International Journal of Forest, Animal and Fisheries Research (IJFAF)
ISSN: 2456-8791
[Vol-10, Issue-1, Jan-Mar, 2026]
Issue DOI: https://dx.doi.org/10.22161/ijfaf.10.1
Article DOI: https://dx.doi.org/10.22161/ijfaf.10.1.2

Application of Pearson Correlation Analysis in Assessing Nutrient-Phytoplankton Dynamics: A Methodological Overview for Aquatic Managers

Ilakkiya S[1]*, Surulivel L[2], Dona Merli A R[3]

[1]PG Scholar, Department of Aquatic Environment Management, Dr. M.G.R Fisheries College and Research Institute, Ponneri, Thiruvallur, Tamil Nadu, India
[2]Assistant Professor, Department of Fisheries Extension, Economics and Statistics, Dr. M.G.R Fisheries College and Research Institute, Ponneri, Thiruvallur, Tamil Nadu, India
[3]PG Scholar, Department of Aquaculture, Dr. M.G.R Fisheries College and Research Institute, Ponneri, Thiruvallur, Tamil Nadu, India
Corresponding Author Mail: IlakkiyaS12@gmail.com

Received: 03 Jan 2026; Received in revised form: 03 Feb 2026; Accepted: 08 Feb 2026; Available online: 15 Feb 2026

***Abstract*— *Nutrient enrichment is a primary driver of phytoplankton production and community shifts in aquatic ecosystems. Quantitative evaluation of nutrient-phytoplankton relationships is fundamental for diagnosing eutrophication, predicting harmful algal blooms, and designing nutrient management strategies. Pearson correlation analysis is widely used as an initial statistical approach to examine linear associations between nutrient concentrations and phytoplankton metrics such as chlorophyll-a, cell density, and biovolume. This article provides a comprehensive methodological overview of Pearson correlation analysis in aquatic ecosystem assessment. It discusses theoretical foundations, assumptions, statistical workflow, interpretation strategies, ecological considerations, and limitations. The paper also outlines how aquatic managers can integrate correlation analysis into broader monitoring and decision-making frameworks.*

Keywords*— *Pearson correlation analysis, Nutrient limitation, Phytoplankton biomass, Eutrophication assessment, Aquatic ecosystem management

I. INTRODUCTION

Phytoplankton form the base of aquatic food webs and play a central role in nutrient cycling and primary productivity. Their growth is regulated by the availability of nitrogen (N), phosphorus (P), silica (Si), light, temperature, and hydrodynamic processes (Reynolds, 2006; Wetzel, 2001). Anthropogenic nutrient loading from agriculture, aquaculture, sewage discharge, and urban runoff often disrupts natural nutrient balances, leading to eutrophication and harmful algal blooms (Paerl et al., 2011; Smith & Schindler, 2009).

To understand these processes, aquatic monitoring programs routinely measure:

- Total nitrogen (TN)
- Nitrate (NO_3^-)
- Ammonium (NH_4^+)
- Total phosphorus (TP)
- Orthophosphate (PO_4^{3-})
- Dissolved silicate (SiO_2)
- Chlorophyll-a
- Phytoplankton density or biomass

Statistical tools are required to evaluate associations between nutrient availability and phytoplankton responses. Pearson correlation analysis remains one of the most commonly applied methods for this purpose due to its simplicity, interpretability, and strong theoretical foundation (Zar, 2010).

II. ECOLOGICAL BASIS OF NUTRIENT–PHYTOPLANKTON INTERACTIONS

Phytoplankton growth is frequently limited by nutrient availability. The concept of nutrient limitation is rooted in Liebig's Law of the Minimum, which states that growth is controlled by the scarcest essential resource (Wetzel, 2001). In freshwater systems, phosphorus often acts as the primary limiting nutrient, whereas nitrogen limitation is more common in marine environments (Smith & Schindler, 2009).

Additionally:

- High N:P ratios may promote phosphorus limitation.
- Low N:P ratios may favor nitrogen-fixing cyanobacteria.
- Silicate availability strongly influences diatom abundance (Reynolds, 2006).

