Rapid and agile ocean forecasting with surrogate modelling

Antti Westerlund, Benoit Espinola,
Aleksi Nummelin

Finnish Meteorological Institute

Publications of the Scientific Advisory Board for Defence
2024:5

Publications of the Scientific Advisory Board for Defence 2024:5

Rapid and agile ocean forecasting with surrogate modelling

Antti Westerlund, Benoit Espinola, Aleksi Nummelin

Finnish Meteorological Institute

Ministry of Defence Helsinki 2024

Publication distribution

Institutional Repository for the Government of Finland Valto

julkaisut.valtioneuvosto.fi

Ministry of Defence

ISBN pdf: 978-951-663-186-1
ISSN pdf: 2984-102X

Layout: Government Administration Department, Publications

Helsinki 2024 Finland

Description sheet
8 April 2024

Rapid and agile ocean forecasting with surrogate modelling

Publications of the Scientific Advisory Board for Defence 2024:5
Publisher Ministry of Defence

Author(s) Antti Westerlund, Benoit Espinola, Aleksi Nummelin
Group author Finnish Meteorological Institute
Language English **Pages** 22

Abstract

Marine safety and security, including military situational awareness, create a need for short-term ocean forecasts of the ocean state. At present, such forecasts are done using numerical models on high performance computers. These models, although robust, often provide information at relatively coarse resolution for local/regional application, are difficult to deploy rapidly, and require connection between the data centre and user in the field. However, the rapid development of data-driven methods has opened a possibility for statistical emulators that can be trained with numerical model data to produce similar information but can be evaluated on a laptop computer in the field in a matter of seconds. To this end, we demonstrate creating an emulator to predict thermocline depth, an important parameter for sonar weather, up to 10 days in advance in the Archipelago Sea, Baltic Sea. We find that for this purpose, multiple linear regression produces the best results as more complex architectures suffered from limited training data or limited memory in the training phase.

Provision This publication is part of the implementation of research funding of the Scientific Advisory Board for Defence (MATINE). (www.defmin.fi/matine) The content is the responsibility of the producers of the information and does not necessarily represent the view of the Defence Ministry.

Keywords national defence, research, comprehensive defence approach, modelling

ISBN PDF 978-951-663-186-1 **ISSN PDF** 2984-102X

URN address https://urn.fi/URN:ISBN:978-951-663-186-1

Kuvailulehti
8.4.2024

Nopeita ja ketteriä meriennusteita surrogaattimalleilla

Maanpuolustuksen tieteellisen neuvottelukunnan julkaisuja 2024:5
Julkaisija Puolustusministeriö

Tekijä/t Antti Westerlund, Benoit Espinola, Aleksi Nummelin
Yhteisötekijä Ilmatieteen laitos
Kieli Englanti **Sivumäärä** 22

Tiivistelmä

Meriturvallisuus ja merenkulun turvaaminen, mukaan lukien sotilaallinen tilannetietoisuus, luovat tarpeen lyhyen aikavälin meren tilan ennusteille. Tällä hetkellä tällaiset ennusteet tehdään käyttämällä numeerisia malleja suuritehotietokoneilla. Vaikka nämä mallit ovat sinänsä luotettavia, ne tuottavat usein tietoa suhteellisen karkealla resoluutiolla paikallista/alueellista soveltamista varten, niitä on vaikea ottaa nopeasti käyttöön ja ne edellyttävät yhteyttä tietokeskuksen ja kentällä olevan käyttäjän välillä. Datapohjaisten menetelmien nopea kehitys on kuitenkin avannut mahdollisuuden tilastollisiin emulaattoreihin, jotka voidaan kouluttaa numeeristen mallien tietojen avulla tuottamaan samanlaista tietoa kuin itse mallit, mutta joita voidaan ajaa kentällä kannettavalla tietokoneella muutamassa sekunnissa. Tässä työssä kehitettiin merimalliemulaattori, jolla voidaan ennustaa termokliinin syvyyttä, joka on tärkeä parametri kaikuluotaussäälle, jopa 10 päivää etukäteen Saaristomerellä. Moninkertainen lineaarinen regressio tuotti parhaat tulokset, sillä monimutkaisemmat menetelmät kärsivät rajallisesta koulutusdatasta tai rajallisesta muistista koulutusvaiheessa.

