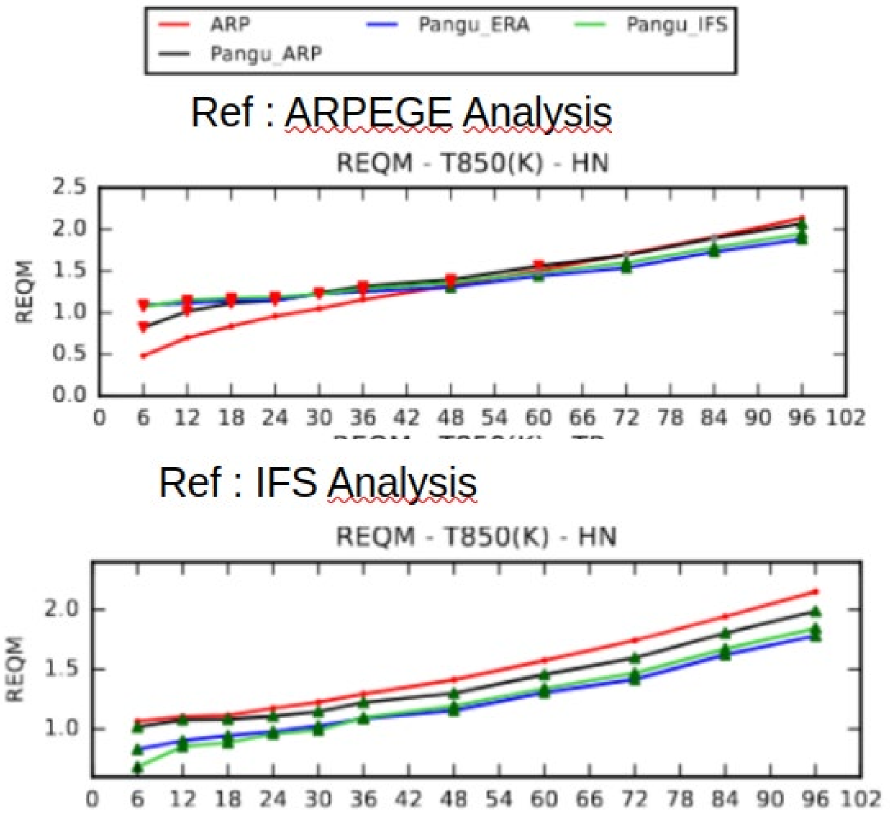
AI and Data Assimilation at Météo-France

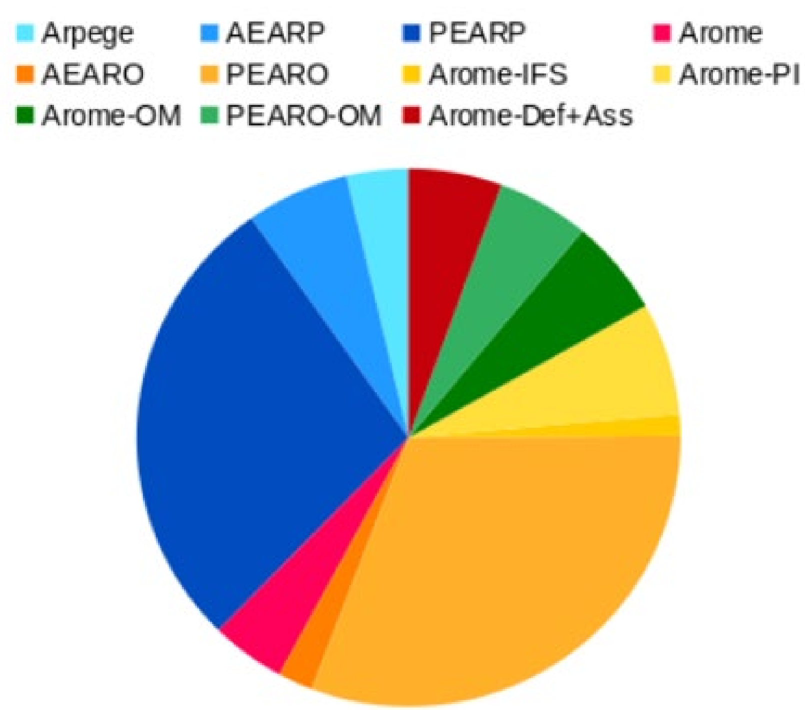
Vincent Chabot, Loïk Berre, Laure Raynaud, Etienne Arbogast, Jocelin Boely, Laurent Descamps, and others

