Large deviation theory applied to study rare and extreme events in turbulence, atmosphere, and climate dynamics - Lecture III

F. BOUCHET – CNRS and ENS de Lyon

Boulder summer school 2022 – Hydrodynamics Across Scales – 07/2022

ENS DE LYON

European Research Council
Established by the European Commission

SIMONS FOUNDATION

Outline

I) Introduction to large deviation theory and its applications to dynamical problems (Wednesday)
II) Large deviation theory for kinetic theories, geostrophic turbulence, and atmosphere dynamics (Thursday)
III) Rare and extreme events in climate dynamics: sampling using rare event algorithms and machine learning (Friday)

RARE AND EXTREME EVENTS IN CLIMATE DYNAMICS: SAMPLING USING RARE EVENT ALGORITHMS AND MACHINE LEARNING

Freddy Bouchet (CNRS and ENS de Lyon)

With:

a) **Eric Simonnet and Joran Rolland** **(Jupiter's abrupt climate change)**
b) **Francesco Ragone and Jeroen Wouters** **(rare event algorithms in climate models)**
b) **Dario Lucente, George Miloshevich and Francesco Ragone** **(extreme heat waves)**
c) **Valerian Jacques-Dumas, George Miloshevich, Francesco Ragone, Pierre Borgnat and Patrice Abry** **(prediction of extreme heat waves with deep neural networks)**
d) **Dario Lucente, Joran Rolland and Corentin Herbert** **(coupling rare event algorithms and machine learning)**

Utrecht, workshop Critical Earth, April 2022

cnrs dépasser les frontières

ENS DE LYON

Lab PHYS ENS de LYON

erc European Research Council Established by the European Commission

anr

https://labos1point5.org

Labos 1point5

Collectif Recherches Vidéos Outils Ressources Actualités

Reducing the environmental footprint of our research activities

GES 1POINT5

Labos 1point5 launches a national scientific study on the **carbon footprint** of the **French public research** to provide food for thought on the levers of action to reduce its impact on the climate and the environment.

Why use GES 1point5?

The answer in 5 points.

#GHG INVENTORIES	#LABORATORIES	#INITIATIVES
809	476	89

THE COLLECTIVE

Labos 1point5 is a collective of members of the academic world whose activities are organized within a research group and a think tank.

To participate in the collective's activities, you can join the GDR and the think tank.

Subscribers follow the collective's activities with the newsletter.

#RG MEMBERS	#THINK TANK MEMBERS	#SUBSCRIBERS
237	457	3567

FOLLOW US

Subscribe to our **newsletter**. This does not commit you to anything, but allows you to stay informed of the various activities of the collective.

Email *

Email *

Send

https://labos1point5.org

GDR « Theoretical challenges for climate sciences »

- **Identify and work on key theoretical issues that need to be solved for improving the quantitative predictions in climate sciences.**

- **A multidisciplinary consortium: climate sciences, mathematics, physics, computer sciences, statistical physics, data sciences.**

- **Examples :** i) How to reduce the uncertainty about climate sensitivity? ii) How to reduce uncertainty when quantifying probabilities of climate extreme events? iii) How to integrate data and theoretical constraints, using machine learning, to build the next generation of climate models? iv) How to make quantitative the study of future and past climate? v) How to build effective coarse-grained descriptions of climate processes?


Institut des Mathématiques pour la Planète Terre
IMPT

How to build a strategy for the ecology transition at ENS de Lyon

Proposition by the ecology transition group at ENS de Lyon:

Chapter 1 Ecology transition at ENS de Lyon : how to build together a cultural change?

Chapter 2 Quantifying the environmental impact of ENS de Lyon

Chapter 3 How to reduce the impact of building and infrastructures

Chapter 4 Impact of travels

Chapter 5 Daily life and environment (recycling, wastes, transport, bikes, ...)

Chapter 6 Environmental impact of digital technologies

Chapter 7 Teaching, research, and the environment

Lien Web : lettre d'interpelation des collegues et de la direction (01/2021)

Lien Web : construire une stratégie (01/2021)

Rare event algorithms for climate dynamics
Outline

I) Introduction: rare events do matter - rare event algorithms

II) Rare events algorithms for predicting extreme heat waves

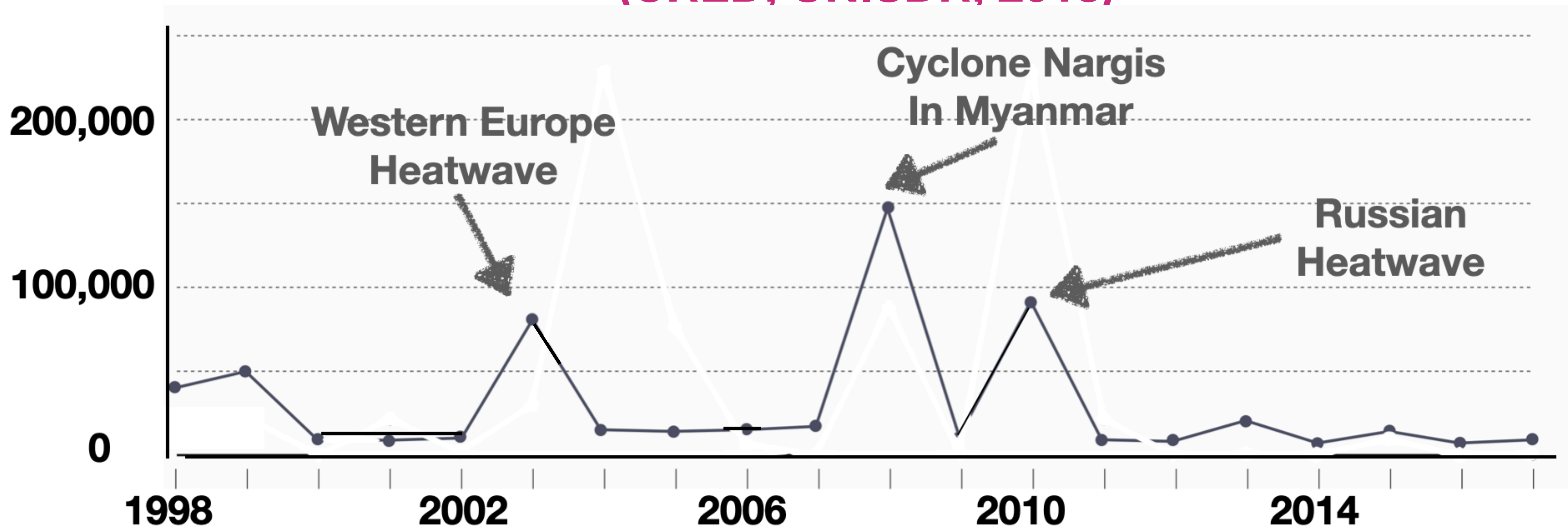
III) Predicting extreme heat waves and committor functions using deep neural networks

IV) Coupling rare event algorithms with machine learning

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

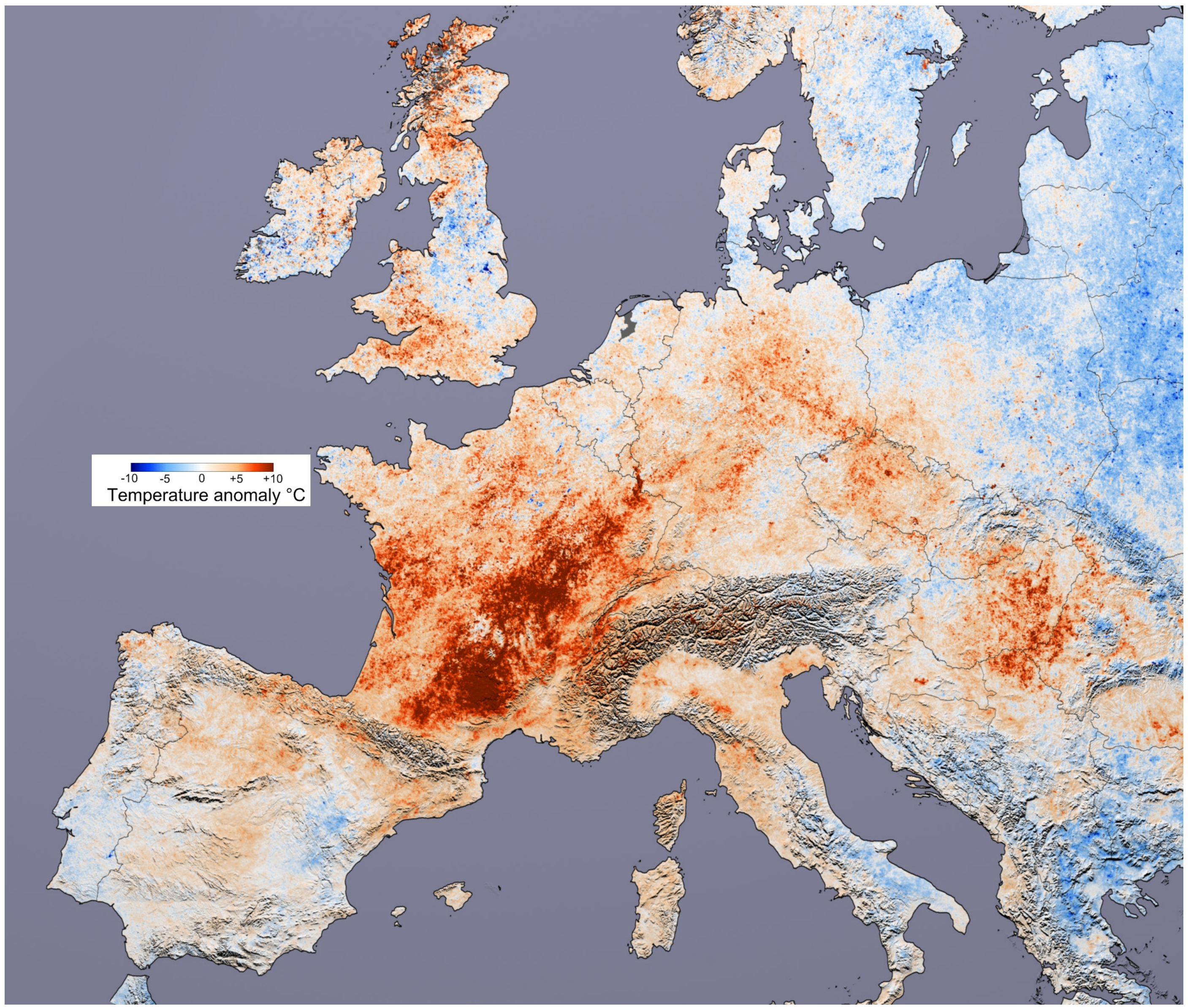
Annual deaths by major climate related disaster

(CRED, UNISDR, 2018)



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

What is the probability (return time) of the 2003 Europe heatwave ?



