iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** XI **Month of publication:** November 2023

DOI: https://doi.org/10.22214/ijraset.2023.56636

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue XI Nov 2023- Available at www.ijraset.com

Utilizing Machine Learning within Artificial Intelligence to Enhance Dissolved Oxygen Estimation in the Mississippi River via Temperature-Driven Polynomial Regression

Akshay Dumbre[1], Dishant Koli[2], Pritee Vaivude[3], Shravani Dumbre[4]

Abstract: In this research, a comprehensive examination of Dissolved Oxygen (DO) levels in the Mississippi River is undertaken, employing a Polynomial Regression model drivenby temperature data for predictive estimation. Dissolved oxygen serves as a pivotal indicator of water quality and the health of aquatic ecosystems, making its accurate forecasting crucial for effective environmental monitoring and management. By utilizingtemperature as a primary predictor, this study seeks to advance our understanding of the intricate relationship between temper- ature and DO within the context of the Mississippi River. The research involves an extensive dataset comprising measurements of water temperature and dissolved oxygen collected over an extended duration. A Polynomial Regression model is employed to establish a mathematical link between temperature and DO, thus providing a predictive tool for estimating DO levels at specific locations along the river. The model's performance is subjected to a rigorous evaluation, involving the assessment of diverse statistical metrics and validation techniques. The research outcomes yield valuable insights into the dynamics of DO in the Mississippi River, emphasizing the pivotal role of temperature as a primary driver of DO fluctuations. This study introduces a practical and efficient method for the monitoring and prediction of DO levels, which can be instrumental in the preservation and sustainable management of this vital aquatic ecosystem. Moreover, it makes a meaningful contribution to the broader realm of water quality assessment and has the potential to inform policies and practices aimed at ensuring the environmental well- being of the Mississippi River.

Index Terms: Research, Comprehensive examination, Dis- solved Oxygen (DO), Mississippi River, Polynomial Regression model, Temperature data, Predictive estimation, Water quality, Aquatic ecosystems, Environmental monitoring, Management, Primary predictor, Relationship, Dataset, Measurements, Mathe- matical link, Performance evaluation, Statistical metrics, Valida- tion techniques, Dynamics, Fluctuations, Monitoring, Prediction, Preservation, Sustainable management, Aquatic ecosystem, Wa- ter quality assessment, Policies, Practices, Environmental well- being, Artificial Intelligence.

I. INTRODUCTION

The Mississippi River, one of the most significant river sys- tems in North America, plays a pivotal role in the ecological and environmental landscape of the region. Understanding and monitoring the key factors influencing water quality in this vast river system is essential for effective resource management and environmental stewardship. Among the various parameters that influence water quality, Dissolved Oxygen (DO) is a critical indicator, reflecting the river's capacity to support aquatic life and overall ecosystem health. The relationship between water temperature and dissolved oxygen levels in rivers has long been recognized as a fundamental aspect of aquatic ecology.

Temperature directly affects the solubility of oxygen in water, with warmer water generally holding less dissolved oxygen. This relationship is particularly crucial in river systems like the Mississippi, where variations in tempera- ture are influenced by seasonal and climatic changes. Ac- curate estimation of dissolved oxygen levels in such dynamic and ecologically significant water bodies is therefore a subjectof great importance. In recent years, the development of predictive models for estimating dissolved oxygen levels from temperature data has gained momentum. Polynomial regres- sion, a versatile and flexible modeling technique, has shown promise in capturing the complex, nonlinear relationships between Temperature and Dissolved Oxygen. This research endeavors to contribute to this area of study by developing a polynomial regression model tailored to the Mississippi River. Leveraging a comprehensive dataset collected over three years,this study aims to provide a robust tool for the estimationof dissolved oxygen levels based on temperature, offeringvaluable insights for ecological management and water qualityassessment in this iconic river system.

This paper outlinesthe methodology, data collection, and analysis techniques employed in developing the polynomial regression model. It also discusses the potential implications of the research for environmental conservation, aquatic ecosystem management, and the broader field of water quality assessment. With afocus on both the technical aspects of model development and the practical applications in ecological preservation, this studyseeks to enhance our understanding of the intricate interplay between temperature and dissolved oxygen in the Mississippi River.

II. LITERATURE REVIEW

Water quality and the factors influencing it are of paramountconcern in managing the health of aquatic ecosystems. The Mississippi River, a colossal river system, is no exception to this concern, as it serves as a vital conduit for both commerce and ecology. Among the many parameters used to evaluate water quality, Dissolved Oxygen (DO) stands out as a critical indicator of the river's ecological health. The extent to which temperature affects DO levels in river systems has been a subject of scientific inquiry for decades, with various machinelearning and neural network methodologies employed to modelDO content in surface waters, as noted by [1], [3], [4]. This literature review explores the background and key findings in the context of predicting DO levels from temperature data,with a focus on the Mississippi River.

