NEURAL NETWORKS VS. BIOWIN AS A MECHANISTIC TOOL: A STUDY ON PREDICTIVE ACCURACY OF AMMONIA LEVELS IN WASTEWATER TREATMENT SYSTEMS *

Fatima-zahra Ezahraoui, Ghassem Tofighi, Rania Ahmed Hamza
Toronto Metropolitan University
Toronto
{fezzahraoui, gtofighi, rhamza}@torontomu.ca

ABSTRACT

This study presents a novel approach to predicting effluent quality in wastewater treatment plants through Machine Learning (ML) models, utilizing a large dataset generated from the BioWin 6.2 simulation tool. The dataset encompasses both influent and effluent characteristics, offering a comprehensive perspective on treatment dynamics. We compared the performance of three different models: Linear Regression, Simple Recurrent Neural Network (RNN), and Long Short-Term Memory (LSTM) networks. Our analysis focused on evaluating each model's predictive accuracy and reliability with regards to ammonia levels prediction. The results demonstrate that LSTM outperformed the others in terms of prediction accuracy and consistency. This work not only underscores the potential of integrating mechanistic models with ML for enhanced predictive accuracy, cost-effectiveness, and risk reduction in model training but also suggests scalable and adaptable solutions for real-time optimization and enhanced environmental impact.

Keywords Wastewater treatment · LSTM · RNN · Machine learning · Biowin modeling

1 Introduction

Wastewater management is a vital component of public health and environmental protection, as it eliminates water pollution, pathogens, excess nutrients, and toxic chemicals. Globally, over 80% of sewage is released either in untreated or inadequately treated forms, hence the urgent need for an improvement in sewage treatment methods that could positively impact human health and environmental conservation. One of the pressing global issues is the increasing demand for potable water, driven primarily by population growth, climate change, elevating living standards, and the deterioration of water quality. The advantages of recycling domestic wastewater are dual-pronged: firstly, it prevents the release of raw sewage into nature, which contaminates water bodies; and secondly, it helps reduce the need for freshwater, especially in industry and agriculture. The presence of high nutrient levels in the effluent can harm water quality and the environment. High pollutant levels, namely nitrogen, can cause algae blooms, which block sunlight from underwater ecosystems, leading to ecological damage. These blooms may also foster harmful algae and bacteria growth, decreasing oxygen levels in the water, and ending aquatic life. For agricultural use, excess nitrates can leach into the soil, contaminating groundwater. In wastewater treatment plants (WWTPs), high nitrate levels can lead to 'rising sludge,' where nitrogen gas accumulates and causes sludge to escape with the effluent. In the WWTPs, nitrogen removal involves nitrification, which converts ammonia to nitrate, and denitrification, which converts nitrate to nitrogen gas under low-oxygen conditions. Predicting ammonia and nitrate levels could help optimize oxygen supply to the reactors, maintaining the treatment process's efficiency [1][2][3]. Amidst this backdrop, the integration of mathematical modeling into wastewater treatment processes presents an innovative avenue for advancing the field. Mechanistic and data-driven modeling have been employed to detect operational abnormalities in WWTPs and help manage effluent quality. Traditional mechanistic models rely on establishing fundamental physical, biological, and

* *Citation*: **Ezzahraoui, F.-z., Tofighi, G., Hamza, R. A. (2023). Neural Networks vs. Biowin as a Mechanistic Tool: A Study on Predictive Accuracy of ammonia levels in Wastewater treatment Systems. Toronto Metropolitan University.**

