IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received June 30, 2021, accepted August 23, 2021, date of publication August 30, 2021, date of current version September 9, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3109216

Sustainable Marine Ecosystems: Deep Learning for Water Quality Assessment and Forecasting

ÁNGEL FERNÁNDEZ GAMBÍN, EDUARD ANGELATS, JESÚS SORIANO GONZÁLEZ, MARCO MIOZZO, AND PAOLO DINI
Centre Tecnològic de Telecomunicacions de Catalunya (CTTC/CERCA), 08860 Castelldefels, Barcelona, Spain

Corresponding author: Ángel Fernández Gambín (angel.fernandez@cttc.es)

ABSTRACT An appropriate management of the available resources within oceans and coastal regions is vital to guarantee their sustainable development and preservation, where water quality is a key element. Leveraging on a combination of cross-disciplinary technologies including Remote Sensing (RS), Internet of Things (IoT), Big Data, cloud computing, and Artificial Intelligence (AI) is essential to attain this aim. In this paper, we review methodologies and technologies for water quality assessment that contribute to a sustainable management of marine environments. Specifically, we focus on Deep Leaning (DL) strategies for water quality estimation and forecasting. The analyzed literature is classified depending on the type of task, scenario and architecture. Moreover, several applications including coastal management and aquaculture are surveyed. Finally, we discuss open issues still to be addressed and potential research lines where transfer learning, knowledge fusion, reinforcement learning, edge computing and decision-making policies are expected to be the main involved agents.

INDEX TERMS Sustainable coastal management, sustainable aquaculture, remote sensing, artificial intelligence, machine learning, water quality, blue economy.

I. INTRODUCTION

Marine habitats are of major importance and their ecological sustainability is increasingly threatened by natural but also cumulative and mainly anthropogenic pollutants from land-based sources. Pollution can be localized, including discharge from a shore-based industrial wastewater treatment plant, a ship or other offshore structure (e.g., an oil platform), or coming from many diffuse sources such as stormwater runoffs or atmospheric depositions [1]. In any case, all these elements can have devastating effects in coastal ecosystems, where poor water quality puts these habitats at risk, directly impacting on their communities who depend on them to sustain and support economic livelihoods [2].

The proportion of the world's population that lives in coastal regions and other water bodies is over 40% and increasing [3]. Changes in these delicate areas due to human activities can endanger the aquatic organisms habitats and threaten the long-term ecosystem sustainability. In this way, water quality assessment has become one of the key pillars to guarantee a sustainable society, where the concentration of contaminants in water bodies needs to be taken under more precise control. Moreover, coastal zones are highly dynamic and productive areas with high ecological and economical value. Their management is of vital importance, having to safeguard the social, economic and environmental functions of the ecosystem. The monitoring and forecasting of coastal ecosystems must involve multidisciplinary and multi-scale observing systems [4]. Accordingly, the development of a sustainable coastal ecosystem represents an urgent short-term priority in order to preserve its integrity, functioning, and natural resources.

Sustainable coastal management, despite increased binding legislation, scientific efforts, and investment, still presents numerous implementation problems, including working with incomplete data for the region under study, insufficient understanding of the scale at which coastal systems operate and aggregating simultaneous monitoring of different activities. Proposed solutions point to a need for more integrated approaches supported by coordinated and multidisciplinary processes [5]. Extensive research about coastal sustainability issues has highlighted opportunities for Earth Observation (EO) data and technologies to support coastal management efforts, particularly in relation to environmental monitoring, including Harmful Algal Blooms (HABs) detection, disaster response supported by early warning systems, flood

The associate editor coordinating the review of this manuscript and approving it for publication was Chao Shen.

monitoring, maritime safety and security issues such as vessel and oil spill detection, and sea-state forecasting [5].

Moreover, climate change is posing additional pressure on coastal ecosystems due to variations in water bio-geochemical and physico-chemical parameters (e.g., pH, salinity) and leading to aquatic ecosystem degradation. These delicate environments are very sensitive to changes in climate. In recent years, they are suffering different impacts including loss of habitat forming species (e.g., coral reefs, seagrasses) [6], ocean productivity decline and changes in the geographic distribution of marine organisms [7]. The increase of sea surface temperature and, thus, the ice melting in polar regions are the main causes of these effects, unleashing changes in marine currents and variation in water characteristics that may exceed the ecosystems tolerance for sustainability. In this sense, the assessment of the environmental status of marine waters is essential to understand how marine ecosystems respond to climate change. However, this research is still far behind in comparison with terrestrial ecosystems, due to the dimension, complexity and variability of seas and to the lack of long time series of relevant variables [7]. Besides, while the only true way to mitigate climate change is to reduce the reliance on carbon-based sources of energy, research is needed to assess where it may be possible to ameliorate the climate effects through management of local stressors (e.g., reduction of nutrient pollution) or biological communities (e.g., marine protected areas and kelp re-seeding to increase resilience) [8].

The agri-food industry plays also a fundamental role in here. During the last decades, seafood consumption has more than doubled, from 9.9kg per capita in the 1960s to 20.2kg on average [9]. Finfish, shellfish and algae farming is one of the world's fastest growing food sectors, providing the planet with about half of all the fish we eat. This surge is due to the nutritional importance of fish, together with technological advances, which enable easier access to seafood products. However, geographic demand is not aligned with geographic production. In fact, aquaculture is heavily concentrated in Asia, whereas United States and Europe are two of the top three markets of seafood consumption. This trend indicates a large opportunity for aquaculture expansion, that has to be achieved preserving the marine resources. For instance, the European Commission have suggested that 48% of the European Maritime and Fisheries Funds will be used to fund sustainable aquaculture and fisheries, which is one of the main pillars of the European Blue Economy and the European Green Deal to avoid over-exploitation and biodiversity related problems of marine environments [10]. In this way, the development of a sustainable blue economy, and so a sustainable aquaculture, is a key priority of the European Commission and the United Nations Sustainable Development Goals (SDGs) (SDG 14 directly and other SDGs indirectly).

Therefore, in order to achieve sustainable marine ecosystems through an appropriate management of the available resources in coastal regions and aquaculture, *water quality assessment and forecasting* are the cornerstone. Water quality refers to the physical, chemical and biological content of water and may vary geographically and seasonally affected by many factors such as nutrient inputs, land mineral runoff, contamination from farming practices or changes in the flow regimes and hydrology [11]. Accomplishing this task in an efficient and smart manner involves several cross-disciplinary approaches. Recent advances in satellite technology, digital mapping, ecological modeling, open data, and connectivity mean that global-level water quality monitoring and planning systems for coastal and aquacultures areas may now be possible. In recent years, the European Commission promoted and financed the Copernicus Marine Environment Monitoring Service (CMEMS) and the European Marine Observation and Data Network (EMODnet) to facilitate the access to that information. Hence, a combination of Remote Sensing (RS) techniques integrated with Internet of Things (IoT), EO Big Data, cloud computing, and Artificial Intelligence (AI) are a powerful tool chain to tackle all the challenges present in marine ecosystems. They provide real-time data collection, quantitative decision-making, intelligent control, precise investment and personalized services, allowing the aggregation and advanced data analytics that leads to the ability of making smart decisions, preventing natural and anthropogenic disasters that put sustainability at risk [12].

In this way, data-driven intelligence methods are able to transform the collected data into manageable valuable information. Specially, Machine Learning (ML) is the essential technology in this regard given its ability to process data automatically and its unique characteristics pertaining to classification, modeling and forecasting [13]. Nevertheless, the traditional ML has a strong dependency on manually-extracted features, i.e., a human-expertise intervention. Moreover, due to the high spatio-temporal variability of the water quality parameters, with complex nonlinear intrinsic characteristics, and the improvements in data collection, a higher measurement frequency is needed, and now possible thanks to the aforementioned technologies. This translates into massive volumes of data to be processed. In this regard, conventional ML algorithms are not sufficient to provide a reasonable performance in terms of accuracy, mainly due to their lack of scalability. As a breakthrough in AI, the cutting-edge Deep Learning (DL) has overcome these limitations. It improves the data processing by automatically extracting highly nonlinear, complex and hidden features via sequences of multiple layers rather than requiring handcrafted optimal feature representations for a particular type of data based on domain knowledge [14]. Besides, it is actually designed to work with large datasets. This provides advanced analytical tools for a better understanding of the enormous amounts of information collected about marine waters, and a powerful tool for coastal and marine planning and management.

A. PAPER CONTRIBUTIONS

The goal of this paper is to review methodologies and technologies for water quality assessment that contribute to a sustainable management of marine environments. Specifically, we focus on DL schemes for water quality estimation and forecasting, presenting also different applications within coastal areas and aquaculture sites. In this regard, the main contributions of this work are:

- We stress the problematic situation that marine habitats are facing due to mainly human activities, and the need to go for a sustainable exploitation of these ecosystems through an intelligent management, where water quality is essential.
- We review techniques for assessment, estimation, and forecasting of marine water quality on several scenarios, focusing on coastal regions, estuaries and aquaculture sites. Our survey leverages on an integrated use of RS, in situ measurements and AI through deep learning architectures.
- We discuss open research issues still to be addressed in this topic and present future research lines to handle the spotted challenges.

B. RELATED WORK DISCUSSION

Throughout the entire paper, related literature is reviewed on each section to provide a better understanding of the discussed subject. Moreover, some tables are supplied as visual support in this regard. Apart from that, several surveys and reviews are proposed as references to the reader for further research on near-related matters. It should be noted that in these cases, the addressed topics are not intended to be the focus of the present article, and therefore, these works do not represent an obstacle to the novelty of this manuscript. For instance, the authors in [15] and [16] discuss about how DL can be integrated with RS in other matters rather than water quality, including image fusion and scene classification. Besides, the authors in [17] and [18] tackle water quality management by applying AI techniques. However, they both focus on fresh waters such as rivers, lakes and groundwaters. These topics are out of our scope, which is only focused on marine ecosystems, coastal and ocean waters. A cross-disciplinary review of DL for water resources is presented in [19], where the goal is to provide scientists and hydrologists with a technical overview, progress update, and a source of inspiration about the relevance of DL to water. Nonetheless, water quality is not tackled. Finally, the authors in [12] and [20] elaborate on the possibilities that DL can offer to aquaculture, where water quality is briefly mentioned and not the main focus.

The rest of the paper is organized as follows. In Section II, we provide a general perspective on the mainstream DL tools. Several water quality considerations are discussed in Section III. An outline about sensing technologies for water quality assessment is discussed in Section IV. Scenarios of applications are reviewed in Section V. In Section VI, open issues and future research lines are addressed. Finally, conclusions are provided in Section VII.

II. DEEP LEARNING OVERVIEW

Machine Learning created new opportunities to unravel, quantify, and understand data-intensive processes. ML is defined as a scientific field that seeks to give machines the ability to learn without being strictly programmed [12]. In recent years, Deep Learning has abruptly interrupted in many fields like computer vision, speech recognition, and natural language processing. Actually, DL is an old branch of AI based on Artificial Neural Networks (ANNs) that has been renewed due to factors like algorithmic advancements, high-performance computing, and Big Data [21]. The idea of DL is simple: the machine is learning automatically the features and decision making, versus a human-tailored learning system.

Conventional ML approaches utilize a processing chain that usually starts with human-coded feature extraction, a feature optimization stage, and then processing on the extracted features. These architectures are mostly ''shallow'', i.e., they usually had only one to two processing layers between the input features and the output. Shallow learners, including Support Vector Machines (SVMs), Gaussian Mixture Models, Hidden Markov Models, Conditional Random Fields, have been the backbone of traditional research efforts for many years [21] and still they provide remarkable results in many problems. In contrast, the strongest and potential characteristic of DL architectures is the learning based on multiple levels of representation, which can obtain a rich variety of highly complex, nonlinear and hierarchical hidden features automatically from raw inputs, and transform the lower-level representation into a higher more abstract one [20].

The goal of this section is to provide a general overview about ML, to better understand DL concepts. It is not intended to provide thorough details, but just general descriptions and terminology to facilitate the reading of the paper. For more details, we refer the reader to [14], [22]–[25].

ML is essentially a form of applied statistics with increased emphasis on the use of computers to statistically estimate complicated functions and a decreased emphasis on proving confidence intervals around these functions [14]. A ML algorithm is compound of a dataset, a model, a cost function and an optimization algorithm, and is able to learn from data [14]. A *dataset* is defined as a collection of inputs, which are in turn collections of *features* that have been quantitatively extracted from an event or object that we want the algorithm to process. A *model* is defined as the goal of the algorithm itself and is constructed by the learning process, i.e., the training procedure. A *cost function* measures the model performance for a given dataset. For example, in regression problems, it quantifies the error between predicted and expected values. Depending on the problem, it can be formed in many different ways, and the purpose is to be either: (i) minimized, finding the model parameters that provide the

TABLE 1. Summary of ML categories.

Learning	Supervised	Unsupervised	Hybrid	Reinforcement
Problem type	Classification, regression	Clustering, association	Clustering & Classification	Reward-based
Data	Labeled	Unlabeled	Both	No predefined
Training	External supervision	No supervision	Both	No supervision
Approach	Maps inputs to outputs	Finds patterns	Find patterns for labeling to ease mapping	Follows exploitation vs exploration

smallest possible returned value (usually called cost, loss or error); (ii) maximized, finding the model parameters that provide the highest possible returned value (reward). Finally, an optimization algorithm refers to the learning procedure, i.e., training the algorithm to find the optimal values for the model parameters that minimize/maximize the cost function. Most traditional ANNs schemes and thus DL algorithms, are based on an optimization algorithm called *stochastic gradient descent*. A recurrent problem in ML is that large training datasets are necessary for good generalization, that is, the algorithm performs well with every type of unseen data, but these large datasets are also more computationally expensive. At the same time, most of the used cost functions in ML decompose as a sum over training inputs. This translates into a prohibit computation time just to compute a single step in the gradient descent when the dataset barely grows [14]. Hence, the key point about the stochastic gradient descent is that the gradient is obtained as an expectation, and then it can be approximated using a reduced amount of samples per time step. Although, it was consider slow or unreliable, it has demonstrated remarkable results for ML and DL algorithms. Actually, it may not be guaranteed to arrive at even a local minimum in a reasonable time, but if often finds a very low value of the cost function quickly enough to be very useful [14]. A fast algorithm to perform the gradient descent, and the workhorse of learning in ANNs and DL schemes, is called the Backpropagation (BP) algorithm. It is probably the most fundamental building block in a neural network. It was first introduced in 1960s and almost 30 years later popularized thanks to the work in [26]. The algorithm is used to effectively train a neural network through a method called *chain rule*. In simple terms, after each forward pass through a network, backpropagation performs a backward pass while adjusting the model's parameters, known as *weights* and *biases*.

