

Exploring SCALE Weather Data via Koopman Modes

Zhicheng Zhang¹ & Yoshihiko Susuki¹ & Atsushi Okazaki²

¹*Department of Electrical Engineering, Kyoto University*

²*Institute of Advanced Academic Research, Chiba University*

第67回自動制御連合会 姫路

2024・11・23

KYOTO UNIVERSITY

京都大学



Purpose and Outline of Presentation

- Report a trial of Koopman mode decomposition (KMD) of **SCALE simulation data**
- Question 1: Can we capture a dominant (e.g., **transient**) Koopman mode (i.e., linear evolution of an observable) embedded in a **short-term** simulation data?
- Question 2: Which DMD (KMD algorithm) type is good to the simulation data?
and so on...

1. Review of Koopman operator framework: KMD and DMD
2. Its application to **vorticity magnitude (scalar-field)** in SCALE simulation data
3. Discussion on **humidity ratio field and low-dim. system representation** .
4. Summary

Previous Works on Weather and Climate Systems



Extended-range statistical ENSO prediction through operator-theoretic techniques for nonlinear dynamics

Xinyang Wang¹, Joanna Slawinska² & Dimitrios Giannakis^{3*}

Forecasting the El Niño-Southern Oscillation (ENSO) has been a subject of vigorous research due to the important role of the phenomenon in climate dynamics and its worldwide socioeconomic impacts. Over the past decades, numerous models for ENSO prediction have been developed, among which statistical models approximating ENSO evolution by linear dynamics have received significant attention owing to their simplicity and comparable forecast skill to first-principles models at short lead times. Yet, due to highly nonlinear and chaotic dynamics (particularly during ENSO initiation), such models have limited skill for longer-term forecasts beyond half a year. To resolve this limitation, here we employ a new nonparametric statistical approach based on analog forecasting, called kernel analog forecasting (KAF), which avoids assumptions on the underlying dynamics through the use of nonlinear kernel methods for machine learning and dimension reduction of high-dimensional datasets. Through a rigorous connection with Koopman operator theory for dynamical systems, KAF yields statistically optimal predictions of future ENSO states as conditional expectations, given noisy and potentially incomplete data at forecast initialization. Here, using industrial-era Indo-Pacific sea surface temperature (SST) as training data, the method is shown to successfully predict the Niño 3.4 index in a 1998–2017 verification period out to a 10-month lead, which corresponds to an increase of 3–8 months (depending on the decade) over a benchmark linear inverse model (LIM), while significantly improving upon the ENSO predictability “spilling barrier”. In particular, KAF successfully predicts the historic 2015/16 El Niño at initialization times as early as June 2015, which is comparable to the skill of current dynamical models. An analysis of a 1300-yr control integration of a comprehensive climate model (CCSM4) further demonstrates that the enhanced predictability afforded by KAF holds over potentially much longer leads, extending to 24 months versus 18 months in the benchmark LIM. Probabilistic forecasts for the occurrence of El Niño/La Niña events are also performed and assessed via information-theoretic metrics, showing an improvement of skill over LIM approaches, thus opening an avenue for environmental risk assessment relevant in a variety of contexts.

Kernel approach on ENSO

- In this presentation, we will focus on:
 - 1) SCALE weather simulation data (rather than SST)
 - 2) non-steady-state (or non-stationary) behavior
 - 3) short-time dynamics (transient pattern)
 - 4) Koopman mode decomposition (KMD) and model-reduction (a low-dim. linear system representation of original high-dim. nonlinear systems)

KYOTO UNIVERSITY



ARTICLE
<https://doi.org/10.1038/s41467-021-26357-x> OPEN
Spectral analysis of climate dynamics with operator-theoretic approaches
Gary Froyland¹, Dimitrios Giannakis^{2,3,8*}, Benjamin R. Lintner⁴, Maxwell Pike⁴ & Joanna Slawinska

The Earth's climate system is a classical example of a multiscale, multiphysics dynamical system with an extremely large number of active degrees of freedom, exhibiting variability on scales ranging from micrometers and seconds in cloud microphysics, to thousands of kilometers and centuries in ocean dynamics. Yet, despite this dynamical complexity, climate dynamics is known to exhibit coherent modes of variability. A primary example is the El Niño Southern Oscillation (ENSO), the dominant mode of interannual (3–5 yr) variability in the climate system. The objective and robust characterization of this and other important phenomena presents a long-standing challenge in Earth system science, the resolution of which would lead to improved scientific understanding and prediction of climate dynamics, as well as assessment of their impacts on human and natural systems. Here, we show that the spectral theory of dynamical systems, combined with techniques from data science, provides an effective means for extracting coherent modes of climate variability from high-dimensional model and observational data, requiring no frequency prefiltering, but recovering multiple timescales and their interactions. Lifecycle composites of ENSO are shown to improve upon results from conventional indices in terms of dynamical consistency and physical interpretability. In addition, the role of combination modes between ENSO and the annual cycle in ENSO diversity is elucidated.

