e-ISSN: 2582-7464

Computational Intelligence and Machine Learning
Volume 6, Issue 2, October 2025

Artificial Intelligence-Based Time Series Forecasting for Operational Risk Management in Healthcare: A Comparative Study of ARIMA, SARIMA, and Prophet on the EPA Platform

Mateus Lacerda Alves Silva Marinho[1], Isadora Stéfany Rezende Remigio Mesquita[2], Lucas Elias Cardoso Rocha[3], Daniel Alves Vieira[4], Taciana Novo Kudo[5], Renata Dutra Braga[6*]

[1,2,3,5,6] Center of Excellence in Artificial Intelligence, Federal University of Goiás, Goiânia, Goiás, Brazil
[4] Simeon Strategy, Development and Data Processing LTDA, Goiânia, Goiás, Brazil
*Corresponding Author: renatadbraga@ufg.br

Abstract

This study presents the development and evaluation of time series forecasting models for operational risk management in healthcare institutions, using real data from the EPA Platform. The objective was to predict the number of risk-related occurrences based on historical records, comparing the performance of ARIMA, SARIMA, and Prophet. The models were applied to both daily and monthly granularities, which enabled the identification of the most effective forecasting strategies for each temporal context. The approach included seasonal pattern analysis and hyperparameter optimization through Grid Search. Experimental results highlighted each model's ability to capture complex temporal dynamics while maintaining low computational cost. The findings supported strategic decision-making aimed at risk mitigation and patient safety, and demonstrated the practical applicability of AI-based forecasting in institutional healthcare environments.

Keywords

Artificial Intelligence, Time Series Forecasting, ARIMA, SARIMA, Prophet, Healthcare Risk Management

INTRODUCTION

Risk management in healthcare institutions is a strategic activity that directly affects patient safety, organizational efficiency, and service quality. The growing complexity of clinical and administrative operations, multiple data sources, and variability in record keeping make early risk identification and preventive measures particularly challenging. In this context, Artificial Intelligence (AI)-based solutions have gained prominence due to their ability to extract patterns from large datasets and support evidence-based decision-making. As highlighted in previous studies [1], research in this area is still evolving, with big data applications lacking clear direction and integration. A deeper understanding of these challenges and exploration of new opportunities is essential to improve safety, reliability, and efficiency across complex systems.

Despite these advances, many healthcare organizations still rely on conventional systems that cannot effectively anticipate adverse events. Without intelligent tools, institutions remain unprepared for recurring risks and cannot effectively reuse knowledge from past analyses. Root cause analysis also remains highly dependent on expert judgment, resulting in subjectivity and low standardization [2].

This study was conducted as part of an innovation project by the Center of Excellence in Artificial Intelligence at the Federal University of Goiás (CEIA/UFG), in partnership with Simeon Estratégia, Desenvolvimento e Processamento de Dados LTDA. Simeon develops the EPA Software — Estratégia para Ação (Strategy for Action) — a corporate platform used by Brazilian healthcare institutions for planning, quality management, incident control, and policy execution. The goal was to enhance EPA's Quality Policy module using AI techniques to support risk and incident management processes [3].

The proposed solution compares three statistical models for time series forecasting—Autoregressive Integrated Moving Average (ARIMA) [4], Seasonal Autoregressive Integrated Moving Average (SARIMA) [5], and Prophet—for forecasting operational risks at daily and monthly granularities. Model selection and calibration were guided by stationarity tests, autocorrelation and partial autocorrelation functions (ACF/PACF), and hyperparameter tuning via grid search. Although detailed results are presented later, our experiments demonstrated both the practical applicability and computational feasibility of the proposed approach.

From a software engineering perspective, the solution was built using a modular architecture with Machine Learning Operations (MLOps) pipelines, integrated with EPA's operational database via RESTful APIs. The user interface, developed in React/Next.js, provides controlled access to dashboards that display unit-level risk visualizations and forecasts. The system supports multiple clients in a multi-tenant model, ensuring security, scalability, and interoperability.

This paper is organized as follows: Section II reviews related works. Section III presents the background on time series forecasting. Section IV describes the dataset and experiments. Section V explains the methodology. Section VI reports the results. Section VII discusses limitations and error analysis. Section VIII concludes with final remarks and directions for future research.

RELATED WORKS

The adoption of AI in healthcare risk management has increased in recent years, driven by the complexity of care systems and the demand for faster, data-informed decisions. Machine learning techniques have been applied to tasks such as demand forecasting, adverse event detection, and clinical decision support. However, few studies integrate predictive modeling with automated causal analysis and recommendations for corrective actions in institutional healthcare environments [6].