Temperature-DO Relationship in Rivers: A Historical Per- spective. The relationship between temperature and DO levelsin rivers has long been recognized as an essential aspect of aquatic ecology. It is well-established that as water temperaturerises, the solubility of oxygen decreases, leading to lower DO concentrations. This negative correlation between temper- ature and DO is fundamental and serves as the foundationfor understanding the oxygen dynamics in rivers. What role does temperature play in changing dissolved oxygen levels, according to early research [2].

Modeling Approaches for Temperature-DO Relationships Numerous modeling approaches have been employed toquantify the temperature-DO relationship in river ecosystems.Traditional linear regression models have been used, andwhile they can capture the basic relationship, they oftenfall short in capturing the complex, nonlinear nature of thisinteraction. This limitation has led to the exploration of moresophisticated modeling techniques. Among these, polynomialregression models have shown great promise [9]. Thesemodels allow for the incorporation of higher-order termsand are better suited to capturing the nonlinear effects oftemperature on DO.

Polynomial Regression Models in Water Quality Studies Polynomial regression models have been applied effectively invarious water quality studies, offering flexibility in modelingcomplex relationships, which is particularly valuable inecosystems as dynamic as the Mississippi River [10], [11].

Challenges and Limitations in Temperature-DO Modeling While polynomial regression models offer advantages, it'simportant to acknowledge the challenges and limitations.One challenge is the potential for over-fitting when usinghigh-order polynomial terms. Careful model selection andvalidation techniques are crucial to ensure the robustnessand reliability of the model. Additionally, variations in localconditions and seasonal patterns in the Mississippi River mayintroduce complexities that require careful consideration inmodeling [5], [13].

Significance of the Present Study The development of a polynomial regression model to predict DO levels from tem- perature data in the Mississippi River can contribute to the ongoing efforts to monitor and manage water quality in this ecologically significant system. This research, conducted overa three-year period, aims to provide a robust and reliablepredictive tool for estimating DO levels, incorporating ad- vanced techniques like neural networks combined with fuzzy computing [12]. This contributes to the wider understanding of how temperature and dissolved oxygen interact within riverecosystems and may guide policy and management decisions aimed at protecting the environmental health of the MississippiRiver [6], [7], [8].

III. METHODOLOGY

A. *Data Collection*

The cornerstone of meaningful research hinges upon the acquisition and meticulous preparation of top-tier data. In our investigation, we conducted an exhaustive review of the waterquality database specific to the Mississippi River, accessi- ble through the U.S. Geological Survey (USGS). USGS is renowned for its extensive coverage of water quality metrics and graciously provided us with comprehensive records span- ning the years 2020 to 2022. Our research centers on the Mississippi River at Baton Rouge, LA, with a specific focuson the monitoring station denoted as USGS 07374000. This USGS water database stands as an invaluable asset, significantly enriching our comprehension of water quality, streamlining its effective management, and shouldering a pivotal role in the preservation of the nation's water resources.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue XI Nov 2023- Available at www.ijraset.com

It acts as an essential cornerstone that bolsters scientific research, informs the development of well-founded policies, and advocates for responsible water resource management across the expanse of the United States. The availability of such painstakingly detailed and high-quality data has been instrumental in easing our nuanced analysis, empowering us to delve into the intricate connections between temperature and Dissolved Oxygen within the realm of water quality analysis and forecasting.

Fig. 1.

Figure 1 shows the location of USGS monitoring station 07374000 Mississippi River at Baton Rouge, LA

B. *Significance and Precision of this Dataset*

The Mississippi River's water quality database functions as a detailed account of the river's state, presenting metic- ulously curated data that exposes shifting trends and the river's holistic well-being throughout the documented years. Given the Mississippi River's central role in U.S. hydrology, gaining a thorough understanding of its water quality holds a subtle yet indispensable significance. Consequently, this dataset emerges as a precious asset, delicately offering the potential for valuable insights and data to guide the efforts of researchers, policymakers, and environmental enthusiasts.

TABLE I

Date	Mean Turbidity(FNU)	Mean Gage Height(feet)	Mean Temperature (°C)	Mean Specific Conductance (µS/cm) at 25°C	Mean Salinity (parts per thousand)	PH Median	Mean. Discharge(cubic feet per second.)	Mean Dissolved Oxygen(milligrams per liter)
01/01/2020	94.8	27.56	7.9	346	0.2	7.8	666000	10.9
01/02/2020	97	27.78	8	336	0.2	7.8	672000	10.8
01/03/2020	89.3	28.39	8.3	322	0.2	7.8	689000	10.7
01/04/2020	109	28.81	8.6	311	0.2	7.8	703000	10.6
01/05/2020	96.7	28.9	8.9	309	0.2	7.8	705000	10.4

Table 1 shows the first 5 rows of our time series dataset.

