

## Operational sea level forecasting in Estonia

Priidik Lagemaa, Jüri Elken and Tarmo Kõuts

Marine Systems Institute at Tallinn University of Technology, Akadeemia tee 15a, 12618 Tallinn, Estonia; priidik26@gmail.com

Received 4 May 2011, in revised form 17 October 2011

**Abstract.** The operational sea level forecasting system in Estonia, based on the HIROMB model forecasts and on 11 online sea level observation stations, is described and validated. The system is operational since 08.08.2005. Statistical analysis for the period 2006–2008 data is performed to investigate the properties of low-frequency sea level error and error estimation of high sea level events. A 7-day backwards moving average filter is the most appropriate for Estonian waters to correct the raw sea level forecast. For the time period 2006–2008 the forecast error of critically high sea level events is  $\pm 25$  cm within  $\pm 3$  h between the observed and forecasted maxima. Taylor skill assessment procedures are applied to the data, covering the period 2009–2011. The data is divided into three forecast sub-ranges (low, medium and high sea level), to investigate the possibilities of the online error estimation. Smaller errors are present for medium and low sub-range and larger errors for the high sub-range. The necessities for further development are outlined.

**Key words:** sea level, water level, modelling, HIROMB, Baltic Sea, Estonia, operational forecasting.

### 1. INTRODUCTION

Sea level variations around the decadal mean values are of great importance for the coastal population. The eastern Baltic Sea has very weak tides but strong and quite often damaging storm surges, combined with other types of sea level variation. Variable isostatic Earth crust vertical motion pattern over the region [1], combined with the global eustatic sea level change [2] results in long-term trends of observed sea level, yielding a maximum decrease  $-8.2$  mm yr<sup>-1</sup> in the northern part of the Bothnian Bay [3], a slight decrease from  $-0.5$  to  $-2.8$  mm yr<sup>-1</sup> in Estonia [4] and a slight increase in the southern part of the sea [5].

Due to the limited transport capacity of the Danish Straits, the variability of the Baltic sea level can be generally decomposed [6] into the external variations,

due to the change in mean sea level of the basins, forced by the outside Kattegat sea level and the basins freshwater budget, and into the internal variations around the mean state, dominated by wind-driven long rotational gravity waves, influenced by the complex coastline and topography. The internal component has also smaller contributions from saline and freshwater pulses and variations in water density and air pressure. Although the external sea level component is in principle oscillatory (the “harbor” or semi-open bay mode [<sup>7</sup>]), at time scales of forcing weather patterns (periods less than a month) the Baltic sea level response is damped and delayed [<sup>6</sup>]. Observed sea levels are well correlated along the coastline [<sup>8</sup>]. Stronger westerly winds and related larger inflows occur usually during autumn and winter. As an example, during the major inflow in January 1993 [<sup>9</sup>], 310 km<sup>3</sup> of North Sea water entered the Baltic in 21 days, raising the mean sea level by 70 cm. Such rapid increases in the sea volume during the storms contribute also to the actual storm surges [<sup>10</sup>].

Internal sea level variations are influenced by semi-isolated seiche systems, Belt Sea–Baltic Proper–Gulf of Finland (passing the Gulf of Bothnia) and Belt Sea–Baltic Proper–Gulf of Bothnia, with self-oscillation periods about 25–27 h and 31–39 h [<sup>11</sup>]. However, due to the complicated multi-basin topography, the seiches do not have high contribution to the sea level power spectrum [<sup>12</sup>]. Jönsson et al. [<sup>13</sup>] have recently discussed that instead of full-basin seiches, the internal sea level variability is dominated by weakly coupled semi-open bay modes. For the Gulf of Riga, such open bay mode results in persistent 24-h current oscillations in the connecting Irbe Strait [<sup>14,15</sup>], while the internal (closed bay) seiche period is about 5 h [<sup>10</sup>].

Short-term prediction of the sea level in marginal seas and coastal areas is a continuous issue of oceanographic research and technological developments. Compared to the earlier semi-empirical forecast methods, a revolutionary approach, numerical modelling based on two-dimensional (2D) shallow water equations, was proposed for sea level modelling and forecasts in the 1950s [<sup>16,17</sup>]. However, first realizations of numerical modelling in the everyday forecast services started only in the 1980s [<sup>18</sup>]. Following the experiences of operational numerical sea level forecasting, based on 2D shallow water equations, Germany’s Federal Maritime and Hydrographic Agency (BSH) implemented a three-dimensional (3D) baroclinic forecast model BSHcmod at the beginning of 1990s [<sup>19</sup>]. That model became a core of the family of operational models, run in addition to BSH also at the Swedish Meteorological and Hydrological Institute (SMHI) and Danish Meteorological Institute (DMI). An overview of different model versions and set-up features has been recently given in [<sup>20</sup>]. The model developments are coordinated by the HIROMB (High-Resolution Operational Model for the Baltic Sea) consortium. A broader Baltic-wide cooperation frame in operational oceanography is provided by the BOOS (Baltic Operational Oceanographic System) [<sup>21</sup>] that takes care, among many other activities, of on-line exchange of observational and modelled data and dissemination of operational products. Further development of the Baltic marine forecasting is

presently going on within the EU-funded MyOcean (Development and pre-operational validation of upgraded GMES Marine Core Services and capabilities) project. It aims at the merging of the best features of different model versions into one harmonized operational model system HIROMB-BOOS, including assimilation of the real-time observational data and launching the forecast component for biochemical state variables.

Introduction of numerical sea level forecasts in Estonia was delayed in the official service level until the highest known storm surge occurred in January 2005 at the western coast of Estonia [<sup>22,23</sup>]. Operational sea level gauge at Pärnu recorded highest sea level height of +275 cm, the highest observed so far since the beginning of instrumental observations in 1923. The coastal towns Pärnu and Haapsalu were heavily flooded and considerable damages and economic loss occurred. The Estonian Ministry of Environment initiated an implementation of numerical ocean forecasting and related observational services in Estonia in cooperation with the Marine Systems Institute (MSI) at Tallinn University of Technology and the Estonian Meteorological and Hydrological Institute (EMHI). The primary goal was a considerable reduction of the errors in short-term forecasting of the extreme sea levels (both the high sea levels causing floods in the coastal areas, and low sea levels stopping the ship traffic between the Estonian mainland and the western larger islands). Numerical sea forecasts are needed also for many other practical problems like harmful algal blooms, drift of surface and subsurface substances and objects, or ice conditions. Therefore it was decided to implement an advanced 3D forecast system.

Among the Baltic-wide oceanographic service providers, SMHI was chosen as a core provider for Estonia, since their operational model HIROMB has the highest horizontal grid resolution (1 NM). The Estonian coastline is very fragmented, with numerous small bays, peninsulas and islands, therefore high horizontal resolution is very important for producing detailed information on sea level, currents, temperature, salinity and ice conditions. The numerical sea level forecasts started in autumn 2005 and the forecast system is presently in a quite mature state. Nevertheless, gradual developments towards the next-generation forecast system are progressing on the national, Baltic-wide and European levels.

The aim of the paper is to give an overview of the present Estonian sea level forecasting system, present the practice of applied forecast production methods and estimate the system performance and statistical accuracy. Following the forecast system description, specific aspects of the low-frequency error correction, examples of forecast performance during extreme sea levels and statistical evaluation of the forecast accuracy along the different coastal areas and sea level variation range are analysed in detail. Finally, conclusions and outlook of further investigations and developments are presented.

## 2. OPERATIONAL FORECASTING SYSTEM DESCRIPTION

### 2.1. The operational model

The operational oceanographic forecast models, belonging to the HIROMB consortium, have been running for the Baltic Sea since the 1990s with the primary purpose of giving short-term (up to 48 or 60 h) predictions of the sea conditions, in order to handle oil spills, storm surges, support navigation etc. The core of the model system is a 3D baroclinic eddy-resolving circulation model, based on the original BSHcmod [<sup>19</sup>] that calculates also currents, temperature, salinity and turbulence in the water column. The model contains also a sea ice module.

The SMHI version, HIROMB (or HIROMB-SMHI) has been running since 1995 in a pre-operational mode, and since 1999 the model is fully operational [<sup>24</sup>]. The model is forced mainly by the data from atmospheric circulation model SMHI-HIRLAM with a horizontal resolution of 22 km and with a 1 h time step. For the freshwater inflow, daily data from the river runoff model HBV is used. At the outer open ocean boundaries a storm surge model (NOAMOD) is used for the water levels together with tides, climatologic salinity and temperature data. The sea surface wind stress is calculated by the common quadratic formulation from the corresponding 10 m height wind speed components  $W_\lambda$  and  $W_\varphi$  as  $\tau_\lambda = \rho_a c_D W_\lambda W_{10}$ ,  $\tau_\varphi = \rho_a c_D W_\varphi W_{10}$ , where  $\lambda$  and  $\varphi$  are longitude and latitude respectively,  $\tau_\lambda$  and  $\tau_\varphi$  are the wind stress components,  $\rho_a$  is the air density,  $W_{10} = \sqrt{W_\lambda^2 + W_\varphi^2}$  is the wind speed, and  $c_D$  is the surface drag coefficient.

During the study period, there had been several model code upgrades, starting from version 3.0 to 4.2. The main parameters of the model's operational versions are presented in Table 1. The model version 3.0, started at 15.11.2005, was set up in three nested grids with horizontal resolution of 12, 3 and 1 NM, respectively. The model domain of highest resolution covers the whole Baltic Sea area with grid step 1' along latitudes and 5/3' along longitudes (Fig. 1a,b). The model presented the Baltic Sea by 16 vertical layers, with 4 m thickness in the upper 12 m and increasing values towards the greater depths. The linear formulation was used to find the surface drag coefficient  $c_D = (0.7 + 0.09W_{10}) \times 10^{-3}$ . The value of bottom friction coefficient  $r$  was 0.0028. In the course of the model development, improved descriptions were introduced for vertical turbulence and drag coefficients on the surface and on the bottom [<sup>20,25-27</sup>].