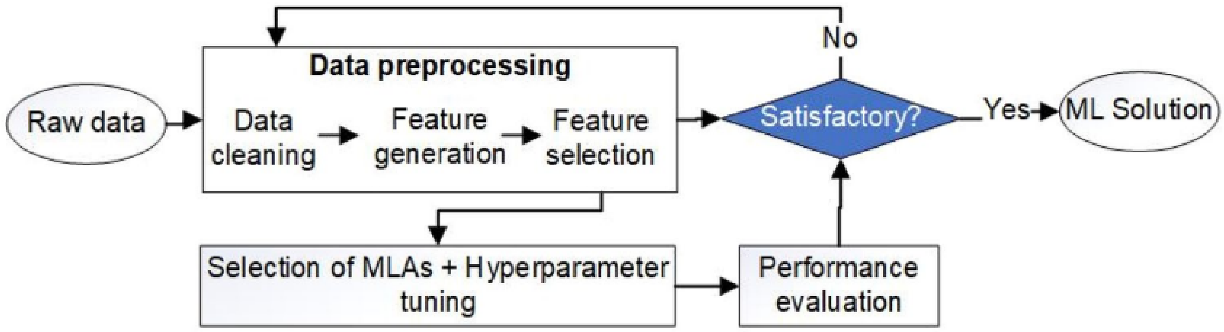

Figure 1: Typical ML model workflow [1]

chemical relationships that exist between the system components, which could be linear, non-linear, feedback loops, differential equations, etc. While these models are instrumental in understanding and predicting system behaviors, they often require deep knowledge of system dynamics and can be limited by the complexity and variability of wastewater treatment processes. In contrast, Machine Learning Algorithms (MLAs) offer a flexible and data-driven approach, capable of capturing complex patterns without the need for explicit system modeling. This study examines the synergy between mechanistic and ML approaches, exploring their combination potential to lead to more accurate, cost-effective, and efficient wastewater treatment solutions. Machine Learning Algorithms (MLAs) like Recurrent Neural Networks (RNN) are highly effective in predicting the behavior of wastewater treatment plants, thanks to their unique capability to integrate the concept of "yesterday, today, and tomorrow" which account for the influence the past and present conditions on the future effluent quality. This temporal understanding enables these algorithms to make more accurate and reliable predictions about the treatment plant's performance, ensuring better management and optimization of wastewater treatment processes.

The potential of ML, particularly in the form of Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) networks, lies in their ability to understand temporal sequences, effectively considering past and present data to predict future outcomes. This feature is particularly relevant in wastewater treatment, where the influence of past and present conditions plays a crucial role in determining effluent quality. Linear Regression, a fundamental statistical technique for predictive analysis, serves as the baseline model in this study. Its simplicity and interpretability make it a common starting point in predictive modeling. However, its effectiveness is often limited in complex, non-linear systems like wastewater treatment plants. On the other hand, Simple RNNs, a form of neural networks specifically designed for handling sequential data, provide a more nuanced approach by considering the temporal dependencies in the data. They are particularly adept at modeling time-series data, making them suitable for predicting dynamic changes in wastewater quality. LSTM networks, an advanced form of RNNs, further enhance this capability by addressing the limitations of traditional RNNs, such as the vanishing gradient problem, thereby allowing the model to learn from long-term dependencies more effectively.

The novelty of this study lies in its comprehensive approach – comparing the efficacy of these diverse modeling techniques using a dataset generated from the BioWin 6.2 simulation tool. This sets the stage for future research to explore hybrid models that blend mechanistic understanding with ML's data-driven insights, holding the promise of enabling real-time optimization and significantly reducing environmental impacts.

ML models, particularly simple RNN and LSTM networks, have shown increasing promise. According to recent studies, LSTM models have been increasingly applied in predicting effluent quality and operational parameters, showcasing significant improvements in prediction accuracy compared to traditional models.This trend highlights a growing recognition of the potential of advanced ML techniques in environmental engineering, marking a shift towards more data-driven, intelligent wastewater treatment solutions [1] [2].

