The Journal of Open Source Software

Interfere: Studying Intervention Response Prediction in Complex Dynamic Models

D. J. Passey [1], **Alice C. Schwarze** [2], **Zachary M. Boyd** [3], **and Peter J. Mucha** [2]

1 University of North Carolina at Chapel Hill, United States **2** Dartmouth College, United States **3** Brigham Young University, United States

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor: Open Journals

Reviewers:

- @openjournals

Submitted: 01 January 1970
Published: unpublished

Summary

The Interfere package is designed to research *intervention response prediction*, or, more specifically, predicting how a complex, dynamic system will respond to a never-before-seen intervention. When developing predictive methods capable of solving this problem, ideal benchmarking data comes exclusively from carefully controlled, longitudinal experiments. Unfortunately, unavoidable confounding, prohibitive costs and ethical boundaries often stymie attempts to run controlled experiments for many important problems. Synthetic data can be used as an alternative to experimental data and has been employed to benchmark predictive algorithms [cite causal graph etc]. However, in practice, the synthetic data used for benchmarking is often simplistic, and the simulation models used to generate it show a range of familiar shortcomings: static, linear relationships, overly independent noise profiles, and a lack of the complex mechanistic relationships and feedback loops that we find in the real world. As a solution, we propose a middle ground: Scientists in many disciplines, such as economics, ecology, systems biology and others, employ mechanistic dynamic models as an integral part of their research and application toolkit (Baker et al., 2018; Banks et al., 2017; Brayton et al., 2014; Izhikevich & Edelman, 2008). Many of these models are accompanied by a body of evidence demonstrating that they do indeed capture important characteristics of the real world problems they emulate. Instead of using simple synthetic data for benchmarking, we propose using datasets generated by scientific models of complex systems as a ground truth against which we can benchmark and test methods that predict intervention response. The logic is this: if a scientific model captures important characteristics of the world and a method can accurately predict intervention response *for that model* then the method ought to be able to predict intervention response for similar systems in the real world.

Interfere offers the first steps towards this vision by combining (1) a general interface for simulating the effect of interventions on dynamic models, (2) a suite of predictive methods and cross validated hyper parameter optimization tools, and (3) the first known extensible benchmark data set of dynamic intervention response scenarios see Figure 1.

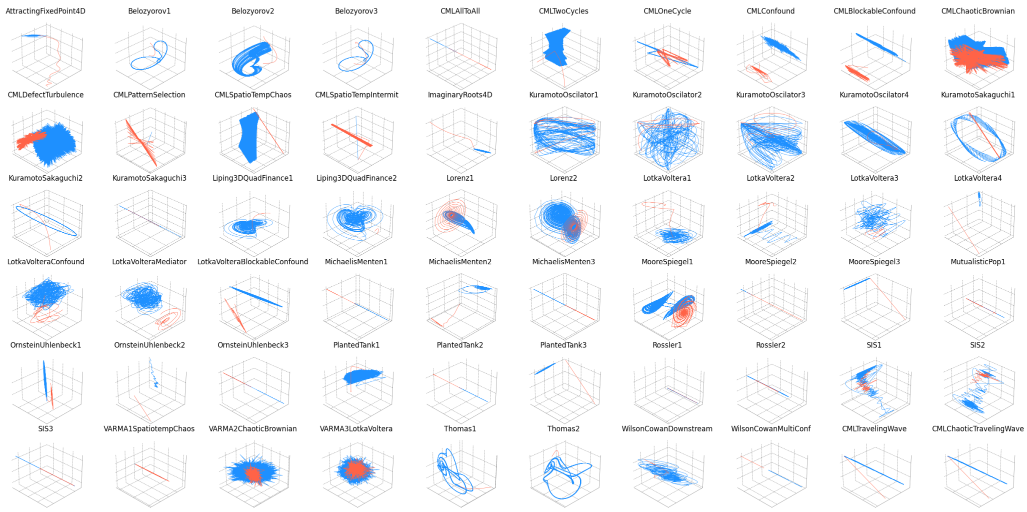

Figure 1: Three-dimensional trajectories of sixty scenarios simulated with the Interfere package. The models simulated here are either differential equations or discrete time difference equations. For each system, the trajectory in blue represents the natural behavior of the system and the red depicts how the system responds to a specified intervention. Many of the models pictured have more than three dimensions (in such cases, only the three dimensions of the trajectory with the highest variance are shown). These sixty scenarios make up the Interfere Benchmark 1.1.1 for intervention response prediction which is available online for download.

