DOI: 10.2478/eces-2024-0031 ECOL CHEM ENG S. 2024;31(4):473-491

Jarosław ZAWADZKI[1*] and Piotr FABIJAŃCZYK[1]

APPLICATION OF GEOSTATISTICAL METHODS IN THE STUDY OF ALGAE IN WATER RESERVOIRS: A REVIEW

Abstract: The growing pollution of aquatic environments, primarily of anthropogenic origin, combined with global climate change, has led to significant increases in eutrophication. This process often results in harmful algal blooms (HABs) of phytoplankton and algae in various water bodies, including inland lakes, marshes, rivers, seas, and oceans. These blooms pose a serious threat not only to aquatic ecosystems but also to human health. Understanding phytoplankton and algal blooms is inherently complex, as these phenomena manifest on multiple spatial and temporal scales. Comprehensive studies of phytoplankton and algae require the collaboration of scientists from diverse scientific disciplines, including biology, ecology, and environmental science. One of the critical tools in this multidisciplinary approach is geostatistics, an advanced and continuously evolving branch of statistics that specialises in analysing spatial and temporal phenomena. Geostatistics is particularly well-suited for the study of phytoplankton and algal blooms due to its ability to handle data that varies across different scales and locations. This review presents and discusses selected studies that employ geostatistical methods to investigate plankton and algae in various water bodies. It highlights the most significant scientific works that, in the authors' opinion, represent milestones in the application of these studies. Furthermore, various geostatistical methods are explored, ranging from variography to spatiotemporal modelling, providing insights into spatial and temporal patterns and their variability of phytoplankton and algal blooms in aquatic ecosystems.

Keywords: phytoplankton, algae, algae blooms, geostatistics, spatial statistics, water reservoirs, time-space phenomena, kriging, co-kriging, environment, water pollution, climate change

Introduction

Algae living in both marine and inland waters are crucial elements of ecosystems at various spatial scales. Algae show great morphological diversity and can be classified into two main groups: macroalgae and microalgae, i.e. phytoplankton. They are the most important primary producers in most water bodies, assimilating carbon dioxide from the atmosphere in the process of photosynthesis. It is estimated that most of the Earth's atmospheric oxygen is produced by the ocean algae [1, 2]. They play a fundamental role in the global carbon cycle [3], and are food sources for many marine organisms, including whales. It is crucial for many biogeochemical transformations, such as nitrogen and sulphur circulation, and affects the quality and health of aquatic ecosystems. Excessive water pollution, often of anthropogenic origin, both industrial and agricultural, climate change, and the associated increase in eutrophication of the aquatic environment may lead to phytoplankton and algae blooms. Such blooms reduce the amount of available oxygen in the water and may also produce toxins that are harmful to aquatic animals and humans.

[1] Faculty of Building Services, Hydro and Environmental Engineering, Warsaw University of Technology, ul. Nowowiejska 20, 00-661 Warszawa, Poland, ORCID: JZ 0000-0003-2842-0018, PF 0000-0002-0769-2324
[*] Corresponding author: jaroslaw.zawadzki@pw.edu.pl

Harmful Algae Blooms (HABs) are a rapidly increasing threat to ecosystems and human health occurring in many countries [4-10]. An example of the complexity of problems related to HABs may be the still insufficiently effective attempts to prevent the spread and control of Golden Algae in both marine and inland waters [11-13]. Toxins released by this type of algae can, under certain conditions, cause the death of many aquatic organisms in large areas, leading to huge economic losses. Phytoplankton and algal blooms are complex spatial and temporal phenomena that require the cooperation of scientists representing many scientific disciplines, including biologists, ecologists, geneticists, geophysicists, geochemists, hydrologists, specialists in remote observations, including satellite ones, marine ecologists, climate scientists, specialists in geo and bioinformatics. However, too little attention is still paid to the possibilities of using modern geostatistical methods, which may be particularly useful in the study of phytoplankton and algae in water reservoirs. The work aims to review the use of geostatistical and related methods in the studies of phytoplankton and algae in water reservoirs, and thus to stimulate the use of these methods in HABs studies. Due to the enormous scope of research on phytoplankton and algae in the aquatic environment, they will not be discussed in this article. In turn, the description of geostatistical methods developed since the mid-twentieth century and used in many fields, including geology and environmental studies will necessarily be limited to the minimum necessary, in such a way that the interested reader who does not know these issues well can understand their essence to such an extent as to follow their specific applications at the interface of these two fields. In return, chosen bibliographic items in the field of geostatistics will be provided. It is worth emphasising that geostatistical methods are constantly being intensively developed, which gives new impetus to their applications also in studies of algae and phytoplankton.

Outline of geostatistical methods

The vast majority of environmental datasets demonstrate spatial continuity. In contrast, conventional statistical methods often disregard spatial information, despite its crucial relevance to various natural phenomena. This pivotal information is frequently overlooked in analyses, leading to its complete loss. An autonomous branch of statistical methods, known as geostatistics, specialises in studying random phenomena while considering their spatial dimension.

Geostatistics uses "regionalised variables" related to location. Regionalised variables have properties intermediate between a random variable and a deterministic variable. Geostatistics proposes methods for describing spatial continuity and uses statistical relationships to study spatial phenomena. The basic issues that geostatistics deals with include the interpretation and prediction (estimation or simulation) of the spatial distribution of the studied phenomena. Geostatistics is therefore a natural complement to traditional statistical analysis, especially useful in various environmental studies, in which the spatial aspect most often plays an important role. Geostatistics has unique methods, e.g. cokriging methods, which allow the use of measurements of one or several additional variables to improve the quality of estimating the main variable under study. This means that geostatistical methods allow, for example, the integration of various types of environmental measurements performed using various measurement techniques or the integration of quantities of different natures or originating from different sources. For example, using cokriging methods, geochemical measurements can be integrated with geophysical and even biological measurements. You can also, for example, integrate direct measurements with remote measurements performed from a certain distance, e.g. physical measurements performed in the field with aerial or satellite measurements. One of the intensively developing branches of geostatistics is the integration of measurements, taking into account their temporal and spatial nature. An extremely important advantage provided by geostatistical methods is the ability to integrate historical environmental data with current data under certain conditions. Measurements of the same quantities are very often performed in independently conducted studies by different research teams. It should be emphasised, however, that due to the completely different dependencies of natural data on spatial coordinates and time, this is a very complex issue that should not be confused with the automatic extension of the analysis to another dimension. Geostatistics has special, advanced methods for analysing spatio-temporal regionalised variables.

From the point of view of environmental research including algae and phytoplankton studies, geostatistical methods have another extremely valuable advantage, namely they allow for a precise assessment of the error of the estimates made, as well as the spatial distribution of this error. For the above-mentioned reasons, these methods have numerous practical applications wherever it is necessary to study spatial phenomena. These include, among others: various studies of the natural environment described, e.g. in [14-19]. The central tool of geostatistics is the semivariance (semivariogram, variogram) function (Fig. 1), which is a measure of spatial continuity. The following formula is used for the semivariance calculations:

$$\gamma(\boldsymbol{h}) = \frac{1}{2N} \sum_{i=1}^{N} [Z(\boldsymbol{x}_i + \boldsymbol{h}) - Z(\boldsymbol{x}_i)]^2 \tag{1}$$

In the above equation, $\boldsymbol{x}_i$ represents the data location, $\boldsymbol{h}$ stands for a lag vector, $Z(\boldsymbol{x}_i)$ denotes the data value at location $\boldsymbol{x}_i$, and N refers to the quantity of data pairs separated by a distance and direction of h units.

Semivariance functions are usually characterised by three parameters:

1) Range (range of influence or correlation), A - the distance above which two data are uncorrelated.
2) Nugget effect, C_0 - the vertical discontinuity at the origin. The nugget effect is a combination of sampling error and short-scale variation that occurs at a scale smaller than the closest sample spacing.
3) Sill, C - the plateau that the semivariogram reaches. Sill is equal to the nugget effect, C_0. Structured variability is equal to $C_1 = C - C_0$ [15, 19].

