International Journal of Simulation, Optimization & Modelling
ISSN: 3083-967X

Statistical Modelling Approaches for Analyzing Patterns and Predicting Behavior in Complex Systems

Cut Fadhilah[1], Humasak Simajuntak[2], Rina Febrina[3]

[1]Faculty of Computer and Multimedia, Universitas Islam Kebangsaan Indonesia, Aceh 24251, Indonesia
[2]Departemen of Information System, Institut Teknologi Del, Sumatera Utara, Medan 22381, Indonesia
[3]Department of Civil Engineering, Universitas Malahayati, Indonesia

Corresponding Author: cutfadhilahzakaria@gmail.com

Abstract
Understanding and predicting behaviours in complex systems is a critical challenge across various fields, including climate science, financial markets, biological systems, and epidemiology. This study evaluates classical and modern statistical modelling approaches to analyze patterns and forecast outcomes in such systems. Classical methods, such as regression analysis and time series modelling, show high interpretability (effectiveness scores of 9 and 8, respectively) but are limited in handling nonlinearity and uncertainty (scores as low as 2–4). In contrast, modern techniques like machine learning models and Bayesian networks demonstrate superior performance in managing complexity and uncertainty (scores of 7–9), though they introduce greater computational demands. Using examples from real-world applications, including General Circulation Models (GCMs) in climate science and stochastic SEIR models in epidemiology, the study highlights that higher model complexity (score 8 in climate modelling) does not always guarantee higher prediction accuracy (score 7). In contrast, moderate complexity in epidemiological models (score 6) achieves excellent predictive performance (score 9). Model selection, validation, and interpretability challenges are discussed, emphasizing the trade-offs between complexity and practical usability. This research provides new insights by systematically comparing the predictive effectiveness and challenges across different application domains. The findings suggest that hybrid and domain-informed modelling strategies offer the best potential for improving prediction and understanding in complex systems. In conclusion, effective statistical modelling requires a balanced approach, integrating domain expertise, model adaptability, and ongoing validation to maximize interpretability and predictive accuracy.

Article Info
Received: 08 April 2025
Revised: 10 May 2025
Accepted: 11 May 2025
Available online: 21 May 2025

Keywords
Complex Systems
Statistical Modeling
Machine Learning
Model Validation
Predictive Analytics

1. Introduction

Complex systems are networks of numerous interconnected components whose interactions produce emergent, often unpredictable behaviours. These systems are pervasive across disciplines, encompassing ecosystems with dynamic species interactions, financial markets with nonlinear trader behaviour, and engineering networks such as power grids and transportation systems. Complex systems are characterized by nonlinearity, feedback loops, adaptability, and self-organization, making their study essential for understanding and managing critical societal infrastructures (Bahagia, Nizar, Yasin, Rosdi, & Faisal, 2025; Barabási, 2009; Nizar et al., 2025; Yana, Mufti, Hasiany, Viena, & Mahyudin,

©2025 The Author(s). Published by Scholar Publishing. This is an open access article under the CC BY license. Available online https://e-journal.scholar-publishing.org/index.php/ijsom

2025). In ecology, for instance, the dynamics of predator-prey relationships can shift dramatically with minor environmental changes, highlighting the delicate balance and sensitivity intrinsic to complex systems. Understanding complex systems' profound influence on stability, resilience, and functionality across sectors is essential. Financial crises (e.g., the 2008 collapse) have been attributed to hidden instabilities within economic networks (Almardhiyah, Mahidin, Fauzi, Abnisa, & Khairil, 2025; Gani, Saisa, et al., 2025; Haldane & May, 2011; Muzakki & Putro, 2025). At the same time, failures in engineered systems like power grids have been linked to cascading effects initiated by small perturbations (Buldyrev, Parshani, Paul, Stanley, & Havlin, 2010; Havlin et al., 2012; Irhamni, Kurnianingtyas, Muhtadin, Bahagia, & Yusop, 2025; Pranoto, Rusiyanto, & Fitriyana, 2025). Thus, unravelling the patterns and predictive mechanisms within complex systems is vital for pre-emptive interventions, policy formulation, and technological innovation across diverse fields.

One of the primary challenges in analyzing complex systems is their inherent high dimensionality and dynamic, evolving nature. Traditional linear models often fail to capture the intricate dependencies and emergent phenomena observed in these systems. Slight variations in initial conditions can lead to vastly different outcomes, a phenomenon known as sensitive dependence or the butterfly effect (Gani, Zaki, Bahagia, Maghfirah, & Faisal, 2025; Maghfirah, Yusop, & Zulkifli, 2025; Mufti, Irhamni, & Darnas, 2025; Toroczkai, 2010). Moreover, data scarcity, noise, and incomplete information frequently hamper efforts to build comprehensive models that generalize across different contexts and time frames. Feedback loops and adaptive behaviours within complex systems further complicate predictive modelling. For example, in financial markets, models must consider that agents adjust their strategies based on past outcomes and emerging information (Farmer & Foley, 2009; Rosdi, Maghfirah, Erdiwansyah, Syafrizal, & Muhibbuddin, 2025; Selvakumar, Gani, Xiaoxia, & Salleh, 2025; Selvakumar, Maawa, & Rusiyanto, 2025). Similarly, biological systems exhibit adaptability that renders static modelling approaches insufficient.

