International Journal of Geo-Information

MDPI

Article

Assessing and Predicting Nearshore Seawater Quality with Spatio-Temporal Semivariograms: The Case of Coastal Waters in Fujian Province, China

Wei Wang [1,*], Wenfang Cheng [2] and Jing Chen [1]

[1] College of Information Technology Engineering, Tianjin University of Technology and Education, Tianjin 300222, China; chenjing0828@tute.edu.cn
[2] Polar Research Institute of China, Ministry of Natural Resources of the People's Republic of China, Shanghai 200136, China; chengwenfang@pric.org.cn
* Correspondence: wangw@tute.edu.cn

Abstract: The scientific assessment and prediction of nearshore water quality are crucial for marine environment protection efforts. This study is based on a comprehensive analysis of existing assessment and prediction methods and considers the regular and random characteristics of nearshore seawater quality due to both natural and anthropogenic influences. It proposes a new method that applies the kriging interpolation algorithm to empirically generated spatio-temporal semivariograms to assess and predict seawater quality. The application of this method in Fujian coastal areas shows that it is able to flexibly and scientifically estimate the variations in various indicators in the region. Combined with GIS spatial data overlay analysis operations, it can be used to quantitatively evaluate different qualities of seawater and provide scientific guidance for marine environmental protection.

Keywords: seawater quality; spatio-temporal kriging; semivariogram; overlay analysis

Citation: Wang, W.; Cheng, W.; Chen, J. Assessing and Predicting Nearshore Seawater Quality with Spatio-Temporal Semivariograms: The Case of Coastal Waters in Fujian Province, China. *ISPRS Int. J. Geo-Inf.* **2024**, *13*, 292. https://doi.org/10.3390/ijgi13080292

Academic Editor: Wolfgang Kainz

Received: 16 May 2024
Revised: 29 July 2024
Accepted: 16 August 2024
Published: 17 August 2024

1. Introduction

The quality of seawater has a direct impact on marine ecosystems and human health [1]. Pollutants and contaminants in seawater can seriously harm marine life and ultimately disrupt the stability of the entire marine ecosystem. Pollutants in seawater can also enter the human body through the food chain, posing a potential threat to human health [2,3]. The comprehensive monitoring of changes in the quality of seawater, especially coastal seawater, along with the accurate and quantitative scientific assessment and prediction of the quality of seawater, can provide a scientific basis for rational and effective ocean management. However, due to the joint influences of natural factors—such as tides, ocean currents, hurricanes, and river inflows—and human factors—such as industrial discharges, agricultural nonpoint source pollution, domestic sewage, and marine debris, changes in the quality of water in nearshore areas exhibit both regularity and randomness. This poses numerous challenges for timely scientific and quantitative assessments and predictions of the quality of nearshore seawater [4–6].

Based on the regularity of changes in nearshore seawater quality, many recent studies have introduced machine learning (ML) and artificial intelligence (AI) technologies for assessing and predicting nearshore seawater quality. By analyzing and studying a large volume of existing sample data, ML and AI technologies can reveal underlying patterns of changes in water quality and thereby assess and predict the quality of water. The study in [7] uses AI technology that includes multi-layer perceptrons and support vector machines to predict aspects of water quality. The study in [8] improves the accuracy of various ML models through optimization and parameter adjustment. The study in [9] proposes advanced AI algorithms to predict the water quality index (WQI) and classify water quality based on its robustness. The study in [10] presents an AI-based binary

classification model, Easy Ensemble, that uses class imbalance learning to predict "very bad" situations, and it achieves accurate predictions for beach closing times. In general, the current applications of ML and AI technologies in water quality prediction focus on exploring internal regularities within large-scale data to achieve scientific quantitative evaluations of changes in water quality. However, due to the influence of various human factors, changes in coastal seawater quality also exhibit a certain degree of randomness within regions. Thus, the exploration of regularity alone does not fully reflect the actual changes in seawater quality. Furthermore, due to the inability to provide large-scale sample data for ML or AI in practice, the predicted values obtained using ML or AI methods cannot obtain the required accuracy.

Therefore, many studies on assessing and predicting changes in seawater quality also consider the random components that are influenced by human activities and sudden pollution events based on small-scale data sets. Although there is relatively little research on the application of geostatistical methods in predicting seawater quality, such methods have been widely used in monitoring and predicting environmental quality factors. These studies are mostly based on statistical analysis and techniques from the field of geostatistics, such as regression analysis, time series analysis, and factor analysis, using the theory and technology of regionalized variables. The study in [11] combines chemometrics and geostatistics to characterize soil geochemistry and microspatial variability along groundwater flow paths, illustrating that chemometric and geostatistical analyses can be powerful tools for studying biogeochemical cycles and sources of groundwater pollutants. The study in [12] uses geostatistical methods to characterize the spatial distribution of groundwater quality parameters and assess the suitability of groundwater for drinking and irrigation. The study in [13] applies comprehensive statistical, geostatistical, and spatial interpolation methods to analyze groundwater hydrochemistry data from 79 boreholes near Leliefontein in the South African Kamiesberg Local Municipality, using inverse distance weighting for major cations and major anions, the sodium adsorption ratio, electrical conductivity, and the WQI for geostatistical and spatial analysis interpolation. The study in [14] comprehensively analyzes the main factors and processes controlling the water quality of some coastal aquifers through a combination of principal component analysis, cluster analysis, geostatistical analysis, and an entropy-based groundwater quality index, along with a detailed hydrogeochemical assessment. In general, the research that uses statistical analyses is based on the theory of regionalized variables, that is, that water quality changes are random in space. The advantage of this theory is that it reveals not only the regularity of water quality distribution, but also the randomness of water quality changes. Introducing the theory and technology of statistical analysis in the assessment and prediction of coastal seawater quality can compensate for the shortcomings of ML or AI technology in detecting random characteristics based on a large-scale data set.

The present study considers both the regularity and the randomness of changes in nearshore seawater quality, incorporates the theory of regionalized variables in geostatistics, and scientifically evaluates and accurately predicts the nearshore seawater quality. Specifically, the objectives of this study are as follows: (1) to use the theory and technology of regionalized variables to study the spatial and temporal correlations of historical seawater quality data; (2) to use historical data on nearshore water quality observations as the research object to uncover both the regular and the random components of water quality changes, thereby achieving trend predictions for changes in water quality; and (3) to conduct a comprehensive evaluation and quantitative analysis of seawater quality through the use of GIS overlay analysis technology. We expect that the results of this study will provide a scientific basis for the accurate evaluation and prediction of coastal seawater quality.

2. Methods

2.1. Semivariograms

Geostatistical methods such as empirical means, empirical covariances, and semivariograms are widely used to map environmental factors and estimate the areal values of environment factors [15]. Investigating spatio-temporal data through empirical means and covariances offers valuable insights [16,17].

To obtain accurate predictions for a spatio-temporal process, it is imperative to understand its joint spatio-temporal dependence structure [18]. Therefore, empirical spatio-temporal covariograms and semivariograms are considered to measure the joint spatio-temporal dependence. The characterization of the covariability within the spatio-temporal data, as a function of specific lags in time and space, is of paramount importance.

