

The effect of density stratification on the prediction of global storm surges

Tsubasa Kodaira¹  · Keith R. Thompson¹ · Natacha B. Bernier²

Received: 16 March 2016 / Accepted: 23 September 2016 / Published online: 6 October 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract With the long-term goal of developing an operational forecast system for total water level, we conduct a hindcast study of global storm surges for Fall 2014 using a baroclinic ocean model based on the NEMO framework. The model has 19 vertical levels, a horizontal resolution of $1/12^\circ$, and is forced by hourly forecasts of atmospheric wind and air pressure. Our first objective is to evaluate the model's ability to predict hourly sea levels recorded by a global array of 257 tide gauges. It is shown that the model can provide reasonable predictions of surges for the whole test period at tide gauges with relatively large tidal residuals (i.e., gauges where the standard deviation of observed sea level, after removal of the tide, exceeds 5 cm). Our second objective is to quantify the effect of density stratification on the prediction of global surges. It is found that the inclusion of density stratification increases the overall predictive skill at almost all tide gauges. The increase in skill for the instantaneous peak surge is smaller. The location for which the increase in overall skill is largest (east coast of South Africa) is discussed in detail and physical reasons for the improvement are given.

Keywords Storm surge · Global prediction · Baroclinic model

1 Introduction

Storm surges are variations of sea level about the tide due to storms. Positive surges can cause severe coastal flooding and negative surges can be navigation hazards for ships operating in shallow water. Needham et al. (2015) recently provided an overview of the global data sources, observations and impacts of historically extraordinary storm surges. To reduce possible casualties and economic losses caused by storm surges, many countries have developed regional forecast systems to provide warnings to coastal communities and marine related industries.

The early development of numerical storm surge prediction systems was reviewed by Heaps (1983) and Bode and Hardy (1997). Two-dimensional shallow water equations have been used extensively in operational regional forecast systems, e.g. SLOSH for the US (Jelesnianski et al. 1992) and CS3 for the UK (Flather 2000). Statistical models are sometimes used in conjunction with a dynamical model (e.g. SLOSH) to compensate for errors in the model or its atmospheric forcing. To take account of uncertainty in the atmospheric forcing, there is now a move towards the development of ensemble-based systems. For example, Flowerdew et al. (2010) have developed an ensemble surge forecast system for the UK and, more recently, Bernier and Thompson (2015) have evaluated an ensemble prediction system for Atlantic Canada.

In contrast to the regional systems mentioned above, the present study focuses on the development of a global, baroclinic storm surge prediction system. The global system is not designed to compete with regional models (it

This article is part of the Topical Collection on the *14th International Workshop on Wave Hindcasting and Forecasting in Key West, Florida, USA, November 8-13, 2015*

Responsible Editor: Kevin Horsburgh

✉ Tsubasa Kodaira
Tsubasa.Kodaira@Dal.Ca

¹ Department of Oceanography, Dalhousie University, Halifax, NS, Canada

² Rechercheen Prévision Numérique Environnementale, RPN-E, Environment and Climate Change Canada, Québec, Canada

only has a resolution of $1/12^\circ$). Its purpose is to improve the predictions of the regional forecast systems. For example, the global model can be used to identify regions with relatively low predictability (due, for example to low predictability in the atmospheric forcing) and also to provide lateral boundary conditions for regional systems.

The first objective of the present study is to quantify the skill of the global model in the prediction of water levels caused by variations in wind and air pressure. In this initial development of the global system, we focus on Fall of 2014 and compare the model's predictions to observations made by a global array of 257 tide gauges. The existing model most relevant to the present study is MOG2D-G (Carrère and Lyard 2003, see also Appendix). This is an excellent global surge prediction model that is used widely to provide a Dynamic Atmospheric Correction (DAC) for dealiasing altimeter data. The DAC data for Fall 2014 are also evaluated against the tide gauge dataset to compare its predictive skill with our model.

The second objective is to quantify how density stratification may influence the predictions of water level and storm surges. One way in which density stratification can influence the surges is by changing the vertical structure of the currents and thus bottom stress (e.g. Gordon 1982). Although some studies (e.g. Davies 1988; Hearn and Holloway 1990; Weisberg and Zheng 2008; Zheng et al. 2013) have examined the effect of vertical structure in barotropic models, the effect of baroclinicity has received less attention. One exception is the modelling study of Minato (1998) of the storm surge in Tosa Bay produced by typhoon Anita. It was shown that the predicted storm surge increased slightly by including density stratification because bottom stress was reduced. Another way in which density stratification can influence surges is by changing the dispersion relation of coastal trapped waves (e.g. Clarke 1977; Gill 1982) and also reducing the amount of backscattering by freely propagating coastal trapped waves as they encounter changes in the width of the shelf (e.g. Wilkin and Chapman 1990). In the present study, we compare model runs with and without realistic density stratification.

The numerical model and its configuration are described in Section 2. The observations used to evaluate the model predictions are described in Section 3. The model predictions are evaluated in Section 4 based on comparison with the observed sea levels, and the effect of density stratification is explored. The results of the study are summarized and discussed in Section 5.

2 Numerical model

We use a baroclinic ocean model based on the NEMO framework to predict storm surges. The model has 19

z-levels in the vertical and a horizontal grid spacing of approximately $1/12^\circ$. To include the effect of mean density stratification, temperature and salinity fields are initialized by, and weakly restored to, an observed climatology. Almost the same numerical model was used recently by Kodaira et al. (2016) in a study of global tidal surface currents associated with M_2 internal tides.

The baroclinic model was further modified to exclude the mean general ocean circulation (e.g. western boundary currents such as the Gulf Stream and Kuroshio) and allow us to focus on current and sea level variations about the mean state. This type of perturbation approach has been used successfully in previous studies of global M_2 internal tides using realistic ocean models (e.g. Niwa and Hibiya 2001). Given surges have similar frequencies to the M_2 tides, the same perturbation approach is used in this study to predict surges in the presence of density stratification. Details are given below.

2.1 Governing equations

The governing equations of our model are as follows:

$$\begin{aligned} \frac{\partial u'_h}{\partial t} + u' \cdot \nabla u'_h + f \times u'_h \\ = -\nabla_h \left[g\eta' + \frac{p_a}{\rho_0} + g \int_z^0 \epsilon ds \right] \\ + A_h \nabla_h^2 u'_h + \frac{\partial}{\partial z} \left(A_z \frac{\partial u'_h}{\partial z} \right) - c_{iw} \bar{u}'_h \quad (1) \end{aligned}$$

$$\nabla \cdot u' = 0 \quad (2)$$

$$\frac{\partial T}{\partial t} + \nabla \cdot (Tu') = K_h \nabla_h^2 T + \frac{\partial}{\partial z} \left(K_z \frac{\partial T}{\partial z} \right) - r(T - T_b) \quad (3)$$

$$\frac{\partial S}{\partial t} + \nabla \cdot (Su') = K_h \nabla_h^2 S + \frac{\partial}{\partial z} \left(K_z \frac{\partial S}{\partial z} \right) - r(S - S_b) \quad (4)$$

Variables with primes denote perturbations about the background mean state. Subscript h refers to horizontal. u'_h is the horizontal velocity vector (u', v') whereas u' is the full three-dimensional velocity vector (u', v', w'). f is the Coriolis parameter multiplied to the upward unit vector. η' is the sea surface fluctuation and p_a is atmospheric pressure at the surface. ρ_0 is a reference density (1025 kg m^{-3}). We define $\epsilon = \rho'/\rho_0$ where ρ' is the perturbation of density about the background state, determined by an equation of state (not shown). The subgrid-scale physics are parameterized by mixing terms as described below. Following Carrère and Lyard (2003), the linear damping term $c_{iw} \bar{u}'_h$ is included to parameterize subgrid-scale internal wave drag acting on

the barotropic motion (\bar{u}'_h is the depth-averaged horizontal velocity vector). The temperature (T) and salinity (S) are maintained by restoring terms in (3) and (4) where r is the restoring coefficient and T_b and S_b are the background mean states to which T and S are weakly restored. The rest of the notation is standard.

