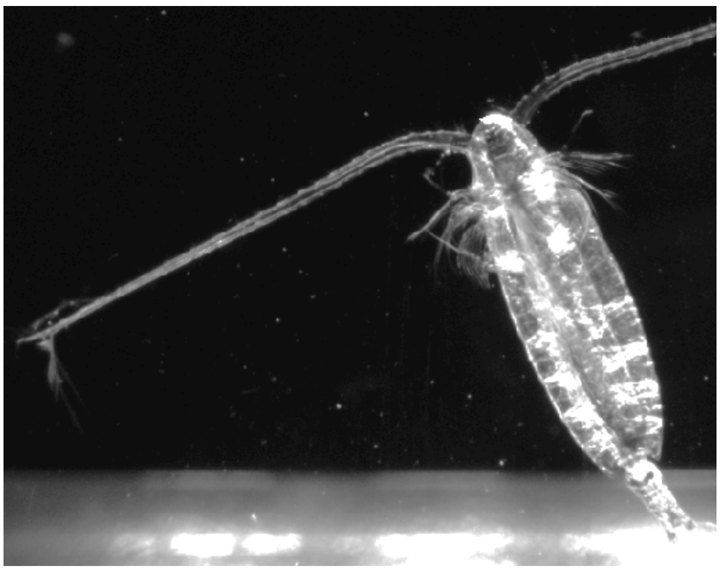
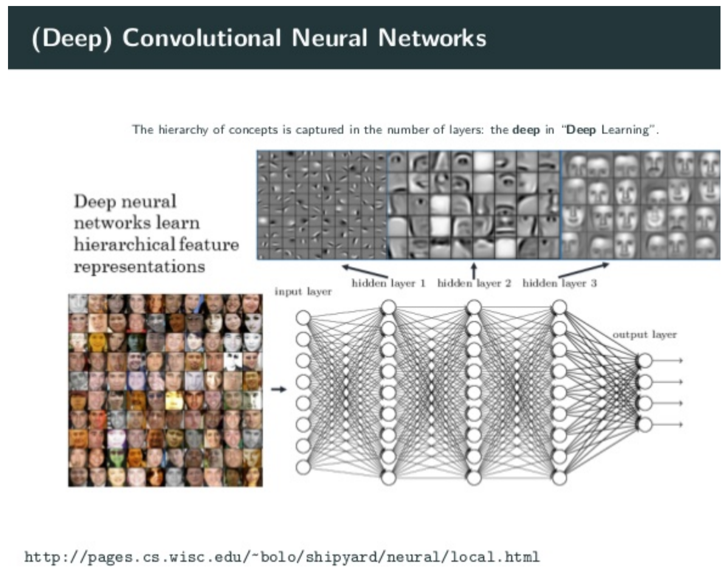
Book of abstracts of the ARTIFACTZ Workshop Artificial intelligence for characterizing plankton traits from images

Frédéric Maps, Jean-Olivier Irisson, Sakina-Dorothée Ayata

ARTIFACTZ Workshop





Book of abstracts of the ARTIFACTZ Workshop: Artificial intelligence for characterizing plankton traits from images

Frederic Maps, Jean-Olivier Irisson, Sakina-Dorothée Ayata (Eds.)

24-26 of April 2019, Villefranche-sur-Mer, France

This workshop gathered researchers interested in applying machine learning to identify and quantify functional traits of aquatic organisms from individual images.

April 24th

Talks

KEYNOTE - IMAGE PROCESSING IN INTERDISCIPLINARY PROJECTS: EXPERIENCE AND LESSON LEARNED

Cédric Pradalier
GeorgiaTech Lorraine-UMI 2958, GeorgiaTech-CNRS, Metz, France

Biology, among other research fields, is a large provider of image databases, quickly going over the available processing capabilities of the scientists collecting these images. On the other hand, image processing and automation are technology fields that can provide answers for the needs of biology, assuming these two can communicate together. Cédric Pradalier, a computer scientist and roboticist currently at GeorgiaTech Lorraine, has been active in interdisciplinary research for the last 10 years, with a specific focus on the application of computer vision and machine learning to environmental science problems. In this talk, he is discussing his experience applying machine vision in various applications based on detection, recognition, tracking and statistical quantification for (mostly) environmental sciences. He will conclude with insights and recommendations for successful interdisciplinary collaborations.

IMAGE-BASED, UNSUPERVISED ESTIMATION OF FISH SIZE FROM COMMERCIAL LANDINGS USING DEEP LEARNING

A. Alvarez, I. Catalán and M. Palmer
Instituto Mediterráneo de Estudios Avanzados, IMEDEA (CSIC–UIB), Esporles, Illes Balears, Spain

We present an example of technical tool for metamorphosing a huge volume of noisy information into useful ecological data. Modelling fish population dynamics is challenging in spite that accurate predictions are essential for providing proper scientific assessment on fishing quotas and fishery management. Unfortunately, scientific assessment has been demonstrated to be inaccurate in some cases, which save been attributed to flawed models inspired in poor understanding of the ecological processes behind. The dynamic of fish size distribution is one of the key inputs for feeding conventional fishery models. Nevertheless, in almost all fisheries, estimating size distribution of landed fish is still made by sand, thus sampling may be precise but it is certainly costly and inefficient, which in turns implies that sample size tends to be small and potentially biased. For the same reason, the temporal unit at which patterns can be described tend to be large. Deep learning offers the possibility for dramatically enlarge sampling size and reduce biases by analyzing images of all the fish cages landed in a given area. Here we present the results obtained in the PLEAMAR project FOTOPEIX (supported by Fundación Biodiversidad). A pretrained Mask-RCNN has been fed wits more than 2000 patterns of fish heads, from which fish size of many individuals in each cage can be estimated. Results show that deep learning techniques allow a segmentation of head fishes inside boxes wits a success index over 90%. This result has encouraged our fish ecology group to start processing all the video and image datasets wits deep learning.

