Machine learning in marine ecology: an overview of techniques and applications

Peter Rubbens, Stephanie Brodie, Tristan Cordier, Diogo Destro Barcellos, Paul Devos, Jose A Fernandes-Salvador, Jennifer I Fincham, Alessandra Gomes, Nils Olav Handegard, Kerry Howell, et al.

ICES Journal of Marine Science, 2023, **80**, 1829–1853
DOI: 10.1093/icesjms/fsad100
Advance access publication date: 3 August 2023
Review Article

Machine learning in marine ecology: an overview of techniques and applications

Peter Rubbens[1,2], Stephanie Brodie[3], Tristan Cordier[4,5], Diogo Destro Barcellos[6], Paul Devos[7], Jose A. Fernandes-Salvador[8], Jennifer I. Fincham[9], Alessandra Gomes[6], Nils Olav Handegard[10], Kerry Howell[11], Cédric Jamet[12], Kyrre Heldal Kartveit[10], Hassan Moustahfid[13], Clea Parcerisas[1,7], Dimitris Politikos[14], Raphaëlle Sauzède[15], Maria Sokolova[16], Laura Uusitalo[17,18], Laure Van den Bulcke[19,20], Aloysius T. M. van Helmond[21], Jordan T. Watson[22,23], Heather Welch[3], Oscar Beltran-Perez[24], Samuel Chaffron[25,26], David S. Greenberg[27], Bernhard Kühn[28], Rainer Kiko[29,30], Madiop Lo[31], Rubens M. Lopes[6], Klas Ove Möller[32], William Michaels[33], Ahmet Pala[34,10], Jean-Baptiste Romagnan[35], Pia Schuchert[36], Vahid Seydi[37], Sebastian Villasante[38], Ketil Malde[10,39], and Jean-Olivier Irisson[29,*]

[1]Flanders Marine Institute (VLIZ), 8400 Oostende, Belgium
[2]Kytos BV, Technologiepark-Zwijnaarde 82, 9052 Gent, Belgium
[3]Institute of Marine Science , University of California Santa Cruz, Santa Cruz, CA 95064, USA
[4]Department of Genetics and Evolution, University of Geneva, 1205 Geneva, Switzerland
[5]NORCE Climate, NORCE Norwegian Research Centre AS, Bjerknes Centre for Climate Research, Jahnebakken 5, 5007 Bergen, Norway
[6]Oceanographic Institute, University of São Paulo, Praça do Oceanográfico, 191, 05508-120, São Paulo, Brazil
[7]Department of Information Technology, Research group WAVES, Ghent University, Tech Lane Ghent Science Park, 126, B-9058 Gent, Belgium
[8]AZTI, Marine Research, Basque Research and Technology Alliance (BRTA). Txatxarramendi Ugartea z/g, 48395 Sukarrieta, Spain
[9]Cefas, Pakefield Road, Lowestoft, Suffolk NR33 0HT, UK
[10]Institute of Marine Research, Nykirkekaien 1, 5005 Bergen, Norway
[11]School of Biological and Marine Sciences, University of Plymouth, Drake Circus, Plymouth PL4 8AA, UK
[12]Université du Littoral Côte d'Opale, CNRS, Univ. Lille, IRD, UMR 8187, LOG, Laboratoire d'Océanologie et de Géosciences, F-62930 Wimereux, France
[13]National Oceanic and Atmospheric Administration, US Integrated Ocean Observing System, Silver Spring, MD 20910, USA
[14]Institute of Marine Biological Resources and Inland, Hellenic Centre for Marine Research, 16452 Argyroupoli, Greece
[15]Sorbonne Université, CNRS, Institut de la Mer de Villefranche, FR3761, F-06230 Villefranche-Sur-Mer, France
[16]Wageningen University and Research, Droevendaalsesteeg 1, Building 107, 6708 PB Wageningen, The Netherlands
[17]Finnish Environment Institute, Latokartanonkaari 11, FI-00790 Helsinki, Finland
[18]Natural Resources Institute Finland (Luke), Latokartanonkaari 9, FI-00790 Helsinki, Finland
[19]Flanders Research Institute for Agriculture, Fisheries and Food, Marine Research, Jacobsenstraat 1, 8400 Ostend, Belgium
[20]Department of Data Analysis and Mathematical Modelling—Knowledge-based Systems Research Group, University of Ghent, Coupure Links 653, 9000 Gent, Belgium
[21]Wageningen University and Research, Wageningen Marine Research, 1976 CP IJmuiden, The Netherlands
[22]Present affiliation: Pacific Islands Ocean Observing System, University of Hawai'i at Mānoa, 1680 East West Road, POST 815, Honolulu HI 96822, USA
[23]Auke Bay Laboratory, National Oceanic and Atmospheric Administration, 17609 Pt. Lena Loop Rd., Juneau, AK 99801, USA
[24]Leibniz Institute for Baltic Sea Research Warnemünde (IOW), Seestrasse 15, 18119 Rostock, Germany
[25]Nantes Université, École Centrale Nantes, CNRS, LS2N, UMR 6004, F-44000 Nantes, France
[26]Research Federation for the Study of Global Ocean Systems Ecology and Evolution, FR2022/Tara Oceans GOSEE, F-75016 Paris, France
[27]Institute of Coastal Systems, Helmholtz-Zentrum Hereon, Max-Planck-Straße 1, 21502 Geesthacht, Germany
[28]Johann Heinrich von Thünen Institute of Sea Fisheries, Herwigstraße 31, 27572 Bremerhaven, Germany
[29]Sorbonne Université, CNRS, Laboratoire d'Océanographie de Villefranche, LOV, F-06230 Villefranche-sur-Mer, France
[30]GEOMAR Helmholtz Centre for Ocean Research Kiel, 24148 Kiel, Germany
[31]Aix Marseille Univ., Univ. Toulon, CNRS, IRD, Mediterranean Institute of Oceanography, F-13009 Marseille, France
[32]Institute of Carbon Cycles, Helmholtz-Zentrum Hereon, Max-Planck-Straße 1, 21502 Geesthacht, Germany
[33]NOAA, National Marine Fisheries Service, Office of Science and Technology, Silver Spring, MD 20910, USA

Received: 29 September 2022; **Revised:** 14 April 2023; **Accepted:** 26 May 2023

[34]Department of Mathematics, University of Bergen, Allégaten 41, 5007 Bergen, Norway
[35]DECOD (Ecosystem Dynamics and Sustainability), IFREMER, INRAe, Institut-Agro-Agrocampus Ouest, rue de L'île d'Yeu, 44311 Nantes Cedex 3, France
[36]Agri-Food & Biosciences Institute (AFBI), Environment and Marine Sciences Division, 18a Newforge Lane, Belfast BT9 5PX, UK
[37]Centre for Applied Marine Science, Bangor University, Menai Bridge LL59 5AB, UK
[38]EqualSea Lab-Cross-Research in Environmental Technologies (CRETUS), Department of Applied Economics, University of Santiago de Compostela, Santiago de Compostela 15782, Spain
[39]Department of Informatics, University of Bergen, Allégaten 41, 5007 Bergen, Norway
* *Corresponding author: tel: +33 (0)4 93 76 38 04;* e-mail: irisson@normalesup.org.

Machine learning covers a large set of algorithms that can be trained to identify patterns in data. Thanks to the increase in the amount of data and computing power available, it has become pervasive across scientific disciplines. We first highlight why machine learning is needed in marine ecology. Then we provide a quick primer on machine learning techniques and vocabulary. We built a database of ~1000 publications that implement such techniques to analyse marine ecology data. For various data types (images, optical spectra, acoustics, omics, geolocations, biogeochemical profiles, and satellite imagery), we present a historical perspective on applications that proved influential, can serve as templates for new work, or represent the diversity of approaches. Then, we illustrate how machine learning can be used to better understand ecological systems, by combining various sources of marine data. Through this coverage of the literature, we demonstrate an increase in the proportion of marine ecology studies that use machine learning, the pervasiveness of images as a data source, the dominance of machine learning for classification-type problems, and a shift towards deep learning for all data types. This overview is meant to guide researchers who wish to apply machine learning methods to their marine datasets.

Keywords: acoustics, ecology, image, machine learning, omics, profiles, remote sensing, review

What is machine learning and why does marine ecology need it?

The term "machine learning" (ML) has become omnipresent in both the scientific literature and everyday news. Its first use dates back to the late 1950s: Regarding a game of checkers, Arthur Samuel, an electrical engineer at IBM, stated that "a computer can be programmed so that it will learn to play a better game of checkers than can be played by the person who wrote the program" by using so-called "machine-learning procedures" (Samuel, 1959, p. 219). In its broadest definition, an ML system improves its performance by extracting information from data (Mitchell, 1997). In contrast to traditional computer programs, which encode a solution designed by the programmer, an ML system can learn to solve a task without being provided an explicit recipe. Instead, the task is learned by providing the system with examples, i.e. data. The ability to produce a solution to a problem that is not representable mechanistically can be extremely powerful, but it depends crucially on selecting an appropriate representation of the problem (an "objective" function) and on having adequate data from which to learn.

Although often used interchangeably in popular literature, ML is a subdomain of the larger field of artificial intelligence (AI), which encompasses knowledge representation, logic models, algorithms, and computational methods capable of intelligent behaviour (Figure 1). Within ML, the subfield of deep learning (DL; LeCun *et al.*, 2015) has advanced rapidly over the last decade. DL systems use large neural networks (Table 1) to extract relevant features from raw data and learn from them, instead of requiring explicit engineering of those features. These data are often complex (such as images or sounds) and big (thousands to millions of records). In this review, we cover ML, therefore, including DL.

The success of ML is associated with the increase in computational power over the last 20 years (Mitchell, 1999), but also with the increasing volume of available data (Jordan and Mitchell, 2015), which led to the development of a broader diversity of algorithms, implemented in widely available software. Scientists from many disciplines outside of computer science are now actively applying ML methods, and marine sciences are no exception, as exemplified by a recent themed set in this journal (Beyan and Browman, 2020). Most examples in this themed set actually relate to ecological questions, within which a central focus is the detection and quantification of the abundance and distribution of living organisms. ML is promising in marine ecology for several reasons. (i) Modern instruments produce large volumes of data (Tanhua *et al.*, 2019; Guidi *et al.*, 2020) that require scaling up their processing; the flexibility and adaptability of ML methods make them a natural choice for such automation. (ii) This automation can also help to reduce the biases necessarily introduced by manual processing (e.g. Culverhouse *et al.*, 2003), hence improving reproducibility. (iii) Finally, ML is adept at handling high degrees of uncertainty (i.e. dealing with noise in the data) associated with unknown underlying mechanisms or with non-stationary processes; therefore, they often yield high predictive power (Baker *et al.*, 2018) and are increasingly used to gain an understanding of ecological processes (Lucas, 2020).

Within marine sciences, ML has been used more extensively in some subdomains. Specialized reviews have already covered some applications. For example, Liu and Weisberg (2011) reviewed the use of Self-Organizing Maps in Oceanography, Culverhouse *et al.* (2006), Benfield *et al.* (2007), and Irisson *et al.* (2022) reviewed ML techniques for the taxonomic

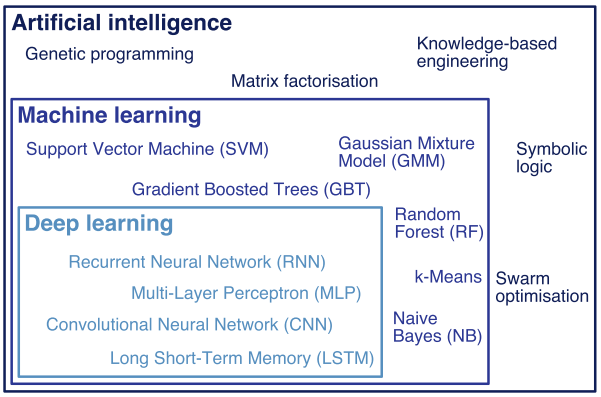

Figure 1. Deep learning is a subdomain of machine learning, which on its own is a subdomain of artificial intelligence, as illustrated. Specific methods are mentioned in each subdomain.

Table 1. Definitions of machine learning algorithms commonly used in marine ecology studies and cited in this review.

Method	Description
Decision tree (DT)	A hierarchy ("tree") of successive decision criteria based on the input variables, in order to label data instances. Popular implementations include the classification and regression trees and C4.5 algorithms.
Random forest (RF)	An ensemble method that combines predictions of multiple decision trees, in which each tree is trained on a bootstrap resample of the data.
Gradient boosted trees, boosted regression trees (BRTs)	An ensemble method that combines decision trees, each working on the residuals of the previous one. Gradient boosting in general is a smart way of combining multiple "weak" learners.
Matrix factorization	Method to find a representation of the input data in fewer variables by decomposing the original data matrix into two latent matrices of fewer dimensions.
k-nearest neighbours	New data points are labelled according to the average/majority label of its k-nearest neighbours. The value of k must be set beforehand.
Linear discriminant analysis (LDA)	A multivariate Gaussian distribution is fitted to the inputs for each class, in which each distribution has its own mean but a shared covariance matrix. New data instances are assigned to the distribution with the highest conditional probability.
Support vector machine (SVM)	An algorithm that tries to find an optimally separating linear boundary in a large transformed space of the input variables.
Naive Bayes (NB)	Classifier that combines Bayes' theorem with the "naive" assumption that all variables are independent from each other; a type of Bayesian network.
Bayesian network	A directed acyclic graph in which each vertex (node) has a probability distribution or a conditional distribution, conditional on the value of its parents.
Gaussian mixture model (GMM)	A distribution (i.e. "mixture") of multiple normal distributions is fitted to the data. Data points are then assigned to the closest normal distribution.
k-means	Data instances are clustered into k groups, in which the within-cluster variance is minimized.
Artificial neural network (ANN)	An algorithm that combines layers of nodes or "neurons". Each neuron receives as input a weighted linear combination of the outputs of neurons in the previous layer. Nodes in the first layer represent the input variables. Weights are updated incrementally using training data, starting from the last layer. Predictions are done by feeding new data through the layers once the weights are set.
Self-Organizing Map (SOM)	An unsupervised artificial neural network, in which nodes in a grid are optimized to match with groups of similar data points.
Convolutional neural network (CNN)	A class of methods within deep learning. A convolutional neural network takes an array as input, and performs convolutions and reductions (pooling) on it to extract features, which are then fed to an artificial neural network.
Region-based convolutional neural network (R-CNN)	R-CNN is a deep learning architecture designed to recognize, localize (bounding box), and classify multiple objects in an image. Mask R-CNN is a variant that is able to recognize, segment, and classify individual pixels to multiple objects in an image.
Deep belief network (DBN)	A form of deep learning in which a generative graphical model is composed out of multiple layers of latent variables. Layers are connected with each other, but the nodes within a layer are not.
Long short-term memory (LSTM)	A type of recurrent neural networks that falls under deep learning. These types of networks are used to predict sequences of data. LSTMs provide feedback connections within the network, while many deep learning architectures only provide feedforward connections.

classification of plankton images, Reichstein *et al.* (2019) gave an overview of DL for earth sciences—including oceanic applications, and Malde *et al.* (2020) provided a brief review on recent developments in DL and highlighted both opportunities and challenges for its adoption in marine sciences.

For researchers whose expertise is outside of computer sciences, ensuring a proper application of ML methods and keeping track of new developments is challenging. The aim of the present review is to serve as a resource for marine ecologists who want to apply ML to their own data. To that effect, the section "A quick primer on machine learning" serves as a starting point for non-practitioners and introduces relevant vocabulary. The section "The setup of the database and its tags" describes our survey of the literature and the resulting structured database, on which the rest of the review is built. From it, we identified that ML is used at two stages in ecological research: (i) to process the raw data collected and extract ecologically meaningful datasets from it and then (ii) to combine these ecology-ready datasets together, and with others, to improve our understanding of ecological systems. Therefore, the section "Machine learning to extract ecological information from observational data" describes applications where ML was used to generate ecological datasets from various raw data types: images and video, optical spectra of single cells, acoustics, omics, geolocation records, and ocean colour imagery and biogeochemical profiles. The section "Machine learning to improve ecological understanding" describes how ML can be used to gain knowledge on the relationships between species and their environment (section "Predicting species abundance and distribution"), among species (section "Capturing dynamic ecological relationships"), and between us, humans, and marine ecosystems (sections "Summarizing ecosystems through regionalization" and "Supporting human decisions on ecosystem management"). Finally, the section "Discussion and perspectives" concludes on the commonalities among ML applications, suggests what is currently limiting them in ecology, and gives a general outlook of the field.

A quick primer on machine learning

In this section, we provide a short overview of the different tasks that ML can achieve, the overall process that ML studies go through in the context of marine ecology, and then present different ML algorithms and software tools that implement them. Interested readers are invited to consult classic texts to deepen their understanding, either in an introductory manner (James *et al.*, 2013) or in a more mathematically oriented one (Friedman *et al.*, 2001).

ML approaches are often divided between supervised and unsupervised. Supervised systems are given a set of input data points and their corresponding output (measurements or labels assigned by experts). The output is often called the target or response variable. In this case, an ML system learns the mapping from the input variables to the output variable (e.g. predict fish diversity from environmental variables; Smoliński and Radtke, 2017). Supervised systems can further be divided into classification, where the output is categorical and the task is to assign a class to input data (e.g. classify plankton taxa from images; Gorsky *et al.*, 2010), and regression, where the output is continuous or at least ordered (e.g. predict nutrient concentrations from hydrological variables; Sauzède *et al.*, 2017). A supervised task relevant to marine ecology is object detection: The ML system locates objects of interest in a form of regression, often of their bounding box (e.g. detect benthic organisms in images of the seafloor; Liu and Wang, 2021). Finally, sometimes, the target variable is only available for a subset of data points, a situation called semi-supervised learning.

Unsupervised systems are given input data only and search for patterns without the availability of a target variable. For instance, unsupervised methods can aim to cluster data points together based on a definition of similarity (e.g. define distinct bioregions based on community compositions; Sonnewald *et al.*, 2020), to define simpler representations for the data while retaining salient properties, also known as dimensionality reduction (e.g. represent correlations between environmental variables through the first two dimensions of a principal component analysis; Zhao and Costello, 2019), or to construct a model for the distribution of the data (e.g. produce a smooth map of the density of active fishing vessels from point records; Kroodsma *et al.*, 2018).

Additional steps can be performed before or within an ML pipeline. An important part of many ML systems is the preprocessing (e.g. feature normalization or smoothing) of input variables, in order to make them as relevant as possible. Feature extraction derives new informative variables from initial, raw ones (e.g. automated extraction of measurements from an image; Hu and Davis, 2005). Feature selection eliminates less relevant variables, either to improve performance or to gain explainability thanks to a simpler system (e.g. removal of correlated variables; Thomas *et al.*, 2018). Finally, the covariance structure in the input variables can be used to impute missing values, which are common in field-collected data, or detect outliers, i.e. values that go beyond the expected range of covariance (e.g. a dissolved oxygen concentration too high given the temperature of the water).

The general process for tackling an ML task is shown in Figure 2, and the successive steps are described in its caption. Of course, depending on the approach and study case, some steps will be modified. For example, target variables are not available in the case of unsupervised learning (step 2). In DL, feature extraction is included in the model (step 4). In many situations, cross-validation is used in lieu of a dedicated validation set (step 6): The training set is split into subsets, the model is trained on all subsets but one, and validated on this remaining one; this process is repeated until each subset has been held out once. Comparisons with an external dataset (step 8), although important, are rarely performed due to the lack of such independent data. Finally, many ML models are never deployed (step 9), but serve to describe and understand a particular dataset.

Diverse ML algorithms have been developed to solve a large variety of tasks. In Table 1, we provide a brief description of those commonly used in marine ecology publications.

Finally, several open-source software libraries implement many ML methods under a consistent interface. Thus, once one understands the general process (as highlighted in Figure 2), exploring various methods is relatively easy and progress can be quick. The better known libraries of relevance for marine ecology are scikit-learn (https://scikit-learn.org/; Pedregosa *et al.*, 2011) and, more recently, TensorFlow (https://www.tensorflow.org; Abadi *et al.*, 2016) and PyTorch (https://pytorch.org; Paszke *et al.*, 2019) in Python, the tidymodels collection of packages in R (https://www.tidymodels.org/; Kuhn and Wickham, 2020), Flux (https://fluxml.ai/) in Julia, or Weka in Java (https://www.cs.waikato.ac.nz/ml/weka/).

