Journal of Marine Systems 83 (2010) 210–220

Contents lists available at ScienceDirect

Journal of Marine Systems

journal homepage: www.elsevier.com/locate/jmarsys

Ensemble forecasting of harmful algal blooms in the Baltic Sea

Petra Roiha *, Antti Westerlund, Aleksi Nummelin, Tapani Stipa

Finnish Meteorological Institute, Erik Palménin aukio 1, P.O.Box 503, FI-00101 Helsinki, Finland

ARTICLE INFO

Article history:
Received 1 October 2009
Received in revised form 12 February 2010
Accepted 16 February 2010
Available online 4 March 2010

Keywords:
Prediction
Algal blooms
Ensemble forecasting
Baltic Sea
(8.7° E 53.85° N)
(30.3° E 65.85° N)

ABSTRACT

Operational marine environmental modelling has been considered notoriously difficult; consequently there are very few operational models of the marine environment. Operational modelling of harmful algal blooms (HABs) requires the modelling of individual species and is therefore harder still. The separation of algal species in models requires detailed knowledge of their behaviour (survival strategy through the life cycle), and physiological ecology.

We present quantitative results of an ensemble approach to HAB forecasting in the Baltic, and discuss the applicability of the forecasting method to bioheochemical modelling. Ensembles were produced by running a biogeochemical model several times and forcing it on every run with different set of seasonal weather parameters from European Centre for Medium-Range Weather Forecasts' (ECMWF) mathematically perturbed ensemble prediction forecasts. The ensembles were then analyzed by statistical methods and the median, quartiles, minimum and maximum values were calculated for estimating the probable amounts of algae. To evaluate the forecast method final results were compared against available and valid in-situ HAB data in a case study. It turns out that quantitative HAB forecasts are possible. Further verification will require expanded observational networks.

1. Introduction

Modelling of non-linear variables always includes uncertainties from different sources. The initial conditions might be inaccurate, the model input has errors and the modelling of ocean conditions is not exact because of truncation errors and inaccuracies in the modelling of subscale phenomena (Leutbecher and Palmer, 2008). Therefore the solution deteriorates in time.

The deterioration of model forecasts with time is a well-known issue in weather forecasting, where the reliable forecast range is about a week. In oceans the predictability of some phenomena is typically longer. For example, the internal weather of the sea, the oceanic mesoscale, includes mainly phenomena which are occurring in temporal scales from days to months and spatial scales from kilometres to hundreds of kilometres (Lermusiaux, 2006). However, ocean predictability is rarely exploited to its useful limits. Most operational ocean forecasting is limited to ten day forecasts in the maximum, or to coupled atmosphere–ocean seasonal forecasts where the focus is nevertheless on the atmospheric forecast.

Physical ocean models are principally built on the same well-known and relied upon equations as atmospheric models, and have similar inherent limitations to their predictive skills. Biogeochemical models, on the other hand do not enjoy a firm basis provided by e.g. the primitive equations of ocean motions. Furthermore, the initial condition for a biogeochemical forecast is often not well observed, and the uncertainties are large. While there are significant constraints to how biogeochemical models are to be constructed (Redfield, 1958), the uncertainties involved in using biogeochemical models for forecasting appear to deserve an explicit treatment. This is particularly true with models intended to predict not only the overall biogeochemical processes, but also the behaviour of individual species.

In harmful algal bloom (HAB) forecasts a relationship between phosphorus concentration and cyanobacterial blooms has been recognized for decades (Niemi, 1979; Niemistö et al., 1989; Kahru et al., 2000). This relationship has been utilized for practical and even operational purposes. These operational estimates of cyanobacterial bloom probability and severeness have, however, been based mostly on the wintertime (January–February) nutrient concentration fields and best and worst case scenarios for summer weather conditions, without accounting for the actual weather development and forecasts.

Janssen et al. (2004) demonstrated with model experiments that a relationship between winter nutrients and summer cyanobacterial blooms, in agreement with the inferences of Kahru et al. (2000, 2007), is replicated with biogeochemical 3-dimensional numerical model.

Biogeochemical ensemble forecastss offer a quantitative tool for the assessment of HAB related environmental risks for a wide range of applications. Ensembles have been an essential tool in meteorology for many years. In comparison with a single deterministic forecast, ensembles offer the benefit of estimates of bias, deviation and range of the modelled variables from real life situation. It is also possible to

* Corresponding author.
E-mail address: petra.roiha@fmi.fi (P. Roiha).

0924-7963/$ – see front matter
doi:10.1016/j.jmarsys.2010.02.015

analyze the ensembles and recognize forecasts with low skill (Buizza, 1997).

There are several ways to dissect the ensemble. Variables can be studied by e.g. calculating the ensemble mean, which provides an estimate of the probabilistic expectation forecast. The ensemble can also be divided into smaller sub-ensembles to make alternative forecasts (Brankovic and Palmer, 1997) and even individual members can be used for prediction purposes. Ensembles can be used as a quantitative tool for risk assessment. The potential economic value can be much higher in many applications than the value of deterministic forecast (Richardson, 2000).

In this work we explore the predictability of ensemble HAB modelling and demonstrate the usefulness of such forecasts in environmental policymaking and risk assessment. The operational ocean model used in this study has been examined closely by Kiiltomäki (2008) and accuracy of the deterministic model has been evaluated against observations. Previously it has been used for bloom forecasting as a part of ad hoc ensembles with the help of expert insight, cf. FIMR and SYKE (2006). Our system widens this approach by formalizing the forecasting process in a computationally sound manner. Our approach brings greater consistency of the forecasting process with actual weather forecasts. Causes for the phenomena seen in the results can be traced more easily, giving more information for forming the final forecast.

2. Materials and methods

2.1. Model configuration

For forecasting we used FMI (Finnish Meteorological Institute) operational 3-dimensional biogeochemical model, Baleco. The model consists of a general circulation model (Marshall et al., 1997a,b) and an ecological module. The model is discretized on a spherical polar grid. The grid size is 0.1° in longitudinal and 0.2° in latitudinal direction (around 11.1 km, or 6 nautical miles) and the model domain reaches 120 grid cells in latitudinal, 108 grid cells in longitudinal and 21 grid cells in vertical direction. The model domain's most south-western corner is located at (53.85° N, 8.7° E). The vertical resolution of the model is concentrated to the euphotic zone so that the top most layer is 3 m, reduced to 2 m in the cells hugging the coast (Kiiltomäki, 2008). The bottom topography (Fig. 1) is from Seifert and Kayser (1995). The spatial discretization is made with minimum filter at 6 nm intervals.

The ecological model is based on ecosystem dynamics formalized by Aksnes et al. (1995) and Tyrrell (1999). The model consists of three phytoplankton groups: diatoms, flagellates and cyanobacteria. These groups have constant mortality rates and they use phosphate, silicate and dissolved inorganic nitrogen. Diatoms are potentially limited by availability of silicate. Cyanobacteria can fix molecular nitrogen and therefore they are not limited by availability on DIN. The flagellates group represents autotrophic flagellates. Altogether the ecosystem model describes the essence of new production in the presence of three functional groups. The growth rates depend on nutrient concentrations, irradiation and temperature (Stipa et al., 2003). For model equations see Appendix A.

Model runs obtained their initial state from FMI's deterministic Baltic Sea forecast for the start date of the run. The deterministic forecast's initial conditions, both physical and chemical, were obtained from winter monitoring data of the HELCOM COMBINE program for the winter 2007–2008. The observations were interpolated in three dimensions with a robust nearest neighbour interpolation and then supplemented by climatological values for the North Sea from the World Ocean Atlas (Boyer et al., 2006).

2.2. Ensemble forecasts

The ensemble prediction system (EPS) is a technique to predict the probability distribution of forecast states, given a probability distribution of random error in inputs and model error. Ensemble forecasts are formed by several slightly perturbed ensemble members.