AUTOMATIC IMAGING, CLASSIFICATION AND MORPHOMETRICS OF CALCAREOUS PLANKTON AND ONGOING DEVELOPMENTS

T. de Garidel-Thoron[1], R. Marchant[1], L. Beaufort[1], M. Tetard[1], B. Bourrel[1], Y. Gally[1], C. Chevalier[2]
1: CEREGE, CNRS, Université Aix-Marseille, Collège De France, IRD, INRA, Aix-en-Provence, France
2: MIO plankton group (MIO, Marseille, France)

The calcifying plankton (pteropods, planktonic foraminifera and coccolithophorids) are critical in the carbon fluxes to the deep ocean, and to the long-term inorganic carbon budget. A unique asset of those taxa is their preservation in deep sea sediments, yielding historical to geological archives of past plankton dynamics on timescales ranging from decades to millions of years. Since more than 20 years, we have at CEREGE developed tools to automatically image, classify and measure (weight, morphometrical parameters) those taxa. Here I will show the latest developments of our group. (i) recently, we have been working on the application of convolutional neural networks (using TensorFlow) to the classification of planktonic foraminifera (project FIRST funded by the ANR). This approach is now implemented in a high-throughput sorting machine of microfossils (MiSo, patent pending), and a couple of softwares have been developed to streamline the image annotation, neural network constructions, and morphometric analyses. (ii) We are currently expanding the application of those tools to pollens (PhD B. Bourrel), radiolarians (Post-Doc M. Tetard), lacustrine diatoms and zooplankton images (LISST-Holo, FlowCam, Zooscan…). The application to plankton images is developed in collaboration with the French MIO group (iii) for the coccolithophorids and coccoliths a suite of tools have been developed that help studying calcification during their culture (to follow the calcification of coccoliths in direct time) as well as in oceanic and fossil samples; neural networks that constitute the core of our pattern recognition system are now applied on 3D image of coccoliths.

AUTOMATED HIERARCHICAL CLASSIFICATION OF ANIMAL SPECIES IN CAMERA TRAP IMAGES

L. Hoebeke[1], *M. Stock*[1], *S. Van Hoey*[2], *J. Casaer*[2], *B. De Baets*[1]
1: Department of Data Analysis and Mathematical Modelling, Ghent University, Belgium
2: Research Institute for Nature and Forest, Belgium

Automatic imaging techniques, such as camera traps, are increasingly being used in biological monitoring. The great advantage of camera traps is that accurate data can be collected without animals being disturbed or researchers being present. However, such imaging frameworks produce high volumes of images, which often need to be reviewed and annotated manually. Convolutional neural networks, nowadays the go-to technique for computer vision problems, can be used to automate this labour-intensive process. The limited number of labelled camera trap images, combined with the difficulty of the classification task, does not allow for detailed classification of all species by the neural network. To overcome this problem, we incorporated hierarchical classification into the network. This way, the network can still reduce the manual workload, while misclassification is being avoided. Depending on the reliability of the classification, the level of detail can be adapted. The classification can be restricted to a higher level, for example family level, if there is insufficient information to classify that image to species level, while other images can still be classified in more detail using the same network. Finally, the network can automatically label images or provide suggestions to users when incorporated into annotation applications to speed up the annotation process. This method of hierarchically classifying camera trap images can easily be extended to other imaging data. This can be especially useful for plankton data, since it comprises a large variety of organisms, whereby detailed classification solely based on imaging data is not always possible.

AUTOMATIC MONITORING OF MARINE FISH USING UNDERWATER BAITED CAMERAS AND MACHINE LEARNING METHODS

A. Lana and J. Alós.
Instituto Mediterráneo de Estudios Avanzados, IMEDEA (CSIC–UIB), Esporles, Illes Balears, Spain

Determining the intensity and distribution of recreational fishing activities, beside the population fish distribution at marine habitats is a key to define its sustainability. A recent computer vision system has been developed to automatically monitor recreational fishing through the use of high-definition cameras to track the boats from the coast. In order to automatically monitor the fish populations at marine habitats, the use of underwater baited cameras to quantify the number of individuals, identification of species and size distribution is one of the most used methodologies. However, the high time-consuming evaluation of the videos or the high cost of the visual censuses with scuba diving, are limitations to monitor the fish species exploited by recreational fishing. The use of new technologies like deep learning can open an opportunity to take advantage of the availability of videos to make the correct map of fish population distribution. A pretrained model was used to identify the presence of fishes in the study area. The use of archived images to train deep learning models will be used to correctly identify fish species exploited by recreational fishing at the marine area of interest.

FORECASTING PLANKTON INTERACTION DYNAMICS USING MACHINE LEARNING AND NONPARAMETRIC ANALYSIS OF HIGH-DIMENSIONAL TRAIT DATA

T. Lorimer, S. Dennis, M. Reyes, C. Ebi, C. Albert and F. Pomati
Swiss Federal Institute of Aquatic Science and Technology (eawag)

Ecological theory suggests that interaction between species or functional groups is mediated by their proximity in trait space; but which trait space? Using a new dual-magnification automatic underwater microscope, we are monitoring relative abundances and physical appearances, in situ, of plankton across three trophic levels in a lake in Switzerland. These high-resolution in-situ image data provide a novel opportunity to define and examine a very large number of morphological traits simultaneously, and relate them to infer population dynamics. We will use a combination of CNN-based automatic species classification, novel graph-based nonparametric measures of high-dimensional trait proximity, and empirical dynamic modelling on trait-proximity time series, to investigate the predictive power of different trait spaces on population dynamics. At this early stage of the project, we will present an overview of our existing and planned workflow (including a custom plankton image labelling tool we have developed for CNN training, that may be of broader use to the community), and of our preliminary investigations into the non-convexity of natural data in high-dimensional trait spaces.

INSIGHTS FROM UVP IMAGES ACROSS THE SOUTH ATLANTIC

A. Rogge[1], *A. M. Waite*[1,2], *J. Karstensen*[3], *P. Prondzinsky*[1], *L. Guidi*[4]
1: Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany
2: Ocean Frontier Institute, Dalhousie University, Halifax, Canada
3: GEOMAR Helmholtz Centre for Ocean Research, Kiel, Germany
4: Laboratoire d'Océanographie de Villefranche, CNRS, Villefranche-sur-Mer, France

The South Atlantic Ocean is still a highly under-sampled region of the world oceans. Its basin is split into a western and eastern sector by the mid-Atlantic ridge and the biogeochemistry is mainly shaped by the oligotrophic South Atlantic Gyre. Thus, particle fluxes are usually thought to be low and strongly impacted by cross shelf transport. During a joint GO-SHIP cruise on board RV Maria S. Merian in early 2017 we performed >120 full-depth UVP profiles, following a transect at 34 °S through the whole basin, to survey the particle and organism distribution in this region. A cyclonic mesoscale eddy in the eastern basin mediated massive particle export caused by enhanced production due to nutrient fluxes driven by vertical inertial wave energy propagation at the eddy flanks, as well as enhanced flux of large particles due to the inertial chimney effect. This caused increased deep carbon flux, which was still threefold at > 3000 m compared to the surrounding ocean. Particle production also led to increased organism abundances in the upper 1000 m of the eddy flanks, i.e. copepods. Interestingly, we could identify an increased abundance of *Rhizaria* above the sea floor of the western basin. Most importantly, a depth and abundance-depending shift in the average particle grey scale was observed throughout the basin. The actual process behind this phenomenon is still ambiguous, but we concluded that a background pool of intensely reflecting particles exists in the Atlantic basin. Our results show up the need for full depth high definition UVP profiles to fully understand major oceanographic processes.

