

Annual Review of Marine Science

Machine Learning for the Study of Plankton and Marine Snow from Images

Jean-Olivier Irisson,¹ Sakina-Dorothee Ayata,¹
Dhugal J. Lindsay,² Lee Karp-Boss,³
and Lars Stemmann¹

¹Laboratoire d'Océanographie de Villefranche, Sorbonne Université, CNRS, F-06230 Villefranche-sur-Mer, France; email: irisson@normalesup.org, sakina.ayata@normalesup.org, lars.stemmann@sorbonne-universite.fr

²Advanced Science-Technology Research (ASTER) Program, Institute for Extra-Cutting-Edge Science and Technology Avant-Garde Research (X-STAR), Japan Agency for Marine-Earth Science and Technology, Yokosuka, Kanagawa 237-0021, Japan; email: dhugal@jamstec.go.jp

³School of Marine Sciences, University of Maine, Orono, Maine 04469, USA; email: lee.karp-boss@maine.edu

Annu. Rev. Mar. Sci. 2022. 14:277–301

First published as a Review in Advance on August 30, 2021

The *Annual Review of Marine Science* is online at marine.annualreviews.org

<https://doi.org/10.1146/annurev-marine-041921-013023>

Copyright © 2022 by Annual Reviews.
All rights reserved

Keywords

plankton, marine snow, ecology, quantitative imaging, machine learning, classifier

Abstract

Quantitative imaging instruments produce a large number of images of plankton and marine snow, acquired in a controlled manner, from which the visual characteristics of individual objects and their in situ concentrations can be computed. To exploit this wealth of information, machine learning is necessary to automate tasks such as taxonomic classification. Through a review of the literature, we highlight the progress of those machine classifiers and what they can and still cannot be trusted for. Several examples showcase how the combination of quantitative imaging with machine learning has brought insights on pelagic ecology. They also highlight what is still missing and how images could be exploited further through trait-based approaches. In the future, we suggest deeper interactions with the computer sciences community, the adoption of data standards, and the more systematic sharing of databases to build a global community of pelagic image providers and users.

ANNUAL REVIEWS CONNECT

www.annualreviews.org

- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

1. INTRODUCTION

Plankton are an extremely diverse group of organisms (de Vargas et al. 2015), and its members are keystone components of Earth's biosphere. Photosynthetic plankton are responsible for approximately half of global primary production (Behrenfeld et al. 2001). The vertical displacement of planktonic organisms and the sinking of their detritus, referred to as marine snow, significantly contribute to the global carbon cycle through a multifaceted biological pump (Boyd et al. 2019). They are also critical components of marine food webs, contributing to the production of organic matter and its transfer to higher trophic levels (Frederiksen et al. 2006) and to the deep ocean and seafloor. Finally, because most planktonic species repeat relatively short life cycles embedded in the same water mass, they are very sensitive indicators of changes in that water mass (Hays et al. 2005). Therefore, the biomass and diversity of plankton were identified as Essential Ocean Variables by the Global Ocean Observing System, Essential Biodiversity Variables by the Group on Earth Observations Biodiversity Observation Network, and Essential Climate Variables by the Global Climate Observing System (Miloslavich et al. 2018, Muller-Karger et al. 2018, WMO et al. 2003).

Investigating planktonic organisms and marine snow at meaningful spatial and temporal scales in the vast and deep ocean is difficult. First, many instruments need to be combined to capture the whole ecosystem, because object sizes range from 0.2–2 μm for bacteria and picoeukaryotes to up to several meters for large siphonophores, and from a few micrometers for small particles to centimeters for marine snow aggregates (**Figure 1**). Second, plankton distribution is characterized by multiscale structuring—from seconds to years in time, from centimeters to hundreds of meters vertically, and from centimeters to hundreds of kilometers horizontally (Davis et al. 1992b, Mackas et al. 1985). Sampling both at high resolution and across large scales is extremely challenging. Yet accurately describing plankton patchiness is key for assessing biological rates and trophic interactions that control ecosystem production and biodiversity (Benoit-Bird & McManus 2012, Priyadarshi et al. 2019, Woodson & Litvin 2015). Finally, common plankton collection methods, such as nets, sediment traps, bottles, and pumps, are not adapted to resolve the smallest scales or to sample fragile organisms and marine snow (Remsen et al. 2004). In addition, they require the manual and often time-consuming sorting of the resulting samples, which makes them difficult to scale up to pan-oceanic observations. These difficulties led to the development of new technologies for plankton studies (Wiebe & Benfield 2003, Lombard et al. 2019).

Among those technologies, quantitative imaging systems have been identified as promising tools to yield consistent observations of plankton at a global scale, thanks, in part, to automation (Benfield et al. 2007, Lombard et al. 2019, Stemmann et al. 2012). These systems acquire a large number of images in a controlled manner, which enables (*a*) the extraction of some features of individual objects (size, opacity, etc.) that are consistent across all images and (*b*) the computation of in situ concentrations from a relatively small imaged volume (**Figure 2**). These characteristics accelerate the measurement and identification of objects in the images and allow the resulting concentrations to be exploited through classic numerical ecology analyses. Without being exhaustive (for more details, see Lombard et al. 2019), we remark that these instruments span a range of technologies and deployment modes. Many take monochromatic images using side illumination [Deep Particle Image Velocimetry (DeepPIV) (Katija et al. 2020) and the Underwater Vision Profiler (UVP) (Picheral et al. 2010)], shadowgraphy [the In Situ Ichthyoplankton Imaging System (ISIIS) (Cowen & Guigand 2008)], or holography (Sun et al. 2008). Some take color images [the Video Plankton Recorder (VPR) (Davis et al. 1992a) and Scripps Plankton Camera System (Orenstein et al. 2020b)], which provides richer information but often at the cost of image sharpness. These images can be light field [the ISIIS and Imaging FlowCytobot (IFCB) (Olson & Sosik 2007)] or dark field [the UVP, the VPR, the Scripps Plankton Camera System, and Lightframe

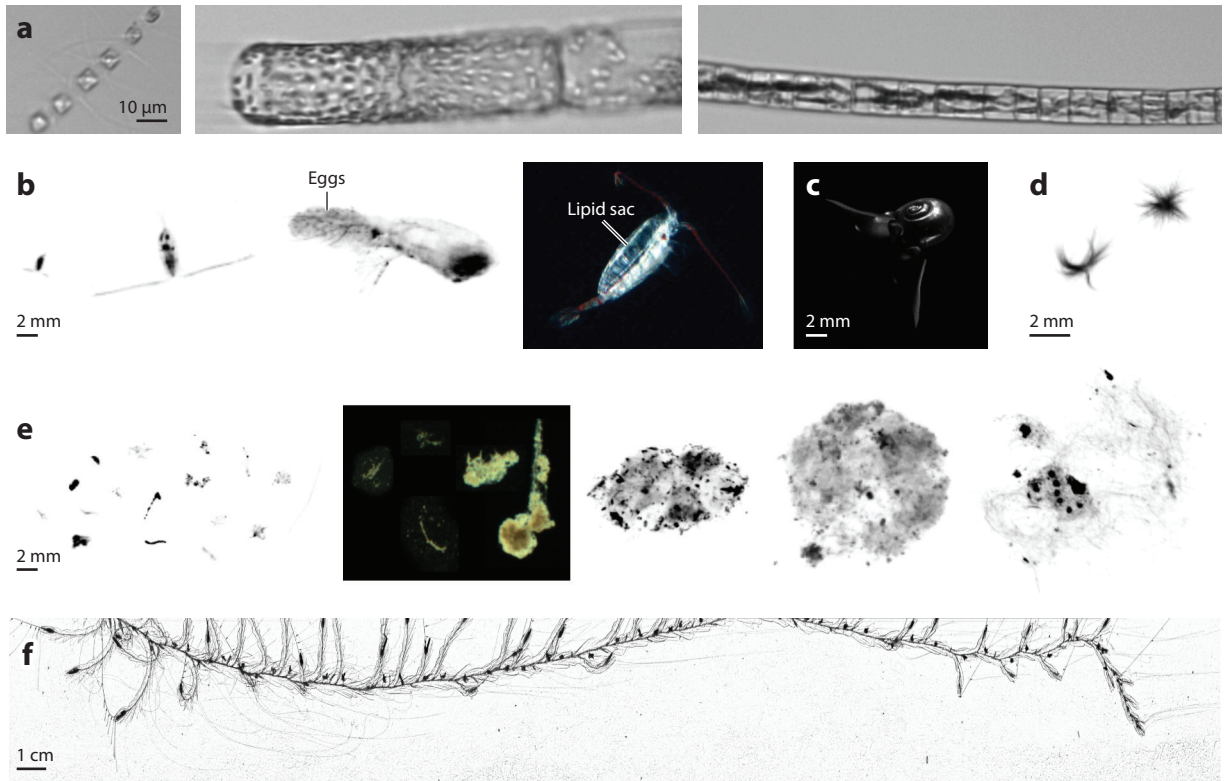


Figure 1

Images of plankton and marine snow particles sampled in situ by various instruments. Note the differences in scale between panels; each scale bar applies to all images in that panel. (a) Phytoplankton from the IFCB: a chain of *Chaetoceros* (with faint spines between cells; the cell size is $\sim 5 \mu\text{m}$) (left), *Guinardia* (center), and a segment of a *Trichodesmium* trichome (right). (b) Three copepods from the UVP (the last one bearing eggs, as labeled) (left) and one *Calanus hyperboreus* from the VPR (with red pigment and a well-developed lipid sac, as indicated) (right). (c) *Limacina* from the UVP (raw image, noninverted). (d) *Trichodesmium* colonies from the UVP. (e) Marine snow, in various shapes and sizes, with heterogeneous content and porosity, from the UVP (grayscale) and VPR (color). (f) A 40-cm-long portion of a siphonophore, likely *Forskalia*, from the ISIIS. Abbreviations: IFCB, Imaging FlowCytobot; ISIIS, In Situ Ichthyoplankton Imaging System; UVP, Underwater Vision Profiler; VPR, Video Plankton Recorder.

On-Sight Key Species Investigation (LOKI) (Schulz et al. 2010)]. Some instruments process samples in the laboratory [the ZooScan (Gorsky et al. 2010) and FlowCam (Sieracki et al. 1998)], some can be used in the laboratory or on in situ moorings [the IFCB and CytoBuoy (Dubelaar et al. 1999)], and some profile open waters (the UVP and ISIIS). In situ systems with low flow field perturbations are more appropriate to image fragile objects and thus allow the simultaneous detection of plankton and marine snow. Finally, many recent developments have been toward smart, miniaturized cameras mounted on remotely operated or autonomous vehicles, the spatiotemporal coverage of which is difficult to achieve with manned submersibles or ships (Durden et al. 2016, Ohman et al. 2019, Stemmann et al. 2012).

