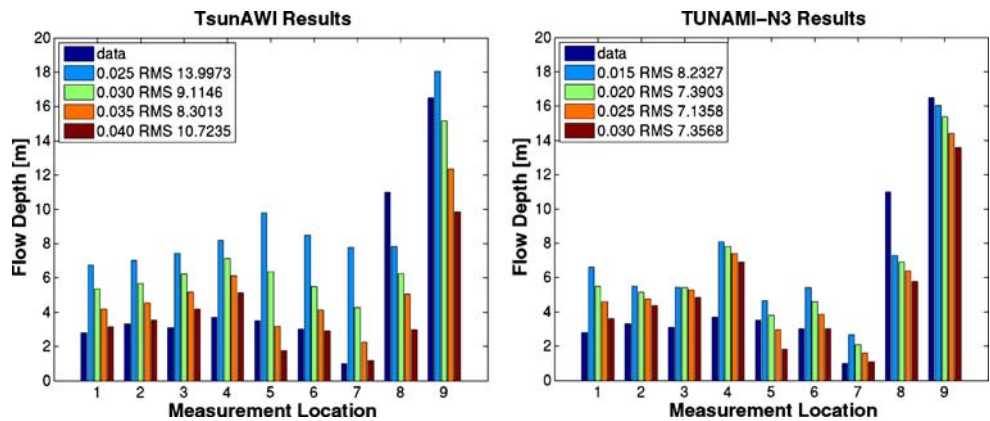


Fig. 7 Inundation results in comparison to field measurements in the nine locations shown in Fig. 8 for different values of Manning coefficient



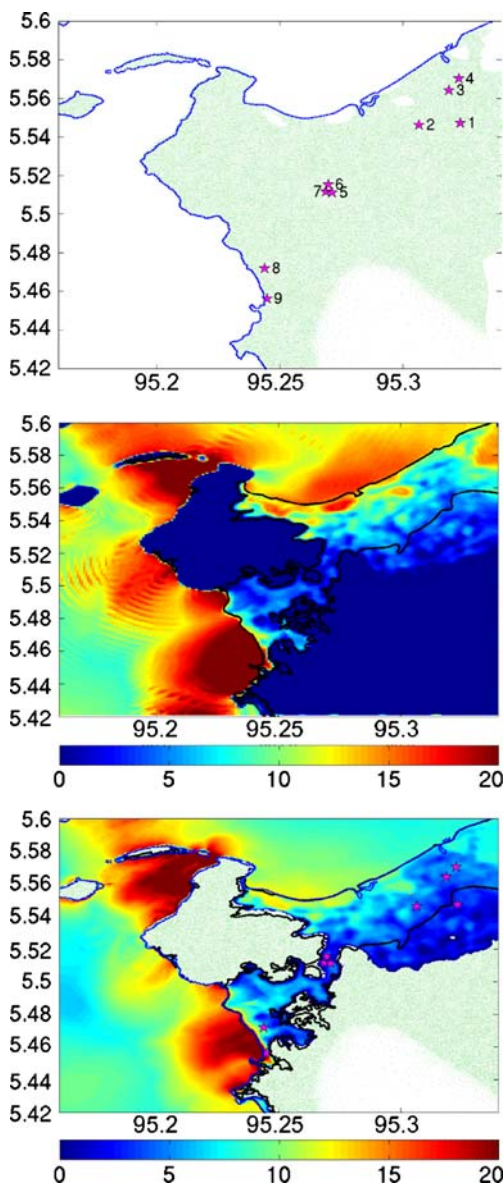


Fig. 8 Maximum flow depth obtained with TUNAMI-N3 (*middle panel*) and TsunAWI. The *black line* gives an approximation to the inundated area obtained by DLR from satellite images. The *upper panel* shows nine positions where field measurements of the flow depth are available from the survey teams ITST-1 and ITST-2. *Green dots* mark land nodes contained in the unstructured grid

4.1.3 Far field tsunami

Since the objective of the finite element modeling approach is a single grid for different scales, we compare the wave propagation to available tide gauge data recorded by stations in the whole Indian Ocean. A comprehensive overview of these stations and an analysis of the tsunami records is published in Rabinovich and Thomson (2007). For the present

study, we chose six stations in the Indian Ocean. These are shown in Fig. 5. The corresponding records of 26 December 2004 are partly available from the Survey of India & National Institute of Oceanography (see <http://www.nio.org/jsp/tsunami.jsp> and Nagarajan et al. (2006)) and partly from the University of Hawaii Sea Level Center within the Global Sea Level Observing System (GLOSS-IOC, for more information see <http://ilikai.soest.hawaii.edu/uhscl/>). The model results of the long time integrations are exemplified as time series in the tide gauge positions in Fig. 6. The arrival times in far away gauges agree very well with the tide gauge records. In these figures, the starting time of the earthquake is assumed at 26 December 2004, 0:59 AM UTC. The lower part of Table 3 compares measured arrival times to the model values and the largest error is 3.2%. In all of these gauges, the bathymetry is merely given by the GEBCO data set. Thus, the local bathymetry around the gauges is not very precise and an accurate wave form cannot be expected from these model runs. The arrival times are very similar to the results obtained from the numerical experiments in Kowalik et al. (2005) (page 54). The wave form of the first crest obtained in Colombo station coincides very well with the numerical results published in Tanioka et al. (2006).

4.2 Inundation

The actual wave heights that occurred in Aceh region were investigated by several survey teams in early 2005 (Borrero et al. 2006; Jaffe et al. 2006). The upper panel of Fig. 8 indicates the locations of field measurements in Aceh region. These values are results of the International Tsunami Survey Teams (ITST-1 and ITST-2) from January to April 2005 [results were presented in Kongko (2007) and published in Kongko et al. (2006)]. These locations contain measurements in the central part of Aceh (stations 5, 6, and 7) where waves from the northern part met those from the south during the tsunami event. All nine stations are not contained in Jaffe et al. (2006) and Borrero et al. (2006).

The measurements together with the corresponding model results are depicted in Fig. 7. Only the parameter of Manning bottom friction is varied between 0.015 and 0.04. Due to different implementations of horizontal diffusion and inundation algorithms in both models, the individual parameter range varies as well. The variations of inundation are larger in TsunAWI. The best agreement between model and data with respect to the RMS error is achieved for Manning friction set to 0.035 (0.025) in TsunAWI (TUNAMI-N3).

The total inundation area for the optimal friction coefficients in both models is shown in the lower panels of Fig. 8. In this figure, the boundary of the actual inundation area after the event is shown as well. This curve was derived from satellite images by the German Aerospace Agency (DLR). A similar inundation area was determined in McAdoo et al. (2007).

5 Conclusions

We simulated the tsunami event generated by the Sumatra–Andaman earthquake with a finite element model (TsunAWI) based on unstructured grids. The flexibility of this approach with respect to a variable resolution is used to resolve both the far field behavior of the tsunami with moderate resolution, as well as inundation processes in a very high resolution.

The model results were compared to available data, i.e., tide gauge records in stations all over the Indian Ocean, satellite altimetry, and field measurements in the inundation area. It turns out that the model is capable of simulating the far field wave, as well as the inundation with high accuracy.

The inundation results are compared to the results of a nested grid model (TUNAMI-N3) in a much smaller domain but similar resolution in the finest grid. Initial conditions were identically defined in a regular grid. However, interpolations to model grids lead to slightly different representations due to different grid resolutions.

Both models agree well in the crucial quantities arrival time of the first wave crest, magnitude of the first crest, and inundation area. Considerable differences occur in later stages of the wave propagation. These are largely influenced by reflections and depend heavily on the coastal representation in the lower resolved parts of the model domain. In this respect, the finite element approach is appropriate due to smooth transitions in resolution and consistent treatment of inundation boundary conditions along the whole coast, in contrast to the nested grid approach with inundation merely implemented in the finest grid.

With respect to CPU time, TsunAWI is still a factor of about 5–10 slower than TUNAMI. This does not apply to wall clock time, since TsunAWI employs OpenMP parallelism and is implemented on a multi-processor shared memory platform. Additionally, the model is yet to be optimized, and, especially in the framework of an early warning system, a consistent accuracy on the broad scale is a necessity. Since the early warning system developed within GITEWS is

based on precomputed scenarios, the CPU time is not the most relevant criterion.

Acknowledgements TsunAWI is the result of joint developments and collaboration of different groups within AWI, namely, Computing center lead by Wolfgang Hiller and the inverse modeling group supervised by Jens Schröter. We acknowledge contributions to code development by Alexey Androsov, Sergey Danilov, Dmitry Sein, and Dmitry Sidorenko. We acknowledge Ralph Kiefl from DLR for providing us with inundation areas estimated from satellite images and NOAA/PMEL/NOAA Center for Tsunami Research for providing satellite altimetry data. Widjo Kongko's contribution is especially acknowledged for providing us with field measurements from Aceh region and discussions on TUNAMI. The tide gauge data were provided by the Survey of India & National Institute of Oceanography and by the University of Hawaii Sea Level Center within the Global Sea Level Observing System (GLOSS-IOC). We acknowledge anonymous reviewers for helpful advice and for pointing us to some literature we were not aware of.

References

- Ammon JC, Ji C, Thio H, Robinson D, Ni S, Hjorleifsdottir V, Lay T, Das S, Helmberger D, Ichinose G, Polet J, Wald D (2005) Rupture process of the 2004 Sumatra–Andaman earthquake. *Science* 308:1133–1139
- Arcement Jr GJ, Schneider VR (1984) Guide for selecting manning's roughness coefficients for natural channels and flood plains. Technical report, US Department of Transportation. Federal Highways Administration Reports no. FHWA-TS-84-204, 62 p
- Banerjee P, Pollitz FF, Nagarajan B, Bürgmann R (2007) Co-seismic slip distributions of the 26 December 2004 Sumatra–Andaman and 28 March 2005 Nias earthquakes from GPS static offsets. *Bull Seismol Soc Am* 97(1A):S86–S102
- Behrens J (2008) TsunAWI—unstructured mesh finite element model for the computation of tsunami scenarios with inundation. In: Habip M (ed) Proceedings of NAFEMS seminar: simulation of complex flows (CFD)—applications and trends, number ISBN 978-1-874376-33-0. NAFEMS Contact DACH & Nordic Countries, Bernau, pp 6–1 ff
- Behrens J, Androsov A, Braune S, Danilov S, Harig S, Schröter J, Sein DV, Sidorenko D, Startseva O, Taguchi E (2007) TsunAWI technical documentation part I: mathematical, numerical and implementation concepts. Tsunami Project Documentation No. 004. Alfred-Wegener-Institut, Bremerhaven
- Borrero JC, Synolakis CE, Fritz H (2006) Northwest Sumatra field survey after the December 2004 great Sumatra earthquake and Indian Ocean tsunami. *Earthq Spectra* 22(S3):S93–S104
- Danilov S, Kivman G, Schröter J (2004) A finite element ocean model: principles and evaluation. *Ocean Model* 6(2):125–150
- Frey WH, Field DA (1991) Mesh relaxation: a new technique for improving triangulations. *Int J Numer Methods Eng* 31:1121–1133
- Furumoto AS, Tatehata H, Morioko C (1999) Japanese tsunami warning system. *Sci Tsunami Hazards* 17:85–105
- The General Bathymetric Chart of the Oceans (GEBCO) (2008) Webpage <http://www.gebco.net>

- Grilli ST, Ioualalen M, Asavanant J, Shi F, Kirby JT, Asce M, Watts P (2007) Source constraints and model simulation of the December 2006, 2004, Indian Ocean tsunami. *J Waterw Port Coast Ocean Eng* 133:414–428
- Geist EL, Titov V, Arcas D, Pollitz FF, Bilek SL (2007) Implications of the 26 December 2004 Sumatra–Andaman earthquake on tsunami forecast and assessment models for great subduction-zone earthquakes. *Bull Seismol Soc Am* 97:S249–S270
- Hayashi Y (2008) Extracting the 2004 Indian Ocean tsunami signals from sea surface height data observed by satellite altimetry. *J Geophys Res* 113, C01001:9
- Hanert E, LeRoux DY, Legat V, Deleersnijder E (2005) An efficient eulerian finite element method for the shallow water equations. *Ocean Model* 10:115–136
- Hirata K, Satake K, Tanioka Y, Kuragano T, Hasegawa Y, Hayashi Y, Hamada N (2006) The 2004 Indian Ocean tsunami: tsunami source model from satellite altimetry. *Earth Planets Space* 58:195–201
- Imamura F, Yalciner AC, Ozyurt G (2006) Tsunami modelling manual. April
- Jaffe BE, Borrero JC, Prasetya GS, Peters R, McAdoo B, Gelfenbaum G, Morton R, Ruggiero P, Hignman B, Dengler L, Eeri M, Hidayat R, Kingsley E, Kongko W, Lukijanto, Moore A, Titov V, Yulianto E (2006) Northwest Sumatra and offshore islands field survey after the December 2004 Indian Ocean tsunami. *Earthq Spectra* 22(S3):S105–S135
- Kongko W (2007) Investigation on tsunami run-up in Dec. 26 2004 Indian ocean earthquake. In: Presentation on the international workshop on recent developments in tsunami modeling, Bremerhaven, 25–27 April 2007
- Kongko W, Istiyanto DC, Irwandi I (2006) Tsunami modeling and field observations of December 26 2004 Indian ocean earthquake. Technical report, Coastal Dynamic Research Center, BPPT Indonesia
- Kowalik Z, Knight W, Logan T, Whitmore P (2005). Numerical modeling of the global tsunami: Indonesian tsunami of 26 December 2004. *Sci Tsunami Hazards* 23:40–56
- Kowalik Z, Knight W, Logan T, Whitmore P (2007) The tsunami of 26 December, 2004: numerical modeling and energy considerations. *Pure Appl Geophys* 164:379–393
- Le Roux DY, Lin CA (1998) Finite elements for shallow-water equation ocean models. *Mon Weather Rev* 126(7): 1931–1951
- Lynett PJ, Wu T-R, Liu PL-F (2002) Modeling wave runup with depth-integrated equations. *Coast Eng* 46:89–107
- McAdoo BG, Richardson N, Borrero J (2007) Inundation distances and run-up measurements from ASTER, quickbird and SRTM data, aceh coast, Indonesia. *Int J Remote Sens* 28:2961–2975
- NOA (2008) Tsunami event – December 26, 2004 Indonesia (Sumatra). Webpage http://nctr.pmel.noaa.gov/indo_1204.html
- Nagarajan B, Suresh I, Sundar D, Sharma R, Lal AK, Neetu S, Shenoj SSC, Shetye SR, Shankar D (2006) The great tsunami of 26 December 2004: a description based on tide-gauge data from the Indian subcontinent and surrounding areas. *Earth Planets Space* 58:211–215
- Okada Y (1985) Surface deformation due to shear and tensile faults in a half space. *Bull Seismol Soc Am* 75: 1135–1154
- Park, J, Anderson K, Aster R, Butler R, Lay T, Simpson D (2005) Global seismographic network records the great Sumatra–Andaman earthquake. *EOS* 86:57, 60–61
- Piatanesi A, Tinti S, Bortolucci E (1999) Finite-element simulations of the 28 December 1908 Messina straits (Southern Italy) tsunami. *Phys Chem Earth (A)* 24: 145–150
- Rabinovich AB, Thomson RE (2007) The 26 December 2004 Sumatra tsunami: analysis of tide gauge data from the world ocean part 1. Indian ocean and south Africa. *Pure Appl Geophys* 164:261–308
- Shewchuk JR (1996) Triangle: engineering a 2d quality mesh generator and delaunay triangulator. In: Lin MC, Manocha D (eds) *Applied computational geometry: towards geometric engineering*. Springer, Heidelberg, pp 203–222
- Stacey MW, Nowak ZP (1995) A numerical model of the circulation in Knight inlet, British Columbia, Canada. *J Phys Oceanogr* 25:1037–1062
- SRT (2007) Shuttle radar topography mission X-SAR / SRTM. Webpage <http://www.dlr.de/srtm/>
- Tinti S, Gavagni I (1995) A smoothing algorithm to enhance finite-element tsunami modelling: an application to the 5 February 1783 calabrian case, Italy. *Nat Hazards* 12: 161–197
- Titov VV, Gonzalez FI (1997) Implementation and testing of the method of splitting tsunami (MOST) model. Technical Report 1927, NOAA Technical Memorandum ERL PMEL-112
- Titov V, Rabinovich AB, Mofjeld HO, Thomson RE, Gonzalez FI (2005) The global reach of the December 2004 Sumatra tsunami. *Science* 309:2045–2048
- Tanioka Y, Yudhicara, Kususeso T, Kathirolu S, Nishimura Y, Iwasaki S-I, Satake K (2006) Rupture process of the 2004 great Sumatra–Andaman earthquake estimated from tsunami waveforms. *Earth Planets Space* 58:203–209
- Walters RA (2006) Design considerations for a finite element coastal ocean model. *Ocean Model* 15:90–100
- Walters RA, Goff J (2003) Assessing tsunami hazard along the New Zealand coast. *Sci Tsunami Hazards* 21:137–153

A multigrid methodology for assimilation of measurements into regional tidal models

Oleg G. Logutov

Received: 14 May 2008 / Accepted: 2 November 2008 / Published online: 20 November 2008
© Springer-Verlag 2008

Abstract This paper presents a rigorous, yet practical, method of multigrid data assimilation into regional structured-grid tidal models. The new inverse tidal nesting scheme, with nesting across multiple grids, is designed to provide a fit of the tidal dynamics to data in areas with highly complex bathymetry and coastline geometry. In these areas, computational constraints make it impractical to fully resolve local topographic and coastal features around all of the observation sites in a stand-alone computation. The proposed strategy consists of increasing the model resolution in multiple limited area domains around the observation locations where a representativeness error is detected in order to improve the representation of the measurements with respect to the dynamics. Multiple high-resolution nested domains are set up and data assimilation is carried out using these embedded nested computations. Every nested domain is coupled to the outer domain through the open boundary conditions (OBCs). Data inversion is carried out in a control space of the outer domain model. A level of generality is retained throughout the presentation with respect to the choice of the control space; however, a specific example of using the outer domain OBCs as the control space is provided, with other sensible choices discussed. In the forward scheme, the computations in the nested domains do not affect the solution in the outer domain.

The subsequent inverse computations utilize the observation-minus-model residuals of the forward computations across these multiple nested domains in order to obtain the optimal values of parameters in the control space of the outer domain model. The inversion is carried out by propagating the uncertainty from the control space to model tidal fields at observation locations in the outer and in the nested domains using efficient low-rank error covariance representations. Subsequently, an analysis increment in the control space of the outer domain model is computed and the multigrid system is steered optimally towards observations while preserving a perfect dynamical balance. The method is illustrated using a real-world application in the context of the Philippines Strait Dynamics experiment.

Keywords Tidal modelling · Inverse methods · Nesting

1 Introduction

A significant number of tidal observations are collected along coasts and in inland waterways with complex coastline geometry and bottom topography. Local small-scale coastal and topographic features are often consequential for the tidal fields measured at coastal locations. For example, coastal measurements in (semi)enclosed bays and estuaries might not be representative of the nearby open-ocean areas, but rather reflect the local characteristics of the shoreline. On the other hand, the computational constraints limit model resolution and often lead to an insufficient representation of small-scale coastal and topographic features

Responsible Editor: Eric Deleersnijder

O. G. Logutov (✉)
Massachusetts Institute of Technology,
77 Massachusetts Ave., rm. 7-321,
02139 Cambridge, MA, USA
e-mail: logutov@mit.edu

in areas where measurements are obtained, such as estuaries, channels, and sills. This presents a difficulty with respect to the assimilation of measurements into regional tidal models. The measurements utilized for assimilation must be representative of the tidal dynamics at a model resolution. An artificial steering of the solution towards unresolved observations can be counter-productive and lead to a degradation of model accuracy. With the unstructured grid tidal models, the mesh resolution can be selectively adjusted around all of the observation locations to fully resolve important coastal features. The structured grid tidal models are less conducive to selective resolution adjustments and, therefore, part of the observational network might be unresolved. A popular way of avoiding the difficulty within the structured grid modeling framework is either an exclusion of the unresolved measurements from the data set or an inflation of the representativeness error covariance around coastal segments unresolved in a model. Both approaches lead to a loss of information from part of the observational network.

The present paper describes a new multigrid data assimilation scheme for regional tidal modeling applications that provides a rigorous way of reducing the representativeness error by employing a multigrid, data-assimilative framework. The representativeness error can be detected by analyzing the observation-minus-forecast residuals and their sensitivity to model resolution. Adaptive methods for data assimilation reviewed in Lermusiaux (2007) can potentially be useful for distinguishing the error of representativeness due to model and bottom topography resolutions vs that due to model formulation. If data-model misfits are found to be consistently larger than the average in certain areas of model domain, the representativeness component in these misfits can be analyzed through refinements to model grid resolution. If misrepresentation due to insufficient resolution is detected, it can be dealt with in the data assimilative component of the modeling system by using the method presented in this paper. High-resolution nested domain(s), coupled to the outer model domain via the open boundary conditions (OBCs), can be setup around the problematic areas and the assimilation can be carried out using these nested computation(s). Thus, with the use of the nested domains, coastal measurements can be assimilated into a modeling system consistently with the resolution. A real-world example of such an application in the context of the Philippines Straits dynamics experiment (PhilEx) is discussed in Section 5 as a demonstration.

An extensive variety of methods exist to constrain regional barotropic tidal estimates with the observational tidal elevation and velocity data (Egbert

and Bennett 1996; Robinson and Lermusiaux 2001; Lermusiaux et al. 2006b). The techniques differ by the choice of an optimization space and by the specifics of implementation. An early important contribution included a variational formulation developed by Bennett and McIntosh (1982) which led to the adjoint techniques of generalized inverse tidal solution, subsequently utilized in a variety of tidal applications (e.g. McIntosh and Bennett 1984; Egbert 1997; Muccino et al. 2008). The representer method (Bennett 1992, 2002) and its reduced-basis alternatives (Egbert and Erofeeva 2002) have been developed following this approach and successfully applied to assimilation of Topex/Poseidon altimeter data (Egbert et al. 1994), estimation of the internal tides (Kurapov et al. 2003), and in other coastal ocean modeling applications (Muccino et al. 2008). The aforementioned methods have one feature in common, namely, an optimization is carried out in the (reduced) data space and requires an adjoint tidal model. Alternative techniques that do not require an adjoint model have also been proposed. These include the steady-state Kalman filter (KF) schemes (Heemink and Kloosterhuis 1990; Sorensen and Madsen 2004) that carry out an optimization in the model state space and rely on reduced-rank error covariance models, defined in the model state space, evolved using stochastic linearized shallow water equations. The KF-based schemes were utilized in a series of realistic two-dimensional inverse problems arising from the assimilation of water level measurements into a regional model of the North Sea (Sorensen and Madsen 2004) and a storm surge prediction model of Danish coastal waters (Canizares et al. 2001). Lynch and Hannah (1998), Xu et al. (2001), and Logutov and Lermusiaux (2008) have developed an inverse scheme that carries out an optimization in the OBC space, also without the need of an adjoint model. The method was specifically developed for regional tidal modeling applications and successfully applied near the Georges Bank, near the Newfoundland and South Labrador shelves, and off the coast of Vancouver Island and in the Hood Canal and Dabob Bay region of WA (Xu et al. 2008), as well as in other regions (Haley et al. 2008). Nudging and optimal interpolation schemes have been proposed and applied in some tidal modeling systems, e.g., Navy PCTides system (Hubbert et al. 2001). Adjoint schemes that utilize bottom topography or model parameters as the control space have also been proposed (e.g., Das and Lardner 1992) but not yet fully demonstrated in real-world applications.

A variety of structured and unstructured grid tidal models are currently in use across various organizations (Davies et al. 1997). The finite-difference

(*e.g.*, LeProvost and Vincent 1986; Davies 1993), the finite-element (*e.g.* Jones and Davies 1996; Walters 2005; Greenberg et al. 2005), and the structured nonorthogonal curvilinear coordinate (*e.g.* George 2007) schemes have been proposed and developed into stand-alone computational systems for regional tidal modeling. Jones and Davies (2005) provide intercomparison between the finite-difference and the finite-element schemes using an example of tidal modeling in the Irish sea. An advantage of the unstructured grid approach is efficiency in representing complex boundary geometries and the potential of employing higher-order approximation schemes, for example, the high-order discontinuous Galerkin method discussed in Bernard et al. (2008). As a result, the unstructured grid tidal models can be made representative of tidal measurements even in areas with complex boundary geometries. However, the generation of unstructured grids, especially for time domain models, is a difficult and time-consuming task that requires considerable human intervention and experience (Blain et al. 2002). In addition, the unstructured grid models are more prone to contamination of the solution with spurious gravity-wave modes, while the finite-difference schemes can be easily made to contain no spurious modes (Walters 2005). For these and other reasons, structured-grid tidal models remain an important component of tidal prediction in both operational and research settings. As an example, the US Navy tidal prediction system currently consists of a time domain finite-element barotropic tidal model ADCIRC (Luettich et al. 1992) and a finite-difference spectral model PCTides (Hubbert et al. 2001). The finite-element system is limited to a number of preexisting computational meshes and is not as easily relocatable as the PCTides system, which provides a rapidly relocatable capability and can be quickly exercised in any region of the world (Blain et al. 2002).

The purpose of this article is to present a rigorous, yet practical, method of data assimilation into regional structured grid tidal models that allows for the use of nested domain(s) resolving the tidal dynamics near the observations that are otherwise unresolved and contain a representativeness error. The computational constraints often make it impractical to fully resolve local topographic and coastal features around all of the observation sites in stand-alone structured grid models. A general discussion of the sensitivity of regional tidal solutions to mesh resolution is provided by Jones and Davies (2007a, b), Walters (2005), Jones and Davies (1996), and LeProvost et al. (1994). The proposed strategy consists of increasing the model resolution in limited-area nested domains in order to improve the

representation of the measurements with respect to the dynamics. In our method, the nested domains are coupled to the outer domain via the OBCs. In the forward system, the computations in the nested domains do not affect the solution in the outer domain, while in the inverse system, observation-minus-forward model residuals computed in the nested domains affect the solution in the outer domain. Data inversion is carried out in a control space of the outer domain model. An adjoint model is not needed in the method because of the low-rank formulation adopted for the representation of the error covariance of the control parameters. The presented approach is tailored towards the applications where the use of strong model constraints is desirable. The multigrid system is steered towards observations via the control parameters of the outer domain model (*e.g.*, OBCs of the outer model) and the inverse solution preserves a perfect dynamical balance. An alternative strategy and a sequential scheme for weakly constrained estimation across nested domains using a KF approach is presented in Barth et al. (2007). The authors have proposed an analysis scheme that uses a single multivariate state vector comprised of the variables from multiple nested models that leads to a reduction of the inconsistencies along the nesting boundaries.

The presence of the representativeness error can, in general, be detected through the sensitivity experiments with model resolution. If the observation-minus-model residuals have a consistent bias in certain areas of the domain and this bias is found sensitive to model resolution, the representativeness error can be suspect. In this case, high-resolution nested domain(s) can be setup around such areas and the assimilation in those areas can be carried out using the nested computation(s). Thus, the representativeness error can be reduced and coastal measurements can be assimilated into a multigrid modeling system more consistent with the resolution.

The organization of this paper is as follows. In Section 2, the model equations, both in continuous and discrete formulations, are described. The model formulation presented is a specific example of the dynamical constraints of the inverse scheme. It is provided here only as an illustration for clarity. Importantly, the method can be extended to use with other structured grid tidal models. Section 3 presents the proposed multigrid assimilation method, with a subsequent discussion in Section 4. Finally, an illustration of the methodology in the context of the Philippines Strait dynamics experiment is presented in Section 5. The goal in Section 5 is not a comprehensive description of tidal dynamics in the Philippines basin but rather an

illustration and an analysis of the described method, as applied in realistic settings. The [Appendix](#) outlines the index notation utilized in Section 2.

2 Model formulation

2.1 Dynamical equations

The governing equations that we utilize as the dynamical constraints of the inverse scheme consist of the shallow water equations forced through the OBCs. In coastal applications, the self-attraction and loading terms, as well as the direct astronomical forcing, are negligible as compared to the open boundary forcing and can be omitted from the equations (Snyder et al. 1979). The development presented here consists of the depth-averaged equations. The 2-D formulation allows us to remove a layer of complexity related to parametrization of the vertical divergence of Reynolds stresses in the momentum equations required in the 3-D models. However, an extension to a 3-D formulation is straightforward if needed. In differential form, the governing equations are expressed in spherical coordinates (λ, ϕ) , where λ and ϕ are the longitude and the latitude, respectively, using the linearized barotropic shallow water equations

$$\frac{\partial}{\partial t} \eta + \frac{1}{a \cos \phi} \frac{\partial}{\partial \lambda} (Hu) + \frac{1}{a \cos \phi} \frac{\partial}{\partial \phi} (Hv \cos \phi) = 0$$

$$\frac{\partial}{\partial t} u - fv + \kappa u = \frac{-g}{a \cos \phi} \frac{\partial}{\partial \lambda} \eta \quad (1)$$

$$\frac{\partial}{\partial t} v + fu + \kappa v = \frac{-g}{a} \frac{\partial}{\partial \phi} \eta$$

subject to the zero normal flow condition

$$\mathbf{n} \cdot \mathbf{u} \Big|_{\partial \Omega_C} = 0 \quad (2)$$

at closed boundaries, and the Dirichlet conditions

$$\eta \Big|_{\partial \Omega_O} = \eta_{obc} \quad (3)$$

at open boundaries. In Eqs. 1–3, η and $\mathbf{u} = (u, v)$ denote tidal elevations and zonal and meridional velocity components, respectively. H is the undisturbed water depth; g , f , and a are the acceleration due to gravity, the Coriolis parameter, and the Earth radius; and κ is the bottom friction parameter. In the subinertial regime (e.g., diurnal tidal constituents in high latitudes),

the depth-integrated shallow water equations (Eq. 1) support coastal-trapped topographic wave response to open boundary forcing and, therefore, errors in the Dirichlet conditions can sometimes cause generation of spurious topographic waves near the adjacent coastal boundaries. Radiation OBCs with external sea surface height data (see an overview by Blayo and Debreu 2005) can then be applied at a subset of open ocean boundaries affected by the errors to effectively remove spurious coastal trapped waves. In the super-inertial regime, the dynamical equations do not support coastal-trapped topographic wave response to errors in the Dirichlet OBCs and condition 3 is always sufficient. For example, in an application presented in Section 5 of this paper, both the semidiurnal and diurnal tidal constituents are super-inertial and, therefore, the Dirichlet condition 3 was used exclusively at all open boundaries for all tidal constituents. The bottom friction term in the momentum equations could be specified following any parameterization of modeler's choice. In this work, the bottom friction is specified following a quadratic parameterization and an iterative approximation method. The bottom friction coefficient is

$$\kappa = \frac{C_D |\mathbf{u}_{(0)}|}{H} \quad (4)$$

where $\mathbf{u}_{(0)}$ is the velocity field obtained by running the model with a fixed value of κ . The nondimensional bottom drag coefficient C_D is a tunable parameter, with values $C_D = 0.002 - 0.003$ typically suggested (e.g. Grenier et al. 1995). Note that $\kappa = \kappa(\lambda, \phi)$ is a spatially varying field.

The kinematic and dissipative nonlinearities in the dynamical equations are capable of introducing new frequencies in tidal spectrum through energy transfer from the astronomical tidal constituents to higher and lower frequency harmonics. In this manner, the overtides, compound tides, and low-frequency tides, with frequencies given by the sums and the differences of the astronomical constituent frequencies, can be created. The nonlinear tidal effects can be simulated by introducing the *shallow-water* tidal constituents, which compensate for the nonlinearities and for wave–wave interactions between the astronomical constituents (LeProvost et al. 1981; Andersen et al. 2006). As opposed to the astronomical constituents defined by a unique set of Doodson numbers, the shallow-water constituents have frequencies determined by the multiples of the astronomical constituent frequencies (overtides) or by sums and differences of frequencies of interacting astronomical constituents (compound tides). For example, the nonlinearity in the continuity equation

manifested through the interaction of the M_2 and the S_2 astronomical constituents gives rise to the MS_4 shallow-water constituent. The nonlinear effects generating the shallow-water constituents can be treated within the linearized modeling framework using a perturbation method described in Snyder et al. (1979) and LeProvost et al. (1981). In this manner, the nonlinear effects can be included following a rigorous approach, and a linearized numerical scheme can be formulated for each astronomical or shallow-water constituent.

By searching for solution of Eqs. 1–3 in the form

$$\{\eta, u, v\}(\lambda, \phi, t) = \mathbb{R} \left\{ \sum_{k=1}^K \{\tilde{\eta}_k, \tilde{u}_k, \tilde{v}_k\}(\lambda, \phi) \exp i\omega_k t \right\}, \tag{5}$$

where ω_k denotes frequencies of the astronomical and shallow-water tidal constituents and $\tilde{\eta}_k, \tilde{u}_k,$ and \tilde{v}_k are complex spatially varying fields; we obtain a spectral representation, with a boundary value problem

$$i\omega_k \tilde{\eta}_k + \nabla \cdot \begin{bmatrix} H\tilde{u}_k \\ H\tilde{v}_k \end{bmatrix} = 0 \tag{6}$$

$$\begin{bmatrix} \tilde{u}_k \\ \tilde{v}_k \end{bmatrix} = \frac{-g}{(i\omega_k + \kappa)^2 + f^2} \begin{bmatrix} i\omega_k + \kappa & f \\ -f & i\omega_k + \kappa \end{bmatrix} \nabla \tilde{\eta}_k, \tag{7}$$

defined for each individual tidal constituent k , subject to the boundary conditions

$$\tilde{\eta}_k \Big|_{\partial\Omega_o} = \zeta_k(x, y) \tag{8}$$

at open boundaries and

$$\mathbf{n} \cdot \begin{bmatrix} \tilde{u}_k \\ \tilde{v}_k \end{bmatrix} \Big|_{\partial\Omega_c} = 0 \tag{9}$$

at closed boundaries. Equations 6–9 can be solved numerically, for example, on an Arakawa-C staggered grid. An outline of the numerical implementation pertinent to the inverse methodology of this paper is discussed next.

