Insights in uncertainty of source parameters : combining historical eyewitness accounts on tsunami-induced wave heights and numerical simulation

J. Rohmer, M. Rousseau, A. Lemoine, J. Lambert, R. Pedreros, B. Aalae

Géosciences pour une Terre durable
brgm

Contexte

Définition

> **Tsunami** ***(du japonais tsu : port et nami : vague)*** **:** une série d'ondes provoquée par une action mécanique brutale et de grande ampleur au niveau d'un lac, d'une mer ou d'un océan. Cette action peut être d'origine :

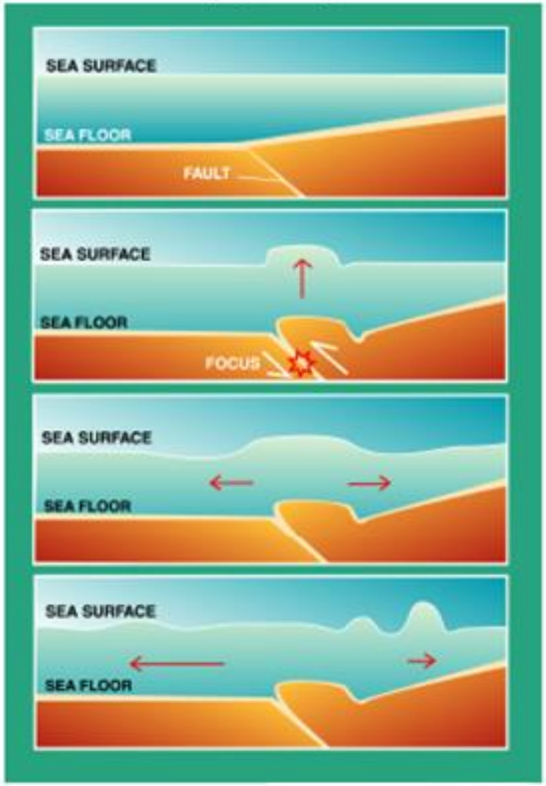
- Sismique (Sumatra 2004, Chili 1960-2010, Japon 2011…)
- Mouvement de Terrain (sous-marin: Papouasie Nouvelle Guinée 1998, Nice 1979; sub-aérien: Fatu-Hiva 1999; …)
- Volcanique (sous-marin: Kick'em Jenny ; sub-aérien, Montagne Pelée 1902;…)
- *Chute Astéroïde (…)*
- *Explosion atomique (…)*

Origine sismique

Génération par :

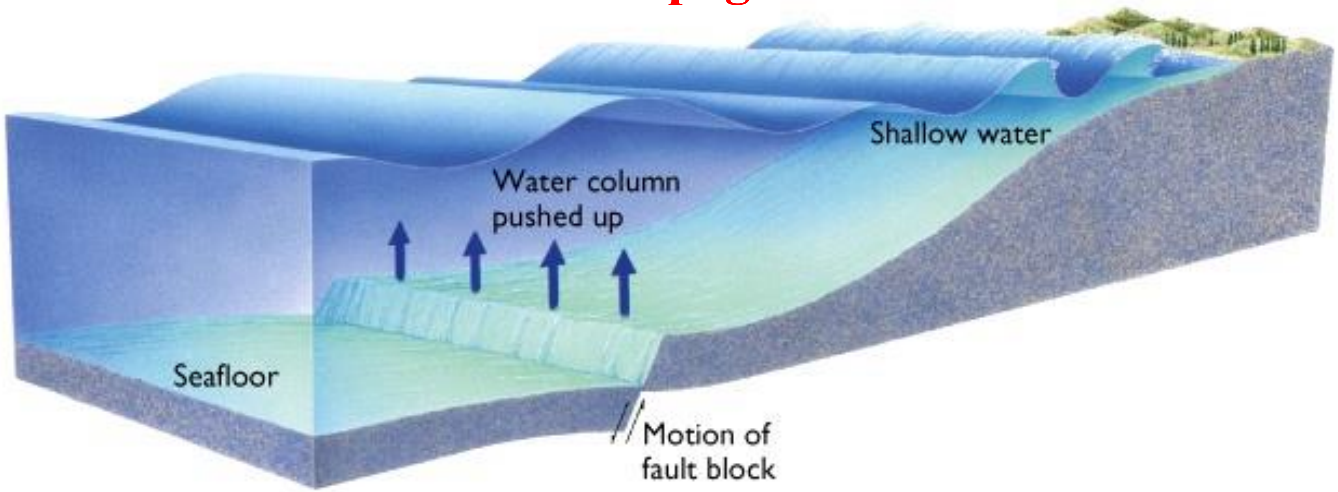
- un **séisme**
- un **glissement de terrain** subaérien ou sous-marin
- une **éruption volcanique**
- une **chute de météorite**

Schéma de la génération d'un tsunami par un séisme (©*ITIC*)

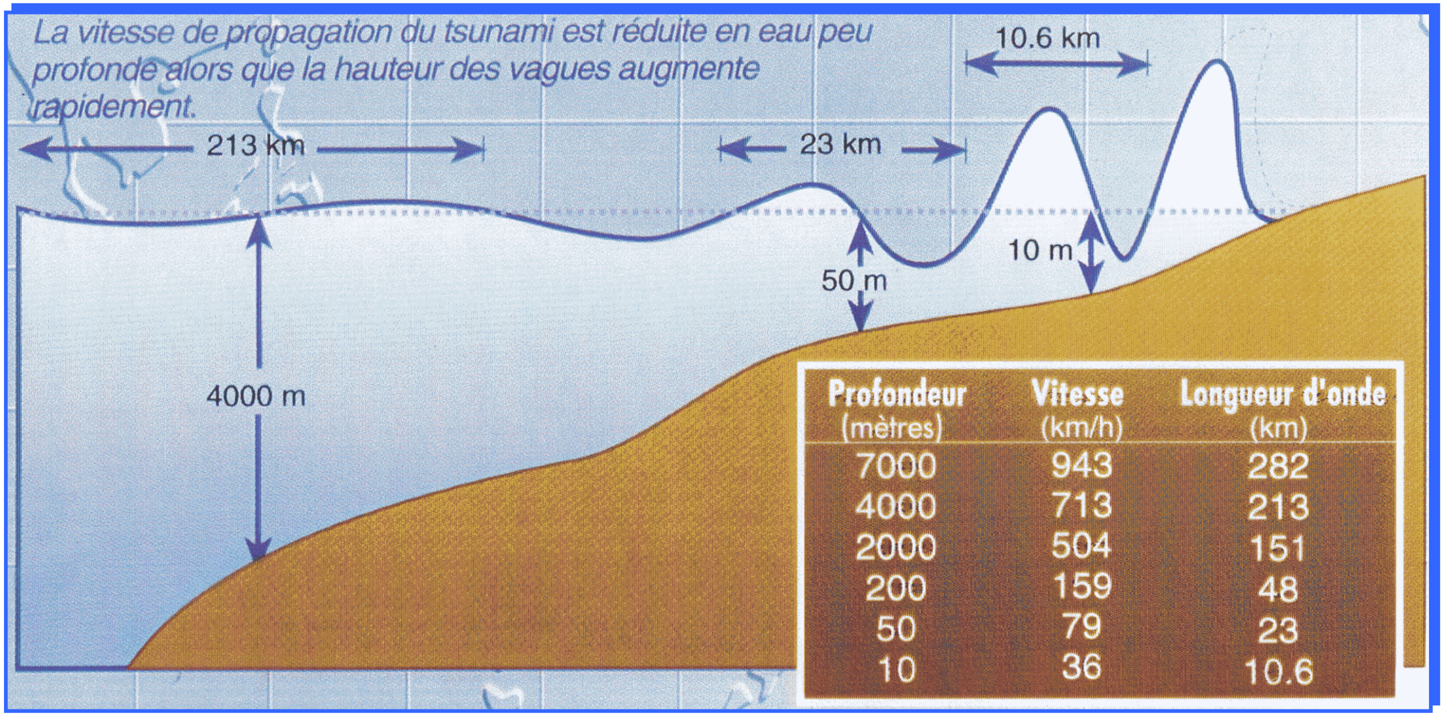


Trois phases

Génération **Propagation** **Inondation**



De la génération à l'inondation



Profondeur (mètres)	Vitesse (km/h)	Longueur d'onde (km)
7000	943	282
4000	713	213
2000	504	151
200	159	48
50	79	23
10	36	10.6

