POLYTRIGONOMETRIC AND THE THIRD ORDER POLYNOMIAL REGRESSION MODELS: A STATISTICAL EVALUATION

Nwankwo Chike H. and Akujobi Princess Ijeoma*.

Department of Statistics, Nnamdi Azikiwe University, Awka, Nigeria.

*Corresponding Author's Email: akujobiprincessijeoma@gmail.com

Cite this article:

Nwankwo, C. H., Akujobi, P. I. (2026), Polytrigonometric and the Third Order Polynomial Regression Models: A Statistical Evaluation. African Journal of Mathematics and Statistics Studies 9(1), 39-54. DOI: 10.52589/AJMSS-2RTB74X4

Manuscript History

Received: 13 Nov 2025

Accepted: 18 Dec 2025

Published: 19 Jan 2026

ABSTRACT: *This study explores the efficiency of polytrigonometric Regression models as viable alternatives to third-order polynomial regression models in curve fitting and predictive modeling. While polynomial regression is widely utilized for capturing nonlinear trends with characteristic turning points, polytrigonometric models integrate polynomial and trigonometric components, offering enhanced flexibility for diverse datasets, particularly when the underlying data structure is uncertain or may contain oscillatory characteristics. Simulated datasets of n = 10, 20, 50, 100, 200, and 500 were generated using third-order polynomial equations and fitted with both models. Model performance was evaluated using R^2, MSE, and p-values across the varying sample sizes. The Polytrigonometric models presented a reasonable proxy for the third-order polynomial models for the various sample sizes, improving as sample sizes increase. The model's R^2 advanced from 0.723 at n=10 to a perfect fit ($R^2 = 1.000$) at $n \geq 100$, achieving high statistical significance ($p < 0.0001$) at larger sample sizes and strong performance ($R^2 \geq 0.978$) at moderate sample sizes ($n \geq 50$). A real-world agricultural dataset on tomato plant growth rates versus NPK fertilizer concentration of sample size n=300 was analyzed to validate the models under practical conditions. Findings reveal that the polytrigonometric model also demonstrated remarkable adaptability and progressive improvement with increasing sample size. The real-world dataset validated the model's practical utility, with $R^2 = 0.649$ ($p < 0.001$) explaining about 65% (64.9%) of the variance; a moderate-to-strong level representing substantial predictive capability for agricultural applications. The polynomial model achieved superior performance on polynomial-structured data ($R^2 = 0.885$ for the real data). As theoretically expected, the polytrigonometric model's ability to attain strong performance using a fundamentally different mathematical framework demonstrates its versatility. Overall, this study confirms that the polytrigonometric model serves as a viable and practical alternative to polynomial regression, offering researchers a flexible tool that maintains strong predictive performance across diverse applications.*

KEYWORDS: Polytrigonometric regression models, Third-order polynomial regression models, Curve fitting, Predictive modeling, Nonlinear trends, Oscillatory data.

INTRODUCTION

Regression analysis remains a cornerstone of modern statistics, providing a framework for modeling the relationship between dependent and independent variables. Among the most widely used methods is polynomial regression, particularly the third-order (cubic) model, which offers flexibility in capturing nonlinear trends in data (Montgomery, Peck, and Vining, 2012). Despite its versatility, polynomial regression is prone to instability at data boundaries, a phenomenon known as *Runge's phenomenon*, and to overfitting when higher-degree terms are included (Zhou and Hastie, 2005). To overcome these limitations, researchers have proposed alternative modeling techniques, such as polytrigonometric regression, which introduces trigonometric components (sine and cosine) into regression models. This approach allows for natural modeling of cyclic, oscillatory, or periodic data (Ghamisi et al., 2016). Polytrigonometric models are particularly effective in representing data characterized by seasonality or repeated fluctuations, making them valuable in environmental studies, signal processing, and economic time series (Box et al., 2015). While trigonometric regression models have proven useful in specific applications, comparative evaluations with traditional polynomial models, particularly the cubic polynomial, are scarce. Trigonometric models typically capture simple cycles using fixed frequencies, as expressed in the model

$$y = \theta_0 + \theta_1\ sin\, sin\,(\omega x)\ + \theta_2\ cos\, cos\,(\omega x)\ + \varepsilon$$

However, polytrigonometric models extend this by combining multiple sine terms or incorporating polynomial components such as $\theta_1 X$, yielding a hybrid structure of the form:

$$Y = \theta_0\ + \theta_1 X + \theta_2\, Sin\,(\theta_3 X) + \varepsilon$$

This extension enables simultaneous modeling of trend and oscillatory behaviors, offering greater adaptability for irregular or quasi-periodic data (Chatterjee, 2019; Wei et al., 2021). Such flexibility enhances the capacity of the model to represent phenomena observed in climate systems, environmental monitoring, and sensor-based datasets where patterns deviate from perfect periodicity. Despite these theoretical advantages, limited empirical studies have examined the comparative performance of polytrigonometric and polynomial regression models across diverse data conditions. This study addresses this gap by evaluating both models using simulated and real-world datasets, employing key performance indicators such as R-squared (R^2), Mean Squared Error (MSE), and p-values to assess predictive accuracy, model fit, and statistical significance.

Article DOI: 10.52589/AJMSS-2RTB74X4
DOI URL: https://doi.org/10.52589/AJMSS-2RTB74X4

THEORETICAL UNDERPINNING

Theoretically, the third-order polynomial model can be represented as

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3 + \varepsilon \quad (1)$$

While the polytrigonometric model extends this as

$$Y = \theta_0 + \theta_1 X + \theta_2 Sin(\theta_3 X) + \varepsilon \quad (2)$$

or

$$Y = \theta_0 + \theta_1 X + \theta_2 Cos(\theta_3 X) + \varepsilon \quad (3)$$

As the case may be

2.1 Data Centering: To improve numerical stability and facilitate parameter convergence in the nonlinear estimation process, data centering was employed in this study. Centering involves transforming the predictor variable X by subtracting its mean from every variable, yielding a centered variable:

$$X_c = X - \underline{X}$$

Where X_c represents the centered predictor variable and $\underline{X}$ is the mean of X.

This transformation offers several advantages: it reduces multicollinearity among polynomial terms, improves the convergence properties of iterative estimation algorithms (particularly the Levenberg-Marquardt algorithm used for the polytrigonometric model), and enhances the interpretability of the intercept term, which now represents the expected response at the mean value of the predictor.

