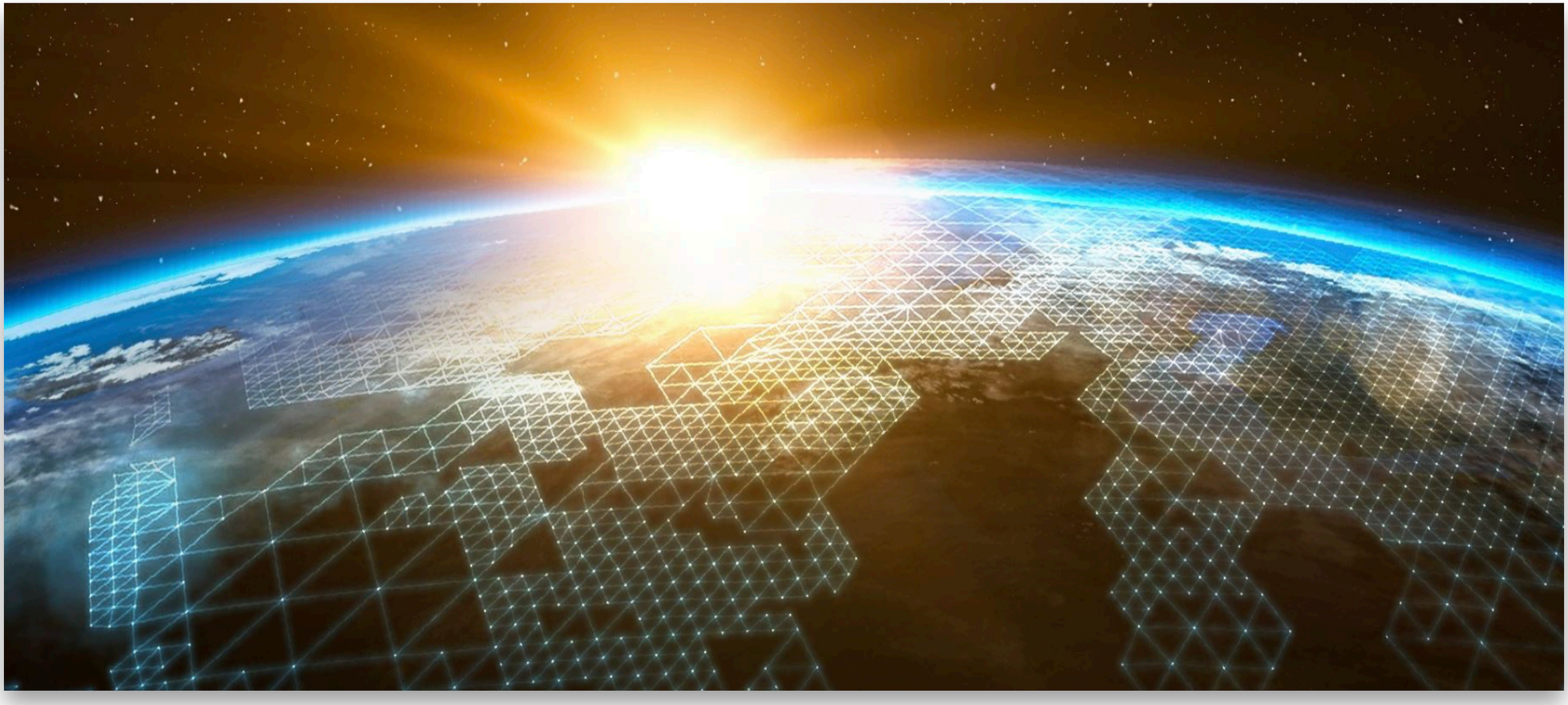


Webinar Barcelona Computing Center

History Matching for the tuning of climate models

Redouane Lguensat

Postdoc @ LOCEAN / LSCE / Sorbonne University

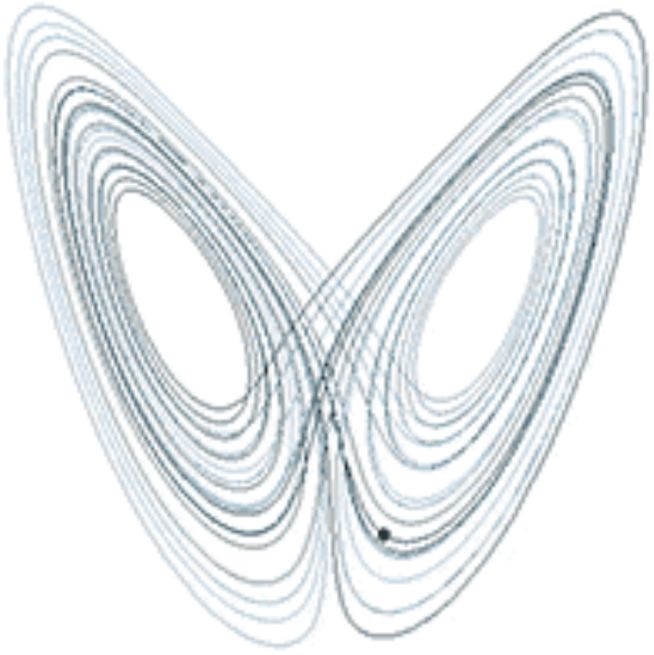
Funded by HRMES-MOPGA

The butterfly effect

A popular hypothetical situation that illustrates how small initial differences may lead to large unforeseen consequences over time.

