Analog forecasting errors from a dynamical systems point of view

Paul Platzer (feb. 10, 1993)

IMT-A & RIKEN workshop :
Statistical Modeling and Machine Learning
in Meteorology and Oceanography, feb. 10, 2020

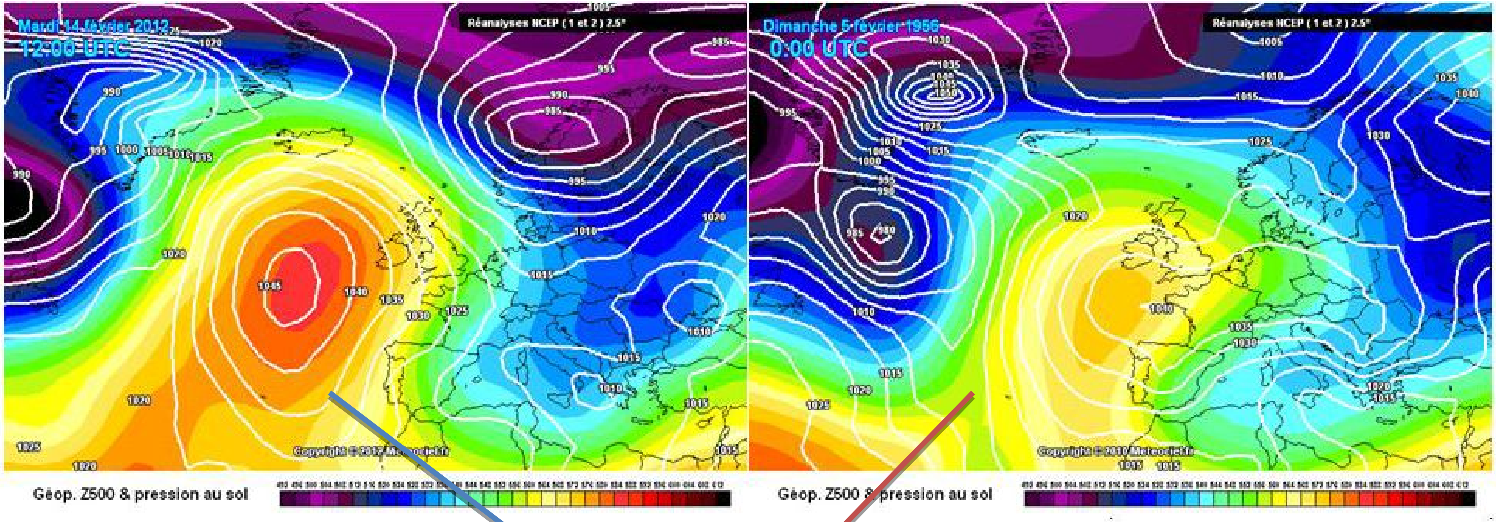
Analog forecasting errors from a dynamical systems point of view

Pierre Tandeo,
Yicun Zhen
AnDA

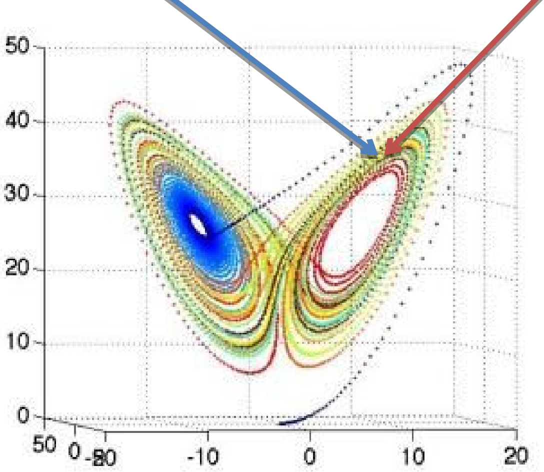
Marc Schoenauer,
Naonori Ueda,
Said Ouala,
Maha Mdini…
→ bridge the gap between
data-driven and process-driven
→ « 5th science »

Paul Platzer (feb. 10, 1993)
IMT-A & RIKEN workshop :
Statistical Modeling and Machine Learning
in Meteorology and Oceanography, feb. 10, 2020

What is an analog ?



→ Introduced by Lorenz in 1969 for atmospheric predictability

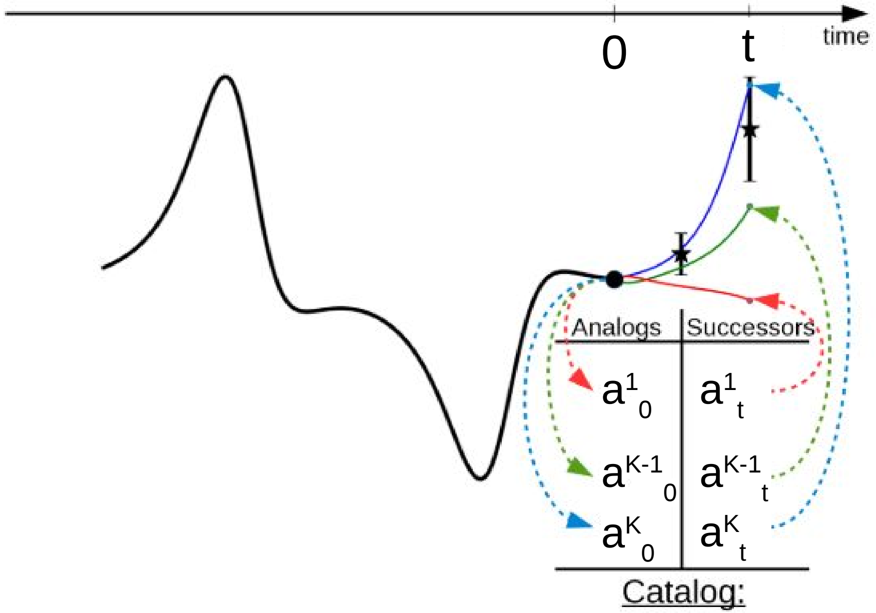


Two analog maps correspond to close points in phase space

(Courtesy of D. Faranda and P. Yiou)