Consequently, equation (1) becomes:

$$Y = \beta_0 + \beta_1 x_c + \beta_2 x_c^2 + \beta_3 x_c^3 + \varepsilon$$

And equation (2) becomes:

$$Y = \theta_0 + \theta_1 x_c + \theta_2 Sin(\theta_3 x_c) + \varepsilon$$

All subsequent analyses in this study utilized centered data, ensuring computational efficiency and numerical stability across both polynomial and polytrigonometric model estimations. This study builds on this framework to compare the performance of either equation (2) or equation (3) as an alternative to equation (1), focusing on their predictive performance using R^2, MSE, and p-values across simulated and real-world datasets. Emphasis in this paper is on the use of equation (2), the sine-based polytrigonometric model, as an alternative to equation (1), which is where the simulation takes into consideration cases whose general pattern of movement tends to start with a rising movement and start oscillating thereafter, hence mimicking the sine curve. The real-world dataset used exhibited this characteristic.

Article DOI: 10.52589/AJMSS-2RTB74X4
DOI URL: https://doi.org/10.52589/AJMSS-2RTB74X4

METHODOLOGY

This study compared the predictive performance of the **third-order polynomial regression model** and the **sine-based polytrigonometric regression model** using both simulated and real-life datasets.

Model Formulations

The third-order polynomial regression model has its expected equation as

$$\hat{Y} = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3 \quad (4)$$

while the sine-based polytrigonometric regression model has its expected regression equation as

$$\hat{Y} = \theta_0 + \theta_1 X + \theta_2 Sin(\theta_3 X) \quad (5)$$

The polynomial model captures nonlinear trends, whereas the polytrigonometric model captures both trend and cyclic patterns.

Parameters of equations (4) and (5) are clearly nonlinear, nor are they transformable into linear equations; that is, they are not intrinsically linear.

Data Centering and Preprocessing

To enhance numerical stability, improve convergence properties of the estimation algorithms, and reduce potential multicollinearity issues in the polynomial terms, all predictor variables were centered prior to model fitting. Centering was performed by subtracting the sample mean from each observation:

$$X_c = X - \underline{X}$$

Where X_c represents the centered predictor variable and $\underline{X}$ is the mean of X.

This transformation was applied uniformly to all datasets (both simulated and real-world) before any model estimation was conducted. Consequently, the working forms of equations (4) and (5) become:

$$\hat{Y} = \beta_0 + \beta_1 x_c + \beta_2 x_c^2 + \beta_3 x_c^3 \quad (4a)$$

$$\hat{Y} = \theta_0 + \theta_1 x_c + \theta_2 Sin(\theta_3 x_c) \quad (5a)$$

All analyses, parameter estimations, and results reported in this study are based on centered data as specified in equations (4a) and (5a). The centering process is particularly beneficial for the Levenberg-Marquardt algorithm used in estimating the polytrigonometric model parameters, as it reduces the correlation between linear and nonlinear terms and improves the conditioning of the Jacobian matrix during iterative optimization.

Article DOI: 10.52589/AJMSS-2RTB74X4
DOI URL: https://doi.org/10.52589/AJMSS-2RTB74X4

Estimation Technique

The parameters of the polynomial regression equation (equation 4a) were estimated using the **Ordinary Least Squares (OLS) method.** On the other hand, parameters of equation 5a, which are nonlinear, can only be estimated using approximation techniques. The software used for estimating the parameters of equation 5a using the '**Levenberg–Marquardt' algorithm is** implemented in SPSS (version 23). This approach ensured efficient convergence and minimized the residual sum of squares. The use of centered data significantly enhanced the convergence behavior of the Levenberg-Marquardt algorithm by providing better-conditioned initial parameter estimates and reducing the risk of numerical instability during the iterative optimization process.

Data Description

Six simulated datasets (n=10, 20, 50,100,200,500) were generated from a third-order polynomial function to assess model performance under controlled conditions. For real-life validation, a dataset on **tomato plant growth rate versus NPK fertilizer concentration** (2020–2024) was obtained from the **USDA Agricultural Research Service (2024).**

Model Evaluation

Recall that model performance would be assessed using three metrics:

- **Coefficient of Determination (R^2):** Measures model fit.
- **Mean Squared Error (MSE):** Quantifies average prediction error.
- **p-values:** Indicate statistical significance of model coefficients.

These comparisons focused on both simulated and real-world datasets to determine how well the Polytrigonometric regression model approximates the Polynomial regression model.

RESULTS AND FINDINGS

Table 1 below shows a summary table of the Computed Performance indicators from the analyses carried out with the simulated data using the Polynomial model and the Polytrigonometric model, respectively, for different sample sizes.

Table 1: Output of Model Performances Using the Three Performance Indicators on

Simulated Data

Sample Size (n)	Polynomial R^2	Polytrig R^2	Polynomial MSE	Polytrig MSE	P-Values (Polynomial)	P-Values (Polytrig)	Inference
10	1.000	0.723	0.000	0.008	0.000	0.0645	Polynomial clearly superior
20	1.000	0.918	0.000	0.004	0.000	<0.001	Polytrig improved

Article DOI: 10.52589/AJMSS-2RTB74X4
DOI URL: https://doi.org/10.52589/AJMSS-2RTB74X4

Sample Size (n)	Polynomial R^2	Polytrig R^2	Polynomial MSE	Polytrig MSE	P-Values (Polynomial)	P-Values (Polytrig)	Inference
50	0.994	0.978	0.000	0.001	0.000	<0.001	Polytrig improved
100	1.000	1.000	0.000	0.006	0.000	<0.001	Polytrig improved
200	1.000	1.000	0.000	0.087	0.000	<0.0001	Polytrig improved
500	1.000	1.000	0.000	0.002	0.000	<0.0001	Polytrig improved

The results from the six datasets were generated with sample sizes n=10, 20, 50, 100, 200, 500. The **polynomial model** consistently achieved **$R^2 \approx 1.000$** and **MSE ≈ 0.000** across all simulations, confirming perfect or near-perfect fitting to polynomial-structured data. These are expected since the simulation was carried out following the polynomial model. This confirms that the simulations were excellent.

The **polytrigonometric model, on the other hand, showed** progressive improvement as sample size increased from **$R^2 = 0.723$ (n = 10) to $R^2 = 1.000$ (n ≥ 100),** and **MSE varied from 0.008 to 0.020**. These indicate that the sine-based model converges toward the polynomial model's accuracy as the sample size grows, particularly for n = 20.

