European Journal of Interdisciplinary Studies ♦ ISSN 2067-3795
Vol. 18, Issue 1, 2026

A Comparative Forecasting Analysis of Clean Energy Stocks using Recurrent Neural Networks

Maria Elena DE GIULI
Department of Economics and Management, University of Pavia, Italy
0000-0002-5221-0005
elena.degiuli@unipv.it

Luca DI PERSIO
Department of Computer Science - College of Mathematics, University of Verona, Italy
0000-0002-0317-4351
luca.dipersio@univr.it

Alessandra TANDA
Department of Economics, University of Insubria, Italy
0000-0002-6885-899X
alessandra.tanda@uninsubria.it

Thi Nhu Y TRAN
Department of Computer Science - College of Mathematics, University of Verona, Italy
thinhuy.tran@studenti.univr.it

Abstract
Climate change represents one of the most pressing existential threats of our time, requiring coordinated, cross-domain responses that integrate technological, financial, and policy-oriented knowledge. This paper investigates the behavior of selected clean energy stock indices before, during, and after the COVID-19 crisis and applies advanced machine learning methodologies, specifically Recurrent Neural Networks (RNNs) and Gated Recurrent Units (GRUs), to predict clean energy stock prices. The results provide new insights into the nonlinear dynamics of financial markets linked to the clean energy sector and show that both LSTM and GRU models outperform VAR in stock price forecasting, delivering superior accuracy. This research highlights the effectiveness of integrating traditional statistical models with deep learning techniques to improve forecasting performance. It promotes a deeper understanding of the behavior of this crucial industry, providing a bridge between finance, technology, and sustainability topics, necessary to achieving a resilient and equitable low-carbon economy.

Keywords: Forecasting; Gated Recurrent Unit (GRU); Recurrent neural networks (RNNs); Energy stocks; clean energy.

JEL Classification: G19; Q49

DOI: https://doi.org/10.24818/ejis.2026.03

1. Introduction

Climate change represents one of the most pressing existential threats of our time, demanding coordinated efforts across scientific, economic, and policy domains. In the realm of energy markets, renewable energy and energy crises have gained interest from academia and policymakers (Triantafyllidou et al., 2024). Advancing knowledge in this field would accelerate the transition toward sustainable energy systems (Leiria & Martins, 2022). The number of studies on the financial aspects of the clean energy sector is cumbersome. Research has focused

Received: 18 March 2026; Revised: 10 April 2026; Accepted: 20 April 2026

on the technical and operational dimensions of the energy transition, including energy efficiency improvements, energy-saving behaviors, demand-side management, and prediction and control of energy consumption peaks (Agakishiev et al., 2025). Many studies aim to optimize the balance between production and consumption, contributing to the broader goal of sustainable energy system design (Haghaniat et al., 2024; Mewada et al., 2024; Prajapati, 2023). Within the financial literature, numerous studies focus on the determinants and prediction of energy market prices. In the context of the transition to cleaner and more sustainable energy production systems, green or clean energy is gaining relevance (Murărașu, 2023), although fossil fuels still play a significant role, directly or indirectly influencing the prices of clean energy producers (Cevik et al., 2024).

Many contributions investigate the connectedness of clean and traditional energy sectors, employing established methods such as Granger causality or Diebold-Yilmaz risk (volatility) spillover (Charda & Pendaraki, 2023; Dai et al., 2022; T. Liu & Gong, 2020; Sharma et al., 2022). While traditional literature, such as Yadav et al. (2024), has focused on co-movements and network-based connectedness between energy commodities and green bonds, research is increasingly pivoting toward advanced predictive modeling built on artificial intelligence (AI). Indeed, traditional methods like Granger causality – employed by Diebold and Yılmaz (2014) and Baruník and Křehlík (2018) to map volatility spillovers – often fail to capture the non-linear complexities inherent in climate-related markets. To address these limitations, Recurrent Neural Networks (RNNs) and Gated Recurrent Units (GRUs) have emerged as superior tools for forecasting, as they are specifically designed to handle sequential dependencies in time-series data. Nadirgil (2023) exemplifies this shift by employing RNNs to analyze the causality between GHG-emitting industries and EUA (European Union Allowance) prices, showing how neural networks can better identify the specific sectors able to influence carbon pricing. Compared to traditional approaches – such as VaR, BEKK, DCC, and Markov regime switching – RNNs and GRUs offer a more robust framework for price prediction and are therefore increasingly employed (e.g., Xiao et al., 2024; Sulistio et al., 2023). In fact, despite the historical prevalence of threshold cointegration and wavelet methods, the integration of these advanced AI-based methodologies represents a critical frontier in energy market analysis, even if their application remains less frequent than traditional models.

The statistical approach predicts energy sector stocks based on historical time series data. Time series models are well-suited for forecasting when data is collected at regular intervals and has temporal dependencies. These models would capture trends, seasonality, and cyclic patterns typical in energy production data (e.g., seasonal fluctuations in energy demand). In particular, ARIMA (Auto-Regressive Integrated Moving Average) works well for stationary data and can handle seasonality if appropriately tuned (Zhang, 2003); SARIMA (Seasonal ARIMA), a variant of ARIMA that accounts for seasonality, is ideal for energy data that exhibits seasonal patterns, like higher demand during winter months, and VAR (Linear Autoregressive models) is employed to detect Granger Causality in time series analysis (Ediger & Akar, 2007; X. Liu et al., 2021). However, the performance declines as the prediction time steps increase.

As mentioned, among the innovative AI approaches, Artificial Neural Networks (ANN) (Hochreiter & Schmidhuber, 1997) and Recurrent Neural Networks (RNN), have been successfully applied to the energy sector (Ahmad et al., 2014; Raza & Khosravi, 2015; Soman et al., 2010). This approach is suitable for predicting the intermittent nature of energy sector stocks due to their non-linearity. RNN models are proposed by Tank et al. (2021) as a novel class of nonlinear approaches that utilize structured neural network architectures. By applying convex group-lasso penalties, the model encourages specific weights to shrink to zero, effectively revealing the Granger causal structure. This framework has notable advantages over traditional approaches, as it efficiently captures long-range dependencies through the inherent structure of RNNs or by automatically selecting relevant time lags.

In recent years, deep learning has experienced rapid growth. Literature shows that, compared to standard neural networks, deep learning can explore more inherent, hidden patterns in data and exhibit higher accuracy for forecasting. RNN is designed to handle time sequences due to the memory units in the neurons, which can retain historical information. However, training RNN is difficult because the RNN model may not converge (Bengio et al., 1994; Pascanu et al., 2013). Long Short-Term Memory (LSTM) is one of the improved variations of RNN, which is a much faster RNN when dealing with time sequence data and is easier to converge (Liang et al., 2018).

The Gated Recurrent Unit (GRU) is introduced to modify the LSTM cell. GRU was developed by Cho (2014) to model time series, creating a mechanism that complements the ability to predict long-term dependencies with improved integration of short-term information. The aim is to enable adaptive modelling of dependencies over different time horizons. Compared to LSTM, GRU has a simplified cell structure that operates based on a gating system, but only has an update and reset gate. The main difference to LSTM is that the cell state can be completely revised at each iteration and updated with short-term information via the reset gate. LSTM, on the other hand, provides a mechanism that limits the change gradient that can be realized at each iteration. Hence, LSTM does not allow past information to be completely discarded, whereas GRU does (Lindemann et al., 2021).

