Machine Ocean

Combining Machine Learning and Earth Observations to Improve Simulations of Turbulent Behavior in the Earth System

1. Excellence

1.1 State of the Art, Knowledge, Needs and Project Objectives

The climate system is based, at a global scale, on a simple energy conservation mechanism: solar energy is received by the Earth, and either stored in the Earth system, or for the vast majority of it, emitted back into space. On the other hand, on a regional or local scale, the climate system is very complex. Since the solar energy input is not uniform with latitude, and the proportion of energy absorbed by the Earth and reflected into space depends strongly on the local environment and atmospheric layers, net energy flux gradients arise. These gradients, in turn, are a driving force for many atmospheric and oceanic phenomena, such as winds and sea currents, which are key to both weather and climate dynamics.

Due to the large scale of geophysical systems, the level of nonlinearity and complexity of winds and currents is considerable. This is visible through the value, among others, taken by the Reynolds number $Re = UL/\nu$ in geophysical flows, where U is the typical velocity of a flow, L its typical length scale, and ν is the viscosity of the fluid. The Reynolds number quantifies the relative importance of inertial and viscous effects in a fluid flow which, in essence, is a measure of its level of non-linearity and complexity. Typical values of Re in the atmospheric boundary layer are within $10^6 - 10^9$, while values as high as 10^{12} can be obtained in tropical cyclones. Such high values imply that simulation costs of these systems are huge and that one cannot, in any foreseeable future, simulate the atmosphere and ocean down to the smallest relevant scales. Therefore, one has to rely on closure models to estimate phenomena happening at the small scales of the atmosphere and oceans systems when performing numerical simulations for either weather or climate purposes. Said otherwise, while the detailed equations that govern the motion of the air in the atmosphere and of the water in the oceans (namely, the Navier–Stokes equations together with the thermodynamics laws for each fluid) have been known since the mid XIXth century, it is in practise impossible to simulate atmospheric or oceanic phenomena from those 'exact' equations in a bottom-up approach. To solve this problem, physicists and meteorologists resort to simplified closure models to represent small-scale dynamics at a reduced computational cost in their models. As a consequence, weather or climate simulations are only as good as the closure models on which they rely for representing small-scale dynamics.

The shear stress $\boldsymbol{\tau}$ across the sea–atmosphere interface (also called momentum transfer) is one such quantity that cannot, in practise, be calculated from a bottom-up approach relying on discretization of the Navier-Stokes equations simulated on full scale geophysical systems. In addition, while it is possible to measure $\boldsymbol{\tau}$ directly by measuring air horizontal and vertical velocity fluctuations, this requires sophisticated instrumentation, processing and platforms. This is particularly true over the world's oceans where the velocity measurements made on ships and buoys require motion correction to remove the platform contamination before computing $\boldsymbol{\tau}$. As a result, oceanic stress estimates are costly and difficult to produce, and have only been measured in a few sparsely distributed places over the ocean.

As a consequence, scientists and engineers have developed a variety of closure models for $\boldsymbol{\tau}$ over the last 150 years. These are based on a combination of simple physical and scaling arguments, fitting to experimental data both from direct measurements in the ocean and from carefully designed experiments in the laboratory, and heuristics. While models for $\boldsymbol{\tau}$ are well known in some controlled conditions, for example the "law of the wall" with ad hoc values over a smooth, flat wall, obtaining such a model at the sea–atmosphere interface over a variety of wind and wave spectra is very challenging. As a consequence, a variety of models are available, ranging from the relatively simple Charnock's roughness length model formulated in 1955, which relies on a single parameter, to much more complex parametrizations featuring as many as 7 native parameters (COARE 3.5 by Edson et al. (2013)). While impressive and complex, none of these models are perfect, and finding better closure models for quantities such as $\boldsymbol{\tau}$ remains a question of both high scientific interest and paramount practical importance for both weather and climate models.

Therefore, in this project we want to investigate a new approach to the old, challenging question of closure models formulations. For this, we will be using modern methods at the intersection between Meteorology, Physical Oceanography, and Machine Learning (ML). More specifically, we want to develop closure models for $\boldsymbol{\tau}$ based on a ML approach applied on large datasets that have become recently available thanks to both in-situ experimental campaigns, and satellite observations such as provided by the Sentinel program. This would be a significant paradigm shift, attempting to replace the approach based on physical arguments and heuristics by a fully data-driven approach. In this context, we want to both develop a new closure model for $\boldsymbol{\tau}$, compare it to existing ones, validate and measure its accuracy, and extract physical knowledge and understanding about the underlying physics by comparing the output of the new ML models with more traditional parametrizations. Finally, we will deploy these new ML parametrizations for $\boldsymbol{\tau}$ in storm surge operational models, and investigate how this can lead to better surge predictions and, therefore, participate in improving the protection of populations, the environment, and economical activities, against natural disasters. In addition to obtaining these results, our work will, on a broader perspective, set milestones in the application of ML-based methods to meteorology, ocean modeling, and climate science, by both developing new methodologies and guidelines, building expertise at the interface between several scientific communities, and bringing to light both strengths and weaknesses of ML-driven approaches compared with more traditional physics-based models.

