University of Reading
UNIVERSITY OF SURREY
Scenario DOCTORAL TRAINING PARTNERSHIP
UKRI Natural Environment Research Council

Complex network approach to improve marine ecosystem modelling and data assimilation - CONECDA

Lead Supervisor: Alberto Carrassi (University of Reading, Department of Meteorology and National Centre for Earth Observation, Reading)
Email: a.n.carrassi@reading.ac.uk

Co-supervisors: Jozef Skakala (National Centre for Earth Observation and Plymouth Marine Laboratory, Plymouth); **Stefano Ciavatta** (National Centre for Earth Observation and Plymouth Marine Laboratory, Plymouth); **Ross Bannister** (University of Reading, Department of Meteorology and National Centre for Earth Observation, Reading)

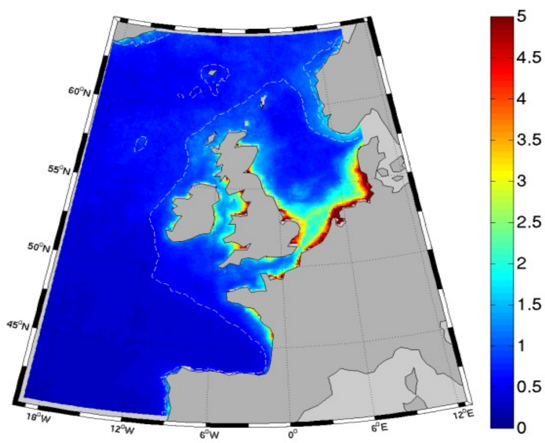

Fig 1: *Surface chlorophyll concentration, which is a proxy of marine life on the NWES (the shelf boundary is marked by the dashed line, from Ciavatta et al (2016)).*

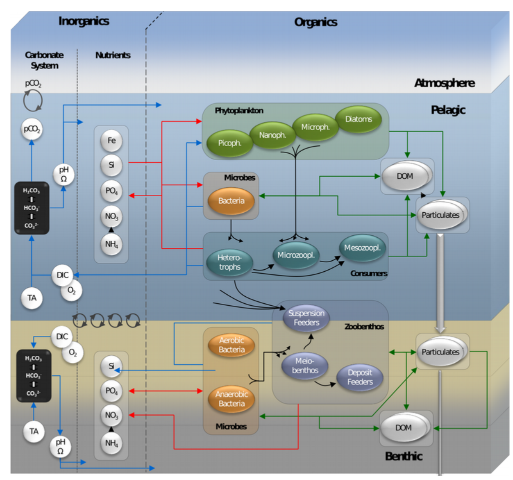

Fig.2: *The schematic representation of the European Regional Ecosystem Model (ERSEM, from Buttenschon et al. (2016)), which is the high complexity ecosystem model of the NWES.*

One of the most important questions in marine ecology is the ecosystem resilience in the face of human pressures on the natural environment, especially anthropogenic climate change (*e.g.* Ives and Carpenter, 2007). From UK policy-making perspective it is crucial to focus on marine ecosystems in the North-West European Shelf (NWES), which is a region of high importance for the carbon cycle and the European economy (*e.g.* Legge *et al*, 2020). Addressing questions about ecosystem sensitivity, or tipping points, using present ecosystem models is computationally costly, as one typically needs to run a large ensemble of simulations with varying model parameters and/or external forcing (*e.g.* Zhang and Pu, 2010). It is therefore highly desirable to bypass the computational cost of the full-complexity models by using creative ideas originating from diverse branches of modern mathematics, artificial intelligence and statistics, such as machine learning and emulators (see *e.g.* Hsieh, 2009; Tokmakian *et al.*, 2012; Brajard *et al.* 2020), or explore contexts in which one can reduce model complexity (*e.g.* Skakala and Lazzari, 2020).

Complex networks (*e.g.* Boccaletti *et al.*, 2006) are a highly efficient mathematical approach to represent connectivity between the degrees of freedom within a diverse range of complex phenomena: from trophic relationships in ecology to representing climate through a network of teleconnected local oscillators, all the way to the financial markets, social networks and the web, neuroscience, epidemic modelling, Ising model, transportation, security and further (for an overview see Costa *et al.*, 2011). We propose here to use complex networks to understand the pathways through which the antrophogenic signal propagates on the NWES, both between different ecosystem variables and between different geographical regions. This will enable us to identify the key variables and the NWES sub-regions having the largest impact on the NWES marine ecosystem dynamics. Such analysis will provide insight into the ecosystem vulnerability, or resilience, however it will also deliver crucial information on which model degrees of freedom are redundant, providing a guide on how to reduce the complexity of the NWES ecosystem model. On this basis we will construct a low-complexity model emulator, relying upon state-of-the-art machine learning tools. In particular the emulator will be designed to improve our current NWES data assimilation system, used to combine marine ecosystem observations with the model forecasts to produce the best possible estimate of the ecosystem state.

