ERC Synergy Grant 2020

Stochastic Transport in Upper Ocean Dynamics (STUOD)

Imperial College London, Inria (Rennes), Ifremer (Brest)

Abstract:
71% of Earth is covered by ocean. The ocean has absorbed 93% of the heat trapped by human’s greenhouse gas emissions. The ocean’s future responses to this heat absorption and to continued global warming are uncertain.

The STUOD project aims to deliver new capabilities for assessing variability and uncertainty in upper ocean dynamics. Its results should provide decision makers a means of quantifying the effects of local patterns of sea level rise, heat uptake, carbon storage and change of oxygen content and pH in the ocean. Moreover, its multimodal monitoring aims to enhance the scientific understanding of marine debris transport, tracking of oil spills and accumulation of plastic in the sea.

The STUOD approach accounts for transport on scales that are currently unresolvable in computer simulations, yet are observable by satellites, drifters and floats. Four scientific capabilities will be engaged: (i) observations at high resolution of upper ocean properties such as temperature, salinity, topography, wind, waves and velocity; (ii) large scale numerical simulations; (iii) data-based stochastic equations for upper ocean dynamics that quantify simulation error; and (iv) stochastic data assimilation to reduce uncertainty.

These four scientific capabilities must tackle a network of joint tasks achieved through cooperation of three world-calibre institutions: IFREMER (ocean observations, reanalysis); INRIA (computational science); and Imperial College London (mathematics, data assimilation). Our complementary skill sets need to combine to produce a single synergistic effort towards four interlinked goals:
(1) Coordinate and interpret high-resolution satellite and in situ upper ocean observations.

(2) Extract correlations from data needed for the mathematical model.

(3) Perform an ensemble of computer simulations using our new stochastic partial differential equations (SPDE) which are derived by matching the observed statistical properties.

(4) Apply advanced data assimilation and computer simulations to reduce model uncertainty. We believe that to achieve these goals we must rely on synergy in our combined expertise.

Aim
Our project aims to produce a new systematic capability for dealing with the changing regimes of uncertainty in upper ocean fluid transport. Our approach is (i) driven by data and new methods for its analysis, (ii) informed by mathematical modelling, (iii) quantified in concert with computer simulation and then (iv) optimized by using our newly developed methods of data assimilation. Our project aims to deliver profound capabilities for dealing with the uncertainty in prediction of upper ocean transport of heat, salinity, acidity and chemical concentration induced by the computationally unresolvable scales of dynamics currently missing from computer predictions and essential for proper estimates of the uncertainty of these predictions. To achieve this sort of breakthrough, we believe that synergy of our capabilities will be required, comprising (i) guided analysis of high-resolution observations; (ii) large-scale numerical simulations; (iii) new data-based mathematical approaches to fluid dynamics which properly account for uncertainty in observations and simulations; and (iv) recent breakthroughs in stochastic data assimilation techniques which deal with high dimensional data.

Why focus on the upper ocean observations? Answer: its importance to humanity and life on Earth.
The upper ocean is of intense interest. Almost all human interactions with the ocean occur in its first few hundred meters in depth. Likewise, most ocean life is concentrated in the upper ocean, because light is available there for photosynthesis. Air-sea interaction processes in the upper ocean are essential factors for humanity. Humans release over 22 million tons of CO₂ into the atmosphere each day. When carbon dioxide is dissolved in water, it produces carbonic acid. This alters pH conditions for sea life. Air-sea interactions are also essential factors in determining weather and climate. This makes the upper ocean an important arena for science which transcends the boundaries of physics, chemistry, biology, meteorology and climatology.
**Context** To meet these urgent societal needs, project STUOD must take advantage two timely and complementary scientific developments:

*First, the golden age of upper ocean observations is beginning.* Observing and monitoring technologies have proved their worth for oceanography and climate science (e.g., satellite altimetry, Argo profilers, TAO moorings) and have reached maturity for modern coupled data assimilation. In the EU Copernicus ‘Eyes on the Earth’ programme, the French space agency CNES (Centre National d’Etudes Spatiales) will launch (in 2021) the Surface Water Ocean Topography (SWOT) mission. SWOT will map for the first time the surface mesoscale field and a large fraction of the submesoscale field by radar interferometry. Mesoscale eddies contain 90% of the ocean’s kinetic energy and submesoscale eddies and fronts dominate the vertical velocities in the upper ocean.45 Thus, the new regime in observational resolution will determine whether the turbulence seen in high-resolution ocean models is realistic.

