Decadal Forecasts with ResDMD: a Residual DMD Neural Network

Eduardo Rodrigues

July 23rd, 2021

Context and objective

IBM

time

Climate
(many decades,
e.g. CMIP)

1-10 years
Support decisions over the next seasons

Season
(typically 3 to 9 months
e.g. TWC probabilistic seasonal forecast

Weather
(up to 15 days,
e.g. TWC daily/hourly forecast)

Use data-driven methods to bridge the gap between seasonal and climate predictions

Linear Inverse Model / Dynamic Mode Decomposition

It is run to make decadal forecasts:


Physical Sciences Laboratory   About   People   Research   Data

Home » Decadal LIM Forecasts » Current LIM Decadal Forecasts

Decadal LIM Forecasts : 1-9 year leads

(Experimental NOAA/ESRL PSL and University of Colorado/CIRES Forecast)

New! **Decadal LIM VERSION 1.0**: Forecast indices now available.
Forecast page is out of beta since paper is published, but some parts of the site may still be incomplete.

LIM/DMD seeks to fit a model (linear dynamical model) to data, that is, to fit a model such:

$$\frac{d}{dt}\mathbf{x} = \mathbf{A}\mathbf{x}$$

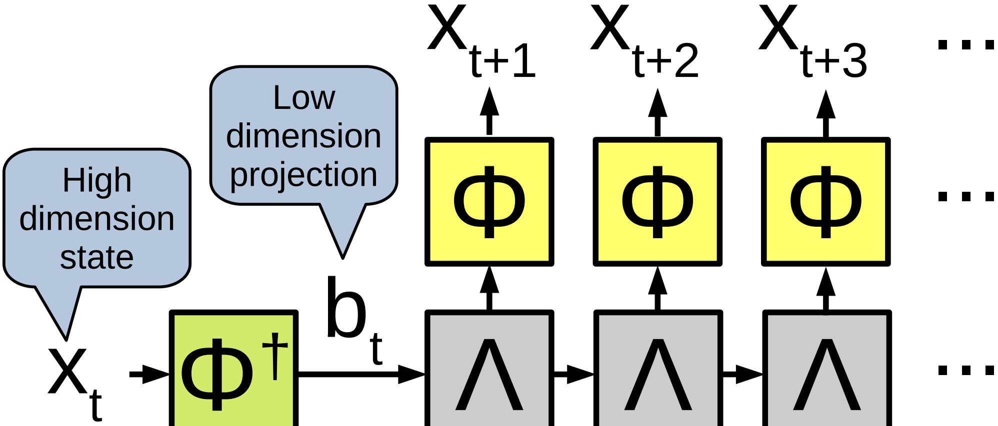
Which has an exact solution given by:

$$\mathbf{x}(t_0 + t) = e^{\mathbf{A}t}\mathbf{x}(t_0)$$

The dynamics are characterized by the eigenvalues and eigenvectors of **A**.

Everything is linear

Solution block diagram



Proposed extension for non-linear terms

The standard procedure assumes that the dynamics can be approximated linearly:

2. Linear inverse modeling

Consider an atmospheric state vector $\mathbf{X}$. We define anomalies as $\mathbf{x} = \mathbf{X} - \overline{\mathbf{x}}$, where $\overline{\mathbf{x}}$ is some base state, typically a time mean. Then the evolution of $\mathbf{x}$ may be represented as

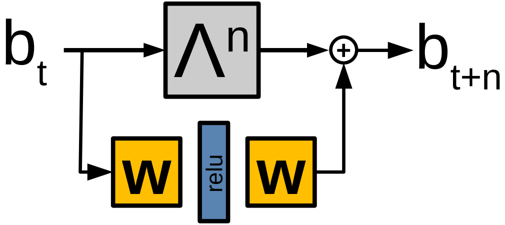
$$\frac{d\mathbf{x}}{dt} = \mathbf{L}\mathbf{x} + \mathbf{N}(\mathbf{x}), \qquad (1)$$

