BOREAL ENVIRONMENT RESEARCH 15: 453–479
ISSN 1239-6095 (print) ISSN 1797-2469 (online)

Helsinki 29 October 2010

Validation of three-dimensional hydrodynamic models of the Gulf of Finland

Kai Myrberg[1)]*, Vladimir Ryabchenko[2)], Alexei Isaev[2)], Roman Vankevich[3)], Oleg Andrejev[1)], Jørgen Bendtsen[4)], Anders Erichsen[5)], Lennart Funkquist[6)], Arto Inkala[7)], Ivan Neelov[2)], Kai Rasmus[8)], Miguel Rodriguez Medina[9)], Urmas Raudsepp[10)], Jelena Passenko[10)], Johan Söderkvist[4)], Alexander Sokolov[9)], Harri Kuosa[11)], Thomas R. Anderson[12)], Andreas Lehmann[13)] and Morten D. Skogen[14)]

[1)] *Finnish Environment Institute, Marine Research Centre, P.O. Box 140, FI-00251 Helsinki, Finland (*corresponding author's e-mail: kai.myrberg@ymparisto.fi)*
[2)] *St. Petersburg Department of the P. P. Shirshov Institute of Oceanology, Russian Academy of Sciences, 30, 1 Liniya, Vasilievskiy Ostrov, 199053 St. Petersburg, Russia*
[3)] *Russian State Hydrometeorological University, 98 Malookhtinsky pr., 195196 St. Petersburg, Russia*
[4)] *National Environmental Research Institute, University of Aarhus, P.O. Box 358, DK-4000 Roskilde, Denmark*
[5)] *DHI, Agern Allé 5, DK-2970 Hørsholm, Denmark*
[6)] *Swedish Meteorological and Hydrological Institute, Folkborgsvägen 1, SE-601 76 Norrköping, Sweden*
[7)] *Environmental Impact Assessment Centre of Finland, Tekniikantie 21 b, FI-02150 Espoo, Finland*
[8)] *Finnish Environment Institute, P.O. Box 35, FI-40014 University of Jyväskylä, Finland*
[9)] *Baltic Nest Institute, Stockholm Resilience Centre, Stockholm University, SE-106 91 Stockholm, Sweden*
[10)] *Marine Systems Institute at Tallinn University of Technology, Akadeemia Tee 21, EE-12618, Tallinn, Estonia*
[11)] *Tvärminne Zoological Station, University of Helsinki, J. A. Palménin tie 260, FI-10900 Hanko, Finland*
[12)] *National Oceanography Centre, University of Southampton, Waterfront Campus, European Way, Southampton SO14 3ZH, United Kingdom*
[13)] *Leibniz Institute of Marine Sciences at Kiel University (IFM-GEOMAR), Düsternbrooker Weg 20, D-24105 Kiel, Germany*
[14)] *Institute of Marine Research, Box 1870 Nordnes, N-5817 Bergen, Norway*

Received 21 Jan. 2009, accepted 7 Nov. 2009 (Editor in charge of this article: Harri Koivusalo)

Myrberg, K., Ryabchenko, V., Isaev, A., Vankevich, R., Andrejev, O., Bendtsen, J., Erichsen, A., Funkquist, L., Inkala, A., Neelov, I., Rasmus, K., Rodriguez Medina, M., Raudsepp, U., Passenko, J., Söderkvist, J., Sokolov, A., Kuosa, H., Anderson, T. R., Lehmann, A. & Skogen, M. D. 2010: Validation of three-dimensional hydrodynamic models of the Gulf of Finland. *Boreal Env. Res.* 15: 453–479.

A model-intercomparison study was conducted, the first of its kind for the Baltic Sea, whose aim was to systematically simulate the basic three-dimensional hydrographic properties of a realistic, complex basin. Simulations of the hydrographic features of the Gulf of Finland for the summer–autumn of 1996 by six three-dimensional hydrodynamic models were compared. Validation was undertaken using more than 300 vertical hydrographic profiles of salinity and temperature. The analysis of model performance, including averaging of the ensemble results, was undertaken with a view to assessing the potential suitability of the models in reproducing the physics of the Baltic Sea accurately enough to

serve as a basis for accurate simulations of biogeochemistry once ecosystem models are incorporated. The performance of the models was generally satisfactory. Nevertheless, all the models had some difficulties in correctly simulating vertical profiles of temperature and salinity, and hence mixed layer dynamics, particularly in the eastern Gulf of Finland. Results emphasized the need for high resolution in both vertical and horizontal directions in order to resolve the complex dynamics and bathymetry of the Baltic Sea. Future work needs to consider the choice of mixing and advection schemes, moving to higher resolution, high-frequency forcing, and the accurate representation of river discharges and boundary conditions.

Introduction

The Baltic Sea is the second largest brackish water mass in the world, with a total area of about 390 000 km^2. It is also very shallow, with a mean depth of only 54 m. Water exchange with the North Sea is restricted due to the shallow and narrow Danish Straits that lie in between the two seas. Permanent stratification is a further notable feature, where lighter and fresher waters overly the saltier waters below (e.g., *see* Leppäranta and Myrberg 2009). This occurs both due to salty waters entering the southwestern Baltic from the North Sea in the near-bottom layer, and also because of a freshwater surplus in the northeast due to river runoff.

An accurate and reliable knowledge of the complex hydrodynamics in the Baltic Sea is important not only for modeling physical processes, but is also a necessary prerequisite for the reliable estimation of nutrient cycling and biological processes. This is especially so in the Baltic which is characterized by the occurrence of harmful algae blooms, zones of anoxia, etc., which are linked closely to upwelling and stratification conditions (e.g., *see* Wulff *et al.* 2001). The vertical stratification in the Baltic Sea is unusual (the thermocline and halocline are usually separated) with a pronounced and relatively stable halocline, whereas the temperature stratification exhibits a marked seasonality. Stagnation in the bottom layer frequently leads to anoxia and sedimentary release of phosphorus to the water column. The large freshwater input from rivers significantly affects the stratification as well as provides input of nutrients.

Numerical modeling of the Baltic Sea using three-dimensional simulations was carried out in order to investigate the physical circulation of the region (e.g., Lehmann and Hinrichsen 2002, Lehmann *et al.* 2002, Meier 2003, Omstedt *et al.* 2004). Nevertheless, many issues remain unresolved, in particular the parameterization of vertical mixing. Even if mixing is described using sophisticated k-ε turbulence models then, when coupled with three-dimensional models (Meier 2001), several aspects are still poorly represented, notably the breaking of surface waves, air bubbles, Langmuir circulation, internal waves and wind conditions over the open sea (Omstedt *et al.* 2004).

Simulation of so-called Major Baltic Inflows, particularly the one in 1993, has been a focus of various modeling studies. Lehmann (1995), for example, used a model with a horizontal resolution of 5 km and 21 vertical levels, along with realistic wind forcing. A realistic distribution of salinity was obtained, although the depth of the mixed layer was underestimated leading to the vertical gradient of the salinity across the halocline being too weak. A good agreement was however found by Meier (1996) in his model between simulated and observed water volume and salt transport, emphasizing the importance of the Drogden Sill in major inflows. One problem that remains an issue is the parameterization of slope convection in models where a z-coordinate system is used in the vertical (e.g., Beckmann and Döscher 1997). The horizontal resolution in the Baltic Sea models should necessarily be very high due to the small internal Rossby-radius of deformation (Fennel *et al.* 1991, Alenius *et al.* 2003), being between 3 and 10 km and even smaller in the Gulf of Finland (hereafter GoF).

High resolution has been achieved in models with the help of nested-grid approaches, especially when modeling the GoF (*see* e.g. Andrejev

et al. 2004a, 2004b, Neelov *et al*. 2003, Tamsalu *et al*. 2003, Korpinen *et al*. 2004, Soomere *et al*. 2004). The issue of resolution applies also to the meteorological forcing fields. Recently, the HIRLAM model (High Resolution Limited Area Model), achieved a horizontal resolution of 9 km in operational mode, enabling a realistic description of local meteorological phenomena. The high resolution meteorological forcing was, however, not used in this case, because no such data were available for the year 1996 during which our extensive oceanographic data set was collected.

We focus our study on the GoF (GoF), which is an elongated estuary in the northeastern Baltic with a mean depth of 37 m. The western part of the Gulf may be considered a part of the Baltic Proper, whereas the eastern end receives the largest single fresh water inflow to the Baltic Sea, the Neva River. The result is a strong east-west gradient in salinity. The vertical distribution of salinity is variable in both space and time, whereas temperature is largely controlled by the seasonal variability of the incoming solar radiation. Buoyancy-driven currents thus play an important role in the circulation, together with the wind-driven circulation and that induced by the sea-level gradient (*see* Alenius *et al*. 1998, Soomere *et al*. 2008). The salinity increases from east to west and from north to south. Sea-surface salinity decreases from 5‰ to 6.5‰ in the western GoF to about 0‰–3‰ in the easternmost part of the Gulf where the role of the Neva River is most pronounced (Alenius *et al*. 1998). In the western GoF, a quasi-permanent halocline is located at a depth of 60–80 m. Salinity in that area can reach values as high as 8‰–10‰ near the sea bed due to the advection of saltier water masses from the Baltic Proper. The bottom salinity also shows significant spatio-temporal variability due to irregular saline water intrusions from the Baltic Proper, as well as from changes in river runoff and the precipitation-evaporation balance. There is no permanent halocline in the eastern GoF, where salinity increases approximately linearly with depth (Nekrasov and Lebedeva 2002, Nekrasov *et al*. 2003). The seasonal cycle of the sea-surface temperature (SST) is also pronounced due to large variations in solar radiation (*see* e.g. Hankimo 1964), with large horizontal gradients occasionally being seen due to local upwelling. A seasonal thermocline starts to develop in May. The surface mixed layer reaches a maximum depth of 15–20 m by mid-summer and erosion of the thermocline starts in late August due to wind mixing and thermal convection.

A model inter-comparison study is a useful means of investigating the ability of different models to reliably reproduce the hydrodynamic fields of the Baltic Sea, and in particular the GoF. It provides a structured approach (e.g. common forcing, initial and boundary conditions) that highlights the strength and weaknesses of modeling the hydrodynamics of the Baltic Sea.

Several inter-comparison studies were carried out in areas other than the Baltic Sea. Such a study was for example carried out for models of the southern North Sea, NOMADS2 (North Sea Model Advection Dispersion Study-2, e.g., *see* www.pol.ac.uk/coin/nomads2/map.html). There is a limited number of previous inter-comparisons of Baltic Sea models. In the GoF, Inkala and Myrberg (2002) compared two hydrodynamic models (the EIA-model and FinEst-model). Recently, Passenko *et al*. (2008) compared two versions of the MIKE3 model (hydrostatic and non-hydrostatic) in the GoF and in the Gotland Sea.

The present study provides a comprehensive comparison of the leading contemporary scientific and engineering models of the Baltic Sea, especially those developed for the primary area of interest, the GoF. As far as we know, this model inter-comparison study is the first one for the Baltic Sea. Here an attempt is made to systematically restore the basic three-dimensional hydrographic properties (including their spatio-temporal variability) of a realistic, extremely complex basin. The emphasis was on studying the present level of accuracy of the available models in reproducing the properties of the marine environment in a transparent and fair way, thus providing an assessment of the current state-of-the-art of hydrodynamic modeling of the GoF. The six different models are compared with each other and to data. Additionally, an ensemble average of the model results is constructed and compared with data, and a skill assessment provided for the models.

The presentation to some extent follows the course of the EUTROPHICATION-MAPS-project. It was focused on two main tasks: firstly inter-comparison of hydrodynamic models, and secondly scenario simulations of cost-effective nutrient load reductions to combat eutrophication by using coupled hydrodynamic-ecological models. Here, we present the results of an inter-comparison of an ensemble of numerical three-dimensional hydrodynamic models (HIROMB, OAAS, SPBM, EIA, COHERENS and MIKE3, *see* below). In particular, we focus on the GoF in 1996. This year was "The Year of the Gulf of Finland", carried out under the umbrella of the trilateral co-operation between Finland, Estonia and Russia (Sarkkula 1997). In this framework, a large number of measurements were carried out in a collaborative effort of these countries. This detailed dataset thus motivated the choice of the GoF as the main study area, providing validation data for the models. Statistical measures are used to compare model performance, including ensemble averaging of results. Results are interpreted in context of whether the models reproduce the physics of the Baltic Sea accurately enough to provide the basis of biogeochemical models involving the implementation of ecosystem models.

Material and methods

Description of the models

Six models were compared regarding their simulation of the hydrodynamics for the GoF (Table 1):

1. HIROMB, the operational model of the Swedish Meteorological and Hydrological Institute (Funkquist 2001).
2. The OAAS model developed by O. Andrejev and A. Sokolov (Andrejev and Sokolov 1989, Sokolov *et al*. 1997, Andrejev *et al*. 2004a, 2004b). This model has been used for operational forecasts by the Finnish Institute of Marine Research (FIMR), as well as by Stockholm University and the State Oceanographic Institute, Russia.
3. SPBM developed by I. Neelov (Neelov 1982, Neelov *et al*. 2003) in the St. Petersburg Branch of the P. P.Shirshov Institute of Oceanology, the Russian Academy of Sciences, and recently used by a consortium of St. Petersburg institutes.
4. EIA (Simons 1980, Koponen *et al*. 1992, Inkala and Myrberg 2002) developed and exploited in the Environmental Impact Assessment Centre of Finland Ltd.
5. COHERENS (Luyten *et al*. 1999). This model was implemented for the Baltic Sea by the National Environmental Research Institute (NERI) in Denmark, and is currently being exploited by NERI and the Finnish Environment Institute (SYKE).
6. MIKE3 (DHI Water and Environment 2000), developed at the DHI (Danish Hydraulic Institute) and used by the DHI and the Marine Systems Institute (MSI) in Estonia.