This dynamic adaptivity requires flexible methodologies, robust to uncertainty, and capable of learning from evolving data streams. Statistical modelling offers a structured framework for capturing uncertainty, identifying underlying patterns, and making informed predictions about complex system behaviours. Classical statistical methods such as time series analysis and multivariate regression have been widely employed to infer relationships and trends from observational data (Fitriyana, Rusiyanto, & Maawa, 2025; Muhibbuddin, Hamidi, & Fitriyana, 2025; Wulff, 2017; Zaki, Adisalamun, & Saisa, 2025). More recently, advanced approaches, including Bayesian networks, hidden Markov models, and machine learning-based statistical inference, have emerged to tackle complex systems' high-dimensional, nonlinear nature (Efremov & Kumarasamy, 2025; Khalisha, Caisarina, & Fakhrana, 2025; Li, Ikram, & Xiaoxia, 2025; Murphy, 2018). Through statistical modelling, researchers can construct probabilistic representations that quantify uncertainty and reveal latent structures within complex systems. Techniques like Bayesian inference allow for updating model beliefs as new data becomes available, providing adaptability essential for dynamic systems (Edition, 2013; Febrina & Anwar, 2025; NOOR, Arif, & Rusirawan, 2025; Sumarno, Fikri, & Irawan, 2025). Additionally, ensemble methods and hybrid models that combine statistical and computational approaches have successfully improved prediction accuracy and resilience to noisy, incomplete datasets (Iqbal, Rosdi, Muhtadin, Erdiwansyah, & Faisal, 2025; Kunapuli, 2023; Rosli, Xiaoxia, & Shuai, 2025; Xiaoxia, Lin, & Salleh, 2025).

This article provides a comprehensive overview of statistical modelling techniques for analyzing patterns and predicting behaviours in complex systems. It aims to bridge classical and modern methodologies, illustrating how evolving statistical tools are being adapted to address the multidimensional challenges presented by real-world systems. This article offers a roadmap for researchers and practitioners interested in leveraging statistical modelling to better understand and manage complexity by reviewing foundational theories, current practices, and emerging innovations. The scope of the discussion spans diverse application domains, including ecological modelling, financial market analysis, and engineered systems reliability studies. In statistical modelling efforts, particular attention is given to model selection strategies, validation techniques, and interpretability considerations. Furthermore, this article highlights case studies that exemplify successful applications of statistical models, shedding light on best practices and future research directions in complex systems analysis.

©2025 The Author(s). Published by Scholar Publishing. This is an open access article under the CC BY license. Available online https://e-journal.scholar-publishing.org/index.php/ijsom

2. Overview of Statistical Modeling Approaches

Classical statistical methods such as regression analysis and time series modelling have long been fundamental tools for analyzing patterns within complex systems. Regression analysis, both linear and nonlinear, enables researchers to quantify relationships between variables, providing insights into the underlying mechanisms driving system behaviours (Chatterjee & Hadi, 2015; Fox, 2015; Montgomery, Peck, & Vining, 2021; Yanti, Simajuntak, & Nurhanif, 2025). Time series models, including ARIMA (Auto Regressive Integrated Moving Average), are widely used to capture temporal dependencies and forecast future trends based on past observations (Wulff, 2017). These classical techniques offer clear interpretability and are computationally efficient, making them especially suitable for systems where relationships are relatively stable and linear approximations are acceptable. However, the limitations of classical methods become apparent when dealing with highly nonlinear or dynamic complex systems. For instance, traditional regression models often assume independence among observations and linearity, assumptions that rarely hold in real-world networks characterized by feedback loops and emergent behaviours (Shmueli & Koppius, 2011). Although capable of handling some forms of nonstationarity, time series models struggle when faced with abrupt regime changes or chaotic dynamics common in ecological, economic, and technological systems. Consequently, the need for more flexible and adaptive modelling approaches has led to the development of modern statistical techniques.

Modern statistical approaches, such as machine learning-based models and Bayesian networks, have emerged to address the complexities that classical methods cannot fully capture. Machine learning models, including random forests, support vector machines, and deep learning architectures, can model high-dimensional, nonlinear relationships without requiring explicit functional form specifications (John Lu, 2010; Khayum, Goyal, & Kamal, 2025). Bayesian networks, on the other hand, provide a probabilistic graphical framework that models conditional dependencies among variables, allowing for robust reasoning under uncertainty (Jalaludin, Kamarulzaman, Sudrajad, Rosdi, & Erdiwansyah, 2025; Sucar, 2021). These methods offer considerable advantages in uncovering latent structures, making dynamic predictions, and adapting to new information. Despite their strengths, modern statistical techniques also present challenges, particularly regarding interpretability and computational demands. Machine learning models of profound neural networks are often criticized as "black boxes," making it difficult to extract meaningful causal interpretations, a critical aspect for decision-making in many fields (Lipton, 2018; Muhtadin, Rosdi, Faisal, Erdiwansyah, & Mahyudin, 2025). Bayesian networks require extensive domain knowledge for proper structure learning and can become computationally intensive as the system's complexity increases (Heckerman, 1998). Therefore, selecting an appropriate modelling approach necessitates balancing predictive performance, interpretability, data availability, and computational resources based on the system under study.

