

FORCOAST



Earth Observation Services For Wild Fisheries, Oystergrounds
Restoration And Bivalve Mariculture Along European Coasts

PROJECT DELIVERABLE REPORT

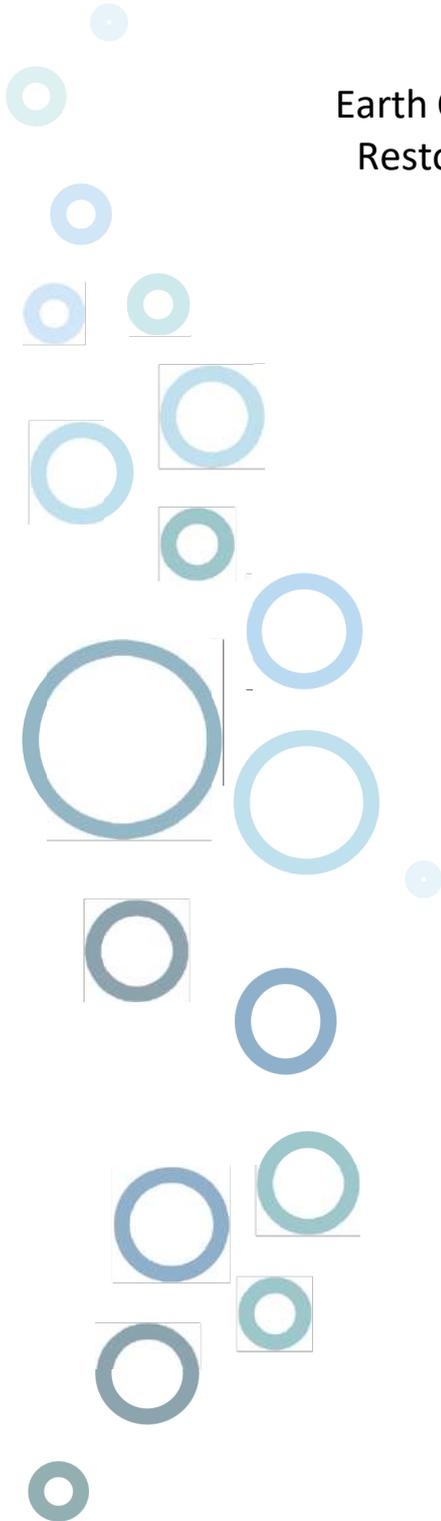
Deliverable Number: 5.4

Deliverable Title: Final Coordinated Pilot Model
Evaluation Report

Author(s): D. Pereiro, T. Dabrowski, L. Barbut,
A. Caballero, F. Campuzano, M. Maar, J. Murawski,
S. Querin, J. She, J. Staneva, L. Vandenbulcke

Work Package Number: 5

Work Package Title: Service Operationalization,
Demonstration and Validation



FORCOAST Project Information	
Project full title	Earth Observation Services For Wild Fisheries, Oystergrounds Restoration And Bivalve Mariculture Along European Coasts
Project acronym	FORCOAST
Grant agreement number	870465
Project coordinator	Ghada El Serafy, Deltares
Project start date and duration	1 st November 2019, 36 months
Project website	https://forcoast.eu/

Deliverable Information	
Work package number	5
Work package title	Service Operationalisation, Demonstration & Validation
Deliverable number	D5.4
Deliverable title	Final Coordinated Pilot Model Evaluation Report
Description	This report provides details the validation of the different coastal models used to generate ocean data and provide services under the FORCOAST project. In particular, the different observational platforms are described, and the quality of the model performance is assessed through the comparison against observed data, which is presented as both graphical information and Estimated Accuracy Numbers, following the guidelines proposed in Deliverable D5.3.
Lead beneficiary	MI_IE
Lead Author(s)	D. Pereiro – Marine Institute, Ireland T. Dabrowski – Marine Institute, Ireland
Contributor(s)	L. Barbut - Royal Belgian Institute of Natural Sciences A. Caballero - AZTI F. Campuzano - Instituto Superior Técnico M. Maar - Aarhus University, Denmark J. Murawski - Danish Meteorological Institute S. Querin - OGS Italy J. She - Danish Meteorological Institute J. Staneva - HZG, Germany L. Vandenbulcke - Jiloo, Romania

Revision number	2
Revision date	15/10/2022
Status (Final (F), Draft(D), Revised Draft (RV))	RV
Dissemination level (Public (PU), Restricted to other program participants (PP), Restricted to a group specified by the consortium (RE), Confidential for consortium members only (CO))	PU

Document History			
Revision	Date	Modification	Author
0	29/04/2022	1 st draft	D. Pereiro, T. Dabrowski
0	31/05/2022	Pilot contributions	Pilots
1	28/06/2022	Review	D. Pereiro
2	03/10/2022	Added model validation conclusions table in the Conclusions section.	D. Pereiro, T. Dabrowski

Approvals				
	Name	Organisation	Date	Signature (initials)
Coordinator	Ghada El Serafy	Deltares		
WP Leaders	Tomasz Dabrowski	MI_IE		

Reviewer comments and reply		
Date	Comment	Reply
Sep 2022	<p>The deliverable is expected to be updated to provide clear quantification of the performance compared to reference data.</p> <p>Currently, it is difficult to be sure what has been proposed in D5.3 has been achieved. Therefore, to obtain an improved overview of the validation efforts, which are described in detail, it would be useful to have a summary table like that now included in D5.3. This would allow us to see a synthesis of the validation results, with the values for each statistical measure listed in D5.3's summary table being provided.</p>	<p>Summary table on model validation conclusions has been included on the Conclusions section.</p>

PROPRIETARY RIGHTS STATEMENT

This document contains information, which is proprietary to the FORCOAST consortium. Neither this document, nor the information contained within may be duplicated, used or communicated except with the prior written permission of the FORCOAST coordinator.



Executive Summary

This report provides details about the validation of the different coastal models used to generate ocean data and provide services under the FORCOAST project. In particular, the different observational platforms are described, and the quality of the model performance is assessed through the comparison against observed data, which is presented as both graphical information and Estimated Accuracy Numbers, following the guidelines proposed in Deliverable D5.3.

Table of Contents

Executive Summary.....	4
1. Introduction	15
2. Coastal model validation per Pilot site	16
2.1 Pilot 1: Portugal.....	16
2.1.1 Hindcast validation.....	22
2.1.2 Forecast validation.....	23
2.1.3 Process-oriented validation	25
2.2 Pilot 2: Spain	28
2.2.1 Forecast validation.....	32
2.2.2 Process-oriented validation	38
2.3 Pilot 3: Bulgaria	44
2.3.1 Hindcast validation.....	44
2.4 Pilot 4: Belgium	49
2.4.1 Hindcast validation.....	51
2.4.2 Forecast validation	83
2.4.3 Process-oriented validation	85
2.5 Pilot 5: Ireland.....	86
2.5.1 Hindcast validation.....	86
2.5.2 Forecast validation.....	95
3.5.3 Process-oriented validation	101
2.6 Pilot 6: Denmark	102
2.6.1 Hindcast validation.....	104
2.6.2 Forecast validation	115
2.6.3 Process-oriented validation	117
2.7 Pilot 7: Romania	122
2.7.1 Hindcast and Forecast validation.....	122
2.7.2 Process-oriented validation	132
2.8 Pilot 8: Italy	132
2.8.1 Forecast validation.....	133
3. Conclusions	135
4. References	140



List of Figures

Figure 1.1 Portuguese Coast - PCOMS (right) and Lisbon Metropolitan Area - LisOcean (left) modelling domains. Spatial resolutions are approximately 5 km and 280 m respectively. The red box in the PCOMS and LisOcean domains indicates the boundaries of the subsequent lower domain. The LisOcean domain indicates the tidal gauges location that will be used for validation and the location of the ExporSado production area where a monitoring tower was installed. 17

Figure 1.2 Portuguese environmental Agency (APA on its Portuguese acronym) monitoring stations in the Sado estuary. . 17

Figure 1.3 Aerial image of the Sado Estuary with the main place names (Source: Google Earth) 18

Figure 1.4 Average surface current speed and direction for April and November 2020 and January-April 2021 in the Sado estuary from the results of the LisOcean numerical model. Vectors represent the intensity and direction of the current and are represented every two cells. 19

Figure 1.5 Mean surface temperature in the Sado estuary from the LisOcean numerical model results every two months for the period May 2020-March 2021 The vectors represent current intensity and direction and are represented every two cells. 21

Figure 1.6 Mean surface salinity in the Sado estuary from the LisOcean numerical model results every two months for the period May 2020-March 2021 The vectors represent current intensity and direction and are represented every two cells. 22

Figure 1.7 Time series (left) and scatter plot cloud (right) of the tidal height observed by the Sesimbra tide gauge (orange) and simulated by the LisOcean model (blue line) in the period October 2018 - January 2021 (left). 22

Figure 1.8 Time series (left) and scatter plot cloud (right) of the tidal height observed by the Setúbal-Troia tide gauge (orange) and simulated by the LisOcean model (blue line) in the period October 2018 - January 2021 (left). 23

Figure 1.9 Time series of the tidal height observed by tide gauge (orange) and simulated by the LisOcean model (blue line) in the Setúbal-Troia (left) and Sesimbra (right) for April 2022. 23

Figure 1.10 Water temperature time series observed (orange dots) and simulated by the LisOcean model (blue line) in the estuary mouth (top) and near the Alcacer Channel (below) since November 2018. 24

Figure 1.11 Water temperature time series observed by the monitoring platform (orange dots) and simulated by the LisOcean model (blue dots) in the ExporSado production area since December 2021. 25

Figure 1.12 Processed environmental conditions (monthly average) for the years 2019-2020: temperature (top left), wind intensity (top right), precipitation (bottom left) and Sado river flow (bottom right). The meteorological conditions were calculated from the results of the IPMA numerical model and the river flows correspond to those of São Romão do Sado station. 26

Figure 1.13 Monthly wind analysis verified during the years 2019-2020, in terms of direction and intensity (results from the IPMA meteorological model with 2.5 km spatial resolution for the Setubal city area). 28

Figure 2.1 CROCO model spatial coverage and temperature forecast for 2021-07-08-01:00. 28

Figure 2.2 Location of the Donostia buoy (red star) over CROCO model coverage area (cyan). 30

Figure 2.3 Diagram of the Donostia Buoy with instrumented line and mooring line. 31

Figure 2.4 EuskOOS HFR coverage (small black squares) and its antennas location (Magenta and Red stars: Matxitxako and Higer stations respectively), as well as the coverage of the coastal model (cyan square). 31

Figure 2.5 CROCO model coverage area (transparent red area inside the rectangle) over the SST (°C) obtained from the SST_EUR_PHY_L4_NRT_010_031 product of CMEMS on August 21, 2021. 32

Figure 2.6 Model and in situ buoy TS (up) and U-V velocity (down) plots, for 2021-06-18 + 3 day (96 hours) horizon. 33

Figure 2.7 (Up) Model and in situ T, S 24-hour forecast average values, on a daily basis, from June to September. (down) Model and in situ U and V velocities 24-hour forecast average values, on a daily basis, from June to September. 34

Figure 2.8 EuskOOS HFR coverage (small black squares) and its antennas location (Magenta and Red filled dots: Matxitxako and Higer stations respectively). The red line indicates the limits of the common HFR - model spatial comparison footprint area..... 36

Figure 2.9 Daily surface currents from (up) EuskOOS HFR system and (down) CROCO model forecast. Left figures are from 2021-July-28 and right figures are from 2021-July-05. 37

Figure 2.10 East-West (up) and Sount-North (down) surface velocity components ME (m/s) for a 24 hours forecast. 37

Figure 2.11 RMSE values of zonal (up) and meridional (down) daily surface velocity components, for the 24 hours forecast. 38

Figure 2.12 RMSE (a) and ARMAE (b) maps between observed and modelled daily SST gradients. 39

Figure 2.13 Maps of the fronts observed (in red) and modelled (black) for 14 days where observations were almost cloud free. 41

Figure 2.14 The different score values corresponding to each test are depicted for the analyzed 14 days. The radius and weights of each test are depicted in Tables 2.3-2.7 42

Figure 3.1 Black Sea bathymetry and the WBS WAM wave model nest (black square). 45

Figure 3.2 Scatter plots of significant wave height (Hs) of all in-situ data available. Also shown are the estimated bivariate probability density (coloured area), the linear slope-fit regression of modelled and observed wave heights (red line), specific quantiles taken from the empirical cumulative density function (black line), and the diagonal (blue line). Furthermore, summary statistics and skill scores are included. R: reference (satellite) data, M: model data. 47

Figure 3.3 Scatter plots of significant wave height (Hs) for Q1 (JFM) 2020 of 7 different satellites and for the whole period using all satellites merged (lower right). See the labels for the satellite names. Also shown are the estimated bivariate probability density (coloured area), the linear slope-fit regression of modelled and observed wave heights (red line), specific quantiles taken from the empirical cumulative density function (black line), and the diagonal (blue line). Furthermore, summary statistics and skill scores are included. R: reference (satellite) data, M: model data. 48

Figure 4.1 Map of the different stations in the North Sea used in the validation..... 49

Figure 4.2 Map of the different stations close to the Belgian coast used in the validation. 50

Figure 4.3 Summary statistics of validation metrics per stations, comparing the sea level CSM forecasts to observations. The stations Aberdeen, Barmouth and Bournemouth have only three data points, and are removed from the two last graphs since they strongly deviate from other points. 52

Figure 4.4 Summary statistics of validation metrics per stations, comparing the sea level NOS forecasts to observations. From top to bottom: (1) Ratio of the standard deviation of the observations by the standard deviation of the forecasts; (2) Correlation; (3) Bias; (4) RMSE 54

Figure 4.5 Summary statistics of validation metrics per stations, comparing the sea level BCZ forecasts to observations. From top to bottom: (1) Ratio of the standard deviation of the observations by the standard deviation of the forecasts; (2) Correlation; (3) Bias; (4) RMSE. 55

Figure 4.6 Position of the stations for the validation of SST forecasts. 58

Figure 4.7 Taylor diagrams of the comparison of the sea surface temperature model forecasts with observations. From top to bottom: Helgoland, Den Helder, Ijmuiden, Hoek van Holland and Europlatform. 60

Figure 4.8 Distribution of SST observations and NOS forecasts at Helgoland for the month 201402. Left: Observations. Right: Model..... 62

Figure 4.9 Time series of the monthly ratio (standard deviation of observations/model forecasts) and correlation for the sea surface temperature at the station Helgoland. Green lines are the months July..... 62

Figure 4.10 Time series of the monthly bias and Q25, Q50, Q75 quantiles comparing the sea surface temperature forecasts by NOS and observations at station Helgoland. The yellow lines highlight the months of July..... 63

Figure 4.11 Time series of the sea surface temperature as forecast by NOS or observed at Den Helder, 96 hours long. 64

Figure 4.12 Time series of the sea surface temperature as forecast by NOS or observed at Hoek van Holland, 96 hours long. 64

Figure 4.13 Time series of the monthly ratio (standard deviation of observations / model forecasts) and correlation for the sea surface temperature at the station Den Helder. Green lines are the months July. The month 201703 is withdrawn for lack of data..... 64

Figure 4.14 Time series of the monthly bias and Q25, Q50, Q75 quantiles comparing the sea surface temperature forecasts by NOS and observations at station Den Helder. The yellow lines highlight the months of July. 65

Figure 4.15 Time series of the monthly bias and Q25, Q50, Q75 quantiles comparing the sea surface temperature forecasts by NOS and observations at station Hoek van Holland. The yellow lines highlight the months July..... 65

Figure 4.16 Time series of the monthly ratio (standard deviation of observations / model forecasts) and correlation for the sea surface temperature at the station IJmuiden. Green lines are the months July. The month 201703 is withdrawn for lack of data..... 66

Figure 4.17 Time series of the monthly bias and Q25, Q50, Q75 quantiles comparing the sea surface temperature forecasts by NOS and observations at station IJmuiden. The yellow lines highlight the months July. 67

Figure 4.18 Time series of the sea surface temperature as forecast by NOS or observed at Europlatform, 96 hours long. ... 68

Figure 4.19 Time series of the monthly ratio (standard deviation of observations / model forecasts) and correlation for the sea surface temperature at the station Europlatform. Green lines are the months July. The month 201502 is removed for lack of data (367 points)..... 68

Figure 4.20 Time series of the monthly bias and Q25, Q50, Q75 quantiles comparing the sea surface temperature forecasts by NOS and observations at station Europlatform. The yellow lines highlight the months of July. 69

Figure 4.21 Position of the Optos and CMEMS stations for the comparison of sea surface salinity values. 70

Figure 4.22 Relative position of the Optos (Vlakte van de Raan) and CMEMS (VlaktevdRaan) stations by the Scheldt estuary. 70

Figure 4.23 Relative position of the optos (Hoek Van Holland) and CMEMS (Hoek Van Holland NAP) stations by the Rhine estuary. 71

Figure 4.24 Relative position of the optos (Helgoland) and CMEMS (Helgoland Ferry Box) stations by the Elbe estuary. 71

Figure 4.25 General performance in predicting the salinity at Helgoland by NOS (top) and Vlakte van de Raan by NOS (mid) and BCZ (bottom). 73

Figure 4.26 Position of the stations K13, Europlatform, Wandelaar..... 74

Figure 4.27 Taylor diagrams of the comparison of the significant wave height from Hypas forecasts with CMEMS observations at Europlatform,(top), K13 (mid) and Wandelaar (bottom) from 2013 to 2017. 76

Figure 4.28 Time series of monthly bias, median and quantiles Q25 & Q75 comparing Hypas significant wave height forecasts with observations at three stations, from top to bottom Europlatform, K13, Wandelaar. 78

Figure 4.29 Significant wave height forecasts by Hypas and observations for the month 201712 at Wandelaar..... 79

Figure 4.30 Position of the stations used for the validation of the average zero up-crossing frequency. 79

Figure 4.31 General performance of the models Hypas and Refra at forecasting the zero up-crossing frequency at (from top to bottom): Wandelaar, Europlatform, K13, Akkaert and Westhinder. 82

Figure 4.32 Time series of the monthly bias between average zero up-crossing frequency model forecasts and observation at various stations..... 83

Figure 4.33 Screenshot of the presentation of forecasting skill at the NOOS website. 84

Figure 4.34 Daily averaged differences of SST between forecast products of the ensemble members and the MME median (<http://noos.eurogoos.eu/increasing-noos-awareness/community-tasks/multi-model-ensemble-of-forecast-products>) ... 84

Figure 4.35 Direct comparison of daily averaged sea surface temperature observations with modelled daily averaged sea surface temperature observations from June till October 2018 at Westdiep (51.18°N, 2.67°E). 85

Figure 5.1 Observational platforms in Galway Bay: (a) Galway Port tide gauge shown as a yellow marker, ADCP moorings (green markers, A, B, C from north to south), CPT loggers (red markers, north: Killeenaran, south: Kinvara West and Kinvara East) and MUR-SST 0.01° grid (dots) in the Galway Bay..... 86

Figure 5.2 Observed (blue) and modelled (red) time series of the sea level at Galway Port from November 2019 to September 2020. 87

Figure 5.3 Scatter plot of the observed vs. the modelled sea level at Galway Port. $y = -0.013 + 1.039x$; $R2 = 0.975$ 88

Figure 5.4 Observed (blue) and modelled (red) time series of the currents at the ADCP A mooring from spring-summer 2018: u-component (top) and v-component (bottom) at the surface level are shown..... 89

Figure 5.5 Observed (blue) and modelled (red) time series of the currents at the ADCP B mooring from spring-summer 2018: u-component (top) and v-component (bottom) at the surface level are shown..... 90

Figure 5.6 Observed (blue) and modelled (red) time series of the currents at the ADCP C mooring from spring-summer 2018: u-component (top) and v-component (bottom) at the surface level are shown..... 91

Figure 5.7 2019-2020 temperature time series at Killeenaran (53°11'52.2"N 08°56'32.2"W) according to the Galway Bay model (red), the Killeenaran CPT logger (dots) and the weekly measurements at the Killeenaran pier with the CO 310 hand-held sensor (green) 92

Figure 5.8 CPT logger and CO 310 temperature measurements at Killeenaran (53°11'52.2"N 08°56'32.2"W) vs Galway Bay model predictions. CPT logger: $y = -0.370 + 1.021x$; $R2 = 0.976$. CO 310: $y = 0.077 + 0.957x$; $R2 = 0.934$ 93

Figure 5.9 2019-2020 salinity time series at Killeenaran (53°11'52.2"N 08°56'32.2"W) according to the ROMS Galway Bay model, the Killeenaran CPT logger and the weekly measurements at the Killeenaran pier with the CO 310 hand-held sensor. 93

Figure 5.10 CO 310 salinity measurements at Killeenaran (53°11'52.2"N 08°56'32.2"W) vs ROMS Galway Bay model predictions. $y = 6.348 + 0.814x$; $R2 = 0.834$ 94

Figure 5.11 2019-2020 temperature time series at Kinvara (53°09'36.0"N 08°57'06.1"W) according to the ROMS Galway Bay model (black: vertically-averaged, pink: surface), the Kinvara CPT logger (blue) and the MUR-SST remote-sensing product (green)..... 94

Figure 5.12 CPT logger temperature measurements at Kinvara (53°09'36.0"N 08°57'06.1"W) vs Galway Bay model predictions. $y = -0.657 + 1.041x$; $R2 = 0.987$ 95

Figure 5.13 CTD sampling on the 18th of May 2021 near Killeenaran. Observed (blue) and Galway Bay model forecasts (F-1, F-2, F-3) for temperature and salinity (red) are shown..... 96

Figure 5.14 CTD sampling on the 18th of August 2021 near Killeenaran. Observed (blue) and Galway Bay model forecasts (F-1, F-2, F-3) for temperature and salinity (red) are shown. 97

Figure 5.15 CTD sampling on the 5th of November 2021 near Killeenaran. Observed (blue) and Galway Bay model forecasts (F-1, F-2) for temperature and salinity (red) are shown. F-3 was not produced for the 5th of November. 98

Figure 5.16 CTD vs. ROMS Galway Bay model temperature for forecasts F-1 (left), F-2 (center) and F-3 (right) taking the three samplings (spring, summer, autumn), every station and all depths into account. 99

Figure 5.17 CTD vs. ROMS Galway Bay model salinity for forecasts F-1 (left), F-2 (center) and F-3 (right) taking the three samplings (spring, summer, autumn), every station and all depths into account. 99

Figure 5.18 CPT logger vs. ROMS Galway Bay model temperature at Killeenaran (53°11'52.2"N 08°56'32.2"W) for forecasts F-1, F-2 and F-3..... 100

Figure 5.19 CPT logger vs. ROMS Galway Bay model temperature at Kinvara (53°09'36.0"N 08°57'06.1"W) for forecasts F-1, F-2 and F-3. 101

Figure 5.20 CPT logger vs. ROMS Galway Bay model temperature at Kinvara East (53°09'48.24''N 08°56'59.28''W) for forecasts F-1, F-2 and F-3..... 102

Figure 5.21 (a) ROMS Galway Bay model vs. CO 310 weekly measurements of salinity at Killeenaran (53°11'52.2''N 08°56'32.2''W), highlighting the threshold $S = 30$.(b) ROC curve analysis..... 102

Figure 6.1 Limfjord in-situ observation stations for sea level: tide gauge stations (blue circles), salinity and temperature regular cruise data (green filled circles) and temperature observations (green circle) at aquafarming site KFO1 (Kulturfelt Follup Odde). 103

Figure 6.2 Centralized Root-Mean-Square-Error in units of centimeter (cm) at Limfjord tide gauge stations. 106

Figure 6.3 Pearson Correlation Coefficient between modelled and observed sea level time series at Limfjord tide gauge stations (2015-2019). 107

Figure 6.4 Comparison of modelled and observed salinity and temperature dynamic in the western Limfjord (Oddesund), central Limfjord (Nykøbing Mors) and eastern Limfjord (Nibe Bredning), near the surface (red) and near the sea bed (blue). Observations ne 108

Figure 6.5 Temperature statistics, Bias and centralized Root-Mean-Square-Error for NOVANA environmental monitoring cruise stations (30 stations in the entire Limfjord)..... 109

Figure 6.6 Salinity statistics, Bias and centralized Root-Mean-Square-Error for NOVANA environmental monitoring cruise stations (30 stations in the entire Limfjord). 109

Figure 6.7 Temperature validation statistics: bias and centralized Root-Mean-Square-Error (cRMSE) for all Limfjord stations at the surface (left) and at 4.5m-5m depth (right). 110

Figure 6.8 Vertical profiles of model validation statistics: bias, centralized Root-Mean-Square-Error (cRMSE), correlation coefficient (CC) of Salinity (a.) and temperature (b.). Shown are the horizontally averaged model validation statistics for all Limfjord stations..... 110

Figure 6.9 Mean modelled surface a) DIN concentration (NO_x + NH₄) b) Phosphate concentration c) Chlorophyll concentration and d) Secchi depth. Red numbered dots mark the stations where measurements were carried out 111

Figure 6.10 Seasonal variability of DIN at the four stations shown in Figure 6.1. Each grey line represents one of the years from 2011 to 2017 and all data points are shown for the same years..... 112

Figure 6.11 Seasonal variability of PO₄ at the four stations shown in Figure 6.1. Each grey line represents one of the years from 2011 to 2017 and all data points are shown for the same years..... 113

Figure 6.12 Seasonal variability of modelled and measured surface chlorophyll concentration at the four stations in Limfjord (Figure 1) from 2011 to 2017. Red crosses mark observations of mixotrophs that are not described in the model..... 113

Figure 6.13 Seasonal variability of the bottom oxygen concentration at six stations from 2011 to 2017 (Figure 6.9d). 114

Figure 6.14 Taylor plots showing Pearson correlation, normalized standard deviation and ME for time series of a) DIN, b) PO₄ and c) Chlorophyll for the years 2011 to 2017. The numbers refer to the stations shown in Figure 6.1. The perfect fit is located at CORR = 1 and normalized STD = 1, at the black dot. The colours show the mean error at each station. 115

Figure 6.15 Surface temperature in Oddesund (left) and central Limfjord (right). Circles denote model data with specified forecast range, triangles – profile observations, where only data with depth <1 m are selected, line – is satellite SST data. 116

Figure 6.16 Surface temperature at Thisted (left) and Nibe Bredning (right). Circles denote model data with specified forecast range, triangles – profile observations, where only data with depth <1 m are selected, line – is satellite SST data. **Error!**

Bookmark not defined.

Figure 6.17 Bias (left) and central RMSD (right) of modelled temperature in Limfjord as average of stations depending on forecast range. 116

Figure 6.18 Modelled versus observed temperature in the upper layer for stations within Limfjord. Left – profile observations, right – satellite SST temperature. Lines denote deviation by 1 degree from observed value. 116

Figure 6.19 Water level at Nykoebing Mors. Circles denote model data with specified forecast range..... 117

Figure 6.20 Central RMSD (left) and bias (right) of water level in stations of Limfjord as function of forecast range. 117

Figure 6.21 Modelled (blue) and observed sea level (black) in the historical period 2015-2019, at Lemvig tide gauge station. The local warning level (grey) of 1.35 m and the sea level of a 20 years event (red) of 1.82 m are shown as well. 118

Figure 6.22 Modelled (blue) and observed sea level (black) during storm Dagmar and Egon (9th and 10th of January 2015) at Lemvig tide gauge station. The local warning level (grey) of 1.35m and the sea level of a 20 years event (red) of 1.82m are shown as well..... 118

Figure 6.23 Left: Sea level anomalies of extreme events, i.e., storm surges, exceeding warning levels of 1.35m of observed sea levels at Lemvig tide gauge station. All events lay in the range of permitted error values for operational assessment, which is 10% of the observed peak value, with a minimum of 10cm. Right: Error-Range dependent assessment of miss-rate, i.e., percentage of forecasts that are outside the permitted error-range of the observations. 119

Figure 6.24 Water temperature near the sea bed at KFO1 station at the Aquafarming site. The shallow water radiation scheme improved the temperature forecast (red), which is much closer to the observations (black) than the previous version (blue). 120

Figure 6.25 Diurnal signal of sea bottom temperature, i.e., the monthly mean temperature, for each hour of the day, at KFO1 station, for the 4 months: June to September 2019..... 121

Figure 7.1 (Top) Spatial average of SST bias between model and L4 satellite observations during 2020 and 2021 (bottom) RMSE..... 123

Figure 7.2 (Top) spatial average of SST bias between model and L3S satellite observations during 2020 and 2021 (bottom) root mean square error..... 124

Figure 7.3 Monthly spatial maps of SST root mean square error between Eforie model and satellite observations for 2020 125

Figure 7.4 Monthly spatial maps of SST root mean square error between CMEMS BS model and satellite observations for 2020..... 126

Figure 7.5 (Top) spatial average of chlorophyll-a bias between model and satellite observations during 2021 (bottom) root mean square error..... 127

Figure 7.6 Monthly spatial maps of chlorophyll-a root mean square error between Eforie model and satellite observations 128

Figure 7.7 Monthly spatial maps of chlorophyll-a root mean square error between CMEMS model and satellite observations 129

Figure 7.8 Current directions in the Eforie area, obtained by (left) ADCP measurements (right) equivalent points in the Eforie model, during 2019, for all measured depths 130

Figure 7.9 Current directions in the larger Eforie area, obtained by (left) ADCP measurements (right) equivalent points in the NWS model, during 2019..... 130

Figure 7.10 Current directions in the Eforie area, obtained by (left) ADCP measurements (right) equivalent points in the Eforie model, during 2021..... 131

Figure 7.11 Current directions in the larger Eforie area, obtained by (left) ADCP measurements (right) equivalent points in the NWS model, during 2021 131

Figure 8.1 Time series of surface chlorophyll (+/- 1 standard deviation) from satellite (green) and model (black) - upper plots. Statistics of bias and root mean square difference (RMS) are computed daily for the first day of forecast - central plots. Number of points in each L3 satellite image - lower plots. Areas shallower and deeper than 20 meters are defined as COAST (left column) and OPEN SEA (central column), respectively. The results for the WHOLE BASIN are plotted on the right column. Image taken from the MedEAF webpage..... 134

Figure 8.2 Surface chlorophyll [mg m⁻³] maps derived from MITgcm-BFM model output (simulation date: 2022-04-12; midday of the first day of forecast - left) and remote sensing data provided by CNR-GOS (right). Image taken from the MedEAF webpage..... 135

Figure 8.3 Sea surface temperature maps derived from MITgcm-BFM model output (simulation date: 2022-04-12; midnight of the first day of forecast [°C] - left) and remote sensing data provided by CNR-GOS [K] (right). Image taken from the MedEAF webpage..... 135

List of Tables

Table 2.1 List of datasets used for model validation.	29
Table 2.2 RMSE and CORR summary table at different depth and horizon timelines.	34
Table 2.3 Configuration of Test 1, with different numbers and distances of rings and with different weights allocated to each ring.....	42
Table 2.4 Configuration of Test 2, with different numbers and distances of rings and with different weights allocated to each ring.....	42
Table 2.5 Configuration of Test 3, with different numbers and distances of rings and with different weights allocated to each ring.....	42
Table 2.6 Configuration of Test 4, with different numbers and distances of rings and with different weights allocated to each ring.....	43
Table 2.7 Configuration of Test 5, with different numbers and distances of rings and with different weights allocated to each ring.....	43
Table 2.8 Mean distance, its STD and their sum (SUM) between each modelled front point and the nearest observed point for each date. Also, the percentage of the cases where the nearest observed point falls within the radii defined by SUM is shown.	43
Table 4.1 List of data providers per stations/parameter	50
Table 4.2 Statistics of the validation metrics from the comparison between observations at the Europlatform station and the CSM forecasts.	53
Table 4.3 Comparison of the statistics of the validation metrics of three models at Newport.	56
Table 4.4 Summary statistics of the model performance, all stations.....	56
Table 4.5 Summary statistics of the model performance, seven common stations.	57
Table 4.6 Summary statistics of the analysis of the error between NOS sea surface temperature forecasts and observations at various stations.....	61
Table 4.7 Statistics of the monthly RMSE between significant wave height forecasts by Hypas and observations, at different stations.	76
Table 4.8 Statistics of the monthly bias between Hypas Significant wave height forecasts and observations at three stations.	78
Table 4.9 Overview of the metrics when assessing the skill of the BCZ model (described in section x.x) at predicting the temperature at Westdiep (51.18°N, 2.67°E).....	85
Table 5.1 Galway Port observed sea level vs. ROMS Galway Bay predictions statistics (ME = Mean Error, MAE = Mean Absolute Error, RMSD = Root Mean Squared Difference, CORR = correlation, n = number of data points).....	87
Table 5.2 CORR, RMSD and ARMAE considering all depths and for the u-component, v-component and speed. An observational error of 1 cm/s has been used for the computation of the ARMAE value (Dabrowski et al., 2016).	88

Table 5.3 Killeenaran observed temperature vs. ROMS Galway Bay predictions statistics (ME = Mean Error, MAE = Mean Absolute Error, RMSD = Root Mean Squared Difference, CORR = correlation, n = number of data points)..... 91

Table 5.4 Killeenaran observed salinity (CO 310 sensor) vs. ROMS Galway Bay predictions statistics (ME = Mean Error, MAE = Mean Absolute Error, RMSD = Root Mean Squared Difference, CORR = correlation, n = number of data points). 92

Table 5.5 Kinvara West observed temperature vs. ROMS Galway Bay predictions statistics (ME = Mean Error, MAE = Mean Absolute Error, RMSD = Root Mean Squared Difference, CORR = correlation, n = number of data points)..... 94

Table 5.6 Observed vs. modelled temperature statistics (ME = Mean Error, MAE = Mean Absolute Error, RMSD = Root Mean Squared Difference, CORR = correlation, n = number of data points) for the CTD surveys conducted in Galway Bay in spring, summer and autumn and for the three forecasts (F1, F2, F3) available for each day..... 98

Table 5.7. Observed vs. modelled salinity statistics (ME = Mean Error, MAE = Mean Absolute Error, RMSD = Root Mean Squared Difference, CORR = correlation, n = number of data points) for the CTD surveys conducted in Galway Bay in spring, summer and autumn and for the three forecasts (F1, F2, F3) available for each day 99

Table 5.8 Observed vs. modelled temperature statistics (ME = Mean Error, MAE = Mean Absolute Error, RMSD = Root Mean Squared Difference, CORR = correlation, n = number of data points) for the three CPT loggers in Galway Bay and for the three forecasts (F1, F2, F3) available for each day. 100

Table 6.1 List of data sets used for model validation..... 104

Table 6.2 Statistics calculated from the spatial distribution of the biogeochemical tracers in the fjord..... 112

Table 7.1 Error statistics of the current velocity computed in the Eforie and NWS models for the 2019 and 2021 campaigns. 132

1. Introduction

Together with remote-sensing products, models play an essential role in delivering services within FORCOAST. Hydrodynamic, biogeochemical, particle-tracking models, etc. are used in FORCOAST for different purposes directly related to the delivery of services, such as predicting the path followed by larvae or contaminants, finding areas with best suitability for aquaculture farming, or for identifying ocean frontal activity relevant to the fisheries sector. For this reason, it is of great importance to provide end users with a qualified assessment of the performance of the coastal models used within the FORCOAST project.

The objective of this report is to present the validation work carried out at the different Pilot sites to provide an estimation of the accuracy of the coastal models and their ability to reproduce the dynamics and the processes in the areas under consideration. To achieve this goal, the guidelines detailed in Deliverable D5.3 are followed, with three different types of validation efforts:

1. **Hindcast validation**, where the focus is on the model's ability to reproduce the conditions in the past.
2. **Forecast validation**, where the objective is to evaluate the model's predictions of a future state.
3. **Process-oriented validation**. The process-oriented validation should focus on events that are of interest to the end users and, at the same time, characterize the Service Module that is being developed by the Pilot, if any.

It is important to notice that not every Pilot has accomplished these three types of validation due to varying reasons, including:

- (a) The lack of archiving of hindcasts or forecasts. For instance, archiving of hindcasts has been deemed unnecessary for SM-F2 Front Detection, where the focus is on forecasting mesoscale fronts and the purpose of this service is to provide a forecast of best fishing areas for the upcoming days.
- (b) When using the same forcing in both hindcast and forecast modes, a good performance in hindcast mode probably means that the performance of the model would be good in forecast mode as well, and vice versa.

For a reference to the foreseen validation efforts at each site, please consult the table in **D5.3 Report describing common methodology for numerical model validation**, Section 6.

2. Coastal model validation per Pilot site

2.1 Pilot 1: Portugal

Hydrodynamics and water properties in coastal areas are a consequence of physical, chemical and biological processes with different spatial and time scales. This makes it difficult to describe them based exclusively on field data which are discrete in space and time. Inside the estuary, the tide and discharges of terrestrial origin are the main mechanisms generating variability, while in the inlet zone the variability is determined by the interaction with the continental shelf, being due essentially to the seasonality of the flow, which determines the interaction between the surface layer of the ocean and the deep water.

The exchanges between the estuary and the ocean are determinant for the properties of water in its interior, and consequently the estuary model has to be coupled to a platform model. The greater spatial variability recorded inside the estuary requires the use of a model with higher resolution. This is not required for the area on the platform and therefore the system has to be based on a system of coupled models as indicated in Figure 1.1, using a downscaling technique where information moves from the coarsest models to the most refined models.

The system used in this study is composed of three embedded models. The solution of the larger scale model (regional model of the western zone of the Iberian Peninsula - PCOMS, Mateus et al., 2012; Campuzano et al., 2018) runs in operational mode and provides boundary conditions (currents, sea level, temperature and salinity) to a model for the Lisbon metropolitan area. This includes the Sado and Tagus estuaries, with a resolution of 280 meters. The regional model, with spatial pitch of about 5 km, is the first to run and results for the central region of Portugal are stored at intervals of 15 minutes. These results are subsequently used to force the Estuary model. The whole system is based on the MOHID numerical modelling system, whose development was initiated at MARETEC-IST and is currently maintained by the main users. This modelling system allows simulating physical and biogeochemical processes. The model will be validated using:

- Continuous observations from the new ExporSado Longa monitoring station, recording sea level, seawater temperature and salinity, pH, suspended sediments, dissolved oxygen and chlorophyll concentration.
- IH (hydrographic institute) tide gauges sea level series.
- Portuguese environmental Agency (APA in the Portuguese acronym) data collected under the Water Framework Directive programme, in several areas of the estuary, including water temperature, salinity, oxygen, etc.
- Remote-sensing sea surface temperature and chlorophyll.

In addition, a description of the residual circulation and essential ocean variables will be used to explain the system processes.

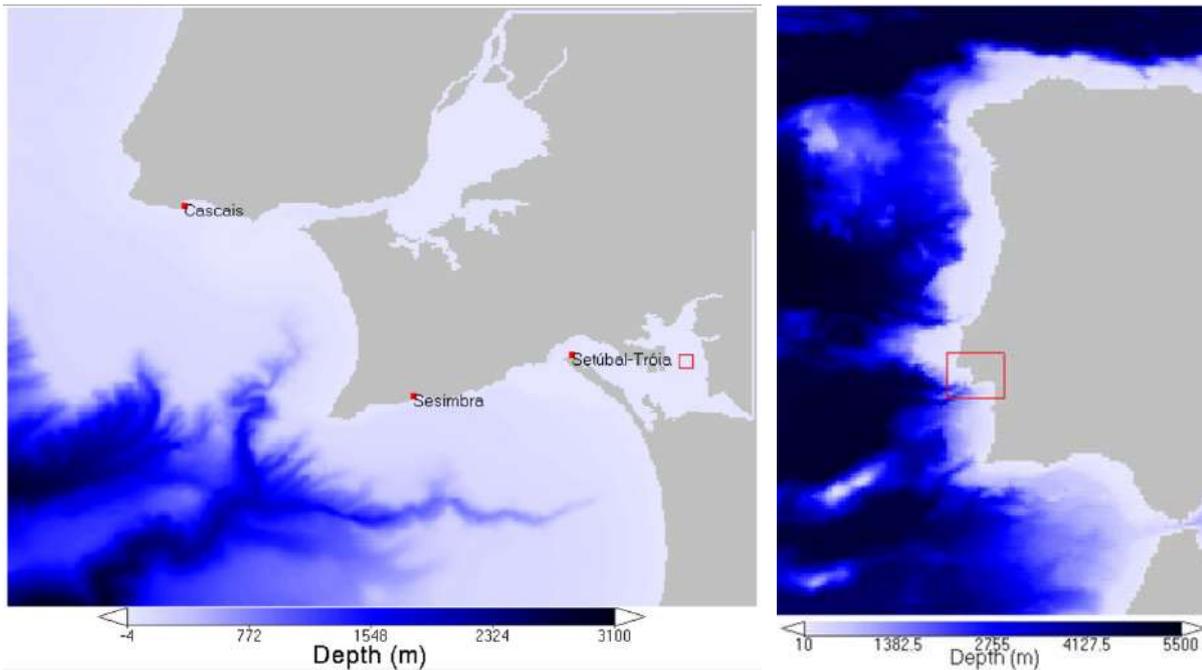


Figure 1.1 Portuguese Coast - PCOMS (right) and Lisbon Metropolitan Area - LisOceon (left) modelling domains. Spatial resolutions are approximately 5 km and 280 m respectively. The red box in the PCOMS and LisOceon domains indicates the boundaries of the subsequent lower domain. The LisOceon domain indicates the tidal gauges location that will be used for validation and the location of the ExporSado production area where a monitoring tower was installed.

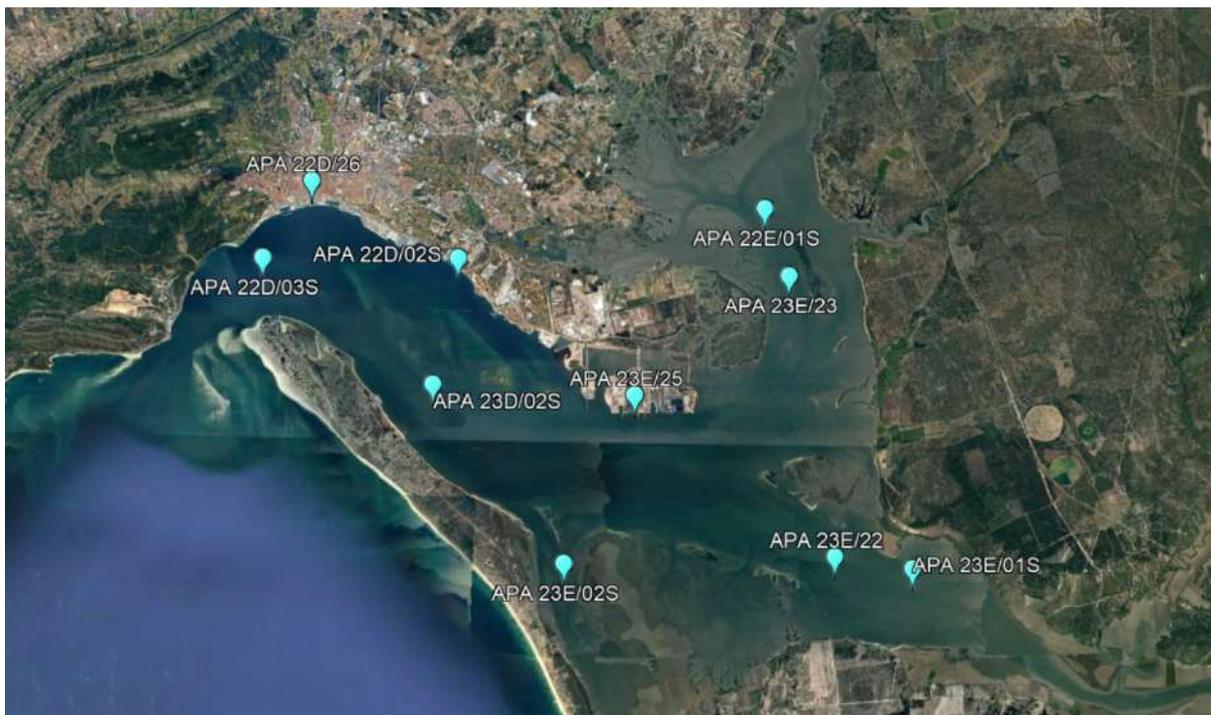
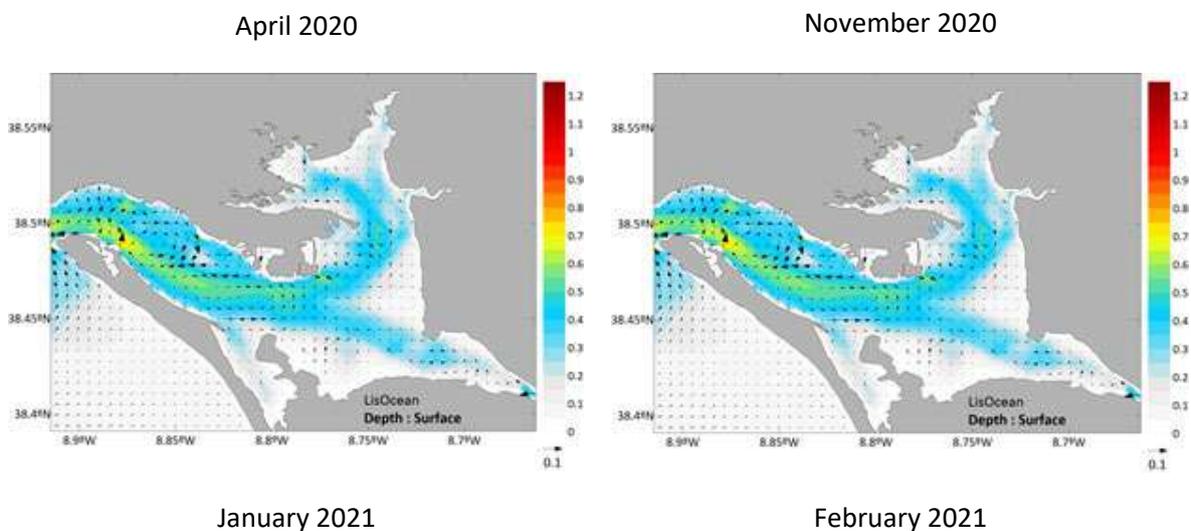


Figure 1.2 Portuguese environmental Agency (APA on its Portuguese acronym) monitoring stations in the Sado estuary.



Figure 1.3 Aerial image of the Sado Estuary with the main place names (Source: Google Earth)

Flow in the Sado estuary is forced mainly by the tide, coming from the ocean. Wind has a secondary contribution to the flow pattern, and river discharge is important mainly in the Alcácer Channel. Consequently, the velocity field has an oscillatory regime with maximum velocity intensities in the middle of floods and ebb tides and minimum velocities at the estuaries. Figure 1.4 shows the estuarine circulation pattern is maintained throughout the year. Only in February and March 2021, we can clearly see the increase in intensity of the velocities coming from the Sado River, due to the large increase in flow, which was up to 100 times the typical flow. Residual circulation shows that:



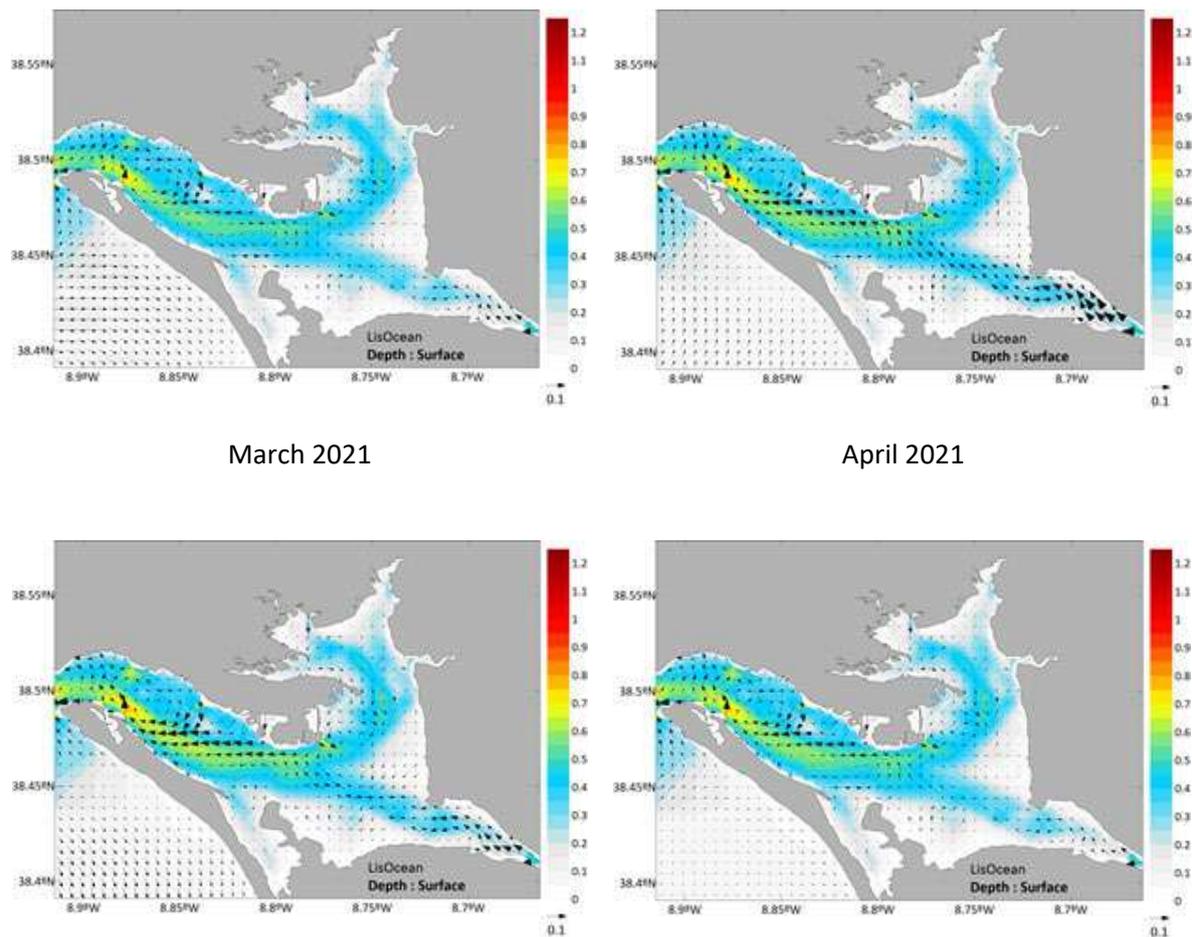


Figure 1.4 Average surface current speed and direction for April and November 2020 and January-April 2021 in the Sado estuary from the results of the LisOcean numerical model. Vectors represent the intensity and direction of the current and are represented every two cells.

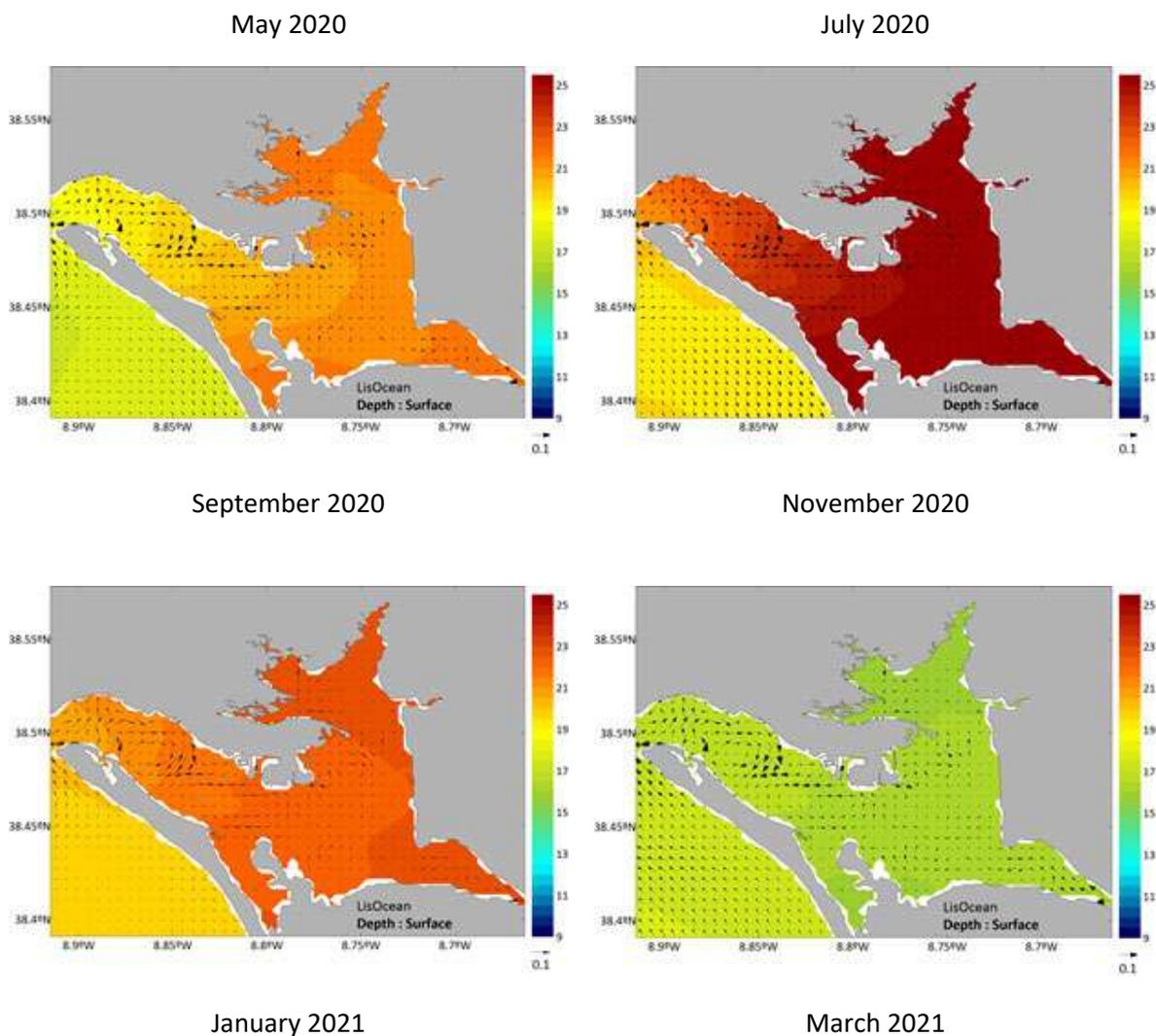
- At the estuarine mouth, the flow is preferably flooding on the Troia side and ebb from the Comenda side.
- The flood flow dominates against the ebb in the North Channel.
- The residence time in the Marateca-Carrasqueira region must be high, because there are no clear flow directions.
- An intense, large vortex enters the South Channel. This vortex separates the region of the estuary dominated by the interaction with the sea, downstream, and the upstream region of lagoon character.