1.1.1 Primary Objective

To investigate whether it is possible to gain further insights into the estimation (including uncertainty quantification) of the vertical momentum transfer $\boldsymbol{\tau}$ between the atmosphere and ocean. This will be done by developing parametric models using recent Machine Learning algorithms trained on the enormous set of Copernicus Earth Sentinel Observations as well as direct measurements from buoys and masts.

1.1.2 Secondary Objectives

1. To investigate whether Machine Learning can improve storm surge forecasting.
2. To build up knowledge and skills in Machine Learning methods within the meteorology, oceanography and climate research communities through research and publications on problems of direct relevance to these communities.
3. To efficiently harness the vast information generated by state-of-the-art Earth Observations programs to investigate changes in the Earth System from a new perspective, namely a primarily data driven perspective, rather than the traditional model driven one.

1.2 Research Questions and Hypotheses, Theoretical Approach and Methodology

In addition to the state-of-the-art described in section 1.1, we observe that the fields of geophysics, meteorology, and climate science are experiencing a revolution due to the amount of data and computational power that became recently available. Therefore, it now becomes possible to take advantage of these to bypass traditional closure model formulations, and our project is based on the following hypotheses:

- It is possible to formulate better closure models based on data than on physical arguments and heuristics, if enough data are available. This is due to the inherent complexity of the underlying equations, their mathematical intractability, and how it contrasts with the comparatively simple analytical models usually formulated to address the closure problems.
- To formulate ML driven closure models, large amounts of data and computational power are needed. These have recently become available for the first time in the quantity and quality needed, and, therefore, it is now the right time to invest into this direction.
- Once developed and thoroughly tested, ML based closure models can be used to both enhance our understanding of the underlying physics ("low level" value), as well as our models of the ocean, weather and climate ("high level" value), and to predict and mitigate the effect of natural disasters / climate change ("societal and environmental" value).

- Testing these hypotheses, and obtaining useful results, will rely mainly on two ingredients: ML algorithms and methodologies, and data.

1.2.1 Machine Learning Methodology

ML methodologies will be (together with data, see next paragraphs) the fundamental building blocks of the Work Packages (WPs). Regarding the ML aspect of the project, recent developments in machine learning and Deep Neural Network (DNN) methods in general, and Convolutional Neural Networks (CNN) in particular, allow for developing parametric and predictive models of physical processes without detailed description of the underlying physics ("model-free" approach). In this project, we will gradually increase the complexity of the Machine Learning models used, starting from simpler NN models, self-organizing maps (Chapman and Charantonis, 2017), and Random Forests, and moving towards complex models such as CNNs or recurrent NNs. In all parts of this work, depth and size of the models will be gradually extended, exploring the effect of hyperparameters, until a robust, but simple parameterization is found. In addition, when we include spatial information from satellite data to develop a more advanced parameterization, we expect to rely primarily on CNNs, as they are at the core of many recent breakthroughs owing to their properties of translation invariance which make them particularly effective on the analysis of structured data such as obtained from satellite images (Lguensat et. al., 2018). Thanks to a culture of sharing, new ML algorithms are typically made available quickly through open source packages, therefore, testing a variety of different ML approaches is easily feasible within this project. More specifically, we will build on state-of-the-art, open source ML tools such as TensorFlow, Keras, and Scikit-Learn.

The difficulty in applying these already developed ML algorithms stems mainly from two points. Firstly, and most challenging, is the need for large and quality controlled training datasets. However, this risk is mitigated thanks to both the level of expertise of the participants to this project with the data needed, and the heavy emphasis on data gathering (see the next paragraph and WPs descriptions). Secondly, and slightly less challenging, is the need for metaparameter tuning. This includes finding the right Machine Learning algorithms, finding the size of the network when an NN is used, and determining several meta-parameters such as the learning rate. While there is no theoretical difficulty in this parameter tuning phase, it is largely based on empiricism and experience and may require a significant amount of time and expert knowledge. This is reflected in the structuring of WPs, and there also the team includes several people with strong Machine Learning experience, which makes us confident in our ability to succeed.

