

# Estimating Dynamical Models using Machine Learning and Data Assimilation

Dr. Pierre Tandeo, associate professor

IMT Atlantique, Brest, France

August 28, 2023

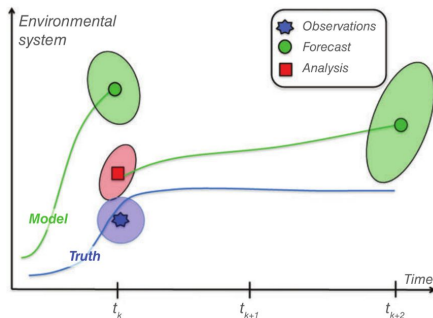


**IMT Atlantique**  
Bretagne-Pays de la Loire  
École Mines-Télécom



*IMT-Atlantique & Kyoto University & RIKEN joint Data Assimilation workshop*

# Data assimilation (graphical representation)



## Principle:

- mix between a **dynamical model** and **observations**
- taking into account their **uncertainties**

## Goal:

- to estimate accurately the **initial condition**
- in order to get the **best forecast**

# Data assimilation (mathematical formulation)

## State-space model:

$$\dot{\mathbf{x}} = f(\mathbf{x}) + \boldsymbol{\eta}$$

$$\mathbf{y} = h(\mathbf{x}) + \boldsymbol{\epsilon}$$

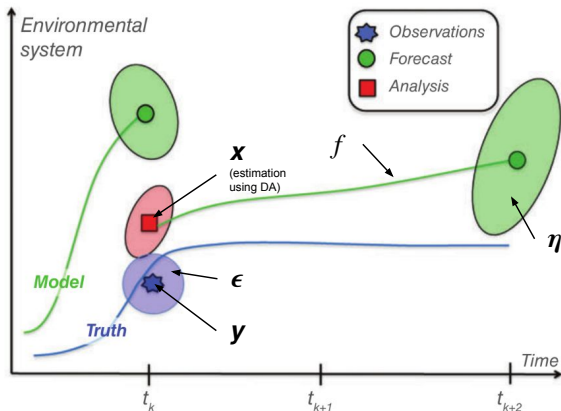
- $\mathbf{y}$  → observations (what is provided)
- $\mathbf{x}$  → state (what is unknown)
- $h$  → link between  $\mathbf{x}$  and  $\mathbf{y}$
- $f$  → **dynamical model**
- $\boldsymbol{\eta}$  and  $\boldsymbol{\epsilon}$  → additive Gaussian errors

# Data assimilation

## State-space model:

$$\dot{\mathbf{x}} = f(\mathbf{x}) + \boldsymbol{\eta}$$

$$\mathbf{y} = h(\mathbf{x}) + \boldsymbol{\epsilon}$$



Source: [Carrassi et al., 2018]

# Examples of dynamical models

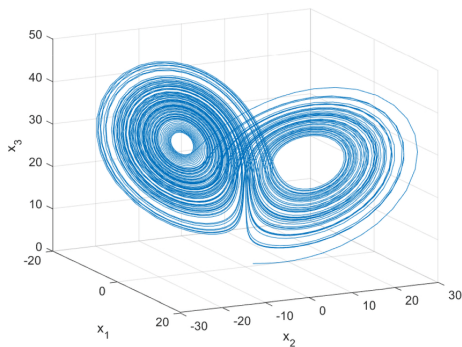
## The Lorenz system:

$$\dot{x}_1 = \sigma(x_2 - x_1)$$

$$\dot{x}_2 = x_1(\rho - x_3) - x_2$$

$$\dot{x}_3 = x_1x_2 - \beta x_3$$

with  $x_1$  the **rate of convection**,  $x_2$  the **horizontal temperature gradient**, and  $x_3$  the **vertical temperature gradient**.