COMBINATION OF MACHINE LEARNING METHODS AND IMAGING-IN-FLOW SYSTEMS (FLOWCAM, CYTOSENSE/SUB) FOR PHYTOPLANKTON DETECTION, CLASSIFICATION AND ENUMERATION

G. Wacquet, A. Louchart, C. Blondel, P.-A. Hébert, E. Poisson-Caillault, F. Gomez, A. Lefebvre, N. Neaud-Masson, C. Belin, P. Grosjean, L. F. Artigas

Université de Lille, CNRS, ULCO, UMR 8187 LOG, Wimereux, France.

Improvements in automated data acquisition techniques have been carried out in order to characterize and quantify phytoplankton communities during oceanographic campaigns or in the frame of monitoring networks. However, these acquisitions and digitization techniques, including those concerning «imaging-in-flow» systems, still generate a more and more important quantity of data which cannot be processed manually in an acceptable time laps. Moreover, a full manual quantification of the particles based on a simple visual inspection can be tedious and consequently lead to erroneous or missing identifications. For this purpose, different semi-automated tools were and are still being developed to allow greater automation in data analysis and classification while permitting a limited user-interaction during the process. The combination of few expert knowledge and some machine learning algorithms can be applied at different levels: (i) to define and adapt a training set to the phytoplankton communities generally encountered in the studied area; (ii) to detect and partially validate the "most suspect" predictions which can represent until 90% of the global error; (iii) to automatically estimate the number of cells for each colonial form through the building of specific predictive models. The proposed methods were applied on the *in vivo* image datasets acquired with the FlowCam and CytoSense instruments during two cruises in the English Channel, in order to evaluate their operational ability to automatically monitor the diversity of samples and to allow the provision of various phytoplankton functional traits.

Posters

A NOVEL MULTIPLE-OBJECTS HIGH-RESOLUTION TRACKING SYSTEM TO MEASURE HUMAN SPATIAL BEHAVIOR IN COASTAL SOCIAL-ECOLOGICAL SYSTEMS

J. Alós, A. Lana

Instituto Mediterráneo de Estudios Avanzados, IMEDEA (CSIC–UIB), Esporles, Illes Balears, Spain

Successful marine spatial planning relies on understanding patterns of human use, with accurate, detailed, and up-to-date information about the human behaviour and space use. In commercial vessels, high resolution tracking through Vessel Monitoring System (VMS) or Automatic Identification System (AIS) have helped to maintain and enhance biodiversity of areas with high human impact through spatial management. Unfortunately, there is no regulation regarding location systems such as VMS or AIS for the recreational fleet. Obtaining spatial data of recreational activities like fishing or anchoring can be difficult and time-expensive given the widespread and variable nature of the fleet. Remote camera systems are increasingly used to overcome cost limitations of these conventional methods. Here we show a novel high-resolution tracking system based on video recording and state-of-the-art computer vision algorithms to obtain precise trajectories of vessels performing recreational activities in coastal areas. Our method allows determining the intensity and distribution of recreational activities, key to define the sustainability of the activity. Moreover, our new method produces a large amount of data composed by thousands of boats trajectories that can be used to test classical and novel ecological questions such as optimal search behaviour, collective movement or human decision-making in social-ecological systems.

FROM OPTICS TO EXPORT: IMAGING THE BIOLOGICAL CARBON PUMP

S. Giering

National Oceanography Centre, Southampton, UK

Sinking particles transport organic carbon to the deep sea, where they form the base of life. The magnitude of particle export and the rate at which particles are consumed (collectively called the 'biological carbon pump') determine carbon sequestration in the oceans and directly influence atmospheric carbon dioxide concentrations and global climate. Organic particle flux measurements from optical instruments, such as in-situ camera systems, are emerging as an important route of understanding the ocean carbon cycle. Optical instruments can be used from ships or installed on autonomous platforms, promising much greater spatial and temporal coverage than traditional techniques such as sediment traps. Yet, while technologies to image particles have advanced greatly during the last two decades, techniques to analyze the often immense datasets remain a challenge. One shortcoming is the translation of optical particle properties (e.g. the image) into particle characteristics such as particle type, carbon content and sinking velocity. Moreover, different devices often measure different optical properties, leading to difficulties in comparing results. Here I provide a practical overview of the challenges and potential of using optical instruments to understand particulate organic carbon fluxes.

AUTOMATIC DIATOM IDENTIFICATION USING A DEEP LEARNING APPROACH

M. Laviale[1], *C. Pradalier*[2]
1: Université de Lorraine, UMR CNRS 7360 LIEC, Metz, France
2: GeorgiaTech Lorraine-UMI 2958, GeorgiaTech-CNRS, Metz, France

Diatoms are ubiquist microalgae inhabiting all aquatic environments. They are widely used organisms in freshwater ecological assessment. However, the most efficient diatom-based diversity indices require a high level of taxonomic knowledge, which involves time and expert training. Diatoms are single-celled organisms which exhibit a highly ornamented silica shell (frustule) made up of two valves, which fit into each other like a Petri dish. Traditional taxonomic identification of these organisms is based on the observation of both shape and ornamentation of the frustule using conventional optical methods. These morphological features/traits are not always easy to characterize, depending on the experience of the operator, the quality of the image and the continuous evolution of diatom classification. Thus, the estimation of diatom-based diversity indices is often prone to multiple biases, implying that the development of a genuine tool is needed for supporting in decision-making in diatom identification. Automatic classification of diatoms has remained a challenge since the first attempts during the 90s, but the recent development of deep learning approaches are promising for solving image-based diatom classification problems. In this context, we aim at developing algorithms for an automatic identification of diatoms using image information derived from handcrafted and/or unsupervised morphological features. As a first attempt, this tool will be applied to the calculation of the Biological Diatom Index (IBD), which has been used routinely in surveillance networks since 2000 in France within the application of the European Water Framework Directive (WFD). On the long term, this tool could improve the efficiency of biomonitoring campaigns, in particular by allowing increased sampling effort.

