

Rethinking Time Series with Kolmogorov-Arnold Networks: A Systematic Review

Antoni Mól

Kozminski University

Dariusz Jemielniak

Kozminski University

Leon Ciechanowski

lciechanowski@kozminski.edu.pl

Kozminski University

Research Article

Keywords: Kolmogorov–Arnold Networks, Time series forecasting, Model interpretability, Concept drift, Systematic review, Deep learning.

Posted Date: July 14th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10247622/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Rethinking Time Series with Kolmogorov-Arnold Networks: A Systematic Review

Antoni Mól^a, Dariusz Jemielniak^{a,b,*}, Leon Ciechanowski^{a,c,*}

^aManagement in Networked and Digital Societies (MINDS) department, Kozminski University, Warsaw, Poland

^bBerkman-Klein Center for Internet and Society, Harvard University, Cambridge, MA, USA

^cMIT Center for Collective Intelligence, Massachusetts Institute of Technology, Cambridge, MA, USA

*Corresponding authors: Leon Ciechanowski lciechanowski@kozminski.edu.pl, Dariusz Jemielniak djemielniak@cyber.harvard.edu

ORCID iDs: Dariusz Jemielniak 0000-0002-3745-7931; Leon Ciechanowski 0000-0002-4569-7222

Acknowledgements

This work was supported by Narodowe Centrum Nauki (National Science Centre, Poland) under Grant 2020/38/A/HS6/00066. Part of the work for the research paper was conducted by Leon Ciechanowski when he was a Research Scholar at MIT (September 2023 - September 2024).

Abstract

Kolmogorov–Arnold Networks (KANs) have gained attention as a promising alternative to Multi-Layer Perceptrons for function approximation and interpretable neural modeling. Their edge-based learnable univariate transformations offer a distinctive architecture with potential applications in scientific prediction, industrial modeling, and other domains where transparency is important. This systematic review examines the suitability of KANs for time series analysis, focusing on their performance relative to established models, interpretability, computational efficiency, and ability to address challenges such as concept drift. We synthesize evidence from recent empirical studies across energy systems, environmental monitoring, biomedical signals, finance, infrastructure, and general forecasting benchmarks. The review finds that KANs often perform well on structured, short-horizon tasks characterized by smooth temporal patterns and benefit from compact parameterizations and visualizable learned functions. However, the evidence base remains preliminary, heterogeneous, and unevenly reported, with limitations in long-range forecasting, abrupt regime shifts, financial markets, and rigorous validation of interpretability claims. Beyond surveyed applications, we identify social media bot detection as a promising but challenging direction because it combines multi-scale temporal behavior with adversarial dynamics. Overall, KANs appear most

suitable for selective deployment in domains where their architectural properties align with problem structure, while future work should prioritize standardized benchmarks, statistical testing, concept-drift evaluation, and user-centered interpretability studies.

Keywords: Kolmogorov–Arnold Networks; Time series forecasting; Model interpretability; Concept drift; Systematic review; Deep learning.

1 Introduction

The analysis of time series data plays a critical role in decision-making across diverse domains including finance, healthcare, energy, climate science, and manufacturing. Time series data – sequences of observations recorded at successive time intervals presents unique challenges that require specialized forecasting approaches. Financial institutions rely on time series models for stock price prediction and risk assessment, healthcare providers monitor patient vital signs to predict adverse events, and energy companies forecast electricity demand to optimize resource allocation and maintain grid stability (Zhao et al. 2024). Despite its widespread importance, time series analysis confronts persistent challenges that existing methodologies struggle to address effectively.

One fundamental challenge is the dynamic nature of time series data. Real-world series frequently exhibit non-stationarity, where statistical properties such as mean and variance change over time. This leads to phenomena such as concept drift, where the underlying data distribution evolves, making it difficult for models to maintain accuracy over extended periods (Gama et al. 2014). Traditional statistical methods like ARIMA, while foundational to time series analysis, impose restrictive linearity and stationarity assumptions that rarely hold in practice. These methods rely on predefined functional forms and cannot adaptively learn complex non-linear relationships from data. Moreover, ARIMA’s differencing mechanism proves insufficient when series exhibit structural breaks, regime changes, or other forms of non-stationarity that cannot be addressed through simple transformations (Kontopoulou et al. 2023).

Deep learning offered promise for overcoming these limitations through representation learning, yet recurrent neural networks and Long Short-Term Memory (LSTM) architectures introduce their own fundamental constraints. Despite gating mechanisms designed to address gradient flow, RNNs and LSTMs suffer from vanishing and exploding gradient problems that limit their ability to capture long-term dependencies (Pascanu et al. 2013). The sequential nature of recurrent processing prevents parallelization, leading to computational inefficiencies that become prohibitive for long sequences. Furthermore, LSTMs are susceptible to catastrophic forgetting when updated with new data, resulting in loss of previously learned knowledge – a critical limitation in non-stationary domains requiring continual adaptation (Ao and Fayek 2023).

More recently, Transformer architectures have been widely explored for time series forecasting. However, these models face fundamental architectural mismatches with temporal data. The permutation-invariant nature of self-attention mechanisms conflicts with the ordered structure inherent to time series, potentially resulting in temporal information loss despite the use of positional encodings (Zeng et al. 2022). Additionally, Transformers exhibit quadratic computational

complexity with sequence length, require large datasets for effective training, and often overfit on the smaller datasets typical of time series applications. Critical evaluations have demonstrated that simple linear models can outperform complex Transformer architectures on standard benchmarks, raising questions about whether the complexity of these models is justified for many time series tasks (Zeng et al. 2022).

Beyond architectural limitations, real-world time series present challenges that no existing method adequately addresses. Concept drift – changes in underlying data-generating processes due to evolving market conditions, policy shifts, or environmental factors causes systematic forecast degradation across all model types (Gama et al. 2014). Financial markets exemplify this challenge, where investor psychology, regulatory changes, and economic conditions cause price patterns to evolve continuously. Models trained on historical data become progressively less accurate as distributions shift, yet existing adaptation methods either require impractical online retraining or environment labels that are unavailable in standard forecasting settings.

These converging limitations – inability to learn arbitrary non-linear transformations, failure to maintain genuine long-term dependencies, lack of adaptation mechanisms for concept drift, quadratic scaling with problem complexity, and poor interpretability define critical gaps that next-generation architectures must address. Kolmogorov-Arnold Networks (KANs), with their basis in function approximation theory and learnable univariate activation functions, represent a fundamentally different modeling paradigm that warrants investigation for time series forecasting. The following sections examine KANs’ mathematical foundations and explore whether they can overcome the systematic limitations documented across existing approaches.

2 Research Questions and Objectives

Time series analysis was selected as the focus of this work due to the promising potential of Kolmogorov–Arnold Networks (KANs) in this domain. The spline-based structure of KANs aligns naturally with the requirements of time series modeling, where smooth functional approximations and adaptive behavior are essential. Moreover, as a comparatively new neural architecture, exploring KANs across different modeling scenarios is necessary to better understand their capabilities and limitations. Their functional, node-level representation also offers an avenue for qualitative interpretability, which can be particularly valuable in time series applications where understanding model behavior is often as important as achieving high predictive accuracy.

To establish a solid foundation for this study, we identified 243 papers related to Kolmogorov–Arnold Networks and their applications published between April 2024 and November 2025. From this corpus, 36 papers were selected for detailed examination based on their relevance to time series analysis and their coverage of key aspects such as model behavior, interpretability, and responsiveness to dynamic or evolving data. The breadth of topics represented in the reviewed literature highlights the growing versatility of KANs and underscores the importance of assessing their potential specifically within the context of time series forecasting.

Guided by this review, the present work aims to systematically investigate the suitability of KANs for time series analysis. The study addresses the following research questions:

- (RQ1) Under what conditions do Kolmogorov–Arnold Networks outperform established architectures for time series forecasting?
- (RQ2) How does the qualitative interpretability provided by KAN visualizations and symbolic formulas compare to post-hoc explainability methods (SHAP, LIME) applied to traditional architectures, when evaluated using interpretability metrics such as user comprehension, perceived trustworthiness, and the actionability of the resulting explanations?
- (RQ3) What are the computational efficiency–accuracy trade-offs of Kolmogorov–Arnold Networks compared with conventional time series models, considering factors such as training complexity, inference behavior, and model size?

3 Information Sources

Regarding the field of this research and best practices in the field, four databases were selected for this research including ScienceDirect, IEEE Xplore, Springer and arXiv. The inclusion of arXiv was particularly important, as the topic is relatively new and many relevant papers have not yet undergone peer review but are available as preprints.

3.1 Search Strategy

To identify relevant studies on the application of Kolmogorov–Arnold Networks (KANs) in time series analysis, a comprehensive search was conducted across four major academic databases, selected based on relevance to computer science, artificial intelligence, and data analysis. The databases searched were: ScienceDirect, IEEE Xplore, Springer and arXiv.

The search strategy was developed using a combination of keywords and Boolean operators to capture all variations of the key concepts. The following search string was adapted to the syntax of each database:

("Kolmogorov–Arnold Network" OR "Kolmogorov Arnold Network" OR "KAN") AND ("time series" OR "temporal data" OR "sequence data") AND ("analysis" OR "forecasting" OR "prediction" OR "modeling")

In arXiv, which includes preprints that are often not yet peer-reviewed, special attention was paid to the most recent developments due to the emerging nature of KANs. This was necessary because many high-quality and relevant articles on this topic had not yet undergone formal peer review at the time of the search.

3.2 Inclusion and exclusion criteria

To ensure the relevance and rigor of the literature review, studies were selected according to pre-defined inclusion and exclusion criteria.

3.2.1 Inclusion Criteria

1. Studies that implement or analyze Kolmogorov–Arnold Networks (KANs) or closely related functional decomposition architectures.
2. Research applying KANs specifically to time series forecasting, modeling, or related temporal prediction tasks.
3. Publications reporting empirical results on KAN performance, including predictive accuracy, generalization, adaptability to non-stationary data, or computational considerations.
4. Studies published between April 2024 and November 2025, reflecting the most recent advances in KAN methodologies.
5. Peer-reviewed journal articles, conference papers, or high-quality preprints with sufficient methodological transparency to allow assessment of experimental design, data usage, and model evaluation.

3.2.2 Exclusion Criteria

1. Studies that focus solely on theoretical aspects of KANs without empirical evaluation or application to time series.
2. Publications lacking adequate methodological detail, including insufficient information on datasets, model configuration, or evaluation procedures.
3. Non-English language studies, duplicates, editorials, commentaries, or opinion pieces.
4. Studies that apply KANs to domains other than time series where results are not generalizable to temporal modeling tasks.

This systematic approach ensured that only studies directly relevant to the evaluation of KANs in time series forecasting were included, enabling a focused and high-quality synthesis of current evidence on their performance.

3.3 Selection Process

The search process was refined by reading the abstract and skimming the content, as a result we selected 57 studies. After the second round of pruning (in-depth screening process-reading full text) 36 papers remained for the final analysis. The study-selection flow chart diagram shown in Fig. 1 represents the process of inclusion/exclusion visually for the reader.

Study-selection flow summary. The search identified 243 records. After abstract screening, 57 studies were retained for full-text assessment; 21 were excluded as outside the scope of the review. The final corpus included 36 studies and 36 corresponding reports.