These instruments produce a huge amount of information—for example, 1 billion pixels containing ~ 2 million objects per year for a single ZooScan (~ 100 of which are now distributed worldwide) or 100 billion pixels corresponding to ~ 11 million objects per hour of deployment for the ISIIS. They have been progressively deployed throughout the world's oceans (Figure 3). The scientific use of this wealth of data became limited by our ability to recognize the imaged

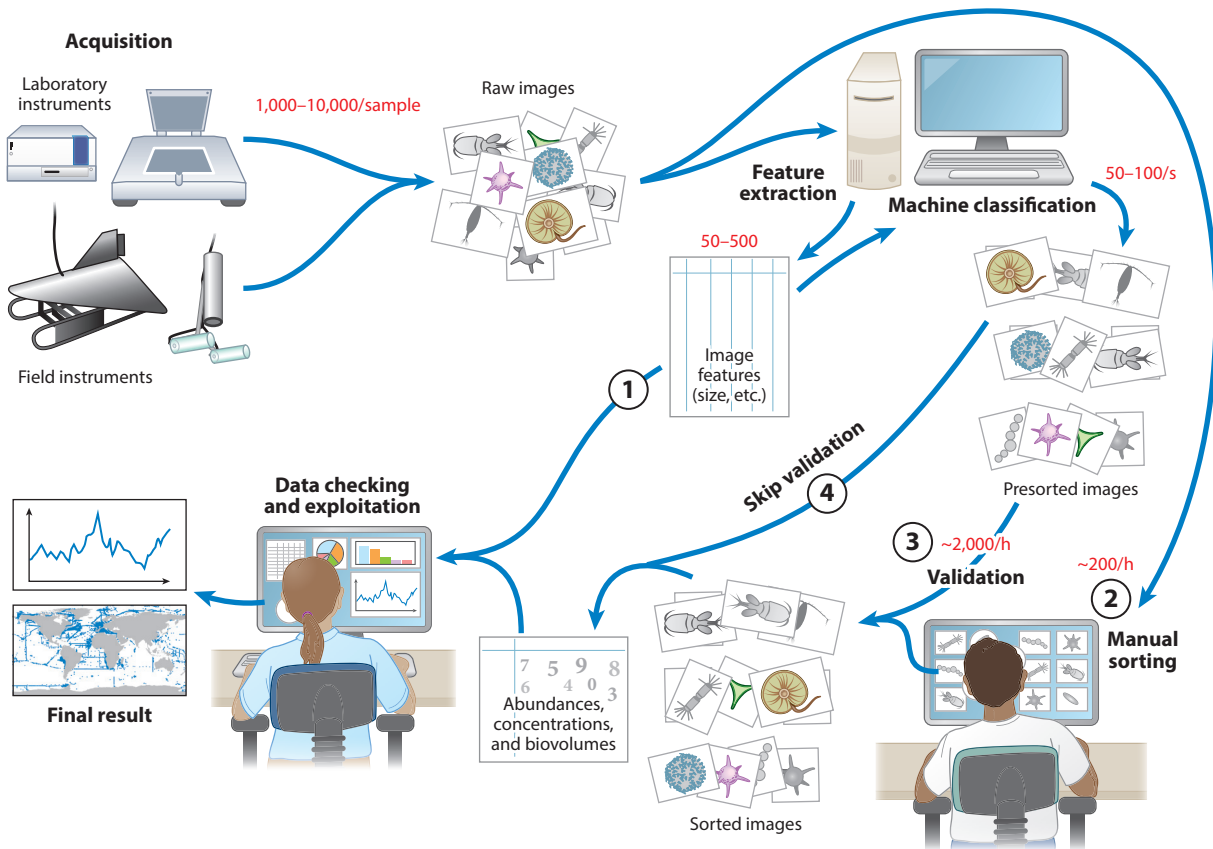


Figure 2

The plankton quantitative imaging process; the circled numbers identify four strategies for processing the raw images. High throughput (*red numbers*) from laboratory- or field-based image acquisition to ecological exploitation is allowed by tight interactions between instruments, computers, and human operators.

objects accurately and rapidly enough. Enter machine learning. The techniques are detailed in the next section, but, briefly, several strategies have emerged (as shown in **Figure 2**): (①) use of the features (size in particular) of all objects, with no classification at all (Lampitt et al. 1993); (②) manual identification of images by experts (Dennett et al. 2002); (③) machine classification followed by visual inspection and validation/correction of some taxa (Garcia-Comas et al. 2011) or all images (Biard et al. 2016); and (④) machine classification only (Irigoiien et al. 2009). In most cases, marine snow is treated as a single heterogeneous entity, despite its overwhelming numerical dominance, because defining objective categories is difficult (but see Schroeder et al. 2020).

The progress of automated techniques and the development of software to facilitate their use by non-computer scientists enabled the use of imaging data sets for research and ecosystem monitoring purposes. In this article, we take a step back and review the body of work at the intersection of machine learning and quantitative pelagic imaging. In particular, we (*a*) survey the literature on machine classification of plankton and marine snow images and compare a classic and a modern approach in a realistic case study, (*b*) examine a few studies to explain how machine learning applied to images has improved our understanding of pelagic ecosystems and highlight the promising use

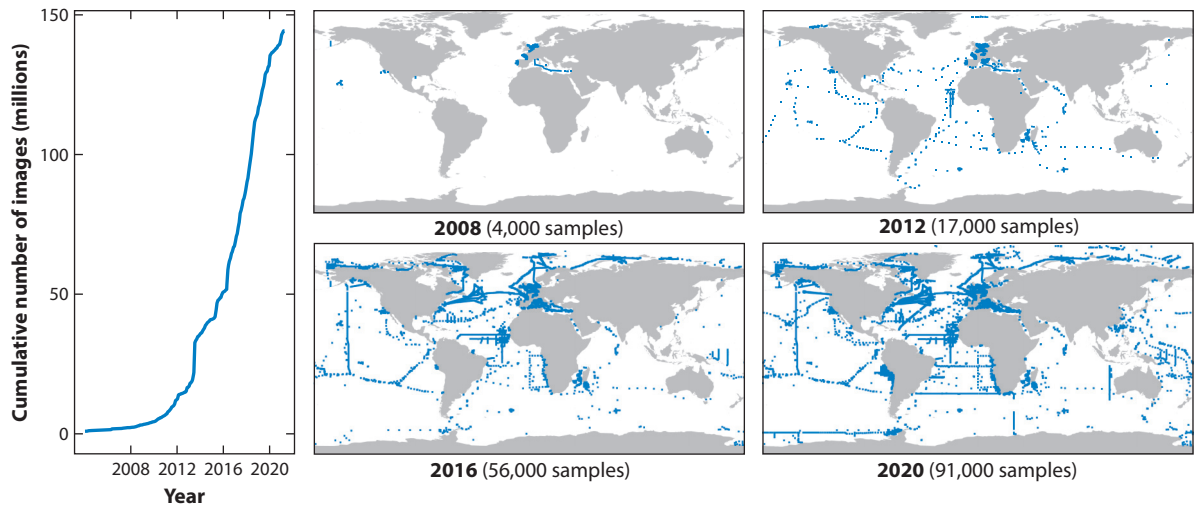


Figure 3

Accumulation of images and samples contained in a single repository (EcoTaxa; Picheral et al. 2017). Image data sets collected with different instruments (the UVP, IFCB, FlowCam, ZooScan, ISIIS, etc.) have been uploaded by an international community of users from more than 350 organizations. Abbreviations: IFCB, Imaging FlowCytobot; ISIIS, In Situ Ichthyoplankton Imaging System; UVP, Underwater Vision Profiler.

of image-derived traits to enable new functional ecology approaches, and (c) propose next steps for the future of plankton classifiers, the coordinated acquisition of massive imaging data sets, and the development of tools for quantitative imaging.

2. THE CLASSIFICATION OF PELAGIC IMAGES

2.1. A History of Machine Learning Approaches

Machine learning covers all techniques that learn patterns from a training data set and can then find the same patterns in another, independent, test data set. To classify images taxonomically and access the underlying ecological information (e.g., concentrations/biomass per taxon), most studies have used supervised classifiers, which learn to classify (i.e., give a name to) new images based on a set of images already classified by human experts.

We now tend to separate classic machine learning from deep learning (LeCun et al. 2015). In the classic approaches, the images are first processed by deterministic algorithms that extract information from them—the size of the organism, its average color or gray level, the complexity of its shape, its symmetry, and so on. Those features are said to be handcrafted because they indeed need to be crafted by a practitioner, who must assess or guess what is relevant to tell the various taxa apart. Then, the classification algorithm, such as a support vector machine (Cortes & Vapnik 1995) or a random forest (RF) (Breiman 2001), learns which combinations of feature values are associated with which taxonomic label. Deep learning for image classification is based on convolutional neural networks (CNNs) (Krizhevsky et al. 2012, Russakovsky et al. 2015). The first part of the network extracts features from the input image by computing convolutions (i.e., multiplication by a filter) over it; convolutions increase contrast, highlight edges, and so on. After several steps of convolution and reduction, the image is transformed into a vector of numbers: its deep features. These features are then used by a classifier, just like with classic learning; the classifier here is an artificial neural network. The main difference from classic machine learning is that the feature

Training data set: in the context of plankton images, a set of images classified into categories by experts, from which the algorithm will learn

Test data set: in the context of plankton images, an independent set of expert-classified images on which the predictive performance of the classifier is evaluated

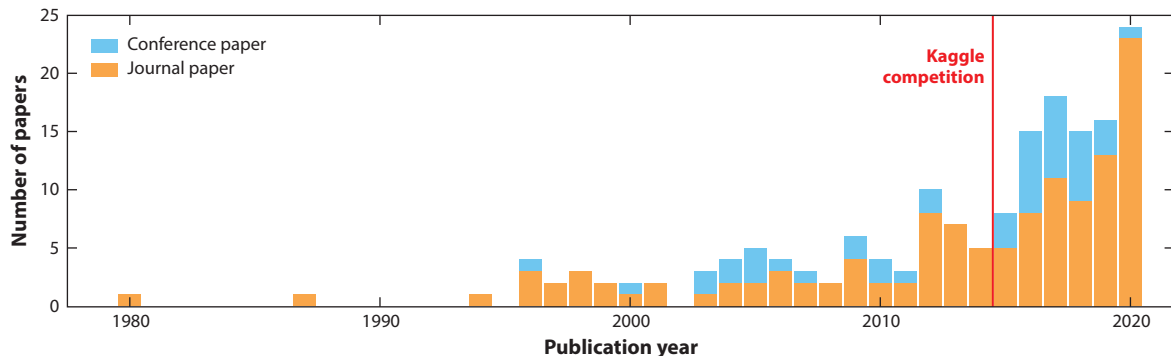


Figure 4

Number of papers dealing with the automated classification of plankton images published each year from 1980 through 2020 (174 total, 138 from journals and 36 from conferences), collected through a systematic literature review.

extractor and classifier are learned at the same time, from the training set. Therefore, the model tends toward an optimal combination of features for the classifier, and less domain knowledge is necessary to craft the very best features. This approach has proved hugely successful for a variety of problems.

To paint a picture of the history of machine learning for plankton image classification, we conducted a systematic literature review that yielded 174 publications (see **Supplemental Appendix 1** and the accompanying **Supplemental Data** containing the references in a commonly readable bibliographic file format). The first paper that used machine learning to classify plankton images dates from 1980 and performed pattern extraction on digital microscopy images to classify five genera of phytoplankton (Schlimpert et al. 1980). The literature started to build in the mid-1990s (**Figure 4**) but grew at a rate similar to that of the field of oceanography as a whole until 2012 (i.e., the yearly number of papers standardized by the total number in the field is flat). Starting in 2012, the yearly number of journal articles increased well beyond the expected growth of the field. After 2015, the number of conference proceedings papers, particularly in computer vision, increased significantly, likely because a competition on plankton image classification hosted by the popular Kaggle website (<https://kaggle.com/c/datasciencebowl>) provided an open data set and put it on the radar of the computer sciences community at large.

Before 2015, all papers used classic machine learning approaches. For example, Culverhouse et al. (1996) developed handcrafted features, based on transformations of the image (two Fourier transforms and an edge detector) and on its texture, to train an artificial neural network and classify microscopy images into 23 taxa of dinoflagellates. Trained and tested on a set of 5,000 images, the model achieved 83% accuracy, while human taxonomists achieved 85%. Tang et al. (1998) used image moments that describe global properties of the object (size, average color, etc.), a Fourier descriptor of the contour, and pattern features with a learning vector quantization classifier trained on a set of 1,000 images to sort 1,000 VPR images into six taxa with ~90% accuracy. However, the same technique applied to a real data set of ~20,000 images containing seven taxa resulted in only ~60–70% accuracy, because the proportions of taxa changed between the training set and the realistic test set (Davis et al. 2004). Switching to features based on gray-level co-occurrence matrices and support vector machines as the classifier increased the accuracy, to 70%, but only when compared with the worse model from 2004 (Hu & Davis 2005).

