Addressing Irregularities and Enhancing Forecasting for Water Quality Metrics by Integrating Ordinary Differential Equations with BiLSTM

Md Abdullah Al Mazid, Liangdong Deng and Naphtali Rishe
School of Computing and Information Sciences
Florida International University

***Abstract*—Environmental monitoring data often have irregularities, like missing values or unevenly spaced observations. These irregularities are not just data anomalies; rather, they mirror the complexity and unpredictability of natural ecosystems and the challenges of collecting data in such settings. Preserving these irregularities is essential because artificial smoothing or interpolation can obscure critical insights and misrepresent natural processes. In this work, we propose a novel method that incorporates Neural Ordinary Differential Equations (Neural ODEs) with Long Short-Term Memory (LSTM)-based models, as well as feature engineering, to address these irregularities in water quality metrics forecasting. To ensure the authenticity of real-world dynamics, our methodology retains the dataset's inherent irregularities. While Long Short-Term Memory (LSTM)-based models excel in capturing sequential patterns and dependencies, Neural ODEs excel at handling uneven sampling by offering a continuous-time perspective. To make the models more adaptable, we incorporate features such as time differences between observations, timestamps, and other relevant water quality metrics. For this study, we use 30 years of data from five monitoring stations along Florida's Saint Lucie River, carefully maintaining the irregularities inherent in the dataset. This study shows that our proposed model significantly outperforms state-of-the-art Neural ODEs and recurrent architectures. This framework provides a reliable way to forecast water quality metrics while handling irregularities in time series forecasting, offering new insights and better predictions for environmental monitoring and management.**

***Index Terms*—Irregular Time Series, Neural ODEs, LSTM, BiLSTM, Water Quality, Forecasting**

I. Introduction

Time-series data is pivotal in understanding, analyzing, and predicting patterns across numerous domains, e.g., finance, healthcare, and environmental monitoring. In environmental monitoring, like water quality assessment, time series analysis carries a critical role in monitoring ecological health, paving the way for effective water resource management and policy decisions. These datasets capture the complex dynamics in aquatic systems. Typically water quality time series datasets contain information on various physical, chemical, and biological metrics across various spatiotemporal points. However, real-world time series data often contains irregularities, such as missing values and unevenly spaced sampling intervals and this characteristic introduces significant challenges for traditional forecasting models. It is highly common to handle those irregularities by various smoothing or interpolation techniques, but that may distort critical information and introduce bias into the dataset. Researchers have determined that the methodologies used for interpolation or bias correction in datasets can lead to significant forecast errors and inaccuracies [1], [2]. Thus, preserving the authenticity of environmental dynamics, which are captured in the irregularities of data, and building models to accurately analyze those data is essential for maintaining the integrity of data analysis.

Conventional models for time series forecasting, such as Long Short-Term Memory (LSTM) networks, are well-suited for capturing sequential dependencies in regularly sampled data [3]. However, they are not very effective with datasets that contain irregular intervals or missing observations. On the other hand, models like Neural Ordinary Differential Equations (Neural ODEs) excel in providing continuous-time representations, which makes it suitable for handling irregular sampling patterns [4]. Despite their strengths, Neural ODEs alone struggle to fully capture the intricate temporal dependencies and multi-variable interactions, particularly in water quality datasets, which are not fully continuous in nature and often exhibit discrete sampling patterns.

To address these challenges, the present study introduces a *novel hybrid approach* that combines Neural ODEs with LSTM-based architectures, particularly Bidirectional LSTM (BiLSTM) networks. The intuition is to leverage the strengths of both models, the capability of Neural ODEs to manage irregular sampling, and the strength of BiLSTM architectures to learn complex sequential dependencies in discrete data. We also incorporate advanced feature engineering by including time differences, timestamps, and supplementary water quality matrices to enhance the model's adaptability and predictive performance. We focus on five crucial water quality metrics e.g. dissolved oxygen, specific conductance, pH, salinity, and water temperature to capture intricate relationships among water quality parameters and improve forecasting accuracy.

