Comparative Analysis of ARIMA and BSTS Models for Electricity Price Forecasting in Major West Coast Cities

Ziyang Zhang*

Department of Mathematics, University of California, Santa Barbara, United States

*Corresponding author: zhang212@ucsb.edu

Abstract. This study presents a detailed comparative analysis of two prominent statistical models, the AutoRegressive Integrated Moving Average (ARIMA) and Bayesian Structural Time Series (BSTS), in the context of forecasting electricity prices in major West Coast cities of the United States, namely Los Angeles, San Francisco, and Seattle. Utilizing historical electricity price data obtained from the Bureau of Labor Statistics, the research meticulously trains and tests both models, aiming to evaluate their predictive accuracy and reliability in dynamic urban energy markets. The core objective of this research is to ascertain the effectiveness of ARIMA and BSTS models in the realm of electricity price forecasting, with a particular emphasis on their applicability in urban settings characterized by rapid market evolution. Through this analysis, the study offers valuable insights that could significantly influence energy market forecasting and policy formulation. The findings are particularly relevant for stakeholders in the energy sector, including policymakers and market strategists, as they navigate the complexities and challenges of modern energy landscapes. This research contributes to a more comprehensive understanding of statistical modeling techniques in energy market analysis and provides a foundation for more informed decision-making in energy policy and strategy development.

Keywords: ARIMA model; BSTS model; electricity price forecasting.

1. Introduction

Forecasting electricity prices in urban centers such as Los Angeles, San Francisco, and Seattle is pivotal in understanding and navigating the multifaceted dynamics of energy policies, market structures, and technological advancements. Accurate electricity price forecasting is critical for various stakeholders, impacting the expenses of energy-intensive industries, adjusting profits for electricity retailers, and influencing the strategic implementation of a country's national energy strategy.

The transformation of the energy sector from a monopolistic and government-controlled entity to a deregulated market has introduced complexities due to the growing installed power of renewable energy resources, the introduction of smart grids, and the increasing electricity system complexity. These changes necessitate robust and adaptive forecasting models to ensure efficient energy system planning and operations. This need is highlighted in the research conducted on the Lithuanian electricity market, emphasizing the significance of forecasting models based on statistical methods and previous variable values in adjusting national energy strategies [1].

In this context, this study employs the AutoRegressive Integrated Moving Average (ARIMA) and Bayesian Structural Time Series (BSTS) models to forecast electricity prices. The ARIMA model, which has shown promising results in various settings, demonstrates its precision and competitiveness with current forecasting techniques. Concurrently, the BSTS model, known for its adaptability and integration of external influences, is posited to provide enhanced accuracy in forecasting, especially in the dynamically changing urban energy markets. This approach aligns with the findings of Elsaraiti et al., who illustrated the effectiveness of the ARIMA model in forecasting electricity consumption in a dynamic market environment [2].

This research specifically focuses on the major West Coast cities of the United States, each with a unique energy landscape. Los Angeles and San Francisco, influenced by California's ambitious renewable energy targets and greenhouse gas reduction commitments, and Seattle, with its significant hydropower resources and commitment to carbon neutrality, present distinct challenges and opportunities for statistical modeling in electricity price forecasting. Through a decade's worth of data analysis, this study seeks to evaluate and compare the forecasting capabilities of ARIMA and BSTS models, offering insights that could guide decision-making processes and contribute to the development of effective energy policies.

2. Methods

2.1. Data Collection and Preprocessing

The study utilized electricity price data from Los Angeles, San Francisco, and Seattle, sourced from the Bureau of Labor Statistics. As shown in Figure 1, this dataset, covering the period from 2017 to 2023, was chosen for its relevance and recency, offering insights into the urban electricity pricing trends [3]. Before analysis, the data was preprocessed to address non-stationarity, a common challenge in time series analysis. This involved first-order differencing, a technique that stabilizes the mean of the time series over time and is essential for models that presume a stationary series [4].

Fig. 1 Monthly electricity price changes in LA, SF, and Seattle from 2013 to 2023 (Photo/Picture credit: Original).

2.2. Model Selection and Implementation

2.2.1 ARIMA model

The ARIMA model, a staple in time series forecasting, excels in predicting linear trends. The model parameters—*p* for autoregressive, *d* for differencing, and *q* for moving average—play crucial roles in its formulation:

$$Y_t = \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} - \theta_1 \epsilon_{t-1} - \ldots - \theta_q \epsilon_{t-q} + \epsilon_t \tag{1}$$

Here, Y_t represents the time series, ϕ are coefficients of the autoregressive terms, θ are coefficients of the moving average terms, and ϵ_t is the error term. This model is known for its flexibility and has been widely applied in energy market studies [5]. The auto. The Arima function in R's 'forecast' package was employed to automatically determine the optimal parameter based on the Akaike Information Criterion, ensuring the model's appropriateness for the dataset [6].

