Assessment of model skill in the Black Sea

1. Description of the model system and the hindcast experiments

The Black Sea model system consists of three models, namely a physical model, a lower-trophic-level model (LTL) and a higher-trophic level model (HTL), coupled in an end-to-end fashion. The physical model used for the Black Sea is an implementation of the Princeton Ocean Model (POM) and its domain includes the whole of the Black Sea except the Sea of Azov. It incorporates parameterisations for the water fluxes at the Bosphorus and the Kerch straits and includes nine of the largest rivers that flow into the Black Sea. The model was forced using DMI reanalysis atmospheric fields and river discharge rates obtained from the Black Sea Commission's river database. The physical model was initialized using World Ocean Atlas fields, spun up for five years and then the hindcast simulation was performed for the period 1990-2009. The resulting physical fields were stored daily in order to force the LTL model.

The LTL model contains thirty state variables that include four phytoplankton types, four zooplankton types, oxygen, hydrogen sulphide, inorganic nutrients and detritus in both nitrogen and phosphorus currencies as well as the carbonate system variables (see Figure 1). The model was initialized using typical winter profiles and spun up for five years in order to reach a statistically steady state. Nutrient loads obtained from the Black Sea Commission's river database were used to prescribe nitrate, ammonium and phosphate fluxes at the river mouths. Unlike the physical model in which the entire water column is represented, the LTL model only considers the water column down to 200m depth. Phosphate, ammonium, hydrogen sulphide and dissolved carbon-dioxide fluxes were prescribed at the open boundary at 200m. The atmospheric partial pressure of carbon dioxide required for the carbonate system was obtained from the Mauna Loa Observatory. The hindcast simulation was performed for the period 1990-2009 using the physical forcing generated by POM.

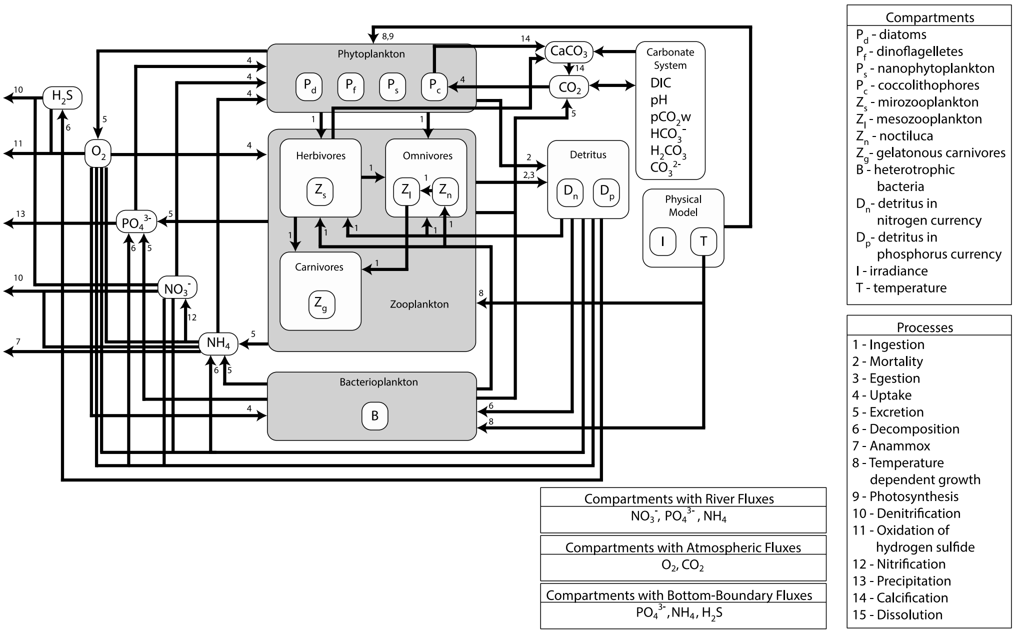


Figure 1. The lower-trophic-level compartments and their interactions within the Black Sea model.

The HTL model used in this project for the Black Sea is a FORTRAN implementation of the Ecopath with Ecosim (EwE) model developed at IMS-METU. The model includes 13 fish species and is forced using spatial averages of the required LTL compartments. Figure 2 shows the interaction between the LTL and HTL compartments.

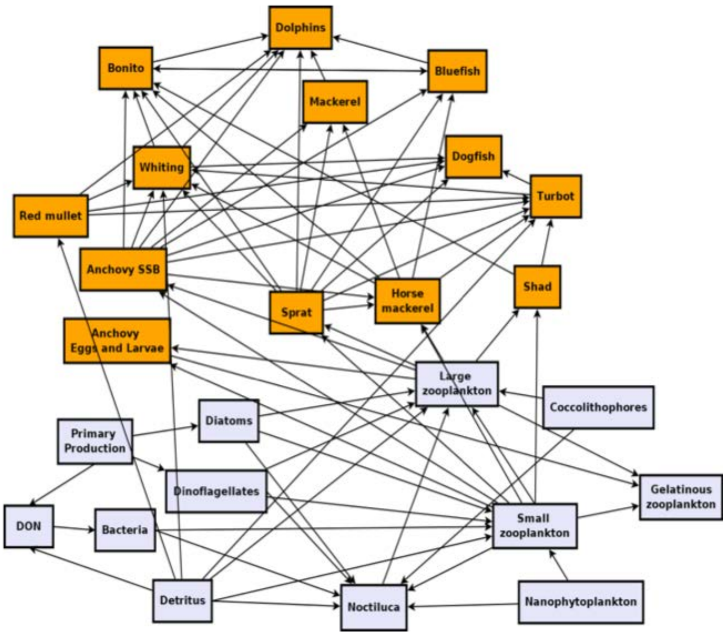


Figure 2. The interactions between the higher-trophic-level (orange) and the lower-trophic-level (blue) compartments.

2. Validation

2.1 Validation Data and Procedure

The hindcast run was validated using a database that was collated using data from IMS-METU cruises as well as fields form the Black Sea Ocean Database. The validated physical variables are temperature, salinity and the density derived variables, potential energy anomaly and mixed layer depth. The validated LTL variables are nitrate, phosphate, ammonium, dissolved oxygen, hydrogen sulphide as well as in-situ chlorophyll. The reference data was interpolated onto model grid and the monthly averages of both the model output and the reference data was fed into the benchmarking tool for validation. The validation for the HTL model was performed by comparing model output against the VPA Virtual Population Analysis (VPA) biomass estimations and catch statistics.

2.2 Validation of the Physical Model

Every single data point in the BS database has been compared with the monthly average of the grid cell closest to the sampling location and a series of model-data comparison metrics have been calculated. Model validation over the entire basin using 191551 and 192982 data points is displayed in Taylor and Target diagrams for temperature and salinity (Figures 3 & 4), respectively. The Taylor diagrams show that the model reproduces both temperature and salinity quite well. This is an essential result to highlight the capability of the model for ecosystem forecast studies.

Model performance indices indicate a high reliability (1.09 for temp and 1.01 for salinity). The model shows a lower standard deviation than observations for both temperature and salinity but the correlation coefficients are very high (0.904 for temperature and 0.88 for salinity). The unbiased RMSD is 2.1 for temperature and 0.66 for salinity. The target diagrams (on the right) shows that the normalized model bias is low and negative for temperature as well as salinity, with a normalized total RMSD of 0.45 and 0.52, respectively.

