Réunion ANR Melody

Analog data assimilation of 1D extremes

Paul Platzer
IMT-Atlantique, LSCE, FEM

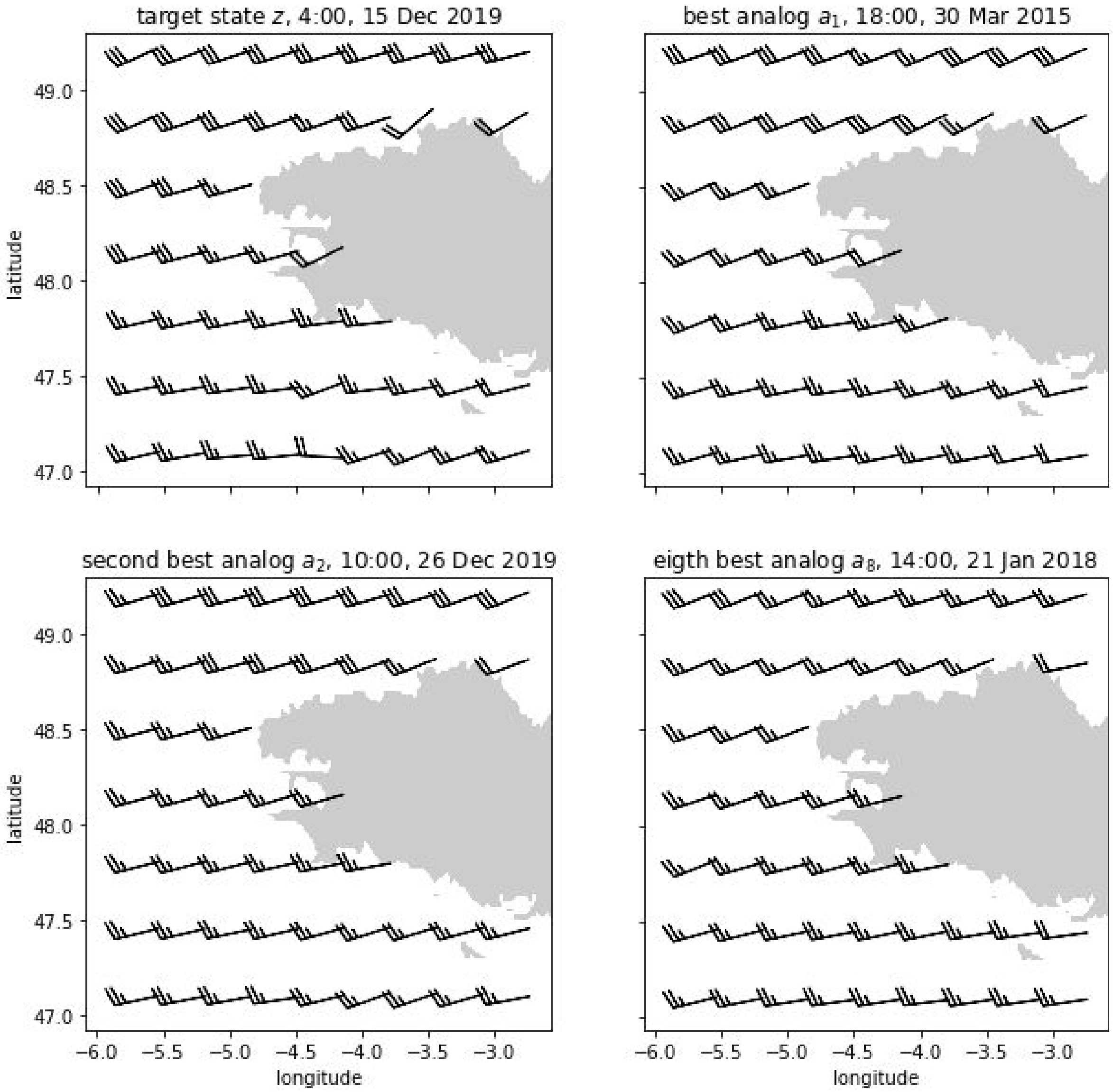
Joint work with Philippe Naveau, Jean-François Filipot, Pierre Tandeo, Pierre Ailliot, Nicolas Raillard, Pascal Yiou

What are analogs ?

