Model Prediction XGBoost for Perspective of Closing Cernavoda Reactor 1 and the Impact on the Romanian Energy System

Andreea CREȚULESCU*
Bucharest, Romania
**Corresponding author, andreea.kretulescu@gmail.com*

Alexandru POPESCU CREȚULESCU*
Bucharest, Romania
**Corresponding author, alexandru.popescucretulescu@gmail.com*

Abstract. E*mploying artificial intelligence (AI) and machine learning (ML) technologies to analyze the progression of energy markets is imperative, as these methodologies facilitate the modeling of diverse domain-specific data to forecast developments in an exceedingly volatile market. Utilizing machine learning algorithms to model data concerning the Romanian energy sector proves beneficial, particularly in circumstances where significant changes necessitate adaptation to internal consumption demands, while mitigating substantial fluctuations in electricity prices. Such fluctuations can profoundly impact crucial national economic indicators.. In this paper, we aim to simulate such a situation for the moment when Romania will be obliged to shut down and modernize Reactor 1 at Cernavodă due to reaching the standard operating time for the CANDU model. We will use an XGBoost for Regression algorithm in which we will input data about the national energy system (consumption, energy quantities according to the energy system's production structure by resource types) from the period 2020-2023, and then we will simulate the shutdown of the reactor to analyze how the national energy market is influenced. Our findings not only reveal the indispensable role of predictive analysis through ML/AI models in anticipating market changes, but also provide practical insights for decision-makers aiming to optimize the national energy system so that when Reactor 1 at Cernavodă is shut down, the domestic market is affected as little as possible, considering that this reactor has supplied the national system with approximately 700 MWe/h over the past 20 years.*

Keywords: dataset, artificial intelligence, nuclear, energy, machine learning, xgboost, algorithm, Cernavodă Reactor 1

Introduction:

Romania operates a single nuclear power plant located in Cernavodă. Construction works began in 1976, following a Romanian-Canadian feasibility study for the CANDU system in Romania. Although initially designed to have five nuclear reactors, currently only two units are operational. The site is situated near the first lock of the Danube-Black Sea navigable canal, bordered by the Cismelei Valley to the north and DJ 223 and the secondary railway line to the southwest (SNN, S.A. CNE Cernavoda, Nuclear Security Report for Unit 1 at Cernavodă Nuclear Power Plant, 2023). The CNE Cernavodă platform resulted from excavations at the former limestone quarry of Ilie Barza, with the current elevation at +16.00 m above mean sea level relative to the Baltic Sea. The CNE Cernavodă site houses nuclear facilities, including operational Unit 1 and Unit 2, constructions and facilities of Units 3 and 4 in conservation, and Unit 5, repurposed to serve as a useful support during the operational lifespan of Units 1 and 2 and the future Units 3 and 4. Additionally, facilities such as the Spent Fuel Intermediate Storage (DICA) and the Intermediate

DOI: 10.2478/picbe-2024-0238
© 2024 A. Crețulescu; A. Popescu Crețulescu, published by Sciendo.
This work is licensed under the Creative Commons Attribution 4.0 License.

Radioactive Solid Waste Storage (DIDSR) are present (SNN, S.A. CNE Cernavoda, Nuclear Security Report for Unit 1 at Cernavodă Nuclear Power Plant , 2023)

Unit 1 features a CANDU-6 PHWR nuclear reactor using heavy water as a moderator and coolant. The heat generated by nuclear fuel is transferred through the primary and secondary cooling circuits, including the steam generators, to produce electrical power. The Unit 1(figure1) turbogenerator provides a gross electrical power of 706.5 MWe. Cooling is achieved using water from the Danube-Black Sea canal, and the generated electrical energy is evacuated through the 400 kV substation into the National Power System. The nuclear fuel used is natural uranium in the form of ceramic pellets assembled into fuel bundles. The primary heat transport system is a closed circuit with two independent loops that transfer heat from the fuel to the steam generators. Electrical energy is produced by a turbogenerator, and the steam is condensed using cooling water from the Danube-Black Sea canal (SNN, S.A. CNE Cernavoda, Nuclear Security Report for Unit 1 at Cernavodă Nuclear Power Plant, 2023).

