Working groups @ AI4OAC: Organisation and Guidelines
January 20-23 2020, Brest

Broadly speaking, the working groups will aim to organize discussion and numerical experiments to explore the potential of data-driven and AI-based frameworks to explore open questions in ocean, atmospheric and climate sciences. Working groups shall target the following specific objectives:

- AI-based formulations for ocean/atmosphere/climate challenges, including the specification and possibly preparation of associated datasets and evaluation framework;
- The preparation and implementation of benchmarking experiments for OAC-related data challenges.

The workshop will support working group outcomes in the form of: (i) the distribution of open data challenges, (ii) the preparation of short papers for workshops and conferences such as Climate Informatics, ML4PS (Machine Learning for Physical Sciences), ICLR workshop on AI for Earth Sciences,...

Preliminary list of working groups

We will consider both methodological and topical working groups:

- Methodological working groups shall address the development of new learning-based and data-driven models and schemes for ocean-atmosphere-climate processes, especially to account for specific features of geophysical dynamics;
- Topical working groups shall address a topical ocean-atmosphere-climate challenge from a data-driven/learning-based perspective.

Proposed list of methodological working groups (global supervision by R. Fablet)

- Learning-based approaches for reduced-order (stochastic) modeling (Leader: G. Tissot, S. Ouala)
- Learning and Space-Time Extremes (Leader: P. Naveau, P. Ailliot)
- Generative modeling for OAC processes (Leader: L. Drumetz)
- Automated NN generation from PDEs (Leader: O. Pannekoucke)
- Dynamical Systems, Optimal Transport & Physics-informed Learning (Leaders: F. Rousseau, N. Papadakis)

Proposed list of topical working groups (global supervision by J. Le Sommer)

- Learning closures for ocean dynamics (Leaders: J. Le Sommer)
- Learning-based separation of wave and eddy processes (Leaders: R. Lguensat, C. Ubelman)
- Space-time interpolation of geophysical dynamics (Leaders: M. Beauchamp, E. Cosme)
- Learning-based schemes for biogeochemical dynamics (Leaders: L. Memery, E. Martinez)

Organisation

WG leaders will shortly introduce each WG (one slide, 2'+questions) during the plenary Open Forum session (Monday 20, 5pm-6pm), so that each participant can choose the WGs he/she wants to participate to.

To strengthen interactions between methodological and topical aspects, the two types of working groups will run sequentially and not in parallel, meaning that any participant may participate to at least one methodological WG and one topical WG. For each WG session, the first one-hour slot will be dedicated to methodological WGs and the second one-hour slot to topical ones. At any time, any participant may leave a WG and move to another one.

Short description of each Working Group

Learning-based approaches for reduced-order modeling (Leader: G. Tissot, S. Ouala)
The aim of this working group is to discuss about the link between Koopman operator and Neural Network or Kernel-based methodologies to learn the non-linear dynamics of a system. In particular, the choice/discovery of coordinates to represent the state space is determinant for the learning performances and different strategies to guide the learning process will be compared. What kind of mathematical, physical or structural guidance could be considered to adapt the learning for oceanic dynamics? Is it possible to use information from the encoder/decoder in the modelling process? What is the gain of these non-linear dimensionality reduction techniques with respect to classical spectral decomposition methods?

Keywords: Neural networks, RKHS, Koopman operator
Specific objectives: (i) review and link existing techniques, (ii) try to guide the learning with the dynamical model, the physics and/or mathematical considerations, (iii) define datasets and evaluation metrics for benchmarking experiments.

Learning and Space-Time Extremes (Leader: P. Naveau, P. Ailliot)
Coming soon

Generative modeling for OAC processes (Leader: L. Drumetz)
Generative modeling has been one of the most enthusing advances in deep learning in the past few years. Contrary to classical machine learning models, they are able to learn the probability density of the training data and to sample from them. Popular examples of such models include Generative adversarial networks (GANs), Variational Autoencoders and their extensions (Conditional and Cycle GANs), Normalizing flows... The goal of this group is to explore the potentialities of generative models for Ocean, Atmospheric and Climate-related Processes. The handling of uncertainties around the variables involved in OAC processes is of utmost importance, and hence being able to learn and propagate probability distributions through a dynamical model is a key challenge that generative models may help tackle. Conditional synthesis applications are also particular very impressive and have far reaching implications. They are able to generate new samples from the dataset while simultaneously conditioning on another variable. For example we can imagine generating an SST field matching a known SSH field. These networks can also be used for super-resolution and component separation applications which are common problems in OAC sciences.