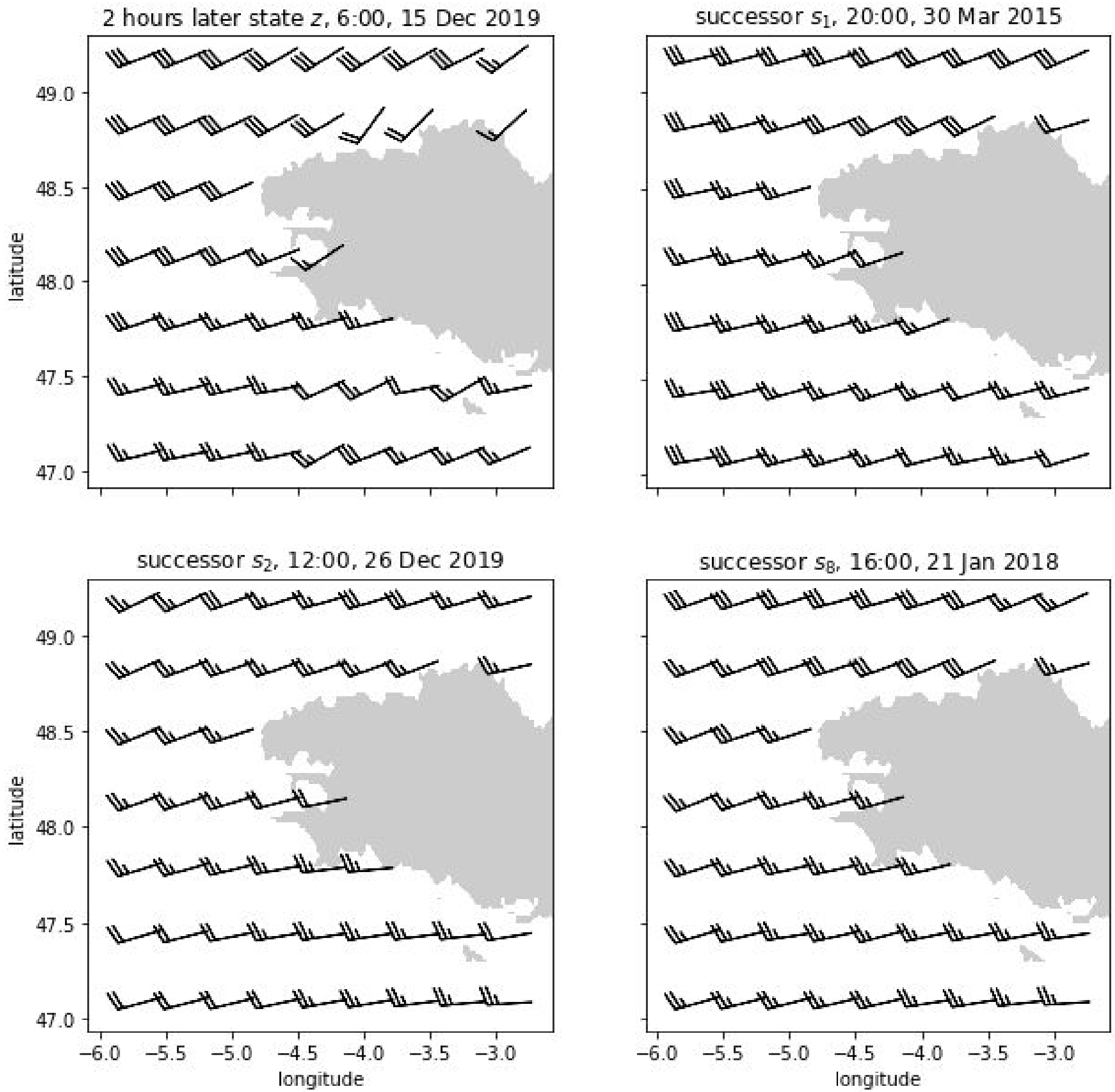
- “k nearest neighbors” in geosciences
  → dynamical systems, recurrences

What are they good for ?

