Ocean Sci. Discuss., 12, 355–401, 2015
www.ocean-sci-discuss.net/12/355/2015/
doi:10.5194/osd-12-355-2015

Ocean Science
Discussions
Open Access

Evaluation of numerical models by FerryBox and Fixed Platform in-situ data in the southern North Sea

M. Haller[1], F. Janssen[2], J. Siddorn[3], W. Petersen[1], and S. Dick[2]

[1]Institute of Coastal Research, Helmholtz-Zentrum Geesthacht GmbH, Max-Planck-Str. 1, 21502 Geesthacht, Germany
[2]Bundesamt für Seeschifffahrt und Hydrographie, Bernhard-Nocht-Straße 78, 20359 Hamburg, Germany
[3]Met Office, Exeter, UK

Received: 14 January 2015 – Accepted: 5 February 2015 – Published: 27 February 2015

Correspondence to: M. Haller (michael.haller@hzg.de)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

FerryBoxes installed on ships of opportunity (SoO) provide high-frequency surface biogeochemical measurements along selected tracks on a regular basis. Within the European FerryBox Community, several FerryBoxes are operated by different institutions. Here we present a comparison of model simulations applied to the North Sea with FerryBox temperature and salinity data from a transect along the southern North Sea and a more detailed analysis at three different positions located off the English East coast, at the Oyster Ground and in the German Bight. In addition to the FerryBox data, data from a Fixed Platform of the MARNET network are applied. Two operational hydrodynamic models have been evaluated for different time periods: results of BSHcmod v4 are analysed for 2009–2012, while simulations of FOAM AMM7 NEMO have been available from MyOcean data base for 2011 and 2012. The simulation of water temperatures is satisfying; however, limitations of the models exist, especially near the coast in the southern North Sea, where both models are underestimating salinity. Statistical errors differ between the models and the measured parameters, as the root mean square error (rmse) accounts for BSHcmod v4 to 0.92 K, for AMM7 only to 0.44 K. For salinity, BSHcmod is slightly better than AMM7 (0.98 and 1.1 psu, respectively).

The study results reveal weaknesses of both models, in terms of variability, absolute levels and limited spatial resolution. In coastal areas, where the simulation of the transition zone between the coasts and the open ocean is still a demanding task for operational modelling, FerryBox data, combined with other observations with differing temporal and spatial scales serve as an invaluable tool for model evaluation and optimization. The optimization of hydrodynamical models with high frequency regional datasets, like the FerryBox data, is beneficial for their subsequent integration in ecosystem modelling.

OSD
12, 355–401, 2015

1 Introduction

The North Sea is a coastal ocean that has one of the highest densities of ship traffic in the world. It is an economically important coastal region, sustaining commercial fisheries, wind farming, oil production and tourism (Kannen, 2012; OSPAR, 2010). A major part of the North-West European continental shelf, the North Sea has a mean depth of 90 m. Bathymetry varies, and while the southern part is shallow (15–50 m), the northern part deepens to 100–200 m, in the Norwegian Trench to well below 200 m (see Fig. 1). The southeast part of the North Sea is known as the German Bight with the Wadden Sea at its coastal margins. Because of freshwater inflow from several rivers in the Southern North Sea (e.g. Rhine, Maas, Elbe), salinity near the coasts is only about 15–25, and increases to 35 in the central parts of the North Sea (Janssen et al., 1999; OSPAR, 2000). The general circulation of North Sea waters is strongly governed by several factors like bathymetry, density distribution and wind stress (Queste et al., 2013). An anti-clockwise circulation dominates the North Sea with North Atlantic water entering the North Sea at its north-west boundary near the Shetland Islands, travelling along the Scottish and English coast and leaving along the Norwegian coasts (Turrell, 1992), see Fig. 2. Only a minor fraction of the North Atlantic water entering from the north reaches the southern North Sea, while the greater part circulates north of the Dogger Bank (Thomas et al., 2005). A much smaller flow of North Atlantic enters through the Dover Strait, and travels to the east Baltic Sea, where the waters are entrained into the North Sea water through the Skagerrak. This water inflow has a significant influence on southern North Sea water. The relatively salty English Channel water is mixed on its way along the south east way of the North Sea coasts with freshwater from several rivers, passes the German Bight and enters the Norwegian Trench region, mixing with the northern branch of the North Sea circulation. The residence time of North Sea water has been estimated to be less than one year (Jickells, 1998; Lenhart and Pohlmann, 1997; Thomas et al., 2003). This residence time is shorter than

OSD
12, 355–401, 2015

the decade-long residence time of the Baltic Sea, but it is longer than other open shelf seas like e.g. the South Atlantic Bight at the US East Coast (Jickells, 1998).

Given the importance of the North Sea to the European economy and to the coastal communities, it is vital to monitor and understand its current ecological state. The FerryBox system is a valuable tool to provide regular high-frequency scientific measurements of ecologically important parameters, including temperature and salinity. A FerryBox is an operational measurement system installed on ships of opportunity (SoO), as well as on fixed onshore stations near harbours, river banks or estuaries. It is a flow-through system that continuously measures biogeochemical parameters every 10 s. FerryBoxes are operated in European coastal regions like the North Sea, Baltic Sea, and Norwegian Sea and some of them cover a time period of more than 20 years. FerryBoxes are a valuable platform to test and operate new developed oceanographic sensors in a sheltered environment (e.g. ship or container) without limitation of energy. However, the operation of mobile FerryBoxes is strongly dependent on shipping companies and their decisions for operation of a certain shipping line. Developed 20 years ago, the FerryBox system was launched on a broader European basis in the FerryBox project.

During the FerryBox project from 2002–2005 (FerryBox, 2014; Petersen et al., 2005), cooperation between several national oceanographic institutions was launched which also targeted developments of new sensors and observing strategies and best practices in quality control, maintenance and biofouling prevention (Hydes et al., 2009; Petersen et al., 2005, 2007).

Shelf seas are complex regions governed by many processes. Along with the operational monitoring using in situ and satellite observing systems, numerical simulation has long been acknowledged to be important for understanding the hydrodynamics of coastal regions. The first attempts in numerical modelling started in the 1960s, with the development of 2-D barotropic models, which then were forced by constant winds (Pohlmann, 2006). Since the 1980s, baroclinic 3-D-models have been developed to prognostically model water temperature and salinity variations (Pohlmann, 2006). For

OSD
12, 355–401, 2015

this study, two different hydrodynamic models, BSHcmod and FOAM AMM7 NEMO, were used. These models are commonly applied, e.g. for ecosystem modelling (Maar et al., 2011) and predicting wave-tide-current interactions (Pleskachevsky et al., 2009) in the North Sea. Other models have been used for different applications for the North Sea, e.g. the General Estuarine Transport model GETM (Burchard, 1998; Burchard and Bolding, 2002), the HAMSOM model (Backhaus, 1985) and the COHERENS model (Luyten et al., 1999), but are not included in this study.

The German Federal Maritime and Hydrographic Agency (Bundesamt für Seeschifffahrt und Hydrographie, BSH) developed the hydrodynamic model BSHcmod for operational use in the North and Baltic Sea (Dick et al., 2001). The coupled Forecasting Ocean Assimilation Model (FOAM) consists of a hydrodynamical (O'Dea et al., 2012) and an ecosystem (Edwards et al., 2012) part. The hydrodynamics are provided by the Nucleus for European Modelling of the Ocean (NEMO, Madec, 2008), while the ecosystem part is supplied by the European Regional Seas Ecosystem Model (ERSEM, Baretta et al., 1995; Blackford et al., 2004). The FOAM is a regional model, nested to the UK Met Office global ocean model (Blockley et al., 2014).

Besides the FerryBoxes, several other measurement networks are available in the North Sea, including the coastal observing system COSYNA (COSYNA, 2014; Grayek et al., 2011; Riethmuller et al., 2009; Stanev et al., 2011). Besides COSYNA, other observation networks like MARNET (BSH, 2014) and the SmartBuoys network (Cefas, 2015; Mills et al., 2003), measure temperature and salinity on buoys and other fixed platforms. Satellite imagery is somewhat limited regarding the time resolution and restricted to certain parameters. Moreover, satellite coverage is limited in coastal regions and in the vicinity of land as well as by cloud coverage, e.g. when using visible parts of the spectrum (Petersen et al., 2008; Volent et al., 2011).

FerryBox data may bridge the gap between the existing in-situ observations typically used for data assimilation, as they provide reliable and high resolution in-situ data for transects in the North Sea (Petersen et al., 2008). However, the influence of these transect data is limited to grid points along a transect. To overcome this limitation,

OSD
12, 355–401, 2015

Wehde et al. (2006) and Petersen et al. (2011) applied a water transport model for comparison of FerryBox measurements with other operational observations. The aim of this study is to compare numerical model data with in-situ measurements of different monitoring systems (FerryBox, fixed platforms). The goal of this study was to evaluate how reliable modelled water temperature and salinity data are in different areas of the North Sea and identify the limitations and weaknesses of the operational models AMM7 and BSHcmod v4.

The first section of our study describes the data sets and the applied methods. Then, data from a complete transect data set are compared with model results. Point comparisons of model data and observations are then presented.

2 Materials and methods

2.1 FerryBox system

In general, all European FerryBox systems have a similar design. The differences are in the design of the flow-through system, the degree of automation and biofouling prevention as well as the possibilities of supervision and remote control. The FerryBox systems designed and manufactured for operation by 4H-Jena engineering GmbH and Helmholtz-Zentrum Geesthacht (HZG) have specifications as follows:

The water is pumped from a subsurface inlet (in 2–7 m depth, depending on the ship geometry) into the measuring circuit of multiple sensors. A debubbling unit removes air bubbles, which may enter the system during heavy seas. At the same time coarse sand particles are removed as well. Coupled to the debubbler, there is an internal water loop in which the seawater is circulated with a constant velocity of about 1 $m\,s^{-1}$. This already decreases the tendency for building biofilms on sensors and tube surfaces. For a reliable unattended operation the system is supervised by a computer which can shut-off the system in case of severe errors and operates automatic cleaning cycles, e.g. in harbour. At certain positions along the transects, water samples are collected

OSD
12, 355–401, 2015

in an automated cooler sampler for subsequent laboratory analyses. Measurement locations are evaluated via GPS positioning.

Biofouling is prevented by cleaning of the sensors with tap water and rinsing with acidified water (using sulfuric acid) in the harbour after each cruise. All flow rates are supervised by the system. More information about the FerryBox system can be found e.g. in Petersen et al. (2003, 2011).

For this study, the data sets of two different commercial ships have been used. The data are available at FerryBox database at Helmholtz-Zentrum Geesthacht (HZG) (http://ferrydata.hzg.de). In Fig. 1, the transect of the *Tor Dania* (which was in service until April 2012), from now on referred to as TD, is shown. TD travels on the route between Cuxhaven (GER) and Immingham (GB) every 2–3 days with an average cruising speed of 12 kn. The temporal resolution of TD measurements is 10 seconds. The cargo vessel *Lysbris* (IMO number 9144263, from now called herein LB) operates on a route going along the coasts of the North Sea (see Fig. 1).