By learning, the definition from [27] can be used: "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks T, as measured by P, improves with experience E". ML tasks are usually described in terms of how the learning system should process an input dataset. Many kinds of tasks can be solved with ML, including classification, regression, association among others. Regarding the performance measure, they are specific to the carried out task, but the main idea is to understand how well the algorithm is performing on data that it has not seen before, i.e. real case scenario. This is why the input data for a ML algorithm is divided between *training* and *test* datasets, and the latter is used to estimate the real case scenario. The choice of a performance measure that exactly corresponds to the desired behavior of the system is often difficult. Depending on what kind of experience the ML algorithms are allowed to have during the learning process, i.e. what type of input dataset they are dealing with, four main categories can be classified:

Supervised Learning: The training dataset is labeled, i.e., a collection of samples tagged with a specific desired outcome that the learning algorithm should come up with on its own. There are two main areas where supervised learning is useful: classification and regression. Classification problems ask the algorithm to predict a discrete value, identifying the input data as the member of a particular class or group. Regression problems look at continuous data.

Unsupervised Learning: The training dataset is unlabeled, i.e., a collection of samples without a specific desired outcome. The learning algorithm then attempts to automatically find patterns in the data by extracting useful features and analyzing its structure. Several areas are identified: clustering, Anomaly Detection (AD), association. Clustering models look for training data that are similar to each other and groups them together. AD models can be used to flag outliers in a dataset. Finally, association models can predict other attributes which are commonly associated at a certain data point.

Hybrid or Semi-Supervised Learning: The training dataset has both labeled and unlabeled data. This method is particularly useful when extracting relevant features from the data is difficult, and labeling examples is a time-intensive task for experts.

Reinforcement Learning: There is no predefined dataset. The algorithm is learning how to map situations to actions within a certain environment on their own, so as to maximize a numerical reward. In this way, it is not told which actions to take, but instead must discover which actions yield the most reward by trying them. In the most challenging cases, actions may affect not only the immediate reward but also the long-term one. The tradeoff between trial-and-error search and delayed reward, also well known as "exploitation vs exploration", is the most important feature of reinforcement learning [23]. Table 1 summarizes the aforementioned ML categories.

Most of the reviewed papers in this survey are addressing supervised learning, i.e., there exist a dataset, labeled usually through in situ measurements, and performing regression tasks. Besides, there are also some unsupervised cases where for example anomaly detection is tackled. However,

TABLE 2. Summary of presented DL architectures.

Architecture	Main Characteristics
MLP	Feed-forward neural network, nonlinear activation functions, tabular data, fixed input length
RNN	Recursive inference, LSTM & GRU cells, sequential data, variable input length, parameter sharing, good for time series
CNN	Feed-forward neural network, convolutional kernels, image data, fixed input length, parameter sharing, spatial relationship, good for images
AE	Feed-forward neural network, encoder & decoder blocks, combined with MLPs & CNNs & RNNs, good for dimensionality reduction
DBN	Generative graphical model, RBMs kernels, good for images combined with CNNs
GAN	Generative model, generator & discriminator, Nash equilibrium game theory, good for data generation

we believe reinforcement learning within marine environments is still an open research topic to be undertaken where valuable applications can be developed. More about this will be discussed in Section VI. Some examples of the most common architectures in DL for the aforementioned categories are Multilayer Perceptrons (MLPs), Recurrent Neural Networks (RNNs), Convolutional Neural Networks (CNNs), RBMs, Deep Belief Networks (DBNs), Autoencoders (AEs) and General Adversarial Networks (GANs). In the following, we briefly present them. A summary with key ideas for each one of the architectures is provided in Table 2.

A. MULTILAYER PERCEPTRON

A feed-forward (the input signal goes only one way within the network) ANN is used to generate a set of outputs from a set of inputs through a set of extracted features. It is formed by an input layer, a certain number of intermediate hidden layers, and an output layer, and represents an efficient way to learn linear and nonlinear relationships between input and output pairs. Neural networks, when formed by many stacked layers, can represent complex features in later layers by using simpler representations formed by earlier layers in the network [18]. Each layer within an ANN comprises at least one neuron. An ANN is a network of these neurons connected to each other with some weights where these neurons run specific functions, called activation functions, mapping its input to an output. Stacked on top of each other, the series of functions runs over the input of the network and translates the input to the output in the output layer. Typically, each neuron within a layer runs the same activation function, defining the layer type. The network type is determined by the combination of used layers and how neurons are connected to each other within and between layers [18]. An MLP is the simplest feed-forward ANN, characterized by, at least, three layers of nodes: an input, a hidden layer and an output, and these layers are connected as a directed graph between the input and output layers. Except for the input nodes, each node is a neuron that uses a nonlinear activation function to produce an output. MLP utilizes the BP for training, i.e. to find the weights for each neuron within the network that minimize a certain error objective function. It can distinguish data that is not linearly separable [28].

B. RECURRENT NEURAL NETWORK

RNNs are a type of neural network that contain loops, i.e., working recursively, allowing information to be stored within the network and devised for handling temporal and predictive problems. The difference of a recurrent layer from a regular fully connected hidden layer is that neurons within a recurrent layer could also be connected to each other, i.e. the output of a neuron is connected both to the neuron(s) within the next layer and to the next neuron within the same layer [18]. Long Short-Term Memory (LSTM) networks are a particular kind of RNN, with strong capabilities in learning long-term dependencies. The neurons in the hidden layers of an LSTM are Memory Cells (MCs). A MC has the ability to store or forget information about past network states by using structures called *gates*, which consist of a cascade of a neuron with sigmoidal activation function and a pointwise multiplication block. Thanks to this architecture, the output of each memory cell possibly depends on the entire sequence of past states, making LSTMs suitable for processing time series with long time dependencies [29]. Gate Recurrent Unit (GRU) is a similar method to LSTM units but simpler to compute and implement [30], while keeping the same efficacy. Although this does not mean GRU always performs better than LSTM. A typical GRU cell is composed by two gates: reset and update gate. The reset gate is used to decide how much past information to forget. The update gate acts similar to the forget and input gate of an LSTM. It decides what information to throw away and what new information to add. It thus helps the model to determine how much of the past information (from previous time steps) needs to be passed along to the future. More details can be found in [31].

C. CONVOLUTIONAL NEURAL NETWORK

CNNs are feed-forward deep neural networks differing from fully connected multilayer networks for the presence of one or more convolutional layers. At each convolutional layer, a number of kernels is defined. Each of them has a number of weights, which are convolved with the input in a way that the same set of weights, i.e., the same kernel, is applied to all the input data, moving the convolution operation across the input span. Note that, as the same weights are reused (shared weights), and each kernel operates on a small portion of the input signal, it follows that the network connectivity structure is sparse. This leads to advantages such as a considerably reduced computational complexity with respect to fully connected feed-forward neural networks [28]. Frequent layers in CNN architectures are Rectified Linear Unit (ReLU), used as activation layer to recognize nonlinear correlations, and Max Pooling, which reduces the input size while keeping the positional knowledge intact. CNNs are good architectures for DL tasks with images or image-like objects as inputs, with important breakthroughs in the fields of object detection, super-resolution, image classification, and computer vision [18].

D. AUTOENCODER

An autoencoder is a network designed to learn useful features from unsupervised data. One of the first applications of AEs was dimensionality reduction, which is required in many RS applications. By reducing the size of the adjacent layers, the AE is forced to learn a compact representation of the data. An AE can be divided into two parts: encoder and decoder. Each part can be regarded as several hidden layers between an input and an output layer. The encoder reduces the dimension of the input high-dimensional sample data to output the low-dimensional encoded data. This output is then taken as the decoder input. The output result, which has the same dimension as the input high-dimensional data of the encoder, is obtained through the dimension-raising operation of the decoder. BP algorithm is used to update the weight of the hidden layer to make the AE output as close as possible to the input. This condition is also used as a criterion to evaluate AE performance [32]. Note that previous DL architectures, e.g. RNNs and CNNs, can be used as autoencoders to build more complex models for specific applications. Some examples can be found in [33] and [34].

E. DEEP BELIEF NETWORK

A Deep Belief Network is a type of generative graphical model, the join between probability and graph theory. Classic DBN contains several RBMs and ends with a backpropagation layer. RBMs are stochastic ANNs: as opposed to assigning discrete values, the model assigns probabilities based on the Boltzmann distribution. Moreover, they are shallow, i.e., each RBM includes just a visible layer and a hidden layer. Each layer contains a certain number of neurons. Connections are found between RBMs but none between units within layers. That is, there is no intra-layer communication, which is the restriction that the name implies. To build a DBN, the upper hidden layer of a RBM serves as the next RBM visible layer input. During the training, the DBN is trained layer by layer, and the weights of each layer are fixed layer by layer to obtain the approximate weight. Finally, the BP algorithm is used to fine-tune the weight optimization and attain the final DBN model [32]. DBNs are used to recognize, cluster and generate images, video sequences and motion-capture data.

F. GENERAL ADVERSARIAL NETWORK

General Adversarial Networks are a powerful class of generative models introduced in 2014 [24]. The main idea of GAN comes from the Nash equilibrium in game theory, inspired by the two-player zero-sum game, in which the total gains of two players are zero. They use two neural networks, pitting one against the other, to generate new, synthetic instances of data that can pass for real data. GANs often comprise a generator and a discriminator that learn simultaneously. The generator tries to capture the potential distribution of real samples, and generates new data samples. The discriminator is often a binary classifier, discriminating real samples from the generated samples as accurately as possible. Both the generator and the discriminator can adopt the structure of deep neural networks. The GAN optimization process is a minimax game with the goal to reach Nash equilibrium. [35]. They are widely used in image, video and voice generation.

III. WATER QUALITY CONSIDERATIONS

Water quality evaluation is the process of determining the chemical, physical and biological characteristics of water bodies and identifying the possible contamination sources that degrade the quality of water, including waste discharges, pesticides, heavy metals, nutrients, microorganisms, and sediments [36]. In the following, we highlight the most relevant water parameters used to estimate water quality.

A. CHLOROPHYLL-*a*

Chlorophyll-a (Chl-a) is a specific form of chlorophyll used in oxygenic photosynthesis, and is found in plants, algae and cyanobacteria. Chl-a is the major indicator of trophic state because it acts as a link between nutrient concentration, particularly phosphorus, and algal production. While mainly reflecting green, Chl-a absorbs most energy from wavelengths of violet-blue and orange-red light, whose reflectance causes chlorophyll to appear green. Algal blooms, defined as high concentrations of phytoplankton (algae), are directly related to Chl-a concentration since it is an essential point in photosynthesis [36]. HABs are problematic algal blooms that cause associated environmental impacts and even toxicity in some cases. HABs have been a significant world-wide research topic over three decades, and they still continue to be of major concern, not only due to their considerable environmental and societal impact but also a recent significant increase in frequency reported around the world [37]. Many factors have been cited as causes of HABs, but are generally caused by favorable environmental conditions, including increasing nutrient levels, light availability, water column stratification and/or changes in water temperature [37].

B. COLORED DISSOLVED ORGANIC MATTERS

Colored Dissolved Organic Matter (CDOM) consists of naturally occurring, water-soluble, biogenic, heterogeneous organic substances that are yellow to brown, which exist in both fresh and saline waters. These compounds can color the water yellowish brown in high concentrations, and together with Chl-a and other sediments dominate the water color. CDOM absorbency spectrum overlaps the chlorophyll absorption, thus the increase in CDOM concentration mainly affects the reflectance values in the blue and green regions. This effect can complicate the use of Chl-a retrieval algorithms and phytoplankton production models [36].

C. TURBIDITY

Turbidity is the amount of particulate matter that is suspended in water. Turbidity measures the scattering effect that suspended solids have on light: the higher the intensity of scattered light, the higher the turbidity. Materials that cause water to be turbid include clay, silt, finely divided organic and inorganic matter, soluble colored organic compounds, plankton, microscopic organisms and others [38].

D. WATER TEMPERATURE

Water temperature regulates physical, chemical, and biological processes in water and air-water interactions. It influences the solubility, and thus availability of various chemical constituents in water. Most importantly, this parameter affects dissolved oxygen concentrations, thus oxygen solubility decreases with increasing water temperature. Moreover, since metabolism of aquatic organisms is directly related to water temperature, this parameter affects the ability of living organisms to resist certain pollutants.

E. SALINITY

Salinity is the dissolved salt content in a water body. Salinity together with temperature are important factors to identify the density of seawater, and in turn, density is a critical component driving the currents in the oceans. The role of ocean currents in moderating the climate is crucial, and thus, salinity is also critical to determine the global water balance, productivity forecast models, as well as evaporation rates. For example, when the salinity is relatively low, the mixed layer will be more stable, and the nutrient pump may be partially inhibited, possibly leading to reduced productivity or a delay in the onset of spring and autumn phytoplankton blooms [36].

F. DISSOLVED OXYGEN

Dissolved oxygen is a crucial water quality parameter that influences the living conditions of all aquatic organisms that require oxygen. The level of dissolved oxygen in water bodies can be affected by anthropogenic activities and natural occurrences. Rapidly moving water, such as in mountain streams or large rivers, tends to contain a lot of dissolved oxygen. Bacteria in water can consume oxygen as organic matter decays. Thus, excess organic material can cause an oxygen-deficient situation, driving hard living conditions for aquatic life [38].

G. SUSPENDED SEDIMENTS

Suspended sediments is the amount of soil moving along within a water stream. It is highly dependent on the speed of the water flow, as fast-flowing water can pick up and suspend more soil. An excess of sediment can harm the water quality of a stream [38]. The more suspended particles, the more difficult for light to travel through the water and therefore, the higher the water's turbidity. The complex nature of suspended substances in water changes the reflectance of the water body and causes variation in color. Hence, interpretation of remotely sensed data just based on the water color is not adequate and accurate. Turbidity and suspended matters are considered as important variables due to their linkage with incoming sunlight that in turn affects photosynthesis for growth of macrophytes and plankton [36].

H. pH

pH is a measure of the relative amount of free hydrogen and hydroxyl ions in water. Water that has more free hydrogen ions is acidic, whereas water that has more free hydroxyl ions is basic. The values of pH range from 0 to 14. Values less than 7 indicate acidity, whereas greater indicate a base. The presence of chemicals in the water, affects its pH, which in turn can harm the living organisms. For example, an even mildly acidulous seawater environment can harm shell cultivation [38].

I. ELECTRICAL CONDUCTIVITY

Conductivity is a measure of the ability of water to carry an electrical current. It is highly dependent on the amount of dissolved solids (such as salt) in the water. The more salt, the higher the conductance, e.g., marine waters has high electrical conductivity.