Three papers have largely shaped the field, because they are highly cited both by other machine learning papers and by ecology papers using the instruments and techniques they present

Supplemental Material >

Accuracy: the proportion of objects in a data set for which identifications are correctly predicted; it should be computed on an independent test set

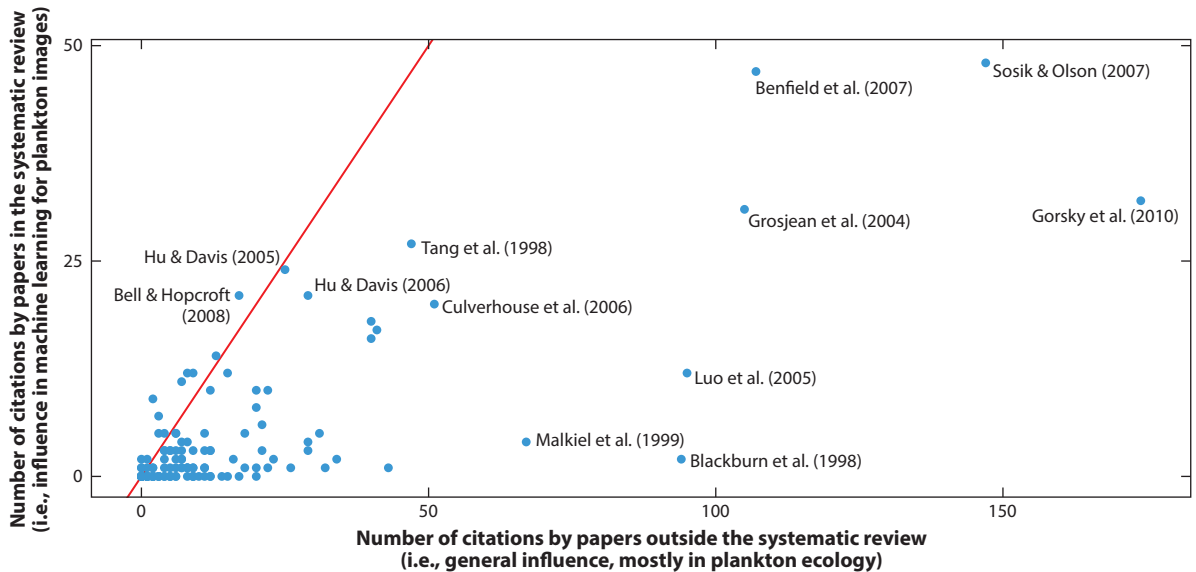


Figure 5

Distribution of the 138 journal articles within the systematic review of machine learning methods applied to plankton images, according to (*x* axis) how many times they were cited by papers outside of these 138 articles (a proxy for their overall influence, mostly in the field of plankton ecology) versus (*y* axis) how many times they were cited by another of the 138 articles (a measure of how influential they are in the specific field of machine learning for plankton images). Most papers are cited fewer than 10 times. Papers far to the right of the red 1:1 line were particularly influential, for reasons that go beyond their presentation of machine learning techniques.

(**Figure 5**): the pair of papers by Grosjean et al. (2004) and Gorsky et al. (2010) that presented the ZooScan and the paper by Sosik & Olson (2007) that presented the IFCB. The classification of ZooScan images is based mostly on global object features derived from image moments. Grosjean et al. (2004) tested many classifiers, and RF achieved ~85% accuracy on 8 broad taxa and ~75% on 29 more detailed ones, as measured by cross-validation on a data set of ~1,000 images. Gorsky et al. (2010) further showed that accuracy decreased from 80% to 65% when going from 5 to 35 categories while controlling for the size of the training data set (300 images per category). Sosik & Olson (2007) derived a much larger set of features based on image moments, general shape and symmetry, and gray-level co-occurrence matrices as well as specialized approaches aimed at detecting spines radiating from phytoplankton cells. Using a support vector machine classifier trained on 3,300 images in 23 taxa, they achieved 88% accuracy on a test set of 3,330 other images.

Overall, in their review, Benfield et al. (2007) concluded that 70% and 80% accuracy could be reached on data sets with 20 and 10 taxa, respectively. The following years saw marginal technical improvements and mostly ecological applications of these now mature approaches.

The next step in the classification of plankton images was the use of CNNs, publicized by the Kaggle competition of 2015. Many CNNs for plankton images were described in conference papers, which are often difficult to access. The first journal article was by Luo et al. (2018), who used a CNN trained on 42,000 images from the ISIIS that were sorted into 108 taxa. The model was evaluated on 75,000 images, for which the 108 taxa were regrouped into 38 broader ones. The overall accuracy was 87%, which looked excellent, but the precision was only 55% for biological categories, which led the authors to postprocess the results for ecological applications. As explained above, a CNN is essentially a feature extractor combined with a classifier. González et al. (2019) and Orenstein & Beijbom (2017) exploited feature extractors trained on nonplankton

Precision: for a taxon *T*, the proportion of images predicted to be *T* that actually are *T*

images combined with classifiers retrained on plankton data sets and showed that such external deep features can be very effective for classifying plankton over multiple data sets of several thousand images. This really is the main progress that CNNs bring: One can forgo the considerable domain expertise required to craft appropriate features, use a pretrained feature extractor, and get results that are equally good, if not better.

The discrepancies in the performance metrics reported, the nature and size of the data sets, and the number of taxa classified among the 174 papers within this systematic review make it impossible to assess progress over the last 40 years in a simple way. Actually, classification performance seems stable, hovering around 70–80% in most papers. Still, some conclusions can be reached. First, a richer set of features produces better classifiers, and CNNs are an evolution along that trend because they produce numerous (and optimized) features. Second, for a given classifier, performance degrades when the number of classified taxa increases. This explains the apparent stability in reported accuracy: Authors, justifiably, tackled tasks for which they could produce useful classifiers (i.e., having “good enough” accuracy). In the early days, this meant classifying 5 taxa, then ~20 in the 2000s, and now ~100. Therefore, the classifiers did improve, but this improvement cannot be demonstrated by a simple plot of accuracy over time. Finally, classifier performance saturates after a sufficient number of training examples per category is reached. The size of training (and testing) data sets increased over the years, probably together with the increasing number of taxa being classified.

Still, most data sets have remained rather small compared with those commonly collected for pelagic ecology studies, raising questions about their representativity. In addition, those data sets were rarely made public for other studies to reuse, making comparisons between studies impossible. This is changing, however, with the release of the WHOI-Plankton data set (Sosik et al. 2021), PlanktonSet (Cowen et al. 2015), and ZooScanNet (Elineau et al. 2018). In addition, all plankton data sets are severely unbalanced: One category dominates and is often made up of detritus or unidentifiable objects. The way this dominant category is handled can completely change the accuracy (the metric most often reported because it is easy to understand) and is a source of discrepancies between studies. For example, the accuracy of a classifier trained and evaluated on data sets with equal proportions of all categories is representative of its raw performance, and hence is a perfectly valid metric for machine learning, but may differ considerably from its accuracy on a real data set, depending on the performance on the dominant class. Other metrics, such as per-taxon statistics and their average across taxa of ecological interest, should be less sensitive and more comparable across studies.

Overall, there is a need for (a) large, public data sets of identified images, containing realistic proportions of the different taxa, and (b) systematic comparisons of different classifiers on these reference data sets.

2.2. Fully Automated Versus Manual Classification

To illustrate the difference between classic approaches and CNNs, as well as assess the difference a few points in accuracy make for the ecological interpretation of results, we ran a numerical experiment. A plankton data set collected along a time series was classified manually and then reclassified using both a classic classifier (RF on handcrafted features) and a modern CNN (for details, see **Supplemental Appendix 2**). The data set is quite typical: 850,000 images in 60 categories, significantly dominated by detritus (350,000 images, i.e., 41% of the data set). It has been checked by several independent experts and has already been used scientifically, highlighting relevant patterns (Feuilloley et al. 2021). Therefore, this data set allowed us to investigate whether patterns found in the human-validated series could be detected the same way in series predicted purely through machine learning.

For the classic RF classifier, the results on the test set were 67% accuracy, 71% mean precision across taxa, and 50% mean recall. Many images from biological taxa were predicted to be in the detritus class, resulting in the bad recall. With the CNN, the results were 76% accuracy, 69% mean precision, and 78% mean recall. The largest improvement was in the classification of biological taxa: Fewer images from biological taxa were predicted to be in the detritus category, hence increasing the recall. Looking at a few representative taxa (**Figure 6**) illustrates that automatic classification worked well for organisms that are abundant and/or have a distinctive shape (Calanidae and Chaetognatha), but both classifiers, particularly the RF classifier, had more trouble with rare taxa (Echinodermata) and smaller subgroups of similar-looking organisms (Temoridae, which were overestimated, among the copepods). Performance was similar, and very good, for Chaetognatha and Phaeodaria, yet this translated into quite different representativity. For Chaetognatha, the general trends were similar across the three methods, despite some errors. On the other hand, Phaeodaria were well predicted at the end of the time series, when they were most abundant, but were overestimated at the beginning, when they were rarer or absent. Feuilloley et al. (2021) observed that the concentrations of most copepod taxa were lower after 2015, while other taxa, such as tunicates (not shown here but very well predicted), increased in concentration. Only the RF classifier would bias this observation, because of the overestimation of some copepods, and only slightly. For Phaeodaria, the intensity of the stepwise increase in 1999 and then 2008 is dampened in the two predicted series because of the overestimation at low concentrations. Finally, the clear increases in the concentration of larvae of echinoderms around 2002–2003 and after 2010 would be completely missed by the RF prediction.

The take-home message from this experiment, which echoes our experience, is that classic approaches were already very good at predicting plankton when they are abundant. From this point of view, CNNs yield little improvement beyond being easier to use and more portable, thanks to libraries such as Keras/TensorFlow and PyTorch. In addition, CNNs do better at low concentrations (i.e., rare taxa but also regions or periods of low concentrations of abundant taxa), when even a few errors have disproportionately large consequences. However, anticipating where and when errors are sufficiently small is difficult, even based on per-category statistics (e.g., as with Phaeodaria above). The problematic portions are usually regions or periods of data set shift: Because of the small sample size, the relative concentrations of taxa change compared with those in the training set, confusing the classifier. In addition to better classifiers, quantification approaches, where the goal is not to predict the category of every single image but to predict overall concentrations per category, can help alleviate this problem (Orenstein et al. 2020a).

Systematic comparisons on reference data sets are still necessary, as explained above. For the time being, conclusions drawn from fully automated predictions regarding high-concentration events can likely be trusted regardless of which classification approach is chosen, as long as the training set is large and representative enough. If finer-scale patterns and small differences in concentration matter, CNNs help but cannot be considered a silver bullet; manual validation, ideally by several different operators, is still required.

3. NEW INSIGHTS ON PLANKTON ECOLOGY FROM AUTOMATED PELAGIC IMAGING

3.1. Resolving Plankton Temporal Patterns

Imaging systems can be deployed in situ and acquire data at a high rate. This creates opportunities for monitoring, as long as the images can be processed in near real time with sufficient accuracy. For example, an early warning system for blooms of the red-tide-forming dinoflagellate *Karenia brevis* was set up by combining trihourly samples taken by the IFCB with several one-versus-all

Recall: for taxon T , the proportion of images that are actually T and are predicted to be T