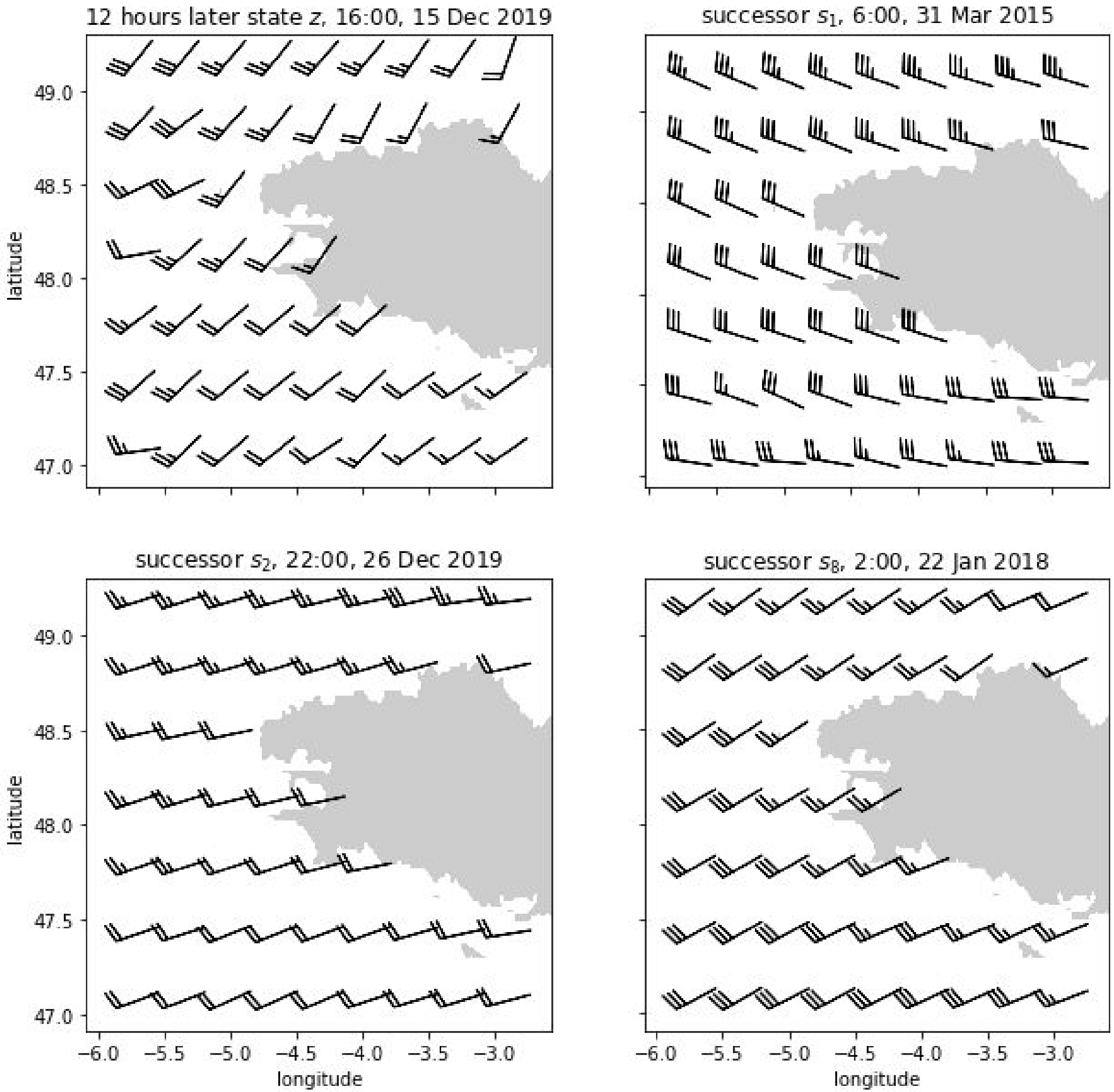
- Forecast (ensembles)
- Downscaling, upscaling
- *Attribution* of extreme events
- Predictability estimation



target state z, 4:00, 15 Dec 2019
best analog a_1, 18:00, 30 Mar 2015
second best analog a_2, 10:00, 26 Dec 2019
eigth best analog a_8, 14:00, 21 Jan 2018
latitude
longitude
49.0
48.5
48.0
47.5
47.0
−6.0
−5.5
−5.0
−4.5
−4.0
−3.5
−3.0



2 hours later state z, 6:00, 15 Dec 2019
successor s_1, 20:00, 30 Mar 2015
successor s_2, 12:00, 26 Dec 2019
successor s_8, 16:00, 21 Jan 2018
latitude
longitude
49.0
48.5
48.0
47.5
47.0
−6.0
−5.5
−5.0
−4.5
−4.0
−3.5
−3.0



12 hours later state z, 16:00, 15 Dec 2019
successor s_1, 6:00, 31 Mar 2015
successor s_2, 22:00, 26 Dec 2019
successor s_8, 2:00, 22 Jan 2018
latitude
longitude
49.0
48.5
48.0
47.5
47.0
−6.0
−5.5
−5.0
−4.5
−4.0
−3.5
−3.0

Why extremes ?

- Important consequences
- Seldom observed → dedicated statistical framework → extreme value theory

Why analogs of extremes ?