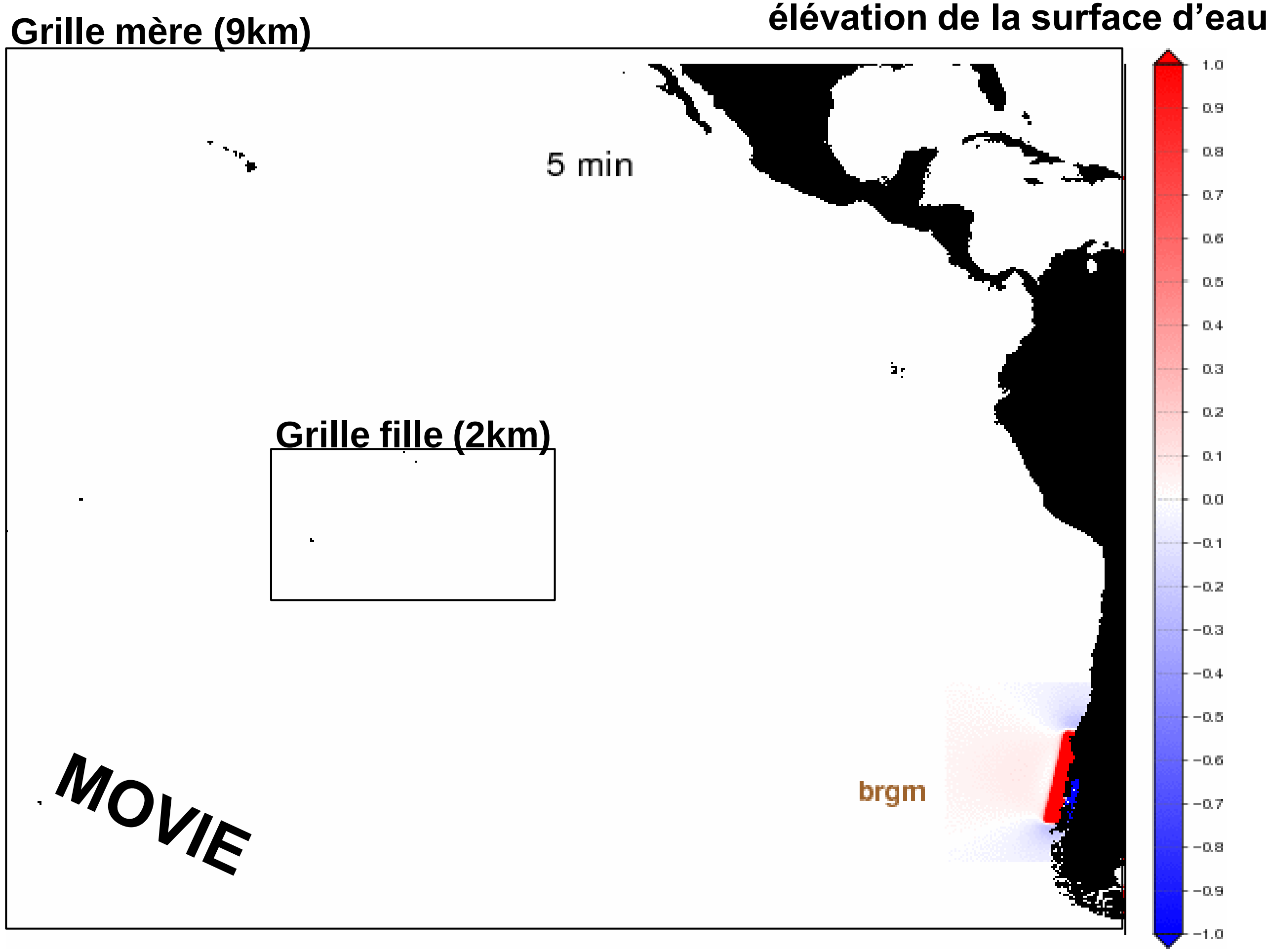
Schéma de propagation d'un tsunami depuis le milieu profond vers la côte

Source : http://www.prh.noaa.gov/itic/fr/library/pubs/great_waves/tsunami_great_waves.html

- **Longueur d'onde : km – centaines de km**
- **Période : quelques minutes – une heure**

Simulation of the tsunami generated in Chili in 1960

Generation (Okada 1985) and propagation (Funwave-TVD) up to French Polynesia



Recent observations due to tsunamis in France

Tsunamis observés en France

Source : BRGM / tsunamis.brgm.fr

- 1755 (1er novembre) : **Atlantique ; tsunami transocéanique lié au séisme dit « de Lisbonne » (Portugal).**
- 1867 (18 novembre) : **Mer des Caraïbes ; tsunami lié au séisme des Iles Vierges (Saint-Thomas, US Virgin Islands).**
- 1887 (23 février) : **Méditerranée ; tsunami lié au séisme de la côte ligure (Italie).**
- 1902 (8 mai) : **Mer des Caraïbes ; tsunami lié à l'éruption de la Montagne Pelée (Martinique).**
- 1979 (16 octobre) : **Méditerranée ; tsunami lié au glissement sous-marin dans la baie de Nice (France).**
- 2003 (21 mai) : **Méditerranée ; tsunami lié au séisme de Boumerdès (Algérie).**
- 2003 (13 juillet) : **Mer des Caraïbes ; tsunami lié à l'éruption de l'île Montserrat.**

Context and problem definition

In tsunami hazard assessment,

- **Major** source of **uncertainty** = **earthquake source** parameters
- Those parameters strongly influence the tsunami-induced sea surface elevation SSE

→ **Combining simulations with observations can improve knowledge and constrain those uncertainties**

Challenges:

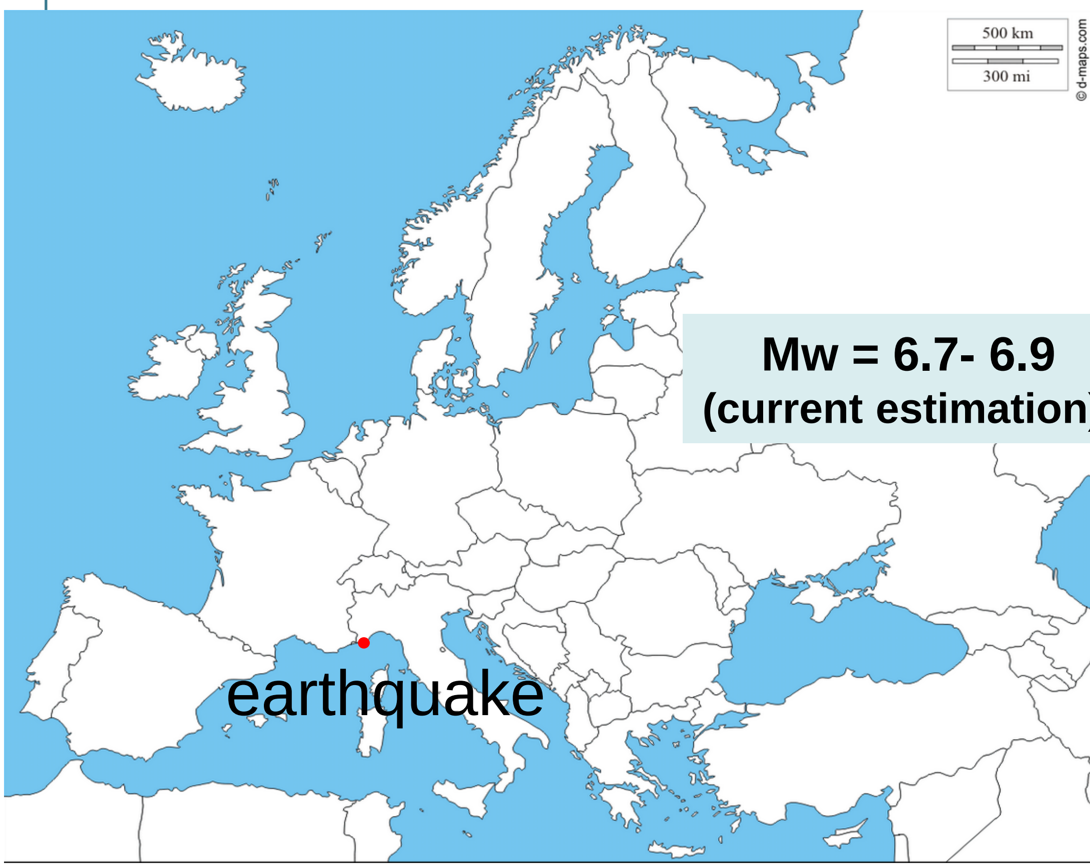
1. Observations on past events → **scarce and imprecise**
2. CPU time cost of one simulation → **several hours**

