

Quantifying uncertainty in climate reanalyses

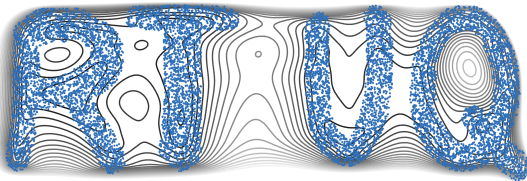
Pierre Tandeo, Jules Guillot, Pierre Ailliot,
Rakesh Teja Konduru, Takemasa Miyoshi

IMT Atlantique, Brest, France

Workshop on **Uncertainty Quantification for Climate Science**

November 13, 2025

Institut Henri Poincaré, Paris, France

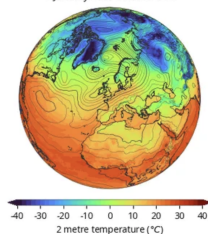


Context of the climate reanalyses

What are climate reanalyses?

- ERA, NCEP, GLORYS, etc.
- state of the atmosphere or the ocean
- over a long period of time
- used in a lot of studies

ERA5 2 metre temperature and Mean sea level pressure
1 January 2023 at 00:00 UTC

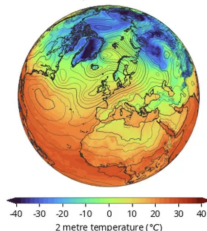


Context of the climate reanalyses

What are climate reanalyses?

- ERA, NCEP, GLORYS, etc.
- state of the atmosphere or the ocean
- over a long period of time
- used in a lot of studies

ERA5 2 metre temperature and Mean sea level pressure
1 January 2023 at 00:00 UTC



How are they computed?

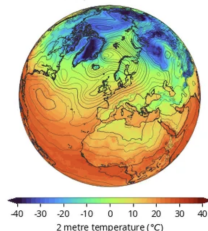
- merge physical model and available observations
- based on data assimilation algorithms
- output the ensemble mean + ensemble spread

Context of the climate reanalyses

What are climate reanalyses?

- ERA, NCEP, GLORYS, etc.
- state of the atmosphere or the ocean
- over a long period of time
- used in a lot of studies

ERA5 2 metre temperature and Mean sea level pressure
1 January 2023 at 00:00 UTC



How are they computed?

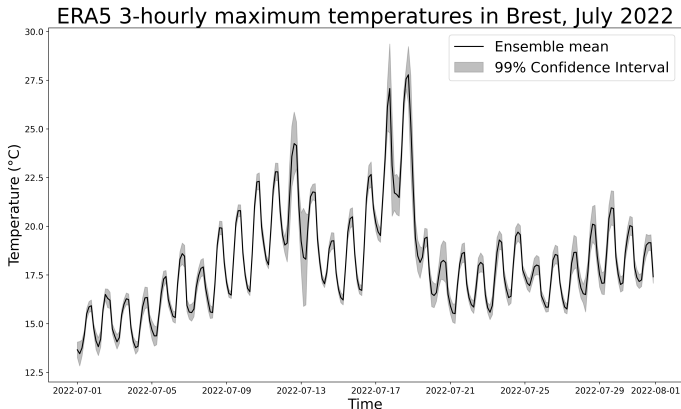
- merge physical model and available observations
- based on data assimilation algorithms
- output the ensemble mean + ensemble spread

⇒ **Are climate reanalyses accurate (mean) and reliable (spread)?**

Example of ERA5 in Brest

Is ERA5 reliable?

