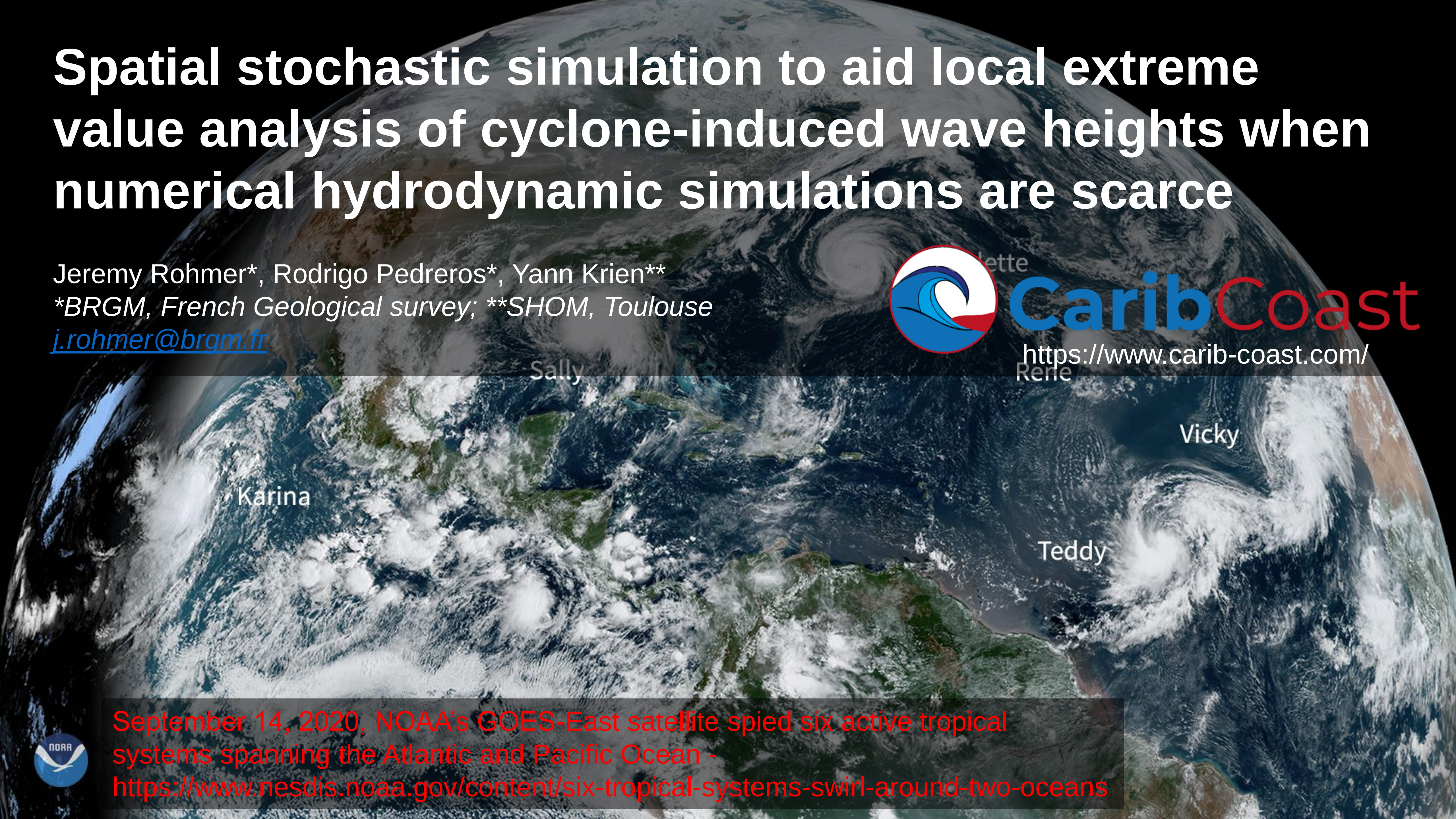
Spatial stochastic simulation to aid local extreme value analysis of cyclone-induced wave heights when numerical hydrodynamic simulations are scarce

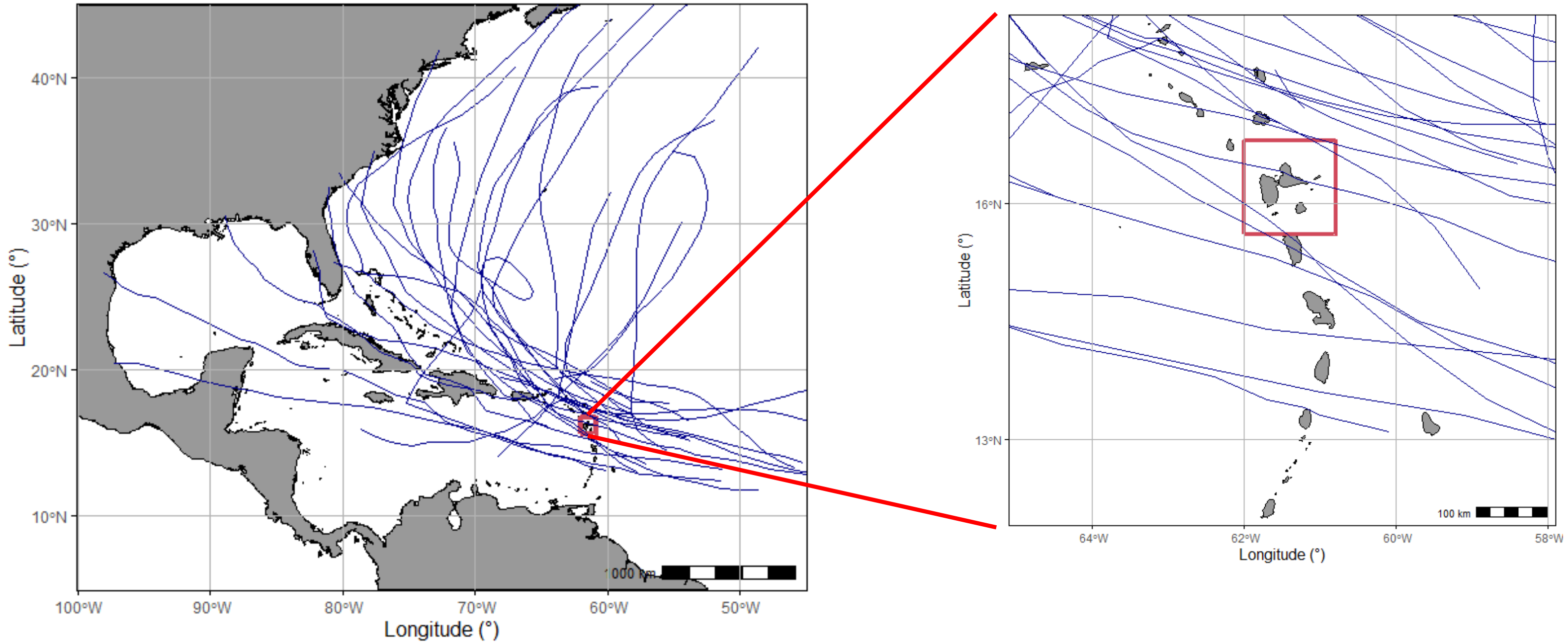
Jeremy Rohmer*, Rodrigo Pedreros*, Yann Krien**
*BRGM, French Geological survey; **SHOM, Toulouse
j.rohmer@brgm.fr

CaribCoast
https://www.carib-coast.com/



September 14, 2020, NOAA's GOES-East satellite spied six active tropical systems spanning the Atlantic and Pacific Ocean - https://www.nesdis.noaa.gov/content/six-tropical-systems-swirl-around-two-oceans

Motivating test case: cyclone-induced extreme waves at Guadeloupe archipelago (French West Indies)



Cyclone Track (HURDAT database 1979-2019) at a distance <400km from the Island

Objective: estimate extreme wave heights (Hs) induced by tropical cyclones TC along the Guadeloupe coast

Difficulty: historical data are relatively scarce (28 TCs)

