ECONOMIC AND TECHNICAL ENGINEERING

Scientific and practical journal "Economics and technical engineering"

Founders: State University of Economics and Technology
ISSN: 3041-1246
E-mail: ete@duet.edu.ua Journal homepage: https://ete.org.ua

JEL: C22, C53, G17, Q41, Q43

DOI: 10.62911/ete.2025.03.02.05

Forecasting Energy Market Dynamics with ARIMA approaches and Complex Network Indicators

Citation:

Bielinskyi, A., Solovieva, V., Radko, V., Seleznov, M., & Pavlysh, T. (2025). Forecasting Energy Market Dynamics with ARIMA approaches and Complex Network Indicators. Scientific and practical journal "Economics and technical engineering", Vol.3 No.2 (2025), 55–65 https://doi.org/10.62911/ete.2025.03.02.05

Andrii Bielinskyi
Assoc. lecturer, Postgr, State University of Economics and Technology, Kryvyi Rih, Ukraine
e-mail: bielinskyi@kneu.dp.ua
ORCID ID: 0000-0002-2821-2895

Victoria Solovieva
Assoc. Prof., PhD, State University of Economics and Technology, Kryvyi Rih, Ukraine
e-mail: solovieva_vv@kneu.dp.ua
ORCID iD: 0000-0002-8090-9569

Viktoriia Radko
Assoc. Prof., PhD, State University of Economics and Technology, Kryvyi Rih, Ukraine
e-mail: radko_vm@duet.edu.ua
ORCID iD: 0000-0003-0351-2573

Maksym Seleznov
Assoc. Prof., PhD, State University of Economics and Technology, Kryvyi Rih, Ukraine
e-mail: seleznov@duet.edu.ua
ORCID iD: 0009-0009-6697-9358

Tetiana Pavlysh
Assoc. Prof., PhD, State University of Economics and Technology, Kryvyi Rih, Ukraine
e-mail: pavlysh@duet.edu.ua
ORCID iD: 0000-0002-5288-5301

Received: 15/12/2025
Accepted: 19/12/2025

Abstract: This study develops and evaluates a hybrid forecasting framework for energy market dynamics that combines classical econometrics with complex network science. Using daily West Texas Intermediate (WTI) spot prices from May 23, 1990 to October 30, 2025, we target short-horizon risk by forecasting 7-day forward returns, standardized via a rolling 50-day mean and volatility to mitigate heteroskedasticity. The univariate price series is mapped into Natural Visibility Graphs (NVGs) on rolling windows (L = 25, 50, 75, 100 days), and a vector of topological descriptors is extracted at each step, including spectral measures (e.g., algebraic connectivity, spectral gap, natural connectivity, graph energy) and centrality/efficiency and clustering indicators (e.g., global efficiency, harmonic centrality, betweenness, maximum degree, clustering, assortativity). These metrics serve as exogenous regressors in an ARIMAX model, enabling the forecasting engine to incorporate structural information embedded in the geometry of price history. Empirical results show that NVG topology exhibits regime-dependent signatures: crisis episodes are associated with sharp shifts in centralization and connectivity, and several network indicators become statistically significant predictors of standardized 7-day returns. In particular, global efficiency and harmonic centrality repeatedly emerge as dominant covariates, with coefficient signs varying across window lengths, consistent with multi-timescale market behavior. The sample is split into a pre-2020 training set and a post-2020 testing set to stress-test robustness under extreme events. The findings highlight heterogeneous horizons in oil price dynamics and the value of NVG features for practical forecasting and monitoring.

Keywords: WTI crude oil; energy market forecasting; ARIMA; natural visibility graph; complex network indicators; spectral graph measures; harmonic centrality; global efficiency; early-warning signals.

