Article

Machine Learning-Based Framework to Predict the Combined Effects of Climate Change and Floating Photovoltaic Systems Installation on Water Quality of Open-Water Lakes

Nagavinothini Ravichandran [1],* and Balamurugan Paneerselvam [2]

1 Center of Excellence in Interdisciplinary Research for Sustainable Development, Faculty of Engineering, Chulalongkorn University, Bangkok 10330, Thailand

2 Department of Community Medicine, Saveetha Medical College, Saveetha Institute of Medical and Technical Sciences (SIMATS), Chennai 602105, India; balamuruganp.smc@saveetha.com

* Correspondence: vino.civil35@gmail.com or nagavinothini.r@chula.ac.th

Abstract: Floating photovoltaic (FPV) systems represent a promising advancement in renewable energy technology; however, a comprehensive understanding of their environmental impacts is essential. The effects of FPV installation on lake water temperature remain unclear, potentially hindering the development of the technology due to associated negative implications for aquatic ecosystems. Furthermore, the rise in water temperature associated with climate change poses additional threats to open-water bodies. In this context, the current study endeavors to develop a machine learning (ML)-based framework to assess the combined impact of climate change and the installation of FPV systems on the water quality of open-water lakes. This framework involves the creation of three predictive models and a forecasting model utilizing various ML algorithms, concentrating on temperature and water quality predictions. The framework was applied to a case study assessing the impact of installing three distinct FPV systems on the water quality of Oostvoornse Lake in the Netherlands, employing water quality data available in the literature. The findings indicate a temporal increase in both air and water temperatures at the site, underscoring the ramifications of climate change. Additionally, the results suggest that FPV installations can influence lake thermal dynamics, leading to variations in water temperature and dissolved oxygen concentration, which presents both opportunities and challenges in addressing the impacts of climate change. The proposed framework will be an effective tool for evaluating the effects of FPV systems on water quality throughout their operational lifespan while addressing significant climate change issues.

Keywords: climate change; dissolved oxygen; floating photovoltaic systems; lake water quality; machine learning; temperature

1. Introduction

A significant increase in renewable energy production through advanced technologies is essential to effectively reduce carbon emissions in the near future. One such promising technology that can play a crucial role in achieving global energy objectives is solar photovoltaics. In recent years, the installation of floating photovoltaic (FPV) systems in inland waters and offshore locations has increased considerably. This trend is mainly due to the limitations associated with land-based photovoltaic (PV) systems in countries with dense populations and the cooling effect provided by water surfaces, which helps improve the efficiency of PV panels and reduces water evaporation through shading [1]. Since the

development of FPV technology, substantial economic and technical advancements have occurred in the renewable energy sector. Sustainability emerges as a critical element in the discussion surrounding FPV systems and their impact on water quality. By utilizing underused water bodies, FPV systems can help combat climate change by reducing greenhouse gas emissions associated with fossil fuel energy production. This aligns with broader sustainability goals, including the reduction of environmental footprints and the promotion of clean energy technologies. Additionally, FPV systems may present opportunities for sustainable urban development. In urban areas where land competition is high, deploying FPV systems on bodies of water can conserve valuable land resources while contributing to cleaner energy production. This innovative approach allows for the coexistence of energy generation and the maintenance of aquatic ecosystems, promoting broader sustainability objectives such as biodiversity conservation and ecosystem resilience.

However, several aspects of FPV technology and its environmental impact still require thorough assessment. One significant concern is the impact of FPV systems on water quality and the surrounding aquatic ecosystems when installed in open-water lakes [2,3]. The unresolved effects of deploying FPV systems in these environments raise serious concerns about potential harmful impacts [3]. While there is literature discussing the benefits of FPV systems in relation to the energy–water–food nexus, there is still a lack of empirical evidence regarding their effects on water quality [4].

Research on the impact of FPV systems on water quality has identified light availability, water temperature, and dissolved oxygen concentration as the three primary factors integral to water quality assessment [5–8]. Notably, water temperature exhibits a strong correlation with various climatic factors and parameters related to FPV systems, such as size and water coverage. It serves as a critical parameter for investigating the physical and chemical properties of water bodies [9]. Consequently, it is essential to conduct site-specific and system-specific studies to accurately determine the variations in water temperature attributable to FPV systems [10,11]. Moreover, fluctuations in wind speed, light availability, and water temperature lead to corresponding changes in dissolved oxygen (DO) concentration in lake water. Such alterations can profoundly affect aquatic life and may disrupt the associated food chain [12]. Studies have also pointed out the beneficial effects of FPV installations, including the management of excessive algal growth, reduction of eutrophication, and enhancement of DO levels [2]. Nonetheless, the overall impact of the competing heat transfer mechanisms in lakes equipped with FPV systems remains uncertain due to the decreased or increased upper surface water temperature in different seasons [5,13,14]. Existing literature also indicates that larger-scale FPV systems, noted for their expansive coverage, possess the capability to significantly mitigate algal blooms. However, it is also important to recognize that these systems may lead to substantial economic losses in hydropower generation by restricting the operational capacity of hydropower plants [2,15]. The broader implications of FPV systems on water quality present a nuanced perspective, with conclusions varying significantly based on factors such as regional context, coverage of FPV systems, and seasonal variations [5,13]. Moreover, the installation of FPV systems contributes to a reduction in wind stress and the creation of a microclimate in the airspace between the panels and the water surface. Consequently, the net effects of shading and the microclimate layer on water temperature and quality, as analyzed through detailed three-dimensional modeling approaches documented in the literature, also differ based on the assumptions employed in the model development [10,11].

Another important issue that has a striking consequence on the nutrient recycling and biodiversity of lake ecosystems is climate change, which causes long-term shifts in temperatures and weather patterns. In recent years, the adverse effects of climate change have contributed to increased degradation of natural ecosystems and disruption of var-

ious economic sectors [16]. Several major consequences of climate change are directly associated with the rise in global temperatures. The Intergovernmental Panel on Climate Change (IPCC) projects that global temperatures may increase by 2 to 4 °C by the end of the 21st century, contingent upon socio-economic development pathways and greenhouse gas emissions [17]. In the case of open-water lakes, water temperature serves as a crucial environmental factor and indicator of climate change and human activities [18]. Alterations to lake water temperature resulting from external sources can profoundly affect the physical, biological, and chemical processes within the lake's ecological environment. Specifically, the average lake water temperature in numerous lakes globally has risen by approximately 0.34 °C per decade [19]. This increase in temperature is likely to exacerbate ecological issues, including heightened eutrophication and an extended growth period for cyanobacteria, in addition to environmental challenges such as increased thermocline depth and strength, enhanced hypoxia at the bottom of lakes, and a prolonged period of lake thermal stratification. Furthermore, research has indicated that an elevation in lake surface water temperature by 0.1 °C per decade may destabilize lake ecosystems, reducing fish yields by as much as 30% [20]. Hence, region-specific detailed assessments are necessary to understand the variations in water temperature due to climate change.

Developing numerical models that accurately reflect water temperature and energy balance while considering varying meteorological conditions is a complex task. Consequently, recent investigations have concentrated on analyzing long-term measured data collected at specific sites to elucidate the effects of FPV systems on water temperature [5,21]. With the availability of such extensive datasets, the application of data-driven artificial intelligence tools will facilitate a more effective exploration of the relationship between water temperature and FPV systems. Given the anticipated increase in the number of water bodies hosting FPV systems, there is an urgent necessity to comprehend their impact on water quality through precise modeling.

The utilization of machine learning within civil engineering has surged in recent years, particularly in forecasting climate change impacts, which aids in conducting comprehensive environmental impact assessments of structures. The application of machine learning to assess the impact of FPV systems will enhance the understanding of the intricate, interconnected effects of these systems on open-water lakes. The appropriate integration of field monitoring data with historical data via a machine learning approach promises to yield valuable insights into FPV systems in the future. In this context, the present study proposes a machine learning-based framework to assess the combined effects of FPV system installation and climate change on the water quality of open-water lakes, a topic that has yet to be explored in the existing literature. Additionally, the proposed framework has been applied to a case study to validate its effectiveness in impact assessment and to demonstrate the promising results that can be achieved through its proper implementation.

The objective of the proposed framework is to establish a practical methodology for assessing water quality utilizing minimal available data. This study does not aim to evaluate the advantages or disadvantages of installing FPV systems. The results derived from this framework will be highly specific to individual cases, and the outcomes of the case study should not be interpreted as representative of general conditions applicable to all open-water lakes. The primary focus of the current investigation is developing a machine learning-based framework that facilitates researchers in addressing the environmental impacts of FPV systems. The innovative aspect of this study lies in applying machine learning models for framework development to evaluate the effects of climate change and the implementation of FPV systems on lake ecosystems.

The organization of the paper is as follows: Section 2 discusses the machine learning framework developed for the combined impact analysis of climate change and FPV imple-

mentation on water quality. Section 3 introduces the specifics of the case study conducted for framework implementation. Lastly, Section 4 presents the results of the case study, highlighting the most effective machine learning model and the detailed impact analysis.

2. Materials and Methods

2.1. Proposed Machine Learning Framework

The developed framework for combined impact analysis, as illustrated in Figure 1, utilizes water quality data collected from pilot projects or existing FPV systems at the site of interest, along with the historical temperature data. This framework comprises the creation of four machine learning models: (i) the FPV water temperature model, (ii) the dissolved oxygen model, (iii) the temperature forecasting model, and (iv) the air–water temperature model.
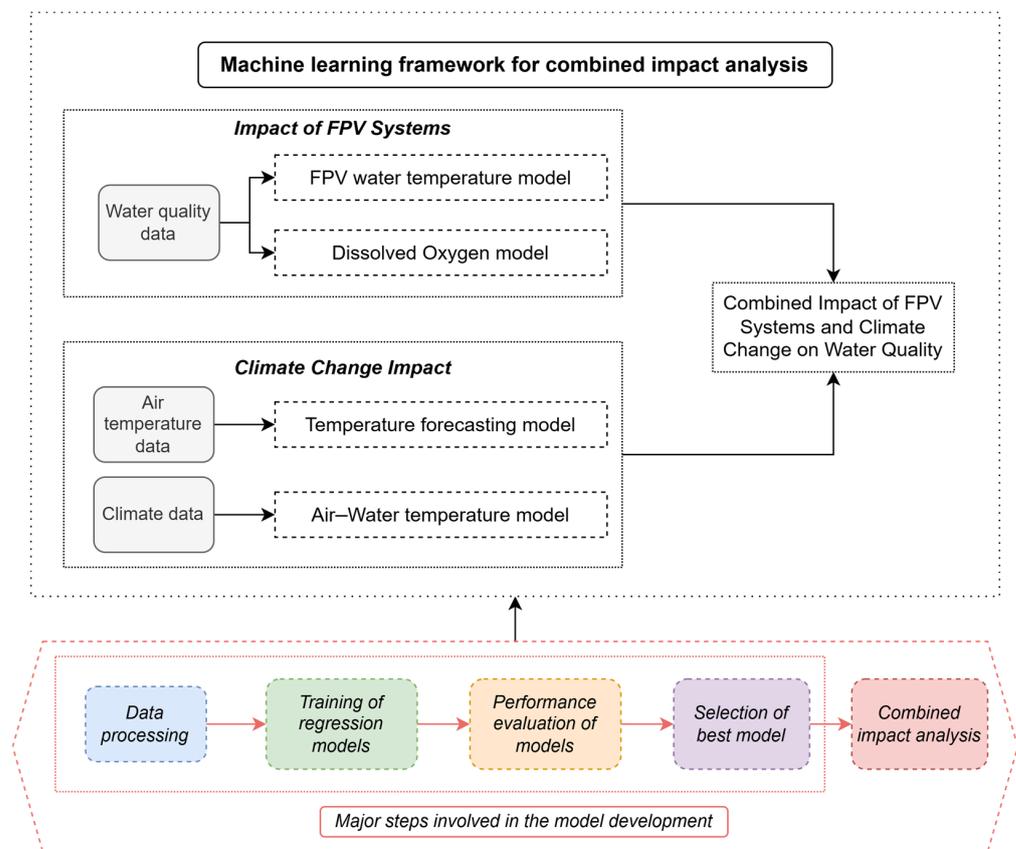
Figure 1. Machine learning-based framework for prediction of water quality affected by FPV systems and climate change.