2.2 Discrete operators and representations

Equations 6–9 discretized on a structured grid lead to a sparse system of equations for a gridded vector η of tidal sea surface elevations

$$\mathbf{A}_{\eta \leftarrow \eta} \eta = \mathbf{0}, \tag{10}$$

where matrix $\mathbf{A}_{\eta \leftarrow \eta}$ is a discrete version of the operator

$$\mathcal{A}\{\cdot\} = \left(i\omega_k - \nabla \cdot \frac{gH}{(i\omega_k + \kappa)^2 + f^2} \times \begin{bmatrix} i\omega_k + \kappa & f \\ -f & i\omega_k + \kappa \end{bmatrix} \nabla \right). \tag{11}$$

The numerical details of assembling the sparse system matrix $\mathbf{A}_{\eta \leftarrow \eta}$ on an Arakawa-C grid are provided by Logutov and Lermusiaux (2008). In their scheme, $\mathbf{A}_{\eta \leftarrow \eta}$ is nine-diagonal. In the notation followed throughout this paper, the subscripts are utilized to indicate mappings provided by the discrete operators between various staggered grids or variables. For example, the finite-difference operators implementing the gradient differential operator, $g\nabla\{\cdot\}$, utilized in Eq. 7, can be denoted as $\mathbf{G}_{u \leftarrow \eta}, \mathbf{G}_{v \leftarrow \eta}$, which indicates that the matrices $\mathbf{G}_{u \leftarrow \eta}$ and $\mathbf{G}_{v \leftarrow \eta}$ act on gridded vectors defined at an η grid and provide mappings to u and v Arakawa-C grids, respectively. Similarly, the finite-difference operators implementing the divergence, $(\nabla \cdot)$, can be consistently denoted as $\mathbf{D}_{\eta \leftarrow u}$ and $\mathbf{D}_{\eta \leftarrow v}$ to indicate that the vectors defined at u - and v -staggered grids constitute the domain of these operators, respectively, while the vectors defined at an η grid constitute the range of these operators. Following this notation, the tidal velocity vectors \mathbf{u} and \mathbf{v} are obtained from the tidal elevation vector η using Eq. 7

$$\begin{aligned} \mathbf{u} &= \mathbf{U}_{u \leftarrow \eta} \eta \\ \mathbf{v} &= \mathbf{U}_{v \leftarrow \eta} \eta, \end{aligned} \tag{12}$$

where the linear operators $\mathbf{U}_{u \leftarrow \eta}$ and $\mathbf{U}_{v \leftarrow \eta}$ are sparse matrices, assembled following a finite-difference scheme of modeler’s choice for the gradient and the matrix multiplication acting on the gradient of tidal elevations, $\tilde{\eta}_k$, in Eq. 7.

Further, let vector $\zeta \in \mathbb{C}^{n_s}$ denote a control space of a tidal model, that is, the partition of the state-space and/or tunable parameters that we would like to utilize to steer the model fields towards observations as part of data assimilation procedures. The exact composition of the control space ζ might vary across different tidal modeling systems and applications. For example, the control space might include the OBCs. Alternatively, a parameterized increment to model bathymetry or bottom friction parameters can be utilized, among other options. To provide a level of generality, we leave the choice of the control space to the discretion of the modeler and only require that a model can be linearized with respect to the control parameters for the purposes of data assimilation. This latter requirement amounts to

an assumption that the control space is separable from the state-space following

$$\mathbf{A}_{(\mathbf{x} \leftarrow \mathbf{x})} \mathbf{x} = \mathbf{B}_{(\mathbf{x} \leftarrow \zeta)} \zeta, \tag{13}$$

where $\mathbf{x} \in \mathbb{N}^{n_x}$ is the model state-space. In order to provide a specific example of how Eq. 13 is obtained, suppose that the OBCs are chosen as the control space. The sea surface elevation vector $\eta \in \mathbb{C}^N$ in Eq. 10 contains active, masked, and open boundary grid-nodes. Let the indices of active and open boundary η -points be denoted as $\mathbf{i}_x \in \mathbb{N}^{n_x}$ and $\mathbf{i}_\zeta \in \mathbb{N}^{n_\zeta}$, respectively, following the index notation in the Appendix. We can separate the partitions of the system matrix $\mathbf{A}_{\eta \leftarrow \eta}$ corresponding to active and open boundary nodes as

$$\mathbf{A}_{(\mathbf{x} \leftarrow \mathbf{x})} \equiv (\mathbf{A}_{\eta \leftarrow \eta})_{\mathbf{i}_x, \mathbf{i}_x}, \quad \mathbf{A}_{(\mathbf{x} \leftarrow \eta_\zeta)} \equiv (\mathbf{A}_{\eta \leftarrow \eta})_{\mathbf{i}_x, \mathbf{i}_\zeta}. \tag{14}$$

In the foregoing, the index notation in the Appendix is followed. With these partitions, the computational discretization of the dynamical equations (Eq. 10) at active grid-points is given by

$$\mathbf{A}_{(\mathbf{x} \leftarrow \mathbf{x})} \mathbf{x} = -\mathbf{A}_{(\mathbf{x} \leftarrow \zeta)} \zeta \tag{15}$$

where vector ζ contains the values of the OBCs. The right-hand side in Eq. 15 represents ocean open boundary forcing and the overall system is a specific form of Eq. 13, with the OBCs providing the control space. Alternatively, the bottom friction parameters can be added to the control space. With such a choice, Eq. 7 needs to be linearized with respect to $\kappa(\lambda, \phi)$ and the first variation of the dynamical equations with respect to κ needs to be included in the right-hand side of Eq. 13. By seeking to control the model through a chosen set of the control parameters, the inverse solution is fitted to regional observations in accord with the dynamics. After assimilation, the interior and the boundary inverse tidal fields satisfy the barotropic shallow water dynamical equations exactly.

System 13 provides the dynamical constraints for the inverse method. At high resolution, Eq. 13 is solved using iterative methods. We use the preconditioned conjugate gradient method, with the left and the right preconditioners obtained by computing an incomplete LU decomposition of $\mathbf{A}_{(\mathbf{x} \leftarrow \mathbf{x})}$. Such preconditioning leads to an acceleration of convergence of the method and allows for a state-space with a dimensionality of up to $\mathcal{O}(10^6)$. The solution of Eq. 13 will hereby be formally denoted as

$$\mathbf{x} = \mathbf{M}_{(\mathbf{x} \leftarrow \zeta)} \zeta, \tag{16}$$

where

$$\mathbf{M}_{(\mathbf{x} \leftarrow \zeta)} = \mathbf{A}_{(\mathbf{x} \leftarrow \mathbf{x})}^{-1} \mathbf{B}_{(\mathbf{x} \leftarrow \zeta)}. \tag{17}$$

In practice, the inverse $\mathbf{A}_{(\mathbf{x} \leftarrow \mathbf{x})}^{-1}$ is never formed explicitly. Equation 17 merely represents an iterative solution of a linear system $\mathbf{A}_{(\mathbf{x} \leftarrow \mathbf{x})} \mathbf{X} = \mathbf{Y}$ for the unknowns $\mathbf{X} \in \mathbb{C}^{n_x \times p}$ and the right-hand side $\mathbf{Y} = \mathbf{B}_{(\mathbf{x} \leftarrow \zeta)} \mathbf{Z}_\zeta$, ($\mathbf{Z}_\zeta \in \mathbb{C}^{n_\zeta \times p}$).

3 Methodology

3.1 Setup and logistics

The methodology presented in this paper is designed to employ nested domains around coastal segments where tidal observations are collected but model resolution is insufficient to fully resolve the tidal dynamics. The outer model domain will be referred to as the *control* domain throughout this paper, consistent with the definition of the control parameters in that domain. The variables and the parameters defined in the control domain are labeled throughout using the subscript 0. The nested model domains are setup within the control domain and, without loss of generality, are assumed to be nonoverlapping. To maintain generality, it is assumed that there are m nested domains. The variables and the parameters defined in the nested domains will be labeled throughout using the subscripts $i \in \{1, 2, \dots, m\}$. Figure 1 presents a schematic of the

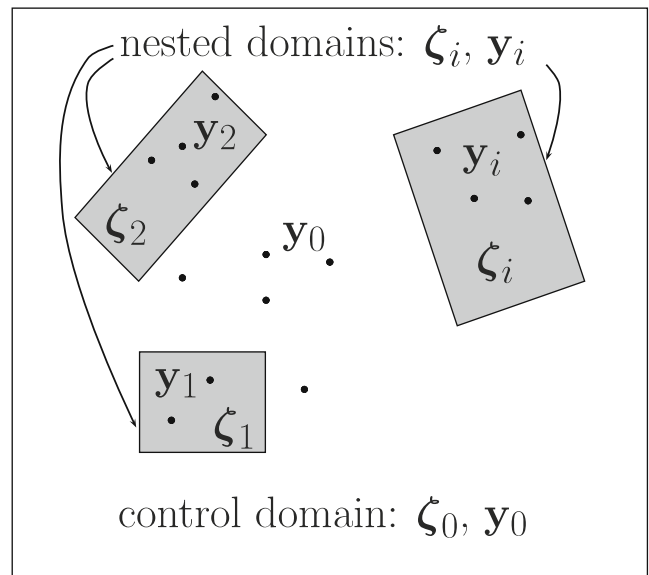


Fig. 1 Schematic of the control and nested domains. The data assimilation scheme tunes the control parameters ζ_0 , defined in the control domain, based on the observational data y_0 , y_1, y_2, \dots, y_m . The data-model misfits for $y_i, i \in \{1, 2, \dots, m\}$, are obtained based on the nested model computations. A nested domain i is coupled to the outer domain via the OBCs, ζ_i

control and the nested domains and the observations. The control space ζ is defined in the outer domain and, therefore, will be denoted as ζ_0 . The nested domains are coupled to the control domain through the OBCs, $\zeta_i, i \in \{1, 2, \dots, m\}$. In the forward problem, the control domain provides the OBCs for the nested domains but the computations in the nested domains do not affect the computation in the control domain. In the inverse problem, the computations in the nested domains affect the solution in the control domain through adjustments made to ζ_0 based on the data-nested model misfits. The goal of this section is to provide an optimal scheme for tuning the control parameters, ζ_0 , based on the observation-minus-forecast residuals in the outer and the nested domains.

3.2 Observational data

Tidal observations may come in the form of sea surface height (SSH) and velocity measurements. Such measurements are obtained from a variety of instruments and platforms, such as moorings, coastal tide gauges, bottom mounted tide gauges, and satellite altimetry, among others. The spectral modeling framework described in Section 2 assumes that the observations are harmonically analyzed and converted to harmonic amplitudes $\{\tilde{\eta}_k, \tilde{u}_k, \tilde{v}_k\}$ for each individual tidal constituent. Details of such a conversion are provided in Logutov and Lermusiaux (2008). The spectral domain data are related to the time domain variables via Eq. 5. Let vectors $\boldsymbol{\eta}_{obs} \in \mathbb{C}^{n_{obs}^\eta}$, $\boldsymbol{u}_{obs} \in \mathbb{C}^{n_{obs}^u}$, and $\boldsymbol{v}_{obs} \in \mathbb{C}^{n_{obs}^v}$ denote the observed values of the tidal elevations $\tilde{\eta}_k$ and the zonal and meridional barotropic tidal velocities \tilde{u}_k and \tilde{v}_k at observation locations. The total observational vector $\mathbf{y} \in \mathbb{C}^{n_{obs}}$ obtained by concatenation

$$\mathbf{y} = \begin{bmatrix} \boldsymbol{\eta}_{obs} \\ \boldsymbol{u}_{obs} \\ \boldsymbol{v}_{obs} \end{bmatrix}_{n_{obs}} \tag{18}$$

is collected throughout the control domain including the areas covered and not covered by the nested domains, as illustrated in the schematic in Fig. 1. The nested domains are setup at a high resolution sufficient to ensure that the collected measurements are representative of the tidal dynamics of the nested models. By permutating the order of entries in \mathbf{y} , we hereby introduce the partitions of the observational data, $\mathbf{y} = \{\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_m\}$, based on their location with respect to the nested domains. The observational data inside the i th nested domain is denoted as \mathbf{y}_i , while

the observational data inside the control domain but outside of any of the nested domains is denoted as \mathbf{y}_0 (Fig. 1)

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_m \end{bmatrix}_{n_{obs}}, \tag{19}$$

with the ordering of the SSH and the velocity measurements in each \mathbf{y}_i given in Eq. 18. Errors associated with measurements can typically be assumed independent between observation locations leading to a diagonal observational error covariance model. We will here assume only that the observational error covariance has a block-diagonal structure, with the measurement error uncorrelated across the domains

$$\mathbf{R} \equiv \mathcal{E} \left\{ \boldsymbol{\epsilon}_y \boldsymbol{\epsilon}_y^H \right\} = \begin{bmatrix} \mathbf{R}_0 & & & \\ & \mathbf{R}_1 & & \\ & & \ddots & \\ & & & \mathbf{R}_m \end{bmatrix}_{n_{obs} \times n_{obs}}. \tag{20}$$

Each block \mathbf{R}_i can be specified based on any practical considerations with respect to measurement error in each domain or estimated using the methods of error covariance parameter estimation from data-model misfits (e.g., Dee 1995). Finally, let $\mathbf{H}_{\mathbf{y}_i \leftarrow \mathbf{x}_i}$ denote an observational operator relating the model state-space, \mathbf{x}_i , in the i th model domain to the observation vector in that domain. In its simplest form, an observational operator represents an interpolation from a model grid onto the observation locations. For the velocity measurements, an observational operator includes the linear mappings $\mathbf{U}_{u \leftarrow \eta}$ and $\mathbf{U}_{v \leftarrow \eta}$ defined in Eq. 12, in addition to interpolation. Following the notation, data-model misfits corresponding to a tidal solution obtained in each model domain are given by

$$\mathbf{d}_i = \mathbf{y}_i - \mathbf{H}_{\mathbf{y}_i \leftarrow \mathbf{x}_i} \mathbf{x}_i, \tag{21}$$

where \mathbf{x}_i denotes the model state-space in the i th model domain, with the control domain counted as the 0th model domain.

3.3 Forward model error covariance

The control parameters ζ_0 can be assigned some a priori values. For example, the a priori values of the OBCs can be specified from a global tidal model. Let $\hat{\zeta}_0$ denote the a priori values of the control parameters, with the corresponding error covariance \mathbf{P}_0 . Theoretical statements about the sources of error in $\hat{\zeta}_0$ combined with practical considerations lead to a specific

form of \mathbf{P}_0 . In practice, \mathbf{P}_0 is not well known and merely reflects the degree of accuracy that a modeler associates with the a priori estimate $\hat{\boldsymbol{\zeta}}_0$. Given the lack of objective information about \mathbf{P}_0 , it can be sufficiently expressed using a low-rank approximation

$$\mathbf{P}_0 = \mathbf{Z}_0 \mathbf{Z}_0^H, \tag{22}$$

with $\mathbf{Z}_0 \in \mathbb{C}^{n_\zeta \times p}$ containing p dominant orthogonal control parameter error subspaces. The error subspaces \mathbf{Z}_0 can be obtained, for example, via a singular value decomposition (svd) of a valid parametric covariance form \mathbf{G}_0 ,

$$\mathbf{G}_0 = \mathbf{U}_0 \boldsymbol{\Lambda}_0 \mathbf{U}_0^T, \quad \mathbf{Z}_0 \equiv \mathbf{U}_0 \boldsymbol{\Lambda}_0^{1/2}, \tag{23}$$

e.g. an svd of a Gaussian parametric form given in Eq. 39. Alternatively, \mathbf{Z}_0 can be specified via the Gramm–Schmidt orthogonalization of an ensemble of the control parameters. The number, p , of orthogonal error subspaces retained in \mathbf{P}_0 is at the discretion of the modeler and can be adjusted as needed. It will be shown below that the orthogonal matrix \mathbf{Z}_0 provides a linear basis for an analysis increment of the control parameters. Therefore, rank p should be chosen such as to ensure that \mathbf{Z}_0 provides a sufficient linear basis depending on a specific application. With the OBCs utilized as the control space, the a priori values $\hat{\boldsymbol{\zeta}}_0$ can be specified from a global tidal model and $p \leq 100$ is typically sufficient. Following Eq. 16, the model error subspace induced by \mathbf{Z}_0 is given by

$$\mathbf{Z}_{\mathbf{x}_0} = \mathbf{M}_{(\mathbf{x}_0 \leftarrow \boldsymbol{\zeta}_0)} \mathbf{Z}_0. \tag{24}$$

As before, the operator $\mathbf{M}_{(\mathbf{x}_0 \leftarrow \boldsymbol{\zeta}_0)}$ represents an iterative solution of Eq. 13, with the right-hand side given by $\mathbf{B}_{(\mathbf{x}_0 \leftarrow \boldsymbol{\zeta}_0)} \mathbf{Z}_0$. Equation 24 can be solved efficiently even for large ($10^3 - 10^4$) values of p by, firstly, carrying out an incomplete LU factorization of $\mathbf{A}_{(\mathbf{x} \leftarrow \mathbf{x})}$ with small values of drop tolerance and, secondly, using the computed LU factors as preconditioners of the conjugate gradient solver. $\mathbf{Z}_{\mathbf{x}_0} \in \mathbb{C}^{n_x \times p}$ is the matrix square root of the forward model error covariance

$$\mathbf{P}_{\mathbf{x}_0} = \mathbf{Z}_{\mathbf{x}_0} \mathbf{Z}_{\mathbf{x}_0}^H, \tag{25}$$

induced by the control parameter error propagated from the control space to the model state-space in the control domain. The formalism can be further applied to obtain the matrix square roots of forward model error covariances in the nested domains, in addition to the control domain. The nested domains are coupled to the control domain through the OBCs, $\boldsymbol{\zeta}_i$. With $\mathbf{H}_{\boldsymbol{\zeta}_i \leftarrow \mathbf{x}_0}$ denoting an interpolation from the state-space in the

control domain onto the open boundaries of the i th nested domain

$$\boldsymbol{\zeta}_i = \mathbf{H}_{\boldsymbol{\zeta}_i \leftarrow \mathbf{x}_0} \mathbf{x}_0, \tag{26}$$

the square root of the OBC error covariance in the i th nested domain is given by

$$\mathbf{Z}_i = \mathbf{H}_{\boldsymbol{\zeta}_i \leftarrow \mathbf{x}_0} \mathbf{Z}_{\mathbf{x}_0}. \tag{27}$$

The subspaces \mathbf{Z}_i , can, next, be propagated through the corresponding nested tidal models, similarly to Eq. 24,

$$\mathbf{Z}_{\mathbf{x}_i} = \mathbf{M}_{(\mathbf{x}_i \leftarrow \boldsymbol{\zeta}_i)} \mathbf{Z}_i. \tag{28}$$

Equation 28 represents an uncertainty propagation from the OBC space of the nested domains to the state-space of the nested domains. As in Eq. 16, operator $\mathbf{M}_{(\mathbf{x}_i \leftarrow \boldsymbol{\zeta}_i)}$ is a formal representation of the numerical procedures solving the dynamical equations in the nested domains, with the open boundary forcing provided by \mathbf{Z}_i . Solution of the dynamical equations (Eq. 28) yields the square roots of the forward model error covariances in the nested domains. In conjunction with Eqs. 24 and 27, the procedure amounts to propagating an uncertainty from the control space $\boldsymbol{\zeta}_0$ to the state-spaces of the nested tidal models.

3.4 Estimation of control parameters

An optimization problem is posed using a quadratic cost functional penalizing the (weighted) variance of the observation-minus-forecast residuals in the control and the nested domains

$$\hat{\boldsymbol{\zeta}}_* = \arg \min_{\boldsymbol{\zeta}} J(\boldsymbol{\zeta}), \tag{29}$$

$$J(\boldsymbol{\zeta}) = (\boldsymbol{\zeta} - \hat{\boldsymbol{\zeta}}_0)^H \mathbf{P}_0^{-1} (\boldsymbol{\zeta} - \hat{\boldsymbol{\zeta}}_0) + \sum_{i=0}^m (\mathbf{y}_i - \mathbf{H}_{\mathbf{y}_i \leftarrow \mathbf{x}_i} \mathbf{x}_i)^H \mathbf{R}_i^{-1} (\mathbf{y}_i - \mathbf{H}_{\mathbf{y}_i \leftarrow \mathbf{x}_i} \mathbf{x}_i).$$

In Bayesian interpretation, the first term in Eq. 29 represents the prior information about the control parameters. It can also be viewed as a regularization term added to penalize a deviation of the control parameter values from their a priori values. The definition (Eq. 29) of a penalty functional is equivalent to the minimum error variance estimation in the control parameter space (Logutov 2007). The minimization problem (Eq. 29) is subject to the linear constraints

$$\mathbf{x}_i = \mathbf{M}_{(\mathbf{x}_i \leftarrow \boldsymbol{\zeta}_i)} \boldsymbol{\zeta}_i, \quad \forall i \in \{0, 1, \dots, m\}, \tag{30}$$

with $\mathbf{M}_{(x_i \leftarrow \zeta_i)}$ described in Eq. 17 and the OBCs ζ_i obtained via Eq. 26. Equations 29–30 fall into a general class of quadratic minimization problems with linear constraints.

To provide a practical method of solution of Eqs. 29–30, we, first, establish the notation that unifies the linear constraints (Eq. 30). By combining Eqs. 30 and 26, a nested state-space, \mathbf{x}_i , is related to the control parameters via

$$\mathbf{x}_i = \mathbf{M}_{(x_i \leftarrow \zeta_0)} \zeta_0 \tag{31}$$

with the operator $\mathbf{M}_{(x_i \leftarrow \zeta_0)}$ given by

$$\mathbf{M}_{(x_i \leftarrow \zeta_0)} \equiv \mathbf{M}_{(x_i \leftarrow \zeta_i)} \mathbf{H}_{\zeta_i \leftarrow x_0} \mathbf{M}_{(x_0 \leftarrow \zeta_0)}. \tag{32}$$

Consistently with the notation, the operator $\mathbf{M}_{(x_i \leftarrow \zeta_0)}$ is a formal representation of the following numerical procedures: solution of the dynamical equations in the outer domain, interpolation of the solution to open boundaries of the i th nested domain, followed by a solution of the dynamical equations in the i th nested domain. Finally, let

$$\mathbf{M}_{(y \leftarrow \zeta_0)} \equiv \begin{bmatrix} \mathbf{H}_{y_0 \leftarrow x_0} \mathbf{M}_{(x_0 \leftarrow \zeta_0)} \\ \mathbf{H}_{y_1 \leftarrow x_1} \mathbf{M}_{(x_1 \leftarrow \zeta_0)} \\ \dots \\ \mathbf{H}_{y_m \leftarrow x_m} \mathbf{M}_{(x_m \leftarrow \zeta_0)} \end{bmatrix} \tag{33}$$

denote the total dynamical operator providing the multigrid model values of the tidal fields at observation locations given the values of the control parameters. The observational operators $\mathbf{H}_{y_i \leftarrow x_i}$ are described in the text leading to Eq. 21. The linear dynamical constraints are now unified through $\mathbf{M}_{(y \leftarrow \zeta_0)}$. With these dynamical constraints substituted directly into Eq. 29, the penalty functional is expressed

$$J(\zeta_0) = (\zeta_0 - \hat{\zeta}_0)^H \mathbf{P}_0^{-1} (\zeta_0 - \hat{\zeta}_0) + (\mathbf{y} - \mathbf{M}_{(y \leftarrow \zeta_0)} \zeta_0)^H \mathbf{R}^{-1} (\mathbf{y} - \mathbf{M}_{(y \leftarrow \zeta_0)} \zeta_0), \tag{34}$$

with the block-diagonal \mathbf{R} given in Eq. 20. The quadratic functional Eq. 34 is convex and has a unique minimum. The solution can be derived in closed form by expressing the first variation of Eq. 34 with respect to the control parameters ζ and setting it to zero. The minimum is reached at

$$\hat{\zeta}_* = \hat{\zeta}_0 + \mathbf{P}_0 \mathbf{M}_{(y \leftarrow \zeta_0)}^H (\mathbf{M}_{(y \leftarrow \zeta_0)} \mathbf{P}_0 \mathbf{M}_{(y \leftarrow \zeta_0)}^H + \mathbf{R})^{-1} \times (\mathbf{y} - \mathbf{M}_{(y \leftarrow \zeta_0)} \hat{\zeta}_0). \tag{35}$$

With the a priori error covariance \mathbf{P}_0 provided by Eq. 22, the analysis equation (Eq. 35) can be equivalently expressed as

$$\hat{\zeta}_* = \hat{\zeta}_0 + \mathbf{Z}_0 \beta. \tag{36}$$

As was eluded to in Section 3.3, the orthogonal matrix \mathbf{Z}_0 provides a linear basis for an analysis increment of the control parameters. Complex coefficients $\beta \in \mathbb{C}^P$ are obtained from the observation-minus-forward model residuals

$$\beta = \tilde{\mathbf{Z}}^H (\tilde{\mathbf{Z}} \tilde{\mathbf{Z}}^H + \mathbf{R})^{-1} (\mathbf{y} - \mathbf{M}_{(y \leftarrow \zeta_0)} \hat{\zeta}_0) \tag{37}$$

$$\tilde{\mathbf{Z}} = \mathbf{M}_{(y \leftarrow \zeta_0)} \mathbf{Z}_0. \tag{38}$$

Matrix $\tilde{\mathbf{Z}}$ is the square root of the error covariance propagated from the control parameter space to observation locations through a tidal dynamical model in the multigrid settings. Analysis equations (Eqs. 36–38) provide a practical method of data assimilation using multigrid resolving computations. The steps of the inverse solution are summarized below.

3.5 Summary of the inverse solution

First, a choice of the square root matrix \mathbf{Z}_0 is made, for example, by following Eq. 23 or via the Gram–Schmidt orthogonalization of an ensemble of the control parameters. Numerical solution of the dynamical equations (Eq. 13) is, next, carried out in the outer domain, with forcing provided by $\mathbf{B}_{(x \leftarrow \zeta)} \mathbf{Z}_0$. The result, the square-root \mathbf{Z}_{x_0} , is interpolated onto the open boundaries of the nested domains to provide the OBCs for the nested computations. The dynamical equations in the nested domains are subsequently solved and the square-roots \mathbf{Z}_{x_i} are generated. Finally, the observational operators $\mathbf{H}_{y_i \leftarrow x_i}$ are applied to \mathbf{Z}_{x_i} in each domain and the square-root $\tilde{\mathbf{Z}}$ in Eq. 38 is obtained. The analysis increment in the control space is computed via Eqs. 36–38, and the multigrid tidal modeling system is run again, with the analysis values of the control parameters, to yield the inverse tidal solution in the control and in the nested domains.

4 Discussion

The key point of divergence of various data assimilation schemes is the method by which an uncertainty is estimated and propagated through a dynamical model (Lermusiaux et al. 2006a). Adjoint methods seek to optimize a model state-space trajectory with respect to measurements and the a priori boundary and initial conditions by solving the coupled forward and adjoint dynamical equations arising from the principles of calculus of variations (Bennett 1992). The

representer method refers to a specific way of solving the coupled forward/adjoint systems by their decoupling of each other, suitable for linearized dynamical systems (Egbert et al. 1994). In these methods, various sources of uncertainty are accounted for by specifying the error covariances or convolution kernels affecting the forcing terms of the forward dynamical system. In contrast, the KF based assimilation algorithms evolve an uncertainty in the model state-space by integrating the linearized error covariance evolution equations, potentially embedded in the stochastic environment (Heemink and Kloosterhuis 1990). Stochastic primitive equation ocean models propagate and model an uncertainty by adding random perturbations, with specified characteristics, to a deterministic primitive equation ocean model to represent the effect of various error sources (Lermusiaux 2006). In the method presented in this paper, data assimilation is carried out, with the mediation of nested computation(s), by propagating an uncertainty from a control space in the outer domain to model tidal fields at observation locations in the outer and in the nested domains using low-rank error covariance representations (Lermusiaux and Robinson 1999). An analysis increment in the control space is computed by minimizing the (weighted) variance of the observation-minus-forecast residuals constructed using nested resolving computations.

The methodology relies on steering the model state-space towards observations by adjusting the values of the parameters in a chosen control space. Because an assimilation scheme is implemented by means of a chosen set of the control parameters, the inverse tidal fields are in a perfect dynamical balance, uniform across the multigrid system. The specific choice of the model control space is left to the discretion of the modeler and depends on a specific application. A level of generality was intentionally retained throughout this presentation with respect to the choice of the control space. The only requirement imposed was that the model dynamics can be linearized with respect to the control parameters for the purposes of data assimilation. As an example, the OBCs provide a sensible control space in regional tidal modeling applications (Logutov and Lermusiaux 2008). In coastal waters, the tidal forcing occurs mainly through the OBCs and constitutes a significant source of uncertainty. The OBCs of regional tidal models are typically prescribed from global tidal estimates, for example, from TPXO (Egbert and Erofeeva 2002) or FES95 (LeProvost et al. 1994) global models. The resolution of global models is insufficient to fully resolve regional topographic and coastal features, and, therefore, tuning the OBCs consistently with the regional tidal dynamics and the measurements is sensible. The

choice of the OBCs as the control space is adopted for the illustration of the method in the next section. Alternatively, the bottom friction parameters or bottom topography parameters can be utilized as the control space. Linearization of the dynamical equations with respect to the OBCs is straightforward, presented in Section 2. Linearization with respect to the bottom friction parameters is also feasible, briefly described in Section 2 but not followed upon. The next section provides a real-world example of an application of the described methodology in the context of the PhilEx.

5 PhilEx case study

The PhilEx project, an ongoing effort at the time of this writing, was designed as a multi-institutional, wide-range study of spatial and temporal variability of the oceanographic processes in and around the straits of the Philippine Archipelago, with an overall goal of better understanding the archipelago dynamics and improving the capability to model and predict regional and coastal ocean processes in areas with diverse topographic conditions. A wide variety of topographic conditions, including enclosed and semienclosed seas connected through a network of channels and sills, make tidal modeling in the waters of the Philippines an exceptionally challenging task. The logistics of the modeling component of PhilEx required that the tidal computations were carried out in a domain covering the Philippines Archipelago, as well as portions of the Pacific and Indian oceans shown in Fig. 2. Computational constraints and other considerations led us to use a structured grid spectral barotropic tidal model (Logutov and Lermusiaux 2008) at a 5-min resolution for that domain. The 5-min resolution domain will be referred to hereafter as the *control* or *outer* domain, consistently with the previous notation. The bottom topography was specified from 1-min resolution (Smith and Sandwell 1997) (version 9.1) bathymetry data set, the highest resolution bathymetry product available to us. Appropriate smoothing was applied to the bathymetric data to avoid aliasing when specifying the model grid at 5-min resolution. The a priori values of the OBCs were obtained from the 1/4-degree resolution TPXO7.0 global tidal model (Egbert and Erofeeva 2002).

