"Ovidius" University Annals, Economic Sciences Series
Volume XXV, Issue 1 /2025

Large Language Models and Transformer Architecture in Electricity Price Forecasting

Adela Bâra
Simona-Vasilica Oprea
Bucharest University of Economic Studies,
Department of Economic Informatics and Cybernetics, Romania
bara.adela@ie.ase.ro
simona.oprea@csie.ase.ro

Abstract

Our research explores the application of Large Language Models (LLMs) and transformer-based architectures to the task of forecasting electricity prices in Romania's Day-Ahead Market (DAM) and Intraday Continuous Market (IDC). Using Amazon Chronos, a deep learning framework built on the T5 transformer architecture, we demonstrate the model's ability to capture complex temporal patterns, nonlinear dynamics and seasonal variations in electricity price time series. The research utilizes hourly price data from June 2024 to February 2025 and evaluates performance using standard metrics such as MAE, RMSE, MSE, MAPE and R^2. Chronos models, particularly Chronos T5 variants, are shown to produce highly accurate forecasts and well-calibrated prediction intervals, even in volatile market conditions. The self-attention mechanism and temporal embeddings enable effective learning of long-range dependencies, while quantile regression supports probabilistic forecasting. Results indicate that Chronos achieves accuracy in DAM forecasting and performs competitively in the more volatile IDC, providing both point estimates and uncertainty ranges.

Key words: price forecast, electricity markets, large language models, transformer architecture
J.E.L. classification: C53, P18, D47

1. Introduction

Accurate forecasting of electricity prices in short-term markets is essential for energy trading, grid operations and the deployment of flexible assets such as storage systems (De la Torre et al., 2024), (Zhang and Fu, 2024). In this research, we apply Amazon Chronos, a deep learning time series forecasting framework accessed through the ChronosPipeline interface, to model and predict electricity prices on both the Day-Ahead Market (DAM) and the Intraday Continuous Market (IDC) in Romania (Bâra and Oprea, 2024), (Bâra, Oprea and Tudorică, 2023). The Chronos models effectively capture the temporal dynamics of electricity prices in both DAM and IDC. Forecasts closely track actual price trends, including volatile fluctuations and daily seasonality patterns. The analyzed dataset includes hourly electricity prices from the DAM and IDC over a multi-month period (Jun. 2024-Feb. 2025). The methodology consists of selecting appropriate model presets (chronos_tiny for CPU or chronos_small for GPU), training Chronos models (Salahuddin et al., 2022) with a prediction horizon of 24 hours and evaluating results using standard performance metrics: MSE, RMSE, MAE, MAPE and R^2.

Built on the foundation of transformer architectures (Arroni et al., 2023), Chronos models are specifically designed to handle the complexities of temporal data, offering accurate forecasts across various domains such as energy, finance and retail. The core architecture leverages self-attention mechanisms, allowing the model to learn dependencies across different time steps and effectively capture long-range patterns, seasonal variations and trends in the data. One of the distinctive features of Chronos is its use of temporal embeddings, which help encode information related to time, such as hour of the day or day of the week, into the model. This temporal context improves the model's ability to detect cyclical behaviors and seasonal shifts. Chronos also supports probabilistic forecasting through quantile regression, which enables it to provide uncertainty estimates in the form of prediction intervals. These intervals are particularly useful in applications where understanding the range of possible outcomes is as important as the forecast itself.

The Chronos family includes several model variants that differ in size and computational requirements. The smallest variant, chronos_tiny, is optimized for CPU use and provides fast training and inference, making it suitable for lightweight applications or environments with limited computational resources. Chronos_mini offers a balance between speed and accuracy, while chronos_small, chronos_base and chronos_large are deeper models requiring GPU acceleration, each offering progressively higher forecasting accuracy. For scenarios demanding high robustness, Chronos also supports ensemble variants that combine multiple models to improve generalization. Chronos models are effective in scenarios where traditional time series models may fall short, such as in the presence of non-linear dynamics, multiple seasonal patterns or high volatility. In the context of electricity markets, Chronos has been used to forecast prices in both the DAM and the IDC. These markets are characterized by rapid price changes and complex temporal structures, making them ideal candidates for Chronos-based forecasting.

