Short Thesis for the Degree of Doctor of Philosophy (PhD)

Time-Series with Multiple Seasonal Periods: Modeling and Forecasting

By Chudo Solomon Buke
Supervisor: Prof. Dr. Gyorgy Terdik

UNIVERSITY OF DEBRECEN
Doctoral School of Informatics

Debrecen, 2025

Contents

1 Introduction

This dissertation addresses the significant issue of predicting time series data that display numerous, intricate seasonal patterns, a prevalent feature in contemporary datasets from fields such as energy management and epidemiology. Even though foundational models like the multiplicative Seasonal Auto Regressive Integrated Moving Average (SARIMA) give us a way to deal with periodic parts [31], and state-space frameworks have been made to combine different seasonal complexities [11], there are still big methodological and technical gaps. Current sophisticated methodologies, such as Box-Cox transformation, ARMA errors, Trend and Seasonal components (BATS), Trigonometric seasonality, Box-Cox transformation, ARMA errors, Trend and Seasonal components (TBATS) [11], and Seasonal and Trend decomposition using Loess (STL) [8], provide flexibility but frequently do not utilize frequency-domain insights for the accurate determination of predominant seasonal intervals.

In order to fill this gap, the present study proposes an approach that is based on spectral density and is systematic in nature for the purposes of detecting seasonality and configuring models. The methodology that has been put forward relies on the periodogram analysis that Shumway has described [27]. Prior to incorporating them into the modeling process, the suggested method rigorously detects major seasonal frequencies. This methodology is utilized to improve the configuration and performance of BATS, TBATS, and STL models, with assessment carried out using standard forecasting accuracy criteria [20].

The effectiveness of this method is substantiated through empirical datasets, encompassing electricity consumption—which exhibits well-documented sub-daily, half-daily, daily, weekly, or other cycles affected by operational and climatic variables [7]—and epidemiological data from the COVID-19 pandemic (World Health Organization, 2021).

The present research establishes a unified theoretical foundation by deriving state-space representations for core time series models (AR, MA, ARMA, ARIMA, Exponential Smoothing), following the framework of [19]. This forms the basis of an important methodological advancement: a revised and organized strategy that uses spectral density analysis using the periodogram to determine and incorporate predominately seasonal frequencies into multi-seasonal models [27]. This method improves model setup using spectral guidance by identifying main periodicities, like the sub-daily, half-daily, daily, weekly, or other cycles common in energy use [7].

The efficacy of this integrated approach is demonstrated through two detailed case studies. First, using World Health Organization (WHO) data from 2020–2021, a SARIMA model is developed to forecast daily COVID-19 fatalities in Hungary. The fitted model
$SARIMA(1,1,2)(1,0,1)_{[7]}$ yields significant diagnostics and normally distributed residuals, forecasting a declining trend in mortality. Second, the analysis turns to hourly energy consumption data from Brazil, which exhibits multiple seasonalities [6]. Here, a Double Seasonal Holt-Winters (DSHW) model enhanced with ARMA(3,1) errors is shown to success-

fully address residual autocorrelation and outperform the standard model across key metrics [20]. Furthermore, when modeling the identified multi-seasonal patterns, a hybrid STL+ETS(A,N,N) method produces the most accurate forecasts, outperforming BATS and TBATS models.

In summary, this dissertation makes four important contributions: (1) it creates state-space representations for basic time series models; (2) it improves SARIMA forecasting for epidemiological data; (3) it improves the combined DSHW-ARMA modeling approach; and (4) it creates a strong, spectrally informed framework for choosing the dominant frequency and modeling multiple seasons. By carefully choosing and customizing models for unique seasonal structures, this work gives professionals in dynamic domains better, more reliable, and easier-to-understand tools for making decisions based on data [6, 5, 7].

2 Statement of the Problem and Research Objectives

To make sound decisions about things like public health and energy management, it's crucial to be able to effectively assess and predict time-series data. The successful application of statistical methodologies depends on two persistent challenges: first, the selection and specification of models that precisely capture the underlying dynamics of a dataset, including trends, seasonality, and noise; second, the adaptation of standard models to fit complex, real-world data characterized by multiple or nested seasonal patterns and residual autocorrelation.

This dissertation tackles these challenges by an intensive analysis in two distinct domains. The challenge of precisely forecasting the track of an epidemic is examined through the case of COVID-19 mortality in Hungary in the context of public health. Traditional forecasting methods might not be able to fully account for the distinctive serial dependencies and seasonal trends that are common in epidemiological time-series data, which limits their usefulness for making policy decisions. One major problem in the subject of energy systems is being able to reliably predict how people will use energy when there are many seasonal cycles that overlap, including sub-daily, daily, weekly, and so on. Conventional forecasting techniques often face considerable complexity, leading to inadequate predictions that hinder efficient grid management and resource distribution. Specifically, classic seasonal models require enhancement by addressing residual autocorrelation and meticulously identifying the primary seasonal frequencies in high-frequency data [25].

To make time-series analytics more useful in the real world, we need to solve these difficulties. This will give those who work in public health and energy better forecasting tools that are more reliable, detailed, and useful.

Scientific Goals

To advance time series forecasting methodologies by modifying approaches capable of handling seasonal patterns in time series data.

Specific Objectives

- Identifying and constructing the most suitable time series model for predicting future confirmed cases and mortality rates of the COVID-19 outbreak.
- Employing a new approach to state and observation equations for formulating the State Space ARIMA model derived from the conventional ARIMA model.
- Creating an advanced modeling framework for the DSHW model, integrating ARMA (p,q) error corrections to improve forecasting precision.
- Enhancing Time Series Forecasting using the DSHW and ARMA Error Corrections.
- Presenting a new methodology utilizing periodograms derived from spectral density analysis to ascertain significant seasonal periods.
- Forecasting and modeling time series with varying degrees of seasonality.
- Comparing the forecasting performance of seasonal models with other advanced time series forecasting techniques.

3 Methodology

State Space Model formulations

A State Space model offers a mathematical framework by representing dynamic systems and time series data through unobserved (latent) states that evolve over time. This is a mathematical depiction of a state-space model. In this equation, y_t

signifies the observed time series, $\mathbf{F}_t$ represents the observation equation matrix, and $\mathbf{G}$ indicates the state transition equation matrix. Furthermore, $\mathbf{x}_t$ denotes the state vector. The observation and state disturbances are denoted by the vectors x_t and w_t, respectively. The equation for the observation [13, 5] is as follows:

$$y_t = \mathbf{F}\mathbf{x}_t + v_t \tag{1}$$

In this equation, $\mathbf{F}$ is a matrix that delineates the contribution of state variables to the observed variable, whereas v_t represents the observation disturbance. The matrices $\mathbf{G}$ and $\mathbf{R}$ are presumed to be initially known. The equation for state transition [5, 13] is provided as follows:

$$\mathbf{x}_{t+1} = \mathbf{G}\mathbf{x}_t + \mathbf{R}w_t \tag{2}$$

There are a number of benefits to using state-space ARIMA models. These include being able to incorporate additional components like seasonality and exogenous variables, being efficient with computationally effective estimation using the Kalman filter, naturally handling missing data, and providing confidence intervals for forecasts to quantify uncertainty.
Combining state space with ARIMA, the state space ARIMA (p, d, q) model is a sophisticated time series framework. We model and project trend, seasonality, and stochastic data. For stationarity the model employs p term for AR, d term for differencing, and q term for MA [14]. For parameter estimation, forecasting, and missing data handling, state-space ARIMA models apply the Kalman filter. Reference [5] refers mainly to the ex-

ample and state-space formulations, basically, it shows the real applications of the ARIMA model in state-space form.
Combining state-space models with seasonality forms a state-space SARIMA model. Time-series data would be perfect for this strong framework. This model expands based on the classic state-space ARIMA as another representation of the classical ARIMA model. An expansion of the ARIMA paradigm used by the State Space SARIMA includes seasonal dynamics directly in the state vector [28]. To better show periodic patterns in time series data, SARIMA is a development of ARIMA that uses seasonal AR/MA components and seasonal differencing. The next Section discuss SARIMA [6] in detail.

3.1 Multiplicative Seasonal ARIMA Model

Seasonality is a pervasive characteristic of real-world time series, representing periodic fluctuations tied to specific intervals (e.g., daily, monthly, or yearly cycles). This pattern is formally defined by the correlation between observations separated by a fixed seasonal period, SS, a feature common across fields from economics to environmental science [18]. To model such data, the classical non-seasonal ARIMA framework is insufficient, as it cannot capture these recurring cyclical patterns.
To address this core limitation, Box and Jenkins [4] fundamentally extended the ARIMA model, introducing the multiplicative Seasonal Autoregressive Integrated Moving Average (SARIMA) model. This extension, denoted as $SARIMA(p, d, q)(P, D, Q)_s$, systematically incorporates seasonal

autoregressive (P), differencing (D), and moving average (Q) components alongside their non-seasonal counterparts (p,d,q). While the model structure expands to accommodate seasonality, the essential model-fitting methodology—encompassing identification, estimation, and diagnostic checking—remains consistent with the foundational Box-Jenkins procedure for standard ARIMA models [26]. This elegant generalization has established SARIMA as a principal and widely adopted tool for analyzing and forecasting seasonal time series.
The mathematical formula for the SARIMA model can be expressed as follows [2]:

$$\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D y_t = \theta_q(B)\Theta_Q(B^s)\varepsilon_t$$

$$(1-\phi_1 B-\phi_2 B^2-\cdots-\phi_p B^p)(1-\Phi_1 B^s-\Phi_2 B^{2s}-\cdots-\Phi_P B^{Ps})(B^s)(1-B)^d(1-B^s)=(1+\theta_1 B+\theta_2 B^2+\cdots+\theta_q B^q)(1+\Theta_1 B^s+\Theta_2 B^{2s}+\cdots+\Theta_Q B^{Qs})\varepsilon_t$$

3.2 DSHW Model and ARMA Error Corrections

The standard Holt-Winters exponential smoothing technique does not include an algorithm that supports several seasonal patterns. This is because there is no such algorithm. Reference [30] made several modifications to the double season technique in 2003. Both additive trend and multiplicative seasonality are utilized in the Holt-Winters methodology. Multiplicative seasonality involves multiplying two seasonal factors together. Based on the information provided by [30], the Holt-Winters approach with a combination of our newly proposed approach for double multiplicative seasonality are presented. I followed

the procedures to create my modified model, which I named a Combined Model. This model combines a Double Seasonal Holt-Winters (DSHW) framework with an ARMA process on the residuals.
Combined Forecasting Formula:

$$\widehat{y_t} = (\widehat{l_t} + \widehat{b_t})\widehat{s}_t^{(1)}\widehat{s}_t^{(2)} + \widehat{w}_t \tag{3}$$

Where:

$\widehat{l}_t, \widehat{b}_t, \widehat{s}_t^{(1)}$, and $\widehat{s}_t^{(2)}$ are the forecasts from the DSHW model.

$\widehat{w}_t$ is the forecast from the ARMA model.

OR

$\widehat{y}_{t+h}(Combined) = \widehat{y}_{t+h}(DSHW) + \widehat{w}_{t+h}$ .... forecast on horizon

h

Combining the structural forecast from the DSHW model with the ARMA model's error adjustments produces a more accurate overall projection.

3.3 BATS, TBATS, and STL Model

3.3.1 BATS and TBATS Models

Exponential smoothing, a fundamental class of forecasting methods, is mathematically grounded in the framework of state-space models through non-linear adaptations. This foundational connection, as established by [11], provides a powerful lens for understanding and extending these techniques. The following is an extension of the DS model that incorporates T seasonal patterns, ARMA errors, and a Box-Cox transformation [11, 7, 18] for BATS and TBATS models:

$$y_t^{\omega} = \begin{cases} \frac{(y_t^{\omega})}{\omega} - \frac{1}{\omega}, & \text{if } \omega \neq 0 \\ \log(y_t), & \text{if } \omega = 0 \end{cases}$$

Here, ω refers Box Cox transformation.
The general formula for y_t^{ω}, l_t, and $s_t^{(j)}$ are given as:

$$y_t^{(\omega)} = l_{t-1} + (b_{t-1})\phi + s_{t-m_1}^{(1)} + s_{t-m_2}^{(2)} + \cdots + s_{t-m_T}^{(T)} + w_t,$$

as formulated by [11].
As noted by de Livera et al. [11], the TBATS model employs a trigonometric representation of seasonal components, utilizing a Fourier series basis. This formulation allows for the modeling of complex, potentially non-integer seasonal patterns that are difficult to capture with conventional methods.

3.3.2 STL Model

According to [9], one common way to break down time series into its component parts [1] is using the STL model. The MSTL+EST model extends the STL model with an exponential smoothing trend component and many seasonal periods. We combine exponential smoothing forecasting capabilities with the STL model's ability to accommodate many seasons. Here is a description of the MSTL + EST model equation:

$$y_t = l_t + b_t + s_t^{(1)} + s_t^{(2)} + \cdots + s_t^{(n)} + \varepsilon_t$$

$$y_t = l_t + b_t + \sum_{i=1}^{n} s_t^{(i)} + \varepsilon_t$$

4 Results

4.1 Thesis One: State Space ARIMA Model Formulation and Its Application in Energy

Time series forecasting is essential for modeling dynamic systems in fields like economics and energy planning. While the classic ARIMA model [3] is widely used, it often struggles to capture complex, evolving patterns in data where underlying states change over time.
State-space models [22, 16] address this by formally modeling latent, time-varying states. This research integrates these approaches by developing the State-Space ARIMA (SSARIMA) model, which recasts ARIMA within a flexible state-space framework to better adapt to system dynamics.
The dissertation makes two core contributions:

A rigorous derivation of SSARIMA, detailing its state and observation equations.