Starting from the operational model version 3.3 (upgraded at 18.09.2007), the breaking surface waves, the water-ice roughness parameter and the variable surface drag coefficient were introduced. The surface drag formulation was changed to account for the atmospheric stability. Actually, the improved surface drag formulation was introduced already in model version 3.1, but this version never got operational and the new formulation was delivered to the users with version 3.3. To improve the representation of surface current velocities, the neutral surface drag was defined as  $c_D = 1.3 \times 10^{-3}$  at  $W_{10} < 8$  m/s and  $c_D = (0.84 + 0.058W_{10}) \times 10^{-3}$

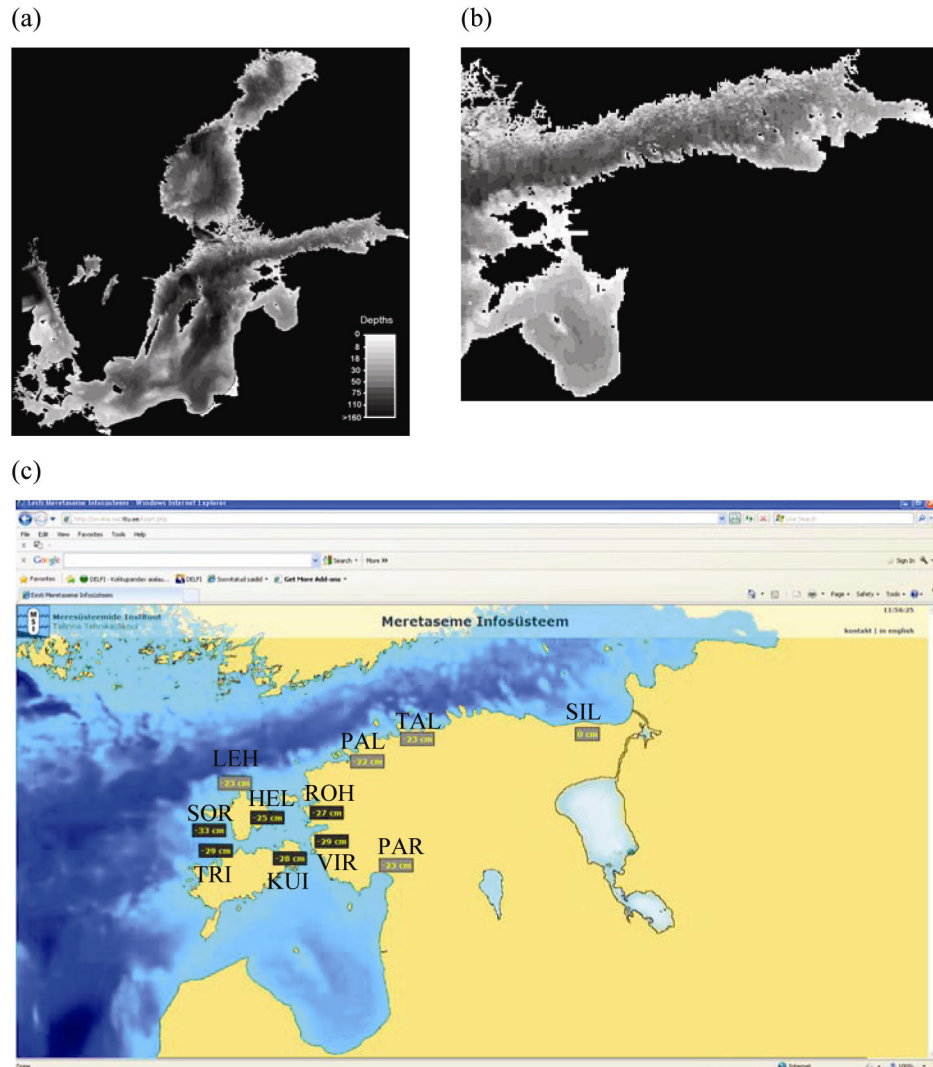
**Table 1.** Operational versions of HIROMB and changes affecting the sea level

| Version                           | 3.0   | 3.3        | 4.0  | 4.2        |
|-----------------------------------|---|------------|--|------------|
| Release date                      | 15.11.2005  | 18.09.2007 | 08.12.2009   | 15.09.2010 |
| Horizontal grid resolution, NM    | 1; 3; 12  |            | 1; 3   |            |
| Vertical grid resolution, m       | 24 layers (3 × 4, 3 × 6, 3 × 10, 2 × 15, 2 × 20, 2 × 30, 2 × 40, 7 × 60)          |            | 50 layers (20 × 4, 5 × 5, 3 × 6, 2 × 7, 8, 9, 10, 11, 14, 17, 21, 26, 32, 39, 9 × 40)                          |            |
| Bottom friction coefficient       | $r = 0.0028$  |            | $c_{D,\text{bottom}} = \frac{\kappa^2}{\log^2\left(\frac{h_b}{2z_{0,b}} + 1\right)}, z_{0,b} = 0.01 \text{ m}$ |            |
| Meteorological forcing            | HIRLAM (SMHI), horizontal resolution 22 km till 30.05.2011 after that date 11 km. |            |  |            |
| Wind drag coefficient formulation | $c_D = (0.7 + 0.09W_{10}) \times 10^{-3}$   |            | $c_D = 1.3 \times 10^{-3}$ if $W_{10} < 8$<br>$c_D = (0.84 + 0.058W_{10}) \times 10^{-3}$ if $W_{10} \geq 8$   |            |

in case of higher values of  $W_{10}$ . The drag coefficient was then slightly modified, according to the stability of the atmosphere, as calculated from the air-water temperature difference. Other modifications in the model code were not relevant to the sea level. The bottom friction coefficient, grid resolution and forcing (i.e. atmospheric, open sea boundary, rivers) parameters remained unchanged.

Starting from 9.07.2008 the model run frequency was increased from 2 to 4 times per day, i.e. the model was re-run every 6th hour whereas all other model parameters remained unchanged. However, the forecasting system still used only the +00 h forecast files until 17.12.2008 when the increased re-run frequency was introduced to the Estonian forecasting system.

With the introduction of version 4.0, upgraded officially on 08.12.2009, the vertical resolution of the model improved significantly and the coastline was also improved at some regions. Now the model presents the Baltic Sea by 50 vertical layers, with the layer's thickness of 4 m in the upper 80 m, and slowly increasing towards the greater depths. The 3 NM grid boundaries were extended to 65°53'30"N, 4°9'10"W, giving nearly the same coverage as the earlier 12 NM grid. The coverage of the 1 NM grid remained unchanged. The formulation of the surface drag coefficient was not changed, but the local bottom drag coefficient formulation was introduced to calculate the momentum transfer from water to bottom, now enabling the calculation of local drag coefficient  $c_{D,\text{bottom}}$  from the local thickness of the bottom cell  $h_b$  and the bottom roughness parameter  $z_{0,b}$ . Though, in the operational model the bottom friction coefficient was kept constant (in 1 NM grid  $z_{0,b} = 0.01$  m and in 3 NM grid  $z_{0,b} = 0.03$  m). Also the



**Fig. 1.** Maps of the HIROMB-SMHI BS01 model grid for the whole Baltic Sea (a) and zoomed for the Estonian sea areas (b). Coastal on-line stations with used station name abbreviations, added to the web page (<http://on-line.msi.ttu.ee/kaart.php>) screenshot showing sea level observation results (c).

corrected air-ice stress and improved ice-ocean stress were introduced and the Successive Corrections data assimilation scheme for temperature and salinity was replaced by the Optimal Interpolation method. The Estonian forecasting system started to use version 4.0 model data on 10.12.2009.

Starting from 15.09.2010, the official version of HIROMB is 4.2. That contains no new parameterizations affecting the sea level. The changes in the

HIROMB have been continued. From 20.05.2011 the 11 km horizontal resolution HIRLAM forcing was taken into use, the forecast length was increased from 48 to 60 h and the amount of data used in data assimilation has been increased.

## 2.2. Online coastal observations

The coastal observational network, run by MSI, has undergone a significant evolution since 2005 and presently consists of 11 on-line stations (Fig. 1c, Table 2). A typical configuration of the automatic sea level station contains a staff gauge as an installation platform and a submerged piezoresistive pressure sensor at zero level of the staff. We use high-precision pressure sensors from Keller Ltd series 36WX and 46X with measuring amplitude of 5 m. Automatic temperature (resolution 0.1 °C) and air pressure compensation ensure an accuracy of 1 cm in terms of water column above the sensor. The pressure sensor is connected via a RS485 interface with the data logging, processing and transmitting device, which calculates sea level as a 30 s average water column height, as well as the basic wave parameters locally at the station, and sends the data with 5 min interval over GPRS communication protocol to a ftp server at MSI. On the server side, on-line data transfer is handled by GPRS Gateway software.

**Table 2.** Online sea level stations operated by MSI

| Station name and acronym | WGS84 position         | Sea level sensor                                  | Reference            | Last levelling | Zero level on staff gauge, cm | In operation since |
|--------------------------|------------------------|---|----------------------|----------------|-------------------------------|--------------------|
| Heltermaa<br>HEL         | 58°52.0'N<br>23°02.8'E | Keller 36WX<br>pressure sensor                    | 3 m staff<br>gauge   | 27.05.2010     | 144                           | Oct 2008           |
| Kuivastu<br>KUI          | 58°34.0'N<br>23°24.0'E | 2×Keller 46X<br>pressure sensor                   | 3 m staff<br>gauge   | 28.05.2010     | 122                           | Jan 2010           |
| Lehtma<br>LEH            | 59°04.1'N<br>22°41.8'E | HMS-1820 sea<br>level station                     | 3 m staff<br>gauge   | 17.04.2009     | 107                           | Jun 2007           |
| Paldiski<br>PAL          | 59°20.1'N<br>24°04.8'E | Keller 36WX<br>pressure sensor                    | 3 m staff<br>gauge   | 05.05.2010     | 128                           | Aug 2005           |
| Pärnu<br>PAR             | 58°23.3'N<br>24°29.2'E | Aanderaa Data<br>Instruments sea<br>level station | 4 m staff<br>gauge   | 23.04.2010     | 119                           | Jul 2000           |
| Rohuküla<br>ROH          | 58°54.3'N<br>23°25.5'E | Keller 36WX<br>pressure sensor                    | 3 m staff<br>gauge   | 27.05.2010     | 144                           | Jan 2009           |
| Sillamäe<br>SIL          | 59°25.4'N<br>27°44.4'E | Keller 36WX<br>pressure sensor                    | —                    | —              | —                             | Jun 2007           |
| Sõru<br>SOR              | 58°41.5'N<br>22°31.3'E | 2×Keller 46X<br>pressure sensor                   | 3 m staff<br>gauge   | 27.05.2010     | 130                           | Feb 2010           |
| Tallinn<br>TAL           | 59°26.7'N<br>24°45.8'E | Keller 36WX<br>pressure sensor                    | 3 m staff<br>gauge   | 05.05.2010     | 140                           | Mar 2004           |
| Triigi<br>TRI            | 58°35.5'N<br>22°42.2'E | 2×Keller 46X<br>pressure sensor                   | 2.6 m staff<br>gauge | 28.05.2010     | 161                           | Jan 2010           |
| Virtsu<br>VIR            | 58°34.6'N<br>23°30.5'E | Keller 36WX<br>pressure sensor                    | 3 m staff<br>gauge   | 28.05.2010     | 121                           | Jan 2009           |

Messages, coming from a number of GPRS modems in the field are converted into ASCII files and transmitted to ftp server(s), and also to the ftp box at MSI, from where the international data exchange occurs. After the initial quality control (obvious outliers are removed), the data is automatically fed into the BOOS system. The data is further delivered also to the EU project MyOcean.

Automatic sea level measurements are regularly checked by the readings of the staff gauge. It is a scale with 1 cm resolution for visual observations of the sea level (Fig. 2). The staff gauge is connected with the geodetic height system by means of high precision leveling, which is repeated regularly, typically once per year. The purpose of the comparison of the data from two independent measurements is the analysis of the long term trends in sensor performance and monitoring of local geodetic peculiarities, coming from vertical movements of hydrotechnical constructions, which the sea level station is fixed to.

The sea level sensors work quite reliably, showing only 2%–3% missing data without taking the major failures as powering of the station or sensor break-downs into account. The 97% up-time of the observation station is within the limits of the institutional goal but the procedures for major failures could be improved. Some stations are very old and should be replaced to decrease the amount of major failures. It would be ideal to replace at least the sensor of the station after a 5 years service. Though today, in case of a major failure of the



**Fig. 2.** Typical sea level observation station. The sensors are submerged behind and below the staff gauge and connected by cable to the box of the data handling unit (right above the staff gauge).



station, first the finances are to be found and then the station can be replaced, which takes much time and causes long gaps in measurements. Major sensor failures occurred in SOR, TRI (both from 20.07.2010 to 9.09.2010) and TAL stations (19.10.2009 to 14.11.2009, from 3.04.2010 to 19.04.2010 and from 22.10.2010 to 22.12.2010). In case of failures in the operational data transmission, the data is retrieved for climatological studies from the SD memory cards of on-site data loggers. The redundancy of the system is very much dependent on the sensors performance. In most of the stations there is only one sea level sensor and if it fails, no observation data is recorded. In some stations sensors are duplicated.