Consider the empirical spatio-temporal covariance function for various space and time lags. The empirical spatio-temporal covariogram for spatial lag h and time lag τ is then given by

$$\widehat{C}_z(h;\tau) = \frac{1}{|N_s(h)|}\frac{1}{|N_t(\tau)|}\sum_{s_i,s_k\in N_s(h)}\sum_{t_j,t_{\updownarrow}\in N_t(\tau)}(Z(s_i;t_j)-\widehat{\mu}_{z,s}(s_i))(Z(s_k;t_{\updownarrow})-\widehat{\mu}_{z,s}(s_k)), \quad (1)$$

where

$\widehat{\mu}_{z,s}(s_i) = (1/T)\sum_{j=1}^{T} Z(s_i;t_j)$,

$N_s(h)$ refers to the pairs of spatial locations with a spatial lag within some tolerance h,

$N_t(\tau)$ refers to the pairs of time points with a time lag within some tolerance τ, and

$|N(\cdot)|$ refers to the number of elements in $N(\cdot)$.

A semivariogram is used to quantify the spatial and/or temporal autocorrelation between sample points in a spatial data set. It can characterize the degree of spatial and/or temporal variation in the variables of a geographical process within a given distance and/or over a certain period of time. In short, it can help to show how sample data are correlated with each other as distance and/or time change [19,20]. The semivariogram is defined as

$$\gamma_z(s_i,s_k;t_j,t_{\updownarrow}) \equiv \frac{1}{2}\mathrm{var}(Z(s_i;t_j)-Z(s_k;t_{\updownarrow})). \quad (2)$$

When the covariance depends only on the displacements in space and differences in time, this can be written as

$$\gamma_z(h;\tau) \equiv \frac{1}{2}\mathrm{var}(Z(s+h;t+\tau)-Z(s;t)), \quad (3)$$

where

$h = s_k - s_i$ is a spatial lag and

$\tau = t_{\updownarrow} - t_j$ is a temporal lag.

If the covariance function of a process is well defined, then the semivariogram is generally characterized by the nugget effect, the sill, and the partial sill. The nugget effect is given by $\gamma_z(h;\tau)$ when $h \to 0$ and $\tau \to 0$, the sill is $\gamma_z(h;\tau)$ when $h \to \infty$ and $\tau \to \infty$, and the partial sill is the difference between the sill and the nugget effect. Figure 1 shows these components of a semivariogram as a function of the spatial distance ||h||.

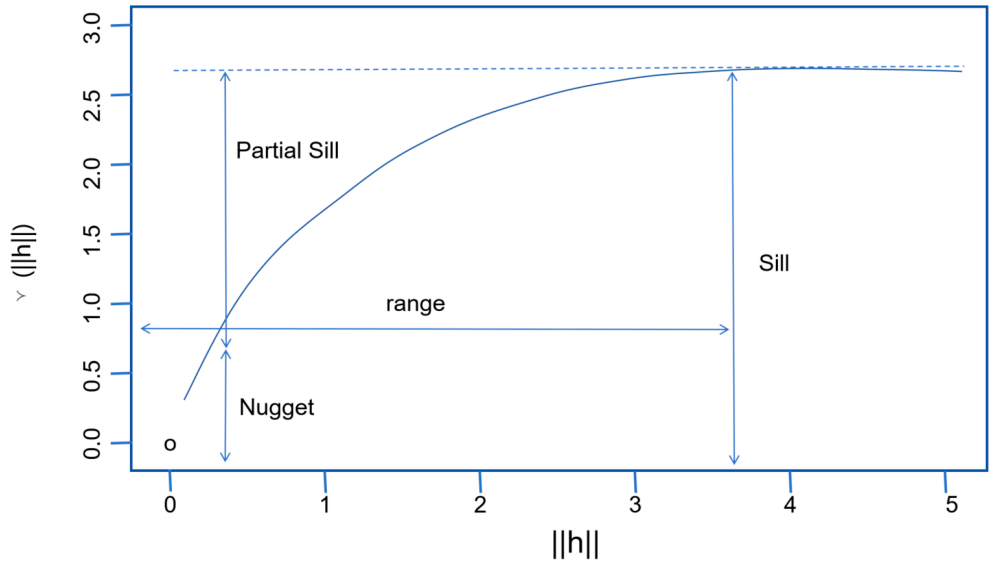

Figure 1. Parameters of a semivariogram [21].

The shape of the semivariogram function is critical for understanding and modeling the correlation structures of the spatial and temporal data, which is essential for accurately predicting values at unsampled locations through interpolation methods [19]. In addition, semivariograms can help identify outliers, trends, periodicity, and other structural features within the data set. More generally, modeling spatio-temporal data by combining covariance functions with semivariograms can provide a more complete understanding of the intrinsic relationships within the data, which, in turn, facilitates the effective prediction and simulation of complex spatio-temporal processes [22].

2.2. Spatio-Temporal Separable Model

In this study, the spatial and temporal influences are independent events, and their joint effect is expressed by multiplication, thereby capturing the overall correlation of the spatio-temporal factors. The spatio-temporal separable model can be described by the following formula [23,24]:

$$C_{sep}(h, u) = C_s(h) \cdot C_t(u), \tag{4}$$

where

$C_s(h)$ represents the spatial semivariogram, which changes with the difference in the spatial distance h, and

$C_t(u)$ represents the temporal semivariogram, which changes with the difference in the time u.

Fitting a theoretical variogram to an empirical variogram is a key step in spatio-temporal data prediction. First, it helps in understanding and quantifying the spatial relationships between samples and the variability of sample variables through a parameterized theoretical function. In the spatio-temporal separable theoretical model, the parameters involved include the nugget, sill, and range. Second, a description of the variability that is valid for the entire data set can be obtained through model fitting. Compared to the empirical semivariogram, which can be affected by data sampling errors and may be noisy and discontinuous, the theoretical variogram provides a smooth, continuous estimate. Next, predictions of values are made by interpolation operations in the spatio-temporal domain. The interpolation is based on the weighted average of adjacent sample values, with weights determined by the semivariogram relationship between samples, all of which require a theoretical model. Finally, the variogram model can be used to assess the uncertainty of predictions for unknown sample locations. From the characteristics of the variogram, the variance in the prediction error can be calculated in order to understand the reliability of the prediction. A well-fitted theoretical variogram provides a standardized way to describe

the spatial correlation of a data set, rendering results more comparable across studies and applications.

2.3. Overlay Analysis with Customized Rules

According to the seawater quality evaluation standard, when evaluating seawater quality, it is necessary to make give comprehensive consideration to multiple observation indicators at a certain point. In order to ensure the accuracy and scientificity of the seawater quality assessment, the evaluation area is first divided into a number of $n \times n$ small cell grids, and the center point of each grid is regarded as the unknown point to be sought in the subsequent interpolation prediction. Next, using the acquired observation point data, the spatio-temporal Kriging interpolation algorithm is used to calculate a certain estimated value of the unknown point. For each small cell within the evaluation area, the data of multiple indicators are analyzed one by one in accordance with the seawater quality evaluation criteria, and the seawater quality grade within each small square is finally determined.

Since the overlay analysis function provided by general GIS software cannot easily meet the complexity of the seawater quality evaluation standards, it is necessary to define the overlay rules through customized methods to perform an overlay analysis of multiple indicators, and ultimately determine the seawater quality level in each small area. The process of overlay analysis is depicted in Figure 2.