The setup of the database and its tags

As a basis for this paper, we built a database of literature references covering the application of ML methods to marine data (supplemented by a few additional works, outside of this scope, but providing context and cited in this review). In its broadest definition, ML covers a wide array of methods and data types. Because many methods have been applied in marine ecology, it is extremely challenging to make an exhaustive inventory. Therefore, the goal of this database is instead to showcase the diversity of ML applications to marine data. To do so, multiple keyword-based searches in various scholar databases were performed by the authors, including the keywords "machine learning", "marine", "ecology", and variations thereof. The results were complemented with the personal libraries of the authors, who span a range of specialties. This (already large) nucleus of papers was further grown using the references cited within them, starting from the most recent and going backwards in time. This procedure was iterative (the references of the newly added papers being also examined) and the search was stopped after several rounds of such tentative additions did not yield any new reference.

After assembling this large body of potentially relevant literature, the authors screened the suitability of each paper for its inclusion in the database according to the following criteria: (i) the paper is peer-reviewed, (ii) its "Methods" section describes the ML approach used, and (iii) it applies it to a marine dataset. Because some classical statistical techniques can be perceived as ML, we further reduced the scope to studies that follow the general process of Figure 2 (i.e. include a validation and/or a test dataset). Then, papers were organized through tags, defining the type of data they analyse ("data:∗" tag), ML tasks achieved ("task:∗"), the algorithms used ("method:∗"), and other useful characteristics (e.g. availability of code and/or data; "meta:∗"). The content of the resulting database, organized according to data type, is summarized in Figure 3.

This selection process still yielded over 1000 papers, which cannot all be described in this review. To decide which ones to cite, we considered the following additional criteria: the paper (i) has been widely adopted by the research community (e.g. is cited very often, defines a method widely applied), (ii) is easily reproducible because its methodology is well-described and/or code and data are publicly available, or (iii) is representative of a body of work not covered by criteria (i) or (ii).

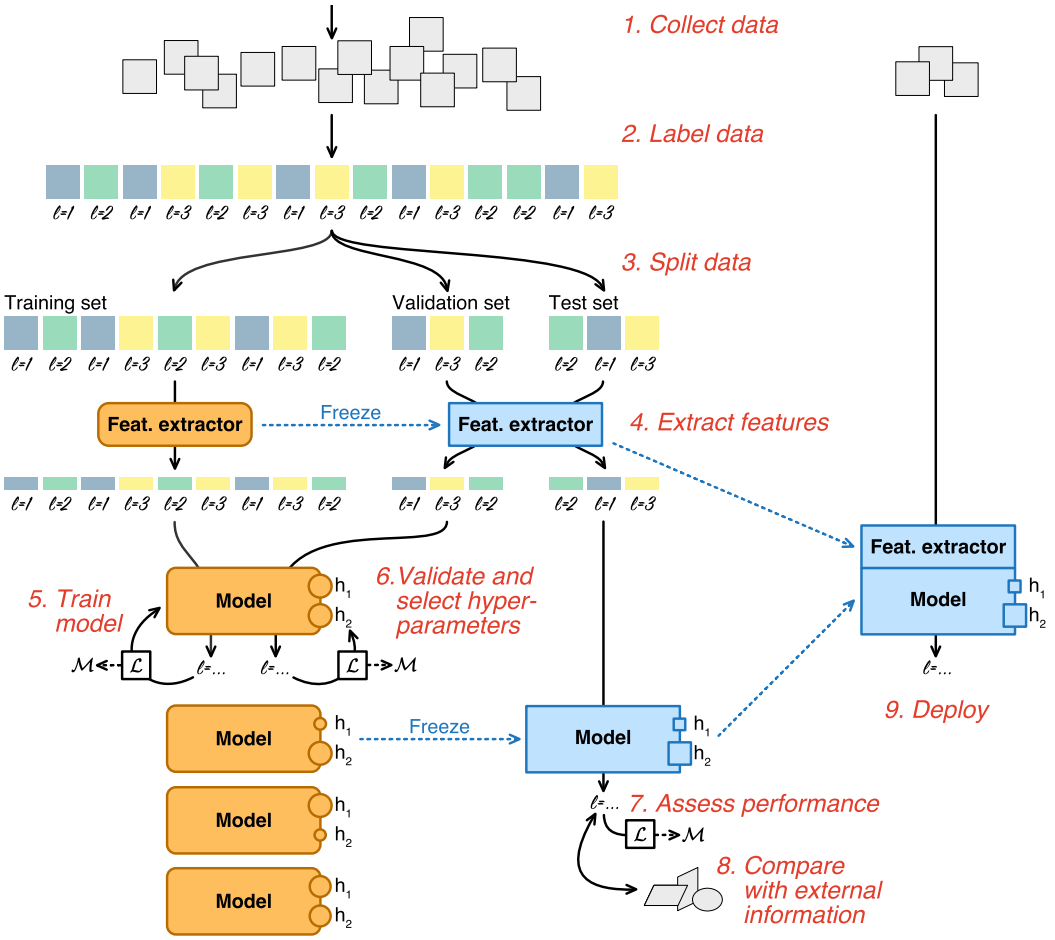

Figure 2. The general process of (supervised) machine learning. After being collected (1), data need to be labelled (2), which means associating the inputs with a number or a name as output ($l = 1$, 2, or 3 in the example). The data are then split into training, validation, and test datasets (3) while taking its structure into account (e.g. ensure that all labels are represented in each dataset). Each input in the training set can be summarized into features (4). The (transformed) training set is used to train the model (5), by minimizing a loss function (L) that computes the value of one or several performance metrics (M). The validation set undergoes the same transformation as the training set, if any, and is then used to evaluate the predictive performance of the model, ideally with the same metric(s) (6). Several versions of the model can be trained with different hyperparameters (i.e. settings, noted h_*) of the machine learning system, and the one with the best performance on the validation set is retained. At this point, the model is frozen and its final performance is assessed on the test set (7). If external information, different from the original data, is available, it should be used to ensure that model predictions are reasonable, in addition to achieving a given performance (8). Finally, the model is ready to be deployed and used with newly collected data (9).

While this database is not exhaustive, the methodical approach described above should avoid overt biases and large omissions. We therefore consider it representative of the diversity of approaches and of the relative volume of research in various domains. More importantly, we hope its use will become continuously maintained and updated by its users. To do so, users can browse the library online (https://www.zotero.org/groups/2325748/wgmlearn/library) and, if they wish to contribute, register to the WGMLEARN Zotero group (https://www.zotero.org/groups/2325748/wgmlearn/), indicating what their contribution would be. The library in its state at the time of submission is available as Supplementary Material (S1).

Machine learning to extract ecological information from observational data

A first set of extensive and successful applications of ML is the processing of raw inputs (images, sounds, sequences, etc.) into ecologically meaningful data, often in the form of tables with samples (locations, times, etc.) as rows and variables (taxa densities, biogeochemical quantities, etc.) as columns.

Quantifying marine objects from images and video

Methods to segment and classify objects of interest from images or video are not sensitive to whether the object is a fish, a bird, or a piece of plastic debris. Yet, the processing of this dominant (Figure 3) type of data has a long history that is often siloed within specific communities, sometimes with reason. For example, object segmentation is very different for benthic objects lying over a complex background than for pelagic ones, imaged over a rather uniform background. Therefore, the literature is presented separately for benthos, marine macrolitter, nekton, and plankton. The commonalities among the methods used for these data, and others, are highlighted in the "Discussion and perspectives" section.

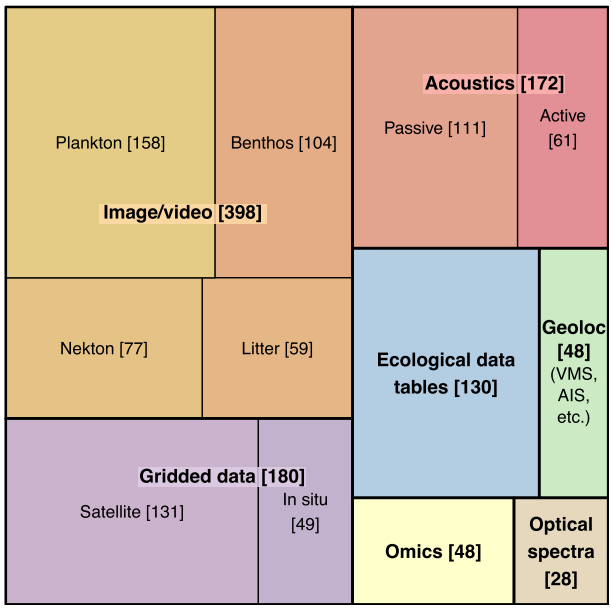

Figure 3. Treemap representation of the papers in the database that can be categorized according to the type of data they use. The area of each rectangle is proportional to the number of papers (written in brackets). The broad data types are bold and coloured with a given hue. Sub-types, when they exist, are in variations of the same hue.

Benthos

Underwater imaging of the benthic environment has grown considerably in the last few decades. We reviewed over 100 papers that used ML to process such data and all were published after 2000 (the earliest is Soriano *et al.*, 2001). Almost half of these focused on habitat mapping (e.g. Porskamp *et al.*, 2018) and coral reefs and their inhabitants (e.g. Villon *et al.*, 2018), followed by studies focussing on the detection of benthic invertebrates (e.g. Kiranyaz *et al.*, 2010, 2011). The most used algorithms included support vector machines (SVMs), random forest (RF), convolutional neural networks (CNNs), k-nearest neighbours (kNN), and classification and regression trees (CARTs). Those were used mostly for image/pixel classification (in more than half of the studies) on their own or with other algorithms, as previously pointed out by Lopez-Vazquez *et al.* (2020). More recently, object-based classification has replaced pixel-based classification (Zhang *et al.*, 2013), especially using CNNs, which reached much higher performance (Gómez-Ríos *et al.*, 2019; Piechaud *et al.*, 2019).

Growth in this field has naturally been accompanied by an increase in the number of images of benthic fauna and habitats. Though ML offers promise towards unlocking the catalogue of unused benthic images, many challenges remain. There are growing concerns regarding the identification of available data for training, the pre-training of deep nets, and the handling of class imbalance in training datasets. For example, of the millions of images acquired each year on coral reef surveys, just 1–2% are labelled (Beijbom *et al.*, 2012). Lumini *et al.* (2020) compared several CNN architectures and found that combinations of several models (i.e. ensembles) were the most successful for image classification of coral (and plankton) datasets. Fincham *et al.* (2020), who classified images from across multiple benthic habitats, found an imbalance in their training data due to the frequency of habitat occurrence, which was countered by using data augmentation to artificially expand the training by flipping, scaling, and rotating images.

The challenge in accessing high-quality training datasets is now being addressed through developments such as standardized reference catalogues (Althaus *et al.*, 2015; Fisher *et al.*, 2016; Howell *et al.*, 2019), wide adoption of specialized annotation software such as BIIGLE 2 (Langenkämper *et al.*, 2017), SQUIDLE+ (Williams and Friedman, 2018), and CoralNet (Beijbom *et al.*, 2015), and annotated image databases e.g. FathomNet (Boulais *et al.*, 2020). In addition, the development of user-friendly software such as VIAME and Superannotate is making ML more accessible to benthic ecologists. However, for researchers to best apply these tools, much remains to be learned regarding model performance under different conditions (e.g. depending on the number of classes used), on training dataset size, on the use of single models versus ensembles of models, etc. (Durden *et al.*, 2021).

Macrolitter

Each year, tonnes of human-created waste litters the sea surface, seafloor, and shorelines and poses a major threat to oceanic ecosystems and coastal communities (NOAA, 2014). Extensive surveys and research are conducted worldwide to assess litter distributions and concentrations in coastal areas and the open sea, to identify litter accumulation zones through numerical models, and to design management actions to promote litter removal and recycling (NOAA, 2016; Madricardo *et al.*, 2020). To quantify marine litter, video and photography-based monitoring is increasingly adopted and deployed on bottom trawl or nets, autonomous underwater vehicles, remotely operated vehicles, unmanned aerial systems, and drones. However, litter identification is mainly done by humans, which is time-consuming, costly, and often very subjective, creating the need for automatic approaches (Canals *et al.*, 2020).

Region-based convolutional neural networks (R-CNNs), designed for object detection, have been increasingly applied to automatically detect and classify beached (Watanabe *et al.*, 2019), floating (Lieshout *et al.*, 2020), and seafloor (Politikos *et al.*, 2021) macrolitter items. Additionally, traditional CNN classifiers have been used to categorize litter types from segmented images (Garcia-Garin *et al.*, 2021). Such studies have generally shown that the classification and detection performance of neural networks is high for floating litter (>80%) but often lower for underwater and seafloor litter, which can be attributed to the challenges of underwater imagery (various camera angles, zoom levels, light shadings, litter buried in the seabed). Several authors have used open and experimental datasets for their analysis, focusing mainly on the predictive performance of the algorithms. The applicability of ML for marine macrolitter research has been recently reviewed in more detail (Politikos *et al.*, 2023).

Ultimately, DL has the potential to support monitoring of marine litter by providing automatic, rapid, and scalable solutions. Nevertheless, a collection of images and video recordings from real-world environments and more effective algorithms are needed to support litter assessment goals set by stakeholders (Politikos *et al.*, 2023). Finally, new imaging technologies such as infrared detection (Inada *et al.*, 2001) or Raman imaging (Gallager, 2019), which can identify plastics at least in a laboratory setting, could be implemented and integrated with ML techniques for improved results.

Nekton

Monitoring of nekton informs decision-making for biodiversity conservation and sustainable fisheries management. Imaging surveys constitute a non-invasive complement to conventional monitoring. However, it yields large datasets and ML has come into play to automate and speed up the data processing. Nekton monitoring from images is challenging due to the diversity of tasks that need to be solved (e.g. species classification but also morphometric estimations) and the very different conditions in which images are collected (e.g. both underwater and on ships).

In early fish imaging studies, classic ML methods were used with data obtained in controlled, experimental setups. For example, in Storbeck and Daan (2001), the image acquisition system consisted of a camera and a laser, which allowed obtaining images but also information on fish volume. They classified six species of fish with 95% accuracy using a shallow artificial neural network (ANN) based on fish contour features. In Zion *et al.* (2007), three edible fish species were sorted using a minimum Mahalanobis distance classifier that combined geometric features and object contours as inputs to yield an accuracy of >96%.

In situ monitoring of nekton is largely focused on fish as well, but those studies present additional challenges due to the wide variations in observation conditions. Datasets are typically collected with underwater cameras (Fisher *et al.*, 2016), but larger organisms, such as marine mammals, are also monitored via satellite images. Here also, before the development of CNNs, global features were used together with background modelling to detect and track objects under water. For example, Spampinato *et al.* (2010) developed a Gaussian mixture model (GMM) and moving average algorithm followed by an adapting mean-shift algorithm to detect and track fish in *in situ* videos, with an 85% success rate. Hu *et al.* (2012) reached >97% accuracy in the classification of fish images based on texture and colour features, using two kinds of SVMs. Another approach to handle the change in appearance of objects underwater is to consider the information from a sequence of frames in a video rather than from only one frame, as done in Shafait *et al.* (2016) where, for ten species, accuracy ranged from 71 to 100%. Finally, artificial alterations in images (i.e. data augmentation) are a common way to improve performance and generalization in CNNs. Allken *et al.* (2019) trained an Inception 3 architecture on 5000 data-augmented images per species to reach 94% accuracy on a test set, while the baseline model, trained on the 70 original images, reached an accuracy between 50 and 71%. Bogucki *et al.* (2019) used a combination of three CNNs to detect and identify North Atlantic right whales in aerial and satellite images. The first CNNs located the whales in satellite images, the second detected key points on the whales' heads in aerial survey images, and the third identified the whales.

Images of nektonic organisms are also collected outside of the water, by camera systems deployed on fishing vessels (known as electronic monitoring), which have replaced some on-board fishery observers. Deep models have become valuable tools to process the videos collected (Helmond *et al.*, 2020). Specifically, Mask R-CNN was provided pixel-level masks and bounding boxes around organisms to automatically monitor catches on-board fishing vessels (Tseng and Kuo, 2020). Such masks can be used to automatically measure the length of the detected objects. In Garcia *et al.* (2020), segmentation performance was assessed with the intersection over union (IoU) metric computed between the predicted and ground truth masks. In this study, 1605 images were used to train a Mask R-CNN model, and an average IoU of 0.89 was obtained on 200 independent test images. Notably, in more challenging images where fishes overlapped, the IoU was lower, as expected.

Plankton

Because planktonic organisms are often micrometric to millimetric, high magnification is needed to image them, which implies a short depth of field and can lead to many out-of-focus objects. In addition, *in situ*, images are dominated by morphologically diverse detrital particles that are similar in size to living organisms. Finally, the organisms themselves are also incredibly diverse. Therefore, the automatic classification of such images is a difficult and interesting ML problem, and remains a major bottleneck for their exploitation.

The first attempts at machine-based classification of plankton images derived various features from the images: statistical moments (which capture size, average lightness, etc.), Fourier transforms of the contour of the object, texture patterns (Tang *et al.*, 1998), and, later, grey-level co-occurrence matrices (Hu and Davis, 2005). Those features were input into a classifier, often an ANN (Tang *et al.*, 1998), an SVM (Hu and Davis, 2005), or a combination of both to classify images into a limited (mostly fewer than ten) number of taxa.

These approaches matured and the next decade saw the rise of their application for numerous ecological studies. The most influential papers of this period are associated with popular instruments and software. For instance, Grosjean *et al.* (2004) and Gorsky *et al.* (2010), while presenting the ZooScan, highlighted that (i) the performance of different classifiers is largely similar and therefore mostly determined by the original features, (ii) this performance decreases strongly when the number of taxa to classify increases, and (iii) with 8 taxa, predictive power saturates beyond 300 example images per taxon in the training set. Sosik and Olson (2007) presented the Imaging FlowCytoBot and described in detail the reasoning and process to derive features particularly relevant for phytoplankton, from the original images. Despite the large number of papers, applications of those techniques at broad spatial and temporal scales are still rare (but see Irigoien *et al.*, 2009).

The next evolution in this research was the increasing use of CNNs, particularly since 2015, owing to a plankton image classification competition run on Kaggle.com (Robinson *et al.*, 2017; Figure 4). However, the thoroughness of papers using this technique is inconsistent and many are published in conference proceedings that are difficult to access. By contrast, Ellen *et al.* (2019) provide an extensive overview of the setup of a CNN from scratch, including the choice of its parameters, the inclusion of classic image features and other external information into the CNN's classifier, and compare the CNN's performance with the more classical approaches described above.

CNNs will likely be increasingly relied upon in the future. Their implementation within dedicated plankton imaging software such as EcoTaxa (Picheral *et al.*, 2017) or the IFCB dashboard will facilitate their routine use by a wide community of ecologists (https://ifcb-data.whoi.edu/dashboard). The separation of their feature-extraction part from their classification part seems like a promising avenue for transfer

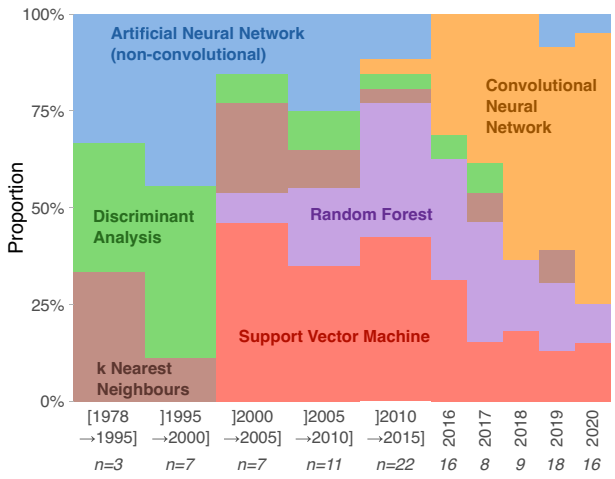

Figure 4. Classifiers used for plankton image recognition through time. The plot displays the proportions rather than absolute numbers. The time bins on the *x*-axis are not regular. The plot highlights the quick adoption of support vector machines and their current decline, the rise and fall of random forests, the increase of convolutional neural networks (particularly since 2015), and their current dominance.

learning (i.e. using a model initially trained on one, often general, dataset to quickly "fine-tune" it on a another dataset, here a plankton one; Orenstein and Beijbom, 2017), unsupervised classification (Schroeder *et al.*, 2020), and active learning procedures, whereby only few images representative of the diversity of the dataset are shown to the user (Bochinski *et al.*, 2019).