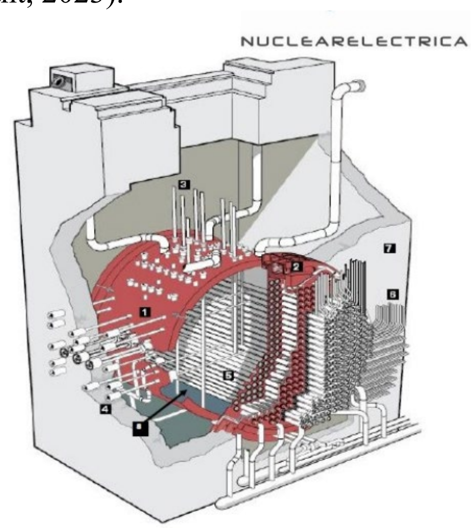
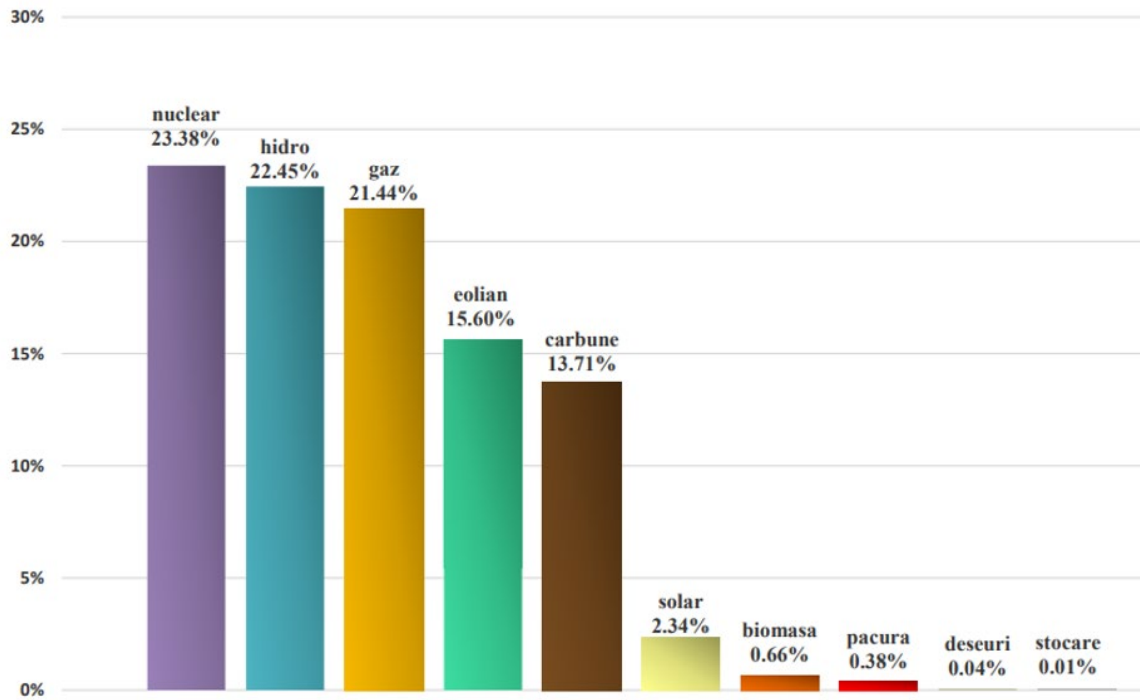


Figure 1. Overview of Reactor Unit 1, CANDU-6 Type

Source SNN, S.A. CNE Cernavoda, Nuclear Security Report for Unit 1 at Cernavodă Nuclear Power Plant.

According to the technical specifications imposed by the constructor and international regulations related to the safe use of nuclear energy (IAEA, 2019), Romania is obliged to shut down the operation of Reactor 1 at Cernavodă for revision and modernization by 2028. Since its commissioning, the nuclear energy supplied by Reactor Cernavodă 1 (approx. 700 MWe) has been a fundamental component for powering the national energy system, a system that operates according to precise and strict rules to provide energy for household and industrial consumption. Additionally, it ensures compliance with established contracts that involve importing or exporting energy (Atanasoae, P. 2020; Vaida, V. et Coroiu, N. 2014). The nuclear energy supplied by the two operational reactors at Cernavodă has influenced the main functional markets in Romania (Vaida, V. et Coroiu, N. 2014), such as the next day market (PZU); bilateral contracts market; balancing market (PE), system services market; cross-border capacity allocation market; capacity market; green certificates market, and intraday market.

Collectively, these Cernavodă Unit 1 et Unit 2 contribute to over 18% of the nation's total electricity production (ANRE, Annual Report, 2023). The Cernavodă Unit 1 reactor is slated for a comprehensive refurbishment project (SNN, Annual Raport, 2020), organized into three distinct phases, with the actual shutdown scheduled between December 2026 and December 2028. The initial phase (2018-2022) focuses on project organization and the development of necessary nuclear safety documentation to extend Unit 1's operational lifespan by another 30 years. This includes increasing its operational hours based on an assessment by Candu Energy and completing a feasibility study for the refurbishment project. The second phase (2022-2026) involves preparing for the refurbishment works, securing financial resources, and acquiring all requisite approvals and permits. The third phase (2027-2028) will see the shutdown of Unit 1 and the execution of the refurbishment works within the unit's facilities. For Unit 2, which commenced commercial operations in 2007, a similar refurbishment process is anticipated to begin post-2037. With the closure of Reactor 1 at Cernavodă estimated for 2027-2028 (SNN, Annual Report, 2020), the national energy mix will undergo a significant change. According to the ANRE report from 2023, the two reactors at Cernavodă have accounted for over 23% of Romania's energy needs (Figure 2).



