TRANSNAV
http://www.transnav.eu
the International Journal
on Marine Navigation
and Safety of Sea Transportation
Volume 17
Number 3
September 2023
DOI: 10.12716/1001.17.03.04

Aggregating Sea Surface Hydrodynamic Forecasts From Multi-Models for European Seas

V. Frishfelds, J. She, J. Murawski & J.W. Nielsen
Danish Meteorological Institute, Copenhagen, Denmark

ABSTRACT: Maritime information services supporting European agencies such as the FRONTEX require European-wide forecast solutions. Following a consistent approach, regional and global forecasts of the sea surface conditions from Copernicus Marine Service and national met-ocean services are aggregated in space and time to provide a European-wide forecast service on a common grid for the assistance of Search and Rescue operations. The best regional oceanographic model solutions are selected in regional seas with seamless transition to the global products covering the Atlantic Ocean. The regional forecast models cover the Black Sea, Mediterranean Sea, Baltic Sea, North Sea and combine the North Sea – Baltic Sea at the Danish straits. Two global models have been added to cover the entire model domain, including the regional models. The aggregated product is required to have an update frequency of 4 times a day and a forecasting range of 7 days, which most of the regional models do not provide. Therefore, smooth transition in time, from the shorter time-range, regional forecast models to the global model with longer forecast range are applied. The set of parameter required for Search and Rescue operations include sea surface temperature and currents, waves and winds. The current version of the aggregation method was developed for surface temperature and surface currents but it will be extended to waves in latter stages. The method relies on the calculation of aggregation weights for individual models. For sea surface temperature (SST), near real-time satellite data at clear-sky locations for the past days is used to determine the aggregation weights of individual forecast models.
A more complicated method is to use a weighted multi-model ensemble (MME) approach based on best forecast features of individual models and possibly including near real time observations. The developed method explores how satellite observations can be used to assess spatially varying, near real time weights of different forecasts. The results showed that, although a MME based on multiple forecasts only may improve the forecast, if the forecasts are unbiased, it is essential to use observations in the MME approach so that proper weights from different models can be calculated and forecast bias can be corrected. It is also noted that, in some months, e.g., June in Baltic Sea, even SST was assimilated, the forecast still show quite high error. There are also visible difference between different Copernicus Marine Environment Monitoring Service (CMEMS) satellite products, e.g. OSTIA and regional SST products, which can lead different forecast quality if different SST observation products are assimilated.

1 INTRODUCTION

Safety at sea is depends greatly on the sea surface conditions. Oceanographic models are utilized to produce forecasts of sea surface information, enabling the prediction of these conditions. European agencies, such as FRONTEX, require access to met-ocean information for pan-European seas, including the

Baltic Sea, North Sea, Mediterranean Sea, Black Sea, Norwegian Sea and parts of Atlantic and Arctic oceans (see Figure 1), for coordinated maritime activities such as search and rescue operations and planning. For the FRONTEX search and rescue service, the forecast period is required to be seven days long and updated four times per day. Currently, none of the community or national forecast services is able to meet this requirement. Global forecasts produced by organizations like Copernicus Marine Environment Monitoring Service (CMEMS), UK Met Office and CMCC (Euro-Mediterranean Centre for Climate Change), cover the pan-European seas, but they are only updated once a day, and their forecast skills have not been specifically tuned for the European seas. Regional oceanographic forecasts, both from national and CMEMS forecast services, have higher resolution and update frequency but they typically cover only part of the European seas, and their forecast ranges sometimes are less than seven days. Therefore it is necessary to combine different oceanographic forecasts to generate a European wide, quality ensured model product that meets all the user needs. All the important hydrodynamic quantities for the operations, such as sea surface temperature and currents, should have seamless transition between the different seas and oceanographic areas.

There are two ways to make this seamless European Sea forecast. One is using static weights that ensure a smooth transition between regional products, the other is using dynamic weights to aggregate products from different models to a Multi-Model-Ensemble (MME) product. The first method uses smooth, static spatial weighting functions for the individual forecasts to aggregate them smoothly at the boundaries of the forecasting area. Since some forecasts have ranges shorter than seven days, a temporal smoothing should also be performed. To generate seamless transition in space and time, only one forecast is needed for a given region. Hence for each region and update time, a “best forecast” should be selected. This approach is referred as simple aggregation.

A more complicated method is to use a weighted multi-model ensemble (MME) approach based on best forecast features of individual models and possibly including near real time observations. Such kind of method has been developed in the atmospheric science [1, 2, 3] showing improved forecasts [4]. For ocean forecasting, the MME approach has been used to generate sea level forecast at tide gauge stations in the Baltic-North Sea. For ocean field forecast, the MME forecast has applied same weights for different models, instead of using observations to determine models’ weights [5].

In this paper, both approaches will be implemented. The simple aggregation method has been set-up to perform operational forecasts. The MME aggregation was developed to replace it, and is currently available for SST forecasts in the Baltic-North Seas as well as in Mediterranean Sea and Black Sea, where it has been implemented, tested and validated. The developed method explores how satellite observations can be used to assess spatially varying, near real time weights of different forecasts.

The paper is organized as follows: section 2 describes methodology and input data, section 3 analyses results and section 4 is conclusions.