All the participating models were set up for the whole Baltic Sea using horizontal grids that have the same resolution $4' \times 2'$ (Fig. 1) based on the grid created by Seifert and Kayser (1995). The models were assigned a common setup in terms of initial and boundary conditions and forcing fields (Table 2). The calculation of the surface energy balance was thus treated as a part of each model system and consequently the calculations were carried out using variable parameterizations and using the model given SST.

The combination contains only one non-hydrostatic model (MIKE3). For the hydrostatic models, convection was parameterized by convective adjustment or, if there was no convective adjustment, the case of unstable stratification was included into the turbulence model.

Five models use z-coordinates in the vertical direction, COHERENS being the only one to use σ-coordinates. The models differ in their vertical resolution, vertical turbulence schemes, methods to approximate advective terms, parameterizations of heat fluxes, precipitation–evaporation balance at the sea surface, and their equations of state. The inclusion of ice dynamics is essential for multi-year simulations, but is not of major importance in this study which focused on the summer-autumn period during 1996, during which the GoF was ice-free.

Table 1. The basic features of hydrodynamic modules of participating models.

Model ID	HIROMB	OAAS	SPBM	EIA	COHERENS	MIKE3
Horizontal grid and resolution	Spherical Arakawa C grid, 4′ × 2′	Spherical, Arakawa C grid, 4′ × 2′	Spherical Arakawa C grid, 4′ × 2′	Spherical Arakawa C grid, 4′ × 2′	Spherical, Arakawa C grid, 4′ × 2′	Spherical, Arakawa C grid, 4′ × 2′
Vertical grid and resolution	z-coordinate 78 levels, min Δz = 2 m	*z*-coordinate 78 levels, min Δz = 2 m	z-coordinate 78 levels, min Δz = 2 m	z-coordinate 20 levels, min Δz = 2.5 m	σ-coordinate, 50 levels	*z*-coordinate, 120 levels, min Δz = 2 m
Vertical turbulence scheme	*k*-ω model	Kochergin scheme (1987)	*k*-*l* model	*k*-ε model	*k*-ε model	*k*-ε model
Horizontal turbulence scheme for momentum	Smagorinsky (1963)	Smagorinsky (1963)	Smagorinsky (1963)	Smagorinsky (1963)	none	Smagorinsky (1963)
Horizontal turbulence scheme for *T* and *S*	Smagorinsky (1963)	Smagorinsky (1963)	K_l = const = 10^6 cm^2 s^{-1}	none	none	Smagorinsky (1963)
Advection scheme for momentum	Conservative and fully 3D scheme (Zalesak 1979)	Upwind scheme	3rd order scheme (Fujii and Obayashi 1989)	TVD superbee scheme (Roe 1985)	Upwind scheme	3rdorder scheme QUICKEST (Vested *et al.* 1992)
Advection scheme for tracers (*T*, *S* and others)	Conservative and fully 3D scheme (Zalesak 1979)	TVD superbee scheme (Roe 1985)	3rd order scheme (Fujii and Obayashi 1989)	TVD superbee scheme (Roe 1985)	TVD-superbee scheme (Roe 1985)	3rd order scheme QUICKEST (Vested *et al.* 1992)
Convection	Hydrostatic model, convective adjustment	Hydrostatic model, convective adjustment	Hydrostatic model, modelling k_z = 1 m^2 s^{-1} for unstable stratification	Hydrostatic model, modelling k_z = 1 m^2 s^{-1} for unstable stratification	Hydrostatic model, convective adjustment	Non-hydrostatic model
Equation of state	UNESCO (1981)	Millero and Kremling (1976)	Millero and Kremling (1976)	UNESCO (1981)	UNESCO (1981)	UNESCO (1981)
Sea surface heat fluxes:						
1. Short-wave radiation	Funkquist (2001)	Rosati and Miyakoda (1988)	Zillmann (1972)	Kennedy (1944), Klein (1948)	Luyten *et al.* (1999)	Reed (1977)
2. Long-wave radiation	Idso and Jackson (1969)	Gill (1982)	Berlyand (1956)	Iziomon *et al.* (2003)	Luyten *et al.* (1999)	Brunt (1932)
3. Sensible heat flux	Liu *et al.* (1979)	Luyten *et al.* (1999)	Bulk formulation, $C_D = 1.75 \times 10^{-3}$	Bowen (1926)	Luyten *et al.* (1999)	Bulk formulation, $C_D = 1.4 \times 10^{-3}$
4. Latent heat flux	Liu *et al.* (1979)	Luyten *et al.* (1999)	Bulk formulation (Bryan *et al.* 1996)	Marciano and Harbeck (1954)	Luyten *et al.* (1999)	Bulk formulation (Bryan *et al.* 1996)

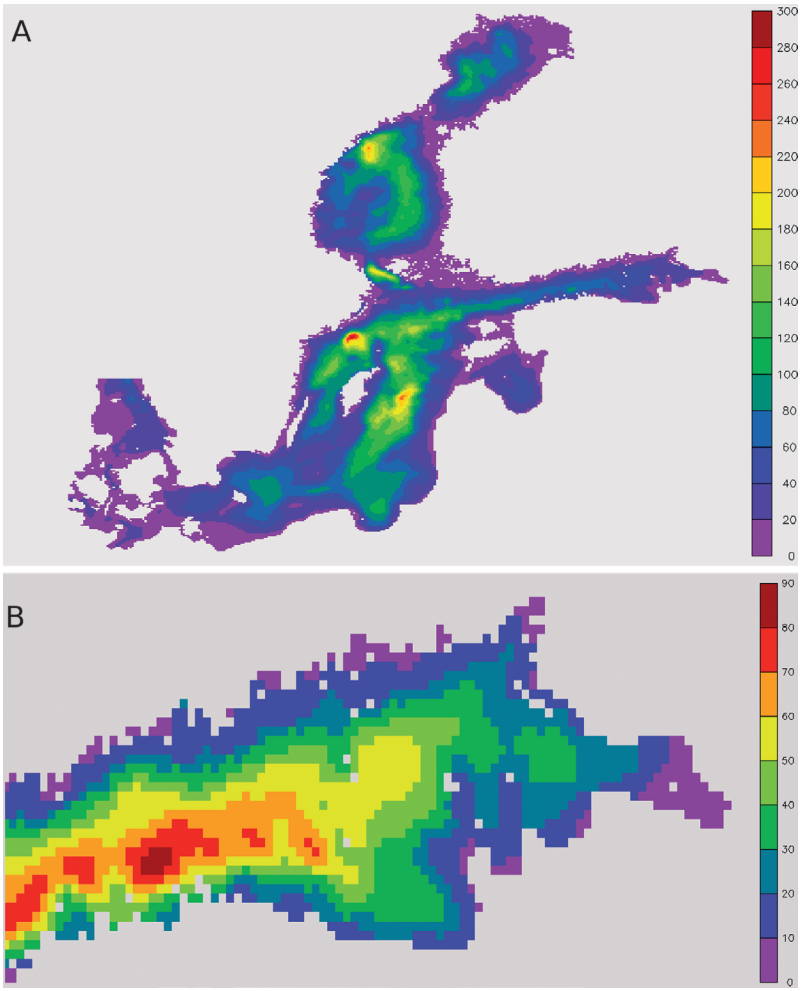

Fig. 1. Bathymetry (m) of the model domain: (**a**) for the whole Baltic Sea and (**b**) for the region of interest, the Gulf of Finland.

Simulation setup

The models were set up to simulate the period from 1 April to 1 November 1996. Initial distributions of temperature (Fig. 2A and B) and salinity fields (Fig. 2C and D) in the Baltic Sea were constructed from the data available in the Baltic Environmental Database (Sokolov *et al*. 1997) for January–March 1995 and 1996. By using data for two three-month periods, a satisfactory coverage of the Baltic Sea was achieved, providing reasonable initial fields (if only 1996 data were used, some parts of the Baltic Sea, including the GoF, had areas without data coverage).

The meteorological forcing (wind speed and direction, air temperature, relative humidity, cloudiness and precipitation) was taken from the SMHI gridded data set with a resolution of 1° (Table 2). Preliminary analysis of these data showed that the geostrophic wind-velocity fields contain some unrealistically high values; a correction was therefore made such that any wind speeds exceeding 40 m s^{-1} were adjusted to equal this value. From the geostrophic wind, the near-surface wind (10 m) was calculated using a standard procedure in which the wind speed was multiplied by 0.6 and the direction turned 15° to the left (Bumke and Hasse 1989). Precipitation for all models was taken from the SMHI data except for HIROMB where precipitation was instead set to equal evaporation.

Monthly mean river discharges were obtained from Bergström and Carlsson (1994). Preliminary work showed that prescribing a usual "no

heat flux" condition at the mouth of the Neva leads to predicted water temperatures in the Neva Bay and the easternmost part GoF being overestimated. In order to overcome the discrepancy, the water temperature in the Neva was modelled using available observational data. The "no heat flux" condition was kept for the other rivers. The open boundary condition at the Danish Straits was taken from model results of HIROMB, and used in all the participating models, rather than using the available scarce data (Table 2).

In order to be sure that all differences seen in the model results were due to the diversity of the models themselves, assimilation of observed data was prohibited except in the case of HIROMB (which was used to produce the open boundary conditions that were used by the other models). The HIROMB model assimilated temperature and salinity data once a month from HELCOM stations BY07, BY09, BY15, and BY29, which are located in the Baltic Proper (for location details *see* http://www.helcom.fi/groups/monas/CombineManual/PartA/en_GB/main/). As for the GoF, situated more than 150 km from the nearest station (BY29), it can be assumed that the assimilation had little impact even on the results for this area.

If the initial distributions of temperature and salinity are realistic with respect to the seasonal characteristics and the inclined halocline, Andrejev *et al*. (2000) indicated a spin-up time of about one month for models of the Baltic Sea. This is sufficient to remove irregularities in the initial temperature and salinity fields (*see* Fig. 2) and to adjust the model to the external forcing. Here, a spin-up of two months was used from

Table 2. Conditions for the short-period simulation: bathymetry, forcing, boundary and initial conditions for the whole Baltic Sea

Parameter	Description	Period	Data source
Sea depth	Depths on the grid $4' \times 2'$ with the SW corner having coordinates 53.8°N, 9.45°E	n/a	Seifert and Kayser (1995)
Atmospheric forcing (wind velocity, air temperature, relative humidity, cloudiness, precipitation, pressure)	SMHI gridded data, temporal resolution 3h, spatial resolution 1°	1 April 1996 to 31 October 1996	K. Boqvist (pers. comm.)
River discharge	Monthly mean values for Baltic Sea rivers	Monthly climatology	Bergström and Carlsson (1994)
Conditions for salinity and temperature in river mouths: $S = 0$, zero heat flux in all rivers excepting Neva, $T = T(t)$ in Neva	Temperature values in Neva averaged over 10-day periods	1 April 1996 to 31 October 1996	Valery Tsepelev (pers. comm.)
Boundary conditions in Danish Straits: current velocity (U, V), temperature T, salinity S	Model results for one grid point of 75 m depth, temporal resolution 3h, 11 levels with min $\Delta z = 4$ m	1 April 1996 to 31 October 1996	Results of HIROMB prepared by L. Funkquist
Initial conditions: temperature (T), salinity (S), zero values for: current velocity, sea level, ice thickness and concentration	Averaged values for winter (January–March) of two years (1995–1996)	1 April 1996	Baltic Environment Database at Stockholm University

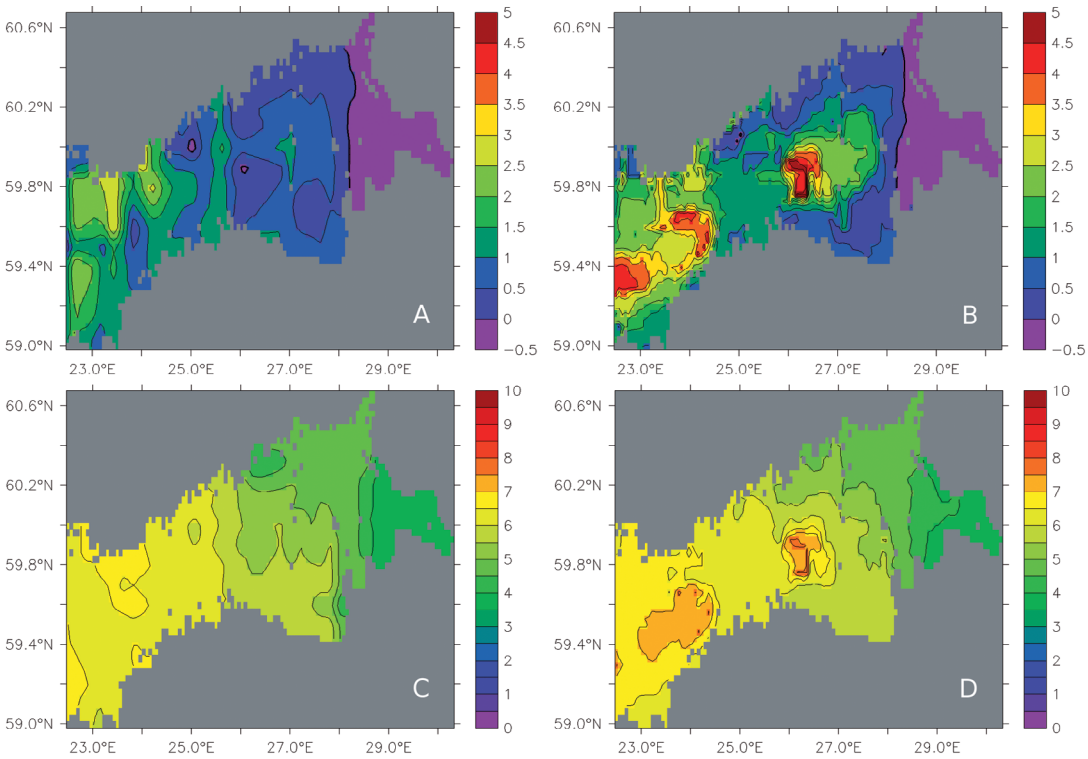

Fig. 2. The initial (winter) distributions of temperature (°C) and salinity (‰) in the Gulf of Finland taken from http://data.ecology.su.se/Models/bed.htm. (**A**) the sea-surface temperature, (**B**) near-bottom temperature, (**C**) sea-surface salinity, (**D**) near-bottom salinity.

the beginning of the run. The comparison of the model results with observations is undertaken for the summer–autumn period starting from 1 June.