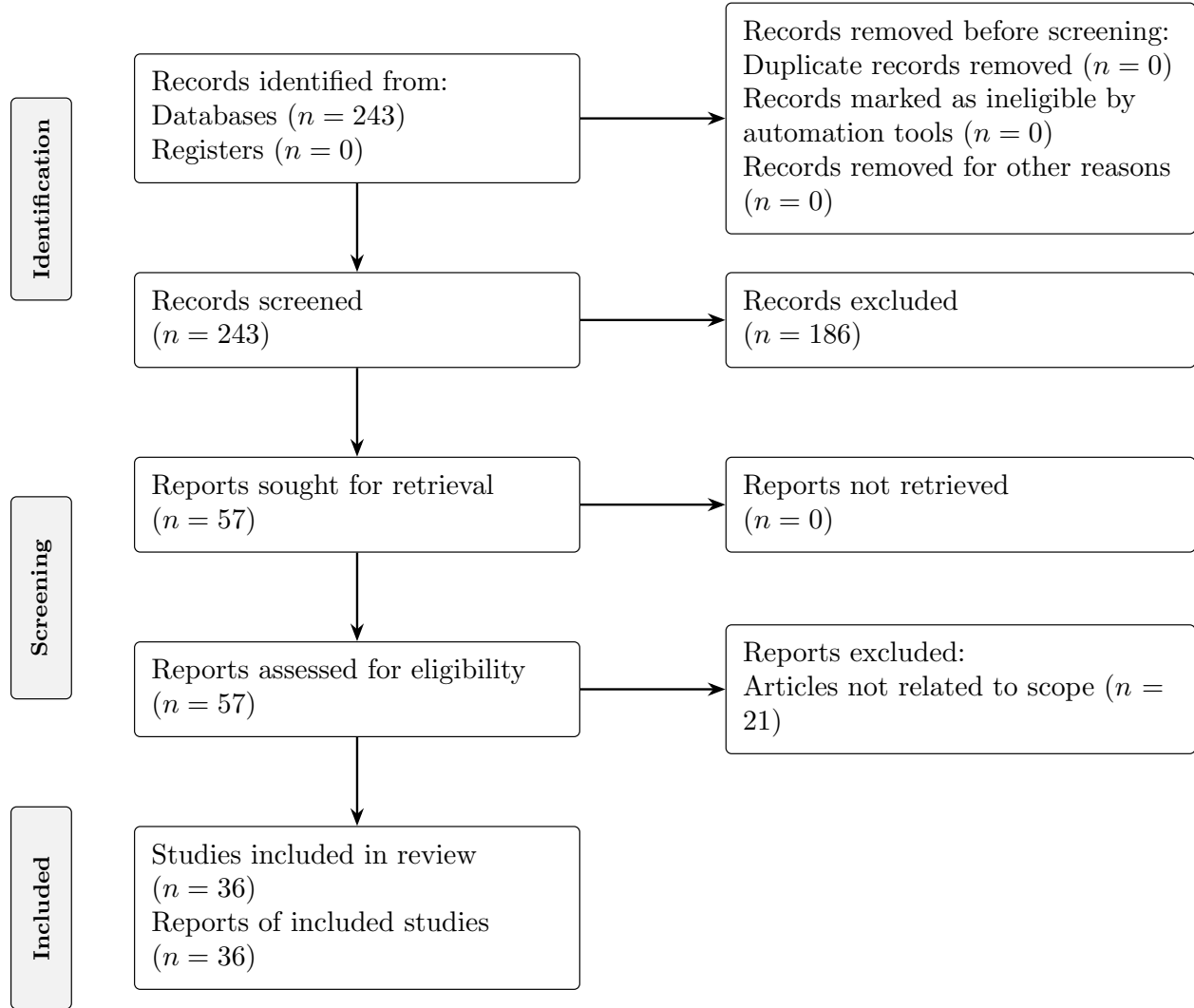


Fig. 1 Study-selection flow diagram showing the study identification, screening, eligibility assessment, and inclusion process.

4 Structure of the Paper

The paper begins with an introduction to time series challenges and the motivation for using Kolmogorov-Arnold Networks (KANs). It then outlines the theoretical and mathematical foundations of KANs, highlighting their unique architecture. This is followed by a benchmarking section comparing KANs to traditional models, with a focus on performance and interpretability. The next sections review KAN applications across various domains, including energy, environment, biomedical, and finance. Concept drift is then examined in detail, assessing how KANs adapt to evolving data, followed by a critical evaluation of KANs interpretability. The paper concludes with a discussion of key findings and implications for future research.

5 The Theoretical Foundations of KANs

Kolmogorov-Arnold Networks are a novel class of neural network architectures based on the Kolmogorov-Arnold representation theorem. This theorem states that any continuous multi-variable function can be decomposed into a finite sum of compositions of continuous univariate functions. In essence, it provides a framework for representing complex, high-dimensional functions through simpler, one-dimensional components.

Traditional neural networks, such as Multi Layer Perceptrons (MLPs), rely on fixed activation functions at each neuron and adjust linear weights during training to approximate target functions. In contrast, KANs innovate by assigning learnable, non-linear activation functions to the connections (edges) between neurons, effectively replacing linear weights with parameterized univariate functions, often modeled as B-splines. This architectural shift allows KANs to adaptively learn the necessary transformations for function approximation, aligning closely with the decomposition principle outlined in the Kolmogorov-Arnold theorem (Liu et al. 2024).

KANs process input data by passing it through a series of learnable functions that are assigned to the edges of the network. Instead of traditional fixed weights, each connection applies a univariate transformation to its input before passing it forward. The outputs from multiple edges are then combined at each neuron, allowing the network to build complex functions from simpler ones. This design provides greater flexibility, as the network can adapt the activation functions themselves instead of relying on predefined ones.

5.1 Mathematical Foundations

The foundation of KANs rests on the Kolmogorov-Arnold representation theorem, which states that any continuous function $f : [0, 1]^n \rightarrow \mathbb{R}$ can be expressed as:

$$f(x_1, \dots, x_n) = \sum_{q=0}^{2n} \Phi_q \left(\sum_{p=1}^n \psi_{q,p}(x_p) \right) \quad (1)$$

where (Φ_q) and $(\psi_{q,p})$ are continuous univariate functions. This decomposition suggests that high-dimensional mappings can be broken down into nested compositions of single-variable functions, forming the theoretical backbone of KANs (Schmidt-Hieber 2021).

In the original KAN construction, each directed edge from input coordinate i in layer ℓ to output coordinate j in layer $\ell + 1$ is associated with a learnable univariate function $\phi_{\ell,j,i}$. A KAN layer is therefore more precisely written as:

$$x_{\ell+1,j} = \sum_i \phi_{\ell,j,i}(x_{\ell,i}). \quad (2)$$

Thus, the node aggregates transformed edge outputs rather than first forming a scalar-weighted preactivation and then applying a fixed node-level nonlinearity, as in a conventional MLP. The functions $\phi_{\ell,j,i}$ are typically parameterized using splines, often combined with residual basis functions in specific implementations, and their parameters are learned during training. This distinction is central to KANs’ expressiveness and interpretability: nonlinearity resides primarily on edges as learned univariate transformations, while any additional residual basis terms, gates, or node nonlinearities should be understood as variant-specific design choices rather than defining properties of KANs in general (Liu et al. 2024).

By leveraging the Kolmogorov-Arnold representation theorem, KANs offer a structured approach to approximating high-dimensional functions with improved efficiency and transparency. Their capacity to learn and adapt activation functions, as opposed to relying on predefined ones, positions them as a compelling alternative to traditional neural network architectures (Liu et al. 2024).

6 Benchmarking Kolmogorov-Arnold Networks for Time Series Modeling

Kolmogorov-Arnold Networks (KANs) have quickly positioned themselves as a serious contender in the landscape of time series modeling, not only because of their theoretical grounding in the Kolmogorov-Arnold representation theorem but also due to consistent empirical performance across diverse domains. Rather than relying on fixed nonlinearities like ReLU or tanh, KANs learn univariate transformations along edges, enabling highly adaptable function approximation. This structure has been repeatedly shown to improve performance and stability in sequential tasks marked by noise, autocorrelation, or structural irregularity (Dong et al. 2024; Han et al. 2024a; Livieris 2024).

KANs are especially effective in time series tasks where the patterns are complex but still follow some hidden structure such as repeating cycles, overall trends, or local shapes. Han et al. (2024a) and Livieris (2024) both found that KAN-based models could match or even outperform leading transformer models, particularly on datasets where the data had smooth or locally structured patterns. In Han’s study, the RMoK model, which combines several specialized KANs, performed very well across different time series benchmarks. Similarly, the C-KAN model developed by Livieris et al. showed strong results on non-stationary and noisy datasets, using a KAN decoder trained

with the DILATE loss function. The strengths of KANs are not limited to forecasting alone. Dong et al. (2024) demonstrated that even basic KANs significantly outperformed MLPs across 128 univariate time series classification tasks in the UCR archive, highlighting their ability to capture complex temporal patterns. This complements findings from Lee et al. (2024), who proposed a hybrid model called HIPPO-KAN that maintained a constant parameter count while achieving superior time series prediction results.

This benchmarking overview shows that Kolmogorov-Arnold Networks (KANs) consistently perform well across a range of time series tasks, including forecasting and classification. Compared to traditional models and transformers, KANs offer strong accuracy, robustness to noise, and lower parameter complexity. Variants like HiPPO-KAN further improve scalability.

6.1 Application of Kolmogorov-Arnold Networks in Energy Systems Forecasting

Energy load forecasting is a fundamental task in modern power systems, directly impacting operational efficiency, cost management, and the integration of renewable energy sources. Traditional load forecasting approaches which include statistical models and deep learning techniques such as Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks, have demonstrated varying degrees of success. However, these models often struggle with issues like capturing non-linear dependencies, adapting to changing consumer behaviors, and maintaining interpretability.

The integration of multiple energy sources, including electricity, heating, and cooling, further complicates the forecasting task. Furthermore, energy production and consumption are highly impacted by factors such as human habits, environment and weather what makes the prediction even more complex (Liu et al. 2025). In integrated energy systems (IES), these variables exhibit strong interdependencies, making it essential to develop forecasting models capable of handling complex, multi-dimensional relationships (Liu et al. 2025).

Kolmogorov-Arnold Networks (KANs) have recently emerged as a promising alternative for load forecasting, leveraging their ability to approximate complex functions using learnable univariate transformations. Unlike traditional neural networks, which rely on fixed activation functions, KANs place learnable activation functions on the network’s edges, parameterized as B-splines. This architectural shift provides several advantages, including improved interpretability, better handling of non-linearity, and enhanced adaptability to different forecasting scenarios (Jiang et al. 2025).

This chapter synthesizes recent advancements in KAN-based models applied to various load forecasting tasks, emphasizing interpretability, computational efficiency, and adaptability across contexts.

One of the key studies in this domain introduced a correlation feature self-assigned KAN for multi-energy load forecasting in Integrated Energy Systems (Liu et al. 2025). This model integrates a multi-decoder Informer architecture, enabling it to extract feature relationships from various energy loads, such as electricity, heat, and cooling demand. The results demonstrated that although the proposed algorithm shows a slightly higher MAPE and RMSE for heat load by 10.792% and 4.757%

respectively compared to the Transformer model, it achieves significant improvements in other areas. Specifically, it reduces the MAPE and RMSE for electrical load by 23.547% and 20.451%, respectively, and lowers the MAPE and RMSE for cooling load by 34.920% and 34.316%. Overall, the KAN-based model achieves a lower average RMSE across predicted electrical, cooling, and heating loads compared to other models.

Another significant advancement is the introduction of the Kolmogorov-Arnold Recurrent Network (KARN) for short-term load forecasting across diverse consumer types (Danish and Grolinger 2024). This approach extends the traditional KAN framework by incorporating recurrent structures, enabling it to capture temporal dependencies in energy demand. By leveraging learnable activation function on edges, KARN deals with the traditional limitations of deep learning models. KARN was evaluated on multiple real-world datasets, including residential homes, student residences, and industrial buildings. The model consistently outperformed conventional RNNs and achieved superior accuracy compared to LSTM and Gated Recurrent Unit (GRU) models in most scenarios. For example, KARN outperformed RNN and GRU models in the individual houses scenario achieving a SMAPE of 0.2873, in comparison to RNN’s 0.4039, LSTM’s 0.3171 and GRU’s 0.3203. In the student residences scenario the proposed model similarly outperformed the RNN’s model which is described as KARN’s equivalent deep learning model (Danish and Grolinger 2024). A related implementation of Kolmogorov-Arnold Networks in energy systems forecasting has been successfully demonstrated in predicting chiller energy consumption in commercial buildings (Sulaiman et al. 2024a). Using real-world data collected from a commercial building the KAN model achieved a superior performance reaching RMSE of 6.1023 compared to RMSE of 6.2892 for TLBO-DL model. Moreover, the KAN demonstrated more consistent and lower prediction error variance, with residuals tightly centered around zero and most predictions falling within a $\pm 10\%$ error range.

A study on short-term probabilistic load forecasting (SPLF) using KANs further addressed the challenge of uncertainty in load forecasting (Jiang et al. 2025). By modeling input uncertainties using normal distributions and introducing a multi-interval linearized SPLF technique, the study improved the accuracy of probabilistic forecasts compared to conventional artificial neural networks (ANNs). However, in the present analysis, the Quantile-ANN-based approach achieved substantially lower forecasting errors across seasons, with an average RMSE of 344.874 compared to 1004.198 for the KAN-based model. These results indicate that, despite the promising probabilistic forecasting capabilities and computational efficiency reported for KANs in previous research, the Quantile-ANN model demonstrated considerably stronger predictive performance under the evaluated seasonal conditions.

In summary, KAN-based models demonstrate strong potential in energy load forecasting by improving accuracy across diverse energy sources, capturing complex correlations, and addressing interpretability concerns. However, not all studies consistently confirm these advantages across all settings, as some results show variability in performance depending on the dataset, forecasting horizon, and evaluation conditions. The continued refinement of these models suggests they may become a useful tool for energy system operators seeking robust and adaptive forecasting techniques, although their superiority is not yet universally established.

6.2 Kolmogorov–Arnold Networks in Battery State Estimation

Kolmogorov–Arnold Networks (KANs) have emerged as a compelling alternative to traditional neural architectures in battery diagnostics, offering improved accuracy, interpretability, and efficiency. Leveraging the Kolmogorov–Arnold representation theorem, KANs replace fixed activation functions with learnable ones – often B-splines – enabling better adaptation to complex nonlinear patterns typical in lithium-ion battery behavior. Recent research has applied KANs to various battery state estimation tasks, such as state-of-charge (SoC), state-of-health (SoH), core temperature, and even broader smart-grid contexts.

