



ESA Contract No.	4000135758/21/I-EF
Project Name	BIOMONDO
Towards Earth Observation supported monitoring of freshwater biodiversity	

Deliverable	D2.3 Product Validation Report
Short description	A detailed specification of the validation methods, metrics and description of the reference data.
Work package No.	WP2
Lead Partner	BC
Distribution	ESA, AB, BEs, EAs
Version	V2
Submission date	2022-04-21
Contributors	Jorrit Scholze (Brockmann Consult) Carsten Brockmann (Brockmann Consult) Tamara Keijzer (PBL) Miguel Dionisio Pires (Deltares) Tineke Troost (Deltares) Petra Philipson (Brockmann Geomatics) Jelle Lever (EAWAG) Daniel Odermatt (EAWAG)

Table of Contents

1	Introduction	6
2	Scope of this document	6
3	Pilot 1 – Eutrophication	8
3.1	<i>Model</i>	8
3.2	<i>Remote Sensing</i>	12
3.2.1	In situ Measurements	12
3.2.2	Matchup Analysis	12
3.2.3	Timeseries Analysis	13
3.2.4	Spatial Variability	14
4	Pilot 2 – Heat Tolerance of Fish	19
4.1	<i>Model</i>	19
4.2	<i>Remote Sensing</i>	19
4.2.1	In situ Measurements	20
4.2.2	Matchup Analysis	20
4.2.3	Timeseries Analysis	22
5	Pilot 3- River Dams	24
5.1	<i>Pareto front optimization and Connectivity Model</i>	24
5.2	<i>Remote Sensing</i>	24
5.2.1	In situ Measurements	25
5.2.2	Timeseries Analysis	25
5.2.3	Matchup Analysis	26
5.2.4	Visual Inspection	27
6	References	29

List of Figures

Figure 1 Matchup of model data against in situ measurements for chl-a-a in Lake Markermeer..... 9

Figure 2 Comparison between model and in situ data for gross primary production in Lake Markermeer for the year 2020..... 10

Figure 3 Comparison between model and in situ data for gross primary production in Lake Markermeer for the year 2021..... 10

Figure 4 *Target diagrams for the modelled gross primary production in all areas in lake Markermeer in which measurements were taken, both for 2020 (upper panel) and 2021 (lower panel). The X-axis shows the normalised unbiased Root Mean Square Deviation, on the Y-axis the normalised Bias is displayed.* 11

Figure Scatterplot of EO and modelled chl-a for two stations (central and nearshore) with matchups within one day..... 12

Figure Scatterplot of EO and in situ for two stations (Pampus Oost and Lelystad Haven) with matchups within one day..... 13

Figure Timeseries of Pampus Oost for chlorophyll-a EO, in situ and Delft3D in 2016. ... 13

Figure Timeseries of the central station for chlorophyll-a EO, in situ and Delft3D in 2016..... 14

Figure Maps for two dates in 2016 comparing chl-a for modelled chl-a (left, Delf3D) and EO chl-a (right, ESA CCI Lakes). 15

Figure Maps for two dates in 2016 comparing the forcing fields of water temperature on basis of in situ data (left) and on basis of EO data (right). Note that the y-axis differs between the two dates. 16

Figure Maps for two dates in 2016 comparing PP modelled on basis of in situ temperature (left) and on basis of EO temperature (right)..... 17

Figure Correlation plot of GPP versus temperature for all areas of the lake in which measurements were taken, showing both the in situ measurements (black dots) as well as the model outputs (coloured dots) in 2021..... 18

Figure SLU in situ stations for LSWT of Lake Mälaren. Basemap: Openstreetmap. 20

Figure Scatterplot for Lake Mälaren-Prästfjärden of EO DINEOF LSWT and in situ LSWT. 21

Figure Scatterplot for Lake Mälaren-Granfj. Djurgårds Udde of EO DINEOF LSWT and in situ LSWT..... 21

Figure Scatterplot for Lake Marken-Central Station of EO DINEOF LSWT and in situ LSWT. 22

Figure Timeseries of 2013 and 2015 of the central station in Lake Marken. DINEOF EO LSWT is shown in blue and in situ LSWT in orange..... 23

Figure Timeseries of the in situ station Savannakhet in the Mekong, Laos, blue shows the EO retrieved TUR and brown the in situ TSS measurements. Data provided by Mekong River Commission and reproduced with permission. 26

Figure Timeseries of the in situ station Tan Chau in the Mekong, Vietnam, blue shows the EO retrieved TUR and brown the in situ TSS measurements. Data provided by Mekong River Commission and reproduced with permission.	26
Figure Scatterplot of EO TUR data in situ TSS data provided by Mekong River Commission and reproduced with permission.	27
Figure Heatmap for EO TUR of the Kg. Chhnang station in Cambodia.	28
Figure Heatmap for EO TUR of the Savannakhet station in Laos.	28

List of Tables

Table 1 References of related documents.....7
Table 2 Information about the used stations of the MRC. 25

1 Introduction

Freshwater ecosystems are some of the most important ecosystems on the planet, providing a range of ecosystem services to humans, including clean water, food, and recreation. They are also home to a vast array of biodiversity, including many species that are found nowhere else on earth. Unfortunately, freshwater ecosystems are under threat from a range of human activities, including pollution, habitat destruction, and climate change. Current biodiversity policies and strategies as well as assessments of progress towards set targets, point out that there has been a general failure to halt the negative trend of biodiversity loss and that different approaches are needed to reverse the situation. This includes revision of targets and the indicators that inform the targets and a greater emphasis on the links between biodiversity, ecosystems and their services and people.

The European Space Agency (ESA) activity called Biodiversity+ Precursors is a contribution to the joint EC-ESA Earth System Science Initiative launched in February 2020 to jointly advance Earth System Science and its response to the global challenges that society is facing. The ESA Biodiversity+ Precursors include four projects on different themes and BIOMONDO is the freshwater project, and has a focus on biodiversity in lakes, wetlands, rivers, and streams.

BIOMONDO aims to improve our understanding of freshwater biodiversity around the world and to support freshwater biodiversity monitoring through development of solutions that integrate EO data and state-of-the-art biodiversity modelling using advanced data science and information and communications technology. Three BIOMONDO Pilots have been developed and will demonstrate how novel Earth Observation and Biodiversity modelling products can be integrated to enhance scientific understanding and support decision systems for biodiversity monitoring addressing policy priorities such as the EU Biodiversity Strategy for 2030.

To develop a broad outlook on ongoing changes in freshwater biodiversity and how these changes can be monitored using EO data, our three BIOMONDO pilots each address pilot objectives and knowledge gaps corresponding to one of the following three drivers of global environmental change in freshwater ecosystems: ‘pollution and nutrient enrichment’ (Pilot 1 - Eutrophication), ‘climate change’ (Pilot 2 – Heat Tolerance of Fish), and ‘habitat change’ (Pilot 3 – River Connectivity). The resulting combination of data from Earth Observation, in-situ measurements and model outputs is made available to the scientific and policy community through the BIOMONDO Freshwater Biodiversity Laboratory.

2 Scope of this document

Within the BIOMONDO project, data from various sources has been collected and own data products have been produced to develop the three BIOMONDO pilot studies for freshwater system biodiversity observation and monitoring.

This document describes the validation of the products collected and processed by the BIOMONDO consortium. Validation activities are performed on different levels. The al-

gorithm validation is performed in parallel to the algorithm development and is often an iterative process. Validation was performed for EO products using collected in situ data from different data sources. The in situ data sources are described respectively in this document. Furthermore, the model output data products are validated with EO data products, and the approach is described in detail.

The following validation approaches are mainly used to assess the quality of the data products:

- Visual inspection
- Match-up analyses
- Data ranges and statistics
- Time series
- Maps

Table 1 shows the reference documents where additional information can be accessed for the description of the algorithms and the experimental datasets used in the pilot studies.

Table 1 References of related documents

Document	Version	Short description
Algorithm Theoretical Baseline Document	D2.2 v2.0	A detailed specification of the final versions of the algorithms/models.
Experimental Datasets	D2.4 v1.0	A detailed description of the Experimental Datasets.

3 Pilot 1 – Eutrophication

Nutrient concentrations have increased substantially in lakes and rivers throughout the world, resulting in eutrophication, harmful algal blooms, loss of submerged macrophytes affecting sedimentation and turbidity, and biodiversity loss.

In this BIOMONDO pilot we explore the possibilities of integrating EO data into the Delft3D model suite to investigate the potential contribution by EO data to the model performance. Within the validation study we compared EO data, in situ data and the model outputs.

3.1 Model

Validation of the Delft3D model was done for chlorophyll-a at three monitoring sites of Lake Markermeer: Pampus Oost, (southern part of the lake) Lelystad Haven (north-east), and the central point. Figure 1 shows the results of the comparison between model and in situ data. The RMSE in all three graphs is high indicating that for this parameter, the model does not accurately estimate the concentrations.

The model performance is better when considering primary production (PP) instead of chlorophyll. For both years 2020 and 2021 the modelled PP values match quite well with the seasonal patterns and absolute ranges of the observed PP-values (see Figure 2 and Figure 3). Target diagrams show that the model results fall in the area within the unit circle for which the model fit can be qualified as 'reasonable' (Figure 4).

The seeming mismatch in model performance regarding chlorophyll-a and PP can be (at least partly) explained by the fact that the model only calculates active chlorophyll that is contributing to primary production, while the observed chlorophyll in winter is not active (as can be derived from the low primary production rates observed in winter, see Figures 2 and 3). The underlying explanation is that during spring/summer the algae attach to suspended solids and then sink to the bottom where they accumulate. Resuspension of these algae-solids particles in winter leads to higher chlorophyll-a concentrations but thus not to higher primary production in winter.