An important problem in plankton image datasets, like in many other biological ones, is class imbalance (a few classes dominate the samples). Among several solutions, generating synthetic images in the rare classes using a generative adversarial network (GAN) was recently tested (Li *et al.*, 2021). Alternatively, quantification approaches, which do not aim to perfectly classify each individual object but rather to directly derive concentration estimates for each class, deal intrinsically with the distribution among classes (Gonzalez *et al.*, 2019).

Identifying microorganisms from single-cell spectra

Flow cytometry has been used since the 1990s to study marine microbial communities. In flow cytometry, scatter and fluorescent properties of individual particles are measured at very high rates (i.e. hundreds to thousands of particles per second). Although most researchers manually analyse the resulting "cytograms", automated methods have become available for the analysis of such microbial flow cytometry data (Rubbens and Props, 2021).

Since the 1990s and early 2000s, artificial neural networks (ANNs) have been developed to identify up to 72 lab-grown phytoplankton species using flow cytometry (Boddy *et al.*, 2001). Supervised single-cell classifiers were then successfully applied for the identification of heterotrophic bacteria as well, by combining flow cytometry with a nucleic acid stain in most cases. Besides ANNs, support vector machines (SVMs), linear discriminant analysis (LDA), and random forests (RFs) have been successfully applied in this setup (Rajwa *et al.*, 2008; Rubbens *et al.*, 2017). These lab-based studies have demonstrated the usefulness of the information captured by flow cytometry for bacterial and phytoplankton identification. However, it is difficult to transfer this knowledge directly to samples taken from the field. As the identity of species present is often unknown, labels are not available to train supervised models. When analysing field samples, unsupervised clustering approaches are therefore used to group together cells that have similar optical properties. Examples include Gaussian mixture models (GMMs), graph-based clustering, and self-organizing maps (SOMs) (Hyrkas *et al.*, 2016; Sgier *et al.*, 2016; Bowman *et al.*, 2017).

In some cases, cell populations do not form distinct patches that can be isolated by clustering: when the complexity of microbial communities is high (i.e. many taxa) or the resolution is limited (e.g. due to the instrumental setup or when studying heterotrophic organisms). Cytometric fingerprinting approaches do not try to identify cell populations; instead, they focus on modelling the multivariate distribution of the cytometric data, by defining informative regions in this distribution and recording cell counts or densities in those regions. Often, binning approaches are employed, although more advanced strategies have become available as well, e.g. by overclustering the data using a GMM (Rubbens *et al.*, 2021) or by an automated deleting, merging, and shrinking of Gaussian mixtures (Bruckmann *et al.*, 2022).

A few hybrid approaches have been proposed for freshwater samples, in which information from laboratory cultures is used to analyse natural samples. RF classification was used to differentiate noise from signal using lab-grown cultures and then used to remove the noise in natural samples (Thomas *et al.*, 2018). Learned representations of lab-grown cultures can also be used as proxies to describe the dynamics of a microbial community in a natural sample (Özel Duygan *et al.*, 2020).

Raman spectroscopy is an alternative, information-rich, single-cell technology for the identification of marine microorganisms. Spectra typically contain many more variables than traditional flow cytometry data; therefore, the use of convolutional neural networks (CNNs) should be beneficial to summarize this information and get to single-particle identification. When a CNN was trained on Raman spectroscopy data, it resulted in high classification accuracy for 13 marine microorganisms (~95%) but similar to that of SVM and LDA, probably due to a low sample size (Liu *et al.*, 2020).

Describing ecosystems with acoustics

Light attenuates faster in water than in air, limiting cameras to observing a small volume (albeit at high resolution). Sound propagates over long distances and is used to monitor the ocean interior. Sound also samples larger volumes of water than towed nets and can be used in areas that are otherwise difficult to reach, such as deep water and rough bathymetry. Both active sensors, which emit sound and measure the returned echoes, either from organisms in the water column or from the seabed, and passive sensors, which just "listen", are commonly used in marine science. The following text is organized along those categories.

Active acoustics for target classification

Active acoustics are widely used in fisheries and aquaculture to evaluate the spatial and temporal distributions of organisms, measure their size distribution, and calculate population structure, as well as characterize the behaviours of species. In all cases, the analysis starts with the identification

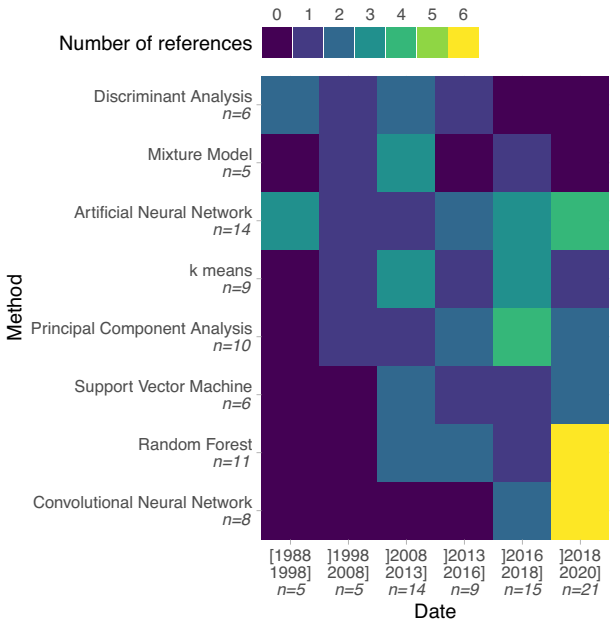

Figure 5. Evolution of the methods used for target classification from active acoustics data, through time. The labels give the total number of references per row or column of the plot. The colour is proportional to the number of references published in the time period of the column and using the method of the row. A single reference can appear in several rows if it uses several methods.

of the returned echo, also called target classification (Korneliussen, 2018). This process frequently involves manually checking, cleaning, processing, and scrutinizing the echogram features. Target objects are then delineated and ascribed to species using "expert" knowledge gained from biological samples. This heavy dependence on manual operations makes the process time-consuming and vulnerable to bias; scalable and reproducible methods, such as ML-based approaches, are therefore needed.

Early attempts to automate target classification typically used deterministic features computed from the data, using details in individual echo pulses (Rose and Leggett, 1988) and/or school-based features like shape or energy, as well as auxiliary information like location or depth; this information was passed to a range of classifiers. Artificial neural networks (ANNs) were used early on (Cabreira *et al.*, 2009; Figure 5). Random forests (RFs) were used with school-based features and auxiliary information, for fish identification (Fallon *et al.*, 2016). Peña (2018) recently reviewed clustering techniques for acoustic data and concluded that expectation-maximization (EM) clustering is the only technique that properly separates acoustic signatures (and noise), after a supervised initialization.

Wideband or multi-frequency echosounders added the frequency dimension to the data, which allowed for improved discriminatory power. Using the frequency response usually involved averaging over certain ping- or range-bins and comparing the scatter distributions to the properties of known aggregations. Using the full broadband echo spectrum, an RF classifier was successful in classifying individual fishes (Gugele *et al.*, 2021).

More recently, convolutional neural networks (CNNs) were used to classify the entire echogram and identify the primary species on patches of echo (Hirama *et al.*, 2017; Figure 5). Shang and Li (2018) used simulated data to compare different classifiers using different features and CNNs reached the best performance. Regions of interest, identified on real data, were more accurately identified by CNNs with various architectures (ResNet, DenseNet, Inception) than by a support vector machine (SVM) classifier working on traditional manual features (Rezvanifar *et al.*, 2019). CNNs have also been used for pixel-level predictions (i.e. segmentation) on raw acoustic data, using a U-net architecture (Brautaset *et al.*, 2020) or Mask-Regional CNN (Marques *et al.*, 2021), trained on manually labelled data. Such supervised methods require large amounts of training data, while recently developed semi-supervised methods allowed only ~10% of the training data to be labelled (Choi *et al.*, 2021).

Active acoustics for seabed and sediment mapping

Active acoustics are also used to map seabed topography and sediment cover, which condition the type of benthic biological community that can develop. Various methods reach high spatial resolutions and accuracy, such as single-beam echosounders, sidescan sonars, and reflection sismographs, but multi-beam echosounders (MBES) are the most cost-effective for mapping large areas (Anderson *et al.*, 2008; Brown *et al.*, 2011). Bathymetry and backscatter data (and their derivatives) are interpreted in order to characterize the type of seabed substrate. For a thorough description of conventional sea bottom classification systems, see the extensive work of Hamilton (2001).

One major challenge for seabed mapping is that the manual interpretation of seabed features from acoustic data is very time-consuming and highly subjective. This explains the increased interest for automated approaches, including inversion algorithms, image-processing techniques, and, mostly, ML (Brown *et al.*, 2011; Stephens and Diesing, 2014).

In the 1990s, the early ML approaches were ANN, e.g. Stewart *et al.* (1994), who successfully classified three different seafloor types based on sidescan sonar data. Dartnell and Gardner (2004) used hierarchical decision trees (DTs) trained on four types of images (backscatter intensity and three variance images). Using 60 ground truth sediment samples, they predicted seafloor types in Santa Monica Bay with an accuracy of 72%, which was better than other automated classification methods at the time.

Since then, a variety of ML methods have been scrutinized through comparative studies (Ierodiaconou *et al.*, 2011; Stephens and Diesing, 2014; Shao *et al.*, 2021). The classification algorithms were very diverse, covering tree-based methods (DT, random forest—RF, Quick Unbiased and Efficient Statistical Tree—QUEST, and Classification Rule with Unbiased Interaction Selection and Estimation—CRUISE, etc.), SVMs, maximum-likelihood classifiers (MLCs), and ANNs. In many cases, ML-based approaches were not significantly different from one another but were vastly superior to the usual manual interpretation procedures.

Recently, Cui *et al.* (2021) demonstrated how a deep belief network (DBN) based on fuzzy ranking feature optimization can be used to map sediment distribution over large areas. Fuzzy ranking is a technique used to identify the feature combination, derived from the MBES data, that is most appropriate for the DBN to correctly classify the seabed sediment type. The accuracy of the DBN proved higher than that of five other supervised classification models (DT, RF, SVM, MLC, and ANN).

Passive acoustics monitoring

Passive acoustic recordings are a reliable and cost-effective method to monitor habitat use, distribution, density, and behaviour of species over space and time. They can be obtained from boats, autonomous devices (either fixed or moving ones), cabled stations, and animal tags, making them usable in a variety of situations (Kowarski and Moors-Murphy, 2021). However, because of their relative ease of use, hydrophones quickly generate large datasets that require automation to extract information from them (Gibb *et al.*, 2019).

The most common approach to process acoustic data is to detect and classify specific acoustic events in a supervised manner. Sound source classification studies have primarily focused on shipping (Zaugg *et al.*, 2010) and mammals' vocalizations (66 out of the 101 references we recorded; Bittle and Duncan, 2013). In the latter, detection and classification algorithms have been used to identify species (Bermant *et al.*, 2019), specific calls (Bergler *et al.*, 2019), or even dialects and individuals (Brown *et al.*, 2010). ML can also be used to localize the position of, or estimate the range to, a certain source without the need to model the sound propagation (Niu *et al.*, 2017), outperforming conventional matched field processing methods. Another application is to relate properties of the source with characteristics of the sound, through regression; these properties included the size of male sperm whales (Beslin *et al.*, 2018) or fish abundance (Rowell *et al.*, 2017). In addition, ML can be used for acoustic source separation, a problem known as the cocktail party problem (Bermant, 2021). Finally, approaches to characterize entire habitats from their soundscape have also been explored (Lin *et al.*, 2019).

A common approach is to extract human-engineered features from the sound and use them as input for an ML algorithm. These features can be derived from the time, frequency, or cepstral domain (transformation of the data to highlight periodic signals), or based on the full image of the spectrogram, a visual representation of sound intensity per frequency as a function of time (Sharma *et al.*, 2020). The algorithms used for classification include SVMs (Jarvis *et al.*, 2008), RFs (Malfante *et al.*, 2018), Gaussian mixture models (GMMs; Roch *et al.*, 2011), and k-means (Weilgart and Whitehead, 1997), among others. More focus has been put on identifying which features are relevant for the classification and characterization of sound events than on which classifier performs best. Often these features or other rule-based signal processing techniques are also used to first segment the data and then ML is used to classify the detected segments.

Advances in image and speech-recognition algorithms have been applied to underwater sound, reducing the amount of preprocessing and improving performance and generalizations (Schröter *et al.*, 2019). In DL approaches, sound is often converted into a spectrogram, which is considered as an image and input into a convolutional neural network (CNN) for classification, regression, or feature extraction and clustering (Bermant *et al.*, 2019; Thomas *et al.*, 2020). However, recently some models have been developed that are applied directly on the waveform (Roch *et al.*, 2021).

In the marine context, sounds of interest can be very sparsely occurring and datasets can comprise long periods of time. This leads to highly imbalanced datasets. This imbalance is usually solved by first detecting and then classifying the detected sounds, where the detection step is a rule-based signal-processing algorithm and the classification step is a DL approach (Stowell, 2022). However, the biggest limitation for the application of ML to passive acoustic recordings is the lack of knowledge regarding which sounds are produced by which species, because visual surveys to associate sound with images of the species are often impossible. This leads to a lack of data annotation and limits the usage of supervised ML approaches. To compensate for the lack of ground-truth data, unsupervised clustering algorithms are being developed to acquire general information about the ecology of certain habitats (Ozanich *et al.*, 2021).

Profiling biological communities with environmental genomics

The study of nucleic acids obtained from an environmental sample is coined as environmental genomics (or meta-omics). In marine ecology studies, the genetic information usually comes from a community of organisms rather than from a single specimen, which is our focus here. Metabarcoding (amplification by polymerase chain reaction and sequencing of a taxonomically informative gene) allows documenting biological communities in terms of species presence and proportions. Metagenomics (shotgun sequencing of a complex mixture of genomic DNA) provides information of random sections of genomes, allowing us to gain insight into both taxonomy and functions. Metatranscriptomics (shotgun sequencing of isolated RNA transcripts) provides similar information for genes active at the time of sampling.

ML approaches have long been used for genomics data analysis. This includes both translating raw signals into nucleotides using base-calling algorithms (Wick *et al.*, 2019) and sequence data analysis. For instance, hidden Markov models have been extensively used for functional annotations, multiple sequence alignments (Yoon, 2009), and more recently for viral signatures detections in metagenomic datasets (Ponsero and Hurwitz, 2019). However, few studies have applied ML to strictly marine meta-omics data. We therefore provide a general overview of the analysis of metabarcoding data and highlight some ML applications to marine data.

Metabarcoding datasets are usually processed by well-established bioinformatics software, e.g. QIIME 2 (Bolyen *et al.*, 2019), which translates raw sequences into statistically exploitable species-to-sites count matrices. Sequences are often grouped into operational taxonomic units (OTUs) or amplicon sequences variants (ASVs) based on their similarity. These sequence units then serve as a proxy for species/strains to document biodiversity changes. Current algorithms to cluster sequences into OTUs or ASVs are VSEARCH (Rognes *et al.*, 2016), which relies on an arbitrary similarity cutoff to delineate OTUs (e.g. 97%), SWARM (Mahé *et al.*, 2015), which aggregates neighbouring sequences to abundant, supposedly genuine, seed sequences, or DADA2 (Callahan *et al.*, 2016), which uses base calling values to separate spurious from genuine sequences. These two latter methods find more "natural" boundaries of OTUs and, as such, can be considered as unsupervised approaches. Some OTUs are then assigned a taxonomic name based on similarities with known sequences from curated databases (e.g. PR2, Guillou *et al.*, 2013; SILVA, Quast *et al.*, 2013). To this end, several ML-based methods have been developed, including naive Bayes (NB) classifiers (the RDP classifier, Wang *et al.*, 2007) and classification trees

using k-mers distributions across sequences (Murali *et al.*, 2018). More recent work successfully applied convolutional neural networks (CNNs) to process and taxonomically annotate raw metabarcoding data faster, without relying on operational OTUs or ASVs (Flück *et al.*, 2022).

Resulting OTU-to-site count matrices are then amenable to biodiversity analysis using compositionality-aware multivariate statistics (Quinn *et al.*, 2019). For example, ML allows routine monitoring of the impact of industries on marine biodiversity. Based on metabarcoding datasets labelled with ecological states obtained by conventional methods, random forest (RF) models can be trained to assess the ecological status of new samples, based on their metabarcoding profiles alone. This is faster and more cost-effective than conventional morpho-taxonomy approaches, enabling scaling up the spatio-temporal scales of biomonitoring programs (Cordier *et al.*, 2018; Frühe *et al.*, 2021).

Network ecology research has been developed on interactions between macro-organisms (e.g. plant-pollinator interaction networks), but many interactions remain difficult to observe and validate. This is especially true within microbial communities, for which statistical frameworks have been developed to detect co-occurrence patterns and include them into more holistic ecological studies. ML techniques can be used to predict species interactions (Vacher *et al.*, 2016; Bohan *et al.*, 2017) and can outperform the identification of trait-matching combinations compared to generalized linear models (Pichler *et al.*, 2020). Microbial networks can be inferred from genomics data (Faust and Raes, 2012; Lima-Mendez *et al.*, 2015) as a means to predict putative biotic interactions, which opens new avenues for understanding the links between marine microbial communities and the large-scale functioning of marine ecosystems (Guidi *et al.*, 2016; Chaffron *et al.*, 2021). Finally, ML is expected to contribute to improve our capacity to analyse massive meta-datasets composed of numerous collated cross-study genomics data, by controlling for covariates (Wirbel *et al.*, 2021).

Quantifying and mapping fishing pressure from geolocation data

Fishing and shipping activities are putting important pressure on marine ecosystems. They are often tracked using vessel monitoring systems (VMSs) or the automatic identification system (AIS), which transmits vessel locations at regular intervals (Thoya *et al.*, 2021). VMSs are required by fisheries management agencies for many commercial fishing vessels and the data are often confidential. AIS is designed for maritime safety, for any type of vessel, and the data are more broadly accessible. These data are often extensively processed using ML to identify vessel and gear types (Russo *et al.*, 2011; Marzuki *et al.*, 2018; Taconet *et al.*, 2019).

Many studies have classified fishing vs. non-fishing behaviours using artificial neural networks (ANNs; Bertrand *et al.*, 2008; Russo *et al.*, 2014) and random forests (RFs; Ducharme-Barth and Ahrens, 2017; Behivoke *et al.*, 2021). To do so, the movement characteristics of vessels across space, time, and habitats are often studied and summarized before being provided to the ML classifier. Kroodsma *et al.* (2018) trained convolutional neural networks (CNNs) with AIS data to identify fishing vs. non-fishing behaviours and fishing gear types, producing the first map of the global footprint of fisheries (Taconet *et al.*, 2019).

The outputs of these models have been used not only to assess fishing pressure but also in ecological studies to estimate noise impacts (Allen *et al.*, 2018), assess marine spatial planning or monitor conservation areas (Robards *et al.*, 2016; White *et al.*, 2020), identify species distribution (Le Guyader *et al.*, 2016), minimize mammal strike risk (Fournier *et al.*, 2018), and mitigate bycatch (Richards *et al.*, 2021).