The constancy of the residual circulation pattern is a consequence of the hydrodynamics of the estuary being forced mainly by the tide, which shows a little variable pattern throughout the year.

Figure 1.5 shows the spatial distributions of monthly mean temperature. A large variability in surface temperature can be observed, with a very extensive region with values above 25°C in the interior of the estuary in July and temperatures of the order of 10°C in January. Temperature is a good tracer of the mixing dynamics in the estuary when the difference between the temperature at sea and in the

interior of the estuary is high, which is the case for the months of July and January. The distributions for these months show that the mixing is dominated by the two large residual vortices existing in the estuary analysed in the previous paragraph. Upstream of these vortices (Maratec, Carrasqueira, Canal de Alcácer) the temperature gradient is very low, showing that in this region the water has time to mix, i.e., has a high residence time.

Figure 1.6 shows the salinity distribution with values higher than 35 PSU in most estuary. Only in the Alcácer Channel there are typical estuarine gradients. This is a consequence of the low fresh water flows received by the Sado Estuary, which makes it a special estuary. Typically, estuaries receive large amounts of freshwater and show clear gradients between the mouth and the river entry points. In March 2021, the salinity values in Marateca/Carrasqueira are much lower than in other months. This is a consequence of the high flows recorded in the month of February. The fact that the signature of that water still persists in the following month and that the water is well mixed in that region shows that the residence time is high in the upstream zone of the estuary.



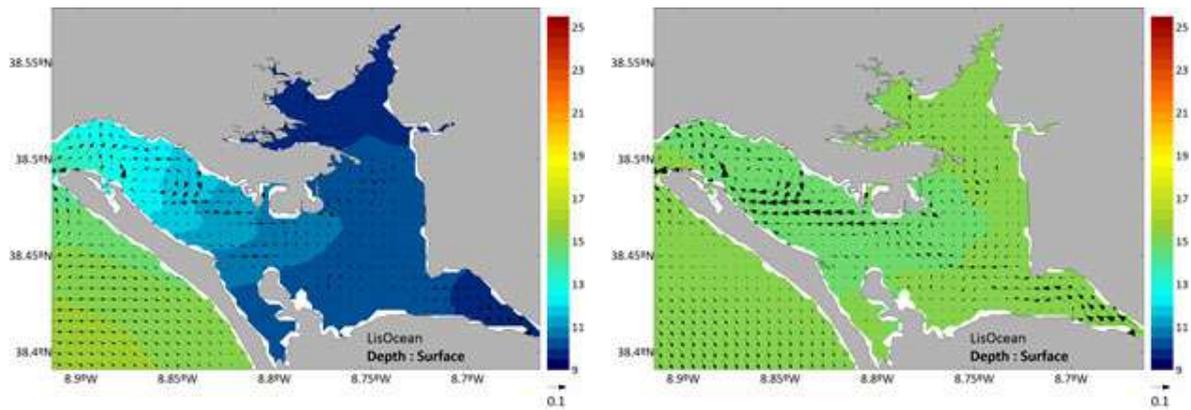
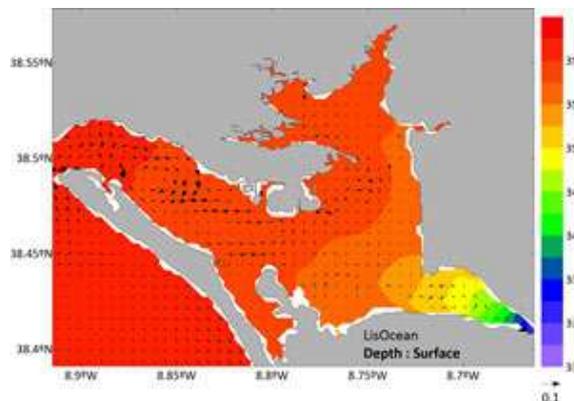
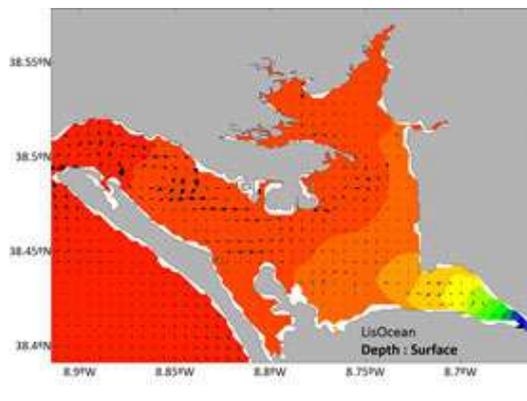


Figure 1.5 Mean surface temperature in the Sado estuary from the LisOcean numerical model results every two months for the period May 2020-March 2021. The vectors represent current intensity and direction and are represented every two cells.

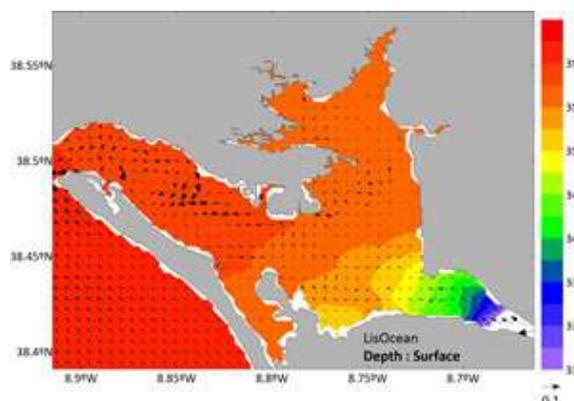
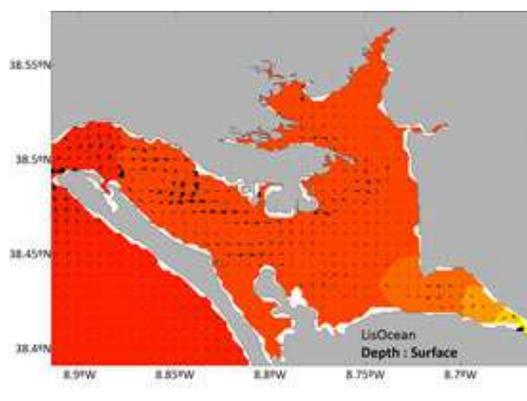
May 2020

July 2020



September 2020

November 2020



January 2021

March 2021

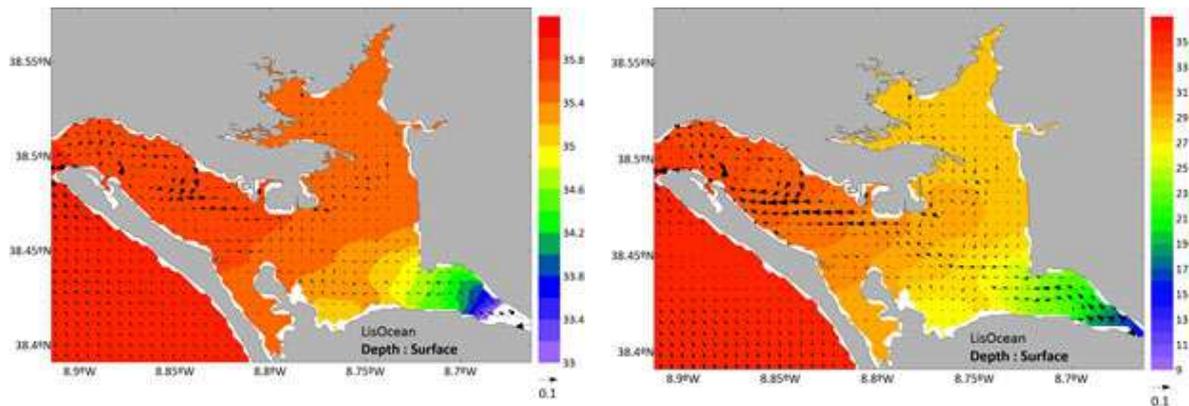


Figure 1.6 Mean surface salinity in the Sado estuary from the LisOcean numerical model results every two months for the period May 2020-March 2021. The vectors represent current intensity and direction and are represented every two cells.

2.1.1 Hindcast validation

The main tidal components are M_2 , due to the Moon, with a period of 12h25 and S_2 due to the Sun with a period of 12 hours. The difference in periods of these components generates a biweekly tidal cycle of spring tide/dead tide. When the Sun and Moon forcing are in phase (full Moon and new Moon) we have a spring tide and when they are 90° apart (rising and waning quarters) we have a neap tide. At low tide, the sea level at high tide is around 3 metres and at high tide, it is around 4 metres.

To validate the tidal propagation in the Sado estuary we used observations from the Sesimbra (38.44°N, 9.11°W) and Setúbal-Troia (38.49°N, 8.90°W) tide gauges of the Hydrographic Institute for the years 2018-2020. Figure 1.7 and Figure 1.8 show those of sea levels observed by the tide gauges of Sesimbra and Setúbal-Troia, respectively. Next to these time series are shown the results obtained with the LisOcean domain. Next to the time series, the scattering clouds of the pairs of observed and modelled values are shown, and the linear regression for each tide gauge.

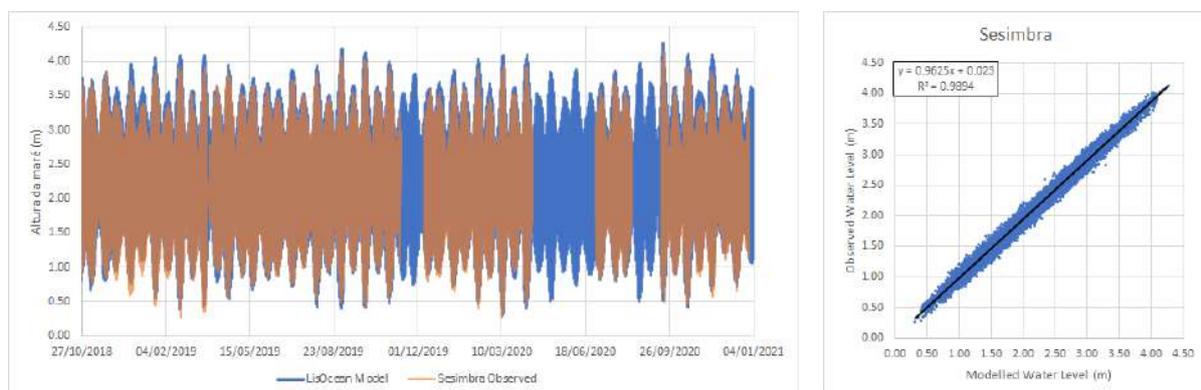


Figure 1.7 Time series (left) and scatter plot cloud (right) of the tidal height observed by the Sesimbra tide gauge (orange) and simulated by the LisOcean model (blue line) in the period October 2018 - January 2021 (left).

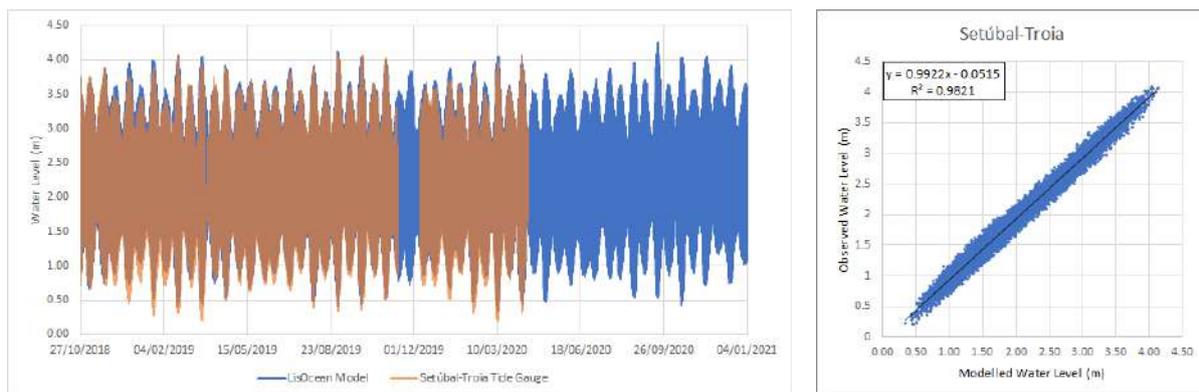


Figure 1.8 Time series (left) and scatter plot cloud (right) of the tidal height observed by the Setúbal-Troia tide gauge (orange) and simulated by the LisOcean model (blue line) in the period October 2018 - January 2021 (left).

2.1.2 Forecast validation

The validation process is being performed continuously and download from tidal gauge stations has been operationalised in order to compare continuously the model performance against observed data (Figure 1.9).

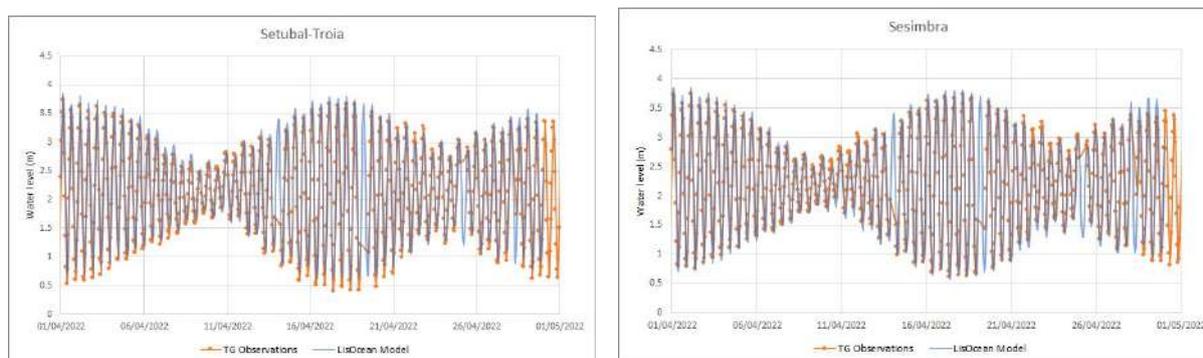


Figure 1.9 Time series of the tidal height observed by tide gauge (orange) and simulated by the LisOcean model (blue line) in the Setúbal-Troia (left) and Sesimbra (right) for April 2022.

In addition, the model is being compared with the APA stations. Figure 1.10 shows an example of model results with observations available for the area.

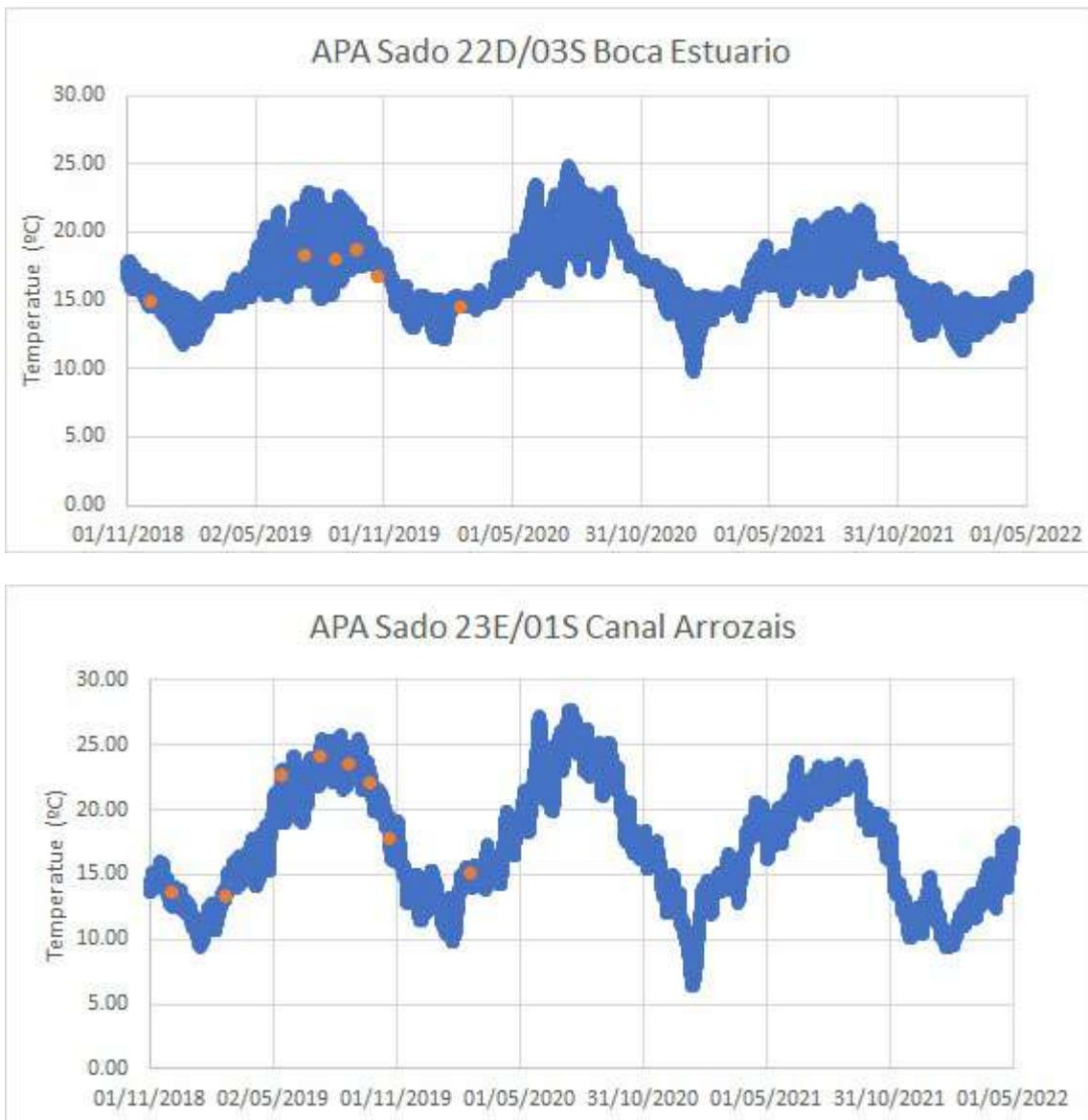


Figure 1.10 Water temperature time series observed (orange dots) and simulated by the LisOcean model (blue line) in the estuary mouth (top) and near the Alcacer Channel (below) since November 2018.

In order to obtain continuous observations in the production area, a monitoring platform was installed. Fig 1.11 shows a first comparison with water temperature.

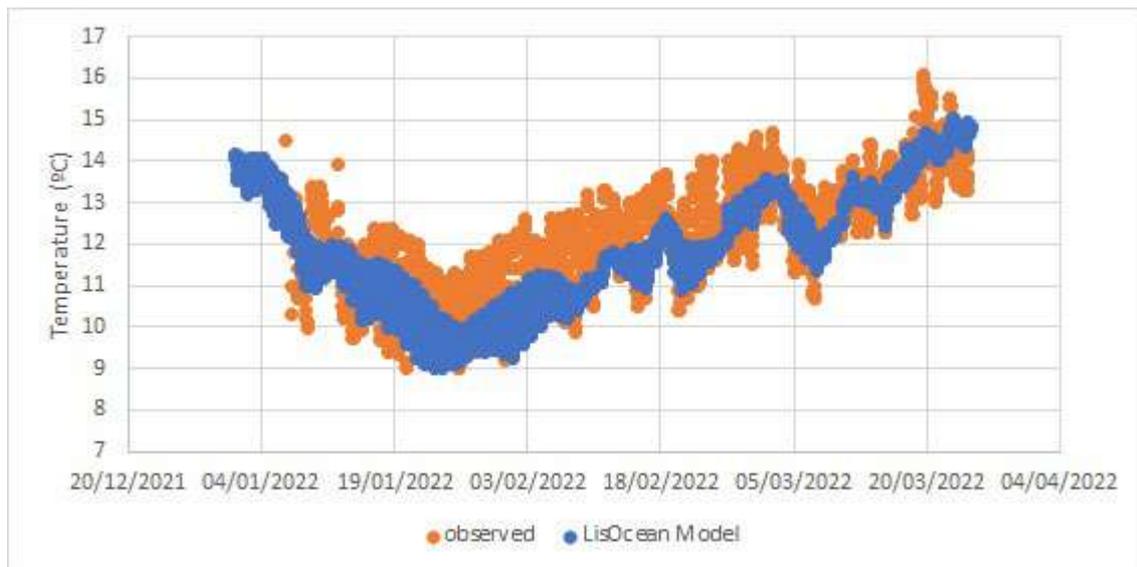


Figure 1.11 Water temperature time series observed by the monitoring platform (orange dots) and simulated by the LisOcean model (blue dots) in the ExporSado production area since December 2021.

2.1.3 Process-oriented validation

The temperature and salinity of the estuarine water determine its density and therefore contribute to the hydrodynamics of the estuary. These variables are also important for the buoyancy of the outfall plume and therefore their simulation is essential to simulate the dilution of the plume. Temperature and salinity depend on meteorological conditions and river flows, where the Sado River is the most important.

The meteorology and the Sado River flow have great annual variation, as shown in the time series represented in Figure 1.12. These monthly average series were obtained from hourly results of IPMA meteorological model and from the SNIRH flow time series at São Romão do Sado station, obtained from the EMODnet physics portal (<https://portal.emodnet-physics.eu/>).

The temperature has a typical seasonal evolution, with daily average values of 10 °C in January and 22 °C in July and August. Precipitation shows two peaks, one in April/May and another in autumn. The river flow is normally low, but increases throughout the autumn, with its maximum value of 10 m³/s reaching in December. In April, after the spring rainfall peak, the river flow reaches values around 2 m³/s. The average flow is of the order of 1 m³/s.

Although the average flow rate of the Sado river is low, it may present important peaks, with flows higher than 120 m³/s registered in February 2021 as a consequence of precipitation peaks and the extension of the hydrographic basin across an area of about 7700 km², being the Portuguese hydrographic basin with the largest surface area.

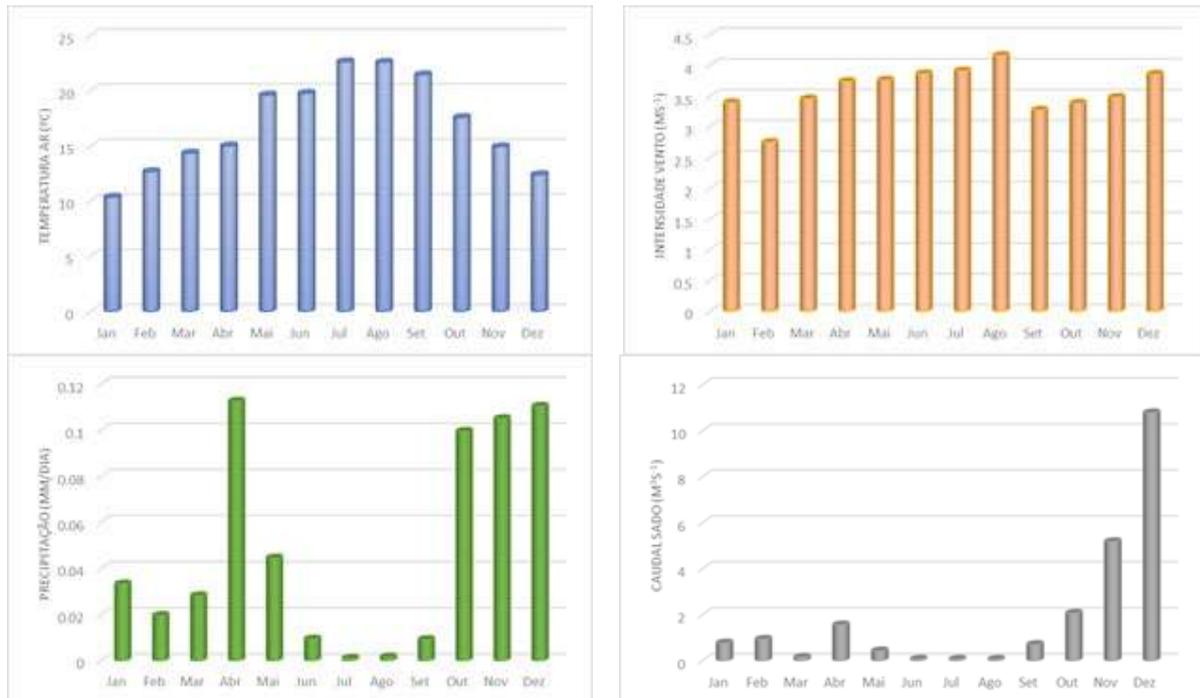
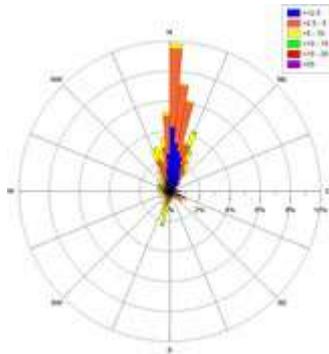


Figure 1.12 Processed environmental conditions (monthly average) for the years 2019-2020: temperature (top left), wind intensity (top right), precipitation (bottom left) and Sado river flow (bottom right). The meteorological conditions were calculated from the results of the IPMA numerical model and the river flows correspond to those of São Romão do Sado station.

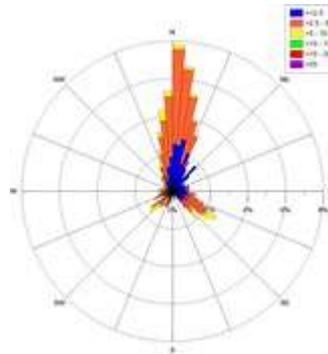
Wind speed is one of the forcing agents of the flow. It has some direct influence on the hydrodynamics, but is particularly important for heat exchange by convection and evaporation between the estuary and the atmosphere. The average wind intensity is fairly uniform throughout the year (Figure 1.12) and does not exceed 5 m/s with the maximum mean value being obtained in the month of August. Therefore, the model must be forced using instantaneous values. Figure 1.13 presents a monthly wind analysis for the Navigator outfall area from the results of the AROME meteorological model in the period 2019-2020 produced by IPMA, with 2.5 km spatial resolution.

In terms of intensity and direction, the wind presents the typical pattern of the Portuguese west coast with predominant north winds during most of the year. The months from June to August are dominated by moderate winds (2.5m/s to 10 m/s) of North component. During winter (December to February), the North component still dominates, but with lower intensities. During the months of September to November and April, the winds with S-SW component gain importance, reaching intensities above 10 m/s in some cases. The winds of this component appear in other months with less frequency, but also associated with higher intensities, such as in December.

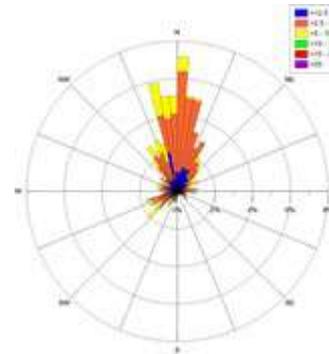
January



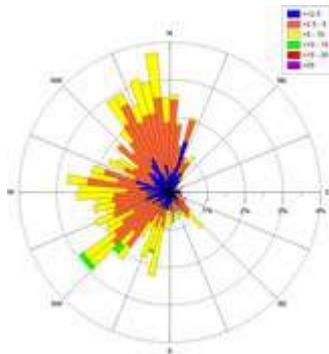
February



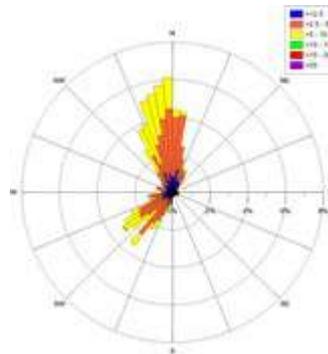
March



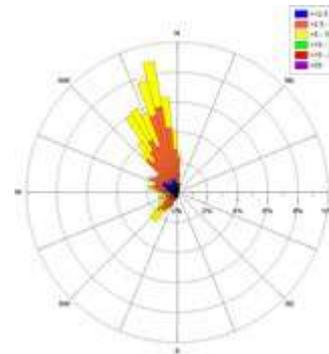
April



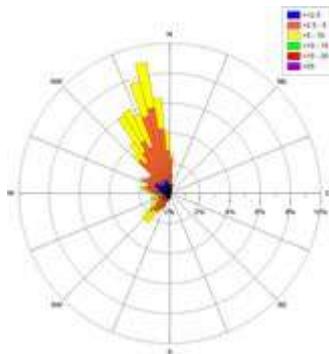
May



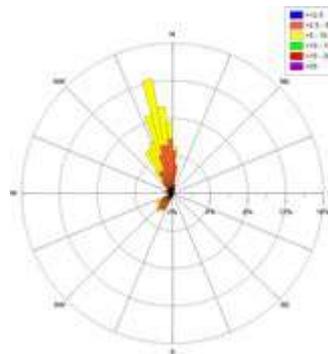
June



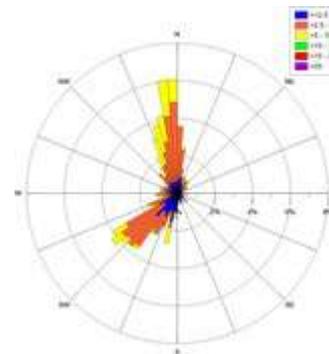
July



August



September



October



November



December



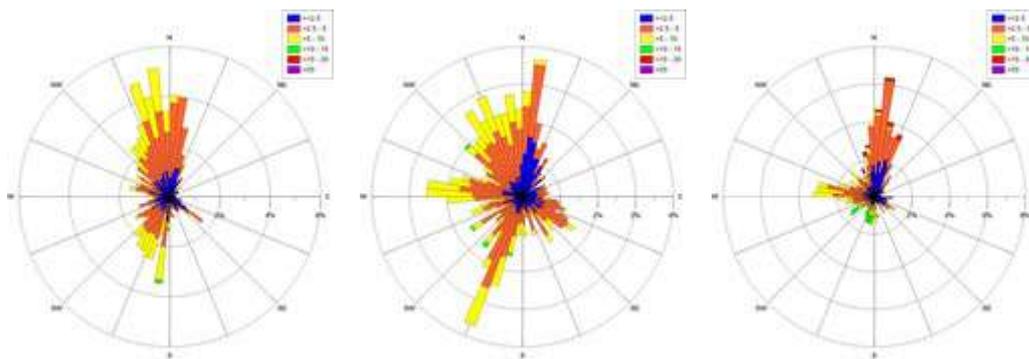


Figure 1.13 Monthly wind analysis verified during the years 2019-2020, in terms of direction and intensity (results from the IPMA meteorological model with 2.5 km spatial resolution for the Setubal city area).

2.2 Pilot 2: Spain

Pilot 2 is covering the south-eastern Bay of Biscay and is targeted to the fishing sector. The Service Module proposed in this pilot is called “Front Detection” that consists in monitoring and forecasting mesoscale sea surface fronts, as a proxy to areas of higher fishing opportunities. Within the context of the Front Detection Service Module, the effort will be focused on validating the surface and subsurface sea temperature (T) and salinity (S). Since the surface current variability at different scales are also of relevant interest to the end-users and the potential transferability of other Service Modules to this Pilot, surface and subsurface model currents will be also validated.

CROCO (Coastal and Regional Ocean COmmunity model, <https://www.croco-ocean.org>) is the numerical modelling tool used in EuskOOS (Basque Operational Oceanography System) to estimate the spatiotemporal evolution of the main physical variables of the marine environment in the south-eastern Bay of Biscay. This model is based on a new non-hydrostatic and non-Boussinesq solver, developed within the former ROMS kernel. CROCO is an extension of ROMS from which it inherited the robustness and efficiency of its time-splitting implementation, the accuracy of high-order methods, including its pressure gradient scheme for terrain-following coordinates, and computing performances. The CROCO domain used in EuskOOS covers the south-eastern Bay of Biscay, extending from 43.24°N to 44°N and from 3.4°W to 1.3°W, with a mean horizontal resolution of approximately 670 m. Vertically, the water column is divided into 32 sigma-coordinate levels. These levels are more concentrated within the surface waters, where most of the ocean variability occurs. The present CROCO configuration provides 96 hours forecasts of the sea conditions (i.e. T, S, sea surface height

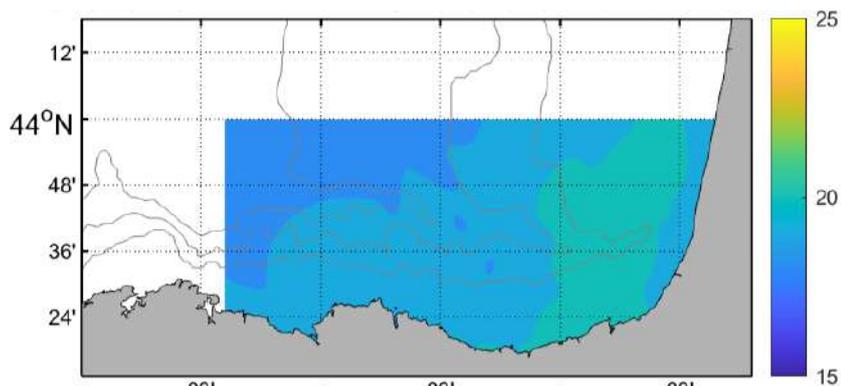


Figure 2.1 CROCO model spatial coverage and temperature forecast for 2021-07-08-01:00.

and ocean currents). The modelling strategy in this Pilot is to produce 96 hours forecast on a daily basis, using a high spatial resolution configuration of the CROCO model (Fig. 2.1) initiated and nested into the corresponding operational Copernicus Marine Service (CMEMS) IBI Forecasts available in the forecast window. For more details about the performance of the CMEMS IBI Forecast product, the reader is referred to Sotillo et al. (2015).

The high-resolution model is daily restarted, therefore the validation strategy will be to compare the forecast data for each day (i.e., 96 hours) with T, S and currents from one slope-moored buoy and surface currents from a High Frequency Radar (HFR, see section 2.2.1). In this way, the evolution of the quality of the modelled data can be analysed over 96 hours. In addition, since there is an interest to know the performance of the model for reproducing sea surface temperature (SST) fronts (aligned with the Front Detection Service Module), a process-oriented validation has been also carried out by using SST maps from moderate resolution satellite products (section 2.2.2).

Table 2.1 shows the details of the datasets used for the validation of the model. All the datasets were recorded for the period 2021-06-10 to 2021-09-04; being this period the one used for the validation. As stated previously, the results of the CROCO model will be validated with a forecast horizon of 96 hours.

In the following subsections, there is a brief description of the datasets and their corresponding platforms used for the validation.

Table 2.1 List of datasets used for model validation.

Parameter	Platform	Spatial resolution	Frequency	Source
Salinity, Temperature	A deep-water buoy	Profile data at fixed station (Figs. 2.2 and 2.3)	Hourly	Directorate of Emergency and Meteorological Services (DAEM), Basque Government
Surface currents	HFR	5 km X 5 km (Fig. 2.4)	Hourly	DAEM
Sea surface temperature	Satellite	0.02° x 0.02° (Fig. 2.5)	Daily	CMEMS

Slope Donostia buoy. Hourly TS and velocity data from the Donostia buoy will be used to validate the CROCO model. This buoy is located at a depth of approximately 570 m, at 15 nm to the north of San Sebastian (see Fig 2.2); with collection at Pasaia Port. It measures TS and current velocities at different levels, from 10 to 200 m depth. The buoy characteristics are:

- Model: FUGRO OCEANOR SEAWATCH WAVESCAN.

- Material: Polyethylene, aluminium, stainless steel..
- Total buoy height: 5.6 m.
- Buoy diameter: 2.8 m.
- Mast height (on water): 3.5 m.
- An underwater-instrumented line (Fig. 2.3) includes: ADCP current sensor, 6 CT sensors and 1 CTD SBE 37IM sensor, installed on an inductive cable (16mm galvanized with plastic mesh), 200 m long and an acoustic release (Benthos).
- Weight: buoy (approx. 930 kg) + instrumented line (approx. 350 kg) + bottom weight (1200 kg).

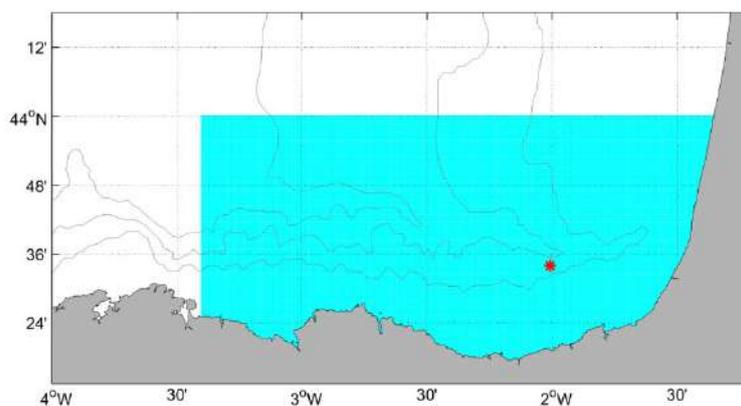


Figure 2.2 Location of the Donostia buoy (red star) over CROCO model coverage area (cyan).

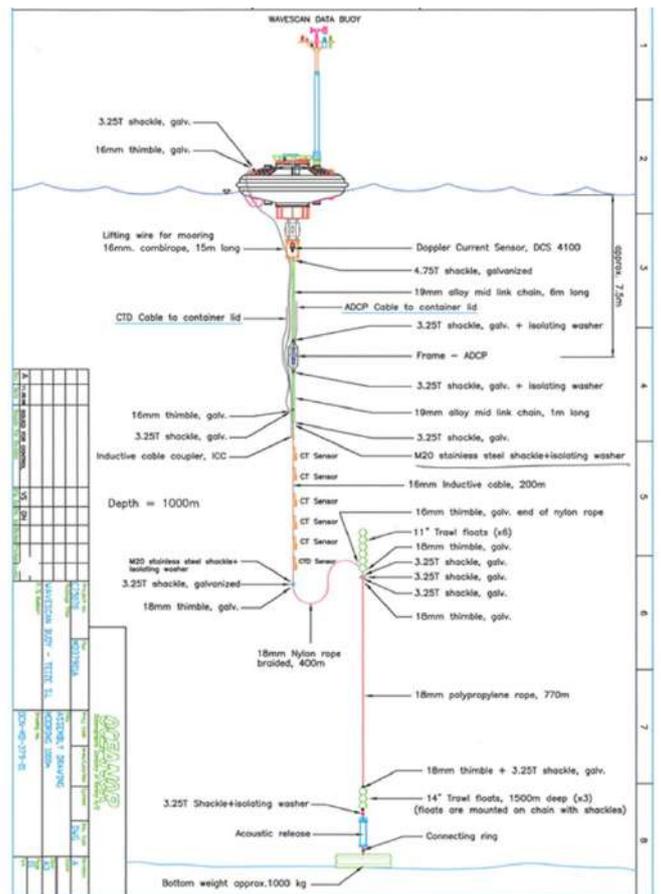


Figure 2.3 Diagram of the Donostia Buoy with instrumented line and mooring line.

In addition to the Donostia buoy currents comparison, data from the HFR installed on the Basque coast will be used to validate the surface currents velocities. The geographical coverage of EuskOOS HFR is shown in Fig. 2.4. The HFR antennas emit at a central frequency of 4.5 MHz and a 40 kHz bandwidth. The range coverage of radial data is 150 km, with 5 km radial resolution, and hourly total surface current vectors are obtained in a regular grid of 5 km x 5 km spatial resolution. For the model validation, daily mean surface current maps will be performed, in order to compare with the model output.

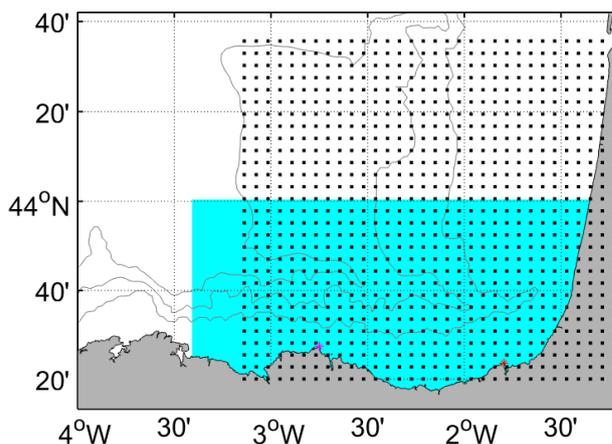


Figure 2.4 EuskOOS HFR coverage (small black squares) and its antennas location (Magenta and Red stars: Matxitxako and Higer stations respectively), as well as the coverage of the coastal model (cyan square).

Satellite-derived SST data was used to validate model SST gradients and fronts. Given that the south-eastern Bay of Biscay is a cloudy area, Level 4 dataset was used instead of Level 3. Specifically, the SST_EUR_PHY_L4_NRT_010_031 product of CMEMS was selected (see an example of this product below a model SST in Fig. 2.5), that has a spatial resolution of $0.02^\circ \times 0.02^\circ$ (CMEMS-SST-PUM-010-009-031, CMEMS-SST-QUID-010-031). Note that the Level 3 version of the product was checked to evaluate the reliability of the filling in the Level 4 product.

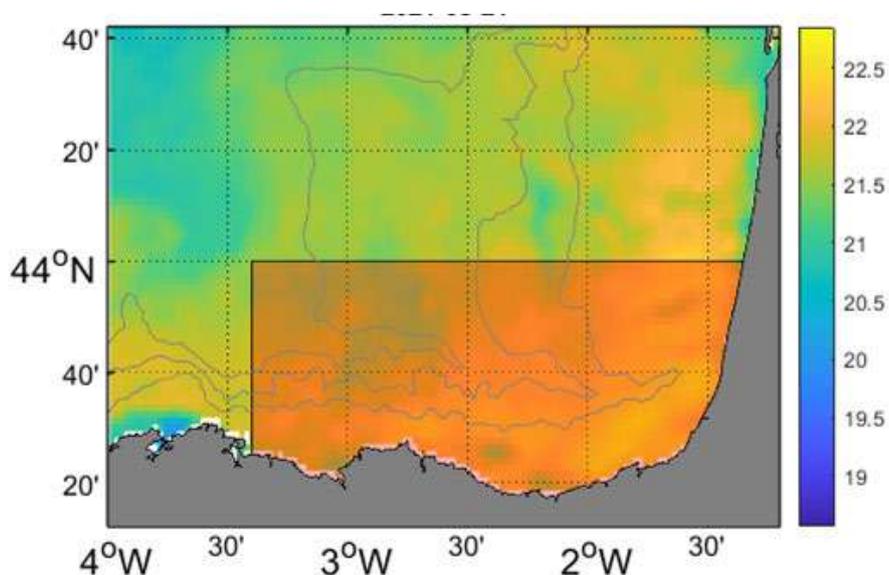


Figure 2.5 CROCO model coverage area (transparent red area inside the rectangle) over the SST ($^\circ\text{C}$) obtained from the SST_EUR_PHY_L4_NRT_010_031 product of CMEMS on August 21, 2021.

With regard to the metrics used for the validation strategy, a set of comparison metrics (listed below) were proposed in “Model validation strategy per pilot” section, in the D5-3 Deliverable”:

- Mean error: ME
- Root mean square error: RMSE
- Correlation Coefficient: CORR
- Adjusted Relative Mean Absolute Error (ARMAE)

2.2.1 Forecast validation

2.2.1.1 Comparison between model and slope buoy

The slope buoy of Donostia, is located at $2,007^\circ\text{W}-43,566^\circ\text{N}$ (Fig. 2.6) and at a depth of approximately 500 m. It includes TS measurements at 10, 20, 30, 50, 75, 100 and 200 m depths, as well as zonal and meridional current velocities at 12, 20, 36, 52, 100 and 204 m depths. Buoy (*in situ*) and model 4-day horizon data examples can be observed in Fig. 2.6, where model outputs have been interpolated to the same buoy depth positions.

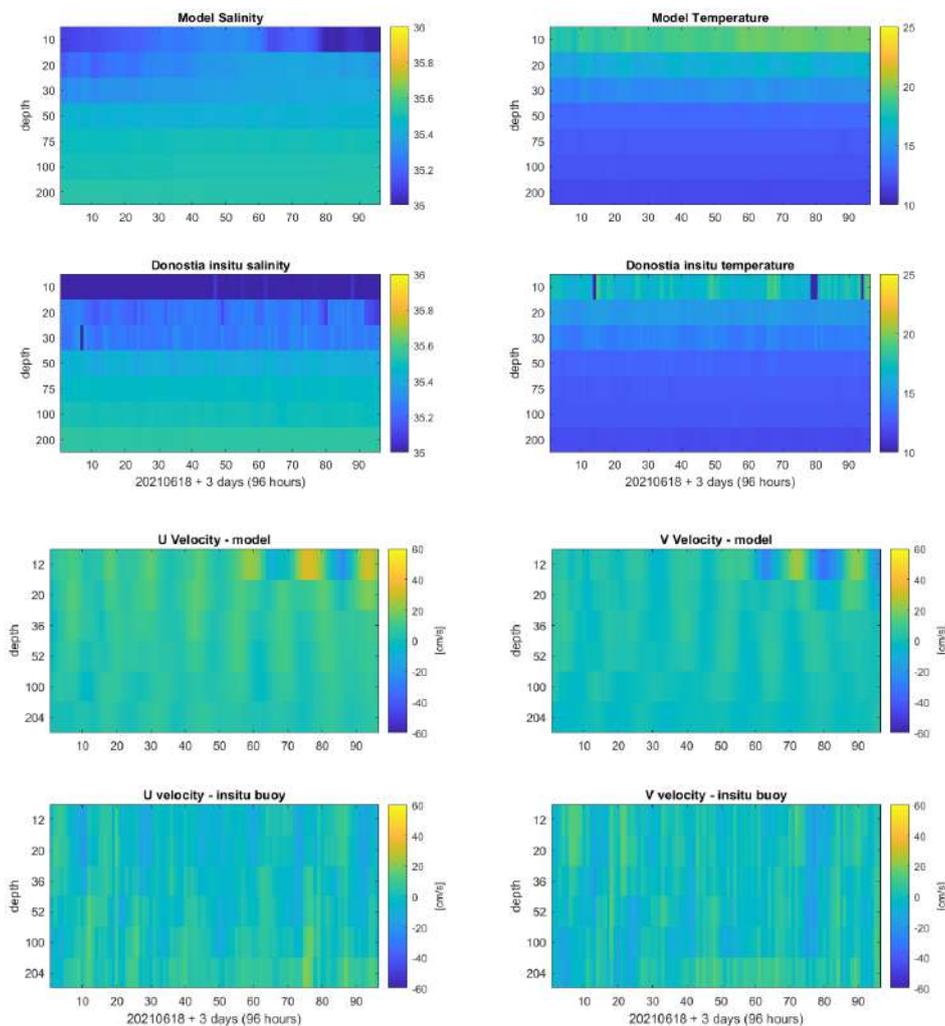


Figure 2.6 Model and in situ buoy TS (up) and U-V velocity (down) plots, for 2021-06-18 + 3 day (96 hours) horizon.

Daily mean values have been calculated and plotted for TS and velocity, at the *in situ* measured depths. Fig. 2.7 shows the daily mean of the first 24 h horizon values but 2nd, 3rd and 4th days' mean values have also been calculated, in order to analyse the detrimental effects with longer forecast horizon windows.

CORR and RMSE values between the model and *in situ* daily mean comparison analysis are summarized in Table 2.2, for different depths. The CORR and RMSE values vary significantly for the different variables and depths. But the values for the different horizon windows (first, second, third and fourth forecasted day) are similar for each variable. This means that the model capabilities for the 4 days forecasted window do not show significant quality decrease. With this in mind, we will focus on the first day forecast, to simplify the number of comparison analysis.

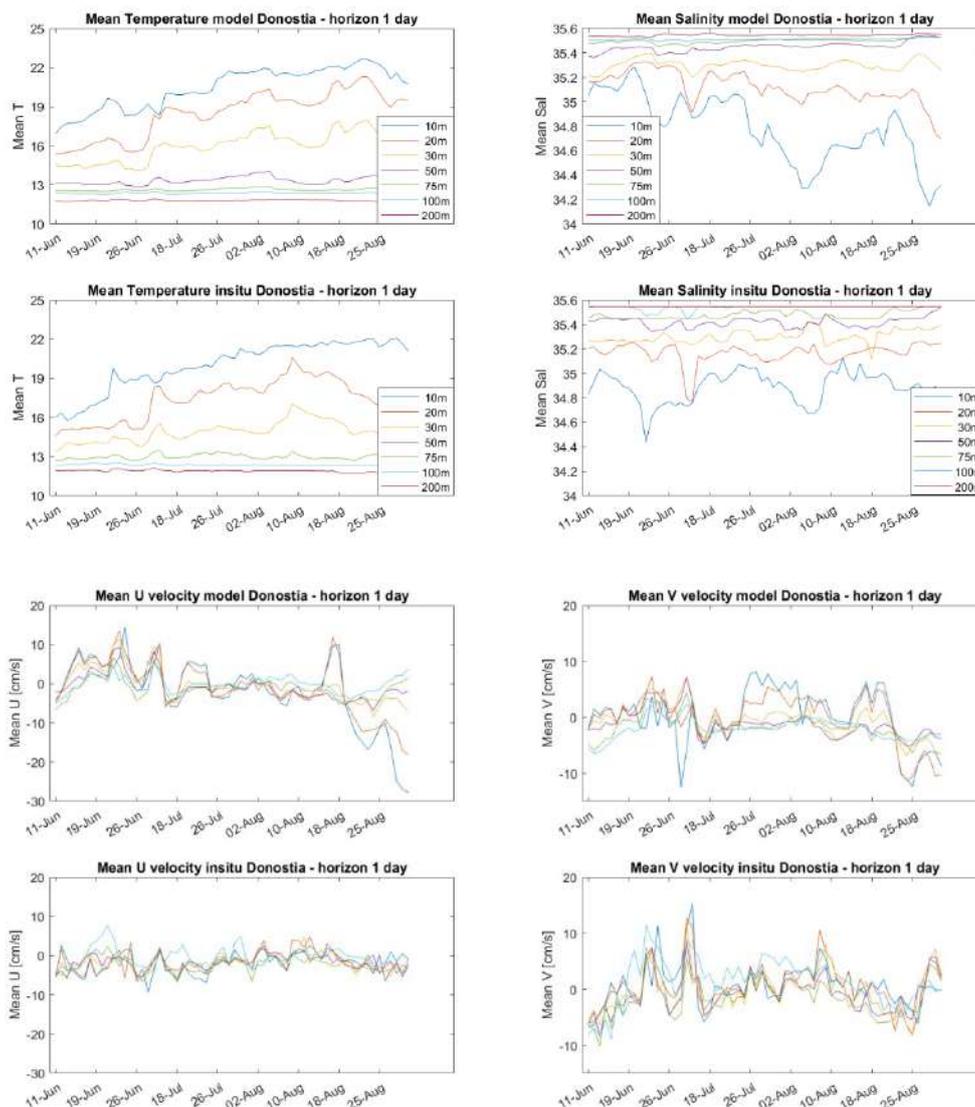


Figure 2.7 (Up) Model and in situ T, S 24-hour forecast average values, on a daily basis, from June to September. (down) Model and in situ U and V velocities 24-hour forecast average values, on a daily basis, from June to September.

Temperature correlation values vary from 0.93 to -0.09 for the different depths. Correlation results for the salinity range from 0.62 to 0.07. RMSE values for temperature are between 0.08 - 1.56°C , and 0.007 - 0.31 ppt for salinity. Although the results are not as good as expected, Fig. 2.7 (up) shows that the model is predicting the variability of both temperature and salinity, which is important for the fisheries sector. Model and *in situ* velocity correlation values also show low values [-0.13 , -0.81] m/s and high RMSE values, which vary between [2.58 , -4.5] m/s. These low values may be related to the high variability associated with local currents, due to local winds, local river flows, coastal affection, etc.

Table 2.2 RMSE and CORR summary table at different depth and horizon timelines.

	Depth (m)	CORR				RMSE			
		Day 1	Day 2	Day 3	Day 4	Day 1	Day 2	Day 3	Day 4
Temperature	10	0.936	0.930	0.921	0.914	0.794	0.799	0.776	0.692

	20	0.784	0.775	0.759	0.723	1.561	1.618	1.595	1.639
	30	0.558	0.535	0.517	0.492	1.414	1.471	1.470	1.455
	50	0.386	0.390	0.274	0.168	0.477	0.540	0.568	0.569
	75	0,319	0.304	0.214	0.139	0,392	0.376	0.388	0.401
	100	-0,095	-0.197	-0.234	-0.308	0.081	0.093	0.095	0.099
	200	0,591	0.637	0.637	0.652	0.103	0.091	0.098	0.107
Salinity	10	0,119	0.203	0.259	0.252	0.311	0.325	0.318	0.326
	20	0,094	0.144	0.108	-0.007	0.155	0.172	0.183	0.193
	30	0,079	-0.046	-0.139	-0.188	0.067	0.076	0.080	0.087
	50	0,443	0.458	0.496	0.528	0.054	0.055	0.057	0.057
	75	0,623	0.629	0.623	0.702	0.033	0.034	0.035	0.035
	100	0,445	0.385	0.393	0.414	0.033	0.033	0.032	0.033
	200	-	-	-	-	0.007	0.008	0.008	0.008
E-W velocity	12	-0,136	-0.160	-0.108	-0.118	3.244	3.206	3.183	3.037
	20	0,069	0.063	0.065	0.099	3.137	3.112	4.053	3.062
	36	0,240	0.193	0.229	0.212	2.756	2.678	2.613	2.448
	52	0,311	0.175	0.193	0.072	2.749	2.669	2.604	2.533
	100	0,498	0.355	0.339	0.269	3.017	2.895	2.831	2.706
	204	0,199	0.125	0.174	0.117	2.581	2.490	2.480	2.448
N-S Velocity	12	0,202	0.198	0.201	0.123	4.021	4.416	4.281	4.183
	20	0,314	0.284	0.274	0.334	4.140	4.644	4.582	4.563
	36	0,251	0.215	0.134	0.237	3.804	3.684	3.677	3.636
	52	0,390	0.268	0.209	0.284	3.517	3.439	3.251	3.313
	100	0,592	0.485	0.471	0.502	3.667	3.573	3.306	3.375
	204	0,816	0.718	0.714	0.720	4.507	4.305	4.340	4.276

2.2.1.2 Comparison between model and High Frequency Radar data

Part of the area covered by CROCO model overlaps with the HFR footprint (see Fig. 2.8). HFR system is providing hourly sea surface current velocity maps. These maps offer a great opportunity for the CROCO validation.