An empirical demonstration of its superior forecasting accuracy compared to standard ARIMA, applied to electricity consumption data in Hungary.

This work provides a more powerful modeling tool for forecasting complex real-world time series.
The state-space approach can be applied to various models [5]. As an illustrative example, consider the state-space formulation

of the ARIMA(2,1,2) model, given by:
The following section details the formulation of the ARIMA(p, d, q) model in state space form, which is then applied to forecast the energy data in this thesis.
In time series, ARIMA(p,d,q) model formula is given by:

$$y_t^* = \phi_1 y_{t-1}^* + \phi_2 y_{t-2}^* + ... + \phi_p y_{t-p}^* + w_t + \theta_1 w_{t-1} + ... + \theta_q w_{t-q} \quad (4)$$

Where: $y^* = \Delta^d y_t$

The formulation of the ARIMA(p, d, q) model in state space form of Equation 4 is:

$$\mathbf{x}_{t+1} = \begin{bmatrix} \mathbf{t}_{d\times d} & \mathbf{1}_{d\times 1} & \mathbf{0}_{d\times p} \\ \mathbf{0}_{(m-1)\times d} & \boldsymbol{\phi}_{(m-1)\times 1} & \mathbf{I}_{(m-1)\times p} \\ \mathbf{0}_{1\times d} & \phi & \mathbf{0}_{1\times p} \end{bmatrix} \mathbf{x}_t + \begin{bmatrix} \mathbf{0}_{d\times 1} \\ 1 \\ \boldsymbol{\theta}_{(m-1)\times 1} \end{bmatrix} w_t \quad (5)$$

Where:

$$\mathbf{t}_{d\times d} \to \begin{cases} 1, & \text{if } d \in \triangle, \text{ where } \triangle \text{ is upper triangle of the block} \\ 0, & \text{if } d \notin \triangle \end{cases},$$

it is $d \times d$ matrix.
The state space formulation offers a powerful framework for time series analysis and forecasting by systematically integrating autoregressive, differencing, and moving average components. Representing an ARIMA(p, d, q) model in this form clearly captures the underlying dynamics of the series. This approach facilitates efficient estimation, filtering, and prediction through two core equations: the observation equation, which relates the observed data to a latent state vector, and the state

equation, which governs the evolution of that state over time. Together, these equations distinguish between the systematic structure of the series and its random disturbances.

N.B. The formal definitions of the state space, its associated metrics and vectors, and the governing state equations are established in Sections 2 and 3. Detailed derivations and illustrative examples are provided therein.

This thesis introduces a robust and new state-space formulation approach of the Autoregressive Integrated Moving Average (ARIMA) model, termed the State Space ARIMA (SSARIMA). The proposed framework leverages the structural versatility of state-space representations to enhance the standard ARIMA methodology. By explicitly modeling the data-generating process through state and observation equations, the SSARIMA provides a more refined and interpretable characterization of time series dynamics, capturing both observed data and latent state variables. The model's forecasting performance is empirically evaluated against the conventional ARIMA model using monthly Hungarian energy consumption data. Results indicate that the SSARIMA achieves superior predictive accuracy, even when benchmarked against a well-specified ARIMA model. These findings suggest that the SSARIMA is a powerful tool for time series analysis, particularly in domains where understanding underlying structural dynamics is critical. The framework effectively bridges classical time series modeling with more complex approaches, offering a synergistic blend of interpretability and enhanced forecasting performance, thereby presenting a valuable new avenue for both researchers and practitioners.

The dissertation has developed and presented an original state-space formulation of the ARIMA model. The SSARIMA framework advances time series analysis by integrating a state-space perspective, which facilitates a deeper structural understanding of series dynamics. A comparative performance analysis, conducted using monthly Hungarian power consumption data, demonstrates that the SSARIMA model yields superior forecasting accuracy relative to its conventional ARIMA counterpart. The enhanced efficacy of the SSARIMA, particularly in capturing the intrinsic dynamics of complex systems, underscores the significant advantage of incorporating state-space representations into traditional time series models.
The results establish the SSARIMA as an effective and versatile forecasting tool with broad applicability across disciplines such as economics, statistics, engineering, and computer science, where state-space modeling can provide a more profound comprehension of underlying processes. Future research may focus on extending the application of SSARIMA models to diverse data types and further refining the model architecture to augment its predictive capability and robustness.
The main publication related to this thesis is [J2].

4.2 Thesis Two: Multiplicative Seasonal ARIMA Modeling and Forecasting

The Hungarian government and the WHO collaborated to compile coronavirus statistics in Hungary for the COVID-19 pandemic country-wise profile. From January 3, 2020, to Septem-

ber 1, 2021, there were 812,531 confirmed cases of COVID-19, with 30,059 deaths, according to the WHO study. This data is available on the website;
https://ourworldindata.org/coronavirus/country/hungary [24]. The modeling in this work was based on the daily death data of COVID-19 from October 4, 2020, to May 12, 2021, which comprises approximately 221 observations [6].
My modeling approach involves a comparison of three time-series models: SARI, SMA, and SARIMA. The first step is to characterize the data's autocorrelation using ACF and PACF plots. This critical step informs both the choice of the best-fitting model and the determination of its optimal lag structure.The ACF plot reveals that the ARI (1) and IMA (1) seasonal parts of the model are the ones that are being considered. $SARIMA(1,1,2)(1,0,1)_{[7]}$ is the candidate model that can be identified based on estimates [6]. Within the confidence bounds, the analysis discovered that the
$ARIMA(1,1,2)(1,0,1)_{[7]}$ model's estimated ACF and residuals' normal distribution are acceptable. Figure 1 shows that there is no noticeable pattern in the ACF and a normality of the residuals, indicating that the model is of excellent quality. With a significant p-value for each of the lags tested, the Ljung-Box test also gives its stamp of approval [6].

Therefore, it would be a good idea to use the model for our forecasts.The estimated coefficients show the predicted values for the $ARIMA(1,1,2)(1,0,1)_{[7]}$ model. The model has an estimated sigma squared value of 0.028, a log likelihood of 80.4,

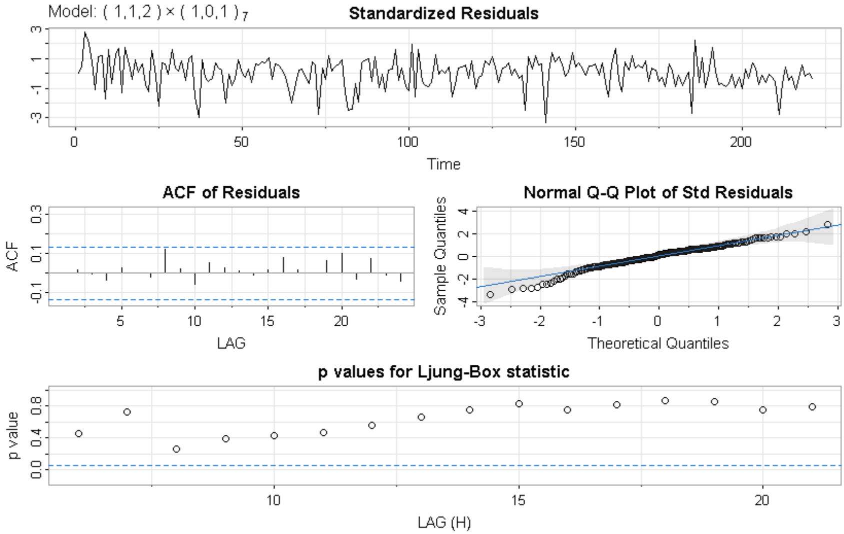

Figure 1: Plot of standardized residuals for the ARIMA model with parameters (1, 1, 2) × (1, 0, 1) [7]

and an AIC value of -148.8 [6]. This AIC value is the lowest compared to all other models for this dataset.