1.2.2 Data Gathering Methodology

One can observe vertical turbulent fluxes in the atmosphere (which correspond to direct measurements of $\boldsymbol{\tau}$) directly with sonic anemometers sampling at typically 20 Hz, which resolves the turbulent fluctuations. It is then rather straightforward to calculate averaged (typically over 10–30 minutes) velocity fluctuations covariances such as $\overline{u'w'}$, which corresponds to the vertical flux of momentum when multiplied by ρ. Over land, the sensors can be mounted on masts fixed on solid ground to yield reliable observations. Over the ocean, things are more complicated due to the movements of the sea surface, which must be corrected for before computing the covariance (Edson et al. 1998; Flügge et al. 2016). Another alternative is to install a mast on a land point very close to the sea. This imposes limits to the wind directions where valid observation can be obtained, and also the land disturbances must be carefully examined and accounted for (Rutgersson et al., 2018). Such data are available from the site on Östergarnsholm island, which has been shown to represent open sea conditions for certain wind direction sectors. The 30 m instrumented tower has been used for measuring fluxes, atmospheric turbulence, mean parameters and water properties semi-continuously for the past 20 years (Rutgersson et al. 2008; Högström et al. 2008). Similar data are available from a land based mast at the edge of the Atlantic Ocean, and a 22 m Air–Sea Interaction Tower (ASIT) located 3 km offshore, that are part of the Martha's Vineyard Coastal Observatory (Austin *et al.* 2002). That land mast has collected stress estimates for approximately 15 years. In addition, several years of data from three Swedish lake sites provide complementary information about surface fluxes from smaller water bodies.

Data from two additional towers, and five open ocean buoy deployments that produced turbulent flux measurements during the RASEX, MBL, CBLAST, CLIMODE, SPURS-1 and SPURS-2 campaigns (see Mahrt *et al.* 1996, Hristov *et al.* 2003, Edson *et al.* 2007, Marshall *et al.* 2009, Edson *et al.* 2013 and Clayson *et al.* 2019), will be available for our project. We will also use several years' worth of data from four flux buoys deployed as part of the Ocean Observatories Initiative (OOI) in the Irminger Sea, Southern Ocean, Mid-Atlantic Bight and off the Oregon Coast. Other buoy deployments are planned during the period of this proposal and will likely be available for this project. All of these buoys sample relatively close to the air–sea interface, which requires careful interpretation of the data (i.e., they effectively observe wave-influenced atmospheric boundary layer), and motion correction. However, the CLIMODE buoy and tower measurements were shown to give consistent results in the development of the COARE 3.5 algorithm in Edson *et al.* (2013), which gives us good hope that we will be able to apply such corrections. Preliminary results from the other research and OOI buoys are also consistent with COARE 3.5, but with some interesting differences due to differing wave conditions. The combination of these data will provide the largest in situ $\boldsymbol{\tau}$ measurement dataset ever used for parameterization development.

Finally, remote sensing data from several satellite missions (including Sentinel I) will be used as inputs to the models, as visible in the WPs. The Norwegian Meteorological Institute has both experience and the infrastructure needed to select and analyze such data. Satellite measurements of backscatter are available from two main types of observations: Synthetic Aperture Radar (SAR) and scatterometers. The common denominator of the SAR and scatterometer sensors is that they emit microwave pulses and measure the backscatter from the surface. Both types of sensors see centimeter-scale ripples; the amount of energy scattered back to the sensor will depend on the surface roughness. The surface roughness over the open ocean is largely due to small waves (on the order of cm) set up by local winds (Liu, 2015). These small waves scale with the dominant wind waves, ranging from cms to tens of meters support most of the momentum exchange with the atmosphere (e.g., Edson *et al.* 2013; Sullivan *et al.* 2014). These wind-waves are driven by the surface stress and we expect there to be a strong correlation between $\boldsymbol{\tau}$ and radar backscatter. Regarding SAR backscatter data, the Sentinel 1, 2 and 3 satellites in the European Union's Earth Observation Programme, Copernicus (http://www.copernicus.eu), constitute a revolution in horizontal resolution for Earth Observations, and they provide freely available data in near real time. MET Norway is developing and hosting the National Ground Segment (NGS) for these data on behalf of the Norwegian Space Centre. SAR data used to be rare and expensive, however with the Copernicus service the volume of data received has exploded. The Sentinel 1 mission consists of two polar orbiting satellites which offers a six day exact repeat cycle at the equator. Due to the near polar orbit, the repeat cycle increases with latitude, e.g. the area around Svalbard has daily coverage. The horizontal resolution is less than 100 meters.

Work Packages

Following these hypothesis and methodologies, we formulate the following Work Packages (WPs):

WP1: Applying Machine Learning Methods to Improve Representations of $\boldsymbol{\tau}$

General Objectives

In this work package, we want to provide better parameterization models for $\boldsymbol{\tau}$. For this, we will investigate several potentially promising directions. These are complementary and will, together, allow us to investigate 1) whether a better parametrization of $\boldsymbol{\tau}$ is possible, and more specifically, 2) whether current models are still missing some of the physics and, in this case, 3a) provide evidence of what physics is missing and 3b) how classical parametrizations could be improved. To perform this work, we will rely on both pure *in situ* data (Part I) and a combination of *in situ* and remotely-sensed data (Part 2), together with a ML approach.