For water temperature and salinity measurements in the FerryBox system, the thermosalinograph Citadel TS-NH, manufactured by Teledyne Technology Company is used. The basic salinity instrument measures inductive conductivity, the according temperatures are measured by a thermistor in close proximity. The resulting practical salinity is measured with an accuracy of ±0.015 psu and temperatures are accurate to ±0.005 K. For validation of the Citadel sensor, water probes taken on board are analysed in the laboratory. The common laboratory sensor, the Autosal Guildline Salinometer 8400B, has been used until 2012, when the lab validation procedure of the FerryBox salinity sensor has been changed to OPTIMARE precision Salinometer system. The accuracy is declared to be ±0.002 psu. The Optimare is accurate to ±0.003 psu, verified in laboratory tests with OSIL standard in 2012. In this study, salinity values are quoted in dimensionless numbers, even though FerryBox measurements are performed in practical salinity units.

OSD
12, 355–401, 2015

2.2 MARNET observing system

The MARNET station network consists of several measurement sites in the German coastal parts of the Baltic Sea and the North Sea (BSH, 2014). It is also part of the COSYNA observing system in North and Arctic Seas. MARNET has a long tradition of monitoring of coastal waters and is operated by BSH. This study uses data from the North Sea MARNET station "Deutsche Bucht" (German Bight), located west of the island Helgoland. It is an unmanned light vessel, located at position 54°10′ N, 07°27′ E in the German Bight. The observations at that station started in 1989 and encompass oceanographic observations in seven depths (3, 6, 10, 15, 20, 25, and 30 m) and meteorological measurements in 14 m height. The data from 6 m depth are used, where water temperature, salinity, oxygen and density are measured. The data are provided in hourly time resolution. For measurement of water temperature and salinity in 6 m depth, a CTD SBE 37-SIP MicroCAT, manufactured by Sea-Bird Electronics Inc. The probe provides water temperatures with accuracy of ±0.002 K and conductivity values with an accuracy of ±0.0003 mS.

2.3 BSHcmod v4

The BSH model is an operational, three-dimensional hydrodynamic ocean circulation model for the North Sea and Baltic Sea (Dick et al., 2001), called BSHcmod v4. The Reynolds-averaged Navier–Stokes equations are discretized on a geographical Arakawa-C grid and on z coordinates with 36 vertical layers. In the German Bight, only 25 vertical layers are used. Internally, BSHcmod v4 is calculating on layers with variable thickness, depending on the tides (in the English Channel 8 m, in the German Bight 1–2 m thickness). Archived BSHcmod data have been interpolated on a coarser and constant vertical layer resolution. Thus, the data applied here are from the upper layer which has a thickness of 5 m. The horizontal grid size differs in the North Sea between 900 m in the German bight and the western part of the Baltic Sea (focus region) and 5 km in the other parts of the North and Baltic Sea. The temporal resolution

OSD
12, 355–401, 2015

of BSHcmod v4 is 15 min. The BSHcmod v4 is in operation since 2009 and contains general vertical coordinates. On a daily basis, it predicts water levels, currents, water temperatures, salinity and ice cover for the next three days ahead. For the German Bight the high variability of daily surface circulation patterns as well as its statistical distribution is presented at BSH (2015). The monthly mean surface circulation for the whole North Sea as simulated by the BSH circulation model is published in several reports e.g. at Loewe (2009) and Loewe et al. (2013), showing a pronounced seasonal as well as inter-annual variability strongly related to the atmospheric circulation pattern over the North Sea. This has also been reviewed in detail by Otto et al. (1990).

Meteorological and wave forecasts are provided by German Weather Service (Deutscher Wetterdienst, DWD) (Doms and Schättler, 1999). The 10 m wind components are extrapolated from the lowest pressure level height data, considering also the stability conditions in the Prandtl layer. The input of freshwater into the North Sea by River runoff is daily averaged data for 5 German rivers, derived from river measurements. For remaining rivers, freshwater runoff climatologies are used. The BSH model simulates the tides based on 14 harmonic constituents which are provided at the open boundaries in the northern part of the North Sea and the western part of the English Channel. Falling dry and flooding of tidal flats is also taken into account.

2.4 FOAM AMM7 NEMO

Besides the BSHcmod v4, also results from the AMM7 system have been evaluated. The AMM7 includes a three-dimensional hydrodynamic component based upon the NEMO model, which is included as part of the Met Office Forecasting Ocean Assimilation Model (FOAM) suite of forecast systems that run daily operationally and include the assimilation of observations (Siddorn et al., 2007). The AMM7 system also contains the ecosystem model ERSEM. The model domain encompasses the European North-West continental Shelf on a regular lat-lon grid resolved on a 1/15° (lat) times 1/9° (lon) grid. To get the correct vertical resolution of the terrain, hybrid s-sigma terrain following coordinates are applied with 50 equally spaced levels. The NEMO model itself is a community model particularly developed in Europe (Madec, 2008). Though it has been developed for the deep ocean, it has then been modified for usability for shelf seas. Details of the model and its implementation are given in O'Dea et al. (2012). The model assimilates in situ and satellite SST using an Optimal Interpolation scheme (Martin et al., 2007). At the open boundaries, AMM7 is one-way nested into the Met Office operational FOAM 1/12° deep ocean model (Storkey et al., 2010). Additionally, the climatological river runoff of more than 300 rivers all around the North-West shelf is included as well as the input from the Baltic Sea. The atmospheric forcing is provided by the Met Office Numerical Weather Forecast model. For this study, AMM7 data set is provided by MyOcean data base (Myocean, 2011) in hourly time resolution and 7 km grid resolution. Data are taken from the surface layer.

2.5 Statistical measures

For the analysis of model performance a variety of statistical measures is commonly applied which will be here shortly introduced.

If the observations are denoted as obs and model predictions as sim, the bias can then described as the mean difference between simulations and observations, i.e.

$$\text{bias} = \overline{\text{sim}} - \overline{\text{obs}}.$$

Thus, negative (positive) bias means model underestimation (overestimation). The standard deviation of error is calculated by

$$\text{stde} = \sqrt{\overline{\left(\text{sim} - \overline{\text{sim}} - (\text{obs} - \overline{\text{obs}})\right)^2}}.$$

The root mean square error (rmse) is then calculated from of bias and stde, namely

$$\text{rmse} = \sqrt{(\text{bias})^2 + (\text{stde})^2},$$

OSD
12, 355–401, 2015

and the skill variance (skvar) is the ratio of standard deviation σ of simulation and observation, i.e.

$$\text{skvar} = \frac{\sigma_{\text{sim}}}{\sigma_{\text{obs}}}.$$

Besides the above described statistical measures, also the index of agreement (IOA) or skill is used in the statistical analysis of this study. The IOA has been first described by Willmott (1981) and it is described as

$$\text{IOA} = 1 - \frac{\sum(\text{obs} - \text{sim})^2}{\sum\left(|\text{sim} - \overline{\text{obs}}| + |\text{obs} - \overline{\text{obs}}|\right)^2}.$$

It is a standardized measure of the degree of model prediction error and varies between 0 and 1. A value of 1 indicates a perfect match, and 0 indicates no agreement at all. The index of agreement can detect additive and proportional differences in the observed and simulated means and variances; however, it is overly sensitive to extreme values due to the squared differences (Legates and McCabe, 1999).

The cost function (cf) field, introduced by Berntsen and Svendsen (1999) and later adapted by Søiland and Skogen (2000) is a measure for discrepancies of parameter F between model and observations, normalized by the standard deviation of the observations $F_{\text{obs}}^{\text{SD}}$, i.e.

$$\text{cf} = \frac{F_{\text{obs}} - F_{\text{mod}}}{\max\left(F_{\text{obs}}^{\text{SD}}, F_{\text{min}}^{\text{SD}}\right)},$$

where $F_{\text{min}}^{\text{SD}}$ denotes the minimum allowed amount of the standard deviation, which then prevents cf to go into infinity. The cost function is the mean of the absolute cost function values of the field the analysis has been applied to. A value of 0.5 means that the model is on average 0.5 times the standard deviation off the observations.

OSD
12, 355–401, 2015

The benefit of fixed platform time series is the high temporal resolution they provide operationally. This has been used for evaluating the model performance in regard of spectral densities. This has been done via Fourier analysis of time series of MARNET and both involved model simulations of BSHcmod and AMM7, respectively.

2.6 Methods

The southern North Sea has different regions with different characteristics. To take that into account, three positions for detailed investigation have been selected for the time period of 2006–2013 (Fig. 1):

i. English coast Point at 53.553° N 0.241° E (p1).

ii. Oyster Ground point at 54.04° N 5° E (p2).

iii. German Bight Point 54.17° N 7.45° E (p3).

Position p1 is situated near to the coast and not far from the mouth of the river Humber. It is dominated by mainly southerly flowing relatively cold Scottish coastal water current, originating from the North Atlantic (OSPAR, 2000), see Fig. 2.

The second point (p2) is located near the Oyster ground, a region with water depths up to 40 m. It is thermally stratified in the summer season, in contrast to the well-mixed shallow southern North Sea. Due to spring algae bloom, the stratification in summer leads to low oxygen concentration (hypoxia). This has been described in Queste et al. (2013) Together with salty water (> 35 psu) from the English Channel, frontal zones are forming in this region, as has been observed e.g. by FerryBox measurements (Petersen et al., 2011).

The German Bight area, where point (p3) is situated, is governed by the Continental Coastal Water current, the input of freshwater, and the complex exchange processes between the Wadden Sea and the North Sea. It is also the region with one of the highest tidal amplitudes of the North Sea (> 4 m).

OSD
12, 355–401, 2015

From the FerryBox database, time series for these points have been extracted, adopting a search radius for ship measurements of up to 10 km around the respective positions. Then the BSHcmod v4 and FOAM AMM7 NEMO model data have been interpolated on these positions, using the nearest-neighbour approach. For BSHcmod, data from the surface box down to 5 m depth have been taken, with instantaneous grid values at 15 min time resolution. AMM7 is taken from the surface box of the model and has instantaneous values for a 7 km grid box mean every hour. As the model uses an S-coordinate vertical discretisation, the depth of these data varies as a function of the total water depth from about 10 cm in the shallowest waters to approximately 1 m in the deeper parts of the southern North Sea.

For the evaluation of the complete transect between UK and Germany, FerryBox data from 5 m depth of the complete transect with time resolution of 10 seconds have been sampled on a longitudinal grid with intervals of 0.02° length for the time period of three years. FerryBox data from 5 m depth of the complete transect with time resolution of 10 seconds have been sampled on a longitudinal grid with intervals of 0.02° length. So, for each position of this track, a time series with hourly resolution has been created. In a two-fold algorithm loop the according simulation values of both models have been interpolated with nearest-neighbour approach.