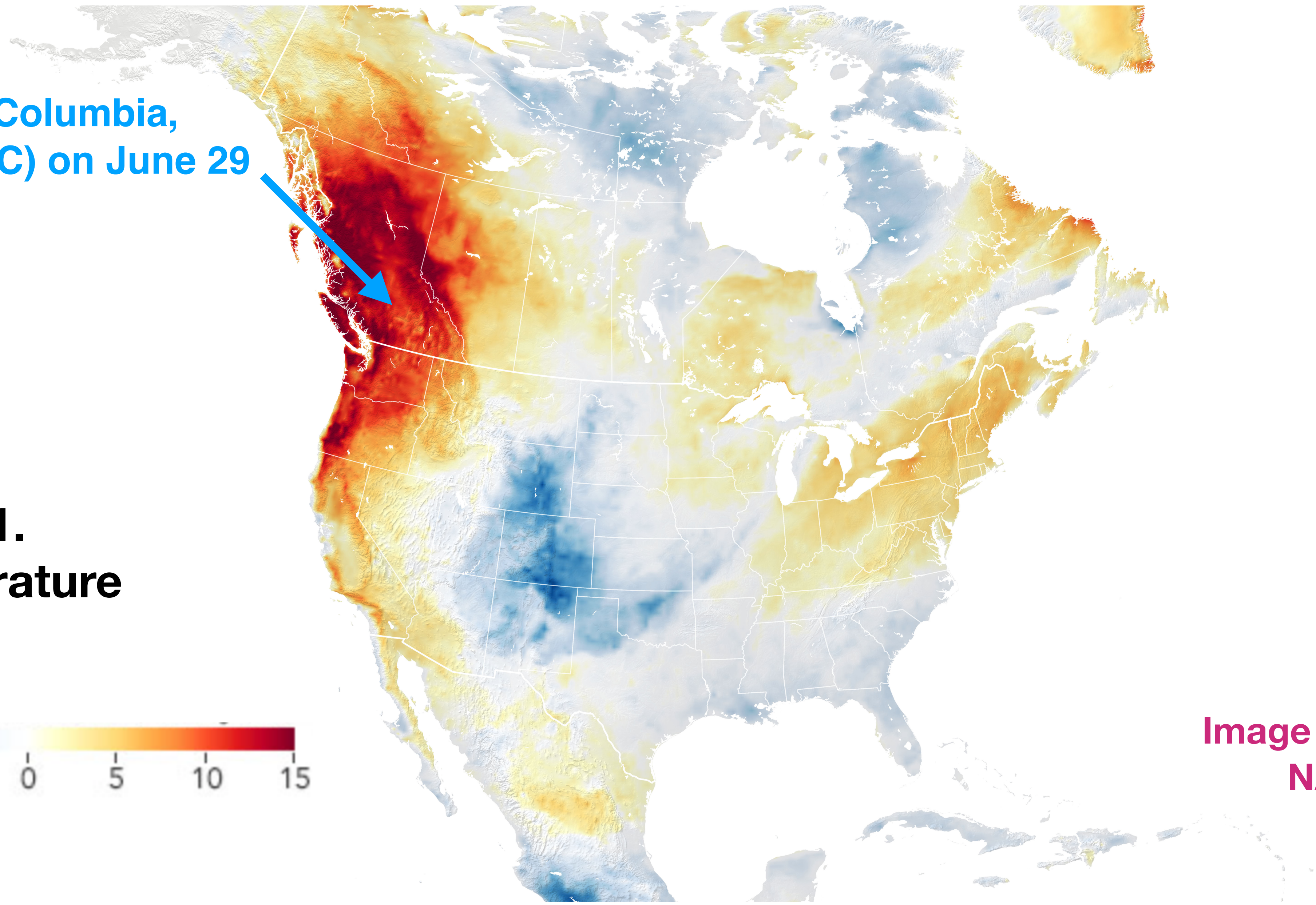
July 20 2003-August 20 2003 land surface temperature minus the average for the same period for years 2001, 2002 and 2004 (TERRA MODIS).

Why are return times so hard to estimate?

i) lack of observation data, ii) model biases,

iii) because of rareness, gathering good model statistics is too costly.

The 2021 northwest America heatwave - Unprecedented



Lytton, British Columbia, hit 121°F (49.6°C) on June 29

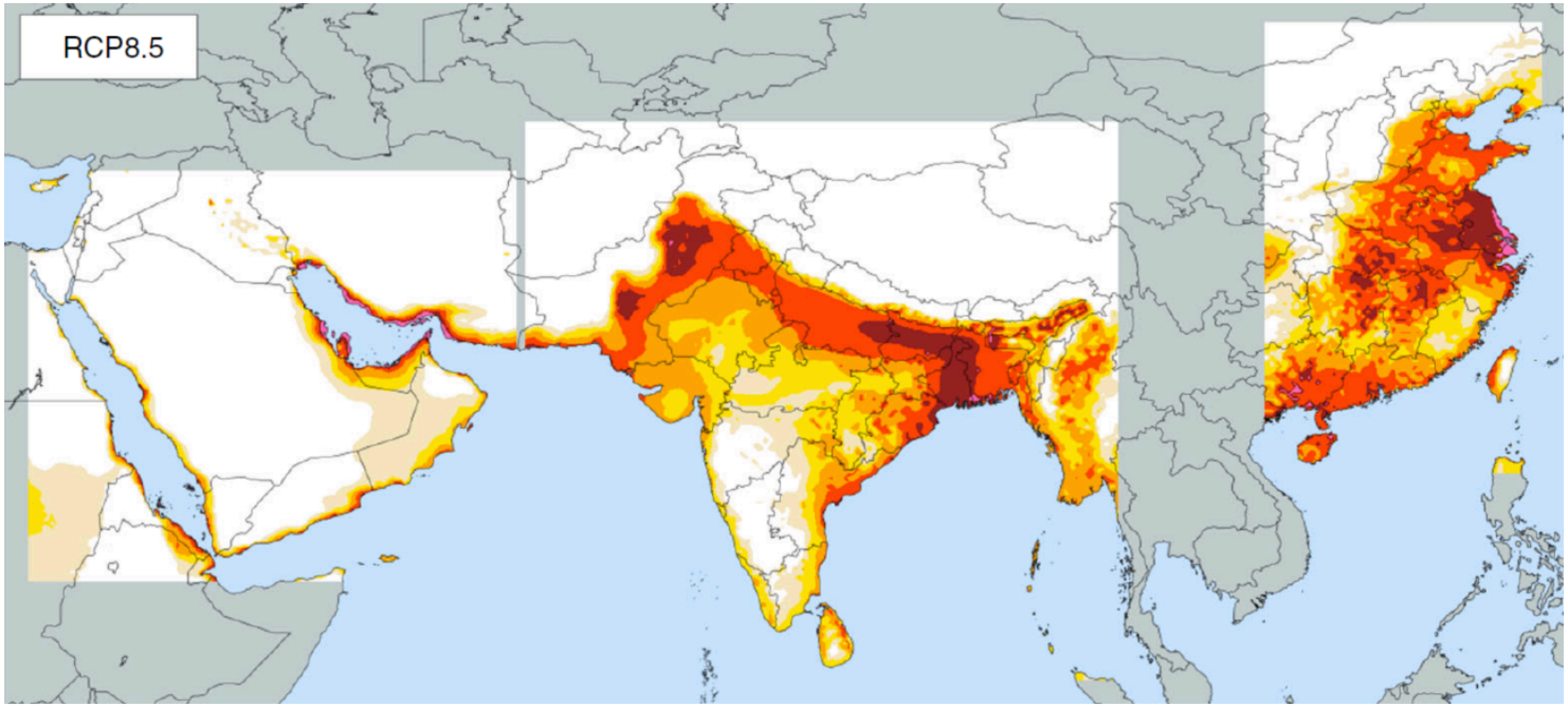
June, 27, 2021. 2m air temperature anomaly (°C).

Image Joshua Stevens
NASA, GEOS

How often shall we expect this to happen?

We do not know (see WWA study).

Potential impacts of global warming and extreme events

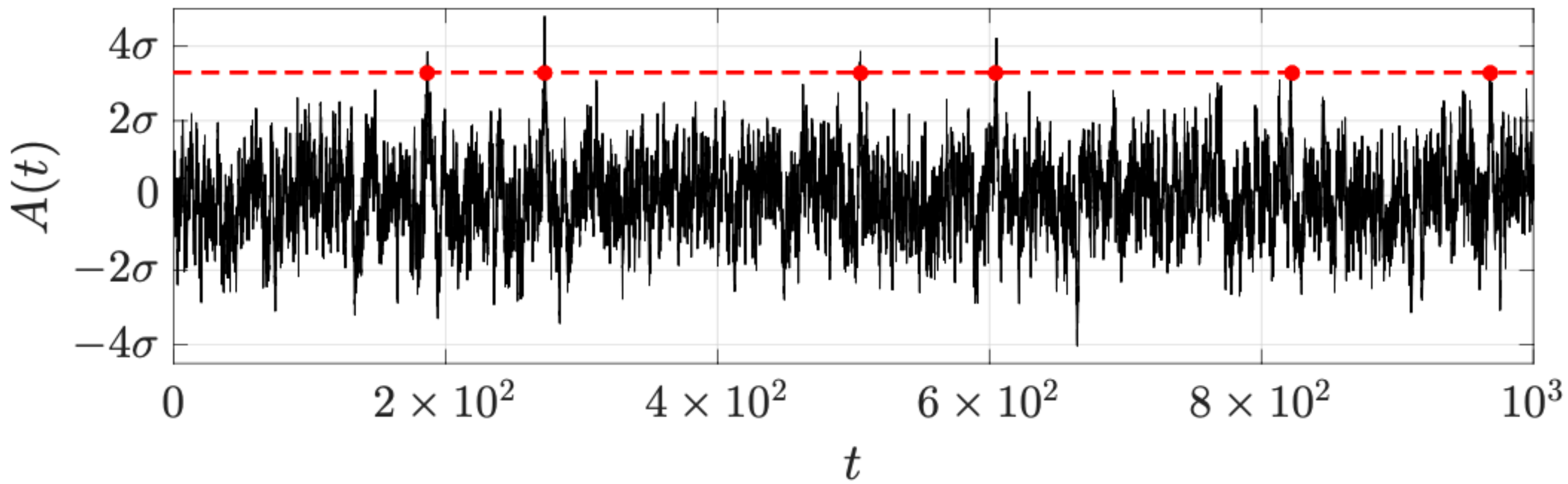


Maximal wet bulb temperature (red color =31-32°C), in 2070, with the RCP8.5 scenario.

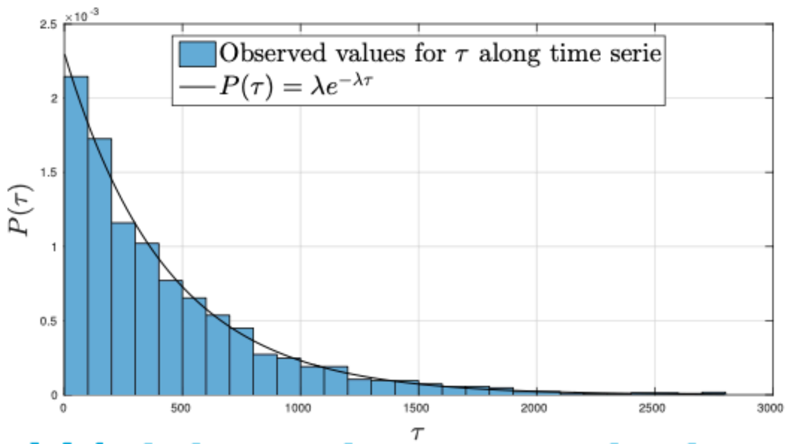
(Kang, Elfatih and Eltahir, 2018)

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.

Extreme Events, Poisson Statistics, and Return Times



Ornstein–Uhlenbeck extremes



Waiting time statistics

For systems with a single state, rare enough events are uncorrelated and have a Poisson statistics

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 they are, the more computationally costly they are.

- 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?

How to study a 10 000 year heat wave with a 200 year simulation ?

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

- **Rare event algorithms:** Kahn and Harris (1953).