Tidal observations available to us consisted of the Topex-Poseidon (TP) satellite altimeter data, with tracks shown in Fig. 2, and the velocity and the SSH data from two moorings in the Panay (A1) and the Dipolog (A2) straits. The A1 and A2 moorings were

Fig. 2 Bathymetry [m] in PhilEx 5-min resolution model domain and the observational data. The *black bounding box* shows the nested 1-min resolution domain. The observational network consisted of the Topex/Poseidon altimetry and two moorings, *A1* and *A2*. The observational data are partitioned into y_0 (*yellow*) and y_1 (*red*) as explained in text

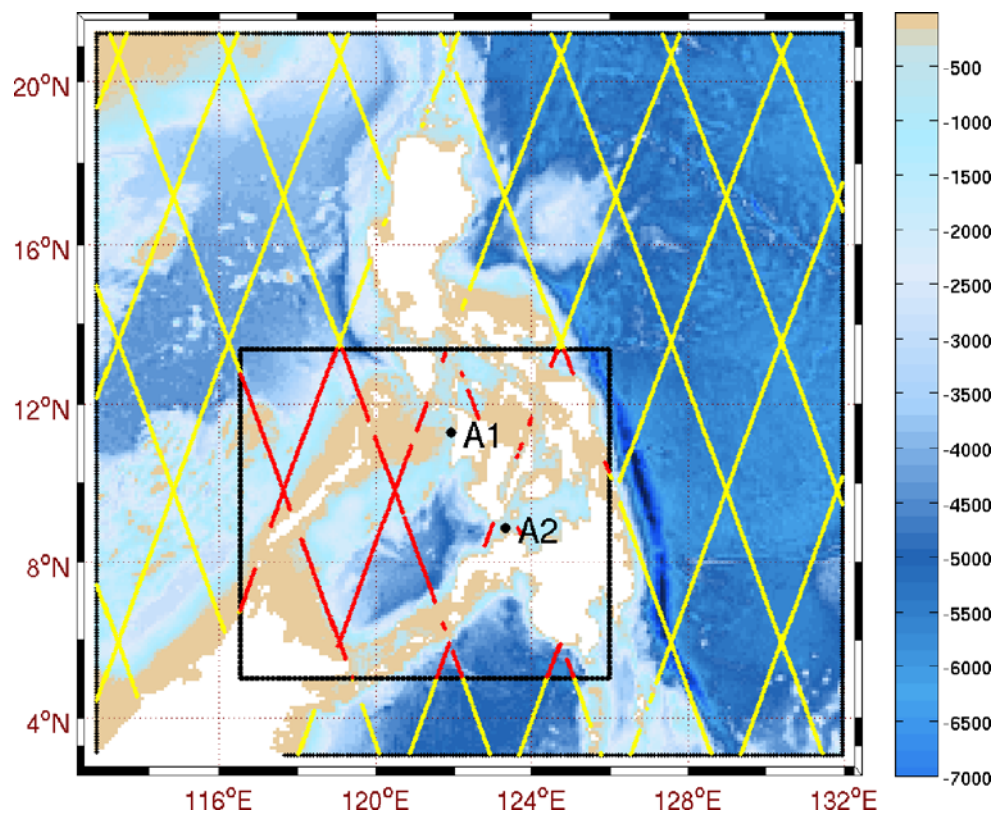
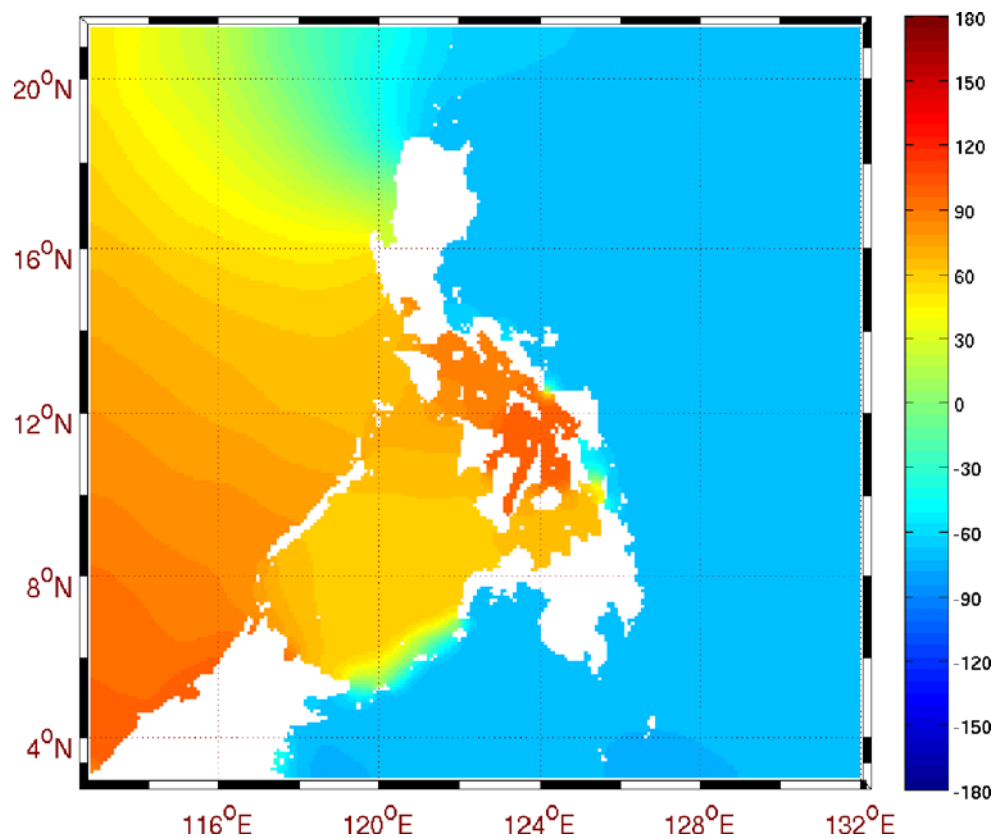


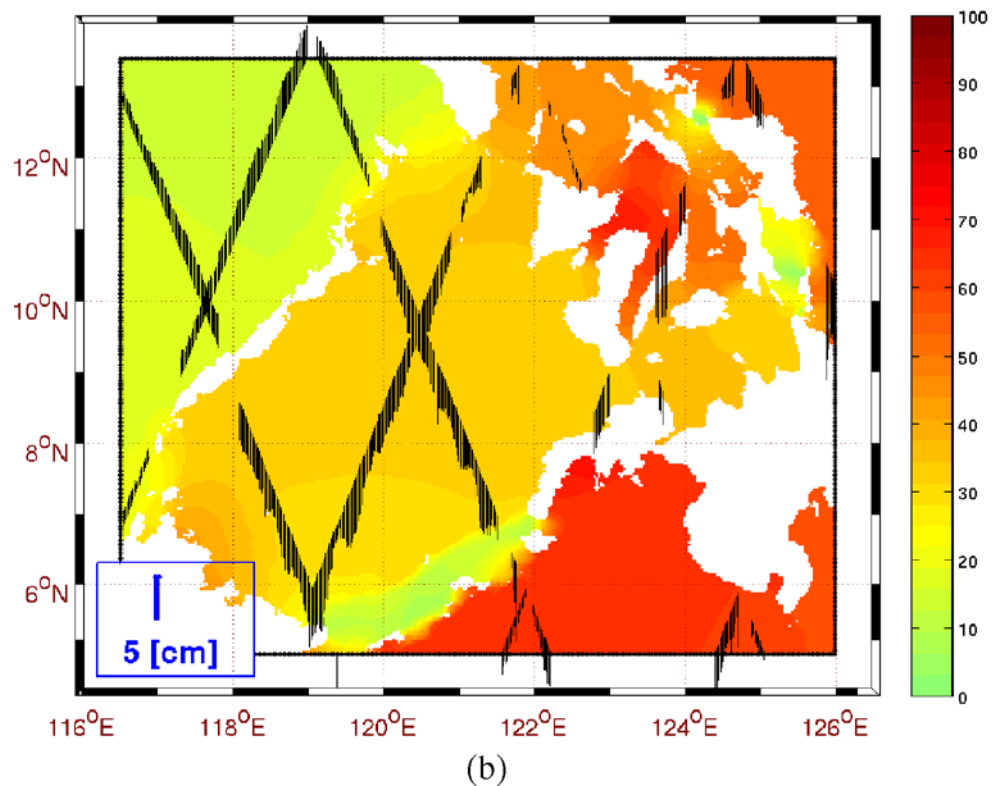
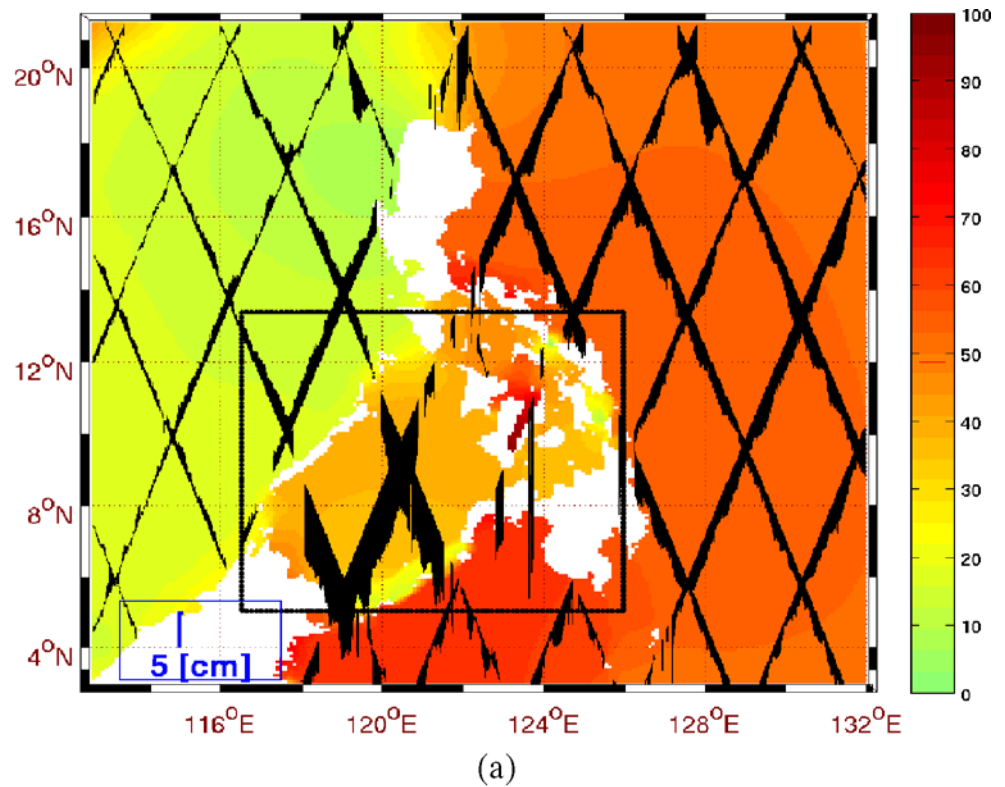
Fig. 3 M_2 Greenwich phase [in deg] based on the forward computations in the outer domain



deployed and operated by the team led by Janet Sprintall of the Scripps Institution of Oceanography as part of PhilEx observational program. In addition to the

acoustic doppler current profilers (ADCPs), the moorings were carrying pressure gauges, suitable for inferring the SSH. The TP data, collected from the launch of

Fig. 4 Data-forward model misfits for M_2 . Plotted: SSH amplitude in the control and nested domains. The misfits are plotted as *arrows* originating at TP tracks and pointing up for positive misfits and down for negative. **a** Five-min resolution computation. **b** One-min resolution computation. Arrow scale is given in *bottom left corner*



the satellite in 1992, were provided to us in the form of tidal constituents (SSH amplitudes and phases) by Dr. Richard Ray of the Planetary Geodynamics Laboratory at the NASA Goddard Space Flight Center (Schrama and Ray 1994; Ray 2001). The TP data were utilized to constrain the model to the observed sea-level elevations; the mooring ADCP and pressure measurements were utilized to validate tidal model outputs. Since the mooring data were utilized for validation, they were not used for assimilation.

Tidal modeling in the Philippine basin is difficult because of the substantial phase and amplitude transitions

in tidal constituents across the straits on the eastern boundary of the Archipelago. The transitions occur in the Surigao and the San Bernardino straits connecting the Bohol Sea and the Visayan Sea with the Pacific Ocean, and in the straits of the Sulu Archipelago connecting the Sulu and the Celebes seas. Figure 3 shows the cotidal chart for M_2 , the dominant tidal constituent in the basin, based on the forward model computations. The M_2 tide is out of phase across the Surigao and the San Bernardino straits and has about 100 degrees phase difference across the straits of the Sulu Archipelago. Under these circumstances, the computations are

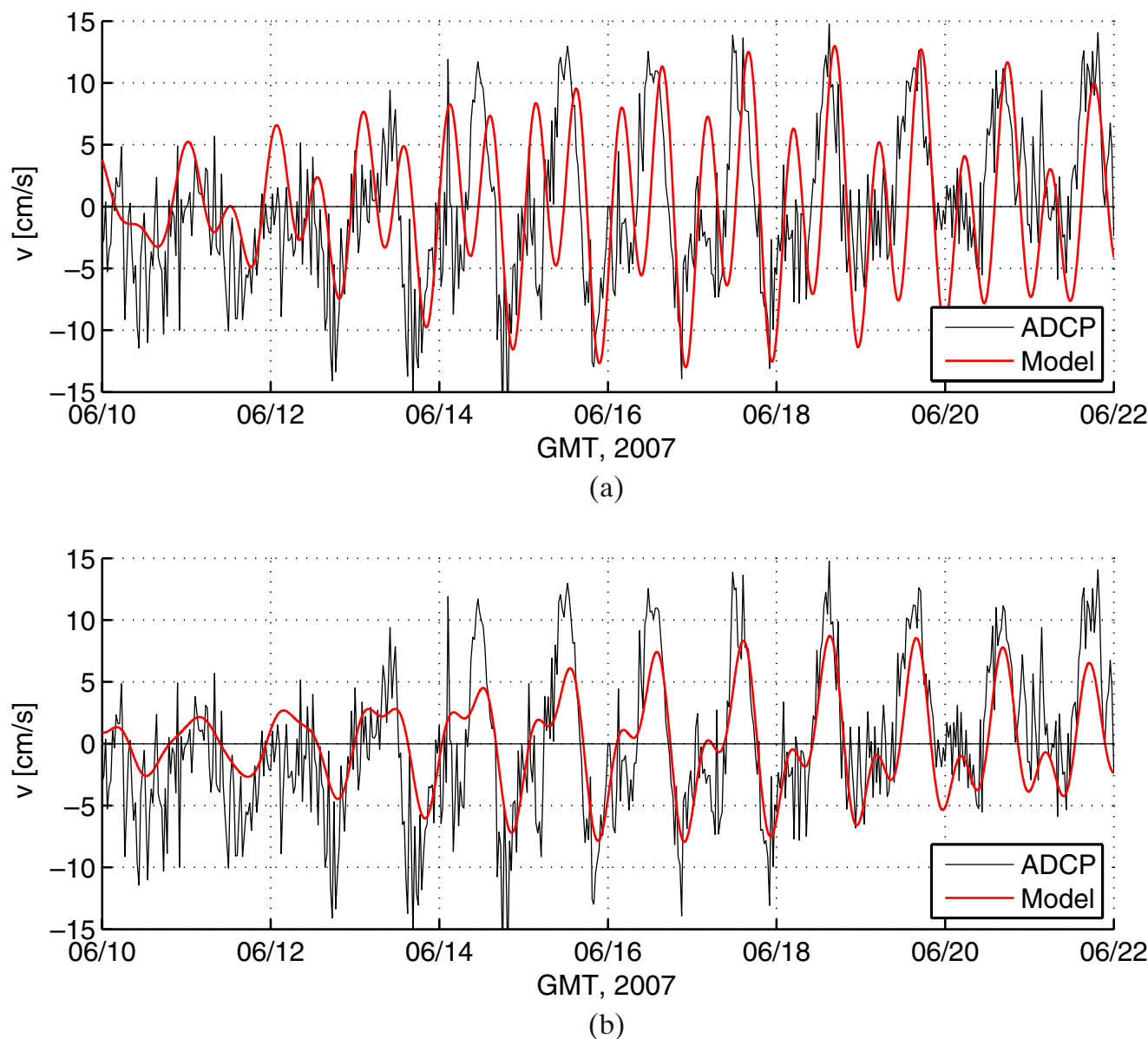
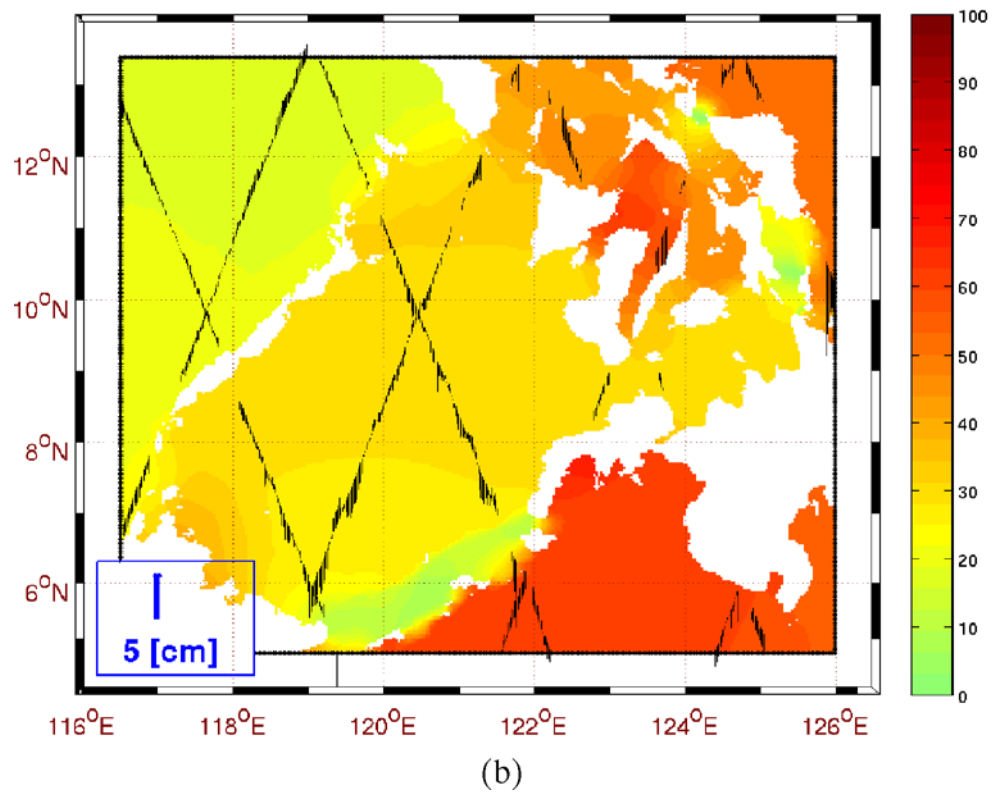
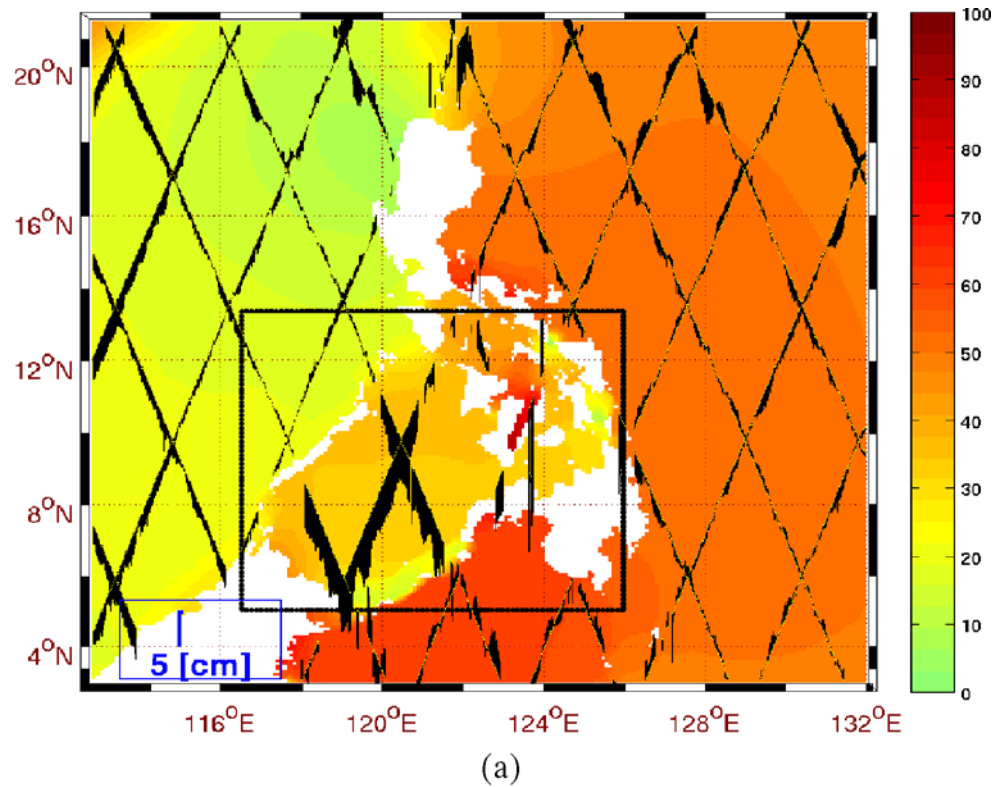


Fig. 5 Observed and forward model velocity at A1. Observed meridional depth-averaged velocity, with mean removed (*black*). Model velocity (*red*). **a** At 5-min resolution; **b** at 1-min resolution

sensitive to model resolution because the resolution affects the representation of bottom topography and, therefore, tidal transports through the straits. At 5-min

resolution, the data-model residuals in the Sulu, Bohol, Visayan, and Sibuyan seas were higher than the average across the model domain. A trial use of a higher-

Fig. 6 Same as Fig. 4 but for the multigrid inverse solution

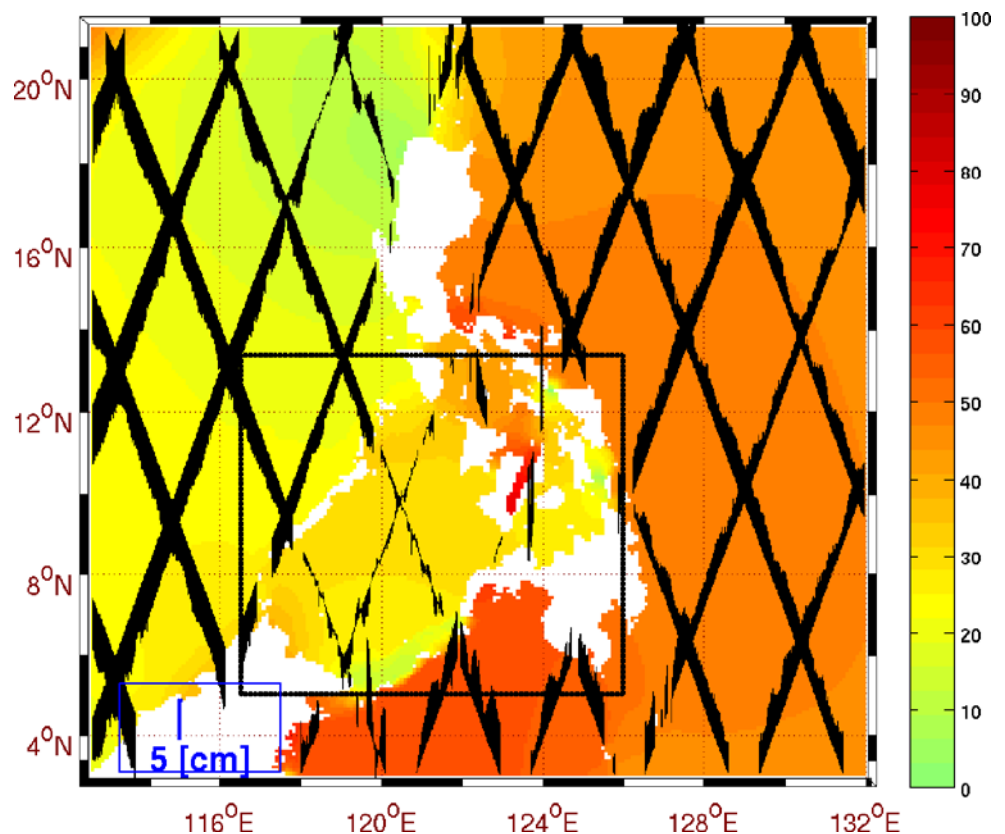


resolution nested domain covering these basins confirmed that the computations were sensitive to model resolution. Figure 4 shows the SSH amplitudes of the M_2 tide modeled at 5-min (panel a) and 1-min (panel b) resolution. Plotted also are the misfits of the observed and the model M_2 amplitude. The misfits are plotted as arrows originating at TP tracks and pointing up or down depending on the sign of a misfit, positive, up, and negative, down. Thus, an upward arrow indicates that an observed value is higher than a model value, and vice versa. The misfits are within the 5-cm range at 5-min model resolution, except in the Sulu sea and the adjacent basins. The Sulu, Bohol, Visayan, and Sibuyan seas had model values consistently higher than the observed SSH by an average of 10 cm, with even larger misfits at several locations. The computations at 1-min resolution, however, showed a substantial reduction of data-model misfits. The reduction indicated to us the presence of the representativeness error in these areas. A 1-min resolution domain was, therefore, setup in the Sulu, Bohol, Visayan, and Sibuyan sea basins, shown by the bounding box in Fig. 2, and a task of assimilating the measurements into the 5-min resolution model via the mediation of the nested 1-min resolution domain was presented. The methodology described in this paper

allowed us to achieve that task. Figure 5 shows the time series of the observed and the forward model velocity at mooring A1 in the Panay strait. The mooring was located about 3 km off the coast of Panay Island, near a coastal segment with predominantly meridional orientation. The meridional velocity plotted in Fig. 5 corresponds to the along-slope direction. Local bottom topography is a major factor determining a velocity field near coastal moorings. The observation-minus-forecast residuals at A1 were found to be highly sensitive to model resolution. Figure 5 shows the velocities at A1 obtained from the 5-min resolution and the 1-min resolution computations compared against the depth-averaged ADCP velocity measurements. The forward solution in the control domain exhibited significant errors in both amplitude and phase. However, the 1-min resolution nested computations were in closer agreement with the measurements.

To illustrate the utility of our method, the information from the TP data was assimilated into the outer model domain with and without the use of the nested domain. The primary source of tidal forcing in the PhilEx region is provided through the open boundaries. Therefore, the OBCs were chosen as the control space for the purposes of data assimilation. The a priori

Fig. 7 Same as Fig. 6a but for the stand-alone inverse computation, with all TP data assimilated directly into the 5-min resolution model



error covariance of the control parameters was specified using a Gaussian two-dimensional parametric form

$$\mathbf{G}_-(\mathbf{r}, \mathbf{r}_0) = \sigma_0^2 \exp \frac{-|\mathbf{r} - \mathbf{r}_0|^2}{2L^2}, \quad |(\mathbf{r}, \mathbf{r}_0) \in \partial\Omega_O, \quad (39)$$

where \mathbf{r} and \mathbf{r}_0 are coordinates of two points on the control domain open boundary. The length scale and the variance parameters in Eq. 39 were chosen $L = 100$ km and $\sigma_0^2 = (10 \text{ cm})^2$. The svd of Eq. 39, with $p = 50$ dominant singular vectors retained, was carried out and Eq. 23 was utilized to obtain \mathbf{Z}_0 in Eq. 37. The ob-

servations were assumed to have spatially uncorrelated error, with variance $\sigma_o^2 = (10 \text{ cm})^2$. Figures 6 and 7 are provided for comparison of the inverse solution for the M_2 constituent in the multigrid system and the stand-alone system, respectively. The former shows the SSH amplitude of the multigrid inverse solution, computed using the methodology presented in this paper. The latter presents the amplitude in the stand-alone inverse computation, with all the TP data assimilated directly into the 5-min resolution model. These computations show that fitting the TP data in the Sulu, Bohol, Visayan, and Sibuyan seas directly into the 5-min

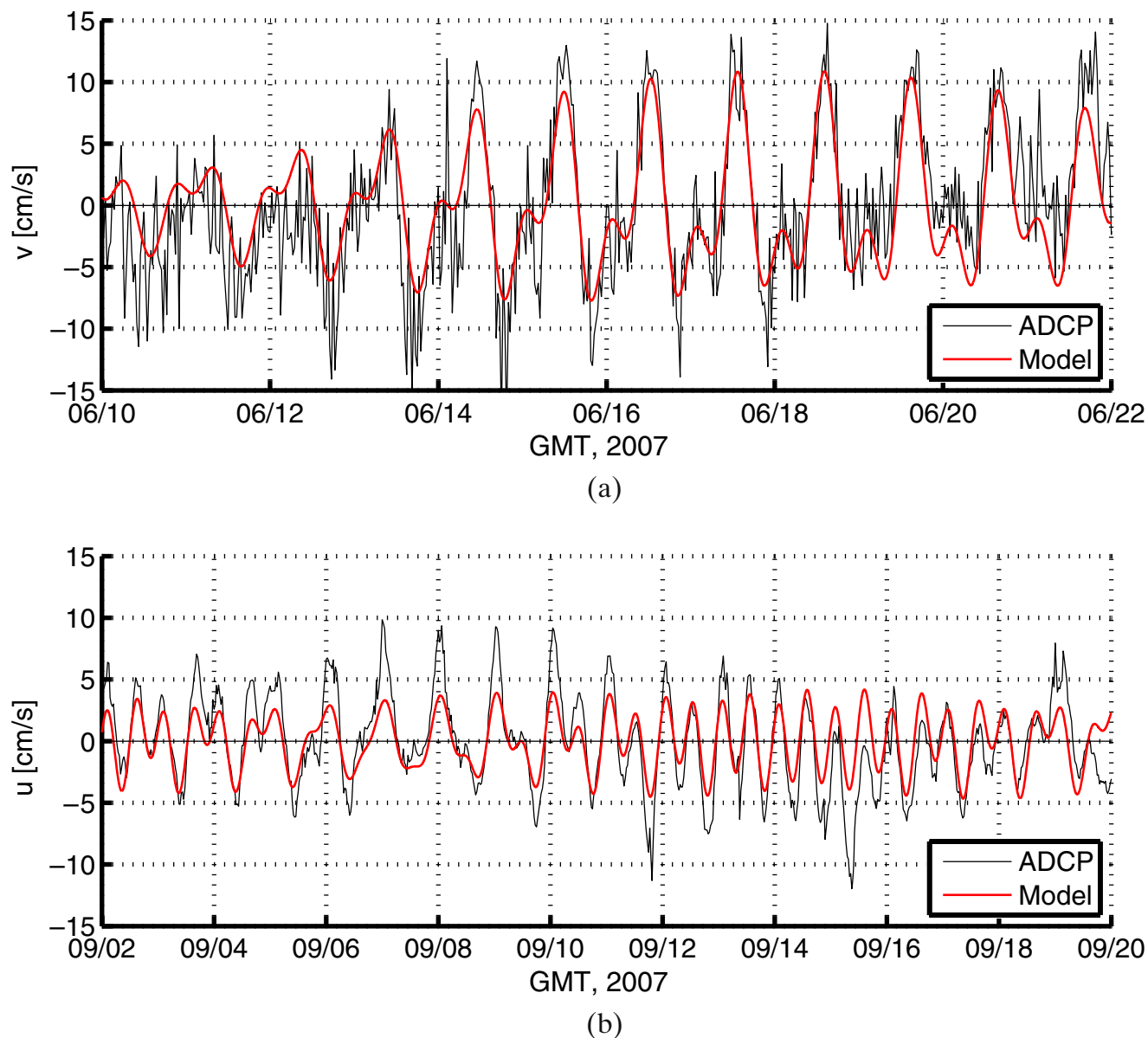
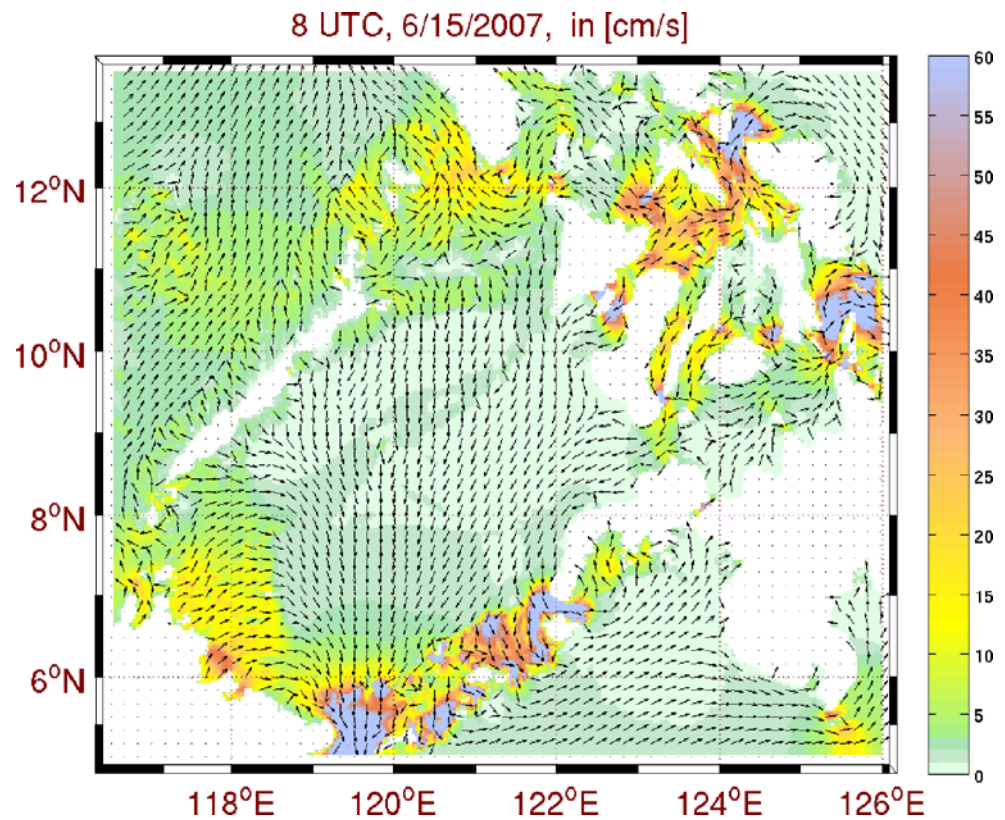


Fig. 8 Observed vs inverse model velocities at A1 and A2. Plotted are the depth-averaged velocities, with mean removed. **a** Meridional velocity at A1; **b** zonal velocity at A2

Fig. 9 Example of a tidal velocity forecast obtained using the inverse scheme. Total tidal velocity valid at 8 UTC, June 15, 2007



resolution model was counter-productive and led to a degradation in the overall model fit to the altimetry. On the contrary, assimilation of the named data subset via the use of the resolving 1-min computation led to a consistent decrease of the overall error. The inverse solution depicted in Fig. 6 provides a fit of the barotropic tidal dynamics to data consistently with the dynamical scales resolved in models and in observations. Figure 8 shows the comparison of the velocity field of the multigrid inverse solution against the ADCP data. An improvement in the velocity field estimates through the use of the multigrid inverse can be analyzed by comparing Figs. 8a and 5b. The match of the ADCP measurements to the inverse model velocity has been substantially improved as compared to the forward model velocity in Fig. 5. On the contrary, the velocity field of the stand-alone inverse solution was in worse agreement (figure not shown) with the ADCP data than the forward solution shown in Fig. 5b. Figure 9 gives an example of the final output of the inverse scheme: the tidal velocity forecast at high resolution, tuned to measurements. The use of the multigrid data-assimilative computations has allowed us to generate reliable barotropic tidal velocity forecasts and demonstrate a good skill in modeling the barotropic tidal circulation despite the challenges presented by the complexity of coastlines and bottom topography.

6 Summary and conclusions

The main focus of this paper was a multigrid data-assimilative framework for assimilation of measurements into regional tidal models. A new technique for inverting the observational data in a multigrid setting has been described. The methodology allows to employ nested domains around coastal segments where tidal observations are collected but model resolution is insufficient to fully resolve the tidal dynamics. As was argued in the paper and demonstrated using a real-world application in the context of the PhilEx dynamical experiment, assimilation of measurements inconsistent with the resolved model dynamics can be counter-productive. In regional tidal modeling applications, coastal measurements are often affected by local small-scale coastal and topographic features. For example, the SSH measured in a small bay or in an estuary can have tidal variations substantially different, both in amplitude and in phase, from the nearby open ocean areas, depending on the local characteristics of a coastline and bottom topography. In areas where model resolution is insufficient to resolve the important characteristics of the waterways, the observational data are compounded with the representativeness error. The inverse procedures presented in this paper are designed to suppress the representativeness error by employing

nested resolving computations. This capability is expected to find applications in the context of structured grid regional tidal modeling applications.

The method seeks to assimilate the observational data into a multigrid modeling system by using a set of control parameters chosen at the discretion of the modeler, such as the OBCs. Presently, the control parameters are defined in the outer domain. With this strategy, a model is fitted to data consistently with the dynamics defined uniformly across the entire multigrid computational system. The inverse fields are in a perfect dynamical balance and a transition of the inverse solution across the domain boundaries is seamless and satisfies the unified dynamical balance. An assimilation is carried out with the mediation of nested computation(s) resolving all important topographic and coastal features. The data inversion is carried out by propagating an uncertainty associated with the a priori values of the control parameters to the multigrid model state-space at observation locations. The inverse does not require an adjoint model and is highly practical. In the future, one could consider the idea of having a scheme that corrects for the control parameters defined in each or some of the nested domains, in addition to the outer domain. Such an approach would be desirable in applications where a solution in the nested domains cannot be fully controlled through the OBCs only. A disadvantage of such an approach is that it would result in a discontinuity of the dynamical balance across the domain boundaries. The inverse method discussed in this paper can also become a useful basis for future investigations of data assimilation into models of physical processes occurring on multiple scales. One of our future research directions would include an extension of the method to other models and processes in environmental prediction, monitoring, and planning (Lermusiaux et al. 2007). In addition, other sources of uncertainty could be allowed, for example, to account for the inaccuracies related to nonlinear and frictional effects in the momentum equations and parameterization of the barotropic-to-baroclinic conversions. Work in this direction is currently in progress.

Acknowledgements This study was supported by the Office of Naval Research under grants N00014-07-1-0473 (PHILEX), N00014-07-1-0501 (AWACS), N00014-07-1-0241 (core ONR), and N00014-07-1-0241 (QPE) to the Massachusetts Institute of Technology (MIT), with P.F.J. Lermusiaux as principal investigator. We are very grateful to Dr. Janet Sprintall and her team of the Scripps Institution of Oceanography, who provided us with the mooring data from the Panay and the Dipolog straits, instrumental for this study. We also thank the whole PHILEX group of scientists, as well as the team of mooring engineers and technicians involved in the observational program, including the planning, deployment, and recovery of the instruments.

