Réunion ANR MOPGA-HRMES, 29/01/2021

History Matching for the tuning of climate models

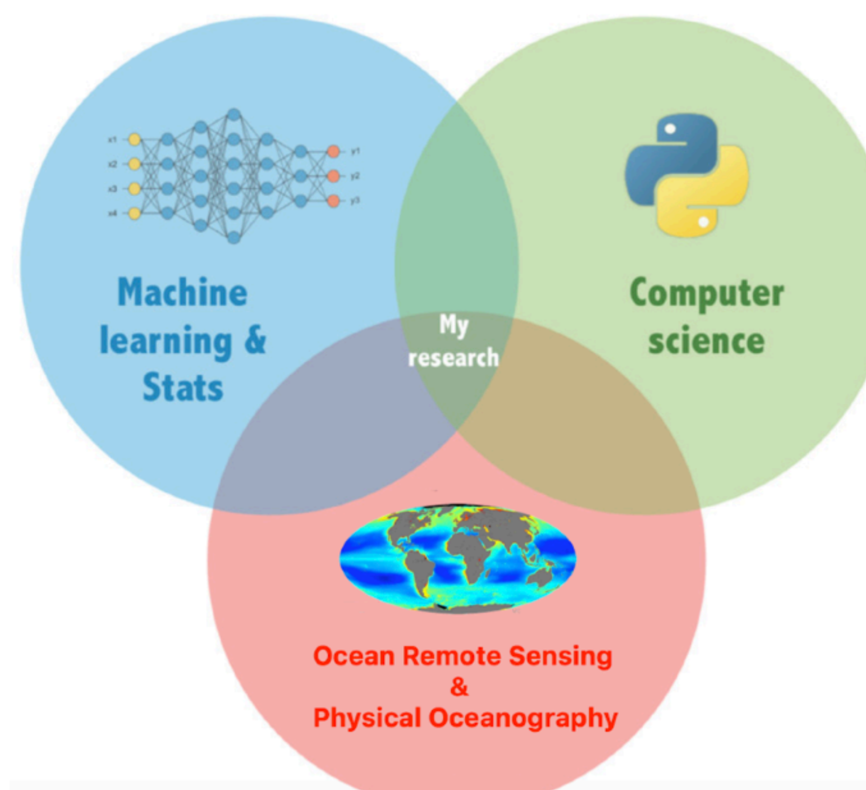
Redouane Lguensat
Postdoc @ LOCEAN-LSCE/Sorbonne University

Background

Interdisciplinary researcher

Education:

* Statistical Signal/Image Processing (a.k.a Data Science) engineer
* PhD IMT Atlantique (Grande Ecole) 2017: Machine Learning for Ocean Science

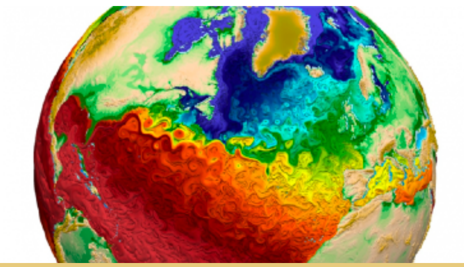


Research interests:

* 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

Climate model tuning

Parameters



Climate model simulation code

$f(x)$

Metrics, ex:

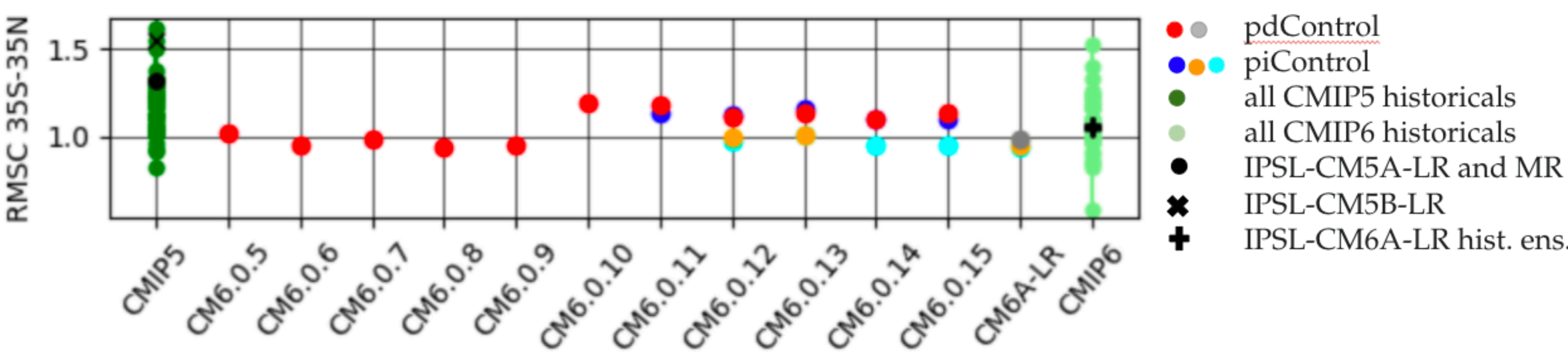
* Meridional Heat Transport
* AMOC
* Temperature profiles
* Etc…

