

RESEARCH ARTICLE

Genetic algorithm-based deep reinforcement learning model for estimating chlorophyll-a concentration from remote sensing data

Xinmei Lan^{1,2,*}, Yang Ye³, Jianfeng An⁴, Xinxin Fang^{1,2}, Yuanda Liu^{1,2}, Guokun Xue³

¹School of Geomatics, ²School of Geospatial Information Service, Liaoning Technical University, Fuxin, Liaoning, China. ³Dalian Huangbohai Marine Surveying Data Information Co., Ltd, Dalian, Liaoning, China.

⁴College of Geomatics, Xi 'an University of Science and Technology, Xi 'an, Shaanxi, China

Received: July 19, 2023; accepted: September 4, 2023.

Chlorophyll concentration relies on the plankton present on the ocean surface that is affected by climatic changes. Remote sensing applications deploy sensing devices on the ocean surface for detecting the characteristics of plankton behaviors. Such a process generates time-dependent data for analyzing the chlorophyll concentrations. This article introduced a Hybrid Chlorophyll Concentration Estimation Scheme (HC2ES) using remote sensing application data. The hybrid process was employed by using a genetic algorithm and deep reinforcement learning for data segregation and concentration analysis. The genetic process assimilated the low and high chlorophyll (sensed) data for identifying the climatic impact on vegetation. The no-concentration data was segregated by using the mutation operation of the genetic process. The mutated data was recurrently analyzed by using current and previous concentration levels (data) in the reinforcement layers, which provided further classifications on high and low chlorophyll concentrations for providing insights into remote sensing expansions. The joint hybrid process filtered the fewer concentration data preventing its analysis from identifying chlorophyll and its associated inputs. The proposed scheme was validated by using analysis rate, classification, concentration detection, and computation time.

Keywords: chlorophyll concentration; data analysis; deep reinforcement learning; generic algorithm; remote sensing.

*Corresponding author: Xinmei Lan, School of Geomatics, Liaoning Technical University, Fuxin, Liaoning 123000, China. Email: xinmei_lan@outlook.com.

Introduction

Chlorophyll-a (Chl-a) concentration is an indicator of phytoplankton abundance in coastal areas and ocean surfaces. Chl-a concentration calculates the exact unit of chlorophyll-a which is presented in coastal water [1] and measures the absorbance of the extract in particular wavelengths. The actual role of Chl-a is that it releases chemical energy to the surface which provides essential elements for photosynthesis [2]. Chlorophyll-a concentration estimation is a crucial task to evaluate the actual potential

presence of phytoplankton and algae in coastal areas [3]. A quantum computing (QC) assisted support vector regression (SVR) is used for the Chl-a concentration estimation process, which mainly analyzes the potential capabilities of Chl-a concentration over coastal areas [4]. Xu *et al.* designed a SVR algorithm-based Chl-a prediction model to identify the Chl-a sensitivity level of the rivers [5]. Temporal and spatial variation and characteristics of Chl-a were identified by using the SVR algorithm. Remote sensing (RS) technology is used in QC to solve the complexity of the data identification process. The QC-based

SVR increases the feasibility level in the Chl-a concentration estimation process [6]. An absorption-specific approach is used to estimate the Chl-a concentration range on the coastal surface, which classifies the types of water based on the heterogeneous datasets captured *via* RS and estimates the exact clear level of Chl-a concentration on coastal water [7, 8]. Chl-a concentration estimation technique based on remote sensing is widely used in coastal areas to identify the Chl-a range in seawater [9] and the concentration level on the ocean surface. Remote sensed data are used in the estimation to reduce the latency in the computation process [10]. Chu and He developed a remote-sensing water quality estimation approach by using sparse representation for the retrieval process to identify the important spectral features from remotely sensed data, which produced necessary information for the water quality estimation process [11]. Vostokov *et al.* proposed a method to investigate the seasonal Chl-a variability based on remotely sensed data in the Caspian Sea to classify the seasonal variability level of Chl-a. The remotely sensed data was used to produce optimal measurement values and features for the classification process [12]. A genetic algorithm (GA) and artificial neural network (ANN) algorithms are used in the sensing technique to extract Chl-a concentration level and retrieve necessary information to estimate the level of water quality in coastal areas [13]. The remote sensing estimation technique increases the accuracy of the Chl-a concentration estimation process [14]. A ground-based remote sensing system (GRSS) is also used to calculate the Chl-a concentration ratio in coastal areas, which identifies the condition and blooming range of phytoplankton in coastal areas [15]. The blooming range provides optimal high-frequency data for the estimation process to reduce the latency in gathering relevant information for the concentration estimation process. The GRSS improves the accuracy of the Chl-a estimation process and enhances the performance level in further management systems [16, 17]. Bai *et al.* developed a novel chlorophyll (Chl) concentration estimation model to detect maize

actual Chl level in coastal areas [18]. However, the system consumed high computation complexity while analyzing the high Chl concentration level. Nguyen *et al.* established a multi-task convolutional neural network (MCNN) for Chl-a concentration estimation process to evaluate the Chl-a concentration range in ocean water [19], while Chen *et al.* built a multi-source data fusion method for Chl-a concentration estimation process [20]. Moreover, Zhou *et al.* and Cuartero *et al.* developed a water Chl estimation method for the urban canal system and a method of atmospheric correction (C2-NET) for Chl-a estimation in small reservoirs, respectively [21, 22]. However, all those methods were in need of improvement of computation problems and system's flexibility, availability, and accuracy.

Machine learning (ML) algorithms are mostly used for estimation and detection processes and are used for Chl-a concentration estimation process. An artificial neural network (ANN) algorithm-based estimation approach has been used to predict the Chl-a concentration level in coastal areas. The ANN algorithm is mainly implemented to detect the exact Chl-a concentration range in ocean waters by extracting the important features from remotely sensed data [8, 23], which minimizes the computational cost ratio of the estimation process. A random forest (RF) model is also used for Chl-a concentration estimation, which assesses the key values relevant to the Chl-a concentration variations and predicts the actual concentration level of Chl-a in coastal areas. RF detects the input variables from remotely sensed data which reduces the latency in the computation process [17, 24]. Asim *et al.* proposed a ML model for Chl-a concentration estimation in sea areas, which was a neural network model to predict the exact ocean color and spectra index level of Chl-a [25]. Ding *et al.* proposed a deep learning (DL) model for coastal Chl-a forecasting systems to forecast the accurate Chl-a concentration level of coastal waters, which provided optimal information to the forecasting system and increased the overall

capability range of forecasting systems [26]. Karimian *et al.* also proposed a new ML based framework for Chl-a concentration estimation, which used multi-source big data in the framework to provide essential indicators for further processes [27]. In addition, Kallio *et al.* introduced a spatiotemporal analysis model for the Chl-a concentration estimation process [28] by using satellite and *in-situ* data to reduce the latency in the data identification process. A Kalman filter was implemented in the model to filter the *in-situ* data and provide feasible data for the analysis process. Satellite data produced the exact content and characteristics of Chl-a level in the water.

Chlorophyll concentration detection relies on distinguishable factors including climate, region, and plankton densities. Analyzing invariant data may cause degradation of the approximate output, leading to varying results for null data, which has been addressed through previous and current levels of observations from time lined remote data. Therefore, possibility-based data handling and output assessment is required for improving concentration estimation. This study combined genetic algorithms with deep reinforcement learning to improve the accuracy and efficiency of environmental monitoring by optimizing parameters, extracting complex characteristics, and addressing the complexities of water quality interactions and to improve remote sensing-based Chl-a concentration estimation. A boosting trees (BST) algorithm-based model for Chl-a concentration estimation was developed to synchronize the variables with inputs that identified the absolute level of concentration, which enhanced the performance range of the estimation process [29-31]. The results of this study would potentially improve the precision and utility of Chl-a concentration estimation and provide a novel way for environmental monitoring. In addition, this study would be significant for making educated decisions in resource management, pollution control, and ecosystem preservation.