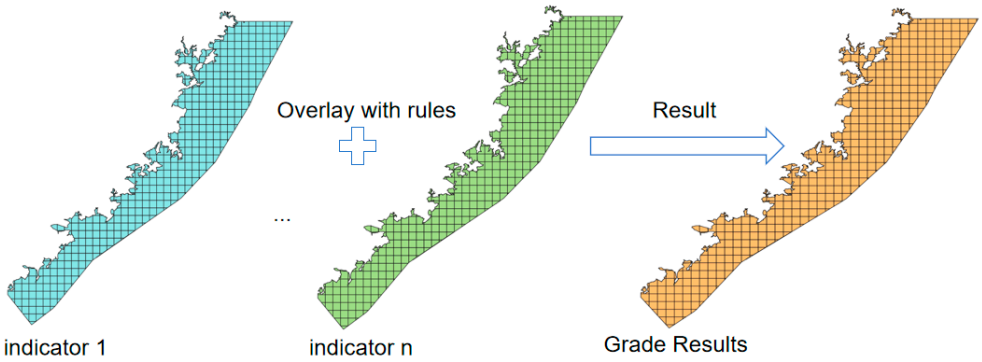

Figure 2. Overlay analysis process.

3. Implementation

3.1. Study Area

This article focuses on the coastal waters of Fujian Province, China, as the area of the study. The Fujian marine area includes subtropical ocean and a shallow continental shelf sea, forming a junction of cold and warm currents. The coastal waters under study cover an area of 36,000 square kilometers, with a diversified marine ecosystem that is rich in various marine resources, including abundant fishery resources and seabed minerals. The locations of the region and of the water quality monitoring stations are shown in Figure 3.

The Fujian Sea area, influenced by the Kuroshio Current and the diluting effects of the Yangtze River's waters, possesses robust seawater exchange capabilities for self-purification [25]. However, given the rapidly developing marine economy and the industrialization and urbanization of coastal areas, the seawater quality of coastal waters is significantly affected by the byproducts of human activities, including industrial discharges, agricultural non-point source pollution, domestic sewage, and marine debris. Seawater quality therefore faces tremendous challenges [26,27]. In order to ensure the sustainable utilization of marine resources, protect the ecological environment of coastal waters, and ensure the diversity of marine life and the long-term development of fisheries, Fujian Province has established 235 water quality monitoring sites to monitor multiple chemical indicators in real time in its coastal waters.

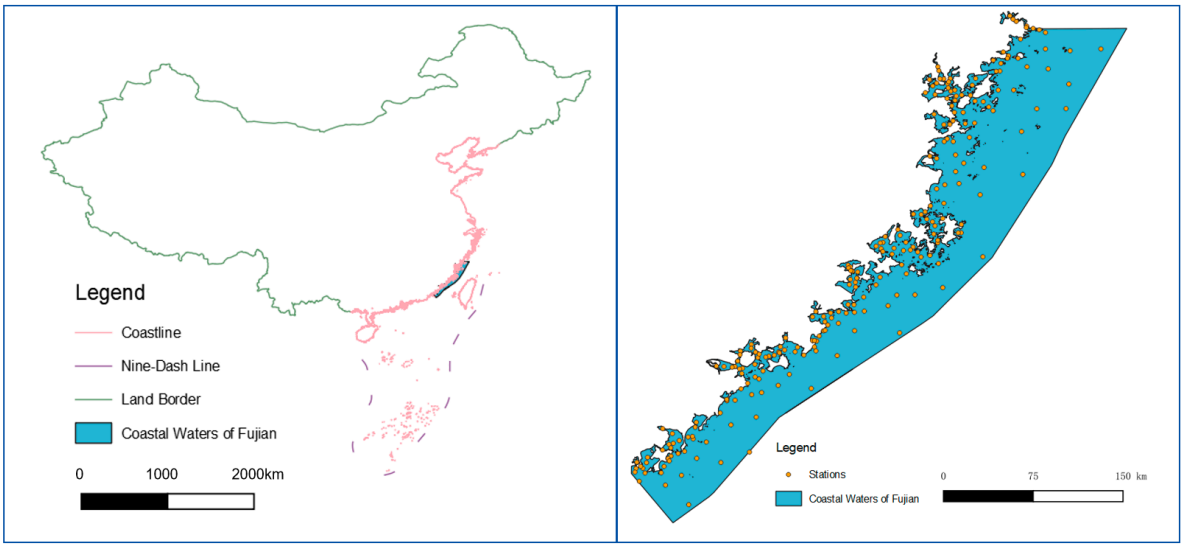

Figure 3. Research area and distribution of observation sites.

3.2. Spatio-Temporal Data Set

The water quality monitoring data used in this study come from the 235 water quality monitoring points along the coast of Fujian Province that collect data on eight indicators: the concentrations of suspended solids, dissolved oxygen, reactive phosphates, chemical oxygen, nitrite nitrogen, nitrate nitrogen, ammonia nitrogen, and total inorganic nitrogen. To produce quarterly seawater quality indicator data, the province's Seawater Quality Control Division collects seawater samples at various locations on a quarterly basis and measures the concentrations of these indicators in a laboratory environment. The quarterly data are averaged by indicator to form a data set that reflects the annual change in water quality. We use the last five years (2019–2023) of annual water quality data as the research data set for this study. Figures 4 and 5 show the overall spatial distributions of the annual water quality monitoring data for reactive phosphate and inorganic nitrogen, respectively.

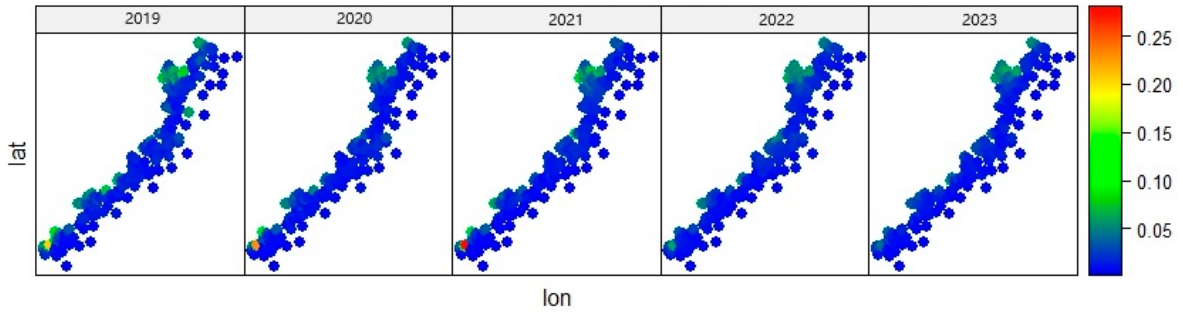

Figure 4. Temporal and spatial distributions of reactive phosphate data from 2019 to 2023.