Comparison with the data

Output for each model was interpolated on a unified grid coinciding in a horizontal plane with a sea-depth grid (*see* Table 2) that has 50 levels in the upper 100 m with $\Delta z = 2$ m starting from $z_1 = 1$ m to $z_{50} = 99$ m and 27 levels below with $\Delta z = 5$ m, $z_{51} = 102.5$ m, $z_{52} = 107.5$ m, …, $z_{77} = 232.5$ m. Three-dimensional distributions of water temperature T and salinity S were averaged in the GoF over five days for the period from 1 June to 1 November. Observations of temperature and salinity for the GoF in 1996 were used for comparison with model results, including both satellite and ship data (Table 3). The reason for averaging over a five-day period is due to following factors. Technically it was impossible to save three-dimensional model fields of six models with a high frequency (an interval of several hours) which would be needed to obtain model-derived profiles coinciding exactly with the times when corresponding observations were made. Moreover, the averaging allows us to filter out high-frequency, artificial (numerical) noise from the model solutions. It should be emphasized that the purpose of this study is to estimate the quality of model performance in relation to the seasonal variability of temperature and salinity fields rather than their synoptic variability. This cannot be simulated using the model resolution employed here, i.e., the horizontal resolution (2 × 2 nautical miles) of models used is not high enough to resolve meso-scale eddies because the internal Rossby-radius of deformation in the GoF is between 2 and 4 km (Alenius *et al*. 2003).

At first, the performance of the models was visually assessed by comparing the results with

available data. For this purpose, the data were reconstructed into sea-surface and bottom maps of temperature and salinity, as well as vertical temperature and salinity sections across and along the Gulf. Time-depth plots of temperature and salinity were also constructed for three of the Finnish intensive monitoring stations (SYKE), permitting analysis of the seasonal evolution of temperature and salinity in the Gulf. Sea-surface temperature (SST) averaged over five-day periods were obtained from daily mean SST derived from satellite measurements (PODAAS). Comparison was performed for each model individually and also for the ensemble mean.

Statistical analysis of the differences between model outputs and the data was performed for three sets of high-resolution vertical profiles of temperature and salinity: (1) all available r/v *Aranda* CTD data (Finnish Institute of Marine Research) collected in the western GoF between late June and early July (Fig. 3A), (2) r/v *Aranda* CTD data collected in the western GoF during mid-July (Fig. 3A), and (3) r/v *Nikolay Matusevich* CTD data collected in the eastern GoF, also in mid-July (Fig. 3B). Other available data (from Sokolov *et al*. 1997 and the SYKE intensive monitoring stations) were not used in the statistical analysis because measurements were undertaken with only a low vertical resolution and the length of corresponding data series of these vertical profiles was not long enough.

In the statistical analysis, vertical temperature and salinity distributions (both observed and modeled) were considered as series of data. Observed data x_j ($j = 1 \ldots M$) at the instant t_o at a particular hydrographic station were sampled at depths z_j with $\Delta z = 1$ m and $\Delta z = 2$ m respectively for stations of r/v *Aranda* and r/v *Nikolay Matusevich*, z_1 being equal to 1 m in both cases. The calculated data y_i at depths z_i with $\Delta z = 2$ m ($z_1 = 1$ m, $i = 1 \ldots N$) from a model were taken at the grid point nearest to the station considered at the instant t_c which was chosen from the condition $|t_o - t_c| \le \Delta t/2$ where $\Delta t = 5$ days is the interval of saving of averaged calculated fields. In order to compare observed and calculated data, the former were taken only at those depths where calculated data were available, i.e. the observed data x_j ($j = 1, 3, 5, \ldots$) was matched with the calculated values y_i ($i = 1, 2, 3, \ldots, N$) in the case of

Table 3. Observation data for the Gulf of Finland in 1996 used for comparison with model results

Parameter	Short description	Period	Data source
Sea surface temperature (SST)	*AVHRR Pathfinder SST estimation; method involves the regression of satellite-sensed brightness temperatures against in situ SST observations; daily data have spatial resolution of 4 km	1 May–31 August 1996	NASA, Physical Oceanography Distributed Active Archive Center (http://poet.jpl.nasa.gov/)
Ship data combined into a special database			
Temperature and salinity (inventoried in http://data.ecology.su.se/Models/bed.htm)	Stations performed in the GoF	1 May–31 August 1996	http://data.ecology.su.se/Models/bed.htm
Temperature and salinity (not inventoried in http://data.ecology.su.se/Models/bed.htm)	Four Finnish coastal stations of intensive monitoring (Haapasaari, Huovari, Längden, Länsi-Tonttu)	1 May–31 August 1996	SYKE (Finnish Environment Institute)
Temperature and salinity (not inventoried in http://data.ecology.su.se/Models/bed.htm)	Russian stations performed in the eastern GoF	1 May–31 August 1996	Russian State Hydrometeorological University, Russian North-West Hydrometeorological Service

*Physical Oceanography DAAC, AVHRR Pathfinder sea surface temperature (NOAA/NASA) ver. 5.0. NASA JPL Physical Oceanography DAAC, Pasadena, CA.

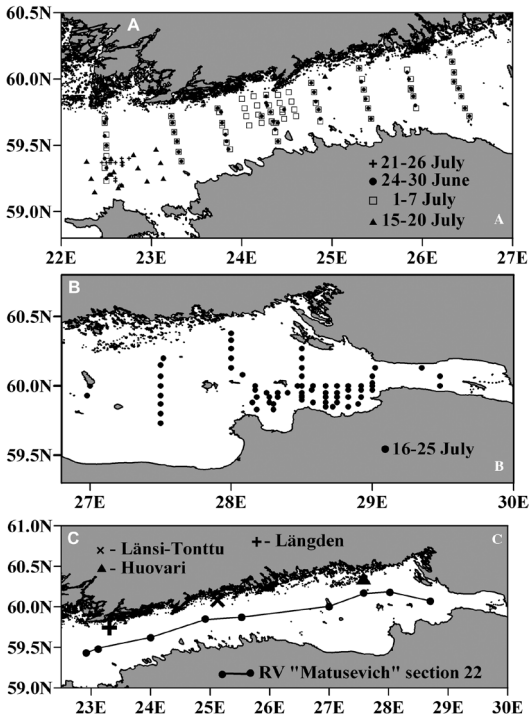

Fig. 3. (**A**) Locations of r/v "*Aranda*" stations in the western GoF in June–July 1996. (**B**) Locations of r/v *Nikolay Matusevich* stations in the eastern GoF in July 1996. (**C**) The section along the GoF from where data were collected by r/v *Nikolay Matusevich* during 11–12 August 1996, and the locations of SYKE intensive monitoring station Huovari (60°23.30´N, 27°39.49´E), Länsi-Tonttu (60°04.99´N, 25°07.39´E), Längden (59°46.60´N, 23°15.98´E).

Aranda stations, and observed data x_j ($j = 1, 2, 3, \ldots$) with y_i ($i = 1, 2, 3, \ldots, N$) at *Nikolay Matusevich* stations. In other words, no interpolation of observed data in a z-direction was done.

The following statistical characteristics were calculated: means for observed and simulated values (Eqs. 1 and 2, respectively):

$$\overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i, \qquad (1)$$

$$\overline{y} = \frac{1}{N}\sum_{i=1}^{N} y_i, \qquad (2)$$

standard deviations for observed and simulated values (Eqs. 3 and 4, respectively):

$$\sigma_{xx} = \sqrt{\frac{\sum (x_i - \overline{x})^2}{N-1}}, \qquad (3)$$

$$\sigma_{yy} = \sqrt{\frac{\sum (y_i - \overline{y})^2}{N-1}}, \qquad (4)$$

correlation coefficient (R)

$$R_{xy} = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{(N-1)\sigma_{xx}\sigma_{yy}}, \qquad (5)$$

mean absolute error:

$$\mathrm{MAE} = \frac{\sum |x_i - y_i|}{N}, \qquad (6)$$

root mean square error (RMSE)

$$\mathrm{RMSE} = \sqrt{\frac{\sum (x_i - y_i)^2}{N}}, \qquad (7)$$

bias (mean error):

$$B = \overline{y} - \overline{x}, \qquad (8)$$

and spread:

$$S = \sqrt{\frac{1}{N-1}\sum \left[(x_i - \overline{x}) - (y_i - \overline{y}) \right]^2}. \qquad (9)$$

The length of the series (N) varied between 15 and 30.

The statistical characteristics for 213 stations of r/v *Aranda* (western GoF) and for 69 stations of r/v *Nikolay Matusevich* (eastern GoF) were compared with the corresponding results from each of the six models, separately for temperature and salinity. Correlation coefficients, mean absolute errors, root mean square errors and spreads, defining the fit of model results to observations, were averaged over stations from each of these three regions (Tables 4 and 5). A statistical analysis of the differences between the outputs of all six models was also performed for vertical profiles of temperature and salinity in the open GoF.

According to Kattsov and Meleshko (2004), the procedure of ranking the six models based on normalized RMSE (resulting from dividing Eq. 7 by Eq. 3) and R of selected variables was used. The normalized RMSE and R were calculated for each vertical profile — depths z_i with $\Delta z = 2$ m ($z_1 = 1$ m, $i = 1 \ldots N$) — of temperature and salinity and averaged over two regions (western and eastern GoF) separately for temperature and salinity.

Table 4. Statistics for the model results compared with observations for temperature in the western and eastern Gulf of Finland.

Model	Correlation coefficient (*R*)	MAE (‰)	RMSE (‰)	Spread (‰)
The western Gulf of Finland,				
24 June–4 July 1996, 172 profiles of r/v *Aranda*				
HIROMB	0.86	2.05	2.64	1.92
OAAS	0.95	1.16	1.55	1.19
SPBM	0.93	1.19	1.52	1.22
EIA	0.92	1.70	2.03	1.57
COHERENS	0.95	1.06	1.39	1.18
MIKE3	0.87	2.11	2.62	1.81
The western Gulf of Finland,				
15–26 July 1996, 41 profiles of r/v *Aranda*				
HIROMB	0.96	1.49	1.85	1.22
OAAS	0.89	2.16	2.77	1.88
SPBM	0.92	1.92	2.46	1.64
EIA	0.94	1.59	1.85	1.40
COHERENS	0.92	1.67	2.10	1.62
MIKE3	0.74	3.34	4.38	3.38
The eastern Gulf of Finland,				
16–25 July 1996, 69 profiles of r/v *Nikolay Matusevich*				
HIROMB	0.91	1.28	1.52	0.97
OAAS	0.82	1.92	2.20	1.28
SPBM	0.84	2.08	2.20	0.98
EIA	0.72	1.76	1.95	1.65
COHERENS	0.81	1.56	1.94	1.39
MIKE3	0.80	3.62	4.30	3.82

Table 5. Statistics for the model results compared with observations for salinity in the western and eastern Gulf of Finland

Model	Correlation coefficient (*R*)	MAE (‰)	RMSE (‰)	Spread (‰)
The western Gulf of Finland,				
24 June–4 July 1996, 172 profiles of r/v *Aranda*				
HIROMB	0.87	0.32	0.36	0.27
OAAS	0.97	0.34	0.38	0.33
SPBM	0.93	0.32	0.37	0.32
EIA	0.92	0.36	0.41	0.34
COHERENS	0.94	0.33	0.37	0.32
MIKE3	0.95	0.31	0.35	0.28
The western Gulf of Finland,				
15–26 July 1996, 41 profiles of r/v *Aranda*				
HIROMB	0.87	0.25	0.31	0.22
OAAS	0.90	0.23	0.28	0.21
SPBM	0.89	0.22	0.28	0.23
EIA	0.78	0.34	0.42	0.38
COHERENS	0.87	0.30	0.36	0.29
MIKE3	0.89	0.27	0.33	0.26
The eastern Gulf of Finland,				
16–25 July 1996, 69 profiles of r/v *Nikolay Matusevich*				
HIROMB	0.90	0.74	0.79	0.34
OAAS	0.77	0.62	0.66	0.37
SPBM	0.81	0.63	0.67	0.31
EIA	0.82	0.78	0.85	0.38
COHERENS	0.78	0.82	0.88	0.36
MIKE3	0.88	0.51	0.59	0.40

Results

When comparing modeled and observed hydrographic characteristics, it is important to be sure that the model solutions are not affected by the modelled initial values. The initial distribution of temperature in the GoF (Fig. 2A and B) is typical for the winter period. SST increases gradually from near 0 °C in the easternmost basin to 1.5 °C in the west, whereas near-bottom temperature varies from 0 °C in the east to 4 °C in the western end of the Gulf, with a local maximum of about 4 °C at 26°E. The thermocline is absent throughout the Gulf, the water masses being well-mixed in the east whereas in the west temperature increases at depths below 30–40 meters due to the advection of warmer water masses from the Baltic Sea Proper.

