Rare event simulations in climate dynamics

Freddy Bouchet (ENS de Lyon and CNRS)

With:

a) Francesco Ragone and Jeroen Wouters (rare event algorithms in climate models)
b) Dario Lucente, George Miloshevich and Francesco Ragone (extreme heat waves)
c) Dario Lucente and Corentin Herbert (committor functions for ENSO)
d) Valerian Jacques-Dumas, Francesco Ragone, Pierre Borgnat and Patrice Abry (committor functions: prediction of extreme heat waves with deep neural networks)

RESIM, Paris, May 2021

I) Introduction

Large deviation theory and study of rare events

- **III a) Path large deviations for kinetic theories.**

F. Bouchet, J. Stat. Phys., 2020: path large deviations for the Boltzmann equation and the irreversibility paradox.

O. Feliachi and F. Bouchet, sub. to J. Stat. Phys., 2020: path large deviations for the plasma and the Vlasov equation.

- **III b) Rare events for the Solar System (planet collisions).**

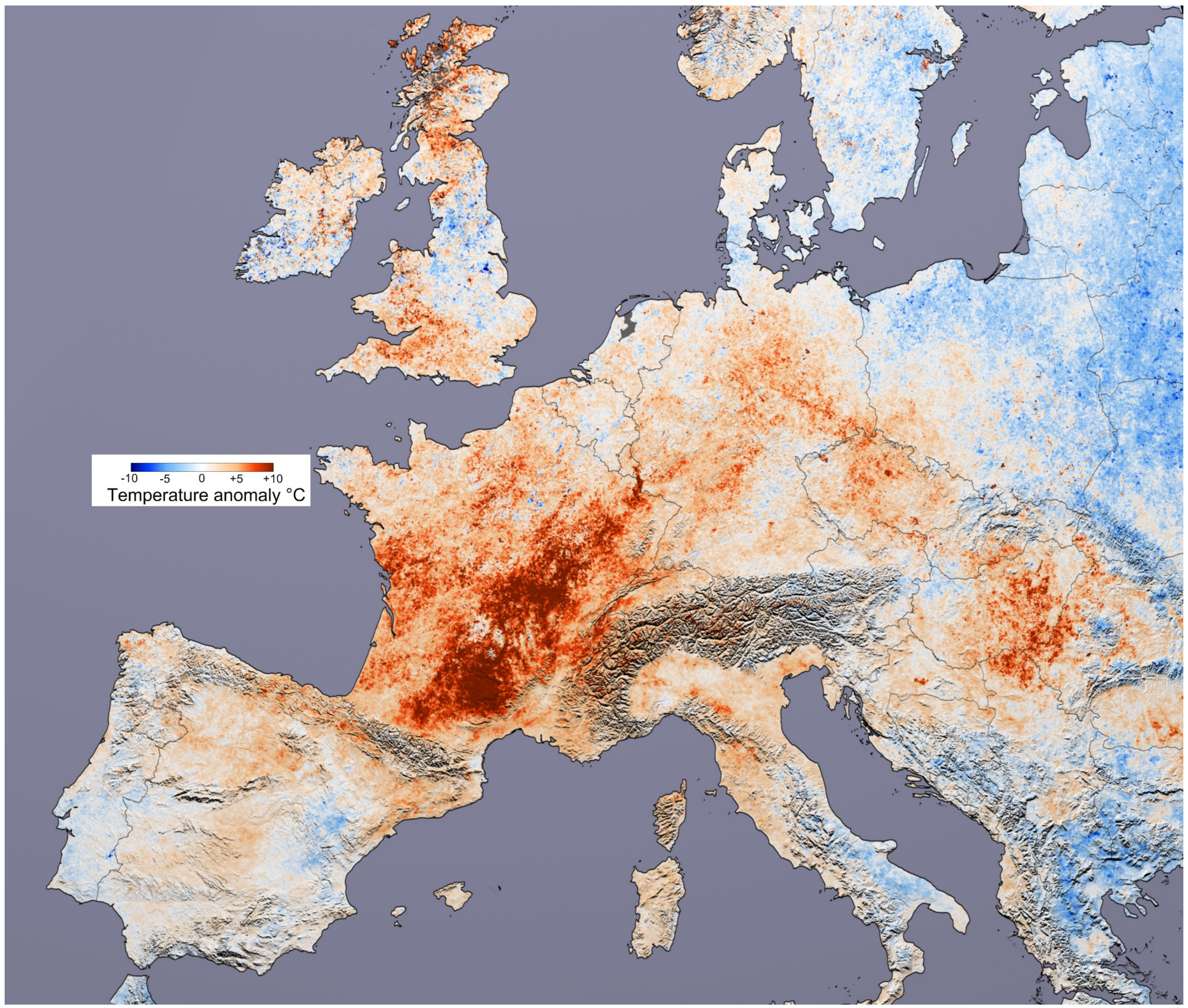
F. Bouchet and E. Woillez, PRL, 2020.

RESEARCH HIGHLIGHTS

Nature Reviews Physics | **The path to the Solar system's destabilization**

16 July 2020

Rare events matter
1- When they have a huge impact

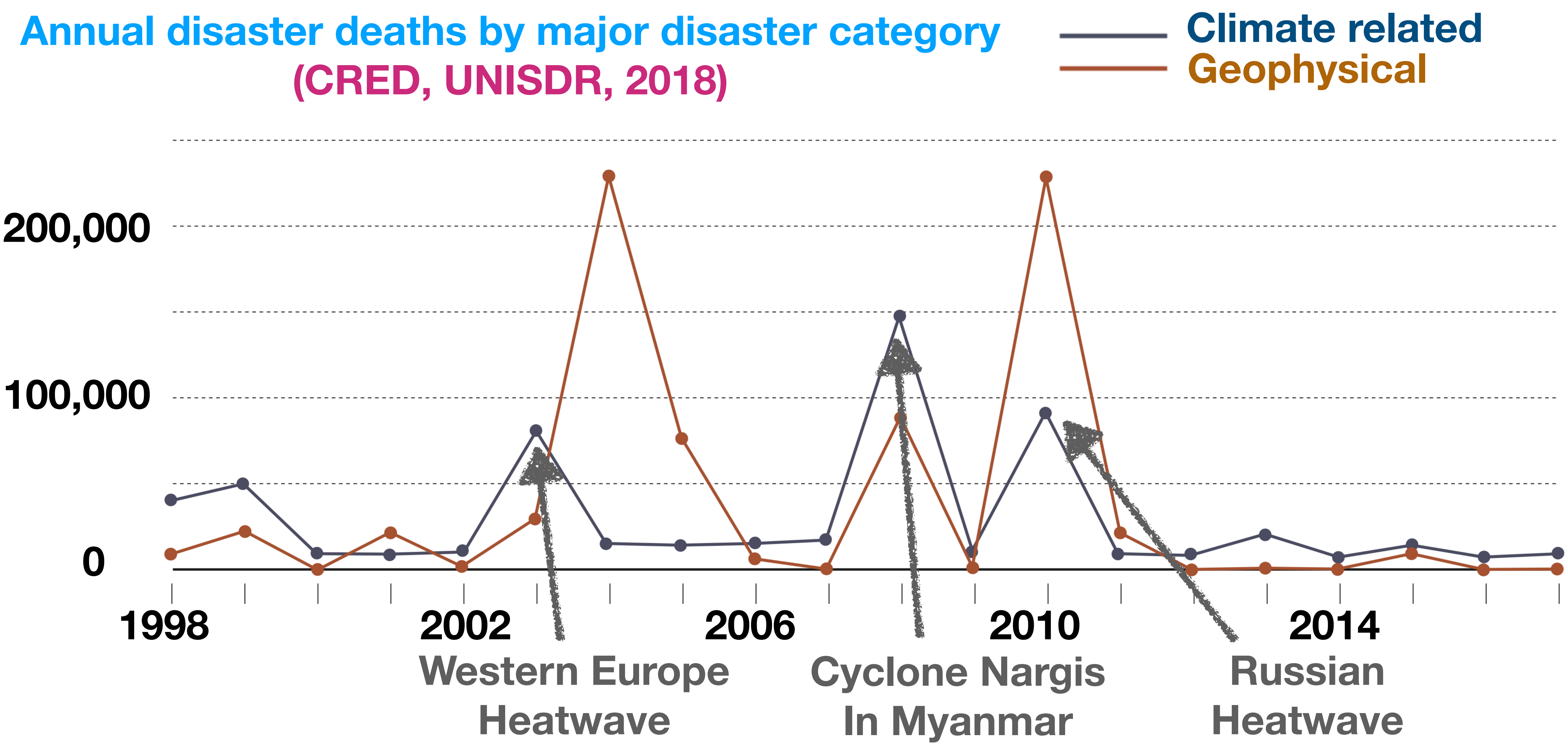


July 20 2003-August 20 2003 land surface temperature minus the average for the same period for years 2001, 2002 and 2004 (TERRA MODIS).

2003 heat wave over western Europe - 70 000 deaths.

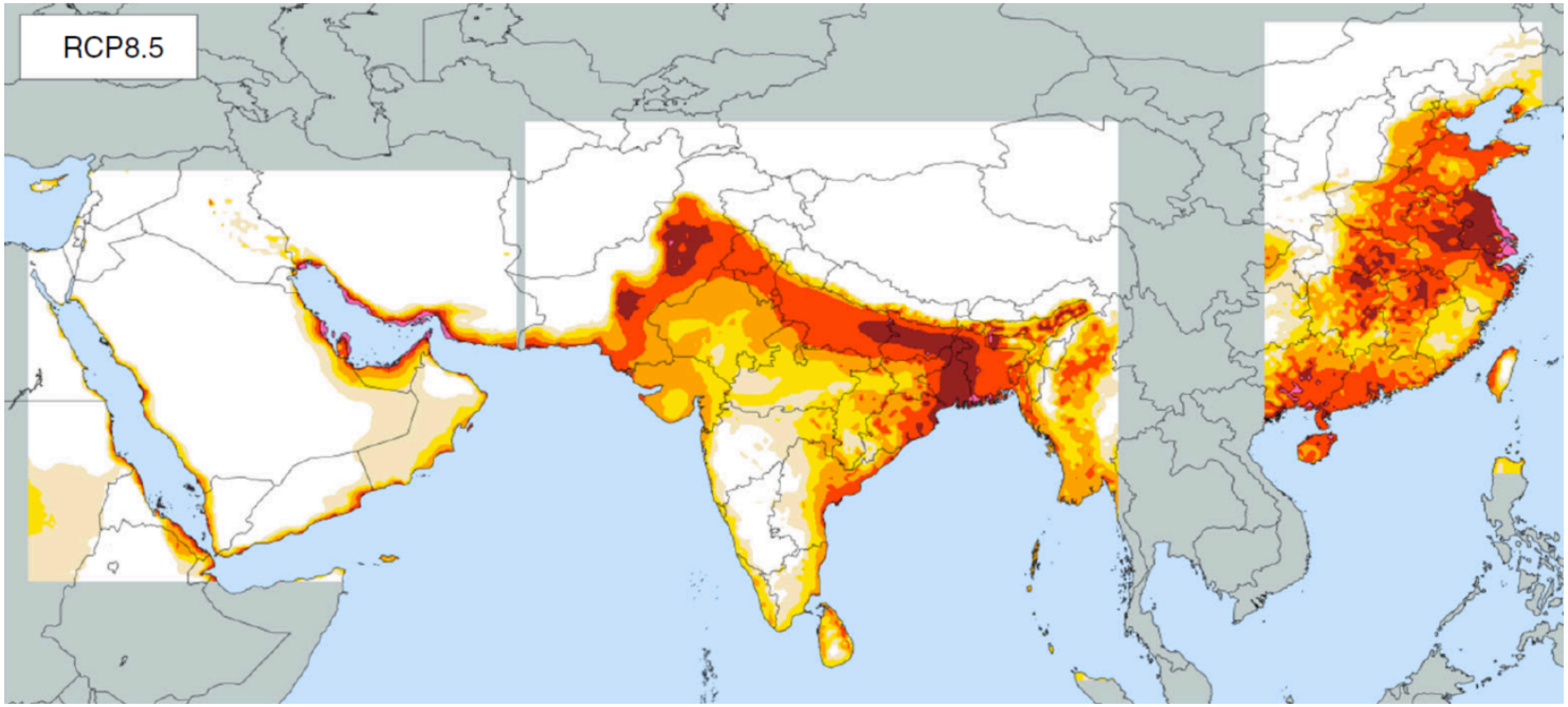
What are the probabilities (return times) and dynamics of extreme events?

The few most extreme climate events have more impact than all the others



**We need to study extremely rare events.
This is a serious scientific challenge.**

Potential impacts of global warming and extreme events

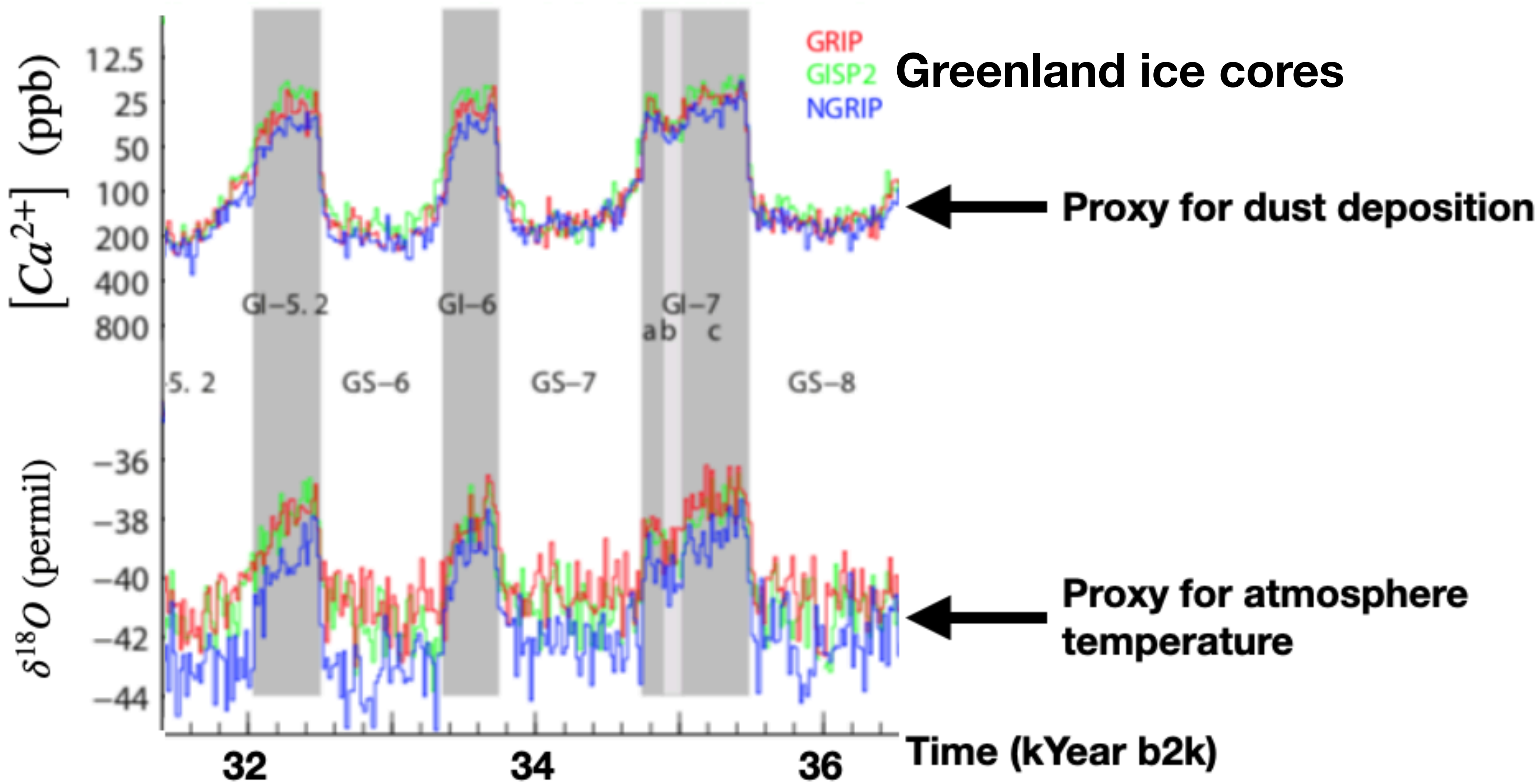


Maximal wet bulb temperature (red color =31-32°C), in 2070, with the RCP8.5 scenario.

(Kang, Elfatih and Eltahir, 2018)

Hundreds of thousands of people leave now in area of the world that will become inhabitable before the end of the century if we do not halt global warmings. Thinking of these phenomena in a classical economic framework does not make any sense.

Rare events matter
2 - When they produce structural changes



Climate abrupt transitions during the last glacial period.
(S. Rasmussen et al, 2014)

Our work on abrupt transitions to superrotating atmospheres
(C. Herbert, R. Caballero, and F. Bouchet, JAS, 2020)

Three key problems in the study of climate extreme events

- The historical records are way too short to make any meaningful predictions for the rarest events (those that matter the most).

- Climate models are wonderful tools, but they have biases. The more precise, the more computationally costly.

- Because they are too rare, the most extreme events cannot be computed using direct numerical simulations (the needed computing times are often unfeasible).

The practical questions: How to sample the probability and dynamics of rare events in complex models? How to build effective models which are relevant for estimating the probability of rare events?

Rare event algorithms for climate dynamics

The practical questions: How to sample the probability and dynamics of rare events (in complex dynamics - climate models)?

Outline :

I) Introduction

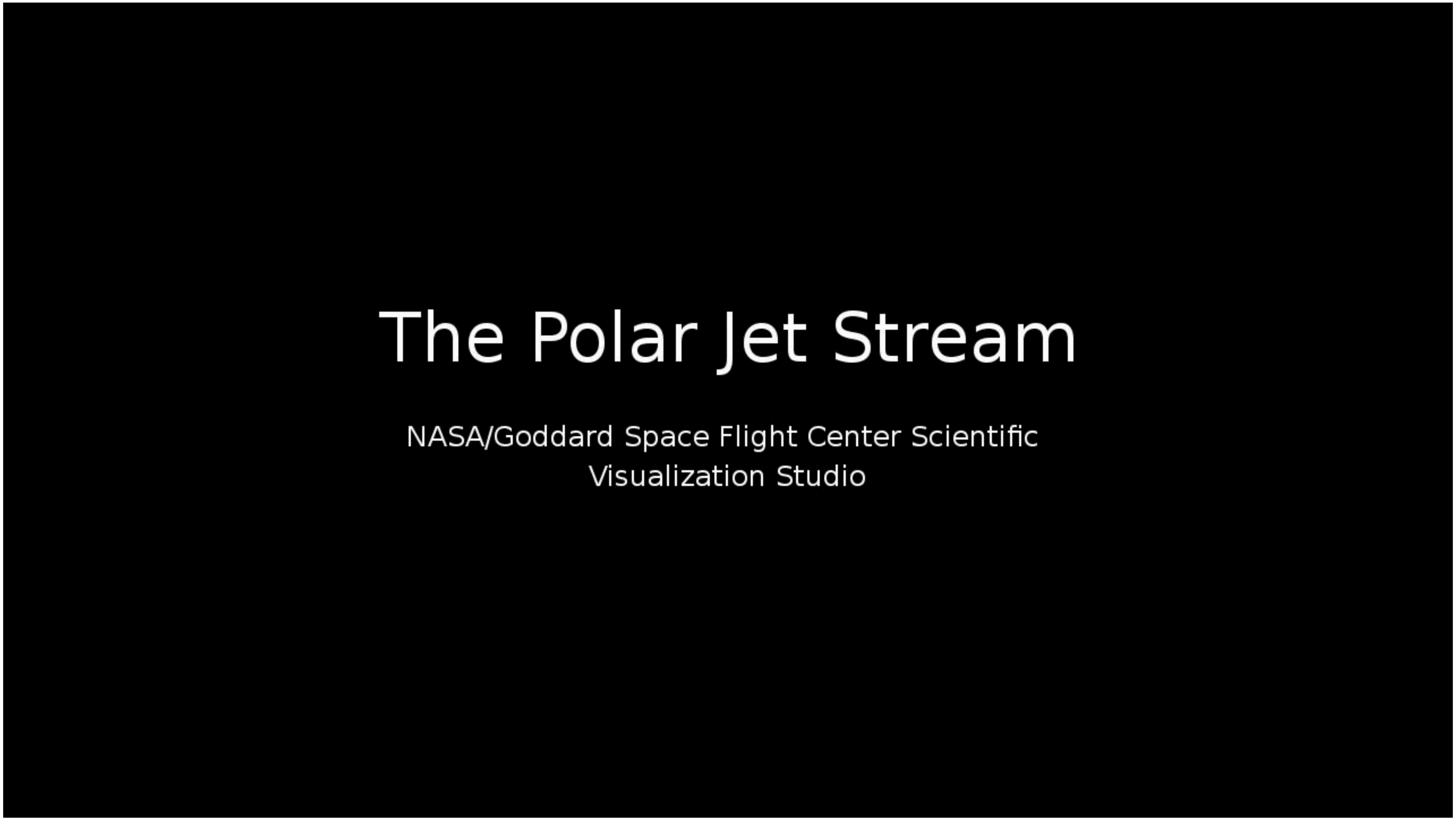
II) Rare events algorithms and teleconnection patterns for extreme heat waves

III) The challenges to go further: estimating good score functions for complex dynamics