« In 1961, Lorenz was running a numerical computer model to redo a weather prediction from the middle of the previous run as a shortcut. He entered the initial condition **0.506** from the printout instead of entering the full precision **0.506127** value. The result was a completely different weather scenario » - Gleick, James (1987)

Edward Lorenz (1917-2008)



Lorenz attractor

Edward Lorenz the « data scientist »

Empirical Orthogonal Functions
and Statistical Weather Prediction

by
EDWARD N. LORENZ

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
DEPARTMENT OF METEOROLOGY
Cambridge, Massachusetts

DECEMBER 1956

Scientific Report No. 1
STATISTICAL FORECASTING PROJECT

EDWARD N. LORENZ

EOF

THE RESEARCH REPORTED IN THIS DOCUMENT HAS BEEN SPONSORED BY THE GEOPHYSICS RESEARCH DIRECTORATE OF THE AIR FORCE CAMBRIDGE RESEARCH CENTER, AIR RESEARCH AND DEVELOPMENT COMMAND, UNDER CONTRACT NO. AF19(604)1566.

636 JOURNAL OF THE ATMOSPHERIC SCIENCES VOLUME 26

Atmospheric Predictability as Revealed by Naturally Occurring Analogues

EDWARD N. LORENZ

Dept. of Meteorology, Massachusetts Institute of Technology, Cambridge, Mass.[1]
(Manuscript received 2 April 1969)

ABSTRACT

Two states of the atmosphere which are observed to resemble one another are termed *analogues*. Either state of a pair of analogues may be regarded as equal to the other state plus a small superposed "error." From the behavior of the atmosphere following each state, the growth rate of the error may be determined.

Five years of twice-daily height values of the 200-, 500-, and 850-mb surfaces at a grid of 1003 points over the Northern Hemisphere are procured. A weighted root-mean-square height difference is used as a measure of the difference between two states, or the error. For each pair of states occurring within one month of the same time of year, but in different years, the error is computed.

There are numerous mediocre analogues but no truly good ones. The smallest errors have an average doubling time of about 8 days. Larger errors grow less rapidly. Extrapolation with the aid of a quadratic hypothesis indicates that truly small errors would double in about 2.5 days. These rates may be compared with a 5-day doubling time previously deduced from dynamical considerations.

The possibility that the computed growth rate is spurious, and results only from having superposed the smaller errors on those particular states where errors grow most rapidly, is considered and rejected. The likelihood of encountering any truly good analogues by processing all existing upper-level data appears to be small.

Analogs

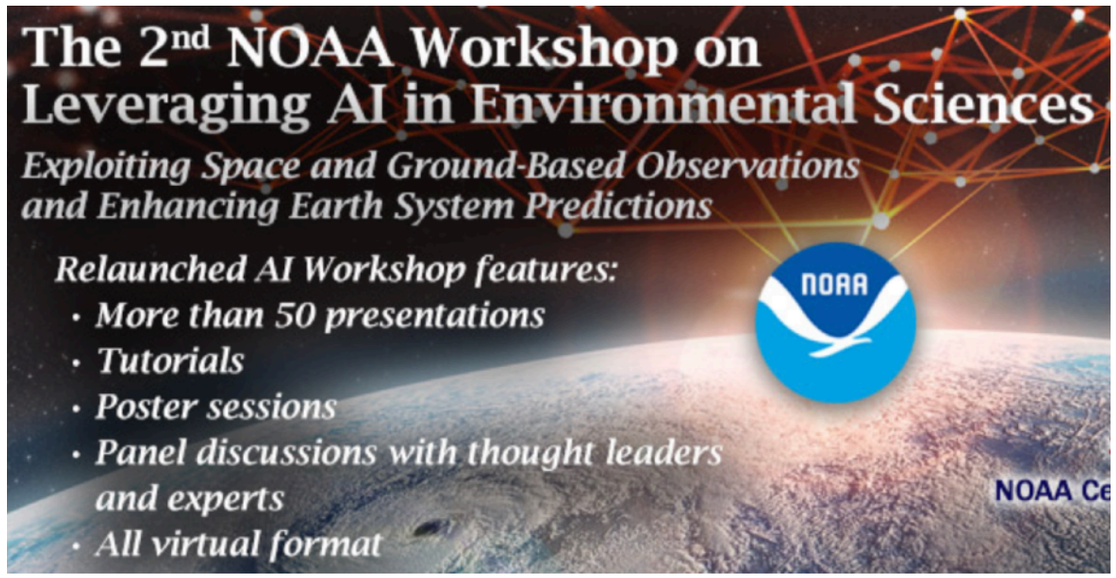
- Edward Lorenz was among the first researchers using Machine Learning for weather time series forecasting

