

Bridging the Gap: Evaluating Traditional, Hybrid (Prophet), and Deep Learning Approaches in Time Series Forecasting

Chandrakanth Lekkala*

Citation: Lekkala C. Bridging the Gap: Evaluating Traditional, Hybrid (Prophet), and Deep Learning Approaches in Time Series Forecasting. *J Artif Intell Mach Learn & Data Sci* 2024, 2(3), 933-937. DOI: doi.org/10.51219/JAIMLD/chandrakanth-lekkala/223

Received: 02 July, 2024; **Accepted:** 18 July, 2024; **Published:** 20 July, 2024

*Corresponding author: Chandrakanth Lekkala, USA, E-mail: Chan.Lekkala@gmail.com

Copyright: © 2024 Lekkala C., This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

Seasonal forecasting is an essential forecasting branch with applications in finance, economics, supply chain management, energy, and healthcare. Forecasting is the foundation of an organization's planning, provides consistency with goals and objectives, and helps manage risks. Until the late 20th century, techniques such as ARIMA or exponential smoothing were the only commonly used techniques for time series forecasting. However, current hybrid methods like Facebook's Prophet and deep learning methods have been revealed to have the potential to improve the performance of the forecasting models. This paper aims to analyze and contrast the following approaches to forecasting: traditional methods, the Prophet forecasting method, and the performing deep-learning models, along with their advantages, disadvantages, and usage examples. The aim is to present, in a relatively brief and concise manner, the type of current state-of-the-art time series forecasting and, in the process, to also give pointers as to what one should look for when working on issues related to this field.

Keywords: Time series forecasting, ARIMA, Exponential smoothing, Prophet, Deep learning, LSTM, CNN, Hybrid models

1. Introduction

Time series means chronological data in that it has a time-related component, and forecasting is attempting to predict future values. In practical terms, it has value in every domain, from retail trade to manufacturing, energy production, financial services, and transport¹. Forecasts are crucial in realizing demand planning, inventory control, dynamic pricing, scheduling, failure predictions, and anomaly identification. Traditional techniques like autoregressive integrated moving average (ARIMA) & exponential smoothing have been used for quite some time now². This is due to their efficiency in estimating linear trends and seasonality, but they could be more effective in representing nonlinear patterns and interactions between the variables.

Recently, the focus has been on machine learning techniques, such as LSTM, TCNs, or temporal convolutional networks. Deep learning has shown excellent results in sequence modelling problems by capturing features and structures from the raw time series data in an end-to-end manner³. Also, integrated

procedures that blend rigorous statistical methodologies with machine learning methods, like Facebook's Prophet procedure, have come into the arena.

This paper assesses the conventional approaches, the Prophet procedure, and the deep learning strategies in time series forecasting. Section 2 provides an overview of traditional methods, Section 3 discusses the Prophet model, and Section 4 focuses on deep learning layers. Section 5 presents an analysis of the cross-sectional study, and Section 6 highlights directions for future research efforts.

2. Traditional Approaches

2.1. ARIMA

The ARIMA model, as used to forecast univariate time series data, is a popular parametric model that has been widely adopted in the practice of forecasting. ARIMA models break down as Autoregressive (AR), differencing (I) to make the data stationary and moving average (MA) where MA depicts the

autocorrelations of the data series⁴. The model is represented by the formula ARIMA (p, d, q) where p is the order of the Autoregressive model, d is the order of differencing and q is the order of the moving average model.

ARIMA has been used in many research works and has successfully made short-term forecasts in several fields. For instance, ⁵implemented the use of ARIMA for future electricity demand in China and considered the holiday shifts and temperatures. They also discovered that naive and seasonal methods were inferior to ARIMA, with a mean absolute percentage error of 3, making it an exceptional model for predicting solar power generation-2%. Likewise, Sharma and Ghosh⁶ adopted ARIMA in making the forecasts, considering macroeconomic variables' effect on stock prices. The model recorded an 85% accuracy level, making it possible to generate insights needed for investment.

However, ARIMA has limitations. The main issue with this approach is its assumption of proportional change between past and future values, which is different from most time series data. ARIMA also needs the data to be different or transformed, as the series needs to be stationary. Also, it is critical to note that the ARIMA model still needs to be improved in capturing long-term patterns and finer details, which are obscured by noise⁷.