IV) Committor functions for climate dynamics

II) Extreme heat waves - Rare event algorithms and teleconnection patterns

Jet stream dynamics



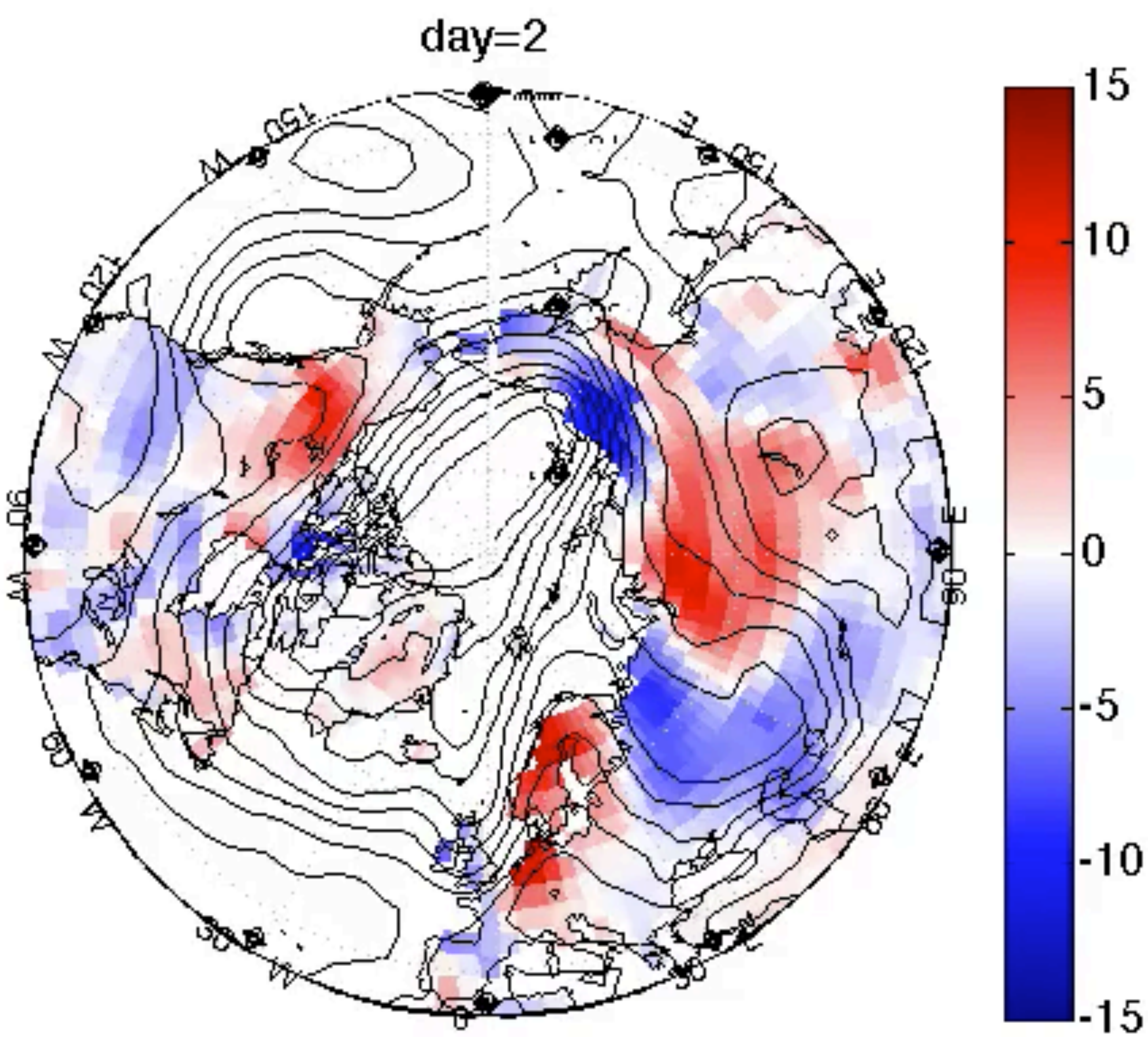
Higher troposphere wind speed. (NASA/Goddard Space Flight Center Scientific Visualization Studio, MERRA reanalysis dataset)

II-1) Rare event algorithms to study extreme heat waves with climate models

Francesco Ragone
RMI, Bruxelles, Belgium

Jeroen Wouters
University of Reading, UK

General Circulation Model



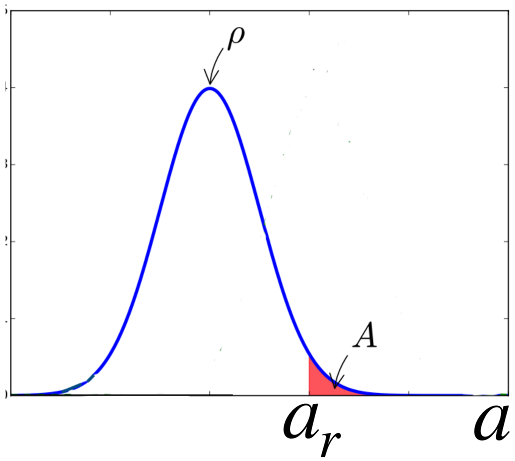
Surface temperature (T_s, colors) and 500 hPa geopotential height (Z_g, lines) anomalies

- **Plasim and CESM climate models.**
- **Global.** Coupled atmosphere/land/ocean/vegetation.

Long lasting summer heat waves

We will study extremes of the time averaged temperature anomalies:

$$a = \frac{1}{D}\int_0^D \mathrm{d}t \ \frac{1}{|\mathscr{A}|}\int_{\mathscr{A}} \mathrm{d}\mathbf{r}\ \mathrm{T_s}(\mathbf{r}, t)$$



- Duration $D =$ one week, a few weeks, a month, or a season.
- Area $\mathscr{A} =$ Scandinavia, Europe, France, Alberta, Russia, …
- **Climate models** (CESM or PLASIM) or reanalysis datasets.

The Giardina—Kurchan (Del-Moral—Garnier) rare event algorithm

- With $A[X](t) = \dfrac{1}{|\mathscr{A}|}\displaystyle\int_{\mathscr{A}} \mathrm{d}\mathbf{r}\ \mathrm{T_S}(\mathbf{r}, t)$, we sample the tilted path-distribution

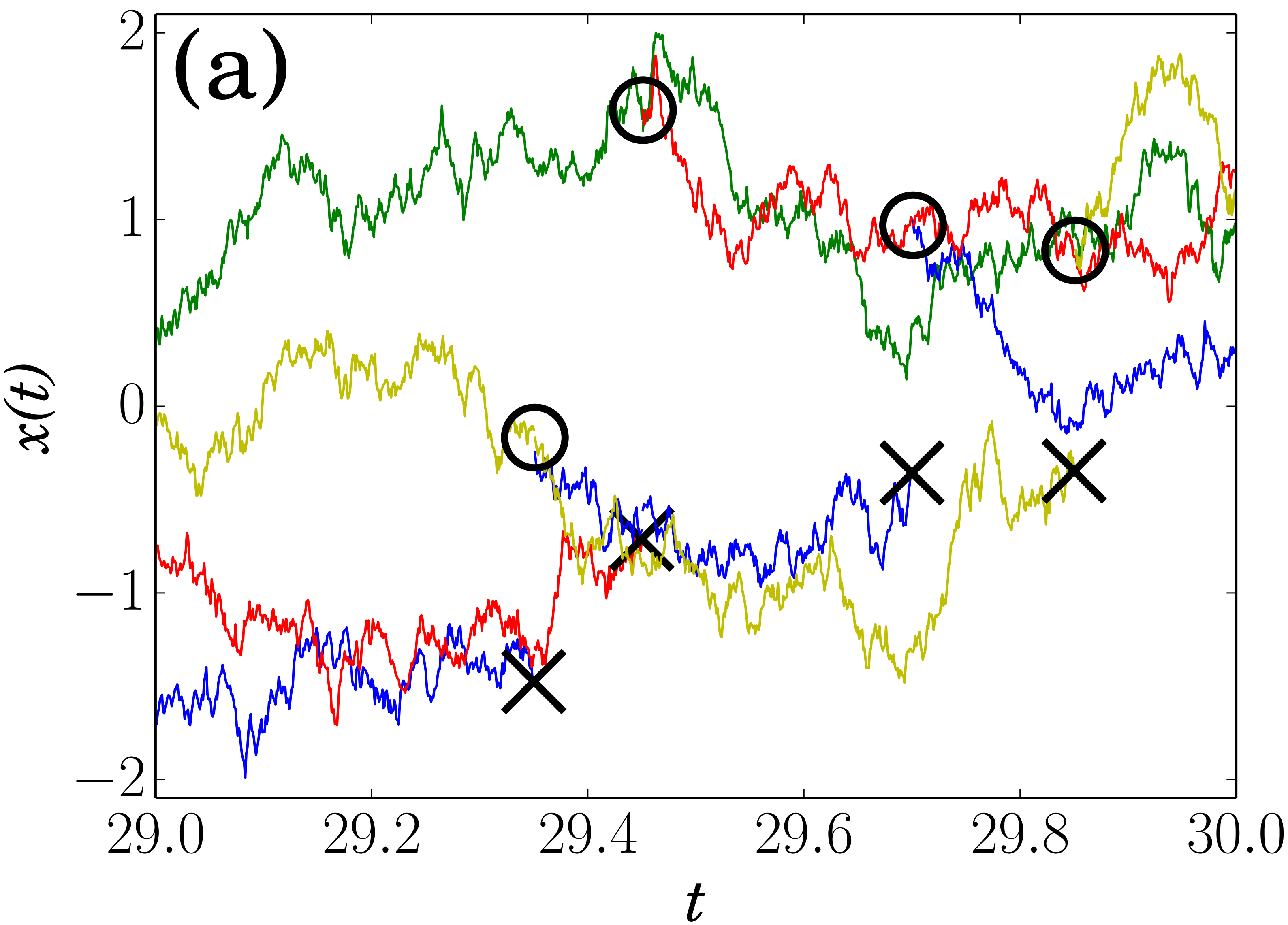
$$\tilde{P}_k\left(\left\{X(t)\right\}_{0\le t\le T}\right) = \frac{1}{\exp(T\lambda(k))} P_0\left(\left\{X(t)\right\}_{0\le t\le T}\right) \exp\left[k\int_0^T A[X](t)\,\mathrm{d}t\right].$$

- We simulate an ensemble of N trajectories $x_n(t)$. At each time step $t_i = i\tau$, each trajectory can be killed or cloned according to the weights

$$\frac{1}{W_i(k)} \exp\left(k\int_{t_{i-1}}^{t_i} A[x_n](t)\,\mathrm{d}t\right) \quad \text{with} \quad W_i(k) = \sum_{n=1}^{N} \exp\left(k\int_{t_{i-1}}^{t_i} A[x_n](t)\,\mathrm{d}t\right).$$

- Algorithm: Giardina et al. 2006. Mathematical aspects: Del Moral's book (2004).

Genealogical algorithm: selecting, killing and cloning trajectories

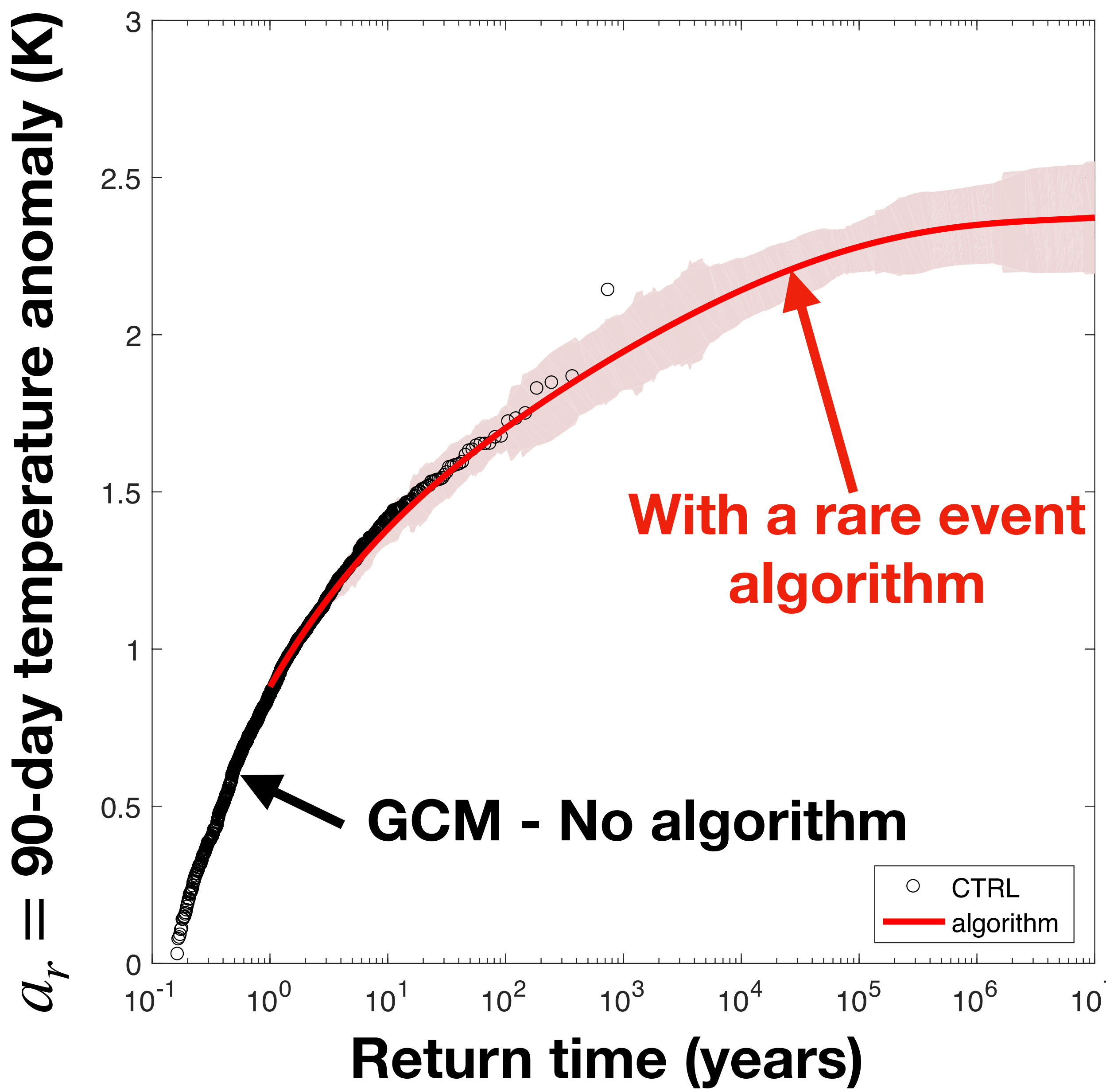


The trajectory statistics is tilted towards the events of interest.

Sample paths of the Giardina Kurchan algorithm

(from Bouchet, Jack, Lecomte, Nemoto, 2016)

Return time plot computed using a rare event algorithm (PLASIM)



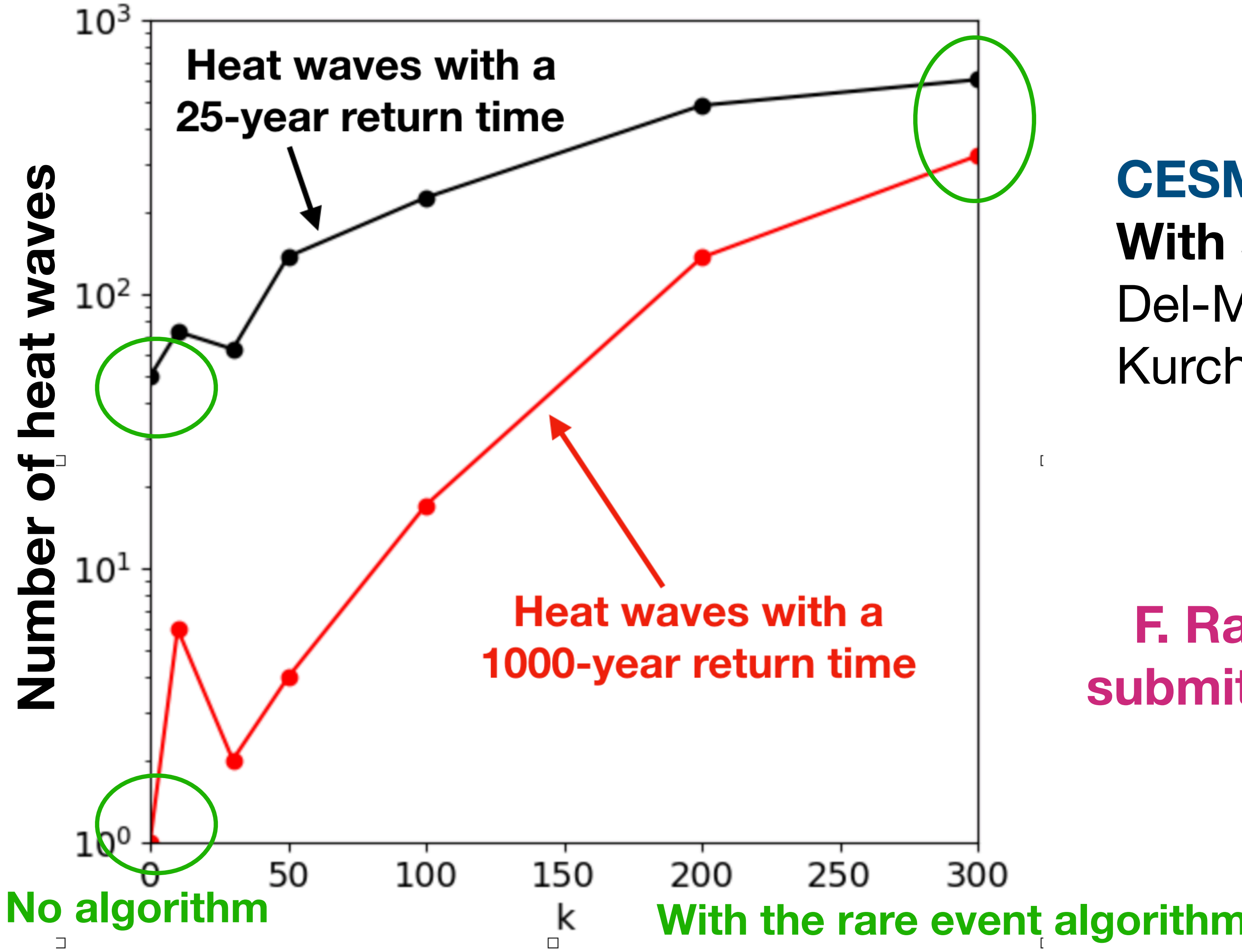
Extremes of 90-day Europe heat waves

PLASIM model.
No seasonal cycle.
Del-Moral—Garnier (or Giardina—Kurchan) algorithm.

F. Ragone, J. Wouters, and F. Bouchet, PNAS, 2018

At a fixed numerical cost, we can study events which are several orders of magnitude rarer with the rare event algorithm than with the control run.

Oversampling of extreme event using a rare event algorithm (CESM)



CESM model

With seasonal cycle.

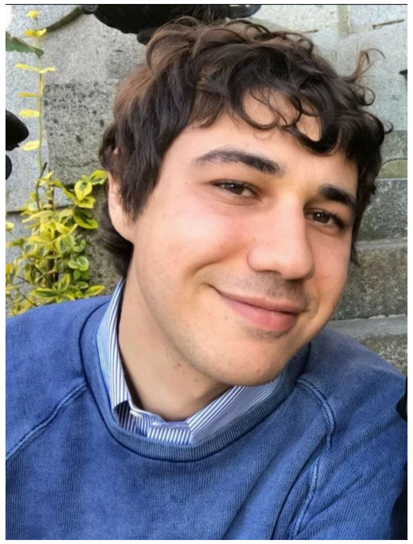
Del-Moral—Garnier (or Giardina—Kurchan) algorithm.

F. Ragone, and F. Bouchet, 2021, submitted to GRL, arXiv:2009.02519.

Number of observed heat waves for 1,000 of simulations

We get several hundreds more heat waves with a return times of 1000 years than with the direct numerical simulation.