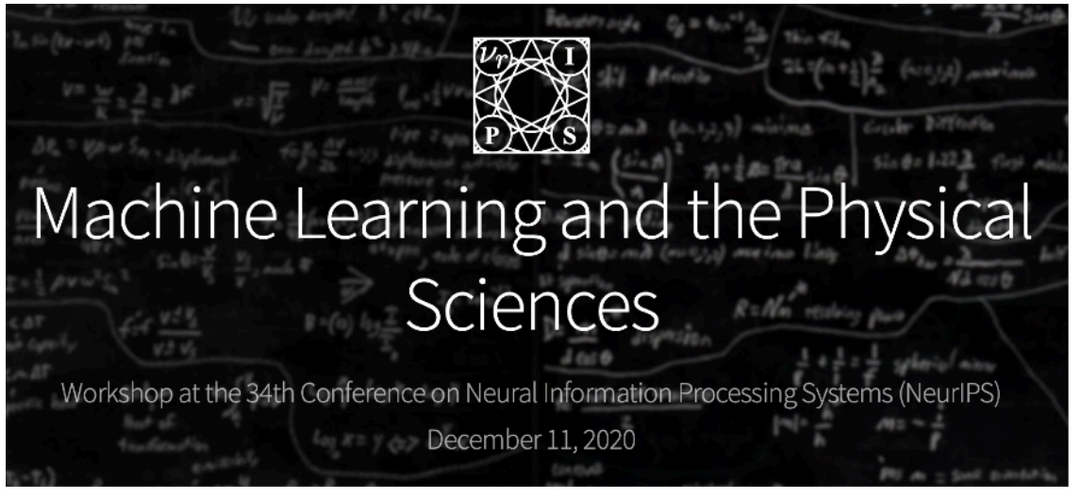
Note: **EOF=Principal Component Analysis** & **Analogs=Nearest Neighbors**

Was Lorenz not in contact with the ML community at that time?

The rise of AI4ES

A vivid and exciting field





+ special sessions in EGU, AGU, etc.

History Matching for the tuning of climate models

Background

Interdisciplinary researcher

Research interests:

* Revive ML methods for earth science, ex: Analog methods
* Use recent ML techniques for earth science problems, ex: Deep Learning
* Develop ML codes easy to adopt by earth scientists

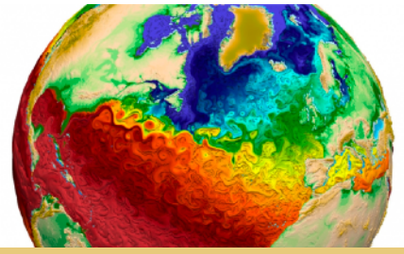
Education:

* Statistical Signal/Image Processing engineer
* PhD IMT Atlantique (French « Grande Ecole ») 2017: Machine Learning from Ocean Remote Sensing

Current position:

* MOPGA postdoc under the project **HRMES** (High Resolution Modeling of the Earth System), P.I: *V. Balaji*

Climate model tuning

Parameters →  → $f(x)$

Climate model simulation code

Metrics, ex:

* Meridional Heat Transport
* AMOC
* Temperature profiles
* Etc…

Compare with

z

Observations

Tuning: the process of estimating uncertain parameters in order to reduce the mismatch between specific observations and model results

RESEARCH ARTICLE | 31 MARCH 2017

The Art and Science of Climate Model Tuning

Frédéric Hourdin; Thorsten Mauritsen; Andrew Gettelman; Jean-Christophe Golaz; Venkatramani Balaji; Qingyun Duan; Doris Folini; Duoying Ji; Daniel Klocke; Yun Qian
Florian Rauser; Catherine Rio; Lorenzo Tomassini; Masahiro Watanabe; Daniel Williamson

Bull. Amer. Meteor. Soc. (2017) **98** (3): 589–602.

https://doi.org/10.1175/BAMS-D-15-00135.1 Article history

« Why such a lack of transparency? This may be because tuning is often seen as an unavoidable but dirty part of climate modeling, more engineering than science, an act of tinkering that does not merit recording in the scientific literature. »

→ It is important that modeling groups communicate their tuning strategy

Ex: when comparing models on a given metric, it is essential to know whether some models used this metric as a tuning target.

Tuning: how to?