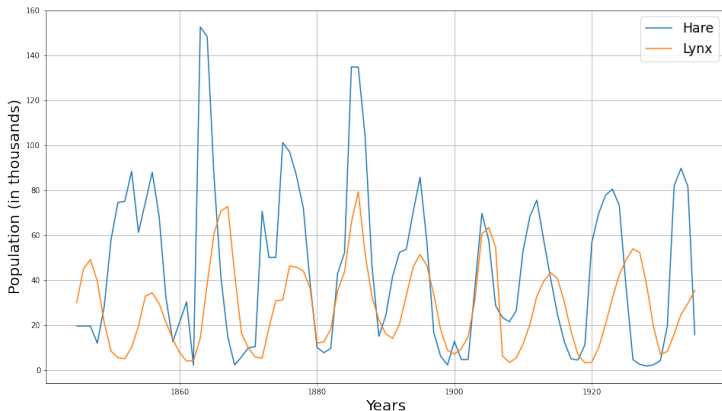
# Examples of dynamical models

Lotka–Volterra equations:

$$\dot{x}_1 = \alpha x_1 - \beta x_1 x_2$$

$$\dot{x}_2 = \delta x_1 x_2 - \gamma x_2$$

with  $x_1$  the **preys** and  $x_2$  the **predators**.



Dynamical model  $f$ :

$$\dot{\mathbf{x}} = f(\mathbf{x})$$

Questions:

- when **equations are available**, is it possible to **improve  $f$** ?
- when **equations are not available**, is it possible to **learn  $f$** ?

Answers:

- yes, using **available observations** (i.e., noisy and partial)
- combining **data assimilation** and **machine learning**

# Combine DA and ML to estimate $f$

## Different case studies:

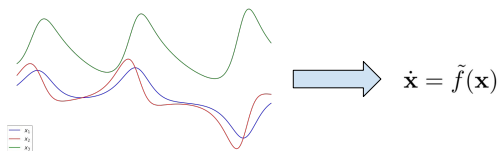
- ①  $f$  is available but **parameters are unknown**

$$\dot{x}_1 = \sigma(x_2 - x_1)$$

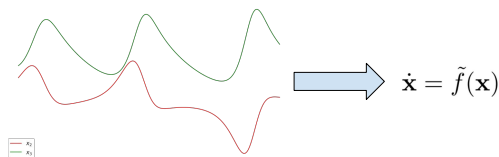
$$\dot{x}_2 = x_1(\rho - x_3) - x_2 \quad \Longrightarrow \quad \theta = [\sigma, \rho, \beta]$$

$$\dot{x}_3 = x_1x_2 - \beta x_3$$

- ②  $f$  is approximated using a **catalog** of the **full system**



- ③  $f$  is approximated using a **catalog** of the **partial system**





# (1) $f$ is defined by equations but parameters are unknown

$$\dot{x}_1 = \sigma(x_2 - x_1)$$

$$\dot{x}_2 = x_1(\rho - x_3) - x_2 \quad \longrightarrow \quad \theta = [\sigma, \rho, \beta]$$

$$\dot{x}_3 = x_1x_2 - \beta x_3$$

State-space model (with augmented state) :

$$\dot{\mathbf{x}} = f(\mathbf{x}) + \eta$$

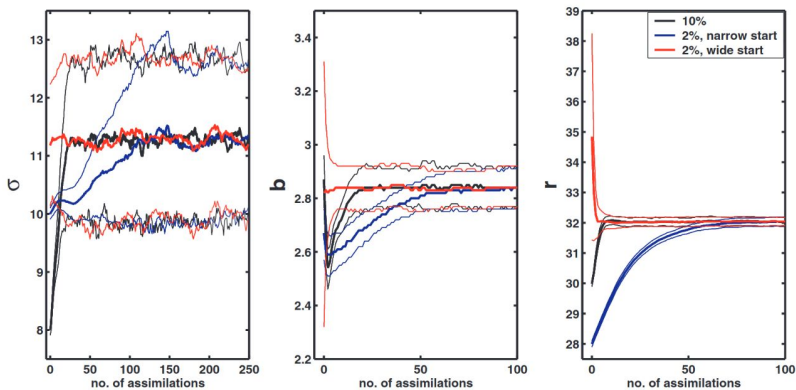
$$\mathbf{y} = h(\mathbf{x}) + \epsilon$$