THE ARTIFACZ PROJECT: MACHINE LEARNING APPLICATION TO THE IDENTIFICATION AND MEASUREMENT OF FUNCTIONAL TRAITS OF ZOOPLANKTON FROM INDIVIDUAL IMAGES

F. Maps[1], *D. Laurendeau*[2], *É. Debreuve*[3], *L. Guidi*[4], *S.-D. Ayata*[4], *J.-O. Irisson*[4]
1: Département de Biologie, Université Laval / UMI Takuvik CNRS – Université Laval, Québec, Canada
2: Département de genie éléctrique et logiciel, Université Laval, Québec, Canada
3: Laboratoire Informatique, Signaux et Systèmes de Sophia Antipolis, UCA / CNRS / Inria, France
4: Laboratoire d'Océanographie de Villefranche, CNRS, Villefranche-sur-Mer, France / Sorbonne Université, Paris, France

The Arctic is strongly impacted by the consequences of the ongoing climate changes, so much that this region is warming at a rate three times higher than elsewhere on the planet. One major consequence is the fast and irreversible retreat of its sea ice cover that will lead to profound and yet unpredictable perturbations of Arctic marine ecosystems. The base of Arctic marine ecosystems is supported by the growth of planktonic organisms that have adapted the extreme environmental conditions. These adaptations, often shared by many species, represent “functional traits” that influence the fitness of individuals and the ecosystem functioning as well. A better understanding of these traits appears crucial to predict the responses of marine ecosystems to the unprecedented changes affecting the Arctic Ocean. Several traits are associated to morphological features (e.g. size, egg sac, lipid stores, etc.), hence allowing to detect and measure them automatically from images. Imaging methods for plankton studies have multiplied and rapidly improved for the past decade. Thus, our main objective is to develop new tools combining imaging methods and machine learning algorithms in order to automatically measure important functional traits of planktonic organisms. This approach aims at a better understanding of the structure and functioning of Arctic marine ecosystems currently under pressure, but it will be possible to expand it to marine ecosystems that benefit from growing monitoring programs, including individual images of organisms. On the long run, functional traits automatically measured from individual images could be used as sentinels of ecological changes in the Arctic and all over the world.

IDENTIFY AND MEASURE A KEY MORPHOLOGICAL TRAIT FROM INDIVIDUAL IMAGES OF ARCTIC COPEPODS

A. Mercier[1], *D. Laurendeau*[1], *F. Maps*[2]
1: Département de genie éléctrique et logiciel, Université Laval, Québec, Canada
2: Département de Biologie, Université Laval / UMI Takuvik CNRS – Université Laval, Québec, Canada

Copepods are the linchpin of trophic networks and ecological services in Arctic marine ecosystems owing to their crucial role of matter and energy hub between the phytoplankton primary production and secondary producers, from fish larvae to marine mammals. In Arctic seas in particular, copepod species have evolved to reach large body sizes that allow for extensive lipid stores. These lipids are concentrated into oil sacs that the individuals mobilize to survive starvation during the harsh environmental conditions of the long winter time. Modern *in situ* imaging systems allow for thousands of individual images of high quality to be acquired at a high vertical and temporal resolution in a single vertical cast, in conjunction with environmental variables. A current bottleneck in the ability of marine ecologist to extract information from these images is the necessity for human identification of morphological traits, including the oil sac area (volume). We are testing a machine learning approach (Convolutional Neural Networks) to automatically identify and quantify the lipid content of large *Calanus* spp individuals sampled in the Baffin Bay, in order to enhance our understanding of plankton ecology and its impact on arctic marine ecosystems.

ZOOPLANKTON COMMUNITIES STRUCTURE FROM AN UPWELLING COASTAL ZONE TO AN OLIGOTROPHIC GYRE IN THE ALBORAN SEA (SW MEDITERRANEAN SEA)

N. Valcárcel Pérez
Instituto Espanol de Oceanografia (IEO), Fuengirola, Spain

The structure of biological communities in the Alboran Sea is strongly influenced by hydrodynamics. The entrance of the Atlantic Jet through the Strait of Gibraltar results in an oligotrophic anticyclonic gyre that dominates the western basin. During summer 2014, we sampled 10 stations from the Spanish coast towards the gyre, by day and night. Zooplankton was obtained by WP2 net (200 µm) vertical tows and fixed with buffered formalin. Scanned samples were processed with Zooprocess and imported to Ecotaxa to make the prediction and validation of the identification. Preliminary results show some shifts in the structure of the community between the coastal stations and the oligotrophic ones. Cladocerans tend to have high numbers in the gyre while Calanoid copepods and Appendicularians had lower relative abundance in those stations. However, as the Alboran Sea is a very productive area, our samples have high percentages of detritus (around 30% of our total vignettes), making difficult to achieve high precision in the identification of some groups. Further work is required to increase precision of our training set.

ARCTIC ZOOPLANKTON DISTRIBUTION IN BAFFIN BAY: TAXONOMIC AND MORPHOLOGICAL TRAIT-BASED APPROACHS

L. Vilgrain[1], S.-D. Ayata[1], J.-O. Irisson[1], F. Maps[2]
1: Laboratoire d'Océanographie de Villefranche, CNRS, Villefranche-sur-Mer, France / Sorbonne Université, Paris, France
2: Département de Biologie, Université Laval / UMI Takuvik CNRS– Université Laval, Québec, Canada

Functional traits are individual characteristics influencing an organism's fitness. They complement with quantitative information the usual taxonomic approach used to assess ecosystems' structure and functions. We studied zooplankton community in the surface waters of an Arctic sea during sea ice break-up. Morphological descriptors (area, darkness, complexity, etc.) were measured at the individual scale on about 28,000 copepod images taken by the Underwater Vision Profiler. A statistically defined morphological space allows to synthesize morphological information into interpretable continuous traits (i.e., size, transparency, appendages visibility). Traits spatial distribution revealed that large copepods are associated with ice-covered waters in the West while smaller individuals are present in open waters in the East. Copepods of the eastern part also seem to have higher feeding activity, as inferred by appendage visibility. High phytoplankton concentrations and probable strong visual predation pressure on larger copepods in well-lit open waters could be responsible for these traits' distributions, in addition to water masses circulation. Furthermore, copepods located at the ice edge appeared more opaque, presumably because of a full gut and/or a strong red pigmentation. Morphological trait-based approach revealed important ecological patterns that would have been inaccessible otherwise, including the role of copepod behaviour and ecological interactions in Arctic planktonic ecosystems.