All parameter estimations show statistical significance, while there is no correlation between the residual series (i.e. white noise), and the model passes the diagnostic testing and estimation phases. The daily death series may now be predicted using a well-fitting multiplicative seasonal $ARIMA(1,1,2)(1,0,1)_{[7]}$ [6]. Making predictions about future daily death rates in Hungary based on known time series data for COVID-19 is known as forecasting.

This project's data only extends from 2020-10-04 to 2021-05-12; hence, Figure 2 displays the projected 14 daily fatalities in

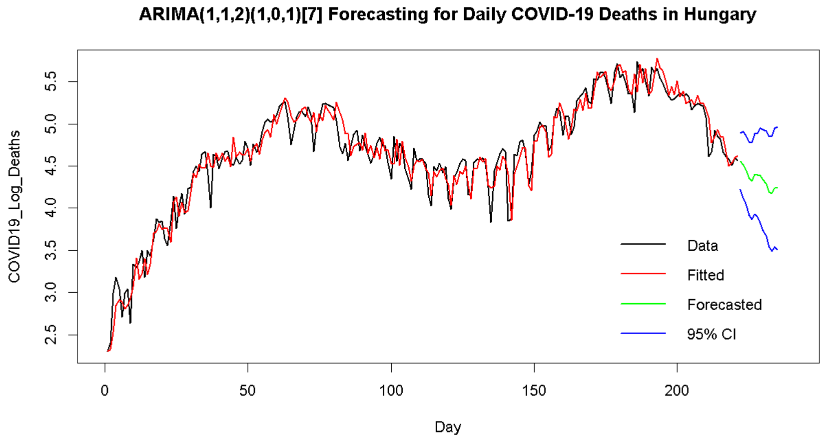

Figure 2: Predicting the number of fatalities in Hungary on a daily basis using $ARIMA(1,1,2)(1,0,1)_{[7]}$ model

Hungary between May 13, 2021 to May 26, 2021. We used the $ARIMA(1,1,2)(1,0,1)_{[7]}$ fitted model to make the predictions. The Figure 2 shows the original data plot on the black line, the fitted line on the red line, the 14-day forecast of daily deaths using the fitted model on the green line, and the 95 percent CI for future daily deaths with COVID-19 on the blue line. The selected model is $ARIMA(1,1,2)(1,0,1)_{[7]}$ [6]. Over the specified date window, the forecast generally showed that the number of daily deaths [21] in Hungary caused by COVID-19 was decreasing.

Thesis one forecasts COVID-19 daily deaths in Hungary using a Seasonal $ARIMA(1,1,2)(1,0,1)_{[7]}$ model. Analysis of WHO data from October 2020 to May 2021 confirmed weekly seasonality. The fitted model, validated by its residuals, indicates a

decreasing trend in daily deaths.
The main publication related to this thesis is [C1].

4.3 Thesis Three: Time Series Forecasting with DSHW Model and ARMA Error Corrections

Thesis two used the most recent three months of Brazilian hourly energy utilization data [29] from 2021 to 2022. This data clearly shows seasonality and a trend pattern. Before analyzing time series, we should consider models that change seasonality and trend patterns. We proposed to adjust models and procedures to accommodate seasonality of the data. Seasonality must be considered or adjusted to model time series data. These must be taken into account before analyzing. Due to my data's numerous seasonal trends, We advocate advanced multiple seasonal time series models like DSHW, BATS, TBATS, SARIMA, and STL. This chapter uses the DSHW model with ARIMA errors to handle complicated seasonal dynamics and increase forecast accuracy.
In Figure 3, the analysis of the DSHW model's residuals reveals significant issues with its fit. The residuals are not normally distributed, show a non-constant variance, and exhibit strong autocorrelation (as seen in the ACF plot, Q-Q plot, and a significant Ljung-Box test result). To address these problems, we propose and evaluate an improved model. This modified model integrates the DSHW model with an ARMA(p,q) component specifically to correct for the autocorrelated errors in the residuals. The goal of this hybrid model is to achieve better pre-

dictive accuracy than the original DSHW model or a DSHW model with only an AR(1) error correction [30].

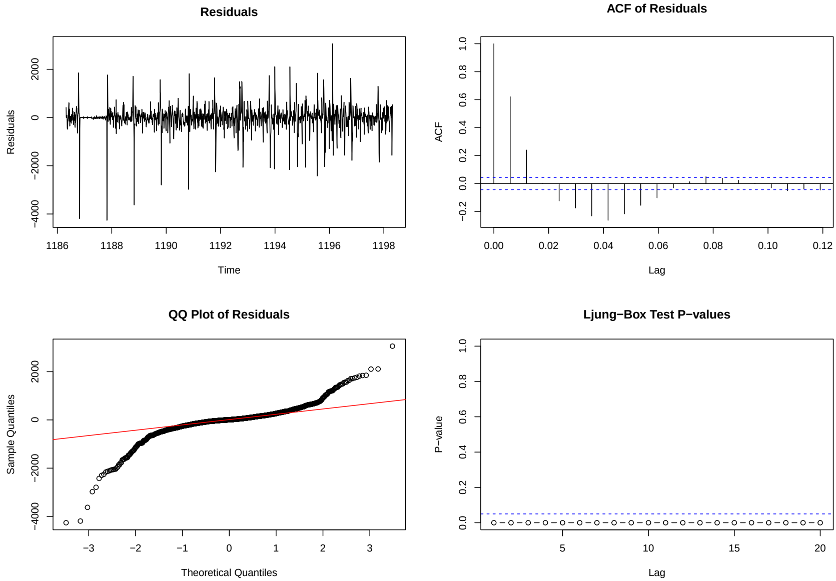

Figure 3: Plot for Residuals

Based on performance metrics Table 1, the revised DSHW model with ARMA errors is the most accurate. It shows lower error values across all measures (ME, RMSE, MAE, etc.) [18] than the standard DSHW model with AR(1) error, demonstrating superior predictive performance.