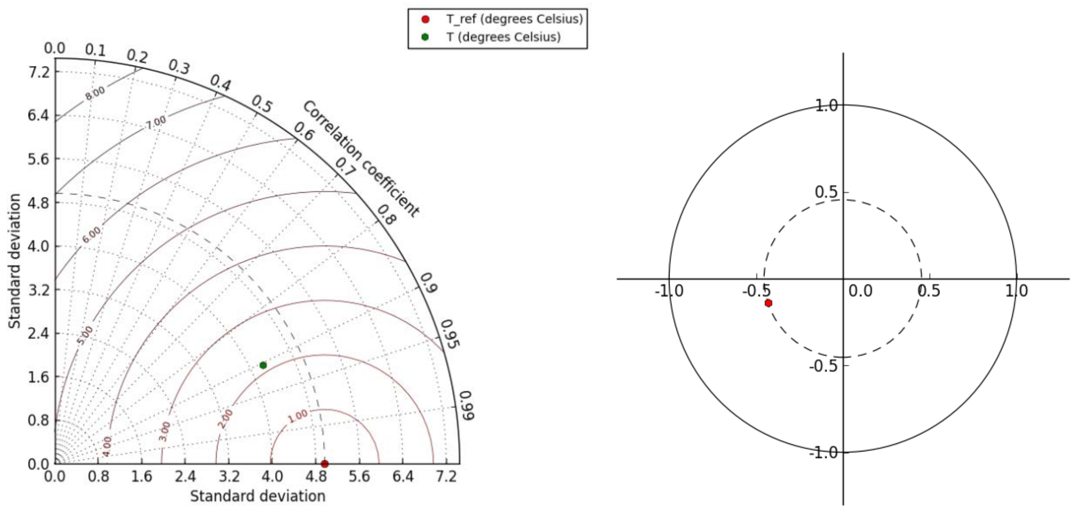


Figure 3. Outcomes of point to point validation of the hindcast temperature on the entire domain vs. the in situ Black Sea database (N=191551). On the right the Taylor diagram, on the left the target diagram.

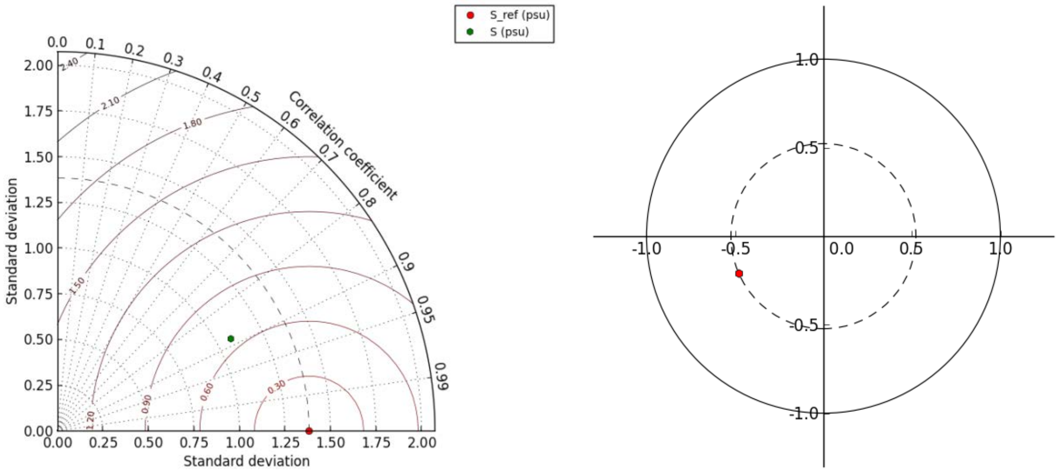


Figure 4. Outcomes of point to point validation of the hindcast salinity on the entire domain vs. the in situ Black Sea database (N= 192982). On the right the Taylor diagram, on the left the target diagram.

The model performs slightly less well with respect to mixed layer depth (MLD) and potential energy anomaly (PEA) with the reliability index climbing to 1.44 and 1.56, respectively.

MLD provided by the model has a lower standard deviation than observations, hence does not reproduce the observed variability. The correlation coefficient is 0.58 and the RMSD of the mixed layer depth lies at 12% of the natural variability of this parameter (Figure 5). The Target diagram shows that there is small positive bias with a normalized total RMSD of 0.81.

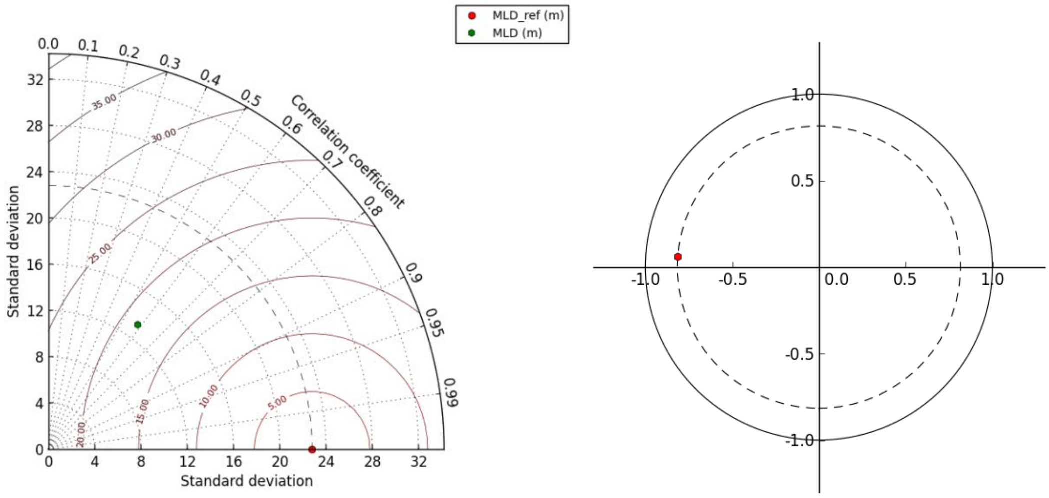


Figure 5. Outcomes of point-to-point validation of the hindcast mixed layer depth on the entire domain vs. the in situ Black Sea database (N= 3625). On the right the Taylor diagram, on the left the target diagram.

PEA as calculated by the model also shows a lower standard deviation than observations, hence does not reproduce the observed variability. The correlation coefficient is high with 0.79 and the RMSD of the mixed layer depth lies at 15% of the natural variability of this parameter (Figure 5). The Target diagram shows that there is very small negative bias with a normalized total RMSD of 0.61.