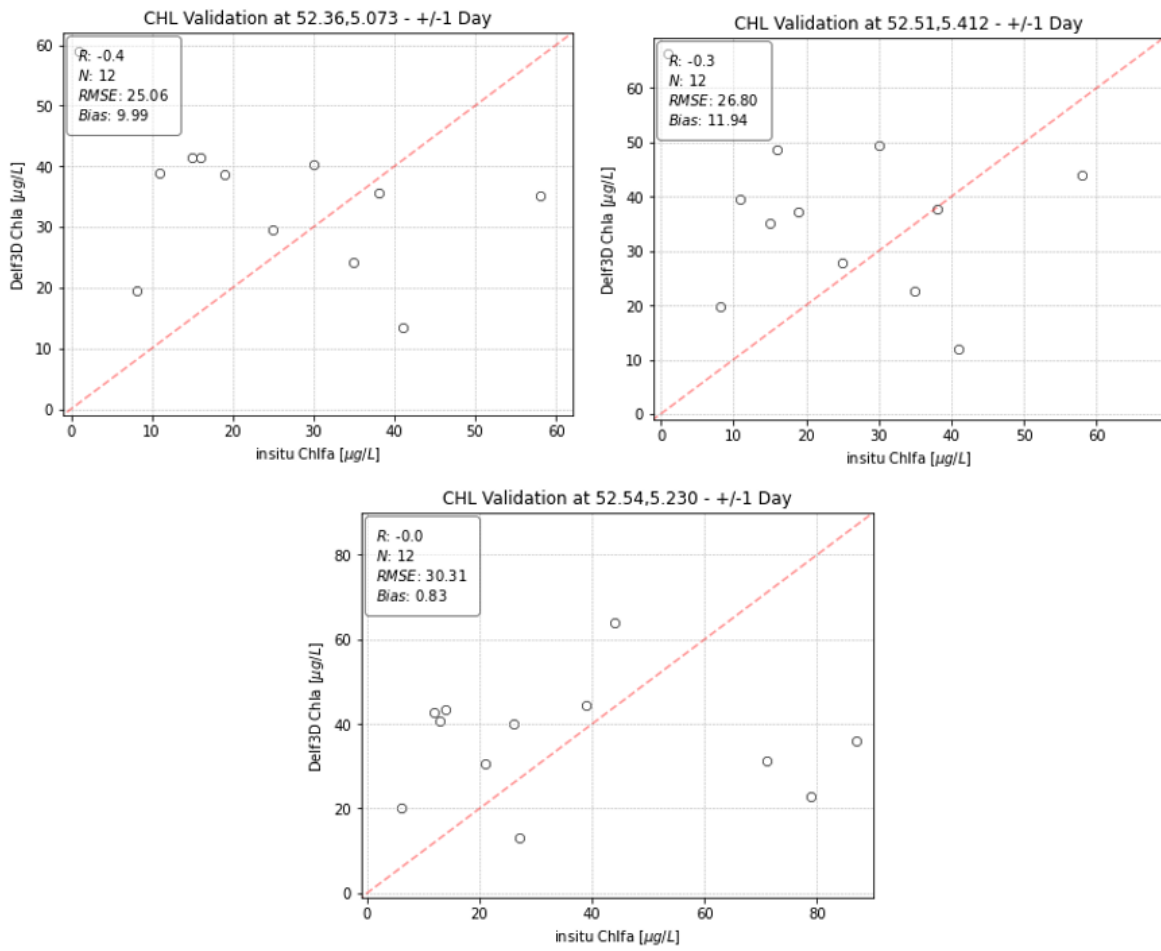


Figure 1 Matchup of model data against in situ measurements for chl-a-a in Lake Markermeer.

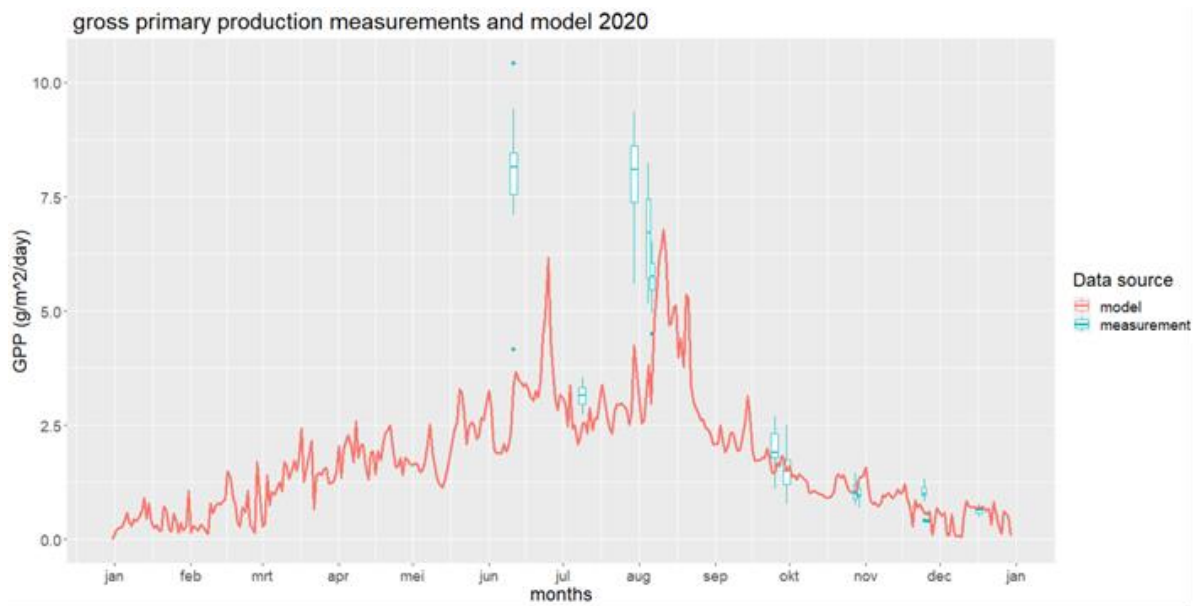


Figure 2 Comparison between model and in situ data for gross primary production in Lake Markermeer for the year 2020.

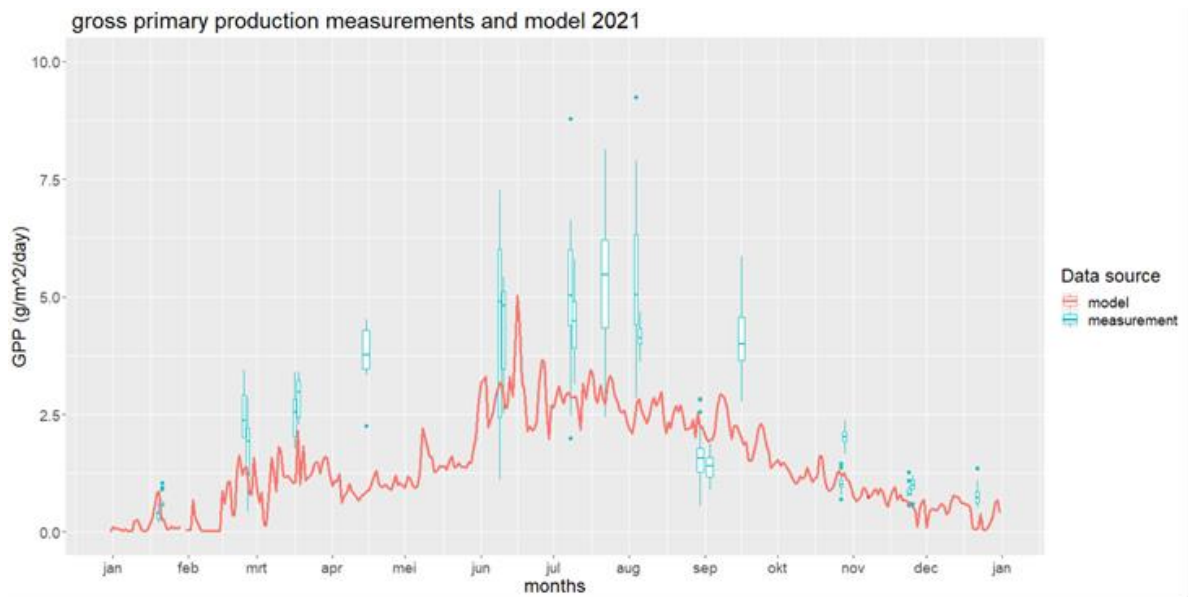


Figure 3 Comparison between model and in situ data for gross primary production in Lake Markermeer for the year 2021.

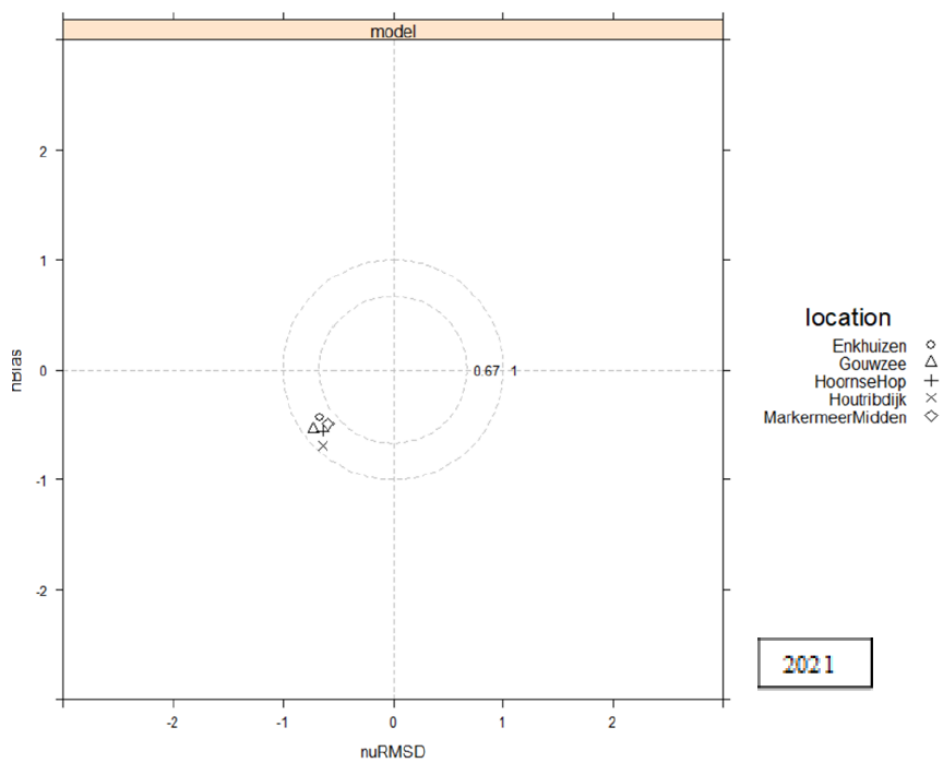
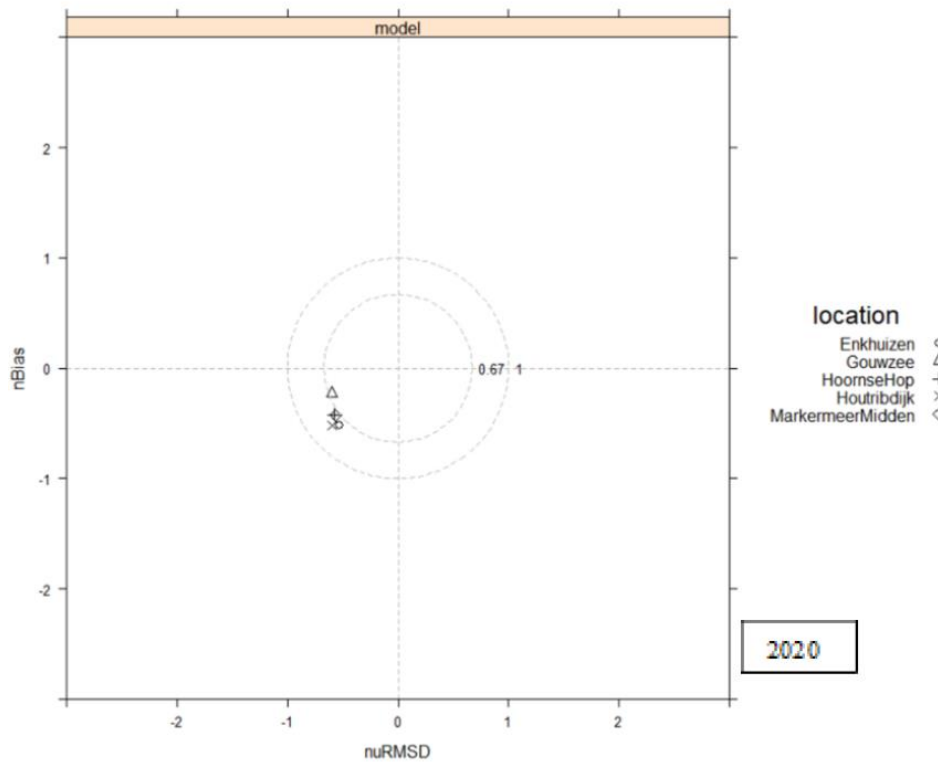