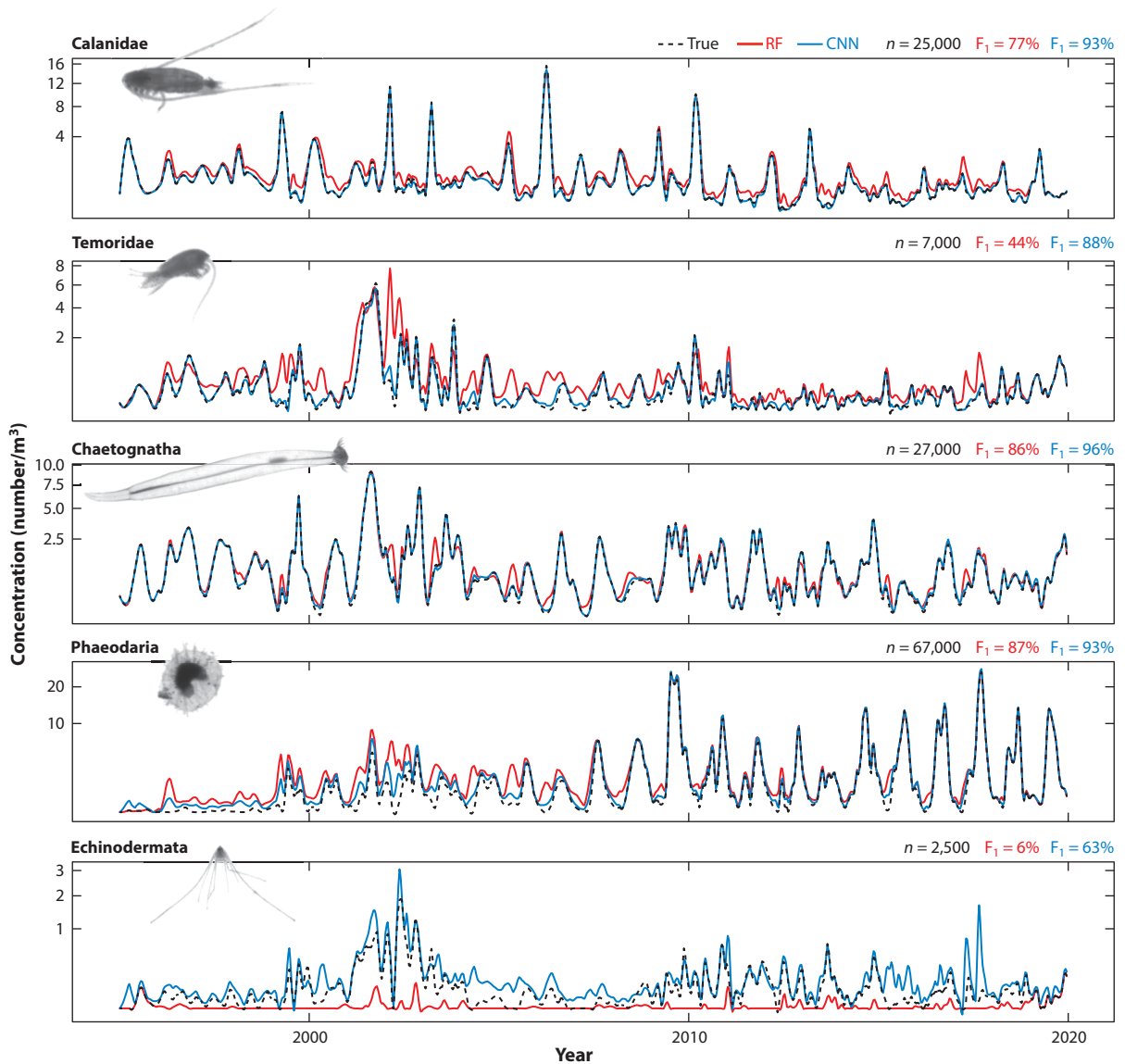


Figure 6

Comparison of time series of the concentrations of planktonic organisms (Point B, Villefranche-sur-Mer, northwestern Mediterranean Sea; IMEV et al. 2020). The true time series (*dashed*) is compared with the series predicted by a classifier based on classic handcrafted features and RF (*red*) and by a CNN (*blue*). The *y* axis is square-root transformed to better show fluctuations at low concentrations. The weekly series are smoothed by a local polynomial with a window size of approximately one season to highlight ecologically relevant patterns. For each taxon, the number of images over the whole series and the F₁ score for each classifier (the harmonic mean of recall and precision, i.e., a number that balances and summarizes both metrics) are given in the top right. Abbreviations: CNN, convolutional neural network; RF, random forests.

support vector machine classifiers, one being focused on *K. brevis*. The classifier was coupled with a notification system that sent an email message once cell counts exceeded a defined threshold (Campbell et al. 2013). This system was recently extended to use CNNs and include two other target species (Henrichs et al. 2021).

Over seasonal scales, changes in plankton assemblages are described as ecological successions (Sommer 2012). Most studies focus on the succession within a specific taxonomic group because sampling and observation instruments target a given size range and because scientific expertise for the identification of species is also split along those lines. By contrast, Romagnan et al. (2015) used a combination of flow cytometry, the FlowCam, and the ZooScan to study the succession of the whole plankton community. Image analysis of weekly samples over 10 months was made possible by a combination of machine learning prediction and human validation. It also fostered the bringing together of disparate expertise to propose an original mechanism for plankton succession: a stepwise stacking of four broad trophic communities, from microbes to predators, complementing the more traditional view of a progressive replacement of one community by the other throughout the season.

Studies of the multidecadal evolution of plankton communities are rare not only because sampling at that scale is difficult, but also because the subsequent identification effort is tremendous. For example, along a time series in the northwestern Mediterranean Sea, the traditional microscope counts of zooplankton performed on some target groups from 1966 to 1993 (Molinero et al. 2008) were supplemented by digitalization and classification of whole samples. By combining machine learning classifiers and human curation over 12 more years of data, researchers determined that the changes initially described as a trend were actually periodic oscillations in the whole zooplankton community (Garcia-Comas et al. 2011, Vandromme et al. 2011). This data acquisition and identification procedure is still in effect today. Furthermore, the archive of images, together with improved machine learning tools, make it possible to regularly review images to deepen the level of taxonomic classification (~10 groups in 2011, >80 now). This suggests that imaging and machine learning could enable the rescue of data from other samples, preserved in formaldehyde for decades at many institutions, to build historical plankton data sets at a global scale.

3.2. Resolving Plankton Spatial Patterns

As with time series, quantitative imaging is particularly adept at resolving small scales in space. The earliest plankton imaging instrument developed (the VPR) was targeted at this problem, showing consistent patchiness at scales of meters to tens of meters horizontally (Davis et al. 1992b). Along the vertical, changes are even more abrupt and require extremely high frequency acquisition. Greer et al. (2013) used the ISIIS to characterize the meter-scale distribution of zooplankton in relation to a thin phytoplankton layer. In the deep ocean, the fine-scale patchiness of filter-feeding larvae around a hydrothermal vent could be related to particle distribution through the analysis of video sequences (Lindsay et al. 2015). None of those studies relied on machine classification, however, possibly because of a lack of trust in the ability of classifiers to reveal such fine patterns. But, as a consequence, all studies have a limited spatiotemporal scope.

The mesoscale is dominated by fronts and eddies, which drive three-dimensional transport and nutrient supply into the euphotic zone, thereby influencing plankton. This scale can be explored through traditional net samples, but imaging can improve the resolution of spatial patterns. For example, a clear association between the filamentous cyanobacteria *Trichodesmium*, identified in VPR images, and anticyclonic eddies was observed across a trans-Atlantic survey (Davis & McGillicuddy 2006). Phytoplankton communities during the spring bloom in the North Atlantic were described through 380,000 images from the IFCB that had been reviewed by humans after

an initial machine learning prediction (Bolaños et al. 2020). They enabled the detection of a clear association between communities and water masses at the mesoscale; the additional size data they provided showed that small cells, rather than diatoms, dominate biomass during the bloom.

Resolving global patterns requires the compilation of thousands of data points—a huge task for which quantitative imaging and machine learning are uniquely suited, because the consistency of identifications in the compiled data set can be checked. Thus, the relationship between the midwater zooplankton community and Longhurst’s biogeochemical provinces could be described from a few hundred stations (Stemmann et al. 2008), and later on, the latitudinal distribution of rhizarians could be illuminated from almost 1,000 points (Biard et al. 2016).

3.3. Estimating Carbon Flux in the Twilight Zone

Because particles of marine snow are substantially more numerous than plankton, in situ instruments mostly image these particles. The images can therefore be used to count and characterize these carbon-rich particles at different depths in the ocean, infer a flux, and estimate its attenuation with depth. The ability to observe the mesopelagic zone, a key layer for the transfer of carbon from the surface to the deep sea, has revealed seasonal particle flux patterns (Lampitt et al. 1993, Stemmann et al. 2002) and hot spots of particle flux in the open ocean (Guidi et al. 2007, Kiko et al. 2017). Because marine snow particles are so numerous and present an apparently continuous spectrum of sizes and shapes (**Figure 1e**), most studies have not used machine learning to classify them and have focused only on measuring their size. Key variables for carbon flux, such as sinking speed, have then been inferred through allometric relationships with size (Giering et al. 2020a, Stemmann & Boss 2012). However, points are often spread widely around those relationships because other particle properties influence sinking speed, such as ballasting, porosity, or the roughness of their exteriors (Cael et al. 2021). Such properties could be inferred from images, and particles could therefore be objectively categorized using machine learning techniques to better assess their contribution to the carbon flux through per-category relationships.

3.4. Seeing Plankton In Situ and (Re-)Discovering Important Taxa

As mentioned in Section 1, plankton nets damage the organisms they sample. The modern nylon nets damage soft-bodied zooplankton even more than the silk nets used by the natural historians of the nineteenth century (40% of all the ctenophore species described in the twentieth century were described in its first decade, from silk net samples; Haddock 2004). The advent of crewed submersibles and remotely operated vehicles has led to a mini-resurgence in descriptions (Robison et al. 2017). With Declaration 45 of the International Commission on Zoological Nomenclature stating that descriptions of new species are permitted based on photographic evidence alone when “justified by special circumstances” (ICZN 2017, p. 96), descriptions relying heavily on images (Lindsay 2017) or even based on images alone (Ford et al. 2020) can be expected to become more commonplace.

Furthermore, some taxa thought to be rare based on net samples have actually been found to be quite common through the use of in situ cameras. This was the case for gelatinous macrozooplankton [Lindsay et al. (2008) reported more than 60 occurrences of the anthomedusa *Pandea rubra* during a single remotely operated vehicle cruise, whereas fewer than 6 had been recorded over the previous 100 years using nets], *Trichodesmium* colonies (Davis & McGillicuddy 2006, Sandel et al. 2015), subgroups of Rhizaria (Biard et al. 2016, Brisbin et al. 2020, Nakamura et al. 2017), and polychaetes (Christiansen et al. 2018). Better estimates of the biomass of these taxa, through in situ imaging combined with machine learning, are relevant because they are significant

players in primary production (*Trichodesmium* and Collodaria), nitrogen fixation (*Trichodesmium*), and vertical carbon flux (Phaeodaria, polychaetes, and gelatinous macroplankton).

3.5. Using Images to Infer Functional Traits

Beyond taxonomic identification, images of plankton and marine snow can provide rich morphological information at the individual level. Such data enable trait-based analyses to investigate the functioning of planktonic communities or biogeochemical fluxes of marine snow, where a trait is a characteristic of an individual organism or object that affects its fitness or role in the ecosystem. Size has been termed the master trait and is readily measured individually to infer the dynamics of phytoplankton (Bolaños et al. 2020), zooplankton (Romagnan et al. 2016), and marine snow (Guidi et al. 2009). Quantitative imaging has recently been used to assess the definitions of phytoplankton size classes, which are classically based on diagnostic pigment analysis. Individual measurements from images showed that the pigment approach tends to overestimate the contribution of microphytoplankton and underestimate the contribution of nanophytoplankton to the total phytoplankton biomass (Chase et al. 2020). But images—those taken in situ in particular—capture many more traits of plankton that could conceivably be measured automatically (Ohman 2019), to investigate feeding (e.g., tentacles of cnidarians; **Figure 1f**), defense (e.g., spines of phytoplankton cells; **Figure 1a**), and reproduction (e.g., egg volume; **Figure 1b**) and to characterize the density or contour shape of marine snow particles (**Figure 1e**). In fact, half of the functional traits proposed for aquatic organisms by Martini et al. (2020) have a visual signature. Few studies explicitly use this optical information, but they show great promise.

From 1.2 million images of organisms and marine snow captured in situ, in the Arctic Ocean, Vilgrain et al. (2021) identified approximately 28,000 copepods using a combination of machine learning and human validation. Visual features were measured on every object, and the authors used dimensionality reduction through principal component analysis to define three main continuous functional traits related to size, opacity, and appendage visibility (**Figure 7a**). This individual-level information revealed a complex response to sea ice melting in the spring: Larger individuals with less visible appendages (i.e., in a resting posture) were under the ice; individuals at the ice edge were darker, likely because of their fuller guts and richer pigment content in this blooming and well-lit region; and smaller, probably younger, individuals with extended antennae (suggesting that they were active) were in the ice-free, postbloom, open waters.

Trudnowska et al. (2021) used a similar approach on 2 million marine snow images from the same region, with an additional step of unsupervised clustering to objectively distinguish different morphotypes of particles (**Figure 7b**). The forms of these particles can illuminate their nature (fecal pellets, dead cells, molts, aggregates, etc.) and therefore their significance for the vertical flux. The authors traced the dynamics of two ice-edge phytoplankton blooms through the relative abundances of the marine snow morphotypes they produced and estimated a faster carbon export by smaller, darker, probably denser particles during the bloom, followed by a significant but slower export by larger and fluffier aggregates after the bloom.