IV. SENSING TECHNOLOGIES

In this section, we provide a general outline about sensing techniques, including in situ and remote sensing, for acquiring and deriving water quality parameters. Moreover, in subsection IV-C, we review several methods based on the integration of RS, in situ measurements and data-driven computer processing (e.g., DL algorithms) for water quality assessment.

A. IN SITU SENSORS

In situ methods allow for measuring variables directly in the environmental medium in continuous or semi-continuous time intervals. Here, we present them depending on the type of device and the relevant water parameters that can be monitored. They have been used for years to measure physical-based parameters, such as oxygen, pH, conductivity, depth, and temperature in marine waters. Arrays of sensors are typically used in automated systems, either deployed from a ship or on a mooring. In situ sampling offers high-resolution and reliable measurements. Low-cost and commercially available sensors can predict total suspended solids concentrations based on high-frequency time series of turbidity, conductivity, and other water-level data. Spectrometers and water quality probes containing fluorescent detectors can be used to measure down-welling spectral irradiance, surface-water levels of dissolved nutrients, Chl-a, fluorescence, and turbidity at transects, while a towed body probe can be used in underway sampling. Surface light sensors can also be used to monitor profile-based ambient fluctuations [39].

The most common and widely available solute for chemical sensors is pH and nitrate. Traditionally, hydrochemistry monitoring has been conducted through automatic water-samplers, yet these are costly due to the need for regular sample collection and laboratory analysis, and such methods are limited by performance and reagent wastes. For other solutes, wet analytical chemistry remains the most viable method. There are also deployable optical sensors, such as fluorimeters, that are capable of measuring photosynthetic pigments and organic matters like Chl-a. Conventionally, the pH value is measured by measuring the potential difference between the working pH probe and the reference electrode. There is a direct correlation between the voltage output of the electrode and the pH value of the water sample [39].

Electrochemical sensors and biosensors are also potentially viable methods for water quality monitoring. Commercially available instruments use conductometric electrodes to measure salinity, sulfur, and potentiometer methods to detect oxygen and nitrous oxide. Micro-electro mechanical systems conjugated with microelectrode array sensors have been developed for phosphate detection and showed precise in situ measurements with a small amount of samples. However, due to the configuration of the array sensors, it may be fragile under high-flow or turbulent waters. The microfluidic devices can be integrated with electrochemical and optical sensors for monitoring heavy metals, nutrients, or pathogens. These devices require a small volume of samples, better processing control, reduced waste generation, and system compactness. Both electrochemical sensors and biosensors are highly specific, sensitive, and can work in a variety of matrices, but the industrial production and long-term deployment are often complicated by calibration and validation difficulties. Biosensors are not widely used in environment sensing, but have some potential regarding drinking water purification in water treatment plants for the detection of live organisms [39].

In spite of in situ sensing widespread use, there are many disadvantages that stem from manual handling and instrumentation required to collect the samples. Probes and plug-in devices are expensive and bulky. Also, in situ sampling is usually undertaken on vessels that require substantial costs in time, human efforts and financial support. Precision calibration equipment, e.g., stable reagents, supporting infrastructure such as flow systems, and interval-frequency consistency are among the major issues that affect in situ monitoring. Moreover, it is difficult to maintain calibration parameters during long-term deployments [39].

As a result of advances in several Information and Communications Technologies (ICT) technologies such as digital systems and micro-controllers, Wireless Sensor Networks (WSNs) have emerged as a powerful technology to automatically collect data in high volumes, thanks to the deployment of a large number of self-organized sensors within the areas of interest. Moreover, IoT has been developed in parallel to WSNs and provides interconnectivity and communication among the different agents, including people and objects such as sensors, actuators and computers, within the monitored system. The aims of intelligently identifying, monitoring, locating, tracking and controlling things are achieved by IoT [40]. In this way, the combination of both poses a perfect candidate to be used for sensing and monitoring purposes in marine environments. Some examples are provided next. The prototype and proof of concept of a distributed monitoring system of dissolved oxygen, pH and temperature is presented in [41] for water quality monitoring in an aquaculture pond. As a future work, the authors suggest the use of AI within the sensor network to implement an early-warning system based on real time events detection. A similar idea is proposed in [42] where a narrow band IoT system is used to monitor temperature, pH and dissolved oxygen within an aquaculture pond in Changzhou, China. Results show that the system provides real-time and accurate data transmission, which can meet the actual production needs and provide strong data and technical support for further water quality regulation and aquaculture production management. A broader concept is discussed in [43] where an aquaculture monitoring framework is devised including carbon outflow control, water quality check, environmental and power monitoring, and a web surveillance platform. A reconfigurable system based on a Field Programmable Gate Array (FPGA) design board and Zigbee wireless communication is presented in [40]. Analogous concept using Raspberry Pi technology is proposed in [44]. An IoT Waspmote microcontroller board is used in [45] to build a smart water quality monitoring system for Fiji islands. Moreover, a multiparameter oceanographic sensor package measuring pH, dissolved oxygen, salinity, temperature, and water depth module is depicted in [46]. Finally, authors in [47] propose the Arduino platform as the core controller.

B. REMOTE SENSING

With advances in space and computer sciences over recent decades, RS techniques have become useful tools to monitor and manage the water quality, by covering large scale regions and water bodies in a more effective and efficient manner. Moreover, the collected data is digital and, therefore, easily readable in computer processing. RS has been used since the 1970's and continue to be widely used in water quality assessment nowadays [36]. A useful classification of various RS observing sensors, located in satellites and airborne systems, commonly used in water quality assessments is presented in [36], along with their spectral properties including spatial resolution, spectral bands, and revisit interval. Also, a microwave radiometers categorization is proposed. The key point here is that the observing sensors of different types [1], [36] and located in different platforms, measure the amount of radiation reflected by the water surface. Thanks to that, the spectral characteristics of water and its included pollutants can be achieved, since they are function of hydrogeological, biological and chemical properties of the water. Hence, water quality can be assessed.

Sun's energy is either reflected (e.g., from the water's surface), as it is for visible wavelengths, or absorbed and then re-emitted, as it is for thermal infrared wavelengths. Remote sensing systems which measure energy that is naturally available are defined as *passive* sensors. In RS, passive sensors measure sea-surface salinity, sea-surface temperature and water-leaving radiance using microwave/infrared radiometers and optical sensors [48]. *Active* sensors, on the other hand, provide their own energy source for illumination. The sensor emits radiation which is directed toward the target to be investigated. The radiation reflected from that target is detected and measured by the sensor. Advantages for active sensors include the ability to obtain measurements anytime, regardless of day time or season. Active sensors can be used for examining wavelengths that are not sufficiently provided by the sun, such as microwaves, or to better control the way a target is illuminated. However, active systems require the generation of a fairly large amount of energy to adequately illuminate targets. In RS, active sensors measure sea-surface height (altimeter), sea-surface roughness (Synthetic-Aperture Radar (SAR)), sea-surface wind (scatterometer and SAR) [48].

A typical example is Ocean Color RS, which is related with the intensity and spectral distribution of visible light reflected out of the water to the biological and biogeochemical processes that influence the optical properties of the water column. For the last few decades, satellite data have been used to estimate large-scale patterns of phytoplankton biomass and productivity across the global ocean from daily to interannual timescales [49]. Moreover, ocean color measurements are increasingly being used for environmental monitoring of HABs, critical coastal habitats (e.g. seagrasses, kelps), eutrophication processes, oil spills, and a variety of hazards in the coastal zone that will be discussed throughout this paper. Regarding ocean color, two types of water are distinguished depending on its components and concentration [50]. Case-I waters are those whose optical properties are determined mainly by phytoplankton (Chl-a concentration as principal pigment) and other bio-genic components, while other water elements such as CDOM and mineral particles are not that relevant. In contrast, case-II waters are water bodies where excessive CDOM and suspended matter such as mineral particles are often present, which do not co-vary with bio-genic particles such as phytoplankton. Case-I waters typically include oligotrophic, open oceans and seas. Case-II waters often include turbid coastal waters and inland waters such as lakes. Thus, case-II waters are associated to coastal zones, where the earth-water interaction is important and the runoff of continental matter propitiates a more heterogeneous and complex aquatic environment where yellow substances and sediments contributions to water color have to be taken in account.

Remote sensing data can be obtained from multiple sources depending on the platform on which the observing sensors are located, such as satellite, sub-orbital (e.g., an aircraft, a balloon or a drone), and ground-based platforms. Satellites provide unique Earth observation capabilities because the cost and complexity of working with satellite data is often reduced given the significant infrastructure already implemented by space agencies and downstream data providers. This is often not the case for other platforms (e.g., sub-orbital). For example, flight operations are generally limited in duration and spatial coverage, and their application is usually limited to dedicated field campaigns of significant events. Further, sub-orbital data acquisition can be a far more complex proposition for users in terms of dealing with operating costs and complexities in scheduling, weather, and flight logistics, as well as crucial data processing, calibration, and validation activities. Similarly, there are also significant costs associated with installation, operation, and maintenance of ground-based platforms [1].

C. WATER QUALITY ASSESSMENT

In this subsection, we review several works that estimate water quality by combining RS, in situ measurements and DL. Moreover, we provide different examples focused on the retrieval of water constituents that define water quality, analyzing as well different issues present in remote sensing, such as gap-filling or atmospheric correction. A summary of the main reviewed works can be found in Table 3.

Ocean color measured from satellites provides daily global, synoptic views of spectral water-leaving reflectances that can be used to generate estimates of marine Inherent Optical Properties (IOP). These reflectances, namely the ratio of spectral upwelled radiances to spectral downwelled irradiances, describe the light exiting a water mass that defines its color. IOPs are the spectral absorption and scattering characteristics of ocean water and its dissolved and particulate constituents. Because of their dependence on the concentration and composition of marine water constituents, IOPs can be used to describe the contents of the upper ocean mixed layer, i.e. satellite ocean color data is used because its variability is primarily driven by biological processes related and correlated in complex, nonlinear relationships with the physical processes of the upper ocean. This information is critical to further our scientific understanding of biogeochemical oceanic processes, such as organic carbon production and export, phytoplankton dynamics, and responses to climatic disturbances [72]. Despite the current widespread availability of ocean color observations, mapping of ocean color is spatio-temporally limited and challenged by inconsistent information due to mainly cloud covers, particularly in polar regions [73]. These regions are usually covered by dense clouds throughout the year, limiting the valid range of satellite observations. As such, many gaps appear in the collected data for these areas. As a result, the demand for continuous

TABLE 3. Summary of water quality assessment works.

Task	Application	DL architecture	Ref.
Water params retrieval	Gap filling	Deep ANN	[51], [52]–[54]
Water params retrieval	Ocean color RS	Deep ANN	[55], [56]
Water params retrieval	Atmospheric correction	Deep ANN	[57], [58]
Water params retrieval	RTM inversion	Deep Gaussian Processes	[59]
Water params retrieval	Chl-a estimation	Deep ANN	[60]
Water params retrieval	Chl-a, turbidity, suspended solids estimation	Deep ANN	[61], [62]
Water params retrieval	Chl-a estimation	Mixture Density Network	[63]
Coastal management	Chl-a estimation	Deep ANN	[64], [65]
Coastal management	Seagrass conservation	Deep Capsule network	[66]
Coastal management	Shellfish reef conservation	CNN	[67]
Coastal management	Coral reef conservation	Deep ANN	[68]
Coastal management	Coral reef conservation	CNN	[69]
Coastal management	Oil spill detection	CNN	[70], [71]
Coastal management	Ocean color RS	CNN & AE	[48]

ocean color data on various spatial and temporal scales in the polar regions has increased. ML early studies, and later some DL ones, have attempted to reconstruct these data gaps. These approaches are more reliable, cost-effective and flexible than conventional parametric models because of its ability to handle non-linear relationships and complex interactions, which are often present in ecological data. The use of ANNs to fill gaps in satellite-derived ocean color data is introduced in [51], linking satellite with in situ physical observations. An ANN transfer function is trained using a two-year dataset from the Joint Polar Satellite System Visible Infrared Imaging Radiometer Suite, where daily Chl-a fields are used interpolated with the 9km resolution provided by NASA. Moreover, temperature and salinity profiles are obtained from the International Pacific Research Center at Hawaii. The study demonstrates that employing ANNs can provide an accurate, computationally cheap method for filling gaps in satellite observation fields and time series.

An early work dealing with ocean color remote sensing in coastal waters within the Yellow Sea is presented in [55]. The authors propose three ANN-based algorithms to retrieve the concentration of water constituents, including Chl-a, suspended matter and CDOM from remotely sensed data. Remote sensing analysis with neural networks is reviewed in [56]. The authors present an overview of the main concepts underlying ANNs, as well as main tasks that involve ANNs in RS. Another retrieval algorithm of ocean color remote sensing products from atmospheric-corrected Sentinel-3 Ocean and Land Colour Instrument (OLCI) is introduced in [58]. The algorithm consists of several specialized ANNs with task-optimized architectures (OLCI Neural Network Swarm). The products contain concentrations of water constituents, inherent and apparent optical properties and a sea color index. The algorithm makes use of a comprehensive fuzzy logic classification scheme. These two studies, as most ocean color remote sensing algorithms, retrieve the water constituent information by first performing atmospheric correction, then deriving the results from the achieved remote sensing reflectances. However, atmospheric correction models are still a challenging task in coastal waters due to their specific complexities such as absorbing aerosols. Actually, the authors of these papers claim that with better atmospheric correction models, their proposed algoithms could be improved. In order to overcome this, the authors in [57] propose an ANN-based reflectance model used to estimate ocean bidirectional reflectance. The results show that their algorithm could expedite the retrieval process and improve efficiency for use in global satellite observations.

Since spaceborne instruments can only measure the properties of electromagnetic waves emitted or scattered by the Earth, a prior knowledge is needed to understand where these waves originate from, how they interact with the environment, and how they propagate towards the sensor. Radiative Transfer Modeling (RTM) is developed to do so, assuming that everything is known about the radiation sources. This is the so-called direct problem. However, RS is an inverse problem, i.e., how to derive the environment properties, given the value of the electromagnetic measurements gathered in space. Currently, the most widely adopted techniques are based on statistical models, mostly nonlinear and non-parametric ML algorithms, applied to invert RTM simulations. The complexity of RTMs, highly nonlinear, and typically hierarchical, makes that very often shallow models cannot capture complex feature relations for inversion. This motivates the use of deeper architectures. In [59], the authors introduce the use of deep Gaussian Processes for bio-geo-physical model inversion, providing an efficient solution that scales well to big datasets. Their results show empirical evidence of performance for the estimation of surface temperature, Chl-a, inorganic suspended matter, and CDOM data acquired by the Sentinel-3 OLCI sensor.