Table 1: Model Performance Metrics

Model	**ME**	**RMSE**	**MAE**	**MPE**	**MAPE**	**ACF1**	**Theil U**
DSHW	8440.580	9567.033	8486.932	12.873	12.969	0.965	4.395
Combined	584.189	3227.526	2532.749	0.865	4.030	0.951	1.561

Based on performance metrics Table 1 and visual analysis Figure 4, the new Combined Model (DSHW with ARMA error correction) outperforms the standard DSHW model with AR(1) error. The Combined Model's forecasts are more accurate, as shown by lower error values and a better fit to test data. Recent studies have shown that residual analysis can improve the accuracy of models [15]. Reference [19] investigated the potential for seasonal decomposition and ARIMA models to mitigate residual autocorrelations. This improvement is achieved by using an ARMA(3,1) model to correct residual autocorrelations, a more effective approach than the simpler AR(1) correction. The study concludes that the Combined Model is a more precise and reliable tool for forecasting electric power consumption and other time-series data.

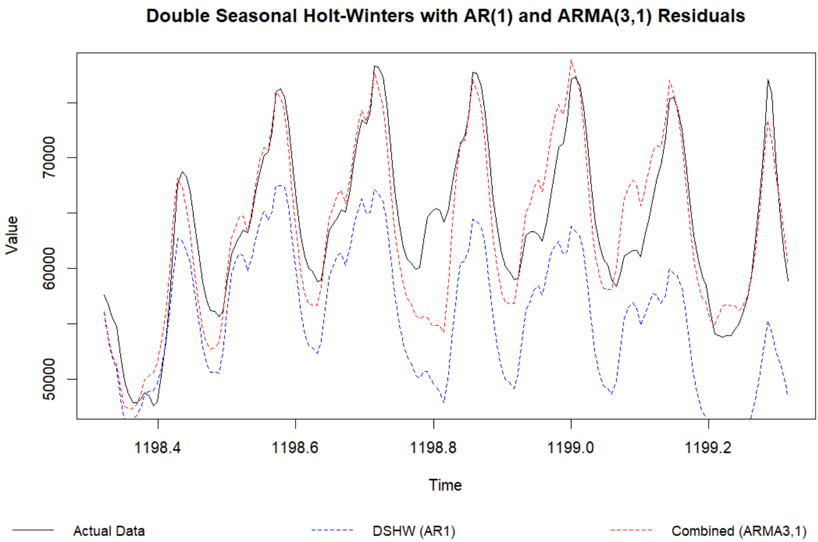

Figure 4: Comparison of DSHW model with AR (1) and ARMA (3,1)

Within the context of forecasting and modeling time-series data, thesis two demonstrates how the incorporation of ARMA (3,1) errors into the DSHW framework might be of use. Compared to the standard DSHW model, the Combined Model performed significantly better on each and every one of the metrics that were analyzed. For the sake of this performance analysis, we decided to use ME, RMSE, MAPE, and a few other metrics. Through incorporation of ARMA-based errors into the model, we are able to rectify residual autocorrelation and enhance the accuracy of the forecast. The results illustrate that the model is capable of predicting the amount of energy that will be consumed. In addition to this, it can be utilized to deal with data that require careful management due to the fact that it contains a number of seasonalities and errors that are connected to among themselves.

The publication related to this thesis are [J3, C2, C4].

4.4 Thesis Four: Utilizing Periodograms for Modeling and Forecasting Time Series with Multiple Seasonal Patterns

Thesis three examined 201,318 observations over 23 years about hourly electric energy consumption in Brazil [29]. We used the latest 7 years of data on hourly electricity consumption in Brazil from January 1, 2015, to January 1, 2023 for our further analysis since the number of total data points is too large.

The first step in choosing the top 5 most essential seasonality in a batch of time series data is identifying the most notable periodic trends in it. This stage is very important for reducing the underlying structure of the data in cases with many seasonal trends. Developed from spectral density analysis, the power spectrum - which gauges the intensity of every data frequency component - forms the basis of the selection. Table 2 shows the selected five dominant frequencies from the above general ideas [12].

Table 2: Top 5 Dominant Frequencies in the Power Spectrum

	Frequency	**Spectrum (dB)**	**Period**
1	0.041 666 67	76.214 37	24.000 00
2	0.083 333 33	73.704 71	12.000 00
3	0.083 349 61	66.175 03	11.997 66
4	0.166 666 67	65.026 88	6.000 00
5	0.083 365 89	62.786 32	11.995 31

For demonstration purposes, Table 2 is plotted as figure 5.

The periodogram and the top five seasonality components will assist us in understanding the periodic tendencies of Brazil's power consumption statistics. Data interpretation is made easier by the plotly constructed interactive labels and highlights each dominating frequency. Research and forecasting benefit much from the hourly lengths of the reported seasonal components. This approach helps to grasp periodograms better and is ready for advanced time series modeling and decomposition. A big step forward in time series analysis is the TBATS and BATS models, which can handle many seasons. Reference [10] proposed this new model that merges BATS and trigonometric seasonality. By expressing the seasonality terms trigonometri-

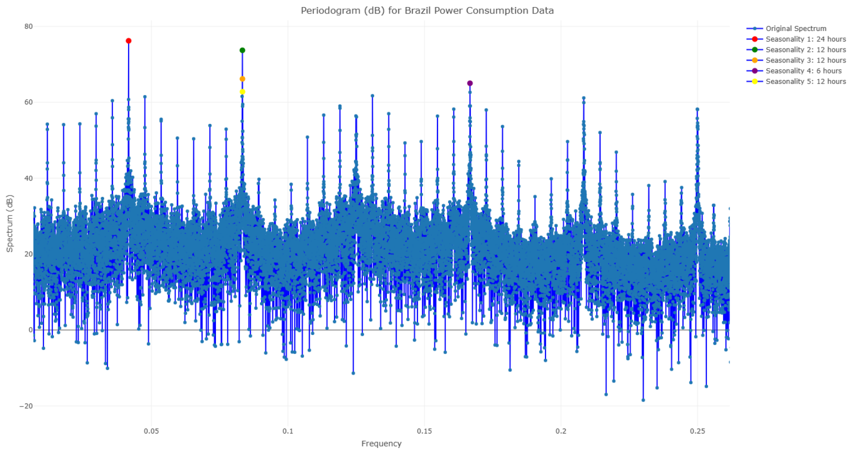

Figure 5: Top 5 dominant frequencies

cally, the model becomes better at handling complex seasonality. We use the residual ARMA adjustment to account for residual autocorrelation after we correct for non-linearity with the Box-Cox transformation.

Table 3 shows the TBATS model paramerter estimations and their output using R.

The TBATS model is a valuable tool for forecasting Brazil's hourly energy consumption [7]. The fitted model, TBATS[1, (3, 4), 0.829, (6, 2), (12, 1), (24, 1)], details of which are in Table 3 and Figure 6, shows strong predictive power. It features a dampened trend (phi=0.829) and uses Fourier terms to model the 6, 12, and 24-hour seasonal periods. With an AIC of 9977.474, this model provides practical insights for energy consumption forecasting.