Materials and Methods

A Hybrid Chlorophyll Concentration Estimation Scheme (HC2ES) that utilized remote sensing application data to analyze and estimate chlorophyll concentrations in the ocean was developed through this study. The concentration of chlorophyll is closely related to the presence and behavior of plankton which in turn is influenced by climatic changes. The HC2ES approach combined genetic algorithms and deep reinforcement learning to process and analyze the time-dependent data obtained from remote sensing devices deployed on the ocean surface. The genetic algorithm helped in determining the high and low chlorophyll based on the repercussions of climatic factors on the ocean surfaces (plankton) and the consequences of the climatic characteristics on vegetation. The genetic algorithm happened between the cross-over and mutation process to estimate the maximum data depending on the time interval of the sensing data. The mutation procedure helped in extracting the data with no concentration in the further process. The acquired mutated data was determined recurrently by consolidating the previous and existing concentration levels of the sensed data. The data which had no chlorophyll concentration on the plankton was identified by using the genetic mutation procedure. The acquired mutated data was repeatedly analyzed in the reinforcement layers by consolidating both the existing and previous concentration levels, which helped in determining the concentration level by using deep reinforcement learning through the comparison to the previous concentration levels. These efficacious operations aided in extracting the lower concentrations by enhancing the mutation and cross-over process. The actual concentration and then the reinforced level were obtained as the input to the further classification process for the determination of lower concentrations. By analyzing the genetic data, the algorithm provided insights into the chlorophyll concentration in the plankton, contributing to the understanding of the ocean ecosystem. The process of identifying the sensing devices for the

analyzing process of the ocean surface was described by Equation (1) below:

$$\left. \begin{aligned}
 Q_t &= Q[q_t + \alpha q_{t+1} + \alpha^2 q_{t+2} + \dots] \\
 &= \alpha [\sum_{n=0} \alpha q_{t+n}] \\
 Q(\alpha, \beta) &= [\alpha_t = \alpha_1, \beta_t = \beta_1] \\
 &= \beta [\sum_{n=0} \beta q_{t-n}] \\
 Q\beta_{n+1}(\alpha, \beta) &= \sum_{n=0} |\alpha_t = \alpha_1, \beta_t = \beta_1| \\
 &= \sum_{n=0} [Q_{t+1} + \alpha_{t+1} + \beta_{t+1}]
 \end{aligned} \right\} \quad (1)$$

where Q was the sensing devices for the analyzing process, α was the information captured by the sensing device, β was the chlorophyll levels detected by the sensing device, t was the aggregation of the plankton plants in the ocean surface. The time-sensed data was collected from the sensing devices of the ocean surface to extract the concentration of chlorophyll in the plankton plants, which helped in extracting the chlorophyll concentration on the ocean surface and provided the information of the plankton according to the vegetation characteristics. The sensed data aided in the classification of the genetic algorithm process which included the cross-over and mutation operation. The extraction of these data from the ocean surface was utilized in the detection of the impacts of climatic changes on the growth and then distribution of the plankton on the ocean surface, which was significant in evaluating the health of the ocean surface where the chlorophyll determination was happening and helped in detecting the overall productivity of the plankton on the ocean surface. Based on the time-sensed data, the process of determining the time sensed data from the sensing devices for the further genetic algorithm procedure was described by the Equation (2):

$$\left. \begin{aligned}
 \alpha_i &= [\alpha^{(1)}, \alpha^{(2)}, \dots, \alpha^{(n)}] \\
 \hat{\beta} &= Q(\alpha) \\
 &= \sum_{t=1} \alpha_1(\alpha_1, \alpha_2) \\
 \beta(\alpha_1, \alpha_2) &= \sum_{n=1} (\alpha_{(t)}^1 - \alpha_{(t)}^2) \\
 (\alpha\beta)^T &= \sqrt{\sum_{t=1} \frac{(\beta_1 - \beta_2)^2}{n}} \\
 &= \frac{1}{n} \sum_{t=1} |\hat{\beta}_1 - \beta_2| \\
 \beta^2 &= 1 - \frac{\sum_{t=1} (\beta_t - \hat{\beta}_t)^2}{\sum_{t=1} (\beta_t - \bar{\beta})^2}
 \end{aligned} \right\} \quad (2)$$

where i was the time sensed data from the sensing devices of the ocean surface. The low-level chlorophyll concentration was determined based on the time-sensed data by using the cross-over and mutation process in the genetic algorithm, which was associated and enhanced the new plankton with the high-level chlorophyll concentration on the ocean surface. This mutation procedure aided in the effective understanding of the low-concentration behavior of the plankton and the variations of the plankton could be determined by the outcome of the cross-over operation. By associating cross-over and mutation processes, the genetic algorithm process helped in better analysis of time-sensed data and also the efficacy understanding of the low chlorophyll concentrations in the plankton. The process of extracting the low chlorophyll sensed data was shown in Equations (3) and (4).

$$\left. \begin{aligned}
 \beta &= W\alpha + \gamma \\
 W &= \frac{\sum_{t=1} (\alpha_t - \bar{\alpha})}{\sum_{t=1} \alpha^2 - \frac{1}{n} (\sum_{n=1} \alpha_{(1)}^2)^2} \\
 \beta &= t_0 + \sum_{n=1} \alpha_i + \sum_{n=1} \beta_i + \sum_{n=1, n=1} (\alpha_i \beta_i) * t \\
 \beta &= t_0 + \sum_{n=1} (\alpha_1, \beta_1) \\
 &= \sum_{n=1} \beta_t (w_t \cdot \alpha_t + \beta_t) \\
 &\text{where } t = 1, 2, \dots, n
 \end{aligned} \right\} \quad (3)$$

$$\left. \begin{aligned}
 \sum_{j=1}^n |0 - t| &= 0 \\
 \sum_{j=1}^n \beta_i (\alpha_i + \beta_i) &= t_j \\
 \sum_{j=1}^n \beta_i (Q + W) &= t_j \\
 \alpha\beta &= T \\
 \tilde{\beta} &= Q * T \\
 Q_i &= \frac{n}{Q_i} \\
 \beta_i &= \frac{n_i}{\sum_{n=1} q_i}
 \end{aligned} \right\} \quad (4)$$

where W was the low chlorophyll sensed data, which was extracted by the genetic algorithm procedure, γ was the variations in the concentration of the chlorophyll in the plankton. The high chlorophyll concentration was determined by using the genetic algorithm and the remote sensing data. The genetic algorithm

played a crucial role in determining high chlorophyll concentrations in plankton by using remote sensing data, which provided information about the ocean surface and the presence of plankton. The genetic algorithm identified the patterns and relationships associated with high chlorophyll levels. The algorithm utilized the principles of natural selection and evolution to iteratively assess and refine the data, allowing for the identification of regions or periods with high chlorophyll concentration in the plankton, which enabled us to gain insights into areas of the ocean with rich vegetation, contributing to the understanding of the factors influencing chlorophyll distribution and its ecological significance in marine ecosystems. The process of determining the high chlorophyll sensed data was explained by the Equation (5):

$$\left. \begin{aligned}
 &F_{ij} = F_{ij}(1 - \beta) + \alpha_{ij}\beta \\
 &F_{it} = F_{it}(1 - \alpha) + \beta_{ij}\alpha \\
 \alpha_{ij} &= \begin{cases} \alpha_{ij} + (\alpha_{ij} - \alpha_t) * F(Q) & \beta_1 \geq t \\ \alpha_{ij} + (\alpha_t - \alpha_{ij}) * F(Q) & \beta_1 < t \end{cases} \\
 &F(Q) = \beta_2 * Qt \\
 \alpha_{it} &= \begin{cases} \alpha_{it} + (\alpha_{it} - \beta_t) * F(Q) & \beta_1 \geq t \\ \alpha_{it} + (\beta_t - \alpha_{it}) * F(Q) & \beta_1 < t \end{cases}
 \end{aligned} \right\} \quad (5)$$