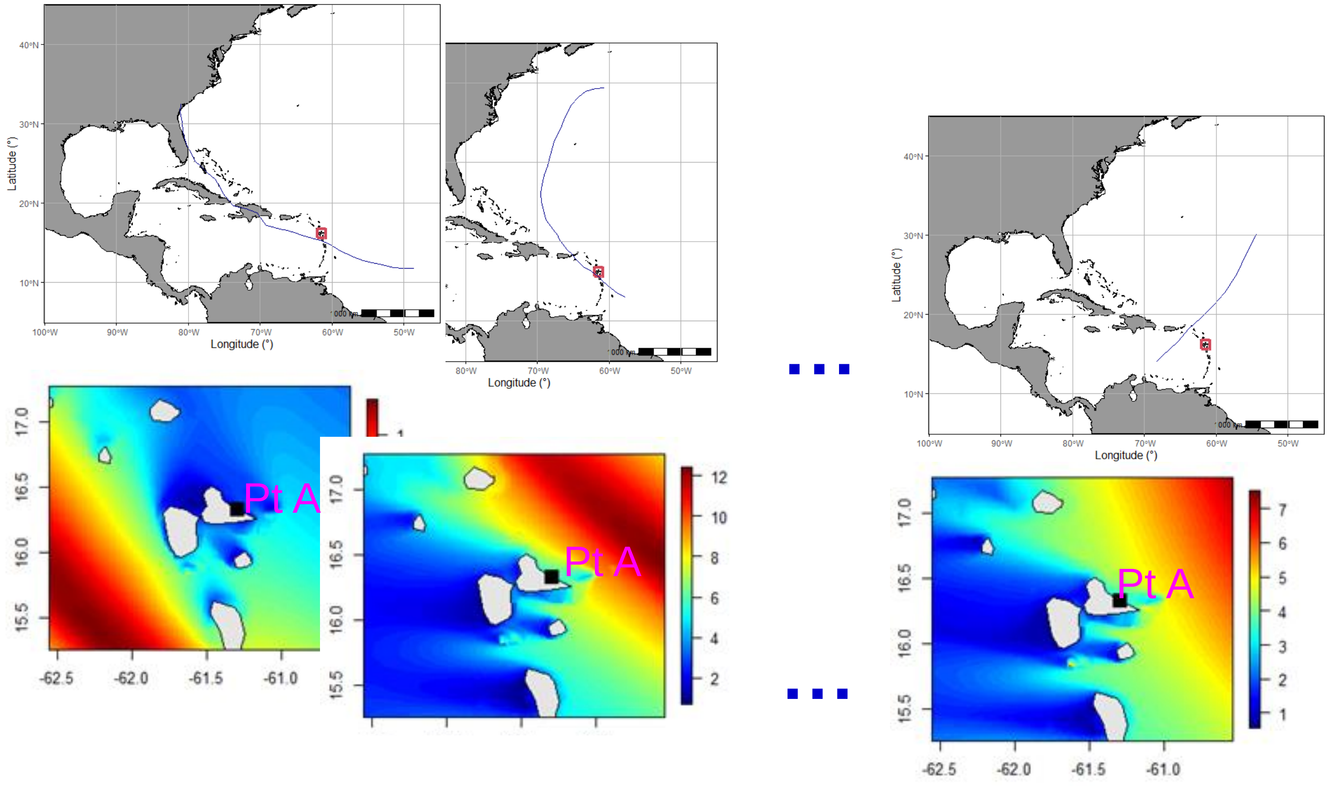
HURDAT: https://www.aoml.noaa.gov/hrd/hurdat/Data_Storm.html

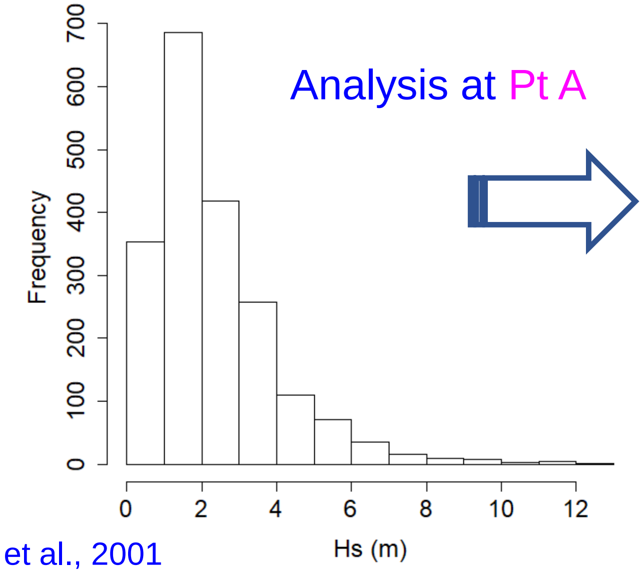
Synthetic TC approach

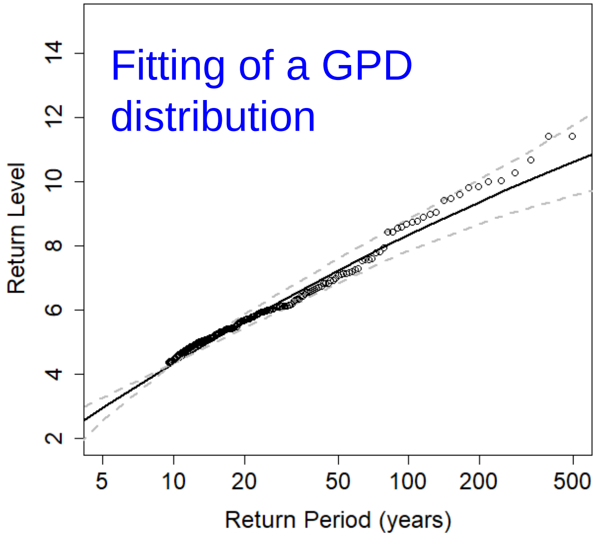
1 **Randomly generate** a large number of synthetic TCs using stochastic generator [1] **2,000 TCs**

2 **Compute** the TC-induced wave heights using numerical hydrodynamic simulators (e.g. ADCIRC-SWAN [2])

3 **Extract Hs maximum values** around the coasts of Guadeloupe and perform **extreme value analysis** [3] to evaluate the return levels of interest (e.g. 100y Hs RL)



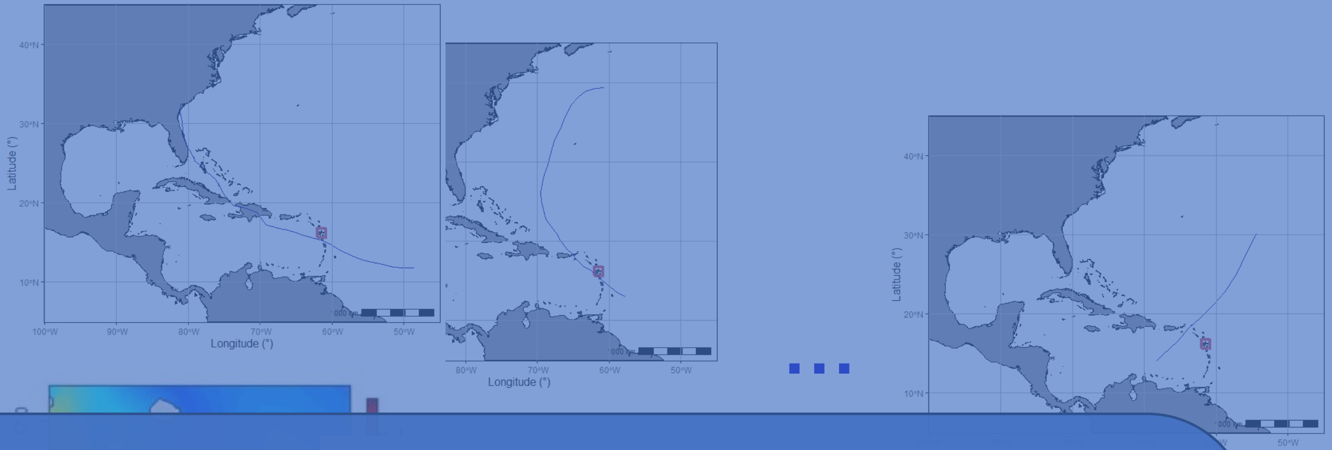




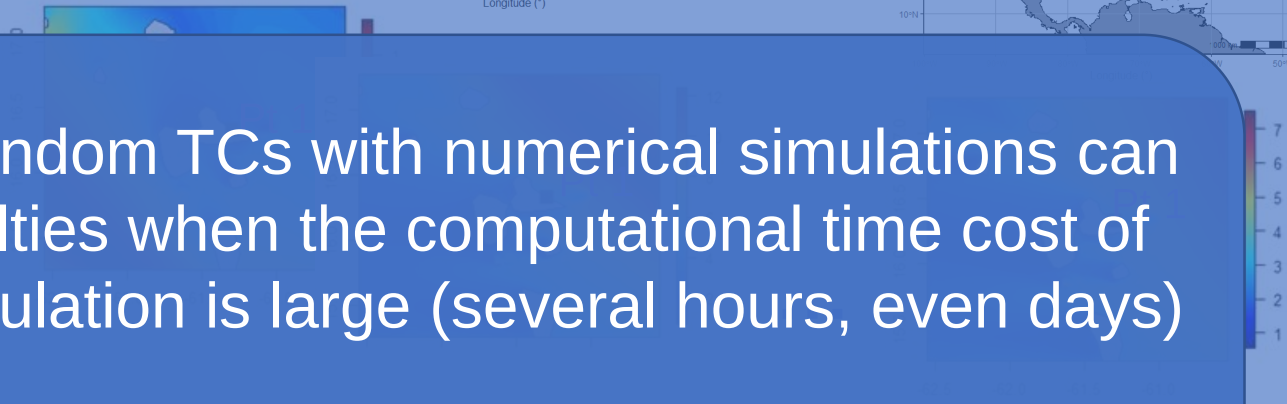
[1]: Emanuel et al. BAMS 2006; [2]: Krien et al., NHESS 2015; [3]: Coles et al., 2001