Figure 4 Target diagrams for the modelled gross primary production in all areas in lake Markermeer in which measurements were taken, both for 2020 (upper panel) and 2021 (lower panel). The X-axis shows the normalised unbiased Root Mean Square Deviation, on the Y-axis the normalised Bias is displayed.

3.2 Remote Sensing

For validation of the EO data in Pilot 1 we compared the ESA CCI Lakes v2.0.1 chlorophyll-a concentration (see Experimental Datasets D2.4) variable to in situ measurements. In addition, a comparison of chlorophyll-a from EO, in situ and the model output from the Delft3D was performed, as it is available for all three methods.

3.2.1 In situ Measurements

Rijkswaterstaat is part of the Ministry of Infrastructure and Water Management of the Netherlands. Its role is the practical execution of the public works and water management, including the construction and maintenance of waterways and roads, and flood protection and prevention. The chlorophyll -a concentration sampled at a depth of 100 cm, from multiple stations in the Markermeer, was used for the validation.

3.2.2 Matchup Analysis

The Scatterplots in Figure 5 shows the correlation between EO and the modelled chlorophyll-a concentration for 2016 with matchups within one day for two points (central and nearshore). With a RMSE of > 25 and a bias of $> 12 \mu\text{g/l}$ the correlation between both methods is very low. Especially for the model predicted values of $30\text{-}50\mu\text{g/l}$ the EO chlorophyll-a values are lower with ranges between $5\text{-}20\mu\text{g/l}$.

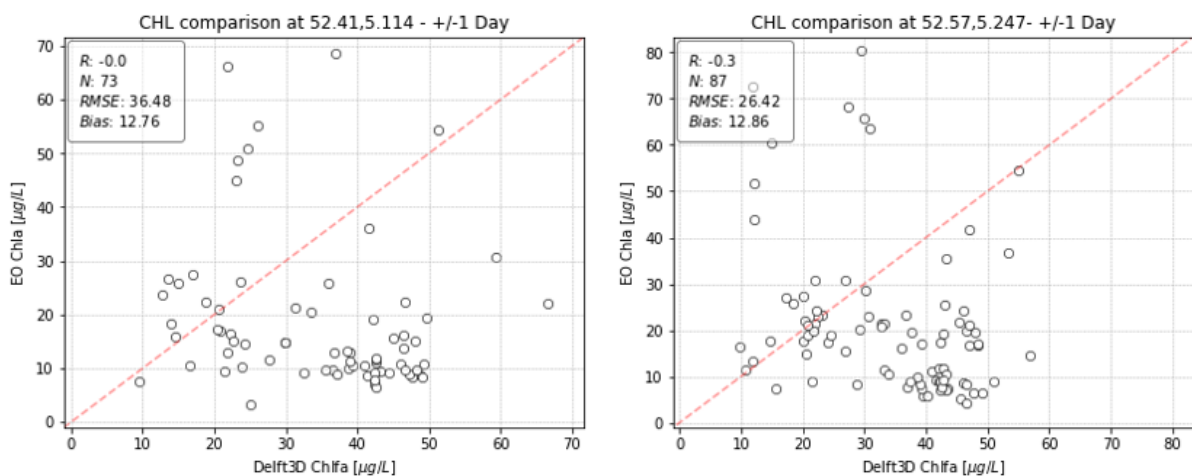


Figure 5 Scatterplot of EO and modelled chlorophyll-a for two stations (central and nearshore) with matchups within one day.

Figure 6 shows the correlation between EO and in situ measurements for two stations with matchups within one day. The RMSE is < 12.5 and the bias is $< 5\mu\text{g/l}$. Compared to the correlation between EO and modelled chlorophyll -a of the Delft 3D this shows a higher correlation. Both stations show outlier with over and underestimations for EO chlorophyll-a, which might happen for natural reasons as aquatic environments shift quickly and as the in situ sample from 1 meters depth not always matches to the situation at the surface, especially when cyanobacteria is present.

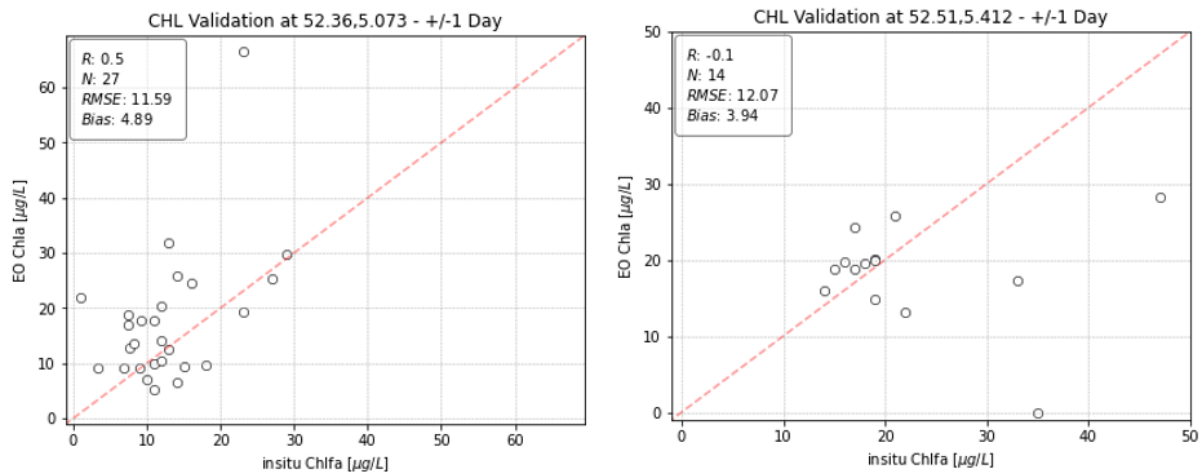


Figure 6 Scatterplot of EO and in situ for two stations (Pampus Oost and Lelystad Haven) with matchups within one day.

3.2.3 Timeseries Analysis

Timeseries comparing all three methods were created for Pampus Oost and central station (Figure 7 and Figure 8). Both timeseries show that the modelled chlorophyll-a from the Delft 3D is higher in the summer month compared to the other methods. For the Pampus Oost station the in-situ measurements and the EO chlorophyll-a are showing similar trends with a spring bloom around March and a summer bloom around September. For the central station the spring bloom in April predicted by the model is not present for the other two methods. EO data from Nov-Feb is not included in the graph as the number of observations and quality of data is reduced during winter due to high cloud cover and low sun angles.

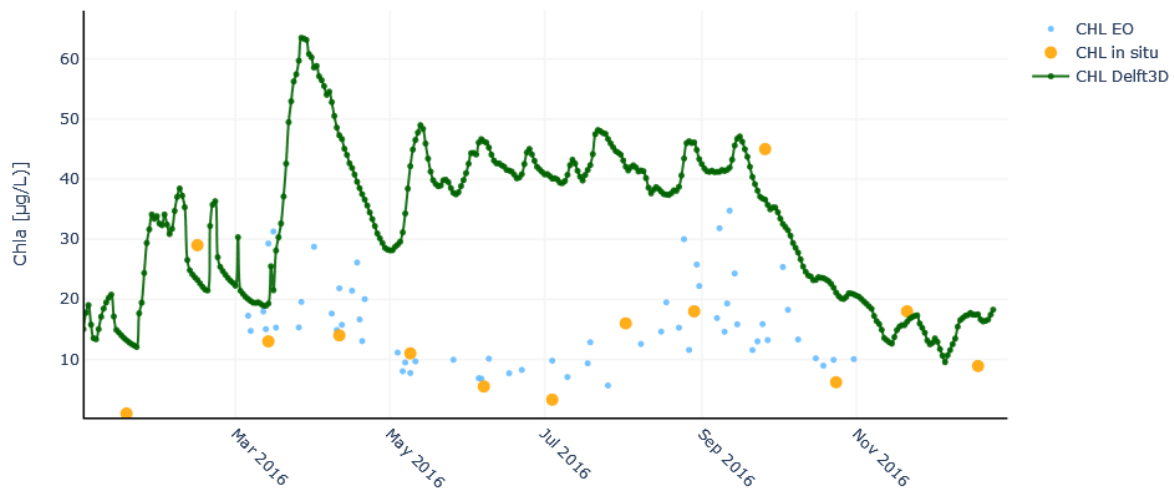


Figure 7 Timeseries of Pampus Oost for chlorophyll-a EO, in situ and Delft3D in 2016.