Temperature: Water temperature is a critical parameter in the river systems of the United States, playing a central role in shaping the health and sustainability of these vital aquatic ecosystems. The temperature of river water is far from a passive variable; it exerts a profound influence on various facets of riverine environments, with its effects extending far beyond the simple reading on a thermometer. Understanding the importance of water temperature in the USA's river systems is pivotal for the effective management and preservation of these essential water bodies. One of the most significant conse-quences of water temperature in river systems is its impact on the levels of dissolved oxygen. This relationship is paramount to the health of aquatic life. As water temperature rises, its capacity to hold oxygen decreases, leading to lower dissolved oxygen concentrations. Dissolved oxygen is a lifeline for fish, invertebrates, and other aquatic organisms, serving as the oxygen source they depend on for respiration. A decrease in dissolved oxygen levels due to elevated water temperature can have detrimental effects on aquatic biodiversity and the overall health of river ecosystems.

Water temperature is a key indica- tor in the assessment of water quality in river systems. Rapid or sustained increases in temperature can signify pollutionsources, such as industrial discharges or urban development, which can lead to reduced dissolved oxygen levels. Monitoring water temperature is indispensable for the early detection of environmental stressors and the implementation of effective management and conservation strategies. Water temperature isundeniably one of the linchpins of the ecological intricaciesof river systems in the USA. Its effects on dissolved oxygen, seasonal dynamics, and its role in responding to climate change emphasize the importance of meticulous monitoring and understanding of this variable. To maintain the health and resilience of these crucial ecosystems, water temperature should be recognized as a cornerstone of river system man- agement and conservation.

Figure 2 shows the frequency of days from 2020 to 2022 with respect to the mean temperature ranges (in degree Cel- sius)

Dissolved Oxygen: Dissolved oxygen is a critical parame- ter in the river systems of the United States, holding profound significance for the health and sustainability of these vitalaquatic ecosystems.

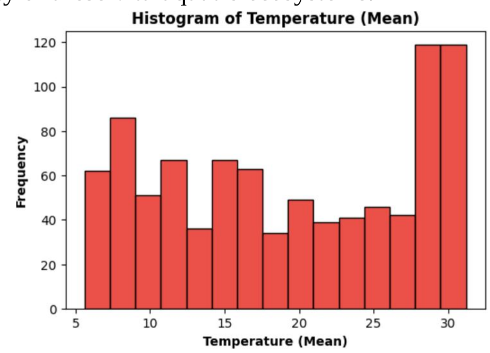

Fig. 2.

It refers to the amount of oxygen present in water, essential for the respiration and survival of aquatic organisms. Understanding the importance of dissolved oxygen in the USA's river systems is fundamental for the management and preservation of these invaluable water bodies. Adequate levels of dissolved oxygen are vital for sustaining a diverse array of aquatic life, from fish to invertebrates. It plays apivotal role in supporting healthy food chains and ecological processes, making it a cornerstone of river ecosystem health. Monitoring dissolved oxygen levels in river systems is es- sential for assessing water quality, as reduced concentrations can signal pollution or ecological stress. Ensuring that river systems maintain sufficient dissolved oxygen levels is not onlycrucial for the well-being of aquatic organisms but also for theoverall balance and sustainability of these ecosystems, which play a central role in the environmental and economic fabricof the United States.

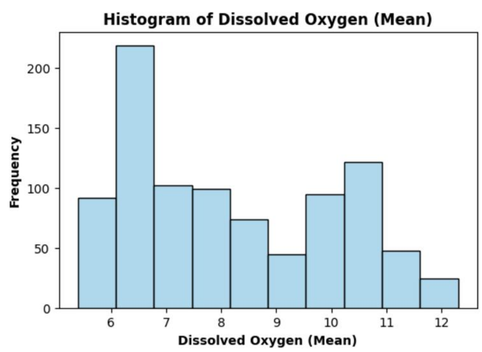

Fig. 3.

Figure 3 shows the frequency of days from 2020 to 2022 with respect to the mean Dissolved Oxygen (milligrams per liter)

C. *Methodology for Data Collection*

Methodology for collecting water quality data from the USGS website involved accessing the comprehensive records spanning 2020 to 2022. This data, primarily sourced fromthe Mississippi River at Baton Rouge, LA (USGS 07374000), served as the foundation for our study. The data collection pro- cess emphasized precision and reliability, while data integrity was ensured through initial checks to identify and addressmissing values.

D. Data Integrity and Preprocessing:

Ensuring the integrity of data stands as a paramount con- cern. Initial checks were carried out to detect any instancesof missing values, outliers, or irregularities. During the pre- processing phase, identified missing values within the dataset were thoughtfully managed through exclusion to maintain the consistency and dependability of our analyses.