- $\mathbf{x} = [\mathbf{x}, \theta] \rightarrow$  **augmented state**
- $\theta \rightarrow$  **unknown parameters**
- $f \rightarrow$  dynamical model for  $\mathbf{x}$  and  $\theta$
- see [Ruiz et al., 2013] for more details

# (1) $f$ is defined by equations but parameters are unknown

Example of the Lorenz system:

- $\mathbf{x} = [x_1, x_2, x_3, \sigma, \rho, \beta]$
- $f$  corresponds to the system of 3 ODEs to predict  $[x_1, x_2, x_3]$
- $f$  also corresponds to random walks for  $[\sigma, \rho, \beta]$

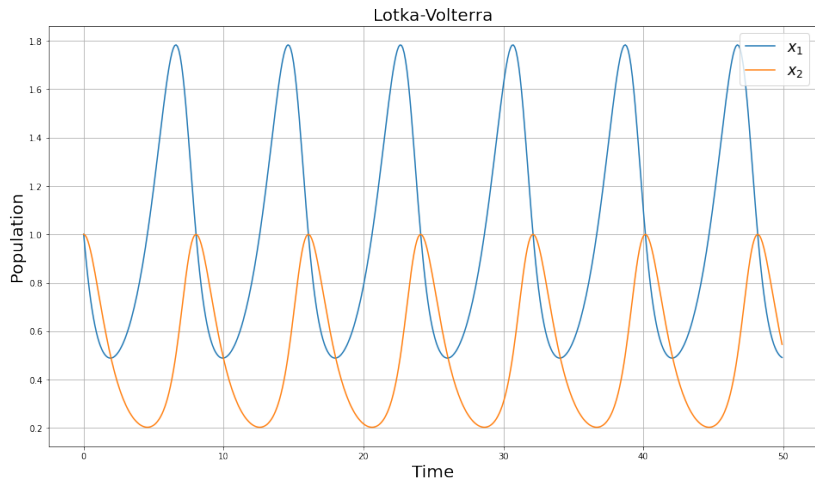


Source: [Annan and Hargreaves, 2004]

# (1) $f$ is defined by equations but parameters are unknown

## Example of Lotka-Volterra:

- $\mathbf{x} = [x_1, x_2, \alpha, \beta, \gamma, \delta]$
- $f$  corresponds to the system of 2 ODEs to predict  $[x_1, x_2]$
- $f$  also corresponds to random walks for  $[\alpha, \beta, \gamma, \delta]$

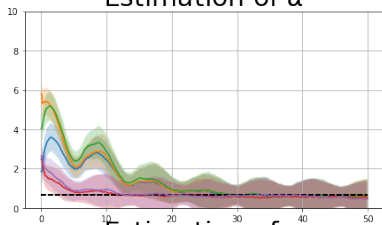


# (1) $f$ is defined by equations but parameters are unknown

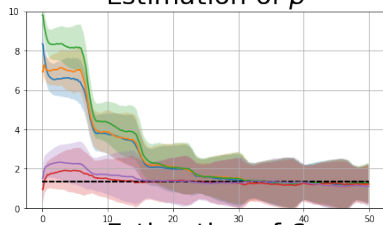
## Example of Lotka-Volterra:

- $x_1$  and  $x_2$  are observed
- $\alpha$ ,  $\beta$ ,  $\gamma$ , and  $\delta$  are initialized using random values

