



On Explaining the Surprising Success of Reservoir Computing Forecaster of Chaos? The Universal Machine Learning Dynamical System with Contrasts to VAR and DMD suggests a Next Generation Reservoir Computing – NG–RC

Erik Bollt¹

Keywords: Reservoir Computing; Vector autoregression (VAR); Machine Learning; Dynamical System; Dynamic Mode Decomposition (DMD); Chaos.

MSC2010 codes: 62M10, 37N30, 68Q32

Abstract

Machine learning has become a widely popular and successful paradigm, including for data-driven science. A major application problem is forecasting complex dynamical systems. Artificial neural networks (ANN) have evolved as a clear leading approach, and recurrent neural networks (RNN) are considered to be especially well suited for. In this setting, the echo state networks (ESN) or reservoir computer (RC) have emerged for simplicity and computational advantages. Instead of a fully trained network, an RC trains only read-out weights. However, why and how an RC works at all, despite randomly selected weights is perhaps a surprise. To this end, we analyze a simplified RC, where the internal activation function is an identity function. We explicitly connect the RC with linear activation and linear read-out to well developed time-series literature on vector autoregressive averages (VAR) that includes theorems on representability through the WOLD theorem, which already perform reasonably for short term forecasts. In the case of a linear activation and now popular quadratic read-out RC, we explicitly connect to a nonlinear VAR (NVAR), which performs quite well. Further, we associate this paradigm to the now widely popular dynamic mode decomposition (DMD), and thus these three are in a sense different faces of the same concept. We illustrate our observations in terms of popular benchmark examples including Mackey-Glass differential delay equations and the Lorenz63 system. There is also a beautiful connection to concepts of concentration of measure about the geometry of data in high dimensions and the random projection theory that leads to the Johnson-Lindenstrauss theorem. With these principles in mind, we will describe a dramatically simpler, more efficient and faster variation as an implicit RC, that we recently introduced as NG-RC.

¹Department of Electrical and Computer Engineering and C^3S^2 the Clarkson Center for Complex Systems Science, Clarkson University, Potsdam, New York 13699, USA. Email: ebollt@clarkson.edu