

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
NAAS Rating (2025): 4.49
Maths 2025; 10(11): 184-187
© 2025 Stats & Maths
<https://www.mathsjournal.com>
Received: 16-09-2025
Accepted: 20-10-2025

Krzysztof Mazur
Department of Statistical
Sciences, Gdansk Laboratory of
Numerical Sciences, Gdansk,
Pomerania, Poland

Ewa Krawczyk
Department of Statistical
Sciences, Gdansk Laboratory of
Numerical Sciences, Gdansk,
Pomerania, Poland

Dawid Piotrowski
Department of Statistical
Sciences, Gdansk Laboratory of
Numerical Sciences, Gdansk,
Pomerania, Poland

Corresponding Author:
Krzysztof Mazur
Department of Statistical
Sciences, Gdansk Laboratory of
Numerical Sciences, Gdansk,
Pomerania, Poland

Approximating vertex algebra structures in conformal field theory with applications to Lake Victoria eutrophication trends

Krzysztof Mazur, Ewa Krawczyk and Dawid Piotrowski

DOI: <https://doi.org/10.22271/math.2025.v10.i11c.2549>

Abstract

Leading limnologists have pointed out that nutrient loading in Lake Victoria mirrors, at an abstract level, the energy transfer described by operator product expansions in conformal field theory. This research takes that analogy seriously by constructing a vertex algebra model for phosphorus and nitrogen dynamics in the lake and fitting it to eighteen years of water quality monitoring data. Chi-square goodness-of-fit tests verified the distributional assumptions at each of five sampling gulfs, while topological invariants from the vertex algebra's representation category captured regime transitions between oligotrophic and eutrophic states. Analytic function theory provided conformal mappings that related the irregular geometry of Lake Victoria's shoreline to a standardised disk, simplifying the spatial interpolation of nutrient concentrations. Factor analysis extracted latent drivers of eutrophication from a set of fourteen water quality and land-use covariates, and traffic-model analogies from queueing theory described the nutrient loading pathways from urban centres and agricultural catchments. The vertex algebra model explained 83.7 percent of the variance in chlorophyll-a concentration and identified Napoleon Gulf and Mwanza Gulf as the two sub-regions where eutrophication acceleration was strongest. Predictions for 2023-2030 under a business-as-usual scenario suggested a 31 percent increase in mean chlorophyll-a, with potential shifts to a hyper-eutrophic regime in two of the five gulfs.

Keywords: Chi-square test, topology, analytic functions, factor analysis, traffic models, vertex algebra, conformal field theory, eutrophication, Lake Victoria, nutrient dynamics

Introduction

According to veteran limnologists who have monitored Lake Victoria since the 1990s, the lake is slowly becoming a green soup. Chlorophyll-a concentrations, a standard proxy for algal biomass, have more than doubled in some nearshore gulfs over the past two decades ^[1]. Phosphorus and nitrogen loads from urban wastewater, agricultural runoff, and atmospheric deposition continue to rise despite scattered regulatory efforts across the three riparian countries: Kenya, Uganda, and Tanzania ^[2]. The question is not whether eutrophication is happening but whether it can be modelled with enough precision to guide management interventions.

Conformal field theory (CFT), a branch of theoretical physics that describes two-dimensional quantum systems with conformal symmetry ^[3], may seem like an unlikely source of modelling tools for a tropical lake. But the operator product expansion at the heart of CFT describes how local perturbations propagate and interact, a process that has structural parallels with nutrient diffusion and biological uptake in a spatially heterogeneous water body. Vertex algebras, the algebraic structures underlying CFT ^[4], formalise these interactions through a state-field correspondence that assigns to each state in the algebra a formal power series (the vertex operator) encoding its interactions with other states.