Tsunami Ligure 1887

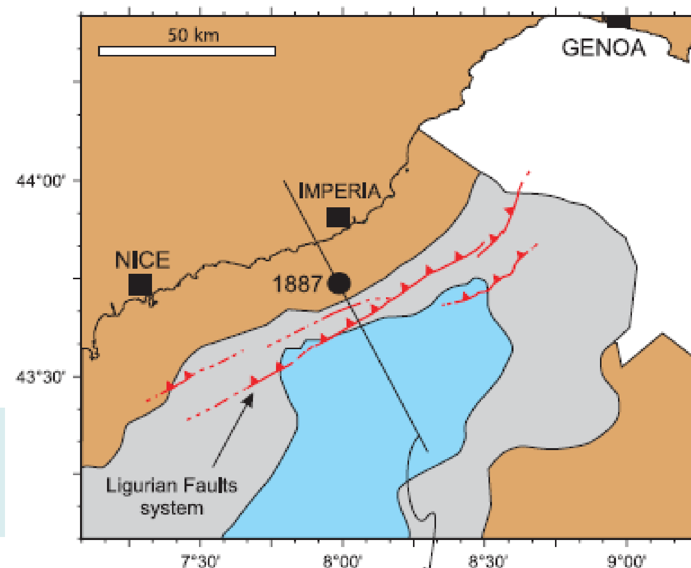
Motivating case study

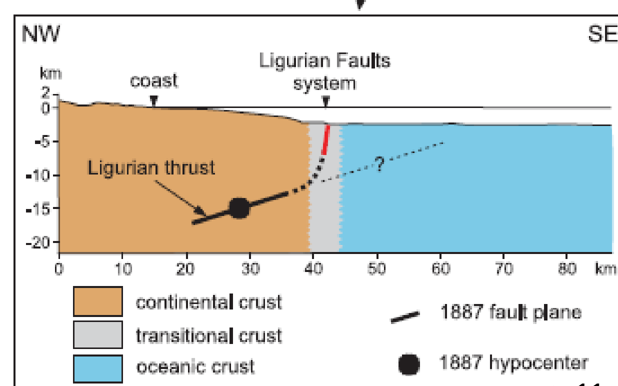
The 1887 Ligurian event

Damaging historical earthquake occuring at the junction between the southern French-Italian Alps and Ligurian basin induced tsunami



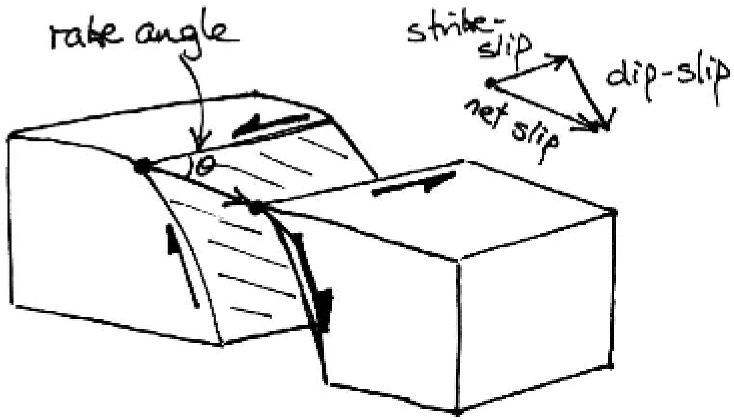
Mw = 6.7- 6.9
(current estimation)





Larroque et al. 2012

Uncertainty characterisation

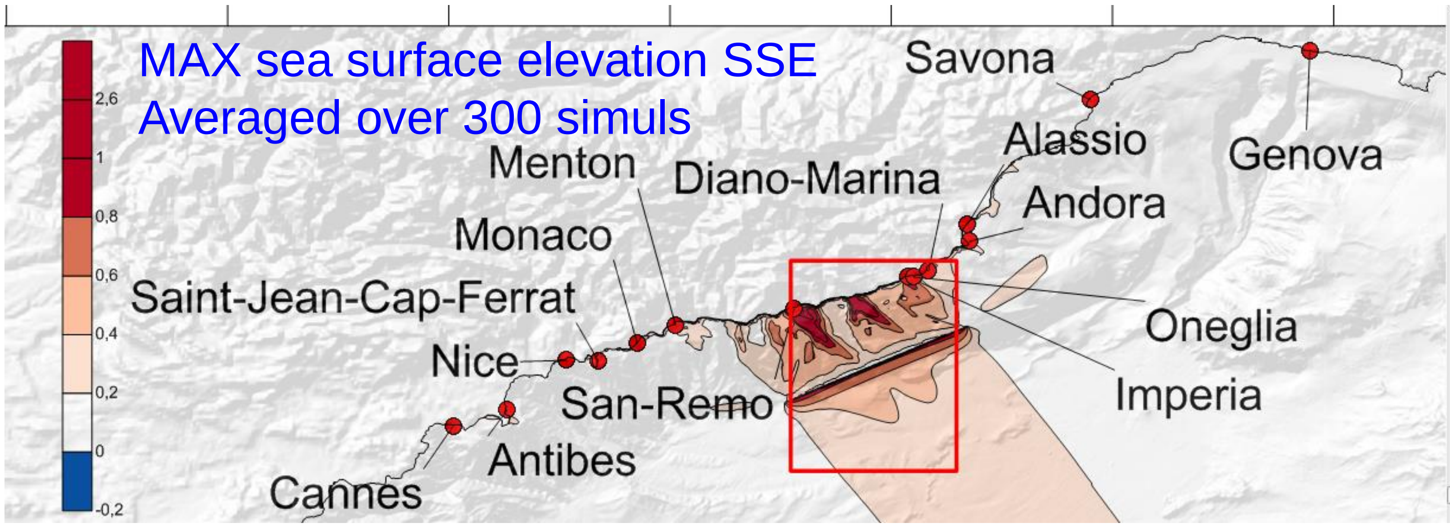


Location			orientation				geometry	
Lon (°)	Lat (°)	Depth Z (m)	Azim (°) 2 scenarios North **(South)** dipping	DIP (°)	Rake (°)	Average displacement D (m)	Length (m)	Width (km)
7.78	43.45	5	220 **(40)**	10	70	0,3	20	10
8.15	43.92	20	260 **(80)**	70	110	2	57	22

Based on Larroque et al. (2012) and Ioualalen et al. (2014)

Some illustrative results

- 300 configurations were randomly generated (latin hypercube sampling and uniform distribution).