WP1, Part 1: Using *in situ* Measurements

The first part of this work package will focus on using *in situ* measurements of "predictor quantities" such as wind speed, surface currents, waves statistics and state variables (e.g., temperature, pressure and specific humidity) to build ML-based parametrizations from *in situ* measurements of $\boldsymbol{\tau}$. This is a direct application of the Supervised Learning paradigm, where the predictor quantities are used as input to the Artificial Neural Networks (ANN) while $\boldsymbol{\tau}$ is the quantity to learn and is used as output. Training in this case aims at finding the set of weights of the ANN that provides the most accurate mapping from the predictor input to the output, $\boldsymbol{\tau}$, both of which are directly measured *in situ*.

WP1, Part 1, Approach 1: Direct Learning

In this Part 1, two more specific methodologies will be used. First, we will perform "direct learning" from input to output, in the most traditional fashion in ML. The scientific value of this approach will be twofold: On the one hand, we will be able to compare the quality of the ML approach to the current state-of-the-art parametrizations. In the case where the ML approach is found to provide lower errors, this will be a strong hint that the current parametrizations can be further improved. On the other hand, if the traditional parameterizations are better than results obtained with this ML approach, this may suggest that either present models are already optimal, or the ML approach is not yet at maturity to compete with these, e.g., because even larger datasets are needed. Second, by performing a wide array of trainings while changing the nature and number of predictors, we may identify which predictors have the most significance for $\boldsymbol{\tau}$, and possibly suggest new parameters that should be used in models for $\boldsymbol{\tau}$.

WP1, Part 1, Approach 2: Residual Learning

The second approach will consist in performing "residual learning". There, we will continue using the same predictors as in direct learning, but the output to match will be the residual error of state-of-the-art parametrizations for $\boldsymbol{\tau}$, i.e., the discrepancy between the output of these classical models and the actual value of $\boldsymbol{\tau}$ on each individual measurement, rather than $\boldsymbol{\tau}$ itself. If the ANN is able to perform a statistically significant prediction of the residual, this will again provide evidence that better models should be attainable. Moreover, analyzing the sensitivity of the trained ANN (i.e., our model of the residual) on its inputs will provide guidance in further attempts to improve classical parametrizations, by pinpointing in a systematic way how their error depends on the input parameters. Adopting a residual approach may also provide some benefits from an operational point of view, as previous studies have shown that this is a promising avenue for combining the robust but possibly less accurate approach (i.e., the classical models based on general physics and heuristics), with the ML approach which has been found to usually be more accurate, but also more brittle.

WP1, Part 2: Introducing Co-Located Remote Sensing Data

In the second part of this work package, we want to introduce co-located remote sensing (RS) data inside the set of predictors used. The motivation here is twofold: First, adding RS data that estimates physical quantities that are not measured in-situ (e.g., surface roughness via backscatter) naturally enriches the predictors set, which is expected to allow for improved ML parametrizations. Second, *in situ* measurements of $\boldsymbol{\tau}$ are only available at a limited number of physical locations. Therefore, the use of ML to directly estimate $\boldsymbol{\tau}$ from the surface roughness measured by active sensors (e.g., SAR and scatterometers) will provide many more global observations of stress than from in situ alone. An advantage of this approach lies in the fact that the *in situ* values of $\boldsymbol{\tau}$ used to develop the parameterization naturally include the effects of surface currents and stability. This fact is the reason why, ideally, the use of scatterometry to estimate winds requires training by stability adjusted neutral winds measured relative to the sea surface. In reality, few of the buoys used in training have all of the measurements required to perform stability correction, and fewer still have measurements of the surface currents. The direct use of the surface stress gets around these problems. Therefore, we expect a relatively robust parameterization of $\boldsymbol{\tau}$ from backscatter without the need for additional predictors. If successful, this approach would greatly expand our estimates of $\boldsymbol{\tau}$ from a few *in situ* measurements to global estimates. These estimates could be used directly as a surface boundary condition in atmospheric and oceanic models.

While there is no conceptual difficulty there, since this is – given enough data – an application of Supervised Learning broadly similar to the one of part 1, there are technical and practical challenges to this approach. Namely, the nature of the RS data we will use, i.e. satellite data, means that the amount of co-located (in space and time) data available will be much lower than solely from *in situ* measurements. Indeed, RS data is only available as a series of snapshots for each *in situ* measurement position as a satellite passes by, rather than as co-located time-series. To attack this problem, we will investigate two different approaches.

WP1, Part 2, Approach 1: Direct Machine Learning of τ Using Remote Sensing Data

While acknowledging that co-location will reduce the amount of data available for ML, our initial approach will attempt to apply Supervised Learning to develop a parameterization between RS data and τ. For this, we will first generate a data-set containing all the RS data collocated in time and space with the in situ data. In this approach, we will use the coincident in situ/RS dataset to determine a new geophysical model function to relate the RS measured surface roughness directly to τ. This new geophysical model function can then be used with all of the RS data to provide a global multi-yearly τ dataset.