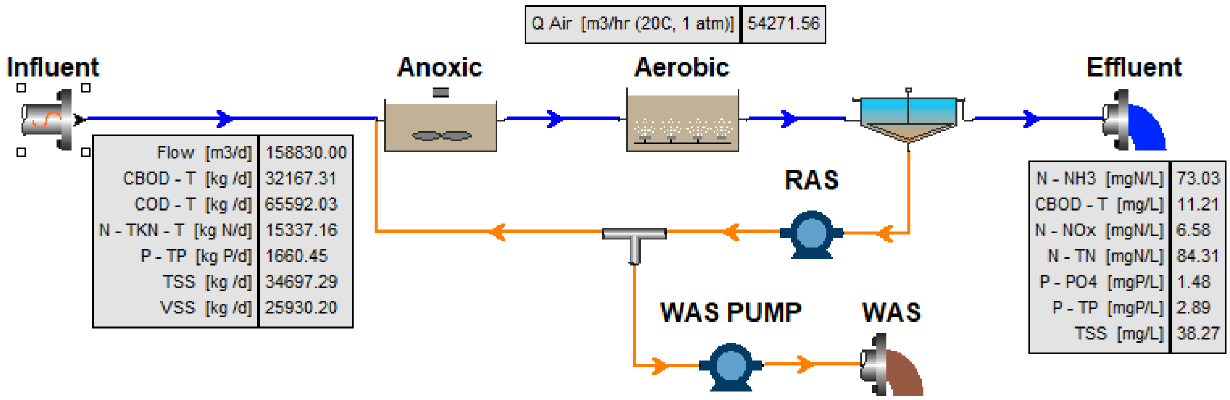

Figure 2: WWTP Biowin setup used to simulate the water quality dataset

2 Methods

2.1 Data preparation

A dataset of more than 37,000 data points has been generated using typical influent data as described in table 2.1. The Biowin setup used to generate the effluent data is a 100% biological activated sludge wastewater treatment plant, with a Dissolved Oxygen level of 2mg/L in the aeration tank, a liquid temperature of 20 °C, and a sludge recirculation rate of 100 MLD (figure 2.1).

Parameter	Influent							
	Flow rate	COD	BOD	TKN	P	NH3	TSS	NO2
Average	147.5	425.6	208.7	100.1	8.3	66.1	220.7	6.5
STD	13.1	109.5	49.5	20.4	2.1	13.5	50.1	0.3
Min	125	250	122.6	80.2	3.3	39.7	312.1	0.4
Max	170	600	294.2	139.8	13.4	92.3	128.9	0.7
	Effluent							
Average	144.1	52.1	9.3	44.6	5.9	40.3	30.3	0.3
STD	13.1	6.8	1.5	10.5	1.2	10.1	4.9	0.4
Min	121.6	34.6	5.9	17.4	2.7	42.6	21.4	1.7
Max	186.5	74.3	14.6	72.7	9.4	101.1	46.4	2.1

Table 1: Characterization of the dataset used for the LSTM model (Flowrates in MLD and concentrations in mg/L)

2.1.1 Model features engineering

Models such as LSTM are resource intensive, therefore, in the process of developing our predictive model for ammonia levels, careful consideration was given to feature selection—an important aspect of feature engineering to keep the features that have significant impact on effluent ammonia, namely Chemical Oxygen Demand (COD), Biochemical Oxygen Demand (BOD), Total Kjeldahl Nitrogen (TKN), Phosphorus (P), and the influent flow rate (Q_{inf}). . The attributes we selected for estimation were exclusively numerical in nature, comprising influent quality and flow rate measures. Parameters such as Total Suspended Solids (TSS) and pH were excluded from the model. TSS was dropped because it showed a high correlation with both COD and BOD, suggesting that its predictive information was already captured by these other attributes. The pH level was omitted due to its negligible variation across the dataset, which translated into a low impact on ammonia prediction in this context.