MODEL CHARACTERISTICS

- 2,650,000 mesh cells
- CPU time of a single simulation ***30 min*** with ***4*24 cores***
- Funwave_TVD Boussinesq ocean surface wave propagation + Okada for initial deformation

Historical observations

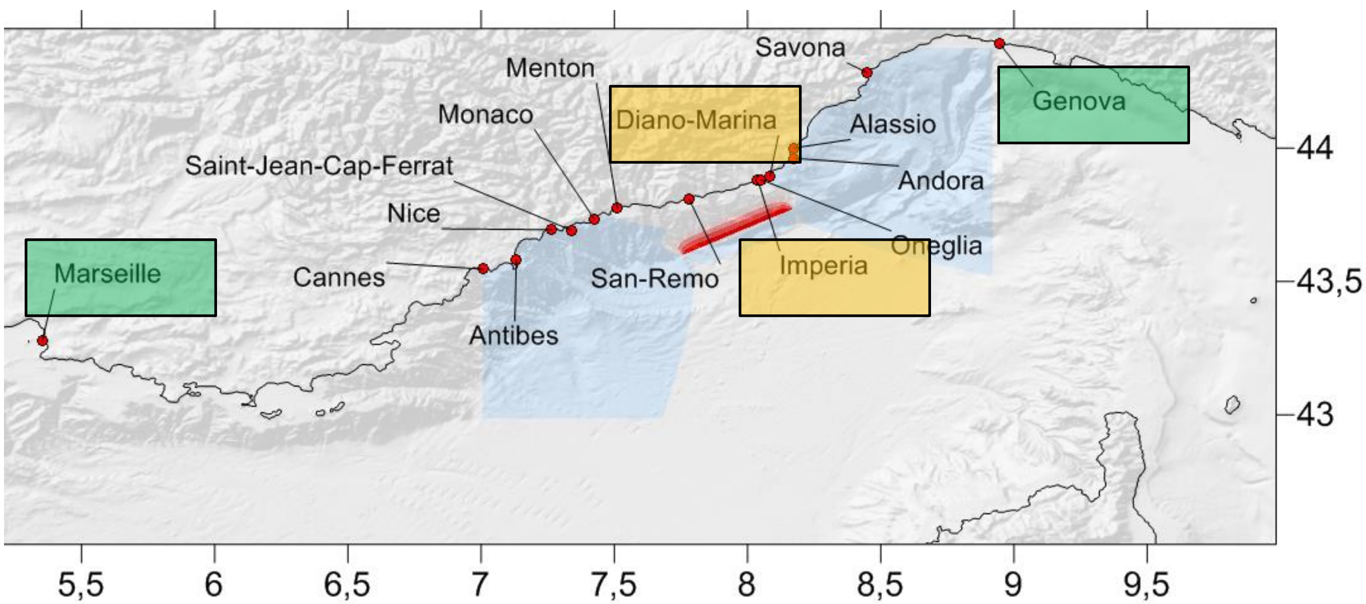
Location	Observation	Value	Confidence
Marseille	Tsunami intensity	0	High
Diano Marina	Maximum SSE	1 m	Moderate
Diano Marina	Minimum SSE	-1m	Moderate
Genova	Maximum SSE	0.07 m	High
Porto Maurizio (Imperia)	Minimum SSE	-1m	Moderate

Based on Lambert (2016) http://tsunamis.brgm.fr/

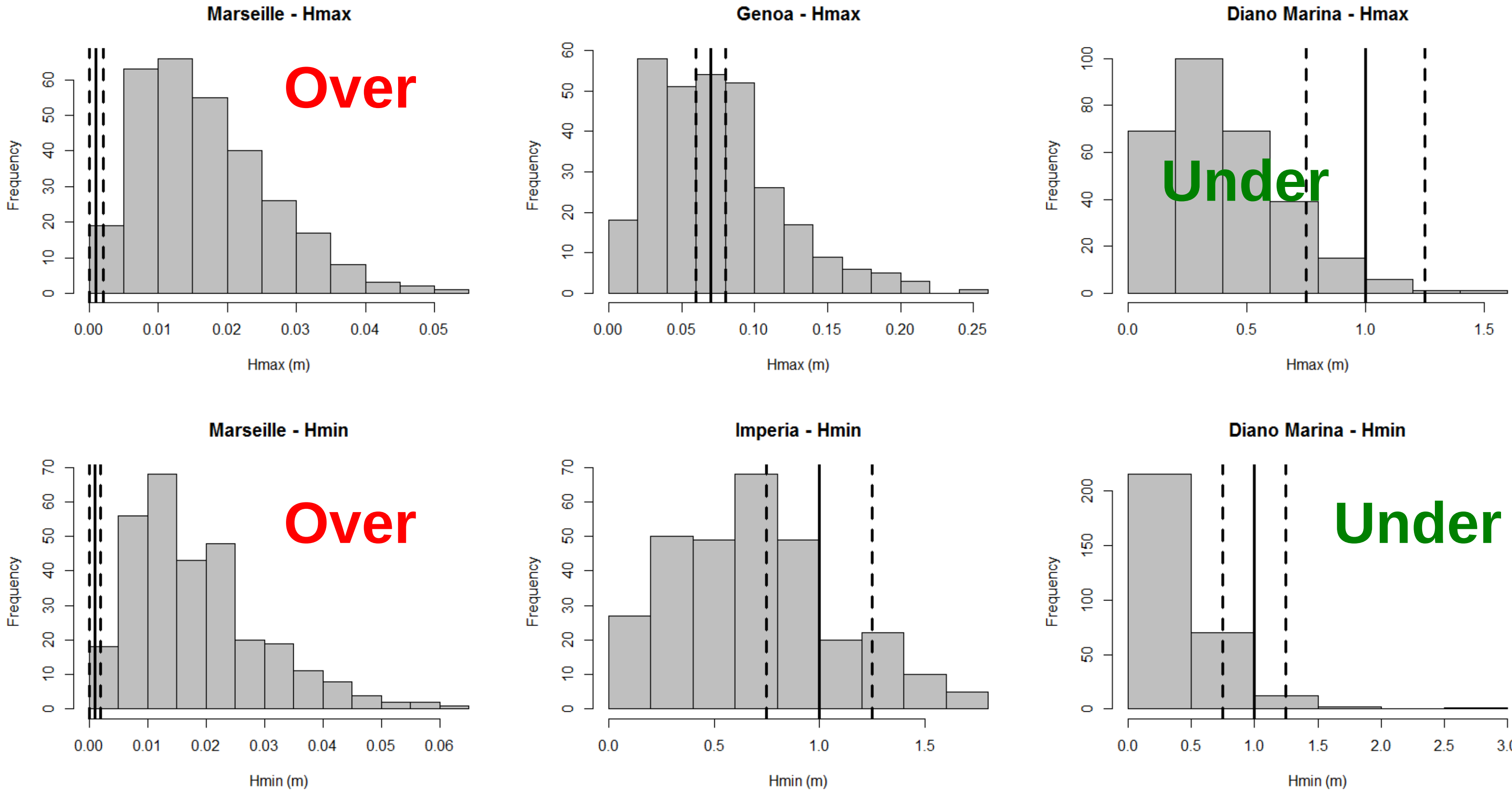
Confidence

High

Moderate



Simulation results (a priori)