2. Literature review

The authors of (Pinhão, Fonseca and Covas, 2022) introduced a novel methodology by forecasting the underlying market dynamics, specifically, the supply and demand curves derived from auction bids. Addressing the complexity of hourly market data with numerous block bids, the proposed model incorporated multiple seasonal effects and key market variables such as wind generation and load. Results showed that this approach outperforms traditional benchmarks and enhances ensemble model performance, underscoring the value of including bid-level information in electricity price forecasting. Another research introduced a multivariate elastic net model for forecasting German quarter-hourly electricity prices, focusing on the less-studied IDC and call-auction markets (Kath and Ziel, 2018). It also examined the influence of early DAM EXAA prices and links forecast accuracy to economic value, showing that simple trading strategies can yield substantial gains when paired with reliable predictions.

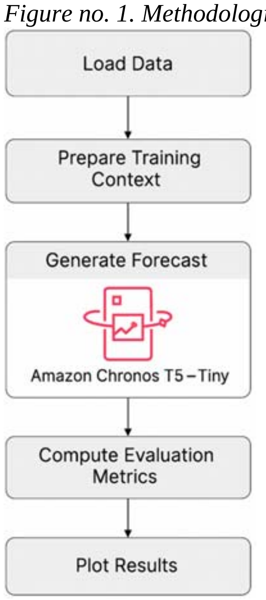
A novel approach to energy forecasting was proposed (Hubicka, Marcjasz and Weron, 2019) demonstrating that averaging DAM electricity price forecasts across multiple calibration windows (ranging from 28 to 728 days) outperforms using a single „optimal" window. Further improvements in accuracy were achieved by averaging forecasts from a few strategically chosen window lengths. Additionally, (Bâra, Oprea and Ciurea, 2024) explored the evolution of electricity prices in Romania's Balancing Market, aiming to improve forecasting in the current economic and geopolitical climate. An AI-driven approach was proposed, combining classifiers and regressors to first predict the imbalance sign, then forecast prices. Two prediction strategies were introduced: one based on averaging five ensemble machine learning models, and another using weighted combinations via linear regression or decision trees. The results provided actionable insights for market participants, supporting the development of optimal bidding strategies through the integration of supervised and unsupervised learning techniques.

Another research presented a short-term electricity price forecasting framework that combines frequency analysis with price spike oversampling (Zhang, Fu and Gong, 2023). Variational Mode Decomposition (VMD) and Extended Discrete Fourier Transform (EDFT) were used to model price dynamics in the frequency domain. To enhance spike prediction, oversampling techniques like ESPO and SMOTE for regression were applied. Results across multiple markets showed the framework accurately forecasts both regular prices and spikes. The impact of rising renewable energy penetration on the variability of day-ahead electricity prices and the need for adaptable forecasting models were further investigated (Lehna, Scheller and Herwartz, 2022). A comparative study was conducted using four forecasting approaches for the German DAM: (S)ARIMA(X), LSTM, CNN-LSTM and a two-stage multivariate VAR model. Incorporating external factors such as load, fuel prices, CO_2 emissions, solar radiation and wind speed, the models were evaluated over multiple forecast horizons. Results showed that LSTM performs best on average, while the two-stage VAR excels in short-term forecasts.

3. Research methodology

When using Amazon Chronos through the ChronosPipeline interface, the process of time series transformation is streamlined for effective forecasting. This transformation involves preparing the input data into a format that the pre-trained Chronos T5 model can understand, along with normalization and structuring that aligns with the model's transformer-based architecture. The core transformation begins with selecting a segment of the historical time series to serve as the training context (as in Figure 1). The Price_DAM values are extracted from the DataFrame and converted into a PyTorch tensor. This tensor, representing a continuous univariate time series, becomes the input context from which the model generates future predictions. The model is not retrained; instead, it performs zero-shot or few-shot inference, using its pre-trained understanding of time series behavior. Internally, the Chronos model expects the input to be a fixed-length sequence. It applies positional encoding to represent the order of time steps and captures temporal patterns through self-attention mechanisms, characteristic of transformer architectures. The forecasting task here is formulated as a sequence continuation problem, where the model generates the next *n* steps (defined by prediction_length) based on the provided context. Importantly, the transformation is more about sequence extraction and structuring than traditional preprocessing like frequency resampling or feature engineering. Unlike the AutoGluon API, the ChronosPipeline expects raw numerical values and does not require explicit timestamp columns or item identifiers. Time is handled implicitly by the model's positional encoding rather than through timestamp-based features. Figure 1 shows the proposed methodological flow.