Synthetic TC approach

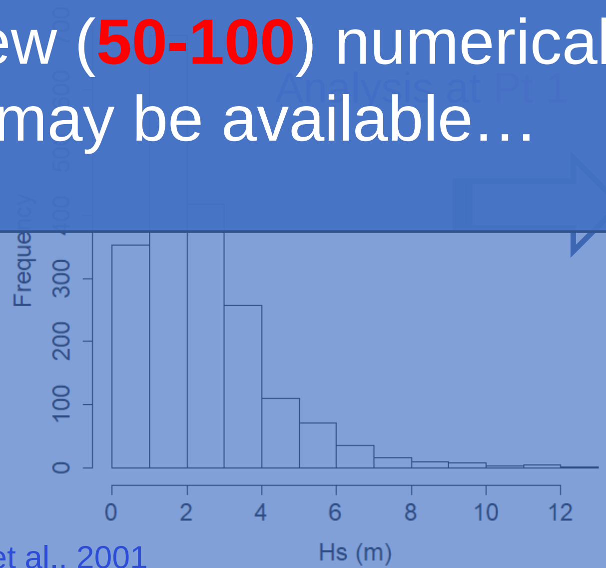
1 **Randomly generate** a large number of synthetic TCs using stochastic generator [1] **2,000 TCs**

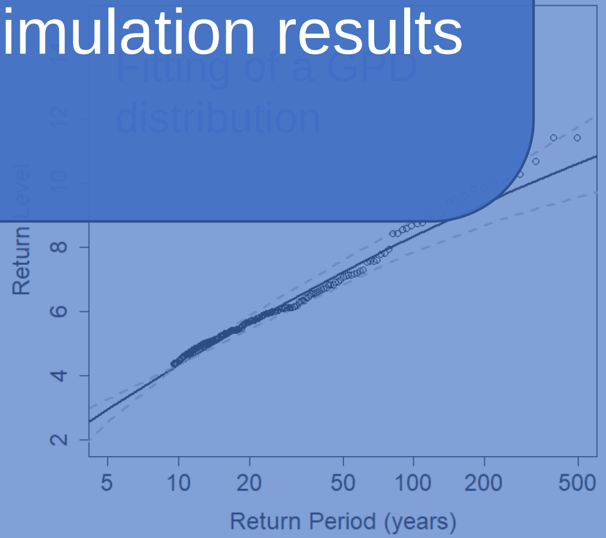


2 **Compute** the TC-induced wave heights using numerical hydrodynamic simulators (e.g. ADCIRC-SW



3 **Extract Hs maximum values** around the coasts of Guadeloupe and perform **extreme value analysis** [3] to evaluate the return levels of interest (e.g. 100y Hs return level)





The combination random TCs with numerical simulations can face severe difficulties when the computational time cost of each numerical simulation is large (several hours, even days)

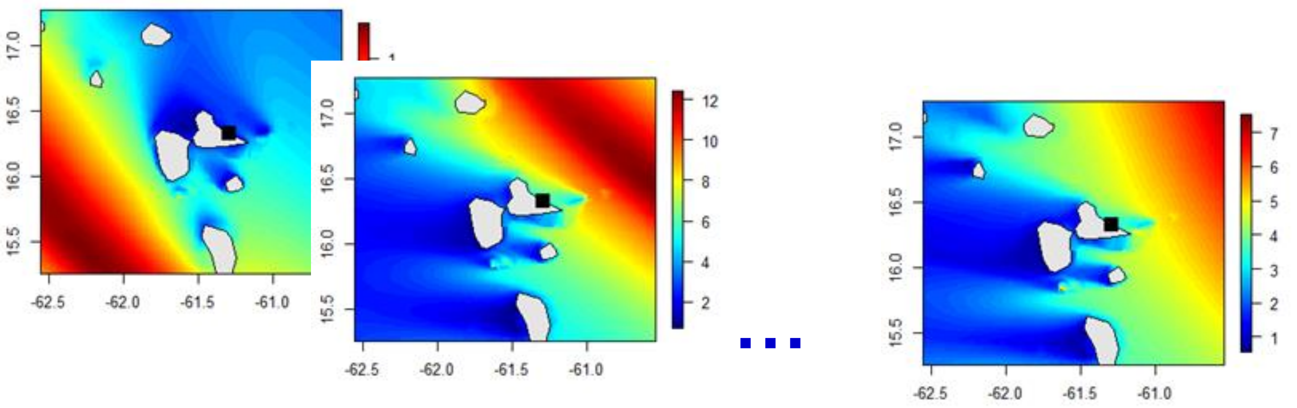
In practices, only a few (**50-100**) numerical simulation results may be available…

[1]: Emanuel et al. BAMS 2006; [2]: Krien et al., NHESS 2015; [3]: Coles et al., 2001

Proposed approach

1

Compute the TC-induced maxmum Hs for the limited number of available TCs (M=50-100)

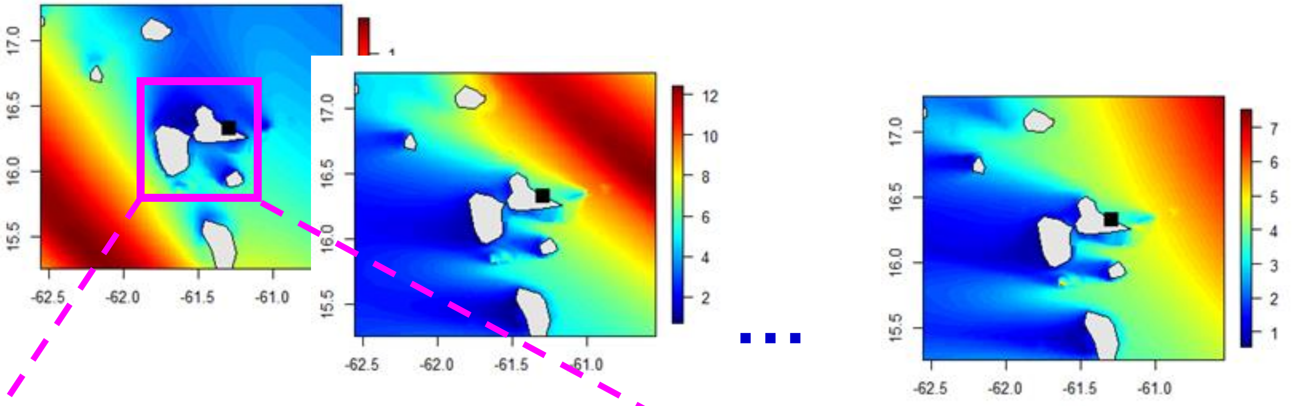


...

Proposed approach

1 **Compute** the TC-induced maxmum Hs for the limited number of available TCs (M=50-100)

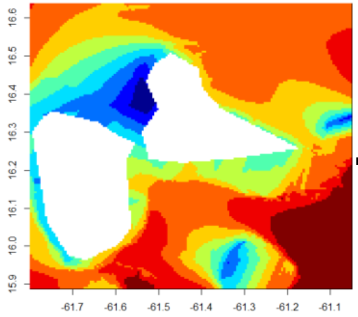
Map (1)



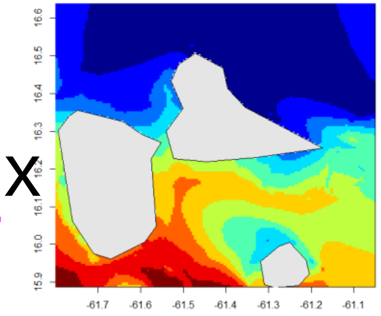
2 Decompose the Hs maps onto suitable basis functions using **Empirical Orthogonal Functions [1]**