This is an interdisciplinary project at the interface of mathematics and environmental science, in which the fast evolving and timely mathematics of complex networks and machine learning methods are combined to deliver new profound insights into marine ecosystems in the shelf seas around the UK. The project can potentially lead to a step-change in the state-of-the-art modelling and data assimilation tools used to represent complex marine ecosystems on the NWES.

Training opportunities:

The student will be partly placed at Plymouth Marine Laboratory (PML), a centre of excellence for marine science and marine ecosystem modelling, where they will be trained in complex networks, NWES ecosystem processes and the NWES ecosystem numerical model. At PML they will also benefit from some training courses on computer programming languages, data processing and machine learning. PML is participating in a MARINT COST proposal, which, if granted, will offer the student the opportunity to participate at various machine learning/emulator building workshops, summer schools and training events. Furthermore, through the National Centre for Earth Observation (NCEO) the student will have access to the large portfolio of the NCEO training courses and will attend its annual conference and young researcher's forum. At the University of Reading (UoR) the student will be exposed to a thriving academic environment and will attend training on dynamical systems and data assimilation. The student will be also given the opportunity to attend international summer or winter school on data assimilation and machine learning, such as the school offered by the University of Utrecht (NL), and also to virtual training organized by our collaborators in Europe and USA. The student will also attend a couple of conferences per year, both in geosciences (*e.g.* EGU or AGU) as well as in machine learning (*e.g.* ICCS).

Student profile:

The student should have a strong honors degree in a quantitative discipline, ideally mathematics, physics, computer science, or engineering. Some understanding of oceanography is a bonus, but it is not strictly required. The student should be confident with developing and editing computer codes (*e.g.* in python, R, Fortran), and should have the personal skills to work in a group of people and work towards the project deadlines.

Funding particulars:

NCEO may co-sponsor this project (50%) - TBC.

References:

- Boccaletti S., et al., Complex networks: Structure and dynamics. Physics Reports, 424(4-5):175–308, 2006.
- Brajard J., et al., Combining data assimilation and machine learning to emulate a dynamical model from

sparse and noisy observations: a case study with the Lorenz 96 model. Journal of Computational Science, 44: 101171, 2020.

- Butenschon, M., et al. "ERSEM 15.06: a generic model for marine biogeochemistry and the ecosystem dynamics of the lower trophic levels." *Geoscientific Model Development* 9.4 (2016): 1293-1339.
- Ciavatta, S., et al. "Decadal reanalysis of biogeochemical indicators and fluxes in the North West European shelf-sea ecosystem." *Journal of Geophysical Research: Oceans* 121.3 (2016): 1824-1845.
- Costa L.d.F., et al., Analyzing and modeling real-world phenomena with complex networks: a survey of applications. Advances in Physics, 60(3):329–412, 2011.
- Hsieh W.W. Machine learning methods in the environmental sciences: Neural networks and kernels. Cambridge university press, 2009.
- Ives A.R. and Carpenter S.R.. Stability and diversity of ecosystems. Science, 317(5834):58–62, 2007.
- Legge O., et al. Carbon on the northwest European shelf: Contemporary budget and future influences. Frontiers in Marine Science, 7:143, 2020.
- Skakala J. and Lazzari P., Keep it simple: A low-complexity model to study scale dependence of phytoplankton dynamics in the tropical Pacific. submitted to Physical Review E, arXiv:2009.03629v1, 2020.
- Tokmakian R., et al., On the use of emulators with extreme and highly nonlinear geophysical simulators. Journal of Atmospheric and Oceanic Technology, 29(11):1704–1715, 2012.
- Zhang H. and Pu Z., Beating the uncertainties: ensemble forecasting and ensemble-based data assimilation in modern numerical weather prediction. Advances in Meteorology, 2010.

https://research.reading.ac.uk/scenario/