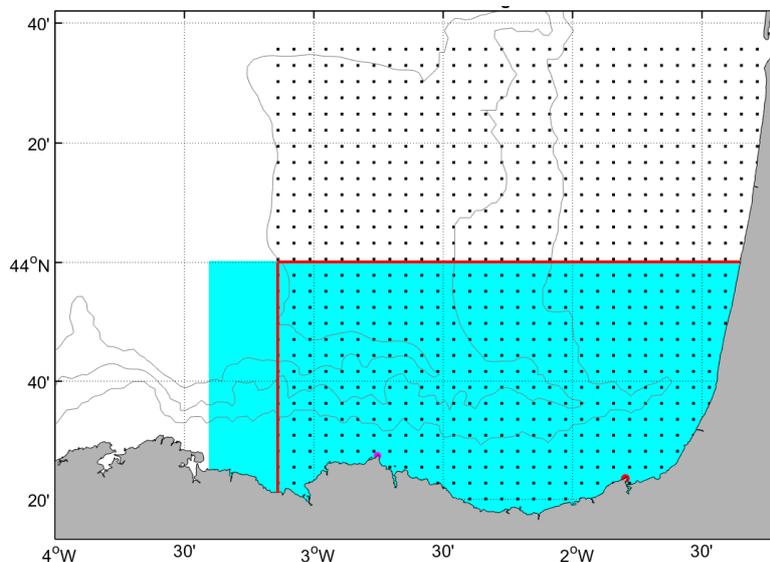
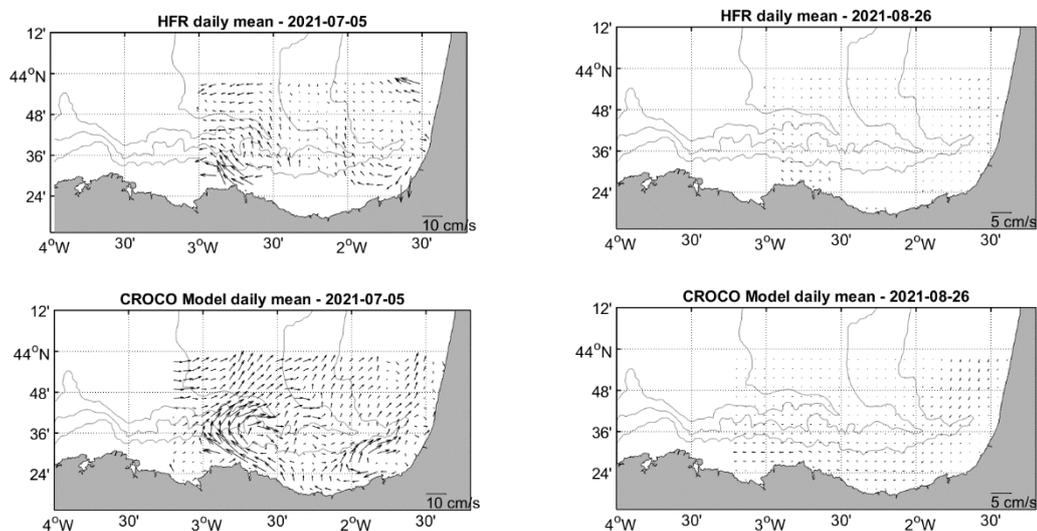


Figure 2.8 EuskOOS HFR coverage (small black squares) and its antennas location (Magenta and Red filled dots: Matxitxako and Higer stations respectively). The red line indicates the limits of the common HFR - model spatial comparison area.

In the previous section, low correlation values have been observed between the model and the buoy currents, which are probably associated with the high variability of the surface currents for point-to-point comparisons. For this reason, daily mean surface current maps have been performed, from both HFR and CROCO model for the comparison (Fig. 2.9).



With regard to the metrics used for the 2D surface currents validation strategy, Mean Error 2D Maps and 2D RMSE maps have been calculated (Figs. 2.10 and 2.11, respectively), following the metrics proposed in “Model validation strategy per pilot” section, in the D5-3 Deliverable”, to see the 2D capabilities of CROCO model. In these figures, there are comparison points without values. This is because in those comparison points, there is no good HFR data, even if it is the footprint of the available dataset grid.

The error maps show that the absolute error values are up to 0,2 m/s, presenting the highest errors in the zonal velocity component, especially in the meridional of the comparison area. The central part of the comparison area shows error values around 0.05 m/s, which are not bad results, taking in account the existing high variability of the surface currents in the study area.

It is worth it to note the capability of the model to reproduce both small velocity and high velocity

Figure 2.9 Daily surface currents from (up) EuskOOS HFR system and (down) CROCO model forecast. Left figures are from 2021-July-28 and right figures are from 2021-July-05. The model was capable of reproducing small currents event on August 26 while it forecasted higher values on July 5. Although the results for the 2, 3 and 4 day horizon forecast are not shown, the capabilities of the model are similar (with an insignificant time degradation) as occurred with the Donostia buoy 1 point comparisons.

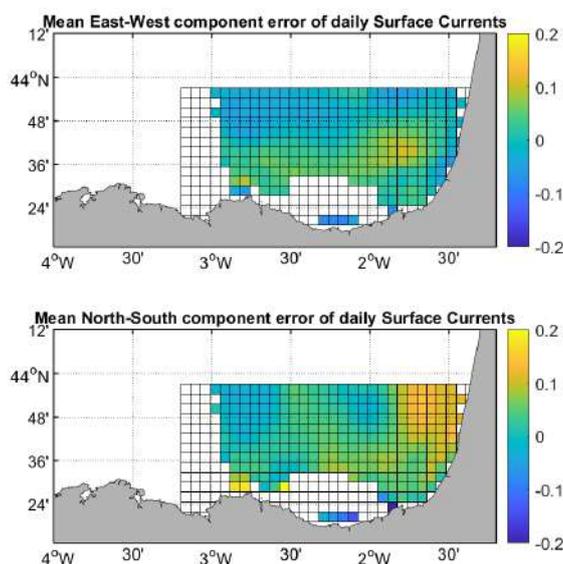


Figure 2.10 East-West (up) and Sount-North (down) surface velocity components ME (m/s) for a 24 hours forecast.

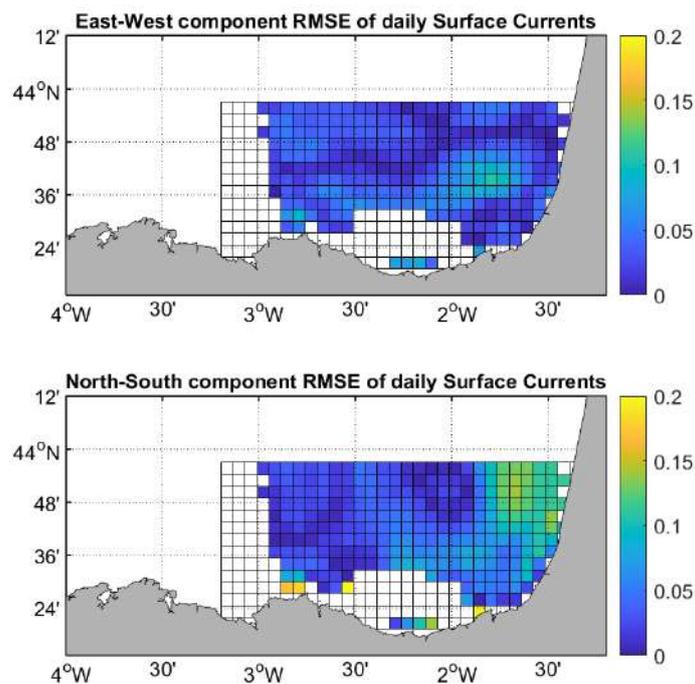


Figure 2.11 RMSE values of zonal (up) and meridional (down) daily surface velocity components, for the 24 hours forecast.

2.2.2 Process-oriented validation

Daily SST data were used to estimate and then compare gradients and fronts between model and observations. These fronts were estimated from SST gradients for estimating fronts based on the work of Cayula and Cornillon (1992). A daily temporal resolution was considered due to the daily observations provided by the SST_EUR_PHY_L4_NRT_010_031 product of CMEMS that we used, therefore, the model outputs were daily averaged. Given that the model capabilities do not change much for the forecasted 4 days (as shown before), only the first day forecast data were considered. These comparisons allowed assessing the skill of the model to predict fronts since these are areas of higher fishing probability.

A series from 10-06-2021 to 04-09-2021 was analyzed from where simultaneous model and observational data were available for 87 days. The assessment was carried out for dates where observations covered almost the entire study area. Due to cloudy conditions, 14 days out of 87 were selected for the comparisons.

Concerning the SST gradients, RMSE and ARMAE maps were computed (Fig. 2.12). It is observed that the biggest discrepancy areas are the Adour river mouth and the area between the Urdaibai and Nervion river mouths probably induced by the accuracy of the river discharge data used in the model. Also, an area with the form of the coastline in the eastern part of the study area shows high RMSD and ARMAE values.

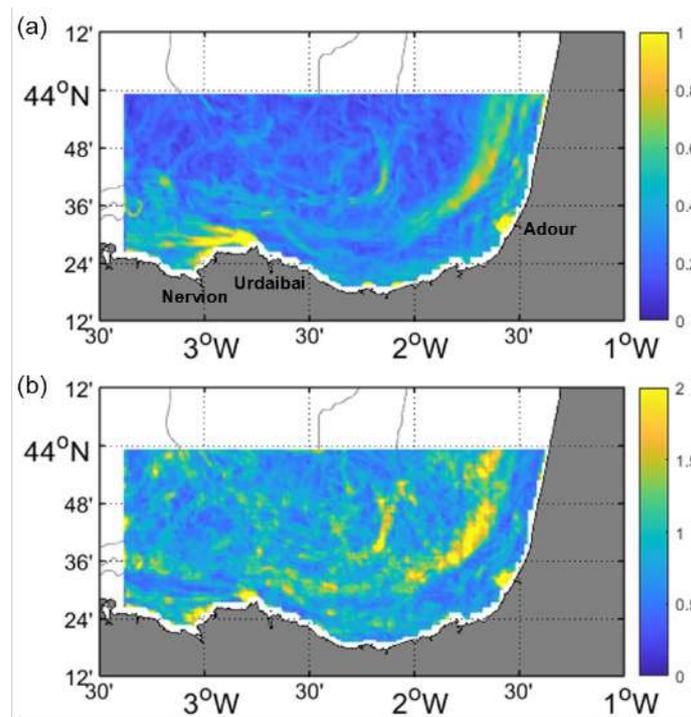


Figure 2.12 RMSE (a) and ARMAE (b) maps between observed and modelled daily SST gradients.

In order to assess the front detection skill of the model, a skill score was computed. The skill score algorithm is as follows:

- 1) For each point where there is an observed front (hereinafter “original point”) the number of points where there are fronts nearby are counted for both observations (OBS) and model (MOD) inside different rings of distance (from the original point).
- 2) For each ring a skill score value (S) is estimated: $S=1-D/(Max+1)$, where:
 - $D=|number\ of\ points\ with\ fronts\ for\ OBS - number\ of\ points\ with\ fronts\ for\ MOD|$
 - $Max=$ the maximum number of points between OBS and MOD
 - The “+1” is to avoid that when the MOD number is 0, S to be 0 whatever is the value of D
- 3) This procedure is applied to all the points with observed fronts and the values for each ring are saved.
- 4) Then the value obtained for each ring is averaged and a weight is allocated to each mean value which is smaller as we get further from the original point (the weights’ summation is 1 in order not to bias the results).
- 5) Finally, all the values are summed, obtaining a final score for each map.

Note that the more fronts the higher the probability of finding front points and therefore for obtaining higher S values. To penalize this effect, we multiply the score times the percentage of points without fronts. The maximum value of the skill score is 1 whereas the minimum is 0. We tested different configurations of rings and weights (see Tables 2.3-2.7) to see which one provided the most representative results. The general pattern from one date to another is the same regardless of the

configuration with maximum values on 19-20 July and minimums on 25 June (Fig. 2.13). The maximum range obtained is of 0.15, and though the values do not change very much the highest (lowest) ones correspond to dates where highest (lowest) qualitative agreement is observed (Fig. 2.14). For example, this can be observed if the maps and score values of 19 July and 25 June are compared. There are also big areas where there are not observed nor detected fronts, thus showing a good performance of the model.

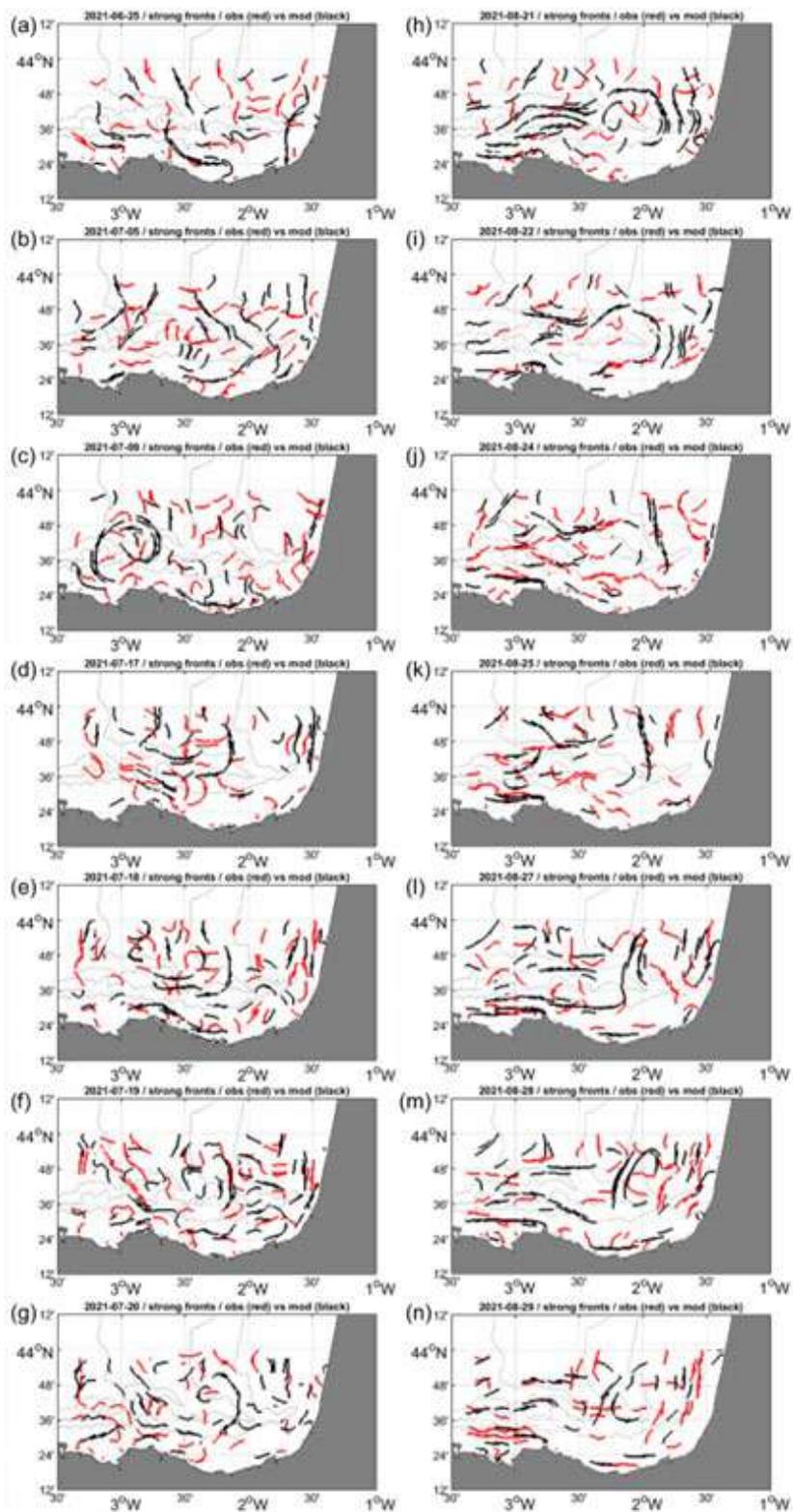


Figure 2.13 Maps of the fronts observed (in red) and modelled (black) for 14 days where observations were almost cloud free.

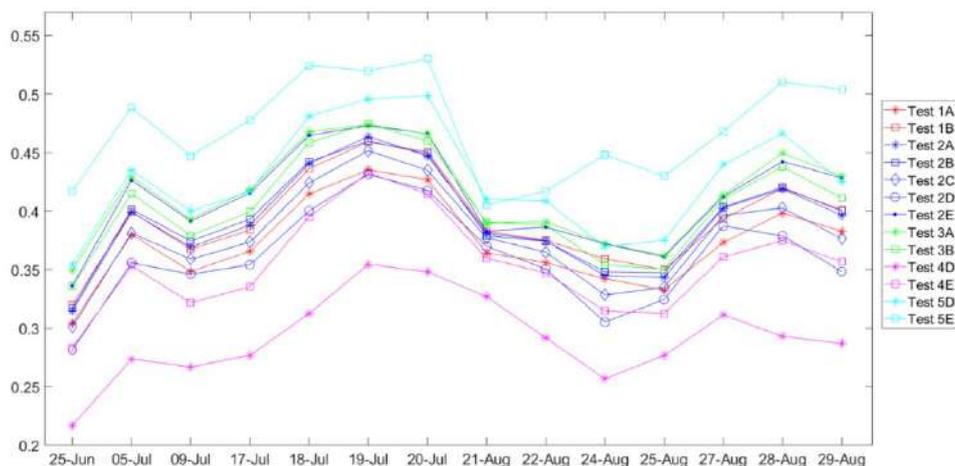


Figure 2.14 The different score values corresponding to each test are depicted for the analyzed 14 days. The radius and weights of each test are depicted in Tables 2.3-2.7

Table 2.3 Configuration of Test 1, with different numbers and distances of rings and with different weights allocated to each ring.

Test 1		
Radius (Km)	Weights	
	A	B
3	0.3	0.3
6	0.25	0.3
9	0.2	0.2
12	0.15	0.15
15	0.1	0.1

Table 2.4 Configuration of Test 2, with different numbers and distances of rings and with different weights allocated to each ring.

Test 2					
Radius (Km)	Weights				
	A	B	C	D	E
5	0.5	0.5	0.6	0.75	0.37
10	0.35	0.3	0.3	0.2	0.33
15	0.15	0.2	0.1	0.05	0.3

Table 2.5 Configuration of Test 3, with different numbers and distances of rings and with different weights allocated to each ring.

Test 3		
Radius (Km)	Weights	
	A	B
6	0.4	0.5
10	0.3	0.35

13	0.2	0.1
15	0.1	0.05

Table 2.6 Configuration of Test 4, with different numbers and distances of rings and with different weights allocated to each ring.

Test 4		
Radius (Km)	Weights	
	A	B
3	0.75	0.37
6	0.2	0.33
10	0.05	0.3

Table 2.7 Configuration of Test 5, with different numbers and distances of rings and with different weights allocated to each ring.

Test 5		
Radius (Km)	Weights	
	A	B
7	0.75	0.37
14	0.2	0.33
20	0.05	0.3

Another metric, more intuitive than the previous skill score, to assess the performance of the model is to measure the distance between the modelled front points and the nearest observed points. Thus, for each date (i.e. for each map in Fig. 2.13), all the modelled points are considered and the mean nearest distance to an observed front point is estimated. The shorter the distance the better the performance. Moreover, the standard deviation (STD) of this mean provides information about the dispersion of the mean, that is, the lower the STD the more representative the mean distance is. If we sum both quantities, we obtain a distance (SUM) where it is probable to find a real front point from a modelled one, thus setting a relatively high probability radius for each date. The results for the 14 studied periods (Table 2.8) show a good agreement with the results of the skill score, with a high distance value for 2021-06-25 when the skill score is the lowest and a low distance for 2021-07-19 when the skill score is the highest. In general, values between 5-11 km are obtained and most of the observed front points are located at shorter distances showing a fairly good performance of the model for finding areas of higher fishing probability.

Table 2.8 Mean distance, its STD and their sum (SUM) between each modelled front point and the nearest observed point for each date. Also, the percentage of the cases where the nearest observed point falls within the radii defined by SUM is shown.

Date	Mean (km)	STD (km)	SUM (km)	% within SUM
2021-06-25	6.10	4.04	10.14	80
2021-07-05	5.03	3.91	8.94	84
2021-07-09	5.31	3.94	9.25	80
2021-07-17	5.15	4.08	9.23	87

2021-07-18	4.2	3.14	7.34	85
2021-07-19	3.29	2.48	5.77	84
2021-07-20	4.14	3.04	7.18	81
2021-08-21	4.55	3.97	8.51	86
2021-08-22	4.94	4.29	9.24	82
2021-08-24	5.55	4.45	10.00	85
2021-08-25	5.48	4.90	10.38	81
2021-08-27	3.90	3.40	7.31	83
2021-08-28	3.97	2.90	6.87	85
2021-08-29	4.10	3.01	7.11	82

2.3 Pilot 3: Bulgaria

The Fishing Suitability Index developed within this service module relies on sea surface temperature and sea surface salinity measurements from the BLKSEA_MULTIYEAR_PHY_007_004 (https://resources.marine.copernicus.eu/product-detail/BLKSEA_MULTIYEAR_PHY_007_004/DOCUMENTATION) product available through Copernicus Marine Service (CMEMS). Quality information and validation activities for the BLKSEA_MULTIYEAR_PHY_007_004 product can be found in the following document: <https://catalogue.marine.copernicus.eu/documents/QUID/CMEMS-BS-QUID-007-004.pdf>. In addition, a coastal, high-resolution wave model for the Western Black Sea downscaled from the CMEMS wave product is implemented and validated here.

2.3.1 Hindcast validation

In this section, we assess the quality of the western Black Sea (WBS) wave hindcast system for the Pilot 3. The wave data that are included are generated using the WAM Cycle 6 Black Sea model (Staneva et al., 2020), which includes the most recent wave model developments. The CMEMS wave data are available from CMEMS catalogue under the BLKSEA_NRT_WAV_007_003. Here however the model runs in a shallow water mode in the domain shown in Fig. 3.1 (black rectangle), with a spatial resolution of about 0.9 km, ($1/120^\circ$ in latitude and longitude direction). The western Black Sea model area and the corresponding bathymetry are shown in Fig. 3.1. WAM calculates the two-dimensional energy density spectrum at each of the ~ 200 k active model grid points in the frequency and directional space. The solution of the energy balance equation is provided for 24 directional bands at 15° each, starting at 7.5° and measured clockwise with respect to true north, and 30 frequencies logarithmically spaced from 0.042 Hz to 0.66 Hz at intervals of $\Delta f/f = 0.1$. Therefore, the prognostic part of the wave model covers periods from approximately 25 to 1.5 seconds. In order to include the important contribution of higher frequency waves to wave growth/dissipation processes and for the output wave characteristics a parametric tail is fitted for frequencies above the spectral maximum.

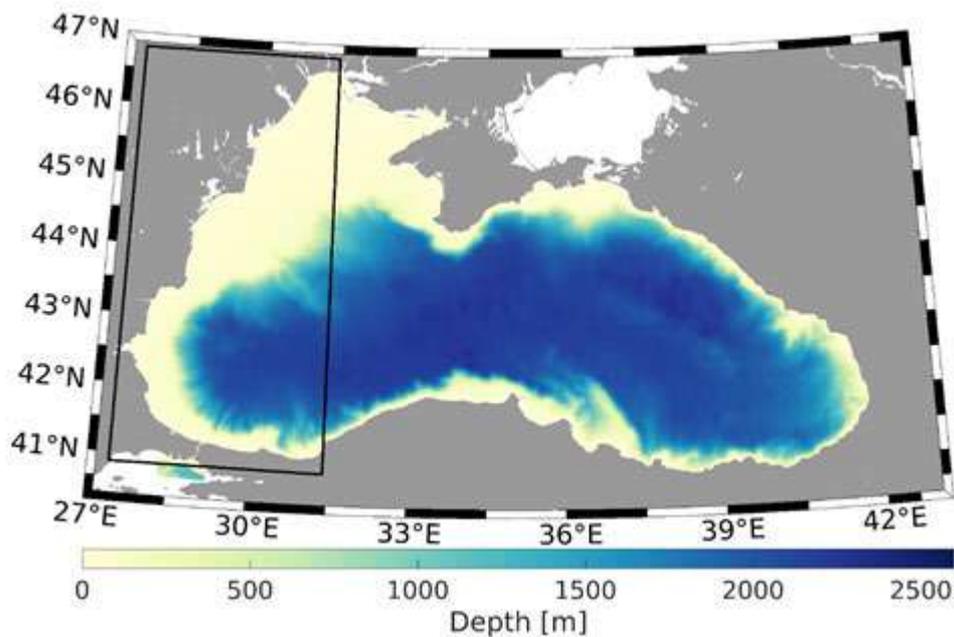


Figure 3.1 Black Sea bathymetry and the WBS WAM wave model nest (black square).

In order to evaluate and assure the quality of the WBS product, the system has been integrated into hindcast mode for the period 01/10/2021 - today. All CMEMS satellite measurements that are available for the entire two-year period (Sentinel-3a, Sentinel-3b, Cryosat-2, SARAL/Altika, Jason-3, CFOSat, and HaiYang-2b; product WAVE_GLO_WAV_L3_SWH_NRT_OBSERVATIONS_014_001) have been used to compare the significant wave height with the corresponding wave model results. As a precondition to enable these comparisons, the satellite data has to be correlated with the wave model data in space and time. The satellites need up to two minutes to cross the Black Sea, and the measurements recorded by the radar altimeter have been compared with the computed results of the nearest model output time. For each of the individual measurements with its unambiguous assignment to longitude and latitude, the computed values of the nearest model grid point in space have been used to compare with.

Since the radar altimeter of the satellites measures wind speed (available from February 2020) and significant wave height, the only integrated wave parameter that can be used for validation is the significant wave height (SWH). The measured data undergoes quality control to make sure that unrealistic values are not taken into account. Such values can occur when the satellite passes the transition zone between land and sea at the coasts. Usually, the satellites pass the Black Sea once a day, sometimes twice.

Although in-situ wave measurements from moored wave buoys are available from CMEMS In Situ Thematic Assemble Centre (CMEMS INS TAC) (product INSITU_BS_NRT_OBSERVATIONS_013_034) for the Black Sea, their locations are restricted to coastal areas of the southwestern basin. The corresponding water depths are mostly 17 m or 20 m and in the model domain, these positions are located at the land-sea boundary or only a few grids in the ocean. As the wave parameters are very

sensitive to the water depth, the buoy positions in the model were slightly shifted (1-2 grids) in order to place them at a more realistic water depth. Time series at these grids have been used for validation.

Statistical analysis

Detailed statistics following the CMEMS PQWG-Waves recommendations have been calculated for all comparisons between modelled and measured data recorded by the radar altimeter of the different satellites.

For the quarters of the considered period (June 2019 – May 2021), the analysis for the significant wave heights is presented as a QQ (Quantile-Quantile) -scatter plot including statistical parameters (Fig. 3.3). These include the RMSD, bias, Scatter Index (SI), Pearson correlation coefficient (CORR), and best-fit Slope (SLOPE). The SI, defined here as the standard deviation of errors (model minus observations) relative to the observed mean of the significant wave, being dimensionless, is more appropriate to evaluate the relative closeness of the model output to the observations at different locations compared with the RMSD, which is representative of the size of a ‘typical’ error. The SLOPE corresponds to a best-fit line forced through the origin (zero intercept). In addition to these core metrics, merged Density Scatter and Quantile-Quantile (QQ) plots are provided.

In general, the statistics show a slight underestimation of the measured data (“R”) by the wave model results (“M”) occurring consistently among all satellites (bias ranges between -18.9 and -9.7 cm). The CORR is always above 0.92. The bias and RMSD over all satellites are -15.8 cm and 26.8 cm, respectively. This model performance of the wave setup is considered good.

The validations against the available in-situ observations show that the model is capable of simulating the significant wave height near the coastal areas. This is especially a challenge for the wave modelling and substantiates the need for higher resolution data near the coastal areas. The bias is very low (ca. 1.5 cm), the correlation is 0.9 and the RMSD is 17 cm.

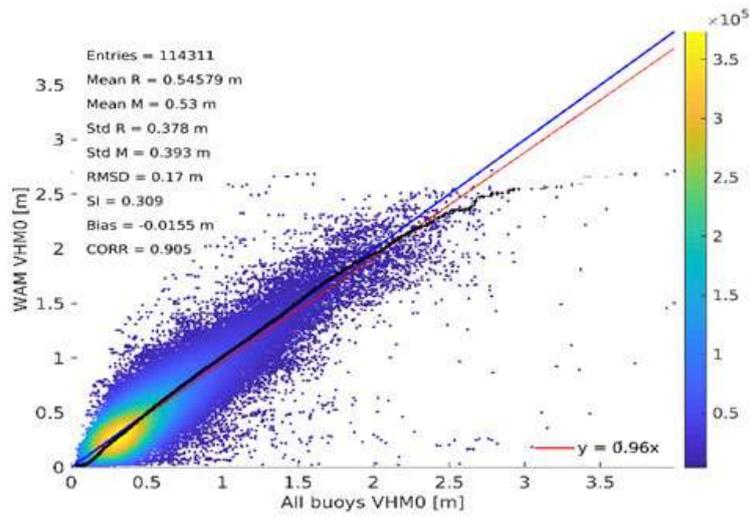


Figure 3.2 Scatter plots of significant wave height (H_s) of all in-situ data available. Also shown are the estimated bivariate probability density (coloured area), the linear slope-fit regression of modelled and observed wave heights (red line), specific quantiles taken from the empirical cumulative density function (black line), and the diagonal (blue line). Furthermore, summary statistics and skill scores are included. R: reference (satellite) data, M: model data.

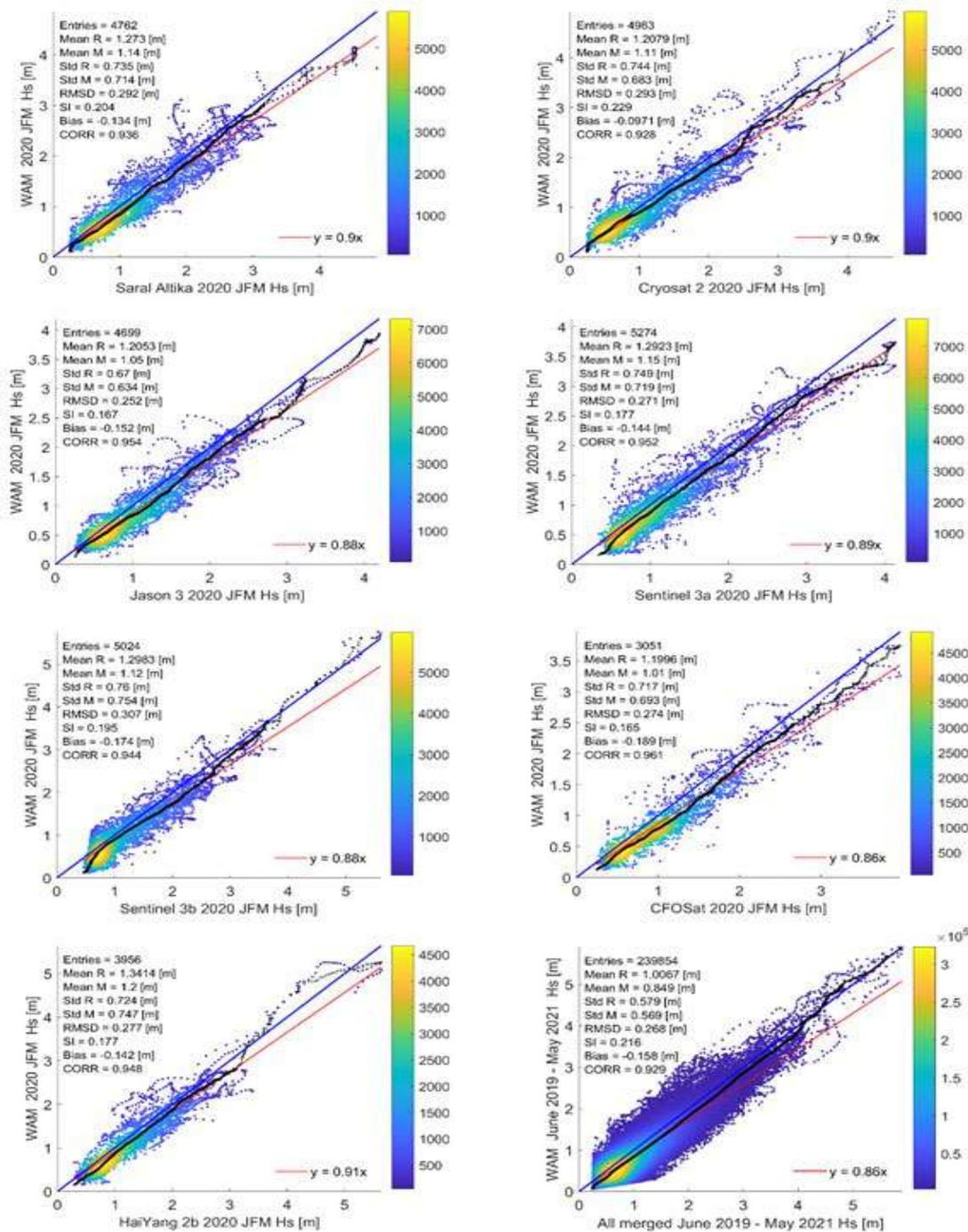


Figure 3.3 Scatter plots of significant wave height (Hs) for Q1 (JFM) 2020 of 7 different satellites and for the whole period using all satellites merged (lower right). See the labels for the satellite names. Also shown are the estimated bivariate probability density (coloured area), the linear slope-fit regression of modelled and observed wave heights (red line), specific quantiles taken from the empirical cumulative density function (black line), and the diagonal (blue line). Furthermore, summary statistics and skill scores are included. R: reference (satellite) data, M: model data.



Figure 4.2 Map of the different stations close to the Belgian coast used in the validation.

The data itself are provided by different organisms, depending on the station locations and the parameter. The table 4.1 lists the stations, parameter codes and providers. The reader should be aware that when comparing wave height frequencies, the models compute them by a spectrum analysis and the observations attain them by a wave analysis.

Table 4.1 List of data providers per stations/parameter

Station	Param.	Provider
Aberdeen	SLEV	National Oceanography Centre (United_Kingdom)
Akkaert	VTZA	Agency for Maritime Services and Coast Department Coast (MDK)
Barmouth	SLEV	National Oceanography Centre (United_Kingdom)
Bournemouth	SLEV	National Oceanography Centre (United_Kingdom)
Cadzand	SLEV	Rijkswaterstaat Water Traffic and Environment
	VHMO	Rijkswaterstaat Water Traffic and Environment
Cuxhaven	SLEV	Waterways and Shipping Office Cuxhaven
Den Helder	SLEV	Rijkswaterstaat Water Traffic and Environment
	TEMP	Rijkswaterstaat Water Traffic and Environment
Europlatform	VAVH	Rijkswaterstaat Water Traffic and Environment
	VHMO	Rijkswaterstaat Water Traffic and Environment
	VTZA	Rijkswaterstaat Water Traffic and Environment
	TEMP	Rijkswaterstaat Water Traffic and Environment
	SLEV	Rijkswaterstaat Water Traffic and Environment
Helgoland	PSAL	Alfred-Wegener-Institute for Polar- and Marine Research (AWI)

	TEMP	Fed. Maritime and Hydrographic Agency - Dept. Oceanography
	SLEV	Fed. Maritime and Hydrographic Agency - Dept. Oceanography
Hoek van Holland	TEMP	Rijkswaterstaat Water Traffic and Environment
	SLEV	Rijkswaterstaat Water Traffic and Environment
IJmuiden	TEMP	Rijkswaterstaat Water Traffic and Environment
	SLEV	Rijkswaterstaat Water Traffic and Environment
K13	VAVH	Rijkswaterstaat Water Traffic and Environment
	VTZA	Rijkswaterstaat Water Traffic and Environment
Newport	SLEV	Agency for Maritime Services and Coast Department Coast (MDK)
Ostend	SLEV	Agency for Maritime Services and Coast Department Coast (MDK)
Roompot buiten	SLEV	Rijkswaterstaat Water Traffic and Environment
Stavanger	SLEV	Norwegian Hydrographic Service
Vlakte van de Raan	PSAL	Rijkswaterstaat Water Traffic and Environment
	SLEV	Agency for Maritime Services and Coast Department Coast (MDK)
Vlissingen	SLEV	Rijkswaterstaat Water Traffic and Environment
Wandelaar	VAVH	Agency for Maritime Services and Coast Department Coast (MDK)
	VTZA	Agency for Maritime Services and Coast Department Coast (MDK)
Westhinder	VTZA	Agency for Maritime Services and Coast Department Coast (MDK)
Zeebruges	SLEV	Agency for Maritime Services and Coast Department Coast (MDK)

Depending on the parameter, the validation starts in 2013, or 2014 up to December 2017. The data time series are not always continuous: there can be gaps in either the observations or optos data. Only matching date-times are taken into account.

2.4.1 Hindcast validation

2.4.1.1 Sea level

There are too many station/model combinations to examine all of them in detail. The general statistics of the model/station pairs are provided below. The skill metrics used for validation are the ratio of the standard deviations, the correlation, the model bias and RMSE. Validation was executed on a monthly basis and the average and quantiles Q25 and Q75 were computed for all stations listed in table 4.1

CSM. The average and quantiles Q25 and Q75 computed on all stations for the 4 monthly validation metrics are shown in Figure 4.3.



Figure 4.3 Summary statistics of validation metrics per stations, comparing the sea level CSM forecasts to observations. The stations Aberdeen, Barmouth and Bournemouth have only three data points, and are removed from the two last graphs since they strongly deviate from other points.

Figure 4.3 strikes by the consistency of the distribution for each station, the range Q75-Q25 per metric is very narrow and not visible for three of the metrics. The ranges for the Europlatform station is given in Table 4.2, for confirmation purposes. Figure 4.3.a shows that the model struggles to capture the variability of the observation at some stations. On the other hand, the correlation is very strong except for Bournemouth (three data points) and Stavanger (at the edge of the domain). The bias demonstrates a wider distribution, but the quantiles Q25 and Q75 remain globally in the range [-0.1 m, 0.1 m]. The RMSE is very stable across stations, except for the outlier Cuxhaven.

Table 4.2 Statistics of the validation metrics from the comparison between observations at the Europlatform station and the CSM forecasts.

CSM at Europlatform				
Metric	Ratio	Correlation	Bias	RMSE
Q25	0.984	0.965	-0.003	0.132
Average	1.00	0.972	0.043	0.167
Q75	1.01	0.981	0.100	0.193

NOS. The average and quantiles Q25 and Q75 computed on all stations for the 4 monthly validation metrics are shown in Figure 4.4. Figure 4.4 shows the same consistency as the figure 4.3 for the CSM validation. The ratio of the standard deviations illustrates that NOS tends to overestimate the variability of the sea level along the Dutch coast. On the other hand, the offshore Europlatform station and the stations along the Belgian coast have a good variability.

The correlation is very strong. The weakest performance is at IJmuiden, which does significantly less well than the neighbouring stations Hoek van Holland and Den Helder. The lower correlation at the station IJmuiden is linked to a shift in the phase prediction of the tide. This is apparent in Figure 4.4: the x-y scatter plot is broader at IJmuiden than at Den Helder. This is also confirmed by the phase analysis: the phase forecast can shift above one hour at IJmuiden.

Figure 4.4.c shows that the distribution of the bias in NOS, just as for the csm model, is wider than the other metrics. The bias is globally positive and the Q25; Q75 cover the range [-0.02; 0.20]. The average bias of stations Cuxhaven⁴⁵ and Den Helder⁴⁶ are significantly different from all other stations. Interestingly, the RMSE of Cuxhaven is very high: this indicates that the bias varies largely around zero. At all the other stations, the RMSE remains below 0.4 m. The analysis of the 4 metrics allows to conclude that NOS performs best at the stations close to the Belgian coast (Cadzand, Newport, Ostend,

Vlakte van de Raan, Vlissingen and Zeebrugge), with a ratio above 0.9, correlation above 0.97, moderate bias below 0.1 and RMSE below 0.3.

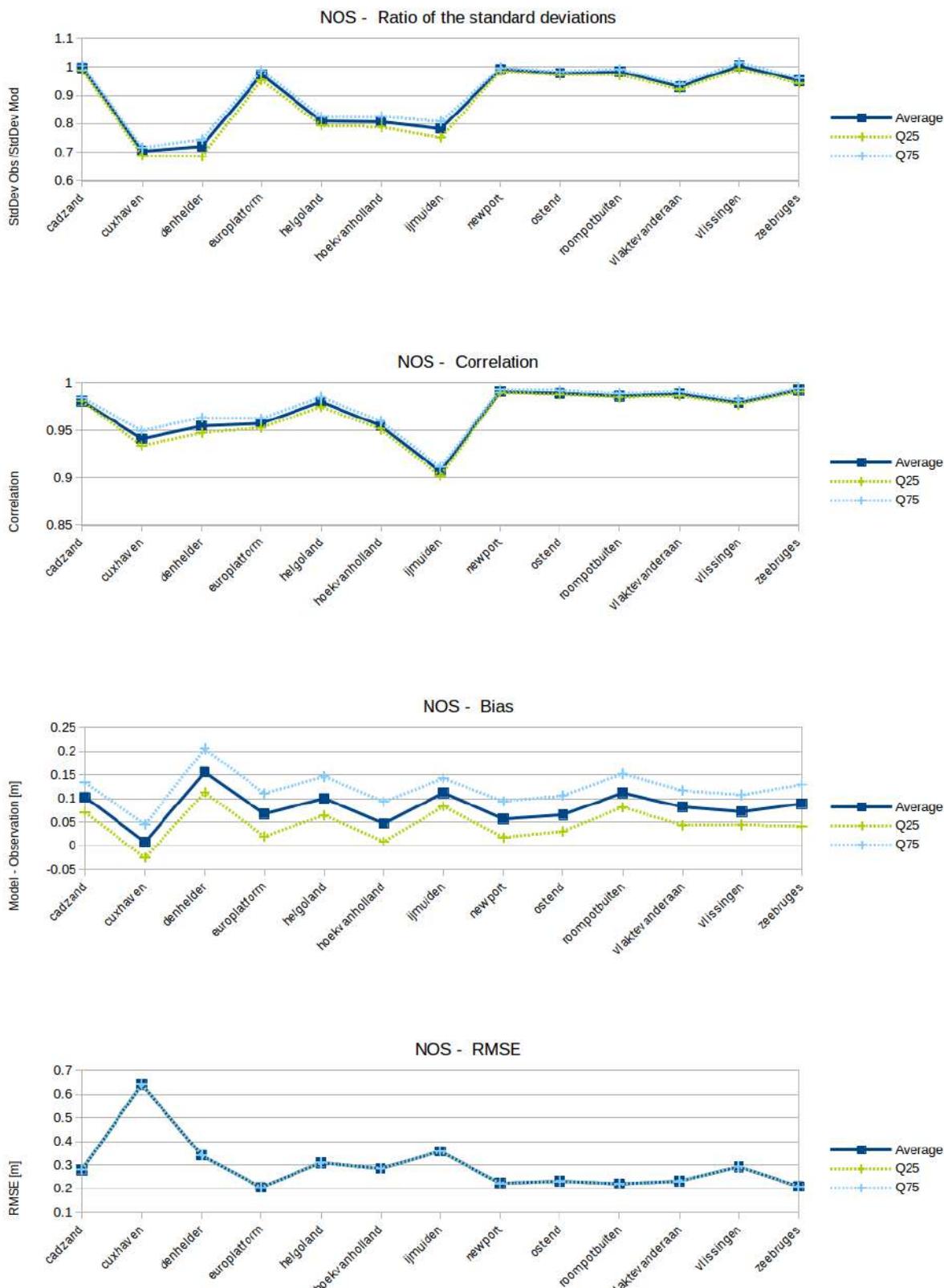


Figure 4.4 Summary statistics of validation metrics per stations, comparing the sea level NOS forecasts to observations. From top to bottom: (1) Ratio of the standard deviation of the observations by the standard deviation of the forecasts; (2) Correlation; (3) Bias; (4) RMSE

BCZ. The results of the validation metrics are illustrated in Figure 4.5.

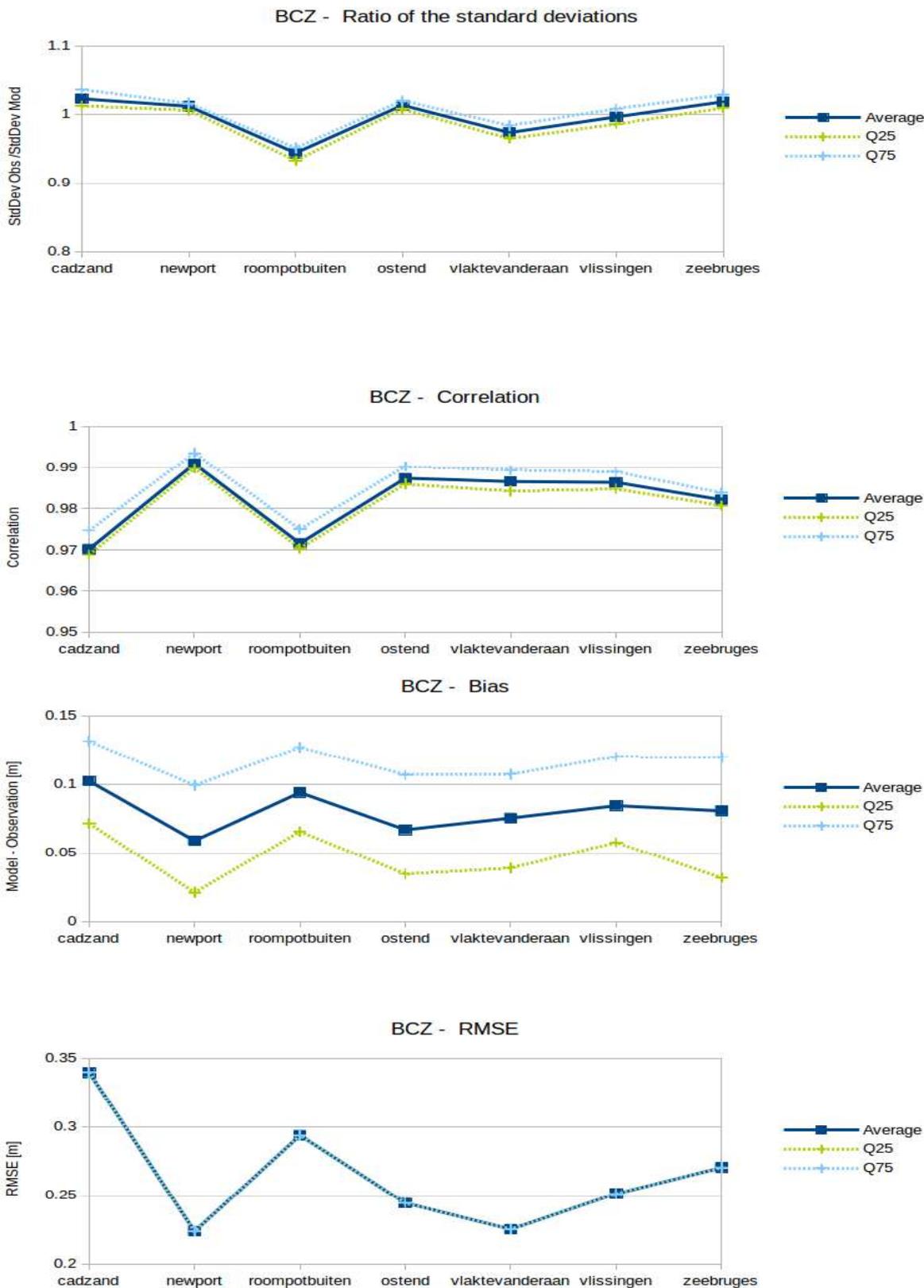


Figure 4.5 Summary statistics of validation metrics per stations, comparing the sea level BCZ forecasts to observations. From top to bottom: (1) Ratio of the standard deviation of the observations by the standard deviation of the forecasts; (2) Correlation; (3) Bias; (4) RMSE.

Figure 4.5 shows the excellent performance of BCZ at predicting the sea level. All metrics remain below satisfactory thresholds. Among all the stations, BCZ does best at Newport and Ostend. For the sake of illustration, Table 4.3 provides the statistics of the validation metrics at Newport for the three models.

Table 4.3 Comparison of the statistics of the validation metrics of three models at Newport.

Newport								
Model	Ratio		Correlation		Bias		RMSE	
	Average	Q75-Q25	Average	Q75-Q25	Average	Q75-Q25	Average	Q75-Q25
CSM	1.03	0.01	0.992	0.004	0.018	0.074	0.21	0.05
NOS	0.99	0.01	0.991	0.003	0.06	0.077	0.22	0.05
BCZ	1.01	0.01	0.991	0.004	0.06	0.078	0.22	0.05

In conclusion, the sea level forecasts are very accurate, with a standard deviation ratio close to 1, and a high correlation. The computed metrics are summarized in Table 4.4, for all the stations available for each model.

Table 4.4 Summary statistics of the model performance, all stations.

Model	Ratio		Correlation		Bias		RMSE	
	Average	Q75-Q25	Average	Q75-Q25	Average	Q75-Q25	Average	Q75-Q25

CSM	0.987	0.145	0.951	0.046	-0.390	0.083	0.092	0.598
NOS	0.895	0.176	0.969	0.034	0.083	0.037	0.295	0.087
BCZ	0.997	0.031	0.982	0.010	0.080	0.018	0.263	0.047

To compare the performance of the three models, the same statistics for the seven stations that the three models have in common are shown in Table 4.5.

Table 4.5 Summary statistics of the model performance, seven common stations.

Model	Ratio		Correlation		Bias		RMSE	
	Average	Q75-Q25	Average	Q75-Q25	Average	Q75-Q25	Average	Q75-Q25
CSM	1.016	0.029	0.990	0.002	0.031	0.018	0.201	0.012
NOS	0.976	0.030	0.987	0.006	0.083	0.027	0.242	0.034
BCZ	0.997	0.030	0.982	0.010	0.080	0.018	0.264	0.047

2.4.1.3 Sea surface temperature

Sea Surface Temperature data are available at the stations Den Helder, Europlatform, Helgoland, Hoek van Holland and IJmuiden. The position of the stations is shown in Figure 4.6. This section will give an overview of the model skill for predicting SST. The results show in general a better performance of the

model for the offshore Europlatform station. As temperature is important for the Belgian pilot study, the results for each individual station will be discussed later in the section.

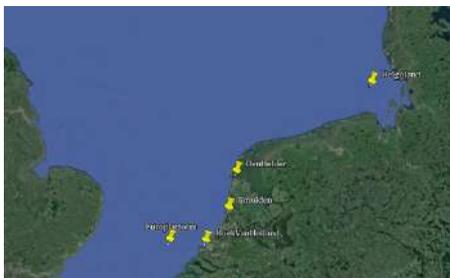
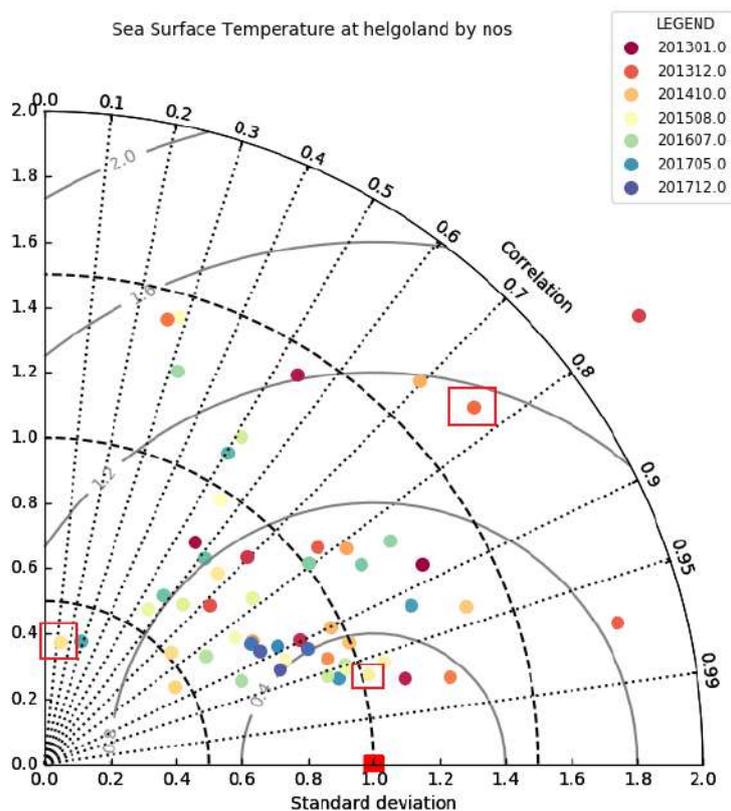
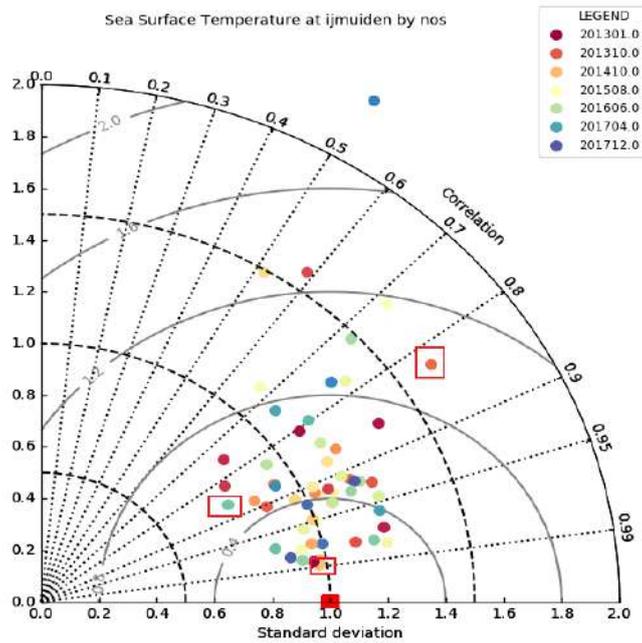
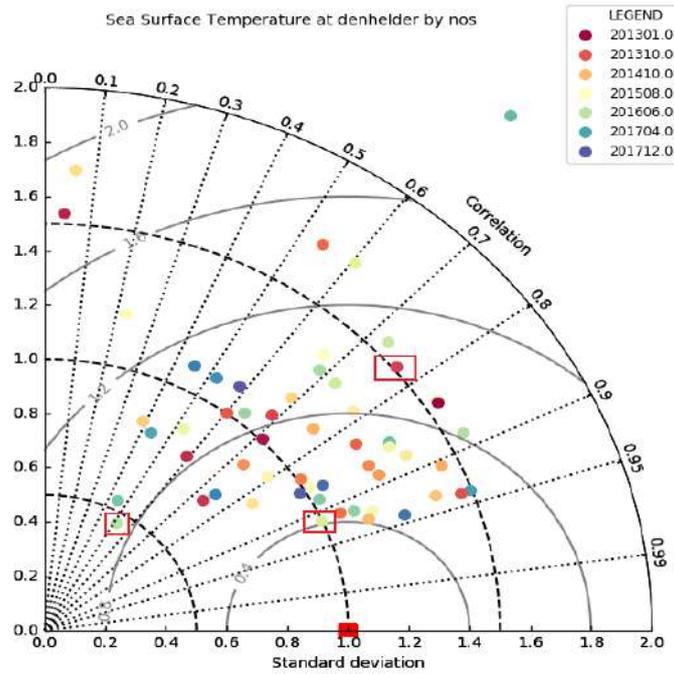


Figure 4.6 Position of the stations for the validation of SST forecasts.