WP1, Part 2, Approach 2: Transfer Learning / Fine-Tuning of the WP1 Part 1 ML Models

If the co-located dataset is too small to perform learning of a τ closure model from scratch, we will apply classical techniques from the ML community. These techniques rely on transfer learning and fine tuning of models obtained in Part I to integrate RS information into the ML approach. More specifically, we will use fully trained models obtained in Part I, and apply a second step of training, this time using the initial predictors from Part I extended with the RS data. Since the underlying τ-closure model to learn is still identical to the task of Part I, this means that all the network will have to do is to slightly fine-tune its weights to take into account the RS information, which is much less data-intensive than learning a new set of weights from scratch. We have, therefore, good hope that such a transfer learning / fine-tuning approach will be fruitful. This can be performed in the context of both direct and residual learning, and the points discussed in Part I are still valid and will be investigated here also. The understanding gained through this approach will provide further information regarding the usefulness of surface roughness for understanding whether the residual errors in the bulk flux parameterization are systematically associated with surface roughness. This information can then be used as an update within the bulk flux parameterization.

WP2: Using Machine Learning Approaches to Enhance Storm Surge Forecasts

General Objectives

In this work package, our goal is to use Machine Learning to improve operational storm surge forecasts for the Norwegian coast. This is a critical mission for safety of both people and infrastructures, especially in a general context of increased storm intensity and future rising sea levels due to climate change, and this work package will thus have a large societal impact. In addition, while the datasets we have and, therefore, the ML models we will obtain, may be at least in part specific to the North Sea and Norwegian coast, we expect the methodology we will develop to be applicable to many other endangered coastal regions around the world.

Storm surge modeling is made challenging due to several reasons. First, results are dependent on the exact bathymetry of the domain and the interaction between topography, tides and atmospheric forces like wind and pressure. This is challenging to render accurately in models, as the computational costs imply that simulations are performed at a grid size of orders of kilometers, much larger than most bathymetry features. In the project we will be working with a storm surge model of considerably higher horizontal resolution. Second, storm surge depends on the momentum transfer τ which is, as discussed previously, challenging to estimate. As a consequence, there is promising potential in combining traditional approaches, which are well understood and have been used for a long time, and ML techniques, which may allow to alleviate some challenges encountered by these traditional methods. Therefore, to perform this work, we will study three non-mutually-exclusive approaches. First, we will investigate the potential of ML models for predicting storm surges, both directly and through a residual approach, similar to the work of WP1. Second, we will investigate how current numerical models can use the τ parametrization from WP1, and how this may be used to improve the results. Third, we will combine both previous methods and investigate how they may, together, provide even better results.

WP2, Part 1, Methods 1 and 2: Perform ML-Based Storm Surge Predictions

The first task of this WP will use a training dataset to perform ML-based storm surge predictions. In this approach, we will first perform direct prediction (method 1), where traditional modeling is bypassed altogether. This may produce, similarly to the results of WP1, a hint of how optimal, or lacking, traditional modeling is. The training dataset consists of records of sea level from the 24 regular sea level measurement stations along the Norwegian coast, along with ustream data from stations around the rest of the North Sea (around 30 stations), together with a set of predictors (including air pressure, wind speed and waves). While the sea level data consist of relatively few measurement locations, the extent of the time series available (spanning over 60 years), as well as the measurement frequency (one measurement per station each 10 minutes), mean that the dataset available is large enough to apply ML techniques. In addition, the datasets generated will be used for performing residual learning of the error between traditional surge models and observations (method 2), similarly to the methodology used in WP1. The main implications are similar to what is discussed in WP1, and the reader is redirected to this section for further details.

WP2, Part 2: Using the $\boldsymbol{\tau}$ Closure Model from WP1 to Run Existing Surge Models

Since the momentum transfer $\boldsymbol{\tau}$ is known to be both one of the key factors that influence storm surge amplitudes (Sætra et al., 2007), and a quantity that is challenging to measure or estimate (see the whole WP1), we expect that storm surge modeling will be improved by using ML-enhanced $\boldsymbol{\tau}$-closure models as an input. Therefore, we will combine the results of the WP1 together with existing storm surge modeling, and investigate how this may allow to improve the quality of predictions.

WP2, Part 3: ML of Errors in Traditional Storm Surge Forecasts Using the $\boldsymbol{\tau}$ Model from WP1

Finally, one can observe that the tasks WP2 Part 1 and WP2 Part 2 are independent approaches. Therefore, it appears natural to attempt to combine first surge predictions using traditional models taking as input ML-estimates of $\boldsymbol{\tau}$, and second on top of that to perform residual learning of the errors observed compared with field measurements. This will, therefore, allow to effectively distinguish the different possible sources of errors (incomplete, noisy, challenging $\boldsymbol{\tau}$ estimates on the one hand, and physical model imperfections on the other hand), and to provide a separate correction for these sources of errors.