NWP Context at Météo France

PanguWeather Scores



HPC Shared for our NWP System



From : **L. Descamps**

Outline

- AI development for forecasting

- Hybrid AI-DA perspectives

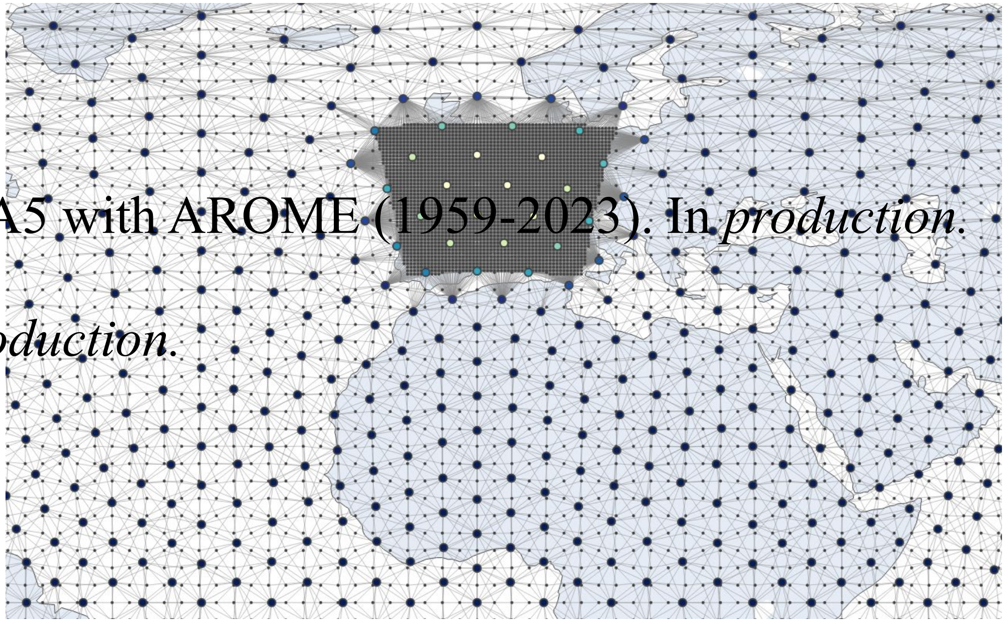
- AI-DA perspectives

AROME AI emulator

Moving from py4cast to anemoi to develop an operational streched-grid approach, following t

Several training datasets to be considered :