There are also other more specific measures of spatial continuity such as the covariance function, correlation function, madogram, and many others. Comprehensive data on the topics analysed here can be found in the literature [11, 12, 19-21]. The obtained experimental semivariogram is used to fit an appropriate theoretical model, such as the spherical, exponential, etc. [15, 19, 21, 22]. The main geostatistical method is kriging [23, 24]. It produces minimum-variance estimates Z^* taking into account the spatial (spatiotemporal) structure of the variable, represented by the semivariogram. Kriging uses a linear estimator given by the following equation:

$$Z^* = \sum_{i=1}^{n} \lambda_i z_i \tag{2}$$

where λ_i are the weights, and z_i are the known data values. The kriging weights are computed by minimising the estimation variance by applying the non-biased condition. The quality of point estimation methods can be determined using various criteria. The kriging method, known as B.L.U.E. signifies the best, linear, and unbiased estimation approach. Kriging employs linear, weighted data combinations and aims for zero mean errors, ultimately minimising error variance [23].

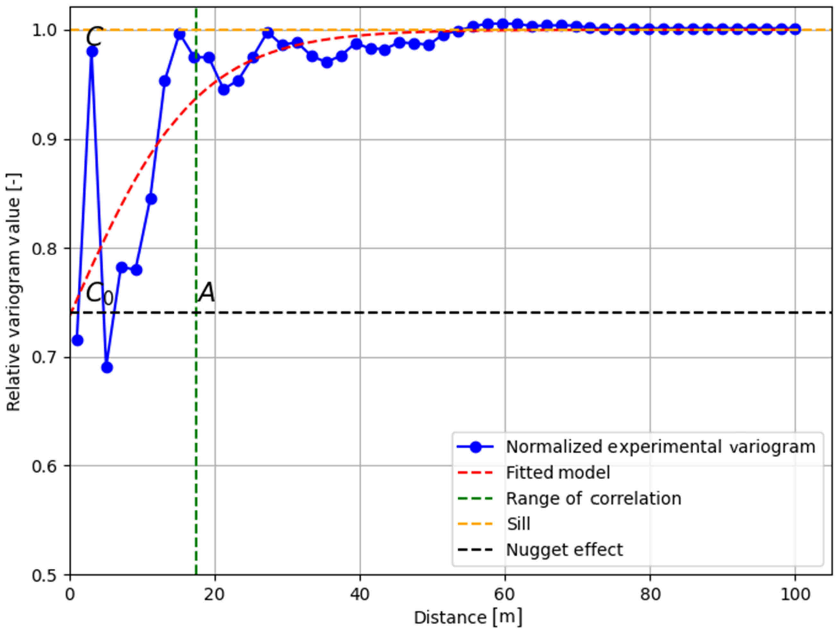

Fig. 1. An exemplary semivariogram. The blue solid line and dots represent the empirical semivariance, and the dashed red line shows the model of this semivariance

Many other types of kriging have also been developed, including cokriging, which uses auxiliary variables to estimate the main variable. This method minimises the variance of estimation errors of the main variable by using cross-correlations between this variable and the secondary variable(s). A spatial measure of variability between two variables Y and Z is analogous to semivariance is the cross-semivariance given by the equation:

$$\boldsymbol{\gamma}(\boldsymbol{h}) = \sum_{i=1}^{N} [(\boldsymbol{Y}(\boldsymbol{x}_i) - \boldsymbol{Y}(\boldsymbol{x}_i + \boldsymbol{h})) \cdot (\boldsymbol{Z}(\boldsymbol{x}_i) - \boldsymbol{Z}(\boldsymbol{x}_i + \boldsymbol{h}))] \tag{3}$$

Other important types of kriging often used in environmental studies are kriging with a trend model (universal), kriging with an external drift, regression kriging, block kriging, indicator kriging, probability kriging, or disjunctive kriging. Due to lack of space, they cannot be discussed in detail here. However, information about these types of kriging can be found in numerous books and publications, e.g. [15, 23, 25-30]. Yet, it is worth mentioning that in recent decades, geostatistical methods have developed rapidly in many different directions. For the discussed research on phytoplankton and algae, the development of spatiotemporal geostatistical methods should be emphasised, which enable estimation and simulation of the studied phenomena both in time and space, as well as the development of geostatistical methods based on the Bayesian methodology [19, 30-32]. Both spatiotemporal and Bayesian modelling is computationally intensive. These difficulties were largely overcome thanks to the use of Markov Chain Monte Carlo methods (MCMC) [32] implemented e.g. in the WinBUGS software [33, 34]. Over the last dozen or so years, the Integrated Nested Laplace Approximation (INLA) approach has been developed for these purposes, which is even more computationally efficient than MCMC [35, 36]. The development of the above-mentioned methods allowed for wider use of geostatistical methods in complex environmental problems, including the first works on the spatiotemporal dynamics of algae in large water reservoirs. It should also be added that cokriging methods using the above-mentioned methods play and will continue to play a particularly important role in phytoplankton research. methodology. This is due to both the purpose of these methods, which is the integration of various data sets, and the specificity of phytoplankton and algae research, which uses a variety of very different data sets, such as measurements on the surface and in the water column, satellite observations, measurements from various centers, and even performed at different times. It can be expected that the above-mentioned methods, combined with various observation techniques, will lead to great progress in the study of phytoplankton and algae.

Results and discussion

From the regional to the global scale, environmental and technological context

The most characteristic studies of chlorophyll or algae in water reservoirs using geostatistical methods will be described below. Until the beginning of the twenty-first century, geostatistical studies of phytoplankton and algae could be considered an initial stage, and studies based on in-situ measurements at the local scale dominated. The main methods in this period were variogram analysis and kriging. Due to the diversity of environmental conditions and spatial scales, it was decided to arrange the presented research chronologically. This choice seemed the most natural, and it also allowed the reader to trace the development of geostatistical methods themselves which in turn enabled more advanced applications in the discussed subject. Gradually, more advanced geostatistical methods are emerging that allow the integration of various surface or underwater measurements with those made from drones or satellites. At the beginning of the twenty-first century, research on phytoplankton and algae was intensified on larger scales, from the regional to the global scale. The reason for this was, among others, the rapid development of satellite techniques that made it possible to observe algae and algae blooms in large areas of seas and oceans, as well as in coastal zones. At the same time, there was rapid technological development of various water measurements, which provided more in-situ data. This significantly increased the amount of data processed and the need to use greater computing power, although traditional geostatistical methods were still widely used. At the same time, geostatistical methods were still used in the study of smaller water reservoirs. Development in this area resulted from the improvement of measurement techniques, which allowed for more detailed research, and their extension to new research areas, such as climate change, rapid spread of HABs, etc. The most important works from this period will be presented below.

From early variography through kriging and cokriging to advanced geostatistical simulations

The presented articles concern research carried out in 1) inland lakes and freshwater marshes, 2) rivers, and 3) seas, oceans, and coastal waters using methods that can be broadly divided as a) variogram analyses (variography), b) geostatistical estimations spatial and c) geostatistical spatial simulations. Therefore, to facilitate reading the article, the presented publications are grouped accordingly in Tables 1 and 2. The most important measurement techniques are given in round brackets below each work.