This paper aims to predict clean energy stocks before, during, and after the COVID-19 crisis and to overcome the limitations of linear models it employs, differently from previous studies, ML methodology, namely a Recurrent Neural Network (RNN). We contribute to the literature on energy prices forecasting by applying more sophisticated methods to understand the interrelations between stocks that can determine the return and riskiness of portfolios for a global investor. Applied to stock index data from energy companies, this approach enables a more precise examination of causal relationships in complex, nonlinear financial networks, providing insights into interactions within the energy sector.

Beyond the methodological contribution, this work adopts a transdisciplinary research perspective that integrates insights from financial markets and energy economics within the broader framework of sustainable development. This integrative framework reflects the complex nature of energy transition phenomena, where technological, financial, and environmental dimensions interact dynamically. By bridging these disciplines, the study contributes to a broader understanding of how data-driven methods can support sustainable investment decisions and enhance the resilience of clean energy markets in the face of systemic shocks.

This paper is organized as follows: Section 2 reviews the models employed for forecasting exercises, drawing on the econometrics and financial mathematics literature; Section 3 presents the data and results; and Section 4 concludes.

2. Methodology

2.1 Linear Autoregressive Models and Granger Causality

To establish a robust baseline for our forecasting comparison, we first consider a traditional linear model. Among these, the Vector Autoregression (VAR) model is a well-known fundamental tool for capturing dynamic interdependencies within a multivariate time series

(Stock and Watson, 2001). By modeling each variable as a function of its own past values and the past values of all other variables in the system, the VAR framework allows us to identify Granger causality, a statistical concept used to determine if one time series is relevant to forecasting another time series.

Let $x_t \in R^p$ be a p-dimensional stationary time series and assume we have observed the process at T time points, $(x_1, \ldots, x_T)$. We use the VAR model to detect Granger causality in time series analysis. In this model, the time series at time t, x_t, is assumed to be a linear combination of the past *K* lags of the series.

$$x_t = \sum_{k=1}^{K} A^{(k)} x_{t-k} + e_t \qquad (1)$$

where $A^{(k)}$ is a $p \times p$ matrix that specifies how lag k affects the future evolution of the series, and e_t is zero mean noise. In this model, time series j does not Granger–cause time series i if and only if for all k, $A_{ij}^{k} = 0$.

A Granger causal analysis in the VAR model thus reduces to determining which values in $A^{(k)}$ are zero over all lags. In higher-dimensional settings, this may be determined by solving a group lasso regression problem,

$$min_{A^{(1)},\ldots,A^{(K)}} \sum_{t=K+1}^{T} ||x_t - \sum_{k=1}^{K} A^{(k)} x_{t-k}||_2^2 + \lambda - \sum_{ij} ||A_{ij}^{(1)}, \ldots, A^{(K)}||_2 \qquad (2)$$

Where

- $||.||_2$ denotes the standard L_2 norm,
- $||A_{ij}^{(1)}, \ldots, A^{(K)}||_2$ jointly shrinks all A_{ij}^{k} parameters to zero across all lags *k*. The sum over all (i,j) L_2 norms of $A_{ij}^{(1)}$,..., $A_{ij}^{(K)}$ is known as a group lasso penalty, and jointly shrinks all A_{ij}^{k} parameters to zero across all lags *k.* The hyperparameter $\lambda > 0$ controls the level of group sparsity.

The group penalty in the Equation (2) could be substituted with a structured hierarchical penalty that automatically selects the lag for each Granger causal interaction.

2.2 Gated Recurrent Neural Networks

2.2.1 Long Short-Term Memory Unit (LSTM)

The Long Short-Term Memory (LSTM) unit was initially proposed by Hochreiter and Schmidhuber (1997). Since then, several minor modifications have been made to the original LSTM unit. We follow the implementation of LSTM as used in Graves (2013). Unlike the recurrent unit, which computes a weighted sum of the input signal and applies a nonlinear function, each j-th LSTM unit maintains a memory c_t^j at time *t*. The output h_t^j, or the activation, of the LSTM unit, is then

$$h_t^j = o_t^j \ tanh(c_t^j) \qquad (3)$$

where o_t^j is an output gate that modulates the amount of memory content exposure, the output gate is computed by:

$$\sigma_t^j = \sigma(W_x x_t + U_h h_{t-1} + V_o c_t^j)$$

where σ is a logistic sigmoid function, and V_o is a diagonal matrix.

The memory cell c_t^j is updated by partially forgetting the existing memory and adding a new memory content $\tilde{c}_t^j$:

$$c_t^j = f_t^j c_{t-1}^j + i_t^j \tilde{c}_t^j \qquad (4)$$

where the new memory content is

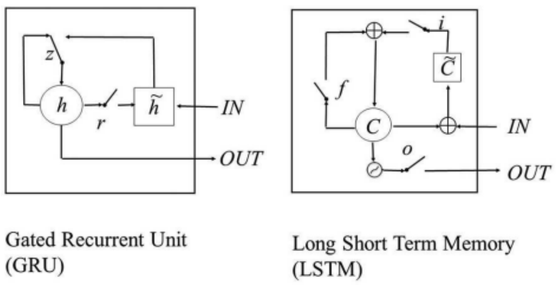
$$\tilde{c}_t^j = tanh(W_x x_t + U_h h_{t-1})^j$$

The extent to which the existing memory is forgotten is modulated by a forget gate f_t^j, and the degree to which the new memory content is added to the memory cell is modulated by an input gate i_t^j. Gates are computed by:

$$f_t^j = \sigma(W_x x_t + U_h h_{t-1} + V_f c_{t-1}^j)$$
$$i_t^j = \sigma\,(W_x x_t + U_h h_{t-1} + V_i c_{t-1}^j)$$

Note that V_f and V_i are diagonal matrices. In a more general setting, W_x and U_h in the equations are the parameters W_o and $U_o$for the output gate, $W_f$and U_f for the forget gate, W_i and U_i for the input gate, and W_c and U_c for the candidate memory. Unlike the traditional recurrent unit, which overwrites its content at each time-step, an LSTM unit is able to decide whether to keep the existing memory via the introduced gates. Intuitively, if the LSTM unit detects an important feature from an input sequence at an early stage, it easily carries this information (the existence of the feature) over a long distance, hence capturing potential long-distance dependencies (see Chung et al.,2014). See Fig. 1 for the graphical illustration.

Figure 1. Gated recurrent unit (GRU; left) *r* = reset gate, *z* = update gate, *h* = activation, $\tilde{h}$ = candidate activation, long short-term memory (LSTM; right) with input (*i*), forget (*f*), output-gate (*o*), memory cell (*c*), new memory cell ($\tilde{c}$).



Source: Scheme cited after Rana (2016, p 3); Chung et al. (2014, p 3).

2.3 Gated Recurrent Unit (GRU)

A gated recurrent unit (GRU) was proposed by Cho et al. (2014) to make each recurrent unit adaptively capture dependencies of different time scales. Similar to the LSTM unit, the GRU has gating units that modulate the flow of information inside the unit; however, without having a separate memory cell (Chung et al., 2014).