### 2.3. Initial system development, low-frequency error correction

Starting from 08.08.2005, the HIROMB data was downloaded daily, retrieving model results with 48 h forecast length. Shortly after this the first version of the forecasting system became operational with the forecast update interval  $L = 24$  h and forecast length  $M = 48$  h. The system was then continuously monitored and developed to improve its stability and reliability. After 17.12.2008, the forecast update interval was decreased to 6 h and the analysis was performed for the time period from 2006 to 2008.

A modelled sea level  $\eta_{mod}$  has always a bias relative to a geodetic reference system. After some months of running the forecasting system, a comparison of the time series of the model output  $\eta_{mod}$  and the observed sea level  $\eta_{obs}$  revealed that the model error has a low-frequency part that varies from location to location and is changing slowly in time. As an example, in the PAR station the “zero drift” varied during the three first implementation months (from September to November 2005) up to 25 cm (it dropped from 55 to 30 cm in about a month). It is obvious, that such a distortion of the forecast is unacceptable. The initial system running period was too short for a comprehensive statistical analysis, therefore a simple method for backward moving average (not centered, since future observations are not known during the forecast) of the model errors over 7 days was applied for the correction of the low-frequency errors. The maximum difference between the centered and the backward moving average in PAR was 8 cm, which is still small compared to the sea level variation range (from – 71 to 173 cm in 2006–2008). A statistical analysis for the period from February 2006 to March 2008 confirmed the usefulness of the applied procedure, details of which are presented below.

We have discrete sea level observations  $\eta_{obs}(n)$  at the specific point (station) and raw model forecasts  $\eta_{mod}(n, p)$  progressing in time, where  $n$  is the time index corresponding to a time  $t$  and  $p$  is the index denoting the forecasts with different lead time. While  $\eta_{obs}(n)$  is updated in real time after each time step (in our case 1 h),  $\eta_{mod}(n, p)$  is updated incrementally over  $L$  time steps (model restart interval, in our case 24 h in period 2006–2008) extending  $M$  steps forward (forecast length, in our case  $M = 48$ ). Usually  $M > L$  and several

forecasts with different lead time are available for the time  $n$ . This way for one observation time series we can assemble  $P = \max(p) = M/L$  forecast time series, including the forecasts with the lead time from  $(p-1)L+1$  to  $pL$ . In our case we have  $P=2$  forecast time series  $\eta_{mod}(n, 1)$  and  $\eta_{mod}(n, 2)$ , denoted with the forecast indices  $p=1$  and  $p=2$ , with the lead times 1–24 and 25–48, respectively. In the operational procedure we have a new forecast after each time step, that starts from the time  $N$  and extends to  $N+M$ . The forecast start time indices  $m(n, p)$  are incrementing with a step  $L$  and can be found as

$$m(n, p) = L \cdot \lfloor (n-S)/L \rfloor + S - (p-1)L, \quad (1)$$

where the floor brackets denote the down-rounded integer operator and  $S \leq n$  is the time span between the start time of the model calculations and the last observation taken into account. Note that due to the time, needed for the model calculations (first HIRLAM and then HIROMB) and data transfer, we have an approximately 7 h long time lag in the operational procedure between model calculations start times  $L \cdot \lfloor n/L \rfloor$  and the calculation of the forecast. Therefore the time, corresponding to the number of observations  $N$ , is usually not the same as model calculations start time ( $N \notin L \cdot \lfloor n/L \rfloor$ ) and we have more observations to use by the time the forecast is calculated. We have chosen to use  $S=4$  h for stability of the system (not 7, since sometimes the model data is received sooner and/or the observation data is delayed). The model results with the shortest lead time  $p=1$  (also called “best available forecast”, the lead time from 1 to  $L$ ) together with the observations  $\eta_{obs}(n)$ ,  $n = M+1, \dots, N$ ,  $N \gg (M, L)$  allow an analysis of the model errors  $d(n, 1) = \eta_{mod}(n, 1) - \eta_{obs}(n)$ . The analysis of an autocorrelation of  $d(n, 1)$  time series over the 2 years period at the PAR station (Fig. 3) shows that the error properties are far from the ideal white noise. The correlations above 0.2 appear in a quite long (up to 4000 h) time lags indicating some systematic error component.

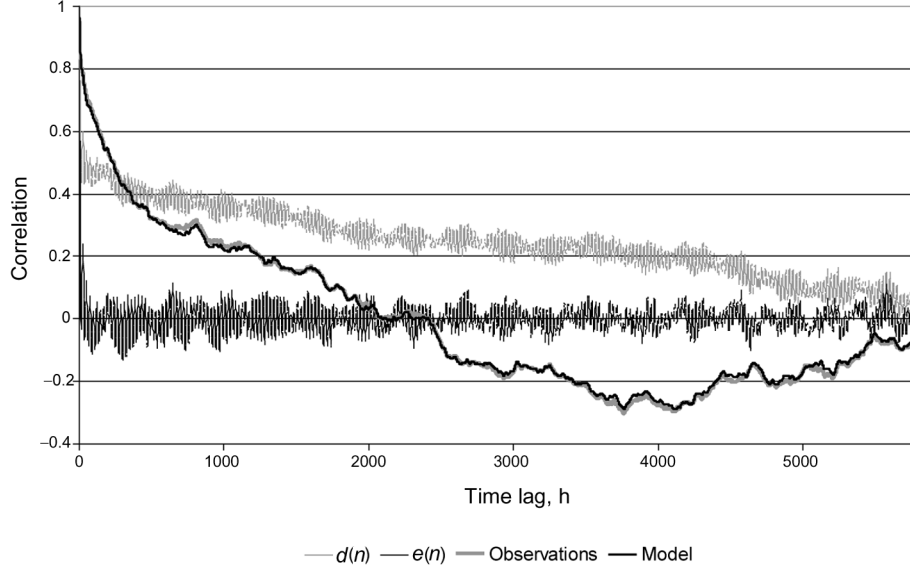
For finding the improved forecast  $\eta_{fc}(n, p)$ , we apply a correction

$$\eta_{fc}(n, p) = \eta_{mod}(n, p) - b(n, p), \quad (2)$$

where

$$b(n, p) = \frac{1}{K} \sum_{k=1}^K d(m-k) = \frac{1}{K} \sum_{k=1}^K [\eta_{mod}(m-k) - \eta_{obs}(m-k)], \quad (3)$$

is the low-frequency part of the model error (or dynamic bias), found using the backward moving average of the initial model error. The indices  $m$  should be found according to Eq. (1). In the operational procedure we use only the dynamic bias from the best available forecast  $b(N+1, 1)$  to correct all the forecasts within the interval  $n = (N+1, N+M)$ . For the performance optimization, it is necessary to choose the filter length  $K$  so that the forecast errors  $e(n, p) = \eta_{fc}(n, p) - \eta_{obs}(n)$  would be minimized. The selected cost function is the least



**Fig. 3.** The autocorrelation of observations, corrected model forecast, raw model error  $d(n)$  and forecast error  $e(n)$  at PAR station.

mean square error (MSE) estimate and we search for the best  $K$  that minimizes the expression

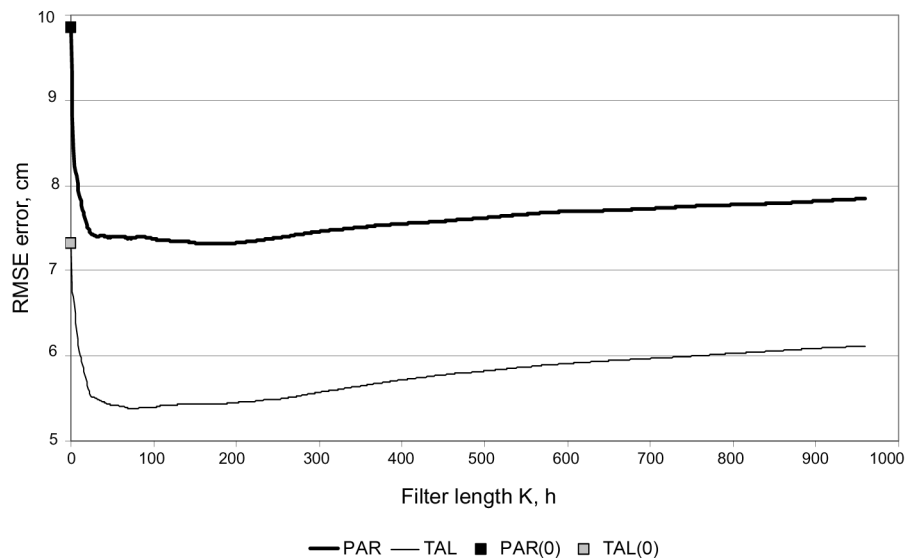
$$\begin{aligned}
 E(K) &= \frac{1}{N-A} \sum_{n=A}^N [\eta_{fc}(n) - \eta_{obs}(n)]^2 = \frac{1}{N-A} \sum_{n=A}^N [\eta_{mod}(n) - b(n) - \eta_{obs}(n)]^2 \\
 &= \frac{1}{N-A} \sum_{n=A}^N \left\langle \left\{ \eta_{mod}(n) - \frac{1}{K} \sum_{k=1}^K [\eta_{mod}(m-k) - \eta_{obs}(m-k)] \right\} - \eta_{obs}(n) \right\rangle^2 \\
 &\rightarrow \min,
 \end{aligned} \tag{4}$$

where  $A = \max(K) + 1$  is the maximum filter length, chosen for the search ( $\max(K) \ll N$ ) and  $m$  can be found from Eq. (1).

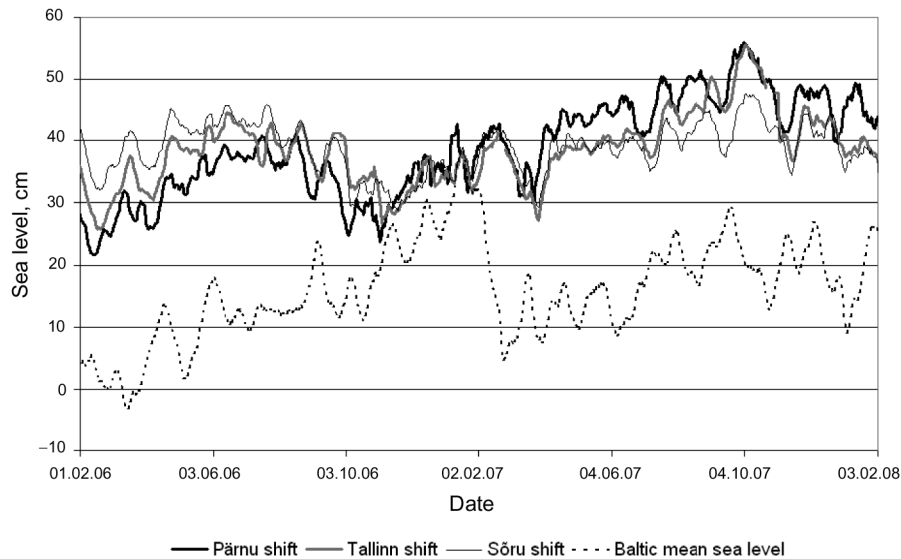
The method (1)–(4) is similar to the autoregressive moving average (ARMA) method, but differs from it by the non-unit increments  $L$  and  $M$ . Instead of the determination of  $K$  from the autocorrelation functions, we used numerical evaluation of Eq. (4) for  $K = 1, \dots, 960$  with  $L = 24$ . Both at the PAR and the TAL stations, the value of  $\sqrt{E}$  has a minimum in the band of moving average filter lengths from about 50 to 200 h. Therefore, considering possible data gaps,  $K = 168$ , adopted during the system start-up period, is a good choice, decreasing the RMSE by about twice compared to the raw forecast ( $K = 0$ ). As a result, the autocorrelation of the forecast error  $e(n)$  drops rapidly close to zero (Fig. 3) and is much closer to the white noise properties than the autocorrelation of the raw forecast error  $d(n)$ .