Time series models have traditionally been used for epidemiological forecasting. For example, one study [7] analyzed monthly SINAN data (Brazil's Notifiable Diseases Information System) to estimate tuberculosis incidence through 2030. The authors used ARIMA and SARIMA models with stationarity tests (ADF, KPSS) and information criteria (AIC, BIC). They achieved a MAPE of 4.5% in non-pandemic scenarios, showing the accuracy of these methods for capturing trends and seasonality. Despite this robustness, the study did not address intra-organizational risk management or integration with preventive actions.

At the operational level, some authors modeled hourly emergency care visits in Florianópolis using Prophet, ARIMA, and Generalized Regression Neural Networks (GRNN). Their goal was to support real-time staff and supply allocation. The GRNN achieved the best performance (RMSE = 17.71), demonstrating the feasibility of high-temporal-resolution forecasting. However, the study focused only on demand forecasting, without addressing explainability, automation, or integration with quality management platforms [8].

Other studies linked predictive modeling to clinical risk mitigation. For instance, [9] applied an AI model in a proactive care program for high-risk patients using interrupted time series analysis. The intervention reduced potentially avoidable admissions by 27%, reinforcing AI's role in prevention.

In a different use case, [10] implemented a no-show prediction model for primary care appointments, achieving 86% accuracy and reducing missed appointments by over 50%. This improved both patient flow and average service times.

Additionally, [11] proposed a big data and machine learning framework for chronic disease management and post-discharge prevention, combining interoperability, supervised learning, and recommendation mechanisms.

A comprehensive review by [2] highlighted forecasting, anomaly detection, decision support, and risk mitigation as the main domains of AI in healthcare safety. The authors emphasized modular architectures, continuously updated pipelines, and hybrid approaches involving time series, natural language processing, and deep learning. They noted the lack of solutions that combine forecasting with actionable insights and prescriptive recommendations, especially in non-clinical contexts such as institutional risk management systems.

Against this backdrop, there is a gap in the literature regarding end-to-end solutions that combine: (i) event forecasting with tactical granularity; and (ii) evidence-based recommendations for actionable responses. This study addresses that gap by comparing ARIMA, SARIMA, and Prophet models embedded in the real-world context of the EPA Platform, integrated into its Quality Policy module. This ensures practical applicability, regulatory alignment, and strong potential for technology transfer.

BACKGROUND

Time series forecasting estimates future values from past observations ordered chronologically. In healthcare, it is essential for anticipating clinical, operational, or epidemiological trends—such as patient volume, disease incidence, adverse events, and institutional risks [12]. Forecasting supports resource allocation, proactive mitigation, and evidence-based decisions.

Hospital discharge prediction has been highlighted as a key use case. One study compared Prophet and SARIMA, showing Prophet's superior accuracy across datasets from two hospitals, with stable forecasts over a one-year horizon [13].

The ARIMA family of models captures autoregressive (AR), moving average (MA), and differencing (I) components, representing dependence on past values, residual errors, and the need for stationarity adjustments. ARIMA is effective for univariate linear series without strong seasonality. The modeling process involves ACF/PACF analysis, stationarity testing, and parameter tuning to minimize AIC or BIC. In healthcare, ARIMA has been used to forecast daily radiotherapy patient volumes, supporting resource planning [4].

SARIMA extends ARIMA with seasonal components to capture recurring patterns such as monthly or annual disease

COMPUTATIONAL INTELLIGENCE AND MACHINE LEARNING
e-ISSN: 2582-7464
Computational Intelligence and Machine Learning
Volume 6, Issue 2, October 2025

cycles. It has been applied to dengue incidence prediction, demonstrating its value for epidemiological surveillance [5].

Prophet, developed by Meta, uses an additive model to decompose time series into trend, seasonality, and holiday effects. It is robust to missing data, trend shifts, and outliers, while requiring less complex parameterization than ARIMA or SARIMA. Prophet also allows explicit inclusion of domain knowledge (e.g., national holidays). In healthcare, it has been applied to COVID-19 forecasts, achieving accuracy comparable to ARIMA with easier deployment [14].

Model selection depends on series characteristics, data granularity, seasonality, and forecast horizon. In this study, ARIMA, SARIMA, and Prophet were applied to institutional risk data at daily and monthly granularities to evaluate their predictive accuracy and suitability.

DATASET

The dataset was structured into two granularities: daily and monthly. For each monitored risk, two independent time series were generated, allowing comparison of model performance across different levels of temporal aggregation. This approach was essential to assess the sensitivity of algorithms to structural variations and to identify the most effective configuration for short-term monthly forecasting — the main objective of this study [15]. Figure 1 shows the monthly-aggregated time series, while Figure 2 illustrates the same series at daily granularity.

