When continuous is not the question

Filippo Ficarra [1]

Abstract

Time series forecasting is a critical tool in medicine, enabling predictions that support clinical decision-making, optimize resource allocation, and improve patient outcomes. However, medical time series data often exhibit significant challenges, including irregular sampling, missing data, and abrupt discontinuities caused by acute health events, treatment changes, or external disruptions. These characteristics make standard forecasting models inadequate, as they frequently assume continuous and uniformly sampled data. This paper addresses the challenge of jump discontinuities in medical time series forecasting by extending transparent the forecasting framework (Kacprzyk et al., 2024) to incorporate discontinuous curves. Our approach models abrupt changes without relying solely on imputation, thereby preserving clinically meaningful patterns and enhancing model interpretability.

FilippoFicarra/discontinuos_timeview

1. Introduction

Time series forecasting has become a cornerstone in various scientific and industrial applications, with particular significance in healthcare where jump discontinuities frequently occur in patient data. While traditional forecasting enables predictions across domains like finance (Kaastra & Boyd, 1996; Kumar Dubey et al., 2021), meteorology (Craddock, 1965; Ray et al., 2021; Faisal et al., 2022), and medicine (Juang et al., 2017; McCoy et al., 2018; Kaushik et al., 2020), the presence of sudden, finite changes in medical data presents unique challenges that conventional models struggle to address effectively (Aminikhanghahi & Cook, 2017; Ebrahimzadeh et al., 2019; Ibrahim et al., 2020; Schaffer et al., 2021; Popovic et al., 2024). Jump discontinuities are especially prevalent in medical time series due to several intrinsic factors. Acute medical events introduce sharp, instantaneous changes in physiological measurements. Additionally, the initiation or termination of treatments can cause immediate shifts in monitored parameters. While other types of discontinuities exist in medical data, such as missing appointments or gradual treatment responses, these sudden jumps represent critical information that most traditional forecasting methods fail to capture effectively (Roggero, 2012). Despite advances in machine learning and statistical modeling, existing approaches often focus on smoothing or interpolating across jump discontinuities, treating them as anomalies rather than significant clinical events. While various imputation methods can provide continuous approximations, they risk masking these clinically crucial sudden changes, potentially leading to misinterpretation of patient trajectories. Furthermore, the lack of transparency in handling jump discontinuities raises concerns about the interpretability and reliability of such forecasts in critical care scenarios. To address this gap, we propose extending transparent time series forecasting tools to explicitly model and predict jump discontinuities while acknowledging other types of data irregularities. By integrating discontinuous curves into forecasting models and their visualizations, with particular attention to sudden finite changes, we aim to capture these abrupt transitions without compromising interpretability. This approach enhances both the accuracy of predictions and provides clinicians with clear insights into significant clinical events. In medicine, properly addressing jump discontinuities is crucial for several reasons. First, it ensures that critical transitions, such as immediate responses to interventions or sudden onset of complications, are accurately captured in predictive models. Second, it enables the development of robust systems that can distinguish between different types of discontinuities, from sudden jumps to gradual changes. Finally, it promotes transparency in machine learning by explicitly representing these significant events, fostering trust in clinical practice.

This paper focuses on extending the transparent forecasting framework TIMEVIEW (Kacprzyk et al., 2024) to account for jump discontinuities in time series data.

2. Background

Time series analysis traditionally assumes underlying continuity in data. A function f is considered continuous ($f \in \mathcal{C}^0$) if it has no abrupt jumps or breaks. Jump dis-

[1]ETH Zürich, Zürich, Switzerland. Correspondence to: Filippo Ficarra <fficarra@ethz.ch>.

continuities represent the most dramatic violation of this continuity, occurring when a function experiences a sudden, finite change in value. These discontinuities are particularly significant in medical contexts, as they often signal crucial clinical events or treatment responses.

Medical time series exhibit various types of discontinuities, with **jump discontinuities** being the primary focus of our work, characterized by sudden, finite changes in value.