Observation

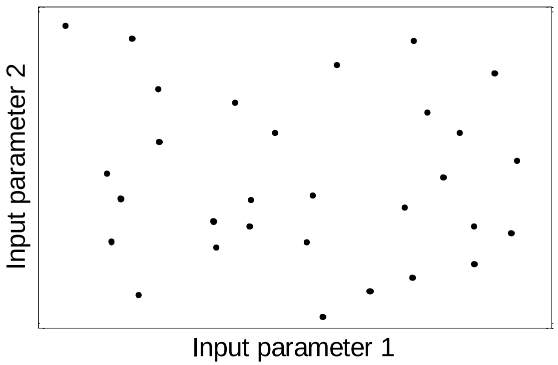
Assumed range of uncertainty

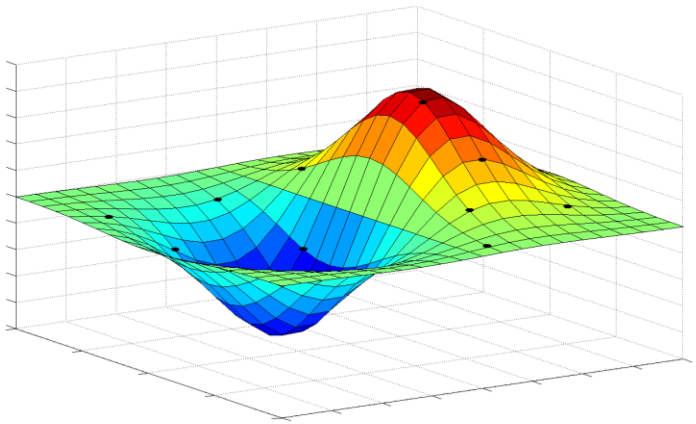
Metamodelling principles

Meta-modelling principles

Basic idea:

replace the computationally intensive G model by an approximation g (meta-model)





Points to perform simulation
Very limited set
(300)

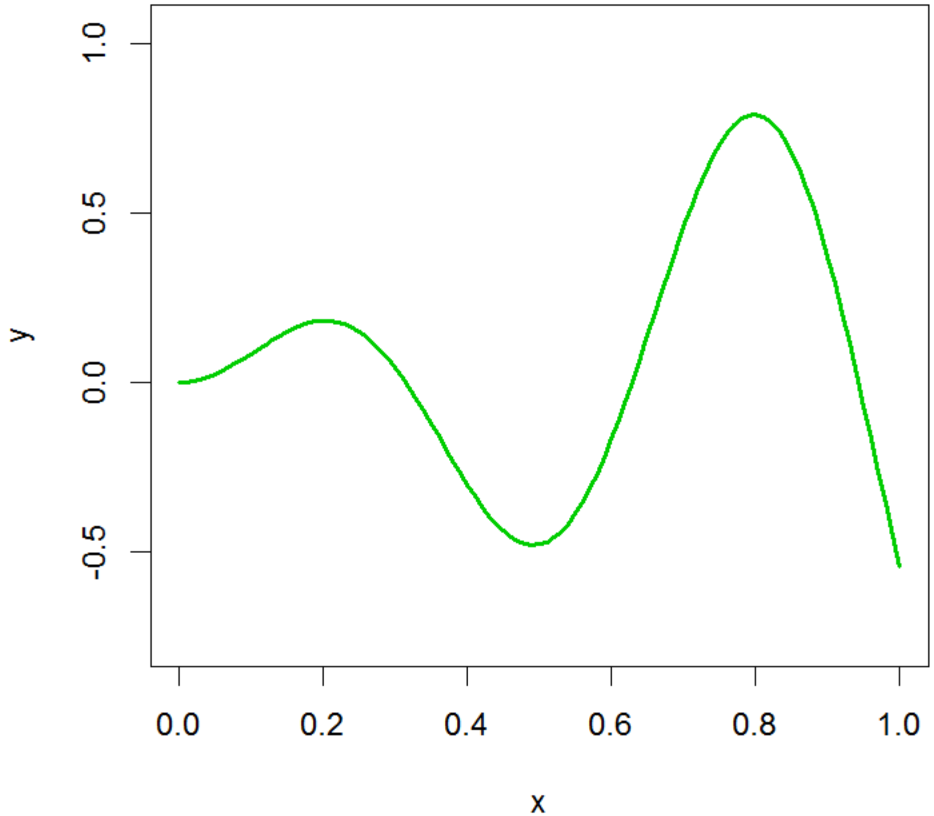
→ **SIMULATION** →

Meta-model
Approximation g

Different meta-models can be used; choice: **Kriging**

Kriging-based meta-modelling

> Consider ***f* the unknown function** *Hs*=*f*(***x***) with *Hs* the response, ***x*** the uncertain cyclone characteristic;



Kriging-based meta-modelling

> **Without any knowledge** (no simulation), assume:

$$H_S(\boldsymbol{x}) = m + Z(\boldsymbol{x})$$

with m the **mean**; *Z*(.) is a centered stationary **gaussian** process *GP* with C the **covariance function**:

$$\mathrm{C}(\boldsymbol{x}^{(1)};\boldsymbol{x}^{(2)}) = \sigma^2\,\mathrm{R}(\boldsymbol{x}^{(1)} - \boldsymbol{x}^{(2)})$$

with σ^2 the process total variance and *R* the correlation function

e.g., [Forrester et al. 2008]

Kriging-based meta-modelling

> **Without any knowledge** (<u>no simulation</u>), assume:

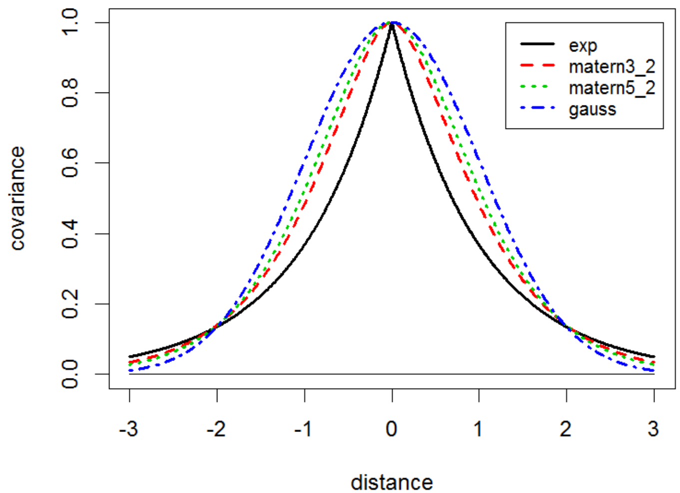
$$H_S(\boldsymbol{x}) = m + Z(\boldsymbol{x})$$

with m the **mean**; *Z*(.) is a centered stationary **gaussian** process *GP* with C the **covariance function**:

$$\mathrm{C}(\boldsymbol{x}^{(1)};\boldsymbol{x}^{(2)}) = \sigma^2\,\mathrm{R}(\boldsymbol{x}^{(1)} - \boldsymbol{x}^{(2)})$$

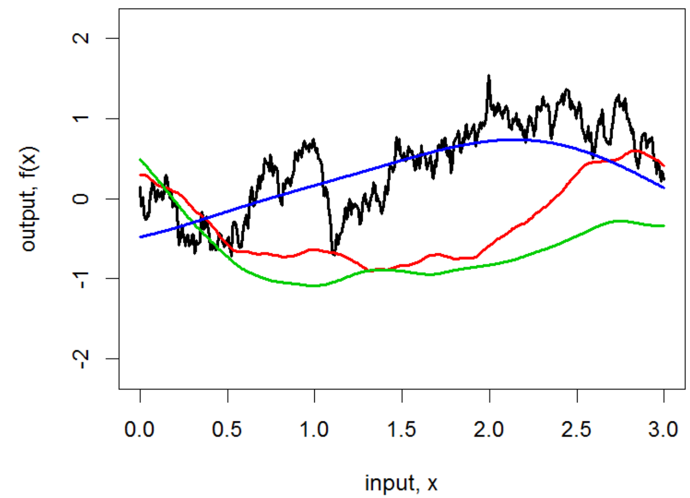
with σ^2 the process total variance and *R* the correlation function