Statement of Need

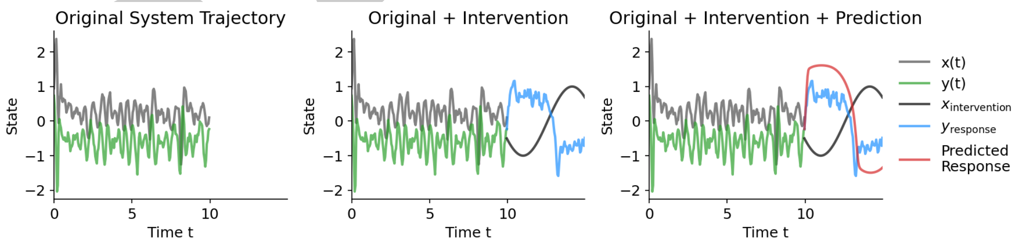

Figure 2: Original System Trajectory (Left): The natural, uninterupted evolution of the quaratic Belozyorov system (Belozyorov, 2015) simulated using the Interfere package with a small amount of stochastic noise. **System Trajectory After Intervention (Center):** The effect that taking exogenous control of $x(t)$ by via $\text{do}(x(t) = \sin(t))$ beginning at time $t = 10$ has on y. The intervention (black) and response (blue), depict a clear departure from the natural behavior of the system (green and gray). **Intervention Response Prediction (Right):** Example of forecasting the response of the Belozyorov system to the sinusoidal intervention. The ground truth response, $y(t)$ for $t > 10$ is plotted in blue. Here, an equation discovery algorithm, SINDy, (Brunton et al., 2016) is fit to the data that occurs prior to the intervention and makes an attempt to predict the intervention response (red curve).

Over the past twenty years, the scientific community has experienced the emergence of multiple frameworks for identifying causal relationships in observational data (Imbens & Rubin, 2015; Pearl, 2009; Wieczorek & Roth, 2019). The most influential frameworks are probabilistic and, while it is not a necessary condition for identifying causality, historically a static, linear relationship has often been assumed. However, when attempting to anticipate the response of complex dynamic systems in the medium and long term, a linear approximation of the dynamics can be insufficient. Therefore, researchers have increasingly begun to employ non-linear, dynamic techniques for causal discovery and forecasting (e.g. Runge, 2022). Still, there are relatively few techniques that are able to fit causal dynamic nonlinear models to data. Because of this, we see an opportunity to bring together the insights from recent advancements in causal inference with historical work in dynamic modeling and simulation.

In order to facilitate this cross pollination, we focus on a key problem — predicting how a complex system responds to a previously unobserved intervention — and designed the Interfere package for benchmarking tools aimed at intervention response prediction. The dynamic models contained in Interfere present challenges for computational methods that can likely only be addressed with the incorporation of mechanistic assumptions alongside probabilistic frameworks for causality. The Interfere package is a toolbox that allows researcher to validate predictive dynamic methods against simulated intervention scenarios. As such, the Interfere package encourages an opportunity for cross pollination between the probabilistic causal inference community and the modeling and simulation community.

Primary Contributions

The Interfere package provides three primary contributions. (1) Dynamically diverse counterfactuals at scale, (2) cross disciplinary forecast methods, and (3) comprehensive and extensible benchmarking.

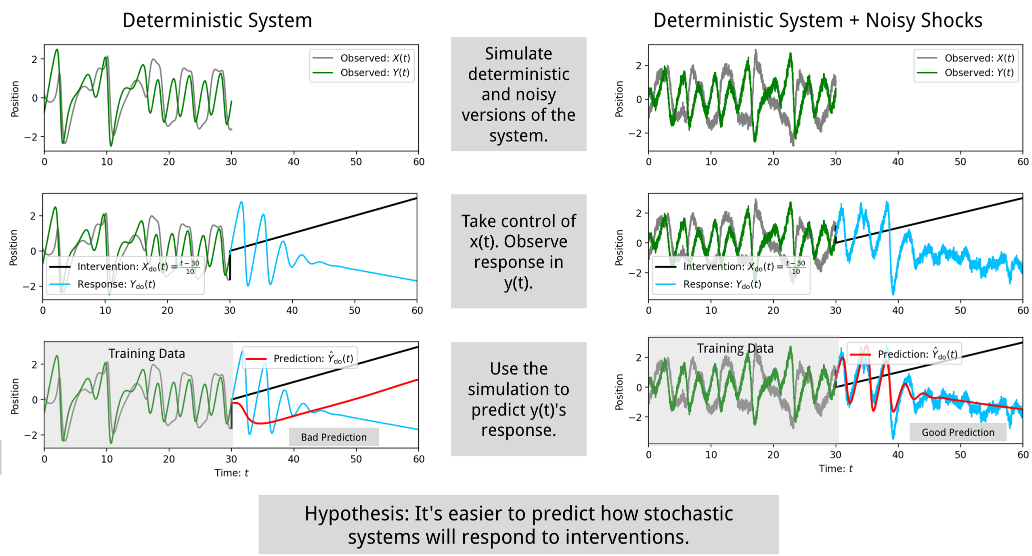

Figure 3: Example experimental setup possible with Interfere: Can stochasticity help reveal associations between variables? Interfere can be used to compare intervention response prediction for deterministic and stochastic versions of the same system.