Klausuuli Tämä julkaisu on toteutettu osana Maanpuolustuksen tieteellisen neuvottelukunnan (MATINEn) tutkimusrahoituksen toimeenpanoa. (www.defmin.fi/matine) Julkaisun sisällöstä vastaavat tiedon tuottajat, eikä tekstisisältö välttämättä edusta puolustusministeriön näkemystä.

Asiasanat maanpuolustus, tutkimus, kokonaismaanpuolustus, mallintaminen

ISBN PDF 978-951-663-186-1 **ISSN PDF** 2984-102X

Julkaisun osoite https://urn.fi/URN:ISBN:978-951-663-186-1

Presentationsblad
8.4.2024

Snabba och smidiga havsprognoser med surrogatmodellering

Publikationer av försvarets vetenskapliga delegation 2024:5

Utgivare	Försvarsministeriet		
Författare	Antti Westerlund, Benoit Espinola, Aleksi Nummelin		
Utarbetad av	Meteorologiska institutet		
Språk	Engelska	**Sidantal**	22

Referat

Maritim säkerhet och trygghet, inklusive militär situationsmedvetenhet, skapar ett behov av kortsiktiga prognoser för havets tillstånd. För närvarande görs sådana prognoser med hjälp av numeriska modeller på högpresterande datorer. Dessa modeller är visserligen robusta, men ger ofta information med relativt grov upplösning för lokala/regionala tillämpningar, är svåra att använda snabbt och begär netvärkanslutning mellan datacentret och användare på fältet. Den snabba utvecklingen av datadrivna metoder har dock öppnat en möjlighet för statistiska emulatorer som kan tränas med numeriska modelldata för att ge liknande information, och som kan utvärderas på en bärbar dator på fältet inom några sekunder. För detta ändamål demonstrerar vi hur man skapar en emulator för att förutsäga termoklindjuphet, en viktig parameter för sonarväder, upp till 10 dagar i förväg i Skärgårdshavet, Östersjön. Vi konstaterade att multipel linjär regression ger det bästa resultatet för detta ändamål, eftersom mer komplexa arkitekturer lider av begränsad träningsdata eller begränsat minne i träningsfasen.

Klausul Den här publikation är en del i genomförandet av forskningsfinansiering av Försvarets vetenskapliga delegation. (www.defmin.fi/matine) De som producerar informationen ansvarar för innehållet i publikationen. Textinnehållet återspeglar inte nödvändigtvis statsrådets ståndpunkt.

Nyckelord försvaret, forskning, totalförsvaret, modellering

ISBN PDF	978-951-663-186-1	**ISSN PDF**	2984-102X
URN-adress	https://urn.fi/URN:ISBN:978-951-663-186-1		

Contents

1 Introduction 7

2 Research objectives 9

3 Materials and methods 10
3.1 Numerical model and training dataset 10
3.2 Thermocline calculation from NEMO 10
3.3 Machine learning model 11
3.4 Training data and feature engineering 12

4 Results 14
4.1 Pilot tool 14
4.2 Thermocline depth predictions 15

5 Discussion 17
5.1 Time and resource needs 17
5.2 Comparison and evaluation of approaches 17
5.3 Suitability for use 18

6 Conclusions 20

7 Scientific publishing and other reports produced by the research project 21

References 22

1 Introduction

Operational and security applications of marine environmental data are numerous (e.g., Burnett et al., 2014). General applications can include support for naval operational activities, maritime accident management, and oil spill management. Also, awareness of the marine physical conditions is important for managing coastal hazards, such as seawater flooding and nuclear power plant safety.

However, especially on short timescales, the physical conditions in the sea are not well described by observations which lack resolution both in space and time. Numerical models, based on equations describing the physical system, can provide a more comprehensive overview of prevailing conditions, with a higher spatial and temporal resolution and range. These models are well established and have been used for decades. However, their use requires expert input and careful tuning of parametrizations to run and produce results. These simulations are computationally expensive, and the results come after long runtimes (order of hours). The Archipelago Sea in the Baltic Sea is an example of an especially challenging region for these numerical models because of its fragmented coast, where high spatial resolution is required. Overall, all these limitations can make numerical models challenging to use in a rapid operational response context.