The activation h_t of the GRU at time t is a linear interpolation between the previous activation h_{t-1} and the candidate activation $\tilde{h}_t$:

$$h_t = (1 - z_t)h_{t-1} + z_t \widetilde{h_t} \tag{5.1}$$

Or in vector notation

$$h_t = (1 - z_t) \odot h_{\{t-1\}} + z_t \odot h_{t_{tilde}} \tag{5.2}$$

an update gate z_t decides how much the unit updates its activation or content. The update gate is computed by:

$$z_t = \sigma (W_z x_t + U_z h_{t-1} + b_z) \tag{6}$$

This procedure of taking a linear sum between the existing and newly computed states is similar to the LSTM unit. The GRU, however, does not have any mechanism to control the degree to which the state is exposed; it exposes the whole state each time. The candidate activation $\tilde{h}_t$ is computed similarly to that of the traditional recurrent unit and as in Bahdanau et al. (2014):

$$\widetilde{h_t} = tanh (W_x x_t + U_h(r_t \; h_{t-1}) + b_h) \tag{7}$$

where r_t is a set of reset gates and $\odot$ is an element-wise multiplication. When r_t is close to 0, the reset gate effectively makes the unit act as if it is reading the first symbol of an input sequence, allowing it to forget the previously computed state.

The reset gate r_t is computed similarly to the update gate:

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \tag{8}$$

See Fig. 1 for the graphical illustration of the GRU.

2.4 Hyperparameter Optimization for GRU and LSTM Models

To ensure the optimal performance of the Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models, hyperparameter optimization was conducted using Keras-Tuner's RandomSearch method. This process systematically explored the configuration space for key parameters critical to the performance of GRU and LSTM models. The following parameters were optimized:

- Number of Units: The number of units in each GRU/LSTM layer defines the model's ability to learn temporal dependencies.
  – First Layer: Units were tested within the range of 50 to 200, incremented by 50 for both GRU and LSTM.

– Second Layer: Units were tested within the range of 25 to 100, incremented by 25 for GRU and LSTM.

Combining two layers allowed the model to learn high-level and detailed temporal patterns in the data.

- Dropout Rate: Dropout regularization was applied to each GRU/LSTM layer to mitigate overfitting by randomly deactivating neurons during training. Dropout rates were optimized within the range of 0.1 to 0.5, incremented by 0.1.
- Learning Rate: The learning rate controls the step size of the Adam optimizer during backpropagation. A logarithmic search was conducted over 10-5 to 10-2, ensuring the exploration of both fine and broad learning adjustments.
- Batch Size: Batch size determines the number of samples used in each gradient descent step. Although not extensively tuned in this study, the model was tested with values such as 10 to strike a balance between computational efficiency and model accuracy.

Each GRU and LSTM model underwent 30 training epochs during hyperparameter tuning, with validation data used to select the configuration yielding the lowest validation loss. This comprehensive tuning process enhanced the GRU and LSTM models' ability to capture complex temporal dependencies in energy sector stock data while maintaining generalization for unseen test sets.

3. Data, Results and Discussions

3.1 Sample description

The data on Energy stock indices were collected from S&P Global, Bloomberg, and Refinitiv LSEG. The dataset covers the period from April 30, 2014, to March 25, 2024, that includes the COVID crisis and the Russian invasion of Ukraine in early 2022, that also produced an increase in energy prices and a peak in geopolitical tension and uncertainty on the markets. Companies in the energy sector engage in various forms of energy production and are typically categorized based on the type of energy they generate. These companies are classified into two main categories:

- Non-renewable energy sources: Petroleum products and oil, natural gas, gasoline, diesel fuel, heating oil, nuclear energy, and coal.
- Renewable energy sources: Hydropower, biofuels (e.g., ethanol), wind power, solar power, and hydroelectric energy.

Among the available indices on the market, based on previous literature and data availability, we selected the main indices tracking the performance of renewable and non-renewable energy producers (Table 1).

Fig. 2 shows the evolution of the selected indices and commodity indicators from 2014 to 2024. Each subplot represents a different index, suggesting a comparative analysis of market sectors, commodities, and volatility over the decade.

Clean Energy & ESG-related indices (SPGTCLEN, CELS, ECO, ERIX, MIWD000IiPUS) show a strong upward trend from around 2019 to 2021, followed by sharp declines through 2022–2024. The peak corresponds to the post-pandemic "green transition" boom, when clean energy stocks surged, followed by corrections as interest rates rose and enthusiasm cooled.

Volatility increased after 2020, suggesting sector sensitivity to macroeconomic policy and energy price fluctuations.

Global & Regional Equity Indices (MIWO00000PUS, MIWD00000PUS) exhibit steady upward trajectories, interrupted briefly by the 2020 pandemic dip, followed by a recovery until around 2022, and then mild corrections.

Beside energy indices, we include the S&P 500, the U.S. Market Index (SPX), as a proxy for general market conditions. SPX exhibits a steady growth pattern, marked by a noticeable decline in 2020 due to the COVID crisis and a subsequent recovery to new highs around 2021–2022. Additionally, we include the Volatility Index (VIX) to capture market sentiment on expected riskiness. The index exhibits spikes in 2020 (during the COVID crisis) and smaller peaks in 2022 (likely due to interest rate increases and inflation concerns, as well as rising political uncertainty), consistent with market stress periods.

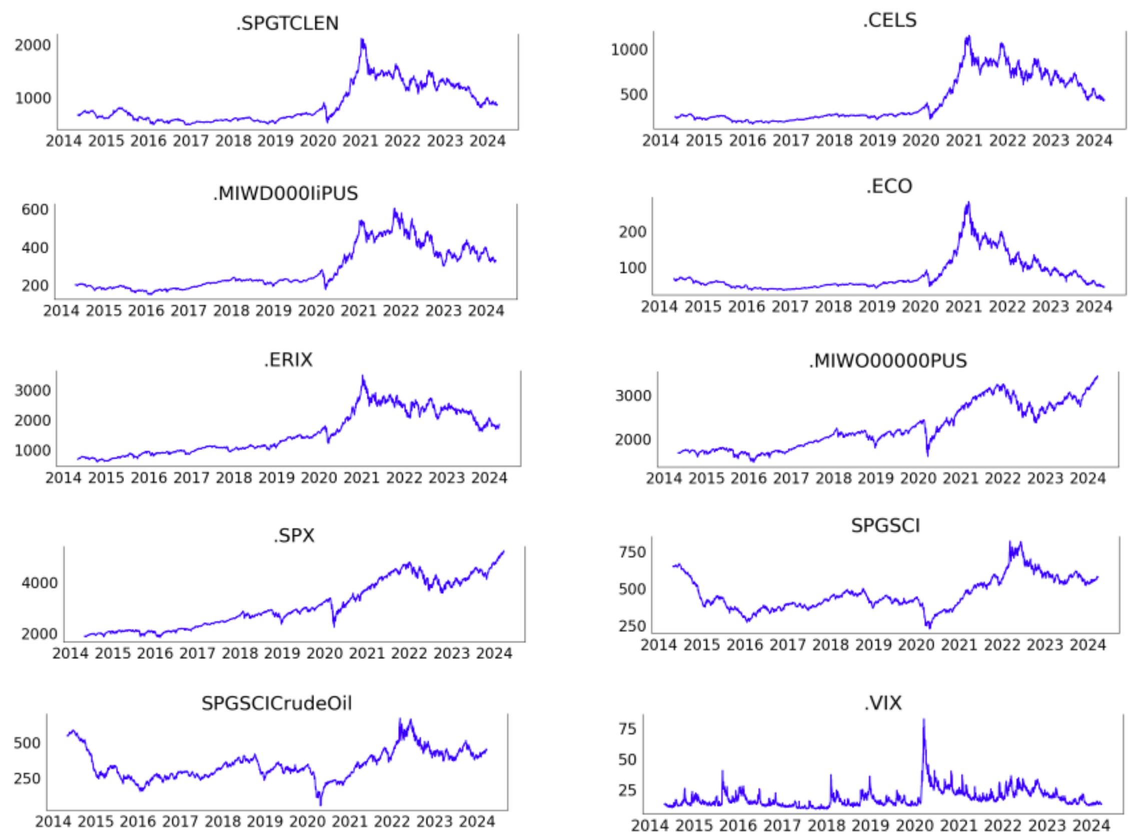
The general pattern aligns with global equity market trends, where growth persisted despite macro headwinds.