Ensemble forecasts include more information than a single deterministic forecast and therefore the analysis can give us a deeper insight to many phenomena. Every ensemble member represents one possible evolution of the system state in time and space. Therefore the variety of applicable analysis approaches is very wide, when compared to deterministic forecasts.

One of the oldest and simplest analysis methods is to calculate ensemble mean to define the mean trajectory. It is also possible to calculate some statistical values for analysis. These values can be, for example, minimum and maximum values, which indicate the extreme values, and 25% and 75% quartiles of the ensemble spread. Another useful approach is to compute the percentage of ensemble members for which a given variable exceeds some limiting value, which can then be interpreted as a probability of the event.

Ensembles in our study were created from an unperturbed initial condition by running the model several times with different sets of weather boundary conditions. The weather ensembles were made by ECMWF using singular vector method (Molteni et al., 1996). Weather parameters used as external forcing for ocean model were 6 hourly 10 m winds and 2 m dew point temperature and 2 m temperature, 12 hourly surface solar radiation and surface thermal radiation. These weather ensembles include 50 perturbed ensemble members and an additional deterministic unperturbed control run. These were used to generate a 28-day ensemble runs for the Baltic Sea for June, July and August of 2008. ECMWF makes new monthly forecasts available once a week, so for this three month period this meant altogether 13 ensemble runs, of which some were chosen for further analysis.

The ensemble forecast produced in this work gives the probability of cyanobacteria concentrations. The forecast shows where the harmful algal concentrations are high and thus the appearance of blooms is more likely.

2.3. Chlorophyll-a conversion

As the model gives the amount of cyanobacteria in molar amount of nitrogen (N) we have used a special *C*:*N* ratio, the molar mass number of carbon (12.01) and *C*:*Chla* ratio to calculate the chlorophyll-a concentrations in mg m^{-3}. Since this work concentrates on the Baltic Sea, we have used *C*:*N* = 6.3, which is based on studies made in the area: Walve and Larsson (2007) found out that *C*:*N* ratio in the cyanobacteria in Western Gotland basin was 6.2–6.4 ± 0.3, the highest ratio being 7.3. Another study made by Nausch et al. (2009) supports these values as they found that the mean *C*:*N* ratio was 6.2 in the Eastern Gotland basin.

The Carbon:Chlorophyll-a ratio η is also variable, and depends on properties of the study area, the state of the bloom and most of all on the algal species studied, e.g. it is known that *C*:*Chla* ratio is usually larger in cyanobacteria than in other algae as they have also other pigments that can be used in light harvesting (Geider et al., 1997). Geider et al. found the minimum *C*:*Chla* to be approximately 38–67 for different kinds of cyanobacteria, although the study was not done in the Baltic Sea. Eker-Develi et al. (2008) found the mean *C*:*Chla* ratio for cyanobacteria to be 33 in Southern Baltic Sea, although there was significant variability, standard deviation being 35. In addition to these findings Engström-Öst et al. (2002) reported a high *POC*:*Chla* (<10 μm) ratio, 427 ± 185, during cyanobacterial bloom decay, although POC does not equal C exactly. Based on these results the *C*:*Chla* ratio is a significant source of uncertainty. This is further discussed in Section 4.3.

With these configurations the chlorophyll-a (*Chla*) concentration in mg m^{-3} was calculated from the model results with the formula

$$Chla = \frac{N_c \cdot 6.3 \cdot 12.01}{\eta} \approx \frac{N_c}{\eta} \cdot 76, \qquad (1)$$

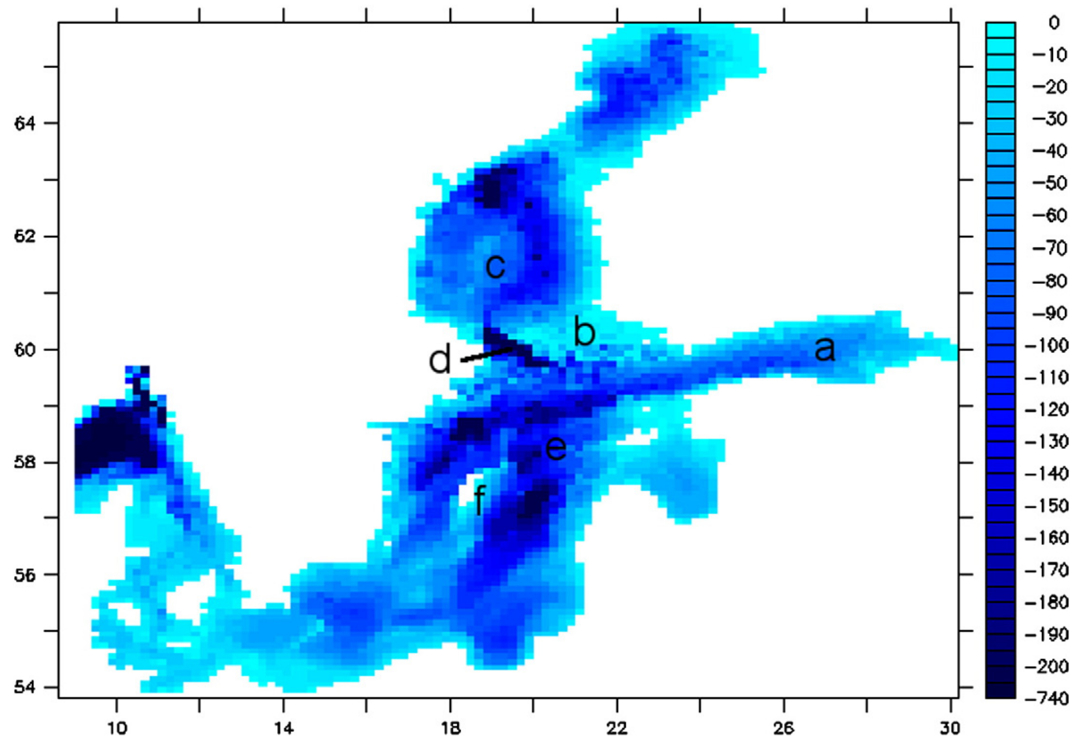

Fig. 1. Bottom topography (m) of the Baleco model. Also indicated on the map are the following geographic references used in the article: a) Gulf of Finland, b) Sea of Archipelago, c) Bothnian Sea, d) Sea of Åland, e) Baltic Proper and f) Gotland.

where N_c is the molar mass of cyanobacteria produced by the model. The C:$Chla$ ratio, η, is given for specific events separately. As a limit for the cyanobacterial occurrence we used chlorophyll-a concentration of 2.0 mg m^{-3} which is based on studies by Seppälä and Balode (1999), Nausch et al. (2004) and Kutser et al. (2006).

3. Results

3.1. Case study: harmful algal blooms and upwelling in the Baltic Sea, June 2008

The summer weather of 2008 was cool and windy and thus large scale HABs were absent. The overall algal bloom situation can be seen in Fig. 2. In June and August blooms were milder than usual, in July the situation was average (HELCOM, 2009). In the beginning of July the sea water temperature increased a few degrees and the concentration of cyanobacteria increased in the Gulf of Finland. Cyanobacteria were mainly mixed in the water column. The Sea of Åland had some blooms. From the middle July to end of July the algal blooms increased first in southern Sea of Archipelago, Sea of Åland, eastern Gulf of Finland and southern Bothnian Sea. In the end of July the blooms were more frequent and small blooms were present all along of the Gulf of Finland. The summer bloom culminated in the end of July, when the maximum extent (approximately 180,000 km^2) of algae blooms was observed. However the normalized bloom indices for bloom extent (6575 km^2), duration (4.9 days) and intensity (32,651 km^2 days) were lower than the mean for the period 1997–2007.

In our case study we saw an upwelling on the east coast of Gotland on the 24th of June. We show here how different weather scenarios affect ocean conditions and how ensemble forecasts see different biogeochemical phenomena.