In the project “Nopeita ja ketteriä meriennusteita surrogaattimalleilla” (“Rapid and agile ocean forecasting with surrogate modelling”) a completely new type of pilot marine forecasting system was developed based on a surrogate model predicting the thermocline depth for the sea areas surrounding the island of Utö in the Archipelago Sea. A surrogate model is a data-driven approach, learning relationships between the numerical model’s output data and its forcing. Based on the relationships between the forcing and numerical model’s output, the surrogate model then predicts an outcome in a way similar to the traditional numerical model. The training of the surrogate model requires the use of high-performance computers; however, once trained, these surrogate models do not require intensive computational resources and can yield results in a very short time (order of seconds).

The island of Utö is located on the southern edge of the Archipelago Sea. The Archipelago Sea is a sub-region of the Baltic Sea located east from the Åland islands and surrounded by the Baltic proper (in the Southern boundary), the Åland Sea (in the Eastern boundary), the Bothnian Sea (in the Northern boundary), and the Finnish mainland (in the Western boundary). The Archipelago Sea is characterized by a highly fragmented coastline with thousands of islands, shallow underwater topography, with several deeper, mostly north-south oriented, canyons/trenches. The complex topography together with small characteristic scales of dynamics, makes the Archipelago Sea a challenging physical system to simulate.

The Archipelago Sea temperature profile is affected by a seasonal thermocline. After the winter period, when the water column is generally mixed (and of constant temperature) all the way to the bottom (or to the halocline where that exists), the incoming solar radiation increases, heating up the surface layer of the sea. The warming surface layer is then separated from the colder waters below by thermocline, a region of an enhanced vertical temperature gradient. Physical processes such as wind, waves, and surface cooling, work against the solar radiation to weaken the stratification and deepen the mixed layer, and eventually in the fall they break the thermal stratification, and the thermocline is mixed away. Unlike the deeper parts of the Baltic Sea, the Archipelago Sea is shallow enough to be mostly above the region of more persistent saline stratification, halocline, and, consequently, thermal stratification dominates the density, and speed of sound, in the water column. Therefore, depth of the thermocline is an essential variable for sonar weather and sonar calibration.

2 Research objectives

In the project, machine learning methods were used to create a surrogate model to predict the depth of the thermocline in the next 7 days. We aimed to create a model that:

- could be trained in a reasonable time on a HPC cluster,
- could be run on a laptop,
- produced results in 10 minutes or less.

Our surrogate model was a pilot model for the Underwater Warfare Data Center of the Finnish Naval Academy. Our objectives were to produce a product that can be useful for the customer with a good enough quality to support customer operations. Our priority was to predict the thermocline depth in the region surrounding Utö during the open water season with 1 to 4 days lead time. Through this work, we also gathered information on the feasibility of this approach, as well as the most suitable methods for this task.

3 Materials and methods

3.1 Numerical model and training dataset

The machine learning model was trained with data from a traditional hydrodynamic model. The training set consists of results from the NEMO V4.0 hydrodynamic model for the Archipelago Sea and the Åland Sea (Westerlund et al. 2022, Miettunen et al. 2024) for the years 2013 to 2017 with a 0.25 NM horizontal resolution (~500 m), up to 1 m vertical resolution, 200 vertical levels, z* vertical coordinates. The bathymetry used comes from VELMU and BSBD (same configuration as in Westerlund et al. 2022). We used Copernicus Climate Change Service's ERA5 atmospheric reanalysis for the meteorological forcing (Hersbach et al., 2023) and a regional reanalysis configuration provided by the Copernicus Marine Environment Monitoring Service (CMEMS) for the lateral boundary and initial conditions (CMEMS, 2022).

3.2 Thermocline calculation from NEMO

To calculate the depth of the thermocline, we used the temperature ($\boldsymbol{T} \in \mathbb{R}^4$) output from NEMO, with coordinates (t, z, y, x) where t is the time, z is the depth, y is the longitude, and x is the latitude.