Wu and Mitsch [37] studied spatial and temporal patterns of metaphyton, periphyton, phytoplankton, and water-column primary productivity in two newly constructed riparian freshwater marshes at the Olentangy River Wetland Research Park at The Ohio State University in Columbus, Ohio, USA. The first one was planted with macrophytes and the second one was unplanted. During the period spanning July through November 1994, researchers sampled surface cover, mid-summer metaphyton biomass, periphyton and phytoplankton biomass and production, and water-column primary productivity. Subsequently, the variogram and correlogram were employed to determine both the spatial pattern and temporal changes of metaphyton. The authors found that the detected spatial arrangement and its fluctuations significantly impact the spatial distribution of periphyton, phytoplankton, and water-column primary productivity by altering nutrient distribution. The way nearby algae and snail grazing interact could be a crucial factor causing differences in periphyton distribution, even in uniform environments. Researchers explored this interaction by observing the distribution of filamentous algae in controlled tanks, with and without grazing snails [38]. Analysing the data using semivariograms they revealed that higher grazing pressure led to increased spatial variation. Interestingly, the size of these varied patches did not simply align with the number of snails present. Instead, there was a tendency for patch size to grow with snail density. The study also employed computer simulations to better understand the factors influencing spatial heterogeneity. Simulation outcomes indicated that grazing effects tend to concentrate grazers in specific areas, forming patches. Additionally, the findings suggested that neighbouring periphyton's beneficial impact on algae growth could contribute to the formation of these patches. Simulation outcomes indicated that grazing effects tend to concentrate grazers in specific areas, forming patches. Additionally, the findings suggested that neighbouring periphyton's beneficial impact on algae growth could contribute to the formation of these patches. A similar work from the geostatistical point of view was published much later by Yan et al. [10], in which authors utilised the kriging method to assess the Qingshan reservoir's eutrophication level, water quality, and planktonic algal communities.

An example of a similar type of research is also the work of Zhao and Cai [39] in which chlorophyll level in freshwater ecosystems was analysed. The study was conducted at Meiziya Reservoir in Hubei province, China. This reservoir is a third-order tributary to the Yangtze River. The area of Meiziya Reservoir is about 30 ha with a mean depth ranging between ca 4 m and 10 m. Standardised directional variograms of the chlorophyll content in a water reservoir and isotropic variograms based on 43 sampling points in three consecutive months May, June, and July 1997 were investigated. In these months, the highest annual rainfall from 1100 mm to 1300 mm, and the largest changes in the water level in the water reservoir were observed. The study also included a fractal analysis of Chl-a. This was carried out based on a variogram which allowed the calculation of the fractal dimension, which is a measure of the spatial complexity of the examined object (Chl-a). The fractal dimension, D is determined from the slope of the logarithm of the semivariogram and the logarithm of distance, [39, 40] based on the following relationship:

$$D = 2 - \frac{m}{2} \tag{4}$$

where D is the Hausdorff-Besicovitch statistic and m is the slope of a variogram on $\log\gamma$ - $\log h$ graph (γ is the variogram, and h is the distance). The fractal dimension of flat objects takes values from 0 to 2. Spatial heterogeneity is strong if $1 \le D < 1\,{}^{1}/_{3}$, moderate if $1\,{}^{1}/_{3} \le D < 1 < {}^{2}/_{3}$, and weak if $1\,{}^{2}/_{3} \le D < 1 \le 2$.

Thus, geostatistical analysis from a relatively small number of measurement points allowed authors to obtain kriging maps for the three months examined. Moreover, the analysis of semivariance parameters and fractal dimension allowed them to find a relationship between the spatial distribution of chlorophyll in water and precipitation. A noteworthy observation emerged regarding the surface distribution of chlorophyll: it transitioned from aggregated to random following minor rainfall events, reverting to an aggregated distribution after a substantial rainfall event [39].

Nonetheless, constructing a comprehensive model for phytoplankton content in water demands a distinct consideration of processes occurring horizontally and vertically. This is due to variations in factors influencing phytoplankton, like temperature, water mixing, and nutrient levels, which differ significantly in these directional aspects. To better account for these processes, Welty and Stein developed a model of phytoplankton levels based on covariance and variogram functions [41]. They used in situ chlorophyll fluorescence measurements [42] to evaluate phytoplankton levels. Analysing this data posed challenges, notably in determining and specifying the covariance model within Michigan Lake. The measurements, conducted about six times annually from 1998 to 2000, were part of the EEGLE (Episodic Events Great Lakes Experiment Program) [41, 43]. The authors developed the appropriate covariance models in horizontal and vertical directions. Their findings highlighted the crucial role of accurately interpolated values for chlorophyll fluorescence. These values play a decisive role in calibrating fluorescence data with water samples, enabling the precise determination of chlorophyll content distribution in water reservoirs. An interesting application of geostatistical methods in the study of marine plankton was presented by Muller [44]. He studied the dynamics of phytoplankton distribution in the coastal waters of the North Sea using low-resolution MERIS satellite images and their algal_2 product. Along the southern coast, high chlorophyll concentrations prevail throughout the year. Phytoplankton blooms were observed in spring and autumn. The satellite observations were hampered by cloud cover. Kriging was used to estimate missing satellite data on the chlorophyll content in water. To meet the assumptions of the intrinsic hypothesis [15, 19], kriging was not directly on the data, but on the residuals after subtracting the spatial trend. Therefore, the chlorophyll estimation process could include only locations where a trend was available. The authors showed that the range of kriging variance obtained through cross-validation characterises the data distribution and can help identify those estimation locations that are either surrounded by the data points or lie close to the data points. Moreover, the examined example showed the number of reliable estimates resulting from the spatial distribution of the above-mentioned approaches. A year later, Wang and Liu [45] applied two-dimensional block kriging and spatial structure analysis to investigate severe eutrophication of Chl-a in Taihu Lake (China), which had been under significant anthropogenic pressure from rapid economic growth and increasing pollution in the Taihu Catchment. Additionally, an eutrophication assessment map was generated using water quality evaluation criteria. Recently, Zhang et al. investigated the spatial distribution of Chl-a, focusing specifically on the occurrence and movement of a localised eutrophication area in Lake Hongze, using the ordinary kriging method [46].

A significant advancement in the early application of geostatistical methods to algal studies was achieved by Ludovisi et al. [47]. They investigated the spatial autocorrelation structures of plankton population data from Lake Trasimeno (Italy), implemented a multivariate spatial model for the joint study of plankton species, and developed an ecological interpretation of the spatial distributions of the emerging factors. Notably, they employed a geostatistical linear factor model to analyse the horizontal spatial distribution of plankton. Their approach effectively captured the multivariate spatial structure of the plankton community across various scales.

In the study by Bucas [48] research on the attached form of *F. lumbricalis* in the hostile environment of the exposed Baltic Sea coast of Lithuania was carried out. A combination of remote underwater video and novel parameters such as the minimum distance between the seabed and holdfast of alga on substrate made it possible to reveal the effects of hydrodynamic on the red alga as one of the main factors limiting the species distribution in the studied area. Among others, the impact of environmental factors on the shape distribution of *F. lumbricalis* within a local habitat (at the microscale, i.e., 10 cm^2 - 100 m^2) and between habitats within a coastal zone (at the mesoscale, i.e., 1000 m^2 - 10 km^2) was studied as well as changes in the species distribution and standing stock. The continuous distribution of *F. lumbricalis,* both occurrence and cover, was modelled using two methods, namely, geostatistical interpolation (natural neighbour interpolation) and statistical approach (regression analysis). All predictions were performed in rasters with pixel dimensions of 50 m x 50 m. Spatial models predicting the occurrence and cover of *F. lumbricalis* were chosen from simulations using various methods based on their superior goodness-of-fit scores. These scores were evaluated using a cross-validation approach, where subsets comprising an equal number of randomly selected data points were created from the entire dataset.

Examining the spatial heterogeneity within intertidal systems at two South African sites, Diaz et al. [49] focused on analysing the variations observed among different shore levels and sites, particularly concerning macroalgae and grazers. The analysis involved using semivariograms and fractal analysis on data gathered from transects. Additionally, researchers examined the correlations between algae and grazers using cross-semivariograms. This work is particularly important because only a limited number of studies have investigated the spatial structure within intertidal systems. Moreover, these studies demonstrated the particular usefulness of geostatistical methods in this type of research. The key findings highlighted that specific organisms, such as algal groups and grazers, displayed varying spatial distributions at shorter distances within the intertidal area. Moreover, a decrease in this spatial variability within the intertidal was observed due to a cascading effect driven by perturbations.

Tapia et al. [50] described the development of spatial patterns depicting algal species richness across two different spatial extents: regionally covering the Gulf of Mexico and the Caribbean, and globally across coastal zones.

To do so the authors used ordinary kriging and the inverse distance method, used for the sake of comparison. The appropriates of these methods to generate the spatial layers was evaluated using the leave-one-out cross-validation technique. Although the validation results were not ideal, it was possible to determine the spatial variability of the process and produce the studied spatial distributions.