Source:

Figure 2. Romania, energy system's production structure by resource types (2023)

ANRE, energy system's production structure by resource types, 2023.

The XGBoost for Regression model will utilize complete official data encompassing energy prices and quantities supplied to the national grid from 2020 to 2023 for each energy source (such as nuclear, hydro, wind, coal, gas, biomass, etc.) within the Romanian energy infrastructure. The simulation will proceed under the assumption that Reactor 1 at Cernavodă will cease operations in 2023. By analyzing three years of data preceding the shutdown, our aim is to assess the repercussions of the closure on the country's daily logistical operations and energy expenditures.

Methodology

In this research, we utilize XGBoost, a machine learning regression model, to accurately predict price changes across different scenarios. XGBoost is chosen for its advanced ability to construct regressor trees that minimize prediction errors using the squared error method, as demonstrated in various studies (Chen, G et al, 2022; Deepanraj, B et al, 2022, Abbasimehr, H et al., 2023). This approach ensures high accuracy and relevance in comprehending price fluctuations. To enhance the reliability of our simulation, the dataset underwent necessary preprocessing adjustments (Chen, G et al, 2022). These modifications were crucial in preparing the dataset for analysis, ensuring it could provide meaningful insights. Our methodology, centered on the strategic utilization of XGBoost and meticulous data processing, aims to establish a robust framework for analyzing price changes. This approach underscores our commitment to standardizing methods and contributes to the existing knowledge base on variable price analysis.

Data Sources: For this research, the study harnessed data from two primary online resources to gather comprehensive insights into the energy sector, specifically focusing on electricity production and pricing within Romania and its neighboring EU countries, Hungary and Bulgaria. Production Data: Hourly electricity production data from January 2020 to December 2023 was extracted from "*Sistemul Energetic*", a platform providing detailed insights into Romania's energy production. This dataset includes the total energy generated within the country, categorized by different energy sources (sistemulenergetic.ro, database, 2020-2023). Price Data: Hourly electricity price data for the same period was sourced from Ember-Climate's European Wholesale Electricity Price Dataet OPCOM. This dataset provided electricity prices for Romania and its EU neighbors, Hungary, and Bulgaria, allowing for a comparative analysis of pricing dynamics across borders.

The rationale for Including Neighboring Countries Energy markets is inherently interconnected, especially within the European Union, where energy policies and infrastructures promote cross-border electricity trade. Electricity cannot be stored efficiently at a large scale; thus, any surplus production is often exported to neighboring countries, while deficits are met through imports. This dynamic interplay affects electricity prices, as prices in interconnected markets tend to converge over time due to the physical realities of electricity flow and the regulatory frameworks designed to facilitate energy trade (Vaida, V. et Coroiu, N. 2014). Including Hungary and Bulgaria in the analysis provides a more detailed perspective the dynamics of the local energy market has, in particular how cross-border trade affects prices.

Data Merging and Preprocessing: The collected datasets were meticulously merged to create a final, usable dataset encompassing hourly electricity production and price data across the three countries. Special attention was given to aligning the data according to UTC time to ensure consistency and accuracy in temporal analysis, considering the time zone differences and daylight-saving time adjustments. Understanding Hourly Prices, The study focused on the trade prices, which are determined a day before the actual delivery of electricity. The day-ahead market is crucial for the energy sector as it provides a platform for producers and consumers to hedge against the uncertainty of next-day electricity demand and production (Atanasoae, P. 2020). These prices are indicative of the market's expectation of supply and demand balance and are influenced by various factors, including production capacities, anticipated consumption, and availability of cross-border transmission capacities (Vaida, V. et Coroiu, N. 2014). The study aims to capture short-term fluctuations in the electricity market by analyzing hourly prices, providing insights into how changes in production and cross-border trade affect location-based pricing the impact of the research.

Data Preprocessing: Cyclical Feature Engineering Electricity demand and supply exhibit cyclical patterns due to natural human activities and environmental factors. To model these patterns effectively, we applied sine and cosine transformations to month, day, and hour features. This transformation is based on the formula:

$$x_{sin} = \sin\left(\frac{2\pi x}{\max(x)}\right)$$
$$x_{\cos} = \cos\left(\frac{2\pi x}{\max(x)}\right)$$

where *x* represents the original feature (month, day, hour) and *max*(*x*) is its maximum value. These transformations help preserve the cyclical nature of the data, facilitating the model's ability to recognize and predict patterns based on time.

The normalization of features, excluding the target variable (electricity price), was achieved using the *StandardScaler*. This process adjusts the features so that they have a mean of 0 and a standard deviation of 1. Mathematically, it transforms a feature *X* to *Z* using the formula:

$$Z = \frac{(X - \mu)}{\sigma}$$

where μ is the mean of the feature and σ is the standard deviation. Normalization is critical for models like XGBoost, as it ensures that all features contribute equally to the model's learning process, preventing bias towards variables with higher magnitude.