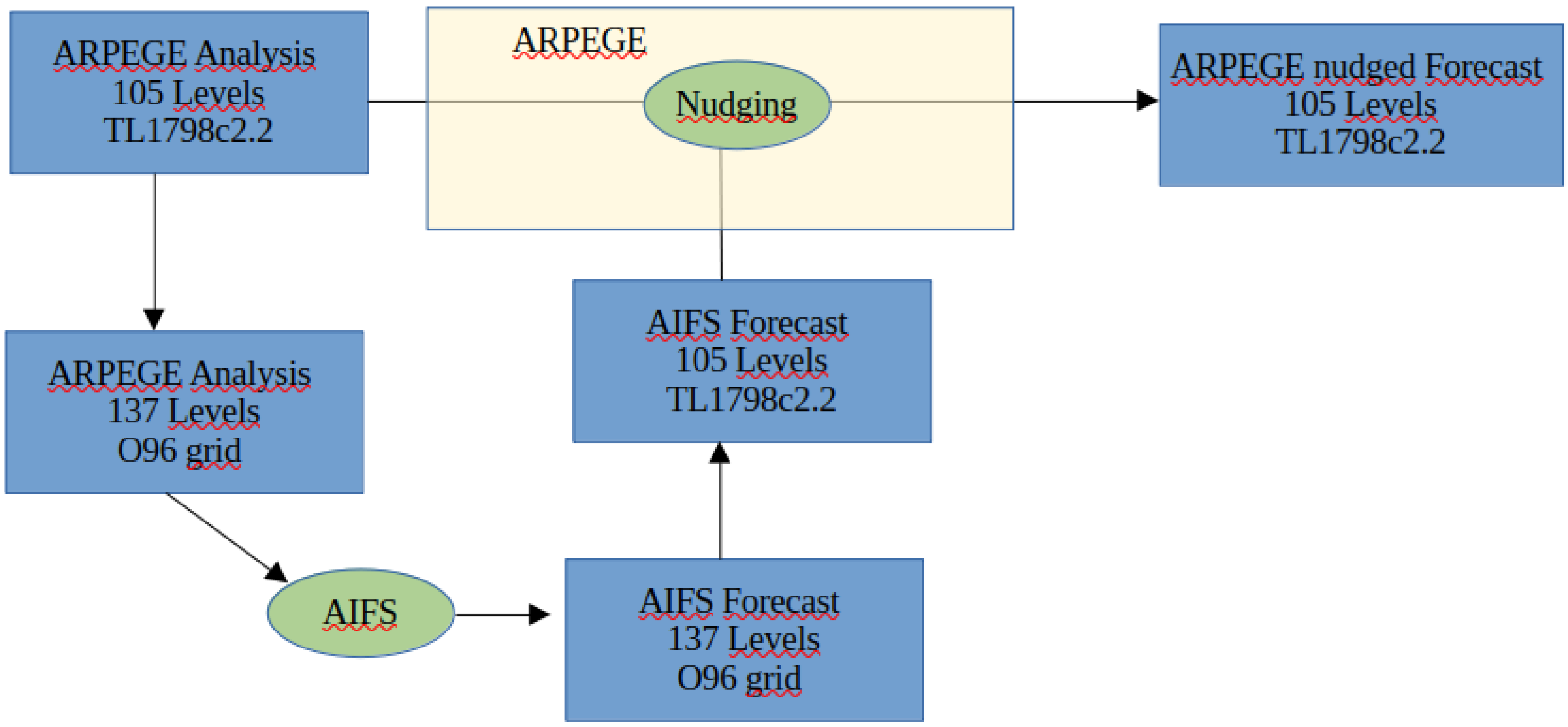
- **Operational AROME** archive (~4years)
- **ERA5-Arome :** Dynamical downscaling of ERA5 with AROME (1959-2023). In *production*.
- **ARRA :** An Arome analysis (1970-2020). *In production*.



From AI-NWP Team

Hybrid forecasting : AIFS nudging with ARPEGE

Following *Hussain et al. 2024* and ECMWF approach to perform a spectral nudging between AIFS on model level (thanks to ECWMF to provides this) initiated by ARPEGE analysis and ARPEGE forecast (streched grid).



Outline

- AI development for forecasting

- Hybrid AI-DA perspectives

- AI-DA perspectives

Hybrid AI Data Assimilation

$$J = \|x^a - x^b\|_B + \|\mathcal{H}(\mathcal{M}(x^a)) - y^o\|_R$$

- Tuning and/or filling R
- Observation operator, e.g. Radar observation operator (ANR PRAISE (submitted))
- *Using the TL/AD of a model emulator for 4DVar ?*
- *Using a model emulator to produce the ensemble forecast for 4DEnVar ?*

Observation error variances estimation

Use deep learning in order to solve the R tuning problem:

Find $\tilde{\sigma}_o = NN(X)$ with X a set of variables giving information on the error

