WORKING GROUP ON MACHINE LEARNING IN MARINE SCIENCE (WGMLEARN; Outputs from 2021 meeting)

VOLUME 4 | ISSUE 15

ICES SCIENTIFIC REPORTS

RAPPORTS SCIENTIFIQUES DU CIEM

ICES INTERNATIONAL COUNCIL FOR THE EXPLORATION OF THE SEA
CIEM CONSEIL INTERNATIONAL POUR L’EXPLORATION DE LA MER

International Council for the Exploration of the Sea
Conseil International pour l'Exploration de la Mer

H.C. Andersens Boulevard 44-46
DK-1553 Copenhagen V
Denmark
Telephone (+45) 33 38 67 00
Telefax (+45) 33 93 42 15
www.ices.dk
info@ices.dk

ISSN number: 2618-1371

ICES Scientific Reports

Volume 4 | Issue 15

WORKING GROUP ON MACHINE LEARNING IN MARINE SCIENCE (WGMLEARN; Outputs from 2021 meeting)

Recommended format for purpose of citation:

ICES. 2022. Working group on machine learning in marine science (WGMLEARN; Outputs from 2021 meeting).
ICES Scientific Reports. 4:15. 16 pp. http://doi.org/10.17895/ices.pub.10060

Editors

Jean-Olivier Irisson • Ketil Malde

Authors

Oscar Beltran • Arjay Cayetano • Jose A. Fernandes-Salvador • Alessandra Gomes
Cédric Jamet • Rainer Kiko • Bernhard Kühn • Hassan Moustahfid • Klas Ove Möller
Dimitris Politikos • Jean-Baptiste Romagnan • Raphaëlle Sauzede • Vahid Seydi
Maria Sokolova • Jordan Watson

ICES CIEM
International Council for the Exploration of the Sea
Conseil International pour l'Exploration de la Mer

Contents

i Executive summary

The WGMLEARN group was formed to explore the use of machine learning in the marine sciences, and work towards increasing knowledge of and competence with relevant methods among marine scientists. The specific objectives were to review methods, applications, and implementations, to gather knowledge about them from a wide array of scientists, to address the implications of these methods for data management, and to highlight how they can be applied more/better in the future.

To achieve those objectives, we performed an extensive literature survey, gathering around 900 published works, and categorized them to extract trends in the usage of methods or data types. Based on this, we drafted three manuscripts.

The first describes the history of machine learning for marine ecology and highlights the dominance of images and acoustics as data sources, as well as the rise of deep learning methods. The second aims to guide new users towards these deep learning methods and, based on examples, shows their potential for a wide array of questions in marine sciences. The third focuses on approaches that are of particular relevance for fisheries science and shows that machine learning can be relevant at all scales of fisheries studies. Overall, we recognize a continued need to accelerate automation and effective data processing, and suggest new activities aimed at training, data management, infrastructure, and outreach, necessary to achieve this acceleration.

ii Expert group information

Expert group name	Working group on machine learning in marine science (WGMLEARN)
Expert group cycle	Multiannual fixed term
Year cycle started	2019
Reporting year in cycle	3/3
Chair(s)	Ketil Malde, Norway
	Jean-Olivier Irisson, France
Meeting venue(s) and dates	Oostende, Belgium, May 22-24, 2019 (25)
	Online, November 27-December 3, 2020 (30)
	Online, October 25-29, 2021 (22)

1 Introduction

Marine observation technologies and platforms are rapidly advancing and the volume of data collected is increasing by the day. The processing and analysis of this large data requires new approaches. Simultaneously, there is a need for better and more accurate models of complex systems in geophysics, ecology, and societal interactions around the ocean. Machine learning is a field at the intersection of computer science and statistics which provides a toolset that is rapidly becoming indispensable for scaling up automation of data processing.

Programs that learn from data are especially suited for complex cases where it is difficult to specify a precise solution. Instead, the model is shaped through examination of the data (typically adjusting parameters until a good solution is found). For traditional approaches, it is often necessary to first convert high-dimensional data to manually designed *features*, which can be used as inputs to the model. Designing appropriate features can be difficult, but with *deep learning* models, it is now possible to learn directly from even very complex raw data.