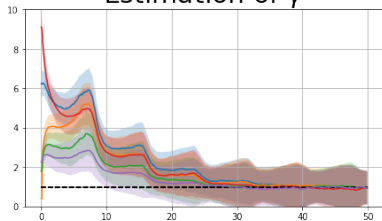
Estimation of  $\alpha$



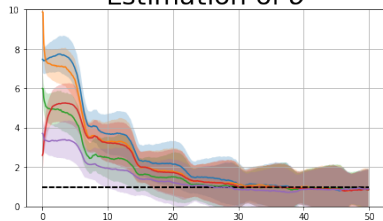
Estimation of  $\beta$



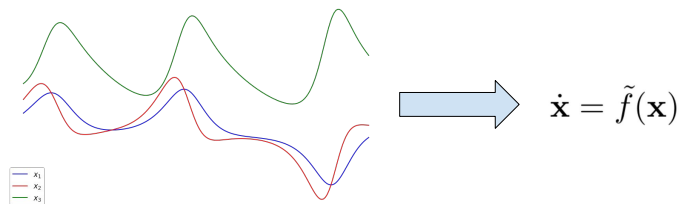
Estimation of  $\gamma$



Estimation of  $\delta$



## (2) $f$ is approximated using a catalog of the full system



State-space model (with emulated model) :

$$\dot{\mathbf{x}} = f(\mathbf{x}) + \eta$$

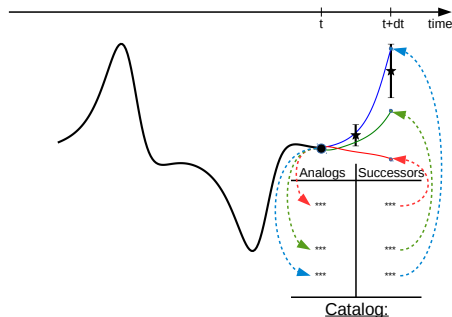
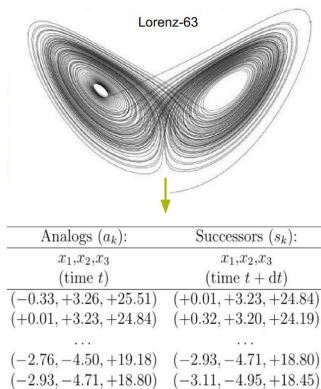
$$\mathbf{y} = h(\mathbf{x}) + \epsilon$$

- $f = \tilde{f} \rightarrow$  **emulated dynamical model**
- $\tilde{f} \rightarrow$  build using a **full catalog** and **regression methods**

## (2) $f$ is approximated using a catalog of the full system

More details about the Analog Data Assimilation (AnDA):

- $\tilde{f}$  is learned using a **catalog of analogs and successors**
- **forecasts** from  $\tilde{f}$  are compared to **observations**



Source: [Tandeo et al., 2015] and [Lguensat et al., 2017]

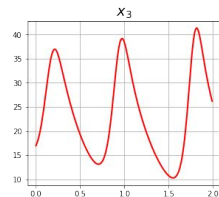
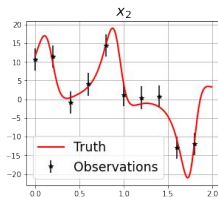
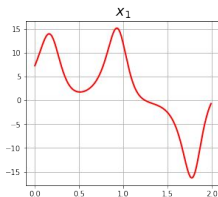
## (2) $f$ is approximated using a catalog of the full system

### Example of the Lorenz system:

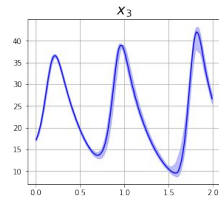
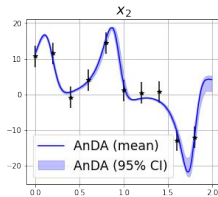
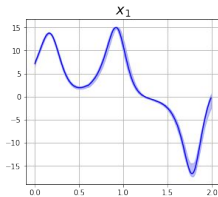
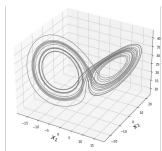
- $\tilde{f}$  is build using a catalog of  $\mathbf{x} = [x_1, x_2, x_3]$
- $x_2$  is observed, but never  $x_1$  nor  $x_3$