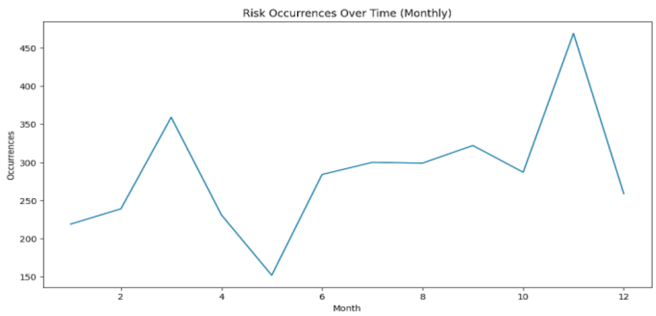

Figure 1. Monthly-aggregated time series displaying the frequency of monthly occurrences for the selected operational risk.

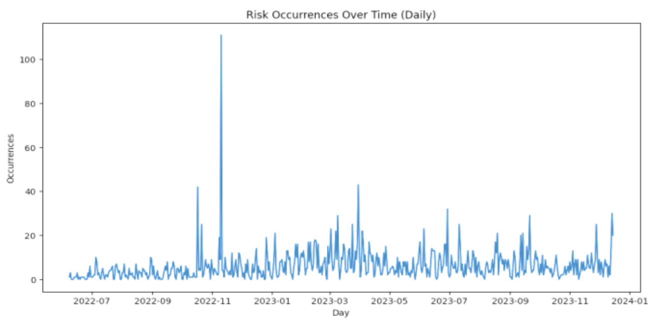

Figure 2. Daily time series of the same risk, illustrating finer-grained fluctuations and seasonal patterns.

This dual approach was justified by complementary advantages. Monthly granularity provided a compact and manageable structure, reducing the number of forecasts required for a three-month horizon (three predictions per risk). This favored operational simplicity and statistical stability, especially when record volume was low or interday variability limited [15]. Figure 3 presents monthly box plots by risk type, while Figure 4 shows daily box plots with more pronounced outliers and seasonal variability.

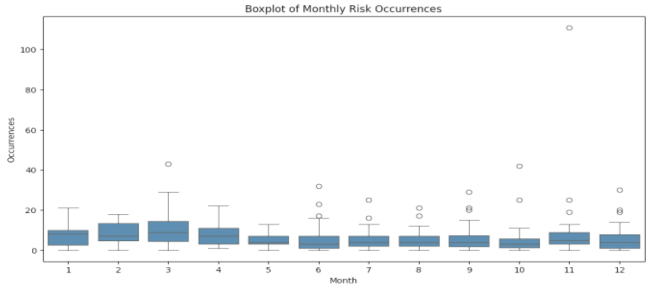

Figure 3. Boxplot of monthly occurrence distribution across different risk types.

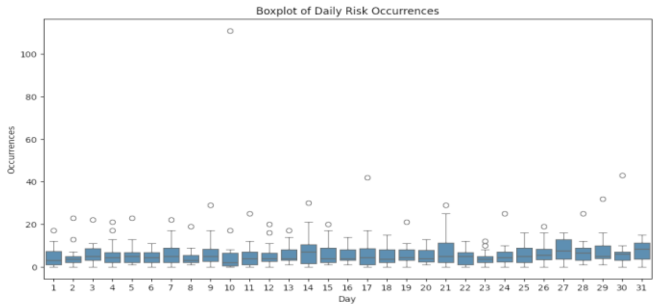

Figure 4. Daily boxplot of occurrence distribution, emphasizing outliers and seasonal variability.

The daily dataset, in contrast, increased information density, allowing models to capture seasonal trends, weekly cycles, and sporadic fluctuations in greater detail. Although this granularity required 90–92 forecasts for a three-month horizon, the gain in sensitivity often offset the higher computational cost [15]. Figure 5 highlights daily operational cycles using a heatmap of occurrences by weekday and month.

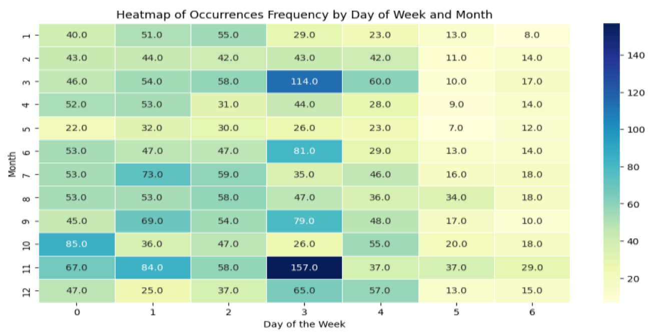

Figure 5. Heatmap of daily occurrences by weekday and month, highlighting operational cycles.