The surface values of the initial winter distribution of salinity (Fig. 2C and D) increase from east to west from about 4‰ (because of the lack of data the values are overestimated) in the easternmost part of the GoF to 6.0‰–6.5‰ in the western GoF and from north to south by 0.5‰–1.5‰. In the near-bottom layer, the pattern is approximately the same with slightly higher salinity (about 7‰) in the western Gulf and a local maximum 7‰ in the deep central part of the GoF. A weak halocline exists only in the western GoF at a depth of 50–80 m. In other parts of the Gulf the vertical salinity distribution is almost homogeneous. Due to the lack of data, especially in the eastern GoF, the salinity distribution, based on observations, does not reflect the real winter situation, mainly due to underestimation of the influence of the Neva River. Modeled temperature and salinity distributions for summer (*see* Figs. 4–9) differ strongly from the distributions for winter used to initialize the models. This feature allows us to conclude that the model solutions for the summer–autumn period are to a large extent adjusted to the external forcing.

Horizontal distributions

According to satellite and ship measurements, SST during summer warming (26–30 June 1996) was between 12 and 14 °C (Fig. 4A and B). A gradual increase in temperature by about two degrees was seen between the western part of the Finnish coastal zone and the southeastern Gulf. This increase has been practically identified by ship-based observations, although no ship data exist in the eastern Gulf. The gross features of the observed SST were reproduced by all the models (Fig. 4C–H). However, the EIA model (Fig. 4F) underestimated SST in the central part of the Gulf by 2–3 °C due to the low vertical resolution of the model and problems with vertical mixing.

According to observations, SST was rather homogeneous in the Gulf from 11–15 August when a high pressure system covered the Baltic region with weak winds at that time (Fig. 5A and B). SST varied between 18 and 20 °C, with lowest temperatures (17 °C) near the western part of the Estonian coast. Four of the models (HIROMB, OAAS, EIA and COHERENS; Fig. 5C, D, F, G) predicted a more or less homogeneous distribution of SST in the Gulf, consistent with the observations to within a few degrees. The SPBM model (Fig, 5E) underestimated SST by about 3 °C. All the models showed a weak upwelling in the western Gulf near the coast of Finland which was not evident in the measurements. This feature was most prominent in the results of the MIKE3 model (Fig. 5H), which overestimated the decrease in SST by several degrees in that region.

The distribution of observed surface salinity averaged over the June–August period showed an increase from close to 0‰ in the Neva mouth to about 6‰ in the western part of the GoF (Fig. 6A). All models reproduced the general pattern of a cyclonic mean circulation (Fig. 6C–H), although greater in magnitude than that of the observations, giving rise to lower salinities near the coast of Finland as compared with corresponding values off the Estonian coast (this difference being at most about 0.5‰). The impact of the Neva River on salinity was slightly overestimated by some models (HIROMB, COHERENS, MIKE3; Fig. 6C, G, H), with predicted salinities that were lower than the measured ones near the coast of Finland and in the easternmost part of the Gulf. Opposite results were produced by the OAAS and SPBIO models for this region (Fig. 6D and E). The results of the EIA model

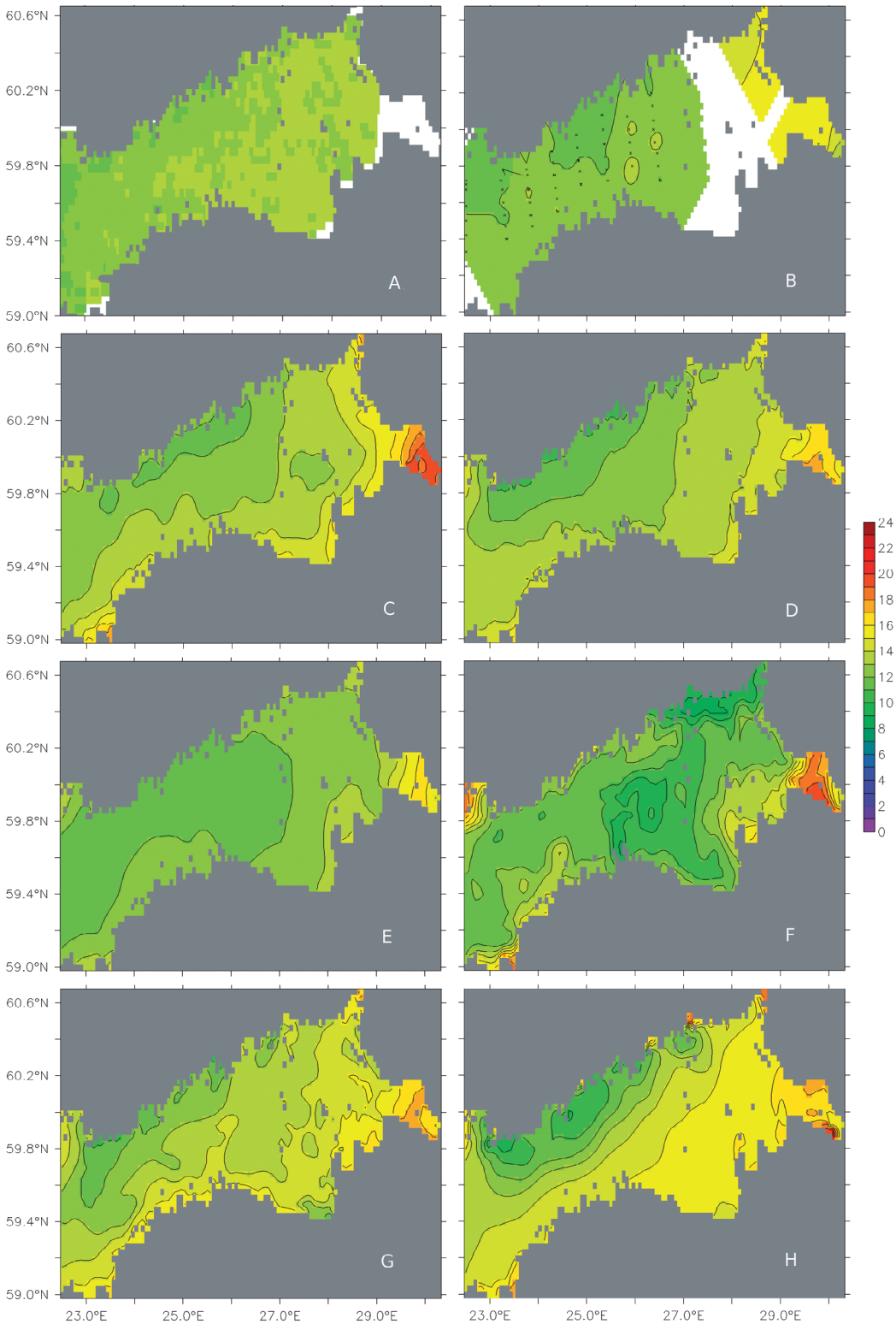

Fig. 4. Sea-surface temperatures (SST, °C) in the Gulf of Finland during 26–30 June 1996 according to (**A**) satellite data, (**B**) ship data, (**C**) HIROMB, (**D**) OAAS, (**E**) SPBM, (**F**) EIA, (**G**) COHERENS, and (**H**) MIKE3.

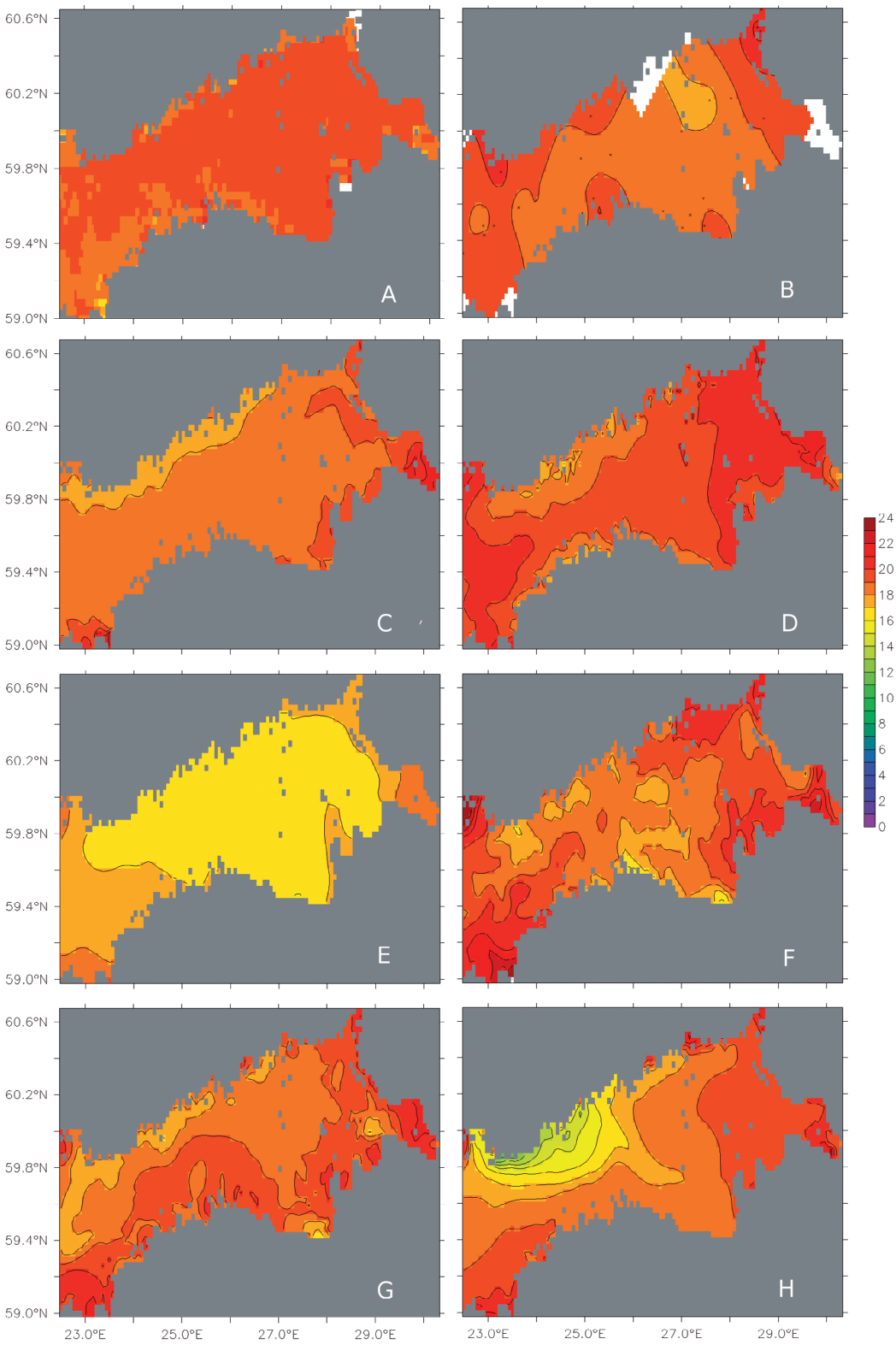

Fig. 5. Sea-surface temperatures (SST, °C) in the Gulf of Finland during 11–15 August 1996. Notation as in Fig. 4.