WP3: Expertise Development in Meteorology and Oceanography

In this WP, we will develop expertise and spread the knowledge generated by the project by 1) co-organizing a summer school on ML in meteorology, oceanography, and climate, 3) develop educational material based from the work done in the two first points, 3) write a detailed review and summary on ML applications to meteorology and oceanography.

WP4: Project Management

Annual meetings, reporting, and general administration.

General Project Considerations

<u>Risks</u>: The project is planned as an agile, iterative process, which means that dead ends are expected and the course of the project is expected to change. A risk is that there is not enough data available today to improve on today's bulk parameterizations. In this case, we will focus on the low risk/medium gain goal of developing the basic machine learning methods needed, address what types of data are needed for their successful training, and publish these results. We will also mitigate this risk, by supplementing observations with systematic physics-based computer simulations to provide training data, which gives us a low risk, medium gain outcome. Another risk is that the search for effective machine learning methods to provide a good parameterization is too extensive. This will be mitigated by starting from simple, but coarse models, and then gradually refining and increasing the complexity of the models.

<u>Interdisciplinarity</u>: this project is naturally interdisciplinary, at the interface between geosciences, meteorology, fluid mechanics, and machine learning.

<u>Gender</u>: the project has good gender balance, with a female lead and three female project partners. We will put special emphasis on recruiting a female PhD candidate.

<u>Ethics / societal value:</u> there is no ethical cost or risk associated with the project. In addition to the intrinsic value of the science that will be carried on, the knowledge output of this project will be used to improve surge models, which is a useful function for the protection of society and the environment.

1.3 Novelty and Ambition

As pinpointed in the previous sections, this project is both ambitious and novel, as it 1) aims to address a current shortcoming on weather and climate models (namely, the limitations associated with closure models for $\boldsymbol{\tau}$), which is known as a harduous problem 2) will use for this a novel ML approach, at the cutting edge of recent technical developments, capitalizing on the expertise available at the Norwegian Meteorological Institute and its partner organizations, 3) will apply the knowledge generated to a complex problem, i.e. storm surge modeling, 4) has a strong international component with 2 domestic and 2 foreign partner institutions.

2. Impact

2.1 Potential for Academic Impact of the Research Project

The project will have a major impact on the geophysics, meteorology, and fluid mechanics communities as highlighted in the previous paragraphs. Indeed, the formulation of closure models is an outstanding problem, and 1) defining such models by using a novel ML approach, 2) generating physical understanding of existing models by comparing with the ML results will be of major academic and practical value, both from a theoretical point of view (better understanding of ocean / atmosphere couplings), and an operational point of view (better surge and, on the long term, weather and climate models). In this context, we will for the first time to our knowledge introduce a cutting edge (ML) methodology.

2.2 Potential for Societal Impact of the Research

As pointed out in the previous paragraphs, this research can lead to improving weather and climate models. This, in turn, allows more effective risk mitigation for both human populations and economical activities in the context of stronger storms and rising sea levels following climate change. In addition, this project will also allow, through better modeling, to improve our understanding of the climate system, and, therefore, to help provide hard data to further raise public awareness on climate change threats. As a consequence, Machine Ocean is particularly relevant for UN SDGs 13 and 14, and indirectly to goal number 2.

2.3 Measures for Communication and Exploitation

This project will produce results of interest for a broad range of scientists within geosciences, meteorology, and fluid mechanics. In addition, we will put special emphasis on communication and dissemination, as highlighted by the content of WP 3. This will include holding a summer school, participating in the hosting of a workshop, and creating educational material for students at both the University of Oslo and other institutions. In addition, a PhD student will be trained in the course of the project.

3. Implementation

3.1 Project Manager and Project Group

The project will be led by physical oceanographer Dr. **Cecilie Mauritzen** (MET Norway). Mauritzen has extensive experience with leading complex multidisciplinary research projects, including EU, NFR and industry projects. Mauritzen was Lead Author for IPCC's 4th and 5th Assessment Reports. The co-lead of the project will be Dr. **Jean Rabault** (MET Norway, starting July 2020). Rabault is a young scientist who has already been a pioneer in the use of machine learning and Artificial Neural Networks (ANN) in fluid mechanics, both for image analysis for direct measurements of velocity fields (Rabault et. al., 2017), and for modeling and investigation of the active control of the NS equations (Rabault et. al., 2019a,b). In addition, Rabault has the necessary background in both Fluid Mechanics and geosciences, which places him at the intersection of the domains of competencies needed for this project. The rest of the research team is described below (organized under their home institution).