Keywords: generative modeling, GANs, Variational Autoencoders, probability distribution function learning and propagation

Specific objectives: Categorize the different tasks that can be tackled by generative models. Find applications to concrete problems in OAC sciences and the associated case studies, learn and propagate probability distribution by dynamical models...

Automated NN generation from PDEs (Leader: O. Pannekoucke)
The aim of the working group is to explore the potential of a NN generator, PDE-NetGen. Given a set of evolution equations (PDE), PDE-NetGen renders a NN code able to integrate the system of equations. This is of particular interest when a part of the dynamics is unknown: translated as a NN, it is possible to design a NN parametrization of the unknown part and focus the training on it. This tool is attractive for exploring NN applications in geophysics, with emerging questions. What are the numerical properties of a NN implementation ? How to combine the NN representation with the data assimilation workflow ? How to improve the time scheme depending on the problem ? Can this improve our understanding of existing architecture or trained NN ?

Keywords: PDE/ODE/ResNet ; symbolic computation ;
Specific objective(s): Applications in forecasting and data assimilation in 2D dynamics (eg barotropic dynamics); Application to uncertainty quantification ; Select a list of relevant applications for automated NN generations ;

Dynamical Systems, Optimal Transport & Physics-informed Learning (Leaders: F. Rousseau, N. Papadakis)
This working group is about the links between neural network architectures and discretization of differential equations. Recent works propose stable NN architectures allowing the simulation of dynamic systems. We will discuss the use of such networks for predicting complex physical processes.

Keywords: ODEs/ResNets; FFJORD; Real NVP
Specific objectives: (i) reviewing/selecting relevant network architectures to model ocean dynamics and mass transport, (ii) defining specific problems where stable dynamical laws could be learned from data.

Learning closures for ocean dynamics (Leaders: J. Le Sommer, ...)
Coming soon.

Learning-based separation of wave and eddy processes (Leaders: R. Lguensat, C. Ubelman)
This working group aims to investigate machine learning techniques for the separation of wave and eddy processes present in ocean-derived data, in particular Sea Surface Height (SSH).
A recent preliminary study shows that this issue can be casted into a supervised learning problem (Lguensat et al. submitted) and uses eNATL60 numerical simulation as a testbed to measure the performance of a convolutional ResNet. Data and example Jupyter notebooks will be provided to participants.

Keywords: Sea Surface Height, Convolutional ResNets, eNATL60, SWOT altimeter
Specific objectives: Designing proper metrics for quantifying the quality of the separation, thinking about relevant loss functions to optimize for this problem. Launching a data challenge based on the different collected ideas.

Space-time interpolation of geophysical dynamics (Leaders: M. Beauchamp, E. Cosme)
The space-time interpolation of fragmented observations to make 2D+time or 3D+time gridded fields is still an open problem in many geophysical applications. Two reasons are (i) the observations and observed variables can be of different nature, and (ii) geophysical variables are driven by PDEs. For such variables, it seems natural to rely upon numerical models discretizing these equations, what gives shape to data assimilation. But in practice, the intricate and often nonlinear nature of the underlying dynamics can limit the efficiency of data assimilation techniques.

Keywords: data assimilation, irregular sampling, end-to-end learning
Specific objectives: (i) Try to review the nature of different interpolation problems met in geophysics, and particularly in oceanography. (ii) Make connexions with ideas from the methodological workgroups. (iii) Design realistic interpolation problems to be proposed as data challenges, possibly including the implementation of baseline algorithms.

Learning-based schemes for biogeochemical dynamics (Leaders: L. Memery, E. Martinez)
This working group aims to investigate machine learning techniques to address open questions related to the understanding of marine biogeochemistry/ecology functioning. For instance, we will discuss how/which Deep learning (DL) and DL-based interpretation methods may help identify important spatial (e.g. mesoscale) and temporal (e.g. regime shift) biogeochemical patterns in the ocean in relation with their physical environment? Can we improve (and how) our understanding of physics-to-biogeochemistry/biology/ecology interactions? Can we learn (and how)stable dynamical laws from data?

Keywords: Neural networks, physics-ecosystems interactions and functioning
Specific objectives: Reviewing/selecting relevant network architectures and biogeochemical key questions in order to launch a data challenge.