Data-driven identification of geophysical dynamics: incorporating stability constraints in neural networks models

Said Ouala[1], Ronan Fablet [1], Bertrand Chapron[2], Ananda Pascual[3], Fabrice Collard[4], Lucile Gaultier[4]
(1) IMT Atlantique, Lab-STICC, Brest, France
(2) Ifremer, LOPS, Brest, France
(3) IMEDEA, UIB-CSIC, Esporles, Spain
(4) ODL, Brest, France

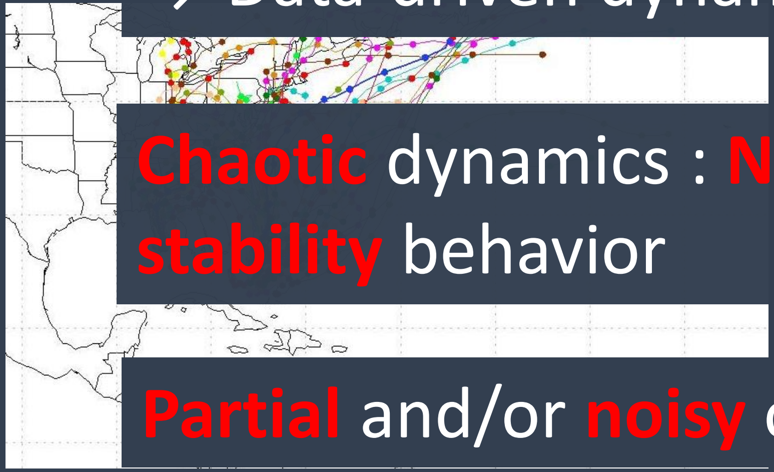
Problem statement

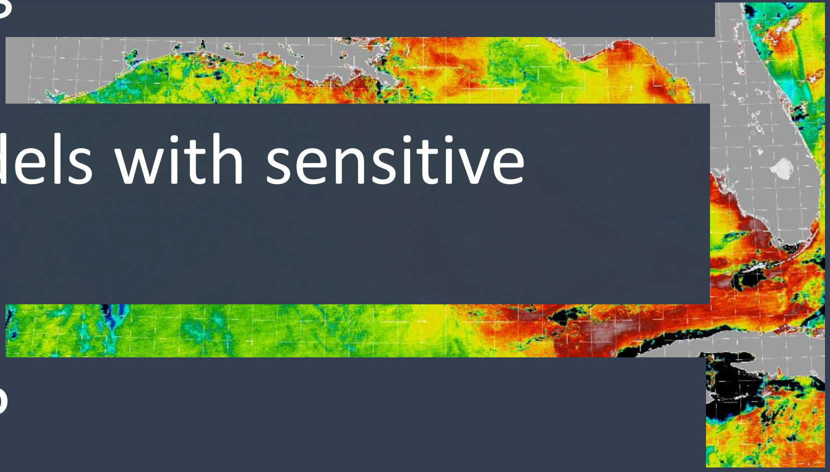
General context : **Modeling** ocean dynamics

Forecasting and **Reconstruction** Applications

Often equations are **too complicated** or **unknown**

→ Data-driven dynamical modeling





Chaotic dynamics : **Non linear** models with sensitive **stability** behavior

Partial and/or **noisy** observations ?

My thesis as an optimization problem

Optimize

- **How to derive a data driven model from a collection of observations ?**
- **Which data driven formulation to use ?**
- **How to evaluate the performances of a data driven model**
- **How to use our data driven model in reconstruction applications**
- **What do we want to forecast/reconstruct : small scale structures <100km or large scale tendency ?**

Subject to

- **Low forecasting cost**
- **Applyability in data assimilation**
- **Model Interpretability**
- **Long term simulation**

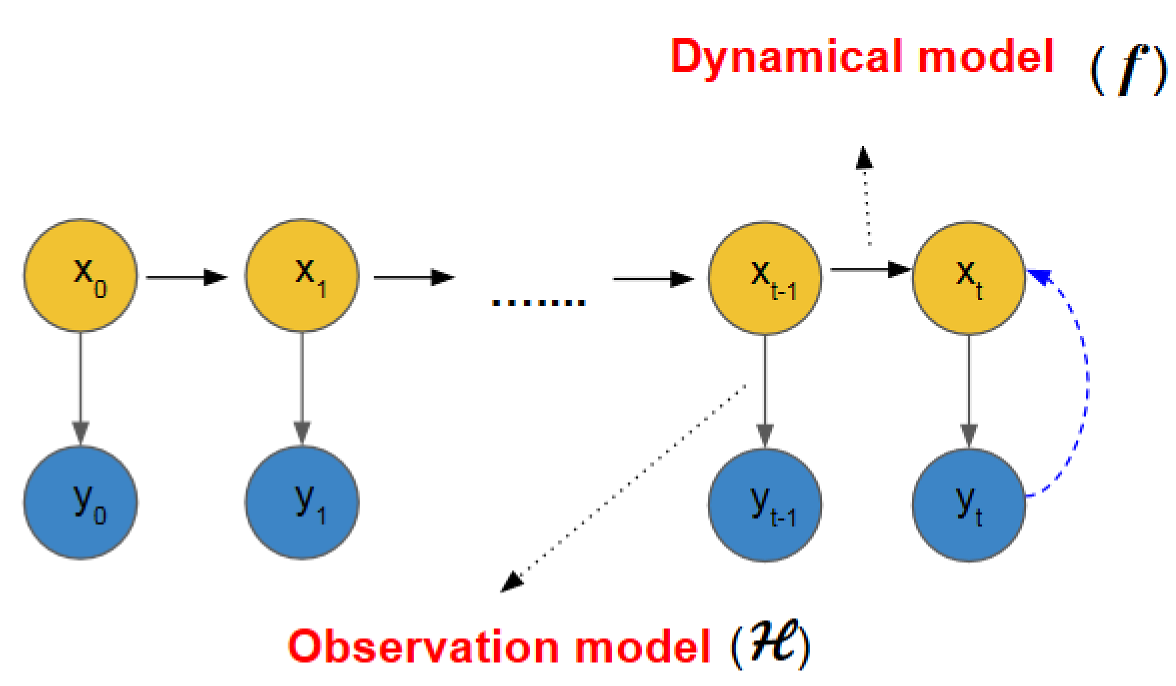
Problem statement
General Framework : State space formulation

X(t) : Hidden states
Y(t) : Observations (noisy and/or sparse)

$$\mathbf{x}_{t+\delta} = f(\mathbf{x}_t) + \omega_{t+\delta}$$
$$\mathbf{y}_t = \Phi_t(\mathcal{H}(\mathbf{x}_t) + \epsilon_t)$$

True State

Observation



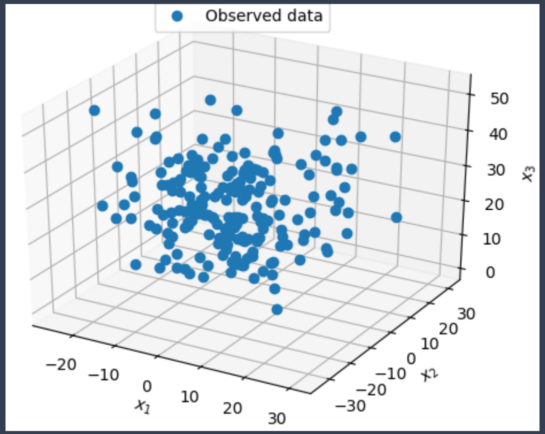
Problem statement
General Framework : State space formulation

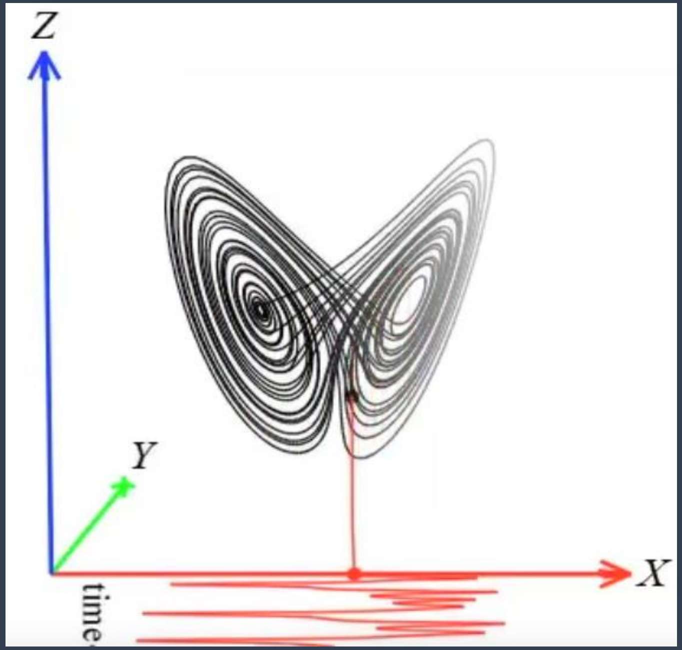
$$\mathbf{x}_{t+\delta} = f(\mathbf{x}_t) + \omega_{t+\delta}$$

$$\mathbf{y}_t = \Phi_t(\mathcal{H}(\mathbf{x}_t) + \epsilon_t)$$

Deterministic or stochastic models

Partial and noisy observations





Outline

- Problem statement
- Data driven model identification from a sequence of observations
  - Direct observations of the state space
  - Noisy and partial observations
  - Temporally sparse data
  - Partially observed systems
- Discussion
- Perspectives

Outline

- Problem statement
- Data driven model identification from a sequence of observations
  - Direct observations of the state space
  - Noisy and partial observations
  - Temporally sparse data
  - Partially observed systems
- Discussion
- Perspectives