To capture temporal dependencies and inform parameter selection, we analyzed the ACF and PACF of the daily series. The ACF showed correlations with lagged values, while the PACF isolated direct lag effects. Figures 6 and 7 illustrate these functions for a representative series, revealing recurrent patterns essential for specifying ARIMA and SARIMA models.

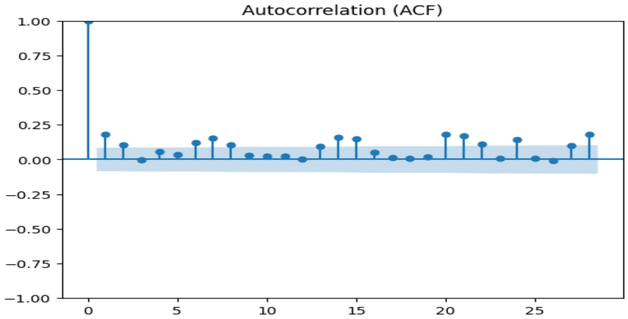

Figure 6. Autocorrelation Function (ACF) of a representative daily time series, evidencing lag dependencies.

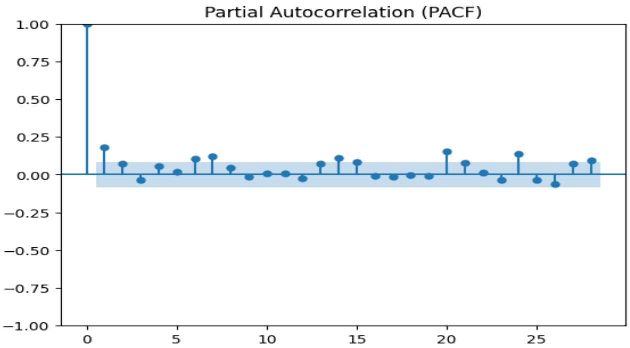

Figure 7. Partial Autocorrelation Function (PACF) of the same series, isolating significant lag effects.

The use of both granularities was crucial to assess model adaptability. It enabled evaluation of how data structure impacts forecasting accuracy and provided evidence to support parameterization and real-world deployment. Figure 8 compares observed and predicted values for one monitored risk over a 90-day forecast window, showing the relative performance of the three models.

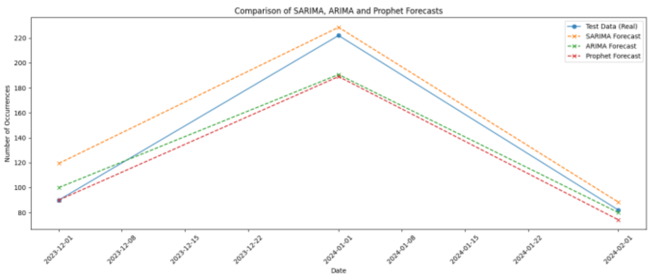

Figure 8. Forecast comparison over a 90-day prediction window, showing observed versus predicted values for ARIMA, SARIMA, and Prophet.

In summary, the multi-granular dataset allowed a comprehensive evaluation of each model's strengths and limitations. This preparation increased experimental robustness and established the empirical foundation for validating forecasting algorithms in real-world institutional risk management. The following sections describe the modeling methods, calibration, validation, and comparative procedures.

METHODOLOGY

The predictive modeling was conducted with a focus on forecasting operational risks in healthcare quality management, using EPA Platform records. Each monitored risk (risk ID) was represented by an independent time series, aggregated at both monthly and daily levels from 2021 to 2024. Three classes of models were evaluated: ARIMA, SARIMA, and Prophet [7].

ARIMA Modeling Approach

The definition of the ARIMA model parameters — autoregressive (p), differencing (d), and moving average (q) — was guided by classical methods for analyzing temporal dependence [16]. For each monthly time series, the ACF and PACF functions were computed in order to identify lag structures and linear dependencies. Formally, for a time series yt, the autocorrelation of order k is given by:

$$\rho k = \frac{\sum_{t=1}^{T-k} (yt - \hat{y})(yt+k - \hat{y})}{\sum_{t=1}^{T} (yt - \hat{y})^{2}}$$

The ACF analysis was used to suggest the appropriate number of moving average terms (q), while the PACF helped determine the number of autoregressive terms (p) by removing the influence of intermediate lags. The PACF was computed based on the Yule-Walker equations. Significant peaks in the ACF indicated relevant lags for the MA components, whereas peaks in the PACF pointed to the AR components.