Upwelling is very often strong enough for the sea-surface temperature to be affected, and thus low temperatures in a thin strip near the coasts are a signature of upwelling. Upwelling can also be seen in the colour of the water and in the abundance of sea life (Gill, 1982).

Temperature and nutrient ensembles (Fig. 3a, c and d) and cyanobacterial concentrations (Fig. 4) show an upwelling event in the area. It can be seen that some weather scenarios cause upwelling while other possible scenarios do not. The cyanobacteria ensemble (Fig. 3b) and biomass observations (Fig. 5) from Algaline's automated ferrybox sampling (Ruokanen et al., 2003) show evidence of upwelling. The lower cyanobacteria biomass concentrations observed near the coast are captured by a large number of ensemble forecast members.

3.2. Comparison between observations and forecast

It is desirable to evaluate the quality of a probabilistic prediction system not only in terms of the intrinsic quality of the results it produces but also in terms of cost efficiency (Talagrand et al., 1999). In HAB validation the available observational data is very sparse and often qualitative in nature. Therefore the validation is especially challenging.

Fig. 6 shows the observed situation in the end of July 2008, when the HAB was at its peak. The observations are done by volunteers, environmental authorities and Finnish border guards. Several kinds of observations are combined in this figure. Observations are mainly visual especially near the Finnish coast and there is no data about biogeochemical variables during the phenomenon.

Contrast this with Fig. 7, showing HAB probability maps produced from an ensemble forecast for the same time. Unlike the observations, these maps are based on probabilities of quantitative values of cyanobacteria chlorophyll-a as predicted by the model. Therefore, it is not predicting directly the concentration of chlorophyll-a but rather the possibility of the blooms. This is further illustrated in Fig. 8 showing single members of the ensemble.

From Figs. 6 and 7 it can be seen that in the Gulf of Finland there are several observations of HABs in the areas where they were predicted. We can also see that in the Northern Baltic Proper the predicted bloom area is considerably to the East of the observed blooms, although with lower values of η the edge does move westward. Furthermore, there

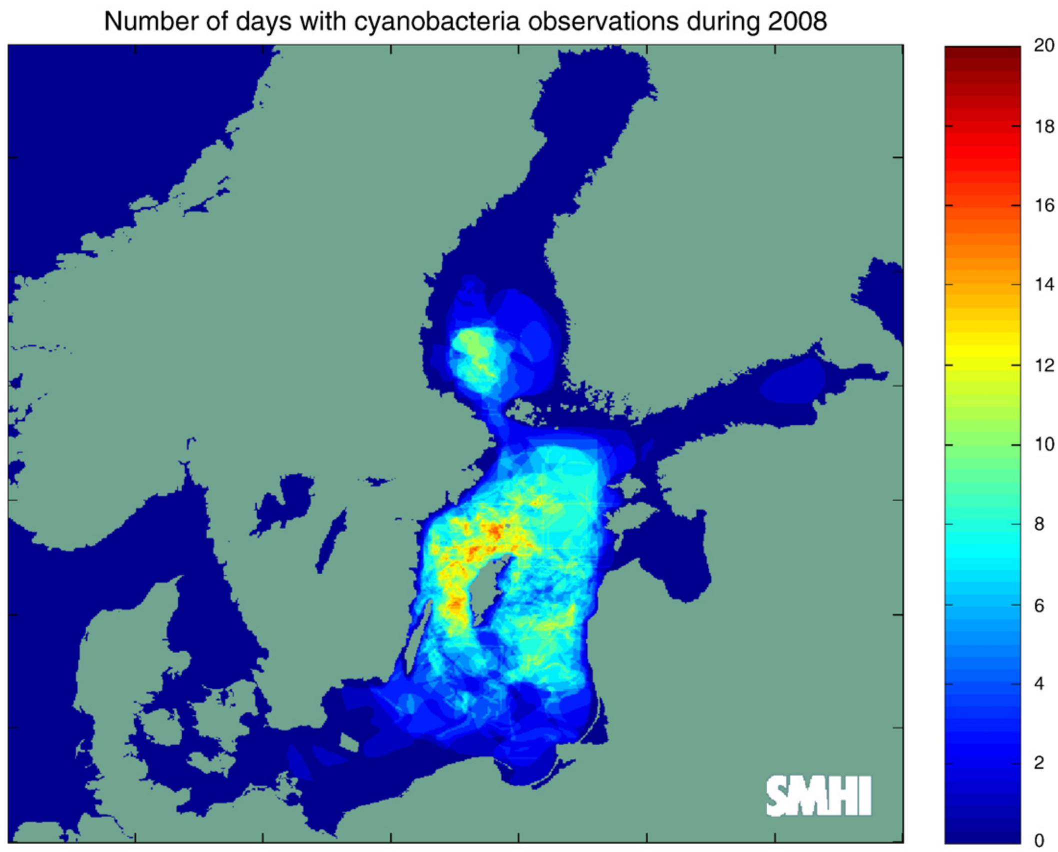

Fig. 2. Number of days with cyanobacteria observations by NOAA-AVHRR satellite imagery during 2008 (HELCOM, 2009). Courtesy: SMHI.

are no blooms forecasted for the Sea of Archipelago, yet there are several observations of blooms.

4. Discussion

Large parts of Baltic Proper and Gulf of Finland are nitrogen-limited, resulting in excess phosphorus in the surface layer in late summer. This excess DIP pool has the potential to stimulate blooms of nitrogen-fixing cyanobacteria. That is, the potential of late summer bloom is determined as early as February by the excessive DIP concentration in the surface layer. This correspondence has been studied by Kahru et al. (2000). In later studies Janssen et al. (2004) have reached the same conclusion with the computational models and Lilover and Stips (2008) based on analysis of observations.

Our ensembles demonstrate the sensitivity of HABs to nutrient fields, especially phosphate. Similarities can be seen in the distribution of the computational initial phosphate field after the spring bloom at the end of May (Fig. 9b) and probability based forecast (Fig. 7a). The observed initial phosphate field (Fig. 9a) has similar high DIP concentrations in the Baltic Proper and western parts of Gulf of Finland, but concentrations in the Bothnian Sea are higher only in the initial winter DIP field.

The initial nutrient field plays an important role in algae growth and it appears at first that spatially the nutrient field is a more dominant factor in algal blooms than the weather conditions. However, during the summer the meteorological variables have a great impact on timing, duration and intensity of the blooms as can be seen in Fig. 3. For example Fig. 3b shows that unfavourable weather conditions can delay the cyanobacteria growth for weeks. It is also seen that the concentration of cyanobacteria chlorophyll is strongly dependent on weather conditions.

This dependance on weather conditions led Kahru et al. (2007) to suggest that only basin wide forecasts of frequency of cyanobacterial accumulations are useful. As this work deals with forecasting probability of blooms instead of bloom frequency and we use a circulation model with actual weather forecasts as inputs, it is possible that our approach is suitable on sub-basin scales. This notion would require further investigation, however.

The biological component of the model used in this study is a simplified yet robust representation of the diverse natural ecosystem, thus it has only three different algal groups and three different nutrients related to each other with relatively simple equations shown in Appendix A. However, the verification work of Kiiltomäki (2008) has showed that the biogeochemical component is fit for this kind of work. Furthermore, it has been shown in Section 3.1 that the model responds well to the changing weather conditions and the dynamics are accurate enough as the decrease in cyanobacterial concentration during the upwelling event is reproduced by the model. Furthermore, the benefit of a more complex biological component is not apparent because of other uncertainties such as the strong effect of the initial nutrient field. It would also require more computing time which is not an advantage for an operational model, especially when it comes to computationally demanding ensemble forecasts. Therefore it stands to reason for this purpose that it might be more useful to increase the accuracy of initial nutrient field and nutrient input than to further develop the relatively accurate biogeochemical component.

As the initial condition data was based on somewhat sparse wintertime observations it would be interesting to further investigate whether perturbating the initial nutrient conditions could enable the quantification of the resulting uncertainty in the forecasts. However, a perturbation scheme for marine biogeochemical forecasts remains to be developed.