2 METHOD AND INPUT DATA

This study focus on examining feasibility of using Copernicus Marine Environment Monitoring Service (CMEMS) and national forecasts, both global and regional, for making a pan-European Sea aggregated forecast. The aggregation forecast experiments are made and assessed for two periods: a 7-month period during September 3, 2022 – March 15, 2023 including all forecasts with focusing on the general performance of the method, and a one year period from May 1, 2021 – April 30, 2022, with focusing on seasonal variability. However, only CMEMS and DMI forecasts are used in this period as CMCC global product is not available for this period.

2.1 *Input data*

Both multiple forecast products of SST and sea surface currents and satellite SST observations are used in this study.

2.1.1 *Forecast products*

The forecast products used in this study consist of global forecast from CMEMS and CMCC, and regional forecast from CMEMS and DMI. A summary of the products is shown in Table 1.

DMI product DKSS uses sub-domains of higher resolution in the Danish straits.

There are a few things to note about the different forecast products that are used in the aggregation.

Data assimilation: All forecasts in Table 1 use SST data assimilation except for DMI forecast. Therefore, all the forecast products except DMI’s may feature smaller temperature bias. It should also be noted that global and regional model systems assimilating different SST products, for example, CMEMS global forecast assimilates OSTIA L4 SST while most of CMEMS regional forecast system assimilates regional L3S SST. The differences in the OSTIA and regional SST data will affect the weights of different forecasts in the MME forecast aggregation.

Currents: all regional forecasts have tides included in the model, except for the two global forecast products (CMEMS, CMCC). In order to obtain the total currents from the global oceanographic products, separate forecasts of tidal currents are obtained from CMEMS global product.

Some areas not covered by the CMCC (Euro-Mediterranean Centre for Climate Change) global product like inland parts of fjords are excluded also from CMEMS global forecast. CMEMS Mediterranean, CMEMS Black Sea and CMCC global forecasts do not cover Marmara Sea, Azov Sea, the narrow Dardanelles Strait at the eastern part but they are still covered by CMEMS global forecast.

Table 1. Individual forecast products used for aggregation forecast (SST and surface currents only)

Area	Provider	Spatial resolution	Temporal resolution	Update time (h)	Forecast range
48-66N, 4W-30E	DMI (Baltic-North Sea)	0.05 deg.	Hourly	00, 06, 12, 18	5days
11-73N, 43W-43E	CMEMS (Global)	0.083 deg.	Hourly	12	10days
46-62.75N, 16W-13E	CMEMS (NW Shelf) [6]	0.03 deg.	Hourly	12	6 days
26-56N, 19W-5E	CMEMS (Bay of Iberian-Biscay-Ireland)	0.028 deg.	Hourly	12	5 days
53-66N, 9-30E	CMEMS (Baltic Sea)	0.018 deg.	Hourly	00, 12	6 days
30.18-45.98N, 17.29W-36.30E	CMEMS (Mediterranean Sea) [7]	0.042 deg.	Hourly	12	5days
40.5-47.0N, 27.25-41.1E	CMEMS (Black Sea) [8]	0.025 deg.	Hourly	12	5days
11-73N, 43W-43E	CMCC (Global)	0.0625 deg.	Hourly	12	6days

Eastern part of Atlantic Ocean is covered by CMEMS regional Iberian-Biscay-Ireland (IBI) model. It is not yet implemented in the operational aggregated forecast product, but it is included in SST validation results in order to estimate whether to include it in updated operational aggregation.

2.1.2 *Satellite SST*

Regional SST satellite observations of Level 3 from CMEMS are preferred for forecast aggregation in regional seas, i.e., Baltic Sea, North Sea, Mediterranean Sea and Black Sea. This includes North Sea/Baltic Sea - Sea Surface Temperature Analysis L3S product, Mediterranean Sea - High Resolution and Ultra High Resolution L3S Sea Surface Temperature product [9], Black Sea - High Resolution and Ultra High Resolution L3S Sea Surface Temperature product [9], and CMEMS global Level 3 product. In the areas where Level 3 SST is not available in CMEMS, e.g., in the Arctic Ocean and Atlantic open sea, Level 4 SST from CMEMS SST Thematic Assembly Center (TAC) is used. As a consequence, CMEMS Arctic Level 4 [10] product is well suited for the Arctic area.

Modelled SST as well as surface currents depends on the thickness of upper layer used in the model. In general, regional models have higher vertical resolution and thus SST could be resolved better. Observed SST of weather satellites often relates to near surface part of the sea surface and is typically at lower depth than the first layer depth of a forecast model. Some of CMEMS SST Level 3 observation products have both SST and adjusted SST as Mediterranean Sea, Black Sea and global ocean products. Adjusted SST observations have a lower influence from daily SST variations and are used for validation of the individual models here. SST satellite observations are often stated as daily average value but depend on exact timing of pathing satellites in their orbits. Therefore, SST observations come with its natural deviations which are typically lower than 1 degree of temperature.

Satellite Sea surface temperature products are used for the calculation of dynamic and static MME aggregation weights. Sea surface temperature is one of the best products that satellite monitoring can deliver for the oceanographic conditions. Existing regional SST satellite observations are better parametrized for corresponding seas: North Sea, Baltic Sea, Mediterranean Sea, Black Sea and Arctic Ocean. Therefore, global SST observations are used only in areas outside the regional ones. CMEMS SST observations of Level 3 type are used everywhere except the Arctic where Level 4 satellite observations are used. Level 4 observations in Arctic Ocean are filtered in such a way that only observations with standard deviation between 0.001 K and 0.2 K are selected, that approximately corresponds to valid Level 3 observations. Observations with existing ice mask are excluded as these SST observations may be less accurate in this situation than the modelled result.