Computer vision, sometimes using machine learning, could be used to investigate many more functional traits. For example, Kenitz et al. (2020) used a binary support vector machine classifier to distinguish single-cell from chain-forming diatoms and generate a daily-resolved time series of their abundance over 2.5 years. They showed that chain-forming diatoms dominated during bloom conditions, whereas solitary diatoms prevailed during the rest of the year, suggesting the existence of a trade-off between resource acquisition and predator defense in diatoms. The computer sciences community provides readily available solutions for more complex problems. Object detectors or semantic segmentation algorithms could detect relevant portions of an organism's

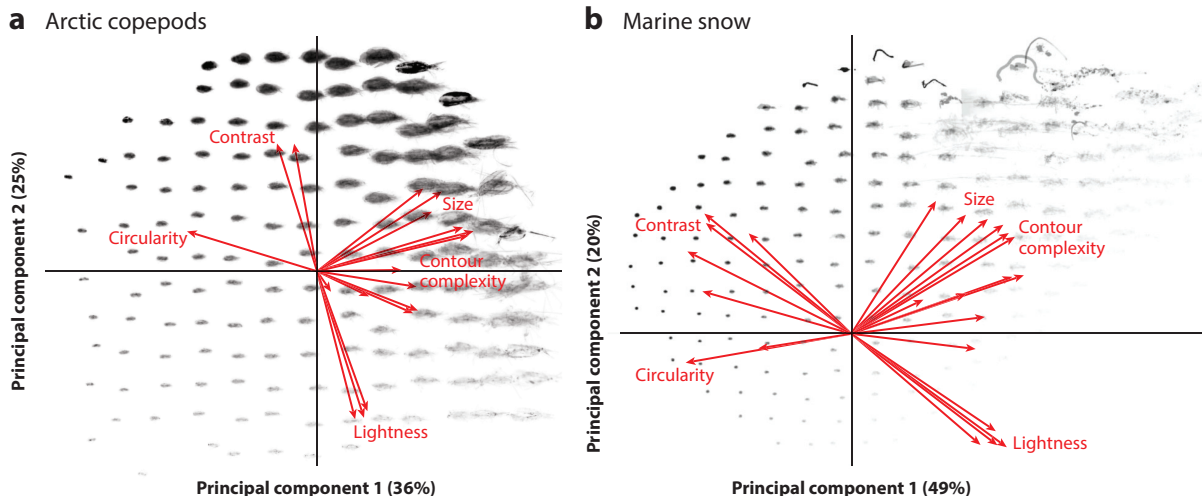


Figure 7

Morphological trait-space representation of (a) Arctic copepods and (b) marine snow. Morphological descriptors computed from UVP images were summarized using principal component analysis, as represented here. The arrows are the original descriptors (whose general meaning is given in the figure), the axes are the principal components, and example images are displayed at their projected locations in the trait space. For both copepods and particles, the first principal component is related to size (small on the left, large on the right) and the second to opacity (dark, and hence opaque, at the top; light, and hence transparent, at the bottom). Abbreviation: UVP, Underwater Vision Profiler. Panel a adapted from Vilgrain et al. (2021) under a CC BY-NC 4.0 license (<https://creativecommons.org/licenses/by-nc/4.0>); panel b adapted from Trudnowska et al. (2021) under a CC BY 4.0 license (<https://creativecommons.org/licenses/by/4.0>).

body, such as lipid sacs in copepods (**Figure 1b**) or gonads visible through the transparent body of jellies. Object detectors combined with tracking would allow investigation of the swimming behavior from in situ videos. Deep regression could estimate a quantity (such as the amount of lipid reserves or number of eggs; **Figure 1b**) directly from an image. Color space manipulation could ease the quantification of the intensity of a given color, to automatically resolve fine-scale changes such as the diel variation in color and transparency of *Sapphirina nigromaculata* documented by Takahashi et al. (2015), which was interpreted as a strategy to find a mate. Finally, in situ live staining combined with object detectors could provide information on trophic functions (Brownlee et al. 2016) or physiological states or traits. All of these approaches would allow the extraction of even more ecologically relevant information from the abundance of pelagic images collected.

4. TOWARD A GLOBAL NETWORK OF IMAGES

4.1. Improved Plankton Image Classification

Several of the studies described above, including recent ones (Greer et al. 2015), used manual classification only, even though supervised classifiers were theoretically available to perform at least part of the work. A likely explanation is that image processing and the coding of pipelines for data management are not what plankton ecologists are trained to do. This highlights the need for tools that give ecologists and taxonomists easy access to machine classifiers, even if those are a few points of accuracy below the state of the art, as long as the user interface abstracts the technicalities and can handle thousands of images quickly.

Of course, better classifiers are always welcome, and the future clearly lies in a more widespread use of CNNs. But other, existing advances in computer sciences could be leveraged more systematically for plankton imaging. For example, simultaneous classification and tracking of organisms in videos is a problem similar to moving-object tracking in self-driving cars (Badue et al. 2021). Learning from a data set containing incorrect identifications is a common problem, known as label noise, for which solutions exist (Fefilyat'ev et al. 2012). Active learning procedures can guide taxonomists toward the most relevant images to improve a training set or the most uncertain ones to assess the performance of a classifier (Luo et al. 2005). Plankton image classification also has some properties that should make it an interesting research question for computer scientists: the intrinsic imbalances of taxa abundances, the small and often grayscale images that lend themselves to the development of smaller networks for embedded devices, the existence of a phylogeny of taxa derived from information external to the image, and a relative wealth of human-classified data. Achieving this transfer of knowledge across disciplines will require collaboration between researchers from both fields and training of students at their interface.

Still, machine learning models are only as good as the data they are trained on. A sure way to improve classifiers is through large, diverse, and high-quality training data sets. However, the number of trained taxonomists is ever decreasing (Pearson et al. 2011), and the number of those who can reliably identify lower taxa based on images alone is even smaller. In addition, manually labeling a large number of images is cumbersome and not gratifying for those trained taxonomists. As a result, images are classified by nonspecialists, often interns, who, unsurprisingly, make mistakes. For example, in the most advanced research paper to date on the use of CNN classifiers applied to plankton images (Ellen et al. 2019), some example images presented as representative are not correctly labeled: The acantharians include spumellarian radiolarians, the fish larvae include copepods, the Narcomedusae include what appear to be Anthomedusae, and so on. Because the sorting of each category was consistent, the names applied to them do not bear on the results of the machine learning endeavor. But the fact that experts could not be consulted to verify the labels and that the reviewers for a leading journal did not catch those mistakes is telling. This deficiency of taxonomic expertise is a challenge among almost all research groups making use of imaging technology. Furthermore, this problem is, of course, not restricted to images of plankton and has recently been shown to hinder the assessment of the performance of all image classifiers (Northcutt et al. 2021). Even GenBank, the go-to database for genetic information, contains many inaccurate taxonomic identifications (Lindsay et al. 2017), so combinations of imaging and sequencing proposed to alleviate the taxonomic bottleneck (Riedel et al. 2013) are also problematic. Therefore, collaboration with trained taxonomists to build and share gold-standard plankton image data sets appears necessary to fully harness the power of modern machine learning techniques.

4.2. Coordinated Acquisition of Massive Data Sets

The increasing rate of acquisition of plankton images (**Figure 3**) has been fueled by the recognition that they can provide biological information at an unprecedented scale, for research and monitoring purposes, and by the wide distribution of commercial instruments (Lombard et al. 2019, Stemmann et al. 2012). This acquisition is quickly growing beyond the capacities of most plankton ecology laboratories, even for classic stand-alone studies: When each deployment yields millions of images to classify and terabytes of raw data to store, personal computers are not enough; data centers become necessary.

Section 3 highlighted that few data sets cover long periods or large spatial scales. With the deployment of imaging sensors on automated moorings (Orenstein et al. 2020b), on autonomous

instruments (Ohman et al. 2019), or as part of large-scale programs such as the Global Ocean Ship-Based Hydrographic Investigations Program (GO-SHIP) (Lombard et al. 2019), such data sets should become easier to acquire, and this would be a boon for understanding plankton dynamics at those scales. However, those data sets would be even larger than those of the typical ecological study and will require adequate infrastructure.

Another way of accessing information over long/large scales, or to investigate local processes but at several times/places, is by aggregating individual data sets. For this to succeed, data sets need to be interoperable, consistent, and shared.

For plankton and marine snow image data sets, interoperability probably starts with a common taxonomy. The World Register of Marine Species (Costello et al. 2013) is the reference for the ocean. While its coverage of protists as well as its basal taxa could be improved, and it does not deal with nonorganismal objects (i.e., marine snow), it is an ongoing work, used as a taxonomic backbone of the Ocean Biodiversity Information System, the Global Biodiversity Information Facility, the Encyclopedia of Life, and so on. But taxonomic identification from images is often not as definitive as identification from physical samples. Therefore, taxonomists may be unsure and/or stop at a coarse level; whether this is due to a lack of meristic information or done by choice, to limit the human effort required, is unclear. Open nomenclature qualifiers were recently recommended to designate cases of incomplete or uncertain identification in the specific case of images (Horton et al. 2021). Beyond taxa names, other qualifiers are often added to identifications (such as life stage), and some are specific to in situ images (such as whether the organism is complete—e.g., an appendicularian body, or just its mucus house, or both). The vocabularies of the British Oceanographic Data Centre (BODC 2021), used by the Ocean Biodiversity Information System, and of the Video Annotation and Reference System knowledge base (MBARI 2019), geared toward deep-sea images, provide standard terms for some but not all of these cases. Those vocabularies also cover the documentation of other essential metadata, including not only the longitude, latitude, and time but also the volume sampled, instrument name, pixel size of the image, and so on, which are essential to aggregate data sets in a meaningful way. They are currently being extended to better accommodate quantitative imaging data sets.

Consistency also starts with taxonomy. As explained above, the nature of image-based identification (at a coarse level, from a single view of the organism that may not display all diagnostic characters) makes it more prone to subjectivity. The difficulty of differentiating smaller objects from the largely dominant and diversely shaped marine snow makes errors common and possibly numerous. But images can at least be reviewed by several human operators, and machine learning can also help identify visual outliers in a data set, which may be incorrect classifications (Fefilat'ev et al. 2012). Still, proper quality control procedures remain to be defined, like they were for other global programs (e.g., Argo floats; Gaillard et al. 2009). In addition, for training sets to be transferable and with the advent of trait-based approaches, the images themselves and the features measured on them would also need to be kept as consistent as possible. This involves intercalibration of instruments, which is rarely done (but see Picheral et al. 2010), and of the software that processes the images (Giering et al. 2020b). Currently, even size measurements must be postprocessed to align between instruments (Forest et al. 2012, Stemann et al. 2012).

Finally, individual data sets need to be shared in order to be aggregated. A few data aggregation initiatives exist, such as the Marine Ecosystem Biomass Data (MAREDAT) initiative (Buitenhuis et al. 2013) and the aforementioned Ocean Biodiversity Information System, but they are often hindered by the lack of high-quality, documented, open data sets. Even with legal and cross-reference tools that should ensure proper traceability and citability (Creative Commons licenses, digital object identifiers), the cost and effort to collect ecological data sets in the ocean are invoked to justify the disinclination toward sharing them. But although DNA sequencing used to

be very expensive (and is still not cheap), public release of sequences used in articles has been mandatory for some time (Kaye et al. 2009). The ImageNet data set (Deng et al. 2009) represents a tremendous curation effort of 14 million images from the internet, in ~22,000 categories, yet is freely available for anyone to download. In both cases, the public release of data has considerably accelerated scientific research, because teams can build on one another's work. Biological oceanographers need to embrace data sharing to reap the same benefits, and images are rather easy to share. Yet the infrastructure for sharing images is in its infancy and requires commitment for continuous financial support.

4.3. Customized Tools for Quantitative Imaging

Most hurdles between us and a worldwide network of plankton image curation seem to be in process and mentalities, not in technology. Indeed, many instruments and machine learning techniques are mature and already in use. It is the connections between them, the overall data flow, and the final release of data that are often missing.