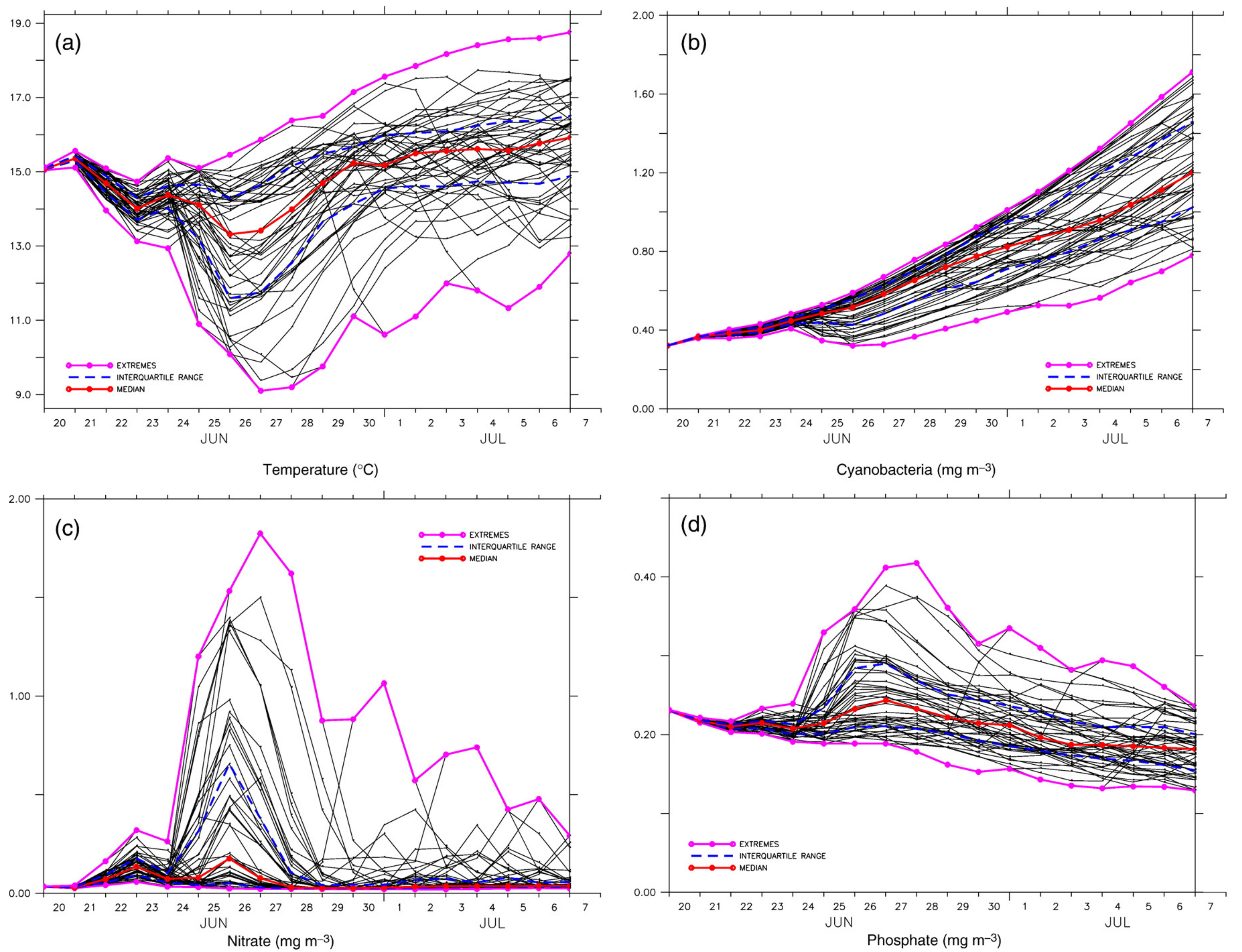

Fig. 3. Monthly ensemble forecast plumes of biogeochemical parameters from the beginning of the forecast 20th of June to 7th of July 2008 on the east coast of Gotland (18.80° E 57.25° N).

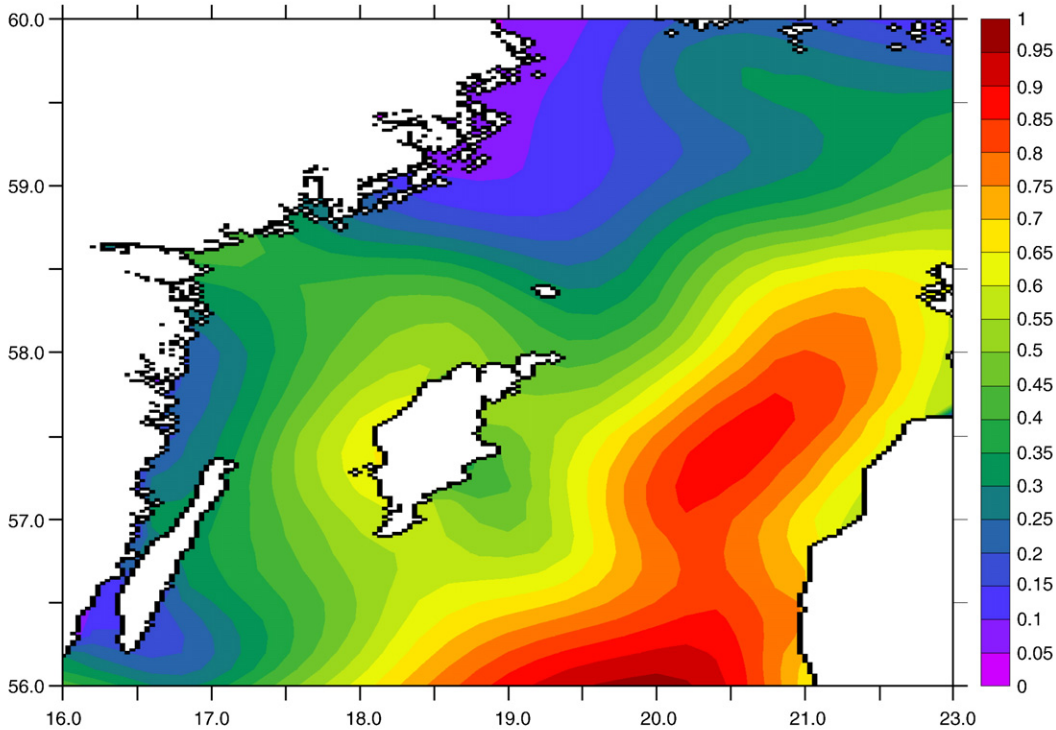

Fig. 4. Cyanobacterial ensemble mean forecast (mg m^{-3}) for the 1st of July from the run beginning from 20th of June ($\eta = 100$). Lower concentrations on the coast of Gotland (18.80° E 57.25° N) indicate upwelling.

4.1. Applications

Ensemble forecasts of the marine environment have a variety of possible applications.

Environmental indicators, derived quantities describing state of the environment in an understandable manner, provide information needed for decision makers to mitigate environmental problems. The range of decision makers varies from individual consumers to politicians, and the indicators should be helpful in making environmental decisions.