*Second, computational simulations have a limiting “irreducible imprecision” compared to measured quantities in the turbulent regimes of the upper ocean.*45 This limiting feature explains the observed irreproducibility among different model schemes which are supposed to be solving the same problem. The imprecision of simulations is due to the variety of independent selections of different numerical algorithms, model parameterizations, and representations of couplings among the different processes. The STUOD project must recognize the reality of computational imprecision and introduce a new approach for computing uncertainty in a controlled fashion. Our new approach is based on a new form of stochastic transport, informed by spatial correlations determined from high-resolution upper ocean observations. As a further step, the STUOD project seeks to reduce this uncertainty to its minimum by comparing with high-resolution data and using recent breakthroughs in particle filtering for data assimilation. Thus, our project seeks to deliver new stochastic parameterizations of small-scale processes while also improving the ocean forecast models based on the improvements of the upper ocean satellite observations now being put into place.

**Scientific Impact.** The STUOD project aims to contribute to the oceanographic community by bringing together and coordinating the expertise of a transdisciplinary group at the interfaces of mathematics, data assimilation, high precision satellite ocean observations and large-scale scientific computation. Building on our previous experience involving fusion of observational data in the Globcurrent project, we aim to apply advanced methods of mathematical modelling, scientific computation and data assimilation to extract and combine data from satellites, drifters and floats to establish a consistent 3D description of the upper ocean dynamics. The STUOD project aims to provide solutions for collecting, coordinating and interpreting data from observations by satellites, drifters and floats of wind, waves, currents, colour and elevation at the ocean surface, as well as temperature, salinity, and velocity in the upper ocean. We plan to make all of these solutions widely available for oceanographic community research.

**Societal Impact.** The STUOD project aims to contribute to the wider society, as well. It aims to provide decision makers with the capability to predict local patterns of sea level rise, as well as heat uptake, carbon storage and change of oxygen content60 and pH in the ocean. Our solutions for processing high-precision information about wind, waves and currents near the ocean surface are planned to be included in ocean modelling software at Inria (NEMO), which is used by many meteorological offices around the world. Our solutions may also benefit international commerce, by providing high-resolution information for use in route optimization systems to reduce fuel consumption and greenhouse gas emissions of ships. This multimodal monitoring of the currents may provide valuable information for tracking the accumulation of plastic in the sea, and locating other debris drifting in the ocean. Our project is also aligned with one of the missions proposed in Horizon Europe47, dedicated to ‘A plastic-free ocean’.

**Why stochastic transport?** A great portion of the data which will be taken in the new regime of high-resolution observations of upper ocean dynamics will not be predictable by existing deterministic simulations. Consequently, the next generation of simulation schemes will face new features that are beyond the capabilities of current deterministic prediction. Answering this challenge will require dealing with uncertainty through stochastic modelling and mathematical analysis, applied in concert with high-resolution Earth observations, computational simulations and analysis of large datasets. Our project aims to deliver essential new capabilities in the prediction of oceanic transport induced by temporal and spatial scales of dynamics that will be unresolvable in direct computer simulations. Delivering these new capabilities will require us to coordinate four types of research endeavour. Starting from observations we must: (i) Compile, interpret and
enhance the new types of high-resolution data from observations of the ocean’s eddies and currents at high resolution; (ii) Derive and analyse new stochastic mathematical models of upper ocean dynamics which are designed to cope with uncertainty in both observations and simulations; (iii) Develop and analyze the computer simulations obtained using stochastic models of the dynamics underlying the observations; and (iv) Apply our new stochastic data assimilation techniques which are intended to reduce the uncertainty in the predictions made in these computer simulations. (See Figure above).

The STUOD organises a cycle of sequential joint tasks achieved through the close cooperation of three world-calibre institutions: IFREMER (ocean observations, inverse problems and oceanography); INRIA (computational science and stochastic modelling); and Imperial College London (mathematical modelling, analysis and stochastic data assimilation).

We do not expect to find new deterministic laws of fluid motion in our work. However, there can be new statistical approaches to the physical and mathematical descriptions of upper ocean dynamics. In fact, the recognition of transport by stochastic (probabilistic) vector fields as the basic paradigm in fluid dynamics has long been the province of turbulence modelling. The use of stochastic transport has also been recognized in estimating statistical model uncertainty in ocean dynamics. From this viewpoint, quantifying the uncertainty and variability of the predictions are crucial aspects of the solution. Stochastic methods offer systematic approaches toward quantitative estimates of uncertainties due to model error and inaccuracy in data assimilation, as well as improved estimates of long-term variability, including estimates of the probability of extreme events. Thus, stochasticity must be incorporated in the fundamental design of physical process parameterizations. The purpose of this proposal is to offer new approaches at this fundamental level for oceanic data assimilation.