EOF1 EOF2 EOF3

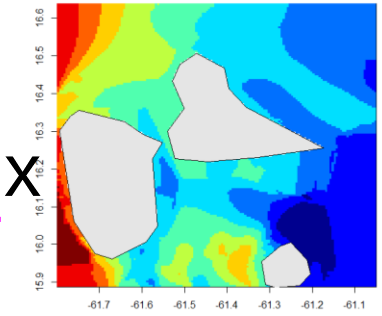
Map(1)= $\alpha_{1,1}$x



+$\alpha_{2,1}$x



+$\alpha_{3,1}$x

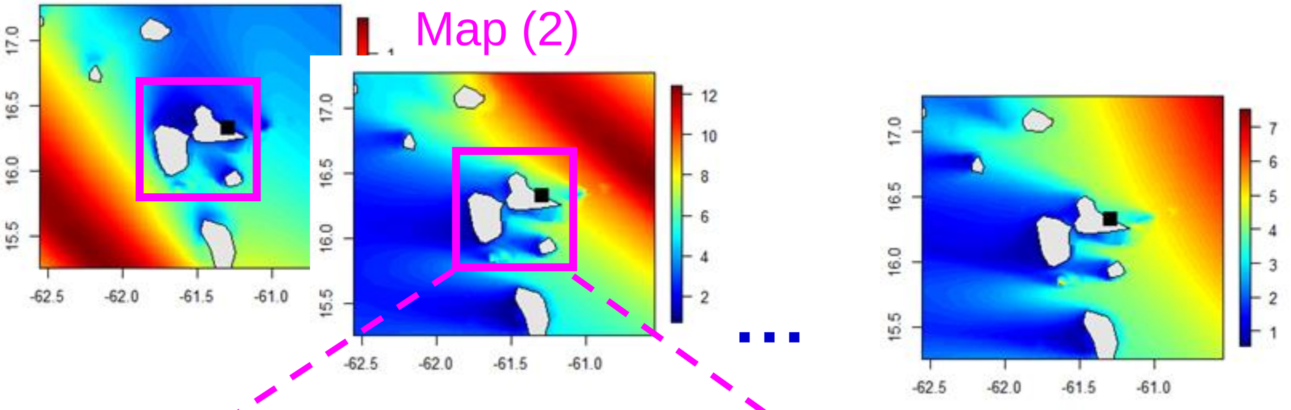


...

[1] Jolliffe & Richman, 1987

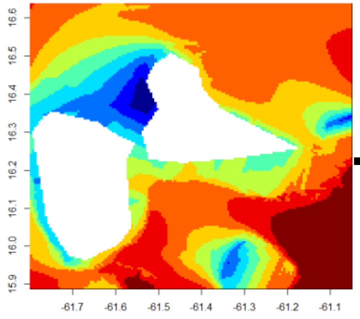
Proposed approach

1. **Compute** the TC-induced maxmum Hs for the limited number of available TCs (M=50-100)

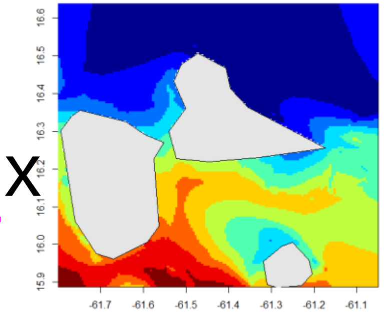


2. Decompose the Hs maps onto suitable basis functions using **Empirical Orthogonal Functions [1]**

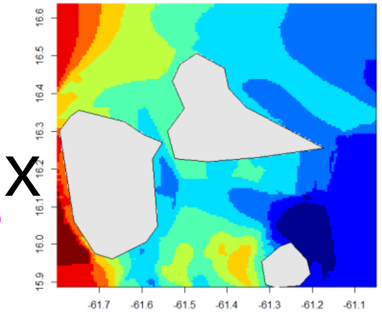
EOF1

$\text{Map}(2) = \alpha_{1,2} \times$ 

EOF2

$+ \alpha_{2,2} \times$ 

EOF3

$+ \alpha_{3,2} \times$ 

...

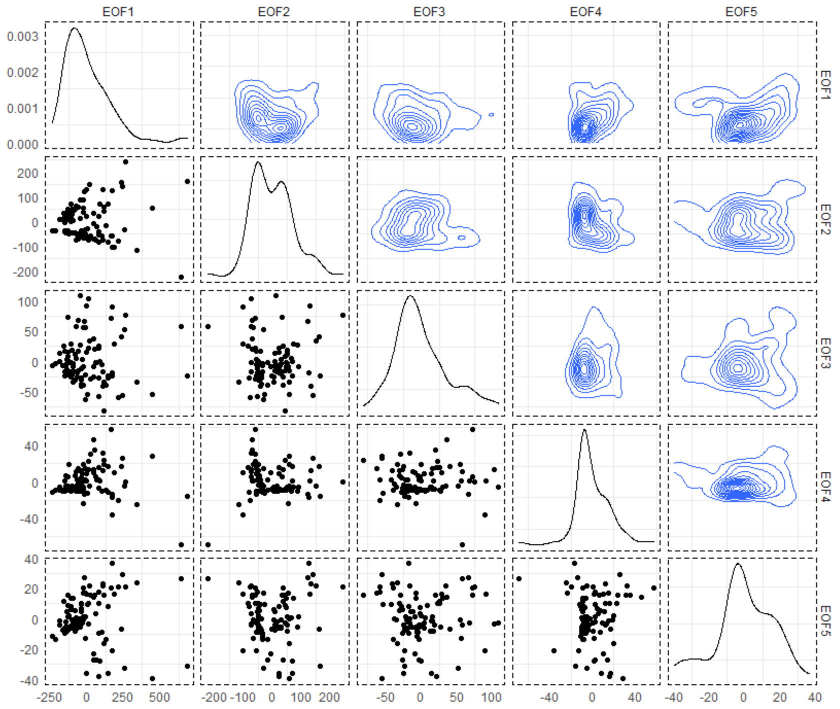
Etc…

[1] Jolliffe & Richman, 1987

Proposed approach

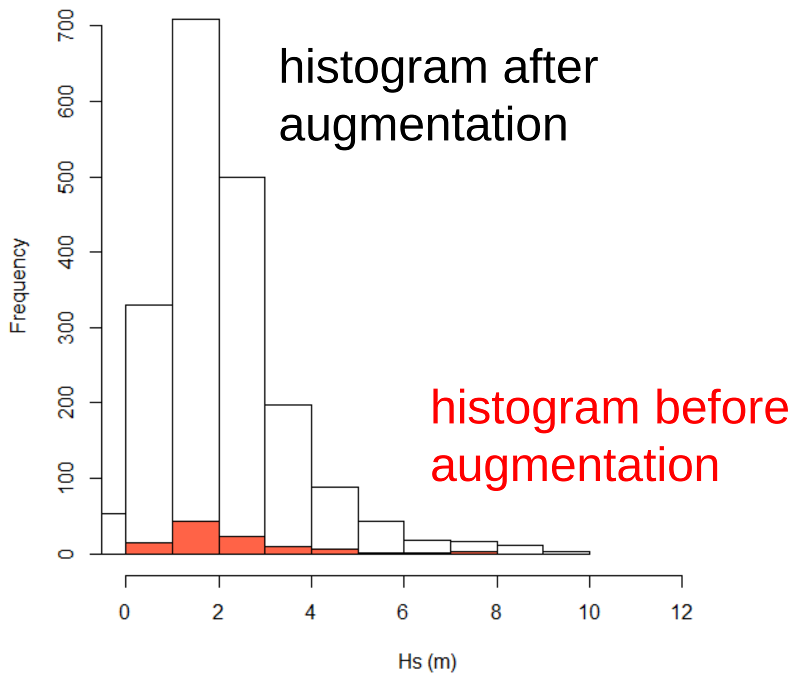
3

Learn the joint statistical law of the expansion coefficients $\alpha_{k,i}$ $k = 1 \ldots n, i = 1 \ldots M$ with n=the number of EOFs and M the number of maps
Method: multivariate Gaussian kernel smoothing



4

Sample from the joint law and reconstruct the synthetic Hs maps. Then, **augment** the observations at the location of interest with the synthetic data

