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 **ROR** 2 Dartmouth College, United States **ROR** 3 Brigham Young University, United States **ROR**

Summary

The Interfere package focuses on *intervention response prediction*—forecasting how complex, dynamic systems respond to novel interventions. While ideal data for benchmarking predictive methods for such problems would come from controlled, longitudinal experiments, such experiments are often infeasible due to confounding, cost, and ethical constraints. In contrast, many synthetic benchmarks rely on models with static linear relationships and/or independent noise profiles that lack the complex interdependent feedback present in the real world. However, many disciplines, including economics, ecology, and systems biology, employ mechanistic dynamic models (Baker et al., 2018; Banks et al., 2017; Brayton et al., 2014; Izhikevich & Edelman, 2008) with evidence demonstrating they capture important real-world characteristics. We propose using datasets generated by such complex systems to benchmark methods that predict intervention response. If a model captures important real-world characteristics and a method can accurately predict intervention response *for that model* then the method might also predict intervention response for similar real-world systems. Interfere works towards this vision by combining (1) a general interface for simulating interventions on dynamic models, (2) a suite of predictive methods and cross-validated hyperparameter optimization tools, and (3) an extensible benchmark data set of dynamic intervention response scenarios (see Figure 1).
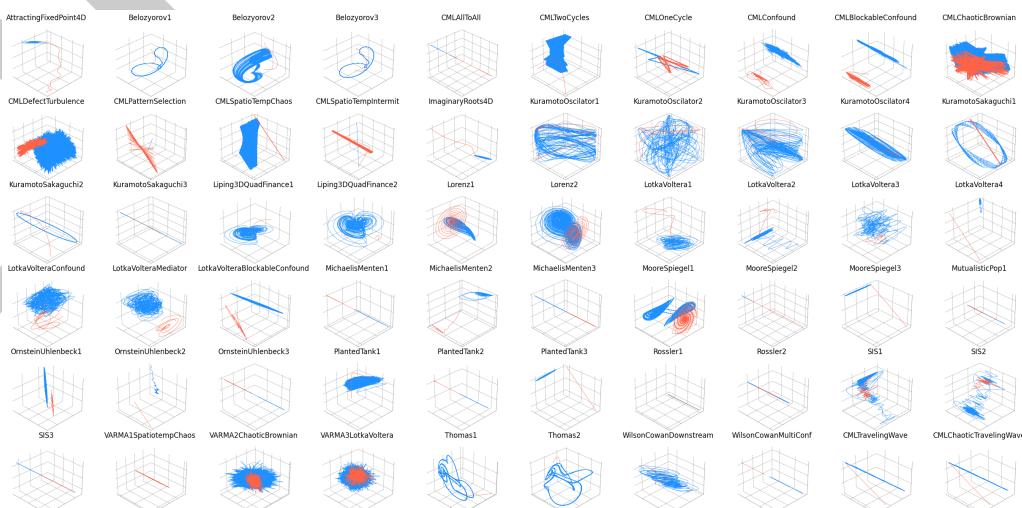
Figure 1: Three-dimensional trajectories of sixty scenarios from differential equation and discrete time difference equation models. Blue trajectories represent natural behavior of the system. Red depicts response to a specified intervention. The three dimensions with highest variance are shown. These sixty scenarios comprise the downloadable Interfere Benchmark 1.1.1.

Statement of Need
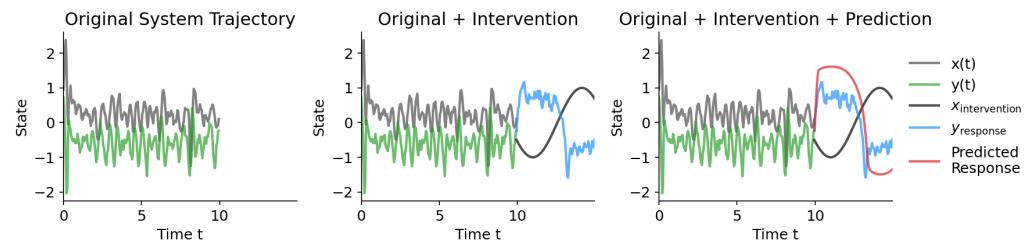
Figure 2: Original System Trajectory (Left): The natural evolution of the quadratic Belozyorov system (Belozyorov, 2015) with small stochastic noise. **System Trajectory After Intervention (Center):** Exogenous control of $x(t)$ by $do(x(t) = \sin(t))$ for $t > 10$ affects $y(t)$. The intervention (black) and response (blue) indicate clear departure from natural behavior (green and gray). **Intervention Response Prediction (Right):** Equation discovery by SINDy (Brunton et al., 2016) is fit to data prior to the intervention to predict intervention response (red).