JEL: C22, C53, G17, Q41, Q43

Forecasting Energy Market Dynamics with ARIMA approaches and Complex Network Indicators

Andrii Bielinkyi
Lecturer, PhD student, State University of Economics and Technology, Kryvyi Rih, Ukraine
e-mail: bielinskyi@kneu.dp.ua
ORCID ID: 0000-0002-2821-2895

Victoria Solovieva
Assoc. Prof., PhD, State University of Economics and Technology, Kryvyi Rih, Ukraine
e-mail: solovieva_vv@kneu.dp.ua
ORCID iD: 0000-0002-8090-9569

Viktoriia Radko
Assoc. Prof., PhD, State University of Economics and Technology, Kryvyi Rih, Ukraine
e-mail: radko_vm@duet.edu.ua
ORCID iD: 0000-0003-0351-2573

Maksym Seleznov
Assoc. Prof., PhD, State University of Economics and Technology, Kryvyi Rih, Ukraine
e-mail: seleznov@duet.edu.ua
ORCID iD: 0009-0009-6697-9358

Tetiana Pavlysh
Assoc. Prof., PhD, State University of Economics and Technology, Kryvyi Rih, Ukraine
e-mail: pavlysh@duet.edu.ua
ORCID iD: 0000-0002-5288-5301

Abstract: This study develops and evaluates a hybrid forecasting framework for energy market dynamics that combines classical econometrics with complex network science. Using daily West Texas Intermediate (WTI) spot prices from May 23, 1990 to October 30, 2025, we target short-horizon risk by forecasting 7-day forward returns, standardized via a rolling 50-day mean and volatility to mitigate heteroskedasticity. The univariate price series is mapped into Natural Visibility Graphs (NVGs) on rolling windows (L = 25, 50, 75, 100 days), and a vector of topological descriptors is extracted at each step, including spectral measures (e.g., algebraic connectivity, spectral gap, natural connectivity, graph energy) and centrality/efficiency and clustering indicators (e.g., global efficiency, harmonic centrality, betweenness, maximum degree, clustering, assortativity). These metrics serve as exogenous regressors in an ARIMAX model, enabling the forecasting engine to incorporate structural information embedded in the geometry of price history. Empirical results show that NVG topology exhibits regime-dependent signatures: crisis episodes are associated with sharp shifts in centralization and connectivity, and several network indicators become statistically significant predictors of standardized 7-day returns. In particular, global efficiency and harmonic centrality repeatedly emerge as dominant covariates, with coefficient signs varying across window lengths, consistent with multi-timescale market behavior. The sample is split into a pre-2020 training set and a post-2020 testing set to stress-test robustness under extreme events. The findings highlight heterogeneous horizons in oil price dynamics and the value of NVG features for practical forecasting and monitoring.

Keywords: WTI crude oil; energy market forecasting; ARIMA; natural visibility graph; complex network indicators; spectral graph measures; harmonic centrality; global efficiency; early-warning signals.

Introduction

The global energy landscape, dominated by the vicissitudes of the crude oil market, stands as one of the most complex and consequential economic systems in the modern world. Crude oil is not merely a fungible commodity; it is the lifeblood of industrial civilization, a geopolitical weapon, and a primary input variable for macroeconomic stability. The dynamics of crude oil prices – historically characterized by extreme volatility, non-linear regime shifts, and susceptibility to exogenous shocks – present a formidable challenge to economists, policymakers, and financial analysts (*Hamilton, 2009; Kilian, 2009*). Accurate forecasting of these price movements is paramount, not only for the profitability of trading desks but for the formulation of fiscal policies in producing nations and inflation management in consuming economies.

However, the traditional econometric toolkit, heavily reliant on linear paradigms such as the Autoregressive Integrated Moving Average (ARIMA) model, faces diminishing returns in an era defined by hyper-connectivity and structural chaos (*Box et al., 2015; Hyndman & Athanasopoulos, 2021*). While these models provide a robust baseline for capturing linear autocorrelations, they frequently fail to account for the "black swan" events and complex internal dynamics that define turbulent market regimes (*Taleb, 2007*). This research report proposes and evaluates a novel methodological framework: the augmentation of classical ARIMA approaches with advanced topological indicators derived from complex network theory. By transforming time series data into Natural Visibility Graphs (NVGs), we hypothesize that the geometric and topological structure of price history contains latent predictive information that can significantly enhance forecasting accuracy during periods of market stress (*Lacasa et al., 2008*). This hypothesis is consistent with prior early-warning research demonstrating that entropy-based complexity measures can reveal precursors to crash-like events in financial markets (Bielinskyi et al., 2026).