In general, the monthly model validation of SST shows a large spread of the model skill in terms of standard deviation, correlation and RMSE as visually shown by the Taylor diagrams in Figure 4.7. The best model performance was at the offshore Europlatform station. The cause of this spread will be discussed for the individual stations in the next paragraphs.





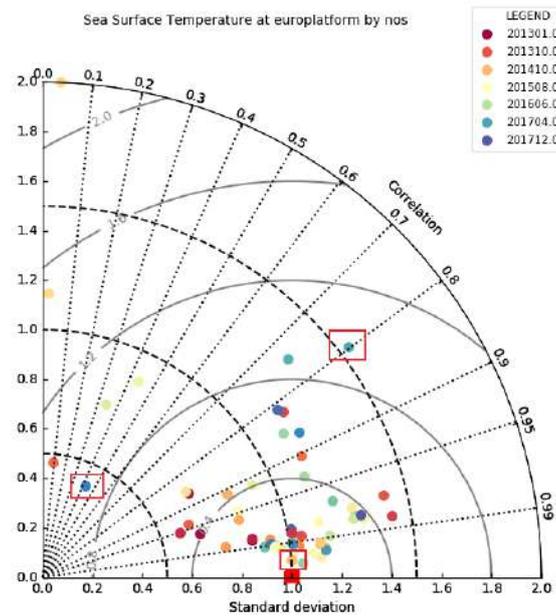
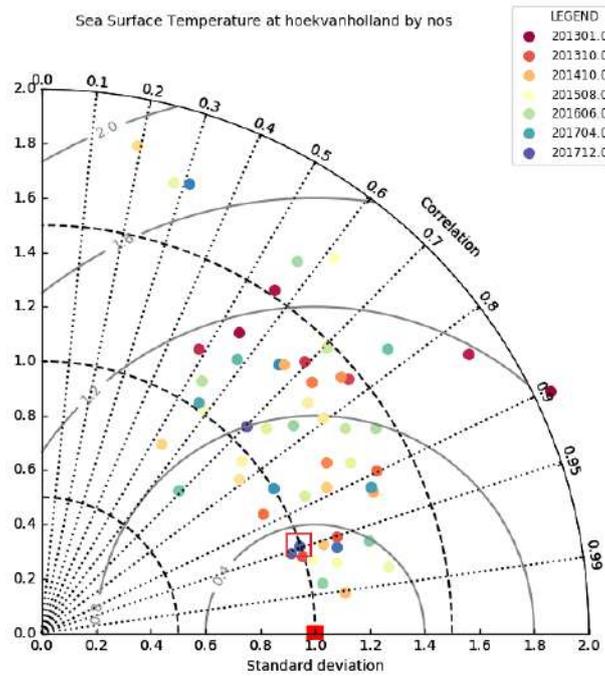


Figure 4.7 Taylor diagrams of the comparison of the sea surface temperature model forecasts with observations. From top to bottom: Helgoland, Den Helder, Ijmuiden, Hoek van Holland and Europlatform.

The statistics of the monthly biases are gathered in Table 4.6. The analysis of the variability in the previous paragraph already hinted that the performance of the model at predicting the sea surface temperature is better at the offshore station Europlatform than at the coastal stations. This conclusion holds for the analysis of the error too. The model forecasts are closer to the observations at Europlatform than at the coastal stations and the bias ranges are more compact at this station. This conclusion is not overly surprising since the offshore station has a better geo-physical adequacy to its model representation. The next paragraphs provide a more in-depth discussion of the model skill at each station.

Table 4.6 Summary statistics of the analysis of the error between NOS sea surface temperature forecasts and observations at various stations.

Station	Avg. monthly bias [°]	Range monthly bias [°]	Range span [°]	Average (Q75-Q25) [°]
Helgoland	-0.169	[-3.17 ; 1.61]	4.78	0.681
Den Helder	-1.21	[-3.38 ; 0.27]	3.65	0.985
Hoek van Holland	-0.915	[-2.19 ; -0.00]	2.13	0.750
IJmuiden	-1.77	[-2.10 ; -0.53]	2.63	0.503
Europlatform	0.181	[-0.09 ; 1.35]	1.44	0.268

Helgoland. The Taylor diagram shows that for some months the model performance is satisfactory, for other months the model performs less well. The main cause of this spread is that the model tends to simulate a strong tidal signal in the sea surface temperature that is not present in the observations. More in depth analysis of the observations seems to confirm the existence of a tidal signal, though much weaker as was predicted by the model. The difference in dominance of the background signal is well illustrated when comparing the distributions of SST in the observations and the forecasts (Figure 4.8) for the month 201402, which has a high correlation and high ratio between observations and model results. The temperature prediction background remains stable, with additional variations around this stable mean. The observations reveal an increase from a lower to a higher temperature value. This increase leads to a double distribution and hence a large standard deviation.

There is no clear seasonal tendency in the model performance for temperature, as shown in Figure 4.9. There does not seem to be a systematic deviation of the ratio or correlation with the season.

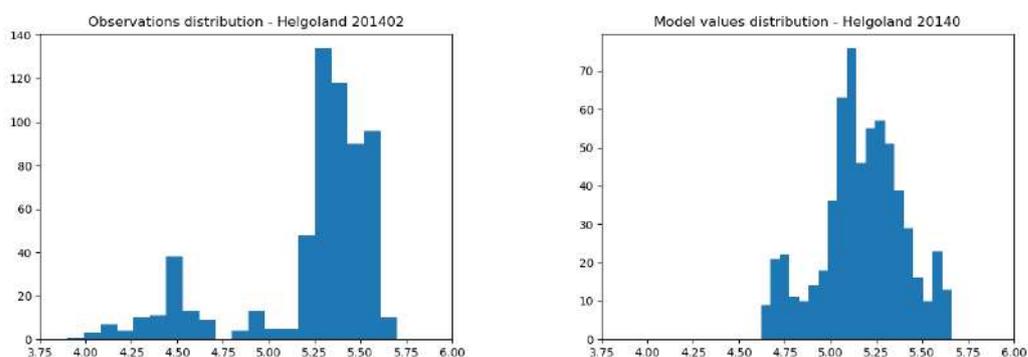


Figure 4.8 Distribution of SST observations and NOS forecasts at Helgoland for the month 201402. Left: Observations. Right: Model

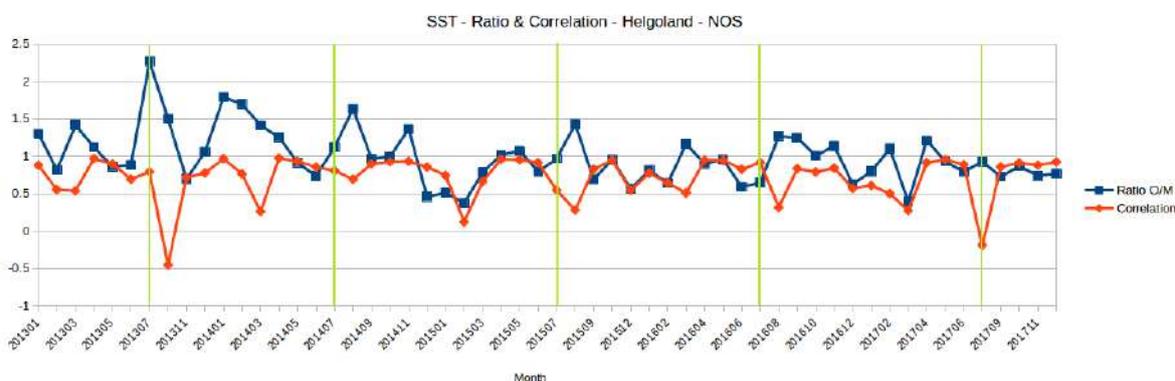


Figure 4.9 Time series of the monthly ratio (standard deviation of observations/model forecasts) and correlation for the sea surface temperature at the station Helgoland. Green lines are the months July.

Figure 4.10 shows the monthly bias and quantiles for the station Helgoland, the results seems to indicate there has been an evolution in the error from 2013 to 2017. The year 2013 is characterized by a systematic underestimation of the sea surface temperature. On the contrary, the years 2014-2017 display an oscillation around 0. There also seems to be a form of yearly trend in the error. The yellow lines pointing to the months of July hint at a stronger underestimation in spring. The years 2014-2017 exhibit a pattern of underestimation in winter/spring and an overestimation in summer/fall.

Den Helder and Hoek van Holland. Den Helder and Hoek van Holland are discussed together as both stations seem to have a similar validation performance (see Figure 4.7) which is physically backed up. The station Den Helder is in a mouth of the Wadden Sea into the North Sea, while the station Hoek van Holland in the mouth of the Rhine (see Figure 4.23). Both stations experience tidal effect of alternating water masses: the North Sea water flowing or ebbing in contact with Rhine or Wadden Sea water.

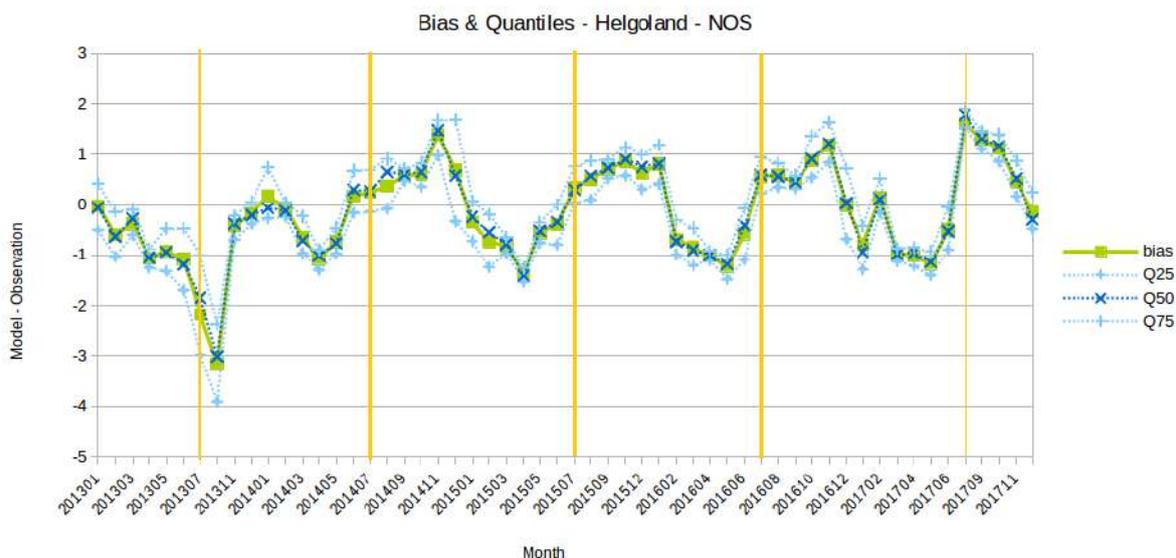


Figure 4.10 Time series of the monthly bias and Q25, Q50, Q75 quantiles comparing the sea surface temperature forecasts by NOS and observations at station Helgoland. The yellow lines highlight the months of July.

The observations at Den Helder have a tidal (12-hourly) signal while the model forecasts have a daily (24 hourly) signal, corresponding to the daily cycle of the temperature of the atmosphere. The model forecasts the average temperature on a 5x5 km cell of seawater, while the observations render the ebb and flow of the tide in the Wadden Sea mouth. This local phenomenon cannot be reproduced by the model. The inability of the model to reproduce the tidal signal seen in the observations may be explained by the fact that the model computes neither the physical state of nor the flow through the Wadden Sea. Contrary to the station Den Helder, the model has a tidal signal at Hoek van Holland. This might be linked to the fact that the ebb and flow into the Rhine mouth is modelled through and open boundary. The reproduction of the tidal signal by the model has fewer high frequencies (see figure 4.12).

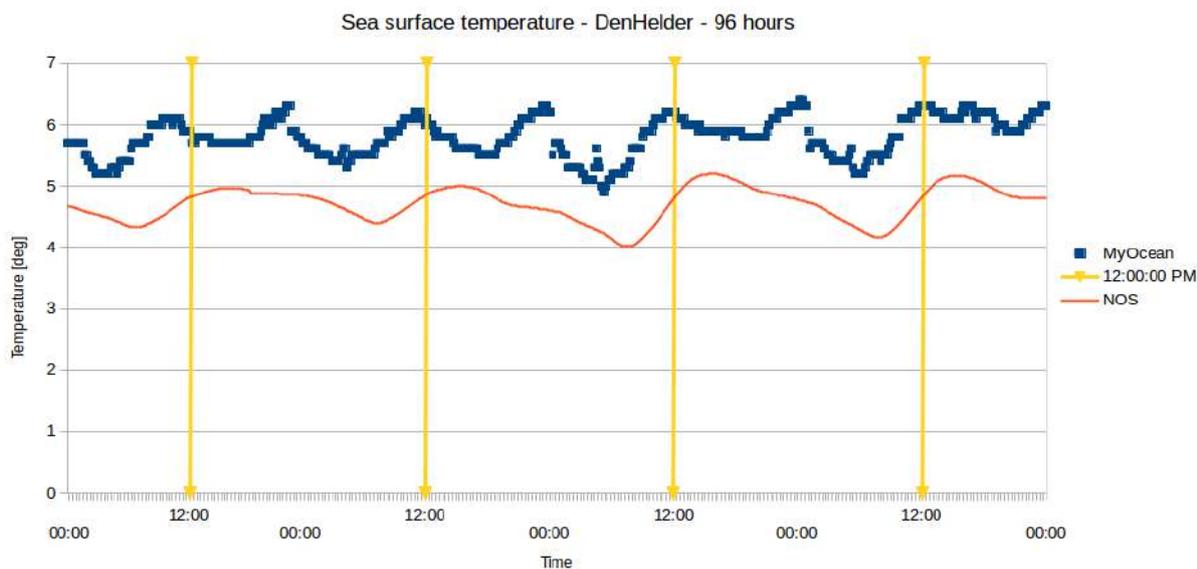


Figure 4.11 Time series of the sea surface temperature as forecast by NOS or observed at Den Helder, 96 hours long.

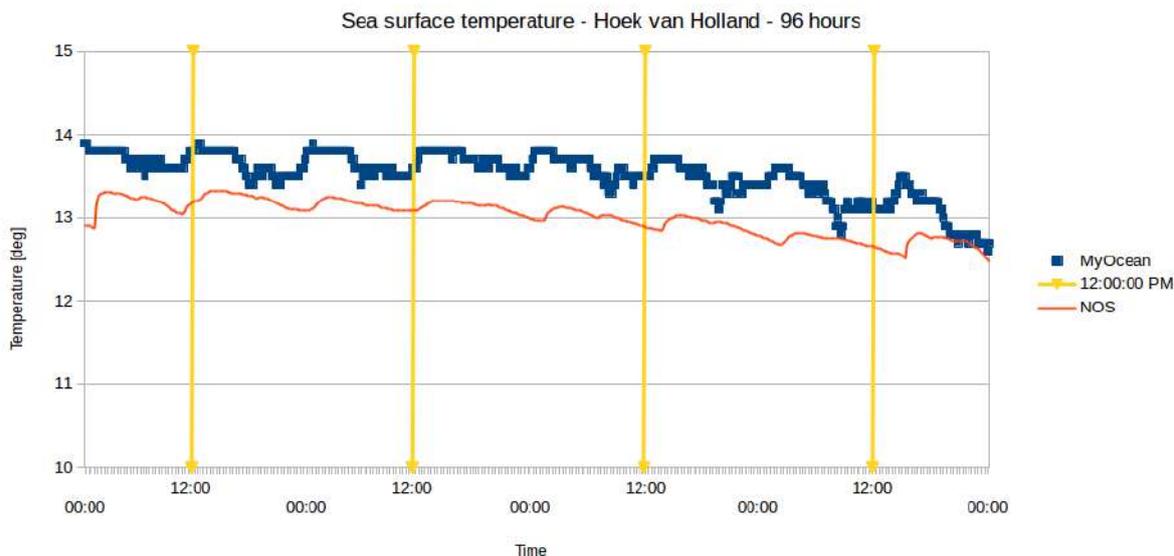


Figure 4.12 Time series of the sea surface temperature as forecast by NOS or observed at Hoek van Holland, 96 hours long.

The time series of the monthly ratio and correlation for Den Helder are shown in Figure 4.13.

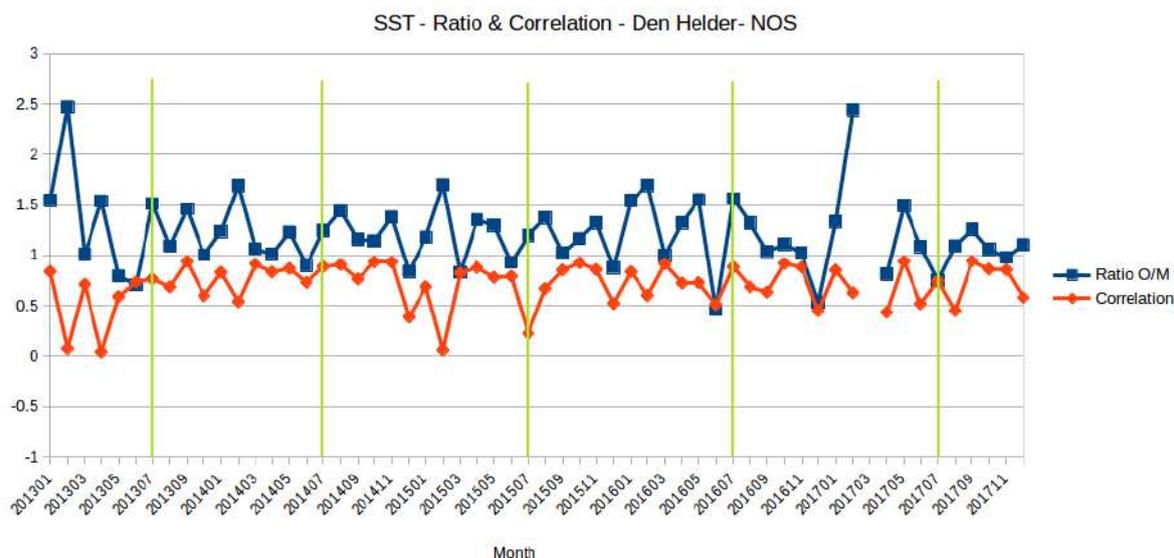


Figure 4.13 Time series of the monthly ratio (standard deviation of observations / model forecasts) and correlation for the sea surface temperature at the station Den Helder. Green lines are the months July. The month 201703 is withdrawn for lack of data

Figures 4.14 and 4.15 show the evolution of the monthly bias and quantiles for the station Den Helder and Hoek van Holland respectively. The monthly distributions of the error remains compact, the average difference between Q75 and Q25 is 0.822 degrees. Similar as observed for the station Helgoland, the stations show a yearly reoccurring dip in model skill. However, this cycle shifted towards negative values: the forecasts almost systematically underestimate the sea surface

temperature. This underestimation is strongest in spring, where the monthly bias drops to -2 to -3 degrees and more extreme at Den Helder than at Hoek van Holland.

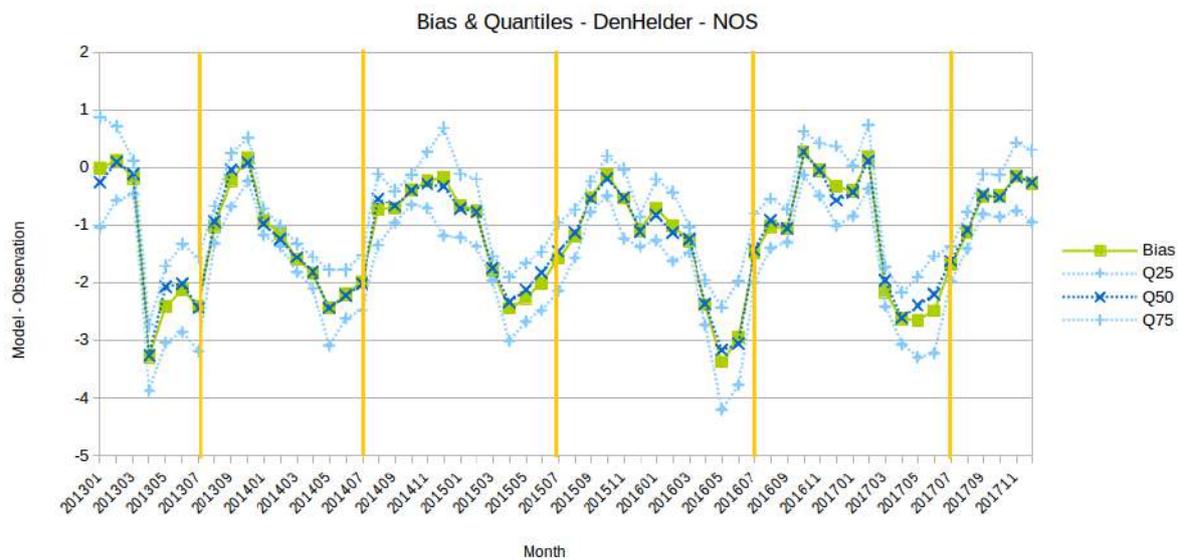


Figure 4.14 Time series of the monthly bias and Q25, Q50, Q75 quantiles comparing the sea surface temperature forecasts by NOS and observations at station Den Helder. The yellow lines highlight the months of July.

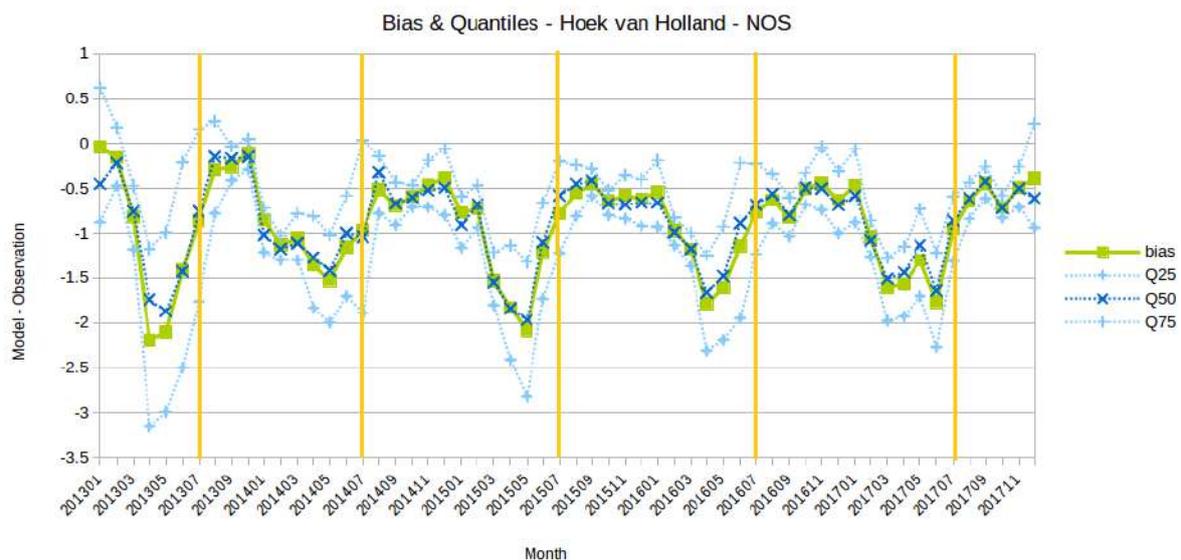


Figure 4.15 Time series of the monthly bias and Q25, Q50, Q75 quantiles comparing the sea surface temperature forecasts by NOS and observations at station Hoek van Holland. The yellow lines highlight the months July.

IJmuiden. The Taylor diagram of the station IJmuiden is visually more compact than the stations Helgoland, Den Helder and Hoek van Holland. The station does not have points with a very low ratio (less than 0.5). The correlation at IJmuiden is high and is linked to the ability of the model to timely reproduce the background variation, it possibly hints at a regular (and weak) river discharge at IJmuiden. Figure 4.16 further illustrates this distribution of the correlation at IJmuiden. Similarly to the

graphs of Helgoland or Den Helder, there is no apparent seasonal tendency in the ratio or the correlation.

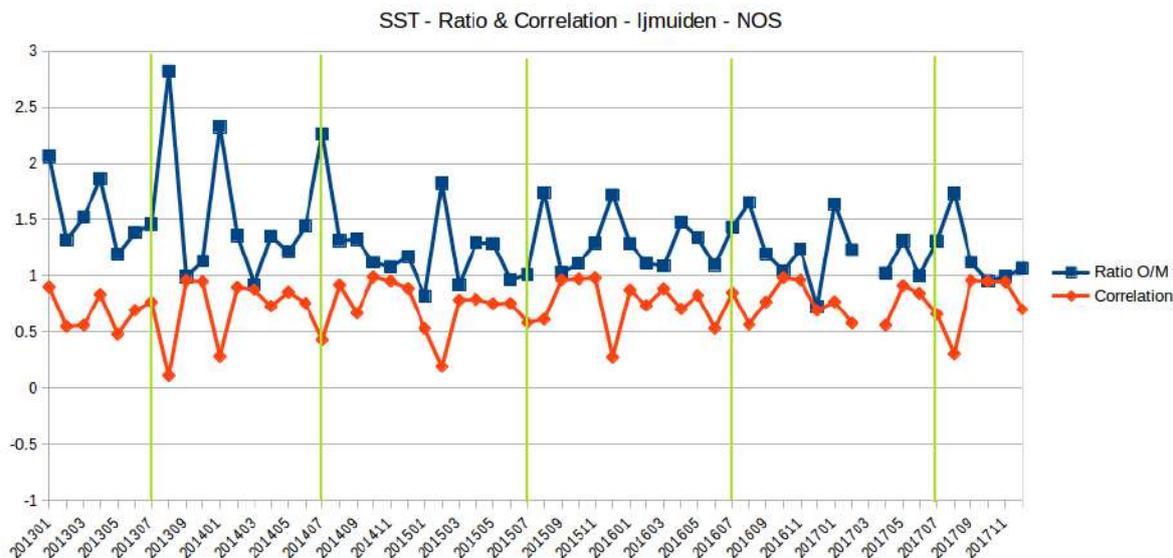


Figure 4.16 Time series of the monthly ratio (standard deviation of observations / model forecasts) and correlation for the sea surface temperature at the station IJmuiden. Green lines are the months July. The month 201703 is withdrawn for lack of data

Figure 4.17 illustrates the time series of monthly bias and quantiles at the station IJmuiden. Here as well, the distribution of the error is compact with an average difference of Q75-Q25 of 0.503 degrees. The temperature forecasts are systematically below the observations, by a larger offset than the other stations. The average monthly bias is -1.77 degrees. The yearly cycle of stronger underestimation in spring is present at IJmuiden also. At this station, the cycle is almost sinusoidal, switching between increasing and decreasing periods of approximately 6 months each.

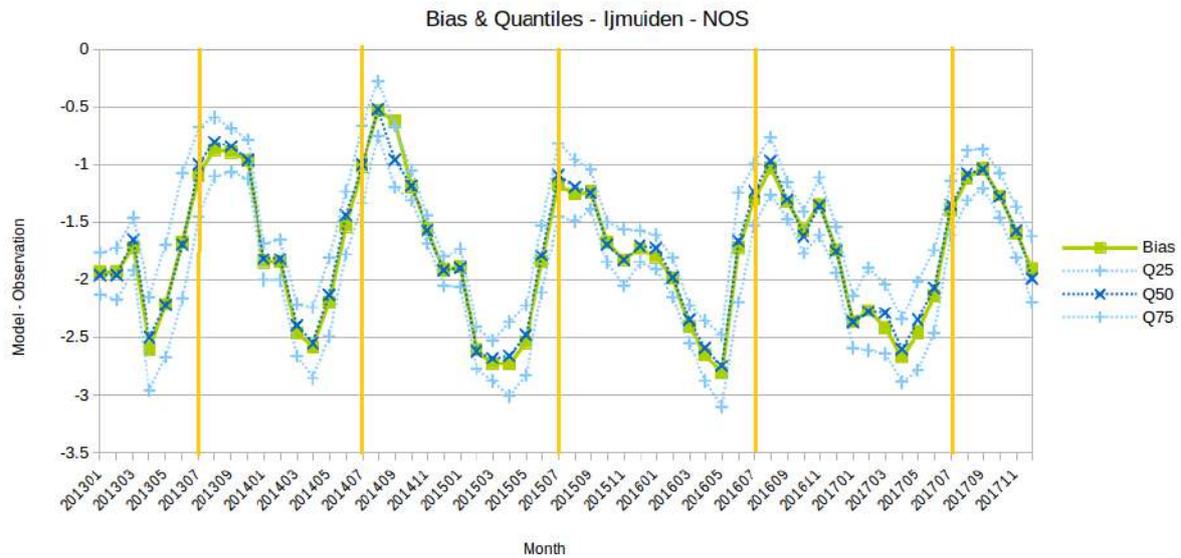


Figure 4.17 Time series of the monthly bias and Q25, Q50, Q75 quantiles comparing the sea surface temperature forecasts by NOS and observations at station Ijmuiden. The yellow lines highlight the months July.

Europlatform. The Europlatform station is the only offshore station, far from any river influence. The Taylor diagram in Figure 4.6.e shows a relatively broad ratio distribution, but a very high correlation, even higher than Ijmuiden. On the other hand, the model tends to create a daily signal, not present in the observations (see Figure 4.18.). The variations in the forecast have a weak daily component that appears to correspond with a slight warming up of the surface layer before noon. This daily component does not seem to be present in the data, possibly hinting at an overestimation of the heat transfer at the surface by the model. The figure 4.19 illustrates the absence of yearly tendency in the ratio and the correlation that are stable in time. The correlation is usually very close to 1, as is the ratio of the standard deviations. The model seems to be better at forecasting the sea surface temperature at Europlatform in terms of the variability than at the other stations. This might indicate a difficulty to model the inflow of non-modelled water mass (e.g. the river at Hoek van Holland or the Wadden Sea at Den Helder).

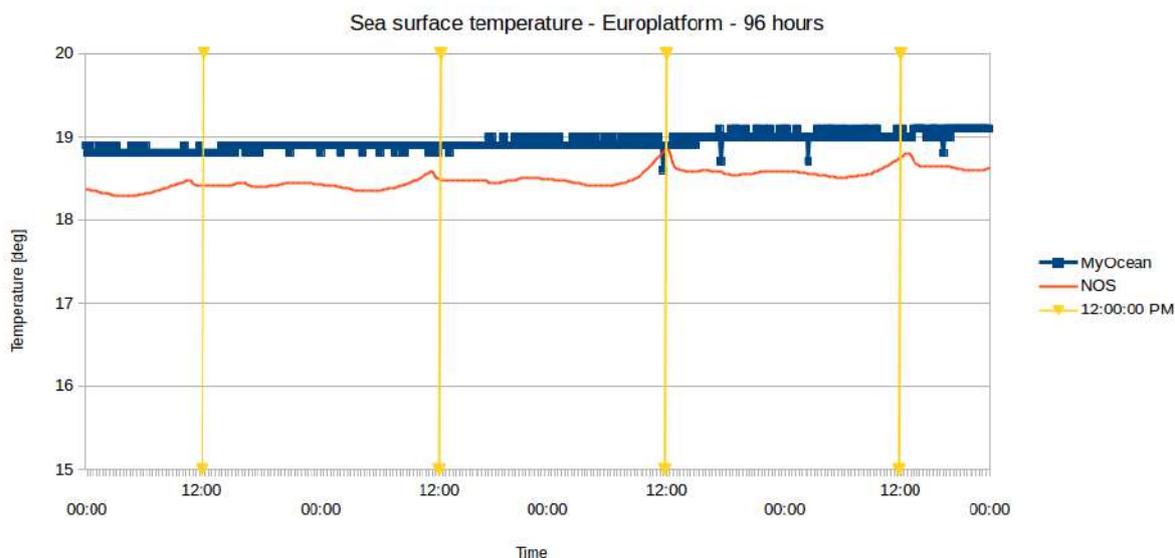


Figure 4.18 Time series of the sea surface temperature as forecast by NOS or observed at Europlatform, 96 hours long.

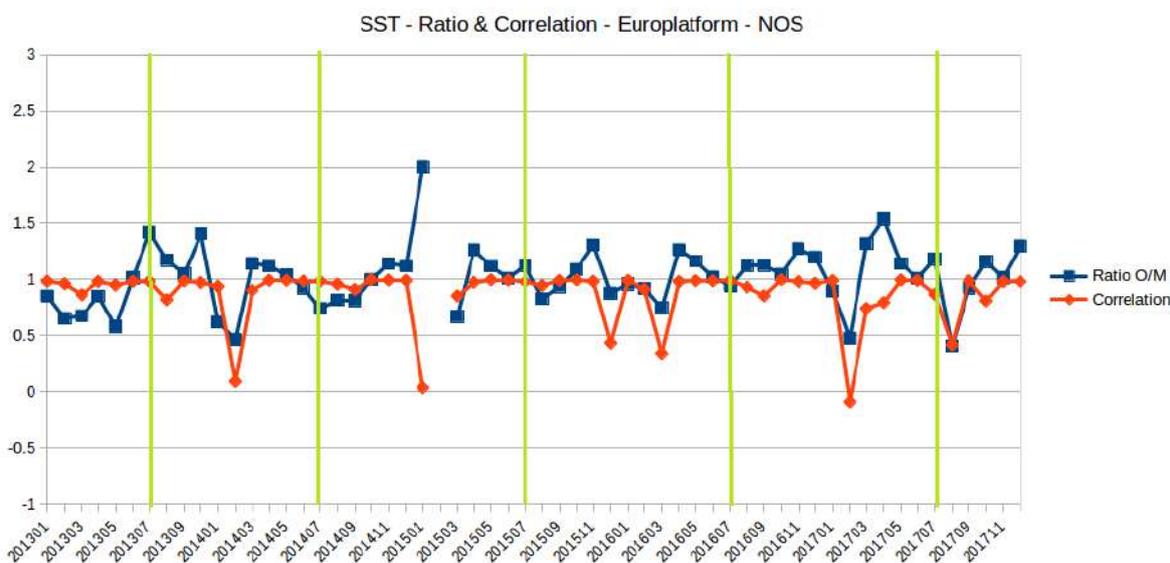


Figure 4.19 Time series of the monthly ratio (standard deviation of observations / model forecasts) and correlation for the sea surface temperature at the station Europlatform. Green lines are the months July. The month 201502 is removed for lack of data (367 points).

The time series of the monthly bias and quantiles at Europlatform is shown in Figure 4.20. This figure illustrates how the offshore station Europlatform differs from the coastal stations. First, the error distributions are very narrow: the average difference Q75-Q25 is 0.268, much lower than the coastal stations. Second, the existence of a yearly cycle is much less obvious for Europlatform. The range of the monthly bias is also narrower, its span covering 1.5 degrees. Third, the bias is more frequently positive than negative at Europlatform, so no systematic underestimation of the temperature is modelled.

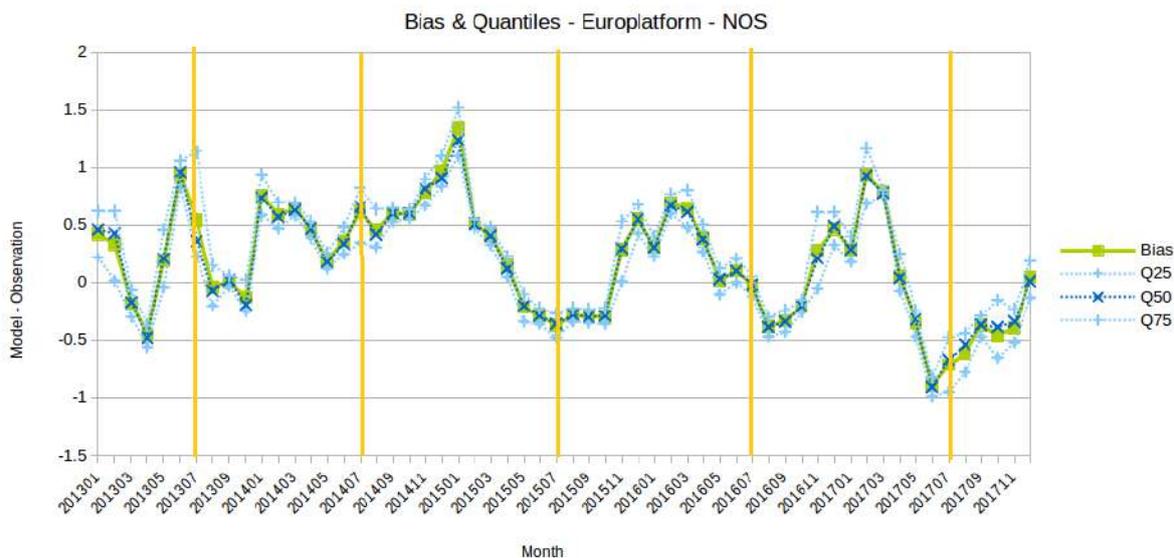


Figure 4.20 Time series of the monthly bias and Q25, Q50, Q75 quantiles comparing the sea surface temperature forecasts by NOS and observations at station Europlatform. The yellow lines highlight the months of July.

2.4.1.4 Sea surface salinity

The sea surface salinity can be compared at 3 different stations: Vlakte van de Raan, Helgoland and Hoek van Holland. The Optos stations and the CMEMS positions are shown in Figure 4.21. In general, the model shows a poor skill performance for SSS prediction. As for temperature the un-modelled inflow of river water masses seems to be the major culprit, hence a more in depth analysis between the relation of the positions of the stations in the model and the observations is given below. As salinity is not as important as temperature for the Belgian pilot study and it is a problematic parameter to validate, only the Taylor diagrams and the conclusions of this analysis will be discussed in this section. Both the implementation of a better representation for inflowing river masses and bathymetry and of additional observations are in progress to improve the model skill of this parameter.



Figure 4.21 Position of the Optos and CMEMS stations for the comparison of sea surface salinity values.

The salinity at these stations is very much influenced by the river plumes, and small deviations in the modelled hydrodynamic flows will potentially lead to large errors. The Vlakte van de Raan station is in the Scheldt estuary, Hoek van Holland by the Rhine estuary and Helgoland in a bay influenced by the Elbe plume. The precise relative positions of the stations are detailed on the maps 4.22, 4.23 and 4.24.



Figure 4.22 Relative position of the Optos (Vlakte van de Raan) and CMEMS (VlaktevdRaan) stations by the Scheldt estuary.

Figure 4.22 shows that the two stations at Vlakte van de Raan are some distance apart. They are actually about 5km distant from one another. This distance has important implications in terms of validation. The salinity is computed by the models NOS and BCZ. The salinity at a particular point must be seen as the value averaged over a grid cell. For BCZ, the grid side is approximately 850m, for NOS approximately 5km. On the other hand, the CMEMS value is retrieved from a mooring buoy and very local. Given the very strong influence of the Scheldt plume in this area, errors are to be expected for both the BCZ and the NOS model, for different reasons:

- NOS will deviate because of its averaging over a 5x5km area.
- BCZ will deviate partly because of its averaging (over a much smaller area), but also because of its distance to the actual CMEMS station. The Scheldt plume will oscillate with the tide, and pass the BCZ and CMEMS station at different times.



Figure 4.23 Relative position of the optos (Hoek Van Holland) and CMEMS (Hoek Van Holland NAP) stations by the Rhine estuary.

Figure 4.23 illustrates the positions of the stations at Hoek van Holland. The main problems arise from the fact that the CMEMS station is in the Rhine itself, while the optos station is at sea. In consequence, a marked offset between the two stations can be foreseen, especially since the Optos values are generated by NOS (average over a 5x5km cell).

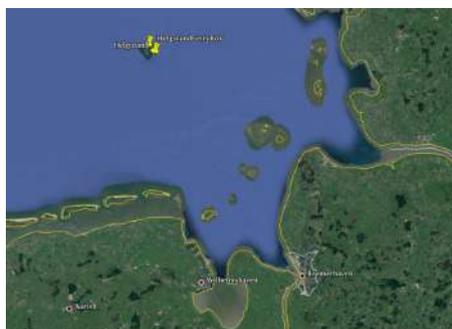
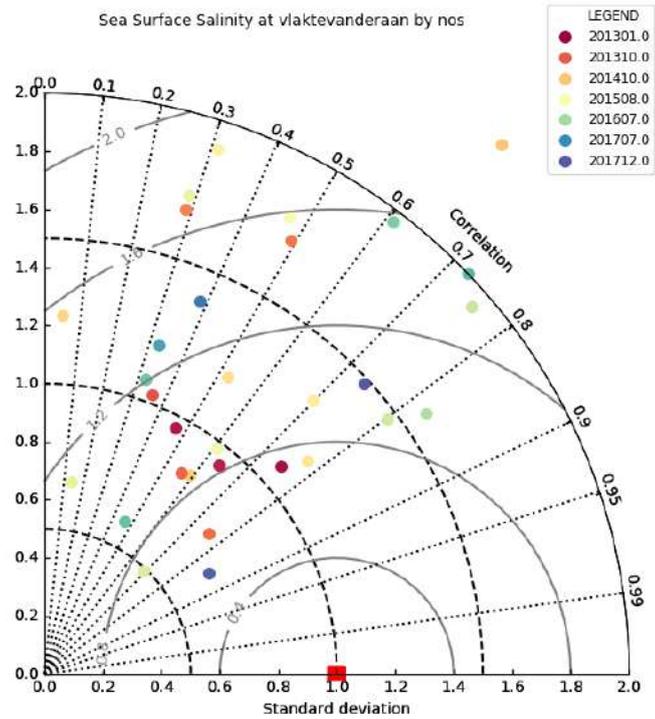
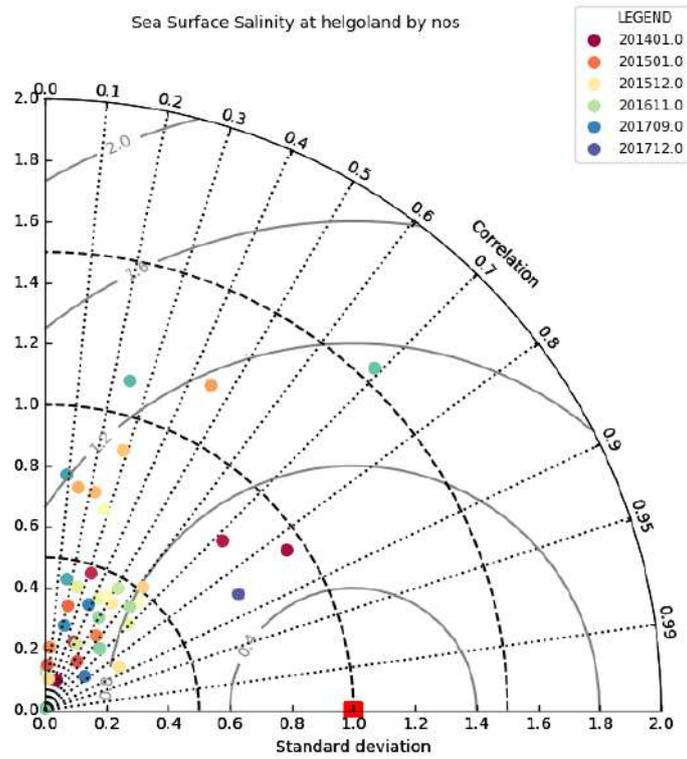


Figure 4.24 Relative position of the optos (Helgoland) and CMEMS (Helgoland Ferry Box) stations by the Elbe estuary.

Figure 4.24 shows that the Helogland stations are very close to one another, the only problem that can arise in the comparison is comparable to the station Vlakte van de Raan: the observations is very localized while the NOS data represents an average over a large area.

The ability of the models at reproducing the variability of the salinity is shown in Figure 4.25



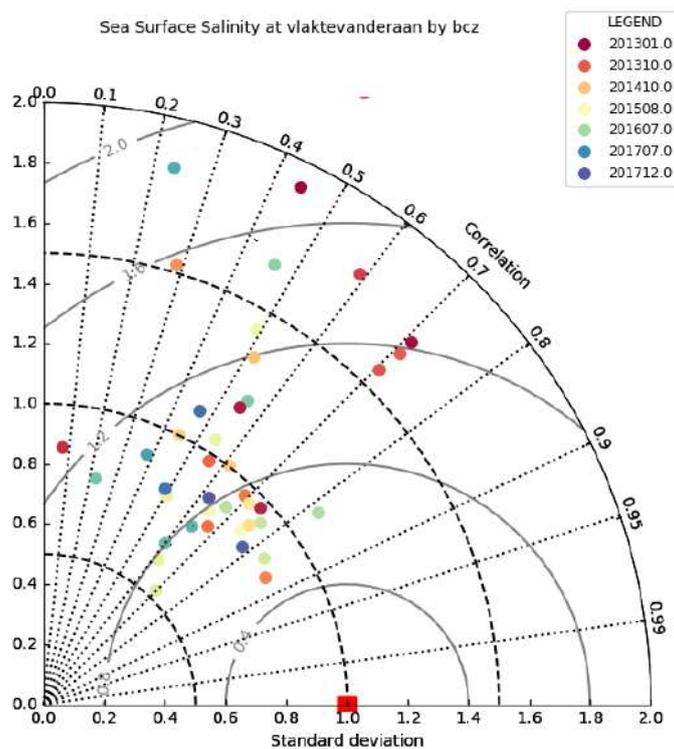


Figure 4.25 General performance in predicting the salinity at Helgoland by NOS (top) and Vlakt van de Raan by NOS (mid) and BCZ (bottom).

Figure 4.25 shows a lot of spreading of the data points on the Taylor diagrams. The analysis of Figure 4.25 demonstrated that Taylor diagrams are ill-fitted for the validation of the salinity. The spreading of the points on the graph is in fact misleading and the variables are uninformative. In addition, the spreading on the graph is so strong that some points are even not present on the graph. Computing averages of the correlation and standard deviation is pointless.

The salinity computed at Hoek van Holland cannot be validated: the measures are taken in the Rhine mouth while the NOS model computes the salinity at sea. Consequently, the data and the forecasts have little coherence when compared together.

At Helgoland, the model seems to overestimate the Elbe discharge and influence: it forecasts too low values, and the computation of the tide introduces a structure that does not exist in the data.

At Vlakte van de Raan, the data shows a monthly tendency with an additional tidal signal. These two features are reasonably well reproduced by the forecasts, but with a bias. The superposition of the background and tidal structures of salinity simulated by the model versus the measurements distort the metrics.

When studying the background sea surface salinity (filtering out the tidal component), the models show a stable offset in salinity. This offset is best measured by the median of the error distribution. The correlation between data and forecasts is poor in the sense that the river discharge is considered constant. Hence, the models do not have the ability to forecast a drop of salinity due to increased river

discharge. On the other hand, a variation in salinity due to hydro-dynamical conditions is reproduced. Statistical tests indicate that the BCZ model forecasts the background salinity with a better accuracy than NOS at Vlake van de Raan.

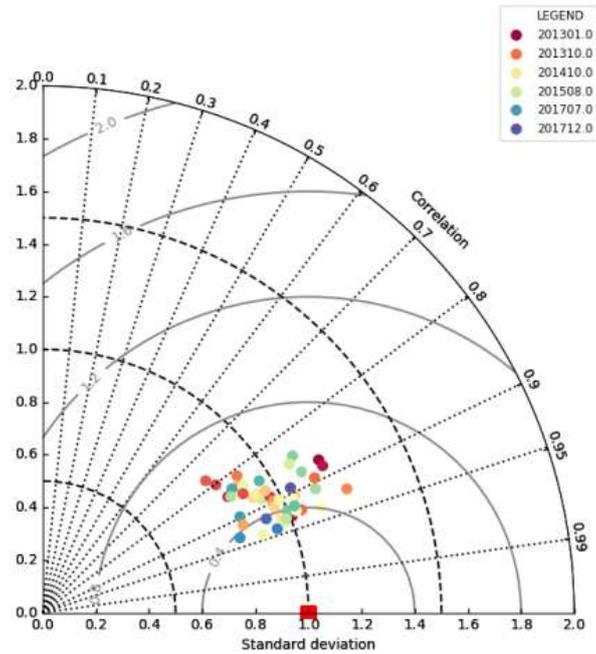
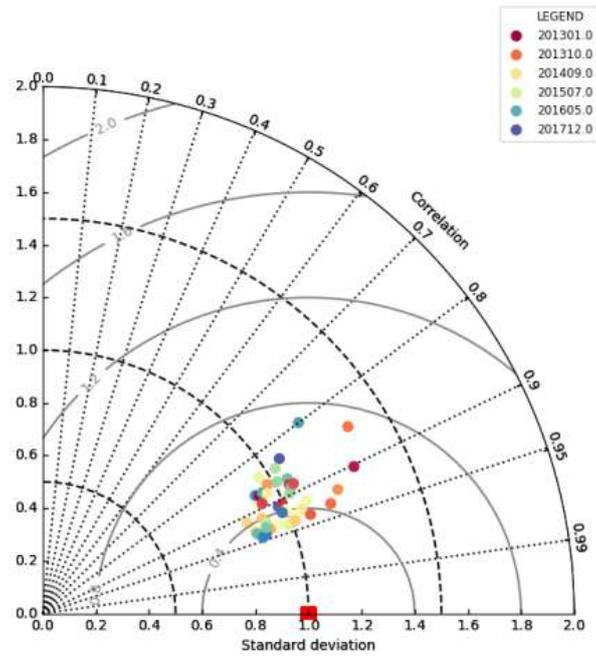
2.4.1.5 Significant wave height: frequency analysis

Hypas delivers forecasts of significant wave height at different stations, validation with observations is possible at Europlatform, K13 and Wandelaar. As a general rule, there is no clear seasonal tendency in the time series of the significant wave height error. The biases, RMSE and correlations are stable across the seasons. The average correlation of Hypas with respect to the observations is 0.885 while the average bias is 0.109. A more in depth discussion is provided in the next paragraphs of this section.

The general performance in terms of variability of the data is good, as is shown in the figures 4.26.a to 4.26.c. At all three stations, the ratios of the standard deviation of observations to the standard deviation of model forecasts are centred on 1. The correlation remains around 0.9 and the standard deviation of the error in the range [0.4-0.8]. These graphs show that the significant wave height variability is globally well reproduced by the model. The model behaves similarly at the 3 stations. The statistics are presented in Table 4.7.



Figure 4.26 Position of the stations K13, Europlatform, Wandelaar.



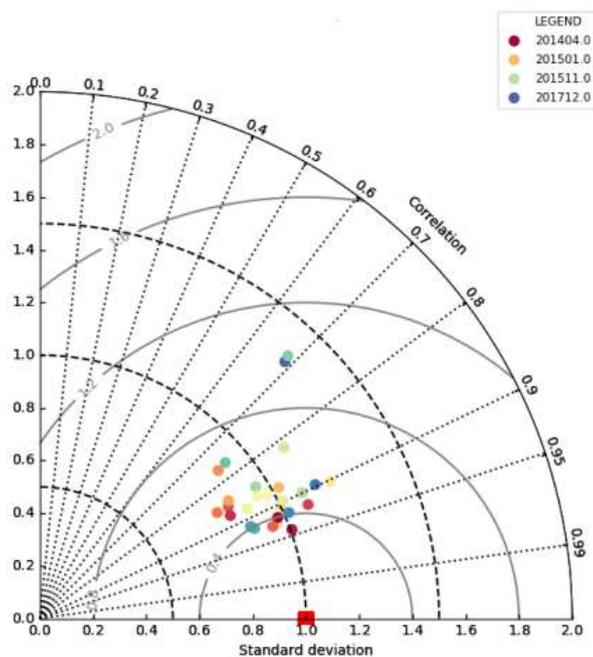


Figure 4.27 Taylor diagrams of the comparison of the significant wave height from Hypas forecasts with CMEMS observations at Europlatform,(top), K13 (mid) and Wandelaar (bottom) from 2013 to 2017.

Table 4.7 Statistics of the monthly RMSE between significant wave height forecasts by Hypas and observations, at different stations.

RMSE	Mean	Median	Standard Deviation
Europlatform	0.303	0.302	0.070
K13	0.344	0.388	0.060
Wandelaar	0.260	0.255	0.060

The time series of the bias and quantiles for the three stations are shown here below in Figure4.28. The bias is globally stable: there is no trend in the evolution of time; except for the station Europlatform and K13 where a slight increase of the bias in time might be seen. The variability of the bias reflects a shift in the distribution of the error: the quantiles move in a similar fashion. The bias is

limited to the range [-.01, 0.25] m. Table 4.8 shows that the average monthly bias remains close to 0.10 m. The Table 4.7 hints that there might be a difference in performance between the stations Europlatform and Wandelaar, this difference is however not significant. Figure 6.16 shows that the forecasts reproduce relatively well the variations of the data. The comparably low correlation is actually due to suspicious observations values, in orange circles. When those 5 points are removed from the data set, the correlation improves from 0.689 to 0.813.



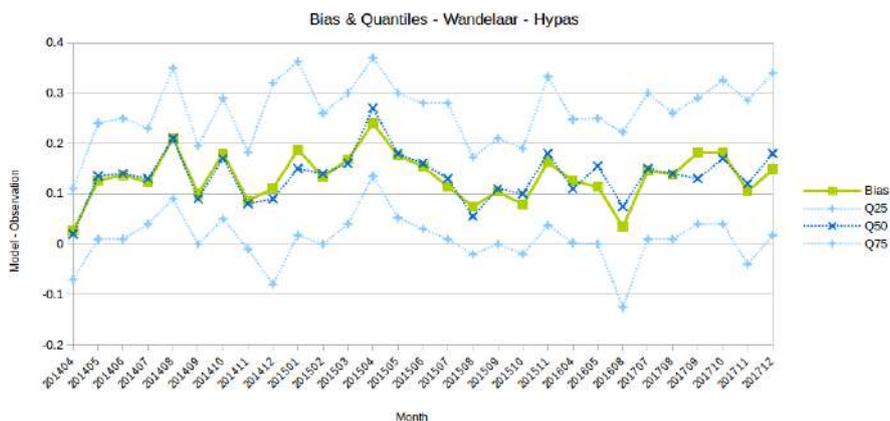


Figure 4.28 Time series of monthly bias, median and quantiles Q25 & Q75 comparing Hypas significant wave height forecasts with observations at three stations, from top to bottom Europlatform, K13, Wandelaar.