To integrate fishing activity with the rest of the ecosystem, ML efforts on fishery geolocation data have used an expanded suite of predictor variables. For example, several studies used boosted regression trees (BRTs) to relate fishing locations with environmental information (e.g. sea surface temperature) and then predict dynamic maps of fishing activity from environmental data (Soykan *et al.*, 2014; Crespo *et al.*, 2018). Other studies added bio-economic considerations into fisher location-choice frameworks, with ANNs (Dreyfus-Leon and Kleiber, 2001; Russo *et al.*, 2019). By characterizing fishing behaviours using these broader features (e.g. environment, bio-economics), ML approaches provide a valuable foundation for operational, dynamic, ocean management tools that support ecosystem-based fishery management in near real-time (Hazen *et al.*, 2018).

Deriving biogeochemical variables from satellite images and floats profiles

Historically, most *in situ* measurements used for the characterization of ocean biogeochemical processes were acquired using ships, resulting in critical undersampling at a global scale. Advances in remote sensing (by ocean colour satellites) and *in situ* robots now allow sampling marine bio-optical variables at unprecedented spatio-temporal resolution (Claustre *et al.*, 2020).

Yuan *et al.* (2020) provide a review of applications of DL to environmental remote sensing for estimating atmospheric, land, and oceanic physical, chemical, optical, and biogeochemical variables. One section is dedicated to the use of ML for remotely sensed ocean colour parameters retrieval, mainly focussed on the estimation of the chlorophyll-a concentration. However, ML has also been applied to remote-sensing data to derive fields of inherent optical properties of the seawater (Ioannou *et al.*, 2011, 2013), pCO_2 (Landschützer *et al.*, 2015), primary production (Mattei *et al.*, 2018), phytoplankton community composition (Stock and Subramaniam, 2020), particulate organic carbon (Liu *et al.*, 2021), dissolved inorganic carbon (Roshan and DeVries, 2017), and nitrogen fixation rate (Tang *et al.*, 2019), as well as perform atmospheric correction (Jamet *et al.*, 2005; Brajard *et al.*, 2012) (Figure 6).

One remarkable example of using ML for ocean science is the synergy between satellite observations and *in situ* profiles, in particular from the Argo programs (>100000 currently). Sauzède *et al.* (2016) used a multi-layer perceptron to extend surface bio-optical properties to depth. This produces four-dimensional (i.e. longitude, latitude, depth, and time) fields of biogeochemical variables at global or regional scales, which fill *in situ* observational gaps. Such continuous fields are particularly valuable for the initialization and validation of biogeochemical models. They are now reaching operational status since four-dimensional fields of chlorophyll-a concentration and particulate organic carbon generated by these methods have recently been made publicly available on the European online portal Copernicus Marine Environment Monitoring Service.

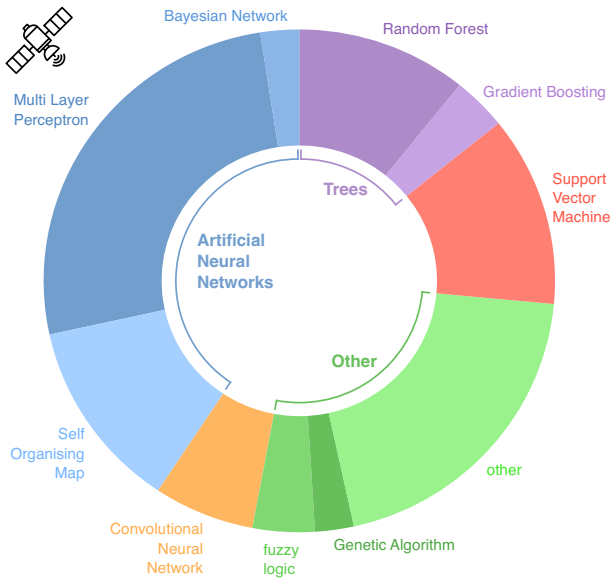

Figure 6. Machine learning methods used with satellite imagery data. Artifical neural networks (in blue shades), and, in particular, multi-layer perceptrons, dominate the literature that was reviewed.

Finally, ML methods are also used to estimate the more scarcely measured biogeochemical variables from the more commonly measured physical ones. For example, an ANN was trained to predict nutrient concentrations and carbonate system variables from over 250000 profiles of pressure, temperature, salinity, and oxygen concentration (Bittig *et al.*, 2018). The predictor variables can be measured with very high accuracy by autonomous floats and now ANN-based methods can spatially and temporally populate the fields of nutrients and carbon variables, which were previously loosely resolved. MLPs have also been used to predict the phytoplankton community composition from profiles of fluorescence of chlorophyll-a (Sauzède *et al.*, 2015a), making it possible to gather and homogenize tens of thousands of fluorescence profiles available from historical databases, which could not be integrated in global analyses before (Sauzède *et al.*, 2015b).

Machine learning to improve ecological understanding

Once ecology-ready tables of data have been extracted from raw sources (see section "Machine learning to extract information from observational data"), they can be analysed to gain a better understanding of socio-ecological marine systems (this section). Such studies traditionally use statistics, often multivariate, and modelling to capture relationships between observed variables; this task is also amenable to ML. In this section, we highlight how ML techniques are used to relate species to their environment and, in particular, predict species distributions, detect dynamic interactions involving several species, and, finally, inform ecosystem management by partitioning the environment in easier-to-understand units through regionalization and fueling monitoring and decision-support tools.

This field is even more difficult to map through literature searches than the more technical studies presented in the previous section. Some searches with relevant keywords yielded >10000 results, while others with minor differences yielded only hundreds. Therefore, in this section, even more than in the previous one, we really focus on presenting papers that showcase different approaches.

Predicting species abundance and distribution

The ability for ML approaches to capture complex and non-linear relationships, as well as their ability to work with missing and heterogeneous data, has driven their popularity for the analysis of species–environment relationships.

When data are sparse or heterogeneous, often also leading to high uncertainty, Bayesian ML methods have proven useful. Fernandes *et al.* (2010) predicted fish recruitment using a naive Bayes (NB) classifier relying on spawning stock biomass, climate, and weather data. Fernandes *et al.* (2013) used multi-dimensional Bayesian networks for a similar task and found that predicting three species simultaneously doubled the chance of being correct, compared to three single-species models. Lehikoinen *et al.* (2019) used tree-augmented NB models to evaluate the influence of various environmental factors, all heterogeneous in type and in spatio-temporal resolution, on coastal fish abundance. They note that some environmental factors are not relevant to predict average abundances, but are important for extreme ones.

Tree-based ensemble models such as random forests (RFs) and boosted regression trees (BRTs) have also proven useful with ecological data thanks to their versatility and ease of use. Knudby *et al.* (2010) found tree-based methods superior to linear models in predicting species richness, biomass, and diversity in coral reefs based on habitat variables. Suikkanen *et al.* (2021) used RF regression to analyse the relationships of zoo- and phytoplankton (particularly cyanobacteria) in multi-decadal (but relatively sparse) monitoring data to find whether relationships found in experiments could also be seen in field data.

Species distribution models (SDMs) are frequently applied to perform spatially explicit analyses of ecological data. They quantify the relationship between species occurrence or abundance and their environment and can be then used to predict their potential geographical distribution (Guisan and Thuiller, 2005; Elith and Leathwick, 2009). A significant body of literature compared the performance of ML-based SDMs with multivariate linear regression or climate envelope methods, generally finding that ML methods yield better predictive performance but are prone to overfitting (e.g. Derville *et al.*, 2018).

The most widely applied ML method for SDMs is Maxent, with over 6000 published papers, which showcases the power and broad applicability of ML for ecological inference (Phillips and Dudík, 2008; Elith *et al.*, 2011). Maxent works with records of a species present at given points in space and iteratively maximizes the probability of presence at these points, predicted from functions of environmental variables at the same points (Phillips *et al.*, 2006). But, many other ML approaches are also used in species distribution modelling, such as decision trees (DTs; Hunt *et al.*, 2020), BRTs (Elith *et al.*, 2008; Cimino *et al.*, 2020), RFs (Reiss *et al.*, 2011), support vector machines (SVMs; Knudby, 2010; Vestbo *et al.*, 2018), and artificial (Benkendorf, 2020) and convolutional neural networks (CNNs; Deneu *et al.*, 2021). These models have been applied to resolve a diverse range of ecological and conservation issues, including understanding species ecology (Brodie *et al.*, 2018), responses to current and future environmental

change (Hindell *et al.*, 2020), threat overlap (Welch *et al.*, 2018), and the design and evaluation of spatial management scenarios (Stock *et al.*, 2020; Smith *et al.*, 2021). Across all applications, communicating the uncertainty of SDMs to stakeholders is critical. In general, estimating uncertainty within ML-based SDMs is difficult, and most solutions underestimate model uncertainty (Beale and Lennon, 2012; Watling *et al.*, 2015; Brodie *et al.*, 2020). However, new approaches, such as Bayesian additive regression trees, are emerging and improving our estimation of uncertainty (Carlson, 2020).

Capturing dynamic ecological relationships

As climate variability and long-term change drive non-stationarity in ecosystems, more research is needed to see how ML approaches can improve our ability to predict and forecast potentially changing species relationships with their environment and other species. Latent (hidden) variable modelling provides one way to detect an underlying systemic change, or to approximate an ecosystem component that is not represented in the dataset. Trifonova *et al.* (2015) modelled the North Sea ecosystem using dynamic Bayesian networks with hidden variables (DBN-HVs), and concluded that a hidden variable in the model managed to learn the zooplankton biomass variations in all modelled areas. Trifonova *et al.* (2017) used this model to predict ecosystem responses under different scenarios. Uusitalo *et al.* (2018) and Maldonado *et al.* (2019) created a DBN-HV model for the central Baltic Sea food web and found that the hidden variables replicated the regime shift, i.e. the drastic change in the ecosystem organization that has been reported by Alheit *et al.* (2005) and others. These studies exemplify the ability to combine data analytics and domain knowledge through ML to provide explanatory models that provide new insight into ecosystem functioning. Sander *et al.* (2017) used DBNs to infer ecological relationships, but note that presence–absence data may not provide enough signal for these models. Pichler *et al.* (2020) evaluated the ability of multiple ML methods to infer species interactions in the terrestrial domain, but similar approaches could be applied to marine data.

Summarizing ecosystems through regionalization

In recognition that the ocean is spatially and temporally heterogeneous, its division into various types of regions (bioregions, ecoregions, provinces, essential habitats, etc.) provides a means of simplifying and summarizing this heterogeneity into units amenable to further analysis and management. Pioneering this approach was Longhurst *et al.* (1995), who defined 57 biogeochemical provinces mainly using regional variation of remotely sensed chlorophyll-a. In more recent years, ML techniques have been adopted to provide more objective classifications. For example, bioregions have been defined based on chlorophyll-a dynamics using k-means clustering (Mayot *et al.*, 2016) and hierarchical Iso Cluster classification (Welch *et al.*, 2016). Multiple biophysical variables have been used as input to multivariate unsupervised clustering to define pelagic habitats (Hobday, 2011; Reygondeau *et al.*, 2018) or track the spatial variability of ocean water masses (Phillips *et al.*, 2020). The concentration of biological organisms derived from survey data (Santora, 2012), ecosystem models (Sonnewald *et al.*, 2020), and species distribution models (Welch and McHenry, 2018) has also been integrated into classifiers to define ecoregions. Such ecoregions can be useful for spatial planning purposes since they are quite close to the biological targets of such management procedures (Douglass *et al.*, 2014).

Supporting human decisions on ecosystem management

Finally, we also need to evaluate human–ecosystem interactions and define management strategies that support the health and sustainable use of marine ecosystems. These strategies are often defined in intergovernmental texts (e.g. the EU Marine Strategy Framework Directive) that summarize them in terms of quantifiable objectives; ML can help assess those objectives. For example, the likelihood to reach the goals set by the European Union's Water Framework Directive in Finland was modelled using Bayesian networks (Fernandes *et al.*, 2012). In another example, the accuracy of the automatic classification of plankton images was assessed by checking whether it could provide zooplankton indicators for the EU's Marine Strategy Framework Directive (Uusitalo *et al.*, 2016).

Early warning regarding specific health indices or potentially harmful species is another area where the fast throughput of ML approaches can improve our practice. A major effort has been spent in predicting algal blooms affecting recreational activities, fisheries, and shellfish farming (Campbell *et al.*, 2013; Fernandes-Salvador *et al.*, 2021). But, similar approaches are used for predicting fish recruitment (Dreyfus-León and Chen, 2007; Fernandes *et al.*, 2010) or forecasting litter accumulations on beaches (Granado *et al.*, 2019; Hernández-González *et al.*, 2019).

An international commitment to protect 10% of the ocean by 2020 showcased the importance of spatial planning as a management tool for marine resources (Grorud-Colvert *et al.*, 2019). ML methods, such as automated plankton image classification, are used to monitor and inform the creation of marine protected areas (Muñoz *et al.*, 2017; Benedetti *et al.*, 2019). Dedman *et al.* (2017) developed a tool to simplify the use of marine spatial planning tools based on boosted regression trees. Bayesian networks in combination with geographical information systems are being used to analyse conflicting uses, e.g. how to reallocate aquaculture and different fishing fleets with minimal harm (Coccoli *et al.*, 2018; Gimpel *et al.*, 2018), to plan the locations of new activities such as wind energy (Pınarbaşı *et al.*, 2019), or to consider social and economic aspects in addition to environmental ones (Pınarbaşı *et al.*, 2017; Laurila-Pant *et al.*, 2019).

The efficient management of marine ecosystems would require taking decisions that are informed by the current and future states of these systems. ML can be used to build such decision support tools. For example, fish abundance and recruitment are good indicators of the status of fish stocks, and are used to set fishing regulations. But, small pelagic fish recruitment does not follow traditional stock–recruitment relationships, which is why environmental conditions were used to forecast recruitment using ML-based regression (Chen and Ware, 1999; Fernandes *et al.*, 2015) and to influence fisheries advice (Fernandes *et al.*, 2009). The ML-based species distribution models described above have been integrated into operational, dynamic, ocean management tools (Hazen *et al.*, 2018; Abrahms *et al.*, 2019), in which management and policy recommendations update regularly in response to changes in biological, environmental, economic, and societal conditions (Welch *et al.*, 2019).

Discussion and perspectives

General trends in machine learning applications: data, methods, and tasks

The diversity in the sections above shows that ML is now used in many fields of marine ecology, albeit at different levels of advancement. Several factors can account for the success of the application of ML in a given scientific domain. Based on the examples above, a major one seems to be the type of data: Applications of ML were more successful when they could rely on techniques developed and tested in other fields, which could be repurposed to marine ecology because data were of the same type. This contributes to explaining the disproportionate number of applications of ML to images and videos from cameras, which constitute ~45% of the references in the database to which ~15% of references using satellite imagery can be added (Figure 3). Many of those applications benefited from advances in ML motivated by the ubiquity of images in everyday life. For example, several CNN architectures were developed to classify general-purpose image datasets (often ImageNet; Deng *et al.*, 2009), and when they were successful at this task, they also proved relevant for marine applications; for example, the ResNet architecture alone is used in at least 60 papers in the database. Beyond architectures, the weights that result from training CNNs on such large generic datasets are freely distributed by companies (to promote their technology) and can be slightly modified by a short retraining on a marine dataset to yield domain-specific tools (e.g. detect fishes in recordings from underwater cameras). This is called fine-tuning and requires much fewer resources than training from scratch, while yielding very good results. This general approach, called transfer learning, is ubiquitous in the applications of CNNs reviewed above. On the other hand, single-cell spectra obtained from cytometry, for example, constitute a very peculiar type of data and therefore do not benefit from ready-made models; applications of ML to such data are therefore more difficult and scarcer. While sequences of nucleic acids are not common in everyday life, their analysis could still benefit from architectures and pre-trained weights designed for Natural Language Processing, since both are sequences of tokens (e.g. Quang and Xie, 2016). However, practically, omics often rely on well-established bioinformatics pipelines, which are not specific to questions in marine ecology and in which some steps do not involve ML; this contributes to explaining the relative scarcity of references from this large field here.

In terms of methods, the four most used algorithms in marine ecological research were, in increasing order of popularity, support vector machines (SVMs), random forests (RFs), convolutional neural networks (CNNs), and non-convolutional artificial neural networks (ANNs; mostly multilayer perceptrons). ANNs have been used for a long time, which partly explains why they top the list of algorithms; SVMs, and then RFs, came after 2000; since 2013, the usage of CNNs has increased steeply and now they are the ML method most commonly found in new publications (Figure 7). The timing of the usage of those methods in marine ecology largely reflects their appearance or popularization in general: 1995 for SVMs (Cortes and Vapnik, 1995), 2001 for RFs (Breiman, 2001), and 2012 for CNNs (Krizhevsky *et al.*, 2012); this highlights an early adoption of ML innovations by the marine ecology community. In addition, after the initial adoption, the proportion of studies using them among all

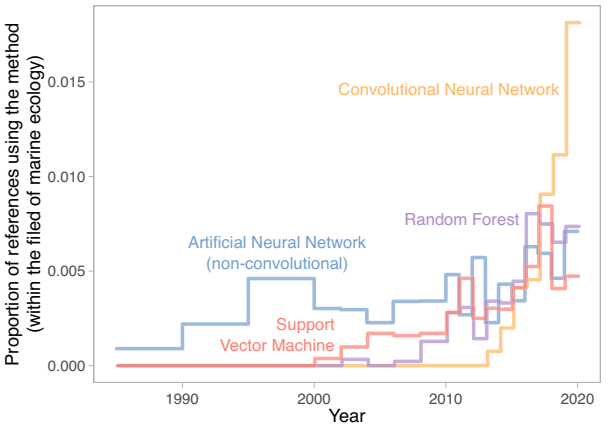

Figure 7. Amount of references per time period using one of the four most common ML methods in the database. To avoid being misled by the global increase in the number of scientific publications, in any field, the amount is expressed as the proportion of the total number of references published in marine ecology in each time period (defined as the result of the query "WC = (Ecology) AND TS = (marine OR sea OR ocean)" on the Web of Science, i.e. Web of Science category is Ecology and title, abstract, or keywords contain "marine", "sea", or "ocean"). All curves increase through time, which means that ML is becoming more common within the field of marine ecology.

studies in marine ecology has grown steeply (Figure 7), which is further evidence of a particular interest for ML approaches in this community. The growth of CNNs, which have progressed the fastest, is associated with their popularity for several data types. Indeed, CNNs take so-called "tensors" as input: multidimensional arrays of numbers. Any type of data that can be made to look like an array within which the proximity between similar numbers is meaningful is amenable to being processed by CNNs. For example, while sounds can be treated as such, most acoustics records can also be represented as spectrograms (intensity as a function of time and frequency), which are tensors and can be processed with models initially designed for images (Stowell, 2022). Finally, depending on the output shape and the loss function used, the same network architecture can be used for regression, classification, object detection, etc. (Goodwin *et al.*, 2022).

Among the papers tagged in the database, ML algorithms are most often used to perform classification (~60% of references) or regression (~20%), and, finally, object extraction (detection or segmentation, ~15%). Yet, the classification of signals, at least, first requires their extraction from the original data (e.g. the detection of an event in a continuous acoustic recording, the segmentation of an organism from an image), so the discrepancy in usage is puzzling. Actually, most automated signal extraction is performed using rules deterministically applied to the raw data. Those rules can be as simple as thresholding (e.g. considering all adjacent dark pixels in an image as objects of interest) but are often much more complex and require both domain expertise to design and signal processing know-how to implement. This hindered the development of automated solutions and explains why objects of interest were (and are still) often extracted manually from underwater videos or acoustics recordings in operational deployments (e.g. Solsona-Berga *et al.*, 2020). DL should enable ecologists to forgo some of the expertise in signal processing and allow extracting signals of interest only from labels placed on a subset of the data. The relative scarcity of their application has

likely several explanations. First, deep models for object detection/segmentation are newer (Girshick *et al.*, 2014) than for classification (Lecun *et al.*, 1998) and their applications lag accordingly. Second, they are a bit more complex to set up than classifiers: Drawing bounding boxes or segmentation masks is more time-consuming than sorting files into folders, training classifiers can often start from just this set of sorted raw files, while object detectors/segmenters require text files in a specific format containing the labels linked to the raw data files, etc. However, as labelling tools (e.g. Labelbox), architectures, and reference datasets (e.g. Katija *et al.*, 2022) continue to improve, such applications are likely to explode in the future.