Spectral analysis of dynamics

APRIL 2021

NAVARRA ET AL.

1227

15 AUGUST 2024

NAVARRA ET AL.

4095

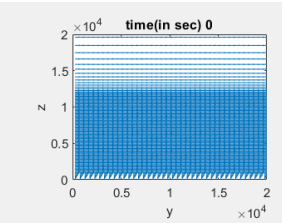
Estimation of Koopman Transfer Operators for the Equatorial Pacific SST

ANTONIO NAVARRA,^{a,b} JOE TRIBBIA,^c AND STEFAN KLUS^d
^a Centro Euromediterraneo sui Cambiamenti Climatici, Bologna, Italy
^b Department of Biological, Geological, and Environmental Sciences (BIGEA), University of Bologna, Bologna, Italy
^c National Center for Atmospheric Research, Boulder, Colorado
^d Department of Mathematics, University of Surrey, Guildford, United Kingdom
(Manuscript received 7 May 2020, in final form 5 January 2021)

ABSTRACT: In the last years, ensemble methods have been widely popular in atmospheric, climate, and ocean dynamics investigations and forecasts as convenient methods to obtain statistical information on these systems. In many cases, ensembles have been used as an approximation to the probability distribution that has acquired more and more a central role, as the importance of a single trajectory, or member, was recognized as less informative. This paper shows that using results from the dynamical systems and more recent results from the machine learning and AI communities, we can arrive at a direct estimation of the probability distribution evolution and also at the formulation of predictor systems based on a nonlinear formulation. The paper introduces the theory and demonstrates its application to two examples. The first is a one-dimensional system based on the Niño-3 index; the second is a multidimensional case based on time series of monthly mean SST in the Pacific. We show that we can construct the probability distribution and set up a system to forecast its evolution and derive various quantities from it. The objective of the paper is not strict realism, but the introduction of these methods and the demonstration that they can be used also in the complex, multidimensional environment typical of atmosphere and ocean applications.

KEYWORDS: Atmosphere; Ocean; Statistical techniques; Superensembles

Koopman/transfer operator for SST



The controllability of climate or weather systems is gaining increasing attention.

Variability of SST through Koopman Modes

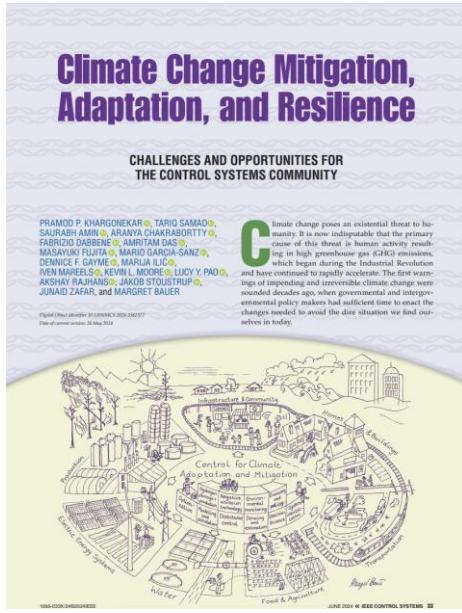
ANTONIO NAVARRA,^{a,b} JOE TRIBBIA,^c STEFAN KLUS,^d AND PAULA LORENZO-SANCHEZ^{a,b}
^a Centro Euromediterraneo sui Cambiamenti Climatici, Bologna, Italy
^b Dipartimento di Scienze Geologiche, Biologiche e Ambientali, Università di Bologna, Bologna, Italy
^c National Center for Atmospheric Research, Boulder, Colorado
^d School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh, United Kingdom
^e Dipartimento di Fisica e Astronomia, Università di Bologna, Bologna, Italy
(Manuscript received 2 June 2023, in final form 26 February 2024, accepted 22 April 2024)

ABSTRACT: The majority of dynamical systems arising from applications show a chaotic character. This is especially true for climate and weather applications. We present here an application of Koopman operator theory to tropical and global sea surface temperature (SST) that yields an approximation to the continuous spectrum typical of these situations. We also show that the Koopman modes yield a decomposition of the datasets that can be used to categorize the variability. Most relevant modes emerge naturally, and they can be identified easily. A difference with other analysis methods such as empirical orthogonal function (EOF) or Fourier expansion is that the Koopman modes have a dynamical interpretation, thanks to their connection to the Koopman operator, and they are not constrained in their shape by special requirements such as orthogonality (as it is the case for EOF) or pure periodicity (as in the case of Fourier expansions). The pure periodic modes emerge naturally, and they form a subspace that can be interpreted as the limiting subspace for the variability. The stationary states therefore are the scaffolding around which the dynamics takes place. The modes can also be traced to the Niño variability and in the case of the global SST to the Pacific decadal oscillation (PDO).