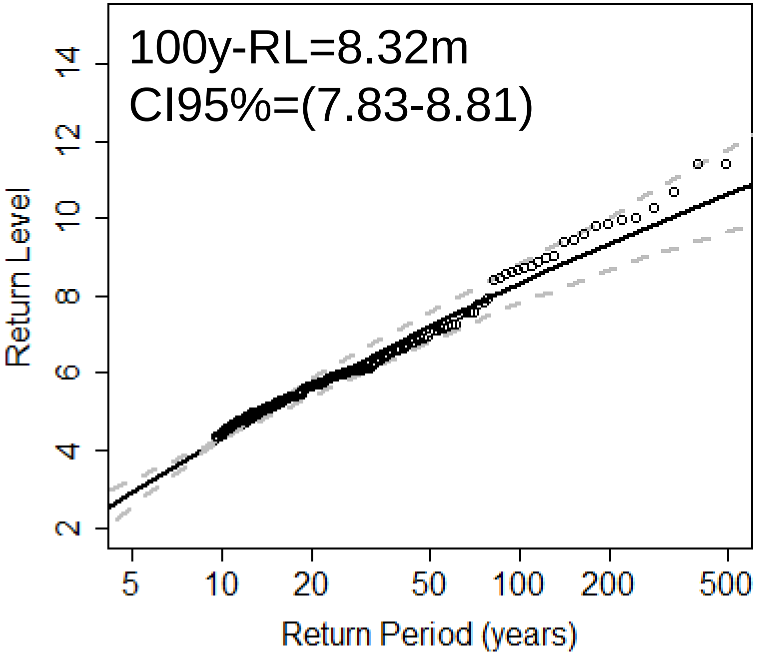


At point Pt A

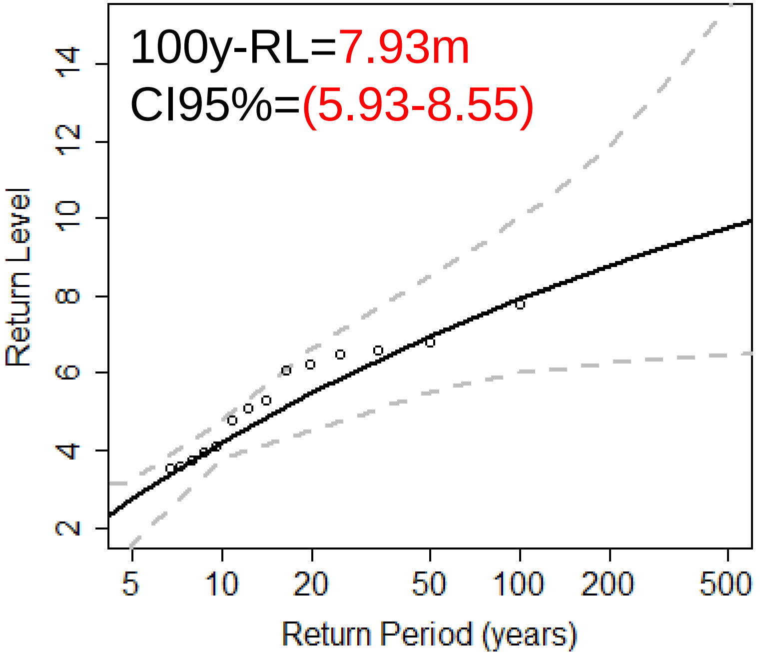
Géosciences pour une Terre durable
brgm

At point Pt A using M=100 Hs maps and n=4 EOFs

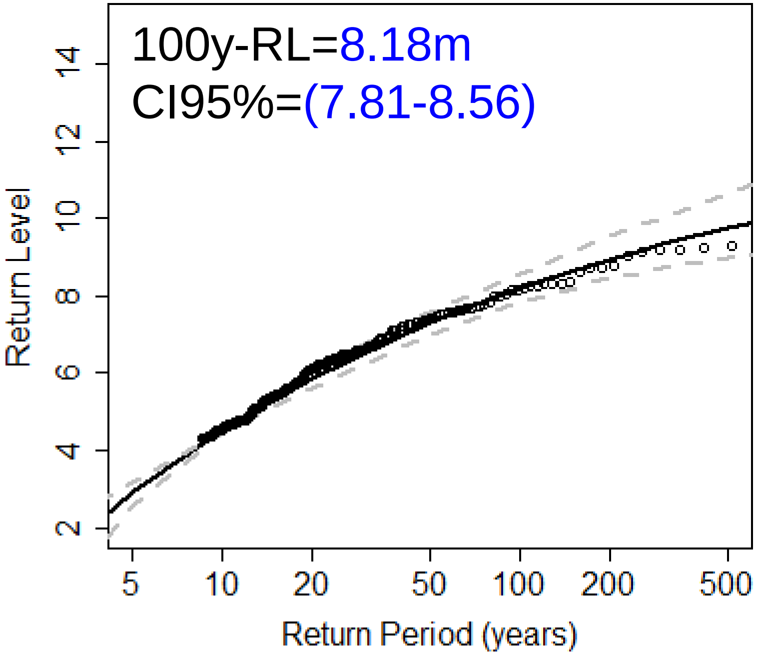
Comparison of Return Levels



With 1971 synthetic TCs
(**Full** Synthetic TC approach)



With 100 synthetic TCs
(**Sparse** Synthetic TC approach)



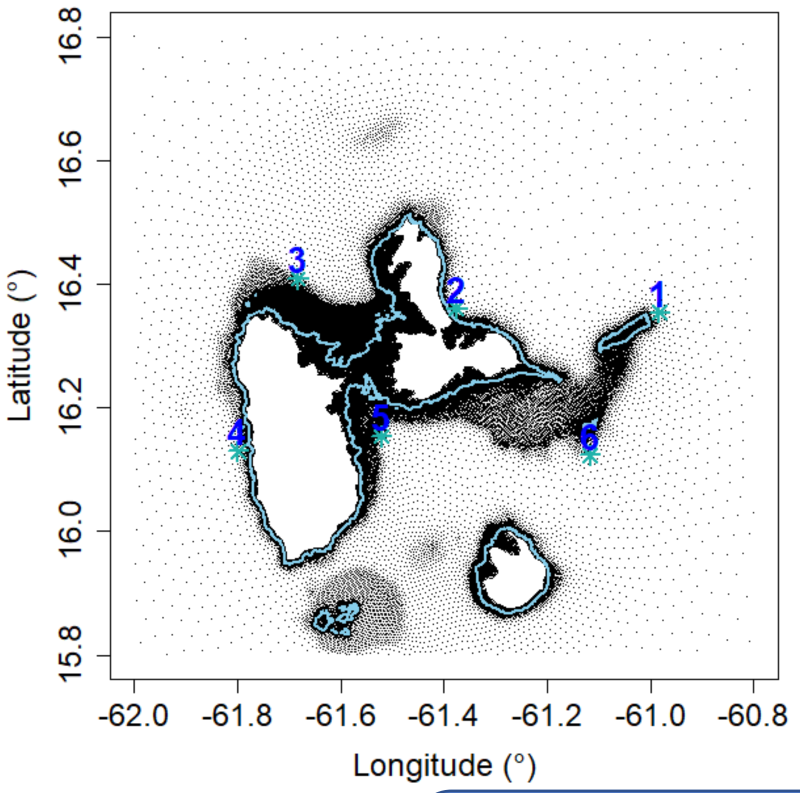
With data augmentation using EOF-based reconstructed Hs maps

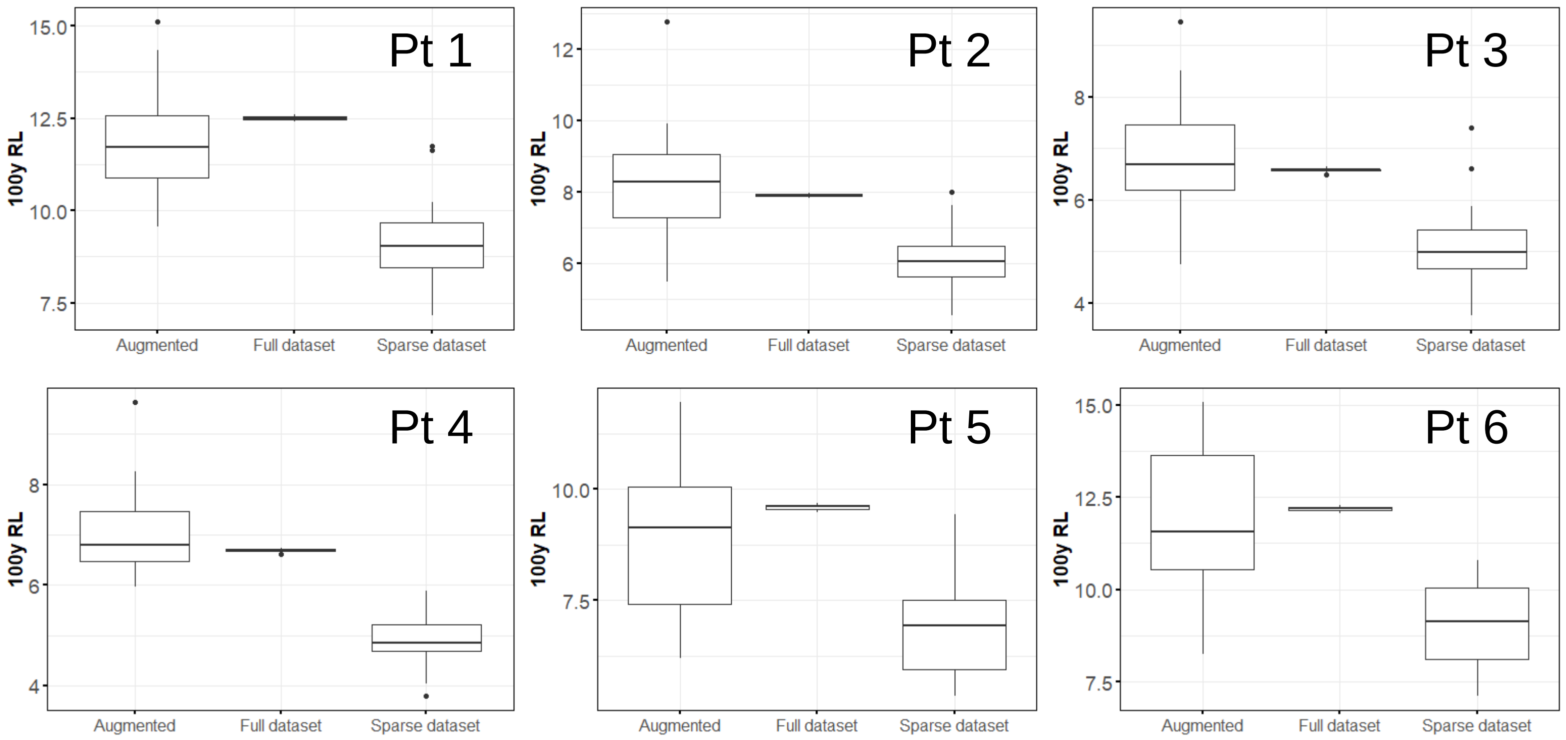
Data augmentation appears to decrease the bias and the uncertainty (95% Confidence Interval width)