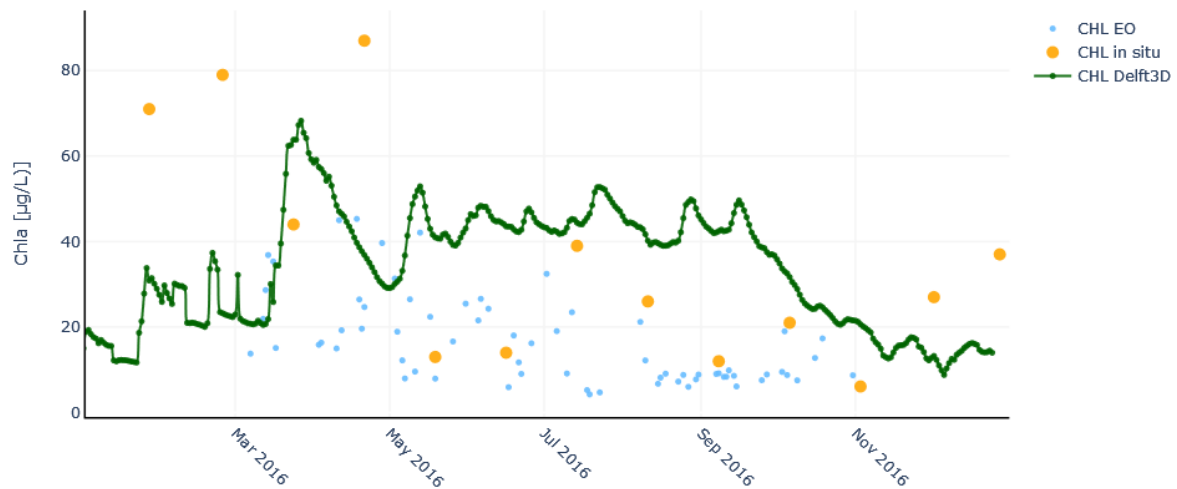


Figure 8 Timeseries of the central station for chlorophyll-a EO, in situ and Delft3D in 2016.

As explained in section 2.1, a potential reason for the mismatch between modelled chlorophyll-a and in situ chlorophyll-a in early spring is because the model only accounts for chlorophyll-a that is contributing to primary production while the observed chlorophyll-a in winter is probably not part of live cells anymore. Algal pigments can still be detectable after cells die which is why they were still measured in situ and by satellites despite not contributing to primary production anymore.

Chlorophyll-a that does not contribute to primary production has only little value for the higher food web. This makes primary production a better indicator for ecosystem functioning than chlorophyll-a. However, fluxes like primary production are difficult to measure, both in situ and remotely.

3.2.4 Spatial Variability

To compare the spatial variability the modelled chlorophyll-a and the EO chlorophyll-a maps were visually inspected. Figure 9 shows two dates comparing both methods. The map of the winter/early spring date of 2016-03-12 shows similar concentration level. The EO chlorophyll-a shows higher concentrations for the Hoornse Hop on 2016-03-12. The increased chlorophyll-a concentrations in this area are possibly the result of resuspension of 'dead chlorophyll-a' (see chapter 3.2.3). Comparing the summer date of 2016-06-05 the concentrations for the modelled chlorophyll-a are higher with a range of 15-20µg/L. Both methods show different spatial patterns. The modelled chlorophyll-a shows increased concentrations in the Gouwzee (western part of Lake Marken) and nearby the Marker Wadden (both are shallow areas). In contrast, in the EO chlorophyll-a map the highest chlorophyll-a levels are found in the center (deepest part) of the Markermeer.

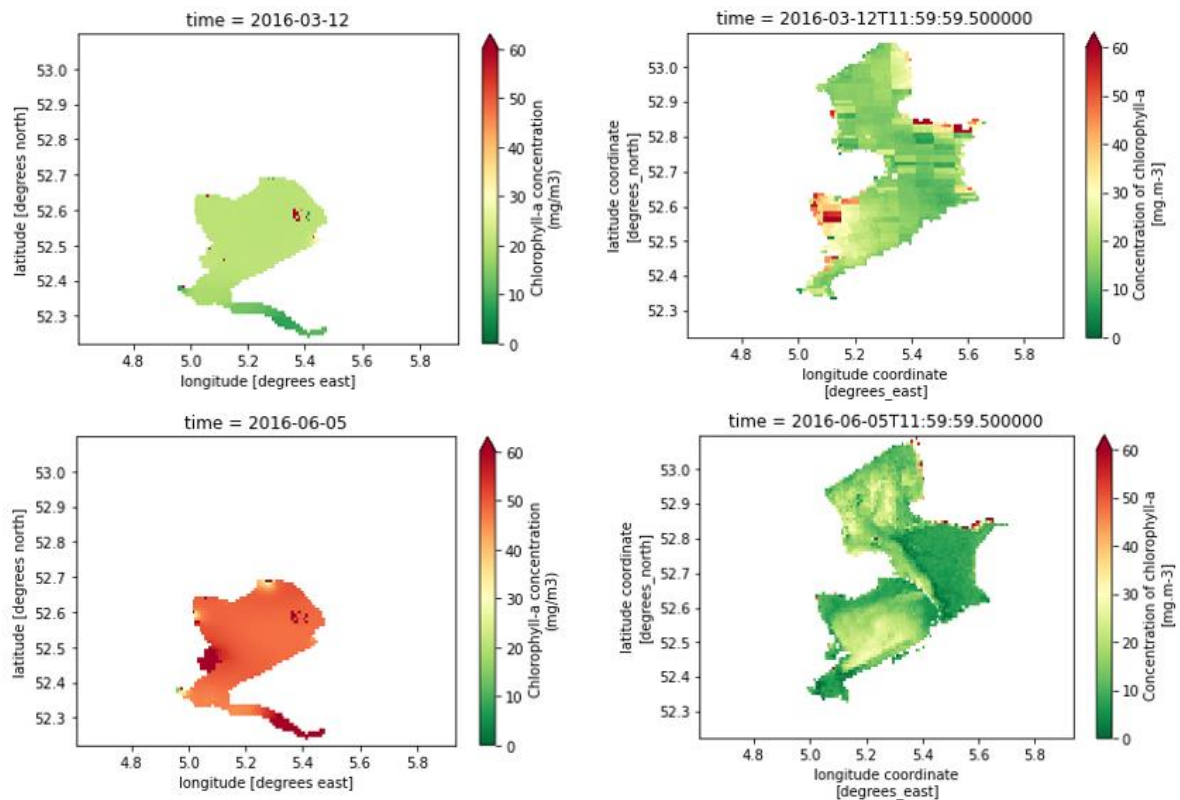


Figure 9 Maps for two dates in 2016 comparing chlorophyll-a for modelled chlorophyll-a (left, Delf3D) and EO chlorophyll-a (right, ESA CCI Lakes).

Figure 10 show the spatial variability of the in situ (left) and EO based (right) temperature used as forcing, and Figure 11 the resulting modelled PP. The shown dates are selected because at these dates the two forcings lead to large differences in modelled PP. The temperature maps make clearly visible that the in situ based forcing is spatially homogeneous, while the EO based forcing does show some spatial variability, with slightly lower temperatures in the centre of the lake (Figure 10). Yet, the PP maps of September 14 (Figure 11) show rather similar patterns, with the same local PP hotspots (e.g. in the Southwest) and slightly increased PP values in the centre of the lake. In contrast, the PP maps of March 10 (Figure 10, upper panels) show a very different pattern, with the EO-based PP values being much higher than those of the in-situ based values. However, these differences are clearly unrelated to the spatial variation in temperatures described above. Instead, they are caused by the absolute difference in the forced temperatures, which allow for an earlier start of the bloom. Furthermore, both model variants show spatial PP gradients along the shores which partly reflect the local depth gradients (PP depends on depth since it is expressed in units of m^2 and is thus integrated over the vertical). However, these gradients are slightly steeper when the temperatures are forced on basis of EO-data, reflecting the local temperature artefacts in those forcings with null temperatures along the shorelines (see Figure 10). EO data at the transition between water and land should always be treated carefully and further masking and filtering might be necessary to generate an optimal EO based input to the model. Although the maps shown suggest that the higher spatial variability in EO based temperatures may not be causing the large differences and peaks in the modelled PP, it may still explain

part of the variability observed in the in situ measurements of PP that is missing in the model based on in situ temperatures (see Figure 2 and 3). Yet, an analysis of the relation between the forced temperature and the modelled primary production shows that the range in measured PP values is larger than can be just explained on basis of the spatial variability in temperatures (Figure 12): for example, with the range of temperatures on September 14th (i.e. between 20 and 25 °C), the model predicts a range of GPPs between 2.5 and 6 gC/m²/d, while the observed GPPs over that same temperature range fall between 1 and 7.5 gC/m²/d. (Note that Figure 12 is based on the year of 2021, since most in situ measurements of PP were available for that year).