Another example of coastal water parameters estimation can be found in [60]. In this case, RS data from MultiSpectral Imager (MSI) on board Sentinel-2 is used for retrieving Chl-a and suspended sediments along the Adriatic and Tyrrhenian coasts in Italy. The proposed algorithm is an MLP feed-forward ANN. This study confirms the potential of Sentinel-2 MSI products for coastal water monitoring, but it also highlights key issues to be further tackled such as the atmospheric correction impact, the need of reliable

TABLE 4. Summary of water quality forecasting works.

Task	Application	DL architecture	Ref.
Time series analysis	HAB	Deep ANN	[81]
Time series analysis	Early warning system	Deep ANN	[38]
Time series analysis	Chl-a concentration	Multi-wavelet ANN	[82], [83]
Time series analysis	Sea surface temperature	RNN	[89]
Water params forecasting	HAB	CNN & LSTM RNN	[37], [74], [75]
Water params forecasting	HAB	DBN	[76], [77]
Water params forecasting	HAB	Deep ANN	[78], [79], [80]
Water params forecasting	Aquaculture - Dissolved oxygen	LSTM RNN	[90], [91]
Water params forecasting	Aquaculture - Dissolved oxygen	LSTM RNN & sparse AE	[92]
Water params forecasting	Aquaculture - Dissolved oxygen	Deep ANN	[93]
Water params forecasting	Aquaculture - pH & water temp	Deep bi-directional SRU	[94]
Water params forecasting	Aquaculture - pH, water temp...	LSTM RNN	[95], [96], [97]
Water params forecasting	Sea surface temperature	LSTM RNN	[98], [99]
Water params forecasting	Sea surface temperature	GRU RNN	[100]
Water params forecasting	Sea surface temperature	GRU AE	[101]
Water params forecasting	Sea surface temperature	GRU CNN	[102]
Water params forecasting	Sea surface temperature	LSTM & CNN	[103]
Water params forecasting	Sea surface temperature	MLP & LSTM	[104]
Water params forecasting	Sea surface temperature	CNN	[105]
Anomaly detection	Coral reef protection	deep ANN	[106]
Anomaly detection	Coastal management	deep ANNs & Extreme Learning Machine (ELM)	[107]
Anomaly detection	Coastal management	Wavelet ANN	[108]

in situ measurements, and possible bathymetry effects near the shores. The study in [61] evaluates the potential of remote sensing using ML techniques for improving water quality estimation over the coastal waters of Hong Kong. Concentrations of suspended solids, Chl-a, and turbidity were estimated using an ANN. Random Forest (RF), Cubist regression, and Support Vector Regression (SVR) models were used as benchmarks. Satellite and in situ reflectance data are compared, where ANN-based methods provide best results. Similar approach is presented in [62] where the authors attempt to estimate Chl-a and suspended particulate matter concentrations, in coastal environments on the west coast of South Korea using Geostationary Ocean Color Imager (GOCI) satellite data. Finally, retrievals of near-surface Chl-a concentration are achieved in [63] through the Mixture Density Network technique.

Most of the surveyed works in this subsection make use of RS as the primary tool to obtain water parameter values, but nonetheless they combine it usually with in situ measurements to provide better performance, reinforcing the key idea that an integrated system is the best approach. Moreover, simple DL architectures, i.e. MLPs and ANNs with reduced number of hidden layers, are the most common strategy to solve the aforementioned problems. Besides, there is extensive literature dealing with these challenges where conventional ML approaches are used satisfactorily. This is mainly due to the reduced datasets available at the time of those studies, which makes the use of DL useless or even impracticable. Nevertheless, with the massive increase in data collection, we believe more works will appear in these years exploiting the benefits of DL in this regard. Finally, we refer the reader to [15] and [16] for further research on how DL can be integrated with RS in other matters rather than water quality assessment, e.g., image fusion, scene classification, object detection and land classification among others.

V. SCENARIOS OF APPLICATION

In this section, we have selected common use cases of sustainable marine ecosystems where DL plays a key role, and namely algal blooms, ecosystem preservation, oil spill and climate change issues including sea surface temperature and aquaculture. A summary of the reviewed works dealing with water quality assessment/estimation can be found in Table 3. Moreover, literature related to water quality forecasting is summarized in Table 4. As for this last item, a large volume of related research can be found. For example, the authors in [17] have conducted a comprehensive survey about ANN-based water quality prediction considering feed-forward, recurrent, and hybrid architectures. However, they mainly focus on fresh waters such as rivers, lakes, groundwater, and streams. Besides, another extensive review about DL techniques for hydrology and water resources is presented in [18]. They primarily focus on fresh water resources management, providing also insights about water quality for human consumption and urban environments. Floods, land and soil, groundwaters and weather forecasting are also tackled. These topics are out of our scope, but these two previous references can help the reader to delve deeper into topics related to this article. In our study, we focus only on marine ecosystems, coastal and ocean waters.

A. ALGAL BLOOMS

Harmful algal blooms are a global problem that is increasing worldwide in frequency and location. One way to mitigate their impacts on people's health and livelihoods is to develop early-warning systems. Models to predict and manage HABs typically make use of complex multi-model structures incorporating satellite imagery and frequent monitoring data with different levels of detail into hydrodynamic models. An example can be found in [37], where a detection and prediction system for spatio-temporal HAB events and using remote sensing data is proposed. They combine CNNs, LSTM components together with RF and SVM classification methods. Results provide high accuracy for detection and prediction in the proposed case study within the coastal waters of Florida. However, these sophisticated methods cannot be applicable when limited data are available. In these cases, empirical statistical models can be simpler alternatives but successful for HAB forecasting. In [74], an LSTM RNN is presented as a way to predict the occurrence time of Margalefidinium polykrikoides blooms in South Sea of Korea. Satellite data is used to extract sea surface temperature and photosynthetically available radiation, factors known to be related to HABs occurrence. Another example can be found in [75] where the authors present an early-warning system for the prediction of two types of HABs, fish kill and toxic bloom occurrences, in Bolinao-Anda, Philippines, using only in situ data. Results show that the most important predictive variable was a decrease in dissolved oxygen. Fish kills were more likely during higher salinity and temperature levels, whereas the toxic blooms occurred more at relatively lower salinity and higher chlorophyll conditions. A five-layered DBM to predict algal blooms based on phytoplankton density is presented in [76]. Their architecture is a generative model, i.e., a DBN that stacks several RBMs. This case study is conducted in coastal waters of East China using in situ data and a traditional ANN as a benchmark for comparison. Results show that the DL model yields better generalization and greater accuracy in predicting algal blooms.

In [77], the authors analyze related factors of the HABs disasters. Based on the forecasting ability of Autoregressive Integrated Moving Average (ARIMA) model and the powerful expression ability of DBN on nonlinear relationships, an hybrid model that combines both is proposed for HAB forecasting. The corresponding ARIMA model is built for each environmental factor in different coastal areas to describe the temporal correlation and spatial heterogeneity. The DBN serves to capture the complex nonlinear relationship between the environmental factors and the HAB biomass. Furthermore, Particle Swarm Optimization (PSO) is introduced to speed up the training phase. Finally, ship monitoring data collected in East China coastal areas is used as the experimental dataset. ANNs and Genetic Programming are used in [78] for the selection of input variables to predict dynamics of algal blooms in coastal waters of Hong Kong. The authors found that Chl-a is the most significant variable in predicting algal blooms, and in contrast to several previous studies, they claim that the use of algal biomass data alone as the input is good enough to perform forecasting, which might reduce the dependency on expensive equipment in algal bloom warning systems. A similar idea is developed in [79], where ANNs and SVMs are implemented, and improved by introducing different hybrid learning algorithms, to accurately forecast algal growth and eutrophication in Tolo Harbour, Hong Kong. Results reveal that the used methods could ensure robustness to learn complicated relationships between algal dynamics and different coastal environmental variables. The authors in [80] leverage on chemical analytical toxin data from coastal waters of Maine, USA to propose a high-resolution forecasting of paralytic shellfish toxin accumulation. In this regard, an ANN is used to provide weekly site-specific toxicity forecasts. The algorithm was trained on images constructed from a chemical fingerprint at each site including toxic compound data series, representing also past conditions. Results show good performance under several configurations. Besides, time horizon tests indicate a decline in accuracy beyond three-week forecast time.

ANNs and SVMs are used in [64] to develop an optimal Chl-a estimation model for coastal waters from Landsat-8 Operational Land Imager satellite images in the middle of the South Sea of Korea. This study provides practical information about effective monitoring systems for coastal algal blooms. An interesting approach is presented in [65] where spatial anisotropy caused by strong coastal-inland environmental gradients is investigated. A directional geographically spatial proximity ANN-based architecture is proposed to address the nonlinear effects of spatial anisotropy. A Chl-a dataset from Zhejiang coastal areas of China in the spring over 2015-2017 are used to evaluate the model. Insightful results allow a better understanding of the main drivers of HABs in the area. Moreover, the impacts of river discharges and ocean currents on Chl-a patterns could be characterized. These findings are quite conductive to formulate algae bloom mitigation strategies for managing coastal ecosystems.

The authors in [81] characterize and forecast the spatial and temporal variations of three sea water features, i.e., Chl-a, a photosynthesis index called fluorescence line height, and sea surface temperature to detect algal blooms, in the Arabian Gulf, and based on Moderate Resolution Imaging Spectroradiometer (MODIS) satellite data. They explore Seasonal ARIMA, ANNs, and linear regression. Moreover, they cover different types of water varying in depth and turbidity. Linear regression and ANNs are found to be the best at predicting Chl-a in all types of water (turbid and shallow). Meanwhile, the seasonal ARIMA model provides the best prediction of the two other considered water features. Prediction of future values for water quality variables, based on under-water sensors data, is performed in [38]. Several ML algorithms and the effect of including values from a varying number of past days are evaluated. As an interesting future work, they plan to integrate their water quality prediction scheme within an intelligent alert system to release early warnings based on predicted hydrological parameters.

Ensemble modeling is a suitable technique to decrease the bias and variance within predictions. Forecasts of different individual models are combined to achieve the desired output. In [82], two techniques including Bates-Granger and least square methods are applied to minimize the error of the forecasting models, which are based on multi-wavelet ANNs. A wavelet system is a set of building blocks to construct a certain signal or function. It has become a trendy analytical tool due to its ability to simultaneously elucidate both spectral and temporal information within the signal, providing valuable insights about the physical data structure. In this sense, it provides a time-frequency representation of a signal within many different time periods. This property overcomes the basic shortcoming of Fourier analysis, i.e., the Fourier spectrum contains only globally averaged data. The models are aimed to predict up to three days in advance the Chl-a concentration as well as water salinity within the coastal waters of Hilo Bay, Hawaii utilizing in situ data. To do so, previous daily time series up to three lags are decomposed via different wavelet functions to be applied as input parameters for the ANN. Results are promising, and as expected, while increasing the time horizon, the reliability and accuracy of the models decrease. Similar approach is presented in [83], where the discrete wavelet transform is combined with an ANN to forecast one month ahead Chl-a levels from in situ data collected within the South San Francisco Bay, USA. Moreover, multi linear regression and genetic algorithm SVR models are also investigated. Decomposed time series thanks to the wavelet transform are used as input data. The authors suggested as further research to improve the model by considering other input variables, e.g., PO4 or NO3, and to predict Chl-a in the second, third, or following months.

B. ECOSYSTEM PRESERVATION

Vegetated (marshes, mangroves, seagrasses, etc) and animal-derived (shellfish reefs, corals, etc) habitats are key components in coastal systems. They provide ecosystem services, including essential nursery habitat, food and shelter for fish and marine organism, carbon sequestration, sea bottom stabilization, improved water quality, and shoreline protection. However, these landscapes are at risk from the combined stress of climatic disasters, such as typhoons and rainfalls, and direct human-driven changes, such as the release of pollutants, making effective management and conservation increasingly crucial [67]. Capsule ANNs were proposed in [84] to address several limitations that CNNs present, such as the invariance caused by pooling and the inability to understand spatial relationships between features. The primary model comprises only one convolution layer and one fully-connected capsule layer [85]. The authors in [66] present a deep Capsule network for classification of seagrass in high-resolution multispectral satellite images within Florida coastal waters. The proposed model outperforms CNN and SVM. Moreover, they propose an interesting deep transfer learning scheme with the goal of detecting seagrass at any location in the world. Same idea is pursued in [86]. In this case, they detect seagrasses from in situ underwater images collected within shallow coastal waters of Western Australia. They present a faster Region-based CNN (R-CNN) model to do so. R-CNNs are a pioneering approach that applies deep models to object detection, by first selecting several proposed regions from an image and label them. Then, they use a CNN to perform forward computation to extract features from each proposed area and predict their categories. The main performance bottleneck of an R-CNN model is the need to independently extract features for each proposed region. As these regions have a high degree of overlap, independent feature extraction results in high computation time. Fast R-CNN improves on the R-CNN by only performing CNN forward computation on the image as a whole [87]. Other examples focusing on automatic species detection based on DL classification of underwater images are [67] where the authors detect and delineate oyster reefs, and [68], [69] where annotation of marine coral species is performed.

C. OIL SPILLS

Sea oil pollution is considered a major threat to oceanic and coastal ecosystems. Accidents at offshore oil drilling platforms or oil pipeline networks can provoke severe oil spills. Yet, illegal discharges of ballast and tank cleaning oily residues from oil tankers and ships are the main sources of relative pollution events [70]. The detection of oil slicks and early warning of the corresponding authorities is vital to attenuate the environmental disaster. Remote sensing has a crucial role towards this objective. In this regard, SAR sensors are commonly used for this objective due to their capability for operating efficiently regardless of the weather and illumination conditions. In particular, SAR sensors can be successfully used to measure sea surface roughness. In fact, sea surfaces covered with oil films appear dark in SAR images because the capillary waves and short gravity waves, that contribute to the sea surface roughness, are damped by the surface tension of oil films [71]. However, black spots, probably related to oil spills, can be clearly captured by SAR sensors. Yet their discrimination from look-alikes poses a challenging objective. A variety of different methods have been proposed to automatically detect and classify these dark spots. The authors in [70] propose semantic segmentation with deep CNNs as an efficient approach. Moreover, a publicly available SAR image dataset is introduced, aiming to provide a benchmark for future oil spill detection methods. Similar approach is presented in [71] where a CNN architecture is used to detect oil spills within the Bohai Bay of China. Further research on this topic, that extremely affect the water quality, can be obtained from this review [88]. Finally, the authors in [48] review two DL frameworks that carry out ocean remote-sensing-image classifications. Both combine CNN layers and AEs among the main features. Moreover, several ocean applications are discussed, ranging from oil spill, coastal inundation, green algae, ship detection and coral reef mapping. This can be a good additional reference to the reader for further investigation in the aforementioned use cases.