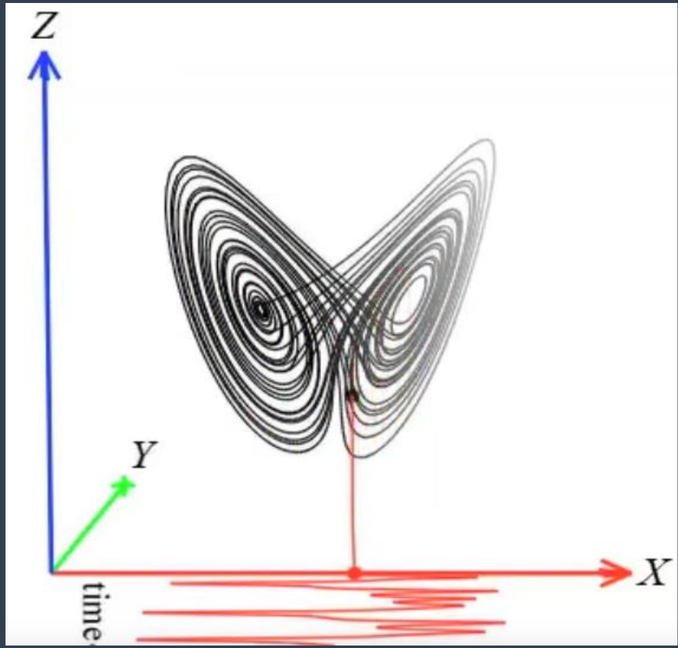
Partially observed systems

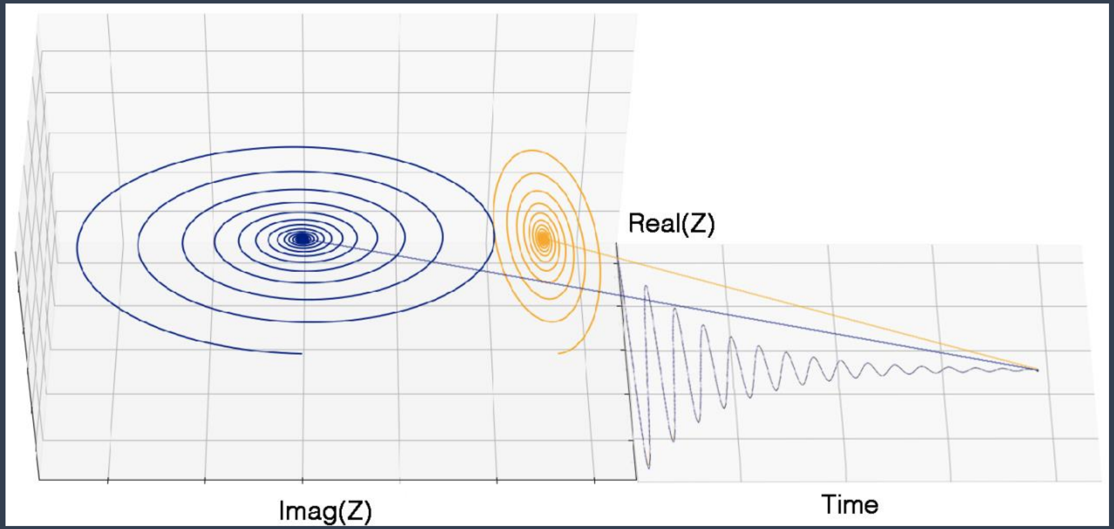
$$\mathbf{x}_{t+\delta} = f(\mathbf{x}_t) + \omega_{t+\delta}$$

$$\mathbf{y}_t = \Phi_t(\mathcal{H}(\mathbf{x}_t) + \epsilon_t)$$

Unobservable Components

No Noise





UNOBSERVABLE COMPONENTS

Motivation

Sea surface tracers evolution depends on a lot of unseen variables (subsurface variables, fine-scale sea structures ...etc.).

The EOF decomposition gives only orthogonal modes, the model may depend on non-orthogonal variables.

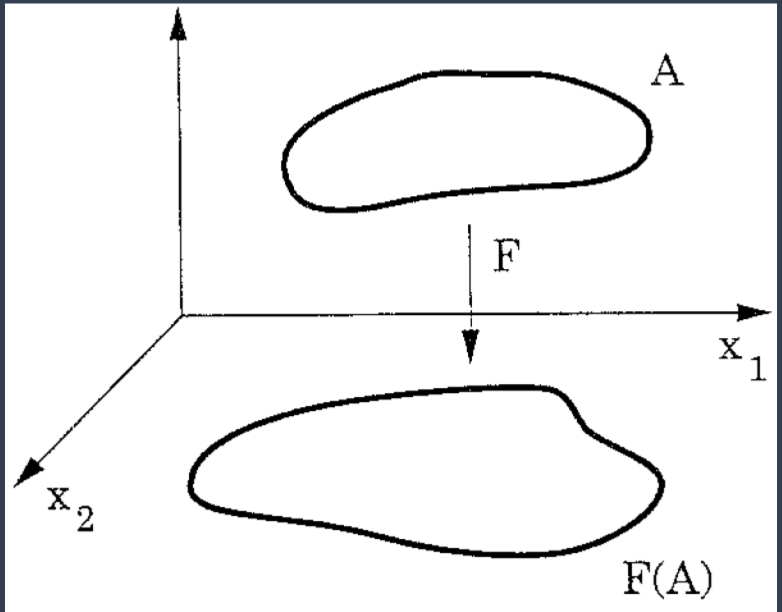
From a dynamical system POV, we only observe few components of a higher dimensional governing equation.

ODE in the observation space => the observations are an embedding of the higher dimensionnal space

UNOBSERVABLE COMPONENTS

Motivation

An observation operator is an **embedding** of the **hidden state** space if :



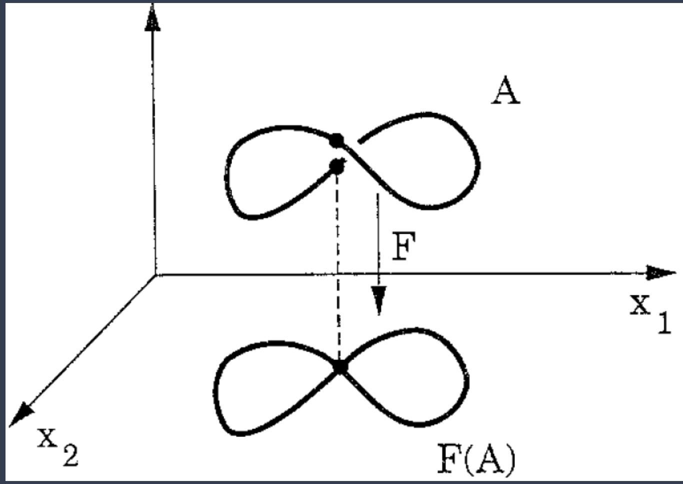
Sauer et al. 1991

The mapping doesn't collapse points (**one-to-one**) or tangent directions (**immersion**)

UNOBSERVABLE COMPONENTS

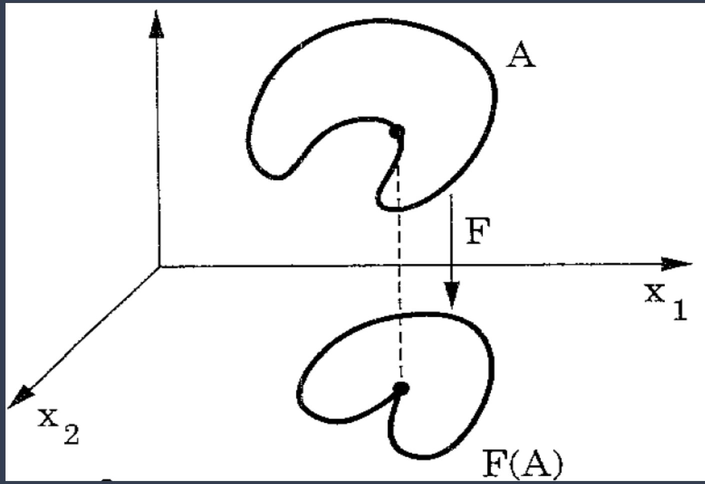
Motivation

An observation operator is an **embedding** of the **hidden state** space if :



Sauer et al. 1991

The mapping is an **immersion** but fails to be **one-to-one**

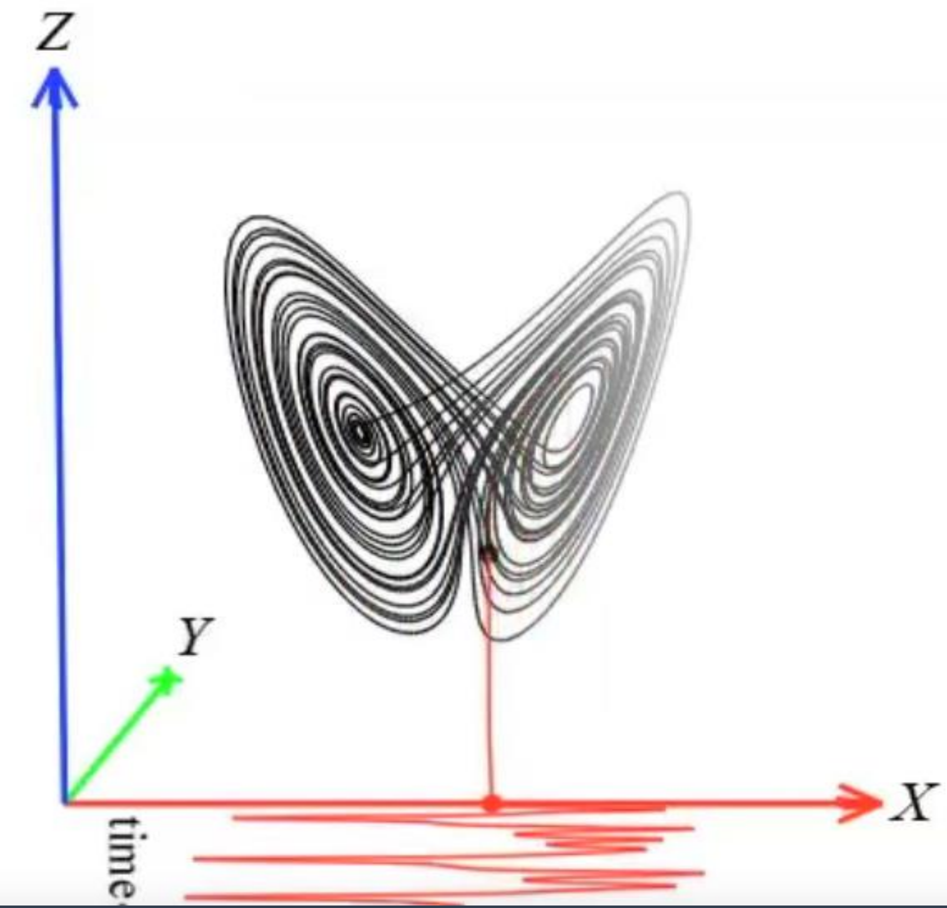


Sauer et al. 1991

The mapping is **one-to-one** but fails to be an **immersion**

UNOBSERVABLE COMPONENTS

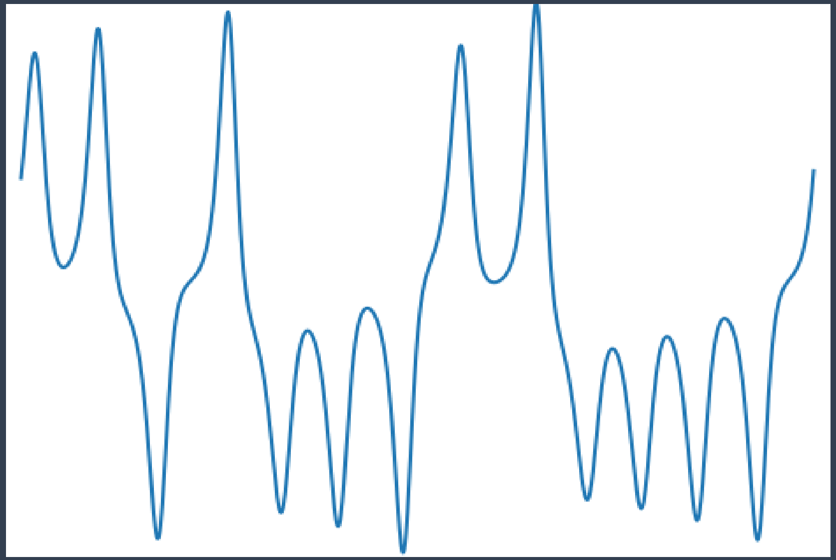
Motivation



What happens when the observation operator doesn't form an **Embedding** of the **hidden state** space ?

