*Exploring the Universe with Artificial Intelligence (UniversAI)*
*Proceedings IAU Symposium No. 397, 2026*
*I. Liodakis & V. Efthymiou, eds.*
doi:10.1017/S1743921325104730

# Probabilistic Interpolation of Sagittarius A*'s Multi-Wavelength Light Curves Using Machine Learning

**G. Sasseville[1,2,3], J. Hlavacek-Larrondo[1,2,3], D. Haggard[2,3,4,5], A. Adam[1,2,6], H. Paugnat[7] and G. Witzel[8]**

[1]Department of Physics, University of Montreal, Montreal, QC H2V 0B3, Canada
email: gabriel.sasseville@umontreal.ca

[2]Ciela - Montreal Institute for Astrophysical Data Analysis and Machine Learning, Montreal, QC H2V 0B3, Canada

[3]Centre de recherche en astrophysique du Québec (CRAQ), Canada

[4]Department of Physics, McGill University, 3600 rue University, Montréal, QC H3A2T8, Canada

[5]Trottier Space Institute at McGill, 3550 rue University, Montréal, QC H3A2A7, Canada

[6]Mila - Quebec Artificial Intelligence Institute, Montreal, Canada

[7]Department of Physics and Astronomy, UCLA, Los Angeles, CA 90095-1547, USA

[8]Max Planck Institute for Radio Astronomy, Bonn 53121, Germany

**Abstract.** Understanding the variability of Sagittarius A* (Sgr A*) requires coordinated, multi-wavelength observations. However, these observations suffer from irregular sampling, noise, and large data gaps, limiting our ability to identify flares and infer cross-wavelength time lags. We address this challenge using machine learning models designed for sparse, multivariate time series. In particular, we introduce a diffusion-based generative model, representing the first application of score-based diffusion models to astronomical time series. We also use a transformer-based model, and benchmark both against a multi-output Gaussian Process. The models are trained on simulated light curves that replicate the cadence, noise, and cross-band correlations of real Sgr A* data, including observations from ALMA, Spitzer, GRAVITY, and Chandra. Our results show that the diffusion model outperforms other methods in accuracy and uncertainty calibration, on both synthetic and real data. This work highlights the potential of generative models for reconstructing astrophysical variability from incomplete observations.

**Keywords.** Interdisciplinary astronomy, Machine learning, Black holes, Sagittarius A*

## 1. Introduction

Understanding the variability of Sagittarius A* (Sgr A*), the supermassive black hole (BH) at the center of our galaxy, the Milky Way, is crucial for probing accretion physics near the event horizon. Multi-wavelength light curves, spanning X-ray (2-8 keV), near-infrared (NIR; 2.2 $\mu$m), infrared (IR; 4.5 $\mu$m), and submillimeter (submm; 340 GHz) wavelengths, reveal its stochastic variability, composed of rapid flares superimposed on longer-term trends (e.g., Boyce et al. 2022; Witzel et al. 2021). However, these crucial light curves suffer from irregular sampling, significant data gaps, and wavelength-dependent noise due to observational constraints. These limitations severely hinder precise analyses, such as robust flare identification and accurate measurement of cross-band time lags.

© The Author(s), 2026. Published by Cambridge University Press on behalf of The International Astronomical Union. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S1743921325104730 Published online by Cambridge University Press

**Table 1.** Model Comparison on Simulated Light Curves: MSE and CRPS by Wavelength.

| Model | X-ray | | NIR | | IR | | Submm | | Total Average | |
|---|---|---|---|---|---|---|---|---|---|---|
| | MSE | CRPS | MSE | CRPS | MSE | CRPS | MSE | CRPS | MSE | CRPS |
| MOGP | 0.169 | 0.234 | 0.116 | 0.207 | 0.091 | 0.195 | 0.690 | 0.439 | 0.266 | 0.269 |
| Tripletformer | 0.218 | 0.230 | 0.081 | 0.126 | 0.060 | 0.107 | 0.263 | 0.284 | 0.155 | 0.186 |
| CSPD | 0.011 | 0.021 | 0.026 | 0.112 | 0.026 | 0.101 | 0.186 | 0.320 | **0.062** | **0.138** |

Addressing these challenges requires robust methods for probabilistic data interpolation, which can reconstruct missing information while quantifying uncertainty. While Gaussian Processes (GPs) offer probabilistic outputs and have been applied to Sgr A* light curves (e.g., Paugnat et al. 2024), they scale poorly with large, multivariate datasets (e.g., Álvarez and Lawrence 2011). Recent deep learning approaches, such as Transformers, show promise for time series but often lack calibrated uncertainty estimates essential for inference (e.g., Boersma et al. 2024).