To assess the stationarity of the time series, the Augmented Dickey-Fuller (ADF) and Kwiatkowski–Phillips–Schmidt–Shin (KPSS) tests were applied. The differencing parameter d was defined based on the results of these tests, as well as on visual inspection of the series. Approximately 45.5% of the series were found to be stationary, requiring no differencing ($d = 0$), while the remaining series required first-order differencing ($d = 1$). The average length of the series was 21 time points.

The final selection of parameters was refined using the Akaike Information Criterion (AIC), and the models were fitted individually for each risk ID. An automated scanning algorithm was employed to identify statistically significant lags, using $2/\sqrt{n}$ as the threshold for significance. Model validation was performed using rolling cross-validation, enabling the assessment of predictive robustness across successive time windows. The model's predictive performance was assessed using the Mean Absolute Error (MAE) metric, yielding an MAE of 10.3 for daily forecasts and 9.94 for direct monthly forecasts.

SARIMA Extension

To capture seasonal patterns, the SARIMA model extension was employed, incorporating the seasonal hyperparameters (P, D, Q) and the length of the seasonal cycle (*m*). For monthly time series, $m=12$ (annual cycle); for daily time series, $m=7$ (weekly cycle).

The seasonal parameters were optimized using a grid search procedure that explored different combinations of (*p, d, q, P, D, Q, m*), evaluated according to performance metrics such as MAE, RMSE, and AIC. This step aimed to validate and refine the parameters initially suggested by the ACF, PACF, and stationarity tests, ensuring that the model was well-adapted to the specific characteristics of each time series [17].

An example of the final configuration for one of the tested models is as follows:

- Non-seasonal order: (2, 2, 1)
- Seasonal order: (0,1,0,7)

Prophet Model

As a modern and flexible alternative, the Prophet model, developed by Meta, was also evaluated. This model is based on the additive decomposition of time series components: trend, seasonality, and holidays. In this study, the configuration included activation of weekly seasonality and deactivation of monthly seasonality, in accordance with the temporal characteristics of the series under analysis. National holidays were incorporated as relevant external regressors [18].

The hyperparameter *changepoint_prior_scale* was adjusted from its default value of 0.05 to 0.1 to allow greater flexibility in adapting to abrupt changes in trend. The model's predictive performance was assessed using the MAE metric, yielding an MAE of 2.48 for daily forecasts and 3.0 for direct monthly forecasts for a specific test case—superior performance to ARIMA in cases involving temporal discontinuities [19].

Figure 9 presents a comparison between the predicted values generated by the Prophet model and the actual observed values for one of the monitored risks, including the estimated confidence interval.

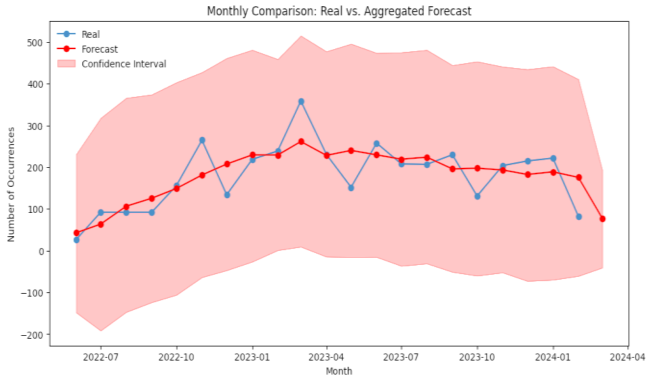

Figure 9. Prophet forecast for a monitored risk at monthly granularity, including actual values and 95% confidence intervals (2022–2024).

In addition to core evaluation metrics, 95% confidence intervals were estimated to quantify uncertainty, and significance tests were applied when relevant. This ensured results were statistically supported rather than relying solely on point estimates.

To enable continuous integration and delivery (CI/CD), an automated MLOps pipeline was developed. It includes: (i) automated training and hyperparameter tuning; (ii) validation with cross-validation and statistical tests; (iii) monthly scheduled deployment; and (iv) continuous monitoring for drift detection. When drift occurs, retraining is triggered, ensuring forecasts adapt to evolving risk patterns. This pipeline supports scalable, reproducible, and auditable deployment across healthcare organizations.

RESULTS

The evaluation was conducted in a realistic computational setting with a multi-tenant architecture comprising multiple databases. Each database contained independent time series for different risk types. Unlike academic datasets with long and balanced series, this study faced challenges typical of production environments: short windows (21 points on average), heterogeneous structures, and the need for scalable, maintainable solutions [20].

Quantitatively, Prophet achieved an average MAE of 3.98 for daily forecasts, below ARIMA's 3.49. SARIMA, though sensitive to parameter tuning, delivered robust results for well-defined seasonal patterns, with an average MAE of 2.24 for daily forecasts aggregated into monthly results.