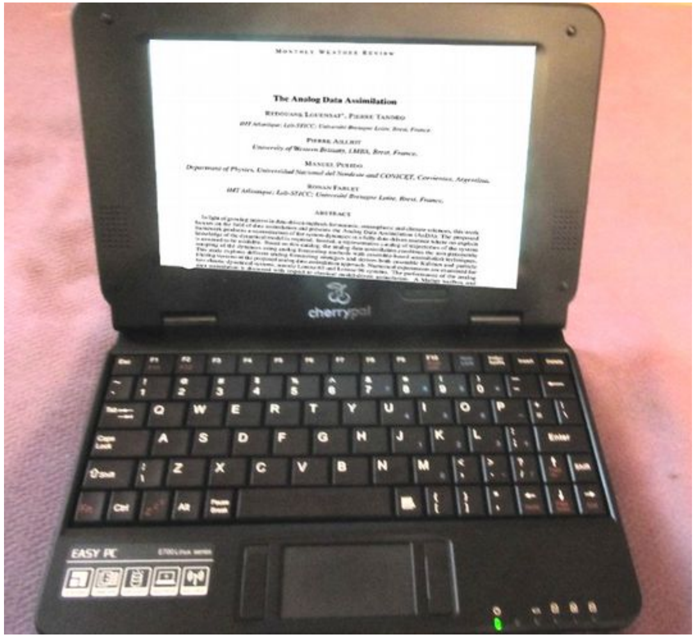
Analog forecasting : the idea

- a^k_0 = analogs of x_0 : « close » to x_0
- a^k_t successors (time *t*) → estimate x_t



(Lguensat et al. 2017)