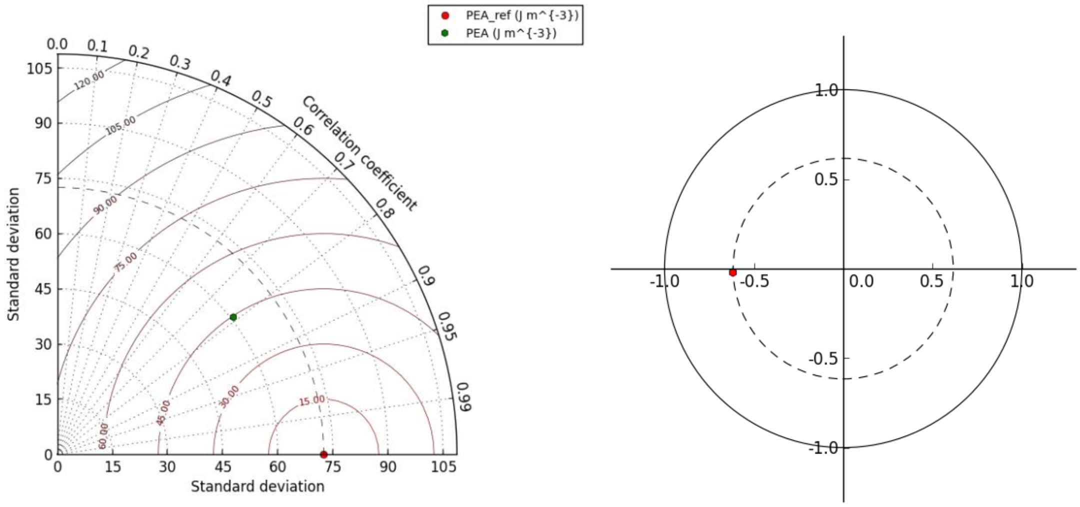
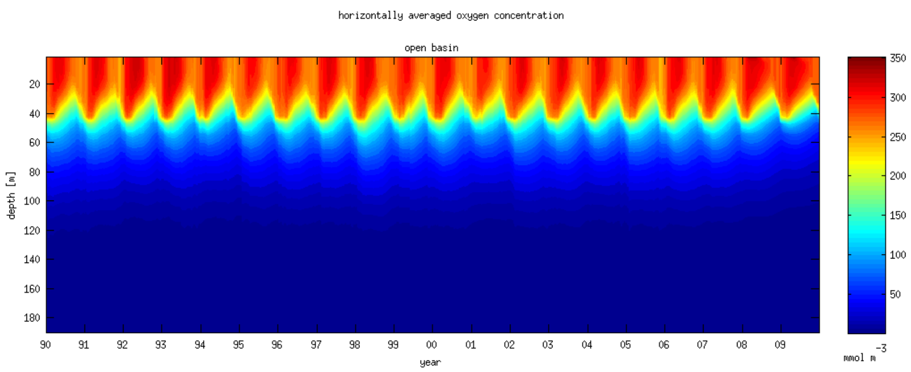


Figure 6. Outcomes of point-to-point validation of the hindcast potential energy anomaly on the entire domain vs. the in situ Black Sea database (N= 3625). On the right the Taylor diagram, on the left the target diagram.

2.3 Validation of the LTL Model

The presence of oxygen is a crucial factor in determining the biogeochemical reactions that take place within the Black Sea, and therefore it is important that the model is able to predict oxygen and hydrogen sulphide concentrations relative accurately. Looking at Figures 7 & 8, we can see that the model does quite well with respect to oxygen and hydrogen sulfide concentrations. Figure 69 shows that the model is able to capture the average profiles of these concentrations as well as their seasonal variations. The results of the benchmarking show that the model output is highly correlated with the reference data (correlation coefficient of 0.9) and the RMSE is less than 50% of standard deviation of the reference data.



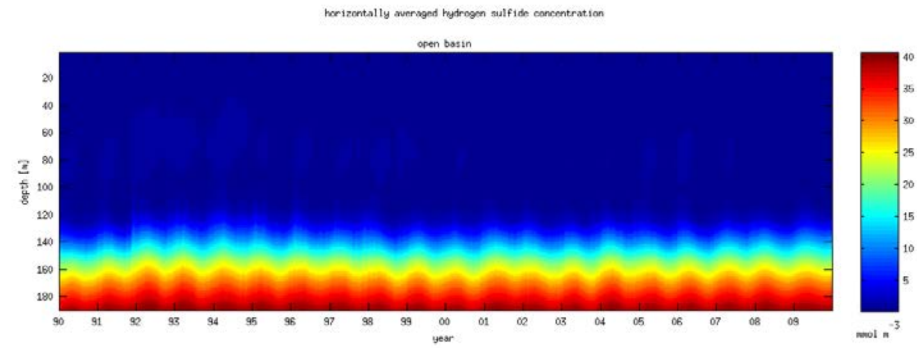


Figure 7. Horizontally averaged oxygen and hydrogen sulphide concentrations within the deep Black Sea (>200m) as a function of depth and time.

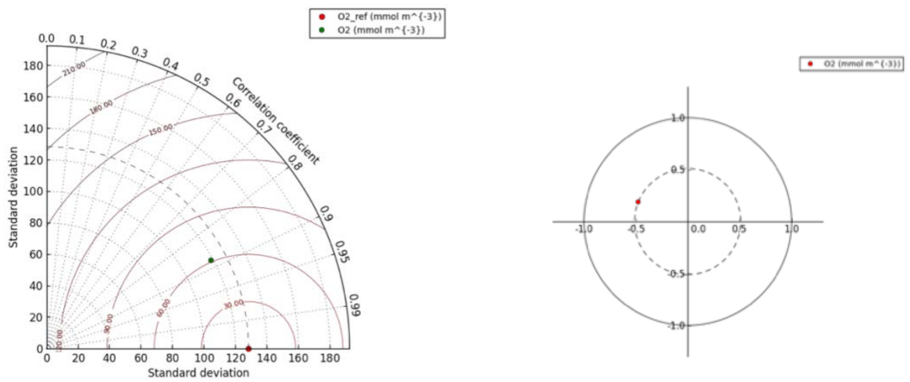


Figure 8. Taylor and Target diagrams for oxygen concentration.

Figures 9 & 10 show the evolution of the average nitrate and ammonium profiles within the deep parts of the Black Sea and the overall structure is consistent with the observations. Nitrate peak is situated at approximately 75m depths and it vanished at around 120m. Ammonium concentrations are high at high depths at vanish at around 120m. Despite the fact that the model does well in terms of capturing the overall profile of nitrate the peak concentrations (2-3 mmol m^{-3}) are lower than what they should be in reality (4-5 mmol m^{-3}). This mismatch is reflected in the benchmarking (Figure 10) whereby the model's correlation with the data is much lower than that of oxygen (correlation coefficient of 0.5) and the RMSE is more than 50% of the standard deviation of the reference data. It is also evident that the model exhibits much less variability than the reference data.