The Table 4.8 summarizes the statistics of the bias for the three stations.

Table 4.8 Statistics of the monthly bias between Hypas Significant wave height forecasts and observations at three stations.

Bias [m]	Mean	Median	Std Dev
<i>Europlatform</i>	0.0917	0.103	0.0889
<i>K13</i>	0.103	0.104	0.0839
<i>Wandelaar</i>	0.133	0.133	0.0485

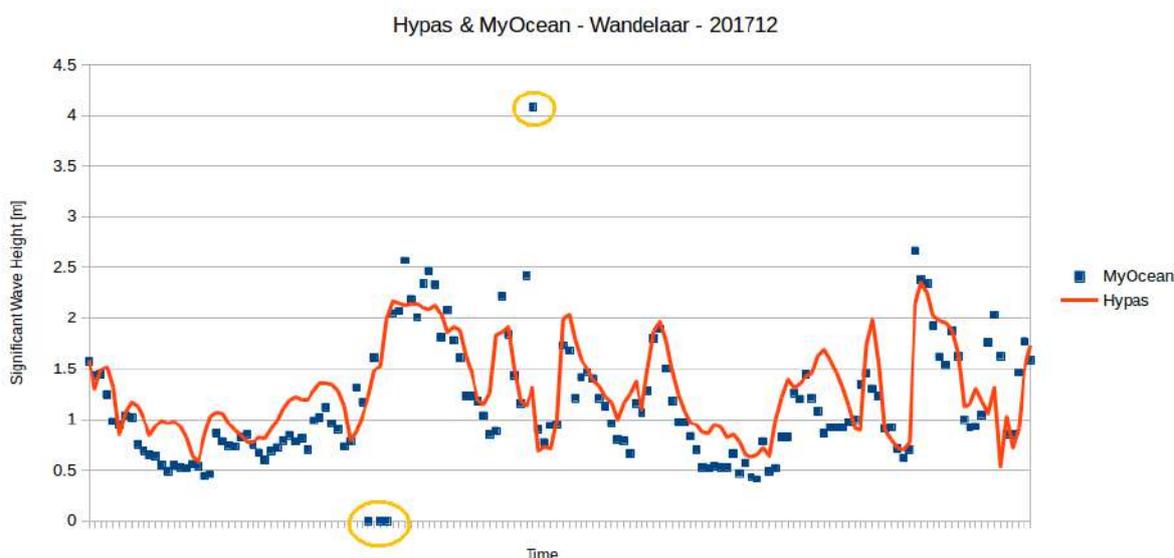


Figure 4.29 Significant wave height forecasts by Hypas and observations for the month 201712 at Wandelaar.

2.4.1.6. Zero up-crossing frequency

The localisation of the stations are displayed on map 4.30. The stations can be divided in two groups: Europlatform and K13 on the one hand, and Akkaert, Wandelaar and Westhinder on the other hand. This clustering is confirmed by the statistical differences: Europlatform and K13 cannot be said to differ, while Akkaert and Westhinder cannot be said to differ either. The case of Wandelaar is less clear, in part because of its lack of data points (29 points for Wandelaar, >40 for the other stations). In any case, the bias is larger at K13 and Europlatform than at Akkaert and Westhinder.

The average wind direction in this part of the North Sea is usually south-westerly, this may explain the tendency of the Hypas model to generate too much high frequencies when the fetch increases, or again to generate too much wind-sea.

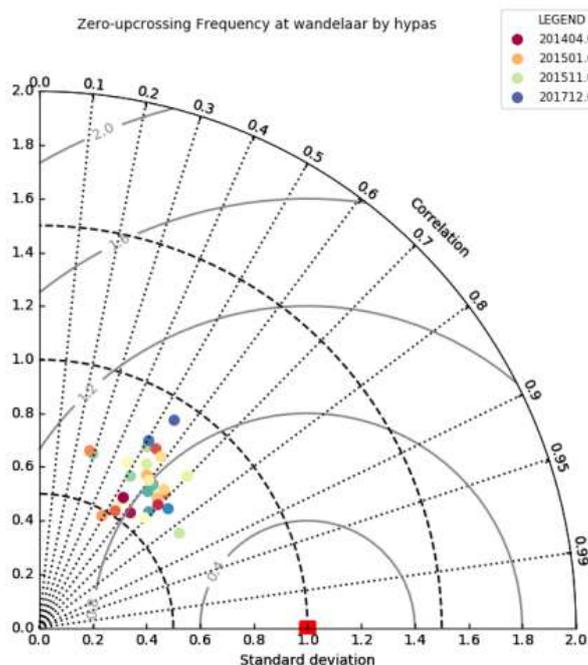


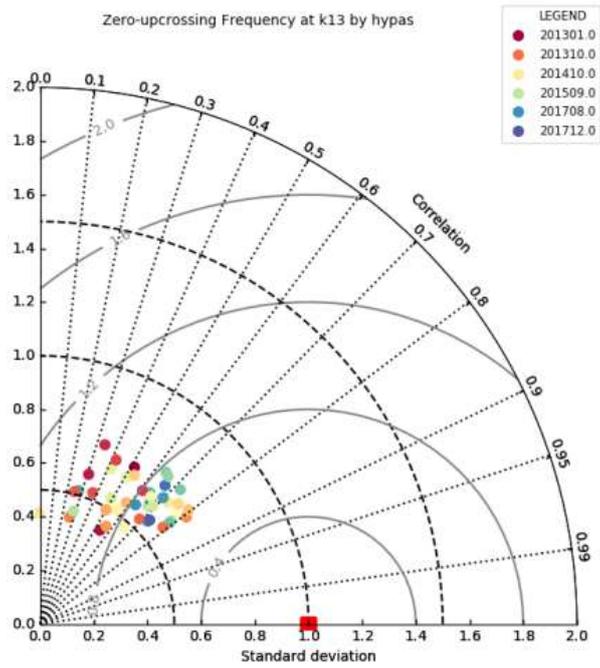
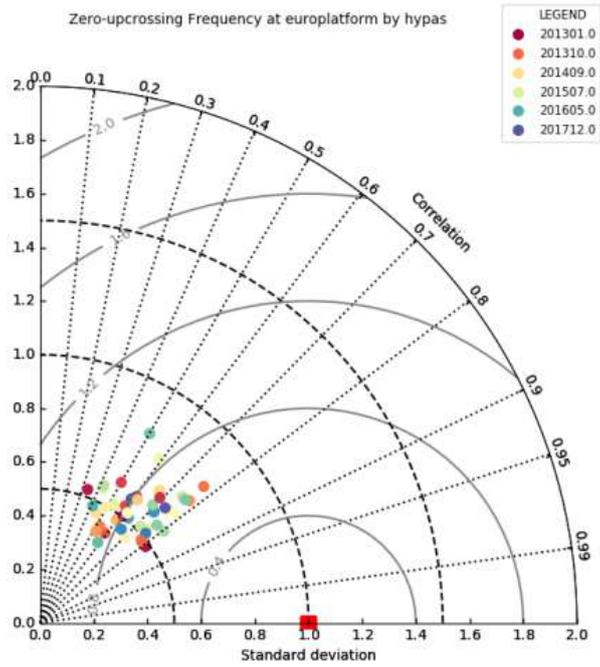
Figure 4.30 Position of the stations used for the validation of the average zero up-crossing frequency.

The general performance of the models at predicting the zero up-crossing frequency is seen at a glance in Figure 4.31. Figure 4.31 illustrates that the performance of the Hypas and Refra model is globally consistent at the different stations. The ratio of the standard deviations usually remains in the range [0.5-1] and the correlation in the range [0.3-0.8]. There are few outliers. Globally, the correlation is relatively poor, and ranges from 0.3 to 0.8 with an average of ~0.55. The ratio of the standard deviations is usually below 1: the models introduce more variability than is present in the data.

Figure 4.32 shows the time series of the monthly bias for the different stations and models, there does not seem to be a seasonal effect, nor a global tendency in the data. On the other hand, the average bias seems to vary with the station. In terms of bias, the stations seem to differ, creating a group of offshore stations (Europlatform, K13) with larger bias or coastal stations (Akkaert, Wandelaar and Westhinder) with more moderate bias. This might reflect the impact of the fetch, given that the wind is usually southwesterly in that part of the North Sea. The analysis of time series of actual data indicates that the models tend to generate events of high average frequency that do not reflect the reality. On the other hand, the observations show the zero up-crossing frequency varies rapidly in the short term; variability that is not reproduced by the model.

The RMSE is usually larger than the Bias. The RMSE seems to be sensitive to the overall overestimation of frequencies, and especially to the high frequencies events. There is no clear clustering of stations like was the case for the bias, the tests on the differences in average RMSE are usually not conclusive.





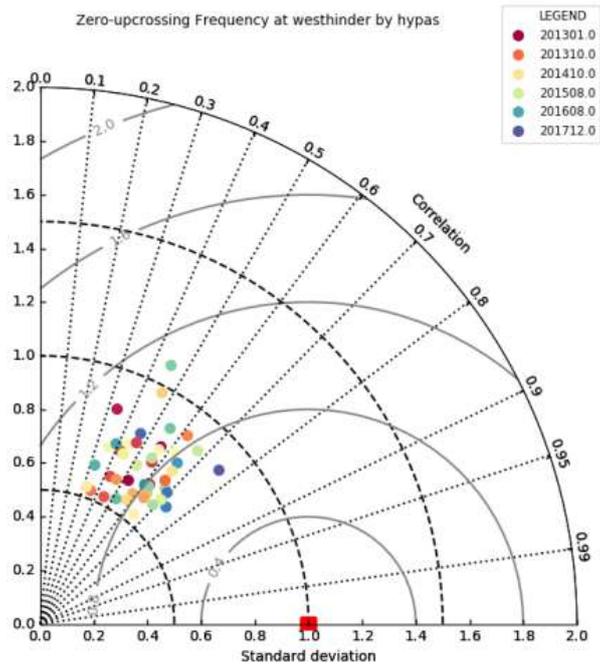
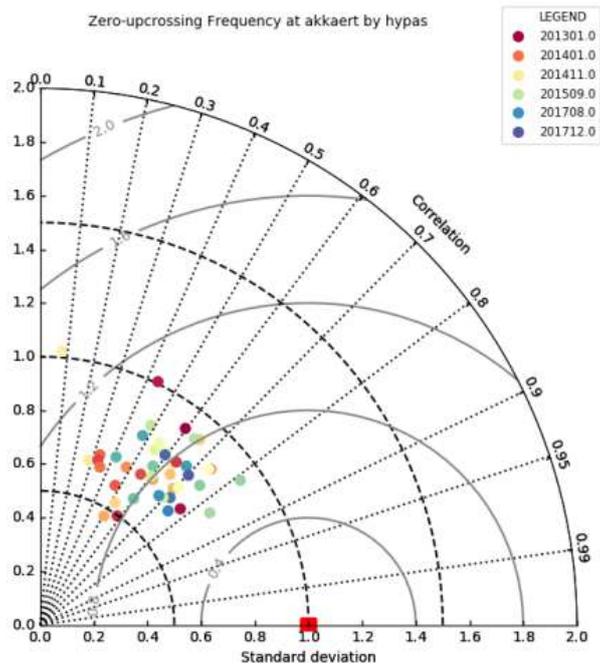


Figure 4.31 General performance of the models Hypas and Refra at forecasting the zero up-crossing frequency at (from top to bottom): Wandelaar, Europlatform, K13, Akkaert and Westhinder.

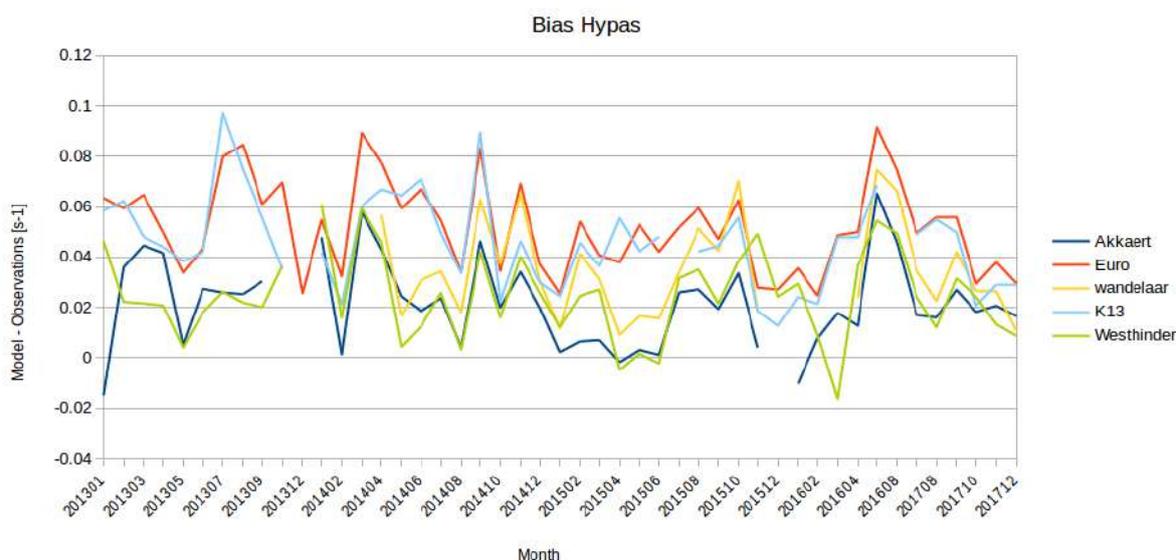


Figure 4.32 Time series of the monthly bias between average zero up-crossing frequency model forecasts and observation at various stations.

2.4.2 Forecast validation

As mentioned in section 2.4.1, the analysis of the Belgian pilot study relies on its input data on products provided by the Marine Forecasting Centre (MFC) of the Royal Belgian Institute of Natural Sciences (RBINS). MFC is part of EuroGOOS (European Global Ocean Observation System) and delivers their forecasts results on a daily basis to NOOS (North West Shelf Operational Oceanographic System) where the data of most oceanographic national centres bordering the North Sea are gathered, compared and processed for a multi-ensemble model (MME) forecast. The Multi Model Ensemble (MME) of forecast products is based on the outputs from several operational ocean forecasting models contributed by NOOS partners. The development of MME is coordinated by BSH and supported by the Copernicus Marine Environment Monitoring Service. Hence, the Belgian pilot study would like to forward the reader looking for information regarding the forecasting skill of the MFC's operational models to the website of NOOS (<http://noos.eurogoos.eu/>). An example of how the data can be consulted is shown in Figure 4.32. Figure 4.33 shows the resulting surface temperature distribution of the ensemble model produces by NOOS 18th of February 2022 together with how the different models are related to this ensemble model forecast. The results for MFC are denoted as RBINS_OPTOS_NOS.

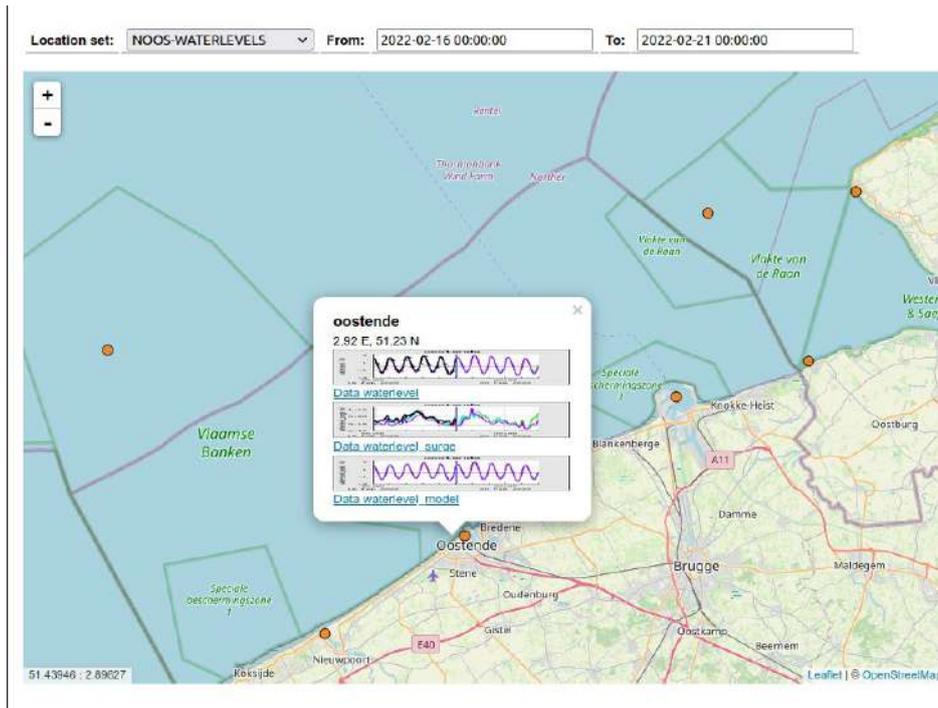


Figure 4.33 Screenshot of the presentation of forecasting skill at the NOOS website.

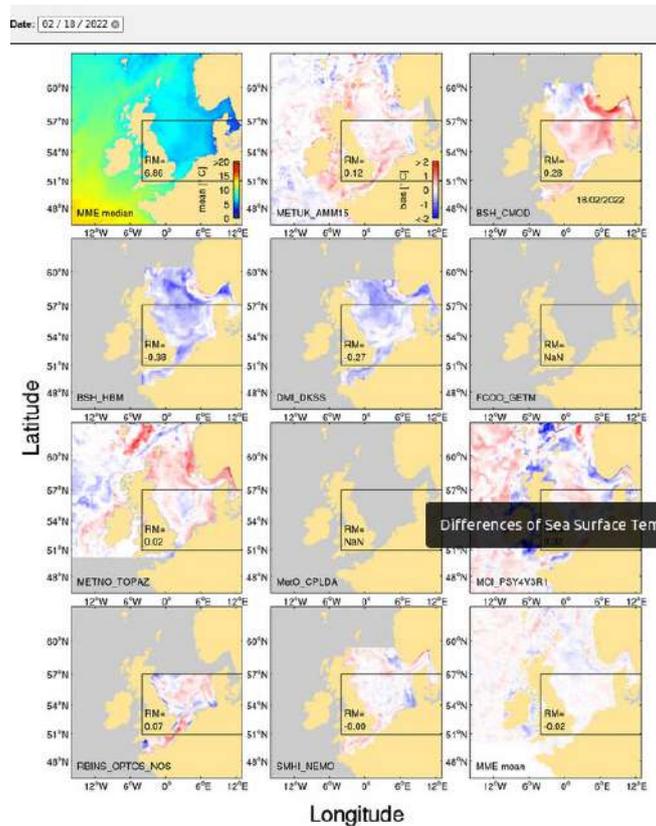


Figure 4.34 Daily averaged differences of SST between forecast products of the ensemble members and the MME median (<http://noos.eurogoos.eu/increasing-noos-awareness/community-tasks/multi-model-ensemble-of-forecast-products>)

2.4.3 Process-oriented validation

This section focuses on the temperature validation at Westdiep (51.18°N, 2.67°E). A buoy was installed in the pilot site to record temperature. Table 4.9 shows the metrics for the monthly temperature validation of the model at Westdiep. It becomes clear from this table that the variability is very well captured by the model (standard deviation ratio model/observations is centred around 1 and correlation coefficient reaching 1), however there is a consequent underestimation of the measured temperature by the model (negative bias of -0.66 °C for the whole period of measurements). Figure 4.35 shows the comparison of the available measure data and the modelled data. The negative bias comes as no surprise as the model calculates the average temperature of the upper layer of the model. Hence, the colder water masses at deeper sea levels are also taken into account, whereas the measures are taken at one point. The direct comparison of the model (Figure 4.35) shows that the bias becomes smaller during July and August. This supports the idea that the cause of this bias is due to point measurements of the observations vs. depth-integrated upper layer results in the model, since the temperature in the North Sea is typically better mixed during the warmer summer months.



Figure 4.35 Direct comparison of daily averaged sea surface temperature observations with modelled daily averaged sea surface temperature observations from June till October 2018 at Westdiep (51.18°N, 2.67°E).

Table 4.9 Overview of the metrics when assessing the skill of the BCZ model (described in Section 4 at predicting the temperature at Westdiep (51.18°N, 2.67°E).

<i>Period</i>	<i>RMSE</i>	<i>Std. dev. model</i>	<i>Std. dev. obs.</i>	<i>Std. dev. ratio</i>	<i>Correlation</i>	<i>Bias</i>
<i>Jun-18</i>	<i>0.63</i>	<i>0.54</i>	<i>0.71</i>	<i>1.33</i>	<i>0.99</i>	<i>-0.60</i>
<i>Jul-18</i>	<i>0.84</i>	<i>0.88</i>	<i>0.82</i>	<i>0.93</i>	<i>0.96</i>	<i>-0.80</i>
<i>Aug-18</i>	<i>0.41</i>	<i>1.16</i>	<i>1.11</i>	<i>0.96</i>	<i>0.98</i>	<i>-0.35</i>
<i>Sep-18</i>	<i>0.76</i>	<i>0.71</i>	<i>0.80</i>	<i>1.12</i>	<i>1.00</i>	<i>-0.75</i>
<i>Oct-18</i>	<i>1.04</i>	<i>0.28</i>	<i>0.45</i>	<i>1.61</i>	<i>0.53</i>	<i>-0.98</i>

2.5 Pilot 5: Ireland

The Galway Bay model is a ROMS-based hydrodynamic prediction system providing a 10-year (2012-2021) hindcast and 3-day forecasts of the physical state of the waters in the Galway Bay east of $09^{\circ}12'43.2''\text{W}$, including temperature, salinity, sea surface height and ocean currents. Different observational platforms are available to assess the quality of the model predictions: (a) the Galway Port tide gauge; (b) the Multi-Scale Ultra-High Resolution (MUR) Sea Surface Temperature (SST); (c) three ADCP moorings deployed in the Galway Bay during spring and summer 2018; (d) three moored CPT loggers located at important farming sites within the bay and operated by Cuan Beo, recording temperature and salinity, together with temperature and salinity measurements from a hand-held sensor at one of the CPT logger locations; and (e) CTD casts recording temperature and salinity measurements throughout the bay on a quarterly basis, starting on May 2021.

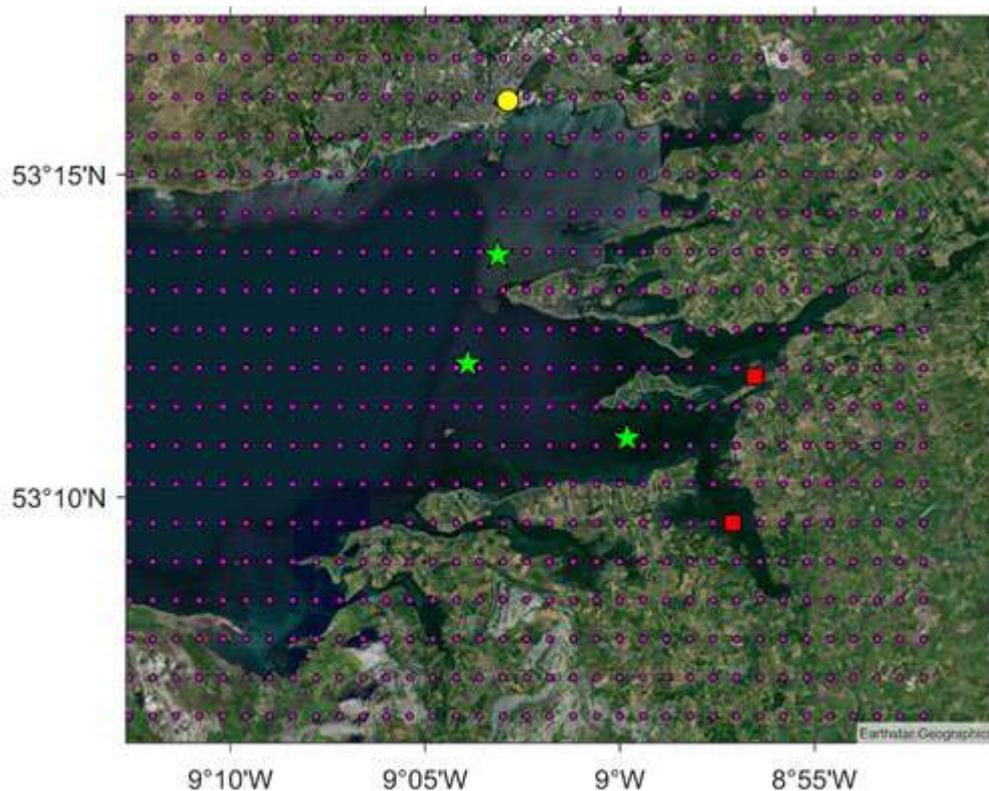


Figure 5.1 Observational platforms in Galway Bay: (a) Galway Port tide gauge shown as a yellow marker, ADCP moorings (green markers, A, B, C from north to south), CPT loggers (red markers, north: Killeenaran, south: Kinvara West and Kinvara East) and MUR-SST 0.01° grid (dots) in the Galway Bay.

2.5.1 Hindcast validation

Even though a 10-year hindcast for the Galway Bay has been produced, the hindcast validation here focuses on one particular year, from October 2019 to September 2020, because this is the first

hindcast that has been made available. In addition, the spring and summer 2018 period is also examined for the ADCP data.

2.5.1.1 Sea level from the Galway Port tide gauge

The Irish National Tide Gauge Network operates several tide gauges around the coastline of Ireland. Sea level recordings from the Galway Port tide gauge (53°16'08.4''N 09°02'52.8''W, see Fig. 5.1) were compared against the Galway Bay model predictions (Figs. 5.2, 5.3).

An excellent correlation (CORR = 0.988, see Table 5.1) reveals that the model accurately predicts the tidal phase. In general, the tide amplitude is also well reproduced, although there is an underestimation of the spring tides during high waters, which results in a negative mean error of -1.3 cm.

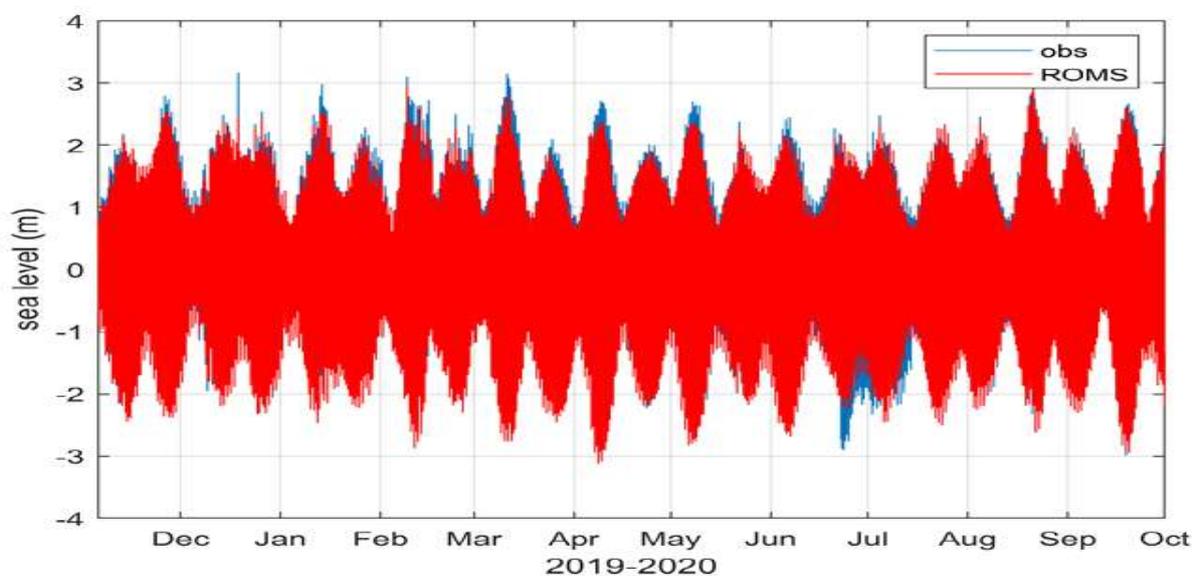


Figure 5.2 Observed (blue) and modelled (red) time series of the sea level at Galway Port from November 2019 to September 2020.

Table 5.1 Galway Port observed sea level vs. ROMS Galway Bay predictions statistics (ME = Mean Error, MAE = Mean Absolute Error, RMSD = Root Mean Squared Difference, CORR = correlation, n = number of data points).

ME	MAE	RMSD	CORR	n
-0.013 m	0.145 m	0.202 m	0.988	93709

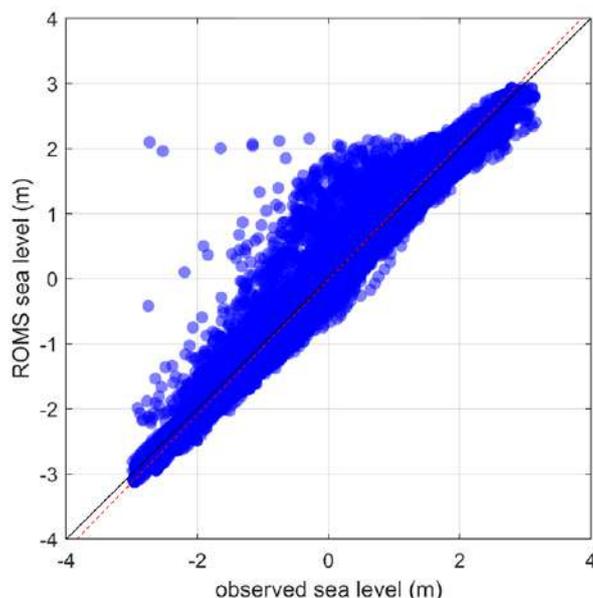


Figure 5.3 Scatter plot of the observed vs. the modelled sea level at Galway Port. $y = -0.013 + 1.039x$; $R^2 = 0.975$

2.5.1.2 ADCP

Three ADCP moorings were deployed by the Marine Institute in the Galway Bay during spring and summer 2018, providing time series of marine currents at 23 different vertical levels. The locations of the three ADCP moorings are shown (Fig. 5.1). Table 5.2 provides general statistics for the three ADCP moorings and for the u-component, the v-component and the current speed independently.

Table 5.2 CORR, RMSD and ARMAE considering all depths and for the u-component, v-component and speed. An observational error of 1 cm/s has been used for the computation of the ARMAE value (Dabrowski et al., 2016).

	CORR	RMSD	ARMAE
ADCP A (u)	0.861	0.066	0.376
ADCP A (v)	0.470	0.055	0.886
ADCP A (speed)	0.619	0.060	0.652
ADCP B (u)	0.827	0.076	0.447
ADCP B (v)	0.294	0.071	0.794
ADCP B (speed)	0.495	0.066	0.730
ADCP C (u)	0.899	0.079	0.348

ADCP C (v)	0.175	0.053	1.008
ADCP C (speed)	0.704	0.072	0.638

In terms of the correlation and ARMAE, it can be seen that better results are obtained for the u-component of the current than for the v-component. This is because of the strong tidal nature of the current field within the Galway Bay, and the tide propagates along the east-west direction following the main axis of the bay. As a result, the u-component of the flow is stronger and its variability almost depends entirely on the tidal forcing, which is easier to predict. The match between the ROMS Galway Bay model predictions and the ADCP observations for the three locations A, B, C can be assessed from the figure below (Figs. 5.4, 5.5, 5.6).

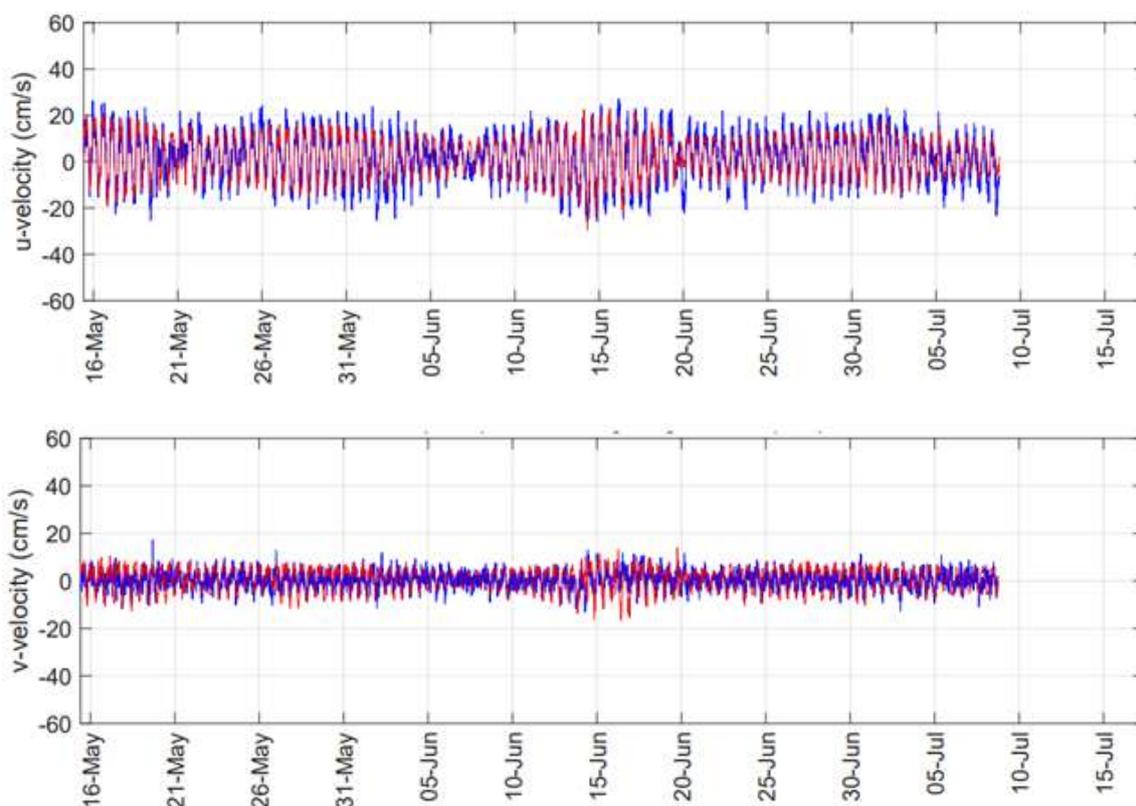


Figure 5.4 Observed (blue) and modelled (red) time series of the currents at the ADCP A mooring from spring-summer 2018: u-component (top) and v-component (bottom) at the surface level are shown.

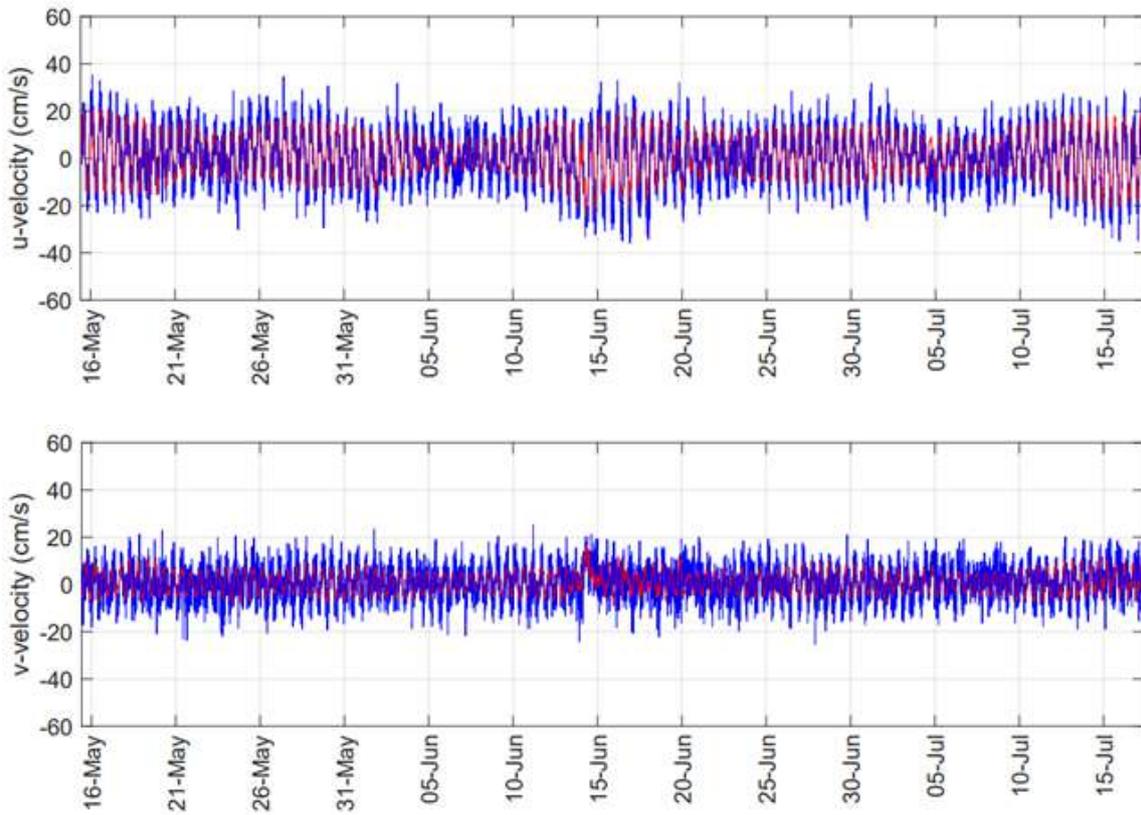
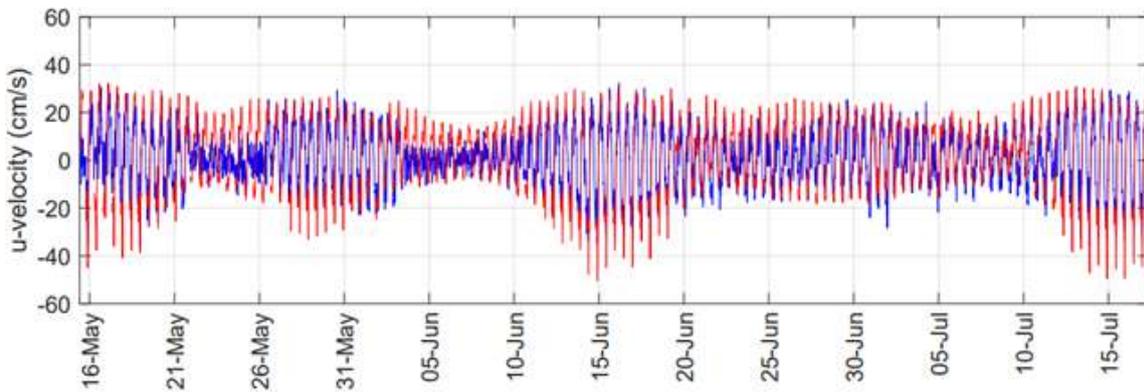


Figure 5.5 Observed (blue) and modelled (red) time series of the currents at the ADCP B mooring from spring-summer 2018: u-component (top) and v-component (bottom) at the surface level are shown.



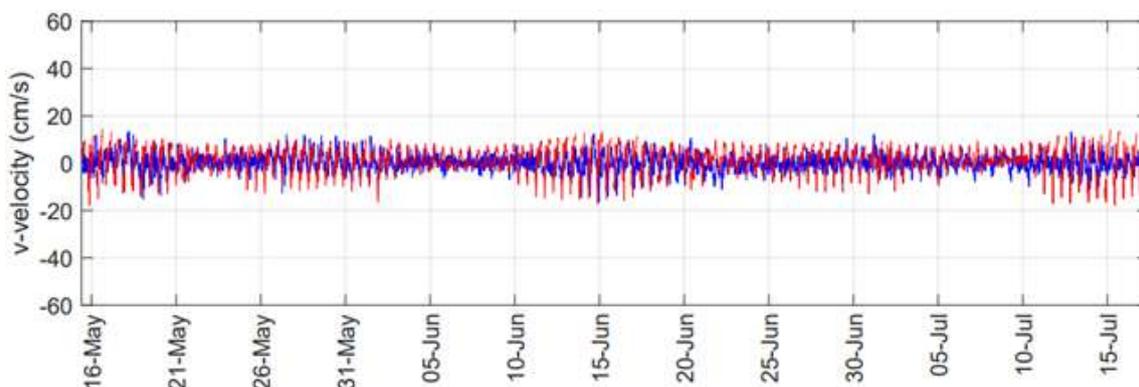


Figure 5.6 Observed (blue) and modelled (red) time series of the currents at the ADCP C mooring from spring-summer 2018: u-component (top) and v-component (bottom) at the surface level are shown.

2.5.1.3 CPT loggers

Two CPT logger moorings providing time series of water temperature and salinity are operated by Cuan Beo. These measurements are particularly interesting since the sensors are located at farming sites in Killeenaran (53°11'52.2"N 08°56'32.2"W) and Kinvara (53°09'36.0"N 08°57'06.1"W). Unfortunately, only the temperature series are applicable for the purpose of model validation, since it has become clear that the salinity measurements from these stations are no longer representative of the actual salinity conditions. Instead, weekly measurements from a CO 310 hand-held sensor are presented for the Killeenaran farming area. Same as the CPT loggers, this sensor also provides both temperature and salinity measurements.

2.5.1.3.1 Killeenaran farming site

The temperature seasonal cycle and small-scale fluctuations are accurately predicted by the ROMS Galway Bay model (Fig. 5.7), in particular when compared to the CPT logger temperature measurements, which yields an excellent correlation (CORR = 0.988) and a low error (RMSD = 0.611 °C). Although the error is higher when compared to the CO 310 sensor (RMSD = 1.106 °C), the correlation is still high (CORR = 0.966).

Table 5.3 Killeenaran observed temperature vs. ROMS Galway Bay predictions statistics (ME = Mean Error, MAE = Mean Absolute Error, RMSD = Root Mean Squared Difference, CORR = correlation, n = number of data points).

	ME	MAE	RMSD	CORR	n
CPT logger	-0.123 °C	0.469 °C	0.611 °C	0.988	7024
CO 310	-0.431 °C	0.746 °C	1.106 °C	0.966	52

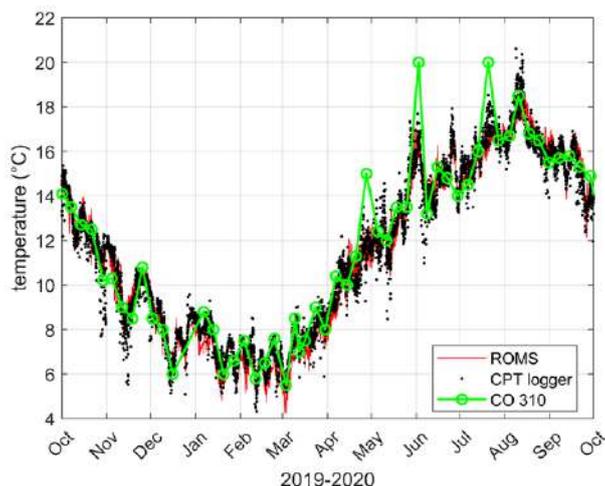


Figure 5.7 2019-2020 temperature time series at Killeenaran (53°11'52.2"N 08°56'32.2"W) according to the Galway Bay model (red), the Killeenaran CPT logger (dots) and the weekly measurements at the Killeenaran pier with the CO 310 hand-held sensor (green)

Regarding salinity, the wintertime freshening is reproduced by the ROMS Galway Bay model. As far as coastal modelling is concerned, this is probably the most challenging process in the Galway Bay, with very dynamical salinity fields affected, on the one hand, by the inflow and outflow of the tides and, on the other hand, by the rainfall and the numerous freshwater sources that discharge into the bay. In addition, submarine groundwater discharges flow from the southern shore due to the existence of a karst system.

The model accurately predicts the measurements obtained with the CO 310 sensor at the Killeenaran pier (Fig. 5.9), reproducing even short-scale variations of salinity that occur due to the effect of heavy rainfall events. This high accuracy in reflecting short-term salinity fluctuations can be appreciated through the high correlation value (CORR = 0.913). On the other hand, the large mismatch between the CPT logger and the CO 310 sensor make it clear that the CPT logger salinity measurements have largely drifted and are thus of no interest for the purpose of validating the Galway Bay model.

Table 5.4 Killeenaran observed salinity (CO 310 sensor) vs. ROMS Galway Bay predictions statistics (ME = Mean Error, MAE = Mean Absolute Error, RMSD = Root Mean Squared Difference, CORR = correlation, n = number of data points).

ME	MAE	RMSD	CORR	n
1.386	2.023	2.706	0.913	52

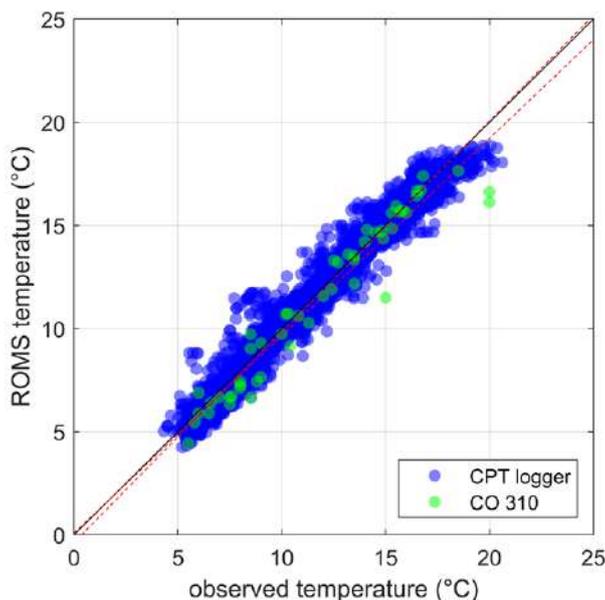


Figure 5.8 CPT logger and CO 310 temperature measurements at Killeenaran (53°11'52.2"N 08°56'32.2"W) vs Galway Bay model predictions. CPT logger: $y = -0.370 + 1.021x$; $R^2 = 0.976$. CO 310: $y = 0.077 + 0.957x$; $R^2 = 0.934$.

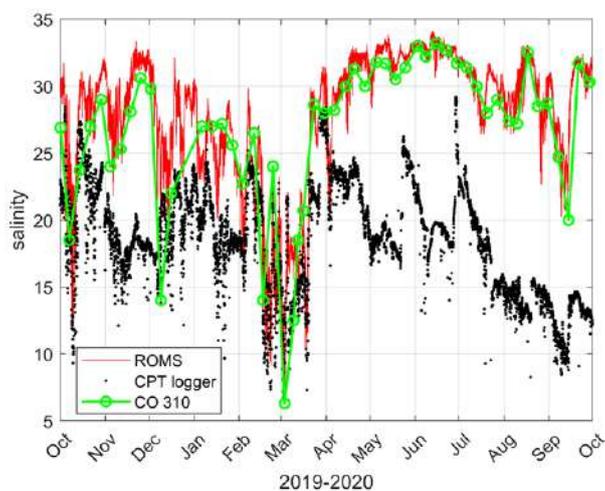


Figure 5.9 2019-2020 salinity time series at Killeenaran (53°11'52.2"N 08°56'32.2"W) according to the ROMS Galway Bay model, the Killeenaran CPT logger and the weekly measurements at the Killeenaran pier with the CO 310 hand-held sensor.

2.5.1.3.2 Kinvara site

Similarly as in Killeenaran, there is an excellent match between the ROMS Galway Bay model estimations and the Kinvara temperature measurements, not just in the ability of the model to reproduce the seasonal cycle, but also in the capacity to predict short-scale fluctuations ($CORR = 0.994$, see Table 5.5). Surface and vertically-averaged temperature series from the model, together with the observed dataset are shown (Fig. 5.11), and it can be easily appreciated that the difference between the observed series and the model vertically-averaged series is small ($RMSD = 0.486$ °C). The ROMS Galway Bay model surface series is noisier, reflecting the natural higher dependency of the surface layer on the diurnal air temperature oscillations. In summary, there is a strong agreement between

both the ROMS Galway Bay model and the *in situ* observations, and this is reflected by a regression line that is very close to that of a perfect agreement (Fig. 5.12).

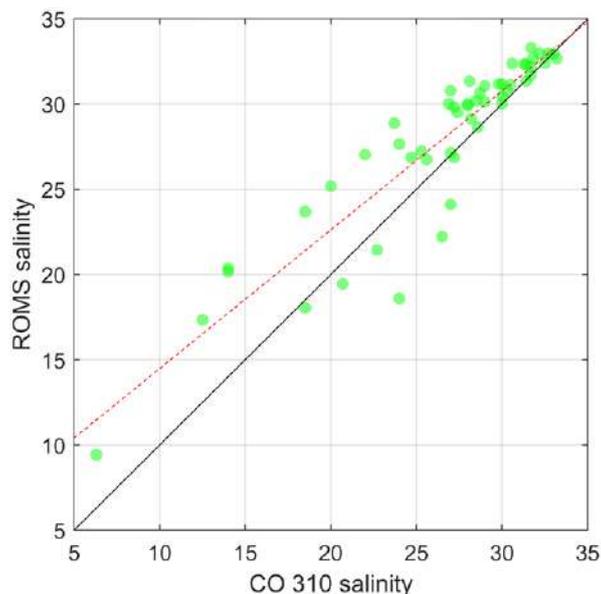


Figure 5.10 CO 310 salinity measurements at Killeenaran (53°11'52.2''N 08°56'32.2''W) vs ROMS Galway Bay model predictions. $y = 6.348 + 0.814x$; $R^2 = 0.834$.

Table 5.5 Kinvara West observed temperature vs. ROMS Galway Bay predictions statistics (ME = Mean Error, MAE = Mean Absolute Error, RMSD = Root Mean Squared Difference, CORR = correlation, n = number of data points).

ME	MAE	RMSD	CORR	n
-0.179 °C	0.405 °C	0.486 °C	0.994	8555

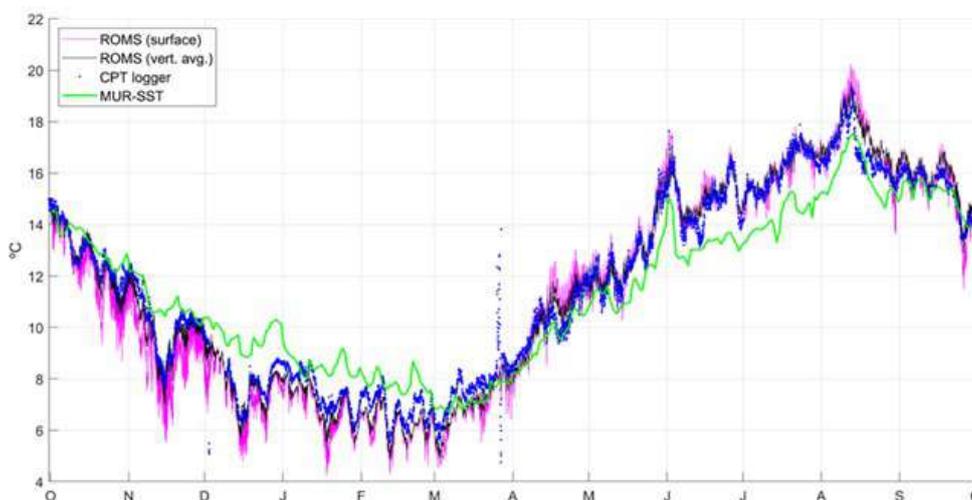


Figure 5.11 2019-2020 temperature time series at Kinvara (53°09'36.0''N 08°57'06.1''W) according to the ROMS Galway Bay model (black: vertically-averaged, pink: surface), the Kinvara CPT logger (blue) and the MUR-SST remote-sensing product (green).

In addition, the temperature time series from the 1-km resolution Multi-Scale Ultra-High Resolution Sea Surface Temperature remote sensing product (Chin et al., 2017) was extracted for the nearest

node and it is shown alongside the modelled and in-situ time series (Fig. 5.11). It is clearly seen that there is a much larger mismatch between the remote-sensing dataset and the CPT logger data. The remote sensing product estimates warmer temperatures during the wintertime and colder waters during the summertime. It is possible that these differences are motivated by the fact the remote-sensing is providing the *foundation* temperature, or the temperature of the surface layer of the ocean devoid of diurnal oscillations, and this may differ from those measurements provided by the *in situ* instrument. On the other hand, remote-sensing products may not be accurate in such near-coastal waters, and this is why the MUR-SST product was not further applied on the validation of the Galway Bay model.

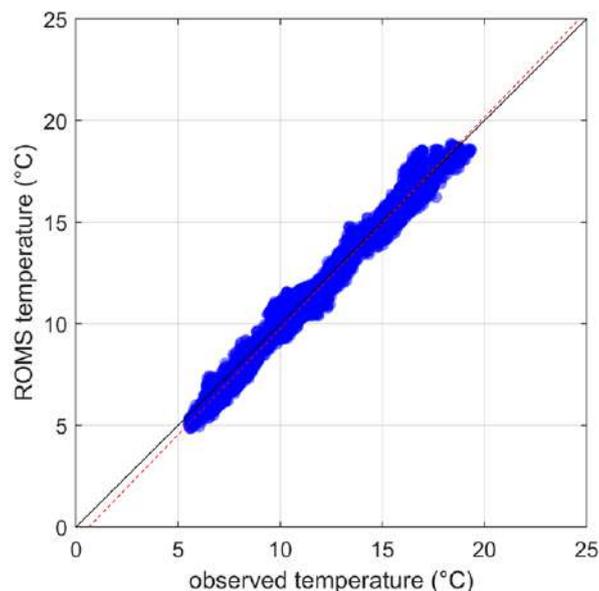


Figure 5.12 CPT logger temperature measurements at Kinvara (53°09'36.0"N 08°57'06.1"W) vs Galway Bay model predictions. $y = -0.657 + 1.041x$; $R^2 = 0.987$.

2.5.2 Forecast validation

To carry out the forecast validation activities required within the FORCOAST project, 3-day forecasts have been stored from January 2021 until present. In this section, quality of forecasts is compared across the length of the forecast period.

2.5.2.1 CTD casts

Quarterly CTD samplings have been carried out by the Marine Institute throughout the Galway Bay using the CastAway CTD. Samplings were carried out on 18-May, 18-Aug and 05-Nov 2021, and about 30 stations were sampled each time. Here, the temperature and salinity sampled profiles closest to the Killeenaran farming site (Fig. 5.1) are shown.

The procedures presented in this section provide an example of *forecast validation* –in contrast to the classical, hindcast validation–, where the performances of three different forecasts for the same day are discussed. These forecasts are named as follows:

- F-1: latest forecast, produced on the same day.
- F-2: forecast produced the day before.
- F-3: forecast produced two days before.