For several years computational models have been used as a decision support tool for policy makers. End users have found the

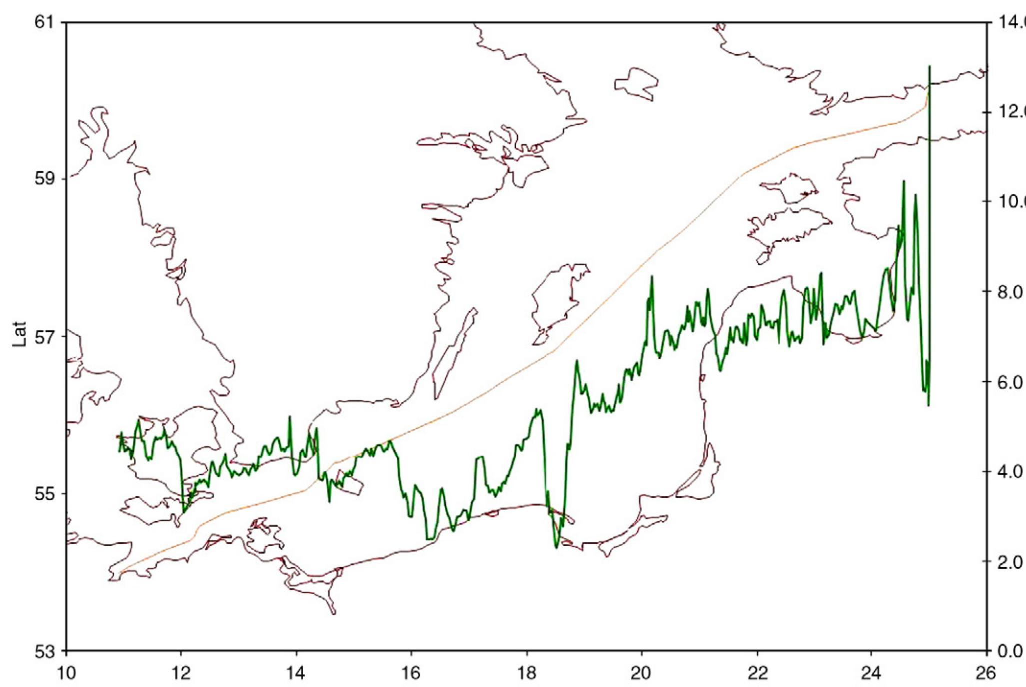

Fig. 5. Algaline fycocyanin (blue–green algal biomass) observations (PC fluorescence) on 1st and 2nd of July 2008 on the route from Helsinki to Travemünde. On the east coast of Gotland the values are clearly lower.
Courtesy: Finnish Institute of Marine Research

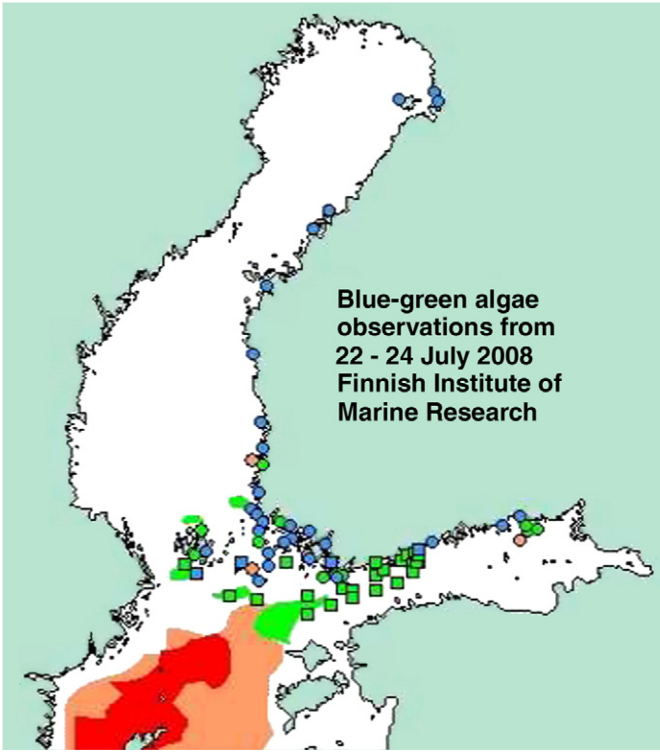

Fig. 6. Combined harmful algal bloom observations from 22nd to 24th of July 2008. Red colour is for very abundant blooms, orange is for abundant blooms, green means there is some algae and blue that there is no algae. This composite picture is based on qualitative visual observations, there are no concentration data included.

products to be useful and fit for that purpose. The added value of the ensemble approach comes from the probabilities, which give insight to what is likely to happen.

With ensemble forecasts it is also possible to illustrate how weather conditions affect HABs. It is a well-known fact that cold and windy weather prevents heavy blooms while sunny and calm weather promotes them. HAB maps can illustrate the probability of blooming instead of absolute amount of cyanobacterial biomass.

Applications developed by working with the users and recognizing their individual needs are essential when adding value for ensemble forecasts.

4.2. Illustrations of probability based forecasts

Because of the large information content of ensemble forecasts, one major challenge is to communicate the results and their uncertainties to users. In deterministic forecasting these kind of problems simply do not appear because there is less information available on the uncertainties.

Recreational users need different kinds of information than, for example, policy makers or professional users. These differences should be taken into account in communication with user groups. There are different ways to solve the problem with probability based forecasts and in this work we have presented some of them. For instance, professional users may benefit from the more detailed information provided in Fig. 3, while recreational users might appreciate the more easily approachable map based presentation like the one in Fig. 7a. Customization of communicated message to suit end user needs often works for the benefit of all partners.

4.3. Carbon:nitrogen and Carbon:chlorophyll-a ratios

Carbon:nitrogen:phosphorus stoichiometry is one of the most discussed topics in marine biogeochemistry and no final agreement on this relationship has yet been found. The most widely used ratio is so called Redfield ratio, C:N:P = 106:16:1. There have been many studies on the subject and it has been pointed out by Arrigo (2005) that the Redfield ratio is more an average than a universal constant, and a single measurement, especially if made in a coastal region, can differ significantly from it. Because cyanobacteria are phosphorus limited in the Baltic Sea, as they can fix nitrogen, there is no need to take phosphorus into account in the chlorophyll-a conversion since the model treats algae as a nitrogen reservoir. As discussed in Section 2.3, we used C:N ratio of 6.3 for chlorophyll-a conversions, which is slightly lower than Redfield's ratio which gives C:N = 106:16 = 6.625.

Although the carbon:nitrogen ratio has its effects on chlorophyll-a conversion, a more significant source of error is the carbon:chlorophyll-a ratio η. As illustrated in Fig. 7 the value of this ratio affects the results considerably. Because this ratio is so poorly constrained in this region it is hard to argue which of the forecasts shown would be the most proper one. Achieving a more accurate C:$Chla$ ratio would require information at least on the mean η of cyanobacteria in the Baltic Sea. This could then be used as a best prior estimate and changed according to the state of the bloom, although another question is how the state of the bloom can be determined. Geider et al. (1997) have presented some modelling approaches that could be used to resolve this variability. However the determination of η is not on the scope of our article, but we note the significance and uncertainty related to it.

The uncertainties in the nitrogen to chlorophyll-a conversion has its effects also on the accuracy of the forecast when verified against observations. Variation in conversion values adds uncertainties which could not be reproduced by this model setup. A biased conversion value will also cause bias to the results irrespective of actual model skill.

4.4. Limiting value for HABs

In Section 2.3, we determined a chlorophyll-a concentration of 2 mg m^{-3} as a limit for a possible visible cyanobacterial occurrence (see Section 2.3) Defining such a limit, however, for cyanobacterial bloom is difficult due to the lack of standards and especially because most bloom observations are based on visual approximation. However there are some studies where measurements have been done during cyanobacterial blooming in the Baltic Sea area (Seppälä and Balode, 1999; Nausch et al., 2004; Kutser et al., 2006; Mazur-Marzec et al., 2006). Kutser et al. (2006) suggested that blooming can be defined as chlorophyll-a concentration exceeding 4 mg m^{-3}. Mazur-Marzec et al. (2006) found that chlorophyll-a concentration was round 10 mg m^{-3} or more during blooming in Gulf of Gdańsk summer 2004. Therefore, taking into account the uncertainties in the nitrogen-chlorophyll conversion, 2 mg m^{-3} is a conservative limit for a level of biomass that could be perceived as a harmful or nuisance bloom.