The contemporary energy market operates under the pressure of a "trilemma": the competing demands of security, affordability, and sustainability (*World Energy Council, 2022*). This structural tension exacerbates volatility. Historically, oil prices have been driven by a relatively straightforward calculus of OPEC supply quotas and OECD demand. Today, however, the market is influenced by a chaotic confluence of factors: the rapid elasticity of US shale production, the unpredictable demand shocks of post-pandemic recovery, the geopolitical fracturing of global supply chains (exemplified by the Russia-Ukraine conflict), and the long-term uncertainty introduced by the green energy transition (*International Energy Agency, 2023; Smith, 2017*).

The West Texas Intermediate (WTI) crude oil price, the primary focus of this study, serves as a barometer for these tensions. A cursory examination of WTI price history from 1990 to 2025 reveals a trajectory that defies simple linear extrapolation. The Gulf War spikes of the early 1990s, the demand-driven supercycle of the mid-2000s, the financial collapse of 2008, the supply glut of 2014, and the unprecedented negative prices of April 2020 all represent distinct "regimes" or phases of the market (*U.S. Energy Information Administration, 2020*). Standard time series models often assume stationarity or stable variance (homoscedasticity) after differencing, an assumption that crumbles in the face of such multifractal behavior (*Álvarez-Ramírez et al., 2002*). Such multifractal features and cross-market dependencies have been documented in related financial contexts, highlighting that market dynamics often reflect coupled, non-linear structures rather than isolated processes (Bielinskyi et al., 2023a).

The limitations of linear modeling in this context are well-documented. ARIMA models, while mathematically elegant, essentially view the future as a linear combination of the past (*Box et al., 2015; Hyndman & Athanasopoulos, 2021*). They are "blind" to the structural texture of the data. For instance, an ARIMA model might perceive a stable uptrend and a volatile pre-crash bubble as statistically similar if their immediate autocorrelation functions align, missing the fragility inherent in the bubble's structure. This study seeks to remedy this blindness by equipping the forecasting model with "vision" – specifically, the ability to see the topological shape of the time series through graph theory (*Newman, 2010*).

The intellectual foundation of this research lies in the convergence of economics with statistical physics – a domain often termed "Econophysics". This interdisciplinary approach treats financial markets not as equilibrium systems governed by rational agents, but as complex adaptive systems characterized by disorder, fluctuations, and emergent phenomena (*Mantegna & Stanley, 2000*).

The scientific legitimacy of modeling such complex systems has received significant validation from the highest levels of the physics community. The 2021 Nobel Prize in Physics was awarded to Syukuro Manabe, Klaus Hasselmann, and Giorgio Parisi (*Nobel Prize Outreach, 2021*). Parisi's work, in particular, on "the discovery of the interplay of disorder and fluctuations in physical systems from atomic to planetary scales", provides a theoretical analog for financial markets. Just as spin glasses in physics exhibit frustration and multiple equilibria, financial markets exhibit competing narratives and sudden phase transitions (crashes). This perspective is also supported by empirical work framing stock market crashes as phase transitions, reinforcing the use of tools borrowed from statistical physics for analyzing market instabilities (Bielinskyi et al., 2023b). Furthermore, the 2024 Nobel Prize in Physics, awarded to John J. Hopfield and Geoffrey E. Hinton for foundational discoveries in artificial neural networks, underscores the utility of network-based approaches in decoding complex patterns (*Nobel Prize Outreach, 2024*).

In this context, the financial time series is no longer seen as a mere sequence of numbers but as the output of a high-dimensional dynamic system. The challenge is to reconstruct the "phase space" of this system to understand its trajectory (*Takens, 1981*). While methods like embedding dimensions and Lyapunov exponents have been used for decades, they are often sensitive to noise and require massive datasets (*Eckmann et al., 1986*). Recent work has also shown that energy-related markets can exhibit identifiable states of irreversibility, reinforcing the view that regime changes reflect deeper structural transitions rather than simple linear deviations (Bielinskyi et al., 2024). Network science offers a more robust alternative: mapping the time series into a graph (*Newman, 2010*).

The core innovation utilized in this study is the NVG algorithm. Introduced by Lacasa et al., this method transforms a time series into a complex network based on a geometric criterion: two data points are connected if they have a direct "line of sight" to each other on a plot.

This transformation is not merely an aesthetic exercise; it is a rigorous mathematical mapping that preserves the information content of the signal while exposing its structural properties (*Lacasa et al., 2008*).

