Incorporating Long Memory into Echo State Networks

Shinjini Paul[1], **Tanujit Chakraborty**[2,3]

[1]Indian Institute of Science Education and Research, Pune, India
[2]Sorbonne University, Abu Dhabi, UAE
[3]Sorbonne Center for Artificial Intelligence, Paris

shinjini.paul@students.iiserpune.ac.in, tanujit.chakraborty@sorbonne.ae

Abstract

The 'echo state' approach developed by Jaeger [2001] has proven to be an effective method for analyzing and training recurrent neural networks, addressing the limitations of traditional methods in capturing complex temporal dynamics. ESNs are efficient at capturing short-term dependencies in time series data, but they lack the ability to model long-range dependencies, where autocorrelations decay at a polynomial rate. In this work, we propose a novel ESN architecture that incorporates Fractionally Differenced Autoregressive Integrated Moving Average (ARFIMA) filters to model long-memory processes. The input layer is augmented with wavelet decomposition to separate high- and low-frequency components, which has been shown to improve time series prediction (Soltani [2002]). The low-frequency components, which carry long-range dependencies, are processed through ARFIMA filters before being fed into the reservoir. Our approach retains the computational advantages of traditional ESNs while significantly improving their performance on tasks involving long-memory time series.

1 Introduction

Echo State Networks (ESNs) are a subclass of recurrent neural networks (RNNs) that are widely used due to their efficient training process. However, their intrinsic short-term memory characteristics make them less suited for tasks requiring the modeling of long-range temporal dependencies, such as those found in certain financial or climate data. This paper addresses this limitation by integrating Fractionally Differenced Autoregressive Integrated Moving Average (ARFIMA) filters with ESNs to enhance their long-memory capabilities.

Long-memory processes, characterized by slowly decaying autocorrelations, are typically found in systems where past events continue to influence future outcomes over extended time intervals. Traditional ESNs fail to model such processes effectively, as their reservoirs are optimized for short-term memory dynamics. By incorporating ARFIMA filters and wavelet decomposition, we aim to extend the memory capacity of ESNs while maintaining their computational efficiency.

2 Methodology

The core innovation in our proposed architecture lies in preprocessing the input data using wavelet decomposition to separate high- and low-frequency components. This method builds on previous findings that highlight the significance of incorporating autoregressive structures in neural networks for improved time series prediction as in Hornik et al. [2000]. Low-frequency components, which are indicative of long-range dependencies, are processed through ARFIMA filters before being fed into the ESN's reservoir. This preprocessing allows the reservoir to focus on short-term dynamics while the ARFIMA filters capture the long-term dependencies.

1st Conference on Applied AI and Scientific Machine Learning (CASML 2024).

The reservoir itself remains unchanged from the traditional ESN architecture, ensuring that the computational efficiency of the model is not compromised. By combining the strengths of ARFIMA and ESN, our approach creates a hybrid model capable of handling both short- and long-memory dynamics.

2.1 Wavelet Decomposition

Wavelet decomposition is used to decompose the input signal into different frequency bands. We focus on retaining the low-frequency components that typically carry long-range dependencies, while discarding high-frequency noise.

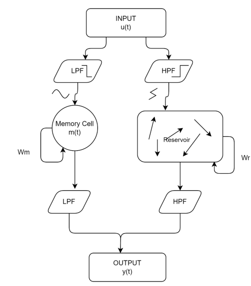

Figure 1: wESN

2.2 ARFIMA Filtering

The low-frequency components are processed through ARFIMA filters, which are specifically designed to handle long-memory processes. ARFIMA allows us to model the autocorrelations in the data more effectively than traditional ESNs, which are limited to short-memory processes. Using the methodology given in Granger and Joyeux [1980] we implement this in our model.

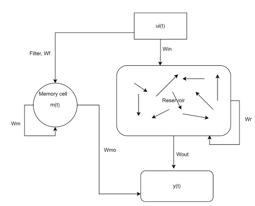

Figure 2: fESN

3 Results and Evaluation

To evaluate the performance of our proposed model, we tested it on both synthetic datasets, designed to exhibit long-range dependencies, and real-world time series datasets. We compared the performance of our enhanced ESN model against traditional ESNs and standalone ARFIMA models. We expect full results by the time of the conference.

4 Conclusion

In this paper, we proposed a novel Echo State Network architecture that integrates ARFIMA filters and wavelet decomposition to model long-memory processes in time series data. Our approach effectively extends the memory capacity of ESNs while preserving their computational efficiency. The experimental results on both synthetic and real-world datasets demonstrate the model's superiority over traditional ESNs and ARFIMA models. Future work will explore the application of this model to more complex, multi-dimensional time series data.

References

Clive W. J. Granger and Roselyne Joyeux. An introduction to long memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1(1):15–29, 1980.

Kurt Hornik, Adrian Trapletti, and Friedrich Leisch. Stationary and integrated autoregressive neural network processes. *Neural Computation*, 12:2427–2450, 02 2000. doi: 10.1162/089976600300015006.

Herbert Jaeger. The "echo state" approach to analysing and training recurrent neural networks—with an erratum note. Technical report, German National Research Center for Information Technology, Bonn, Germany, 2001. Available at: https://www.researchgate.net/publication/220350686.

Skander Soltani. On the use of the wavelet decomposition for time series prediction. *Neurocomputing*, 48:267–277, 2002. Received 31 October 2000, Accepted 6 June 2001, Available online 29 August 2002.