Multiple frameworks have emerged for identifying causal relationships from data (Imbens & Rubin, 2015; Pearl, 2009; Wieczorek & Roth, 2019). Many frameworks assume a static, linear relationship. Because a linear approximation can be insufficient for anticipating the response of complex dynamic systems at medium and long times, non-linear, dynamic techniques have been developed for causal discovery and forecasting (e.g. Runge, 2022). Nevertheless, there are relatively few intervention oriented techniques to fit dynamic nonlinear models to data.

Interfere's benchmarking approach is enabled by a unified toolkit for simulating intervention response, standardized method evaluation, and systematic cross-validation across diverse dynamic systems. By lowering barriers to method evaluation, Interfere facilitates development of prediction methods that translate from theory to real-world application.

Primary Contributions

Interfere provides (1) dynamically diverse counterfactuals, (2) cross-disciplinary forecast methods, and (3) comprehensive and extensible benchmarking.
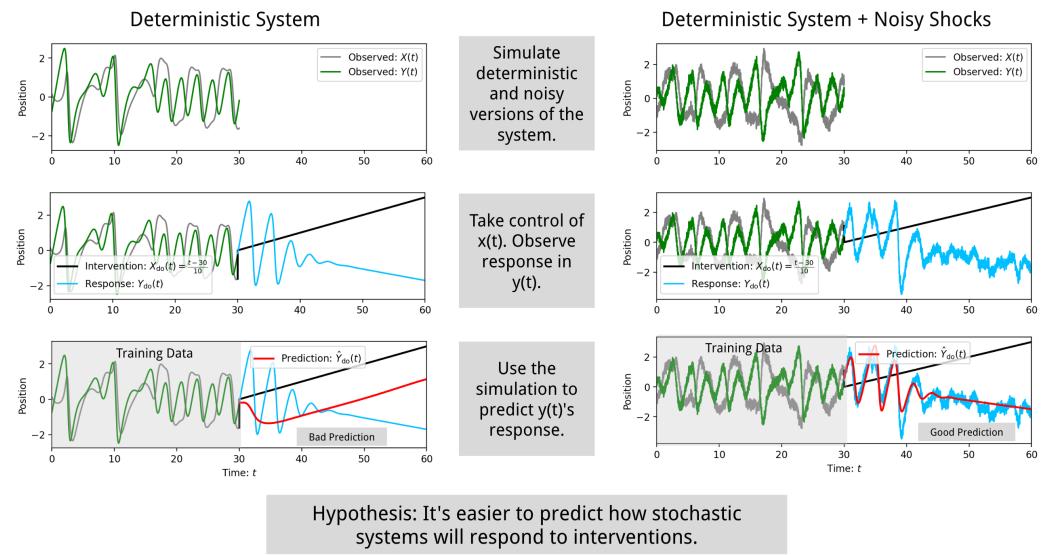
Figure 3: Can stochasticity help reveal associations between variables? In this example, Interfere compares intervention response prediction for deterministic and stochastic versions of the same system.

1. Dynamically Diverse Counterfactuals at Scale

Whereas many predictive methods are typically benchmarked on < 10 systems (Brunton et al., 2016; Challu et al., 2023; Pathak et al., 2018; Prasse & Van Mieghem, 2022; Vlachas et al., 2020), Interfere's "dynamics" submodule contains > 50 models, including linear, nonlinear, chaotic, continuous-time, discrete-time, stochastic, and deterministic models, from disciplines including finance, ecology, biology, neuroscience and public health. Most importantly, Interfere is built for studying interventions: the inherited "BaseDynamics" type allows exogenous control of any observed state, added measurement noise, and, for most models, stochasticity parameterized by a scalar or covariance matrix. Interfere thus offers a user-friendly framework to produce complex dynamic intervention response and forecasting scenarios at scale.