Because these interactions can manifest as measurable statistical relationships, correlation analysis is frequently used as an exploratory tool to detect patterns between nutrient concentrations and phytoplankton biomass indicators.

III. STATISTICAL FOUNDATION OF PEARSON CORRELATION

The Pearson product-moment correlation coefficient (r) measures the degree of linear association between two continuous variables (Zar, 2010).

Mathematically, it is expressed as:

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$$

Where:

- x_i represents nutrient values
- y_i represents phytoplankton measurements
- $\bar{x}$ and $\bar{y}$ are sample means

The value of r ranges from −1 to +1:

- $r > 0$ indicates a positive linear relationship
- $r < 0$ indicates a negative linear relationship
- $r = 0$ indicates no linear association

The statistical significance of r is tested using a t-distribution, considering sample size (Zar, 2010).

IV. ASSUMPTIONS AND DATA CONSIDERATIONS IN AQUATIC SYSTEMS

Proper application of Pearson correlation requires careful examination of assumptions.

4.1 Linearity

The relationship between nutrient concentration and phytoplankton response must be approximately linear. Many ecological processes are nonlinear, especially under nutrient saturation or threshold conditions (Reynolds, 2006). Scatter plots should always be examined before computing correlation coefficients.

4.2 Normality

Pearson correlation assumes that both variables are approximately normally distributed. Nutrient data are often right-skewed and may require logarithmic transformation (Wetzel, 2001).

4.3 Homoscedasticity

Variance of phytoplankton response should remain relatively constant across nutrient gradients. Heteroscedastic data may distort correlation estimates.

4.4 Independence of Observations

Temporal monitoring datasets often contain autocorrelation. For example, monthly chlorophyll-a measurements may not be independent of previous months. Ignoring autocorrelation can inflate statistical significance (Legendre & Legendre, 2012).

V. METHODOLOGICAL WORKFLOW FOR MANAGERS

Step 1: Monitoring Design

- Establish consistent sampling stations.
- Ensure standardized depth and timing.
- Maintain adequate sample size.

Long-term monitoring strengthens inference reliability (Wetzel, 2001).

Step 2: Data Screening

- Check for outliers.
- Evaluate descriptive statistics.
- Apply transformations where necessary.

Step 3: Exploratory Visualization

Construct scatter plots with trend lines to verify linearity.

Step 4: Computation

Use statistical software (R, SPSS, SAS) to calculate r and associated p-values.

Step 5: Interpretation

Interpret both magnitude and ecological plausibility. A statistically significant correlation does not necessarily imply ecological importance (Zar, 2010).

VI. INTERPRETATION IN NUTRIENT MANAGEMENT CONTEXT

6.1 Identifying Limiting Nutrients

A strong positive correlation between TP and chlorophyll-a may indicate phosphorus-driven biomass accumulation, supporting phosphorus reduction strategies (Smith & Schindler, 2009).

6.2 Harmful Algal Bloom Prediction

Positive associations between ammonium and cyanobacterial biomass may suggest nitrogen forms influencing bloom formation (Paerl et al., 2011).

6.3 Species Composition Shifts

Strong correlations between silicate and diatom abundance may indicate silica-mediated community structure regulation (Reynolds, 2006).

VII. LIMITATIONS OF PEARSON CORRELATION IN AQUATIC ECOSYSTEMS

7.1 Correlation Does Not Establish Causation

Statistical association alone cannot confirm causality. Temperature, light, and hydrodynamics may influence both nutrients and phytoplankton simultaneously (Wetzel, 2001).

7.2 Nonlinear Dynamics

Phytoplankton often exhibit threshold responses or saturating growth patterns. Pearson correlation underestimates such relationships (Legendre & Legendre, 2012).

7.3 Multicollinearity

Nutrients are frequently interrelated. TN and TP may increase simultaneously, making it difficult to isolate independent effects through simple correlation.

7.4 Seasonal and Spatial Variability

Stratification, mixing events, and spatial heterogeneity can alter nutrient availability patterns, complicating interpretation (Reynolds, 2006).