II-2) Heat wave dynamics and global teleconnection patterns for extremes

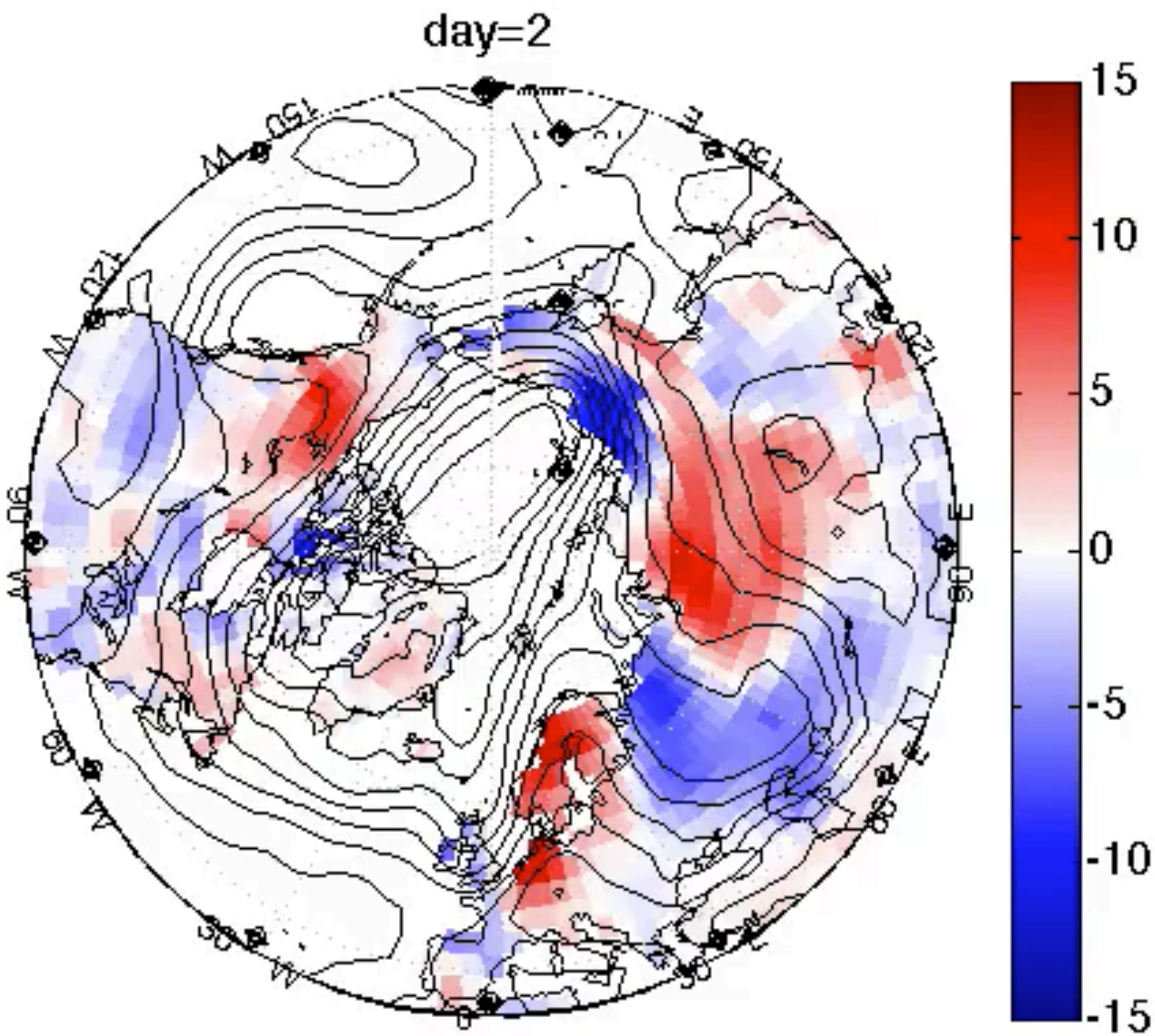


Dario Lucente

George Miloshevich

Francesco Ragone

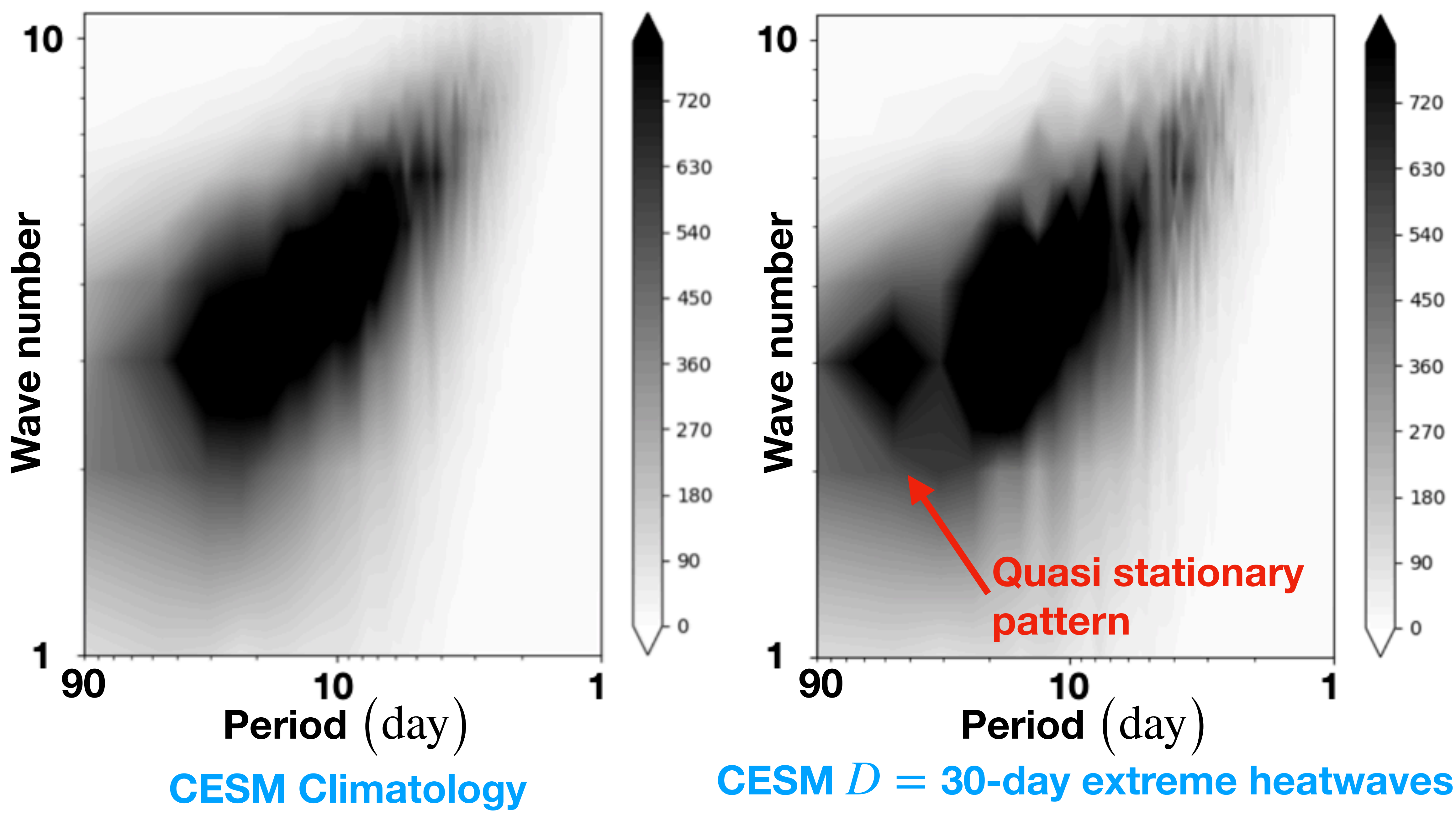
Heat wave dynamics



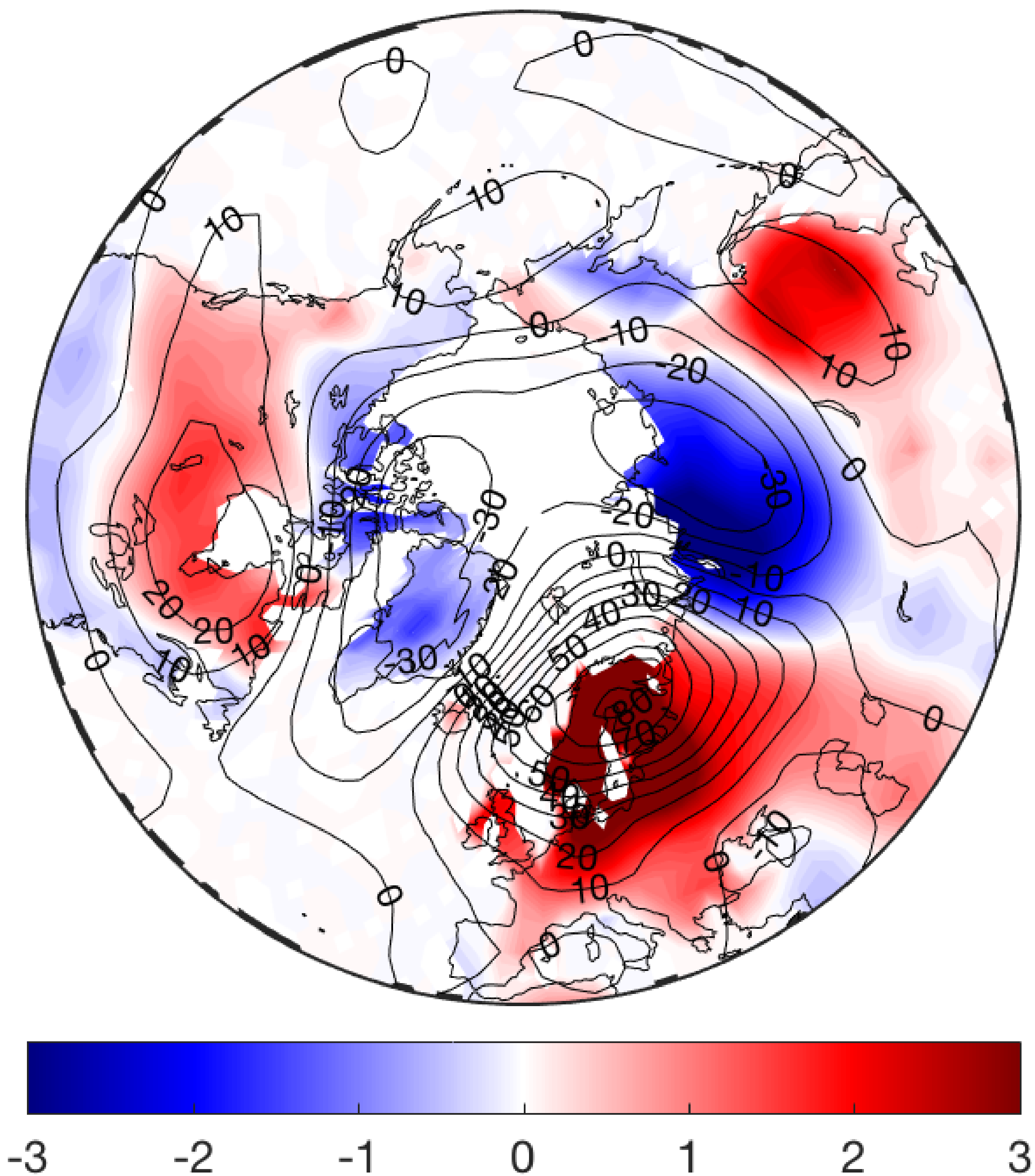
Plasim heat wave over Scandinavia

Heat wave = unusual quasi stationary pattern + progressive Rossby wave

Hayashi spatio-temporal spectrum for eastward waves - CESM model (from the 500 hPa geopotential height over a latitudinal band $55^o - 75^oN$)



Extreme teleconnection pattern



500 hPa geopotential height and temperature anomalies

Extreme teleconnection patterns = conditional averages with

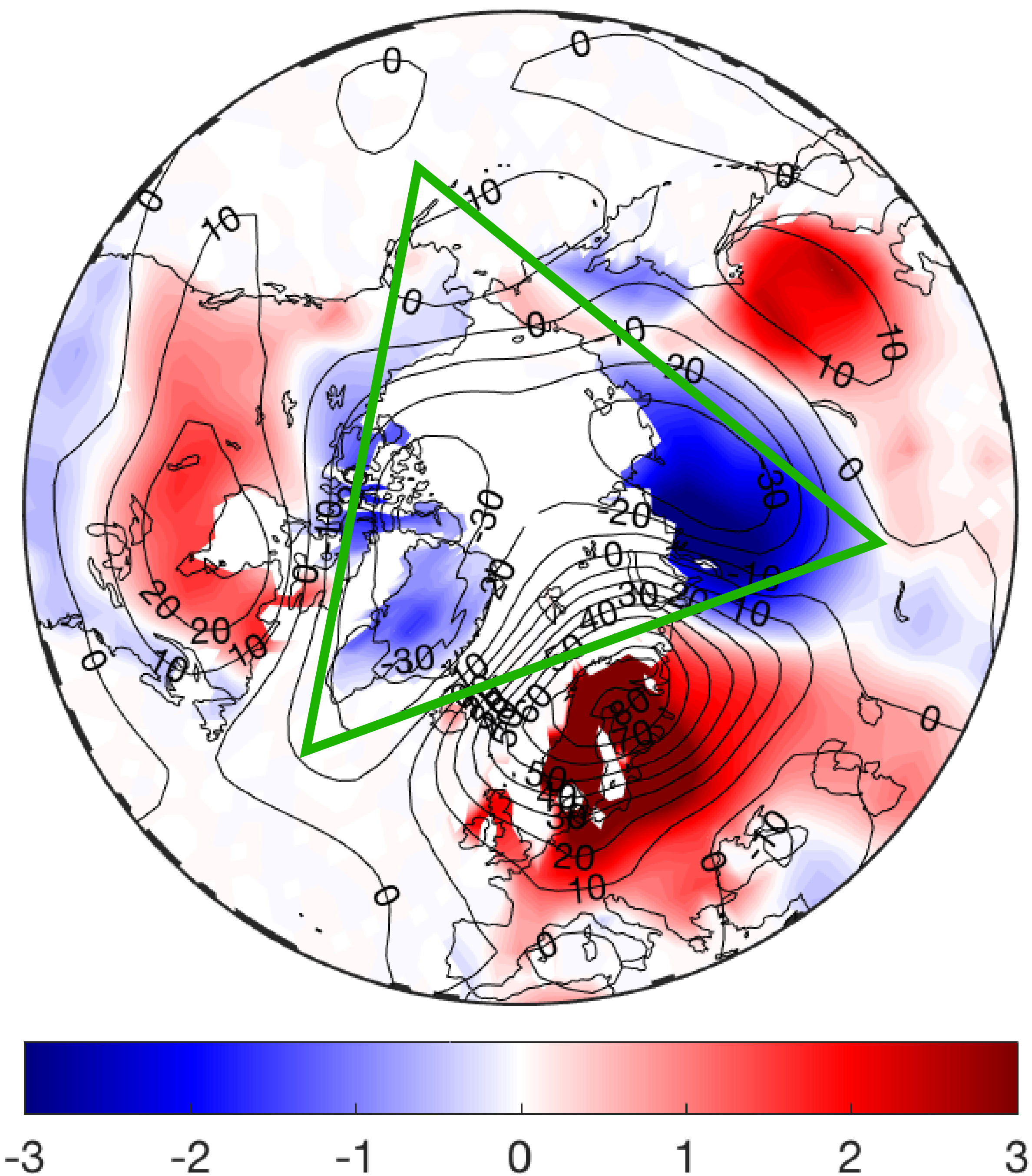
$$\frac{1}{D}\int_0^D \mathrm{d}t \frac{1}{|\mathscr{A}|}\int_{\mathscr{A}} \mathrm{d}\mathbf{r}\ \mathrm{T}_S(\mathbf{r}, t) > 2\ \mathrm{K}$$

and $D = 40$ days.

Plasim model. Summer Scandinavian heat waves.

F. Ragone, J. Wouters, and F. Bouchet, PNAS, 2018

Extreme teleconnection pattern



500 hPa geopotential height and temperature anomalies

Extreme teleconnection patterns = conditional averages with

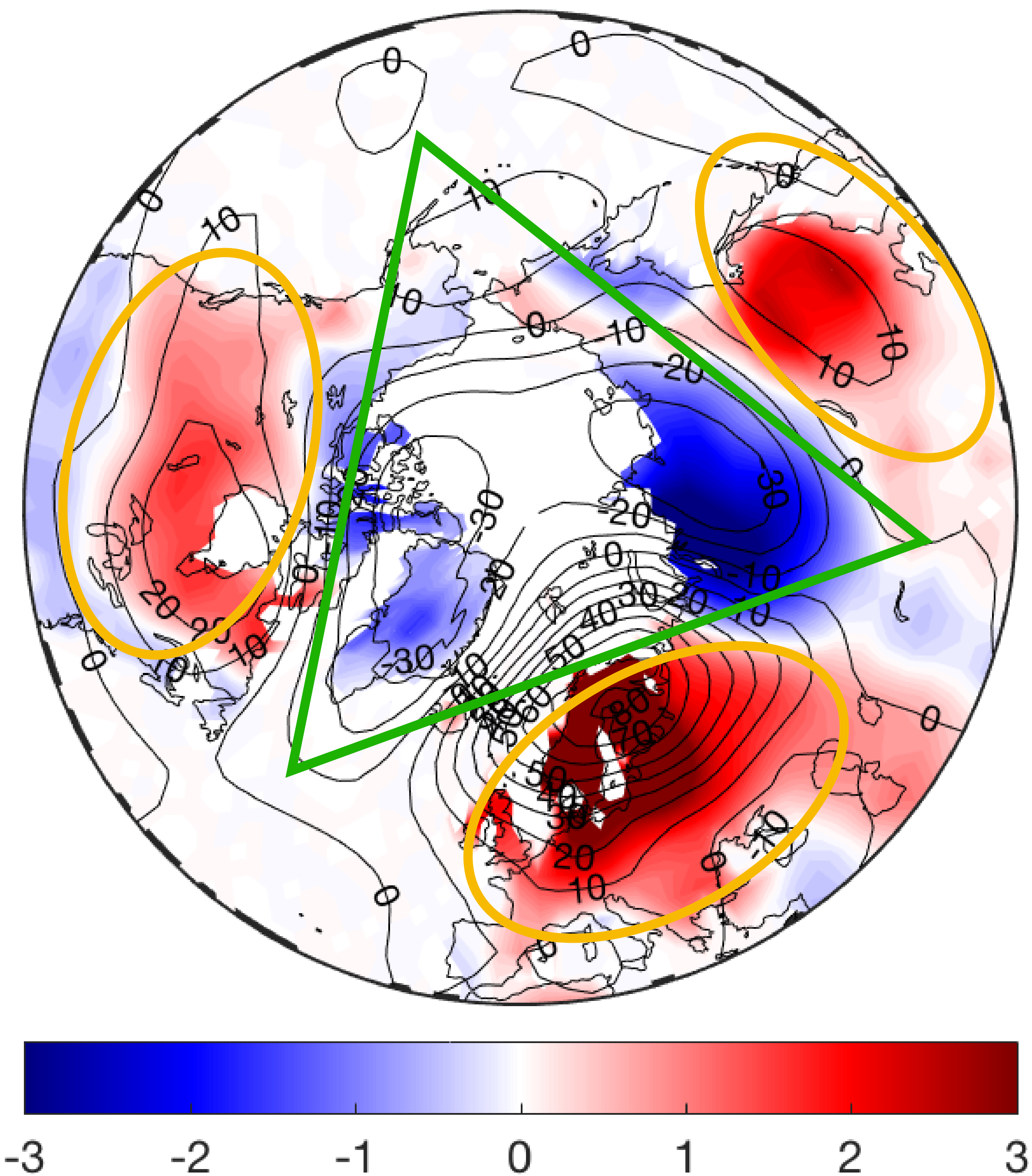
$$\frac{1}{D}\int_0^D \mathrm{d}t \frac{1}{|\mathscr{A}|}\int_{\mathscr{A}} \mathrm{d}\mathbf{r}\; \mathrm{T_S}(\mathbf{r},t) > 2\ \mathrm{K}$$

and $D = 40$ days.

Plasim model.
Summer Scandinavian heat waves.

F. Ragone, J. Wouters, and F. Bouchet, PNAS, 2018

Extreme teleconnection pattern



500 hPa geopotential height and temperature anomalies

Extreme teleconnection patterns = conditional averages with

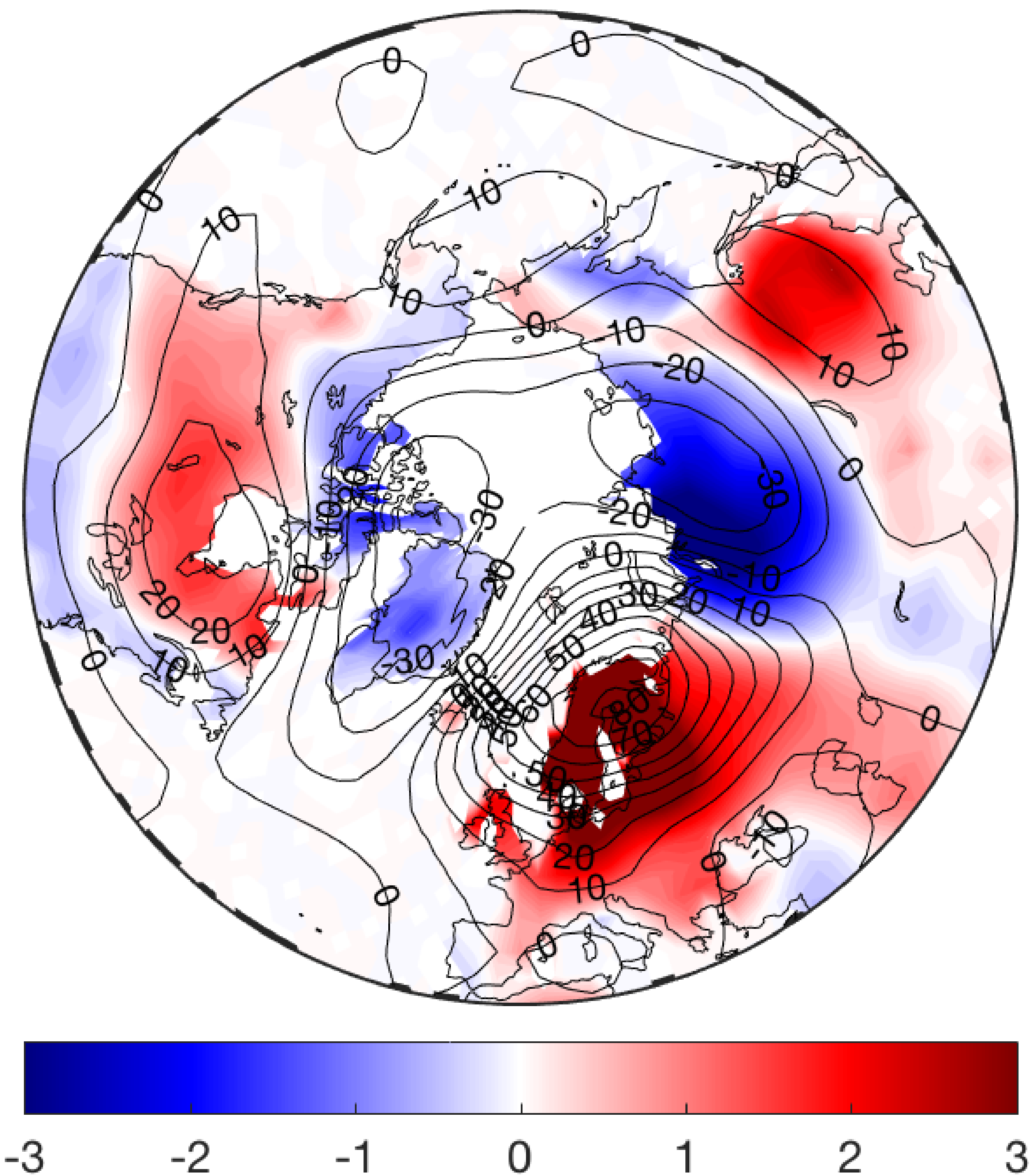
$$\frac{1}{D}\int_0^D \mathrm{d}t \frac{1}{|\mathscr{A}|}\int_{\mathscr{A}} \mathrm{d}\mathbf{r}\ \mathrm{T_S}(\mathbf{r}, t) > 2\ \mathrm{K}$$

and $D = 40$ days.