2. Cross-Disciplinary Forecast Methods

Interfere integrates dynamic forecasting methodologies from deep learning (LSTM, NHITS), applied mathematics (SINDy, Reservoir Computers) and social science (VAR). The "ForecastingMethod" class is expressive enough to describe, fit and predict with multivariate dynamic models, and apply interventions to model states.

3. Comprehensive and Extensible Benchmarking

The Interfere Benchmark is a comprehensive and extensible set containing 60 intervention response scenarios for testing, each simulated with different levels of stochastic noise. Each scenario is housed in a JSON file, with metadata annotation, documentation, versioning and commit hashes marking the Interfere commit used to generate data.

Scenarios were reviewed by hand with exogenous input such that none of the key variables settle to steady state. Interventions were chosen so target variable response significantly departs from prior behavior. We aim for this benchmark to facilitate future progress towards predicting how complex systems respond to never before seen situations.

Related Software and Mathematical Foundations

Predictive Methods

Interfere draws from the Nixtla open source ecosystem for time series forecasting, including cross validation and hyperparameter optimization. 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). Interfere includes predictive methods from the PySINDy (Kaptanoglu et al., 2022) and StatsModels (Seabold & Perktold, 2010) packages. We also include ResComp, a reservoir computing method for forecasts (Harding et al., 2024). Hyperparameter optimization is designed around the Optuna framework (Akiba et al., 2019).

Integrating a method into Interfere requires that the method supports (1) multivariate endogenous dynamic forecasting, (2) exogenous variables, and (3) flexible length forecast windows or recursive predictions. While few forecasting methods meet these criteria, we hope Interfere will encourage development of additional methods.

Acknowledgements

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. https://doi.org/10.1142/S0218127415500480

Brayton, F., Laubach, T., & Reifschneider, D. (2014). The FRB/US model: A tool for macroeconomic policy analysis. *FEDS Notes*, *2014-04*, 03. https://doi.org/10.17016/2380-7172.0012

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

105 forecasts in reservoir computers. *Chaos: An Interdisciplinary Journal of Nonlinear Science*,
106 *34*(2), 023136. https://doi.org/10.1063/5.0181694

107 Imbens, G. W., & Rubin, D. B. (2015). *Causal Inference for Statistics, Social, and Biomedical*
108 *Sciences: An Introduction* (1st ed.). Cambridge University Press. https://doi.org/10.1017/
109 CBO9781139025751

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

113 Kaptanoglu, A. A., Silva, B. M. de, Fasel, U., Kaheman, K., Goldschmidt, A. J., Callaham, J.,
114 Delahunt, C. B., Nicolaou, Z. G., Champion, K., Loiseau, J.-C., Kutz, J. N., & Brunton, S. L.
115 (2022). PySINDy: A comprehensive python package for robust sparse system identification.
116 *Journal of Open Source Software*, *7*(69), 3994. https://doi.org/10.21105/joss.03994

117 Olivares, K. G., Challú, C., Garza, A., Canseco, M. M., & Dubrawski, A. (2022). *NeuralForecast:*
118 *User friendly state-of-the-art neural forecasting models.* PyCon Salt Lake City, Utah, US
119 2022. https://github.com/Nixtla/neuralforecast

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

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

125 Prasse, B., & Van Mieghem, P. (2022). Predicting network dynamics without requiring the
126 knowledge of the interaction graph. *Proceedings of the National Academy of Sciences*,
127 *119*(44), e2205517119. https://doi.org/10.1073/pnas.2205517119

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

130 Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and statistical mod-
131 eling with python. *9th Python in Science Conference*. https://doi.org/10.25080/
132 majora-92bf1922-011

133 Vlachas, P. R., Pathak, J., Hunt, B. R., Sapsis, T. P., Girvan, M., Ott, E., & Koumoutsakos,
134 P. (2020). Backpropagation algorithms and Reservoir Computing in Recurrent Neural
135 Networks for the forecasting of complex spatiotemporal dynamics. *Neural Networks*, *126*,
136 191–217. https://doi.org/10.1016/j.neunet.2020.02.016

137 Wieczorek, A., & Roth, V. (2019). Information Theoretic Causal Effect Quantification. *Entropy*,
138 *21*(10), 975. https://doi.org/10.3390/e21100975