Figure no. 1. Methodological flow



Source: Authors' contribution

The output of the model is a probabilistic forecast, a distribution over multiple future scenarios. These forecasts are post-processed using numpy.quantile() to extract the 10th, 50th (median) and 90th percentiles, allowing for interval-based forecasting. This is an essential feature of Chronos, as it enables point predictions and also uncertainty quantification, which is especially valuable in volatile domains like electricity markets. Therefore, time-series transformation in this Chronos setup involves selecting and slicing the appropriate historical window, converting it to a numerical tensor and interpreting the model's output in a probabilistic way. The process is minimal and model-centric, relying on Chronos's pre-trained representations and transformer-based architecture to extract and learn time dependencies, rather than on manual feature design or extensive preprocessing.

Large Language Models (LLMs), particularly models like Amazon’s Chronos T5, are applied to numerical time series forecasting, linking the core concepts of language modeling to the domain of temporal prediction. Originally designed for Natural Language Processing (NLP) tasks like translation, summarization and text generation, LLMs have proven surprisingly effective in modeling structured numerical sequences as well, including time series data, by leveraging the same foundational concepts that underpin language understanding. At the core of an LLM like T5 (Text-to-Text Transfer Transformer) lies the transformer architecture, which uses self-attention mechanisms to capture dependencies across sequential data. In natural language, this means learning relationships between words; in time series, it means learning patterns, trends and dependencies across time steps. When adapted for forecasting, time series values are treated similarly to words in a sentence, each point in the series is a "token" with contextual meaning shaped by its position in the overall sequence. Chronos T5 follows a naming convention similar to that of the original T5 models, offering a range of variants as in Table 1 that trade off between computational efficiency and forecasting accuracy.

Table no. 1 Chronos T5 variants

Model Variant	**Parameters (approx.)**	**Description**	**Ideal Use Case**
chronos-t5-nano	~25 million	Extremely lightweight version, very fast and resource-light	Quick tests, edge devices, micro-batches
chronos-t5-tiny	~60 million	Small and efficient; good balance for CPU inference	Short-horizon forecasting on CPUs
chronos-t5-mini	~100 million	Larger context and more accurate forecasts	Standard applications, longer sequences
chronos-t5-small	~250 million	Higher accuracy with broader temporal dynamics	Multi-step, volatile time series
chronos-t5-base	~500 million	Strong baseline model with powerful pattern learning	Energy, finance, and industrial data
chronos-t5-large	~770 million	Most powerful version for highly accurate long-horizon forecasts	GPU-only, high-resource environments
chronos-t5-ensemble	Varies	Ensemble of multiple variants with quantile averaging	Max accuracy, production forecasting

Source: Authors’contribution

In the Chronos T5 model, the numerical time series is tokenized and embedded, allowing the transformer layers to attend over historical values and generate coherent forecasts. This process can be seen as a form of sequence-to-sequence learning, where the “input sentence” is a segment of past observations and the “output sentence” is the model’s prediction of future values. The model learns the local temporal structure and also global patterns such as seasonality and trends, much like how it learns grammatical rules and context in language. One of the key advantages of using LLMs in this setting is their ability to generalize across different types of time series without task-specific tuning. Trained on a large and diverse corpus of temporal data (similar to how LLMs are pretrained on vast text corpora), Chronos T5 develops a form of “temporal literacy”, it learns abstract representations of time and variability that transfer across domains, including energy markets, e-commerce, finance, retail and more.