We then define the thermocline depth as the depth of the maximum gradient of the temperature profile along the z-axis, complemented with the stratification index and thermocline strength as suggested by Fiedler (2010). Additionally, we constrained the thermocline to $z \in [2, z_{bathy}]$, where z_{bathy} is the depth of the bathymetry at the coordinate (x, y). When the thermocline depth is the same as the depth of the local bathymetry, the thermocline depth is set to zero.

3.3 Machine learning model

For our surrogate model architecture, we used multiple linear regression (MLR) which is a generalization of linear regression for higher dimensions (Marill, 2014). MLR consists of a linear combination of the independent variables (also known as features) with coefficients (also known as weights) vector. Additionally, a constant term known as the intercept is added to this linear combination. When fitting a MLR to data, an error term is added. This error term is a distance metric (also known as cost) between the dependent variable (also known as the response variable), in our case the depth of the thermocline, and the prediction given by the error-free MLR. This error is aimed to be minimized when fitting the training data.

$$\boldsymbol{y} = \beta_0 + \boldsymbol{\beta X} + \boldsymbol{\varepsilon}$$

Where y is the dependent variable (thermocline depth), β_0 is the intercept, $\boldsymbol{\beta X}$ is the linear combination between the coefficients vector ($\boldsymbol{\beta}$) and the independent variables ($\boldsymbol{X}$), and $\boldsymbol{\varepsilon}$ is the model error term.

Changes in thermocline depth are non-linear and depend on non-linear processes such as wind shear and stress (e.g., Krauss, 1981). Furthermore, when the thermocline reaches the bottom, the thermocline is no longer present and its depth is set to zero, which is also a non-linear change. The fragmented islands and coastline in our domain further introduce discontinuity, this time spatially induced.

Discontinuities represent challenges for linear machine learning algorithms such as the MLR. We partially address these challenges by using a custom loss function and through feature engineering.

To tackle the discontinuity introduced by the thermocline reaching the bottom, we developed a custom error term where the distance is measured as follows:

$z_a = \min(z_T, z_P)$ $\quad z_b = \max(z_T, z_P)$ $\quad z_a \leq z_b \therefore z_b - z_a \geq 0$

$\varepsilon_1 = z_b - z_a$ $\quad \varepsilon_2 = z_{bathy} - (z_b - z_a)$

$$\varepsilon = \min(\varepsilon_1 , \varepsilon_2)$$

Where z_T is the depth of the thermocline in the training set (true value), z_P is the depth of the thermocline predicted by the MLR and z_{bathy} is the depth of the bathymetry. ε_1 represents the Euclidian distance between the prediction and the true value of the thermocline depth and ε_2 is a distance metric between the deepest value and the bathymetry combined with the depth of the shallow most point. The approach used for ε_2 considers that points near the bathymetry are close to points at the surface. The error ε becomes the minimum error between these two components.

For the loss function, we used the mean absolute error:

$$MAE = \frac{1}{N}\sum_{i=0}^{N}|\varepsilon|$$

To train the model, we aim at finding the weight vector and intercept that minimizes the MAE.

3.4 Training data and feature engineering

The training data contains a total of 127 variables including ERA5 forcing key variables, the thermocline depth at the previous time step (as initial conditions) and engineered features. The following variables come from the ERA 5 forcing: 10 m U and V wind components (two variables), 2 m dewpoint temperature, 2 m air temperature, mean sea level pressure, mean snowfall rate, mean surface downward short and long-wave radiation fluxes (two variables), mean total precipitation rate, surface air pressure, and total precipitation.

We created the following variables to attempt to linearize the relationship between the raw data and the depth of the thermocline: squared wind speed at

10 m, fourth root of the downward short and long-wave radiation fluxes (two variables), and temperature at 2 m over the squared wind speed at 10 m. We introduced a rolling mean, median, standard deviation, sum, minimum and maximum along the time-axis for the past 6h for the following variables: 10 m U and V wind components (two variables), 2 m dewpoint temperature, 2 m air temperature, mean sea level pressure, mean snowfall rate, mean surface downward short and long-wave radiation fluxes (two variables), mean total precipitation rate, and surface air pressure, squared wind speed at 10 m, fourth root of the downward short and long-wave radiation fluxes (two variables), and temperature at 2 m over the squared wind speed at 10 m. Additionally, we introduced rolling means along the time-axis for the 2 m air temperature including the past 24h and 7 days (a total of two rolling means).