- Never done before because “it shouldn’t work”
- Rare events → few analogs → ?
- Urge to clarify why, when and how “it doesn’t work”

The experiment

- 1D state-space model with heavy-tailed variables

latent-state: model

$$\mathrm{X}_t = \mathcal{M}_t(\mathrm{X}_{t-1}),$$

observation: noise

$$\mathrm{Y}_t = \sigma \frac{\mathrm{E}_t}{\mathrm{X}_t},$$

The experiment

- <u>Latent-state</u> : Gamma-distributed

$$\mathrm{X}_t \sim \mathrm{Ga}(\xi^{-1}, \xi^{-1}) \qquad \mathrm{Corr}(\mathrm{X}_t, \mathrm{X}_s) = \rho^{|s-t|}$$

$$\mathcal{M}_t^{\mathrm{AR1}}(\mathrm{X}_{t-1}) = \rho\, \mathrm{X}_{t-1} + \eta_t\,,$$

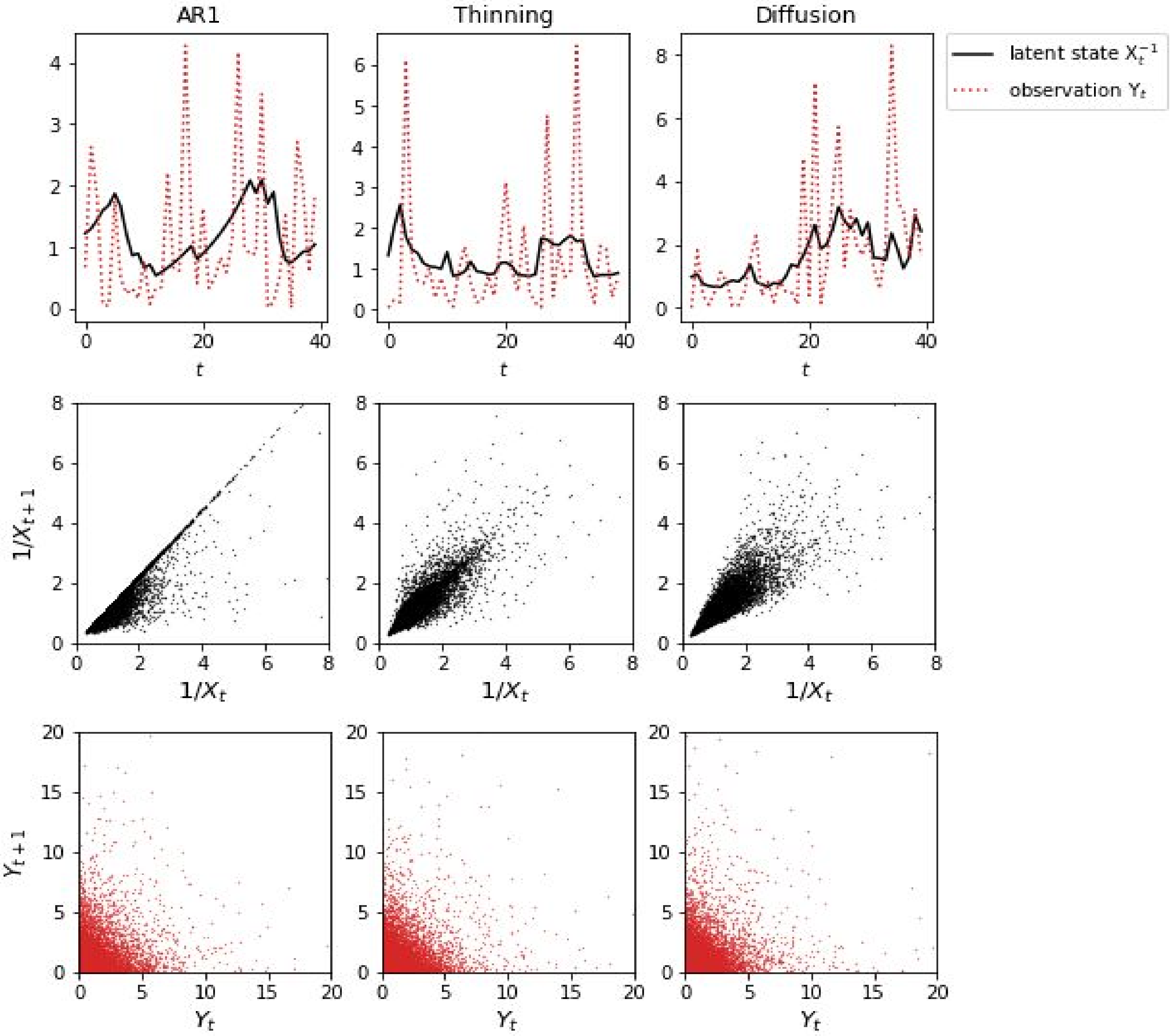
$$\mathcal{M}_t^{\mathrm{THIN}}(\mathrm{X}_{t-1}) = \mathrm{B}_t\, \mathrm{X}_{t-1} + \zeta_t\,,$$

$$\mathcal{M}_t^{\mathrm{OU}}(\mathrm{X}_{t-1}) = \mathrm{X}_{t-1} + \lambda \int_{t-1}^{t} (\mathrm{X}_s - 1)\mathrm{d}s + \sqrt{2\lambda\xi} \int_{t-1}^{t} \sqrt{\mathrm{X}_s}\mathrm{dW}_s\,,$$