2.2 Model setup

The governing equations are solved numerically using the Nucleus for European Modeling of the Ocean (NEMO) framework (Madec 2008). A mode-splitting technique is used to calculate the fast barotropic mode efficiently.

The model domain is global apart from a closed boundary at 78°S. A tri-polar ORCA grid with a nominal horizontal resolution of 1/12° is used along with the ORCA12 bathymetry V3 (available, with documentation, from <http://servdap.legi.grenoble-inp.fr/meom>). The minimum and maximum water depths are taken to be 10 and 6000 m, respectively, in part to avoid numerical instability. The vertical grid has 19 z-levels with the level spacing increasing from approximately 15 m near the surface to 900 m in the deep ocean. Partial cells are used for the bottom cells to represent the bathymetry more accurately.

A modified UNESCO formula (Jackett and McDougall 1995) is used to calculate the density ρ and its perturbation ρ' . To include the effect of density stratification on vertical mixing, the vertical eddy viscosity and diffusivity coefficients are computed using the TKE-type turbulent closure model introduced by Gaspar et al. (1990) and subsequently improved by Madec et al. (1998).

The coefficient c_{iw} is calculated following the formulation of Jayne and St. Laurent (2001), and its global distribution is given by Kodaira et al. (2016). Internal wave drag is only applied in areas deeper than 1000 m and so its effect on storm surge prediction is limited. The drag is however important in suppressing basin-scale barotropic motions caused by wind stress. The bottom boundary condition includes quadratic bottom friction with a constant drag coefficient of 2.5×10^{-3} .

The time steps for the internal and external mode calculations are 180 and 6 s, respectively. The horizontal viscosity and diffusion coefficients are $A_h = 100 \text{ m}^2 \text{ s}^{-1}$ and $K_h = 10 \text{ m}^2 \text{ s}^{-1}$. The restoration coefficient for temperature and salinity is $r^{-1} = 5$ days.

The model was run for 3 months, from August 1, 2014 to October 30, 2014. The model results for the first 2 weeks are not used in the present study. Due to the large amount of model output, only sea surface height, steric sea level (e.g. Marshall and Plumb 1965) defined here as $\eta_s = -\int_{-h}^0 \epsilon dz$, and bottom stress were stored for each forecast hour.

2.3 Initial conditions

The ocean is assumed to be initially at rest. Initial conditions for temperature and salinity were interpolated from the 1/4° annual climatology of the World Ocean Atlas 2013 (Locarnini et al. 2013; Zweng et al. 2013). Although the original climatology is statically stable, this is not necessarily true of the interpolated fields. For this reason, the interpolated temperature and salinity fields were initially diffused for 2 days without advection. If unstable density profiles were detected, the vertical diffusivity was increased to $5 \text{ m}^2 \text{ s}^{-1}$. During the diffusion process, the horizontal diffusion coefficient was set to $2 \times 10^3 \text{ m}^2 \text{ s}^{-1}$, which mixes the ocean by approximately 25 km over 2 days. The diffused seasonal climatologies are then used as the background temperature and salinity fields (T_b and S_b) to which the model is restored.

The model run that is initialized and restored to the diffused climatology is baroclinic and henceforth referred to NEMO-Bc. The model with constant temperature and salinity is referred to as NEMO-Bt because it is barotropic.

2.4 Atmospheric forcing

Following Bernier and Thompson (2015), the magnitude of the wind stress is assumed to be of the form $C_d(W)W^2$ where W is the wind speed and the drag coefficient is

$$10^3 C_d = \begin{cases} 1.2 & W < 11 \text{ m s}^{-1} \\ 0.49 + 0.065W & W \geq 11 \text{ m s}^{-1} \end{cases} \quad (5)$$

The surface air pressures and winds used to drive the model came from the control run of the global ensemble atmospheric prediction systems (GEPS) developed by Environment Canada (Houtekamer et al. 2014). The initial conditions for GEPS are the result of a recently updated ensemble Kalman filter. GEPS operates with a grid spacing of approximately 66 km. Forecast runs are renewed every 12 h, and the prediction continues for 10 days for each model run.

To concatenate the time sequence of GEPS forecasts of wind and air pressure, temporal blending is applied to the first five forecast hours of the atmospheric forcing fields as follows:

$$b(t_i) = c(t_i)a_o(t_i) + [1 - c(t_i)]a_n(t_i) \quad (6)$$

where $b(t_i)$ is the blended forecast for hour t_i ($1 \leq t_i \leq 5$), $a_o(t_i)$ and $a_n(t_i)$ are the old and new forecasts, and $c(t_i)$ is the blending coefficient for time t_i (defined by $c = [0.50, 0.35, 0.25, 0.15, 0.05]$). The forecasts, and the blending step, are carried out every 12 h as indicated by the schematic shown in Fig. 1. The same blending was applied by Bernier and Thompson (2015). It was used to avoid

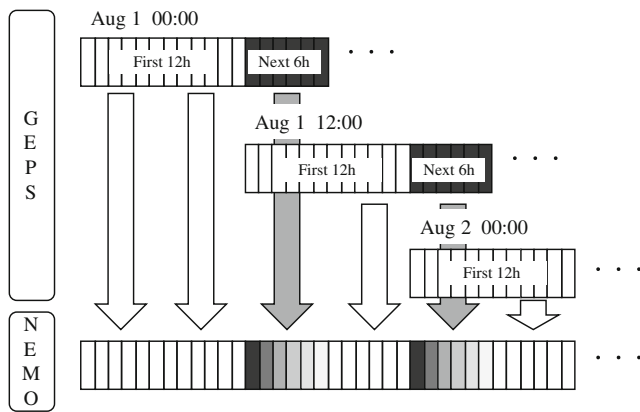


Fig. 1 Schematic of blending process of the atmospheric forcing (wind stress and pressure) from GEPS. The dates denote the starting time of each GEPS run. The large grey arrows show the blending process between adjacent runs. (see text for details)