In this work, we use a continuous stochastic process diffusion (CSPD) model for interpolating sparse, multi-wavelength Sgr A* time series, marking the first application of score-based models to astronomical time series (e.g., Biloš et al. 2023). We compare CSPD against two established approaches that serve as robust probabilistic benchmarks: a multi-output GP (MOGP) and the Tripletformer, a transformer-based model designed to interpolate time series with uncertainty estimates (e.g., Yalavarthi et al. 2023). Our models are trained on realistic simulated Sgr A* light curves that capture observational characteristics and multi-wavelength correlations. The following sections summarize our findings, evaluating their performance on both synthetic and real Sgr A* data.

## 2. Model Performances on Simulated Data

To begin, our models are trained on realistic simulations of Sgr A* multi-wavelength light curves, designed to faithfully mimic the statistical variability, sampling characteristics, and noise properties of real Sgr A* observations. We then evaluate their performance on a reserved test set, assessing how well CSPD, Tripletformer, and MOGP reconstruct these light curves.

Quantitative performance metrics on the simulated test set (Table 1) demonstrate clear distinctions between the models. For both mean squared error (MSE) and continuous ranked probability score (CRPS), lower values indicate better performance, corresponding to more accurate point predictions and better-calibrated uncertainty estimates, respectively. The CSPD model achieves the lowest overall average MSE (0.062) and CRPS (0.138), significantly outperforming both MOGP and Tripletformer. Its performance is notably strong in the X-ray and NIR bands, where it shows a substantial advantage.

The Tripletformer exhibits competitive performance in the NIR, IR, and submm bands, generally outperforming the MOGP across these modalities. Its relatively low CRPS values suggest reasonable uncertainty calibration for the simulated data. In contrast, the MOGP model consistently performs the poorest, with the highest average MSE (0.266) and CRPS (0.269). Its largest errors are particularly pronounced in the submm band (MSE of 0.690), likely attributed to our choice of a radial basis function kernel, which often produces overly smooth interpolations, and its assumption of homoscedastic noise (e.g., Bonilla et al. 2007).

## 3. Model Performances on Observational Data

To evaluate model generalizability, we tested CSPD, Tripletformer, and MOGP on real Sgr A* observational data from a July 2019 campaign. Importantly, we didn't train

https://doi.org/10.1017/S1743921325104730 Published online by Cambridge University Press