Compare with

z

Observations

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

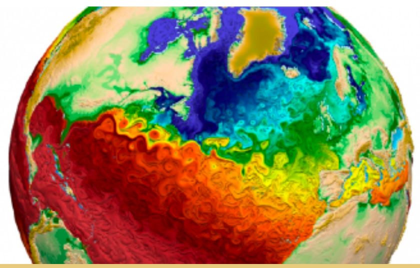


Evolution of RMSC SST averaged over 35S-35N tropical band along the course of IPSL-CM6A-LR tuning (Figure extracted from Mignot et al. 2021, sub.)

History Matching
for climate model tuning

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

History Matching:

* Has roots in Uncertainty Quantification (UQ)
* 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

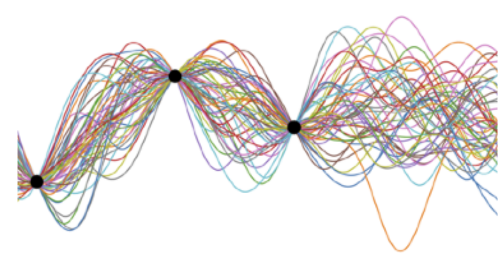
Ideally we would individually check every possible parameter setting for the input: **Impossible** (climate models are expensive to run)

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

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

HM algorithm:

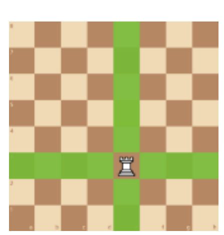
Initial guess of parameter space



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

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

Ex: using Latin Hypercube Sampling



Ex: using Gaussian Process Regression

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
Open Access
EGU

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?

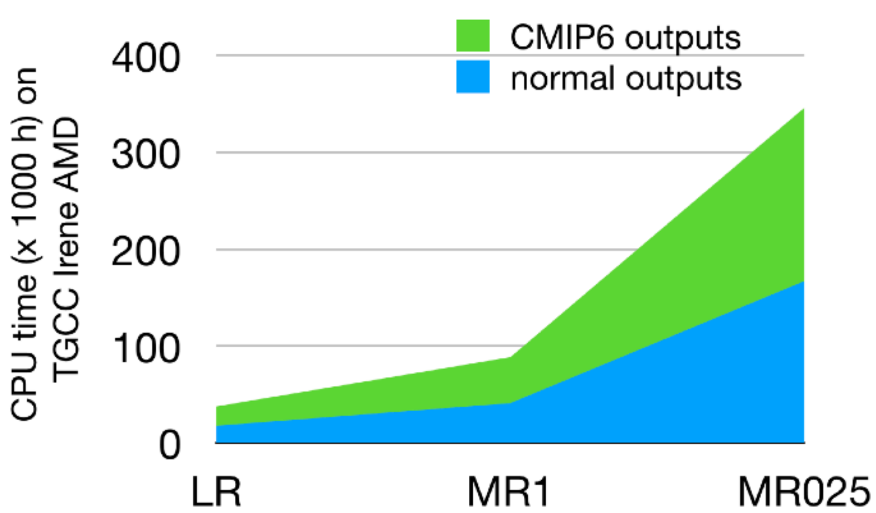
HRMES

Coupling

<u>Example:</u> 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 and V. Balaji