2.2 *Simple aggregation method*

To obtain an aggregated forecast, the simple aggregation method is used to smoothly merge multiple forecasts with different spatiotemporal resolutions and coverages into a unified grid, coverage and forecast range. This aggregation involves several stages.

First, spatial and temporal grid is constructed to cover the forecast area (11° N - 73° N, 43° W - 43° E) with the requested resolution of 0.1 degree, and a 7-day forecast period with an hourly interval. That grid acts as the basis for the aggregation method, and the aggregated product with the specified grid is delivered to FRONTEX users.

Next, a common land-sea mask is constructed on the base grid to ensure the same representation for all forecast products. Disconnected small water bodies in the base grid are removed, and most of inland fjords are also removed as they are represented only in one or two forecast products. The CMEMS global product, with a spatial resolution of 0.083 degree and the best spatial and time coverage, is used as the basis for the land-sea mask in order to ensure that all grid points are represented for the entire forecast period.

Then, a weighting function is constructed for each forecast source on the base grid to ensure smooth transition in space and time. Space and time variables are separated in the weighting function for each forecast source *i*:

$$w_i\left(lat,lon,t\right)=g_i\left(lat,lon\right)\cdot h_i\left(t\right) \tag{1}$$

where *lat* is latitude and *lon* is longitude, and *t* is time. The spatial function $g_i(lat,lon)$ has a transition zone around the boarders where regional and global forecasts overlap. This ensures continuous spatial transition from a regional solution to a global one. A scaling coefficient ai is added to put a higher weight on a regional forecast source:

The next step is to perform a linear interpolation of source forecast fields from their model grids to the base grid. Each variable (SST and components of surface currents) of each model source results in separate interpolated field. However, linear interpolation does not work at coastal locations, so a nearest neighbour method is used, and if it fails to get a value, then a moving average is selected with a window size not exceeding 0.3 degrees. After that, the common land-sea mask is applied to cut off unnecessary points, and NAN (not any number) values are used outside the area covered by a source forecast.

Then, a weighting function is constructed for each source on the base grid to ensure smooth transition in space and time. Space and time variables are separated in the weighting function for each forecast source i:

$$w_i(lat, lon, t) = g_i(lat, lon) \cdot h_i(t) \tag{2}$$

where *lat* is latitude and *lon* is longitude, *t* is time. Weighting function is not normalized at this stage. The final weights are derived when all the sources are considered. The spatial function $g_i(lat,lon)$ has a buffer zone in open waters for the regional forecast sources to ensure continuous spatial transition from a regional solution to the global one. A constant scaling coefficient a_i is added to put a higher weight on a regional source:

$$g_i(lat, lon) = a_i \cdot \eta_i(lat, lon) \tag{3}$$

where weighting function $\eta_i(lat,lon)$ of forecast i changes from 0 to 1, see Figure 1. If it is 0, the source is disregarded at the given location, and if it is 1, then there is maximal effect of the corresponding source on the aggregated product at the given location. Coefficients a_i are chosen according to validation results which are typically better for regional forecast sources with higher resolution.

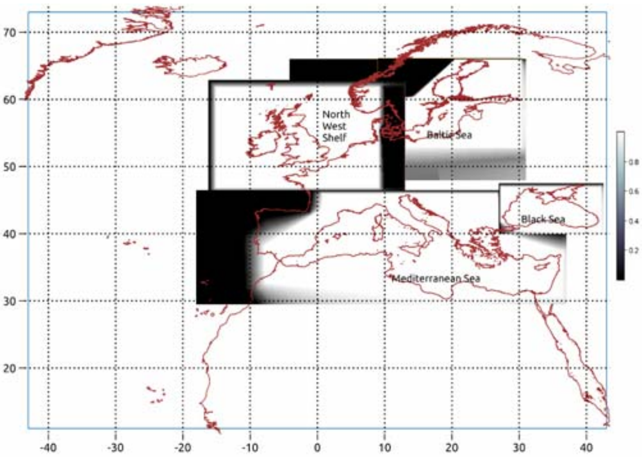

Figure 1. Blue rectangle - area of aggregation. Shaded rectangles: unit weighting function $\eta_i(lat,lon)$ for regional model areas (Northwest Shelf Sea, Baltic Sea, Mediterranean Sea, Black Sea)

Most of the sources have a forecast period of less than seven days. Therefore, time function $h_i(t)$ is constructed, which is 1 when the corresponding hour is covered in the source forecast and 0 when it is not. However, this would result in step-like jump in time. Hence, function $h_i(t)$ is chosen in such a way that it continuously changes from 1 to 0 at the final day of the period covered by forecast source i.

If a specific forecast source may not be available, it is disregarded by setting its weighting coefficient a_i to zero. Finally, a value $V(lat,lon,t)$ in aggregated forecast product is obtained from combined values $v_i(lat,lon,t)$ of weighted sources on the base grid:

$$V(lat, lon, t) = \frac{\sum_i v_i(lat, lon, t) w_i(lat, lon, t)}{\sum_i w_i(lat, lon, t)} g_i(lat, lon) \cdot h_i(t) \tag{4}$$

where summation occurs over sources with non-NAN values. The V value can represent any scalar value, e.g., SST or components of the surface currents.