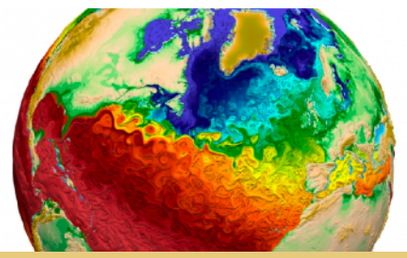
Two main strategies to perform tuning in an objective and reproducible way (better than naive trial-and-error):

- Optimization of a **cost function** measuring the distance of model simulations to a observations (gradient/gradient-free based techniques)
- Uncertainty Quantification (UQ): Bayesian approach that provides **uncertainty** for the parameters using a statistical model relating the climate model to observations

We are interested here in a UQ technique: **History Matching**

History Matching

$\boldsymbol{x}$
Parameters



Climate model simulation code

$\boldsymbol{f}(\boldsymbol{x})$
Metrics, ex:

* Meridional Heat Transport
* AMOC
* Temperature profiles
* Etc…

Compare with

$\boldsymbol{z}$
Observations

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}} \|z - \boldsymbol{f}(\boldsymbol{x})\|_f$$

But ! tuning to a handful of metrics may risk achieving improved performance in those metrics at the expense of unphysical behavior in metrics or processes that were not used in tuning

——> **Overtuning**

« *Overtuning is a real concern and the raison d'être for Bayesian UQ methods* » Hourdin et al. 2017

<u>History Matching:</u>

* Closely related to Approximate Bayesian Computation (ABC): used in particle physics, molecular dynamics, population genetics, neuroscience, epidemiology, ecology, astrophysics and recently climate science
* Has always benefited from ML advances

Instead of looking for THE best set of parameters that solves an optimization problem. History Matching uses observed data to rule-out any parameter settings which are ``**implausible**''.

History Matching and ML

Ideally we would individually check every possible parameter setting for the input: **Impossible** (climate models are expensive to run)

→ Need for **space-filling designs** to cover the space of parameter search

→ Need for replacing the expensive simulator with a rapid and cheap **emulator**

We suppose that we have the results of some simulation runs

Set of parameters → Simulation code → Metrics

Supervised learning

Set of parameters → Emulator → Estimated metrics

→ **Uncertainty**

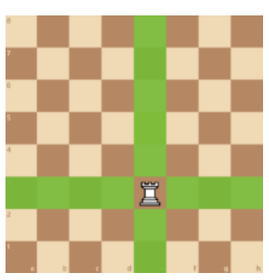
History Matching pipeline

Initial guess of parameter space

$$I(\boldsymbol{\theta}) = \frac{|E[f(\boldsymbol{\theta})] - z|}{\sqrt{\mathrm{Var}[f(\boldsymbol{\theta})]}}$$

Space-filling design → Calculate metrics → Training an Emulator → Calculate Implausibility

Ex: using Latin Hypercube Sampling



Ex: using Gaussian Process Regression

Rule out if I >3

(Pukelsheim's 3-sigma rule)

Set of plausible parameters

HM is an iterative process: done in *waves*

HM for climate modeling

Tuning atmospheric models:

HM was used to tune **atmospheric models**, ex: LMDZ (Hourdin et al. 2020, Couvreux et al. 2020):

* Using single-column models (SCMs) they afford to run several simulations with different set of parameters
* Short timescales

JAMES | Journal of Advances in Modeling Earth Systems

Research Article | Open Access

Process-based climate model development harnessing machine learning: I. a calibration tool for parameterization improvement

Fleur Couvreux, Frédéric Hourdin, Daniel Williamson, Romain Roehrig ... See all authors

First published: 07 December 2020 | https://doi.org/10.1029/2020MS002217

JAMES | Journal of Advances in Modeling Earth Systems

Research Article | Open Access

Process-based climate model development harnessing machine learning: II. model calibration from single column to global

Frédéric Hourdin, Daniel Williamson, Catherine Rio, Fleur Couvreux ... See all authors

First published: 04 December 2020 | https://doi.org/10.1029/2020MS002225

Tuning ocean models:

HM was used to tune **ocean models**, ex: NEMO ORCA 2° (Williamsson et al. 2017):

* Using an available ensemble of 400 NEMO simulations ran for 150 years.
* Long timescales

Geosci. Model Dev., 10, 1789–1816, 2017
www.geosci-model-dev.net/10/1789/2017/
doi:10.5194/gmd-10-1789-2017

Geoscientific Model Development | EGU | Open Access

Tuning without over-tuning: parametric uncertainty quantification for the NEMO ocean model

Daniel B. Williamson[1], Adam T. Blaker[2], and Bablu Sinha[2]

[1]College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, UK
[2]National Oceanography Centre, Southampton, SO14 3ZH, UK

Correspondence to: Daniel B. Williamson (d.williamson@exeter.ac.uk)

Received: 20 July 2016 – Discussion started: 30 August 2016
Revised: 24 November 2016 – Accepted: 30 January 2017 – Published: 27 April 2017

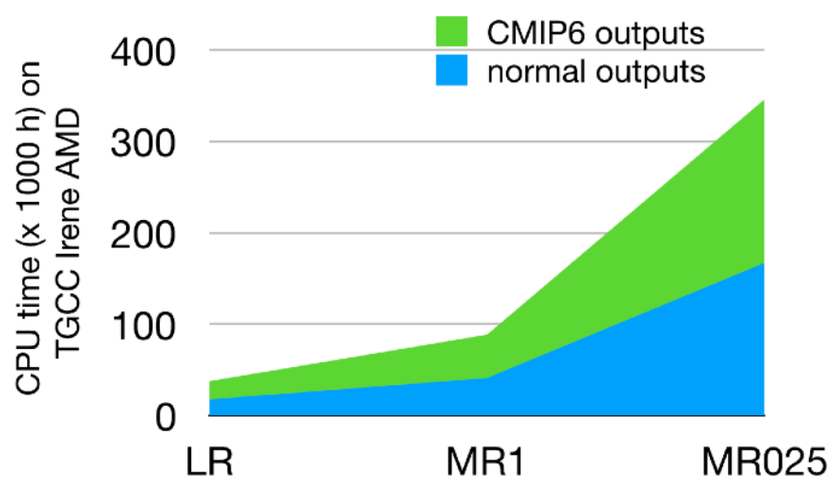
What happens when coupling independently tuned components?