The low-frequency error correction method, described above (and used in the forecasting system), is designed to minimize the errors of the sea level forecast for the operational purposes only. The method is purely statistical, giving a temporary solution until the actual model physics, numerical implementation and forcing data are improved. Regarding more comprehensive correction techniques, there are some test studies for the sea level data assimilation in the North and Baltic Sea region [28], but such assimilation is not yet implemented in the operational practice.

There have been several discussions during the last decade concerning the low-frequency forecast errors of the sea level. The origin of such a dynamic bias is still not clear; there could be several reasons like distortions in the boundary conditions provided by the “outer” storm surge model NOAMOD, density inaccuracies in the model response, inaccurate volume changes in the Baltic Sea and its sub-basins due to the errors in volume transports in the straits, errors in the freshwater budget, etc. Although the analysis of these reasons is out of the scope of this paper, an attempt was made to find the relationship between the dynamic bias and the Baltic Sea volume change during the optimization of the low-frequency error calculation method. The Baltic-wide mean sea level time series were found and compared with the dynamic sea level bias time series in different stations (Fig. 5). Although several variations of  $b(n, 1)$  and horizontally mean sea level occur at the same time, their phases vary significantly and the statistical correlation is practically absent. Therefore it was found that the forecast of the horizontally mean sea level cannot be used for the prediction of  $b(n, 1)$ .



**Fig. 4.** The RMSE dependence on the backward moving average filter length  $K$ .



**Fig. 5.** Time series of low-pass filtered (7 days moving average) sea level difference between model raw output and sea level observations (shift) for Pärnu, Tallinn and Sõru and model raw output for Baltic mean sea level.

The low-frequency error variations at PAL station during a later period (03.2009–03.2011) are presented later. There were two model upgrades during this period, which can be distinguished from the dynamic bias curve. Although the version 4.2 did not introduce any changes, affecting directly sea level, the variability range of the dynamic bias seemingly increased, a feature that requires further investigation.

## 2.4. Forecast production

Results from the HIROMB-SMHI model with 1 NM resolution (grid BS01) are used in the production of the sea level forecasts for the Estonian coast. The forecasting system is operationally run 4 times per day with cron scheduler in Linux cluster. The system has 4 main steps, all covered with error handling procedures:

- 1) download the model data;
- 2) download the observational data;
- 3) process the model and observation data;
- 4) publish the forecast and archive the data.

The model data is downloaded from the SMHI ftp server, saved and unzipped. The data is provided in GRIB format, having separate file for each time step and resulting in  $M$  model files. Each file is downloaded and checked separately. The system automatically assesses the quality of the file and either re-downloads the file or starts downloading the next file. The re-download is repeated maximum

5 times per file after which the error message is sent to the administrator. The downloading state is the most time consuming and fragile step in the forecasting system due to the size of the 3D files and network stability. However, there has been significant progress in the network stability and speed. Although the file sizes were approximately twice smaller in 2006, the download took about 2 h, while today it completes usually within 10 to 20 min.

The crucial element for the operational forecasting is the sea level observations. New observation is available every 5 min, as described in Section 2.2. The forecasting system, however, only downloads the data with a 1 h interval, assembling separate 1 h interval observations data file, which is not always with the same quality as the data in observational system itself. Similarly with the model data, in case of errors in the observational data the re-download is performed maximum 5 times, after which the data is defined missing for the forecasting system. For example, if the data is too late in the observational ftp, then this data is defined as missing, although files of this data exists in observational system archive. For the offline analysis, the differences between the observational and the forecasting system's data archive could be decreased by creating a direct link between the subsystems, but this has not been implemented yet.

In the data processing step the model data is first extracted from the 3D model data files and appended to the model extractions time series files  $\eta_{mod}(n, p)$  for each station. After constructing the model extractions and the observation data, a simple automated quality control is performed: the unreasonable high or low values and data gaps are flagged. However, the manual revision of the data during the current study showed the need for some improvement of this quality control, although only 0.7% of data was considered "bad" during manual revision. Before adding more sophisticated quality controls to the operational forecasting system, these controls have to be well calibrated and validated, to avoid filtering out actual rapid sea level changes, which naturally happen mostly in the areas and periods of high flooding risk.

After the bad quality data has been flagged and removed from further calculations, the low-frequency error  $b(N+1,1)$  (from Eq. (3)) and the corrected forecasts are calculated as  $\eta_{fc}(n, p) = \eta_{mod}(n, p) - b(N+1,1)$  for each station, where  $n = N + (p-1)L + 1, \dots, N + pL - S$  (the special case of Eq. (2)). Although it is possible to use Eq. (4) operationally for finding the best filter length  $K$  for each forecast separately, in order to ensure stability of the long-term statistics the recalculation procedure is not implemented operationally. The processing step continues with gathering the forecasts of all lead times and stations into one forecast file. Due to up to 7 h time span between the start times of the model and the forecast calculations (see also the explanation of Eq. (1)), only the forecast lead times greater than 5 h are written into forecast file, which unfortunately means that only the 6th hour forecast from the "best available" forecast reaches the users.

The data gaps are usually the most problematic in the operational systems. The forecasting system has some simpler redundancy functions in case the model

or observation data is missing. In case the model data is missing for some reason, the data from the previous forecast is used to produce the forecast; so in principle the first  $L$  records are deleted from the previous forecast file and used as the new forecast file. This means that the replacement  $\eta_{fc}(n, p) = \eta_{fc}(n, p+1)$  is performed and, of course, the forecast length  $M$  is decreased by  $L$ . If the model data is missing also during the next forecasting system run time, the replacement  $\eta_{fc}(n, p) = \eta_{fc}(n, p+2)$  is performed etc., until the model data has been missing for  $M = 48$  h in a row. After that period the forecast will be provided with missing value codes and an alert message will be sent to the administrator. Another source for missing forecast is gaps in observations. In case there are less than 48 acceptable observation-forecast pairs in the last 7 days or the observation data has been missing for the last 48 h, the forecast will be again provided with missing value codes.

In connection with the forecast production, an automated high/low sea level checking is done for each forecast. Critical values for every coastal station are defined (Table 3) and when the forecast is out of the defined limits, an automated high/low sea level warning message is sent to the users. The skill of the high and low sea level forecast and warning system was evaluated by comparing the number of high/low sea level events forecasted and observed (Table 3). In general the amount of high sea levels within the period from 2009 to 2011 was low and no coastal flood was observed, but there were several low sea level events which actually influenced the ship traffic between small islands and mainland. Most of the occurred high/low sea level events were alerted properly, but still 18% of high/low events were underestimated by the forecast and the warning message was not sent. The reason behind the missed warning messages was that the sea levels were slightly below the warning limits. Thus it can be concluded that no really important high/low sea level event was missed by the forecasting system. The users can choose how often they want to get the warning message. By default, the maximum warning message frequency is once in 24 h. Without this option the system could send out up to 8 messages for the same

**Table 3.** Low and high sea level warning limits and the performance of the warning system. If forecast value is outside the limits the warning message will be sent to the users by the system

| Station acronym                                       | HEL      | KUI     | LEH    | PAL    | PAR    | ROH      | SIL    | SOR    | TAL    | TRI    | VIR     |
|---|----------|---------|--------|--------|--------|----------|--------|--------|--------|--------|---------|
| Limit of low sea level warning, cm                    | -50      | -60     | -50    | -60    | -90    | -50      | -60    | -50    | -60    | -50    | -60     |
| Limit of high sea level warning, cm                   | 70       | 70      | 50     | 50     | 90     | 70       | 80     | 50     | 70     | 50     | 70      |
| Number of low sea level warnings forecasted/observed  | 1<br>2   | 0<br>0  | 2<br>3 | 3<br>3 | 1<br>1 | 2<br>2   | 3<br>3 | 2<br>2 | 1<br>1 | 0<br>1 | 1<br>3  |
| Number of high sea level warnings forecasted/observed | 12<br>10 | 8<br>13 | 5<br>7 | 4<br>4 | 2<br>2 | 13<br>13 | 6<br>5 | 3<br>7 | 4<br>4 | 3<br>5 | 9<br>13 |

high/low sea level event. Only one warning per one high/low sea level event is recommended for a typical end-user.

As a final step of the forecasting system, all the data is archived and the forecast is published to the users (via ftp and web). Both the measurements of the sea level along the Estonian coast and the forecasts are presented in an open access webpage Sea Level Information System <http://on-line.msi.ttu.ee/kaart.php?en>. In normal situation, when the sea level is around 0 cm, the number of clicks in the system per day is about 1000, but during storms it increases 1000 and even 10 000 times, resulting in 10 million clicks per day. The users of the system are mainly the people living in the coastal areas, but also larger institutions like Estonian Rescue Service. This system has been used for issuing warnings about high sea level in the Estonian coastal sea, and by national and local authorities for decision-making. The company that operates ferries between the Estonian mainland and islands relies on the forecasts of critically low sea level.

## **2.5. Analysis methods**

Standard statistical methods were used to study the archived operational results of the observations and forecasts. An analysis was made for two different time periods: 01.10.2006–01.03.2008 and 1.03.2009–1.03.2011.

The RMSE minimizing algorithms described above and the autocorrelation function technique were used mainly for the time period 2006–2008 data to find the optimal algorithm for low-frequency error calculations. Storm surge events were analysed to investigate the skill of the forecasting system for predicting the high sea level events. The events were handled separately to find the forecast error with respect to observations. Also the temporal errors of the storm peaks (time span between the maximal sea level in the forecast and in the observations) were found and recorded manually.

Starting from 17.12.2008, the forecasting system was improved to use model data and produce forecast 4 times per day, i.e., after each 6 h. Also the observation network was extended to 11 stations. For the analysis period from 2009 to 2011, different statistics were calculated, mainly for the forecasts with lead time +01 h to +06 h. The time percentage analysis was done to evaluate the forecasting system up-time and different sources of the forecasting system failures. For statistical calculations all data was revised, gaps and corrupted data were removed and the data time series (model, forecast and observation) were trimmed to the same length for each station separately. For example, if the observation data for certain time was missing, then both the forecast and the model data were excluded from the analysis, although one of those could have been quality data. This trimming procedure was essential to find comparable statistics.

Mean absolute error (MAE) was calculated by taking arithmetic mean of the absolute deviations between each observation and forecast and monthly mean absolute error (MMAE), describing seasonal variation of the forecast error, was found as the arithmetic mean of MAE over the one month period. The root mean



square difference (RMSD), standard deviation (STD) and correlation (CORR) were found between forecast and observed data to describe the forecast error, amplitude and phase respectively. Mean error (ME), describing the conformity of the error distribution to the normal distribution, was found as arithmetic mean over the deviations between the observation and the forecast. To evaluate the accuracy of the forecasting system in greater detail, the whole range of sea level variations was divided into three sub-ranges at each station: high, medium and low. The forecasted sea levels were rounded to the nearest integer and sorted with corresponding forecast errors in ascending order. Then MAE was calculated for each forecasted sea level and the sea level sub-ranges were found. Both the whole range and the sub-range statistical characteristics were analysed. Taylor diagram [29] was chosen to summarize the statistical results.