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

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

HM on the L96

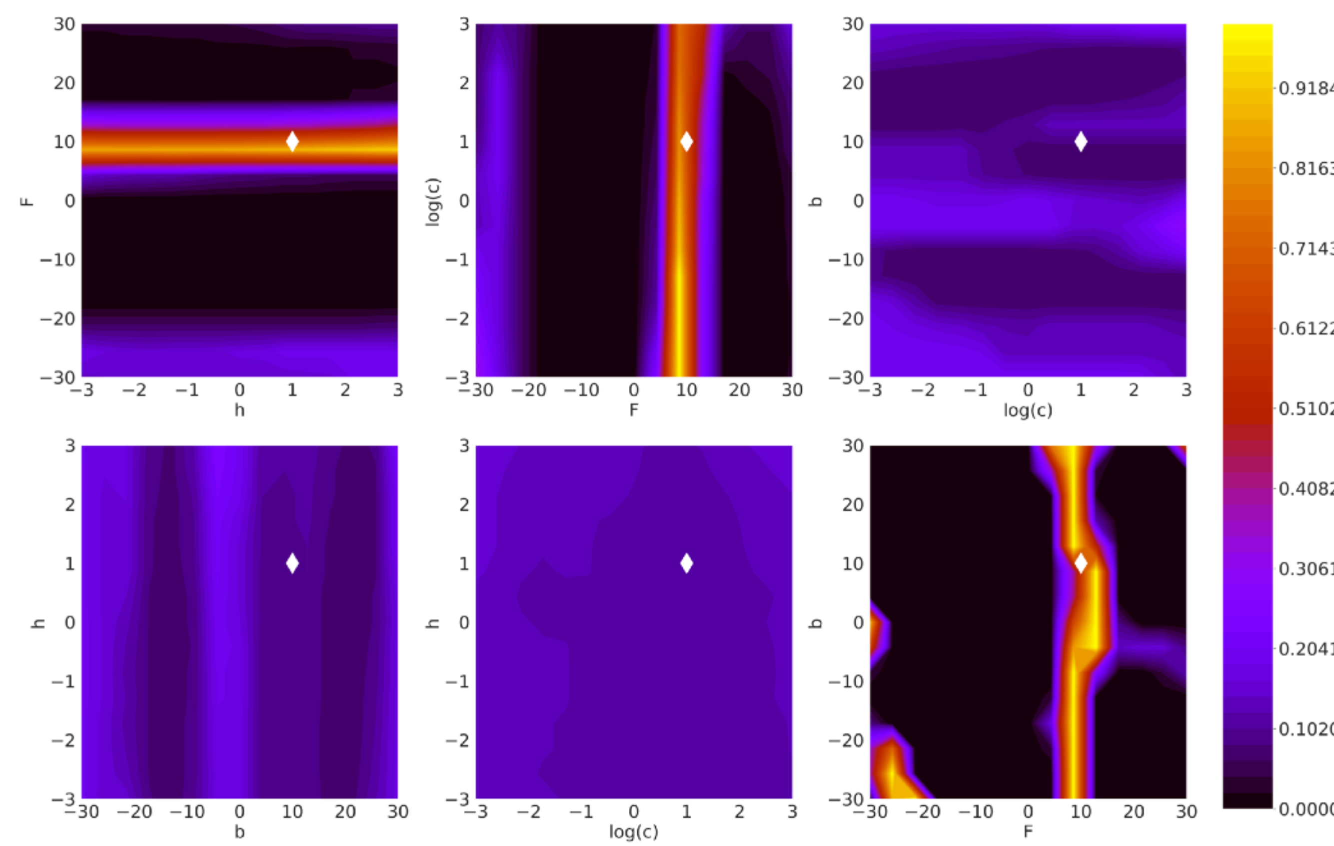
* **Metrics**: long-term time means to mimic climatological quantities
* **Ground Truth**: perfect setting K=36 X variables each coupled with J=10 Y variable. F=10, h=1, c=10, b=10, chaotic behavior.
* **HM code:** Python code run on **Jean-Zay** cluster + Parallel computation + ML models can be trained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

Parameter F is highly constrained since the first HM wave !