So, for each position of this track, a time series with hourly resolution has been created. In a two-fold algorithm loop the according simulation values of both models have been interpolated with nearest-neighbour approach.

2.7 Validation of FerryBox data

Before model evaluation results of our study are presented, the reliability of our measurements has been analysed with cross-checks and laboratory comparison of water samples.

For validation of FerryBox salinity measurements, a control check with bottle samples has been performed. On both ships – TD and LB – water samples have been taken at predefined positions and analysed at the institute's lab. For FerryBox water

OSD
12, 355–401, 2015

temperature measurements, it is not feasible to compare them with water samples that will be analysed later in the laboratory, so instead a cross-check between the TD FerryBox and MARNET observations has been done. In future, it would be suggestive to cross-check the FerryBox instruments with an additional certified temperature probe on board.

In Fig. 3, comparisons of FerryBox measurements and lab analyses of salinity for TD and LB are shown. The water samples are taken regularly along the respective FerryBox transect from 2007 to 2011 and from 2009 to 2012, respectively. The data correspond in both cases very well, only few outliers are observed. Note the different scales of salinity in the graphs. In case of LB, a higher range of salinity values is covered. This is due to the included FerryBox route section in the Elbe river estuary up to Hamburg port. The correlation is well above 0.9 (0.96 for TD and 0.99 for LB), which indicates a high reliability of FerryBox salinity measurements. The RMSE is for LB salinity regression model slightly lower than for TD, which may be suggesting a lower uncertainty because of a higher range of salinity values.

For the evaluation of water temperature reliability, a comparison has been done for the German Bight region checking the correspondence of MARNET measurements and FerryBox observations. TD passes every second day the MARNET position (p3) on its way between Cuxhaven (GER) and Immingham (UK). Only data of TD, which have been recorded less than 10 km away, have been used. Data are available from 2007 until 2011. So, 5 annual cycles could be analysed.

For both parameters, a good agreement can be observed in Fig. 4. The scatterplot for water temperatures on the left confirms the good agreement. There, TD measurements are higher than corresponding MARNET observations. Linear regression shows an offset of 0.33 K, the mean bias (FerryBox-MARNET) accounts to 0.37 K. The stde of the FerryBox in regard to the “truth” that MARNET provides accounts to 0.42 K. It is believed that the bias is due to the relatively long way of the water pumped from outside into the FerryBox whereas MARNET temperature sensors are in close contact to the

OSD
12, 355–401, 2015

water body. Subsequently, FerryBox water temperature measurements have been bias-corrected by a simple additive correction method (Sperna Weiland et al., 2010).

Grayek et al. (2011) also compared FerryBox data to MARNET observations and the OSTIA satellite data package (Donlon et al., 2009) and found similar agreement between the temperature data sets.

The time series of salinity are also in good agreement (Fig. 4). However, the figure shows a higher scattering than for water temperature, with MARNET station observing higher values. The standard deviation of the difference accounts to 0.57 psu; the determination coefficient accounting to 0.82 is, thus, not as high as for water temperatures. So, agreement of salinity observation between FerryBox and MARNET is good; however, the influence of time lapse and displacement of up to 10 km together with the high influence of tides and river discharge of low salinity water may explain lower correlation. A further analysis shows that this stands not only for the one station *Deutsche Bucht* but also for the MARNET station *Ems* where correlation coefficient for salinity accounts to about 0.9 (not shown).

All in all, a high reliability of different FerryBox sensor observations as well as high agreement between different measurement systems (FerryBox and MARNET) can be ensured. Both analysed FerryBox parameters, water temperature and salinity, are well-suited for comparison with model data, which will be described in the next section.

3 Results and discussion

3.1 Transect comparisons for southern North Sea

Together with model output of BSHcmod v4 and AMM7, the complete TD transect between Germany and England has been analysed regarding differences of simulated and observed water temperatures and salinity.

OSD
12, 355–401, 2015

3.1.1 Water temperature

In Fig. 5, the differences of water temperatures from 2009–2011 for BSHcmod (left) and from 2011 and 2012 for AMM7 (right) are shown. Note the different time scales of the model comparisons in both figures. Positive (negative) differences indicate too high (too low) simulated temperatures. The differences have been marked according to the double stde of the FerryBox data which has been described above. Thus, values beyond ±0.84 K for water temperatures and beyond ±0.8 for salinity are statistically significant. Gaps in the data mainly in the years 2009 and 2010 are due to FerryBox malfunction.

BSHcmod

On first glance, water temperature differences range around ±1.5 K for BSHcmod. Several spacial aspects can be determined in combination with Fig. 6, which shows the temporal mean along the transect of bias, standard deviation of error (stde), root mean square error (rmse) and skill variance (skvar). In Fig. 5, differences are mainly positive in winter months and mainly negative in the summer months. The bias in Fig. 6 ranges around −1 and +0.4 K, the stde ranges around 1 K. Thus, the rmse is around 1 K. The mean bias for the whole transect is at −0.02 K (Table 1), the mean rmse accounts to 0.92 K. The skill variance evaluates the model's ability to reproduce the variability of the data. The optimal value is 1. In Fig. 6, the overall skvar varies between 0.7 and 1.2.

Close to the English coast, temperature differences are systematically negative, dropping down to −2 K, showing the largest negative bias values for the transect. In Fig. 6 the corresponding bias is also at −2 K for this region. The stde also reaches 2 K and the skvar has its minimum accounting to 0.8. From there more eastwards between 0.5 and 2° E, biases range around 0.5 K, stde is at 0.8 K and the rmse at 1 K. Figure 5 indicates there is a seasonal cycle in the bias, with differences in all three years in late summer around 1° E dropping below −1 K, and small positive normalized differences in the winter. This seems to be a systematic error in the model, caused probably by

OSD
12, 355–401, 2015

too weak vertical mixing in the Scottish coastal water current in this time of year or it could mean that the flow of (warmer) Atlantic water is overestimated by the model. It should be noted that the TD transect crosses the southern North Sea approximately along the transition zone between the stratified and well mixed regions of the southern North Sea and therefore small errors in the position of the seasonal front will cause biases in this region. While in the central parts of the transect the differences as well as the according measures bias, stde and skvar show average values, in the German Bight east of 6° E stde is above the average amounting to 1 K, that being confirmed by higher differences showing in Fig. 5. The skvar reveals an overestimation of simulated water temperature variability near the German coast.

AMM7

Results for AMM7 (Fig. 5, right) show general good agreement to FerryBox observations for 2011 to 2012, as the bias for the whole transect amounts to 0.19 K (Table 1). However, some weaknesses are also revealed in AMM7 simulations of water temperatures off the English East coast near 0.5° E and in the German Bight in 2011. The differences rise to as high as 2 K at the English coast and in the German Bight from −1 to +1 K, depending on the seasons (overestimation in summer, underestimation in winter).

The statistical measures for AMM7 are present in Fig. 6, underlying the results in Fig. 5: stde and bias show two local extreme positions; near the English coast and in the German Bight. The skvar is around 1 or slightly above, reflecting a good model performance for water temperature variability.

Overestimation of water temperatures near the English coast in 2011 at around 0.5° E indicate that the FerryBox observed a drop of temperature in this area of around 1–2 K in respect to the surrounding areas that AMM7 did not catch entirely. In this region, the cooler Scottish coastal water current extends towards the south which seems to be underrepresented in AMM7. For the German Bight, one could argue, that the coarse resolution of 7 km is not sufficient enough to reflect the highly variable temper-

OSD
12, 355–401, 2015

ature field in this complex area. While bias is around zero, the stde peaks to over 1 K, reflecting the seasonal dependence of differences for AMM7 at German Bight.

3.1.2 Salinity

As for water temperatures, the error of simulated salinity of BSHcmod and AMM7 has been calculated for the whole transect and shown in Fig. 7. Positive (negative) values show too high (too low) simulated salinity values. For both models, differences can be divided in three sectors all over the TD transect. Both coastal zones (English east coast and German Bight) are dominated by high negative differences whereas in the central parts absolute differences are significantly lower, negative for AMM7 in most parts, positive for BSHcmod. However, they are not significant as they are lower than twofold std of FerryBox.

BSHcmod

For BSHcmod, positive differences occur in the western part of the transect between 0.5 and 5° E, meaning an overestimation of salinity by the BSHcmod v4 model. They range between 0 and 1. While the underestimation of BSHcmod v4 at the western transect part is restricted close to the English coast, negative values reach in the German Bight until 6° E. For the coastal parts of the transect, also the bias of BSHcmod in Fig. 8 is negative and the stde increases above the mean value of 0.68, whereas for the central part of the transect, the bias accounts to around 0.3, the stde to only 0.2. The mean salinity bias for BSHcmod is slightly negative (−0.17, see Table 1). The BSHcmod salinity variability is well below 1 in the western part from 0 to 3° E, having a minimum of only 0.15 at 0.5° E; the mean of this sector accounts to 0.52. East of this sector, the Skill variance varies between 0.5 and 1.5.

OSD
12, 355–401, 2015

AMM7

The salinity is generally underestimated by AMM7, except for the region between 3 and 6.5° E from April to June 2011. Near the coasts (west of 1° E and east of 7.5° E the differences are significant. Between 5 and 7.5° E, differences are only significant from September to December 2011. The mean bias of AMM7 accounts to −0.89; near the coasts, the bias is higher, accounting up to −3. This also holds for the stde, which exceeds 2 psu near the coasts. In the central parts of the transect, the bias (stde) shows low variation and is between −1 and 0 (between 0 and 1). The AMM7 salinity skvar is between 0.3 and 1.2 over the whole transect, so in total no spatial dependences could be found.

A combination of several factors seems to be responsible for the underestimation of salinity in the German Bight for both models. First of all the runoff from river Elbe and thus the freshwater input into the region seems to be overestimated, although in BSHcmod v4 daily averaged runoff rates of German rivers are included. For AMM7, climatological runoff is provided. An underestimation of vertical mixing in the BSHcmod v4 simulation possibly contributes to the underestimation of the salinity by mixing bottom water with higher salinity into the top layer sampled by the Ferrybox systems. Also, a numerical problem in the nesting scheme of BSHcmod cannot be ruled out as a possible reason at the moment. The western boundary of the high resolution grid nested into the coarse North Sea grid is located at 6°10′25″ E which coincides with the boundary of the region with underestimation in salinity. Further studies of vertical (and horizontal) mixing as well as investigations of the interactive coupling scheme have to be carried out. AMM7, on the other hand, is more limited near the coast in terms of special resolution than BSHcmod. The combination of poor representation of the river inputs along the German coastline with relatively coarse resolution and no representation of the wetting-and-drying limits the AMM7 model in these regions.

OSD
12, 355–401, 2015

3.2 Long-term measurement time series

In this section, time series of measurements and model simulations for the time period of 2009 to 2012 are presented. The observations have been recorded by the FerryBox of TD and LB. To address the different results along the transect between UK and Germany that have been described in the previous sections, three single positions in the southern North Sea have been selected, as it was mentioned above (see also Fig. 1). Two of them are located near the coasts of England and Germany, respectively.