The present study applies the proposed hybrid model to univariate water quality forecasting, focusing on key water quality indicators mentioned above. We use 30 years of real-world data from the Saint Lucie River of Florida, which is sourced from the *DBHydro environmental database*. The dataset's inherent irregularities are preserved to maintain the authenticity of ecological dynamics. The data is structured to align the aforementioned water quality metrics within the same spatio-temporal context. The experimental evaluations

SoutheastCon 2025 | 979-8-3315-0484-7/25/$31.00 ©2025 IEEE | DOI: 10.1109/SOUTHEASTCON56624.2025.10971531

include Symmetric Mean Absolute Percentage Error (sMAPE), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE). This study demonstrates that our model significantly outperforms standalone Neural ODEs and traditional recurrent models like LSTM and BiLSTM and achieves higher accuracy for forecasting tasks. The contributions of this work are summarized as follows:

- **Hybrid Architecture:** Integration Neural ODE with BiLSTM and development of a novel model tailored for handling irregularities in discrete time series data for forecasting purposes.
- **Feature Engineering:** Incorporation of the time difference, timestamp features, and additional water quality matrices to enhance the model's ability to capture multivariable dependencies.
- **Application to Water Quality:** Demonstration of the model's effectiveness in forecasting critical water quality metrics using real-world datasets.
- **Comparative Analysis:** Comprehensive benchmarking against standalone state-of-the-art models and showcasing the advantages of hybrid architectures for handling irregularities in time series data especially in the environmental domain.

II. Related Works

Forecasting discrete time series with irregularities is a challenging task. Traditional methods and recent advancements in machine learning and deep learning have attempted to address these challenges. However, they often fall short when applied to real-world discrete datasets with irregularities. One of the real-world discrete datasets with irregularities is water quality data. This section provides an overview of related works on time series forecasting for data with irregularities, hybrid modeling approaches, and forecasting of water quality metrics in environmental studies.

A. Irregular Time Series Forecasting

Irregularities in time series data have motivated the development of specialized forecasting methods. Classical statistical approaches, such as ARIMA and its variations [5] rely on assumptions of regular time steps. Those approaches often require data imputation techniques to handle irregular intervals. Although these methods are computationally efficient, they are not well suited for capturing complex, nonlinear relationships in the data.

Recent advancements in deep learning have produced various machine learning models for various time series-related tasks. Among those Long Short-Term Memory (LSTM) network-based models have shown promise in handling sequential dependencies in time series data [6], [7]. However, standard LSTM models typically assume a uniform time step [8], [9] and struggle to accommodate irregular sampling. To address this issue, Neural Ordinary Differential Equations (Neural ODEs) have been proposed [10], [11] which allows integration of irregular time intervals directly into the forecasting process as a continuous-time modeling approach. The effectiveness of Neural ODEs in irregular time series has been demonstrated by Rubanova et al. [12], particularly in healthcare data which is continuous in nature.

B. Hybrid Modeling Approaches

Hybrid models that combine the strengths of multiple architectures have emerged as a promising direction for time series forecasting [13]. For example, studies integrating RNNs with attention mechanisms [14] or RNNs with graph convolutional networks have demonstrated improved performance in capturing both global trends and local dependencies [15]. In environmental applications, hybrid models such as LSTM-CNNs and LSTM-Attention networks have been used to forecast air quality and climate variables [16], [17]. These approaches highlight the potential of hybrid architectures to address domain-specific challenges in time series forecasting.

Neural ODEs have also been combined with LSTM-based models in recent works to leverage their complementary strengths. For example, recent studies have shown that integrating time-aware features into hybrid ODE-RNN, ODE-LSTM models significantly improves time series modeling capacity [18], [19] for continuous data.

C. Water Quality Forecasting

Water quality forecasting plays a critical role in environmental management, public health, and policy decision-making. Metrics such as dissolved oxygen, water temperature, salinity, pH, and specific conductance are widely used as indicators of water quality [20]. Traditional methods for water quality prediction usually rely on physical models or statistical regression techniques. Those approaches often require extensive domain knowledge and are not suitable for handling nonlinear relationships [21].