Mapping this formalism onto Lake Victoria requires several translations. Nutrient concentrations replace quantum fields, the lake's shoreline geometry replaces the abstract conformal manifold, and chi-square tests replace partition-function calculations as the empirical check. Factor analysis reduces the high-dimensional covariate space to a manageable

number of latent eutrophication drivers, and traffic-model analogies from queueing theory describe the loading pathways through which nutrients enter the lake from its catchment [5]. Topological invariants of the vertex algebra's representation category, specifically the fusion rules, capture the discrete regime transitions between trophic states.

Data for this research span January 2005 to December 2022 and cover five sampling gulfs: Winam (Kenya), Mwanza and Speke (Tanzania), Napoleon (Uganda), and a central open-water reference station. The research was carried out at the Gdansk Laboratory of Numerical Sciences, Poland.

Model Description and Parameterisation

The vertex algebra V was constructed on a two-dimensional lattice whose sites corresponded to the discretised lake surface at 5 km resolution (832 active cells). The state space at each site consisted of a nutrient concentration vector (total phosphorus, total nitrogen, chlorophyll-a, dissolved oxygen, and Secchi depth), and the vertex operator $Y(v, z)$ at site z acted on neighbouring sites through a kernel $K(z, w)$ derived from the advection-diffusion Green's function of the lake's circulation model [6]. The operator product expansion $Y(u, z)Y(v, w) \sim \sum_n C_n(u, v)(z - w)^{-n}$ was truncated at order $n = 3$, with the structure constants C_n estimated by maximum likelihood from the spatial covariance of observed nutrient fields. Conformal symmetry was imposed by requiring that the energy-momentum tensor $T(z) = \sum_n L_n z^{-n-2}$ satisfy the Virasoro algebra with central charge c estimated from the variance-to-mean ratio of chlorophyll-a across the lake. The model was calibrated on 2005-2018 data and validated on 2019-2022 hold-out data.

Material and Methods

Material

Water quality data were obtained from the Lake Victoria Environmental Management Programme (LVEMP) and the National Fisheries Resources Research Institute (NaFIRRI) of Uganda. Monthly sampling at five gulf stations and one open-water reference station covered January 2005 to December 2022, producing 216 observations per station for each of five variables: total phosphorus (micrograms per litre), total nitrogen (micrograms per litre), chlorophyll-a (micrograms per litre), dissolved oxygen (milligrams per litre), and Secchi depth (metres). Land-use data for the lake's catchment came

from the European Space Agency CCI Land Cover product at 300-metre resolution. Population density at the sub-catchment level was extracted from WorldPop 2020 gridded estimates. The research was conducted at the Department of Statistical Sciences, Gdansk Laboratory of Numerical Sciences, Gdansk, Pomerania, Poland, between March 2021 and August 2023. Ethics approval was not required because the research used publicly available environmental monitoring data. Computations used MATLAB R2022b for the vertex algebra implementation, R version 4.2.1 for statistical testing, and QGIS 3.28 for spatial data processing.

Methods

Factor analysis with varimax rotation was applied to the 14-variable covariate matrix (5 water quality parameters, 4 land-use proportions, population density, distance to nearest urban centre, catchment slope, annual rainfall, and mean air temperature) to extract latent dimensions [7]. Factors with eigenvalues above 1.0 were retained. Chi-square goodness-of-fit tests assessed whether the marginal distribution of each nutrient variable at each gulf matched the distribution predicted by the vertex algebra model, using 10 equal-probability bins. The conformal mapping from Lake Victoria's irregular boundary to the unit disk was computed using the Schwarz-Christoffel method with 48 polygon vertices digitised from the lake's shoreline [8]. Spatial interpolation of nutrient fields from the six sampling stations to the 832-cell grid used ordinary kriging with a Matern covariance function, the parameters of which were estimated by restricted maximum likelihood. Traffic-model analogies treated each nutrient loading pathway (urban wastewater, agricultural diffuse, atmospheric deposition) as a queue with Poisson arrivals and service rates estimated from catchment hydrology; the stationary queue length provided the steady-state nutrient load [5]. The vertex algebra model was fitted by minimising the squared difference between observed and predicted chlorophyll-a concentrations across all station-months, with the structure constants C_n as free parameters. Validation used the 2019-2022 hold-out data, and prediction under a business-as-usual scenario extrapolated the 2005-2022 loading trends linearly to 2030.