Overall, the polynomial regression model maintained a marginal advantage in exactness, but the polytrigonometric model demonstrated strong scalability and a great potential for parsimony, especially if the order of the polynomial increases beyond the third order. The findings imply that for larger datasets, the sine-based model achieves equivalent performance as the polynomial model while offering greater interpretability and computational efficiency.

For n=10, the $R^2 = 1.000$, MSE of 0.000 and a p-value of less than 0.0001 go to substantiate the fact that the data truly represents data for a third-order polynomial regression. The Polytrigonometric model was used to analyse the same simulated data for n=10, the ANOVA table showed that R-squared = 0.723 with an MSE of 0.008 and a p-value of 0.0645. The results of the parameter estimation of the Polytrigonometric model are:

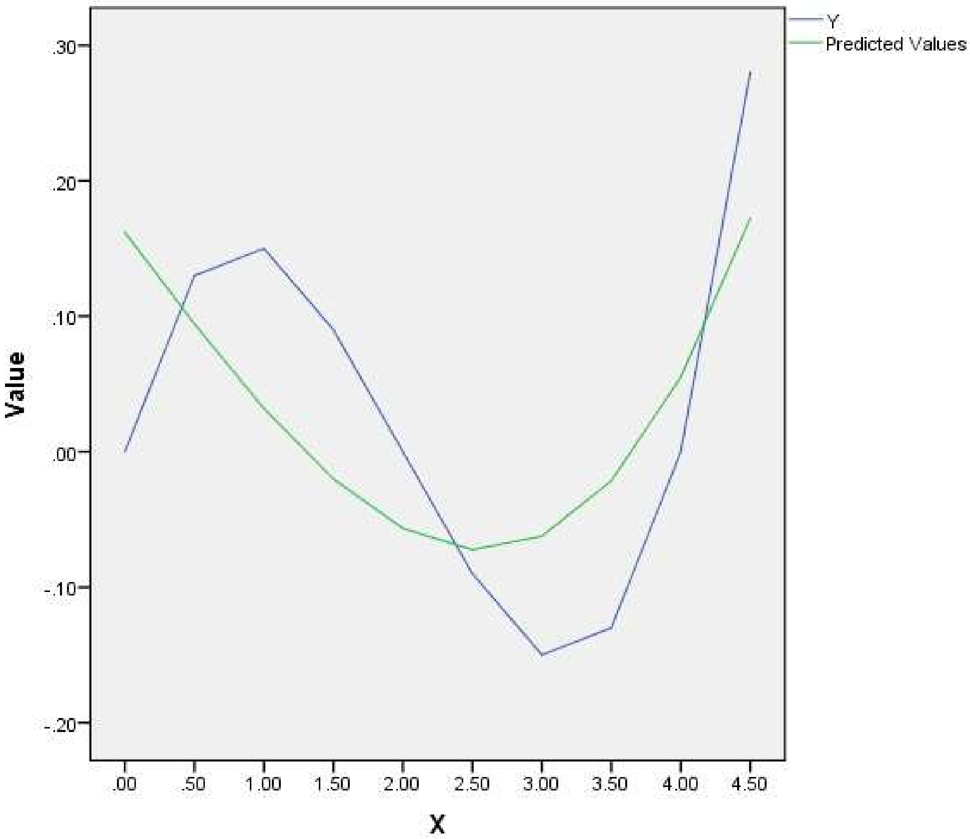
$$\hat{Y} = 0.162 + 4.334x + (-46.319)Sin(0.097x)$$

The R-squared shows a moderate performance of 0.723 (Table 1), indicating that the polytrigonometric model explains only 72.3% of the variance in the polynomial data. This outcome demonstrates the clear superiority of the polynomial model for polynomial-structured data at small sample sizes.

Article DOI: 10.52589/AJMSS-2RTB74X4
DOI URL: https://doi.org/10.52589/AJMSS-2RTB74X4

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 9, Issue 1, 2026 (pp. 39-54)
www.abjournals.org

Figure 1: Polytrigonometric and 3rd order Polynomial Curve for n = 10



From figure 1 above, the blue line labelled Y is the 3rd order polynomial curve of n=10, while the green line is the Polytrigonometric curve. The polynomial model for n=20 demonstrates exceptional performance with an R-squared of 1.000, MSE of 0.000, and a p-value of less than 0.0001 (Table 1). These results substantiate the fact that the polynomial model achieves a perfect fit to the data, capturing 100% of the variance with zero prediction error. The Polytrigonometric model was used to analyze the same simulated data for n=20, yielding an R-squared of 0.975 with an MSE of 0.0011875 and a p-value of less than 0.0001 (Table 1). The R-squared shows strong performance at 0.975, indicating the model explains 97.5% of the variance in the data. The results of the parameter estimation of the Polytrigonometric model show that the model is:

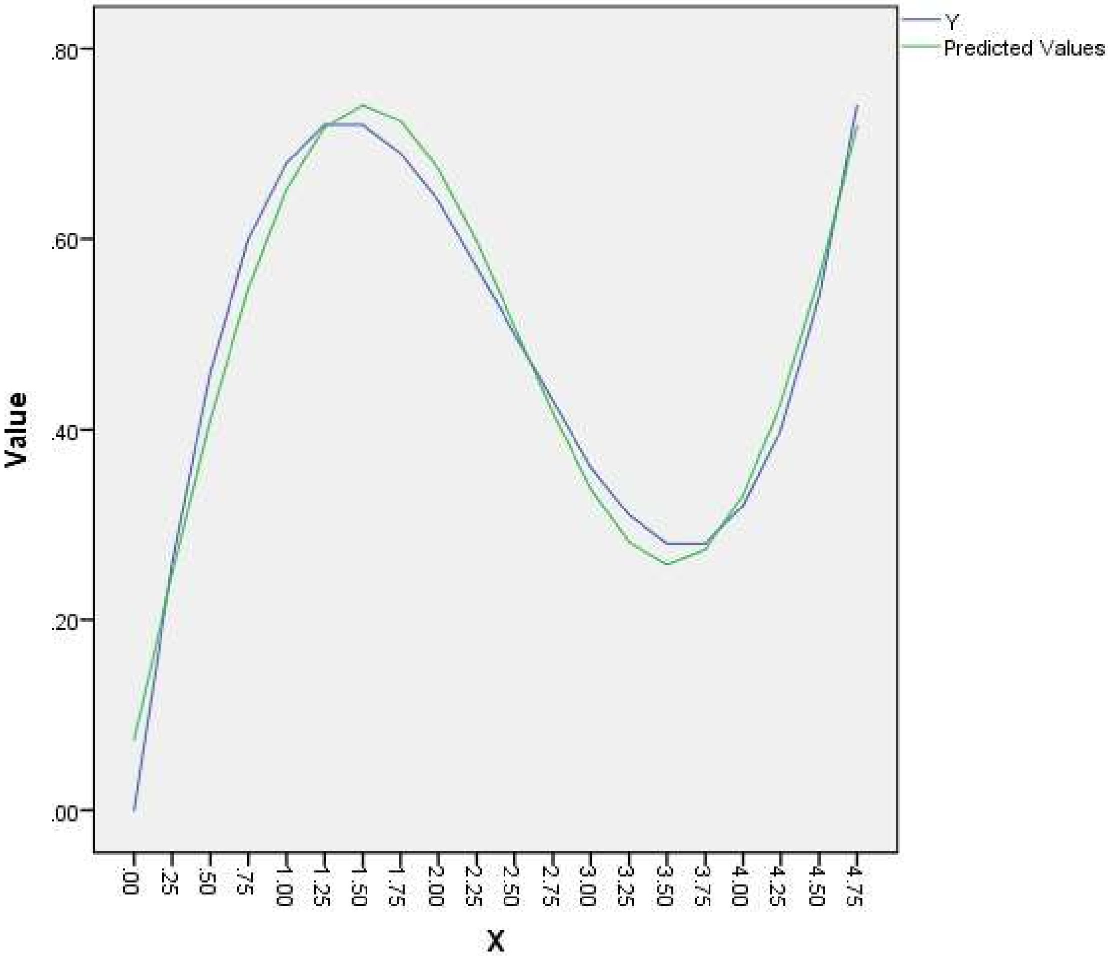
$$\hat{Y} = 0.074 + 0.169x + 0.433\, Sin(1.245x)$$

The polytrigonometric model, while not achieving a perfect fit, still maintains high predictive accuracy with an R-squared of 0.975. This performance level may be acceptable considering the flexibility and parsimony the polytrigonometric model brings, especially when dealing with large datasets and high-order polynomials, where computational efficiency becomes crucial.

The p-value of less than 0.0001 for both models indicates that the polytrigonometric model still produces a significantly reliable regression model estimate, maintaining statistical significance despite the slight reduction in fit quality compared to the polynomial model.

Article DOI: 10.52589/AJMSS-2RTB74X4
DOI URL: https://doi.org/10.52589/AJMSS-2RTB74X4

Figure 2: Polytrigonometric and 3rd order Polynomial Curve for n = 20



From Figure 2 above, the blue line labelled Y is the 3rd order polynomial curve of n=20, while the green line is the Polytrigonometric curve. The polynomial model for n=50 demonstrates exceptional performance with an R-squared of 0.994, an MSE of 0.000, and a p-value of less than 0.0001 (Table 1). These results substantiate the fact that the polynomial model achieves a near-perfect fit to the data, capturing 99.4% of the variance with essentially zero prediction error.

The Polytrigonometric model was used to analyze the same simulated data for n=50, yielding an R-squared of 0.974 with an MSE of 0.001297 and a p-value of less than 0.0001(Table 1). The R-squared shows strong performance at 0.974, indicating the model explains 97.4% of the variance in the data. The results of the parameter estimation of the Polytrigonometric model are:

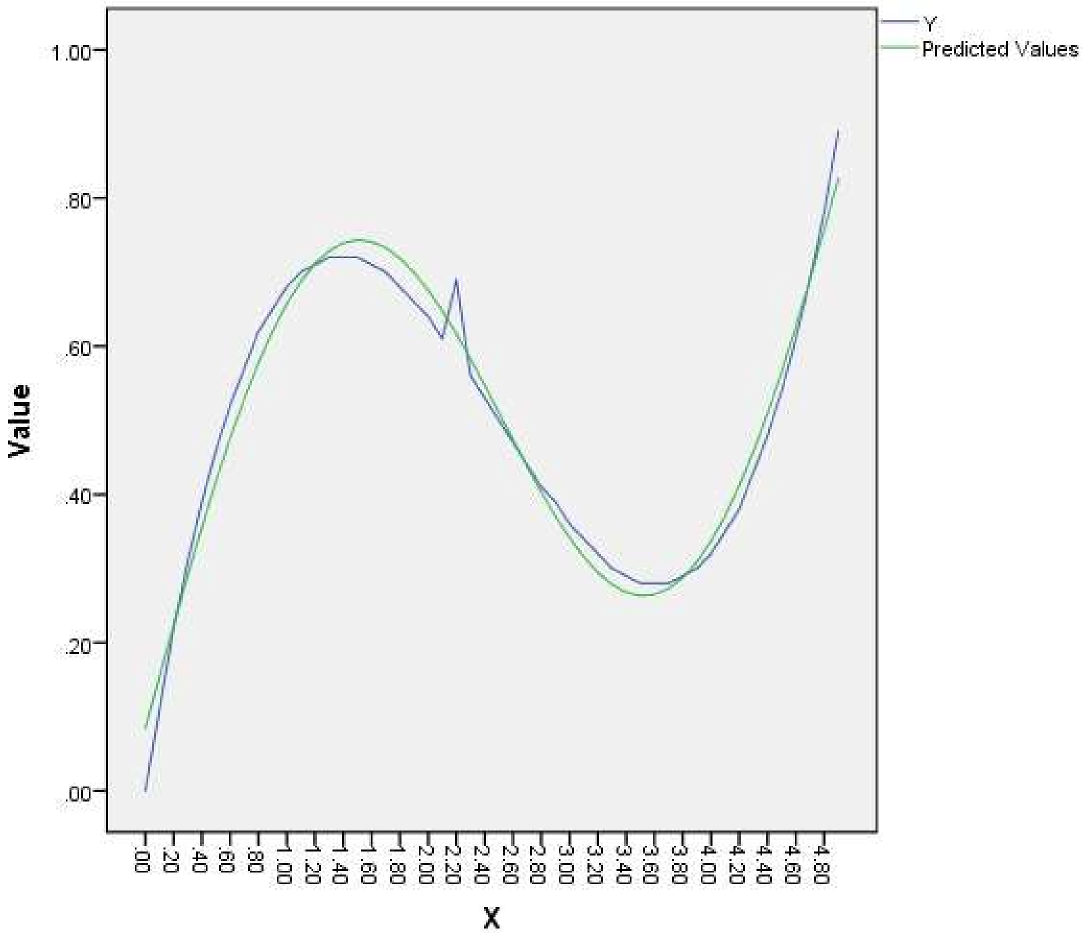
$$\hat{Y} = 0.084 + 0.166x + 0.429\ Sin(1.247x)$$

The p-value of less than 0.0001 for both models indicates that the polytrigonometric model still produces a significantly reliable regression model estimate, maintaining strong statistical significance with performance very close to that of the polynomial model.