Exploit the fact that we know/postpone :

$$d_o^b = \mathcal{H}(x^b) - y^o \sim \mathcal{N}\left(0, \sqrt{\sigma_b^2 + \sigma_o^2}\right)$$

Negative Log Likelihood (act on a sample)

$$\frac{1}{2}\left[\log \tilde{\sigma_o}^2 + \frac{d^2}{\tilde{\sigma}_o^2 + \sigma_b^2}\right]$$

Wasserstein distance (act on a batch)

Distance between the normalized departure and

$\mathcal{N}(0,1)$

Observation error variances estimation (Synthetic example)

Noise Generation (50 « global » samples)

$$X \sim Log\mathcal{N}(1, 0.8) \qquad Y \sim Log\mathcal{N}(1.2, 0.8)$$

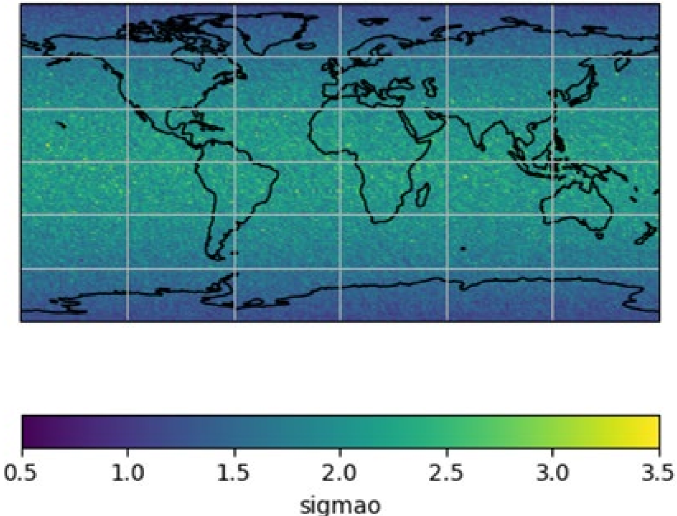
$$\sigma_o = 0.2 + \frac{\alpha}{20} f(lat) \sqrt{X \frac{(X+Y)}{2} (|X-Y| + 0.1)}$$

$$\sigma_b = 0.1 + \frac{\beta}{20} f(lon) \sqrt{X^2 \left(\left| X - \frac{Y}{2} \right| + 0.1 \right)}$$

$$\alpha \sim U[0.5, 2] \qquad \beta \sim U[0.5, 2]$$

$$d_o^b \sim \mathcal{N}(0, \sqrt{\sigma_o^2 + \sigma_b^2})$$

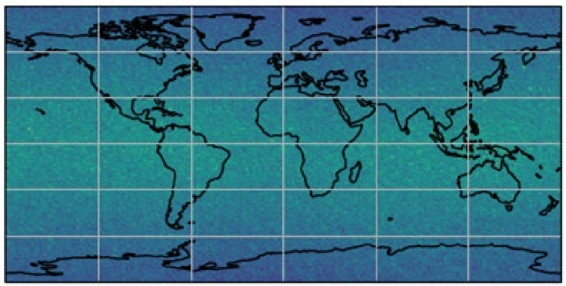
Spatial structure (mean over samples)