3.2.1 English East Coast

The time series of English Coast point at 53.553° N 0.241° E (p1) for the difference of water temperatures are shown in Fig. 9, encompassing the years 2009 to 2011. The figure contains FerryBox data of TD and LB as well as model data of BSHcmod v4 and AMM7. However, the time series of the different sources contain some gaps. The time series of TD are present from 2009 to 2012, but shows some data gaps in 2009. The time series of LB generally has many gaps, because the vessel is only every two weeks at same positions. AMM7 data are available only from 2011 to 2012.

It can be seen in Fig. 9 (upper) that both models show similar behaviour, except the bias. The bias of AMM7 temperature accounts to 0.43 K, which is surprising as this model assimilates SST and in most other evaluations of the SST against in situ observations the bias has been an order of magnitude smaller. For example MyOcean (2011) documents a bias in the southern North Sea as a whole of 0.02 K, and for a buoy in the German Bight of 0.01 K. The BSHcmod v4 bias is below zero, accounting to −0.28 K. The temperature variability is well matched by both models, as the simulated std is nearly the same as the observed (skvar around 1). In 2011, the seasonal variation in sea surface temperature of 13 K was observed (not shown). This is also simulated by AMM7; however, BSHcmod v4 slightly overestimated the winter low water temperatures in January 2011 and underestimated the summerly temperature maximum, resulting in a positive differences in winter and negative differences in summer. This can also be

OSD
12, 355–401, 2015

observed in 2010. Though, the IOA of BSHcmod v4 is at 0.99, as well as for AMM7. This is also visually demonstrated by the high level of agreement shown in the scatterplot of Fig. 9 (upper right).

Results of comparison between salinity observations and simulations for East England Coast are shown in Fig. 9 (lower left), statistical measures in Fig. 9 (lower right). In the time period of 2009–2012, observations range between 30 and 35, resulting in a mean value of 33.03. Some low salinity events occur below 30, mainly in winter months (not shown). These low salinity events are not entirely reproduced by BSHcmod in 2010 and 2011, resulting in high positive differences. AMM7 results starting in 2011 give salinity values between 30 and 34 (mean at 32.39), with a bias at −0.99. The mean FerryBox salinity for the AMM7 period is at 33.38. The skvar is 1.19, which is better than BSHcmod which has a skvar of 0.45. The BSHcmod v4 salinity values range around 33.67 and does not capture the high variability seen in the observations, with the variation mainly showing oscillatory changes as would be expected from water mass movements due to tidal fluctuations in the English coastal waters. The IOA as well as the correlation coefficient is slightly higher for BSHcmod than for AMM7, but in some cases it may be right for the wrong reasons. Agreement is achieved when observed salinity happens to be in same range than tidally varying model values; otherwise there is no agreement. Concluding, BSHcmod v4 results show too salty water, AMM7 results are too fresh. This is also pictured by the different sign of the cost function (cf) results (negative for BSHcmod v4, positive for AMM7).

The general lack of agreement in both models can be for the most part explained by the model forcing concerning fresh water discharges. For most rivers entering the North Sea and the Baltic Sea BSHcmod uses either measurements of current river runoff or hydrological predictions. However BSHcmod uses constant climatological, for British rivers in the operational model, explaining the lack of seasonality. Therefore, the model is presently not able to simulate the high fluctuations observed. The AMM7 model also uses climatological runoff data for British rivers, but monthly variations are included and this is visible in the results of Fig. 9 (lower left).

3.2.2 Oyster Ground

The second position is situated in the southern North Sea off the Dutch coast in the Oyster Ground area, in Fig. 1 labelled as p2. TD travels along the German and Dutch coasts to England and back. In Petersen et al. (2011), this point has been previously selected for analysis of low saline water of fluvial origin, which have been observed by two FerryBox transects crossing at this point. Model simulations of BSHcmod v4 are available for the time period of 2009–2012, for AMM7 from 2011–2012.

BSHcmod and AMM7 simulations of water temperatures are in line with observations for most parts, getting the annual cycle and the amplitude in the correct phase in their respective time period resulting in mainly good agreement except of several peaks high differences around ±1.5 K (Fig. 10, upper left). Agreement is visible also in Fig. 10 (upper right). Statistical measures are in a similar range as for the English coast (p1), giving 0.99 for correlation coefficients, 0.99 for IOA and near 1 for skill variance (skvar). The bias is for both models on a low level, slightly negative for BSHcmod v4 (−0.02 K), positive for AMM7 at 0.15 K.

In Fig. 10 (lower), the time series of salinity difference for the Oyster Ground point p2 are shown. The mean level of observed salinity (mean value = 34.43) has been slightly overestimated by BSHcmod v4 (mean value = 34.68) and underestimated by AMM7 (mean value = 34.11). This is visible in Fig. 10 (lower left) which shows mainly positive differences for BSHcmod and mainly negative forAMM7. The observed variability cannot be reached by both models. Although AMM7 skvar is around 1, the IOA is only 0.3 (for BSHcmod v4 at 0.53), the BSHcmod (AMM7) correlation coefficient accounts only to 0.4 (−0.04). The visual inspection of lower panel in Fig. 10 confirms that agreement between models and observations is low.

As it was already described in Petersen et al. (2011), low salinity intrusions can be observed in that North Sea region, often originating from the Rhine/Maas river estuary. The salinity dropped in 2011 to a level of 33.5. In 2008, an even more pronounced salinity drop to 32 has been observed (not shown). The drop event of 2011 has been recognized by BSHcmod v4 and AMM7, however, the amplitude has been underestimated resulting in high differences between model and in-situ data. This has also been visible in Fig. 7 (right) by positive values between April and June 2011 for AMM7. However, subsequent to the observed salinity drop, AMM7 simulated a second, even more pronounced drop in summer 2011 which has not been observed by the FerryBox and by BSHcmod v4 at that position. It could be argued, that a second drop has been there, but at a shifted position which could not be detected by the FerryBox.

So both models seem to be capable to simulate riverine influence in the North Sea not only directly at the river mouth. However, the mixing of coastal and estuarine water is probably underestimated in the systems. It is known for example that the AMM7 model underestimates flushing in the German Bight (probably due to the lack of representation of wetting and drying processes in this shallow region) and this is at least partially responsible for the underestimates of salinity in the region. Another explanation could be an overestimation of vertical mixing in the model. A mixing of colder bottom water with higher salinity into the top layer would lead to lower temperatures and higher salinities in the surface layer. Moreover, freshwater eddies far away from the coasts are supposed to cover only small areas which are hard to detect and set to the correct time and position by FerryBox and models, respectively.

3.2.3 MARNET station German Bight

For analysis of time series of MARNET station point, the annual cycle of water temperatures for MARNET, FerryBox on TD and both models are shown in Fig. 11 (upper left) The time period encompasses three years from 2009–2012. The amplitude of water temperature varies from year to year, but is generally higher than at the other two positions p1 and p2. The highest one is observed in 2010, with an 18 K seasonal SST range and summer SST around 19 °C. In 2011, the summer water temperatures were lower, reaching only 16–18 °C. Unfortunately, in 2009 FerryBox data contain many gaps; so, the full annual cycle cannot be compared to BSHcmod v4, however, for existing observations, BSHcmod v4 simulation results are in agreement with the observations. In

OSD
12, 355–401, 2015

2010 and 2011, BSHcmod v4 agrees well with MARNET and FerryBox observations, except for two time episodes: in July 2010, the observed temperature maximum is also recognized correctly by the model; however, the temperatures in September are too low. In 2011, the summer maximum in July and August is underestimated by up to 2 K. It is clearly seen in Fig. 11 (upper left) that not only the annual cycle, but also smaller variation is present in the model. In Fig. 11 (upper right) scattering of water temperatures reflects the overall good agreement between observations. There is also good agreement between AMM7 and observations for the year 2011, as shown also in Fig. 11 (upper left).

Consequently, the bias of AMM7 water temperatures is at −0.24 K and, thus, is in the range of the other two areas. The bias of BSHcmod v4 is at 0.02; however, the stde of BSHcmod v4 is higher (1.2 K) than for English coast (0.99 K) and for the Oyster Ground (0.87 K), while for AMM7 the stde is slightly lower, at 0.98. The skvar and the IOA are for both models near the ideal value of 1. Also the cost function is near zero for both models, meaning that simulations are well within the standard deviation of observations. However, both values are higher than for p1 and p2.

Figure 11 (lower left) shows the time series of salinity in a manner as it is described for water temperatures in Fig. 11 (upper). The time series are dominated by three large salinity drops below 31 in June 2010, January 2011 and May 2011. The first one lasts over a long time period of more than one month and is recognized by BSHcmod v4, but not before July. The next event in January 2011 is recognized timely but the salinity values are obviously underestimated. The third event is recognized by BSHcmod v4, but not on the correct date. It is also underestimated by AMM7. After the third observed drop event, BSHcmod v4 and AMM7 salinity simulations show a further salinity drop from June 2011 on for a time period of several months, which is not present in the observations. It is not clear why both models should both predict freshening in the summer 2011 which in fact did not occur.

In comparison to FB observations, shown in Fig. 11 (lower right), the bias of salinity simulation is negative for BSHcmod v4 (−0.33) at German Bight, while being positive

OSD
12, 355–401, 2015

at English Coast and Oyster Ground. The bias for AMM7 is on same level than for BSHcmod v4, i.e. negative at −0.39. This is in line with the other positions. It should be noted that the statistics for the AMM7 and BSHcmod are calculated from different length time series, so despite the differences in statistics shown here both models behave similarly over the period in which data is available for both. Skvar and IOA are much less for salinity than for water temperature. Yet, both models are mostly better for German Bight than for the other regions.

For extending the analysis of temporal variation in observations and model simulations, the spectral densities for water temperatures of MARNET measurements and BSHcmod v4 and AMM7 simulations have been evaluated. In Fig. 12 the power spectra of the three time series are shown. The frequency is shown in s^{-1}. No smoothing has been applied to the spectra.

The spectral densities of each time series are located in the same range. In all spectra, the density peak at the diurnal cycle is present, at the BSHcmod v4 model simulations more sharpened than at MARNET observations and AMM7 simulations. Also the tide peak at 12.43 h is recognizable in every spectrum.

Altogether, it can be observed that both models contain similar spectrum features than observations. Indeed, these features are supposed to be basic dynamics in the simulation of North Sea dynamics and that has been reviewed in Otto et al. (1990).

All in all, the finer grid size as well as a more correct simulation of Elbe River discharge results in slightly better simulation of salinity at the German Bight point than for other positions. Moreover, the German Bight point is more distant to the coast than the English coast point, so the influence of near coastal processes is supposed to be less pronounced.