Commodities (SPGSCI, SPGSCICrudeOil) show major fluctuations. For general commodities, there was a sharp decline in early 2020, in line with the pandemic-induced demand shock, followed by a recovery in 2021–2022 amid inflation and supply constraints. In recent years, we have witnessed a general decline. These patterns reflect the cyclicality of commodity markets. In particular, S&P GSCI Crude Oil Index is a measure of crude oil market performance, but it is able to reflect energy price cycles, global demand recovery, and geopolitical risk.

Table 1. List of Indices included in the sample

Code	Index name	Tag
SPGTCLEN	S&P Global Clean Energy Transition Index	Clean Energy & ESG-related indices
CELS	Nasdaq Clean Edge Green Energy	Clean Energy & ESG-related indices
MIWD000IiPUS	MSCI Global Environment Index	Clean Energy & ESG-related indices
ECO	Wilder Hill Clean Energy Index	Clean Energy & ESG-related indices
ERIX	European Renewable Energy Index	Clean Energy & ESG-related indices
MIWO00000PUS	MSCI International World Price Index	Global & Regional Equity
SPX	S&P 500	Global & Regional Equity
SPGSCI	S&P Goldman Sachs Commodity Index	Commodities
SPGSCICrudeOil	S&P GSCI Crude Oil Index	Commodities
VIX	CBOE Volatility Index	Uncertainty

Figure 2. Indices price time series for the period 2014-2024



3.2 Data Preprocessing

The data preprocessing steps include handling missing values and normalization of the dataset. To address missing values, the forward filling method was applied, where null values were filled with the corresponding nearest available values from the previous period. Specifically, the `fillna(method='ffill')` function was used to replace any missing data. After filling in the missing values, duplicate rows were identified and removed using the `duplicated()` function. This ensures that the dataset is free of redundant entries.

The dataset was then partitioned into training and test sets. Next, the training partition underwent normalization to bring the variables onto a standard scale. Normalization was performed using the `StandardScaler` from the `sklearn.preprocessing` library, which scales each feature so that its mean becomes zero and its standard deviation becomes 1.

3.3 VAR and Granger Causality

We test for Granger causality using a maximum lag of 12 on each pair of indices. The causality analysis reveals significant bidirectional dependencies across the majority of the time series, with most p-values falling below the 0.05 threshold. For instance, the reciprocal causal relationship between CELS and SPGTCLEN is statistically significant, rejecting the null hypothesis of non-causality. While most variables within the system exhibit strong interconnectedness, MIWD000IiPUS presents several non-significant relationships (p-value > 0.05).

Besides, we control for stationarity of time series using Augmented Dickey Fuller (ADF) test (Dickey & Fuller, 1979) and the KPSS test (Kwiatkowski et al., 1992). Preliminary ADF and KPSS tests indicated that the raw series were non-stationary. To achieve stationarity, a first-differencing transformation was applied; subsequent KPSS testing confirmed that all differenced series were stationary, satisfying the fundamental requirements for VAR estimation.

To determine the optimal lag order for the Vector Autoregression (VAR) model, we primarily utilized the Akaike Information Criterion (AIC) (Akaike, 2003; Lütkepohl, 2013), selecting the order that minimized information loss and subsequently testing incremental increases to ensure model robustness. While AIC served as the primary benchmark, other fit diagnostics were also evaluated for comparative consistency, including the Bayesian Information Criterion (BIC) (Schwarz, 1978), the Final Prediction Error (FPE) (Akaike, 1979), and the Hannan-Quinn Information Criterion (HQIC) (Hannan & Quinn, 1979).

Regarding the experimental setup, the model was trained using two distinct data-splitting strategies: a standard 70/30 train-test split and a rolling window analysis, the latter of which accounts for the dynamic evolution of time series parameters over time.

In the rolling window analysis, the model predicts values for 15 days, while the remaining data are used for training. A comparison between the predicted and actual values of the test dataset is presented in Fig. 3. The model's forecasting performance improves with increasing lag order, but the forecasts do not change significantly once the lag order reaches 35. We obtain similar forecasts between the lag orders of 35 and 40.
Choosing the time period for the data is an important factor in building the model. In additional tests, the data is taken from 2020 onwards and analyzed using the rolling window method for time series models. The model forecasts improve as we increase the lag orders with forecast horizons of 15 and 30. The model has shown improvement especially in predicting the stock indices SPGSCI-CrudeOil and SPGSCI (see the example in Fig. 4).

3.4 Implementation

This study utilizes Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks for time series forecasting, with both models optimized using KerasTuner's Random Search to identify the best hyperparameters. Each model consists of two recurrent layers followed by a dense output layer. The first recurrent layer contains 50 to 200 units, returning sequences to maintain long-term dependencies, while the second recurrent layer has 25 to 100 units to refine feature extraction. Dropout layers, with rates ranging from 0.1 to 0.5, are applied after each recurrent layer to mitigate overfitting. A fully connected Dense layer with one neuron serves as the output layer, and the Adam optimizer is used with a tunable learning rate ranging from 1e-5 to 1e-2 for optimal convergence.

Hyperparameter tuning is performed using KerasTuner's Random Search, exploring configurations of 50 to 200 units for both LSTM and GRU, dropout rates between 0.1 and 0.5, and a learning rate sampled logarithmically between 1e-5 and 1e-2. Each model is trained for 30 epochs with a batch size of 10, optimizing validation loss to select the best-performing configuration.

Figure 3. The actual and forecasted values of the VAR model with the order-lags: 5 and 35, the first difference forecast, and forecast horizon = 15.