Coupling

Example: QUEST project (PIs : J. Deshayes, F. Hourdin, J. Mignot, supported by PRACE)

History Matching to LMDZ in 144x142 configuration (atmosphere only, with observed SST and sea-ice)
→ 5 new tunings to 250yr piCtrl coupled LR configurations
→ excessive cold biases in T2m and sea-ice cover in Northern Hemisphere
→ additional tuning of sea-ice and ocean parameters !

* Applying HM directly on coupled models is still not explored



However, since coupled models are expensive we opted to use a toy model as a testbed to investigate HM strengths and challenges

History Matching for Lorenz96

Joint work with J. Deshayes (LOCEAN-IPSL)
and V. Balaji (Princeton Univ./LSCE-IPSL)

Lorenz 96 as a test model

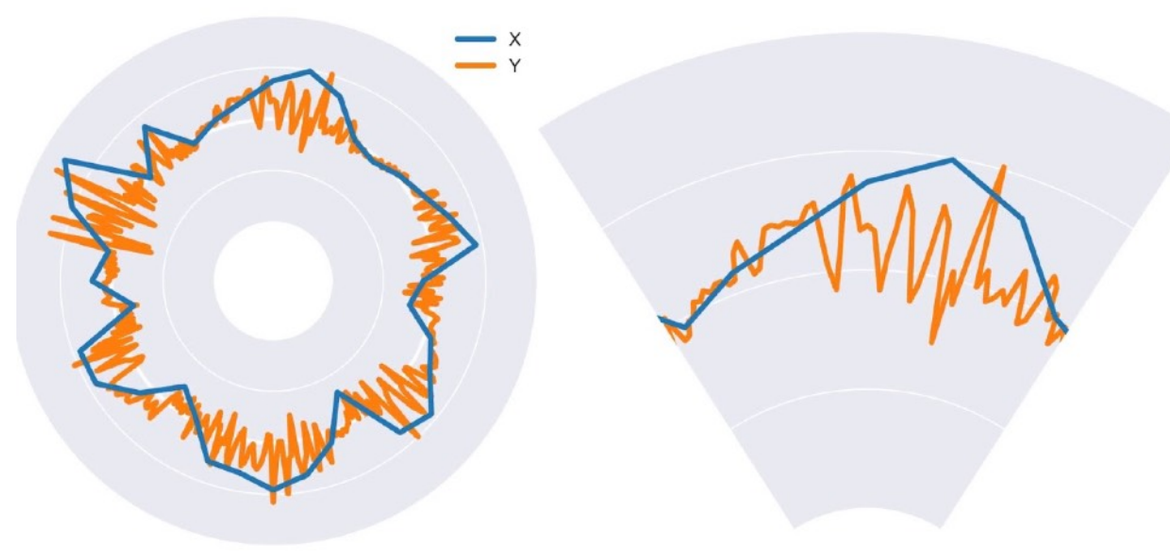
* Periodic system of K (k=1,…,K) ODEs
* **Two-level version:** add periodic variable **Y** with its own set of ODEs.
* The **X** and **Y** ODEs are linked through coupling terms. Each **X** has J **Y** variables associated with it.

$$\frac{dX_k}{dt} = \underbrace{-X_{k-1}(X_{k-2} - X_{k+1})}_{\text{Advection}} \underbrace{-X_k}_{\text{Diffusion}} \underbrace{+F}_{\text{Forcing}} \underbrace{-hc\bar{Y}_k}_{\text{Coupling}}$$

$$\frac{1}{c}\frac{dY_{j,k}}{dt} = \underbrace{-bY_{j+1,k}(Y_{j+2,k} - Y_{j-1,k})}_{\text{Advection}} \underbrace{-Y_{j,k}}_{\text{Diffusion}} \underbrace{+\frac{h}{J}X_k}_{\text{Coupling}}$$

Analogy with coupled ocean-atmosphere models:

Atmosphere	Fast dynamics	Y
Ocean	Slow dynamics	X



Credits: Stephan Rasp

Experiment: HM for the tuning of parameters (F, h, c, b)

HM on the L96

* **Metrics**: long-term time means to mimic climatological quantities
* **Ground Truth:** perfect setting K=36 X variables each coupled with J=10 Y variable. F=10, h=1, c=10, b=10, chaotic behavior.
* **HM code:** Python code + Parallel computation + GPR models can be trained on **GPU**

$$f(X,Y) = \begin{pmatrix} X \\ \bar{Y} \\ X^2 \\ X\bar{Y} \\ \bar{Y}^2 \end{pmatrix}$$