Analog forecasting : cheap ML



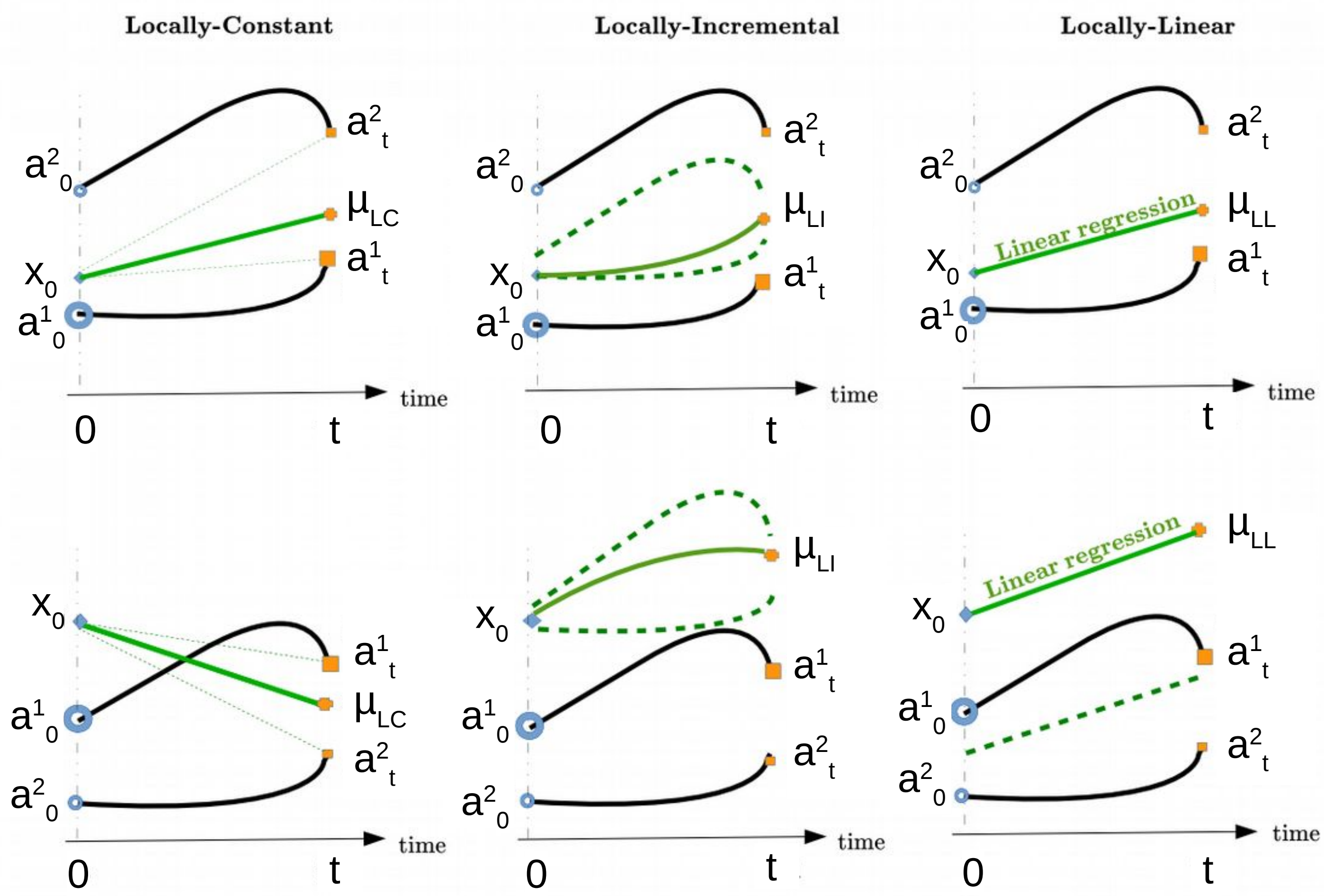
Analog forecasting : in practice

- Large database called « catalog » (reanalysis, model output)
- Select K nearest neighbors of x_0 (fast thanks to ML libraries)
- Assign weights to the analogs (discard bad analogs)
- Apply a given analog forecasting operator

$$\mathcal{A} : \mathbf{x}_0 \rightarrow \hat{\mathbf{x}}_t \begin{cases} \sim \sum_k \omega_k \boldsymbol{\delta_k} \\ \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \end{cases}$$

• • •

Analog forecasting operators



$$\mu_{LC} = \sum_k \omega_k \mathbf{a}_t^k$$
$$\mu_{LI} = \mathbf{x}_0 + \sum_k \omega_k (\mathbf{a}_t^k - \mathbf{a}_0^k)$$
$$\mu_{LL} = \mathbf{S}\,(\mathbf{x}_0 - \mu_0) + \mathbf{c}$$

(figure from Lguensat et al. 2017)

Dynamical systems

- Hypotheses :

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{f}(\mathbf{x})\,,$$

$$\forall k,\ \frac{\mathrm{d}\mathbf{a}^k}{\mathrm{d}t} = \mathbf{f}^a(\mathbf{a}^k)\ ; \quad \mathbf{f}^a = \mathbf{f} + \delta\tilde{\mathbf{f}}\,.$$

$$\delta \rightarrow 0$$

- Distance between successor and future state :

$$\mathbf{a}_t^k - \mathbf{x}_t = t\delta\tilde{\mathbf{f}}(\mathbf{x}_0) + \left[\mathbf{I} + t\boldsymbol{\nabla}\mathbf{f}|_{\mathbf{x}_0}\right](\mathbf{a}_0^k - \mathbf{x}_0) + \mathcal{O}\left(t^2\,,\ ||\mathbf{a}_0^k - \mathbf{x}_0||^2\,,\ \delta||\mathbf{a}_0^k - \mathbf{x}_0||\right)$$

$$t \rightarrow 0 \qquad \mathbf{a}_0^k \rightarrow \mathbf{x}_0 \qquad \delta \rightarrow 0$$

Dynamical systems

- Here → looking at local error
  → similar to Nicolis et al. (2009)
- Zhao and Giannakis (2016)
  → looking at global dynamical properties

Dynamical systems : mean error

- Hypotheses :

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{f}(\mathbf{x}),$$

$$\forall k,\ \frac{\mathrm{d}\mathbf{a}^k}{\mathrm{d}t} = \mathbf{f}^a(\mathbf{a}^k)\ ; \quad \mathbf{f}^a = \mathbf{f} + \delta\tilde{\mathbf{f}}.$$

$$t \to 0 \qquad \mathbf{a}_0^k \to \mathbf{x}_0 \qquad \delta \to 0$$

$$\boldsymbol{\mu}_0 = \textstyle\sum_k \omega_k \mathbf{a}_0^k$$

- Mean error of analog forecasting operators :

$$\boldsymbol{\mu}_{\mathrm{LC}} - \mathbf{x}_t = t\delta\tilde{\mathbf{f}}(\mathbf{x}_0) + \left[\mathbf{I} + t\boldsymbol{\nabla}\mathbf{f}|_{\mathbf{x}_0}\right](\boldsymbol{\mu}_0 - \mathbf{x}_0) + \mathcal{O}\left(t^2,\ \sum_k \omega_k ||\mathbf{a}_0^k - \mathbf{x}_0||^2,\ \delta \sum_k \omega_k ||\mathbf{a}_0^k - \mathbf{x}_0||\right)$$

$$\boldsymbol{\mu}_{\mathrm{LI}} - \mathbf{x}_t = t\delta\tilde{\mathbf{f}}(\mathbf{x}_0) + \left[t\boldsymbol{\nabla}\mathbf{f}|_{\mathbf{x}_0}\right](\boldsymbol{\mu}_0 - \mathbf{x}_0) + \mathcal{O}\left(t^2, \sum_k \omega_k ||\mathbf{a}_0^k - \mathbf{x}_0||^2,\ \delta \sum_k \omega_k ||\mathbf{a}_0^k - \mathbf{x}_0||\right)$$

$$\boldsymbol{\mu}_{\mathrm{LL}} - \mathbf{x}_t = t\delta\tilde{\mathbf{f}}(\mathbf{x}_0) + \mathcal{O}\left(t^2,\ \sum_k \omega_k ||\mathbf{a}_0^k - \mathbf{x}_0||^2,\ \delta \sum_k \omega_k ||\mathbf{a}_0^k - \mathbf{x}_0||\right)$$

Dynamical systems : mean error

- Hypotheses :