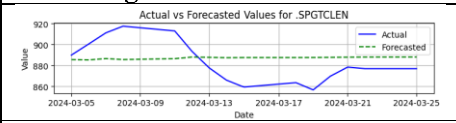
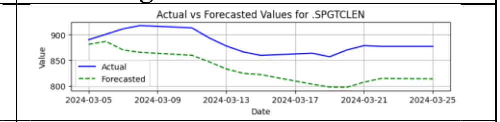
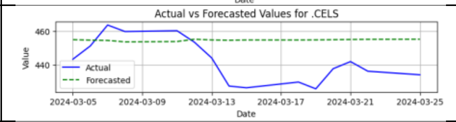
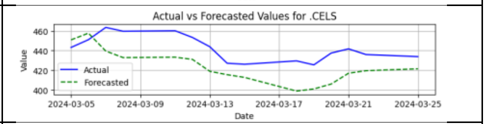
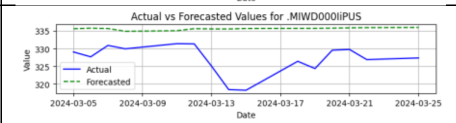
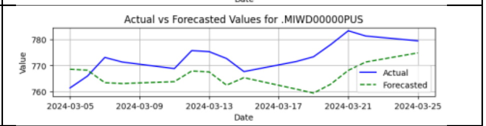
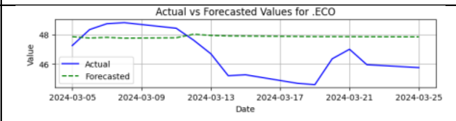
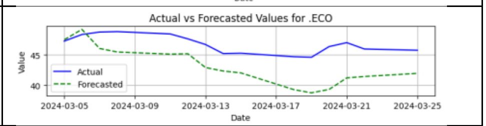
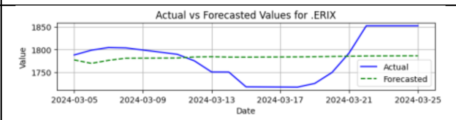
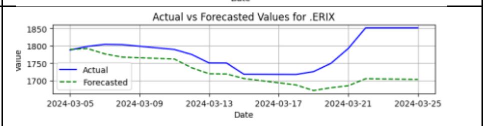
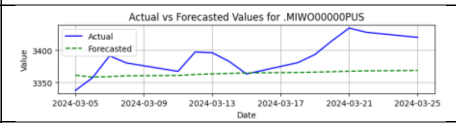
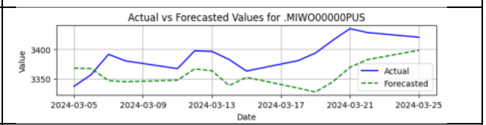
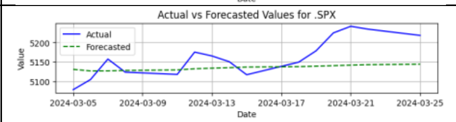
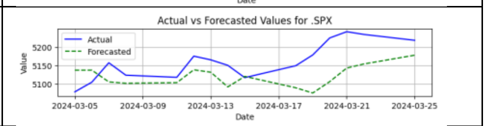
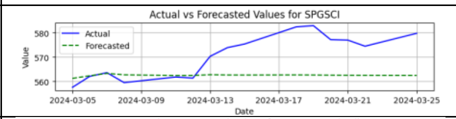
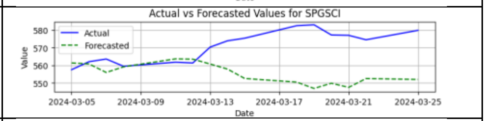
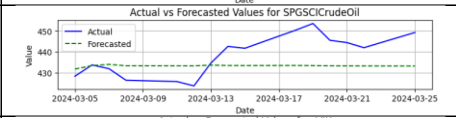
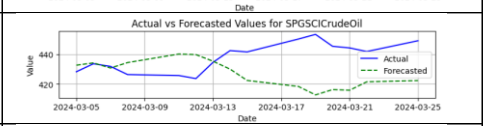
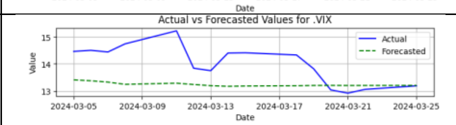
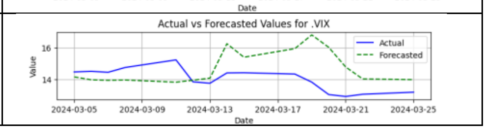
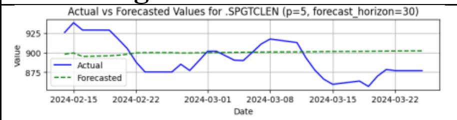
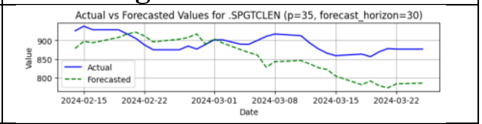
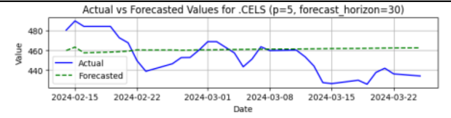
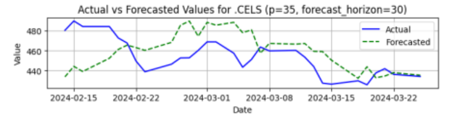
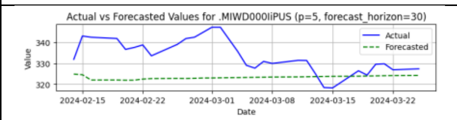
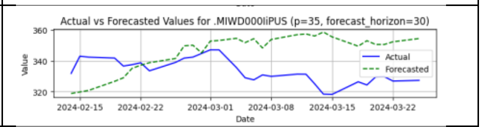
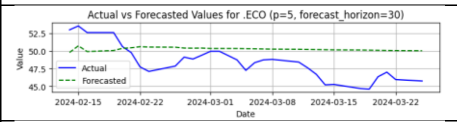
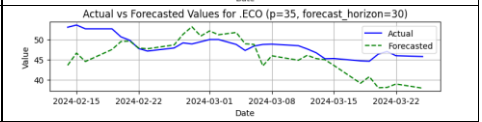
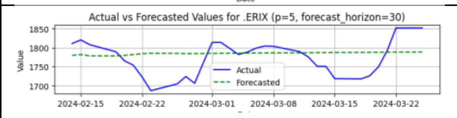
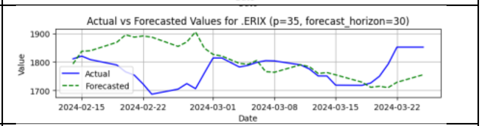
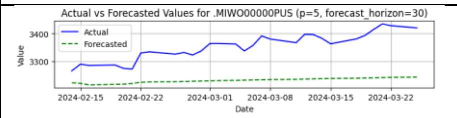
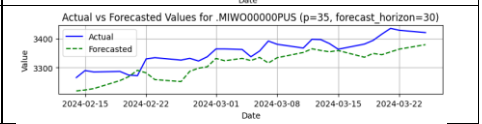
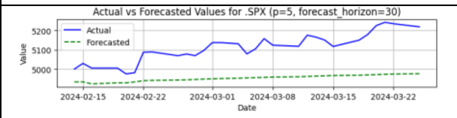
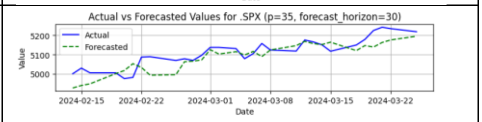
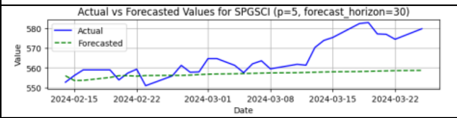
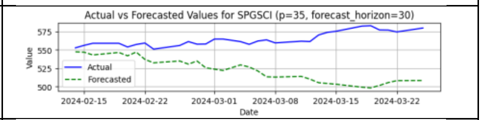
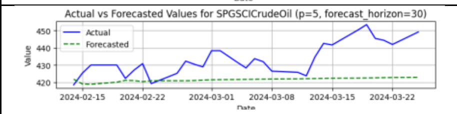
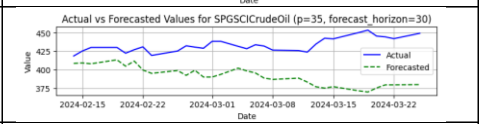
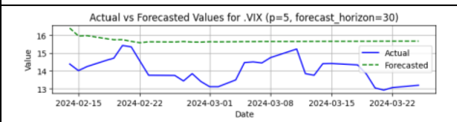
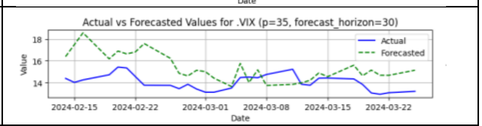
Index	Order lag 5	Order lag 35
SPGTCLEN		
CELS		
MIWD000IiPUS		
ECO		
ERIX		
MIWO00000PUS		
SPX		
SPGSCI		
SPGSCICrudeOil		
VIX		

Figure 4. The actual and forecasted values of the VAR model with the dataset taken from 2020, with lag orders of 5 and 35, the first difference applied, and a forecast horizon = 30.