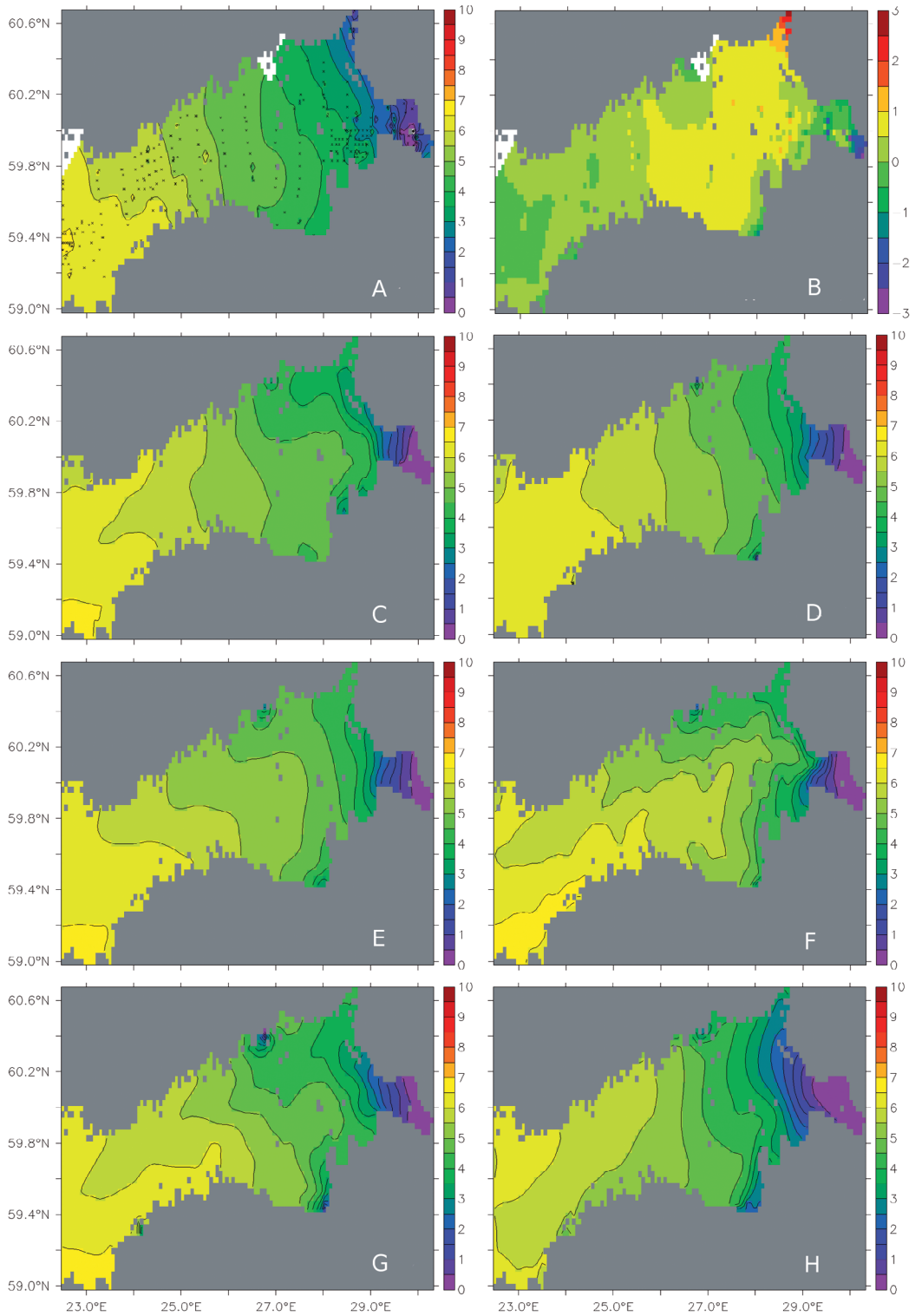

Fig. 6. Sea-surface salinities (*S*, ‰) in the Gulf of Finland averaged for 1 June–1 September 1996 according to (**A**) ship data, (**B**) difference between ensemble mean of the models and data, (**C**) HIROMB, (**D**) OAAS, (**E**) SPBM, (**F**) EIA, (**G**) COHERENS, and (**H**) MIKE3.

(Fig. 6F) exhibited a pattern of salinity in the easternmost part of the Gulf that was contradictory to the cyclonic circulation.

The ensemble mean prediction of the models (Fig. 6B), shows a difference in salinity of only 0.5‰–1‰ as compared with the observations in the open Gulf. However, the model ensemble predicted penetration of saline water too far eastwards in the upper layer relative to that observed; this leading to overestimation of salinity by 0.5‰–1.0‰. In the western Gulf the model ensemble simulated the surface salinity successfully, with an error of not more than 0.5‰.

A distinctive feature of the observed near-bottom salinity distribution during June–August was a tongue of water with relatively high salinity penetrating into the central part of the GoF from the Baltic Proper (Fig. 7A). The salinity values in this tongue were 8.5‰–9.0‰ at the entrance to the Gulf and remained high (7.0‰–7.5‰) at 27.5°E. The gulf became shallower eastward of this longitude and had no specific bottom layer with a halocline. All models (Fig. 7C–H) made good predictions of the penetration of this salt water tongue as far east as 27.5°E, the most successful being MIKE3 (Fig. 7H). The salinity as predicted by the models was however not as high as that seen in the data, the underestimation being 0.5‰–1‰. The predicted vertical gradient of salinity was thus smoother than that observed.

The discrepancy in bottom salinity between the model results and observations can be explained by the fact that the depth levels in the model grid are shallower than the real depths, due to the fact that the resolution of the model grid is 2 × 2 nautical miles. The near-bottom salinity near the entrance of the GoF was underestimated by SPB (by 0.5‰), EIA (by 1.0‰) and COHERENS (by 0.5‰, Fig. 7E, F, G, respectively). The ensemble-mean difference between the model results and the measurements was usually 0.5‰–1‰ in the open Gulf — the model ensemble underestimating the salinity (Fig. 7B). Larger discrepancies were found only in coastal regions where the depth of the model grid differed significantly from reality. It was not possible to estimate the accuracy of salinity near the coasts because of lack of data.

Sections along the GoF

A temperature section along the central axis of the GoF was measured on board the r/v *Nikolay Matusevich* during 11–12 August 1996 (*see* Fig. 3C), between longitudes 23°E and 29°E. The corresponding model simulations covered the period between 11 and 15 August. By this time, warming of the upper layer was strong and a weakly mixed layer could be identified in the upper 15 m where temperature drops from 18–19 °C to 15 °C (Fig. 8A). A strong thermocline was present between 10 and 35 m. At the depths of more than 40 m, water temperature was usually close to 4 °C although the westernmost Gulf hosted slightly warmer water masses below 85 m originating from the Baltic proper. Note that along the section, the sea depth in the common output grid is shallower than the real sea depth.

The HIROMB model predicted a realistic upper mixed layer, but the predicted thermocline (Fig. 8B) was thicker than that observed such that there was no homogeneous layer below 40 m. In contrast, the OAAS model (Fig. 8C) produced no clear upper mixed layer and, as a result, the thermocline was located continuously between the surface and 40 m. A homogenous bottom layer exists below this depth where predicted temperature was some 2 °C lower than that observed. The SPBM model (Fig. 8D) produced a realistic upper mixed layer of 5–7 m in this area. The predicted thermocline extended to the 35-m depth, consistent with observations, although the simulated temperature was about 1 °C lower than that observed. An intermediate cold water layer between 40 and 65 m was simulated by the SPBM model, with temperatures that were 2 °C lower than those observed. The model did however capture the near-bottom penetration of warmer water from the Baltic Sea. The EIA model (Fig. 8E) failed to reproduce the upper mixed layer, although the predicted depth of the thermocline was realistic. Nevertheless, the simulated water mass in the bottom layer was too warm. Similarly, an upper mixed layer could barely be identified in the results of the COHERENS model (Fig. 8F) which predicted that the thermocline extended down to 45 m. The predicted temperature in deep water did however

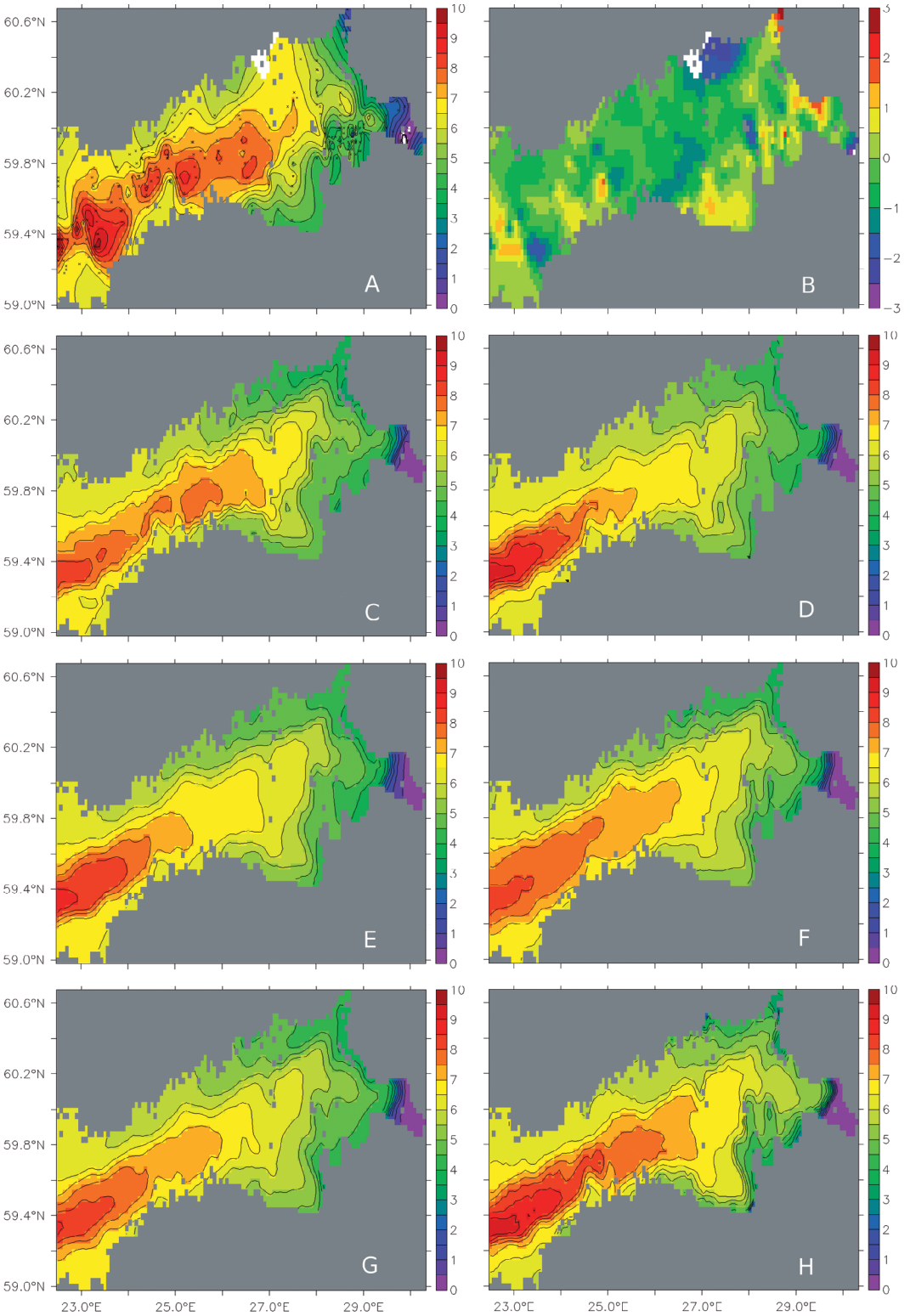

Fig. 7. Near-bottom salinities (*S*, ‰) in the Gulf of Finland averaged for 1 June–1 September 1996. Notation as in Fig. 6.

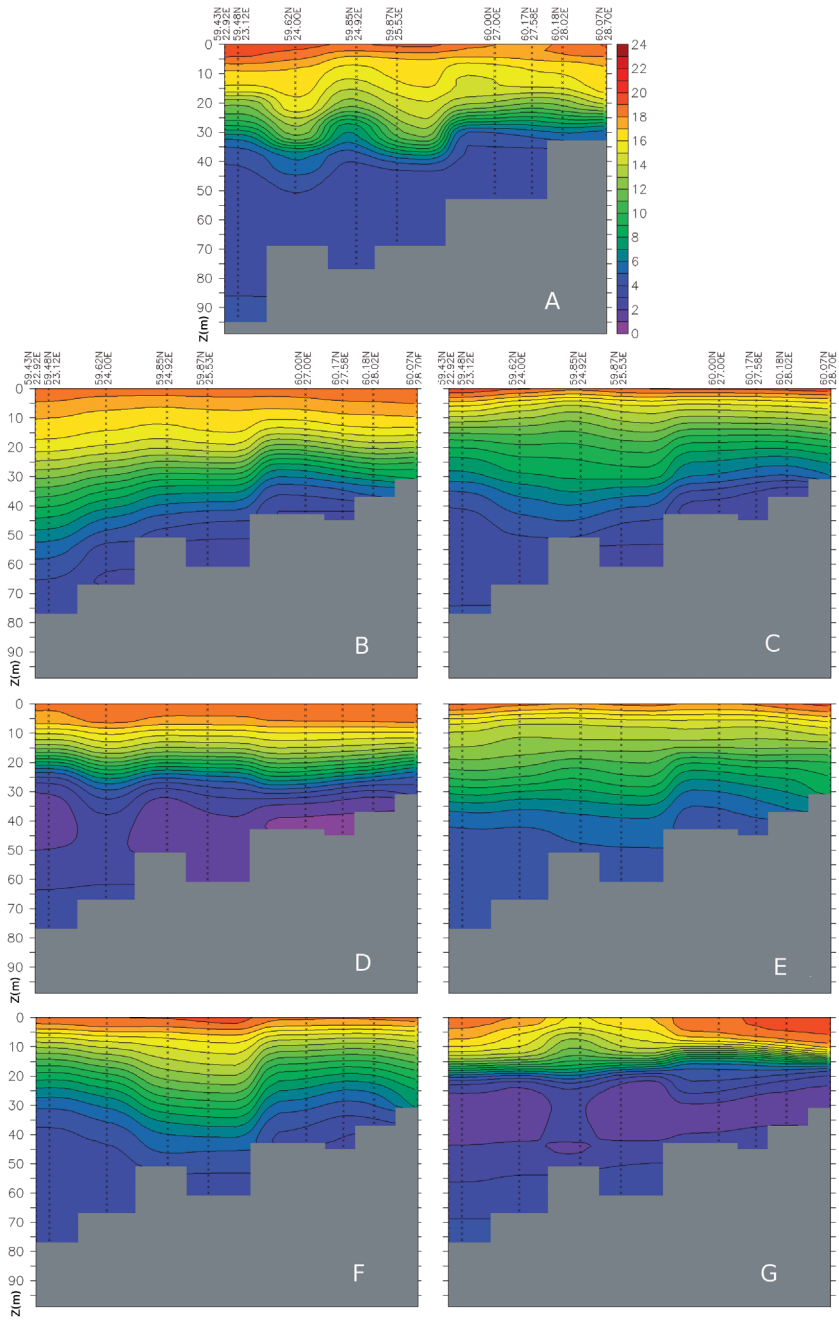

Fig. 8. Temperature (°C) section along the GoF (*see* Fig. 3) (**A**) according to the observations carried out onboard r/v *Nikolay Matusevich* during 11–12 August 1996 and modelled for the period of 11– 15 August 1996 using (**B**) HIROMB, (**C**) OAAS, (**D**) SPBM, (**E**) EIA, (**F**) COHERENS, and (**G**) MIKE3.

fit well with observations. In the prediction of the mixed layer structure by the MIKE3 model (Fig. 8G), the thermocline was too thin, the intermediate cold-water layer too shallow and the predicted water temperature was underestimated by 2 °C. An upwelling region was predicted in the western Gulf, which did not appear in the observations.

