

Modeling and Inference of Mixed Dynamics and Detection of Causal Emergent Features

William Casey^{1,*}, Leigh Metcalf², Shirshendu Chatterjee³, Heeralal Janwa⁴, Ernest Battifarano⁵, April Edwards¹, and Yaron Gvili⁶

¹United States Naval Academy

²Carnegie Mellon University

³City University of New York, Graduate Center and City College

⁴University of Puerto Rico, Rio Piedras

⁵New York University

⁶Independent Scholar

*corresponding wcasey@usna.edu

ABSTRACT

Many real-world problems feature nonlinear dynamic processes. Classical mathematical models may be adequate to describe a single dynamic process in isolation, but can be easily undermined by two natural and simple kinds of phenomenological variations: the emergence (or activation) of an additional dynamic process, and events that affect the parameters of an active process. COVID-19 data offers an important case study expressing these phenomenological variations that deeply challenge the classical SIR epidemiological model, and call for novel mathematical methods to detect and adapt to these critical variations.

We address the modeling issues with a novel mathematical framework that reenvision data as a mixture of multiple causal generating processes, each subject to possible parameter change points. The new viewpoint extends nonlinear classical models in a manner that overcomes many of these types of phenomenological variations and enables a highly adaptive modeling closely linked to causal events. The new model space unifies a wider class of dynamics and is particularly effective at fitting multi-surge data and explaining key causal events related to surge origination. To demonstrate, we construct a mixture of logistic models termed the Adaptive Logistic Model (ALM), and then formulate appropriate nonlinear least squares optimization and regularization goals, and then apply ALM to data. To validate the approach, we return to COVID-19 forecasting (for case count), and compare ALM directly to other forecasting methods. ALM forecast accuracy is competitive with all leading forecast methods, but its greatest utility may be in how it detects changing dynamics (change-points) and retains far fewer but more interpretable parameters relating naturally to cause and intervening change. The method can be applied more generally as it adapts well to the multi-generative nature of many time series data problems. We demonstrate ALM robustness through data experiments in hydrology, economics, cybersecurity, and social media.