The Norwegian Meteorological Institute (MET Norway)

MET Norway has the national responsibility for operational ocean and weather forecasting. The institute collects and processes real time earth observations on a massive scale, and is among the world-leading institutions in both remote sensing, weather and climate modeling, and management/processing of large datasets. The Machine Ocean team includes: Dr. **Veronica Berglyd Olsen**, who has a background in computational physics and high energy physics, with a focus on numerical modelling and high performance computing. Dr. **Martin Lilleeng Sætra**, who has extensive experience in scientific computing on massively parallel architectures. In his PhD work he focused on efficient shallow-water simulation and visualization on Graphical Processing Units (GPUs). He is currently working on GPU-accelerated particle filters for predicting drift in the ocean. Dr. **Øyvind Sætra**, an expert in ocean and wave modelling. He has 20 years' experience on research and development of operational weather and ocean forecasting systems for ECMWF and MET Norway. **Nils Melsom Kristensen**, responsible for operational storm surge modelling at MET Norway, and has extensive knowledge and experience in high resolution coastal ocean modelling. Prof. **Kai Håkon Christensen** heads the Division for Ocean and Ice at MET Norway. He specializes in upper ocean dynamics and oceanic transport.

University of Oslo, Norway

Professor **Morten Hjorth-Jensen** has a strong track record in applied quantum mechanics and computational physics. He is presently building the Machine Learning group at the Department of Physics. Regarding the training of the PhD student, as well as the organization of the summer school and workshop, this will be performed in cooperation between the Meteorological Institute and the University of Oslo. Both institutions have very strong track records on these points.

Woods Hole Oceanographic Institution (WHOI), Massachusetts, USA

Dr. **James Edson** has been the driving force behind the development of the COARE bulk parameterization (Edson et al., 2013), the most commonly used turbulent flux parameterization in climate modelling today. His research focus includes developing instrumentation and techniques to compute and parameterize atmospheric turbulent fluxes within the marine boundary layer including momentum and energy exchange across coupled boundary layers. He has produced most of the over ocean data to be used during the learning process. Dr. **Carol Anne Clayson** is the Director of the Center for Air–Sea Interaction and Marine Atmospheric Sciences. Her research focuses on air–sea interaction using satellite remote sensing and modeling, with the aim to make more accurate estimates of surface turbulent fluxes (heat, water vapor, momentum) at high resolution. She has used machine learning to develop satellite based air-sea flux products (see seaflux.org).

Uppsala University, Sweden

Professor **Anna Rutgersson** is focusing on air-sea interaction and the impact of surface gravity waves on the atmosphere. Her group has maintained the micrometeorological site Östergarnsholm for direct measurements of vertical turbulent energy transfer over the ocean since 1995, and will provide critical data for the learning tasks. Dr. **Erik Nilsson** has long experience on running LES with a moving lower boundary (representing a surface gravity wave) and also in transition zones between different surface roughnesses and temperatures.

3.2 Project Organisation and Management

The timeline is described in the attached GANTT diagram (Table 1).

		2020		2021				2022				2023				2024		
		3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3
WP1	Part 1	Collect data		Build ML model		Iterate on model and compare vs closure schemes												
	Part 2							Structure data		Direct learning		Transfer learning		Compare to 1.1				
WP2	Part 1	Setup model		direct learning		Analyze results												
	Part 2				integrate from WP1.1			Analyze/compare										
	Part 3								Combine WP2.1 and 2.2				Evaluate robustness					
WP3	Summ. school									Summer school								
	Educ. material						Prepare material						Integrate to teaching					
	Dissemination									Review impact								
WP4	Meetings	Virtual				In-person				Virtual				In-person			Virtual	
	Administration																	
PhD		Recruit		Join WP1 and 2						Summer School							defense	

Table 1: GANTT planning for the project. See the WP descriptions for more details about the tasks.

MET Norway's infrastructure will be used to run the modelling workloads and for public sharing of project results. The local operational infrastructure has been built to fulfil MET's needs for reliability, scalability, throughput and data consistency. The storage infrastructure consists of two file systems of approximately 13 PB in total, connected with up to 100 GB/s aggregated bandwidth. The infrastructure scales and fulfils the need for parallel read and write operations according to requirements. Forecast model data and most pre, post and reprocessing tasks are performed on the Post-Processing Infrastructure (PPI). The system is equipped with hardware to perform serial and parallel computing tasks. This includes four Nvidia P100 GPUs that are highly efficient for machine learning workloads. MET Norway provides computing infrastructure for small to medium scale simulation workloads. Low latency/high throughput is implemented using Infiniband and Omnipath. MET Norway also hosts a virtualization environment based on OpenStack. In addition to MET Norway's own infrastructure, we intend to apply for computational resources at Sigma2. Sigma2's GPU nodes run on the same hardware as MET's own nodes, which makes it easier to move workloads between them.