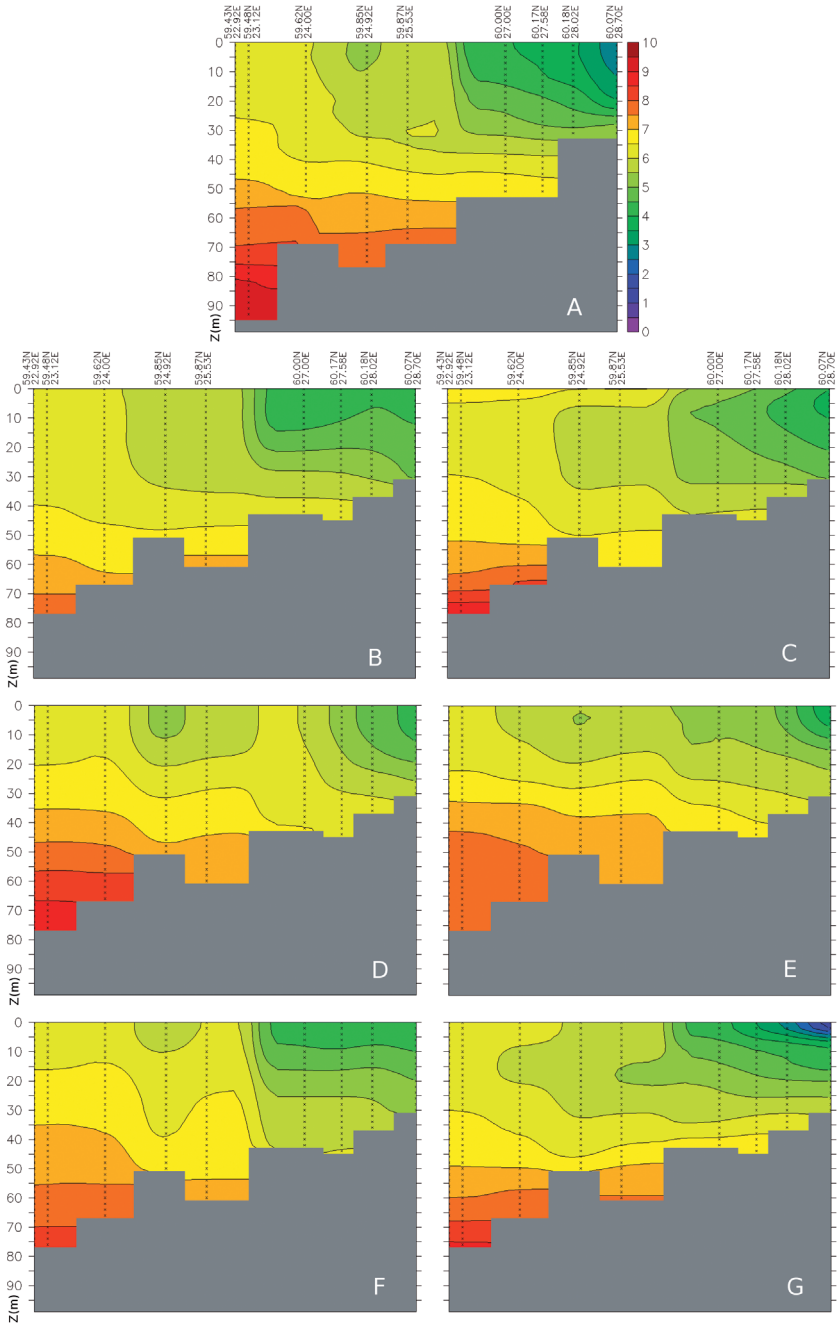

Fig. 9. Salinity (*S*, ‰) section along the GoF (*see* Fig. 3). Notation as in Fig. 8.

Measurements used for the estimation of the performance of the models were made along a corresponding salinity section across the central axis of the Gulf (Fig. 9A). Salinity increased from 3‰ at the sea surface in the eastern part of the GoF to about 9‰ near the bottom of the western part. A front, located at about 26.5°E on the section, divided fresher water ($S < 4.5$‰) and saline water ($S > 5.5$‰). A homohaline layer was located in the central part of the Gulf between 0 and 35 m, below which there was a seasonal halocline. Here, the models performed well (Fig.

9B–G). Near-bottom salinity in the western part of the section was accurately reproduced, with the exception of HIROMB (Fig. 9B) and MIKE3 (Fig. 9G). Model results differed from observations by about 0.5‰ at a depth of 70–75 m. In the eastern part of the section, all models gave 1‰–1.5‰ higher surface salinities except for MIKE3 (Fig. 9G) which produced a salinity that was about 1.5‰ lower than observed. The location of the salinity front was accurately reproduced by the HIROMB, COHERENS and MIKE3 models (Fig. 9B, F, G, respectively), although the last one overestimated the salinity towards the seaward side of the front. The other models underestimated the westward extent of the relatively fresh water. The eastward penetration of saline water in the bottom layer (isoline of S = 6.5‰ is taken as a marker) was reasonably well reproduced by SPBM, EIA, MIKE3 (Fig. 9D, E, G) and to some extent by HIROMB (Fig. 9B). The SPBM and COHERENS models successfully reproduced the westward spreading of relatively fresh water between 25.0°E and 25.5°E. Results from the OAAS and MIKE3 models (Fig. 9C and G) predicted westward intrusion of this less saline water into the thermocline forming an inverse salinity distribution, which is not supported by observations. The inverse vertical structure in salinity was balanced by a stable structure in temperature (*see* Fig. 8C and G) such that the density stratification was stable.

The above analysis shows that the models had difficulties in describing the pronounced estuarine two-layer circulation, i.e. the westward transport of fresh water in the upper layer and eastward transport of saline water beneath.

Temporal evolution in the Finnish coastal zone

The modeled temporal evolution of SST was compared with that measured at three monitoring stations in the Finnish coastal zone (Fig. 10A–C). Stations Huovari, Länsi-Tonttu and Längden are located in the eastern, central and western parts of the GoF, respectively (*see* Fig. 3C). The models successfully reproduced the qualitative trends in SST at all three stations. The temperature difference between the various models was sometimes several degrees: Some models overestimated and some underestimated the temperature. All the models failed to reproduce the high temperature observed in early June at Länsi-Tonttu (Fig. 10B), as well as the maximum summer temperature and the low temperature in mid-October at Längden (Fig. 10C). As the local SST is strongly influenced by the heat flux from the atmosphere and turbulent mixing at the sea surface, the mismatches between the observations and the model results were probably caused by the coarse resolution of the meteorological fields used in the study. Intuitively, it is obvious that inconsistent data about cloudiness or about small-scale effects such as local coastal winds, which are not described by the coarse-resolution meteorological forcing used, may lead to large uncertainties of modeled data (measured SST is about 9 °C in Huovari and Länsi-Tonttu, but about 6 °C in Längden, *see* Fig. 10). The differences in SST simulated by the various models were larger during the warming period than later in the summer and during the autumn cooling (*see* Fig. 11). No single model was consistently best for all periods and all places. For instance, during the warming period the OAAS model accurately reproduced measured values of SST at Längden (Fig. 10C), COHERENS at Länsi-Tonttu (Fig. 10B) and HIROMB at Huovari (Fig. 10A). During the cooling period, the HIROMB model showed good agreement with data at all three stations, except on 15 October in Längden (Fig. 10C) when all models failed.

The accuracy of the SST hindcast is higher in terms of the model ensemble (Fig. 11). Except for the three occasions where all models failed, and for the failure to reproduce maximum summer temperature in Länsi-Tonttu (Fig. 11B), the ensemble mean fits observations very well: the maximum difference of about 1.5 °C is reached on a few occasions only.

All the models, with the exception of MIKE3, successfully simulated the magnitude and timing of the maximum SST. The MIKE3 model reproduced the maximum temperature, but lags behind in its timing at Länsi-Tonttu and Längden (Fig. 10B and C). This failure is appar-

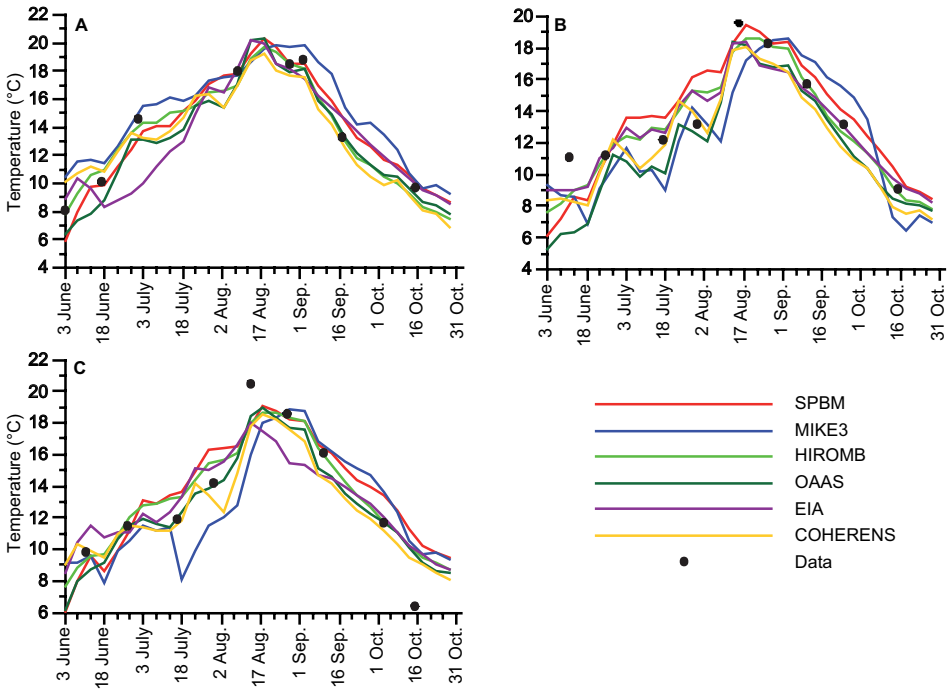

Fig. 10. Time evolution of sea-surface temperature (SST, °C) at Finnish intensive monitoring stations (**A**) Huovari, (**B**) Länsi-Tonttu and (**C**) Längden during 3 June–31 October 1996 (model results against data).

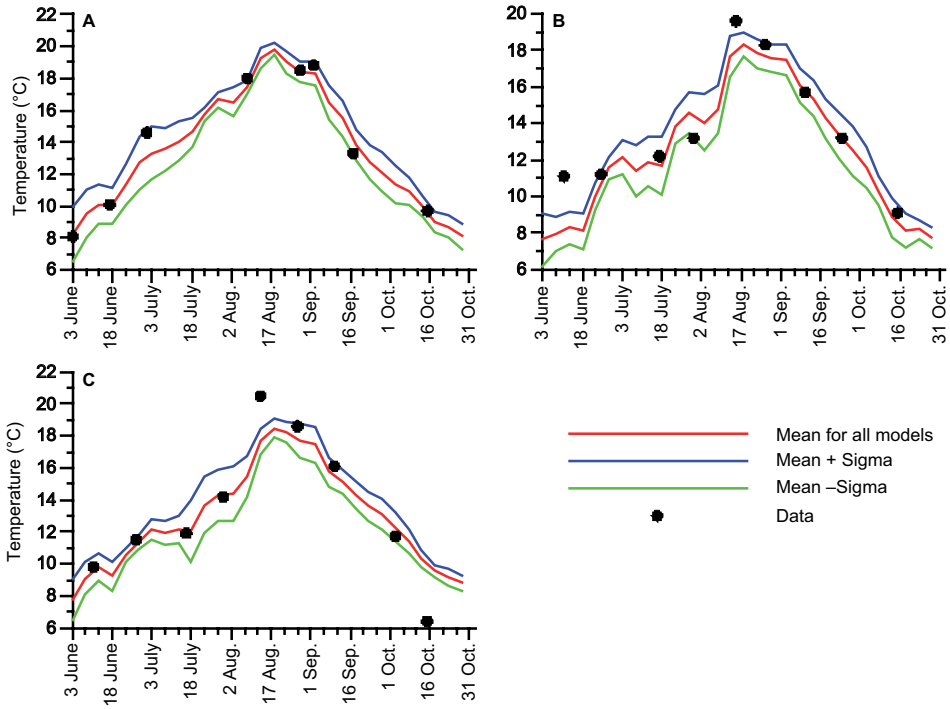

Fig. 11. Time-evolution of the model ensemble average of sea-surface temperature (SST, °C) at Finnish intensive monitoring stations (**A**) Huovari, (**B**) Länsi-Tonttu and (**C**) Längden during 3 June–31 October 1996.

ently due to the unrealistic upwelling simulated in the western GoF in mid-July. The maximum deviations from the observed SSTs were –1.5°C for Länsi-Tonttu and –2.4 °C for Längden.