where F was the high-concentration chlorophyll data, j was the efficiency of the remote sensing data. To handle the data with no concentration, the genetic process employed a mutation operation which selectively modified the genetic information associated with the no-concentration data, allowing for its segregation from the rest of the dataset. By introducing random variations through mutation, the genetic algorithm separated and identified the data points that did not exhibit significant chlorophyll concentrations. This segregation step ensured that the subsequent analysis and estimation focused on the relevant data, enhancing the accuracy of chlorophyll concentration determination in the plankton. The process of segregating the no-concentration data was shown in Equation (6):

$$\left. \begin{aligned}
 &\eta = \sqrt{\frac{\sum \alpha_i^2}{n}} \\
 &\text{where } i = 1,2,3, \dots, n \\
 &\eta_1 = \frac{\sum \alpha_i}{n} \\
 &\text{where } i = 1,2,3, \dots, n \\
 &\eta_2 = \sqrt{\frac{\sum \beta_i^2}{n}} \\
 &\text{where } i = 1,2,3, \dots, \infty \\
 &\eta_3 = \frac{\sum \beta_i}{n} \\
 &\text{where } i = 1,2,3, \dots, \infty
 \end{aligned} \right\} \quad (6)$$

where η was the no-concentration data which was determined by the mutation procedure of the genetic algorithm. The γ analysis for the year and mutated output ($t + i$) was performed in addition to η for the varying years in handled. Considering the η as $(\eta_1 + \eta_2 + \dots + \eta_i)$, the variation in any instance was checked for its forward interval of i or t or both. Therefore, the j differentiation for $\eta \in (F, W)$ was used for computing the difference between successive intervals. The identified ($t + i$) was utilized for segregating η across t . In this case, $\beta_1 < t$ or $\beta_1 \geq t$ was further classified for preventing new no-concentration data. The acquired no-concentration data was then sent as the input to the concentration analyzing process by using deep reinforcement learning. The extraction of the current concentration of chlorophyll in plankton was facilitated through the use of deep reinforcement learning, which built upon the outcomes of the genetic algorithm process. Deep reinforcement learning involved training a neural network to make sequential decisions based on past and current concentration levels. By consolidating the information obtained from the genetic algorithm process, the neural network classified and estimated the current chlorophyll concentration in the plankton. This iterative analysis by utilizing both historical and real-time data enabled the model to learn patterns and make predictions about the current chlorophyll levels. Deep reinforcement learning (DRL) enhanced the accuracy and understanding of chlorophyll concentration dynamics in the plankton population, contributing to more comprehensive insights into oceanic ecosystems.

The process of determining the current chlorophyll concentration by using DRL was shown in Equations (7) and (8).

$$\left. \begin{aligned}
 M_{i,n} &= \{(\alpha_1, T_1)(\alpha_2, T_2) \dots (\alpha_n, T_n)\} \\
 M(t) &= \alpha_{i+n} + \left(\frac{\alpha_i + \alpha + \alpha_{i+\beta}}{T_1 + T_2}\right) \\
 Ft &= \eta(\omega_n + \alpha_t) \\
 Q_t &= \{(Ft + \alpha_t) + (\beta_t * \alpha_t)\} \\
 \beta_t &= \sum_{n=1} \frac{\eta_t}{\omega_{t+1}} \\
 \alpha(t) &= \sum_{i=1} F_i(t) + \sum_{j=1} F_j(t)
 \end{aligned} \right\} \tag{7}$$

$$\left. \begin{aligned}
 \sigma_1 &= \frac{1}{n} \sum_{i=1} |M_i - F_i| \\
 \sigma_2 &= \frac{1}{n} \sum_{i=1} (M_i - F_i)^2 \\
 \sigma_3 &= \sqrt{\frac{1}{n} \sum_{i=1} (M_i - F_i)^2} \\
 \sigma_4 &= \frac{1}{n} \sum_{i=1} \left| \frac{M_i - F_i}{M_i} \right| \\
 \sigma_5 &= \frac{1}{n} \sum_{i=1} \|M_i - F_i\|^2 \\
 \sigma_6 &= \frac{1}{n} \sum_{i=1} [\|M_i\| + \|F_i\|]
 \end{aligned} \right\} \tag{8}$$

where M was the current concentration of the chlorophyll in the planktons, σ was the outcome of the genetic algorithm process. The previous concentration data level was determined by using the DRL technique by considering the outcome of the genetic algorithm process. By comparing the previous concentration levels, the existing concentration levels of the chlorophyll on the plankton were then determined. The DRL helped in analyzing the historical concentration levels of the plankton on the ocean surface for the further classification process, which reinforced the input to the upcoming classification process. Based on both actual and reinforced inputs, the actual concentration and no concentrations were identified. With the help of the genetic algorithm process output, the precision of the chlorophyll concentrations was detected and extorted the assessments of the chlorophyll levels over the various time intervals. Hence the accurate chlorophyll concentrations were determined depending on the previous concentration levels of the plankton. The

determination of previous chlorophyll level was calculated by Equation (9).

$$\left. \begin{aligned}
 Z_t &= \sqrt{\frac{1}{n} \sum_{i=1} (\sigma_1 - \sigma_2)} \\
 Z_i &= \frac{1}{n} \sum_{n=1} \frac{|\sigma_1 - \sigma_2|}{\sigma_i} \\
 &= 1 - \sum_{n=1} \frac{(\sigma_1 - \sigma_2)^2}{(\sigma_1 + \sigma_2)^2} \\
 \alpha_t &= F(\omega_1 \alpha_t + \omega_2 \alpha_{t-1}) \\
 Z_{i,n}(\alpha_t) &= \sum_{n=1} \left| \frac{\sigma_t}{\sigma_n} \right|
 \end{aligned} \right\} \tag{9}$$

where Z was the previous chlorophyll concentration level in the plankton. The mutated data was identified during the determination of the current and previous concentration levels recurrently. The mutated data, which had undergone segregation through the genetic algorithm process, was recurrently analyzed within the reinforcement layers where the mutated data was examined alongside previous concentration levels. Through the consideration of the historical data, the reinforcement layers enabled the model to capture temporal patterns and dependencies. This iterative process enhanced the understanding of chlorophyll concentration dynamics over time and facilitated more accurate estimation and classification of chlorophyll levels in plankton, contributing to a comprehensive analysis of the ocean's ecosystem. The process of analyzing the mutated data was determined by Equation (10).

$$\left. \begin{aligned}
 E_\sigma &= \left[\left(\gamma \cdot \frac{F_m}{F_m - F_0} \right) - 1 \right] \cdot F_m \\
 Q_\sigma &= \frac{F_n^t - FQ}{F_n^t - F_0} \\
 Q_\eta &= \left(\frac{1}{F_m} - \frac{1}{F_n^t} \right) \\
 &= \left(\frac{1}{F_0} - \frac{1}{F_m} \right) \\
 E_n &= \alpha + \beta_1 \left(\frac{m_1}{m_2} \right)
 \end{aligned} \right\} \tag{10}$$

where E was the mutated data analyzed recurrently. The deep reinforcement learning process was illustrated in Figure 1. The learning process for classifying E_σ and M was performed by using W and F . Z was used for $(i * \sigma)$ provided $(M(t))$ satisfied $Ft = \eta(\omega_n + \alpha_t)$

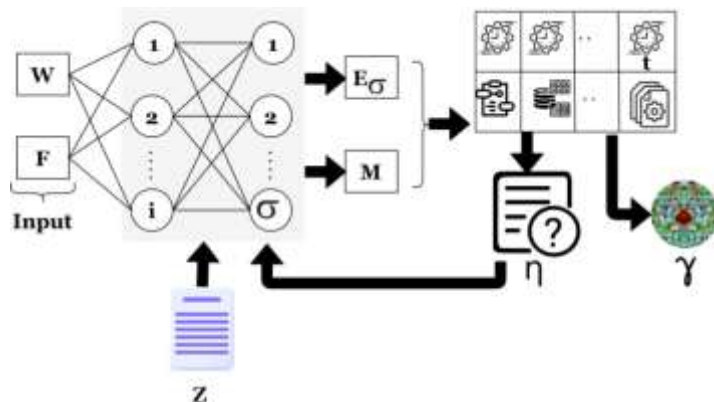


Figure 1. Reinforcement learning process.