- **Statistical mechanics:** diffusion Monte-Carlo, Wang Landau algorithms, go with the winners, …

- **Applied Mathematics:** Chandler, Vanden-Eijnden, Schuss, Del Moral, Dupuis, Lelièvre, Guyader, …

- **For turbulence and climate applications:** J. Weare and D. Abbot, R. Grauer and T. Grafke, E. Vanden-Eijnden, Lyon group, …

Rare event algorithms for climate dynamics
Outline

I) Introduction: rare events do matter - rare event algorithms

II) Rare events algorithms for predicting extreme heat waves

III) Predicting extreme heat waves and committor functions using deep neural networks

IV) Coupling rare event algorithms with machine learning

III) Rare event algorithms to study extreme heat waves with climate models

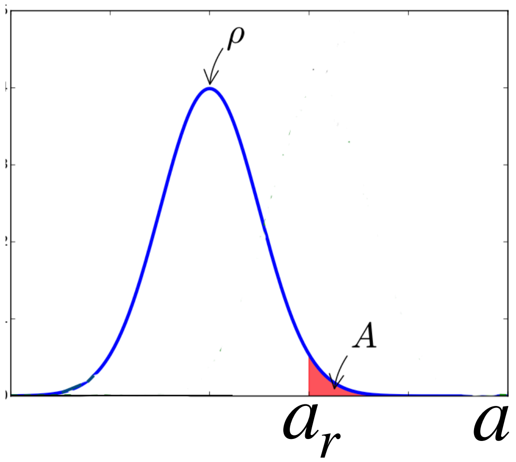
Francesco Ragone
RMI, Bruxelles, Belgium

Jeroen Wouters
University of Reading, UK

Long lasting summer heat waves

We will study extremes of the time averaged temperature:

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



- $\mathscr{A} =$ Scandinavia, Europe, France, Alberta, Russia, …
- $T =$ one week, a few weeks, a month, or a season.
- **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

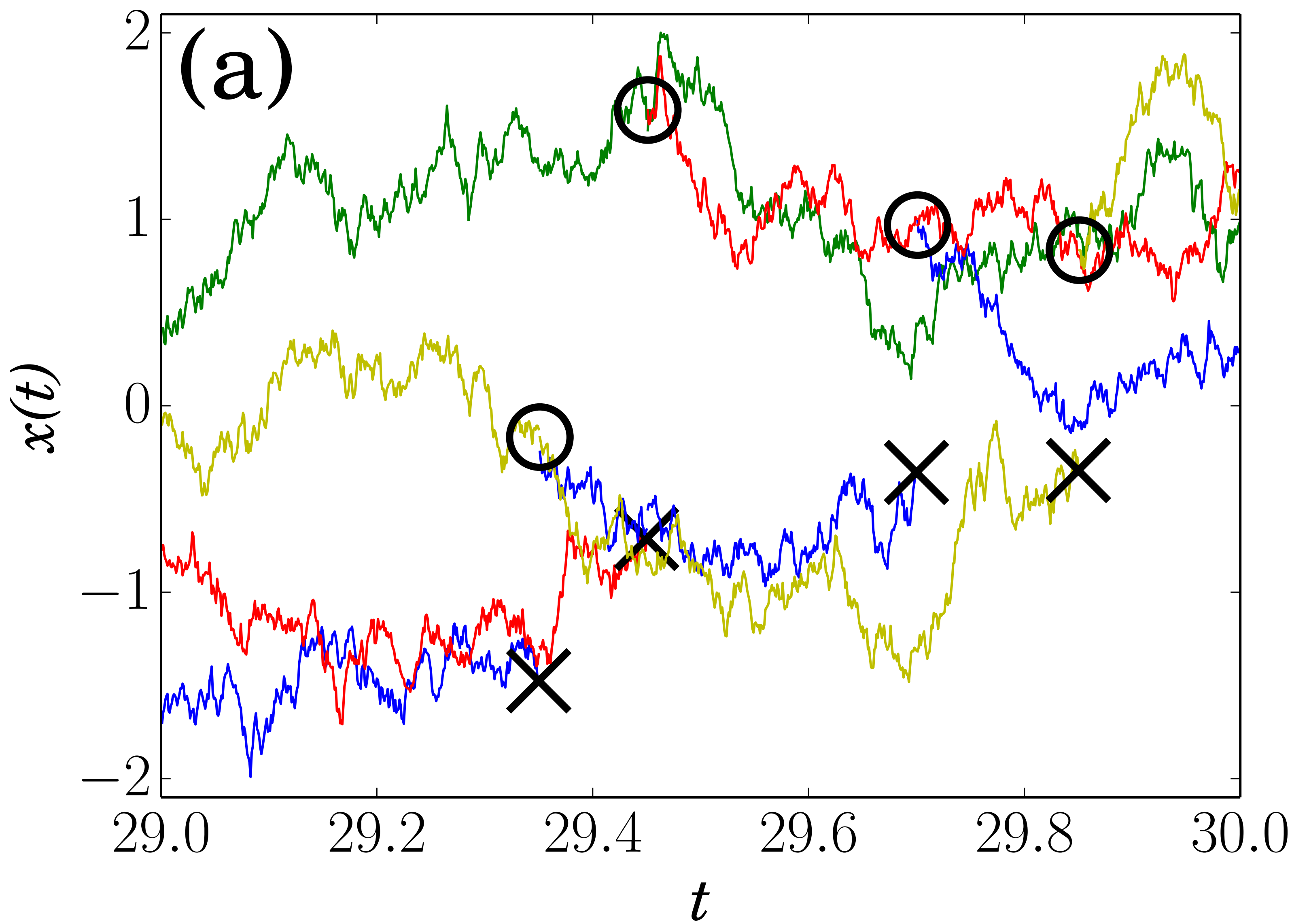
$$\tilde{P}_k\left(\{X(t)\}_{0\leq t\leq T}\right) = \frac{1}{\exp(T\lambda(k))} P_0\left(\{X(t)\}_{0\leq t\leq T}\right) \exp\left[k\int_0^T A[X](t)\,\mathrm{d}t\right].$$

- 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

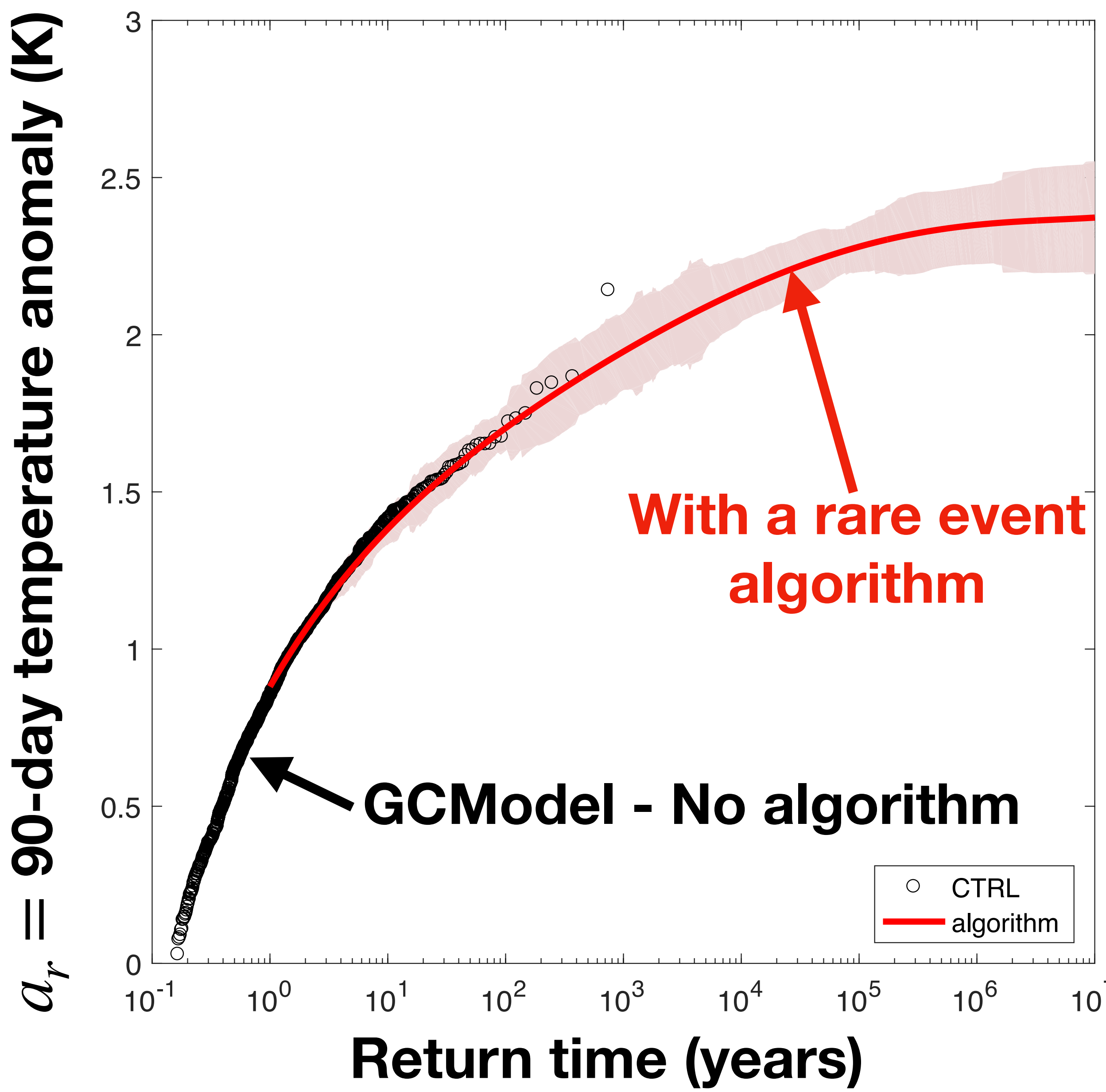


The trajectory statistics is tilted towards the events of interest.

Sample paths of the Giardina Kurchan algorithm

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

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



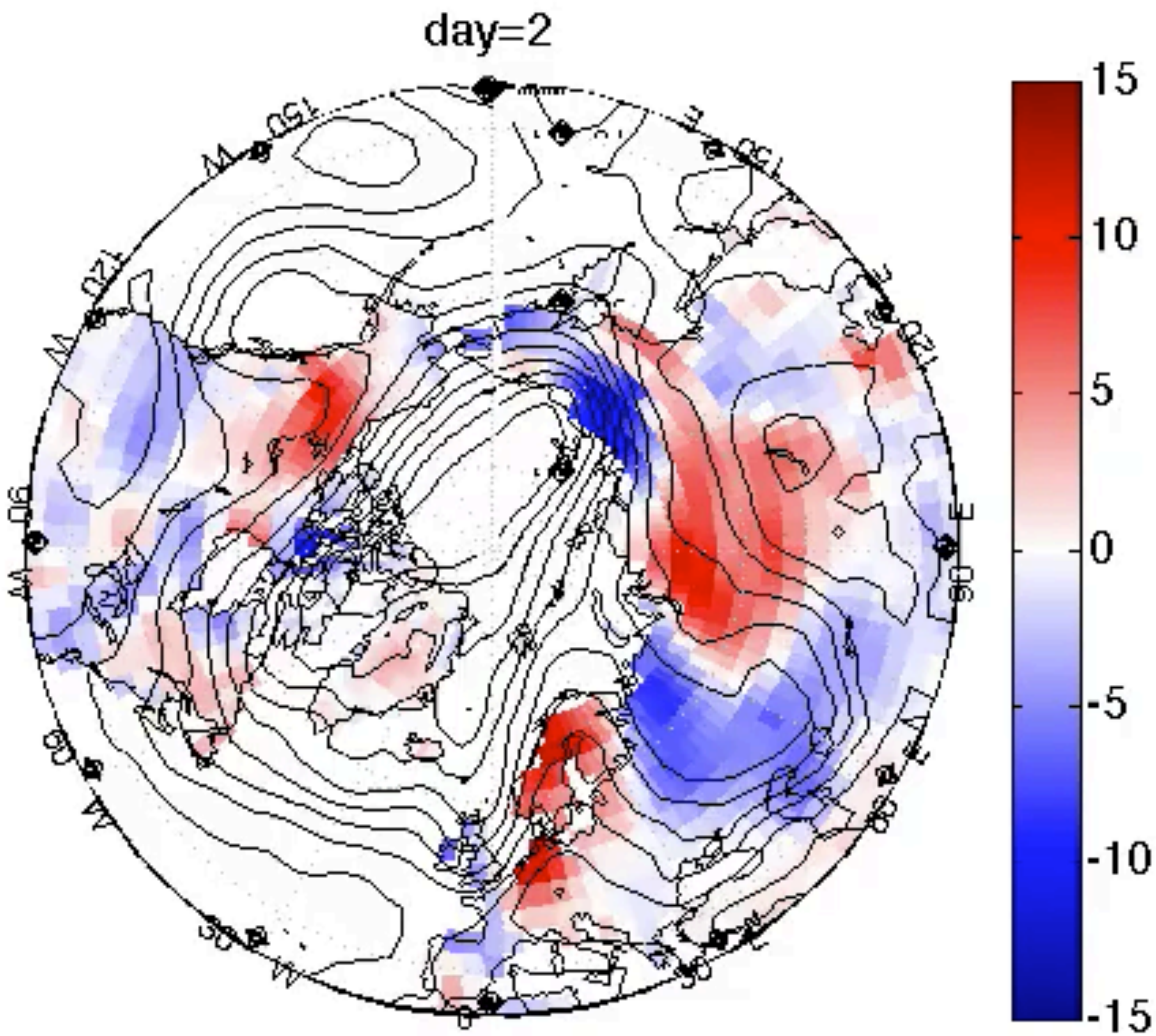
Extremes of 90-day Europe heat waves

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.