Plasim model.
Summer Scandinavian heat waves.

F. Ragone, J. Wouters, and F. Bouchet, PNAS, 2018

Extreme teleconnection pattern



Extreme teleconnection patterns = conditional averages with

$$\frac{1}{D}\int_0^D \mathrm{d}t \frac{1}{|\mathscr{A}|}\int_{\mathscr{A}} \mathbf{dr}\ \mathrm{T_S}(\mathbf{r},t) > 2\ \mathrm{K}$$

and $D = 40$ days.

Plasim model.
Summer Scandinavian heat waves.

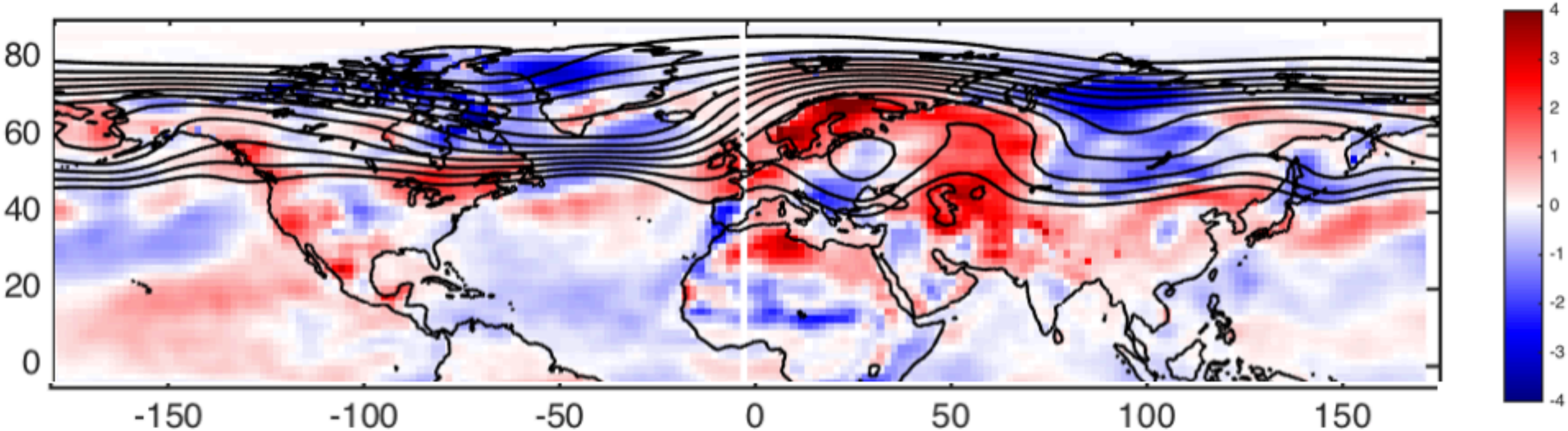
F. Ragone, J. Wouters, and F. Bouchet, PNAS, 2018

500 hPa geopotential height and temperature anomalies

Extreme teleconnection patterns differ from teleconnections for typical fluctuations and are not characterized by a single wavenumber but are much constrained by geography.

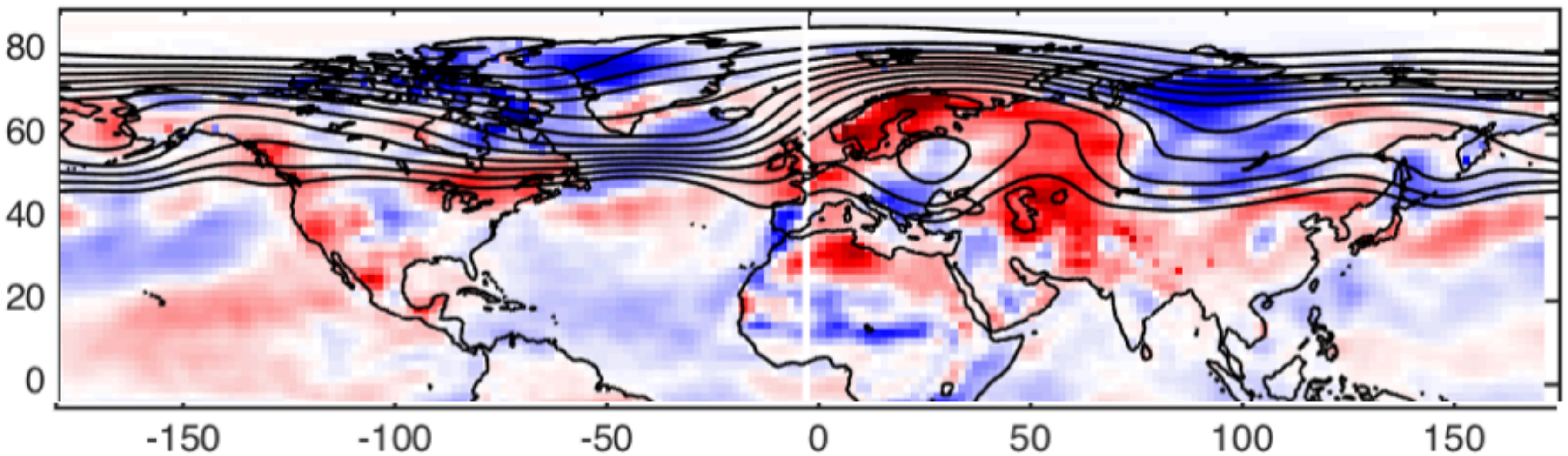
2018 heat wave over Scandinavia

Era 5 reanalysis dataset



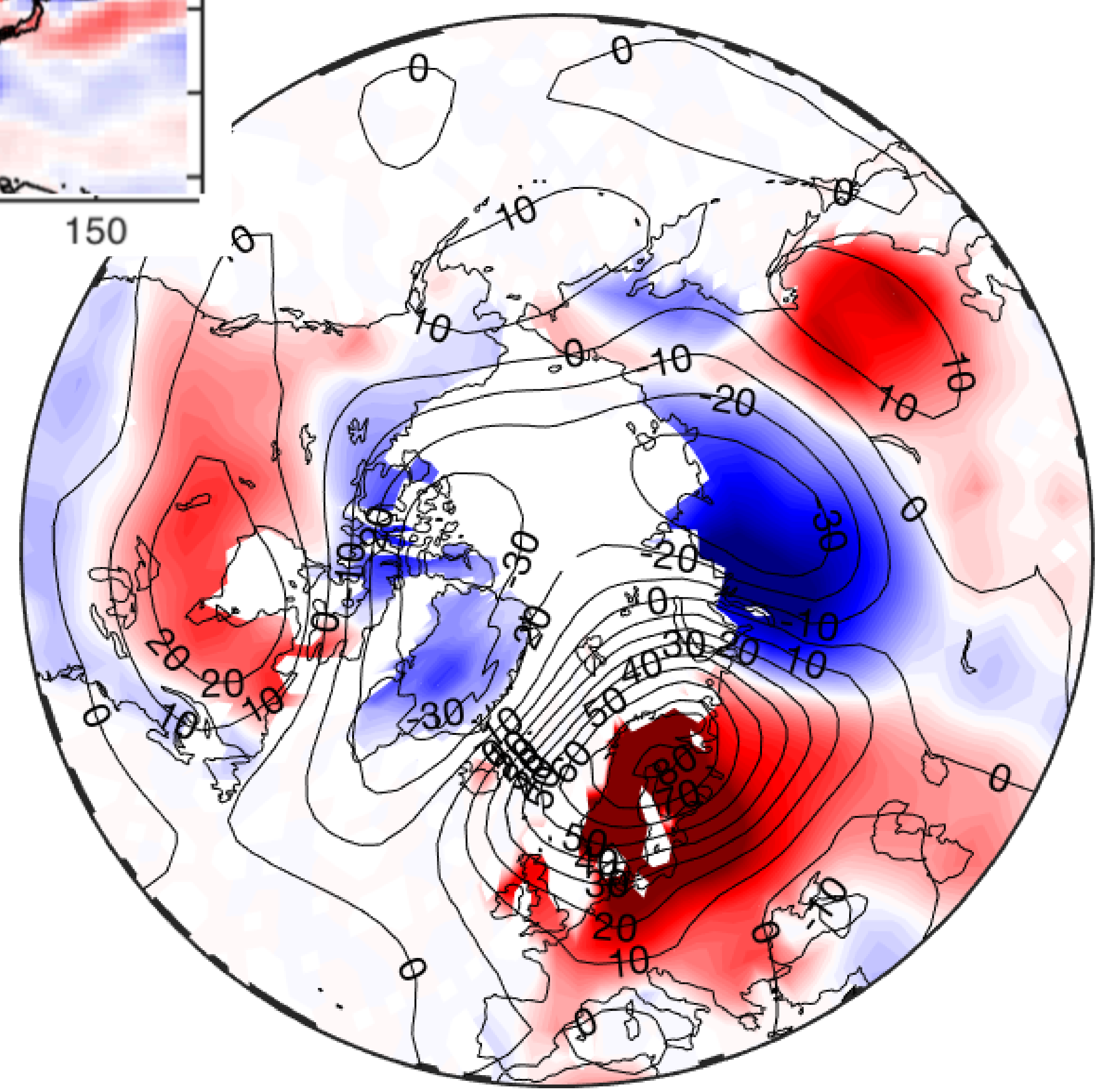
Temperature anomalies and 500 hPa geopotential height in July 2018

2018 heat wave over Scandinavia



Observed in July 2018

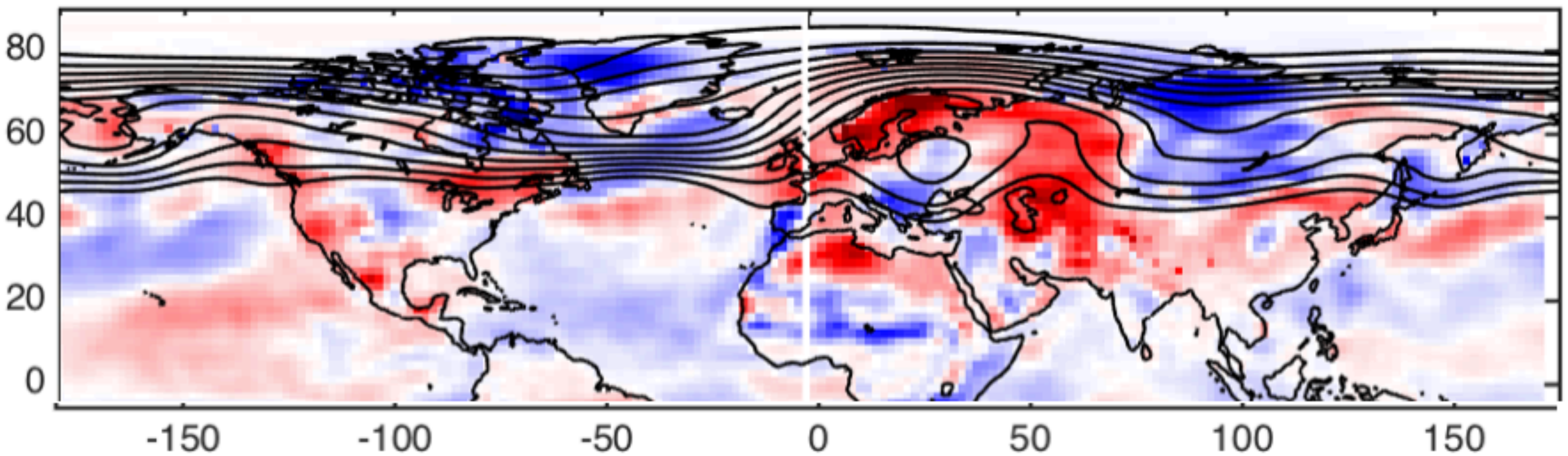
Published in January 2018



Is this just by chance?

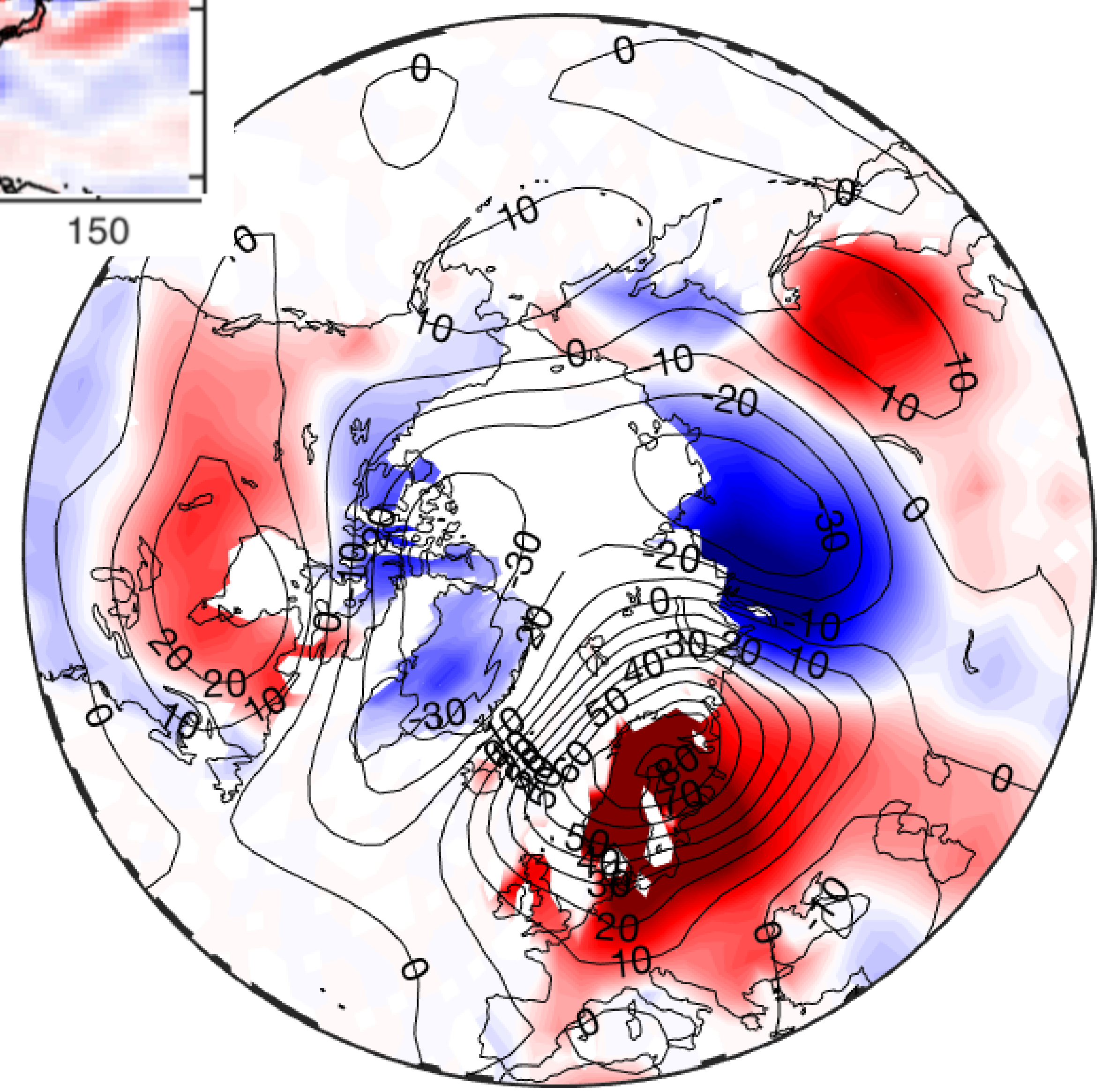
How good are the models for predicting those patterns?

2018 heat wave over Scandinavia



Observed in July 2018

Published in January 2018

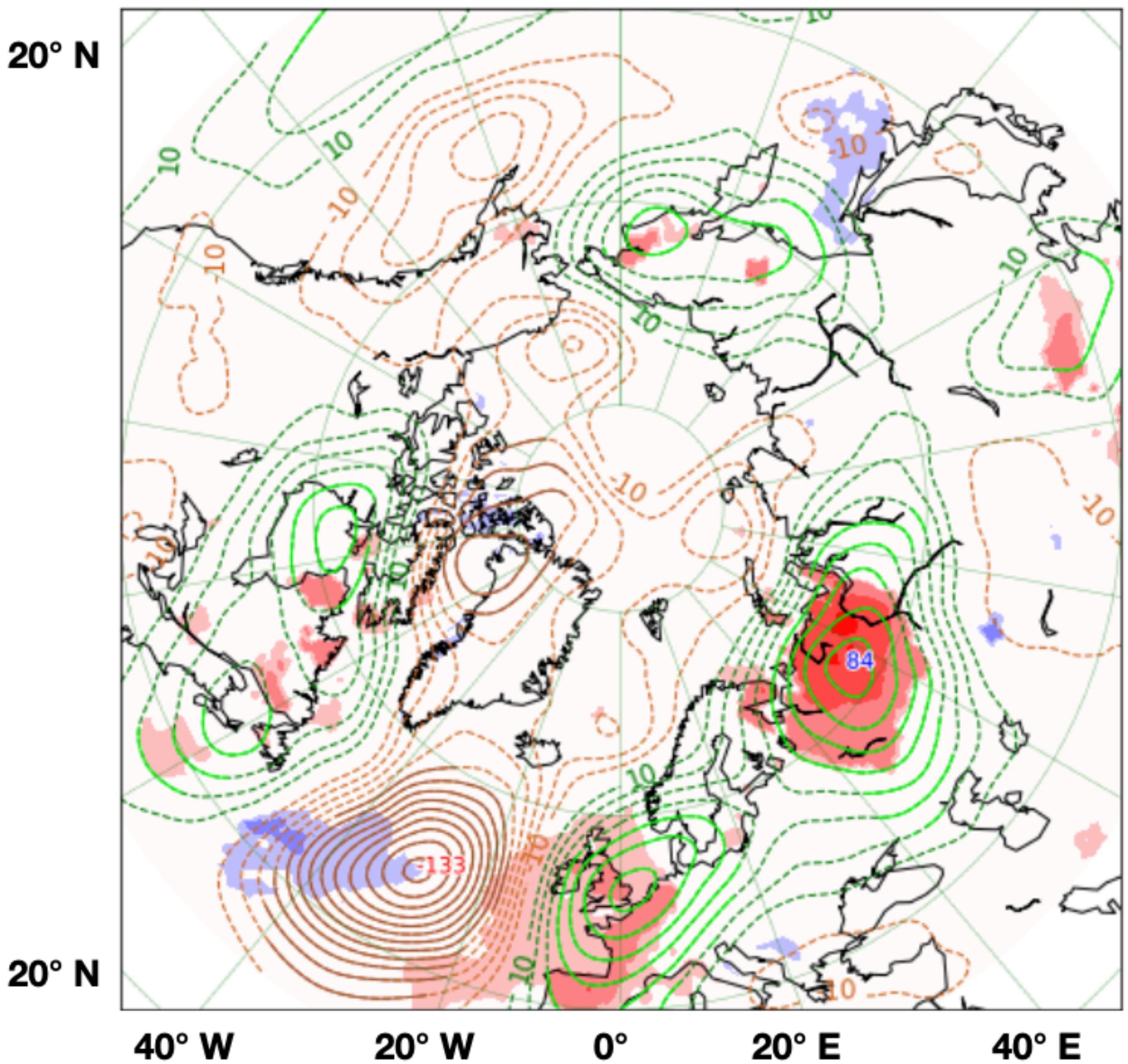


Is this just by chance?

How good are the models for predicting those patterns?

CESM patterns (2020) are similar to Plasim ones (2018).

Teleconnection patterns for moderate heat waves over France - ERA5



ERA5 reanalysis

4-year return time.
$D = 14.$
At day $\tau = 0.$

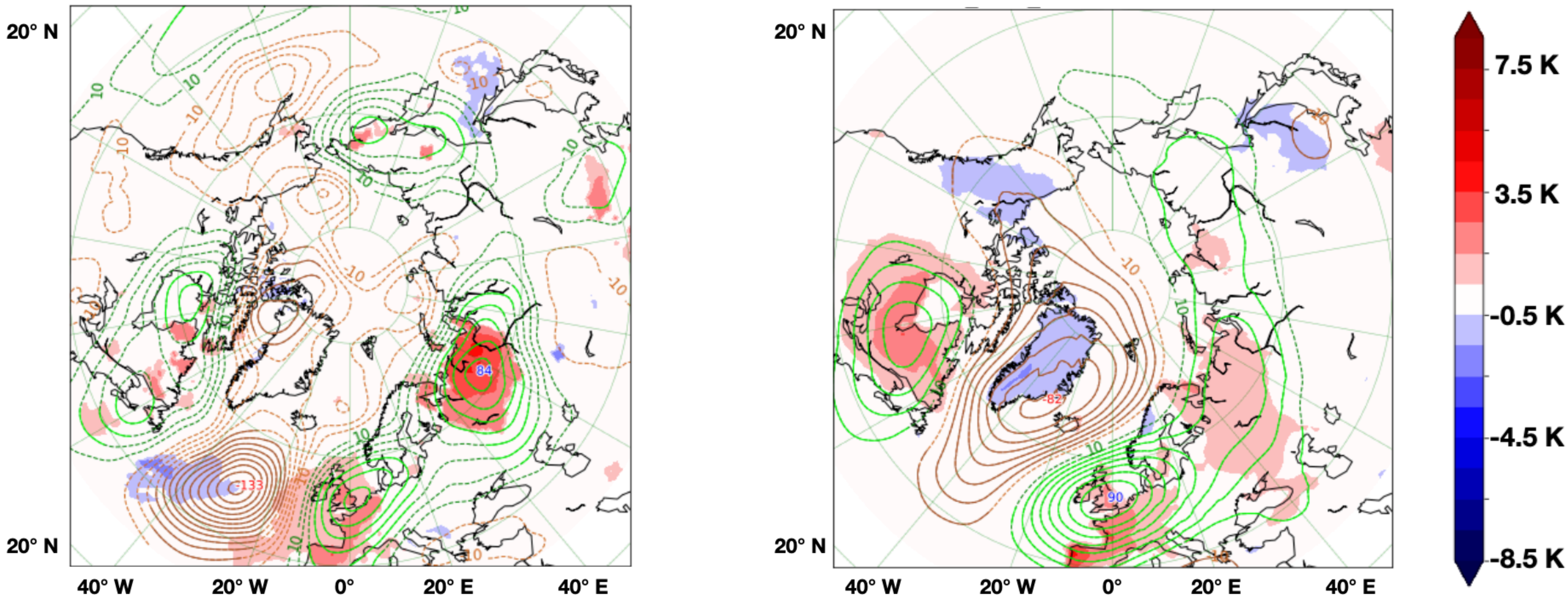
Dashed = not statistically significant for a t>2 student test.