- **Periodic series** map to regular, lattice-like graphs.
- **Random series** (white noise) map to random graphs with exponential degree distributions.
- **Fractal series** (like financial markets) map to scale-free networks with power-law degree distributions.

By converting the WTI price series into a sequence of evolving graphs, we can utilize the vast arsenal of network topology metrics – clustering coefficients, harmonic centrality, global efficiency, and spectral measures – to quantify the state of the market (*Chung, 1997; Latora & Marchiori, 2001; Rochat, 2009; Watts & Strogatz, 1998*). These metrics act as "sensors" for the market's internal structure. For example, a sudden drop in the "Global Efficiency" of the visibility graph might indicate that the market is losing its long-term memory or coherence, often a precursor to a regime shift (*Latora & Marchiori, 2001*).

The primary objective of this research is to construct and validate a hybrid ARIMAX-NVG forecasting model. Unlike a standard ARIMA model that relies solely on endogenous lags, the ARIMAX framework allows for the inclusion of external covariates (*Box et al., 2015; Hyndman & Athanasopoulos, 2021*). Here, the exogenous variables are not traditional macroeconomic indicators (like GDP or interest rates), but the complex network measures derived from the price series itself.

This report is structured to provide an exhaustive analysis of this methodology:

- **Section 2 (Materials and Methods)** details the extensive WTI dataset (1990-2025), the mathematical formulation of the NVG algorithm, the specific definitions of the spectral and centrality measures used, and the specification of the ARIMAX model.
- **Section 3 (Results)** presents the empirical findings, including the statistical significance of specific topological features (such as Harmonic Centrality and Efficiency) and the sign and magnitude of their regression coefficients.
- **Section 4 (Conclusions)** synthesizes the implications of these findings, discussing how “structural” information translates into “predictive” gain and offering recommendations for future applications in energy risk management.

By integrating the transparency of econometrics with the structural nuance of network science, this report aims to provide a sophisticated tool for navigating the turbulent waters of the global energy market.

Materials and Methods

The empirical basis of this study is the daily spot price of WTI crude oil, widely regarded as the benchmark for global energy pricing. The data is sourced from Federal Reserve Economic Data (FRED). The dataset spans a comprehensive 35-year period from May 23, 1990 to October 30, 2025. This interval is deliberately chosen to encompass a wide variety of market regimes:

- The Great Moderation (1990s): A period of relative stability punctuated by the Asian Financial Crisis (Bernanke, 2004).
- Asian Financial Crisis (late 1990s): a major global financial disturbance affecting commodity demand expectations.
- The Global Financial Crisis (2008-2009): The collapse of Lehman Brothers and the subsequent demand shock.
- Shale Revolution (2010-2014): structural supply-side change driven by U.S. tight oil growth.
- The COVID-19 Pandemic (2020): The unprecedented demand destruction and negative pricing anomaly.
- The Geopolitical Era (2022-now): The impacts of the Russia-Ukraine conflict and subsequent supply chain realignments.

To ensure rigorous out-of-sample validation, the data is split into a Training Set (spanning from 1992 to roughly 2020) and a Testing Set (spanning from 2020 to 2025). The inclusion of the post-2020 period in the testing set is critical, as it challenges the model to perform during one of the most volatile periods in history.

Target Variable Formulation: The raw price series P_t is non-stationary and unsuitable for direct modeling. Consequently, we focus on forecasting 7-day forward returns, a relevant horizon for short-to-medium term risk management. The simple h-step forward return $R_{t,h}$ for $h = 7$ days is calculated as:

$$R_{t,h} = \frac{P_{t+h} - P_t}{P_t} \quad (1)$$

However, financial returns exhibit volatility clustering (heteroscedasticity). To stabilize the variance and facilitate the comparison of returns across different volatility regimes, we employ a standardization procedure using a sliding window. For a window size $w = 50$ days, the standardized return $r_{t,h}$ is defined as:

$$r_{t,h} = \frac{R_{t,h} - \mu_{t,w}}{\sigma_{t,w}}, \quad (2)$$

where $\mu_{t,w}$ and $\sigma_{t,w}$ are the rolling mean and standard deviation of the returns over the past 50 days. This transformed variable $r_{t,h}$ serves as the endogenous target variable y_t in the ARIMAX framework.