### **3. RESULTS AND DISCUSSION**

#### **3.1. Forecasting high sea levels in 2006–2008**

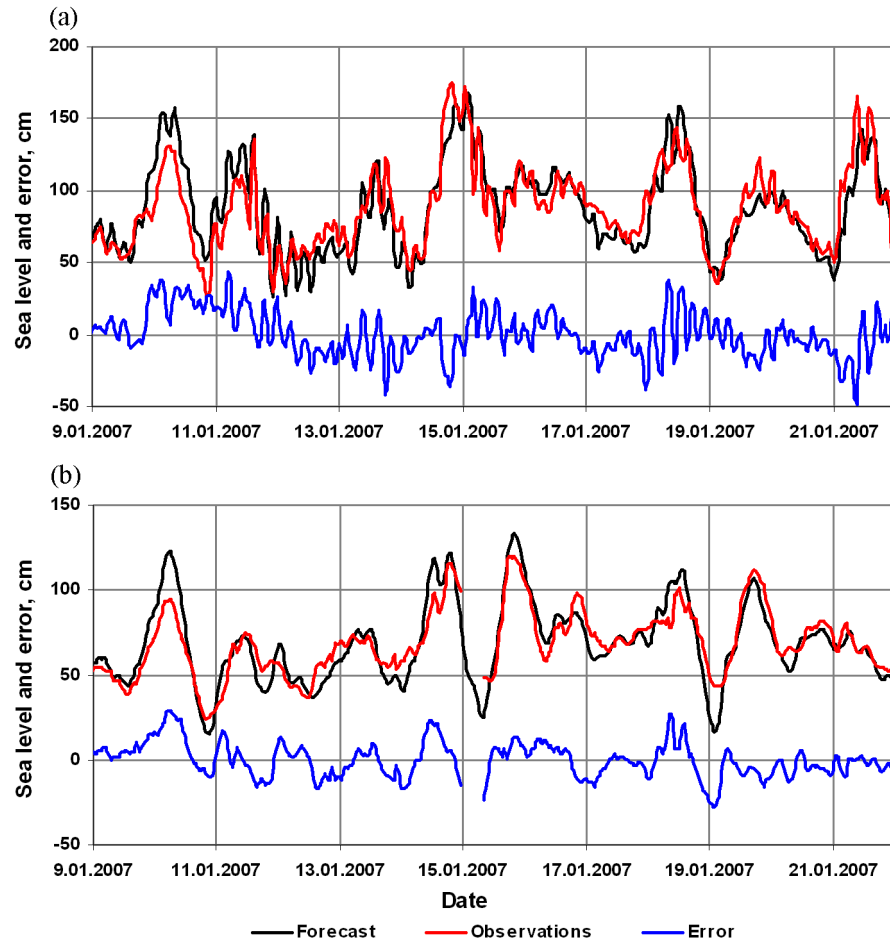
The period 2006–2008 included several storm surges with high sea levels, observed and forecast in many coastal areas. Since these cases were extensively reported in mass media and published partly in [30,31], only a short note is given here. The purpose of including this period here is that the second (2009–2011) dataset does not contain significant storm surges.

Figure 6 presents examples of high sea level events observed during January 2007 at PAR and TAL station. The sea level can increase over 40 cm within an hour (see PAR), which could lead to large instantaneous errors, if there is a time lag between the model and the observations. In case of flood risk, the end-users are interested in maximum sea levels during the risk period, normally about half a day.

High sea levels were found for two stations, PAR and TAL, for the time period 10.2006–03.2008. The criteria for a high sea level event were chosen 130 cm in PAR station and 90 cm in TAL station, with at least a 24 h time span between different high sea level events. There were 5 and 6 such events at PAR and TAL stations, respectively (Table 4). The forecast error was lower than  $\pm 25$  cm for all these events and the time shift between observed and forecasted sea level maximum was within  $\pm 3$  h.

#### **3.2. Statistical analysis of sea level forecasts in 2009–2011**

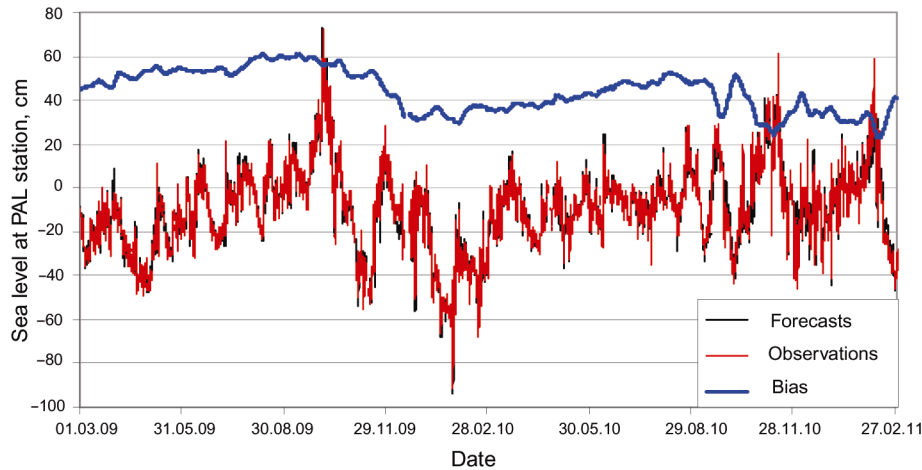
Statistical analysis for 11 coastal stations (Table 2) was performed for sea level forecasts within time period 1.03.2009–1.03.2011, although different time periods were available for different stations. A sample plot of forecasts and observations of the whole time period at PAL station is given in Fig. 7. Notice that the forecast is the output of the forecasting system, which is related to the model through the low-frequency error correction (see Eq. (2)).



**Fig. 6.** High sea level events in January 2007 at the stations PAR (a) and TAL (b).

**Table 4.** Observed high sea level events at stations PAR and TAL within the time period 10.2006–03.2008

| Date       | Time  | Observed sea level, cm | Forecast error, cm | Time shift, h | Station |
|------------|-------|------------------------|--------------------|---------------|---------|
| 10.01.2007 | 10:00 | 132                    | 24                 | 1             | PAR     |
| 11.01.2007 | 18:00 | 138                    | −3                 | 0             | PAR     |
| 15.01.2007 | 04:00 | 175                    | −5                 | 1             | PAR     |
| 18.01.2007 | 14:00 | 144                    | 16                 | 1             | PAR     |
| 21.01.2007 | 12:00 | 166                    | −19                | 2             | PAR     |
| 10.01.2007 | 09:00 | 94                     | 25                 | 0             | TAL     |
| 14.01.2007 | 22:00 | 117                    | 6                  | 0             | TAL     |
| 15.01.2007 | 23:00 | 121                    | 14                 | 0             | TAL     |
| 16.01.2007 | 23:00 | 98                     | −9                 | 0             | TAL     |
| 18.01.2007 | 15:00 | 101                    | 13                 | 1             | TAL     |
| 19.01.2007 | 20:00 | 112                    | −1                 | 0             | TAL     |



**Fig. 7.** Sea level observation and output of the forecasting system at the PAL station. The original model data can be found by adding forecast (black) and dynamic bias (blue). Investigation period included one storm surge event (over 65 cm) on 5.10.2009 and one low sea level event (above -80 cm) on 28.01.2010.

### 3.2.1. The data gaps and the up-time of the forecasting system

Two stations, KUI and TRI went online in Nov 2009 and the PAR observation station broke down on 24.05.2010, so these stations had less data compared to other stations. The observation data at SOR, TAL and TRI contained major data gaps. The stations SOR and TRI were out of order from 20.07.2010 to 9.09.2010 and the station TAL was down from 19.10.2009 to 14.11.2009, from 3.04.2010 to 19.04.2010 and from 22.10.2010 to 22.12.2010. These gaps were taken into account in up-time calculations of the forecasting system.

The percentage of missing data (Table 5) shows the data amount that was actually used in statistical calculations. Notice that the amount of data, actually used in calculations, was lower than the amount of actual forecasts, delivered to the users. The difference comes from the system's ability to forecast sea level even if observation data is not available for a short time. Statistical analysis used needs the data without any gaps, which means that the percentage of missing data amount, not used in the calculations, is higher by short-time (below 48 h) missing observation data.

The data gaps for the best available (forecast lead times from 1 to 6 h) forecast were analysed in greater detail. Three reasons for the forecasting system failures were identified and the amount of these failures was calculated. First, and the most obvious, is that the forecasting system can not run in case of hardware failures (cluster down, no electrical power, etc). This kind of error was found 1% of the time. Secondly, if the model results are missing then the forecasting system fails, which was the case in 2.6% of time. This failure appears mainly due to the network errors. It is also possible that the network is working, but the model data is uploaded to ftp-box too late (for example, the model calculations take too

**Table 5.** Main statistics for the whole range of data within time period 1.03.2009–1.03.2011. Stations KUI, TRI and PAR have shorter time period, i.e. 6.11.2009–1.03.2011 for KUI and TRI and 01.03.2009–24.05.2010 for PAR

| Station acronym                    | HEL   | KUI   | LEH   | PAL   | PAR   | ROH   | SIL   | SOR   | TAL   | TRI   | VIR   |
|------------------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| Missing data, %                    | 7     | 7     | 8     | 5     | 11    | 6     | 5     | 11    | 19    | 16    | 6     |
| Long-term missing observation, %   | 3     | 1     | 3     | 0     | 9     | 1     | 0     | 7     | 9     | 12    | 2     |
| Forecast down-time for end-user, % | 1.4   | 1     | 3     | 1     | 4     | 1     | 1     | 7     | 8     | 11    | 2     |
| ME, cm                             | 0.08  | 0.10  | 0.11  | 0.05  | 0.24  | 0.07  | 0.08  | 0.00  | 0.07  | 0.25  | 0.07  |
| MAE, cm                            | 2.92  | 3.98  | 3.65  | 2.69  | 4.79  | 3.47  | 3.85  | 2.77  | 2.94  | 3.32  | 3.77  |
| Mean dynamic bias, cm              | 43.25 | 45.34 | 37.66 | 44.30 | 44.15 | 42.50 | 32.09 | 40.47 | 41.55 | 35.83 | 41.51 |
| STD of observation, cm             | 18.91 | 21.44 | 19.94 | 18.69 | 26.35 | 21.26 | 22.20 | 19.69 | 19.60 | 18.79 | 21.59 |
| STD, cm                            | 18.62 | 20.79 | 19.14 | 18.44 | 26.16 | 21.21 | 21.89 | 19.23 | 19.17 | 18.08 | 21.29 |
| RMSD, cm                           | 3.82  | 5.26  | 4.88  | 3.53  | 6.37  | 4.55  | 5.01  | 3.72  | 3.85  | 4.32  | 4.98  |
| Normalized STD                     | 0.98  | 0.97  | 0.96  | 0.99  | 0.99  | 1.00  | 0.99  | 0.98  | 0.98  | 0.96  | 0.99  |
| Normalized RMSD                    | 0.20  | 0.25  | 0.24  | 0.19  | 0.24  | 0.21  | 0.23  | 0.19  | 0.20  | 0.23  | 0.23  |
| CORR, %                            | 98    | 97    | 97    | 98    | 97    | 98    | 97    | 98    | 98    | 97    | 97    |

much time in case of heavy ice conditions), but the forecasting system has to send out the forecast. The percentages in time of the first two failures were obviously equal for every station. The third reason for the failures is long-term missing observations (Table 5). As explained above, the forecast is defined missing in case there are less than 24 successful observations present within the last three days (72 h). At most of the stations this failure was below 3% of time, while at three stations (SOR, TAL, TRI) it was higher. The reason was breakdown of the observation station or very long network error between the forecasting system and the observation station. The analysis was made intentionally with the data stored by the forecasting system, to evaluate the performance of the actual forecast that has been delivered to the end-user. To summarize the best available forecast data gaps analysis, the general down-time was about 6%. The down-time of the forecasting system itself would have been about 4%.