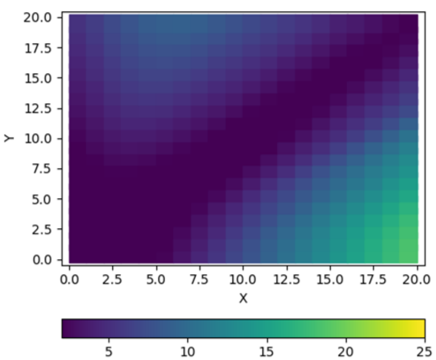


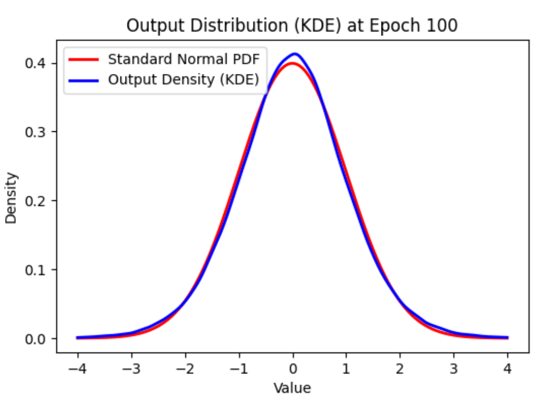
Neural Network inputs :
- X, Y
- cos(lat), sin(lat), cos(lon), sin(lon)

Observation error variances estimation (Synthetic example)

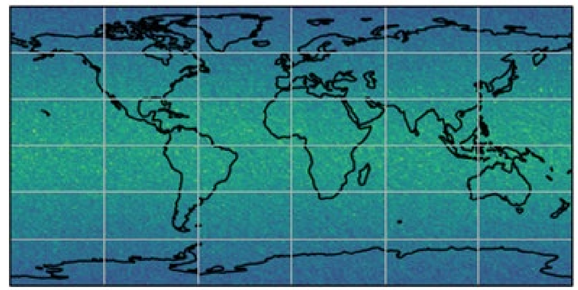
Wasserstein

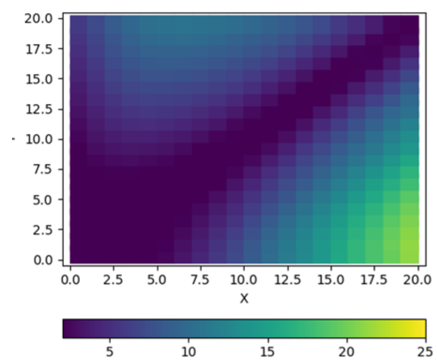


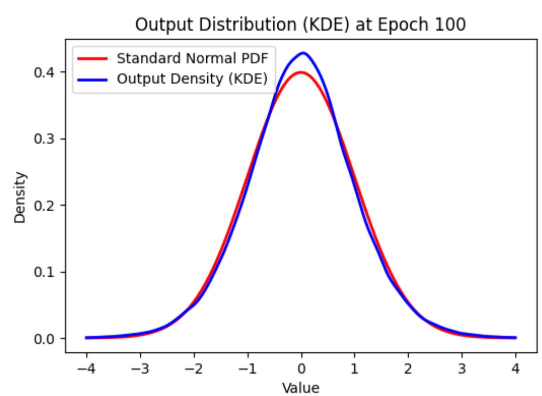




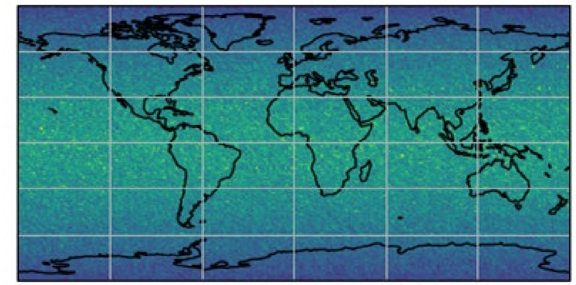
Negative Log Likelihood

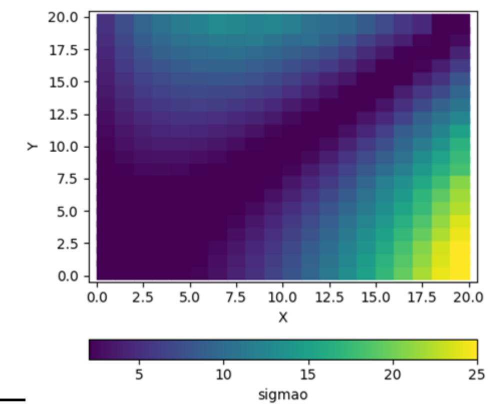






True Mean





Mean structure in XY space

	MSE	$\left\|ln\left(\frac{\sigma_o}{\tilde{\sigma}_o+\epsilon}\right)\right\|$
Wasserstein	1.62	0.23
NLLL	1.35	0.22
MSE (TRUE)	1.18	0.31