Buelo et al. [51] tested the applicability of spatial indicators to algal blooms using a nutrient-phytoplankton spatial model developed in [52]. They found that standard deviation and autocorrelation successfully distinguished bloom state and proximity to transitions, while skewness and kurtosis were more ambiguous. Thus the results obtained suggest that certain spatial indicators apply to aquatic ecosystems even if dynamic physical-biological interactions reduce detectable signals. The authors concluded that the growing capacity to collect spatial data on algal biomass presents an exciting opportunity for the application and testing of spatial indicators to the study and management of blooms.

Fox [53] utilised chlorophyll fluorescence to assess the variability of phytoplankton biomass and productivity in the northwest European shelf seas. He investigated the variability of phytoplankton physiology and biomass across a range of light and nutrient environments. He found that growth and physiological responses differ in geographically distinct regimes of nutrient limitation across the North Sea as well that variable nutrient stoichiometry across the region could therefore lead to alternative growth and productivity rates that reflect the populations suited to the different conditions. Besides, through geostatistical approaches, the spatial distribution of phytoplankton was characterised during a spring bloom. Fox showed that both in situ and satellite sampling approaches possess the ability to capture mesoscale variability in phytoplankton distribution, but ocean colour estimates lose accuracy in highly heterogenous conditions.

Kim et al. [54] employed geographically weighted regression to extract the spatial pattern of cyanobacterial concentration from satellite imagery data and calibrate a 2D numerical model in South Korea's Nakdong River. The authors showed that the geographically weighted regression model yielded more accurate results than a traditional multiple linear regression one. The calibrated 2D numerical model accurately reproduced the buildup of cyanobacteria along the shoreline. This accumulation was due to favorable conditions for cyanobacterial growth, stemming from extended water residence times linked to decreased flow velocity. Through simulations investigating the changes of HABs using the refined 2D numerical model, a significant reduction was observed in both the blooming area and the maximum cyanobacterial concentration within the study area. This decline correlated with an increased inflow discharge originating from a weir located upstream of the study zone.

The substantial increase in inflow discharge to nearly sixfold its initial level resulted in the almost complete elimination of the cyanobacterial blooming area. This was primarily due to a decreased water residence time resulting from reduced flow velocity. Furthermore, the escalated shear velocity caused by rapid flows effectively restrained the localised expansion of cyanobacteria in shallow, slow-moving zones within the study area by facilitating the lateral spread of water characteristics.

Significant research was done by Pinkerton et al. [55] who used low resolution satellite data of NASA's MODIS-Aqua to observe a large coastal mussel farm to evaluate its impact on water quality, focusing on Chl-a concentration (as a phytoplankton proxy), turbidity, and sea surface temperature [55]. Utilising approximately 890 single-pass satellite images, each covering less than 50 %, they analysed data from 2002 to 2017. To determine the impact of the farm, the mussel farm area in the analysed satellite images was removed and this area was then filled using the kriging method with parameters obtained from semivariogram analysis of each remaining part of the satellite imagery using linear and exponential models. The authors concluded that the approach could be potentially used for other low-resolution satellite imagery.

Son et al. [56] demonstrated the usefulness of spatial statistics for anticipatory actions to massive HABs prevention, through which before the algal bloom season one can identify prone areas based on spatial estimations. The study was carried out the downstream of Nam and Nakdong River confluence, South Korea, which preceded three months of algal bloom season monitored afterward by unmanned aerial vehicles. The HABs-prone areas were derived from temperature, depth, flow velocity, and sediment concentration data based on acoustic Doppler current profilers without using additional data collection. The following methods were used for these hot-spots detection and analysis: K-means clustering, the Getis-Ord G^*, statistics combined with Moran's I spatial autocorrelation as well the local index of spatial association. The visual inspection demonstrated that the comparison between the predicted HABs-prone regions using spatial statistics and unmanned aerial vehicle observations of actual algal blooms was satisfactory [56].

Geostatistical methods are very useful in chlorophyll studies on large scales because they allow to take into account problems related to deteriorated data quality, e.g. filling data gaps resulting from the presence of clouds during satellite observations, instrument noise, biophysical processes on a subsample scale, and allow for the proper assessment of the measurement scale or range of spatial correlation. These issues have been discussed in the series of works by Doney et al. [57, 58], Wallis et al. [59], and Glover et al. [60]. Their objective involved evaluating the mesoscale biological patterns present using global SeaWiFS and/or MODIS/Aqua ocean colour data for Chl-a. The study scrutinised how specific geostatistical measures, such as unresolved variability (nugget effect or noise), resolved variability (partial sill), and spatial range, varied across geographical locations, seasons, and between different years.

One of the essential findings was the close similarity in resolved variability between both instruments. This suggests that geostatistical techniques can effectively capture a reliable measure of biophysical mesoscale variability, regardless of the sensor used. In turn, the nugget effect in variograms calculated from MODIS/Aqua data was substantially lower than that calculated for SeaWiFS, especially in oligotrophic waters which was attributed to the SeaWiFS sensor noise characteristics. Both datasets demonstrated a statistically significant correlation between sills and the horizontal gradient magnitude of the low-pass filtered chlorophyll field. This consistency aligns with the impact of physical stirring on large-scale gradients. Regional discrepancies in the ranges might indicate the scale-dependent nature of biological mechanisms, contributing to variability directly at the mesoscale. The unresolved sub-mesoscale variability could potentially arise from non-geophysical noise linked to instruments, atmospheric aerosols, constraints within the algorithms, and finer-scale variability.

Moreover, geostatistical methods complement eddy-centric compositing techniques. These issues were partially considered in the following works [61-64]. Eveleth et al. [65] continued this subject and evaluated the simulated mesoscale variability of the high-resolution ocean biophysical model Community Earth System Model (CESM). Such models are used at basin and global scales to help deepen understanding of the mechanistic coupling of physical and biological processes at the mesoscale. The authors calculated 2D variograms and map of variogram parameters (resolved-variability, i.e., sill, and range), as well as coefficients of variation to quantify the magnitude and spatial scale of chlorophyll patchiness in the North Atlantic Ocean, where there is a large seasonal phytoplankton bloom. The results obtained using 1/10th-degree eddy-resolving coupled CESM simulations [66] agreed with those obtained using satellite (MODIS, SeaWiFS) observations.

The HABs and Hypoxia Research and Control Amendments Act of 2017 (U.S. Congress, 2017) emphasised coordination and integration to prevent unnecessary duplication of effort among programs. This stimulated the spatiotemporal analysis of the HABs studies over large areas of the western basin of Lake Erie (approximately 3000 km^2) during the HABs season (June-October) from 2008 to 2017 which were carried out in the frame of five independent monitoring programs [67]. Substantial efforts have been made to track these blooms' intensity and extent using in situ chlorophyll sampling and remote sensing observations. Many scientific institutions were involved in this study, namely the National Oceanic and Atmospheric Administration, the Great Lakes Environmental Research Laboratory, the Ohio State University, Stone Laboratory, the U.S. Geological Survey, Great Lakes Science Center, University of Toledo, Lake Erie Center, and Environment and Climate Change Canada. Each monitoring program had different analytical methods to measure Chl-a concentrations, but all used filtration to capture phytoplankton on a glass fiber filter and an organic solvent to extract Chl-a from cells. In addition, ancillary data such as bathymetry and wind speed to check the potential chlorophyll response to mixing conditions were used. However, despite extensive measurements, complete comprehension of HABs' spatial and temporal dynamics was not fully achieved. This is attributed to the constraints imposed by discrete shipboard sampling across vast regions and the impact of clouds and winds on remote sensing estimations.