2.2.2 BSTS model

The Bayesian Structural Time Series (BSTS) model is renowned for its ability to handle structural changes and include external covariates. Its composition involves several components:

$$Y_t = \mu_t + S_t + \beta X_t + \epsilon_t \tag{2}$$

where Y_t is the observed time series, μ_t denotes the trend component, S_t represents the seasonal component, and βX_t is regression components with covariates. This model is particularly effective in dynamic markets like urban energy sectors due to its adaptability [7]. The BSTS model was implemented using the ‘bsts’ package in R. It included ‘AddLocalLinearTrend’ to model the trend and ‘AddSeasonal’ for seasonal effects. Parameters like ‘niter’ were carefully selected to balance computational efficiency with model accuracy [8].

2.3. Training and Testing

Following standard practice in time series forecasting, the dataset was split into a 70% training set and a 30% testing set. This split not only aligns with the methodological rigor required in predictive modeling but also allows for an extensive evaluation of the models’ performance on recent, unseen data [9].

2.4. Evaluation Metrics

RMSE provides a measure of the average magnitude of the forecast errors, while MAPE offers a perspective on the relative accuracy, making them complementary metrics in forecasting analysis. These metrics are critical in comparing the models’ effectiveness, especially in contexts with variable data scales [10]. This combination of metrics allows for a comprehensive assessment of the models’ performance. As highlighted by Bernardi and Petrella, the use of such metrics is vital for evaluating the quality of forecasts, especially when dealing with complex seasonal data [9].

3. Results

3.1. Analytical Outcomes from ARIMA and BSTS Model Implementations

Following the methodological framework outlined earlier, this section delineates the analytical outcomes derived from the application of ARIMA and BSTS models to the decade-spanning electricity price data for Los Angeles, San Francisco, and Seattle.

3.1.1 ARIMA Model Analysis

Los Angeles: Implementing the ARIMA model, as structured by the predefined parameters (p, d, q), the Los Angeles dataset yielded a Mean Error (ME) of 0.0293, Root Mean Squared Error (RMSE) of 0.0306, and Mean Absolute Percentage Error (MAPE) of 11.18%. The forecast values, with comparison to actual data, are shown in Figure 2. The diagnostic checks, specifically the residual analysis, indicated an absence of significant autocorrelation, thus affirming the model’s adequacy in capturing the city’s electricity price trends.

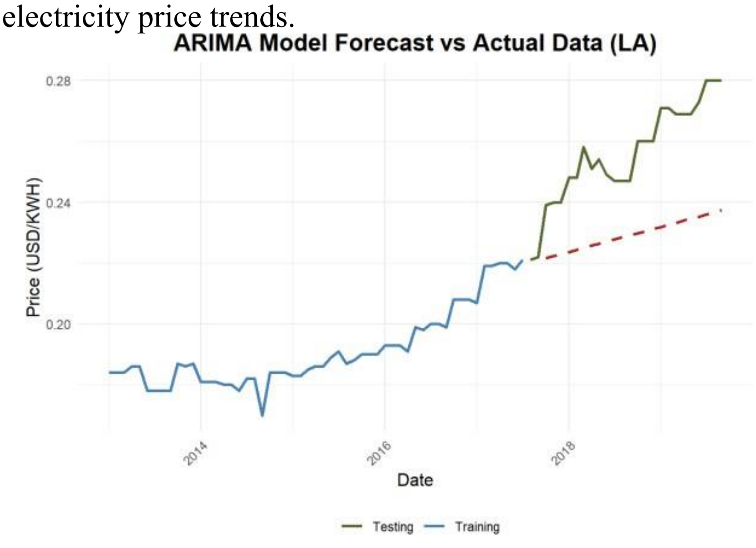

Fig. 2 ARIMA model forecast of LA electricity price (Photo/Picture credit: Original).

San Francisco: In San Francisco, the ARIMA model resulted in an ME of 0.0299, RMSE of 0.0351, and MAPE of 9.91%. The Ljung-Box test, aligning with the methodological rigor, exhibited a p-value of 0.525, thereby validating the model’s fit by confirming the independence of residuals. And the forecast values, with comparison to actual data, are shown in Figure 3.

