



## UvA-DARE (Digital Academic Repository)

### Tackling the Multiscale Challenge of Climate Modelling

Crommelin, D.T.; Edeling, W.; Jansson, F.

**Publication date**

2020

**Document Version**

Final published version

**Published in**

ERCIM News

**License**

CC BY

[Link to publication](#)

**Citation for published version (APA):**

Crommelin, D. T., Edeling, W., & Jansson, F. (2020). Tackling the Multiscale Challenge of Climate Modelling. *ERCIM News*, 121, 15-17. <https://ercim-news.ercim.eu/en121>

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

figure 1, [1]), but going beyond domain decomposition, with possibly moving boundaries and ensuring conservative properties wherever it proves feasible. The main outcome is to devise an optimal strategy for further developments for coastal ocean modelling at Inria, from regional to nearshore.

#### Uncertainty reduction and numerical schemes

Ocean models are built on a combination of numerical schemes and choices of physical terms. The variety of available models in SURF makes it possible to understand the relative strengths and weaknesses of these combinations, and this could enable more appropriate choices to be designed. Moreover, some specific issues like how to best discretize vertically or how to properly deal with moving boundaries are still open questions [2].

#### Uncertainty propagation

A numerical forecast is most useful if associated to a measure of its uncertainties. This is commonly achieved using an ensemble of forecasts representing the range of possible future states. The key is to well represent this range. Inria recently developed a way to represent uncertainties associated to unresolved processes (processes that live at smaller scales than the discretization grid) through stochastic modelling. It is more

consistent with the approximation actually made but requires the stochastic process at an early stage of the equations' derivation [3]. Within SURF, we first develop and validate this approach for our hierarchy of models and second use the same approach to describe the uncertainties associated to ocean atmosphere coupling processes.

With the multitude of scientific questions it raises, SURF is an ambitious research project that requires the mobilisation of varied skills and thus involves complementary Inria teams. Airsea (in Grenoble) will contribute its expertise on oceanic modelling and the coupling of models at very different spatial and temporal scales. Cardamom (in Bordeaux) will complete the numerical methods and the associated uncertainties for the models. Ange in Paris and Lemon in Sophia will be asked to model coastal flows in shallow waters, or on the coupling between ocean and river models. Fluminance (in Rennes) will develop statistical image processing techniques to supply the sub-mesh models. Finally, Defi in Saclay and Mingus in Rennes will contribute theoretical skills on the quantification of uncertainties and partial differential equations, respectively.

During the four years of the project, the Inria teams will also collaborate with

experts from the Bureau of Geological and Mining Research (BRGM), the Hydrographic Service of the Navy (SHOM) and IFREMER on both coastal security issues and the development and validation of the models.

#### Link:

[L1] <https://project.inria.fr/surf/>

#### References:

- [1] M. Tayachi, et al.: "Design and analysis of a Schwarz coupling method for a dimensionally heterogeneous problem", *Int. Journal for Numerical Methods in Fluids*, Wiley, 2014, 75 (6).
- [2] T. Song, et al.: "The shifted boundary method for hyperbolic systems: Embedded domain computations of linear waves and shallow water flows", *Journal of Computational Physics*, 369, 2018
- [3] V. Resseguier, E. Mémin, B. Chapron: "Geophysical flows under location uncertainty", Part I, II & III, *Geophysical and Astrophysical Fluid Dynamics*, 111 (3), pp.149-176, 2017.

#### Please contact:

Arthur Vidard  
Univ. Grenoble Alpes, Inria, CNRS,  
Grenoble INP, LJK, France  
[Arthur.Vidard@inria.fr](mailto:Arthur.Vidard@inria.fr)

## Tackling the Multiscale Challenge of Climate Modelling

by Daan Crommelin (CWI; Korteweg-de Vries Institute for Mathematics, University of Amsterdam), Wouter Edeling (CWI) and Fredrik Jansson (CWI)

***The atmosphere and oceans are key components of the climate system, each involving a wide range of space and time scales. Resolving all relevant scales in numerical simulations with atmosphere/ocean models is computationally not feasible. At CWI, we are tackling this longstanding multiscale challenge by developing new algorithms, including data-based and stochastic methods, to represent small, unresolved scales.***

Simulating the climate system on a computer presents a formidable challenge. A major difficulty is the multiscale nature of the key components of the climate system - the atmosphere and oceans. They possess physical and dynamic processes that occur across a range of spatial and temporal scales. Some aspects of the global atmospheric circulation operate at the planetary

scale, in the order of  $10^4$  km, yet it is also significantly affected by atmospheric convection and cloud formation, processes taking place at scales of order 10-100 m. Similarly disparate scales play a role in oceanic circulation.