To overcome these constraints, Fang et al. [67] devised an innovative spatiotemporal geostatistical framework to estimate HABs intensity. Within this model, observed Chl-a concentrations, $z(x, y, t)$, are viewed as instances of a spatiotemporal random field, where x, y, and t denote longitude, latitude, and time, respectively. The formulation of the space-time geostatistical model can be depicted as:

$$\boldsymbol{z}(x,y,t) = \boldsymbol{\mu}(x,y,t) + \boldsymbol{\eta}(x,y,t) + \boldsymbol{\varepsilon}(x,y,t) \tag{5}$$

Equation (5) comprises three components: $\boldsymbol{\mu}(x,y,t)$ represents a deterministic space-time trend function derived from the matrix product $\boldsymbol{X\beta}$, wherein $\boldsymbol{X}$ denotes a matrix encompassing explanatory variables such as bathymetry and sampling type, and $\boldsymbol{\beta}$ represents a vector of estimated regression coefficients; $\boldsymbol{\eta}(x,y,t)$ embodies spatiotemporally correlated stochasticity, within the context of this study this component follows a Gaussian space-time process with a mean of zero and a Matern covariance function; $\boldsymbol{\varepsilon}(x,y,t)$ represents the independent error term following a Gaussian distribution $N(0, \tau^2)$, where τ^2 captures spatially uncorrelated residual variance arising from microvariability and measurement error (also referred to as the 'nugget effect'). Chlorophyll data underwent log transformation to ensure normality in the residuals [68].

Based on the Bayesian information criterion for model selection, trend variables explained bloom expansion, wind effects over depth, and variability among sampling methods. Cross-validation results demonstrated that space-time kriging explained on average over half of the variability in daily, location-specific chlorophyll observations. Conditional simulations provided, for the first time, comprehensive estimates of overall bloom biomass based on depth-integrated concentrations as well as surface aerial extent with quantified uncertainties. Thus, the first space-time geostatistical modelling framework for estimating HABs intensity has shown great potential both in modelling algal blooms over large areas, as well as in integrating data obtained from various programs. This has become possible thanks to the development of Bayesian methods and progress in the development of spatio-temporal geostatistics in recent decades.

To understand the spatial heterogeneity and temporal dynamics of cyanobacterial blooms in Clear Lake, California, which is a shallow, polymictic, naturally eutrophic lake with a long record of episodic cyanobacteria blooms, Sharp et al. [69] evaluated a satellite remote sensing tool for estimating coarse cyanobacteria distribution with coincident, in situ measurements at varying scales and resolutions. Two Cyanobacteria Index (CI) remote sensing algorithms (a classical one based on the spectral shape at wavelength 681 nm, [70] and a refined one which incorporates an exclusionary criterion for the spectral shape at 665 nm, [71]) were used to estimate cyanobacterial abundance in the top portion of the water column from data acquired from the Ocean and Land Color Instrument (OLCI) sensor on the Sentinel-3a satellite. The authors also collected hyperspectral data using various tools: a handheld spectroradiometer for Chl-a and phycocyanin measurements, multispectral imagery from small Unmanned Aerial System (sUAS) flights, and Autonomous Underwater Vehicle (AUV) measurements for Chl-a, turbidity, and coloured dissolved organic matter. Additionally, they obtained meteorological and lake temperature data. Analyses of the high-resolution AUV and sUAS data revealed a Critical Scale of Variability (or practical range, according to geostatistical nomenclature [15, 25]) for cyanobacterial blooms ranging from 70 m to 175 m, which is finer than satellite resolution. This suggests substantial spatial variability within each 300-meter satellite pixel. The field spectroscopy data assessed both the original and revised CI algorithms; however, the revised algorithm did not effectively estimate cyanobacterial abundance for their study site. Nevertheless, the authors concluded that satellite-based remote sensing tools provide consistent and cost-effective observations over large areas. These findings emphasise the need for continuous advancements in remote sensing techniques crucial for monitoring Harmful Algal Blooms (HABs) in lakes and reservoirs [71].

The study authored by Szatmari et al. [72] merits attention as well. Their study encompassed a comprehensive spatial modelling endeavour focusing on nutrients and their sedimentary ratios within Hungary's Balaton Lake. They aimed to assess the efficacy of employing multivariate geostatistics, particularly when multiple variables require spatial modelling across diverse scales within water ecosystems. This is a pivotal issue because the quality and quantity of nutrients stored in sediments, and their relative content may be an important factor that has an important impact on eutrophication. The authors modelled the spatial distribution of the nutrients nitrogen (N) and phosphorus (P) and their ratio (i.e., N : P) in the sediments of the lake and then provided spatial predictions at different scales (i.e., point, basin, and entire lake) with the associated uncertainty. It was found that in Balaton Lake increased phosphorus loads can cause a growth in the mass of blue-green algae (*Cylindrospermopsis raciborskii*) with the ability to fix atmospheric N_2 and thus reduce the nitrogen limitation in water bodies [73, 74]. The authors [72] calculated variograms and cross-variograms of the normal score transformed N and P data to explore, analyse, and model the spatial continuity of each nutrient separately and the spatial interdependence between the nutrients. Variography confirmed that there is a spatial interdependence between the nutrients. Thus, the results revealed that multivariate geostatistics allows this interdependence to be taken into account and exploited to provide coherent and accurate spatial models. Additionally, stochastic realisations, reproducing the joint spatial variability, can be generated that allow providing spatially aggregated predictions with the associated uncertainty at various scales. The study highlighted that it is worthy of applying multivariate geostatistics in case spatial modelling of two or more variables, which jointly vary in space, is targeted in water ecosystems.

The zooplankton community is widely used as a bioindicator for the biological assessment in rivers. Simultaneously, phyto-zooplankton interactions and spatially varying river environment parameters perceivably govern their spatial distribution. Sahu et al. [75] proposed a geostatistical framework to predict zooplankton abundance along the river course while decoupling the phyto-zooplankton relationship from spatial dependency. To do so they used the nonlinear logistic regression kriging. The study was carried out on the river Narmada which is a large tropical river in India. It was found that the correlation between phytoplankton and zooplankton predominantly accounted for zooplankton variations, exhibiting some residual spatial dependence. The suggested approach led to smooth predictive spatial distributions of zooplankton abundance. The study also enabled approximation of zooplankton abundance in regions where direct sampling was not feasible.

Table 1

The main measurement methods and data used for geostatistical analyses in the presented research on algae and phytoplankton

Geostatistical methods	**Inland lakes, freshwater marshes**	**Rivers**	**Seas, oceans, and coastal waters**
Variography, fractal analysis, etc.	[37] (Field measurements included water-column primary productivity, phytoplankton, peri, and meta phyton) [41] (Fluorescence measurements) [69] (Handheld spectroradiometer, Unmanned Aerial System, Autonomous Underwater Vehicle, meteorological forcing and water temperature data, Sentinel-3A/ OLCI)		[49] (Fast Repetition rate data) [53] (Fluorescence measurements, VIIRS) [57-59, 65] (CESM model, and/or MODIS, and/or SeaWiFS)
Spatial estimations	[39] (Unmanned Aerial Vehicles, acoustic Doppler current profilers) [45] Water samples were collected below the surface, and chemical analyses were conducted for Chl-a and other parameters, including Secchi disk (SD) depth, water temperature, pH, and suspended solids (SS) [46] (Water samples were collected from the water surface, and total nitrogen (TN), total phosphorus (TP), chemical oxygen demand (COD), and Chl-a were analysed in the laboratory) [47] Macrophyte coverage at each site was estimated through visual observation and converted to the Braun-Blanquet scale [76]. Zooplankton samples were collected from the bottom to the surface using a plankton pump and filtered through nets with varying mesh sizes. The quantitative counting and measurement of zooplankton were performed using stereomicroscopy.	[56] (Acoustic Doppler current profiler) [75] (Plankton data and water quality parameters)	[44] (MERIS) [50] (Literature and AlgaeBase online data) [48] (SCUBA diving and underwater video) [60] (SeaWiFS, MODIS/Aqua [55] (MODIS/Aqua)

Geostatistical methods	Inland lakes, freshwater marshes	Rivers	Seas, oceans, and coastal waters
Spatial simulations	[67] (In situ chlorophyll sampling, bathymetry, wind speed, and remote sensing observations). [72] (N and P measurements)	[54] (Fluorescence probe, RapidEye satellite data).	[51] (Spatial model of algal blooms with known critical transitions).