Justified by energy conservation constraints, check Schneider et al. 2017 for details (ESM 2.0 paper)

180-dimensional vector

Initial guess of parameter space

Params	Prior	True
F	[-20,20]	10
h	[-2,2]	1
c	[0,20]	10
b	[-20,20]	10

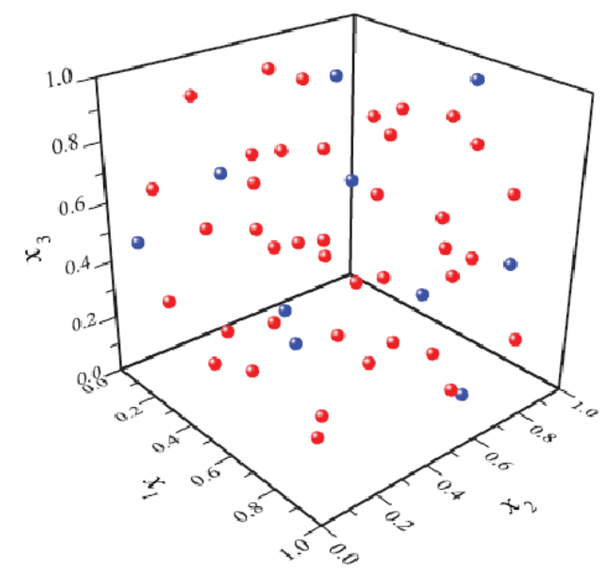
HM on the L96

Initial guess of parameter space

Params	Prior	True
F	[-20,20]	10
h	[-2,2]	1
c	[0,20]	10
b	[-20,20]	10

40 samples from a LHS

Space filling design

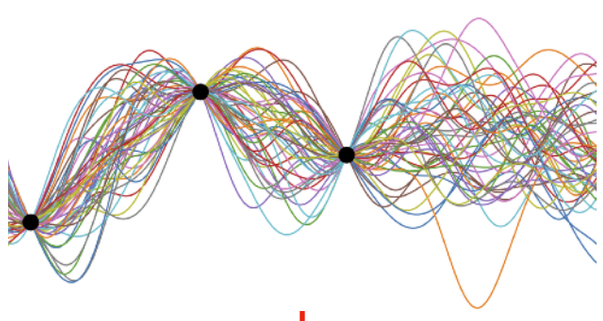


Run the L96 model

Build a training database for the emulator:
x_train.size=(40,4)
y_train.size=(40,180)