E. Data Analysis

The process of translating raw data into actionable insights necessitates a structured and methodical approach to analysis. Our research employed a series of analytical techniques, with each step building upon the insights gained in the preceding phase.

F. Data Exploration and analysis

The dataset employed for this research was sourced from ”United States Geological Survey(USGS)” and primarily fo- cuses on the relationship between the mean dissolved oxygen levels and mean temperature values from the Mississippi River.

G. Preliminary Data Exploration

Upon initial review, the dataset consisted of multiple vari- ables, with particular emphasis placed on the average measure-ments of dissolved oxygen and temperature. The structure of the dataset was carefully examined, revealing a considerable volume of records. Subsequent scrutiny detected instances of missing data, which were systematically excluded from the dataset to ensure the integrity of the analysis. This refinement of the dataset was imperative for conducting a rigorous and dependable statistical assessment.

H. Statistical Analysis

Upon initial review, the dataset was found to contain amultitude of attributes, with investigative efforts particularly honed in on the mean values for dissolved oxygen and tem- perature. A thorough examination of the dataset's framework was conducted, revealing a substantial number of records. Thisin-depth evaluation also uncovered a series of missing entries, necessitating their removal to preserve the dataset's analytical validity. Such meticulous data curation was fundamental to ensuring the subsequent analysis would be both precise and credible.

The Correlation Heatmap (Fig. 4) clarified the connections among the different attributes. A significant finding was the strong correlation between the average dissolved oxygen and average temperature, affirming the choice to delve deeper intothis particular association.

Scatter plots are a powerful tool for illustrating the nuances of the relationship between pairs of numerical variables. The scatter plot detailing the relationship between the mean dis- solved oxygen and the mean temperature, in particular, sheds light on the nature of their correlation. By plotting each data point, it provides a compelling visualization that highlights

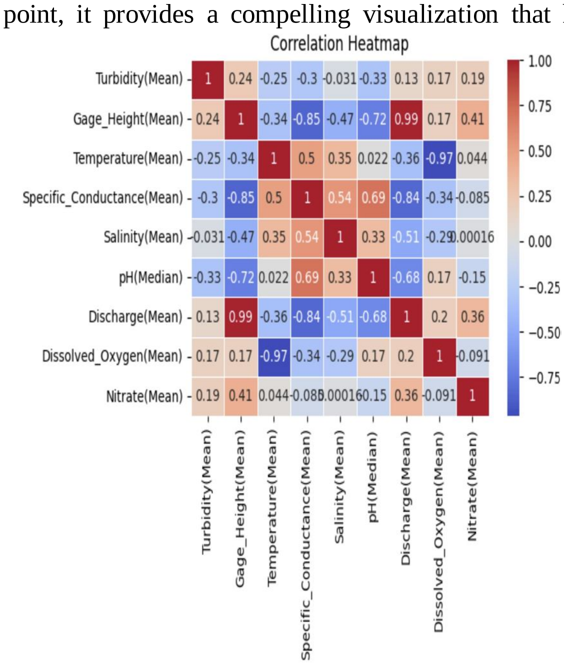

Fig. 4.

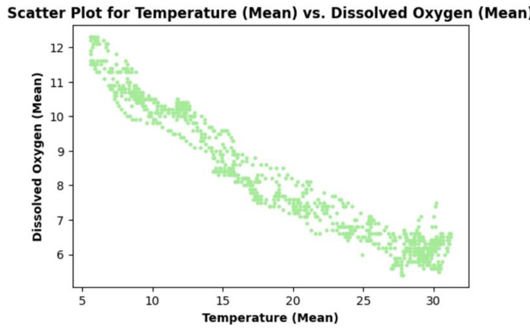

Fig. 5. Scatter Plot of mean of dissolved oxygen and mean of temperature

patterns or trends, such as the apparent direct correlation showcased in Fig. 5, offering a deeper insight into how these variables might influence one another.

I. *Visual Data Exploration*

Several visualizations were crafted to explore the data further:

1) *Scatter Plots*: Scatter plots are instrumental in visualiz- ing the distribution and relationship of data points acrosstwo numerical dimensions. The scatter plot comparing mean dissolved oxygen to mean temperature (Fig. 5)was particularly revealing, as it displayed a discernible pattern that suggested a potential relationship. This type of plot is crucial for identifying trends, outliers, andclusters within the data, which can lead to a deeperunderstanding of the factors at play.
2) *Histograms:* Histograms are a critical analytical in- strument for depicting the frequency distribution of individual variables. The histograms developed for the mean dissolved oxygen and mean temperature (shown inFig. 6 and Fig. 2, respectively) shed light on the data's distribution patterns within the dataset. Not only did theyoffer a clear picture of the central tendencies — the mostcommon values — but they also detailed the variability of the data, highlighting any potential asymmetry or peakedness through skewness and kurtosis in these key variables.