Machine learning methods e.g. Support Vector Machines (SVM), Random Forests, and Gradient Boosting, have been applied to water quality forecasting [22]. Deep learning models, particularly LSTMs, have shown superior performance by capturing temporal dependencies in multivariate water quality datasets [23], [24].However, irregularities in discrete water quality data, such as missing observations due to equipment failure or sampling variability, pose significant challenges. Few studies have explored deep learning architectures specifically tailored for such irregularities in discrete time series.

D. Research Gaps and Motivations

Although Neural ODEs and LSTMs individually demonstrate their capabilities in irregular and sequential data modeling, there are gaps for handling irregularities in discrete data without deploying smoothing or interpolation techniques. Furthermore, existing works often neglect the potential of incorporating multiple metrics as features. Incorporating multiple metrics as features can improve the robustness and accuracy of forecasting. This study bridges these gaps by proposing a novel model that leverages both the continuous-time representation of Neural ODEs and the temporal dependency-capturing abilities of BiLSTM networks. The proposed model captures multi-variable dependencies and addresses irregularities in discrete time series data more effectively than the state-of-the-art methods.

III. Proposed Methodology

A. Overview

This study presents a novel framework combining Bidirectional Long Short-Term Memory networks (BiLSTM) with Neural Ordinary Differential Equations (nODE) to predict the next-step value of key water quality parameter. The model integrates sequential modeling capabilities of BiLSTM with continuous time dynamics via neural ODEs to handle irregular time series data.

B. Problem Formulation

The dataset consists of time series measurements $X = \{x_1, x_2, \ldots, x_T\}$, where each $x_t \in \mathbb{R}^7$ contains:

- Main water quality metric (the target variable),
- Four associated water quality metric (predictor variables),
- Time difference ((Δt), elapsed time since the last measurement),
- Timestamp (absolute time).

Given $X = \{x_1, x_2, \ldots, x_T\}$, the goal is to predict the value of the main metric for the next time step, X_{T+1}, while accounting for the irregular intervals (Δt) in the time series.

C. Model Architecture

1) BiLSTM Network: The input sequence X is processed by an BiLSTM network to capture temporal dependencies. The hidden state output H for each time step is computed as:

$$H_t, C_t = \text{LSTM}(x_t, (H_{t-1}, C_{t-1})),$$

$$H_t = \text{concat}(\overrightarrow{H_t}, \overleftarrow{H_t}),$$

where $\overrightarrow{H_t}$ and $\overleftarrow{H_t}$ represent forward and backward passes, respectively.

The BiLSTM outputs the hidden states for the entire sequence:

$$H = \{H_1, H_2, \ldots, H_T\}, \quad H_t \in \mathbb{R}^H,$$

where H is the hidden dimension.

The final hidden state H_{last} corresponding to the last time step is extracted for further processing:

$$H_{\text{last}} = H_T,$$

2) Neural ODE Integration: The hidden state H_{last} is considered as the initial condition for the ordinary differential equation (ODE) and is passed through a Neural ODE block to capture continuous-time dynamics. The ODE function is parameterized as follows:

$$\frac{dH}{dt} = f_\theta(H, t),$$

where f_θ is a neural network that governs how the hidden state evolves over time. In this implementation, f_θ is defined by a single linear layer followed by a ReLU activation:

$$f_\theta(H, t) = \text{ReLU}(W_H H + b),$$

The Neural ODE solver integrates the dynamics i.e. ODE over a time interval $[t_{\text{start}} = 0, t_{\text{end}} = 1]$ using a fifth-order Runge–Kutta solver, specifically the *Dormand–Prince* method (often referred to as "dopri5").

$$H_{\text{ode}}(t) = \text{odeint}(f_\theta, H_{\text{last}}, t),$$

The resulting hidden state at t_{end} which is denoted by $H_{\text{ode}}(t_{\text{end}})$ captures the *continuous-time* evolution of H_{last}. This final integrated state $H_{\text{ode}}(t_{\text{end}})$ is fed into into subsequent layer.

3) Final Prediction Layer: The Neural ODE output $H_{\text{ode}}(t_{\text{end}})$ is combined with time features of X_{T+1}:

$$Z = [H_{\text{ode}}(t_{\text{end}}), X_{T+1,\text{time difference}}, X_{T+1,\text{timestamp}}],$$

where $Z \in \mathbb{R}^{B \times (H+2)}$.