Beyond MAE, RMSE and MAPE were also analyzed in representative cases, reinforcing the relative performance of models. Although MAE was the primary metric for its interpretability and robustness in short, heterogeneous series, complementary measures confirmed result consistency.

ARIMA performed satisfactorily for stationary or differenced monthly series, with lower residual variance, but was less responsive to nonlinear cycles. Model choice therefore depended on data structure, seasonality, and forecasting horizon. The technical architecture of the solution (Figure 10) includes the following components:

- Back-end in Laravel, connecting AI models, database, and interface;
- Front-end in React/Next.js, offering dashboards for visualization, predictions, and alerts;
- Automated MLOps pipeline, retraining and updating models monthly;
- Relational database, storing time series, forecasts, and historical risk records.

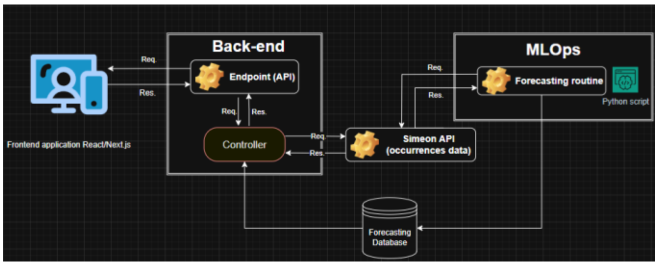

Figure 10. Technical architecture of the forecasting solution, including back-end, front-end, relational database, and automated MLOps pipeline.

Figure 11 shows the dashboard interface in production, displaying forecasts, alerts, and historical analyses segmented by unit and risk type. Validation with Simeon's team and client institutions confirmed several benefits: operational dashboard, faster incident response, higher accuracy in identifying causes and corrective actions, full automation of updates, and achievement of Technology Readiness Level 4 (TRL 4).

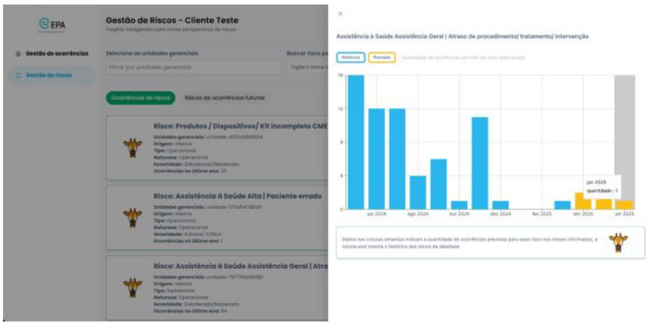

Figure 11. EPA Platform dashboard interface, showing integrated risk forecasts, alert levels, and historical analyses segmented by unit and risk type.

A comparative analysis with related studies (Table 1) contextualized the outcomes in the EPA Platform relative to other healthcare forecasting applications. Results confirmed the trade-offs between statistical accuracy and operational viability: SARIMA excelled with strong seasonal data but required complex tuning; Prophet offered simpler configuration and scalability for multi-tenant use.

Table 1. Comparison between the present study and related works.

Study	Models	Granularity	Domain	Metrics	Optimization Methods
Marinho & Mesquita (2025)	ARIMA, SARIMA, Prophet	Daily, Monthly	Hospital risks	MAE: 2.24 (SARIMA), 3.49 (ARIMA), 3.98 (Prophet)	Grid Search, ACF/PACF, ADF/KPSS
Silva & Galvão (2024)	ARIMA, SARIMA	Monthly	Tuberculosis	MAPE: 8.0% (with COVID-19), 4.5% (without)	AIC, BIC, ADF/KPSS
Weber et al. (2024)	ARIMA, Prophet	Hourly	Emergency KPIs	RMSE: 18.11 (ARIMA), 21.50 (Prophet), 17.71 (GRNN)	AIC/BIC (ARIMA), internal tuning (Prophet), supervised training (GRNN)

Beyond predictive accuracy, a key contribution of this work was the automation of the end-to-end forecasting process via MLOps, enabling monthly model updates and scalable deployment. This ensures that the solution remains operational and continuously aligned with evolving data patterns, making it a practical and sustainable tool for healthcare institutions seeking to improve risk governance through data-driven strategies.

From a scalability perspective, the solution was designed to handle multi-hospital datasets by leveraging the modular MLOps pipeline, which allows parallel training and forecasting across independent time series. The orchestration of models ensures that new risk categories and institutions can be incorporated with minimal manual intervention. Performance tests showed that the system scales efficiently for dozens of concurrent series, but memory demands from SARIMA and the overhead of drift monitoring remain critical bottlenecks. These findings highlight the need for distributed execution and proactive resource allocation in real-world multi-tenant environments.