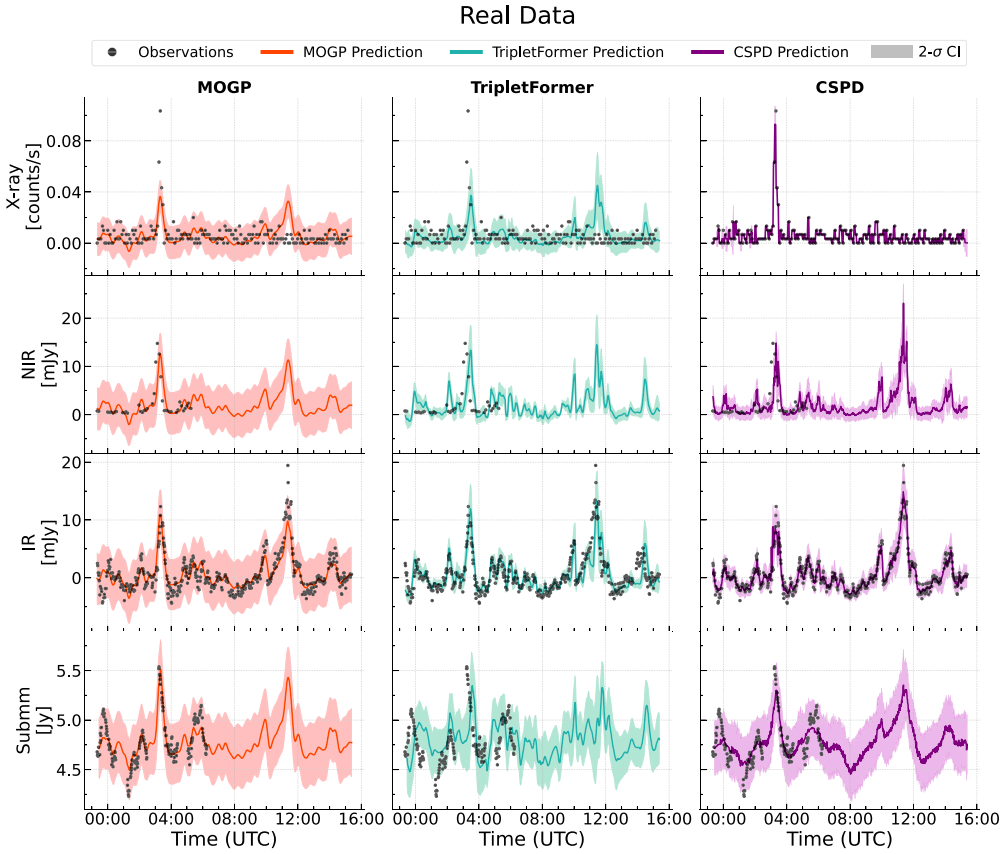



**Figure 1.** Model interpolation results on real observational data of Sgr A*. Each column corresponds to a different model (MOGP, TripletFormer, CSPD), and each row represents a different wavelength band (X-ray, NIR, IR, submm). Observational data upon which the models are conditioned are shown as black circles. Solid lines denote the mean predictions of each model while shaded regions denote $2\sigma$ credible intervals. While all models generally follow the observed data, the CSPD model produces sharper predictions with more selective uncertainty and avoids false-positive flares hallucinated by the other models.

or fine-tune any models on this dataset, ensuring a true test of their ability to transfer knowledge from purely simulated data to real-world astrophysical variability.

As shown in Figure 1, CSPD and Tripletformer generally produce mean predictions that more closely follow the observed data compared to the MOGP, especially for rapid flares in the NIR and IR bands. MOGP's smoother reconstructions align with its radial basis function kernel, as seen in our simulated data results (Section 2).

A critical difference lies in uncertainty estimates. MOGP exhibits very broad credible intervals, reflecting its conservative Gaussian assumptions (e.g., Bonilla et al. 2007). CSPD, conversely, yields sharply defined uncertainty bounds, though they may appear overly narrow in low signal-to-noise regions like the X-ray band, potentially indicating overconfidence. Tripletformer offers intermediate uncertainty.

A notable finding is the CSPD model's ability to avoid hallucinating a false X-ray flare between 10:00 and 13:00 UTC, an event spuriously predicted by both MOGP and Tripletformer despite observational data indicating no such activity. This suggests that CSPD better disentangles inter-band correlations from wavelength-specific trends, likely internalizing the physical understanding that not all NIR flares have X-ray counterparts

https://doi.org/10.1017/S1743921325104730 Published online by Cambridge University Press

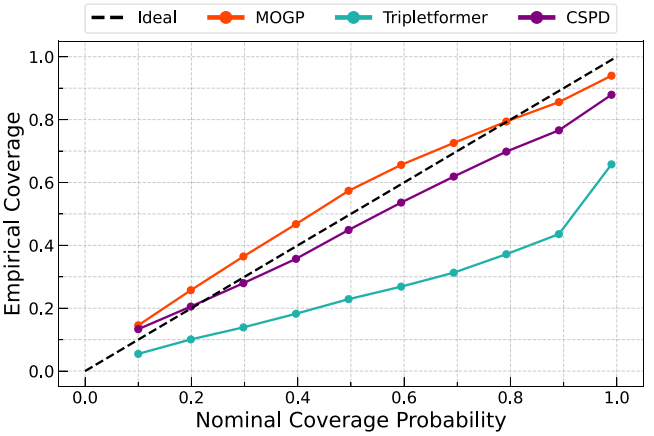



**Figure 2.** Coverage analysis of predictive uncertainties across a range of nominal confidence levels, computed on real observational data. Each curve shows the empirical coverage achieved by a model as a function of the nominal coverage level, aggregated across all wavelength bands. The dashed diagonal line indicates perfect calibration. The MOGP (orange) shows the best overall calibration, closely following the ideal line. CSPD (purple) slightly underestimates uncertainty but remains well-calibrated, meanwhile the TripletFormer (teal) consistently undercovers.

(e.g., Witzel et al. 2021). This indicates CSPD's learning of a more complex generative structure.

While all models capture the ∼3:30 UTC IR flare, there's a slight time delay in their predicted peaks for X-ray, submm, and NIR. This might stem from model limitations in explicitly encoding inter-band time lags or from training solely on a single physical model.

## 4. Uncertainty Calibration

We assessed the calibration of each model's uncertainty estimates using real Sgr A* observational data. This involved masking random subsets of data to compute the empirical coverage, the fraction of true values falling within the predicted credible intervals, across nominal confidence levels from 10% to 99%). A perfectly calibrated model would align with the identity line, where empirical coverage equals nominal coverage.