The final prediction is computed using a fully connected layer:

$$\hat{y}_{T+1} = W_Z Z + b_Z,$$

where $W_Z \in \mathbb{R}^{(H+2)\times 1}$ and $b_Z \in \mathbb{R}$.

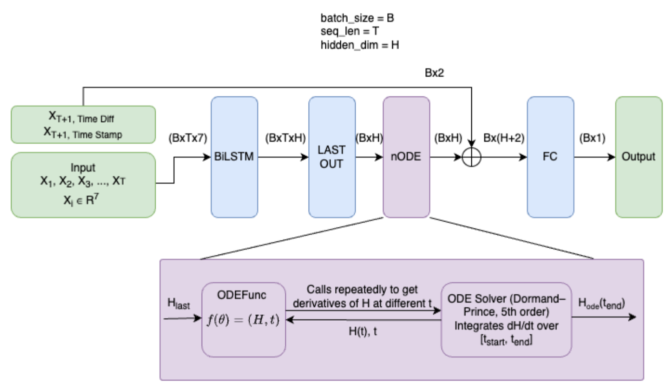

Fig. 1. Proposed Models Architecture

D. Loss Function

The model is trained to minimize the Symmetric Mean Absolute Percentage Error (sMAPE), defined as:

$$\text{sMAPE} = \frac{100\%}{N} \sum_{i=1}^{N} \frac{|y_i - \hat{y}_i|}{(|y_i| + |\hat{y}_i|)/2},$$

where:

- y_i: ground truth value of the main analyte,
- $\hat{y}_i$: predicted value of the main analyte,
- N: number of samples in the batch.

E. Code Availability

The implementation of the proposed models and experiment is publicly available at the following GitHub repository: https://github.com/mazid-rafee/irregular-ts-water-quality-ml.

IV. Experimental Setup

A. Dataset Downloading

The dataset is obtained from the DBHydro Water Quality Database which is maintained by the South Florida Water Management District (SFWMD). This database provides long-term water quality monitoring data collected from various stations across Florida. We develop a C#, .NET desktop command-line application to interface with the DBHydro API which enables efficient retrieval of the last 30 years of water quality time series data of numerous monitoring stations across Florida.

The downloaded raw data is structured with each row representing a single water quality metric observed at a specific station and timestamp. However, this raw format is not suitable for direct analysis or model training. Hence, extensive preprocessing to prepare the data for time series forecasting tasks is done.

B. Data Preprocessing

The raw data is processed through several preprocessing steps before feeding into the model.

1) Unit Harmonization: Water quality metrics in the raw data are recorded in varying units (e.g., *Specific Conductance* is observed in *uS/cm*, *mS/cm*, and *umho/cm*). These along with other water quality metrics are converted to consistent units using appropriate scaling factors to ensure uniformity across all observations. For Dissolved oxygen, we keep the unit mg/L, for Specific Conductance uS/cm and PSU and Deg C for Salinity and Water Temperature respectively.

2) Data Restructuring: Each row in the raw dataset represents a single metric for a specific station at a specific timestamp which makes it unsuitable for other water quality metrics to be used as features in our model. To resolve this, rows are grouped by *Station ID* and *Sample ID* (a unique identifier provided in the dataset for observations from the same sampling event). The grouped data is pivoted so that each row contains multiple water quality metrics observed at the same station and timestamp.

3) Station and Metric Selection: Stations with the most extensive data coverage are prioritized to ensure robust temporal patterns for forecasting. The top five stations with the highest data counts are selected: SE 01, SE 02, SE 03, SE 06, SE 09. Similarly, five key water quality metrics relevant to environmental monitoring and forecasting are chosen: *Dissolved Oxygen (mg/L)*, *pH*, *Water Temperature (°C)*, *Specific Conductance (uS/cm)*, and *Salinity (PSU)*.