In each model incorporated within the framework, a systematic approach is employed, consisting of four critical steps: data processing, training of machine learning models, performance evaluation of these models, and the selection of the most effective model. The concluding step of the framework involves the integration of all four developed models to ascertain the forecasted water quality parameters, thereby delineating the combined effects of the deployment of FPV systems and climate change. By enhancing the understanding and management of water quality through these predictive models, the framework supports environmentally responsible use of resources and promotes sustainable practices in the installation of the FPV system, ensuring that advancements in renewable energy do not compromise the ecological integrity of water bodies. This section provides a comprehensive overview of the principal steps involved in the framework's development.

2.2. Data Processing

The framework developed for assessing water quality necessitates acquiring three critical datasets: water quality data, historical temperature records, and climate conditions. The FPV water temperature and dissolved oxygen models were established using water quality data collected from the study site. To effectively evaluate the impact of FPV system deployment on the water quality of open-water lakes, it is essential to consider several key parameters: the water temperature in open water, the water temperature beneath the FPV systems, the area of FPV systems, the extent of water coverage, and the concentrations of dissolved oxygen [3,5]. For the temperature forecasting model, it is imperative to obtain historical temperature data pertinent to the selected location. Moreover, comprehensive climate data (including wind conditions, heat fluxes, and air and water temperature) must be sourced consistently to facilitate the development of the air–water temperature model [20,22,23]. This process requires integrating site-specific water temperature data with relevant climate data acquired from various databases. It is crucial to emphasize that a well-structured dataset is vital for successfully developing machine learning models. Consequently, meticulous data acquisition and thorough data preprocessing are essential for conducting a proper impact analysis. The steps involved in data processing are illustrated in Figure 2.
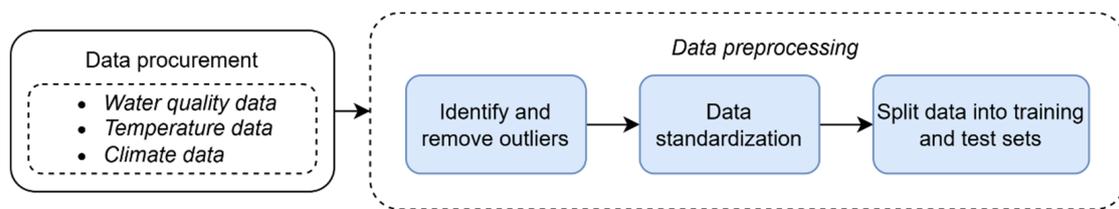
Figure 2. Steps involved in data procurement and preprocessing.

Date preprocessing is a critical step in enhancing the accuracy of machine learning models. Data collected often contain outliers, which may arise from monitoring errors, misinterpretation, or malfunctions in measuring equipment. It is essential to eliminate such anomalous data prior to model training to improve predictive accuracy. After removing erroneous entries from the original dataset, the next step involves standardizing the dataset. Standardization entails rescaling the data to achieve a mean of zero and a standard deviation of one. This process diminishes the influence of outliers and facilitates the assessment of feature importance. Subsequently, the processed data are partitioned into two sets: (i) the training set and (ii) the test set. Typically, 70% of the dataset is allocated for training the models, while the remaining 30% is reserved for evaluating the performance of the developed model utilizing standard performance metrics.

2.3. Training of Regression Models

Machine learning (ML) regression models leverage statistical analysis of the provided data to inform model training based on historical experiences [24]. Accordingly, ML models require data on predictor and response variables to facilitate effective model implementation and future predictions. During the training phase, the ML models undergo a process wherein they are trained using the dataset to optimize the cost function by appropriately adjusting model parameters. The regression models offer significant advantages in elucidating the relationships between predictor or input variables and response or output variables.

In the current study, regression models are formulated as a component of the framework for predicting water temperature in floating photovoltaic (FPV) systems, dissolved oxygen concentrations, lake water temperature, and general temperature forecasting.

Figure 2 illustrates the various models incorporated within the framework, along with the associated predictor and response variables relevant to the development of ML models.

The FPV water temperature model was developed to predict the water temperature beneath the installed FPV system (output) based on several input variables, including the area of the FPV system, water coverage, and the temperature of the uncovered water. Although the existing literature provides limited studies on water quality, the model currently incorporates only three primary input parameters. Nonetheless, additional water quality monitoring data may be integrated through appropriate data preprocessing techniques to enhance the performance of the model. The general predictive equation for the FPV water temperature models is as follows:

$$Y_{FPV\_T} = f(X_A, X_{WC}, X_{UWT}) \tag{1}$$

where Y_{FPV_T} is the water temperature under the FPV system in °C, X_A is the area of the FPV system in m^2, X_{WC} is the water coverage ratio, and X_{UWT} is the uncovered water temperature in °C. It is also important to mention that supplementary meteorological data can be included as input variables for the FPV water temperature model. Notably, wind speed, which influences the energy transfer processes between the atmosphere and the lake, represents a significant parameter for inclusion in the model [25]. Based on the available data, the model can be designed to accommodate hourly or daily average data, provided that all input and output variables are sampled over the same time period.

The dissolved oxygen model has been developed to ascertain the concentrations of dissolved oxygen (output) based on water temperature (input). The database may encompass results from both FPV-covered and uncovered locations, provided that the data are sourced from the same site utilized for the FPV water temperature model. The general predictive equation for the dissolved oxygen model is as follows:

$$Y_{DO} = f(X_T) \tag{2}$$

where Y_{DO} is the dissolved oxygen concentration in mg/L, and X_T is the water temperature in °C. When available, additional parameters such as anthropogenic nutrient inputs, wind speed, and rainfall, which significantly influence dissolved oxygen concentrations, may be incorporated into this model.

The air–water temperature model has been developed to elucidate the relationship between air temperature and water temperature at the designated location. While historical air temperature data can be readily obtained from various open-source databases, specific water temperature data for the lake may not be accessible. Consequently, water temperature data gathered through monitoring at the site will be integrated with climate data to project future water temperatures of the lake. The primary objective of this model is to predict lake water temperature (output) based on several input variables, including air temperature, wind speed, skin temperature of the earth, shortwave downward irradiance, and longwave downward irradiance at the selected location. The general predictive equation for the air–water temperature model is as follows:

$$Y_{WT} = f(X_{AT}, X_{EST}, X_{WS}, X_{SDI}, X_{LDI}) \tag{3}$$

where Y_{WT} is the water temperature in °C, X_{AT} is the air temperature in °C, X_{EST} is the earth's skin temperature in °C, X_{WS} is the wind speed in m/s, X_{SDI} is the shortwave downward irradiance in kW-h/m^2/day, and X_{LDI} is the longwave downward irradiance in kW-h/m^2/day.

The temperature forecasting model implemented within the current framework necessitates access to historical air temperature data pertinent to the selected location. Such data can be obtained from established databases, and it is crucial to utilize a comprehensive dataset to enable accurate forecasting. In the development of the forecasting model, the response variable within the temperature data will be established, and the predictor variable will be generated by temporally shifting the response variable backward. Hence, the general predictive equation for the temperature forecasting model is as follows:

$$Y_{WT+1} = f(X_{WT}) \tag{4}$$

where Y_{WT+1} and X_{WT} are the air temperature in °C at time steps t + 1 and t, respectively.

The models developed in this study can be broadly categorized into two types: prediction models and forecasting models, as illustrated in Figure 3. A range of machine learning regression models was employed for supervised learning utilizing the available output variables. In the realm of prediction modeling, eight regression models were implemented, including linear regression (LR), stepwise linear regression (SLR), ridge regression (RR), decision tree (DT), random forest (RF), support vector machine (SVM), Gaussian process regression (GPR), and artificial neural networks (ANNs). These models were utilized to identify variables such as FPV water temperature, dissolved oxygen concentration, and lake water temperature.
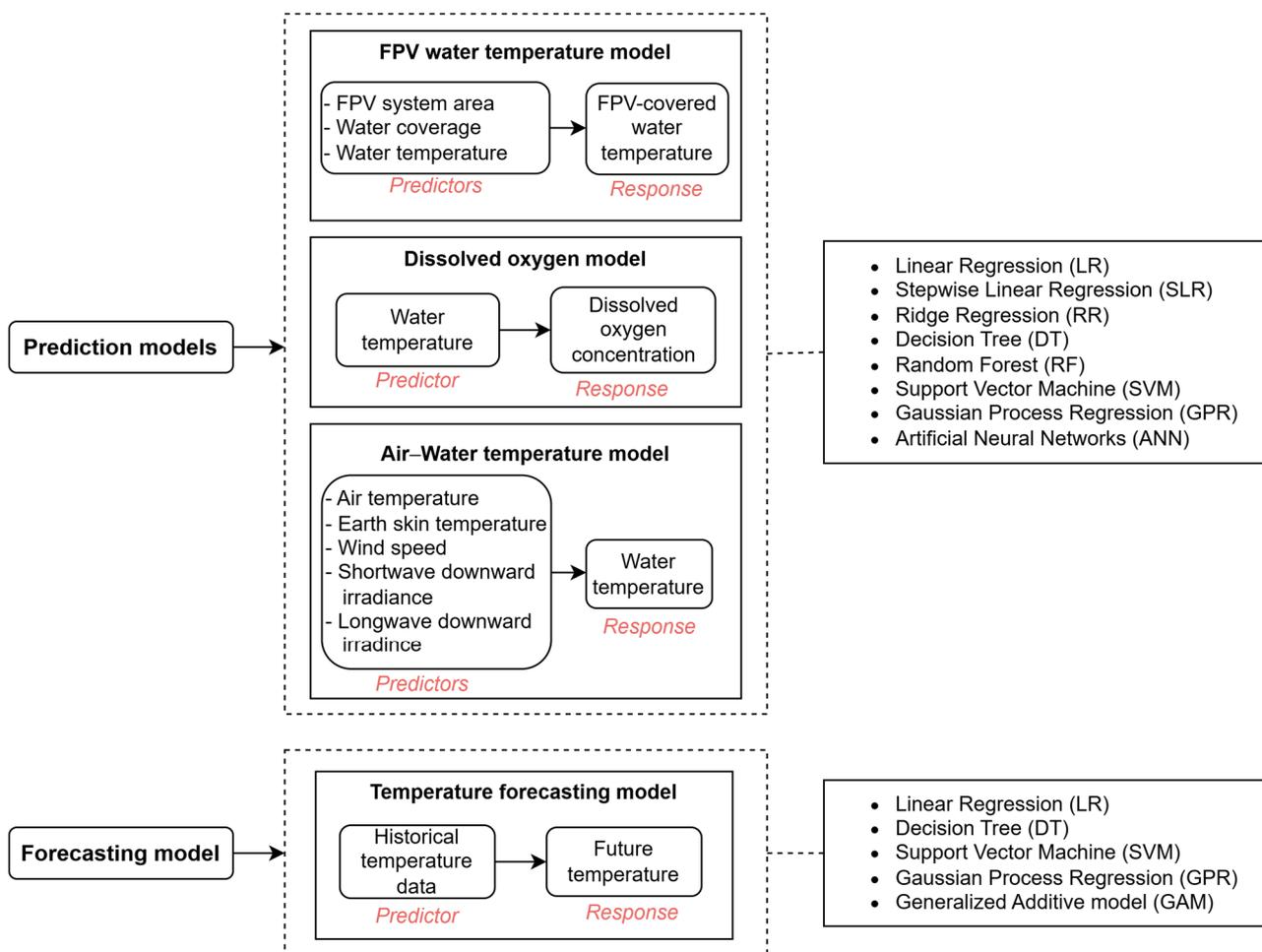
Figure 3. Prediction and forecasting models used in the present study.