condition, which was required for $Z_{i,n}(\alpha_t)$ suppressing γ in consecutive intervals that were required for identifying η in any E_σ . Therefore, the new augmentation in t was utilized for classifying γ from Z with M . The classification process took place with the outcome of the current and previous concentration determination process. The recurrent analysis of the mutated data and incorporating both current and previous concentration levels took place in the reinforcement layers. This iterative process enabled further classifications of high and low chlorophyll concentrations, offering valuable insights for the expansion of remote sensing capabilities, and enhancing the understanding of chlorophyll dynamics in the context of oceanic ecosystems. The classification operation was described by Equation (11).

$$\left. \begin{aligned} R^2 &= 1 - \frac{\sum_{i=1} (Q_1 - Q_2)^2}{\sum_{i=1} (Q_1 - Q)^2} \\ R_t^2 &= \sqrt{\frac{1}{n} \sum_{n=1} (E_i - E_t)^2} \\ R_\eta &= \frac{\frac{1}{n} \sum_{i=1} |\beta_i - \beta_t|}{\frac{1}{n} \sum_{i=1} |\alpha_i - \alpha_t|} \\ R_{\alpha,\beta} &= \frac{1}{n} \sum_{i=1} |\alpha_t, \beta_t| \\ R(\alpha * \beta) &= \frac{\sum_{n=1} |\alpha_t|}{\sum_{n=1} |\beta_t|} \end{aligned} \right\} \quad (11)$$

where R was the classification procedure. The current concentration of the plankton's chlorophyll produced the actual input for the evaluation of the actual concentration in the

plankton, which provided a reference point for comparison and analysis. Assessing the current chlorophyll concentration helped in evaluating the state of the plankton's chlorophyll content, making inferences about its ecological significance, understanding the overall health and productivity of the plankton population and its chlorophyll concentration. The process was expressed as follows:

$$\left. \begin{aligned} S^2 &= \frac{(\sum_{i=1} \beta_i - \beta)^2}{\eta} \\ S_t &= \frac{(\beta_i - \alpha_t)^2}{\|\alpha_t - \alpha_n\|^2} \\ \eta S t &= \sqrt{\frac{\sum_{n=1} (\beta_i - \alpha_t)^2}{n}} \\ \eta S(E_t) &= \frac{E_t}{\eta E_{n=1}} \\ E(t, n) &= \frac{\sum_{n=1} (E_t)}{\sum_{n=1} \left\| \frac{\alpha_i - \alpha_j}{\eta} \right\|} \end{aligned} \right\} \quad (12)$$

where S was the actual concentration. The no-chlorophyll data was determined by the reinforced input which was produced by the previous concentration of DRL. The DRL model provided valuable insights into the absence of chlorophyll in certain instances. The actual and analyzed concentrations with γ were calculated. The computed concentration was derived from R and E_σ across $E(t, n)$. The conditions $\beta_1 < t$ and $\beta_1 \geq t$ were required to be satisfied for variations and concentration assessments. Depending on the multiple inputs, the σ was

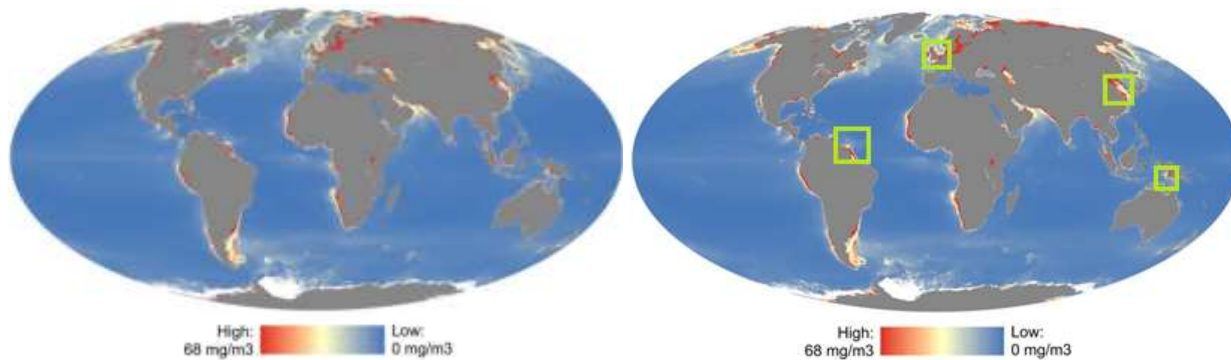


Figure 2. Chlorophyll concentrations over the world. The highlighted areas were the selected concentration detection areas.

required for γ detection due to which R_η was pre-estimated. This pre-estimation was a predictive analysis for defending minimum variations. Analyzing the reinforced input helped to identify and classify data points that exhibited no significant chlorophyll concentration, which helped in distinguishing and segregating the no-chlorophyll data from the dataset, enabling a focused analysis of the plankton samples where chlorophyll was not present. This step contributed to a more comprehensive understanding of chlorophyll dynamics in the plankton population and was performed by Equation (13).

$$\left. \begin{aligned} \rho_t &= \frac{Q_t \cdot \alpha_t}{Q_t \cdot \beta_t} \\ \rho_n &= \frac{w_{i=1}(t \cdot n)}{w_{i=1}(\frac{t}{n})} \\ \rho_n &= \frac{Q_{n=1}}{Q_{t=1}} \left[\frac{1}{n} \right] \end{aligned} \right\} \quad (13)$$

where ρ was the no-chlorophyll data identified by the reinforced input. The joint hybrid process effectively filtered out data points with lower concentrations, ensuring no interfere with the analysis aimed at identifying chlorophyll and its associated concentrations. By removing the data with fewer concentrations, the process improved the accuracy and reliability of chlorophyll identification and concentration analysis in the plankton population. To validate the proposed scheme, various measures including analysis rate, classification accuracy, concentration detection capability, and computation time were

used to assess the performance and effectiveness in analyzing chlorophyll concentrations in plankton.

Results and discussion

HC2ES detected the remote sensing data from the ocean surface according to the vegetation factors, and then, the low and high chlorophyll concentrations were determined. This detection procedure established the efficaciousness of the proposed method in detecting the chlorophyll concentrations on the ocean surface (plankton). Remote sensing devices helped sense the ocean surfaces for the further classification process and extract information about the plankton, and thus helped in the detection of the present chlorophyll concentration levels in the ocean surface by comparing with the previous levels. This information could be used to determine the concentration of chlorophyll in plankton. The genetic algorithm identified the sensed data, both low and high chlorophyll levels, and the relationship between climatic changes and vegetation. The low and high chlorophyll concentrations over the world’s ocean surface were analyzed between 2009 and 2013 by using yearly merged chlorophyll concentration (mg/m^3) data obtained at depth of 9 km submerged planktons with the region selected for concentration detection restricted to the certain ocean surface (Figure 2) [32]. The genetic process for assimilating low- and high-