VIII. COMPLEMENTARY ANALYTICAL APPROACHES

Because Pearson correlation is bivariate, it should be integrated with multivariate methods such as:

- Multiple regression analysis
- Principal component analysis (PCA)
- Canonical correspondence analysis (CCA)
- Time-series modelling

Multivariate approaches allow managers to account for interacting environmental drivers (Legendre & Legendre, 2012).

IX. MANAGEMENT IMPLICATIONS

When carefully applied, Pearson correlation analysis helps aquatic managers:

- Diagnose nutrient-driven eutrophication
- Prioritize nutrient reduction strategies
- Develop nutrient threshold guidelines
- Strengthen monitoring program design
- Identify early warning signals of bloom formation

However, interpretation must always integrate ecological theory, watershed knowledge, and hydrodynamic context.

X. WORKED STATISTICAL EXAMPLE USING A SIMULATED DATASET

To demonstrate how Pearson correlation is applied in real-world aquatic monitoring, this section presents a simulated monthly dataset from a freshwater reservoir experiencing moderate nutrient enrichment.

10.1 Scenario Description

An inland reservoir is monitored monthly for one year. The following parameters are measured:

- Total Phosphorus (TP, µg L^{-1})
- Total Nitrogen (TN, mg L^{-1})
- Chlorophyll-a (Chl-a, µg L^{-1}) as a proxy for phytoplankton biomass

The management objective is to determine whether phosphorus or nitrogen is more strongly associated with phytoplankton biomass.

10.2 Simulated Dataset (12 Months)

Month	TP (µg L^{-1})	TN (mg L^{-1})	Chl-a (µg L^{-1})
Jan	22	0.85	8
Feb	25	0.88	9
Mar	28	0.92	12
Apr	35	1.05	18
May	42	1.15	25
Jun	48	1.22	31
Jul	55	1.30	38
Aug	60	1.34	44
Sep	52	1.25	35
Oct	40	1.10	22
Nov	30	0.95	14
Dec	24	0.90	10

Visual inspection suggests increasing chlorophyll-a with increasing TP and TN.

10.3 Step 1: Exploratory Visualization

Scatterplots (conceptually described):

- TP vs Chl-a → Appears strongly linear and positive.
- TN vs Chl-a → Positive but slightly less steep relationship.

Linearity assumption appears satisfied.

10.4 Step 2: Calculation of Pearson Correlation

Using the Pearson correlation formula:

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$$

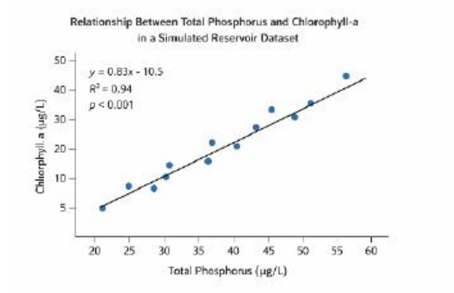
Correlation Between TP and Chlorophyll-a

After computation:

- r (TP, Chl-a) = 0.97
- $p < 0.001$

Interpretation:

This indicates a very strong positive linear relationship between total phosphorus and phytoplankton biomass.

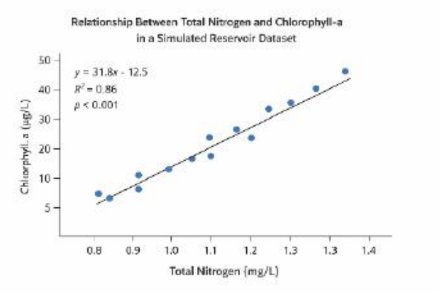


Correlation Between TN and Chlorophyll-a

After computation:

- r (TN, Chl-a) = 0.93
- $p < 0.001$

Interpretation:
There is also a very strong positive relationship between nitrogen and phytoplankton biomass, though slightly weaker than phosphorus.