Other notable discontinuities are:

- **Derivative Discontinuities**: Sudden changes in rate while maintaining $\mathcal{C}^0$ continuity
- **Curvature Discontinuities**: Abrupt but smooth changes in trend
- **Removable Discontinuities**: Isolated deviations from patterns

While the TIMEVIEW framework (Kacprzyk et al., 2024) offers promising advances in transparent forecasting, its assumption of underlying continuity limits its ability to handle these crucial sudden changes.

This paper builds upon these foundations by developing a methodology that explicitly models jump discontinuities.

3. Proposed Method

While Kacprzyk et al. propose a methodology predicated on the assumption of continuous curves—employing cubic B-splines (degree $= 3$) with a predetermined knot sequence for time series approximation—our framework extends to the more general case of potentially discontinuous curves. This generalization addresses two fundamental limitations in existing approaches: the implicit continuity assumption and the requirement for a priori knot placement.

3.1. Framework

First, we develop a robust methodology for detecting temporal discontinuities in the input time series $\mathbf{x} \in \mathbb{R}^M$. Let $\mathbf{y} = \{y_t\}_{t=1}^T$ be a time series where y_t represents the observation at time t. We define a discontinuity point τ_i as a time-stamp where the curve exhibits a significant structural break, characterized by:

$$\lim_{h \to 0^+} |y_{\tau_i+h} - y_{\tau_i-h}| > \delta$$

where δ is a threshold determined through statistical analysis of the time series' local variation. We define the set of discontinuity points $\mathcal{T} = \{\tau_1, \dots, \tau_k\}$, effectively partitioning the time domain into $k+1$ continuous segments.

Second, we introduce a novel curve reconstruction framework that preserves detected discontinuities while maintaining optimal smoothness within continuous segments. For each segment $[\tau_i, \tau_{i+1}]$, we construct a local spline approximation $s_i(t)$ such that:

$$s_i(t) = \sum_{j=1}^{m_i} c_{ij} B_{j,p}(t), \quad t \in [\tau_i, \tau_{i+1}]$$

where $B_{j,p}(t)$ are B-spline basis functions of degree p, m_i is the umber of basis functions and c_{ij} are control points optimized to minimize:

$$\mathcal{L} = \sum_{t \in [\tau_i, \tau_{i+1}]} \|y_t - s_i(t)\|^2 + \lambda \mathcal{L}_{L2}(s_i)$$

This formulation ensures optimal smoothness within segments while allowing for discontinuities at the detected break points τ_i. Unlike previous approaches that enforce global continuity constraints, our method adapts to the inherent structure of the data, providing a more faithful representation of time series exhibiting abrupt changes or regime shifts.

The resulting piecewise-continuous approximation $s(t)$ is defined as:

$$s(t) = s_i(t) \quad \text{for} \quad t \in [\tau_i, \tau_{i+1}]$$

This representation maintains high fidelity to the original time series while explicitly accounting for structural breaks.

3.2. Predict discontinuities

Let $\mathbf{h} : \mathbb{R}^M \to \mathbb{R}^T$ represent the encoder function that maps an input feature vector to a vector of probabilities indicating the likelihood of discontinuities at each timestamp. Specifically:

- $\boldsymbol{d} = \mathbf{h}(\mathbf{x})$: the vector of predicted probabilities for discontinuities at each timestamp.
- $\boldsymbol{\chi} \in \mathbb{R}^N$: the ground truth binary vector indicating whether each timestamp is a discontinuity.

$$\mathcal{L}_{\text{BCE}} = -\frac{1}{T} \sum_{t=1}^{T} \left[\chi_t \log(d_t) + (1-\chi_t) \log(1-d_t)\right].$$

The encoder $\mathbf{h}$ has been implemented as an MLP with 3 layers.