LR, SLR, and RR are parametric models that establish a linear relationship between predictor and response variables [26,27]. These linear models were initially utilized due to their interpretative capacity, which facilitates understanding the relationship between the predictor and outcome variables. However, instances may arise where the relationship between these variables is nonlinear, necessitating the use of non-parametric models. The DT model constructs a tree-like structure by recursively partitioning the regression region into a hierarchy of simple decisions until a predetermined stopping criterion is met [26,28]. The optimal split for the model is determined through the minimization of node error, and the high variance associated with the model can be mitigated using either boosting or bagging techniques. The RF method, an ensemble-based approach, comprises several trees derived from a bootstrap sample of the training dataset [29]. The SVM model functions as a binary classifier, employing an optimal hyperplane to distinguish data points; it identifies nonlinear boundaries by transforming features into a higher-dimensional space [30–32]. The GPR model is a kernel-based probabilistic approach, whereby predictions are made by projecting inputs into a p-dimensional feature space [33]. The ANN model, drawing inspiration from human cognitive processes, captures responses through the nonlinear functional relationships of predictor variables [34]. Each model is associated with specific hyperparameters that require optimization.

Several non-linear models were investigated in the present study to elucidate the intricate relationship between the variables, particularly the correlation between dissolved oxygen concentration and water temperature. Furthermore, the study incorporated SVM, RF, DT, and NN models due to their robust adaptability and learning capabilities in accommodating diverse data types. Additionally, the KNN model was employed owing to its effectiveness in handling smaller datasets, which may prove advantageous for the current analysis.

Furthermore, this study implements a multistep forecasting model using a direct strategy for temperature forecasting, where a distinct trained regression model supports each forecasting time step. For the forecasting component, five machine learning models were utilized: linear regression (LR), decision tree (DT), support vector machine (SVM), Gaussian process regression (GPR), and generalized additive model (GAM). Except for GAM, which extends multiple linear models by substituting the linear component with a smooth nonlinear function, the LR, DT, SVM, and GPR models operate by the previously described methodologies [35].

2.4. Performance Evaluation of Models

The performance of each machine learning model can be evaluated using widely accepted performance metrics, including the coefficient of determination (R^2), adjusted R^2, and root mean square error (RMSE) for both the training set and the test set. The coefficient of determination serves as a statistical measure that indicates the extent to which a statistical model can predict an outcome, reflecting the goodness of fit. It represents the proportion of variance in the response variable that is accounted for by the model. The R^2 value ranges from 0 to 1, where 0 denotes that the model fails to predict the outcome, while 1 signifies perfect prediction. It can be obtained as follows:

$$R^2 = 1 - \frac{SSE}{SST} \tag{5}$$

where SSE is the sum of squared error, and SST is the sum of squared total. The R^2 value increases with the increase in predictor variables in the regression model. The adjusted R^2 value considers the number of predictor variables in the regression model, rendering it

more applicable for models with a large number of predictors. The adjusted R^2 value can be calculated as follows:

$$R^2_{adj} = 1 - \left(\frac{n-1}{n-p}\right) - \frac{SSE}{SST} \qquad (6)$$

where n is the number of observations, and p is the number of regression coefficients. Root mean square error (RMSE) quantifies the average discrepancy between the predicted and actual values generated by the statistical model. It is the standard deviation of the residuals, and it can be calculated as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}|Y_A - Y_P|^2} \qquad (7)$$

where n is the number of observations, Y_A is the actual value of the response variable, and Y_P is the predicted value of the response variable. The RMSE value ranges from 0 to positive infinity and is expressed in the same units as the response variable. An RMSE value of 0 indicates a perfect match between the predicted and actual values, whereas higher RMSE values imply greater error and less accurate predictions.

Consequently, smaller RMSE values and higher R^2 values indicate superior model performance. Such models are typically selected for further combined impact analysis. Given that the accuracy of predictions relies heavily on model performance, a judicious selection of models based on evaluation metrics is imperative. Additionally, careful optimization of hyperparameters associated with each model is essential.

2.5. Combined Impact Analysis

Following the selection of optimal prediction and forecasting models, these models were employed to evaluate the combined effects of temperature increase and the installation of FPV systems in open-water lakes over their operational lifespan. Figure 4 illustrates the concluding steps involved in this combined impact analysis.
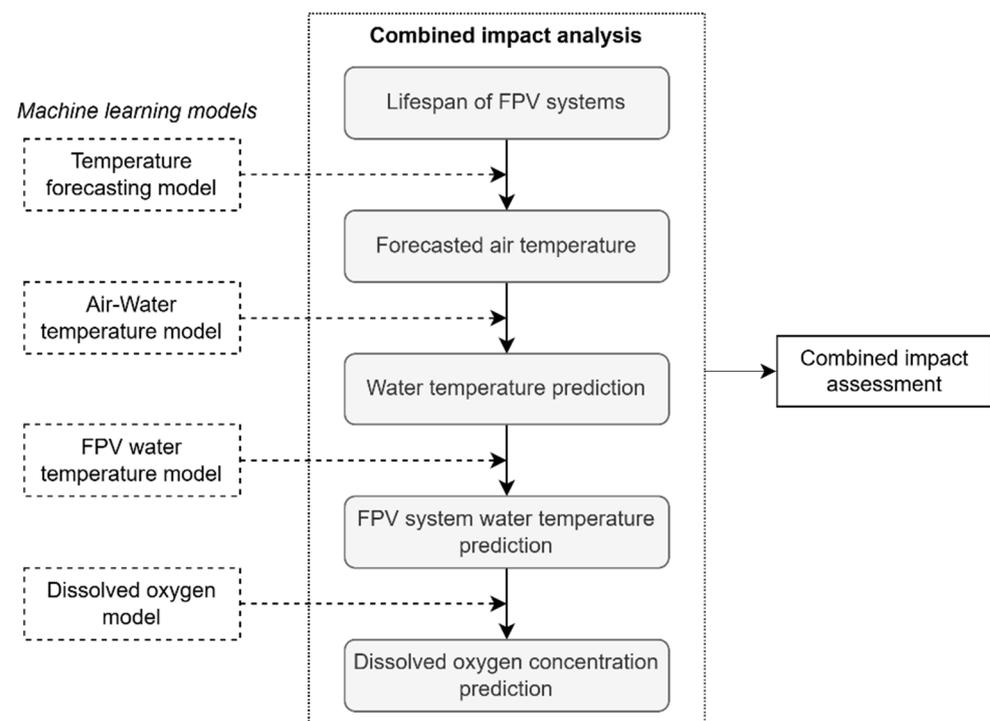
Figure 4. Combined impact analysis from the developed ML models.

The lifespan of FPV systems installed in recent years ranges from 20 to 30 years [8,36]. Therefore, it is essential to determine the operational lifespan of the proposed FPV system at the selected site to calculate the forecasted temperature and water quality parameters initially. Subsequently, utilizing the temperature forecasting model, the air temperature at the designated location was predicted for the lifespan of the FPV systems, thereby indicating the impact of climate change on the area in question. The water temperature was then determined using the air–water temperature model corresponding to the forecasted air temperature. Finally, the FPV water temperature and dissolved oxygen models were applied to ascertain the water temperature under the FPV system and the concentration of dissolved oxygen in the lake water derived from the forecasted lake water temperature.

3. Case Study—Oostvoornse Lake

In the present study, water quality data obtained from Bax et al. [5] were utilized to implement the developed framework aimed at predicting the cumulative effects of rising temperatures and the installation of FPV systems in open waters. The geographical focus of this investigation was Oostvoornse Lake, a brackish lake in the Netherlands, as illustrated in Figure 5. The surface area of the lake encompasses approximately 20 ha, with average and maximum water depths recorded at 20 m and 40 m, respectively. The lake experiences an influx of freshwater from the adjacent dune regions, contributing to increased algal growth. Consequently, supplementary saltwater inflow has been facilitated to the lake through underground pipelines originating from Mississippi Haven in order to preserve the ecological biodiversity of the area.

Figure 5. Location of Oostvoornse Lake and the FPV pilot project.

Bax et al. analyzed water quality data collected over a one-year pilot project, during which three distinct FPV systems were established in the northwestern section of the lake, with respective installed capacities of 41.93 kWp, 39.42 kWp, and 50.7 kWp. Key characteristics of the FPV systems are detailed in Table 1, and the schematic representation of the three FPV systems is shown in Figure 6. Notably, the FPV 3 system occupies a larger area, while the FPV 2 system provides greater water surface coverage than the other systems. The water surface coverage refers to the extent to which the FPV systems fully shade the water surface. Systems with closely arranged floaters typically achieve greater coverage. Continuous measurements were conducted to collect data on light intensity, water temperature, dissolved oxygen concentration, pH levels, turbidity, electrical conductivity, and oxygen reduction potential. Further information regarding the study area, FPV systems, and water quality data is available in reference [5].

Table 1. Features of FPV systems established in Oostvoornse Lake [5].

FPV Systems	Shape	Material	Area (m^2)	Water Coverage (%)
FPV 1	Roughly rectangular	High-density polyethylene tubes connected by aluminum frames	350	75
FPV 2	Rectangular	Polypropylene floaters	400	100
FPV 3	Roughly circular	Metal frames mounted on polypropylene floaters	600	75

(a) (b) (c)
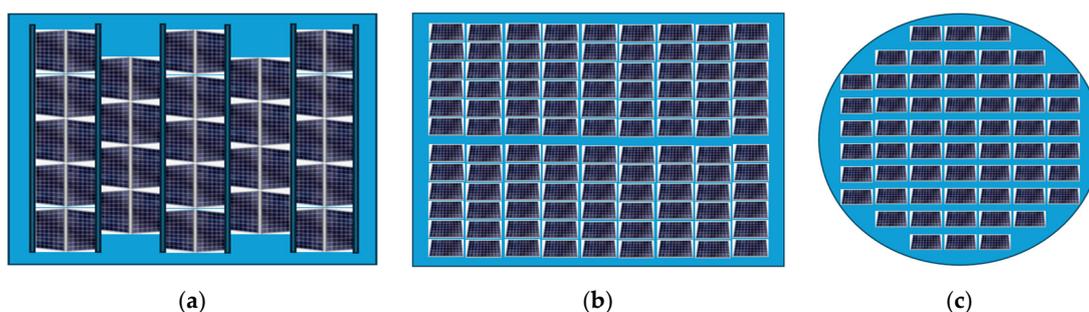
Figure 6. Schematic top view of the FPV systems at Oostvoornse Lake: (**a**) FPV 1, (**b**) FPV 2, and (**c**) FPV 3.