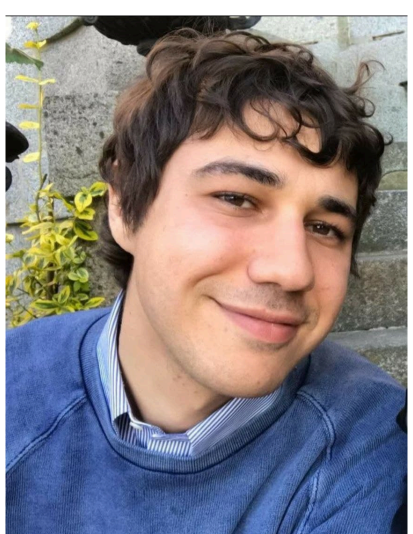
Heat wave dynamics



Plasim heat wave over Scandinavia

500 hPa geopotential height and temperature anomalies

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



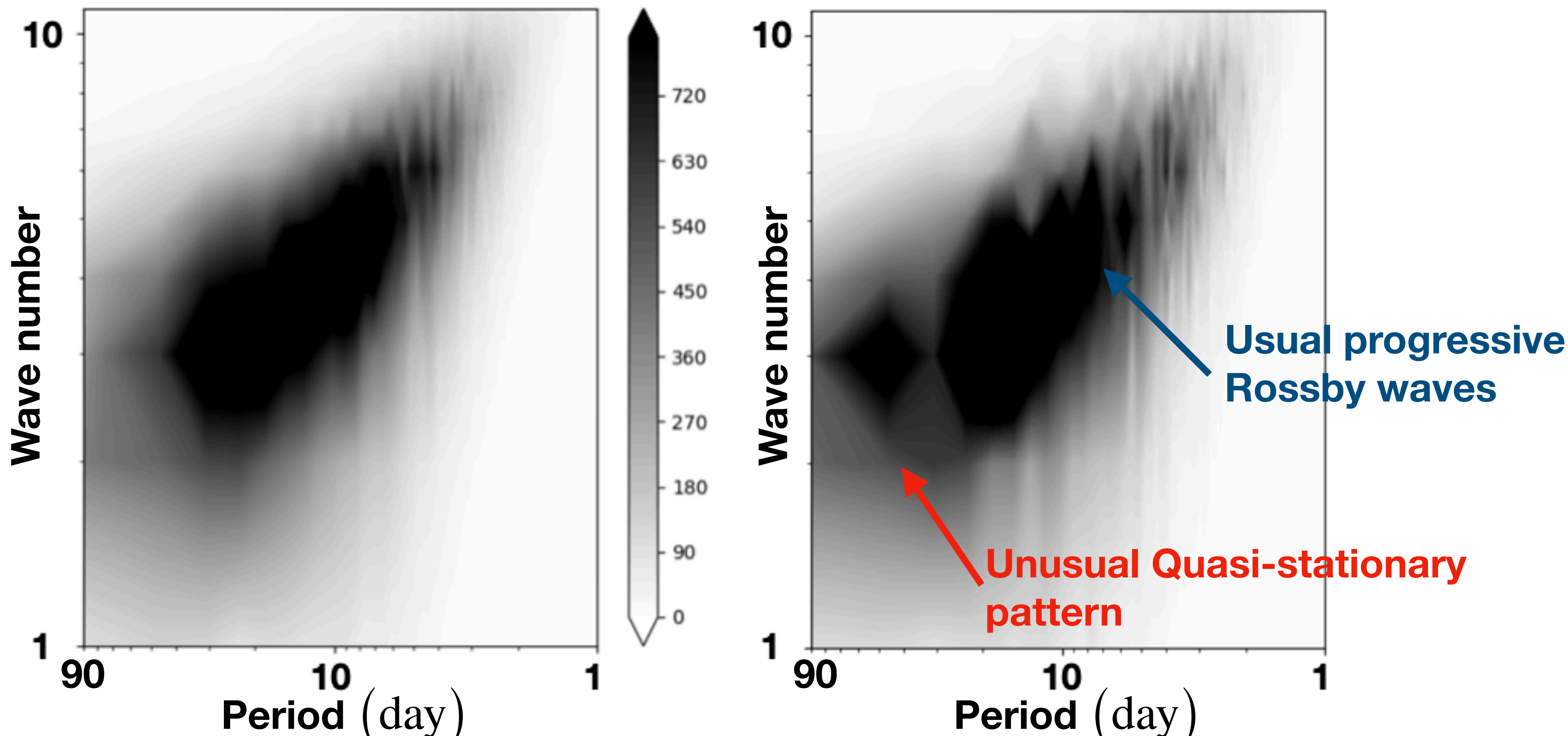
Dario Lucente

George Miloshevich

Francesco Ragone

Heat wave = unusual quasi stationary pattern + progressive Rossby wave

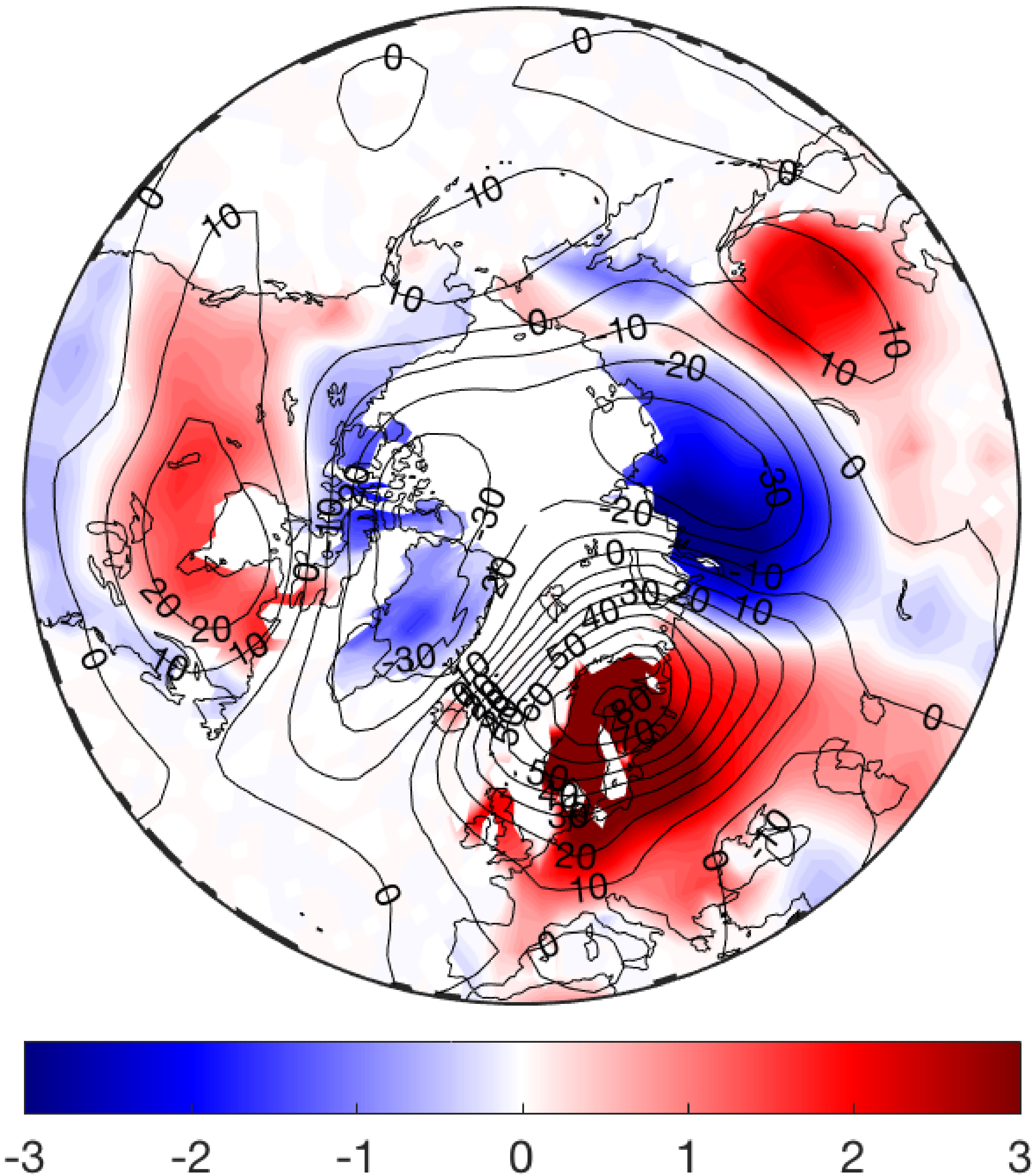
Hayashi spatio-temporal spectrum for eastward waves - CESM model (from the 500 hPa geopotential height over a latitudinal band $55^o - 75^o N$)