In SoC estimation which is a crucial factor for optimization of vehicle performance and efficient battery management, KANs outperform both artificial neural networks (ANNs) and hybrid deep models. For instance, the best performing KAN-Model 2 achieved a RMSE of 3.8545 and R^2 of 0.9755 compared to a classical ANN and BMO-DL models which achieved a RMSE of 4.3041 and 5.6610 respectively (Sulaiman et al. 2024a). In a study conducted by Le et al., 2024 concerning the classification of EV energy consumption KANs also demonstrated superiority over benchmarked models. Using a dataset of 3,345 Volkswagen EVs, a KAN classifier reached an average accuracy of 0.893 and best-case accuracy of 0.938 with only 3,537 parameters. This outperformed a deeper MLP (5,572 parameters), which averaged 0.858 accuracy. KANs also showed a smaller variability of results, as the training size decreased (Le et al. 2024). These results suggest that KANs could be well-suited to addressing the challenges of EV batteries.

In parallel, KANs have shown promise in SoH estimation which is a crucial component of safe and efficient battery-powered systems. It helps assess the battery’s remaining capacity and lifespan, and plays a key role in optimizing charging and discharging to prolong battery life and avoid early failure. Peng et al. (2024) used a combination of features that reflect both electrochemical behavior and battery degradation trends specifically, patterns identified from incremental capacity (IC) curves along with measurements of internal resistance. These inputs were fed into a range of machine learning models, including BP, LSTM, and hybrid proposed architectures such as LSTM-KAN and TCN-KAN. Among all the tested models, the LSTM-KAN model achieved the best performance, with a root mean square error (RMSE) of 0.00570 compared to TCN’s 0.00607. Importantly, the model does not need full charge/discharge cycles, making it practical for real-world applications where rapid and reliable SOH estimation is needed.

Beyond charge and health metrics, KANs are also proving useful for thermal state estimation. Karnehm et al. (2024) used a KAN model to estimate the internal core temperature of cylindrical lithium-ion batteries. The model achieved a temperature prediction error (MAE) of only $\pm 0.5^\circ\text{C}$, which makes it suitable for onboard use. Unlike conventional methods, the model did not need surface temperature as input, thanks to KAN’s ability to learn internal thermal behavior from electrical signals alone.

Together, these studies underscore KANs as a robust and efficient architecture for battery state estimation. From predicting degradation to internal temperature and operational classifications, KANs deliver both precision and interpretability—qualities that position them as future mainstays in smart battery management systems.

6.3 Kolmogorov–Arnold Networks in Environmental and Infrastructure Forecasting

KANs’ recent applications span hydrological streamflow prediction, wind nowcasting, structural monitoring, energy resource forecasting, and air transportation resilience analysis – domains previously dominated by deep learning models like LSTMs and CNNs (Granata et al. 2024).

In streamflow forecasting, Granata et al. (2024) evaluated the performance of KANs against Transformer models across four major Central European rivers: the Rhine, Danube, Elbe, and Oder. Using long-term daily flow records and derived features such as moving averages and rate changes, KANs consistently outperformed Transformers for 1- to 3-day forecasts. For example, in 1-day forecasts, the KAN achieved R^2 values of 0.975 (Rhine), 0.956 (Danube), 0.992 (Elbe), and 0.969 (Oder), compared to Transformer values of 0.955, 0.923, 0.984, and 0.971, respectively. Beyond 3 days, the performance of both models converged. The spline-based adaptability of KANs enabled them to more accurately capture short-term hydrological variability, particularly under dynamic flow conditions, offering improved predictive performance that supports effective flood risk mitigation and water resource management (Granata et al. 2024).

Forecasting infrastructure behavior has also benefitted from hybrid KAN architectures. Xu et al. (2024c) introduced a dual-stage attention model combining LSTM with KAN (DA-LSTM-KAN) for predicting dam deformation. Applied to a concrete gravity dam, the model achieved superior accuracy over traditional models, including linear regression and standard LSTM variants. Specifically, it achieved an accuracy of 1.000 within the accepted error threshold. The model’s interpretability, enabled by its attention mechanisms, provided clear insights into which factors and time periods most influenced dam deformation, enhancing trust in the predictions and supporting informed engineering decisions (Xu et al. 2024c).

KANs have also been leveraged for production forecasting in energy infrastructure. Hu et al. (2024) proposed a hybrid model integrating Temporal Convolutional Networks (TCNs) with KANs, optimized via the Whale Optimization Algorithm (WOA), for oil well production forecasting. Using production data from the Volvo oil field, this TCN-KAN model consistently outperformed all other models, achieving the highest R^2 score (0.9815) and the lowest error values across all metrics. Residual plots and cross-plots further demonstrated the model’s robustness, with predictions closely tracking actual values and minimal error variance across different production regimes (Hu et al. 2024). Beyond accuracy, the hybrid structure of the model also provided interpretability benefits. By separating the roles of temporal and nonlinear processing, the architecture made it easier to understand which components of the input data contributed to specific predictive behaviors.

Moreover, KANs have shown strong potential in enhancing airport forecasting and resilience. At Madeira International Airport, Alves et al. (2024) used a KAN model to correct Global Forecast System wind predictions, achieving a 48.5% reduction in mean squared error and lowering the 3-hour forecast MAE. Crucially, the model provided interpretable formulas that linked wind errors to terrain effects. This not only improved understanding of atmospheric dynamics but also enhanced the transparency and interpretability of the applied machine learning approach. In a complementary study, Song et al. (2024) applied KANs to assess how weather factors affect operational resilience at Xi’an Xianyang Airport. The KAN model achieved $R^2 = 0.99806$ and $\text{MAE} = 60.7$. The authors report that it reliably forecasted resilience scores under varying weather conditions. Both

studies highlight KANs’ ability to accurately model complex, localized weather effects and produce transparent insights for aviation decision-making.

In summary, Kolmogorov-Arnold Networks offer a robust and highly adaptable framework for forecasting across diverse environmental and infrastructural domains. Whether deployed standalone or as part of hybrid models, KANs consistently yield improvements in accuracy over traditional neural architectures. Their success across hydrology, meteorology, structural engineering, energy systems, and transportation demonstrates not only the breadth of their applicability but also their potential to reshape forecasting practices in critical societal systems.

6.4 Biomedical Applications of Kolmogorov–Arnold Networks

Across recent studies, KANs have been deployed for tasks such as electrocardiogram (ECG) abnormality detection, infectious disease modeling, and anomaly detection in biomedical sensor data. The results are promising but preliminary.

A prime example of KANs’ potential lies in cardiac monitoring. Huang et al. (2024) developed a lightweight KAN model with a single hidden layer of 64 neurons for classifying ECG abnormalities using the TNMG dataset and testing on the CPSC 2018 dataset. The model achieved an F1-score of 0.75 on TNMG, and generalization scores of 0.62 (F1) on the out-of-distribution CPSC data. Notably, the model maintained acceptable performance when reduced to a single ECG lead, a critical advantage for edge deployment in wearable medical devices. Compared to multilayer perceptrons (MLPs) and Neural Circuit Policies (NCPs) which are bio-inspired neural networks, the KAN model exhibited superior parameter efficiency and interpretability, crucial for clinical acceptance while maintaining robustness under missing-channel conditions (Huang et al. 2024).

Beyond classification, KANs have also been adapted for anomaly detection in time-series physiological data through the KAN-AD framework. Zhou et al. (2024) redesigned KANs using Fourier series instead of splines to better capture global signal patterns while suppressing noise-induced local peaks and drops, a common issue in biosignal monitoring. On four benchmark datasets, KAN-AD increased anomaly detection F1-scores by 29% and reduced inference time by 36% compared to the original KAN model. On the UCR dataset, which includes biomedical time series such as ECG and internal bleeding signals, KAN-AD achieved the highest average F1 score (0.7789) while maintaining exceptional efficiency with only 274 parameters compared to LSTMAD 0.6808 F1 score and 10421 parameter count. While not originally applied in clinical studies, these results underscore KAN-AD’s strong potential for deployment in real-time biomedical monitoring systems. (Zhou et al. 2024).

On a broader systems level, KANs have also been integrated with mechanistic models in epidemiology. Ma et al. (2024) introduced KAN-UDEs hybrids combining KANs with universal differential equations to model infectious disease spread. Using simulated time-series data from SIR-type compartment models with four types of incidence functions, KAN-UDEs demonstrated faster convergence and more accurate functional reconstruction than MLP-based UDEs. This integration preserved mechanistic interpretability while boosting predictive accuracy, offering a path toward hybrid clinical modeling systems that combine theoretical disease dynamics with flexible, data-driven pattern recognition (Ma et al. 2024).

Together, these studies illustrate how KANs are advancing biomedical AI by bridging the gap between data-driven learning and interpretable, efficient modeling. Whether deployed in wearable ECG monitors, clinical anomaly detection systems, or epidemic forecasting platforms, KANs provide a foundation for reliable, low-latency, and explainable health analytics.

6.5 Financial Forecasting with Kolmogorov-Arnold Networks

Kolmogorov-Arnold Networks (KANs), rooted in the Kolmogorov-Arnold representation theorem, offer a novel framework for approximating multivariate functions, making them particularly suitable for time series analysis in financial forecasting. Financial markets, characterized by volatility, non-stationarity, and complex temporal dependencies, pose significant challenges for traditional models like Multi-Layer Perceptions (MLPs), Long Short-Term Memory (LSTM) networks, and Gated Recurrent Units (GRUs) (Zhao et al. 2024). KANs, with their ability to represent complex functions have shown promise in enhancing prediction accuracy and stability across various financial domains, including cryptocurrency and stock markets. This part synthesizes findings from five recent studies, focusing on the application of KAN-based models, their experimental results, and the datasets employed, to provide a comprehensive overview of their role in financial forecasting.

Recent advancements in time series forecasting have emphasized models that balance theoretical robustness, nonlinear expressiveness, and practical performance. A central development in this direction is the Kolmogorov-Arnold Network (KAN) which allows any multivariate continuous function to be decomposed into sums of univariate functions. This property underpins a range of modern KAN-based architectures designed for sequence modeling, particularly in financial contexts (as cited in Genet and Inzirillo 2024a,b; Inzirillo and Genet 2024a,b; Liu et al. 2024).

Temporal Kolmogorov-Arnold Networks (TKANs) extend KANs with memory structures similar to LSTMs, enabling them to model both short- and long-term dependencies. When applied to multivariate financial time series from Binance (2020–2022), TKAN achieved an R^2 of 0.1378 at a 6-step prediction horizon, outperforming a GRU model (Genet and Inzirillo 2024b). The Temporal Kolmogorov-Arnold Transformer (TKAT) further improved these results, embedding TKANs within an attention-based encoder-decoder structure. TKAT achieved an R^2 of 0.1796 at 6 steps, indicating that the attention mechanism enables more effective modeling of long-range dependencies (Genet and Inzirillo 2024a). Another approach, Signature-Weighted KANs (SigKANs), introduces learnable path signatures-geometric descriptors derived from rough path theory to weight KAN outputs. SigKANs matched or surpassed alternatives on longer horizons ($R^2 = 0.1636$ at 6 steps), outperforming TKAN but remaining slightly behind TKAT (Inzirillo and Genet 2024b).

Complementing these studies, Yao (2024) proposed a hybrid LSTM-KAN model tailored to stock price prediction on the Shanghai Stock Exchange Composite Index. Compared to a standalone LSTM (RMSE = 0.0655), the LSTM-KAN reduced RMSE to 0.0082 and mean absolute error (MAE) to 0.0057. This striking improvement demonstrates the capacity of KAN layers to refine LSTM outputs via nonlinear functional decomposition, significantly enhancing accuracy in volatile financial environments.

Together, these models show that KANs can enhance deep learning approaches in financial forecasting. By combining theoretical flexibility with practical performance, KANs offer a promising

direction for capturing the complex patterns that define modern financial time series.

6.6 Performance Comparison Table Across Various Studies and Domains

Table 1 summarises the reviewed performance comparisons across domains, including each KAN variant, comparator model, dataset, reported predictive metrics, and computational-resource information where available.

Table 1: Performance comparison of KAN-based models and baseline methods across reviewed domains. The table reports the application domain, study, KAN variant, baseline model, dataset, primary performance metrics for both model types, and reported computational resources where available.

Domain	Study	KAN Model	Baseline	Dataset	KAN Performance	Baseline Performance	Computational Resources KAN	Computational Resources Baseline
Energy	State of Health Estimation of Li-Ion Battery via Incremental Capacity Analysis and Internal Resistance Identification Based on Kolmogorov–Arnold Networks (Peng et al. 2024)	LSTM-KAN	TCN	SOH Battery Estimation	RMSE: 0.00570	RMSE: 0.00607	Training time: 15s	Training time: 22.5s
Energy	Core Temperature Estimation of Lithium-Ion Batteries Using Long Short-Term Memory (LSTM) Network and Kolmogorov-Arnold Network (KAN) (Karnehm et al. 2024)	KAN	LSTM	CC-CV Dataset	RMSE Core: 0.814 RMSE Surface: 0.519 R^2 Core: 0.99808 R^2 Surface: 0.99911 *in Celsius degrees	RMSE Core: 0.875 RMSE Surface: 0.520 R^2 Core: 0.99779 R^2 Surface: 0.99910	2.92+- 0.45 ms *on CPU Intel Xeon Gold 6338 2G	0.55+- 0.04 ms *on CPU Intel Xeon Gold 6338 2G
Energy	Battery state of charge estimation for electric vehicle using Kolmogorov-Arnold networks (Sulaiman et al. 2024a)	KAN-Model 2	ANN	State of Charge Battery Estimation	RMSE : 3.8545 R^2 : 0.9755	RMSE: 4.3041 R^2 : 0.9693	Not Available	Not Available

Continued on next page

Table 1: Performance comparison of KAN-based models and baseline methods across reviewed domains (continued).