The BATS model integrates Box-Cox transformation, ARMA errors, and exponential smoothing for forecasting [10, 11]. While

Table 3: TBATS Model Estimated Parameters

Parameter	**Value**
Model Specification	TBATS(1, {3,4}, 0.829, {(6,2), (12,1), (24,1)})
Alpha	−0.11546
Beta	0.02284
Damping Parameter	0.82868
Gamma-1 Values	0.00083, 0.00034, 0.00014
Gamma-2 Values	−0.00136, 0.00035, 0.00063
AR coefficients	−0.22261, −0.30933, 0.37104
MA coefficients	0.90512, 0.64211, −0.15285, −0.31634
Sigma	28.6432
AIC	9977.474

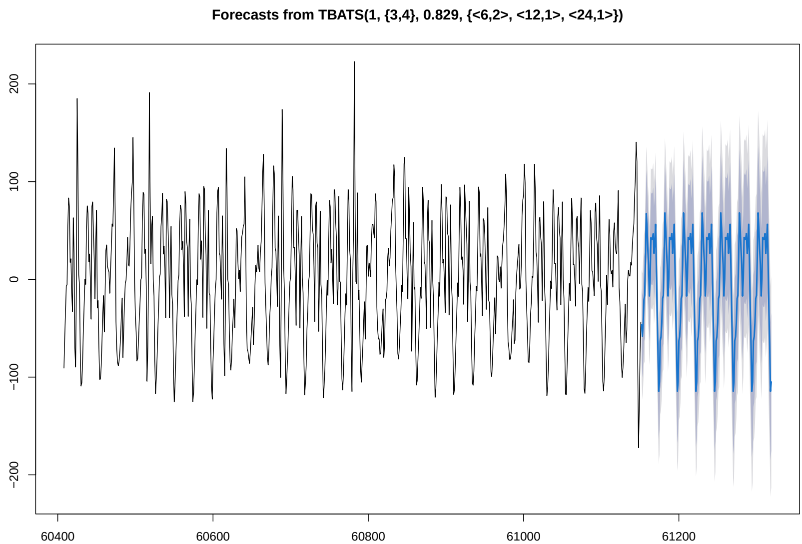

Figure 6: Hourly electricity power consumption forecasting by TBATS Model

it can struggle with complex seasonality, it often outperforms simpler state-space models [17]. For Brazil's hourly energy consumption, the fitted model BATS (1, (5,0), 0.998, (6, 12, 24)) [7], detailed in Table 4 and Figure 7, achieved an AIC of 9779.919, demonstrating its practical application.

Table 4: BATS Model Estimated Parameters

Parameter	**Value**
Model Specification	BATS(1, {5,0}, 0.998, {6,12,24})
Alpha	0.02925
Beta	0.00005
Damping Parameter	0.99835
Gamma Values	0.05369, 0.00622, −0.01774
AR coefficients	0.82257, −0.29863, 0.00289, −0.09035, 0.17738
Sigma	24.13716
AIC	9779.919

The STL decomposition model [9], known for its robust seasonal, trend, and residual filtering [23], was applied to Brazil's hourly energy data. Using the most recent 744 observations and seasonal periods of 6, 12, and 24 hours, the resulting STL + ETS(A,N,N) model produced the forecasts [7] shown in Figure 8.

Table 5: Performance Metrics Comparison

Model	**MAE**	**MPE**	**MAPE**	**Theil's U**
TBATS_forecast	22.27283	163.60077	266.3995	0.62375
BATS_forecast	19.66056	-127.02622	280.3458	0.99675
STL_forecast	19.54752	-86.53971	248.1057	0.44585

Table 5 compares the STL, BATS, and TBATS models to the Brazilian electric power hourly consumption data. we evaluated

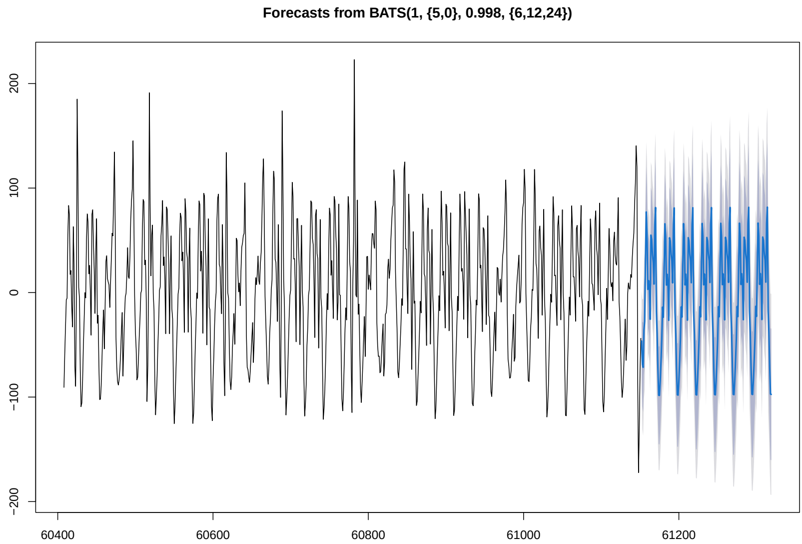

Figure 7: Hourly electricity power consumption forecasting by BATS Model

these combination models by focusing on those with the lowest Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) values. Since MAPE is straightforward, we use it to evaluate predicting strategies using available data. The least significant MAPE shows that STL models outperform the BATS and TBATS models. TBATS is the second best model. In the previous sections, we demonstrated how the: BATS, TBATS, and STL models' fitted models forecasted their individual results. Figure 9 illustrates the forecast performance and compares the models available for different seasonal data sets in Brazil. This study employs a number of STL models, as well

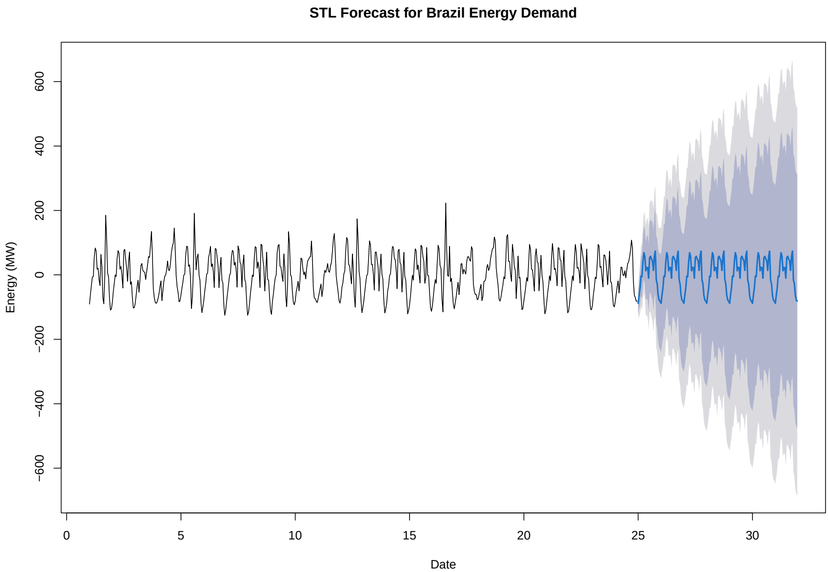

Figure 8: Hourly electricity power consumption forecasting by STL Model

as BATS and TBATS, to examine the multi-seasonal data sets. In projecting the different seasonal data, the STL (cyan color) models perform better than the BATS (olive color) and TBATS (purple color) models [7], as shown in Figure 9 and Table 5. Thesis three introduced a spectral density analysis (periodogram) to detect multiple seasonalities: daily, half-daily, and sub-daily in Brazil's hourly energy consumption. It compared the forecasting performance of TBATS, BATS, and STL models on this multi-seasonal data. Results showed that the STL+ETS(A,N,N) model outperformed the others, achieving the lowest error metrics. The thesis concludes that STL models are superior for forecasting this multi-seasonal time series based on the peri-