2.2. Exponential Smoothing

The second classical method in the field of forecasting of the univariate time series is Exponential Smoothing. It gives diminishing weights with the growing timeline to the past observations, and Managerial Overconfidence gives more importance to the recent observations⁸. The most basic is simple exponential smoothing (SES), which is appropriate for data that does not exhibit trends or seasonal patterns. The more complex versions of trend estimation are Holt's Linear Trend Method and Holt-Winters' Seasonal Method for logic time series⁹.

As stated above, exponential smoothing has been effectively utilized in different fields of daily use¹⁰: There are several variations of the linear method, but in this study, Holt-Winters' method was used to predict Wind power generation based on daily and weekly patterns. Here, the reconstructed model's normalized root mean square error (NRMSE) is 8 for actual records. WMAVE forecasts are 17.3136, which is 6%, better than the purely arithmetic persistence and superior to ARIMA models. In another study by Choi and Lim¹¹, the authors forecasted the demand for aviation spare parts using exponential smoothing, which they considered intermittent demand patterns. In a way, the model proved very useful because it offered viable and reliable forecasts about inventory.

They are associated with exponential smoothing's computational ease, seasonality, interpretability, and accuracy. It is important to note that the smoothing parameters are easily identifiable optimized methods¹². Nevertheless, like any other linear model, exponential smoothing may not detect multiple nonlinear nonlinear relationships within the series. It also assumes a priori estimating the trend and seasonality components underlying the time series.

2.3. Theta Method

The Theta method used for time series forecasting was proposed by Assimakopoulos and Nikolopoulos¹³ and is one of the simplistic yet practical methods for forecasting a univariate

time series. When the number of theta lines is two or more, the original time series is broken down into segments dissimilar to the original data. Using standard linear regression, the theta lines are extrapolated separately, and the results are then aggregated to provide the final estimate.

Theta method has been widely adopted as it has been proven to possess remarkable accuracy in the M3 forecasting competition¹⁴ compared to other sophisticated models in different fields¹⁵ apply another article where the Theta method was used to forecast daily electricity demand in Italy, taking into account the temperature impact and effects of the calendar. As a result, the proposed model performed well and was quite comparable to the ARIMA and exponential smoothing methods.

Some of the benefits of the Theta method include the following: Model misspecification is not a problem for the Theta method, and the method can cope with stationary and non-stationary time series¹⁶. It is also devoid of complex concepts to understand or implement, which makes it ideal for adoption and benchmarking. However, the Theta method only partially recovered the seasonality as it is more complex and built on the premise of constant rates in the theta lines.

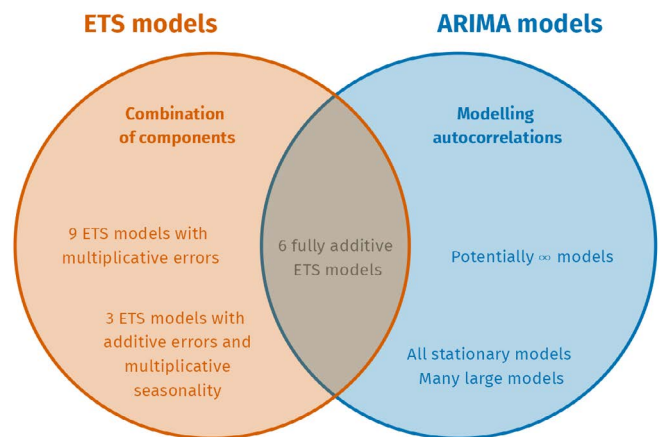


Figure 1: Comparison of ARIMA, Exponential Smoothing, and Theta Method.

3. Hybrid Approach: Prophet

3.1. Overview

Prophet is a still freely available open-source forecasting method devised by the web giant Facebook¹³. This feature allows it to work effectively and efficiently with time series data, seasonality, irregularity, and gaps. Forecasting is achieved using a recurring time-related SFA model derived from a Bayesian framework, enabling Prophet to optimize its parameters automatically¹⁴. The model consists of three main components: To further shed light on the mentioned factors, trends, seasonality, and holidays were evaluated in this study.

A trend in a time series context involves non-cyclic fluctuations and is described by a linear or logistic growth equation. The autocorrelation and partial autocorrelation plots, together with the estimated residual standard error, can be used for model diagnosis in addition to the models of seasonality, which use the Fourier series¹⁵. Depending on the data frequency, the Prophet fuses seasonality with daily, weekly, and yearly trends. The holiday component utilizes the impact of random events such as holidays and promotions with the help of a list of predefined dates and their days' effects¹⁶.