CESM Climatology

CESM $T =$ 30-day extreme heatwaves

Extreme teleconnection pattern



500 hPa geopotential height and temperature anomalies

Extreme teleconnection patterns = conditional averages with

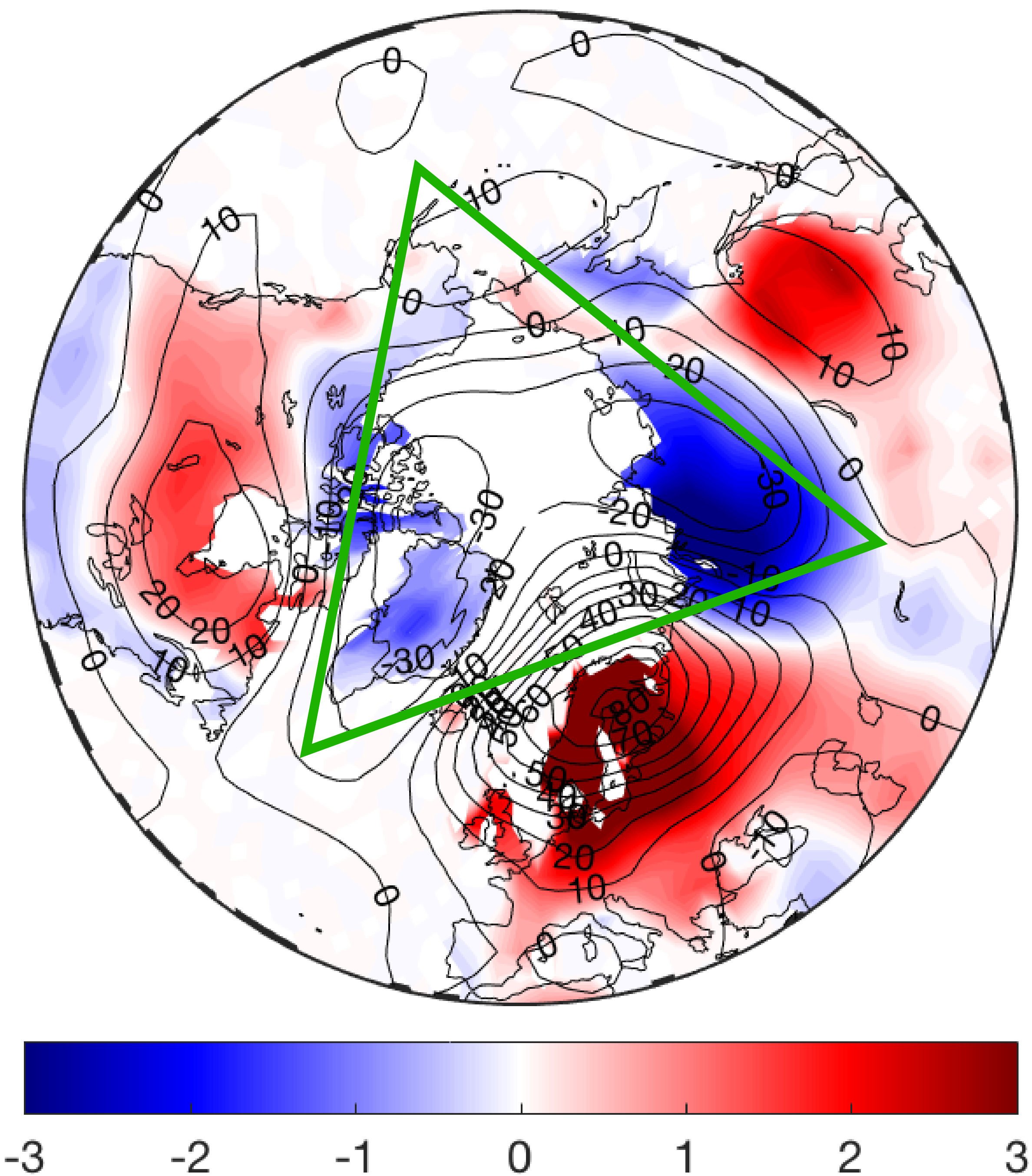
$$\frac{1}{T}\int_0^T \mathrm{d}t \frac{1}{|\mathscr{A}|}\int_{\mathscr{A}} \mathrm{d}\mathbf{r}\ \mathrm{T_S}(\mathbf{r},t) > 2\ \mathrm{K}$$

and $T = 40$ 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

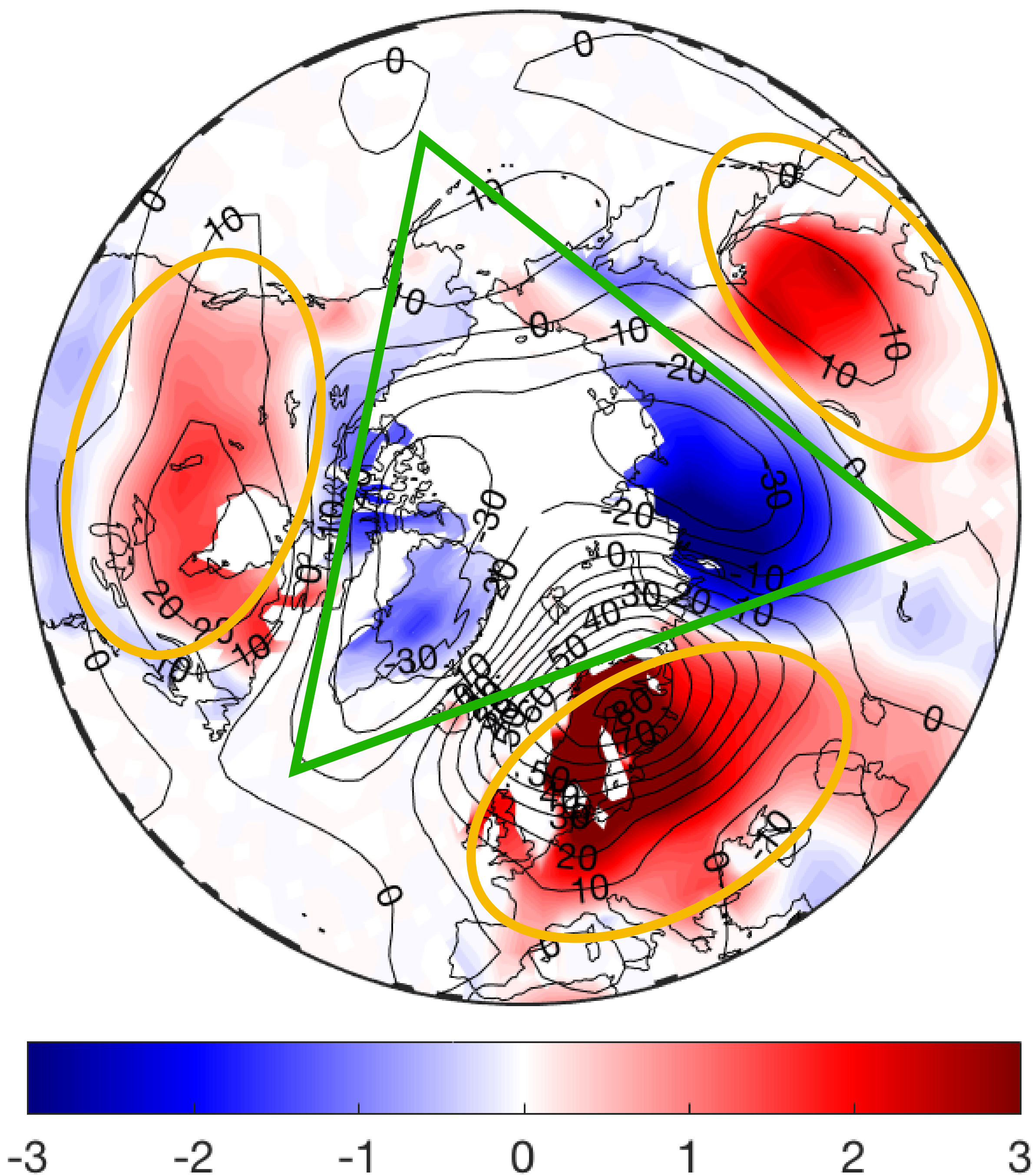
$$\frac{1}{T}\int_0^T \mathrm{d}t \frac{1}{|\mathscr{A}|}\int_{\mathscr{A}} \mathrm{d}\mathbf{r}\ \mathrm{T_S}(\mathbf{r},t) > 2\ \mathrm{K}$$

and $T = 40$ 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

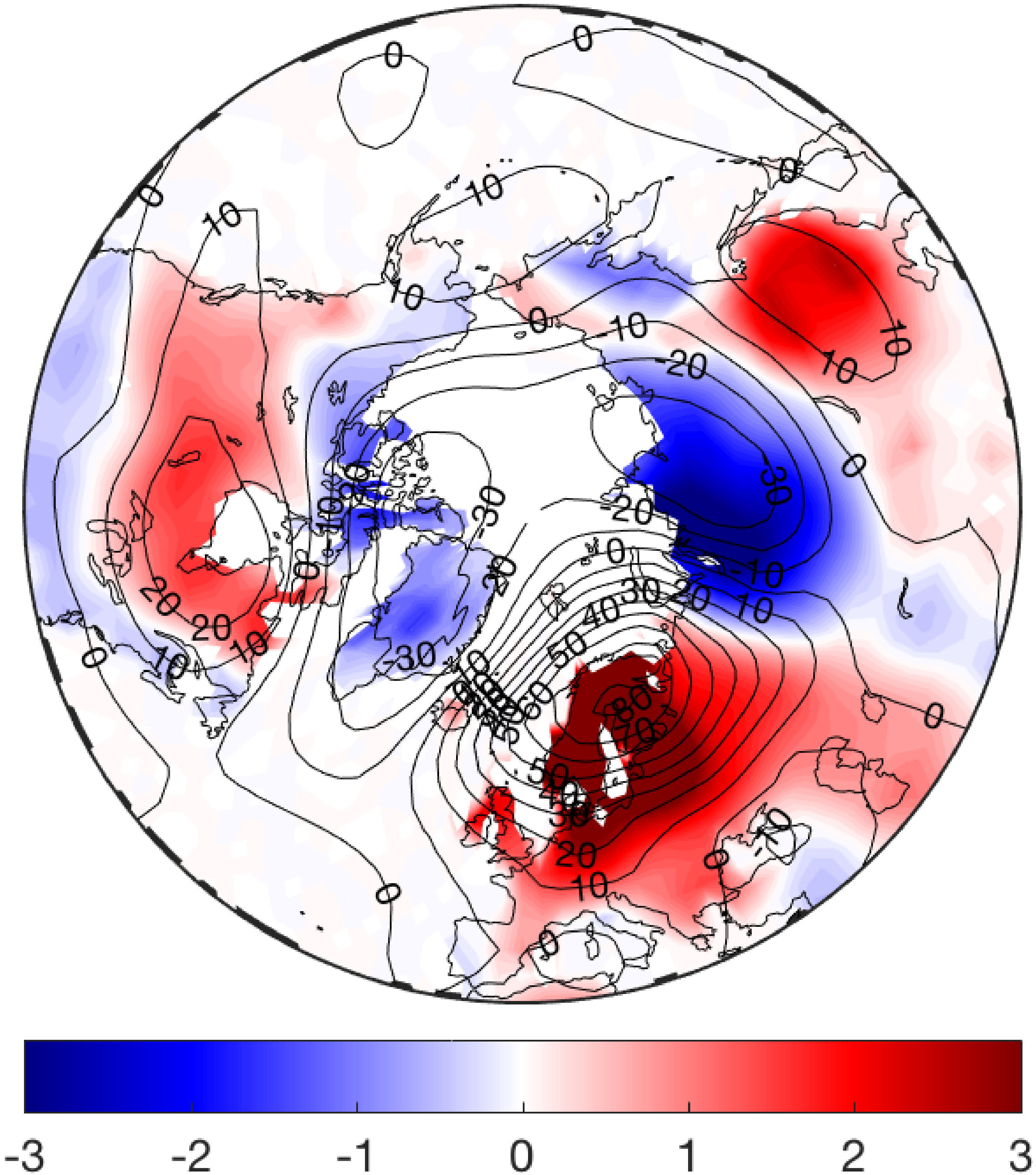
$$\frac{1}{T}\int_0^T \mathrm{d}t \frac{1}{|\mathscr{A}|}\int_{\mathscr{A}} \mathrm{d}\mathbf{r}\ \mathrm{T_S}(\mathbf{r},t) > 2\ \mathrm{K}$$