4) Data Sanitizing and Normalization: To ensure data quality and consistency, we apply outlier removal and normalization during preprocessing. Outliers are identified and removed using the Interquartile Range (IQR) method, where the IQR is defined as IQR $= Q3 - Q1$, with $Q1$ and $Q3$ representing the first and third quartiles, respectively. Data points lying outside the range $[Q1 - 1.5 \times \text{IQR}, Q3 + 1.5 \times \text{IQR}]$ are considered outliers and excluded from the analysis. We also remove those rows where there are any missing values for the selected water quality metrics.

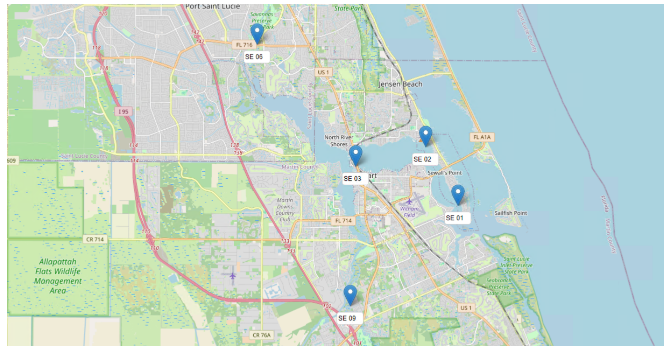

Fig. 2. Map showing the locations of the 5 stations

TABLE I
Data-points after data sanitization

	DO	SP	PH	SL	TM	Avg CV(Δt)
SE 01	3274	3272	3250	3282	3275	3.49
SE 02	2539	2524	2521	2524	2524	2.99
SE 03	4063	4080	4080	4080	4080	3.84
SE 06	2371	2364	2370	2369	2369	5.36
SE 09	2142	2149	2149	2149	2138	3.87

In the TABLE I, DO stands for Dissolved Oxygen, SP for Specific Conductivity, PH for pH, SL for Salinity, and TM for Water Temperature. This table shows the number of data points in each monitoring stations for each water quality metric. It also contains Avg CV which represents the magnitude of irregularities in the dataset. The Coefficient of Variation (CV) is calculated as the ratio of the standard deviation to the mean of time gaps, given by CV $= \frac{\sigma_{\Delta t}}{\mu_{\Delta t}}$, where $\sigma_{\Delta t}$ is the standard deviation and $\mu_{\Delta t}$ is the mean of the time gaps. From Avg CV it's evident that all five monitoring stations contain high level of irregularities in data in terms of time difference (Δt).

We also normalize the data using min-max scaling to ensure all features lie within the range $[0, 1]$, calculated as $x' = \frac{x - \min(x)}{\max(x) - \min(x)}$. Here x is the original value, $\min(x)$ and $\max(x)$ are the minimum and maximum values of the feature, respectively, and x' is the normalized value. This preprocessing approach ensures that the data is free from extreme values and scaled appropriately for further analysis.

C. Model Training and Configuration

The dataset is split into 80% for training and 20% for testing. A sequence length of 100 (T) is used to capture temporal dependencies, with a batch size of 32 (B) for computational efficiency. Our proposed model is consisted of a single bi-directional recurrent layer with 50 (H) hidden units and are trained for 20 epochs using the Adam optimizer to minimize the Symmetric Mean Absolute Percentage Error (sMAPE) loss function.

D. Evaluation Metrics

The performance of the models is assessed using sMAPE, MAE, and RMSE. We choose those metrics because each sheds light on a different aspect of our model's forecasting performance. Thus, it gives us well-rounded view of its strengths and weaknesses. sMAPE focuses on relative error which makes it useful when the data can vary significantly in scale for different water quality parameters and it's also less sensitive to outliers compared to typical percentage-based metrics. MAE, on the other hand, is easy to interpret because it provides the average magnitude of error in the same units as the original data and it helps us see how far off predictions are in practical terms. Finally, RMSE punishes larger errors more heavily by squaring the residuals, which can be especially important if big misses are costlier or riskier in the real world. Together, these three metrics paint a comprehensive picture of how well our model handles irregular time-series data.