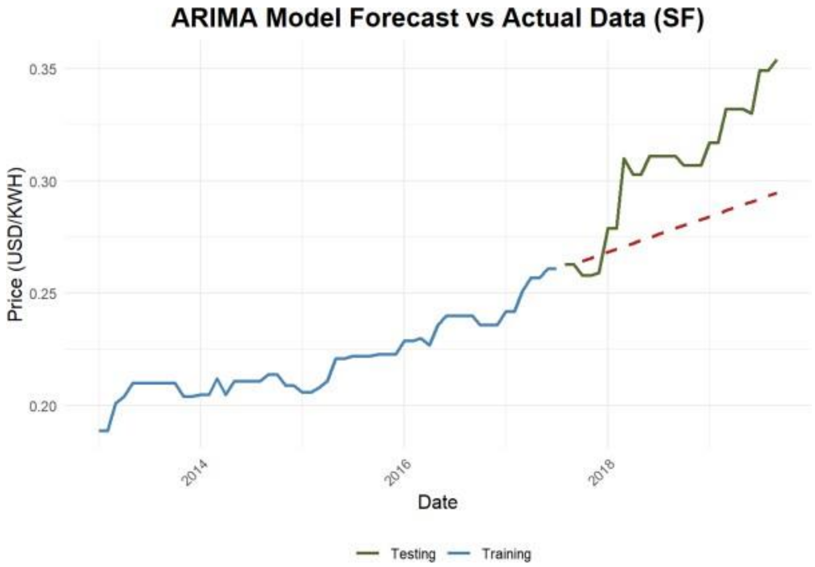

Fig. 3 ARIMA model forecast of SF electricity price (Photo/Picture credit: Original).

Seattle: The model's execution for Seattle's data revealed a notably higher accuracy, as shown in Figure 4, with an ME of 0.0040, RMSE of 0.0060, and a MAPE of 3.28%. Consistent with the methodological expectations, the Ljung-Box test p-value of 0.993 substantiated the model's suitability, as evidenced by the independent distribution of residuals.

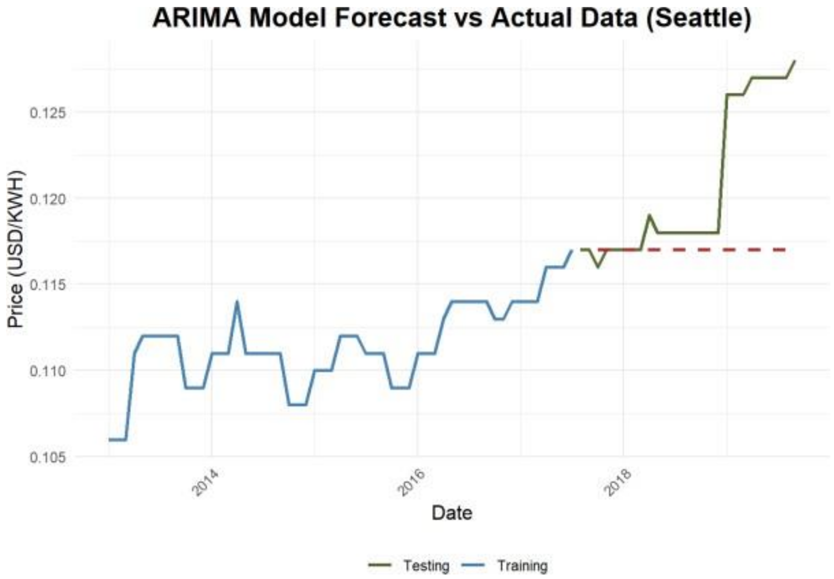

Fig. 4 ARIMA model forecast of Seattle electricity price (Photo/Picture credit: Original).

3.1.2 BSTS Model Evaluation

Adhering to the systematic approach for BSTS model application, the results underscored the model's enhanced ability to accommodate the intricate dynamics of urban electricity markets.

Los Angeles: The BSTS model, when applied to the Los Angeles dataset, demonstrated a significant refinement in forecasting accuracy, as shown in Figure 5, with an ME of 0.0125, RMSE of 0.0133, and a MAPE of 4.78%. These metrics not only surpass the ARIMA model's performance but also highlight the BSTS model's advanced adaptability to the city's energy market.