1. Dynamically Diverse Counterfactuals at Scale

The "dynamics" submodule in the Interfere package contains over fifty dynamic models. It contains a mix of linear, nonlinear, chaotic, continuous time, discrete time, stochastic, and deterministic models. The models come from a variety of disciplines including finance, ecology, biology, neuroscience and public health. Each model inherits the from the Interfere BaseDynamics type and gains the ability to take exogenous control of any observed state and to add measurement noise. Most models also gain the ability to make any observed state stochastic where magnitude of stochasticity can be controlled by a simple scalar parameter or fine tuned with a covariance matrix.

Because of the difficulty of building models of complex systems, predictive methods for complex dynamics are typically benchmarked on less than ten dynamical systems (Brunton et al., 2016; Challu et al., 2023; Pathak et al., 2018; Prasse & Van Mieghem, 2022; Vlachas et al., 2020). As such, Interfere offers a clear improvement over current benchmarking methods for prediction in complex dynamics.

Most importantly, Interfere is built around interventions: the ability to take exogenous control of one or several state variables in a complex system and observe the response. Imbuing a suite of scientific models with general exogenous control is no small feat because models can be complex and are implemented in a variety of ways. Interfere offers the ability to produce complex dynamic intervention response and standard forecasting scenarios at scale. This unique feature enables large scale evaluation of dynamic causal prediction methods—tested against systems with properties of interest to scientists. For example, we can simulate the change in concentration of ammonia based on the nitrogen cycle and an exogenous fertilizing schedule.

2. Cross Disciplinary Forecast Methods

A second contribution of Interfere is the integration of dynamic *forecasting* methodologies from deep learning (LSTM, NHITS), applied mathematics (SINDy, Reservoir Computers) and social science (VAR). The Interfere "ForecastingMethod" class is expressive enough to describe, fit and predict with multivariate dynamic models and apply interventions to the states of the models during prediction. This cross disciplinary mix of techniques has the potential to produce new insights into the problem of intervention response prediction among others. For example, experiments using this package have revealed that cross validation error does not correlate with well with prediction error when LSTM and NHITS attempt to predict intervention response.

3. Comprehensive and Extensible Benchmarking

The third major contribution of Interfere is the collection of dynamic scenarios organized into the Interfere Benchmark. The Interfere Benchmark is a comprehensive and extensible set of dynamic scenarios that are conveniently available for testing methods that predict the effects of interventions. The benchmark set contains 60 intervention response scenarios for testing, each simulated with different levels of stochastic noise. Each scenario is housed in a JSON file, complete with full metadata annotation, documentation, versioning and commit hashes marking the commit of Interfere that was used to generate the data. The scenarios were reviewed by hand with some systems exposed to exogenous input to ensure that none of the key variables settle into a steady state. Additionally, all interventions were chosen in a manner such that the response of the target variable is a significant departure from its previous behavior.

The Interfere package enables researchers from various backgrounds to systematically study the problem of predicting intervention response on simulated data from a wide range of disciplines. It thereby facilitates future progress towards correctly anticipating how complex systems will respond in new, never before seen scenarios.

Related Software and Mathematical Foundations

Predictive Methods

The Interfere package draws from the Nixtla open source ecosystem for time series forecasting. We implemented intervention support for LSTM and NHITS from the NeuralForecast package, and for ARIMA from the StatsForecast package (Azul Garza, 2022; Olivares et al., 2022). We followed Nixtla's example for cross validation and hyperparameter optimization approaches. We integrated predictive methods from the PySINDy (Kaptanoglu et al., 2022) and StatsModels (Seabold & Perktold, 2010) packages. We also include ResComp, a reservoir computing method for global forecasts from (Harding et al., 2024). Hyperparameter optimization is designed around the Optuna framework (Akiba et al., 2019).

While other forecasting methods exist, integrating a method with Interfere requires that the method is capable of (1) multivariate endogenous dynamic forecasting, (2) support for exogenous variables, and (3) support for flexible length forecast windows or recursive predictions. Few forecasting methods meet these criteria, and it is our hope that this package can encourage the development of additional methods.

Dynamic Models

The table below list the dynamic models that are currently implemented in the Interfere package, plus attributions. These dynamic models in were implemented directly from mathematical descriptions except for two, "Hodgkin Huxley Pyclustering" and "Stuart Landau Kuramoto" which adapt existing simulations from the PyClustering package (Novikov, 2019).

Acknowledgements

The work described here was supported by an NSF Graduate Research Fellowship (DJP) and by award W911NF2510049 from the Army Research Office. The content is solely the responsibility of the authors and does not necessarily represent the official views of any agency supporting this research.