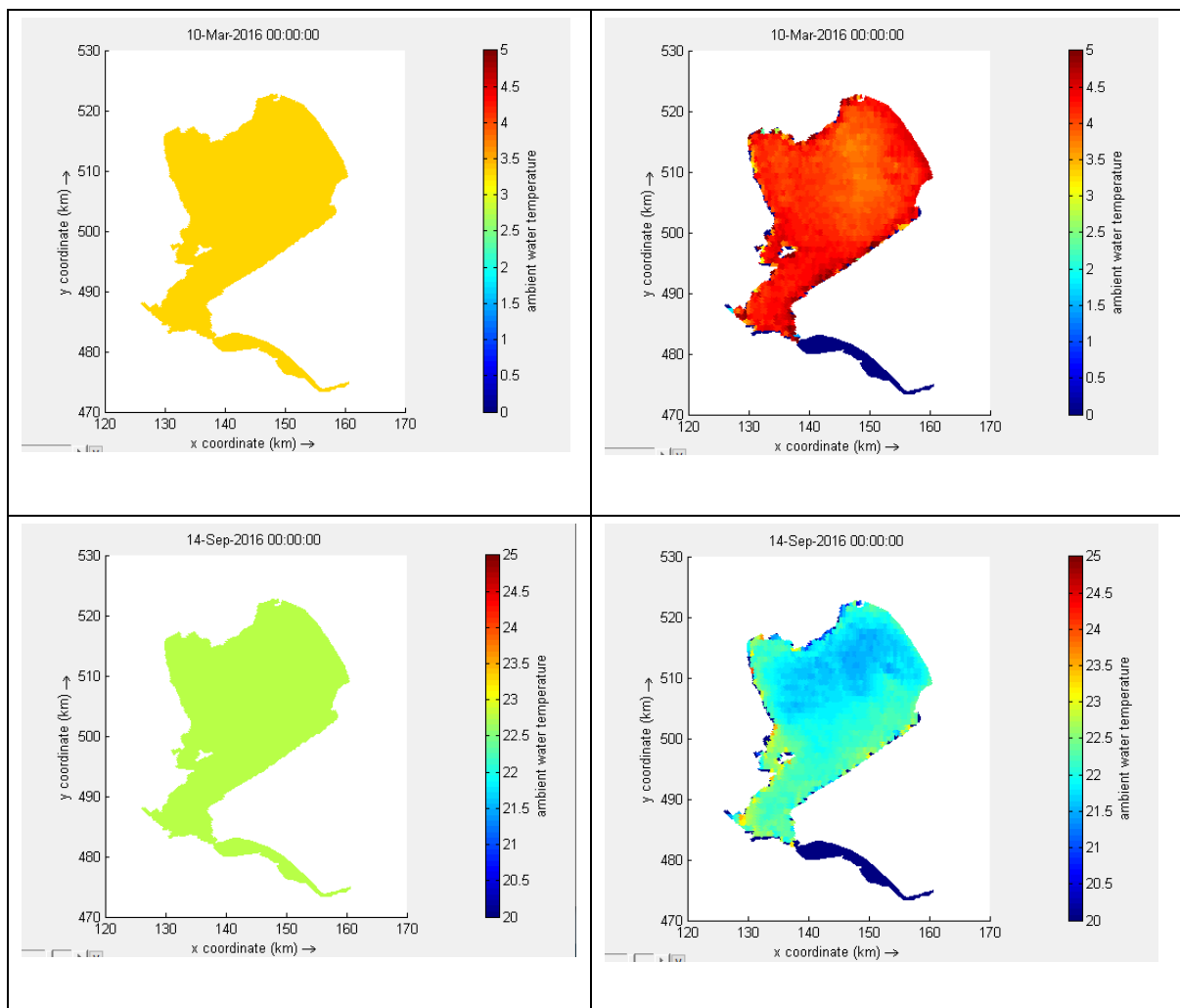


Figure 10 Maps for two dates in 2016 comparing the forcing fields of water temperature on basis of in situ data (left) and on basis of EO data (right). Note that the y-axis differs between the two dates.

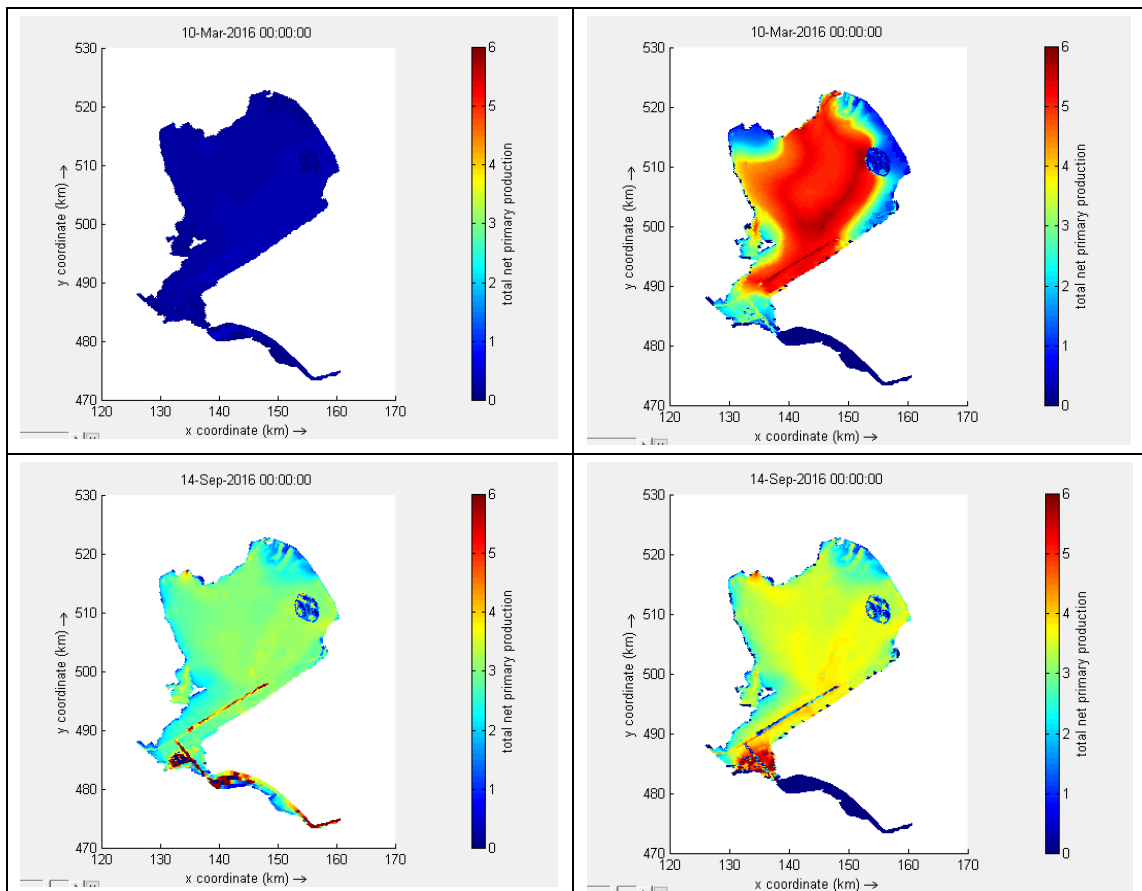


Figure 11 Maps for two dates in 2016 comparing PP modelled on basis of in situ temperature (left) and on basis of EO temperature (right).

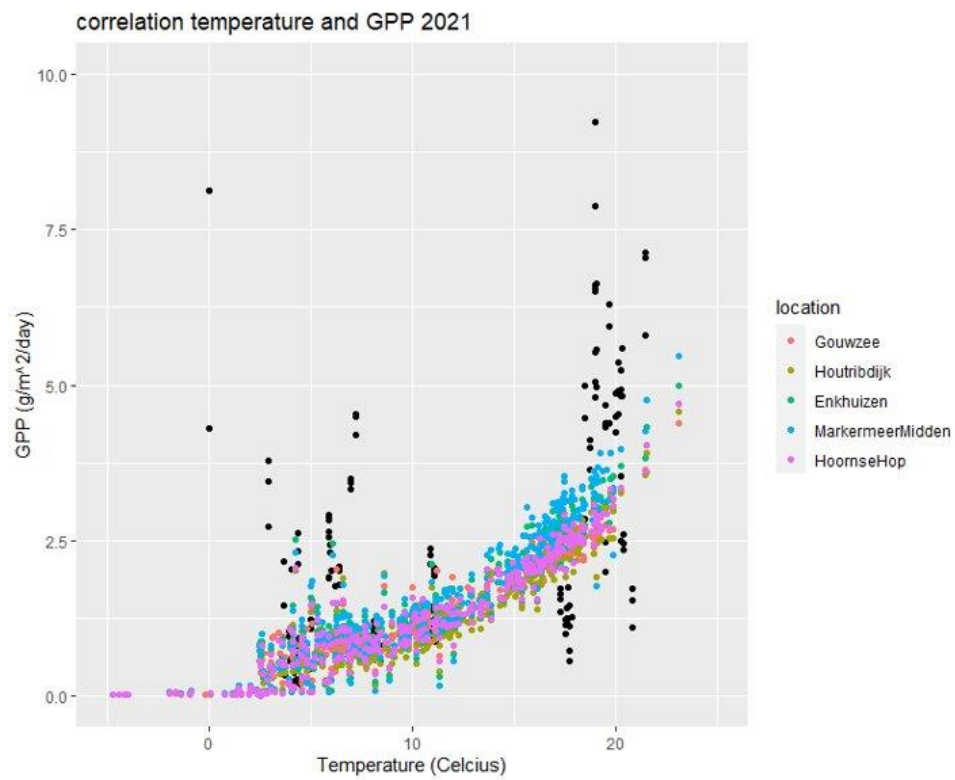


Figure 12 Correlation plot of GPP versus temperature for all areas of the lake in which measurements were taken, showing both the in situ measurements (black dots) as well as the model outputs (coloured dots) in 2021.

4 Pilot 2 – Heat Tolerance of Fish

Lake surface water temperatures have been rising rapidly globally. Additionally, lake heat waves intensity and duration are expected to increase with future climate change, exacerbating the effects of long-term warming. Lake ecosystems are vulnerable to these temperature changes: directly by pushing to or exceeding species and ecosystems limits of resilience, and indirectly through for example decreasing amount of oxygen in the water, altering stratification or algae blooms altering oxygen availability.

The objective of this BIOMONDO pilot is to explore the possibilities of using a combination of EO data on LSWT and thermal tolerance of freshwater fish species to quantify the impacts of increases in temperature and heat waves on freshwater fish diversity.

4.1 Model

Validation of the model results would require observations of fish kills known to be caused by the direct effect of high temperatures (i.e. due to exceedance of the heat tolerance) at the pilot location. However, even if noticed, fish kills registration is not a standard procedure. Additionally, it can be hard to appoint the exact cause of the fish kills. Even when fish kills or change in abundance occurs simultaneously with high temperatures/heat waves, this could also be a result of for example indirect effects as algae blooms and oxygen depletion. Records of fish kills with temperature assigned as the direct cause were not available for the pilot sites. We have received information from news articles, local experts on fish kill occurrences and about what fish species would be more susceptible to heat. We also received abundance data from fish surveys (Lake Marken: WMR; Lake Mälaren: Axenrot, T. (2020)). Although we cannot use this information to truly validate the results, we did compare it with our results (see ATBD document).

4.2 Remote Sensing

Within Pilot 2 daily LSWT data is needed to run the heat tolerance model. Daily LSWT gap filled L4 products are generated using the DINEOF (Data Interpolating Empirical Orthogonal Functions). The data source for the L4 products is the ESA CCI Lakes v2.0.1 LSWT variable (see Experimental Datasets D2.4) with quality level 4 and 5. For the L3 LSWT product a temperature bias of 0.15 to 0.25 °C is expected according to Simis et al. (2020). With the L4 DINEOF products we will not be able to achieve the requirements for the L3 products from the CCI Lakes dataset, therefore we expect a median bias of 1K due to uncertainties and assumptions during the application of the DINEOF interpolation model.

4.2.1 In situ Measurements

For Lake Mälaren, in situ LSWT measurements are available for various stations. Figure 13 shows the locations of all available stations. The data is collected within the Swedish national monitoring program and provided by the Swedish University of Agricultural Sciences (SLU), the national data host for lakes and watercourses (Miljödata-MVM (2022)). Only measurements with a maximum sample depth of 50cm were selected to ensure comparability between both measuring methods.