This particular study [5] was selected to implement the proposed framework due to the availability of field monitoring data encompassing various FPV systems. The findings indicated that the shading effects resulting from the FPV systems are minimal, attributable to their relatively small size in comparison to the expanse of the lake. Additionally, no significant adverse impacts on water quality were identified, nor were there consistent trends indicating increases or decreases in water temperature or dissolved oxygen concentration. The variations noted in water temperature and dissolved oxygen levels were slight and tended to diverge from the reference measurements. The limited impact on water quality caused by the FPV systems may be explained by the relatively high water flow velocities, currents, and extensive water mixing within the lake, as noted in reference [5]. These data are utilized in the current study to underscore the effectiveness of the proposed framework, and the minimal impact on water quality does not impede the capability of the framework to determine the combined effects of climate change and water quality. The appropriate application of temperature and climate data through the proposed framework is expected to yield reliable outcomes, irrespective of the direct effects associated with the monitored data.

3.1. Date Procurement and Processing

The four distinct datasets required for developing the ML models were sourced from existing literature and widely recognized databases. The water quality data pertinent to the FPV system temperature model and the dissolved oxygen model encompassed variables such as water coverage of FPV systems, area of FPV systems, dissolved oxygen concentration, open-water temperature, and FPV system temperature. These data were obtained from the study conducted by Bax et al. [5]. The field monitoring data, encompassing the summer and fall seasons, were available from March 2021 to October 2021 for every 30 min. However, technical complications resulted in several gaps within the reported data, and various extreme events hindered the accurate measurement of water quality parameters. Consequently, outliers within the dataset were eliminated during the data preprocessing stage by data visualization and interquartile range methods to enhance the fitting of machine learning models. The field-monitored data were utilized directly for the

temperature model of the FPV system, while the daily average values were employed for the dissolved oxygen model.

Historical temperature data for the temperature forecasting model were acquired from the National Aeronautics and Space Administration (NASA) Langley Research Center (LaRC) through the Prediction of Worldwide Energy Resource (POWER) Project, which is funded by the NASA Earth Science/Applied Science Program [37]. The monthly air temperature data span the period from 1981 to 2022.

Moreover, the climate data necessary for the air–water temperature model were also sourced from the NASA POWER project. This dataset included daily average air temperature, wind speed, earth's skin temperature, shortwave downward irradiance, and longwave downward irradiance relevant to the same timeframe as the water quality data (March to October 2021). In this manner, the water temperature data from the study by Bax et al. [5] were integrated with the climate data for the same period to facilitate the development of the air–water temperature model. As delineated in the framework, the collected datasets were subsequently organized as predictor and response variables essential for developing the models needed for impact analysis. The outliers, identified as erroneous data points, were eliminated from the datasets, followed by the standardization process in the data preprocessing stage. It is essential to emphasize that the varied time intervals associated with the data utilized in the development of the models for this study will not impact the efficacy of the ML models. This is due to the fact that the time interval or frequency of the data points is not included as a factor or input in the model development process itself. Consequently, the performance of the models remains unaffected by the timing of the data collection, allowing for a more robust and consistent evaluation of their capabilities.

3.2. Development of Machine Learning Models

As illustrated in Figure 3, various ML algorithms were employed to ascertain the most suitable model for the development of prediction and forecasting frameworks. The ML models were developed using MATLAB R2024b. This section delineates the model parameters and optimized hyperparameters and identifies the most effective model based on a comparative analysis of performance metrics.

3.2.1. FPV System Water Temperature Model

A model for predicting water temperature beneath FPV systems was developed to ascertain water temperature at a depth of 2.0 m based on the specific type of FPV system and the ambient open-water temperature, measured every 30 min. The processed dataset, which includes three predictor variables and one response variable, comprises 12,000 data points. The dataset was standardized and subsequently partitioned into training (70% of the total data) and test (30% of the total data) sets. Eight ML models were trained using the training set, and the outcomes were assessed against both the training and test sets.

The LR and SLR models were developed utilizing a quadratic character vector. The RR model was formulated with a ridge parameter value of 2. The DT model was developed using a minimum of 27 leaf node observations and 1120 decision splits. The RF model was designed employing the bagging method, incorporating 31 learning cycles, a minimum leaf size of 2, and allowing for a maximum of 8155 splits. The SVM model was implemented with a Gaussian kernel function, characterized by a kernel scale value of 2.073, a box constraint of 9.510, and an epsilon value of 0.096. The GPR model was established utilizing an exponential kernel function with a sigma value of 3. The ANN model was created with seven layers, employing a sigmoid activation function and a lambda value of 1.895×10^{-9}. The results obtained from the ML models are presented in Table 2.

Table 2. Performance metrics of FPV system temperature models.

Model	R^2		Adjusted R^2		RMSE (°C)	
	Training Set	**Test Set**	**Training Set**	**Test Set**	**Training Set**	**Test Set**
LR	0.999	0.999	0.999	0.999	0.137	0.150
SLR	0.999	0.999	0.999	0.999	0.136	0.149
RR	0.999	0.999	0.999	0.999	0.137	0.150
DT	0.999	0.999	0.999	0.999	0.136	0.157
RF	0.999	0.999	0.999	0.999	0.116	0.154
SVM	0.999	0.999	0.999	0.999	0.136	0.149
GPR	0.999	0.999	0.999	0.999	0.123	0.150
ANN	0.999	0.999	0.999	0.999	0.135	0.148

The analysis of the results revealed that the R^2 and adjusted R^2 values were consistently observed to be equal at 0.999 across all algorithms. Notably, there was a slight variation in *RMSE* values. The RF model exhibited the lowest *RMSE* value of 0.116 °C within the training set, followed by the GPR and ANN models. However, when applied to the test set, the *RMSE* value for the RF model was marginally higher than that of the other regression models. In contrast, the ANN model demonstrated the lowest *RMSE* value of 0.148 °C in the test set.

It is essential to recognize that predicting extreme temperature variations in the future entails temperature values that may surpass the range of data present in the current dataset. Given the significance of accurately forecasting extreme temperature values, cross-validation tests employing ten folds were conducted utilizing the superior ML models, specifically RR, GPR, and ANN. The results indicated that the cross-validation error of the ANN model was lower than that of the other models, suggesting superior accuracy in predicting the water temperature of FPV systems under extreme conditions. Consequently, the ANN model was designated as the optimal choice for water temperature prediction beneath the FPV systems. A comparison of the actual versus predicted outcomes of the ANN model is illustrated in Figure 7a.

(a)

(b)
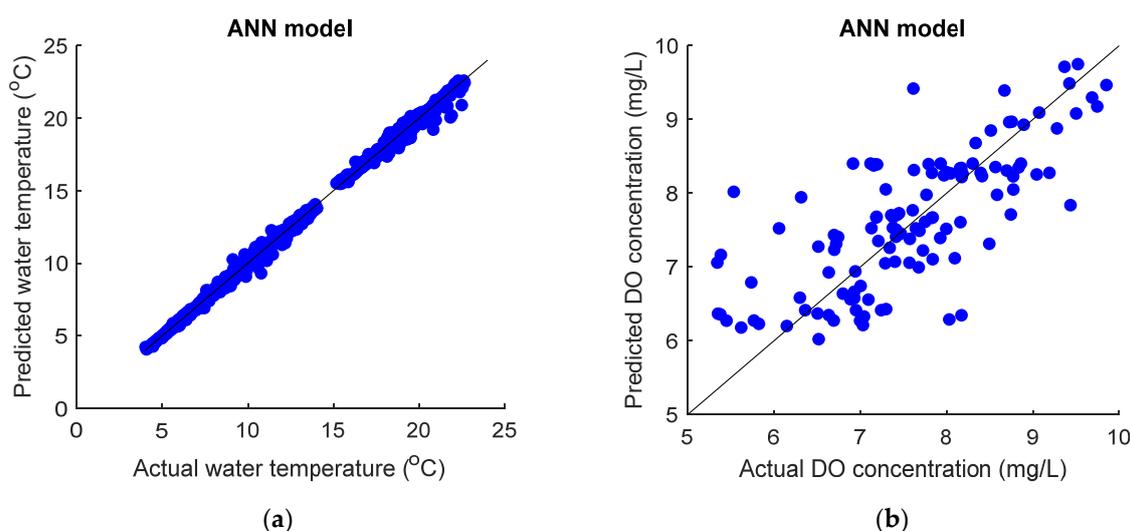
Figure 7. *Cont.*
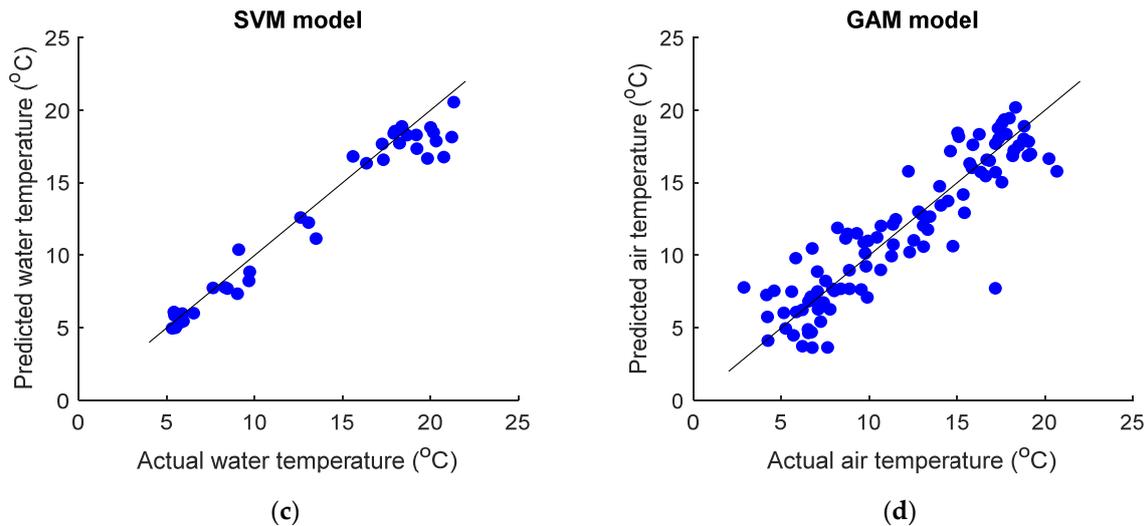
Figure 7. Actual and predicted values of response variables from the best model: (**a**) FPV system water temperature model, (**b**) dissolved oxygen model, (**c**) air–water temperature model, and (**d**) temperature forecasting model.

3.2.2. Dissolved Oxygen Model

A dataset consisting of one-day average values for water temperature and DO concentration was utilized for the dissolved oxygen model, encompassing 400 data points. It is essential to emphasize that utilizing input and output data recorded at identical time stamps is crucial for developing ML models. Given the discrepancies in the time stamps associated with the monitored lake water temperature and DO concentration, daily average values are employed in the dissolved oxygen model. Furthermore, in instances where the field-monitored data for both water temperature and dissolved oxygen concentration are collected concurrently and at uniform time intervals, such data can be directly employed for future model development, thereby enhancing the precision and applicability of ML applications. The water quality data collected from open-water areas and regions occupied by FPV systems served as the foundation for developing the ML models. Eight regression models were constructed using the pre-processed dataset, each featuring one predictor variable and one response variable.