Domain	Study	KAN Model	Baseline	Dataset	KAN Performance	Baseline Performance	Computational Resources KAN	Computational Resources Baseline
Energy	Kolmogorov-Arnold networks for supervised learning tasks in smart grids (Le et al. 2024)	KAN	MLP	Volkswagen Electric Golf Vehicles - Battery Capacity Prediction	Regression task: a) Train size 80% Avg Accuracy: 0.983 b) Train size 60% Avg Accuracy: 0.885 c) Train size 40% Avg Accuracy: 0.875	Regression task: a) Train size 80% Avg Accuracy: 0.858 b) Train size 60% Avg Accuracy: 0.827 c) Train size 40% Avg Accuracy: 0.789	3537 parameters	5572 parameters
Biomedical	Integrating Kolmogorov-Arnold Networks with Ordinary Differential Equations for Efficient, Interpretable and Robust Deep Learning: A Case Study in the Epidemiology of Infectious Diseases (Ma et al. 2024)	KAN-UDE	MLP-UDE	SIR - generated data of new infections and diseases	Accuracy: 1.000 LOSS: 8.59 x 10 ⁻⁷	Accuracy: 1.000 LOSS: 6.45 x 10 ⁻⁴	Training time: 1264 s Iterations to full accuracy: 1800	Training time: 670 s Iterations to full accuracy: 4600

Continued on next page

Table 1: Performance comparison of KAN-based models and baseline methods across reviewed domains (continued).

Domain	Study	KAN Model	Baseline	Dataset	KAN Performance	Baseline Performance	Computational Resources KAN	Computational Resources Baseline
Biomedical	Abnormality Detection in Time-Series Bio-Signals using Kolmogorov-Arnold Networks (KANs) Based Models for Resource-Constrained Devices (Huang et al. 2024)	CKAN	CCfC	TNMG and CPSC datasets	Recall: 0.838 Precision: 0.840	Recall: 0.806 Precision: 0.859	Not Available	Not Available
Biomedical	KAN-AD: Time Series Anomaly Detection with Kolmogorov-Arnold Networks (Zhou et al. 2024)	KAN-AD	LSTMAD	KPI, TODS, WSD, UCR Datasets	Average F1 Score Across All Datasets: 0.7789	Average F1 Score Across All Datasets: 0.6808	For UCR Dataset: GPU Time: 42s CPU Time: 36s Parameters: 274	For UCR Dataset: GPU Time: 73s CPU Time: 267s Parameters: 10421
General	KOLMOGOROV-ARNOLD NETWORKS FOR TIME SERIES: BRIDGING PREDICTIVE POWER AND INTER-PRETABILITY (Xu et al. 2024b)	MT-KAN	LSTM	Financial Time Series Dataset	RMSE: 0.0075	RMSE: 0.0076	2132 parameters	32071 parameters

Continued on next page

Table 1: Performance comparison of KAN-based models and baseline methods across reviewed domains (continued).

Domain	Study	KAN Model	Baseline	Dataset	KAN Performance	Baseline Performance	Computational Resources KAN	Computational Resources Baseline
Energy/Load Forecasting	TIMEKAN: KAN-BASED FREQUENCY DECOMPOSITION LEARNING ARCHITECTURE FOR LONG-TERM TIME SERIES FORECASTING (Huang et al. 2025)	TimeKAN	iTransformer	Weather, ETTh1, ETTh2, ETTm1, ETTm2 and Electricity for long-term forecasting.	Average MSE across all datasets: MSE: 0.197	Average MSE across all datasets: MSE: 0.178	In all studies KAN-model has the lowest amount of parameters, here results for ETTH1 dataset: 12.84K parameters	In all studies KAN-model has the lowest amount of parameters, here results for ETTH1 dataset: 841.57K parameters
Energy/Load Forecasting	Integrating Kolmogorov–Arnold Networks with Time Series Prediction Framework in Electricity Demand Forecasting (Zhang et al. 2025)	Transformer_KAN	Transformer	ETTh1, ETTm1	Results for ETTh1: RMSE: 0.541 R^2 : 0.9435	Results for ETTh1: RMSE: 0.682 R^2 : 0.9136	Not Available	Not Available
Energy/Load Forecasting	A Primer on Kolmogorov–Arnold Networks (KANs) for Probabilistic Time Series Forecasting (Vaca-Rubio et al. 2025)	KAN Student-t	MLP Student-t	European 5GSTARDUST GEO Satellite Traffic	RMSE: 2.085	RMSE: 4.584	90000 parameters	252972 parameters

Continued on next page

Table 1: Performance comparison of KAN-based models and baseline methods across reviewed domains (continued).

Domain	Study	KAN Model	Baseline	Dataset	KAN Performance	Baseline Performance	Computational Resources KAN	Computational Resources Baseline
Energy/Load Forecasting	A novel correlation feature self-assigned Kolmogorov-Arnold Networks for Multi-energy load forecasting in integrated energy systems (Liu et al. 2025)	Adp-3EncInformer-KAN	Transformer	Tempe campus of Arizona State University Dataset	Average RMSE across all datasets: RMSE: 592.424	Average RMSE across all datasets: RMSE: 821.447	Training time: 843.104s	Training time: 244.513s
Energy/Load Forecasting	A Novel Interpretable Short-term Probabilistic Load Forecasting Method based on Kolmogorov-Arnold Networks (Jiang et al. 2025)	KAN	ANN	Switzerland actual load data	Average RMSE across all periods: RMSE: 1004,198	Average RMSE across all periods: RMSE: 344,874	Average inferring time: 200,67s	Average inferring time: 0.1s
Energy/Load Forecasting	Kolmogorov-Arnold recurrent network for short term load forecasting across diverse consumers (Danish and Grolinger 2024)	KARN	LSTM	Energy consumption datasets	RMSE for Residence 1 for Case study 1: RMSE: 29.70	RMSE for Residence 1 for Case study 1: RMSE: 29.85	Training time: 70.5 min Testing time: 0.18 min	Training time: 39.9 min Testing time: 0.13 min

Continued on next page

Table 1: Performance comparison of KAN-based models and baseline methods across reviewed domains (continued).

Domain	Study	KAN Model	Baseline	Dataset	KAN Performance	Baseline Performance	Computational Resources KAN	Computational Resources Baseline
Energy/Load Forecasting	Utilizing the Kolmogorov-Arnold Networks for chiller energy consumption prediction in commercial building (Sulaiman et al. 2024b)	KAN	TLBO-DL	Chiller energy consumption prediction	RMSE: 6.1023	RMSE: 6.2892	Not Available	Not Available
Environmental Infrastructure Forecasting	The Importance of Weather Factors on the Resilience of Airport Flight Based on Kolmogorov-Arnold Networks (KANs) (Song et al. 2024)	KAN	no baseline	Xi'an Xianyang International Airport dataset	R^2 : 0.99806		Not Available	Not Available
Environmental Infrastructure Forecasting	Predicting the Deformation of a Concrete Dam Using an Integration of Long Short-Term Memory (LSTM) Networks and Kolmogorov-Arnold Networks (KANs) with a Dual-Stage Attention Mechanism (Xu et al. 2024c)	DA-LSTM-KAN	DA-LATM	Concrete Dam Deformation Prediction	RMSE: 0.1618	RMSE: 0.2023	Not Available	Not Available

Continued on next page

Table 1: Performance comparison of KAN-based models and baseline methods across reviewed domains (continued).

Domain	Study	KAN Model	Baseline	Dataset	KAN Performance	Baseline Performance	Computational Resources KAN	Computational Resources Baseline
Environmental Infrastructure Forecasting	Deep insight: An Efficient Hybrid Model for Oil Well Production Forecasting Using Spatio-Temporal Convolutional Networks and Kolmogorov-Arnold Networks (Hu et al. 2024)	WOA-TCN-KAN	LSTM	Equinor (2018) oil dataset	R^2 : 0.9815 RMSE: 9.93	R^2 : 0.9177 RMSE: 14.42	Not Available	Not Available
Environmental Infrastructure Forecasting	On the use of kolmogorov-arnold networks for adapting wind numerical weather forecasts with explainability and interpretability: application to madeira international airport (Alves et al. 2024)	KAN	GFS	Madeira airport wind dataset	RMSE: 1.72	RMSE: 1.81	Not Available	Not Available
Environmental Infrastructure Forecasting	Advanced streamflow forecasting for Central European Rivers: The Cutting-Edge Kolmogorov-Arnold networks compared to Transformers (Granata et al. 2024)	KAN	Transformer	Rhine river flow volume dataset	R^2 : 0.975 RMSE: 166.8	R^2 : 0.955 RMSE: 224.6	Not Available	Not Available

Continued on next page

Table 1: Performance comparison of KAN-based models and baseline methods across reviewed domains (continued).

Domain	Study	KAN Model	Baseline	Dataset	KAN Performance	Baseline Performance	Computational Resources KAN	Computational Resources Baseline
Financial	KASPER: Kolmogorov Arnold networks for stock prediction and explainable regimes (Oad et al. 2025)	KASPER	DQS	Yahoo Finance Dataset	Sharpe: 12.02	Sharpe: 3.65	Not Available	Not Available
Financial	Advancing real estate forecasting: A novel approach using Kolmogorov-Arnold Networks (KANs) (Viktoratos and Tsadiras 2025)	KAN	XGBoost	California house pricing	RMSE: 44191	RMSE: 42477	Not Available	Not Available
Financial	Lstm model enhanced by kolmogorov-Arnold network: Improving stock price prediction accuracy (Yao 2024)	LSTM-KAN	LSTM	Shanghai Stock Exchange Composite Index dataset	RMSE: 0.008226 MAE: 0.005742	RMSE: 0.065467 MAE: 0.052623	Not Available	Not Available
Financial	A Gated Residual Kolmogorov-Arnold Networks for mixtures of Experts (Inzirillo and Genet 2024a)	KAMoE	MoE	BTC Binance dataset	R^2 : 0.795	R^2 : 0.792	Same number of parameters	Same number of parameters

Continued on next page

Table 1: Performance comparison of KAN-based models and baseline methods across reviewed domains (continued).

Domain	Study	KAN Model	Baseline	Dataset	KAN Performance	Baseline Performance	Computational Resources KAN	Computational Resources Baseline
Financial	Signature-Weighted Kolmogorov-Arnold Networks for Time Series (Inzirillo and Genet 2024b)	SigKAN	GRU	Cryptocurrencies dataset (for 6 step horizon)	R^2 : 0.1349	R^2 : 0.1360	Parameters: 114711	Parameters: 93015
Financial	TKAN: Temporal Kolmogorov-Arnold Networks (Genet and Inzirillo 2024b)	TKAN	GRU	Cryptocurrencies dataset (for 6 step horizon)	R^2 : 0.1378	R^2 : 0.0825	Not Available	Not Available
Financial	A Temporal Kolmogorov-Arnold Transformer for Time Series Forecasting (Genet and Inzirillo 2024a)	TKAT	GRU	Cryptocurrencies dataset (for 6 step horizon)	R^2 : 0.1795	R^2 : 0.0825	Trainable parameters: 1561704	Trainable parameters: 99930
General	A multivariate time series prediction model based on the KAN network (Long and Qin 2025)	KANMTS	MLP-based SOFTS	Weather, ETTh1, ETTh2, ETTm1, ETTm2 and Electricity	For ETTm2: Avg MSE: 0.277	For ETTm2: Avg MSE: 0.287	256.01 MB, 27.96 s	274.67 MB, 10.98s
General	TSKANMixer: Kolmogorov-Arnold Networks with MLP-Mixer Model for Time Series Forecasting (Hong et al. 2025)	TSKANMixer	Informer	Weather, ETTh1, ETTh2, ETTm1, ETTm2, NN5_daily, NN5_weekly, CIF_2016, Hospital, Exchange, FRED_MD	For ETTh1: MSE: 0.285	For ETTh1: MSE: 0.337	Training time: 4885.08 s	Not Available

Continued on next page

Table 1: Performance comparison of KAN-based models and baseline methods across reviewed domains (continued).