SIGNIFICANCE STATEMENT: We compute the Koopman modes for the tropical and global SST, demonstrating that significant dynamical modes can be identified also in complex high-dimensional datasets with continuous spectra. The Koopman modes are then used to identify the stationary subspace, namely, the limit subspace for the invariant evolution of the system.

KEYWORDS: Sea surface temperature; Interannual variability; Tropical variability; Data science; Support vector machines

Continuous spectrum for SST



京都大学

Weather Dynamics & Observables

- Nonlinear CT/DT dynamical system:

$$x^+ = F(x_t), \quad x \in \mathbb{X} \subseteq \mathbb{R}^n \quad t = 0, 1, 2, \dots$$

State equation

State/Phase space



- Output equation

$$y_t = g(x_t),$$

Observable space

$$\forall g \in \mathcal{O}$$

❖ Hilbert / Banach space
(complete inner product / normed space)

Q: How to realize system representation?

❖ Fact: true weather dynamical systems often go beyond the mathematical toy model, and can be unknown

$$F: \mathbb{R}^n \rightarrow \mathbb{R}^n$$

- ❑ **Nonlinear map**: describes the time evolution of the **state**, such as the nonlinear weather or climate systems

EX: Lorenz 96 model, chaotic dynamics

$$g: \mathbb{R}^n \rightarrow \mathbb{R} \text{ or } \mathbb{C}$$

- ❑ **Observable**: the function of the state that can be **measured** directly from **data**
- ❑ E.g., scalar, vortex, pressure, temperature, humidity ratio fields in SCALE Simulation
- ❑ **Importance**: The choice of dictionaries of observable (linear/nonlinear) **Open issue!**

Koopman Operator: A (Bounded) Linear Operator

Koopman Framework

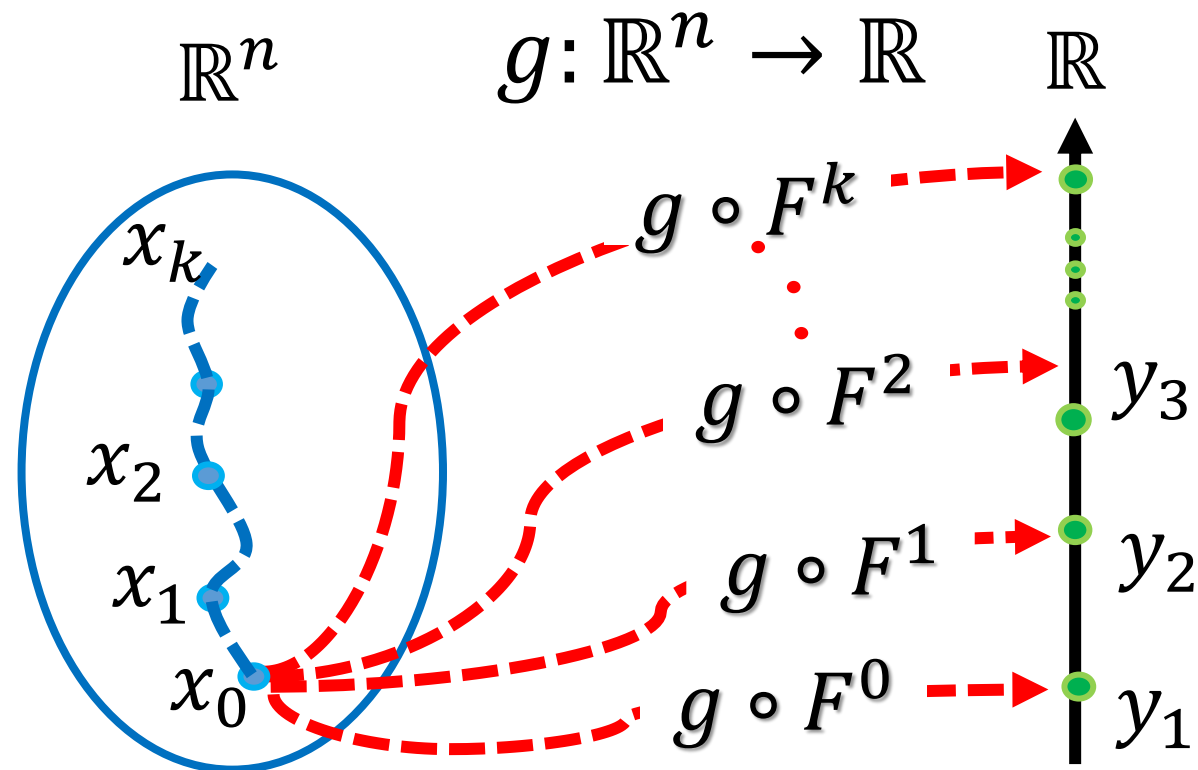
$$x_{t+1} = F(x_t)$$

REALIZATION

LIFTING

$$\mathcal{K}^t g = g \circ F$$

$$g_{t+1} = \mathcal{K}^t g_t = g \circ F = g(x_{t+1})$$



State space (Finite dim. Nonlinear dynamics) $\xrightarrow{\text{LIFTING}}$ Observable space (Infinite dim. linear evolution)