Finally, supervised ML approaches are much more common than unsupervised ones. This is partly linked with the dominance of classification tasks in the references reviewed. Supervised classification is the archetype of task where ML techniques outperform all others: mimic a simple human action, learn it only from examples generated by humans, and be evaluated almost solely on the quality of the prediction.

Limitations for the application of machine learning

Machine learning is particularly effective when the primary concern of ecologists aligns with the performance metric optimized by the technique (e.g. how many images are classified as the correct species?). Conversely, when focus is on both performance and explainability (e.g. how does yearly recruitment intensity depend on environmental variables?), conventional statistics are often chosen over ML. Indeed, ML approaches are commonly qualified as "black boxes", while people trust models more when they understand the "why" and "how" of their results (Shin, 2021). So, when it comes to decision making at least, inherently explainable models are preferred (Rudin, 2019). However, those black boxes can be studied, by investigating the importance of each input variable or data point independently through randomization, for example (Lucas, 2020). These developments are not unique to ecology and "explainable AI" is an active research domain (Barredo Arrieta *et al.*, 2020).

Another limitation, inherent to ecological data, is the long-tailed distribution of almost everything in the natural world (Preston, 1948). Ecosystems are dominated by some species and some processes, yet many others are present at low abundance/frequency and can be key in the response of the system to changes. Such distributions bias the usual loss functions or evaluation metrics (e.g. least-square error in regression, accuracy in classification) and wide data tables (number or variables larger than the number of observations) favour overfitting to the training dataset, which many ML techniques are already prone to. Dealing with imbalanced datasets is a current research topic in ML: In 2020 and 2021, dozens of papers targeted long-tailed distribution and/or imbalance at the Conference on Computer Vision and Pattern Recognition (CVPR), the major conference in the field (e.g. https://openaccess.thecvf.com/CVPR2021, searching for "long-tail" or "imbalance"). Some of these solutions have been implemented for marine applications, such as rebalancing the training data using data augmentation (Fincham *et al.*, 2020) or generative adversarial networks (GANs; Li *et al.*, 2021), ensembles of several models (Kerr *et al.*, 2020), transfer learning from models trained on balanced data (Lee *et al.*, 2016), etc., but the problem is not solved in the general case.

A consequence of this imbalance is that some of the, numerous, rare classes can largely change in proportion from one data sample to the next, causing a mismatch between the class distribution in the training dataset (usually an average of several samples) and in the new data on which the model will ultimately be applied. This problem, known as "dataset shift" or "concept drift" (Moreno-Torres *et al.*, 2012), is a very common pitfall in the application of ML to marine ecology problems (e.g. Langenkämper *et al.*, 2020) and for the trustworthiness of ML models in general (D'Amour *et al.*, 2020). Indeed, it leads to poor predictive performance that is not necessarily detected when the model is evaluated. Detecting it requires specific validation methods, such as computing evaluation metrics per sample, to capture the inter-sample variability in distribution (Gonzalez *et al.*, 2017). When such a shift is detected, retraining the model with a new training set incorporating more of the natural variability (Langenkämper *et al.*, 2020) or discarding low confidence predictions (Plonus *et al.*, 2021) can help reduce its effect. For classification-type problems, the transition towards quantification approaches, which estimate abundance per class directly, rather than classifying each object, and use the class distribution, can help alleviate it (Gonzalez *et al.*, 2019; Orenstein *et al.*, 2020). Overall, the transferability of a model learned on a given dataset to a dataset with different characteristics is called "domain adaptation" and is also an active field of research (Kouw and Loog, 2019).

Finally, ML models are only as good as the datasets they are trained on. Those training datasets are generated by humans, who can make mistakes. For example, in the hard task of discriminating among six dinoflagellate species with large intra-species morphological variations from images, trained scientists achieved 67–83% self-consistency and only 43% consensus (Culverhouse *et al.*, 2003). For the estimation of benthic cover from quadrat pictures in coral reefs, self-consistency of experts ranged from 50 to 90% depending on the type of cover (Beijbom *et al.*, 2015). One upside is that, in both cases, ML models trained on a reference dataset reached performance similar to or higher than human labellers when their inconsistency is taken into account. Another use of ML can be to resolve ambiguous labels and present such potential mistakes to new experts (Schmarje *et al.*, 2022). Finally, a way of alleviating the effect of inconsistent labels is to use additional, independent "ground truth" validation information. For example, in Lekunberri *et al.* (2022), estimations of species abundance and size inferred with ML were compared with independent samplings and counting at port when fish were landed. While it is impossible to know which is best between the model-generated or on-the-ground estimates, their discrepancies allow narrowing down on potential biases or difficulties for the experts to accurately label the data.

General outlook

As remarked above, so far, applications of ML in marine ecology have closely followed the development of techniques in computer sciences (Figure 7). However, innovation in ML is accelerating and it may be difficult for marine ecologists to keep track of it. Current developments include transformer-based architectures and diffusion models. Transformers can be seen as an alternative to LSTM recurrent neural networks (Table 1) for sequential inputs, such as language; a well-known everyday application is ChatGPT (Chat Generative

Pre-Trained Transformer). Their extensions to images are called vision transformers (ViTs), which can be considered as alternatives to CNNs; they have been topping the ImageNet classification challenge since their release (Dosovitskiy *et al.*, 2021). The combination of text and vision models can be used to learn the relationships between images and their captions. New sets of unlabelled images and potential labels can then be placed in the space created by these relationships to label images without any retraining (i.e. zero-shot learning); an operational example is CLIP (Contrastive Language–Image Pre-training). Diffusion models are improved alternatives to generative adversarial networks (GANs) and variational autoencoders to create synthetic images. They can be used to increase the resolution of input images or create completely new images from text input; a popular example is Stable Diffusion.

Now, how could these innovations percolate to marine ecology? Some applications are straightforward. For example, CNNs can simply be swapped for ViTs in image classification tasks to yield better results (Kyathanahally *et al.*, 2022); similarly, GANs could be swapped for diffusion models. Other applications would require more testing: There is no guarantee that the text-to-image relationships learned by CLIP on images from the internet are relevant enough for specific tasks, such as fish species classification from underwater images, for example. Yet ML models have often been discovered to generalize outside their initial domain: Features extracted by a CNN trained on generic images (from ImageNet) were found to be effective for plankton image classification tasks (Orenstein and Beijbom, 2017). Still, the question whether the potential improvements brought by these new developments are relevant to solving marine ecology problems remains. Improved performance comes at the cost of larger models (25 M parameters for ResNet50 and 632 M parameters for the ViT-H vision transformer; https://paperswithcode.com/sota/image-classification-on-imagenet), which require more data to train (Dosovitskiy *et al.*, 2021). Such massive datasets and the computing power to train on them are often only available in large private companies (ChatGPT and CLIP are from OpenAI, the first ViT is from Google, etc.). While the performance benefit is measurable on well-defined challenges (8% increase in accuracy on ImageNet between the two models above), the actual gains on the smaller, noisy, imbalanced datasets of marine ecology, for which global accuracy may not even be a relevant metric, remain to be demonstrated; effort may turn out to be better spent elsewhere.

Actually, the relative performance of existing solutions is already difficult to assess in most subdomains described above because of the lack of standard benchmarks (see also Irisson *et al.*, 2022 for plankton imaging; and Politikos *et al.*, 2023 for macrolitter). Such benchmarks depend on the availability of published (and labelled) datasets. The field is progressing on that end, with the release of datasets on e.g. plankton (Sosik *et al.*, 2015) and fish (Fisher *et al.*, 2016) images, remotely sensed images (Kikaki *et al.*, 2022), or ship noise (Santos-Domínguez *et al.*, 2016). However, other datasets, such as images from electronic monitoring on ships, are gathered by private companies that aim to use them to develop and sell electronic monitoring solutions, which are either made mandatory by authorities or desired by fishing companies to reduce costs compared to human observers. For researchers, the effort of gathering and labelling a dataset consistently is often huge and makes some people reluctant to distributing the result openly, although the availability of referenceable repositories (e.g. Zenodo) and citation tracking via Digital Object Identifiers helps. After releasing datasets, the next steps would be to define evaluation metrics following guidelines for proper benchmarking (Weber *et al.*, 2019) and to provide tools to easily track the results of those, now comparable, studies. So, overall, releasing high-quality public datasets, defining benchmarking studies, and centralizing their results are necessary to assess (i) the current state of ML tools in marine ecology and (ii) the tradeoff between gains from new architectures and their cost in complexity.

Transferring innovations from computer sciences to marine ecology also depends largely on efficient collaboration across disciplines. However, establishing interdisciplinary research teams is difficult and takes a long time (Haapasaari *et al.*, 2012). Once again, public datasets are an efficient first step for marine ecologists to garner interest from computer scientists. For example, after the WHOI-Plankton dataset was released (Sosik *et al.*, 2015), it was used in many papers on this topic in computer science conferences. In the assembled database, about 15% of references are from such computer science conferences or engineering journals, but very few are from high-level ones. This can indicate that marine ecology questions have not gained enough interest from the ML research community to generate significant new developments that would be published at high-profile conferences, unlike other applications such as face recognition, customer tracking, or self-driving cars. It could also simply reflect differences in overall funding for research in those fields, linked to the potential commercial applications of some research. On the other hand, publishing highly technical ML papers in ecology journals can also be challenging, because of the scarcity of editors and reviewers who can assess both the importance of the ecological questions and the relevance of the methods used to address them. Still, problems such as estimating the stock sizes of fish species that feed human populations, the distribution of the litter we create, the composition of plankton that forms the basis of oceanic food webs, or the global export of the excess carbon we produce seem no less important than designing targeted ads; they should generate proportionate interest and funding (Blair *et al.*, 2019). Therefore, ways forward for ML in marine ecology include (i) long-term digitalization strategies by funding agencies to scale efforts to the stakes we face and (ii) raising awareness of those stakes among the public in general and computer scientists in particular.

Another way to advance the interplay between ML and marine ecology is to train a new generation of scientists at the intersection of these fields. Then, long-term changes in the strategies for funding allocation and career evaluation would be needed to foster such hybrid profiles. Indeed, garnering the simultaneous interest, on a common problem, of researchers currently specialized in either marine ecology or computer sciences is difficult. Challenges raised by marine ecologists can be perceived as not novel or generic enough to constitute research questions for computer scientists. New developments in computer science are often not immediately actionable by marine ecologists, as seen above for large transformers or image-text encoders. So, computer scientists may feel like service providers for ecologists and ecologists as simple data providers for computer scientists, which is satisfactory for neither. Recent reviews and perspectives, in ecology as a whole, actually show that the interaction can be beneficial for both parties. Several also point towards the need to train ecologists

in computer sciences, not the opposite (Olden *et al.*, 2008 for an older one; Christin *et al.*, 2019 for a recent one), notably because computer science students rarely choose careers in ecology and environment, in part due to differences in financial compensation or job security. Overall, we argue that interdisciplinary training and career paths are potential solutions to many of the current shortcomings of ML applications in marine ecology.

Acknowledgements

All authors acknowledge the support of ICES through the Working group on Machine Learning in Marine Science (WGMLEARN).

Supplementary data

Supplementary material is available at the *ICESJMS* online version of the manuscript.

Conflict of interest

No conflict of interest was reported.

Authors contributions

AP, AG, BK, CJ, CP, DSG, DP, DDB, HM, JBR, JOI, JF, JTW, JAF, KM, KOM, KHK, LU, LVdB, ML, ODBP, PR, PS, RS, RML, SC, TC, and WM built and labelled the publications database. Contributions are then broken down by section, listing the main providers of content: section "What is machine learning and why does marine ecology need it?"—HM, JOI, KM, PR, and RK; section "A quick primer on machine learning"—JOI, KM, PR, and VS; section "The setup of the database and its tags"—JOI and PR; section "Benthos" AG, JF, and KH; section "Macrolitter"—DP and SV; section "Nekton"—HM, JBR, MS, and AT-MvH; section "Plankton"—HM, JOI, KOM, and RK; section "Identifying microorganisms from single-cell spectra"—PR; section "Active acoustics for target classification"—AP, HM, and NOH; section "Active acoustics for seabed and sediment mapping"—KHK; section "Passive acoustics monitoring"—CP, DDB, PD, and VS; section "Profiling biological communities with environmental genomics"—HM, KM, LVdB, SC, and TC; section "Quantifying and mapping fishing pressure from geolocation data"—JAF; section "Deriving biogeochemical variables from satellite images and floats profiles"—CJ and RS; section "Predicting species abundance and distribution"—SB and HW; section "Capturing dynamic ecological relationships"—LU and JAF; section "Summarizing ecosystems through regionalization"—SB and HW; section "Supporting human decisions on ecosystem management"—JAF; and section "Discussion and perspectives"—JOI. In addition, DSG, DP, HM, HW, JOI, JTW, JAF, ODBP, PR, RK, RML, SC, and SB reviewed the whole text. PR, KM, and JOI led the overall project.

Funding

TC acknowledges support from the Swiss National Science Foundation (#31003A_179125), the European Research Council (#818449 AGENSI), and the Horizon Europe programme (#101094924 ANERIS). JAF-S has received funding from the project H2020 FutureMARES (#869300) and SusTunTech (#869342). NOH acknowledges support from the CRIMAC centre funded by the Research Council of Norway #309512. KH is supported by the Mission Atlantic project funded by the European Union's Horizon 2020 Research and Innovation Programme (#862428). MS acknowledges funding from the European Union's H2020 programme #7553521 (SMARTFISH); European's Maritime and Fisheries Fund and the Danish Fisheries Agency, #33112-I-19-076 (AutoCatch); and Fully Documented Fisheries, funded by the European Maritime and Fisheries Fund (EMFF). LVdB acknowledges support from the Sand Fund of the Federal Public Service Economy. SC, RML, and SV acknowledge support from the H2020 project AtlantECO (#862923). RML acknowledges support from CNPq, Brazil (grant number 315033/2021-5). RK acknowledges support via a "Make Our Planet Great Again" grant of the French National Research Agency within the "Programme d'Investissements d'Avenir" (#ANR-19-MPGA-0012), by the Heisenberg programme of the German Science Foundation (#469175784), and from NOAA (#NA21OAR4310254). JBR acknowledges funding from the IFREMER Scientific Direction project DEEP. KM's participation was funded by the Norwegian Ministry of Trade, Industry and Fisheries. JOI acknowledges funding from the Belmont Forum project WWWPIC (#ANR-018-BELM-0003–01).

Data availability

The data underlying this article are available in the article and in its online supplementary material.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S *et al.* 2016. TensorFlow: large-scale machine learning on heterogeneous distributed systems. http://arxiv.org/abs/1603.04467.

Abrahms, B., Welch, H., Brodie, S., Jacox, M. G., Becker, E. A., Bograd, S. J., Irvine, L. M *et al.* 2019. Dynamic ensemble models to predict distributions and anthropogenic risk exposure for highly mobile species. Diversity and Distributions, 25: 1182–1193.

Alheit, J., Möllmann, C., Dutz, J., Kornilovs, G., Loewe, P., Mohrholz, V., and Wasmund, N. 2005. Synchronous ecological regime shifts in the central Baltic and the North Sea in the late 1980s. ICES Journal of Marine Science, 62: 1205–1215.

Allen, A. S., Yurk, H., Vagle, S., Pilkington, J., and Canessa, R. 2018. The underwater acoustic environment at SGaan Kinghlas–Bowie seamount marine protected area: characterizing vessel traffic and associated noise using satellite AIS and acoustic datasets. Marine Pollution Bulletin, 128: 82–88.

Allken, V., Handegard, N. O., Rosen, S., Schreyeck, T., Mahiout, T., and Malde, K. 2019. Fish species identification using a convolutional neural network trained on synthetic data. ICES Journal of Marine Science, 76: 342–349.

Althaus, F., Hill, N., Ferrari, R., Edwards, L., Przeslawski, R., Schönberg, C. H. L., Stuart-Smith, R *et al.* 2015. A standardised vocabulary for identifying benthic biota and substrata from underwater imagery: the CATAMI classification scheme. PLoS One, 10: e0141039.

Anderson, J. T., Holliday, D. V., Kloser, R., Reid, D. G., and Simard, Y. 2008. Acoustic seabed classification: current practice and future directions. ICES Journal of Marine Science, 65: 1004–1011.

Baker, R. E., Peña, J.-M., Jayamohan, J., and Jérusalem, A. 2018. Mechanistic models versus machine learning, a fight worth fighting for the biological community? Biology Letters, 14: 20170660.

Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S *et al.* 2020. Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 58: 82–115.

Beale, C. M., and Lennon, J. J. 2012. Incorporating uncertainty in predictive species distribution modelling. Philosophical Transactions of the Royal Society B: Biological Sciences, 367: 247–258.

Behivoke, F., Etienne, M.-P., Guitton, J., Randriatsara, R. M., Ranaivoson, E., and Léopold, M. 2021. Estimating fishing effort in small-scale fisheries using GPS tracking data and random forests. Ecological Indicators, 123: 107321.

Beijbom, O., Edmunds, P. J., Kline, D. I., Mitchell, B. G., and Kriegman, D. 2012. Automated annotation of coral reef survey images. IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, pp. 1170–1177.

Beijbom, O., Edmunds, P. J., Roelfsema, C., Smith, J., Kline, D. I., Neal, B. P., Dunlap, M. J *et al.* 2015. Towards automated annotation of benthic survey images: variability of human experts and operational modes of automation. PLoS One, 10: e0130312.

Benedetti, F., Jalabert, L., Sourisseau, M., Beker, B., Cailliau, C., Desnos, C., Elineau, A *et al.* 2019. The seasonal and inter-annual fluctuations of plankton abundance and community structure in a North Atlantic Marine Protected Area. Frontiers in Marine Science, 6: 214.

Benfield, M. C., Grosjean, P., Culverhouse, P. F., Irigoien, X., Sieracki, M. E., Lopez-Urrutia, A., Dam, H. G *et al.* 2007. RAPID: research on automated plankton identification. Oceanography, 20: 172–187.

Benkendorf, D. 2020. Effects of sample size and network depth on a deep learning approach to species distribution modeling. Ecological Informatics, 60: 101137.

Bergler, C., Schmitt, M., Cheng, R. X., Schröter, H., Maier, A., Barth, V., Weber, M *et al.* 2019. Deep representation learning for orca call type classification. *In* Text, Speech, and Dialogue, pp. 274–286. Ed. by K. Ekštein. Springer International Publishing, Cham.

Bermant, P. C. 2021. BioCPPNet: automatic bioacoustic source separation with deep neural networks. Scientific Reports, 11: 23502.

Bermant, P. C., Bronstein, M. M., Wood, R. J., Gero, S., and Gruber, D. F. 2019. Deep machine learning techniques for the detection and classification of sperm whale bioacoustics. Scientific Reports, 9: 12588.

Bertrand, S., Díaz, E., and Lengaigne, M. 2008. Patterns in the spatial distribution of peruvian anchovy (*Engraulis ringens*) revealed by spatially explicit fishing data. Progress in Oceanography, 79: 379–389.

Beslin, W. A. M., Whitehead, H., and Gero, S. 2018. Automatic acoustic estimation of sperm whale size distributions achieved through machine recognition of on-axis clicks. The Journal of the Acoustical Society of America, 144: 3485–3495.

Beyan, C., and Browman, H. I. 2020. Setting the stage for the machine intelligence era in marine science. ICES Journal of Marine Science, 77: 1267–1273.

Bittig, H. C., Steinhoff, T., Claustre, H., Fiedler, B., Williams, N. L., Sauzède, R., Körtzinger, A *et al.* 2018. An alternative to static climatologies: robust estimation of open ocean CO_2 variables and nutrient concentrations from T, S, and O_2 data using Bayesian neural networks. Frontiers in Marine Science, 5: 328.

Bittle, M., and Duncan, A. 2013. A review of current marine mammal detection and classification algorithms for use in automated passive acoustic monitoring. *In* Proceedings of Acoustics: Science, Technology and Amenity, pp. 1–8. Ed. by T. McMinn, Australian Acoustical Society, Victor Harbour, South Australia.