Machine learning techniques can be applied across *most domains*, and the topic should therefore be relevant to a number of *ICES working groups*. For example, we identified WGZE (Zooplankton Ecology) where identification and measurement of organisms from images is taking an increasing importance, WGFTFB (Fishing Technology and Fish Behaviour) for similar purposes as well as the monitoring of fishing gear performance, WGFAST (Fisheries Acoustics, Science and Technology) for automated detection and classification of targets in echograms (i.e. fish detection and classification), or WGSHIP (Shipping Impacts in the Marine Environment) for mapping and classification of human impacts.

One of our central deliverables is an extensive literature database, which we believe will be useful to the renewed bibliographic efforts of ICES.

2 Progress with respect to the ToRs

2.1 ToR a) Review new method developments in machine learning, current applications of machine learning methods in marine science, and their implementations and deployments in advisory and scientific processes

Machine learning methods have a long history of use in the marine sciences, as in many other fields. In order to develop an understanding of the current status in the field, WGMLEARN did an extensive survey of the scientific literature to identify relevant works. From this effort, we derived four deliverables: a literature database implemented in Zotero, and three scientific papers, which we aim to publish in peer reviewed journals.

2.1.1 Literature database

The first step to reviewing the current *applications of machine learning* in marine sciences was to *collect* them in a central place. Zotero is a free online service for literature database management, which also has clients for all major operating systems. It also features collaborative *group library* functionality, which we used to build a shared database. The papers were collected from online searches on public databases and complemented with works from the references of those papers. The search criteria were that work (i) is *peer reviewed*, (ii) applies a *machine learning* technique to (iii) a *marine dataset*. This effort uncovered ~900 references, with an additional 150 related and interesting works but that did not match all three criteria. The groups library is browsable at zotero.org: https://www.zotero.org/groups/2325748/wgmlearn/library

To extract information about trends and usage of machine learning among those references, they were *tagged* according to three main characteristics: the type of *data* (images, sounds, etc.), the machine learning *task* (regression, classification, etc.), and the machine learning *method* (support vector machines, neural networks, etc.). Additionally, *key papers* were identified, based on their influence in the field, the clarity with which they present a method etc.

This database highlighted a *dominance of images* and, to a lesser extent, acoustic records, in the type of data to which machine learning is applied in marine ecology, but a relative paucity of applications to omics data for example. For almost all data types and tasks, it also showcased a *shift towards deep learning* techniques.

2.1.2 Review of machine learning in marine ecology

Machine learning is a prominent tool in marine ecology and has advanced so rapidly in recent years that keeping up with its evolution has become a challenge. In addition, despite its importance, beginning to work with machine learning can still be a daunting task for a large proportion of marine scientists and ecologists. We sought to reduce this barrier by providing an *overview of most applied techniques* and the latest achievements in machine learning to marine ecology, as well as describing and providing examples of how this tool has been used thus far.

Drawing from the expertise in our working group, we included *sections on several data types* including imagery, acoustics (active and passive), omics, geolocation records, spectroscopy, remote sensing, time-series, and biogeochemical profiles. Within these data types, we further detailed applications. For example, we covered the use of machine learning in imagery for classification of nekton, plankton, benthos, and marine litter or for estimating bio-optical or biogeochemical parameters from remote sensing, while the use of machine learning associated with acoustics data were reported in works focusing on fishes, seabed and benthic community. Omics data processed with machine learning tools allowed the study of microbiomes at a detailed taxonomic level. At the macroscale, we examined applications of *integrated ecological and environmental datasets* for analysis of species and fishing fleet distributions, and how such landscape-level analyses can be used for decision support in conservation and management.

Broadly, we sought to compile a set of examples and resources that ecologists can add to their quiver of analytical tools. Combined with the database, the review paper provides not only a *primer on machine learning methods* but also on some of their potential applications in marine ecology. In addition, the review can introduce newcomers to machine learning for marine science, as well as driving scientists with different levels of knowledge on machine learning methods towards the most suited technique to their type of research problem.

2.1.3 A "guided tour" of deep learning in marine science

Over the last decade, *deep learning* has emerged, within machine learning, as a set of powerful technologies for data analysis. With sufficiently deep neural networks, containing several layers of "neurons", models can learn to extract information directly from high-dimensional raw data, such as images or echograms. Although we are starting to see adoption in marine domains, it is important to develop a better understanding of this technology among marine scientists.

To begin addressing this need, WGMLEARN has drafted a research paper which gives an *introduction* to the *methods* and algorithms involved, and *illustrates their applications* with published examples from the marine problem domains. Where possible, examples are chosen to highlight both typical use of a method, and also more unusual or creative applications.