Domain	Study	KAN Model	Baseline	Dataset	KAN Performance	Baseline Performance	Computational Resources KAN	Computational Resources Baseline
General	TFKAN: Time-Frequency KAN for Long-Term Time Series Forecasting (Kui et al. 2025)	TFKAN	ATFNet	ILI, Weather, ETTh1, ETTh2, ETTm1, ETTm2 and Electricity, Air	For ILI: RMSE: 0.211	For ILI: RMSE: 0.216	Parameters: 16.33M GPU Memory: 250.65	Parameters: 19.20M GPU Memory: 1718.67
General	Exploring Kolmogorov–Arnold Networks for Interpretable Time Series Classification (Barašin et al. 2025)	Efficient KAN	MLP	UCR Datasets	Precision Mean: 0.72 Recall Mean: 0.70	Precision Mean: 0.73 Recall Mean: 0.66	Parameters: 257649	Parameters: 344518
General	KAN4Drift: Are KAN Effective for Identifying and Tracking Concept Drift in Time Series? (Xu et al. 2024a)	WormKan	AutoPlait	Motion Capture Data	F1-score: 0.90	F1-score: 0.87	Not Available	Not Available
General	HiPPO-KAN: Efficient KAN Model for Time Series Analysis (Lee et al. 2024)	HiPPO-KAN	LSTM	BTC-USDT dataset	MSE: 3.26 * 10 ⁻⁷	MSE: 4.69*10 ⁻⁷	Parameters: 4384	Parameters: 4513
General	Are KANs Effective for Multivariate Time Series Forecasting? (Han et al. 2024b)	MMK	DLinear	Weather, ETTh1, ETTh2, ETTm1, ETTm2, ECL, Traffic, Weather	For ETTh1: MSE: 0.374	For ETTh1: MSE: 0.386	Parameters: 139K	Parameters: 972K

Continued on next page

Table 1: Performance comparison of KAN-based models and baseline methods across reviewed domains (continued).

Domain	Study	KAN Model	Baseline	Dataset	KAN Performance	Baseline Performance	Computational Resources KAN	Computational Resources Baseline
General	C-KAN: A New Approach for Integrating Convolutional Layers with Kolmogorov–Arnold Networks for Time-Series Forecasting (Livieris 2024)	C-KAN	Seq2Seq	Bitcoin, Gold, Synthetic, ECG5000, Solar datasets	For Bitcoin: RMSE: 1246.08	For Bitcoin RMSE: 1465.83	Not Available	Not Available
General	Kolmogorov-Arnold Networks (KAN) for Time Series Classification and Robust Analysis (Dong et al. 2024)	KAN	MLP	UCR Datasets	Test accuracy: 0.7750	Test accuracy: 0.7815	Not Available	Not Available

Note. For each experiment reported in the table, we present the best-performing KAN model and the best-performing baseline model. Model selection was conducted independently within each group, and only models with the most comparable architectural capacity and hyperparameter configurations were considered. This ensures a fair comparison under closely matched experimental conditions.

6.6.1 Preliminary Note on Data Availability

A meaningful observation that emerges before any substantive analysis is the widespread absence of computational benchmarking data across the reviewed studies. Of the 36 studies that reported quantitative performance comparisons, only a minority provided detailed information on training time, inference latency, or memory consumption. Specifically, training time data were available in 8 studies, inference time in 3, GPU memory consumption in 2, and parameter count in 12. The remaining studies either stated that this information was "Not Available" or omitted it entirely. This systematic gap reflects a field-wide tendency to prioritise predictive accuracy in publication over reproducibility and deployment considerations - a pattern consistent with the early adoption phase of a novel architecture. Consequently, the analysis of computational trade-offs presented below is necessarily based on a partial subset of the reviewed literature, and any conclusions should be understood as indicative rather than definitive. The absence of such reporting is, in itself, a finding that carries implications for the maturity and practical readiness of KAN-based methods.

6.6.2 Model Size and Parameter Efficiency

Among the dimensions of computational efficiency that were most consistently reported, model size—measured in terms of the number of trainable parameters—provides the clearest evidence in favour of KAN architectures. Across the 12 studies that supplied parameter counts for both KAN and baseline models, 9 showed that the KAN variant required substantially fewer parameters to achieve comparable or superior performance.

The reductions are, in several cases, dramatic. Zhou et al. demonstrated that the KAN-AD anomaly detection model achieved strong performance across multiple benchmark datasets with only 274 parameters, compared to 10,421 in the LSTM-based baseline—a reduction of approximately 97%. Similarly, the MT-KAN model for financial time series operated with 2,132 parameters against 32,071 for an LSTM baseline (93% fewer), while TimeKAN achieved competitive long-term forecasting performance with 12,840 parameters versus 841,570 in the iTransformer baseline (98% fewer). In the energy domain, a KAN model for smart grid battery capacity prediction required only 186 parameters compared to 481 for an equivalent MLP (61% fewer), and the MMK model for multivariate forecasting used 139,000 parameters against 972,000 for DLinear (86% fewer). These results suggest that the functional expressiveness afforded by learnable activation functions, a core property of KANs grounded in the Kolmogorov–Arnold representation theorem, can substitute for parameter depth in a way that feed-forward architectures cannot.

Nevertheless, the parameter efficiency advantage is not universal. Two notable exceptions were identified. SigKAN, a signature-weighted KAN variant designed for complex non-stationary financial signals, employed 114,711 parameters against 93,015 for its GRU baseline — a 23% increase. More strikingly, the Temporal Kolmogorov–Arnold Transformer (TKAT) used over 1.56 million parameters compared to 99,930 for GRU, a 15-fold increase attributable to the additional attention mechanism embedded in the architecture. These cases indicate that hybrid architectures combining KAN layers with transformers or signature methods may incur significant parameter overhead, negating the efficiency advantage observed in leaner standalone KAN models.

6.6.3 Training Complexity

The picture with respect to training time is considerably less favourable. Of the 8 studies that reported training duration for both the KAN model and its baseline, only two: the SOH battery estimation study and KAN-AD model showed a training time advantage for the KAN variant (15 seconds versus 22.5 seconds for TCN, a 33% reduction and 42 seconds vs 73 seconds respectively). In all remaining cases, KAN models required longer training. The Adp-3EncInformer-KAN for multi-energy load forecasting trained in 843 seconds versus 244 seconds for the baseline Transformer (a factor of $3.45\times$). The KARN model for short-term load forecasting required 70.5 minutes compared to 39.9 minutes for LSTM ($1.77\times$). The KANMTS multivariate prediction model trained in 27.96 seconds versus 10.98 seconds for MLP-SOFTS ($2.55\times$). In the biomedical domain, the KAN-UDE epidemiology model required 1,264 seconds compared to 670 seconds for MLP-UDE ($1.89\times$), though this must be contextualised against the fact that KAN-UDE reached 100% accuracy in 1,800 training iterations versus 4,600 for the MLP baseline – a 61% reduction in the number of gradient steps required. One additional study reported a total training time of 4,885 seconds for TSKANMixer, though without a comparable baseline figure, interpretation is limited.

Importantly, one outlier in the dataset warrants separate treatment. Study 2 reported KAN training on a CPU (Intel Xeon Gold 6338) at 2.92 seconds versus the LSTM baseline at 0.55 seconds – a factor of $5.3\times$. However, given the hardware discrepancy, this comparison cannot be treated as evidence of KAN’s intrinsic training overhead and should be interpreted cautiously.

The consistent pattern of longer training time is not unexpected from an architectural standpoint. KANs replace fixed activation functions with B-spline curves whose coefficients are learned during training, introducing additional computation per forward and backward pass compared to element-wise non-linearities such as ReLU or tanh. As the field matures, GPU-optimised implementations of spline operations are expected to reduce this overhead, but as of the studies reviewed, training efficiency remains a tangible limitation.

6.6.4 GPU Memory Footprint

GPU memory consumption was reported in only two studies, but the results are noteworthy. The TFKAN model, evaluated on the ILI and benchmark forecasting datasets, consumed 250.65 MB of GPU memory compared to 1,718.67 MB for the ATNet baseline – a reduction of approximately 85%. This aligns conceptually with the parameter efficiency results: a smaller model with fewer weights requires less memory for forward and backward activations. The KANMTS study reported a marginal reduction (256.01 MB versus 274.67 MB), which, given the nearly identical training times in that comparison, suggests that memory and speed scale differently and that memory efficiency may be a more consistent KAN advantage than training speed.

6.6.5 Challenges of Time Series: Concept Drift

Concept drift in time series refers to the evolving nature of the data-generating process, where the statistical properties of the series change over time in ways that may be abrupt, nonlinear,

or cyclical. This phenomenon undermines the assumptions of stationarity that underpin many traditional forecasting models, presenting a significant challenge in domains where adaptability is crucial such as financial modeling, clinical monitoring, and energy infrastructure management (Gama et al. 2014; Xu et al. 2024a). Many classical and deep learning-based approaches struggle to effectively model these shifts, often due to static architectures or a limited capacity to differentiate between transient noise and meaningful structural change (Pascanu et al. 2013).

Kolmogorov-Arnold Networks (KANs) offer a mathematically grounded and practically flexible approach that may help address some of these challenges. Unlike conventional neural networks, KANs replace static weight matrices with learnable, spline-parametrized univariate functions, allowing them to capture complex, nonlinear patterns with enhanced interpretability and potential adaptability (Liu et al. 2024; Xu et al. 2024b). This architectural property suggests KANs might adjust internal representations as data distributions evolve, making them potentially well-suited to scenarios involving concept drift.

WormKAN, a KAN-based autoencoder architecture, specifically targets concept drift identification and representation in co-evolving time series. It leverages latent representations shaped by a self-representation layer and temporal smoothness constraints, enabling the model to segment time series into distinct behavioral regimes. These transitions are described as "wormholes" - abrupt shifts in latent space that signify the onset of new conceptual patterns (Xu et al. 2024a). Experiments on financial market data and synthetic drift scenarios demonstrated that WormKAN can effectively identify concept transitions, though systematic evaluation against benchmark drift scenarios remains limited.

T-KAN and MT-KAN extend this line of research by incorporating symbolic regression and temporal modeling for time series forecasting. These variants aim to provide interpretable representations of how input-output relationships change over time, potentially enabling both detection and understanding of drift dynamics (Xu et al. 2024b). Initial results show competitive forecasting performance compared to traditional methods, though their explicit drift adaptation capabilities require further investigation.

In operational domains, KARN (Kolmogorov-Arnold Recurrent Network) integrates the spline-based adaptability of KANs with recurrent modeling capabilities to capture temporal dependencies across diverse consumer types in load forecasting tasks (Danish and Grolinger 2024). The model demonstrated improved accuracy over traditional RNNs in short-term load forecasting across residential, commercial, and industrial consumers, suggesting robustness to consumer-specific temporal patterns. However, the study does not explicitly evaluate performance under controlled concept drift scenarios or sudden distributional shifts.

6.7 Current State of Evidence and Future Research Directions

While the initial applications of KANs to time series problems show promise, the evidence base for their performance under concept drift conditions remains in an early stage of development. A systematic review of the literature reveals that only a small subset of KAN-related studies - primarily the work by Xu and colleagues (2024; 2024) directly evaluate drift detection and adaptation capabilities. The majority of existing applications demonstrate KAN effectiveness in forecasting

tasks under relatively stable or mildly non-stationary conditions, but do not systematically assess performance when confronted with distributional shifts characteristic of concept drift scenarios.

This limitation in the current evidence base points to several critical gaps that must be addressed to fully characterize KAN capabilities in drift-prone environments. First, the absence of systematic evaluation using controlled synthetic benchmarks such as those with predefined drift points, magnitudes, and transition patterns makes it difficult to establish performance boundaries or to compare drift detection latency and accuracy across different approaches. Standard benchmarks used in the concept drift literature, including synthetic streams with rotating hyperplanes, SEA concepts, or other well-characterized drift patterns, have not yet been employed to evaluate KAN-based methods.

Second, and perhaps most importantly, none of the reviewed studies compare KAN-based approaches against established concept drift detection and adaptation frameworks. Methods such as ADWIN (Adaptive Windowing), DDM (Drift Detection Method), EDDM (Early Drift Detection Method), and ensemble-based drift adaptation strategies have been extensively validated in the concept drift literature (Baena-García et al. 2006; Bifet and Gavaldà 2007; Gama et al. 2004). Without direct empirical comparison, it remains unclear whether KAN architectures offer fundamental advantages over these established approaches or simply represent an alternative modeling framework with different trade-offs.

Third, the types of drift evaluated in existing KAN studies are relatively narrow in scope. The WormKAN framework, for instance, focuses primarily on regime changes in financial time series and co-evolving data streams. Performance across a broader spectrum of drift phenomena including sudden versus gradual transitions, recurring versus novel concepts, real drift (changes in $P(y|X)$) versus virtual drift (changes in $P(X)$), and varying rates of drift remain largely uncharacterized. This limits our understanding of the boundary conditions under which KAN-based approaches are most effective.

The computational efficiency of KAN-based drift adaptation also warrants careful investigation. Concept drift scenarios often require rapid model updates or online learning capabilities to maintain predictive accuracy as distributions shift. The spline-based architecture of KANs, while offering enhanced interpretability and flexibility, introduces computational overhead relative to traditional neural architectures. Whether this overhead is acceptable in streaming data contexts, and how it compares to lightweight drift adaptation methods designed specifically for resource-constrained environments, has not been systematically evaluated.