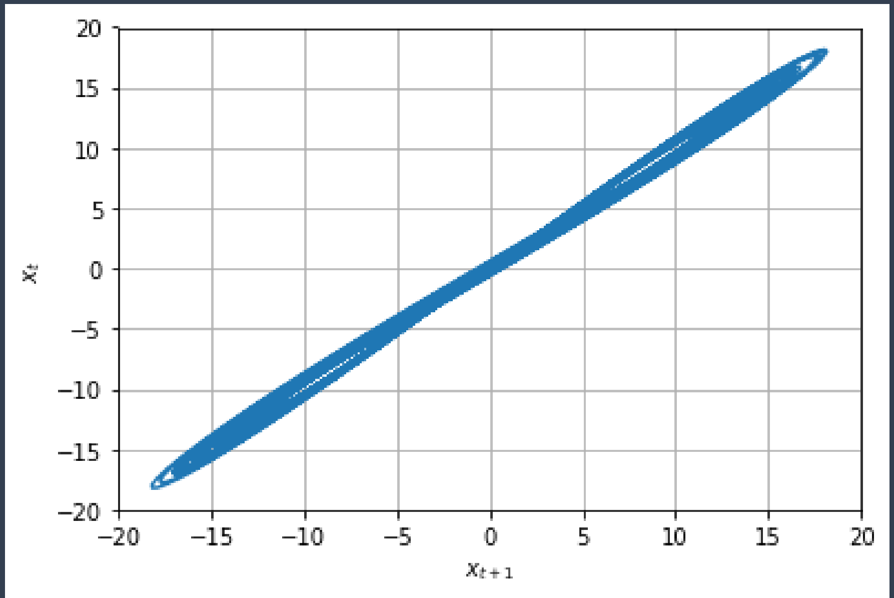
UNOBSERVABLE COMPONENTS : Issues with classical approach

Classical approach : $dx/dt = f(x)$?



UNOBSERVABLE COMPONENTS : Issues with classical approach

Data are not



Classical approach : dx/dt = f(x) ?

The associated mapping $\varphi_t(x(0)) = \int_{t_0}^{t} f(x(t'))dt'$ is one-to-one
$\varphi_t(x(0)) = \varphi_t(x(u)) \Rightarrow x(0) = x(u)$

UNOBSERVABLE COMPONENTS : Conclusion

In several Real world problems, we are not (at all) provided with some components directly or indirectly influencing the variability of our underlying system

We need to find a higher dimensional space where our data will be an embedding of the true states

Synergie studie then include several other observations and hope for the best

Physically guided data driven : find the most appropriate manifold in a higher dimension that fits the best our obs

Find an embedded representation :

Our approach : Project our observation x into a high dimensional space u, with $u = [x, l1, l2, \dots, ln]$

Fit :

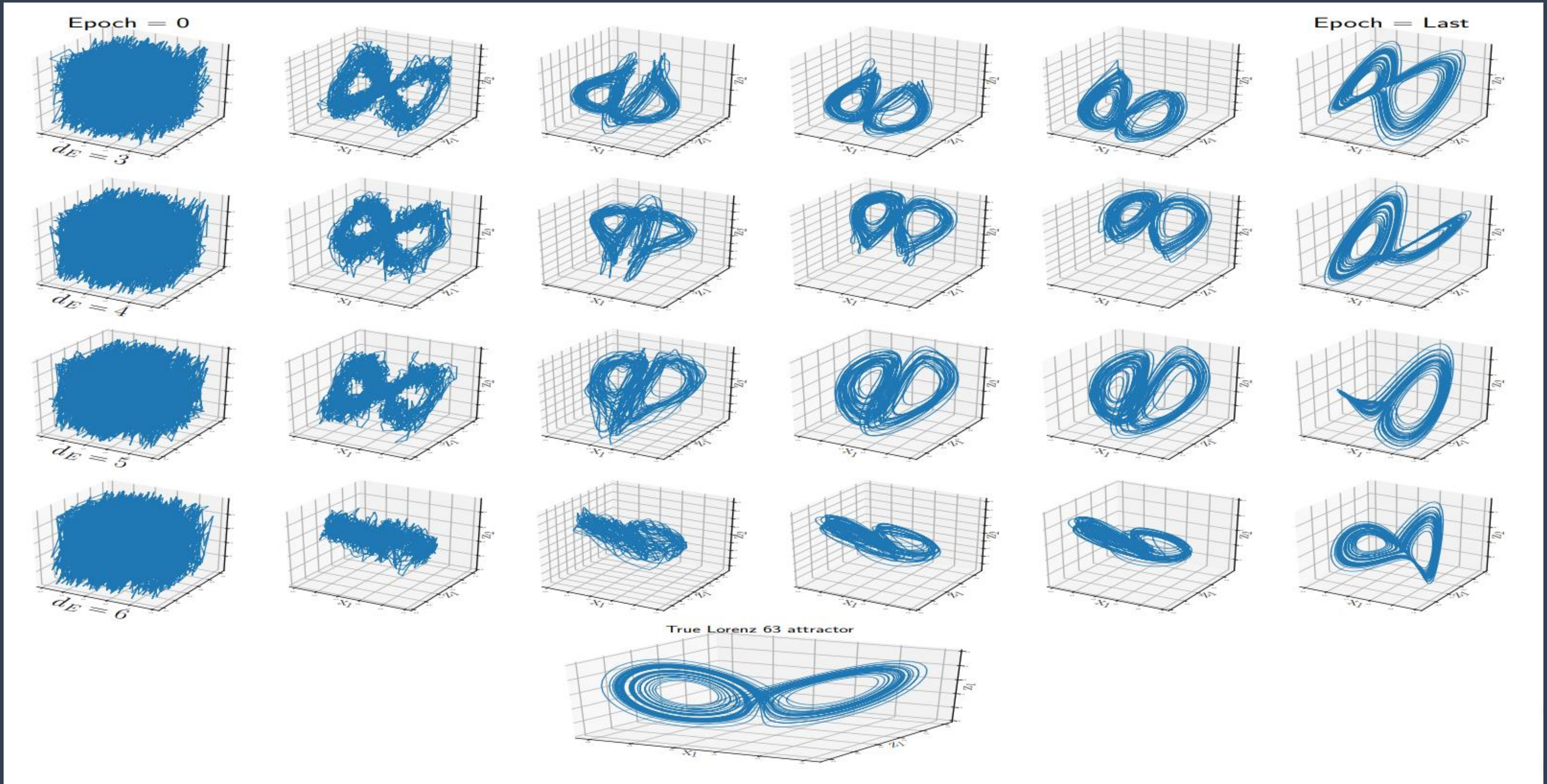
$$\frac{du}{dt} = f_\theta(u)$$

With

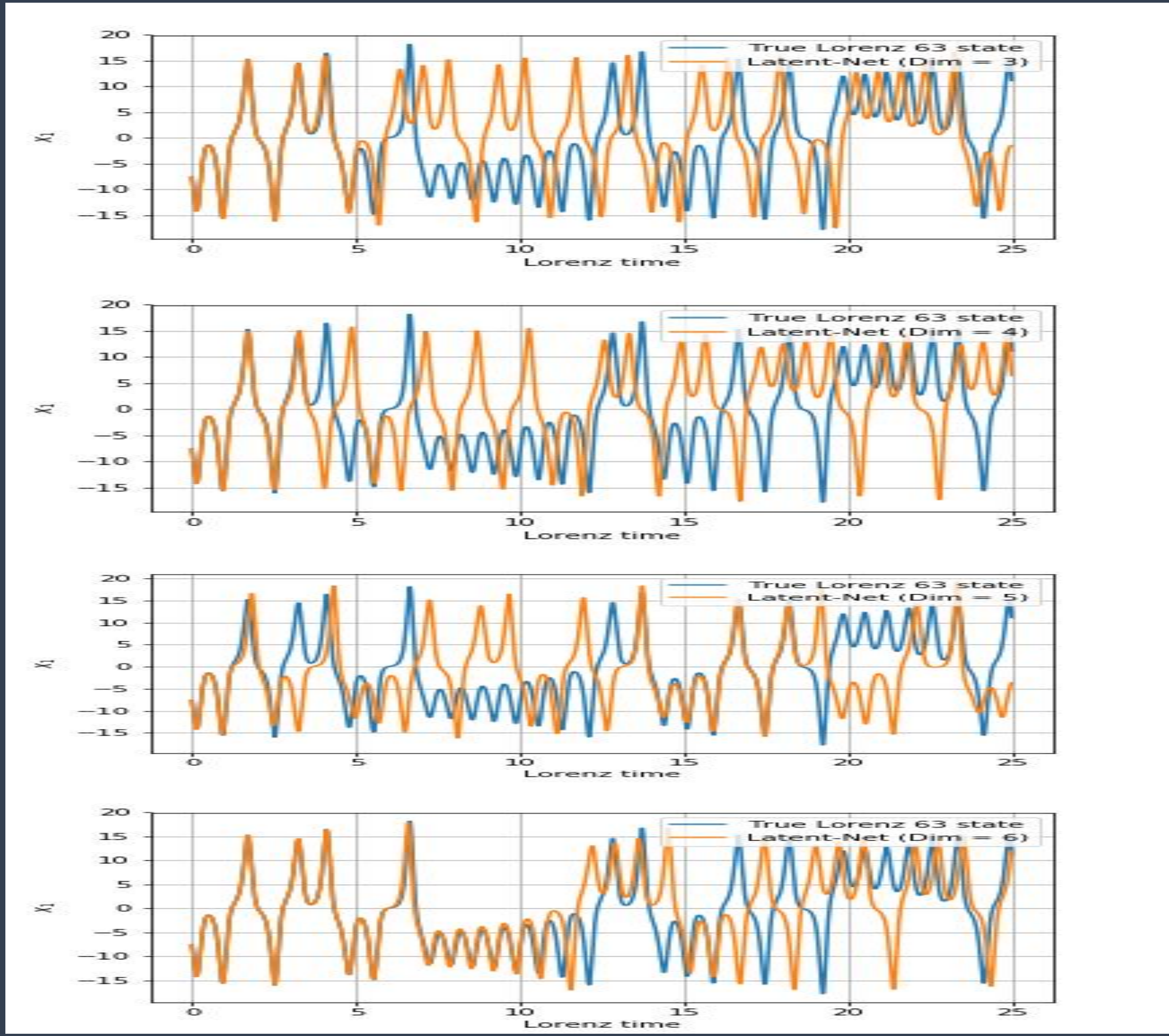
$$u = [x, l1, l2, \dots, ln]$$

$$\theta, l = argmin_{\theta,l}\{\alpha|x(t) \ - G(\int_{t-1}^{t} \ f(u(t'))dt')| + (1-\alpha)|u(t) - \int_{t-1}^{t} \ f(u(t'))dt' \ |\}$$