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{f}(\mathbf{x}),$$

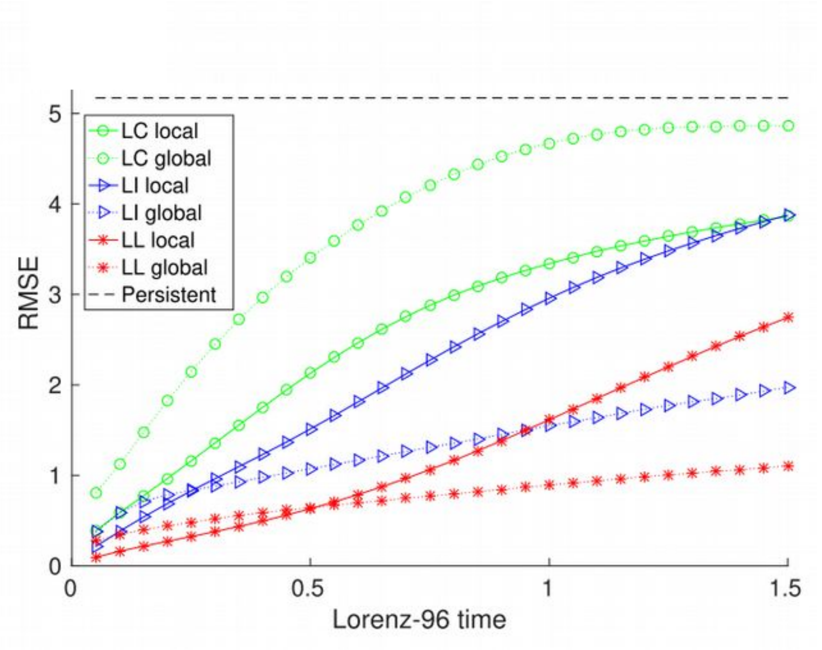
$$\forall k, \; \frac{\mathrm{d}\mathbf{a}^k}{\mathrm{d}t} = \mathbf{f}^a(\mathbf{a}^k) \; ; \quad \mathbf{f}^a = \mathbf{f} + \delta\tilde{\mathbf{f}}.$$

$t \to 0$ $\quad$ $\mathbf{a}_0^k \to \mathbf{x}_0$ $\quad$ $\delta \to 0$

$$\boldsymbol{\mu}_0 = \sum_k \omega_k \mathbf{a}_0^k$$

- Mean error of analog forecasting :

→ *experiments on L96 with δ=0*



(figure from Lguensat et al. 2017)

Dynamical systems : mean error

- Hypotheses :

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{f}(\mathbf{x}),$$

$$\forall k,\ \frac{\mathrm{d}\mathbf{a}^k}{\mathrm{d}t} = \mathbf{f}^a(\mathbf{a}^k)\ ; \quad \mathbf{f}^a = \mathbf{f} + \delta\tilde{\mathbf{f}}\,.$$

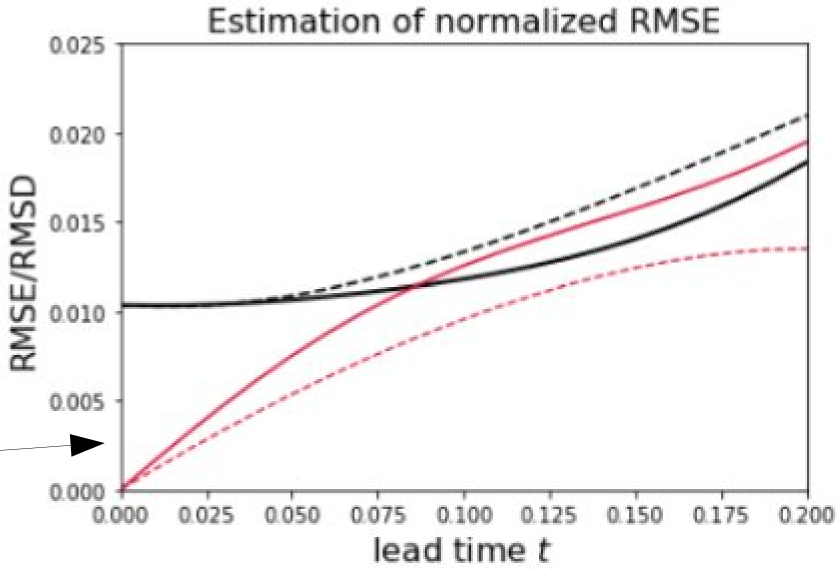
$t \to 0$ $\quad$ $\mathbf{a}_0^k \to \mathbf{x}_0$ $\quad$ $\delta \to 0$

$$\boldsymbol{\mu}_0 = \sum_k \omega_k \mathbf{a}_0^k$$

- Mean error of analog forecasting :

→ *experiments on L63 with δ=0*

— LC empirical error
-- LC error estimation
— LI empirical error
-- LI error estimation



(figure from Platzer et al. 2019)

Dynamical systems : mean error

- Hypotheses :

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{f}(\mathbf{x}) \,,$$

$$\forall k, \; \frac{\mathrm{d}\mathbf{a}^k}{\mathrm{d}t} = \mathbf{f}^a(\mathbf{a}^k) \; ; \quad \mathbf{f}^a = \mathbf{f} + \delta \tilde{\mathbf{f}} \,.$$