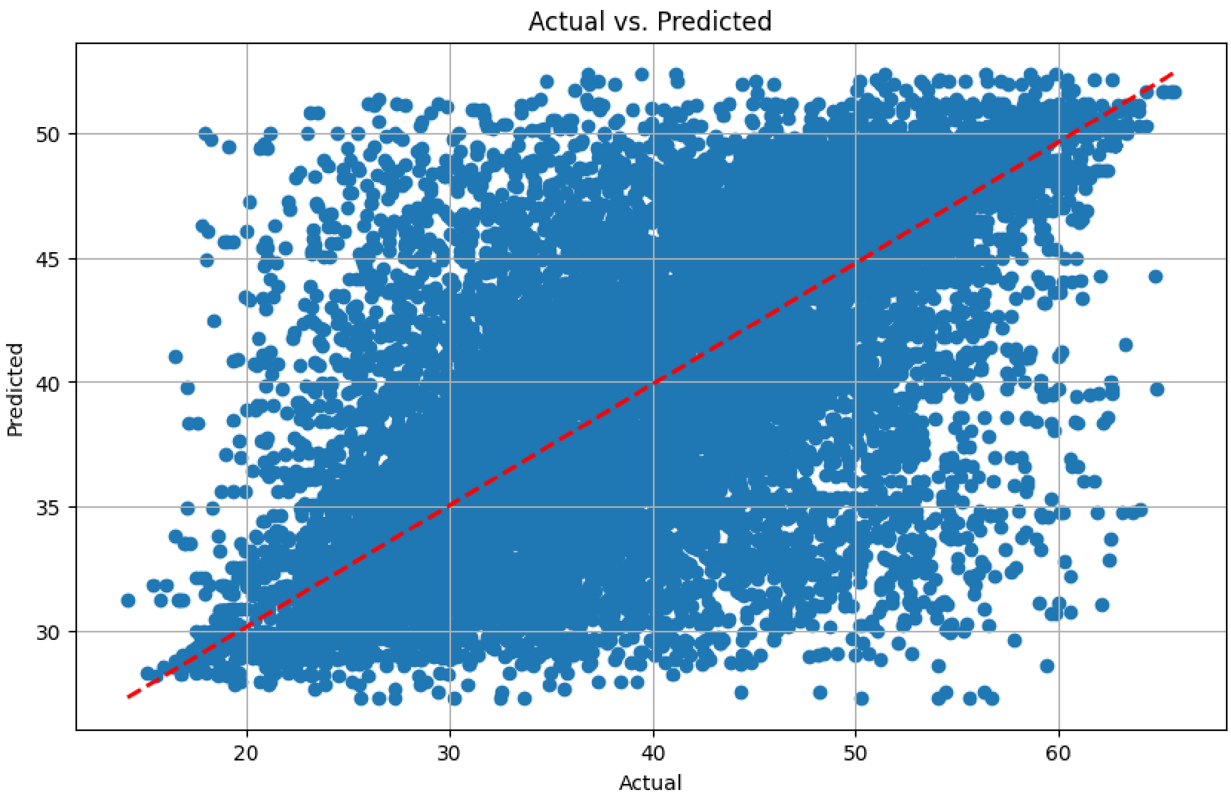

Figure 3: Effluent ammonia prediction results using Linear Regression

The approach to feature selection aimed at simplifying the model to improve computational efficiency and avoid data over-fitting, while retaining the most influential variables for ammonia estimation. This set of features allowed our models to focus on the key factors affecting ammonia levels.

2.1.2 Performance metric

In this work focusing on the prediction of ammonia levels in the effluent, we opted for Root Mean Squared Error (RMSE) and the Coefficient of Determination (R-squared, R^2) as our primary evaluative metric. RMSE was deemed crucial for its capacity to emphasize relatively large prediction errors, a key consideration in environmental applications where accurate ammonia prediction is important for compliance with regulatory standards.

On the other hand, we employed R^2 to offer a normalized indication of the model's predictive strength, and indicate how well our model's predictions correlated with actual Biowin calculated ammonia concentrations.

3 Results and discussion

The model was developed in Python using the Keras library [3] with a TensorFlow backend [4]. The training process took place on Google Colab utilizing an Intel(R) Xeon(R) CPU.

3.1 Linear regression

The scatter plot in figure 3 indicates a linear regression analysis of actual vs. predicted ammonia concentrations. There is a dense clustering of points along the diagonal, with a red dashed line that indicates the ideal 1:1 prediction line, which suggests that the regression model has a reasonable fit for the central range of values. However, the dispersion of points increases as the actual ammonia values increase, inferring that the model's accuracy relatively diminishes at higher ammonia levels. This could indicate that the model struggles with higher concentrations, possibly due to variability in the data not captured by the model, or other factors not accounted for in the regression.