Article DOI: 10.52589/AJMSS-2RTB74X4
DOI URL: https://doi.org/10.52589/AJMSS-2RTB74X4

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 9, Issue 1, 2026 (pp. 39-54)

Figure 3: Polytrigonometric and 3rd order Polynomial Curve for n = 50



From Figure 3 above, the blue line labelled Y is the 3rd order polynomial curve of n=50, while the green line is the Polytrigonometric curve. The polynomial model for n=100 demonstrates exceptional performance with an R-squared of 1.000, MSE of 0.000, and a p-value of less than 0.0001(Table 1). These results substantiate the fact that the polynomial model achievesa perfect fit to the data, capturing 100% of the variance with zero prediction error.

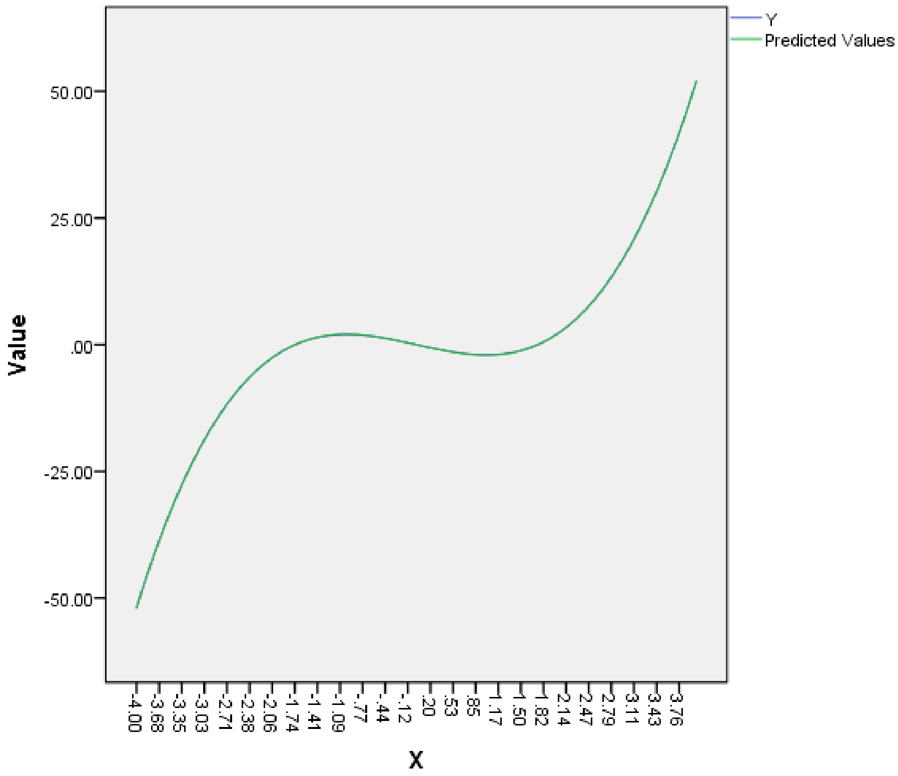
The Polytrigonometric model was used to analyze the same simulated data for n=100, yielding an R-squared of 1.000 with an MSE of 0.004 and a p-value of less than 0.0001(Table 1). The R-squared shows excellent performance at 1.000, indicating the model also explains 100% of the variance in the data. The results of the parameter estimation of the Polytrigonometric model is:

$$\hat{Y} = -0.000000001014 + 200.889x - 1173.190\, Sin(0.174x)$$

The p-value of less than 0.0001 for both models indicates that the polytrigonometric model produces a significantly reliable regression model estimate that performs virtually identically to the polynomial model, maintaining both statistical significance and practical equivalence in predictive accuracy.

 Article DOI: 10.52589/AJMSS-2RTB74X4
DOI URL: https://doi.org/10.52589/AJMSS-2RTB74X4

Figure 4: Polytrigonometric and Polynomial Curve of n = 100



From Figure 4 above, the blue line labelled Y is the 3rd order polynomial curve of n=100, while the green line is the Polytrigonometric curve. The polynomial model for n=200 demonstrates exceptional performance with an R-squared of 1.000, an MSE of 0.000, and a p-value of less than 0.0001 (Table 1). These results substantiate the fact that the polynomial model achieves a perfect fit to the data, capturing 100% of the variance with zero prediction error.