Image analysis is perhaps the archetypical application of deep learning, and for tasks like species identification and abundance estimation, the user can adapt existing models with relative ease. Other areas are less developed, and for instance, active acoustics produce data in the form of regular arrays (tensors) suitable for deep learning methods, but are only beginning to see applications. The *lack of good training datasets* and the difficulty of reliable labeling are likely to be a factor.

2.1.4 Machine learning for fisheries

Fisheries science aims to understand and sustainably manage aquatic renewable resources and the fisheries they support. It relies on resource intensive tasks associated with analyses that are often completed in a *laborious* manner. The emergence of artificial intelligence systems holds great promise not only to *accelerate* current workflows that deal with the collection and analysis of samples, but also to *help linking different components* in their ecosystem context and gain deeper understanding of their dynamics.

To help fisheries scientists get a grasp of current machine learning applications in their field, WGMLEARN has drafted a research paper focusing on *current applications*, with an emphasis not only on the *advantages* they have over the traditional approaches but also their *limitations*, especially when it comes to operationality and potential robustness concerns. In a complementary manner to the “Review” and “Guided Tour” papers described above, we tried to organize the paper around thematic fields in the realm of fisheries science, rather than around specific data types or machine learning tasks, to make it easier for non-machine learning practitioners to identify their field of research.

The thematic fields of machine learning applications in fisheries were organized *from small to large scales* encompassing omics, individuals (catch items), aggregations of different species in situ (swarms), on-board processing, stock/populations assessment and dynamics, spatial mapping, fishing-related organizational units and ecosystem dynamics. The span of the applied methods varies from traditional statistical methods to data-specific approaches utilizing a higher level of semantics in the analysis. For example, machine learning tools applied to image analysis show great promise in accelerating routine work e.g. dealing with the analysis of biometric data. However, each field faces its own *challenges* ranging from pre-processing steps, the quality and extent of training data, and the need for proper model validation together with awareness of potential weaknesses of machine learning algorithms (e.g. extrapolation).

The paper concludes with a discussion of the effects that machine learning applications could have on management decisions and summarizes benefits and challenges associated with these methods in fisheries.

2.2 ToR b) Invite presentations and review data or analysis challenges in order to discuss possible methods, approaches and technologies

The initial composition of the group was rich in fisheries and general marine ecology scientists. To cover a wide range of methods, approaches, and technologies, we sought *presenters from various fields* in addition to these core members of the group. In particular, in the fields of omics, remote sensing, and terrestrial ecology. This highlighted *convergence of tools* across fields but also a *disparate advancement of the application of machine learning* among various topics. For example, machine learning based methods have reached an operational state for images of some organisms (plankton, fish in the context of electronic monitoring, etc.) but remain more explanatory for other data types, such as active acoustics records or ocean color remote sensing.

Notably, many invited *presenters then became chair-invited members*, enriching and balancing the topics that the group could cover.

Another important challenge for machine learning is its deployment at scale and its application for the advisory process, or decision making in general. To address those questions, speakers were invited to present *strategic initiatives* in the EU (European Marine Board) and the US (NOAA Artificial Intelligence Strategy). The overall conclusion is that *machine learning* (often under the moniker of "AI") is considered *strategic at many levels*, and that large *structuring initiatives* are being set up but that the field is highly dynamic and *rapidly evolving*.

2.3 ToR c) Communicate with DIG and the ICES Data Centre on data organization and requirements related to machine learning analysis

For machine learning to percolate through the ICES processes, the data that it requires needs to be accessible and structured appropriately. To anticipate how this can be done at ICES' scale, discussions with the Data and Information Group and the ICES Data Center are necessary.

A presentation of the ICES data and services raised awareness among group members about the large amount of data that ICES maintains and makes accessible through Application Programming Interfaces. Information about ICES' Data and Information Group was also provided by its former chair, highlighting that internal discussions are going on within ICES regarding whether the members states, rather than ICES, should be responsible for the storage of the raw data; in which case ICES would only store the derived data necessary for its advisory process. This would impact the applicability of some machine learning methods, in particular deep learning, which typically works directly on the raw data.

Yet, these interactions were too limited to allow significantly linking WGMLEARN and the data handling groups within ICES and such an effort should continue.