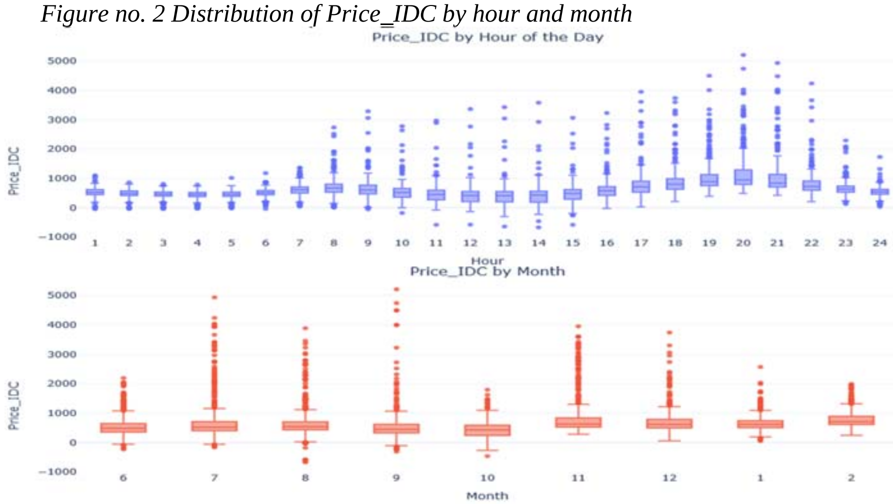
Moreover, the generation process in Chronos T5 is probabilistic, akin to language generation models like GPT. Rather than predicting a single deterministic output, it can generate multiple plausible future sequences from which prediction intervals (e.g., 10th, 50th, and 90th percentiles) can be derived. This is essential in domains like electricity price forecasting, where uncertainty and volatility are inherent to the system. Importantly, the shift from classic statistical models or task-specific deep learning architectures to LLM-based models also changes the way forecasting is approached. Instead of explicitly engineering features (e.g., lag values, holiday flags), the model implicitly learns what matters through attention weights, leveraging its massive parameter space and pretraining to detect subtle dynamics in the data.

Thus, the application of LLM concepts to numerical time series forecasting leverages the same strengths that enhanced NLP: self-attention, transfer learning, generative modeling and scalability. Amazon Chronos T5 treats time series as a structured “language” of numbers, enabling it to generate highly expressive and context-aware forecasts, a significant advancement for complex prediction tasks such as electricity market modeling.

4. Findings

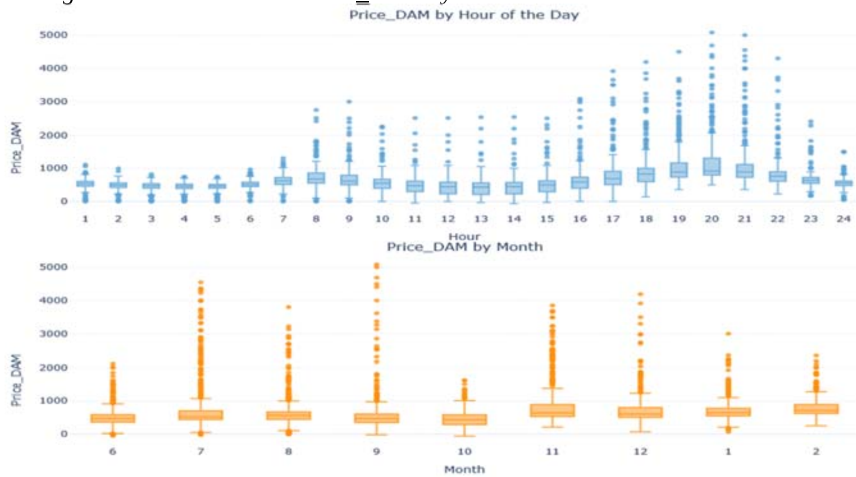
Figure 2 (top) illustrates the distribution of Price_IDC across different hours of the day. It shows a clear pattern where prices start relatively low and stable during the night and early morning hours. As the day progresses, particularly from around 8 AM onward, both the median prices and the variability increase. The highest prices and the widest spreads occur in the evening hours, especially between 6 PM and 9 PM. This suggests that demand spikes during these hours, contributing to price surges and a higher frequency of extreme outliers. After 9 PM, the prices begin to stabilize and gradually decrease toward midnight. Figure 2 (bottom) displays the distribution of Price_IDC across the months. The summer months: June, July, August and September, exhibit the highest volatility and price levels. Figure 3 (top) presents the distribution of Price_DAM across the 24 hours of the day. A distinct pattern emerges where prices are relatively low and stable during the early morning hours, from around midnight to 6 AM. Starting from 7–8 AM, the prices gradually increase, reaching their peak during the evening hours, especially between 6 PM and 9 PM. This is also when the price spread becomes most pronounced, indicating higher volatility and frequent extreme values. The behavior suggests a clear alignment with daily demand cycles, where consumption tends to peak during the evening. Figure 3 (bottom) illustrates the distribution of Price_DAM by month. During the summer months: June through September, there is greater price dispersion and higher median levels, pointing to increased market activity or stress. September shows particularly high volatility with a substantial number of outliers.