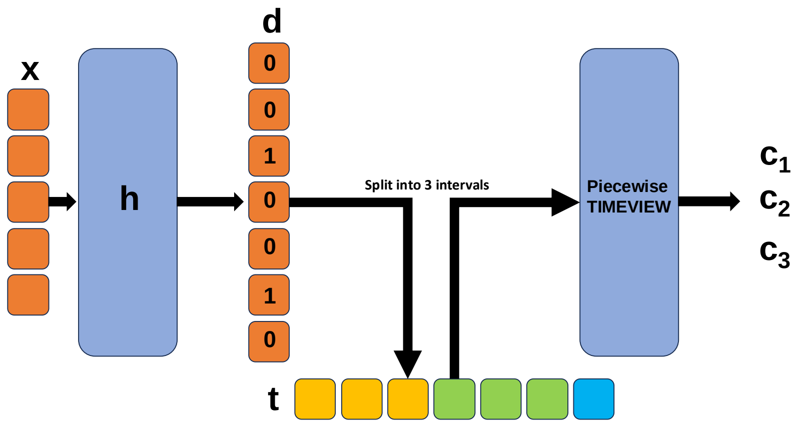

Figure 1. Augmented TIMEVIEW architecture with jump discontinuity prediction and piecewise reconstruction.

3.3. Approximate the intervals

Each segment is approximated using a B-spline as in TIMEVIEW. For each interval, we compute the control points $\mathbf{c}$ that weight the basis functions. **Piecewise TIMEVIEW** takes as input the feature vector $\mathbf{x}$ and the number of discontinuities. We set a the number of knots (equally spaced) for the whole time interval when 0 discontinuities are present. Let f be the minimum number of knots for each interval and K_{max} be the maximum amount of discontinuities that we hypothesize our function to have. Let $f = 4$ and $K_{max} = 4$, it follows that $a = 16$ such that if we have the maximum number of knots we can fit the minimum number of knots inside each interval. Then we model **Piecewise TIMEVIEW** simply as an encoder that outputs the concatenated control points for each interval.

4. Dataset Creation

To evaluate the proposed methodology for handling discontinuities in time series forecasting, we generated a synthetic dataset designed to mimic the characteristics of real-world medical data. This section provides an explanation of the data generation process, highlighting key steps and underlying principles.

4.1. Synthetic functions

The dataset was created by simulating time series data with both smooth trends and abrupt discontinuities. The steps for generating each sample are as follows:

1. **Input Features:** Each time series sample is associated with a feature vector $\mathbf{x} \in \mathbb{R}^N$, where N is the number of features. These features are sampled from a uniform distribution:

$$\mathbf{x} \in \mathbb{R}^N, \ \mathbf{x} \sim \text{Uniform}(-1, 1).$$

2. **Temporal Component:** A uniformly spaced time grid $\mathbf{t} \in [0, T]$ with M points is defined for each sample.

3. **Smooth Trend Modeling:** The smooth component of the time series is generated using sinusoidal and cosine functions:

$$g(t, \mathbf{x}) = \sin(\|\mathbf{x}\|_2 \cdot t), \quad h(\mathbf{x}) = \cos\left(\frac{\|\mathbf{x}\|_2}{\sqrt{N}}\right).$$

4. **Discontinuity Introduction:** Abrupt changes are introduced at specific time points determined by a discontinuity detection function $T_k(\mathbf{x}, k)$. For a time series with K discontinuities, the time of each discontinuity is defined as:

$$T_k(\mathbf{x}) = k \cdot \frac{\|\mathbf{x}\|_2}{\sqrt{N}} + 2k \cdot \max(\mathbf{x}).$$

At each discontinuity point, a sharp change in value is introduced:

$$\Delta_k(\mathbf{x}) = \frac{1}{k} \exp\left(\frac{\|\mathbf{x}\|_1}{N}\right)$$

where k is the k-th discontinuity.

5. **Time Series Construction:** The final time series $y(t)$ is computed as the combination of the smooth components and the introduced discontinuities:

$$y(t) = g(t, \mathbf{x}) + h(\mathbf{x}) + \sum_{k=1}^{K} \mathbf{1}_{[t \geq T_k(\mathbf{x})]} \Delta_k(\mathbf{x}),$$

where $\mathbf{1}$ is an indicator function ensuring that each discontinuity affects only subsequent time points.