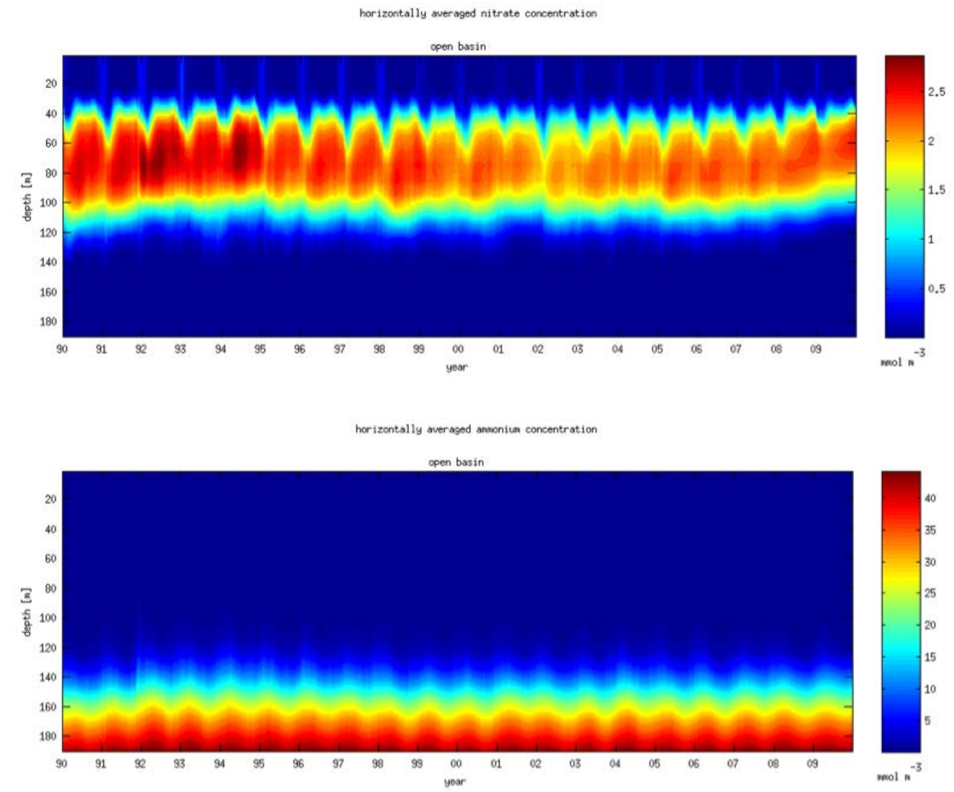


Figure 9. Horizontally averaged nitrate and ammonium concentrations within the deep Black Sea (>200m) as a function of depth and time.

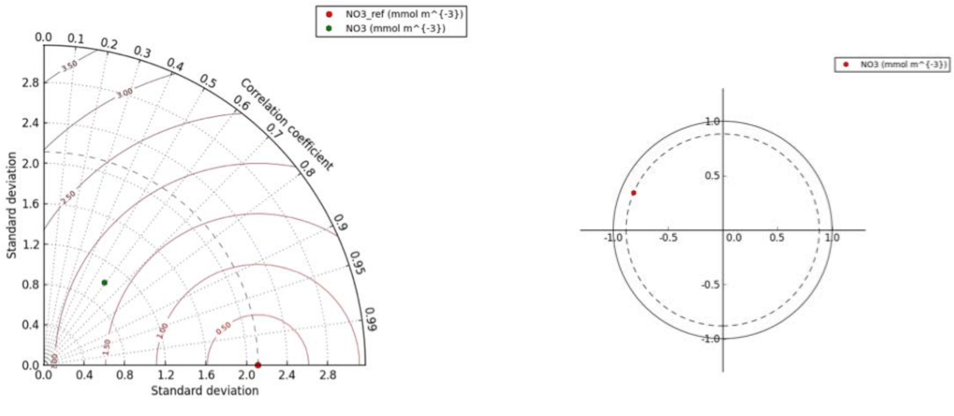


Figure 10. Taylor and Target diagrams for nitrate concentration.

The model does worse in predicting phosphate concentrations than nitrate concentrations. The results of the benchmarking show that the correlation between the model and the reference data for phosphate is about 0.4 and the RMSE is more than 1.75 mmol m^{-3} which is approaching 100% of the standard deviation in the reference data.

Figure 11 compares the horizontally averaged surface chlorophyll concentrations within the deep (>200m) parts of the Black Sea with the data obtained from the GlobColour product. It shows that there is a mismatch between the maximum bloom timing in the model and the reference data. The maximum bloom timing in the model is in January-February while that in the reference data is

November-December. Furthermore, the model consistently underestimates surface chlorophyll (Figure 11). The mismatch in bloom timing results in a low correlation coefficient of 0.45 between the data and model (Figure12). We believe the model's inability to accurately represent surface chlorophyll concentrations is the result of the fact that the physical model is not able to reproduce the mesoscale dynamics present with the Black Sea. A new setup for the physical model that yields better results in terms of capturing the mesoscale features will be used for the simulations in WP3 and WP4.

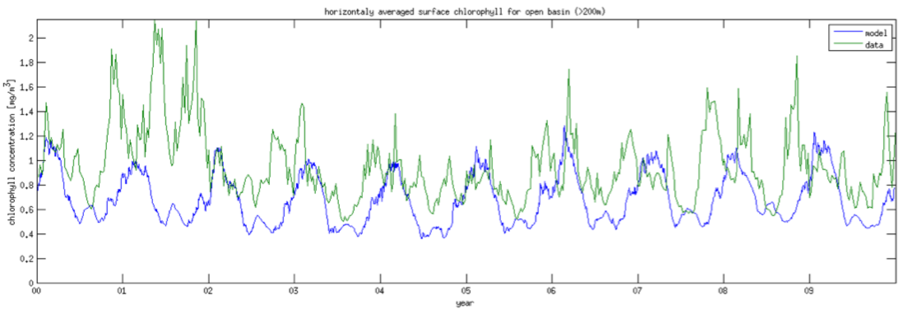


Figure 11. Comparison of the horizontally averaged surface chlorophyll in the model and the data obtained from the GlobColour product.

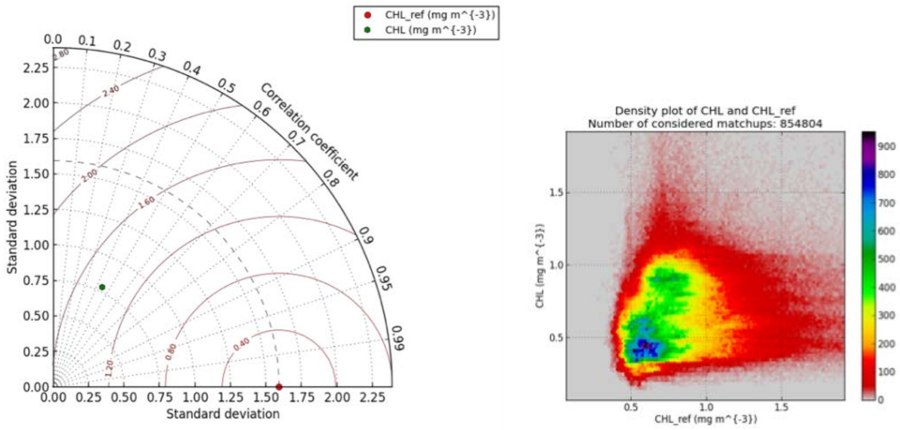


Figure 12. Taylor diagram and density plot for surface chlorophyll.

In-situ measurements of carbonate system properties are extremely scarce for the Black Sea. As a result, using the benchmarking tool to validate the models performance with respect to the carbonate system properties is not a useful exercise. Figures 13-16 compare the model results with six profiles obtained between 1988-2003. These show that the model captures the main features of the carbonate system properties very well.