and $T = 40$ 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

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

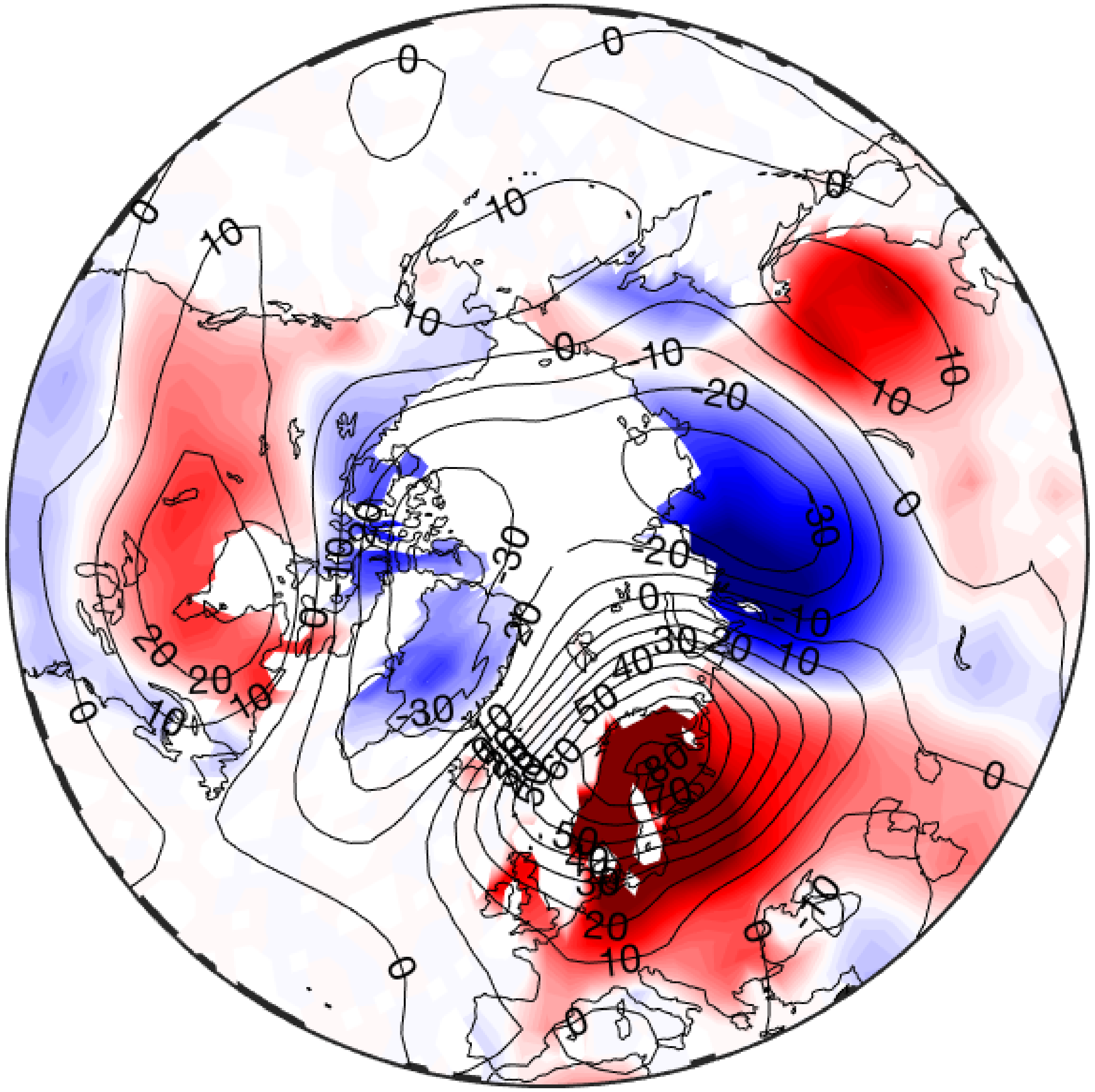
and $T = 40$ days.

Plasim model.
Summer Scandinavian heat waves.

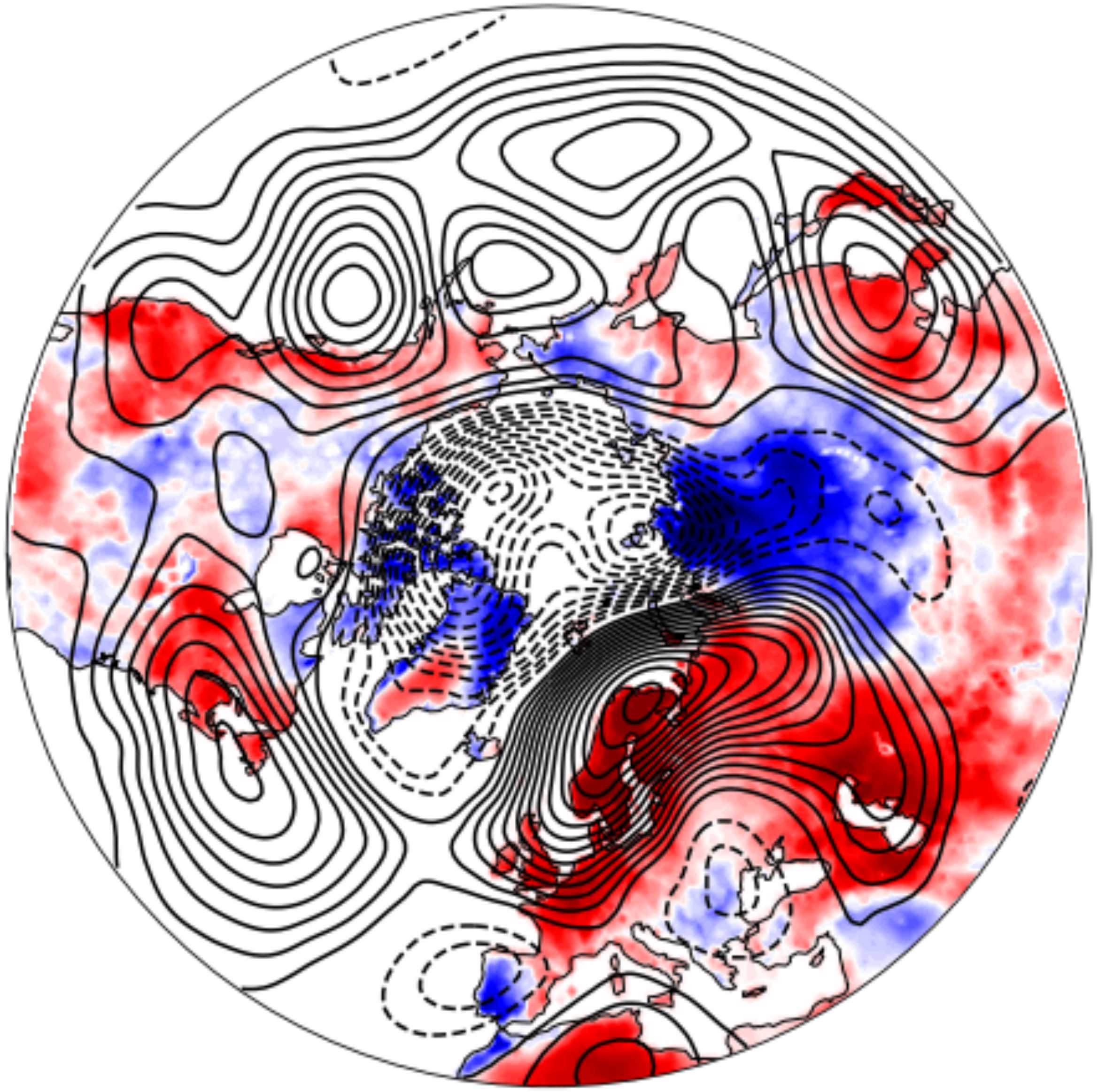
F. Ragone, J. Wouters, and F. Bouchet, PNAS, 2018

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



Published in January 2018
(PLASIM model)



Observed in July 2018 (ERA5)

Climate models correctly predict extreme teleconnection patterns.

Rare event algorithms for climate dynamics
Outline

I) Introduction: rare events do matter - rare event algorithms

II) Rare events algorithms for predicting extreme heat waves

III) Predicting extreme heat waves and committor functions using deep neural networks

IV) Coupling rare event algorithms with machine learning

IV) Predicting extreme heat waves and committor functions using deep neural networks

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

Valerian Jacques-Dumas

George Miloshevich

Machine learning, climate, and weather forecast models

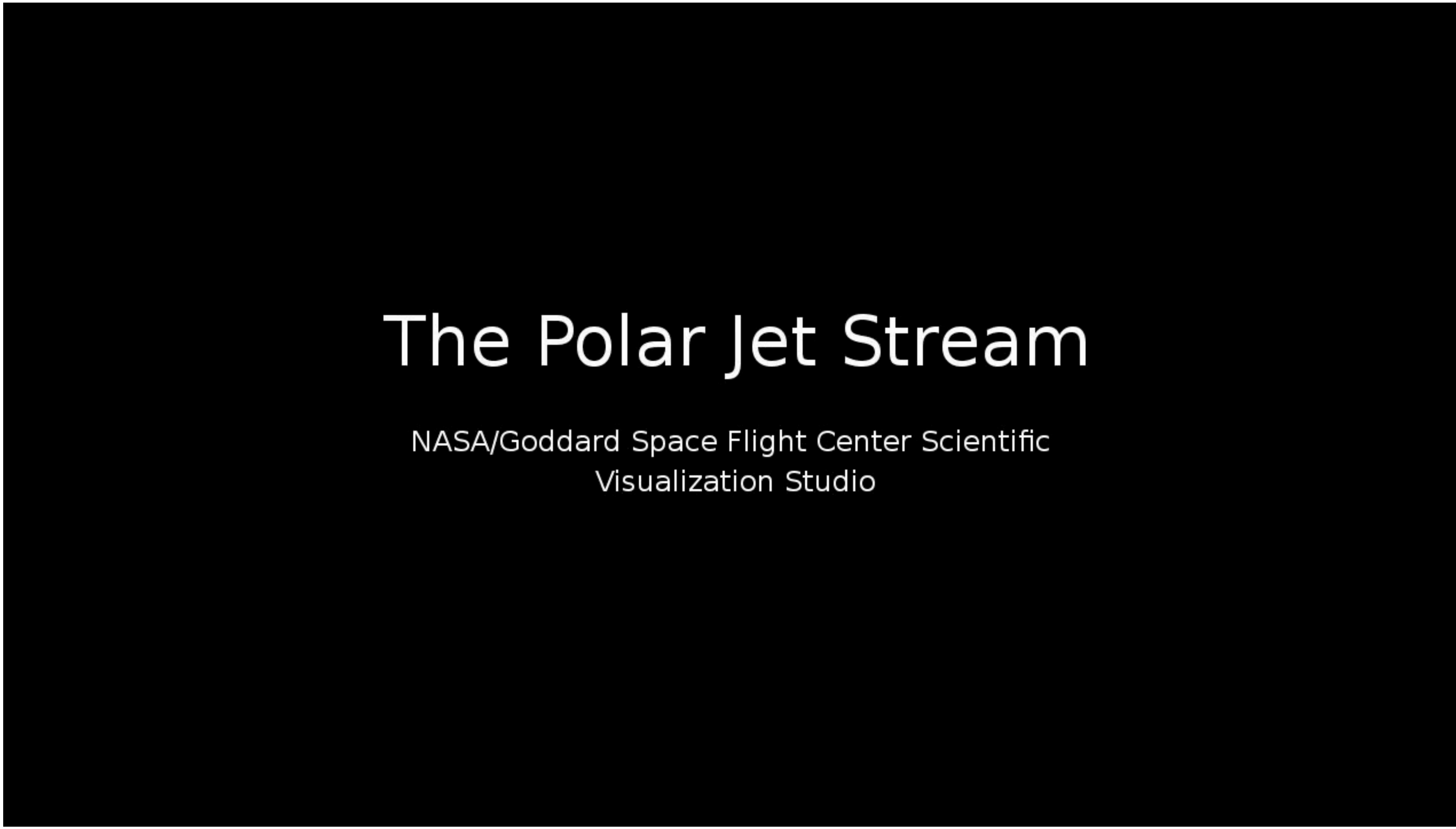
- The Earth (atmosphere, ocean, land, etc.) is the most observed system with an exponentially growing dataset.

- Those observations are coupled to physical models through data assimilation techniques (a very old and very smart machine learning scheme for physically based data integration).

- Machine learning and deep neural networks enter in many different ways for both weather forecast and climate dynamics.