April 25th

Talks

KEYNOTE – FROM IMAGE TO ECOSYSTEM FUNCTION: TRAIT-BASED APPROACHES TO OCEAN LIFE

Thomas Kiørboe
Centre for Ocean Life, DTU Aqua, Technical University of Denmark. Denmark

In situ images of plankton are typically used to identify species, but one may alternatively obtain trait information of individuals captured on images. Relevant traits include size, body plan, and characteristics that may inform about the foraging strategy of the organism. Such information is relevant because organisms are distributed in the ocean according to their traits, not their taxonomic affiliation. And the function of an ecosystem depends on the traits of the organisms making up the ecosystem, not on their taxonomy. Here, I will describe the trait-based approach focusing on traits that can be easily obtained from images. The approach consists of 4 elements, viz., (i) the identification of key traits and a quantification of their trade-offs for the main life forms in the ocean; (ii) the development of trait-based community and ecosystem models based on key traits and associated trade-offs; (iii) testing the models by comparing predicted trait distributions with those that can be observed in nature, e.g. from in situ imaging; and (iv) infer ecosystem key functions from observed or predicted trait distributions. Throughout, I will show examples to illustrate the approach and, hence, the entire avenue from image to assessment of ecosystem functions.

KEYNOTE - DEEP LEARNING FOR UNDERSTANDING THE IMAGE FORMATION PROCESS

Jean-François Lalonde
Département de genie éléctrique et logiciel, Université Laval, Québec, Canada

In the last decade, the renaissance of convolutional neural networks has enabled tremendous progress in various computer vision and image understanding tasks such as object detection and image segmentation. More recently, however, researchers have demonstrated that deep learning can also be used in inverse problems, which attempt to recover scene information (e.g. 3D geometry) by inverting the image formation process. Images are formed through a series of complex interactions between the lighting, the reflectance properties of surfaces, the geometry of objects present in the scene, and the camera capturing the image. Because the image formation process conflates these interactions together, recovering one (or several) of these components from images is a severely ill-posed problem. In this talk, I will describe novel approaches for estimating components of the image formation process (such as lighting, reflectance and geometry) from images. These approaches all have in common that they employ deep learning priors, trained on large quantities of specially designed data, to help constrain the inverse problems. The talk will be split into three main sections. First, I will present a new method that automatically estimates the lighting conditions of an outdoor image, simply from the image itself. The method yields robust estimates in all weather conditions, ranging from sunny to completely overcast. Second, I will introduce a method for tracking the position and orientation of 3D objects in 6 degrees of freedom. The approach is robust to occlusions, and runs in real time on commodity hardware. Third, I will (more briefly) present two recent methods that learn to model 1) spatially varying surface reflectance; and 2) the camera itself, with deep learning models. While very few mentions of plankton will be provided, it is the speaker's hope that these techniques inspire applications to marine biology!

A FIRST GLIMPSE INTO THE GLOBAL PATTERNS OF ZOOPLANKTON FUNCTIONAL DIVERSITY FROM THE TARA IMAGING DATASETS

F. Benedetti
Institute of Biogeochemistry and Pollutant Dynamics, ETH Zürich, Switzerland

Traits are characteristics of various natures (morphological/ physiological or behavioural) that mediate the performance of the fundamental Darwinian missions of organisms: survival, feeding and reproduction. Because traits are often difficult to measure at the level of the individual, patterns of functional diversity have usually been investigated through changes in the composition of species and their associated mean traits values. Therefore, intra-species and inter-individual variability in traits expression are often not accounted for when examining functional diversity patterns and drivers. This is a potential great limitation when studying systems such as the plankton where environmental variability is strongly coupled with intra-species variability in traits expression. For the first time, plankton imaging techniques applied to the samples of the TARA expeditions allowed to measure the morphometric traits of individual particles in an automated fashion, and investigate traits variations at the global scale. Here, we present a preliminary study of the spatial patterns of global zooplankton functional diversity from the TARA imaging datasets. Quantitative morphometric traits (size, biovolume and transparency) are coupled with categorical traits related to feeding and reproduction to ordinate zooplankton organisms in a global functional space. Based on the latter, convex hulls are used to quantify the functional volume occupied by the zooplankton community at each sampling station and to estimate functional diversity. The latitudinal pattern and the biological and environmental drivers of functional diversity are evidenced. Our results highlight the potential sensitivity of high latitude ecosystems to future environmental changes.

IMPROVING PLANKTON IMAGE CLASSIFICATION USING CONTEXT METADATA

J. S. Ellen[1], *C. Graff*[2], *M. D. Ohman*[3]
1: Department of Computer Science, University of California San Diego, USA
2: Department of Computer Science, University of California Irvine, USA
3: Scripps Institution of Oceanography, University of California, San Diego, USA

Deep Learning methods such as Convolutional Neural Networks (CNNs) show marked improvement over traditional feature-based supervised machine learning algorithms, but only consider the image pixels, ignoring any potentially diagnostic information regarding the environment where the image was obtained. We boost the performance of CNN classifiers by incorporating metadata of different types and illustrate how to assimilate metadata beyond simple concatenation. We utilize both geotemporal (e.g., sample depth, location, time of day) and hydrographic (e.g., temperature, salinity, chlorophyll-a) metadata and show that either type by itself, or both combined, can substantially reduce error rates. Incorporation of context metadata also boosts performance of the feature-based classifiers we evaluated: Random Forest, Extremely Randomized Trees, Gradient Boosted Classifier, Support Vector Machines, and Multilayer Perceptron. For our assessments, we use an original data set of 350,000 in situ images (roughly 50% marine snow and 50% non-snow sorted into 26 categories) from a novel in situ *Zooglider*. We document asymptotically increasing performance with more computationally intensive techniques, such as substantially deeper networks and data sets artificially augmented by rotation or offset. We find that metadata inclusion on our largest data set provides an improvement of the same order as augmentation, 1.5 points, but metadata add an insignificant amount of computational burden, whereas augmentation increases runtime by 5-8x. Combining augmentation with metadata boosts accuracy above either individually, and achieves 92.3 % classification accuracy with our 27-class dataset.

MACHINE LEARNING FOR BENTHIC IMAGERY PROCESSING: CAN WE ASPIRE TO MEASURE FUNCTIONAL TRAITS?