where $\mathbf{L}$ is the linearized part of the dynamical equations and $\mathbf{N}(\mathbf{x})$ represents nonlinear terms. Note that $\mathbf{L}$ and $\mathbf{N}(\mathbf{x})$ depend upon $\overline{\mathbf{x}}$. In general, the statistics of $\mathbf{x}$ may not be understood without a detailed knowledge of $\mathbf{N}(\mathbf{x})$. In some highly nonlinear systems, however, for suitable temporal and/or spatial averaging $\mathbf{N}(\mathbf{x})$ may be approximated as $\mathbf{N}(\mathbf{x}) \approx \mathbf{T}\mathbf{x} + \mathbf{F}_s$, where $\mathbf{T}$ is a linear operator and $\mathbf{F}_s$ is noise that is white in time but that may be spatially correlated. Thus under this averaging (1) may be expressed as

$$\frac{d\mathbf{x}}{dt} = \mathbf{B}\mathbf{x} + \mathbf{F}_s \qquad (2)$$

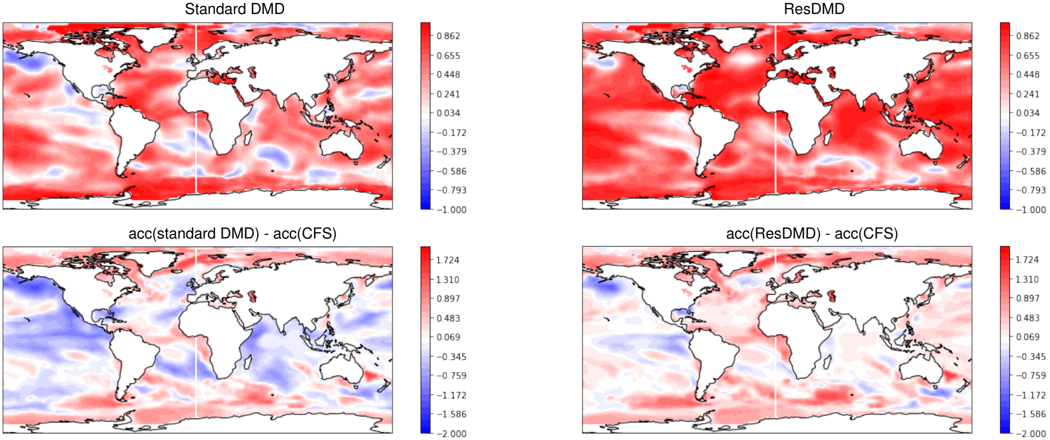
The proposed extension aims to approximate the non-linear term:

$$\frac{d}{dt}\mathbf{x}(t) = L\mathbf{x}(t) + N(\mathbf{x}(t))$$

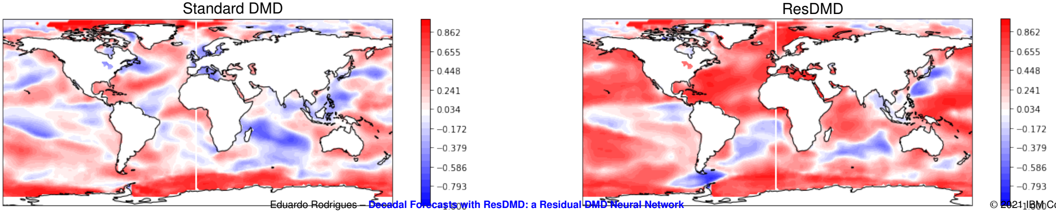


- Similar to Resnet, but the residual represents the non-linear term in the dynamics
- Uses an idea I had explored previously (DeepDownscale), in which **before the SGD training, network gives a sensible result** (in this case the LIM/DMD one)

Monthly SST - testing data (2010 to 2020) - 5 month leadtime



Monthly SST - testing data (2010 to 2020) - 1 year leadtime



Final remarks

- proposed a novel extension to the DMD method
- started with SST since it is slow varing but still important variable
- intend to fuse (probabilistically) data-driven with physically driven methods (e.g. CFS)
- computing SVD can be done in parallel

https://arxiv.org/abs/2106.11111

Eduardo Rodrigues
edrodri@br.ibm.com