The visual clarity provided by these histograms is in- valuable. It facilitates a rapid and intuitive understanding of where the majority of data points lie, the tails ofthe distribution, and any deviations from the expected normal distribution. By revealing outliers and anoma- lies, histograms guide researchers in making informed decisions about which statistical methods are appropriate for further analysis. They are a foundational step in data exploration, offering a preliminary check on the assump-tions underlying parametric statistical tests. Moreover, these visual aids can direct the researcher's attention to underlying phenomena that warrant more sophisticated modeling techniques, and they can also serve as an aid incommunicating findings to a broader audience, making complex data more accessible.

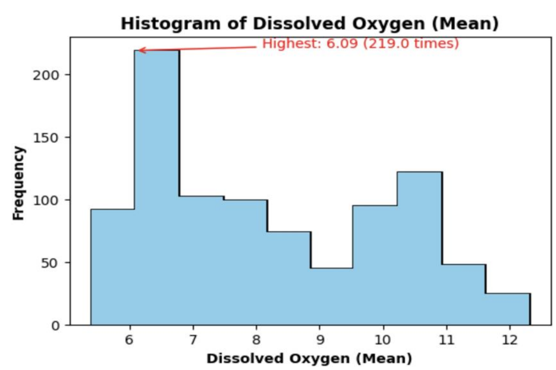

Fig. 6. Histogram of mean of dissolved oxygen

J. *Polynomial Regression Analysis*

In an effort to explore the intricacies of the relationship between the average dissolved oxygen and average tem- perature, a polynomial regression model was utilized. This approach entailed the generation of 3rd-degree polynomial features to better capture any nonlinear asso- ciations between the variables. The dataset was divided into training and test sets to facilitate an objective assessment of the model's predictive performance.

IV. RESULTS

We present the following polynomial regression equation that was developed to predict dissolved oxygen levels based on temperature.

$$y = 1.57437389 \times 10^{-4}x^3$$
$$-2.84628947 \times 10^{-3}x^2$$
$$-0.301763021 \times x$$
$$+13.5032488$$

The analysis yielded a robust and statistically significantmodel that can effectively capture the non-linear rela- tionship between these two variables in the context of theMississippi River. The polynomial regression equation, as shown above, provides a powerful tool for accu-rately estimating dissolved oxygen levels, which is of paramount importance for understanding and managing water quality in the region.

Once the model was trained, the evaluation yielded the following metrics:

- Mean Squared Error (MSE): 0.14
- Root Mean Squared Error (RMSE): 0.38
- Mean Absolute Error (MAE): 0.31
- R-squared (R^2) Score: 0.96

These results indicate a model that performs well,particularly highlighted by the high R^2 score, whichimplies that a substantial proportion of the variance inthe dependent variable is accounted for by the model. The model's predictions were visualized against theactual data points for further analysis:

Each figure serves to illustrate a different aspect of the model's fit and predictive capabilities. The test dataset figure assesses how well the model can predict new data,the training dataset figure shows the model's fit to the data on which it was trained, and the entire dataset figuredemonstrates the comprehensive trend captured by the model.

- Test dataset (Fig. 7)

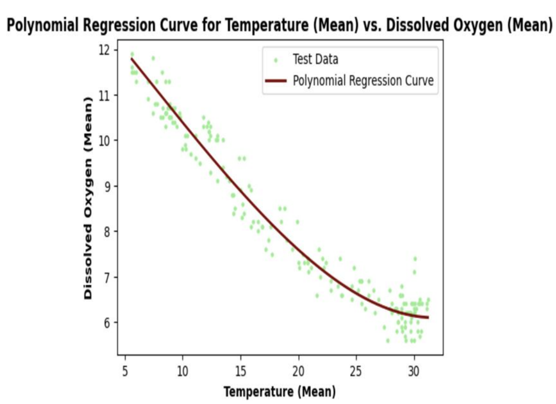

Fig. 7. Polynomial Regression Fit for Test Dataset

- Train dataset (Fig. 8)

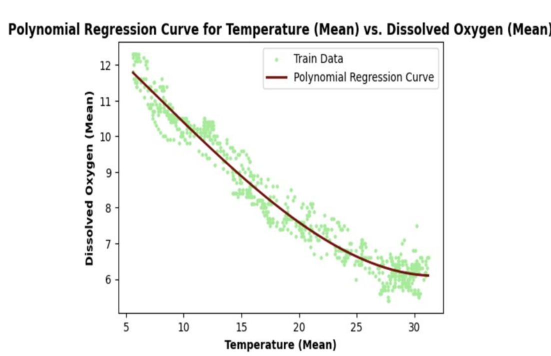

Fig. 8. Polynomial Regression Fit for Training Dataset

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue XI Nov 2023- Available at www.ijraset.com