5. Conclusions

Ensemble forecasting appears to be a promising tool in operational oceanography. The probability based approach illuminates the uncertainty of modelled phenomena. Stable conditions create more unanimous ensembles and vice versa.

Spring-time phosphorus fields are a relatively good predictor for the spatial, basin-scale distributions of HABs in the summer. The spatial variation of forecasted blooms is relatively small.

The weather conditions, however, clearly have an impact on timing, duration and intensity of HABs. This variation can be observed from and quantified with the ensemble forecasts in a manner that

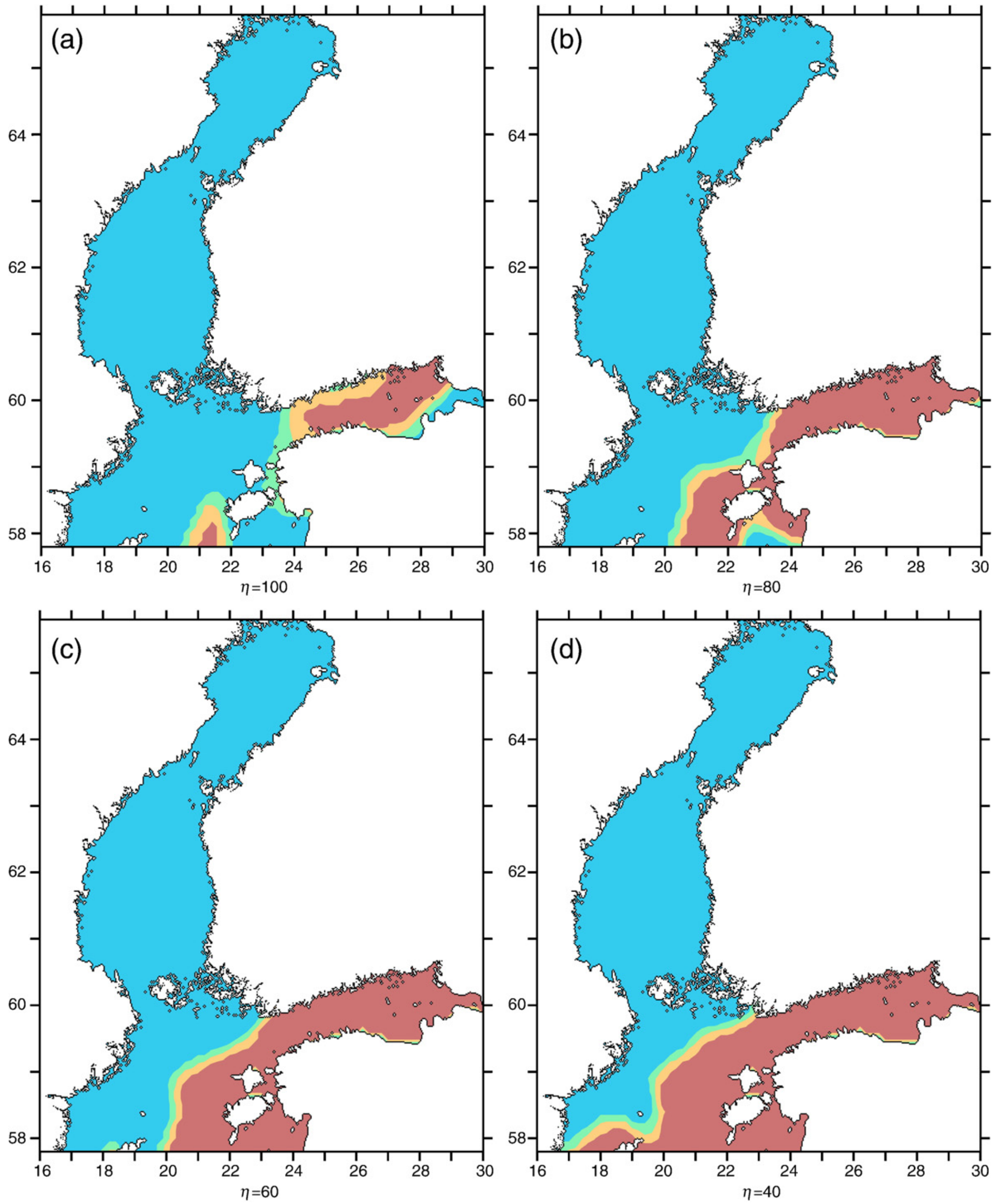

Fig. 7. Harmful algal concentration ensemble forecast with limiting value of 2 mg m^{-3} for blooming, with different values of *C*:*Chla* ratio (from 100 in 6 to 40 in 6). This two week harmful algal probability forecast was formed from an ensemble run starting from 10th of July (cf. Fig. 8). Red colour indicates high probability of blooms (>75%) at the end of the forecast, yellow considerable probability (50%–75%), green moderate probability (25%–50%) and blue low probability (<25%). This harmful algal forecast depends among other things on the value chosen for *C*:*Chla* ratio η as shown in Eq. (1). The forecast map changes notably with different values of η ranging from 100 to 40.

supports pre-emptive actions, at least against exposure to adverse health effects.

When developing tools for marine policy making, the quality of the modelled results should be well known and the quantification of errors should be considered. Quantitative verification of the HAB ensemble forecasts, however, is very challenging because of the limited amount of observational data, and the difficulties in matching the observed variables with the predicted variables.

Both the model's sensitivity to initial conditions and the challenges faced with verification suggest that HAB modelling would greatly benefit from an increased amount of relevant observations. Verification would become much easier if quantitative information about biomass concentrations in the Baltic Sea during the summer was available. Gaps in the winter time phosphate monitoring measurements can lead to notable shortcomings in summer's bloom forecasts, and would need either more observations or a highly sophisticated

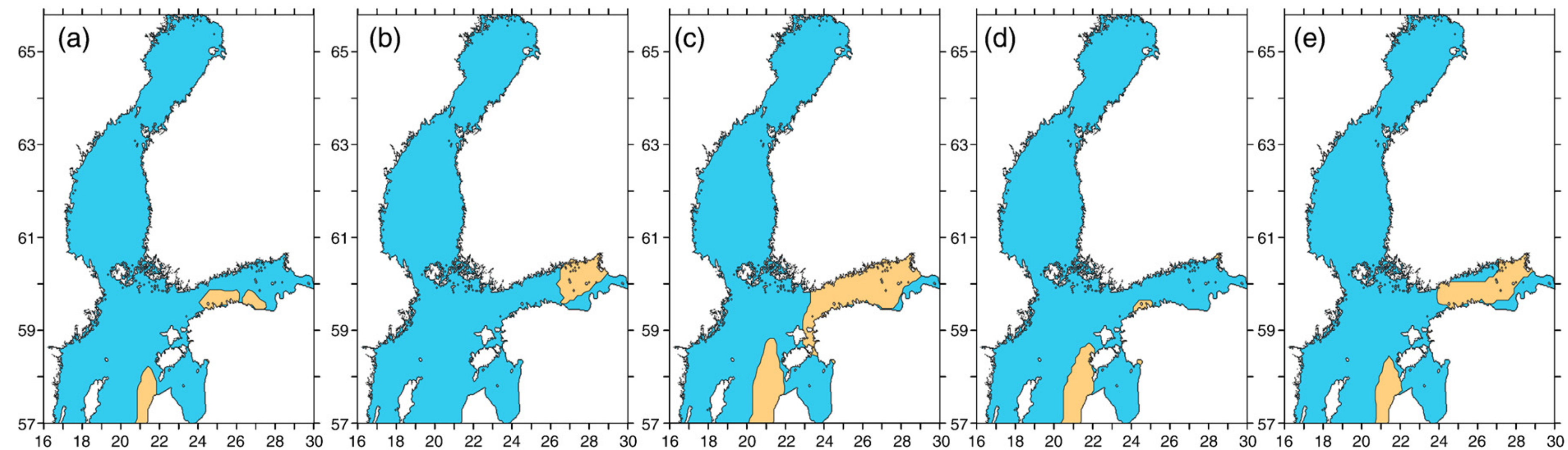

Fig. 8. Probabilistic ensemble harmful algal forecasts, such as those in Fig. 7, are formed from multiple ensemble members. Percentage of ensemble members that exceed chosen limiting concentration are interpreted to have potential for bloom formation during the forecast interval. This figure shows a selection of forecasts from an ensemble of 50, showing in yellow areas where concentration exceeds the chosen limit of 2 mg m^{-3}. These ensemble members (Fig. 8a, b, c, d and e) have been selected to represent the typically observed variance within ensembles. It can be seen how the extent of possible bloom areas depends on weather conditions. For instance, some ensemble members predict no bloom potential in the Gulf of Finland (Fig. 8c), while others show almost all of the basins having potential for blooms during the forecast period (Fig. 8d, for example).