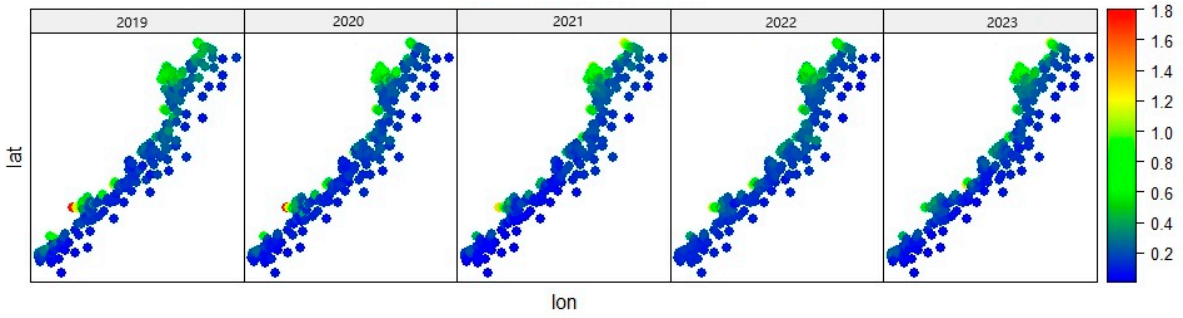

Figure 5. Temporal and spatial distributions of inorganic nitrogen data from 2019 to 2023.

3.3. Empirical Semivariogram

Using the variogramST () method of the R language package gstat, we can compute an empirical spatio-temporal semivariogram (or sample variogram) from spatio-temporal data [28,29]. Figure 6 provides a multi-view visualization of the empirical spatio-temporal semivariogram. Such visualizations are typically used to show how semivariograms vary with distance and time, providing insights into the spatial and temporal continuity of the monitored environmental variables.

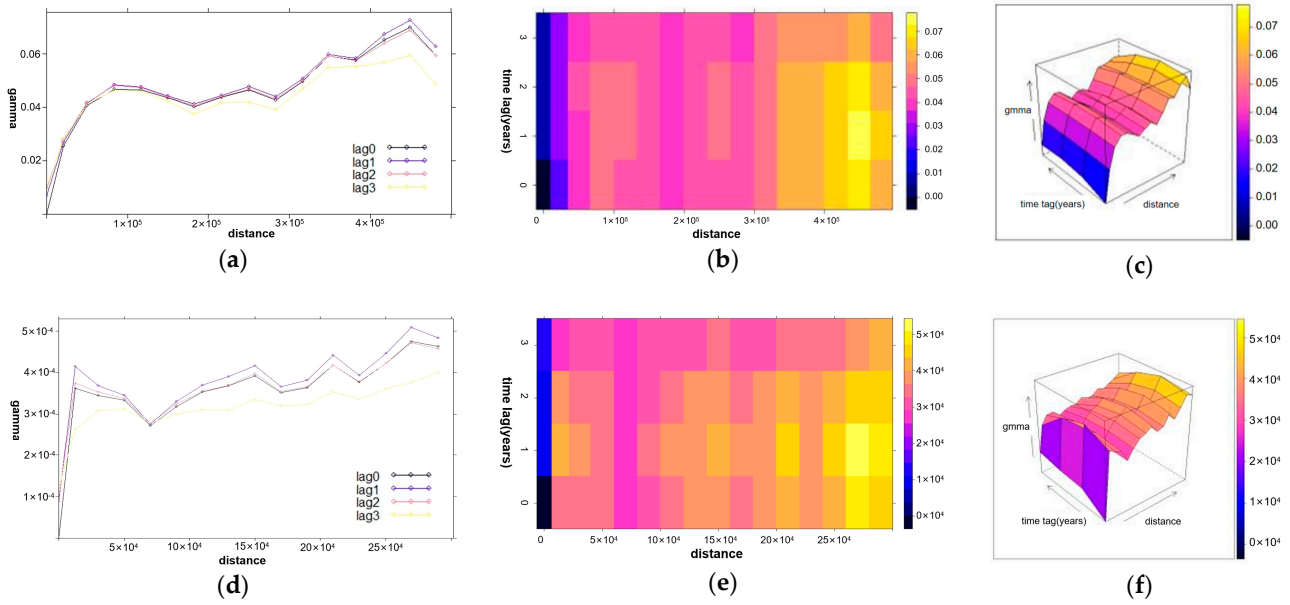

Figure 6. Multi-view visualization of empirical semivariogram of water quality monitoring data: (**a**) curve for inorganic nitrogen; (**b**) 2D map for inorganic nitrogen; (**c**) 3D cube for inorganic nitrogen; (**d**) curve for reactive phosphate; (**e**) 2D map for reactive phosphate; (**f**) 3D cube for reactive phosphate.

In the empirical variograms of inorganic nitrogen content shown in Figure 6a,d, it can intuitively be observed that there is a correlation in the seawater concentrations of inorganic nitrogen and reactive phosphates along the spatial dimension. As the value on the horizontal axis (distance) increases, the value on the vertical axis (gamma) increases significantly. For inorganic nitrogen, the semivariogram value reaches its maximum when the spatial distance reaches about 80 km and remains flat thereafter. This indicates that the inorganic nitrogen monitoring data have significant spatial correlation characteristics within a range of 80 km, that is, spatial location has a significant effect on the inorganic nitrogen within an 80 km range. For reactive phosphates, the semivariogram value increases significantly within a 10 km range, indicating a significant spatial correlation characteristic for reactive phosphates within a 10 km range.

Turning to the temporal dimension, it can intuitively be observed in Figure 6b,c,e,f that the seawater quality observation data also show a certain degree of correlation over the years. As the number of years in each time lag increases, the data correlation (gamma) shows some degree of change. This suggests that changes in the seawater quality show trends over time.

3.4. Semivariogram Fitting

The key to fitting a theoretical semivariogram model lies in estimating the range, sill, and nugget values. In this study, the empirical semivariogram and the method of minimizing the Mean Squared Error (MSE) index are used to determine the nugget value, sill value, and range value. Firstly, the empirical semivariogram is used to make an initial estimate of these values. The nugget value can be approximated as the intercept on the y-axis, i.e., when x = 0, the value on the y-axis represents the estimated nugget value.

The sill value corresponds to the y-axis value when the curve tends to stabilize, which typically occurs at a larger distance h, where the semivariogram value no longer increases significantly with increasing distance. The range can be estimated by finding the x-axis value at which the curve rises from the origin to half the sill value. By observing the empirical semivariogram plots, Figure 6a,d, and referring to the methodology shown in Figure 1, an initial range for the parameters is determined. Subsequently, the parameters are fine-tuned and the MSE is minimized to determine the final and accurate values of the range, sill, and nugget.

Figure 7 shows the visualization effect of the theoretical variogram fitted to the empirical semivariogram using a separable spatio-temporal model.

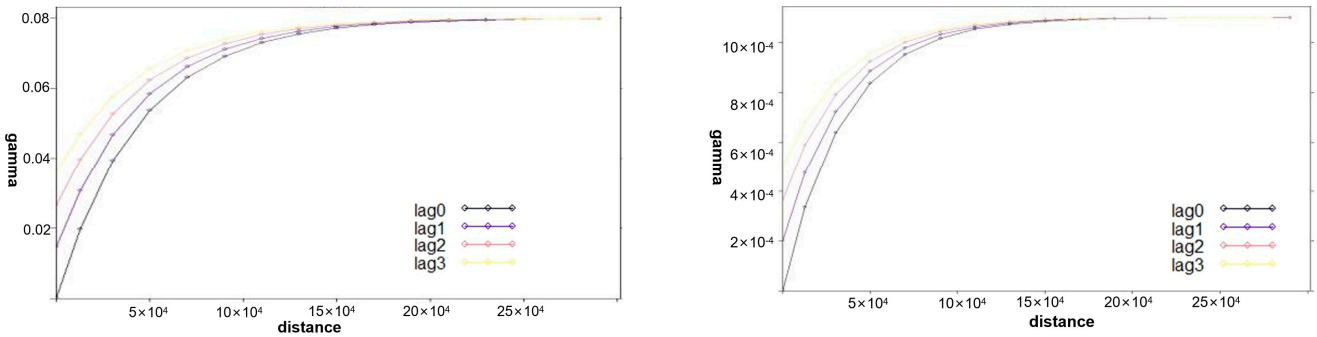

Figure 7. Semivariogram fitting results: (**a**) theoretical variogram function curve fitted for the inorganic nitrogen empirical variogram model; (**b**) theoretical variogram function curve fitted for the reactive phosphate empirical variogram model.