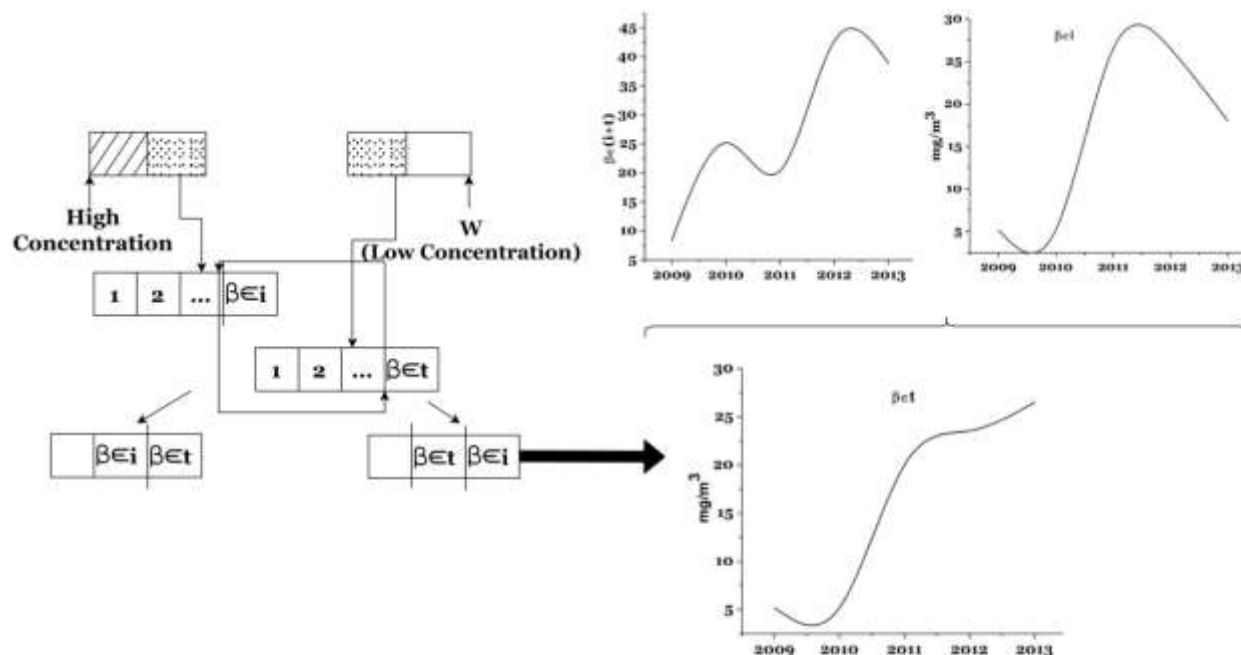


Figure 3. Low and high concentration data assimilation.

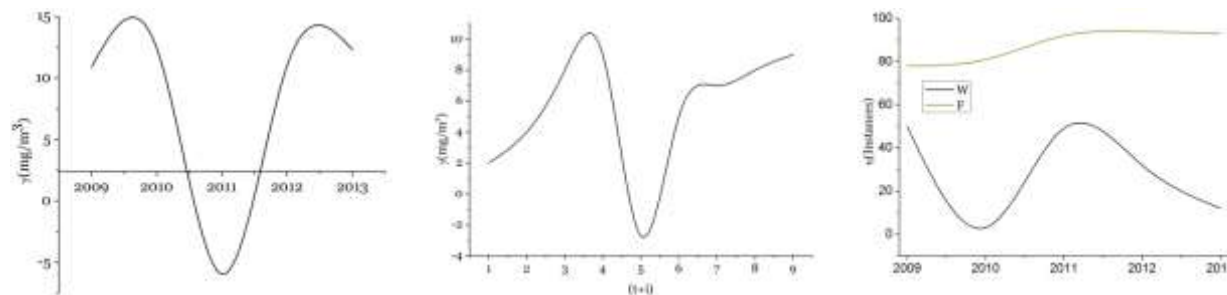


Figure 4. Variations and no-concentration data analysis.

concentration data was shown in Figure 3. The high and low concentration data was utilized for identifying nullified data (concentration), which was prevented by identifying multiple (i, t) instances using $\alpha \in \beta$. The levels were inflated based on climatic changes for which the γ between the $\beta \in i$ and $\beta \in t$ was suppressed for no concentration data, which was required for identifying high concentration and preventing further information falsifications. The $\beta \in i$, $\beta \in t$, and $\beta \in (i + t)$ (mutated) were analyzed, from which the unavailable concentrations were filled. Thus, a slighter hike in concentration detection was observed. The variations and no-

concentration data obtained from this study were shown in Figure 4. The current concentration based on the analytics of the actual and identified concentrations with the variation were analyzed (Figure 5), while the ρ analysis for different seasons between 2009 and 2013 was presented in Figure 6. The variables of S1, S2, S3, and S4 represented the 4 seasons in a year. The seasons were considered commonly for different years. The no-concentration data was obtained as an error in computation or actual unavailability from the sensed source, which streamlined the reinforcement training for η (alone) preventing errors. Such γ segregations

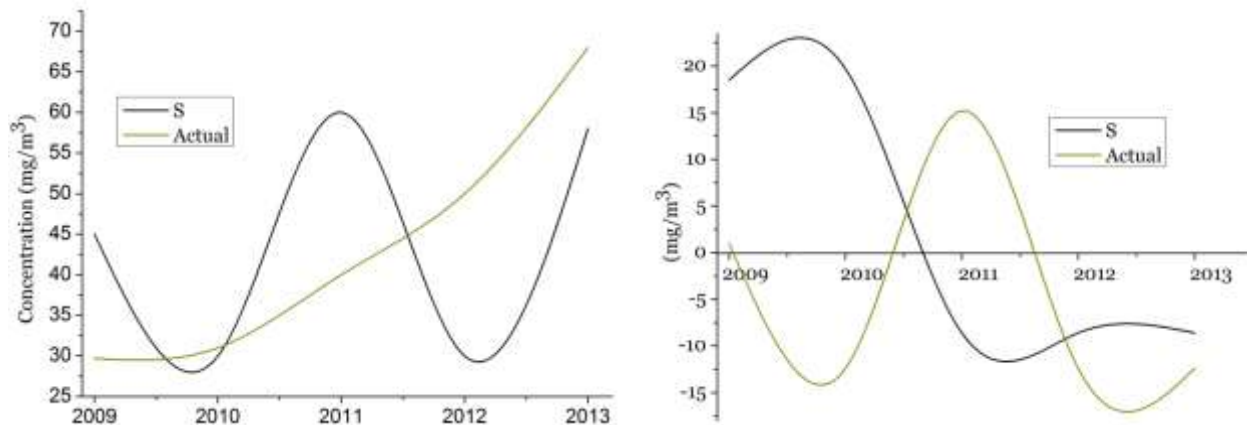


Figure 5. Current concentration analysis.

were listed for multiple M extracting E_{σ} . Thus, the approximated ρ for different seasons were analyzed.

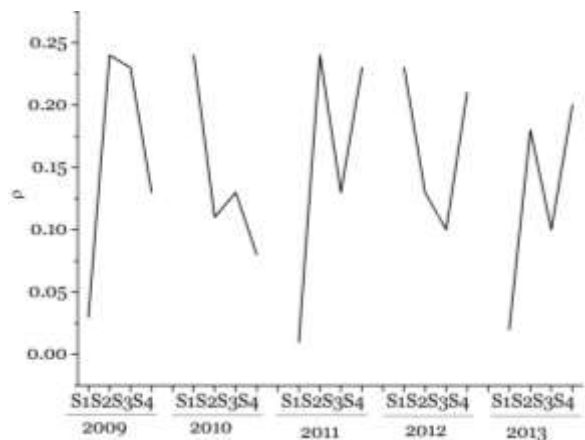


Figure 6. ρ analysis between 2009 and 2013.

The DRL’s concentration classification for current and previous β analyzed was shown in Figure 7. The DRL process was intended to perform classification by using (F, W) for preventing γ and concentration errors. In this classification, $\beta < t$ and $\beta \geq t$ were independently used for γ detection and ρ matching, which relied on t and i before the mutation process. Post the assessment the $\beta \in (t + i)$ was required for increasing the classifications with the previous and current i . Therefore, the η was suppressed at any β condition failing instance, reducing the

error chances.