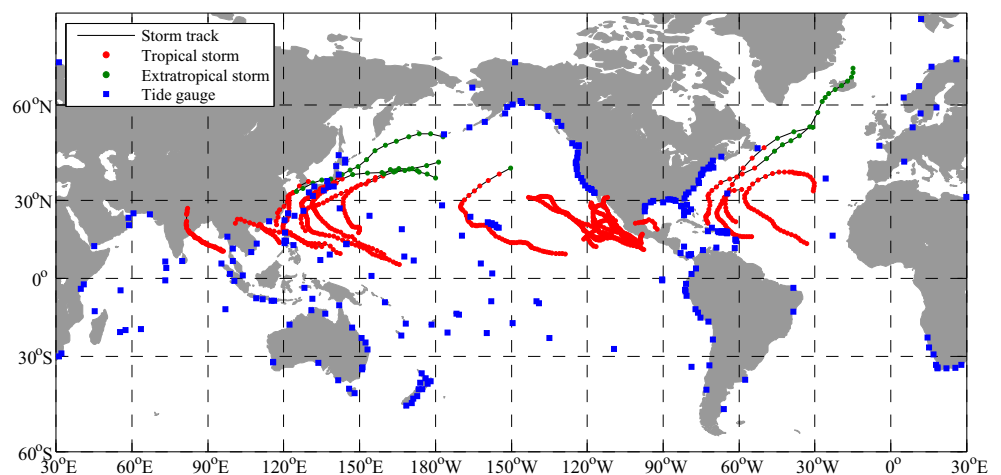
shocks to the surge model when adjacent atmospheric forecasts differ significantly. The blending ensures a smooth transition between the surge forecasts which can be sensitive to the perturbations applied to the atmospheric model.

3 Observations

3.1 Tropical storms

The International Best Track Archive for Climate Stewardship (IBTrACS) provides an inventory of reported tropical storms worldwide (Knapp et al. 2010). The position, maximum sustained winds, central pressure, storm name, radius of maximum winds and additional properties are provided by multiple agencies. Figure 2 shows the tracks of tropical storms, and their extratropical extensions, observed during the study period. The tracks were used to choose tide gauges to evaluate the storm surge predictions for specific case studies as detailed below.

Fig. 2 Global distribution of storms observed during Fall 2014 based on IBTrACS, and the 257 tide gauge stations used for model evaluation



3.2 Coastal sea levels

Hourly sea levels of research quality were obtained from the Sea Level Centre, University of Hawaii for a globally distributed array of tide gauges. Tide gauges were excluded if (1) the tide gauge is more than 25 km from the nearest model grid point, or (2) observations are missing for more than 10 % of the study period. The result of applying the two selection criteria is a global distribution of 257 tide gauges for model evaluation (Fig. 2).

Tides were removed from each hourly sea level record using the `t.tide` package (Pawlowicz et al. 2002). Hourly observations for all of 2014 were used to determine the amplitude and phase of the tidal constituents for each station. In addition to removing tides, a high pass Butterworth filter was used to suppress variability with period exceeding 20 days. This was done to suppress lower frequency motions not resolved by the NEMO ocean model.

4 Results

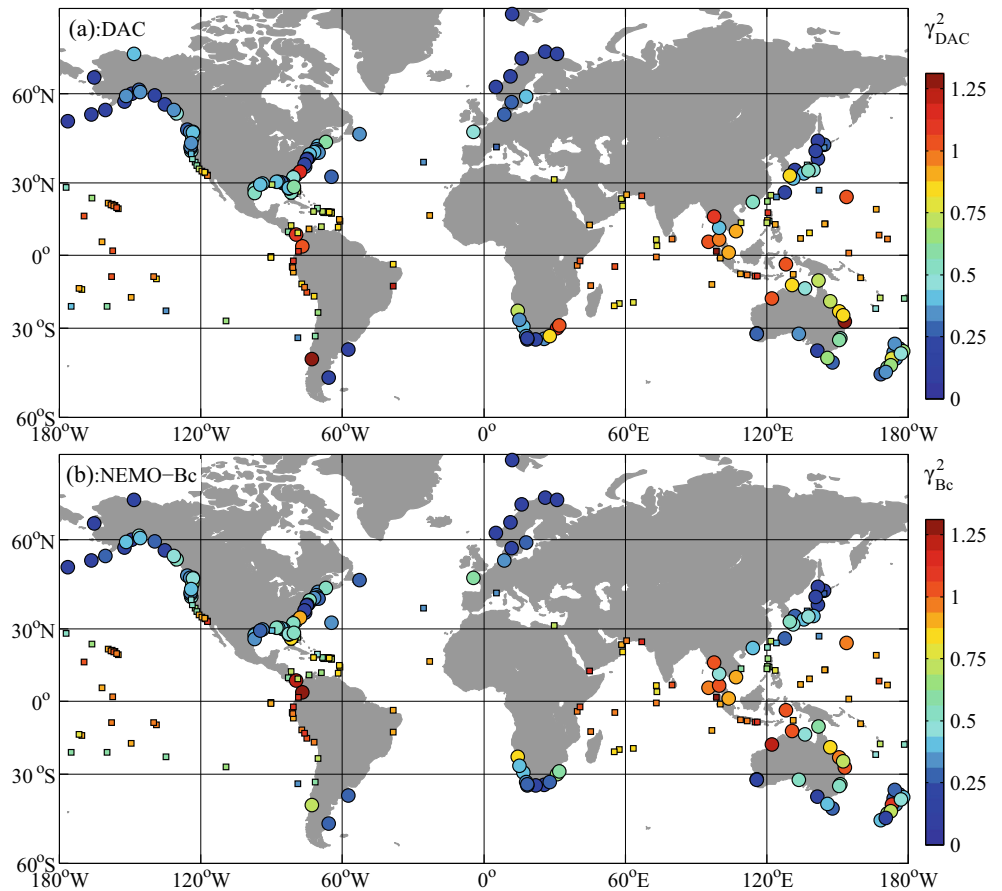
4.1 Global sea level prediction

Sea levels predicted at the grid point nearest to each tide gauge station are used for model evaluation. Before comparison, the predicted hourly sea levels were filtered using the same filtering process applied to the tide gauge records (see Section 3.2). Although the astronomical tidal potential is not included in the model, motions at tidal frequencies can be excited by atmospheric pressure as previously reported by Arbic (2005).

The filtered predictions from August 15 to October 30 are quantitatively assessed using the following γ^2 statistic:

$$\gamma^2 = \frac{\text{Var}(\tilde{\eta}_m - \tilde{\eta}_o)}{\text{Var}(\tilde{\eta}_o)} \quad (7)$$

Fig. 3 Global distribution of γ^2 for **a** the Dynamic Atmospheric Correction based on the model of Carrère and Lyard (2003) and **b** the baroclinic version of the NEMO model. The marker type indicates the standard deviation of $\tilde{\eta}_o$ for each tide gauge; standard deviations larger(smaller) than 5 cm are shown by a *circle (square)*. Low values of γ^2 indicate good model fit



where the tilde denotes a filtered quantity (see Section 3.2), subscripts m and o refer to model predictions and observations respectively and Var denotes the variance operator.