The analysis of the empirical semi-variogram functions of inorganic nitrogen and active phosphate data determined that the parameters of the theoretical variogram function models of inorganic nitrogen and active phosphate are as shown in Table 1.

Table 1. Theoretical variogram function parameters.

Category	Parameters		
	Range	**Partial Sill**	**Nugget**
Inorganic nitrogen (spatial)	80,000 m	0.0001	0.0001
Inorganic nitrogen (temporal)	5 years	0.00005	0.00001
Reactive phosphate (spatial)	40,000 m	0.0001	0.0001
Reactive phosphate (temporal)	5 years	0.00005	0.00001

The different values of the range parameters in Table 1 indicate that the diffusion capabilities of reactive phosphate and inorganic nitrogen in seawater are different.

3.5. Generation of a Prediction Grid

Given that the sea area under evaluation near the coast of Fujian is 36,000 square kilometers and considering the evaluation accuracy requirements of this study, it was determined that the monitoring sea area should be divided into 500 m × 500 m grids. Using the center points of the grids as interpolation points, the spatio-temporal kriging interpolation algorithm was applied to compute the values at unknown points. We used the "raster()" method from the raster package of the R language to generate the prediction grid.

4. Results

4.1. Spatio-Temporal Kriging Interpolation

In this research, we utilize the krigeST method from the gstat package for interpolating monitoring data on bioactive phosphates, inorganic nitrogen, and other substances within

the spatio-temporal domain. KrigeST can manage data points that include both temporal and spatial coordinates, and it can interpolate predictions for locations in time and space that have not been sampled [28]. The annual fluctuations in the concentrations of bioactive phosphates and inorganic nitrogen in the coastal waters of Fujian from 2019 to 2024 are depicted in Figures 8 and 9, respectively. Importantly, these figures include predicted data for the year 2024.

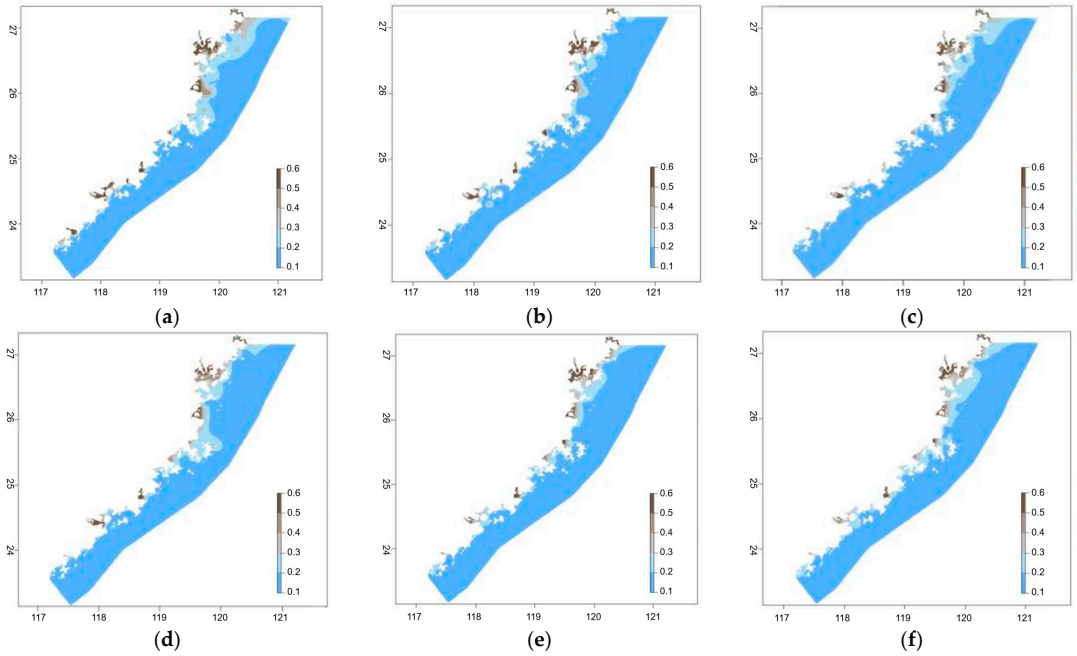

Figure 8. Visualization of changes in the inorganic nitrogen content in the nearshore waters of Fujian Province from 2019 to 2024: (**a**) 2019; (**b**) 2020; (**c**) 2021; (**d**) 2022; (**e**) 2023; (**f**) 2024.

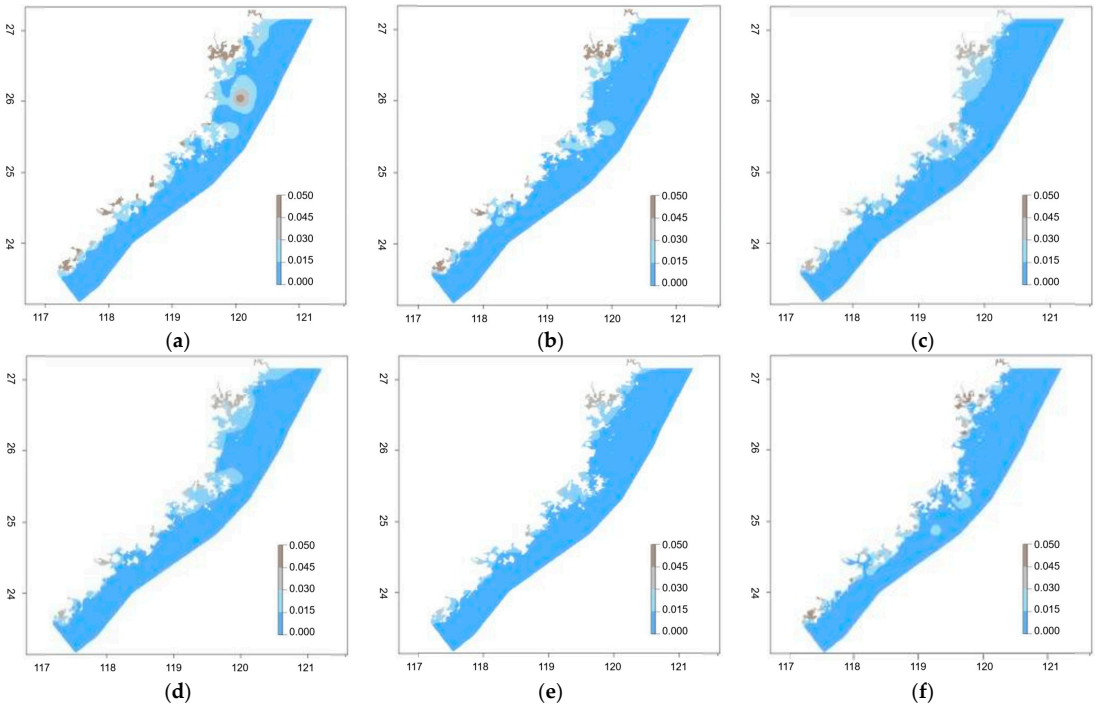

Figure 9. Visualization of changes in the reactive phosphate content in the nearshore waters of Fujian Province from 2019 to 2024: (**a**) 2019; (**b**) 2020; (**c**) 2021; (**d**) 2022; (**e**) 2023; (**f**) 2024.