- Entire dataset (Fig. 9)

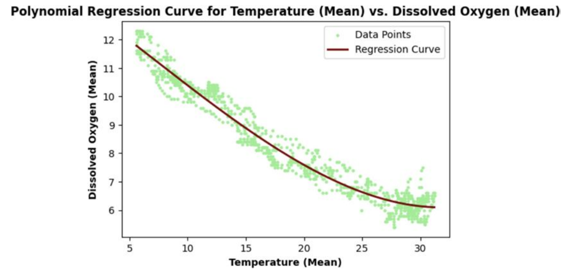

Fig. 9. Polynomial Regression Fit for Entire Dataset

A. Residual Analysis

Residual plots serve as a diagnostic tool to verify the assumptions of a regression model. For the polynomial regression model developed in this study, Figures 10 and11 display the residual plots for both the training andtest datasets. These plots are critical for assessing the model's accuracy in predicting dissolved oxygen levels based on water temperature. A key feature of an effective model is the random scatter of residuals, which suggeststhat the model captures the underlying pattern without systematic errors. The uniform distribution of residuals around the horizontal axis in our plots suggests that the variances of the error terms are constant (homoscedastic-ity) and that the residuals are independent of each other. This indicates that the model accounts well for the non- linear relationship between temperature and dissolved oxygen without overfitting to the training data or failing to capture the variance in the test data.

Moreover, the absence of distinct patterns in the resid- ual plots implies that the polynomial regression model does not suffer from omitted variable bias, where the exclusion of a relevant variable would otherwise lead to systematic errors in prediction. The plots do not exhibit any clear signs of heteroscedasticity or autocorrelation, further supporting the model's appropriateness for the data at hand. The homogeneity in variance across the range of predicted values is particularly important when dealing with environmental data, which often contain outliers or high-leverage points due to natural variability. The examination of residual plots also offers insights into the potential for model improvement. While thecurrent model performs well, any outliers or patterns thatcould emerge in residual plots with larger datasets couldindicate the need for additional explanatory variablesor transformation of the data. Nonetheless, the current residual plots corroborate the polynomial regression model's capability to generalize beyond the observed data, making it a valuable tool for predicting dissolved oxygen levels in the Mississippi River. This analytical strength underpins the model's utility for environmental scientists and policymakers who rely on accurate wa- ter quality predictions to make informed decisions for ecosystem management and conservation efforts.

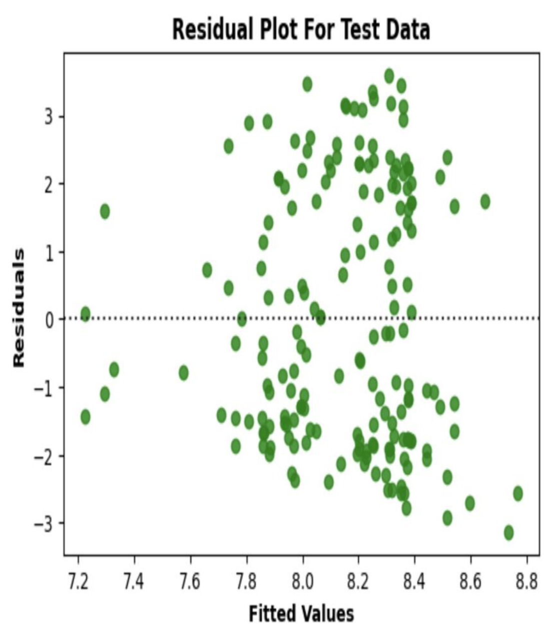

Fig. 10. Residual Plot Test Dataset

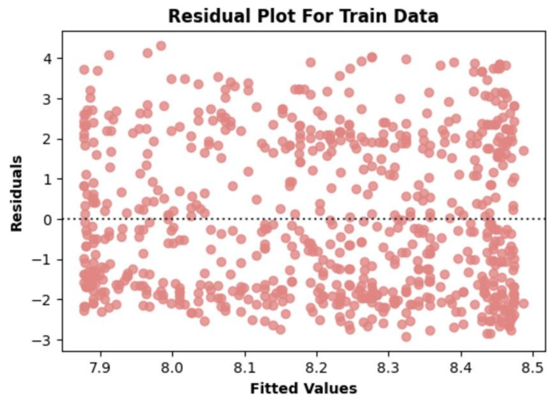

Fig. 11. Residual Plot Train Dataset

B. Distribution Plot

The distribution plot for mean dissolved oxygen levels, as seen in the Mississippi River dataset, is a pivotaltool in exploratory data analysis. This visual representa- tion offers immediate insight into the central tendency, variability, and the form of the data distribution. By assessing the plot, researchers can determine the degree of skewness and kurtosis, which are measures of data asymmetry and the peakedness of the distribution, re- spectively. These metrics are crucial because they inform the suitability of various data analysis techniques and guide the choice of predictive models.