Physically guided data driven : State reconstruction



Physically guided data driven : Forecasting



Lyap = 0,82

Lyap = 0,96

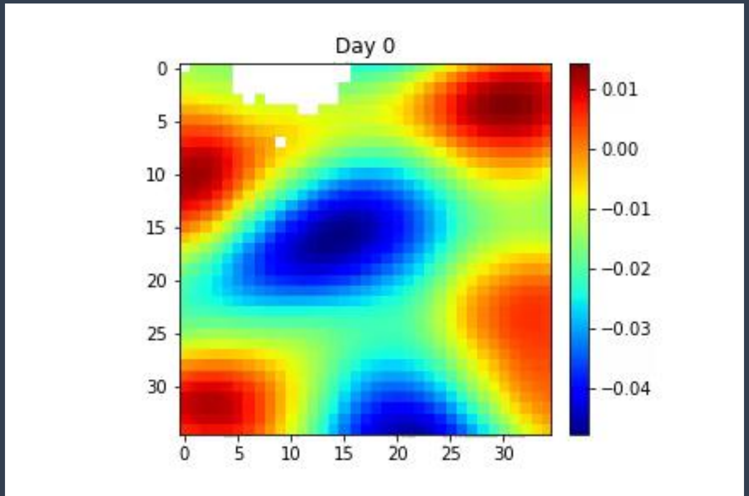
Lyap = 0,82

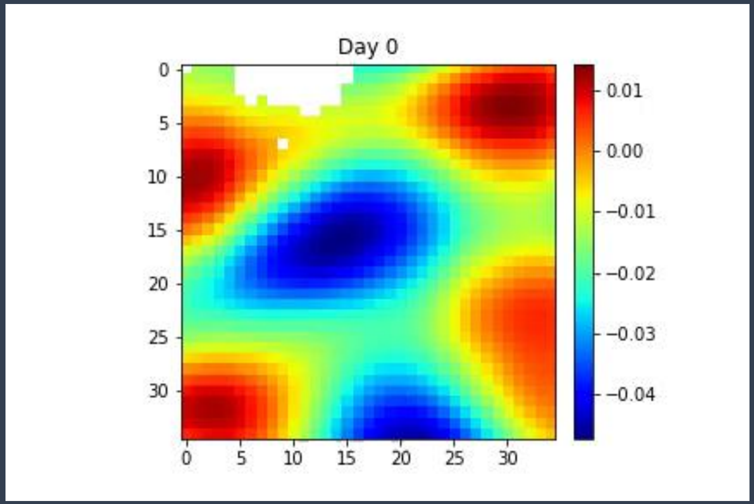
Lyap = 0,87

Forecasting SLA :

- OSSE based on realistic high-resolution ocean simulation data in the Western Mediterranean sea from WMOP configuration (Juza et al. 2016).
- The Spatial resolution of 0.05 and a temporal resolution dt=1 day.
- The data from January 2009 to December 2014 were used as training and we tested our approach on the first 347 days of the year 2015.
- The EOF space dimension n=15, which amounts to capture 95% of the total variance.

Forecasting SLA realistic forecast up to ~ 170 days !

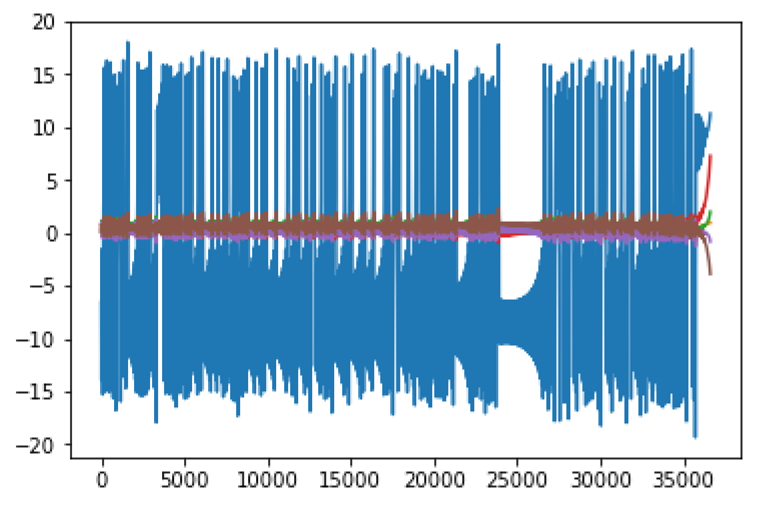


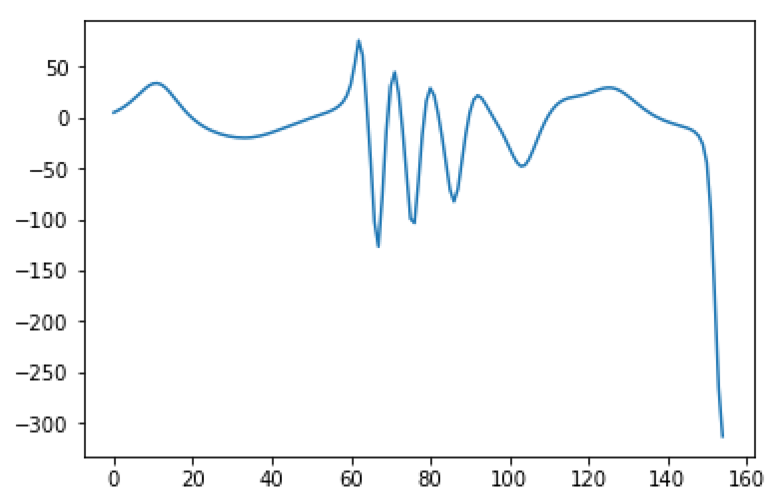


UNOBSERVABLE COMPONENTS

Issue

Our spawned manifold is not dense in the phase space :/





UNOBSERVABLE COMPONENTS

Idea

Boundedness constraints : Constraint the trajectories of the dynamical system to live in a closed ball in the phase space

- The dynamical system is by construction stable (will not diverge)
- The training data will (hopfully) make the model converge to the true limit cycle

UNOBSERVABLE COMPONENTS

Idea

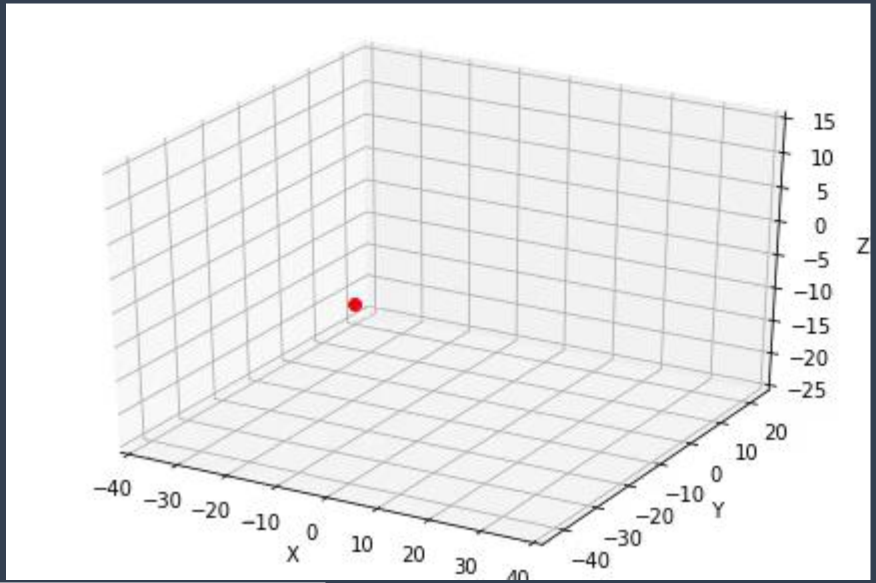
Boundedness constraints : Constraint the trajectories of the dynamical system to live in a closed ball in the phase space

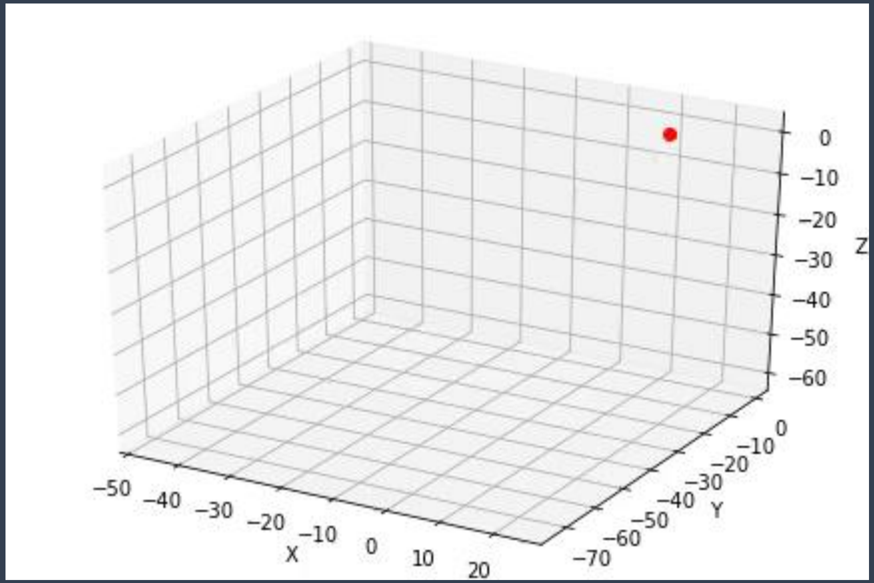
In practice : Energy preserving non linearity + Negative eigenvalues of the linear part of (a shifted version of) the model (Schlegel et al. 2013):

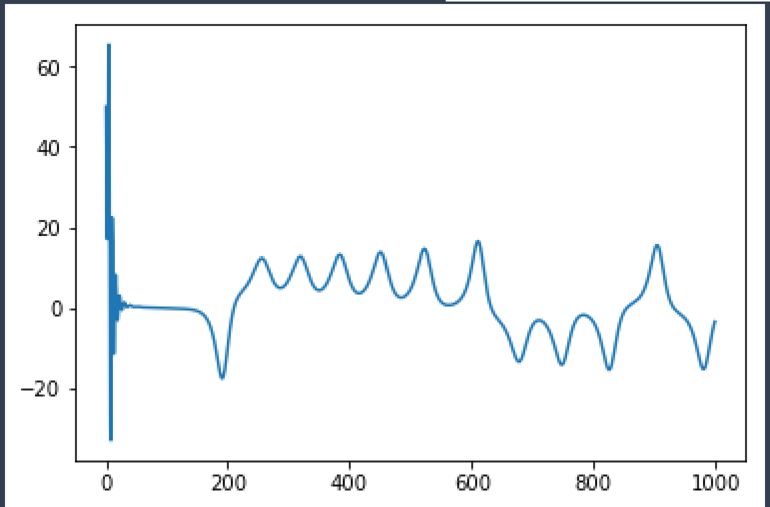
$$\begin{aligned}\hat{\theta}, \mathbf{y}_{1:T} = \arg\min_{\theta} \min_{\{\mathbf{y}_t\}_t} \sum_{t=1}^{T} \|\mathbf{x}_t - G\left(\Phi_{\theta,t}\left(\mathbf{u}_{t-1}\right)\right)\|^2 \\ + \lambda\|\mathbf{u}_t - \Phi_{\theta,t}(\mathbf{u}_{t-1})\|^2 \\ + \lambda_2\|\mathbf{u}_t\mathcal{N}(u_t)\|^2 \\ + \lambda_3\|\mathrm{Relu}(\alpha)/\mathrm{Relu}(\alpha+1)\|^2\end{aligned}$$