The proposed HC2ES scheme was evaluated by comparing it to the methods of MCNN [19], DFS+KF [28], and MSBD+MLF [27] in analysis rate, chlorophyll classification accuracy, the ability of concentration detection, and computation time. The analysis rate was efficacious in this study by combining genetic algorithm and deep reinforcement learning technique. The genetic algorithm efficiently processed and assimilated the sensed data, enabling the identification of patterns and relationships between climatic changes and vegetation, which provided a foundation for the subsequent application of deep reinforcement learning, which further refined the estimation process. The reinforcement learning model leveraged the learned knowledge from the genetic algorithm, allowing for faster and more accurate analysis of current and previous concentration levels. By using the two techniques of cross-over and mutation in the genetic algorithm process, the analysis of the chlorophyll concentration was effective based on the time-sensed data obtained from the sensing devices. (Figure 8). The classification of the actual chlorophyll concentration and no-chlorophyll data was performed by using deep reinforcement learning. By using the outcomes of the genetic algorithm process, the previous and then the current chlorophyll concentrations were extorted. After

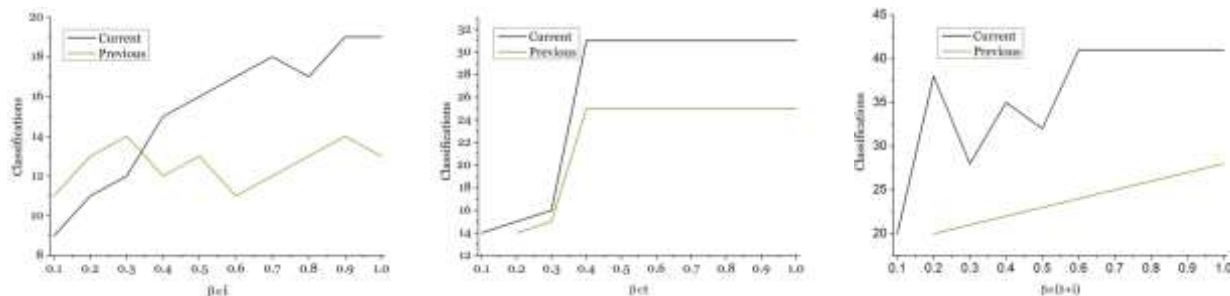


Figure 7. The DRL’s concentration classification analysis.

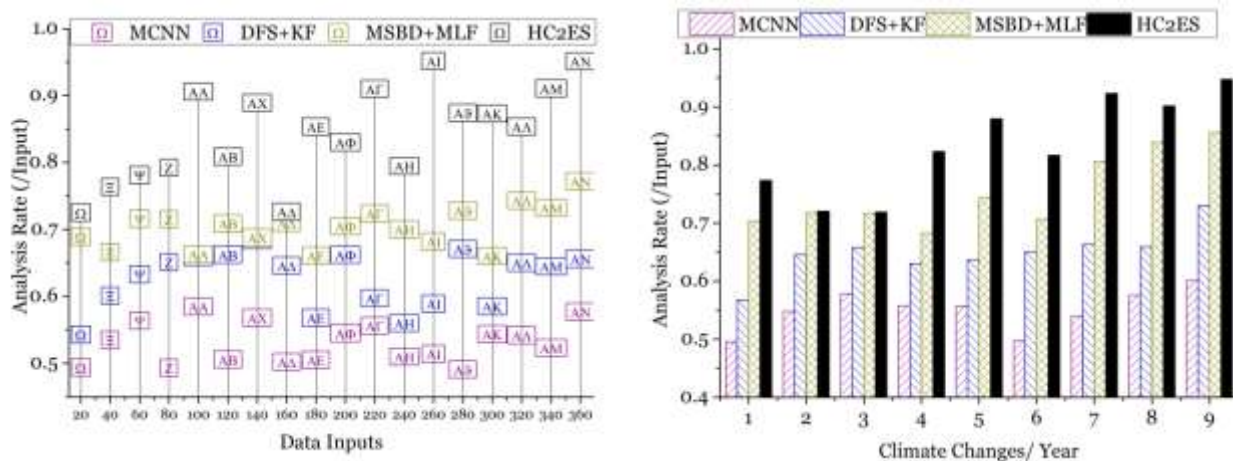


Figure 8. Analysis rate comparison of the chlorophyll concentration based on the time-sensed data obtained from the sensing devices.

determining the concentration levels, the actual and then the reinforced input were given to the classification process. The maximum data was extracted based on the time-sensed information provided by the sensing devices from the ocean surface, while the lower concentration data was determined, and further steps were taken to enhance the chlorophyll concentration level on the plankton. By consolidating the results from those processes, the classification of the actual chlorophyll and no concentration data was effective and the overall productivity of the plankton on the ocean surface was determined. (Figure 9). The concentration detection was high in this method by using DRL for the identification of the current and previous concentration of the chlorophyll in

the plankton after determining the low and high chlorophyll concentrations by using the genetic algorithm. According to the vegetation factors on the ocean surface, the time-sensed data was collected for the chlorophyll concentration extraction process. The DRL helped in establishing the actual and reinforced input to the classification operation in which the actual and no concentrations in the plankton were identified. By leveraging these acquired time-sensed data and then the outcome of the genetic algorithm process, the concentration detection process was conducted to produce the precise information of the chlorophyll concentration and enhance the accuracy of the detection (Figure 10). The time taken for the computation was less in this study. The computation process was performed with time-sensed data which produced precise information about the ocean

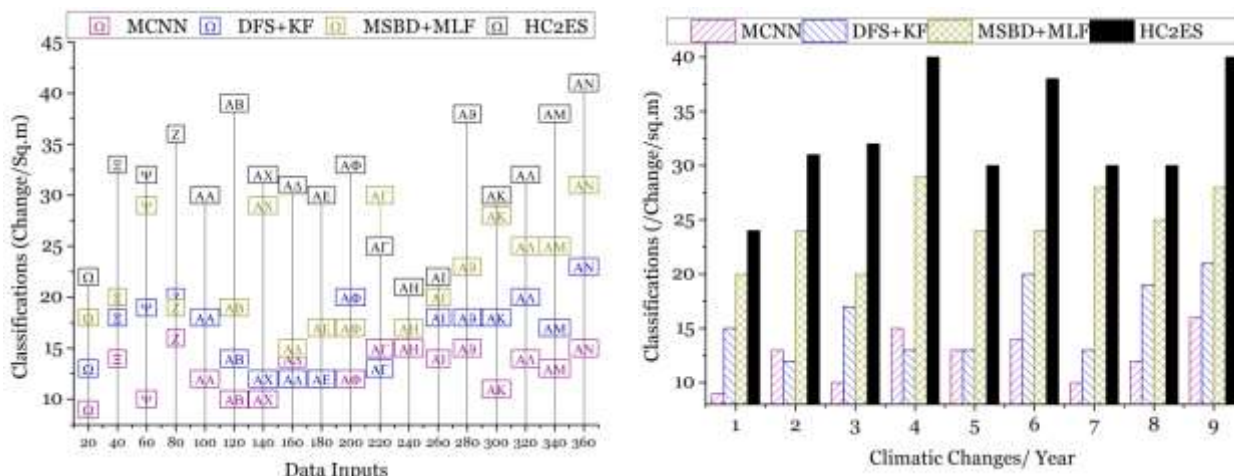


Figure 9. Classification analysis comparison of the actual chlorophyll and no concentration data in different methods.