The LR, SLR, and RR model parameters are similar to those applied in the previously referenced FPV system water temperature model. The DT model was configured with a minimum leaf size of 16 and included 83 decision splits. The RF model was established employing the bagging method, consisting of 12 learning cycles, a maximum of 27 splits, and a leaf size set at 10. The SVM model was formulated utilizing a Gaussian kernel function with a kernel scale of 3.547. The GPR model was designed with a rational quadratic kernel function and a sigma value of 0.258. The ANN model was constructed using a sigmoid activation function and comprised 25 layers. The performance metrics of the various ML algorithms are presented in Table 3.

In contrast to the FPV system water temperature model, the performance of various ML algorithms for predicting DO concentrations was observed to be less accurate, characterized by higher $RMSE$ values and lower R^2 values. The results revealed that the GPR model demonstrated higher R^2 and reduced $RMSE$ values within the training dataset than all other models. However, the GPR model exhibited suboptimal performance in the test dataset, suggesting an occurrence of overfitting. Following the GPR model, the RF model achieved lower $RMSE$ values in the training dataset. Moreover, the DT and ANN models exhibited enhanced performance compared to linear models.

Table 3. Performance metrics of dissolved oxygen models.

Model	R^2		Adjusted R^2		RMSE (mg/L)	
	Training Set	Test Set	Training Set	Test Set	Training Set	Test Set
LR	0.483	0.224	0.469	0.176	0.898	0.963
SLR	0.483	0.224	0.469	0.176	0.898	0.963
RR	0.342	0.243	0.325	0.196	0.964	1.080
DT	0.646	0.348	0.637	0.307	0.743	0.883
RF	0.679	0.373	0.671	0.334	0.708	0.866
SVM	0.528	0.595	0.516	0.57	0.817	0.790
GPR	0.992	0.539	0.991	0.51	0.109	0.843
ANN	0.585	0.392	0.574	0.354	0.805	0.853

To comprehensively assess the model performance with novel data, a 10-fold cross-validation test was conducted using the developed models. The findings indicated that the ANN model resulted in a lower error rate, thereby underscoring its superior performance. As a result, the ANN model was designated as the most appropriate model for the current analysis. It is also imperative to note that the inadequacy in the performance of all the ML models in predicting DO concentrations can be attributed to the limited number of data points available for training. A comparative analysis of the actual and predicted responses of the ANN model is illustrated in Figure 7b.

3.2.3. Air–Water Temperature Model

The air–water temperature model was developed to ascertain water temperature at a depth of 2.0 m based on air temperature and climate data. Similar to the dissolved oxygen model, the present study utilized daily average values for both predictors and responses. Consequently, a total of 126 data points were employed for model development. The ML models were trained utilizing five predictor variables alongside one response variable.

A quadratic character vector was applied to the LR and SLR models. The RR model was constructed with a ridge parameter value of 2. The DT model included 13 decision splits and maintained a minimum leaf size of 2. The RF model was created using the LSBoost method, comprising ten learning cycles and a leaf size of 6. The SVM model was established with a Gaussian kernel function, with a kernel scale of 15.261, an epsilon of 0.01061, and a box constraint of 889.2. The GPR model employed a squared exponential kernel function. Additionally, the ANN model utilized two layers with a sigmoid activation function. The results of the performance metrics are detailed in Table 4.

Table 4. Performance metrics of air–water temperature models.

Model	R^2		Adjusted R^2		RMSE (°C)	
	Training Set	Test Set	Training Set	Test Set	Training Set	Test Set
LR	0.977	0.918	0.974	0.898	0.842	1.327
SLR	0.975	0.937	0.973	0.922	0.868	1.161
RR	0.959	0.936	0.956	0.921	1.025	1.458
DT	0.988	0.843	0.987	0.805	0.600	1.836
RF	0.998	0.838	0.997	0.799	0.274	1.862
SVM	0.981	0.948	0.980	0.935	0.865	1.316
GPR	0.975	0.941	0.973	0.927	0.798	1.397
NN	0.974	0.955	0.972	0.945	0.888	1.218

The results indicate that the RF model demonstrated superior performance with higher R^2 and lower *RMSE* values when evaluated on the training set. However, the performance of the model on the test set did not surpass that of other ML models. Following the RF model, the SVM, GPR, and ANN models displayed more favorable performance outcomes. When considering the performances of the models across both training and test sets, the SVM model emerged as the most appropriate choice, characterized by minimal overfitting. Consequently, the SVM model was selected as the optimal model for predicting the water temperature of open lakes. A comparison between the actual and predicted responses of the SVM model is illustrated in Figure 7c.

3.2.4. Temperature Forecasting Model

The temperature forecasting model was developed utilizing historical air temperature data from the site from 1980 to 2022, comprising a total of 504 data points. This forecasting model was employed to determine the minimum, mean, and maximum temperatures over the projected lifespan of the FPV systems, which is set at 30 years. As indicated in the case study, water quality monitoring for the FPV pilot project took place in 2022. Accordingly, the temperature data were forecasted until 2052 to identify future patterns in air temperature, thereby highlighting the implications of climate change. The performance metrics for the various ML models utilized in the forecasting process are presented in Table 5.

Table 5. Performance metrics of temperature forecasting models.

Model	R^2		*RMSE* (°C)	
	Training Set	**Test Set**	**Training Set**	**Test Set**
LR	0.601	0.483	3.077	3.443
DT	0.953	0.799	1.052	2.147
SVM	0.963	0.043	0.932	4.888
GPR	0.837	0.852	1.964	1.844
GAM	0.997	0.843	0.283	1.894

The DT and GAM models demonstrated superior performance relative to other forecasting models. Based on the lower *RMSE* values and the larger R^2 values, the GAM model was identified as the most optimal model for forecasting air temperature. The comparison between actual and predicted responses of the GAM model is illustrated in Figure 7d. The projected air temperature for the lifespan of the FPV systems, extending to the year 2052, is presented in Figure 8.
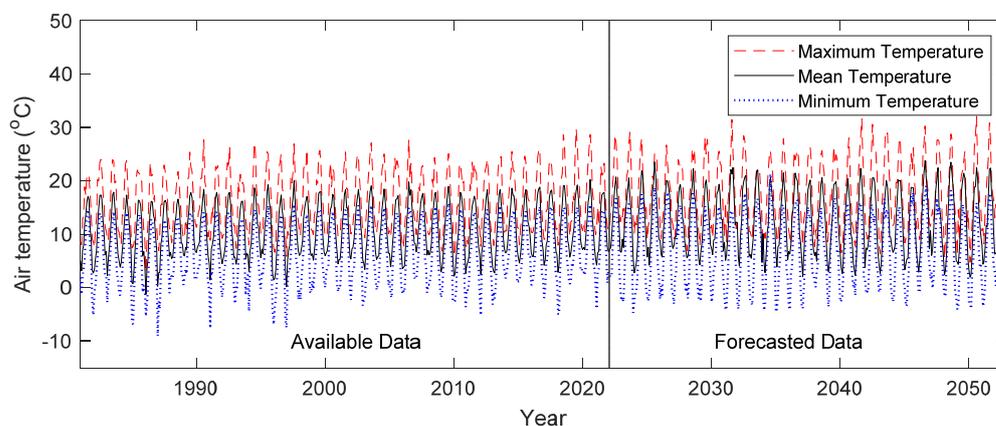
Figure 8. Forecasted air temperature for the lifespan of FPV systems.

Moreover, the comparison of mean surface air temperature for three distinct forecasting periods such as 2023–2032, 2033–2042, and 2043–2052, relative to the reference period of 2022, is depicted in Figure 9. The positive shift observed in the 10-year mean monthly air temperature from the reference period to subsequent future time periods indicates a trend of rising temperatures in the summer season. Significant variations were recorded during the winter months (February and March) and summer months (August and September). The highest recorded mean air temperature during the reference period was determined to be 20.68 °C, while the maximum values for the future forecasting periods were 20.88 °C, 21.14 °C, and 21.64 °C respectively. These findings suggest that the maximum average air temperature may increase by approximately 0.96 °C over the next three decades.

(**a**)

(**b**)

(**c**)
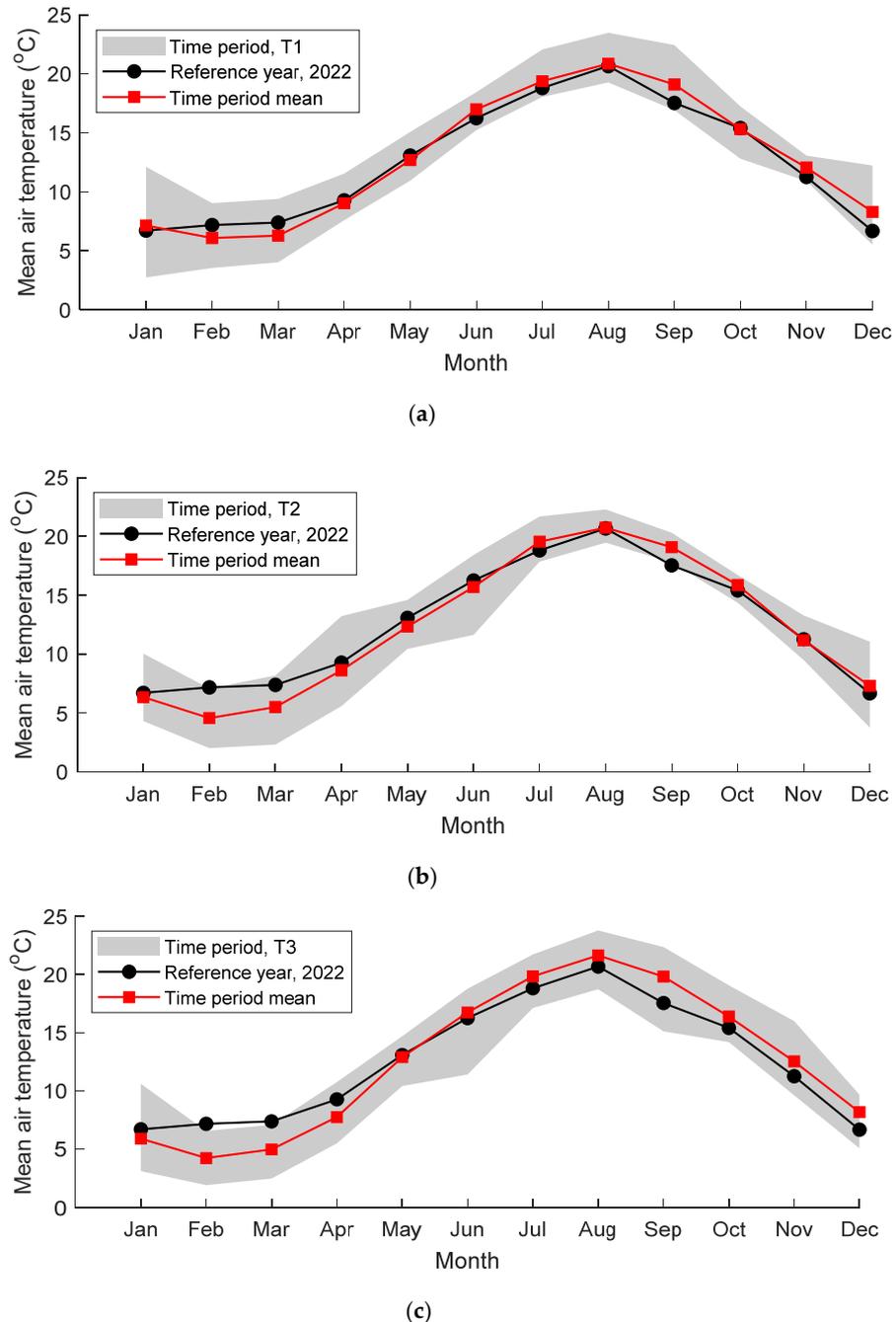
Figure 9. Comparison of monthly mean air temperature of reference period with the 10-year mean air temperature at different time periods: (**a**) 2023–2032, (**b**) 2033–2042, and (**c**) 2043–2052.