Resolving all these scales at once in numerical simulation is computationally unfeasible. Therefore, global

models employ simplified representations, or "parameterizations" of the effect that the unresolved processes have on the resolved-scale processes. Formulating such parameterizations is difficult, and the limitations of common, existing methods of doing so are well-known. The uncertainties and errors of parameterizations are a major source of uncertainty in climate change

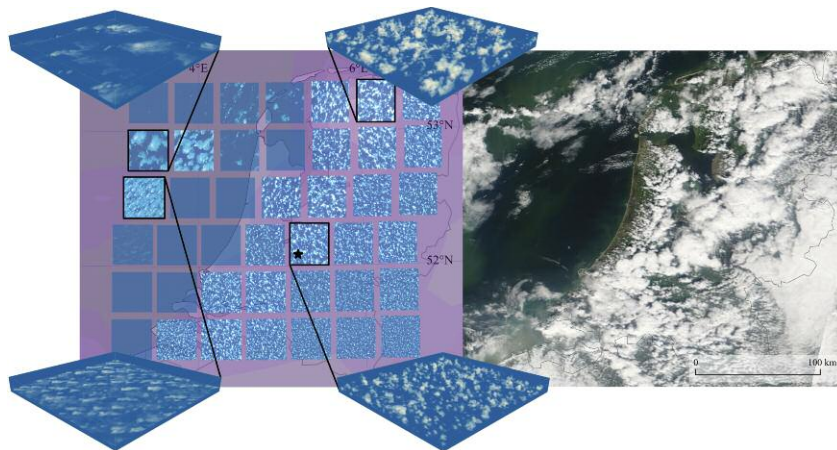


Figure 1: Superparameterized weather simulation over the Netherlands, compared to a satellite image from Terra/MODIS. Each blue tile represents one local, cloud-resolving model, connected to a global model (shown as the purple background). From ref. [1].

simulations (e.g., through uncertainties in the cloud-climate feedback); see also [1] for more background information and references.

New methods and approaches for parameterization are vital. In the Scientific Computing group at CWI, we are working on this topic along two related research lines. One is focused on superparameterization, a computational approach to multiscale modelling and simulation of atmosphere and ocean, in which high-resolution local models (i.e., which cover a small area) are nested in the model columns (vertically stacked numerical discretization boxes) of a coarse-resolution global model that covers the entire earth. Importantly, it concerns a two-way nesting, in which the global model state drives the local models while the local models also feed back onto the global model. It effectively replaces traditional parameterizations based on physical insights and intuition by a computational model based on first principles.

Superparameterization is computationally very expensive, as in principle it would involve high-resolution local models that collectively cover the entire earth (one local model nested within each global model column). The set-up is very well suited for massive parallelization, because the local models do not directly interact with each other; only with the global model. Notwithstanding, it is still much too expensive run roughly  $10^5$  local models in parallel, each with a horizontal domain of, say, 25 km x 25 km and grid resolution of 50 m so that they can resolve atmospheric convection and cloud formation explicitly. To reduce computational costs, previous superparameterization studies have reduced

either the grid resolution or the domain size of the local models. In a joint project between CWI, the Netherlands eScience Center and Delft University of Technology, we have taken a different approach and developed a method in which the local models are only nested in a selected geographical region [1]. The selection is flexible and made by the user, based on factors such as research interest and available computational resources. Outside the selected region, traditional parameterizations are used.

Comparing results from simulations with and without superparameterization, clear differences were observed in the height and vertical extent of the cloud layers. Using superparameterization resulted in higher clouds, in good agreement with ground-based LIDAR observations. Figure 1 (reproduced from [1]) shows a snapshot of the modelled cloud fields over the Netherlands next to a satellite image.

In another research line at CWI, we are developing methods to train a data-based parameterization scheme using data from high-resolution models, such as the local models used in superparameterization, or from observations. By inferring or training parameterizations from such data one can circumvent ad-hoc physical assumptions for formulating parameterizations while also avoiding running high-resolution models for the entire duration of climate simulations (although clearly, a certain amount of computational effort is needed to generate the training data, unless these can be obtained from observations).

Our focus is on data-based methods for stochastic parameterization. The feed-

back from unresolved scales is intrinsically uncertain (e.g., because of chaotic dynamics) and this uncertainty can be represented with stochastic methods for parameterization [2,3]. In [2], we explored several methods to parameterize unresolved scales with stochastic models trained from data. The methods were tested on a multiscale test model (the Kac-Zwanzig heat bath model) that has its origins outside the climate domain yet forms a suitable test bed. One approach, making use of data resampling (or bootstrapping) for parameterization, was shown to be particularly effective.

Building on the results from [2], we use the resampling approach for parameterizing subgrid scales in a simple ocean model in [3], with positive results. The data-driven parameterization approach can be viewed as a methodology for surrogate modelling, as the parameterization is meant to replace (i.e., serve as a surrogate of) the expensive high-resolution model that generated the data. Stochastic (as opposed to deterministic) methods for surrogate modelling have not been explored much to date; our resampling approach is such a stochastic surrogate modelling method. We are currently making this approach part of the software toolkit VECMAtk, under development in the EU H2020 project VECMA (Verified Exascale Computing for Multiscale Applications). Furthermore, we are strengthening our methods by using machine learning methods to build stochastic surrogates with wider capacity.