4 Summary and conclusions

The hydrodynamic model simulations agreed well with the continuous operational FerryBox and MARNET in situ water temperature observations along the FerryBox route

OSD
12, 355–401, 2015

from England to Germany, as well as in detail for three positions also situated along the transect.

Besides capturing seasonal variations, smaller scale temporal variations, like tidal fluctuations, are also present in BSHcmod v4 and AMM7. Statistical tests indicate that AMM7 could be improved by reducing the offset of mean temperature levels (AMM7 0.19 K, see Table 1), the correct representation of summer maximum temperatures in case of BSHcmod v4 and the level of variation.

Both model results predict poorly the variations of water temperatures and salinity near the coasts, and in particular in the cold Scottish coastal current. It could be argued, that this is due to weak vertical mixing, especially at the end of summer (only for BSHcmod v4). In the German Bight, scattering of modelled and observed data is higher than for central parts of the North Sea.

Spectral analysis reveals good agreement between simulations and observations. The strong tidal signal in the German Bight is well represented in both models.

Comparisons of salinity show much higher differences between observations and simulations and reveal geographical dependencies of the model performance. Altogether, both models show certain limitations:

BSHcmod does not capture properly the variability, or the correct salinity range in the German Bight east of 6° E. This may be due to a deficient model input of fresh water river forcing. Otherwise BSHcmod v4 generally accurately captures salinity for the open North Sea. Similarly, AMM7 generally performs well in the central parts of the North Sea, but misrepresents the salinity distribution near the coasts.

Low salinity events occurring in the southern North Sea are caught by BSHcmod v4 and AMM7 to some extent. In order to improve salinity values in the model it is for one thing recommended using validated daily fresh water input data for all main rivers entering the North Sea.

And for another thing, the poor model representation of the vertical mixing that occurs within the Continental coastal water, along the south-east edge of the North Sea, or within the Southern North Sea, where freshwater entering the coastal ocean from the

OSD
12, 355–401, 2015

major rivers Rhine, Elbe, Ems, etc. affects both salinity and temperature distributions. This process and the model misfits have to be further studied. In addition, for BSHcmod v4 also the nesting process of different grid sizes has to be further evaluated.

FerryBox measurements, routinely validated for accuracy and precision using external checks and laboratory analyses, can serve as a reliable proxy for the state of the surface temperature and salinity variations in the North Sea. The operational FerryBox measurements are routinely checked against water probes. Water temperature and salinity measurements are validated against laboratory analyses and revealed good results. The FerryBox and the MARNET measurements were also in a good agreement. There is a bias of 0.37 K, in the water temperature measurements from the FerryBox, most likely caused by warming inside of the system. While the FerryBox measurements are done along transects throughout the North Sea and other parts of the European coastal oceans, fixed stations provide longer term time series at a particular site, but lack of spatial information for the neighbouring regions. In this study, using the FerryBox and the MARNET datasets, both of these types of measurements were combined.

Through European Union projects like JERICO and MyOcean, collaboration between the institutions operating the different FerryBox lines has been improved in terms of maintenance, best practices of calibration, biofouling and installation of new FerryBox routes. With the improved FerryBox data consistency, there is currently even more interest in how the monitored coastal data can be assimilated and used for validation of (operational) models. Especially for ecosystem models, reliable physical observation data are essential (Artioli et al., 2012).

Previously, FerryBox transect data have been successfully assimilated in North Sea models, as has been demonstrated by Schulz-Stellenfleth and Stanev (2010), Stanev et al. (2011) and Grayek et al. (2011). The latter have shown, that FerryBox measurements are comparable to satellite-derived SST data (extracted from the OSTIA data set, Donlon et al., 2009) and to other measurements from fixed stations. For the Aegean Sea, Korres et al. (2009) also have assimilated FerryBox sea surface salinity data together with AVHRR sea surface temperature data into a hydrodynamic model.

OSD
12, 355–401, 2015

They showed that the assimilation of satellite SST data enhanced the model performance and that the addition of FerryBox salinity data helped to improve model results even more, by significantly decreasing the RMS error statistics for the southern Aegean Sea.

The AMM7 model already assimilates SST from SoO managed under the Joint WMO-IOC Technical Commission for Oceanography and Marine Meteorology, where telecommunications have been established to transmit data via the Global Telecommunications System (GTS) in real-time. Ferrybox data, like the ones used in this study could also be included relatively easily, if communications allowed it.

Data assimilation of FerryBox data is performed in most cases using a Kalman filter approach to extrapolate one-dimensional data on 2-D fields. Since the influence of the assimilated FerryBox data are restricted to a rather shallow area around the FerryBox track, one method for data assimilation could be the use of particle tracking algorithms for (approximately) conservative parameters like temperature and salinity in combination with 2-D North Sea current fields, e.g. of operational BSHcmod. This method has been successfully used by Klein and Dick (1999). A data assimilation scheme for operational use is under development at BSH (Losa et al., 2014, 2012). So far this method has been tested during the assimilation of satellite-derived SST data along with vertical temperature and salinity profiles.

The operational implementation of FerryBox data is one of the next steps for completion of the scheme. An important next step is overcoming the delayed mode limitation of FerryBox measurements for assimilation into operational forecast modelling systems. This is partly achieved already, mainly at the recently installed FerryBoxes using satellite communication. For the operational assimilation, also operational post-processing of FerryBox data for data quality assessment is necessary and has also been partly established.

Acknowledgements. We want to thank the shipping company DFDS Tor Line and DFDS Lys Line and the crews on *TorDania* and *Lysbris*. It is much work to maintain our FerryBoxes to keep them running so we are especially thankful to our technical and data management staff.

This work was partly funded within the 7th Framework Program of the European Union under grant number 262584 in the project JERICO and partly by the Coastal Observing System for Northern and Arctic Seas (COSYNA), financed by the Helmholtz-Zentrum Geesthacht Centre for Materials and Coastal Research GmbH (HZG).

The service charges for this open-access publication have been covered by a research centre of the Helmholtz Association.

References

Artioli, Y., Blackford, J. C., Butenschon, M., Holt, J. T., Wakelin, S. L., Thomas, H., Borges, A. V., and Allen, J. I.: The carbonate system in the North Sea: sensitivity and model validation, J. Marine Syst., 102–104, 1–13, 2012.

Backhaus, J. O.: A three-dimensional model for the simulation of shelf sea dynamics, Deut. Hydrogr. Z., 38, 165–187, 1985.

Baretta, J. W., Ebenhöh, W., and Ruardij, P.: The European regional seas ecosystem model, a complex marine ecosystem model, Neth. J. Sea Res., 33, 233–246, 1995.

Berntsen, J. and Svendsen, E.: Using the SKAGEX dataset for evaluation of ocean model skills, J. Marine Syst., 18, 313–331, 1999.

Blackford, J. C., Allen, J. I., and Gilbert, F. J.: Ecosystem dynamics at six contrasting sites: a generic modelling study, J. Marine Syst., 52, 191–215, 2004.

Blockley, E. W., Martin, M. J., McLaren, A. J., Ryan, A. G., Waters, J., Lea, D. J., Mirouze, I., Peterson, K. A., Sellar, A., and Storkey, D.: Recent development of the Met Office operational ocean forecasting system: an overview and assessment of the new Global FOAM forecasts, Geosci. Model Dev., 7, 2613–2638, doi:10.5194/gmd-7-2613-2014, 2014.

BSH: http://www.bsh.de/en/Marine_data/Observations/MARNET_monitoring_network/ (last access: 17 December 2014), 2014.

BSH: http://www.bsh.de/en/Marine_data/Observations/Daily_circulation_pattern_German_Bight/index.jsp (last access: 17 February 2015), 2015.

Burchard, H.: Presentation of a new numerical model for turbulent flow in estuaries, Copenhagen, Denmark, 24–26 August 1998.

OSD
12, 355–401, 2015

Evaluation of numerical models by FerryBox and Fixed Platform data in the North Sea

M. Haller et al.

Burchard, H. and Bolding, K.: GETM – a general estuarine transport model, Scientific Documentation, European Commission, 157 pp., 2002.
Cefas: SmartBuoys, available at: http://www.cefas.defra.gov.uk/our-services/monitoring-and-mapping/autonomous-monitoring/smartbuoys.aspx (last access: 25 February 2015), 2015.
COSYNA: http://www.hzg.de/institutes_platforms/cosyna/index.php.en (last access: 17 December 2014), 2014.
Dick, S., Kleine, E., Müller-Navarra, S. H., Klein, H., and Komo, H.: The operational circulation model of BSH (BSHcmod) – model description and validation, Berichte des BSH, vol. 29, 49 pp., 2001.
Doms, G. and Schättler, U.: The non-hydrostatic limited-area model LM (Lokalmodell) of DWD. Part I: Scientific Documentation, 1999.
Donlon, C. J., Casey, K. S., Robinson, I. S., Gentemann, C. L., Reynolds, R. W., Barton, I., Arino, O., Stark, J., Rayner, N., LeBorgne, P., Poulter, D., Vazquez-Cuervo, J., Armstrong, E., Beggs, H., Llewellyn-Jones, D., Minnett, P. J., Merchant, C. J., and Evans, R.: The GODAE high-resolution sea surface temperature pilot project, Oceanography, 22, 34–45, 2009.
Edwards, K. P., Barciela, R., and Butenschön, M.: Validation of the NEMO-ERSEM operational ecosystem model for the North West European Continental Shelf, Ocean Sci., 8, 983–1000, doi:10.5194/os-8-983-2012, 2012.
FerryBox: available at: http://www.ferrybox.org, 2014.
Grayek, S., Staneva, J., Schulz-Stellenfleth, J., and Petersen, W.: Use of Ferrybox surface temperature and salinity measurements to improve model based state estimates for the German Bight, J. Marine Syst., 88, 45–59, 2011.
Hydes, D. J., Kelly-Gerreyn, B. A., Colijn, F., Petersen, W., Schroeder, F., Mills, D. K., Durand, D., Wehde, H., Sorensen, K., and Morrison, G.: The way forward in developing and integrating FerryBox technologies, Venice, Italy, 21–25 September 2009.
Janssen, F., Schrum, C., and Backhaus, J. O.: A climatological data set of temperature and salinity for the Baltic Sea and the North Sea, Ger. J. Hydrogr., 9, 5–245, 1999.
Jickells, T. D.: Nutrient biogeochemistry of the coastal zone, Science, 281, 217–222, 1998.
Kannen, A.: Challenges for marine spatial planning in the context of multiple sea uses, policy arenas and actors based on experiences from the German North Sea, Reg. Environ. Change, 9, 1–12, doi:10.1007/s10113-012-0349-7, 2012.

OSD
12, 355–401, 2015
Evaluation of numerical models by FerryBox and Fixed Platform data in the North Sea
M. Haller et al.