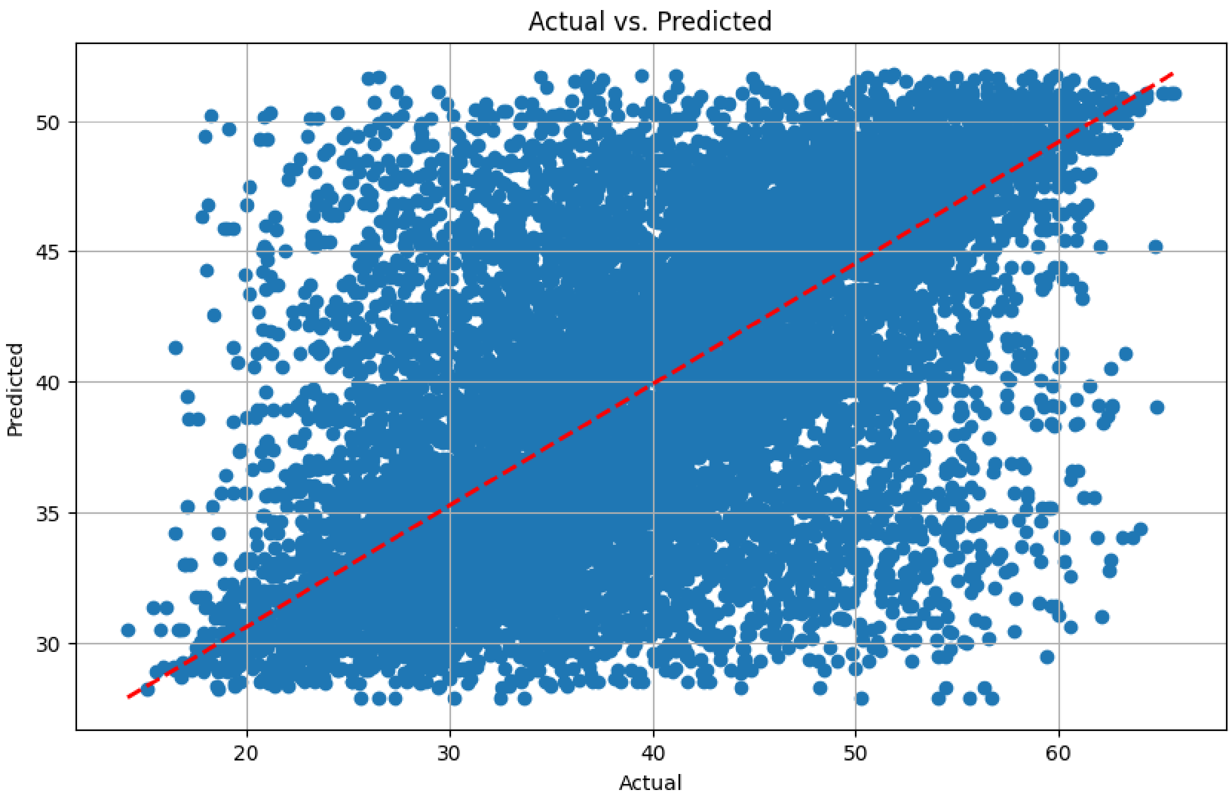

Figure 4: Effluent ammonia prediction results using Simple RNN

3.2 Simple RNN

A first round of RNN training and testing gave a poorly performing model. Hence the necessity to tune the model hyperparameters. Keras was used to experiment different configurations of the model To find the best Recurrent Neural Network (RNN) model using Keras, you can employ hyperparameter tuning to experiment with different configurations of the model. Table **??** describes the final model parameters

3.3 LSTM

Figure 6 represents a scatter plot comparing the actual values to the predicted values from the LSTM model. The density of points along the line of perfect prediction suggests that the model has a degree of predictive accuracy, as many predictions are close to the actual values. However, there is a spread of points that deviate from the line, indicating errors in the predictions. These errors appear to be larger for higher actual values, similarly to LR and LSTM mdoels, but slightly better performance as indicated in table 2. There is no distinct pattern indicating systematic bias; the points are dispersed around the line fairly evenly across the range of values. The LSTM model has a higher predictive accuracy and a better fit for the data, with lower mean squared error, and R-squared value than the other two models.