**Linking observational data with the theoretical models** The cornerstone of our project is its direct link from the observational data to the stochastic model. This link will be achieved by calibrating the model to data; namely, 25 years of Globcurrent products. The key step is to decompose the flow velocity into its resolved slow component and a stochastic term. This decomposition represents the effects of the rapidly fluctuating components of the velocity that cannot be resolved at the slow temporal and spatial scales. The effects of the rapidly fluctuating components will be extracted (in terms of basis or shape functions, say) from state-of-the-art high-resolution oceanic simulations and multimodal observations using improved data analysis methodologies which combine kernel methods, Koopman operator methods, diffusion maps, and principle component analysis.

**Modeling external factors** Upper ocean dynamics is influenced by many external factors. This includes variable atmosphere fluxes, rain, ice, river runoff, surface and internal waves, Langmuir circulation, mixing and biological processes. Instead of simulating at high resolution each of these complex interaction mechanisms individually, their combined effect will be incorporated in our data-based models through the shape and the amplitude of the random terms, as in reality the observations themselves bear the full probabilistic signatures of all of these complex effects. Thus, extracting the stochastic oceanic representation statistically will include the action of a multitude of physical phenomena without modelling each of them explicitly; rather, representing their effects statistically.

**State of the Art** The design of stochastic representations of geophysical flows for ocean science and meteorology in the past has been problematic. Several strategies have been proposed for that purpose. For example, turbulence models that aimed principally at the effects of ‘backscatter’ of energy from the fast/small scales to the larger scales were devised and found to be influential. Related models have focused greater attention on oceanography and weather forecasting. They have been shown to improve the system forecasting skill.

However, difficulties remain because incorrect results in variance has often led to ad hoc approaches to corrections via stochastic parameterization. Moreover, to estimate the energy dissipation due to the unresolved scales, but also to slow the increase in variance, these models are often artificially damped by ‘eddy viscosity’ diffusion. This concept, introduced via analogy with the molecular viscosity mechanism, dates back to the work of Boussinesq and Prandtl. However, eddy viscosity neglects high frequencies, whose effects on transport are needed to quantify variance, uncertainty and error in numerical simulations.

To the extent that one ignores the nonlinear transport of the rapid fluctuations, the validity of representing their effects on transport via

<table>
<thead>
<tr>
<th>Work packages:</th>
</tr>
</thead>
<tbody>
<tr>
<td>WP1 - Multi-modal ocean data acquisition, analysis and interpretation</td>
</tr>
<tr>
<td>WP2 - Uncertainty representation and stochastic parameterization</td>
</tr>
<tr>
<td>WP3 - Numerical schemes for stochastic fluid transport and diffusion</td>
</tr>
<tr>
<td>WP4 - Multiscale ensemble Data Assimilation and forecasting methods</td>
</tr>
</tbody>
</table>

All the PIs will contribute to all the WPs in a synergetic collaboration as identified above. The WPs cannot be completed without applying the expertise of the entire team.
standard methods such as eddy viscosity is questionable. Moreover, eddy viscosity models are reported to increase forecast error for oceanic models. In addition, careful ‘tuning’ of the noise parameter together with an adapted viscosity to stabilize the model is required for implementing these models successfully, and the success of this ‘tuning’ often does not extend into new flow regimes.

**Stochastic Transport in Upper Ocean Dynamics (STUOD).** Our new stochastic methods will use the observed oceanic velocity correlation statistics with rapid variations in space and time to complete the description of the net transport current velocity as the sum of the resolvable motion plus the observed correlation modes of the sub-grid scales, each evolving by its own random process. Our new mathematical methods use **stochastic calculus** to unify the effects of the ocean transport velocity at both coarse and fine scales by representing the fine (unresolvable) scales by random processes whose statistics are determined from observed datasets and then deriving the equations for the coarse scales, given that their net coarse and fine scale velocity statistics must match the statistics of the observed data.