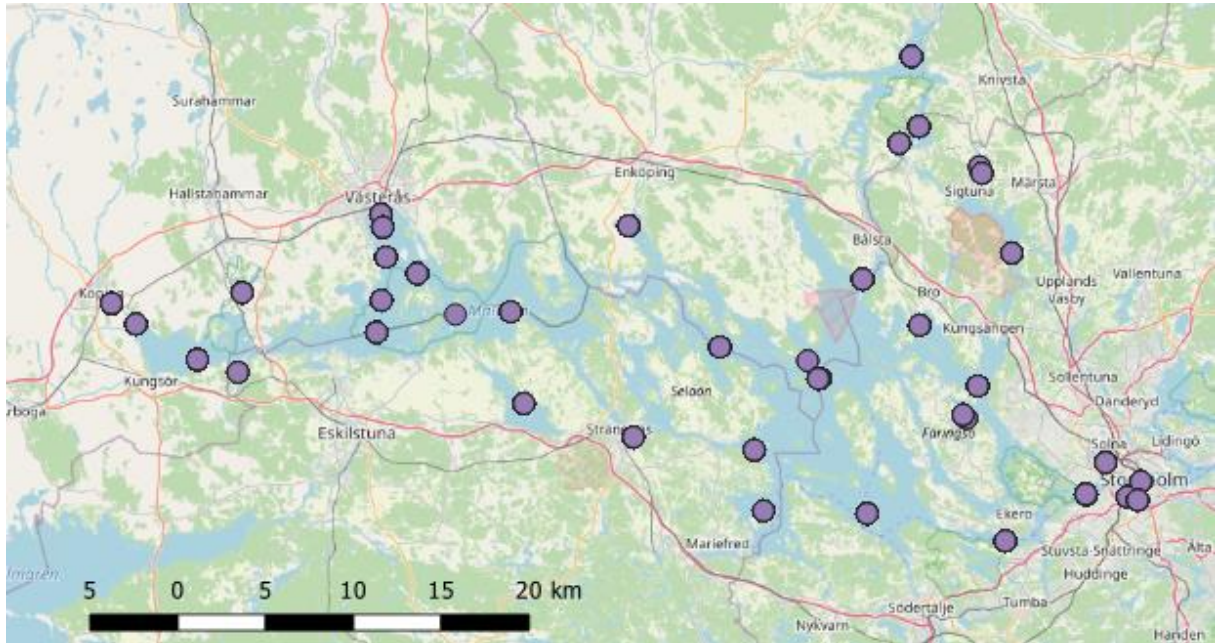


Figure 13 SLU in situ stations for LSWT of Lake Mälaren. Basemap: Openstreetmap.

For Lake Marken daily average LSWT in situ measurements are available from Rijkswaterstaat for the central station of the lake. Only measurements with a maximum sample depth of 100cm were selected to ensure the comparability between both measuring methods.

4.2.2 Matchup Analysis

Scatterplots for Lake Mälaren are shown in Figure 14 and Figure 15 for two central stations (Prästfjärden and Granfj. Djugaruds Udde). The matchups are generated for measurements at the same day. Both stations show a bias $< 0.2^{\circ}\text{C}$ and a RMSE < 1.3 . The scatterplots show that EO LSWT rather underestimated lower temperatures than higher temperatures.

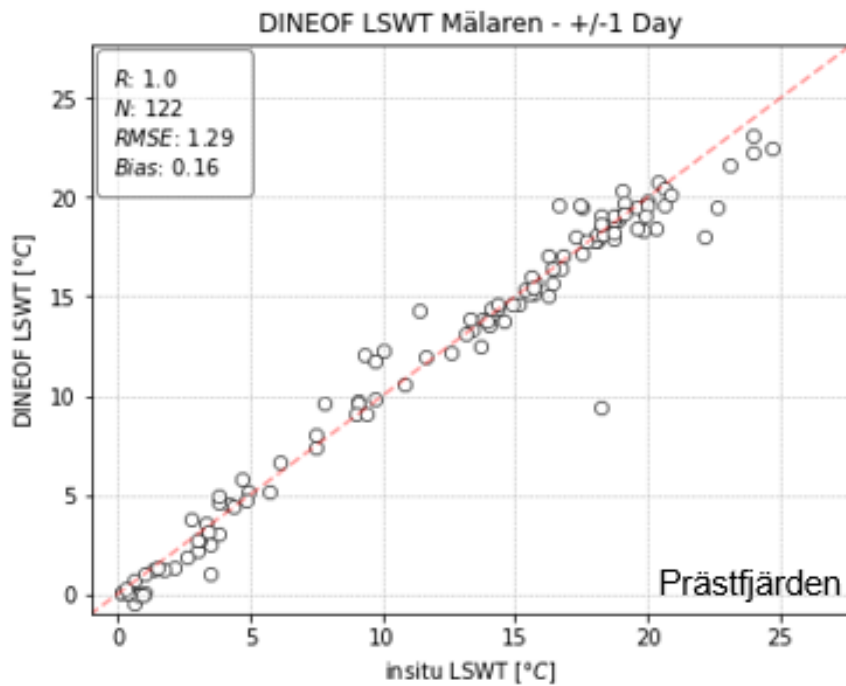


Figure 14 Scatterplot for Lake Mälaren-Prästfjärden of EO DINEOF LSWT and in situ LSWT.

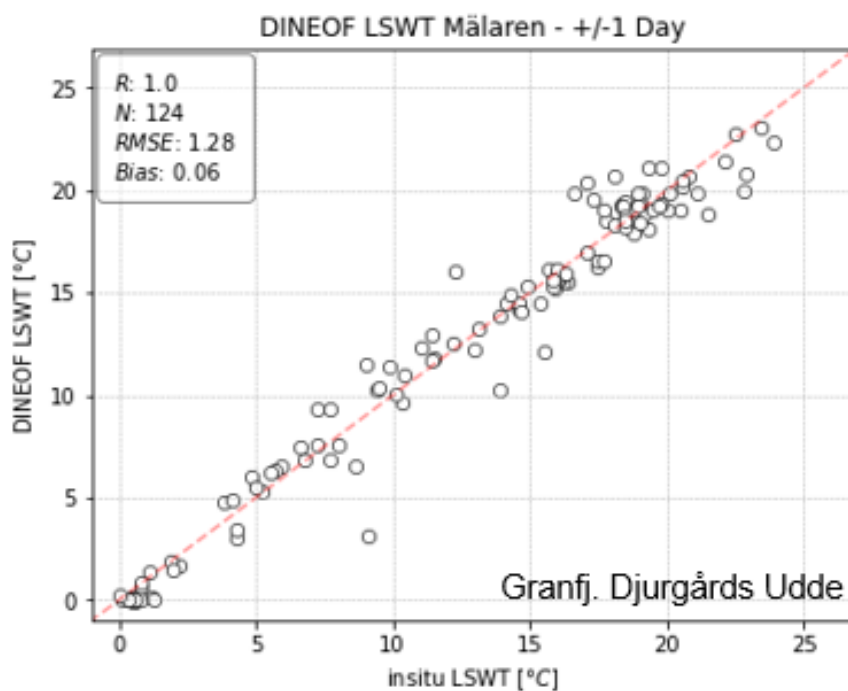


Figure 15 Scatterplot for Lake Mälaren-Granfj. Djurgårds Udde of EO DINEOF LSWT and in situ LSWT.

The scatterplot of Lake Marken is shown for the central station in Figure 16. Due to daily averages of the LSWT in situ measurements a total number of 2357 matchups from the

same day with an average bias of 0.48 and a RMSE of 1.14 are shown. Furthermore, this scatterplot shows the similar behaviour of underestimating lower temperature values.

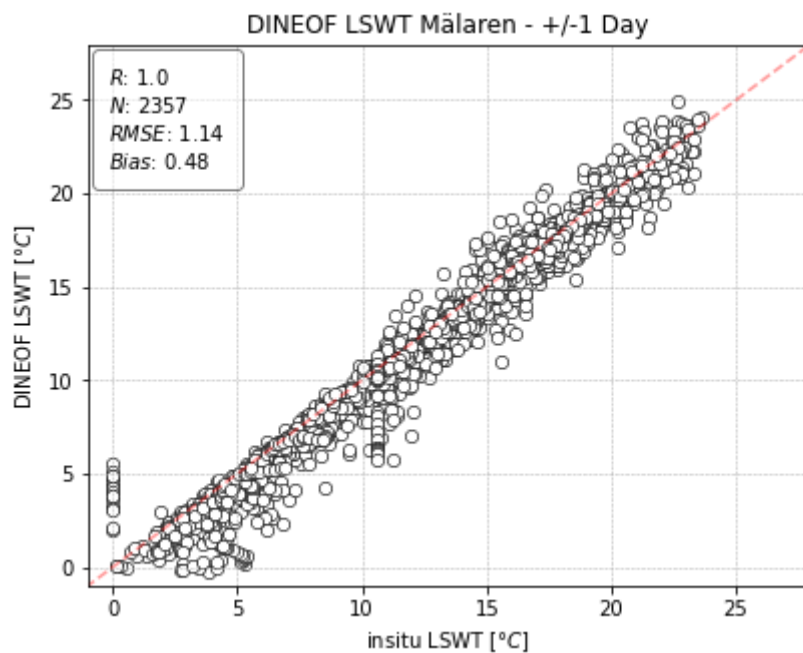


Figure 16 Scatterplot for Lake Marken-Central Station of EO DINEOF LSWT and in situ LSWT.

The validation scatterplots show that the DINEOF interpolation is applicable for the CCI Lakes LSWT dataset and is showing good results with overall biases lower than 1K. The higher uncertainty and underestimation of lower temperatures is neglectable for the Biomondo project because within Pilot 2 heat wave events and the effect on fish species abundances is investigated.

4.2.3 Timeseries Analysis

Timeseries for the central station in Lake Marken are shown in Figure 17 to investigate the coverage of heat waves during the summer month. To showcase the investigation the years 2013 and 2015 were selected. Both years show that both methods correlate over the summer months and that all periods with increased temperatures are covered by the interpolated DINEOF EO LSWT. The investigation of the timeseries support the statement that lower temperatures are underestimated, showing lower peaks for periods with lower temperatures.



Figure 17 Timeseries of 2013 and 2015 of the central station in Lake Marken. DINEOF EO LSWT is shown in blue and in situ LSWT in orange.