Furthermore, we introduced 3 dummy variables for winter (air temperature ≤ 0 ºC), transition (0 ºC < air temperature < 10 ºC) and summer seasons (air temperature ≥ 10 ºC). Another 4 season variables were introduced where winter is represented by December, January and February; spring is March, April and May; summer is June, July and August; and autumn is September, October and November.

For time, we introduced 4 variables: sine and cosine of the time of the day and sine and cosine of the day of the year:

$$c_{time} = \cos\left(\frac{t_{day}}{24\ h}\right),\ s_{time} = \sin\left(\frac{t_{day}}{24\ h}\right)$$

$$c_{date} = \cos\left(\frac{t_{year}}{days\ in\ year}\right),\ s_{date} = \sin\left(\frac{t_{year}}{days\ in\ year}\right)$$

Finally, we included 6 h, 1 day, 1 day 18 h and 2.5 days lags for the following variables: 2 m air temperature, the downward short and long-wave radiation fluxes, and the thermocline (from the initial conditions).

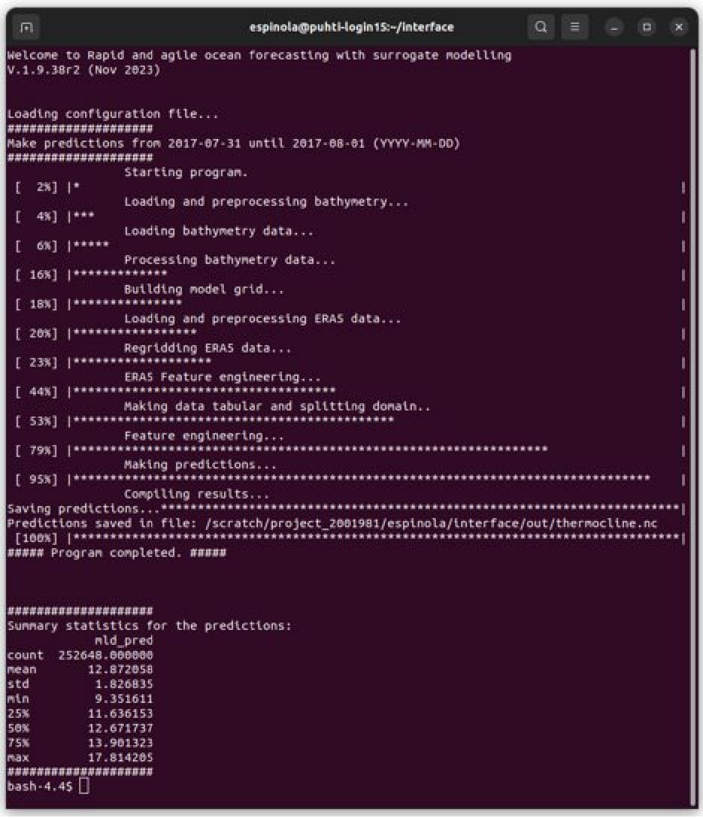
The original ERA5 data has a time resolution of 1h. The data has been regridded to match output grid of the NEMO data (with 6h time resolution).

4 Results

4.1 Pilot tool

We have developed a pilot tool using Python 3.10.13 that uses the command line as an interface together with a configuration text file. The tool generates a NetCDF file containing the predictions obtained by using our model.

Kuvio 1. Screenshot of the pilot tool's command line interface



The tool, presented in Figure 1, enabled us to understand the requirements of implementing this new kind of pilot modelling system, including the resource requirements of the tool. The runtime for training our model is less than an hour on a high-performance computing centre and the results come in less than 5 minutes using a laptop. Our model's computational requirements are a fraction of the needs required by traditional hydrodynamic models.