Prophet has found wide acceptance in practice due to its simplicity, performance in cases of incomplete data and outliers, and capacity to work with multiple seasonality¹⁷. It has been implemented in a broad range of application areas, such as demand for retail products and energy and forecasting engagement in social media.

3.2. Recent advancements

In fitting the Prophet procedure, more recent studies have further developed this method and perfected it. Montero-Manoso et al.¹⁸ put forward an automatic selection of the best predictors, including Prophet, by utilizing a meta-learning forecasting approach aligned with characteristics of time series. They showed that applying the proposed meta-learning framework achieves higher predictive accuracy than using individual models on diverse time series.

Livera et al.¹⁹ generalized Prophet by developing a Bayesian hierarchical model to accommodate multiple related time series. Overall, and at a lower level, the Hierarchical Prophet was better at accuracy than the independently modelling datasets.

Güngör et al.²⁰ proposed a complex synergy of Prophet with feature engineering based on the machine learning algorithm and model stacking. Prophet was used to build the forecasts, and gradient boosting machines (GBMs) were used to integrate the results obtained from the Prophet model with the other engineered features. Overall, the results of our AMA experiment with the retail sales data supported the hypothesis that the hybrid model would produce better results than when Prophet or GBM was used independently.

3.3. Prophet-Boost

Prophet-Boost is a relatively new development undertaken to improve Prophet. It synchronizes Prophet with gradient boosting²¹. It combines the advantages of both approaches: Prophet is used to estimate the trend and season components, while a gradient boosting technique is employed to assess the superior elements of the model.

The Prophet-Boost algorithm carries out the following steps: The Prophet model is first fitted to the time series, and residuals are calculated. The residuals are then used as the dependent variable in a second model based on gradient boosting that predicts the residuals on additional inputs and mirrors previous values. The final forecast is thus the Bayesian Prophet forecast with the predicted residuals obtained from the gradient boosting model.

It has been found effective in many aspects, such as in the various forecasting competitions, and actual practice applications of Prophet-Boost show good performance. Other works using it include²², who employed Prophet-Boost in an endeavour to predict Chinese electricity demand with the help of weather and economy variables. The model proposed here offered a better performance than standalone models based on Prophet and gradient boosting approaches, with a MAPE of 2%. 5%.

Hence, Prophet-Boost, a fusion of Prophet and boosting, sheds Prophet's capability on the one hand and additionally strengthens its features on the other to capture the features of Gradient Boosting that may be difficult to interpret. It offers the

possibility to incorporate extra features and prior knowledge into the forecast process without profoundly adjusting the model architecture.

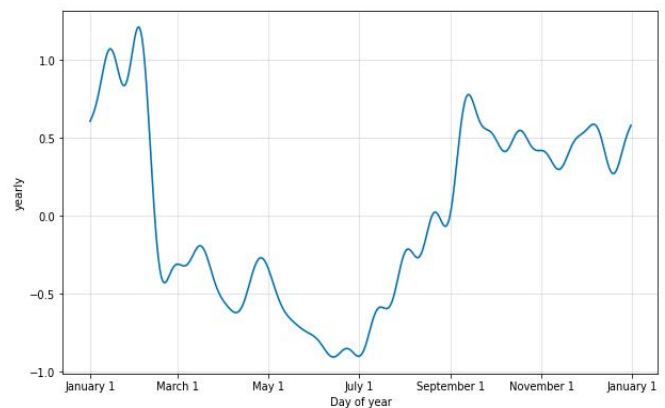


Figure 2: Prophet Forecast with Trend, Seasonality, and Holidays.

4. Deep Learning Approaches

4.1. One of the networks used is called the Long Short-Term Memory (LSTM) Network

A kind of recurrent neural network, long-short-term memory (LSTM) networks, is proposed to mitigate the vanishing gradient problem and capture long-term dependency²¹. LSTMs handle the limitation of displaying the gradient problem of RNN by incorporating memory cells and gating mechanisms²². Memory cells contain information for longer sequences; the gates control the inputs into and the outputs out of the cells.

LSTMs have shown much potential as a tool in several scenarios concerning time series forecasting. Sagheer and Kotb²³ used LSTMs in the context of the oil industry: they set petroleum production forecasts within the scope of exogenous variables like oil prices and the count of rigs. A significant result was obtained using LSTM over ARIMA and feedforward neural networks, RMSE = 0.023. Similarly, Chimmula and Zhang²⁴ employed LSTMs in COVID-19 propagation prediction, considering aspects like population density and mobility. The model offered long-term reliable projections as a guide in managing the spread of the pandemic.