- **Symmetric Mean Absolute Percentage Error (sMAPE):** Evaluates the relative accuracy of the forecasts and is defined as:

$$\text{sMAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{(|y_i| + |\hat{y}_i|)/2} \times 100, \tag{1}$$

  where y_i is the observed value, $\hat{y}_i$ is the predicted value, and n is the total number of observations.
- **Mean Absolute Error (MAE):** Measures the average magnitude of errors in predictions and is expressed as:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|, \tag{2}$$

  where y_i and $\hat{y}_i$ are the observed and predicted values, respectively.
- **Root Mean Square Error (RMSE):** Provides a measure of the spread of prediction errors and is given by:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}. \tag{3}$$

V. Results and Comparative Analysis

A. Overview of Results

The performance of the proposed hybrid model (nODE-BiLSTM) is evaluated against baseline models (LSTM, BiLSTM, and NeuralODE). The experiments are conducted across five monitoring stations along Saint Lucie River of Florida and we focus on five water quality metrics e.g. dissolved oxygen, specific conductance, pH, salinity, and water temperature. For each metric, we consider it as the main metric and treated the other four as associated metrics.

B. Performance Comparison

Table II summarizes the key performance metrics for each model, each water quality metric and each evaluation metric, averaged across all five stations:

In the TABLE II, DO stands for Dissolved Oxygen, SP for Specific Conductivity, PH for pH, SL for Salinity, and TM for Water Temperature. The table shows that our proposed model yields lower error rate (better accuracy) for all five water quality metrics across all three evaluation metrics.

TABLE II
Performance Metrics for Different Models

	LSTM	**BiLSTM**	**nODE**	**nODE-BiLSTM**
DO (sMAPE)	5.958	6.028	5.754	**5.382**
DO (RMSE)	0.534	0.535	0.532	**0.515**
DO (MAPE)	6.107	6.189	5.911	**5.754**
SP (sMAPE)	16.292	18.206	17.996	**15.112**
SP (RMSE)	4658.7	4814.7	5415.1	**4613.4**
SP (MAPE)	187.9	188.6	202.6	**182.9**
SL (sMAPE)	0.808	0.74	0.724	**0.682**
SL (RMSE)	3.008	2.994	2.967	**2.885**
SL (MAPE)	69.288	57.989	65.692	**56.351**
PH (sMAPE)	20.078	18.49	20.204	**16.54**
PH (RMSE)	0.096	0.095	0.094	**0.093**
PH (MAPE)	0.804	0.736	0.719	**0.687**
TM (sMAPE)	3.022	3.0118	2.542	**2.062**
TM (RMSE)	1.021	1.064	0.955	**0.905**
TM (MAPE)	2.583	2.634	2.246	**1.743**

C. Station-wise Observations

In this section station wide results for each water quality metrics is shown only for sMAPE evaluation metric. From Fig. 3, it is evident that our model achieves not only lower error rates (higher accuracy) on average across all stations but also consistently outperforms in individual station evaluations.

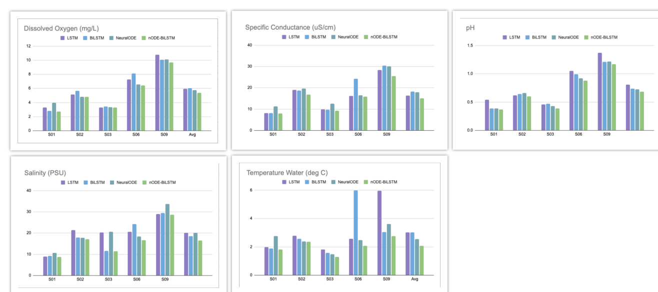

Fig. 3. Station-wise result (sMAPE)

D. Summary of Findings

Our proposed model consistently outperforms all other base models in terms of all three evaluation metrics (sMAPE, RMSE, MAPE). The models architecture demonstrates its robustness in handling irregularities in discrete time series data and offers a significant improvement over standalone state of the art models.

VI. Conclusion and Future Work

This paper proposes integration of Neural ODE with Bi-LSTM for addressing irregularities in discrete time series data and improves forecasting capabilities with a focus on various water quality metrics. By combining the continuous-time dynamics of Neural ODEs with the temporal modeling capabilities of LSTM-based architectures, the model demonstrates better performance compared to the state-of-the-art baseline methods. The integration of time-related features and auxiliary metrics further enhances the models' robustness and accuracy and also addresses the challenges posed by irregularly sampled environmental data. Experimental results show that the proposed model achieves lower error metrics across multiple stations. These findings highlight the effectiveness of integrating neural ODE with LSTM-based model for capturing temporal dependencies with irregular sampling. Thus, it makes the model highly suitable for handling irregularities in discrete time series data for forecasting in environmental and other domains.