Classical methods offer the advantage of simplicity, transparency, and well-established theoretical foundations, making them highly effective for systems where linearity, stationarity, and independence assumptions are approximately valid. Their interpretability facilitates hypothesis testing and model validation, crucial steps for scientific inquiry (Fox, 2015). However, their rigidity in handling nonlinearity and complex interactions limits their applicability to many modern problems involving dynamic feedback, emergent behaviours, or chaotic phenomena. In contrast, modern machine learning-based and Bayesian approaches excel in handling complexity, nonlinearity, and uncertainty, offering powerful tools for prediction and system understanding. These models are particularly valuable in high-dimensional settings where classical assumptions break down. Nevertheless, they often require large amounts of data for practical training, significant computational resources, and careful attention to avoid overfitting or spurious associations (Domingos, 2012; Rudin & Carlson, 2019). Thus, researchers must carefully consider the trade-offs between model complexity, interpretability, and practicality when choosing statistical methods to analyze and predict behaviours in complex systems.

Table 1. Comparison of Statistical Modeling Approaches

Modelling Approach	Strengths	Limitations
Regression Analysis	Interpretability, Simplicity	Assumes Linearity, Independence

©2025 The Author(s). Published by Scholar Publishing. This is an open access article under the CC BY license. Available online https://e-journal.scholar-publishing.org/index.php/ijsom

Modelling Approach	Strengths	Limitations
Time Series Modeling	Captures Temporal Patterns	Struggles with Regime Changes
Machine Learning Models	Handles Nonlinearity, High Dimensionality	Requires Large Data, Low Interpretability
Bayesian Networks	Models Uncertainty, Dependencies	High Computational Cost, Needs Domain Knowledge

Table 1 presents a comparative overview of four major statistical modelling approaches: Regression Analysis, Time Series Modeling, Machine Learning Models, and Bayesian Networks, highlighting their strengths and limitations when applied to complex systems. Regression analysis is praised for its interpretability and simplicity, making it an excellent choice when relationships between variables are expected to be linear and straightforward. However, its major limitation lies in the underlying assumptions of linearity and independence among observations, which are often violated in complex systems. Similarly, time series modelling is grounded in classical statistics and excels at capturing temporal patterns and making short-term forecasts. Yet, it struggles when faced with regime changes or structural breaks, which are standard in dynamic, evolving systems.

Modern methods, such as machine learning models, can handle nonlinearity and high-dimensional data, significantly enhancing prediction accuracy in complex settings. Nonetheless, they typically require large volumes of data for practical training and often suffer from low interpretability, posing challenges to understanding the causal mechanisms behind the predictions. Bayesian networks provide a robust framework for modelling uncertainty and dependencies among variables, supporting reasoning under incomplete information. However, they are computationally intensive and require substantial domain knowledge for accurate model specification and inference. While classical methods offer transparency and ease of use, they are less equipped to deal with the intricacies of highly nonlinear, adaptive systems. In contrast, modern statistical techniques, although powerful, demand greater computational resources and expertise to implement and interpret effectively.