Here, one GP Per output

Train the emulator then use it for inference on a large number of samples

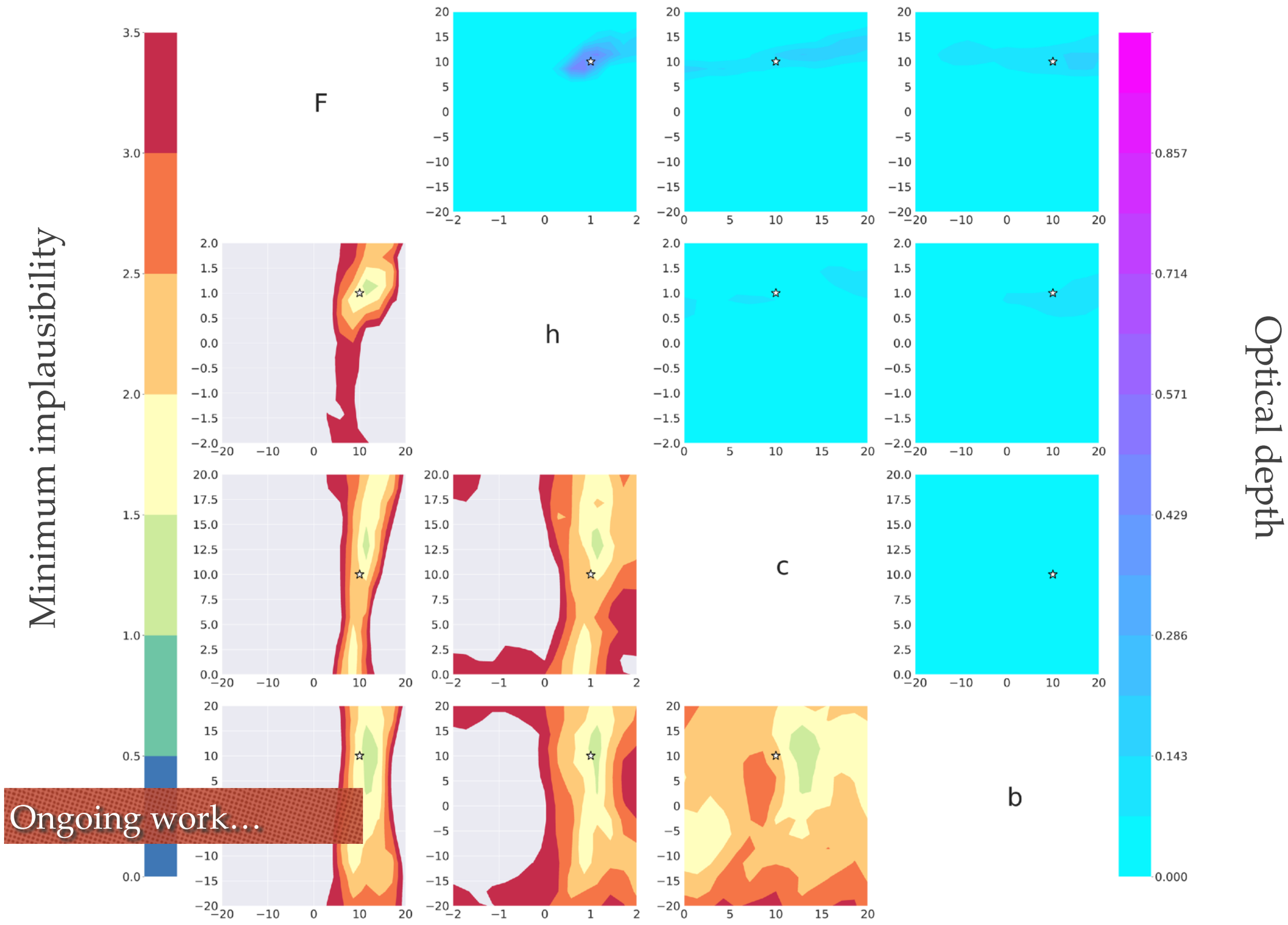


Calculate Implausibility

$$I(\boldsymbol{\theta}) = \frac{|E[f(\boldsymbol{\theta})] - z|}{\sqrt{\mathrm{Var}[f(\boldsymbol{\theta})]}}$$

Wave 1: NROY

$$f(X,Y) = \begin{pmatrix} X \\ \bar{Y} \\ X^2 \\ X\bar{Y} \\ \bar{Y}^2 \end{pmatrix}$$



Ongoing work…

How can you contribute as a Statistian and/or SP?

- How to sample the NROY region? Space-filling designs in disconnected regions
- Is it maybe more interesting to choose next points to use for next waves using some sort of metric, like in Bayesian Optimization?
- Relationships with MCMC methods?
- For the first wave, we could maybe use techniques like Ensemble Kalman Sampling (Cleary et al. 2020)
- Efficient techniques to find good kernel functions and mean function fo Gaussian Processes
- Data visualisation?

Discussion and Future work

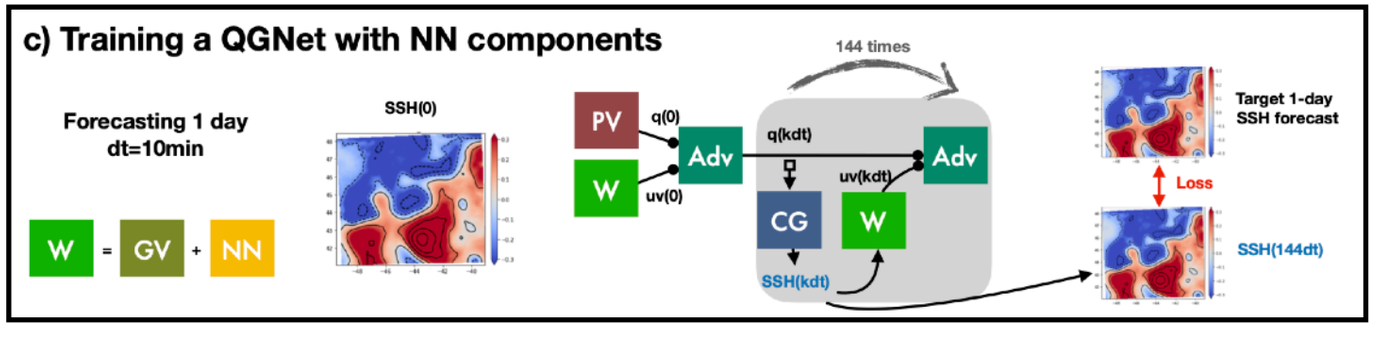
- HM is a powerful technique that would save much time and energy spent by modeling centers when tuning their models
- While presented as an automatic tool, HM depends on various choices made by the modeler (metrics, priors, implausibility thresholds…)
- Keeping humans in the loop is rather encouraged — connections with active learning?