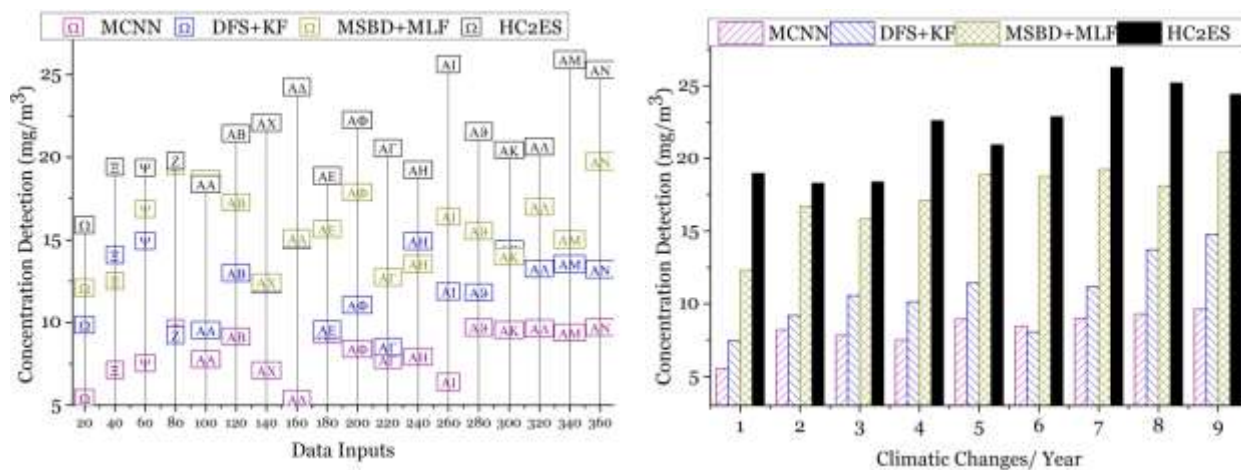


Figure 10. Comparison of chlorophyll concentration detection.

surface and the productivity of the plankton according to the vegetation factors. Hence, by establishing the accurate relationship between the ocean surface and plankton production, the computation time needed in this study was lesser than the other methods. With this genetic algorithm output, the further classification process was taken by using the DRL technique, which resulted in the faster and more effective determination of both low and high chlorophyll levels in the plankton. Using the effective operation in the genetic algorithm, the computation time was reduced and hence provided accurate information of the chlorophyll concentrations in the planktons. (Figure 11). The

errors in the classification were also lesser in this process than the others. The utilization of the DRL technique, coupled with the outcomes of the genetic algorithm process led to reduced errors in the classification of chlorophyll concentration levels in plankton. The genetic algorithm served as a preliminary analysis step, identifying patterns and relationships between chlorophyll concentration and other factors. The DRL model then consolidated the knowledge learned from the genetic algorithms' outputs to improve the accuracy of classification. By leveraging the strengths of both techniques, the DRL model refined the classification process, effectively distinguishing between low and high levels of

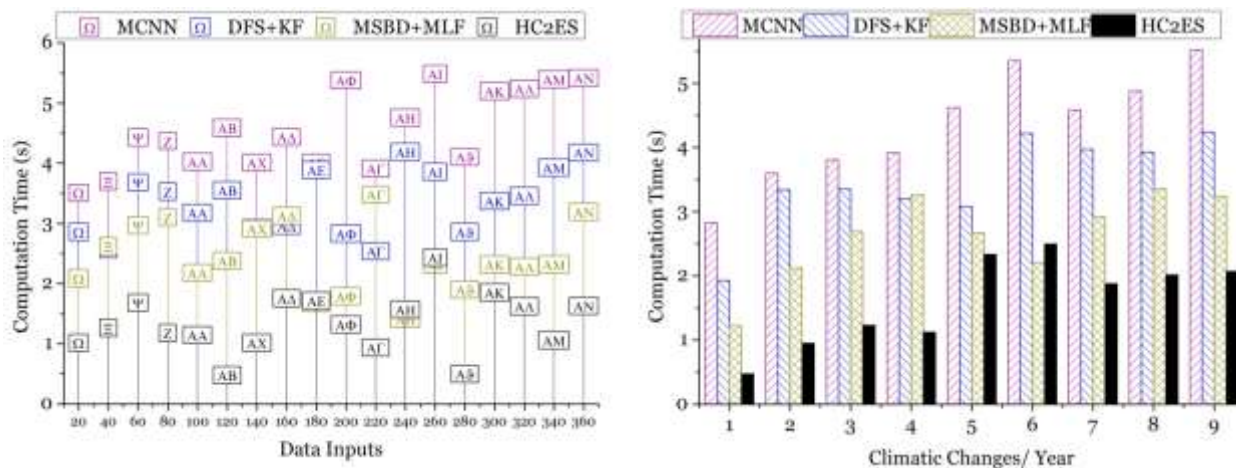


Figure 11. Comparison of computation time.

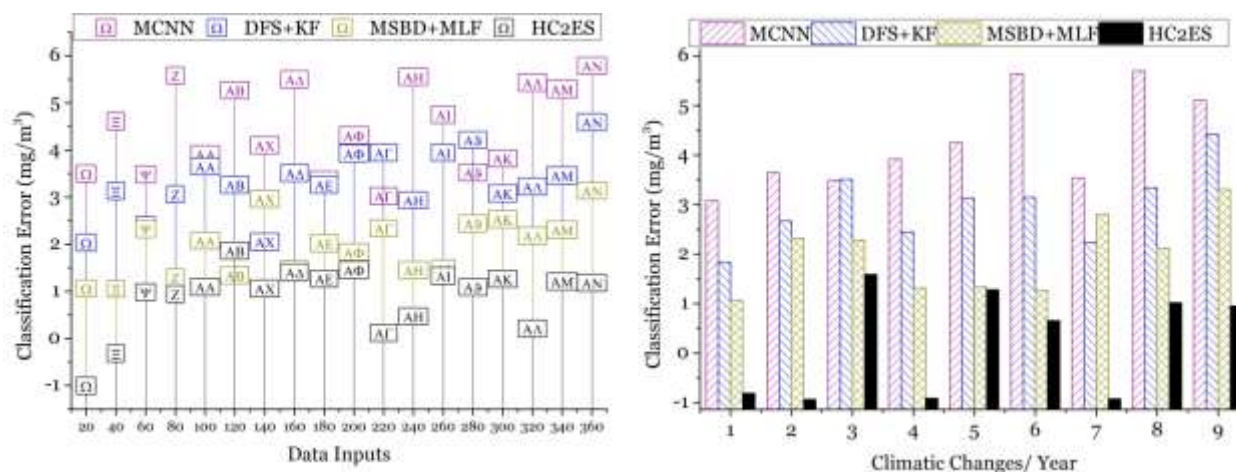


Figure 12. Comparison of classification error.

chlorophyll concentration in plankton. The iterative learning and feedback loop within the DRL framework contributed to a reduction in errors, ensuring more reliable and precise classification results for better understanding and monitoring of chlorophyll dynamics in plankton populations (Figure 12).

Conclusion

This study introduced an effective method for chlorophyll concentration assessment from the ocean surface by using the hybrid scheme harmonizing genetic algorithm and deep reinforcement learning. The proposed scheme

identified the low and high concentrations with the maximum possible combinations by using a genetic algorithm. The mutation process of the genetic algorithm reduced the nullified detection from different sensed intervals preventing erroneous classifications. The low, high, and null concentration classifications provided various insights on further remote sensing inputs. The mutated concentration input was utilized by reinforcement learning for verifying the current and previous concentration levels, which was required for validating the efficacy of the proposed scheme across different sensing intervals. The learning paradigm was trained by using the available high-concentration data without increasing the possible combinations.

The proposed scheme improved classifications by 15.28% and reduced the error by 12.93% under different climatic changes/ year. The method developed in this study could overcome the difficulty of deducing chlorophyll-a concentration from satellite images. The model was meant to optimize parameters and automatically extract significant features, enhancing estimation precision through the combination of genetic algorithms and deep reinforcement learning. This multidisciplinary strategy demonstrated a great promise of better comprehending the dynamics of water quality through cutting-edge AI methods. The results of this study provided a novel combination of cutting-edge AI methods with environmental science, which had the potentials to open up new avenues of interdisciplinary investigation, lead to more comprehensive model development, and advance both theory and practice.

Acknowledgement

This work was supported by National Natural Science Foundation of China (NSFC) (Project No. 42071343).