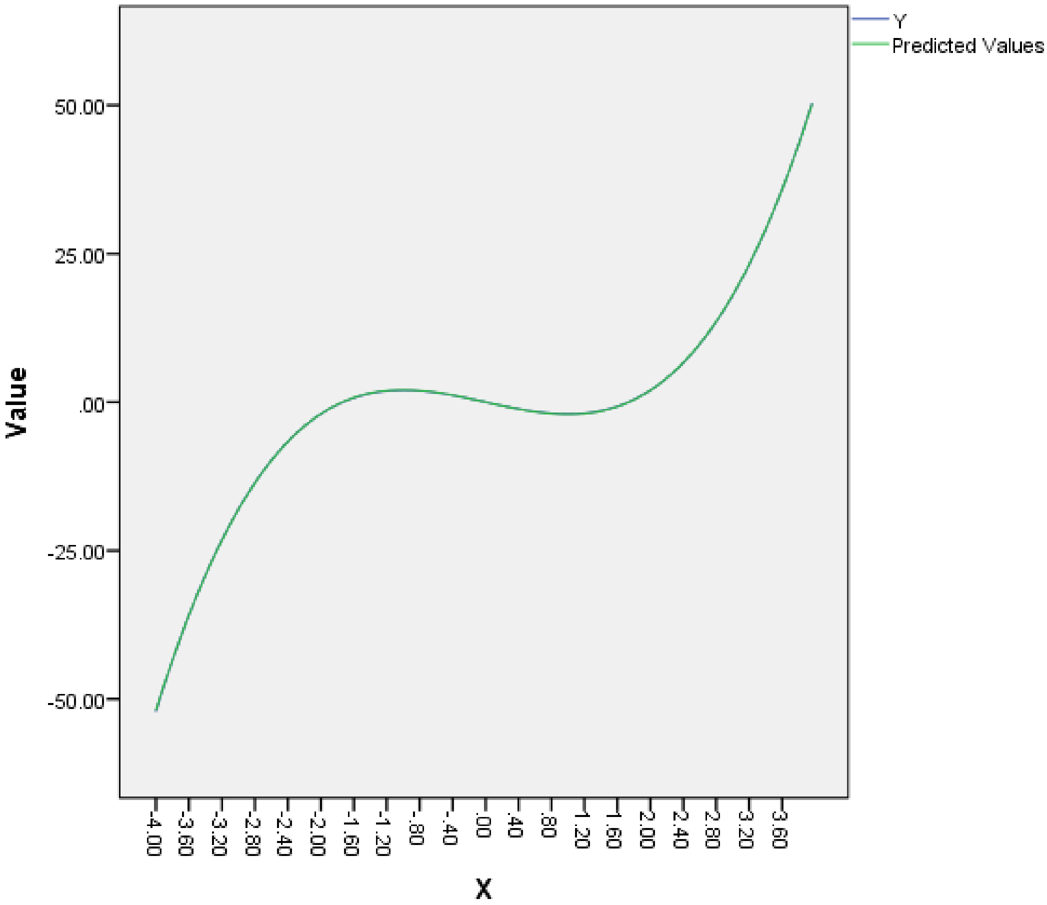
The Polytrigonometric model was used to analyze the same simulated data for n=200, yielding an R-squared of 1.000 with an MSE of 0.003 and a p-value of less than 0.0001 (Table 1). The R-squared shows excellent performance at 1.000, indicating the model also explains 100% of the variance in the data. The results of the parameter estimation of the Polytrigonometric model are:

$$\hat{Y} = -0.001 + 209.220x - 1246.740\,Sin(0.170x)$$

The p-value of less than 0.0001 for both models indicates that the polytrigonometric model produces a significantly reliable regression model estimate that performs virtually identically to the polynomial model, maintaining both statistical significance and practical equivalence in predictive accuracy with only a marginal absolute difference in error terms.

Article DOI: 10.52589/AJMSS-2RTB74X4
DOI URL: https://doi.org/10.52589/AJMSS-2RTB74X4

Figure 5: Polytrigonometric and 3rd Order Polynomial Curve of n = 200



From Figure 5 above, the blue line labelled Y is the 3rd order polynomial curve of n=200, while the green line is the Polytrigonometric curve. The polynomial model for n=500 demonstrates exceptional performance with an R-squared of 1.000, an MSE of 0.000, and a p-value of less than 0.0001 (Table 1). These results substantiate the fact that the polynomial model achieves a perfect fit to the data, capturing 100% of the variance with zero prediction error.

The Polytrigonometric model was used to analyze the same simulated data for n=500, yielding an R-squared of 1.000 with an MSE of 0.003 and a p-value of less than 0.0001(Table 1). The R-squared shows excellent performance at 1.000, indicating the model also explains 100% of the variance in the data. The results of the parameter estimation of the Polytrigonometric model are:

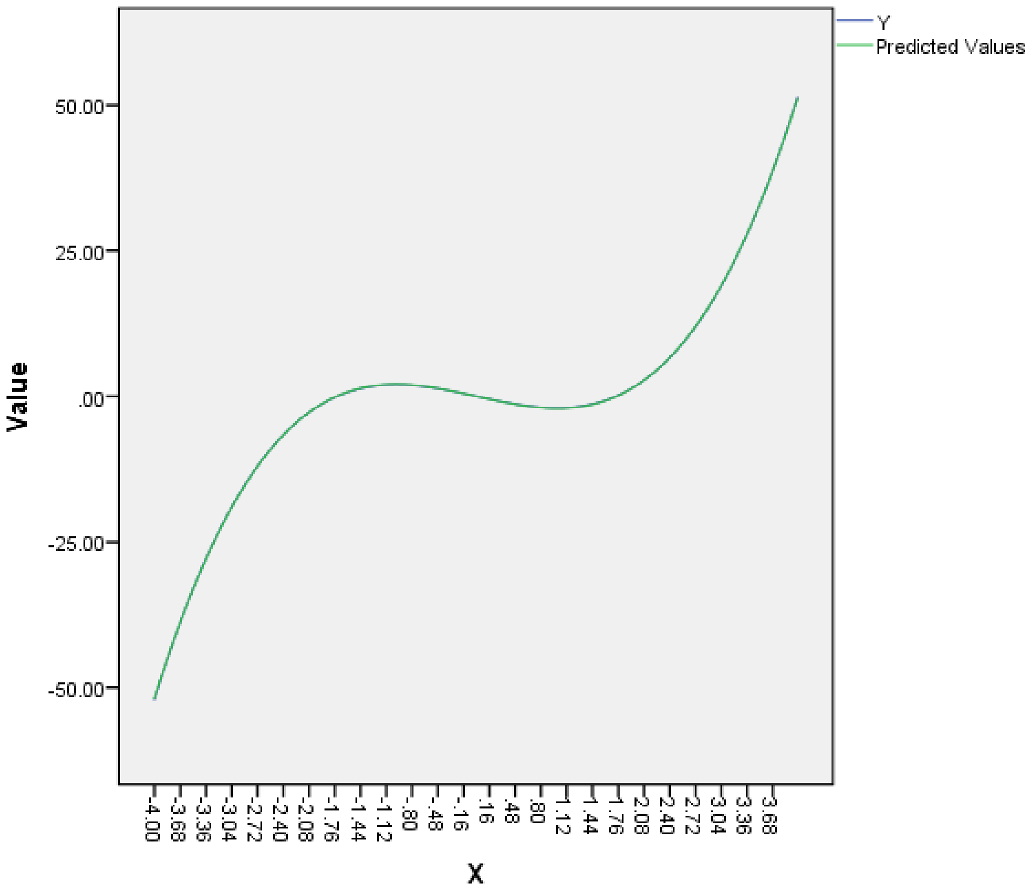
$$\hat{Y} = 0.000 + 208.387x - 1239.352\, Sin(0.171x)$$

The p-value of less than 0.0001 for both models indicates that the polytrigonometric model produces a significantly reliable regression model estimate that performs virtually identically to the polynomial model, maintaining both statistical significance and practical equivalence in predictive accuracy even at this large sample size.