Figure no. 2 Distribution of Price_IDC by hour and month



Source: Authors’ contribution

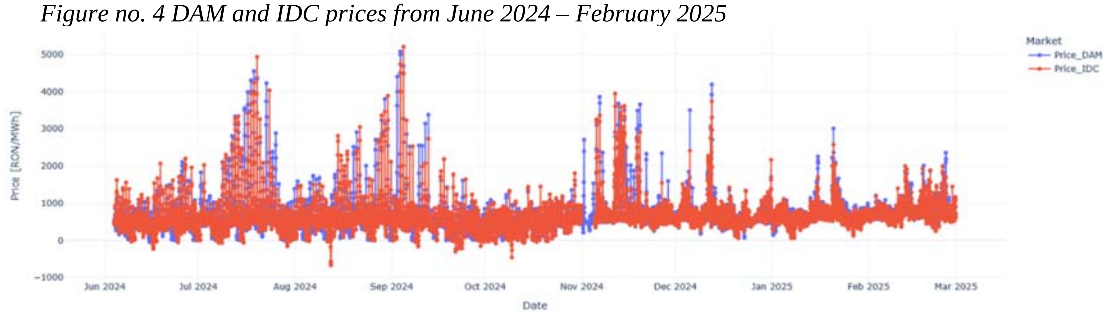
Figure no. 3 Distribution Price_DAM by hour and month



Source: Authors'contribution

Figure 4 shows the time series of Price_DAM, in blue and Price_IDC, in red from June 2024 to Feb. 2025. Both markets follow similar overall trends, with price spikes frequently aligning across the two series. However, Price_IDC appears slightly more volatile, exhibiting sharper fluctuations, especially during high-stress periods like July, August and September 2024. These months show the most extreme peaks, where prices often reach or exceed 4000–5000 RON/MWh. Starting from October 2024, both markets show a visible reduction in volatility and fewer extreme spikes. This calmer behavior continues throughout the winter months. Nonetheless, occasional price jumps are still observed, especially during colder periods (e.g., in December and January), driven by increased heating demand or market imbalances.

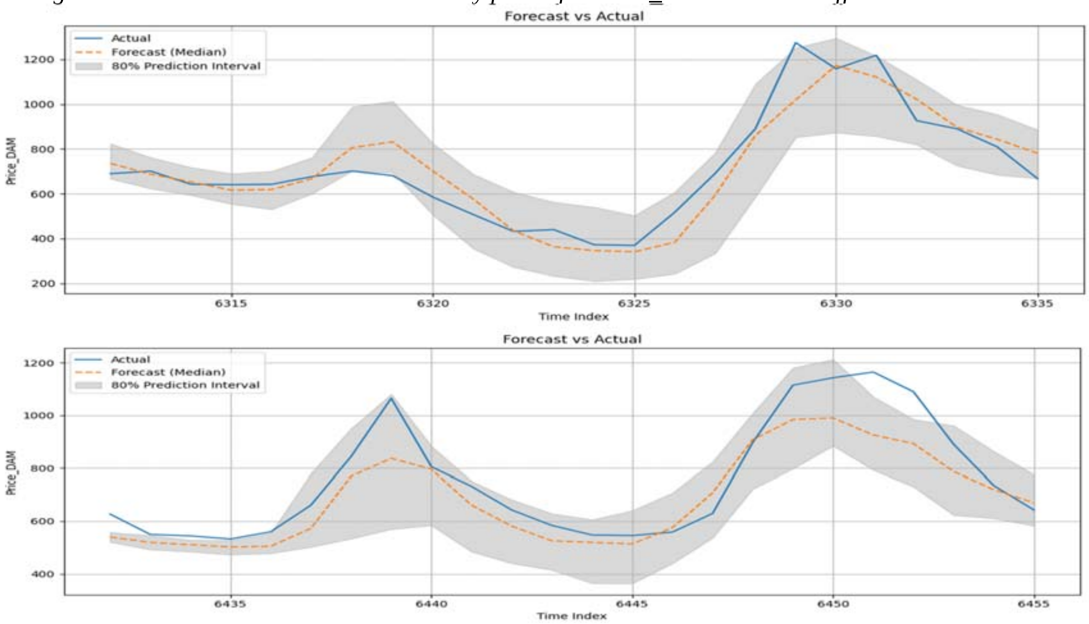
Figure no. 4 DAM and IDC prices from June 2024 – February 2025