References

Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). Optuna: A next-generation hyperparameter optimization framework. *The 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2623–2631.

Azul Garza, C. C., Max Mergenthaler Canseco. (2022). *StatsForecast: Lightning fast forecasting with statistical and econometric models*. PyCon Salt Lake City, Utah, US 2022. https://github.com/Nixtla/statsforecast

Baker, R. E., Peña, J.-M., Jayamohan, J., & Jérusalem, A. (2018). Mechanistic models versus machine learning, a fight worth fighting for the biological community? *Biology Letters*, *14*(5), 20170660. https://doi.org/10.1098/rsbl.2017.0660

Banks, H. T., Banks, J. E., Bommarco, R., Curtsdotter, A., Jonsson, T., & Laubmeier, A. N. (2017). Parameter estimation for an allometric food web model. *International Journal of Pure and Apllied Mathematics*, *114*(1). https://doi.org/10.12732/ijpam.v114i1.12

Belozyorov, V. Y. (2015). Exponential-algebraic maps and chaos in 3D autonomous quadratic systems. *International Journal of Bifurcation and Chaos*, *25*(04), 1550048.

Brayton, F., Laubach, T., & Reifschneider, D. (2014). The FRB/US model: A tool for macroeconomic policy analysis. *FEDS Notes*, *2014-04*, 03.

Brunton, S. L., Proctor, J. L., & Kutz, J. N. (2016). Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, *113*(15), 3932–3937. https://doi.org/10.1073/pnas.1517384113

Challu, C., Olivares, K. G., Oreshkin, B. N., Ramirez, F. G., Canseco, M. M., & Dubrawski, A. (2023). NHITS: Neural Hierarchical Interpolation for Time Series Forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence*, *37*(6), 6989–6997. https://doi.org/10.1609/aaai.v37i6.25854

Harding, S., Leishman, Q., Lunceford, W., Passey, D. J., Pool, T., & Webb, B. (2024). Global forecasts in reservoir computers. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *34*(2), 023136. https://doi.org/10.1063/5.0181694

Imbens, G. W., & Rubin, D. B. (2015). *Causal Inference for Statistics, Social, and Biomedical*

Sciences: An Introduction (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139025751

Izhikevich, E. M., & Edelman, G. M. (2008). Large-scale model of mammalian thalamocortical systems. *Proceedings of the National Academy of Sciences*, *105*(9), 3593–3598. https://doi.org/10.1073/pnas.0712231105

Kaptanoglu, A. A., Silva, B. M. de, Fasel, U., Kaheman, K., Goldschmidt, A. J., Callaham, J., Delahunt, C. B., Nicolaou, Z. G., Champion, K., Loiseau, J.-C., Kutz, J. N., & Brunton, S. L. (2022). PySINDy: A comprehensive python package for robust sparse system identification. *Journal of Open Source Software*, *7*(69), 3994. https://doi.org/10.21105/joss.03994

Novikov, A. V. (2019). PyClustering: Data mining library. *Journal of Open Source Software*, *4*(36), 1230. https://doi.org/10.21105/joss.01230

Olivares, K. G., Challú, C., Garza, A., Canseco, M. M., & Dubrawski, A. (2022). *NeuralForecast: User friendly state-of-the-art neural forecasting models.* PyCon Salt Lake City, Utah, US 2022. https://github.com/Nixtla/neuralforecast

Pathak, J., Hunt, B., Girvan, M., Lu, Z., & Ott, E. (2018). Model-Free Prediction of Large Spatiotemporally Chaotic Systems from Data: A Reservoir Computing Approach. *Physical Review Letters*, *120*(2), 024102. https://doi.org/10.1103/PhysRevLett.120.024102

Pearl, J. (2009). *Causality* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511803161

Prasse, B., & Van Mieghem, P. (2022). Predicting network dynamics without requiring the knowledge of the interaction graph. *Proceedings of the National Academy of Sciences*, *119*(44), e2205517119. https://doi.org/10.1073/pnas.2205517119

Runge, J. (2022). *Discovering contemporaneous and lagged causal relations in autocorrelated nonlinear time series datasets.* arXiv. https://doi.org/10.48550/arXiv.2003.03685

Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and statistical modeling with python. *9th Python in Science Conference*.

Vlachas, P. R., Pathak, J., Hunt, B. R., Sapsis, T. P., Girvan, M., Ott, E., & Koumoutsakos, P. (2020). Backpropagation algorithms and Reservoir Computing in Recurrent Neural Networks for the forecasting of complex spatiotemporal dynamics. *Neural Networks*, *126*, 191–217. https://doi.org/10.1016/j.neunet.2020.02.016

Wieczorek, A., & Roth, V. (2019). Information Theoretic Causal Effect Quantification. *Entropy*, *21*(10), 975. https://doi.org/10.3390/e21100975