(a) (b)

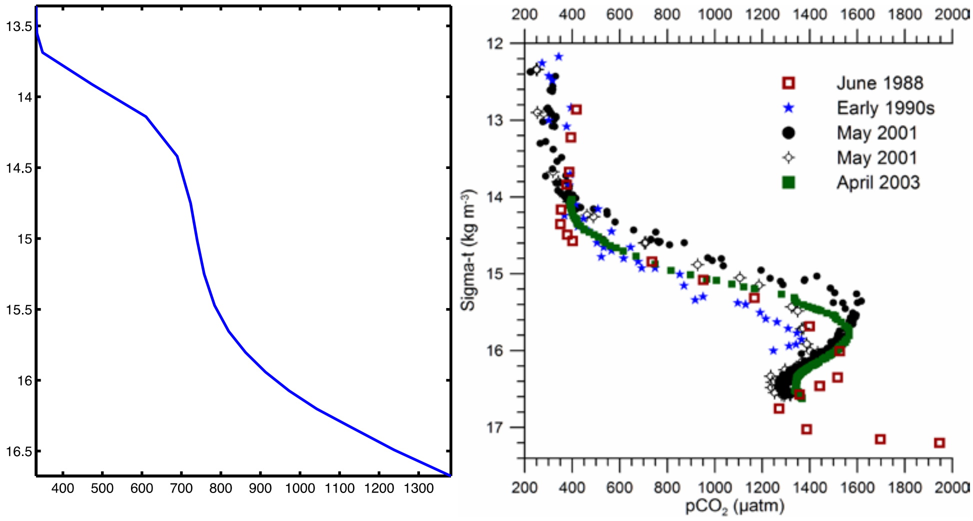


Figure 13. A comparison of Sigma-t vs. partial pressure of carbon dioxide in (a) the model and (b) data compiled from the interior basin during the early 1990s, June 1988, May 2001 and April 2003 Knorr measurements conducted within the west-central basin, and along the northeastern Black Sea at station 44.516oN, 37.872oE during July 2, 2002 and 44.415oN, 37.317oE during January 26, 2004. The profile in (a) is the average profile of the deep basin (>1000m depth) for the entire simulation period.

(a) (b)

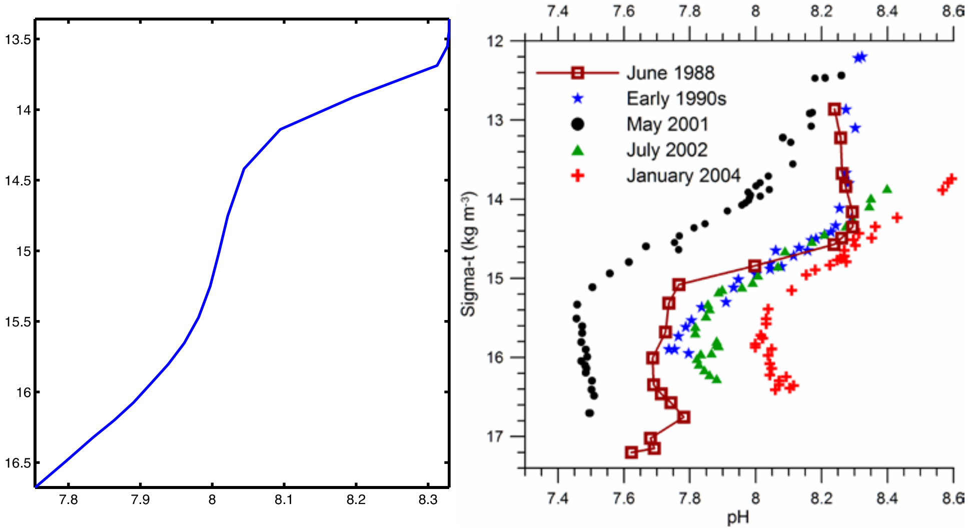


Figure 14. A comparison of Sigma-t vs. pH in (a) the model and (b) data described in the caption of Figure 13.

(a) (b)

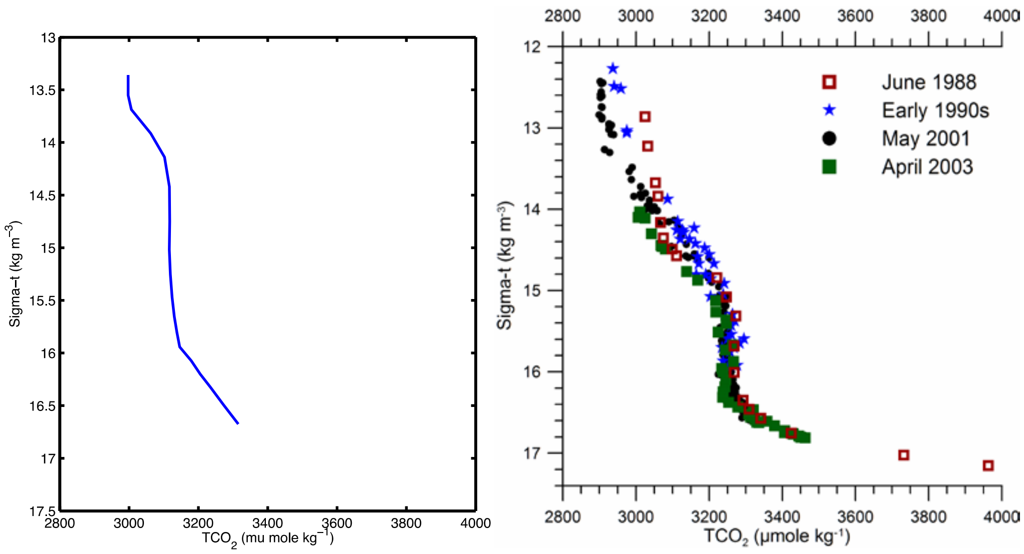


Figure 15. A comparison of Sigma-t vs. total dissolved inorganic carbon in (a) the model and (b) data described in the caption of Figure 13.

(a) (b)

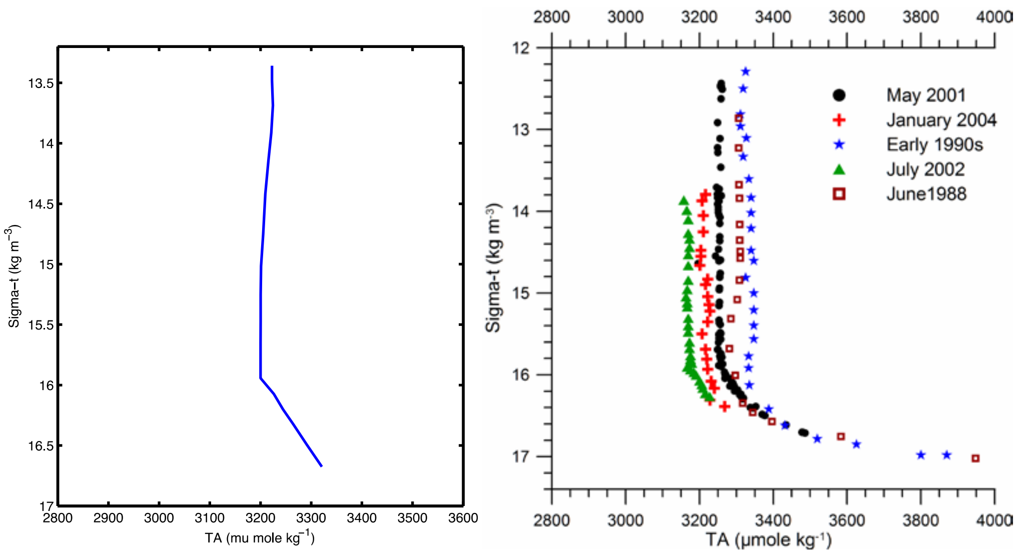


Figure 16. A comparison of Sigma-t vs. total alkalinity in (a) the model and (b) data described in the caption of Figure 13.

2.4 Validation of the HTL Model

The basin-wide VPA biomass estimations and the statistical survey results of fisheries landings in the Black Sea have been compared with the annual average of the basin-wide hindcast simulation results of biomass and catches of fish groups and a series of comparison metrics have been calculated. Model validation results over the entire basin using 20 data points are displayed in Taylor and target diagrams for Anchovy SSB, Anchovy Eggs and Larvae, Sprat, Shad, Horse mackerel, Whiting, Turbot, Red mullet, Spiny dogfish, Atlantic mackerel, Bluefish and Bonito (Figures 17-20), respectively. The Taylor diagrams show that the model reproduces biomass and catch values reasonably. However, further fitting and parameter tuning exercises are still necessary before proving the capability of the model for ecosystem forecast studies. The overall model performance indices indicate a medium reliability (~ 1.4).