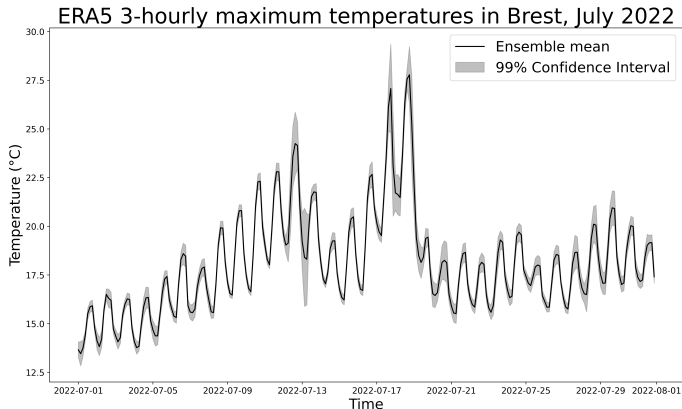
- ground temperature in Brest, France
- during strong heat wave in 2022
- 39.3 degrees on July 18



Example of ERA5 in Brest

Is ERA5 reliable?

- ground temperature in Brest, France
- during strong heat wave in 2022
- 39.3 degrees on July 18



⇒ **39.3 is not included (at all) in the 99% confidence interval**

Observation, model, and state:

- we use observations \mathbf{y} (partial, noisy)
- we use a model \mathcal{M} (general circulation model)
- we want to estimate the state \mathbf{x} (everywhere, anytime)

Observation, model, and state:

- we use observations \mathbf{y} (partial, noisy)
- we use a model \mathcal{M} (general circulation model)
- we want to estimate the state \mathbf{x} (everywhere, anytime)

State-space model:

$$\begin{aligned}\mathbf{x}(t + dt) &= \mathcal{M}(\mathbf{x}(t)) + \boldsymbol{\eta}(t) \\ \mathbf{y}(t) &= \mathbf{H}\mathbf{x}(t) + \boldsymbol{\epsilon}(t)\end{aligned}$$

Error terms:

$$\begin{aligned}\boldsymbol{\eta}(t) &\sim \mathcal{N}(\mathbf{0}, \mathbf{Q}(t)) \\ \boldsymbol{\epsilon}(t) &\sim \mathcal{N}(\mathbf{0}, \mathbf{R}(t))\end{aligned}$$

Data assimilation formulation

Observation, model, and state:

- we use observations \mathbf{y} (partial, noisy)
- we use a model \mathcal{M} (general circulation model)
- we want to estimate the state \mathbf{x} (everywhere, anytime)

State-space model:

$$\begin{aligned}\mathbf{x}(t + dt) &= \mathcal{M}(\mathbf{x}(t)) + \boldsymbol{\eta}(t) \\ \mathbf{y}(t) &= \mathbf{H}\mathbf{x}(t) + \boldsymbol{\epsilon}(t)\end{aligned}$$

Error terms:

$$\begin{aligned}\boldsymbol{\eta}(t) &\sim \mathcal{N}(\mathbf{0}, \mathbf{Q}(t)) \\ \boldsymbol{\epsilon}(t) &\sim \mathcal{N}(\mathbf{0}, \mathbf{R}(t))\end{aligned}$$

⇒ **Calibration of Q and R covariance matrices is a key point**¹

¹Dee, D. P. (1995). On-line Estimation of Error Covariance Parameters for Atmospheric Data Assimilation. *Monthly Weather Review*

Illustrative example

Equations:

$$x(t+1) = 0.95x(t) + \eta(t)$$

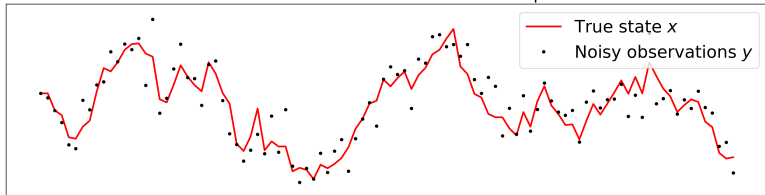
$$y(t) = x(t) + \epsilon(t)$$

Error terms:

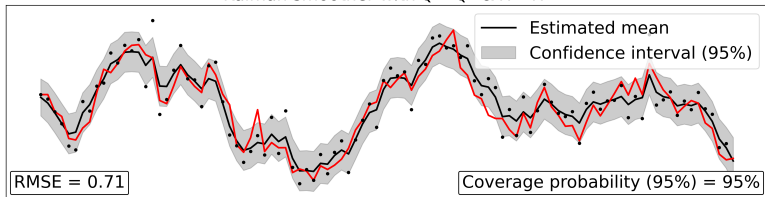
$$\eta(t) \sim \mathcal{N}(0, Q^t = 1)$$

$$\epsilon(k) \sim \mathcal{N}(0, R^t = 1)$$

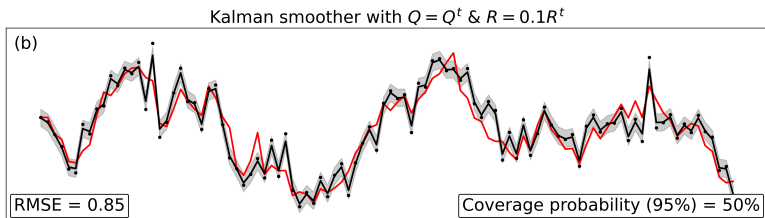
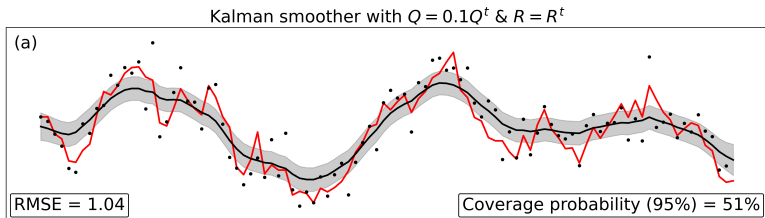
Simulated data from a linear Gaussian state-space model



Kalman smoother with $Q = Q^t$ & $R = R^t$



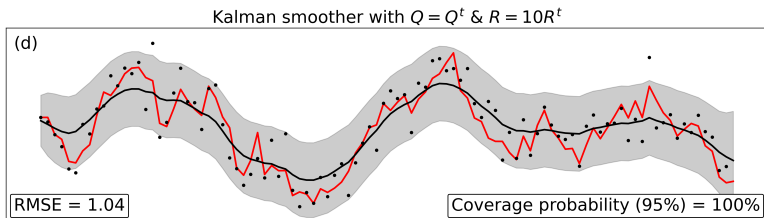
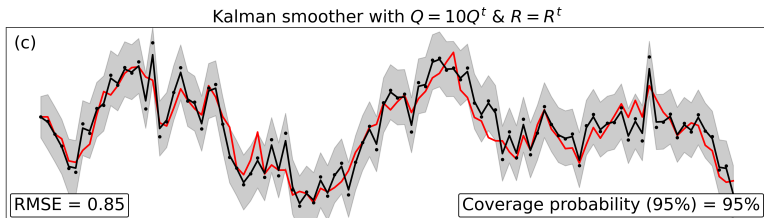
Bad Q/R ratio (under-calibration of Q or R)



Bad Q/R ratio:

- (a) $Q/R = 0.1 \Rightarrow$ **Too confident in the model**
- (b) $Q/R = 10. \Rightarrow$ **Too confident in the observations**

Bad Q/R ratio (over-calibration of Q or R)

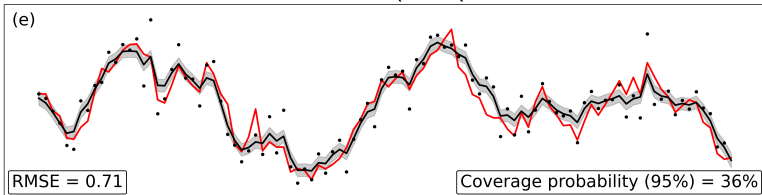


Bad Q/R ratio:

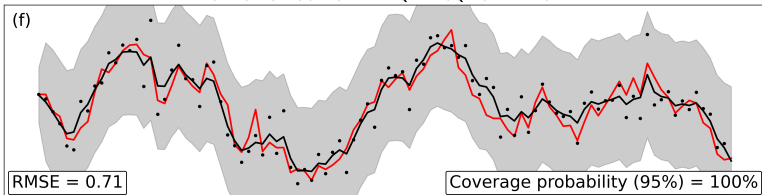
- (c) $Q/R = 10$. \Rightarrow **Same RMSE as figure (b)**
- (d) $Q/R = 0.1$ \Rightarrow **Same RMSE as figure (a)**

Good Q/R ratio

Kalman smoother with $Q = 0.1Q^t$ & $R = 0.1R^t$



Kalman smoother with $Q = 10Q^t$ & $R = 10R^t$



Good Q/R ratio:

- (e) Q and R too low \Rightarrow **Overconfident in the results**
- (f) Q and R too high \Rightarrow **Underconfident in the results**

Estimation of both Q and R

State-of-the-art paper¹:

- review methods from different communities
- clear mathematical formulation and comparison



¹Tandeo et al. (2020). A Review of Innovation-based Methods to Jointly Estimate Model and Observation Error Covariance Matrices in Ensemble Data Assimilation. *Monthly Weather Review*

Maximization of the innovation likelihood

Definition of the innovation likelihood:

$$\mathcal{L}(\mathbf{y}(t)) = \mathcal{N}\left(0; \mathbf{y}(t) - \mathbf{H}\mathbf{x}^f(t), \mathbf{H}\mathbf{P}^f(t)\mathbf{H}^\top + \mathbf{R}(t)\right)$$

with:

- $\mathbf{x}^f(t)$ the mean and $\mathbf{P}^f(t)$ the covariance of the forecast
- $\mathbf{P}^f(t)$ is depending on the model error covariance $\mathbf{Q}(t)$

Maximization of the innovation likelihood

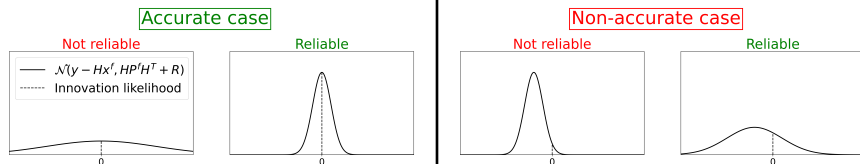
Definition of the innovation likelihood:

$$\mathcal{L}(\mathbf{y}(t)) = \mathcal{N}\left(0; \mathbf{y}(t) - \mathbf{H}\mathbf{x}^f(t), \mathbf{H}\mathbf{P}^f(t)\mathbf{H}^\top + \mathbf{R}(t)\right)$$

with:

- $\mathbf{x}^f(t)$ the mean and $\mathbf{P}^f(t)$ the covariance of the forecast
- $\mathbf{P}^f(t)$ is depending on the model error covariance $\mathbf{Q}(t)$

Graphical representation of the innovation likelihood:



Maximization of the innovation likelihood

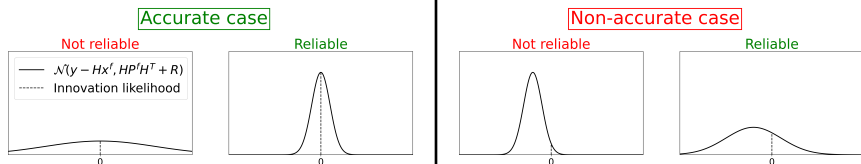
Definition of the innovation likelihood:

$$\mathcal{L}(\mathbf{y}(t)) = \mathcal{N}\left(0; \mathbf{y}(t) - \mathbf{H}\mathbf{x}^f(t), \mathbf{H}\mathbf{P}^f(t)\mathbf{H}^\top + \mathbf{R}(t)\right)$$

with:

- $\mathbf{x}^f(t)$ the mean and $\mathbf{P}^f(t)$ the covariance of the forecast
- $\mathbf{P}^f(t)$ is depending on the model error covariance $\mathbf{Q}(t)$

Graphical representation of the innovation likelihood:



⇒ **The innovation likelihood $\mathcal{L}(\mathbf{y}(t))$ depends on $\mathbf{Q}(t)$ and $\mathbf{R}(t)$**

Inflation and localization of \mathbf{P}^f

Issues:

- estimate \mathbf{Q} and \mathbf{R} jointly is a complicated task
- need several iterations of data assimilation (e.g., EM algorithm)

Inflation and localization of \mathbf{P}^f

Issues:

- estimate \mathbf{Q} and \mathbf{R} jointly is a complicated task
- need several iterations of data assimilation (e.g., EM algorithm)

Alternative:

- inflate (λ) and localize (l) the covariance \mathbf{P}^f
- using a $L(\lambda, l)$ matrix and Schur product

Inflation and localization of \mathbf{P}^f

Issues:

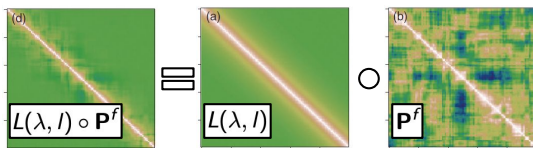
- estimate \mathbf{Q} and \mathbf{R} jointly is a complicated task
- need several iterations of data assimilation (e.g., EM algorithm)

Alternative:

- inflate (λ) and localize (l) the covariance \mathbf{P}^f
- using a $L(\lambda, l)$ matrix and Schur product

Formula:

$$\mathbf{P}^f(t) \leftarrow L(\lambda, l) \circ \mathbf{P}^f(t)$$



Inflation and localization of \mathbf{P}^f

Issues:

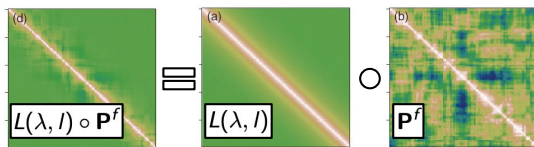
- estimate \mathbf{Q} and \mathbf{R} jointly is a complicated task
- need several iterations of data assimilation (e.g., EM algorithm)

Alternative:

- inflate (λ) and localize (l) the covariance \mathbf{P}^f
- using a $L(\lambda, l)$ matrix and Schur product

Formula:

$$\mathbf{P}^f(t) \leftarrow L(\lambda, l) \circ \mathbf{P}^f(t)$$



\Rightarrow **The goal is to find, at each time t , $\hat{\lambda}(t), \hat{l}(t) = \operatorname{argmax}_{\lambda(t), l(t)} (\mathcal{L}(\mathbf{y}(t)))$**

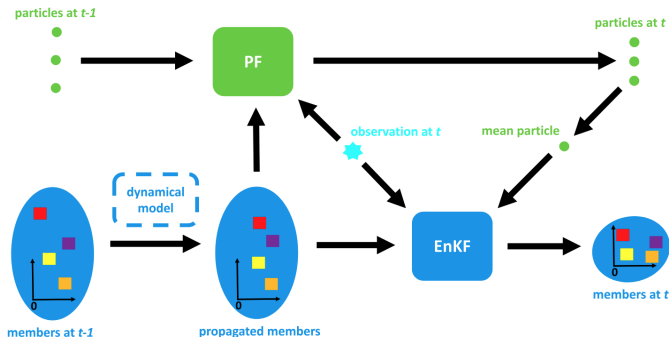
Combination of two data assimilation systems (PF-EnKF¹)

Ensemble Kalman Filter (EnKF):

- for the state (\mathbf{x})
- provide an estimate of $\mathbf{P}^f(t)$

Particle Filter (PF):