The missing observation data was divided into two categories. A missing observation was defined as long-term in case there was less than 24 h not-missing data within the 72 h period, preceding the missing observation under evaluation. All other missing data was defined as short-term. The sub-category of long-term missing data was major data gaps, i.e., the data was missing over 2 weeks in a row. The observation was missing on average 7% of time, of that 4% long-term and 3% short-term.

Due to the forecasting system's ability to restore a forecast from previous forecasts, the up-time of the forecasting system for the end-user is higher than the best available forecast up-time. In case the model data is missing, the forecasting system uses data from the previous model run and delivers 6 h shorter forecast to the end-user. If the model data has been missing for 48 h or more, the forecasting system is "down" for the end-user. The percentage of down-time for end-user is given in Table 5 for every station. Mean down-time of all stations is 4%. If the major data gaps from SOR, TAL, TRI stations are excluded, the mean up-time of the forecasting system for the end-user would be 98%.

The institutional goal for the end-user up-time is to keep it higher than 96% while the forecasting system can not be down longer than 48 h in a row. As described above, the first part of this goal is fulfilled, but there are major data gaps in the observations that should be avoided. The network link from the observation stations to the forecasting system can be improved by software changes (multiple download attempts, other transfer protocols like 3G or satellite communication, etc). The second part needs improved observation stations maintenance routines, which could prevent the breakdown and more effective removing of the failures.

### 3.2.2. Statistics for the whole dataset

In order to evaluate the dependence of the performance of the forecasting system on the forecast update interval  $L$ , the statistics for forecasts, made with  $L=6$  (used operationally since about 2009) and  $L=24$  (used operationally until about 2009) were calculated and compared. The best available forecasts for both update intervals for the period 2009–2011 were used to obtain comparable results, i.e., the  $L=6$  data was obtained from the operational runs and the  $L=24$  data was simulated with 2009–2011 data. The improvement was calculated as the decrease in the mean RMSD and increase in the mean correlation with regards to the increase in the forecast frequency. The mean improvements were found 8% and 0.5% for RMSD and correlations, respectively. Correlation improvement percentage for high and low sea level sub-ranges (see below) was 3.1% and 4.3%, respectively.

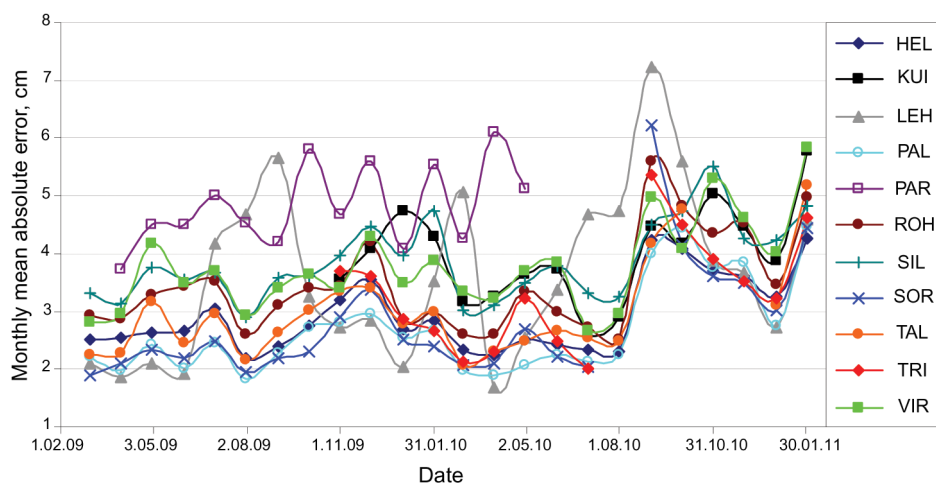
As described before, a 7 day filter was used in the forecasting system to correct the low-frequency error. The mean dynamic bias (Table 5), shows a range from 32 to 46 cm. Most of the stations have the mean dynamic bias over 40 cm, the lowest mean dynamic bias is at SIL station, which can be explained by the geographical distance from other stations and closeness to the Neva and Narva rivers, discharging fresh water into that region. Unexpectedly, low values of mean dynamic biases were found also at LEH and TRI stations, which can point to bathymetry inaccuracies in the model domain. The mean forecast errors are near zero at all stations, which shows a rather good performance of the dynamic bias correction algorithm.

Monthly mean absolute errors at different stations are presented in Fig. 8. In general there is no seasonal difference in MMAEs. The MMAE at station PAR is

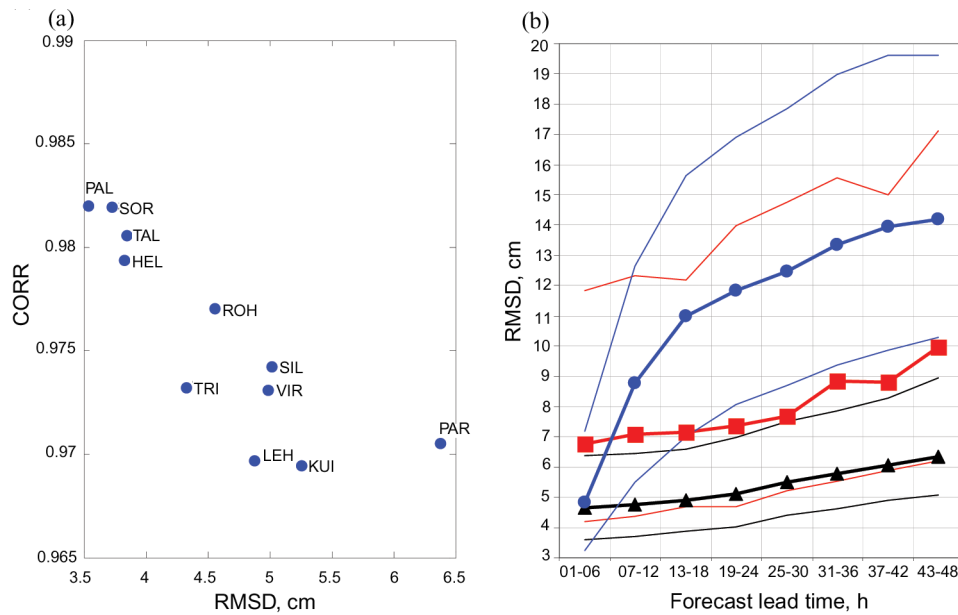
higher than for the other stations, reaching 6.1 cm in April 2010, while MMAE of the LEH station distinguishes with high variability of MMAE. There is a significant increase in MMAEs, starting from 09.2010, which is the same time as upgrade to HIROMB v. 4.2 (released on 15.09.2010), although there were no changes introduced in v. 4.2, which should influence the sea level. The sea level had also high variance in spring 2010 (Fig. 7), which amplified the increase in MMAEs. Previous upgrade to the model version 4.0 on 10.12.2009 is not so clearly visible, although the MMAEs are slightly increased in that period.

The RMSD is one of the most important statistics for describing the performance and accuracy of the forecasting system. While it combines standard deviation of forecasts and correlation between observations and forecasts, it also has clear statistical meaning. Considering that RMSD is equivalent to a standard deviation of the forecast error (in case of normal distribution), it can be interpreted as the absolute error level within the 68% probability level. The RMSDs of the best available sea level forecasts, varied from 3 cm (PAL) to 7 cm (PAR), giving evidence of rather accurate forecasts.

To summarize the performance of investigated sea level forecasts, RMSD vs. correlation plot (Fig. 9a) was used instead of the commonly used normalized Taylor diagram. The reason is the high similarity of statistics between different stations (notice the scales of Fig. 9a) which would lead to a poorly readable Taylor diagram. Nevertheless, from the values of STDs, normalized with STDs of observations (Table 6) it is concluded that forecasts tend to underestimate the variance of the sea level. The best general performance in the means of RMSD/CORR rate is at PAL station with 98.2% correlation. Correlation is high (over 97%) at all stations, which points to small overall phase errors of the forecasting system.



**Fig. 8.** Monthly mean absolute errors of the forecast at different stations. Mean errors are increased after release of the new HIROMB v. 4.2 on 15.09.2010.



**Fig. 9.** Forecast correlation and RMSD with respect to observation for the whole dataset (a) and RMSD dependence on forecast length for different data ranges (b). On (b) the black, red and blue bold lines show the mean RMSD over all stations of the whole dataset, high sub-range and persistency forecast, respectively. Thin lines with corresponding color show minimal and maximal individual curve from the most accurate and least accurate station, respectively. Forecast is most accurate at the PAL station and least accurate at the PAR station.

To investigate the forecast accuracy dependence on the forecast length, the RMSD was calculated for each forecast lead time range (from  $(p-1)L+1$  to  $pL$ ) and for each station (Fig. 9b). The minimal boundary is from the station PAL and the maximal boundary from the station PAR as could be expected from Fig. 9a. The mean accuracy of the forecast decreases only up to 2 cm for the 48 h forecast length and the error for the 48 h forecast is up to 2.5 cm higher than the best available forecast error, which is mostly presented in this paper.

As an illustration of the forecast precision, the persistency forecasts were carried out, using available observational data, and compared to the results of the operational forecast. The persistency forecast is one of the most primitive forecast methods and it is based on the assumption that the last observed situation will continue over the forecast period. For example, in 1–6 h persistency forecast it is assumed that the most recent observed sea level is constant over the next 6 h. The accuracy of the persistency forecast (Fig. 9b) is close to the accuracy of the forecasting system forecast up to 6 h, but in longer lead times, up to 48 h, the advantage of the forecast system is clearly evident.