In general, more accurate predictions have smaller values of γ^2 . For example, if $\gamma^2 = 0$, the model predicts the observed signal perfectly and if $\gamma^2 = 1$ the

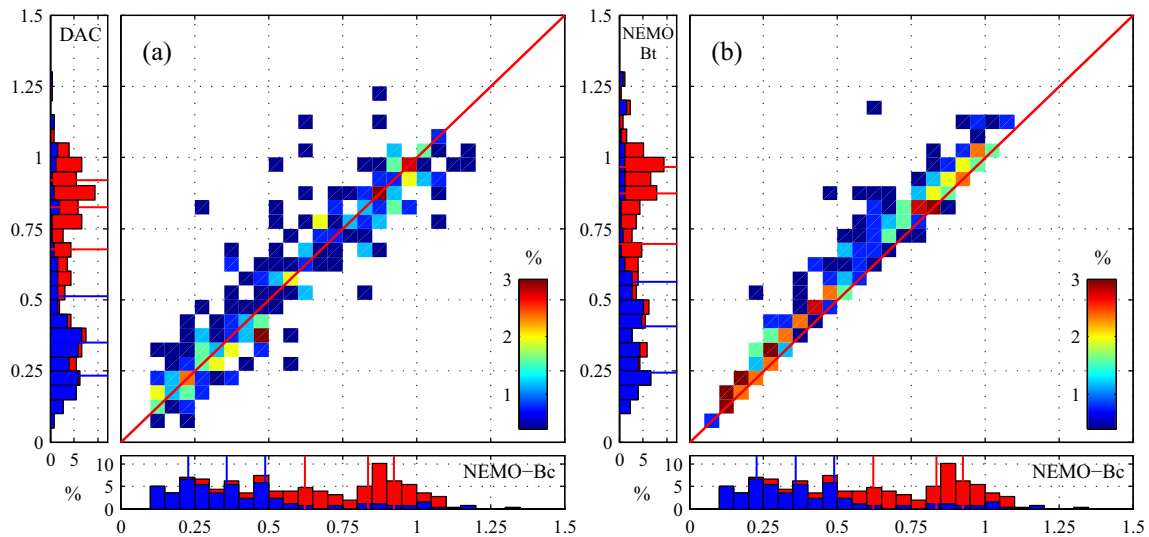
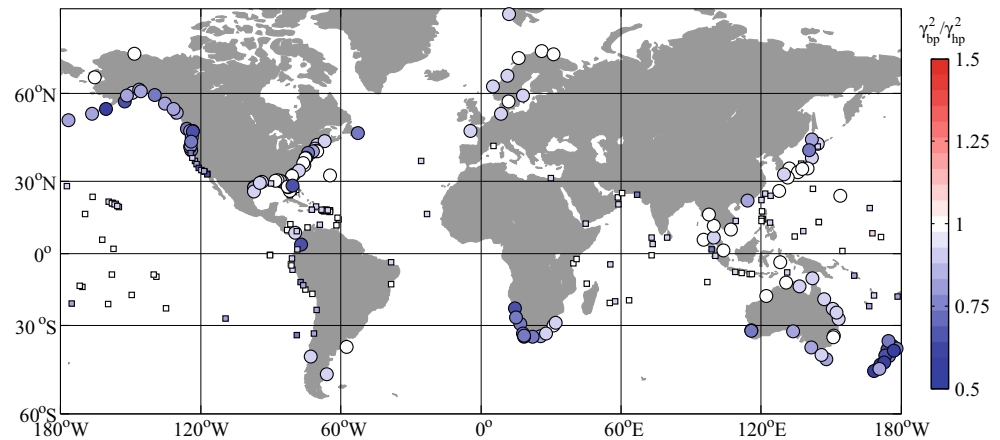


Fig. 4 Comparison of γ^2 for **a** NEMO-Bc and DAC and **b** NEMO-Bc and NEMO-Bt. The main panels show bivariate histograms for all 257 tide gauge stations. The *red line* is the 1:1 line. The side panels show the marginal histograms of γ^2 . The *blue* and *red* histograms are

for stations with the relatively large (≥ 5 cm) and small (< 5 cm) standard deviations of $\tilde{\eta}_o$, respectively. The *blue* and *red* lines in each histogram show the first quartile (25 % below this value), the median, and the third quartile (25 % above this value) for each group

Fig. 5 Effect of removing variations with periods less than 6 h on the fit of the NEMO-Bc predictions. Global distribution of the ratio of γ^2 based on the band-pass filtered predictions and γ^2 based on the high pass filtered predictions (i.e. $\gamma_{bp}^2/\gamma_{hp}^2$). Blue shading means the removal of the high frequency variability has improved the model fit



prediction is no better than constant. Note γ^2 can exceed 1.

The global distributions of γ^2 from the baroclinic run (NEMO-Bc) and the global surge model of Carrère and Lyard (2003) are shown in Fig. 3. In general, γ^2 is less than 0.5 poleward of 30°. Within 30° of the equator, relatively large γ^2 values are evident but we note the standard deviation of $\tilde{\eta}_o$ is less than 5 cm at most of these tide gauge stations. We also note that the model will not resolve well the bathymetry in the vicinity of small islands and so we expect to find large discrepancies in the model predictions (and thus large γ^2) for such locations. Large γ^2 values are also evident along the west coast of South America, between the Indian Ocean and Pacific Ocean and the north coast of Australia.

The γ^2 for NEMO-Bc and the DAC are compared in Fig. 4a. The marginal histograms for stations with large (≥ 5 cm) and small (< 5 cm) standard deviations of $\tilde{\eta}_o$ are fundamentally different for both model runs. The histograms for stations with small standard deviations indicate relatively poor predictions: the median γ^2 values are approximately 0.83 for both runs. For large standard deviations of $\tilde{\eta}_o$, the fits are much better: the median (interquartile) of γ^2 for NEMO-Bc and the DAC are 0.36 (0.26) and 0.35 (0.28), respectively.

The DAC is based on six-hourly atmospheric forcing whereas NEMO-Bc is forced hourly. This means that the atmospheric forcing for NEMO contains more high-frequency variability. To reduce this discrepancy, we filtered the observed and predicted time series to eliminate variability with period less than 6 h and examined the effect on predictive skill.

The high pass filtering led to a decrease in γ^2 (i.e. increase in skill) at 88 % of the tide gauge stations. The effect is clearly evident in the global distribution of the γ^2 ratio (Fig. 5). Removing variability shorter than 6 h has significantly improved the fit along the west coast of North America, the south coasts of Africa and Australia and

New Zealand. One possible explanation is that the blending of the atmospheric forecasts (see (6)) was not completely effective leading to spurious high frequency sea level variability. Another possibility is the limited horizontal resolution of GEPS. This results in the failure of NEMO-Bc to accurately predict higher frequency signals associated with relatively small spatial scales. (Note that the filtering used to remove variability with period less than 6 h is not used in the rest of the paper.)