References

- Chen J, Chen S, Fu R, Wang C, Li D, Peng Y, *et al.* 2021. Remote sensing estimation of Chlorophyll-A in Case-II waters of coastal areas: Three-band model versus genetic algorithm-artificial neural networks model. *IEEE J Sel Top Appl Earth Obs Remote Sens.* 14:3640-3658.
- Wang WK, Shi Y, Zhang N, Li X, Sun D, Zhang Y, *et al.* 2022. A ground-based remote sensing system for high-frequency and real-time monitoring of phytoplankton blooms. *J Hazard Mater.* 439:129623.
- Free G, Bresciani M, Pinardi S, Simis S, Liu X, Albergel, C. *et al.* 2022. Investigating lake chlorophyll-a responses to the 2019 European double heatwave using satellite remote sensing. *Ecol Indic.* 142:109217.
- Jiang G, Loiselle SA, Yang D, Ma R, Su W, Gao C. 2020. Remote estimation of chlorophyll a concentrations over a wide range of optical conditions based on water classification from VIIRS observations. *Remote Sens Environ.* 241:111735.
- Xu L, Hao G, Li S, Song F, Zhao Y, Guo P. 2023. Prediction and sensitivity analysis of chlorophyll a based on a support vector machine regression algorithm. *Environ Monit Assess.* 195(6):698.
- Tong Y, Feng L, Zhao D, Xu W, Zheng C. 2022. Remote sensing of chlorophyll-a concentrations in coastal oceans of the greater bay area in China: Algorithm development and long-term changes. *Int J Appl Earth Obs Geoinf.* 112:102922.
- Souza APD, Teodoro PE, Teodoro LPR, Taveira AC, de Oliveira-Júnior JF, Della-Silva JL, *et al.* 2021. Application of remote sensing in environmental impact assessment: A case study of dam rupture in Brumadinho, Minas Gerais, Brazil. *Environ Monit Assess.* 193(9):606.
- Papenfus M, Schaeffer B, Pollard AI, Loftin K. 2020. Exploring the potential value of satellite remote sensing to monitor chlorophyll-a for US lakes and reservoirs. *Sci Inventory.* 192(12):808.
- Ye X., Liu J, Lin M, Ding J, Zou B, Song Q. 2021. Global ocean chlorophyll-a concentrations derived from COCTS onboard the hy-1c satellite and their preliminary evaluation. *IEEE Trans Geosci Remote Sens.* 59(12):9914-9926.
- Pasetto E, Riedel M, Melgani F, Michielsen K, Cavallaro G. 2022. Quantum SVR for chlorophyll concentration estimation in water with remote sensing. *IEEE Geosci Remote Sens Lett.* 19:1-5.
- Chu HJ, He YC. 2023. Remote sensing water quality inversion using sparse representation: Chlorophyll-a retrieval from sentinel-2 MSI data. *Remote Sensing Applications: Society and Environment.* 31:101006.
- Vostokov SV, Sahling IV, Vostokova AS, Gadzhiev AA, Lobachev EN, Behrouz A, *et al.* 2023. Seasonal chlorophyll variability in the Caspian Sea based on remote sensing data. *Dokl Earth Sci.* 508:47-52.
- Zhu X, Guo H, Huang JJ, Tian S, Zhang Z. 2023. A hybrid decomposition and machine learning model for forecasting chlorophyll-A and total nitrogen concentration in coastal waters. *J Hydrol.* 619:129207.
- Jang MTG, Alcântara E, Rodrigues T, Park E, Ogashawara I, Marengo JA. 2022. Increased chlorophyll-a concentration in barra bonita reservoir during extreme drought periods *Sci Total Environ.* 843:157106.
- Chan SN, Fan YW, Yao XH. 2022. Mapping of coastal surface chlorophyll-a concentration by multispectral reflectance measurement from unmanned aerial vehicles. *J Hydro-Environ Res.* 44:88-101.
- Liu X, Steele C, Simis S, Warren M, Tyler A, Spyarakos E, *et al.* 2021. Retrieval of chlorophyll-a concentration and associated product uncertainty in optically diverse lakes and reservoirs. *Remote Sens Environ.* 267:112710.
- Rodríguez-López L, Duran-Llacer I, González-Rodríguez L, Abarca-del-Río R, Cárdenas R, Parra O, *et al.* 2020. Spectral analysis using LANDSAT images to monitor the chlorophyll-a concentration in Lake Laja in Chile. *Ecol Inform.* 60:101183.
- Bi K, Xiao S, Gao S, Zhang C, Huang N, Niu Z. 2020. Estimating vertical chlorophyll concentrations in maize in different health states using hyperspectral LiDAR. *IEEE Trans Geosci Remote Sens.* 58(11):8125-8133.
- Nguyen MV, Lin CH, Syariz MA, Le H, Blanco AC. 2021. Multi-task convolution neural network for season-insensitive chlorophyll-A estimation in inland water. *IEEE J Sel Top Appl Earth Obs Remote Sens.* 14:10439-10449.

20. Chen C, Chen Q, Li G, He M, Dong JW, Yan H, *et al.* 2021. A novel multi-source data fusion method based on bayesian inference for accurate estimation of chlorophyll-a concentration over eutrophic lakes. *Environ Model Softw.* 141:105057.
21. Zhou X, Chen J, Rakstad TE, Ploughe M, Tang P. 2021. Water chlorophyll estimation in an urban canal system with high-resolution remote sensing data. *IEEE Geosci Remote Sens Lett.* 18(11):1876–1880.
22. Cuartero A, Cáceres-Merino J, Torrecilla-Pinero JA. 2023. An application of c2-net atmospheric corrections for chlorophyll-A estimation in small reservoirs. *Remote Sensing Applications: Society and Environment.* 32:101021.
23. Stepanova IE. 2021. Analysis of the correlation between the phytoplankton development level, assessed by chlorophyll a, and nutrient concentrations in rybinsk reservoir. *Biol Bull Russ Acad Sci.* 48(2):193–198.
24. Hu C, Feng L, Guan Q. 2020. A machine learning approach to estimate surface chlorophyll a concentrations in global oceans from satellite measurements. *IEEE Trans Geosci Remote Sens.* 59:4590-4607.
25. Asim M, Brekke C, Mahmood A, Eltoft T, Reigstad M. 2021. Improving chlorophyll-A estimation from sentinel-2 (MSI) in the barents sea using machine learning. *IEEE J Sel Top Appl Earth Obs Remote Sens.* 14:5529-5549.
26. Ding W, Zhang C, Shang S, Li X. 2022. Optimization of deep learning model for coastal chlorophyll a dynamic forecast. *Ecol Modell.* 467:109913.
27. Karimian H, Huang J, Chen Y, Wang Z, Huang J. 2023. A novel framework to predict chlorophyll-a concentrations in water bodies through multi-source big data and machine learning algorithms. *Environ Sci Pollut Res.* 30(32):79402–79422.
28. Kallio K, Malve O, Siivola E, Kervinen M, Koponen S, Lepistö A, *et al.* 2023. Spatiotemporal analysis of lake chlorophyll-a with combined in situ and satellite data. *Environ Monit Assess.* 195(4):465.
29. Cao Z, Ma R, Duan H, Pahlevan N, Melack J, Shen M. 2020. A machine learning approach to estimate chlorophyll-a from landsat-8 measurements in inland lakes. *Remote Sens Environ.* 248:111974.
30. Xin H, Li Y, Li X, Xu M, Sun L. 2022. Estimation of chlorophyll-a concentration in Lake Taihu from Gaofen-1 wide-field-of-view data through a machine learning trained algorithm. *J Meteorol Res.* 36(1):208-226.
31. Hadid NB, Goyet C, Chaar H, Maiz NB, Guglielmi V, Shili A. 2021. Machine learning modeling techniques for forecasting the trophic level in a restored south mediterranean lagoon using chlorophyll-a. *Wetlands.* 41(8):111.
32. Ocean Data Viewer. a.n.d. Data.unep-Wcmc.org. Accessed August 18, 2023. <https://data.unep-wcmc.org/datasets/37>.