F. Ferrario
Québec-Océan / Université Laval, Québec, Canada

The advent of Autonomous Underwater Vehicles, self-propelled panoramic cameras, as well as simple diver-operated cameras, are flooding the benthic scientific community with data-rich imagery of the benthoscape. However, despite some machine learning (ML) solutions for benthic images exist, the automation of data extraction seems to be advancing at a different pace compared to its planktonic counterpart and facing a peculiar set of challenges. For example, a proper species identification is not always possible and organisms are classified using morphological or functional groups (e.g. "encrusting algae", "branched coral"). Moreover, adopted data extraction approaches are still often derived from sampling methods used by scientists in the field (e.g. quadrats, point-counts) and designed to return traditionally used variables (e.g. abundances, % covers). Consequently, the training of ML algorithms to quantify individual functional traits could be more complicate than for plankton. At the same time, photogrammetric and computer vision methods are being more frequently applied to benthic imagery to produce 3D models and orthoimagery of the bottom and biological communities. The opportunity to gather new layers of data (e.g. 3d structure and volumes) from these new products could drive the adoption of different data extraction approaches paving the way to individual functional traits estimation. In this contribution, I will present examples of benthic imagery and of current ML solutions for data extraction, discuss the need of different data extraction approaches given the new technological opportunities and discuss their potential for the estimation of the organism's functional traits.

GET IT FROM THE PICTURE: EXTENDING THE SCOPE OF PLANKTON AND PARTICLE IMAGING BEYOND DISTRIBUTION PATTERNS

K. O. Möller,
Institute of Coastal Research / Helmholtz-Zentrum Geesthacht, Germany

Advances in imaging technology have led to the development of new optical-imaging devices which have become widely used in plankton ecology and the observation of particle fluxes in the ocean. Underwater imaging is non-invasive and provides temporal and spatial high-resolution *in-situ* observations of plankton and particles and, hence, can overcome the limitations of traditional sampling techniques. There has been a considerable number of studies using imaging approaches for the observation of particle and plankton distribution patterns and abundances, yet. However, there is far more information that can be explored in the image content. We here present case studies where images reveal crucial information on behaviour (e.g. predator-prey behaviour, trophic interactions, diurnal vertical migration) as well as physiological state and population conditions from which certain rates can be inferred (e.g. lipid content and egg production rates of ovigerous copepods). Our examples highlight the potential significance of images beyond the scope of distribution patterns and how this might be used to gain information in the context of monitoring programs and ecosystem health assessment.

DISCERNING ZOOPLANKTON CHARACTERISTICS IN SITU FROM A NEW AUTONOMOUS *ZOOGLIDER*

M. D. Ohman
Scripps Institution of Oceanography, San Diego, USA

Conventional zooplankton sampling and fixation invariably distort morphological structures, alters opacity, induces shrinkage, and removes organisms from their 4-dimensional ocean context, making it difficult to infer morphological and behavioural traits of adaptive significance. A plethora of submersible imaging devices has proliferated in recent years, enabling in situ digital recording of organisms when such instruments are deployed from shipboard profilers, towed instruments, or fixed platforms or piers. We recently introduced a novel *Zooglider*, a fully autonomous vehicle that is

disconnected from surface vessels or bottom platforms. *Zooglider* dives repeatedly in the upper 400 m of the water column for mission durations potentially reaching 7 weeks. *Zooglider* employs a shadowgraph imaging system to image zooplankton and marine snow particles in situ with minimal hydrodynamic or visual disturbance to the surrounding environment. Acoustic backscatter, Chl-*a* fluorescence, and CTD measurements are made concurrently, providing important context information for the zooplankton images. With this minimally-invasive device we are able to record morphological structures and behaviours of potential significance to organismal fitness in situ, including body size, transparency, 3-D orientation, encounter volumes of feeding tentacles or rhizopods, morphological defence mechanisms, appendicularian house morphology, dormancy, copepod ovigerous/non-ovigerous state, and other traits.

STEP UP YOUR SORTING GAME! INCREASING EFFICIENCY WITH MACHINE LEARNING

S.-M. Schröder
Multimedia Information Processing Group, Department of Computer Science, Kiel University, Kiel, Germany

The increasing quantity of data brings the opportunity of new insights but also puts a strain on the people that process these data. The handling of millions of images is only feasible if it is carried out with the aid of pattern recognition solutions. The prevalent approach is to pre-sort all images using a classifier trained on the available classes and then correct for errors afterward. One obvious approach to increase efficiency is therefore the improvement of the classifier. However, sorting into a fixed set of categories neglects the fact that the distribution of the data at hand might differ from the training distribution (e.g. variability in number or functional traits). We developed a new way of sorting images. Using unsupervised clustering, we can shift the unit of work from individual images to groups of similar images. This increases consistency and multiplies the throughput. Starting from clusters ensures that the actual characteristics of the dataset are captured. Beyond that, operators are not restricted to a predefined list of categories. Rather, they can explore the dataset and adapt the granularity of their taxonomy to the respective application.

FROM BILLIONS OF IMAGES TO NEW INSIGHTS IN PLANKTON ECOLOGY

H. Sosik
Biology Department, Woods Hole Oceanographic Institution, Woods Hole, USA

Automated imaging of microscopic organisms such as enabled by Imaging FlowCytobot (IFCB), a submersible imaging-in-flow cytometer, provides an exceptional window into the ecology of marine plankton. Because this approach produces large data sets (billions of images already accumulated with ongoing observations at >10,000 images per hour), automated analysis is essential to realize the full potential. To address this challenge, we have developed a workflow for near real time generation of taxon-resolved assessments of plankton abundance, biomass and trait distributions from IFCB images. Essential steps include image segmentation to separate targets from background, extraction of geometric features and other traits for targets, target classification by supervised machine-learning, and estimation of abundance and biomass concentration across taxa and traits. Our strategy includes extensive evaluation of each step in the workflow and emphasizes open source solutions adopted by IFCB users worldwide, including providing open access to images and image products. In this presentation, I will highlight examples of products and their validation for characterizing event scale, seasonal, and multi-year patterns of change in plankton communities at the Martha's Vineyard Coastal Observatory.

ARTIFICIAL NEURAL NETWORKS FROM THE POINT OF VIEW OF NON-LINEAR DYNAMICS: LOCAL MINIMA AND CONVERGENCE PROBLEMS

A. Tsygvintsev
Unité de mathématiques pures et appliquées, ENS de Lyon, France

In this talk we discuss the fundamentals of artificial neural network theory from the mathematical point of view while underlying the strong link with theory of non-linear dynamic systems. We discuss some links with the qualitative theory of differential equations and introduce the overfly algorithm to tackle the local minima and convergence problem. Our approach is based on the existence of first integrals of the generalized gradient system with build-in dissipation.

