Proceedings of the 2025 Winter Simulation Conference
E. Azar, A. Djanatliev, A. Harper, C. Kogler, V. Ramamohan, A. Anagnostou, and S. J. E. Taylor, eds.

UNCERTAINTY-AWARE DIGITAL TWIN OF BUSINESS PROCESSES VIA BAYESIAN CALIBRATION AND POSTERIOR-PREDICTIVE SIMULATION

Guilhem Nespoulous[1,2]

[1]Université de Technologie de Troyes, LIST3N research unit, Troyes, FRANCE
[2]QAD Process Intelligence, Courbevoie, FRANCE

ABSTRACT

Event logs are finite, partial views of a latent stochastic process. We present a Bayesian digital twin based on probabilistic, random-effects, event-by-event generators that utilize historical logs and propagate uncertainty. After calibration with Hamiltonian Monte Carlo, each posterior draw is a parameter vector that defines a complete simulator, with which we generate or continue event logs by sequentially sampling the next activity, the inter-event time, and new case arrivals conditional on history and crowding (congestion). Computing KPIs (cycle time, cost, directly-follows counts, etc.) on the simulated logs and aggregating over all posterior draws yields posterior-predictive KPI distributions. Validation compares these distributions to bootstrap baselines from the observed log using distributional distances. The result is a scenario-ready (Aalst 2010) process twin that reports outcomes as distributions, enabling risk-aware decisions.

1 INTRODUCTION

Discrete-event simulation is routinely used to analyze and redesign business processes. However, standard calibrations often ignore epistemic uncertainty stemming from limited logs and from the gap between a chosen simulation structure and the real system. We target an *uncertainty-aware* digital twin: a simulation model calibrated in a Bayesian way so that (i) parameters are inferred with full posterior uncertainty and (ii) decisions are supported by predictive distributions of KPIs rather than point estimates.

Two design principles drive our work: (i) treat the log as one realization among many of an underlying stochastic process and (ii) make the twin *generative at the event level*, so that activities and timestamps are predicted conditional on the complete case history and on instantaneous system state (e.g., crowding), not only on the last event.

2 CALIBRATION AND MODEL EVALUATION

Parameters are calibrated with the NUTS (Hoffman and Gelman 2014) variant of HMC in `Stan` after a warm-start stage that fits some of the parameter distributions via variational inference to obtain data-driven priors. Hyperprior design encodes parsimony to favor simpler effects unless strongly supported by data. We monitor convergence with trace plots, $\widehat{R}$, and effective sample sizes.

Model criticism relies on posterior-predictive checks (PPCs): we draw parameter vectors $\{\vartheta^{(s)}\}$ from the posterior, run event-level simulations, and compare summaries to the observed log. Crucially, we assess *distributional* agreement for KPIs (cycle time, throughput, activity costs, directly-follows counts) via Wasserstein or energy distances between empirical distributions, alongside coverage of observed KPIs by posterior-predictive intervals.

3 PREDICTIVE SIMULATION, KPIs, AND SCENARIOS

Each posterior draw defines a complete simulation parameterization; aggregating across draws yields KPI distributions that integrate parameter and process variability. Scenario analysis modifies simulation parameters (e.g., directly-follows activity probabilities) to represent policies. Because results are distributions, we can trade off expected performance against risk with posterior-predictive intervals. The generation of complete event logs decouples simulation from analysis: KPIs may be selected or revised after the fact without re-executing the simulator. The digital twin is an a posteriori mixture of simulators:

$$\mathbb{Q}^{\text{pred}}(\,\cdot \mid D) = \int \mathbb{P}_{\theta}(\,\cdot\,)\,\Pi(\mathrm{d}\theta \mid D), \qquad \widehat{\mathbb{Q}}^{\text{pred}}(\,\cdot \mid D) = \frac{1}{S}\sum_{s=1}^{S} \mathbb{P}_{\theta^{(s)}}(\,\cdot\,), \quad \theta^{(s)} \sim \Pi(\cdot \mid D).$$

KPIs are computed on simulated logs $L^{(s)} \sim \mathbb{P}_{\theta^{(s)}}$ and aggregated over s to obtain posterior-predictive distributions.

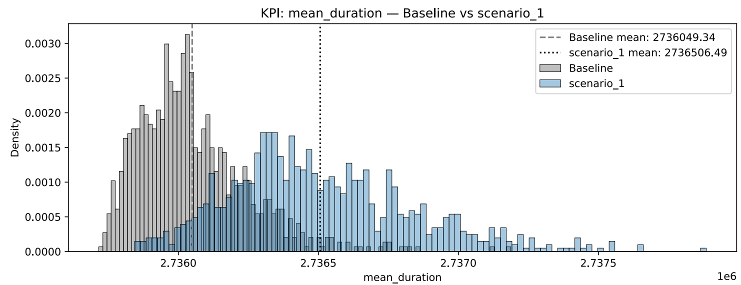

Figure 1: Effect of adding an offset δ to directly-follows transition probabilities on the case mean duration KPI, Helpdesk dataset (Verenich 2016).

4 EVALUATION PROTOCOL

We compare three targets on held-out horizons: (i) observed KPI distributions; (ii) posterior-predictive KPI distributions from the digital twin; and (iii) bootstrap distributions computed from the observed log. Agreement is quantified with distributional distances. This positions the Bayesian twin against a strong, assumption-light baseline.

5 CONCLUSION

The proposed framework turns event logs into an *uncertainty-calibrated* digital twin able to produce full predictive KPI distributions and principled, risk-aware scenario analyses. Ongoing work focuses on richer state representations (resource calendars, priorities), faster inference via amortization for warm starts, and sequential updating for streaming logs.

ACKNOWLEDGMENTS

The author thanks Frédéric Bertrand (Conservatoire national des arts et métiers) and Myriam Maumy (École des hautes études en santé publique), Ph.D. directors, and Yoann Valéro, industrial scientific tutor, for their guidance and feedback. This work was supported by a CIFRE industrial Ph.D. fellowship with QAD Inc.

REFERENCES

Aalst, W. 2010. “Business Process Simulation Revisited”. In *Enterprise and Organizational Modeling and Simulation*, Volume 63, 1–14 https://doi.org/10.1007/978-3-642-15723-3_1.

Hoffman, M. D., and A. Gelman. 2014. “The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo”. *Journal of Machine Learning Research* 15(1):1593–1623.

Verenich, I. 2016. “Helpdesk” https://doi.org/10.17632/39bp3vv62t.1.