5 Pilot 3- River Dams

Obstacles such as dams and other human-made waterworks heavily alter and interrupt dispersal routes for many species, including aquatic invertebrates, and plants. In addition to this, river dams and other human-made waterworks change the natural flow regimes and habitats of aquatic and semi-aquatic species in rivers and river floodplains. Other effects of dams on biota occur via water quality deterioration and reduction of sediment transport to coastal wetlands.

This BIOMONDO pilot explores the possibilities for combining EO data and biodiversity modelling for monitoring and assessing the impact of dam construction and removal on biodiversity, including the effects on Habitat fragmentation and dispersal routes, Changes in habitat extent and water quality (e.g. through influences on sedimentation and turbidity).

5.1 Pareto front optimization and Connectivity Model

The ultimate aim of this pilot is to contribute to a decision framework that helps to weigh the pros (i.e. in terms of energy production) and cons of (different types) individual river dams and their placement within a river basin (e.g. as in Schmitt *et al.* 2018). In order to do this, we estimate the impact of river dams as the probability that the dam brings highest energy benefits for the lowest (i.e. connectivity) impacts we by determining the 'pareto front' (see Schmitt *et al.* 2018).

The impact of individual river dams on the connectivity of the Mekong basin will be determined following a procedure described in Barbarossa *et al.* (2020). This procedure estimates the impact of river dams on the geographic range connectivity of ~10.000 fish species living partially or exclusively in rapidly flowing freshwater (e.g. rivers) for the entire Mekong basin (which can be averaged for all fish species or for certain taxonomic groups). Following the same procedure, changes in the overall connectivity over the past 50 years were already determined for the Mekong basin (see D2.2 ATBD v1.0).

Our optimization procedure can be expanded to include other impacts of river dams such as sedimentation processes and land cover changes as well.

5.2 Remote Sensing

Nechad *et al.* (2009, 2010) provided the theoretical basis and estimation of turbidity (TUR) as a function of reflectance (RRS) at a single band, and provided calibration coefficients for all wavelengths, λ , between 520 nm and 885 nm so that the same basic algorithm can easily be applied to any sensor with a single red or NIR band. Within the Pilot 3 the TUR was calculated from Sentinel-2 based on the 665 nm band and the coefficients provided by Van der Zande *et al.* (2022).

5.2.1 In situ Measurements

The Mekong River Commission (MRC) is an intergovernmental organisation for regional dialogue and cooperation in the Lower Mekong River Basin, established in 1995 based on the Mekong Agreement between Cambodia, Lao PDR, Thailand, and Vietnam. The organisation serves as a regional platform for water diplomacy and a knowledge hub of water resources management for the sustainable development of the region. For validation of EO TUR the parameter Total Suspended Solids (TSS) were extracted from 6 stations in the Mekong and side arms of the Mekong. The stations are distributed in Laos, Vietnam, Cambodia, and Thailand. Table 2 shows the stations used in the validation process and the adjustments performed on the location of the measuring stations.

Table 2 Information about the used stations of the MRC.

Id	Name	River	Latitude	Longitude	Country	Comment
13401	Savannakhet	Mekong	16.561993	104.742904	Lao PDR	Coordinates moved about 50 m offshore
19803	Tan Chau	Mekong	10.803	105.243	Vietnam	Coordinates moved about 50 m offshore
20103	Kg. Chhnang	Tonle Sap	12.260297	105.678530	Cambodia	Coordinates moved into the center of a sidearm
39803	Can Tho	Bassac	10.033317	105.789687	Vietnam	Coordinates moved about 50 m offshore
350101	Ban Keng Done	Se Bang Hieng	16.185	105.317	Lao PDR	Coordinates moved to the center of the pond.
380127	Kaeng Saphu Tai	Nam Mun	15.24	105.248	Thailand	

5.2.2 Timeseries Analysis

For all stations the seasonal pattern is in good agreement between in situ TSS and the EO retrieved TUR. During November – April both data sources show low values. For the rainier season starting in May the concentrations increase until August and decrease after the peak until November. In Figure 18 and Figure 19 two example stations are shown. The station in Laos shows higher deviations in the years 2016-2017.

Deviations can be explained by inaccuracies in locations of the in situ measuring stations and therefore different measuring points. Furthermore, turbid rivers show high level of patchiness in the spatial distribution of TSS, indicating that accurate measuring time and location is crucial for a qualitative comparison.

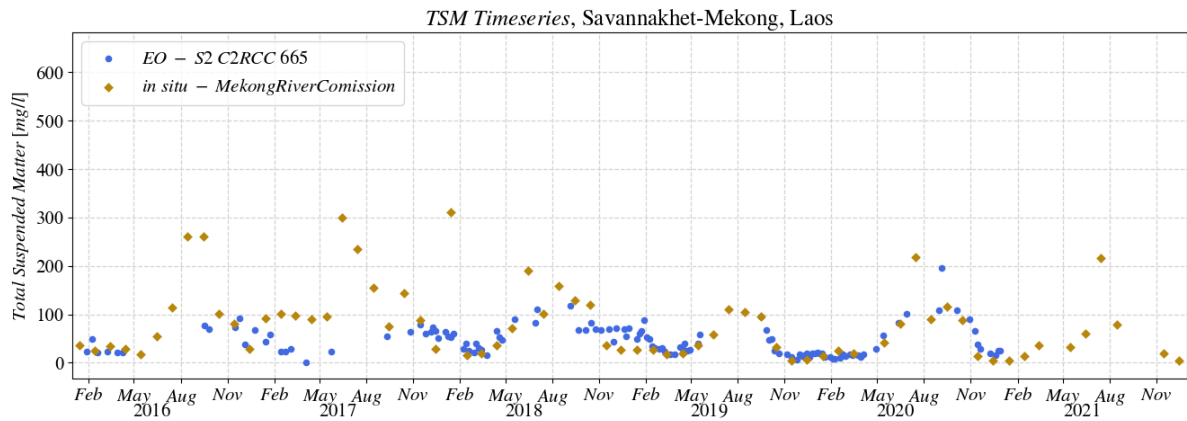


Figure 18 Timeseries of the in situ station Savannakhet in the Mekong, Loas, blue shows the EO retrieved TUR and brown the in situ TSS measurements. Data provided by Mekong River Commission and reproduced with permission.

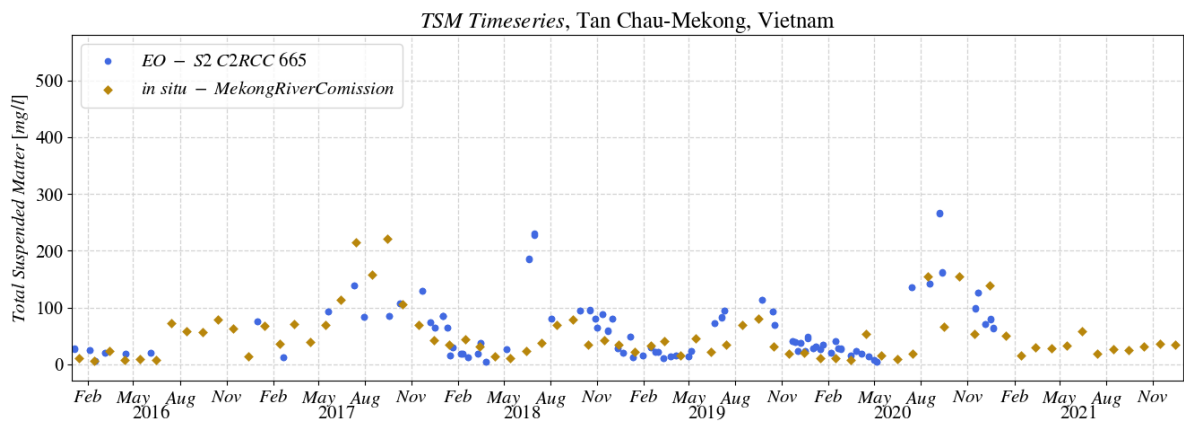


Figure 19 Timeseries of the in situ station Tan Chau in the Mekong, Vietnam, blue shows the EO retrieved TUR and brown the in situ TSS measurements. Data provided by Mekong River Commission and reproduced with permission.

5.2.3 Matchup Analysis

The scatterplot in Figure 20 shows a $R=0.3$ and $RMSE=40.27$ indicating a low correlation between EO TUR and the in situ TSS data. The matchups were extracted for measurements at the same day. The visual inspection shows that the low correlation can be explained with single outliers, showing very high EO TUR concentrations and low in situ TSS concentrations. Furthermore, some matchups show the opposite behaviour. When investigating these deviations, the flagging for the EO TUR data showed no exceptional behaviour and for the in situ data no additional information about the measuring methodology was given. Except for the outliers the scatterplot shows that a correlation between EO TUR and in situ TSS can be found. Especially taking the time series into consideration and evaluation both validation methods together.