Blair, G. S., Henrys, P., Leeson, A., Watkins, J., Eastoe, E., Jarvis, S., and Young, P. J. 2019. Data science of the natural environment: a research roadmap. Frontiers in Environmental Science, 7: 121.

Bochinski, E., Bacha, G., Eiselein, V., Walles, T. J. W., Nejstgaard, J. C., and Sikora, T. 2019. Deep active learning for in situ plankton classification. *In* Pattern Recognition and Information Forensics, pp. 5–15. Ed. by Z. Zhang, D. Suter, Y. Tian, A. Branzan Albu, N. Sidère, and H. Jair Escalante. Springer International Publishing, Cham.

Boddy, L., Wilkins, M. F., and Morris, C. W. 2001. Pattern recognition in flow cytometry. Cytometry, 44: 195–209.

Bogucki, R., Cygan, M., Khan, C. B., Klimek, M., Milczek, J. K., and Mucha, M. 2019. Applying deep learning to right whale photo identification. Conservation Biology, 33: 676–684.

Bohan, D. A., Vacher, C., Tamaddoni-Nezhad, A., Raybould, A., Dumbrell, A. J., and Woodward, G. 2017. Next-generation global biomonitoring: large-scale, automated reconstruction of ecological networks. Trends in Ecology and Evolution, 32: 477–487.

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., Alexander, H *et al.* 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature biotechnology, 37: 852–857.

Boulais, O., Woodward, B., Schlining, B., Lundsten, L., Barnard, K., Bell, K. C., and Katija, K. 2020. FathomNet: an underwater image training database for ocean exploration and discovery. http://arxiv.org/abs/2007.00114.

Bowman, J. S., Amaral-Zettler, L. A., J Rich, J., M Luria, C., and Ducklow, H. W. 2017. Bacterial community segmentation facilitates the prediction of ecosystem function along the coast of the western Antarctic Peninsula. The ISME Journal, 11: 1460–1471.

Brajard, J., Santer, R., Crépon, M., and Thiria, S. 2012. Atmospheric correction of MERIS data for case-2 waters using a neuro-variational inversion. Remote Sensing of Environment, 126: 51–61.

Brautaset, O., Waldeland, A. U., Johnsen, E., Malde, K., Eikvil, L., Salberg, A.-B., and Handegard, N. O. 2020. Acoustic classification in multifrequency echosounder data using deep convolutional neural networks. ICES Journal of Marine Science, 77: 1391–1400.

Breiman, L. 2001. Random forests. Machine learning, 45: 5–32.

Brodie, S. J., Thorson, J. T., Carroll, G., Hazen, E. L., Bograd, S., Haltuch, M. A., Holsman, K. K *et al.* 2020. Trade-offs in covariate selection for species distribution models: a methodological comparison. Ecography, 43: 11–24.

Brodie, S., Jacox, M. G., Bograd, S. J., Welch, H., Dewar, H., Scales, K. L., Maxwell, S. M *et al.* 2018. Integrating dynamic subsurface habitat metrics into species distribution models. Frontiers in Marine Science, 5: 219.

Brown, C. J., Smith, S. J., Lawton, P., and Anderson, J. T. 2011. Benthic habitat mapping: a review of progress towards improved understanding of the spatial ecology of the seafloor using acoustic techniques. Estuarine, Coastal and Shelf Science, 92: 502–520.

Brown, J. C., Smaragdis, P., and Nousek-McGregor, A. 2010. Automatic identification of individual killer whales. The Journal of the Acoustical Society of America, 128: EL93–EL98.

Bruckmann, C., Müller, S., and Höner zu Siederdissen, C. 2022. Automatic, fast, hierarchical, and non-overlapping gating of flow cytometric data with flowEMMi v2. Computational and Structural Biotechnology Journal, 20: 6473–6489.

Cabreira, A. G., Tripode, M., and Madirolas, A. 2009. Artificial neural networks for fish-species identification. ICES Journal of Marine Science, 66: 1119–1129.

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. 2016. DADA2: high-resolution sample inference from Illumina amplicon data. Nature Methods, 13: 581–583.

Campbell, L., Henrichs, D. W., Olson, R. J., and Sosik, H. M. 2013. Continuous automated imaging-in-flow cytometry for detection and early warning of *Karenia brevis* blooms in the Gulf of Mexico. Environmental Science and Pollution Research, 20: 6896–6902.

Canals, M., Pham, C. K., Bergmann, M., Gutow, L., Hanke, G., van Sebille, E., Angiolillo, M *et al.* 2020. The quest for seafloor macrolitter: a critical review of background knowledge, current methods and future prospects. Environmental Research Letters, 16(2): 023001.

Carlson, C. J. 2020. embarcadero: species distribution modelling with Bayesian additive regression trees in R. Methods in Ecology and Evolution, 11: 850–858.

Chaffron, S., Delage, E., Budinich, M., Vintache, D., Henry, N., Nef, C., Ardyna, M *et al.* 2021. Environmental vulnerability of the global ocean epipelagic plankton community interactome. Science Advances, 7: eabg1921.

Chen, D. G., and Ware, D. M. 1999. A neural network model for forecasting fish stock recruitment. Canadian Journal of Fisheries and Aquatic Sciences, 56: 2385.

Choi, C., Kampffmeyer, M., Handegard, N. O., Salberg, A., Brautaset, O., Eikvil, L., and Jenssen, R. 2021. Semi-supervised target classification in multi-frequency echosounder data. ICES Journal of Marine Science, 78: 2615–2627,

Christin, S., Hervet, É., and Lecomte, N. 2019. Applications for deep learning in ecology. Methods in Ecology and Evolution, 10: 1632–1644.

Cimino, M. A., Santora, J. A., Schroeder, I., Sydeman, W., Jacox, M. G., Hazen, E. L., and Bograd, S. J. 2020. Essential krill species habitat resolved by seasonal upwelling and ocean circulation models within the large marine ecosystem of the California Current System. Ecography, 43: 1536–1549.

Claustre, H., Johnson, K. S., and Takeshita, Y. 2020. Observing the global ocean with Biogeochemical-Argo. Annual review of marine science, 12: 23–48.

Coccoli, C., Galparsoro, I., Murillas, A., Pınarbaşı, K., and Fernandes, J. A. 2018. Conflict analysis and reallocation opportunities in the framework of marine spatial planning: a novel, spatially explicit Bayesian belief network approach for artisanal fishing and aquaculture. Marine Policy, 94: 119–131.

Cordier, T., Forster, D., Dufresne, Y., Martins, C. I. M., Stoeck, T., and Pawlowski, J. 2018. Supervised machine learning outperforms taxonomy-based environmental DNA metabarcoding applied to biomonitoring. Molecular Ecology Resources, 18: 1381–1391.

Cortes, C., and Vapnik, V. 1995. Support vector machine. Machine learning, 20: 273–297.

Crespo, G. O., Dunn, D. C., Reygondeau, G., Boerder, K., Worm, B., Cheung, W., Tittensor, D. P *et al.* 2018. The environmental niche of the global high seas pelagic longline fleet. Science Advances, 4: eaat3681.

Cui, X., Yang, F., Wang, X., Ai, B., Luo, Y., and Ma, D. 2021. Deep learning model for seabed sediment classification based on fuzzy ranking feature optimization. Marine Geology, 432: 106390.

Culverhouse, P. F., Williams, R., Benfield, M., Flood, P. R., Sell, A. F., Mazzocchi, M. G., Buttino, I *et al.* 2006. Automatic image analysis of plankton: future perspectives. Marine Ecology Progress Series, 312: 297–309.

Culverhouse, P. F., Williams, R., Reguera, B., Herry, V., and González-Gil, S. 2003. Do experts make mistakes? A comparison of human and machine indentification of dinoflagellates. Marine Ecology Progress Series, 247: 17–25.

D'Amour, A., Heller, K., Moldovan, D., Adlam, B., Alipanahi, B., Beutel, A., and Chen, C., et al. 2020. Underspecification presents challenges for credibility in modern machine learning. http://arxiv.org/abs/2011.03395.

Dartnell, P., and Gardner, J. V. 2004. Predicting seafloor facies from multibeam bathymetry and backscatter data. Photogrammetric Engineering and Remote Sensing, 70: 1081–1091.

Dedman, S., Officer, R., Clarke, M., Reid, D. G., and Brophy, D. 2017. Gbm.auto: a software tool to simplify spatial modelling and marine protected area planning. PLoS One, 12: e0188955.

Deneu, B., Servajean, M., Bonnet, P., Botella, C., Munoz, F., and Joly, A. 2021. Convolutional neural networks improve species distribution modelling by capturing the spatial structure of the environment. PLoS Computational Biology, 17: e1008856.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. 2009. ImageNet: a large-scale hierarchical image database. *In* IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255. IEEE.

Derville, S., Torres, L. G., Iovan, C., and Garrigue, C. 2018. Finding the right fit: comparative cetacean distribution models using multiple data sources and statistical approaches. Diversity and Distributions, 24: 1657–1673.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M *et al.* 2021, An image is worth 16 × 16 words: transformers for image recognition at scale. http://arxiv.org/abs/2010.11929.

Douglass, L. L., Turner, J., Grantham, H. S., Kaiser, S., Constable, A., Nicoll, R., Raymond, B *et al.* 2014. A hierarchical classification of benthic biodiversity and assessment of protected areas in the Southern Ocean. PLoS One, 9: e100551.

Dreyfus-León, M., and Chen, D. G. 2007. Recruitment prediction with genetic algorithms with application to the Pacific Herring fishery. Ecological Modelling, 203: 141–146.

Dreyfus-Leon, M., and Kleiber, P. 2001. A spatial individual behaviour-based model approach of the yellowfin tuna fishery in the eastern Pacific Ocean. Ecological Modelling, 146: 47–56.

Ducharme-Barth, N. D., and Ahrens, R. N. M. 2017. Classification and analysis of VMS data in vertical line fisheries: incorporating uncertainty into spatial distributions. Canadian Journal of Fisheries and Aquatic Sciences, 74: 1749–1764.

Durden, J. M., Hosking, B., Bett, B. J., Cline, D., and Ruhl, H. A. 2021. Automated classification of fauna in seabed photographs: the impact of training and validation dataset size, with considerations for the class imbalance. Progress in Oceanography, 196: 102612.

Elith, J., and Leathwick, J. R. 2009. Species distribution models: ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution, and Systematics, 40: 677–697.

Elith, J., Leathwick, J. R., and Hastie, T. 2008. A working guide to boosted regression trees. Journal of Animal Ecology, 77: 802–813.

Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., and Yates, C. J. 2011. A statistical explanation of MaxEnt for ecologists: statistical explanation of MaxEnt. Diversity and Distributions, 17: 43–57.

Ellen, J. S., Graff, C. A., and Ohman, M. D. 2019. Improving plankton image classification using context metadata. Limnology and Oceanography: Methods, 17: 439–461.

Fallon, N. G., Fielding, S., and Fernandes, P. G. 2016. Classification of Southern Ocean krill and icefish echoes using random forests. ICES Journal of Marine Science, 73: 1998–2008.

Faust, K., and Raes, J. 2012. Microbial interactions: from networks to models. Nature Reviews Microbiology, 10: 538–550.

Fernandes, J. A., Irigoien, X., Goikoetxea, N., Lozano, J. A., Inza, I., Pérez, A., and Bode, A. 2010. Fish recruitment prediction, using robust supervised classification methods. Ecological Modelling, 221: 338–352.

Fernandes, J. A., Irigoien, X., Lozano, J. A., Inza, I., Goikoetxea, N., and Pérez, A. 2015. Evaluating machine-learning techniques for recruitment forecasting of seven North East Atlantic fish species. Ecological Informatics, 25: 35–42.

Fernandes, J. A., Kauppila, P., Uusitalo, L., Fleming-Lehtinen, V., Kuikka, S., and Pitkänen, H. 2012. Evaluation of reaching the targets of the water framework directive in the Gulf of Finland. Environmental Science and Technology, 46: 8220–8228.

Fernandes, J. A., Lozano, J. A., Inza, I., Irigoien, X., Pérez, A., and Rodríguez, J. D. 2013. Supervised pre-processing approaches in multiple class variables classification for fish recruitment forecasting. Environmental Modelling and Software, 40: 245–254.

Fernandes, J., Irigoien, X., Uriarte, A., Ibaibarriaga, L., Lozano, J., and Inza, I. 2009. Anchovy recruitment mixed long series prediction using supervised classification. Working document to the ICES Benchmark Workshop on Short-lived Species (WKSHORT).

Fernandes-Salvador, J. A., Davidson, K., Sourisseau, M., Revilla, M., Schmidt, W., Clarke, D., Miller, P. I., *et al.* 2021. Current status of forecasting toxic harmful algae for the north-east Atlantic shellfish aquaculture industry. Frontiers in Marine Science, 8: 666583.

Fincham, J. I., Wilson, C., Barry, J., Bolam, S., and French, G. 2020. Developing the use of convolutional neural networking in benthic habitat classification and species distribution modelling. ICES Journal of Marine Science, 77: 3074–3082.

Fisher, R. B., Chen-Burger, Y.-H., Giordano, D., Hardman, L., and Lin, F.-P. (Eds). 2016. Fish4Knowledge: Collecting and Analyzing Massive Coral Reef Fish Video Data. Intelligent Systems Reference Library. Springer International Publishing, Cham.

Flück, B., Mathon, L., Manel, S., Valentini, A., Dejean, T., Albouy, C., Mouillot, D *et al.* 2022. Applying convolutional neural networks to speed up environmental DNA annotation in a highly diverse ecosystem. Scientific Reports, 12: 10247.

Fournier, M., Casey Hilliard, R., Rezaee, S., and Pelot, R. 2018. Past, present, and future of the satellite-based automatic identification system: areas of applications (2004–2016). WMU Journal of Maritime Affairs, 17: 311–345.

Friedman, J., Hastie, T., and Tibshirani, R. 2001. The Elements of Statistical Learning. Springer, New York, NY.

Frühe, L., Cordier, T., Dully, V., Breiner, H.-W., Lentendu, G., Pawlowski, J., Martins, C *et al.* 2021. Supervised machine learning is superior to indicator value inference in monitoring the environmental impacts of salmon aquaculture using eDNA metabarcodes. Molecular Ecology, 30: 2988–3006.

Gallager, S. M. 2019. System for rapid assessment of water quality and harmful algal bloom toxins. US patent: UW 2019/0293565 A1.

Garcia, R., Prados, R., Quintana, J., Tempelaar, A., Gracias, N., Rosen, S., Vågstøl, H *et al.* 2020. Automatic segmentation of fish using deep learning with application to fish size measurement. ICES Journal of Marine Science, 77: 1354–1366.

Garcia-Garin, O., Monleón-Getino, T., López-Brosa, P., Borrell, A., Aguilar, A., Borja-Robalino, R., Cardona, L *et al.* 2021. Automatic detection and quantification of floating marine macro-litter in aerial images: introducing a novel deep learning approach connected to a web application in R. Environmental Pollution, 273: 116490.

Gibb, R., Browning, E., Glover-Kapfer, P., and Jones, K. E. 2019. Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring. Methods in Ecology and Evolution, 10: 169–185.

Gimpel, A., Stelzenmüller, V., Töpsch, S., Galparsoro, I., Gubbins, M., Miller, D., Murillas, A *et al.* 2018. A GIS-based tool for an integrated assessment of spatial planning trade-offs with aquaculture. Science of the Total Environment, 627: 1644–1655.

Girshick, R., Donahue, J., Darrell, T., and Malik, J. 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Columbus, OH, pp. 580–587.

Gómez-Ríos, A., Tabik, S., Luengo, J., Shihavuddin, A., Krawczyk, B., and Herrera, F. 2019. Towards highly accurate coral texture images classification using deep convolutional neural networks and data augmentation. Expert Systems with Applications, 118: 315–328.

Gonzalez, P., Alvarez, E., Diez, J., Lopez-Urrutia, A., and del Coz, J. J. 2017. Validation methods for plankton image classification systems. Limnology and Oceanography: Methods, 15: 221–237.

Gonzalez, P., Castano, A., Peacock, E. E., Diez, J., Jose Del Coz, J., and Sosik, H. M. 2019. Automatic plankton quantification using deep features. Journal of Plankton Research, 41: 449–463.

Goodwin, M., Halvorsen, K. T., Jiao, L., Knausgård, K. M., Martin, A. H., Moyano, M., Oomen, R. A *et al.* 2022. Unlocking the potential of deep learning for marine ecology: overview, applications, and outlook. ICES Journal of Marine Science, 79: 319–336.

Gorsky, G., Ohman, M. D., Picheral, M., Gasparini, S., Stemmann, L., Romagnan, J.-B., Cawood, A *et al.* 2010. Digital zooplankton image analysis using the ZooScan integrated system. Journal of Plankton Research, 32: 285–303.

Granado, I., Basurko, O. C., Rubio, A., Ferrer, L., Hernández-González, J., Epelde, I., and Fernandes, J. A. 2019. Beach litter forecasting on the south-eastern coast of the Bay of Biscay: a Bayesian networks approach. Continental Shelf Research, 180: 14–23.

Grorud-Colvert, K., Constant, V., Sullivan-Stack, J., Dziedzic, K., Hamilton, S. L., Randell, Z., Fulton-Bennett, H *et al.* 2019. High-profile international commitments for ocean protection: empty promises or meaningful progress? Marine Policy, 105: 52–66.

Grosjean, P., Picheral, M., Warembourg, C., and Gorsky, G. 2004. Enumeration, measurement, and identification of net zooplankton samples using the ZOOSCAN digital imaging system. ICES Journal of Marine Science, 61: 518–525.

Gugele, S. M., Widmer, M., Baer, J., DeWeber, J. T., Balk, H., and Brinker, A. 2021. Differentiation of two swim bladdered fish species using next generation wideband hydroacoustics. Scientific Reports, 11: 10520.

Guidi, L., Chaffron, S., Bittner, L., Eveillard, D., Larhlimi, A., Roux, S., Darzi, Y *et al.* 2016. Plankton networks driving carbon export in the oligotrophic ocean. Nature, 532: 465–470.

Guidi, L., Fernàndez-Guerra, A., Canchaya, C., Curry, E., Foglini, F., Irisson, J.-O., Malde, K *et al.* 2020. Big data in marine science. *In* Future Science Brief 6 of the European Marine Board. Ed. by B. Alexander, J. J. Heymans, A. Muniz Piniella, P. Kellett, and J. Coopman. European Marine Board, Ostend.

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., Boutte, C *et al.* 2013. The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. Nucleic Acids Research, 41: D597–D604.

Guisan, A., and Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. Ecology Letters, 8: 993–1009.

Haapasaari, P., Kulmala, S., and Kuikka, S. 2012. Growing into interdisciplinarity: how to converge biology, economics, and social science in fisheries research? Ecology and Society, 17: 6.

Hamilton, L. J. 2001. Acoustic seabed classification systems. Report no. DSTO-TN-0401, Defence Science and Technolgy Organisation, Aeronautical and Martime Research Lab.

Hazen, E. L., Scales, K. L., Maxwell, S. M., Briscoe, D. K., Welch, H., Bograd, S. J., Bailey, H *et al.* 2018. A dynamic ocean management tool to reduce bycatch and support sustainable fisheries. Science Advances, 4: eaar3001.

Helmond, A. T. M., Mortensen, L. O., Plet-Hansen, K. S., Ulrich, C., Needle, C. L., Oesterwind, D., Kindt-Larsen, L *et al.* 2020. Electronic monitoring in fisheries: lessons from global experiences and future opportunities. Fish and Fisheries, 21: 162–189.

Hernández-González, J., Inza, I., Granado, I., Basurko, O. C., Fernandes, J. A., and Lozano, J. A. 2019. Aggregated outputs by linear models: an application on marine litter beaching prediction. Information Sciences, 481: 381–393.

Hindell, M. A., Reisinger, R. R., Ropert-Coudert, Y., Hückstädt, L. A., Trathan, P. N., Bornemann, H., Charrassin, J.-B *et al.* 2020. Tracking of marine predators to protect Southern Ocean ecosystems. Nature, 580: 87–92.