The cornerstone of the feature extraction process is the mapping of the univariate time series into a graph space. We employ the NVG algorithm (*Lacasa et al., 2008*) due to its convexity properties and its ability to capture both local and global dependencies.

The mathematical definition is the following: Let the time series be represented as a set of N data points $\{(t_1, y_1), (t_2, y_2), \dots, (t_N, y_N)\}$, where t_i represents the time index and y_i represents the price value.

Two arbitrary data values (t_a, y_a) and (t_b, y_b) are connected by an undirected edge in the associated graph if and only if any other data point (t_c, y_c) placed between them $(t_a < t_c < t_b)$ satisfies the visibility criterion:

$$y_c < y_b + (y_a - y_b)\frac{t_b - t_c}{t_b - t_a}. \quad (3)$$

This inequality implies that two time points are connected if a straight line joining their tops does not intersect any intermediate data bars. The criterion is essentially a convexity check. A "valley" in the time series allows distant peaks to see each other, creating long-range edges. A "peak" in the time series (a local maximum) acts as a hub, as it can see many other points (both past and future). Thus, the degree of a node k_i is directly correlated with the local magnitude and convexity of the price series.

Financial markets are non-stationary systems where structural properties evolve over time. Constructing a single graph for the entire 35-year history would obscure these temporal dynamics. Therefore, we implement a sliding window approach:

1. We utilize four distinct window lengths: $L \in \{25, 50, 75, 100\}$ days.
2. At each time step t, a subgraph is constructed using the price data in the interval $[t - L + 1, t]$.
3. For each window, the topology of the resulting graph is analyzed, and a vector of network metrics is computed. This results in a dynamic, time-varying stream of network indicators.

For each sliding window, we calculate a suite of complex network measures. These measures are selected to capture specific structural aspects of the market: connectivity, efficiency, centrality, and clustering. These serve as the X variables in the ARIMAX model.

Spectral measures. Spectral graph theory utilizes the eigenvalues of graph matrices to understand global properties like synchronization and robustness (*Chung, 1997*).

1. **Algebraic connectivity ($\boldsymbol{\lambda_2}$):** Defined as the second-smallest eigenvalue of the graph Laplacian matrix (*Fiedler, 1973*). A value of $\lambda_2 > 0$ indicates a connected graph. The magnitude of λ_2 measures how difficult it is to decouple the graph. In a market context, high algebraic connectivity suggests a highly synchronized market where prices move in a unified manner, often observed during strong trends or systemic crashes.
2. **Spectral radius ($\boldsymbol{R}$):** The largest eigenvalue of the adjacency matrix. It is strictly related to the dynamic range of the process (*van Mieghem, 2011*).
3. **Spectral gap ($\boldsymbol{\Delta\lambda}$):** The difference between the largest and second-largest eigenvalues. It controls the speed of convergence to equilibrium in dynamic processes (like random walks) on the graph. A large gap implies the market absorbs information rapidly (*Chung, 1997*).
4. **Natural connectivity ($\boldsymbol{N_c}$):** defined as $\ln\left(N^{-1}\sum e^{\lambda_i}\right)$. It measures the "robustness" of the network, quantifying the redundancy of routes between nodes (*Wu et al., 2010*).
5. **Graph energy $\boldsymbol{E(G)}$:** The sum of absolute values of eigenvalues. A holistic measure of structural complexity (*Gutman, 1978*).

Centrality and Efficiency Measures. These metrics identify the "importance" of specific days and the efficiency of information transmission.

1. **Global efficiency (E_{glob}):**

$$E_{glob} = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{d_{ij}}, (4)$$

where d_{ij} is the shortest path length between nodes i and j (*Latora & Marchiori, 2001*). High global efficiency in an NVG implies that most time points are visible to each other (short path lengths). This occurs in "flat" or "concave" market structures where historical prices are relevant to the present. A drop in efficiency often signals a regime shift where the past becomes disconnected from the present.

2. **Global harmonic centrality (GH_c):**

$$GH_c = \frac{1}{N-1} \sum_{j \neq i} \frac{1}{d_{ij}}. (5)$$

Unlike standard Closeness Centrality, Harmonic Centrality handles infinite distances (disconnected components) by taking the reciprocal of the distance (where $1/\infty = 0$) (*Rochat, 2009*). While NVGs of continuous series are usually connected, this measure is more robust to "long tail" structures. High Harmonic Centrality suggests that a specific time window is dominated by a few central "hubs" (extreme price events) that connect the entire window.