Géosciences pour une Terre durable
brgm

Validation exercise – 100y RL estimates

Use of **M=100 randomly sampled TCs** among the full dataset of Krien et al. (2015) composed of 1971 TCs (representative of 3,200 years) with **n=4 EOFs (replicated 25 times)**

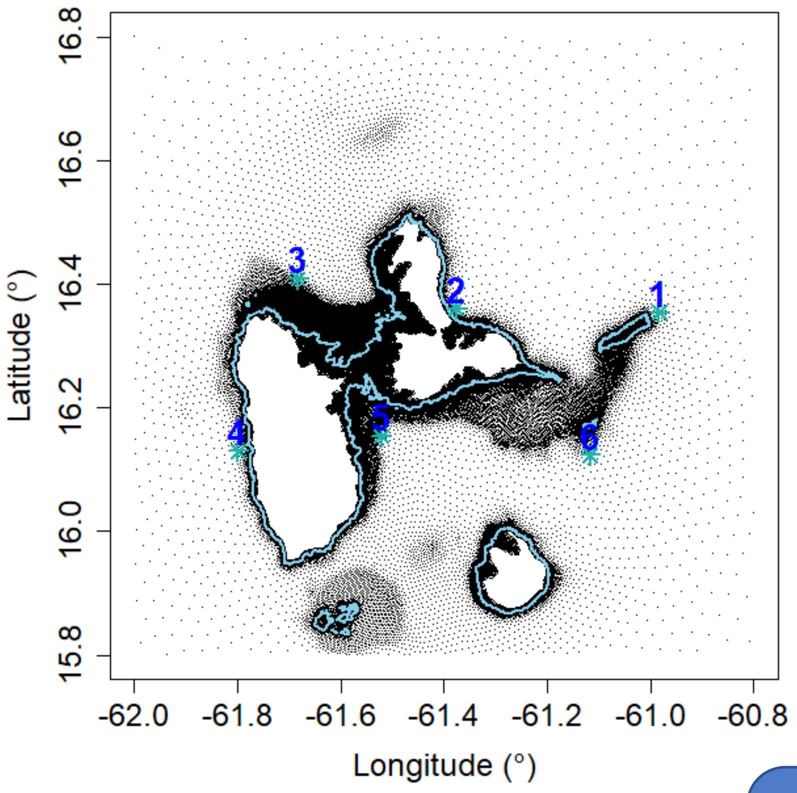


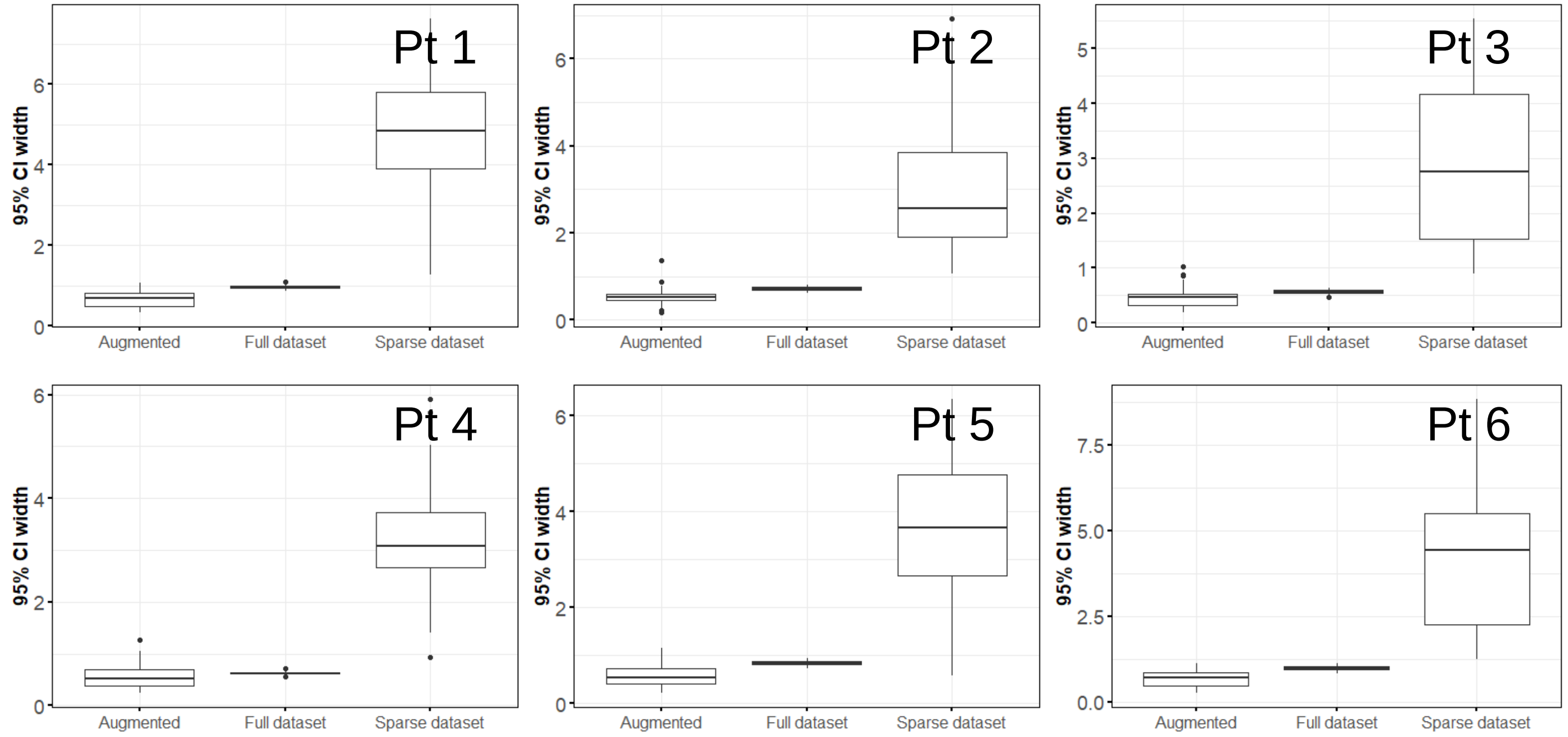


100y RL estimates with data augmentation has the lower bias: absolute percentage error ~12.3% (averaged over the Pts) to be compared to 25.5% without

Validation exercise – 95% confidence interval (CI) width

Use of **M=100 randomly sampled TCs** among the full dataset of Krien et al. (2015) composed of 1971 TCs (representative of 3,200 years) with **n=4 EOFs (replicated 25 times)**