Despite setting K as an hyperparameter that describes the maximum number of possible discontinuities in our function, the maximum possible number of discontinuities k is intrisically dependent on $\mathbf{x}$ given our definition of $T_k(\mathbf{x})$:

$$k < \frac{T}{\frac{\|\mathbf{x}\|^2}{N} + 2\max(\mathbf{x})}. \tag{1}$$

4.2. Dataset Characteristics

The generated dataset includes 100000 —divided into 80/10/10 split—samples, each with 100 time points. The number of discontinuity k per sample is randomly selected from $[0, K_{\text{max}}]$, where $K_{\text{max}} = 4$, under the constraint underlined in Equation (1). Each time series is characterized by its smooth trends and abrupt changes.

4.3. Benefits of Synthetic Data

The synthetic dataset offers the following advantages:

- **Controlled Complexity:** The possibility of creating a synthetic dataset allows precise control over the characteristics of smooth trends and discontinuities.
- **Realism:** The data captures essential features of real-world medical time series, including irregular patterns and abrupt transitions.
- **Reproducibility:** The parameterized generation process ensures that the dataset can be replicated for benchmarking and comparison across methodologies.

5. Results

5.1. Predict discontinuity

Predicting the discontinuity is a pretty simple task for the synthetic dataset that I have created. The MLP was trained for 10 epochs and $lr = 1e-3$ and batch_size = 32.

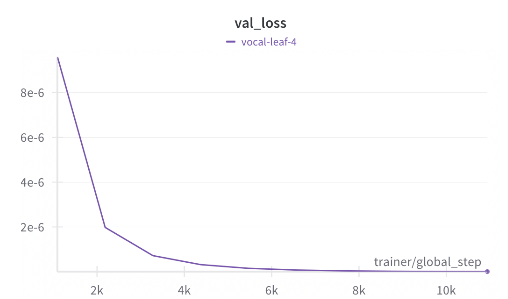

Figure 2. Binary cross-entropy loss for the discontinuity times prediction.

5.2. Approximate the intervals

The code is not yet been developed, hence there is no result to show.

6. Conclusions

The framework approximates functions without making any assumptions about the presence of jump discontinuities. Its adaptive knot selection ensures high resolution for single intervals and maintains consistent resolution across intervals when discontinuities are present. In conclusion, by extending TIMEVIEW, the **Piecewise TIMEVIEW** offers a transparent model that not only clarifies the influence of input features on function shapes but also reveals unexpected behaviors.

7. Future Work

We aim to enhance visualization tools to clearly represent discontinuities, providing healthcare professionals with actionable insights. By directly addressing the issue of discontinuities, our work bridges a crucial gap in time series forecasting for healthcare, opening the way for more robust and transparent predictive systems that align with the complexities of real-world medical data. We further aim to improve this approach by including other discontinuity types. Finally future work would use real world data instead of synthetic data.

8. Limitations

I acknowledge that the set of experiments is poor and would be more complete and rich in a normal research environment of several months of work.

References

Aminikhanghahi, S. and Cook, D. J. A survey of methods for time series change point detection. *Knowledge and Information Systems*, 51(2):339–367, May 2017. doi: 10.1007/s10115-016-0987-z. URL https://pmc.ncbi.nlm.nih.gov/articles/PMC5464762/.

Craddock, J. M. The Analysis of Meteorological Time Series for Use in Forecasting. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 15(2):167–190, 1965. doi: 10.2307/2987390. URL http://www.jstor.org/stable/2987390. Publisher: [Royal Statistical Society, Wiley].

Ebrahimzadeh, Z., Zheng, M., Karakas, S., and Kleinberg, S. Deep learning for multi-scale changepoint detection in multivariate time series. 2019. URL https://arxiv.org/abs/1905.06913.