This study would not be possible without their efforts and the collected data. The work of Dr. Janet Sprintall and her team was supported by ONR through grant N00014-06-1-0690. Dr. Richard Ray of the Planetary Geodynamics Laboratory at the NASA Goddard Space Flight Center has provided us with the harmonically analyzed TP data. His contribution, critical for this study, is hereby gratefully acknowledged. I would like also to thank Prof. Pierre Lermusiaux (MIT) for his general guidance through the project and many useful discussions and recommendations with regard to the material of this paper. Wayne Leslie and Pat Haley have provided some technical and computer support for this work at MIT. Lastly, the anonymous reviewers of this paper have provided useful criticisms and inputs, which were included in the text and improved the quality and clarity of the presented material.

Appendix

Index notation

$\mathbf{A} \in \mathbb{C}^{m \times n}$ complex $m \times n$ matrix,

$\mathbf{i}_K \in \mathbb{N}^K$ a set $\mathbf{i}_K = \{i_k\}_{k=1}^K$, $i_k \in \{1, 2, \dots\}$,

$(\mathbf{a})_{\mathbf{i}_K} \in \mathbb{C}^K$ complex vector of length K containing the \mathbf{i}_K th entries of \mathbf{a} , $(\mathbf{a})_{\mathbf{i}_K} = [a_{i_1}, a_{i_2}, \dots, a_{i_K}]^T$,

$(\mathbf{A})_{\mathbf{i}_K, \mathbf{j}_L} \in \mathbb{C}^{K \times L}$ complex $K \times L$ matrix containing the entries in the \mathbf{i}_K th rows and \mathbf{j}_L th columns of matrix \mathbf{A} ,

$$(\mathbf{A})_{\mathbf{i}_K, \mathbf{j}_L} = \begin{bmatrix} a_{i_1, j_1} & a_{i_1, j_2} & \dots & a_{i_1, j_L} \\ a_{i_2, j_1} & a_{i_2, j_2} & \dots & a_{i_2, j_L} \\ \vdots & \ddots & \dots & \vdots \\ a_{i_K, j_1} & a_{i_K, j_2} & \dots & a_{i_K, j_L} \end{bmatrix}_{K \times L}.$$

References

- Andersen O, Egbert G, Erofeeva S, Ray S (2006) Mapping nonlinear shallow-water tides: a look at the past and future. *Ocean Dyn* 56(14):416–429
- Barth A, Alvera-Azcrate A, Beckers J-M, Rixen M, Vandenbulcke L (2007) Multigrid state vector for data assimilation in a two-way nested model of the Ligurian sea. *J Mar Syst* 65:41–59
- Bennett A, McIntosh P (1982) Open ocean modeling as an inverse problem: tidal theory. *J Phys Oceanogr* 12:1004–1018
- Bennett AF (1992) Inverse methods in physical oceanography. In: Cambridge monographs on mechanics and applied mathematics. Cambridge University Press, Cambridge
- Bennett AF (2002) Inverse modeling of the ocean and atmosphere. Cambridge University Press, Cambridge
- Bernard PE, Remacle JF, Legat V (2008) Boundary discretization for high order discontinuous Galerkin computations of tidal flows around shallow water islands. *Int J Numer Methods Fluids*. doi:10.1002/flid.1831

- Blain CA, Preller RH, Rivera AP (2002) Tidal prediction using the advanced circulation model (ADCIRC) and a relocatable PC-based system. *Oceanography* 15(1):77–87
- Blayo E, Debreu L (2005) Revisiting open boundary conditions from the point of view of characteristic variables. *Ocean Model* 9(3):231–252
- Canizares R, Madsen H, Jensen HR, Vested HJ (2001) Developments in operational shelf sea modelling in Danish waters. *Estuar Coast Shelf Sci* 53:595–605
- Das SK, Lardner RW (1992) Variational parameter estimation for a two-dimensional numerical tidal model. *Int J Numer Methods Fluids* 15:313–327
- Davies AM (1993) A bottom boundary layer-resolving three-dimensional tidal model: a sensitivity study of eddy viscosity formulation. *J Phys Oceanogr* 23:1437–1453
- Davies AM, Jones JE, Xing J (1997) Review of recent developments in tidal hydrodynamic modelling I. Spectral models. *J Hydrol Eng ASCE* 123:278–292
- Dee DP (1995) On-line estimation of error covariance parameters for atmospheric data assimilation. *Mon Weather Rev* 123:1128–1145
- Egbert G, Bennett A, Foreman M (1994) TOPEX/Poseidon tides estimated using a global inverse model. *J Geophys Res* 99(C12):24821–24852
- Egbert GD (1997) Tidal data inversion: interpolation and inference. *Prog Oceanogr* 40:53–80
- Egbert GD, Bennett AF (1996) Data assimilation methods for ocean tides. In: Malanotte-Rizzoli P (ed) *Modern approaches to data assimilation in ocean modeling*, vol 10. Elsevier, Amsterdam, pp 147–179
- Egbert GD, Erofeeva SY (2002) Efficient inverse modeling of barotropic ocean tides. *J Atmos Ocean Technol* 19:183–204
- George K (2007) A depth-averaged tidal numerical model using non-orthogonal curvilinear co-ordinates. *Ocean Dyn* 57:363–374
- Greenberg DA, Shore JA, Page FH, Dowd M (2005) A finite element circulation model for embayments with drying in intertidal areas and its application to the quoddy region of the bay of fundy. *Ocean Model* 10:211–231
- Grenier RR, Luettich RA, Westerink JJ (1995) A comparison of non-linear frictional characteristics of two-dimensional and three-dimensional models of a shallow tidal embayment. *J Geophys Res* 100:13719–13735
- Haley PJ Jr, Lermusiaux PFJ, Robinson AR, Leslie WG, Logutov O, Cossarini G, Liang XS, Moreno P, Ramp SR, Doyle J, Bellingham J, Chavez F, Johnston S (2008) Forecasting and reanalysis in the Monterey Bay/California Current region for autonomous ocean sampling network-II experiment. *Deep Sea Res Part II* (in press)
- Heemink AW, Kloosterhuis H (1990) Data assimilation for non-linear tidal models. *Int J Numer Methods Fluids* 11(12):1097–1112
- Hubbert GD, Preller RH, Posey GDPG, Carroll SN (2001) Software requirements specification for the globally relocatable navy tide/atmosphere modeling system (PCTides). Memorandum Report NRL/MR/7322-01-8266, Naval Research Laboratory, Stennis Space
- Jones JE, Davies AM (1996) A high resolution three dimensional model of the M2, M4, M6, S2, N2, K1 and O1 tides in the eastern Irish Sea. *Estuar Coast Shelf Sci* 42:311–346
- Jones JE, Davies AM (2005) An intercomparison between finite difference and finite element (TELEMAC) approaches to modelling west coast of Britain tides. *Ocean Dyn* 55:178–198
- Jones JE, Davies AM (2007a) On the sensitivity of computed higher tidal harmonics to mesh size in a finite element model. *Cont Shelf Res* 27:1908–1927
- Jones JE, Davies AM (2007b) On the sensitivity of tidal residuals off the west coast of Britain to mesh resolution. *Cont Shelf Res* 27:64–81
- Kurapov AL, Egbert GD, Allen JS, Miller RN, Erofeeva SY, Kosro PM (2003) M2 internal tide off Oregon: inferences from data assimilation. *J Phys Oceanogr* 33:1733–1757
- LeProvost C, Genco ML, Lyard F, Vincent P, Canceil P (1994) Spectroscopy of the world ocean tides from a finite element hydrodynamic model. *J Geophys Res* 99(C12):24777–24798
- LeProvost C, Rougier G, Poncet A (1981) Numerical modeling of the harmonic constituents of the tides, with application to the English channel. *J Phys Oceanogr* 11:1123–1138
- LeProvost C, Vincent P (1986) Some tests of precision for a finite element model of ocean tides. *J Comput Phys* 65:273–291
- Lermusiaux PFJ (2006) Uncertainty estimation and prediction for interdisciplinary ocean dynamics. *J Comput Phys* 217:176–199
- Lermusiaux PFJ (2007) Adaptive modeling, adaptive data assimilation and adaptive sampling. *Physica D* 230:172–196
- Lermusiaux PFJ, Chiu CS, Gawarkiewicz GG, Abbot P, Robinson AR, Miller RN, Haley PJ, Leslie WG, Majumdar SJ, Pang A, Lekien F (2006a) Quantifying uncertainties in ocean predictions. *Oceanography* 19(1):92–105
- Lermusiaux PFJ, Malanotte-Rizzoli P, Stammer D, Carton J, Cummings J, Moore AM (2006b) Progress and prospects of U.S. data assimilation in ocean research. *Oceanography* 19(1):172–183
- Lermusiaux PFJ, Haley PJ, Yilmaz NK (2007) Environmental prediction, path planning and adaptive sampling: sensing and modeling for efficient ocean monitoring, management and pollution control. *Sea Technol* 48(9):3538
- Lermusiaux PFJ, Robinson AR (1999) Data assimilation via error subspace statistical estimation. Part I: Theory and schemes. *Mon Weather Rev* 127(7):1385–1407
- Logutov OG (2007) Multi-model fusion and uncertainty estimation for ocean prediction. PhD thesis, Harvard University
- Logutov OG, Lermusiaux PFJ (2008) Inverse barotropic tidal estimation for regional ocean applications. *Ocean Model* 25:17–34
- Luettich, Jr. RA, Westerink JJ, Scheffner NW (1992) ADCIRC: an advanced three-dimensional circulation model for shelves coasts and estuaries, report 1: theory and methodology of ADCIRC-2DDI and ADCIRC-3DL. In: *Dredging research program technical report DRP-92-6*. U.S. Army Engineers Waterways Experiment Station, Vicksburg, p 137
- Lynch DR, Hannah CG (1998) Hindcasting the Georges Bank circulation. Part I: Detiding. *Cont Shelf Res* 18:607–639
- McIntosh PC, Bennett AF (1984) Open ocean modeling as an inverse problem: M2 tides in Bass strait. *J Phys Oceanogr* 14:601–614
- Muccino JC, Arango HG, Bennett AF, Chua BS, Cornuelle BD, DiLorenzo E, Egbert GD, Haidvogel D, Levin JC, Luo H, Miller AJ, Moore AM, Zaron ED (2008) The inverse ocean modeling system. II. Applications. *J Atm Ocean Tech* 25:1608–1622
- Ray RD (2001) Inversion of oceanic tidal currents from measured elevations. *J Mar Syst* 28:1–18
- Robinson AR, Lermusiaux PFJ (2001) Data assimilation in models. In: *Encyclopedia of ocean sciences*. Academic, London, pp 623–634
- Schrama EJO, Ray RD (1994) A preliminary tidal analysis of Topex/Poseidon altimetry. *J Geophys Res* 99:24799–24808
- Smith WHF, Sandwell DT (1997) Global seafloor topography from satellite altimetry and ship depth soundings. *Science* 277:1957–1962

- Snyder R, Sidjabat M, Filloux J (1979) A study of tides, setup and bottom friction in a shallow semi-enclosed basin. part II: Tidal model and comparison with data. *J Phys Oceanogr* 9:170–188
- Sorensen JV, Madsen H (2004) Efficient Kalman filter techniques for the assimilation of tide gauge data in three-dimensional modeling of the North Sea and Baltic Sea system. *J Geophys Res* 109(C03017):1–14
- Walters RA (2005) Coastal ocean models: two useful finite element methods. *Cont Shelf Res* 25:775–793
- Xu J, Lermusiaux PFJ, Haley PJ, Leslie WG, Logotov OG (2008) Spatial and temporal variations in acoustic propagation during PLUSNet-07 exercise in Dabob Bay. In: Proceedings of meetings on acoustics (POMA), vol 4. 155th Meeting Acoustical Society of America, Paris, 29 June–4 July 2008
- Xu Z, Hendry RM, Loder JW (2001) Application of a direct inverse data assimilation method to the M2 tide on the Newfoundland and Southern Labrador shelves. *J Atmos Ocean Technol* 18:665–690

Multiscale mesh generation on the sphere

Jonathan Lambrechts · Richard Comblen ·
Vincent Legat · Christophe Geuzaine ·
Jean-François Remacle

Received: 15 May 2008 / Accepted: 8 September 2008 / Published online: 8 October 2008
© Springer-Verlag 2008

Abstract A method for generating computational meshes for applications in ocean modeling is presented. The method uses a standard engineering approach for describing the geometry of the domain that requires meshing. The underlying sphere is parametrized using stereographic coordinates. Then, coastlines are described with cubic splines drawn in the stereographic parametric space. The mesh generation algorithm builds the mesh in the parametric plane using available techniques. The method enables to import coastlines from different data sets and, consequently, to build meshes of domains with highly variable length scales. The results include meshes together with numerical simulations of various kinds.

Keywords Mesh generation · Sphere · Ocean modeling

1 Introduction

Finite elements have been used in engineering analysis for several decades. Since the 1990s, geometric domains that are used in finite element analysis and design have been built using computer-aided design (CAD) programs. Today's CAD systems are highly reliable: they deal with most of the complex geometric features of industrial parts or assemblies.

Traditional ocean models are based on finite difference schemes on Cartesian grids (Griffies et al. 2000). It is only recently that finite elements and unstructured meshes have been used in ocean modeling (e.g., Piggott et al. 2007; White et al. 2008; Danilov et al. 2005). One of the advantages of unstructured grids is their ability to conform to coastlines.

As unstructured grid ocean models began to appear, mesh generation algorithms were either specifically developed or simply adapted from classical engineering tools. Le Provost et al. (1994) use the mesh generation tools of Henry and Walters (1993) on several subdomains to obtain a mesh of the World Ocean, aiming at global scale tidal modeling. Further, Lyard et al. (2006) use a higher-resolution version of the same kind of meshes with the state-of-the-art FES2004 tidal model. Hagen et al. (2001) give two algorithms to generate meshes of coastal domains and use them to model tides in the Gulf of Mexico. Legrand et al. (2006) show high-resolution meshes of the Great Barrier Reef (Australia). On the global scale, Legrand et al. (2000) and Gorman et al. (2007) developed specific algorithms to obtain meshes of the World Ocean.

Responsible Editor: Pierre Lermusiaux

J. Lambrechts (✉) · R. Comblen · V. Legat · J.-F. Remacle
Institute for Mechanical, Material and Civil Engineering,
Université Catholique de Louvain,
Louvain-la-Neuve, Belgium
e-mail: jonathanlambrechts@gmail.com

C. Geuzaine
Electrical Engineering and Computer Science,
Montefiore Institute, Liège, Belgium

Our domain of interest is the Earth's surface, i.e., within a sufficiently good approximation, a sphere S centered at the origin and of radius R of about 6,370 km. The World Ocean is bounded by continents' and islands' coastlines. The first aim of the paper is to describe an automatic procedure that enables to build a boundary representation (BRep) of the geometry of the World Ocean within a prescribed accuracy. This procedure takes advantage of various sets of data: high-resolution shoreline databases (Wessel and Smith 1996), global relief data (National Geographic Data Center 2006), local cartographic data, etc.

Even if accurate data are available, it cannot be envisaged to build a BRep with the maximal available resolution everywhere. For example, the current global shoreline database has a resolution of about 50 m, which would lead to a huge number of control points (9,451,331). Our procedure enables to construct a model with an adaptive geometrical accuracy. Some regions of interest of the globe are discretized with the maximal available geometrical accuracy while other regions are approximated in a coarser way. Our technique also allows to mix various data sets as input.

Numerical analysis procedures utilize meshes, i.e., discretized versions of the domains described by CAD models. In this paper, we have decided *not* to develop a new mesh generation algorithm specifically designed for doing meshes that can be used in finite element marine modeling. Here, we have rather decided to build a CAD model that can serve as input for *any* surface mesher. In the last decade, mesh generation procedures have evolved with the objective of being able to interact directly with CAD models (e.g., Beall and Shephard 1997; Haimes 2000). More specifically, some of the authors of this paper have developed Gmsh: a 3D finite element mesh generator with built-in pre- and post-processing facilities. The specific nature of the model—the Earth surface with several thousands of islands, including hundreds of thousands of control points—have led us to greatly improve the meshing procedures implemented in Gmsh. Those specific features are also explained in the paper.

The paper is divided in three sections. The first section deals with the procedure for building CAD models of ocean geometries. The second section describes mesh generation procedures. In the last section, we

provide illustrative examples with diverse simulation results.

2 A geometric model for the World Ocean

Any 3D model can be defined using its BRep: a volume (called *region*) is bounded by a set of surfaces, and a surface is bounded by a series of curves; a curve is bounded by two end points. Therefore, three kinds of *model entity* are used: *model vertices* G_i^0 (dimension 0), *model edges* G_i^1 (dimension 1), and *model surfaces* G_i^2 (dimension 2).

Model entities are topological entities, i.e., they only deal with adjacencies in the model. A geometry has to be associated to each model entity. The geometries of curves and surfaces are their shapes. A parametrization of the shapes, typically a mapping, is usually available.

The geometry of a model edge is its underlying curve defined by the parametrization:

$$t \in \mathcal{R} \mapsto \mathbf{p}(t) \in \mathcal{R}^3.$$

Similarly, the geometry of a model surface is its underlying surface defined by the parametrization:

$$(u, v) \in \mathcal{R}^2 \mapsto \mathbf{p}(u, v) \in \mathcal{R}^3.$$

If a curve is included within a surface, it is usually drawn on the parameter plane (u, v) of the surface:

$$t \in \mathcal{R} \mapsto (u(t), v(t)) \in \mathcal{R}^2 \mapsto \mathbf{p}(u(t), v(t)) \in \mathcal{R}^3.$$

As an illustration, let us consider the surface represented in Fig. 1. Most important features of model entities are highlighted in this example:

- The surface is periodic. A seam curve has been introduced in the list of boundary edges of the surface to define its closure properly.
- The surface is trimmed: it contains four holes and one of those holes is crossed by the seam.
- One of the model edges on the closure of the model face is degenerated. Degenerated edges are used to take into account singularities of the mapping. Such degeneracy is present in many surface geometries: spheres, cones, and other surfaces of revolution.

From an engineering point of view, dealing with the geometry of the ocean is dealing with a trimmed

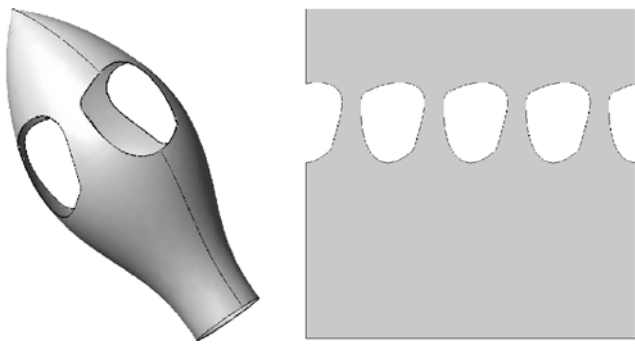


Fig. 1 A model surface in real (*left*) and parametric (*right*) coordinates. The seam of the surface is highlighted in the left plot

sphere—i.e., a surface that is periodic, that is bounded by continents and islands and that has degeneracies at both poles.

2.1 Parametrization of the sphere

Several parametrizations exist for the sphere. CAD systems use spherical coordinates. In geosciences, most of the available data are expressed in the geographic coordinate system, which has the same properties as the spherical coordinate system. Spherical coordinates suffer from all the problems that we have just mentioned before: there exist two singular points in the mapping, leading to the definition of two degenerated edges in the model; one of the coordinate directions is periodic, leading to the introduction of one seam edge; shorelines

may cross the seam edge, leading to complexity in the definition of the geometry. It is indeed impossible to choose the seam edge so that it does not cross any shoreline. In Fig. 2, a mesh of the World Ocean built using the spherical coordinate system is shown. The seam passes through the Bering Strait and crosses the Pacific Ocean, ending somewhere in the coastline of Antarctica.

Moreover, the spherical coordinates, as with most of the parametrizations, are not conformal. A conformal mapping will conserve the angle at which curves cross each other. Consequently, in order to obtain an isotropic mesh in real space, one has to build an anisotropic mesh in the parametric plane. In the case of spherical coordinates, the mapping is highly distorted near the singularities, i.e., near the poles. Robust surface meshers might be able to deal with that issue, as illustrated in Fig. 2. Anyway, it is always better to use a conformal mapping, such as the stereographic projection of Fig. 3.

Let us consider a sphere \mathcal{S} centered at the origin and of radius R , and one point s . This point lies on the surface of the sphere, does not belong to the oceans, and will be the only singular point of the mapping. A suitable choice for s could be a location in the middle of Kazakhstan, but here, we choose $s = \{0, 0, -R\}$. It corresponds to the South Pole. Antarctica being a continent, this choice makes sense for ocean modeling applications. The stereographic projection consists in projecting points p of the sphere on the plane $z = R$.

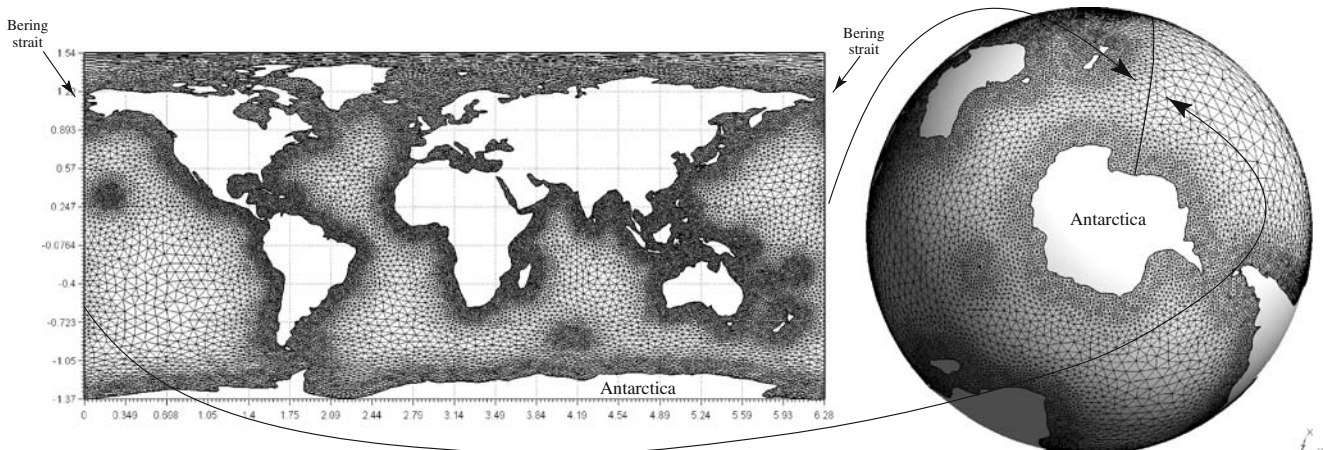


Fig. 2 Mesh of the World Ocean using the spherical coordinate system. The seam edge is visible on the right plot

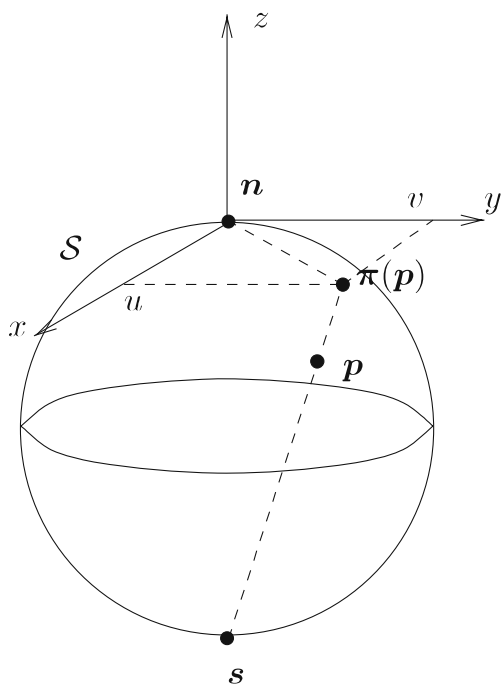


Fig. 3 Stereographic projection

The stereographic projection $\mathbf{u}(\mathbf{x}) = \{u, v\}$ of a point $\mathbf{x} = \{x, y, z\}$ is the intersection of vector $\mathbf{q} - \mathbf{p}$ with $z = R$:

$$\mathbf{u} = \{u, v\} = \left\{ \frac{2R}{R+z}x, \frac{2R}{R+z}y \right\},$$

$$\mathbf{x} = \{x, y, z\} = \frac{4R^2}{u^2 + v^2 + 4R^2} \{u, v, R(4R^2 - u^2 + v^2)\}.$$

Fig. 4 The World Ocean in stereographic coordinates. The North Pole is the center, Antarctica constitutes the external boundary

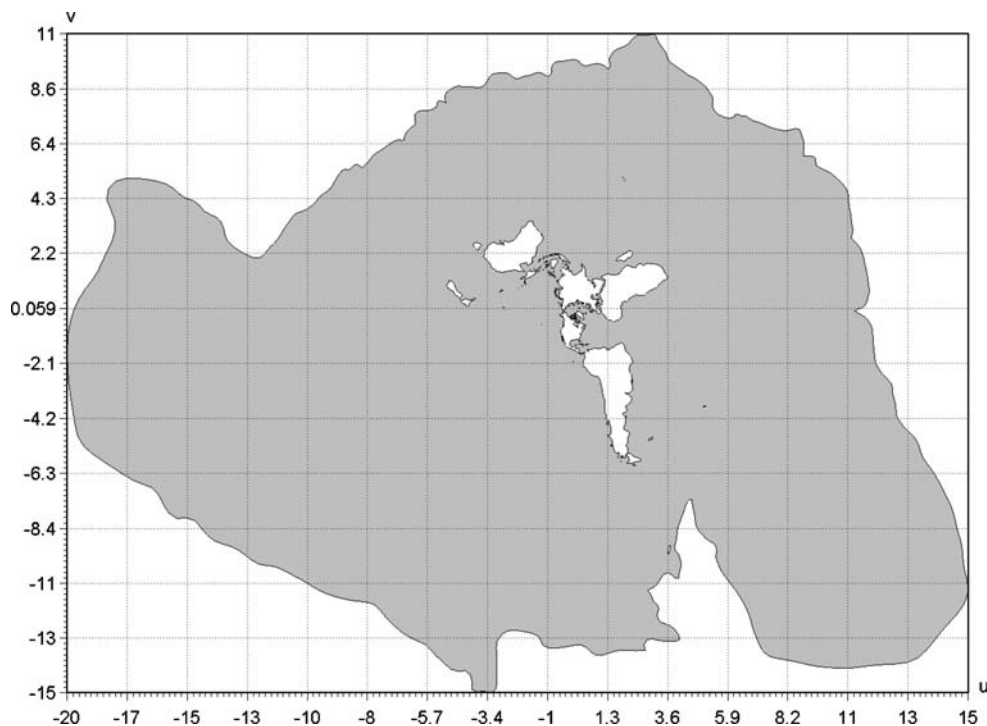


Figure 4 shows the World Ocean in stereographic coordinates $\{u, v\}$. The outside loop surrounding the domain is the stereographic projection of the Antarctica. The radius of the Earth is chosen arbitrarily to $R = 1$. No seam is required to define the overall domain and no singular point exists in the domain of interest.

2.2 Coastlines definition

Today, the most accurate shoreline database is the Global Self-consistent Hierarchical High-resolution Shorelines (GSHHS, (Wessel and Smith 1996)). This data set describes the Earth’s coastlines with a global resolution of about 50 m. GSHHS data are guaranteed to be self-consistent, i.e., coastlines in the database do not intersect themselves. In our approach, we define a size field $\gamma(\mathbf{x})$ that expresses the geometrical requested accuracy of the model at any point \mathbf{x} . The GSHHS data set is coarsened with respect to this size field and every successive point at \mathbf{x} that is closer than $\gamma(\mathbf{x})$ is collapsed.

A coastline is defined in practice as a periodic curve. A first approach could be to use a piecewise linear definition for defining such a curve. A first naive approach would consist in defining such a curve by a piecewise linear interpolation. However, in order to enjoy more flexibility, we use here cubic B-splines with control points taken in the GSHHS database. As

Fig. 5 Great Britain and Ireland with resolutions of 100 km (*top/left*), 15 km (*top/right*), 1.5 km (*bottom/left*), and 150 m (*bottom/right*). Splines control points are depicted on the geometry with the two lowest resolutions



B-splines remain inside the convex hull defined by the control points, it can be shown that, if the piecewise linear representation is convex and consistent, then the curvilinear B-splines representation is also consistent.

In Fig. 5, we generate the coastlines of Great Britain and Ireland with different resolutions. With a resolution $\gamma = 100$ km, we only consider Ireland and Great Britain, and we neglect all smaller islands. With a resolution of $\gamma = 15$ km, 15 contours appear in the domain. Typically, the Isle of Wight is now included. With a resolution of $\gamma = 1.5$ km, the domain contains 152 islands. With a resolution of $\gamma = 150$ m, the domain contains 2,176 islands and a total of 83,277 control points.

3 Mesh generation

As an accurate representation of the boundaries is now available, the next task involves the generation of finite element meshes on curved surface. Two major approaches are available:

- Techniques for which the surface mesh is generated directly in the real 3D space

- Techniques for which the surface mesh is generated in the parametric plane of the surface

When a parametrization of the surface exists, building the mesh in the parametric plane appears to be the most robust choice.

3.1 Definition of a local mesh size field

The aim of the mesh generation process is to build elements of controlled shape and size. Mesh generators are usually able to adapt to a so-called mesh size field. An isotropic mesh size field is a scalar function $\delta(\mathbf{x})$ that defines the optimal length of an edge at position \mathbf{x} of the real space. In the domain of ocean modeling, there exist some heuristics on the way mesh sizes should be distributed in the World Ocean.

The mesh should take into account the bathymetry. The bathymetry $H(\mathbf{x})$ is taken into account in two ways, leading to two fields f_1 and f_2 . Gravity waves move at speed \sqrt{gH} , with g being the acceleration of gravity. The lengthscale λ of a gravity wave is therefore proportional to $\lambda = \mathcal{O}(1/\sqrt{H})$. If we consider that N mesh sizes are necessary to capture one wavelength,

and if λ_{\min} is the smallest wavelength that has to be captured for a reference bathymetry H_{ref} , we define f_1 as

$$f_1(\mathbf{x}) = \frac{\lambda_{\min}}{N} \sqrt{\frac{H_{ref}}{H(\mathbf{x})}}.$$

Another way of taking into account the bathymetry is to force the mesh to capture its variations with a given accuracy (Gorman et al. 2006). Bathymetry can be seen as a scalar field defined at mesh vertices and interpolated piecewise linearly. As the first term of error in its interpolation is supposed to depend on λ_{\max} , the greatest (in absolute value) eigenvalue of the Hessian $\mathcal{H}(x)$,

$$\mathcal{H}(x) = \nabla \nabla \left(\frac{H(x)}{H_{ref}} \right),$$

we define the second field as:

$$f_2(\mathbf{x}) = \frac{1}{\sqrt{\lambda_{\max}}}.$$

In order to represent coastlines well and to capture the small-scale phenomena generated by the friction on the coasts, mesh size should be even smaller near coastlines. This criterion has already been used in the literature (e.g., Legrand et al. 2006). We define a first field $f_3(\mathbf{x})$ as the distance to the closest shoreline:

$$f_3(\mathbf{x}) = d(\mathbf{x}).$$

This field f_3 is also called a *shore proximity function*.

This distance can be computed *in place* using the Approximated Nearest Neighbor Algorithm (Arya et al. 1998).

For each criterion field f_i , a mesh size field δ_i is computed as follows:

$$\delta_i(\mathbf{x}) = \delta_i^{\text{small}} + \alpha_i(\mathbf{x}) \left(\delta_i^{\text{large}} - \delta_i^{\text{small}} \right),$$

where

$$\alpha_i(\mathbf{x}) = \begin{cases} 0 & \text{if } f_i(\mathbf{x}) \leq f_i^{\min} \\ \frac{f_i(\mathbf{x}) - f_i^{\min}}{f_i^{\max} - f_i^{\min}} & \text{if } f_i^{\min} < f_i(\mathbf{x}) < f_i^{\max} \\ 1 & \text{if } f_i(\mathbf{x}) \geq f_i^{\max} \end{cases}$$

with δ_i^{large} and δ_i^{small} as large and small desired mesh sizes and f_i^{\max} and f_i^{\min} as two field values that define the zone of refinement. The final size field is simply computed as the minimum of all size fields:

$$\delta(\mathbf{x}) = \min(\delta_1(\mathbf{x}), \delta_2(\mathbf{x}), \dots).$$

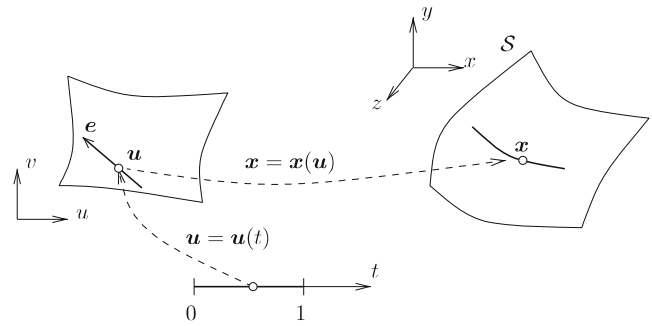


Fig. 6 Three parametrizations of a straight edge

Finally, it is always possible to add other size fields as error estimators that may depend on the finite element solution.

In Fig. 6, we consider a straight edge e described by its vector \mathbf{e} in the parametric plane, where the mesh generation process is performed. Its length L is computed as follows:

$$L = \int_e \sqrt{\|d\mathbf{x}\|^2} = \int_e \sqrt{d\mathbf{u}^T \mathbf{J}^T \mathbf{J} d\mathbf{u}} = \int_0^1 \sqrt{\mathbf{e}^T \mathbf{M} \mathbf{e}} dt$$

where $\mathbf{J} = \partial \mathbf{x} / \partial \mathbf{u}$ is the Jacobian of the mapping and $\mathbf{M} = \mathbf{J}^T \mathbf{J}$ is the metric tensor. In the case of a stereographic projection, both eigenvalues of \mathbf{M} are positive and equal:

$$\lambda(\mathbf{u}) = \left(\frac{4R^2}{u^2 + v^2 + 4R^2} \right).$$

To obtain the mesh in the parametric space but with the right sizing in the real space, a suitable mesh size field $\delta_u(\mathbf{u})$ has to be defined in this parametric plane. As the stereographic projection is a conforming mapping, it can be defined with a simple scaling:

$$\delta_u(\mathbf{u}) = \delta(\mathbf{x}(\mathbf{u})) \frac{1}{\lambda(\mathbf{u})}.$$

3.2 Coastlines mesh generation

Let us consider a curve in the parametric plane $\mathbf{u}(t) : [0, 1] \rightarrow \mathcal{R}^2$. The number of subdivisions N of the curve is the following function of the size field

$$N = \int_0^1 \frac{1}{\delta_u(\mathbf{u}(t))} \|d_t \mathbf{u}\| dt,$$

where $\|d_t \mathbf{u}\| = \sqrt{(\partial_t u)^2 + (\partial_t v)^2}$. The $N + 1$ mesh points on the curve are located at coordinates $\{t_0, \dots, t_N\}$ where t_i is computed using the following rule:

$$i = \int_{t_0}^{t_i} \frac{1}{\delta_u(\mathbf{u}(t))} \|d_t \mathbf{c}\| dt.$$

Integration of those expressions must be performed with an adaptive trapeze rule, as coastlines are discretized with cubic splines that contain a large number of control points. Typically, Europe and Asia are discretized by only one spline with more than 20 thousand control points (Fig. 4).