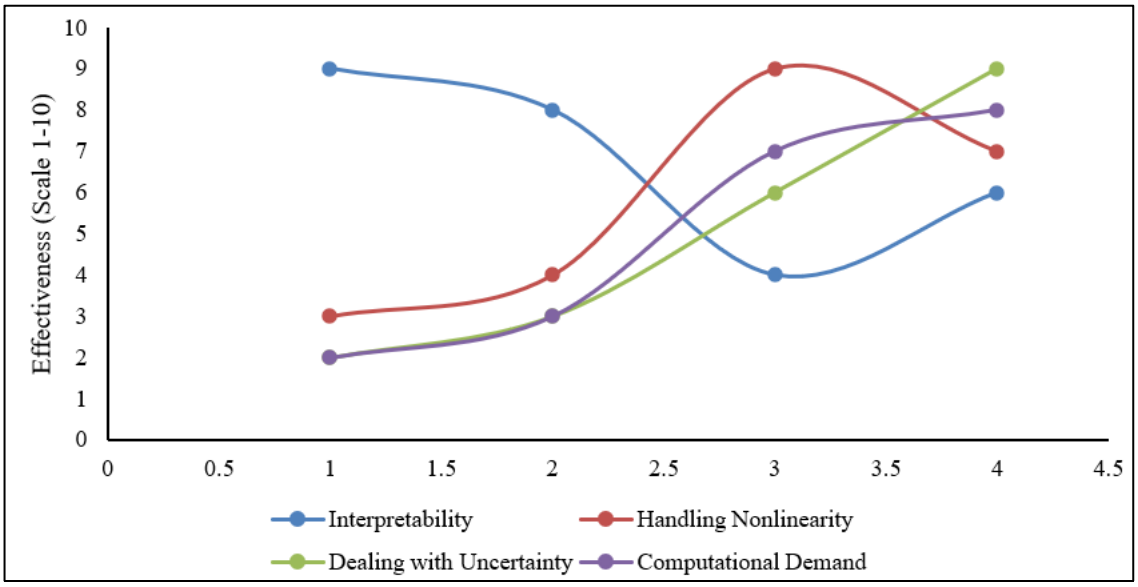

Fig. 1. Comparison of Statistical Modeling Approaches Across Key Dimensions

Fig. 1 illustrates the comparative effectiveness of various statistical modelling approaches across four critical dimensions: Interpretability, Handling Nonlinearity, Dealing with Uncertainty, and Computational Demand. The effectiveness is assessed on a scale from 1 to 10, reflecting the relative strength of each modelling characteristic. The chart shows that interpretability is highest in traditional models such as regression analysis and time series modelling, indicated by their high scores at the beginning of the curve. However, their effectiveness diminishes significantly as we move towards handling nonlinearity and dealing with uncertainty. In contrast, machine learning models and Bayesian

©2025 The Author(s). Published by Scholar Publishing. This is an open access article under the CC BY license. Available online https://e-journal.scholar-publishing.org/index.php/ijsom

networks demonstrate superior performance in handling nonlinearity and managing uncertainty, as reflected by their sharp upward trend in these dimensions. Nonetheless, these modern approaches also exhibit higher computational demand, evident from the increasing values toward the end of the graph. The trade-off between model complexity and ease of interpretation is thus clearly depicted, emphasizing the importance of careful method selection based on the specific needs and constraints of the complex system being studied.

3. Applications and Case Studies

Statistical modelling has been extensively applied in various domains to unravel patterns and predict behaviours in complex systems. General Circulation Models (GCMs) use statistical methods in climate science to simulate atmospheric, oceanic, and land surface processes, providing crucial insights into future climate scenarios under different emission pathways (Canadell et al., 2023). These models integrate large volumes of observational and simulated data to account for feedback loops and uncertainties in climatic systems. In economics, statistical and machine learning models are widely used to predict market behaviour, assess systemic risk, and simulate agent-based interactions, as exemplified by (Farmer & Foley, 2009), who emphasized using agent-based statistical models to replicate market crashes and bubbles. In biological systems, statistical approaches like Bayesian hierarchical models and machine learning techniques have advanced the understanding of genetic regulatory networks, disease progression patterns, and ecological dynamics. Bayesian inference was applied to predict species distribution shifts under climate change, considering habitat interactions and adaptation responses (Cánibe, Titeux, Domínguez, & Regos, 2022; Pearson, Dawson, & Liu, 2004). Similarly, statistical models have been pivotal in epidemiology, where they support real-time forecasting of disease outbreaks, such as the application of SEIR (Susceptible-Exposed-Infectious-Recovered) models enhanced with stochastic elements during the COVID-19 pandemic (Jewell, Lewnard, & Jewell, 2020). Choosing an appropriate statistical model for complex systems analysis remains a non-trivial task due to trade-offs among model complexity, data availability, interpretability, and predictive performance. Traditional criteria like the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) guide model selection by balancing model fit and parsimony (Guthery, 2003). However, cross-validation techniques and ensemble learning methods are often preferred in high-dimensional or nonlinear systems to mitigate overfitting risks and ensure generalizability (Varma & Simon, 2006). The difficulty is further compounded by the need to capture rare events or emergent behaviours that simple models might overlook. Model validation and interpretation present additional challenges, especially for complex systems where ground truth data may be scarce or evolving. Black-box models, such as deep learning architectures, can offer high predictive accuracy but often lack transparency, making it difficult to derive actionable insights (Doshi-Velez & Kim, 2017). Techniques like SHAP (Shapley Additive exPlanations) values and partial dependence plots have been developed to enhance model interpretability, but their use in very high-dimensional settings remains computationally intensive. Moreover, validation must consider predictive accuracy and robustness under different scenarios, an essential aspect when modeling dynamic systems like ecosystems or financial markets.

Successful applications of statistical modelling in complex systems have demonstrated the value of integrating domain knowledge with advanced analytics. For example, the hybrid use of physics-based models and machine learning in climate science has improved regional climate forecasts, addressing the limitations of purely empirical approaches (Reichstein et al., 2019). In financial systems, network-based statistical models have helped regulators identify systemic vulnerabilities by analyzing interconnected exposures, leading to more targeted risk management strategies (Battiston, Caldarelli, May, Roukny, & Stiglitz, 2016). Another key insight from successful applications is the importance of adaptability and model updating. Bayesian updating techniques and online learning algorithms have enabled models to incorporate new data and adjust predictions dynamically, enhancing performance in rapidly changing environments like infectious disease modelling (Brooks et al., 2020). These successes underscore that beyond technical sophistication, the practical effectiveness of statistical models in complex systems depends heavily on continuous validation, interpretability efforts, and alignment with evolving real-world conditions.