2.5.2.1.1. Spring sampling: 18th of May 2021

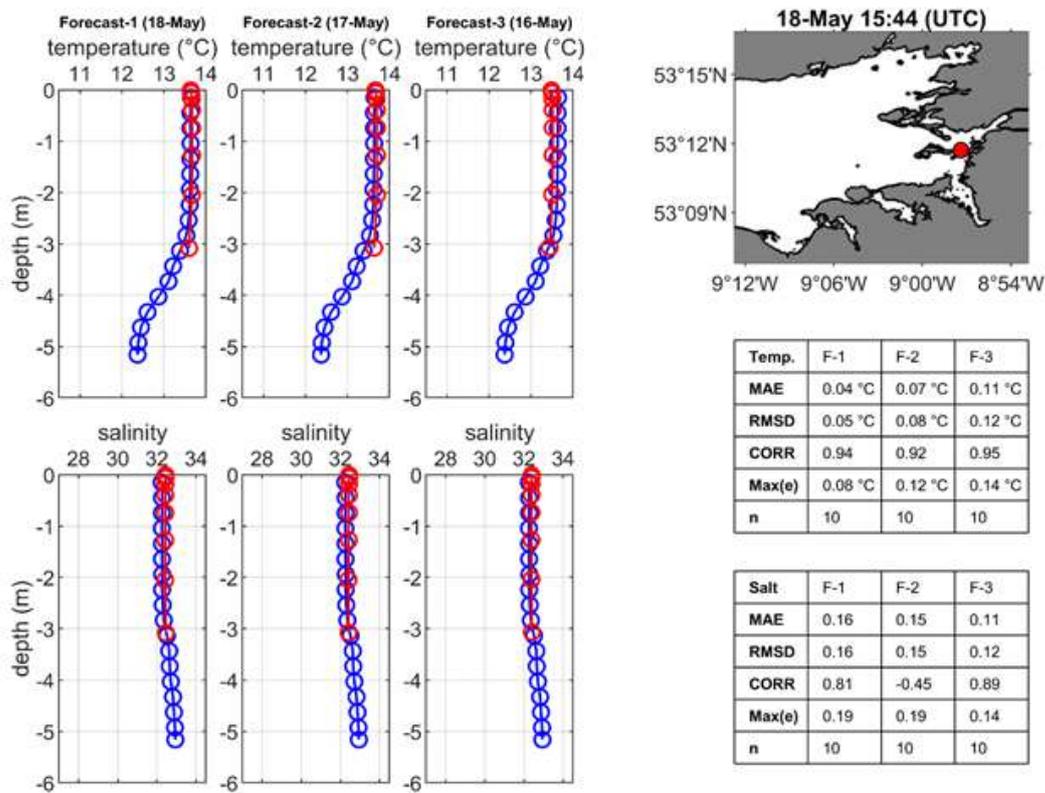


Figure 5.13 CTD sampling on the 18th of May 2021 near Killeenaran. Observed (blue) and Galway Bay model forecasts (F-1, F-2, F-3) for temperature and salinity (red) are shown.

The model successfully captures the temperature and salinity conditions for the 18th of May, with a maximum error in temperature of only 0.14 °C for forecast F-3, and a maximum error in salinity of 0.19 for forecasts F-1 and F-2. The performances of the three forecasts are very similar (Fig. 5.13).

2.5.5.1.2. Summer sampling: 18th of August 2021

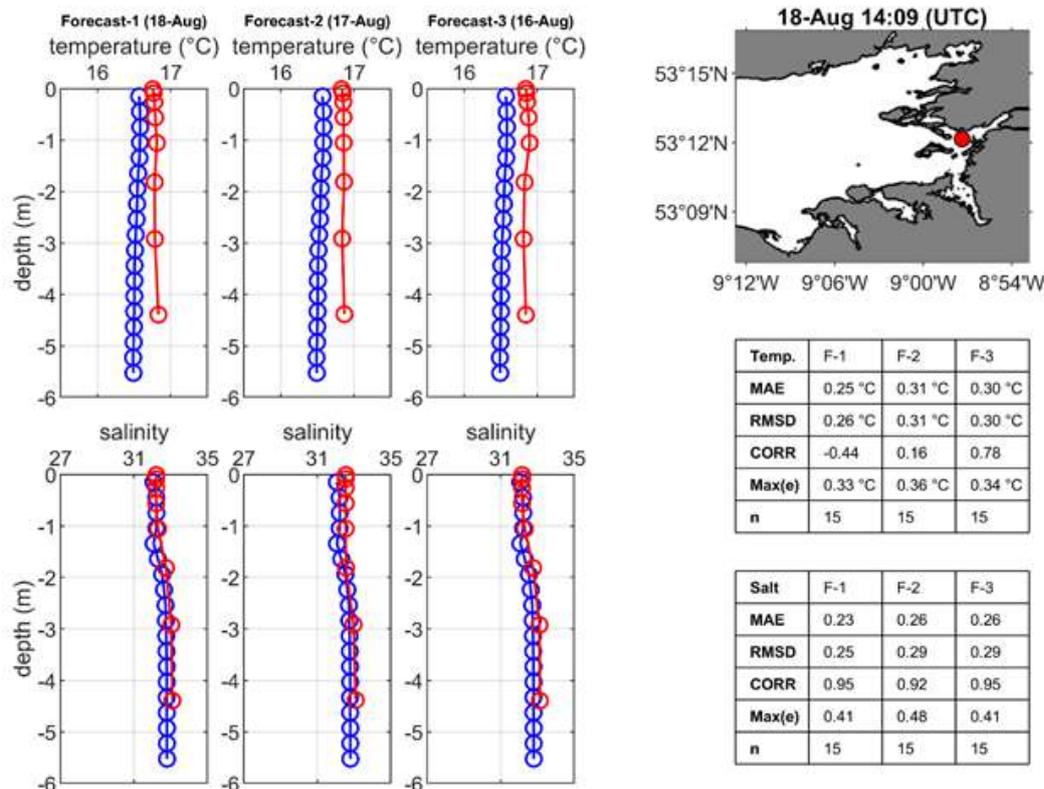


Figure 5.14 CTD sampling on the 18th of August 2021 near Killeenaran. Observed (blue) and Galway Bay model forecasts (F-1, F-2, F-3) for temperature and salinity (red) are shown.

In the summer sampling, the model slightly overestimates the temperature profile at Killeenaran with a maximum error of 0.36 °C for F-2. The salinity profile is accurately predicted, with a maximum error of 0.48 for F-2 (Fig. 5.14), which is probably good enough given that this is an area where large spatial and time gradients of salinity develop under the effect of tides, freshwater runoff and rainfall.

2.5.5.1.3. Autumn sampling: 5th of November 2021

Finally, the autumn conditions proved to be the most challenging for the model, with large salinity gradients across the bay as a result of rainy weather. The maximum temperature error is 1.00 °C for F-2 and maximum salinity error is 2.18 at the Killeenaran profile (Fig. 5.15).

2.5.5.1.4. Forecasts comparison

Here, the performance of the three forecasts is compared for both temperature (Fig. 5.16) and salinity (Fig. 5.17), and considering the total number of observations recorded by the instrument, i.e. unlike the sections above, not only the station nearest to Killeenaran is shown here, but the whole temperature and salinity observations at every station and at all depths.

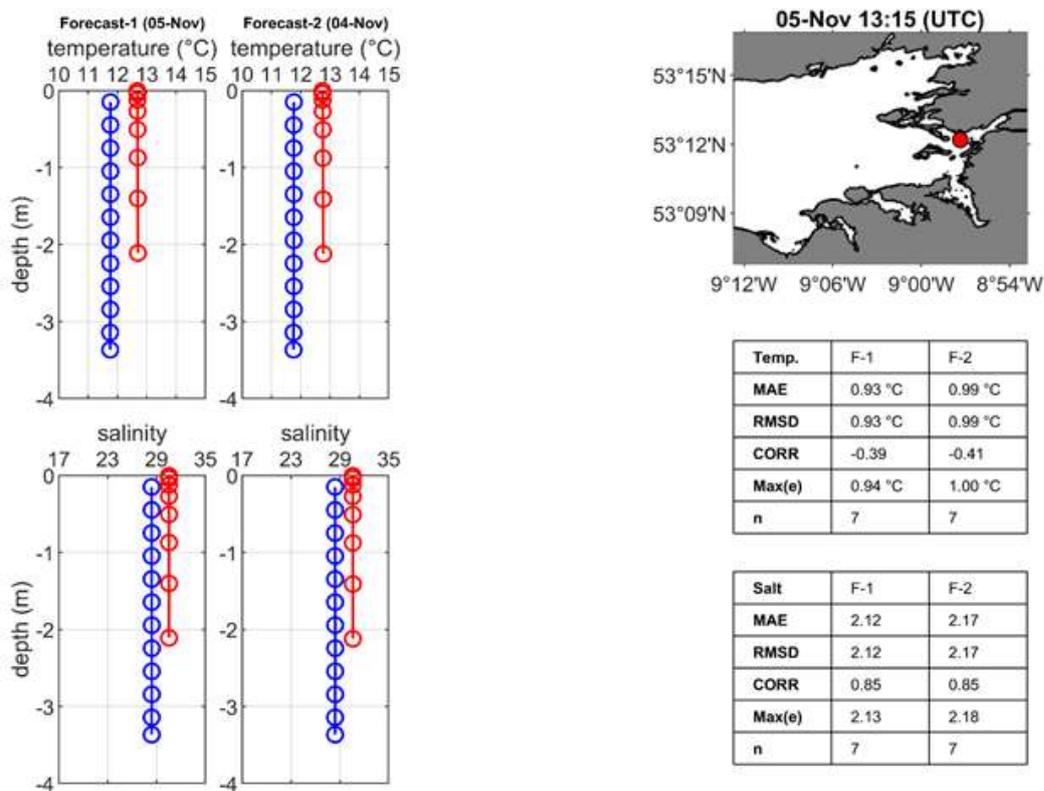


Figure 5.15 CTD sampling on the 5th of November 2021 near Killeenaran. Observed (blue) and Galway Bay model forecasts (F-1, F-2) for temperature and salinity (red) are shown. F-3 was not produced for the 5th of November.

Again, the performance of the temperature forecasts is approximately the same, with very subtle differences in the quality of the predictions (Table 5.6). Both the spring and autumn samplings reveal a slight deterioration of the forecast quality from F-1 to F-3, with a slim increase of the error and decrease in the correlation. This is not observed in the summer sampling, with the F-2 forecast having the highest correlation. In any case, these differences are so small that it can be assumed that the forecasts are of approximately the same quality.

The same general conclusion can be derived when examining the performance of the salinity forecasts, where the correlation is similar among the three forecasts in spring ($0.81 < CORR < 0.88$), summer ($0.82 < CORR < 0.84$) and autumn ($0.88 < CORR < 0.89$).

Table 5.6 Observed vs. modelled temperature statistics (*ME* = Mean Error, *MAE* = Mean Absolute Error, *RMSD* = Root Mean Squared Difference, *CORR* = correlation, *n* = number of data points) for the CTD surveys conducted in Galway Bay in spring, summer and autumn and for the three forecasts (F1, F2, F3) available for each day.

	Spring			Summer			Autumn	
	F1	F2	F3	F1	F2	F3	F1	F2
ME (°C)	0.016	0.019	-0.064	0.045	0.064	0.018	0.492	0.479

MAE (°C)	0.244	0.252	0.267	0.135	0.144	0.141	0.553	0.549
RMSD (°C)	0.299	0.309	0.324	0.179	0.194	0.189	0.724	0.738
CORR	0.913	0.907	0.906	0.613	0.647	0.631	0.859	0.837
n	818	817	817	1150	1150	1150	854	855

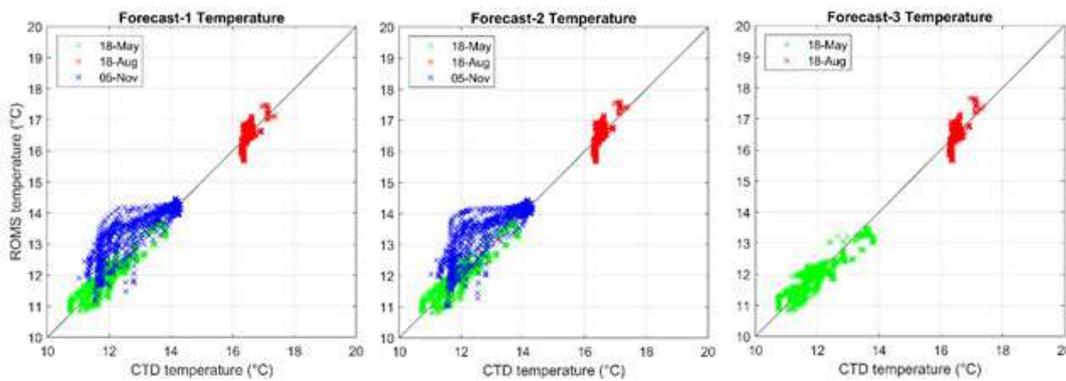


Figure 5.16 CTD vs. ROMS Galway Bay model temperature for forecasts F-1 (left), F-2 (center) and F-3 (right) taking the three samplings (spring, summer, autumn), every station and all depths into account.

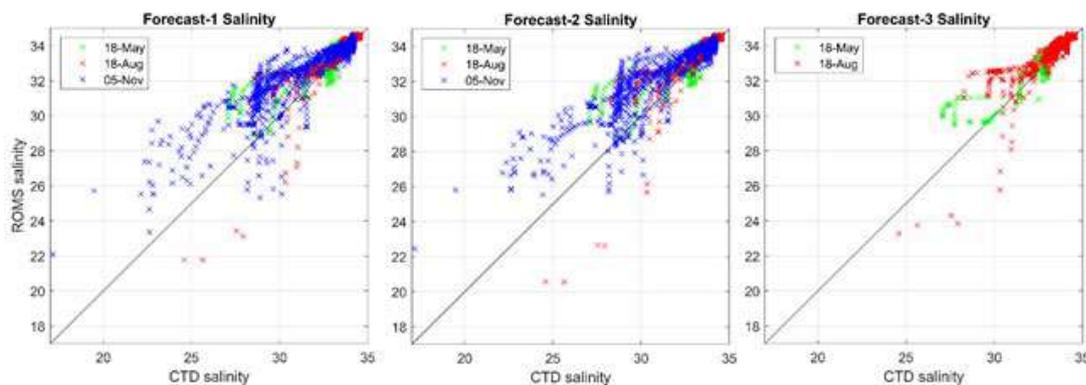


Figure 5.17 CTD vs. ROMS Galway Bay model salinity for forecasts F-1 (left), F-2 (center) and F-3 (right) taking the three samplings (spring, summer, autumn), every station and all depths into account.

Table 5.7. Observed vs. modelled salinity statistics (*ME* = Mean Error, *MAE* = Mean Absolute Error, *RMSD* = Root Mean Squared Difference, *CORR* = correlation, *n* = number of data points) for the CTD surveys conducted in Galway Bay in spring, summer and autumn and for the three forecasts (F1, F2, F3) available for each day

	Spring			Summer			Autumn	
	F1	F2	F3	F1	F2	F3	F1	F2

ME	0.341	0.345	0.289	0.355	0.358	0.390	1.160	1.121
MAE	0.479	0.490	0.436	0.478	0.485	0.502	1.337	1.297
RMSD	0.767	0.782	0.654	0.725	0.743	0.736	1.823	1.777
CORR	0.826	0.817	0.878	0.830	0.824	0.827	0.881	0.889
n	818	817	817	1150	1150	1150	854	855

3.5.2.2 CPT loggers

2.5.2.2.1 Killeenaran farming site

Similarly as with the hindcasts, the performance of the three forecasts in reproducing the temperature series at Killeenaran (Fig. 5.18) is excellent ($CORR = 0.986$) and remains the same during the whole forecast length, with only a slight increase in the error when switching from F-1 to F-3 (Table 5.8).

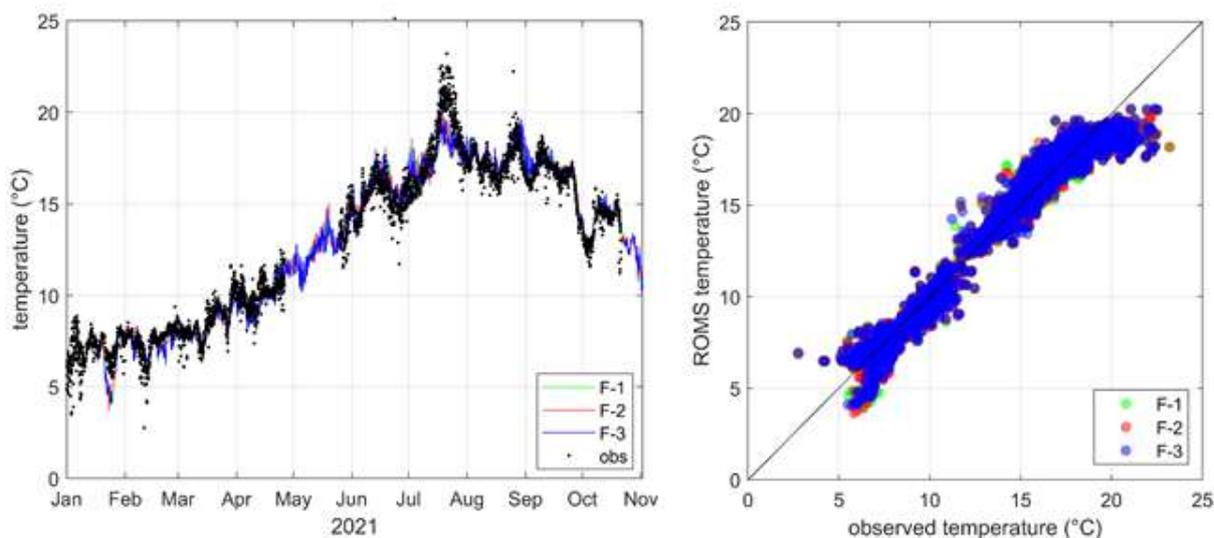


Figure 5.18 CPT logger vs. ROMS Galway Bay model temperature at Killeenaran ($53^{\circ}11'52.2''N$ $08^{\circ}56'32.2''W$) for forecasts F-1, F-2 and F-3.

Table 5.8 Observed vs. modelled temperature statistics (ME = Mean Error, MAE = Mean Absolute Error, $RMSD$ = Root Mean Squared Difference, $CORR$ = correlation, n = number of data points) for the three CPT loggers in Galway Bay and for the three forecasts (F1, F2, F3) available for each day.

	Killeenaran			Kinvara			Kinvara East		
	F1	F2	F3	F1	F2	F3	F1	F2	F3

ME (°C)	-0.045	-0.050	-0.049	-0.039	-0.038	-0.044	0.051	0.053	0.051
MAE (°C)	0.496	0.501	0.515	0.356	0.364	0.374	0.358	0.366	0.372
RMSD (°C)	0.704	0.706	0.717	0.515	0.515	0.516	0.538	0.544	0.540
CORR	0.986	0.986	0.986	0.991	0.991	0.991	0.978	0.978	0.978
n	4518	4538	4549	6387	6400	6400	4463	4468	4468

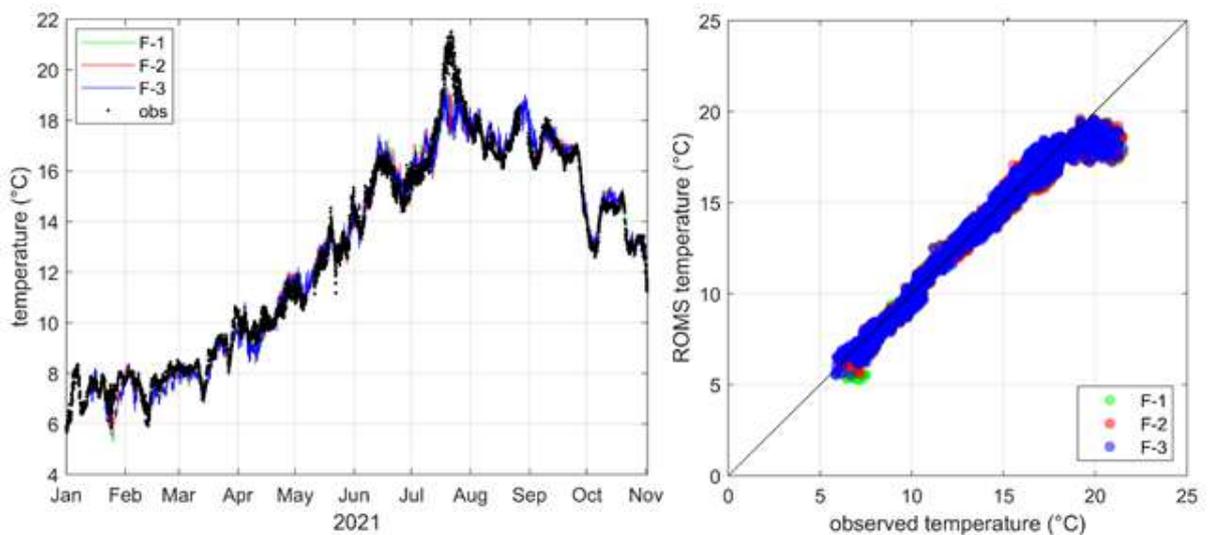


Figure 5.19 CPT logger vs. ROMS Galway Bay model temperature at Kinvara (53°09'36.0"N 08°57'06.1"W) for forecasts F-1, F-2 and F-3.

2.5.2.2.2 Kinvara farming site

Starting from April 2021, a new sensor was deployed at Kinvara Bay, here referred to as Kinvara East (53°09'48.24"N 08°56'59.28"W). The three forecasts accurately represent the temperature series at both stations in Kinvara, with CORR = 0.991 at Kinvara, and CORR = 0.978 at Kinvara East. Examination of Figures 5.19 and 5.20 reveals a nearly perfect agreement between the model and the observed data.

3.5.3 Process-oriented validation

The Receiver Operating Characteristic (ROC) curve analysis, as illustrated in D5.3, Section 2.5, is applied here using the CO 310 salinity measurements at the Killeenaran pier (Fig. 5.21). A total of 313 data points were used. Fig. 5.21.a shows the procedure for the particular case where the salinity

threshold is 30. When the threshold is displaced along the whole range of values, the ROC curve is obtained (Fig. 5.21. b). The area below the curve is 0.906, which reflects an excellent performance of the model regarding its ability to reproduce the observed salinity at Killeenaran.

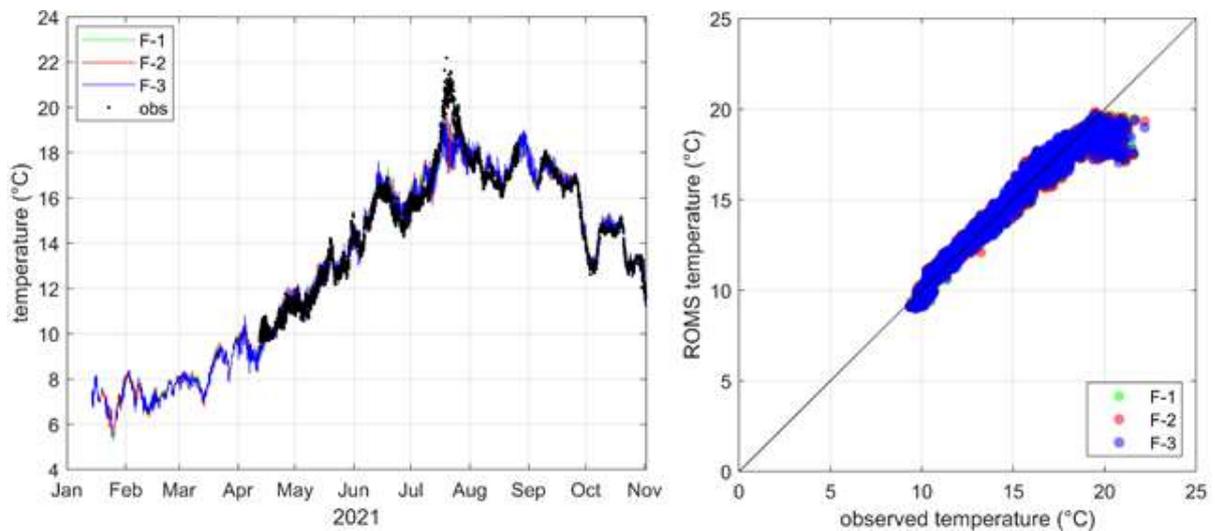


Figure 5.20 CPT logger vs. ROMS Galway Bay model temperature at Kinvara East (53°09'48.24''N 08°56'59.28''W) for forecasts F-1, F-2 and F-3.

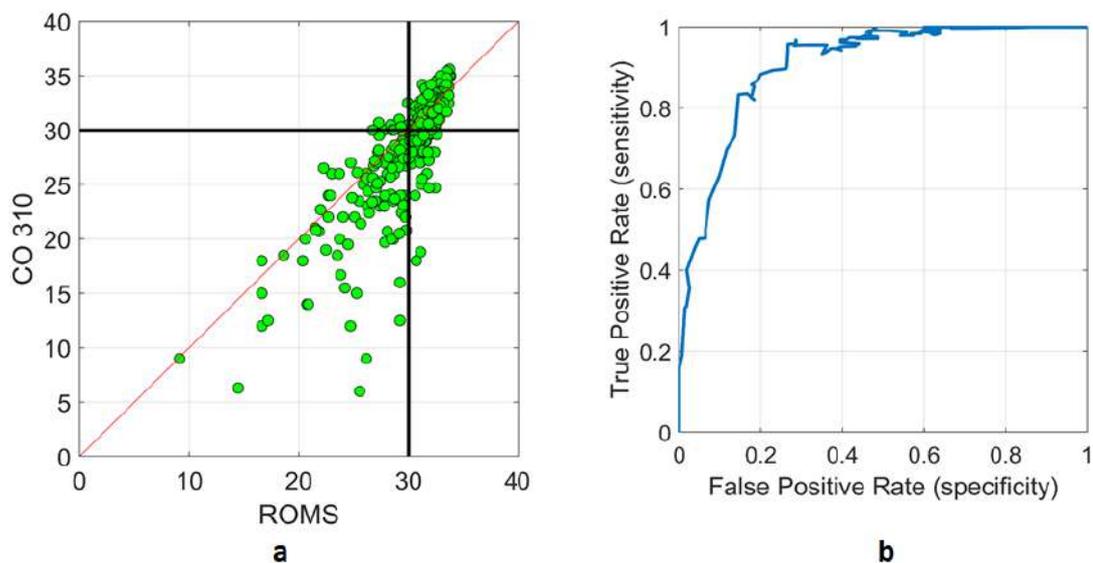


Figure 5.21 (a) ROMS Galway Bay model vs. CO 310 weekly measurements of salinity at Killeenaran (53°11'52.2''N 08°56'32.2''W), highlighting the threshold $S = 30$. (b) ROC curve analysis.

2.6 Pilot 6: Denmark

Two models were applied at pilot 6, the FlexSem-ERGOM model (Aarhus University) for hindcast validation of biogeochemistry and HBM (DMI) for hindcast validation of physics. The FlexSem-ERGOM model provides data to the SM A3 site prospectation and HBM to SM A1 Marine Conditions. HBM provided initial data, surface temperature, and boundary conditions of physics to FlexSem.

The hydrodynamical-ecological model FlexSem-ERGOM for the Limfjord, Denmark, was run for the period 2009 to 2017. We consider the first two years spin-up and carry out the model validation for the years 2011 to 2017. Observed values for nutrients (Dissolved inorganic nitrogen ($DIN = NO_x + NH_4$) and dissolved inorganic phosphate, (PO_4), chlorophyll and bottom oxygen concentrations were downloaded from the Danish National Database (www.odaforalle.au.dk) for stations distributed across the fjord (Figure 1). Observations exist for all seasons of the year and for all years in the modelled period (2011 to 2017). However, a larger number of stations were available for oxygen validation.

Model performance of the FORCOAST physical ocean model HBM has been evaluated in Limfjord for historical periods 2015-2019 and the operational period, since commissioning of the model, 17th of March 2021 to now, for forecast validation. Process-oriented performance evaluations for extreme events are carried out using historical data-set 2015-2019, for better time coverage. The physical ocean parameters include sea level, salinity and temperature. The following document describes the observation data sets used for model validation and presents the results of the model validation study, for historical, forecasting and process-oriented assessments.

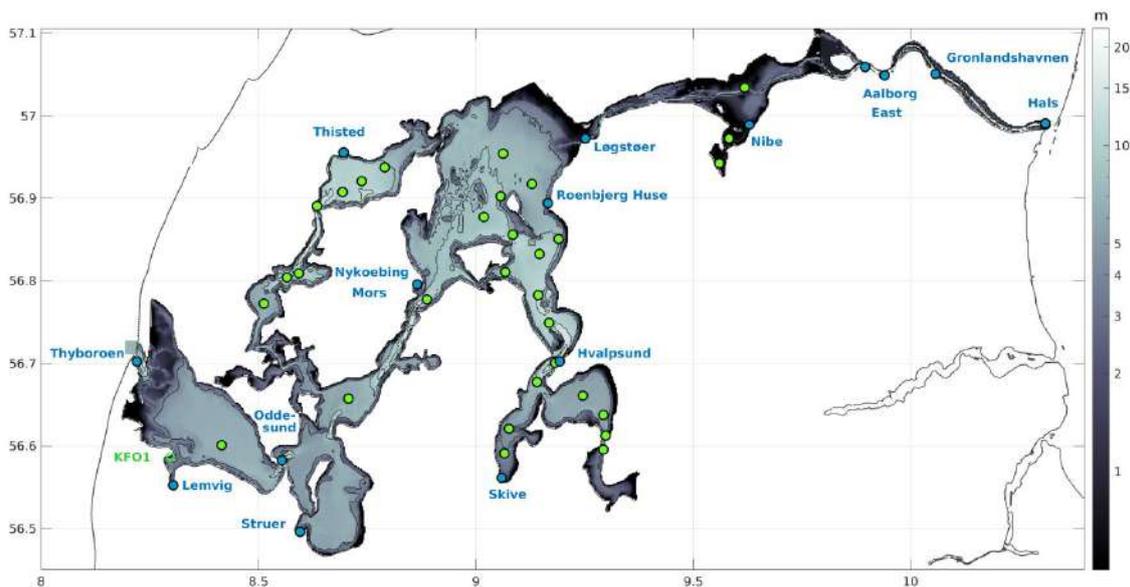


Figure 6.1 Limfjord in-situ observation stations for sea level: tide gauge stations (blue circles), salinity and temperature regular cruise data (green filled circles) and temperature observations (green circle) at aquafarming site KFO1 (Kulturfelt Follup Odde).

Model validation assessment of historical periods 2015-2019 covers the entire Limfjord to provide a better overall picture of the model performance, whereas the forecast assessments and process-oriented assessments focus on the Aquafarming site Kulturfelt Follup Odde (KFO1) in the western Limfjord and nearby stations in the Odde-sund (salinity and temperature) and Lemvig (sea level) (Fig. 6.1).

Table 6.1 List of data sets used for model validation

Parameter	Type of Observations	Position of the monitoring station nearest to the aquafarming site	Number of stations	Period	Frequency	Source
Sea level for hindcast	Tide gauge stations at harbors	Lemvig station 8° 18.925' E 56° 33.15' N	11 stations	2015-2019; 17 th of March 2021 – present	10 minutes	3rd party data from harbor authorities, received through collaboration.
Salinity, Temperature	Profile data from cruises at fixed stations	Oddesund station 8° 24.917' E 56° 36.05' N	30 stations(Fig. 1)	2015-2019	Variable, from weeks to months	NOVANA national monitoring program, from ODA surface water database, Overfladenvandsdatabasen: https://odaforalle.au.dk/
Temperature	Profile data at aquafarming site KFO1, Kulturfelt Follup Odde	KFO1 station: 8° 17.815' E 56° 35.2' N	1 station on the northern side of Lem Vig bay.	01.06.2019 to 01.10.2019	Hourly	Monitoring at Aquafarming sites, provided by the Oysterboat company in the FORCOAST project
Sea surface temperature	Satellite	Limfjorden	N/A	2021 – present	daily	SST_BAL_SST_L3S_NRT_OBSERVATIONS_010_032

Denmark operates a dense network of sea level monitoring stations. The Limfjord is covered by 15 stations, which are run by local harbour authorities. The data is provided to DMI in real-time. Eleven of these stations were selected for statistical analysis (Table 6.1).

Salinity and temperature profiles from NOVANA national environmental monitoring program cruises have been made available by the Danish Center for Environment DCE at the Aarhus University, in collaboration with the Danish Ministry for Environment and Food. Time range and frequency of the observations are variable and change from station to station.

High frequency (hourly) temperature data from a shallow water monitoring station KFO1 at aquafarming site: Kulturfelt Follup Odde, near Lemvig has been provided by the Oysterboat Company in an offline mode. The depth of the surface sensor is at 20cm depth and the depth of the bottom sensor is approximately at 2.3 m depth, 20 cm above the seabed. The model depth at the site is only 1.4 m. The data set used for validation covers the period 3rd of June 2019 to 1th of October 2019.

Satellite data sets: CMEMS provides both L3 and L4 SST products in the Limfjorden, in a horizontal resolution of 0.02 degree and daily temporal resolution. The plan is to inter-compare the CMEMS L3 SST product with in-situ observations and forecast products for selected periods.

2.6.1 Hindcast validation

Hindcast validation of HBM has been carried out for a 5-year period 2015-2019, covering all the Limfjord stations (Fig. 6.1). The model data set was used for the site suitability studies. The scope of the assessments includes long-term statistics for sea level, salinity and temperature. Time series

validation has been carried out using salinity and temperature profiles from NOVANA environmental monitoring cruises and sea level observation at tide gauge stations (Fig. 6.1).

The model hindcast was generated from a two-way nested BalticSea-Limfjord-North Sea model, with a 185-meter resolution in Limfjorden. The hindcast model configuration differs from the one used for forecast and process-oriented validation. The hindcast model is 2-way nested into DMI's operational model and not a stand-alone model using CMEMS boundary conditions, as it is the case for the forecast modeling system. The reason is, that only monthly mean reanalysis data sets from CMEMS were available as boundary forcing for the long-term run covering historical periods. This is insufficient for such a dynamic system as the Limfjord, which is strongly influenced by the boundary signals at the entrance to the fjord.

Model quality assessments use methods of time series validation, i.e., comparison of the modelled and observed time series at certain stations. The time series can either be surface values of the sea level anomaly or profile time series of salinity and temperatures at certain depth. Long-term model quality assessments are based on statistical parameters: the model bias, centralized root-mean-square-error (cRMSE) and correlation coefficient. More information is available in Murawski et al. (2021).

2.6.1.1 Sea level validation using tide gauge stations

Denmark's coastlines are particularly vulnerable to floodings generated by storm surges. These are generated by strong, predominantly westerly winds, which raise the waters along the west coast and at the entrance to the Limfjord. In the Baltic Sea, these winds can generate sea level gradients, that may generate strong seiche-like waves, when the winds abate. In the Limfjord, the hydrodynamics during storm surges is characterized by wind and sea level boundary signal driven water transport through the broads and narrows of the fjord. The road of water transport follows the main axis of the fjord, from west to east, through Kås Bredning, east of the island of Møn. The transport capacity of the fjord is limited by the narrow straits leading to larger observed and modelled sea levels in the central Limfjord. While the observed sea level at the entrance to the fjord rarely exceeds 1.6 to 1.7 meters, the maximum sea level at Logstor in the central Limfjord is 2.1 m. Limited transport through Oddeund in the western Limfjord leads to higher maximum observed sea level in Nissum Bredning and Lemvig, of up to 1.95 m at Lemvig tide gauge station.

The main requirements for a good sea level forecast are accurate meteorological predictions, especially of the winds in the interior of the fjord, which are influenced by the higher roughness of the land surface and the adequate prediction of the sea levels at the model boundaries to the North Sea and the Kattegat. The quality of the bathymetry data and the resolution of the model must be high enough, to resolve the transport through the narrow straits of the fjord. The applied 185.2 m horizontal resolution (6'' x 10'') is good enough to resolve most of the straits with more than 1 grid cell across. However, doubling the horizontal resolution to 92.6 m (3'' x 5'') has proven to increase the quality of the sea level predictions, although it is not feasible to do long-term runs with such a high resolution.

The model validation assessment for hindcast validation of HBM focusses on statistical parameters: the centralized Root-Mean-Square-Error (cRMSE) and the correlation coefficient of the modelled time

series in the 5-years hindcast period: 2015-2019. Peak-Error assessments for storm surge events are part of the process-oriented validation and will be presented in Section 2.6.3.

The model validation focuses on sea level anomalies, which are used for operational storm surge validation as well. The standard way to generate the anomalies is to subtract the long-term mean from the model forecasts and observations. For the current assessment, we have chosen to subtract the long-term mean of the 5 years period 2015-2019.

The cRMSE are within acceptable error limits of 10 cm in the interior of the fjord. Largest cRMSE errors occur at boundary station Thyboron Havn in the west (9.4 cm) and Hals (10.08 cm) in the east. Excluding these stations, cRMSE values are lowest in the western fjord (<8.64 cm), including tide gauge station Lemvig (9.18 cm), near the aquafarming site KFO (Fig. 6.1, 6.2). Largest cRMSE value occur in the central and eastern Limfjord, at tide gauge stations near narrow straits, i.e., at Logstør (9.42 cm), near Aggersund and at Gronlandshavn (9.72 cm) near Hals’s channel. The average cRMSE’s is 7.81cm for western stations (Thyboron, Lemvig, Struer, Thisted), 8.57cm for central stations (Nykobing Mors, Skive, Rønbjerg, Logstør) and 9.13 cm for eastern Limfjord stations (Aalborg East, Grønlandshavn, Hals). This includes eastern boundary station Hals, which features the largest cRMSE error of 10.08 cm.

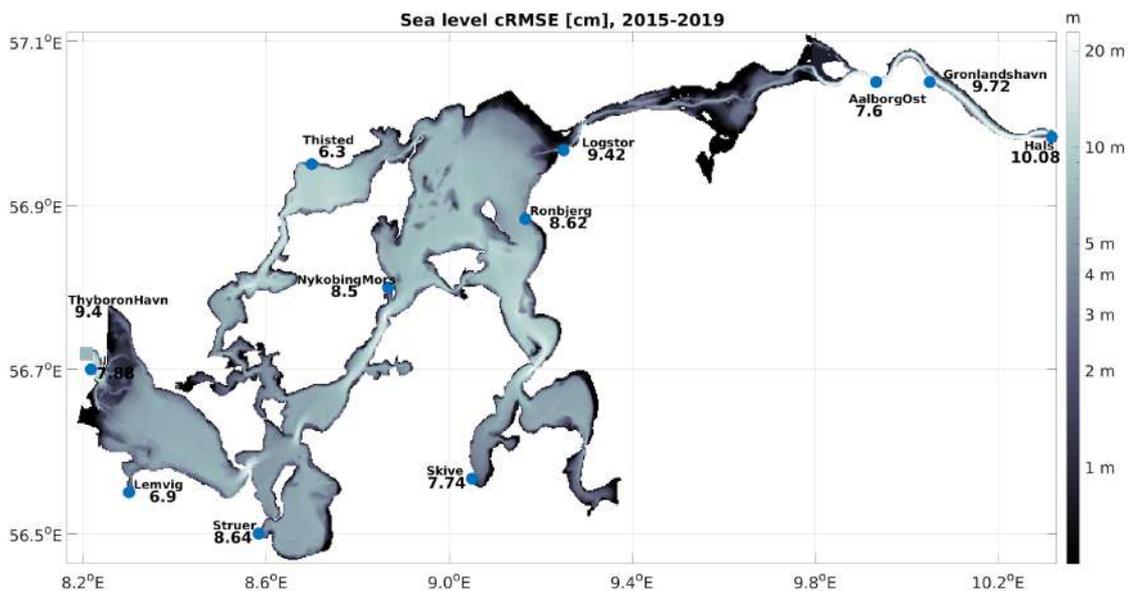


Figure 6.2 Centralized Root-Mean-Square-Error in units of centimeter (cm) at Limfjord tide gauge stations.

The correlation coefficient shows generally high values, above 0.91 in the entire fjord (Fig. 6.3). Lowest values occur near the boundaries, at Thyboron Havn (0.92) and Hals (0.92), and at tide gauge stations near narrow channels, in the central Limfjord, at Logstør (0.91) near Aggersund and in the eastern Limfjord, at Gronlandshavn (0.93) near Hals channel. At Lemvig station, near aquafarming site KFO, the correlation coefficient is relatively high 0.96. This indicates that transport through the narrow straits is modelled adequately, because otherwise, underestimated transport would lead to a lag time

between modelled and observed sea level events, reducing the correlation between the two time series.

2.6.1.2 Temperature and Salinity validation using NOVANA cruise data

Hydrographical conditions in the Limfjord are very dynamic and characterized by a very fast and efficient water exchange with bordering seas, the North Sea in the west and the Kattegat in the East. The water renewal process is controlled by wind forcing, sea level and ocean conditions at the entrances to the fjord (Fig 6.1) and the annual river runoff. The Limfjord does not include any major rivers, but the collected river runoff is enough to replace 34% of the fjords volume (7.72 km³) in an average year. Additional water exchange with the North Sea and with Kattegat, through Thyboron and Hals Channel add to the annual renewal rate. The main water transport is eastward, from the North Sea to the Kattegat, driven mainly by the westerly winds over the North Sea and the Limfjord as well as the sea level boundary conditions in Thyboron and Hals. This affects the fjords hydrography, which is characterized by a strong zonal salinity gradient, from about 32 – 34 psu at the western opening to the North Sea to 19 – 25 psu at the narrow outlet to the Kattegat (Maar et al., 2010).

Figure 6.4 demonstrates the HBM model’s ability to reproduce the variability of salinity and temperature on inter-annual, seasonal and diurnal scales. The three stations at Oddsund, Nykøbing Mors and Nibe Bredning represent the western, middle and eastern Limfjord, respectively. Seasonal and inter-annual variability of water temperature are well reproduced. The model is able to capture the maximum temperatures during summer as well as the minimum temperatures during winter. The performance is nearly identical at the surface and the seabed (Fig 6.4). Shallow thermoclines during summer can be reproduced.

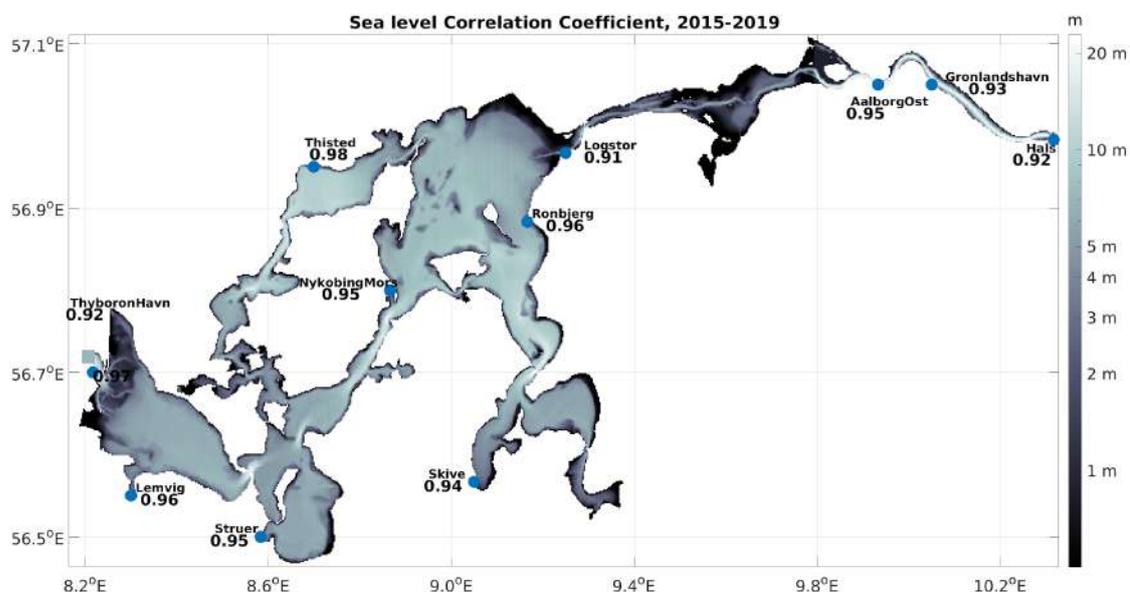


Figure 6.3 Pearson Correlation Coefficient between modelled and observed sea level time series at Limfjord tide gauge stations (2015-2019).

For salinity, the conditions are frequently changing depending on the atmospheric (wind) and transport conditions. There are several events when salinity at Oddesund station drops significantly in April 2019 and November 2019 (Fig 6.2 top). These events are well reproduced. At Nykøbing Mors, in the central Limfjord (Fig 6.4 middle), the variability of salinity is very high. High and low salinities change frequently, depending on the transport conditions. These frequent changes are well reproduced by the model. At station Nibe Bredning, in the eastern Limfjord (Fig 6.4 below), the average conditions are well reproduced. Events of low salinity in April 2019 and November 2019 are captured by the model, even though the extent of the salinity reduction is somewhat underestimated. However, the timing of the events is very well reproduced.

Figure 6.5 shows the horizontal distribution of temperature validation statistics at the surface and in 4.5 m to 5 m depth. The temperature bias is generally negative, with larger values near the surface than at 4.5 m to 5 m depth. Larger negative biases and cRMSE values occur at a very shallow outlet near Nibe bredning in the eastern fjord, which regularly falls dry during lower waters. The cRMSE values are about 0.7 °C to 0.8 °C at this location. High temperature cRMSE values occur also near Virksund, between Lovns bredning and Hjarbæk fjord. These are transport related. At Oddesund station in the western Limfjord, the temperature bias and cRMSE values are in a very acceptable range (Fig 6.5, 6.6).

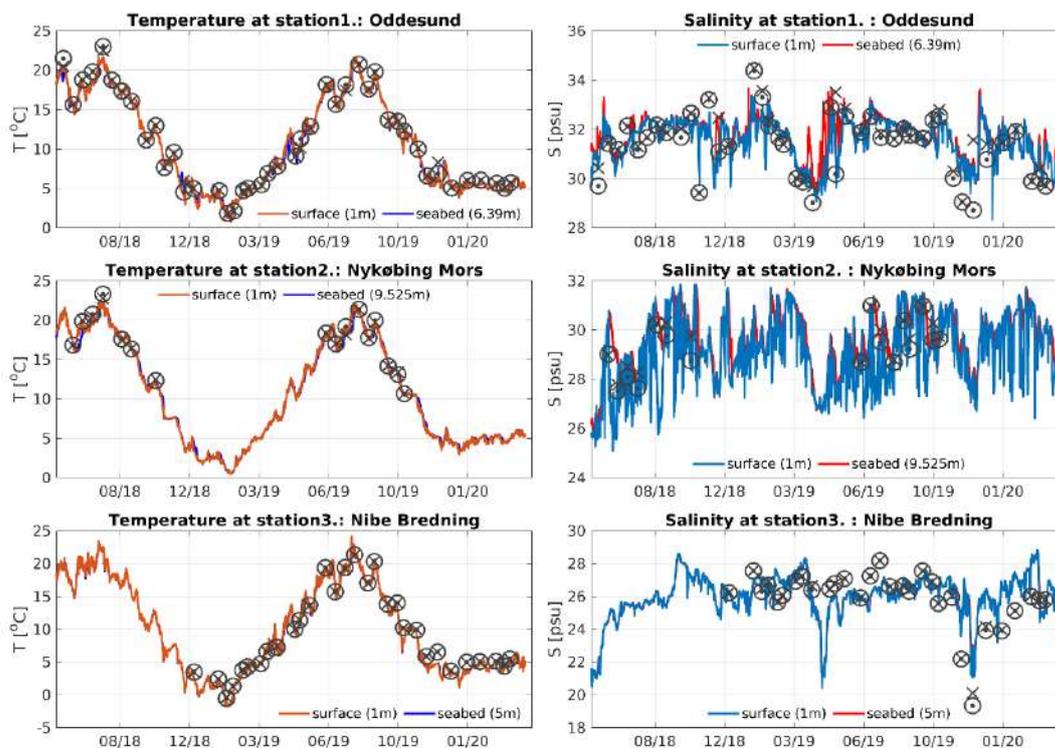


Figure 6.4 Comparison of modelled and observed salinity and temperature dynamic in the western Limfjord (Oddesund), central Limfjord (Nykøbing Mors) and eastern Limfjord (Nibe Bredning), near the surface (red) and near the sea bed (blue). Observations ne

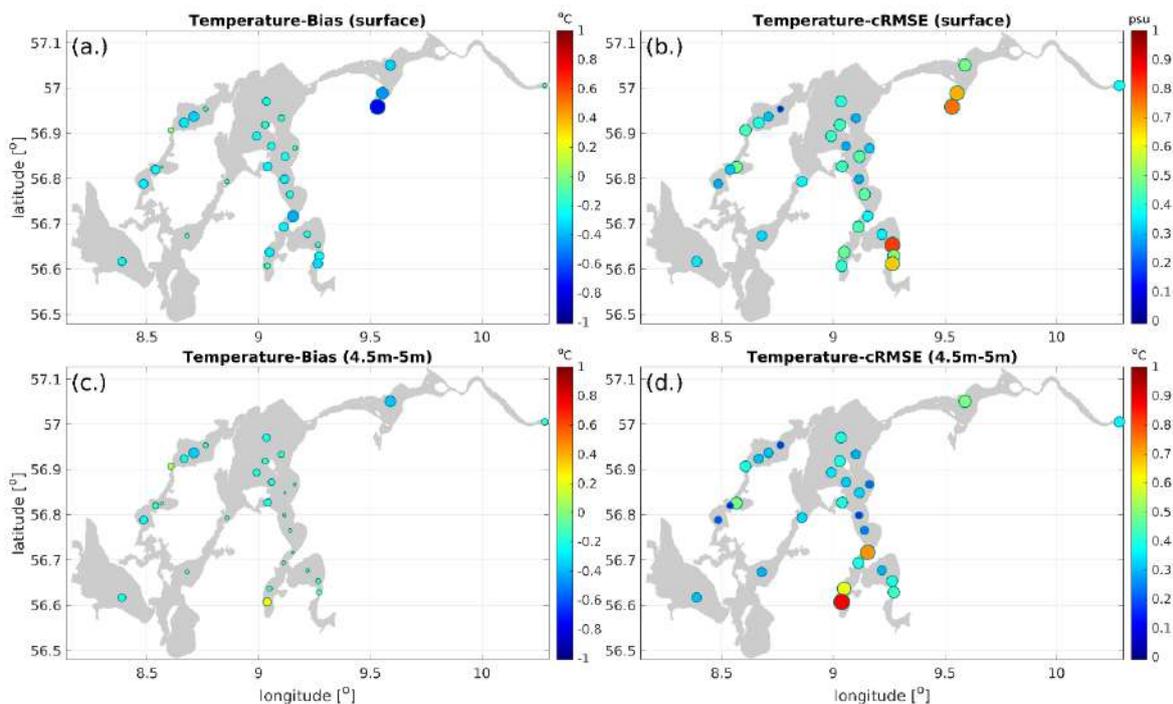


Figure 6.5 Temperature statistics, Bias and centralized Root-Mean-Square-Error for NOVANA environmental monitoring cruise stations (30 stations in the entire Limfjord).

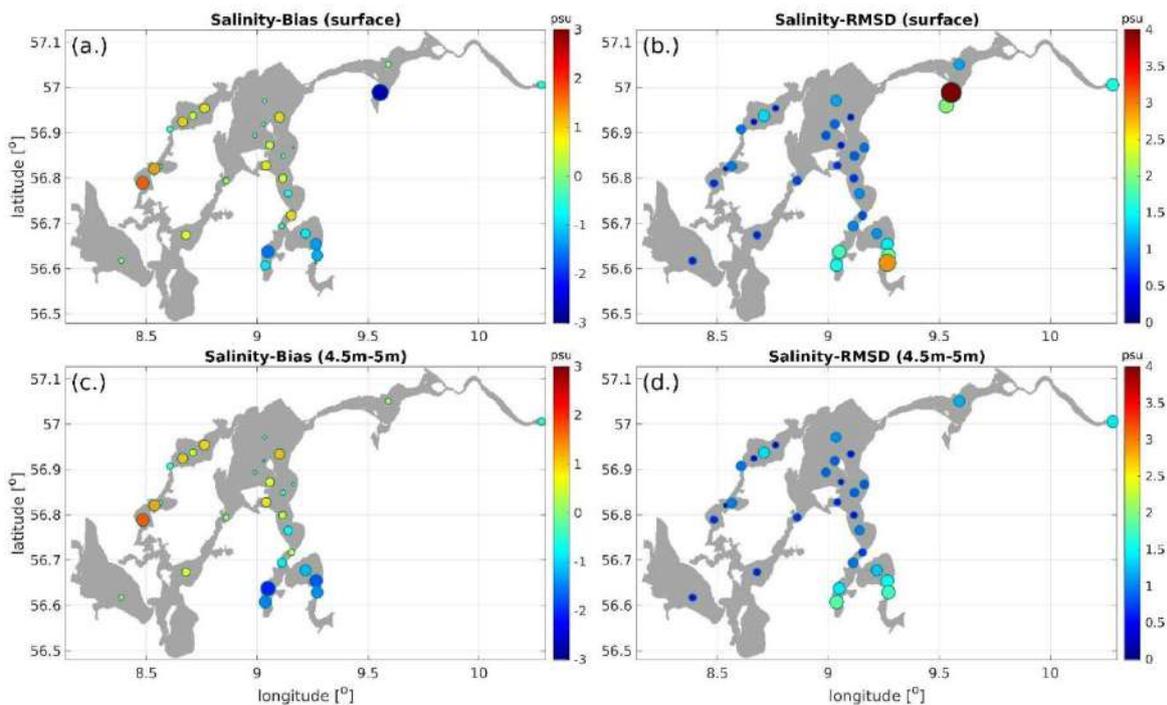


Figure 6.6 Salinity statistics, Bias and centralized Root-Mean-Square-Error for NOVANA environmental monitoring cruise stations (30 stations in the entire Limfjord).