However, this algorithm does not guarantee that, even if the model edges G_j^1 that constitute the boundaries of the domain are nonintersecting, the corresponding 1D meshes do not self-intersect. Figure 7 shows two islands very close to each other. Yet, even if the geometry is itself not self-intersecting, the first 1D generated mesh intersects itself. This can be considered as a critical issue: modifying the mesh size field *by hand* locally cannot be considered when several thousand islands are to be involved. It is therefore mandatory to define a systematic recovery procedure. Such an algorithm, illustrated in Fig. 7, works as follows:

1. A Delaunay mesh that contains all points of the 1D mesh is initially constructed using a divide-and-conquer algorithm (Dwyer 1986).
2. Missing edges are recovered using edge swaps (Weatherill 1990). If a mesh edge e_i that belongs to the 1D mesh is to be swapped for recovering edge e_j , then the mesh edges e_i and e_j that both belong to the 1D mesh intersect.
3. All intersecting edges e_k are split in two segments and the new point is snapped onto the geometry. Then, we go back to the first step until the list of intersecting edges is empty.

If an intersecting edge is smaller than the geometrical tolerance, then an error message is thrown claiming that the geometry is self-intersecting. Note that when a unique mesh edge connects two different islands, those islands are numerically merged if a nonslip boundary condition is applied along their coastlines.

3.3 Surface mesh generation

To generate a mesh on the sphere, three approaches are available in Gmsh software. All of them start with an initial Delaunay mesh that contains all the mesh

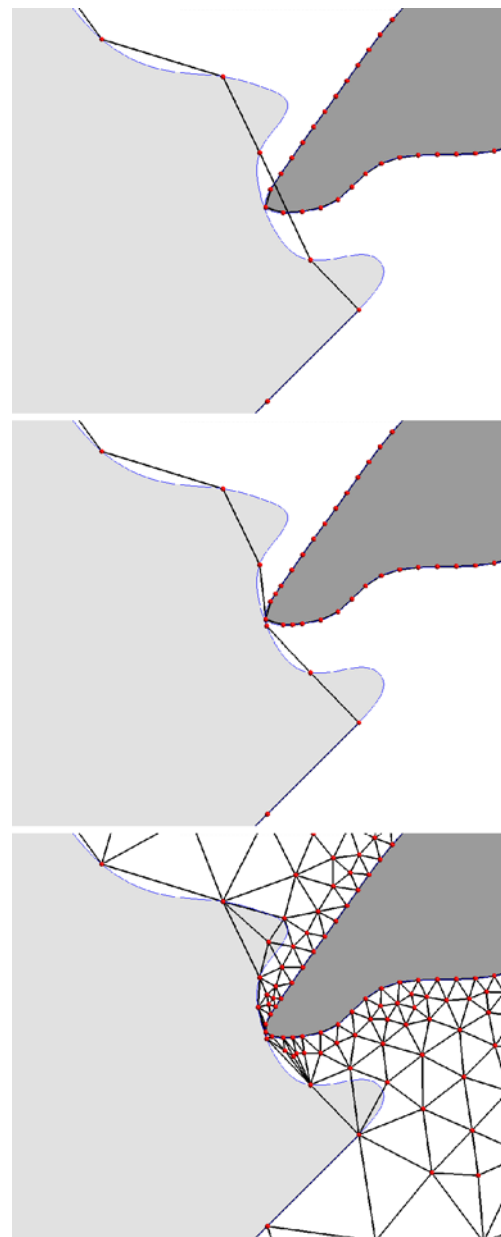


Fig. 7 A geometry with two islands (in light and dark gray) that are very close to each other. The *top image* shows the initial 1D mesh that respects mesh size field. The *middle image* shows the first iteration of the recovery algorithm. The *bottom image* shows the final mesh that was possible to realize after two recovery iterations

vertices of the contours. Then, every mesh edge of the 1D mesh is recovered using edge swaps. Then, internal vertices are iteratively inserted inside the domain. The way points are inserted differently in the three algorithms:

- The *del2d* algorithm is inspired by the work of the GAMMA team at INRIA (George and Frey

2000). New points are inserted sequentially at the circumcenter of the element that has the largest adimensional circumradius. The mesh is then reconnected using an anisotropic Delaunay criterion.

- In the *frontal* algorithm (Rebay 1993), new points are inserted optimally on Voronoi edges. The mesh is then reconnected using the same anisotropic Delaunay criterion as the one in the *del2d* algorithm. Note that this algorithm's implementation only differs slightly from that of algorithm *del2d*.
- The *meshadapt* algorithm is very different from the first two ones. It is based on local mesh modification: This technique makes use of edge swaps, splits, and collapses. Long edges are split, short edges are collapsed, and edges are swapped if a better geometrical configuration is obtained.

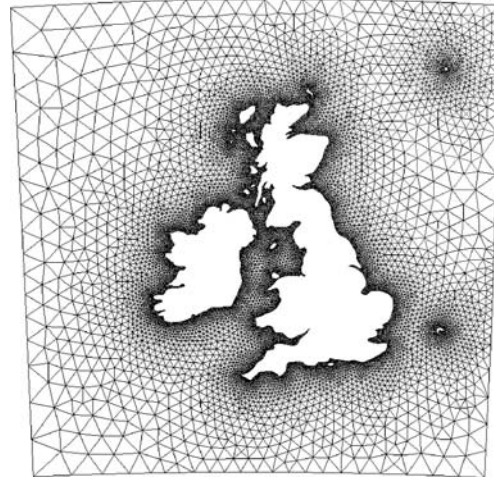
The *frontal* algorithm usually gives the highest-quality meshes while the *del2d* algorithm is the fastest: it produces about five million triangles a minute if the size field δ is not too complex to compute. Figure 8 presents three meshes of one of the models of Fig. 5 for which we have used a shore proximity function as the only size field. Meshes have, respectively, 18,698, 19,514, and 17,154 triangles. The percentage of elements that have an aspect ratio $\rho > 0.9$ is, respectively, 93.2%, 88.1%, and 84.5%. CPU time for generating meshes was, respectively, 0.7, 0.5, and 5.7 s.

Figures 9 and 10 present a mesh of the World Ocean that makes use of all size fields defined in Section 3.1:

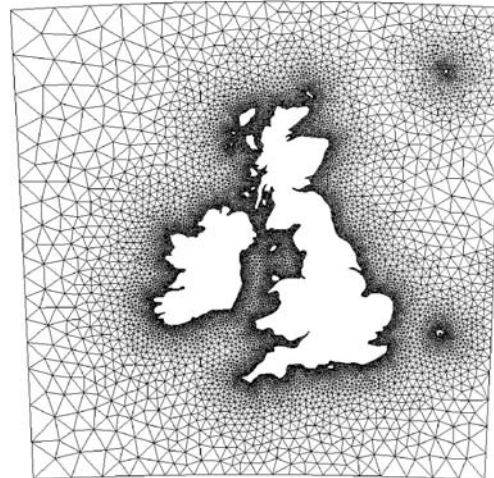
- A shore proximity function f_1 is used with $\delta_1^{\text{small}} = 30$ km, $\delta_1^{\text{large}} = 200$ km, $f_1^{\text{min}} = 0$, and $f_1^{\text{max}} = 500$ km.
- We use f_2 and refine the mesh proportionally to the square root of the ocean depth. The size field δ_2 ranges from 25 to 500 km.
- We use f_3 to capture the bathymetry. The size field δ_3 ranges from 25 to 500 km.

The resulting mesh is generated of 436,409 triangles and the whole mesh generation process (data loading, coastline reduction, 1D mesh generation, 2D mesh generation, output files writing) takes 35 s on a recent laptop. Those timings compare advantageously

a) Mesh done using the *frontal* algorithm



b) Mesh done using the *del2d* algorithm



c) Mesh done using the *meshadapt* algorithm

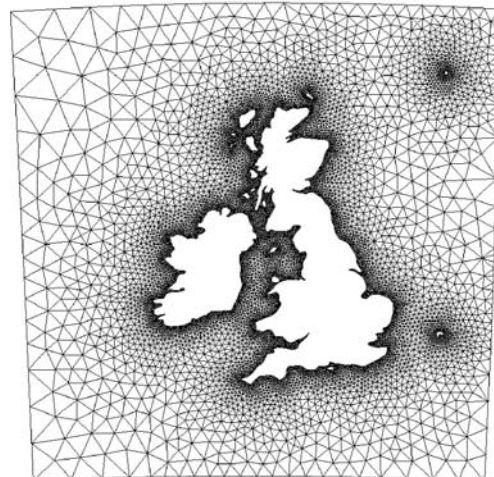


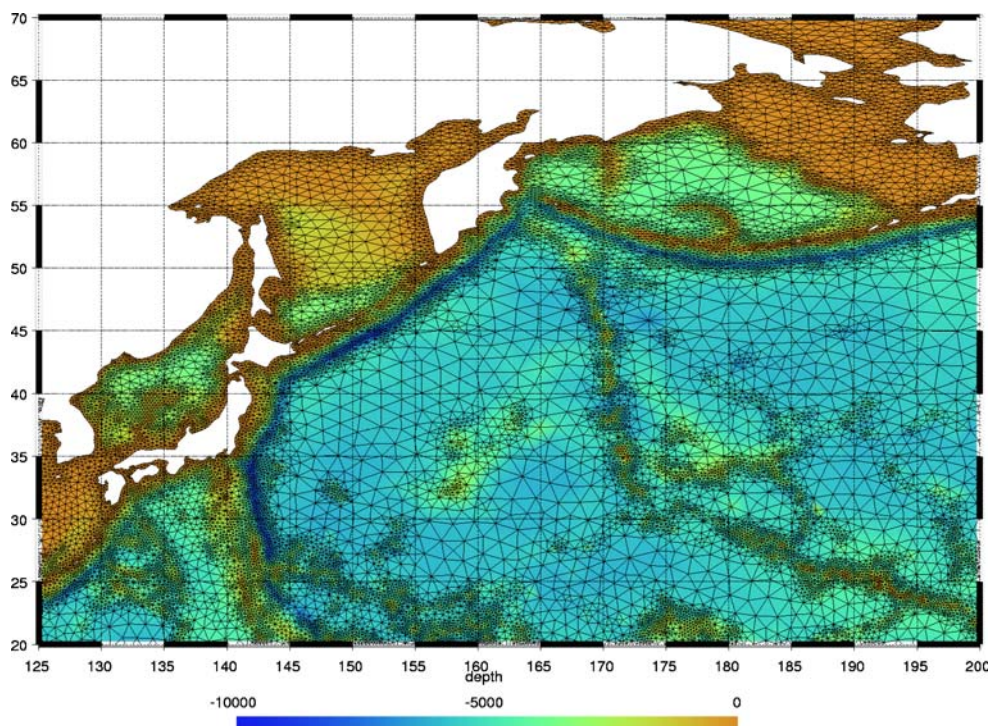
Fig. 8 Meshes of the same domain using three different algorithms (a–c)



Fig. 9 Mesh of the World Ocean. The mesh size field is defined using a shore proximity function, the bathymetry, and its Hessian

with alternative techniques based on mesh decimation (Gorman et al. 2006). The memory footprint of the meshing algorithms is low: about 12 million triangles

Fig. 10 Close up of the mesh of Fig. 9 in the north Pacific Ocean; color levels represent the bathymetry (in meters). The effects of the three refinement rules are clearly visible



(six million nodes) can be generated per gigabyte of memory.

4 Examples

In the mesh generation community, it is assumed that a good paper should present nice pictures of meshes. We will not circumvent that prerequisite. Yet, mesh generation is usually considered as a tool, not as an aim. Therefore, the meshes that we present in this section are accompanied by some simulation results.

4.1 Sea ice modeling

The mesh presented in Fig. 11 was used to investigate the sensitivity of the Arctic sea ice cover features to the resolution of the narrow straits constituting the Canadian Arctic Archipelago. This mesh constitutes of 17,053 triangles with a resolution of 20 km near the islands in the archipelago and 40 km elsewhere. Far from coasts and islands, the resolution decreases up to 300 km. Model results are shown in Fig. 12. A complete description of the model and its validation can be found in Lietaer et al. (2008).



Fig. 11 Mesh of the Arctic region (north of the parallel 50 degrees North) especially refined along coastlines and in the Canadian Arctic Archipelago

4.2 Multiscale model of the Scheldt River

Within the framework of the multidisciplinary project TIMOTHY,¹ our team is presently involved in the development of a 2D hydrodynamic model of the Scheldt Estuary in the Netherlands. Our main goal in this project is to take advantage of the finite element method to study very specific ecological problems, such as the dynamics of fecal bacteria or heavy metals. We also intend to study the characteristic time scales determining the physics and the biology in the estuary. In this model, the tide is forced at the shelf break, which is more than 1,000 km away from the mouth of the estuary, and the upstream boundary is situated in the areas of Antwerp, where the river width is a few hundred meters. The multiscale character of the problem is then one of its main feature and the use of an unstructured grid is thus totally appropriate. Figure 13 presents a mesh of 27,472 elements used in our preliminary runs. Various criteria based on the distance from coasts, islands, and shelf break are used to define the mesh size fields. The element sizes range

¹TIMOTHY, Tracing and Integrated Modeling of Natural and Anthropogenic Effects on Hydrosystems: The Scheldt River basin and adjacent coastal North Sea, <http://www.climate.be/TIMOTHY>.

from 150 m in the Scheldt river near Antwerp to about 50 km away from the coastlines.

4.3 The Great Barrier Reef

Our research team has developed the first multiscale hydrodynamic model of the whole Great Barrier Reef. The Great Barrier Reef is on the continental shelf of the Australian northeastern coastline. There are over 2,500 coral reefs in a strip that is about 2,600 km in length and 200 km in width. The simulation that is presented here makes use of most of the specific mesh generation features that were presented in this paper: a geometric domain with multiple scales, use of a shore proximity function, grid adaptation with respect to the bathymetry, and special refinement in the domain of interest. A complete simulation is described in Lambrechts et al. (2008). Figure 14 shows a mesh built to run small simulations on a single CPU to study a specific region while keeping the boundary conditions of the complete simulation. Around a specific island (Lizard Island), the resolution is sufficient to capture some small-scale hydrodynamic features like tidal jets in small interreef passages and recirculations around islands. Elsewhere on the shelf, the resolution is very coarse, those regions are only used as boundary conditions. Seventy percent of the 20,384 elements are

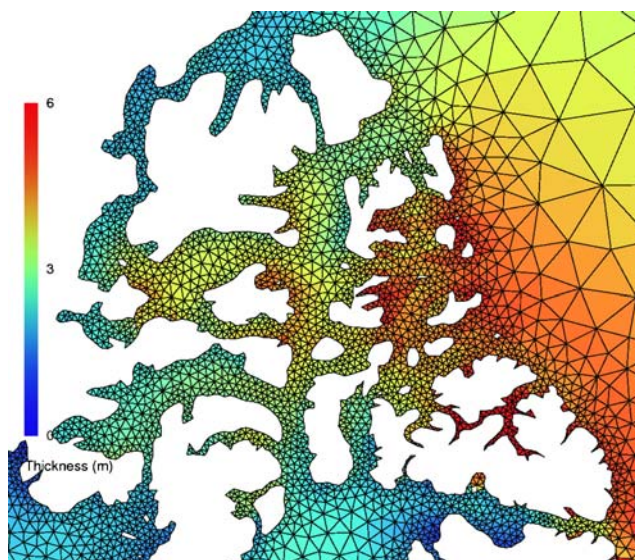
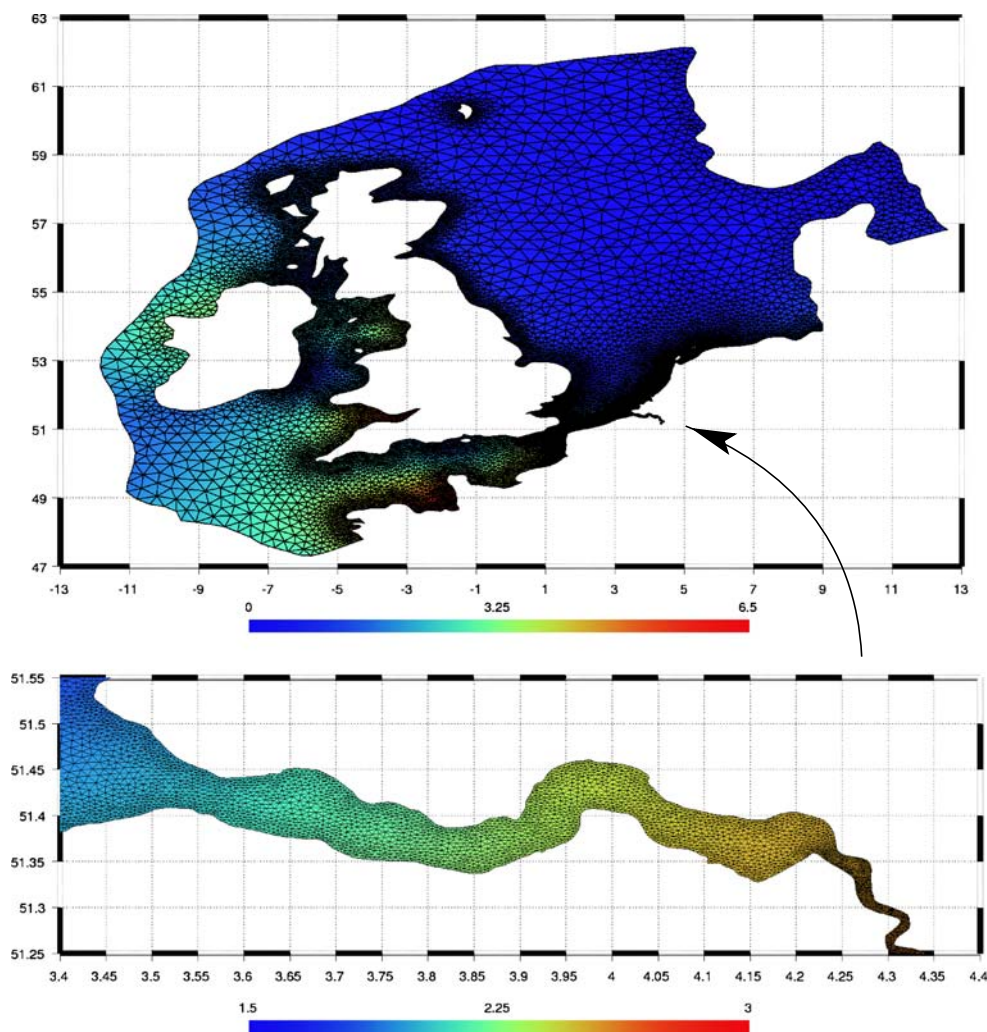


Fig. 12 Detail of the mesh of Fig. 11. Mean March sea ice thickness pattern (in meters) in the western part of the Canadian Arctic Archipelago as computed by the finite element sea ice model (1979–2005)

Fig. 13 Multiscale mesh: North Sea and Scheld River Estuary. *Color levels* represent the amplitude of the M2 tidal component (in meters)



located in the refined region. A plot of velocity vectors is also presented. Tidal jets and eddies due to the interaction of the flow with the topography near the open-sea boundary are clearly visible. Those small-scale features were captured thanks to the accurate description of the bottom topography.

5 Summary

A CAD-based mesh generation procedure for ocean modeling has been developed. The new approach has the advantage of relying on existing well-known engineering mesh generation procedures. The CAD model, based on a smooth BRep of the domain, allows to build a compact, self-consistent, and portable geometric model. Existing robust meshing procedures can be applied to the CAD model. Various meshes can be

constructed based on the same CAD definition, and various meshing algorithms can be used as well. Last but not least, everything that has been described in this paper is now part of Gmsh, a 3D finite element mesh generator with built-in pre- and postprocessing facilities (<http://www.geuz.org/gmsh>). Since Gmsh is open-source (under the GNU General Public License), anyone within the finite element marine modeling community has the opportunity to use this freely.

Acknowledgements The present study was carried out within the scope of the project *A second-generation model of the ocean system*, which is funded by the *Communauté Française de Belgique*, as *Actions de Recherche Concertées*, under contract ARC 04/09-316. This work is a contribution to the SLIM² project.

²SLIM, Second-generation Louvain-la-Neuve Ice-ocean Model, <http://www.climate.be/SLIM>

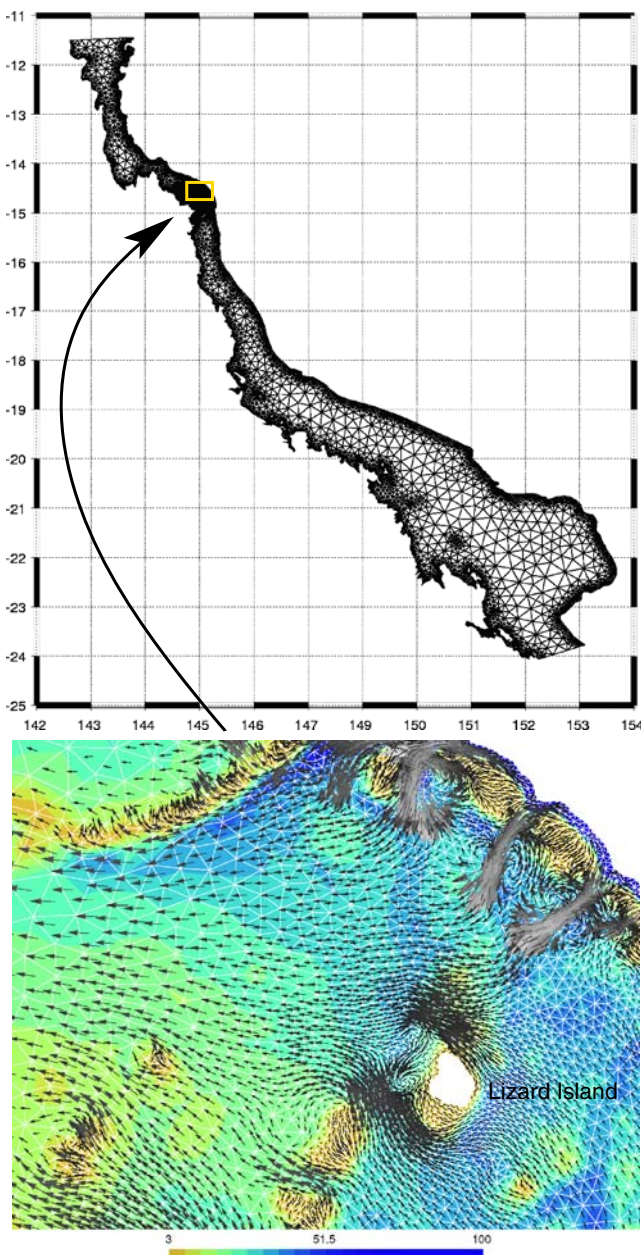


Fig. 14 Coarse mesh of the Great Barrier Reef refined around Lizard Island (*top*). Details of the simulation computed on this mesh in the vicinity of this island (*bottom*). *Color levels* show the depth and the *arrows* indicate the bidimensional velocity field

References

- Arya S, Mount DM, Netanyahu NS, Silverman R, Wu AY (1998) An optimal algorithm for approximate nearest neighbor searching. *J ACM* 45:891–923. <http://www.cs.umd.edu/mount/ANN/>
- Beall MW, Shephard MS (1997) A general topology-based mesh data structure. *Int J Numer Methods Eng* 40(9):1573–1596
- Danilov S, Kivman G, Schröter J (2005) Evaluation of an eddy-permitting finite-element ocean model in the north atlantic. *Ocean Model* 10:35–49
- Dwyer RA (1986) A simple divide-and-conquer algorithm for computing delaunay triangulations in $o(n \log \log n)$ expected time. In: *Proceedings of the second annual symposium on computational geometry*, Yorktown Heights, 2–4 June 1986, pp 276–284
- George P-L, Frey P (2000) *Mesh generation*. Hermes, Lyon
- Gorman G, Piggott M, Pain C (2007) Shoreline approximation for unstructured mesh generation. *Comput Geosci* 33: 666–677
- Gorman G, Piggott M, Pain C, de Oliveira R, Umpleby A, Goddard A (2006) Optimisation based bathymetry approximation through constrained unstructured mesh adaptivity. *Ocean Model* 12:436–452
- Griffies SM, Böning C, Bryan FO, Chassignet EP, Gerdes R, Hasumi H, Hirst A, Treguier A-M, Webb D (2000) Developments in ocean climate modeling. *Ocean Model* 2:123–192
- Hagen SC, Westerink JJ, Kolar RL, Horstmann O (2001) Two-dimensional, unstructured mesh generation for tidal models. *Int J Numer Methods Fluids* 35:669–686 (printed version in Richard's office)
- Haines R (2000) CAPRI: computational analysis programming interface (a solid modeling based infra-structure for engineering analysis and design). Tech. rep., Massachusetts Institute of Technology
- Henry RF, Walters RA (1993) Geometrically based, automatic generator for irregular triangular networks. *Commun Numer Methods Eng* 9:555–566
- Lambrechts J, Hanert E, Deleersnijder E, Bernard P-E, Legat V, Wolanski J-FRE (2008) A high-resolution model of the whole great barrier reef hydrodynamics. *Estuar Coast Shelf Sci* 79(1):143–151. doi:10.1016/j.ecss.2008.03.016
- Le Provost C, Genco ML, Lyard F (1994) Spectroscopy of the world ocean tides from a finite element hydrodynamic model. *J Geophys Res* 99:777–797
- Legrand S, Deleersnijder E, Hanert E, Legat V, Wolanski E (2006) High-resolution, unstructured meshes for hydrodynamic models of the Great Barrier Reef, Australia. *Estuar Coast Shelf Sci* 68:36–46
- Legrand S, Legat V, Deleersnijder E (2000) Delaunay mesh generation for an unstructured-grid ocean circulation model. *Ocean Model* 2:17–28
- Lietaer O, Fichet T, Legat V (2008) The effects of resolving the Canadian Arctic Archipelago in a finite element sea ice model. *Ocean Model* 24:140–152. doi:10.1016/j.ocemod.2008.06.002
- Lyard F, Lefevre F, Letellier T, Francis O (2006) Modelling the global ocean tides: modern insights from FES2004. *Ocean Dyn* 56:394–415
- National Geographic Data Center (2006) ETOPO1 global relief model. <http://www.ngdc.noaa.gov/mgg/global/global.html>.
- Piggott M, Gorman G, Pain C (2007) Multi-scale ocean modelling with adaptive unstructured grids. *CLIVAR Exch*

- Ocean Model Dev Assess 12(42):21–23 (<http://eprints.soton.ac.uk/47576/>)
- Rebay S (1993) Efficient unstructured mesh generation by means of delaunay triangulation and Bowyer-Watson algorithm. *J Comput Phys* 106:25–138
- Weatherill NP (1990) The integrity of geometrical boundaries in the two-dimensional delaunay triangulation. *Commun Appl Numer Methods* 6(2):101–109
- Wessel P, Smith WHF (1996) A global self-consistent, hierarchical, high-resolution shoreline database. *J Geophys Res* 101(B4):8741–8743. <http://www.soest.hawaii.edu/wessel/gshhs/gshhs.html>
- White L, Deleersnijder E, Legat V (2008) A three-dimensional unstructured mesh finite element shallow-water model, with application to the flows around an island and in a wind-driven elongated basin. *Ocean Model* 22:26–47

A multiresolution method for climate system modeling: application of spherical centroidal Voronoi tessellations

Todd Ringler · Lili Ju · Max Gunzburger

Received: 15 July 2008 / Accepted: 6 October 2008 / Published online: 14 November 2008
© Springer-Verlag 2008

Abstract During the next decade and beyond, climate system models will be challenged to resolve scales and processes that are far beyond their current scope. Each climate system component has its prototypical example of an unresolved process that may strongly influence the global climate system, ranging from eddy activity within ocean models, to ice streams within ice sheet models, to surface hydrological processes within land system models, to cloud processes within atmosphere models. These new demands will almost certainly result in the develop of multiresolution schemes that are able, at least regionally, to faithfully simulate these fine-scale processes. Spherical centroidal Voronoi tessellations (SCVTs) offer one potential path toward the development of a robust, multiresolution climate system model components. SCVTs allow for the generation of high-quality Voronoi diagrams and Delaunay triangulations

through the use of an intuitive, user-defined density function. In each of the examples provided, this method results in high-quality meshes where the quality measures are guaranteed to improve as the number of nodes is increased. Real-world examples are developed for the Greenland ice sheet and the North Atlantic ocean. Idealized examples are developed for ocean–ice shelf interaction and for regional atmospheric modeling. In addition to defining, developing, and exhibiting SCVTs, we pair this mesh generation technique with a previously developed finite-volume method. Our numerical example is based on the nonlinear, shallow-water equations spanning the entire surface of the sphere. This example is used to elucidate both the potential benefits of this multiresolution method and the challenges ahead.

Keywords Voronoi diagram · Delaunay triangulation · Climate modeling · Multiresolution

Responsible Editor: Eric Deleersnijder

LA-UR-08-05303.

T. Ringler (✉)
T-3 Fluid Dynamics Group, Theoretical Division,
Los Alamos National Laboratory, Los Alamos,
NM 87545, USA
e-mail: ringler@lanl.gov

L. Ju
Department of Mathematics, University of South Carolina,
1523 Greene Street, Columbia, SC 29208, USA
e-mail: ju@math.sc.edu

M. Gunzburger
School of Computational Science, 400 Dirac Science Library,
Florida State University, Tallahassee, FL 32306-4120, USA
e-mail: gunzburg@scs.fsu.edu

1 Introduction

Climate system models (CSMs) are an increasingly important tool for assessing anthropogenic climate change. CSMs, along with observations and theory, form the basis for the Intergovernmental Panel on Climate Change (IPCC) Working Group 1 Assessment Reports that detail the anticipated consequences of rising concentrations of atmospheric greenhouse gases (International Panel on Climate Change 2007). While CSMs have been highly successful in interpreting observations, confirming theory and providing gross estimates of climate sensitivity, the climate modeling

community will be challenged in the coming decade to extend the utility of CSMs well beyond their current scope.

At least two drivers are pushing CSMs into new and expanding roles. The first is the increasingly urgent need to resolve scales and processes that are far beyond the current scope of these models. There are likely to be unresolved processes and currently misrepresented processes that have significant influence on the global climate system. Every component of the Earth system has its own prototypical example, ranging from eddy activity within ocean models (Hallberg and Gnanadesikan 2006), to ice streams within ice sheet models (Joughin et al. 1999), to surface hydrological processes within land system models (Newman et al. 2006), to cloud processes with atmosphere models (Tomita et al. 2007). All of these processes are not faithfully included in IPCC-class CSMs primarily due to lack of resolution; the degrees of freedom required to comprehensively simulate these processes are computationally prohibitive given the current (and foreseeable) resources.

The second driver pushing the evolution of CSMs is the rapidly growing demand for high-fidelity assessments of regional climate change driven by increasing concentrations of atmospheric greenhouse gases. As appreciation for the possible consequences of anthropogenic climate change improves, we are confronted with the need to characterize the regional aspects of climate change in order to support adaptation and mitigation strategies. As indicated by the last chapter of the IPCC WG1 Fourth Assessment Report (AR4), the push in this direction is already underway (International Panel on Climate Change 2007). To be successful in providing the relevant information regarding regional climate impacts, CSMs will require significant increases in resolution, at least regionally, along with the incorporation of new processes.

The magnitude of the problem that must be addressed by the climate modeling community in order to transition from coarse-grain global CSMs to robust multiresolution CSMs is portrayed in chapter 8 of AR4. Every one of the 23 models contributing to AR4 utilizes an ocean model based on structured quadrilateral grids using low-order (\sim second-order) numerics based on compact finite-difference/finite-volume stencils (See Table 8.2.1 in AR4). In addition, as a group, these 23 models showed a rapid migration in their atmospheric component from global spectral methods to finite-volume methods built on traditional latitude–longitude grids. Not a single model contributing to AR4 utilized unstructured grids or multiresolution methods. So, while the scientific and societal needs for multireso-

lution CSMs are strikingly clear, the path to that end is not at all obvious.

CSM components are presently testing various types of quasiuniform tessellations (also referred to as grids or meshes) to discretize the surface of the sphere. These quasiuniform tessellations are a significant improvement over their predecessor, latitude–longitude grids, by removing both the strong grid-pole singularities and the accompanying numerical filters required to regularize these singularities. Various types of meshes have been proposed as alternatives to the traditional latitude–longitude grid. For example, the cubed-sphere, which offers the same topological structure as the latitude–longitude grid without the strong pole singularities, has been successfully implemented in various efforts (McGregor 1996; Adcroft et al. 2004; Nair et al. 2005). Voronoi tessellations (also referred to as geodesic, icosahedral or hexagonal grids) have sometimes been chosen for their remarkable uniformity and isotropy (Ringler et al. 2000). Finally, closely related to these Voronoi tessellations are the Delaunay triangulations that have been successfully implemented in an idealized setting and are now being integrated into full CSMs (Bonaventura and Ringler 2005; Comblen et al. 2008). While all of these methods have successfully removed the grid pole singularities associated with latitude–longitude grids, it is not clear that any of these methods, as presently formulated, will be able to meet the challenges outlined above.

By their nature, quasiuniform tessellations imply a substantial increase in computational costs with an increase in horizontal resolution. A halving of the nominal grid spacing implies an increase in computational cost of approximately a factor of eight; a factor of four arises from doubling the degrees of freedom in each of the horizontal directions and a factor of two arises due to halving the time step. The computational burden associated with increasing resolution everywhere within the domain quickly exhausts available computational resources. For example, conducting eddy-resolving ocean simulations as a part of century-long coupled climate simulations is impracticable now and will likely continue to be so for at least the next decade or more. The current NCAR coupled CSM (Collins et al. 2006) uses an ocean component model with a 320×384 grid and a time step of approximately 1 h. The eddy-resolving version of this ocean model uses a $3,600 \times 2,400$ grid and a time step of approximately 6 min (Maltrud and McClean 2004). The two configurations differ by a factor of about 1,000 in terms of their computational burden. Similarly daunting computational burdens are found in ice sheet modeling, surface hydrology modeling, and atmospheric modeling.