References

Austin, T., et al., 2002: "A network-based telemetry architecture developed for the Martha's Vineyard coastal observatory." *IEEE J. Oceanic Eng.,* **27**, 228-234.

Chapman and Charantonis, 2017: "Reconstruction of Subsurface Velocities From Satellite Observations Using Iterative Self-Organizing Maps", *IEEE Geosc. and Remote Sensing Lett.*, **14** (5), 617-620.

Charnock, H., 1955: "Wind stress on a water surface." *Quarterly J. of the Royal Meteorol. Soc.* 81.350, 639-640.

Clayson, C. A., J. B. Edson, A. Paget, R. Graham, and B. Greenwood, 2019: "The effects of rainfall on the atmosphere and ocean during SPURS-2." *Oceanography*, **32 (2)**, 86 - 97.

Edson, J. B., and C. W. Fairall, 1998: "Similarity relationships in the marine atmospheric surface layer for terms in the TKE and scalar variance budgets." *J. Atmos. Sci*., **55**, 2311–2328.

Edson, J., et al., 2007: "The coupled boundary layers and air-sea transfer experiment in low winds", *Bull. Amer. Meteor. Soc.,* **88**, 341-356.

Edson, James B., et al., 2013: "On the exchange of momentum over the open ocean." *J. of Phys. Oceanogr.* **43** (8), 1589-1610.

Flügge, M., M.B. Paskyabi, J. Reuder, J.B. Edson, and A.J. Plueddemann, 2016: "Comparison of Direct Covariance Flux Measurements from an Offshore Tower and a Buoy". *J. Atmos. Oceanic Technol*., **33**, 873–890.

Högström, U., Sahlée, E., Drennan, W. M., Kahma, K. K., Smedman, A.-S., Johansson, C., Pettersson, H., Rutgersson, A., Tuomi, L., Zhang, F. and M. Johansson, 2008: "Momentum fluxes and wind gradients in the marine boundary layer – a multi platform study." *Boreal Env. Res.* **13**, 475–502.

Hristov, T. S., S. D. Miller, and C. A. Friehe, 2003: "Dynamical coupling of wind and ocean waves through wave-induced air flow", *Nature,* **422**, 55-58.

Lguensat, Sun,Fablet,Tandeo, Mason and Chen, 2018: "EddyNet: A deep neural network for pixel-wise classification of oceanic eddies", *IGARSS 2018-2018 IEEE Int. Geosci. and Remote Sens. Symp.,* 1764-1767.

Liu, W.T., 2015: "Surface Wind and Stress". *Encycl. of Atmosph. Sci. 2nd Ed., Vol.* **5**.

Mahrt, L., D. Vickers, J. Howell, J. Højstrup, J.M. Wilczak, J.B. Edson, and J.E. Hare, 1996: "Sea surface drag coefficients in RASEX", *J. Geophys. Res.,* **101**, 14327-14335.

Marshall J., et al. 2009: "Observing the cycle of convection and restratification over the Gulf Stream system and the subtropical gyre of the North Atlantic Ocean: preliminary results from the CLIMODE field campaign." *Bull. Amer. Met. Soc.*, **90,** 1337-1350.

Rabault, Kolaas, and Jensen, 2017: "Performing particle image velocimetry using artificial neural networks: a proof-of-concept." *Measurement Sci. and Tech.* **28** (12).

Rabault, Kuchta, Jensen, Reglade and Cerardi, 2019a: "Artificial Neural Networks trained through Deep Reinforcement Learning discover control strategies for active flow control", *J. of Fluid Mechanic*s **865**, 281-302.

Rabault, Kuhnle, 2019b: "Accelerating deep reinforcement learning strategies of flow control through a multi-environment approach", *Physics of Fluids,* **31** *(9), 94-105.*

Rutgersson, A., M. Norman, B. Schneider, H. Pettersson, E. , Sahlée, 2008: "The annual cycle of carbon-dioxide and parameters influencing the air-sea carbon exchange in the Baltic Proper." *J. Mar. Syst.,* **74**, 381-394.

Rutgersson, A, Heidi Pettersson, Erik Nilsson, Hans Bergström, Marcus B. Wallin, E. Douglas Nilsson, Erik Sahlée, Lichuan Wu & E. Monica Mårtensson, 2020: "Using land-based stations for air–sea interaction studies", *Tellus A: Dynamic Meteorology and Oceanography*, **72** (1), 1-23.

Saetra, Ø., J. Albretsen and P.Janssen. 2007: "Sea State Dependent Momentum Fluxes for Ocean Modelling". *J. Phys. Oceanography,* **37** (11), 2714-2725.

Sullivan, P. P., J. C. McWilliams, and E. G. Patton, 2014: "Large-eddy simulation of marine atmospheric boundary layers above a spectrum of moving waves." *J. Atmos. Sci*., **71**, 4001–4027.