The experiment

- Observations : generalized Pareto distribution

$$\begin{cases} Y_t = \sigma \dfrac{E_t}{X_t} \\ X_t \sim \mathrm{Ga}(\xi^{-1}, \xi^{-1}) \\ E_t \sim \exp(1) \end{cases} \longrightarrow \quad Y_t \sim \mathrm{GPD}(\sigma, \xi)$$


AR1
Thinning
Diffusion
latent state X_t^{-1}
observation Y_t
t
$1/X_{t+1}$
$1/X_t$
Y_{t+1}
Y_t

Objectives

- “reconstruct” the latent-state

$$\mathrm{d}\mathbb{P}(\mathrm{X}_{1:t} = x_{1:t} \,|\, \mathrm{Y}_{1:t} = y_{1:t}) \,,$$

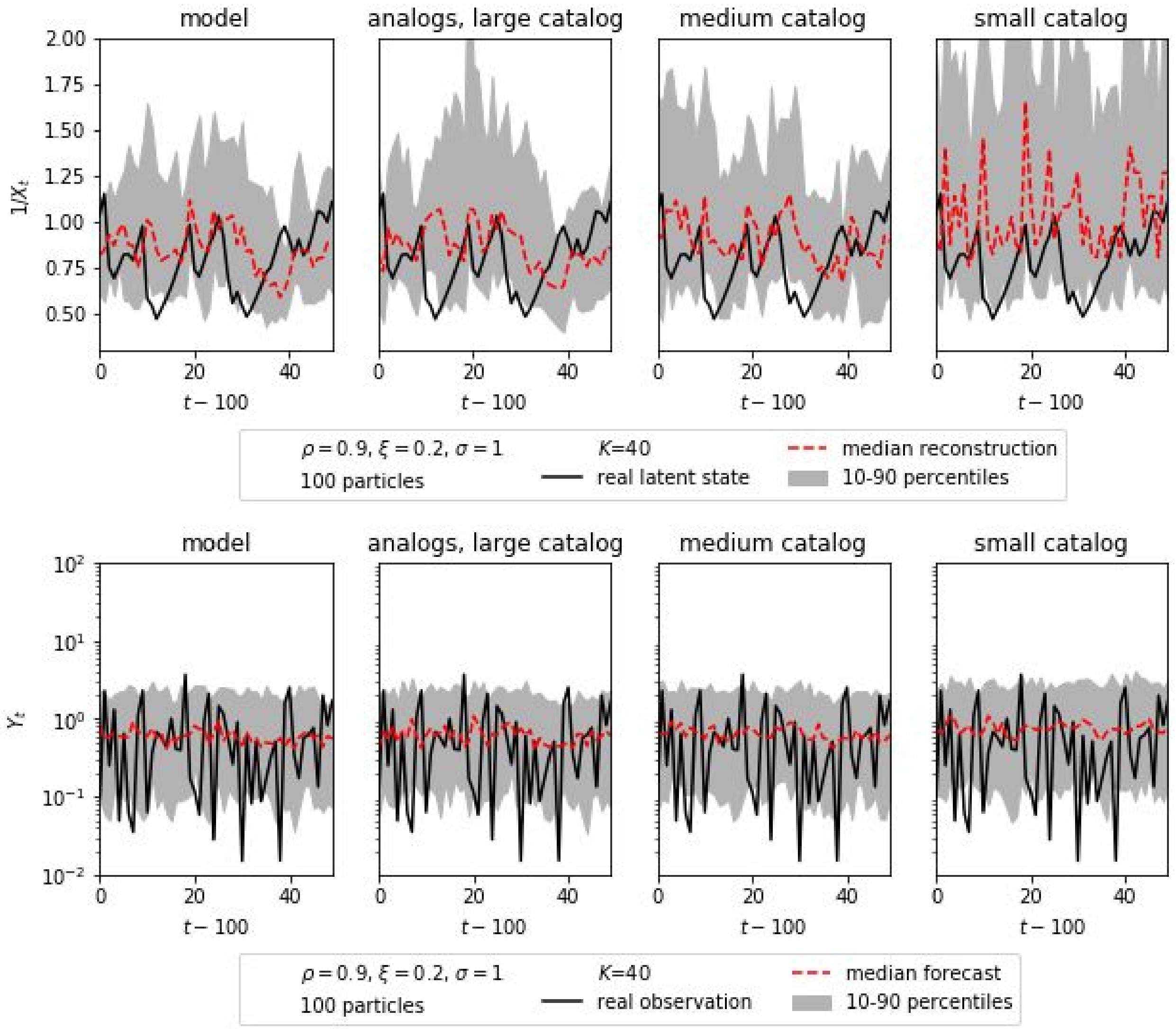
- forecast the next observation

$$\mathrm{d}\mathbb{P}(\mathrm{Y}_{t+1} = y_{t+1} \,|\, \mathrm{Y}_{1:t} = y_{1:t}) \,,$$

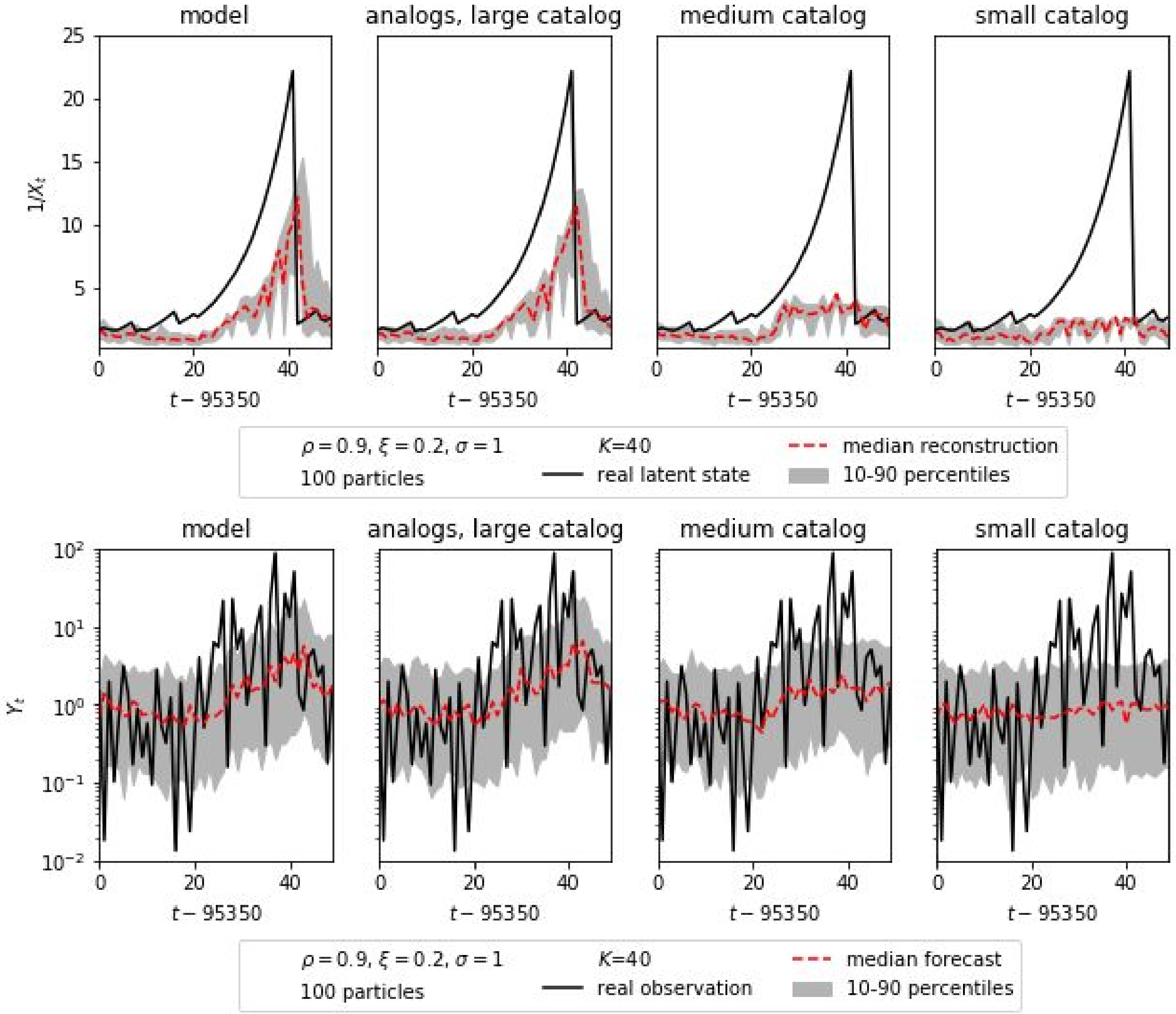
Tools

- Particle filtering or “sequential Monte-Carlo”
- Analogs with “perfect” catalog generated from the models