Software tools that ease the storage, management, annotation, and sharing of images can help. Gomes-Pereira et al. (2016) reviewed 23 such tools, most of which are geared toward benthic images, although SQUIDLE+ (<https://squidle.org>), for example, is also used for plankton. Tools dedicated to plankton images have also been developed by independent teams; for example, Plankton Identifier (Gorsky et al. 2010) and Zoo/PhytoImage (Grosjean et al. 2018) can be used to build training sets and ease access to classic machine learning methods, and the IFCB dashboard (<https://ifcb-data.whoi.edu>) enables users to store and browse data sets from this instrument. Recently, the EcoTaxa web application (Picheral et al. 2017) has gained traction and now hosts ~150 million images, from ~10 different instruments, generated by over 1,000 users from more than 350 organizations worldwide (Figure 3). It integrates image and metadata storage, data browsing and subsetting, classic and deep learning, and data export. Its main goal is to provide an interface that accelerates the work of human curators by using machine learning. Yet, in its current form, it will not scale to the billions of plankton images collected worldwide, it only conforms to a few metadata standards, it has no tools dedicated to quality control, and its machine learning back end could be improved. These shortcomings are being actively worked on, but at the moment, neither EcoTaxa nor any other tool is fulfilling the needs outlined in the paragraphs above. More work should be dedicated to the development of such tools for the benefit of the whole community.

Furthermore, all of those tools are designed for experts, packing many functions into a dense user interface. If plankton observations are to be scaled to the global ocean, these specialized tools could be complemented by more approachable ones, geared toward a global community of citizen scientists. Successful examples already exist for plankton (Kiko et al. 2018, Robinson et al. 2017), and inspiring solutions have been developed by other ecologists, such as iNaturalist (<https://www.inaturalist.org>) and Pl@ntNet (<https://plantnet.org>). In these projects, scientists provided the initial input of data, which served to train the first machine learning classifiers. This then allowed a community of hobbyists to grow, together with improved classifiers, over several years. Now, the records generated by these applications feed back into the Global Biodiversity Information Facility, are presented at major computer sciences conferences (Van Horn et al. 2018), and are used by ecologists to describe the distribution of species (Botella et al. 2018). The community of planktonology hobbyists is likely much smaller than the community of botany hobbyists, so the number of classifications generated may not be impressive at first. But citizen science applied to images is an excellent outreach tool to grow such a community. Plankton are just as pretty as flowers!

SUMMARY POINTS

1. Pelagic imaging instruments and associated machine learning models are now mature and have brought significant insights on ocean ecology and biogeochemistry.
2. However, the progress of machine classifiers is difficult to track because much of the current literature relies on different small data sets, with different numbers of image categories, and often highlights only overall accuracy, which is potentially misleading for unbalanced data sets.
3. Modern convolutional neural networks are easier to set up than classic machine learning approaches and perform better at low abundances.
4. Fully automated taxonomic prediction can be trusted for high-abundance events, but manual validation by experts remains necessary for rare taxa and the locations or periods where common taxa are present in small numbers, even with convolutional neural networks.
5. Despite the apparent success of automation, many studies still rely on manual image classification because of a lack of tools that are approachable for all plankton ecologists.
6. While quantitative imaging and machine learning facilitate the acquisition of large, consistent data sets, studies over long periods and at a global scale are still rare.
7. Beyond taxonomic classification, pelagic images provide rich visual information about plankton and marine snow that enable promising trait-based approaches.

FUTURE ISSUES

1. Plankton and marine snow image data have rare properties that create interesting machine learning problems and should be leveraged to foster deeper collaborations between computer scientists and pelagic ecologists.
2. Systematic exploration of existing and future machine learning methods should be performed on large, public data sets that reflect realistic planktonic ecosystems and should be assessed with metrics that account for the imbalance between categories.
3. Such gold-standard data sets should be curated by expert taxonomists, through collaborative tools, to avoid the errors that plague current training sets.
4. Now that image acquisition and classification are mature, effort should be directed toward the development of significant infrastructure to store and curate images, with the help of machine learning, at a scale that is beyond what plankton ecology laboratories are used to.
5. To leverage such an infrastructure and the data therein, oceanographers should be trained in data science, not only to leverage machine learning methods but also simply to deal with the massive data sets involved.
6. The software tools developed should rely on standard terms and taxonomy, to facilitate data exchange, and should allow the development of quality control procedures, to ensure data consistency.

7. These reference data sets, infrastructures, and tools can only be achieved through global collaboration and a change of mindset toward systematic data sharing.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

The authors wish to thank their hiring institutions for financial support through their salaries. Their work and reflections on the issues discussed in this article were also fostered by the following projects: the World Wide Web of Plankton Image Curation, Belmont Forum (J.-O.I. grant ANR-18-BELM-003, D.J.L. grant 18076935); the Working Group on Machine Learning in Marine Sciences, International Council for the Exploration of the Sea (J.-O.I.); Blue Cloud, European Union (J.-O.I.); JERICO-S3, European Union (J.-O.I., L.S.); FORMAL (From Observing to Modeling Ocean Life), Institut des Sciences du Calcul et des Données, Sorbonne Université (J.-O.I., S.-D.A., L.S.); ARTIFACTZ, Université Côte d'Azur (S.-D.A., J.-O.I.); KAKENHI, Japan Society for the Promotion of Science (D.J.L. grant 20K06806); the Vision in the Ocean chair, CNRS, Sorbonne Université (L.S.); BRIDGES, European Union (L.S., J.-O.I.); and GOPPI, BPI/Pôle Mer Méditerranée (L.S.). We also thank C. Desnos, L. Jalabert, and A. Elineau for their work on the data in **Figure 6**.

LITERATURE CITED

- Badue C, Guidolini R, Carneiro RV, Azevedo P, Cardoso VB, et al. 2021. Self-driving cars: a survey. *Expert Syst. Appl.* 165:113816
- Behrenfeld MJ, Randerson JT, McClain CR, Feldman GC, Los SO, et al. 2001. Biospheric primary production during an ENSO transition. *Science* 291:2594–97
- Bell JL, Hopcroft RR. 2008. Assessment of ZooImage as a tool for the classification of zooplankton. *J. Plankton Res.* 30:1351–67
- Benfield MC, Grosjean P, Culverhouse PF, Irigoien X, Sieracki ME, et al. 2007. RAPID: Research on Automated Plankton Identification. *Oceanography* 20(2):172–87
- Benoit-Bird KJ, McManus MA. 2012. Bottom-up regulation of a pelagic community through spatial aggregations. *Biol. Lett.* 8:813–16
- Biard T, Stemmann L, Picheral M, Mayot N, Vandromme P, et al. 2016. In situ imaging reveals the biomass of giant protists in the global ocean. *Nature* 532:504–7
- Blackburn N, Hagström Å, Wikner J, Cuadros-Hansson R, Bjørnson PK. 1998. Rapid determination of bacterial abundance, biovolume, morphology, and growth by neural network-based image analysis. *Appl. Environ. Microbiol.* 64:3246–55
- BODC (Br. Oceanogr. Data Cent.). 2021. Vocabularies. *British Oceanographic Data Centre*. <https://www.bodc.ac.uk/resources/vocabularies>
- Bolaños LM, Karp-Boss L, Choi CJ, Worden AZ, Graff JR, et al. 2020. Small phytoplankton dominate western North Atlantic biomass. *ISME J.* 14:1663–74
- Botella C, Joly A, Bonnet P, Monestiez P, Munoz F. 2018. Species distribution modeling based on the automated identification of citizen observations. *Appl. Plant Sci.* 6:e1029
- Boyd PW, Claustre H, Levy M, Siegel DA, Weber T. 2019. Multi-faceted particle pumps drive carbon sequestration in the ocean. *Nature* 568:327–35
- Breiman L. 2001. Random forests. *Mach. Learn.* 45:5–32

- Brisbin MM, Brunner OD, Grossmann MM, Mitarai S. 2020. Paired high-throughput, in situ imaging and high-throughput sequencing illuminate acantharian abundance and vertical distribution. *Limnol. Oceanogr.* 65:2953–65
- Brownlee EF, Olson RJ, Sosik HM. 2016. Microzooplankton community structure investigated with imaging flow cytometry and automated live-cell staining. *Mar. Ecol. Prog. Ser.* 550:65–81
- Buitenhuis ET, Vogt M, Moriarty R, Bednaršek N, Doney SC, et al. 2013. MAREDAT: towards a world atlas of MARine Ecosystem DATA. *Earth Syst. Sci. Data* 5:227–39
- Cael BB, Cavan EL, Britten GL. 2021. Reconciling the size-dependence of marine particle sinking speed. *Geophys. Res. Lett.* 48:e2020GL091771
- Campbell L, Henrichs DW, Olson RJ, Sosik HM. 2013. Continuous automated imaging-in-flow cytometry for detection and early warning of *Karenia brevis* blooms in the Gulf of Mexico. *Environ. Sci. Pollut. Res.* 20:6896–902
- Chase AP, Kramer SJ, Haëntjens N, Boss ES, Karp-Boss L, et al. 2020. Evaluation of diagnostic pigments to estimate phytoplankton size classes. *Limnol. Oceanogr. Methods* 18:570–84
- Christiansen S, Hoving H-J, Schütte F, Hauss H, Karstensen J, et al. 2018. Particulate matter flux interception in oceanic mesoscale eddies by the polychaete *Poebius* sp. *Limnol. Oceanogr.* 63:2093–2109
- Cortes C, Vapnik V. 1995. Support vector machine. *Mach. Learn.* 20:273–97
- Costello MJ, Bouchet P, Boxshall G, Fauchald K, Gordon D, et al. 2013. Global coordination and standardisation in marine biodiversity through the World Register of Marine Species (WoRMS) and related databases. *PLoS ONE.* 8:e51629
- Cowen RK, Guigand CM. 2008. *In situ* ichthyoplankton imaging system (ISIS): system design and preliminary results. *Limnol. Oceanogr. Methods* 6:126–32
- Cowen RK, Sponaugle S, Robinson KL, Luo JY. 2015. PlanktonSet 1.0: plankton imagery data collected from F.G. Walton Smith in Straits of Florida from 2014-06-03 to 2014-06-06 and used in the 2015 National Data Science Bowl (NCEI Accession 0127422). *National Centers for Environmental Information.* <https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.nodc:0127422>
- Culverhouse PF, Simpson RG, Ellis R, Lindley JA, Williams R, et al. 1996. Automatic classification of field-collected dinoflagellates by artificial neural network. *Mar. Ecol. Prog. Ser.* 139:281–87
- Culverhouse PF, Williams R, Benfield M, Flood PR, Sell AF, et al. 2006. Automatic image analysis of plankton: future perspectives. *Mar. Ecol. Prog. Ser.* 312:297–309
- Davis CS, Gallagher SM, Berman MS, Haury LR, Strickler JR. 1992a. The Video Plankton Recorder (VPR): design and initial results. *Arch. Hydrobiol. Beib.* 36:67–81
- Davis CS, Gallagher SM, Solow AR. 1992b. Microaggregations of oceanic plankton observed by towed video microscopy. *Science* 257:230–32
- Davis CS, Hu Q, Gallagher SM, Tang X, Ashjian CJ. 2004. Real-time observation of taxa-specific plankton distributions: an optical sampling method. *Mar. Ecol. Prog. Ser.* 284:77–96
- Davis CS, McGillicuddy DJ. 2006. Transatlantic abundance of the N₂-fixing colonial cyanobacterium *Trichodesmium*. *Science* 312:1517–20
- de Vargas C, Audic S, Henry N, Decelle J, Mahé F, et al. 2015. Eukaryotic plankton diversity in the sunlit ocean. *Science* 348:1261605
- Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L. 2009. ImageNet: a large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–55. Piscataway, NJ: IEEE
- Dennett MR, Caron DA, Michaels AF, Gallagher SM, Davis CS. 2002. Video plankton recorder reveals high abundances of colonial Radiolaria in surface waters of the central North Pacific. *J. Plankton Res.* 24:797–805
- Dubelaar GB, Gerritzen PL, Beeker AE, Jonker RR, Tangen K. 1999. Design and first results of CytoBuoy: a wireless flow cytometer for in situ analysis of marine and fresh waters. *Cytometry* 37:247–54
- Durden JM, Schoening T, Althaus F, Friedman A, Garcia R, et al. 2016. Perspectives in visual imaging for marine biology and ecology: from acquisition to understanding. *Oceanogr. Mar. Biol. Annu. Rev.* 54:1–72
- Elineau A, Desnos C, Jalabert L, Olivier M, Romagnan J-B, et al. 2018. ZooScanNet: plankton images captured with the ZooScan. *SEANOE.* <https://doi.org/10.17882/55741>
- Ellen JS, Graff CA, Ohman MD. 2019. Improving plankton image classification using context metadata. *Limnol. Oceanogr. Methods* 17:439–61