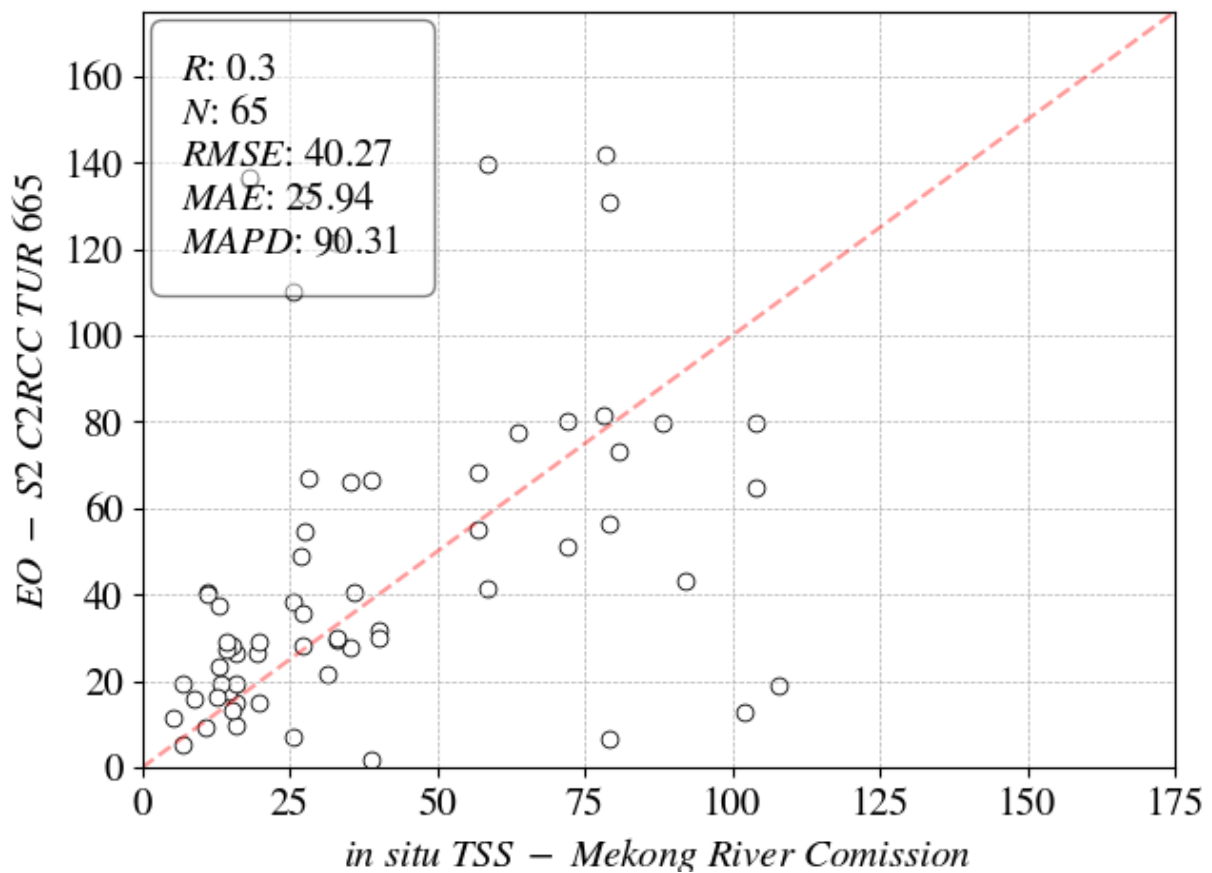


Figure 20 Scatterplot of EO TUR data in situ TSS data provided by Mekong River Commission and reproduced with permission.

5.2.4 Visual Inspection

For visual inspection and validation heatmaps were taken into consideration to analyse the retrieved EO TUR. Figure 21 shows the monthly median EO TUR measurements of the Kg. Chhnang station in Cambodia of the years 2016-2021. Around May and June the TUR is increasing due to monsoon season in the Mekong Basin for all years. The increase of the TUR values is very dominant with values from ~10 FNU to ~90 FNU within a one or two months. Figure 22 shows the monthly median EO TUR measurements of the Savannakhet station in Laos and displays a similar behaviour as the Kg. Chhnang station except that within the monsoon season no EO data is available due to cloud cover. This shows the importance of a combined measuring methodology of in situ and EO data.

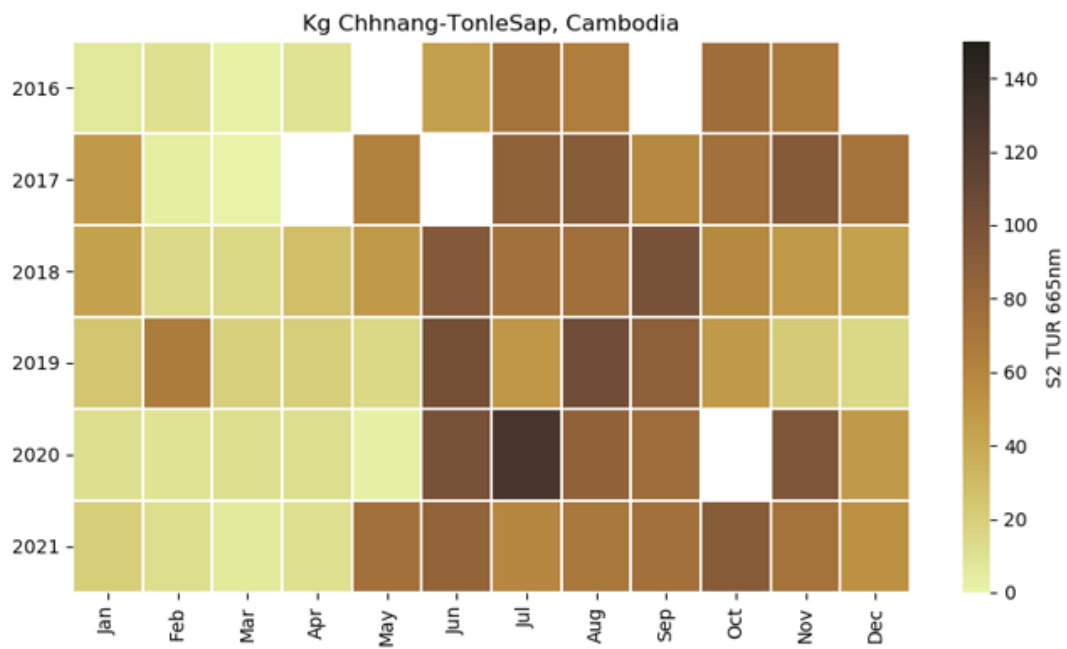


Figure 21 Heatmap for EO TUR of the Kg. Chhnang station in Cambodia.

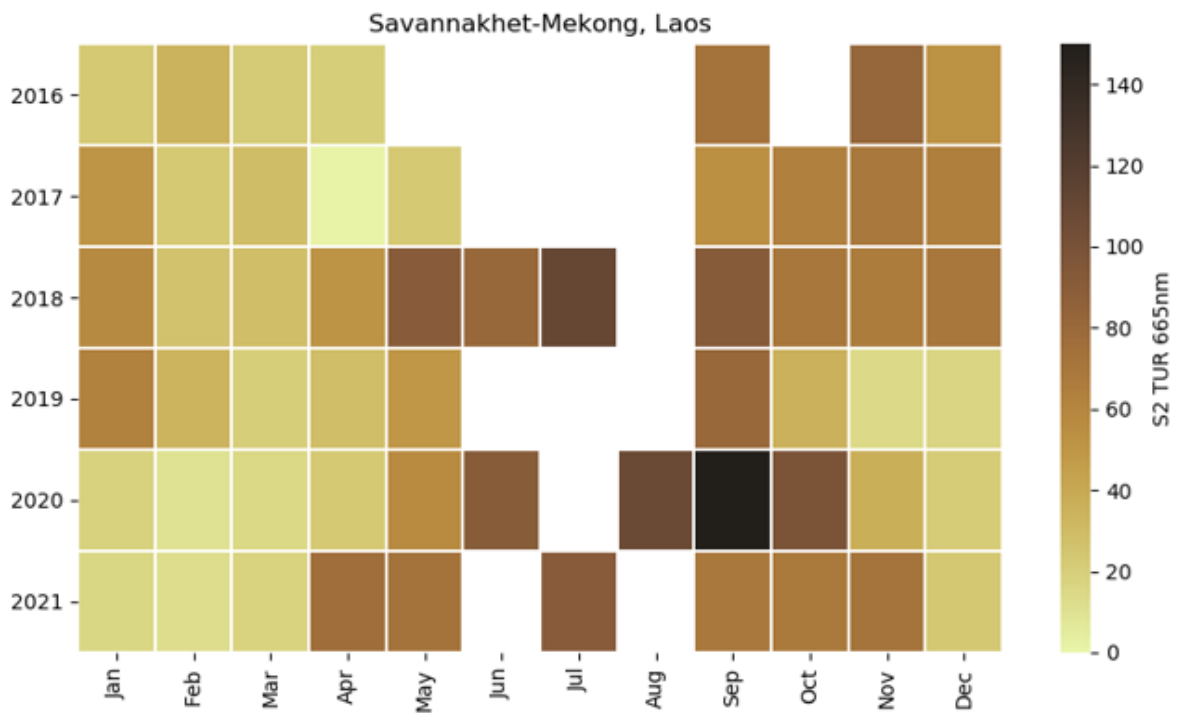


Figure 22 Heatmap for EO TUR of the Savannakhet station in Laos.

6 References

- Axenrot, T. (2020). Hydroakustik i sötvatten. Ett verktyg i fisk- och miljöövervakning. Aqua reports 2020:12. Department of Aquatic Resources (SLU Aqua), Swedish University of Agricultural Sciences. [axenrot t 200924.pdf \(slu.se\)](#)
- Barbarossa, V., Schmitt, R. J., Huijbregts, M. A., Zarfl, C., King, H., & Schipper, A. M. (2020). Impacts of current and future large dams on the geographic range connectivity of freshwater fish worldwide. *Proceedings of the National Academy of Sciences*, 117(7), 3648-3655.
- Ibelings, B.W., M. Vonk, H.F.J. Los, D.T. van der Molen & W.M. Mooij (2003) Fuzzy modeling of cyanobacterial surface waterblooms: validation with NOAA-AVHRR satellite images. *Ecological Applications* 13: 1456-1472.
- Loucks, D.P., van Beek, E. (2017). Water Quality Modeling and Prediction. In: *Water Resource Systems Planning and Management*. Springer, Cham. https://doi.org/10.1007/978-3-319-44234-1_10.
- Miljödata-MVM (2022). Swedish University of Agricultural Sciences (SLU). National data host lakes and watercourses, and national data host agricultural land, <https://miljodata.slu.se/mvm/> (2022-06-15).
- Nechad, B., Ruddick, K. G., & Neukermans, G. (2009, September). Calibration and validation of a generic multisensor algorithm for mapping of turbidity in coastal waters. In *Remote Sensing of the Ocean, Sea Ice, and Large Water Regions 2009* (Vol. 7473, pp. 161-171). SPIE.
- Nechad, B., Ruddick, K. G., & Park, Y. (2010). Calibration and validation of a generic multisensor algorithm for mapping of total suspended matter in turbid waters. *Remote Sensing of Environment*, 114(4), 854-866.
- Schmitt, Rafael JP, et al. "Improved trade-offs of hydropower and sand connectivity by strategic dam planning in the Mekong." *Nature Sustainability* 1.2 (2018): 96-104.
- Simis, S., Liu, X., Yesou, H., Malnes, E., Vickers, H., Blanco, P., Merchant, C., Carrea, L., Duguay, C. (2020). CCI-LAKES-0030-PVP. ESA. https://climate.esa.int/media/documents/CCI-LAKES-0030-PVP_v1.3.pdf
- Van der Zande, D., Stelzer, K., Santos, J., Böttcher, M., Lebreton, C., Vanhellefont, Q., Sterckx, S. (2022). QUID for HR OC Products. CMEMS. <https://catalogue.marine.copernicus.eu/documents/QUID/CMEMS-HR-OC-QUID-009-201to212.pdf>
- WMR Open Data (2021). Toestand vis en visserij in de zoete Rijkswateren – Data 1 december 2021. Wageningen Marine Research – Wageningen UR. Available from: <https://wmropendata.wur.nl/>. Access date: 2022-06-20.