#### With Lorenz equations:

$$\begin{aligned}\dot{x}_1 &= \sigma(x_2 - x_1) \\ \dot{x}_2 &= x_1(\rho - x_3) - x_2 \\ \dot{x}_3 &= x_1x_2 - \beta x_3\end{aligned}$$



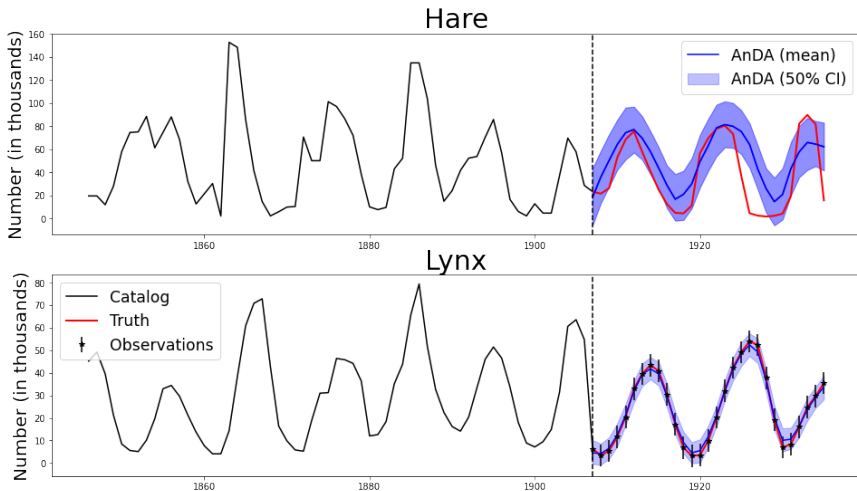
#### With a full catalog:



## (2) $f$ is approximated using a catalog of the full system

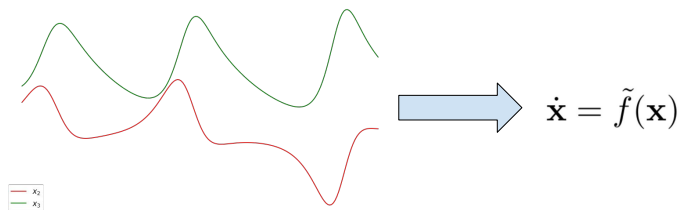
Example of the Hare-Lynx database:

- $\tilde{f}$  is build using a catalog of  $\mathbf{x} = [Hare, Lynx]$  until 1907
- after 1907, Lynx are observed, but never Hares





### (3) $f$ is approximated using a catalog of the partial system



State-space model (with emulated model + augmented state) :

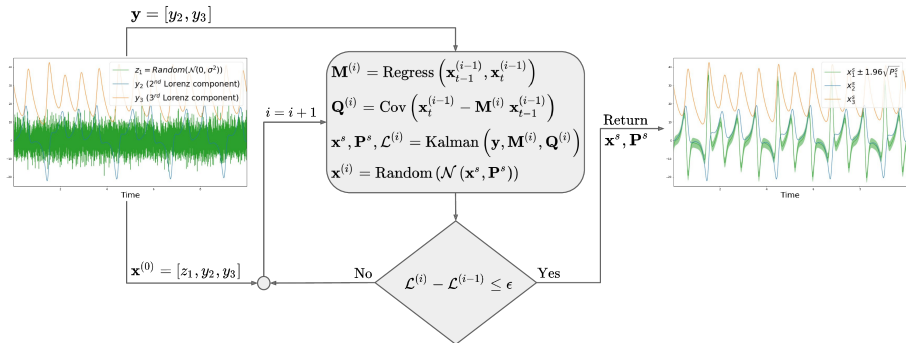
$$\dot{\mathbf{x}} = f(\mathbf{x}) + \eta$$

$$\mathbf{y} = h(\mathbf{x}) + \epsilon$$

- $f = \tilde{f} \rightarrow$  **emulated dynamical model**
- $\tilde{f} \rightarrow$  build using a **partial database** and **regression methods**
- $\mathbf{x} = [\mathbf{x}, \mathbf{z}] \rightarrow$  **augmented state** with latent variables  $\mathbf{z}$