Klein, H. and Dick, S.: Currents at the German Bight lightship: a comparison between ADCP measurements and the BSH forecast model, Deut. Hydrogr. Z., 51, 465–475, 1999.
Korres, G., Nittis, K., Hoteit, I., and Triantafyllou, G.: A high resolution data assimilation system for the Aegean Sea hydrodynamics, J. Marine Syst., 77, 325–340, 2009.
Legates, D. R. and McCabe, G. J.: Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation, Water Resour. Res., 35, 233–241, 1999.
Lenhart, H.-J. and Pohlmann, T.: The ICES-boxes approach in relation to results of a North Sea circulation model, Tellus A, 49, 139–160, 1997.
Loewe, P. (Ed.): System Nordsee – Zustand 2005 im Kontext langzeitlicher Entwicklungen, BSH, Hamburg und Rostock, 2009.
Loewe, P., Klein, H., and Weigelt-Krenz, S. (Eds.): System Nordsee – 2006 & 2007: Zustand und Entwicklungen, BSH, Hamburg und Rostock, 2013.
Losa, S. N., Danilov, S., Schröter, J., Nerger, L., Maßmann, S., and Janssen, F.: Assimilating NOAA SST data into the BSH operational circulation model for the North and Baltic Seas: inference about the data, J. Marine Syst., 105–108, 152–162, 2012.
Losa, S. N., Danilov, S., Schröter, J., Janjić, T., Nerger, L., and Janssen, F.: Assimilating NOAA SST data into BSH operational circulation model for the North and Baltic Seas: Part 2. Sensitivity of the forecast's skill to the prior model error statistics, J. Marine Syst., 129, 259–270, 2014.
Luyten, P. J., Jones, J. E., Proctor, R., Tabor, A., Tett, P., and Wild-Allen, K.: COHERENS – A coupled hydrodynamical-ecological model for regional and shelf seas: user documentation, management unit of the mathematical models of the North Sea, Belgium, 911 pp., 1999.
Maar, M., Moller, E. F., Larsen, J., Madsen, K. S., Wan, Z. W., She, J., Jonasson, L., and Neumann, T.: Ecosystem modelling across a salinity gradient from the North Sea to the Baltic Sea, Ecol. Model., 222, 1696–1711, 2011.
Madec, G.: NEMO Ocean Engine, 300 pp., 2008.
Martin, M. J., Hines, A., and Bell, M. J.: Data assimilation in the FOAM operational short-range ocean forecasting system: a description of the scheme and its impact, Q. J. Roy. Meteor. Soc., 133, 981–995, 2007.
Mills, D. K., Laane, R. W. P. M., Rees, J. M., Rutgers van der Loeff, M., Suylen, J. M., Pearce, D. J., Sivyer, D. B., Heins, C., Platt, K., and Rawlinson, M.: Smartbuoy: a marine environmental monitoring buoy with a difference, in: Elsevier Oceanography Series, edited by: Dahlin, H., Flemming, N. C., Nittis, K., and Petersson, S. E., Elsevier, 2003.

OSD
12, 355–401, 2015

Myocean: Quality Information Document MyOcean V1 System WP7 – North-West Shelf, 2011.
O'Dea, E. J., Arneold, A. K., Edwards, K. P., Furner, R., Hyder, P., Martin, M. J., Siddorn, J. R., Storkey, D., While, J., Holt, J. T., and Liu, H.: An operational ocean forecast system incorporating NEMO and SST data assimilation for the tidally driven European North-West shelf, J. Oper. Oceanogr., 5, 3–17, 2012.
OSPAR: Quality status Report 2000, Region II – Greater North Sea, London, 5–25, 2000.
OSPAR: Quality status report 2010, OSPAR Commission, London, 2010.
Otto, L., Zimmerman, J. T. F., Furnes, G. K., Mork, M., Saetre, R., and Becker, G.: Review of the physical oceanography of the North Sea, Neth. J. Sea Res., 26, 161–238, 1990.
Petersen, W., Petschatnikov, M., and Schroeder, F.: FerryBox Systems for Monitoring Coastal Waters, Amsterdam, the Netherlands, 325–333, 2003.
Petersen, W., Colijn, F., Elliot, J., Howarth, M. J., Hydes, D. J., Kaitala, S., Kontoyiannis, H., Lavin, A., Lips, I., Pfeiffer, K. D., Proctor, R., Ridderinkhof, H., and Sorensen, K.: European FerryBox Project: From Online Oceanographic Measurements to Environmental Information, Brest, France, 2005.
Petersen, W., Colijn, F., Hydes, D. J., and Schroeder, F.: FerryBox: From on-line oceanographic observations to environmental information, EU Project FerryBox 2002–2005, EuroGOOS Publications, 25, 36, 2007.
Petersen, W., Wehde, H., Krasemann, H., Colijn, F., and Schroeder, F.: FerryBox and MERIS – assessment of coastal and shelf sea ecosystems by combining in situ and remotely sensed data, Estuar. Coast. Shelf S., 77, 296–307, 2008.
Petersen, W., Schroeder, F., and Bockelmann, F. D.: FerryBox - Application of continuous water quality observations along transects in the North Sea, Ocean Dynam., 61, 1541–1554, 2011.
Pleskachevsky, A., Eppel, D. P., and Kapitza, H.: Interaction of waves, currents and tides, and wave-energy impact on the beach area of Sylt Island, Ocean Dynam., 59, 451–461, 2009.
Pohlmann, T.: A meso-scale model of the central and southern North Sea: consequences of an improved resolution, Cont. Shelf Res., 26, 2367–2385, 2006.
Queste, B. Y., Fernand, L., Jickells, T. D., and Heywood, K. J.: Spatial extent and historical context of North Sea oxygen depletion in August 2010, Biogeochemistry, 113, 53–68, 2013.
Riethmuller, R., Colijn, F., Krasemann, H., Schroeder, F., Ziemer, F., and IEEE: COSYNA, an Integrated Coastal Observation System for Northern and Arctic Seas, in: Oceans 2009 – Europe, Vols. 1 and 2, Oceans-IEEE, 2009.

OSD
12, 355–401, 2015

Evaluation of numerical models by FerryBox and Fixed Platform data in the North Sea

M. Haller et al.

Schulz-Stellenfleth, J. and Stanev, E. V.: Statistical assessment of ocean observing networks – a study of water level measurements in the German Bight, Ocean Model., 33, 270–282, 2010.
Siddorn, J. R., Allen, J. I., Blackford, J. C., Gilbert, F. J., Holt, J. T., Holt, M. W., Osborne, J. P., Proctor, R., and Mills, D. K.: Modelling the hydrodynamics and ecosystem of the North-West European continental shelf for operational oceanography, J. Marine Syst., 65, 417–429, 2007.
Søiland, H. and Skogen, M. D.: Validation of a three-dimensional biophysical model using nutrient observations in the North Sea, ICES J. Mar. Sci., 57, 816–823, 2000.
Sperna Weiland, F. C., van Beek, L. P. H., Kwadijk, J. C. J., and Bierkens, M. F. P.: The ability of a GCM-forced hydrological model to reproduce global discharge variability, Hydrol. Earth Syst. Sci., 14, 1595–1621, doi:10.5194/hess-14-1595-2010, 2010.
Stanev, E. V., Schulz-Stellenfleth, J., Staneva, J., Grayek, S., Seemann, J., and Petersen, W.: Coastal observing and forecasting system for the German Bight – estimates of hydrophysical states, Ocean Sci., 7, 569–583, doi:10.5194/os-7-569-2011, 2011.
Storkey, D., Blockley, E. W., Furner, R., Guiavarc'h, C., Lea, D., Martin, M. J., Barciela, R. M., Hines, A., Hyder, P., and Siddorn, J. R.: Forecasting the ocean state using NEMO: the new FOAM system, J. Operat. Oceanogr., 3, 3–15, 2010.
Thomas, H., Gattuso, J. P., and Smith, S. V.: Coastal Biogeochemistry at the EGS-AGU-EUG Joint Assembly, LOICZ newsletter 28, 1–4, 2003.
Thomas, H., Bozec, Y., de Baar, H. J. W., Elkalay, K., Frankignoulle, M., Schiettecatte, L.-S., Kattner, G., and Borges, A. V.: The carbon budget of the North Sea, Biogeosciences, 2, 87–96, doi:10.5194/bg-2-87-2005, 2005.
Turrell, W. R.: New hypotheses concerning the circulation of the northern North Sea and its relation to North Sea fish stock recruitment, ICES J. Mar. Sci., 49, 107–123, 1992.
Volent, Z., Johnsen, G., Hovland, E. K., Folkestad, A., Olsen, L. M., Tangen, K., and Sorensen, K.: Improved monitoring of phytoplankton bloom dynamics in a Norwegian fjord by integrating satellite data, pigment analysis, and FerryBox data with a coastal observation network, J. Appl. Remote Sens., 5, 053561, doi:10.1117/1.3658032, 2011.
Wehde, H., Schroeder, F., Colijn, F., Reinke, S., Petersen, W., Schrum, C., Plüß, A., and Mills, D. K.: FerryBox observations in the southern North Sea – application of numerical models for improving the significance of the FerryBox data, Brest, France, 6–9 June 2005, 169–173, 2006.

OSD
12, 355–401, 2015

Willmott, C. J.: On the validation of models, Phys. Geogr., 2, 184–194, 1981.

OSD
12, 355–401, 2015

Evaluation of numerical models by FerryBox and Fixed Platform data in the North Sea

M. Haller et al.

Title Page
Abstract Introduction
Conclusions References
Tables Figures
◀ ▶
Back Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

Table 1. Statistical measures for performance analysis of BSHcmod v4 and AMM7, respectively.

