From LSTM to Transformers: A Comprehensive Analysis of Deep Learning Techniques for Groundwater Contamination Forecasting at the Hanford Site

Rathinaraja Jeyaraj, Aobo Jin, Hilary Emerson, Daniel I. Kaplan, and Hardik A. Gohel

Abstract

Keywords

1. Introduction

Hanford Site

Groundwater contamination

Climate change impact

Importance of AI-based solutions

Time-series forecasting importance

Different time-series models

Advantages of using DL-based models

Different types of DL models for time-series forecasting

Motivation of this manuscript

List of contributions of the manuscript

Reminder of the section

2. Background

Physics based models

Statistics-based models

DL-based models

3. Problem Formulation

Univariate forecasting

Multi-variate forecasting

Single step and multi-step forecasting

Sliding window – self supervised learning

Include definition of seq2seq

Timeseries characteristics (seasonality, trend) and Challenges

4. Deep learning models for time-series forecasting

4.1.1 RNN

4.1.2 LSTM

4.1.3 Bi-LSTM

4.1.4 GRU

4.1.5 Bi-GRU

4.1.6 CNN

4.1.7 TCN

4.1.8 Encoder-decoder

4.1.9 Attention

4.1.10 Transformer

DL architecture	Advantages	Disadvantages	Remarks

5. Results and analysis

5.1 Application area and dataset descriptions

Climate change data

Groundwater contamination data

5.2 Forecast error metrics

Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE)

Coefficient of Determination

5.3 Performance analysis

Sequence length 4

Dataset	RNN		LSTM		Bi-LSTM		GRU		Bi-GRU		CNN		TCN		ED		Attention		MHA	
	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
Dataset1																				
Dataset2																				
Dataset3																				
Dataset4																				
Dataset5																				
Dataset6																				

Loss and accuracy convergence graphs.

Test error distributions

Model	Test error	Dataset1		Dataset2		Dataset3		Dataset4		Dataset5		Dataset6	
		μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
RNN													
LSTM													
BI_LSTM													
GRU													
BI_GRU													
CNN													
TCN													
ED													
Attention													
MHA													

Test error distribution graphs.

Forecast comparison graphs.

Confidence and prediction interval graphs.

Pattern exploration using PCA.

Sequence length 50

Dataset	RNN		LSTM		Bi-LSTM		GRU		Bi-GRU		CNN		TCN		ED		Attention		MHA	
	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
Dataset1																				
Dataset2																				
Dataset3																				
Dataset4																				
Dataset5																				
Dataset6																				

Loss and accuracy convergence graphs.

Test error distributions

Model	Test error	Dataset1		Dataset2		Dataset3		Dataset4		Dataset5		Dataset6	
		μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
RNN													
LSTM													
BI_LSTM													
GRU													
BI_GRU													
CNN													
TCN													
ED													
Attention													
Transformer													

Test error distribution graphs.

Forecast comparison graphs.

Confidence and prediction interval graphs.

Pattern exploration using PCA.