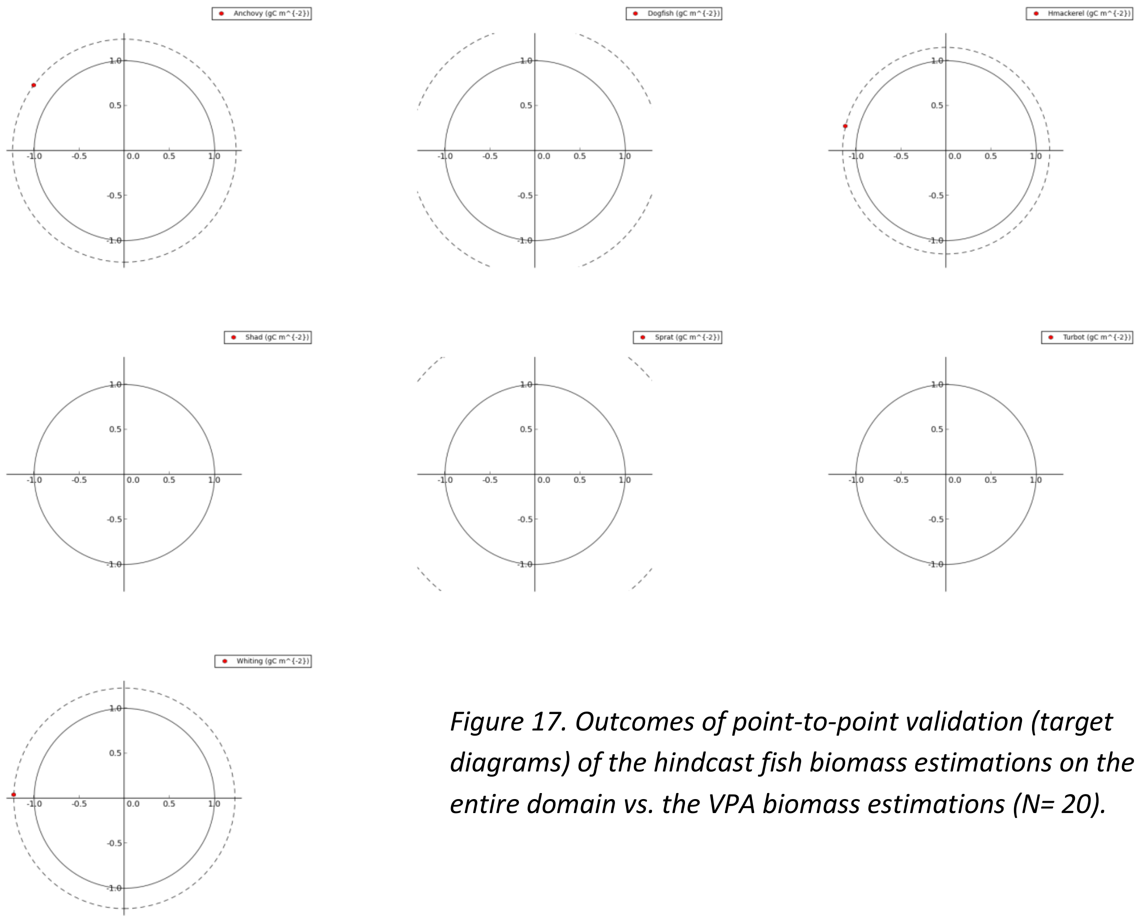


Figure 17. Outcomes of point-to-point validation (target diagrams) of the hindcast fish biomass estimations on the entire domain vs. the VPA biomass estimations (N= 20).

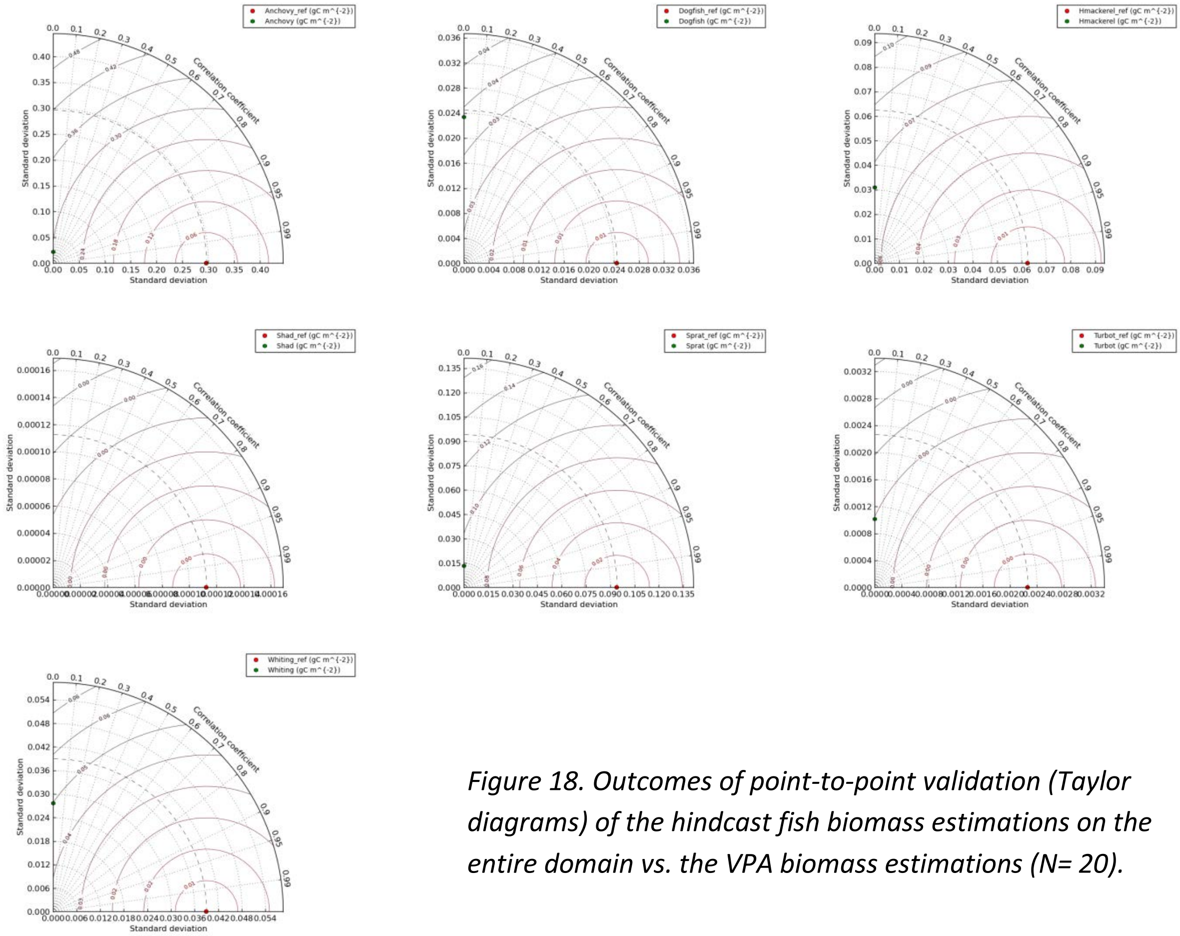
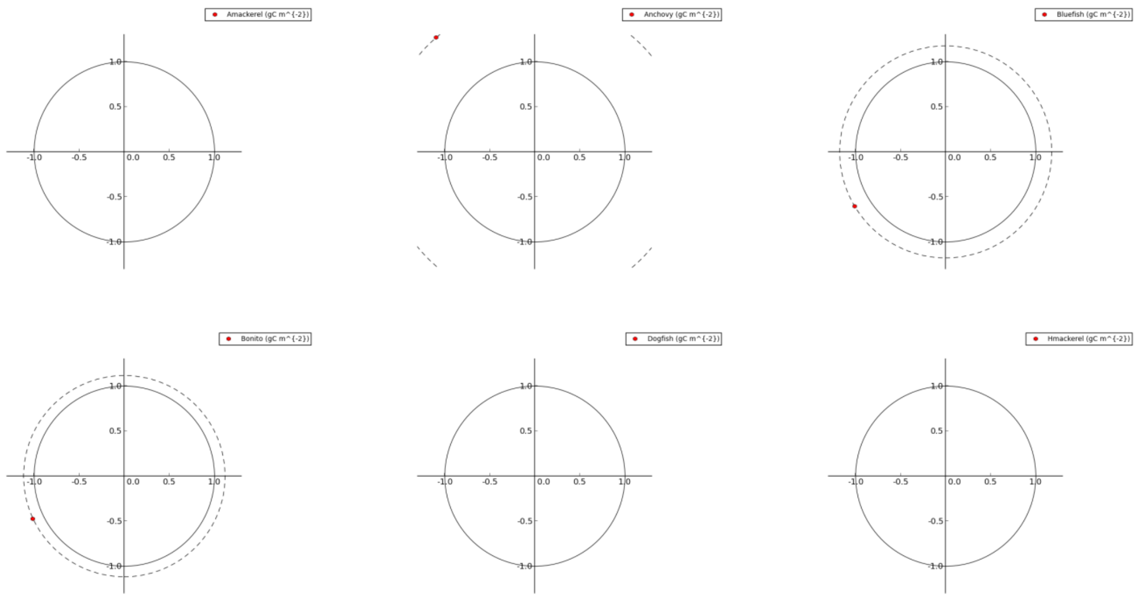