Parameter	BSHcmod v4 WTemp	AMM7 WTemp	BSHcmod v4 Salinity	AMM7 Salinity
Bias	−0.02 K	0.19 K	−0.17	−0.89
STDE	0.72 K	0.38 K	0.66	0.62
RMSE	0.92 K	0.44 K	0.98	1.1
Skill variance	0.93	1.03	0.8	0.77
Cost function	0.07	0.06	1.25	1.75
Index of Agreement	0.94	0.98	0.56	0.19
Correlation	0.93	0.99	0.8	0.19

OSD

12, 355–401, 2015

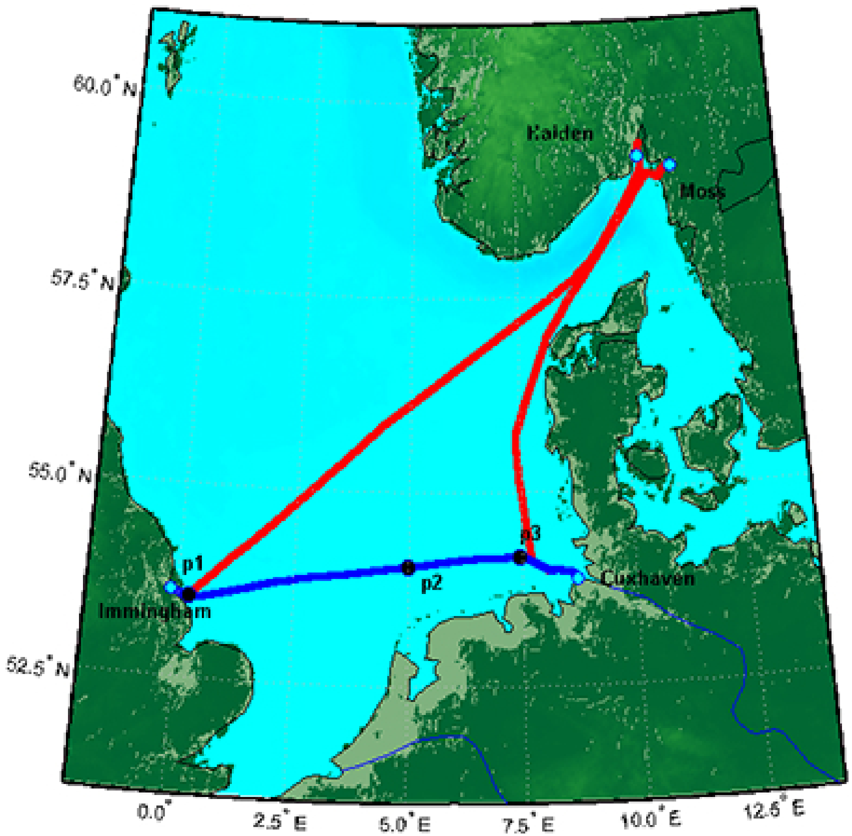

Figure 1. FerryBox routes and crossing points in the North Sea. Contour lines indicate the bathymetry. The green line shows the TC route Amsterdam – Bergen, the blue line marks the TD route Cuxhaven-Immingham and the red lines indicate the LB route England–Norway–Germany. Crossing points of FerryBox routes are indicated by black points and labelled p1, p2, p3, respectively. p1 is situated at the English East Coast. p2 marks the TD/TC meeting point in the Oyster Ground area. At p3, the MARNET station *Deutsche Bucht* is located.

OSD
12, 355–401, 2015

Evaluation of numerical models by FerryBox and Fixed Platform data in the North Sea

M. Haller et al.

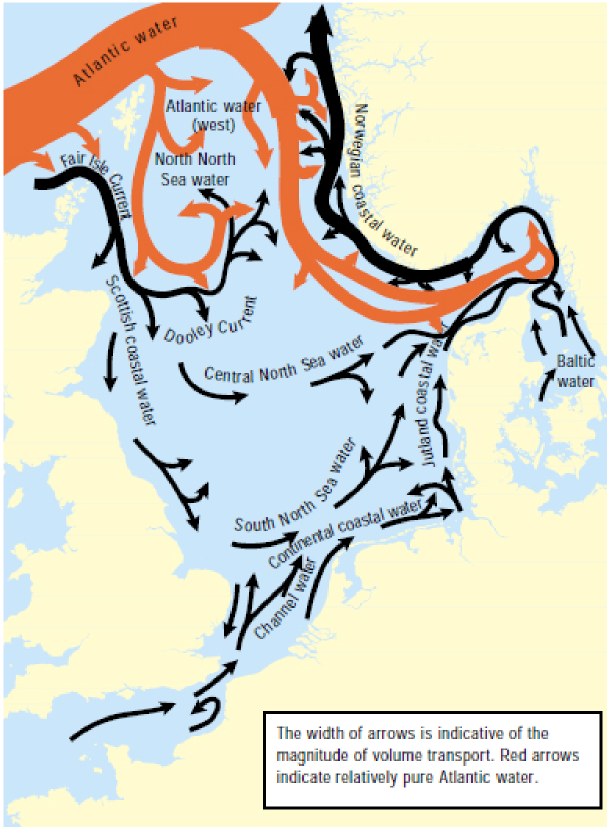

Figure 2. General circulation scheme in the North Sea (from OSPAR, 2000; adapted from Turrell, 1992).

OSD
12, 355–401, 2015

Evaluation of numerical models by FerryBox and Fixed Platform data in the North Sea

M. Haller et al.

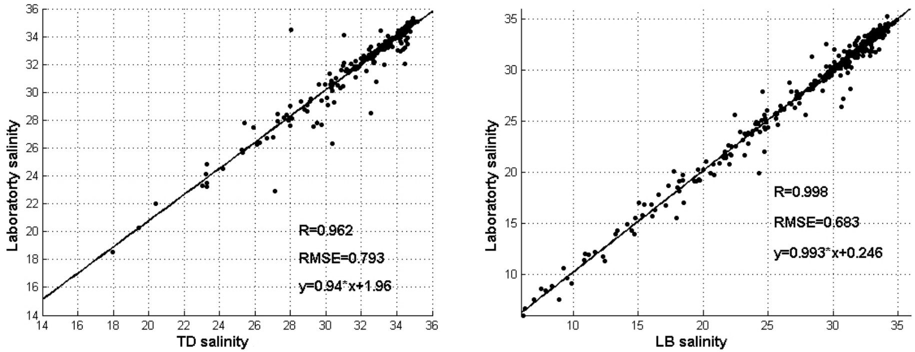

Figure 3. Comparison of FerryBox salinity measurements and water sample analyses in the lab for LB (left) and TD (right).

OSD
12, 355–401, 2015

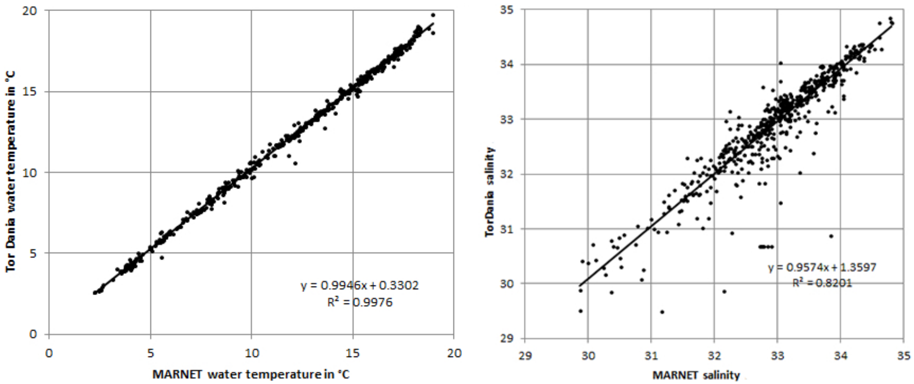

Figure 4. Comparison of water temperature (left) and salinity (right) measurements in German Bight at geographical point p3 from 2007 to 2011.

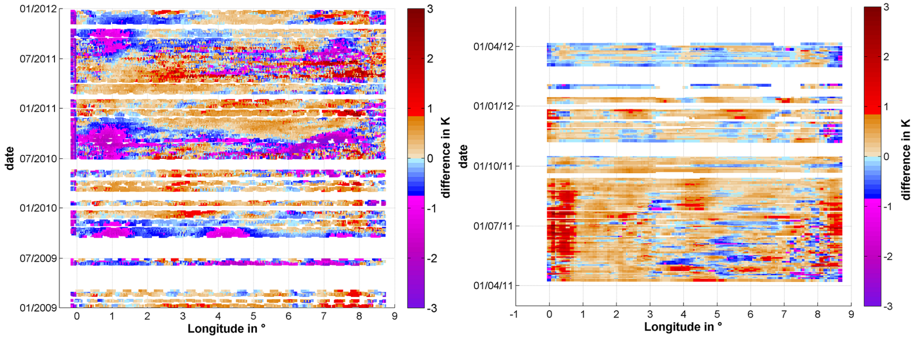

Figure 5. Differences of water temperatures for TD transect (left side BSH-TD 2009–2011, right side AMM7-TD 2011–2012). On the left side is located the East England coast, on the right side the German Bight. Positive values indicate model overestimation. Differences are statistically significant beyond ±0.84 psu (twofold stde of FerryBox).

OSD
12, 355–401, 2015

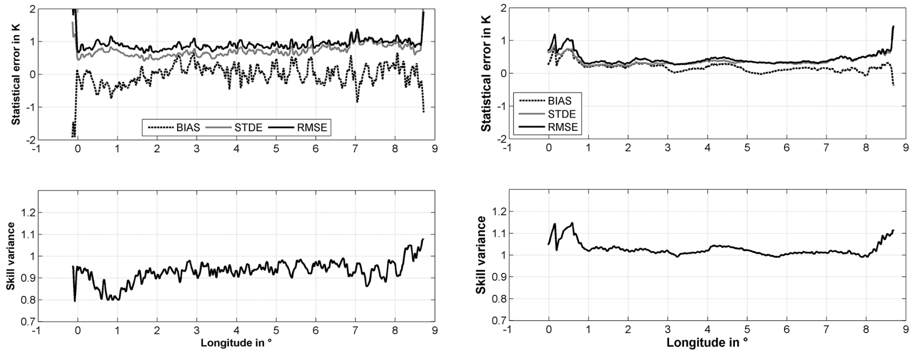

Figure 6. Standard deviation of error (stde), bias and root mean square error (rmse) (up) and skill variance (skvar) (down) of BSHcmod 2009–2011 (left) and AMM7 (right) 2011–2012 for water temperatures and TD transect.

OSD
12, 355–401, 2015

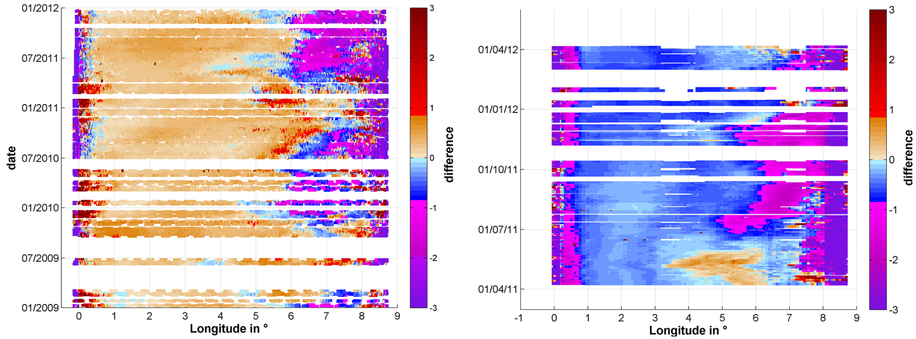

Figure 7. Differences of salinity for TD transect (left side BSH-TD 2009–2011, right side AMM7-TD 2011–2012). On the left side is located the East England coast, on the right side the German Bight. Positive values indicate model overestimation. Differences are statistically significant beyond ±0.8 psu (twofold stde of FerryBox).