Example of medium values of 1/Xt



Example of high values of 1/Xt

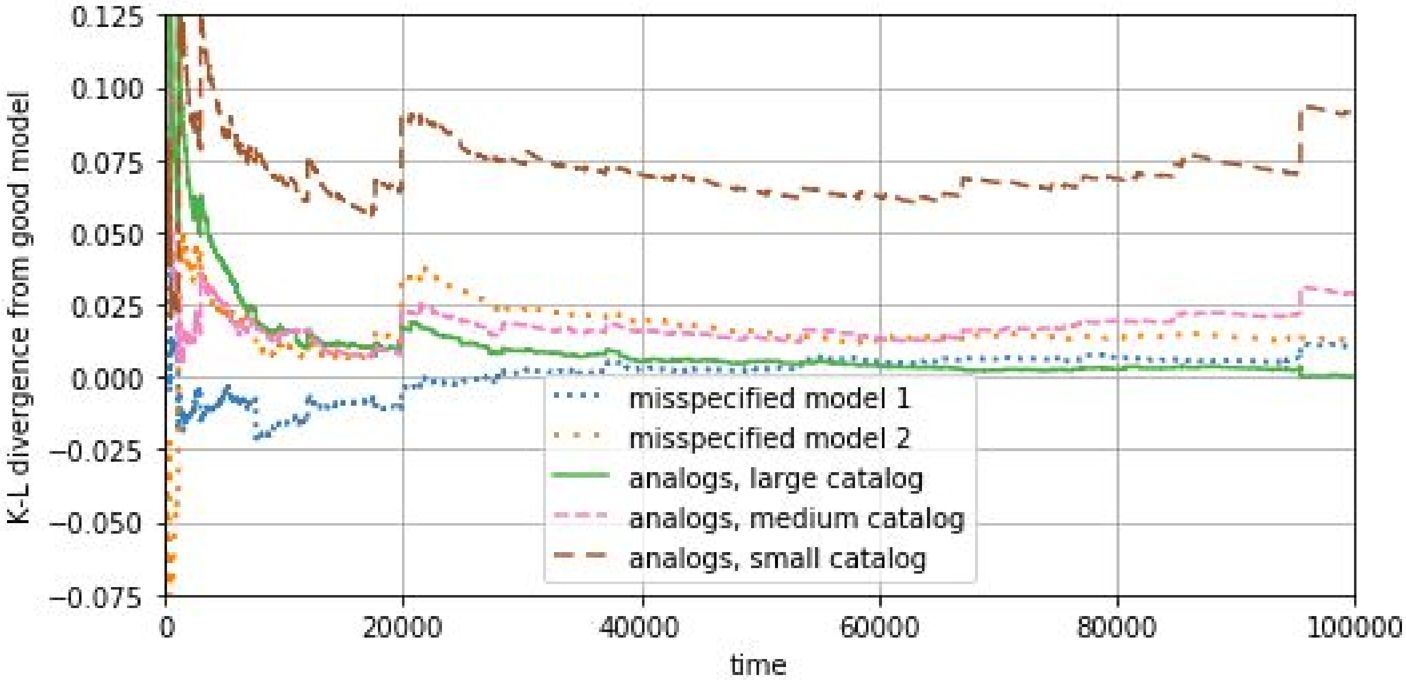


Scores

- log-likelihood $$s\left(\mathrm{d}\tilde{\mathbb{P}}\right) = \frac{1}{t}\sum_{k=1}^{t} \log \mathrm{d}\tilde{\mathbb{P}}(y_{k+1}),$$

better than good model

worse than good model



The chance to find analogs ?

best analog

target

distance

ball of radius *r* around *z*

catalog

size of catalog

$$\mathbb{P}(\|a_1 - z\| < r) = 1 - [1 - \mathbb{P}(a \in B_{z,r} \mid a \in \mathcal{C})]^L$$
$$= 1 - [1 - P_{z,r}]^L$$
$$\sim LP_{z,r}$$

<u>Dynamical systems:</u>

$$P_{z,r} = \mu(B_{z,r}) \approx \rho_z r^{d_z}$$

dimension

scale factor ~ 1

Van den Dool (1994)
Nicolis (1998)
Platzer (2021?)