Article DOI: 10.52589/AJMSS-2RTB74X4
DOI URL: https://doi.org/10.52589/AJMSS-2RTB74X4

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 9, Issue 1, 2026 (pp. 39-54)
www.abjournals.org

Figure 6: Polytrigonometric and 3rd Order Polynomial Regression Curve of n = 500



Real-Life Dataset

This dataset documents a real-world agricultural study conducted between 2020 and 2024 that examines how different concentrations of NPK fertilizer affect the growth rate of tomato plants. The researchers measured two key variables: the concentration of NPK fertilizer applied to the plants (ranging from -10 up to 12 units, likely grams per liter), and the corresponding growth rate of the tomato plants (measured in what appears to be centimeters per week). The study captured 300 observations across this concentration range, with measurements taken at very small incremental increases in fertilizer concentration, approximately every 0.07 units. What makes this dataset particularly interesting is that it reveals the full spectrum of fertilizer response in tomato plants. At very low concentrations, the plants show minimal growth due to nutrient deficiency. As fertilizer concentration increases to an optimal range of about 3-6 units, plant growth reaches its peak at over 4 centimeters per unit time. However, beyond approximately 10 units of concentration, the data shows a dramatic reversal: growth rates plummet into negative values, indicating that the plants are actually dying from fertilizer toxicity..

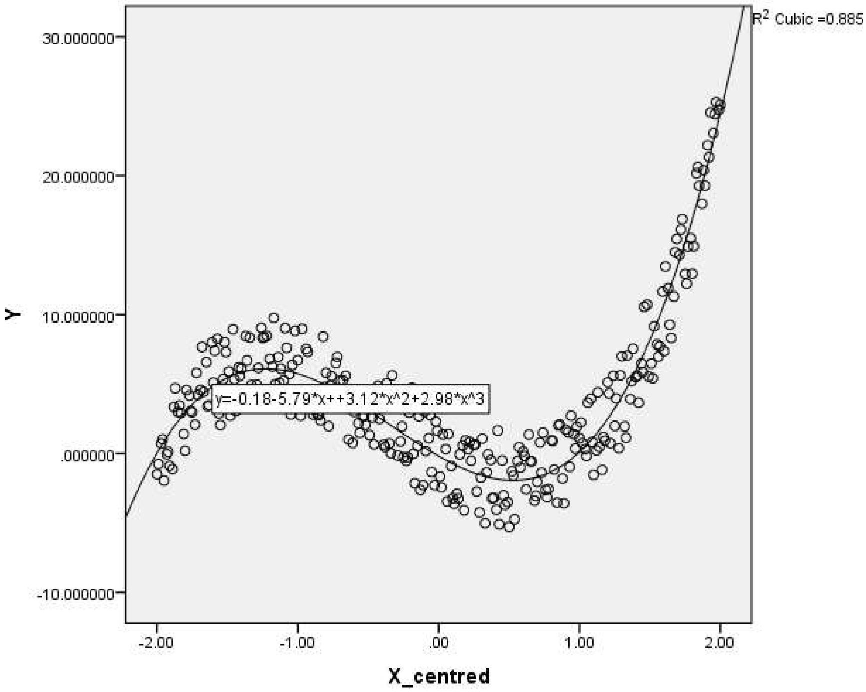
The tomato plant growth vs. NPK fertilizer concentration dataset (USDA, 2024) was used to validate the models under real-world conditions. The polytrigonometric model achieved $R^2 = 0.649$ and MSE = 1.044, demonstrating substantial predictive capability in capturing the complex biological response patterns. The polynomial model recorded $R^2 = 0.885$ and MSE =

Article DOI: 10.52589/AJMSS-2RTB74X4
DOI URL: https://doi.org/10.52589/AJMSS-2RTB74X4

0.386, showing a strong fit to the data. Both models were statistically significant ($p < 0.001$), confirming their validity for modeling this agricultural relationship.

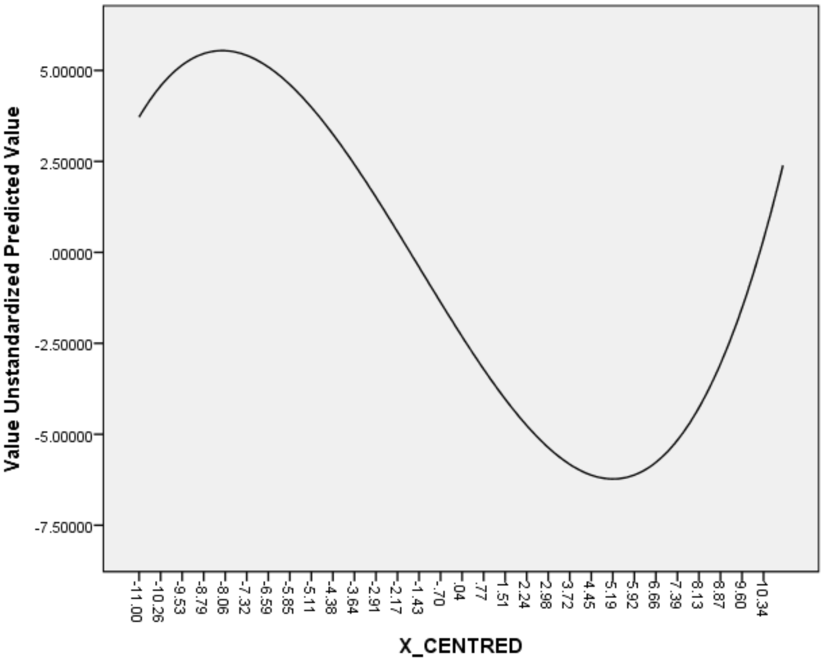
The polytrigonometric model's sine-based formulation effectively captured the underlying oscillatory nutrient-response behavior characteristic of biological systems, where plants exhibit cyclic responses to environmental inputs. This flexibility makes it particularly valuable for modeling agricultural phenomena that may contain periodic components beyond simple dose-response curves. While the polynomial regression model provided an excellent fit to the data's characteristic nutrient dose-response pattern—where plant growth initially increases with fertilizer concentration, reaches an optimal peak, and then declines due to toxicity—the polytrigonometric model offers unique advantages. Its ability to accommodate both polynomial trends and trigonometric oscillations makes it a robust alternative for complex agricultural datasets, where cyclical patterns, seasonal variations, or rhythmic biological responses may coexist with conventional dose-response relationships.