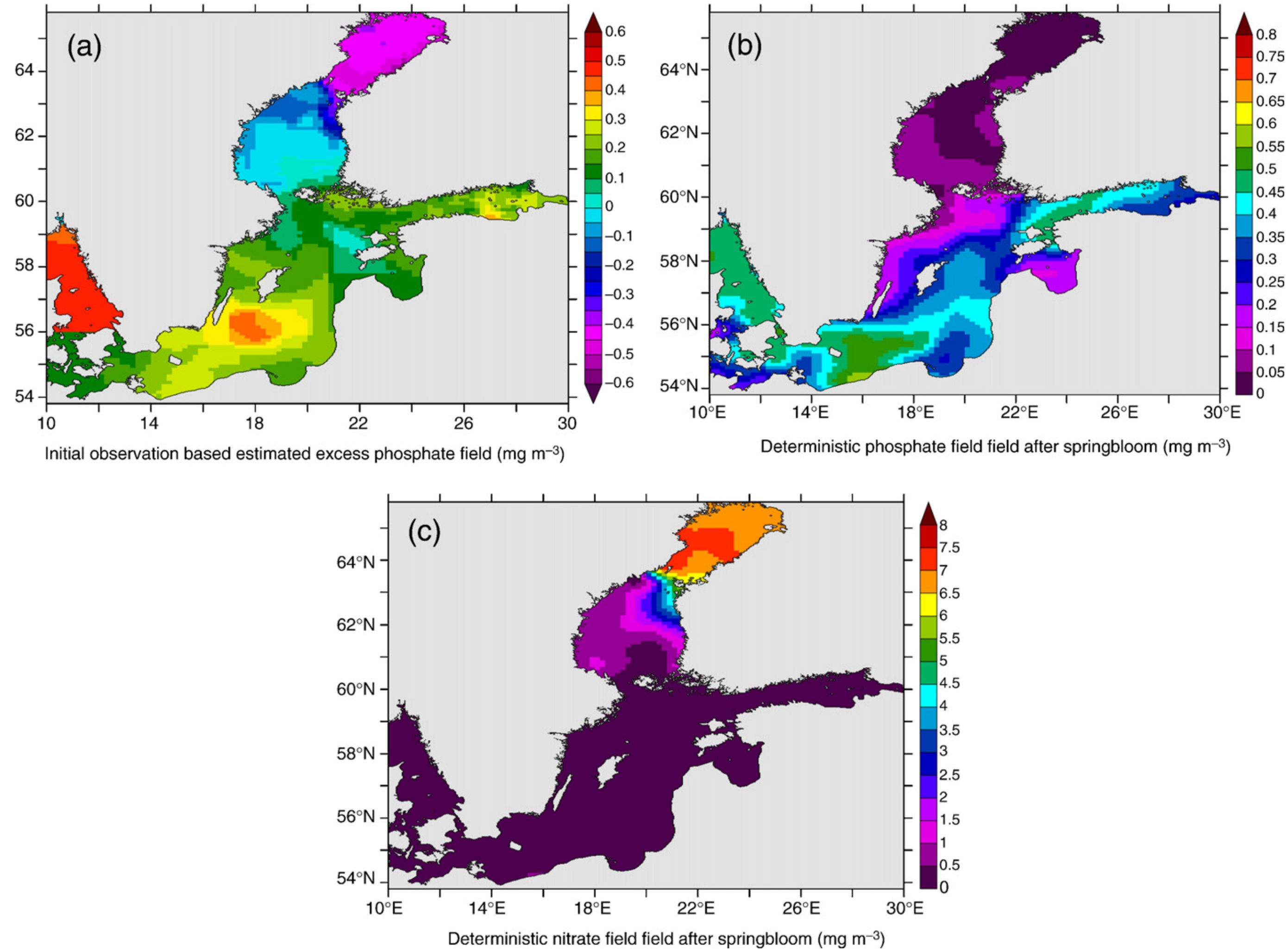

Fig. 9. Fig. 9a shows initial excess phosphate field (DIP-DIN/16) interpolated and estimated from nitrate and phosphate observations made by R/V Aranda in winter 2008. Fig. 9b shows phosphate after spring bloom in the end of May 2008. Field is computed from observed initial fields by deterministic model. Fig. 9c is as Fig. 9b but for nitrate. We can see that spring bloom has depleted nitrate field in areas where cyanobacterial blooms typically occur. Therefore Fig. 9b can in this context be interpreted as a more complex, model generated estimate of the simple calculation shown in Fig. 9a.

assimilation scheme to fill. Should these kind of measurements and better assimilation schemes become available, we expect the benefits of the ensemble approach to become even more pronounced.

Acknowledgements

This work has been partly financed by the EU Commission 6th framework project ECOOP European COastal-shelf sea OPerational Observing and forecasting system (contract number 036355). We thank also the Algaline project for making the ferrybox data available. We would also like to thank our colleague Letizia Tedesco for her valuable comments.

Appendix A. Ecological equations

The equations in the ecological model are

$$\frac{\partial c_d}{\partial t} = c_d(\mu_d - e_d - m_d c_d) \tag{A.1}$$

$$\frac{\partial c_f}{\partial t} = c_f(\mu_f - e_f - m_f c_f) \tag{A.2}$$

$$\frac{\partial c_c}{\partial t} = c_c(\mu_c - e_c - m_c c_c) \tag{A.3}$$

$$\frac{\partial c_N}{\partial t} = -c_d(\mu_d - e_d) - c_f(\mu_f - e_f) - c_c(-e_c) \tag{A.4}$$

$$\frac{\partial c_P}{\partial t} = r_{PN}(-c_d(\mu_d - e_d) - c_f(\mu_f - e_f) - c_c(\mu_c - e_c)) \tag{A.5}$$

$$\frac{\partial c_S}{\partial t} = r_{SN}(-c_d(\mu_d - e_d)), \tag{A.6}$$

where c_d, c_f and c_c are the biomasses of diatoms, flagellate and cyanobacteria, respectively. Concentrations of nitrogen, phosphate and silicate are c_N, c_P and c_S. The constant ratios for cyanobacteria nutrient intake are r_{PN} and r_{SN}. Mortality rates are m_d, m_f and m_c. The specific rates of exudations are dependent exponentially on temperature by equation $e_{d,f,c} = e_{0_2}^{a\ T}$.

The phytoplankton growth rates $\mu_{d,f,c}$ depend on nutrient concentrations, irradiation and temperature:

$$\mu_{f_{max}} = \mu_{f0} a^{dT} \tag{A.7}$$

$$\mu_f = \mu_{f_{max}} \frac{I}{I + k_{fI}} \left(\frac{c_P}{c_P + k_{fP}} \frac{c_N}{c_N + k_{fN}} \right), \tag{A.8}$$

where $I = I(z)$ is the illumination, μ_{f0} is the maximum growth rate at 0 °C. $k_{fP,fN} = \mu_{f_{max}}/\alpha_{fN,fP}$ are the half saturation functions as used by Aksnes et al. (1995) with constant but species and limitation dependent affinities α. $I(z)$ depends on the amount of biomass between the depth z and the surface (shelf-shading).