©2025 The Author(s). Published by Scholar Publishing. This is an open access article under the CC BY license. Available online https://e-journal.scholar-publishing.org/index.php/ijsom

Table 2. Applications of Statistical Modeling in Complex Systems

Domain	Statistical Method Used	Key Insights
Climate Modeling	General Circulation Models (GCMs)	Predicted future climate scenarios under different emissions pathways
Financial Markets	Agent-based Statistical Models	Simulated market crashes and bubbles, identified systemic risks
Biological Systems	Bayesian Hierarchical Models	Predicted species distribution shifts and genetic network behaviours
Epidemiology	Stochastic SEIR Models	Forecasted outbreak dynamics and intervention effects

Table 2 summarises the applications of statistical modelling approaches across various domains, highlighting the methods employed and key insights gained in each field. General Circulation Models (GCMs) have been pivotal in simulating future climate scenarios under different greenhouse gas emission pathways in climate modelling. These models integrate atmospheric, oceanic, and terrestrial processes, enabling researchers to forecast potential changes in temperature, precipitation, and extreme weather events with an increasing degree of confidence. Agent-based statistical models have successfully replicated phenomena such as market crashes and speculative bubbles in financial markets, providing valuable insights into systemic risks arising from interconnected financial institutions and investor behaviours.

In biological systems, Bayesian hierarchical models have facilitated predicting species distribution shifts and modelling complex genetic network behaviours, especially under environmental stress and climate change scenarios. Meanwhile, in epidemiology, stochastic SEIR (Susceptible-Exposed-Infectious-Recovered) models have been widely adopted to forecast outbreak dynamics and assess the potential impact of public health interventions. These models account for randomness in transmission and recovery processes, offering more realistic projections, particularly during fast-evolving epidemics like COVID-19. These applications demonstrate the adaptability and power of statistical modelling in addressing the multifaceted challenges posed by real-world complex systems.

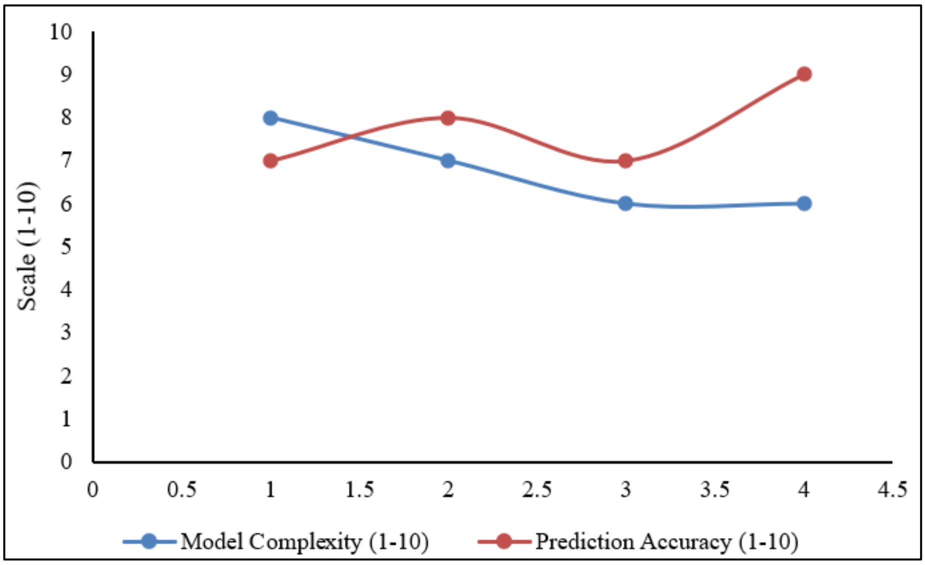

Fig. 2. Comparison of Model Complexity and Prediction Accuracy Across Domains

Fig. 2 compares the model complexity and prediction accuracy across four major application domains: climate modelling, financial markets, biological systems, and epidemiology, using a scale from 1 to 10. The figure shows climate modelling is associated with relatively high model complexity (score 8) but slightly lower prediction accuracy (score 7), reflecting the immense challenges of simulating interdependent atmospheric, oceanic, and terrestrial processes. In financial markets, while model complexity is slightly lower (score 7), prediction accuracy improves (score 8), likely due to the focus on agent-based statistical models tailored for specific market phenomena. Biological systems present a

©2025 The Author(s). Published by Scholar Publishing. This is an open access article under the CC BY license. Available online https://e-journal.scholar-publishing.org/index.php/ijsom

further drop in model complexity (score 6) while maintaining moderate prediction accuracy (score 7), highlighting the challenges of modelling genetic and ecological dynamics with limited data. Epidemiology, particularly during pandemic forecasting, demonstrates high prediction accuracy (score 9) despite moderate model complexity (score 6), attributed to the successful integration of stochastic SEIR models with real-time data. Overall, the graph illustrates that achieving high prediction accuracy does not always require extremely complex models, especially when models are appropriately tailored to the domain characteristics.