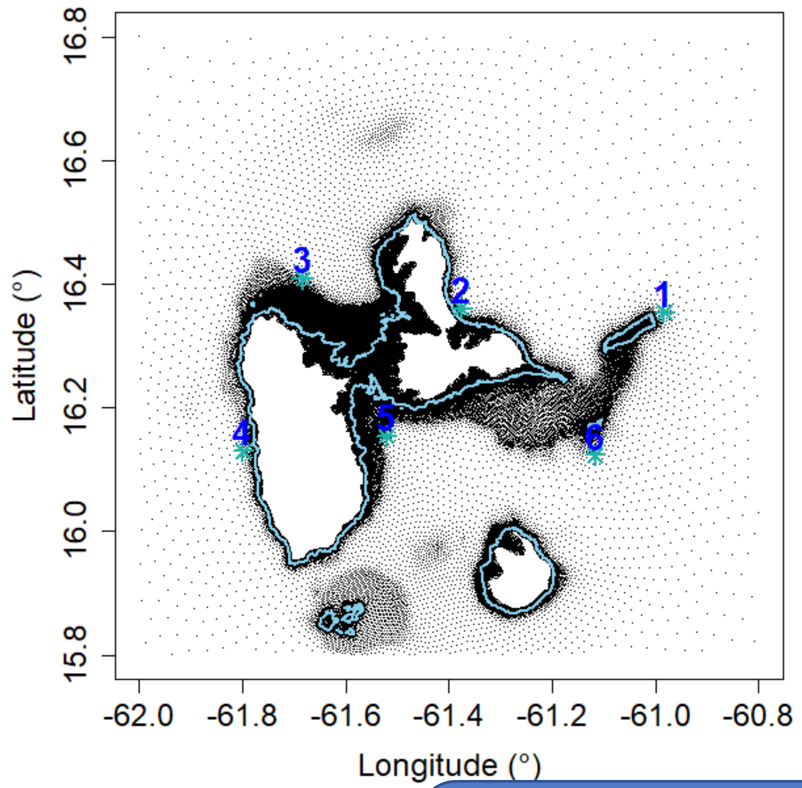


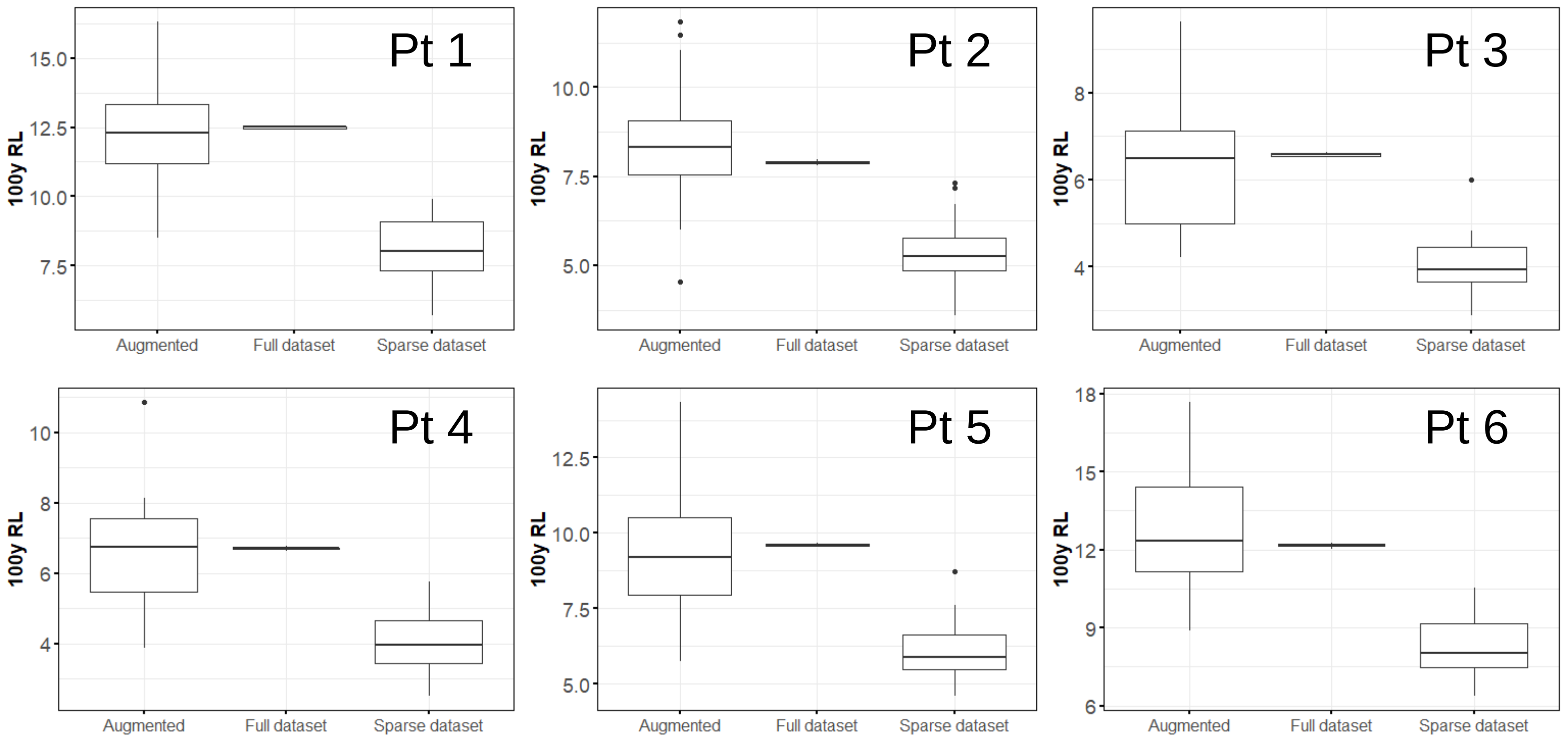
Data augmentation leads to CI width of the same order than the approach with the full dataset

Géosciences pour une Terre durable
brgm

Influence of the number of maps – 100y RL estimates

Use of **M=50 randomly sampled TCs** among the full dataset of Krien et al. (2015) composed of 1971 TCs (representative of 3,200 years) with **n=4 EOFs (replicated 25 times)**