Observation error variances estimation

Plan to try this with real data:

- For tuning NEMOVar system (with M. Chrust)
-
- For Land Surface Temperature observation (S. Marimbordes)
- For microwave observation (all-sky) (A. Dauriac Internship, just starting)
- How to account for crude background error variances estimates in observation space or too large wrt first guess departure ?

Hybrid AI Data Assimilation

$$J = \|x^a - x^b\|_B + \|\mathcal{H}(\mathcal{M}(x^a)) - y^o\|_R$$

- Tuning and/or filling R
- Observation operator, e.g. Radar observation operator (ANR PRAISE (submitted))
- *Using the TL/AD of a model emulator for 4DVar ?*
- *Using a model emulator to produce the ensemble forecast for 4DEnVar ?*

TL control for Lorenz63 (Boely internship)

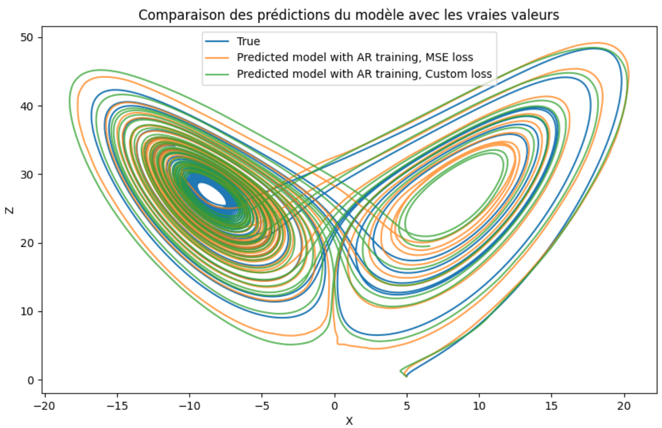
Add a constraint on the Tangent Linear of the emulator $\tilde{M}_x$ when learning the non-linear emula

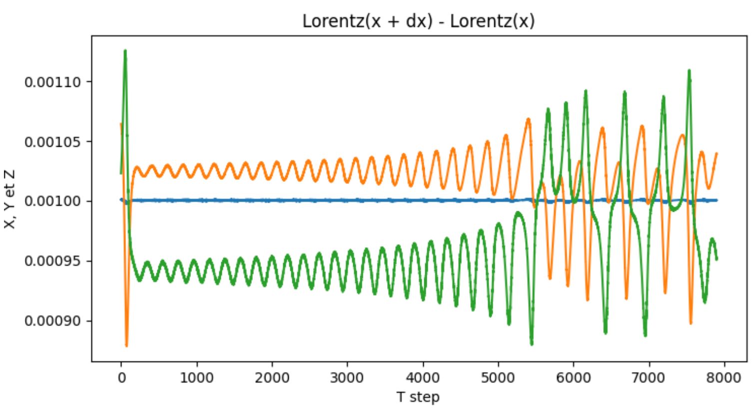
$\tilde{\mathcal{M}}$

$$\alpha \| \tilde{M}_x \, dx - (\mathcal{M}(x+dx) - \mathcal{M}(x)) \|$$

$\tilde{\mathcal{M}}$: Feed forward neural network α: weight in the cost function