4. Conclusion

Statistical modelling has proven to be an indispensable tool in analysing and predicting behaviours within complex systems across diverse fields such as climate science, financial markets, biological systems, and epidemiology. Classical approaches, including regression analysis and time series modelling, offer strong interpretability (effectiveness scores 9 and 8, respectively) but show limitations in handling nonlinearity and uncertainty (scores as low as 2–4). In contrast, modern techniques such as machine learning models and Bayesian networks excel in managing complexity and uncertainty (scores 7–9), although they come with higher computational demands and challenges in interpretability. The analysed case studies highlight statistical models' versatility when adapted to domain-specific challenges. In climate modelling, General Circulation Models (GCMs) demonstrate high model complexity (score 8) but slightly lower prediction accuracy (score 7), reflecting the inherent uncertainty in simulating environmental systems. Financial market models show improved prediction accuracy (score 8) despite slightly reduced model complexity (score 7). In biological systems and epidemiology, stochastic and Bayesian approaches maintain moderate to high accuracy (scores 7 and 9, respectively) even with models of lower structural complexity (score 6), emphasizing that model efficiency and dynamic adaptability are critical for success. The insights gained from successful applications underscore that no single modelling approach is universally optimal. Instead, effective statistical modelling in complex systems requires a careful balance between model complexity, data availability, interpretability, and computational feasibility. As best practices, hybrid approaches, domain-informed model designs, and continuous model validation emerge. Future research should focus on developing interpretable machine learning models, enhancing real-time adaptability, and fostering cross-disciplinary integration to further advance statistical models' predictive power and utility in managing complexity.

Acknowledgement

The authors express their sincere gratitude for the financial support that made this research possible. This work was fully funded through the collective contributions of all authors. Their shared commitment and support were instrumental in completing this study.

References

Almardhiyah, F., Mahidin, M., Fauzi, F., Abnisa, F., & Khairil, K. (2025). Optimization of Aceh Low-Rank Coal Upgrading Process with Combination of Heating Media to Reduce Water Content through Response Surface Method. *International Journal of Science & Advanced Technology (IJSAT)*, 1(1), 29–37.

Bahagia, B., Nizar, M., Yasin, M. H. M., Rosdi, S. M., & Faisal, M. (2025). Advancements in Communication and Information Technologies for Smart Energy Systems and Renewable Energy Transition: A Review. *International Journal of Engineering and Technology (IJET)*, 1(1), 1–29.

Barabási, A.-L. (2009). Scale-free networks: a decade and beyond. *Science*, 325(5939), 412–413.

Battiston, S., Caldarelli, G., May, R. M., Roukny, T., & Stiglitz, J. E. (2016). The price of complexity

©2025 The Author(s). Published by Scholar Publishing. This is an open access article under the CC BY license. Available online https://e-journal.scholar-publishing.org/index.php/ijsom

in financial networks. *Proceedings of the National Academy of Sciences*, 113(36), 10031–10036.

Brooks, L. C., Ray, E. L., Bien, J., Bracher, J., Rumack, A., Tibshirani, R. J., & Reich, N. G. (2020). Comparing ensemble approaches for short-term probabilistic COVID-19 forecasts in the US. *International Institute of Forecasters*, 39.

Buldyrev, S. V, Parshani, R., Paul, G., Stanley, H. E., & Havlin, S. (2010). Catastrophic cascade of failures in interdependent networks. *Nature*, 464(7291), 1025–1028.

Canadell, J. G., Monteiro, P. M. S., Costa, M. H., Cotrim da Cunha, L., Cox, P. M., Eliseev, A. V, … Koven, C. (2023). Intergovernmental Panel on Climate Change (IPCC). Global carbon and other biogeochemical cycles and feedbacks. In *Climate change 2021: The physical science basis. Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change* (pp. 673–816). Cambridge University Press.

Cánibe, M., Titeux, N., Domínguez, J., & Regos, A. (2022). Assessing the uncertainty arising from standard land-cover mapping procedures when modelling species distributions. *Diversity and Distributions*, 28(4), 636–648.

Chatterjee, S., & Hadi, A. S. (2015). *Regression analysis by example*. John Wiley & Sons.

Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, 55(10), 78–87.

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *ArXiv Preprint ArXiv:1702.08608*.

Edition, S. (2013). Bayesian data analysis. CRC Press.

Efremov, C., & Kumarasamy, S. (2025). Optimisation of Microgrid by HOMER Pro Software Design: Innovative Approach and Performance Evaluation. *International Journal of Engineering and Technology (IJET)*, 1(1), 120–130.

Farmer, J. D., & Foley, D. (2009). The economy needs agent-based modelling. *Nature*, 460(7256), 685–686.

Febrina, R., & Anwar, A. (2025). Dynamic Modelling and Optimisation of Heat Exchange Networks for Enhanced Energy Efficiency in Industrial Processes. *International Journal of Simulation, Optimization & Modelling*, 1(1), 33–42.