### (3) $f$ is approximated using a catalog of the partial system

Schematic representation of the proposed iterative algorithm:



$\Rightarrow$  starting from a **Gaussian white noise**, the latent variable automatically adjusts over **iterations**

### (3) $f$ is approximated using a catalog of the partial system

Example of the Lorenz system:

- $\mathbf{x} = [x_2, x_3, z_1]$
- $x_2$  and  $x_3$  are observed, but never  $x_1$

### (3) $f$ is approximated using a catalog of the partial system

Example of the Hare-Lynx database:

- $\mathbf{x} = [\text{Lynx}, z]$
- Lynx are observed, but never Hares

## Conclusions (presented here):

- a dynamical model  $f$  can be **improved or learned**
- using **observations** and appropriate tools (**DA and ML**)

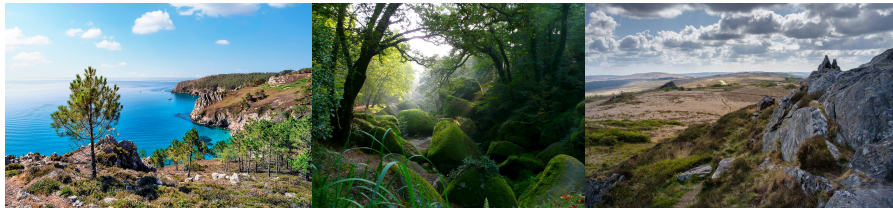
## Important remarks (not presented here):

- estimation of **model and observation errors** is crucial (see [Tandeo et al., 2020] for a review)
- using **optimization algorithms** (e.g., descent gradient or Expectation-Maximization)

ありがとうございます (arigatō gozaimasu)!

Contact:

- pierre.tandeo@imt-atlantique.fr
- www.tandeo.wordpress.com



**Figure:** Nice places to visit near Brest, France: Crozon (left), Huelgoat (middle), Monts d'Arrée (right).

# Bibliography I



Annan, J. and Hargreaves, J. (2004).  
Efficient parameter estimation for a highly chaotic system.  
[Tellus A: Dynamic Meteorology and Oceanography](#), 56(5):520–526.



Carrassi, A., Bocquet, M., Bertino, L., and Evensen, G. (2018).  
Data assimilation in the geosciences: An overview of methods, issues, and perspectives.  
[Wiley Interdisciplinary Reviews: Climate Change](#), 9(5):e535.



Lguensat, R., Tandeo, P., Ailliot, P., Pulido, M., and Fablet, R. (2017).  
The analog data assimilation.  
[Monthly Weather Review](#), 145(10):4093–4107.



Ruiz, J. J., Pulido, M., and Miyoshi, T. (2013).  
Estimating model parameters with ensemble-based data assimilation: A review.  
[Journal of the Meteorological Society of Japan. Ser. II](#), 91(2):79–99.



Tandeo, P., Ailliot, P., Bocquet, M., Carrassi, A., Miyoshi, T., Pulido, M., and Zhen, Y. (2020).  
A review of innovation-based methods to jointly estimate model and observation error covariance matrices in ensemble data assimilation.  
[Monthly Weather Review](#), 148(10):3973–3994.



Tandeo, P., Ailliot, P., Ruiz, J., Hannart, A., Chapron, B., Cuzol, A., Monbet, V., Easton, R., and Fablet, R. (2015).  
Combining analog method and ensemble data assimilation: application to the lorenz-63 chaotic system.  
In [Machine learning and data mining approaches to climate science](#), pages 3–12. Springer.