Temperature anomalies and 500 hPa geopotential height conditioned on moderate heat waves

It is extremely difficult to have statistically significant patterns from reanalysis datasets.

Teleconnection patterns for moderate heat waves over France

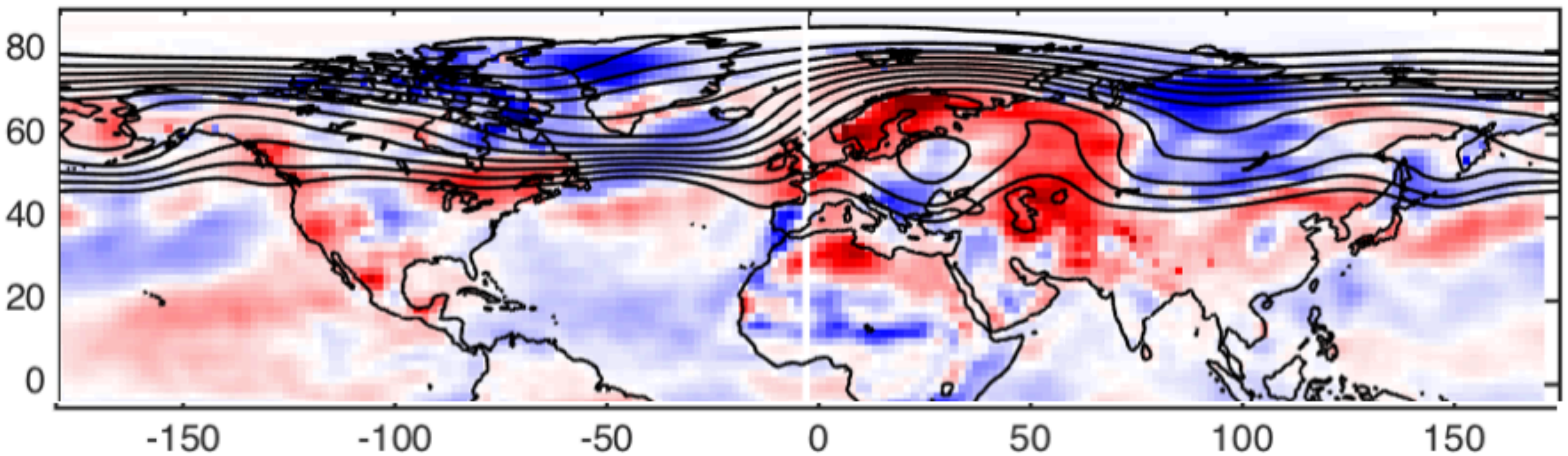
ERA5 reanalysis **CESM model**



Temperature anomalies and 500 hPa geopotential height conditioned on extreme heat waves with a 4-year return time, with $D = 14$, at day $\tau = 0$

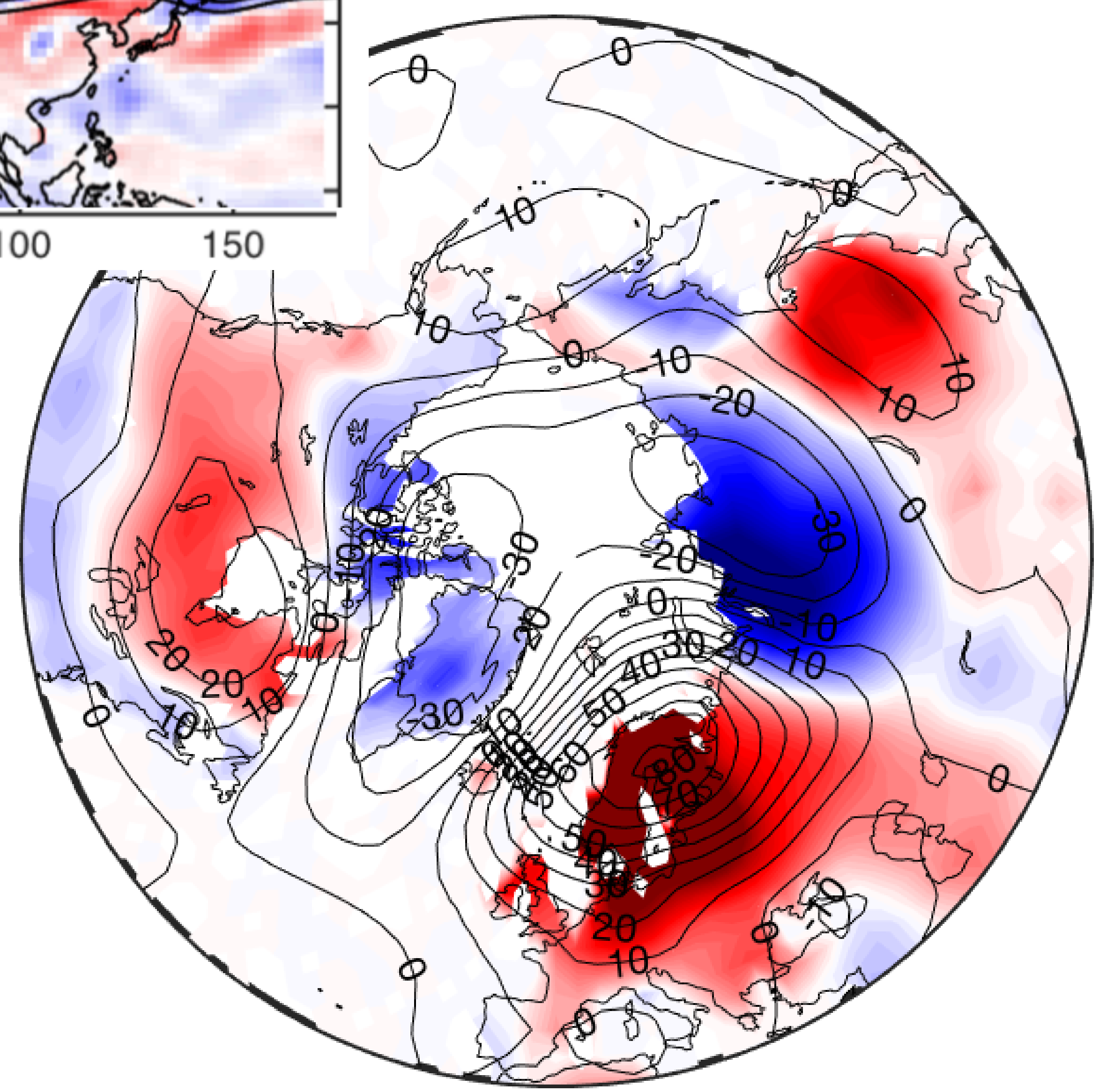
For moderately extreme heat waves, CESM and ERA5 reanalysis dataset are qualitatively consistent.

2018 heat wave over Scandinavia



Observed in July 2018

Published in January 2018



Conclusions:

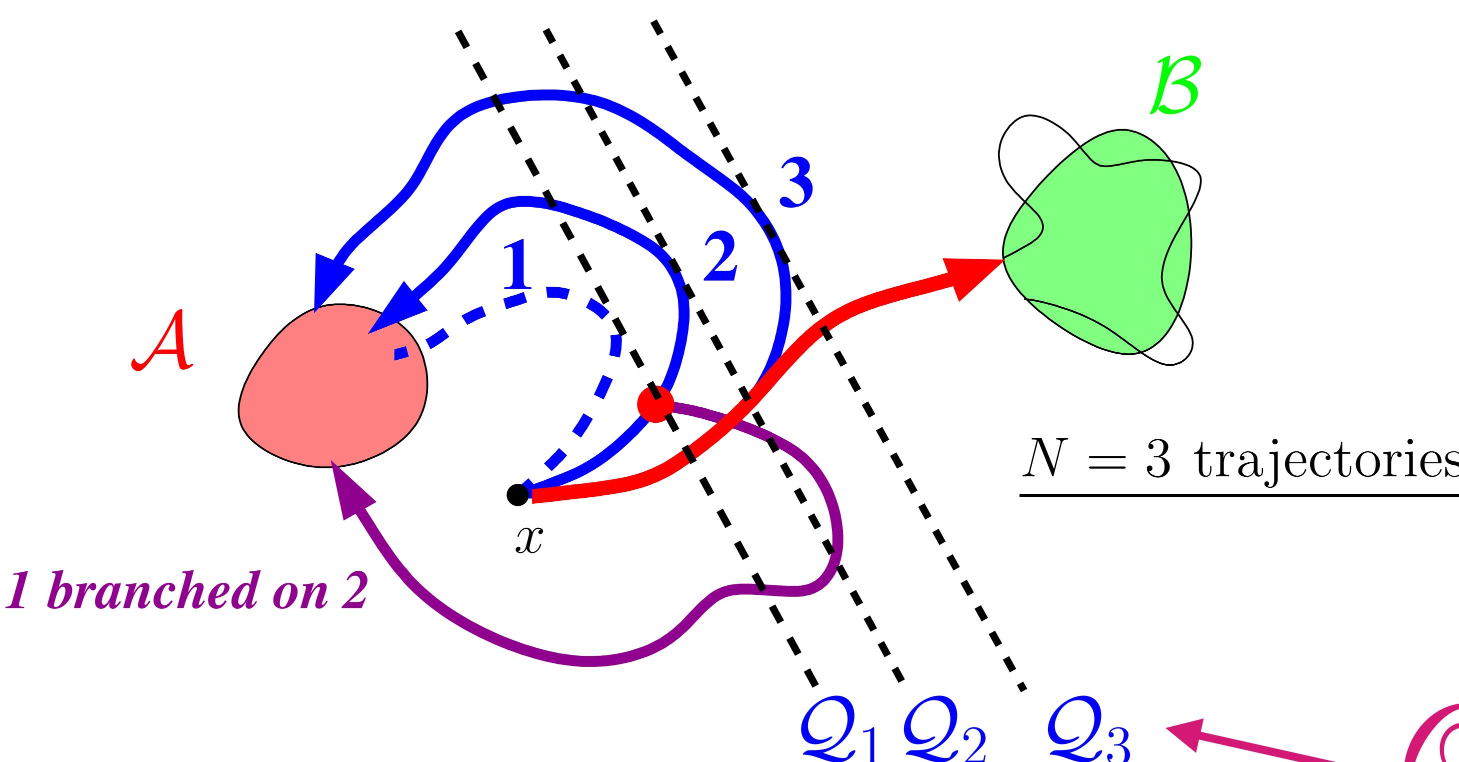
Climate models predict well moderately extreme teleconnection patterns.

This is a hint that they might predict correctly more extreme teleconnection patterns and that what we observed is not a coincidence.

III) The challenges to go further: estimating good score functions for complex dynamics

The Adaptive Multilevel Splitting (AMS) rare event algorithm

Strategy: selection, pruning and cloning.



Probability estimate:

$\hat{p} = (1 - 1/N)^K$,

where N is the clone number and K is the iteration number.

$\mathcal{Q}$ is the score (sélection) function

Cérou, Guyader (2007). Cérou, Guyader, Lelièvre, and Pommier (2011).
PDEs: Rolland, Bouchet et Simonnet (2016) - **TAMS:** Lestang et al (2018)
Atmosphere turbulent jets: Rolland, Bouchet et Simonnet (2019 and 2021).

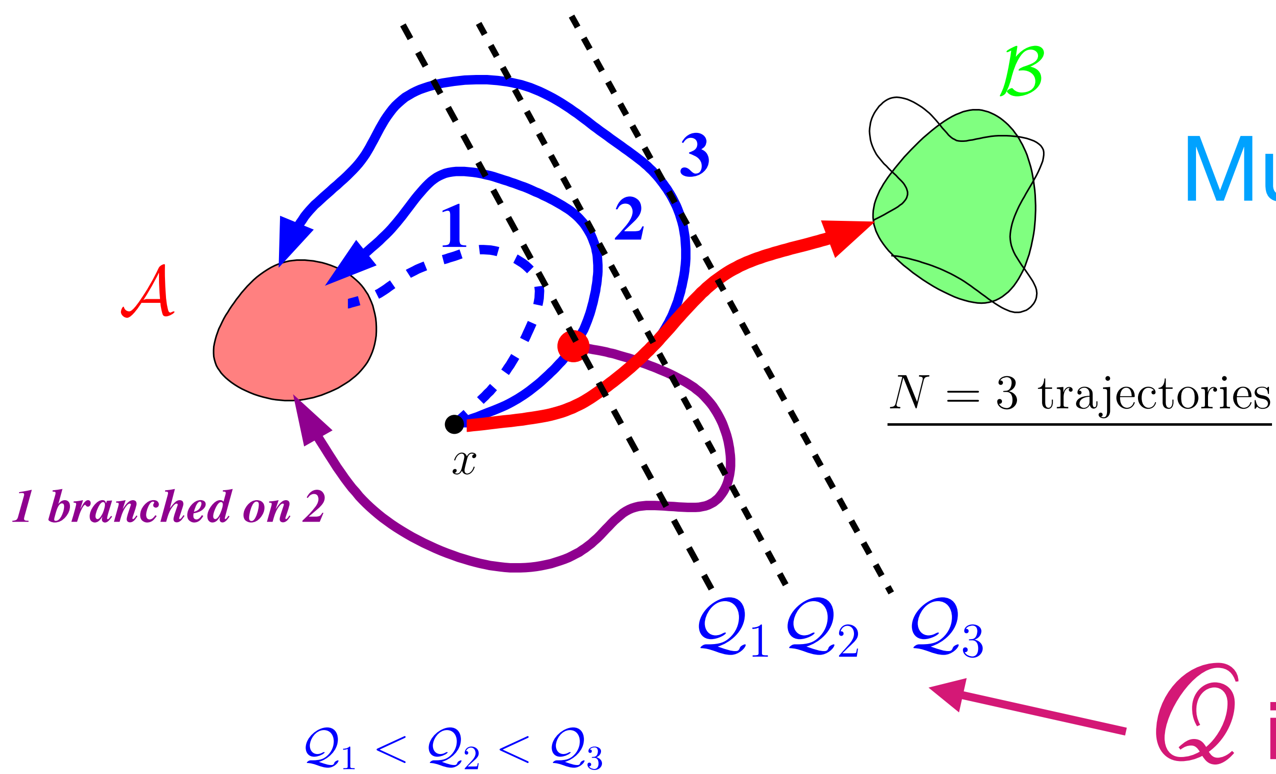
The score function is the key practical problem

- With a good score function, rare event algorithms give excellent results.

- With a poor score function, rare event algorithms are useless.

- How to build good score functions?

Committor functions are optimal score functions for rare event algorithms



The Adaptive Multilevel Splitting algorithm

$\mathcal{Q}$ is the score (sélection) function

The efficiency of the algorithm depends on the choice of the score function.

The optimal score function is the committor function.

Committor function

- $\{X(t)\}_{-\infty \le t < +\infty}$ is a Markov process. A, B are subsets of the phase space.

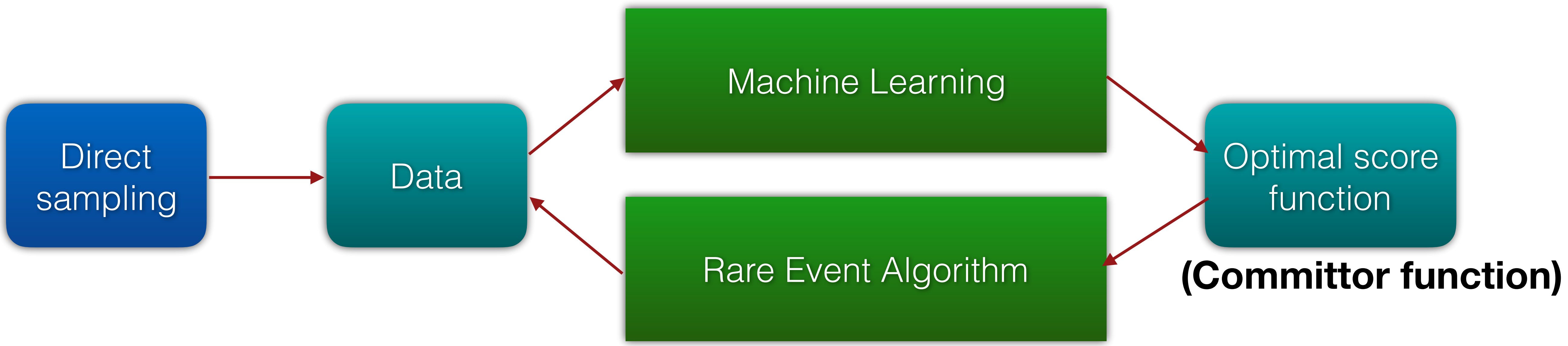
- For a given sample path $\{X(t)\}_{-\infty \le t < +\infty}$, the first hitting time τ_A is $\tau_A = \inf\{t \,|\, X(t) \in A\}$.

- The committor function *q(x)* of the sets A and B is defined as the probability that a trajectory starting at the point *x* reaches the set *B* before the set *A*

$$q(x) = \mathbb{P}_x \left(\tau_B < \tau_A \right) .$$

- How to estimate the committor function? With a rare event algorithm!

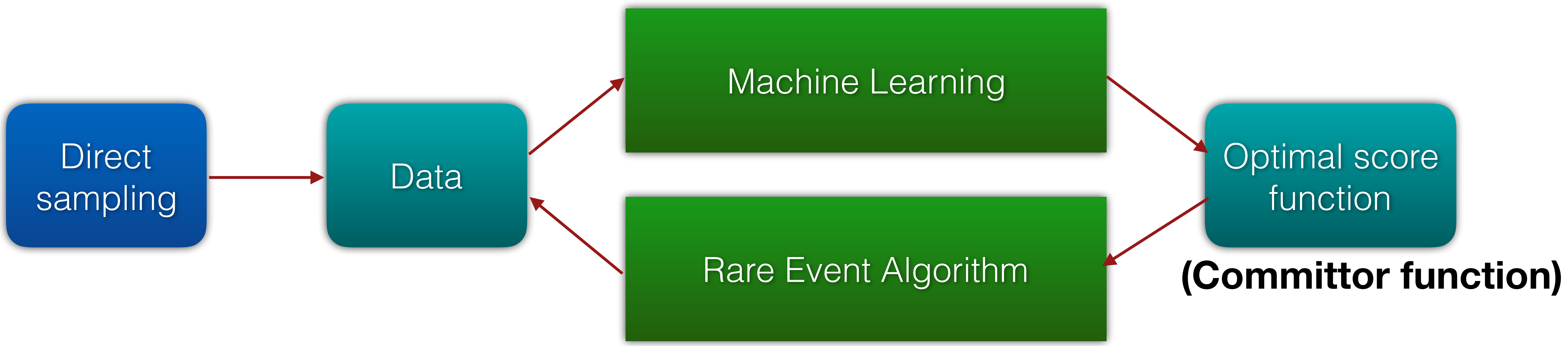
Coupling rare event algorithms with data based learning of committor functions



One example: Bouchet, Jack, Lecomte, Nemoto, PRE, 2016

(For X in dimension 1)

Coupling rare event algorithms with data based learning of committor functions

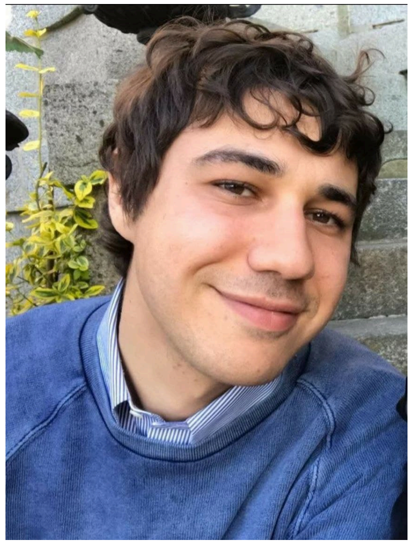


Work in progress for climate models!