□ Linear operator \longrightarrow (Spectral Analysis)

- Lifts the nonlinear evolution of state into a linear evolution of functional observable
- Linearity: $\mathcal{K}^t(\alpha_1 g_1 + \alpha_2 g_2) = \alpha_1 \mathcal{K}^t g_1 + \alpha_2 \mathcal{K}^t g_2$
 $\forall \alpha_1, \alpha_2 \in \mathbb{R}, g_1, g_2 \in \mathcal{O}$

Spectral Analysis: Koopman & Dynamic Mode Decomposition

Dynamic mode decomposition (DMD):

$$y_t = \sum_{j=1}^N \phi_j \lambda_j^t b_j \xrightarrow{\text{SVD}} y_t = \sum_{j=1}^r \phi_j \lambda_j^t b_j$$

$A = Y(Y^+)^{\dagger}$ Approx.

Refs.) Kutz et. al, SIAM'16; Mauroy, Mezic & Susuki, Springer'20

Koopman mode decomposition (KMD):

$$[\mathcal{K}^t g](x_0) = g \circ \mathbf{F}(x) = \sum_j \lambda_j \varphi_j(x) \mathbf{v}_j$$

Point spectra

Eigenpairs: $\{\lambda_j, \varphi_j\}_j^{\infty} \quad \{g_1, g_2, \dots\}$

Finite dim. approx:

$$x_i \in \text{span}\{g_1, g_2, \dots, g_M, \dots, g_N\}$$

state realizations \approx a linear combination of observables

➤ **Least Square Error Optimization (DMD)**

The diagram illustrates the Least Square Error Optimization (DMD) process. It shows the decomposition of the Data matrix $Y \in \mathbb{R}^{n \times N}$ into three components: Spatial Modes Φ , Amplitudes \mathbf{D}_b , and Temporal Modes \mathbf{V}_{and} . The spatial axis is labeled 'spatial' and the temporal axis is labeled 'temporal'. The decomposition is represented as:

$$\begin{bmatrix} | & \dots & | \\ y_1 & \dots & y_N \\ | & \dots & | \end{bmatrix} \approx \begin{bmatrix} | & \dots & | \\ \phi_1 & \dots & \phi_r \\ | & \dots & | \end{bmatrix} \begin{bmatrix} b_1 & & \\ & \ddots & \\ & & b_r \end{bmatrix} \begin{bmatrix} 1 & \lambda_1 & \dots & \lambda_1^{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_r & \dots & \lambda_r^{N-1} \end{bmatrix}$$

The matrices are labeled $Y \in \mathbb{R}^{n \times N}$, Φ , \mathbf{D}_b , and \mathbf{V}_{and} respectively.

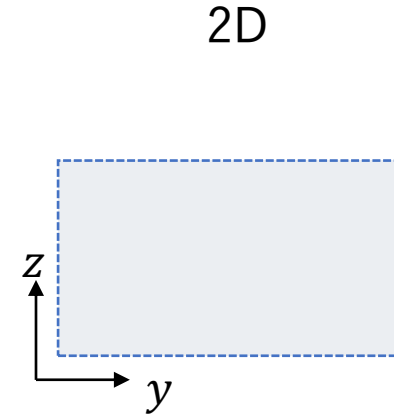
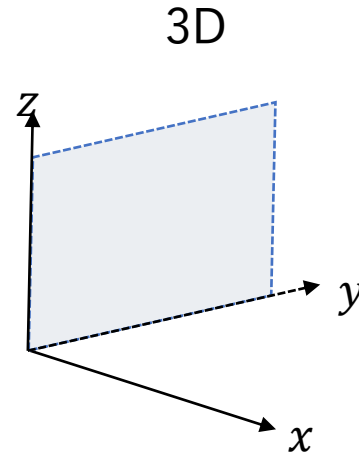
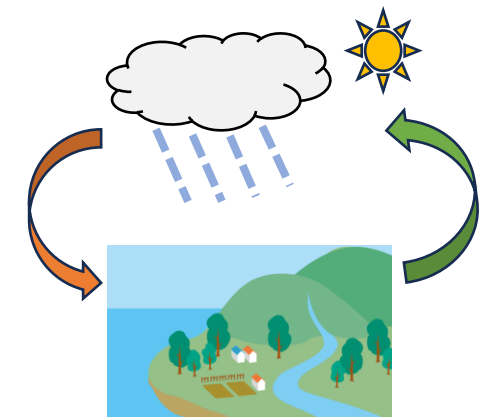
Penalize an ℓ_1 norm on the amplitudes