2.4 ToR d) Summarize current and future needs in marine science and identify how machine learning methods can provide solutions. Work actively to promote adoption of relevant technologies

Much of the group effort has been *dedicated to reviewing* the current uses of machine learning in marine sciences. The deliverables of the group focus on this, with some short recommendations for the future in the conclusions. However, our scientific papers also highlight where methodologies mature enough to be used operationally and provide the basis for policy decisions, for example: machine-based indicators for the Marine Strategy Framework Directive (MSFD), modelling for the objectives of the Common Fisheries Policy (CFP) or United Nations Sustainable Development Goals towards sustainable use of marine resources (e.g. 12, 13 and 14).

In addition, to promote work done in machine learning and encourage its adoption by the community, members of the working group organized a session at the ICES Annual Science Conference in 2019, and are currently planning sessions at ASC (jointly with WGFAST) and at the ASLO Ocean Sciences Meeting 2022.

Furthermore, the thinking process and experience leading to those deliverables allowed us to *gain an overview of the field* and *identify domains where more work is needed* for machine learning to be adopted and provide solutions for marine science.

2.4.1 Capacity building

As mentioned, machine learning is developing very quickly and is becoming a very useful tool for data analysis. However, the methods are often difficult to use without a background in mathematics or statistics. To promote and to develop the use of machine learning, we propose to develop *training courses* on machine learning for marine sciences. Some of these courses could be dedicated to PhD students and postdoctoral researchers in marine science, without a background in machine learning. The courses should be provided by experts in machine learning and marine scientists that are machine learning practitioners. The courses should encompass the *basics of machine learning* and should also contain *practical exercises* where participants address problems and use data relevant to marine sciences. Other courses could also be dedicated to trained computer scientists, to introduce them to the data and problems of marine ecology. As the number of attendees is often limited by practical constraints, the courses could also be *recorded* and made available through a dedicated website to reach a larger audience.

Another gap to fill is to establish *efficient communication between machine learning experts and marine sciences experts*. It can take a long time to understand the vocabulary of the other domain and to explain how to interpret data and ask the right questions. For example, for a computer scientist, a datum used as the input for a machine learning model can be a number, an image, a succession of values and its most important aspects will be its structure and shape. For the same datum, the first focus of a marine scientist will be on its meaning, units and range of values, regardless of its structure. However, both aspects are important for successful and insightful applications of machine learning to marine data. This mutual understanding can be achieved through training courses for both sides, as outlined above, but also through working groups such as WGMLEARN, or joint sessions at international conferences.

Training efforts could also target a *wider audience*, and include participants from Asia, Africa, South America, and the South Pacific. A few existing programmes are dedicated to knowledge transfer to developing countries: (e.g. POGO-SCOR Fellowship, SCOR visiting scholars, IOC oceanteacher) and ICES could coordinate its training with them.

2.4.2 Data Sharing

To support the development of machine learning algorithms, *data* should be *freely accessible* and archived in a consistent manner, including *metadata*. Generally, all research data (and in particular the raw data) should be archived according to the *FAIR principles*[1], i.e. Findable, Accessible, Interoperable, and Reusable. For supervised learning (e.g. most cases of species classification from images) training data with ground-truth labels are also needed. These need to be provided as well, and, if possible, labelling uncertainties (e.g. discrepancy between labellers) should be documented. Ideally, developed code should also be freely accessible, to allow *reproducibility* and *reuse*. One challenge in this regard is competing interests, where data sharing can be perceived as losing a competitive advantage.

[1] https://www.go-fair.org/fair-principles/

Such *qualified reference datasets are still rare*, although there are some examples: the quality-controlled dataset of Argo and Biogeochemical-Argo profiles data[2]; the high-quality nutrient data collected over the last 30 years that compose the GLODAPv2[3] database. Even if *efforts* are made by the community to create such reference data, this should be *expanded* to more databases, to increase their usability, especially for machine learning applications. For this, this working group should develop guidance for the community, such as the definition of *what constitutes AI-'ready' datasets* and best practices to create them. A first step could be to provide a *listing* of (or even host) such datasets, as available for the Weather and Climate Sciences[4] for example. Furthermore, tools need to be developed to enable multi-operator annotations and to address inter-operator differences in data labelling.

However, in order to promote data sharing, researchers should be given incentives, expert support, training, and the infrastructure to make it easy to share data, and therefore, worth their while. Governments, funders, research institutions, libraries, and publishers all have a role to play to unlock the huge potential of research data.