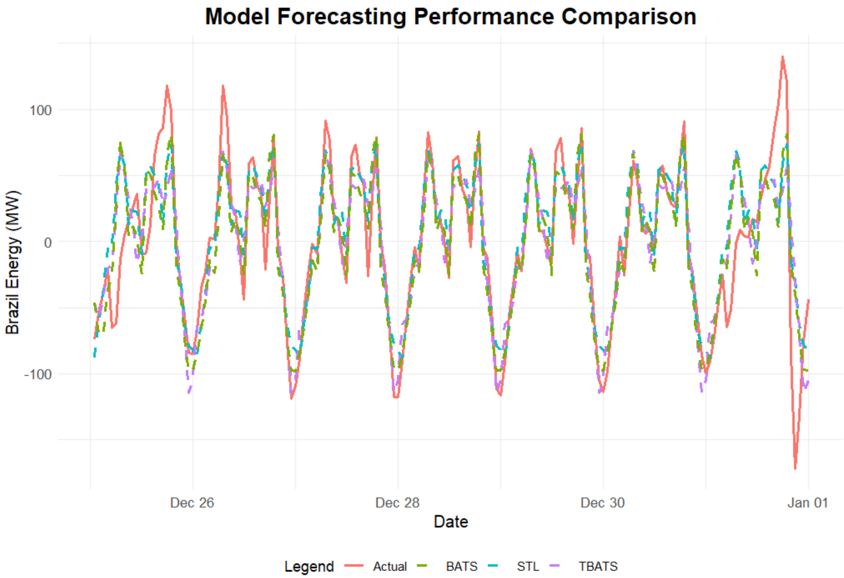

Figure 9: Comparing Forecasting Performance of BATS, TBATS, and STL Models

odogram technique and provide a crucial tool for accurate business planning and operations.

The main publication related to this thesis is [J1].

5 Significance of the Results

The dissertation, a source of these thesis "Time-Series with Multiple Seasonal Periods: Modeling and Forecasting," holds significant scientific and practical importance by addressing critical gaps in forecasting complex seasonal data. Its primary contributions include establishing a unified theoretical framework through state-space representations of core models and introducing key methodological advances: a modified DSHW model with ARMA error correction for enhanced precision, and demonstration of spectral analysis-guided STL+ETS superiority over established methods like TBATS. The research achieved substantial real-world impact through its Seasonal ARIMA model that accurately forecasted COVID-19 mortality in Hungary during the pandemic crisis. By bridging theoretical time series analysis with practical forecasting needs, this work provides essential tools for critical sectors including public health and energy management, enabling more reliable decision-making in data-driven environments.

6 List of Publications

J1. Chudo, S. B., and Terdik, G. (2025). Modeling and Forecasting Time-Series Data with Multiple Seasonal Periods Using Periodograms. (**Econometrics, 13(2), 14. https://doi.org/10.3390/econometrics13020014, Q2**)

J2. Chudo, S. B.” State Space ARIMA Model Formulation and Its Application in Energy,” In Proceedings of IEMTRONICS 2025 (978-981-95-0432-9, 646351-1-En, Chapter 7, ’Lecture Notes in Electrical Engineering, Vol. 1468), 2025, In Press, **Springer Nature, DOI: 10.1007/978-981-95-0433-6, Q4**.

J3. Chudo, S. B. ” Modeling and Forecasting Energy Consumption Using DSHW Model with ARMA Errors: The State Space Approach ,” 2025, (**Submitted to Journal**

of Sustainable Energy, Grids and Networks, Q1).

J4. Chudo, S. B., and D. Gebeyehu, " Statistical Analysis of Determinants of Academic Outcomes of Public TVET Students: A Case Study at Dilla and Hawassa TVET Colleges, Ethiopia," 2025, (**Under the Review in the Journal of Technical Education and Training (JTET), Penerbit UTHM, Q3**).

C1. S.B. Chudo, "Multiplicative Seasonal ARIMA Modeling and Forecasting of COVID-19 Daily Deaths in Hungary," 2022 10th International Conference on Bioinformatics and Computational Biology (ICBCB, ISBN: 978-1-6654-5135-2), Hangzhou, China, 2022, pp. 142-147, **IEEE, DOI: 10.1109/ICBCB55259.2022.9802498 (Best paper Award). Scopus Indexed**.

Conference Presentation and Abstract Publication

C2. Solomon Buke Chudo. (2022). Modeling and Forecasting Time Series with Multiple Seasonal Periods. The 2022 IEEE 2nd Conference on Information Technology and Data Science (CITDS-2022), Debrecen, Hungary (Conference presentation).

C3. Solomon Buke Chudo. (2022). Statistical Analysis of Determinants of Academic Outcomes of Public TVET Students: A Case Study at Dilla and Hawassa TVET Colleges, Ethiopia, 9th International Conference on Social

Sciences and Humanities held on March 19-20, 2022, Burdur, Turkey (Conference presentation and Abstract publication,
https://www.ispecongress.org/sosyal-bilimler, pg. 298).

C4. Solomon Buke Chudo. (2024). " Time Series Models with Multiple Seasonal Periods: An Evaluation of Their Forecasting Performance", ICISDM2024 (The 8th International Conference on Information System and Data Mining), Los Angeles, USA, 2024,
(https://www.icisdm.org/ICKMS Program - June 14. PDF, p.12).