Model Training and Evaluation, XGBoost for Regression:

XGBoost, a supervised machine learning algorithm, is renowned for its performance in regression tasks (Abbasimehr, H et al, 2023). This efficiency stems from its ability to handle large datasets and its flexibility in modeling complex nonlinear relationships. The algorithm optimizes a regularized objective function that combines a convex loss function (to measure prediction error) with a regularization term (to penalize model complexity), thereby balancing the trade-off between bias and variance (Xgboost Release 0.80 xgboost developers, 2018). Mathematical Explanation of XGBoost. XGBoost improves upon the concept of gradient boosting by systematically optimizing the construction of decision trees (Cortes, C et al, 2019). At its core, XGBoost minimizes a loss function L, which is a sum of the loss calculated at each step t over N trees (Xgboost Release 0.80 xgboost developers, 2018):

$$L(\theta) = \sum_{i=1}^{N} l\left(y_i, y_i^{(t)}\right) + \sum_{i=1}^{t} \Omega(f_i)$$

where $l\left(y_i, y_i^{(t)}\right)$ is the loss at step t comparing the true value y_i with the predicted value $y_i^{(t)}$, and $\Omega(f_i)$ represents the regularization term for the tree f_i, which helps prevent overfitting by penalizing the complexity of the model. XGBoost applies the concept of boosting by combining multiple weak learners (typically decision trees) to create a strong learner (Cortes, C et al, 2019). The model is built in a stage-wise manner, where each new tree corrects errors made by previously trained trees. Mathematically, the prediction model at the n^{th} step$\left(y_i^{(n)}\right)$ for a given data point i can be represented as:

$$y_i^{(n)} = y_i^{(n-1)} + \eta f_n(x_i)$$

where $y_i^{(n-1)}$ is the prediction from the previous step, f_n is the new tree added at step n, x_i is the feature vector for the data point i, and η is the learning rate (Cortes, C et al, 2019).

Initial model Analysis A distribution analysis of the prediction errors (percentage deltas) was performed to assess the model's performance characteristics. The deltas were filtered to focus on a specific range, and their distribution was plotted and fitted with a normal distribution to understand the error characteristics. This analysis provided insights into the model's accuracy and the likelihood of prediction errors.

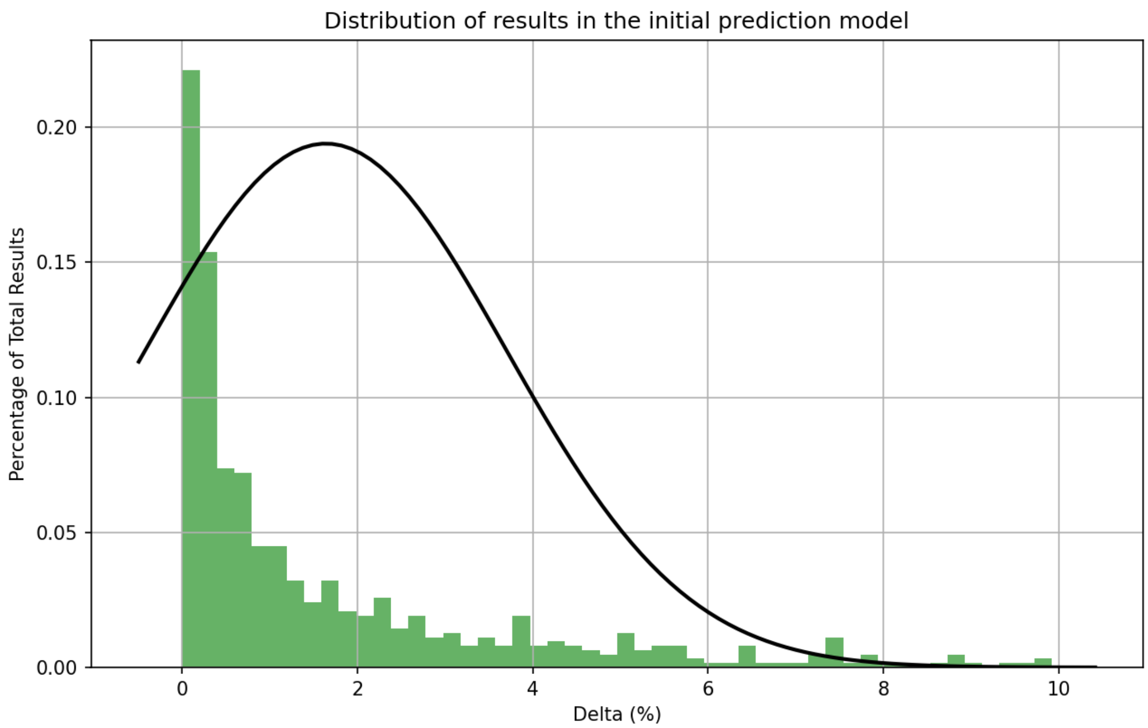


Figure 3. Distribution of the results in the initial prediction model

Source: Own computations

High Prediction Accuracy: The model is excellent at predicting electricity prices with a high level of accuracy. It's so precise that 95% of its predictions are within 10% of the actual prices. Considering that with this limited dataset and the lack of information related to different raw resource prices, weather and socio-economic events the accuracy is high enough so that we can consider changing said data and achieving a good simulation result.