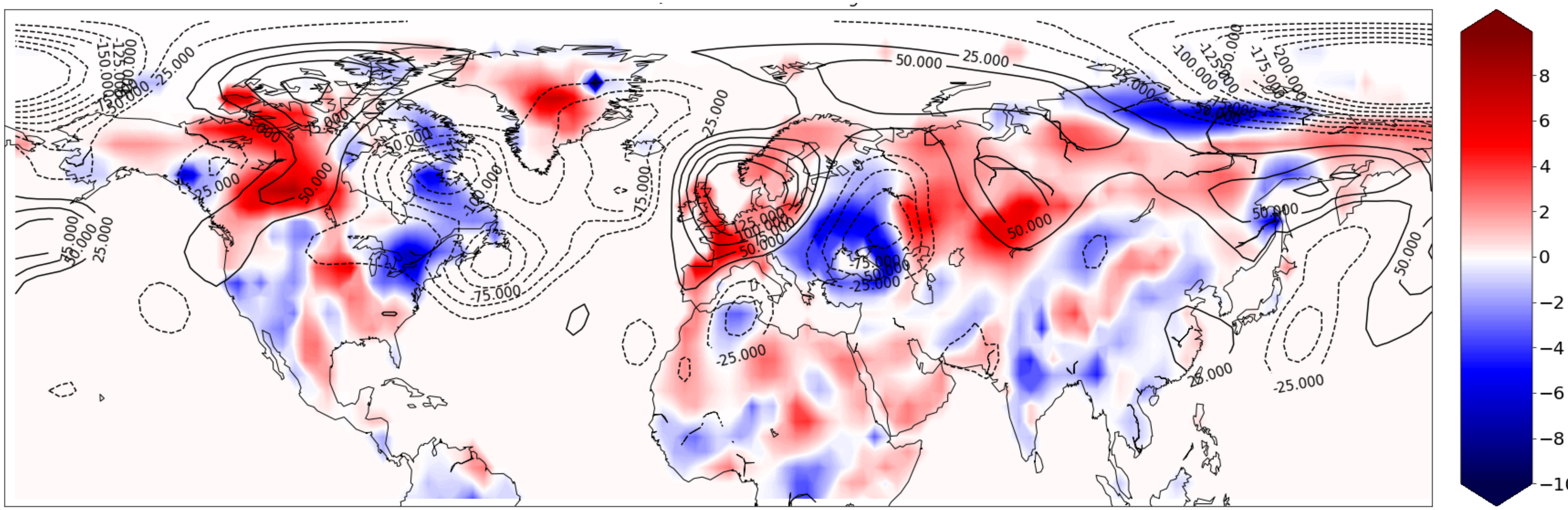
- For many (not all) of these problems, machine learning should be performed in a regime of lack of data. This is key for understanding the challenge for machine learning.

Jet stream dynamics



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

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:**

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

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

- $Z(t) \in \{0,1\}$. 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 T-day heat wave starting τ days from now?

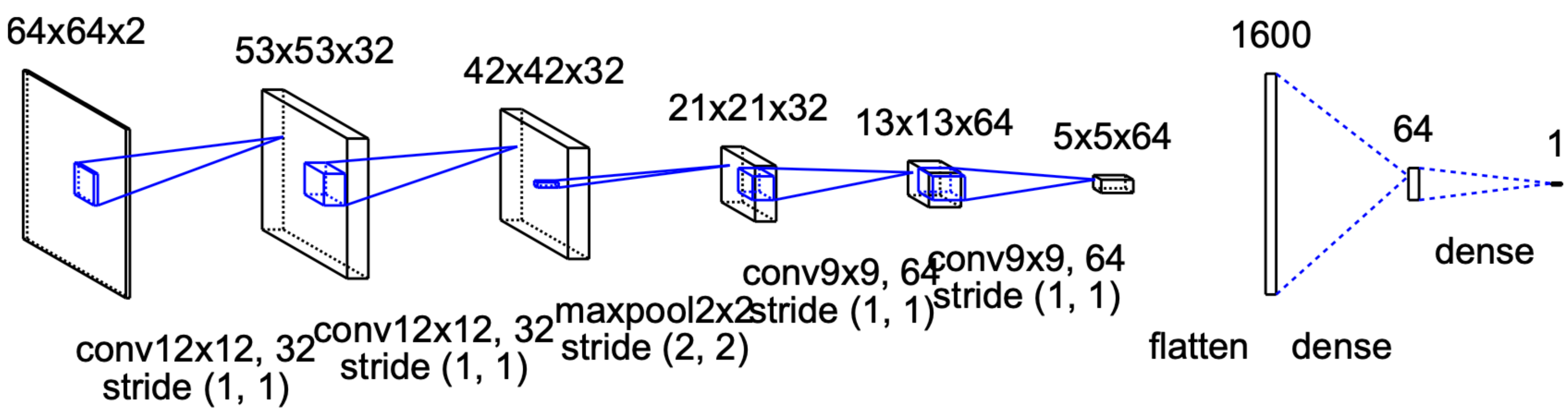
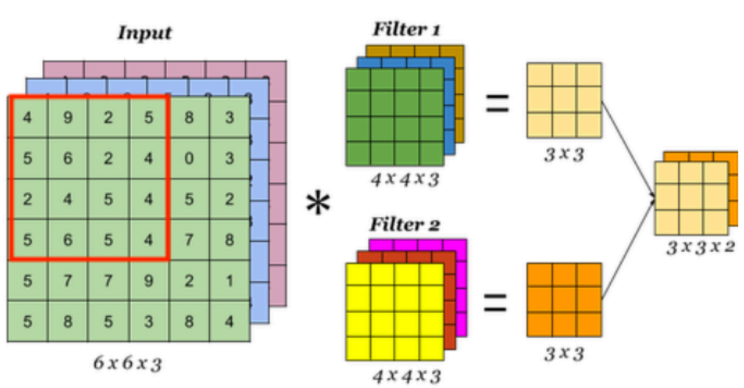
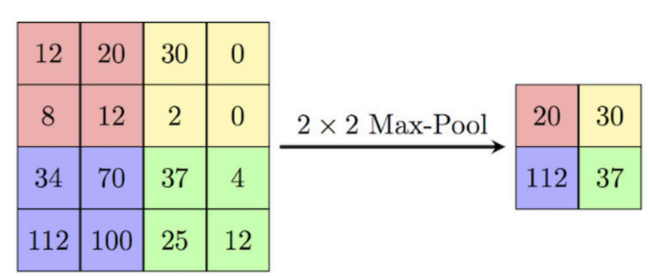


Figure 2: Architecture of the CNN used to forecast extreme heatwaves.



(a) A Convolution layer

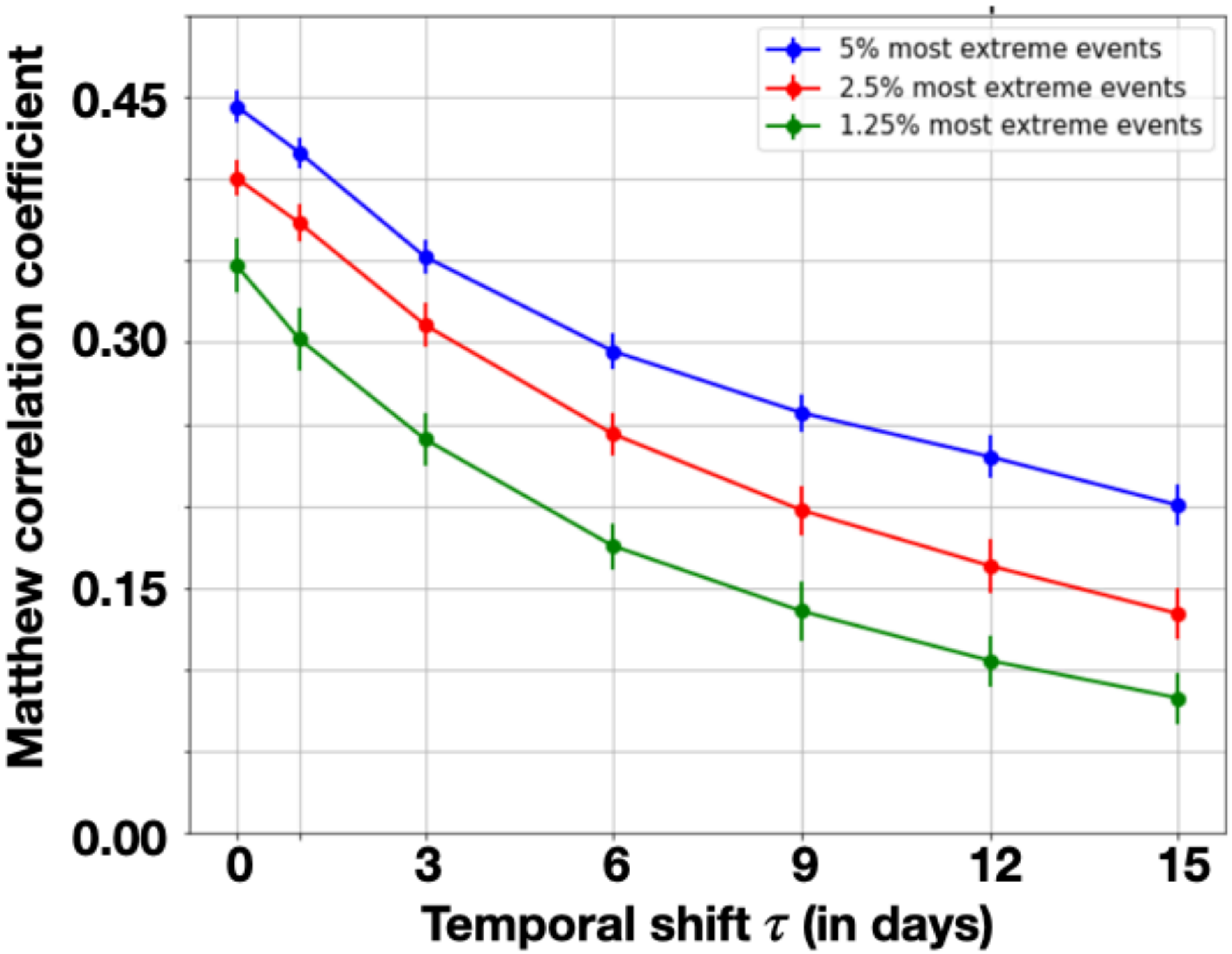


(b) A MaxPool layer

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
$T = 14$

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 $T = 14$-day heatwaves

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

Probabilistic versus deterministic classification

- Recognize cats from dogs with a neural network is an example of a probabilistic prediction for a deterministic classification. We predict the class, in a probabilistic way measuring the level of confidence.

- We use scores like the Mathew correlation coefficient to test the class prediction.

- Predict the occurence of heat waves τ days ahead with a neural network is an example of a probabilistic prediction for a probabilistic classification. We predict the class probability.

- We use scores like the logarithmic or the Brier score, to test the prediction of the class probability.