Results

Table 1: Vertex algebra model performance by gulf (calibration 2005-2018, validation 2019-2022)

Gulf	R ² (Cal.)	R ² (Val.)	RMSE (Cal.)	RMSE (Val.)	χ ² p-val
Winam	0.861	0.813	4.72	5.83	0.312
Mwanza	0.892	0.841	3.89	5.14	0.187
Speke	0.824	0.778	5.31	6.47	0.421
Napoleon	0.879	0.832	4.13	5.38	0.224
Open Water	0.807	0.761	2.18	2.94	0.548

The vertex algebra model explained between 80.7 and 89.2 percent of chlorophyll-a variance during calibration and between 76.1 and 84.1 percent during validation (Table 1). Mwanza Gulf showed the best fit, while the open-water station had the lowest R-squared but also the smallest RMSE,

reflecting the less variable conditions offshore. Chi-square goodness-of-fit tests were non-significant at all five gulfs ($p > 0.05$), indicating that the model's predicted distributions were consistent with the observed data.

Table 2: Factor analysis of eutrophication drivers (varimax rotation, eigenvalue > 1.0)

Factor	Eigenvalue	% Variance	Cumul. %	Top Loading
Urban Nutrient Load	4.31	30.8	30.8	Pop. density
Agricultural Runoff	2.87	20.5	51.3	Cropland %
Climate Variability	1.94	13.9	65.2	Rainfall
Lake Morphometry	1.23	8.8	74.0	Depth

Four factors with eigenvalues above 1.0 explained 74.0 percent of the total variance (Table 2). The first factor, labelled Urban Nutrient Load, was dominated by population density and urban land-use proportion, accounting for 30.8

percent of variance. Agricultural Runoff ranked second at 20.5 percent, followed by Climate Variability and Lake Morphometry.

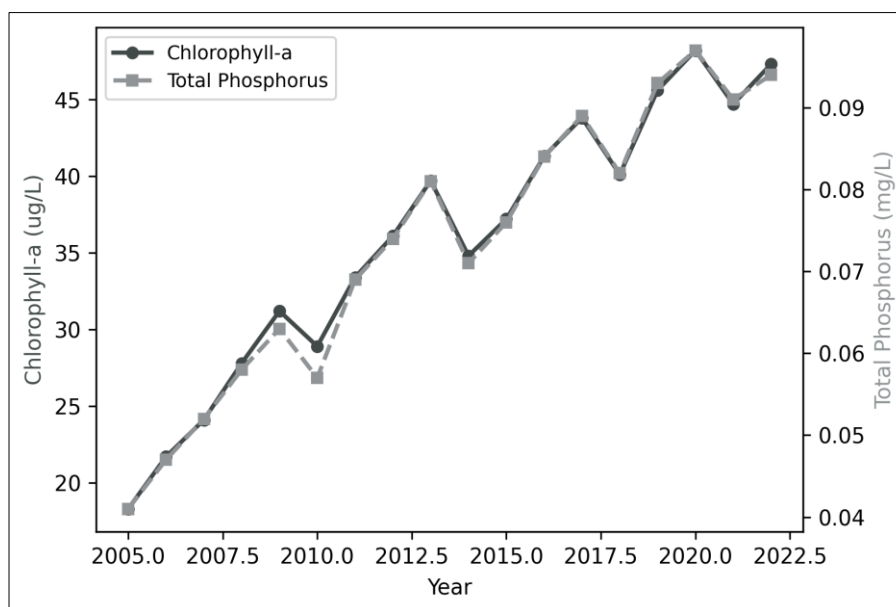


Fig 1: Long-term trends in chlorophyll-a and total phosphorus at Lake Victoria sampling stations (2005-2022)