4. Results and Discussions

Three distinct time periods, namely, 2023–2032, 2033–2042, and 2043–2052, were examined to analyze the combined effects of climate change and the implementation of FPV systems. The projected values for air temperature, lake water temperature (both in open water and beneath the FPV systems), and various water quality parameters were calculated for each of these periods. This analysis aims to elucidate the variations induced by climate change and the installation of FPV systems. As depicted in Figure 4, the methodology follows the established framework designed to identify the projected water quality parameters.

4.1. Forecasted Air Temperature

The temperature forecasting model was utilized to ascertain the projected values of maximum, mean, and minimum temperature data. Given the significance of extreme temperature fluctuations in the analysis, the maximum values from the maximum temperature dataset and the minimum values from the minimum temperature dataset were incorporated into the comprehensive result analysis. A summary of the findings generated by the temperature forecasting model is presented in Table 6, which outlines the projections for the chosen time intervals of 2023–2032, 2033–2042, and 2043–2052. These intervals are referred to as T1, T2, and T3, respectively.

Table 6. Predicted air and water temperatures at the selected time periods during the lifetime of the FPV system.

Parameters	Statistics	T1 (2023–2032)	T2 (2033–2042)	T3 (2043–2052)
Air temperature (°C)	Maximum	31.44	31.63	32.08
	Minimum	−4.71	−4.55	−4.06
Water temperature (°C)	Maximum	27.98	28.03	28.17
	Minimum	−5.22	−5.08	−4.65

The maximum air temperatures were recorded in August and September during the summer, while the minimum temperatures were noted in February and March during the winter. This pattern strongly indicates that the forecasting model has successfully captured the seasonal temperature fluctuations characteristic of the specified location. Moreover, the results show an upward trend in maximum and minimum air temperatures across the successive time frames examined. Notably, the maximum and minimum air temperature values anticipated during the later period, T3, were elevated by 0.64 °C and 0.65 °C, respectively, compared to the measurements recorded during the initial period, T1. The results also indicate an average increase in the air temperature by 0.32 °C, every 10 years.

4.2. Water Temperature Prediction

The water temperature of the lake was forecasted utilizing the air–water temperature model based on the predicted extreme air temperature values for the selected periods. The earth's skin temperature and longwave downward irradiance, essential for the air–water temperature model, were determined using the linear correlations of these parameters with air temperature. Additionally, mean values of wind speed and shortwave downward irradiance were employed for the prediction. The projected water temperature corresponds to the lake water temperature at a depth of 2.0 m from the surface water level. The predicted maximum and minimum water temperatures for the different periods are presented in Table 6.

The findings suggest that the maximum water temperature during summer may increase by approximately 0.19 °C, while the minimum water temperature during winter may rise by about 0.57 °C over the next 30 years. This underscores the necessity of accounting for seasonal variations in assessments related to climate change. Moreover, such increases in water temperature over time could significantly alter the ecological and biochemical processes within aquatic systems [38,39].

A comparative analysis of maximum water and air temperatures for the different time periods is illustrated in Figure 10a. This comparison distinctly illustrates the rise in air and water temperatures over time, alongside a substantial decrease in water temperature relative to air temperature across all instances. The maximum water temperatures recorded during time periods T1, T2, and T3 were 11.01%, 11.38%, and 12.19% lower than the air temperature at the site. Regarding minimum predicted temperatures as shown in Figure 10b, the water temperatures were 9.77%, 10.43%, and 12.69% lower than the air temperatures during the respective time frames. The results also indicate that the difference between air and lake water temperatures increases over time.

(**a**)

(**b**)
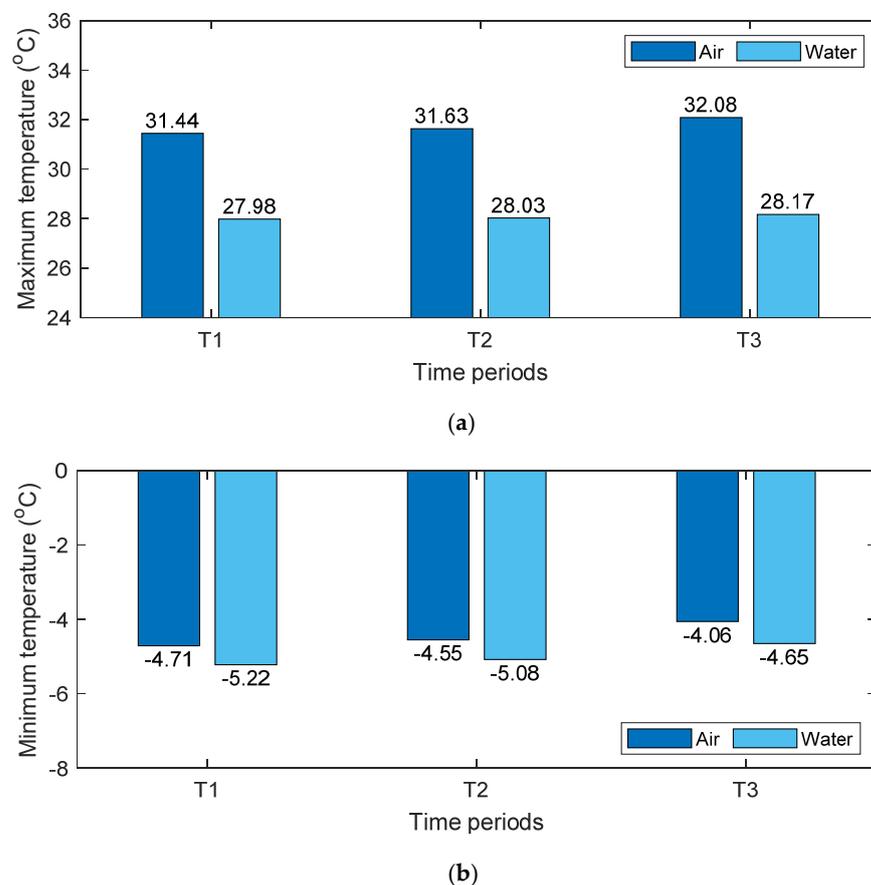
Figure 10. Comparison of predicted air and water temperature for the different time periods: (**a**) maximum temperature and (**b**) minimum temperature.

4.3. FPV System Water Temperature Prediction

The case study concerning the pilot project at Oostvoornse Lake examines three distinct types of FPV systems, each characterized by varying areas and water coverage. Consequently, the water temperature under the different FPV systems was predicted based on the forecasted lake water temperature using the FPV system temperature model. The maximum and minimum values of the water temperature across the different FPV systems are presented in Table 7.

Table 7. Predicted water temperature under the different FPV systems.

FPV Systems	Temperature (°C)	T1 (2023–2032)	T2 (2033–2042)	T3 (2043–2052)
FPV 1	Maximum	27.89	27.92	28.08
	Minimum	−1.69	−1.66	−1.56
FPV 2	Maximum	27.73	27.77	27.94
	Minimum	−1.62	−1.59	−1.48
FPV 3	Maximum	25.71	25.72	25.80
	Minimum	−1.35	−1.32	−1.21

A comparison of the maximum temperature values in open-water and FPV-covered water areas for the various FPV systems and time periods is illustrated in Figure 11a. As the case study focuses on water temperature data measured at a depth of 2.0 m, the predicted water temperature correspondingly reflects conditions at the same depth. Variations with respect to different water depths have not been addressed in this study. The findings indicate that the water temperature beneath all FPV systems is consistently lower than that of open water, a result attributed to the shading effect imposed by FPV systems. Notably, the water temperature under the FPV 3 system is significantly lower than that recorded for FPV 1 and FPV 2, likely due to the greater area associated with the FPV 3 installation. The increased surface area of the FPV system contributes to a more substantial cooling effect on the water temperature across a larger expanse of the surface, underscoring the importance of considering the specific properties of FPV systems in temperature predictions.
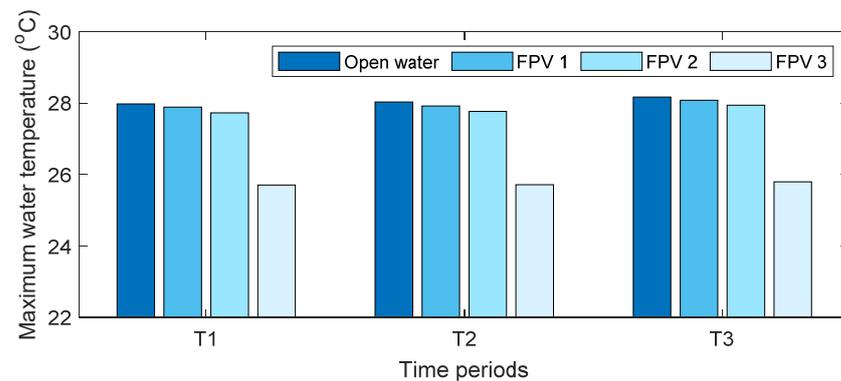
(a)
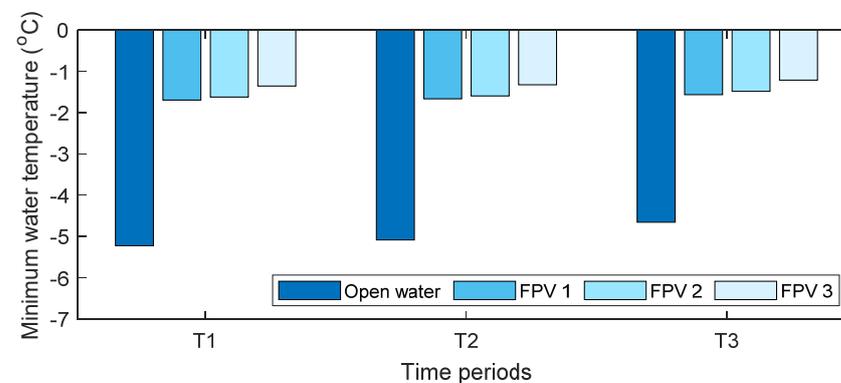
(b)

Figure 11. Comparison of water temperature in open-water and FPV-covered regions for different FPV systems: (**a**) maximum temperature and (**b**) minimum temperature.