The qualitative success of the models in simulating maximum SST at the three stations is shown in Table 6. These characteristics, especially the maximum SST, are indicators not only of the general quality of the models, as they serve also to identify important features of the hydrodynamic models that could have a bearing on the successful implementation of ecological models. In particular, the growth of cyanobacteria is strongly dependent upon temperature (as well as phosphorus) in that blooms are only initiated when SST exceeds a critical value (for some species about 15 °C). Maximum temperature could not be defined for Huovari because there were no measurements between 8 and 29 August, the period within which maximum temperatures were recorded at the other stations.

Model performance

A statistical analysis of the differences between the model outputs and the data was performed for three groups of detailed vertical profiles of temperature and salinity. These data allow us to compare the model performance with observations, mostly in the western Gulf at the beginning of the summer and in the western and eastern parts in the middle of the summer.

The analysis for temperature showed that the performance of the models was better in the western than in the eastern part of the Gulf (Table 4). Correlation coefficients (R) for the west were usually higher than 0.9, with lower values for the eastern Gulf. The MIKE3 model is markedly different from the other models during the second period in the western Gulf due to its predictions of unreasonably strong upwelling. Different models did best in the western Gulf at different times. The HIROMB model, for example, gave particu-

Table 6. Comparison of modeled and observed characteristics of the summer evolution of sea-surface temperature (SST) in the Finnish coastal zone.

Model	Date (t_s), when SST reaches 15 °C	Period (τ, days), when SST exceeds 15 °C	Maximum SST (°C)	Date when SST reaches maximum
Western GoF, station Längden				
Data	7 July	47	20.4	12 August
HIROMB	27 July	50	18.6	17 August
OAAS	5 August	35	19.0	17 August
SPBM	7 July	58	19.1	17 August
EIA	23 July	45	18.0	12 August
COHERENS	7 August	32	18.5	17 August
MIKE3	10 August	48	19.0	28 August
Central GoF, station Länsi-Tonttu				
Data	3 August	44	19.5	17 August
HIROMB	26 July	51	18.6	18 August
OAAS	7 August	35	18.4	12 August
SPBM	23 July	59	19.5	18 August
EIA	26 July	50	18.4	18 August
COHERENS	7 August	33	18.0	18 August
MIKE3	12 August	48	18.5	28 August
Eastern GoF, station Huovari				
Data	6 July	69	18.8	1 September
HIROMB	17 July	58	19.5	17 August
OAAS	21 July	54	20.2	17 August
SPBM	17 July	62	20.2	17 August
EIA	23 July	55	20.1	12 August
COHERENS	18 July	55	19.1	17 August
MIKE3	1 July	83	20.0	22 August

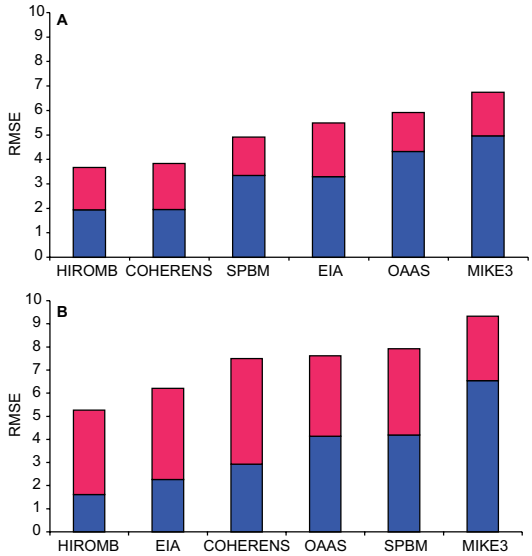

Fig. 12. Normalized RMSE averaged over a region for temperature (blue bars) and salinity (red bars) for each participating model in the (**A**) western and (**B**) eastern parts of the GoF. The sum of the normalised RMSE for temperature and salinity (blue + red bars) characterizes the performance of a model and is used for ranking.

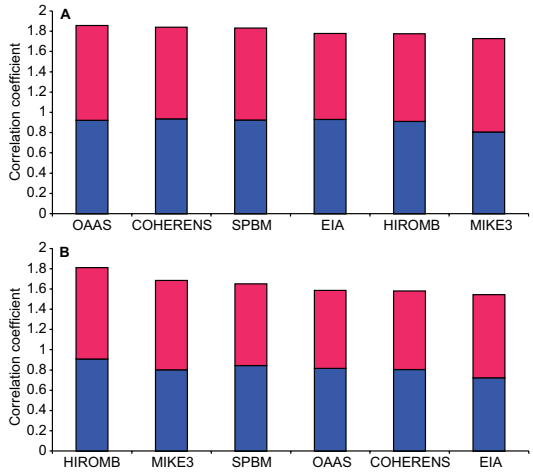

Fig. 13. Normalized correlation coefficients (*R*) averaged over a region for temperature (blue bars) and salinity (red bars) for each participating model in the (**A**) western and (**B**) eastern parts of the GoF. The sum of the normalised correlation coefficients for temperature and salinity (blue + red bars) characterizes the performance of a model and is used for ranking.

larly erroneous predictions between 24 June and 4 July, but then performed well for 15–26 July. With the exception of the difference in correlation coefficients, and the fact that several profiles were excluded from the analysis in the eastern GoF, the results in terms of other statistical characteristics are comparable for the eastern and western Gulf. The ensemble mean of the different models for different periods reproduced the temperature in the GoF with RMSE of less than 2 °C.

The highest correlation coefficients between the modeled and observed salinities were found for the western part of the Gulf at the beginning of the summer. Their values, as a rule, were greater than 0.9, and at times as high as 0.97 (Table 5). The correlation coefficients were lower in late July in the western and eastern parts of the Gulf being, as a rule, higher in the western part. The EIA model had a correlation coefficient of only 0.78 in the western gulf for July, with COHERENS and OAAS performing poorly in the eastern gulf with correlation coefficients of less than 0.8. MEA, RMSE and spread in the western Gulf were usually between 0.2‰–0.4‰. In the eastern Gulf MEA and RMSE were about 0.6‰–0.9‰ and the spread less than 0.4‰. The larger deviations in the eastern Gulf are most probably due to the pronounced salinity gradients that occur there.

The ranking of participating models in accordance with normalized RMSE (Fig. 12) indicates the relative amplitude of the simulated and observed variations, and shows that the best model in both regions considered is HIROMB, and the worst is MIKE3. The differences in model performance is quite large (in the western Gulf, RSME equals 3.8 and 6.8 for HIROMB and MIKE3, respectively; in the eastern Gulf, RSME equals 5.2 and 9.2 for HIROMB and MIKE3, respectively) and all the models work better for the western than for the eastern Gulf.

On the other hand, the ranking in accordance to *R* (Fig. 13) indicates whether the fields have similar patterns of variation regardless of amplitude, and shows that in the western Gulf the best model is OAAS, the worst MIKE3, whereas in the eastern Gulf the best model is HIROMB, the worst being EIA. However, the differences between the performances of the models are not very large.

As was pointed out above, the models that we compared differ in many respects and, there-

fore, despite a common setup in terms of initial and boundary conditions and forcing fields, it is not evident beforehand that the model solutions will demonstrate similar behavior. A statistical analysis of the differences between the outputs of all six models was, therefore, performed for vertical profiles of temperature and salinity in the open GoF, defined as the part of the Gulf with the sea depths greater than 40 m (Table 7). The modeled temperatures are highly correlated with each other: the values of correlation coefficients are usually higher than 0.94, the lowest values of 0.80–0.92 for the MIKE3 model. Predictions of salinity gave slightly lower correlation coefficients, the lowest values of 0.81–0.91 and 0.81–0.93 for OAAS and COHERENCE, respectively.

Discussion

The motivation for this study was to examine and compare the performance of a suite of models in simulating the hydrodynamics of the Baltic Sea. A model inter-comparison exercise of this kind has not been previously undertaken for this area. The relative success of six different models was assessed by detailed comparison of their predictions with measurements, as well as with each other. Such an assessment highlights the strength and weaknesses of the state of the art in the hydrodynamics modeling of the Baltic Sea.

The following six models were compared: HIROMB, OAAS, SPBM, EIA, COHERENS and MIKE3. Care was taken to ensure that forcing functions, and boundary and initial conditions were the same for all of them. Results were compared with measurements carried out during the Estonian–Finnish–Russian "Year of the GoF", including 316 profiles of salinity and temperature, and also at three Finnish monitoring stations, all in the GoF during the summer of 1996. These data allowed us to construct maps of the surface and bottom salinity of the entire GoF for the summer 1996. These maps will be useful for other purposes, such as in forthcoming model verifications or when investigating climatological conditions in the GoF.

Ensemble averaged results of the models showed no systematic over- or underestimation of hydrodynamic parameters. Predicted mean temperatures in various models generally differed from the observations by less than 1–2 °C. The mean error in salinity was less than 1‰. Nevertheless, accurate simulations of the hydrography of the Gulf turned out to be more difficult than anticipated. Despite the application of sophisticated turbulent closure schemes, the main difficulty with all the models was the correct simulation of the dynamics of the mixed layer, notably the accurate simulation of its depth and the properties of the corresponding thermo- and haloclines. None of the models was able to accurately simulate the vertical profiles of temperature and salinity. Various models used different schemes for vertical mixing (Table 1). Further, vertical resolution differed between models, leading to differences in their predictions.

Our findings, thus, indicate that ecosystem modeling studies of the Baltic Sea will need high horizontal resolution: the grid size should not exceed the internal Rossby-radius and a vertical resolution of a few meters is necessary in order to resolve the complex dynamics and topography of the Baltic Sea. One particular aspect that needs specific attention is the description of vertical advection. In some cases, the numerical

Table 7. Correlation coefficients (R) between the models for vertical profiles of temperature (first numbers) and salinity (second numbers) averaged over the open Gulf of Finland (depths greater than 40 m) during 11–15 August. The number of values in each profile was at least 20.

	OAAS	SPBM	EIA	COHERENS	MIKE3
HIROMB	0.94; 0.91	0.96; 0.97	0.96; 0.93	0.96; 0.91	0.80; 0.96
OAAS	–	0.97; 0.89	0.95; 0.84	0.99; 0.81	0.91; 0.89
SPBM	–	–	0.98; 0.96	0.97; 0.93	0.92; 0.94
EIA	–	–	–	0.96; 0.93	0.87; 0.91
COHERENS	–	–	–	–	0.90; 0.89
MIKE3	–	–	–	–	–

schemes may produce artificial diffusion, which may partly explain the problems in reproducing the observed vertical stratification. The horizontal resolution (2×2 nautical miles) of the models in use may in some areas not be high enough to resolve meso-scale eddies because the internal Rossby-radius of deformation in the GoF is between 2–4 km (Alenius *et al*. 2003). This also influences the shape of the vertical profiles. The horizontal grid spacing used in this study was a compromise between the need for accuracy and the simulation run-time. Today, it is still extremely time-consuming to carry out simulations at very fine resolution for the entire Baltic Sea.

Statistically, the models were generally able to hindcast the salinity and temperature structure more accurately in the western than in the eastern Gulf. Comparison of the models and the data showed systematic problems in correctly simulating specific features of the Baltic Sea. All the models produced an upwelling along the coast of Finland between mid-July and early August 1996, but with different intensities depending on the predicted depth of the thermocline. Although satellite imagery shows some decrease in SST along the coast of Finland in August, this feature is not as marked as was simulated by MIKE3 (*see* Fig. 4). Another difficulty was in the simulation of the salinity distribution in Neva Bay, in the very eastern part of the GoF. These difficulties arose from the coarse resolution of the model bathymetry, and in particular the absence of the St. Petersburg dam which was not represented. For the same reason, and also due to problems with atmospheric forcing, SST was not always well simulated at the Finnish coastal stations.

The model inter-comparison exercise was encouraging in terms of the ability of the models to reproduce the general hydrodynamics of the Baltic Sea. Difficulties with simulating salinity and temperature in the eastern GoF led to problems in assessing the effect of the Neva River on the physical circulation. Ecosystem models are only as good as the physical setting in which they are incorporated (Doney 1999). We may expect that the problems encountered by the models in reproducing patterns of vertical stratification and mixing in the eastern Gulf may lead to further difficulties if biogeochemical tracers are added, in particular in the accurate simulation of spring bloom dynamics. The Baltic Sea is, for example, renowned for its blooms of cyanobacteria. Cyanobacterial growth is strongly related to temperature and it is encouraging to note that all the models perform reasonably well in simulating the magnitude and timing of maximum SST in the GoF.