➤ (Regularized) Least Square Error Optimization
Sparsity-Promoting DMD

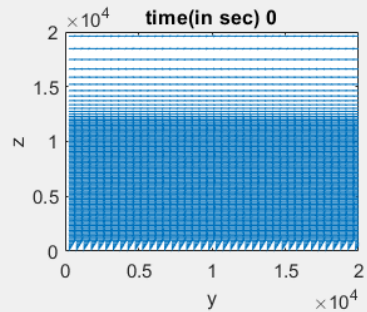
$$J_r(\mathbf{b}) = \min_{\mathbf{b}} ||Y - \Phi \mathbf{D}_b \mathbf{V}_{\text{and}}||_2 + \gamma ||\mathbf{b}||_1$$

Ref.) Jovanovic et. al, Physics of Fluids'14

Model Configuration: SCALE Weather Simulation



- horizontal velocity component in the y-direction (**data.V**)
- vertical velocity component in the z-direction (**data.W**)



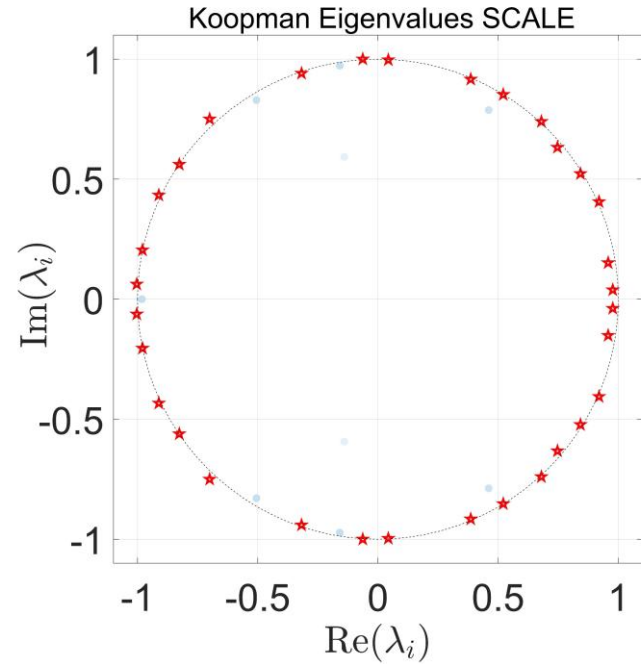
SCALE Weather dynamics Simulation

- Time unit: 30 seconds per interval/iteration
- Original data :
121 snapshots & $40 * 97 = 3380$ grids

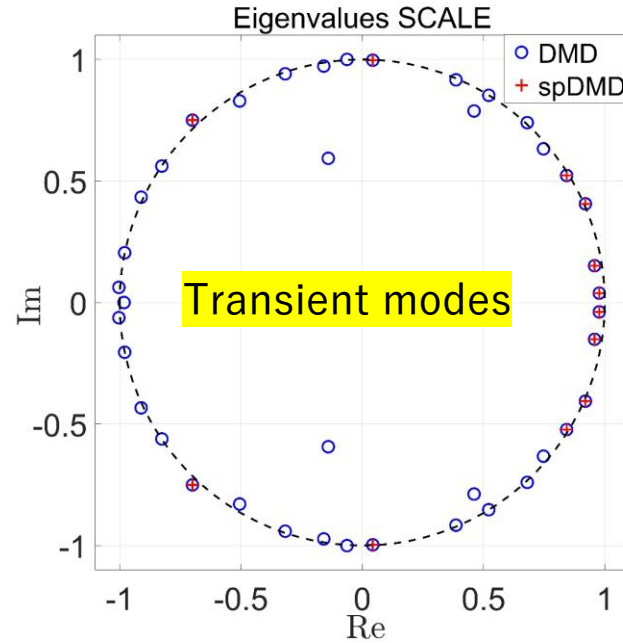
Other Atmospheric factors:

- pressure (**data.PRES**) and temperature (**data.T**)
- relative humidity (**data.QV**) and precipitation (**data.PREC**)