Source: Authors' contribution

Figure 5 compare the forecasted vs. actual electricity prices for Price_DAM over two different time windows, using a median forecast with an 80% prediction interval. The solid blue line represents the actual prices, while the dashed orange line shows the forecasted median values. The shaded grey area indicates the 80% prediction interval, reflecting the uncertainty around the forecast. In the upper chart, the forecast captures the general trend quite well, especially the decline around time index 6320 and the sharp rise approaching index 6330. However, the forecast slightly underestimates the peak at index 6329–6331, although the actual values still fall within the prediction interval, indicating reasonable uncertainty bounds. In the lower chart, the model again successfully tracks the upward and downward movements, although it underestimates the magnitude of the peaks, particularly at index 6439 and again near 6450. Nonetheless, most of the actual values remain within the shaded prediction area, suggesting that the model is capturing the overall shape and variability of the price curve, even if it does not always match the precise amplitude. Overall, the forecast model demonstrates strong short-term accuracy in following daily price trends.

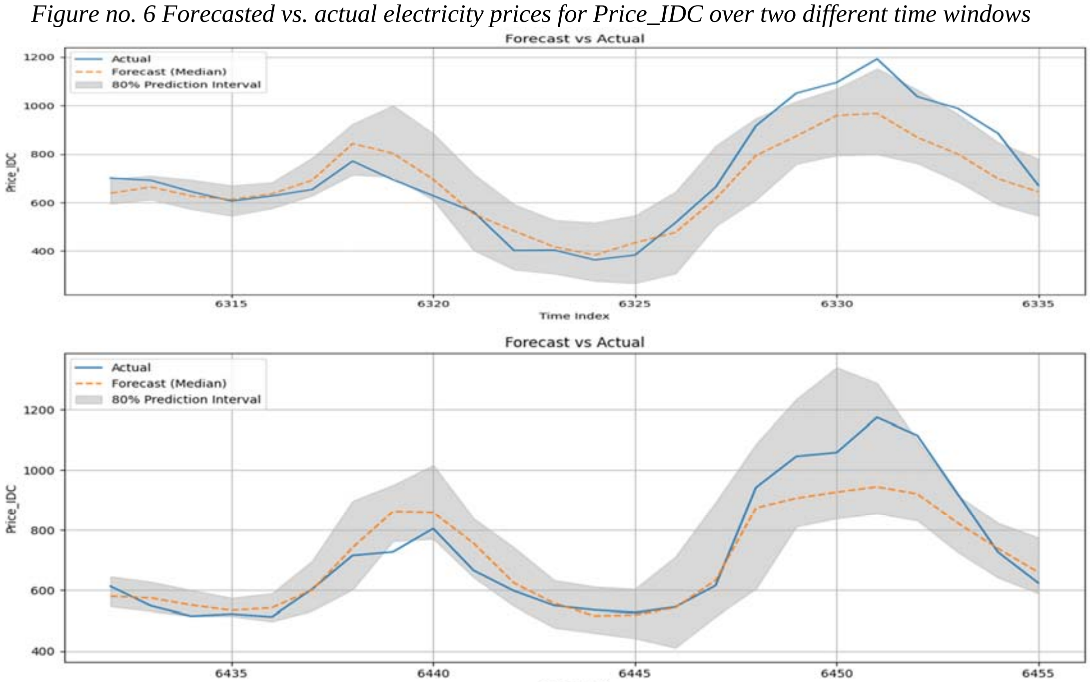
Figure no. 5 Forecasted vs. actual electricity prices for Price_DAM over two different time windows



Source: Authors' contribution

The forecast vs. actual plots for Price_IDC in Figure 6 show that the model captures the general pattern of the actual price quite effectively. In both time windows, the forecast median line follows the ups and downs of the actual prices with reasonable accuracy. The most prominent peaks, particularly around time indices 6330 and 6450, are slightly underestimated by the model, although the actual values still fall within the 80% prediction interval. This suggests that the model accounts for uncertainty well, especially during periods of rapid price change. The prediction interval widens appropriately during more volatile periods, indicating that the model is dynamically adjusting its confidence levels.