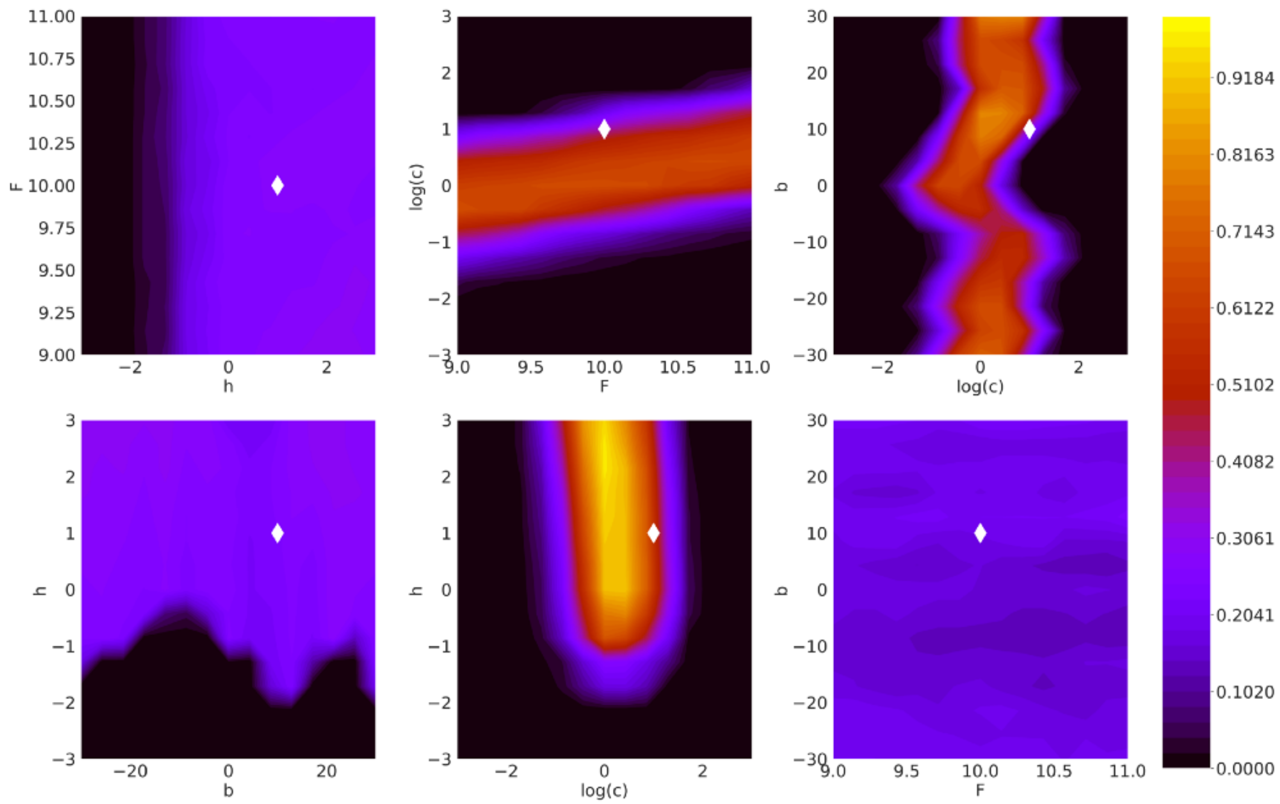
wave 1



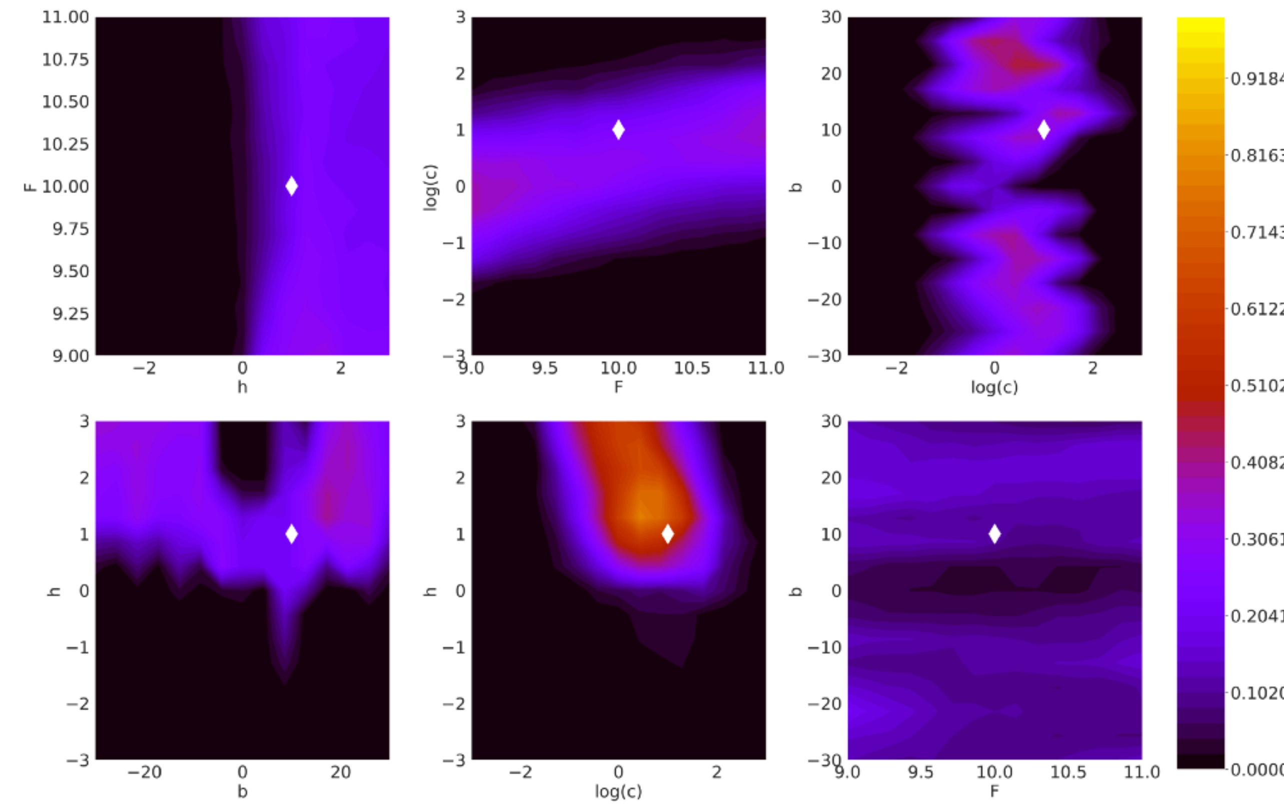
HM on the L96

Change F prior to a uniform distribution between 9 and 11, then reapply HM:

wave 3



wave 4



Parameters associated with the large scale components contribute highly to the metrics, HM focus on them initially, then moves to the parameters associated with small scale component

Ongoing work…