As Figure 2 shows, the MOGP consistently demonstrates the best overall calibration, closely tracking the ideal line with slightly conservative uncertainty estimates. The CSPD model also performs well, staying just below the ideal line, indicating a modest underestimation of uncertainty despite its flexibility in learning complex noise structures (e.g., Biloš et al. 2023). In contrast, the Tripletformer model exhibits significant and consistent undercoverage, meaning it tends to be overconfident in its predictions. This miscalibration likely stems from its simpler, parametric treatment of uncertainty and architectural limitations (e.g., Yalavarthi et al. 2023).

## 5. Future Work

This work demonstrates the promising capability of CSPD for high-fidelity interpolation and uncertainty quantification in sparse, multi-wavelength Sgr A* light curves. This approach significantly enhances our ability to extract information from challenging astrophysical observations.

Moving forward, key areas for improvement include enhancing generalization to real observational data. This involves incorporating more diverse training data, such as a combination of phenomenological and GRMHD-based simulations, and integrating historical

https://doi.org/10.1017/S1743921325104730 Published online by Cambridge University Press

observational datasets to account for realistic sampling and noise complexities. Exploring methods like learned observational priors (e.g., Adam et al. 2022) and advanced domain adaptation techniques will also be crucial. Ultimately, these advancements will enable more precise constraints on time delays, significantly advancing our understanding of accreting BHs.

## References

Adam, A., Coogan, A., Malkin, N., Legin, R., Perreault-Levasseur, L., Hezaveh, Y., & Bengio, Y. 2022. Posterior samples of source galaxies in strong gravitational lenses with score-based priors. arXiv:2211.03812 [astro-ph].

Álvarez, M. A. & Lawrence, N. D. 2011, Computationally Efficient Convolved Multiple Output Gaussian Processes. *Journal of Machine Learning Research*, 12(41), 1459–1500.

Biloš, M., Rasul, K., Schneider, A., Nevmyvaka, Y., & Günnemann, S. Modeling Temporal Data as Continuous Functions with Stochastic Process Diffusion. In *Proceedings of the 40th International Conference on Machine Learning* 2023, pp. 2452–2470. PMLR. ISSN: 2640–3498.

Boersma, O. M., Ayache, E. H., & Leeuwen, J. v. 2024. Transformer models for astrophysical time series and the GRB prompt-afterglow relation. arXiv:2406.12515.

Bonilla, E. V., Chai, K., & Williams, C. Multi-task Gaussian Process Prediction. In *Advances in Neural Information Processing Systems* 2007, volume 20. Curran Associates, Inc.

Boyce, H., Haggard, D., Witzel, G., von Fellenberg, S., Willner, S. P., Becklin, E. E., Do, T., Eckart, A., Fazio, G. G., Gurwell, M. A., Hora, J. L., Markoff, S., Morris, M. R., Neilsen, J., Nowak, M., Smith, H. A., & Zhang, S. 2022, Multiwavelength Variability of Sagittarius A* in 2019 July. *The Astrophysical Journal*, 931, 7. Publisher: IOP ADS Bibcode: 2022ApJ...931....7B.

Paugnat, H., Do, T., Gautam, A. K., Martinez, G. D., Ghez, A. M., Sakai, S., Weldon, G. C., Hosek, M. W., Haggard, Z., O'Neil, K. K., Becklin, E. E., Witzel, G., Lu, J. R., & Matthews, K. 2024, New Evidence for a Flux-independent Spectral Index of Sgr A* in the Near-infrared. *The Astrophysical Journal*, 977, 228. Publisher: IOP ADS Bibcode: 2024ApJ...977..228P.

Witzel, G., Martinez, G., Willner, S. P., Becklin, E. E., Boyce, H., Do, T., Eckart, A., Fazio, G. G., Ghez, A., Gurwell, M. A., Haggard, D., Herrero-Illana, R., Hora, J. L., Li, Z., Liu, J., Marchili, N., Morris, M. R., Smith, H. A., Subroweit, M., & Zensus, J. A. 2021, Rapid Variability of Sgr A* across the Electromagnetic Spectrum. *The Astrophysical Journal*, 917, 73. Publisher: IOP ADS Bibcode: 2021ApJ...917...73W.

Yalavarthi, V. K., Burchert, J., & Schmidt-Thieme, L. Tripletformer for Probabilistic Interpolation of Irregularly sampled Time Series. In *2023 IEEE International Conference on Big Data (BigData)* 2023, pp. 986–995. IEEE Computer Society.

https://doi.org/10.1017/S1743921325104730 Published online by Cambridge University Press