Covariance function



σ^2=1; lengthscale θ =1

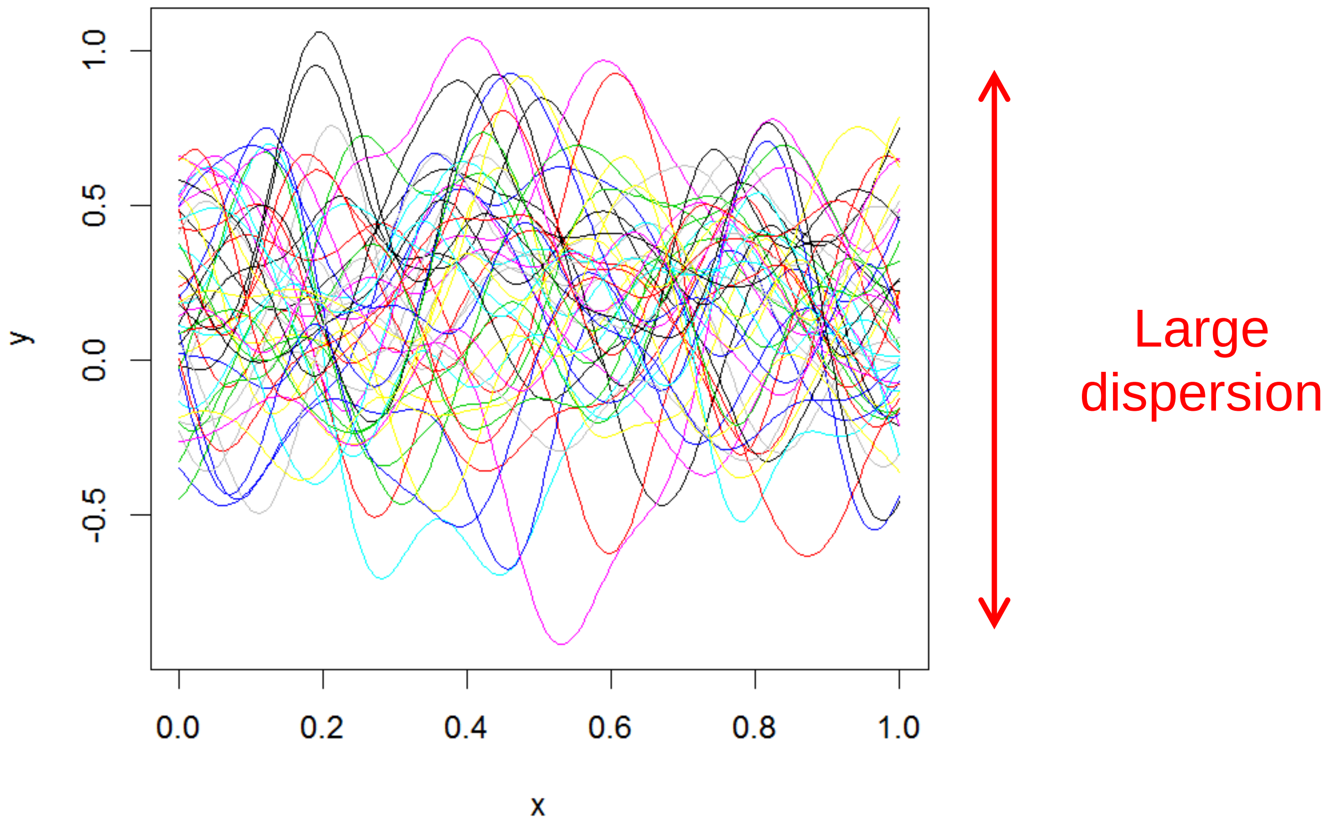
Example of 1 random sample



e.g., [Forrester et al. 2008]

Kriging-based meta-modelling

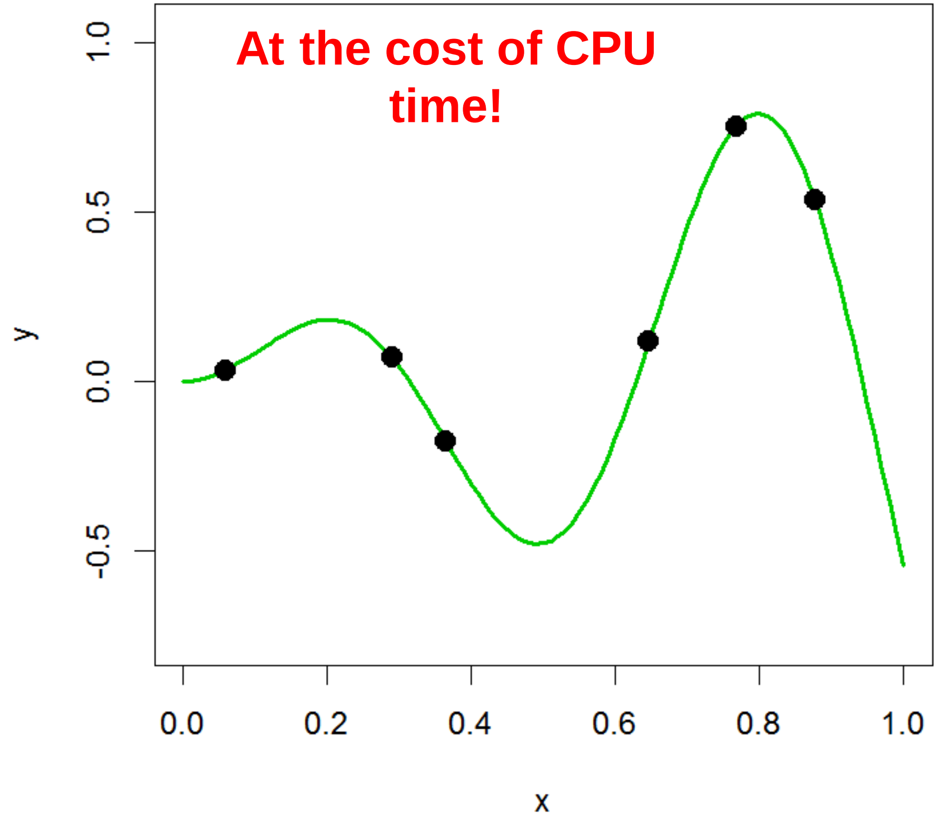
> **A priori knowledge**, 50 random samples



Kriging-based meta-modelling

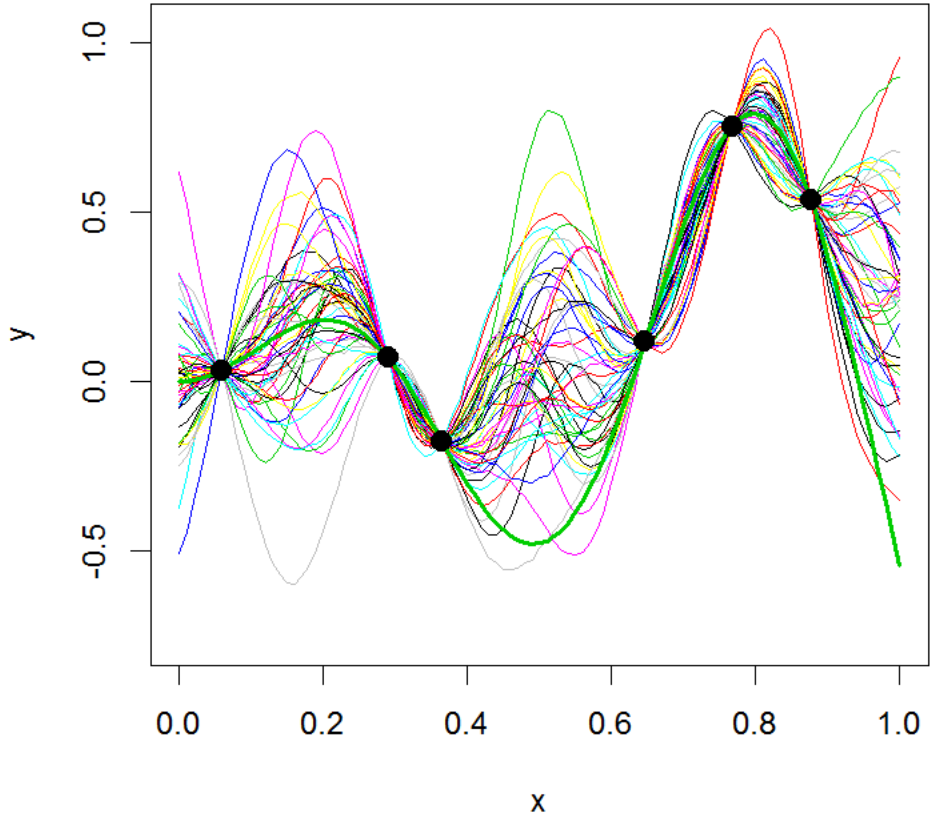
> Let us run **6 simulation scenarios** ***X***;