References

Aksnes, D.L., Ulvestad, K.B., Balio, B.M., Berntsen, J., Egge, J.K., Svendsen, E., 1995. Ecological modelling in coastal waters: towards predictive physical–chemical–biological simulation models. Ophelia 41, 5–36.

Arrigo, K., 2005. Marine microorganisms and global nutrient cycles. Nature 437 (7057), 349–355.

Boyer, T., Antonov, J., Garcia, H., Johnson, D., Locarnini, R., Mishonov, A., Pitcher, M., Baranova, O., Smolyar, I., 2006. World Ocean Database 2005. NOAA Atlas NESDIS 60. U.S Government Printing Office, Washington D.C. version WOD 2001.

Brankovic, C., Palmer, T.N., 1997. Atmospheric seasonal predictability and estimates of ensemble size. Monthly Weather Review 125, 859–874.

Buizza, R., 1997. Potential forecast skill of ensemble prediction and spread and skill distributions of the ECMWF ensemble prediction system. Monthly Weather Review 125, 99–119.

Eker-Develi, E., Berthon, J.-F., van der Linde, D., 2008. Phytoplankton class determination by microscopic and HPLC-CHEMTAX analyses in the southern Baltic Sea. Marine Ecology Progress Series 359, 69–87.

Engström-Öst, J., Koski, M., Schmidt, K., Viitasalo, M., Jonasdottir, S.H., Kokkonen, M., Repka, S., Sivonen, K., 2002. Effects of toxic cyanobacteria on a plankton assemblage: community development during decay of Nodularia spumigena. Marine Ecology Progress Series 232, 1–14.

FIMR, SYKE, 2006. Moderate algal blooms are forecast for the western Gulf of Finland and southern Archipelago Sea stronger blue–green algal blooms are expected in the central and southern parts of the Baltic Sea. http://www.itameriportaali.fi/en/ajankohtaista/uutisia_muualta/2006/en_GB/1143.

Geider, R.J., MacIntyre, H.L., Kana, T.M., 1997. Dynamic model of phytoplankton growth and acclimation: responses of the balanced growth rate and the chlorophyll a: carbon ratio to light, nutrient-limitation and temperature. Marine Ecology Progress Series 148, 15–25.

Gill, A.E., 1982. Atmosphere–Ocean Dynamics. Vol. 30 of International Geophysics. Academic Press, San Diego, USA.

HELCOM, 2009. Activities 2008 — overview. No. 118 in Baltic Sea Environment Proceedings. Helsinki Comission (HELCOM) — Baltic Marine Environment Protection Commission.

Janssen, F., Neumann, T., Schmidt, M., 2004. Inter-annual variability in cyanobacterial blooms in the Baltic Sea controlled by wintertime hydrographic conditions. Marine Ecology Progress Series 275, 59–68.

Kahru, M., Leppänen, J.M., Rud, O., Savchuk, O.P., 2000. Cyanobacteria blooms in the Gulf of Finland triggered by saltwater inflow into the Baltic Sea. Marine Ecology Progress Series 207, 1318.

Kahru, M., Savchuck, O.P., Elmgren, R., 2007. Satellite measurements of cyanobacterial bloom frequency in the Baltic Sea: internannual and spatial variability. Marine Ecology Progress Series 343, 1523.

Kiiltomäki, A., 2008. Development of the Baltic Sea indicator report by Algline measurements and BalEco model results. Master's thesis, University of Helsinki.

Kutser, T., Metsamaa, L., Strömbeck, N., Vahtmäe, E., 2006. Monitoring cyanobacterial blooms by satellite remote sensing. Estuarine, Coastal and Shelf Science 67, 303–312.

Lermusiaux, P.F.J., 2006. Uncertainty estimation and prediction for interdisciplinary ocean dynamics. Journal of Computational Physics 217, 176–199.

Leutbecher, M., Palmer, T., 2008. Ensemble forecasting. Journal of Computational Physics 227, 3515–3539.

Lilover, M.-J., Stips, A., 2008. The variability of parameters controlling the cyanobacteria bloom biomass in the Baltic Sea. Journal of Marine Systems 74, S108–S115.

Marshall, J., Adcroft, A., Hill, C., Perelman, L., Heisey, C., 1997a. A finite-volume incompressible Navier Stokes model for studies of the ocean on parallel computers. Journal of Geophysical Research 102 (C3), 5753–5766.

Marshall, J., Hill, C., Perelman, L., Adcroft, A., 1997b. Hydrostatic, quasi-hydrostatic, and non-hydrostatic ocean modeling. Journal of Geophysical Research 102 (C3), 5733–5752.

Mazur-Marzec, H., Krel, A., Kobos, J., Pliski, M., 2006. Toxic nodularia spumigena blooms in the coastal waters of the Gulf of Gdansk: a ten-year survey. Oceanologia 48, 255–273.

Molteni, F., Buizza, R., Palmer, T., Petroliagis, T., 1996. The ECMWF ensemble prediction system: methodology and validation. Quarterly Journal Royal Meteorological Society 122, 73–119.

Nausch, M., Nausch, G., Wasmund, N., 2004. Phosphorus dynamics during the transition from nitrogen to phosphate limitation in the central Baltic Sea. Marine Ecology Progress Series 266, 15–25.

Nausch, M., Nausch, G., Lass, H.S., Mohrholz, V., Nagel, K., Siegel, H., Wasmund, N., 2009. Phosphorus input by upwelling in the eastern Gotland Basin (Baltic Sea) in summer and its effects on filamentous cyanobacteria. Estuarine, Coastal and Shelf Science 83, 434–442.

Niemi, Å., 1979. Blue–green algal blooms and N:P ratio in the Baltic Sea. Acta Botanica Fennica 110, 57–61.

Niemistö, L., Rinne, I., T., M., Niemi, Å., 1989. Blue–green algae and their nitrogen fixation in the Baltic Sea. Finnish Marine Research 17, 3–20.

Redfield, A.C., 1958. The biological control of chemical factors in the environment. American Scientist 205.

Richardson, D.S., 2000. Skill and relative value of the ensemble prediction system. Quarterly Journal Royal Meteorological Society 126, 649–667.

Ruokanen, L., Kaitala, S., Fleming, V., Maunula, P., 2003. Alg@line-joint operational unattended phytoplankton monitoring in the Baltic Sea. In: Dahlin, H., Flemming, N.C., Nxcvxc, K., Petersson, S. (Eds.), Building the European Capacity in Operational Oceanography, Proceedings of the Third International Conference on EuroGOOS. Elsevier Oceanography Series, vol. 69. Elsevier, pp. 519–522. http://dx.doi.org/10.1016/S0422-9894(03)80083-1.

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

Seppälä, J., Balode, M., 1999. Spatial distribution of phytoplankton in the Gulf of Riga during spring and summer stages. Journal of Marine Systems 23 (1), 51–67.

Stipa, T., Skogen, M., Sehested Hansen, I., Eriksen, A., Hense, I., Kiiltomäki, A., Søiland, H., Westerlund, A., 2003. Short-term effects of nutrient reductions in the North Sea and the Baltic Sea as seen by an ensemble of numerical models. Meri – Report Series of the Finnish Institute of Marine Research 49, 43–70.

Talagrand, O., Vautard, R., Strauss, B., 1999. Evaluation of probabilistic prediction systems. Proceedings of the ECMWF Workshop on Predictability, 20.–22. October 1997. ECMWF, Shienfield Park, Reading.

Tyrrell, T., 1999. The relative influences of nitrogen and phosphorus on ocean primary production. Nature 400, 525–531.

Walve, J., Larsson, U., 2007. Blooms of Baltic Sea aphanizomenon sp. (cyanobacteria) collapse after internal phosphorus depletion. Aquatic Microbial Ecology 49, 57–69.