For example, a distribution with high skewness maysuggest the need for data transformation to meet the assumptions of parametric tests and models, which typi-cally assume data normality. Kurtosis, on the other hand,provides information on the data's tendency to produce outliers, which can have a disproportionate impact on certain statistical analyses.

In the context of this study, the distribution plot of dissolved oxygen levels is more than a preliminary step; it's a foundational aspect that influences the entire modeling process. If the data exhibits a normal distri- bution, this supports the use of polynomial regression models since these models assume that the residualsare normally distributed. On the contrary, if significant deviations from normality are present, it may necessitatea reevaluation of the model choice or the adoption of data transformation techniques.

Moreover, the distribution plot can reveal outliers which, depending on their nature and the goals of the study, might be candidates for further investigation or ex- clusion from the dataset to avoid skewing the results. Outliers can be particularly informative or problematic in environmental data; they could represent rare but ecologically significant events or data errors.

In the analysis of the Mississippi River data, the distribu- tion plot contributes to a comprehensive understanding of the environmental factors at play. It ensures that the subsequent polynomial regression model is constructed on a solid foundation of data that is well-understood and properly conditioned, thereby enhancing the credibility of the model's predictions and the conclusions drawn from it. This is indispensable for researchers and poli- cymakers who rely on accurate predictions of dissolved oxygen levels for effective water quality management and the conservation of aquatic ecosystems.

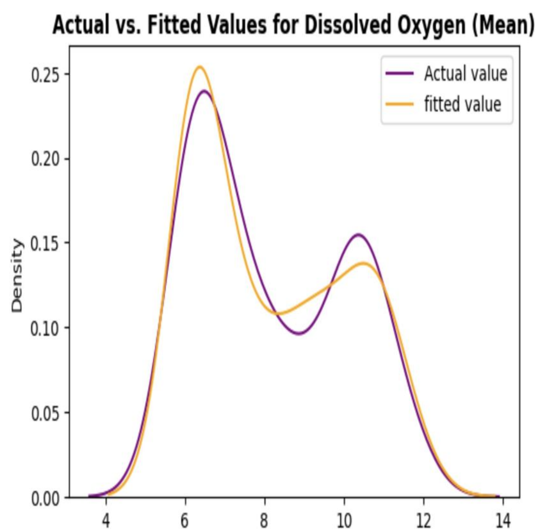

Fig. 12. Distribution Plot

V. DISCUSSION

In our research study, we have implemented the ma- chine learning technique of polynomial regression, while concurrently recognizing the substantial potential for embracing a diverse spectrum of artificial intelligence methods in the domain of water quality analysis. Specifi- cally, artificial neural networks, encompassing advanced deep learning models like Convolutional Neural Net- works (CNNs) and Recurrent Neural Networks (RNNs),have emerged as formidable assets for the tasks of pattern recognition and predictive modeling within the realm of water quality data. CNNs, renowned for their prowess in image analysis and spatial data interpretation, hold the potential to offer invaluable capabilities when assessing water quality, particularly in visually complex scenarios. Their application extends to the analysis of water quality-related images, such as those derived from satellite or drone sources, facilitating the identification of environmental factors, pollution sources, and temporal dynamics within water bodies. In parallel, RNNs exhibitexcellence in processing sequential data, rendering themexceptionally suitable for time series analysis in the context of water quality monitoring. These networksexhibit proficiency in capturing temporal dependencies and variations in water quality parameters, thereby em- powering the detection of prolonged trends and the anticipation of future conditions.

The utilization of artificial neural networks bears a mul- titude of advantages, principally anchored in their apti- tude for discerning intricate relationships from extensiveand multifaceted datasets. This adaptability equips themto harmonize with the distinct nuances present in water quality data. By leveraging the capabilities inherent in deep learning models, water quality predictions stand poised to benefit from elevated precision and enhanced generalization. Moreover, the inherent potential for con- tinuous monitoring conferred by these models equips them for the early identification of anomalies, facilitatinga proactive response conducive to the preservation of water resources and the safeguarding of public health.

