Data-driven Modeling, Prediction and Predictability:
The Complex Systems Framework

Surja Sharma, Erin Lynch and Eugenia Kalnay
University of Maryland, College Park

V. Krishnamurthy
George Mason University, Fairfax

GSFC AI Workshop
Nov. 2018
Nonlinear Dynamics and Complexity

Dynamics  (Lorenz, 1963)
Deterministic dynamics, Chaos
Quantitative results
Weak connection with data

Structure  (Mandelbrot, 1977)
Real objects in nature
(Trees, clouds, coastline, etc.)
Fractals and Multifractals

Dynamics + Structure
Spatio-temporal Dynamics
Data-driven modeling in Complex Systems Framework

Machine Learning / Artificial Intelligence
Reconstruction of Dynamics

“Geometry from a time series”
(Packard et al., 1980)

Embedding theorem (Takens, 1981)

Time series data: \( x(t) \)

Time-delay embedding:
\[
x_k(t_i) = x(t_i + (k-1)\tau)
\]

Reconstructed space:
\[
X_i = \{x_1(t_i), x_2(t_i), x_3(t_i), ..\}
\]


Complex Systems Framework:
- Low-dimensional (data-driven) modeling
- Dynamical prediction (Dynamics)
- Predictability analysis (Statistics)
Space Weather: Prediction and Predictability

Data-driven Modeling:
Phase space reconstruction of driver (solar wind) – response (magnetosphere)
• Storms (Dst)
• Substorms (AL)
• Killer electron flux

First Predictions
[Sharma, Rev. Geophys., 1995]
Early contributions to AI / Machine Learning

Solar wind (VBs)
Past data: auroral electrojet index AL
Predicted AL and Predictability

[Ukhorskiy et al., 2002, 2004].
Predictability of Space Weather

Global or Coherent aspects of the Magnetosphere

Demonstrated by

• Low-dimensionality – reconstruction of phase space [Vassiliadis et al., 1990; Sharma et al., 1993]
• Modeling [Baker et al., 1990]
• Phenomenology [Siscoe, 1991]
• Phase transition-like behavior from data-driven modeling and MHD simulations

implies Predictability

Fundamental contribution based on data-driven modeling
(AI / Machine Learning)

Similar to early results on dynamical behavior of the atmosphere

Transition from higher (orange) to lower (green) level.
AL index data – observed and from global MHD simulations.
Extreme events and Ensemble forecasting

- Data-driven models without governing equations
- Forecasts using Ensemble Transform Kalman Filter (ETKF)
- Ensemble spread as an indicator of extreme events

Data-driven Prediction of Monsoon

Phase space reconstruction model (PSRM).

Rainfall data on 0.25 deg longitude × 0.25 deg latitude grid for 1901-2009 (1800 stations)

Climate Forecasting System (CFSv2)
State of the art numerical model (NOAA)

Modeling by Reconstruction using Rainfall and CFSv2 data.

Improvement of predictability

Comparison of predictions of PSRM and CFSv2

Key results and conclusions:

Intraseasonal oscillations are predictable

Predictability of intraseasonal phenomena such as MJO and midlatitude processes

Data-driven modeling provides higher predictability

Modeling and prediction of spatio-temporal structure of space weather

Spatio-temporal Data: Networks of monitoring stations

Complex Systems Framework

- Data-driven Modeling
- Dynamical prediction (global and spatially extended)
- Characterization of predictability
- Extreme events: Quantification of predictability
- Predictability of extreme events from Big Data
- Quantitative measure of the likelihood of extreme space weather events (data-driven modeling)
- Prediction of Intraseasonal climate (Indian Monsoon)
- Applications and a framework for artificial intelligence, and machine learning
- Fourth paradigm – Data –enabled science