4.2 Thermocline depth predictions

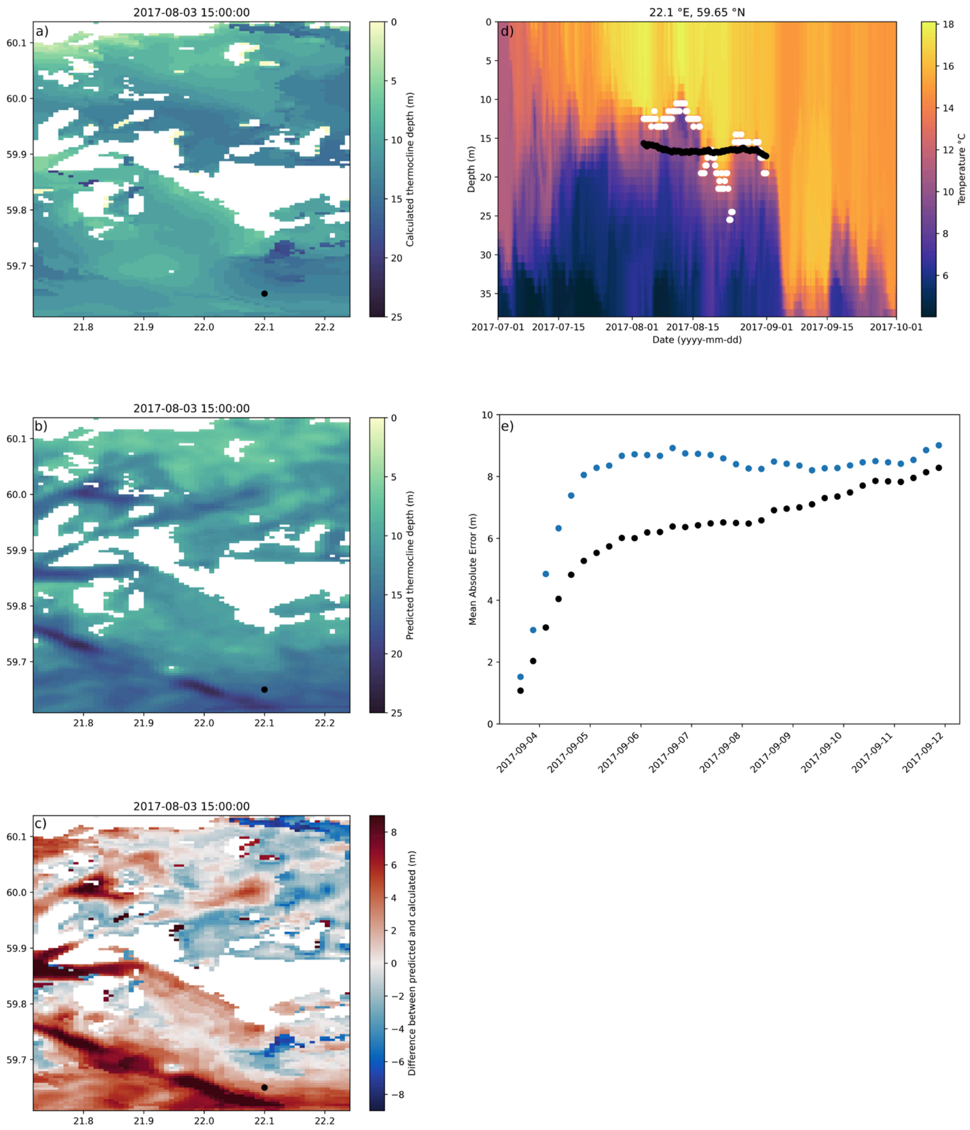
To illustrate the prediction capacity of our model, we plotted maps and time series of our thermocline depth predictions alongside with errors and the thermocline depth coming from the numerical model (see Figure 2). The thermocline depth calculated with NEMO (a) shows variability in the field for 03 August 2017. Our model prediction for the same date (b), captures the general variability.

However, the difference between the two plots (c) suggests a positive bias in the South-West area of our domain and a negative bias in the North-East area of our domain. Extreme values are likely due to the discontinuity of the thermocline depth between the surface and the bottom. The model's mean absolute error (MAE) with respect to lead time (e) suggests the model has degrading skill over lead time with a plateau after 2 to 3 days. The difference between the MAE and the custom MAE

occurs because the custom MAE considers points near the surface close to points near the bottom, tackling the discontinuity of the thermocline depth introduced by the thermocline reaching the bottom.

The time series (d) shows the temperature depth profile calculated from NEMO with the calculated thermocline depth (white points) and our predictions (black points). Here we can see that locally, our model captures large scale variability but not so well the short-term variability.