The obvious implication here is that CSMs will not be able to fulfill their expanding roles by *solely* using quasiuniform tessellations. The corollary to this assertion is that multiresolution schemes will be required if CSMs are to meet the growing challenges over the next decade. In many ways, ocean models are ahead of the other climate system components in developing models amenable to multiresolution modeling (e.g. Stuhne and Peltier 2006; Giraldo and Warburton 2008). Yet, even within the limited context of ocean modeling, a large gap remains between the idealized dynamical-core simulations conducted to date and the goal of full-physics simulations of the real ocean. Clearly, a host of scientific complexities arise as we begin to contemplate the construction of a multiresolution IPCC-class CSM.

The recent work of St-Cyr et al. (2007) clearly indicates that a successful multiresolution scheme requires attention to the combination of method and mesh. The authors develop a multiresolution mesh by implementing an adaptive, hierarchical nesting technique in which quadrilateral elements are bisected to locally increase resolution. When this adaptive meshing technique was used in combination with a high-order spectral method, the resulting multiresolution scheme produced positive results. Alternatively, when the same technique was used in combination with a low-order, finite-volume technique, the results were equivocal at best; adding degrees of freedoms did not reduce numerical solution error. The implication is that robust, multiresolution climate system components will require close attention to both the quality of the variable-resolution meshes and to the numerical techniques we place “on top” of these meshes.

An alternative to hierarchical nesting is to produce a smoothly varying tessellation. By their design, smoothly varying tessellations provide strong control over the spatial patterns of truncation error. While this control may be superfluous when used in combination with high-order methods, it may prove to be critically important when used with the low-order, finite-volume methods that are ubiquitous in IPCC-class component models discussed above. The primary purpose of this paper is to develop a class of robust, variable-resolution meshes, called spherical centroidal Voronoi tessellations (SCVTs), that have the requisite characteristics necessary to meet the present and future challenges of climate system modeling.

SCVTs contain a host of qualities that should produce tangible benefits in the context of climate modeling. First, SCVTs are a superset of the quasiuniform Voronoi tessellations currently being used in the climate modeling community (e.g., Randall et al. 2002; Satoh et al. 2008). Thus, SCVTs are a logical extension

to meshes already being utilized. Second, as discussed in Section 2, even nonuniform SCVTs always produce smoother, more locally uniform meshes as the degrees of freedom are increased. The implication here is clear; SCVTs offer a robust means of producing multiresolution meshes that are guaranteed to increase in quality as computational resources grow. As discussed in Section 3, the technique to produce variable-resolution SCVTs is intuitive and straightforward to implement. Finally, each SCVT is associated with a Delaunay triangulation. The positive attributes associated with the SCVTs are also present in the associated Delaunay triangulation. While a tremendous amount of work is required to translate these positive attributes into robust climate simulations, we begin the task here by taking two steps. First, we develop example meshes for several types of climate system components to demonstrate the method’s ability to produce high-quality, variable-resolution meshes in a diversity of systems. Second, we will demonstrate the ability of these meshes to reduce solution errors, at least locally, in the context of the shallow-water system.

The two primary purposes of this paper are the following: (1) to introduce the climate modeling community to the basic principles of SCVT and (2) to highlight the broad applicability of SCVT to climate system modeling. Section 2 introduces and develops the mathematical foundation for SCVTs. Section 3 develops real-world SCVTs in the context of ice sheet and ocean modeling. Section 4 combines our SCVT technique for the generation of multiresolution meshes with a low-order, finite-volume technique to produce a prototype multiresolution scheme that is broadly applicable to climate system modeling. We look toward future developments of SCVT and draw some conclusions in Section 5.

2 Centroidal Voronoi tessellations

2.1 Definitions

Let Ω denote an open domain or a piecewise smooth hyper-surface in \mathbb{R}^d and $\|\cdot\|$ the corresponding standard Euclidean metric for \mathbb{R}^d . Given a set of distinct points $\{\mathbf{x}_i\}_{i=1}^n \subset \Omega$, we define

$$V_i = \{\mathbf{y} \in \Omega \mid \|\mathbf{x}_i - \mathbf{y}\| < \|\mathbf{x}_j - \mathbf{y}\| \text{ for } j = 1, \dots, n \text{ and } j \neq i\}$$

for $i = 1, \dots, n$. Clearly, $\{V_i\}_{i=1}^n$ forms a tessellation of Ω , i.e., the union of \bar{V}_i spans $\bar{\Omega}$ written as $\cup_{i=1}^n \bar{V}_i = \bar{\Omega}$.

We refer to $\{V_i\}_{i=1}^n$ as the *Voronoi tessellation* or *Voronoi diagram* (Okabe et al. 2000) of Ω associated with the point set $\{\mathbf{x}_i\}_{i=1}^n$. A point \mathbf{x}_i is called a *generator* and a subregion V_i is referred to as the *Voronoi region* corresponding to the generator \mathbf{x}_i . The duality (in a graph-theoretical sense) of a Voronoi tessellation of Ω is the well-known *Delaunay tessellation*, which always consists of triangles/tetrahedra. Algorithms for the construction of corresponding Voronoi diagrams and Delaunay triangulations have been well developed (Okabe et al. 2000; Renka 1999).

Given a density function $\rho(\mathbf{x})$ defined on Ω , for any region $V \subset \Omega$, we call \mathbf{x}^c the *constrained mass centroid* of V with respect to Ω if

$$\mathbf{x}^c = \arg \min_{\mathbf{x} \in V} \int_V \rho(\mathbf{y}) \|\mathbf{y} - \mathbf{x}\|^2 \, d\mathbf{y}. \tag{1}$$

The existence of solutions of Eq. 1 can be easily obtained using the continuity and compactness of the object function; however, solutions may not be unique. It is worth noting that, if Ω is an open domain or a flat hyper-surface, then \mathbf{x}^c coincides with \mathbf{x}^* , the *standard mass centroid* of V defined by

$$\mathbf{x}^* = \frac{\int_V \mathbf{y} \rho(\mathbf{y}) \, d\mathbf{y}}{\int_V \rho(\mathbf{y}) \, d\mathbf{y}}$$

that is much easier to compute.

A Voronoi tessellation of Ω is called a *constrained centroidal Voronoi tessellation* (CCVT) (Du et al. 2003a) if and only if the points $\{\mathbf{x}_i\}_{i=1}^n$ that serve as the generators of the associated Voronoi tessellation

$\{V_i\}_{i=1}^n$ are also the constrained mass centroids $\{\mathbf{x}^c\}_{i=1}^n$ of those regions, i.e., if and only if

$$\mathbf{x}_i = \mathbf{x}_i^c, \quad i = 1, \dots, n. \tag{2}$$

We often refer to the relation (2) as the *centroidal Voronoi tessellation* (CVT) property. The dual Delaunay grid is then called *constrained centroidal Voronoi Delaunay triangulation*. We remark that, when Ω is an open domain in \mathbb{R}^d , $\{\mathbf{x}_i, V_i\}_{i=1}^n$ is often just called a *CVT* (Du et al. 1999).

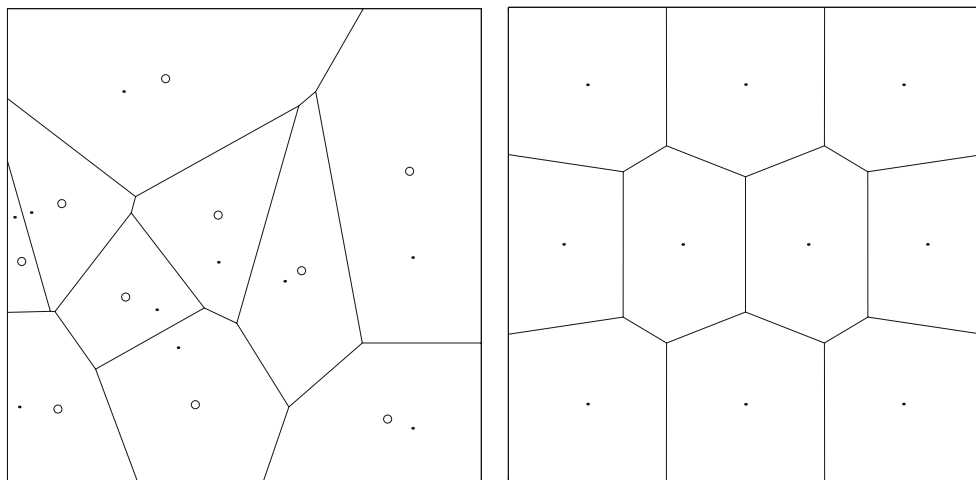
A very important case should be specially addressed for the application of CVT/CCVT to climate system modeling, that is, Ω denotes the surface of a sphere in \mathbb{R}^3 or part of it. In this case, we often refer to $\{\mathbf{x}_i, V_i\}_{i=1}^n$ as a *SCVT*. It is easy to verify that

$$\mathbf{x}^c = r \frac{\mathbf{x}^*}{\|\mathbf{x}^*\|}, \tag{3}$$

where r denotes the radius of the sphere, so that \mathbf{x}^c can be easily computed by first determining \mathbf{x}^* .

General Voronoi tessellations do not satisfy the CVT property; see Fig. 1 for an illustration. A square domain is randomly seeded with ten points (dots in Fig. 1, left panel). These ten points serve as generators for the Voronoi tessellation (cell boundaries in Fig. 1, left panel). For each Voronoi region, the standard mass centroid (open circles in Fig. 1, left panel) is computed. As discussed below in Section 2.3, a simple iterative calculation *regularizes* the initial Voronoi diagram on the left to the diagram shown to the right. In this example, we used a constant density to compute the standard mass centroid; a variable density field would have biased the resulting generator points toward the region of high density. This relationship between density and generator position is the key aspect of this grid-generation method.

Fig. 1 *Left:* a Voronoi tessellation of a square in \mathbb{R}^2 with 10 generators (the dots) randomly selected (the circles denote the centroid of Voronoi region); *right:* a 10-point CVT with a constant density function throughout the domain



2.2 Properties

Given any set of points $\{\tilde{\mathbf{x}}_i\}_{i=1}^n$ on Ω and any tessellation $\{\tilde{V}_i\}_{i=1}^n$ of Ω , we define the *energy* functional

$$\mathcal{K} \left(\{\tilde{\mathbf{x}}_i, \tilde{V}_i\}_{i=1}^n \right) = \sum_{i=1}^n \int_{\tilde{V}_i} \rho(\mathbf{y}) \|\mathbf{y} - \tilde{\mathbf{x}}_i\|^2 d\mathbf{y}.$$

The energy is often referred to as some physical quantity such as *variance*, *cost*, *distortion error*, or *mean square error* in practical applications. A priori, there is no assumed relation between the point set $\{\tilde{\mathbf{x}}_i\}_{i=1}^n$ and the tessellation $\{\tilde{V}_i\}_{i=1}^n$. However, it can be shown that $\mathcal{K}(\cdot)$ is minimized only if $\{\tilde{\mathbf{x}}_i, \tilde{V}_i\}_{i=1}^n$ is a CVT/CCVT (Du et al. 1999, 2003a). Thus, CVTs/CCVTs are Voronoi tessellations for which the generators are, in some sense, *optimally distributed*.

Let us set $\tilde{d} = d$ if Ω is an open domain and $\tilde{d} = d - 1$ if Ω is a hyper-surface in \mathbb{R}^d . Specially, for SCVTs, we have $\tilde{d} = 2$. As a consequence, CVT/CCVT meshes in \mathbb{R}^d have many good geometric properties, including the following (Du et al. 1999, 2003a; Du and Wang 2003):

- For a constant density function, the generators $\{\mathbf{x}_i\}_{i=1}^m$ are uniformly distributed across Ω .
 - Most Voronoi regions are (nearly) congruent (Gersho and Gray 1992; Du et al. 1999). Specially, for SCVTs, they are primarily convex spherical hexagons.
 - The mesh size h (as defined below in Eq. 5) is approximately proportional to $n^{-1/\tilde{d}}$
- For a nonconstant density function, the generators $\{\mathbf{x}_i\}_{i=1}^m$ are still locally uniformly distributed, and it is conjectured (and computationally verified) that, asymptotically,

$$\frac{h_{V_i}}{h_{V_j}} \approx \left(\frac{\rho(\mathbf{x}_j)}{\rho(\mathbf{x}_i)} \right)^{\frac{1}{\tilde{d}+2}}. \tag{4}$$

- The relationship between the relative sizes of Voronoi regions (i.e., grid cells) is controlled *entirely* by the specified density function.
- CVT/CCVT generators tend to accumulate in regions having relatively high values of ρ while remaining locally very regular.
- Thus, in principle, one could control the distribution of generators to minimize the error (either locally or globally) in the solution of a partial differential equation by, e.g., connecting the density function $\rho(\mathbf{x})$ to some a priori or a posteriori error estimates.

It is important to note that we are restricting our analysis and discussion to meshes generated with respect to an isotropic metric. We make this choice because we are currently interested in the methods ability to construct highly uniform, variable-resolution meshes. Extensions to anisotropic meshes have already been developed and are available for use in climate modeling if the need arises (Du and Wang 2005; Du et al. 2005).

2.3 Algorithms

Construction of CVT/CCVT is usually done by either probabilistic methods typified by MacQueen’s random algorithm (MacQueen 1967) (which is a simple iteration between sampling and averaging points) or deterministic methods typified by Lloyd iteration (Lloyd 1982) (which is a simple iteration between constructing Voronoi diagrams and mass centroids). Due to the low convergence rate of MacQueen’s method (MacQueen 1967), much attention has been focused on Lloyd method described below:

Algorithm 1 (Lloyd method) *Given a domain Ω , a density function $\rho(\mathbf{x})$ defined on Ω , and a positive integer n .*

0. *Select an initial set of n points $\{\mathbf{x}_i\}_{i=1}^n$ on Ω ;*
 1. *Construct the Voronoi regions $\{V_i\}_{i=1}^n$ of Ω associated with $\{\mathbf{x}_i\}_{i=1}^n$;*
 2. *Determine the (constrained) mass centroids of the Voronoi regions $\{V_i\}_{i=1}^n$; these centroids form the new set of points $\{\mathbf{x}_i\}_{i=1}^n$;*
 3. *If the new points meet some convergence criterion, return $\{(\mathbf{x}_i, V_i)\}_{i=1}^n$ and terminate; otherwise, go to step 1.*
-

Referring to Fig. 1, the process is as follows: The initial point set, shown as dots in the left panel, represents step 0. Step 1 is shown by the solid cell boundary lines in Fig. 1 (left panel). Step 2, the location of the cell centroids, is shown by the open circles in Fig. 1 (left panel) and forms the new point set from which we compute the new Voronoi diagram. The final result, after satisfying the convergence criterion in step 3, is shown in Fig. 1 (right panel). It should be noted that the mesh generation procedure utilizes, at most, the Voronoi diagram. While no reference to the Delaunay triangulation is required, the properties of smoothness and uniformity convey from the (S)CVT to the Delaunay triangulation.

2.4 Quality measure of Voronoi cells and Delaunay triangles

For the Voronoi cell V_i associated with the generator \mathbf{x}_i , we define its size to be

$$h_{V_i} = 2 \max_{\mathbf{y} \in V_i} \|\mathbf{x}_i - \mathbf{y}\|. \quad (5)$$

Then, h_{\max}/h_{\min} can be used to measure the global nonuniformity of the given Voronoi mesh where $h_{\max} = \max_i h_{V_i}$ and $h_{\min} = \min_i h_{V_i}$. In order to measure the local uniformity or quality of Voronoi cells of the SCVT, we use the following σ measure (Du et al. 2003b). For the Voronoi cell V_i associated with the generator \mathbf{x}_i ,

$$\sigma(V_i) = \frac{\min_j \|\mathbf{x}_i - \mathbf{x}_j\|}{\max_j \|\mathbf{x}_i - \mathbf{x}_j\|}, \quad (6)$$

where \mathbf{x}_j values denote Voronoi neighbors of \mathbf{x}_i . Clearly, $0 < \sigma \leq 1$ and $\sigma = 1$ correspond to the equilateral polygons. We then set

$$\sigma_{\min} = \min_i \sigma(V_i) \quad \text{and} \quad \sigma_{\text{avg}} = \frac{1}{n} \sum_i \sigma(V_i).$$

where n denotes the number of Voronoi cells. σ_{\min} measures the quality of the worst Voronoi cell and σ_{avg} measures the average quality of the Voronoi mesh.

In the Delaunay triangulation, the size of a triangle, T , is defined to be its longest side length, h_T . We apply the commonly used q -measure (Field 2000) to evaluate the quality of the dual triangular mesh (Delaunay triangles), where, for any triangle T , q is defined to be twice the ratio of the radius R_T of the largest inscribed circle and the radius r_T of the smallest circumscribed circle, i.e.,

$$q(T) = 2 \frac{R_T}{r_T} = \frac{(b+c-a)(c+a-b)(a+b-c)}{abc}, \quad (7)$$

where a , b , and c are side lengths of T . Clearly, $0 < q \leq 1$ and $q = 1$ correspond to the equilateral triangle. For a given triangulation, \mathcal{T} , composed of m triangles, we define

$$q_{\min} = \min_{T \in \mathcal{T}} q(T) \quad \text{and} \quad q_{\text{avg}} = \frac{1}{m} \sum_{T \in \mathcal{T}} q(T).$$

where q_{\min} measures the quality of the worst triangle and q_{avg} measures the average quality of the triangular mesh \mathcal{T} .

It is worth noting that the energy \mathcal{H} associated with the Voronoi tessellation $\{(\mathbf{x}_i, V_i)\}_{i=1}^n$ decreases monotonically during the Lloyd iteration if $\{(x_i, V_i)\}_{i=1}^n$ has not reached a CVT/CCVT yet. In certain systems, or with certain methods, we may require nodes to be located on the boundary of the problem domain Ω .

CVTs/CCVTs and the above construction algorithm can be easily generalized so that some of the generators are constrained to lie on the boundary $\partial\Omega$ (Du and Wang 2003; Ju et al. 2006).

3 Example meshes

3.1 Land ice: Greenland

3.1.1 Motivation

The Greenland and Antarctic ice sheets are characterized by a wide range of spatial and temporal scales. In terms of spatial scales, each of these ice sheets spans several thousand kilometers. Interior regions of these ice sheets are characterized by relatively broad spatial scales on the order of 100 km or more. These interior regions are generally areas of net accumulation of mass due to atmospheric precipitation of water. As this net source of water is exported toward the ocean in the form of ice, relatively fast-moving ice streams form within each catchment zone. Not unlike their liquid water counterparts on land, these ice streams are long and thin with along-stream scales of several hundred kilometers and cross-stream scales often less than 10 km (Joughin et al. 1999). In addition, the shear zone separating the fast-moving ice streams from the adjacent nearly stationary ice is characterized by scales of 1 km or less. These ice streams transport the majority of ice volume from Greenland and Antarctica into the surrounding ocean (Rignot et al. 2008). As a result, robust predictions of sea-level rise will require an accurate simulation of ice stream dynamics. In fact, the recent IPCC WG1 AR4 document declined to draw substantive conclusions on the likelihood of rapid sea level rise during the twenty-first century because, in part, ice stream dynamics are not included in current ice sheet models (International Panel on Climate Change 2007).

In addition to the kinematically driven need for locally enhanced resolution, there is also a desire for increased resolution at the ice margin due to significant seasonal ablation. This is particularly relevant for Greenland, where intense melting occurs annually below 1,200 m. The subsequent transport of this meltwater to the bottom of the ice sheet may have a strong impact on basal sliding processes (Bell 2008). In the context of ice sheet modeling, we see both kinematics and physics as driving the need for locally enhanced resolution. Spatial resolutions below 1 km might be required for the accurate representation of these processes. When considering an entire ice sheet, a uniform

mesh of 1 km combined with the emerging three-dimensional Stokes solvers is not computationally tractable. So, instead, we turn to variable-resolution SCVTs to discretize this system.

3.1.2 Proxy for SCVT density

In this example, we will generate a mesh of Greenland that places enhanced resolution in the vicinity of ice streams. Our target resolutions for this grid are $h_{\min} = 2$ km in the vicinity of the ice margin and $h_{\max} = 100$ km in the interior; let $R = h_{\max}/h_{\min}$ represent the ratio of these target resolutions. Our estimates suggest that the resulting mesh will be computationally tractable for climate change simulations, even when used in combination with a full three-dimensional Stokes solver.

Figure 2 shows an observational estimate of surface ice velocity at a spatial resolution of 2.5 km (Bamber et al. 2000). The magnitude of velocity, $\|\mathbf{V}\|$, is presented on a \log_{10} scale and ranges from a minimum of approximately 0.1 km/year along ridgelines to over 10.0 km/year at the outlet of some ice streams. We use this observational data set in two ways. First, this data set allows us to define the location of the ice boundary of Greenland as a set of piece-wise linear loops (not shown) within which we develop the SCVT. Second, within each loop, we define the SCVT density function with the following sequence:

$$x = \log_{10}(\|\mathbf{V}\|), \quad x_{\min} = -0.5, \quad x_{\max} = 3.5 \tag{8}$$

$$x = \max(x, x_{\min}) \tag{9}$$

$$x = \min(x, x_{\max}) \tag{10}$$

$$\rho = \left(\frac{(x - x_{\min})}{(x_{\max} - x_{\min})} * R \right)^4 + 1 \tag{11}$$

We limit the lower bound on the SCVT density function such that all regions with surface velocities less than $x_{\min} = -0.5$ or 0.3 km/year receive the same resolution. We limit the upper bound on the SCVT density function such that all regions with surface velocities more than $x_{\max} = 3.5$ or 3 km/year receive the same resolution. Finally, density is normalized such that it ranges from 1 to R^4 to generate nominal grid cells spacings that vary by a factor of R ; see Eq. 4.

Figure 3 depicts the resulting Voronoi diagram using 25,936 nodes, resulting in a minimum resolution of approximately 4 km. Figure 4 shows the \log_{10} of the Voronoi cell area. We find broad regions of low resolution along the quiescent ice ridgelines with the vast majority of nodes placed in the vicinity of ice streams.

The color scale is saturated for all cell areas greater than 150 km² and for all cell areas less than 10 km². While 90% of the cells are spaced less than 10 km from their neighbors, 10% of the cells with grid spacing greater than 10 km cover approximately 40% of the ice domain.

We progressively add nodes into the domain until our target minimum grid resolution of 2 km is reached. Figure 5 shows the \log_{10} of Voronoi cell area using 101,115 nodes. In this figure, the color scale is saturated for all cell areas greater than 75 km² and for all cell areas less than 5 km². Figure 5 looks identical to Fig. 4; the only noticeable difference between the figures is the scale on the colorbar.

Table 1 presents the global quality metrics for the Greenland SCVTs. The quality histograms of the SCVTs are shown in Fig. 6. The bulk measures of uniformity shown in Table 1 show improvement in every category as resolution is increased. The histograms shown in Fig. 6 indicate a systematic shift toward higher mesh quality with increasing degrees of freedom. It is equally important that, as the number of nodes is increased, the histograms exhibit a noticeable reduction in the proportion of cells residing in the “low-quality” end of the histograms.

3.2 Ocean: North Atlantic

3.2.1 Motivation

Incorporating eddies into IPCC-class global ocean simulations remains a computational challenge. Eddy-resolving simulations typically require grid resolutions of approximately 10 km, implying approximately 5e6 degrees of freedom to span the global ocean surface. This is in stark contrast to typical IPCC simulations that currently use approximately 5e4 degrees of freedom to cover the same extent. The factor of 100 separating the two simulations is compounded by another factor of 10 since eddy-resolving simulations require a significantly shorter time step. The 1,000-fold increase in computational burden to move from resolutions presently used in IPCC-class simulations to global, eddy-resolving resolutions is currently beyond reach and will likely remain so for a decade or more.

An alternative and computationally tractable approach is to employ variable-resolution grids, such as SCVTs, to permit eddy-resolving resolutions at targeted locations. These variable-resolution grids could be employed in limited area domains or as part of a global ocean simulation. The ability to readily generate variable-resolution meshes for the global ocean system allows us to consider the notion of an optimal spatial

Fig. 2 Log base 10 distribution of ice velocity at the surface of the Greenland ice sheet (Bamber et al. 2000). The scale is saturated for all velocities above 10 km/year and for all velocities below 0.1 km/year

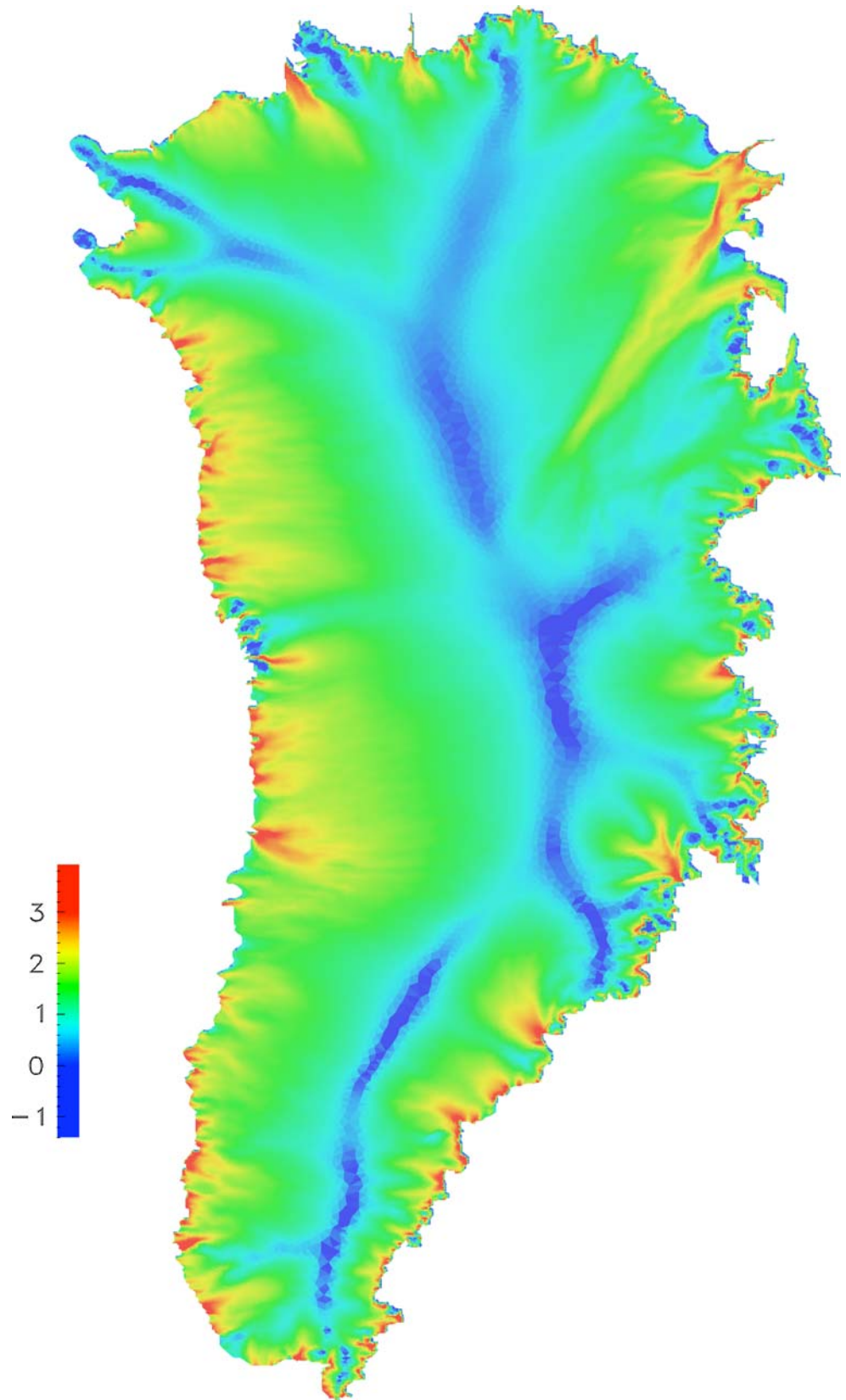


Fig. 3 A SCVT of the Greenland ice sheet using 25,936 nodes based on the density function given in Eq. 11. Note that grid cells within the ice streams are often too small to be visible

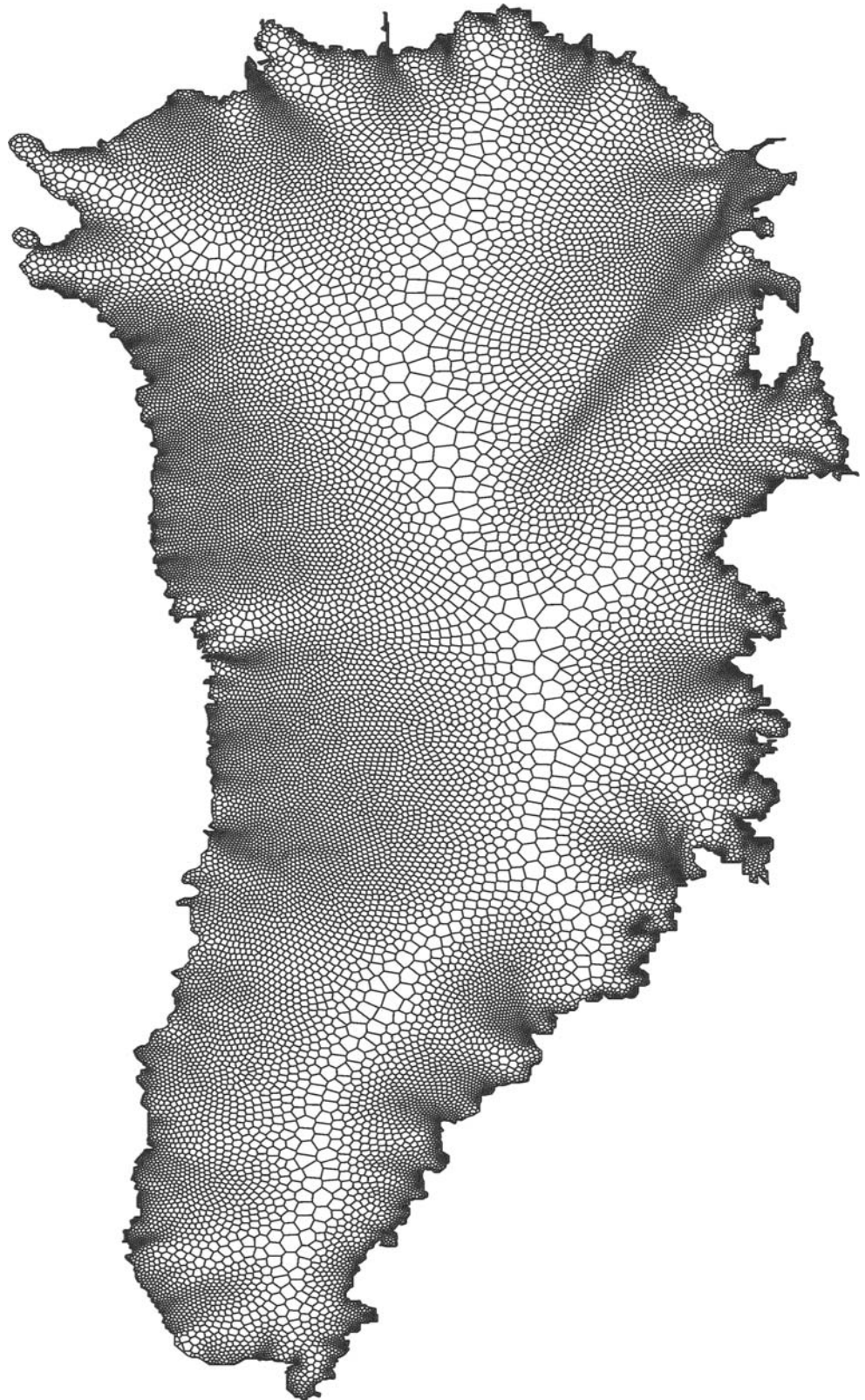


Fig. 4 Log base 10 distribution of Voronoi cell area of the Greenland ice sheet using 25,936 nodes. The scale is saturated for all cell areas above 250 km and for all cell area below 10 km

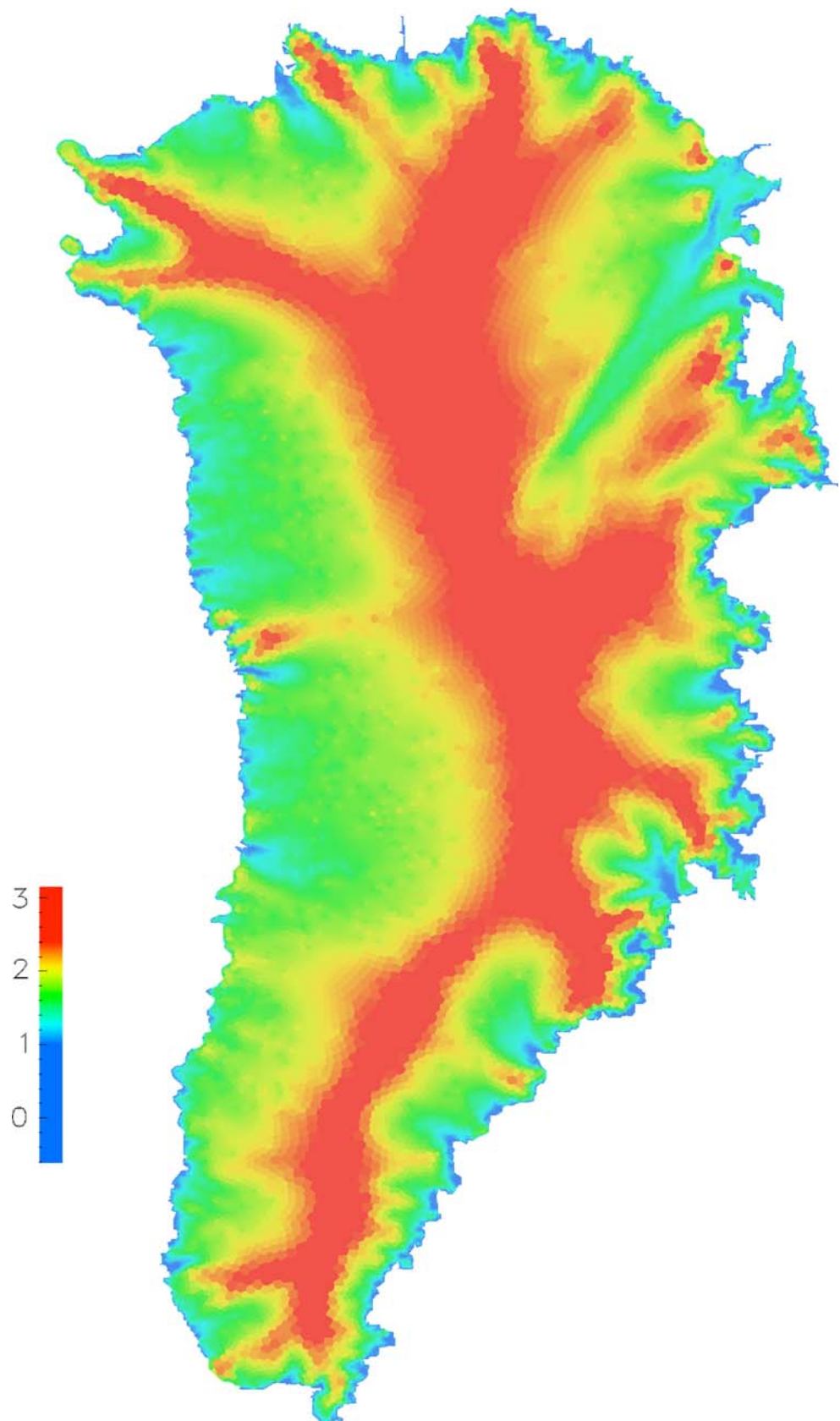


Fig. 5 Log base 10 distribution of Voronoi cell area of the Greenland ice sheet using 101,115 nodes. The scale is saturated for all cell areas above 65 km and for all cell area below 2.5 km