Fitriyana, D. F., Rusiyanto, R., & Maawa, W. (2025). Renewable Energy Application Research Using VOSviewer software: Bibliometric Analysis. *International Journal of Science & Advanced Technology (IJSAT)*, 1(1), 92–107.

Fox, J. (2015). *Applied regression analysis and generalized linear models*. Sage publications.

Gani, A., Saisa, S., Muhtadin, M., Bahagia, B., Erdiwansyah, E., & Lisafitri, Y. (2025). Optimisation of home grid-connected photovoltaic systems: performance analysis and energy implications. *International Journal of Engineering and Technology (IJET)*, 1(1), 63–74.

Gani, A., Zaki, M., Bahagia, B., Maghfirah, G., & Faisal, M. (2025). Characterization of Porosity and Pore Volume in EFB Samples through Physical and Morphological Parameters. *International Journal of Engineering and Technology (IJET)*, 1(1), 90–99.

Guthery, F. S. (2003). Model selection and multimodel inference: a practical information-theoretic approach. JSTOR.

Haldane, A. G., & May, R. M. (2011). Systemic risk in banking ecosystems. *Nature*, 469(7330), 351–355.

Havlin, S., Araújo, N. A. M., Buldyrev, S. V, Dias, C. S., Parshani, R., Paul, G., & Stanley, H. E. (2012). Catastrophic cascade of failures in interdependent networks. In *Complex Materials in Physics and Biology* (pp. 311–324). IOS Press.

Heckerman, D. (1998). A tutorial on learning with Bayesian networks. *Learning in Graphical Models*, 301–354.

Iqbal, I., Rosdi, S. M., Muhtadin, M., Erdiwansyah, E., & Faisal, M. (2025). Optimisation of combustion parameters in turbocharged engines using computational fluid dynamics modelling. *International Journal of Simulation, Optimization & Modelling*, 1(1), 63–69.

Irhamni, I., Kurnianingtyas, E., Muhtadin, M., Bahagia, B., & Yusop, A. F. (2025). Bibliometric Analysis of Renewable Energy Research Trends Using VOSviewer: Network Mapping and Topic Evolution. *International Journal of Engineering and Technology (IJET)*, 1(1), 75–82.

Jalaludin, H. A., Kamarulzaman, M. K., Sudrajad, A., Rosdi, S. M., & Erdiwansyah, E. (2025). Engine

©2025 The Author(s). Published by Scholar Publishing. This is an open access article under the CC BY license. Available online https://e-journal.scholar-publishing.org/index.php/ijsom

Performance Analysis Based on Speed and Throttle Through Simulation. *International Journal of Simulation, Optimization & Modelling*, 1(1), 86–93.

Jewell, N. P., Lewnard, J. A., & Jewell, B. L. (2020). Predictive mathematical models of the COVID-19 pandemic: underlying principles and value of projections. *Jama*, 323(19), 1893–1894.

John Lu, Z. Q. (2010). The elements of statistical learning: data mining, inference, and prediction. Oxford University Press.

Khalisha, N., Caisarina, I., & Fakhrana, S. Z. (2025). Mobility Patterns of Rural Communities in Traveling Traveling from The Origin Area to the Destination. *International Journal of Science & Advanced Technology (IJSAT)*, 1(1), 108–119.

Khayum, N., Goyal, R., & Kamal, M. (2025). Finite Element Modelling and Optimisation of Structural Components for Lightweight Automotive Design. *International Journal of Simulation, Optimization & Modelling*, 1(1), 78–85.

Kunapuli, G. (2023). *Ensemble methods for machine learning*. Simon and Schuster.

Li, D., Ikram, M., & Xiaoxia, J. (2025). A brief overview of the physical layer test system: Development of an IoTbased energy storage and electrical energy distribution system. *International Journal of Engineering and Technology (IJET)*, 1(1), 131–140.

Lipton, Z. C. (2018). The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue*, 16(3), 31–57.

Maghfirah, G., Yusop, A. F., & Zulkifli, Z. (2025). Using VOSviewer for Renewable Energy Literature Analysis: Mapping Technology and Policy-Related Research. *International Journal of Engineering and Technology (IJET)*, 1(1), 83–89.

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2021). *Introduction to linear regression analysis*. John Wiley & Sons.

Mufti, A. A., Irhamni, I., & Darnas, Y. (2025). Exploration of predictive models in optimising renewable energy integration in grid systems. *International Journal of Science & Advanced Technology (IJSAT)*, 1(1), 47–61.

Muhibbuddin, M., Hamidi, M. A., & Fitriyana, D. F. (2025). Bibliometric Analysis of Renewable Energy Technologies Using VOSviewer: Mapping Innovations and Applications. *International Journal of Science & Advanced Technology (IJSAT)*, 1(1), 81–91.

Muhtadin, M., Rosdi, S. M., Faisal, M., Erdiwansyah, E., & Mahyudin, M. (2025). Analysis of NOx, HC, and CO Emission Prediction in Internal Combustion Engines by Statistical Regression and ANOVA Methods. *International Journal of Simulation, Optimization & Modelling*, 1(1), 94–102.