The Normalized and Positively oriented logarithmic score

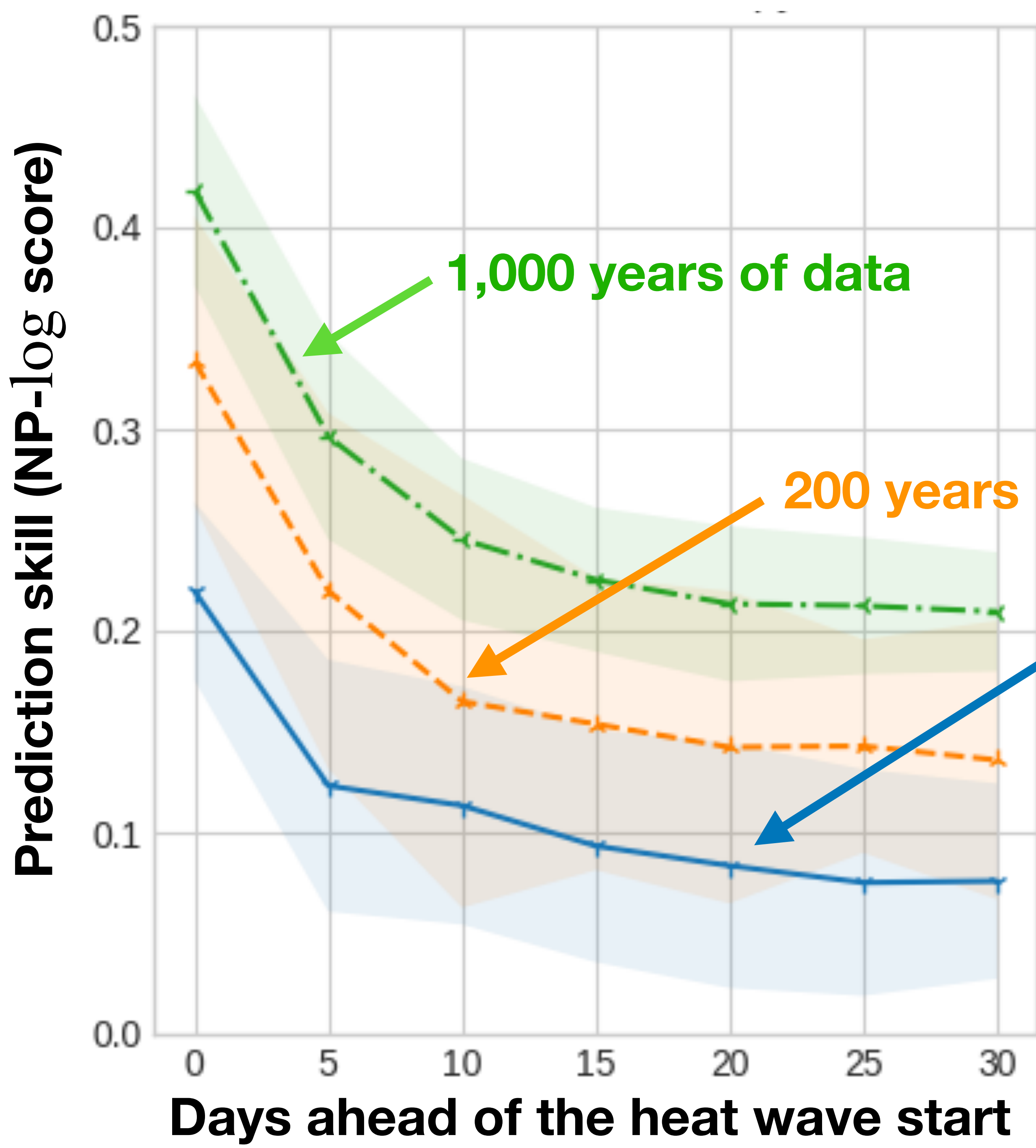
- In order to test the efficiency of the probabilistic prediction of the probabilistic classification, we use the logarithmic score $\mathbb{E}\left\{\log\left[p_{Y_n}\left(X_n\right)\right]\right\}$.

- We define a normalized and positively oriented logarithmic score

$$NP\log = a\mathbb{E}\left\{\log\left[p_{Y_n}\left(X_n\right)\right]\right\} + b,$$

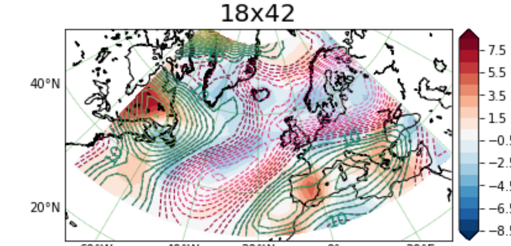
where a and b are such that $NP\log = 0$ for the prediction according to the climatology (prediction using no information on the state X) and $NP\log = 1$ for perfect prediction.

Machine learning for climate applications: a lack of data regime



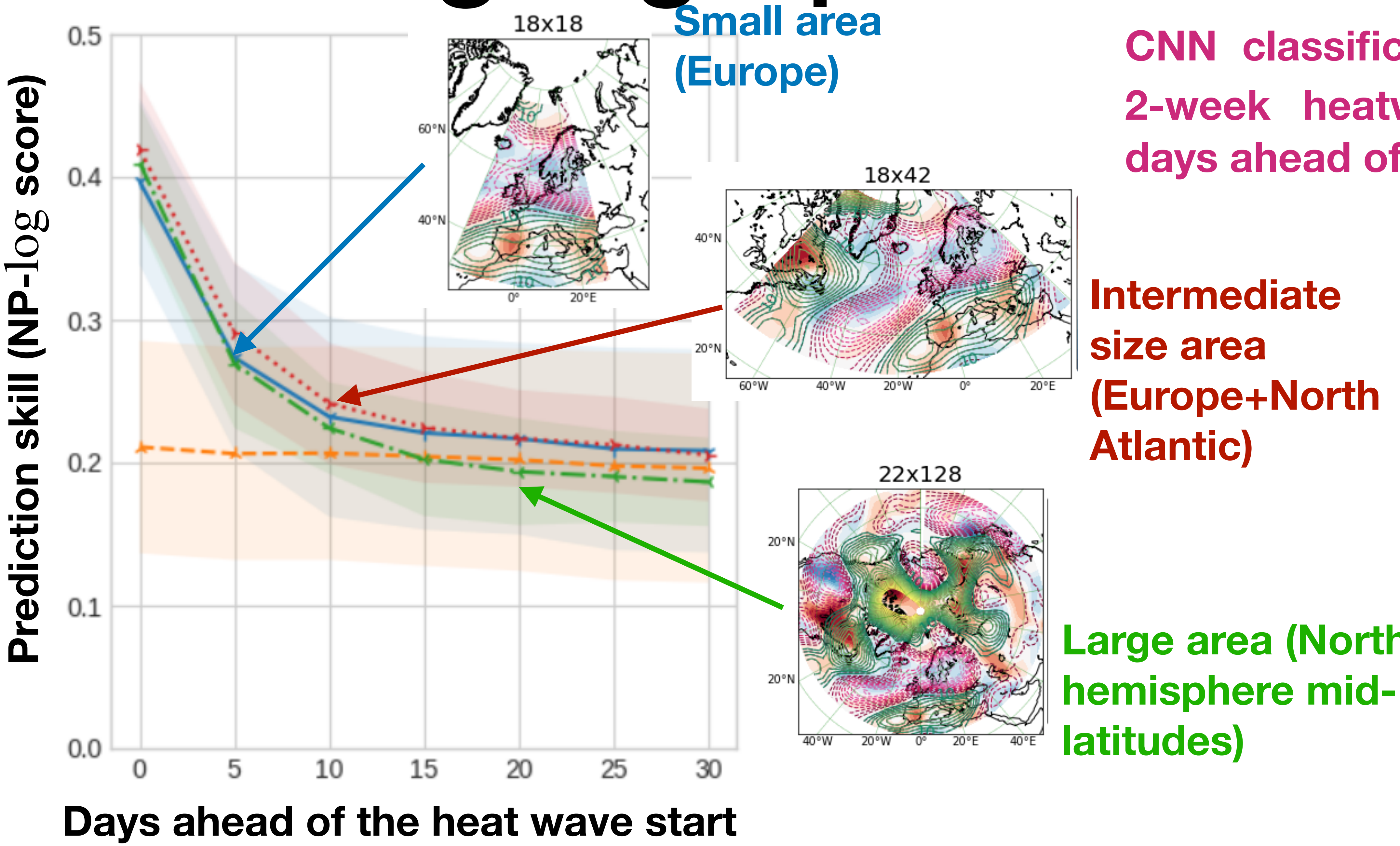
CNN classification of 2-week heatwaves τ days ahead of time

a_5 . The input consists of (t2m,zg500,mrso) with resolution (18,42,3). filtering applied to mrso (no part of the map is set to zero. We were using out-of-the-box CNN



The observation dataset is way too small for good machine learning prediction

Which is the optimal dataset geographical area?



CNN classification of 2-week heatwaves τ days ahead of time

The best performance is obtained for an area of an intermediate size.
This also points to a regime of lack of data for optimal learning.

Conclusions: predicting heat waves with deep neural networks

- Prediction of heat waves is an example of a probabilistic classification problem.
- We use off-the-shelf CNN algorithms, adapted to this situation (probabilistic scores, undersampling, transfer learning).
- Two-week heat waves can be efficiently predicted up to 15 days ahead.
- We are clearly in a regime of lack of data for an optimal prediction.

Rare event algorithms for climate dynamics

Outline

I) Introduction: rare events do matter - rare event algorithm

III) Rare events algorithms for predicting extreme heat waves

III) Predicting extreme heat waves and committor functions using deep neural networks

IV) Coupling rare event algorithms with machine learning

Conclusion: Coupling machine learning with rare event algorithms

- We can learn committor functions from dynamical datasets either using the definition, or first learning an approximate Markov dynamics.

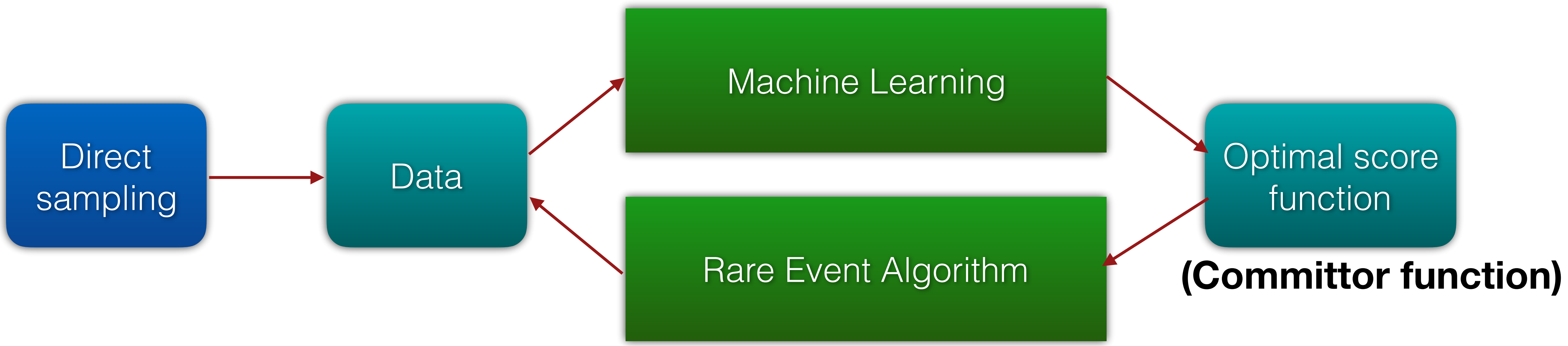
- The analogue Markov chain does not require an impossible discretization of the phase space, and can use any kind of dynamical data, including short trajectories.

- Using learned committor functions is much more efficient than using user-defined score functions with the AMS rare event algorithm.

- The range of applicability of this approach, in terms of system dimension and complexity, is a key question for the future.

D. Lucente, J. Rolland, C. Herbert and F. Bouchet, ArXiv, submitted to. JSTAT

Coupling rare event algorithms with machine learning of committor functions



Work in progress for climate models!

Conclusions: Studying rare and extreme climate events with rare event algorithms and machine learning

- We can use **rare event algorithms** to gather an amazing statistics for extreme heat waves with PLASIM (PNAS, 2018), and CESM (GRL, 2021) models.

- The dynamical mechanism is the birth of **quasi-stationary non zonal global patterns**, which are much affected by the orography and oceans (PNAS, 2018, GRL 2021).

- **With CNN machine learning, we predict the probability of extreme heat waves** up to 15 days ahead of time (Sub. to Frontiers in Climate, 2021).

- **The coupling of learned committor functions with rare-event algorithms is extremely efficient for toy models** (Sub. to JSTAT, 2021). Work in progress for climate models.

Studying rare events is extremely fascinating!

Opened post-doc positions