Metric	Linear Regression	RNN	LSTM
Mean Squared Error (MSE)	65.828	64.929	62.189
Mean Absolute Error (MAE)	6.414	6.291	6.190
Root Mean Squared Error (RMSE)	8.113	8.058	7.886
R-squared	0.356	0.365	0.392

Table 2: Comparison of Error Metrics between Linear Regression, RNN, and LSTM Models

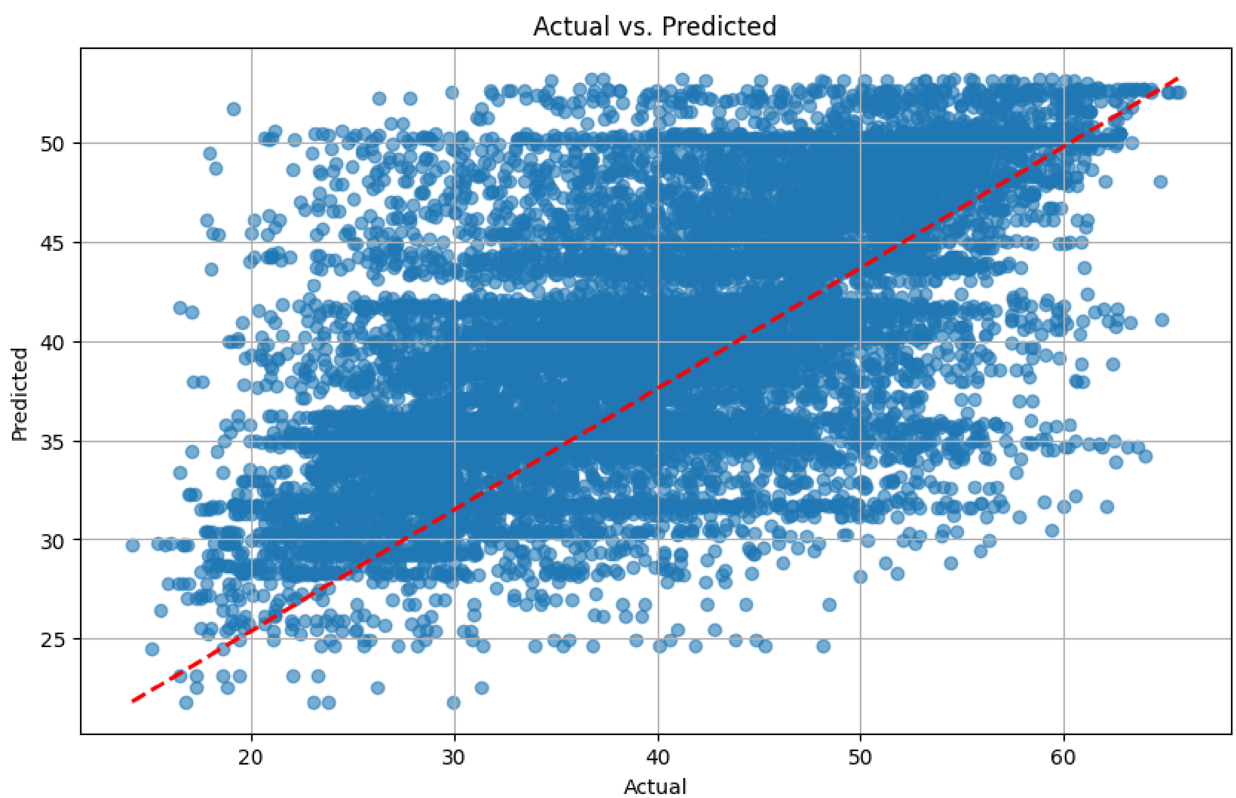

Figure 5: Effluent ammonia prediction results using LSTM

4 Deployment into the Cloud

A cloud-based infrastructure was designed and successfully implemented on the AWS platform, leveraging a multi-component architecture to support a web application. The key components include an Elastic Kubernetes Service (EKS) for orchestration, Elastic Container Registry (ECR) for Docker image storage, and EC2 instances to serve as worker nodes where the Flask app is running the AS LSTM model prediction.