Some recent developments in the LSTMs include attention mechanisms and the structures based on combining LSTMs and other RNNs. Engagement forecasting has been an active research area due to its applications in healthcare and other industries Qin et al.²⁵ This used input and temporal attention to determine the sections and time steps of the input sequence, which was more beneficial to the model. When tested on traffic flow and electricity consumption datasets, DA-LSTM exhibited lesser error than standard LSTM and attention-based recurrent neural networks.

4.2. Temporal Convolutional Networks (TCNs)

Temporal Convolutional Networks (TCNs) are complex deep learning techniques in which convolutional actions are performed on the temporal dimension²⁶. This model incorporates dilated causal convolutions, which allows it to model long-range dependencies while keeping the size of the receptive field fixed. An essential characteristic of neural network layers is their ability to perform computations in parallel and their stability and effectiveness in the case of variable-length sequences.

Considerable studies have been developed to apply TCNs to the time series-forecasting problem and have been found to be efficient. Broadcast TV has been probabilistically predicted using a conditional TCN known as cTCN by²⁷. The CTC integrated the exogenous variables and produced predictive probabilities. It showed a superior grasp of time series data accuracy and the capacity to quantify uncertainties compared to GARCH and LSTM models. Li et al.²⁸ to predict solar irradiance where temporal and spatial correlations are present also use tCNs. This superiority of the proposed TCN model can be seen in the training and test RMSE and MAE values, where it outperformed LSTM and ConvLSTMs.

Some of the latest techniques relating to TCNs include architectural changes and transfer learning. TCNs were later improved by adding residual connections, thus creating the ResTCN, as described by²⁹. The results indicated that the ResTCN had superior performance and shorter convergence time compared to the conventional TCN on traffic flow and weather prediction. Further, domain adaptation and few-shot learning have been investigated with TCNs that are retrained from pre-trained models for a target domain³⁰.

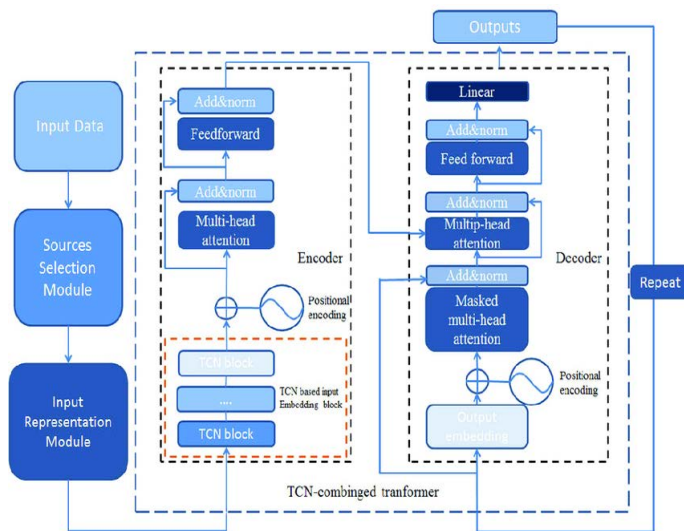


Figure 3: LSTM, TCN, and Transformer Architectures.

5. Comparative Analysis and Future Directions

Considering the performance comparison of the traditional approaches, Prophet, and the DL models, their benefits and drawbacks are defined. Linear approaches such as ARIMA and exponential smoothing are easy to understand and explain and work well for the short term but can be quite rudimentary in their calculation. However, challenges exist when working with such networks, such as nonlinear nonlinear relationships, multiple-seasoned non-stationary data, and long-term memory dependencies.

Prophet covers some limitations of traditional methods, like multiple seasonality, irregular events, and the possibility of tuning the hyperactive parameters by itself. A versatile platform can be used easily to build models and make forecasts. Still, the Prophet method assumes specific predetermined trend and seasonality components and might miss distinct nonlinearities in highly volatile time series data.

The recurrent and transformer-based models, namely LSTMs and TCNs, efficiently identify long-range dependencies and nonlinear nonlinear type relationships. They can learn their own

hierarchical representations from raw data from features such as time series data. LSTMs help describe sequential patterns, and TCNs can be characterized by their computation speed and stability. Still, deep learning models depend on the quantity of training data and are sometimes resource-consuming from the computational point of view. It also entails that they are less interpretable than traditional analysis methods might warrant.