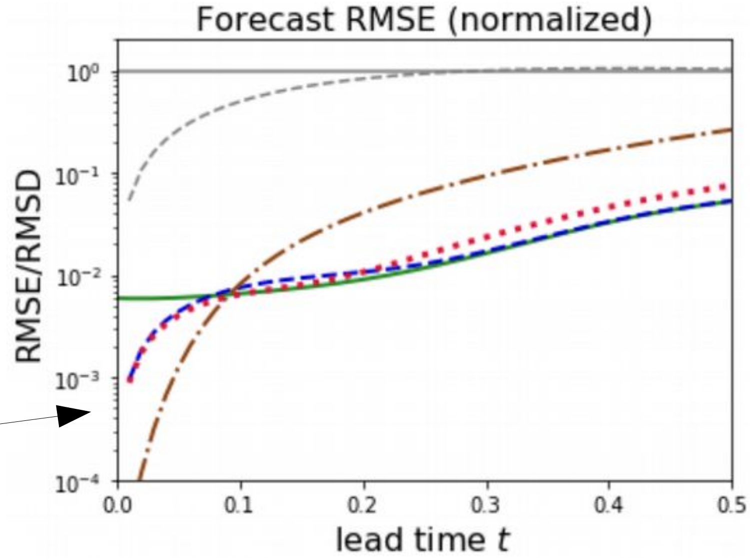
$t \to 0$ $\quad \mathbf{a}_0^k \to \mathbf{x}_0$ $\quad \delta \to 0$

$$\boldsymbol{\mu}_0 = \textstyle\sum_k \omega_k \mathbf{a}_0^k$$

- Mean error of analog forecasting :

→ *experiments on L63 with δ=0*

— LC
-- LI
-- LI+mean correction
-·-· LI+local correction



(figure from Platzer et al. 2019)

Relation between operators

- Locally-linear :

$$\mathbf{a}_t^k = \mathbf{S}(\mathbf{a}_0^k - \boldsymbol{\mu}_0) + \mathbf{c} + \text{residuals}$$

$$\rightarrow \begin{cases} \mathbf{c} = \sum_k \omega_k \mathbf{a}_t^k = \boldsymbol{\mu}_{\mathrm{LC}} \\ \mathbf{S} = \mathbf{I} + t \boldsymbol{\nabla} \mathbf{f}|_{\mathbf{x}_0} + \mathcal{O}\left(t^2 \,,\, ||\boldsymbol{\mu}_0 - \mathbf{x}_0||\right) \end{cases}$$

$$\boxed{\boldsymbol{\mu}_{\mathrm{LL}} = \boldsymbol{\mu}_{\mathrm{LC}} + \mathbf{S}(\mathbf{x}_0 - \boldsymbol{\mu}_0)} \sim_{t \to 0} \boldsymbol{\mu}_{\mathrm{LI}}$$

$\mathbf{a}_0^k \to \mathbf{x}_0$

$\delta \to 0$

$t \to 0$

$\boldsymbol{\mu}_0 = \sum_k \omega_k \mathbf{a}_0^k$

- Locally-constant, locally-incremental :

$$\begin{cases} \boldsymbol{\mu}_{\mathrm{LC}} = \boldsymbol{\mu}_{\mathrm{LL}}|_{\mathbf{S}=\mathbf{0}} \\ \boldsymbol{\mu}_{\mathrm{LI}} = \boldsymbol{\mu}_{\mathrm{LL}}|_{\mathbf{S}=\mathbf{I}} \end{cases}$$