Figure 18. Outcomes of point-to-point validation (Taylor diagrams) of the hindcast fish biomass estimations on the entire domain vs. the VPA biomass estimations (N= 20).



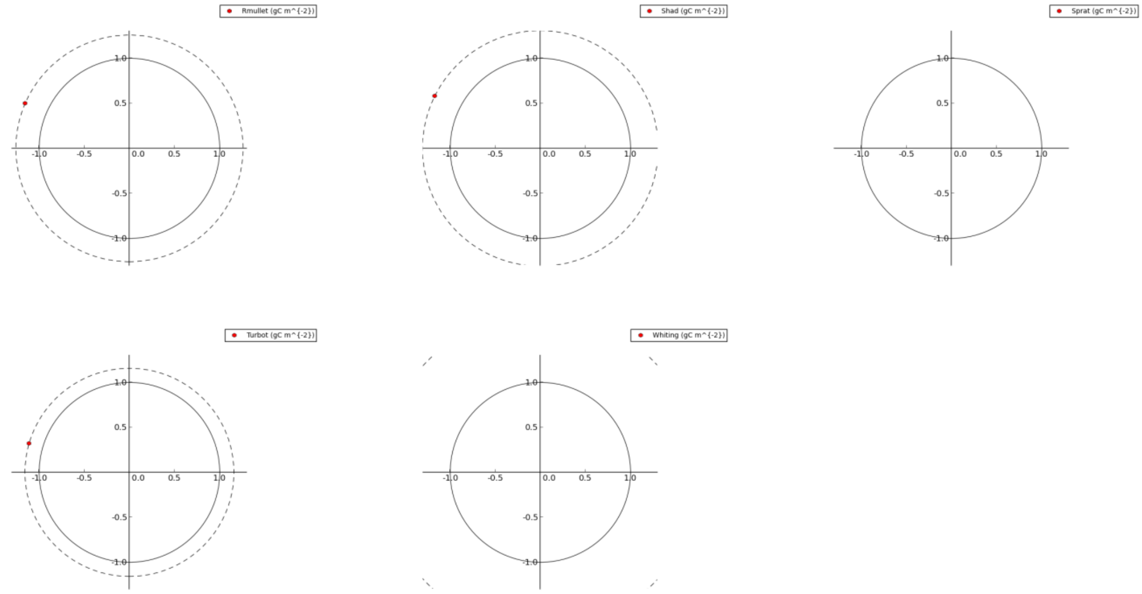
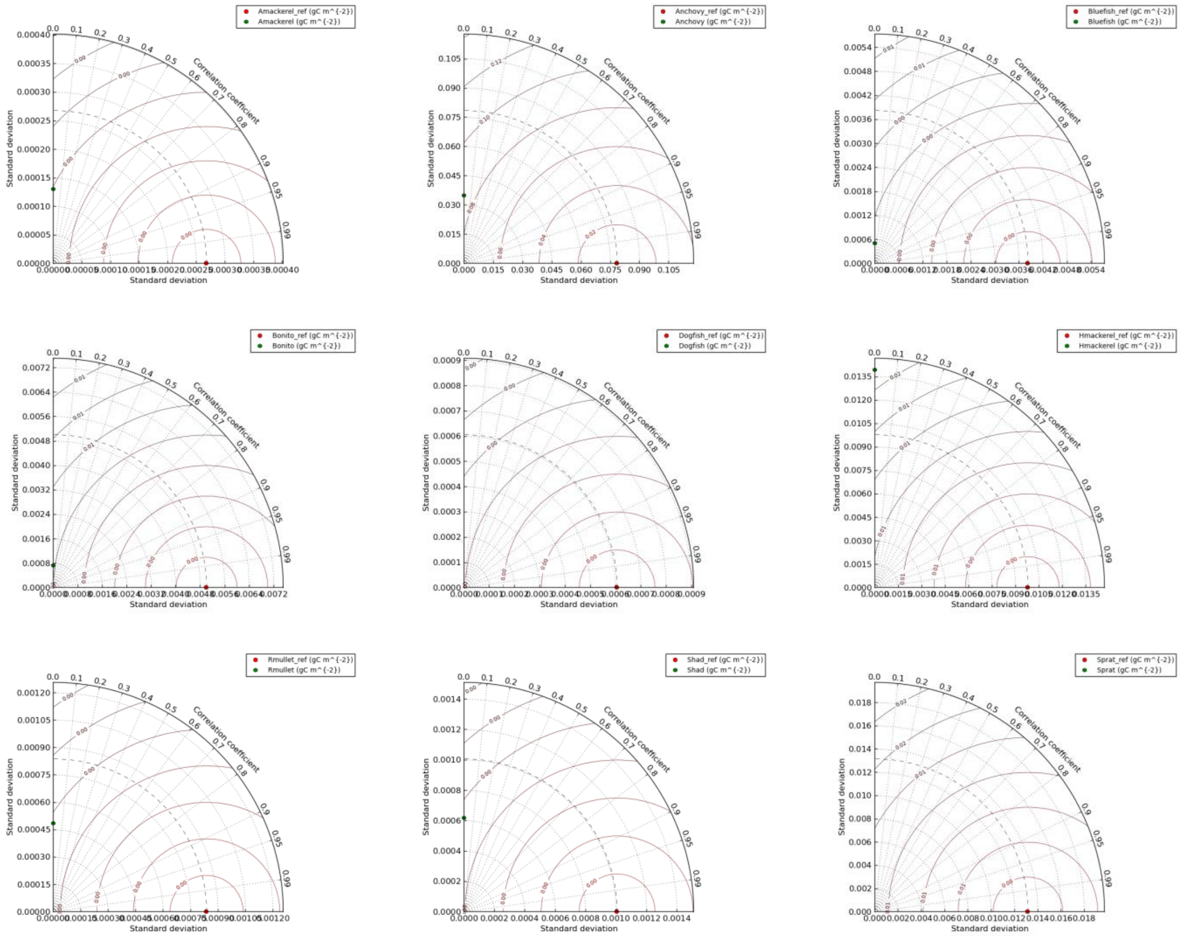


Figure 19. Outcomes of point-to-point validation (target diagrams) of the hindcast fish catch estimations on the entire domain vs. the Black Sea catch statistics (N= 20).