OSD
12, 355–401, 2015

Evaluation of numerical models by FerryBox and Fixed Platform data in the North Sea

M. Haller et al.

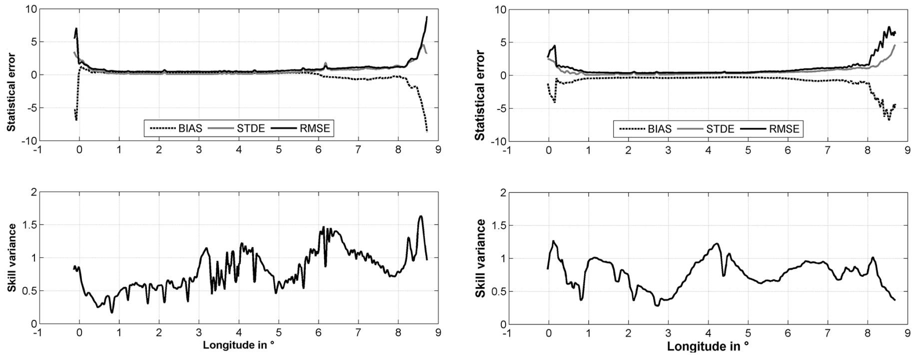

Figure 8. Standard deviation of error (stde), bias and rmse (up) and skill variance (skvar) (down) of BSHcmod (left) 2009–2011 and AMM7 (right) 2011–2012 for salinity and TD transect.

OSD
12, 355–401, 2015

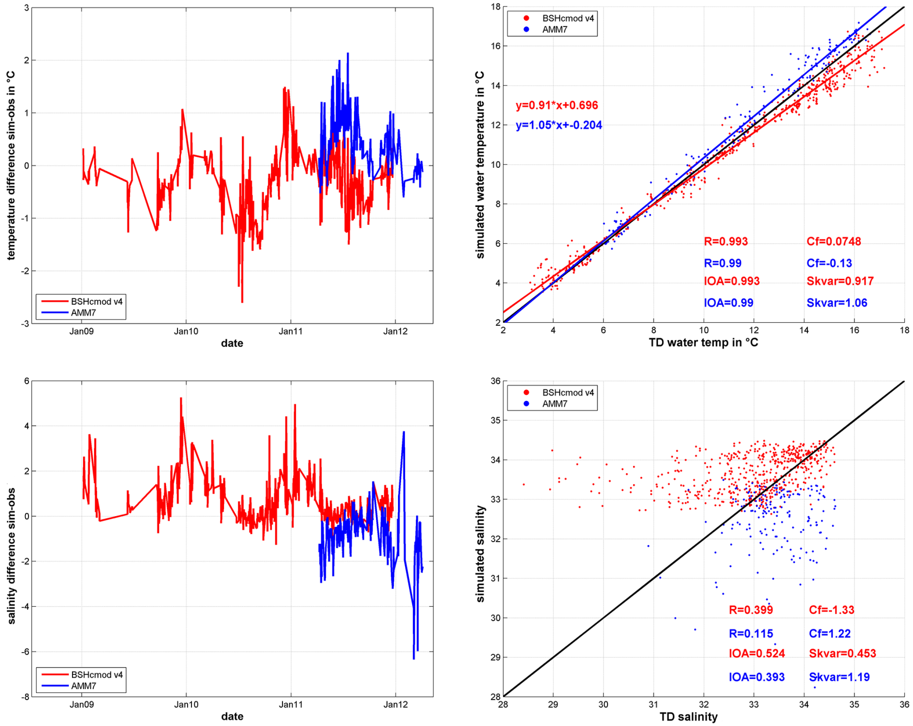

Figure 9. Upper panel: time series of temperature difference (left) and scatterplot of temperatures (right) of FerryBox measurements of TD and LB and model results at East Coast of England. Lower panel: time series of salinity difference (left) and scatterplot of salinity (right) of FerryBox measurements of TD and LB and model results at East Coast of England. Note the trimmed axes for salinity ranging from 28 to 36 psu.

OSD

12, 355–401, 2015

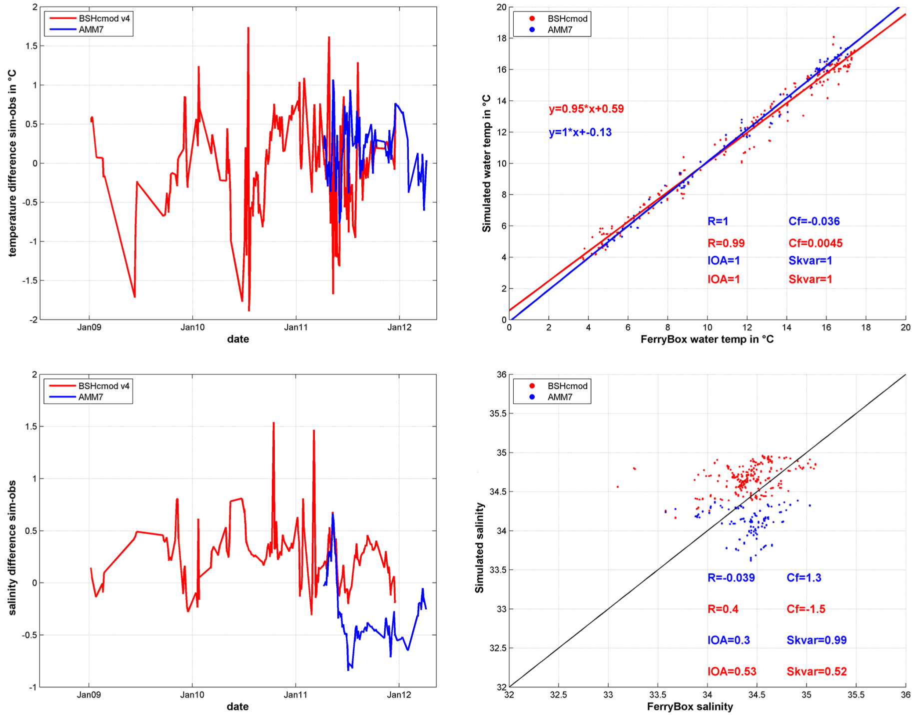

Figure 10. Upper panel: time series of temperature differences of BSHcmod v4 and AMM7 minus FerryBox measurements of TD and LB at crossing point p2 (Oyster Ground). Lower panel: time series of salinity differences of BSHcmod v4 and AMM7 minus FerryBox measurements of TD and LB at crossing point p2 (Oyster Ground). Note the trimmed axes for salinity ranging from 32 to 36 psu.

OSD
12, 355–401, 2015

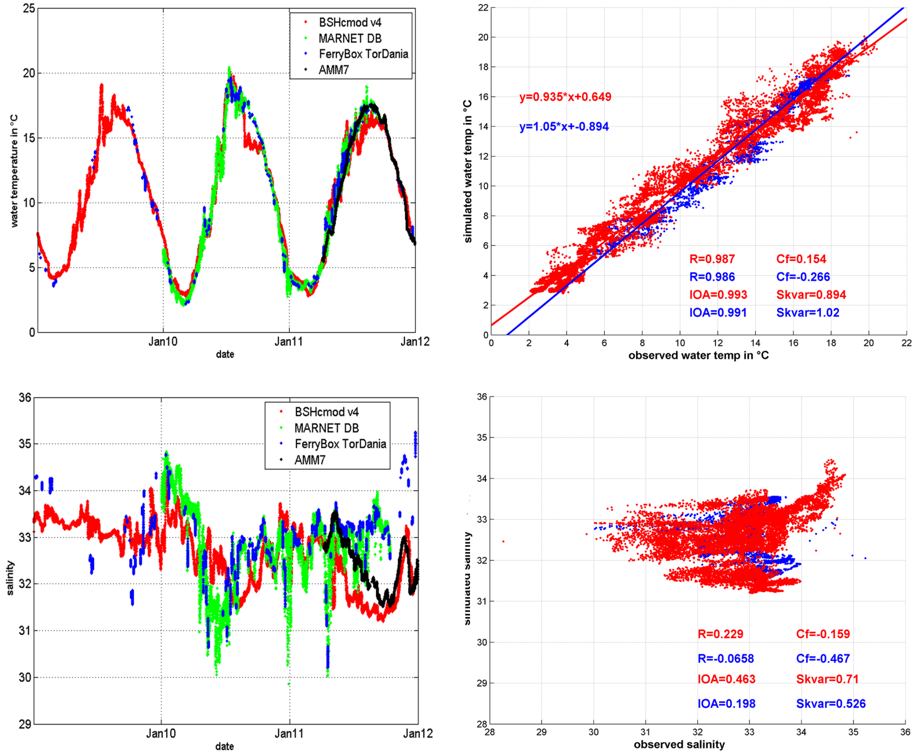

Figure 11. Upper panel: time series (left) and scattering (right) of water temperatures of FerryBox observations, BSHcmod v4 and AMM7 at German Bight (p3). In right figure, red denotes BSHcmod v4, blue AMM7. Lower panel: time series of salinity of FerryBox observations, BSHcmod v4 and AMM7 at German Bight (p3). In right figures, red dots denotes BSHcmod v4, blue dots denote AMM7.

OSD
12, 355–401, 2015

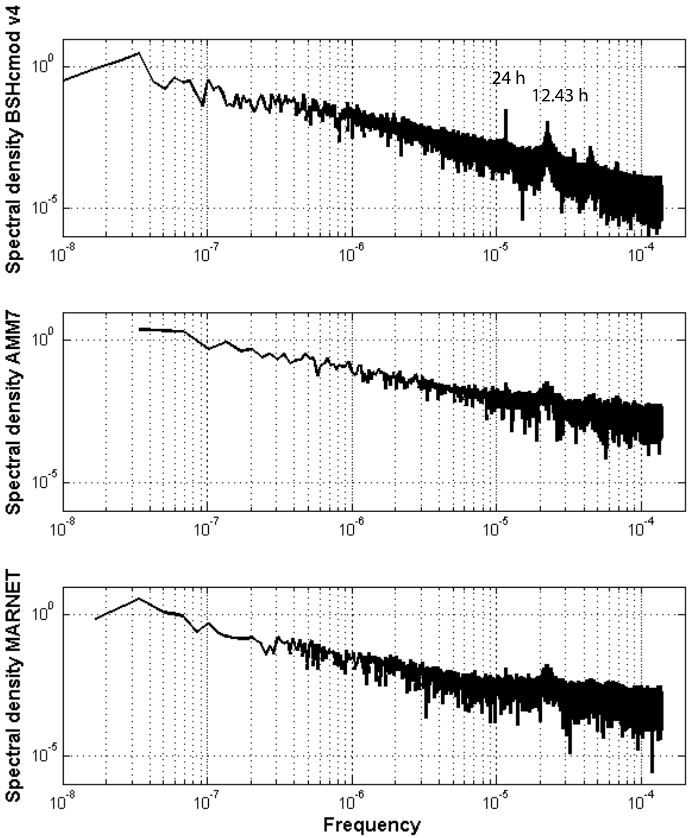

Figure 12. Spectral density of water temperature time series of BSHcmod v4 (up), AMM7 (middle) and MARNET (down). Frequencies are denoted in s^{-1}.

OSD
12, 355–401, 2015

Evaluation of numerical models by FerryBox and Fixed Platform data in the North Sea

M. Haller et al.