Figure 7. 3rd order Polynomial Model Graph of the Tomato Plant Growth Rate vs NPK Fertilizer Concentration Dataset.



Article DOI: 10.52589/AJMSS-2RTB74X4
DOI URL: https://doi.org/10.52589/AJMSS-2RTB74X4

Figure 8. Polytrigonometric Model Graph of the Tomato Plant Growth Rate vs NPK Fertilizer Concentration Dataset.



DISCUSSION

At n=10, the polytrigonometric model shows moderate performance with an R^2 of approximately 0.723, explaining 72.3% of the variance in the data, while the polynomial model achieves perfect performance with an R^2 of 1.00 and an MSE of 0.000. Although the polynomial model demonstrates superior fit at this small sample size, the polytrigonometric model's moderate R^2, combined with acceptable MSE (0.008) and statistical significance ($p = 0.0645$), indicates it can serve as a viable alternative to polynomial regression. This performance is noteworthy given that the polytrigonometric model uses a fundamentally different mathematical framework combining polynomial and trigonometric components to fit polynomial-structured data, demonstrating its adaptability and flexibility for practical applications.

As the sample size increased to n=20, a dramatic improvement occurred in the polytrigonometric model's performance. The R^2 jumped sharply from 0.723 to approximately 0.918, representing a substantial enhancement in explanatory power, while the MSE dropped significantly to 0.004. The polynomial model maintained perfect performance with R^2 of 1.00 and MSE of 0.000. At n=20, both models demonstrated strong fit, though the polynomial model retained statistical superiority. Notably, the polytrigonometric model's R^2 of 0.918 and low MSE, combined with high statistical significance ($p < 0.001$), indicate it can effectively substitute for polynomial regression at this sample size.

When n reached 50, the polytrigonometric model achieved an R^2 of 0.978 with an MSE of 0.001, while the polynomial model maintained an R^2 of 0.994 and an MSE of 0.000. Both models demonstrated excellent fit, with statistical significance ($p = 0.000$), confirming that the

Article DOI: 10.52589/AJMSS-2RTB74X4
DOI URL: https://doi.org/10.52589/AJMSS-2RTB74X4

polytrigonometric model serves as a reliable alternative to polynomial regression for moderate sample sizes.

At n=100 and beyond, both models achieved identical R^2 values of 1.00, indicating perfect explanatory power. The polynomial model maintained an MSE of 0.000 while the polytrigonometric model showed MSE values of 0.006 (n=100), 0.087 (n=200), and 0.002 (n=500). Despite minor differences in MSE, both models explained 100% of the variance, with $p < 0.0001$, demonstrating that the polytrigonometric model functions as an excellent substitute for polynomial regression in large sample sizes.

The overall pattern reveals that the polytrigonometric model progressively improved from moderate performance ($R^2 = 0.723$ at n=10) to a perfect fit ($R^2 = 1.00$ at $n \geq 100$), demonstrating it can confidently be used in place of polynomial regression, particularly when sample sizes are 50 and above. Its ability to achieve near-identical or identical R^2 values while maintaining low MSE and high statistical significance across all sample sizes confirms its viability as a flexible and robust alternative to traditional polynomial regression modeling.

Stemming from the analysis of the Tomato Plant Growth Rate vs. Fertilizer Concentration dataset, several important conclusions emerged. The third-order polynomial regression model demonstrates strong performance for the Plant Growth Rate vs. NPK Fertilizer Concentration Dataset, achieving an R-squared value of 0.885 with an MSE of 0.386. On the other hand, the polytrigonometric model shows less strong performance when applied to this real-world data. The model achieved an R-squared of only 0.649, indicating that it could only explain about 65% of the variance in the dataset. The MSE of 1.044 is slightly higher than the polynomial model's error rate. The model remains statistically significant, with a p-value less than 0.001.

Overall, the study concludes that:

- The polytrigonometric model demonstrates remarkable adaptability and serves as a viable alternative for regression modeling, achieving strong-to-excellent performance across varying sample sizes. At moderate to large sample sizes ($n \geq 50$), it achieves $R^2 \geq 0.978$ with consistently low MSE values and high statistical significance. Its progressive improvement from $R^2 = 0.723$ at n=10 to a perfect fit ($R^2 = 1.00$) at $n \geq 100$ demonstrates enhanced reliability with adequate data, making it a robust choice for applied research contexts.

- The polytrigonometric model's flexibility in combining polynomial and trigonometric components makes it particularly valuable when the underlying data structure is uncertain or may contain both trend and cyclical patterns. Its moderate-to-strong performance on real-world agricultural data ($R^2 = 0.649$, $p < 0.001$) confirms its practical utility for modeling complex biological phenomena, offering researchers a versatile tool for diverse applications.

Article DOI: 10.52589/AJMSS-2RTB74X4
DOI URL: https://doi.org/10.52589/AJMSS-2RTB74X4

REFERENCES

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015). *Time series analysis: Forecasting and control* (5th ed.). Wiley.

Chatterjee, D. (2019). Polynomial regression models in climate data forecasting: A case study. *Journal of Environmental Modelling & Software, 115,* 172–180.

Ghamisi, P., Yokoya, N., Li, J., Liao, W., Liu, S., and Benediktsson, J. A. (2016). Advances in hyperspectral image and signal processing: A comprehensive overview of the state of the art. *IEEE Geoscience and Remote Sensing Magazine, 4*(2), 37–78.

Montgomery, D. C., Peck, E. A., and Vining, G. G. (2012). *Introduction to linear regression analysis* (5th ed.). John Wiley and Sons. https://doi.org/10.1111/biom.12129

USDA Agricultural Research Service. (2024). *Greenhouse tomato fertilizer guidelines and growth optimization studies.* Technical Bulletin 2024-47, Washington, DC.

Wei, Z., Wu, Y., Yang, Y., and Shi, T. (2021). Application of trigonometric polynomial fitting method in simulating the spatial distribution of PM2.5 concentration in South Central China. *Atmosphere, 15*(1), 28.

Zhou, H., and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 67*(2), 301–

Article DOI: 10.52589/AJMSS-2RTB74X4
DOI URL: https://doi.org/10.52589/AJMSS-2RTB74X4