Index	Order lag 5	Order lag 35
SPGTCLEN		
CELS		
MIWD000IiPUS		
ECO		
ERIX		
MIWO00000PUS		
SPX		
SPGSCI		
SPGSCICrudeOil		
VIX		

3.5 Performance Metrics of Forecasts

The performance metrics, including Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE), are analyzed to evaluate the predictive capabilities of both models. The results indicate that the GRU model generally outperforms the LSTM model in error reduction and achieves more accurate forecasts across multiple target variables.

Table 2 presents a quantitative comparison of the LSTM and GRU models. The GRU model consistently achieves lower RMSE and MAE values, signifying better overall accuracy and reduced forecasting errors. Notably, for target variables such as SPGTCLEN, ECO, and MIWD000IiPUS, GRU significantly outperforms LSTM, for instance, in SPGTCLEN, where the RMSE is reduced from 0.755 (LSTM) to 0.568 (GRU), highlighting the efficiency of GRU in capturing complex patterns with lower deviations from actual values. Similarly, in most cases, the GRU model achieves a lower MAE, such as SPGTCLEN (GRU: 0.417 vs. LSTM: 0.921) and MIWD000IiPUS (GRU: 0.261 vs LSTM: 0.655), further reinforcing its superior predictive performance.

We also observe that MAPE takes values exceeding 1000% especially for some selected indices. Although this is a known mathematical artifact of the MAPE formula when the actual price approaches zero, causing the denominator to vanish and the error to scale disproportionately, MAPE is reported for completeness.

Table 2. Performance Metrics for LSTM and GRU Models

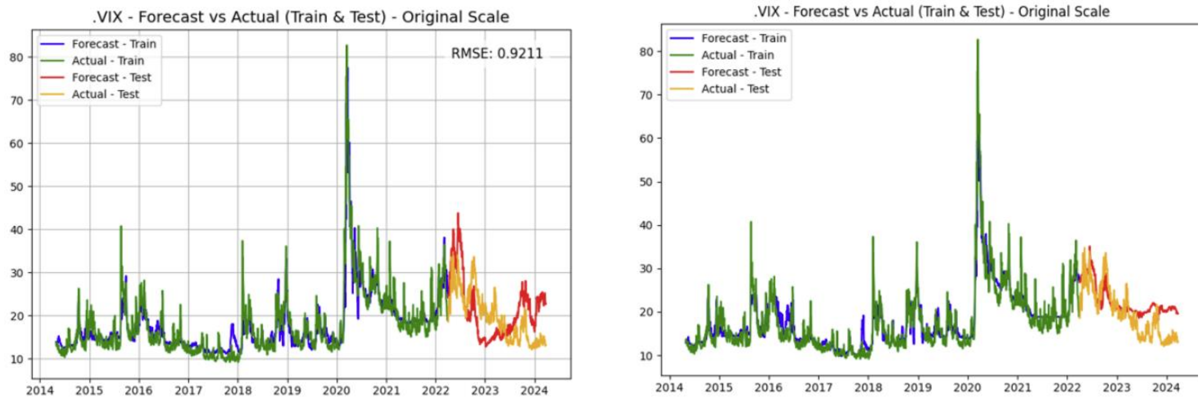
Target Column	LSTM RMSE	LSTM MAE	LSTM MAPE	GRU RMSE	GRU MAE	GRU MAPE
SPGTCLEN	0.755	0.921	1004.160	0.568	0.417	1283.070
CELS	0.476	0.631	349.950	0.390	0.330	379.330
MIWO00000PUS	0.243	0.458	56.350	0.321	0.280	59.580
ECO	0.892	0.923	2660.180	0.835	0.750	2567.060
MIWD000IiPUS	0.713	0.655	115.440	0.332	0.261	76.360
ERIX	0.284	0.469	66.600	0.244	0.192	85.670
SPGSCI	0.173	0.665	60.540	0.142	0.106	56.970
SPGSCICrudeOil	0.306	0.666	88.120	0.148	0.116	73.280
SPX	0.195	0.438	37.730	0.187	0.136	34.720
VIX	0.921	0.958	344.190	0.563	0.480	203.930

However, the LSTM model remains effective in certain scenarios. For example, in the case of MIWO00000PUS, the LSTM model achieves a lower MAPE value than GRU (LSTM: 56.35 vs GRU: 59.58). This suggests that LSTM may be better suited for capturing long-term dependencies in time series data that exhibit subtle but consistent changes. Some other target variables, such as SPX and SPGSCI show nearly equal performance between the two models, with RMSE and MAE values exhibiting minimal differences.

The superiority of GRU over LSTM is further supported by the forecast vs. actual plots (Fig. 5-8). While both models effectively capture long-term trends, the GRU model produces forecasts that are more closely aligned with actual values, particularly during the test phase. The LSTM model, in contrast, exhibits noticeable deviations, especially for highly volatile series such as VIX and ECO, where its RMSE remains high (e.g., 0.921 for VIX). GRU offers better adaptability to fluctuations in these cases, as evidenced by its lower RMSE values (0.563 for VIX).

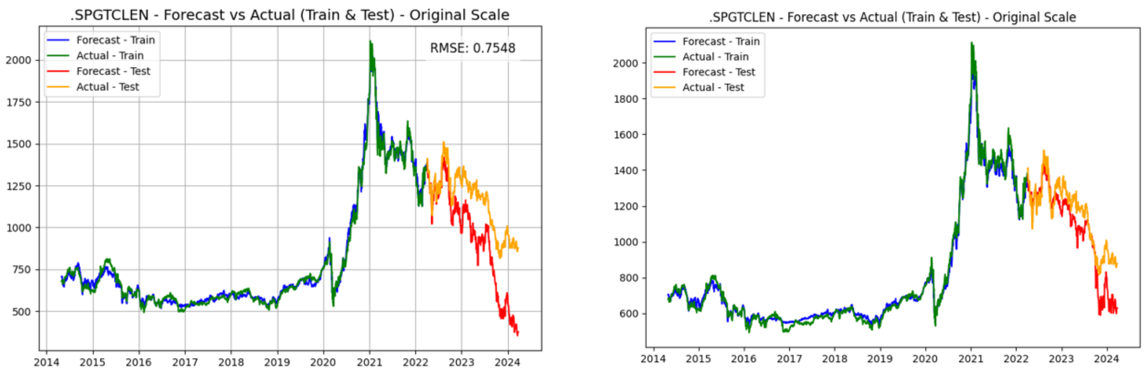
Additionally, the LSTM model occasionally demonstrates overfitting tendencies, performing well on the training set but showing increased errors in the test phase. With its simpler architecture and better convergence properties, the GRU model mitigates this issue by generating smoother and more reliable forecasts. The visualizations reinforce this finding, with GRU demonstrating more stable and less erratic predictions than LSTM, particularly in volatile market conditions.