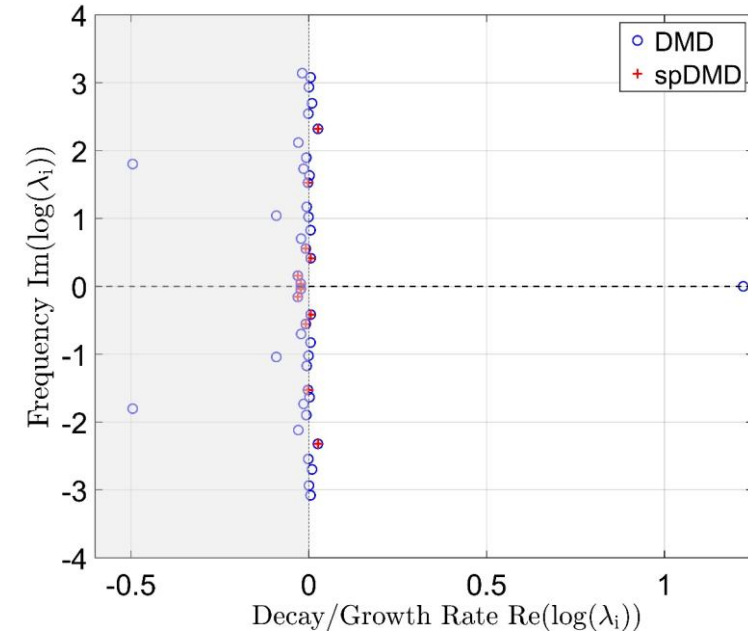
Experiment I: Magnitude of Vorticity Field