UNOBSERVABLE COMPONENTS

L63

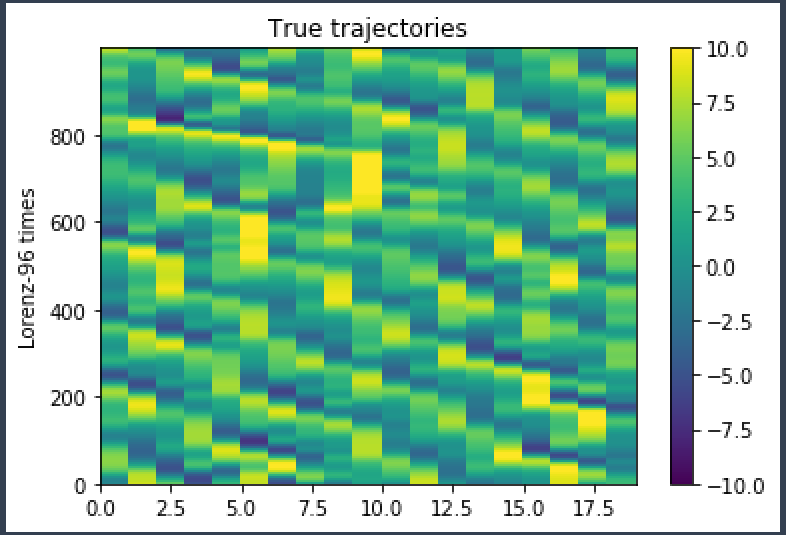


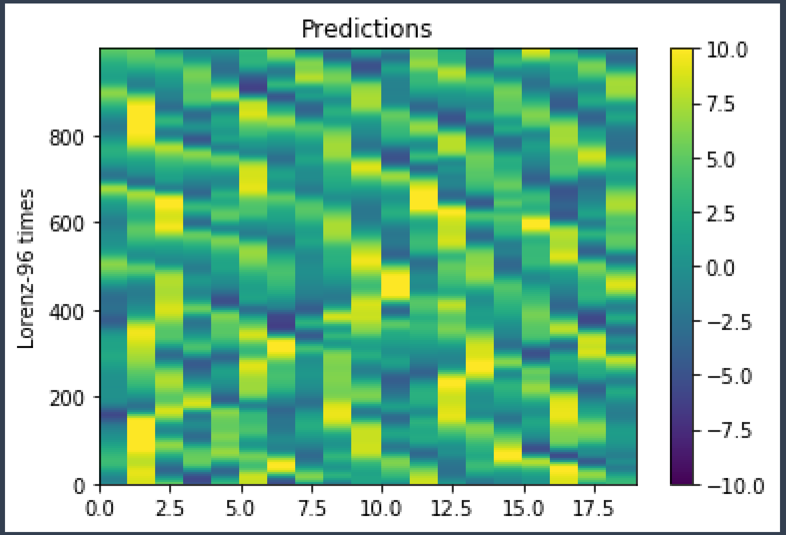


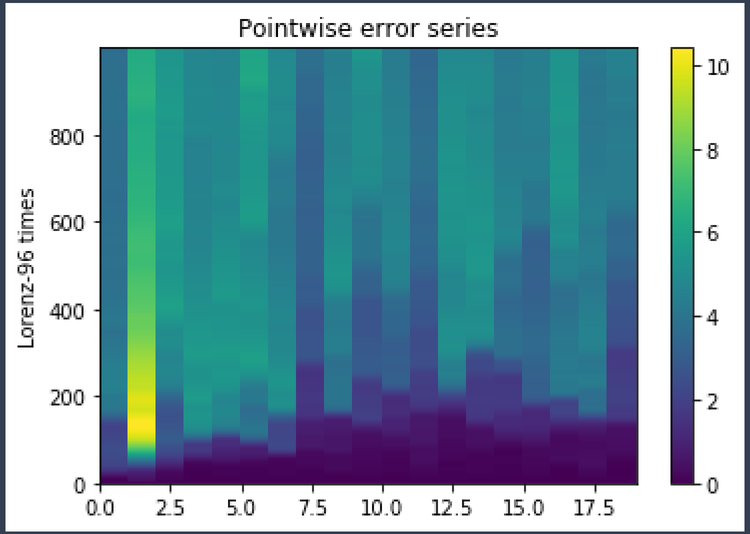


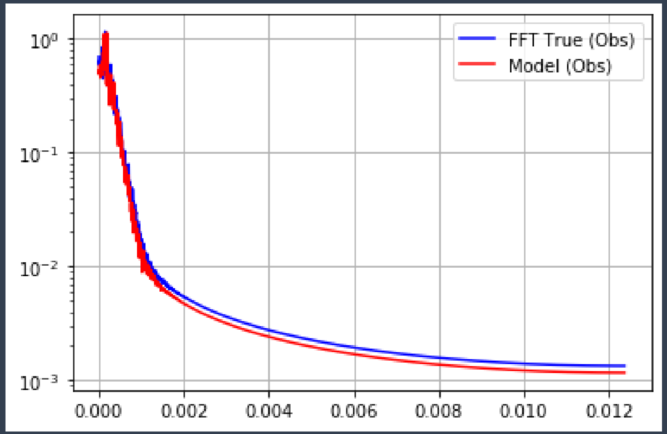
UNOBSERVABLE COMPONENTS

L96





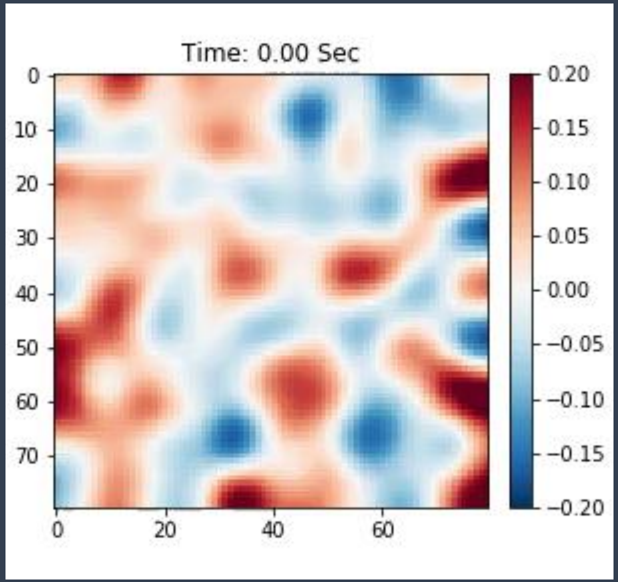




UNOBSERVABLE COMPONENTS

Patch Based Shallow Water Equation

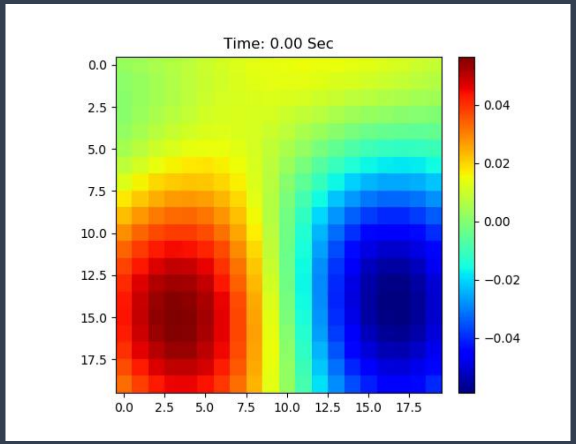
- The length of the domain is set to 1000 km x 1000 km with a corresponding regular discretization of 80 x 80.
- The temporal step size was set to satisfy the Courant–Friedrichs–Lewy condition (dt=20.41 seconds).
- As training data, we took a patch of size 250km x 250km in the center of the 2D domain.
- We use the first 49701 time steps as training data. The training data was projected into an EOF basis with a dimension n=5, which amounts to capture 75% of the total variance.



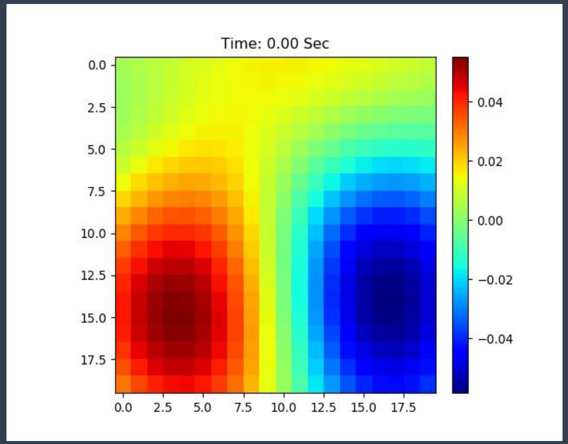
UNOBSERVABLE COMPONENTS

Patch Based Shallow Water Equation

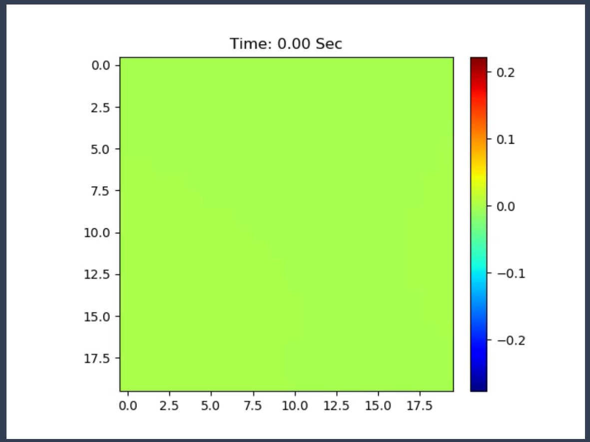
True Shallow water



Model Simulation



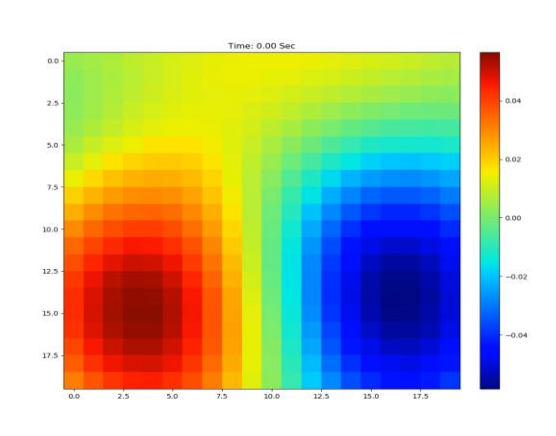
Error (RMSE)