Posters

IMAGING AND TRADITIONAL TAXONOMY: AN INTEGRATED APPROACH FOR COPEPOD BIODIVERSITY ON THE SOUTHWESTERN ATLANTIC OCEAN

É. C. Becker
PPGE-UFSC Universidade Federal de Santa Catarina, Santa Catarina, Brazil

In most oceanic ecosystems, copepods usually dominate for abundance, biomass and diversity in the mesozooplankton communities, performing a pivotal role in the pelagic food webs and biogeochemical cycles. Recent findings suggest a tight linkage between copepod functional traits and ecosystem functioning mainly related to feeding behaviour, spawning strategy and body size. The latter has been considered a master trait that influences several functions and is a major determinant of zooplankton ecological strategies. Latitudinal shifts in copepod diversity and size might explain how changes in biodiversity affect ecosystem functioning. Globally, the seasonal size response to temperature is negative in almost 90% of the copepod species, exhibiting a greater reduction in adult body mass in current-feeding calanoid species compared to ambush-feeding cyclopoids and, a decrease in the mean body size in the extratropical North Atlantic. Size changes in plankton communities driven by the increase of sea surface temperature is likely to become, in a near future, a major cause of ecosystem shift in the marine realm. Given that body size is an important predictor of species fitness and that warming is a prominent feature of climate change, a better evaluation of copepod intra- and interspecific variations of this trait should bring novel insights into the eco-evolutionary drivers of species biogeography and functional niches. Integrating the copepod size-species relationship offers complementary information to explore the intraspecific variability that is generally neglected in species-based studies and produce more effective linkages across methodological perspectives to predict functional diversity within communities.

IN SITU IMAGING IN AN UPWELLING SYSTEM - PAST, PRESENT AND FUTURE PERSPECTIVES

T. Biard[1,2], *M. Ohman*[2]
1: Université du Littoral; Laboratoire d’Océanologie et des Geosciences, CNRS-UMR LOG 8187; France
2: Scripps Institution of Oceanography; USA

Since the early ages of modern oceanography, collection of plankton sample has been mainly performed using traditional sampling methods (i.e., nets, bottles, etc.). Yet, a number of in-situ imaging systems have been made available within the last three decades, radically changing perspectives in plankton sampling. This technological shift has been underway since the late 2000s, within the *California Current Ecosystem Long Term Ecological Research* (CCE- LTER) program. CCE-LTER is located in the Southern California Current System, a productive eastern boundary current coastal upwelling ecosystem. In particular, application of such in situ imaging systems has led to a number of studies with various topics: (i) distribution patterns of particles and zooplankton across frontal systems with increased abundances at the front; (ii) ocean biogeochemistry, in particular silica fluxes to the mesopelagic and carbon flux attenuation by flux-feeding protists; or (iii) vertical niche distribution of fragile protists from the surface down to the mesopelagic ocean. These different types of application will be described briefly, highlighting the strong potential for multidisciplinary approaches with such in situ imaging system. Ultimately, we will discuss the perspectives offered by the release of new cameras designed for low power vectors, such as fully autonomous underway vehicles.

WHAT CAN WE LEARN FROM HIGH-RESOLUTION DISTRIBUTION DATA: ECOLOGICAL INFORMATION ON ARCTIC COPEPOD SPECIES

B. Niehoff
Alfred Wegener Institute for Polar and Marine, Bremerhaven, Germany

The Lightframe On-sight Key species Investigation system (LOKI) is a plankton recorder that provides high-resolution photographs (6 megapixel) of mesozooplankton organisms, taken in the water column during vertical hauls from 1000 m to the surface. A build-in image analysis program automatically recognizes objects in the pictures and saves the respective areas of interest as JPEG files, often more than 20,000 per cast, for later analyses. Synchronously, sensors measure depth, salinity, temperature, oxygen concentration and fluorescence. This allows to studying the small-scale zooplankton distribution patterns in relation to environmental conditions. The organisms on the images are annotated using the web-based application EcoTaxa, often to genus or species level and in some copepod species, also copepodite developmental stages (C) can be determined. Our data show that the youngest stages of *Metridia longa* (CI-III) reside in deep water layers in July in the Fram Strait, between Greenland and Svalbard. During development, this species seems to migrate upwards as the CIV and CV dominated the population in mid-water layers whereas adult females were mainly found close to the surface. In contrast, the developmental of *Calanus* spp. did not show stage-specific distribution patters. Continuous high-resolution optical measurements thus seem to allow for tackling spatial niches of species with yet unprecedented accuracy.

PROSPECTIVE IN THE USE OF MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE APPLIED TO THE ECOSYSTEMIC APPROACH TO FISHERIES AT IFREMER

J.-B. Romagnan & RT DEEP participants
Unité EMH, Ifremer Centre Atlantique, Nantes, France

Ifremer is France's leading actor for providing fisheries assessment and management advice to EU stakeholders and policy makers. Fisheries assessment and advice are based on multidisciplinary integrated surveys and studies linking ecology to socio-economic activities at the fisheries sector level. This is referred to as the Ecosystemic Approach to Fisheries (EAF). EAF data products, scientific and advisory outcomes originate from multiple and heterogeneous data collection and analytic processes. Recent years have seen the generalized deployment of (partially)-automated high throughput acquisition techniques (*i.e.* acoustic, imaging, remote sensing) and the development of complex models, to inform EAF data collection frameworks. Yet, EAF still heavily rely on human expertise for key variables (*i.e.* fish functional traits, assessment of catches on board professional vessels) or when automated techniques do not apply (*i.e.* bentho-demersal fishes). The discrepancies in data (*i.e.* spatio-temporal and biological resolutions, data quality and variability) make their combined analysis hard to achieve. Moreover, larger data fluxes generated by the automated collection often result in bottlenecks in the scrutinizing and analytic processes. Machine-Learning (ML) and Artificial intelligence (AI) are thought to be appropriate to address these Big Data related issues. In this context, Ifremer recently set up an internal network of scientists, engineers and technicians (RT DEEP) to evaluate current and future practises in the use of DL and AI tools in the framework of EAF. We propose to give a glimpse of what Ifremer and its academic and non-academic partners do and intend to do in the near future to promote the application of DL and AI to EAF.