2.4.3 Infrastructure

The development of machine learning applications, and in particular of deep learning, sometimes require (i) a large amount of training data, and (ii) a significant computational power to run on.

For the proof-of-concept developments of algorithms as part of fundamental research, local computing resources can be sufficient, e.g. an off-the-shelf workstation with a dozen Central Processing Unit (CPU) cores and one Graphic Processing Unit (GPU). This is well within the reach of most research laboratories or even single individuals with enough technical knowhow.

However, to *deploy* these algorithms in operational conditions, to handle the ever-growing inflow of observational data (from remote sensing and emerging in situ automated sensors) in quasi-real time, the computing resources need to be scaled up and this requires *significant infrastructure*. Such infrastructure consists of separated and high-capacity CPU, GPU, and storage, linked through high-performance network connections. Their set up and maintenance is costly and complex; they are not within the reach of any single research laboratory. Numerous *efforts are ongoing* to set up such infrastructures, for example the Extreme Science and Engineering Discovery Environment (XSEDE) by National Science Foundation (NSF) , the Big Data Program by National Oceanic and Atmospheric Administration (NOAA) (which is becoming the NOAA Open Data Dissemination program), in the United States (US) ; the European Grid Infrastructure (EGI), the European Open Science Cloud (EOSC) or, within it, the Blue Cloud project and the associated D4Science infrastructure in the European Union (EU). There is a notable *difference in strategy* between these two sets of initiatives. The US is developing *centralized* resources, handled by federal agencies, fostering and leveraging *collaboration* for research and development in artificial intelligence *with large technology companies* that have large stakes in cloud computing (e.g. Amazon, Microsoft, Google, NVIDIA). The EU is developing *federations* of national resources, both in terms of computing and of data handling, with *less reliance* on (often US-based) *technology companies*. A

---

[2] ftp://ftp.ifremer.fr/ifremer/argo
[3] https://www.glodap.info/
[4] http://mldata.pangeo.io/

commonality of these initiatives is that they require *collaboration*, at national or international level, with corporations or between agencies, etc. Because of their scale, they cannot be achieved without these collaborations or partnerships.

Even with such infrastructures in place, a significant challenge remains for practitioners of machine learning: how to *run the same model* on the *local computer* on which it was developed, a university-level *data center*, a *cloud computing* service but also onboard a relatively "disconnected" research vessel. This requires standardization of processes/workflows and data. For processes, a clear trend, to be encouraged, is towards *containers* (e.g. Docker, Kubernetes). This is unfortunately not a technology that marine scientists are familiar with and could become part of the necessary training mentioned above.

2.4.4 Innovation and industry outreach

Scientific and government institutions that engage in scientific and advisory activities depend on industry to provide equipment and services, and offshore and coastal industry both depend on and affect the ecosystems. To stimulate innovation and to better target both technology development and research, *good communication channels between scientists and industry* are crucial. It is our experience that there is considerable interest in machine learning from industry, and it would be mutually beneficial if shared venues or channels could be established. To attract industry, traditional scientific *communication* and dissemination may *need to be augmented* with higher weight placed on innovation, application, and technology development. Increasingly, the outputs of scientific research —especially those involving machine learning— are not only standalone analyses or manuscripts, but data or analytical products that are linked to automation, and operationalization. Such evolution requires that *academic* and *government* scientists *partner with industry* experts that are poised to facilitate such evolving scientific demands.

Public funding is increasingly geared towards *projects that have direct consequences for industry*. In the EU, these are related to the umbrella term of the "Blue Economy". Such projects need to involve industrial partners from the start, to define common needs and goals, and many could involve machine learning. One example, among several, of relevance to ICES is the SusTunTech project which aims to improve the energy efficiency of tuna fishing vessels through a partnership between research, industry and fishing companies. Vessel monitoring and Copernicus data are combined through machine learning to improve the detection of fish distribution, reduce time at sea, and save fuel. The industrial partners are integrating the algorithms developed into commercial products.

2.5 New legislation

In light of the technological advances in artificial intelligence and the challenges it creates, in April 2021, the European Commission adopted a *proposal for an Artificial Intelligence Act*[5] (AIA) harmonizing the rules on the development and the use of AI technologies in the EU. One target of this act is the use of artificial intelligence to achieve sustainable development goals while

[5] Proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM/2021/206 final), Brussels, European Commission.

preserving human safety, in particular relative to human-robot collaboration, autonomous machines or privacy. However, a related communication[6] also highlights important potential benefits such as increased citizen engagement initiatives and new employment opportunities outweighing potential job losses. The AIA *covers many techniques* (machine learning but also logic-based approaches, statistical estimation, etc.). It *defines "artificial intelligence* systems" quite *loosely* as software that has the "ability, for a given set of human-defined objectives, to generate outputs such as content, predictions, recommendations, or decisions which influence the environment with which the system interacts". Therefore, it defines artificial intelligence as a domain even broader than what the group considered.