Murphy, K. P. (2018). Machine learning: A probabilistic perspective (adaptive computation and machine learning series). *The MIT Press: London, UK*.

Muzakki, M. I., & Putro, R. K. H. (2025). Greenhouse Gas Emission Inventory at Benowo Landfill Using IPCC Method. *International Journal of Science & Advanced Technology (IJSAT)*, 1(1), 18–28.

Nizar, M., Syafrizal, S., Zikrillah, A.-F., Rahman, A., Hadi, A. E., & Pranoto, H. (2025). Optimizing Waste Transport Efficiency in Langsa City, Indonesia: A Dynamic Programming Approach. *International Journal of Science & Advanced Technology (IJSAT)*, 1(1), 10–17.

NOOR, C. H. E. W. A. N. M., Arif, F., & Rusirawan, D. (2025). Optimising Engine Performance and Emission Characteristics Through Advanced Simulation Techniques. *International Journal of Simulation, Optimization & Modelling*, 1(1), 10–20.

Pearson, R. G., Dawson, T. P., & Liu, C. (2004). Modelling species distributions in Britain: a hierarchical integration of climate and land-cover data. *Ecography*, 27(3), 285–298.

Pranoto, H., Rusiyanto, R., & Fitriyana, D. F. (2025). Sustainable Wastewater Management in Sumedang: Design, Treatment Technologies, and Resource Recovery. *International Journal of Science & Advanced Technology (IJSAT)*, 1(1), 38–46.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat, F. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, 566(7743), 195–204.

Rosdi, S. M., Maghfirah, G., Erdiwansyah, E., Syafrizal, S., & Muhibbuddin, M. (2025). Bibliometric Study of Renewable Energy Technology Development: Application of VOSviewer in Identifying Global Trends. *International Journal of Science & Advanced Technology (IJSAT)*, 1(1), 71–80.

©2025 The Author(s). Published by Scholar Publishing. This is an open access article under the CC BY license. Available online https://e-journal.scholar-publishing.org/index.php/ijsom

Rosli, M. A., Xiaoxia, J., & Shuai, Z. (2025). Machine Learning-Driven Optimisation of Aerodynamic Designs for High-Performance Vehicles. *International Journal of Simulation, Optimization & Modelling*, 1(1), 43–53.

Rudin, C., & Carlson, D. (2019). The secrets of machine learning: Ten things you wish you had known earlier to be more effective at data analysis. In *Operations research & management science in the age of analytics* (pp. 44–72). Informs.

Selvakumar, P., Gani, A., Xiaoxia, J., & Salleh, M. R. (2025). Porosity and Pore Volume Analysis of EFB Fiber: Physical Characterization and Effect of Thermal Treatment. *International Journal of Engineering and Technology (IJET)*, 1(1), 100–108.

Selvakumar, P., Maawa, W., & Rusiyanto, R. (2025). Hybrid Grid System as a Solution for Renewable Energy Integration: A Case Study. *International Journal of Science & Advanced Technology (IJSAT)*, 1(1), 62–70.

Shmueli, G., & Koppius, O. R. (2011). Predictive analytics in information systems research. *MIS Quarterly*, 553–572.

Sucar, L. E. (2021). *Probabilistic graphical models*. Springer.

Sumarno, R. N., Fikri, A., & Irawan, B. (2025). Multi-objective optimisation of renewable energy systems using genetic algorithms: A case study. *International Journal of Simulation, Optimization & Modelling*, 1(1), 21–32.

Toroczkai, Z. (2010). Complexity: A guided tour. American Institute of Physics.

Varma, S., & Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics*, 7, 1–8.

Wulff, S. S. (2017). Time series analysis: Forecasting and control. Taylor & Francis.

Xiaoxia, J., Lin, D., & Salleh, M. Z. (2025). Mathematical Modelling and Optimisation of Supply Chain Networks Under Uncertain Demand Scenarios. *International Journal of Simulation, Optimization & Modelling*, 1(1), 54–62.

Yana, S., Mufti, A. A., Hasiany, S., Viena, V., & Mahyudin, M. (2025). Overview of biomass-based waste to renewable energy technology, socioeconomic, and environmental impact. *International Journal of Engineering and Technology (IJET)*, 1(1), 30–62.

Yanti, Y., Simajuntak, H., & Nurhanif, N. (2025). Integrated simulation and optimisation of traffic flow management systems in urban smart cities. *International Journal of Simulation, Optimization & Modelling*, 1(1), 70–77.

Zaki, M., Adisalamun, A., & Saisa, S. (2025). Analysis of the concentration microplastics on waste generation in the beach: A case study Banda Aceh City, Indonesia. *International Journal of Engineering and Technology (IJET)*, 1(1), 109–119.

©2025 The Author(s). Published by Scholar Publishing. This is an open access article under the CC BY license. Available online https://e-journal.scholar-publishing.org/index.php/ijsom