Figure no. 6 Forecasted vs. actual electricity prices for Price_IDC over two different time windows



Source: Authors' contribution

The error metrics for DAM and IDC in Table 2 indicate that the forecasting models perform well overall, with some variation across the two index ranges. For the DAM, Index1 (6315-6335) outperforms Index2 (6435-6455). The Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) are both lower in Index1, suggesting more accurate forecasts in terms of both average and standard deviation of errors. Although the Mean Absolute Percentage Error (MAPE) is slightly higher in Index1 at 9.47% compared to 9.33% in Index2, this difference is marginal. The R^2 score is stronger in Index1 at 0.864, indicating that around 86.4% of the variance in DAM prices is explained by the model, which reflects a high level of predictive accuracy. Index2, on the other hand, has a lower R^2 score of 0.7751, showing that the model's ability to capture variability in prices is somewhat reduced during this period, likely due to more unpredictable market behavior. In the case of the IDC, Index2 shows improved performance compared to Index1. The MAE and RMSE are both lower in Index2, with the RMSE dropping to 86.37 from 103.39. The MAPE also shows a notable decrease from 9.99% to 7.09%, indicating that forecasts in Index2 are more accurate in relative percentage terms. Additionally, the R^2 score improves to 0.8259 in Index2, suggesting that the model better captures the variance in IDC prices during this interval. Overall, the models demonstrate strong forecasting capabilities across both markets, with R^2 scores consistently above 0.77. The DAM forecasts are slightly more stable and accurate in Index1, while the IDC forecasts show significant gains in Index2. These results highlight that model performance can vary depending on temporal or market-specific factors.

Table no. 2 Performance metrics for the two-time indexes

	DAM Index1	DAM Index2	IDC Index1	IDC Index2
MAE	65.8337	77.3196	79.1581	59.7805
MSE	7896.998	10296.24	10690.32	7460.52
RMSE	88.8651	101.4704	103.394	86.3743
MAPE	9.47%	9.33%	9.99%	7.09%
R^2 Score	0.864	0.7751	0.798	0.8259

Source: Authors' contribution

The following observations are made: (1) DAM forecasting-performance was robust, especially under stable market conditions, where Chronos provided low MSE and high R^2 values; (2) IDC forecasting-due to higher intra-day volatility, IDC predictions were slightly less accurate, though still within acceptable error margins. Chronos' ability to produce probabilistic forecasts (via quantiles) offers valuable insights into price uncertainty.

5. Conclusions

Based on the provided text, several conclusions can be drawn regarding the use of large language models (LLMs) and transformer architectures, specifically Amazon Chronos T5, for electricity price forecasting in Romania's short-term markets. The Chronos framework demonstrates strong capability in modeling complex temporal patterns within both the Day-Ahead Market (DAM) and the Intraday Continuous Market (IDC). Its transformer-based architecture, which incorporates self-attention mechanisms and temporal embeddings, enables it to effectively capture non-linear dynamics, long-range dependencies and multiple seasonalities in electricity price data. The use of quantile regression for probabilistic forecasting adds further value by providing uncertainty estimates alongside point predictions, a critical feature for highly volatile markets.

The application of Chronos T5 showcases the adaptability of LLMs to numerical time series, leveraging techniques originally developed for natural language processing. The sequence-to-sequence formulation and token-based embedding approach allow the model to treat time series as structured "sentences" of data, enabling generalization across domains and reducing the need for manual feature engineering. This shift from traditional statistical models toward LLM-based forecasting represents a significant advancement in how time series problems are addressed.

Experimental results across two forecasting windows confirm the effectiveness of Chronos in both markets. In the DAM, the model achieved low error rates and high R^2 scores, particularly during stable market periods, indicating strong predictive power. For the IDC, which is inherently more volatile, performance was still robust, with the second index window showing notably improved accuracy and lower forecast errors. These results underscore the importance of probabilistic forecasting, as Chronos was able to provide meaningful prediction intervals even in turbulent conditions.