D. SEA SURFACE TEMPERATURE

Changes in ocean temperature over time have important implications for marine ecosystems and global climate change. Sea surface temperature provides significant predictive information, supplying basis for revealing the spatial distribution of biological environmental factors, and as an indicator to monitor marine disasters. However, due to large variations in heat flux, radiation, and diurnal wind near the sea surface, its prediction has always been a highly uncertain issue [102]. In the following, we present some data-driven models tackling this. The temporal dependence of marine temperature variation at multiple depths is analyzed in [98], where the authors perform time-series prediction based on LSTM. The data used is from the Global Ocean Argo Grid Data Set, that provides annual, monthly and yearly average ocean temperature covering multiple seas, including the Coral Sea, the Equatorial Pacific Region, and the South China Sea, and salinity data. The proposed scheme is able to outperform other conventional benchmarks with a reduced input sequence length. The authors in [99] adopt an LSTM for predicting sea surface temperature over the China Seas for 12-month lead time. Considering the sub-regional feature differences within the study area, they use self-organizing feature maps to classify the data first, and then use the classification results as additional inputs for the DL network. Moreover, ensemble modeling is applied with nine selected models differing in structure and initial parameters to overcome the high variance in the output. Results show that introducing appropriate class labels as auxiliary information can improve the prediction accuracy, but however lacks the capability to predict extreme events. Therefore, as further research, specific area features including, the Indian Ocean Dipole and monsoons shall be considered.

Medium and long-term sea surface temperature prediction models are designed in [100] based on GRU from data collected in the Bohai Sea. This sea is characterized by a large annual temperature difference, and several time scales (monthly and quarterly) are used to verify the practicability and stability of the model. The dataset is from the Optimum Interpolation USA global grid, that combines EO from different platforms, including satellites, ships and buoys. Regardless of whether monthly or quarterly data, the proposed scheme outperforms the LSTM case in terms of stability and accuracy when the prediction horizon increases. A similar scheme is proposed in [101], where a GRU encoder-decoder is used to capture the dynamics on sea surface temperature over the Bohai and South China Seas. As in the previous work, the proposed model outperforms a fully connected LSTM network and SVR on different prediction scales (daily, weekly, and monthly), especially in long-scale and long-term predictions. In addition, relationships between historical and future data are explored, finding that each future daily mean sea surface temperature within the Bohai Sea most strongly correlated with past historical values. Another example of using GRU in combination with CNN can be found in [102]. A complex DL architecture is designed in [103] to predict sea surface temperature on multiple scales over data collected in Yellow and Bohai Seas in China. The model is compound by three blocks: (i) a wavelet transformation plus the addition of Gaussian noise to enhance the robustness of the model for data transformation; (ii) an LSTM and a convolutional layer for feature extraction; (iii) and a fully connected layer for the prediction output. MLP and LSTM, together with a suite of ML models, including linear regression and decision tree are tested in [104] to estimate sea surface temperatures. The authors in [105] develop a DL framework for sea surface temperature forecasting associated with a tropical instability wave within the eastern equatorial Pacific Ocean. A spatio-temporal model based on RNN for forecasting time series of spatial processes is proposed in [89]. The model learns these dependencies through a structured latent dynamical component, and a decoder predicts the observations from the latent representations. It is evaluated on a geo-spatial dataset from the Pacific sea surface temperature and with the goal to predict future temperatures at different spatial locations.

E. AQUACULTURE

Aquaculture is a complex and interactive process that depends on many factors such as water resource, animal, human as well as capital investment. There is a wide variety of aquatic organisms in aquaculture, including fish, shrimp, crab, scallop, coral, jellyfish, aquatic macro-invertebrate and phytoplankton. Good water quality is the essential existence condition for these organisms, affecting directly their growth and diet safety. The measured data influenced by complex environmental agents are usually nonlinear and various, which make it difficult to design accurately control systems. Traditionally, cameras, underwater robots and water quality sensors are mainly used to monitor the entire production process. However, the underwater image data are low quality due to luminosity change, turbidity, complex background and fast-moving aquatic animals.

In this regard, traditional ML has not entirely satisfied the actual requirements. A lot of research has been done in this topic, also focusing on the possibilities that DL can offer. Two main surveys about DL aquaculture are presented next. The authors in [12] classify DL aquaculture research including live fish identification, species classification, behavioral analysis, feeding decisions, size or biomass estimation, and water quality evaluation. Technical details, including data, algorithms, and performance, are also analyzed. Main conclusions are that DL offers important breakthroughs in the implementation of smart fish farming. In particular, the authors claim that DL is expected to expand into new application areas such as fish disease diagnosis; data will become increasingly important; and composite models considering spatio-temporal sequences will represent the main research direction. Another review is presented in [20]. In this case, classification is done by aquatic products, including fish, shrimp, scallop, coral, jellyfish, aquatic macro-invertebrates, phytoplankton and water quality. Regarding water quality estimation, they classify research into three groups based on the used water parameter, i.e., dissolved oxygen, Chl-a and sea temperature. Moreover, advantages and limitations of DL in aquaculture are discussed together with future research directions and challenges. We refer the reader to these two reviews for further details regarding different aquaculture applications where DL is utilized.

Since aquaculture in marine environments is always open to its surroundings, the changes in water quality parameters are usually nonlinear, dynamic, and complex. Due to this, traditional forecasting methods have lots of problems, such as low accuracy, poor generalization, and high-time complexity. Next, different cases where DL is proposed as a way to solve this complex problem are presented. An hybrid model based on Principal Component Analysis (PCA) and LSTM is proposed in [90] to forecast the dissolved oxygen content in an aquaculture pond located in China. The key impact factors are extracted by PCA, including water temperature, solar radiation, wind speed, wind direction, soil temperature and soil moisture. Experimental results show that the proposed scheme outperforms the evaluated benchmarks, i.e., a Backpropagation ANN (BPNN), PSO BPNN, ELM and least squares SVM. A similar approach is proposed in [91] where a kernel PCA and an RNN are used to forecast the trend of dissolved oxygen from data collected in Australia. Moreover, an ANN-based model is proposed in [109] to predict the spatio-temporal distribution of dissolved oxygen and hypoxia condition in Chesapeake Bay, the largest estuary in US. Different from other empirical models, the approach is able to simulate three-dimensional variations of water quality variables without in situ data, but only using external forcings including nutrient loading, river flow, air temperature, solar radiation, and wind speed, as model inputs. A sparse AE and an LSTM combined architecture is presented in [92] to predict the dissolved oxygen content in the following 3, 6, and 72 hours in their experimental case study. This paper verifies that using a sparse AE as a feature extraction pre-training network can enhance the prediction accuracy of LSTM. Moreover, the authors highlight possible improvements: (i) to explore a multivariate output that only trains the model once, so as to realize a better comprehensive prediction of the water body; and (ii) to analyze the effect of using a deeper sparse AE, that may provide better latent features.

A deep ANN is designed in [93] to forecast water quality based on the trend of dissolved oxygen from data collected in Baffle Creek, Australia. High accuracy performances are shown for predicting 90 and 120 min ahead of the last observed data in the wet season, overcoming results from traditional ML schemes. More papers dealing with dissolved oxygen prediction can be found in [12]. Another example of water quality prediction is presented in [94] for pH, water temperature and dissolved oxygen. The authors build a new raw dataset collected in time series from the marine aquaculture base in Xincun Town, China. Their framework is a deep Bi-directional Stacked Simple Recurrent Unit (SRU) learning network, able to achieve higher accuracy than the RNN and LSTM approaches. An LSTM-based model is used in [95] to forecast pH and water temperature from data collected in the mariculture base at Xincun Town, China by deploying sensor devices in a cage. Pearson's correlation coefficient is used to obtain the correlation priors between pH, water temperature, and other water quality parameters to be used in the prediction model. The authors in [96] introduce an architecture composed of an LSTM forecasting model and an IoT system to monitor real time salinity, temperature, pH, and dissolved oxygen water quality from different aquaculture ponds. Since the data is recorded daily, they can build sequential time series that are fed into the forecasting scheme. Results show that the approach could be applied in a real scenario, and provides valuable information regarding shrimp/fish raising. Moreover, the authors claim that further research on more sophisticated models should be taken into account, considering multivariate LSTM. An analogous concept is introduced in [97] where pH and water temperature are predicted. Again, the Pearson's correlation coefficient is employed to extract correlation among the water parameters. This information is used for the Simple Recurrent Unit RNN-based model they propose.

VI. OPEN CHALLENGES

In this section, we discuss the main outcomes found within the reviewed literature and state open challenges. We first present our concluding remarks on the new possibilities offered by DL in water quality assessment and forecasting for sustainable marine ecosystems and discuss the key open issues of this technology. Finally, we list a set of identified new research opportunities to encourage work on those directions.

A. CONCLUDING REMARKS AND LESSON LEARNED

It is undeniable that DL is a remarkable breakthrough and in combination with other technologies such as cloud computing, Big Data and IoT is revolutionizing many aspects of problem solving in different fields. Among the main advantages, we highlight the automatic feature extraction, in contrast with manually selecting features, which is a laborious, heuristic approach, and the final outcome is highly dependent on previous expertise. Moreover, it provides high accuracy and strong stability for target recognition in complex environments, with the ability of obtaining meaningful features automatically using a general-purpose learning procedure. Although DL requires in general more computing power and longer training time than traditional ML strategies, the trained DL models are highly efficient at performing test tasks [12]. This comes from one of the most significant drawbacks of DL, which is the need for large input data. Besides, DL schemes are excessively reliant on sample data and have low interpretability, that translates into gained experience only from a specific dataset. Moreover, when faced with unbalanced training data, most models will tend to ignore some important features.

B. FUTURE RESEARCH DIRECTIONS

1) DECISION MAKING & REINFORCEMENT LEARNING

The use of DL tools simultaneously expand the reach and speed of *decision-making*, while also stripping away layers of context that would possibly be relevant to a human decision-making process. Few DL works within marine environments and water quality include decision-making components, which presents an opportunity for a new research line [18]. The vast majority of reviewed papers in this survey focus on forecasting of certain water parameters given the problem suitability for DL. As future research, efforts between agencies, research institutions and regional governments can be promoted to develop better coordinated models that will yield actionable and reliable information, based even on real-time data, to constitute the future of decision-making systems [18]. The key point is not only providing the evaluation and forecasting of water quality for a certain coastal area or marine ecosystem, but also going beyond that by proposing significant actions based on these evaluations that can have a real impact in the system sustainability. In this way, we believe *reinforcement learning* in conjunction with deep architectures are the perfect match, inferring some intelligence to the system, empowering it with the ability to make early smart decisions and autonomous foresighted control. This is opening a research spot to be tackled. So far, not many efforts within marine environments dealing with this can be found.

2) DATA ACQUISITION & TRANSFER LEARNING

Regarding sensing technologies, the integration of IoT and WSNs together with DL opens new possibilities in the way to automatically collect data in large amounts concerning in situ sensing. Due to these technological advances, data is usually collected from different sources and combining several sensors, thus the integration of these data mainly in space and time is also an open challenge. This is not only affecting in situ measurements, but also remote sensing, and finally their own combination. Great efforts have to be done in this way to extract the optimal profit from the measured data. This is especially important in RS, where the introduction of DL arise many challenges to be addressed. First, limited datasets. Despite the technological advances that have facilitate the data collection in huge volumes, RS still suffers in general from a lack of labeled training data, which is in contrast with the DL prime requirement of large datasets. RS training data is expensive in terms of time, effort and investment, because as usual it requires some expert interpretation to label. The involved dataset sizes are not the only cause, but also the conceptual difficulty in labeling. Besides the challenge of working with a limited training dataset, problems are often under constrained, leading to the possibility of models thought to be of high quality, which perform well in training and even test datasets, but deviate strongly in real-world situations and data outside their valid domain [110].

In this regard, *Transfer learning* can be a potential solution. Conventional ML and DL algorithms, so far, have been traditionally designed to work in isolation. These algorithms are trained to solve specific tasks, thus the models have to be rebuilt from scratch once the feature-space distribution changes. Transfer learning is the idea of overcoming the isolated learning paradigm and utilizing knowledge acquired for one task to solve related ones. If the learning crosses domains, it may be possible to utilize lower to mid-level features learned from one domain into the other. Specially, it is a popular approach in DL due to the requirement of large datasets, and could solve the common problem in RS of limited available data. Some examples of transfer learning across domains within RS can be found in [21].

3) DATA MANAGEMENT

Having the opportunity to build on a widely accepted dataset would provide researchers the possibility to improve the accuracy of their models by taking advantage of previous works created using the same dataset. However, the lack of available open datasets makes difficult the comparison among new DL approaches. Again, efforts between the different marine environment stakeholders need to be done to build *open data* repositories. Moreover, while conventional DL computer vision applications deal with three-channels images, i.e. red, green and blue, as input data, satellite images extend to multiple spectral bands well beyond the visible range, which often induce different statistical properties to those of natural images. This includes spatial dependence and interdependence of variables, violating the important assumption of identically, independently distributed data. Additionally, and as we mentioned before, integrating multi-sensor data is not trivial since different sensors exhibit diverse imaging geometries, spatial and temporal resolution, physical meaning, content and statistics. Sequences of multi-sensor satellite observations also come with several noise sources, uncertainty levels, missing data and often systematic gaps, due to the presence of clouds or snow, distortions in acquisition, storage and transmission [110].

In addition, spectral, spatial and temporal dimensionality raise computational challenges. Data volume is increasing every day. Currently, the biggest meteorological agencies have to process terabytes per day in near real time, often at very high precision. Furthermore, a moderate-resolution (around 1 km) global field has sizes of approximately $40,000 \times 20,000$ pixels, i.e. three orders of magnitude more than in computer vision works [110]. Ocean color observation through RS rely on very complex and highly nonlinear models. If the input data is not adequately understood, the attained outputs can be very inaccurate. Besides, these applications involve scientific end users, who need to understand how the DL systems work. However, a DL system is usually a large and complex structure, seen as a black box that is hard to understand what is happening inside. *Visualization tools* capable of showing what the DL network is learning is an open area of research and would be a great benefit [21].