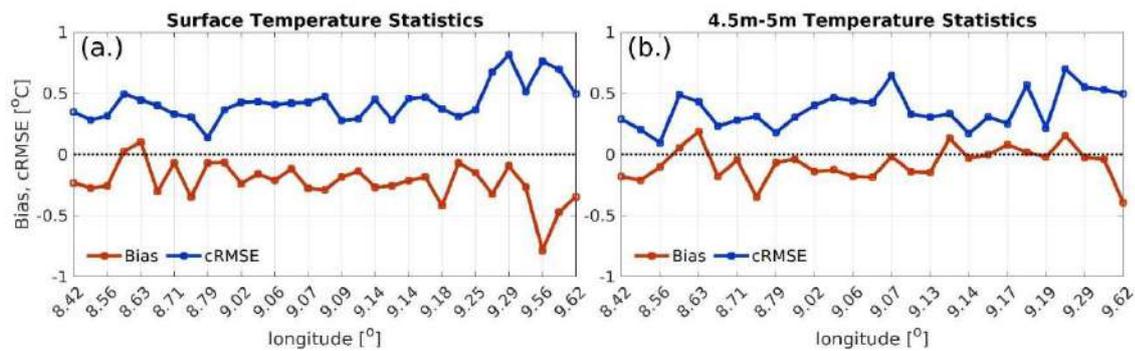


Figure 6.7 Temperature validation statistics: bias and centralized Root-Mean-Square-Error (cRMSE) for all Limfjord stations at the surface (left) and at 4.5m-5m depth (right).

The validation assessment of surface temperatures shows that the Limfjord model features a slight cold bias at most stations (Fig. 6.7a), which leads to a spatial averaged mean surface bias of $-0.2\text{ }^{\circ}\text{C}$ (Fig. 6.8b). Larger biases of $-0.5\text{ }^{\circ}\text{C}$ to $-1.0\text{ }^{\circ}\text{C}$ are found in the area around Sebbersund, which is regularly falling dry during low waters. The centralized Root-Mean-Square-Error (cRMSE) of surface temperature is lower than $0.5\text{ }^{\circ}\text{C}$ except for some stations in the eastern fjord, east of Løgstør. At mid-depth 4-4.5 m, both temperature bias and cRMSE are smaller than at the surface layer (Figs. 6.3b and 6.4b), except for Skive fjord. Averaged cRMSE in the Limfjord is $0.42\text{ }^{\circ}\text{C}$ at the surface and $0.37\text{ }^{\circ}\text{C}$ at 4.5-5 m.

For salinity, the validation statistics shows a general overestimation of salinity in the western and northern part of the fjord (Fig 6.6) and an underestimation in the southern and eastern part of the fjord near Skive and Lovns Bredning as well as near Sæbesund, near Halkær Bredning. These areas are affected by strong river runoff. In general, the cRMSE value is below 1 psu at most stations in the fjord, but at some stations in the southern and eastern part of the fjord, the cRMSE values are larger. On average, the cRMSE value is about 1 psu.

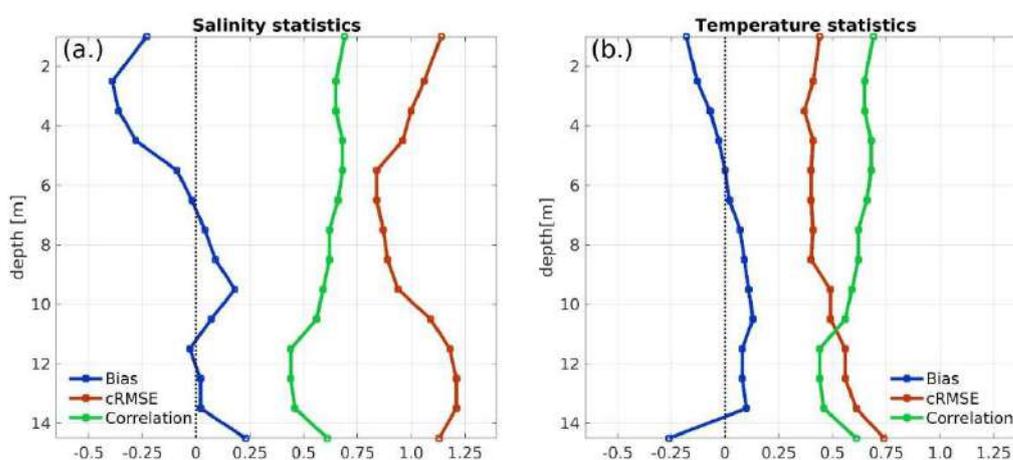


Figure 6.8 Vertical profiles of model validation statistics: bias, centralized Root-Mean-Square-Error (cRMSE), correlation coefficient (CC) of Salinity (a.) and temperature (b.). Shown are the horizontally averaged model validation statistics for all Limfjord stations.

The vertical validation assessment (Fig. 6.8b) for all stations presents vertically averaged model statistics for a number of layers covering a range of up to 14m depth. The model temperatures are colder than the observations in the upper 5 m of the water column. The largest negative temperature bias is found near the surface (-0.2 °C). But at greater depth, below 6 m, the modelled temperatures are slightly warmer than the observations. The cRMSE is almost constant in the upper 8 m. Below 8 m depth, cRMSE increases with depth. The correlation coefficient between the model and observation data is about 0.7 on the surface and decreases with depth (Fig. 6.8b).

The vertical assessment of salinity (Fig. 6.8.a) reveals negative biases in the upper 6m of the water column. At depths of 6m to 11m, the model becomes, on average, saltier than the observations. Below 11m, the average bias is almost zero, but below 14m it slightly increases again. The average cRMSE value of salinity is larger at the surface (1.14 psu) and the bottom (1.21 psu), but it decreases with depth at intermediate layers. The minimum value of 0.84 psu is at the depth of 6 m.

The mean modelled surface nutrient concentrations (DIN and phosphate) are on average relatively high in all areas of the fjord (Figure 6.9 a, b), with the highest concentrations in the Skive-Lovns fjordarm (station 3, Figure 6.9) and in the eastern channel towards the Kattegat, both of which receive significant riverine nutrient loads. Chlorophyll largely follows the same pattern, with elevated values in the Skive-Lovns fjordarm, and lower in areas with a large degree of water exchange with surrounding seas (Figure 6.9 c). The mean bottom oxygen concentration is generally low in areas with high productivity and low degree of ventilation (Figure 6.9 d).

3.6.1.3 Assessment of biogeochemistry in the ecological model ERGOM

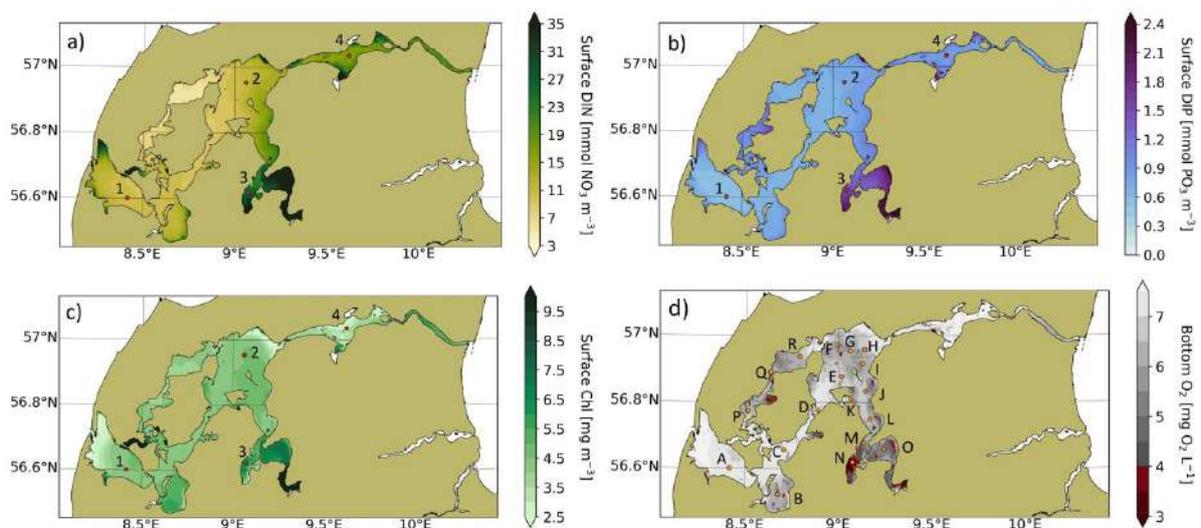


Figure 6.9 Mean modelled surface a) DIN concentration (NO_x + NH₄) b) Phosphate concentration c) Chlorophyll concentration and d) Secchi depth. Red numbered dots mark the stations where measurements were carried out

We assess the validity of the model results by masking out model output to the same time and place as the measured values. Statistics of the spatial distribution in the fjord (Table 6.2) are subsequently calculated using the average value at each station (Figure 6.9). The ME and correlation is low and high, respectively, for all tracers, showing a good fit between model and observations. However, the correlations for chlorophyll are slightly lower than for the nutrients due to the nutrient dependence

of chlorophyll. The RMSE is somewhat high for DIN due to the dependence on the mean. Overall, the statistics show that the spatial distribution of biogeochemical parameters is well captured in the model.

The seasonal variability of DIN is relatively large (Figure 6.10). In summer, the concentration is low at all stations, while a build-up takes place during fall and early winter. Qualitatively, the model follows the seasonal cycle of the observations, though the model shows less short-term variability.

Table 6.2 Statistics calculated from the spatial distribution of the biogeochemical tracers in the fjord.

	DIN	PO ₄	Chl	O ₂
ME	-0.724	-0.41	1.14	0.117
RMSE	2.29	0.454	1.95	1.168
CORR	0.98	0.89	0.91	0.67
ARMAE	0.034	0.90	0.16	0.0175

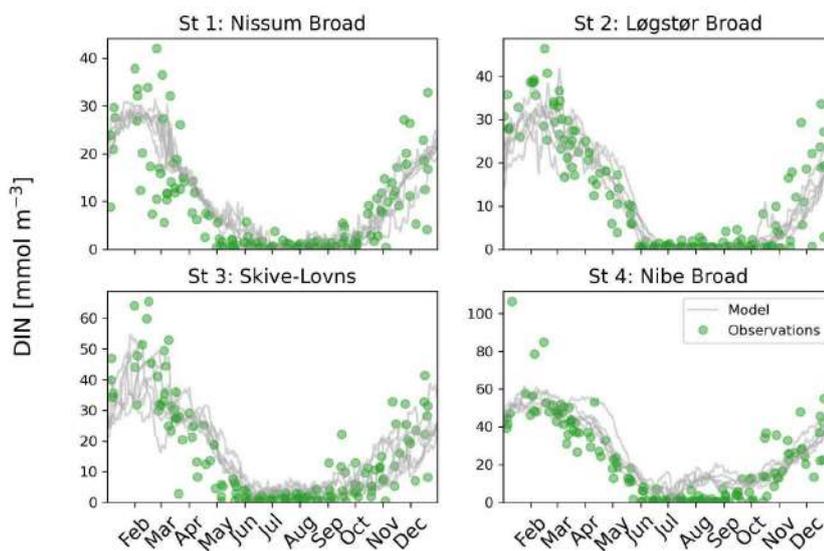


Figure 6.10 Seasonal variability of DIN at the four stations shown in Figure 6.1. Each grey line represents one of the years from 2011 to 2017 and all data points are shown for the same years.

After depletion in spring and early summer, the surface concentration of phosphate increases again during mid-summer at all stations due to hypoxia-induced release from the sediment. The seasonal cycle of the phosphate is well represented in the model (Figure 6.11). However, due to the dependence on hypoxia, the mid-summer release occurs slightly earlier in the model than what is seen in the observations.

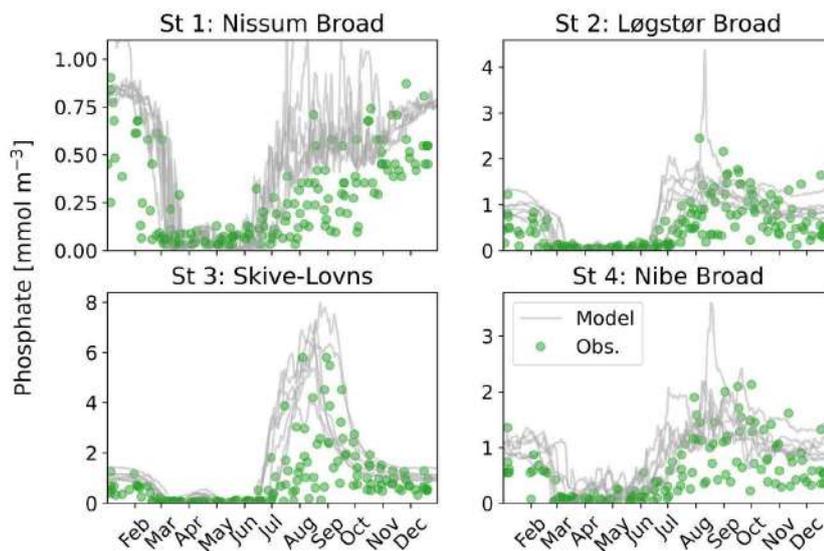


Figure 6.11 Seasonal variability of PO₄ at the four stations shown in Figure 6.1. Each grey line represents one of the years from 2011 to 2017 and all data points are shown for the same years.

The spring bloom, taking place from February to April, is captured in the model with respect to timing and magnitude. The consecutive low summer chlorophyll is likewise represented in the model. However, the fall bloom taking place in the Skive sub-arm (Figure 6.12, St 3) is lacking in the model, likely due to insufficient vertical mixing of nutrients. During fall and winter, the fjord in some years does have blooms of mixotrophs, which are not represented in the model. These blooms have been marked by red crosses in Figure 6.4.

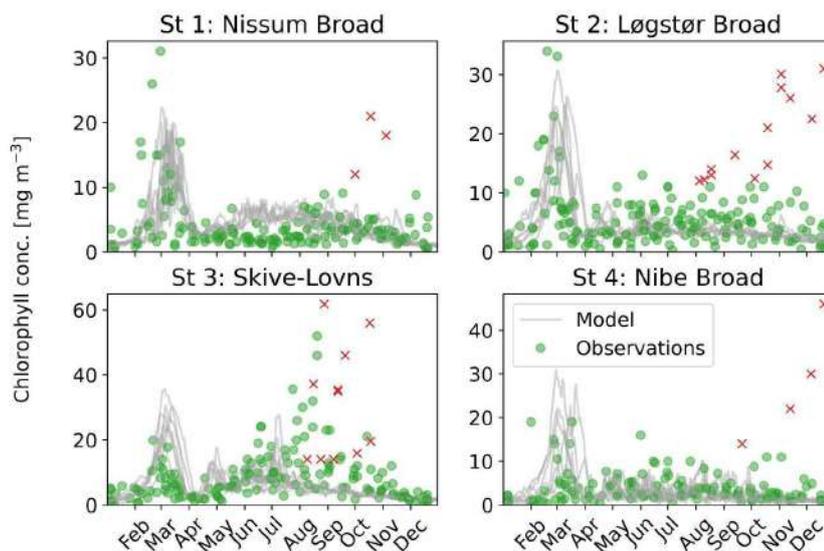


Figure 6.12 Seasonal variability of modelled and measured surface chlorophyll concentration at the four stations in Limfjord (Figure 1) from 2011 to 2017. Red crosses mark observations of mixotrophs that are not described in the model.

The bottom oxygen concentration reaches a maximum in early spring, when wind mixing occurs with high frequency, the water temperature is low and biological consumption is reduced (Figure 6.13). Following the spring bloom (Figure 6.12) and accompanying transport of biological material to the

sediments, the oxygen concentration decreases and reaches a minimum in late summer. The model captures the general trend of the seasonal cycle. However, the short-term variability caused by erosion of the stratification due to wind mixing, is not captured (Figure 6.5, St G, L, N and R).

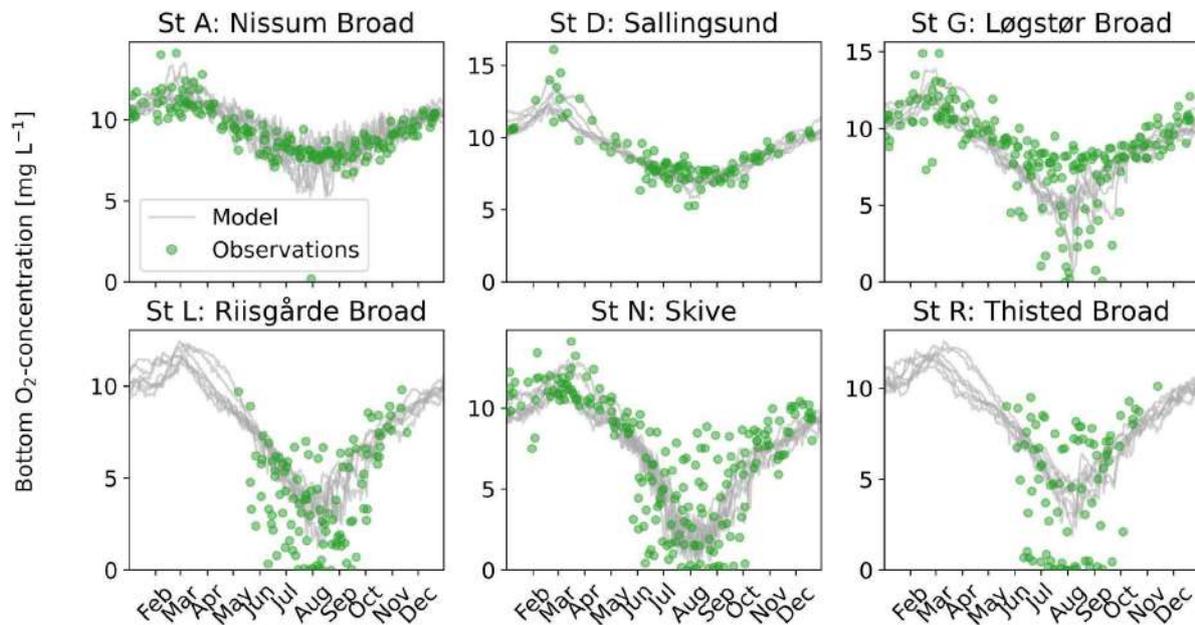


Figure 6.13 Seasonal variability of the bottom oxygen concentration at six stations from 2011 to 2017 (Figure 6.9d).

To quantitatively assess the long-term seasonal cycle, statistics is summarized in Taylor plots for DIN, PO_4 , chlorophyll and oxygen using the full time series from 2011 to 2017 (Figure 6.14). The modelled DIN has correlations higher than 0.8 at all stations, while the normalized standard deviation is smaller than 1, showing that the standard deviation of the observations is slightly higher than in the model. The highest mean error is in the Skive sub-arm (station 3), where the modelled values are underestimated on average. The seasonal cycle plots show that these errors are caused by winter values, which are largely controlled by riverine input.

The PO_4 has correlations between 0.4 and 0.7 at all stations (Figure 6.14b). The normalized standard deviation is relatively high for station 1, where the phosphate concentration is overestimated in the model in late summer, possibly due to localized hypoxia in the model, inducing a release of phosphate from the sediment.

While the modelled chlorophyll captures the seasonal cycle well, it has lower correlations with the observations (Figure 6.14c) due to less short-term variability in the model during summer (Figure 6.12). For station 3, where the fall bloom is underestimated in the model, the model further fails to capture the fall bloom, which is reflected in the overall correlation at this station. The standard deviation is well captured for chlorophyll (Figure 6.14c).

The statistics between modelled and observed oxygen concentration is best for the stations with a full seasonal cycle of observations (station A, B, C, I and N, Figure 6.14d), while the stations with observed summer values only have lower correlation and normalized standard deviation due to the model's lack in variability during summer.

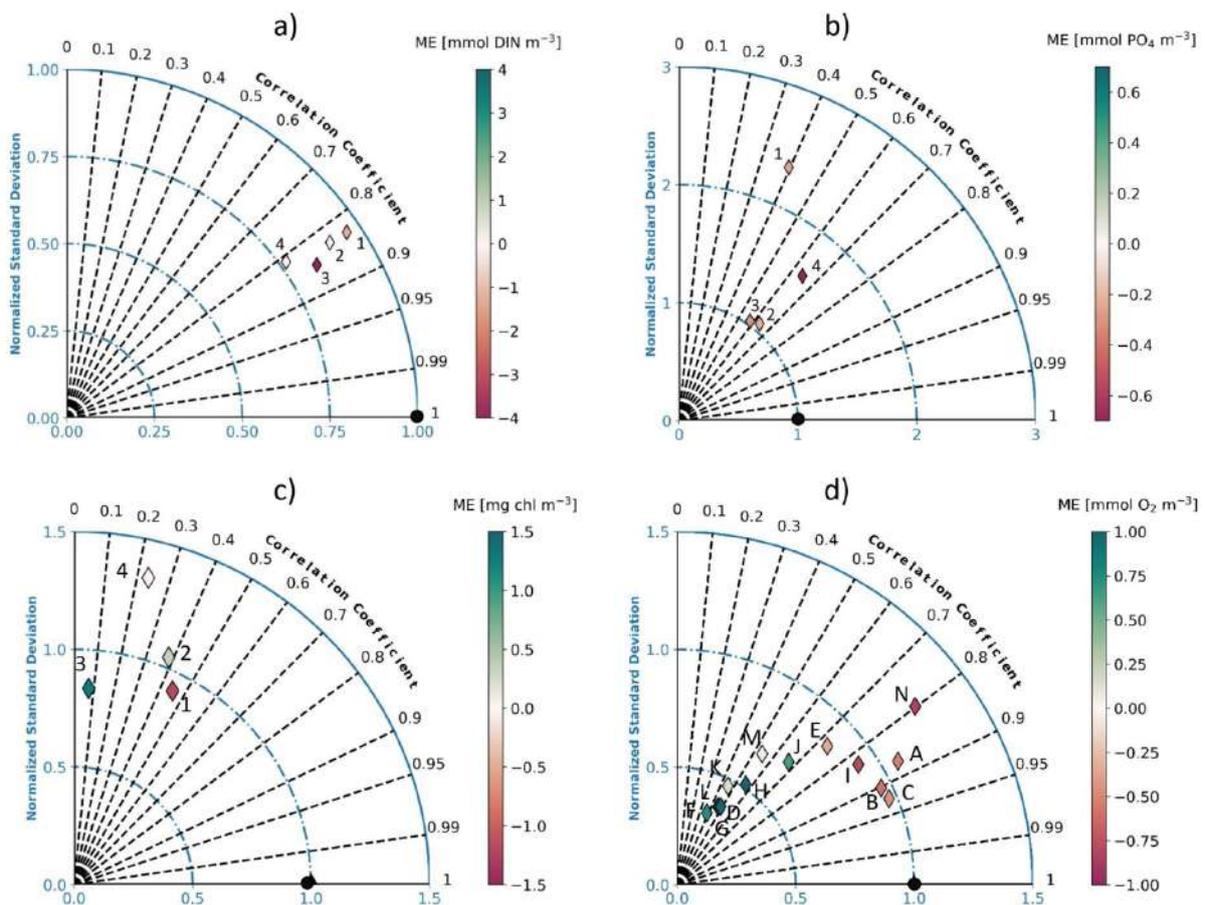


Figure 6.14 Taylor plots showing Pearson correlation, normalized standard deviation and ME for time series of a) DIN, b) PO_4 and c) Chlorophyll for the years 2011 to 2017. The numbers refer to the stations shown in Figure 6.1. The perfect fit is located at $\text{CORR} = 1$ and normalized $\text{STD} = 1$, at the black dot. The colours show the mean error at each station.

3.6.1.4 Conclusion

The morphologically and hydrologically complicated Limfjord is notoriously difficult to model. However, the FlexSem-ERGOM model manages to capture the spatial and seasonal development of the main biogeochemical tracers in the model. The nutrients and oxygen fields are especially well described, though the model's description of chlorophyll and oxygen lacks some summer variability.

2.6.2 Forecast validation

Forecast validation is limited by the availability of in-situ data. The only sources that are available are sea level data from tide gauges, temperature observations from satellites (Fig 6.1 and Table 1) and profile observations. The method uses the entire range of 6 days forecasts for a lag-time dependent validation. The forecasting model runs twice daily, providing a complete forecast of 6 days, which means that each observation is covered by up to 12 forecasts of different lag-time. Forecast-validation allows us to estimate how fast the quality of the forecast decreases with forecasting range, as a result of errors in the model forcing and boundary data, e.g. errors in the numerical weather prediction product.

Figs. 6.15-16 shows temperature time series at selected stations in Limfjord as compared to profile observations and satellite SST data. The agreement is relatively good in summer time meaning that the radiation correction scheme is working properly. However, the modelled temperature is lower in wintertime that results from improper ventilation of Limfjord with standalone setup when inflow events at western and eastern boundaries are underestimated. Therefore, temperature bias is negative, see Fig.6.17 left. Winter has usually more cloudy days than summer meaning that satellite observations are less influenced by negative bias in winter. The increase of central RMSD is small with increased forecast range, see Fig.6.17 right, denoting that forecasted situation is useful even when forecasted 6 days ago. Figure 6.18 shows that winter temperatures are underestimated with respect to both profile observations and satellite SST data.

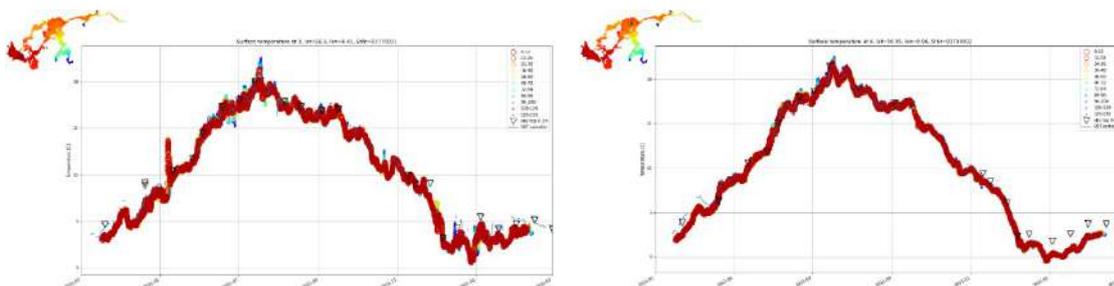


Figure 6.15 Surface temperature in Odde-sund (left) and central Limfjord (right). Circles denote model data with specified forecast range, triangles – profile observations, where only data with depth <1 m are selected, line – satellite SST data.

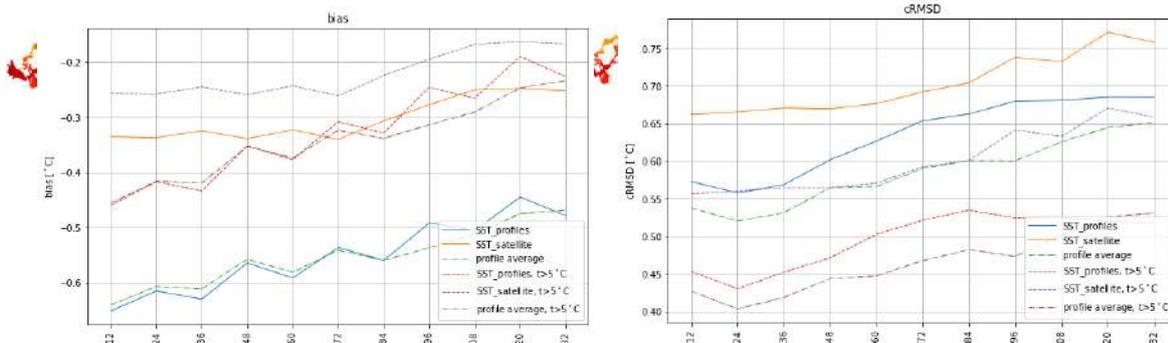


Figure 6.16 Bias (left) and central RMSD (right) of modelled temperature in Limfjord as average of stations depending on forecast range.

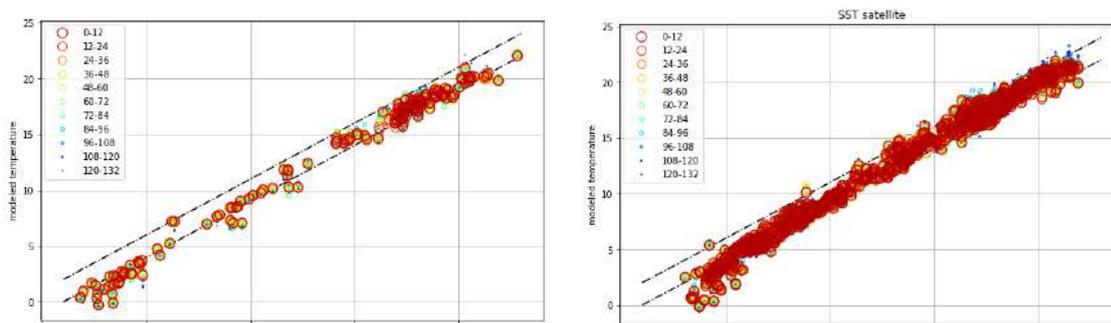


Figure 6.17 Modelled versus observed temperature in the upper layer for stations within Limfjord. Left – profile observations, right – satellite SST temperature. Lines denote deviation by 1 degree from observed value.

The operational product for Limfjord involved just one storm event, i.e. storm Malik in January 2022, when water level was over 1 m (Fig.6.19). The standalone setup of Limfjord works well at areas close to either western and eastern boundaries but underpredicts the extreme events in the middle of the Limfjord. Central RMSD for water level across stations at Limfjord shows that almost the same accuracy is up to 3 day of forecast range (see Fig.6.20 left). For a longer forecast range, the RMSD starts to deteriorate. See level bias is station dependent and does not change with forecast range (Fig. 6.20).

2.6.3 Process-oriented validation

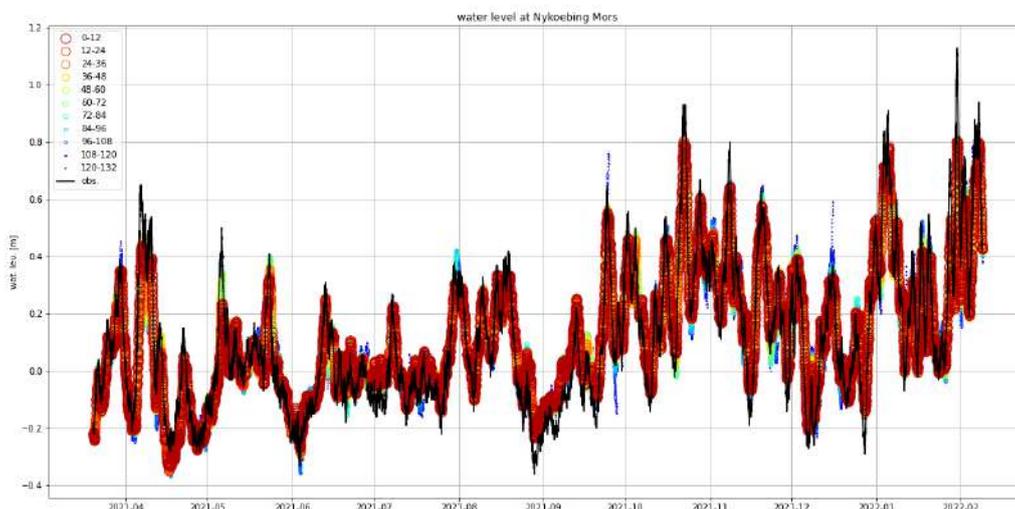


Figure 6.18 Water level at Nykoebing Mors. Circles denote model data with specified forecast range.

Process-oriented validation has been carried out for the 5-year historic period 2015-2019 (Fig. 6.21), for sea levels at Lemvig tide gauge station (Fig. 6.1). The scope of the assessment focuses on the model performance for predicting sea level extreme events, for storm surge warning. Studied are KPI's for sea level peak error and the forecast miss-rate at Lemvig tide gauge station near the aquafarming site.

3.6.3.1 The assessment of sea level extreme events at Lemvig tide gauge station

Process oriented validation of sea level anomalies focuses on the model performance with regards to predicting peak events. Figure 6.22 provides an overview of modelled and observed sea level anomalies at Lemvig tide gauge station, close to the aquafarming site Kulturfelt Follup Odde (KFO). There is generally a good agreement between modelled and observed sea level anomalies.

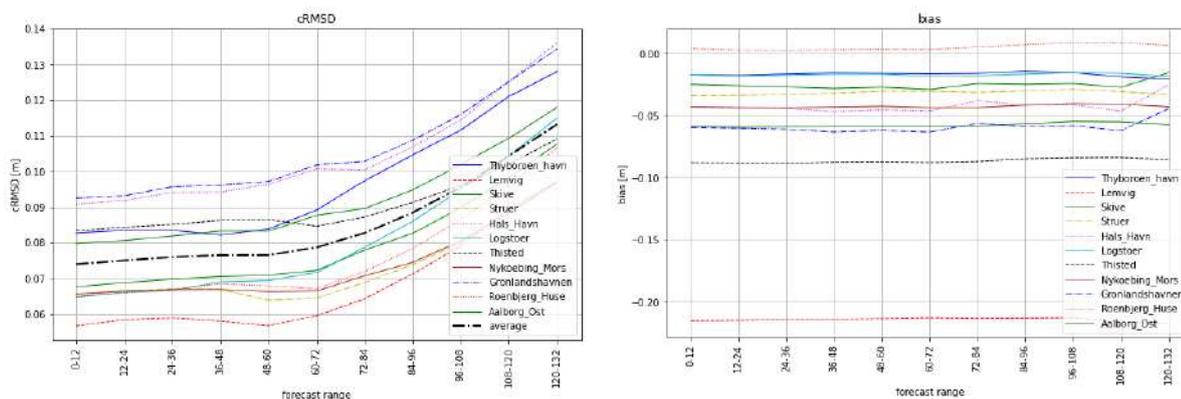


Figure 6.19 Central RMSD (left) and bias (right) of water level in stations of Limfjord as function of forecast range.

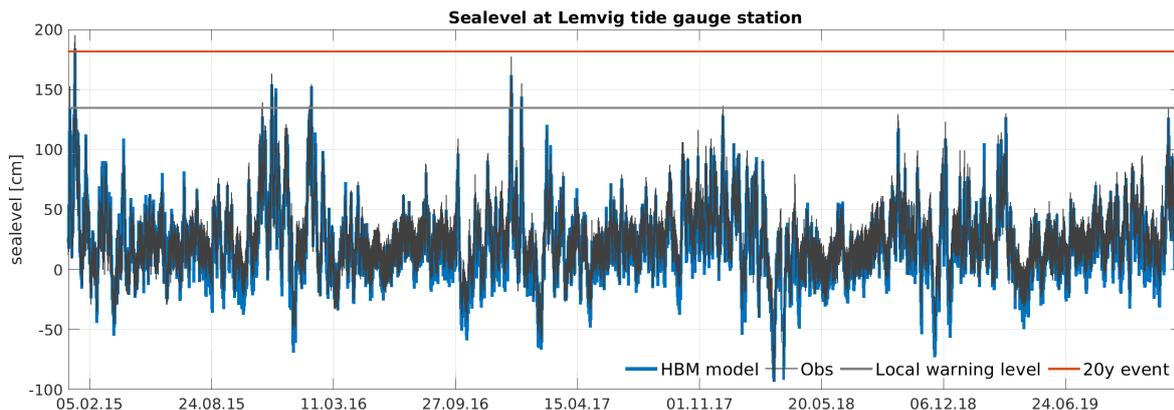


Figure 6.20 Modelled (blue) and observed sea level (black) in the historical period 2015-2019, at Lemvig tide gauge station. The local warning level (grey) of 1.35 m and the sea level of a 20 years event (red) of 1.82 m are shown as well.

The assessment focuses on sea level events exceeding warning levels. These are usually defined by the user of the warning service, i.e., the local authorities operating the harbours. At Lemvig harbour, the warning level is 1.35 meter, which is exceeded 19 times in the course of the 5 years historical period 2015-2019 (Fig 6.21). The sea level value of a 20-years event (1.82 meter) has been added as well. There is only one 20-years event in the entire time series, the January storm in 2015, at the beginning of the time series. These were actually 2 storm events: on the 9th and the 10th of January 2015, which in Denmark carry the names Dagmar and Egon (Fig 6.22).

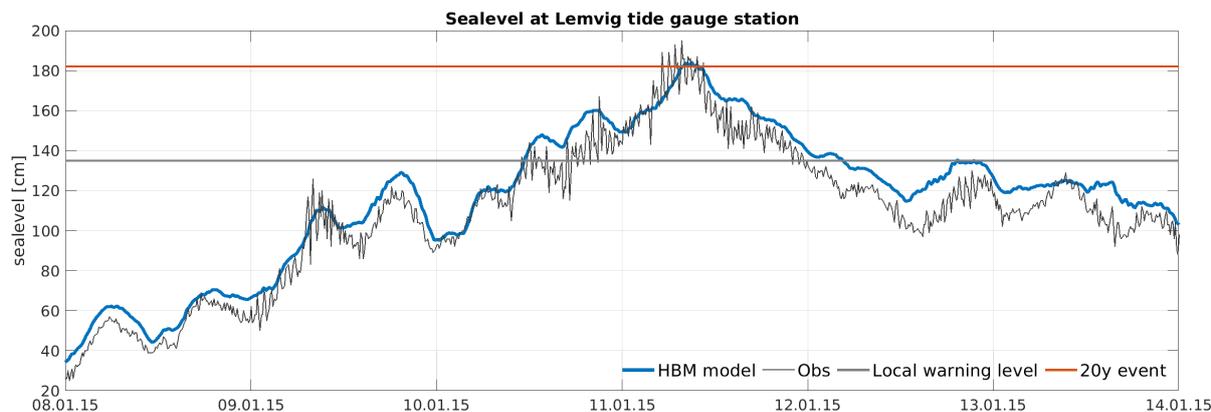


Figure 6.21 Modelled (blue) and observed sea level (black) during storm Dagmar and Egon (9th and 10th of January 2015) at Lemvig tide gauge station. The local warning level (grey) of 1.35m and the sea level of a 20 years event (red) of 1.82m are shown as well.

Sea level peak error assessments of operational products usually study the 10 highest observed sea level events in a 1-year period. For the present study, we have modified this, by studying all the events, when observed sea level exceeds local warning level of 1.35 m at Lemvig tide gauge station in the entire period of 5 years 2015-2019. That happens 19 times (Fig. 6.21). The assessment follows the procedure for operational model validation assessments used for deriving key-performance-indicators KPI's. In a first step, modelled and observed sea level peak values are matched to each other. The method neglects phase errors by allowing for a time shift of up to ± 6 hours between modelled and observed peaks. A separation window of at least 12 hours between peak events ensures that storm surges are not counted twice. The assessment analyses peak errors and compares them with justified error ranges, to calculate a miss rate. If the peak error exceeds the error range, the forecast is called

a “missed warning”. The number of missed warnings is studied as well. The peak error assessments usually work with a justified error range of 10% of the observed peak value, with a minimum of 10cm.

The assessment (Fig. 6.23) focuses on sea level anomalies, which is also the parameter used by the operational storm surge forecasting service. Reason for this is that every tide gauge station has its own reference level, which is not uniformly defined internationally. Sometimes, the information is completely missing. Reference levels might also be a function of time, as stations are replaced or reconstructed. Therefore, the standard is to analyse sea level anomalies, which are generated by subtracting the long-term mean from the model forecasts and observations. For the current assessment we have chosen to subtract the long-term mean of the 5 years period 2015-2019, which at Lemvig station is 25.4 cm for the observations and 20.6 cm for the model data.

The model is able to predict all the sea level extreme events. The miss-rate is 0% for error-ranges larger than 9% of the observed value, meaning that 0% of the observed events are outside that error-range, i.e., that 100% of the peak events were predicted. Validation assessments of the operational system use error-ranges of 10%, with a minimum of 10 cm. So, from an operational point of view, the system has a miss-rate of 0% at Lemvig. This is considered to be a very good performance. But, to analyze this further, the miss-rate has been calculated as a function of the permitted error-range (Fig. 6.23, right). The analysis shows that reducing the permitted error-range to 7% of the observed value, increases the number of missed events only to 10.5%. Further reduction to 4% error range leads to an increase of the miss-rate to 47.4%. On the other hand, the 20-years events, storm Dagmar and Egon, in January 2015 are still within the accuracy limits. The difference between modelled (1.88 m) and observed sea level (1.95 m) for this storm period is within a 3.29% error-range. This is a very good skill score for a 20 years event.

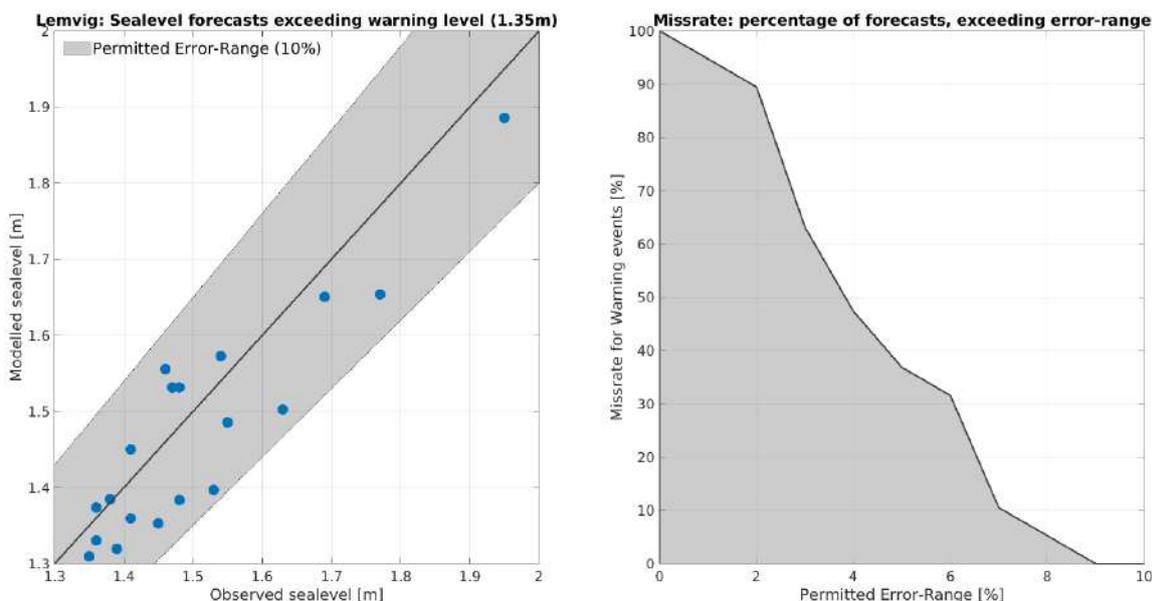


Figure 6.22 Left: Sea level anomalies of extreme events, i.e., storm surges, exceeding warning levels of 1.35m of observed sea levels at Lemvig tide gauge station. All events lay in the range of permitted error values for operational assessment, which is 10% of the observed peak value, with a minimum of 10cm. Right: Error-Range dependent assessment of miss-rate, i.e., percentage of forecasts that are outside the permitted error-range of the observations.

3.6.3.2 The assessment of model performance with regards to diurnal warming in very shallow waters at aquafarming site KFO1

Forecasts of hydrographic conditions for marine conditions services for aquafarming, such as oyster restoration and farming near Lemvig, at Kulturfelt Follup Odde (KFO1) require good model performance, especially with regards to temperature forecast. The diurnal variability can be relatively high in very shallow, coastal waters. Usually, the model is validated using in-situ and ship cruise data in deeper waters (see historical validation). The model has a good performance for temperature forecasts at the environmental monitoring cruise stations of the NOVANA surveys, although in very shallow, coastal waters, we found a consistent cool bias of the model, which was partially removed by improving the radiation scheme in the model.

In shallow water, the seabed becomes an integral part of the ocean thermodynamics. Vertical heat exchange between the ocean and the bottom soil is implemented, but not enough to warm the ocean sufficiently, if the effect of solar radiation that is reflected from the seabed is not included. Here we improved the model by increasing the energy absorption of the radiative flux in shallow waters with less than 5m water depth. Figure 6.24 shows the impact of the shallow water radiation scheme. The new temperature forecasts near the seabed (red) are much closer to the observations (black) than the previous model results (blue).

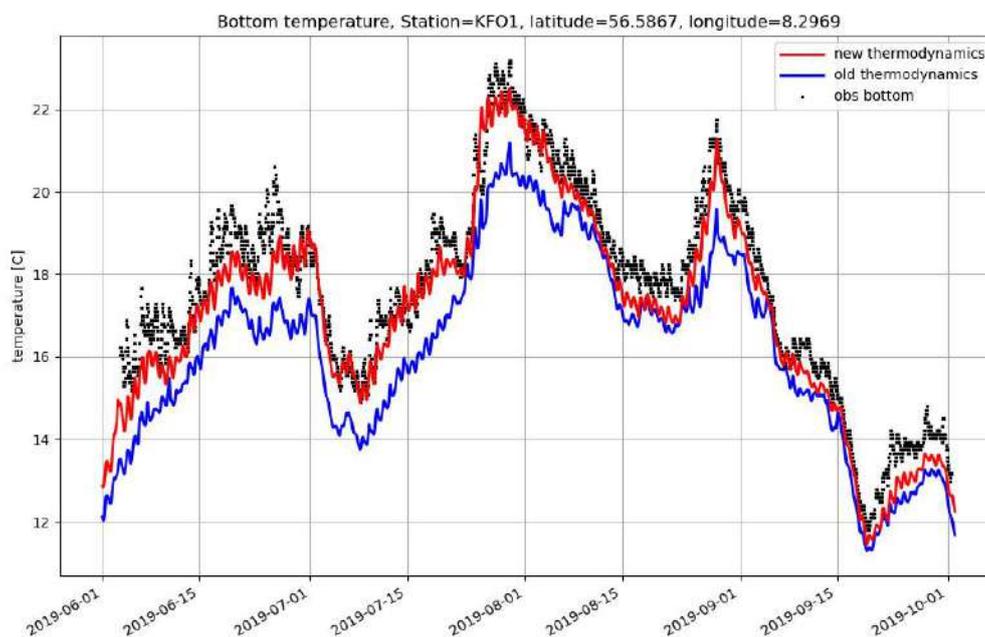


Figure 6.23 Water temperature near the sea bed at KFO1 station at the Aquafarming site. The shallow water radiation scheme improved the temperature forecast (red), which is much closer to the observations (black) than the previous version (blue).

At KFO1 station, temperature data were measured both at the surface and at bottom with a water depth of about 2 m. However, surface data has significant skin effects which are not suitable for model validation. Hence, only the bottom temperature is used for the validation. The new shallow water radiation scheme improves the forecasts of diurnal warming. Figure 6.25 shows the monthly mean temperature for each hour of the day, for the 4 months: June-September 2019. The cool bias is

relatively low, about $-1.0\text{ }^{\circ}\text{C}$ in June, but only $-0.27\text{ }^{\circ}\text{C}$ in July; and $-0.56\text{ }^{\circ}\text{C}$ and $-0.53\text{ }^{\circ}\text{C}$ in August and September. The centralized Root-Mean-Square-Error is generally low: $0.1\text{ }^{\circ}\text{C}$ in June, but only $0.08\text{ }^{\circ}\text{C}$ in July; and $0.07\text{ }^{\circ}\text{C}$ and $0.04\text{ }^{\circ}\text{C}$ in August and September. This shows that the variability of the observations is reproduced. The standard deviation of the observations ranges from $0.28\text{ }^{\circ}\text{C}$ and $0.22\text{ }^{\circ}\text{C}$ in June and July to $0.16\text{ }^{\circ}\text{C}$ and $0.09\text{ }^{\circ}\text{C}$ in August and September. The standard deviation of the model forecast is lower, but in a similar range: $0.2\text{ }^{\circ}\text{C}$ and $0.16\text{ }^{\circ}\text{C}$ in June and July, as well as $0.1\text{ }^{\circ}\text{C}$ and $0.08\text{ }^{\circ}\text{C}$ in August and September. The phase of diurnal warming is well reproduced. The Pearson correlation coefficient is 0.95 in June, above 0.98 in July and August (0.988 and 0.981) and 0.91 in September. Accurate forecasts of the timing of diurnal warming events can be provided.

Process-oriented validation focusses on two physical processes, the forecast of sea level extreme events and the model quality with regards to forecasts of diurnal warming. The model is well tuned and configured. It is able to forecast all extreme events that are exceeding warning levels at Lemvig in the period 2015-2019, with an accuracy of at least 9% of the observed value. The miss-rate, i.e., the percentage of missed forecasts is 0%, which means that all events have been predicted and warnings would have been issued, if the model would have been operational in the considered period (2015-2019).

The model quality with regards to diurnal warming predictions is good. The model bias finds its maximum $-1.0\text{ }^{\circ}\text{C}$, in June. At most other months (July-September) it is below $-0.56\text{ }^{\circ}\text{C}$, roughly half the value. The standard deviation of the modelled product is roughly 30% to low, which could only be improved if the solar radiative input at sea surface would be increased. As it is right now, nearly all of the solar radiation in the very shallow waters are absorbed by the ocean. The timing of the diurnal warming event is very good, as the assessment of the correlation coefficient shows. With this, the model is able to capture diurnal warming events in the very shallow waters at KFO1.

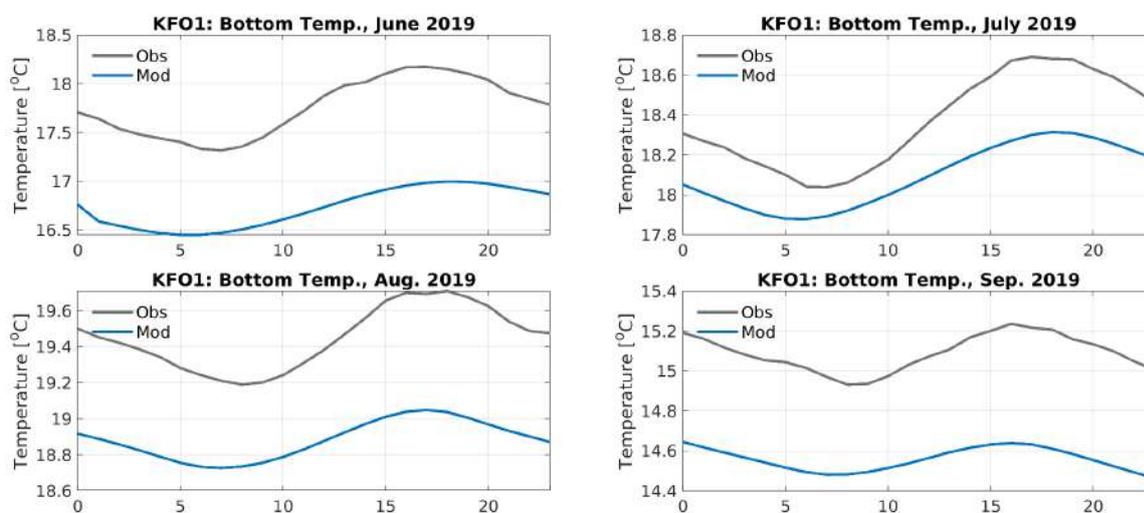


Figure 6.24 Diurnal signal of sea bottom temperature, i.e., the monthly mean temperature, for each hour of the day, at KFO1 station, for the 4 months: June to September 2019.

2.7 Pilot 7: Romania

The “Eforie” forecasting system is a suite of two-way nested, online-coupled physical and biogeochemical models. The physical model is Nemo 3.6, the nesting procedure is done with AGRIF, and the biogeochemical model is the BAMHBI model developed at the University of Liege. The parent model covers the Black Sea with a horizontal resolution of 3 km. The nested grids cover i) (“NWS”) the Black Sea Northwestern Shelf with an horizontal resolution of 1 km, and ii) (“Eforie”) the coastal area in front of the cities of Contanta and Eforie at 200m resolution. All three models share the same vertical grid using 59 unequally-space z-layers. The surface layer has 0.5 m height.

A recurrent problem in the Black Sea is the lack of in situ data, particularly in the coastal area, and regarding biogeochemical variables. This renders both model calibration and validation difficult. However, this lack of data is attenuated by satellite observations and, in the case of FORCOAST, by dedicated *in situ* campaigns conveyed during 2018 and 2020.

During hindcast, the ECMWF IFS analysis fields are used for atmospheric forcings, whereas for forecasts, the ECMWF IFS forecasts are used. Apart from this difference, the model configuration is strictly identical in hindcast and forecast modes. In particular, there is no data assimilation procedure during the hindcasts. Every day, the operational chain specifies that the first day of the forecast (provided the previous day) is overwritten by a new hindcast, and subsequently 5 days are forecasted.

Therefore, it is estimated that forecast and hindcast have similar skill, and the validation presented hereunder is relevant for both modes. The validation is targeted towards the variables that are both relevant for the aquaculture sector (and potentially fisheries), and present sufficient availability. This limits the validation to sea surface temperature (SST), chlorophyll-a (CHL), and currents obtained from ADCP measurements.

2.7.1 Hindcast and Forecast validation

2.7.1.1 Sea surface temperature

In this section, the sea surface temperature (SST) from the Eforie model is compared to satellite measurements. As a reference, the comparison is repeated using the CMEMS Black Sea model; it should be noted that the latter daily assimilates SST observations, hence it is expected that its error is lower. Obviously, this advantage of the CMEMS product compared to the Eforie will be reduced in forecasting mode, especially with longer lead times.

The comparison below is performed using the CMEMS level-4 (L4) images (product 010_006), where gaps in the images, due to clouds, have been filled. The horizontal resolution of the images is 1 km (i.e. coarser than the Eforie model, but finer than the CMEMS Black Sea model) and comparisons have been computed.

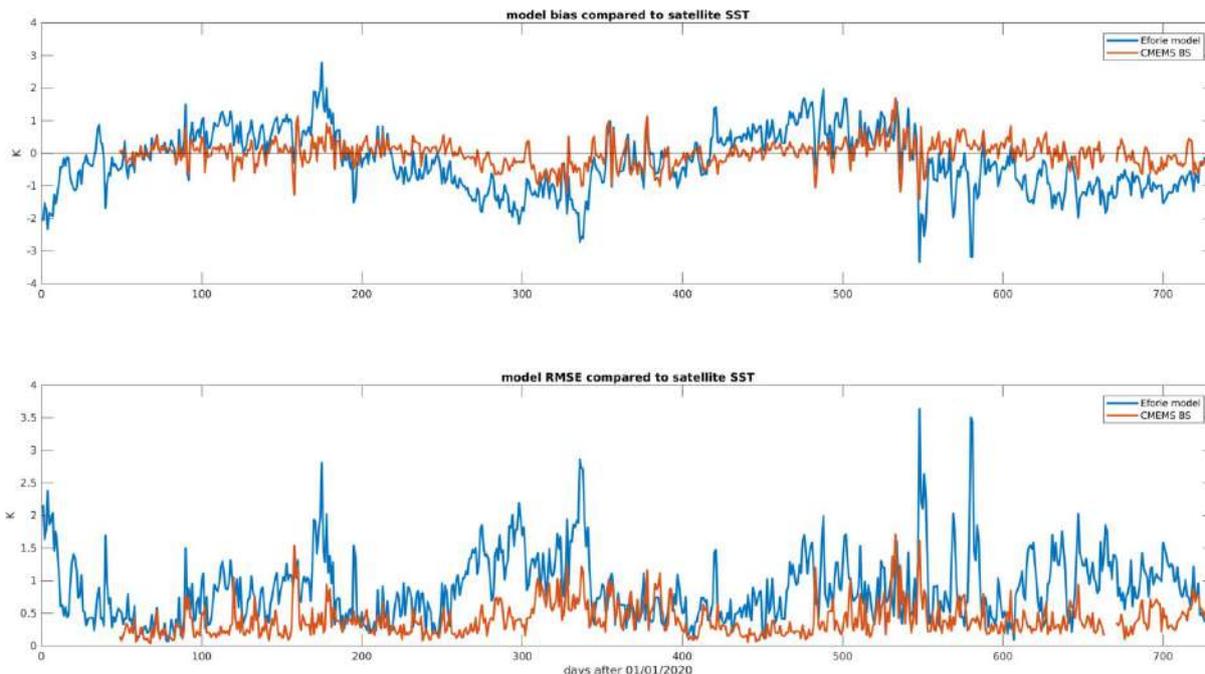


Figure 7.1 (Top) Spatial average of SST bias between model and L4 satellite observations during 2020 and 2021 (bottom) RMSE.