Finally, questions remain regarding KAN performance during extreme events or sharp regime transitions. While some applications in streamflow forecasting demonstrate strong performance under typical hydrological conditions (Granata et al. 2024), the behavior of KAN-based models when confronted with distribution shifts that substantially deviate from training data such as those induced by rare events, structural breaks, or adversarial conditions requires further characterization.

Addressing these gaps will require a coordinated research effort encompassing several key directions. Systematic benchmarking using standard concept drift datasets would enable direct comparison with established baselines and help identify the specific scenarios where KAN architectures offer advantages. Evaluation across diverse drift types: abrupt, gradual, recurring, and incremental would clarify the range of applicability for KAN-based methods. Detailed analysis of detection la-

tency, false positive and false negative rates, and adaptation speed would provide the performance metrics necessary for practitioners to make informed decisions about model selection. Finally, investigation of hybrid approaches that combine KAN architectures with established drift detection mechanisms may offer a path toward leveraging the interpretability benefits of KANs while maintaining the proven reliability of existing drift adaptation frameworks.

Rather than suppressing the tension between stability and change, KAN-based models offer a framework for encoding it, enabling a paradigm shift in time series analysis from static prediction toward learning the logic of change itself. However, realizing this potential will require the rigorous evaluation outlined above, ensuring that KAN-based approaches to concept drift are grounded in systematic empirical evidence and positioned within the broader context of the drift adaptation literature.

7 Critical evaluation of KAN’s interpretability

While Kolmogorov-Arnold Networks (KANs) are frequently touted for superior interpretability compared to traditional multilayer perceptrons (MLPs), the empirical evidence supporting these claims requires careful scrutiny, particularly regarding the distinction between visualization capability and genuine interpretability that enables actionable decision-making. KAN interpretability operates through three primary mechanisms: (1) visualization of learned B-spline functions on network edges that reveal how input features are transformed (Barašin et al. 2025, p. 13); (Long and Qin 2025, Fig. 9), (2) network graph structures where edge thickness indicates connection importance (Barašin et al. 2025, Fig. 9), and (3) symbolic regression that can extract explicit mathematical formulas from trained networks (Long and Qin 2025, Figs. 9–10). For instance, Long and Qin (2025, Fig. 9) demonstrated that a KAN trained on the ETTh1 time series dataset could be pruned and subjected to symbolic regression to yield the explicit mathematical expression

$$X_{t+1} = e^{X_t^2} + \sin(3.14 \cdot X_{t-1}), \quad (3)$$

providing transparent insight into how historical values at times t and $t-1$ nonlinearly combine to generate future predictions. Similarly, Alves et al. (2024, pp. 4–5) applied KANs to wind nowcasting at Madeira International Airport, generating mathematical formulas describing wind components as

$$U\text{-KAN} = 4.0 (0.33 \sin(0.15V_{GFS} + 5.8) + 0.46 \tanh(0.17U_{GFS} + 0.23) + 1)^2 - 3.54, \quad (4)$$

$$V\text{-KAN} = 20.61 (-0.09 \tanh(0.17U_{GFS} + 0.39) - 0.31 \tanh(0.07V_{GFS} - 0.17) - 1)^2 - 20.41, \quad (5)$$

which achieved a 48.5% improvement over baseline Global Forecast System predictions (Alves et al. 2024, Table 2). These examples showcase KANs’ potential for extracting interpretable mathematical relationships from data, a capability distinct from traditional neural networks.

However, the practical utility and validity of this interpretability remains incompletely established across multiple dimensions. When Barašin et al. (2025, pp. 11–13) conducted a rigorous comparison of KAN interpretability against SHAP analysis on time series classification across 117 UCR benchmark datasets, the findings revealed both convergence and divergence that challenge simplistic interpretability claims. On the SmoothSubspace dataset, both methods identified features

x10 and x12 as top contributors for classification decisions, but the methods yielded complementary rather than identical insights: KAN’s B-spline visualizations revealed that x12 values in the range $[0,1]$ were amplified up to 300 while values outside this range were suppressed to between 0 and -200, thereby highlighting nonlinear feature transformations (Barašin et al. 2025, Fig. 9), whereas SHAP quantified marginal contributions across all features using cooperative game theory. Intriguingly, feature x5, which appeared as a top-3 contributor in the KAN graph structure visualization, was not ranked among the top contributors in SHAP analysis, demonstrating that “different interpretability methods can yield complementary insights, each capturing unique aspects of feature importance” (Barašin et al. 2025, p. 13). This discrepancy suggests that KAN’s edge weights and SHAP’s marginal contributions measure fundamentally different aspects of model behavior, raising questions about which metric better captures true interpretability for end-users. Furthermore, while Barašin et al. (2025, p. 15) emphasized that “KAN provides interpretability inherently through its graph structure and learned splines” whereas “MLP requires SHAP or similar tools to expose any interpretability,” this comparison conflates built-in visualization capabilities with genuine interpretability—both architectures ultimately require some form of analysis (either visual inspection of complex graphs or post-hoc SHAP) to extract meaning, and it remains unclear whether examining KAN graphs is cognitively simpler than interpreting SHAP summary plots.

The absence of domain expert evaluation studies represents perhaps the most critical gap in validating KAN interpretability claims. While Xu et al. (2024c, pp. 13–14) applied a hybrid DA-LSTM-KAN model to concrete dam deformation monitoring on the Minjiang River in China and validated that attention weight rankings (hydrostatic components 37.3%, seasonal 41.6%, irreversible 21.1%) aligned qualitatively with SHAP-based feature importance on LSTM models (hydrostatic 37.0%, seasonal 43.5%, irreversible 19.5%) and relative weight analysis from multiple linear regression (hydrostatic 37.9%, seasonal 40.2%, irreversible 21.9%), they conducted only quantitative validation of feature rankings without assessing whether civil engineering practitioners found the KAN attention visualizations more comprehensible, trustworthy, or actionable than existing explainability methods for dam safety decision-making. Similarly, while Jiang et al. (2025, Abstract) demonstrated that KANs could generate analytical expressions for short-term load forecasting in power systems—applications where transparency is critical for regulatory approval and operational trust—they provided no evidence that power system operators actually understood or utilized these symbolic formulas in practice, nor any comparison with how operators currently interpret and trust conventional forecasting models with post-hoc explanations. The literature consistently presents technical demonstrations of interpretability mechanisms without empirical validation from the domain experts who would actually use these models in high-stakes decisions, representing a fundamental disconnect between technical capability and practical utility.

Moreover, interpretability degrades substantially with network complexity, creating a fundamental performance-interpretability tradeoff that limits practical applicability. Barašin et al. (2025, p. 11) explicitly acknowledged this tension, noting that interpretability analysis required using a simplified architecture $[3, 15, 15, 15]$ with grid size $G=5$ that was “amenable to visualization,” yet observed that “smaller configurations lack the complexity required to model diverse time-series data effectively.” Indeed, across 117 UCR datasets, the best-performing interpretable KAN configuration achieved only $F1=0.54\pm0.22$, dramatically underperforming both the Efficient KAN implementation ($F1=0.70\pm0.23$), MLP baselines ($F1=0.64\pm0.25$), and especially state-of-the-art time series classifiers like InceptionTime ($F1=0.84\pm0.18$) and HIVE-COTE 2.0 ($F1=0.85\pm0.17$) (Barašin et al. 2025, Table 5, p. 12). This 30-36% performance gap relative to leading methods poses a critical question: in high-stakes applications, would practitioners accept substantially degraded predictive

accuracy for enhanced interpretability, and if so, what threshold of performance loss is acceptable?

The original KAN implementation performed even worse ($F1=0.30\pm0.21$), requiring 257,649 parameters and 6,146,993 FLOPs while achieving less than half the F1 score of MLPs that used 344,518 parameters and only 688,420 FLOPs (Barašin et al. 2025, Table 4, p. 10), demonstrating that interpretability can come at enormous computational and performance costs. Long and Qin (2025, Fig. 8, p. 11) reported that their dual-layer KAN structure reduced memory usage by 6.8% relative to baseline architectures (from 274.67MB to 256.01MB on ETTm2 dataset), while their ablation study (referencing Table 4) showed that dual-layer KAN improved feature interaction capture capability by 1.65% compared to single-layer design (Long and Qin 2025, p. 2), but inference times remained longer than fixed-activation models due to complex iterative B-spline operations that resist parallelization. These tradeoffs matter immensely in production settings: Alves et al. (2024, pp. 5–6) found that while KAN-enhanced wind forecasts improved accuracy, the models required careful hyperparameter tuning (grid number G , core sequence dimension, model layers L) that varied substantially between datasets—small datasets like ETTm1 needed fewer grids and layers, while large datasets like Weather required more complex architectures—suggesting that achieving good performance while maintaining interpretability demands significant domain expertise and computational resources.

The requirement for substantial post-hoc processing to achieve interpretable symbolic forms fundamentally undermines claims of interpretability “by design.” Long and Qin (2025, Fig. 9) demonstrated symbolic formula extraction from trained KANs, yielding

$$X_{t+1} = e^{X_t^2} + \sin(3.14 \cdot X_{t-1}), \quad (6)$$

but did not detail the pruning methodology required to achieve this simplification. The multi-stage processing needed to extract interpretable symbolic forms raises questions about whether KANs provide interpretability “by design” or simply offer different post-hoc interpretation tools comparable in complexity to applying SHAP or LIME to traditional networks. The fact that raw, unpruned KANs with dense connections and complex B-spline activations are not directly interpretable suggests that the architecture itself does not guarantee interpretability—rather, interpretability emerges only after significant additional processing that may be comparable in complexity to applying SHAP or LIME to MLPs.

The performance characteristics and computational costs of achieving interpretability reveal additional practical limitations. While Xu et al. (2024c, Table 2, p. 12) reported that their DA-LSTM-KAN model achieved superior performance for dam deformation monitoring on the test set (RMSE=0.1618 mm) compared to standalone LSTM (RMSE=0.2766 mm) and multiple linear regression (RMSE=0.3458 mm), they did not report whether the interpretability analysis added computational overhead or required simplified model variants. Jiang et al. (2025, p. 9) found that while KAN-based short-term load forecasting reduced computational costs through multi-interval linearization (partitioning ranges and using Taylor expansion approximations within sub-intervals), this efficiency gain came specifically from the linear approximation strategy rather than inherent KAN interpretability. The computational costs matter particularly for iterative model development, hyperparameter tuning, and real-time applications where inference latency is critical.

Furthermore, while KAN visualizations - particularly B-spline activation function plots showing learned univariate transformations, graph structures with edge thickness encoding connection importance, weight distribution heatmaps showing feature contributions across grid points and spline

orders, and information flow diagrams tracing data transformation through network layers offer valuable diagnostic tools that could facilitate model debugging, validation, and hypothesis generation (Barašin et al. 2025, Figs. 9–11; Long and Qin 2025), the cognitive complexity of interpreting these visualizations remains unquantified. A fully connected 3-layer KAN with even modest dimensions generates dozens to hundreds of B-spline plots that must be examined individually, with the relationship between local edge-level transformations and global network behavior being non-trivial to intuit. Barašin et al. (2025, p. 11) acknowledged this by focusing interpretability analysis on a single dataset (SmoothSubspace with 15 time steps and 3 classes) using a carefully tuned architecture, yet across their 117-dataset evaluation, most datasets had substantially longer sequences (mean length 537.10, maximum 2,844 time steps) (Barašin et al. 2025, Table 1, p. 7) that would require correspondingly more complex architectures and exponentially more visualizations to achieve competitive performance. The scalability of human interpretation of KAN visualizations to realistic problem sizes remains an open empirical question that directly impacts practical utility.

The field would benefit immensely from controlled, pre-registered studies comparing how effectively domain experts can extract actionable insights from different explanation modalities when performing realistic decision-making tasks. Such studies should employ rigorous experimental designs measuring: (1) comprehension accuracy (can users correctly identify which features drive predictions?), (2) trust calibration (do confidence levels appropriately match actual model reliability?), (3) decision quality (do users make better decisions with explanations than without?), (4) cognitive efficiency (how long does interpretation take and how mentally demanding is it?), and (5) actionability (can users effectively debug models, satisfy regulatory requirements, or generate novel hypotheses using the explanations?). These metrics should be assessed across participant groups (novices versus domain experts versus ML practitioners), problem domains (high-stakes medical diagnosis versus exploratory scientific discovery), and model complexities (simple 2-3 layer networks versus production-scale architectures). Until such evidence exists, claims that KAN visualizations are "more interpretable" than SHAP summary plots, LIME local explanations, attention heatmaps, saliency maps, or other established XAI methods remain plausible hypotheses rather than established facts.

The promising results from applied case studies - Alves et al. (2024, Table 2) achieving 48.5% error reduction in wind forecasting with extractable formulas, Xu et al. (2024c, Fig. 13, p. 14) providing attention-based feature importance that aligns with physical intuitions about dam behavior, Long and Qin (2025, p. 2) reporting 6.3% average MSE reduction across five benchmarks (as stated in their contributions) with 6.8% memory reduction (Fig. 8, p. 11) - demonstrate that KANs can simultaneously achieve competitive performance and provide multiple interpretability mechanisms, but without human-subjects validation, the actual interpretability benefit remains speculative. Particularly for high-stakes applications where interpretability directly impacts regulatory approval, operational trust, liability assignment, and human learning from models such as power grid management (Jiang et al. 2025), dam safety monitoring (Xu et al. 2024c), and financial forecasting - the burden of proof for interpretability claims should be correspondingly high, requiring not just demonstrations that models can be visualized or that symbolic formulas can be extracted, but rigorous evidence that these capabilities translate into measurably better human understanding, appropriate trust, and superior decision outcomes compared to existing explainability methods applied to high-performing but ostensibly less interpretable models.