4.2. Overlay Analysis

According to the National Seawater Quality Standard (GB-3097-1997) [30], seawater quality can be divided into four grades: Grade I, Grade II, Grade III, and Grade IV. The numerical values for inorganic nitrogen and bioactive phosphate indicators are shown in Table 2.

Table 2. Inorganic nitrogen and reactive phosphate concentration standards (mg/L).

No.	Item	Grade I	Grade II	Grade III	Grade IV	Below Grade IV
1	Inorganic nitrogen	≤0.20	>0.20, ≤0.30	>0.30, ≤0.40	>0.40, ≤0.50	>0.50
2	Reactive phosphate	≤0.015	>0.015, ≤0.030		>0.030, ≤0.045	>0.045

The evaluation of seawater quality levels requires comprehensive consideration of various indicators. Since inorganic nitrogen and reactive phosphate are the primary factors that exceed the secondary category in the coastal waters of Fujian, the indicators of inorganic nitrogen and reactive phosphate are superimposed to obtain the cumulative comprehensive evaluation results.

In this study, the *overlay* function in the *raster* package of the R language is utilized to superimpose the data layers of two water quality indicators: inorganic nitrogen and reactive phosphate. Specifically, it reads and compares the interpolation results of each pixel point in two TIFF files (i.e., the concentration values of these two water quality indicators at each geographical location). Then, based on the criteria in Table 2, it determines the seawater quality grade of each pixel point and assigns a numerical value between 1 and 5 to these pixel points as the grade label. In this way, the distribution of seawater quality grades in different regions can be visually observed. Figure 10 shows the visualized results of the comprehensive assessment of seawater quality.

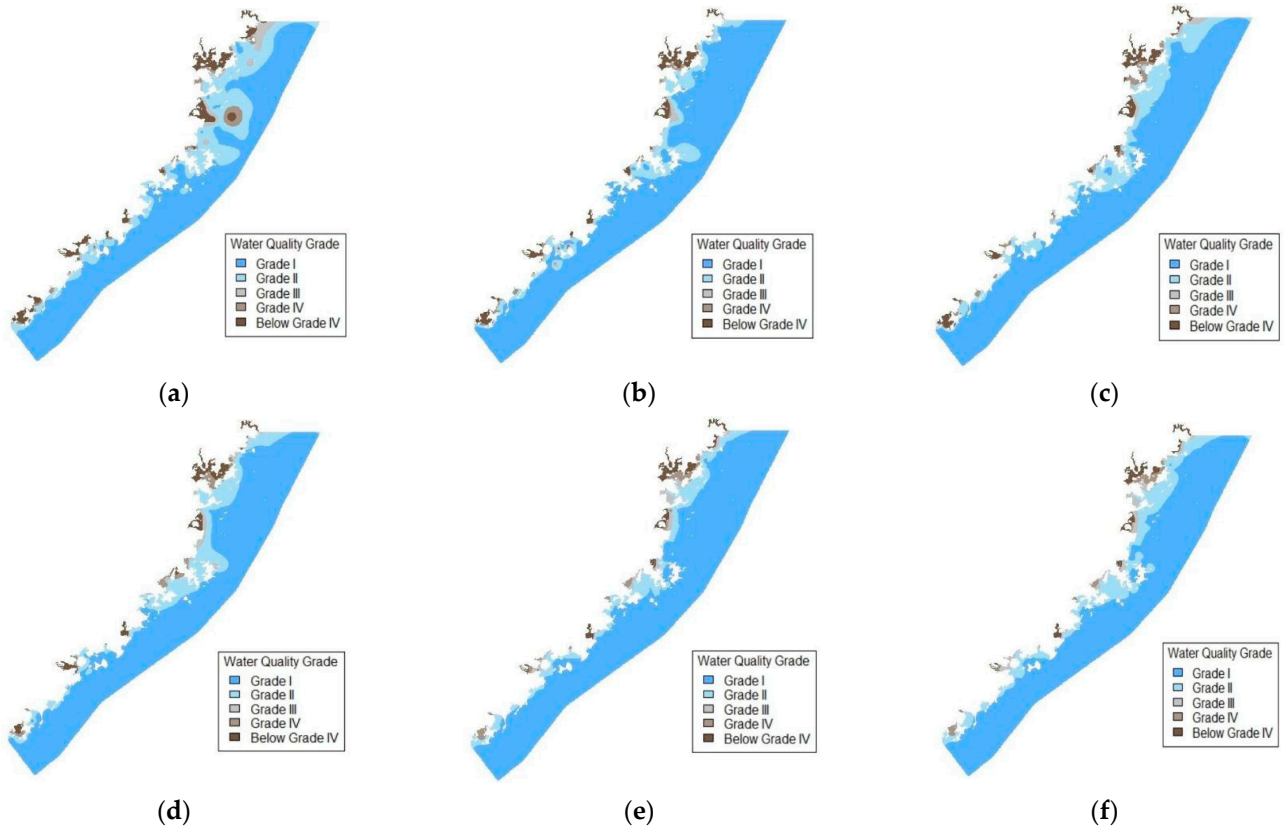

Figure 10. Water quality grade evaluation and prediction: (**a**–**e**) water quality assessments from 2019 to 2023, respectively; (**f**) water quality prediction for 2024.

We extract the number of grid cells representing different seawater quality levels from the raster image in Figure 7 and multiply it by the actual sea area represented by the grid (500 m × 500 m) to calculate the actual area of different classified sea regions. Table 3 shows the actual area conditions of the different classified sea regions in the coastal waters of Fujian Province (annual data for 2019 to 2024).

Table 3. Area of seawater classified by different quality levels (square kilometers).

Year	Grade I	Grade II	Grade III	Grade IV	Below Grade IV
2019	23,270.2	8914.4	1169.2	962.8	2196.8
2020	29,638.6	4415.6	506	437.6	1515.6
2021	27,759.8	6015.6	520.4	734.6	1483
2022	28,129	5877.6	558	698	1250.8
2023	29,907.4	4123.8	807.8	683.6	990.8
2024	27,962.8	6085	967.2	641.2	857.2

Figure 11 shows the trend of seawater quality changes in the coastal areas of Fujian Province from 2019 to 2024. In the figure, it can be seen that the overall trend of seawater quality in the coastal areas of Fujian Province tends to be gentle and generally improving. However, compared to 2023, the forecast for 2024 shows a more obvious decrease in the area of Class I seawater. Correspondingly, the area of class II seawater has increased significantly. This indicates a deteriorating trend in water quality, with coastal marine environmental protection facing more severe challenges.