Handling Price Anomalies: When it comes to managing unusual price events, our model stands out by how it deals with low prices, especially when they drop below 15 EUR/MWh. Such low prices are rare and usually occur on days with high production, holidays, or unusual events. Another valid point in this simulation is whether or not the model's accuracy is valid when these exceptional. If it does, we consider it a valid prediction and our simulation becomes relevant. This method shows that our model doesn't just understand typical market trends it also picks up on rare scenarios.

Implications for Energy Production and Pricing: Our initial analysis establishes a solid foundation for further investigation into the correlation between shifts in energy production and pricing dynamics. By effectively forecasting prices under both typical and unusual circumstances, the model emerges as a valuable tool for stakeholders in the energy sector, including producers, policymakers, and analysts. It provides insights into how alterations in energy output can influence pricing, enabling better planning and strategic decision-making to uphold market stability and address unforeseen developments.

Simulation Overview: In response to the temporary reduction in nuclear energy production, we conducted simulations to evaluate the potential impact of two primary strategies on electricity prices and market dynamics. Romania's energy sector faced a critical decision: whether to import the required energy or increase domestic coal energy production, despite its higher costs.

Case 1: Energy Import

The initial strategy entails importing the required energy to offset the decline in nuclear production. This scenario assesses the logistical feasibility, cost ramifications, and potential influence on electricity prices associated with depending on foreign energy sources. While leveraging international energy markets could provide a short-term remedy due to their interconnected nature, it also subjects the market to fluctuations in global prices and reliance on external suppliers.

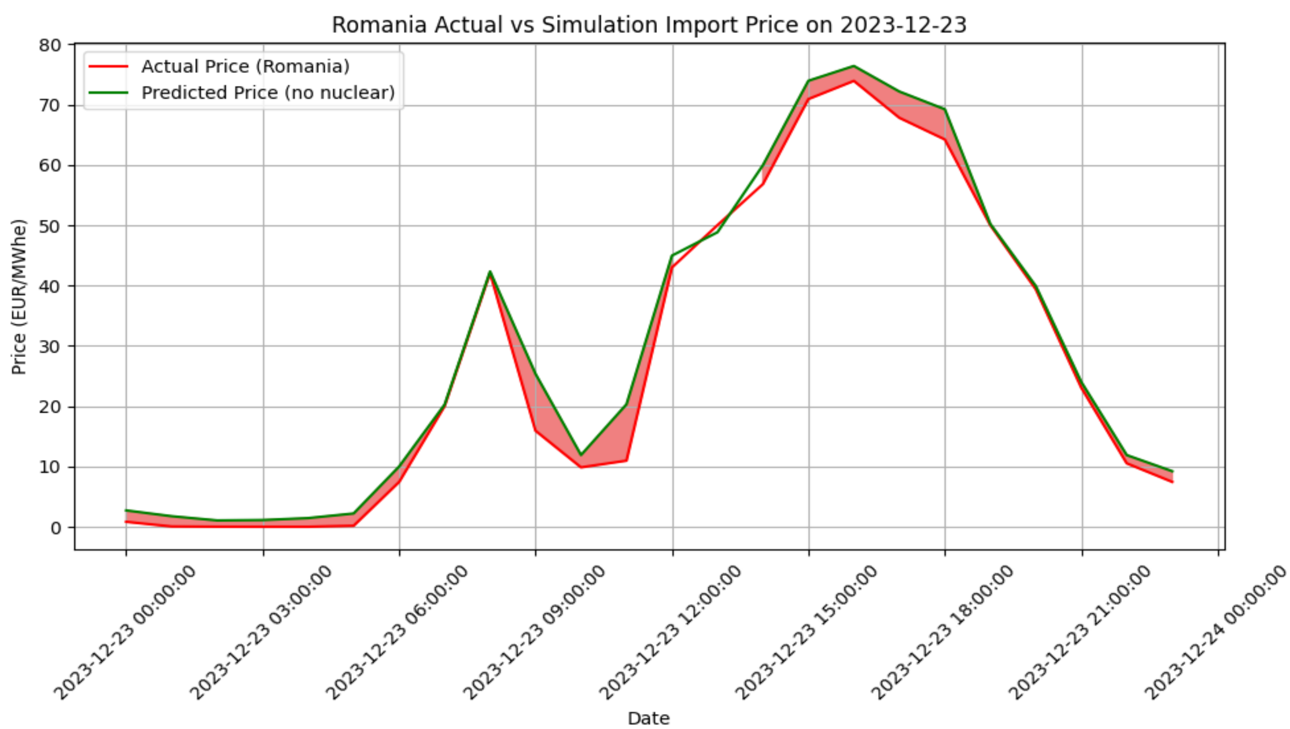


Figure 4. Actual versus simulation, import price analysis

Source: Own computations

Case 2: Increasing Coal Energy Production

The second scenario explores the potential of ramping up coal energy production in Romania. Despite coal being costlier and environmentally more impactful than other energy sources, this approach investigates its viability as a short-term solution to the energy deficit. The assessment encompasses financial considerations, environmental implications, and the effectiveness of boosting coal production in bridging the supply shortfall.