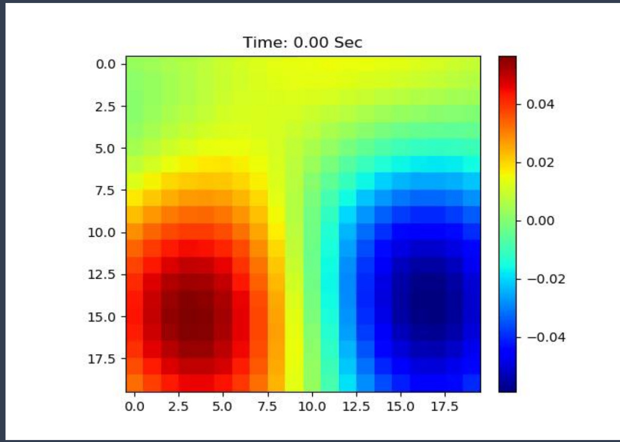
UNOBSERVABLE COMPONENTS

Patch Based Shallow
Water Equation

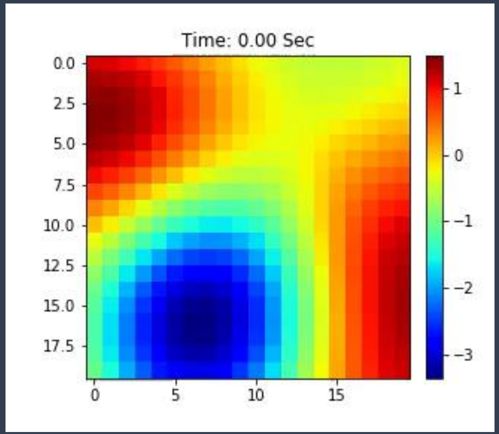
Model Simulation #1



Model Simulation from a perturbed initial condition



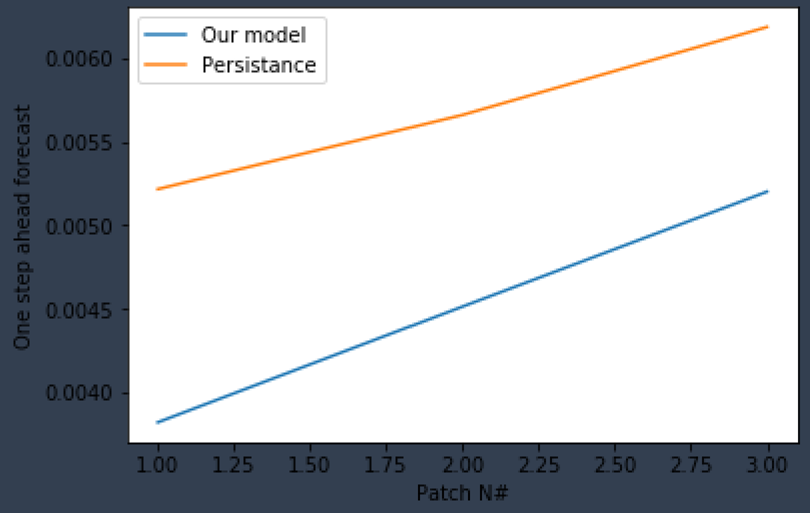
Model Simulation from a far initial condition

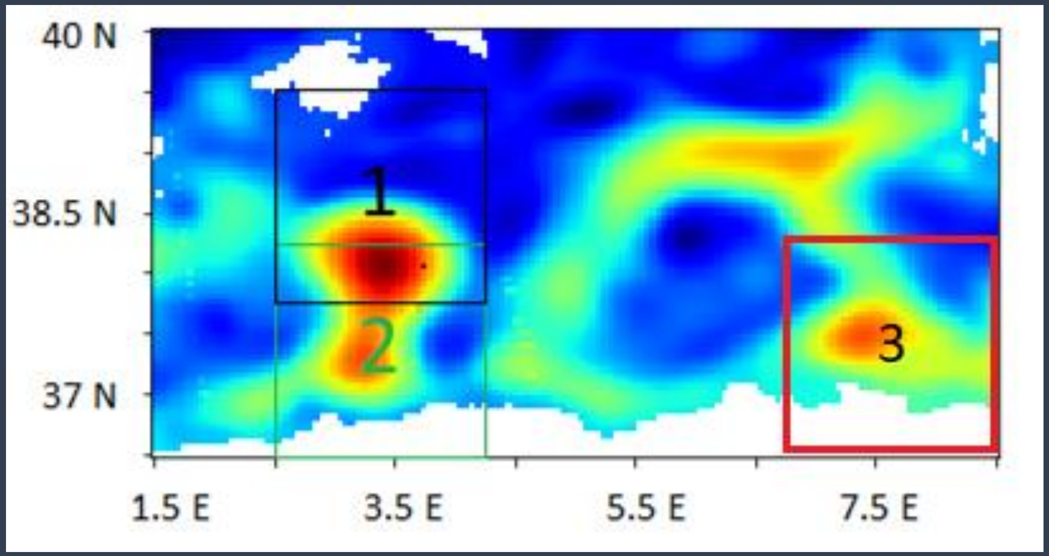


UNOBSERVABLE COMPONENTS

Patch Based Sea Level Anomaly (SLA)

- Patch 1 : Training patch
- Test on test set of patch 1, 2 and 3.

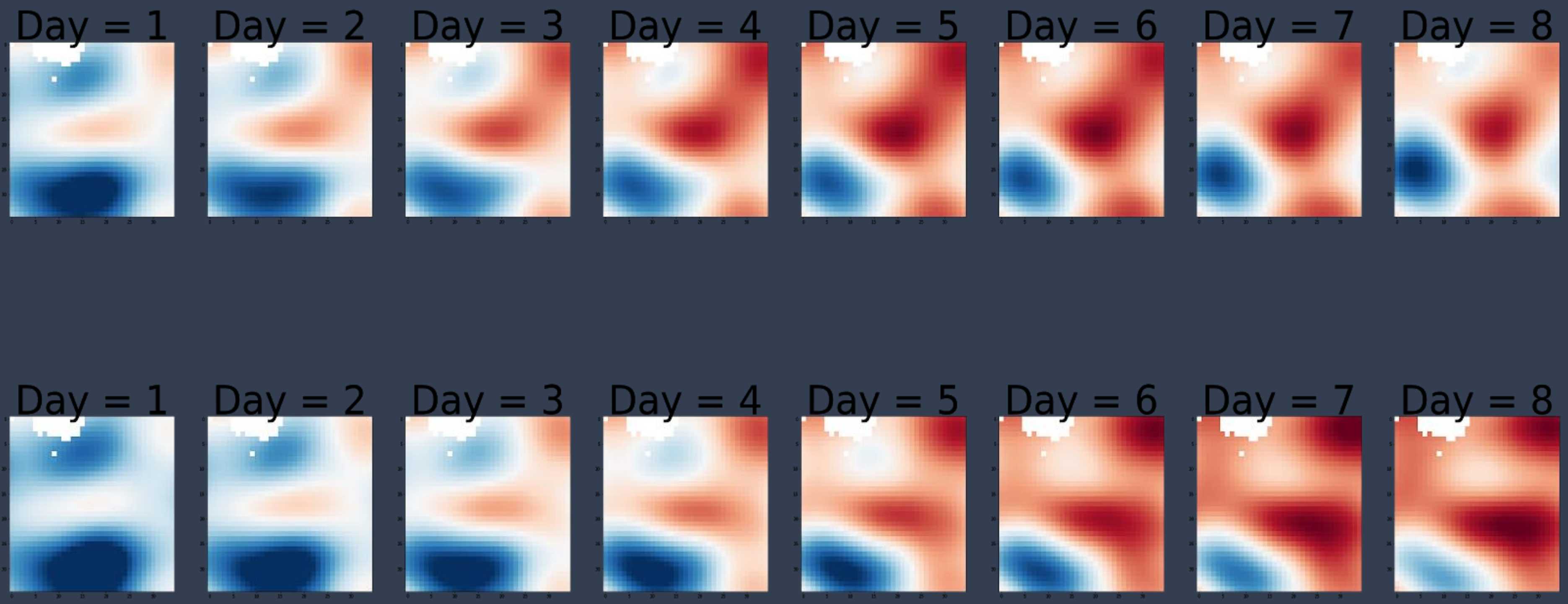




UNOBSERVABLE COMPONENTS

Patch Based Sea Level Anomaly (SLA)

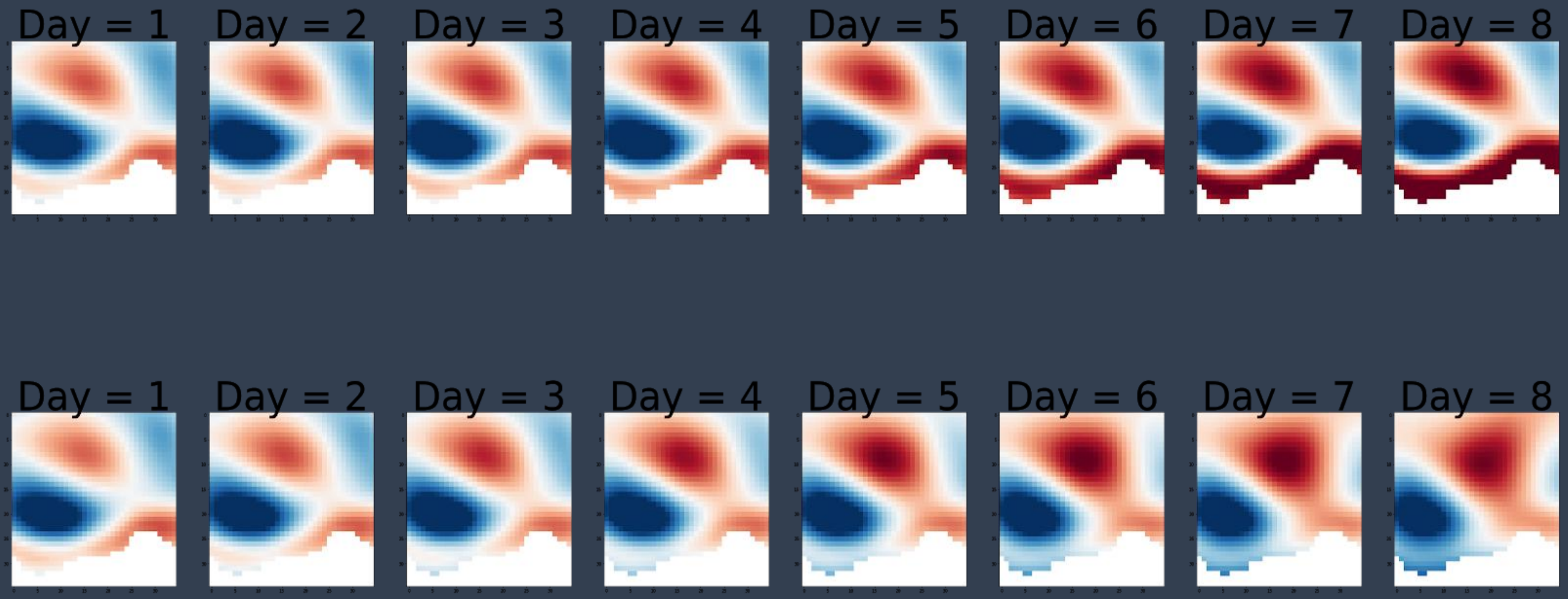
Forecast on test set of patch # 1



UNOBSERVABLE COMPONENTS

Patch Based Sea Level Anomaly (SLA)