- Fefilat'yev S, Shreve M, Kramer K, Hall L, Goldof D, et al. 2012. Label-noise reduction with support vector machines. In *Proceedings of the 21st International Conference on Pattern Recognition*, pp. 3504–8. Piscataway, NJ: IEEE
- Feuilloley G, Fromentin J-M, Saraux C, Irsson J-O, Jalabert L, Stemmann L. 2021. Temporal fluctuations in zooplankton size, abundance and taxonomic composition since 1995 in the North Western Mediterranean Sea. *ICES J. Mar. Sci.* In press
- Ford M, Bezio N, Collins A. 2020. *Duobruchium sparksae* (*incertae sedis* Ctenophora Tentaculata Cydippida): a new genus and species of benthopelagic ctenophore seen at 3,910 m depth off the coast of Puerto Rico. *Plankton Benthos Res.* 15:296–305
- Forest A, Stemmann L, Picheral M, Burdorf L, Robert D, et al. 2012. Size distribution of particles and zooplankton across the shelf-basin system in southeast Beaufort Sea: combined results from an Underwater Vision Profiler and vertical net tows. *Biogeosciences* 9:1301–20
- Frederiksen M, Edwards M, Richardson AJ, Halliday NC, Wanless S. 2006. From plankton to top predators: bottom-up control of a marine food web across four trophic levels. *J. Anim. Ecol.* 75:1259–68
- Gaillard F, Autret E, Thierry V, Galaup P, Coatanoan C, Loubrieu T. 2009. Quality control of large Argo datasets. *J. Atmos. Ocean. Technol.* 26:337–51
- Garcia-Comas C, Stemmann L, Ibanez F, Berline L, Mazzocchi MG, et al. 2011. Zooplankton long-term changes in the NW Mediterranean Sea: decadal periodicity forced by winter hydrographic conditions related to large-scale atmospheric changes? *J. Mar. Syst.* 87:216–26
- Giering SLC, Cavan EL, Basedow SL, Briggs N, Burd AB, et al. 2020a. Sinking organic particles in the ocean—flux estimates from in situ optical devices. *Front. Mar. Sci.* 6:834
- Giering SLC, Hosking B, Briggs N, Iversen MH. 2020b. The interpretation of particle size, shape, and carbon flux of marine particle images is strongly affected by the choice of particle detection algorithm. *Front. Mar. Sci.* 7:564
- Gomes-Pereira JN, Auger V, Beisiegel K, Benjamin R, Bergmann M, et al. 2016. Current and future trends in marine image annotation software. *Prog. Oceanogr.* 149:106–20
- González P, Castaño A, Peacock EE, Díez J, Del Coz JJ, Sosik HM. 2019. Automatic plankton quantification using deep features. *J. Plankton Res.* 41:449–63
- Gorsky G, Ohman MD, Picheral M, Gasparini S, Stemmann L, et al. 2010. Digital zooplankton image analysis using the ZooScan integrated system. *J. Plankton Res.* 32:285–303
- Greer AT, Cowen RK, Guigand CM, Hare JA. 2015. Fine-scale planktonic habitat partitioning at a shelf-slope front revealed by a high-resolution imaging system. *J. Mar. Syst.* 142:111–25
- Greer AT, Cowen RK, Guigand CM, McManus MA, Sevdjian JC, Timmerman AHV. 2013. Relationships between phytoplankton thin layers and the fine-scale vertical distributions of two trophic levels of zooplankton. *J. Plankton Res.* 35:939–56
- Grosjean P, Denis K, Wacquet G. 2018. Zooimage: analysis of numerical plankton images. *CRAN*. <https://CRAN.R-project.org/package=zooimage>
- Grosjean P, Picheral M, Warembourg C, Gorsky G. 2004. Enumeration, measurement, and identification of net zooplankton samples using the ZOOSCAN digital imaging system. *ICES J. Mar. Sci. J. Cons.* 61:518–25
- Guidi L, Stemmann L, Jackson GA, Ibanez F, Claustre H, et al. 2009. Effects of phytoplankton community on production, size, and export of large aggregates: a world-ocean analysis. *Limnol. Oceanogr.* 54:1951–63
- Guidi L, Stemmann L, Legendre L, Picheral M, Prieur L, Gorsky G. 2007. Vertical distribution of aggregates (>110 μm) and mesoscale activity in the northeastern Atlantic: effects on the deep vertical export of surface carbon. *Limnol. Oceanogr.* 52:7–18
- Haddock SH. 2004. A golden age of gelata: past and future research on planktonic ctenophores and cnidarians. *Hydrobiologia* 530:549–56
- Hays GC, Richardson AJ, Robinson C. 2005. Climate change and marine plankton. *Trends Ecol. Evol.* 20:337–44
- Henrichs DW, Anglès S, Gaonkar CC, Campbell L. 2021. Application of a convolutional neural network to improve automated early warning of harmful algal blooms. *Environ. Sci. Pollut. Res.* 28:28544–55

- Horton T, Marsh L, Bett BJ, Gates AR, Jones DOB, et al. 2021. Recommendations for the standardisation of open taxonomic nomenclature for image-based identifications. *Front. Mar. Sci.* 8:620702
- Hu Q, Davis C. 2005. Automatic plankton image recognition with co-occurrence matrices and Support Vector Machine. *Mar. Ecol. Prog. Ser.* 295:21–31
- Hu Q, Davis C. 2006. Accurate automatic quantification of taxa-specific plankton abundance using dual classification with correction. *Mar. Ecol. Prog. Ser.* 306:51–61
- ICZN (Int. Comm. Zool. Nomencl.). 2017. Declaration 45—addition of Recommendations to Article 73 and of the term “specimen, preserved” to the Glossary. *Bull. Zool. Nomencl.* 73:96–97
- IMEV (Inst. Mer Villefranche), LOV (Lab. Océanogr. Villefranche), Sorbonne Univ./CNRS. 2020. *Plankton community in Régent (680 μ m) net, Point B, Villefranche-sur-Mer, France*. Data Set, Eur. Mar. Obs. Data Netw. (EMODnet), updated July 29. <https://www.emodnet-biology.eu/data-catalog?module=dataset&dasid=6506>
- Irigoin X, Fernandes JA, Grosjean P, Denis K, Albaina A, Santos M. 2009. Spring zooplankton distribution in the Bay of Biscay from 1998 to 2006 in relation with anchovy recruitment. *J. Plankton Res.* 31:1–17
- Katija K, Troni G, Daniels J, Lance K, Sherlock RE, et al. 2020. Revealing enigmatic mucus structures in the deep sea using DeepPIV. *Nature* 583:78–82
- Kaye J, Heeney C, Hawkins N, de Vries J, Boddington P. 2009. Data sharing in genomics—re-shaping scientific practice. *Nat. Rev. Genet.* 10:331–35
- Kenitz KM, Orenstein EC, Roberts PLD, Franks PJS, Jaffe JS, et al. 2020. Environmental drivers of population variability in colony-forming marine diatoms. *Limnol. Oceanogr.* 65:2515–28
- Kiko R, Biastoch A, Brandt P, Cravatte S, Hauss H, et al. 2017. Biological and physical influences on marine snowfall at the equator. *Nat. Geosci.* 10:852–58
- Kiko R, Christiansen S, Schröder S-M, Koch R, Stemann L. 2018. PlanktonID – combining deep learning, in situ imaging and citizen science to resolve the distribution of zooplankton in major upwelling regions. In *Proceedings of the 10th International Conference on Ecological Informatics: Translating Ecological Data into Knowledge and Decisions in a Rapidly Changing World*, ed. J Gaikwad, B König-Ries, F Recknagel, p. 174. Jena, Ger.: Friedrich-Schiller-Univ. Jena (Abstr.)
- Krizhevsky A, Sutskever I, Hinton GE. 2012. ImageNet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems 25*, ed. F Pereira, CJC Burges, L Bottou, KQ Weinberger, pp. 1097–105. Red Hook, NY: Curran
- Lampitt RS, Hillier WR, Challenor PG. 1993. Seasonal and diel variation in the open ocean concentration of marine snow aggregates. *Nature* 362:737–39
- LeCun Y, Bengio Y, Hinton G. 2015. Deep learning. *Nature* 521:436–44
- Lindsay DJ. 2017. *Pukia obtusukai* sp. nov., a new species of “cydippid” ctenophore from Japan. *Plankton Benthos Res.* 12:61–65
- Lindsay DJ, Grossmann MM, Bentlage B, Collins AG, Minemizu R, et al. 2017. The perils of online biogeographic databases: a case study with the ‘monospecific’ genus *Aegina* (Cnidaria, Hydrozoa, Narcomedusae). *Mar. Biol. Res.* 13:494–512
- Lindsay DJ, Pagès F, Corbera J, Miyake H, Hunt JC, et al. 2008. The anthomedusan fauna of the Japan Trench: preliminary results from in situ surveys with manned and unmanned vehicles. *J. Mar. Biol. Assoc. UK* 88:1519–39
- Lindsay DJ, Umetsu M, Grossmann M, Miyake H, Yamamoto H. 2015. The gelatinous macroplankton community at the Hatoma Knoll hydrothermal vent. In *Subseafloor Biosphere Linked to Hydrothermal Systems*, J Ishibashi, K Okino, M Sunamura, pp. 639–66. Tokyo: Springer
- Lombard F, Boss E, Waite AM, Vogt M, Uitz J, et al. 2019. Globally consistent quantitative observations of planktonic ecosystems. *Front. Mar. Sci.* 6:196
- Luo JY, Irisson J-O, Graham B, Guigand C, Sarafraz A, et al. 2018. Automated plankton image analysis using convolutional neural networks. *Limnol. Oceanogr. Methods.* 16:814–27
- Luo T, Kramer K, Goldfob DB, Hall LO, Samson S, et al. 2005. Active learning to recognize multiple types of plankton. *J. Mach. Learn. Res.* 6:589–613
- Mackas DL, Denman KL, Abbott MR. 1985. Plankton patchiness: biology in the physical vernacular. *Bull. Mar. Sci.* 37:652–74