Infrastructure Design The infrastructure is structured within a Virtual Private Cloud (VPC) to provide a secure environment. The VPC is divided into two subnets across different Availability Zones (AZs) for high availability:

Public Subnet (AZ A) Hosts a web server EC2 instance with an associated Elastic IP for stable public access. A NAT Gateway is present to enable outbound traffic from private resources.

Private Subnet (AZ B) Contains worker EC2 nodes managed by the EKS control plane, which facilitates Kubernetes orchestration across the cluster. The domain wastewaterai.com was configured with the necessary DNS records to point to the EC2 web server's Elastic IP.

Kubernetes Cluster An EKS cluster was set up, comprising multiple EC2 worker nodes for deploying containerized applications. The cluster management and orchestration are handled by the EKS control plane.

Container Registry The ECR stores two Docker images:

- A Flask backend application, which is a TensorFlow model-serving API.
- A frontend image to serve the user interface components.

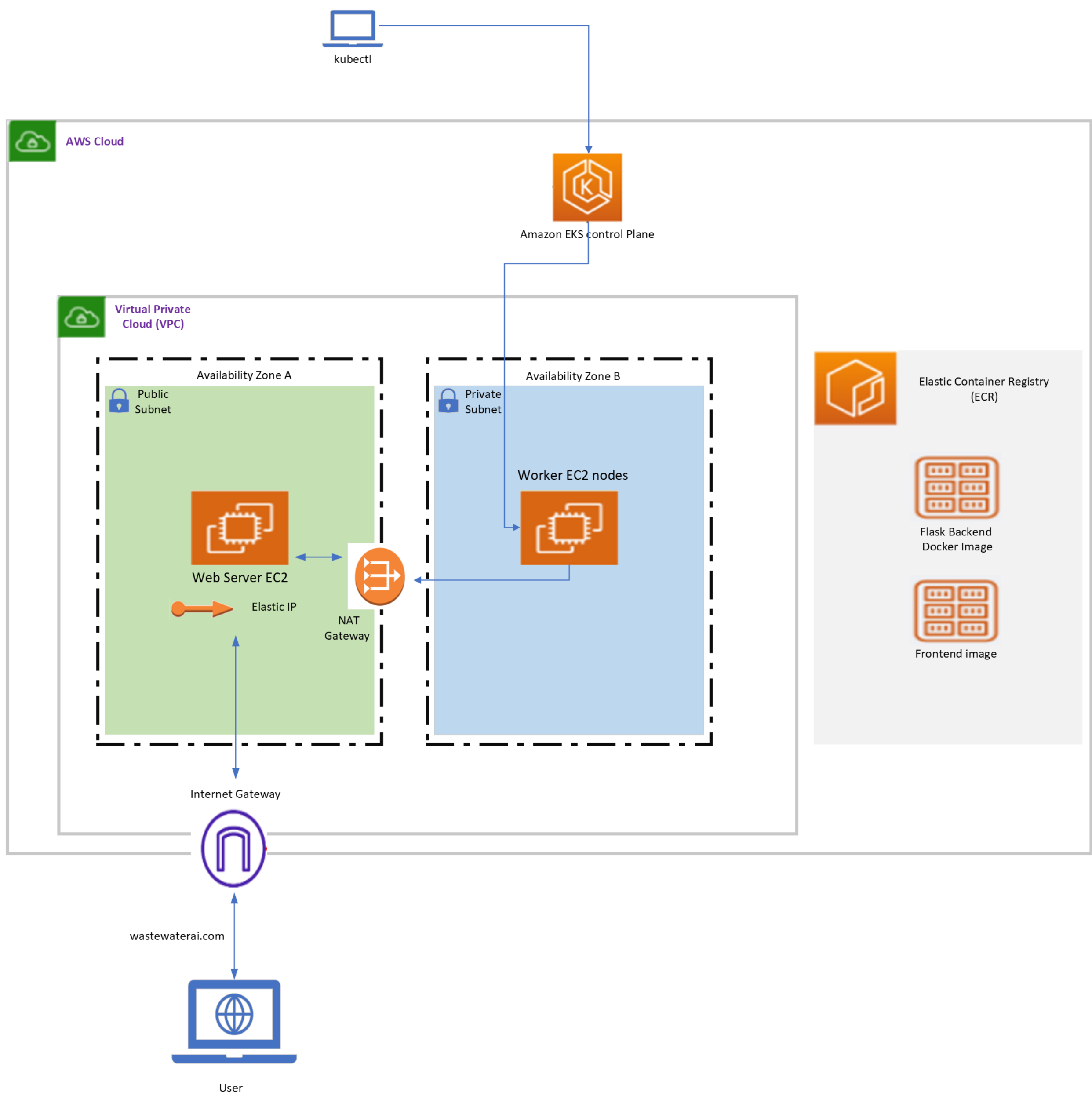
kubectl
AWS Cloud
Amazon EKS control Plane
Virtual Private
Cloud (VPC)
Availability Zone A
Public
Subnet
Web Server EC2
Elastic IP
NAT
Gateway
Availability Zone B
Private
Subnet
Worker EC2 nodes
Elastic Container Registry
(ECR)
Flask Backend
Docker Image
Frontend image
Internet Gateway
wastewaterai.com
User