Figure 7.1 shows the bias and root mean square difference (RMSD) between both models and the satellite product, spatially averaged over the Eforie model domain. It can be seen that the bias is larger in the Eforie model than in the CMEMS BS. In particular, Eforie presents a positive bias during summer and negative bias during winter. The time-average of the bias in the Eforie model is -0.28°C . The corresponding RMSD is 0.87°C , whereas for CMEMS BS it is 0.40°C , or roughly half. As explained before, this is expected, as no data is assimilated in the Eforie model; and in forecasting mode, the errors of both models are expected to be more similar.

Monthly plots RMSD between the Eforie model and the satellite images, show that most of the error concentrates close to the coast during winter months, whereas during spring-summer-autumn, the error is spread over in the whole domain.

Figure 7.2 shows the same time-series, now comparing the models with CMEMS (cloudy) L3S SST images (CMEMS product 010_013). The difference between model and satellite images are more similar; the mean bias is -0.04°C for Eforie and 0.23°C for CMEMS BS; whereas the RMSD is 0.8683°C for Eforie and 0.55°C for CMEMS BS.

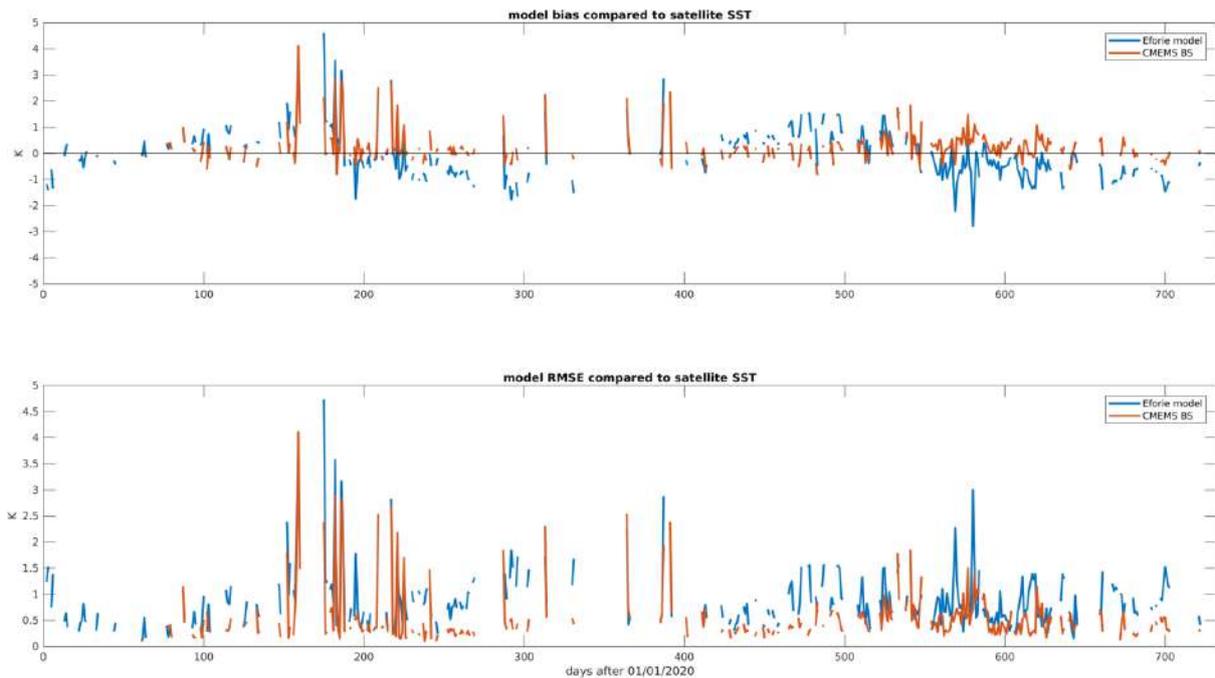


Figure 7.2 (Top) spatial average of SST bias between model and L3S satellite observations during 2020 and 2021 (bottom) root mean square error

The monthly plots of RMSD are shown in Figures 7.3 and 7.4 respectively for Eforie and CMEMS BS.

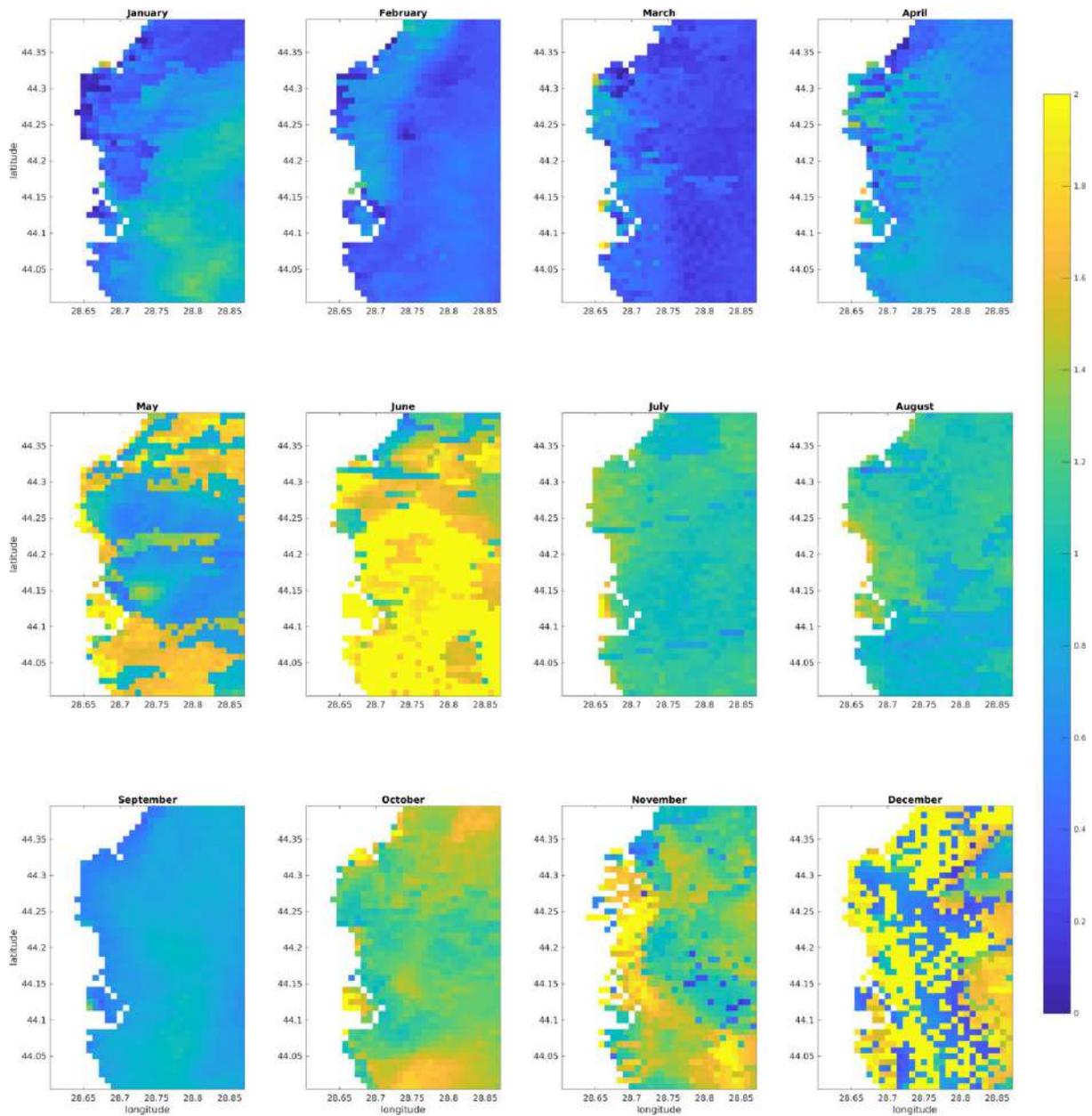


Figure 7.3 Monthly spatial maps of SST root mean square error between Eforie model and satellite observations for 2020

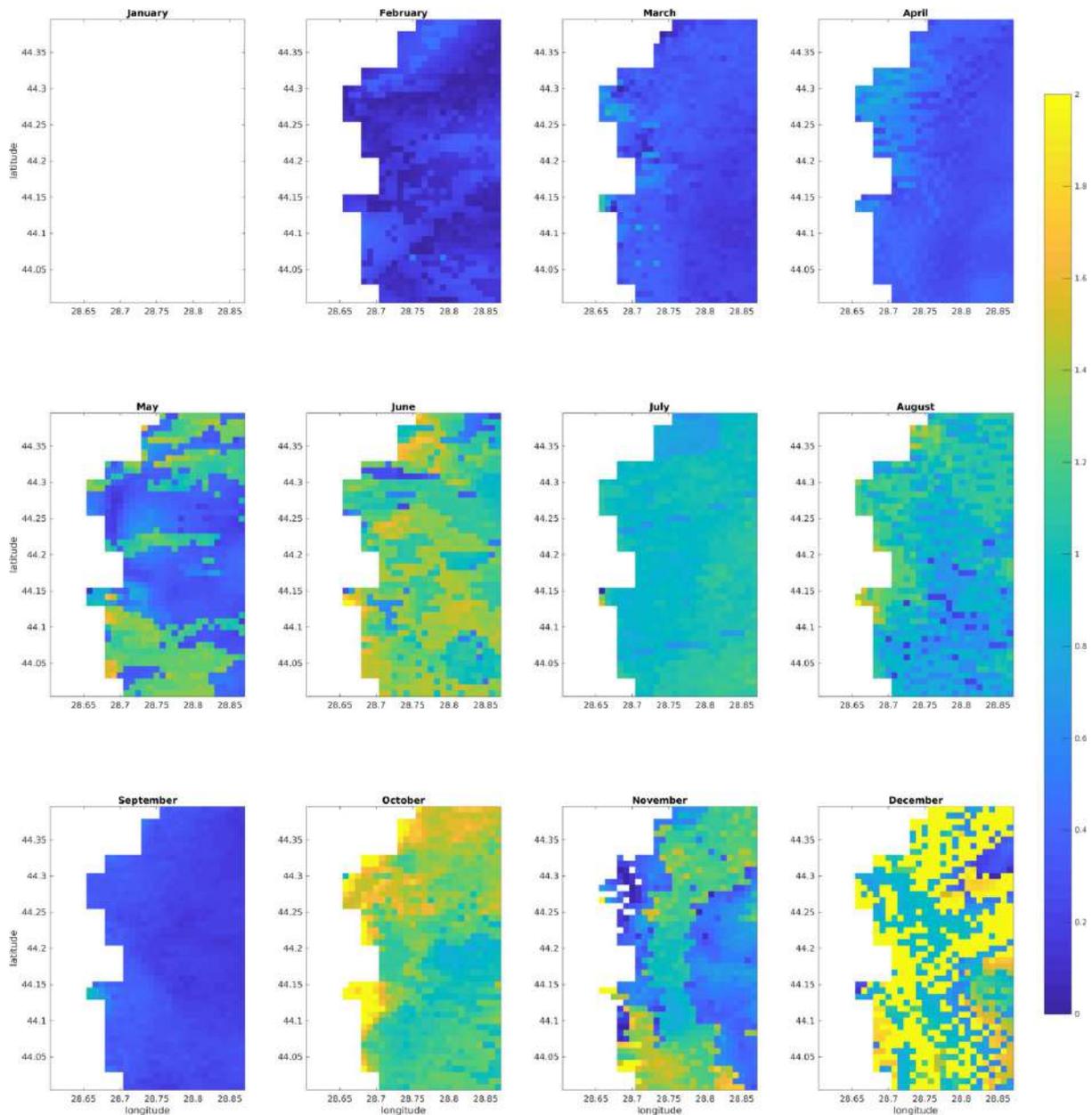


Figure 7.4 Monthly spatial maps of SST root mean square error between CMEMS BS model and satellite observations for 2020

2.7.1.2 Sea surface chlorophyll

For recent years, the Sentinel satellites provide high-resolution and high-quality measurements of surface reflectances in various bandwidths. These measurements are then used to estimate surface chlorophyll-a concentrations at the ocean surface, using algorithms based on ratios between reflectances, or neural networks.

In the Pilot 7 area, chlorophyll-a concentrations are extremely variable due to the Danube river plume which may (or not) pass through the domain. According to the satellite observations, concentrations can reach over 70 mg.m^{-3} .

Figure 7.5 presents a time-series, for each day in 2021, of i) the spatially-averaged bias between the model forecasts and the Sentinel 3 chlorophyll estimates, and of the root mean square difference (RMSD) between them. Missing points are caused by cloudy images. The validation is performed both for the Eforie model (used for Pilot 7), and for the CMEMS Black Sea product. The latter suffers from a lower spatial resolution, but however assimilates the CMEMS multi-satellite product.

Both models have a low bias: they tend to under-estimate chlorophyll-a concentrations during the summer months, but slightly overestimate it during the winter months. Furthermore, it can be seen that, even without data assimilation, the Eforie model has similar bias and RMSD values as the CMEMS modelled product. At periods of very high bias, while it is clear that the Sentinel-3 chlorophyll estimate is very different from the modelled products, it is not clear (by lack of *in situ* measurements) which product is actually closest to the truth.

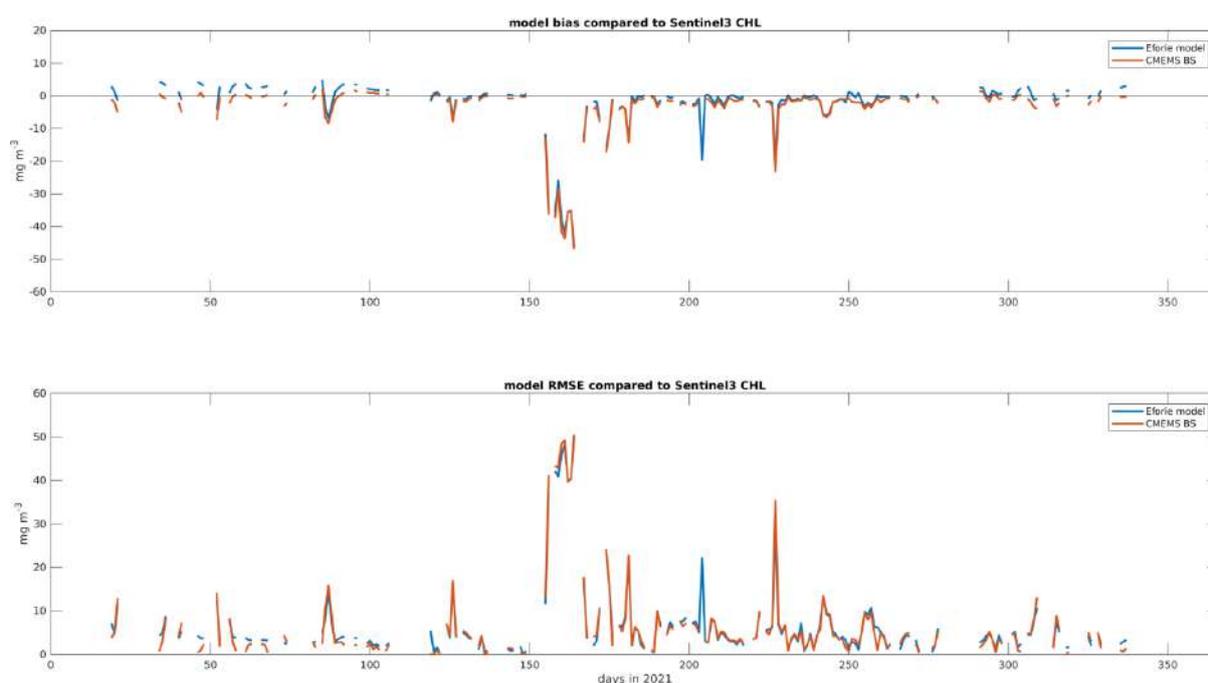


Figure 7.5 (Top) spatial average of chlorophyll-a bias between model and satellite observations during 2021 (bottom) root mean square error

Figures 7.6 and 7.7 present monthly spatial maps of the RMSD between the satellite estimate and, respectively, the Eforie and CMEMS model products. Again, it can be observed that the maps are quite similar, although the Eforie model now presents evidence of small structures (e.g. in the harbour area or very close to the coast) naturally absent in the 3-km resolution CMEMS model. During the summer months (and particularly June), both models present a very large RMSD compared with the Sentinel-3 estimate.

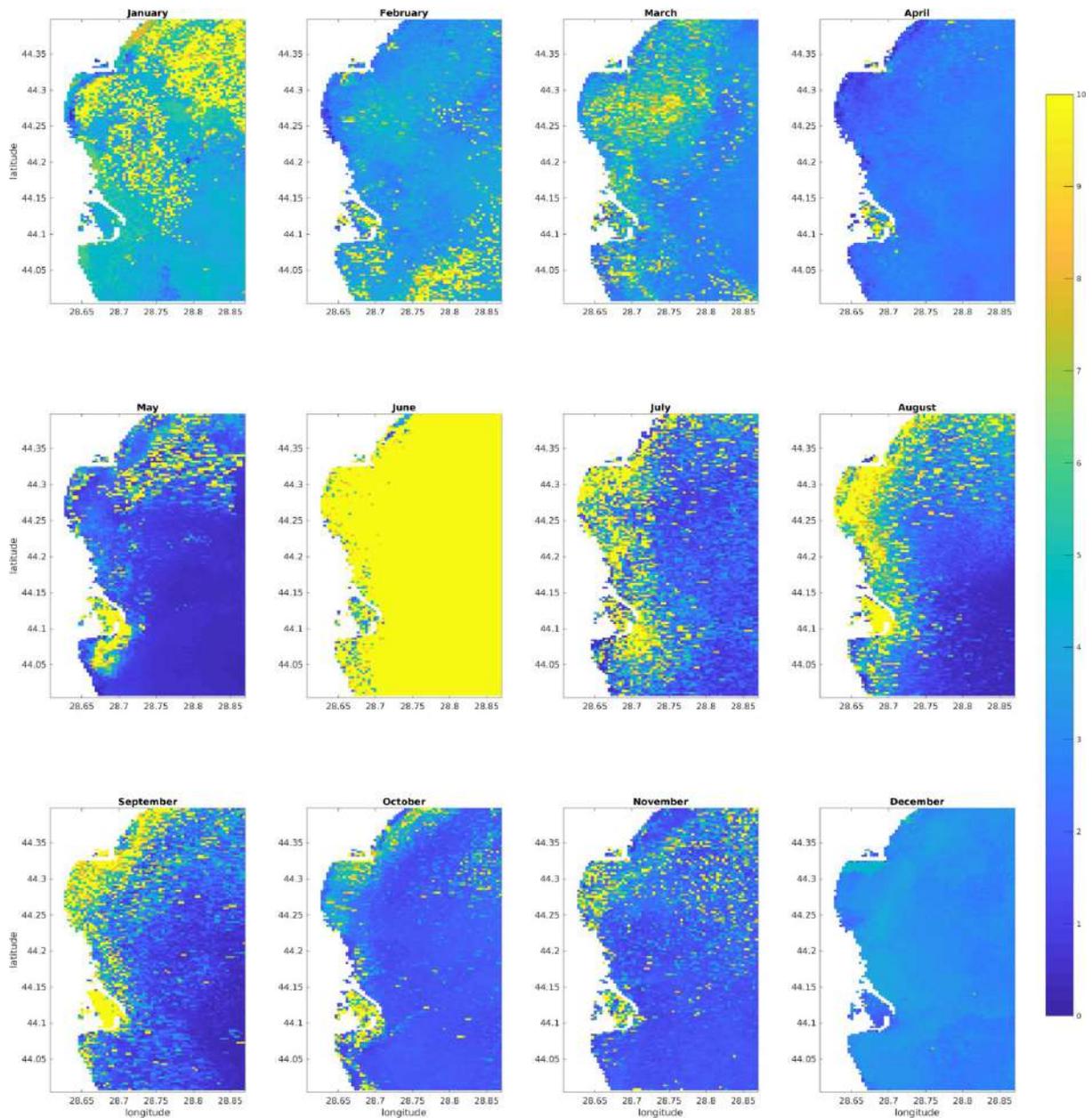


Figure 7.6 Monthly spatial maps of chlorophyll-a root mean square error between Eforie model and satellite observations

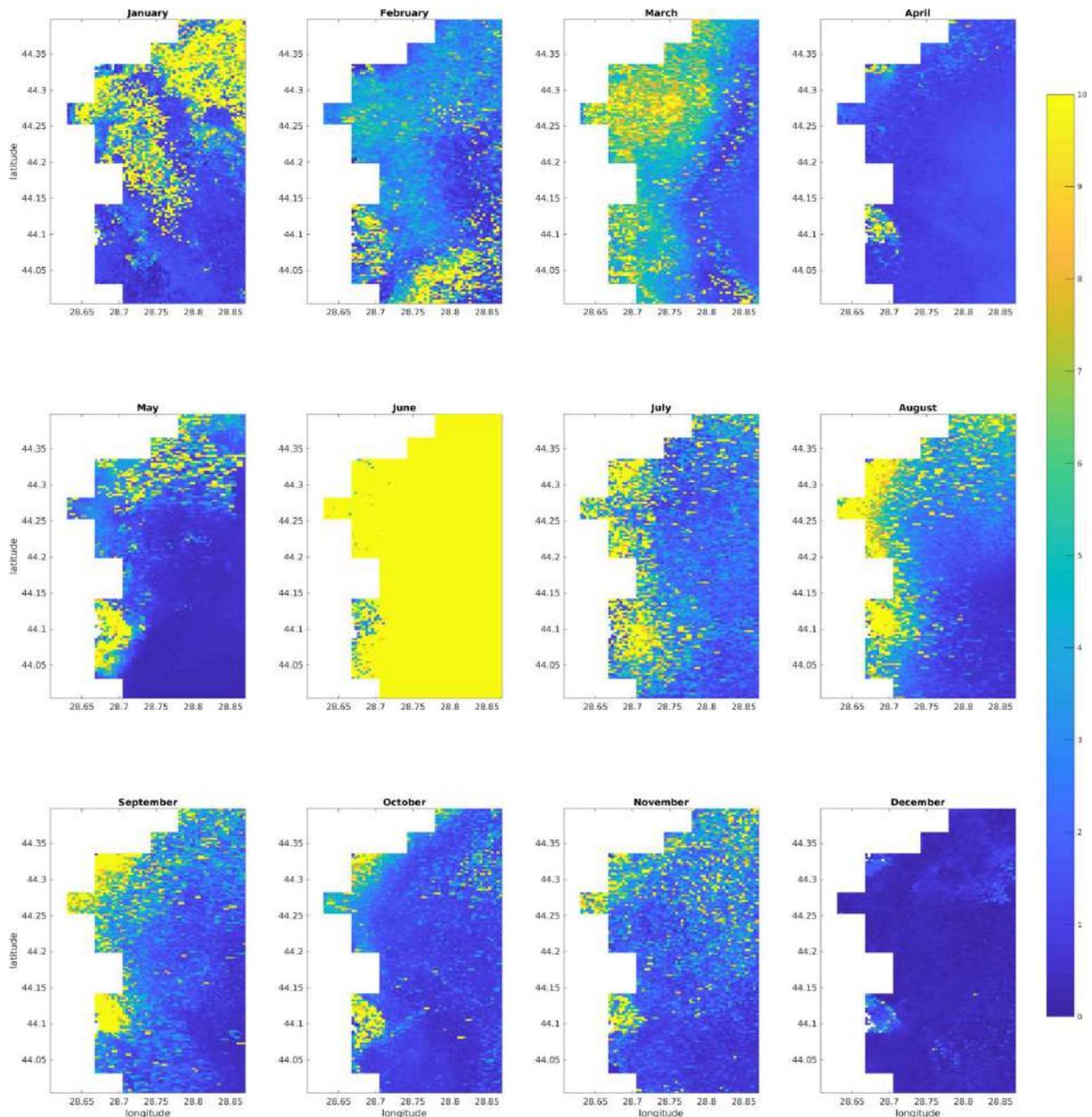


Figure 7.7 Monthly spatial maps of chlorophyll-a root mean square error between CMEMS model and satellite observations

2.7.1.3 Horizontal currents

During the FORCOAST project, specific measurement campaigns were organized in the coastal area of interest. Horizontal current velocity measurements were obtained by ADCPs. Some ADCPs were located inside the domain covered by the innermost nested model (“Eforie”), whereas other measurements were outside this domain and hence they were compared with the intermediate-resolution model covering the Black Sea’s North-Western shelf (“NWS”)

Figures 7.8 to 7.11 present the current directions and intensities during 2018 and 2020, separately for the data points in Eforie and for all the data points.

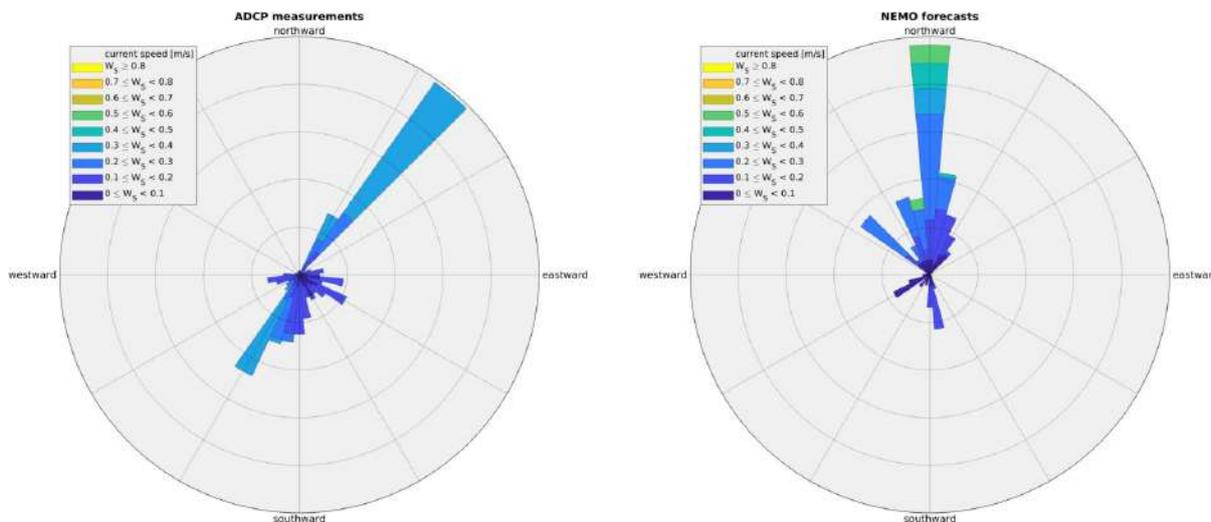


Figure 7.8 Current directions in the Eforie area, obtained by (left) ADCP measurements (right) equivalent points in the Eforie model, during 2019, for all measured depths

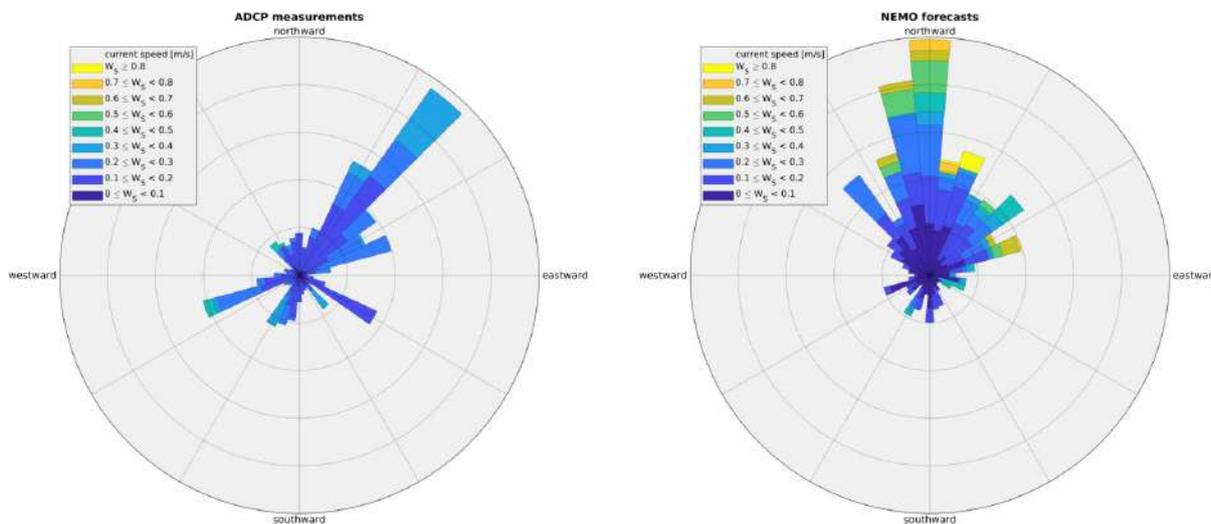


Figure 7.9 Current directions in the larger Eforie area, obtained by (left) ADCP measurements (right) equivalent points in the NWS model, during 2019

During the 2019 campaigns, the current was mostly flowing along the coastline, usually north-ward but sometimes also south-ward. The modelled north-ward current is slightly misplaced (towards the North instead of North-East), and some of the south-ward currents are missing. The current velocity is realistic in both directions.

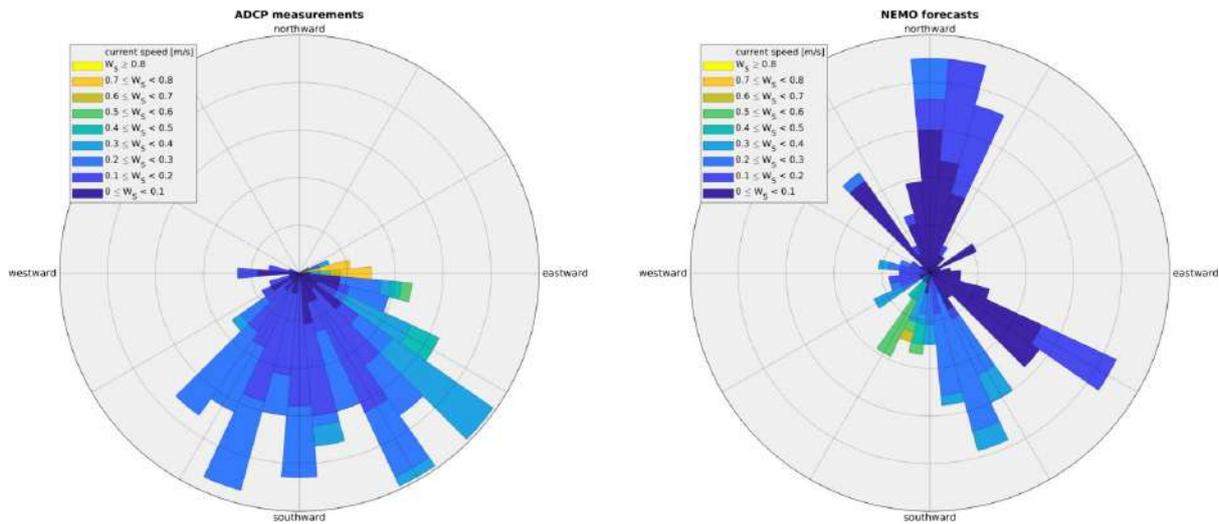


Figure 7.10 Current [ms⁻¹] directions in the Eforie area, obtained by (left) ADCP measurements (right) equivalent points in the Eforie model, during 2021

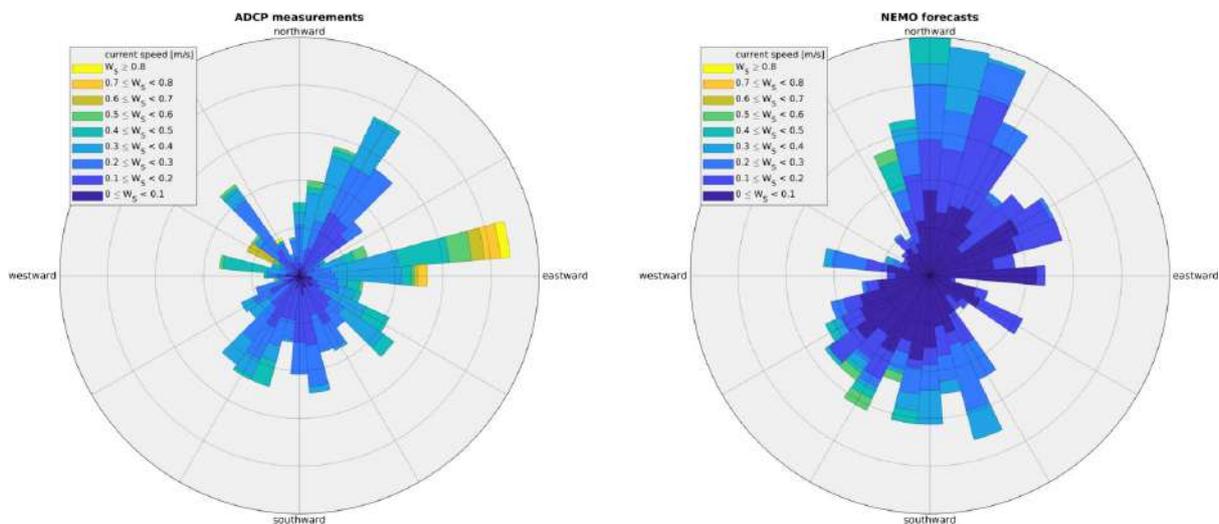


Figure 7.11 Current [ms⁻¹] directions in the larger Eforie area, obtained by (left) ADCP measurements (right) equivalent points in the NWS model, during 2021

During the 2021 campaign, close to the coast, the currents were mostly oriented southward, whereas the model falsely predicted some northward currents as well. When considering all data points (including the ones outside the Eforie model domain), the modelled current direction is closer to the measured one.

Table 7.1 presents the current velocity error statistics corresponding to the 4 figures above. It can be seen that the bias is low, except in 2021 in the NWS model. Model skill seems to be higher in the innermost nest (Eforie) then in the outer one (NWS), with lower (percentage) bias and root mean square (RMS) errors. The ratio between observed and modelled standard deviations is also closer to 1 (the ideal value).

Table 7.1 Error statistics of the current velocity computed in the Eforie and NWS models for the 2019 and 2021 campaigns.

		Eforie	NWS
<i>bias [m/s]</i>	2019	-0.020	-0.005
	2021	-0.051	-0.13
<i>percentage bias [%]</i>	2019	-13.75	-2.8
	2021	-24.45	-50.2
<i>RMS error [m/s]</i>	2019	0.10	0.19
	2021	0.22	0.24
<i>standard deviation ratio [/]</i>	2019	0.84	0.48
	2021	1.06	1.40

From Table 7.1, it can be seen that in the Eforie model, the mean bias is relatively low, and the model and observations have similar ranges (i.e. the standard deviation ratio is close to one). In NWS, the bias is lower in 2019 and higher in 2021, and the ranges of modelled and observed velocities are also more different from one another. Finally, the RMS error is also larger in NWS than in Eforie, especially during the 2019 campaign.

2.7.2 Process-oriented validation

The validation of the current velocity presented above, is particularly relevant for the service modules using currents as input data, e.g. service module A2.

2.8 Pilot 8: Italy

The Italian pilot is focused on the northern Adriatic Sea and it is mainly targeted to the aquaculture sector. The pilot is based on a coupled hydrodynamic-biogeochemical model (MITgcm-BFM) and on remote sensing observations of sea surface temperature (SST) and chlorophyll.

The numerical model is a downscaling of CMEMS (from 1/24° to 1/128° horizontal resolution) and it is currently running in operational mode, providing daily products of the main hydrodynamic (e.g., temperature, salinity, velocity) and biogeochemical (e.g., nutrients, oxygen, chlorophyll) variables. Each daily simulation covers a 10-day period: 7 days of hindcast and 3 days of forecast (from the current date), with hourly temporal resolution. The products are stored for one week on a dedicated THREDDS server (<https://dsecho.ogs.it/thredds/catalog/modelCatalog.html>), while the forecast (72 hours) is

published in the northern Adriatic section of the Mediterranean Ecosystem Analysis and Forecast (MedEAF) webpage (<http://medeaf.ogs.it/adriatic>).

The first Service Module applied to Pilot 8 is the Front Detection, but other services of interest to the end-users could be considered. Since Front Detection is focused on monitoring and forecasting mesoscale sea surface fronts, the validation is targeted to surface forecast products, by comparing model output and satellite data of temperature and chlorophyll. SST data are the L3 ultra-high resolution (1 km) NRT CMEMS products (SST_MED_SST_L3S_NRT_OBSERVATIONS_010_012), while chlorophyll data are the Sentinel3-OLCI L3 full resolution (300 m, single-sensor) NRT CMEMS products (OCEANCOLOUR_MED_CHL_L3_NRT_OBSERVATIONS_009_040).

2.8.1 Forecast validation

The Near Real Time skill performance aims at assessing the quality of physical and biogeochemical forecast products (i.e., to identify main biases and possible suspicious trends in the time series). Time series of surface chlorophyll-a are produced and updated on a daily basis starting from March 2022 (Figure 8.1). MITgcm-BFM model results are compared with satellite observations consisting of the new daily OC regional products from Sentinel3-OLCI full resolution. The analysis has been subdivided into 3 areas, to account for coastal - open sea differences. Areas shallower and deeper than 20 meters have been defined as coastal and open sea, respectively, while the results for the whole basin have also been considered. In general, the model underestimates chlorophyll in the coastal area, but model parameters are currently being updating to account for coastal processes and obtain better accuracy of the results.

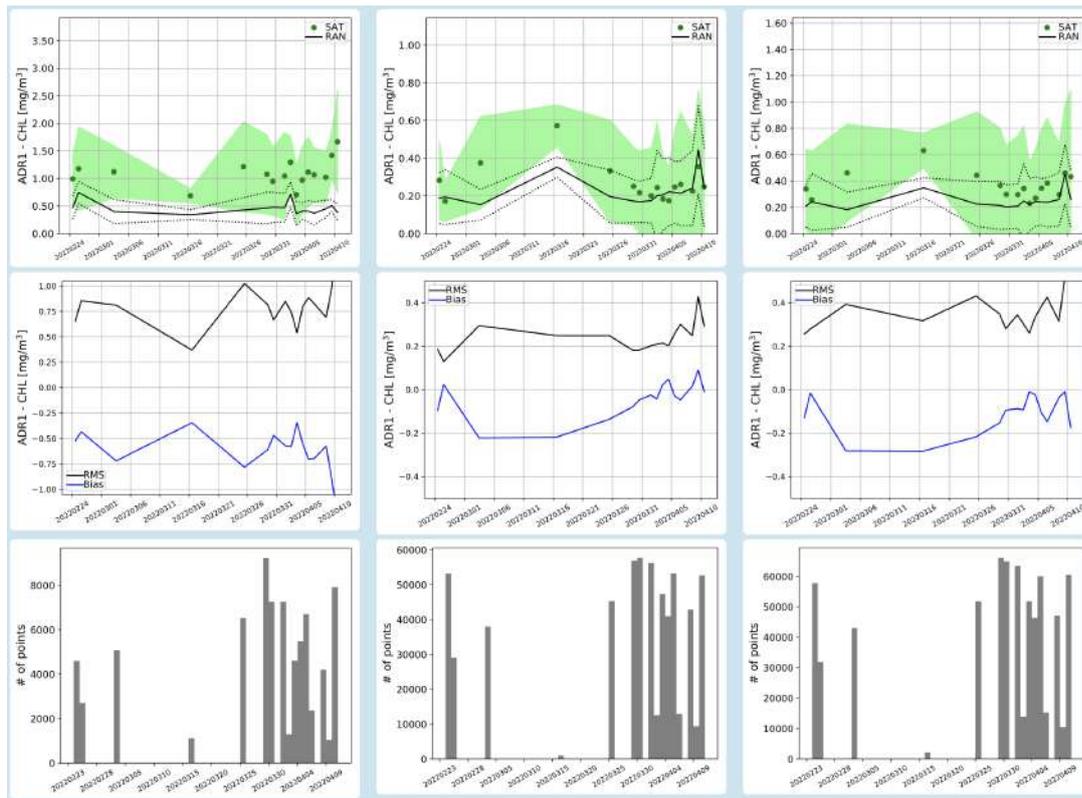


Figure 8.1 Time series of surface chlorophyll (+/- 1 standard deviation) from satellite (green) and model (black) - upper plots. Statistics of bias and root mean square difference (RMS) are computed daily for the first day of forecast - central plots. Number of points in each L3 satellite image - lower plots. Areas shallower and deeper than 20 meters are defined as COAST (left column) and OPEN SEA (central column), respectively. The results for the WHOLE BASIN are plotted on the right column. Image taken from the MedEAF webpage.

The last available maps of both chlorophyll (Figure 8.2) and SST (Figure 8.3) are published for a preliminary visual comparison. Overall, the model correctly reproduces chlorophyll concentration in the open sea, while it tends to underestimate chlorophyll in the shallow Italian coastal areas.

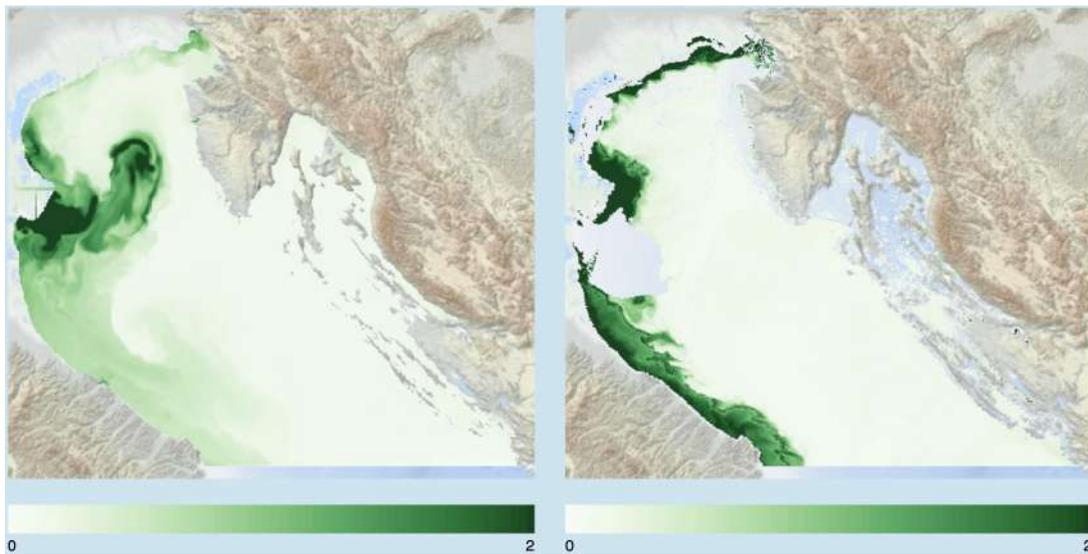


Figure 8.2 Surface chlorophyll [mg m^{-3}] maps derived from MITgcm-BFM model output (simulation date: 2022-04-12; midday of the first day of forecast - left) and remote sensing data provided by CNR-GOS (right). Image taken from the MedEAF webpage.

SST maps, in general, show a better match between model and remote sensing data.

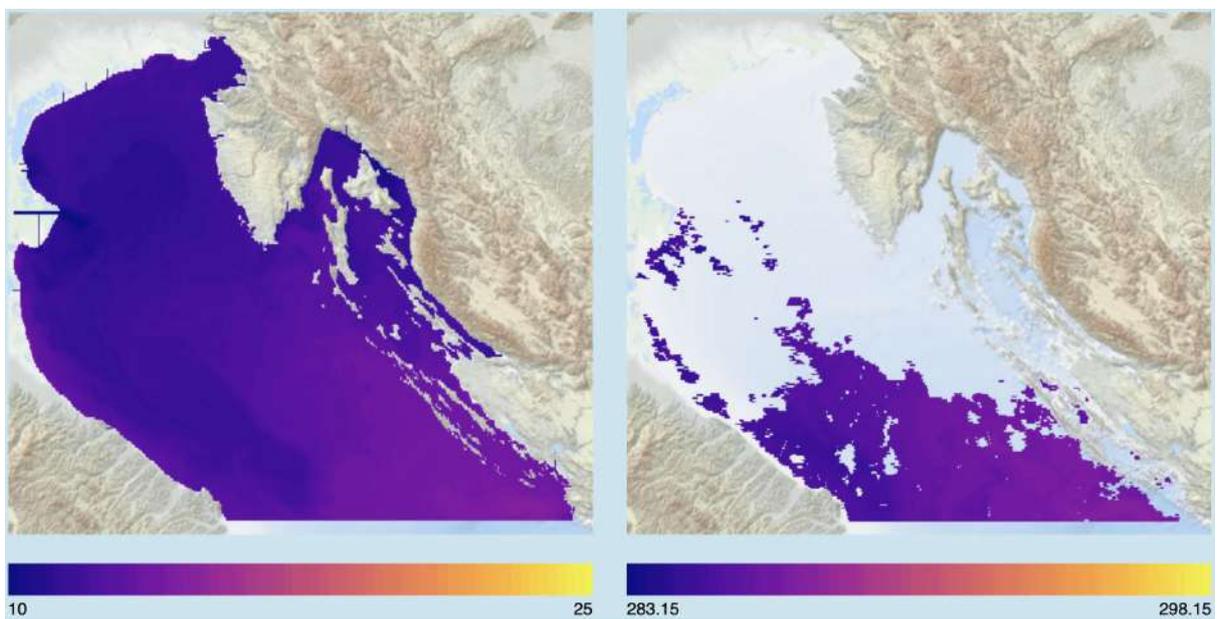


Figure 8.3 Sea surface temperature maps derived from MITgcm-BFM model output (simulation date: 2022-04-12; midnight of the first day of forecast [$^{\circ}\text{C}$] - left) and remote sensing data provided by CNR-GOS [K] (right). Image taken from the MedEAF webpage.

3. Conclusions

Services delivered in FORCOAST heavily rely on the predictions of coastal hydrodynamic and biogeochemical models. In this report, model predictions are compared against measurements from

different observational platforms, to assess the model performance in reproducing the dynamics and processes at the different Pilot sites.

Sections 2.1 to 2.8 above provide the results and conclusions of this effort with detail, providing figures and Estimated Accuracy Numbers to illustrate the goodness-of-fit between observations and predictions. This work is of paramount importance within the FORCOAST project to build up confidence on the coastal models and thus on the services being delivered that rely on such models.

The following table provides the main model validation conclusions for each pilot.

Table 2. Model validation per Pilot

Pilot	Conclusions and metrics
<p>1: Portugal</p>	<p>1. The scatter plots of tide gauge sea level recordings show that the model accurately predicts sea surface height.</p> <p>2. There is a good match between the water temperature recorded at the monitoring platform and the water temperature series predicted by the LisOcean model.</p> <p>3. However, the variability of the observed water temperature series is higher than the variability of the LisOcean predicted series.</p> <p>Metrics: CORR: Figure 1.7, Figure 1.8</p>
<p>2: Spain</p>	<p>1. The model is predicting the variability of both temperature and salinity, which is important for the fisheries sector.</p> <p>2. It is worth noting the capability of the model to reproduce both slow velocity and high velocity fields.</p> <p>3. Most of the observed front points are located at distances shorter than 5-11 km, showing a fairly good performance of the model for finding areas of higher fishing probability.</p> <p>Metrics: CORR: Table 2.2 RMSE: Table 2.2 ARMAE: Figure 2.12</p>
<p>3: Bulgaria</p>	<p>1. In general, the statistics show a slight underestimation of the measured data by the wave model.</p>

	<p>2. The validations against the available in-situ observations show that the model is capable of simulating the significant wave height near the coastal areas.</p> <p>3. This is especially a challenge for the wave modelling and substantiates the need for higher resolution data near the coastal areas.</p> <p>Metrics: SLOPE: Figure 3.2, Figure 3.3 CORR: Figure 3.2, Figure 3.3 RMSD: Figure 3.2, Figure 3.3 Bias: Figure 3.2, Figure 3.3 Scatter Index: Figure 3.2, Figure 3.3</p>
<p>4: Belgium</p>	<p>1. The model seems to be better at forecasting the sea surface temperature at Europlatform in terms of the variability than at the other stations.</p> <p>2. Statistical tests indicate that the BCZ model forecasts the background salinity with a better accuracy than NOS at Vlakte van de Raan.</p> <p>3. The temperature variability at Westdiep is very well captured by the model.</p> <p>Metrics: RMSE: Figure 4.4, Figure 4.5, Table 4.2, Table 4.3, Table 4.4, Table 4.5, Table 4.9 CORR: Table 4.3, Table 4.4, Table 4.5, Table 4.9 Bias: Table 4.3, Table 4.4, Table 4.5, Table 4.6, Table 4.9 Ratio: Table 4.3, Table 4.4, Table 4.5, Table 4.9</p>
<p>5: Ireland</p>	<p>1. Better results are obtained for the u-component of the current than for the v-component due to the predominant east-west direction of propagation of the tide.</p> <p>2. The model successfully captures the temperature and salinity vertical profiles at the Killeenaran farming site.</p> <p>3. The ROC analysis reflects an excellent performance of the model regarding its ability to reproduce the observed salinity at the Killeenaran farming site.</p> <p>Metrics: CORR: Table 5.1, Table 5.2, Table 5.3, Table 5.6, Table 5.7, Table 5.8 ME: Table 5.1, Table 5.3, Table 5.6, Table 5.7, Table 5.8 RMSD: Table 5.1, Table 5.2, Table 5.3, Table 5.6, Table 5.7, Table 5.8</p>

	<p>ARMAE: Table 5.2</p>
<p>6: Denmark</p>	<ol style="list-style-type: none"> 1. The FlexSem-ERGOM model manages to capture the spatial and seasonal development of the main biogeochemical tracers in the Limfjord. The nutrients and oxygen fields are especially well described. 2. The model is able to predict all the sea level extreme events at the Lemvig tide gauge station. 3. The model is able to capture diurnal warming events in the very shallow waters at KFO1. <p>Metrics: CORR: Section 2.6.1.1 Table 6.2 cRMSE: Section 2.6.1.1, Figure 6.5, Figure 6.6, Figure 6.7, Figure 6.8 Bias: Figure 6.5, Figure 6.6, Figure 6.7, Figure 6.8, Figure 6.17, Figure 6.20 RMSE: Table 6.2, Figure 6.17, Figure 6.20 ME: Table 6.2 ARMAE: Table 6.2</p>
<p>7: Romania</p>	<ol style="list-style-type: none"> 1. The Eforie model presents a positive temperature bias during summer and negative temperature bias during winter. 2. RMSD shows that most of the error concentrates close to the coast during winter months, whereas during spring-summer-autumn, the error is spread over in the whole domain. 3. The Eforie model presents evidence of small structures that cannot be reproduced in the 3-km CMEMS model. <p>Metrics: CORR: Figure 6.14, Table 6.2 cRMSE: Figure 6.7, Figure 6.8 RMSE: Table 6.2 ME: Table 6.2 ARMAE: Table 6.2</p>

8: Italy

1. Overall, the model correctly reproduces chlorophyll concentration in the open sea.
2. However, it tends to underestimate chlorophyll in the shallow Italian coastal areas.
3. Model parameters are currently being updated to account for coastal processes and obtain better accuracy of the results.

Metrics:

RMS: Figure 8.1

4. References

- Campuzano, F.**, Juliano, M., Angélico, M. M., Oliveira, P., Neves, R. Western iberia sea surface salinity patterns due to land inputs. In Proceedings of the 5a Jornadas da Engenharia Hidrográfica, Lisboa, Portugal, 19–20 July 2018.
- Chin, T. M.**, Vazquez-Cuervo, J., Armstrong, E. M. (2017). A multi-scale high-resolution analysis of global sea surface temperature. *Remote Sens. Environ.* 200, 154-169.
- Dabrowski, T.**, Lyons, K., Nolan, G., Berry, A., Cusack, C., Silke, J., 2016. Harmful algal bloom forecast system for SW Ireland. Part I: Description and validation of an operational forecasting model. *Harmful Algae* 53:64-76. doi:10.1016/j.hal.2015.11.015
- Maar, M.**, Timmermann, K., Petersen, J. K., Gustafsson, K. E., and Storm, L. M. (2010). A model study of the regulation of blue mussels by nutrient loadings and water column stability in a shallow estuary, the Limfjorden. *J. Sea Res.* 64, 322–333.
- Mateus, M.**, Riflet, G., Chambel, P., Fernandes, L., Fernandes, R., Juliano, M., Campuzano, F., H. de Pablo, Neves, R. (2012). An operational model for the West Iberian coast: products and services. *Ocean Sci.*, 8 (4), pp. 713-732
- Murawski, J.**, She, J., Mohn, C., Frishfelds, V., Nielsen, J. W. (2021). Ocean Circulation Model Applications for the Estuary-Coastal-Open Sea Continuum, *Frontiers in Marine Science*, vol. 8, 2021,
- Sotillo, M. G.**, Cailleau, S., Lorente, P., Levier, B., Aznar, R., Reffray, G., Amo-Baladrón, A., Chanut, J., Benkiran, M., Alvarez-Fanjul, E. (2015) The MyOcean IBI Ocean Forecast and Reanalysis Systems: operational products and roadmap to the future Copernicus Service, *Journal of Operational Oceanography*, 8:1, 63-79.
- Staneva, J.**, Behrens, A., Ricker, M., & Gayer, G. (2020). Black Sea Waves Reanalysis (CMEMS BS-Waves) (Version 2) [Data Set]. *Copernicus Monitoring Environment Marine Service (CMEMS)*.