Hirama, Y., Yokoyama, S., Yamashita, T., Kawamura, H., Suzuki, K., and Wada, M. 2017. Discriminating fish species by an echo sounder in a set-net using a CNN. 21st Asia Pacific Symposium on Intelligent and Evolutionary Systems (IES), Hanoi, pp. 112–115.

Hobday, A. 2011. Defining dynamic pelagic habitats in oceanic waters off eastern Australia. Deep Sea Research Part II: Topical Studies in Oceanography, 58: 734–745.

Howell, K. L., Davies, J. S., Allcock, A. L., Braga-Henriques, A., Buhl-Mortensen, P., Carreiro-Silva, M., Dominguez-Carrió, C *et al.* 2019. A framework for the development of a global standardised marine taxon reference image database (SMarTaR-ID) to support image-based analyses. PLoS One, 14: e0218904.

Hu, J., Li, D., Duan, Q., Han, Y., Chen, G., and Si, X. 2012. Fish species classification by color, texture and multi-class support vector machine using computer vision. Computers and Electronics in Agriculture, 88: 133–140.

Hu, Q., and Davis, C. 2005. Automatic plankton image recognition with co-occurrence matrices and support vector machine. Marine Ecology Progress Series, 295: 21–31.

Hunt, T. N., Allen, S. J., Bejder, L., and Parra, G. J. 2020. Identifying priority habitat for conservation and management of Australian humpback dolphins within a marine protected area. Scientific Reports, 10: 14366.

Hyrkas, J., Clayton, S., Ribalet, F., Halperin, D., Virginia Armbrust, E., and Howe, B. 2016. Scalable clustering algorithms for continuous environmental flow cytometry. Bioinformatics, 32: 417–423.

Ierodiaconou, D., Monk, J., Rattray, A., Laurenson, L., and Versace, V. L. 2011. Comparison of automated classification techniques for predicting benthic biological communities using hydroacoustics and video observations. Continental Shelf Research, 31: S28–S38.

Inada, K., Matsuda, R., Fujiwara, C., Nomura, M., Tamon, T., Nishihara, I., Takao, T *et al.* 2001. Identification of plastics by infrared absorption using InGaAsP laser diode. Resources, Conservation and Recycling, 33: 131–146.

Ioannou, I., Gilerson, A., Gross, B., Moshary, F., and Ahmed, S. 2011. Neural network approach to retrieve the inherent optical properties of the ocean from observations of MODIS. Applied Optics, 50: 3168.

Ioannou, I., Gilerson, A., Gross, B., Moshary, F., and Ahmed, S. 2013. Deriving ocean color products using neural networks. Remote Sensing of Environment, 134: 78–91.

Irigoien, X., Fernandes, J. A., Grosjean, P., Denis, K., Albaina, A., and Santos, M. 2009. Spring zooplankton distribution in the Bay of Biscay from 1998 to 2006 in relation with anchovy recruitment. Journal of Plankton Research, 31: 1–17.

Irisson, J.-O., Ayata, S.-D., Lindsay, D. J., Karp-Boss, L., and Stemmann, L. 2022. Machine learning for the study of plankton and marine snow from images. Annual Review of Marine Science, 14: 277–301.

James, G., Witten, D., Hastie, T., and Tibshirani, R. 2013. An Introduction to Statistical Learning. Springer, New York, NY.

Jamet, C., Thiria, S., Moulin, C., and Crepon, M. 2005. Use of a neurovariational inversion for retrieving oceanic and atmospheric constituents from ocean color imagery: a feasibility study. Journal of Atmospheric and Oceanic Technology, 22: 460–475.

Jarvis, S., DiMarzio, N., Morrissey, R., and Moretti, D. 2008. A novel multi-class support vector machine classifier for automated classification of beaked whales and other small odontocetes. Canadian Acoustics, 36: 34–40.

Jordan, M. I., and Mitchell, T. M. 2015. Machine learning: trends, perspectives, and prospects. Science, 349: 255–260.

Katija, K., Orenstein, E., Schlining, B., Lundsten, L., Barnard, K., Sainz, G., Boulais, O *et al.* 2022. FathomNet: a global image database for enabling artificial intelligence in the ocean. Scientific Reports, 12: 15914.

Kerr, T., Clark, J. R., Fileman, E. S., Widdicombe, C. E., and Pugeault, N. 2020. Collaborative deep learning models to handle class imbalance in flowcam plankton imagery. IEEE Access, 8: 170013–170032.

Kikaki, K., Kakogeorgiou, I., Mikeli, P., Raitsos, D. E., and Karantzalos, K. 2022. MARIDA: a benchmark for marine debris detection from sentinel-2 remote sensing data. PLoS One, 17: e0262247.

Kiranyaz, S., Gabbouj, M., Pulkkinen, J., Ince, T., and Meissner, K. 2010. Network of evolutionary binary classifiers for classification and retrieval in macroinvertebrate databases. IEEE International Conference on Image Processing, pp. 2257–2260. IEEE.

Kiranyaz, S., Ince, T., Pulkkinen, J., Gabbouj, M., Ärje, J., Kärkkäinen, S., Tirronen, V *et al.* 2011. Classification and retrieval on macroinvertebrate image databases. Computers in Biology and Medicine, 41: 463–472.

Knudby, A. 2010. New approaches to modelling fish–habitat relationships. Ecological Modelling, 221: 503–511.

Knudby, A., LeDrew, E., and Brenning, A. 2010. Predictive mapping of reef fish species richness, diversity and biomass in Zanzibar using IKONOS imagery and machine-learning techniques. Remote Sensing of Environment, 114: 1230–1241.

Korneliussen, R. J. (Ed.) 2018. Acoustic target classification. ICES Cooperative Research Report, 344, 104 pp.

Kouw, W. M., and Loog, M. 2019. A review of domain adaptation without target labels. IEEE Transactions on Pattern Analysis and Machine Intelligence, 43: 766–785.

Kowarski, K. A., and Moors-Murphy, H. 2021. A review of big data analysis methods for baleen whale passive acoustic monitoring. Marine Mammal Science, 37(2), pp. 652–673.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. 2012. Imagenet classification with deep convolutional neural networks. *In* Advances in Neural Information Processing Systems, pp. 1097–1105. Ed. by F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinerger. Neural Information Processing Systems Foundation, Inc. (NeurIPS), Lake Tahoe, NV.

Kroodsma, D. A., Mayorga, J., Hochberg, T., Miller, N. A., Boerder, K., Ferretti, F., Wilson, A *et al.* 2018. Tracking the global footprint of fisheries. Science, 359: 904–908.

Kuhn, M., and Wickham, H. 2020. Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles. https://www.tidymodels.org (accessed 29 September 2022).

Kyathanahally, S. P., Hardeman, T., Reyes, M., Merz, E., Bulas, T., Brun, P., Pomati, F *et al.* 2022. Ensembles of data-efficient vision transformers as a new paradigm for automated classification in ecology. Scientific Reports, 12: 18590.

Landschützer, P., Gruber, N., Haumann, F. A., Rödenbeck, C., Bakker, D. C. E., Heuven, S. van, Hoppema, M *et al.* 2015. The reinvigoration of the Southern Ocean carbon sink. Science, 349: 1221–1224.

Langenkämper, D., van Kevelaer, R., Purser, A., and Nattkemper, T. W. 2020. Gear-induced concept drift in marine images and its effect on deep learning classification. Frontiers in Marine Science, 7: 506.

Langenkämper, D., Zurowietz, M., Schoening, T., and Nattkemper, T. W. 2017. BIIGLE 2.0—browsing and annotating large marine image collections. Frontiers in Marine Science, 4: 83.

Laurila-Pant, M., Mäntyniemi, S., Venesjärvi, R., and Lehikoinen, A. 2019. Incorporating stakeholders' values into environmental decision support: a Bayesian belief network approach. Science of The Total Environment, 697: 134026.

Le Guyader, D., Ray, C., and Brosset, D. 2016. Defining fishing grounds variability with automatic identification system (AIS). 2nd International Workshop on Maritime Flows and Networks (WIMAKS'16), p. 96.

LeCun, Y., Bengio, Y., and Hinton, G. 2015. Deep learning. Nature, 521: 436–444.

Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P. 1998. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86: 2278–2324.

Lee, H., Park, M., and Kim, J. 2016. Plankton classification on imbalanced large scale database via convolutional neural networks with transfer learning. 2016 IEEE International Conference on Image Processing (ICIP), pp. 3713–3717.

Lehikoinen, A., Olsson, J., Bergström, L., Bergström, U., Bryhn, A., Fredriksson, R., and Uusitalo, L. 2019. Evaluating complex relationships between ecological indicators and environmental factors in the Baltic Sea: a machine learning approach. Ecological Indicators, 101: 117–125.

Lekunberri, X., Ruiz, J., Quincoces, I., Dornaika, F., Arganda-Carreras, I., and Fernandes, J. A. 2022. Identification and measurement of tropical tuna species in purse seiner catches using computer vision and deep learning. Ecological Informatics, 67: 101495.

Li, Y., Guo, J., Guo, X., Hu, Z., and Tian, Y. 2021. Plankton detection with adversarial learning and a densely connected deep learning model for class imbalanced distribution. Journal of Marine Science and Engineering, 9: 636.

Lieshout, C., Oeveren, K., Emmerik, T., and Postma, E. 2020. Automated river plastic monitoring using deep learning and cameras. Earth and Space Science, 7: e2019EA000960.

Lima-Mendez, G., Faust, K., Henry, N., Decelle, J., Colin, S., Carcillo, F., Chaffron, S *et al.* 2015. Determinants of community structure in the global plankton interactome. Science, 348: 1262073.

Lin, T.-H., Yang, H.-T., Huang, J.-M., Yao, C.-J., Lien, Y.-S., Wang, P.-J., and Hu, F.-Y. 2019. Evaluating changes in the marine soundscape of an offshore wind farm via the machine learning-based

source separation. *In* 2019 IEEE Underwater Technology (UT), pp. 1–6.
Liu, H., Li, Q., Bai, Y., Yang, C., Wang, J., Zhou, Q., Hu, S *et al.* 2021. Improving satellite retrieval of oceanic particulate organic carbon concentrations using machine learning methods. Remote Sensing of Environment, 256: 112316.
Liu, Y., and Wang, S. 2021. A quantitative detection algorithm based on improved faster R-CNN for marine benthos. Ecological Informatics, 61: 101228.
Liu, Y., and Weisberg, R. H. 2011. A review of self-organizing map applications in meteorology and oceanography. *In* Self Organizing Maps—Applications and Novel Algorithm Design. Ed. by J. I Mwasiagi. InTech.
Liu, Y., Xu, J., Tao, Y., Fang, T., Du, W., and Ye, A. 2020. Rapid and accurate identification of marine microbes with single-cell Raman spectroscopy. The Analyst, 145: 3297–3305.
Longhurst, A., Sathyendranath, S., Platt, T., and Caverhill, C. 1995. An estimate of global primary production in the ocean from satellite radiometer data. Journal of Plankton Research, 17: 1245–1271.
Lopez-Vazquez, V., Lopez-Guede, J. M., Marini, S., Fanelli, E., Johnsen, E., and Aguzzi, J. 2020. Video image enhancement and machine learning pipeline for underwater animal detection and classification at cabled observatories. Sensors, 20: 726.
Lucas, T. C. D. 2020. A translucent box: interpretable machine learning in ecology. Ecological Monographs, 90: e01422.
Lumini, A., Nanni, L., and Maguolo, G. 2020. Deep learning for plankton and coral classification. Applied Computing and Informatics, 19: 265–283.
Madricardo, F., Ghezzo, M., Nesto, N., Mc Kiver, W. J., Faussone, G. C., Fiorin, R., Riccato, F *et al.* 2020. How to deal with seafloor marine litter: an overview of the state-of-the-art and future perspectives. Frontiers in Marine Science, 7: 505134.
Mahé, F., Rognes, T., Quince, C., de Vargas, C., and Dunthorn, M. 2015. Swarm v2: highly-scalable and high-resolution amplicon clustering. PeerJ, 3: e1420.
Malde, K., Handegard, N. O., Eikvil, L., and Salberg, A.-B. 2020. Machine intelligence and the data-driven future of marine science. ICES Journal of Marine Science, 77: 1274–1285.
Maldonado, A. D., Uusitalo, L., Tucker, A., Blenckner, T., Aguilera, P. A., and Salmerón, A. 2019. Prediction of a complex system with few data: evaluation of the effect of model structure and amount of data with dynamic bayesian network models. Environmental Modelling and Software, 118: 281–297.
Malfante, M., Mohammed, O., Gervaise, C., Mura, M. D., and Mars, J. I. 2018. Use of deep features for the automatic classification of fish sounds. 2018 OCEANS—MTS/IEEE Kobe Techno-Oceans (OTO), IEEE, pp. 1–5.
Marques, T. P., Rezvanifar, A., Cote, M., Albu, A. B., Ersahin, K., Mudge, T., and Gauthier, S. 2021. Detecting marine species in echograms via traditional, hybrid, and deep learning frameworks. 2020 25th International Conference on Pattern Recognition (ICPR), IEEE, pp. 5928–5935.
Marzuki, M. I., Gaspar, P., Garello, R., Kerbaol, V., and Fablet, R. 2018. Fishing gear identification from vessel-monitoring-system-based fishing vessel trajectories. IEEE Journal of Oceanic Engineering, 43: 689–699.
Mattei, F., Franceschini, S., and Scardi, M. 2018. A depth-resolved artificial neural network model of marine phytoplankton primary production. Ecological Modelling, 382: 51–62.
Mayot, N., D'Ortenzio, F., Ribera d'Alcalà, M., Lavigne, H., and Claustre, H. 2016. Interannual variability of the Mediterranean trophic regimes from ocean color satellites. Biogeosciences, 13: 1901–1917.
Mitchell, T. M. 1997. Machine Learning. McGraw-Hill, New York, NY.
Mitchell, T. M. 1999. Machine learning and data mining. Communications of the ACM, 42: 30–36.
Moreno-Torres, J. G., Raeder, T., Alaiz-Rodríguez, R., Chawla, N. V., and Herrera, F. 2012. A unifying view on dataset shift in classification. Pattern Recognition, 45: 521–530.
Muñoz, M., Reul, A., Vargas-Yáñez, M., Plaza, F., Bautista, B., García-Martínez, M. C., Moya, F *et al.* 2017. Fertilization and connectivity in the Garrucha Canyon (SE-Spain) implications for marine spatial planning. Marine Environmental Research, 126: 45–68.
Murali, A., Bhargava, A., and Wright, E. S. 2018. IDTAXA: a novel approach for accurate taxonomic classification of microbiome sequences. Microbiome, 6: 140.
Niu, H., Reeves, E., and Gerstoft, P. 2017. Source localization in an ocean waveguide using supervised machine learning. The Journal of the Acoustical Society of America, 142: 1176–1188.
NOAA. 2014. Report on the occurrence and health effects of anthropogenic debris ingested by marine organisms. Silver Spring, MD. 19 pp.
NOAA. 2016. Report on modeling oceanic transport of floating marine debris. Silver Spring, MD. 21 pp.
Olden, J. D., Lawler, J. J., and Poff, N. L. 2008. Machine learning methods without tears: a primer for ecologists. The Quarterly Review of Biology, 83: 171–193.
Orenstein, E. C., and Beijbom, O. 2017. Transfer learning and deep feature extraction for planktonic image data sets. *In* 2017 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 1082–1088.
Orenstein, E. C., Kenitz, K. M., Roberts, P. L. D., Franks, P. J. S., Jaffe, J. S., and Barton, A. D. 2020. Semi- and fully supervised quantification techniques to improve population estimates from machine classifiers. Limnology and Oceanography: Methods, 18: 739–753.
Ozanich, E., Thode, A., Gerstoft, P., Freeman, L. A., and Freeman, S. 2021. Deep embedded clustering of coral reef bioacoustics. The Journal of the Acoustical Society of America, 149: 2587–2601.
Özel Duygan, B. D., Hadadi, N., Babu, A. F., Seyfried, M., and van der Meer, J. R. 2020. Rapid detection of microbiota cell type diversity using machine-learned classification of flow cytometry data. Communications Biology, 3: 379.
Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T *et al.* 2019. Pytorch: an imperative style, high-performance deep learning library. Advances in Neural Information Processing Systems, 32.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M *et al.* 2011. Scikit-learn: machine learning in Python. Journal of Machine Learning Research, 12: 2825–2830.
Peña, M. 2018. Robust clustering methodology for multi-frequency acoustic data: a review of standardization, initialization and cluster geometry. Fisheries Research, 200: 49–60.
Phillips, L. R., Carroll, G., Jonsen, I., Harcourt, R., and Roughan, M. 2020. A water mass classification approach to tracking variability in the east Australian current. Frontiers in Marine Science, 7: 365.
Phillips, S. J., Anderson, R. P., and Schapire, R. E. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling, 190: 231–259.
Phillips, S. J., and Dudík, M. 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography, 31: 161–175.
Picheral, M., Colin, S., and Irisson, J.-O. 2017. EcoTaxa, a tool for the taxonomic classification of images. http://ecotaxa.obs-vlfr.fr (accessed 29 September 2022).
Pichler, M., Boreux, V., Klein, A., Schleuning, M., and Hartig, F. 2020. Machine learning algorithms to infer trait-matching and predict species interactions in ecological networks. Methods in Ecology and Evolution, 11: 281–293.
Piechaud, N., Hunt, C., Culverhouse, P. F., Foster, N. L., and Howell, K. L. 2019. Automated identification of benthic epifauna with computer vision. Marine Ecology Progress Series, 615: 15–30.
Pınarbaşı, K., Galparsoro, I., Borja, Á., Stelzenmüller, V., Ehler, C. N., and Gimpel, A. 2017. Decision support tools in marine spatial