**Table 6.** Main statistics for low, medium and high sub-ranges within the time period 1.03.2009–1.03.2011. Stations KUI, TRI and PAR have shorter time period, i.e. 6.11.2009–1.03.2011 for KUI and TRI and 01.03.2009–24.05.2010 for PAR

| Station acronym                            | HEL    | KUI   | LEH    | PAL    | PAR   | ROH    | SIL    | SOR    | TAL    | TRI   | VIR    |
|--|--------|-------|--------|--------|-------|--------|--------|--------|--------|-------|--------|
| Definition of low sea level sub-range, cm  | –50    | –50   | –50    | –60    | –60   | –60    | –50    | –60    | –50    | –60   | –50    |
| Definition of high sea level sub-range, cm | 35     | 15    | 35     | 40     | 50    | 40     | 50     | 40     | 50     | 20    | 40     |
| Number of low sea levels                   | 307    | 693   | 375    | 50     | 222   | 97     | 150    | 46     | 284    | 46    | 322    |
| Number of medium sea levels                | 15 272 | 9 368 | 15 273 | 16 120 | 8 317 | 15 754 | 15 795 | 14 963 | 12 962 | 8 977 | 15 337 |
| Number of high sea levels                  | 232    | 521   | 110    | 111    | 123   | 216    | 334    | 144    | 87     | 245   | 310    |
| ME of low sub-range forecast, cm           | 1.49   | 1.06  | –0.31  | 2.43   | 2.61  | 5.56   | 0.47   | 1.31   | 0.17   | 0.79  | 1.21   |
| STD of low sea level observation, cm       | 7.49   | 8.74  | 6.25   | 9.27   | 8.78  | 9.48   | 14.81  | 7.00   | 8.79   | 8.32  | 7.79   |
| STD of low sea level, cm                   | 7.26   | 7.95  | 6.12   | 9.69   | 8.30  | 7.21   | 12.82  | 4.64   | 7.59   | 5.01  | 6.40   |
| RMSD of low sea level, cm                  | 3.65   | 6.30  | 3.17   | 4.51   | 6.44  | 5.48   | 9.31   | 4.47   | 4.22   | 3.78  | 4.26   |
| Normalized STD of low sea level            | 0.97   | 0.91  | 0.98   | 1.05   | 0.95  | 0.76   | 0.87   | 0.66   | 0.86   | 0.60  | 0.82   |
| Normalized RMSD of low sea level           | 0.49   | 0.72  | 0.51   | 0.49   | 0.73  | 0.58   | 0.63   | 0.64   | 0.48   | 0.45  | 0.55   |
| CORR of low sea levels, %                  | 88     | 72    | 87     | 89     | 72    | 82     | 78     | 78     | 88     | 96    | 84     |
| ME of medium sub-range forecast, cm        | 0.05   | 0.06  | 0.11   | 0.05   | 0.22  | 0.07   | 0.10   | 0.00   | 0.07   | 0.23  | 0.11   |
| STD of medium sea level observation, cm    | 16.48  | 16.17 | 17.93  | 17.83  | 22.92 | 19.66  | 19.70  | 18.62  | 17.57  | 17.51 | 18.80  |
| STD of medium sea level, cm                | 16.06  | 15.00 | 17.06  | 17.52  | 22.37 | 19.40  | 19.31  | 18.10  | 17.07  | 16.77 | 18.14  |
| RMSD of medium sea level, cm               | 3.80   | 4.96  | 4.91   | 3.51   | 6.23  | 4.48   | 4.88   | 3.71   | 3.82   | 4.29  | 4.85   |



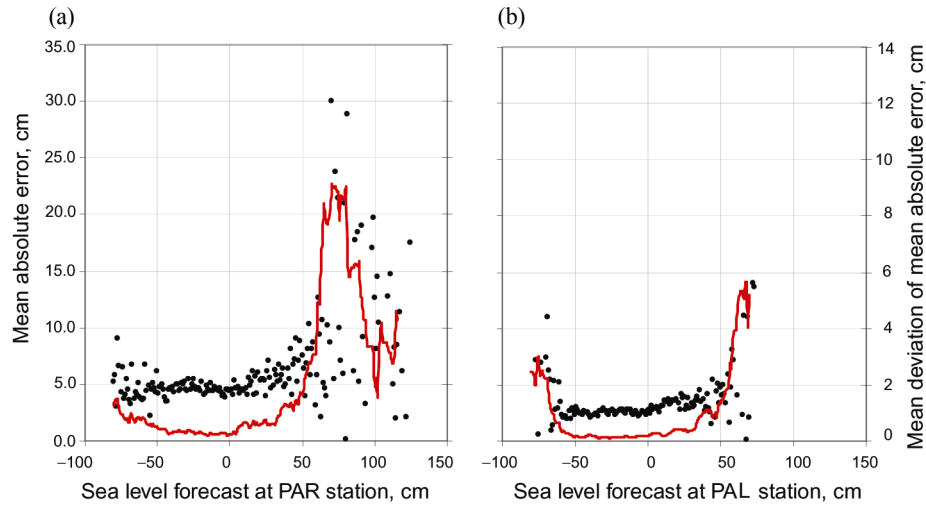
**Table 6. Continued**

| Station acronym                               | HEL   | KUI   | LEH  | PAL   | PAR   | ROH   | SIL   | SOR   | TAL   | TRI  | VIR   |
|---|-------|-------|------|-------|-------|-------|-------|-------|-------|------|-------|
| Normalized<br>STD of<br>medium sea<br>level   | 0.97  | 0.93  | 0.95 | 0.98  | 0.98  | 0.99  | 0.98  | 0.97  | 0.97  | 0.96 | 0.97  |
| Normalized<br>RMSD of<br>medium sea<br>level  | 0.23  | 0.31  | 0.27 | 0.20  | 0.27  | 0.23  | 0.25  | 0.20  | 0.22  | 0.24 | 0.26  |
| CORR of<br>medium sea<br>levels, %            | 97    | 95    | 96   | 98    | 96    | 97    | 97    | 98    | 98    | 97   | 97    |
| ME of high sub-<br>range fore-<br>cast, cm    | 0.10  | -0.44 | 0.37 | -0.64 | -2.81 | -2.32 | -0.71 | -0.52 | -0.88 | 0.71 | -3.39 |
| STD of high sea<br>level observa-<br>tion, cm | 11.40 | 12.05 | 9.08 | 7.75  | 22.80 | 15.29 | 13.19 | 9.85  | 8.48  | 7.10 | 15.87 |
| STD of high sea<br>level, cm                  | 10.95 | 9.62  | 6.08 | 7.14  | 22.28 | 16.28 | 11.82 | 9.03  | 7.61  | 4.89 | 16.35 |
| RMSD of high<br>sea level, cm                 | 5.18  | 8.01  | 5.47 | 5.42  | 12.05 | 6.89  | 7.57  | 4.13  | 5.97  | 5.52 | 9.16  |
| Normalized<br>STD of high<br>sea level        | 0.96  | 0.80  | 0.67 | 0.92  | 0.98  | 1.06  | 0.90  | 0.92  | 0.90  | 0.69 | 1.03  |
| Normalized<br>RMSD of<br>high sea level       | 0.45  | 0.66  | 0.60 | 0.70  | 0.53  | 0.45  | 0.57  | 0.42  | 0.70  | 0.78 | 0.58  |
| CORR of high<br>sea level, %                  | 89    | 75    | 81   | 74    | 86    | 91    | 82    | 91    | 73    | 63   | 84    |

### 3.2.3. Statistics of sub-ranges

Due to the broad range of temporal scales of the sea level variations it is obvious that the statistics of sea level time series greatly depends on the length of the time series and the presence of high/low sea level events. The statistics for mostly medium range sea level fluctuations does not describe extreme events like high and low sea levels sufficiently and can be even misleading, if used in the wrong context. For the end-users, the accuracy of sea level forecast is always much more important during the extreme events than in “normal situations”, therefore evaluation of statistics especially for such events has a great practical relevance.

The relationships between mean absolute errors and the forecast sea levels were found (Fig. 10), to find the statistics, describing in greater detail the extreme events of sea level. Three different sub-ranges were found for every station: low, medium and high sea level sub-ranges (the limits of these sub-ranges should not be confused with high/low sea level warning limits). Limits defining these sub-



**Fig. 10.** Mean absolute errors at PAR (a) and PAL (b) stations. Line indicates mean deviation of 10 following mean absolute errors. Three sub-ranges can be distinguished: low, medium and high sea level sub-range. For the PAL station the low sub-range has more contrast than the PAR station.

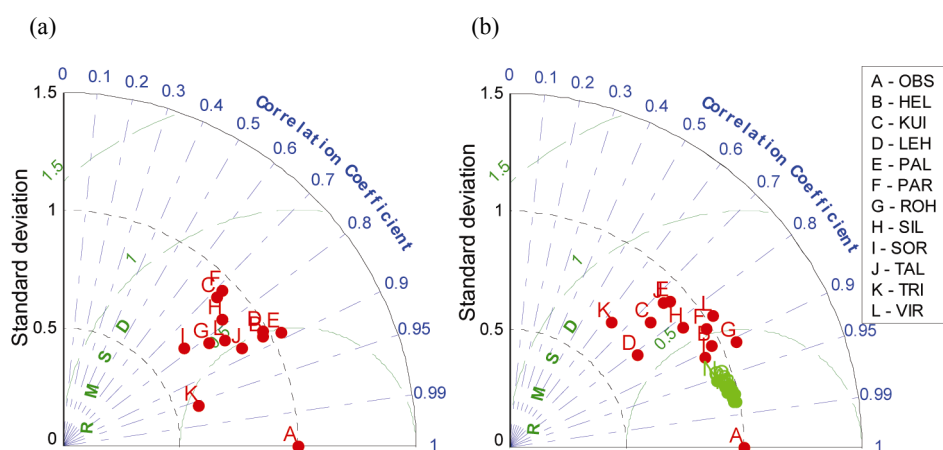
ranges were found for every station (Table 6). These sub-ranges were different from station to station, although an attempt was made to use more general common sub-range limits. The limits were rounded to the nearest 5 cm. Two values for the low sub-range limit were found:  $-50$  cm for 6 stations and  $-60$  cm for 5 stations. The limits of high sub-range varied from  $+15$  to  $+50$  cm. The graphical presentation of the mean deviation of 10 following data points was used to identify the boundaries of different sub-ranges (not shown here). These three sub-ranges were clearly visible for all stations, indicating a relationship between the forecast sea level and error, and making it possible to estimate forecast error while the forecast value is known. While most of the data fall in the medium sub-range and have the lowest errors, the high and low sub-ranges had less data and errors were higher than medium sub-range sea level forecasts. Also, since most of the data was in the medium sub-range, the statistics of this sub-range was very similar to that of the whole range of data.

The forecasting system has a tendency to show an uneven error distribution in case of low and high sea level sub-ranges, while in the medium sub-range the mean errors were near zero, indicating the conformity to normal distribution and adequate low-frequency error correction. Low sub-range error is shifted to the positive side (low sea levels are predicted higher) and high sub-range to the negative side (high sea levels are predicted lower), which should be taken into account when interpreting RMSD values for corresponding sub-ranges. Such a distortion of low and high sea level forecasts was noted also for the northern coast of the Gulf of Finland [20].

From Fig. 10 it could be expected that RMSD of medium sub-range forecasts are significantly lower than in other sub-ranges. Nevertheless, the RMSDs of medium and low sub-ranges are comparable (which is probably the effect of uneven error distribution), while high sub-range forecasts have a higher RMSDs, as expected. The correlations of low and high sub-range forecasts were lower than in the medium sub-range, indicating certain phasing errors in these sub-ranges. The low sub-range had considerably lower STDs, compared to other sub-ranges.

The mean high sub-range forecast error dependence on the forecast length is presented in Fig. 9b. The high sub-range forecast depends slightly more on the forecast lead time than the mean accuracy of the whole dataset. The 48 h forecast accuracy is up to 5 cm worse than the RMSD of the best available forecast, but on average the 6 h forecast is 3 cm more accurate than the 48 h forecast. It could also be noted that the accuracy of the high sub-range forecast is still better than the accuracy of the persistency forecast.

Normalized Taylor diagrams of high and low sub-range sea level forecasts are presented in Fig. 11. The RMSDs and STDs are normalized with the STD of the observations. To illustrate the differences between medium and high/low forecast sub-ranges, the medium sub-range statistics are presented in the same diagram as the high sub-range statistics. Compared to the medium sub-range (and also to the whole data statistics) the correlations and RMSDs are considerably lower. The distribution of STDs is sparser, although the forecast at most stations underestimates the variance of the sea level, following the same trend as general statistics.



**Fig. 11.** Normalized Taylor diagram for low (a) and high (b) sea level forecast sub-range. The green dots in high mode Taylor diagram represent medium sub-range forecast and are shown to illustrate differences between high/low and medium sub-range forecast. Data at each station is normalized with standard deviation of corresponding observations; A indicates observations.