The predictions regarding minimum water temperatures in regions with open water and those covered by FPV systems indicate an increase in water temperature under the FPV-covered areas, as shown in Figure 11b. The most significant effect was observed at FPV 3, highlighting the insulation effect provided by FPV installations during winter. While the shading effect of FPV systems tends to induce cooling during the summer, resulting in lower water temperatures, the opposite effect is noted during the winter months. These phenomena can also be attributed to the combined influences of various meteorological factors, including wind speed, solar radiation, and relative humidity. Furthermore, thermal radiation from the PV panels and the reduction of evaporative heat flux at the water surface contribute to increased water temperature [10]. Reducing water temperatures during the winter can present operational challenges in distribution mains and may increase tensile stresses in the pipes. Thus, the temperature increase induced by FPV panels could alleviate such challenges [40]. Nonetheless, these predictions are site-specific, and the effects may vary across different locations, types of FPV systems, and meteorological conditions. Literature also supports the idea that FPV systems can diminish the diurnal water temperature fluctuations throughout the day [21].

The variation in maximum water temperature is projected to rise by 0.19 °C over the next three decades. A similar increasing trend, albeit with a lesser magnitude, was recorded for the maximum water temperature beneath the FPV systems. Specifically, the increases in water temperature over the next 30 years under FPV systems 1, 2, and 3 are estimated to be 0.19 °C, 0.21 °C, and 0.09 °C, respectively. Despite the rising temperatures beneath FPV systems, these values remain lower than those observed in open-water bodies.

Another significant conclusion drawn from the predicted water temperature results is that, compared to open-water regions, the reduced water temperatures under the FPV systems after 30 years are close to the climate change-induced increased water temperatures of open water in 2022. This suggests that a 30-year warming trend in lakes could be mitigated by implementing FPV systems. Consequently, FPV installations represent a viable option for addressing the impacts of climate change on aquatic ecosystems [3]. However, it is essential to acknowledge that installing FPV systems may influence numerous metabolic processes within water bodies, ultimately leading to alterations in food web dynamics, species interactions, and carbon cycling [3,41].

4.4. Dissolved Oxygen Concentration Prediction

The dissolved oxygen model was utilized to predict the concentration of DO based on water temperature. In this analysis, both open-water temperature and the water temperature under three distinct FPV systems were considered in comparison with DO predictions. The results derived are presented in Table 8. Across all examined locations, both with and without FPV systems, it was observed that DO concentrations decreased over time, indicating an inverse correlation with rising temperatures. Notably, the increase in DO concentration over time was minimal, generally in the range of 0.1 to 0.2 mg/L.

For clarity, the DO concentration results corresponding to maximum and minimum water temperatures were designated as summer and winter seasons, respectively. Upon comparing the DO concentrations across different locations, it was noted that the DO concentration under FPV System 3 was greater than that observed in other scenarios during the summer season. This finding is further illustrated in the maximum DO concentration comparisons shown in Figure 12a. The shading effect produced by FPV systems typically contributes to a decline in DO concentration [42]. However, water temperature emerges as a significant factor that substantially influences DO levels. Variations in DO concentrations within lakes have been associated with both climate change and anthropogenic activities; several studies have documented a decrease in DO concentrations in inland waters over the

past decade [43,44]. Such fluctuations in DO concentration can adversely affect biodiversity and the functionality of aquatic ecosystems.

Table 8. Predicted DO concentration in the lake.

Location	DO Concentration (mg/L)	T1 (2023–2032)	T2 (2033–2042)	T3 (2043–2052)
Open water	Summer	5.70	5.68	5.58
	Winter	13.33	13.29	13.16
FPV 1	Summer	5.74	5.73	5.66
	Winter	12.21	12.20	12.17
FPV 2	Summer	5.81	5.79	5.72
	Winter	12.19	12.18	12.14
FPV 3	Summer	6.68	6.68	6.64
	Winter	12.10	12.09	12.05
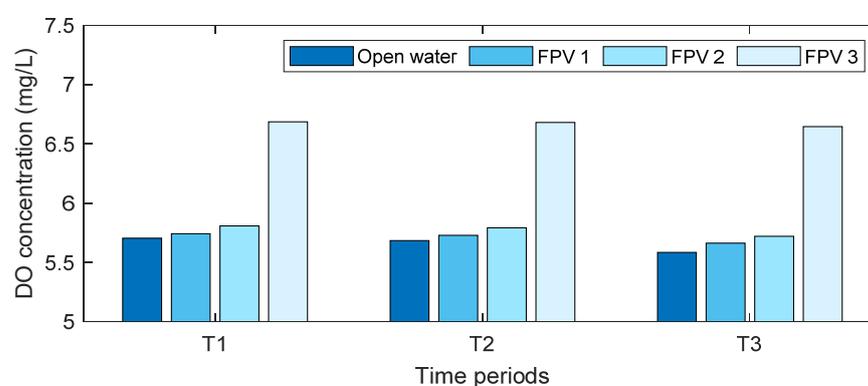
(**a**)
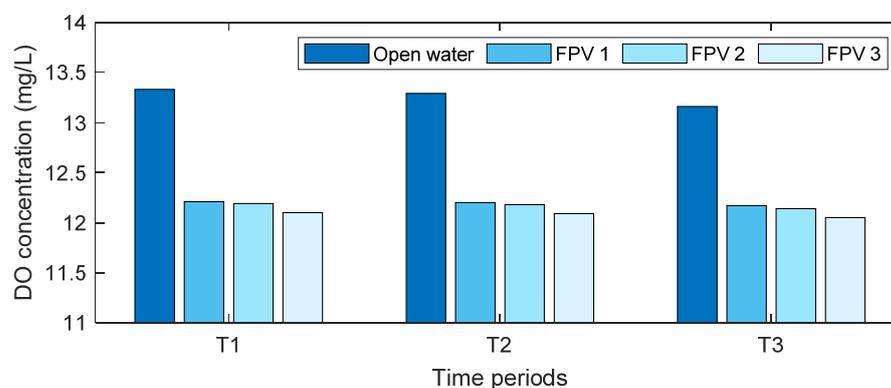
(**b**)

Figure 12. Comparison of DO concentration in open-water and FPV-covered locations under different time periods during (**a**) summer and (**b**) winter seasons.

In this study, the maximum DO concentrations measured under all FPV systems exceeded the open-water concentration during summer. Consequently, the increase in DO concentration attributed to FPV systems may benefit the aquatic environment. Furthermore, the findings indicate that open-water DO concentrations are declining over time, and installing FPV systems may facilitate an increase in these levels, potentially preventing the decline of DO concentrations to below 5 mg/L, a threshold critical for the survival of aquatic organisms [45]. Conversely, when analyzing DO concentrations during the winter season, the FPV systems were observed to reduce DO levels, as depicted in Figure 12b.

Although the magnitude of this reduction is relatively minor, the combined effects of climate change and FPV systems could significantly impact the biodiversity of the lake. Nevertheless, the results suggest that the range of DO concentrations recorded under the FPV systems throughout their operational lifespan does not pose a substantial threat to aquatic life.

5. Conclusions

The evaluation of the combined effects of FPV installations and climate change on the water quality of lakes presents a multifaceted challenge, owing to the multitude of interconnected meteorological, morphological, and quality parameters involved in the energy–water nexus. While the deployment of FPV systems is primarily motivated by the need to decarbonize energy generation, ensuring that these systems are installed and maintained sustainably is paramount to achieving long-term ecological balance. This study aims to develop an ML-based framework to ascertain the interaction between climate change and FPV system installations concerning the water quality in open-water lakes. The framework was applied to a case study derived from existing literature to gain insights into the potential impacts of FPV systems on the specific lake under consideration.

The framework entails the formulation of three predictive models alongside a forecasting model utilizing various ML algorithms. The selection of the most appropriate model for predicting various temperature and water quality parameters is crucial to achieving high accuracy in predictions. The case study chosen for the application of this ML framework is the Oostvoornse Lake in the Netherlands. Water quality data were sourced from the literature, while climate data were obtained from public databases to facilitate the impact analysis utilizing the developed framework. The artificial neural network (ANN), support vector machine (SVM), and generalized additive model (GAM) algorithms were identified as the most suitable for predicting the parameters at the selected site.

The application of the framework unveiled several significant insights regarding the combined impacts of FPV system installations and climate change on the water quality of Oostvoornse Lake. These insights are summarized as follows:

- The impact analysis indicated an overall increase in air temperature at the location, averaging 0.32 °C per decade.
- The findings projected a potential rise in maximum and minimum water temperatures of approximately 0.19 °C and 0.57 °C, respectively, over the next three decades, underscoring the importance of seasonal variations in climate assessments.
- Furthermore, predictions for maximum water temperatures beneath FPV systems demonstrated that temperatures remained consistently lower than those of the open water, indicative of the shading effect imposed by FPV systems. However, it was observed that the minimum water temperatures during winter under FPV systems were greater than those in the open water, attributable to the insulation effect of the FPV systems, which may assist in addressing challenges encountered by distribution systems in colder temperatures.
- In all FPV systems analyzed, maximum dissolved oxygen (DO) concentrations in summer surpassed those observed in open water, suggesting a potential benefit for aquatic environments. Despite a minor reduction in DO levels, overall concentrations beneath FPV systems remained within a range that does not pose a significant threat to aquatic life.

The above results are notably data-centric and specific to the site. The results may be relevant to open-water lakes with similar characteristics, contingent upon analogous climate conditions. A key conclusion from this investigation is the alteration in the thermal dynamics of the water body attributable to FPV systems. The motivation behind deploying

FPV systems is primarily driven by the necessity to decarbonize the energy supply, thereby mitigating the severe repercussions of climate change. However, the alterations in lake water temperature induced by FPV installations could either alleviate or exacerbate the impacts of climate change. The results from the present case study indicate that the effects of FPV system installations may significantly influence positive outcomes, potentially reversing the deleterious effects of climate change. However, it is crucial to underscore that the long-term sustainability of FPV systems could be compromised by their potential to disrupt local biodiversity and influence algae proliferation. These unintended consequences could alter aquatic ecosystems, necessitating further investigation to ensure the environmental benefits outweigh any negative impacts.

Based on the results obtained through the application of the developed framework, it is essential to acknowledge that the predicted outcomes are highly contingent upon the effective utilization of temperature and water quality-related parameters during model development. While temperature data for the location can be readily accessed, collecting comprehensive water quality data necessitates extensive monitoring efforts. Furthermore, incorporating additional meteorological and water quality parameters and the meticulous development of ML models will enhance the accuracy of prediction regarding water quality effects. Despite the substantial efforts dedicated to data development, conducting a thorough assessment of water quality before the commissioning of FPV systems is crucial to mitigating the environmental impacts of FPV projects.

In conclusion, FPV systems may have beneficial effects on climate and water quality dynamics. However, it is essential to continue investigating and addressing the broader environmental implications to secure sustainable outcomes. Achieving a balance between energy production and ecological health will ultimately determine the long-term effectiveness and acceptance of FPV technology in vulnerable lake ecosystems.