> H_D =set of model results (= **knowledge** on the unknown simulator).



Kriging-based meta-modelling

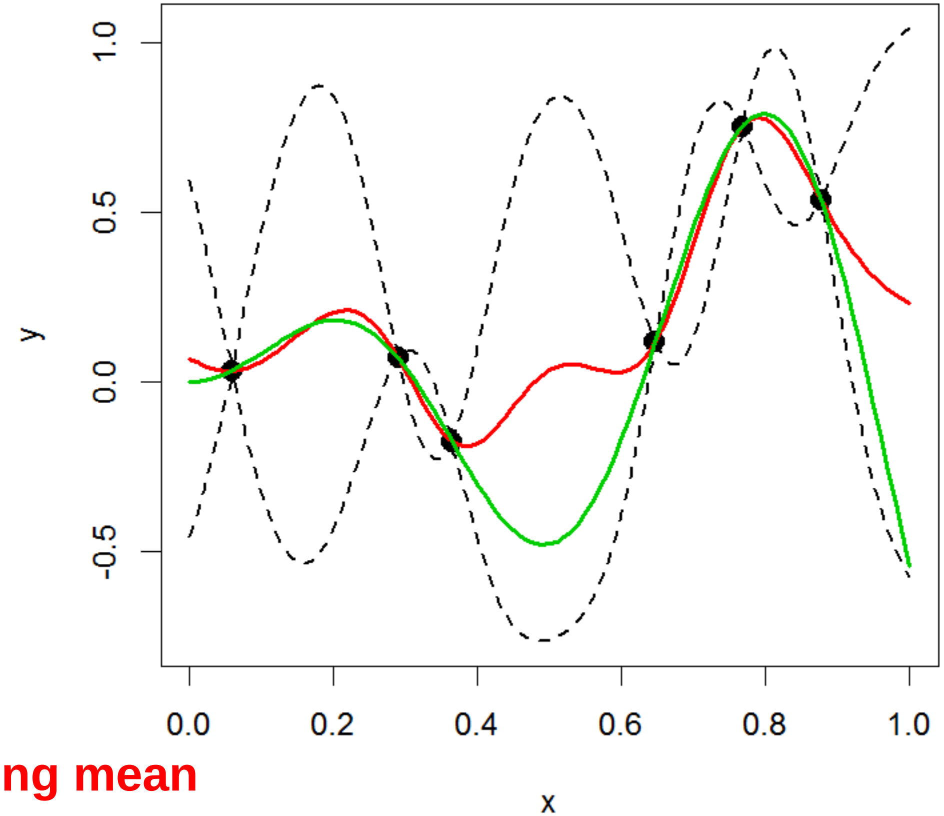
- Update the **gaussian process;**
- New mean m^* and new covariance funct. C^* **conditional on H_D**.



m=0.19, R gaussian; lengthscale θ=0.08; σ^2=0.01 ➔ **via Maximum Likelihood Estimation**

Kriging-based meta-modelling

- Update the **gaussian process;**
- New mean m^* and new covariance funct. C^* **conditional on H_D**



Kriging mean

95% confidence enveloppe (linked to the kriging variance)

m=0.19, R gaussian; lengthscale θ=0.08; σ^2=0.01 ➔ **via Maximum Likelihood Estimation**

Strategy

Proposed Strategy

> Simulate tsunami-induced sea surface elevation H (max / min) using 300 different configurations of earthquake source via Latin Hypercube Sampling

- Use of Promethee environment (IRSN, Y. Richet)
- lhs R package

Proposed Strategy

> Simulate tsunami-induced wave height H (max / min) using 300 different configurations of earthquake source via Latin Hypercube Sampling

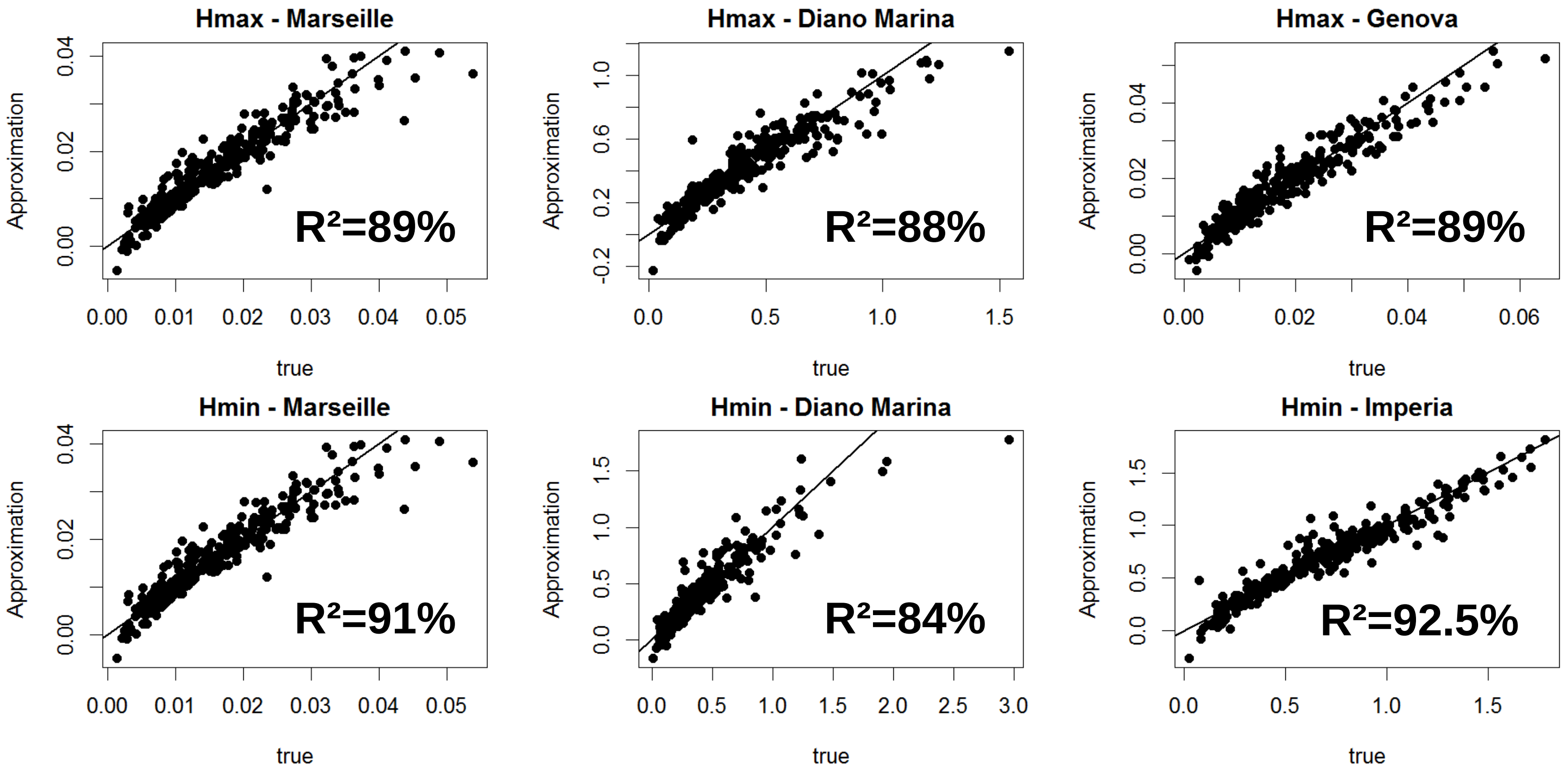
> Build a 9**-dimensional kriging meta-model**