2.3 *Multi-Model-Ensemble (MME) aggregation method*

SST forecasts are typically accompanied with daily satellite SST observations from the preceding days, allowing for the minimization of forecast bias. Furthermore, multiple forecasts are available for a given region, such as global and regional forecasts from CMEMS and national services. This presents the potential to improve SST forecast beyond the simple aggregation method by using a MME aggregation. Satellite SST observations enable to the determination of optimal weights for individual forecasts in the aggregated product. In such a case, one would often expect that regional oceanographic forecasts will outperform the global ones. As a result, higher weights are assigned to the regional products. The MME aggregation will not be applied to locations with ice masks.

For the MME aggregation, the first step is to obtain the forecast error statistics. Then the forecast error is used to calculate the weights for different forecast products and finally the MME aggregated forecast is obtained as a linear, weighted sum of the individual forecasts. SST deviation at location-time (x,t) of a forecast i is:

$$\Delta T_i(x, t) = T_i(x, t) - T_0(x, t) \tag{5}$$

where $T_0(x,t)$ is satellite SST at each grid point $x=(lat, lon)$ and time moment t. $T_i(x,t)$ is modelled SST value of the forecast model i.

In the same way, squared temperature difference is treated. Mean square difference at location-time (x,t) of forecast i is:

$$\Delta T^2{}_i(x, t) = \left(T_i(x + \Delta x, t + \Delta t) - T_0(x + \Delta x, t + \Delta t)\right)^2 \tag{6}$$

In order to have spatially and temporally smooth squared error distribution, rolling mean value is used at each grid point x and time t:

$$\Delta T^2{}_{i,rol}(x, t) = \frac{1}{m} \sum_{\Delta x, \Delta t}^{m} \left(T_i(x + \Delta x, t + \Delta t) - T_0(x + \Delta x, t + \Delta t)\right)^2 \tag{7}$$

where, in practice, Δx runs through 5 × 5 spatial grid points of latitude and longitude centred on grid point x; Δt runs through 7 days: -3 days, -2 days, …., 3 days centred on current day t; $m = 5 \times 5 \times 7 = 175$. Points with no data values at location-time $x+\Delta x, t+\Delta t$ are excluded from averaging.

Using multi model ensemble, local weighting factor for each source i is obtained by:

$$w_{i,rol}(x,t) = \frac{\left(\Delta T^2_{i,rol}(x,t)\right)^{-4}}{\sum_j \left(\Delta T^2_{j,rol}(x,t)\right)^{-4}} \tag{8}$$

The 4th power is used here, in order to have higher weight for the best product. For the same reason, bias is not subtracted at this stage. The resulting deviation at *(x,t)* of multi-model ensemble is

$$\Delta T(x,t) = \sum_i w_{i,rol}(x,t) \cdot \Delta T_i(x,t) \tag{9}$$

where summation occurs over all forecast sources i in the area. The resulting centred Root Mean Square Error (cRMSE) at *(x,t)* of multi-model ensemble is

$$cRMSE(x,t) = \sqrt{\sum_i \left(w_{i,rol}(x,t) \cdot \Delta T^2_i(x,t)\right) - \Delta T^2(x,t)} \tag{10}$$

These formulas yield quantitative estimations of quality of each forecast source and resulting effect on the aggregated product.

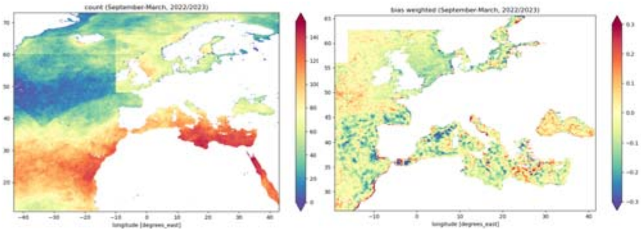

Figure 2. Left: Count of Level 3 observations from regional and global CMEMS sources at each location on 0.1×0.1 grid in September 3, 2022 to March 15, 2023. Right: bias of weighted product (9) of 3 forecast models: CMEMS global, CMCC global and CMEMS regional.

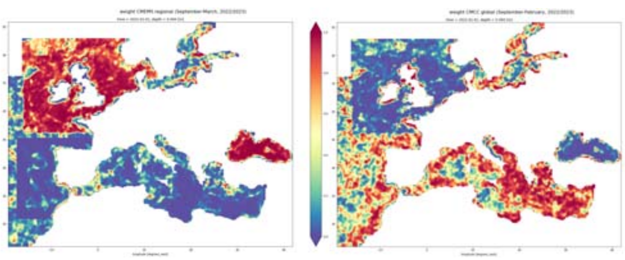

Figure 3. Left: weight of CMEMS regional product (Black Sea, Mediterranean Sea, Baltic Sea, North West Shelf, Iberia Biscay Ireland) with respect to the set of 3 forecast models: CMEMS global, CMCC global and CMEMS regional. Right: weight of CMCC global product in the same configuration