4) KNOWLEDGE FUSION

In addition to optimal prediction, and to achieve models that maximally learn from data, a more challenging task is taking into account physical and biological knowledge. One promising uncharted approach is the integration of physical modeling with DL, i.e. theory-driven with data-driven modeling. In fact, both are complementary, with physical approaches in principle being directly interpretable and offering the potential of extrapolation beyond observed conditions, whereas data-driven strategies are highly flexible in adapting to data and finding hidden patterns. More information can be found in [110]. This is the concept of *knowledge fusion*, where information discovered from different areas of expertise can complement and reinforce each other in order to derive more meaningful insights. DL techniques are considered as good candidates for this [111]. The authors in [112] elaborate on this paradigm, by studying the applicability and limitations of different knowledge fusion techniques. The task of data fusion is to identify the true values of data items among multiple observed values drawn from different sources of varying (and unknown) reliability. In this sense, a timely data fusion and analysis, to enable efficient, reliable, and accurate decision making and management of ubiquitous environments is a great challenge. The authors in [113] review the current research on data fusion for IoT with a particular focus on mathematical tools, such as AI, probabilistic methods and theory of belief. Opportunities and challenges are also discussed, including emerging areas that intrinsically benefit from data fusion as DL, autonomous vehicles and smart cities.

As we review in Section IV-C, several studies have shown strong potential for the application of EO together with RS for deriving water quality estimates over long temporal and spatial scales, but the reliable application of these methods is complicated by the diversity of water types, sensor configurations, and inherent limitations of the used approaches [5]. In this regard, the success of ocean color observation schemes can be limited by the effectiveness of the atmospheric correction applied to the data. Atmospheric correction over coastal waters often exhibits large inaccuracies. In addition to the low reflectance of water reached at the satellite sensor, less than 20% in certain spectrum regions, atmospheric correction in coastal areas is challenged by site-specific effects (e.g. specular reflection of sun-glint, land adjacency, very high turbidity, bottom reflectance) [1], [114]. Multiple sources of noise and the difficulty on their assessment have led some authors to choose partially corrected images, rather than full atmospheric corrected data [115]. This arises the importance of improving the understanding of the limitations posed by ineffective atmospheric correction. Several methods over coastal waters have been addressed, but there is no single model that works well under the different atmospheric, solar, geographic and water scenarios. Given the amount of involved data and its complex interactions, atmospheric correction over coastal waters is still an open issue to be tackled where DL can be a promising solution [5].

5) AQUACULTURE

The use of DL in aquaculture will continue to expand, including potential applications such as fish disease diagnosis, aquatic product quality control and traceability [12]. Moreover, aquaculture is expected to benefit greatly from deep reinforcement learning, mainly regarding unmanned underwater vehicles, aquatic robots (i.e. picking robot, conveyance robot, sorting robot) and water parameter optimization [20]. Currently, there are few articles dealing with it. In addition, smart aquaculture literature appear as a collection of loosely coupled applications with little relation among them. The current research is focused on applying DL to extract superficial knowledge in order to solve a specific problem. However, data from different applications may be correlated and affect each other. For instance, bad water quality conditions can affect fish diseases or behavior. Yet, such correlations have been mostly overlooked in the existing studies. In this sense, the aforementioned concept of knowledge fusion is also applicable here. Specifically, multi-modal deep learning [116], which learn features from multiple sources, is expected to provide great improvements within aquaculture.

To this respect, several aquaculture monitoring platforms have been introduced in the market, e.g., AquaConnect, AquaX Online, JellyX, AquaCloud.ai, providing advices for efficient management of aquacultures by processing mainly in situ data, and designed for specific purposes, e.g., to calculate the growth of shrimp and predict the most profitable harvest periods, or to estimate the abundance of jellyfish and classify their risk of proliferation. Moreover, many European funded research projects have been/are devoted to sustainable management of aquaculture sites based on sensors data. Examples are: NewTechAqua, centered around epidemiological models within aquacultures and Blue-Cloud, which combines distributed marine data resources, computing platforms, and analytical services to find suitable areas for aquacultures. Gain and iFishIENCi are, instead, more on improving the efficiency of aquaculture production by smart feeding algorithms. Finally, CERES investigates the influence of climate change into fish and shellfish resources using remote sensors. Setting the technological basis on these projects and products, a further research goal could be to study, not only the effects of climate change tackled in some of them, but also the intra- and extra-system phenomena in aquaculture to better understand under which conditions the marine environment is resilient and can coexist with environmental changes produced by aquacultures. Extra-system phenomena include exogenous events (e.g., storms), interaction with other systems (e.g., agricultural systems), extreme biological occurrences (e.g., phytoplankton blooms) or other anthropogenic externalities. Intra-system phenomena are those related to the activities within the aquaculture itself and are specially linked with its carrying capacity.

6) EDGE COMPUTING

Consistent and diverse data is the basics to build smart water quality applications. However, as the frequency and type of data resources expand, a centralized system to collect and analyze data simply may not be a viable approach. Furthermore, associated costs with the data transfer from distributed sensors are not irrelevant, putting away further investments in sensor coverage and data reporting interval [18]. As a solution to this, advances in distributed intelligence, whereby content, control, and computation are moved closer to end users/sensors, i.e., to the network edge, has led to the emergence of the *Edge Computing* paradigm, which allows network functions to be virtualized and deployed at the network edge [117]. Essentially, the physical proximity between the computing and information-generation sources promises several benefits compared to the traditional cloud-based computing paradigm, including low latency, energy efficiency, privacy protection, reduced bandwidth consumption, and context awareness [118].

The combination of AI in conjunction with edge computing arises a new concept, the *Edge Intelligence (EI)* [118], which aims to investigate distributed computing of ML models. A popular solution for distributed learning is represented by Federated Learning (FL) [119], which works by combining in a central server the training results of a shared model with certain Stochastic Gradient Descent (SGD) methods, such as the selective SGD [120]. A water quality application of this can be found in [121], where a sensor-based FL model is proposed to monitor algal blooms by using distributed observation data with geologically separated local models. A large IoT network is deployed within the Keum river, South Korea. The authors present a novel FL scheduling algorithm in order to fairly schedule data over edge servers for avoiding the overfitting problem. EI could encourage the ocean conservation and water quality community to innovate novel applications. This is an open research area to be tackled within sustainable marine ecosystems.

7) ANOMALY DETECTION

Anomaly detection is the identification of rare items, events or observations which raise suspicions by differing significantly from the majority of the data, i.e., that do not conform to an expected pattern. Generally, anomalies can be categorized into two types: (i) outliers, data points that differ greatly from others within the dataset; and (ii) anomaly patterns, a small fraction of data different from the majority, including particular fluctuations, shapes and trends, which always provide more information than single points [108]. AD has been traditionally implemented manually, relying mainly on visualization tools. However, marine systems are characterized by complex interactions among different dynamical and random processes, so conventional models cannot accurately identify these special patterns. Data-driven methods, specially DL architectures, can help in this way. We present some examples next.

A framework for automated AD in high-frequency water quality data from in situ sensors, using turbidity, conductivity and river-level data collected in the Great Barrier Reef is presented in [106]. As future work, the authors say it could be applied to other variables such as dissolved oxygen, water temperature and nitrates. These properties are highly dynamic in space and time and so differentiating anomalies shall be more challenging, but they are potential near-real time surrogates for sediment and nutrient concentrations, reducing laboratory analysis and improving the way water quality is assessed. The study in [107] explores the river-flow-induced impacts on the performance of water quality forecasting in the coastal waters in Hilo Bay, Pacific Ocean. For this purpose, hourly recorded water quality parameters of salinity, temperature and turbidity as well as the flow data of the Wailuku River were used as inputs for deep ANNs, ELM and SVR. Results show that river flow made the most and least improvement on the forecasting efficiency of turbidity and water temperature, respectively. The authors in [108] present a wavelet ANN for detecting anomalies in ocean fixed-point in situ time series from the National Ocean Test Site of China. Salinity and surface current speed are the considered water parameters under an unsupervised setting. Results show that the presented method is more tolerant to noise and more sensitive to anomalies with temporal dependencies. However, its spatial particularity and short period of data records may lead to misclassification.

To the best of our knowledge, we believe AD has not yet received the necessary focus within marine environments, and could provide powerful insights thanks to its ability to identify hidden patterns. In this regard, AD can offer prior information and be a useful tool to build early-warning systems with foresighted control that are vital to avoid ecological disasters and to maintain the desired system sustainability. Therefore, we believe it is still an open research area to be undertaken. In this sense, more complex and integrated autonomous systems shall be designed involving water quality assessment and forecasting, early warning through AD, and finally decision-making based on reinforcement learning, with the overall wrap of deep learning.

VII. CONCLUSION

An appropriate evaluation of water quality is key in order to guarantee sustainability to our oceans and coastal regions. Leveraging on a combination of cross-disciplinary technologies including RS, IoT, Big Data, cloud computing, and AI is essential to attain this aim. In this paper, we have conducted a review about methodologies and technologies for water quality monitoring that contribute to a sustainable management of marine environments. Specifically, we have focused on DL strategies for water quality estimation and forecasting. The analyzed works have been classified depending on the type of task, scenario and proposed DL architecture. Moreover, several applications including coastal management and aquaculture have been surveyed. Finally, we have discussed open issues still to be addressed and potential research lines where transfer learning, knowledge fusion, anomaly detection, reinforcement learning, edge computing and decision-making policies are expected to be the main involved agents.

REFERENCES

[1] G. Zheng and P. M. DiGiacomo, "Uncertainties and applications of satellite-derived coastal water quality products," *Prog. Oceanogr.*, vol. 159, pp. 45–72, Dec. 2017.

[2] C. A. Graves, A. Powell, M. Stone, F. Redfern, T. Biko, and M. Devlin, "Marine water quality of a densely populated Pacific atoll (Tarawa, Kiribati): Cumulative pressures and resulting impacts on ecosystem and human health," *Mar. Pollut. Bull.*, vol. 163, Feb. 2021, Art. no. 111951.

[3] M. Visbeck, "Ocean science research is key for a sustainable future," *Nature Commun.*, vol. 9, no. 1, pp. 1–4, Feb. 2018.

[4] M. Ramírez-Pérez, R. Gonçalves-Araujo, S. Wiegmann, E. Torrecilla, R. Bardaji, R. Röttgers, A. Bracher, and J. Piera, "Towards cost-effective operational monitoring systems for complex waters: Analyzing small-scale coastal processes with optical transmissometry," *PLoS ONE*, vol. 12, no. 1, Jan. 2017, Art. no. e0170706.

[5] E. Politi, S. K. Paterson, R. Scarrott, E. Tuohy, C. O'Mahony, and W. C. A. Cámaro-García, "Earth observation applications for coastal sustainability: Potential and challenges for implementation," *Anthropocene Coasts*, vol. 2, no. 1, pp. 306–329, Jan. 2019.

[6] S. A. Wooldridge and T. J. Done, "Improved water quality can ameliorate effects of climate change on corals," *Ecol. Appl.*, vol. 19, no. 6, pp. 1492–1499, Sep. 2009.

[7] J. Rizzi, S. Torresan, A. Critto, A. Zabeo, D. Brigolin, S. Carniel, R. Pastres, and A. Marcomini, "Climate change impacts on marine water quality: The case study of the Northern Adriatic sea," *Mar. Pollut. Bull.*, vol. 102, no. 2, pp. 271–282, Jan. 2016.

[8] T. Wernberg, B. D. Russell, P. J. Moore, S. D. Ling, D. A. Smale, A. Campbell, M. A. Coleman, P. D. Steinberg, G. A. Kendrick, and S. D. Connell, "Impacts of climate change in a global hotspot for temperate marine biodiversity and ocean warming," *J. Experim. Mar. Biol. Ecol.*, vol. 400, nos. 1–2, pp. 7–16, Apr. 2011.

[9] T. O'Shea, R. Jones, A. Markham, E. Norell, J. Scott, S. Theuerkauf, and T. Waters, "Towards a blue revolution: Catalyzing private investment in sustainable aquaculture production systems," Nature Encourage Capital, Arlington, VA, USA, Tech. Rep. 1, May 2019.

[10] *The EU Blue Economy Report*, Eur. Commission, Brussels, Belgium, 2020.

[11] L. García, D. Rodríguez, M. Wijnen, and I. Pakulski, *Earth Observation for Water Resources Management: Current Use and Future Opportunities for the Water Sector*. Washington, DC, USA: The World Bank, 2016.

[12] X. Yang, S. Zhang, J. Liu, Q. Gao, S. Dong, and C. Zhou, "Deep learning for smart fish farming: Applications, opportunities and challenges," *Rev. Aquaculture*, vol. 13, no. 1, pp. 66–90, Jan. 2021.

[13] B. Ferreira, M. Iten, and R. G. Silva, "Monitoring sustainable development by means of earth observation data and machine learning: A review," *Environ. Sci. Eur.*, vol. 32, no. 1, pp. 1–17, Sep. 2020.

[14] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep Learning*, vol. 1, no. 2. Cambridge, MA, USA: MIT Press, 2016.

[15] L. Ma, Y. Liu, X. Zhang, Y. Ye, G. Yin, and B. A. Johnson, "Deep learning in remote sensing applications: A meta-analysis and review," *ISPRS J. Photogramm. Remote Sens.*, vol. 152, pp. 166–177, Jun. 2019.

[16] X. X. Zhu, D. Tuia, L. Mou, G.-S. Xia, L. Zhang, F. Xu, and F. Fraundorfer, "Deep learning in remote sensing: A comprehensive review and list of resources," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 8–36, Dec. 2017.

[17] Y. Chen, L. Song, Y. Liu, L. Yang, and D. Li, "A review of the artificial neural network models for water quality prediction," *Appl. Sci.*, vol. 10, no. 17, p. 5776, Aug. 2020.

[18] M. Sit, B. Z. Demiray, Z. Xiang, G. J. Ewing, Y. Sermet, and I. Demir, "A comprehensive review of deep learning applications in hydrology and water resources," *Water Sci. Technol.*, vol. 82, no. 12, pp. 2635–2670, Aug. 2020.

[19] C. Shen, "A transdisciplinary review of deep learning research and its relevance for water resources scientists," *Water Resour. Res.*, vol. 54, no. 11, pp. 8558–8593, Nov. 2018.

[20] M. Sun, X. Yang, and Y. Xie, "Deep learning in aquaculture: A review," *J. Comput.*, vol. 31, no. 1, pp. 294–319, Jan. 2020.

[21] J. E. Ball, D. T. Anderson, and C. S. Chan, "Comprehensive survey of deep learning in remote sensing: Theories, tools, and challenges for the community," *J. Appl. Remote Sens.*, vol. 11, no. 4, Sep. 2017, Art. no. 042609.

[22] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.

[23] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, 2018.