Our future research will focus on multivariate and long sequence time-series forecasting, enabling the model to handle interdependent variables and better capture complex relationships within the data and also to enhance the model's ability to predict over extended horizons, addressing challenges such as error accumulation and computational efficiency. Additionally, we plan to explore our models efficacy for continuous-time modeling in areas such as healthcare (irregular patient monitoring), finance (asynchronous trading data), and environmental monitoring (multi-sensor networks), where effective handling of data irregularities is crucial for generating accurate long-term predictions. We will also investigate potential applications of our model in diverse domains such as finance and climate science by incorporating transfer learning strategies to handle irregularly sampled data scenarios.

This study demonstrated the potential of integration of Neural ODE and LSTM-based architectures to bypass the necessity of data augmentation techniques for handling irregularities in time series forecasting, thus providing a significant step forward in addressing the challenges of handling irregularities in time series forecasting across various disciplines.

Acknowledgment

This material is based in part upon work supported by the National Science Foundation under Grant Nos. MRI20 CNS-2018611 and MRI CNS-1920182, and the Florida Department of Environmental Protection (FDEP) B-BROIPS Grant C-2104.

References

[1] M. K. Tippett, L. Trenary, T. DelSole, K. Pegion, and M. L. L'Heureux, "Sources of bias in the monthly cfsv2 forecast climatology," *Journal of Applied Meteorology and Climatology*, vol. 57, pp. 1111–1122, 2018, doi: 10.1175/JAMC-D-17-0299.1.

[2] C. Skoulikaris, P. Venetsanou, G. Lazoglou, C. Anagnostopoulou, and K. Voudouris, "Spatio-temporal interpolation and bias correction ordering analysis for hydrological simulations: An assessment on a mountainous river basin," *Water*, vol. 14, p. 660, 2022, doi: 10.3390/w14040660.

[3] H. Kang, S. Yang, J. Huang, and J. Oh, "Time series prediction of wastewater flow rate by bidirectional lstm deep learning," *International Journal of Control, Automation and Systems*, vol. 18, pp. 3023–3030, 2020, doi: 10.1007/s12555-019-0984-6.

[4] P. Kidger, J. Morrill, J. Foster, and T. Lyons, "Neural controlled differential equations for irregular time series," *Advances in Neural Information Processing Systems*, vol. 33, pp. 6696–6707, 2020.

[5] Q. Zhou, J. Hu, W. Hu, H. Li, and G.-z. Lin, "Interrupted time series analysis using the arima model of the impact of covid-19 on the incidence rate of notifiable communicable diseases in china," *BMC Infectious Diseases*, vol. 23, p. 375, 2023, doi: 10.1186/s12879-023-08229-5.

[6] H. Yadav and A. Thakkar, "Noa-lstm: An efficient lstm cell architecture for time series forecasting," *Expert Systems with Applications*, vol. 238, p. 122333, 2024, doi: 10.1016/j.eswa.2023.122333.

[7] O. M. Osama, K. Alakkari, M. Abotaleb, and E.-S. M. El-Kenawy, "Forecasting global monkeypox infections using lstm: A non-stationary time series analysis," in *2023 3rd International Conference on Electronic Engineering (ICEEM)*, 2023, pp. 1–7, doi: 10.1109/ICEEM58740.2023.10319532.

[8] P. Wang, X. Zheng, G. Ai, D. Liu, and B. Zhu, "Time series prediction for the epidemic trends of covid-19 using the improved lstm deep learning method: Case studies in russia, peru and iran," *Chaos, Solitons & Fractals*, vol. 140, p. 110214, 2020, doi: 10.1016/j.chaos.2020.110214.

[9] K. Albeladi, B. Zafar, and A. Mueen, "Time series forecasting using lstm and arima," *International Journal of Advanced Computer Science and Applications*, vol. 14, pp. 313–320, 2023, doi: 10.14569/IJACSA.2023.0140133.