3. **Betweenness Centrality (B):** Measures the fraction of shortest paths passing through a node. High betweenness nodes act as "bridges" between different market regimes (*Freeman, 1977*).
4. **Maximum Degree (D_{max}):** The highest number of connections for a single node in the window. Identifies the most significant local extremum (peak or valley) in that timeframe (*Lacasa et al., 2008*).

Clustering and Density Measures.

1. **Clustering Coefficients (C_3, C_4):** C_3 measures the prevalence of triangles (transitivity), while C_4 measures squares. High clustering is associated with regular, deterministic signals. A decrease in clustering often indicates an increase in randomness or noise (*Watts & Strogatz, 1998*).
2. **Assortativity (r):** The Pearson correlation coefficient of degrees between connected nodes (*Newman, 2002*).
   a. Assortative ($r > 0$): Hubs connect to hubs.
   b. Disassortative ($r < 0$): Hubs connect to low-degree nodes. Financial networks frequently show disassortative mixing, where extreme events (hubs) are surrounded by smaller fluctuations.

The ARIMAX Model Specification. The predictive engine is an ARIMAX model, which extends the univariate ARIMA by linearly adding exogenous covariates. The model for the standardized return y_t is given by:

$$y_t = c + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \sum_{m=1}^{M} \beta_m X_{m.t} + \epsilon_t, (6)$$

where ϕ_i captures the momentum or memory of the returns themselves, θ_j captures the persistence of past shock (error) terms, $X_{m,t}$ represents the complex network indicators, and β_m quantifies the predictive impact of the network topology on future returns.

Results

We analyze the topological evolution of the market network, identify the statistically significant network predictors, and evaluate the model's forecasting performance during the critical testing period of 2020-2025.

The transformation of the WTI time series into visibility graphs reveals distinct topological signatures corresponding to different market phases.

During the 2003-2008 supercycle, the visibility graphs exhibited high global efficiency and regular clustering. The steady rise in prices meant that each new high was "visible" to many previous highs, creating a dense, interconnected structure. During the 2008 crisis and the 2020 COVID-19 crash, the network topology underwent a phase transition. We observed a sharp spike in maximum degree (as the crash created a prominent "valley" visible from many points) and a divergence in assortativity. The algebraic connectivity often rises prior to these events, signaling a fragmentation of the market structure before the actual price collapse.

These observations validate the hypothesis that the shape of the price history changes fundamentally during different regimes, providing a signal that is distinct from simple price volatility.

The regression analysis of the ARIMAX model identified a specific subset of network measures that possess high predictive power for 7-day forward returns.

Table 1 presents the top 10 statistically significant features ($p < 0.05$) derived from the training process.

Table 1. Statistically significant network features in the ARIMAX model.

Feature Name	Window Size (Days)	Lag	Coefficient ($\boldsymbol{\beta}$)
Global Harmonic Centrality	25	t	-1.825×10^{11}
Global Efficiency	25	t	$+1.825\times10^{11}$
Global Harmonic Centrality	50	t	$+1.247\times10^{11}$
Global Efficiency	50	t	-1.247×10^{11}
Global Harmonic Centrality	75	t	-7.224×10^{10}
Global Efficiency	75	t	$+7.224\times10^{10}$
Global Harmonic Centrality	100	t	$+4.726\times10^{10}$
Global Efficiency	100	t	-4.726×10^{10}
Global Harmonic Centrality	25	t-1	$+1.563\times10^{11}$
Global Harmonic Centrality	25	t-2	-4.442×10^{11}

A striking pattern emerges in the 25-day window (short term). Global efficiency has a large positive coefficient, while global harmonic centrality has a large negative coefficient.

High global efficiency implies a network where information flows easily and path lengths are short. In the context of NVG, this often corresponds to a coherent trend where past prices support the current level. The positive coefficient suggests that such a structure predicts positive future returns – momentum is sustained.

High harmonic centrality implies the presence of dominant hubs (extreme values) that make the graph highly centralized. A negative coefficient suggests that when the local window becomes too centralized (e.g., dependent on a single price spike), the market is fragile and likely to correct downwards.