Overall, WGMLEARN members have *not had the opportunity to discuss* the consequences of such recent regulations for scientific research in, and applications of, artificial intelligence. This can be a future objective of the group.

---

[6] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions - Fostering a European approach to Artificial Intelligence (COM/2021/205 final), Brussels, European Commission.

3 Conclusion

Machine learning has made substantial inroads into the marine sciences, and in particular, an extensive scientific literature exists applying or developing methods for application to a variety of datasets. However, its operational deployment is still sparse compared with traditional methodologies. Synergies between different disciplines and with industry seem like the way forward to increase its operationalization. As data continues to be collected in ever larger volumes, these efforts must continue to expand to address current and future data processing and knowledge extraction needs. ICES can play a central role in developing the knowledge and capability needed for continued development. Moving from scientific works to operational systems is a major challenge, which faces technical, organizational, social, economic, and legal obstacles. These hurdles have so far only been outlined, and ICES may want to initiate work towards developing a better understanding both of the challenges and of the possible solutions.

Annex 1: List of participants

2019

Name	Institute	Country (of institute)	Email
Alessandra Gomes	University of São Paulo	Brazil	colombo.gomes@gmail.com
Bernhard Kühn	Thuenen Institute	Germany	bernhard.kuehn@thuenen.de
Cédric Jamet	University of the Littoral Opal Coast	France	cedric.jamet@univ-littoral.fr
David Currie	Marine Institute	Ireland	David.Currie@Marine.ie
Edwin van Helmond	Wageningen University	Netherlands	edwin.vanhelmond@wur.nl
Esha Mohamed	Swedish University of Agricultural Sciences	Sweden	esha.mohamed@slu.se
Jean-Baptiste Romagnan	Ifremer	France	jean.baptiste.romagnan@ifremer.fr
Jean-Olivier Irisson	Villefranche Oceanographic Laboratory	France	irisson@normalesup.org
Jenni Fincham	Centre for Environment, Fisheries and Aquaculture Science	UK	jenni.fincham@cefas.co.uk
Jose Fernandes	Plymouth Marine Laboratory	UK	jfernandes@azti.es
Julien Simon	Ifremer	France	julien.simon@ifremer.fr
Justin Defever	The Flanders Research Institute for Agriculture	Belgium	justin.defever@ilvo.vlaanderen.be
Ketil Malde	Institute of Marine Research	Norway	ketil@malde.org
Laura Hoebeke	Ghent University	Belgium	laura.hoebeke@ugent.be
Laura Uusitalo	Finnish Environment Institute	Finland	laura.uusitalo@syke.fi
Linh Nguyen	Wageningen University & Research	Netherlands	linh.nguyen@wur.nl
Madiop Lo	Observatoire des Sciences de l'Univers	France	lomadiop4@gmail.com
Michiel Stock	Ghent University	Belgium	michiel.stock@ugent.be
Rainer Kiko	Villefranche Oceanographic Laboratory	France	rainer.kiko@obs-vlfr.fr

Raphaelle Sauzede	Villefranche Oceano-graphic Laboratory	France	raphaelle.sauzede@imev-mer.fr
Rubens Lopes	University of São Paulo	Brazil	rubens@usp.br
Sam Vanhoorne	The Flanders Research Institute for Agriculture, Fisheries and Food	Belgium	sam.vanhoorne@ilvo.vlaanderen.be
Sven Kupschus		UK	sven@kupschus.net
Tristan Cordier	University of Geneva	Switzerland	tristan.cordier@unige.ch
William Michaels		USA	WMichaels001@gmail.com