4.2 Selected case studies

We now focus on the ability of the model to predict peak surges and their arrival times for a selection of tropical storms listed in IBTrACS. Although more than 20 tropical storms were observed during the study period, an associated storm surge was not always observed by the global array of tide gauges. More specifically, five storms were detected in

Table 1 Peak surges (in cm) for the case studies defined by four tropical storms listed in IBTrACS and observed by coastal tide gauge stations during the study period

Storm	Station	Date	Obs	Bc	Bt
Kalmaegi	Hong Kong	Sep 11	76.9	66.8(4)	59.0(4)
Phanfone	Kushimoto	Oct 5	60.5	30.4(1)	27.8(1)
Vongfong	Naha	Oct 11	64.7	39.3(1)	35.7(1)
Vongfong	Nagasaki	Oct 12	33.9	29.7(0)	25.9(1)
Vongfong	Kushimoto	Oct 13	37.3	29.5(0)	28.3(0)
Gonzalo	St. John's	Oct 19	54.2	35.1(0)	35.1(0)
Gonzalo	St. John's	Oct 19	61.6	27.4(1)	27.4(1)

The number in parentheses shows the delay (in hours) in the arrival time of the predicted peak surge compared to the time of the observed peak. There are two records for Gonzalo because two peaks are observed at St. John's (see Fig. 8). The columns labelled Bc and Bt refer to predictions by the baroclinic and barotropic versions of NEMO respectively

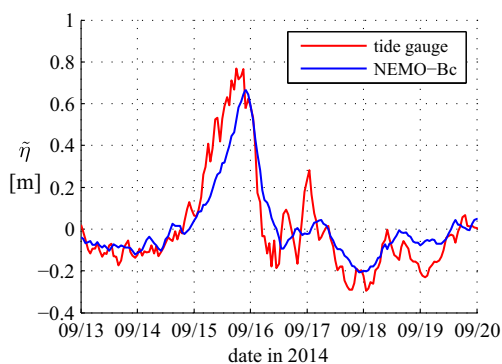


Fig. 6 The observed and predicted storm surge (in m) at Hong Kong caused by Kalmaegi in September 2014

the North Atlantic but only Gonzalo moved close enough to shore to generate a surge that was observed by the tide gauge station at St. John’s (Canada). In the eastern Pacific, 10 storms were detected but storm surges were not observed because of the lack of tide gauges along the southwest coast of North America. The tide gauge at San Diego is closest to the tracks of these storms but its maximum $\tilde{\eta}_o$ was only 12 cm. In the western Pacific, eight storms were detected and surges associated with Kalmaegi, Phanfone and Vongfong were observed. In the Indian Ocean, only one storm was detected but there were no tide gauge stations in the vicinity of its path.

In total, seven storm surges were detected in the sea level records and they are listed in Table 1. The listed peak surge and its arrival time were determined using the local maximum of $\tilde{\eta}_o$. The observed storm surges produced by Kalmaegi, Phanfone, Vongfong and Gonzalo were all smaller than one meter (Table 1). The peak surges predicted by NEMO-Bc are consistently smaller than the observed peak surges. The relatively coarse resolution of the ocean model, and its minimum water depth of 10 m, will contribute to the underestimation of the peak surges.

The underestimation is also due in part to the coarse resolution of GEPS (66 km) leading to underestimation of storm intensity.

The arrival time of the peak surge predicted by NEMO-Bc is in good agreement with the observations except for the surge generated by Kalmaegi (Table 1). The predicted surge arrived at Hong Kong 4 h after the observed peak (Fig. 6). The later arrival of the predicted surge is not only because of high frequency fluctuations near the peak surge but also because of an inaccurate prediction of the shape of the surge.

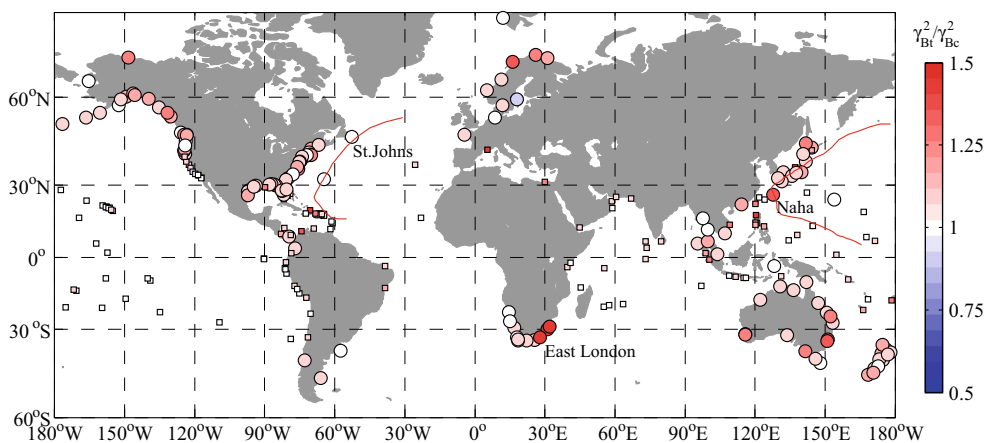
4.3 The effect of density stratification

We first compare γ^2 values from NEMO-Bc and NEMO-Bt. When the density stratification is removed from the model, the fit is deteriorated at almost all stations (Figs. 4b and 7). An alternative metric to γ^2 is the squared correlation (r^2) between observed and predicted sea levels for the same location. This statistic is more forgiving because it allows an arbitrary scaling of the observations and predictions. Using r^2 led to the same conclusion: the inclusion of density stratification increases the predictive skill of the model.

The improvement in model fit is also evident in the predictions of peak surges by the baroclinic, compared to the barotropic, model (Table 1, see columns labelled Bc and Bt respectively). The arrival time of the peak surges however remains almost unchanged.

The three stations labelled in Fig. 7 were selected for further examination of the effect of density stratification. One station is East London (South Africa) because the change in γ^2 is the largest. The other two stations are Naha (Japan) and St. John’s (east coast of Canada) because storms passed over these stations (Fig. 7). We first examine time series for each location, and then the spatial distributions of predicted sea levels to clarify the difference in the predictions of NEMO-Bc and NEMO-Bt. (The filtering described in

Fig. 7 Global distribution of ratio of γ^2 from NEMO-Bt and NEMO-Bc ($\gamma_{Bt}^2/\gamma_{Bc}^2$). Red shading implies the addition of density stratification has improved model fit. The red lines show the tracks of Gonzalo and Vongfong, respectively