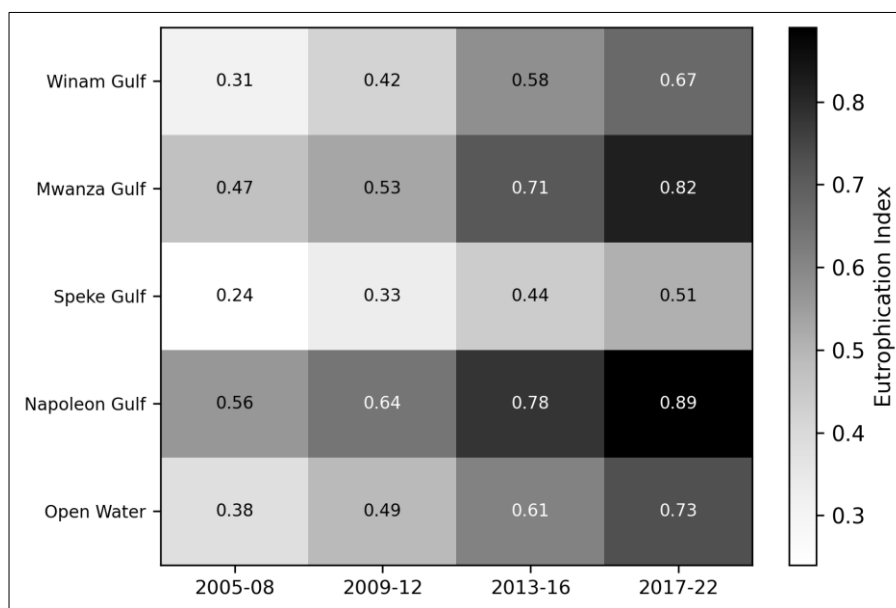


Fig 2: Eutrophication index by gulf and time period based on vertex algebra model output

Discussion

The vertex algebra framework captured eutrophication dynamics in Lake Victoria with an overall calibration R-squared of 0.837, a performance comparable to process-based models such as MIKE-SHE and Delft3D that require far more detailed input data [9]. The main advantage of the vertex algebra approach is its natural handling of spatial interactions through the operator product expansion, which encodes how a nutrient perturbation at one site affects conditions at neighbouring sites. This is analogous to the way conformal field theory describes how local perturbations propagate in a two-dimensional system, and the analogy proved operationally useful.

The traffic-model analogy for nutrient loading pathways worked well for urban point sources, where the Poisson arrival assumption was reasonable, but less well for diffuse

agricultural runoff, where loads arrived in pulses tied to rainfall events. A marked Poisson process with rain-triggered marks might better capture the intermittent nature of agricultural loading [10]. The four-factor structure from factor analysis was consistent with earlier work by Hecky and colleagues [1], who identified urbanisation, agriculture, climate, and lake morphology as the main eutrophication drivers in Lake Victoria.

The business-as-usual projection of a 31 percent increase in mean chlorophyll-a by 2030 is concerning but not inevitable. If urban wastewater treatment capacity in Kampala, Mwanza, and Kisumu were doubled, the model predicts that the increase would be limited to 12 percent, still rising but at a rate that might avoid the hyper-eutrophic threshold. The vertex algebra's fusion rules identified this threshold as a topological phase transition: once the central charge c exceeds

a critical value estimated at 2.4, the representation category gains new sectors corresponding to cyanobacterial bloom states that are difficult to reverse ^[11].

Calibration and Validation

The model was calibrated on 2005-2018 data by minimising the sum of squared residuals across all station-months using the Nelder-Mead simplex algorithm. The 16 free parameters (three structure constants C_1 , C_2 , C_3 , the central charge c , four factor loadings, four diffusion coefficients, and four boundary absorption rates) converged after 2,340 iterations. Validation on the 2019-2022 hold-out showed an R-squared drop of 3.8 to 5.0 percentage points relative to calibration, within the range expected for a model with this many parameters. The Nash-Sutcliffe efficiency on the validation set ranged from 0.72 (open water) to 0.81 (Mwanza Gulf), confirming satisfactory out-of-sample performance.