Figure 5. Comparison of LSTM and GRU Forecasting Performance on VIX



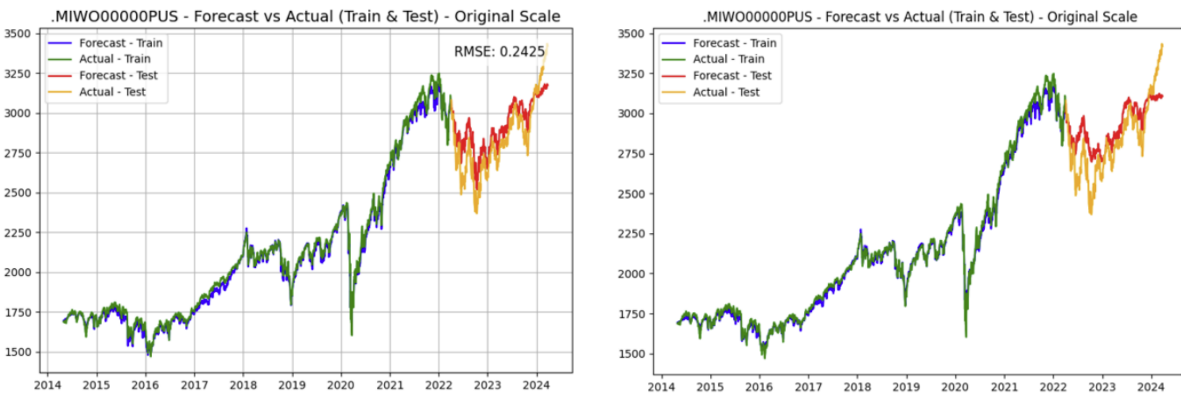
(a) LSTM Forecast for VIX (b) GRU Forecast for VIX

Figure 6. Comparison of LSTM and GRU Forecasting Performance on SPGTCLEN



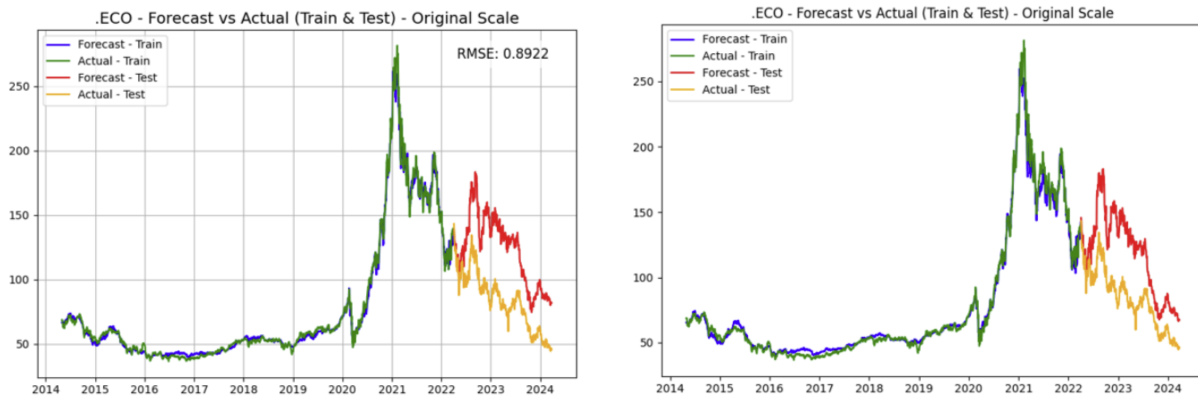
(a) LSTM Forecast for SPGTCLEN (b) GRU Forecast for SPGTCLEN

Figure 7. Comparison of LSTM and GRU Forecasting Performance on MIWO00000PUS



(a) LSTM Forecast for MIWO00000PUS (b) GRU Forecast for MIWO00000PUS

Figure 8. Comparison of LSTM and GRU Forecasting Performance on ECO



(a) LSTM Forecast for ECO (b) GRU Forecast for ECO

4. Conclusions

This study evaluated the predictive capabilities of different machine learning frameworks in the context of the clean energy sector. The comparative analysis reveals that the GRU model consistently outperforms the LSTM model across various performance metrics and experimental settings. GRU's architecture, which combines the forget and input gates into a single update gate, enables it to capture temporal dependencies efficiently while reducing computational overhead. This structural simplicity not only accelerates training but also enhances generalization performance, particularly in environments where data patterns change rapidly. Consequently, GRU demonstrates a superior ability to model short- to medium-term dependencies, making it especially effective for forecasting complex financial time series characterized by high volatility and frequent fluctuations.

Although the LSTM model remains valuable for applications that demand long-term memory retention and nuanced sequence learning, the GRU achieves a more favorable balance between predictive accuracy, training speed, and computational efficiency. The observed reductions in RMSE, MAE, and MAPE across most target variables support the conclusion that the GRU model is the more robust and practical choice, particularly for financial datasets that exhibit strong short-term variability and limited long-term dependencies.

Beyond the technical findings, these results also highlight the importance of adopting transdisciplinary approaches to understanding clean energy markets. By integrating methods and insights from financial econometrics, machine learning, and sustainability science, this research contributes to a more comprehensive understanding of the clean energy transition, that constitutes an essential step in addressing the existential threat posed by climate change.

Furthermore, while much of the existing literature focuses on technical and operational dimensions such as energy efficiency, energy savings, and peak demand management, our findings extend this knowledge to the financial domain, demonstrating how advanced data-driven methods can be useful in predicting with greater accuracy clean energy stock prices. This study therefore underscores the potential of AI-based predictive modeling to inform both investment strategies and policy design aimed at fostering resilient and sustainable energy systems.

Acknowledgements

M. E. De Giuli acknowledges financial support from the European Union – NextGenerationEU, within the framework of the GRINS – Growing Resilient, INclusive and Sustainable (GRINS PE0000018 – Project: SAFE-MLE), and from the Italian Ministry of University and Research (MUR) through the PRIN 2022 program, project “A geo-localized data framework for managing climate risks and designing policies to support sustainable investments” (No. 20229CWYXC).

References

Agakishiev, I., Haerdle, W. K., Kopa, M., Kozmik, K., & Petukhina, A. (2025). Multivariate probabilistic forecasting of electricity prices with trading applications. *Energy Economics*, 141, 108008. https://doi.org/10.1016/j.eneco.2024.108008

Ahmad, A. S., Hassan, M. Y., Abdullah, M. P., Rahman, H. A., Hussin, F., Abdullah, H., & Saidur, R. (2014). A review on applications of ANN and SVM for building electrical energy consumption forecasting. *Renewable and Sustainable Energy Reviews*, 33, 102–109. https://doi.org/10.1016/j.rser.2014.01.069

Akaike, H. (1969). Fitting autoregreesive models for prediction. In Selected Papers of Hirotugu Akaike (pp. 131-135). New York, NY: Springer New York. https://doi.org/10.1007/978-1-4612-1694-0_10

Akaike, H. (2003). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716-723. https://doi.org/10.1007/978-1-4612-1694-0_16

Baruník, J., & Křehlík, T. (2018). Measuring the frequency dynamics of financial connectedness and systemic risk. *Journal of Financial Econometrics*, 16 (2), 271–296. https://doi.org/10.1093/jjfinec/nby001

Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE transactions on neural networks*, 5 (2), 157–166. https://doi.org/10.1109/72.279181

Cevik, E., Cevik, E. I., Dibooglu, S., Cergibozan, R., Bugan, M. F., & Destek, M. A. (2024). Connectedness and risk spillovers between crude oil and clean energy stock markets. *Energy & Environment,* 35 (7), 3319–3339. https://doi.org/10.1177/0958305X231167468