Finally, the overall forecasting framework is streamlined and computationally efficient, especially when using smaller Chronos T5 variants such as chronos_tiny and chronos_small, which balance inference speed with accuracy. This makes the approach scalable and practical for both research and operational deployment in energy trading, grid optimization and flexibility management. The findings highlight the transformative potential of LLMs in time series forecasting and suggest that transformer-based architectures like Chronos T5 are well-suited for modeling the intricacies of electricity markets.

6. Acknowledgement

This work was supported by a grant of the Ministry of Research, Innovation and Digitization, CNCS/CCCDI - UEFISCDI, project number ERANET-ERAMIN-3-ValorWaste-1, within PNCDI IV.

7. References

- Arroni, S., Galán, Y., Guzmán-Guzmán, X., Núñez-Valdez, E.R. and Gómez, A., 2023. Sentiment Analysis and Classification of Hotel Opinions in Twitter With the Transformer Architecture. *International Journal of Interactive Multimedia and Artificial Intelligence*. https://doi.org/10.9781/ijimai.2023.02.005.
- Bâra, A. and Oprea, S.-V., 2024. Predicting Day-Ahead Electricity Market Prices through the Integration of Macroeconomic Factors and Machine Learning Techniques. *International Journal of Computational Intelligence Systems*, [online] 17(1), p.10. https://doi.org/10.1007/s44196-023-00387-3.
- Bâra, A., Oprea, S.-V. and Tudorică, B.G., 2023. From the East-European Regional Day-Ahead Markets to a Global Electricity Market. *Computational Economics*. [online] https://doi.org/10.1007/s10614-023-10416-0.
- Bâra, A., Oprea, S.V. and Ciurea, C.E., 2024. *IMPROVING THE STRATEGIES OF THE MARKET PLAYERS USING AN AI-POWERED PRICE FORECAST FOR ELECTRICITY MARKET*. *Technological and Economic Development of Economy*. https://doi.org/10.3846/tede.2023.20251.
- Hubicka, K., Marcjasz, G. and Weron, R., 2019. A Note on Averaging Day-Ahead Electricity Price Forecasts Across Calibration Windows. *IEEE Transactions on Sustainable Energy*. https://doi.org/10.1109/TSTE.2018.2869557.
- Kath, C. and Ziel, F., 2018. The value of forecasts: Quantifying the economic gains of accurate quarter-hourly electricity price forecasts. *Energy Economics*. https://doi.org/10.1016/j.eneco.2018.10.005.
- De la Torre, J., Rodriguez, L.R., Monteagudo, F.E.L., Arredondo, L.R. and Enriquez, J.B., 2024. Electricity price forecast in wholesale markets using conformal prediction: Case study in Mexico. *Energy Science and Engineering*. https://doi.org/10.1002/ese3.1710.
- Lehna, M., Scheller, F. and Herwartz, H., 2022. Forecasting day-ahead electricity prices: A comparison of time series and neural network models taking external regressors into account. *Energy Economics*. https://doi.org/10.1016/j.eneco.2021.105742.
- Pinhão, M., Fonseca, M. and Covas, R., 2022. Electricity Spot Price Forecast by Modelling Supply and Demand Curve. *Mathematics*. https://doi.org/10.3390/math10122012.
- Salahuddin, M.A., Pourahmadi, V., Alameddine, H.A., Bari, M.F. and Boutaba, R., 2022. Chronos: DDoS Attack Detection Using Time-Based Autoencoder. *IEEE Transactions on Network and Service Management*. https://doi.org/10.1109/TNSM.2021.3088326.
- Zhang, C. and Fu, Y., 2024. Probabilistic Electricity Price Forecast with Optimal Prediction Interval. *IEEE Transactions on Power Systems*. https://doi.org/10.1109/TPWRS.2023.3235193.

- Zhang, C., Fu, Y. and Gong, L., 2023. Short-Term Electricity Price Forecast Using Frequency Analysis and Price Spikes Oversampling. *IEEE Transactions on Power Systems*. https://doi.org/10.1109/TPWRS.2022.3218712.