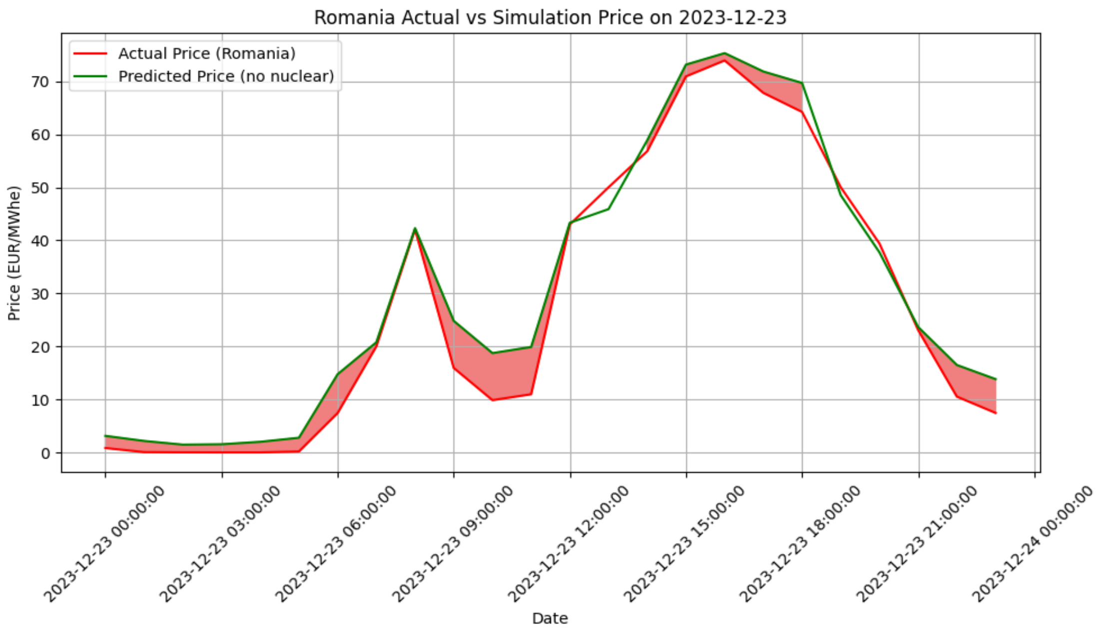


Figure 5. Actual versus simulation, coal price analysis

Source: Own computations

Results and discussions

The simulation produced outcomes with a minimal margin of error, indicating the predictive precision of the model across all three proposed scenarios. The reliability of the simulation results was assessed using real-world data, with special attention given to the unique circumstances of December 25th. During this time, energy consumption experiences a notable decrease due to the Christmas holiday and the closure of many industrial facilities, resulting in trading prices often turning negative.

Limitations and Considerations, Lack of Base Load Energy: The study acknowledges Romania's limited capacity for incorporating additional gas energy, which, although cheaper than coal and capable of instant generation, cannot be relied upon due to infrastructural constraints. Base Load Energy refers to the consistent, reliable energy supply typically provided by sources like nuclear or coal, essential for meeting the base demand level across the grid. Renewable Energy Variability: The unpredictable nature of weather-dependent renewable energy sources, such as wind and solar power, renders them unreliable for compensating the nuclear shortfall under the current circumstances. Hydroelectric Energy Constraints: Although hydroelectric power was considered as an alternative, the simulation determined that redirecting hydroelectric energy to replace nuclear capacity would exceed Romania's annual hydroelectric production capabilities, highlighting the limitations of this energy source in addressing the shortfall.

Impact of December Holiday Season Price Predictions During Holidays: A critical test for the simulation's validity was its ability to predict price adjustments during the Christmas holiday period. In reality, energy consumption patterns shift significantly during this time, with potential reductions in industrial activity and variations in household consumption. The model's accuracy was demonstrated by its prediction of significantly lowered prices, aligning with the expected decrease in demand. Prices dropping within the 15 EUR mark, and even approaching 0 EUR/MWh, confirmed the model's sensitivity to temporal variations and special events, validating its predictive accuracy for periods of atypical consumption.