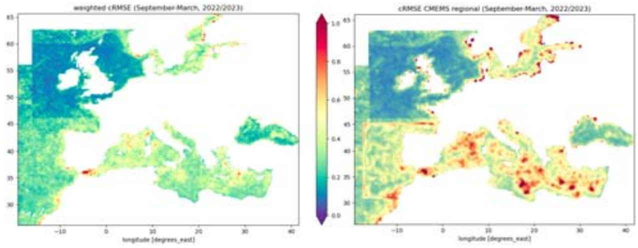

Figure 4. Left: central RMSE of combined set of 3 forecast models: CMEMS global, CMCC global and CMEMS regional (Black Sea, Mediterranean Sea, Baltic Sea, North West Shelf, Iberia-Biscay-Ireland). Right: central RMSE of only CMEMS regional product in respective areas

3 RESULTS

Results of the Multi-Model-Ensemble (MME) aggregation method of have been analysed for 2 periods covering different data products. The first experiment 09/22-03/23 includes all regional and global data sets, but does not cover an entire year. The second experiment 2021-2022 includes all regional, but only one global data set, the Copernicus Marine Environment Monitoring Service (CMEMS) global product and covers one full year. It can therefore be used for the assessment of the seasonal variations of the MME aggregation method. Finally, the operational implementation of the simple aggregation method using static weights is discussed as well.

3.1 *Experiment during 09/22-03/23*

The first experiment studies the spatial pattern of the MME aggregation on European scale. It covers a limited time frame of 7 months (September 2022 to March 2023), because CMCC (Euro-Mediterranean Centre for Climate Change) global data was not available for the period prior to 09/23. Figure 2 shows the number of observations in the considered period September 2022 to March 2023. The highest number of clear-sky days with valid SST observations occurs in Mediterranean Sea, Red Sea, at latitudes of around 25oN. CMEMS regional SST observations in North Sea, Baltic Sea and Arctic Ocean provide higher number of days with valid observations. Lowest number of observations occurs in Atlantics at latitude of 50 degrees due to the specific positioning of the satellite orbits. Also, Arctic coastline has lower number of observations due to ice mask in winter time.

Figures 3-4 show derived weight of CMEMS regional or CMCC global forecasts (Figure 3) and central Root Mean Square Error (cRMSE, Figure 4) when 3 forecasts are involved: CMEMS global, CMCC global and CMEMS regional. CMEMS regional forecast model corresponds either to Black Sea, Mediterranean Sea, Baltic Sea, North West Shelf (NWS) or Iberia-Biscay-Ireland model. Iberia-Biscay-Ireland model is not yet included in operational aggregation but included in validation results. Bias of weighted product (9) is well within half a degree range, see Figure 2 right. That bias is easily removable from the final product.

CMEMS regional models clearly dominate the weights in the aggregation in the North West Shelf and Black Sea, see Figure 3. Actually, some of the models like North West Shelf (NWS) model benefit from the fact that it uses best-estimate products for the days prior to the analysis that incorporate also sea surface temperature (SST) observations. These data are then stored in historical CMEMS database that are used for validation. For this reason, the real operational NWS forecast product from CMEMS is not as accurate as the historical NWS dataset archived at CMEMS. Henceforth, the calculated MME weights that use historical data do not necessary reflect the quality of the forecast data (Figure 3).

Regional Baltic Sea model has good performance in SST, but may slightly over-predict upwelling events which are rather tricky to model in Baltic Sea, see Figure 5. Also, locations with upwelling events may promote creation of low altitude clouds meaning that usable satellite SST observations may not be available. Mediterranean and Iberia-Biscay-Ireland regional models show approximately the same quality with respect to SST observations. Mediterranean and Black Sea models use 3DVAR scheme of OceanVar [11] that assimilates data predominantly in weekly basis. Different assimilation scheme can lead to that SST performance sometimes falls behind CMCC global oceanographic model as in the given time period of 7 months.

Figure 4 shows that cRMSE of the weighted SST product of the mentioned 2 global models and 1 regional model in the respective sea. The result is better in North West Shelf and Black Sea where the weight of regional product dominate. The first one may benefit from hindcasting nature of North West Shelf archived data in CMEMS database. Both of the global modelling products are less effective in Black Sea that leads to dominance of the regional Black Sea product. Some larger deviations occur at Gibraltar. Also the Mediterranean and West African coastlines have a lower combined accuracy that may result from ability of Nemo oceanographic model to handle shallower waters. There is strong gain of using MME in Mediterranean Sea and Baltic Sea rather than using the CMEMS regional product alone, compare Figure 4 left and right.

The quality of the aggregated product deteriorates as the forecast becomes longer. It is because there is lower number of models for the final days and the performance of the models decreases with increasing forecast length in general. Validation results in [12] suggest that relatively good results are obtained for the first 3-4 days of the forecast. The final days of the forecast can be used only as initial estimate of what conditions are expected without a remarkable accuracy. Longer forecasts of SST are generally more accurate than that of the currents, because temperature is essentially a cumulative quantity of resulting from previous conditions.