Table 2

The main results in research on algae or phytoplankton obtained using geostatistical methods

Geostatistical methods	Inland lakes, freshwater marshes	Rivers	Seas, oceans, and coastal waters
Variography, fractal analysis, etc.	[37] (Spatial and temporal patterns of metaphyton, periphyton, phytoplankton in two riparian freshwater marshes. [41] (First use of SeaWiFS imagery to investigate the connection between physical and biological processes in the Great Lakes) [69] (A multiplatform sampling approach to evaluate the spatial variability of cyanobacteria blooms. The approach provided a more holistic view of a cyanobacteria bloom.		[49] (Analysis of the consistency of spatial heterogeneity in several algal functional groups and in grazer biomass between two sites and across the shore of a river) [53] (Determination of the variability of phytoplankton physiology and biomass across a range of light and nutrient environments) [57-59, 65] (Characterisation and determination of mesoscale global patterns in ocean biological variability [58], meso and submesoscale ocean colour variability [57, 59] and mesoscale ocean chlorophyll variability [65])
Spatial estimations	[39] (Kriging maps for the three months examined. Determination of the relationship between the spatial distribution of chlorophyll in water and precipitation. Variography - a powerful method to analyse spatial patterns of ecological factors in freshwater ecosystems) [45] (The spherical model could be applied to fit all experimental variograms of Chl-a content in the lake. The spatial structural patterns of Chl-a were influenced by factors such as the distribution of pollution sources, water flow, and wind). [46] Mass balance calculations for TP demonstrated substantial phosphorus input during the flood season, with the most severe algal blooms occurring	[56] (Development of a practical and rapid countermeasure, enabling a preemptive response to massive algal blooms in a river) [75] (Generation of smooth zooplankton abundance and sustainability predictive maps. These maps have precisely identified the most productive zone of zooplankton sustainability)	[44] (Kriging can be used to estimate missing satellite chlorophyll data in the North Sea) [50] (The effective generation of a database of algal species richness at two spatial scales: regional Gulf of Mexico and the Caribbean, and -global coastal zones) [48] (Assessment of distribution and ecological significance of the red alga *Furcellaria lumbricalis*) [60] (Application of geostatistical analysis techniques allows the variability observed in

Geostatistical methods	Inland lakes, freshwater marshes	Rivers	Seas, oceans, and coastal waters
	afterward. This delay was attributed to hydrodynamic processes in the lake during the flood period. [47] The study demonstrated that zooplankton and phytoplankton exhibited distinct degrees of heterogeneity and spatial structures, necessitating separate modelling approaches. However, the observed similarity in spatial autocorrelation within both zooplankton and phytoplankton communities suggested that, at the scales examined, the horizontal organisation of these components was not significantly influenced by species-specific behaviours.		ocean colour imagery to be partitioned into resolved and unresolved variability) [55] (Satellite information and geostatistical methods can be used to quantify the effect of a large coastal mussel farm on water quality)
Spatial simulations	[67] (Developing a space-time geostatistical approach for modelling algal bloom variability. Modelling simultaneously estimates bloom surface area and overall biomass)	[54] (Retrieving the spatial distribution of cyanobacterial concentration from satellite imagery data for the calibration of the 2D numerical model in a river)	[51] (Analysing a spatially dynamic nutrient-phytoplankton model across a range of nutrient input and physical mixing conditions)

Conclusion

Reliably assessing the spatial distributions of algae and phytoplankton is a pivotal area of scientific interest due to their crucial ecological role and potential risks to aquatic ecosystems and human health. Typically examined in two or three dimensions, and more recently, expanded into four dimensions with the inclusion of time, these analyses demand extensive measurements and considerable effort. Geostatistical methods serve as invaluable tools for studying their distribution in water bodies, enabling a reduction in measurements and research costs. These methods allow data analysis across different scales, from small reservoirs to oceans. Geostatistical methods accurately estimate the spatial distribution of algae and phytoplankton in the water, including kriging variance distributions, in both single and multi-variable studies, utilising approaches such as cokriging. Integration of diverse data types, ranging from underwater measurements to satellite observations, is crucial for studying these organisms. Geostatistical methods also facilitate spatial and space-time simulations. Since the 1980s, the use of geostatistical methods in algae and phytoplankton research has significantly expanded. Further advancements are anticipated, particularly with the adoption of Bayesian, MCMC methods, and the INLA approach for spatiotemporal modelling, overcoming prior computational constraints. The fusion of geostatistical methods and artificial intelligence is also projected. The increasing availability of high-resolution remote observations from satellites and drones, coupled with progress in underwater measurement techniques, is expected to enhance the utilisation of geostatistical methods. This advancement holds particular promise for studying algae and phytoplankton in rivers, where the need for spatiotemporal modelling has previously hindered research efforts.

Acknowledgements

We would like to express our sincere gratitude to the anonymous reviewer for the insightful comments and constructive feedback. The expertise and thoughtful suggestions offered have significantly enhanced the quality of this work.

References

[1] Harris G. Phytoplankton Ecology: Structure, Function and Fluctuation. London, New York: Chapman and Hall; 1986. DOI: 10.1007/978-94-009-3165-7.

[2] Chapman RL. Algae: the world's most important "plants" - an introduction. Mitig Adapt Strateg Glob Change. 2013;18:5-12. DOI: 10.1007/s11027-010-9255-9.

[3] Geider RJ, Moore CM, Suggett DJ. Ecology of Marine Phytoplankton. In: Ecology and the Environment. New York: Springer; 2014:483-531. DOI: 0.1007/978-1-4614-7501-9_23.

[4] Lapointe BE, Burkholder JM, Alstyne KLV. Harmful Macroalgal Blooms in a Changing World: Causes, Impacts, and Management. Chapter 15 in Harmful Algal Blooms: A Compendium Desk Reference. 2018:515-60. DOI: 10.1002/9781118994672.ch15.

[5] Zhang Y, Whalen JK, Cai C, Shan K, Zhou H. Harmful cyanobacteria-diatom/dinoflagellate blooms and their cyanotoxins in freshwaters: A nonnegligible chronic health and ecological hazard. Water Res. 2023;233:119807. DOI: 10.1016/j.watres.2023.119807.

[6] Patino R, Christensen VG, Graham JL, Rogosch JS, Rosen BH. Toxic algae in inland waters of the conterminous United States - A review and synthesis. Water. 2023;15:2808. DOI: 10.3390/w15152808.

[7] Park J, Patel K, Lee WH. Recent advances in algal bloom detection and prediction technology using machine learning. Sci Total Environ. 2024;938:173546. DOI: 10. 1016/j.scitotenv.2024.173546.

[8] Igwaran A, Kayode AJ, Moloantoa KM, Khetsha ZP, Unuofin JO. Cyanobacteria harmful algae blooms: causes, impacts, and risk management. Water Air Soil Pollut. 2024;235:71. DOI: 10.1007/s11270-023-06782-y.

[9] Calomeni-Eck AJ, McQueen AD, Kinley-Baird CM, Clyde Jr T. Identification of cyanobacteria overwintering cells and environmental conditions causing growth: Application for preventative management. Ecol Solut Evid. 2024;5:e12326. DOI: 10.1002/2688-8319.12326.

[10] Yan Y, Xu Z, Yang B, Jiang, He S, Sheng H, et al. Spatio-temporal variations of water quality and planktonic algal communities in Qingshan Reservoir, China. Pol J Environ Stud. 2023;32:2405-16. DOI: 10.15244/pjoes/158907.

[11] Barkoh A, Fries LT. Aspects of the origins, ecology, and control of golden alga Prymnesium parvum: Introduction to the featured collection. JAWRA J Am Water Resources Assoc. 2010;46:1-5. DOI: 0.1111/j.1752-1688.2009.00394.x.

[12] Roelke D, Manning S. Harmful Algal Species Fact Sheet: Prymnesium parvum (Carter) Golden Algae: A Compendium Desk Reference. In: Harmful Algal Blooms. Hoboken, NJ: Wiley; 2018:629-32. DOI: 10.1002/9781118994672.ch16q.