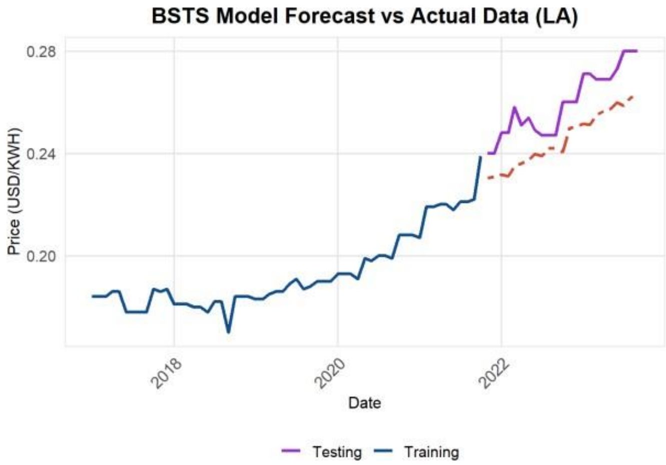

Fig. 5 BSTS model forecast of LA electricity price (Photo/Picture credit: Original).

San Francisco: For San Francisco, the BSTS model furnished an ME of 0.0277, RMSE of 0.0313, and MAPE of 8.92%. This improved precision, particularly in relative error assessment, reinforces the model's competency in capturing the city's specific market fluctuations. And the forecast values, with comparison to actual data, are shown in Figure 6.

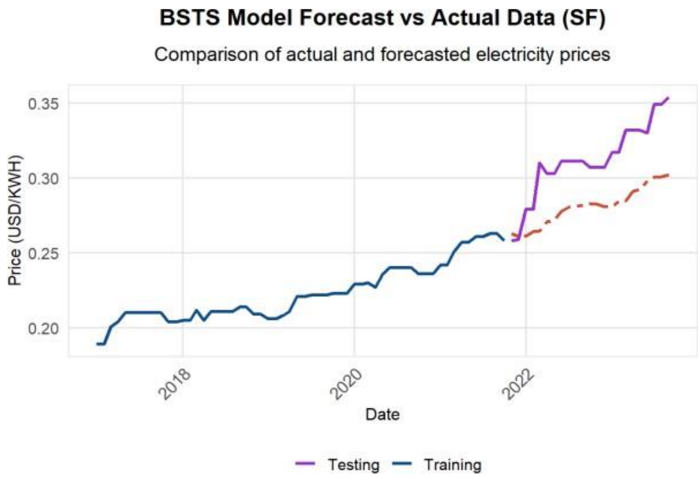

Fig. 6 BSTS model forecast of SF electricity price (Photo/Picture credit: Original).

Seattle: The BSTS model's efficacy was most pronounced in Seattle, as shown in Figure 7, with an ME of approximately 0.00004, RMSE of 0.0023, and a MAPE of 1.71%. These results are indicative of the model's exceptional capability in accurately forecasting within a market predominantly influenced by hydropower.

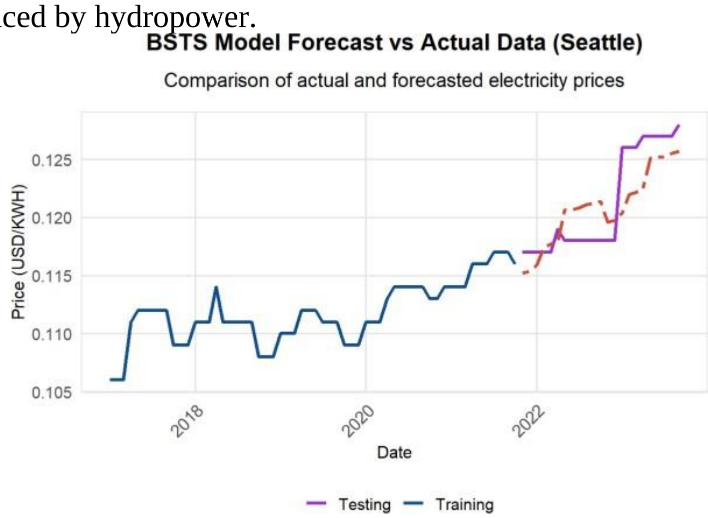

Fig. 7 BSTS model forecast of Seattle electricity price (Photo/Picture credit: Original).

3.2. Comparative and Contextual Analysis

In line with the methodological premise of this study, a comparative analysis between the ARIMA and BSTS models elucidates the latter's superior performance across all cities and metrics. This consistency in outperforming the ARIMA model can be attributed to the BSTS model's sophisticated integration of local trends, seasonal components, and external factors, as outlined in the method section. Such attributes are crucial in dynamic urban settings, characterized by diverse energy policies and renewable energy adoption rates.

3.3. Policy and Strategic Implications

Drawing from the analytical outcomes, this study's findings offer pivotal insights for stakeholders in the energy sector. The demonstrable efficacy of BSTS models in forecasting electricity prices advocates for their preferential adoption in urban contexts with intricate and evolving energy frameworks. For policymakers, market strategists, and urban planners, leveraging the predictive prowess of BSTS models could lead to more informed, accurate, and reliable decision-making processes, thereby facilitating the development of robust and effective energy policies tailored to each city's unique energy landscape.