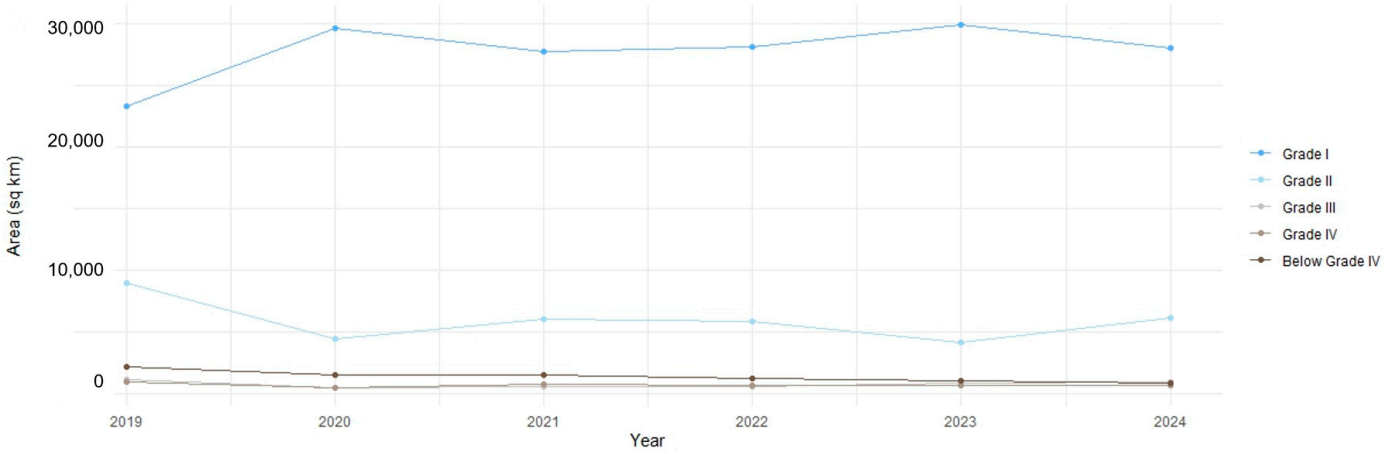

Figure 11. Annual coverage area by water quality grade.

5. Conclusions

The assessment and prediction of seawater quality is of paramount importance in protecting the marine environment, maintaining ecological balance, and achieving sustainable development in the exploitation and use of marine resources. By continuously improving the research and application of assessment and prediction methods, we can effectively guide the scientific management of the marine environment and promote a deeper understanding and protection of our blue planet. The quality of coastal seawater is influenced by both natural factors, such as tides, ocean currents, hurricanes, and river discharges, and anthropogenic factors, such as industrial emissions, agricultural non-point source pollution, domestic sewage, and marine debris. As a result, changes in coastal seawater quality within a region exhibit both regularity and randomness. Given the limited amount of observational data, the use of small-scale data sets to achieve the scientific and quantitative assessment and prediction of coastal seawater quality within a given region has significant practical value in marine environmental management.

This paper presents a seawater quality assessment and prediction method based on the spatio-temporal semivariogram model. This method uses small-scale spatio-temporal data sets to assess and analyze the seawater quality in Fujian coastal waters from 2019 to 2023, and to predict the quality in 2024. By fitting theoretical semivariograms to empirical semivariograms, this method enables the estimation of observed values for indicative seawater quality parameters such as inorganic nitrogen and reactive phosphate at unknown spatio-temporal locations. Through the application of GIS overlay analysis with customized rules, a multi-indicator analysis of seawater quality is achieved, and different seawater quality grades and areas are accurately classified according to industry standards, thereby scientifically achieving the quantitative assessment and prediction of seawater quality and providing scientific guidance for marine environmental protection and management.

An in-depth analysis of the evaluation and prediction results shows that this method accurately captures the trends and randomness of the data across spatial and temporal dimensions, demonstrating its absolute advantage in handling small-scale spatio-temporal data. The spatio-temporal semivariogram deepens the traditional spatial semivariogram by comprehensively incorporating the variability of sample data in both temporal and spatial dimensions. Based on its model theory and parameter foundation, the spatio-temporal kriging interpolation algorithm achieves the optimal linear unbiased prediction of unknown point values. As an innovative application of kriging methods, the spatio-temporal kriging interpolation algorithm takes into account the auto-correlation in both time and space for spatio-temporal data, enabling the accurate prediction of unknown spatio-temporal location values. By exploiting the auto-correlation cues revealed by the spatio-temporal semivariogram, this algorithm can accurately estimate data values at specific spatio-temporal points that have not been directly observed.

The accurate estimation of seawater quality indicators at unknown spatio-temporal locations using this method depends crucially on the precise determination of the range, sill, nugget, and other parameters of the fitted semivariogram. The method described in this paper, which is based on observations of empirical semivariograms and fine-tuning by MSE minimization, requires extensive industry background knowledge and iterative validation processes. Significant research efforts are still needed to accurately and efficiently determine function parameters. In future studies, AI algorithms should be integrated to automate the determination of fitted semivariogram parameters, replacing complex and repetitive manual work and improving parameter accuracy.

In summary, this paper effectively addresses the issues of the quantitative assessment and prediction of seawater quality through the integrated use of semivariogram models and GIS overlay analysis with customized rules. Unlike machine learning and AI techniques that rely on massive sample data, this method can fully reflect the internal regularity and randomness of data in small-scale spatio-temporal datasets, thus providing relatively accurate and scientific quantitative assessment and prediction results.

Author Contributions: Conceptualization, Wei Wang; methodology, Wei Wang; software, Wei Wang; validation, Wei Wang; formal analysis, Wenfang Cheng; investigation, Wenfang Cheng; resources, Wei Wang; data curation, Jing Chen; writing—original draft preparation, Wei Wang; writing—review and editing, Jing Chen; supervision, Wei Wang; project administration, Wei Wang. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Tianjin Jinnan District Bureau of Science and Technology under Grant No. 20220105.