IV-1) Committor functions for climate dynamics

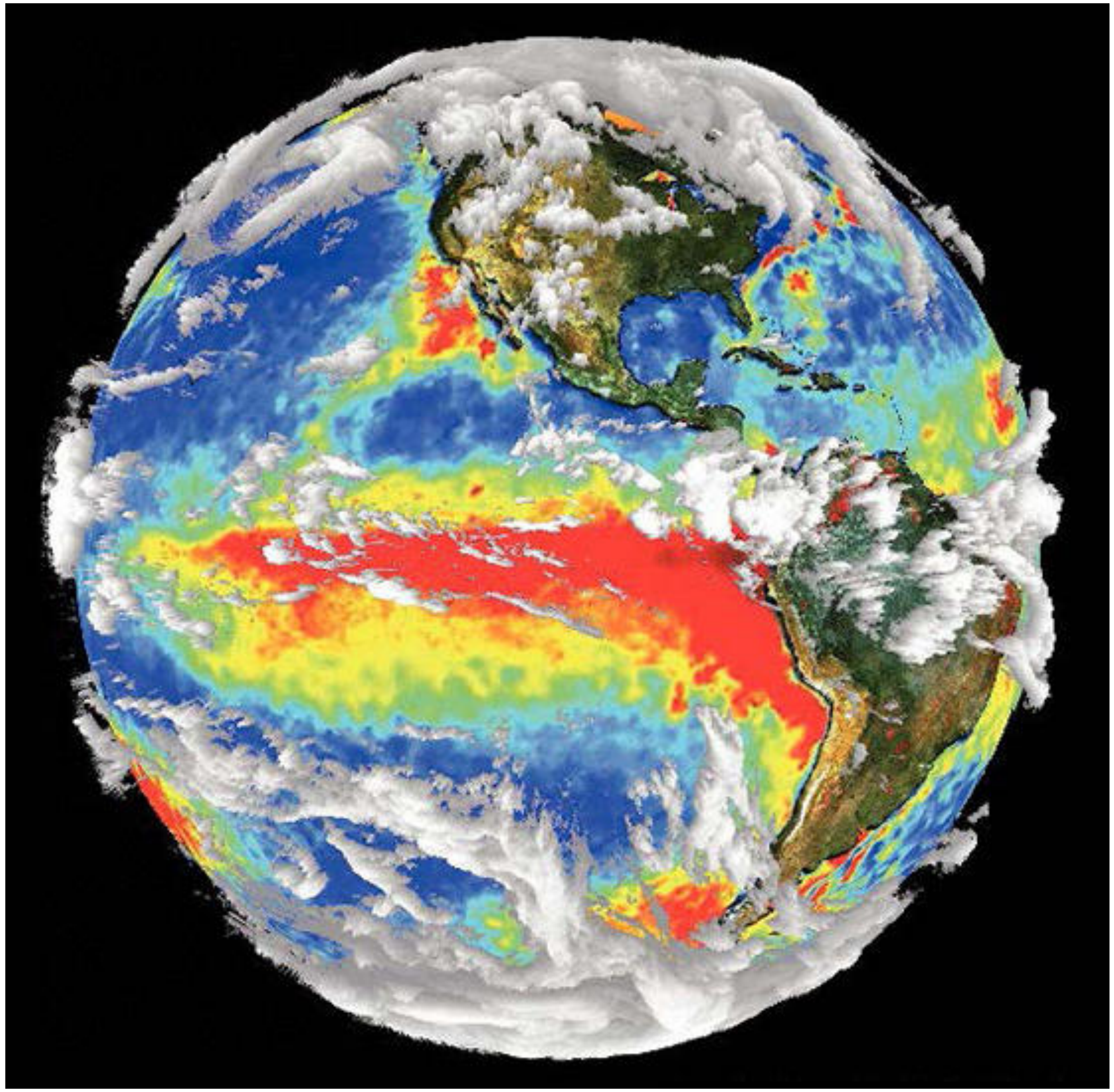
IV-2) Predictions at the predictability margin and committor functions for a model of ENSO (El Niño)

With Dario Lucente, Stefan Duffner, Corentin Herbert and Joran Rolland



Dario Lucente

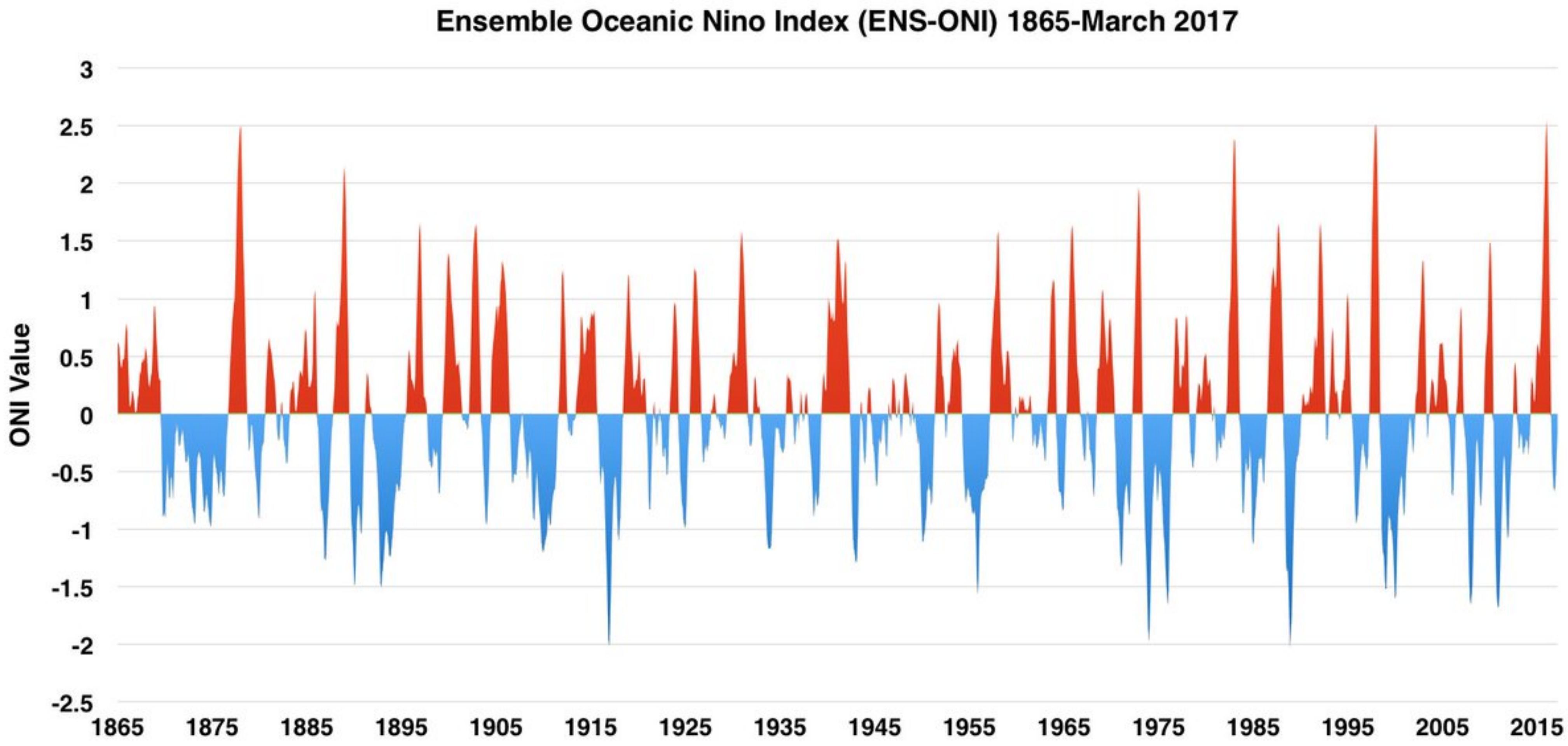
El Niño



Picture R. Houser
Washington university

Satellite observation of sea surface temperature and cloud cover during El Niño 1997-1998 event, the most intense one during the last century

El Niño

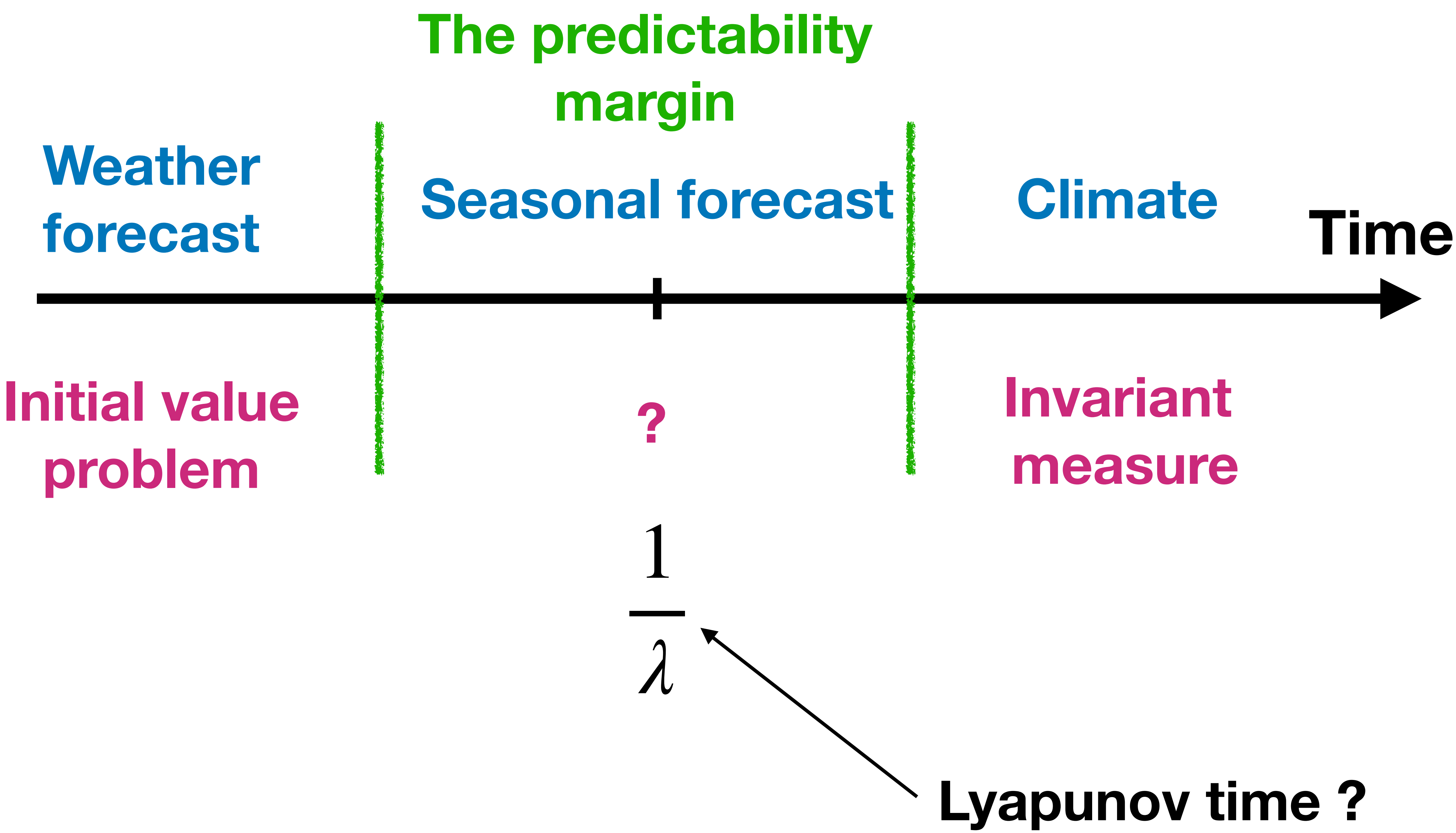


Some prediction problems:

- **Will we have El Niño next year?**
- **When will be the next strong El Niño?**

The predictability time is of the order of the Lyapunov time.

The predictability margin of chaotic dynamical systems or stochastic processes



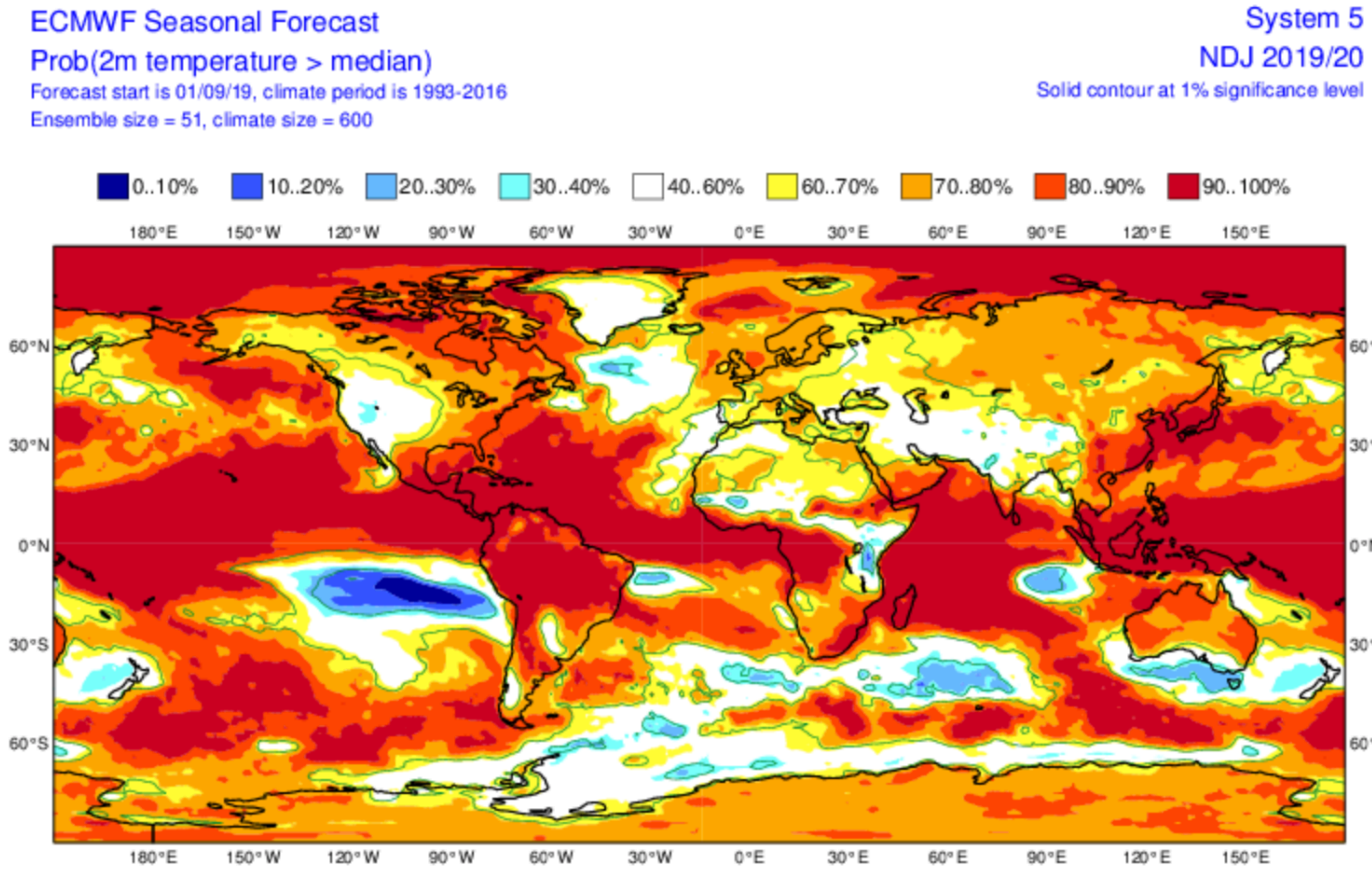
What should be predicted at the predictability margin ?

- Does the question: « Will we have El Niño next year?» make sense? (in the predictability margin range)
- Is it a deterministic forecast problem?

What should be predicted in the predictability margin ?

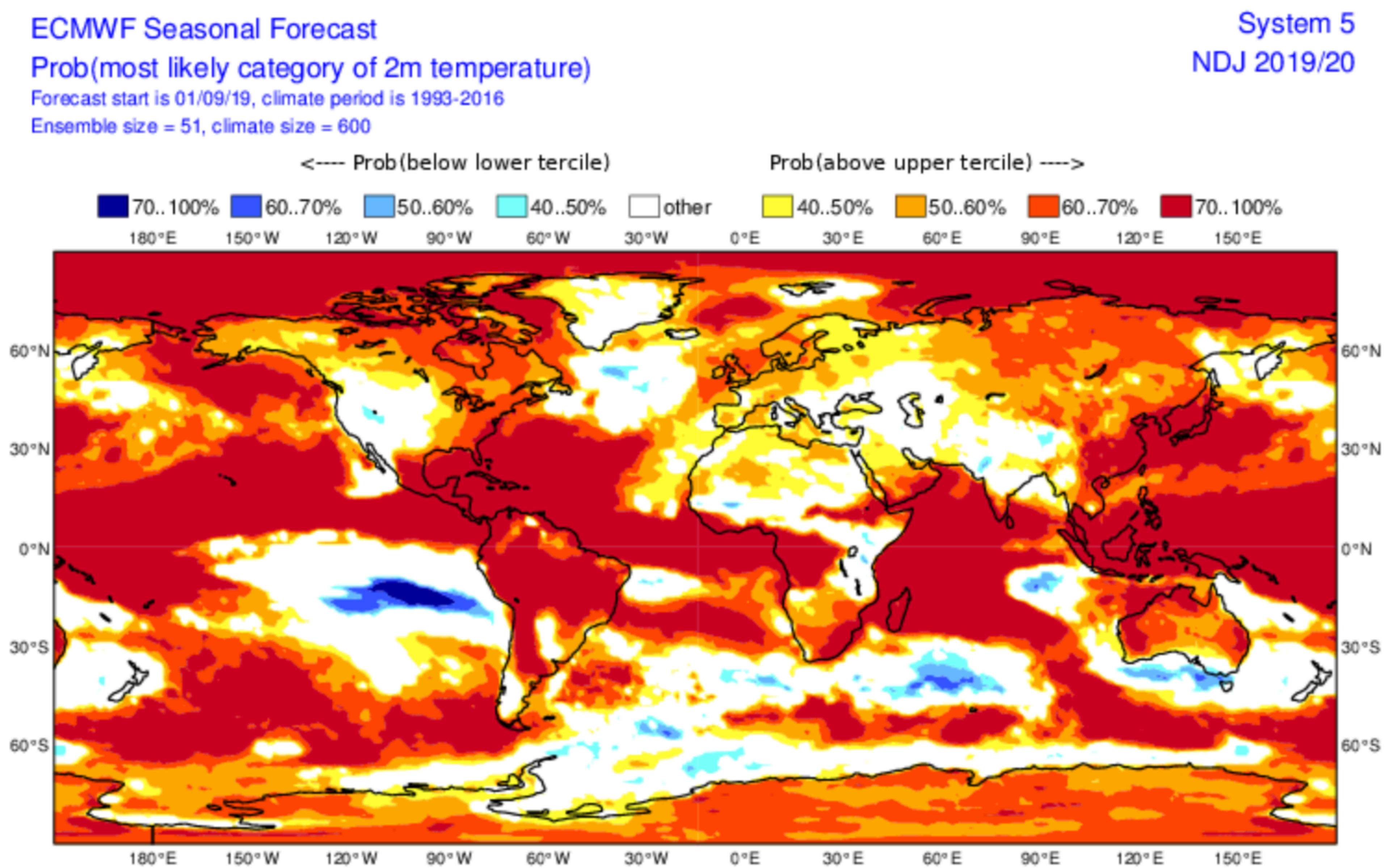
- ~~Does the question: « Will we have El Niño next year?» make sense? (in the predictability margin range)~~
- ~~Is it a deterministic forecast problem?~~
- Of course not.
- It is a probabilistic forecast problem.
- The question is « What is the probability to have El Niño next year? »

Seasonal Forecast is probabilistic



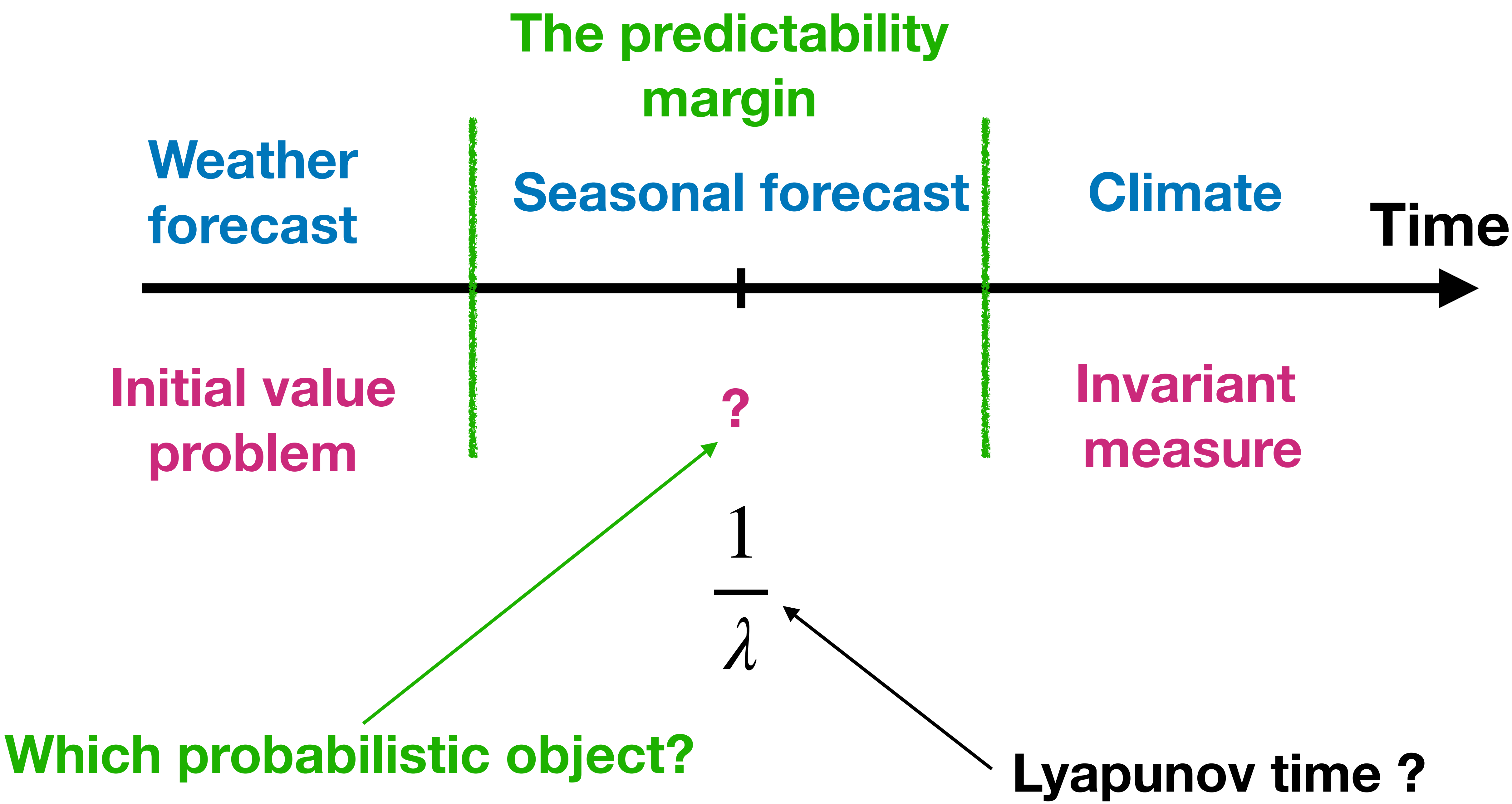
ECMWF seasonal forecast on 01/09/2019 of the 2 meter temperature in November 2019 - Probability of exceeding the median - (from ECMWF SEAS5 website)

Seasonal Forecast is probabilistic



ECMWF seasonal forecast on 01/09/2019 of the 2 meter temperature in November 2019 - Tercile summary - (from ECMWF SEAS5 website)

What is the mathematical concept of the predictability margin?