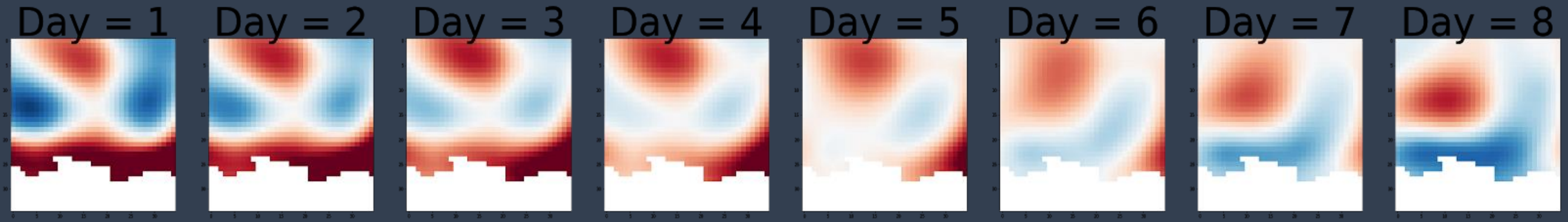
Forecast on test set of patch # 2

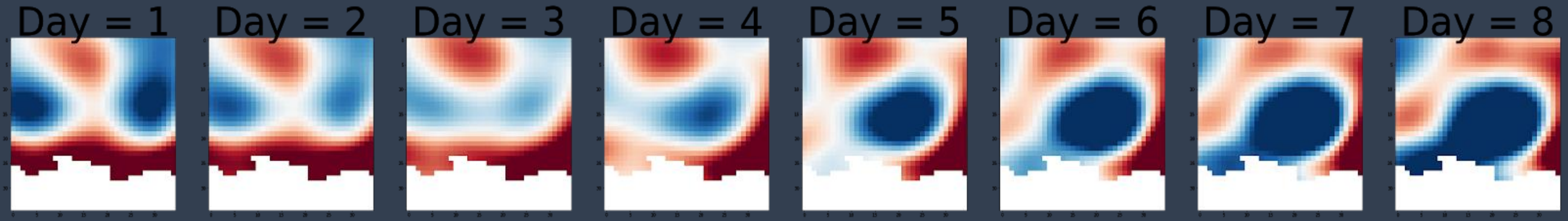


UNOBSERVABLE COMPONENTS

Patch Based Sea Level Anomaly (SLA)

Forecast on test set of patch # 3

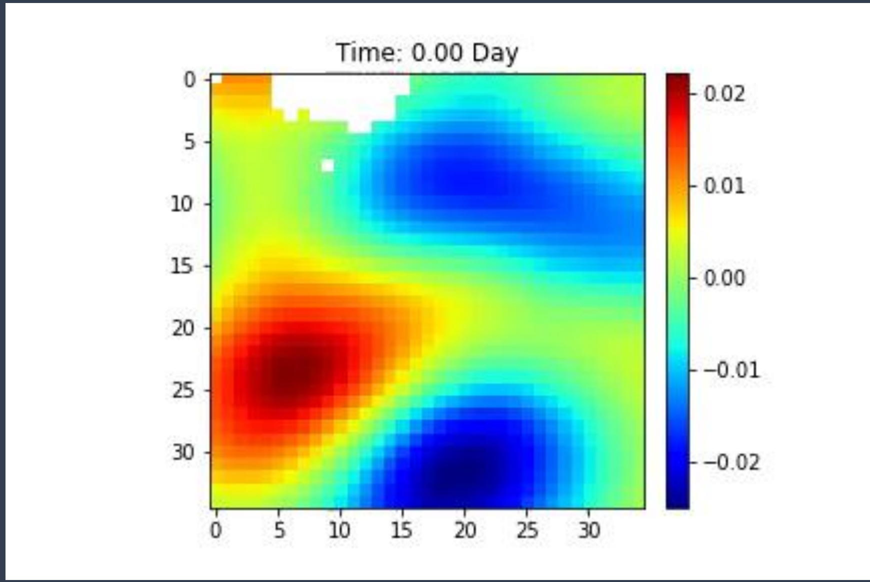




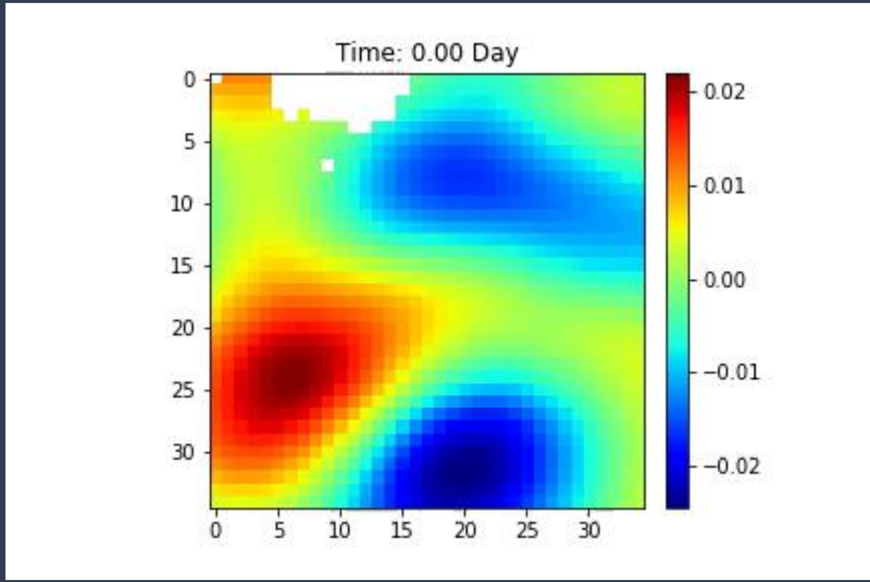
UNOBSERVABLE COMPONENTS

Patch Based Sea Level Anomaly (SLA)

Model Simulation



Model Simulation from a perturbed initial condition



Links to koopman operator theory

Koopman operator : Infinite-dimensional linear operator propagating observables in time.

Observables: a function of the observations.

$$U_t g(x) = g(\Phi_t(x))$$

With
$g(x)$ The observables and U_t the koopman operator.

Find an embedded representation :

Our approach : Project our observation x into a high dimensional space u, with $u = [x, l1, l2, \dots, ln]$

Fit :

$$\frac{du}{dt} = Au$$

With

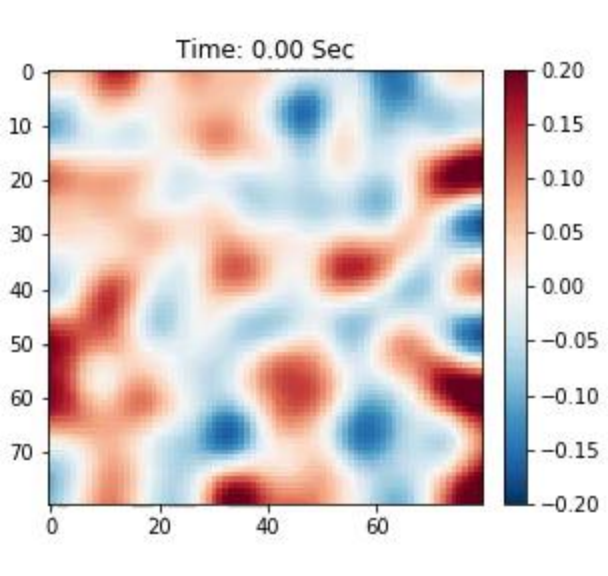
$$u = [x, l1, l2, \dots, ln]$$

$$\theta, l = argmin_{\theta,l}\left\{\alpha \left|x(t) - G\left(\int_{t-1}^{t} Au(t')dt'\right)\right| + (1-\alpha)\left|u(t) - \int_{t-1}^{t} Au(t')dt'\right|\right\}$$