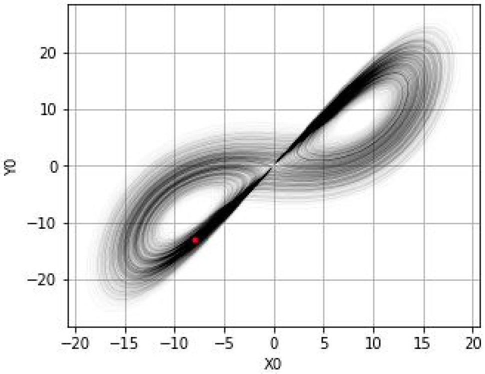
The locally-linear and the Jacobian

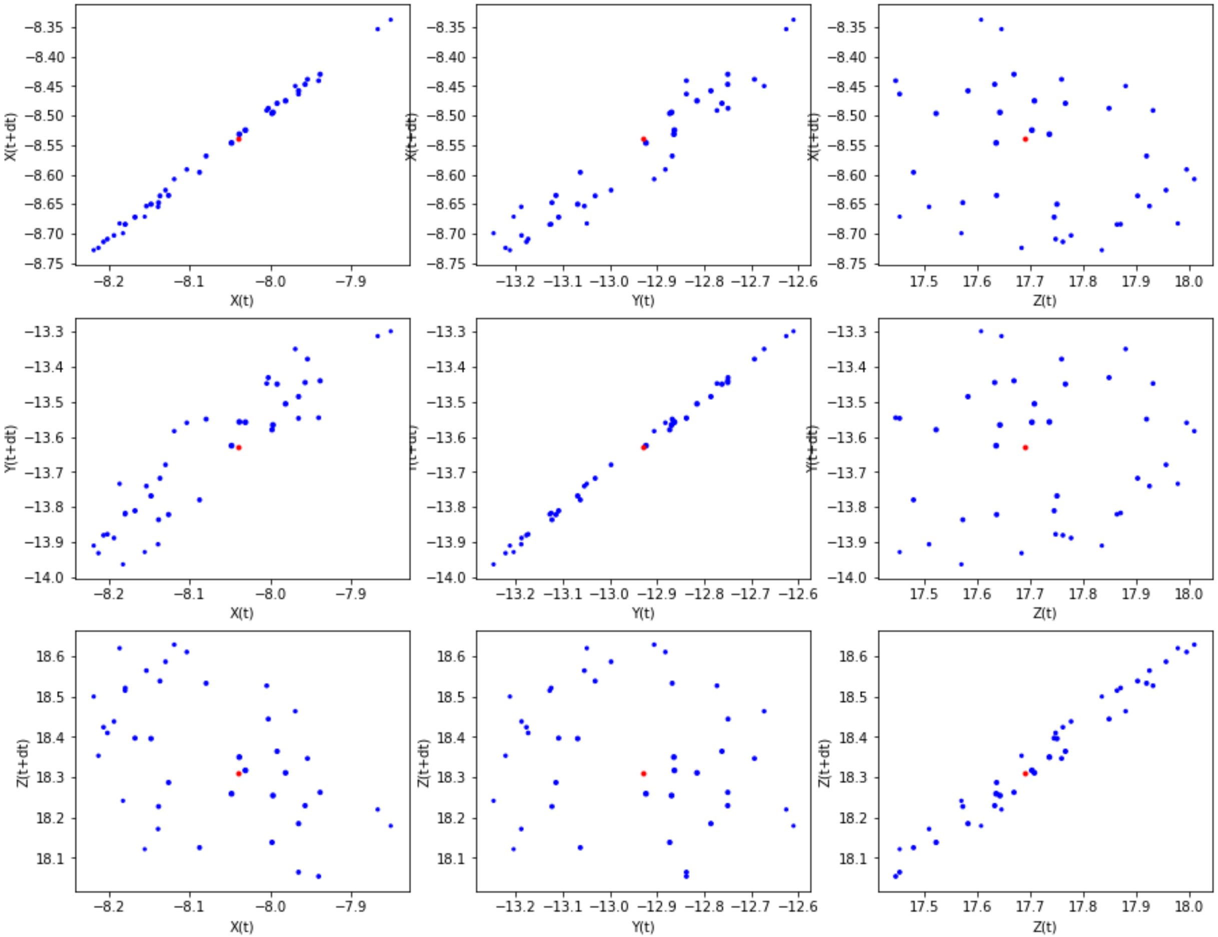
$$\mathbf{S} = \underbrace{\mathbf{I} + t\nabla\mathbf{f}|_{\mathbf{x}_0}} + \mathcal{O}\left(t^2\,,\,||\boldsymbol{\mu}_0 - \mathbf{x}_0||\right)$$

```
array([[ 0.9       ,  0.1       ,  0.        ],
       [ 0.10310034,  0.99      ,  0.08039593],
       [-0.12930831, -0.08039593,  0.97333333]])
```

```
array([[ 0.90948646,  0.09466732,  0.00396134],
       [ 0.09185399,  0.98995536,  0.08224164],
       [-0.12538949, -0.08990802,  0.97069705]])
```

```
dt=0.01
```