[13] Kitsiou D, Karydis M. Coastal marine eutrophication assessment: A review on data analysis. Environ Int. 2011;37:778-801. DOI: 10.1016/j.envint.2011.02.004.

[14] Goovaerts P. Geostatistics for Natural Resources Evaluation. Oxford: University Press; 1997. DOI: 10.1093/oso/9780195115383.001.0001.

[15] Webster R, Oliver MA. Geostatistics for Environmental Scientists. 2nd ed. Chichester: Wiley; 2007: DOI: 10.1002/9780470517277.

[16] Gómez-Hernández JJ. Geostatistics for environmental applications. Math Geosci. 2016;48:1-2. DOI: 10.1007/s11004-015-9627-5.

[17] Christakos G. Modern Spatiotemporal Geostatisics. New York: Oxford University Press; 2000: ISBN: 0195138953.

[18] Zawadzki J, Fabijańczyk P. On the influence of the nugget effect on the efficiency of magnetometric soil surface screening. Ecol Chem Eng S. 2023;29:525-35. DOI: 10.2478/eces-2022-0038.

[19] Chilès JP, Delfiner P. Geostatistics: Modeling Spatial Uncertainty. 2nd ed. Hoboken, NJ: Wiley; 2012: ISBN: 9780470183151.

[20] Clark I, Harper WV. Practical Geostatistics 2000. Columbus Ohio: Ecosse North America; 2000. ISBN: 0970331746.
[21] Zawadzki J. Metody geostatystyczne: dla kierunków przyrodniczych i technicznych (Geostatistical Methods: For Natural and Technical Sciences). Warszawa: Ofic Wyd Polit Warszawskiej; 2011. ISBN: 8372079536 .
[22] Isaaks EH, Srivastava RM. An Introduction to Applied Geostatistics. New York: Oxford University Press; 1990. ISBN: 9780195050127.
[23] Chilès JP, Desassis N. Fifty Years of Kriging. In: Daya Sagar BS, Cheng Q, Agterberg F, editors. Handbook of Mathematical Geosciences: Fifty Years of IAMG. Cham: Springer Int Publishing; 2018:589-612. DOI: 978-0-19-505012-7.
[24] Hengl T. A Practical Guide to Geostatistical Mapping of Environmental Variables. EUR 22904 EN. Luxembourg: Office for Official Publications of the European Communities; 2007. ISBN: 9789279069048.
[25] Cressie N. Statistics for Spatial Data. NY, Chichester, Toronto, Brisbane, Singapore: John Wiley Sons; 2015. ISBN: 9781119115182.
[26] Varouchakis EA. 1 - Geostatistics: Mathematical and Statistical Basis. In: Corzo G, Varouchakis EA, editors. Spatiotemporal Analysis of Extreme Hydrological Events. Elsevier; 2019:1-38. ISBN: 9780128116890.
[27] Kleijnen J. Kriging: Methods and Applications. SSRN Electronic J. 2017; DOI: 10.2139/ssrn.3075151.
[28] Oliver MA, Webster R. A tutorial guide to geostatistics: Computing and modelling variograms and kriging. CATENA 2014;113:56-69. DOI: 10.1016/j.catena.2013.09.006.
[29] Diggle P, Lophaven S. Bayesian geostatistical design. Scandinavian J Statistics. 2006;33:53-64. DOI: 10.1111/j.1467-9469.2005.00469.x.
[30] Pilz J, Spöck G. Why do we need and how should we implement Bayesian kriging methods. Stoch Environ Res Risk Assess. 2008;22:621-32. DOI:10.1007/s00477-007-0165-7.
[31] Montero J, Fernández-Avilés G, Mateu J. Spatial and Spatio-Temporal Geostatistical Modeling and Kriging. Wiley; 2015. DOI: 10.1002/9781118762387.
[32] Gómez-Rubio V, Rue H. Markov chain Monte Carlo with the Integrated Nested Laplace Approximation. Stat Comput. 2018;28:1033-51. DOI: 10.1007/s11222-017-9778-y.
[33] Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility. Statistics Computing. 2000;10:325-37. DOI: 10.1023/A:1008929526011.
[34] Ntzoufras I. Bayesian Modeling Using WinBUGS. Hoboken, NJ: Wiley; 2009: DOI: 10.1002/9780470434567.
[35] Blangiardo M, Cameletti M. Spatial and Spatio-temporal Bayesian Models with R - INLA. Chichester, UK: Wiley; 2015. DOI: 10.1002/9781118950203.
[36] Van Niekerk J, Krainski E, Rustand D, Rue H. A new avenue for Bayesian inference with INLA. Computational Statistics Data Analysis. 2023;181:107692. DOI: 10.1016/j.csda.2023.107692.
[37] Wu X, Mitsch W. Spatial and temporal patterns of algae in newly constructed freshwater wetlands. Wetlands. 1998;18:9-20. DOI: 10.1007/BF03161438.
[38] Kawata M, Hayashi M, Hara T. Interactions between neighboring algae and snail grazing in structuring microdistribution patterns of periphyton. Oikos. 2001;92:404-16. DOI: 10.1034/j.1600-0706.2001.920302.x.
[39] Zhao B, Cai Q. Geostatistical analysis of chlorophyll a in freshwater ecosystems. J Freshwater Ecol. 2004;19:613-21. DOI: 10.1080/02705060.2004.9664742.
[40] Burrough P. Fractal dimensions of landscapes and other environmental data. Nature. 1981;294:240-2. DOI: 10.1038/294240a0.
[41] Welty L, Stein M. Modeling Phytoplankton: Covariance and Variogram Model Specification for Phytoplankton Levels in Lake Michigan. In: geoENV IV - Geostatistics for Environmental Applications. Barcelona: Cluver Academic Publishers; 2004:163-73. DOI: 10.1007/1-4020-2115-1_14.
[42] Falkowski PG, Dubinsky Z, Wyman K. Growth-irradiance relationships in phytoplankton. Limnology Oceanography. 1985;30:311-21. DOI: 10.4319/lo.1985.30.2.0311.
[43] Eadie BJ, Robbins JA, Klump JV, Schwab DJ, Edgington DN. Winter-spring storms and their influence on sediment resuspension, transport, and accumulation patterns in Southern Lake Michigan. Oceanography. 2008;21:118-35. DOI: 10.5670/oceanog.2008.09.
[44] Müller D. Estimation of algae concentration in cloud covered scenes using geostatistical methods. In: Proceedings of Envisat Symposium, 23-27 April 2007, Montreux, Switzerland. 2007. ISBN: 9789292912000.
[45] Wang XJ, Liu RM. Spatial analysis and eutrophication assesment for chlorophyll a in Taihu Lake. Environ Monit Assess. 2005;101:167-74. DOI: 10.1007/s10661-005-9154-9.
[46] Zhang F, Tang H, Jin G, Zhu Y, Zhang H, Stewart RA, et al. Evaluating nutrient distribution and eutrophication pattern in a shallow impounded lake: Exploring the influence of floods. Int J Sediment Res. 2024;39:375-85. DOI: 10.1016/j.ijsrc.2024.04.006.

[47] Ludovisi A, Minozzo M, Pandolfi P, Taticchiet MI. Modelling the horizontal spatial structure of planktonic community in Lake Trasimeno (Umbria, Italy) using multivariate geostatistical methods. Ecol Modelling. 2005;181:247-62. DOI: 10.1016/j.ecolmodel.2004.06.033.

[48] Bucas M. Distribution patterns and ecological role of the Red Alga Furcellaria Lumbricalis (Hudson) J.V.Lamouroux off the exposed Baltic Sea coast of Lithuania. [PhD Thesis]. Klaipėda 2009. Available from: https://www.ku.lt/uploads/documents/files/jti/studijos/MBucas2004-2009.pdf; Access: 16.12.2024.

[49] Diaz E, Erlandsson J, McQuaid C. Detecting spatial heterogeneity in intertidal algal functional groups, grazers and their co-variation among shore levels and sites. J Experimental Marine Biol Ecol. 2011;409. DOI: 10.1016/j.jembe.2011.08.013.