7.1 Additional Critical Considerations

Several additional dimensions merit attention when evaluating KAN interpretability claims. First, the context-dependency of interpretability itself: what constitutes "interpretable" varies dramatically across stakeholders (data scientists versus domain experts versus regulatory auditors versus end-users) and applications (exploratory research versus operational deployment versus regulatory compliance). A mathematical formula like

$$X_{t+1} = e^{X_t^2} + \sin(3.14 \cdot X_{t-1}) \quad (7)$$

may be interpretable to mathematicians but opaque to clinical practitioners, while graph visualizations with edge weights may convey different information than SHAP's marginal contribution decompositions.

Second, the reproducibility and stability of interpretations: do KAN visualizations and extracted symbolic formulas remain consistent across different random initializations, training runs, and hyperparameter configurations? Jiang et al. (2025, p. 8) conducted 20 repeated experiments to assess reliability, but did not examine whether the interpretability artifacts (symbolic formulas, importance rankings) were stable across runs or whether different equally-accurate models yielded contradictory interpretations. If interpretation varies substantially with minor training differences, this undermines claims that KANs reveal "true" underlying relationships versus merely providing one of many equally-valid perspectives.

Third, the quantitative-qualitative interpretability gap: while studies demonstrate quantitative alignment between KAN attention weights and SHAP values (Xu et al. 2024c, Figure 13), or that symbolic formulas can be extracted algorithmically (Long and Qin 2025, Fig. 9); (Alves et al. 2024, pp. 4–5), these technical achievements do not necessarily translate into the qualitative interpretability that enables human insight, learning, and trust. The critical research gap lies not in whether interpretability mechanisms exist, but in whether they enable better human-AI collaboration in practice.

Finally, the opportunity cost consideration: research and engineering effort invested in developing, tuning, and deploying KAN-specific interpretability tools could alternatively be directed toward improving model-agnostic explainability methods (SHAP, LIME, integrated gradients, counterfactual explanations) that work across all architectures. Given that KANs currently underperform state-of-the-art models on many benchmarks (Barašin et al. 2025, Table 5), practitioners face a choice between accepting performance penalties for built-in KAN interpretability versus using best-in-class models with post-hoc explanations. The relative merits of this tradeoff remain an open empirical question requiring domain-specific evaluation.

8 Discussion

This systematic review examined Kolmogorov-Arnold Networks as an alternative architecture for time series analysis. Rather than simply summarizing findings, this discussion synthesizes performance patterns, identifies domain-specific insights, and critically assesses methodological quality to understand when and why KANs offer advantages over conventional approaches.

8.1 Performance Patterns: When KANs Excel and When They Struggle

The reviewed literature reveals clear patterns regarding conditions under which KANs consistently outperform baseline models versus scenarios where they show limitations.

8.1.1 Consistent Superior Performance

KANs demonstrate reliable advantages in short-term forecasting with smooth, structured patterns. In streamflow prediction, Granata et al. (2024) showed KANs substantially outperformed Transformers for 1-3 day forecasts, achieving R^2 values exceeding 0.95. Energy load forecasting consistently achieved 20-34% RMSE improvements (Liu et al. 2025), with Danish and Grolinger (2024) reporting SMAPE reductions from 0.4039 to 0.2873 for residential load prediction. Battery state-of-charge estimation achieved $R^2 = 0.9755$ versus 0.9281 for ANNs (Sulaiman et al. 2024a), while time series classification outperformed MLPs across 128 UCR datasets (Dong et al. 2024). Importantly, KANs maintained performance advantages under data scarcity, with ECG classification achieving F1=0.62 out-of-distribution with single-lead data (Huang et al. 2024).

8.1.2 Systematic Limitations

However, KANs show pronounced struggles in financial forecasting, where market dynamics involve high volatility and abrupt regime changes. The TKAN model achieved R^2 values of only 0.1378-0.1796 for cryptocurrency prediction (Genet and Inzirillo 2024a,b) - substantially below practical deployment thresholds. Performance degradation with extended forecast horizons represents another consistent pattern, with advantages diminishing beyond 3-day forecasts and converging with Transformers at 7-day horizons (Granata et al. 2024). The literature does not systematically evaluate challenges with high-dimensional inputs, very long sequences ($>$ few hundred timesteps), or abrupt distributional shifts beyond gradual drift. Beyond financial markets, isolated underperformance was also observed in probabilistic load forecasting against a Quantile-ANN baseline, in real-estate price prediction against XGBoost, and in one long-term benchmark suite where TimeKAN trailed iTransformer - suggesting that KAN advantages are not guaranteed even within domains where they typically excel.

8.2 Domain-Specific Insights: Why Performance Varies

The contrast between KAN effectiveness in energy forecasting versus financial markets illuminates fundamental relationships between architectural properties and problem characteristics.

Energy load forecasting represents near-optimal application for KANs. Energy demand exhibits strong autocorrelation and cyclical patterns driven by human behavior and weather cycles - structured temporal dependencies aligning with KANs' strength in learning smooth, continuous functions. Physical constraints (conservation laws, infrastructure capacities, thermal dynamics) impose regularity that KAN approximation theorems exploit. Additionally, interpretability has operational

value: utility operators benefit from understanding demand drivers, and KANs’ learnable activation functions provide more transparent feature-to-output mappings than black-box transformers.

In contrast, financial markets present characteristics fundamentally misaligned with KAN assumptions. Markets exhibit adversarial dynamics where participants exploit predictable patterns, eroding stable functions. The resulting non-stationarity involves abrupt regime switching rather than gradual drift. Volatility clustering, fat-tailed distributions, and black-swan events violate smooth-function assumptions underlying spline-based approximation. The low R^2 values achieved suggest that no learnable smooth function adequately captures market dynamics.

Streamflow forecasting, dam deformation monitoring, and wind prediction occupy intermediate positions. KAN success ($R^2 > 0.95$ for short-term streamflow) indicates effectiveness when physical laws constrain the problem space, while degradation at longer horizons reveals limits when systems enter less-constrained regimes. Medical time series show moderate improvements (F1=0.75 on ECG), suggesting advantages but not transformative breakthroughs.

8.3 Methodological Quality Assessment

Critical evaluation reveals significant limitations tempering conclusions about practical utility.

A striking pattern is the near-universal absence of statistical significance testing. The vast majority report point estimates (RMSE, MAE, R^2) without confidence intervals or p-values comparing KANs against baselines. Barašin et al. (2025) reported standard deviations (0.70 ± 0.23 for KAN versus 0.64 ± 0.25 for MLP) but without formal hypothesis testing. This absence is concerning given modest effect sizes when improvements range from 3-10%, understanding whether differences exceed random variation becomes critical.

Many studies employ single train-test splits without k-fold cross-validation. Out-of-sample validation on independent datasets remains rare. Huang et al. (2024) represents a positive exception, evaluating their ECG model on CPSC 2018 after training on TNMG and observing performance degradation (F1 from 0.75 to 0.62). Hyperparameter sensitivity is rarely evaluated systematically, making it unclear whether improvements stem from KAN advantages or more extensive tuning.

Notably, the vast majority of reviewed papers report positive results - of the 35 comparable studies, 27 showed KAN superiority, 2 produced mixed outcomes, and only 5 showed clear baseline advantage. While this distribution still suggests some degree of publication bias, the existence of documented underperformance cases across energy, financial, and classification domains partially mitigates concerns about a uniformly inflated evidence base.

An intriguing potential application emerges in social media bot detection. Bots exhibit temporal behavioral patterns distinguishing them from human users - posting frequency, activity timing, and engagement dynamics show mechanical regularity punctuated by coordinated bursts during campaigns, unlike natural human variability. KAN architectures may offer advantages through multi-scale temporal pattern modeling, interpretability for security analysts inspecting detection logic, and effectiveness under data scarcity for zero-day campaigns. However, bot behavior during coordinated attacks exhibits abrupt regime switching similar to financial markets, and adversarial

dynamics create an arms race that may exceed KAN capacity for learning stable functions. Nevertheless, the convergence of KAN strengths (interpretability, multi-scale modeling, data efficiency) with detection requirements suggests this as a promising direction meriting investigation.

KANs offer genuine advantages for structured, short-term forecasting where data exhibits smooth patterns and interpretability provides practical value. Energy systems, infrastructure monitoring, and select biomedical applications represent strong use cases. However, KANs do not represent universal improvement. Financial forecasting, abrupt regime switches, and very long-range modeling reveal clear limitations. Modest improvements combined with absent statistical testing raise questions about practical impact, while publication bias suggests domains where KANs excel are over-represented.

Future research should prioritize rigorous comparative evaluation using standardized benchmarks, statistical significance testing, and out-of-sample validation. Pre-registered studies with explicit success criteria would help mitigate publication bias. Most critically, interpretability claims require empirical validation through user studies - technical demonstrations that KANs produce symbolic formulas do not automatically translate into improved human understanding or decision-making.

In conclusion, Kolmogorov-Arnold Networks represent a promising architectural innovation with demonstrated advantages in specific domains characterized by smooth, structured temporal patterns. However, the evidence base remains preliminary, methodologically limited, and likely subject to publication bias. Practitioners should adopt KANs selectively for applications where architectural properties align with problem characteristics, while maintaining realistic expectations about performance gains and interpretability benefits.

9 Conclusions

This systematic review examined the suitability of Kolmogorov-Arnold Networks for time series analysis across 36 empirical studies, guided by three research questions addressing conditions of outperformance, interpretability, and computational efficiency.

On the question of when KANs outperform conventional architectures (RQ1), the evidence points to a clear domain dependency. To enable a transparent performance synthesis, we applied the following counting rule: a study was classified as a KAN win if the KAN variant achieved a numerically better value on the primary reported metric (lower error or higher accuracy/ R^2) under matched experimental conditions; studies reporting multiple conflicting metrics were classified as mixed; and the single study without any baseline (airport resilience, $R^2 = 0.99806$) was excluded from the count. Applying this rule to the 35 remaining studies, KANs achieved numerically superior performance in 27 cases (77%), produced mixed results in 2 cases, and underperformed the baseline in 5 cases. The five underperformance cases span distinct domains and warrant explicit acknowledgement: TimeKAN produced higher MSE than iTransformer on the benchmark suite (0.197 vs. 0.178); the KAN-based short-term probabilistic load forecasting model yielded substantially higher RMSE than the Quantile-ANN baseline (1,004 vs. 345); the KAN for real-estate price prediction was outperformed by XGBoost (RMSE 44,191 vs. 42,477); SigKAN marginally underperformed GRU on the 6-step cryptocurrency horizon (R^2 0.1349 vs. 0.1360); and the standalone KAN clas-

sifier on UCR datasets was slightly below the MLP baseline in test accuracy (0.775 vs. 0.782). Across the 27 wins, the median relative improvement on the primary metric was approximately 9.1% (interquartile range: 2.8%–32.4%), though this figure should be interpreted cautiously given the heterogeneity of metrics and domains involved. Advantages were most consistent in energy systems, battery state estimation, environmental monitoring, and biomedical signal analysis—domains characterised by smooth, physically constrained temporal patterns and short forecasting horizons. Financial forecasting represented the clearest limitation, where absolute predictive accuracy remained low regardless of relative gains, reflecting a fundamental mismatch between spline-based function approximation and the regime-switching, adversarially shaped dynamics of financial markets. The counterexamples identified above reinforce this picture: KANs underperformed in tasks involving high-dimensional non-stationary benchmark suites (TimeKAN), heavy-tailed probabilistic load scenarios (SPLF), tree-based competitors on tabular data (real estate), and near-random-walk financial signals (SigKAN, UCR classification). These cases are informative precisely because they define the boundary conditions of KAN applicability.

Regarding interpretability (RQ2), KANs offer technically demonstrable mechanisms - B-spline visualisations, graph-level importance representations, and extractable symbolic formulas that conventional architectures do not provide natively. However, none of the reviewed studies validated whether these mechanisms translate into improved comprehension, trust calibration, or decision quality among domain practitioners. The performance cost of configurations amenable to visual interpretation was substantial in the one rigorous comparative study available (Barašin et al. 2025), raising an unresolved trade-off that warrants empirical investigation through user studies before interpretability claims can be treated as established.

On computational efficiency (RQ3), KANs demonstrated clear parameter efficiency advantages reductions of 61–98% relative to baselines in the majority of studies reporting parameter counts and meaningfully lower GPU memory consumption where measured. Training time, however, was consistently longer, reflecting the overhead of spline coefficient optimisation. Critically, the majority of reviewed studies did not report any computational benchmarks, a gap that limits conclusions about deployment feasibility and represents a call for more rigorous reporting standards in future KAN research.