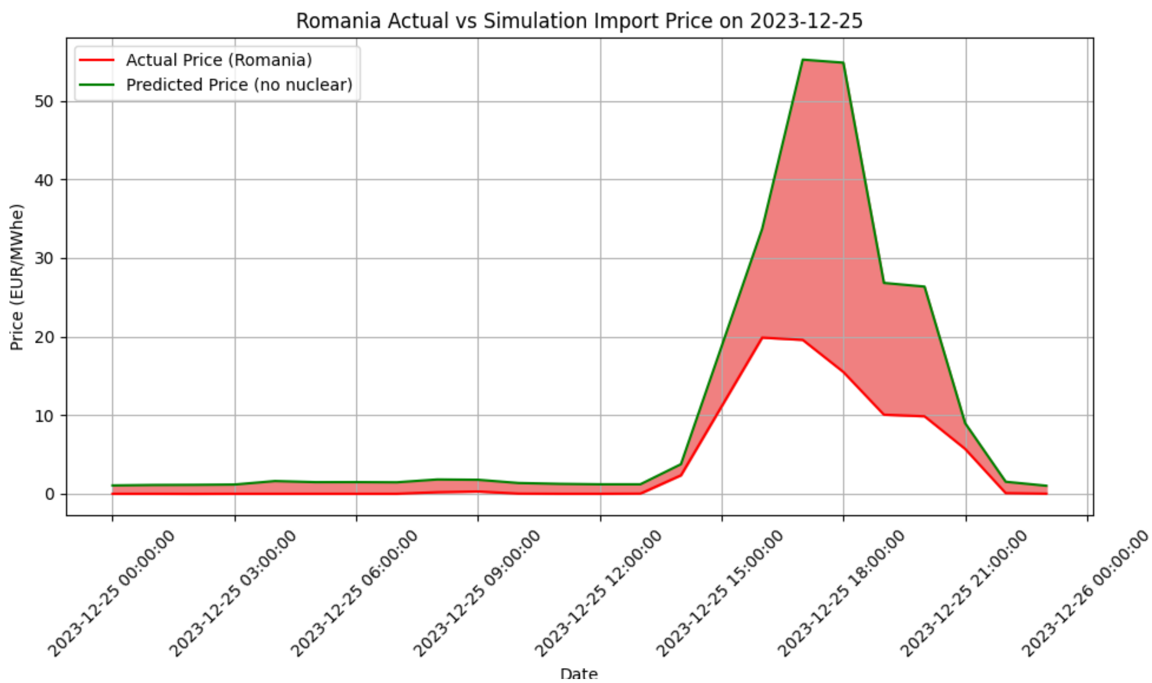


Fig 6. Holiday Season Price dip

Source: Own computation

Weather Conditions: December's weather plays a crucial role in shaping energy production and consumption. The model accounted for: Reduced Solar Impact: Shorter daylight hours and potential cloud cover diminish solar energy's contribution, a factor accurately reflected in the simulation outcomes. Wind Energy Variability: Wind energy production is highly susceptible to weather conditions. The simulation considered scenarios of both insufficient and excessively strong winds, which could limit wind turbine operations, reflecting real-world constraints on wind energy availability. Hydroelectric Power Fluctuations: Weather unpredictability also affects hydroelectric power generation, with precipitation levels influencing water availability for hydro plants.

Energy Demand Considerations The simulation also factored in the increased energy demand associated with December, not only from regular household and industrial consumption but also from holiday-specific activities like lighting displays. This period often sees a mix of low production capabilities and high, albeit fluctuating, consumption demands, further complicated by elevated import prices due to increased demand across Europe.

Simulation Validity: The simulation was thorough, factoring in the subtle impacts of the holiday season, weather conditions, and consumption patterns, which highlighted its accuracy. The model's capability to predict price reductions, like those seen in low-demand periods, and its inclusion of various external factors that influence supply and demand, showcase its advanced grasp of market dynamics.

Analysis Summary Import Scenario: Adopting an import strategy to mitigate the nuclear shortfall results in an average predicted price increase of 2.924%. While seemingly economical, this approach is fraught with risks due to market volatility and infrastructure limitations. Romania's ability to bring in energy from other countries has its limits. In extreme situations where there's no wind, solar, or hydroelectric power available, relying too much on imported energy could mean having to take serious steps. This might include asking big energy users to cut down on their use to keep the power grid stable and prevent problems. Coal Scenario: Increasing coal production as an alternative lead to an average predicted price increase of 3.505%, reflecting the higher costs and environmental toll associated with coal energy. This strategy also faces limitations regarding the availability and cost of coal, as well as the significant initial expenses involved in activating additional energy production facilities.