Committor function

- $\{X(t)\}_{-\infty \le t < +\infty}$ is a Markov process. A, B are subsets of the phase space.
- For a given sample path $\{X(t)\}_{-\infty \le t < +\infty}$, the first hitting time τ_A is $\tau_A = \inf\{t \,|\, X(t) \in A\}$.
- The committor function *q(x)* of the sets A and B is defined as the probability that a trajectory starting at the point *x* at the time 0 reaches the set *B* before the set *A*

$$q(x) = \mathbb{P}_x \left(\tau_B < \tau_A \right) .$$

- We will also consider

$$q(x,t) = \mathbb{P}_x \left(X(t) \in A \right) = \mathbb{E}_x \left(1_{X(t) \in A} \right),$$

which is also a committor function for an auxiliary process.

The Jin and Timmermann model

- In order to explain El Nino, Jin and Timmermann introduced a simple model which accounts for the recharge-discharge mechanism which is at the basis of ENSO.

- The relevant variables are:

1. The western SST of the Pacific Ocean T_1,
2. The eastern SST of the Pacific Ocean T_2,
3. The thermocline depth anomaly of the western Pacific h_1 that links them.

The Jin and Timmermann model

- With dimensionless units, the equations are

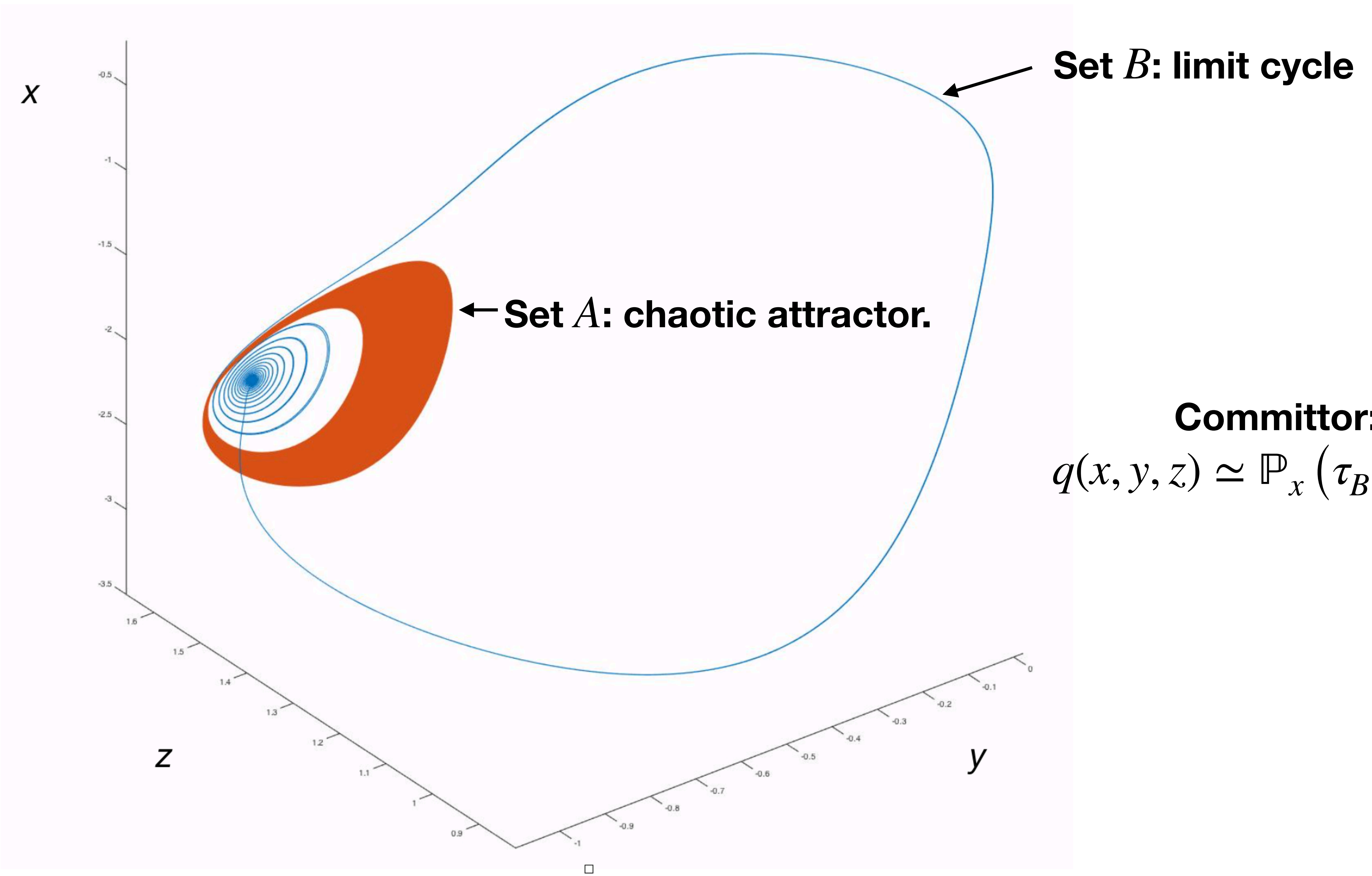
$$\dot{x} = \rho\delta\left(x^2 - ax\right) + x\left[x + y + c - c\tanh(x+z)\right] - D_x(x,y,z)\xi_t$$

$$\dot{y} = -\rho\delta\left(x^2 + ay\right) + D_y(x,y,z)\xi_t$$

$$\dot{z} = \delta\left(k - z - \frac{x}{2}\right)$$

- This is chaotic deterministic dynamics perturbed by a small noise.

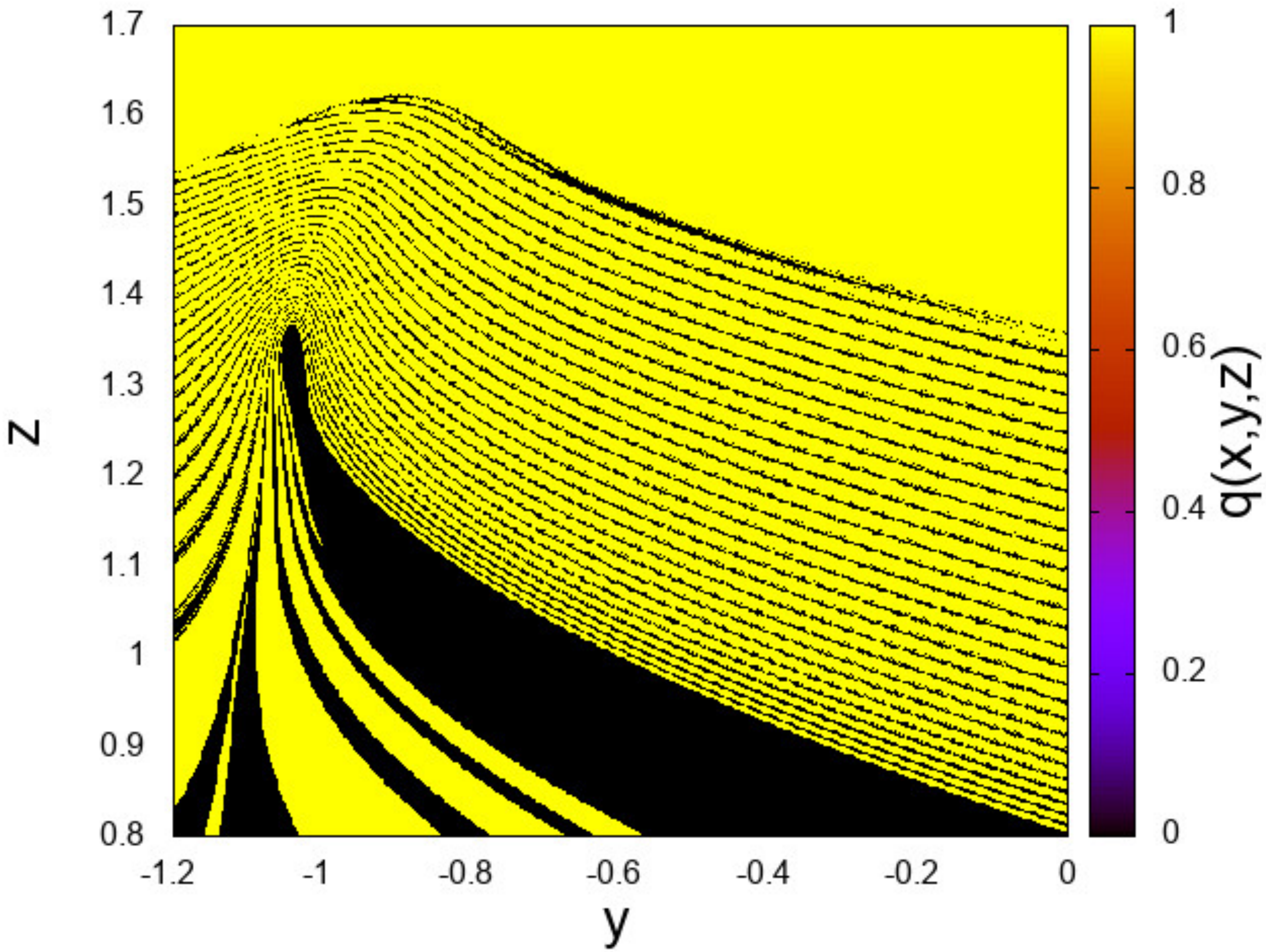
The chaotic attractor and the periodic orbit of the Jin and Timmermann model



Committor:

$$q(x, y, z) \simeq \mathbb{P}_x\left(\tau_B < \tau_A\right) .$$

A cut of the committor function. Deterministic case



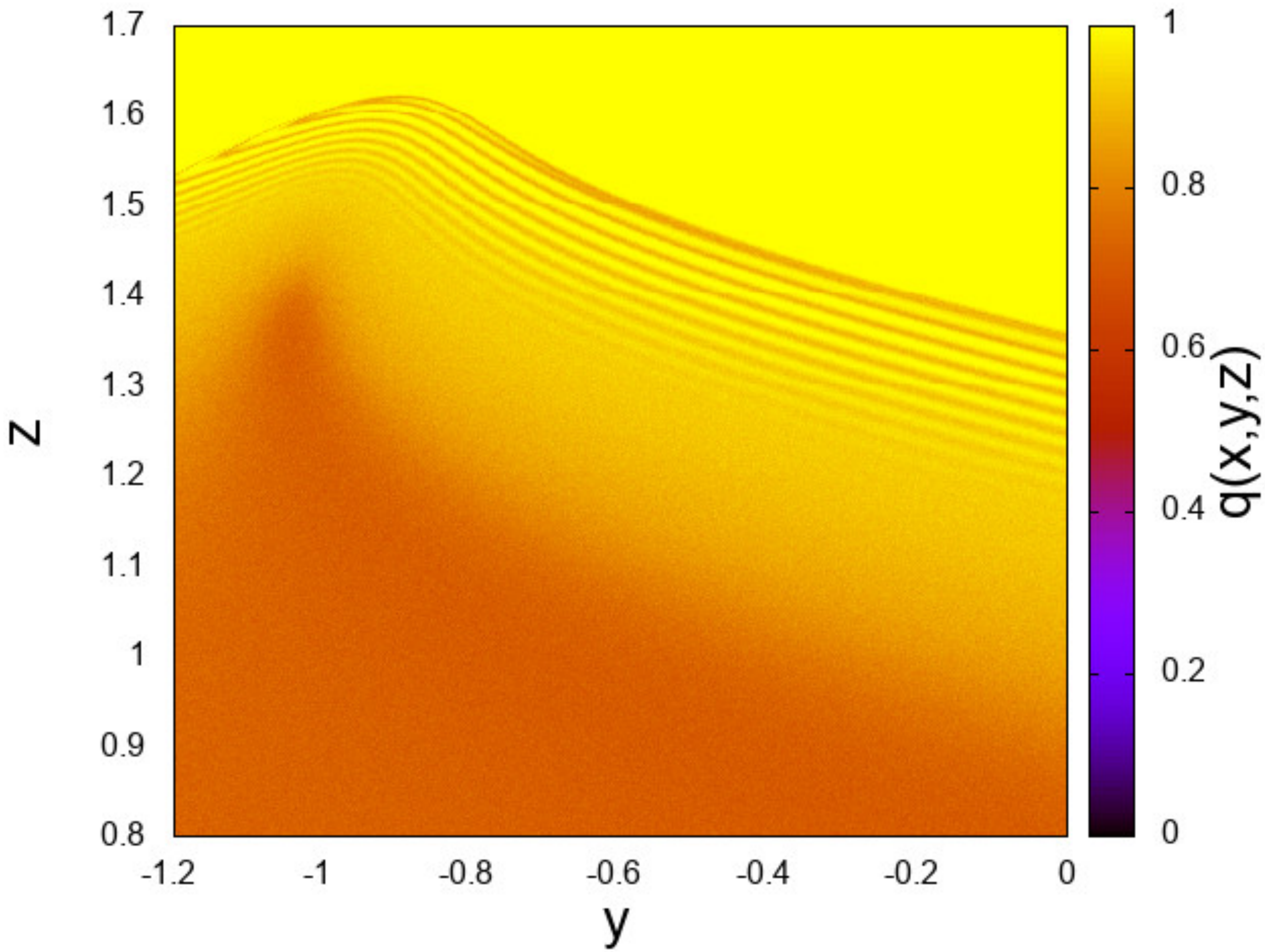
$x = -2.831$

$\sigma = 0.$

Noise strength

Deterministic committor $\simeq$ basins of attraction

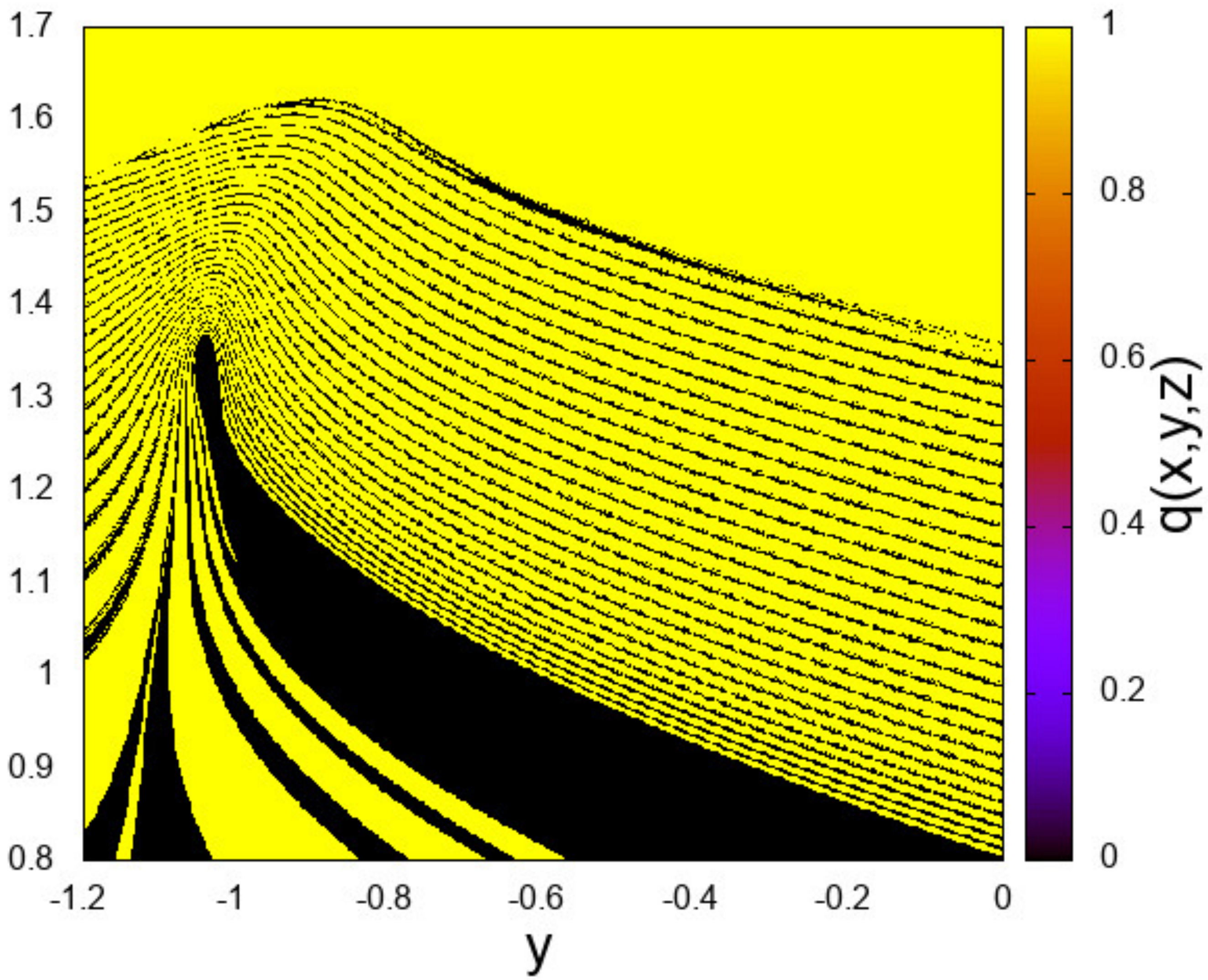
A cut of the committor function. High noise case



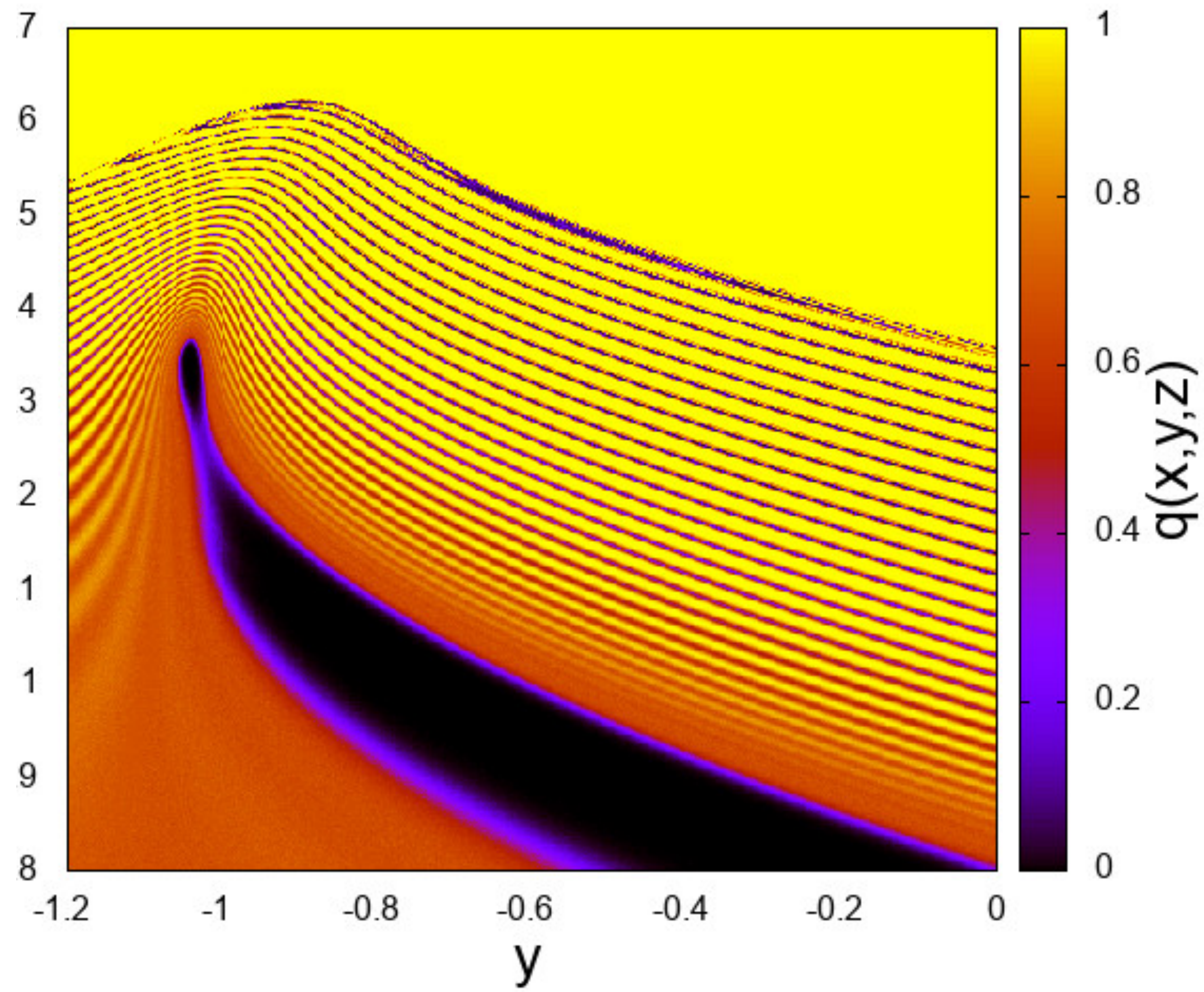
$x = -2.831$

$\sigma = 10^{-3}$.