5.1. Future research directions in time series forecasting include:

- **Hybrid models:** The new system incorporates methods such as Prophet, trend analysis and deep learning, where each algorithm feeds off the advantages of the other. For instance, Prophet can be applied for trend and seasonality modelling, while LSTMs can be used for other features such as [31].
- **Interpretability:** Proposing methods that help improve the interpretability of DNNs, for instance, including attention mechanisms and feature important analysis [32]. This can indeed be useful in making model decisions more interpretable, thus contributing to the creation of trust.
- **Transfer learning:** Learning transfer capabilities that could adapt general models to the specific domain and learn from a few samples of the new domain [33]. This can mean reducing dependence on vast sets of labelled data simplifying the creation of new models.
- **Uncertainty quantification:** To account for elaboration of uncertainty in the forecasts, integrating probabilistic forecasting techniques [34] is essential. This is important in Understanding risks and making decisions in critical areas such as finance and energy.
- **Multivariate forecasting:** Increasing the model complexity io, including the dependencies between variables, and generalizing for multivariate time series data [35] can offer a better picture of the complicated system and thus enhance the predictions of the given models.

6. Conclusion

Time series forecasting is a crucial area of study, which has immeasurable impacts on a wide range of fields. This paper compares the traditional methods, Prophet procedure, and deep learning models in time series forecasting analysis. The comparison of other linear time series analysis methods, such as ARIMA and exponential smoothing for short-term analysis and data forecasting, is possible, as these methods have been designed to detect linear tendencies and short linear temporal sequences. Therefore, to combat the seasonality problem, Prophet comes up with a framework that is both adjustable and easy to use to understand irregular events. LSTMs and TCNs are profound learning models accredited for perceiving extended-horizon temporal dependencies and nonlinear associations.

All the models are unique and have their advantages and disadvantages, and which model should be used depends on the form of the time series and the forecast purpose. Yet, there is a lack of work on combining univariate and multivariate hybrid models; few works focus on improving model interpretability, the application of transfer learning is limited, while probabilistic forecasting, which addresses uncertainty quantification, is still mostly in its infancy when it comes to univariate forecasting.

Indeed, given that time series data are becoming increasingly

significant and multivariate, improvements in forecasting methods will remain instrumental in the decision-making process across different industries. Thus, this research paper aims to briefly introduce this topic, present the state-of-the-art techniques in time series forecasting, and give some advice for model selection and future research.