Future work

- 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

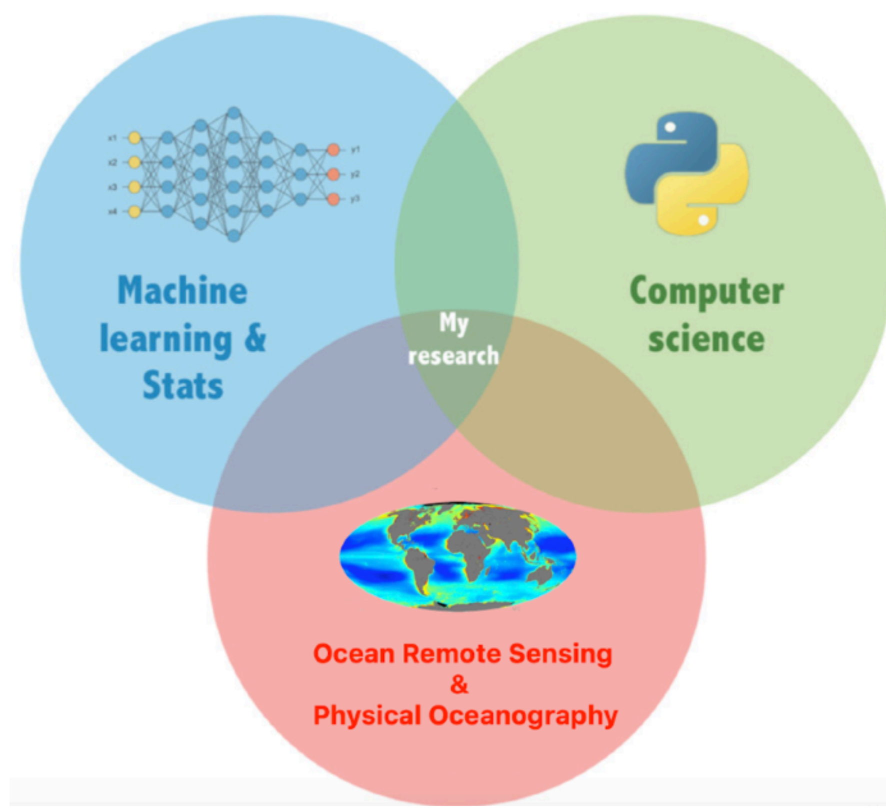
- Apply HM on NEMO ORCA 1°

- Look more into the machine learning component of History Matching (neural networks combined with Gaussian Processes) -> **M2 internship** of Homer Durand

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@lsce.ipsl.fr



@redouanelg redouanelg redouanelg.github.io