MTMT Publication List

UNIVERSITY of DEBRECEN

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

Registry number: DEENK/585/2025.PL
Subject: PhD Publication List

Candidate: Solomon Buke Chudo
Doctoral School: Doctoral School of Informatics
MTMT ID: 10100203

List of publications related to the dissertation

Foreign language scientific articles in international journals (2)

1. **Chudo, S. B.**, Terdik, G.: Modeling and Forecasting Time-Series Data with Multiple Seasonal Periods Using Periodograms.
   Econometrics. 13 (2), 1-19, 2025. EISSN: 2225-1146.
   DOI: http://dx.doi.org/10.3390/econometrics13020014
   IF: 1.4 (2024)
2. **Chudo, S. B.**: State Space ARIMA Model Formulation and Its Application in Energy.
   Lecture Notes in Electrical Engineering (LNEE). [Epub ahead of print] Chapter 7., 1-10, 2025.
   ISSN: 1876-1100.
   DOI: http://dx.doi.org/10.1007/978-981-95-0433-6

Foreign language conference proceedings (1)

3. **Chudo, S. B.**: Multiplicative Seasonal ARIMA Modeling and Forecasting of COVID_19 Daily Deaths in Hungary.
   In: 2022 10th International Conference on Bioinformatics and Computational Biology (ICBCB 2022) /IEEE (ed.), IEEE, Hangzhou, China, 142-147, 2022. ISBN: 9781665451352

Foreign language abstracts (1)

4. **Chudo, S. B.**: Time Series Models with Multiple Seasonal Periods: An Evaluation of their Forecasting Performance.
   In: 8th International Conference on Information System and Data Mining ICISDM 2024 - 7th International Conference on Knowledge Management Systems ICKMS 2024 / Association for Computing Machinery (publ.), Association for Computing Machinery, New York, 12-12, 2024.
   ISBN: 9798400717345

UNIVERSITY of DEBRECEN

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

List of other publications

Foreign language abstracts (1)

5. **Chudo, S. B.**: Statistical analysis of determinants of academic outcomes of public TVET students: A Case Study at Dilla and Hawassa TVET Colleges, Ethiopia.
   In: 9th International Conference on Social Sciences & Humanities : The proceedings book. Ed.: Ihsan Çapcioğlu, Journal ISPEC Publications, Burdur, 298-298, 2022. ISBN: 9786257720779

Total IF of journals (all publications): 1,4
Total IF of journals (publications related to the dissertation): 1,4

The Candidate's publication data submitted to the Tudóstér have been validated by DEENK on the basis of the Journal Citation Report (Impact Factor) database.

12 November, 2025

References

[1] André Bauer, Marwin Züfle, Johannes Grohmann, Norbert Schmitt, Nikolas Herbst, and Samuel Kounev. An automated forecasting framework based on method recommendation for seasonal time series. In *Proceedings of the ACM/SPEC International Conference on Performance Engineering*, pages 48–55, 2020.

[2] George Box. Box and jenkins: time series analysis, forecasting and control. In *A Very British Affair: Six Britons and the Development of Time Series Analysis During the 20th Century*, pages 161–215. Springer, 2013.

[3] George EP Box. G. m. jenkins. *Time series analysis: forecasting and control*, 1970.

[4] George EP Box, Gwilym M Jenkins, Gregory C Reinsel, and Greta M Ljung. *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.

[5] S. B. Chudo. State space ARIMA model formulation and its application in energy. In *Proceedings of the International IOT, Electronics and Mechatronics Conference (IEMTRONICS 2025)*, Lecture Notes in Electrical Engineering, Imperial College London, UK, 2025. Springer. In press, awarded. To be published in `https://www.springer.com/series/7818`.

[6] Solomon Buke Chudo. Multiplicative seasonal arima modeling and forecasting of covid_19 daily deaths in hungary. In

2022 10th International Conference on Bioinformatics and Computational Biology (ICBCB), pages 142–147. IEEE, 2022.

[7] Solomon Buke Chudo and Gyorgy Terdik. Modeling and forecasting time-series data with multiple seasonal periods using periodograms. *Econometrics*, 13(2):14, 2025.

[8] Robert B Cleveland and William S Cleveland. E mcrae, j., and terpenning, i.: Stl: A seasonal-trend decomposition procedure based on loess. *J. Offic. Stat*, 6:3–33, 1990.

[9] Robert B Cleveland, William S Cleveland, Jean E McRae, Irma Terpenning, et al. Stl: A seasonal-trend decomposition. *J. off. Stat*, 6(1):3–73, 1990.

[10] Alysha M De Livera et al. Automatic forecasting with a modified exponential smoothing state space framework. *Monash Econometrics and Business Statistics Working Papers*, 10(10):6, 2010.

[11] Alysha M De Livera, Rob J Hyndman, and Ralph D Snyder. Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American statistical association*, 106(496):1513–1527, 2011.

[12] Ivo D Dinov. Data science and predictive analytics. *Cham, Switzerland*, 2018.

[13] James Durbin and Siem Jan Koopman. *Time series analysis by state space methods*. Oxford University Press (UK), 2012.

[14] Hans Gilgen. *Univariate time series in geosciences.* Springer, 2006.

[15] Frank E Harrell and FE Harrell. *Regression modeling strategies*, volume 54. Springer, 2001.

[16] Andrew C Harvey. Forecasting, structural time series models and the kalman filter. 1990.

[17] Zachary C Herbert, Zeeshan Asghar, and Carlos A Oroza. Long-term reservoir inflow forecasts: enhanced water supply and inflow volume accuracy using deep learning. *Journal of Hydrology*, 601:126676, 2021.

[18] RJ Hyndman and G Athanasopoulos. Forecasting: Principles and practice, 3rd edn, otexts: Melbourne. *OTexts.com/fpp3. Accessed on July, 07th*, 2021.

[19] Rob J Hyndman and Yeasmin Khandakar. Automatic time series forecasting: the forecast package for r. *Journal of statistical software*, 27:1–22, 2008.

[20] Rob J Hyndman and Anne B Koehler. Another look at measures of forecast accuracy. *International journal of forecasting*, 22(4):679–688, 2006.

[21] Mahdi Kalantari. Forecasting covid-19 pandemic using optimal singular spectrum analysis. *Chaos, Solitons & Fractals*, 142:110547, 2021.

[22] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. *Journal of Fluids Engineering*, 1960.

[23] Yakup Tarik Kranda and Ruya Samli. A novel clustering based algorithm to mitigate the demand of forecasting errors for newly deployed lte cells with insufficient historical data. *Computer Communications*, 190:190–200, 2022.

[24] Our World in Data. Coronavirus pandemic (covid-19) – hungary, 2021.

[25] Donald B Percival and Andrew T Walden. *Wavelet methods for time series analysis*, volume 4. Cambridge university press, 2000.

[26] Robert H Shumway and David S Stoffer. Time series analysis and applications, 2016.

[27] Robert H Shumway and David S Stoffer. State space models. In *Time series analysis and its applications: with R examples*, pages 289–384. Springer, 2017.

[28] Robert H Shumway, David S Stoffer, and David S Stoffer. *Time series analysis and its applications*, volume 3. Springer, 2000.

[29] Aruã Souza. 23 years of hourly electric energy demand (brazil). Retrieved from Kaggle, 2022. Accessed: [Insert Date].

[30] James W Taylor. Short-term electricity demand forecasting using double seasonal exponential smoothing. *Journal of the Operational Research Society*, 54(8):799–805, 2003.

[31] Ashton T Williams, Ryan E Sperl, and Soon M Chung. Anomaly detection in multi-seasonal time series data. *IEEE Access*, 11:106456–106464, 2023.