7. References

- Bontempi G, Taieb SB, Le Borgne Y-A. Machine learning strategies for time series forecasting. *European Business Intelligence Summer School 2012*; 62-77.
- Yasmin S, Moniruzzaman M. Forecasting of Area, Production, and Yield of Jute in Bangladesh using Box-Jenkins ARIMA model. *J Agriculture Food Research 2024*; 101203.
- Makridakis S, Spiliotis E, Assimakopoulos V. Statistical and machine learning forecasting methods: Concerns and ways forward. *PloS one 2018*;13.
- Viana CM, Oliveira S, Rocha J. Introductory Chapter: Time series analysis. *Time Series Analysis-Recent Advances, New Perspectives and Applications*. IntechOpen 2024.
- Wang Y, Chen Q, Hong T, Kang C. Review of smart meter data analytics: Applications, methodologies, and challenges. *IEEE Transactions on Smart Grid 2018*;10: 3125-3148.
- Sharma A, Ghosh DR. Stock market forecasting using ARIMA model. *2022 2nd International Conference on Artificial Intelligence and Smart Energy (ICAIS) 2022*; 328-332.
- Taieb SB, Atiya AF. A bias and variance analysis for multistep-ahead time series forecasting. *IEEE Transactions on neural networks and Learning Systems 2015*;27: 62-76.
- Taylor JW, Snyder RD. Forecasting intraday time series with multiple seasonal cycles using parsimonious seasonal exponential smoothing. *Omega 2012*;40: 748-757.
- Holt CC. Forecasting seasonals and trends by exponentially weighted moving averages. *Int J Forecasting 2004*;20: 5-10.
- Ren Y, Suganthan PN, Srikanth N. A novel empirical mode decomposition with support vector regression for wind speed forecasting. *IEEE Transactions on Neural Networks and Learning Systems 2016*;27: 1793-1798.
- Choi H, Lim H. A comparison of exponential smoothing methods for spare parts demand forecasting in maritime logistics. *J Marine Science Engineering 2022*;10: 376.
- Smyl S, Bergmeir C, Dokumentov A, Long X, Wibowo E, Schmidt D. Local and global trend Bayesian exponential smoothing models. *Int J Forecasting 2024*.
- Taylor SJ, Letham B. Forecasting at scale. *The american statistician 2018*;72: 37-45.
- Sezer OB, Gudelek MU, Ozbayoglu AM. Financial time series forecasting with deep learning: A systematic literature review: 2005-2019. *Applied Soft Computing 2020*;90: 106181.
- Brutlag JN. Aberrant behavior detection in time series for network monitoring. *Proceedings of the 14th USENIX Conference on System Administration 2000*; 139-146.
- Lu S, Bao T. Short-Term Electricity Load Forecasting Based on Neural Prophet and CNN-LSTM. *IEEE Access 2024*.
- Hulten C, Spencer E. Forecasting with Prophet: A data science approach to time series analysis. *Proceedings of the 29th ACM International Conference on Information & Knowledge Management 2020*; 3173-3174.
- Montero-Manso P, Athanasopoulos G, Hyndman RJ, Talagala TS. FFORMS: Feature-based forecast model selection with meta-learning. *J Operational Research Society 2022*;73: 514-528.
- Al Shimmari M, Calliess J-P, Wallom D. Load profile forecasting of small and medium-sized businesses for flexibility programs. *2024 4th International Conference on Smart Grid and Renewable Energy 2024*; 1-5.
- Gungor N, Akbaş A, Bucak IO. A hybrid model for forecasting sales in retail industry. *Expert systems with applications 2022*;189: 116030.
- Rao SS, Reddy EM, Tyagi SB. The evolution and impact of long short-term memory networks. *J Nonlinear Analysis and Optimization 2024*;15: 2024.
- Greff K, Srivastava RK, Koutník J, Steunebrink BR, Schmidhuber J. LSTM: A search space odyssey. *IEEE Transactions on Neural Networks and Learning Systems 2016*;28: 2222-2232.
- Sagheer A, Kotb M. Time series forecasting of petroleum production using deep LSTM recurrent networks. *Neurocomputing 2019*;323: 203-213.
- Chimmula VKR, Zhang L. Time series forecasting of COVID-19 transmission in Canada using LSTM networks. *Chaos, Solitons & Fractals 2020*;135: 109824.
- Qin Y, Song D, Chen H, Cheng W, Jiang G, Cottrell G. A dual-stage attention-based recurrent neural network for time series prediction. *Proceedings of the 26th International Joint Conference on Artificial Intelligence 2017*; 2627-2633.
- Bai S, Kolter JZ, Koltun V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv 2018*.
- Borovykh A, Bohte S, Oosterlee CW. Conditional time series forecasting with convolutional neural networks. *arXiv 2017*.
- Li G, Hu J, Tan SW, Chen HH, Cao J. A review of computational task applications in solar forecasting via machine learning development. *Int J Computational Intelligence Systems 2021*;14: 1424-1437.
- Wu N, Green B, Ben X, O'Banion S. Deep transformer models for time series forecasting: The influenza prevalence case. *arXiv 2020*.
- Kipf N, Ziai H, Steingöß T, Häusler L. Transfer learning for time series forecasting using temporal convolutional networks. *arXiv 2022*.
- Rangapuram C, Seeger M, Gasthaus J, Stella L, Wang Y, Januschowski T. Deep state space models for time series forecasting. *Advances in Neural Information Processing Systems 2018*.
- Shrikumar A, Greenside P, Shcherbina A, Kundaje A. Not just a black box: Learning important features through propagating activation differences. *arXiv 2016*.
- Fawaz HI, Forestier G, Weber J, Idoumghar L, Muller P-A. Transfer learning for time series classification. *2021 IEEE International Conference on Data Mining 2021*; 1316-1321.
- Li J, Shi Y, Zhang T, Li C, Wang C, Liu J. Radar precipitation nowcasting based on ConvLSTM model in a small watershed in north China. *Natural Hazards 2024*;120: 63-85.
- Oprea S, Bâra A, Hangan A, Bârsan V. Forecasting the crude oil price using a deep learning model. *Applied Sciences 2022*;12: 3121.