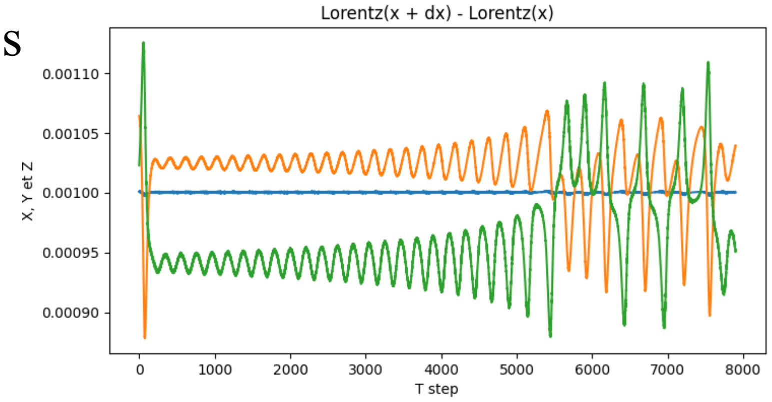
Exemple (Lorenz63)



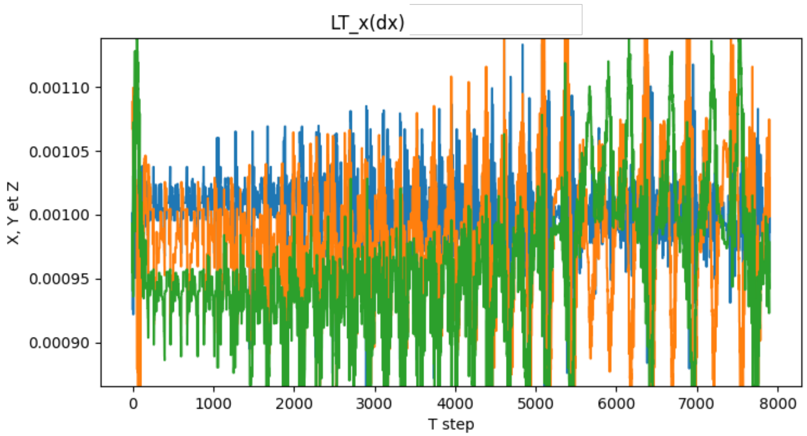


Impact of the penalty term

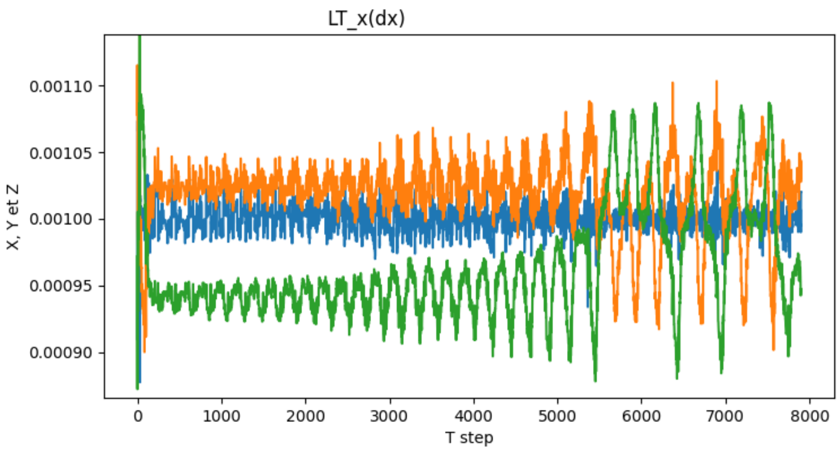
dx is fixed (over the trajectory) in order to see s



Tangent Linear without constraint



With constraint



What we learned (in this simple case)

- Constraint on TL needed since the begining (Finetuning does not work well)
- Relaxing the constraint on TL during the minimsation leads to the bests results.

=> α decrease up to 0.

- Better Lorenz63 emulator (wrt MSE criterion) when adding the TL penalization term
-

Can this be generalized to (some) real application ?