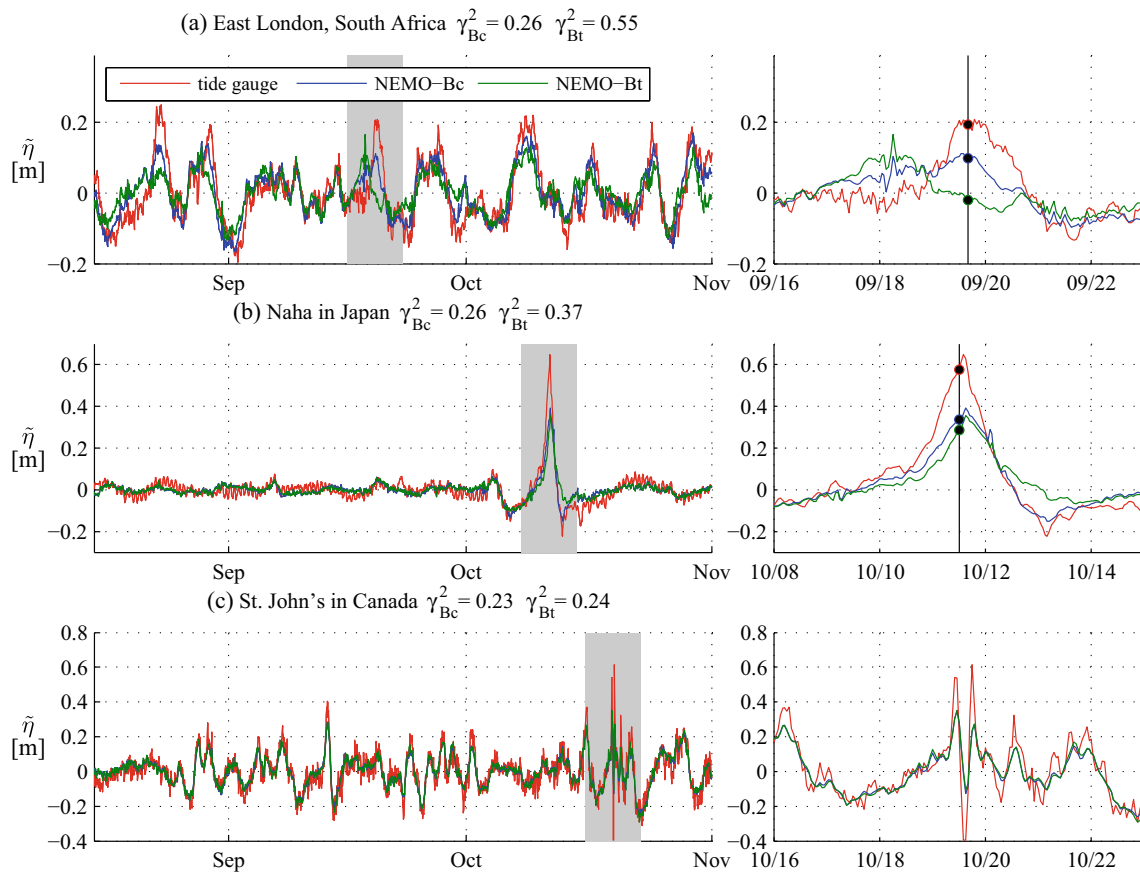


Fig. 8 Time series of observed and predicted sea levels (in m) from August 15 to October 30. **a** East London in South Africa, **b** Naha in Japan and **c** St. John's in Canada. The shaded area in the left panel in

each row is enlarged in the right panel. The black line and dots in the right panels (**a**, **b**) indicate the time of spatial distributions shown in Fig. 11 and Fig. 12

Section 3.2 was not applied to the model output before mapping their spatial distributions.)

EAST LONDON: Time series of filtered sea level show that NEMO-Bc consistently provides better predictions than NEMO-Bt (Fig. 8a). The exclusion of density stratification increases γ^2 from 0.26 to 0.55.

In order to investigate the physical reasons for the different predictions by NEMO-Bc and NEMO-Bt, predicted sea level changes along the coast of South Africa are examined (Fig. 9). The Hovmöller diagram (Fig. 10a) shows the frequent generation of signals that propagate with the coast on their left at a speed of about 4 m s^{-1} , consistent with coastal trapped wave theory. The Hovmöller diagram of the difference between NEMO-Bc and NEMO-Bt (Fig. 10b) shows that NEMO-Bt attenuates the propagation of the signal as it encounters the rapid narrowing of the shelf at $s = 1500 \text{ km}$ (Fig. 9). The phase speed slows for $s > 1500 \text{ km}$, consistent with the expected effect of density stratification on coastal trapped wave propagation (e.g. Gill 1982).

Snapshots of sea level are shown in Fig. 11 for 16:00 on September 19, 2014 when the largest differences in the model predictions is found. It is clear that the sea level

signal is coastally trapped and the baroclinic run predicts stronger transmission past the shelf's rapid narrowing at

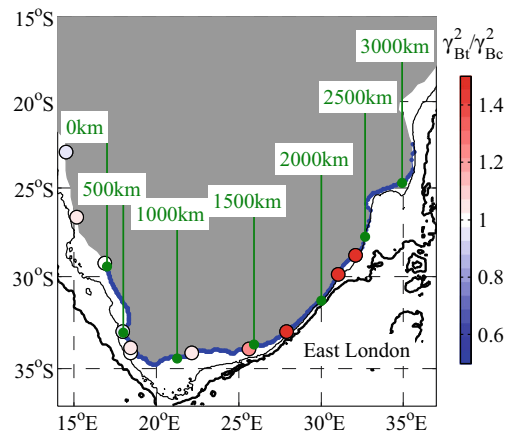
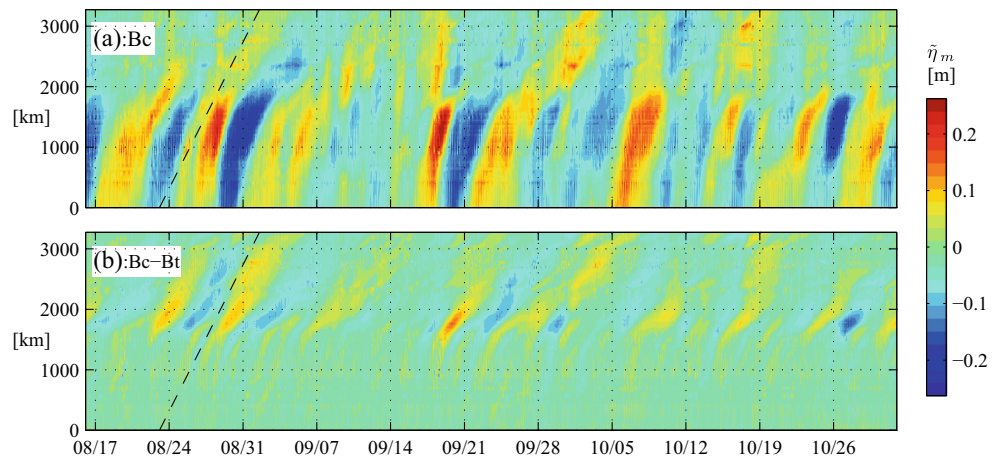


Fig. 9 The along-shore coordinate (s) defined along the coast of South Africa (blue). The green dots and labels show position along the coast. The circles indicate location of tide gauge stations with the color showing the ratio of γ^2 for NEMO-Bt to NEMO-Bc ($\gamma_{Bt}^2 / \gamma_{Bc}^2$). The thin and thick black lines show the 200 and 2000 m isobaths, respectively

Fig. 10 Hovmöller diagrams of the filtered sea levels along the coast of South Africa (Fig. 9). **a** NEMO-Bc, **b** the difference between NEMO-Bc and NEMO-Bt (Bc-Bt). The *black dashed lines* indicate a propagation speed of 4 m s^{-1}



$s = 1500 \text{ km}$. This is consistent with the expected effect of density stratification on coastal trapped wave propagation (e.g. Wilkin and Chapman 1990).