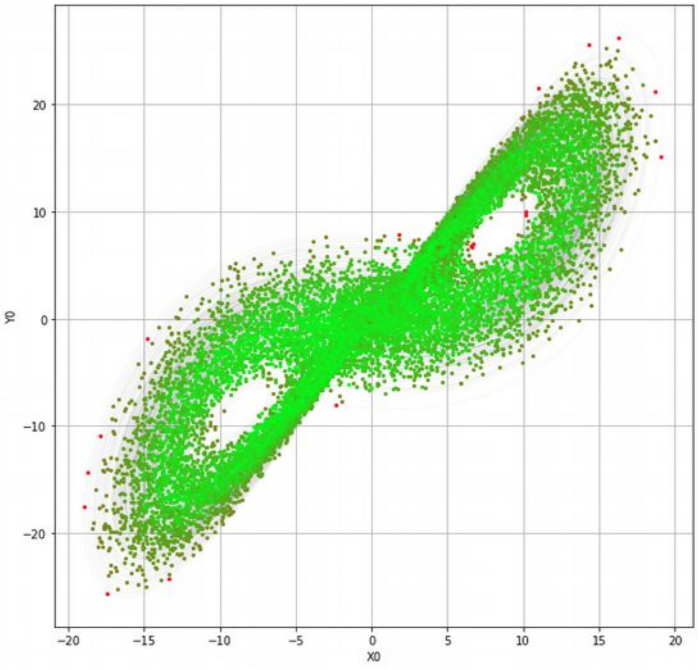
IMT-A & RIKEN
Feb. 10, 2020

The locally-linear and the Jacobian

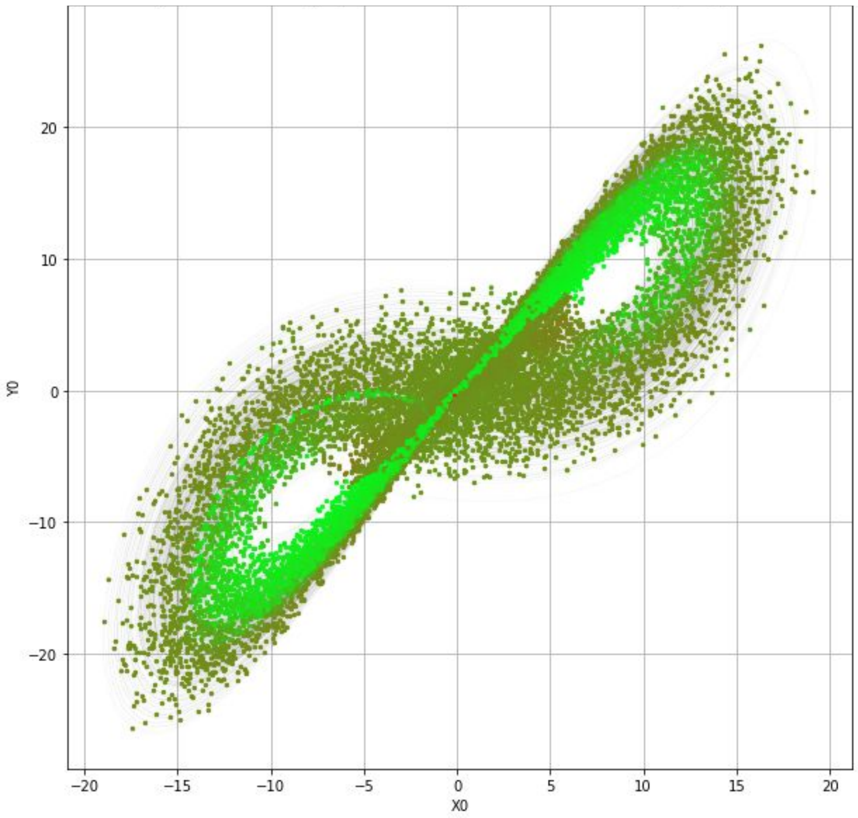
$$\mathbf{S} = \mathbf{I} + t\nabla\mathbf{f}|_{\mathbf{x}_0} + \mathcal{O}\left(t^2\,,\,||\boldsymbol{\mu}_0 - \mathbf{x}_0||\right)$$

```
dt=0.01
```

Mean error of LL forecast



Quality of Jacobian estimation using LL



What about the covariance ?

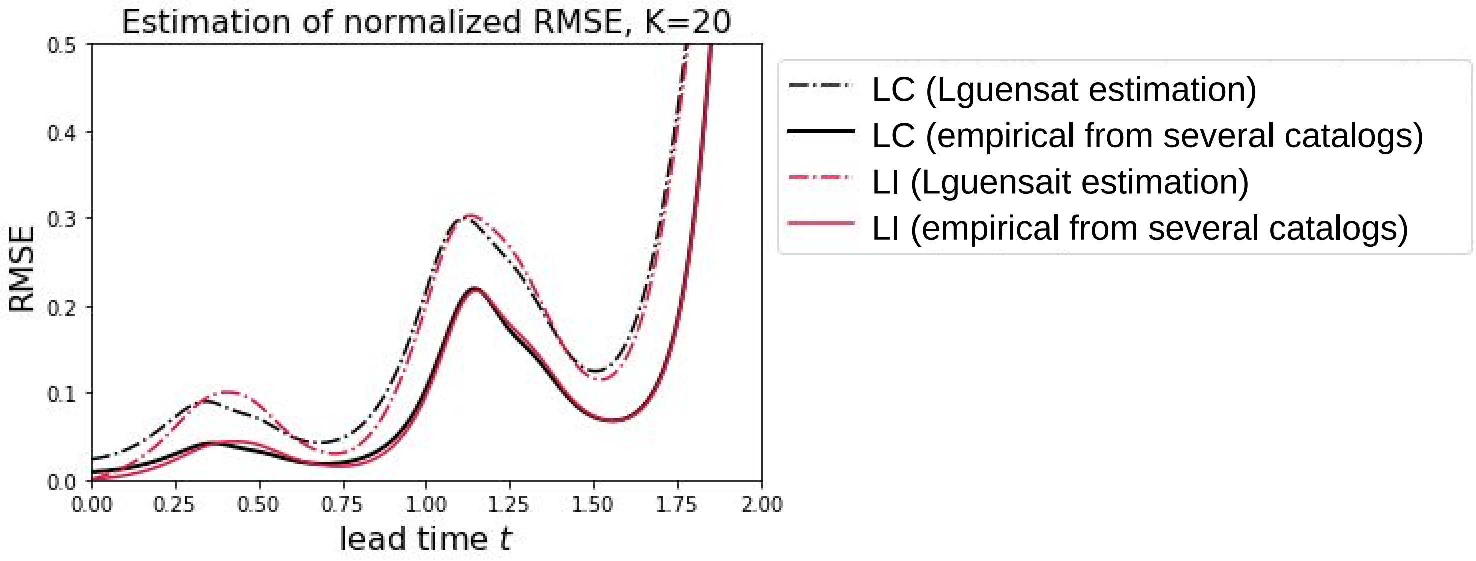
$$\mathcal{A} : \mathbf{x}_0 \to \hat{\mathbf{x}}_t \begin{cases} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \\ \sim \sum_k \omega_k \boldsymbol{\delta}_{\boldsymbol{k}} \end{cases}$$

$\boldsymbol{\Sigma} \to$

- $\boldsymbol{\Sigma}_{\mathrm{LC}}$
- $\boldsymbol{\Sigma}_{\mathrm{LI}}$
- $\boldsymbol{\Sigma}_{\mathrm{LL}}$