Kuvio 2. Thermocline depth calculated by the hydrodynamic model NEMO (a), thermocline depth predicted by the multiple linear regression (MLR) (b), and difference between the MLR and NEMO thermocline depth (c). In panel (c), warm colors suggest the MLR overestimated the thermocline depth while cold colors suggest an underestimation. The black dot indicated in panels (a-c) at 59.65 °N, 22.1 °E represents the location of an example station used in panel (d). Panel (d): temperature profile time series (background) with the thermocline depth calculated from NEMO (white) and the thermocline depth predicted by the multiple linear regression (black) at 59.65 °N, 22.1 °E. Panel (e): Mean Absolute Error (blue) and mean absolute custom error (black) for the entire domain with respect to lead prediction time.



5 Discussion

Our model manages to capture large scale features. It has limitations for short variability both in the temporal and spatial domain.

5.1 Time and resource needs

In this work we created an emulator for a host numerical model with a horizontal domain of size 30 km by 90 km, a horizontal resolution of approximately 500 m, and with varying bottom depth such that the total number of grid cells in the 2D domain is 7710 cells. With the given domain, and the multiple linear regression architecture, the emulator training took less than an hour on the CSC hosted Puhti HPC (with 1 core and 185 GiB memory). We note that changing the architecture can significantly impact the training time, but on the other hand, because training only needs to be done once, even a much longer training time, would still be feasible. A weeklong prediction with the trained emulator then takes about 5 minutes on a regular laptop. The evaluation time has only minor dependence on the emulator architecture. Given these results, we conclude that creating emulators for the marine realm is computationally feasible in with the current HPC resources.

5.2 Comparison and evaluation of approaches

On the top of MLR, we have experimented with other ML approaches. We attempted to train the model using a Random Forest (RF) method, widely used because of its interpretability and performance. However, not only training a RF model requires long runs using an HPC, but the method also itself requires the entire dataset to be loaded in the memory. We have reached the limits of the available resources and had to downscale our dataset. This resulted in poor model performance compared to MLR. To obtain better results we would need a larger dataset, which in turn requires more HPC memory. This approach was therefore considered not to be cost effective and was excluded in this study.

We also attempted to train a Conv 3D U-Net, which is a deep neural network combining 3D convolutional layers with a U-Net architecture, which potentially performs well finding geospatial patterns in the data. However, this architecture is difficult to parametrize and the data we have available was insufficient for us to find a good fit. Additionally, the explainability of deep neural networks is challenging, making the MLR approach more appealing.

In future studies, also the following approaches could be tested: Long-Short Term Memory neural networks (LSTM), which performs well for time series data; LSTM U-Net, for the same reasons with an added spatial component; multi-step methods where we first classify our domain (e.g., using K-Nearest Neighbours or Self Organized Maps) to then fit a model for each class; and an ensemble approach, using a collection of ML models to make predictions and merge these predictions for our final output.

Conclusions for the next versions.

- More training material would be beneficial.
- Tuning of the machine learning architecture takes significant amount of time.
- Memory is an essential resource for training the models.
- The initial state of the system is important.

5.3 Suitability for use

One of the objectives of the project was to evaluate the pilot system together with the customer to on one hand gather feedback about the usefulness of the product, and on the other hand prioritize development tasks.

The following areas for improvement were identified through the discussion:

- Domain spatial extent was refined based on customer expectations.
- For forecast length, the focus was decided on a time horizon of about one week so it would optimally support planning operations.
- The discussion made obvious the need for revision of the boundary condition treatment, which had a positive impact on the quality of the results.

- The discussions improve understanding on the expected accuracy of the results.

This discussion will hopefully continue as part of further co-operation between the customer and the project team, in future projects.

6 Conclusions

The project developed a pilot data-driven model system with an ocean model emulator at its core. Here, the emulator targeted the seasonal thermocline depth which was deemed important for the sonar weather in the Archipelago Sea. The emulator architecture was based on multiple linear regression, and it was trained against data from a regional high resolution ocean model. The emulator training and execution proved computationally efficient and orders of magnitude faster than required by the host models to produce corresponding output.