Next steps:

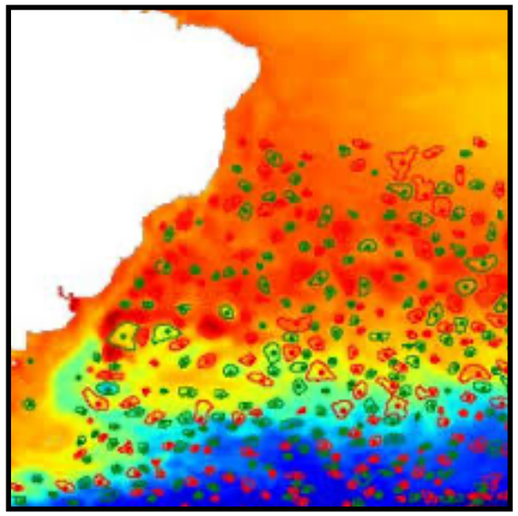
- Tuning both components of the L96 **independently** then looking at the intersection of their tuned space of parameters, investigating the effect on the original **coupled** model
- Investigate the effect of noisy and partial observations
- Weak coupling vs Strong coupling
- Look more into the machine learning component of History Matching (neural networks instead/combined with Gaussian Processes)

Other {ML+Ocean} projects

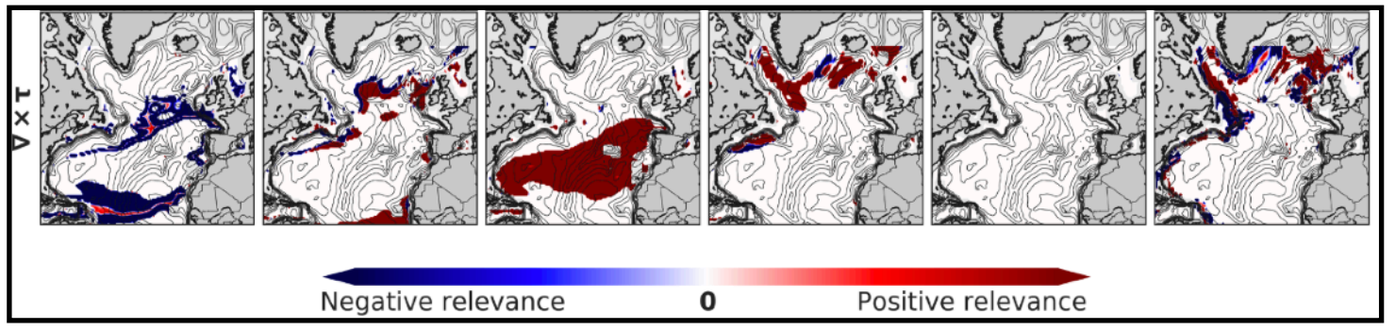
Learning Generalized Quasi-Geostrophic Models Using Deep Neural Numerical Models
Machine Learning and the Physical Sciences workshop, NeurIPS 2019, Vancouver, Canada.



EddyNet: A Deep Neural Network For Pixel-Wise Classification of Oceanic Eddies. IEEE Geoscience and Remote Sensing Symposium (IGARSS) 2018, Valencia, Spain.



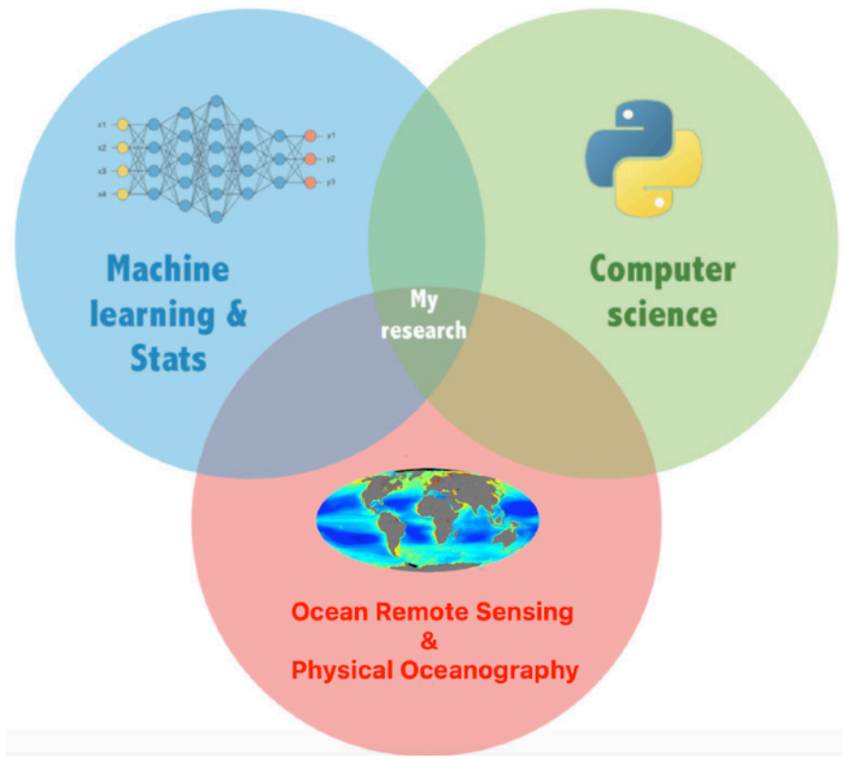
Revealing the impact of global heating on North Atlantic circulation using transparent machine learning. Submitted.



Keep in touch

I work on several applications of machine learning for physical oceanography and ocean remote sensing. If you have questions or have subjects to discuss, please do not hesitate to send me an email:

redouane.lguensat@locean.ipsl.fr



@redouanelg
redouanelg
redouanelg.github.io