#### 4. CONCLUSIONS AND OUTLOOK

The performance of the Estonian sea level forecast system was analysed within two time periods using different statistical methods. The HIROMB BS01 grid model and data from sea level observation stations was used to produce sea level forecast for the Estonian coast. Using the RMSD minimizing algorithms, a 7 day moving average filter was found to be the most accurate for the correction of low-frequency error. Correlation of this error (dynamic bias) with the Baltic Sea volume change was not found. High sea level analysis within the time period of 2006–2008 showed  $\pm 25$  cm accuracy of storm surge forecasts, taking into account the  $\pm 3$  h time shift between the observation and forecast.

Today the update frequency of the sea level forecast is 4 times a day (+00, +06, +12 and +18), which gives on average 8% more accurate forecast compared to 24 h update interval with respect to the RMSD. The accuracy of the forecast is slightly related to the forecast length, resulting in up to 2.5 cm higher RMSD for 48 h forecast than the best available forecast (6 h). Although an input data quality control algorithm is applied to the forecasting system, which can remove obvious outliers, there is still a need for further calibration of online quality control to increase the reliability of the system. Statistical analysis of the whole period 2009–2011 revealed a RMSD less than 7 cm and correlations above 97% at all stations.

The data was divided into three sub-ranges (low, medium and high), which were analysed separately. The limits of sub-ranges and statistics were found for every station, which showed the differences especially in the high sub-range, where the maximum RMSD was found up to 13 cm and a correlation down to 63%. This method for estimating forecast errors revealed the possibility of implementation of an automatic on-line error estimation system.

As an improvement for the forecasting system it is planned to use more observational stations, especially the stations operated by EMHI. The work is already in progress. Also the implementation of MyOcean products could improve the forecasting system, enabling the possibility for ensemble forecasts of the sea level, which provides more information and better online error estimates for the forecast and therefore is becoming more and more popular around the Baltic Sea. There are 9 new sea level stations recently installed by EMHI, which will be merged into the system soon after the test period. The coverage of the Estonian coastal sea with sea level measurement sites increases remarkably, when new stations are added. Nevertheless, more efforts are needed in order to harmonize the measurement methods. The current forecast system has already showed good performance. With the improved on-line measurement station network, the accuracy of the system should improve even more towards better representation of local peculiarities and features, which in turn are very important in certain storm cases at certain locations of the coast.

Today, at 6 on-line stations out of 11, basic wave parameters are estimated, based on the pressure data, and broadcast to the Sea Level Information System (<http://on-line.msi.ttu.ee/kaart.php?en>) already for a period of 1.5 year. The wave data are used as background information for high resolution sea level measure-

ments. As indicated by the analysis of extreme storm surges, the occurrence of critical and above critical sea levels is the weakest point in the forecast system. The role of waves in forming of extreme sea levels is quite obvious and some statistics is already available, based on the existing time series. The HIROMB models do not include a wave module. Thus, taking into account the local wave pattern in actual wind conditions could improve the performance of the forecast system for extreme sea levels.

## ACKNOWLEDGEMENTS

We thank the Environmental Investment Centre, Estonian Meteorological and Hydrological Institute, Estonian Science Foundation (grant No. 7328) for financial support for the development and maintenance of the forecasting system.

## REFERENCES

1. Ekman, M. A consistent map of the postglacial uplift of Fennoscandia. *Terra Nova*, 1996, **8**, 158–165.
2. Ekman, M. Climate changes detected through the world's longest sea level series. *Global Planet Change*, 1999, **21**, 215–224.
3. Johansson, M., Kahma, K. and Boman, H. An improved estimate for the long-term mean sea level on the Finnish coast. *Geophysica*, 2003, **39**, 51–73.
4. Suursaar, Ü. and Sooäär, J. Decadal variations in mean and extreme sea level values along the Estonian coast of the Baltic Sea. *Tellus*, 2007, **59A**, 249–260.
5. Dailidienė, I., Davulienė, L., Tilickis, B., Stankevicius, A. and Myrberg, K. Sea level variability at the Lithuanian coast of the Baltic Sea. *Boreal Environ. Res.*, 2006, **11**, 109–121.
6. Samuelsson, M. and Stigebrandt, A. Main characteristics of the long-term sea level variability in the Baltic Sea. *Tellus*, 1996, **48A**, 672–683.
7. Miles, J. W. Harbor seiching. *Annu. Rev. Fluid. Mech.*, 1974, **6**, 17–35.
8. Raudsepp, U., Toompuu, A. and Kõuts, T. A stochastic model for the sea level in the Estonian coastal area. *J. Marine Syst.*, 1999, **22**, 69–87.
9. Matthäus, W. and Lass, H.-U. The recent salt inflow into the Baltic Sea. *J. Phys. Oceanogr.*, 1995, **25**, 280–286.
10. Suursaar, Ü., Kullas, T., Otsmann, M. and Kõuts, T. Extreme sea level events in the coastal waters of West Estonia. *J. Sea Res.*, 2003, **49**, 295–303.
11. Wübbler, C. and Krauss, W. The two-dimensional seiches of the Baltic Sea. *Oceanol. Acta*, 1979, **2**, 435–446.
12. Johansson, M., Boman, H., Kahma, K. K. and Launiainen, J. Trends in sea level variability in the Baltic Sea. *Boreal Environ. Res.*, 2001, **6**, 159–179.
13. Jönsson, B., Döös, K., Nycander, J. and Lundberg, P. Standing waves in the Gulf of Finland and their relationship to the basin-wide Baltic seiches. *J. Geophys. Res.*, 2008, **113**, C03004.
14. Lilover, M.-J., Lips, U., Laanearu, J. and Liljebladh, B. Flow regime on the Irbe Strait. *Aquatic Sci.*, 1998, **60**, 253–265.
15. Otsmann, M., Suursaar, Ü. and Kullas, T. The oscillatory nature of the flows in the system of straits and small semiencloded basins of the Baltic Sea. *Cont. Shelf Res.*, 2001, **21**, 1577–1603.
16. Hansen, W. Theorie zur Errechnung des Wasserstandes und der Strömungen in Randmeeren nebst Anwendungen. *Tellus*, 1956, **8**, 287–300.

17. Uusitalo, S. The numerical calculation of wind effect on sea level elevations. *Tellus*, 1960, **12**, 427–435.
18. Peeck, H. H., Proctor, R. and Brockmann, C. Operational storm surge models for the North Sea. *Cont. Shelf Res.*, 1982, **2**, 317–329.
19. Kleine, E. Das operationelle Modell des BSH für Nordsee und Ostsee. Konzeption und Übersicht. Bundesamt für Seeschifffahrt und Hydrographie (Manuscript), 1994.
20. Gästgifvars, M., Müller-Navarra, S., Funkquist, L. and Huess, V. Performance of operational systems with respect to water level forecasts in the Gulf of Finland. *Ocean Dynamics*, 2008, **58**, 139–153.
21. Buch, E., Elken, J., Gajewski, J., Håkansson, B., Kahma, K. and Soetje, K. Present status of the Baltic Operational Oceanographic System. In *Proc. Fourth International Conference on EuroGOOS*. Brest, France, 2006 (Dahlin, H., Flemming, N. C., Marchand, P. and Petersson, S. E., eds). Office of Official Publications of the European Communities, 2006, 276–280.
22. Elken, J., Kõuts, T., Raudsepp, U., Laanemets, J. and Lagemaa, P. BOOS/HIROMB-based marine forecasts in Estonia: problems, experiences and challenges. In *Proc. US/EU-Baltic International Symposium*. Klaipeda, Lithuania, 2006, 22.
23. Suursaar, Ü., Kullas, T., Otsmann, M., Saaremäe, I., Kuik, J. and Merilain, M. Cyclone Gudrun in January 2005 and modelling its consequences in the Estonian coastal waters. *Boreal Environ. Res.*, 2006, **11**, 143–159.
24. Funkquist, L. HIROMB, an operational eddy-resolving model for the Baltic Sea. *Bull. Mar. Inst.*, Gdańsk, 2001, **28**, 7–16.
25. Elken, J., Nõmm, M. and Lagemaa, P. Circulation patterns in the Gulf of Finland derived from the EOF analysis of model results. *Boreal Environ. Res.*, 2011, **16**, 84–102.
26. Axell, L. Weaker surface currents in HIROMB 3.1. In *9th HIROMB Scientific Workshop*. Gothenburg, 2006. [www.environment.fi/syke/hiromb](http://www.environment.fi/syke/hiromb), 16.10.2011.
27. HIROMB Scientific Plan 2009–2013, ver. May 2010. <http://balticlagoons.projects.eucc-d.de/plugins/>
28. Sørensen, J. V. T., Madsen, H. and Madsen, H. Efficient Kalman filter techniques for the assimilation of tide gauge data in three-dimensional modeling of the North Sea and Baltic Sea system. *J. Geophys. Res.*, 2004, **109**, C03017.
29. Taylor, K. E. Summarizing multiple aspects of model performance in single diagram. *J. Geophys. Res.*, 2001, **106**, 7183–7192.
30. Elken, J., Kõuts, T., Lagemaa, P., Lips, U., Raudsepp, U. and Väli, G. Sub-regional observing and forecast system for the NE Baltic: Needs and first results. In *US/EU-Baltic International Symposium “Ocean Observations, Ecosystem-Based Management & Forecasting”*. Tallinn, 2008. IEEE Confer. Proc., 2008, 421–429.
31. Elken, J. Mereproгноosisüsteemi HIROMB arendamine, SA KIK Veekaitseprogrammi projekt nr 36 lõpparuanne. Tallinn, 2008. [http://www.msi.ttu.ee/~elken/HIROMB\\_2008\\_Lopparuanne\\_sisuline.pdf](http://www.msi.ttu.ee/~elken/HIROMB_2008_Lopparuanne_sisuline.pdf), 16.10.2011.

## Operatiivne meretaseme prognoosisüsteem Eestis

Priidik Lagemaa, Jüri Elken ja Tarmo Kõuts

On kirjeldatud ja valideeritud operatiivset meretaseme prognoosisüsteemi Eestis. Prognoosisüsteem põhineb HIROMB-mudeli prognoosil ja 11-l reaajas toimival meretaseme mõõtejaamal. Süsteem töötab alates 08.08.2005. Töös on analüüsitud prognoosisüsteemi efektiivsust aastail 2006–2008 ja 2009–2011. On uuritud madalsagedusliku veetaseme prognoosivea omadusi ja hinnatud prognoosiviga ekstreemsete veeseisude esinemise korral, mille täpne ennustamine on

prognoosisüsteemi põhirakenduseks. Uuring näitab, et 7-päevane tagasiulatuva liikuva keskmisega filter on Eesti rannikul sobivaimaks meetodiks, et prognoosimudeli väljundandmetest madalsageduslikku viga välja filtreerida. Ekstreemsete (üle kriitilise taseme) veetasemete prognoosi täpsus on  $\pm 25$  cm ja ajaline maksimaalse meretaseme saabumise täpsus  $\pm 3$  tundi. Värskemate, 2009.–2011. aasta meretaseme prognooside hindamiseks kasutatakse Taylори täpsushindamise meetodit. Reaalajas prognoosivea arvutamise võimalikkuse uurimiseks jagatakse meretaseme prognoosid kolme alamrühma (madal, keskmine ja kõrge). On leitud, et madalamate ja keskmiste veetasemete prognoosid on täpsemad kui kõrgemate veetasemete omad. On vaadeldud prognoosisüsteemi edasiarendamise võimalusi ja põhjendatud vajadust parendada reaalajas toimivat mõõteandmete kvaliteedikontrolli.