Extract 32 modes

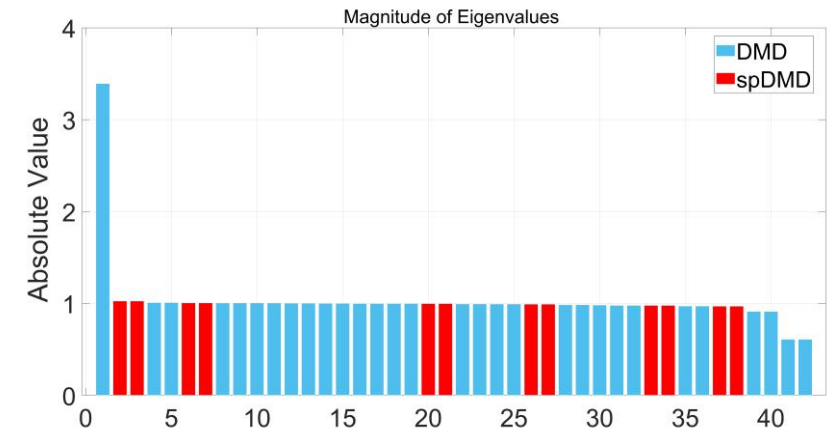


Extract 12 modes

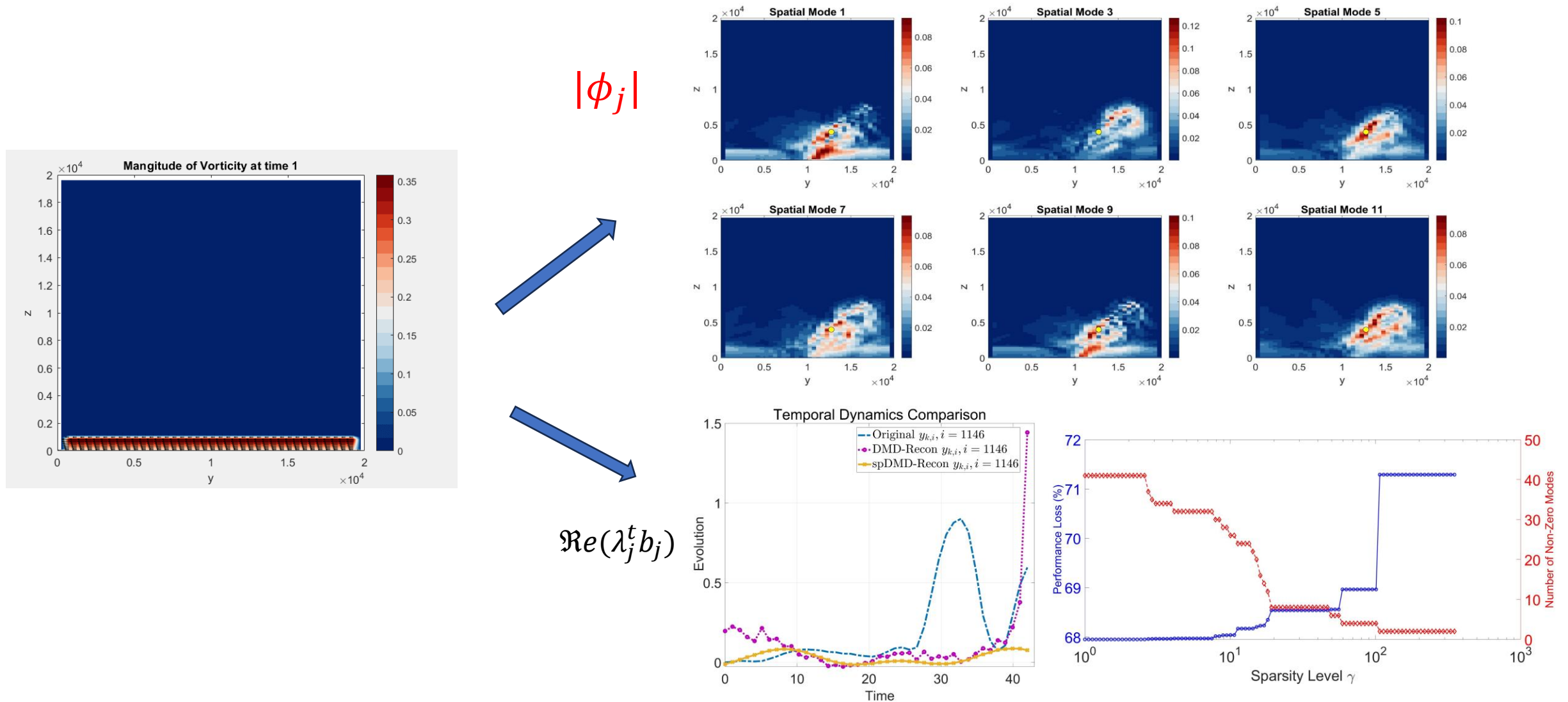


Check *growth/decay rate*
w.r.t. eigenvalues

- Vorticity Data: $\mathbf{Y} \in \mathbb{R}^{3880 \times 43}$
- Only takes the first 43 snapshots and has $40 * 97 = 3380$ grids
- Generate a **tall & short** matrix
- Rich spatial knowledge while less temporal information



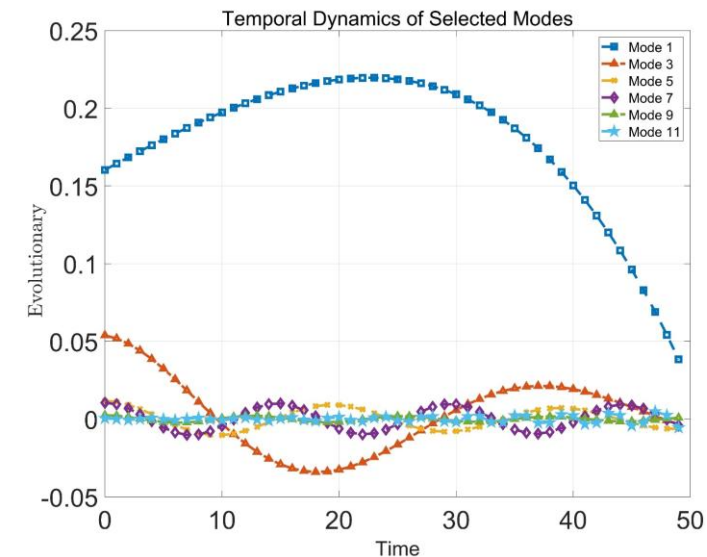
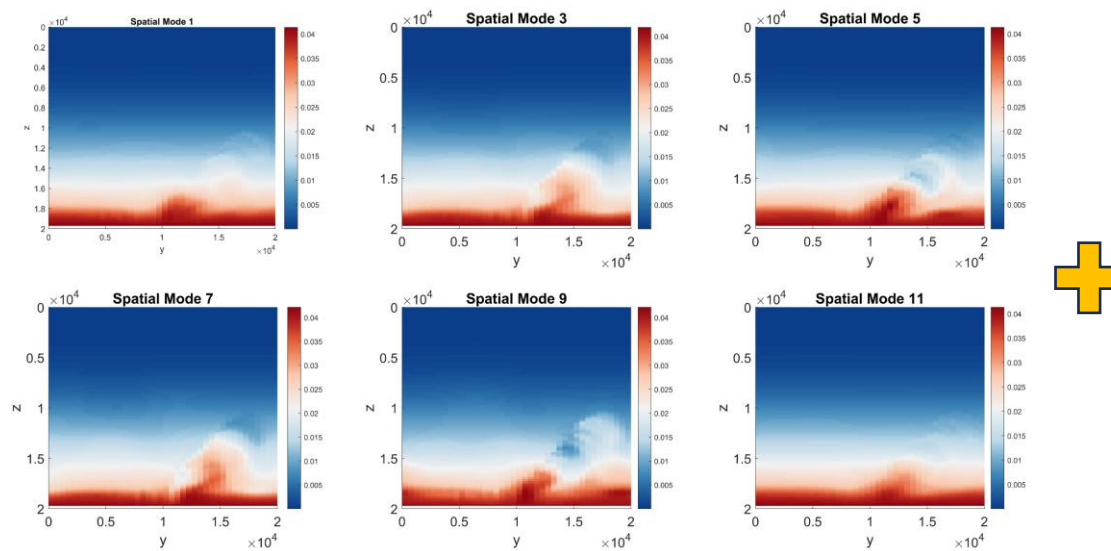
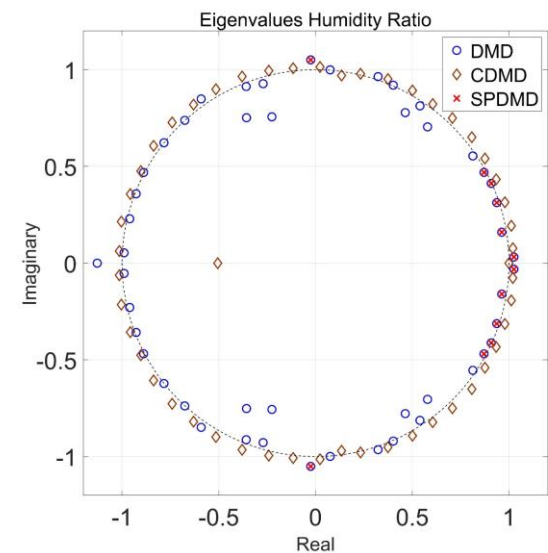
Experiment I: Magnitude of Vorticity Field