- for the parameters (λ, l)
- weights prop. to $\mathcal{L}(\mathbf{y}(t))$

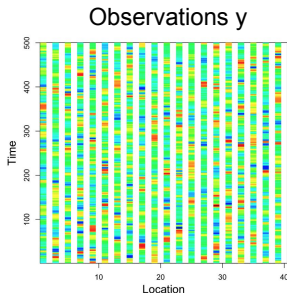
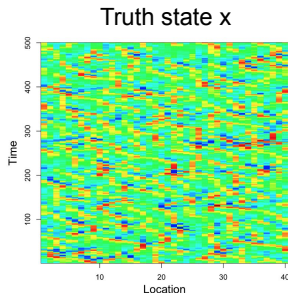


¹Guillot et al. (2025). State and Stochastic Parameters Estimation with Combined Ensemble Kalman and Particle Filters. *Monthly Weather Review*

Application to the Lorenz-96 system

Lorenz-96 equations:

$$\frac{dx_k}{dt} = (x_{k+1} - x_{k-2})x_{k-1} - x_k + 8, \quad \forall k = 1, \dots, 40$$



State-space model:

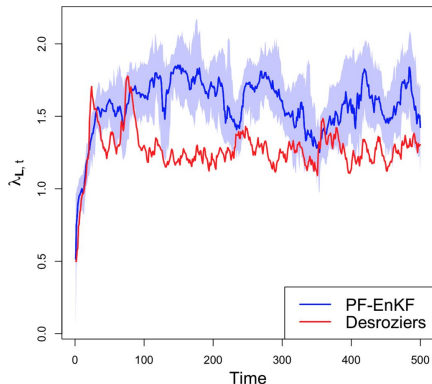
$$\begin{aligned}\mathbf{x}(t + dt) &= \mathcal{M}(\mathbf{x}(t)) + \boldsymbol{\eta}(t) \\ \mathbf{y}(t) &= \mathbf{H}\mathbf{x}(t) + \boldsymbol{\epsilon}(t)\end{aligned}$$

With:

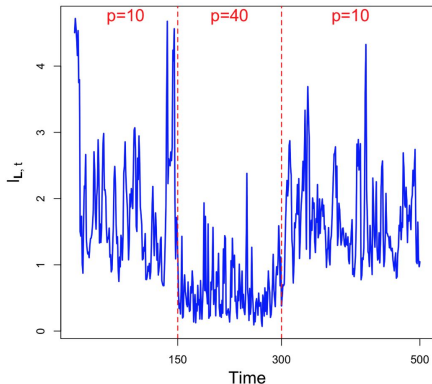
- predefined \mathbf{Q} and \mathbf{R} matrices
- time-varying number of obs.

Estimation of inflation (λ) and localization (l) parameters

Inflation parameter



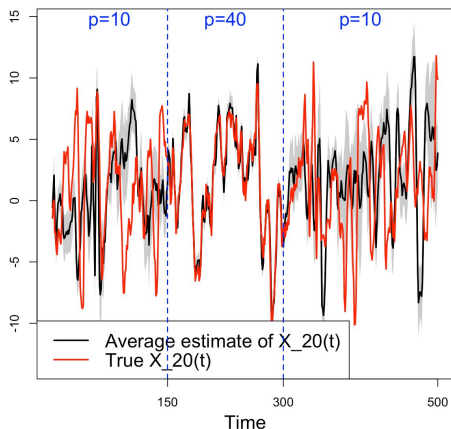
Localization parameter



⇒ **PF-EnKF is able to estimate time-varying parameters**

Impact of PF-EnKF on the estimation of the state

Results of the PF-EnKF algorithm



When $p = 40$ obs. variables:

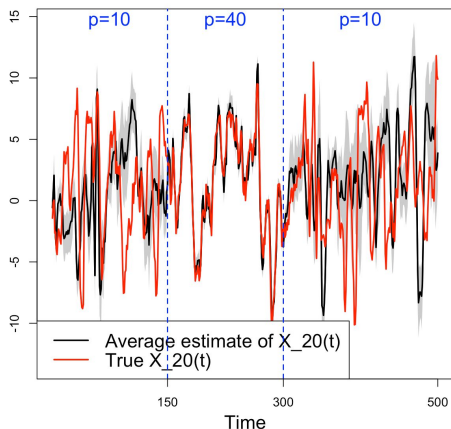
- good accuracy
- small confidence interval

When $p = 10$ obs. variables:

- bad accuracy
- big confidence interval

Impact of PF-EnKF on the estimation of the state

Results of the PF-EnKF algorithm



When $p = 40$ obs. variables:

- good accuracy
- small confidence interval

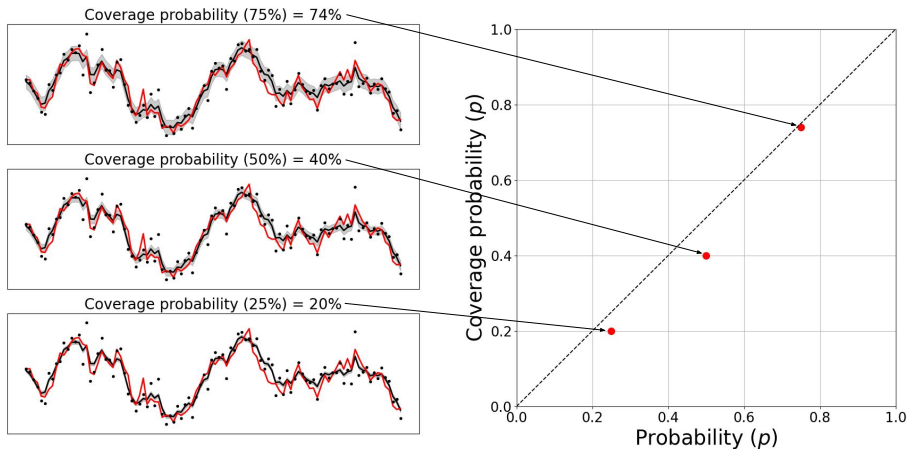
When $p = 10$ obs. variables:

- bad accuracy
- big confidence interval

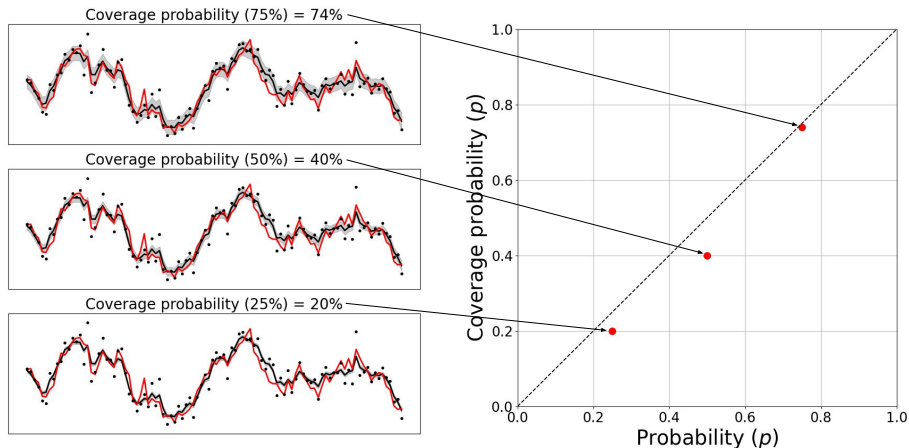
Overall: Cov. prob. (70%) \approx 70%

⇒ **Uncertainty is well calibrated (good coverage probability)**

Confidence intervals and coverage probabilities



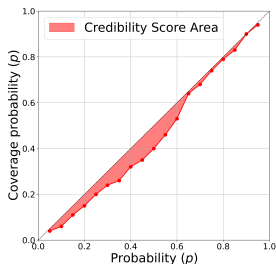
Confidence intervals and coverage probabilities