TL control in ANEMOI (Just starting)

Penalty term memory and GPU costs are dependent of the model (architecture at least).
With the current implementation (a few tricks) and with O96 grid, training with the penalty term

- x3 in GPU RAM
- x2 in GPU Cost

Determinist

$$\alpha \| \tilde{M}_x \, dx - (\mathcal{M}(x+dx) - \mathcal{M}(x)) \|$$

Analysis perturbation

First EDA member (recentered)

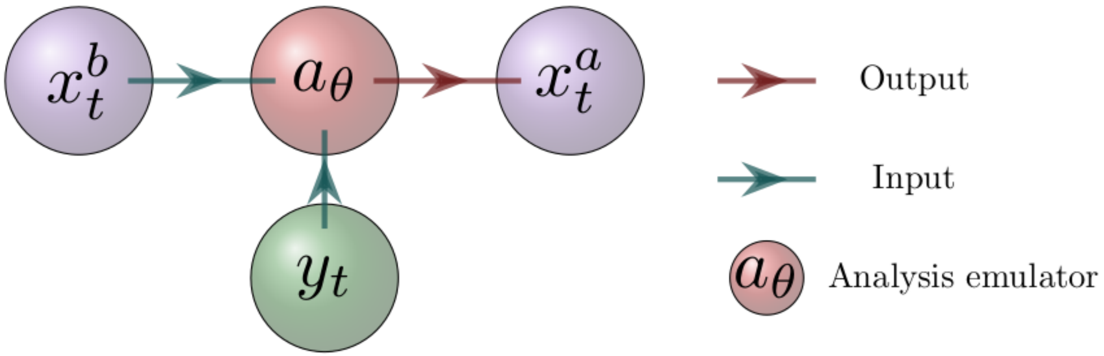
Outline

- AI development for forecasting

- Hybrid AI-DA perspectives

- AI-DA perspectives

Analysis emulation (Ongoing internship)

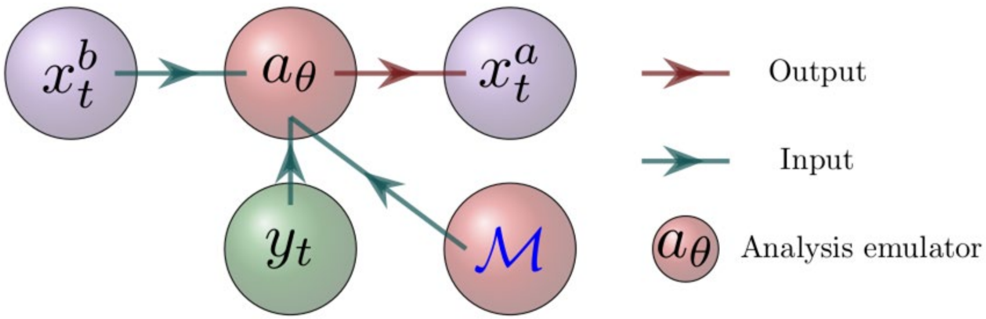


$$\left\| NN_{assim} \begin{pmatrix} y^o \\ x_0^b \end{pmatrix} - x_0^a \right\|^2$$

Dataset 3 years of hourly low resolution for AROME (5.2 km) :

- AROME 3DVar analysis and background fields
- SEVIRI data (without thinning)

Analysis emulation : Changing the view point



$$\|NN_{assim}\begin{pmatrix} y^o \\ x_0^b \end{pmatrix} - x_0^a\|^2$$

$$+\sum_{k=1}^{N}\|\mathcal{M}^k_{AROME}(NN_{assim}\begin{pmatrix} y^o \\ x_0^b \end{pmatrix}) - x_k^a\|^2$$

with $\mathcal{M}^k_{AROME}$ an AI-AROME model.

Is it possible to find an **analysis state** which provides the best forecast for a given emulator ?

Follow and evaluate emerging approaches developed by our academic partners, such as Data Assimilation Network (**DRUIDS** project, **LITCHI** project).

Conclusion / Perspective

- Investigate Hybrid AI-DA and AI-DA perspectives
- Plan to use ANEMOI when we can (e.g. analysis emulation, or TL/AD of emulator)
- AI-DA workforce at Météo-France may mostly comes from research projects
- We will also follow and evaluate theoretical developments from our academic partners.