The stochastic fluid equations which we will derive during the STUOD project aim to provide a new paradigm for the analysis of geophysical flows. They contain a new type of noise that respects the fundamental transformation properties of fluid dynamics. In fact, they represent a new class of stochastic equations for which little is yet known. As a result, new rigorous mathematics will need to be developed to prove their mathematical properties such as well-posedness (WP2). In addition, new methodologies must be constructed (and validated) to solve them numerically (WP3), new types of data assimilation will be needed for determining the optimal solutions (WP4), and new data analysis methods will be needed for extracting the statistics from the data, as input for the stochastic decomposition of the fluid velocity (WP1).

**Our preparation.** Recently we have derived a new class of mathematical models of stochastic transport in fluid dynamics at the most fundamental level (Holm 2015, Memin 2014). This class of models is derived from physical principles according to which the unresolved scales introduce stochastic uncertainty into the transport of fluid parcels. Our models encompass both Newtonian and variational perspectives of mechanics and lead to proper energy conservation and circulation dynamics. They also preserve the fundamental mathematical properties of their deterministic counterparts (Crisan et al. 2018). Moreover, they provide new approaches to subgrid parameterization, expressed both in terms of fluctuation distributions, and spatial/temporal correlations. As such, they introduce **stochastic corrections that are amenable to statistical inference from data** (either observed, or numerical high-resolution data) and yield new data analysis tools and models for turbulent flows. Our new class of models forms the basis for the development of a novel data assimilation technology based on particle filters. Filtering and ensemble techniques require de facto a stochastic representation of the dynamics.

The randomization is most often realized through perturbation of the initial conditions. This approach yields generally insufficient spreading of the ensemble and a poor representation of the error dynamics. The Holm-Mémin class of models establishes a physically relevant randomization via its rigorous derivation at the fundamental level, rather than via ad hoc forcing. See Cotter et al. for another rigorous derivation of the same Holm-Mémin decomposition into mean and fluctuating velocities.

Particle filters are theoretically validated Monte-Carlo based techniques that do not require Gaussian assumptions on the model. They have previously proved successful in solving low-dimensional problems (<100 dimensions). Crisan and his collaborators as well as others have recently found new ways to extend the applicability of particle filters to high dimensional problems. These new theoretically well-founded developments mark the beginning of a fundamental shift in Data Assimilation methodology from Gaussian assumptions to a general (non Gaussian) setup with physical model errors.

The STUOD project develops and uses particle ensemble based Data Assimilation methods with improved forecasting skills and accurate uncertainty estimates. Over the last several years, Bertrand Chapron has pioneered new measurement techniques, and analysis methods implemented in the international data-driven GlobCurrent project. As illustrated above, the Gulf Stream near Labrador Bay is well expressed, colours indicating warmer surface water temperature. One can also track the small-scale erratic paths of drifters on the surface. No existing
data-driven model simulations have ever achieved the accuracy we have achieved in combining the high-resolution observations of the erratic paths with numerical simulations at coarser resolution.

**How will we proceed?** In our synergetic venture, we have organised work packages to achieve in concert the following **iterative sequence of goals** (details of the time lines for these tasks are given in part B2):

1. Compile and interpret unique high-resolution IFREMER’s satellite, float and drifter datasets and high resolution numerical simulation; **(WP1)**
2. Extract the statistical properties from the datasets needed for the mathematical model to separate the velocities of the ocean currents into coarse-scale drift and fine-scale noise; **(WP1+WP2)**
3. Apply the statistics of the fine-scale noise in our new class of mathematical models to derive rigorous equations for the coarse-scale drift velocities in a set of realistic exemplars, discussed below; **(WP2)**
4. Create an ensemble of computer simulations for each exemplar at coarse resolution with the corresponding new stochastic equations calibrated from the statistical properties of the satellite observations; **(WP2+WP3)**
5. Apply newly developed data assimilation methods to this ensemble of simulations to find the most likely solutions among them for the currents and waves, along with their uncertainty and expected variability; **(WP4)**
6. Compare the results with further Satellite Observations; **(WP1)**
7. Design and make improvements in steps (1) and (2) to better formulate the observed statistics; **all WPs**
8. Go to step (3), and complete the cycle again until satisfactory convergence has been achieved.

**Why Synergy?** Synergy (coordinating our expertise, tools, methods, and overlapping collaborator networks) is the key to reaching meaningful convergence of our efforts. To enhance our cooperation, we must establish innovative working arrangements based on virtual institute applications, so that the entire team can be available to each other on a daily basis.

**References**


49. https://www.imperial.ac.uk/ocean-plastic-solutions/ 
50. https://www.nature.com/articles/s41561-018-0152-2