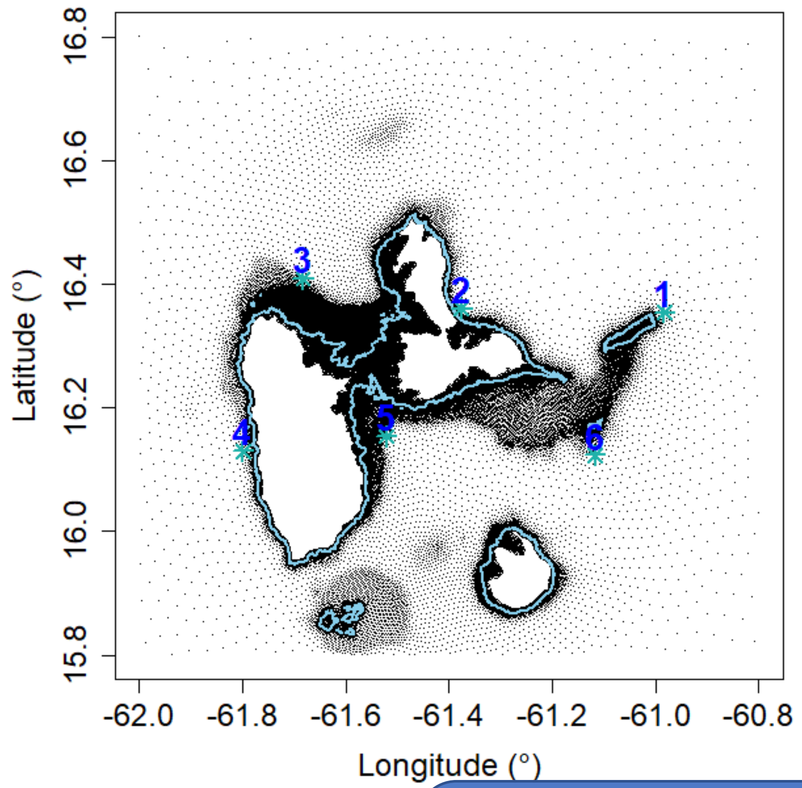


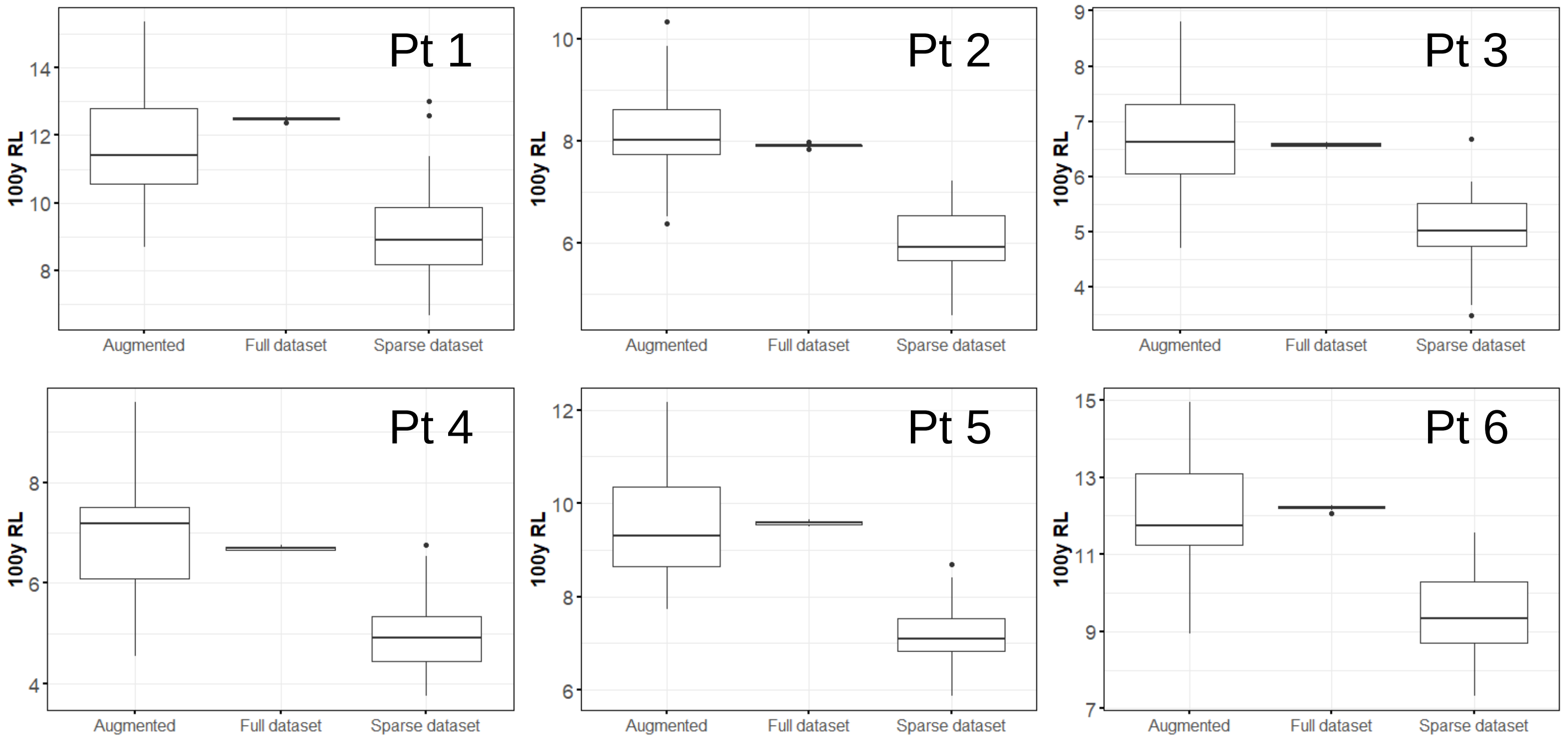


100y RL estimates with data augmentation has the lower bias: absolute percentage error ~16.2% (averaged over the Pts)

Influence of the number of EOFs – 100y RL estimates

Use of **M=100 randomly sampled TCs** among the full dataset of Krien et al. (2015) composed of 1971 TCs (representative of 3,200 years) with **n=3 EOFs (replicated 25 times)**

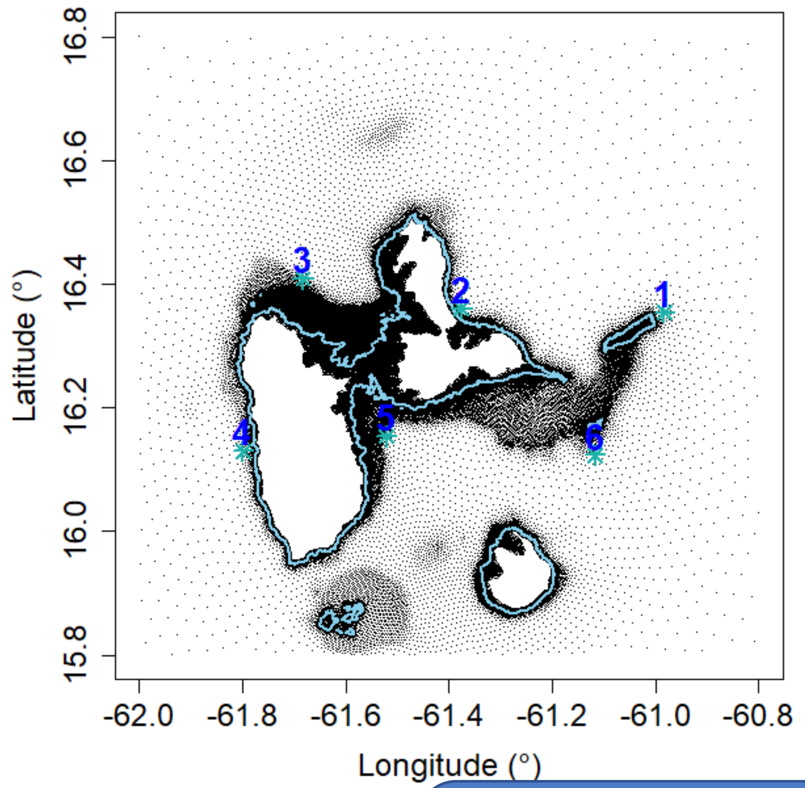


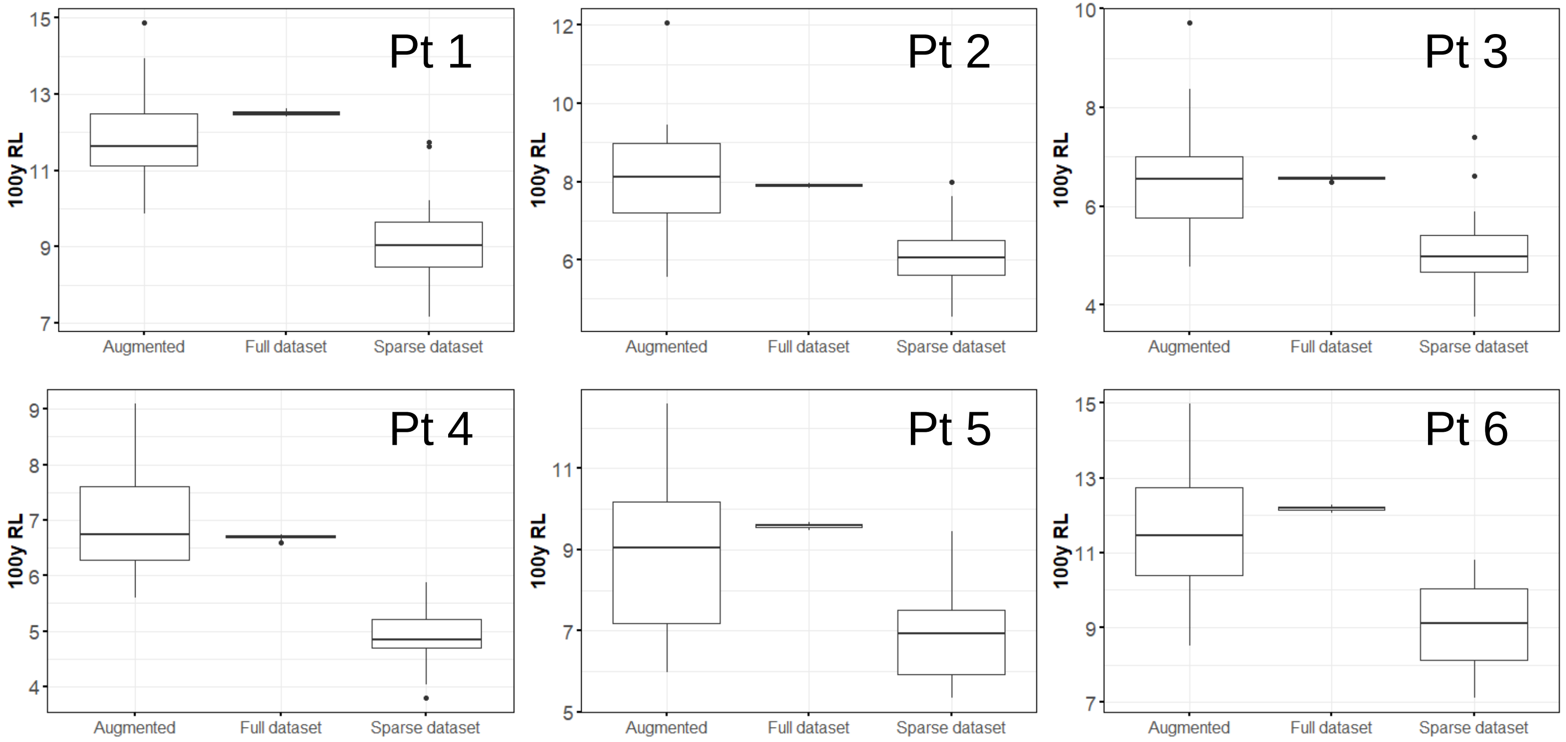


100y RL estimates with data augmentation has the lower bias: absolute percentage error ~12.2% (averaged over the Pts)

Influence of the decomposition method – 100y RL estimates

Use of **M=100 randomly sampled TCs** among the full dataset of Krien et al. (2015) composed of 1971 TCs (representative of 3,200 years) with **Partial Least Squares Regression [1] (replicated 25 times)**





100y RL estimates with Data augmentation has the lower bias: absolute percentage error ~12.3% (averaged over the Pts)

[1] Martens, 2001

Summary and open questions

- Estimates of 100 year Return Level imposes the use of a large dataset (**1,000-2,000**) of computed Hs maps; each of thme corresponding to a synthetic Tropical Cyclone

- To perform the analysis using 50-100 Hs maps, we propose to augment the available database with **reconstructed Hs maps using EOFs**

- Compared to a crude approach using only 50-100 maps, **results suggest that the bias is lowered (~10-15%) as well as the uncertainty estimates (95% confidence interval width)**

- **Open questions**
  - Could a limited number of synthetic tropical cyclones be **selected beforehand?**
  - Does the approach perform well for **larger return periods** (>100 years)?
  - Is the performance equivalent to **other procedures** like the recently developed STM-E approach [1] or resampling techniques specifically developed for extremes [2]?

[1] Wada et al., Ocean Eng. 2018; [2] Opitz et al., spat. Stat., 2020