Faisal, A. F., Rahman, A., Habib, M. T. M., Siddique, A. H., Hasan, M., and Khan, M. M. Neural networks based multivariate time series forecasting of solar radiation using meteorological data of different cities of bangladesh. *Results in Engineering*, 13:100365, 2022. doi: https://doi.org/10.1016/j.rineng.2022.100365. URL https://www.sciencedirect.com/science/article/pii/S2590123022000354.

Ibrahim, Z. M., Wu, H., and Dobson, R. J. B. Modeling rare interactions in time series data through qualitative change: Application to outcome prediction in intensive care units. *CoRR*, abs/2004.01431, 2020. URL https://arxiv.org/abs/2004.01431.

Juang, W.-C., Huang, S.-J., Huang, F.-D., Cheng, P.-W., and Wann, S.-R. Application of time series analysis in modelling and forecasting emergency department visits in a medical centre in southern taiwan. *BMJ Open*, 7(11), 2017. ISSN 2044-6055. doi: 10.1136/bmjopen-2017-018628. URL https://bmjopen.bmj.com/content/7/11/e018628.

Kaastra, I. and Boyd, M. Designing a neural network for forecasting financial and economic time series. *Neurocomputing*, 10(3):215–236, 1996. doi: https://doi.org/10.1016/0925-2312(95)00039-9. URL https://www.sciencedirect.com/science/article/pii/0925231295000399.

Kacprzyk, K., Liu, T., and van der Schaar, M. Towards transparent time series forecasting. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=TYXtXLYHpR.

Kaushik, S., Choudhury, A., Sheron, P. K., Dasgupta, N., Natarajan, S., Pickett, L. A., and Dutt, V. Ai in healthcare: Time-series forecasting using statistical, neural, and ensemble architectures. *Frontiers in Big Data*, 3, 2020. doi: 10.3389/fdata.2020.00004. URL https://www.frontiersin.org/journals/big-data/articles/10.3389/fdata.2020.00004.

Kumar Dubey, A., Kumar, A., García-Díaz, V., Kumar Sharma, A., and Kanhaiya, K. Study and analysis of sarima and lstm in forecasting time series data. *Sustainable Energy Technologies and Assessments*, 47:101474, 2021. doi: https://doi.org/10.1016/j.seta.2021.101474. URL https://www.sciencedirect.com/science/article/pii/S2213138821004847.

McCoy, Thomas H., J., Pellegrini, A. M., and Perlis, R. H. Assessment of time-series machine learning methods for forecasting hospital discharge volume. *JAMA Network Open*, 1(7):e184087–e184087, 11 2018. doi: 10.1001/jamanetworkopen.2018.4087. URL https://doi.org/10.1001/jamanetworkopen.2018.4087.

Popovic, M., Zugna, D., Tilling, K., and Richiardi, L. Regression discontinuity design for the study of health effects of exposures acting early in life. *Frontiers in Public Health*, 12:1377456, 2024. doi: 10.3389/fpubh.2024.1377456. URL https://pmc.ncbi.nlm.nih.gov/articles/PMC11066219/.

Ray, S., Das, S. S., Mishra, P., and Al Khatib, A. M. G. Time Series SARIMA Modelling and Forecasting of Monthly Rainfall and Temperature in the South Asian Countries. *Earth Systems and Environment*, 5(3):531–546, September 2021. ISSN 2509-9434. doi: 10.1007/s41748-021-00205-w. URL https://doi.org/10.1007/s41748-021-00205-w.

Roggero, M. Discontinuity detection and removal from data time series. In *VII Hotine-Marussi Symposium on Mathematical Geodesy*, pp. 135–140, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg. URL https://iris.polito.it/handle/11583/2379396.

Schaffer, A. L., Dobbins, T. A., and Pearson, S.-A. Interrupted time series analysis using autoregressive integrated moving average (ARIMA) models: a guide for evaluating large-scale health interventions. *BMC Medical Research Methodology*, 21(1):58, March 2021. ISSN 1471-2288. doi: 10.1186/s12874-021-01235-8. URL https://doi.org/10.1186/s12874-021-01235-8.