GLOBAL STRUCTURE OF PLANKTONIC POPULATIONS USING QUANTITATIVE IMAGING METHODS

F. Lombard
Laboratoire d'Océanographie de Villefranche, CNRS, Villefranche-sur-Mer, France / Sorbonne Université, Paris, France

Imaging methods provide replicable, standardized and homogeneous ways to quantify diverse planktonic organisms on a routine way. However, a large heterogeneity still remains between devices, each of them targeting a given target in terms of size range and trophic/taxonomic type. Here we demonstrate that, with the use of Ecotaxa, heavily relying on recognition algorithms, we can provide a multi-device overview of the whole planktonic ecosystem, from sub-micron to centimetre sized organisms. Beyond classical abundance and biovolume of the various groups, imaging also gives access to many planktonic traits, either determined directly from images, or assessed from taxonomic identifications, and allows it quantification accordingly to spatial and temporal environmental constrains. The various possibilities of those approaches have been explored mainly on Tara-Ocean expedition but also on temporal observation in Villefranche-sur-Mer and shows, among others, that traits such as size but also functional and trophic structure of planktonic communities, are strongly related to environmental constrains.

PAIRWISE LEARNING TO PREDICT SPECIES INTERACTION NETWORKS

M. Stock
Faculty BioScience Engineering, Ghent University, Belgium

Plankton communities can be described as food webs, specific instances of species interaction networks. These networks, a collection of species as nodes and their interactions as edges, encode the structure of an ecosystem. We study how to use supervised machine learning tools to be able to predict new species interactions. Based on an observed network, we learn a function that takes as inputs the description of two species (e.g. traits, phylogenetic similarity or a morphological description) and predicts whether these two species are likely to interact or not. This framework for pairwise learning is based on kernels and similar methods have been highly successful for predicting molecular networks and for recommender systems, as used by companies such as Netflix and Amazon. During this workshop, we would like to explore how traits can automatically be extracted from (microscopy) imaging data, in order to predict trophic relations between species.

IMPACT FROM ZOOPLANKTON ON THE EFFICIENCY OF THE BIOLOGICAL CARBON PUMP FROM IN SITU OPTICS AND DIRECT VIDEO OBSERVATIONS

H. van der Jagt [1,2], I. Wiedmann [3], N. Hildebrandt [4], B. Niehoff [4], M. H. Iversen [1,2]
1: Helmholtz Young Investigator Group SEAPUMP, Alfred Wegener Institute for Polar and Marine Research, Bremerhaven, Germany
2: MARUM, Center for Marine Environmental Sciences, Bremen, Germany
3: UiT The Arctic University of Norway, Tromsø, Norway
4: Alfred Wegener Institute for Polar and Marine Research, Bremerhaven, Germany

Up to 95% of the primary production in the ocean is recycled within the upper few hundred meters of the water column and less than 1% of the organic matter produced in the surface ocean is exported to the sea floor. Marine snow and fecal pellets in the upper water column are often recycled at rates exceeding those measured for microbial degradation, suggesting zooplankton might be important for upper ocean flux attenuation. However, direct evidence of interactions with zooplankton and settling aggregates are still rare. We used in situ optics and direct video observations to investigate the role of zooplankton aggregate feeding in upper ocean flux attenuation. This was done from direct observations of aggregate ingestion rates and feeding behaviour on settling aggregates by the dominant sub-Arctic filter-feeding copepods *Calanus sp.* and *Pseudocalanus sp.* Both genera were observed to detect and feed on aggregates, but only *Pseudocalanus* seemed to actively and intentionally feed on settling aggregates while *Calanus* seemed to fragment aggregates and primarily ingest the small fragments. Using in situ zooplankton and aggregate abundances in combination with the measured aggregate feeding rates, we calculated that 60-67% of the observed flux attenuation at three subarctic locations could be explained by *Calanus* and *Pseudocalanus* aggregate feeding. When including microbial degradation, we were able to explain 77% of the observed flux attenuation. Our results reveal mesozooplankton as the key organisms for flux attenuation in the sub-Arctic, directly ingesting and fragmenting settling marine snow.

PRESENTERS

Josep Alós	alos@imedea.uib-csic.es
Amaya Alvarez Ellacuria	amaya@imedea.uib-csic.es
Érica Caroline Becker	ericabeckker@gmail.com
Fabio Benedetti	fabio.benedetti@usys.ethz.ch
Tristan Biard	tristan.biard@univ-littoral.fr
Thibault de Garidel-Thoron	garidel@cerege.fr
Jeff S. Ellen	jellen@ucsd.edu
Filippo Ferrario	filippo.f3rrario@gmail.com
Sarah Giering	s.giering@noc.ac.uk
Laura Hoebeke	laura.hoebeke@ugent.be
Morten Iversen	miversen@marum.de
Thomas Kiørboe	tk@aqua.dtu.dk
Jean-François Lalonde	jflalonde@gel.ulaval.ca
Arancha Lana	alana@imedea.uib-csic.es
Martin Laviale	martin.laviale@univ-lorraine.fr
Fabien Lombard	lombard@obs-vlfr.fr
Tom Lorimer	thomas.lorimer@eawag.ch
Frédéric Maps	frederic.maps@bio.ulaval.ca
Alexandra Mercier	alexandra.mercier.4@ulaval.ca
Klas Ove Möller	klas.moeller@hzg.de
Barbara Niehoff	barbara.niehoff@awi.de
Mark D. Ohman	mohman@ucsd.edu
Cédric Pradalier	cedric.pradalier@georgiatech-metz.fr
Jean-Baptiste Romagnan	jean.baptiste.romagnan@ifremer.fr
Simon-Martin Schröder	sms@informatik.uni-kiel.de
Heidi Sosik	hsosik@whoi.edu
Michiel Stock	michiel.stock@ugent.be
Alexei Tsygvintsev	alexei.tsygvintsev@ens-lyon.fr
Nerea Valcárcel Pérez	nerea.valcarcel@ieo.es
Laure Vilgrain	laure.vilgrain@obs-vlfr.fr
Guillaume Wacquet	guillaume.wacquet@univ-littoral.fr
Anya M. Waite	anya.waite@dal.ca

SUPORTING PARTNERS

Sentinelle Nord https://sentinellenord.ulaval.ca

Université Côte d'Azur http://univ-cotedazur.fr/fr

MODELIFE https://modelife.univ-cotedazur.fr/

Laboratoire d'Océanographie de Villefranche http://www.lov.obs-vlfr.fr/

IMEV https://www.imev-mer.fr/web/