⇒ **The coverage probability should equal the probability of the corresponding confidence interval**

New metric for uncertainty quantification: credibility score¹

Computation and visualization of the Credibility Score (CS):



Formula:

$$CS = \int_0^1 (\text{cov. prob.}(p) - p) dp$$

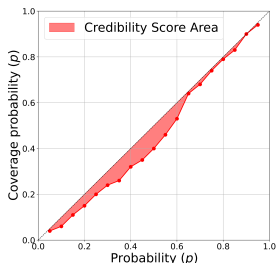
Property:

$$-0.5 \leq CS \leq 0.5$$

¹Konduru, Miyoshi, and Tandeo (in preparation)

New metric for uncertainty quantification: credibility score¹

Computation and visualization of the Credibility Score (CS):



Formula:

$$CS = \int_0^1 (\text{cov. prob.}(p) - p) dp$$

Property:

$$-0.5 \leq CS \leq 0.5$$

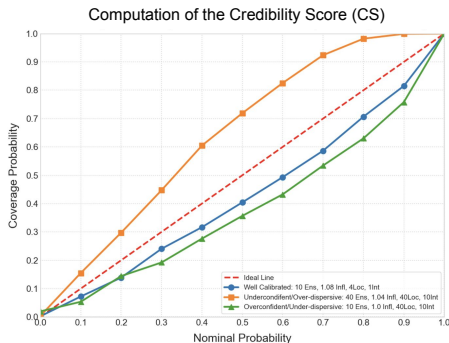
Benefits of the credibility score:

- non-parametric (no need Gaussian assumptions)
- easy to compute (simple counting, variable by variable)
- interpretable ($CS < 0$ for overconfident, $CS > 0$ for underconfident)

¹Konduru, Miyoshi, and Tandeo (in preparation)

Application to the Lorenz-96 system

Experiment	#1	#2	#3
Ens. Memb.	10	40	10
Inflation	1.08	1.04	no
Localization	yes	no	no
CRPS	2.04	2.14	2.16
CS	-0.07	0.14	-0.11



⇒ **Credibility score gives information on accuracy and reliability**

Importance of \mathbf{Q} and \mathbf{R} in reanalysis:

- balance between model and observations
- extremely important, but not easy to calibrate jointly

Importance of \mathbf{Q} and \mathbf{R} in reanalysis:

- balance between model and observations
- extremely important, but not easy to calibrate jointly

PF-EnKF algorithm:

- online estimation of inflation and localization of \mathbf{P}^f
- low-cost produce that can be plugged to a reanalysis algorithm

Conclusions and perspectives

Importance of \mathbf{Q} and \mathbf{R} in reanalysis:

- balance between model and observations
- extremely important, but not easy to calibrate jointly

PF-EnKF algorithm:

- online estimation of inflation and localization of \mathbf{P}^f
- low-cost produce that can be plugged to a reanalysis algorithm

Credibility score:

- new metric to quantify uncertainty (overconfidence, underconfidence)
- could be useful as loss function in ML and AI models

Merci pour votre attention !

Interested in this topic? Collaborations are welcome!

Contact:

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

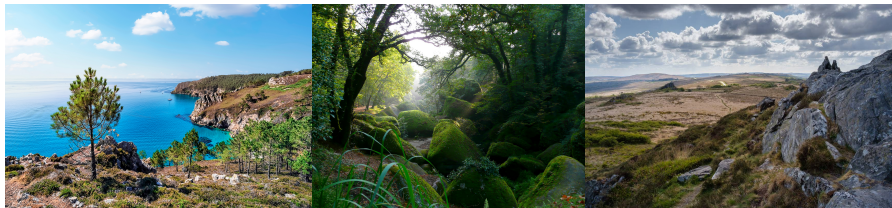


Figure: Nice places to visit near Brest, Bretagne, France