[24] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," Jun. 2014, *arXiv:1406.2661*. [Online]. Available: http://arxiv.org/abs/1406.2661

[25] O. Chapelle, B. Scholkopf, and A. Zien, *Semi-Supervised Learning*. Cambridge, MA, USA: MIT Press, 2006.

[26] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, Oct. 1986.

[27] T. Mitchell, *Machine Learning*. New York, NY, USA: McGraw-Hill, 1997.

[28] H. D. Trinh, A. F. Gambin, L. Giupponi, M. Rossi, and P. Dini, "Mobile traffic classification through physical control channel fingerprinting: A deep learning approach," *IEEE Trans. Netw. Service Manage.*, vol. 18, no. 2, pp. 1946–1961, Jun. 2021.

[29] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[30] R. Fu, Z. Zhang, and L. Li, "Using LSTM and GRU neural network methods for traffic flow prediction," in *Proc. 31st Youth Acad. Annu. Conf. Chin. Assoc. Automat. (YAC)*. Wuhan, China: IEEE, Nov. 2016, pp. 324–328.

[31] K. Cho, B. van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," 2014, *arXiv:1406.1078*. [Online]. Available: http://arxiv.org/abs/1406.1078

[32] Q. Yuan, H. Shen, T. Li, Z. Li, S. Li, Y. Jiang, H. Xu, W. Tan, Q. Yang, J. Wang, J. Gao, and L. Zhang, "Deep learning in environmental remote sensing: Achievements and challenges," *Remote Sens. Environ.*, vol. 241, May 2020, Art. no. 111716.

[33] Y. Pu, Z. Gan, R. Henao, X. Yuan, C. Li, A. Stevens, and L. Carin, "Variational autoencoder for deep learning of images, labels and captions," 2016, *arXiv:1609.08976*. [Online]. Available: http://arxiv.org/abs/1609.08976

[34] Y.-A. Chung, C.-C. Wu, C.-H. Shen, H.-Y. Lee, and L.-S. Lee, "Audio Word2Vec: Unsupervised learning of audio segment representations using sequence-to-sequence autoencoder," 2016, *arXiv:1603.00982*. [Online]. Available: http://arxiv.org/abs/1603.00982

[35] K. Wang, C. Gou, Y. Duan, Y. Lin, X. Zheng, and F.-Y. Wang, "Generative adversarial networks: Introduction and outlook," *IEEE/CAA J. Autom. Sinica*, vol. 4, no. 4, pp. 588–598, Sep. 2017.

[36] M. H. Gholizadeh, A. M. Melesse, and L. Reddi, "A comprehensive review on water quality parameters estimation using remote sensing techniques," *Sensors*, vol. 16, no. 8, p. 1298, Aug. 2016.

[37] P. R. Hill, A. Kumar, M. Temimi, and D. R. Bull, "HABNet: Machine learning, remote sensing-based detection of harmful algal blooms," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 3229–3239, 2020.

[38] E. V. Hatzikos, G. Tsoumakas, G. Tzanis, N. Bassiliades, and I. Vlahavas, "An empirical study on sea water quality prediction," *Knowl.-Based Syst.*, vol. 21, no. 6, pp. 471–478, Aug. 2008.

[39] J. Park, K. T. Kim, and W. H. Lee, "Recent advances in information and communications technology (ICT) and sensor technology for monitoring water quality," *Water*, vol. 12, no. 2, p. 510, Feb. 2020.

[40] C. Z. Myint, L. Gopal, and Y. L. Aung, "Reconfigurable smart water quality monitoring system in IoT environment," in *Proc. IEEE/ACIS 16th Int. Conf. Comput. Inf. Sci. (ICIS)*. Wuhan, China: IEEE, May 2017, pp. 435–440.

[41] C. Encinas, E. Ruiz, J. Cortez, and A. Espinoza, "Design and implementation of a distributed IoT system for the monitoring of water quality in aquaculture," in *Proc. IEEE Int. Conf. Wireless Telecommun. Symp. (WTS)*. Chicago, IL, USA: IEEE, Apr. 2017, pp. 1–7.

[42] J. Huan, H. Li, F. Wu, and W. Cao, "Design of water quality monitoring system for aquaculture ponds based on NB-IoT," *Aquacultural Eng.*, vol. 90, Aug. 2020, Art. no. 102088.

[43] S. Balakrishnan, S. S. Rani, and K. C. Ramya, "Design and development of IoT based smart aquaculture system in a cloud environment," *Int. J. Oceans Oceanogr.*, vol. 13, no. 1, pp. 121–127, Oct. 2019.

[44] N. Vijayakumar and R. Ramya, "The real time monitoring of water quality in IoT environment," in *Proc. Int. Conf. Innov. Inf. Embedded Commun. Syst. (ICIIECS)*. Coimbatore, India: IEEE, Mar. 2015, pp. 1–5.

[45] A. N. Prasad, K. A. Mamun, F. R. Islam, and H. Haqva, "Smart water quality monitoring system," in *Proc. 2nd Asia–Pacific World Congr. Comput. Sci. Eng. (APWC CSE)*. Nadi, Fiji: IEEE, Dec. 2015, pp. 1–6.

[46] P. J. Bresnahan, T. Wirth, T. Martz, K. Shipley, V. Rowley, C. Anderson, and T. Grimm, "Equipping smart coasts with marine water quality IoT sensors," *Results Eng.*, vol. 5, Mar. 2020, Art. no. 100087.

[47] V. V. Daigavane and M. Gaikwad, "Water quality monitoring system based on IoT," *Dept. Electron. Telecommun. Eng., Adv. Wireless Mobile Commun.*, vol. 10, no. 5, pp. 1107–1116, Dec. 2017.

[48] X. Li, B. Liu, G. Zheng, Y. Ren, S. Zhang, Y. Liu, L. Gao, Y. Liu, B. Zhang, and F. Wang, "Deep-learning-based information mining from ocean remote-sensing imagery," *Nat. Sci. Rev.*, vol. 7, no. 10, pp. 1584–1605, Oct. 2020.

[49] H. M. Dierssen and K. Randolph, "Remote sensing of ocean color," *Earth Syst. Monitor.*, vol. 2013, pp. 439–472, Dec. 2013.

[50] H. R. Gordon and A. Y. Morel, *Remote Assessment of Ocean Color for Interpretation of Satellite Visible Imagery: A Review*, vol. 4. Luxemburg: Springer, 2012.

[51] V. Krasnopolsky, S. Nadiga, A. Mehra, E. Bayler, and D. Behringer, "Neural networks technique for filling gaps in satellite measurements: Application to ocean color observations," *Comput. Intell. Neurosci.*, vol. 2016, pp. 1–9, Dec. 2016.

[52] S. Nadiga, V. Krasnopol'sky, E. Bayler, H.-C. Kim, A. Mehra, and D. W. Behringer, "Neural network technique for gap-filling satellite ocean color observations," U.S. Dept. Commerce Nat. Ocean. Atmos. Admin., Washington, DC, USA, Tech. Rep. 483, Aug. 2016.

[53] Z. Shao, J. Cai, P. Fu, L. Hu, and T. Liu, "Deep learning-based fusion of Landsat-8 and Sentinel-2 images for a harmonized surface reflectance product," *Remote Sens. Environ.*, vol. 235, Dec. 2019, Art. no. 111425.

[54] F. Vieira, G. Cavalcante, E. Campos, and F. Taveira-Pinto, "A methodology for data gap filling in wave records using artificial neural networks," *Appl. Ocean Res.*, vol. 98, May 2020, Art. no. 102109.

[55] A. Tanaka, M. Kishino, T. Oishi, R. Doerffer, and H. Schiller, "Application of neural network method to case ii water," in *Remote Sensing of the Ocean and Sea Ice 2000*, vol. 4172. Barcelona, Spain: International Society for Optics and Photonics, Dec. 2000, pp. 144–152.

[56] J. F. Mas and J. J. Flores, "The application of artificial neural networks to the analysis of remotely sensed data," *Int. J. Remote Sens.*, vol. 29, no. 3, pp. 617–663, Dec. 2007.

[57] L. Mukherjee, P.-W. Zhai, M. Gao, Y. Hu, B. A. Franz, and P. J. Werdell, "Neural network reflectance prediction model for both open ocean and coastal waters," *Remote Sens.*, vol. 12, no. 9, p. 1421, Apr. 2020.

[58] M. Hieronymi, D. Müller, and R. Doerffer, "The OLCI neural network swarm (ONNS): A bio-geo-optical algorithm for open ocean and coastal waters," *Front. Mar. Sci.*, vol. 4, p. 140, May 2017.

[59] D. H. Svendsen, P. Morales-Álvarez, A. B. Ruescas, R. Molina, and G. Camps-Valls, "Deep Gaussian processes for biogeophysical parameter retrieval and model inversion," *ISPRS J. Photogramm. Remote Sens.*, vol. 166, pp. 68–81, Aug. 2020.

[60] F. S. Marzano, M. Iacobelli, M. Orlandi, and D. Cimini, "Coastal water remote sensing from Sentinel-2 satellite data using physical, statistical, and neural network retrieval approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 2, pp. 915–928, Feb. 2021.

[61] S. Hafeez, M. Wong, H. Ho, M. Nazeer, J. Nichol, S. Abbas, D. Tang, K. Lee, and L. Pun, "Comparison of machine learning algorithms for retrieval of water quality indicators in case-II waters: A case study of Hong Kong," *Remote Sens.*, vol. 11, no. 6, p. 617, Mar. 2019.

[62] Y. H. Kim, J. Im, H. K. Ha, J.-K. Choi, and S. Ha, "Machine learning approaches to coastal water quality monitoring using GOCI satellite data," *GISci. Remote Sens.*, vol. 51, no. 2, pp. 158–174, Apr. 2014.

[63] N. Pahlevan, B. Smith, J. Schalles, C. Binding, Z. Cao, R. Ma, K. Alikas, K. Kangro, D. Gurlin, N. Hà, B. Matsushita, W. Moses, S. Greb, M. K. Lehmann, M. Ondrusek, N. Oppelt, and R. Stumpf, "Seamless retrievals of chlorophyll—A from Sentinel-2 (MSI) and Sentinel-3 (OLCI) in inland and coastal waters: A machine-learning approach," *Remote Sens. Environ.*, vol. 240, no. 1, Apr. 2020, Art. no. 111604.

[64] Y. Kown, S. Baek, Y. Lim, J. Pyo, M. Ligaray, Y. Park, and K. Cho, "Monitoring coastal chlorophyll—A concentrations in coastal areas using machine learning models," *Water*, vol. 10, no. 8, p. 1020, Aug. 2018.

[65] S. Wu, Z. Du, Y. Wang, T. Lin, F. Zhang, and R. Liu, "Modeling spatially anisotropic nonstationary processes in coastal environments based on a directional geographically neural network weighted regression," *Sci. Total Environ.*, vol. 709, Mar. 2020, Art. no. 136097.

[66] K. A. Islam, D. Pérez, V. Hill, B. Schaeffer, R. Zimmerman, and J. Li, "Seagrass detection in coastal water through deep capsule networks," in *Proc. Chin. Conf. Pattern Recognit. Comput. Vis. (PRCV)*. Nanjing, China: Springer, Nov. 2018, pp. 320–331.

[67] J. T. Ridge, P. C. Gray, A. E. Windle, and D. W. Johnston, "Deep learning for coastal resource conservation: Automating detection of shellfish reefs," *Remote Sens. Ecol. Conservation*, vol. 6, no. 4, pp. 431–440, Feb. 2020.

[68] A. Mahmood, M. Bennamoun, S. An, F. Sohel, F. Boussaid, R. Hovey, G. Kendrick, and R. B. Fisher, "Automatic annotation of coral reefs using deep learning," in *Proc. OCEANS MTS/IEEE Monterey*. Monterey, CA, USA: IEEE, Sep. 2016, pp. 1–5.

[69] A. Raphael, Z. Dubinsky, D. Iluz, J. I. C. Benichou, and N. S. Netanyahu, "Deep neural network recognition of shallow water corals in the Gulf of Eilat (Aqaba)," *Sci. Rep.*, vol. 10, no. 1, pp. 1–11, Jul. 2020.

[70] M. Krestenitis, G. Orfanidis, K. Ioannidis, K. Avgerinakis, S. Vrochidis, and I. Kompatsiaris, "Oil spill identification from satellite images using deep neural networks," *Remote Sens.*, vol. 11, no. 15, p. 1762, Jul. 2019.

[71] K. Zeng and Y. Wang, "A deep convolutional neural network for oil spill detection from spaceborne SAR images," *Remote Sens.*, vol. 12, no. 6, p. 1015, Mar. 2020.

[72] P. J. Werdell, L. I. W. McKinna, E. Boss, S. G. Ackleson, S E. Craig, W. W. Gregg, Z. Lee, S. Maritorena, C. S. Roesler, C. S. Rousseaux, and D. Stramski, "An overview of approaches and challenges for retrieving marine inherent optical properties from ocean color remote sensing," *Progr. Oceanogr.*, vol. 160, pp. 186–212, Jan. 2018.

[73] J. Park, J.-H. Kim, H.-C. Kim, B.-K. Kim, D. Bae, Y.-H. Jo, N. Jo, and S. H. Lee, "Reconstruction of ocean color data using machine learning techniques in polar regions: Focusing on Off Cape Hallett, Ross Sea," *Remote Sens.*, vol. 11, no. 11, p. 1366, Jun. 2019.

[74] J. Shin, S. M. Kim, Y. B. Son, K. Kim, and J.-H. Ryu, "Early prediction of *Margalefidinium polykrikoides* bloom using a LSTM neural network model in the South Sea of Korea," *J. Coastal Res.*, vol. 90, no. SI, pp. 236–242, Sep. 2019.

[75] A. T. Yñiguez and Z. J. Ottong, "Predicting fish kills and toxic blooms in an intensive mariculture site in the Philippines using a machine learning model," *Sci. Total Environ.*, vol. 707, Mar. 2020, Art. no. 136173.

[76] F. Zhang, Y. Wang, M. Cao, X. Sun, Z. Du, R. Liu, and X. Ye, "Deep-learning-based approach for prediction of algal blooms," *Sustainability*, vol. 8, no. 10, p. 1060, Oct. 2016.

[77] M. Qin, Z. Li, and Z. Du, "Red tide time series forecasting by combining ARIMA and deep belief network," *Knowl.-Based Syst.*, vol. 125, pp. 39–52, Jul. 2017.

[78] N. Muttil and K.-W. Chau, "Machine-learning paradigms for selecting ecologically significant input variables," *Eng. Appl. Artif. Intell.*, vol. 20, no. 6, pp. 735–744, Sep. 2007.