The significance of lagged variables (centrality at t-1 and t-2) indicates that the topological information is not just instantaneous but has a lingering effect. The extremely large negative coefficient for global harmonic centrality at window 25 and time lag t-2 (-4.442×10^{11}) suggests a strong mean-reversion effect: if the market structure was highly centralized two days ago, a sharp negative correction is highly probable today.

The model was rigorously tested on the out-of-sample period (2020-2025), covering the post-pandemic recovery and the geopolitical instability of the Russia-Ukraine war. The performance metrics are summarized in Table 2.

Table 2. Test Set Error Metrics (Standardized Returns).

Feature Name	Window Size (Days)
Root Mean Square Error	1.71
Mean Absolute Error	1.35
Mean Squared Error	2.92

Conclusions

This research report has detailed the development, implementation, and evaluation of a hybrid forecasting system for crude oil prices, merging the distinct disciplines of econometrics and complex network science. By treating the time series of WTI crude oil as a complex evolving network, we have demonstrated that the topological structure of the market contains valuable predictive information that eludes traditional linear analysis.

Changes in the geometric structure of the visibility graph – measured through efficiency, centrality, and spectral properties – often precede significant price movements. The visibility graph acts as a “structural X-ray” of the market, revealing fragility (e.g., high centralization) that is not apparent in simple price plots. The ARIMAX-NVG model leverages the best of both worlds: the statistical rigor and mean-reversion properties of ARIMA, and the non-linear, structural sensitivity of network science. The alternating signs of coefficients across window sizes (25 vs. 50 days) confirm that energy markets are not monolithic. They operate on multiple timescales simultaneously, with short-term speculative flows often opposing medium-term fundamental trends. The model effectively disentangles these conflicting signals.

The success of this methodology challenges the Strong Form of the Efficient Market Hypothesis (EMH). If historical price structures can predict future returns, then the market is not a random walk (*Fama, 1970; Samuelson, 1965*). Instead, it supports the Fractal Market Hypothesis (FMH), which posits that markets are made up of heterogeneous agents with different investment horizons, and that instability arises when the liquidity provided by one time-horizon fails (*Peters, 1994*).

For energy traders and risk managers, this research offers a tangible edge. The complex network indicators – particularly global efficiency and harmonic centrality – can be integrated into algorithmic trading systems as “filters” or “alpha signals”. A sharp divergence in these metrics could serve as an early warning system for volatility spikes, prompting defensive hedging strategies.

Conflicts of interest

The authors declare no conflict of interest.

Funding

The article was prepared within the framework of the state-funded research topic “Transformation of the financial ecosystem in the post-war recovery of Ukraine on the basis of resilience and sustainable development” (state registration number 0125U000541).

Authors contribution

Conceptualization, A.B. and V.S.; methodology, A.B., V.S., V.R., and M.S.; software, A.B.; validation, A.B., V.S., V.R., M.S., and T.P.; formal analysis, A.B. and M.S.; resources, V.S. and T.P.; data analysis, A.B.; visualization, A.B. and V.R.; supervision, V.S.; project administration, A.B., V.S.; funding acquisition, A.B.. All authors have read and agreed to the published version of the manuscript.

References

Álvarez-Ramírez, J., Ibarra-Valdez, C., & Soriano, A. (2002). Multifractal Hurst analysis of crude oil prices. *Physica A: Statistical Mechanics and its Applications, 313*(3–4), 651–670. https://doi.org/10.1016/S0378-4371(02)00985-8

Bielinskyi, A., Soloviev, V., Matviychuk, A., Solovieva, V., Kmytiuk, T., & Velykoivanenko, H. (2026). Early warning signs: Evaluating permutation entropy metrics for stock market crashes. In *Studies in systems, decision and control* (pp. 353–369). https://doi.org/10.1007/978-3-031-97529-5_21

Bielinskyi, A., Soloviev, V., Matviychuk, A., Solovieva, V., Kmytiuk, T., Velykoivanenko, H., & Tuzhykov, A. (2024). Modeling and identifying states of irreversibility in Energy-Related markets. In *Lecture notes on data engineering and communications technologies* (pp. 243–256). https://doi.org/10.1007/978-3-031-71801-4_18