VI. CONCLUSION

Our research endeavors to shed light on the relationship between temperature and dissolved oxygen levels in the Mississippi River. We collected and analyzed a substantial dataset, providing a detailed account of the river's water quality dynamics. Through the development and application of a polynomial regression model, we have uncovered valuable insights into the predictive ca- pabilities of temperature as a key variable in estimating dissolved oxygen levels. Our findings emphasize the sig- nificance of temperature as a primary driver of dissolvedoxygen fluctuations in the Mississippi River. By employ- ing this polynomial regression model, we have enhanced our understanding of the intricate interactions within this dynamic ecosystem. This research not only contributes to the body of knowledge in water quality assessment butalso offers a practical and efficient tool for monitoring and predicting dissolved oxygen levels, essential for the preservation and responsible management of this crucialaquatic environment. Furthermore, the implications of this research extend beyond the Mississippi River. The methodologies employed and the insights gained from our study have the potential to inform broader water quality assessment practices, benefiting environmental management and conservation efforts in river systems across the United States. In a time when sustainable water resource management is of utmost importance, ourresearch offers a valuable contribution towards achievingthis goal.

VII. ACKNOWLEDGMENT

The United States Geological Survey (USGS) played a pivotal role as a priceless asset for our research paper.It provided water quality dataset for the Mississippi River at Baton Rouge, LA (USGS 07374000) coveringa span of three years. This extensive dataset formed the cornerstone of our efforts to construct a polynomial regression model aimed at predicting dissolved oxygen levels based on temperature, significantly elevating the precision and reliability of our analysis. The accessibilityof water data from USGS was indispensable in advanc- ing our comprehension of the intricate dynamics within the Mississippi River and, by extension, the broader fieldof water quality assessment in river systems throughout the United States.

REFERENCES

[1] Sˇiljic´ Tomic´, D. Antanasijevic´, M. Ristic´, A. Peric´-Grujic´, V. Pocajt, ”A linear and non-linear polynomial neural network modeling of dissolved oxygen content in surface water: Inter-and extrapolation performance with inputs' significance analysis”,Elsevier, 2017, pp. 9.

[2] W. Li, H. Fang, G. Qin, X. Tan, Z. Huang, F. Zeng, H. Du, S. Li, ”Concentration estimation of dissolved oxygen in Pearl River Basin using input variable selection and machine learning techniques”, Elsevier, 2020, pp. 12.

[3] P. Mart´ı, J. Shiri, M. Duran-Ros, G. Arbat, F. Ram´ırez de Cartagena, J. Puig-Bargue´s, ”Artificial neural networks vs. Gene Expression Programming for estimating outlet dissolved oxygen in micro-irrigation sand filters fed with effluents”, Elsevier, 2013, pp. 10.

[4] A. Csa´bra´gia, S. Molna´ra, P. Tanosa, J. Kova´csb, M. Molna´rc, Szabo´a, I. G. Hatvanid, ”Estimation of dissolved oxygen in riverine ecosystems: Comparison of differently optimized neural networks”, Ecological Engineering, 2019, pp. 17.

[5] H. Wang, M. Hondzo, C. Xu, V. Poole, A. Spacie, ”Dissolved oxygen dynamics of streams draining an urbanized and an agri- cultural catchment”, Elsevier, 2002, pp. 17.

[6] M. Ay, O¨ . Kis¸i, ”Modeling of Dissolved Oxygen Concentration Using Different Neural Network Techniques in Foundation Creek, El Paso County, Colorado”, American Society of Civil Engineers, 2012, pp. 9

[7] J. Li, N. S. Xu, W. W. Su, ”Online estimation of stirred-tank microalgal photobioreactor cultures based on dissolved oxygen measurement”, Elsevier, 2002, pp. 15

[8] K. Galajit, S. Duangpummet, P. Dangsakul, R. Keinprasit, P. Dillon, J. Intha, K. Rungprateepthaworn, J. Karnjana, ”Prediction of Dissolved Oxygen Concentration for Shrimp Farming Using Quadratic Regression and Artificial Neural Network”, Not pro- vided, 2006, pp. 6.

[9] B. Keshtegar, S. Heddam, H. L. Krauss, ”Modeling daily dis- solved oxygen concentration using modified response surface method and artificial neural network: a comparative study”, The Natural Computing Applications Forum, 2017-2018, pp. 12.

[10] S. Nacar, B. Mete, A. Bayram, ”Estimation of dissolved oxygen concentration using conventional regression analysis, multivari- ate adaptive regression splines, and TreeNet techniques”, KSCE Journal of Civil Engineering, 2017, pp. 9.

[11] B. Keshtegar, S. Heddam, H. Hosseinabadi, ”The employment of polynomial chaos expansion approach for modeling dissolved oxygen concentration in river”, Environmental Earth Sciences, 2019, pp. 18.

[12] M. Ay, O¨ . Kis¸i, ”Estimation of Dissolved Oxygen by using Neural Networks and Neuro Fuzzy Computing Techniques”, KSCE Journal of Civil Engineering, 2017, pp. 9.

[13] B. Keshtegar, S. Heddam, H. Hosseinabadi, ”The employment of polynomial chaos expansion approach for modeling dissolved oxygen concentration in river”, Environmental Earth Sciences, 2019, pp. 18.