4. Discussion

This study's exploration of ARIMA and BSTS models for forecasting electricity prices in major West Coast cities reveals significant theoretical and practical insights.

4.1. Theoretical Implications

The BSTS model's superior performance, especially in its univariate form, echoes the findings of contemporary time series forecasting literature. BSTS models typically demonstrate a Mean Absolute Percentage Error (MAPE) below 10%, highlighting their high accuracy and reliability in forecasting [10]. This efficiency in handling complex markets like electricity, where adaptability to structural changes is crucial, underscores the need for robust modeling approaches [6].

4.2. Practical Implications

For policymakers and market strategists, these results underscore the importance of selecting appropriate forecasting models. The ability of the BSTS model to outperform the ARIMA model, particularly in Seattle's hydropower-influenced energy market, signals a need for re-evaluation of current forecasting tools, considering evolving market dynamics and policy environments [7].

4.3. Methodological Considerations

The methodological approach of this study, examining both ARIMA and BSTS models across various urban contexts, offers a comprehensive understanding of these models' applicability. This aligns with the current trend in energy market research, where diverse methodologies are employed to capture market nuances [9]. This approach is particularly pertinent in light of recent global events like the COVID-19 pandemic, which have significantly impacted energy markets and consumption patterns.

4.4. Future Research Directions

Future research could explore incorporating additional variables into the BSTS model and assessing their impact on forecasting accuracy. Although the univariate BSTS model often yields accurate predictions, investigating the effects of adding socioeconomic covariates could provide deeper insights into urban electricity pricing dynamics [10].

5. Conclusion

In conclusion, this study's comprehensive analysis of the ARIMA and BSTS models in forecasting electricity prices has provided substantial contributions to both theoretical knowledge and practical applications in the field of statistical modeling for energy markets. The research not only reaffirms the importance and effectiveness of these models in urban energy market forecasting but also highlights the particular effectiveness of the BSTS model. Its clear superiority, especially in its more straightforward form, has been consistently demonstrated across various urban settings including Los Angeles, San Francisco, and Seattle. These findings are vital for policymakers and market strategists, providing them with a robust tool to better understand and navigate the complexities inherent in modern energy landscapes. The implications of this research are far-reaching, offering guidance that could significantly influence decision-making processes and contribute to the development of more effective and tailored energy policies, resonating with the current global shift towards more dynamic and sustainable energy solutions.

References

[1] Cˇesnaviˇcius M. Lithuanian electricity market price forecasting model based on univariate time series analysis. Energetika (Vilnius, Lithuania), 2020, 66(1).

[2] Elsaraiti M, Ali G, Musbah H, Merabet A, Little T. Time series analysis of electricity consumption forecasting using ARIMA model. 2021 IEEE Green Technologies Conference (GreenTech), 2021, 259-262.

[3] Xie L. The analysis and forecasting of COVID-19 cases in the United States using Bayesian structural time series models. Biostatistics & Epidemiology (Oxford, UK), 2022, 6(1), 1-15.

[4] Kalinina I, Bidyuk P, Gozhyj A. Construction of forecast models based on bayesian structural time series. 2022 IEEE 17th International Conference on Computer Sciences and Information Technologies (CSIT), Lviv, Ukraine, 2022, 180-184.

[5] Carriere T, Kariniotakis G. An integrated approach for ValueOriented energy forecasting and Data-Driven Decision-Making application to renewable energy trading. IEEE Transactions on Smart Grid, 2019, 10(6), 6933-6944.

[6] Conejo A J, Plazas M A, Espinola R, Molina A B. Dayahead electricity price forecasting using the wavelet transform and ARIMA models. IEEE Transactions on Power Systems, 2005, 20(2), 1035–1042.

[7] Bernardi M, Petrella L. Multiple seasonal cycles forecasting model: the Italian electricity demand. Statistical Methods & Applications, 2015, 24(4), 671-695.

[8] Rodr´ıguez Y, D´ıaz Olariaga O. Air traffic demand forecasting with a Bayesian structural time series approach. Transportation Engineering, 2023.

[9] Cuaresma J C, Hlouskova J, Kossmeier S, et al. Forecasting electricity spot-prices using linear univariate time-series models. Applied Energy, 2004, 77(1): 87-106.

[10] Pavlyshenko B. Bayesian regression approach for building and stacking predictive models in time series analytics. Data Stream Mining & Processing: Third International Conference, DSMP 2020, Lviv, Ukraine, August 21–25, 2020, Proceedings 3. Springer International Publishing, 2020: 486-500.