This research was supported by the Netherlands eScience Center, by the Dutch Research Council (NWO) through the Vidi project “Stochastic models for unresolved scales in geo-

physical flows” and by the European Union Horizon 2020 Research and Innovation Programme under grant agreement #800925 (VECMA project). Furthermore, we acknowledge the use of ECMWF’s computing and archive facilities.

**Links:**

<https://kwz.me/hSQ>  
<https://www.vecma.eu/>

**References:**

- [1] F. Jansson et al.: “Regional superparameterization in a global circulation model using Large Eddy Simulations”, *JAMES*, Vol. 11 (2019), 2958-2979.
- [2] N. Verheul, D. Crommelin: “Data-driven stochastic representations of unresolved features in multiscale models”, *Comm. Math. Sci.*, Vol. 14 (2016), 1213 – 1236.

- [3] W. Edeling, D. Crommelin: “Towards data-driven dynamics surrogate models for ocean flow”, in: *Proc. of PASC 2019*, ACM.

**Please contact:**

Daan Crommelin  
CWI, The Netherlands  
[Daan.Crommelin@cwi.nl](mailto:Daan.Crommelin@cwi.nl)

## A State of the Art Technology in Large Scale Underwater Monitoring

by Gaia Pavoni, Massimiliano Corsini and Paolo Cignoni (ISTI-CNR)

*In recent decades, benthic populations have been subjected to recurrent episodes of mass mortality. These events have been blamed in part on declining water quality and elevated water temperatures (see Figure 1) correlated to global climate change. Ecosystems are enhanced by the presence of species with three-dimensional growth. The study of the growth, resilience, and recovery capability of those species provides valuable information on the conservation status of entire habitats. We discuss here a state-of-the art solution to speed up the monitoring of benthic population through the automatic or assisted analysis of underwater visual data.*

Ecological monitoring provides essential information to analyse and understand the current condition and persisting trends of marine habitats, to quantify the impacts of bounded and extensive events, and to assess the resilience of animal and plant species.

The underwater world is a hostile working environment for humans, with researchers’ activities being limited in time and space. Large-scale exploration requires underwater vehicles. The use of autonomous data-driven robotics for acquiring underwater image data is making large-scale underwater imaging increasingly popular. Nevertheless, video and image sequences are a trustworthy source of knowledge doomed to remain partially unexploited. A recent study [1] reported that just 1-2% of the millions of underwater images acquired each year on coral reefs by the National Oceanic and Atmosphere Administration (NOAA) are later analysed by experts. Automated solutions could help overcome this bottleneck.

Underwater photogrammetry represents a useful technology for obtaining reliable measurements and monitoring benthic populations at different spatial scales. Detecting temporal variations in both biotic and abiotic space holders is

a challenging task, demanding a high degree of accuracy and fine-scale resolution. The detailed optical spatial-temporal analysis involves the acquisition of a massive stream of data. Evaluating changes in the benthos at a scale reflective of the growth and dissolution rates of its constituents requires a pixel-wise classification. This task, called semantic segmentation, is highly labor-intensive when conducted manually. Current manual workflows generate highly accurate and precise segmentation for fine-scale colony mapping but they demand about one hour per square metre.

The automatic extraction of information can contribute enormously to environmental monitoring efforts and, more generally, to our understanding of climate change. This operation can be performed automatically by using Convolutional Neural Networks (CNNs). Nevertheless, the automatic semantic segmentation of benthic communities remains a challenging task owing to the complexity and high intra-specific morphological variability of benthic organisms, as well as by the numerous artefacts related to the underwater image formation process.

In our work, we propose to carry out the classification and the outlining of species from seabed ortho-mosaics. From a machine learning perspective, ortho-mosaics displays a reduced variance of distinctive class features, simplifying the task of automatic classification.

While fully automated semantic segmentation can significantly reduce the amount of processing time, current state-of-the-art solutions still lack the accuracy provided by human experts. Besides, the automation of such specific processes requires specific tools to prepare the data and provide control over the results. So, we propose a human-in-the-loop approach in which a skilled operator and supervised learning neural networks cooperate through a user-friendly interface to achieve the required degree of accuracy. This can be seen as an interactive segmentation cycle for the automatic analysis of benthic communities: intelligent tools assist users in the labelling of large-scale training datasets. Exploiting these data, CNNs are trained to classify coral species, and predictions inferred in new areas. In the evaluation of predictions, the human operator intervenes again to correct both ML-based and human annotation errors. This maximises accuracy when it comes to extracting demo-