LIMITATIONS AND ERROR ANALYSIS

Despite the promising results, several limitations should be acknowledged. The time series available for training were relatively short, averaging only 21 points. This restricted length reduces statistical power and increases the risk of overfitting, especially in models with complex seasonal structures.

The models were also sensitive to atypical events, such as the COVID-19 pandemic, which caused abrupt discontinuities that standard parametric approaches (ARIMA/SARIMA) cannot easily capture. Likewise, isolated outliers in daily records had disproportionate effects on error measures and occasionally distorted short-term forecasts.

Finally, although the MLOps pipeline enables monthly retraining, long-term stability is still challenged by data drift. Changes in reporting practices, organizational processes, or external factors may gradually degrade predictive accuracy. Sustaining model reliability in real-world healthcare contexts therefore requires proactive monitoring, anomaly detection, and periodic recalibration.

CONCLUSION

This study presented the development, implementation, and validation of a predictive solution integrated into the Quality Policy Module of Simeon's EPA Platform. By applying ARIMA, SARIMA, and Prophet models to real institutional data, the system demonstrated the feasibility of forecasting operational risks in healthcare environments. Comparative experiments showed that SARIMA achieved the highest accuracy in datasets with clear seasonal patterns, while Prophet offered greater simplicity and adaptability, making it suitable for multi-tenant applications. ARIMA also performed well for stationary or differenced monthly series, underscoring the importance of aligning model choice with the temporal structure of each case.

The solution was validated in a near-operational environment with real client data, integrating forecasting models into an interactive dashboard and an automated MLOps pipeline. Reported benefits included reduced incident response times, improved accuracy in identifying causes and corrective actions, and full automation of the forecasting cycle with monthly updates. The platform successfully reached Technology Readiness Level (TRL) 4, confirming its feasibility for healthcare risk management.

Future research will focus on extending the framework with advanced deep learning architectures such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks, capable of modeling long-range dependencies and nonlinear patterns. Privacy-preserving federated learning approaches will be explored to allow collaborative training across institutions without exposing sensitive data. Dashboards will also be enhanced with real-time monitoring and explainable AI techniques (e.g., SHAP, LIME), reinforcing transparency, user trust, and interpretability. Together, these directions aim to consolidate the EPA Platform as a scalable, secure, and human-centered solution for data-driven healthcare risk management.

Acknowledgments

The authors thank Simeon Estratégia, Desenvolvimento e Processamento de Dados LTDA for its technical collaboration and for validating the proposed solution in a real-world healthcare context, as well as for facilitating access to operational data and integrating the platform for the proof of concept.

This project was jointly funded by the Empresa Brasileira de Pesquisa e Inovação Industrial (EMBRAPII), the Serviço Brasileiro de Apoio às Micro e Pequenas Empresas (SEBRAE), Simeon, and the Centro de Excelência em Inteligência Artificial of the Universidade Federal de Goiás (CEIA/UFG), under the SEBRAE AT Tipo 1 innovation support initiative.

The authors also acknowledge the institutional collaboration of Everso Tecnologia LTDA and Tutoraggio Consultoria em Transformação Digital e Hiperautomação LTDA, which participated in the formal partnership agreement, although not directly involved in the development activities.

Special thanks are extended to the technical and negotiation teams for their contributions to the functional validation, and to the CEIA/UFG research group for their scientific guidance and support in the design, experimentation, and deployment of artificial intelligence models for healthcare risk management.

REFERENCES

[1]. T.-M. Choi, H. K. Chan, and X. Yue, "Recent Development in Big Data Analytics for Business Operations and Risk Management," IEEE Trans. Cybern., vol. 47, no. 1, pp. 81–92, Jan. 2017, doi: 10.1109/TCYB.2015.2507599.

[2]. M. Javaid, A. Haleem, R. P. Singh, R. Suman, and S. Rab, “Significance of machine learning in healthcare: Features, pillars and applications,” Int. J. Intell. Netw., vol. 3, pp. 1–14, 2022, doi: 10.1016/j.ijin.2022.05.002.

[3]. M. Ferrara, G. Bertozzi, N. Di Fazio, I. Aquila, A. Di Fazio, A. Maiese, G. Volonnino, P. Frati, and R. La Russa, “Risk management and patient safety in the artificial intelligence era: A systematic review,” Healthcare (Basel), vol. 12, no. 5, p. 549, Feb. 2024.