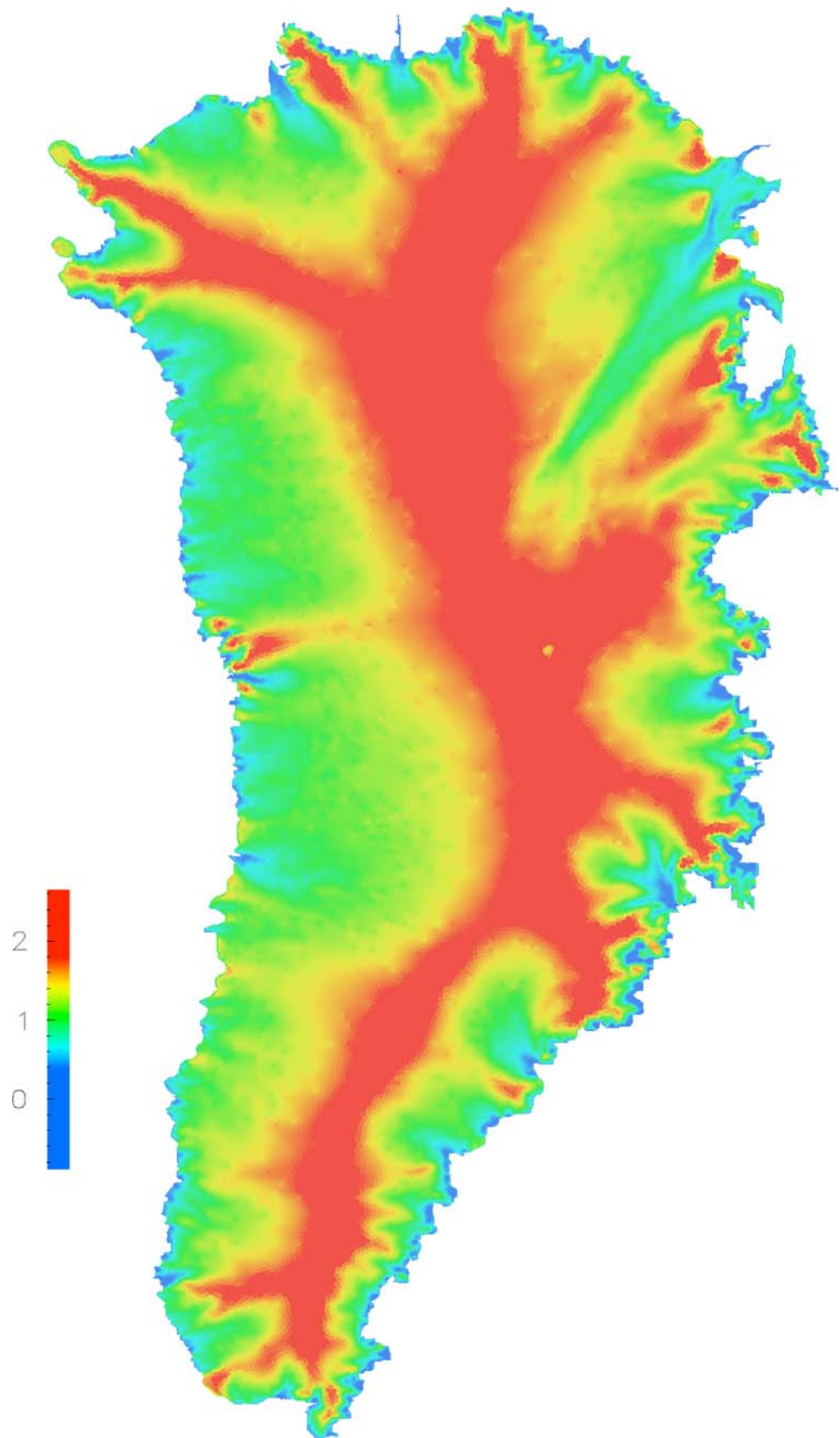


Table 1 Mesh information of SCVTs for the Greenland

Number of generators	σ_{\min}	σ_{avg}	h_{\max}/h_{\min}	Number of triangles	q_{\min}	q_{avg}
25,936	0.094	0.706	40.94	49,244	0.219	0.935
101,115	0.091	0.751	47.81	197,346	0.235	0.948

allocation of computation resources. In addition, the scientific study of many processes would surely benefit from the ability to support eddy activity in certain regions while maintaining a global ocean domain. Two supporting examples include the role of eddies on the meridional overturning circulation (Gnanadesikan 1999) and the role of eddies in mediating the ocean's response to changes in wind stress forcing (Hallberg and Gnanadesikan 2006).

3.2.2 Proxy for SCVT density

In this example, we will derive a variable-resolution mesh of the North Atlantic Ocean with sufficient resolution to resolve eddies within the major current systems. The domain is identical to the nominal 1/10 degree eddy-resolving simulation discussed in Smith et al. (2000). Using data from these simulations, we compute the time-mean kinetic energy of the surface

currents, as shown in Fig. 7. As can be seen in Fig. 7, the regions of intense eddy activity are extremely localized. Based on the kinetic energy, KE, we defined the density function as

$$x = \text{KE}/\text{KE}_{\max}, \quad x_{\min} = 0.1, \quad x_{\max} = 1.0 \quad (12)$$

$$x = \max(x, x_{\min}) \quad (13)$$

$$\rho = x^4 \quad (14)$$

where KE_{\max} is the maximum kinetic energy in the domain. The lower bound of 0.1 insures that a minimum resolution is maintained in the quiescent regions. The ratio $x_{\max}/x_{\min} = 10$ leads to a grid spacing that varies by approximately a factor of 10. In addition, we enhance the density function near the land–sea interface to insure that the boundary is adequately resolved. While we want the mesh to capture these regions of high activity via enhanced resolution, we also recognize the need to expand this region to allow eddies to travel

Fig. 6 Quality histograms of SCVTs of the Greenland ice sheet with 25,936 nodes (*top*) and 101,115 nodes (*bottom*). *Left*: Distribution of quality measurement of Voronoi cells σ ; *right*: distribution of quality measurement of Delaunay triangles q . Note the uniform shift toward higher-quality as the number of nodes is increased

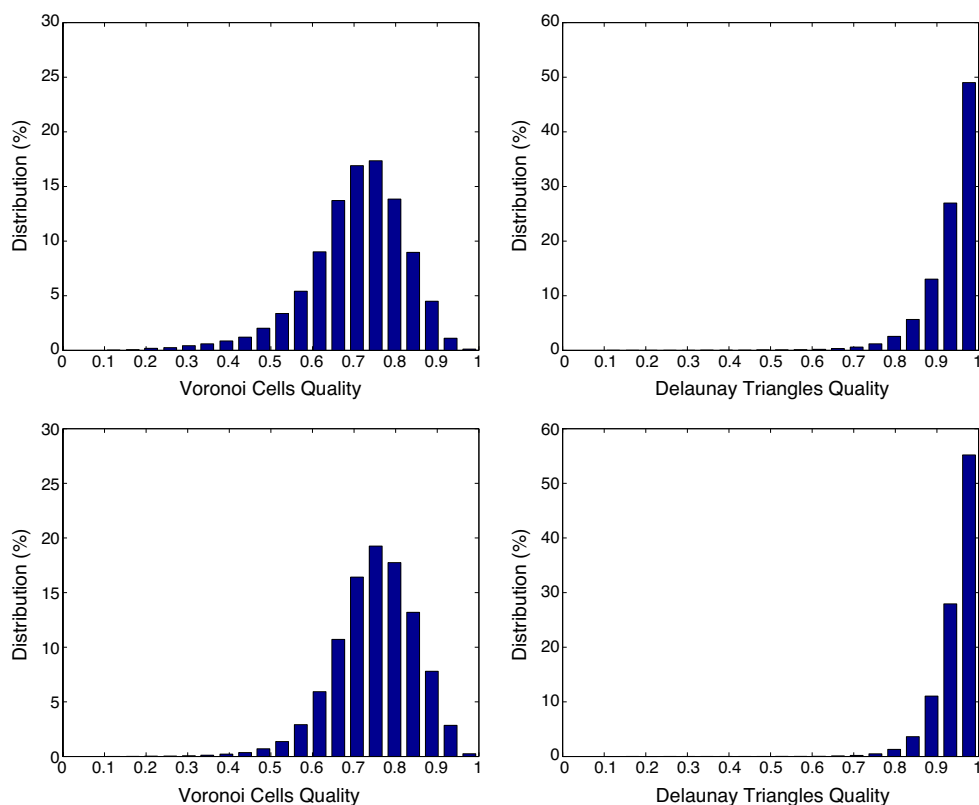
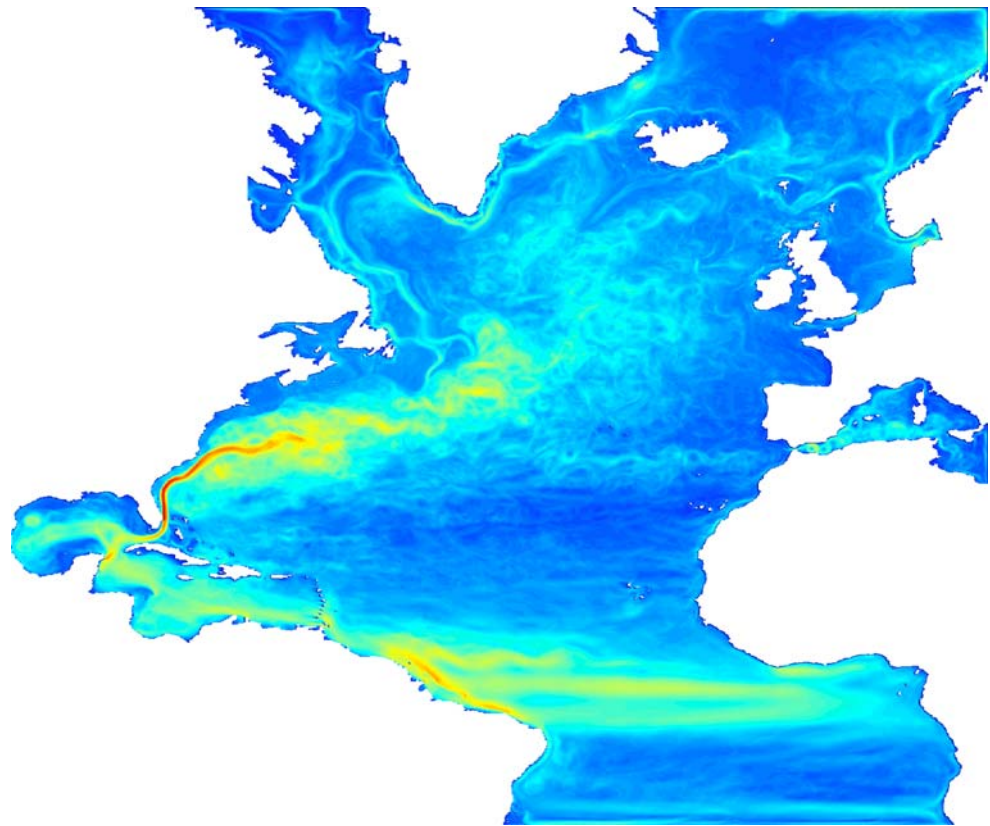


Fig. 7 Time mean kinetic energy from a global 0.1 degree simulation of the North Atlantic Ocean (Smith et al. 2000)



uninhibited by grid resolution. As such, we applied a substantial amount of Laplacian smoothing to our density function (approximately 20 passes) to expand and smooth the regions of enhanced resolution. (Note that the RMS of sea-surface height is also an accurate reflection of mesoscale ocean variability and we have developed global ocean SCVTs based on TOPEX remote sensing of sea-surface height.)

As with the Greenland example, we produced a continuous, piece-wise linear representation of the land–ocean boundary based on the land–sea mask used in the 0.1 degree simulation. This approach also identifies all islands. Islands with a circumference less than 10 km were discarded; the resulting domain contains 58 islands.

The Voronoi diagram shown in Fig. 8 uses 47,305 nodes. This results in a minimum grid resolution of approximately 20 km. We continue to add nodes into the domain until we reach a minimum resolution of 10 km. Closeups of this high-resolution mesh using 183,807 nodes and its low-resolution counterpart are shown in Fig. 9.

Table 2 presents the results of our North Atlantic SCVTs. The corresponding quality histograms are shown in Fig. 10. As with the Greenland example, the

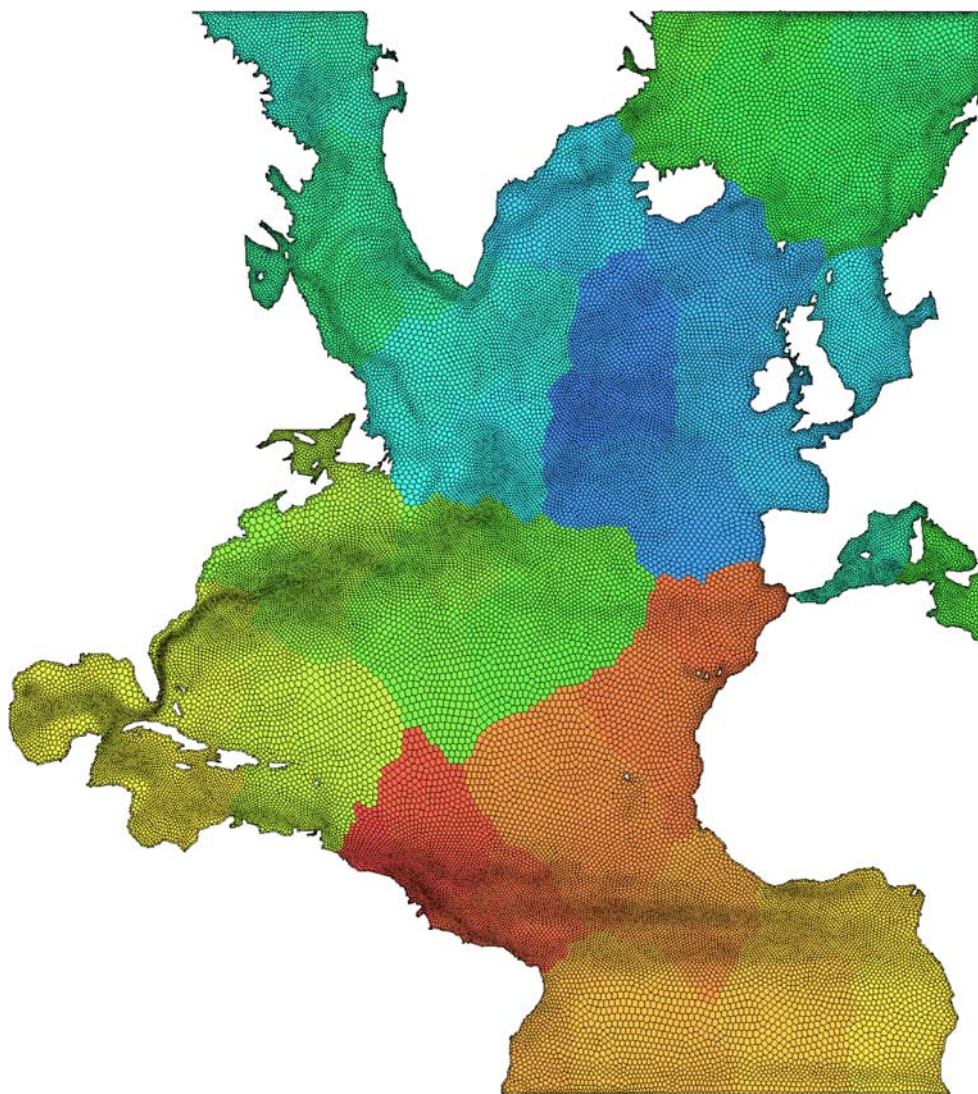
quality measures show a systematic improvement as we increase the degrees of freedom.

3.3 Ocean–ice shelf interaction

3.3.1 Motivation

Our final example couples ocean and ice domains in the context of ocean–ice shelf interaction. Ice shelves are ice flows that become ungrounded and buoyant and rest on top of ocean water. The location at which ice transitions from resting on bedrock to resting on ocean water is referred to as the grounding line. As the ice shelves are pushed outward into the ocean, the ice is either melted along the ice–ocean interface or calves from the main shelf into icebergs. These ice shelves provide a significant buttressing force that resists the flow of upstream, grounded ice. As evidenced by the Larsen B ice-shelf collapse, when the ice-shelf buttressing force is removed, the upstream ice flow can increase by several hundred percent (Rignot et al. 2004). The accurate simulation of ocean–ice shelf interaction is necessary in order to quantify the risk of rapid sea level rise (International Panel on Climate Change 2007).

Fig. 8 SCVT of North Atlantic ocean domain using 47,305 nodes. *Coloring* indicates the mesh decomposition for implementation on distributed memory systems: each *color* represents a separate computational unit



The West Antarctic Ice Sheet (WAIS) is particularly relevant to the study of ocean–ice shelf interaction. Not only are ice shelves the primary outlet of grounded ice, but the grounded ice frequently rests on bedrock that is increasingly below sea level as one moves toward the ice interior region (Schoof 2007). The physical geometry is such that a rapid erosion of the WAIS due to ocean–ice shelf interaction is a plausible scenario for the twenty-first century.

Ice shelves connected to WAIS have spatial extents of more than 1,000 km (e.g., the Ross Ice Shelf) down to less than 50 km (e.g., the Thwaites Ice Shelf). The embayments where this ice flows into the ocean have similar ranges in spatial scale. While the ice shelves are not as dynamically active as the ice streams that feed them, the structure and shape of the ice–ocean interface is a primary factor that drives mixing at this interface (Holland et al. 2008). Grid resolutions of less than 5 km are often used when simulating ocean–ice shelf

coupled dynamics. Furthermore, analysis of the global 1/10 degree ocean simulations in the vicinity of WAIS indicates that the transport of heat into these embayments may be eddy-driven and episodic (Maltrud, personal communication, 2007). Thus, resolving ocean eddies in and around these embayments will likely be required for robust simulations.

The horizontal discretization of this system is difficult because part of the domain will be ice (ice domain), part will be ocean (ocean domain), and part will be both ocean and ice (shelf domain). Furthermore, the characterization of a region as ice, ocean, or shelf will evolve over the time scales of decades to millennia. Due to this complexity and the fact that only limited work has been completed on modeling the coupled ocean–ice shelf system, we will explore techniques to discretize this system in an idealized setting. Figure 11 shows our idealized domain with a spatial extent of 1,100 by 550 km. The domain is characterized by a

Fig. 9 Close-up of Gulf Stream region using 47,305 nodes (*top*) and 183,907 nodes (*bottom*)

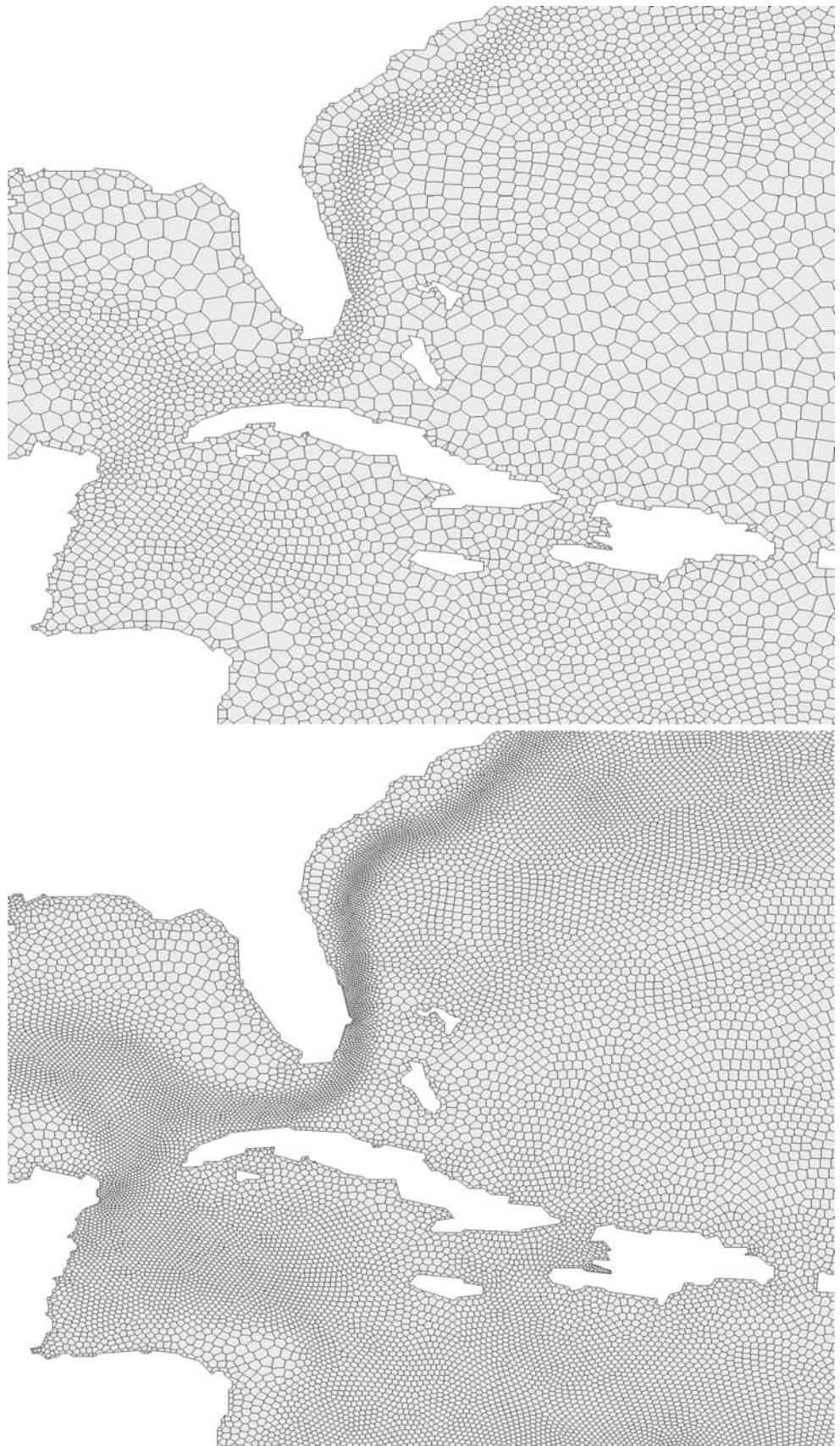


Table 2 Mesh information of SCVTs for the North Atlantic

Number of generators	σ_{\min}	σ_{avg}	h_{\max}/h_{\min}	Number of triangles	q_{\min}	q_{avg}
47,305	0.089	0.712	16.60	89,272	0.104	0.933
183,907	0.105	0.754	15.05	358,577	0.127	0.947

region of grounded ice (to the left), a region of ocean (to the right), and an ice-shelf region (center). The ice domain includes an ice stream that feeds the shelf region. As indicated in the figure, a robust simulation of this system will require enhanced resolution in the vicinity of the ice shelf, ice stream, and region of the ocean in proximity to the ice.

3.3.2 Proxy for SCVT density

In this idealized example, our intent is to produce a tessellation with a minimum grid spacing of 2 km and a maximum grid spacing of 20 km. As opposed to our other examples, we are not building our density function from a physical characteristic of the system. The generated density function has local maxima in the vicinity of the ice stream, in the region of the ice shelf, and along the entire ocean–ice boundary. The resulting Voronoi diagram of this system using 9,359 nodes is shown in Fig. 12 with a close-up of the shelf region at both high and low resolution shown in Fig. 13.

Table 3 presents the results on our SCVTs for the idealized ocean–ice sheet, and corresponding quality histograms are shown in Fig. 14. Yet again, the quality measure show increasing mesh quality with increasing degrees of freedom.

4 Example numerical method

The large majority of numerical methods utilized in IPCC-class climate models were developed in the context of uniform meshes. Successfully implementing these same numerical methods on nonuniform meshes, such as those developed above, will likely prove to be a difficult task, as discussed in St-Cyr et al. (2007). While emerging numerical methods based on spectral elements, discontinuous Galerkin, finite-element, or similar approaches are a more natural choice when considering the multiresolution meshes developed here, these alternative approaches are still relatively new to climate system modeling in comparison to low-order finite-volume methods. The sole purpose of this

Fig. 10 Quality histograms of SCVTs of the North Atlantic with 47,305 nodes (*top*) and 183,907 nodes (*bottom*). *Left*: Distribution of quality measurement of Voronoi cells σ ; *right*: distribution of quality measurement of Delaunay triangles q

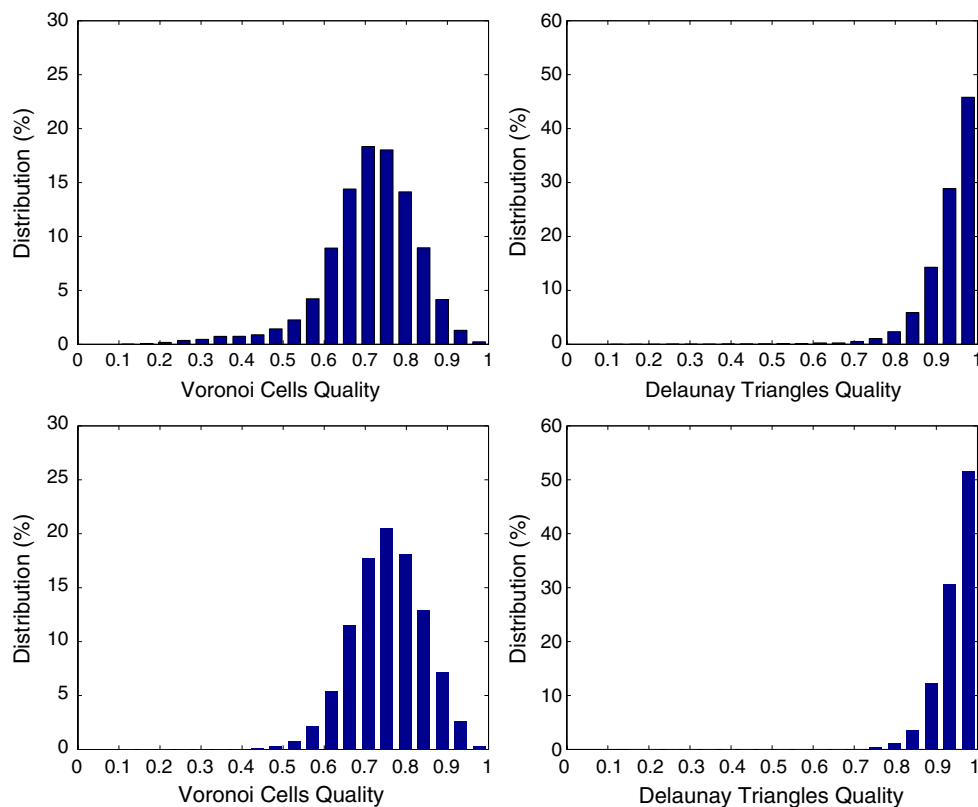


Fig. 11 An idealized ocean–ice shelf system. The ice domain (*left*) flows into the shelf region (*semicircle*) via an ice stream. Enhanced resolution in the vicinity of the ice stream, ice shelf, and ice margin will be required

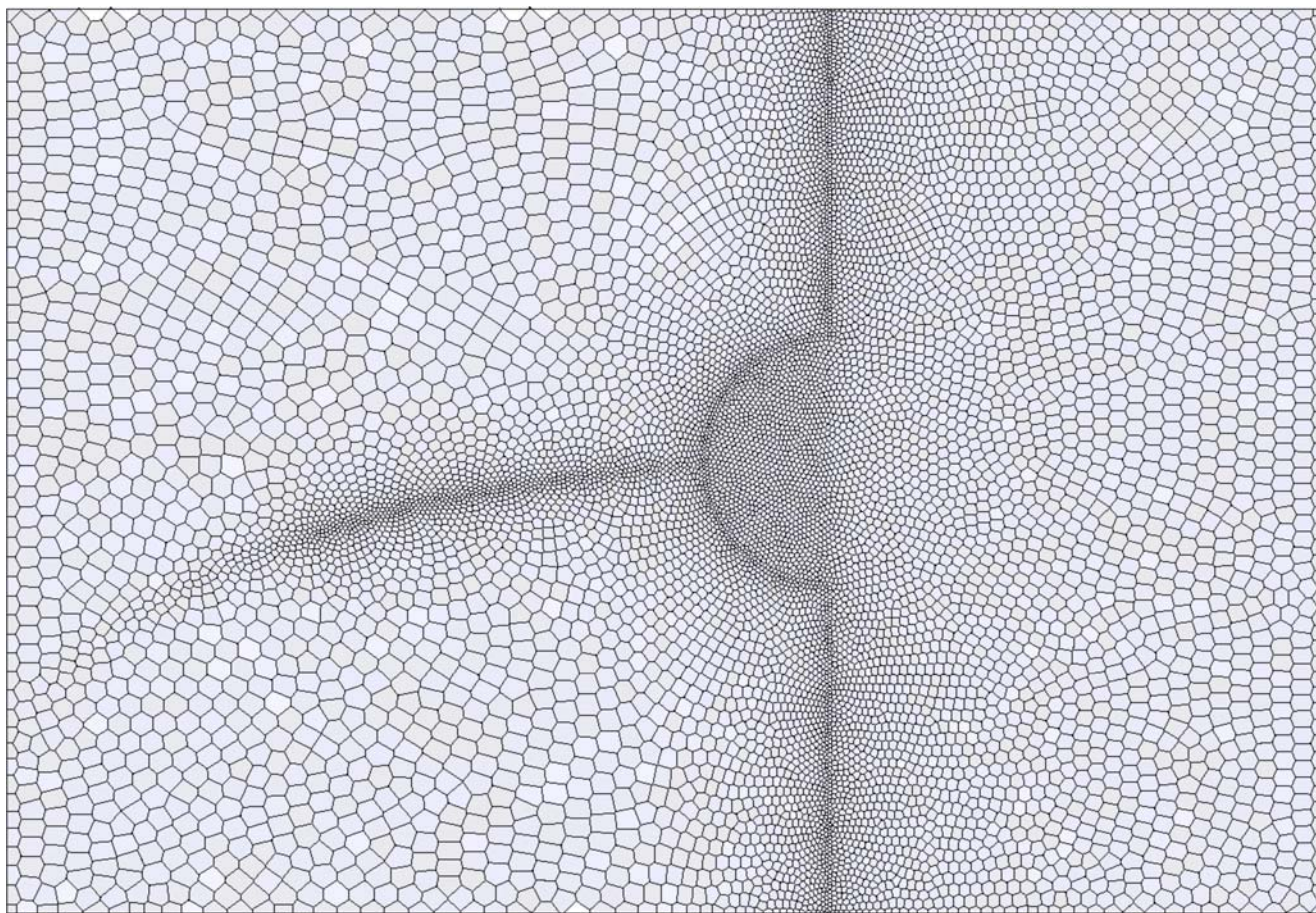
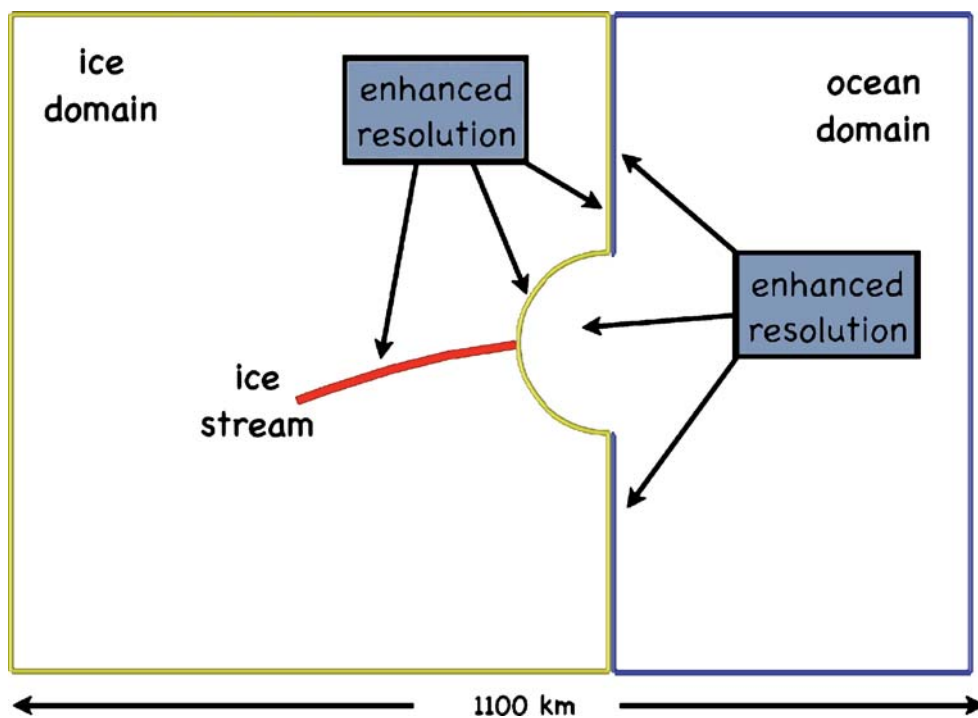


Fig. 12 SCVT of ocean–ice shelf system using 9,359 nodes. Note enhanced resolution in the vicinity of the ice stream, ice shelf, and ocean–ice interface

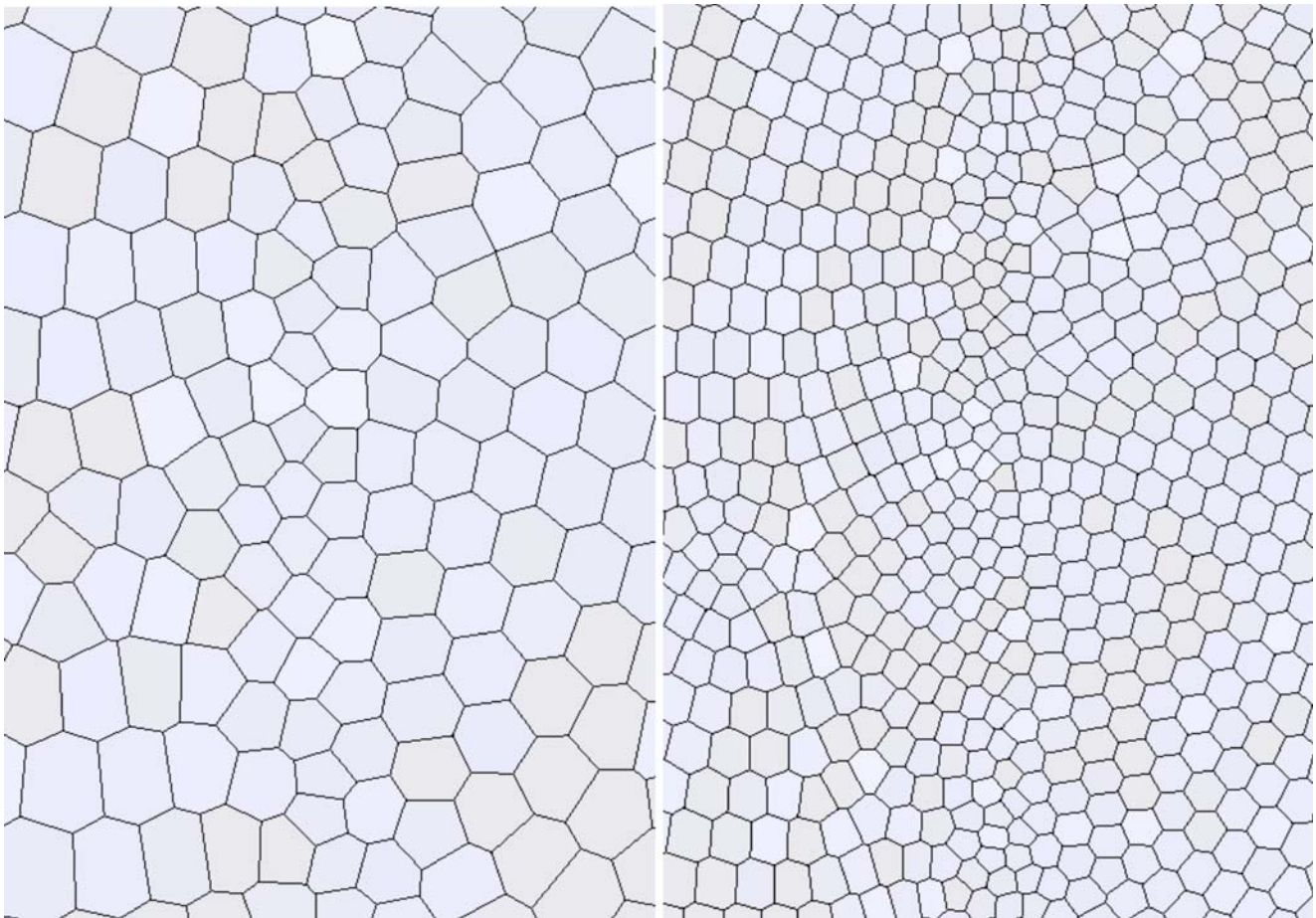


Fig. 13 Close-up of ocean-ice shelf SCVT in the vicinity where the ice stream enters the shelf region. *Left:* SCVT using 9,359 nodes. *Right:* SCVT using 37,157 nodes

section is to exhibit a low-order, finite-volume method capable of producing robust simulations when implemented on nonuniform SCVTs. The implication is that these meshes are immediately applicable to current-generation CSM components. With this purpose in mind, the discussion below is not intended to be exhaustive. In many respects, developing numerical methods that effectively utilize these nonuniform SCVTs is a much richer and more difficult problem than generating the mesh itself. While some efforts to exploit the local uniformity of SCVTs have already been completed (e.g., see Du and Ju 2005), much work remains. We have made significant progress regarding the formulation of finite-volume schemes suitable for implementation in these variable resolution meshes. These results will be detailed at a later time.