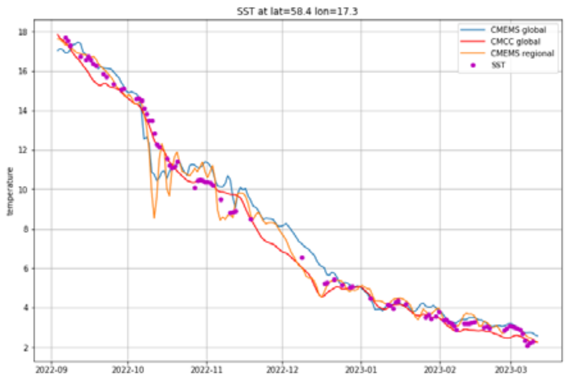

Figure 5. Example of SST time series of 3 different forecast models and observations in Baltic Sea near Swedish coast at latitude=58.4° and longitude=17.3°

3.2 *Experiment in 2021-2022 (Seasonal validation of SST)*

Sea surface temperature features a strong seasonality at northern latitudes. In winter, the SST is close to zero degrees with very small day to day changes as well as small diurnal variations, whereas in summer, SST features strong day to day variations with distinct diurnal cycle. Figure 6 shows monthly cRMSE of the aggregated MME forecast. It is noted that there are exceptionally high forecast errors in June and July, which is originated from all the individual forecast (Figure 7). Consequently, optimal weight of each source model in the aggregation could depend on the season. For example, monthly weight of CMEMS Baltic Sea model is dominant in spring-summer but less dominant in autumn, see Figures 6-7. The main reason why CMEMS regional Baltic Sea model has less accuracy in autumn months is a slight over-prediction of upwelling events resulting in a high penalty. Also in Mediterranean Sea, the CMEMS regional product is better in summer and spring, but it provides less accuracy in autumn, see Figure 7.

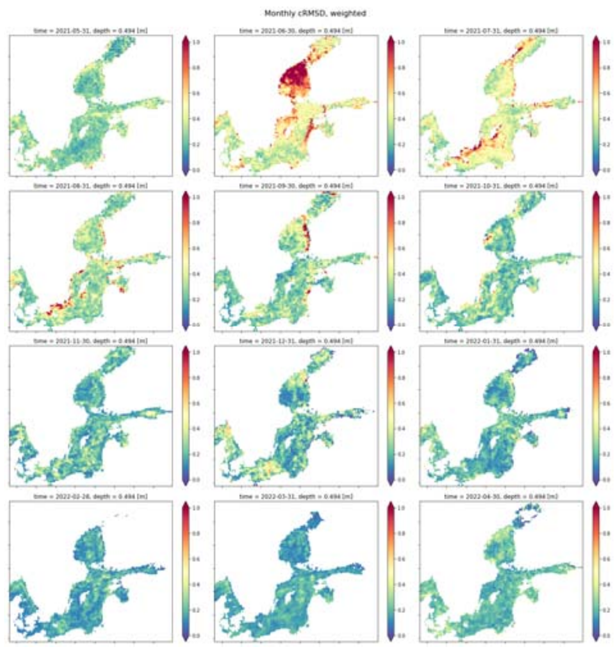

Figure 6. Monthly weighted cRMSE in Baltic Sea in period from May 2021 to April 2022. The included forecast models are CMEMS global, DMI DKSS and CMEMS Baltic Sea model.

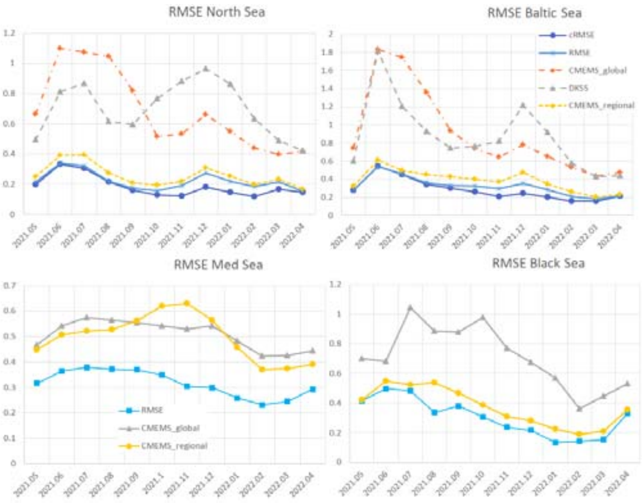

Figure 7. Monthly averaged RMSE of SST in regional seas of CMEMS regional (North Sea, Baltic Sea, Mediterranean Sea, Black Sea) model, DMI DKSS model for North Sea-Baltic Sea (upper row) and CMEMS global ocean model and the weighted model. RMSE and cRMSE represent weighted values of aggregated product.

Figure 7 shows monthly RMSE in regional seas. In general, it is less than half a degree. It is slightly less in winter when typical variations of temperature are lower. The CMEMS regional products clearly dominate here. Use of MME still improves the final result, especially in Mediterranean Sea. CMEMS regional product in Mediterranean Sea has larger deviations in months of autumn. Figure 7 also shows that the aggregated MME product features the lowest RMSE values in all regions.

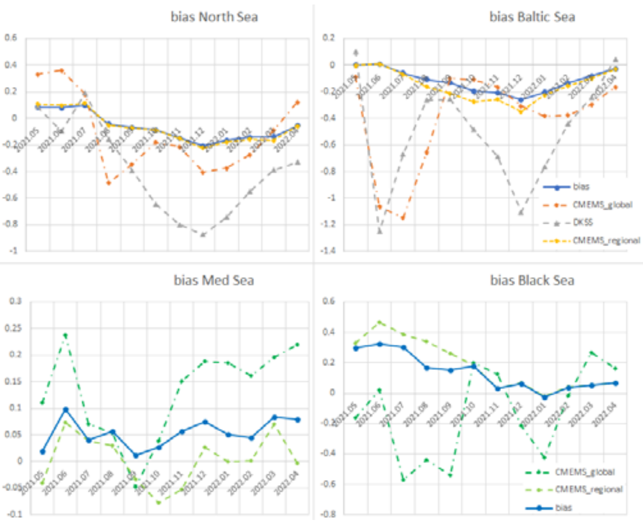

Figure 8. Seasonal bias (in degrees) of SST regional seas: North West Shelf, Baltic Sea, Mediterranean Sea and Black Sea. The included models are CMEMS global model, CMEMS regional model in respective area and DMI DKSS model in North Sea – Baltic Sea. The blue line is bias of the weighted product (9).