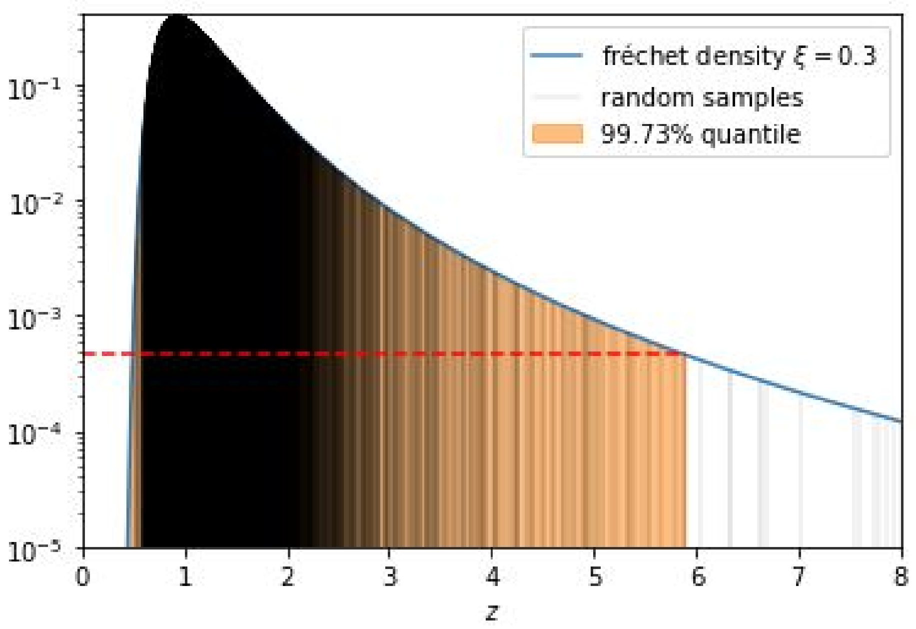
<u>One-dimensional random variable:</u>

$$P_{z,r} = \int_{z-r}^{z+r} f(u)\mathrm{d}u \approx 2rf(z)$$

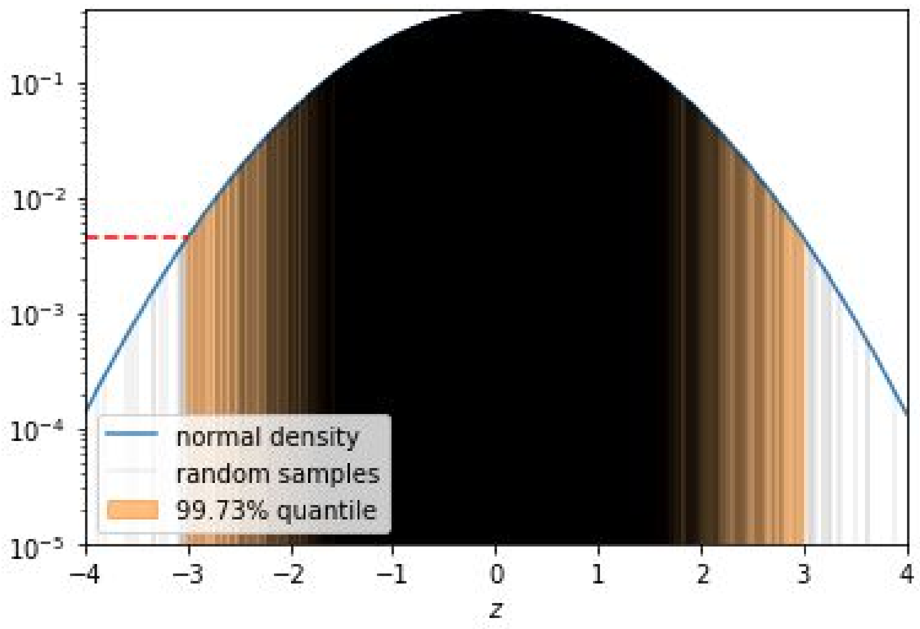
for heavy-tailed variables *f*(*z*) can be small !

The chance to find analogs ?

heavy-tailed



light-tailed



Take-home messages

- Analogs ~ monte-carlo sampling, not limited to a given family of distributions
- Analogs can forecast extreme events, but with a catalog larger than expected for “normal events”
- Dimension is not the only factor of point density

What to do next

- *Adapt* analog forecasting techniques to extreme events for a *fixed catalog length* (2.1)
- Use other machine-learning tools (NN…) with a structure adapted to state-space models, and a cost-function adapted to extremes (2.2)
- Real data