The work presented here has highlighted the need for further model development in several areas in order to improve predictions of the hydrodynamics of the Baltic Sea. Accurate meteorological forcing is needed with a temporal resolution of for example 1 h and a spatial resolution of about 5×5 km. Descriptions of vertical mixing and advection terms should be improved, with particular emphasis on parameterizations specific to the Baltic Sea accounting for its pronounced salinity stratification in comparison with most of the world ocean. Movement towards higher vertical resolution should also be beneficial in accurately simulating the complex and changeable conditions of stratification that occur in the Baltic Sea. A better representation of bathymetry through the increase in both vertical and horizontal resolutions is needed to minimize discrepancies between real and model depths.

This modeling exercise demonstrated that the ensemble averaged results of the models showed no consistent over- or underestimation of hydrodynamic parameters. Thus, the ensemble average of the hydrodynamic models is accurate enough to be used as an input to biogeochemical and ecosystem models.

Conclusions

1. The present hydrodynamic models can reproduce the hydrographic conditions of the GoF reasonably well (temperature is simulated on average with an accuracy of ±1 °C and salinity with an accuracy of ±0.5‰) to give reliable forcing for ecosystem models.
2. There are still problems, especially with the description of vertical mixing. Sophisticated methods should be developed with particular relevance to the Baltic Sea because of the specific conditions of stratification that occur there.

3. There is still work to be done to refine the resolution of the models and to get more accurate initial, boundary and forcing functions for models of the Baltic Sea. Low vertical resolution in the model can lead to severe inaccuracies in the results.
4. According to skill tests of the models, the best model was HIROMB although no model was best/worst in all cases.

Acknowledgement: This study was carried out during the EUTROPHICATION-MAPS project. The funding organizations Nordic Council of Ministers and the Finnish Ministry of Environment are acknowledged. V. Ryabchenko, A. Isaev, R. Vankevich and I. Neelov were partly supported by the Federal Targeted Programme "Scientific and Scientific-Pedagogical Personnel of the Innovative Russia in 2009–2013". The AVHRR Oceans Pathfinder SST data were obtained through the online PO.DAAC Ocean ESIP Tool (POET) at the Physical Oceanography Distributed Active Archive Center (PO.DAAC), NASA Jet Propulsion Laboratory, Pasadena, CA. We also thank the anonymous referees for their constructive criticisms.

References

Alenius P., Myrberg K. & Nekrasov A. 1998. Physical oceanography of the Gulf of Finland: a review. *Boreal Env. Res*. 3: 97–125.

Alenius P., Nekrasov A. & Myrberg K. 2003. The baroclinic Rossby-radius in the Gulf of Finland. *Cont. Shelf Res*. 23: 563–573

Andrejev O. & Sokolov A. [Андреев О. & Соколов А.] 1989. [Numerical modelling of the water dynamics and passive pollutant transport in the Neva inlet]. *Meteorologia i Hydrologia* 12: 75–85. [In Russian].

Andrejev O., Myrberg K. & Lundberg P.A. 2004b. Age and renewal time of water masses in a semi-enclosed basin — application to the Gulf of Finland. *Tellus* 56A: 548–558.

Andrejev O., Myrberg K., Alenius P. & Lundberg P.A. 2004a. Mean circulation and water exchange in the Gulf of Finland — a study based on three-dimensional modeling. *Boreal Env. Res*. 9: 1–16.

Andrejev O., Myrberg K., Andrejev A. & Perttilä M. 2000. Hydrodynamic and chemical modelling of the Baltic Sea — a three-dimensional approach. *Meri* 42: 1–42.

Beckmann A. & Döscher R. 1997. A method for improved representation of dense water spreading over topography in geopotential-coordinate models. *J. Phys. Oceanogr*. 27: 581–591.

Bergström S. & Carlsson B. 1994. River run-off to the Baltic Sea: 1950–1990. *Ambio* 23: 280–287.

Berlyand M.E. [Берлянд М.Е.] 1956. [*Forecasting and control of the heat regime surface air layer*]. Gidrometeoizdat, Leningrad. [In Russian].

Bowen I.S. 1926. The ratio of heat losses by conduction and by evaporation from any water surface. *Phys. Rev*. 27: 779–787.

Brunt D. 1932. Notes on radiation in the atmosphere. *Q. J. Roy. Meteor. Soc*. 58: 389–420.

Bryan F.O., Kauffman B.G., Large W.G. & Gent P.R. 1996. *The NCAR CSM flux coupler*. Technical note TN-425+STR, NCAR.

Bumke K. & Hasse L. 1989. An analysis scheme for determination of true surface winds at sea from ship synoptic wind and pressure observations. *Bound. Lay. Meteor*. 47: 295–308.

DHI Water and Environment 2000. *MIKE 3. Environmental hydraulics*. DHI Software User Guide, Documentation and Reference Manual, Denmark.

Doney S.C. 1999. Major challenges confronting marine biogeochemical modeling. *Global Biogeochem. Cy*. 13: 705–714.

Fennel W., Seifert T. & Kayser B. 1991. Rossby radii and phase speeds in the Baltic Sea. *Cont. Shelf Res*. 11: 23–76.

Fujii K. & Obayashi S. 1989. High-resolution upwind scheme for vortical-flow simulations. *J. Aircraft* 26: 1123–1129.

Funkquist L. 2001. HIROMB, an operational eddy-resolving model for the Baltic Sea. *Bulletin of the Maritime Institute in Gdansk* XXVIII: 7–16.

Gill A.E. 1982. *Atmosphere–ocean dynamics*. International Geophysics Series, vol. 30, Academic Press, Orlando.

Hankimo, J. 1964. Some computations of the energy exchange between the sea and the atmosphere in the Baltic area. *Ilmatieteellisen keskuslaitoksen toimituksia* 57: 1–26.

Idso S.B. & Jackson R.D. 1969. Thermal radiation from the atmosphere. *J. Geophys. Res*. 74: 5397–5403.

Inkala A. & Myrberg K. 2002. Comparison of hydrodynamical models of the Gulf of Finland in 1995: a case study. *Environ. Modell. Softw*. 17: 237–250.

Iziomon M.G., Mayer H. & Matzarakis A. 2003. Downward atmospheric longwave irradiance under clear and cloudy skies: Measurement and parameterization. *J. Atmos. Sol.-Terr. Phy*. 65: 1107–1116.

Kattsov V.M. & Meleshko V.P. 2004. Evaluation of atmosphere-ocean general circulation models used for projecting future climate change. *Izv. Atmos. Ocean. Phy*. 40: 647–658.

Kennedy R.E. 1944. Computation of daily insolation of energy. *B. Am. Meteorol. Soc*. 30: 208–213.

Klein W.H. 1948. Calculation of solar radiation and the solar heat load on man. *J. Meteorol*. 5: 119–129.

Kochergin V. 1987. Three-dimensional prognostic models. *Coastal Estuarine Science Series* 4: 201–208.

Koponen J., Alasaarela E., Lehtinen K., Sarkkula J., Simbierowicz P., Vepsä H. & Virtanen M. 1992. Modelling the dynamics of a large sea area. *Publications of Water and Environment Research Institute* 7: 1–91.

Korpinen P., Kiirikki M., Koponen J., Peltoniemi H. & Sarkkula J. 2004. Evaluation and control of eutrophication in Helsinki sea area with the help of a nested 3D-ecohydrodynamic model. *J. Marine Syst*. 45: 255–265.

Lehmann A. 1995. A three-dimensional baroclinic eddy-

resolving model of the Baltic Sea. *Tellus* 47A: 1013–1031.

Lehmann A. & Hinrichsen H.-H. 2002. Water, heat and salt exchange between the deep basins of the Baltic Sea. *Boreal Env. Res.* 7: 405–415.

Lehmann A., Krauss W. & Hinrichsen H.-H. 2002. Effects of remote and local atmospheric forcing on circulation and upwelling in the Baltic Sea. *Tellus* 54A: 299–316.

Leppäranta M. & Myrberg K. 2009. *The physical oceanography of the Baltic Sea*. Springer-Verlag, Berlin-Heidelberg.

Liu A.J., Katsaros K.B. & Businger J.A. 1979. Parameterization of air–sea exchange of heat and water vapor including the molecular constraints at the interface. *J. Atm. Sci.* 36: 1722–1735.

Luyten P.J., Jones J.E., Proctor R., Tabor A., Tett P. & Wild-Allen K. 1999. *COHERENS — a coupled hydrodynamic–ecological model for regional and shelf seas: user documentation*. MUMM Report, Management Unit of the Mathematical Models of the North Sea.

Marciano J.J. & Harbeck G.E. 1954. *Mass-transfer studies*. Technical Report, USGS Professional Paper 2657, Lake Hefner studies.

Meier H.E.M. 1996. Ein regionales Modell der westlichen Ostsee mit offenen Randbedingungen und Datenassimilation. *Berichte* aus dem Institut für *Meereskunde* der *Universität Kiel* 284: 1–118.

Meier H.E.M. 2001. On the parameterization of mixing in three-dimensional Baltic Sea models. *J. Geophys. Res.* 106(C12): 30997–31016.

Meier H.E.M. 2003. Sensitivity of the Baltic Sea salinity to freshwater supply. *Climate Res.* 24: 231–242.

Millero F. & Kremling I. 1976. The densities of the Baltic Sea deep waters. *Deep Sea Res.* 23: 611–622.

Neelov I.A. [Неелов И.А.] 1982. [A mathematical model of eddies in the ocean]. *Okeanologiya* 22: 875–884. [In Russian].

Neelov I.A., Eremina T.R., Isaev A.V., Ryabchenko V.A., Savchuk O.P. & Vankevich R.E. 2003. A simulation of the Gulf of Finland ecosystem with a 3-D model. *Proc. Estonian Acad. Sci. Biol. Ecol.* 52: 346–359.

Nekrasov A.V. & Lebedeva I.K. 2002. Estimation of baroclinic Rossby radii in Luga–Koporye region. *BFU Research Bulletin* 4–5: 89–93.

Nekrasov A.V., Provotorov P.P., Korovin V.P., Lyakhin Y.I. & Chantsev V.Y. 2003. Hydrographic and hydro-chemical structure of waters in the Luga–Koporye region during the summer period. *ICES Cooperative Research Report* 257: 248–253.

Omstedt A., Elken J., Lehmann A. & Piechura J. 2004. Knowledge of the Baltic Sea physics gained during the BALTEX and related programmes. *Prog. Oceanogr.* 63: 1–28.

Passenko J., Lessin G., Erichsen A. & Raudsepp U. 2008. Validation of hydrostatic and non-hydrostatic versions of the hydrodynamic model MIKE 3 applied for the Baltic Sea. *Estonian Journal of Engineering* 14: 255–270.

Reed R.K. 1977. On estimating insolation over the ocean. *J. Phys. Oceanogr.* 7: 482–485.

Roe P.L. 1985. Some contributions to the modeling of discontinuous flow. *Lectures in Applied Mathematics* 22: 163–192.

Rosati A. & Miyakoda K. 1988. A general circulation model for upper ocean simulation. *J. Phys. Oceanogr.* 18: 1601–1626.

Sarkkula, J. (ed.) 1997. Proceedings of the Final Seminar of the Gulf of Finland Year 1996, March 17–18, 1997, Helsinki. *Suomen ympäristökeskuksen moniste* 105: 331–339.

Seifert T. & Kayser B. 1995. A high resolution spherical grid topography of the Baltic Sea. *Meereswissenschaftliche Berichte* 9: 72–88.

Simons T.J. 1980. Circulation models of lakes and inland seas. *Can. Bull. Fish. Aquat. Sci.* 203: 1–146.

Smagorinsky J. 1963. General circulation experiments with the primitive equations. *Mon. Weather Rev.* 91: 99–164.

Sokolov A., Andrejev O., Wulff F. & Rodriguez Medina M. 1997. The data assimilation system for data analysis in the Baltic Sea. *System Ecology Contributions* 3: 1–66.

Soomere T., Myrberg K., Leppäranta M. & Nekrasov A. 2008. The progress in knowledge of physical oceanography of the Gulf of Finland: a review for 1997–2007. *Oceanologia* 50: 287–362.

Soomere T., Dick S., Gästgifvars M., Huess V. & Nielsen J.W. 2004. *Project plan for implementation of interfacing between Baltic scale models to local (coastal area) models*. Report to the EU project PAPA (available at http://www.boos.org/papa/index.html).

Tamsalu R., Zalesny V.B., Ennet P. & Kuosa H. 2003. The modelling of ecosystem processes in the Gulf of Finland. *Proceedings of the Estonian Academy of Sciences. Biology, Ecology* 52: 332–345.

UNESCO 1981. *Tenth report of the joint panel on oceanographic tables and standards*. UNESCO Technical Papers in Marine Science no. 36, UNESCO, Paris.

Vested H.J., Justesen P. & Ekebjærg L.C. 1992. Advection-dispersion modelling in three dimensions. *Appl. Math. Model.* 16: 506–519.

Wulff F., Rahm L. & Larsson P. (eds.) 2001. *A systems analysis of the Baltic Sea*. Springer-Verlag, Berlin.

Zalesak S.T. 1979. Fully multidimensional flux-correced transport algorithms for fluids. *J. Comput. Phys.* 31: 335–361.

Zillmann J.W. 1972. *A study of some aspects of the radiation and the heat budgets of the southern hemisphere oceans*. Meteorological Studies 26, Bureau of Meteorology, Department of the Interior, Canberra, Australia.