[10] P. Goyal and P. Benner, "Neural ordinary differential equations with irregular and noisy data," *Royal Society Open Science*, vol. 10, p. 221475, 2023, doi: 10.1098/rsos.221475.

[11] Q. Zhu, X. Li, and W. Lin, "Leveraging neural differential equations and adaptive delayed feedback to detect unstable periodic orbits based on irregularly sampled time series," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 33, 2023, doi: 10.1063/5.0143839.

[12] R. T. Chen, Y. Rubanova, J. Bettencourt, and D. K. Duvenaud, "Neural ordinary differential equations," *Advances in neural information processing systems*, vol. 31, 2018.

[13] Z. Hajirahimi and M. Khashei, "Hybridization of hybrid structures for time series forecasting: A review," *Artificial Intelligence Review*, vol. 56, pp. 1201–1261, 2023, doi: 10.1007/s10462-022-10199-0.

[14] N. Zrira, A. Kamal-Idrissi, R. Farssi, and H. A. Khan, "Time series prediction of sea surface temperature based on bilstm model with attention mechanism," *Journal of Sea Research*, vol. 198, p. 102472, 2024, doi: 10.1016/j.seares.2024.102472.

[15] A. Lazcano, P. J. Herrera, and M. Monge, "A combined model based on recurrent neural networks and graph convolutional networks for financial time series forecasting," *Mathematics*, vol. 11, p. 224, 2023, doi: 10.3390/math11010224.

[16] A. Gilik, A. S. Ogrenci, and A. Ozmen, "Air quality prediction using cnn+ lstm-based hybrid deep learning architecture," *Environmental science and pollution research*, vol. 29, pp. 1–19, 2022, doi: 10.1007/s11356-021-16227-w.

[17] H. Zhou, T. Wang, H. Zhao, and Z. Wang, "Updated prediction of air quality based on kalman-attention-lstm network," *Sustainability*, vol. 15, p. 356, 2022, doi: 10.3390/su15010356.

[18] L. Li, J. Yan, Y. Zhang, J. Zhang, J. Bao, Y. Jin, and X. Yang, "Learning generative rnn-ode for collaborative time-series and event sequence forecasting," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, pp. 7118–7137, 2022, doi: 10.1109/TKDE.2022.3185115.

[19] M. Habiba and B. A. Pearlmutter, "Neural ordinary differential equation based recurrent neural network model," in *2020 31st Irish signals and systems conference (ISSC)*, 2020, pp. 1–6, doi: 10.1109/ISSC49989.2020.9180182.

[20] R. Nayar, "Assessment of water quality index and monitoring of pollutants by physico-chemical analysis in water bodies: a review," *International Journal of Engineering Research and Technology*, vol. 9, pp. 180–185, 2020.

[21] T. Rajaee, S. Khani, and M. Ravansalar, "Artificial intelligence-based single and hybrid models for prediction of water quality in rivers: A review," *Chemometrics and Intelligent Laboratory Systems*, vol. 200, p. 103978, 2020, doi: 10.1016/j.chemolab.2020.103978.

[22] M. Y. Shams, A. M. Elshewey, E.-S. M. El-Kenawy, A. Ibrahim, F. M. Talaat, and Z. Tarek, "Water quality prediction using machine learning models based on grid search method," *Multimedia Tools and Applications*, vol. 83, pp. 35 307–35 334, 2024, doi: 10.1007/s11042-023-16737-4.

[23] Z. Gao, J. Chen, G. Wang, S. Ren, L. Fang, A. Yinglan, and Q. Wang, "A novel multivariate time series prediction of crucial water quality parameters with long short-term memory (lstm) networks," *Journal of Contaminant Hydrology*, vol. 259, p. 104262, 2023, doi: 10.1016/j.jconhyd.2023.104262.

[24] S. Khullar and N. Singh, "Water quality assessment of a river using deep learning bi-lstm methodology: forecasting and validation," *Environmental Science and Pollution Research*, vol. 29, pp. 12 875–12 889, 2022, doi: 10.1007/s11356-021-13875-w.