[4]. T. Peerawong, C. Chaichulee, and P. Sangsupawanich, “Forecasting daily radiotherapy patient volumes in a tertiary hospital using autoregressive integrated moving average (ARIMA) models,” Cureus, vol. 16, no. 10, p. e72752, Oct. 2024.

[5]. P. Hasan, T. D. Khan, I. Alam, and M. E. Haque, “Dengue in tomorrow: Predictive insights from ARIMA and SARIMA models in Bangladesh: A time series analysis,” Health Sci. Rep., vol. 7, no. 12, p. e70276, Dec. 2024.

[6]. S. Rao, et al., “Targeted-BEHRT: Deep learning for observational causal inference on longitudinal electronic health records,” IEEE Trans. Neural Netw. Learn. Syst., vol. 35, no. 4, pp. 5027–5038, Apr. 2024.

[7]. T. M. Silva and F. T. Galvão, “Tuberculosis incidence in Brazil: Time series analysis between 2001 and 2021 and

projection until 2030," Rev. Saúde Pública, vol. X, no. Y, pp. Z–Z, 2024, doi: 10.1590/1980-549720240027.

[8]. L. Weber, L. A. Lourenço, K. M. Brehm, M. P. Pereira, and C. F. V. Ramos, "Análise de predições de atendimentos na saúde em Florianópolis," Anais do Congresso Brasileiro de Informática em Saúde (CBIS), v. 16, Especial (CBIS-2024), 2024, doi: 10.59681/2175-4411.v16.iEspecial.2024.1338.

[9]. A. C. Raldow, N. Raja, C. W. Villaflores, S. A. Skootsky, E. A. Jaureguy, H. L. Rosenstein, S. D. Meshkat, S. S. Vangala, and C. A. Sarkisian, "Proactive care management of AI-identified at-risk patients decreases preventable admissions," Am. J. Manag. Care, vol. 30, no. 11, pp. 548–554, Nov. 2024.

[10]. Y. M. AlSerkal, N. M. Ibrahim, A. S. Alsereidi, M. Ibrahim, S. Kurakula, S. A. Naqvi, Y. Khan, and N. P. Oottumadathil, "Real-time analytics and AI for managing no-show appointments in primary health care in the United Arab Emirates: Before-and-after study," JMIR Form. Res., vol. 9, p. e64936, Jan. 2025.

[11]. Y. Wang, R. Deng, and X. Geng, "Exploring the integration of medical and preventive chronic disease health management in the context of big data," Front. Public Health, vol. 13, p. 1547392, Apr. 2025.

[12]. G. E. P. Box and G. M. Jenkins, Time Series Analysis: Forecasting and Control. San Francisco, CA, USA: Holden-Day, 1976.

[13]. T. H. McCoy Jr., A. M. Pellegrini, and R. H. Perlis, "Assessment of time-series machine learning methods for forecasting hospital discharge volume," JAMA Netw. Open, vol. 1, no. 7, p. e184087, Nov. 2018.

[14]. R. Zrieq, S. Kamel, S. Boubaker, F. D. Algahtani, M. A. Alzain, F. Alshammari, F. S. Alshammari, B. K. Aldhmadi, S. Atique, M. A. A. Al-Najjar, and S. C. Villareal, "Time-series analysis and healthcare implications of COVID-19 pandemic in Saudi Arabia," Healthcare (Basel), vol. 10, no. 10, p. 1874, Sep. 2022.

[15]. S. Ghalehnovi, A. M. Moghaddam, and S. I. Mohammadpour, "Modelling low temporal, large spatial data of fatal crashes: An application of negative binomial GSARIMAX time series," Accid. Anal. Prev., vol. 198, p. 107958, 2025, doi: 10.1016/j.aap.2025.107958.

[16]. D. A. Dickey and W. A. Fuller, "Distribution of the estimators for autoregressive time series with a unit root," J. Am. Stat. Assoc., vol. 74, no. 366, pp. 427–431, 1979.

[17]. D. Kwiatkowski, P. C. B. Phillips, P. Schmidt, and Y. Shin, "Testing the null hypothesis of stationarity against the alternative of a unit root," J. Econometrics, vol. 54, no. 1-3, pp. 159–178, 1992.

[18]. R. J. Hyndman and G. Athanasopoulos, Forecasting: Principles and Practice, 2nd ed. Melbourne, Australia: OTexts, 2018. [Online]. Available: https://otexts.com/fpp2/.

[19]. S. J. Taylor and B. Letham, "Forecasting at scale," PeerJ Preprints, vol. 5, p. e3190v2, 2017, doi: 10.7287/peerj.preprints.3190v2.

[20]. P. J. Brockwell and R. A. Davis, Introduction to Time Series and Forecasting, 2nd ed. New York, NY, USA: Springer, 2002.