The results of the project have proven encouraging and will be taken up in other projects. However, we have also identified several action points for further development. For more elaborated approaches (such as deep neural networks), more training data is needed. Additionally, computational resources for training can be a limiting factor, especially the available memory. Choosing methods that accept larger than memory strategies can tackle this limitation.

In general, we also see further potential for data driven applications in understanding the physical state of the sea. For example, the current approach could be extended towards predicting tracers like temperature and salinity, as well as dynamical parameters such as sea surface height. Ocean model resolution could also be enhanced by building an emulator for the sub-gridscale dynamics using data from even higher resolution models or from, e.g., satellite observations. Such an approach would be useful especially in the near shore regions (on an order 10-100 m from the coast) which would be computationally very demanding to resolve. Yet another application is drift calculations e.g., for oil spills and hazardous substances.

7 Scientific publishing and other reports produced by the research project

The results of the project have been presented in several scientific meetings and conferences.

The results were presented in the 54th International Liège Colloquium on Ocean Dynamics organized at the University of Liege 8 to 12 May 2023. The topic of the conference was machine learning and data analysis in oceanography. This colloquium was an excellent opportunity to discuss the developments done in the project with many of the leading experts in the field.

The results were also reported at the BOOS Science Day held at the Finnish Meteorological Institute on 10 May 2023. BOOS (Baltic Operational Oceanographic System) is the alliance of institutes around the Baltic Sea involved in operational oceanographic activities.

Additionally, the project was presented in the annual MATINE seminars in 2022 and 2023.

A scientific manuscript is in preparation and is aimed to be submitted to a peer-reviewed journal in the second quarter of 2024.

References

Baltic Sea Subsurface Temperature trend from Reanalysis. E.U. Copernicus Marine Service Information (CMEMS). Marine Data Store (MDS). DOI: 10.48670/moi-00208 (Accessed on 01-Nov-2022)

Burnett, W., S. Harper, R. Preller, G. Jacobs, and K. LaCroix. 2014. Overview of oper-ational ocean forecasting in the US Navy: Past, present, and fu-ture. Oceanography 27(3):24–31, https://doi.org/10.5670/oceanog.2014.65.

CMEMS, (2022), Baltic Sea Physics Reanalysis BALTICSEA_REANALY-SIS_PHY_003_011, replaced on 30 March 2023 with “BAL-TICSEA_MULTIYEAR_PHY_003_011, E.U. Copernicus Marine Service Information (CMEMS)”, Marine Data Store (MDS) [data set], https://doi.org/10.48670/moi-00013, 2022.

Fiedler, Paul C., (2010), Comparison of objective descriptions of the thermo-cline, Lim-nol. Oceanogr. Methods, 8, doi:10.4319/lom.2010.8.313.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., Thépaut, J-N. (2023): ERA5 hourly data on single levels from 1940 to present. Co-pernicus Climate Change Service (C3S) Climate Data Store (CDS), DOI: 10.24381/cds.adbb2d47 (Accessed on 01-Nov-2022)

Krauss, W. "The erosion of a thermocline." Journal of Physical Oceanography 11.4 (1981): 415-433, https://doi.org/10.1175/1520-0485(1981)011<0415:TEOAT>2.0.CO;2.

Marill, K.A. (2004), Advanced Statistics: Linear Regression, Part II: Multiple Linear Re-gression. Academic Emergency Medicine, 11: 94-102. https://doi.org/10.1197/j.aem.2003.09.006

Miettunen, E., Tuomi, L., Westerlund, A., Kanarik, H., and Myrberg, K.: Transport dynamics in a complex coastal archipelago, Ocean Sci., 20, 69–83, https://doi.org/10.5194/os-20-69-2024, 2024.

Westerlund, Antti, et al. "Refined estimates of water transport through the Åland Sea in the Baltic Sea." Ocean Science 18.1 (2022): 89-108, https://doi.org/10.5194/os-18-89-2022.

Puolustusministeriö
Försvarsministeriet
Ministry of Defence

MATINE

Ministry of Defence

MATINE
the Scientific Advisory Board
for Defence

Eteläinen Makasiinikatu 8, Helsinki

PO Box 31, 00131 Helsinki

defmin.fi

ISSN PDF: 2984-102X

ISBN PDF: 978-951-663-186-1