- Malkiel E, Alquaddoomi O, Katz J. 1999. Measurements of plankton distribution in the ocean using submersible holography. *Meas. Sci. Technol.* 10:1142–52
- Martini S, Larras F, Boyé A, Faure E, Aberle N, et al. 2020. Functional trait-based approaches as a common framework for aquatic ecologists. *Limnol. Oceanogr.* 66:965–94
- MBARI (Monterey Bay Aquar. Res. Inst.). 2019. Video Annotation and Reference System. *Monterey Bay Aquarium Research Institute*. <https://www.mbari.org/products/research-software/video-annotation-and-reference-system-vars>
- Miloslavich P, Bax NJ, Simmons SE, Klein E, Appeltans W, et al. 2018. Essential ocean variables for global sustained observations of biodiversity and ecosystem changes. *Glob. Change Biol.* 24:2416–33
- Moliner JC, Ibanez F, Souissi S, Buecher E, Dallot S, Nival P. 2008. Climate control on the long-term anomalous changes of zooplankton communities in the Northwestern Mediterranean. *Glob. Change Biol.* 14:11–26
- Muller-Karger FE, Miloslavich P, Bax NJ, Simmons S, Costello MJ, et al. 2018. Advancing marine biological observations and data requirements of the complementary essential ocean variables (EOVs) and essential biodiversity variables (EBVs) frameworks. *Front. Mar. Sci.* 5:211
- Nakamura Y, Somiya R, Suzuki N, Hidaka-Umetsu M, Yamaguchi A, Lindsay DJ. 2017. Optics-based surveys of large unicellular zooplankton: a case study on radiolarians and phaeodarians. *Plankton Benthos Res.* 12:95–103
- Northcutt CG, Athalye A, Mueller J. 2021. Pervasive label errors in test sets destabilize machine learning benchmarks. arXiv:2103.14749 [stat.ML]
- Ohman MD. 2019. A sea of tentacles: optically discernible traits resolved from planktonic organisms in situ. *ICES J. Mar. Sci.* 76:1959–72
- Ohman MD, Davis RE, Sherman JT, Grindley KR, Whitmore BM, et al. 2019. Zooglider: an autonomous vehicle for optical and acoustic sensing of zooplankton. *Limnol. Oceanogr. Methods* 17:69–86
- Olson RJ, Sosik HM. 2007. A submersible imaging-in-flow instrument to analyze nano- and microplankton: Imaging FlowCytobot. *Limnol. Oceanogr. Methods* 5:195–203
- Orenstein EC, Beijbom O. 2017. Transfer learning and deep feature extraction for planktonic image data sets. In *2017 IEEE Winter Conference on Applications of Computer Vision*, pp. 1082–88. Piscataway, NJ: IEEE
- Orenstein EC, Kenitz KM, Roberts PL, Franks PJ, Jaffe JS, Barton AD. 2020a. Semi- and fully supervised quantification techniques to improve population estimates from machine classifiers. *Limnol. Oceanogr. Methods* 18:739–53
- Orenstein EC, Ratelle D, Briseno-Avena C, Carter ML, Franks PJS, et al. 2020b. The Scripps Plankton Camera system: a framework and platform for in situ microscopy. *Limnol. Oceanogr. Methods* 18:681–95
- Pearson DL, Hamilton AL, Erwin TL. 2011. Recovery plan for the endangered taxonomy profession. *BioScience* 61:58–63
- Picheral M, Colin S, Irisson J-O. 2017. *EcoTaxa, a tool for the taxonomic classification of images*. <http://ecotaxa.obs-vlfr.fr>
- Picheral M, Guidi L, Stemmann L, Karl DM, Iddaoud G, Gorsky G. 2010. The Underwater Vision Profiler 5: an advanced instrument for high spatial resolution studies of particle size spectra and zooplankton. *Limnol. Oceanogr. Methods* 8:462–73
- Priyadarshi A, Smith SL, Mandal S, Tanaka M, Yamazaki H. 2019. Micro-scale patchiness enhances trophic transfer efficiency and potential plankton biodiversity. *Sci. Rep.* 9:17243
- Remsen A, Hopkins TL, Samson S. 2004. What you see is not what you catch: a comparison of concurrently collected net, Optical Plankton Counter, and Shadowed Image Particle Profiling Evaluation Recorder data from the northeast Gulf of Mexico. *Deep-Sea Res. I* 51:129–51
- Riedel A, Sagata K, Suhardjono YR, Tänzler R, Balke M. 2013. Integrative taxonomy on the fast track-towards more sustainability in biodiversity research. *Front. Zool.* 10:15
- Robinson KL, Luo JY, Sponaugle S, Guigand C, Cowen RK. 2017. A tale of two crowds: public engagement in plankton classification. *Front. Mar. Sci.* 4:82
- Robison BH, Reisenbichler KR, Sherlock RE. 2017. The coevolution of midwater research and ROV technology at MBARI. *Oceanography* 30(4):26–37
- Romagnan J-B, Aldamman L, Gasparini S, Nival P, Aubert A, et al. 2016. High frequency mesozooplankton monitoring: Can imaging systems and automated sample analysis help us describe and interpret changes

- in zooplankton community composition and size structure—an example from a coastal site. *J. Mar. Syst.* 162:18–28
- Romagnan J-B, Legendre L, Guidi L, Jamet J-L, Jamet D, et al. 2015. Comprehensive model of annual plankton succession based on the whole-plankton time series approach. *PLOS ONE* 10:e0119219
- Russakovsky O, Deng J, Su H, Krause J, Satheesh S, et al. 2015. ImageNet large scale visual recognition challenge. *Int. J. Comput. Vis.* 115:211–52
- Sandel V, Kiko R, Brandt P, Dengler M, Stemann L, et al. 2015. Nitrogen fuelling of the pelagic food web of the tropical Atlantic. *PLOS ONE* 10:e0131258
- Schlimpert O, Uhlmann D, Schuller M, Hohne E. 1980. Automated pattern-recognition of phytoplankton – procedure and results. *Int. Rev. Gesamten Hydrobiol. Hydrogr.* 65:427–37
- Schroeder S-M, Kiko R, Koch R. 2020. MorphoCluster: efficient annotation of plankton images by clustering. *Sensors* 20:3060
- Schulz J, Barz K, Ayon P, Luedtke A, Zielinski O, et al. 2010. Imaging of plankton specimens with the light-frame on-sight key species investigation (LOKI) system. *J. Eur. Opt. Soc. Rapid Publ.* 5:10017s
- Sieracki CK, Sieracki ME, Yentsch CS. 1998. An imaging-in-flow system for automated analysis of marine microplankton. *Mar. Ecol. Prog. Ser.* 168:285–96
- Sommer U. 2012. *Plankton Ecology: Succession in Plankton Communities*. Berlin: Springer
- Sosik HM, Olson RJ. 2007. Automated taxonomic classification of phytoplankton sampled with imaging-in-flow cytometry. *Limnol. Oceanogr. Methods* 5:204–16
- Sosik HM, Peacock EE, Brownlee EF. 2021. WHOI-Plankton: annotated plankton images - data set for developing and evaluating classification methods. *Woods Hole Open Access Server*. <http://hdl.handle.net/1912/7341>
- Stemann L, Boss E. 2012. Plankton and particle size and packaging: from determining optical properties to driving the biological pump. *Annu. Rev. Mar. Sci.* 4:263–90
- Stemann L, Claustre H, D’Ortenzio F. 2012. Integrated observation system for pelagic ecosystems and biogeochemical cycles in the oceans. In *Sensors for Ecology: Towards Integrated Knowledge of Ecosystems*, ed. J-F Le Galliard, J-M Guarini, F Gaill, pp. 261–78. Paris: CNRS
- Stemann L, Gorsky G, Marty JC, Picheral M, Miquel JC. 2002. Four-year study of large-particle vertical distribution (0–1000 m) in the NW Mediterranean in relation to hydrology, phytoplankton, and vertical flux. *Deep-Sea Res. II* 49:2143–62
- Stemann L, Hosia A, Youngbluth MJ, Soiland H, Picheral M, Gorsky G. 2008. Vertical distribution (0–1000 m) of macrozooplankton, estimated using the Underwater Video Profiler, in different hydrographic regimes along the northern portion of the Mid-Atlantic Ridge. *Deep-Sea Res. II* 55:94–105
- Sun H, Benzie PW, Burns N, Hendry DC, Player MA, Watson J. 2008. Underwater digital holography for studies of marine plankton. *Philos. Trans. R. Soc. A* 366:1789–806
- Takahashi K, Ichikawa T, Tadokoro K. 2015. Diel colour changes in male *Sapphirina nigromaculata* (Cyclopoida, Copepoda). *J. Plankton Res.* 37:1181–89
- Tang XO, Stewart WK, Vincent L, Huang H, Marra M, et al. 1998. Automatic plankton image recognition. *Artif. Intell. Rev.* 12:177–99
- Trudnowska E, Lacour L, Ardyna M, Rogge A, Irisson J-O, et al. 2021. Marine snow morphology illuminates the evolution of phytoplankton blooms and determines their subsequent vertical export. *Nat. Commun.* 12:2816
- Van Horn G, Mac Aodha O, Song Y, Cui Y, Sun C, et al. 2018. The iNaturalist species classification and detection dataset. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 8769–78. Piscataway, NJ: IEEE
- Vandromme P, Stemann L, Berline L, Gasparini S, Mousseau L, et al. 2011. Inter-annual fluctuations of zooplankton communities in the Bay of Villefranche-sur-mer from 1995 to 2005 (Northern Ligurian Sea, France). *Biogeosciences* 8:3143–58
- Vilgrain L, Maps F, Picheral M, Babin M, Aubry C, et al. 2021. Trait-based approach using in situ copepod images reveals contrasting ecological patterns across an Arctic ice melt zone. *Limnol. Oceanogr.* 66:1155–67
- Wiebe PH, Benfield MC. 2003. From the Hensen net toward four-dimensional biological oceanography. *Prog. Oceanogr.* 56:7–136

- WMO (World Meteorol. Organ.), UNEP (UN Environ. Programme), ICSU (Int. Counc. Sci.), UNESCO (UN Educ. Sci. Cult. Organ.), UNFCCC (UN Framew. Conv. Clim. Change). 2003. *The second report on the adequacy of the global observing systems for climate in support of the UNFCCC*. Rep. GCOS 82, WMO/TD 1143, WMO, Geneva
- Woodson CB, Litvin SY. 2015. Ocean fronts drive marine fishery production and biogeochemical cycling. *PNAS* 112:1710–15



Contents

The Goldilocks Principle: A Unifying Perspective on Biochemical Adaptation to Abiotic Stressors in the Sea <i>George N. Somero</i>	1
Physiological Consequences of Oceanic Environmental Variation: Life from a Pelagic Organism's Perspective <i>Mark W. Denny and W. Wesley Dowd</i>	25
The Balance of Nature: A Global Marine Perspective <i>Constantin W. Arnscheidt and Daniel H. Rothman</i>	49
Ecological Leverage Points: Species Interactions Amplify the Physiological Effects of Global Environmental Change in the Ocean <i>Kristy J. Kroeker and Eric Sanford</i>	75
The Biological Effects of Pharmaceuticals in the Marine Environment <i>Marica Mezzelani and Francesco Regoli</i>	105
The Functional and Ecological Significance of Deep Diving by Large Marine Predators <i>Camrin D. Braun, Martin C. Arostegui, Simon R. Thorrold, Yannis P. Papastamatiou, Peter Gaube, Jorge Fontes, and Pedro Afonso</i>	129
Environmental DNA Metabarcoding: A Novel Method for Biodiversity Monitoring of Marine Fish Communities <i>Masaki Miya</i>	161
The Enzymology of Ocean Global Change <i>David A. Hutchins and Sergio A. Sañudo-Wilhelmy</i>	187
Using Chlorophyll Fluorescence to Determine the Fate of Photons Absorbed by Phytoplankton in the World's Oceans <i>Maxim Y. Gorbunov and Paul G. Falkowski</i>	213
Temporal and Spatial Signaling Mediating the Balance of the Plankton Microbiome <i>Yun Deng, Marine Vallet, and Georg Pohnert</i>	239
The Physiology and Biogeochemistry of SUP05 <i>Robert M. Morris and Rachel L. Spietz</i>	261

Machine Learning for the Study of Plankton and Marine Snow from Images <i>Jean-Olivier Irisson, Sakina-Dorothee Ayata, Dbugal J. Lindsay, Lee Karp-Boss, and Lars Stemmann</i>	277
Earth, Wind, Fire, and Pollution: Aerosol Nutrient Sources and Impacts on Ocean Biogeochemistry <i>Douglas S. Hamilton, Morgane M.G. Perron, Tami C. Bond, Andrew R. Bowie, Rebecca R. Buchholz, Cecile Guieu, Akinori Ito, Willy Maenhaut, Stelios Myriokefalitakis, Nazlı Olgun, Sagar D. Rathod, Kerstin Schepanski, Alessandro Tagliabue, Robert Wagner, and Natalie M. Mahowald</i>	303
The History of Ocean Oxygenation <i>Christopher T. Reinhard and Noah J. Planavsky</i>	331
Organic Matter Supply and Utilization in Oxygen Minimum Zones <i>Anja Engel, Rainer Kiko, and Marcus Dengler</i>	355
Argo—Two Decades: Global Oceanography, Revolutionized <i>Gregory C. Johnson, Shigeki Hosoda, Steven R. Jayne, Peter R. Oke, Stephen C. Riser, Dean Roemmich, Tobsio Suga, Virginie Thierry, Susan E. Wijffels, and Jianping Xu</i>	379
Ventilation of the Southern Ocean Pycnocline <i>Adele K. Morrison, Darryn W. Waugh, Andrew McC. Hogg, Daniel C. Jones, and Ryan P. Abernathy</i>	405
Aquatic Eddy Covariance: The Method and Its Contributions to Defining Oxygen and Carbon Fluxes in Marine Environments <i>Peter Berg, Markus Huettel, Ronnie N. Glud, Clare E. Reimers, and Karl M. Attard</i>	431
Modeling the Morphodynamics of Coastal Responses to Extreme Events: What Shape Are We In? <i>Christopher R. Sherwood, Ap van Dongeren, James Doyle, Christie A. Hegermiller, Tian-Jian Hsu, Tarandeep S. Kalra, Maitane Olabarrieta, Allison M. Penko, Yasbar Rafati, Dano Roelvink, Marlies van der Lugt, Jay Veeramony, and John C. Warner</i>	457

Errata

An online log of corrections to *Annual Review of Marine Science* articles may be found at <http://www.annualreviews.org/errata/marine>