Figure 6:

Networking and Access The architecture allows user access via an Internet Gateway, which routes traffic to the web server EC2 instance. The Kubernetes cluster is configured to be accessible from the control plane, with the kubectl command-line tool used for cluster management operations.

EC2 Worker nodes We opted for self-managed worker nodes which allowed us to maintain a high degree of control over the Kubernetes environment. Self-managed nodes offered the flexibility to customize the operating system, security settings, and installations, ensuring that we could tailor the infrastructure to meet our specific application needs.

EC2 Web Server The architecture includes a strategically placed Apache web server within a public subnet of the VPC. The server's role is to handle incoming HTTP requests from the internet, facilitated by an Elastic IP that provides a stable public IP address for consistent user access. The Apache server is renowned for its performance and reliability, making it a fitting choice for serving static content and acting as a reverse proxy to the containerized applications within the private subnet of the VPC.

Application Deployment A Flask-based application was containerized and deployed within the Kubernetes cluster. The application serves a machine learning model through a RESTful API, responding to /predict endpoints for model inference requests.

Service and Ingress Configuration The application is exposed through a NodePort service, making it accessible within the VPC. An Ingress resource was initially considered for external exposure but was deemed unnecessary due to the single service nature of the application and direct access via the NodePort.

5 Conclusion

This study provided a comprehensive comparison between traditional and advanced modeling techniques for predicting ammonia levels in wastewater treatment systems. Through the application of Linear Regression, RNN, and LSTM models to a dataset derived from BioWin 6.2 simulations, we established that LSTM models offer superior predictive performance. The LSTM model demonstrated a lower RMSE and a higher R-squared value compared to its counterparts, thanks to its enhanced capability in capturing the complex dynamics of WWTPs. This performance underscores the potential of LSTM networks in optimizing wastewater treatment processes, potentially contributing to more sustainable environmental management practices. Future research directions will be to integrate of LSTM models with real-time data acquisition for dynamic system control and further exploration of hybrid models that combine the strengths of mechanistic and machine learning approaches. The outcomes of this study pave the way for the adoption of intelligent modeling techniques in environmental engineering, promising significant advancements in the treatment and management of wastewater.

References

[1] Batsuren Sundui, Olga Alejandra Ramirez Calderon, Omar M Abdeldayem, Jimena Lázaro-Gil, Eldon R Rene, and Uyanga Sambuu. Applications of machine learning algorithms for biological wastewater treatment: updates and perspectives. *Clean Technologies and Environmental Policy*, 23:127–143, 2021.

[2] Francesco Granata, Stefano Papirio, Giovanni Esposito, Rudy Gargano, and Giovanni De Marinis. Machine learning algorithms for the forecasting of wastewater quality indicators. *Water*, 9(2):105, 2017.

[3] François Chollet et al. Keras. `https://github.com/keras-team/keras`, 2015. Original-source software available at `https://github.com/keras-team/keras`.

[4] Tensorflow. `https://www.tensorflow.org/`, 2015. Accessed: 20-11-2023.