[50] Tapia O, Vilchis MI, Sentíes A, Dreckmann K. Mapping of algae richness using spatial data interpolation. International J Remote Sensing. 2015;XL-7/W3. DOI: 10.5194/isprsarchives-XL-7-W3-1005-2015.

[51] Buelo C, Carpenter S, Pace MA modeling analysis of spatial statistical indicators of thresholds for algal blooms. Limnology Oceanography Lett. 2018;3. DOI: 10.1002/lol2.10091.

[52] Serizawa H, Amemiya T, Itoh K. Patchiness in a minimal nutrient - phytoplankton model. J Biosci. 2008;33:391-403. DOI: 10.1007/s12038-008-0059-y.

[53] Fox JE. Utilising chlorophyll fluorescence to assess the variability of phytoplankton biomass and productivity in the north-west European shelf seas. PhD Thesis. University of Essex; 2018. Available from: https://repository.essex.ac.uk/21546/1/Fox_2017_corrected.pdf; Access: 16.12.2024.

[54] Kim JS, Seo IW, Baek D. Modeling spatial variability of harmful algal bloom in regulated rivers using a depth-averaged 2D numerical model. J Hydroenviron Res. 2018;20. DOI: 10.1016/j.jher.2018.04.008.

[55] Pinkerton M, Gall M, Wood S, Zeldis J. Measuring the effects of bivalve mariculture on water quality in northern New Zealand using 15 years of MODIS-Aqua satellite observations. Aquacult Environ Interactions. 2018;10:529-45. Available from: https://www.researchgate.net/publication/328024096_Measuring_the_effects_of_bivalve_mariculture_on_water_quality_in_northern_New_Zealand_using_15_years_of_MODIS-Aqua_satellite_observations; Access: 16.12.2024.

[56] Son G, Kim D, Kim Y, Lyu S, Kim S. A forecasting method for harmful algal bloom(HAB)-prone regions allowing preemptive countermeasures based only on acoustic doppler current profiler measurements in a large river. Water. 2020;12:3488. DOI: 10.3390/w12123488.

[57] Doney SC, Glover D, McCue S, Fuentes M. Mesoscale Variability of SeaWiFS Satellite Ocean Color: Global Patterns and Spatial Scales. Available from: https://www.researchgate.net/publication/2409753_Mesoscale_Variability_of_SeaWiFS_Satellite_Ocean_Color_Global_Patterns_and_Spatial_Scales; Access: 03.12.2024.

[58] Doney SC, Glover D, McCue S, Fuentes M. Mesoscale variability of Sea-viewing Wide Field-of-view Sensor (SeaWiFS) satellite ocean color: Global patterns in spatial scales. J Geophys Res (Oceans). 2003;108. DOI: 10.1029/2001JC000843.

[59] Wallis A, Doney SC, Glover D, Nelson NB. Characterizing Submesoscale Ocean Color Variability in the Sargasso Sea in the Vicinity of the Bermuda Atlantic Time-series Site (BATS): A Geostatistical Approach. Available from: https://www.researchgate.net/publication/252148091_Characterizing_Submesoscale_Ocean_Color_Variability_in_the_Sargasso_Sea_in_the_Vicinity_of_the_Bermuda_Atlantic_Time-series_Site_BATS_A_Geostatistical_Approach; Stand: 03.12.2024.

[60] Glover D, Doney SC, Oestreich W, Tullo A. Geostatistical analysis of mesoscale spatial variability and error in SeaWiFS and MODIS/Aqua Global Ocean Color Data. J Geophys Res: Oceans. 2017;123. DOI: 10.1002/2017JC013023.

[61] Chelton D, Gaube P, Schlax M, Early JJ, Samelson RM. The influence of nonlinear mesoscale eddies on near-surface oceanic chlorophyll. Science. New York. 2011;334:328-32. DOI: 10.1126/science.1208897.

[62] Gaube P, Chelton D, Strutton P, Behrenfeld MJ. Satellite observations of chlorophyll, phytoplankton biomass, and Ekman pumping in nonlinear mesoscale eddies. J Geophys Res: Oceans. 2013;118. DOI: 10.1002/2013JC009027.

[63] Rohr T, Harrison C, Long MC, Gaube P, Doney SC. The simulated biological response to southern ocean eddies via biological rate modification and physical transport. Global Biogeochem Cycles. 2020;34:e2019GB006385. DOI: 10.1029/2019GB006385.

[64] Rohr T, Harrison C, Long MC, Gaube P, Doney SC. Eddy-modified iron, light, and phytoplankton cell division rates in the simulated southern ocean. Global Biogeochem Cycles. 2020;34: e2019GB006380. DOI: 10.1029/2019GB006380.

[65] Eveleth R, Glover DM, Long MC, Lima ID, Chase AP, Doney SC. Assessing the skill of a high-resolution marine biophysical model using geostatistical analysis of mesoscale ocean chlorophyll variability from field observations and remote sensing. Front Marine Sci. 2021;8. DOI: 10.3389/fmars.2021.612764.

[66] Smith R, Jones P, Briegleb P, Brayan F, Danabasoglu G, Dennis J, et al. The Parallel Ocean Program (POP) reference manual: Ocean Component Community Climate System Model (CESM). 2010. Available from: https://www2.cesm.ucar.edu/ models/cesm1.3/ ocean/doc/sci/POPRefManual.pdf. Access: 03.12.2024.

[67] Fang S, Del Giudice D, Scavia D, Binding CE, Bridgeman TB, Chaffin JD, et al. A space-time geostatistical model for probabilistic estimation of harmful algal bloom biomass and areal extent. Sci Total Environ. 2019;695:133776. DOI: 10.1016/j.scitotenv.2019.133776.

[68] Stein ML. Space-time covariance functions. J Am Statistical Assoc. 2005; DOI: 10.1198/016214504000000854.

[69] Sharp SL, Forrest AL, Bouma-Gregson K, Jin Y, Cortes A, Schladow SG. Quantifying scales of spatial variability of cyanobacteria in a large, eutrophic lake using multiplatform remote sensing tools. Front Environ Sci. 2021;9. DOI: 10.3389/fenvs.2021.612934.

[70] Wynne T, Stumpf R, Tomlinson M, Dyble J. Characterizing a cyanobacterial bloom in Western Lake Erie using satellite imagery and meteorological data. Limnol Oceanography. 2010. DOI: 10.4319/lo.2010.55.5.2025.

[71] Lunetta RS, Schaeffer BA, Stumpf RP, Keith D, Scott AJ, Murphy MS. Evaluation of cyanobacteria cell count detection derived from MERIS imagery across the eastern USA. Remote Sensing Environ. 2015;157:24-34. DOI: 10.1016/j.rse.2014.06.008.

[72] Szatmári G, Kocsis M, Makó A, Pásztor L, Bakacsi Z. Joint spatial modeling of nutrients and their ratio in the sediments of Lake Balaton (Hungary): A multivariate geostatistical approach. Water. 2022;14:361. DOI: 10.3390/w14030361.

[73] Présing M, Preston T, Takátsy A, Kovács AW, Vörös L, Kenesi G, et al. Phytoplankton nitrogen demand and the significance of internal and external nitrogen sources in a large shallow lake (Lake Balaton, Hungary). Hydrobiologia. 2008;599:87-95. DOI: 10.1007/s10750-007-9191-1.

[74] Vörös L, Göde PN. Long term changes of phytoplankton in Lake Balaton (Hungary). Int Vereinigung theoretische angewandte Limnologie: Verhandlungen. 1993;25:682-6. DOI: 10.1080/03680770.1992.11900224.

[75] Sahu S, Sarkar S, Gogoi P, Naskar M. A geostatistical framework predicting zooplankton abundance in a large river: Management implications towards potamoplankton sustainability. Environ Manage. 2023;71. DOI: 10.1007/s00267-023-01784-2.

[76] Ivanova N. Global overview of the application of the Braun-Blanquet approach in research. Forests. 2024;15:937. DOI: 10.3390/f15060937.