[79] T. Deng, K.-W. Chau, and H.-F. Duan, "Machine learning based marine water quality prediction for coastal hydro-environment management," *J. Environ. Manage.*, vol. 284, Apr. 2021, Art. no. 112051.

[80] I. Grasso, S. D. Archer, C. Burnell, B. Tupper, C. Rauschenberg, K. Kanwit, and N. R. Record, "The hunt for red tides: Deep learning algorithm forecasts shellfish toxicity at site scales in coastal Maine," *Ecosphere*, vol. 10, no. 12, Dec. 2019, Art. no. e02960.

[81] M. R. Al Shehhi and A. Kaya, "Time series and machine learning to forecast the water quality from satellite data," 2020, *arXiv:2003.11923*. [Online]. Available: http://arxiv.org/abs/2003.11923

[82] S. Shamshirband, E. J. Nodoushan, J. E. Adolf, A. A. Manaf, A. Mosavi, and K.-W. Chau, "Ensemble models with uncertainty analysis for multi-day ahead forecasting of chlorophyll a concentration in coastal waters," *Eng. Appl. Comput. Fluid Mech.*, vol. 13, no. 1, pp. 91–101, Dec. 2018.

[83] T. Rajaee and A. Boroumand, "Forecasting of chlorophyll—A concentrations in South San Francisco Bay using five different models," *Appl. Ocean Res.*, vol. 53, pp. 208–217, Oct. 2015.
[84] S. Sabour, N. Frosst, and G. E Hinton, "Dynamic routing between capsules," Nov. 2017, *arXiv:1710.09829*. [Online]. Available: http://arxiv.org/abs/1710.09829
[85] J. Rajasegaran, V. Jayasundara, S. Jayasekara, H. Jayasekara, S. Seneviratne, and R. Rodrigo, "DeepCaps: Going deeper with capsule networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Long Beach, CA, USA, Jun. 2019, pp. 10725–10733.
[86] M. Moniruzzaman, S. M. S. Islam, P. Lavery, and M. Bennamoun, "Faster R-CNN based deep learning for seagrass detection from underwater digital images," in *Proc. Digit. Image Comput., Techn. Appl. (DICTA)*. Perth, WA, Australia: IEEE, Dec. 2019, pp. 1–7.
[87] R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Santiago, Chile, Sep. 2015, pp. 1440–1448.
[88] S. T. Yekeen and A.-L. Balogun, "Advances in remote sensing technology, machine learning and deep learning for marine oil spill detection, prediction and vulnerability assessment," *Remote Sens.*, vol. 12, no. 20, p. 3416, Oct. 2020.
[89] A. Ziat, E. Delasalles, L. Denoyer, and P. Gallinari, "Spatio-temporal neural networks for space-time series forecasting and relations discovery," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*. New Orleans, LA, USA: IEEE, Nov. 2017, pp. 705–714.
[90] Y. Chen, Q. Cheng, and X. Fang, "Principal component analysis and long short-term memory neural network for predicting dissolved oxygen in water for aquaculture," *Trans. Chin. Soc. Agricult. Eng.*, vol. 34, no. 17, pp. 183–191, Sep. 2018.
[91] Y.-F. Zhang, P. Fitch, and P. J. Thorburn, "Predicting the trend of dissolved oxygen based on the kPCA-RNN model," *Water*, vol. 12, no. 2, p. 585, Feb. 2020.
[92] Z. Li, F. Peng, B. Niu, G. Li, J. Wu, and Z. Miao, "Water quality prediction model combining sparse auto-encoder and LSTM network," *IFAC-PapersOnLine*, vol. 51, no. 17, pp. 831–836, 2018.
[93] Y. Zhang, P. Fitch, M. P. Vilas, and P. J. Thorburn, "Applying multilayer artificial neural network and mutual information to the prediction of trends in dissolved oxygen," *Frontiers Environ. Sci.*, vol. 7, p. 46, Apr. 2019.
[94] J. Liu, C. Yu, Z. Hu, Y. Zhao, Y. Bai, M. Xie, and J. Luo, "Accurate prediction scheme of water quality in smart mariculture with deep Bi-S-SRU learning network," *IEEE Access*, vol. 8, pp. 24784–24798, 2020.
[95] Z. Hu, Y. Zhang, Y. Zhao, M. Xie, J. Zhong, Z. Tu, and J. Liu, "A water quality prediction method based on the deep LSTM network considering correlation in smart mariculture," *Sensors*, vol. 19, no. 6, p. 1420, Mar. 2019.
[96] N. Thai-Nghe, N. Thanh-Hai, and N. Chi, "Deep learning approach for forecasting water quality in IoT systems," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 8, pp. 686–693, 2020.
[97] J. Liu, C. Yu, Z. Hu, Y. Zhao, X. Xia, Z. Tu, and R. Li, "Automatic and accurate prediction of key water quality parameters based on SRU deep learning in mariculture," in *Proc. IEEE Int. Conf. Adv. Manuf. (ICAM)*. Yunlin, Taiwan: IEEE, Nov. 2018, pp. 437–440.
[98] J. Liu, T. Zhang, G. Han, and Y. Gou, "TD-LSTM: Temporal dependence-based LSTM networks for marine temperature prediction," *Sensors*, vol. 18, no. 11, p. 3797, Nov. 2018.
[99] L. Wei, L. Guan, L. Qu, and D. Guo, "Prediction of sea surface temperature in the China seas based on long short-term memory neural networks," *Remote Sens.*, vol. 12, no. 17, p. 2697, Aug. 2020.
[100] Z. Zhang, X. Pan, T. Jiang, B. Sui, C. Liu, and W. Sun, "Monthly and quarterly sea surface temperature prediction based on gated recurrent unit neural network," *J. Mar. Sci. Eng.*, vol. 8, no. 4, p. 249, Apr. 2020.
[101] J. Xie, J. Zhang, J. Yu, and L. Xu, "An adaptive scale sea surface temperature predicting method based on deep learning with attention mechanism," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 5, pp. 740–744, May 2019.
[102] X. Yu, S. Shi, L. Xu, Y. Liu, Q. Miao, and M. Sun, "A novel method for sea surface temperature prediction based on deep learning," *Math. Problems Eng.*, vol. 2020, pp. 1–9, May 2020.
[103] L. Xu, Y. Li, J. Yu, Q. Li, and S. Shi, "Prediction of sea surface temperature using a multiscale deep combination neural network," *Remote Sens. Lett.*, vol. 11, no. 7, pp. 611–619, May 2020.
[104] S. Wolff, F. O'Donncha, and B. Chen, "Statistical and machine learning ensemble modelling to forecast sea surface temperature," *J. Mar. Syst.*, vol. 208, Aug. 2020, Art. no. 103347.
[105] G. Zheng, X. Li, R.-H. Zhang, and B. Liu, "Purely satellite data–driven deep learning forecast of complicated tropical instability waves," *Sci. Adv.*, vol. 6, no. 29, Jul. 2020, Art. no. eaba1482.
[106] C. Leigh, O. Alsibai, R. J. Hyndman, S. Kandanaarachchi, O. C. King, J. M. McGree, C. Neelamraju, J. Strauss, P. D. Talagala, R. D. Turner, and K. Mengersen, "A framework for automated anomaly detection in high frequency water-quality data from *in situ* sensors," *Sci. Total Environ.*, vol. 664, pp. 885–898, May 2019.
[107] M. J. Alizadeh, M. R. Kavianpour, M. Danesh, J. Adolf, S. Shamshirband, and K.-W. Chau, "Effect of river flow on the quality of estuarine and coastal waters using machine learning models," *Eng. Appl. Comput. Fluid Mech.*, vol. 12, no. 1, pp. 810–823, Oct. 2018.
[108] Y. Wang, L. Han, W. Liu, S. Yang, and Y. Gao, "Study on wavelet neural network based anomaly detection in ocean observing data series," *Ocean Eng.*, vol. 186, Aug. 2019, Art. no. 106129.
[109] X. Yu, J. Shen, and J. Du, "A machine-learning-based model for water quality in coastal waters, taking dissolved oxygen and hypoxia in chesapeake bay as an example," *Water Resour. Res.*, vol. 56, no. 9, Sep. 2020, Art. no. e2020WR027227.
[110] M. Reichstein, G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, N. Carvalhais, and M. Prabhat, "Deep learning and process understanding for data-driven earth system science," *Nature*, vol. 566, no. 7743, pp. 195–204, Feb. 2019.
[111] Q. Chen, W. Wang, F. Wu, S. De, R. Wang, B. Zhang, and X. Huang, "A survey on an emerging area: Deep learning for smart city data," *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 3, no. 5, pp. 392–410, Oct. 2019.
[112] X. L. Dong, E. Gabrilovich, G. Heitz, W. Horn, K. Murphy, S. Sun, and W. Zhang, "From data fusion to knowledge fusion," 2015, *arXiv:1503.00302*. [Online]. Available: http://arxiv.org/abs/1503.00302
[113] F. Alam, R. Mehmood, I. Katib, N. N. Albogami, and A. Albeshri, "Data fusion and IoT for smart ubiquitous environments: A survey," *IEEE Access*, vol. 5, pp. 9533–9554, 2017.
[114] M. W. Matthews, "A current review of empirical procedures of remote sensing in inland and near-coastal transitional waters," *Int. J. Remote Sens.*, vol. 32, no. 21, pp. 6855–6899, Aug. 2011.
[115] A. Soriano-González and D. A. Fernández-Tejedor, "First results of phytoplankton spatial dynamics in two NW-mediterranean bays from chlorophyll—*A* estimates using sentinel 2: Potential implications for aquaculture," *Remote Sens.*, vol. 11, no. 15, p. 1756, Jul. 2019.
[116] J. Ngiam, A. Khosla, M. Kim, J. Nam, H. Lee, and A. Y. Ng, "Multimodal deep learning," in *Proc. ICML*, Bellevue, WA, USA, Jun. 2011, pp. 1–8.
[117] W. Shi and S. Dustdar, "The promise of edge computing," *Computer*, vol. 49, no. 5, pp. 78–81, May 2016.
[118] Z. Zhou, X. Chen, E. Li, L. Zeng, K. Luo, and J. Zhang, "Edge intelligence: Paving the last mile of artificial intelligence with edge computing," *Proc. IEEE*, vol. 107, no. 8, pp. 1738–1762, Aug. 2019.
[119] Q. Chen, Z. Zheng, C. Hu, D. Wang, and F. Liu, "Data-driven task allocation for multi-task transfer learning on the edge," in *Proc. IEEE 39th Int. Conf. Distrib. Comput. Syst. (ICDCS)*. Dallas, TX, USA: IEEE, Jul. 2019, pp. 1040–1050.
[120] R. Shokri and V. Shmatikov, "Privacy-preserving deep learning," in *Proc. 22nd ACM SIGSAC Conf. Comput. Commun. Secur.*, Denver, CO, USA, Oct. 2015, pp. 1310–1321.
[121] S. Park, S. Jung, H. Lee, J. Kim, and J.-H. Kim, "Large-scale water quality prediction using federated sensing and learning: A case study with real-world sensing big-data," *Sensors*, vol. 21, no. 4, p. 1462, Feb. 2021.

ÁNGEL FERNÁNDEZ GAMBÍN received the B.Sc. and M.Sc. degrees in telecommunication engineering from the Technical University of Cartagena, Spain, in 2014 and 2016, respectively, and the Ph.D. degree in information engineering from the SIGnal processing and NETworking (SIGNET) Research Group, University of Padua, Italy, in 2020. He was involved as an Early-Stage Researcher with the European H2020-MSCA-ITN SCAVENGE Project. Currently, he is working as a Postdoctoral Researcher at the Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), Barcelona, Spain, working on AI for ocean conservation. His research interests include AI, ocean conservation, sustainable mobile networks, and renewable energies.

EDUARD ANGELATS received the M.Sc. degree in telecommunications engineering from the Technical University of Catalonia (UPC), in 2008. In 2009, he joined as a Research Assistant with the former Institute of Geomatics, where he was involved in the development of innovative trajectory determination systems based on robust and hybrid navigation and focused on alternative applications. Since January 2014, he has been with CTTC, where he is currently a Researcher of the GEON Department as part of the Geomatics Division. In the last years, he has investigated the potential of using multispectral imagery, from RPAS or Sentinel-2, together with advanced processing techniques for the detection of archaeological non-superficial remains and water quality control for aquaculture applications.

JESÚS SORIANO GONZÁLEZ received the B.Sc. degree in marine sciences from the Catholic University of Valencia, Spain, in 2014, the M.Sc. degree in environmental and territorial management techniques from the University of Valencia, Spain, in 2015, and the M.Sc. degree in aquaculture from the University of Barcelona, Barcelona, Spain, in 2017. Currently, he is an Early Stage Researcher at the Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), Barcelona, working on satellite remote sensing for coastal monitoring.

MARCO MIOZZO received the M.Sc. degree in telecommunication engineering from the University of Ferrara, Italy, in 2005, and the Ph.D. degree from the Technical University of Catalonia (UPC), in 2018. After graduated, he worked as a Research Engineer in wireless networking at Consorzio Ferrara Ricerche (CFR), Italy, where he collaborated with the Department of Information Engineering (DEI), University of Padua, Italy. In June 2008, he joined the Centre Tecnològic de Telecomunicacions de Catalunya (CTTC). Currently, he is collaborating with the EU funded H2020 GREENEDGE (MSCA ETN). He participated in several research and development projects, among them 5G-Crosshaul, Flex5Gware, and SANSA, working on environmental sustainable mobile networks with energy harvesting capabilities through learning techniques. His main research interests include sustainable mobile networks, green wireless networking, energy harvesting, multi-agent systems, machine learning, green AI, energy ethical consumerism, transparent, and explainable AI.

PAOLO DINI received the M.Sc. and Ph.D. degrees from the Universita di Roma La Sapienza, in 2001 and 2005, respectively. He served as a Postdoctoral Researcher at the Research Centre on Software Technology (RCOST), Universita del Sannio, and a Contracted Professor at the Universita di Roma La Sapienza, in 2005. He is currently with the Centre Tecnològic de Telecomunicacions de Catalunya (CTTC) as a Senior Researcher. He is also a Contracted Professor with the Universitat Oberta de Catalunya (UOC). He has been the Project Coordinator of the EU-H2020-MSCA-SCAVENGE, which has been nominated as an example of ITN Best Practices, in 2018. He is the author of around 90 publications in scientific journals and conferences. His research interests include sustainable networking and computing, distributed optimization and optimal control, machine learning, and data analytics. He received two awards from Cisco Silicon Valley Foundation for his research on heterogeneous mobile networks, in 2008 and 2011, respectively.