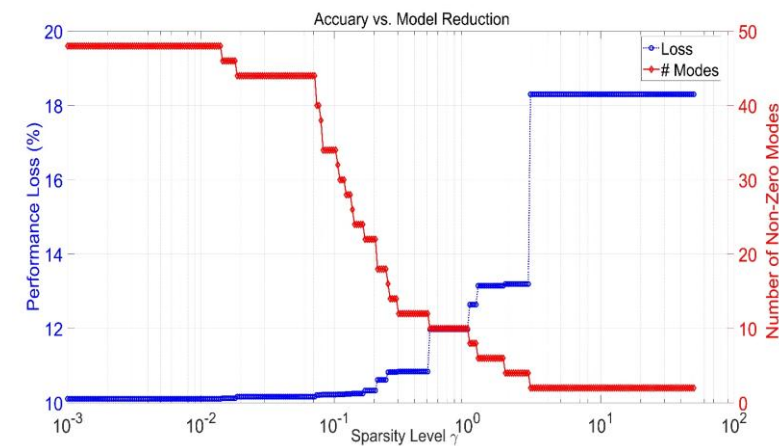
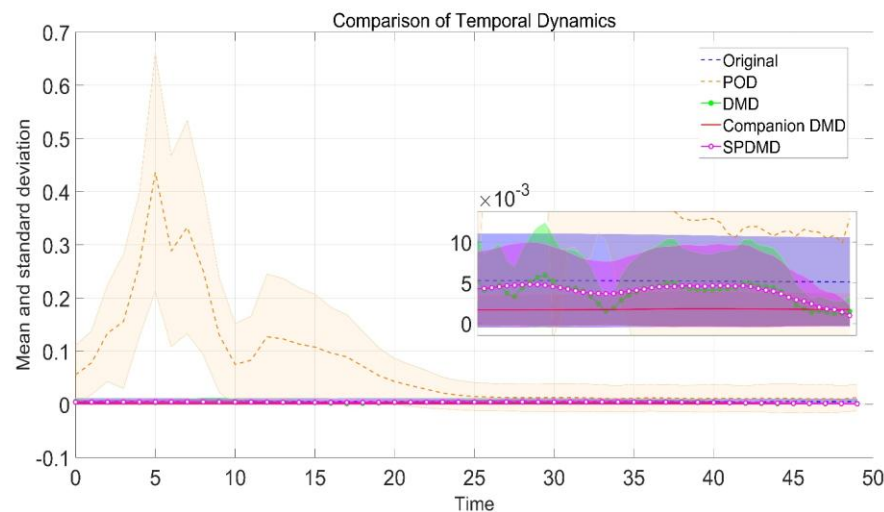
- ❑ DMD generates 42 modes and SPDMD extracts the 6 pairs dominant spatial modes by taking a sparsity level
- ❑ However, SPDMD can not directly capture the transient evolution of *temporal dynamics* (vs. original data)

Experiment II: Humidity Ratio (Relative Humidity) Field

Goal: Extract bubble-like behavior in dynamics (i.e., transient modes)



- ❖ Humidity ratio: $Y \in \mathbb{R}^{3880 \times 51}$
- ❖ Various DMD algorithms (Companion DMD, SPDMD)
- ❖ Extract 6 pairs leading modes



➤ Balancing accuracy and no. of reduced modes

Conclusion:

- ✓ Koopman analysis for SCLAE simulation data by using DMD algorithms to identify the dominant (spatial & temporal) modes and realize the model reduction.
- ✓ Vorticity magnitude data, humidity ratio data fields of SCALE simulations (i.e., *different observable function spaces*) are explored, and analyze the **point spectrum** of the linearized dynamics and find the corresponding leading modes.
- ✓ Use the key modes to represent a low dim. system.

Next...

- ❑ Generate more realistic SACLE simulation data (3D) and explore Koopman modes.
- ❑ Ensemble initial conditions: Koopman operator on **multiple trajectories**.
- ❑ Design a controller to predict the *precipitation* with a low dim. System.

Thank you for your listening !
Q&A