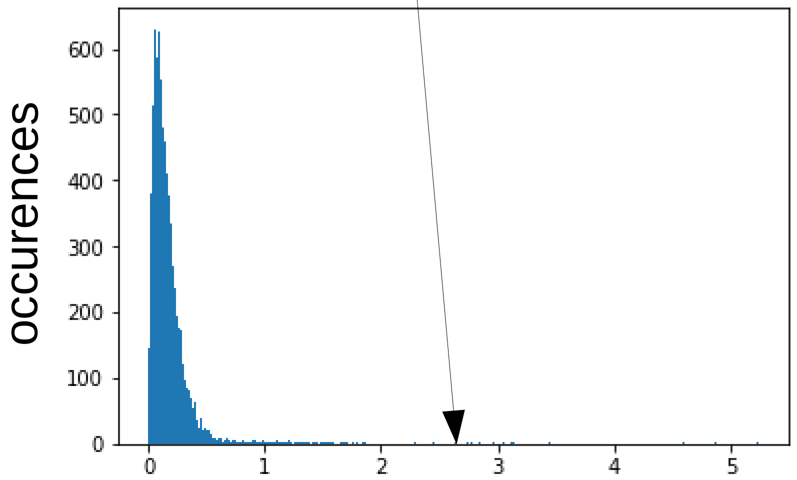
What about the variance ?

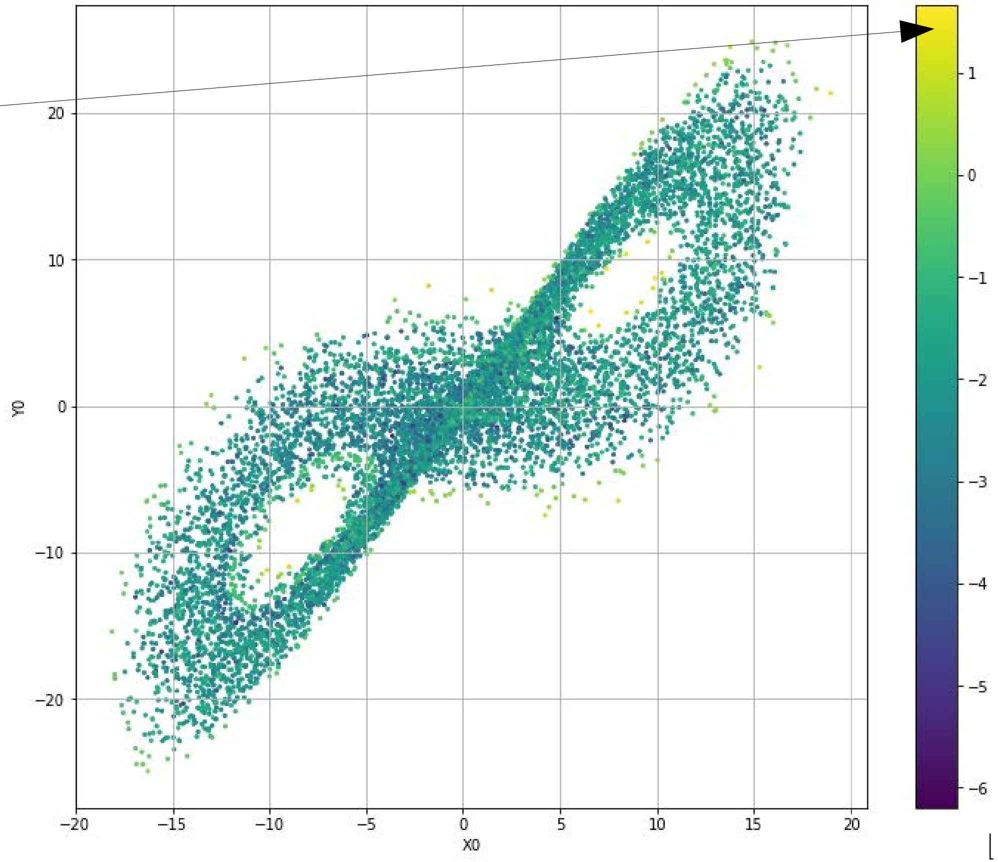
→ *Experiments on L63*



What about the variance ?

$$\frac{|\boldsymbol{\mu}_{\mathrm{LC}} - \mathbf{x}_t|}{\mathbf{Tr}(\boldsymbol{\Sigma})^{1/2}}$$





Conclusion

- Analog forecasting = empirical, simple ML
- Using dynamical systems helps
  → interpretation
  → error estimation
  → data-based + model-based hybrid method ?
  → finer tuning ?
- Combine analogs and NN ? (analogs=first guess)

Thank you !

Bibliography :

• Lguensat, R., Tandeo, P., Ailliot, P., Pulido, M., & Fablet, R. (2017). The analog data assimilation. Monthly Weather Review, 145(10), 4093-4107.
• Lorenz, E. N. (1969). Atmospheric predictability as revealed by naturally occurring analogues. Journal of the Atmospheric sciences, 26(4), 636-646.
• Nicolis, C., Perdigao, R. A., & Vannitsem, S. (2009). Dynamics of prediction errors under the combined effect of initial condition and model errors. Journal of the atmospheric sciences, 66(3), 766-778.
• Platzer, P., Yiou, P., Tandeo, P., Naveau, P., & Filipot, J. F. (2019, October). Predicting Analog Forecasting Errors using Dynamical Systems.
• Tippett, M. K., & DelSole, T. (2013). Constructed analogs and linear regression. Monthly Weather Review, 141(7), 2519-2525.
• Zhao, Z., & Giannakis, D. (2016). Analog forecasting with dynamics-adapted kernels. Nonlinearity, 29(9), 2888.