4.1 Continuous equations

For this demonstration, we choose the nonlinear, shallow-water equations spanning the entire surface of the sphere:

$$\frac{\partial h}{\partial t} + \nabla \cdot (h\mathbf{u}) = 0 \quad (15)$$

$$\frac{\partial \mathbf{u}}{\partial t} + (\omega + f)\mathbf{k} \times \mathbf{u} = -g\nabla(h + h_s) - \frac{1}{2}\nabla\|\mathbf{u}\|^2 \quad (16)$$

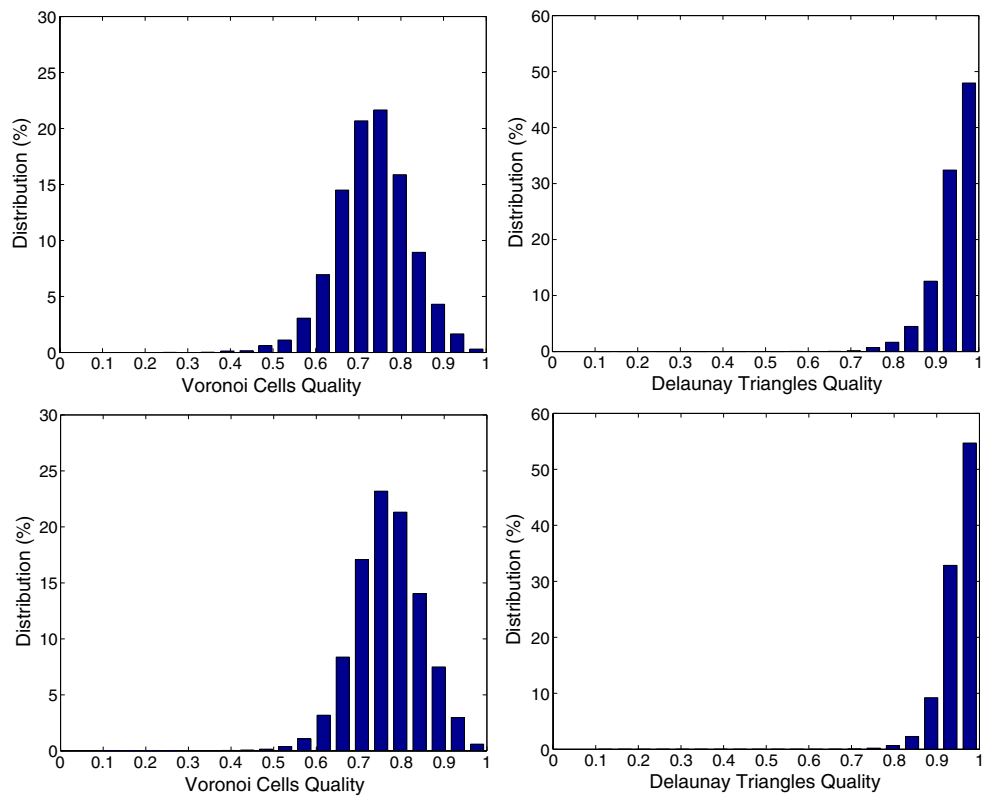
$$\omega = \mathbf{k} \cdot (\nabla \times \mathbf{u}) \quad (17)$$

where h is the fluid thickness, h_s is the height of the lower boundary, \mathbf{u} is the vector velocity orthogonal

Table 3 Mesh information of SCVTs for ocean-ice shelf system

Number of generators	σ_{\min}	σ_{avg}	h_{\max}/h_{\min}	Number of triangles	q_{\min}	q_{avg}
9,359	0.275	0.735	8.91	18,440	0.568	0.942
37,157	0.313	0.769	10.15	73,765	0.626	0.951

Fig. 14 Quality histograms of SCVTs of the idealized ocean–ice sheet with 9,359 nodes (top) and 37,157 nodes (bottom). Left: distribution of quality measurement of Voronoi cells σ ; right: distribution of quality measurement of Delaunay triangles q



to the local normal vector \hat{k} , and f is the Coriolis parameter. The component of relative vorticity in the plane normal to the surface of the sphere, ω , is defined in Eq. 17.

4.2 Discrete equations

We utilize the discrete method developed by Bonaventura and Ringler (2005). While the method developed in Bonaventura and Ringler (2005) is intended for use on multiresolution meshes, to our knowledge, this is the first demonstration. This method uses the Delaunay triangulation as the finite-volume cell for the thickness equation. The vorticity field is defined on the Voronoi diagram. Velocity components normal to the triangle edges are retained as prognostic equations. A schematic of this discretization is shown in Fig. 15. All quantities with overhats are derived fields, with \hat{T} representing the reconstructed tangent velocity required for the Coriolis force and $\hat{\eta}$ representing the absolute vorticity (see Bonaventura and Ringler 2005 for a full discussion). The discrete system is expressed as:

$$\frac{\partial h_i}{\partial t} = \frac{-1}{A_i} \sum_{j=1}^{nedges} \hat{h}_j N_j dl_j \tag{18}$$

$$\frac{\partial N_j}{\partial t} = \hat{\eta}_j \hat{T}_j - \left\{ \left[gh + gh_s + \hat{K} \right]_{iForward} - \left[gh + gh_s + \hat{K} \right]_{iBackward} \right\} / dc_j, \tag{19}$$

where the summation in Eq. 18 is over the edges of each triangle. In terms of solution error, the scheme is nominally second-order accurate in space using centered-in-space reconstructions and fourth-order accurate in time using fourth-order Runge–Kutta time-stepping (see, e.g., Bonaventura and Ringler 2005). The simulations utilize no limiters, filters, or explicit dissipation of any sort. The center-in-space numerics, along with the fourth order Runge–Kutta scheme, is used to minimize any implicit diffusion.

We demonstrate this method on the two SCVTs shown in Fig. 16. Each mesh contains 40,962 nodes. The solid black line indicates the boundary of an orographic feature that is the sole forcing of the simulation (see below). The first mesh (top) is generated with a uniform density function leading to an average grid spacing of 120 km. The second mesh (bottom) is generated with higher densities in the vicinity of the orographic feature. The density function is chosen such that the average grid spacing is three times smaller (40 km) in the vicinity of the mountain than compared to its quasiuniform counterpart. The solid colors indicate our

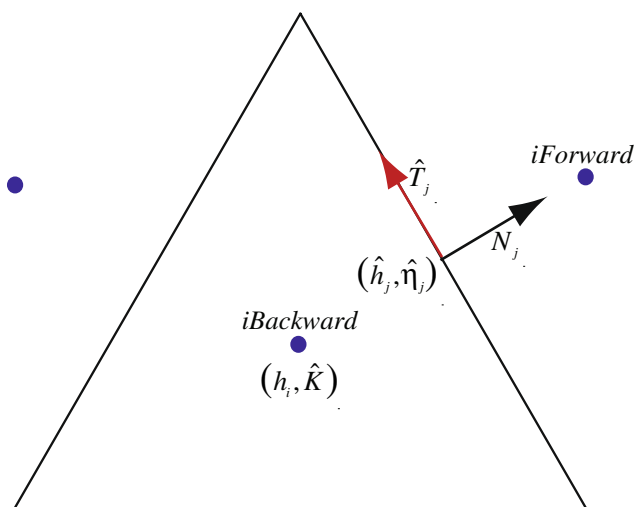


Fig. 15 A schematic of the finite-volume system. Thickness, h , and kinetic energy, \hat{K} , are defined at the center of the triangle. The normal component of velocity, N_j is defined at each cell edge. Vorticity, η , is defined at the triangle vertices. All quantities with overhats are derived fields, see Bonaventura and Ringler (2005) for details

domain-decomposition strategy for efficient implementation on distributed memory systems: each block represents a different computational processor.

4.3 Simulation

We apply this numerical method to one of the standard shallow-water test cases developed by Williamson et al. (1992) referred to as test case 5. In this test case, a flow in geostrophic balance is confronted with a large-scale orographic feature at the start of the simulation, $t = 0$. The transient forcing at $t = 0$ leads to the generation of large-amplitude gravity waves and Rossby waves. The sole forcing mechanism is the presence of the orographic forcing. While no analytical solution is known, results from high-resolution global spectral models are adequate reference solutions for the simulations conducted here (e.g., see Lipscomb and Ringler 2005).

Both simulations are stable over the course of the 15-day integration. The kinetic energy field for each simulation is shown in Fig. 17. Both simulations produce the same large-scale flow structure: an anticyclone dominates in the region of orography with a strong,

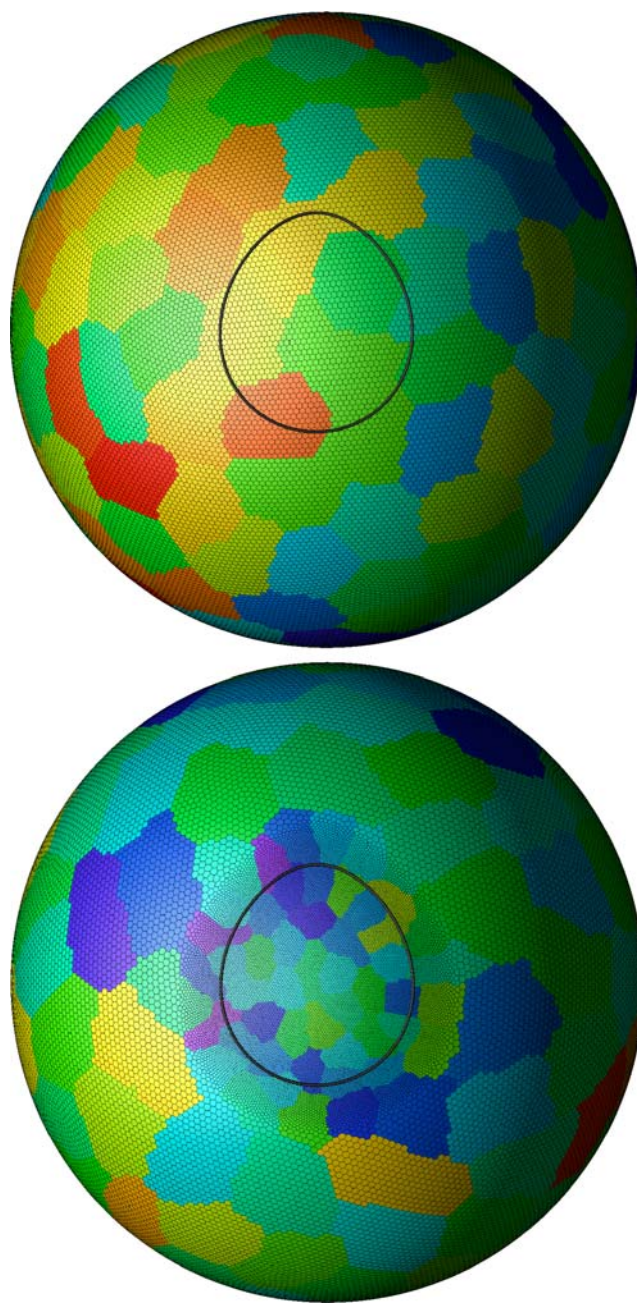


Fig. 16 *Top*: a SCVT using 40,962 nodes with a uniform density function. *Bottom*: a SCVT also using 40,962 nodes but using a nonuniform density function with high values of density occurring in the vicinity of the orography (shown by the solid black line). The variable-resolution mesh results in a minimum grid spacing of approximately 1/3 that found in the quasiuniform mesh. The colored background denotes groups of cells (blocks) that are distributed across multiple processors. The numerical method employed here defines vorticity on the SCVT nodes and mass on the dual Delaunay triangulation

stationary, low-pressure system residing immediately downstream. Both simulations produce velocities in excess of 40 m/s in the jet region.

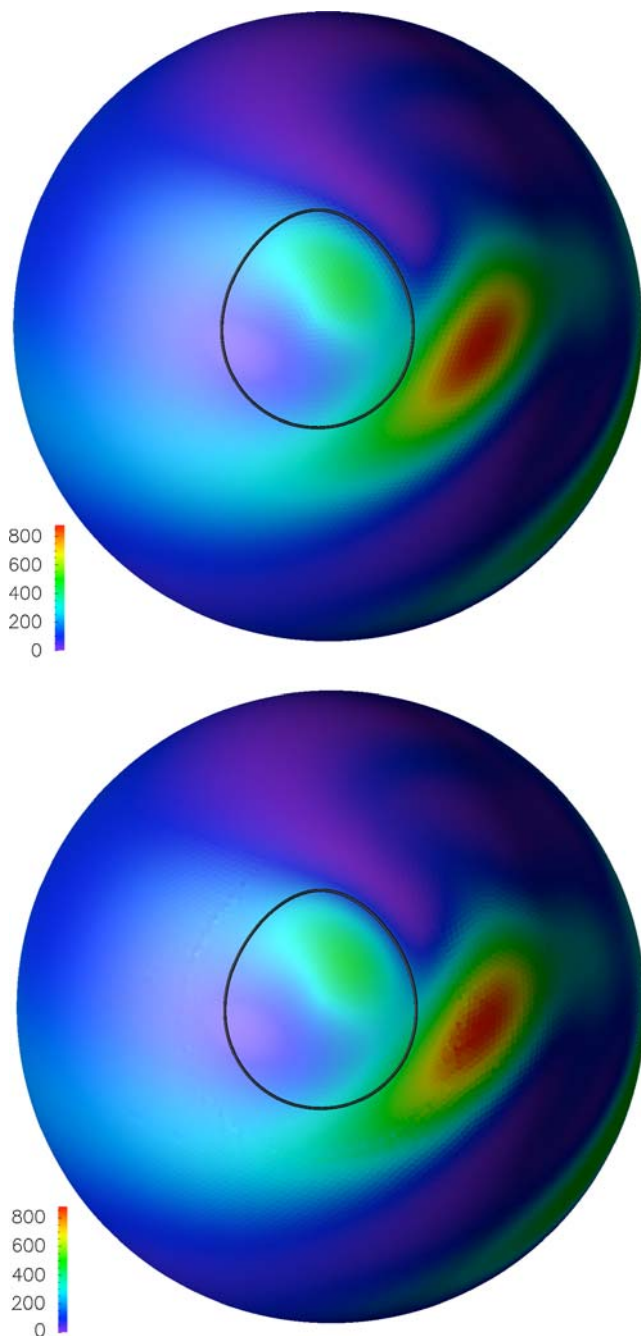


Fig. 17 Kinetic energy field at day 10 of simulation. *Top*: simulation using quasiuniform mesh. *Bottom*: simulation using variable-resolution mesh

Figure 18 shows how the error norms for each of these simulations evolve over the course of the simulation. The error is based on the deviation of the thickness field from high-resolution spectral results. The panel on the top shows the normalized L_2 -error following the procedure in Tomita et al. (2001), Eq. 17. The L_2 error norm is computed for two regions: a global domain and a local domain defined by $h_s > 0$

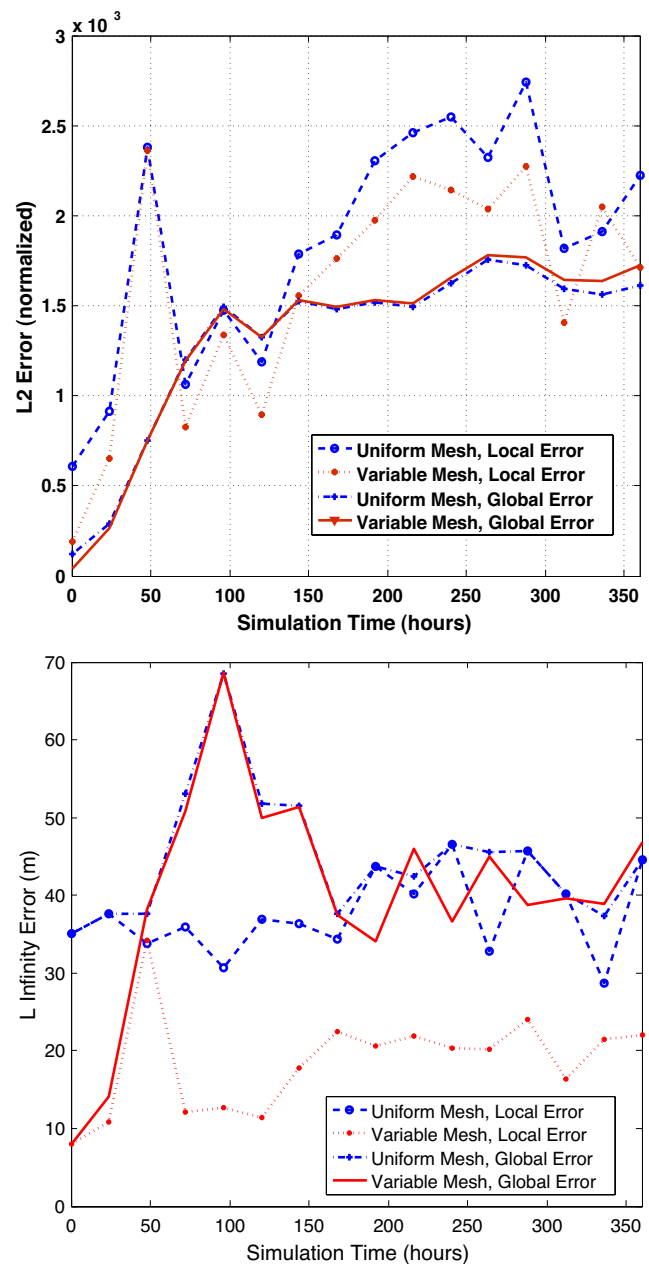


Fig. 18 *Top*: L_2 error norms for both the global domain and the local domain in the vicinity of the orography. *Bottom*: L_∞ norms for the same two domains. Each figure compares the errors produced in the uniform-mesh simulation to the errors produced in the variable-mesh simulation

that is coincident to the region of enhanced resolution. The panel on the bottom depicts the L_∞ -norm with the same layout. Since the L_2 error is normalized by the reference values, we only compare norms within the same averaging domain. When comparing the global L_2 error norms between the simulations, we find that the variable-resolution mesh provides marginal improvements only for times less than 24 h. At the early stages

of the simulation, the benefit of the variable-resolution mesh is due mostly to a better representation of the initial condition. For the remainder of the integration, the two simulations have nearly identical global L_2 error values. When we compare the local L_2 error values, we find a slightly different result; averaged over the duration of the simulation, the variable-resolution mesh reduces the error by approximately 20%. The L_∞ error values (bottom panel) exhibit a similar tendency. The variable-resolution mesh provides limited benefit in the context of global error reduction but does significantly reduce the errors in the vicinity of the orography. In this case, the variable-resolution mesh reduces the local L_∞ error norm by a factor of two as compared to the uniform mesh simulation. The implications of these findings on the merit of multiresolution CSMs will be discussed in the next section.

5 Discussion and conclusions

We have argued that the traditional paradigm of constructing IPCC-class climate models based on quasiuniform meshes will be strained in the coming decade by two mechanisms. First, each CSM component currently has one or more unresolved processes that may play an important role in the dynamics of the global climate system. These processes are either omitted altogether, exemplified by the omission of ice streams in ice sheet models, or highly parameterized, exemplified by the subgrid scale models of eddy activity in the ocean. The current and foreseeable computational resources preclude the notion of resolving these processes everywhere all of the time. Second, IPCC-class climate models will be pressed into the role of simulating regional climate change with the purpose of developing adaptation and mitigation strategies. The resolution and computational resources required for the robust simulation of regional climate change will force the climate modeling community to develop an alternative approach to complement the emerging suite of quasiuniform global CSMs.

One promising approach to meet these new challenges is based on the use of SCVTs. These tessellations (or meshes) offer many attractive qualities in the context of climate system modeling. First, since these meshes are a superset of the commonly used “icosahedral–hexagonal grids,” we can conceptually consider SCVTs to be an extension of meshes already in use today. Second, SCVTs allow for the spatial allocation of nodes in a straightforward, intuitive manner. SCVTs are generated with respect to a user-defined density function where nodes are “clustered” toward

regions of high density and away from regions of low density. Since each SCVT is associated with a Delaunay triangulation, this method is amenable to numerical methods situated on either the Voronoi diagram or the Delaunay triangulation. The proven SCVT properties related to smoothness and uniformity are conveyed to the Delaunay triangulation. If we understand a system well enough to know how to redistribute our degrees of freedom (and, hence, our computational resources), SCVTs offer an easy way to implement this redistribution. Finally, and most importantly, SCVTs are amenable to rigorous analysis from which we can make statements regarding the regularity of a given mesh and how that regularity will improve as we increase the nodes in a given domain.

We demonstrated the potential for this technique by developing example meshes for several different components of the Earth’s climate system: the Greenland ice sheet, the North Atlantic Ocean, and a generic Antarctica ice shelf–ocean interaction. Furthermore, our example numerical method developed a multiresolution mesh that is characteristic of local resolution enhancement in regional atmospheric modeling. In each of these examples, we exhibited the ability to precisely manipulate the regions of enhanced resolution through our choice of the SCVT density function. In two of the examples (Greenland and North Atlantic), the SCVT density function was developed directly from physical characteristics of the system. In the Greenland example, we used the observed ice velocity distribution to develop a SCVT density function that places increased resolution in and around ice streams. In the North Atlantic example, we constructed the SCVT density function in order to obtain meshes that are able to resolve eddy activity associated with the Gulf Stream and North Atlantic current. In each of these examples, we computed metrics that measure the quality of the mesh. In all cases, and in agreement with the theoretical underpinnings of SCVT, we found that increasing the degrees of freedom results in a uniform improvement in mesh quality. We found this consistent improvement in both the Voronoi diagrams, and in the Delaunay triangulations.

While the primary purpose of this work is to demonstrate the potential for SCVTs to produce high-quality, multiresolution meshes for climate system applications, we felt it important to also exhibit a traditional, finite-volume technique that can successfully exploit the benefits of a variable-resolution mesh. While our results in this regard are far from sufficient, we have at least produced one positive example in the context of the global shallow-water equations. Even this simple example has provided some guidance on what we should

and should not expect from multiresolution techniques, such as the one developed here. For instance, given the hyperbolic nature of many of the Earth's climate system components, it will be extremely difficult to reduce formal solution error over a wide range of conditions. Eventually, the error will become dominated by some phenomena (transient or otherwise) occurring in regions of low resolution. While this problem occurs regardless of the numerical method employed, it will likely be particularly evident when using the low-order, finite-volume methods that are ubiquitous in CSMs today. In contrast to reducing formal solution error, our emphasis will be on the formulation of robust numerical methods that produce stable, long-term simulations over a wide class of phenomena without the need for ad hoc filtering or dissipation. The driving purpose for developing multiresolution climate system components will be for the simulation of *new phenomena* requiring enhanced resolution, not necessarily for the formal reduction in solution error.

While this work has demonstrated the ability to generate high-quality meshes for a wide class of problems, the daunting challenge going forward is to develop numerical techniques that can effectively exploit these high-quality, multiresolution meshes.

Acknowledgements This work was supported by the DOE Office of Science Climate Change Prediction Program through DE-FG02-07ER64431, DE-FG02-07ER64432, and DOE 07SCPF152. The authors would like to thank Dr. Sebastien Legrand and two anonymous reviewers for their constructive comments.

References

- Adcroft A, Campin J, Hill C, Marshall J (2004) Implementation of an atmosphere-ocean general circulation model on the expanded spherical cube. *Mon Weather Rev* 132:2845–2863
- Bamber JL, Hardy RJ, Joughin I (2000) An analysis of balance velocities over the Greenland ice sheet and comparison with synthetic aperture radar interferometry. *J Glaciol* 46:67–74
- Bell RE (2008) The role of subglacial water in ice-sheet mass balance. *Nat Geosci* 1:297–304
- Bonaventura L, Ringler T (2005) Analysis of discrete shallow-water models on geodesic Delaunay grids with c-type staggering. *Mon Weather Rev* 133:2351–2373
- Collins WD, Bitz CM, Blackmon ML, Bonan GB, Bretherton CS, Carton JA, Chang P, Doney SC, Hack JJ, Henderson TB, Kiehl JT, Large WG, McKenna DS, Santer BD, Smith RD (2006) The community climate system model version 3 (CCSM3). *J Climate* 19:2122–2143
- Comblen R, Legrand S, Deleersnijder E, Legat V (2008) A finite element method for solving the shallow water equations on the sphere. *Ocean Model*. doi:10.1016/j.ocemod.2008.05.004
- Du Q, Ju L (2005) Finite volume methods on spheres and spherical centroidal Voronoi meshes. *SIAM J Numer Anal* 43:1673–1692
- Du Q, Wang D (2003) Tetrahedral mesh generation and optimization based on centroidal Voronoi tessellations. *Int J Numer Methods Eng* 56:1355–1373
- Du Q, Wang D (2005) Anisotropic centroidal Voronoi tessellations and their applications. *SIAM J Sci Comput* 26:737–761
- Du Q, Faber V, Gunzburger M (1999) Centroidal Voronoi tessellations: applications and algorithms. *SIAM Rev* 41:637–676
- Du Q, Gunzburger M, Ju L (2003a) Constrained centroidal Voronoi tessellations on general surfaces. *SIAM J Sci Comput* 24:1488–1506
- Du Q, Gunzburger M, Ju L (2003b) Voronoi-based finite volume methods, optimal Voronoi meshes and PDEs on the sphere. *Comput Methods Appl Mech Eng* 192:3933–3957
- Du Q, Wang D, Huang Z (2005) Mesh and solver coadaptation in finite element methods for anisotropic problems. *Numer Methods Partial Differ Equ* 21:859–874
- Field D (2000) Quantitative measures for initial meshes. *Int J Numer Methods Eng* 47:887–906
- Gnanadesikan A (1999) A simple predictive model for the structure of the oceanic pycnocline. *Science* 283:2077–2079
- Gersho A, Gray R (1992) Vector quantization and signal compression. Kluwer, Boston
- Giraldo F, Warburton T (2008) A high-order triangular discontinuous Galerkin oceanic shallow water model. *Int J Numer Methods Fluids* 56:899–925
- Hallberg R, Gnanadesikan A (2006) The role of eddies in determining the structure and response of the wind-driven southern hemisphere overturning: results from the modeling eddies in the southern ocean (MESO) project. *J Phys Oceanogr* 36:2232–2252
- Holland PR, Jenkins A, Holland DM (2008) The response of ice shelf basal melting to variation in ocean temperature. *J Climate* 21:2558–2572
- International Panel on Climate Change (2007) Climate change 2007: the scientific basis. International Panel on Climate Change, Valencia
- Joughin I, Gray L, Bindschadler R, Price S, Morse D, Hulbe C, Mattar K, Werner C (1999) Tributaries of West Antarctic ice streams revealed by RADARSAT interferometry. *Science* 286:283–286
- Ju L, Gunzburger M, Zhao W-D (2006) Adaptive finite element methods for elliptic PDEs based on conforming centroidal Voronoi Delaunay triangulations. *SIAM J Sci Comput* 28:2023–2053
- Lipscomb WH, Ringler TD (2005) An incremental remapping transport scheme on a spherical geodesic grid. *Mon Weather Rev* 133:2335–2350
- Lloyd S (1982) Least squares quantization in PCM. *IEEE Trans Inf Theory* 28:129–137
- MacQueen J (1967) Some methods for classification and analysis of multivariate observations. In: Proc. fifth Berkeley symposium on mathematical statistics and probability, vol I. University of California, Berkeley, pp 281–297
- McGregor JL (1996) Semi-Lagrangian advection on a conformal cubic grid. *Mon Weather Rev* 124:1311–1322
- Maltrud ME, McClean JL (2004) An eddy-resolving global 1/10 degree ocean simulation. *Ocean Model* 8:31–54
- Nair R, Thomas S, Loft R (2005) A discontinuous Galerkin global shallow water model. *Mon Weather Rev* 133:867–888
- Newman BD, Wilcox BP, Archer S, Breshears DD, Dahm CN, Duffy CJ, McDowell NG, Phillips FM, Scanlon BR, Vivoni ER (2006) The ecohdrology of arid and semiarid environments: a scientific vision. *Water Resour Res* 42:W06302. doi:10.1029/2005WR004141

- Okabe A, Boots B, Sugihara K, Chiu S (2000) Spatial tessellations: concepts and applications of Voronoi diagrams, 2nd edn. Wiley, Chichester
- Randall DA, Ringer TD, Heikes RP, Jones P, Baumgardner J (2002) Climate modeling with spherical geodesic grids. *Comput Sci Eng* 4:32–41
- Renka R (1999) ALGORITHM 772. STRIPACK: Delaunay triangulation and Voronoi diagrams on the surface of a sphere. *ACM Trans Math Softw* 23:416–434
- Rignot E, Bamber JL, van den Broeke MR, Davis C, Li Y, van de Berg WJ, van Meijgaard E (2008) Recent Antarctic ice mass loss from radar interferometry and regional climate modelling. *Nat Geosci* 1:106–110
- Rignot E, Casassa G, Gogineni P, Krabill W, Rivera A, Thomas R (2004) Accelerated ice discharge from the Antarctic Peninsula following the collapse of Larsen B ice shelf. *Geophys Res Lett* 31(18):L18401. doi:[10.1029/2004GL020697](https://doi.org/10.1029/2004GL020697)
- Ringer TD, Heikes RP, Randall DA (2000) Modeling the atmospheric general circulation using a spherical geodesic grid: a new class of dynamical cores. *Mon Weather Rev* 128:2471–2490
- Satoh M, Matsuno T, Tomita H, Miura H, Nasuno T, Iga S (2008) Nonhydrostatic icosahedral atmospheric model (NICAM) for global cloud resolving simulations. *J Comput Phys* 227:3486–3514
- Schoof C (2007) Ice sheet grounding line dynamics: steady states, stability, and hysteresis. *J Geophys Res* 112(F3):F03S28
- Smith RD, Maltrud ME, Bryan FO, Hecht MW (2000) Numerical simulation of the north atlantic ocean at 1/10. *J Phys Oceanogr* 30:1532–1561
- St-Cyr A, Jablonowski C, Dennis JM, Tufo HM, Thomas SJ (2007) A comparison of two shallow water models with non-conforming adaptive grids. *Mon Weather Rev* 136:1898–1922
- Stuhne G, Peltier W (2006) A robust unstructured grid discretization for 3-dimensional hydrostatic flows in spherical geometry: a new numerical structure for ocean general circulation modeling. *J Comput Phys* 213:704–729
- Tomita H, Tsugawa M, Satoh M, Goto K (2001) Shallow water model on a modified icosahedral geodesic grid by using spring dynamics. *J Comput Phys* 174:579–613
- Tomita H, Miura H, Iga S, Nasuno T, Satoh M (2007) A global cloud-resolving simulation: preliminary results from an aqua planet experiment. *Geophys Res Lett* 32:L08805. doi:[10.1029/2005GL022459](https://doi.org/10.1029/2005GL022459)
- Williamson DL, Drake JB, Hack J, Jacob R, Swartztrauber PN (1992) A standard test for numerical approximation to the shallow water equations in spherical geometry. *J Comput Phys* 102:211–224