2020

Name	**Institute**	**Country (of institute)**	**Email**
Alessandra Gomes	University of São Paulo	Brazil	colombo.gomes@gmail.com
Andreas Hermann	Thünen-Institute of Baltic Sea Fisheries	Germany	andreas.hermann@thuenen.de
Arjay Cayetano	Thünen-Institute of Sea Fisheries	Germany	arjay.cayetano@thuenen.de
Bernhard Kühn	Thünen-Institute of Sea Fisheries	Germany	bernhard.kuehn@thuenen.de
Cédric Jamet	University of the Littoral Opal Coast	France	cedric.jamet@univ-littoral.fr
Dimitris Politikos	Hellenic Centre for Marine Research	Greece	dimpolitik@gmail.com
Hassan Moustahfid	United States National Oceanic and Atmospheric Administration	United States	Hassan.Moustahfid@noaa.gov
Hong Minh Le	Royal Belgian Institute of Natural Sciences	Belgium	hmle@naturalsciences.be
Jean-Olivier Irisson	Villefranche Oceano-graphic Laboratory	France	irisson@normalesup.org
Jenni Raitoharju	Tampere University	Finland	jenni.raitoharju@tuni.fi
Jens Rasmussen	*ICES*	*UK*	*J.Rasmussen@marlab.ac.uk*
Jose Fernandes	Plymouth Marine Laboratory	UK	jfernandes@azti.es
Kerry Howell	University of Plymouth	UK	kerry.howell@plymouth.ac.uk
Ketil Malde	Institute of Marine Research	Norway	ketil@malde.org

Klas Ove Möller	Helmholtz-Zentrum Heroen Institute of Coastal Research	Germany	klas.moeller@hereon.de
Laura Uusitalo	Finnish Environment Institute	Finland	laura.uusitalo@syke.fi
Laure Van Den Bulcke	The Flanders Research Institute for Agriculture, Fisheries and Food	Belgium	Laure.VanDenBulcke@ilvo.vlaanderen.be
Maria Sokolova	DTU Aqua, National Institute of Aquatic Resources	Denmark	msok@aqua.dtu.dk
Nils Olav Handegard	Institute of Marine Research	Norway	nilsolav@hi.no
Peter Rubbens	Flanders Marine Institute	Belgium	peter.rubbens@vliz.be
Pia Schuchert	Agri-food and Biosciences Institute	UK	pia.schuchert@afbini.gov.uk
Rainer Kiko	Villefranche Oceanographic Laboratory	France	rainer.kiko@obs-vlfr.fr
Rubens Lopes	University of São Paulo	Brazil	rubens@usp.br
Samuel Chaffron	University of Nantes	France	samuel.chaffron@univ-nantes.fr
Sheila Heymans	Scottish Association for Marine Science	UK	sheymans@marineboard.eu
Simon-Martin Schröder	Christian-Albrechts-University of Kiel	Germany	sms@informatik.uni-kiel.de
Stanislovas Jonusas	European Commission Directorate-General for Maritime Affairs and Fisheries	Belgium	Stanislovas.jonusas@ec.europa.eu
Sven Kupschus		UK	sven@kupschus.net
Tristan Cordier	University of Geneva	Switzerland	tristan.cordier@unige.ch
William Michaels		USA	WMichaels001@gmail.com

2021

Name	**Institute**	**Country (of institute)**	**Email**
Alessandra Gomes	Univ São Paulo	Brazil	colombo.gomes@gmail.com
Andreas Hermann	Thuenen Inst.	Germany	andreas.hermann@thuenen.de
Arjay Cayetano	Thuenen Inst.	Germany	arjay.cayetano@thuenen.de
Bernhard Kühn	Thuenen Inst.	Germany	bernhard.kuehn@thuenen.de
Cédric Jamet	ULCO	France	cedric.jamet@univ-littoral.fr
Clea Parcerisas	VLIZ	Belgium	clea.parcerisas@vliz.be
Hassan Moustahfid	NOAA	USA	hassan.moustahfid@noaa.gov
Jean-Olivier Irisson	Sorbonne Université	France	irisson@normalesup.org
Jenni Fincham	Cefas	UK	jenni.fincham@cefas.co.uk
Jens Rasmussen	*ICES*	*UK*	*J.Rasmussen@marlab.ac.uk*
Jordan Watson	NOAA	USA	jordan.watson@noaa.gov
Jose Fernandes	AZTI	Spain	jfernandes@azti.es
Ketil Malde	Inst of Marine Res.	Norway	ketil@malde.org
Klas Ove Möller	Hereon	Germany	klas.moeller@hereon.de
Maria Sokolova	DTU	Denmark	msok@aqua.dtu.dk
Mikael Overgård	Swedish University of Agricultural Sciences	Sweden	mikael.ovegard@slu.se
Mohab Badreldin	AASTMT	Egypt	mohab.mahmoud13@aast.edu
Oscar Beltran	IOW	Germany	oscar.beltran@io-warnemuende.de
Pia Shuchert	Queen's University	UK	pia.schuchert1@gmail.com
Rainer Kiko	Sorbonne Université	France	rainer.kiko@imev-mer.fr
Raphaelle Sauzede	CNRS	France	raphaelle.sauzede@imev-mer.fr
Rubens Lopes	Univ São Paulo	Brazil	rubens@usp.br
Vahid Seydi	Bangor University	UK	vahidseydi@gmail.com