Figure 8 shows seasonal bias in regional seas. Negative bias is typical in North West Shelf and Baltic Sea. The CMEMS regional products clearly dominate here, too, as they extensively use SST assimilation scheme. They have most negative bias in winter when the number of clear-sky days is lowest and there is a smaller amount of satellite SST data available for assimilation. The DKSS model is not using SST data assimilation. For this reason, the SST bias of the DKSS model is higher. Thus, the bias of the aggregated product is very close to CMEMS regional product, which is having a larger weight. CMEMS regional product in Mediterranean Sea has monthly bias with amplitude less than 0.1 degree that results from the fact that a larger amount of satellite SST data is available for data assimilation (Fig. 2). CMEMS regional product in Black Sea uses similar model parameterisation as in Mediterranean Sea but has pronounced positive bias in summer.

3.3 *SST bias correction in aggregated forecasts*

The validation results showed that resulting bias of SST taking into account several models is usually less than a degree in regional seas. The weighted bias of aggregated product shown in Figure 2 right and Figure 8 can be subtracted from the final result. It means that SST observations can be used to correct the biases of the model data. However, the near real time SST observations and model forecasts do not overlap in time. Therefore, aggregated forecast data are archived for the past 3 days which are then compared to SST observations of the same days. This yields an initial estimation of the location dependent bias correction function $\Delta_o(x)$ for the start of the forecast. Because, it is unclear how the bias could change on the forecast then the amplitude of initial bias is gradually set to fade for the final days of the forecast. The time dependent fading factor is set to have an exponential decrease

$$\Delta(x,t) = \Delta_o(x)e^{-t/\tau} \tag{11}$$

where τ is 3 days. Bias correction is set to zero in locations masked as cloudy or with an ice mask. Moving average method is used to obtain a smooth bias correction function $\Delta_o(x)$ with spatial window of 0.5 degrees.

Regarding optimal weight maps for the operational forecasts, we cannot use a detailed time-dependent weight maps as in Figure 4 because there are no SST observations in forecast period. Instead, optimal set of weights of individual model sources are derived from validation results of historical data.

3.4 *Operational implementation*

The core procedure of aggregation is carried out in a Python script using Xarray module to work with gridded NetCDF or Grib data. The script has been running operationally four times a day since November 2021 at Danish Meteorological Institute (DMI). The performance of the operational production is monitored in real time using an automatic monitoring tool. If there are errors in downloading and aggregation, a warning will be sent to assigned forecasters with an error report. An example of aggregated SST is shown in Figure 9. As can be seen, seamless spatial transition from regional solutions to global ones is well represented. Also transitions in time are smooth.

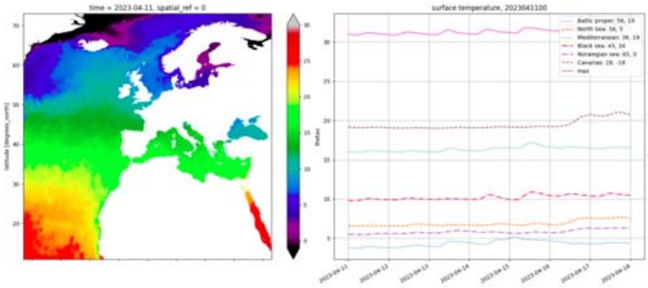

Figure 9. Example of operationally aggregated SST at 2023.04.11 00:00:00 using MME. Left is SST at the start of the forecast, right is SST development in time at selected locations shown by latitude, longitude values in legend.

4 CONCLUSIONS AND DISCUSSIONS

In this study, aggregating met-ocean forecast for pan-European seas is implemented and validated by integrating CMEMS and national global and regional forecasts, using both a static-weight deterministic method and a MME-based dynamic weight method. In the static-weight method, continuous transition in space and time between different models is made using spatial weighting functions with gradual transitions in space in order to produce a four times a day, pan-European sea 7-day forecast. In order to test the performance of aggregated forecast, validation of SST is used. In the MME-based method, the SST error statistics are used to generate optimal weights of each individual forecast model. Based on one year verification results, regional oceanographic models outperform the global ones as they have a higher native resolution. It should be noted that some of the regional models use a strong hindcasting nature of archived data in CMEMS database, e.g., CMEMS North West Shelf model that leads to very good performance for historical dataset but not necessary for the forecast. CMEMS Baltic Sea model uses SST assimilation, but slightly overestimates upwelling events near coastline in autumn. Similarly, CMEMS Mediterranean model has a good performance with respect to SST through almost all of the year except autumn, when CMCC global oceanographic model provides better validation results. It may result from assimilation scheme in CMEMS Mediterranean model which is running on a weekly basis rather than daily one. In order to obtain smooth weighting function, a moving average error analysis is made with small special window. That yields a weight map showing the strong and weak areas for each of the forecast model. That is used to estimate the weights in the operational version. MME approach yields notable benefits over simple aggregation method. It is especially notable in areas where a single model is not dominating as in Mediterranean Sea. The same principles of aggregation will be used also for aggregation of other fields as surface currents and waves.