10.5 Step 3: Statistical Significance Testing

The test statistic for correlation:

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

Where:

- n = 12

- degrees of freedom = 10

For TP:

- $t \approx 12.3$
- Highly significant at $\alpha = 0.05$

For TN:

- $t \approx 8.1$
- Also, highly significant

Thus, both nutrients show statistically significant relationships with phytoplankton biomass.

10.6 Ecological Interpretation

Although both nutrients are strongly correlated with chlorophyll-a, phosphorus shows the stronger association ($r = 0.97$ vs $r = 0.93$).

This suggests:

- The reservoir may be phosphorus-driven.
- Nutrient reduction strategies focusing on phosphorus inputs could yield substantial reductions in phytoplankton biomass.
- The high correlation also suggests a classic eutrophication pattern (Wetzel, 2001; Smith & Schindler, 2009).

However, because TN and TP also increase together, multicollinearity is possible. Further regression analysis would be required to isolate independent effects.

10.7 Testing Assumptions

Linearity

Scatterplots confirm linear trend.

Normality

Log transformation is not required in this simulated dataset because distributions are approximately symmetric. In real datasets, log10(TP) and log10(Chl-a) are often used.

Independence

Monthly time-series data may contain autocorrelation. A Durbin-Watson test or time-series modelling would be recommended for rigorous analysis (Legendre & Legendre, 2012).

10.8 Management Implications from Example

From this analysis, aquatic managers may conclude:

1. Strong nutrient-biomass coupling exists.
2. Phosphorus appears slightly more influential.
3. Bloom prediction models can incorporate TP thresholds.
4. Monitoring frequency should be maintained during peak summer months.

If this were a real reservoir, management actions might include:

- Reducing watershed phosphorus runoff.
- Upgrading wastewater treatment plants.
- Implementing buffer zones in agricultural catchments.

10.9 Extension: Detecting Nonlinearity

If chlorophyll-a plateaued beyond 55 µg L^{-1} TP, Pearson correlation would decrease despite strong ecological linkage. In such cases:

- Non-linear regression
- Generalized additive models (GAMs)

would provide a better model fit.

This highlights that correlation analysis is an exploratory, not definitive, analytical tool.

XI. CONCLUSION

Pearson correlation analysis remains a foundational statistical tool for examining nutrient-phytoplankton interactions in aquatic ecosystems. When applied with proper attention to assumptions, data screening, and ecological interpretation, it provides meaningful preliminary insight into nutrient limitation patterns and eutrophication processes.

The worked example demonstrates how correlation analysis can quantify nutrient-biomass relationships and guide management decisions, such as prioritising phosphorus reduction strategies. However, aquatic ecosystems are influenced by interacting environmental drivers, seasonal variability, and non-linear dynamics. Therefore, correlation results must be interpreted within a broader ecological framework and, where necessary, supplemented with multivariate or time-series approaches.

For aquatic managers, Pearson correlation serves not as a final diagnostic tool but as an evidence-based starting point in adaptive water quality management and nutrient control planning.

REFERENCES

[1] Legendre, P., & Legendre, L. (2012). *Numerical ecology* (3rd ed.). Elsevier.

[2] Paerl, H. W., Hall, N. S., & Calandrino, E. S. (2011). Controlling harmful cyanobacterial blooms in a world experiencing anthropogenic and climatic-induced change. *Science of the Total Environment, 409*(10), 1739–1745. https://doi.org/10.1016/j.scitotenv.2011.02.001

[3] Reynolds, C. S. (2006). *The ecology of phytoplankton*. Cambridge University Press. https://doi.org/10.1017/CBO9780511542145

[4] Smith, V. H., & Schindler, D. W. (2009). Eutrophication science: Where do we go from here? *Trends in Ecology & Evolution, 24*(4), 201–207. https://doi.org/10.1016/j.tree.2008.11.009

[5] Wetzel, R. G. (2001). *Limnology: Lake and river ecosystems* (3rd ed.). Academic Press. https://doi.org/10.1016/C2009-0-02112-6

[6] Zar, J. H. (2010). *Biostatistical analysis* (5th ed.). Pearson.