NAHA: This tide gauge detected a storm surge associated with tropical storm Vongfong as it passed almost directly over the station (Fig. 7). The rapid increase of $\tilde{\eta}$ on October 11 indicates the passage of Vongfong (Fig. 8b). The inclusion of realistic density stratification results in better prediction with a slightly larger storm surge. The difference is more significant after the peak surge because NEMO-Bc predicts the subsequent depression of sea level.

Snapshots of sea level for 12:00 on October 11, 2014, when the storm was directly overhead, are shown in Fig. 12. As expected, the large (order 0.1 m) difference in deep water between the barotropic and baroclinic runs is due to steric changes (not shown). The larger peak surge predicted by the baroclinic model in the shallow water in the vicinity of the tide gauge is likely due to a reduction in bottom stress due to baroclinicity.

ST JOHN'S: Differences between the two model runs are not always evident. For example, the observed time series for St. John's indicates a large surge caused by Gonzalo (Fig. 8c). However, no significant difference between the

two model predictions can be seen. This indicates both models predict a barotropic response, consistent with the rapid translational speed of this particular storm over the wide and shallow Grand Banks offshore of this tide gauge station.

5 Summary and Discussion

A hindcast study of global storm surges has been carried out for Fall 2014 using a baroclinic ocean model based on the NEMO framework. The model has 19 vertical levels and a horizontal resolution of $1/12^\circ$. It was forced by hourly atmospheric wind and pressure forecast fields provided by Environment and Climate Change Canada.

Based on the γ^2 statistic, the present model provides predictions of global water level that are almost as accurate as the DAC. The model has some difficulty in predicting sea level fluctuations shorter than 6 h; filtering out this high frequency variability improves significantly the fit of the model. The possible explanations include inaccurate prediction of high frequency/small-scale atmospheric and oceanic features by GEPS and NEMO-Bc, and the blending process used to generate the atmospheric forcing.

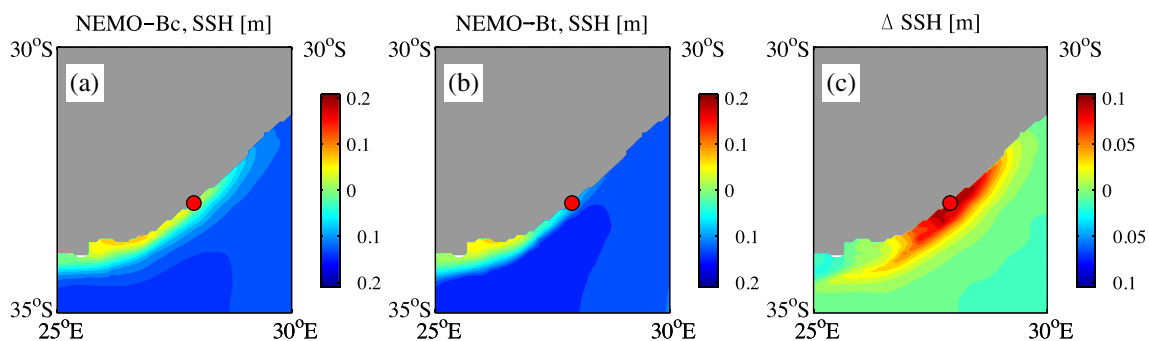


Fig. 11 Spatial distribution of sea level predictions by NEMO-Bc and NEMO-Bt, and their difference (NEMO-Bc - NEMO-Bt), near the East London tide gauge for 16:00 on September 19, 2014. Sea surface

height from **a** NEMO-Bc, **b** NEMO-Bt and **c** their difference. The tide gauge station is shown by the *red dot* in each panel

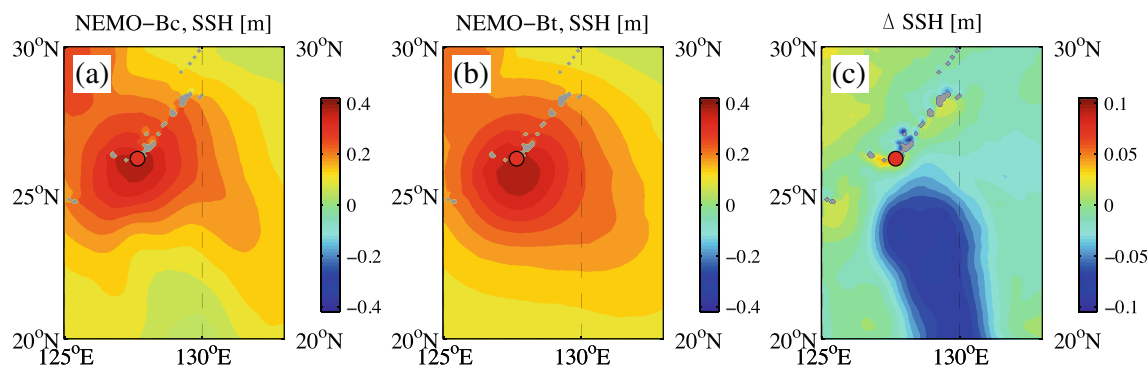


Fig. 12 Same as Fig. 11 but for sea level distributions near Naha at 12:00 on October 11

During the study period, more than 20 tropical storms appeared in the IBTrACS archive. However, the global array of 257 tide gauge stations detected storm surges generated by only four of these storms. The arrival time of the peak surge predicted by the model was typically within 2 h of the observed peak arrival time. The model consistently underestimates the peak surges by about 20 cm.

The effect of density stratification, the main focus of the study, was examined by comparing barotropic and baroclinic runs of the model. It was shown that the inclusion of density stratification generally improved the surge predictions as measured by γ^2 . The inclusion of density stratification however only slightly improved (by several cm) the prediction of peak surges. Case studies suggest that the physical reasons for the improvement include the effect of density stratification on the propagation and transmission of coastal trapped waves, and changes in bottom stress associated with changes in the vertical structure of the currents.

An interesting possibility raised by one reviewer is that a depth-averaged barotropic model may fit the observations as well as the 3D baroclinic model used in this study (NEMO-Bc). The reason is that the dynamical response of the shelf and near shore is represented differently in these two types of model. For this reason, we modified NEMO-Bt to have only one layer (henceforth NEMO-2D). Note NEMO-2D uses partial cells that best fit the bathymetry.

The γ^2 for NEMO-2D and NEMO-Bt were similar (not shown), especially for tide gauges with relatively large tidal residuals. In general, the γ^2 for NEMO-Bc remained the lowest. We note however that the implication that 2D models will not perform as well as baroclinic models needs to be checked using 2D models with higher horizontal resolution and smaller minimum depth. (NEMO-Bc assumes a minimum depth of 10 m.)

With regard to future work, we plan to optimize the choice of wind drag coefficients and also extend the study period. A longer test period will increase the number of storm surges for analysis and make the model evaluation more comprehensive and definitive. Another important next

step is to better understand the physical reasons for the increased skill resulting from the inclusion of density stratification. This will require examination of more case studies and the possible use of higher resolution, nested baroclinic models of regions of particular interest (e.g. South Africa, west of North America). Our longer term goal is to include tides in the surge model and fully utilize GEPS to develop a global ensemble prediction system for total water level.