It was found that, in the Baltic-North Sea, individual models have very high errors in June and July even if SST has been assimilated. The situation can be improved by using MME based aggregation but still give high errors.

The aggregation method of forecasts should be robust and work even if some of the forecast models are missing at the time of aggregation. Therefore, multi-model ensemble is essential to replace a missing model with the other ones. Moreover, some forecast models may be outdated at the time of aggregation. These situations are handled by proper selection of dynamic weights of the individual models in the aggregation method.

Major differences of these models with respect to observations occur in autumn and winter when the skies are cloudier and the amount of data and data quality are both low. Moreover, upwelling events have generally lower forecast accuracy and are more characteristic in autumn. Therefore, it is expected that there could be major deviations of modelled SST in autumn and winter months.

ACKNOWLEDGEMENT

This study is supported by FRONTEX project “Frontex/OP/234/2021/RS Meteorological and Oceanic Services Lot 2 - Oceanographic Data and Visualisation”. The authors acknowledge to Rita Lecci and CMCC colleagues for providing CMCC global forecast product.

REFERENCES

[1] T. N. Krishnamurti, C. M. Kishtawal, Z. Zhang, T. LaRow, D. Bachiochi, E. Williford, S. Gadgil, and S. Surendran, “Multimodel Ensemble Forecasts for Weather and Seasonal Climate,” J. Climate, vol. 13, pp. 4196–4216, 2000, https://doi.org/10.1175/1520-0442

[2] F. J. Doblas-Reyes, R. Hagedorn, T. N. Palmer, “The rationale behind the success of multi-model ensembles in seasonal forecasting—II. Calibration and combination,” Tellus Ser. A-Dyn. Meteorol. Oceanol., vol. 57, pp. 234–252, 2005

[3] A. P. Weigel, M. A. Liniger, C. Appenzeller, “Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts?” Q. J. R. Meteorol. Soc., vol. 134, pp. 241–260, 2008

[4] T. N. Krishnamurti, V. Kumar, A. Simon, A. Bhardwaj, T. Ghosh, R. Ross, “A review of multimodel superensemble forecasting for weather, seasonal climate, and hurricanes,” Rev. Geophys., vol. 54, pp. 336–377, 2016.

[5] I. Golbeck, X. Li, F. Janssen, et al, “Uncertainty estimation for operational ocean forecast products—a multi-model ensemble for the North Sea and the Baltic Sea,” Ocean Dynamics, vol. 65, pp. 1603–1631, 2015, https://doi.org/10.1007/s10236-015-0897-8

[6] D. Bruciaferri, M. Tonani, H. Lewis, J. Siddorn, A. Saulter, J. M. Castillo, N. Garcia Valiente, D. Conley, P. Sykes, I. Ascione, N. McConnell, “The impact of ocean-wave coupling on the upper ocean circulation during storm events,” Journal of Geophysical Research, Oceans, vol. 126, issue 6, 2021 https://doi.org/10.1029/2021JC017343

[7] E. Clementi, A. Aydogdu, A. C. Goglio, J. Pistoia, R. Escudier, M. Drudi, A. Grandi, A. Mariani, V. Lyubartsev, R. Lecci, S. Cretí, G. Coppini, S. Masina, and N. Pinardi, “Mediterranean Sea Physical Analysis and Forecast (CMEMS MED-Currents, EAS6 system) (Version 1),” Data set. Copernicus Monitoring Environment Marine Service (CMEMS), 2021

[8] E. Jansen, D. Martins, L. Stefanizzi, S. A. Ciliberti, M. Gunduz, M. Ilicak, R. Lecci, S. Cretí, S. Causio, A. Aydoğdu, L. Lima, F. Palermo, E. L. Peneva, G. Coppini,

S. Masina, N. Pinardi, A. Palazov, and N. Valchev, "Black Sea Physical Analysis and Forecast (Copernicus Marine Service BS-Currents, EAS5 system) (Version 1)," Data set. Copernicus Monitoring Environment Marine Service (CMEMS), 2022

[9] B. Buongiorno Nardelli, C. Tronconi, A. Pisano, R.Santoleri, "High and Ultra-High resolution processing of satellite Sea Surface Temperature data over Southern European Seas in the framework of MyOcean project", Rem. Sens. Env., vol. 129, pp. 1-16, 2013, doi:10.1016/j.rse.2012.10.012.

[10] J. L. Høyer, P. Le Borgne, and S. Eastwood, "A bias correction method for Arctic satellite sea surface temperature observations," Remote Sensing of Environment, 2014, https://doi.org/10.1016/j.rse.2013.04.020.

[11] A. Storto, P. Oddo, A. Cipollone, I. Mirouze, B. Lemieux-Dudon, "Extending an oceanographic variational scheme to allow for affordable hybrid and four-dimensional data assimilation," Ocean Modelling, vol. 128, pp. 67-86, 2018

[12] J. Murawski, J. She, C. Mohn, V. Frishfelds, J. W. Nielsen, "Ocean Circulation Model Applications for the Estuary-Coastal-Open Sea Continuum," Frontiers in Marine Science, vol. 8, 2021