Annex 2: Resolutions

2018/MA2/EOSG06

A **Working group on machine learning in marine science (WGMLEARN)**, chaired by Ketil Malde, Norway, and Jean-Olivier Irisson, France. The group will work on ToRs and generate deliverables as listed in the Table below.

	Meeting dates	Venue	Reporting details	Comments (change in Chair, etc.)
Year 2019	22-24 May	Ostend, Belgium	Interim report by 1 July, 2019	
Year 2020	1-2 December	Online meeting	Interim report by 14 January, 2021	
Year 2021	25-26 and 28-29 October	Online meeting	Final report by 10 December, 2021	

ToR descriptors

ToR	Description	Background	Science Plan codes	Duration	Expected Deliverables
a	Review 1) new method developments in machine learning, 2) current applications of machine learning methods in marine science, and 3) their implementations and deployments in advisory and scientific processes.	Machine learning holds great potential, but it is necessary for practitioners to keep up with new developments and to gain an understanding of the opportunities and challenges with new methods.	4.1, 4.5, 3.2	1, 2, 3	On-line (live) report
b	Invite presentations (externally and internally) and review data or analysis challenges in order to discuss possible methods, approaches and technologies.	ML experts need to meet with stakeholders and data collection efforts for mutual understanding of data analysis challenges.	4.2, 4.3	1, 2, 3	On-line list of challenges
c	Communicate with DIG and the ICES Data Centre on data organization and requirements related to machine learning analysis.	For effective deployment, ML has to be integrated with data collection and data management efforts.	4.2	1, 2, 3	
d	Summarize current and future needs in marine science and identify how machine learning methods can provide solutions. Work actively to promote adoption of relevant technologies.	Future developments in the marine sciences, including project proposals, need to have an informed and up to date view of the state-of-the-art, in order to make optimal use of the technology.	4.2, 4.3	3	

Summary of the Work Plan

Year 1	Produce the annual overview of recent developments.
Year 2	Produce the annual overview of recent developments.
Year 3	Produce the annual overview of recent developments.

Supporting information

Priority	Machine learning is a prioritized topic by DIG, and was explored in the WKM-LEARN workshop in April 2018, on an initiative by ACOM. The workshop highlighted a need for a centrally organized venue to share methods and best practices between researchers, to attract outside expertise, and to support publication and dissemination of results. Long term engagement is especially needed to support deployment and integration of the new methods.
Resource requirements	The research programmes which provide the main input to this group are already underway, and resources are already committed. The additional resource required to undertake additional activities in the framework of this group is negligible.
Participants	Machine learning is a topic of considerable and broad interest, and is likely to attract participants from outside the traditional ICES organization. We expect some 30 members, similar to the attendance of the WKMLEARN workshop.
Secretariat facilities	None.
Financial	No financial implications.
Linkages to ACOM and groups under ACOM	DIG (Julie could you check does DIG sit under ACOM?, certainly they go to the SCICOM meetings), ICES Data Centre (also I think this sits under the secretariat rather than ACOM), could just be moved to the section below if we are not sure
Linkages to other committees or groups	Close working relationships with other groups that target data collection or analysis. Relevant examples are: WGFTFB (targets non-destructive fisheries sampling) WGNEPS (video surveys to monitor nephrops populations) WGFAST (analysis of acoustics data) WGBIOP and WGSMART A planned WG for electronic monitoring of vessels
Linkages to other organizations	Machine learning is a prioritized topic by DIG, and was explored in the WKM-LEARN workshop in April 2018, on an initiative by ACOM. The workshop highlighted a need for a centrally organized venue to share methods and best practices between researchers, to attract outside expertise, and to support publication and dissemination of results. Long term engagement is especially needed to support deployment and integration of the new methods.