Data Availability Statement: Data is contained within the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Adadzi, P. Geostatistics and Spatial Analysis of Groundwater Hydrochemistry near Leliefontein in the Northern Cape, South Africa. *Inż. Ekol.* **2020**, *21*, 243–260. [CrossRef]
2. Ccoica-López, K.; Pasapera-Gonzales, J.; Jimenez, J. Spatio-Temporal Variability of the Precipitable Water Vapor over Peru through MODIS and ERA-Interim Time Series. *Atmosphere* **2019**, *10*, 192. [CrossRef]
3. De Faveri, J.; Verbyla, A.P.; Culvenor, R.A. Modelling Spatial and Temporal Correlation in Multi-Assessment Perennial Crop Variety Selection Trials Using a Multivariate Autoregressive Model. *Crop Pasture Sci.* **2023**, *74*, 1142–1155. [CrossRef]
4. Deng, X. Influence of Water Body Area on Water Quality in the Southern Jiangsu Plain, Eastern China. *J. Clean. Prod.* **2020**, *254*, 120136. [CrossRef]
5. Deogharia, R.; Sil, S. Reconstructing High-Frequency Radar Derived Ocean Surface-Current Fields Using Spatio-Temporal Kriging. *IEEE J. Ocean. Eng.* **2023**, *48*, 1289–1299. [CrossRef]
6. Frigstad, H.; Andersen, G.S.; Trannum, H.C.; McGovern, M.; Naustvoll, L.-J.; Kaste, Ø.; Deininger, A.; Hjermann, D.Ø. Three Decades of Change in the Skagerrak Coastal Ecosystem, Shaped by Eutrophication and Coastal Darkening. *Estuar. Coast. Shelf Sci.* **2023**, *283*, 108193. [CrossRef]
7. Gräler, B.; Pebesma, E.; Heuvelink, G. Spatio-Temporal Interpolation Using Gstat. *R J.* **2016**, *8*, 204. [CrossRef]
8. Guo, J.; Lee, J.H.W. Development of Predictive Models for "Very Poor" Beach Water Quality Gradings Using Class-Imbalance Learning. *Environ. Sci. Technol.* **2021**, *55*, 14990–15000. [CrossRef] [PubMed]
9. Haghiabi, A.H.; Nasrolahi, A.H.; Parsaie, A. Water Quality Prediction Using Machine Learning Methods. *Water Qual. Res. J. Can.* **2018**, *53*, 3–13. [CrossRef]
10. Hatzikos, E.V.; Tsoumakas, G.; Tzanis, G.; Bassiliades, N.; Vlahavas, I. An Empirical Study on Sea Water Quality Prediction. *Knowl. Based Syst.* **2008**, *21*, 471–478. [CrossRef]
11. Johannesson, K.; Leder, E.H.; André, C.; Dupont, S.; Eriksson, S.P.; Harding, K.; Havenhand, J.N.; Jahnke, M.; Jonsson, P.R.; Kvarnemo, C.; et al. Ten Years of Marine Evolutionary Biology—Challenges and Achievements of a Multidisciplinary Research Initiative. *Evol. Appl.* **2023**, *16*, 530–541. [CrossRef] [PubMed]
12. Johnson, D.C.; Enriquez, C.E.; Pepper, I.L.; Davis, T.L.; Gerba, C.P.; Rose, J.B. Survival of Giardia, Cryptosporidium, Poliovirus and Salmonella in Marine Waters. *Water Sci. Technol.* **1997**, *35*, 261–268. [CrossRef]
13. Júnez-Ferreira, H.E.; Hernández-Hernández, M.A.; Herrera, G.S.; González-Trinidad, J.; Cappello, C.; Maggio, S.; De Iaco, S. Évaluation Des Changements Dans Les Niveaux Régionaux Des Eaux Souterraines Par Krigeage Spatio-Temporel: Application Au Système Aquifère Du Sud Du Bassin de Mexico. *Hydrogeol. J.* **2023**, *31*, 1405–1423. [CrossRef]
14. Landrigan, P.J.; Stegeman, J.J.; Fleming, L.E.; Allemand, D.; Anderson, D.M.; Backer, L.C.; Brucker-Davis, F.; Chevalier, N.; Corra, L.; Czerucka, D.; et al. Human Health and Ocean Pollution. *Ann. Glob. Health* **2020**, *86*, 151. [CrossRef]
15. Sohrabian, B. Geostatistical Prediction through Convex Combination of Archimedean Copulas. *Spat. Stat.* **2021**, *41*, 100488. [CrossRef]
16. Gribov, A.; Krivoruchko, K. Empirical Bayesian Kriging Implementation and Usage. *Sci. Total Environ.* **2020**, *722*, 137290. [CrossRef]
17. Ludwig, G.; Zhu, J.; Reyes, P.; Chen, C.-S.; Conley, S.P. On Spline-Based Approaches to Spatial Linear Regression for Geostatistical Data. *Environ. Ecol. Stat.* **2020**, *27*, 175–202. [CrossRef]
18. Salvaña, M.L.O.; Lenzi, A.; Genton, M.G. Spatio-Temporal Cross-Covariance Functions under the Lagrangian Framework with Multiple Advections. *J. Am. Stat. Assoc.* **2023**, *118*, 2746–2761. [CrossRef]
19. Akram, W.; Hussain, A.; Ali, S.; Hussain, I.; Muhammad, M. Geostatistical Analysis of Spatio-Temporal Variability and Mapping Genus Bactrocera in Apricot Orchard in Northern Pakistan. *Pak. J. Zool.* **2022**, *55*, 765–771. [CrossRef]
20. Awais, M.; Arshad, M.; Shah, S.H.H.; Anwar-ul-Haq, M. Evaluating Groundwater Quality for Irrigated Agriculture: Spatio-Temporal Investigations Using GIS and Geostatistics in Punjab, Pakistan. *Arab. J. Geosci.* **2017**, *10*, 510. [CrossRef]
21. Wikle, C.K.; Zammit-Mangion, A.; Cressie, N. *Spatio-Temporal Statistics with R*; CRC Press: London, UK, 2019; ISBN 9781138711136.
22. Montero, J.-M.; Fernández-Avilés, G. Functional Kriging Prediction of Atmospheric Particulate Matter Concentrations in Madrid, Spain: Is the New Monitoring System Masking Potential Public Health Problems? *J. Clean. Prod.* **2018**, *175*, 283–293. [CrossRef]
23. Lambardi di San Miniato, M.; Bellio, R.; Grassetti, L.; Vidoni, P. Separable Spatio-temporal Kriging for Fast Virtual Sensing. *Appl. Stoch. Models Bus. Ind.* **2022**, *38*, 806–829. [CrossRef]
24. Kyprioti, A.P.; Irwin, C.; Taflanidis, A.A.; Nadal-Caraballo, N.C.; Yawn, M.C.; Aucoin, L.A. Spatio-Temporal Storm Surge Emulation Using Gaussian Process Techniques. *Coast. Eng.* **2023**, *180*, 104231. [CrossRef]
25. Zheng, D.; Chen, D.; Lin, Q. Research on the Influencing Factors of the Development Level of Marine Economy in Fujian Province. *J. Coast. Res.* **2020**, *115*, 434. [CrossRef]
26. Hu, Y.; Zhou, H.Q.; Jin, X.M.; Shen, Y.F.; Yan, Y.Z. Assessing the Resilience of the Marine Economy: A Case Study of Southern China's Marine Economy Circle. *Front. Mar. Sci.* **2022**, *9*, 912462. [CrossRef]
27. Shi, X.; Ye, X.; Zhong, H.; Wang, T.; Jin, F. Sustainable Nitrogen-Containing Chemicals and Materials from Natural Marine Resources Chitin and Microalgae. *Mol. Catal.* **2021**, *505*, 111517. [CrossRef]
28. Mukherjee, S.; Huang, X.; Udpa, L.; Deng, Y. A Kriging-Based Magnetic Flux Leakage Method for Fast Defect Detection in Massive Pipelines. *J. Nondestruct. Eval. Diagn. Progn. Eng. Syst.* **2022**, *5*, 011002. [CrossRef]

29. Pebesma, E.J.; Wesseling, C.G. Gstat: A Program for Geostatistical Modelling, Prediction and Simulation. *Comput. Geosci.* **1998**, *24*, 17–31. [CrossRef]
30. (GB 3097-1997 replacing GB 3097-82 putting into effect as of Jul 1, 1998) Sea Water Quality Standard. Available online: https://english.mee.gov.cn/Resources/standards/water_environment/quality_standard/200710/t20071024_111791.shtml (accessed on 29 July 2024).