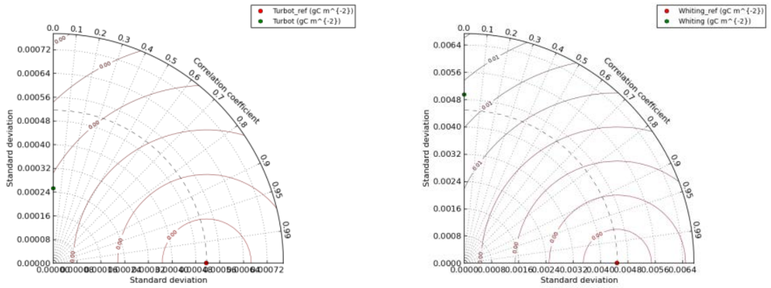


Figure 20. Outcomes of point-to-point validation (Taylor diagrams) of the hindcast fish catch estimations on the entire domain vs. the Black Sea catch statistics (N= 20).

3. Results

Table 1 shows the results of the validation of the hindcast run. Model skill scores are colour coded to give an indication of model skill as follows; Good = green, Moderate = Black, Poor = Red; correlation r good > 0.75, poor < 0.20, ME good = > 0.5, Poor = < 0, Reliability index good 0.8-1.2, poor > 2 or –ve. PBias, good = < 20%, poor = > 100%.

Table 1

Region: Black Sea
Model: BIMS-ECO, EwE
Met forcing: DMI regional hindcast
Hindcast time period: 1990-2009
Reanalysis time period: 2000-2009
Contact: Sinan S. Arkin sinan.arkin@ims.metu.edu.tr, Baris Salihoglu baris@ims.metu.edu.tr

		Hindcast					
Variable	**Unit**	**Model Efficiency**	**% Model Bias**	**Pearson Correlation Coefficient**	**RMSE**	**Unbiased RMSE**	**Reliability Index**
Temperature	Degrees Celcius	0.7944	-6.7145	0.904	2.2516	2.1377	1.0922
Salinity	psu	0.7267	-0.2953	0.885	0.723	0.66	1.0194
Potential Energy Anomaly	$J\ m^{-3}$	0.6204	-1.4318	0.7894	44.67	44.6489	1.5602
Mixed Layer Depth	m	0.2657	3.0786	0.5644	11.9879	11.9815	1.4874
Nitrate	$mmol\ m^{-3}$	-0.531	47.6117	0.1567	6.7533	6.6132	inf
Phosphate	$mmol\ m^{-3}$	0.0404	79.6059	0.5572	1.5939	1.4157	inf
Ammonium	$mmol\ m^{-3}$	-1.8438	0.2769	0.5168	5.7347	5.728	inf
Dissolved Oxygen	$mmol\ m^{-3}$	0.3705	22.5354	0.7638	95.8554	82.5332	inf
Hydrogen Sulfide	$mmol\ m^{-3}$	0.3176	-27.9064	0.715	10.1167	9.8078	inf
Chlorophyll	$mg\ m^{-3}$	0.2439	-40.3049	0.5314	1.1279	1.1027	inf
Anchovy	$gC\ m^{-2}$	-0.5362	66.4437	-0.0006	0.3675	0.2974	1.6169

Biomass							
Dogfish Biomass	gC m^{-2}	-0.9128	-1.0164	0.0003	0.0339	0.0339	1.4136
H. mackerel Biomass	gC m^{-2}	-0.3191	19.5314	-0.0001	0.0718	0.0698	1.6982
Shad Biomass	gC m^{-2}	-52.431	3.9542	0.0004	0.0008	0.0008	1.1574
Sprat Biomass	gC m^{-2}	-1.7703	44.6028	0.0002	0.1567	0.0951	1.3267
Turbot Biomass	gC m^{-2}	-2.4347	50.5212	0.0001	0.0042	0.0025	1.4388
Whiting Biomass	gC m^{-2}	-0.5023	2.4565	0.0002	0.0478	0.0478	1.3972
A. mackerel Yield	gC m^{-2}	-16.2562	-276.357	-0.0012	0.0011	0.0003	2.0183
Anchovy Yield	gC m^{-2}	-1.8106	55.3553	0.0003	0.1316	0.0859	1.5744
Bluefish Yield	gC m^{-2}	-0.3855	-41.68	-0.0001	0.0045	0.0039	1.5887
Bonito Yield	gC m^{-2}	-0.2477	-38.7859	0.0001	0.0056	0.0051	1.4563
Dogfish Yield	gC m^{-2}	-19.9727	-118.671	0.0001	0.0028	0.0019	1.51
H. mackerel Yield	gC m^{-2}	-8.3838	-219.068	-0.0001	0.03	0.017	1.8712
Red mullet Yield	gC m^{-2}	-0.583	20.7015	0	0.0011	0.001	1.2595
Shad Yield	gC m^{-2}	-0.7085	39.8473	0.0002	0.0013	0.0012	1.5431
Sprat Yield	gC m^{-2}	-16.5993	-139.328	0.0004	0.0552	0.0375	1.6153
Turbot Yield	gC m^{-2}	-0.3391	12.4455	-0.0002	0.0006	0.0006	1.2565
Whiting Yield	gC m^{-2}	-2.1018	-42.8198	0.001	0.0079	0.0067	1.3584

4. Summary of hindcast experiment and lessons learnt

The Black Sea hindcast experiment for the 1990-2009 period was performed using a coupled modelling system that was developed at IMS-METU. The modelling system consists of a physical model, a lower-trophic level model and a higher-trophic level model run in an offline, one-way coupled fashion. The hindcast validation demonstrated that the physical model is able to predict temperature and salinity quite well. However, comparison of satellite chlorophyll observations with model output suggests that the physical model lacks the ability to generate mesoscale dynamics. We believe that the upwelling caused by the geostrophic flow at the boundaries of the mesoscale eddies is the major cause of the high chlorophyll concentrations observed by satellite close to the periphery of the Black Sea. We have put a significant amount of effort into developing a new physical model that is able to exhibit mesoscale variability and we plan to employ this new model for the simulations required in WP3 and WP4. Generally, the validation demonstrates that the predictive ability of our models decreases as one proceeds up the chain of models. Therefore the physical model has the highest capability, which is followed by the inorganic nutrients and then the lower-trophic-level compartments. The higher-trophic-level compartments have the least amount of predictability.