planning: present applications, gaps and future perspectives. Marine Policy, 83: 83–91.
Pınarbaşı, K., Galparsoro, I., Depellegrin, D., Bald, J., Pérez-Morán, G., and Borja, Á. 2019. A modelling approach for offshore wind farm feasibility with respect to ecosystem-based marine spatial planning. Science of the Total Environment, 667: 306–317.
Plonus, R.-M., Conradt, J., Harmer, A., Janssen, S., and Floeter, J. 2021. Automatic plankton image classification-Can capsules and filters help cope with data set shift? Limnology and Oceanography: Methods, 19: 176–195.
Politikos, D. V., Adamopoulou, A., Petasis, G., and Galgani, F. 2023. Using artificial intelligence to support marine macrolitter research: a content analysis and an online database. Ocean and Coastal Management, 233: 106466.
Politikos, D. V., Fakiris, E., Davvetas, A., Klampanos, I. A., and Papatheodorou, G. 2021. Automatic detection of seafloor marine litter using towed camera images and deep learning. Marine Pollution Bulletin, 164: 111974.
Ponsero, A. J., and Hurwitz, B. L. 2019. The promises and pitfalls of machine learning for detecting viruses in aquatic metagenomes. Frontiers in Microbiology, 10: 806.
Porskamp, P., Rattray, A., Young, M., and Ierodiaconou, D. 2018. Multiscale and hierarchical classification for benthic habitat mapping. Geosciences, 8: 119.
Preston, F. W. 1948. The commonness, and rarity, of species. Ecology, 29: 254–283.
Quang, D., and Xie, X. 2016. DanQ: a hybrid convolutional and recurrent deep neural network for quantifying the function of DNA sequences. Nucleic Acids Research, 44: e107–e107.
Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J *et al.* 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research, 41: D590–D596.
Quinn, T. P., Erb, I., Gloor, G., Notredame, C., Richardson, M. F., and Crowley, T. M. 2019. A field guide for the compositional analysis of any-omics data. GigaScience, 8: giz107.
Rajwa, B., Venkatapathi, M., Ragheb, K., Banada, P. P., Hirleman, E. D., Lary, T., and Robinson, J. P. 2008. Automated classification of bacterial particles in flow by multiangle scatter measurement and support vector machine classifier. Cytometry Part A, 73A: 369–379.
Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., and Prabhat. 2019. Deep learning and process understanding for data-driven Earth system science. Nature, 566: 195.
Reiss, H., Cunze, S., König, K., Neumann, H., and Kröncke, I. 2011. Species distribution modelling of marine benthos: a North Sea case study. Marine Ecology Progress Series, 442: 71–86.
Reygondeau, G., Guidi, L., Beaugrand, G., Henson, S. A., Koubbi, P., MacKenzie, B. R., Sutton, T. T *et al.* 2018. Global biogeochemical provinces of the mesopelagic zone. Journal of Biogeography, 45: 500–514.
Rezvanifar, A., Marques, T. P., Cote, M., Albu, A. B., Slonimer, A., Tolhurst, T., and Ersahin, K., et al. 2019. A deep learning-based framework for the detection of schools of herring in echograms. arXiv:1910.08215 [cs, eess, stat]. http://arxiv.org/abs/1910.08215 (accessed 6 August 2021).
Richards, C., Cooke, R. S., Bowler, D. E., Boerder, K., and Bates, A. E. 2021. Bycatch mitigation could prevent strong changes in the ecological strategies of seabird communities across the globe. https://doi.org/10.1101/2021.05.24.445481.
Robards, M. D., Silber, G. K., Adams, J. D., Arroyo, J., Lorenzini, D., Schwehr, K., and Amos, J. 2016. Conservation science and policy applications of the marine vessel Automatic Identification System (AIS)—a review. Bulletin of Marine Science, 92: 75–103.
Robinson, K. L., Luo, J. Y., Sponaugle, S., Guigand, C., and Cowen, R. K. 2017. A tale of two crowds: public engagement in plankton classification. Frontiers in Marine Science, 4: 82.
Roch, M. A., Klinck, H., Baumann-Pickering, S., Mellinger, D. K., Qui, S., Soldevilla, M. S., and Hildebrand, J. A. 2011. Classification of echolocation clicks from odontocetes in the Southern California Bight. The Journal of the Acoustical Society of America, 129: 467–475.
Roch, M. A., Lindeneau, S., Aurora, G. S., Frasier, K. E., Hildebrand, J. A., Glotin, H., and Baumann-Pickering, S. 2021. Using context to train time-domain echolocation click detectors. The Journal of the Acoustical Society of America, 149: 3301–3310.
Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. 2016. VSEARCH: a versatile open source tool for metagenomics. PeerJ, 4: e2584.
Rose, G. A., and Leggett, W. C. 1988. Hydroacoustic signal classification of fish schools by species. Canadian Journal of Fisheries and Aquatic Sciences, 45: 597–604.
Roshan, S., and DeVries, T. 2017. Efficient dissolved organic carbon production and export in the oligotrophic ocean. Nature Communications, 8: 2036.
Rowell, T. J., Demer, D. A., Aburto-Oropeza, O., Cota-Nieto, J. J., Hyde, J. R., and Erisman, B. E. 2017. Estimating fish abundance at spawning aggregations from courtship sound levels. Scientific Reports, 7: 3340.
Rubbens, P., and Props, R. 2021. Computational analysis of microbial flow cytometry data. mSystems, 6: e00895–20,.
Rubbens, P., Props, R., Boon, N., and Waegeman, W. 2017. Flow cytometric single-cell identification of populations in synthetic bacterial communities. PLoS One, 12: e0169754.
Rubbens, P., Props, R., Kerckhof, F.-M., Boon, N., and Waegeman, W. 2021. PhenoGMM: Gaussian mixture modeling of cytometry data quantifies changes in microbial community structure. mSphere, 6: e00530–20,.
Rudin, C. 2019. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1: 206–215.
Russo, T., Franceschini, S., D'Andrea, L., Scardi, M., Parisi, A., and Cataudella, S. 2019. Predicting fishing footprint of trawlers from environmental and fleet data: an application of artificial neural networks. Frontiers in Marine Science, 6: 670.
Russo, T., Parisi, A., Garofalo, G., Gristina, M., Cataudella, S., and Fiorentino, F. 2014. SMART: a spatially explicit bio-economic model for assessing and managing demersal fisheries, with an application to italian trawlers in the strait of sicily. PLoS One, 9: e86222.
Russo, T., Parisi, A., Prorgi, M., Boccoli, F., Cignini, I., Tordoni, M., and Cataudella, S. 2011. When behaviour reveals activity: assigning fishing effort to métiers based on VMS data using artificial neural networks. Fisheries Research, 111: 53–64.
Samuel, A. L. 1959. Some studies in machine learning using the game of checkers. IBM Journal of Research and Development, 3: 210–229.
Sander, E. L., Wootton, J. T., and Allesina, S. 2017. Ecological network inference from long-term presence–absence data. Scientific Reports, 7: 7154.
Santora. 2012. Spatial ecology of krill, micronekton and top predators in the central California Current: implications for defining ecologically important areas. Progress in Oceanography, 106: 154–174.
Santos-Domínguez, D., Torres-Guijarro, S., Cardenal-López, A., and Pena-Gimenez, A. 2016. ShipsEar: an underwater vessel noise database. Applied Acoustics, 113: 64–69.
Sauzède, R., Claustre, H., Jamet, C., Uitz, J., Ras, J., Mignot, A., and D'Ortenzio, F. 2015a. Retrieving the vertical distribution of chlorophyll a concentration and phytoplankton community composition from in situ fluorescence profiles: a method based on a neural network with potential for global-scale applications. Journal of Geophysical Research: Oceans, 120: 451–470.
Sauzède, R., Lavigne, H., Claustre, H., Uitz, J., Schmechtig, C., d'Ortenzio, F., Guinet, C. *et al.* 2015b. Vertical distribution of chlorophyll a concentration and phytoplankton community composition from in situ fluorescence profiles: a first database for the global ocean. Earth System Science Data, 7: 261–273.
Sauzède, R., Claustre, H., Uitz, J., Jamet, C., Dall'Olmo, G., D'Ortenzio, F., Gentili, B *et al.* 2016. A neural network-based method for merging ocean color and Argo data to extend surface

bio-optical properties to depth: retrieval of the particulate backscattering coefficient. Journal of Geophysical Research: Oceans, 121: 2552–2571.

Sauzède, R., Bittig, H. C., Claustre, H., Pasqueron de Fommervault, O., Gattuso, J.-P., Legendre, L., and Johnson, K. S. 2017. Estimates of water-column nutrient concentrations and carbonate system parameters in the global ocean: a novel approach based on neural networks. Frontiers in Marine Science, 4: 128.

Schmarje, L., Santarossa, M., Schröder, S.-M., Zelenka, C., Kiko, R., Stracke, J., Volkmann, N *et al.* 2022. A data-centric approach for improving ambiguous labels with combined semi-supervised classification and clustering. Computer Vision—ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Part VIII, pp. 363–380. Berlin, Heidelberg. https://doi.org/10.1007/978-3-031-20074-8_21 (accessed 29 November 2022).

Schroeder, S.-M., Kiko, R., and Koch, R. 2020. MorphoCluster: efficient annotation of plankton images by clustering. Sensors, 20: 3060. Basel.

Schröter, H., Nöth, E., Maier, A., Cheng, R., Barth, V., and Bergler, C. 2019. Segmentation, classification, and visualization of orca calls using deep learning. 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 8231–8235.

Sgier, L., Freimann, R., Zupanic, A., and Kroll, A. 2016. Flow cytometry combined with viSNE for the analysis of microbial biofilms and detection of microplastics. Nature Communications, 7: 11587.

Shafait, F., Mian, A., Shortis, M., Ghanem, B., Culverhouse, P. F., Edgington, D., Cline, D *et al.* 2016. Fish identification from videos captured in uncontrolled underwater environments. ICES Journal of Marine Science: Journal du Conseil, 73: 2737–2746.

Shang, Y., and Li, J. 2018. Study on echo features and classification methods of fish species. 2018 10th International Conference on Wireless Communications and Signal Processing (WCSP), pp. 1–6.

Shao, H., Kiyomoto, S., Kawauchi, Y., Kadota, T., Nakagawa, M., Yoshimura, T., Yamada, H *et al.* 2021. Classification of various algae canopy, algae turf, and barren seafloor types using a scientific echosounder and machine learning analysis. Estuarine, Coastal and Shelf Science, 255: 107362.

Sharma, G., Umapathy, K., and Krishnan, S. 2020. Trends in audio signal feature extraction methods. Applied Acoustics, 158: 107020.

Shin, D. 2021. The effects of explainability and causability on perception, trust, and acceptance: implications for explainable AI. International Journal of Human-Computer Studies, 146: 102551.

Smith, J. A., Tommasi, D., Welch, H., Hazen, E. L., Sweeney, J., Brodie, S., Muhling, B *et al.* 2021. Comparing dynamic and static time-area closures for bycatch mitigation: a management strategy evaluation of a swordfish fishery. Frontiers in Marine Science, 8: 630607.

Smoliński, S., and Radtke, K. 2017. Spatial prediction of demersal fish diversity in the Baltic Sea: comparison of machine learning and regression-based techniques. ICES Journal of Marine Science, 74: 102–111.

Solsona-Berga, A., Frasier, K. E., Baumann-Pickering, S., Wiggins, S. M., and Hildebrand, J. A. 2020. DetEdit: a graphical user interface for annotating and editing events detected in long-term acoustic monitoring data. PLoS Computational Biology, 16: e1007598.

Sonnewald, M., Dutkiewicz, S., Hill, C., and Forget, G. 2020. Elucidating ecological complexity: unsupervised learning determines global marine eco-provinces. Science Advances, 6: eaay4740.

Soriano, M., Marcos, S., Saloma, C., Quibilan, M., and Alino, P. 2001. Image classification of coral reef components from underwater color video. *In* MTS/IEEE Oceans 2001. An Ocean Odyssey. Conference Proceedings (IEEE Cat. No.01CH37295), pp. 1008–1013 vol. 2.

Sosik, H. M., and Olson, R. J. 2007. Automated taxonomic classification of phytoplankton sampled with imaging-in-flow cytometry. Limnology and Oceanography: Methods, 5: 204–216.

Sosik, H. M., Peacock, E. E., and Brownlee, E. F. 2015. WHOI plankton, annotated plankton images—Data set for developing and evaluating classification methods. http://hdl.handle.net/1912/7341 (accessed 10 April 2021).

Soykan, C. U., Eguchi, T., Kohin, S., and Dewar, H. 2014. Prediction of fishing effort distributions using boosted regression trees. Ecological Applications, 24: 71–83.

Spampinato, C., Giordano, D., Di Salvo, R., Chen-Burger, Y.-H. J., Fisher, R. B., and Nadarajan, G. 2010. Automatic fish classification for underwater species behavior understanding. Proceedings of the First ACM International Workshop on Analysis and Retrieval of Tracked Events and Motion in Imagery Streams, pp. 45–50. New York, NY. http://doi.acm.org/10.1145/1877868.1877881 (accessed 7 August 2019).

Stephens, D., and Diesing, M. 2014. A comparison of supervised classification methods for the prediction of substrate type using multibeam acoustic and legacy grain-size data. PLoS One, 9: e93950.

Stewart, W. K., Jiang, M., and Marra, M. 1994. A neural network approach to classification of sidescan sonar imagery from a midocean ridge area. IEEE Journal of Oceanic Engineering, 19: 214–224.

Stock, A., and Subramaniam, A. 2020. Accuracy of empirical satellite algorithms for mapping phytoplankton diagnostic pigments in the open ocean: a supervised learning perspective. Frontiers in Marine Science, 7: 599.

Stock, B. C., Ward, E. J., Eguchi, T., Jannot, J. E., Thorson, J. T., Feist, B. E., and Semmens, B. X. 2020. Comparing predictions of fisheries bycatch using multiple spatiotemporal species distribution model frameworks. Canadian Journal of Fisheries and Aquatic Sciences, 77: 146–163.

Storbeck, F., and Daan, B. 2001. Fish species recognition using computer vision and a neural network. Fisheries Research, 51: 11–15.

Stowell, D. 2022. Computational bioacoustics with deep learning: a review and roadmap. PeerJ, 10: e13152. [CrossRef]

Suikkanen, S., Uusitalo, L., Lehtinen, S., Lehtiniemi, M., Kauppila, P., Mäkinen, K., and Kuosa, H. 2021. Diazotrophic cyanobacteria in planktonic food webs. Food Webs, 28: e00202.

Taconet, M., Kroodsma, D., and Fernandes, J. A., Food and Agriculture Organization of the United Nations, Global Fishing Watch, AZTI-Tecnalia, and Seychelles Fishing Authority. 2019. Global atlas of AIS-based fishing activity: challenges and opportunities. 395pp. www.fao.org/3/ca7012en/ca7012en.pdf (accessed 29 September 2022).

Tang, W., Li, Z., and Cassar, N. 2019. Machine learning estimates of global marine nitrogen fixation. Journal of Geophysical Research: Biogeosciences, 124: 717–730.

Tang, X., Stewart, W. K., Huang, H., Gallager, S. M., Davis, C. S., Vincent, L., and Marra, M. 1998. Automatic plankton image recognition. Artificial Intelligence Review, 12: 177–199.

Tanhua, T., Pouliquen, S., Hausman, J., O'Brien, K., Bricher, P., de Bruin, T., Buck, J. J. H., *et al.* 2019. Ocean FAIR data services. Frontiers in Marine Science, 6: 440.

Thomas, M. K., Fontana, S., Reyes, M., and Pomati, F. 2018. Quantifying cell densities and biovolumes of phytoplankton communities and functional groups using scanning flow cytometry, machine learning and unsupervised clustering. PLoS One, 13: e0196225.

Thomas, M., Martin, B., Kowarski, K., Gaudet, B., and Matwin, S. 2020. Marine mammal species classification using convolutional neural networks and a novel acoustic representation. *In* Machine Learning and Knowledge Discovery in Databases, pp. 290–305. Ed. by U. Brefeld, E. Fromont, A. Hotho, A. Knobbe, M. Maathuis, and C Robardet. Springer International Publishing, Cham.

Thoya, P., Maina, J., Möllmann, C., and Schiele, K. S. 2021. AIS and VMS ensemble can address data gaps on fisheries for marine spatial planning. Sustainability, 13: 3769.

Trifonova, N., Kenny, A., Maxwell, D., Duplisea, D., Fernandes, J., and Tucker, A. 2015. Spatio-temporal Bayesian network models with latent variables for revealing trophic dynamics and functional networks in fisheries ecology. Ecological Informatics, 30: 142–158.

Trifonova, N., Maxwell, D., Pinnegar, J., Kenny, A., and Tucker, A. 2017. Predicting ecosystem responses to changes in fisheries catch, temperature, and primary productivity with a dynamic Bayesian network model. ICES Journal of Marine Science, 74: 1334–1343.

Tseng, C.-H., and Kuo, Y.-F. 2020. Detecting and counting harvested fish and identifying fish types in electronic monitoring system videos

using deep convolutional neural networks. ICES Journal of Marine Science, 77: 1367–1378.

Uusitalo, L., Fernandes, J. A., Bachiller, E., Tasala, S., and Lehtiniemi, M. 2016. Semi-automated classification method addressing marine strategy framework directive (MSFD) zooplankton indicators. Ecological Indicators, 71: 398–405.

Uusitalo, L., Tomczak, M. T., Müller-Karulis, B., Putnis, I., Trifonova, N., and Tucker, A. 2018. Hidden variables in a dynamic Bayesian network identify ecosystem level change. Ecological Informatics, 45: 9–15.

Vacher, C., Tamaddoni-Nezhad, A., Kamenova, S., Peyrard, N., Moalic, Y., Sabbadin, R., Schwaller, L *et al.* 2016. Learning ecological networks from next-generation sequencing data. *In* Advances in Ecological Research, 54: 1–39.

Vestbo, S., Obst, M., Quevedo Fernandez, F. J., Intanai, I., and Funch, P. 2018. Present and potential future distributions of Asian horseshoe crabs determine areas for conservation. Frontiers in Marine Science, 5:164.

Villon, S., Mouillot, D., Chaumont, M., Darling, E. S., Subsol, G., Claverie, T., and Villéger, S. 2018. A deep learning method for accurate and fast identification of coral reef fishes in underwater images. Ecological Informatics, 48: 238–244.

Wang, Q., Garrity, G. M., Tiedje, J. M., and Cole, J. R. 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Applied and Environmental Microbiology, 73: 5261–5267.

Watanabe, J., Shao, Y., and Miura, N. 2019. Underwater and airborne monitoring of marine ecosystems and debris. Journal of Applied Remote Sensing, 13: 1.

Watling, J. I., Brandt, L. A., Bucklin, D. N., Fujisaki, I., Mazzotti, F. J., Romañach, S. S., and Speroterra, C. 2015. Performance metrics and variance partitioning reveal sources of uncertainty in species distribution models. Ecological Modelling, 309–310: 48–59.

Weber, L. M., Saelens, W., Cannoodt, R., Soneson, C., Hapfelmeier, A., Gardner, P. P., Boulesteix, A.-L *et al.* 2019. Essential guidelines for computational method benchmarking. Genome Biology, 20: 125.

Weilgart, L., and Whitehead, H. 1997. Group-specific dialects and geographical variation in coda repertoire in South Pacific sperm whales. Behavioral Ecology and Sociobiology, 40: 277–285.

Welch, H., Hazen, E. L., Bograd, S. J., Jacox, M. G., Brodie, S., Robinson, D., Scales, K. L *et al.* 2019. Practical considerations for operationalizing dynamic management tools. Journal of Applied Ecology, 56: 459–469.

Welch, H., and McHenry, J. 2018. Planning for dynamic process: an assemblage-level surrogate strategy for species seasonal movement pathways. Aquatic Conservation: Marine and Freshwater Ecosystems, 28: 337–350.

Welch, H., Pressey, R. L., Heron, S. F., Ceccarelli, D. M., and Hobday, A. J. 2016. Regimes of chlorophyll-a in the Coral Sea: implications for evaluating adequacy of marine protected areas. Ecography, 39: 289–304.

Welch, H., Pressey, R. L., and Reside, A. E. 2018. Using temporally explicit habitat suitability models to assess threats to mobile species and evaluate the effectiveness of marine protected areas. Journal for Nature Conservation, 41: 106–115.

White, T. D., Ong, T., Ferretti, F., Block, B. A., McCauley, D. J., Micheli, F., and De Leo, G. A. 2020. Tracking the response of industrial fishing fleets to large marine protected areas in the Pacific Ocean. Conservation Biology, 34: 1571–1578.

Wick, R. R., Judd, L. M., and Holt, K. E. 2019. Performance of neural network basecalling tools for Oxford nanopore sequencing. Genome Biology, 20: 129.

Williams, S., and Friedman, A. 2018. SQUIDLE+. http://squidle.acfr.usyd.edu.au (accessed 29 September 2022).

Wirbel, J., Zych, K., Essex, M., Karcher, N., Kartal, E., Salazar, G., Bork, P *et al.* 2021. Microbiome meta-analysis and cross-disease comparison enabled by the SIAMCAT machine learning toolbox. Genome Biology, 22: 93.

Yoon, B.-J. 2009. Hidden markov models and their applications in biological sequence analysis. Current Genomics, 10: 402–415.

Yuan, Q., Shen, H., Li, T., Li, Z., Li, S., Jiang, Y., Xu, H *et al.* 2020. Deep learning in environmental remote sensing: achievements and challenges. Remote Sensing of Environment, 241: 111716.

Zaugg, S., van der Schaar, M., Houégnigan, L., Gervaise, C., and André, M. 2010. Real-time acoustic classification of sperm whale clicks and shipping impulses from deep-sea observatories. Applied Acoustics, 71: 1011–1019.

Zhang, C., Selch, D., Xie, Z., Roberts, C., Cooper, H., and Chen, G. 2013. Object-based benthic habitat mapping in the Florida keys from hyperspectral imagery. Estuarine, Coastal and Shelf Science, 134: 88–97.

Zhao, Q., and Costello, M. J. 2019. Summer and winter ecosystems of the world ocean photic zone. Ecological Research, 34(4): 457–471.

Zion, B., Alchanatis, V., Ostrovsky, V., Barki, A., and Karplus, I. 2007. Real-time underwater sorting of edible fish species. Computers and Electronics in Agriculture, 56: 34–45.

Handling editor: Christopher Whidden