Conclusion

Through the conducted study, we demonstrate the importance of utilizing predictions made with the help of AI/ML (XGBoost for Regression) technologies by processing public databases about the Romanian energy system. Even though our model showed a marginal impact on energy prices in a hypothetical closure of Cernavoda Reactor 1 in 2023, creating this model will be extremely useful for analyzing the impact that the mandatory closure of the reactor by 2028 will have on the market and energy costs in Romania at that time. Furthermore, until the closure of Cernavoda Reactor 1, Romania has committed to abandoning a significant part of its capacities for coal-based energy production (Energy Strategy, 2022), and the country's capacity for energy imports is limited to 3000 MW (Transelectrica, The Development Plan RET for the period 2022 – 2031). The reliance on external energy sources exposes the national grid to market volatility, which can lead to cost increases. In this study by employing a simple form of a prediction model we can somewhat see the consequences of different actions. By considering two hypothetical scenarios: one where energy imports increase and another where coal energy production is ramped up. The analysis of these scenarios aimed to assess the impact on electricity prices. The findings highlight the limitations and risks associated with each strategy, offering insights for energy policy development.

Energy Imports: The strategy of compensating for domestic production shortfalls through energy imports, while initially seeming viable, encounters significant challenges. Risks posed by market volatility and Romania's limited import capacity are considerable. In extreme weather conditions, affecting renewable energy sources such as wind and solar, reliance on imports could necessitate severe consumption restrictions to prevent grid failures.

Energy Coal Production: The option to increase coal energy production, though potentially addressing immediate energy needs, is fraught with financial, environmental, and logistical challenges. The adverse environmental effects, uncertainties regarding coal supply, and the considerable expenses associated with activating additional production capacities deem this alternative impractical.

This research represents a significant stride in grasping the intricate dynamics of energy management and emphasizes the crucial role of predictive analytics in shaping well-informed policy decisions. The study's constraints, particularly the difficulties associated with reliance on energy sources during constrained circumstances, highlight the importance of having a diversified energy mix. The model crafted in this research underscores the importance of employing comprehensive and advanced analytical tools in the energy sector. Such analytical frameworks are indispensable for maximizing energy production, maintaining grid reliability, attaining energy self-sufficiency, and mitigating environmental repercussions.

References

Atanasoae, P. (2020). *Energy market*, Editura Matrix Rom. ISBN: 978-606-25-0195-2

Abbasimehr, H, Paki, R, Bahrini, A (2023*) A novel XGBoost-based featurization approach to forecast renewable energy consumption with deep learning models*, SienceDirect, Vol. 38, https://doi.org/10.1016/j.suscom.2023.100863

ANRE, 2023, REPORT ON THE RESULTS OF MONITORING THE ELECTRIC ENERGY MARKET IN OCTOBER 2023 https://anre.ro/wp-content/uploads/2023/12/Monitoct23.pdf

Chen, G; Yang, H; Yang, D; Guo, K; Feng, M; Wang, Y; Li Fu; Jiang, H, (2022) *A New Rate of Penetration Prediction Model Based on the XGBoost and Explanation Method* , The 56th U.S. Rock Mechanics/Geomechanics Symposium, Santa Fe, New Mexico, USA, June 2022. https://doi.org/10.56952/ARMA-2022-0142

Cortes, C., Mohri, M., & Storcheus, D. (2019). *Regularized Gradient Boosting.* In Proceedings of the 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada.

Deepanraj, B., Senthilkumar,N., T. Jarin, Gurel, A., Sundar, L.S., Anand, A.V. (2022). *Intelligent wild geese algorithm with deep learning driven short term load forecasting for sustainable energy management in microgrids*, SienceDirect, Vol. 36, https://doi.org/10.1016/j.suscom.2022.100813

Energy price data 2020-2023 https://www.opcom.ro/acasa/ro

Energy data RO system 2020-2023 https://sistemulenergetic.ro/

Ember-Climate's European Wholesale Electricity Price Data,

https://ember-climate.org/data/data-tools/europe-power-prices/Transelectrica, energy nuclear data/10 minutes /2023

The National Integrated Plan in the field of Energy and Climate Change 2021-2030, https://energy.ec.europa.eu/system/files/2020-04/ro_final_necp_main_ro_0.pdf

IEA, 2023. *World Energy Outlook 2023*, https://www.iea.org/reports/world-energy-outlook-2023

IAEA, 2019, National Report under the Convention on Nuclear Safety/Romania

SNN, S.A. CNE Cernavoda. (2023) Nuclear Security Report for Unit 1 at Cernavodă Nuclear Power Plant , 2023 https://www.nuclearelectrica.ro/wp-content/uploads/2023/03/RFS-U1-Sumar-2023-von.pdf

SNN, The Project for the Reengineering of Unit 1 at Cernavoda Nuclear Power Plant, (2022) https://www.nuclearelectrica.ro/ir/proiecte-de-investitii/

Vaida,V., Coroiu,N. (2014), *Electricity Market. Creation, Operation, and Development of the Electricity Market in Romania*, ISBN: 978-973-720-534-6

Xgboost Release 0.80 xgboost developers, (2018) Python Module Index, https://xgboost.readthedocs.io/_/downloads/en/release_0.80/pdf/