A cut of the committor function



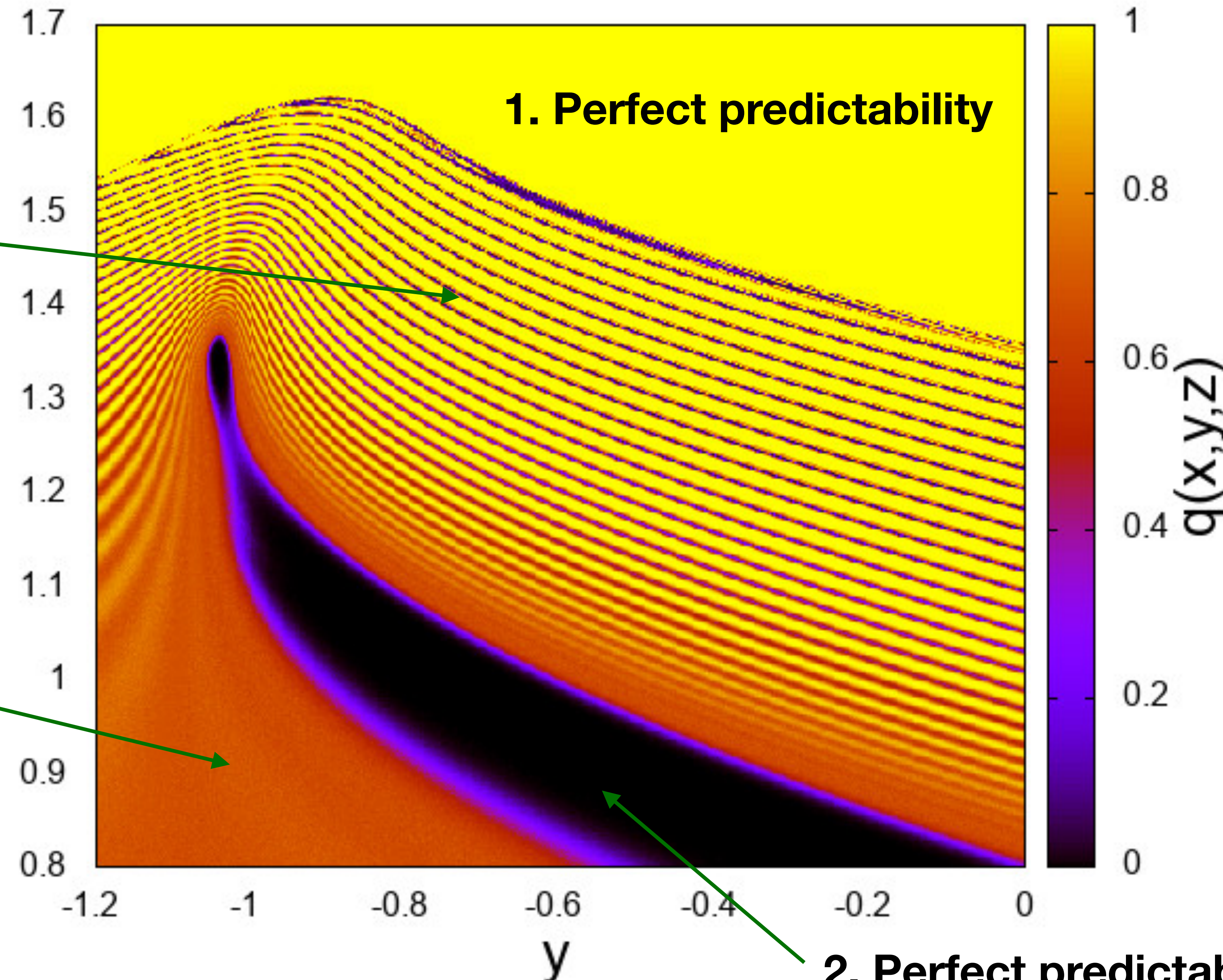
$\sigma = 0.$



$\sigma = 5.10^{-5}.$

Deterministic and intermediate noise cases.

A cut of the committor function. Intermediate noise case

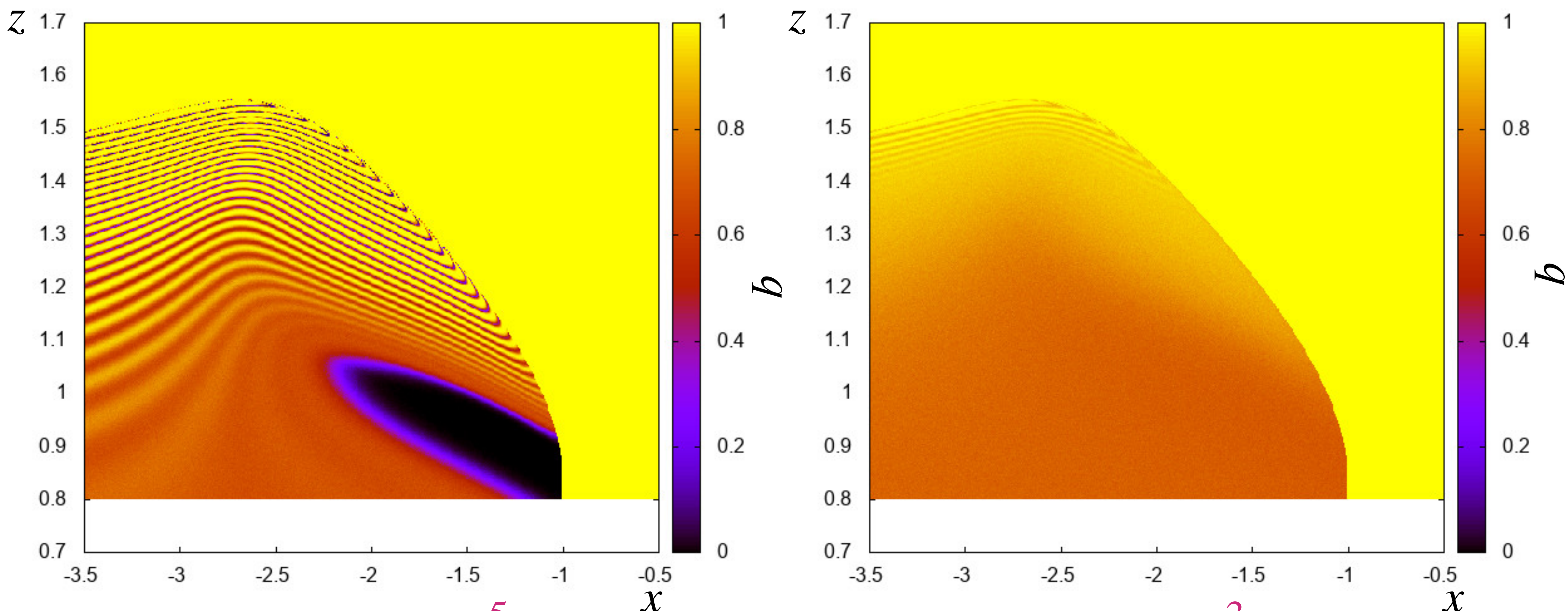


3. Deterministic like sensitivity to initial condition
Hard potential predictability

4. Smooth probability
Easy potential predictability of the probability

2. Perfect predictability

Another cut of the committor function



$\sigma = 5.10^{-5}$ $\sigma = 10^{-3}$

$y = -1.158$

Intermediate and high noise cases.

Conclusions: Predictions at the Predictability Margin and Committor Functions

- At the predictability margin, predictions should be probabilistic. **The committor functions are the proper mathematical objects.**

- For a simple dynamics of El Niño, which is a small stochastic perturbation of a chaotic deterministic system, we have computed a committor function for a transition to occur.

- The committor functions shows **areas of the phase space with hard, respectively easy, probabilistic predictability potential** and quantifies the probability of the event.

- This informs us on what to expect for a predictability problem at the predictability margin, **in a perfect information context.**

D. Lucente, S. Duffner, C. Herbert, J. Rolland, and F. Bouchet, proceeding of Climate Informatics 2019.

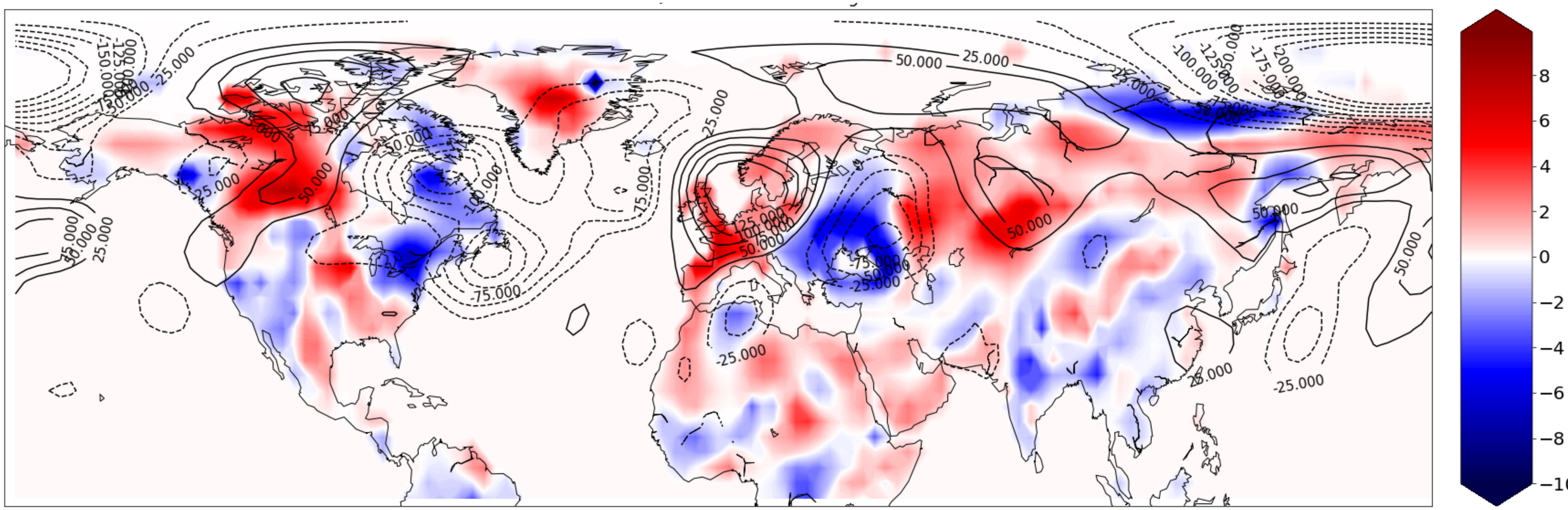
D. Lucente, C. Herbert, and F. Bouchet, to be submitted to Climate Dynamics.

IV-2) Predicting heat waves (committor functions) with deep neural networks

With P. Abry, P. Borgnat, V. Jacques-Dumas and F. Ragone

Valerian Jacques-Dumas

Predicting heat waves with a deep neural network - 1) Data



Surface temperature (T_s, colors) and 500 hPa geopotential height (Z_g, lines) anomalies

- **Plasim and CESM climate models.**
- We use summer (JJA) data: 8 maps/day, 90 days/year, 1000 year = 720 000 maps.
- For Plasim data, each field has a resolution 64×128, restricted to 25×128 above 30^o North.

Heat wave definition

- $X(t) = T_s$ field at time t, or $X(t) = (T_s, Z_g)$ fields at time t.

- $Y(t)$: **time and space averaged surface temperature anomaly within τ days:**

$$Y(t) = \frac{1}{D}\int_{t+\tau}^{t+\tau+D} \frac{1}{|\mathscr{A}|}\int_{\mathscr{A}} T_s(\vec{r}, u)\ \mathrm{d}\vec{r}\ \mathrm{d}u,$$

$$\text{and } Z(t) = 1 \text{ if } Y(t) > a, \text{ and } Z(t) = 0 \text{ otherwise}$$

- $Z(t) \in \{0,1\}$. A heat wave occurs if $Z = 1$.

- We have a classification problem for the data (X, Z). We want to learn the probability $q(x)$ that $Z = 1$ given that $X = x$ (committor function).

- 5% most extreme events: $a = a_5 = 3.08\ \mathrm{K}$. 2.5% most extreme events: $a = a_{2.5} = 3.7\ \mathrm{K}$. 1.25% most extreme events: $a = a_{1.25} = 4.23\ \mathrm{K}$.

Predicting heat waves with a deep neural network

Observing the temperature and geopotential height at 500 hPa today, what is the probability to observe a D-day heat wave starting τ days from now?

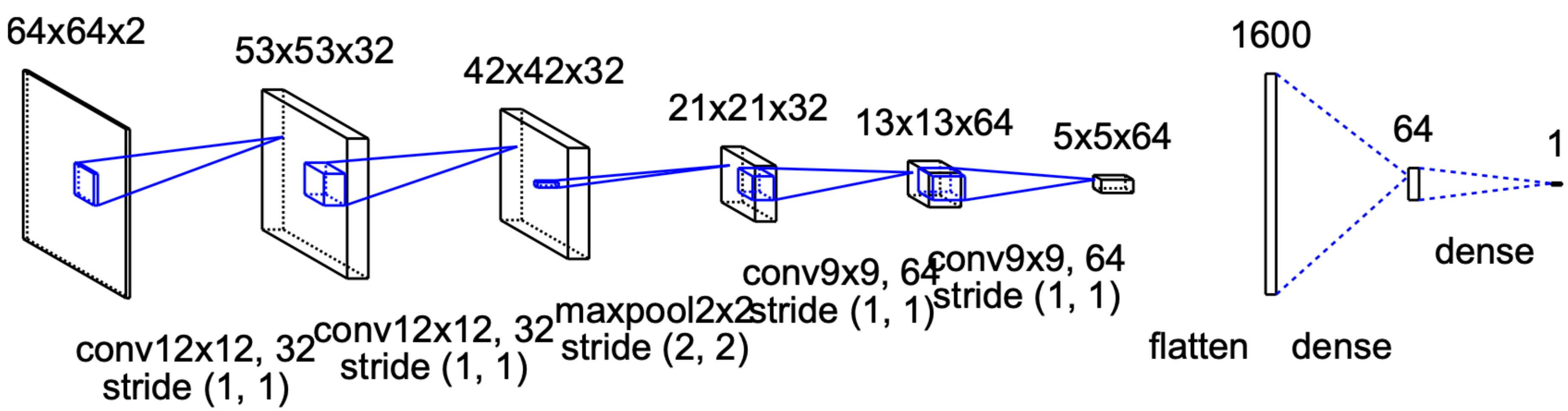
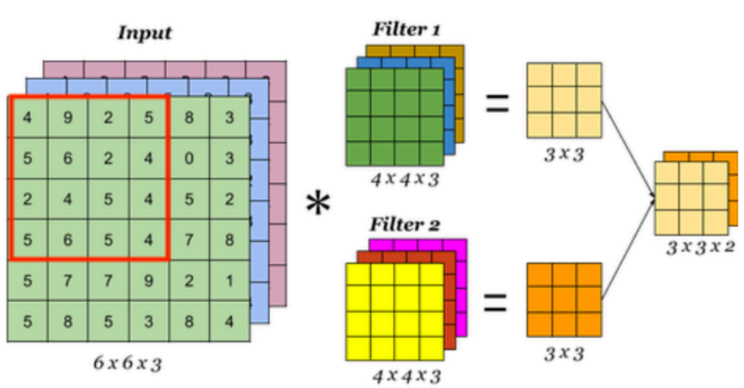
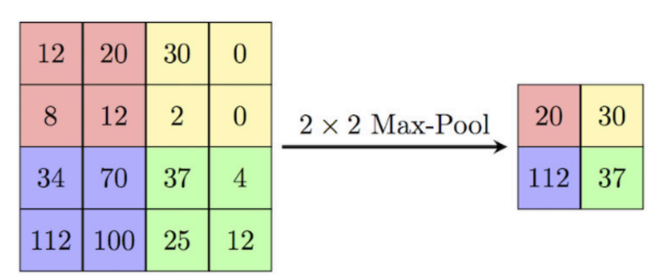


Figure 2: Architecture of the CNN used to forecast extreme heatwaves.



(a) A Convolution layer

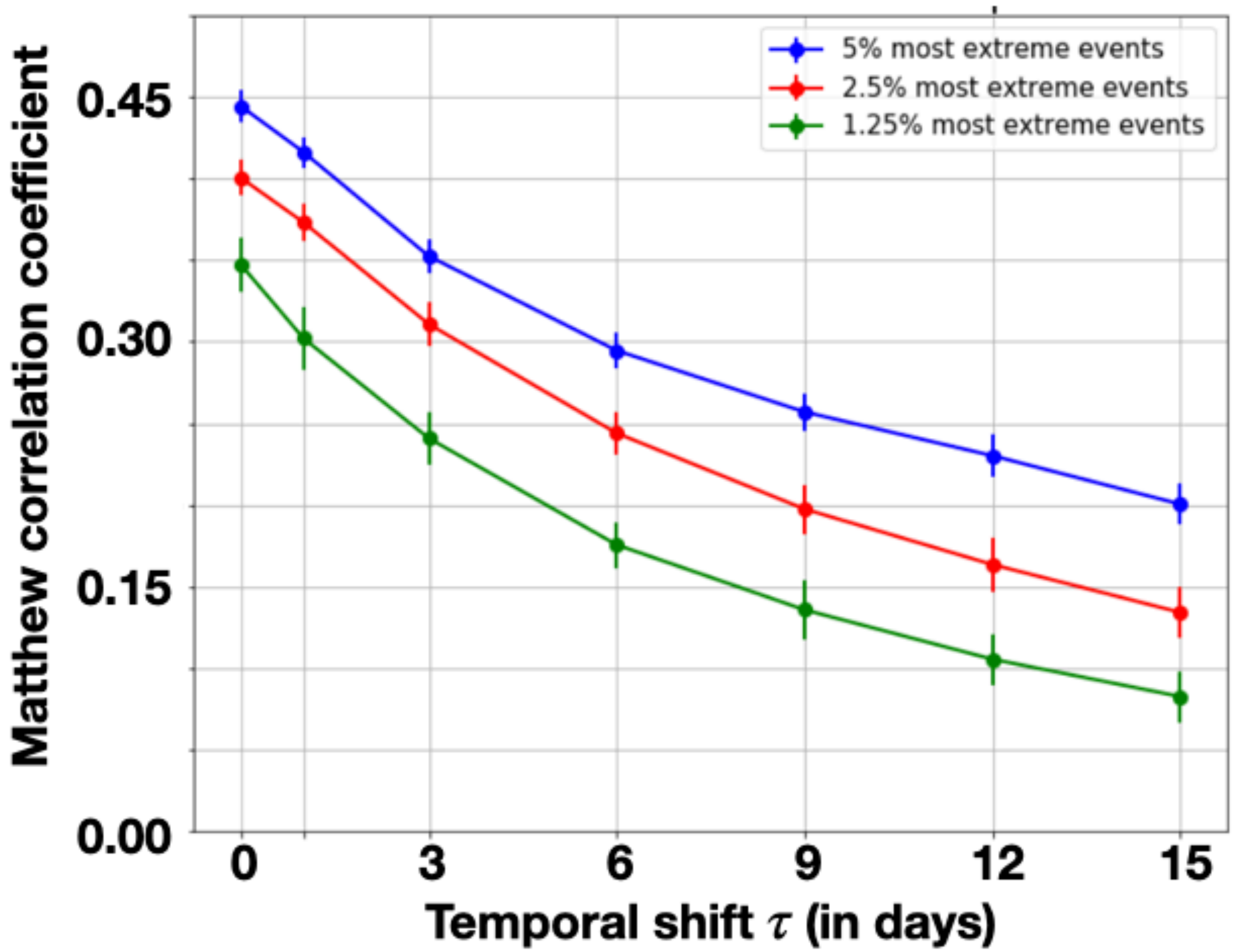


(b) A MaxPool layer

Machine learning for extreme heat waves

- **Supervised learning** from 1,000 years of climate model data (720 000 couples (X, Z)).
- We use **undersampling** to deal with class imbalance.
- We use **transfer learning** between return levels a, first training a deep neural network for less rare events, and then transferring to learn rarer events with less data.

Predicting heat waves



Heat waves over France
$D = 14$

Predictability, τ day ahead, for a 14-day heatwave from the temperature and GPH fields

We have very interesting prediction capabilities up to 15 days ahead of time for $D = 14$-day heatwaves

V. Jacques-Dumas, F. Ragone, F. Bouchet, P. Borgnat and P. Abry, 2021, sub. to IEEE TPAMI + ArXiv

Conclusions

- We can use **rare event algorithms** to gather an amazing statistics for extreme heat waves with Plasim (PNAS, 2018), and CESM (GRL and ArXiv, 2021).
- The dynamical mechanism is the birth of **quasi-stationary non zonal global patterns,** which are much affected by topography and oceans (PNAS, 2018, GRL 2021).
- **Models reproduce correctly those extreme teleconnection patterns** for moderate extremes, compared with the ERA5 reanalysis dataset. **We need model and rare event algorithms to study more extreme heat waves teleconnection patterns.**
- Studying the committor function for El-Nino transitions, we introduced the notions of **hard, versus easy probabilistic predictability potential.** (Climate informatics 2019, Sub. To Climate dynamics 2021).
- **Machine learning** has the potential to give **meaningful statistical predictions for long-lasting heat wave up to 2 weeks ahead of time** (Sub. to IEEE TPAMI, 2021).

Please join us to study climate extreme events

The scientific questions are fascinating!