Acknowledgments The authors thank two anonymous reviewers and the editor for their exceptionally thorough reviews and constructive comments on an earlier draft of the manuscript. We acknowledge Environment and Climate Change Canada for providing access to computing resources. The authors acknowledge financial support from the Marine Environmental Observation Prediction And Response (MEOPAR) network of centres of excellence.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix MOG2D-G and the DAC

The existing numerical model most relevant to the present study is MOG2D-G, a global barotropic model using finite element space discretization (Carrère and Lyard 2003). The grid size ranges from 20 km in the shallow ocean to 400 km in the deep ocean. The model is forced by 6 hourly atmospheric pressure and wind fields predicted by European Centre for Medium-range Weather Forecasts (ECMWF) analysis.

The predicted signals longer than 20 days are not used and are replaced by estimates based on the inverse barometer effect. After the replacement, the predicted fields are made available online as the so-called Dynamic Atmospheric Correction (DAC). The available period of DAC and the data acquisition method are described by Pascual et al.

(2008). We obtained DAC data on a regular $1/4^\circ$ grid every 6 h. After the 6 hourly data were interpolated to hourly values they were filtered in the same way as the observations (Section 3.2).

References

- Arbic BK (2005) Atmospheric forcing of the oceanic semidiurnal tide. *Geophys Res Lett* 32(2):L02610
- Bernier NB, Thompson KR (2015) Deterministic and ensemble storm surge prediction for Atlantic Canada with lead times of hours to ten days. *Ocean Model* 86:114–127
- Bode L, Hardy TA (1997) Progress and recent developments in storm surge modeling. *J Hydraul Eng* 123(4):315–331
- Carrère L, Lyard F (2003) Modeling the barotropic response of the global ocean to atmospheric wind and pressure forcing—comparisons with observations. *Geophys Res Lett* 30(6):1275
- Clarke AJ (1977) Observational and numerical evidence for wind-forced coastal trapped long waves. *J Phys Oceanogr* 7(2):231–247
- Davies AM (1988) On formulating two-dimensional vertically integrated hydrodynamic numerical models with an enhanced representation of bed stress. *J Geophys Res* 93(C2):1241–1263
- Flather RA (2000) Existing operational oceanography. *Coastal Eng* 41(13):13–40
- Flowerdew J, Horsburgh K, Wilson C, Mylne K (2010) Development and evaluation of an ensemble forecasting system for coastal storm surges. *Quart J Roy Meteor Soc* 136(651):1444–1456
- Gaspar P, Grgoris Y, Lefevre J-M (1990) A simple eddy kinetic energy model for simulations of the oceanic vertical mixing: tests at station Papa and long-term upper ocean study site. *J Geophys Res* 95(C9):16179–16193
- Gill AE (1982) *Atmosphere-Ocean Dynamics*. International Geophysics Series. Academic Press, New-York
- Gordon RL (1982) Coastal ocean current response to storm winds. *J Geophys Res* 87(C3):1939–1951
- Heaps NS (1983) Storm surges, 1967-1982. *Geophys J Roy Astron Soc* 74(1):331–376
- Hearn C, Holloway P (1990) A three-dimensional barotropic model of the response of the Australian north west shelf to tropical cyclones. *J Phys Oceanogr* 20(1):60–80
- Houtekamer PL, Deng X, Mitchell HL, Baek S-J, Gagnon N (2014) Higher resolution in an operational ensemble Kalman filter. *Mon Wea Rev* 142(3):1143–1162
- Jackett DR, McDougall TJ (1995) Minimal adjustment of hydrographic profiles to achieve static stability. *J Atmos Oceanic Technol* 12(2):381–389
- Jayne SR, St. Laurent LC (2001) Parameterizing tidal dissipation over rough topography. *Geophys Res Lett* 28(5):811–814
- Jeselnianski CP, Chen J, Shaffer WA (1992) SLOSH: Sea, lake, and overland surges from hurricanes. US Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service
- Knapp KR, Kruk MC, Levinson DH, Diamond HJ, Neumann CJ (2010) The International Best Track Archive for Climate Stewardship (IBTrACS). *Bull Am Meteorol Soc* 91(3):363–376
- Kodaira T, Thompson KR, Bernier NB (2016) Prediction of M_2 tidal surface currents by a global baroclinic ocean model and evaluation using observed drifter trajectories. *J Geophys Res* 121:6159–6183
- Locarnini RA, Mishonov AV, Antonov JI, Boyer TP, Garcia HE, Baranova OK, Zweng MM, Paver CR, Reagan JR, Johnson DR, Hamilton M, Seidov D (2013) *World Ocean Atlas 2013, Volume 1: Temperature*. S. Levitus, Ed., A. Mishonov Technical Ed.; NOAA Atlas NESDIS 73, 40 pp
- Madec G (2008) NEMO ocean engine. Note du Pôle de modélisation. Institut Pierre-Simon Laplace (IPSL), France
- Madec G, Delecluse P, Imbard M, Levy C (1998) OPA 8.1 ocean general circulation model reference manual. Note du Pôle de modélisation
- Marshall J, Plumb RA (1965) *Atmosphere, ocean and climate dynamics: an introductory text, volume 8*. Academic Press
- Minato S (1998) Storm surge simulation using POM and a revisitation of dynamics of sea surface elevation short-term variation. *Oceanograph Lit Rev* 45(7):79–88
- Needham HF, Keim BD, Sathiaraj D (2015) A review of tropical cyclone-generated storm surges: global data sources, observations, and impacts. *Rev Geophys* 53(2):545–591
- Niwa Y, Hibiya T (2001) Numerical study of the spatial distribution of the M_2 internal tide in the pacific ocean. *J Geophys Res* 106(C10):22,441–22,449
- Pascual A, Marcos M, Gomis D (2008) Comparing the sea level response to pressure and wind forcing of two barotropic models: validation with tide gauge and altimetry data. *J Geophys Res* 113(C7):C07011
- Pawlowicz R, Beardsley B, Lentz S (2002) Classical tidal harmonic analysis including error estimates in MATLAB using T_TIDE. *Comput Geosci* 28(8):929–937
- Weisberg RH, Zheng L (2008) Hurricane storm surge simulations comparing three-dimensional with two-dimensional formulations based on an Ivan-like storm over the Tampa Bay, Florida region. *J Geophys Res* 113(C12):C12001
- Wilkin JL, Chapman DC (1990) Scattering of coastal-trapped waves by irregularities in coastline and topography. *J Phys Oceanogr* 20(3):396–421
- Zheng L, Weisberg RH, Huang Y, Luettich RA, Westerink JJ, Kerr PC, Donahue AS, Crane G, Akli L (2013) Implications from the comparisons between two- and three-dimensional model simulations of the Hurricane Ike storm surge. *J Geophys Res* 118(7):3350–3369
- Zweng MM, Reagan JR, Antonov JI, Locarnini RA, Mishonov AV, Boyer TP, Garcia HE, Baranova OK, Johnson DR, Seidov D, Biddle MM (2013) *World Ocean Atlas 2013, Volume 2: Salinity*. S. Levitus, Ed., A. Mishonov Technical Ed.; NOAA Atlas NESDIS 74, 39 pp