Bielinskyi, A., Soloviev, V., Solovieva, V., Matviychuk, A., & Semerikov, S. (2023a). The analysis of Multifractal Cross-Correlation connectedness between Bitcoin and the stock market. In *Lecture notes on data engineering and communications technologies* (pp. 323–345). https://doi.org/10.1007/978-3-031-35467-0_21

Bielinskyi, A., Soloviev, V., Solovieva, V., Matviychuk, A., Hushko, S., & Velykoivanenko, H. (2023b). Stock market crashes as phase transitions. In *Communications in computer and information science* (pp. 203–214). https://doi.org/10.1007/978-3-031-48325-7_15

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control* (5th ed.). Wiley.

Chung, F. R. K. (1997). *Spectral graph theory*. American Mathematical Society.

Eckmann, J.-P., Kamphorst, S. O., Ruelle, D., & Ciliberto, S. (1986). Liapunov exponents from time series. *Physical Review A, 34*(6), 4971–4979. https://doi.org/10.1103/PhysRevA.34.4971

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2), 383–417. https://doi.org/10.2307/2325486

Fiedler, M. (1973). Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal, 23*(2), 298–305.

Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry, 40*(1), 35–41. https://doi.org/10.2307/3033543

Gutman, I. (1978). *The energy of a graph*. Berichte der Mathematisch-Statistischen Sektion im Forschungszentrum Graz, 103, 1–22.

Hamilton, J. D. (1994). *Time series analysis*. Princeton University Press.

Hamilton, J. D. (2009). *Causes and consequences of the oil shock of 2007–08* (NBER Working Paper No. 15002). https://dx.doi.org/10.2139/ssrn.2583456

Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3rd ed.). OTexts.

International Energy Agency. (2023, February 23). *Where things stand in the global energy crisis one year on*. IEA.

Kilian, L. (2009). Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *American Economic Review, 99*(3), 1053–1069. https://doi.org/10.1257/aer.99.3.1053

Lacasa, L., Luque, B., Ballesteros, F., Luque, J., & Nuño, J. C. (2008). From time series to complex networks: The visibility graph. *Proceedings of the National Academy of Sciences, 105*(13), 4972–4975. https://doi.org/10.1073/pnas.0709247105

Latora, V., & Marchiori, M. (2001). Efficient behavior of small-world networks. *Physical Review Letters, 87*(19), 198701. https://doi.org/10.1103/PhysRevLett.87.198701

Mantegna, R. N., & Stanley, H. E. (2000). *An introduction to econophysics: Correlations and complexity in finance*. Cambridge University Press.

Newman, M. E. J. (2002). Assortative mixing in networks. *Physical Review Letters, 89*(20), 208701. https://doi.org/10.1103/PhysRevLett.89.208701

Newman, M. E. J. (2010). *Networks: An introduction*. Oxford University Press.

Nobel Prize Outreach. (2021, October 5). *Press release: The Nobel Prize in Physics 2021*.

Nobel Prize Outreach. (2024, October 8). *Press release: The Nobel Prize in Physics 2024*.

Peters, E. E. (1994). *Fractal market analysis: Applying chaos theory to investment and economics*. John Wiley & Sons.

Rochat, Y. (2009). *Closeness centrality extended to unconnected graphs: The harmonic centrality index*. Applications of Social Network Analysis (ASNA).

Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, 6(2), 41–49.

Smith, J. L. (2017). *The price elasticity of U.S. shale oil reserves* (EIA Working Paper). U.S. Energy Information Administration.

Taleb, N. N. (2007). *The black swan: The impact of the highly improbable*. Random House.

U.S. Energy Information Administration. (2020, April 27). Low liquidity and limited available storage pushed WTI crude oil futures prices below zero. *Today in Energy*.

van Mieghem, P. (2011). *Graph spectra for complex networks*. Cambridge University Press.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of “small-world” networks. *Nature, 393*(6684), 440–442. https://doi.org/10.1038/30918

World Energy Council. (2022). *World Energy Trilemma Index 2022*.

Wu, J., Barahona, M., Tan, Y.-J., & Deng, H.-Z. (2010). Natural connectivity of complex networks. *Chinese Physics Letters, 27*(7), 078902. https://doi.org/10.1088/0256-307x/27/7/078902