References

1. Nagananthini, R.; Nagavinothini, R. Investigation on Floating Photovoltaic Covering System in Rural Indian Reservoir to Minimize Evaporation Loss. *Int. J. Sustain. Energy* **2021**, *40*, 781–805. [CrossRef]
2. Haas, J.; Khalighi, J.; de la Fuente, A.; Gerbersdorf, S.U.; Nowak, W.; Chen, P.-J. Floating Photovoltaic Plants: Ecological Impacts versus Hydropower Operation Flexibility. *Energy Convers. Manag.* **2020**, *206*, 112414. [CrossRef]
3. Exley, G.; Armstrong, A.; Page, T.; Jones, I.D. Floating Photovoltaics Could Mitigate Climate Change Impacts on Water Body Temperature and Stratification. *Solar Energy* **2021**, *219*, 24–33. [CrossRef]
4. Gadzanku, S.; Mirletz, H.; Lee, N.; Daw, J.; Warren, A. Benefits and Critical Knowledge Gaps in Determining the Role of Floating Photovoltaics in the Energy-Water-Food Nexus. *Sustainability* **2021**, *13*, 4317. [CrossRef]
5. Bax, V.; van de Lageweg, W.I.; Hoosemans, R.; van den Berg, B. Floating Photovoltaic Pilot Project at the Oostvoornse Lake: Assessment of the Water Quality Effects of Three Different System Designs. *Energy Rep.* **2023**, *9*, 1415–1425. [CrossRef]
6. Ravichandran, N.; Ravichandran, N.; Panneerselvam, B. Comparative Assessment of Offshore Floating Photovoltaic Systems Using Thin Film Modules for Maldives Islands. *Sustain. Energy Technol. Assess.* **2022**, *53*, 102490. [CrossRef]

7. Ravichandran, N.; Ravichandran, N.; Panneerselvam, B. Review on the Structural Components of Floating Photovoltaic Covering Systems. In *Intelligent Manufacturing and Energy Sustainability: Proceedings of ICIMES 2021*; Springer: Singapore, 2021; pp. 125–133.

8. Ravichandran, N.; Ravichandran, N.; Panneerselvam, B. Performance Analysis of a Floating Photovoltaic Covering System in an Indian Reservoir. *Clean Energy* **2021**, *5*, 208–228. [CrossRef]

9. Yang, K.; Yu, Z.; Luo, Y.; Zhou, X.; Shang, C. Spatial-Temporal Variation of Lake Surface Water Temperature and Its Driving Factors in Yunnan-Guizhou Plateau. *Water Resour. Res.* **2019**, *55*, 4688–4703. [CrossRef]

10. Yang, P.; Chua, L.H.C.; Irvine, K.N.; Nguyen, M.T.; Low, E.W. Impacts of a Floating Photovoltaic System on Temperature and Water Quality in a Shallow Tropical Reservoir. *Limnology* **2022**, *23*, 441–454. [CrossRef]

11. Wang, T.W.; Chang, P.H.; Huang, Y.S.; Lin, T.S.; Yang, S.D.; Yeh, S.L.; Tung, C.H.; Kuo, S.R.; Lai, H.T.; Chen, C.C. Effects of Floating Photovoltaic Systems on Water Quality of Aquaculture Ponds. *Aquac. Res.* **2022**, *53*, 1304–1315. [CrossRef]

12. Château, P.A.; Wunderlich, R.F.; Wang, T.W.; Lai, H.T.; Chen, C.C.; Chang, F.J. Mathematical Modeling Suggests High Potential for the Deployment of Floating Photovoltaic on Fish Ponds. *Sci. Total Environ.* **2019**, *687*, 654–666. [CrossRef] [PubMed]

13. de Lima, R.L.P.; Paxinou, K.; C. Boogaard, F.; Akkerman, O.; Lin, F.-Y. In-Situ Water Quality Observations under a Large-Scale Floating Solar Farm Using Sensors and Underwater Drones. *Sustainability* **2021**, *13*, 6421. [CrossRef]

14. Ziar, H.; Prudon, B.; Lin, F.Y.; Roeffen, B.; Heijkoop, D.; Stark, T.; Teurlincx, S.; de Senerpont Domis, L.; Goma, E.G.; Extebarria, J.G.; et al. Innovative Floating Bifacial Photovoltaic Solutions for Inland Water Areas. *Prog. Photovolt. Res. Appl.* **2021**, *29*, 725–743. [CrossRef]

15. Ravichandran, N.; Paneerselvam, B.; Ravichandran, N. GIS-Based Potential Assessment of Floating Photovoltaic Systems in Reservoirs of Tamil Nadu in India. *Clean. Energy* **2023**, *7*, 671–689. [CrossRef]

16. Nicholls, R.J.; Cazenave, A. Sea-Level Rise and Its Impact on Coastal Zones. *Science* **2010**, *328*, 1517–1520. [CrossRef]

17. Intergovernmental Panel on Climate Change; Stocker, T.F.; Qin, D.; Plattner, G.-K.; Tignor, M.; Allen, S.K.; Boschung, J.; Nauels, A.; Xia, Y.; Bex, V.; et al. *IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*; Cambridge University Press: Cambridge, UK, 2013.

18. Yang, K.; Yu, Z.; Luo, Y. Analysis on Driving Factors of Lake Surface Water Temperature for Major Lakes in Yunnan-Guizhou Plateau. *Water Res.* **2020**, *184*, 116018. [CrossRef]

19. O'Reilly, C.M.; Sharma, S.; Gray, D.K.; Hampton, S.E.; Read, J.S.; Rowley, R.J.; Schneider, P.; Lenters, J.D.; McIntyre, P.B.; Kraemer, B.M.; et al. Rapid and Highly Variable Warming of Lake Surface Waters around the Globe. *Geophys. Res. Lett.* **2015**, *42*, 10773–10781. [CrossRef]

20. O'Reilly, C.M.; Alin, S.R.; Plisnier, P.-D.; Cohen, A.S.; McKee, B.A. Climate Change Decreases Aquatic Ecosystem Productivity of Lake Tanganyika, Africa. *Nature* **2003**, *424*, 766–768. [CrossRef] [PubMed]

21. Liu, Z.; Ma, C.; Yang, Y.; Li, X.; Gou, H.; Folkard, A.M. Water Temperature and Energy Balance of Floating Photovoltaic Construction Water Area—Field Study and Modelling. *J. Environ. Manag.* **2024**, *365*, 121494. [CrossRef] [PubMed]

22. Schmid, M.; Köster, O. Excess Warming of a Central European Lake Driven by Solar Brightening. *Water Resour. Res.* **2016**, *52*, 8103–8116. [CrossRef]

23. Woolway, R.I.; Merchant, C.J. Intralake Heterogeneity of Thermal Responses to Climate Change: A Study of Large Northern Hemisphere Lakes. *J. Geophys. Res. Atmos.* **2018**, *123*, 3087–3098. [CrossRef]

24. Michalski, R.S.; Carbonell, J.G.; Mitchell, T.M. (Eds.) *Machine Learning*; Springer: Berlin/Heidelberg, Germany, 1983; ISBN 978-3-662-12407-9.

25. Oswald, C.J.; Rouse, W.R. Thermal Characteristics and Energy Balance of Various-Size Canadian Shield Lakes in the Mackenzie River Basin. *J. Hydrometeorol.* **2004**, *5*, 129–144. [CrossRef]

26. Breiman, L.; Friedman, J.H.; Olshen, R.A.; Stone, C.J. *Classification And Regression Trees*; Routledge: Abingdon, UK, 2017; ISBN 9781315139470.

27. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning*; Springer: New York, NY, USA, 2009; ISBN 978-0-387-84857-0.

28. Quinlan, J.R. Induction of Decision Trees. *Mach. Learn.* **1986**, *1*, 81–106. [CrossRef]

29. Ho, T.K. The Random Subspace Method for Constructing Decision Forests. *IEEE Trans. Pattern Anal. Mach. Intell.* **1998**, *20*, 832–844. [CrossRef]

30. Bishop, C.M. *Pattern Recognition and Machine Learning (Information Science and Statistics)*; Springer: Berlin/Heidelberg, Germany, 2006; ISBN 0387310738.

31. Cristianini, N.; Shawe-Taylor, J. *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*; Cambridge University Press: Cambridge, UK, 2000; ISBN 9780521780193.

32. Cortes, C.; Vapnik, V. Support-Vector Networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]

33. Rasmussen, C.E. Gaussian Processes in Machine Learning. In *Summer School on Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2004; pp. 63–71.

34. Simon, S. *Haykin Neural_Networks*, 2nd ed.; Pearson Prentice Hall: Bangalore, India, 2004.

35. Chambers, J.M.; Hastie, T.J. (Eds.) *Statistical Models in S*; Routledge: Abingdon, UK, 2017; ISBN 9780203738535.

36. Cromratie Clemons, S.K.; Salloum, C.R.; Herdegen, K.G.; Kamens, R.M.; Gheewala, S.H. Life Cycle Assessment of a Floating Photovoltaic System and Feasibility for Application in Thailand. *Renew. Energy* **2021**, *168*, 448–462. [CrossRef]

37. NASA POWER National Aeronautics and Space Administration (NASA) Langley Research Center (LaRC) Prediction of Worldwide Energy Resource (POWER). Available online: https://power.larc.nasa.gov/data-access-viewer/ (accessed on 10 November 2024).

38. Paerl, H.W.; Paul, V.J. Climate Change: Links to Global Expansion of Harmful Cyanobacteria. *Water Res.* **2012**, *46*, 1349–1363. [CrossRef]

39. Woolway, R.I.; Weyhenmeyer, G.A.; Schmid, M.; Dokulil, M.T.; de Eyto, E.; Maberly, S.C.; May, L.; Merchant, C.J. Substantial Increase in Minimum Lake Surface Temperatures under Climate Change. *Clim. Change* **2019**, *155*, 81–94. [CrossRef]

40. Jesson, D.A.; Le Page, B.H.; Mulheron, M.J.; Smith, P.A.; Wallen, A.; Cocks, R.; Farrow, J.; Whiter, J.T. Thermally Induced Strains and Stresses in Cast Iron Water Distribution Pipes: An Experimental Investigation. *J. Water Supply Res. Technol. AQUA* **2010**, *59*, 221–229. [CrossRef]

41. Kraemer, B.M.; Chandra, S.; Dell, A.I.; Dix, M.; Kuusisto, E.; Livingstone, D.M.; Schladow, S.G.; Silow, E.; Sitoki, L.M.; Tamatamah, R.; et al. Global Patterns in Lake Ecosystem Responses to Warming Based on the Temperature Dependence of Metabolism. *Glob. Change Biol.* **2017**, *23*, 1881–1890. [CrossRef] [PubMed]

42. Chen, X.C.; Kong, H.N.; He, S.B.; Wu, D.Y.; Li, C.J.; Huang, X.C. Reducing Harmful Algae in Raw Water by Light-Shading. *Process Biochem.* **2009**, *44*, 357–360. [CrossRef]

43. Breitburg, D.; Levin, L.A.; Oschlies, A.; Grégoire, M.; Chavez, F.P.; Conley, D.J.; Garçon, V.; Gilbert, D.; Gutiérrez, D.; Isensee, K.; et al. Declining Oxygen in the Global Ocean and Coastal Waters. *Science* **2018**, *359*, eaam7240. [CrossRef] [PubMed]

44. Luo, A.; Chen, H.; Gao, X.; Carvalho, L.; Zhang, H.; Yang, J. The Impact of Rainfall Events on Dissolved Oxygen Concentrations in a Subtropical Urban Reservoir. *Environ. Res.* **2024**, *244*, 117856. [CrossRef]

45. Bozorg-Haddad, O.; Delpasand, M.; Loáiciga, H.A. Water Quality, Hygiene, and Health. In *Economical, Political, and Social Issues in Water Resources*; Elsevier: Amsterdam, The Netherlands, 2021.