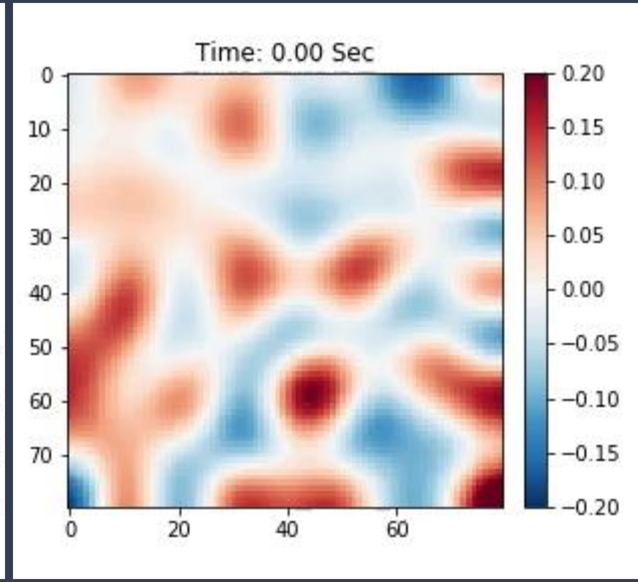
UNOBSERVABLE COMPONENTS

SWE

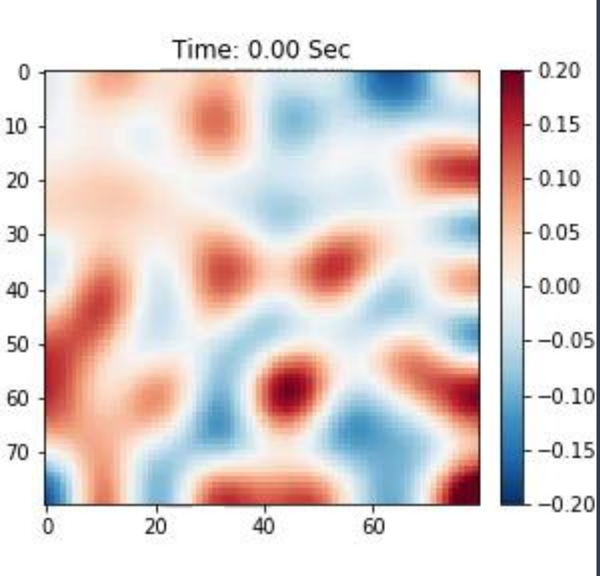
True State



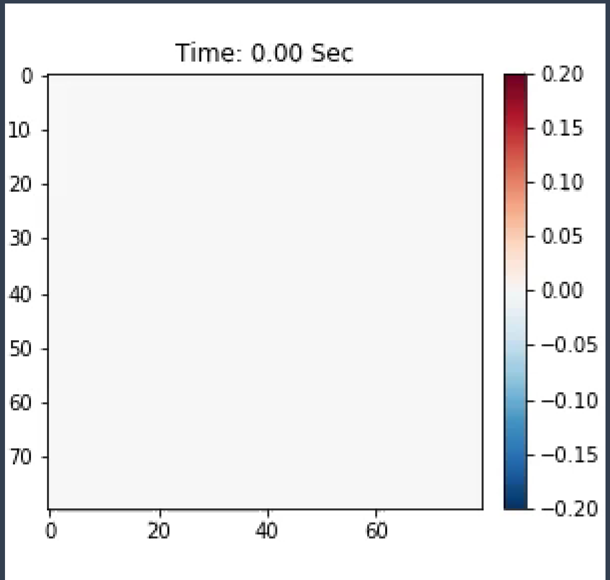
Projection



Koopman

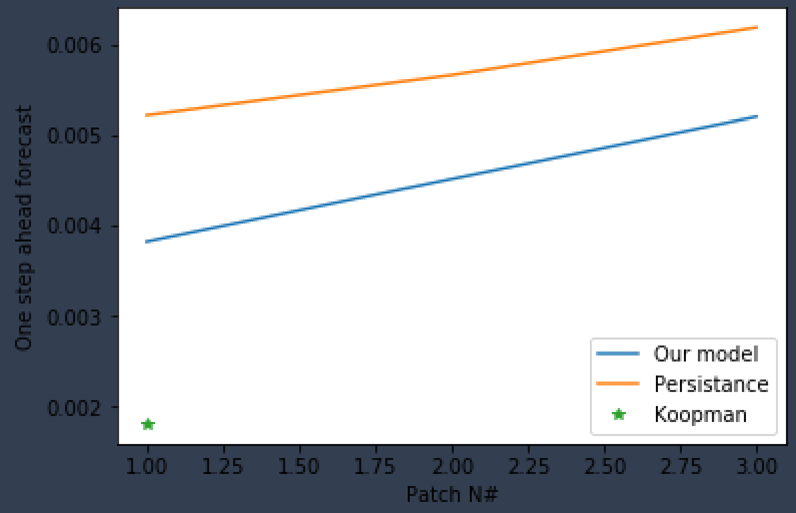


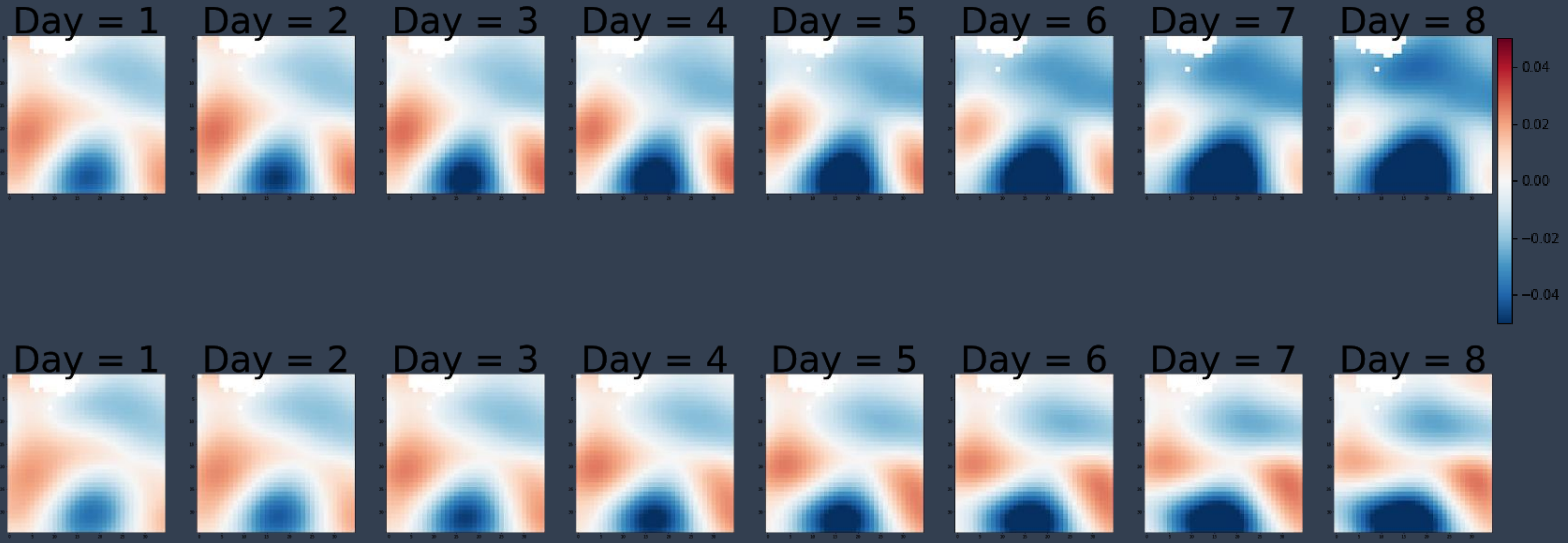
Error (MAE)



UNOBSERVABLE COMPONENTS

SLA (WMOP)





UNOBSERVABLE COMPONENTS

Data assimilation

Reconstruction RMSE :

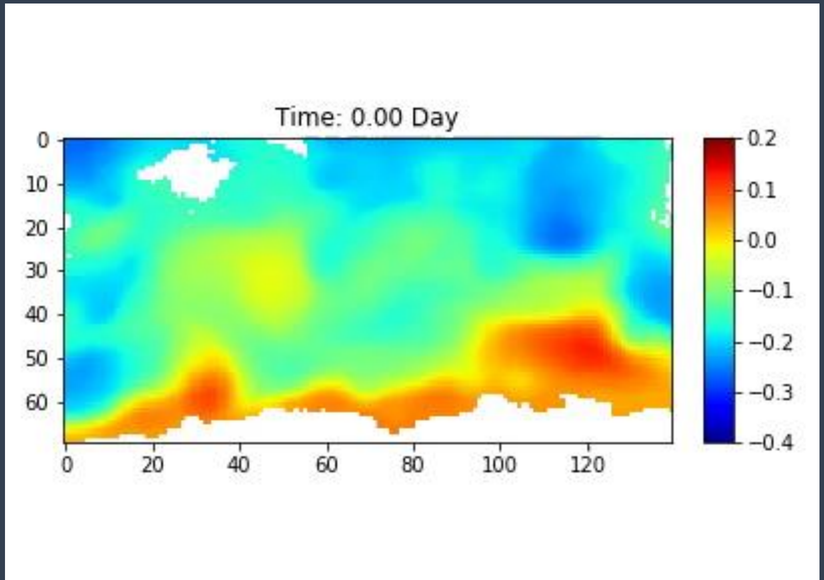
	Model		Forecast ($t_0 + dt$)	Data assimilation
SW	MLP	RMSE	$3.5E-3$	$2.2E-2$
		Correlation	99.90%	74.60%
	AF	RMSE	$4.4E-2$	$1.7E-2$
		Correlation	99.39%	93.96%
	Our model	RMSE	$2.9E-4$	$2.5E-3$
		Correlation	99.99%	99.69%
SLA	MLP	RMSE	$4.0E-3$	$1.6E-2$
		Correlation	96.90%	78.84%
	AF	RMSE	$4.3E-3$	$1.6E-2$
		Correlation	96.44%	78.32%
	Our model	RMSE	$3.0E-3$	$1.4E-2$
		Correlation	98.35%	83.55%

Koopman RMSE : 1.3 E-2

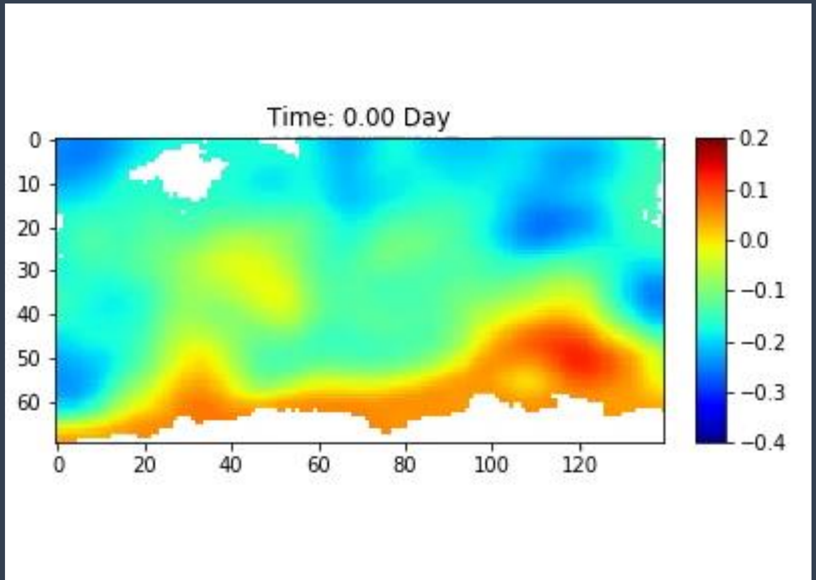
UNOBSERVABLE COMPONENTS

SLA (WMOP)

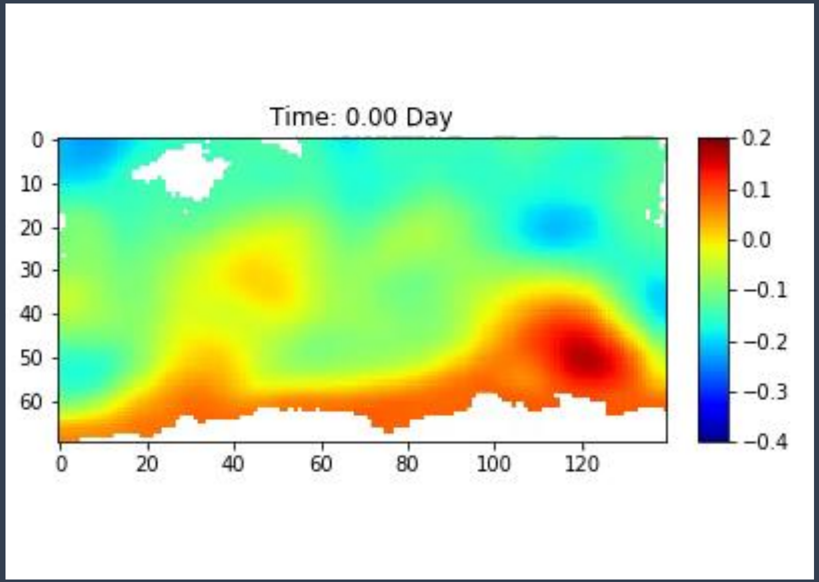
True State



Projection



Koopman



Linear model ?

Pros :

- Simple linear model (can be integrated analytically even without a laptop)
- Easier to train
- More suitable for data assimilation application (modulo good observations)

Cons :

- Can’t simulate the data dynamics (only short term forecast)
- Initial condition needs to be in the attractor space
- No transient reproduction

Outline

- Problem statement
- Data driven model identification from a sequence of observations
    - Direct observations of the state space
    - Noisy and partial observations
    - Temporally sparse data
    - Partially observed systems
- **Discussion**
- **Perspectives**

Discussion

- Can we use data driven models to forcast to ∞ sea surface variables ?

References

- Ouala, Said & Nguyen, Duong & Drumetz, Lucas & Chapron, Bertrand & Pascual, Ananda & Collard, Fabrice & Gaultier, Lucile & Fablet, Ronan. (2019). Learning Latent Dynamics for Partially Observed Systems. doi: 10.13140/RG.2.2.25392.20481/1.
- Sauer, T., Yorke, J. A., & Casdagli, M. (1991). Embedology. Journal of statistical Physics, 65(3-4), 579-616.
- Schlegel, M., & Noack, B. R. (2015). On long-term boundedness of Galerkin models. Journal of Fluid Mechanics, 765, 325-352.
- Juza, M., Mourre, B., Renault, L., Gómara, S., Sebastián, K., Lora, S., Beltran, J.P., Frontera, B., Garau, B., Troupin, C., Torner, M., Heslop, E., Casas, B., Escudier, R., Vizoso, G., Tintoré, J. (2016). SOCIB operational ocean forecasting system and multi-platform validation in the Western Mediterranean Sea. Journal of Operational Oceanography, 9, s155-s166.