Charda, M., & Pendaraki, K. (2023). Investigating Causal Spillovers among International Stock Markets. *European Journal of Interdisciplinary Studies*, 15(1), 81-92. https://doi.org/10.24818/ejis.2023.06

Cho, K. (2014). *Learning phrase representations using RNN encoder-decoder for statistical machine translation*. arXiv preprint arXiv:1406.1078. https://doi.org/10.48550/arXiv.1406.1078

Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). *Empirical evaluation of gated recurrent neural networks on sequence modeling*. arXiv preprint arXiv:1412.3555. https://doi.org/10.48550/arXiv.1412.3555

Dai, Z., Zhu, H., & Zhang, X. (2022). Dynamic spillover effects and portfolio strategies between crude oil, gold and chinese stock markets related to new energy vehicle. *Energy Economics*, 109, 105959. https://doi.org/10.1016/j.eneco.2022.105959

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a), 427-431. https://doi.org/10.2307/2286348

Diebold, F. X., & Yılmaz, K. (2014). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics*, 182 (1), 119–134. https://doi.org/10.1016/j.jeconom.2014.04.012

Ediger, V. S¸., & Akar, S. (2007). ARIMA forecasting of primary energy demand by fuel in Turkey. *Energy policy*, 35 (3), 1701–1708. https://doi.org/10.1016/j.enpol.2006.05.009

Haghaniat, S., Moosavirad, S. H., & Namjoo, M. R. (2024). Prioritizing project risks by a transdisciplinary approach using the grey ordinal priority approach: A case study of an electricity distribution company. *Transdisciplinary Journal of Engineering & Science*, 15. https://doi.org/10.22545/2024/00266

Hannan, E. J., & Quinn, B. G. (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society: Series B (Methodological)*, 41(2), 190-195. https://www.jstor.org/stable/2985032

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9 (8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Kwiatkowski, D., Phillips, P. C., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?. *Journal of econometrics*, 54(1-3), 159-178. https://doi.org/10.1016/0304-4076(92)90104-Y

Leiria, A., & Martins, P. (2022). A transdisciplinary perspective about climate change: A case of Portugal. *Transdisciplinary Journal of Engineering & Science*, 13. https://doi.org/10.22545/2021/00171

Liang, S., Nguyen, L., & Jin, F. (2018). *A multi-variable stacked long-short term memory network for wind speed forecasting*. In 2018 IEEE international conference on big data (Big Data) (pp. 4561–4564). https://doi.org/10.1109/BigData.2018.8622332

Lindemann, B., Müller, T., Vietz, H., Jazdi, N., & Weyrich, M. (2021). A survey on long short-term memory networks for time series prediction. *Procedia Cirp*, 99, 650–655. https://doi.org/10.1016/j.procir.2021.03.088

Liu, T., & Gong, X. (2020). Analyzing time-varying volatility spillovers between the crude oil markets using a new method. *Energy Economics*, 87, 104711. https://doi.org/10.1016/j.eneco.2020.104711

Liu, X., Lin, Z., & Feng, Z. (2021). Short-term offshore wind speed forecast by seasonal ARIMA-A comparison against GRU and LSTM. *Energy*, 227, 120492. https://doi.org/10.1016/j.energy.2021.120492

Lütkepohl, H. (2013). *Introduction to multiple time series analysis.* Springer Science & Business Media.

Mewada, H., Sundar, L. S., Patel, B., Desai, M., & Khan, A. (2024). A lightweight sequential convolutional neural network for smart grid stability analysis. *Transdisciplinary Journal of Engineering & Science*, 15. https://doi.org/10.22545/2024/00259

Murărașu, I. C. (2023). Analysis of the Volatility of Renewable Sources of Electricity in Romania and the Assessment of Their Capacity to Replace the Conventional Sources. *European Journal of Interdisciplinary Studies*, 15(1), 93-105. https://doi.org/10.24818/ejis.2023.07

Nadirgil, O. (2023). The relationship between the contaminating industries and the european carbon price, machine learning approach. *Journal of Cleaner Production*, 426, 139131. https://doi.org/10.1016/j.jclepro.2023.139131

Pascanu, R., Mikolov, T., & Bengio, Y. (2013). *On the difficulty of training recurrent neural networks*. In International conference on machine learning (pp. 1310–1318). https://doi.org/10.48550/arXiv.1211.5063

Prajapati, S. (2023). Management and optimization of Grid-Tied renewable energy for EV charging station. *Transdisciplinary Journal of Engineering & Science*, 14. https://doi.org/10.22545/2023/00236

Rana, R. (2016). Gated recurrent unit (GRU) for emotion classification from noisy speech. *arXiv preprint arXiv:1612.07778*. https://doi.org/10.48550/arXiv.1612.07778

Raza, M. Q., & Khosravi, A. (2015). A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings. *Renewable and Sustainable Energy Reviews*, 50, 1352–1372. https://doi.org/10.1016/J.RSER.2015.04.065

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 461-464. https://www.jstor.org/stable/2958889

Sharma, G. D., Sarker, T., Rao, A., Talan, G., & Jain, M. (2022). Revisiting conventional and green finance spillover in post-COVID world: Evidence from robust econometric models. *Global Finance Journal*, 51, 100691. https://doi.org/10.1016/j.gfj.2021.100691

Soman, S. S., Zareipour, H., Malik, O., & Mandal, P. (2010). *A review of wind power and wind speed forecasting methods with different time horizons.* In North American power symposium 2010 (pp. 1–8). https://doi.org/10.1109/NAPS.2010.5619586

Stock, J. H., & Watson, M. W. (2001). Vector autoregressions. *Journal of Economic perspectives*, 15(4), 101-115. https://doi.org/101-115. 10.1257/jep.15.4.101

Sulistio, B., Warnars, H. L. H. S., Gaol, F. L., & Soewito, B. (2023). Energy sector stock price prediction using the CNN, GRU & LSTM hybrid algorithm. *2023 international conference on computer science, information technology and engineering (ICCoSITE)*, February, 178-182, https://doi.org/10.1109/ICCoSITE57641.2023.10127847

Tank, A., Covert, I., Foti, N., Shojaie, A., & Fox, E. B. (2021). Neural granger causality. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44 (8), 4267–4279. https://doi.org/10.48550/arXiv.1802.05842

Triantafyllidou, A., Polychronidou, P., & Mantzaris, I. (2024). Renewable and Non-Renewable Energy Consumption and Economy: A Systematic Literature Review for Greece. *European Journal of Interdisciplinary Studies*, *16*(2), 155–174. https://doi.org/10.24818/ejis.2024.17

Xiao, J., Deng, T., & Bi, S. (2024). Comparative analysis of LSTM, GRU, and transformer models for stock price prediction. *Proceedings of the International Conference on Digital Economy, Blockchain and Artificial Intelligence,* August, 103-108, https://doi.org/10.48550/arXiv.2411.05790

Yadav, M. P., Ashok, S., Taghizadeh-Hesary, F., Dhingra, D., Mishra, N., & Malhotra, N. (2024). Uncovering time and frequency co-movement among green bonds, energy commodities and stock market*.* *Studies in Economics and Finance*, 41, 638–659. https://doi.org/10.1108/SEF-03-2023-0126

Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, 159–175. https://doi.org/10.1016/S0925-2312(01)00702-0