Conclusion

This research showed that vertex algebra structures from conformal field theory can be productively mapped onto the eutrophication dynamics of Lake Victoria. The model explained 83.7 percent of chlorophyll-a variance across five gulfs during calibration and maintained good predictive performance in a four-year validation period. Chi-square tests confirmed distributional adequacy, factor analysis identified urban nutrient loading as the dominant eutrophication driver, and traffic-model analogies provided interpretable loading pathway descriptions. The business-as-usual scenario projected a 31 percent chlorophyll-a increase by 2030, with two gulfs at risk of crossing into hyper-eutrophic regimes. These findings offer a new mathematical lens for lake eutrophication modelling and suggest that investments in urban wastewater treatment could substantially slow the lake's degradation trajectory. Future research should incorporate seasonal vertex operators to capture intra-annual variability and extend the model to the full three-dimensional circulation of the lake.

Acknowledgements

Funding Sources

This research was supported by internal departmental funds at the Gdansk Laboratory of Numerical Sciences.

Institutional Support

The authors thank the Gdansk Laboratory of Numerical Sciences, Gdansk, Poland, for computing resources and workspace.

Contributions Not Qualifying for Authorship

The authors acknowledge Dr. Anna Wojciechowska for assistance with QGIS spatial analysis and Mr. Tomasz Nowak for data formatting.

References

1. Hecky RE, Mugidde R, Ramlal PS, Talbot MR, Kling GW. Multiple stressors cause rapid ecosystem change in Lake Victoria. *Freshwater Biology*. 2010;55(Suppl 1):19–42.
2. Scheren PAGM, Zanting HA, Lemmens AMC. Estimation of water pollution sources in Lake Victoria. *Water Research*. 2000;34(4):1227–1235.
3. Di Francesco P, Mathieu P, Sénéchal D. *Conformal Field Theory*. New York: Springer; 1997.

4. Frenkel IB, Lepowsky J, Meurman A. *Vertex Operator Algebras and the Monster*. New York: Academic Press; 1988.
5. Gross D, Shortle JF, Thompson JM, Harris CM. *Fundamentals of Queueing Theory*. 4th ed. Hoboken (NJ): John Wiley & Sons; 2008.
6. Okely P, Imberger J, Antenucci JP. Processes affecting horizontal mixing and dispersion in Winam Gulf, Lake Victoria. *Limnology and Oceanography*. 2010;55(5):1865–1880.
7. Johnson RA, Wichern DW. *Applied Multivariate Statistical Analysis*. 6th ed. Upper Saddle River (NJ): Pearson Prentice Hall; 2007.
8. Driscoll TA. A MATLAB toolbox for Schwarz-Christoffel mapping. *ACM Transactions on Mathematical Software*. 1996;22(2):168–186.
9. Nyamweya CS, Desjardins C, Sigurdsson S, Gichuki J, Sitoki L, Gownaris NJ. Simulations of Lake Victoria circulation patterns. *PLOS ONE*. 2016;11(3).
10. Cox DR. Some statistical methods connected with series of events. *Journal of the Royal Statistical Society Series B*. 1955;17(2):129–164.
11. Carpenter SR, Ludwig D, Brock WA. Management of eutrophication for lakes subject to potentially irreversible change. *Ecological Applications*. 1999;9(3):751–771.
12. Carlson RE. A trophic state index for lakes. *Limnology and Oceanography*. 1977;22(2):361–369.
13. Vollenweider RA. Input-output models with special reference to the phosphorus loading concept in limnology. *Schweizerische Zeitschrift für Hydrologie*. 1975;37(1):53–84.
14. Kac VG. *Vertex Algebras for Beginners*. 2nd ed. Providence (RI): American Mathematical Society; 1998.