Taken together, the evidence supports a selective adoption perspective: KANs are a principled and well-performing choice for structured, short-horizon time series tasks where interpretability and model compactness carry operational value. They are not, on current evidence, a universal improvement over state-of-the-art architectures. Realising their full potential will require standardised benchmarking, rigorous evaluation under concept drift conditions, and empirical validation of interpretability claims - methodological investments that the field is well-positioned to make as the architecture matures.

Statements and Declarations

Funding

This work was supported by Narodowe Centrum Nauki (National Science Centre, Poland) under Grant 2020/38/A/HS6/00066.

Competing Interests

The authors have no competing interests to declare that are relevant to the content of this article.

Data Availability

No new datasets were generated during the current study. The evidence and extracted information analysed in this review are derived from the cited published studies.

Author Contributions

Antoni Mól prepared the initial draft and evidence synthesis. Dariusz Jemielniak and Leon Ciechanowski contributed to study conception, supervision, interpretation, and critical revision. All authors read and approved the final manuscript.

Use of AI Tools

Large language model-based AI tools were used for copy-editing and LaTeX coding.

Ethics Approval

Not applicable; this review did not involve human participants, human data, animals, or biological material.

Consent to Participate

Not applicable.

References

- Alves D, Mendonça F, Mostafa SS, Morgado-Dias F (2024) On the Use of Kolmogorov–Arnold Networks for Adapting Wind Numerical Weather Forecasts with Explainability and Interpretability: Application to Madeira International Airport. *Environmental Research Communications* 6:105008. <https://doi.org/10.1088/2515-7620/ad810f>.
- Ao SI, Fayek H (2023) Continual Deep Learning for Time Series Modeling. *Sensors* 23:7167.
- Baena-García M, Campo-Ávila J del, Fidalgo R, Bifet A, Gavaldà R, Morales-Bueno R (2006) Early Drift Detection Method. *Fourth International Workshop on Knowledge Discovery from Data Streams*. Vol. 6, pp 77–86.
- Barašin I, Bertalaníč B, Mohorčič M, Fortuna C (2025) Exploring Kolmogorov-Arnold Networks for Interpretable Time Series Classification. *International Journal of Intelligent Systems* 2025:9553189. <https://doi.org/10.1155/int/9553189>.
- Bifet A, Gavaldà R (2007) Learning from Time-Changing Data with Adaptive Windowing. *Proceedings of the 2007 SIAM International Conference on Data Mining*. SIAM, pp 443–448.
- Danish MU, Grolinger K (2024) Kolmogorov–Arnold Recurrent Network for Short Term Load Forecasting Across Diverse Consumers. *Energy Reports* 13:713–727. <https://doi.org/10.1016/j.egy.2024.12.038>.
- Dong C, Zheng L, Chen W, The University of Adelaide (2024) Kolmogorov-Arnold Networks (KAN) for Time Series Classification and Robust Analysis. <https://arxiv.org/abs/2408.07314>.
- Gama J, Medas P, Castillo G, Rodrigues P (2004) Learning with Drift Detection. *Brazilian Symposium on Artificial Intelligence*. Springer, pp 286–295.
- Gama J, Žliobaitė I, Bifet A, Pechenizkiy M, Bouchachia A (2014) A Survey on Concept Drift Adaptation. *ACM Computing Surveys* 46:1–37. <https://doi.org/10.1145/2523813>.
- Genet R, Inzirillo H (2024a) A Temporal Kolmogorov-Arnold Transformer for Time Series Forecasting. Version 2, 5 Jun 2024. <https://arxiv.org/abs/2406.02486>.
- Genet R, Inzirillo H (2024b) TKAN: Temporal Kolmogorov-Arnold Networks. Version 3, 17 Dec 2024. <https://arxiv.org/abs/2405.07344>.
- Granata F, Zhu S, Di Nunno F (2024) Advanced Streamflow Forecasting for Central European Rivers: The Cutting-Edge Kolmogorov-Arnold Networks Compared to Transformers. *Journal of Hydrology* 645:132175. <https://doi.org/10.1016/j.jhydrol.2024.132175>.
- Han X, Zhang X, Wu Y, Zhang Z (2024a) Are KAN and KAN-Based Models Effective for Time Series Forecasting? KAN4TSF. <https://github.com/2448845600/KAN4TSF>.
- Han X, Zhang X, Wu Y, Zhang Z, Wu Z (2024b) Are KANs Effective for Multivariate Time Series Forecasting? <https://doi.org/10.48550/arXiv.2408.11306>.
- Hong YC, Xiao B, Chen Y (2025) TSKANMixer: Kolmogorov–Arnold Networks with MLP-Mixer Model for Time Series Forecasting. <https://doi.org/10.48550/arXiv.2502.18410>.
- Hu Y, Xin X, Yu G, Deng W (2024) Deep Insight: An Efficient Hybrid Model for Oil Well Production Forecasting Using Spatio-Temporal Convolutional Networks and Kolmogorov-Arnold Networks. *Research Square preprint*. <https://doi.org/10.21203/rs.3.rs-4988083/v1>.
- Huang S, Zhao Z, Li C, Bai L (2025) TimeKAN: KAN-Based Frequency Decomposition Learning Architecture for Long-Term Time Series Forecasting. *International Conference on Learning Representations*. <https://openreview.net/forum?id=wTlc79YNbh>.
- Huang Z, Cui J, Yu L, Herbozo Contreras LF, Kavehei O (2024) Abnormality Detection in Time-Series Bio-Signals Using Kolmogorov-Arnold Networks for Resource-Constrained Devices. *medRxiv preprint*. <https://doi.org/10.1101/2024.06.04.24308428>.
- Inzirillo H, Genet R (2024a) A Gated Residual Kolmogorov-Arnold Networks for Mixtures of Experts. Version 2, 13 Dec 2024. <https://arxiv.org/abs/2409.15161>.

- Inzirillo H, Genet R (2024b) Signature-Weighted Kolmogorov-Arnold Networks for Time Series. Version 2. <https://arxiv.org/abs/2406.17890>.
- Jiang B, Wang Y, Li Y, Wang Q, Geng H (2025) A Novel Interpretable Short-Term Probabilistic Load Forecasting Method Based on Kolmogorov-Arnold Networks. IEEE journal article; code repository. <https://github.com/BozhenJIANG/KAN-for-SPLF>.
- Karnehm D, Samanta A, Rosenmüller C, Neve A, Williamson S (2024) Core Temperature Estimation of Lithium-Ion Batteries Using Long Short-Term Memory (LSTM) Network and Kolmogorov-Arnold Network (KAN). TechRxiv preprint. <https://doi.org/10.36227/techrxiv.172468367.74530953/v1>.
- Kontopoulou VI, Panagopoulos AD, Kakkos I, Matsopoulos GK (2023) A Review of ARIMA vs. Machine Learning Approaches for Time Series Forecasting in Data Driven Networks. Future Internet 15:255.
- Kui H, Zhou X, Zhao S, Guan J, Li W, Ye Y (2025) TFKAN: Time-Frequency KAN for Long-Term Time Series Forecasting. <https://doi.org/10.48550/arXiv.2506.12696>.
- Le N, Ngo AP, Nguyen HT (2024) Kolmogorov-Arnold Networks for Supervised Learning Tasks in Smart Grids. *2024 56th North American Power Symposium (NAPS)*, pp 1–6.
- Lee S, Kim JK, Kim J, Kim T, Lee J (2024) HiPPO-KAN: Efficient KAN Model for Time Series Analysis. <https://arxiv.org/abs/2410.14939>.
- Liu X, Yang Z, Guo Y, Li Z, Xu X (2025) A Novel Correlation Feature Self-Assigned Kolmogorov-Arnold Networks for Multi-Energy Load Forecasting in Integrated Energy Systems. Available at SSRN 4907222. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4907222.
- Liu Z, Wang Y, Vaidya S, Ruehle F, Halverson J, Soljačić M, Hou TY, Tegmark M (2024) KAN: Kolmogorov-Arnold Networks. <https://arxiv.org/abs/2404.19756>.
- Livieris IE (2024) C-KAN: A New Approach for Integrating Convolutional Layers with Kolmogorov-Arnold Networks for Time-Series Forecasting. Mathematics 12:3022. <https://doi.org/10.3390/math12193022>.
- Long Y, Qin X (2025) A Multivariate Time Series Prediction Model Based on the KAN Network. Scientific Reports 15:23621. <https://doi.org/10.1038/s41598-025-07654-7>.
- Ma K, Lu X, Bragazzi NL, Tang B (2024) Integrating Kolmogorov-Arnold Networks with Ordinary Differential Equations for Efficient, Interpretable and Robust Deep Learning: A Case Study in the Epidemiology of Infectious Diseases. medRxiv preprint. <https://doi.org/10.1101/2024.09.23.24314194>.
- Oad A, Wang Y, Fu X, Pudur N (2025) KASPER: Kolmogorov Arnold Networks for Stock Prediction and Explainable Regimes. <https://doi.org/10.48550/arXiv.2507.18983>.
- Pascanu R, Mikolov T, Bengio Y (2013) On the Difficulty of Training Recurrent Neural Networks. *Proceedings of the 30th International Conference on Machine Learning*. PMLR, pp 1310–1318.
- Peng J, Zhao X, Ma J, Meng D, Jia S, Zhang K, Gu C, Ding W (2024) State of Health Estimation of Li-Ion Battery via Incremental Capacity Analysis and Internal Resistance Identification Based on Kolmogorov-Arnold Networks. Batteries 10:315. <https://doi.org/10.3390/batteries10090315>.
- Schmidt-Hieber J (2021) The Kolmogorov-Arnold Representation Theorem Revisited. Neural Networks 137:119–126. <https://doi.org/10.1016/j.neunet.2021.01.020>.
- Song M, Wang J, Li R (2024) The Importance of Weather Factors on the Resilience of Airport Flight Based on Kolmogorov-Arnold Networks (KANs). Preprints.org preprint. <https://doi.org/10.20944/preprints202409.0800.v1>.
- Sulaiman MH, Mustaffa Z, Mohamed AI, Samsudin AS, Rashid MIM (2024a) Battery State of Charge Estimation for Electric Vehicle Using Kolmogorov-Arnold Networks. Energy 311:133417. <https://doi.org/10.1016/j.energy.2024.133417>.

- Sulaiman MH, Mustaffa Z, Saealal MS, Saari MM, Ahmad AZ (2024b) Utilizing the Kolmogorov-Arnold Networks for Chiller Energy Consumption Prediction in Commercial Building. *Journal of Building Engineering* 96:110475. <https://doi.org/10.1016/j.jobe.2024.110475>.
- Vaca-Rubio CJ, Pereira R, Blanco L, Zeydan E, Caus M (2025) A Primer on Kolmogorov-Arnold Networks (KANs) for Probabilistic Time Series Forecasting. <https://doi.org/10.48550/arXiv.2510.16940>.
- Viktoratos I, Tsadiras A (2025) Advancing Real-Estate Forecasting: A Novel Approach Using Kolmogorov-Arnold Networks (KANs). *Algorithms* 18:93. <https://doi.org/10.3390/a18020093>.
- Xu K, Chen L, Wang S (2024a) Are KAN Effective for Identifying and Tracking Concept Drift in Time Series? *NeurIPS 2024 Workshop on Time Series in the Age of Large Models*.
- Xu K, Chen L, Wang S (2024b) Kolmogorov-Arnold Networks for Time Series: Bridging Predictive Power and Interpretability. <https://arxiv.org/abs/2406.02496>.
- Xu R, Liu X, Wei J, Ai X, Li Z, He H (2024c) Predicting the Deformation of a Concrete Dam Using an Integration of Long Short-Term Memory (LSTM) Networks and Kolmogorov-Arnold Networks (KANs) with a Dual-Stage Attention Mechanism. *Water* 16:3043. <https://doi.org/10.3390/w16213043>.
- Yao X (2024) LSTM Model Enhanced by Kolmogorov-Arnold Network: Improving Stock Price Prediction Accuracy. *Trends in Social Sciences and Humanities Research* 2. <https://doi.org/10.61784/tsshr3015>.
- Zeng A, Chen M, Zhang L, Xu Q (2022) Are Transformers Effective for Time Series Forecasting? <https://doi.org/10.48550/arXiv.2205.13504>.
- Zhang Y, Cui L, Yan W (2025) Integrating Kolmogorov-Arnold Networks with Time Series Prediction Framework in Electricity Demand Forecasting. *Energies* 18:1365. <https://doi.org/10.3390/en18061365>.
- Zhao S, Jin M, Hou Z, Yang C, Li Z, Wen Q, Wang Y (2024) HiMTM: Hierarchical Multi-Scale Masked Time Series Modeling for Long-Term Forecasting. <https://doi.org/10.48550/arXiv.2401.05012>.
- Zhou Q, Pei C, Sun F, Han J, Gao Z, Pei D, Zhang H, Xie G, Li J (2024) KAN-AD: Time Series Anomaly Detection with Kolmogorov-Arnold Networks. <https://arxiv.org/abs/2411.00278>.