- DiceKriging R package (Roustant et al., 2012)
- Matern covariance, linear trend, no nugget, MLE estimation

Proposed Strategy

- Simulate tsunami-induced wave height H (max / min) using 300 different configurations of earthquake source via Latin Hypercube Sampling
- Build a 9**-dimensional kriging meta-model**
- **Validate** via LOOCV (Leave-One-Out-Cross-Validation, compute R^2)

Validation of kriging metamodel – N. dipping model



Satisfactory approximation quality (typical threshold at 80%, see e.g., Marrel et al., (2009); Storlie et al. (2009)

Proposed Strategy

> Simulate tsunami-induced sea surface elevation H (max / min) using 300 different configurations of earthquake source θ via Latin Hypercube Sampling

> Build a 9**-dimensional kriging meta-model**

> **Validate** via LOOCV (Leave-One-Out-Cross-Validation, compute R^2)

> Perform **Bayesian Inference** using **ABC methods** (e.g. Marin et al., 2012)

- 1 million of simulations using Sobol' sequence (randtoolbox R package)
- Abc & abctools R package

$$\boxed{H_{observed} = H_{simulated} + \theta_{bias} \approx \hat{H}(\theta_{Source}) + \theta_{bias}}$$

Bayesian inference

Bayesian inference - principles

$$p(\theta|H_{obs}) \propto p(\theta) \cdot p(H_{obs}|\theta)$$

Update of probability information on θ after observing H

Prior knowledge

Likelihood = probability of observing H given θ

Likelihood function usually hard to compute

Approximate Bayesian Computation ABC

1. Draw θ from the prior : $\theta \sim p(\theta)$;

2. Simulate a realisation from the model: $H(\theta) \sim p(H|\theta)$;

3. Accept θ if and only if $\rho(S(H_{obs}), S(H(\theta))) \leq \varepsilon$.

Wilkinson (2015)

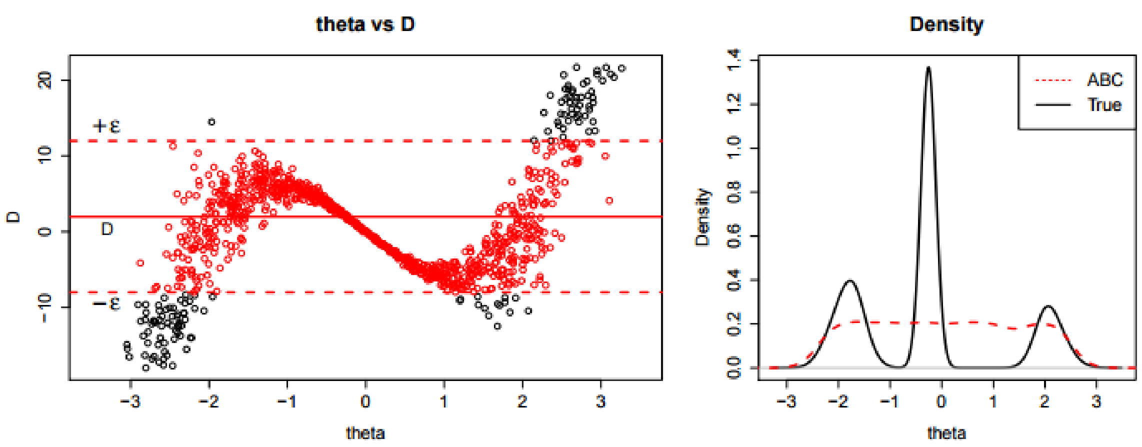
Approximate Bayesian Computation ABC

1. Draw θ from the prior : $\theta \sim p(\theta)$;

2. Simulate a realisation from the model: $H(\theta) \sim p(H|\theta)$;

3. Accept θ if and only if $\rho(S(H_{obs}), S(H(\theta))) \leq \varepsilon$.

$\epsilon = 10$



$$\theta \sim U[-10, 10], \qquad X \sim N(2(\theta+2)\theta(\theta-2), 0.1+\theta^2)$$

$$\rho(D, X) = |D - X|, \qquad D = 2$$

Wilkinson (2015)

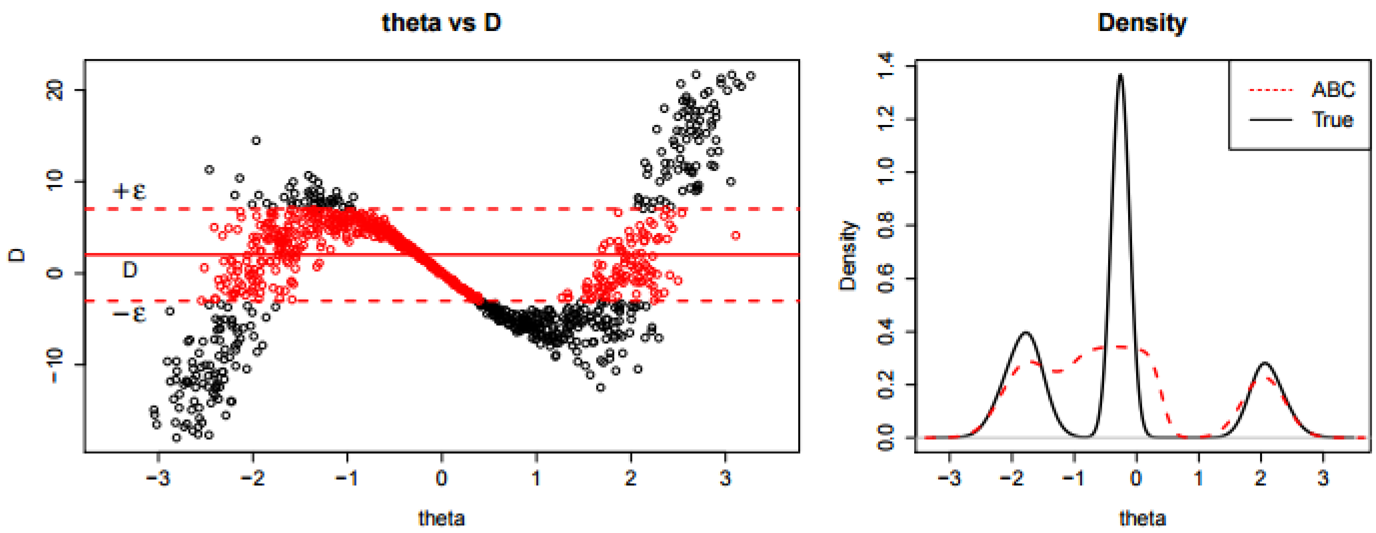
Approximate Bayesian Computation ABC

1. Draw θ from the prior : $\theta \sim p(\theta)$;

2. Simulate a realisation from the model: $H(\theta) \sim p(H|\theta)$;

3. Accept θ if and only if $\rho(S(H_{obs}), S(H(\theta))) \leq \varepsilon$.

ε=5



Wilkinson (2015)

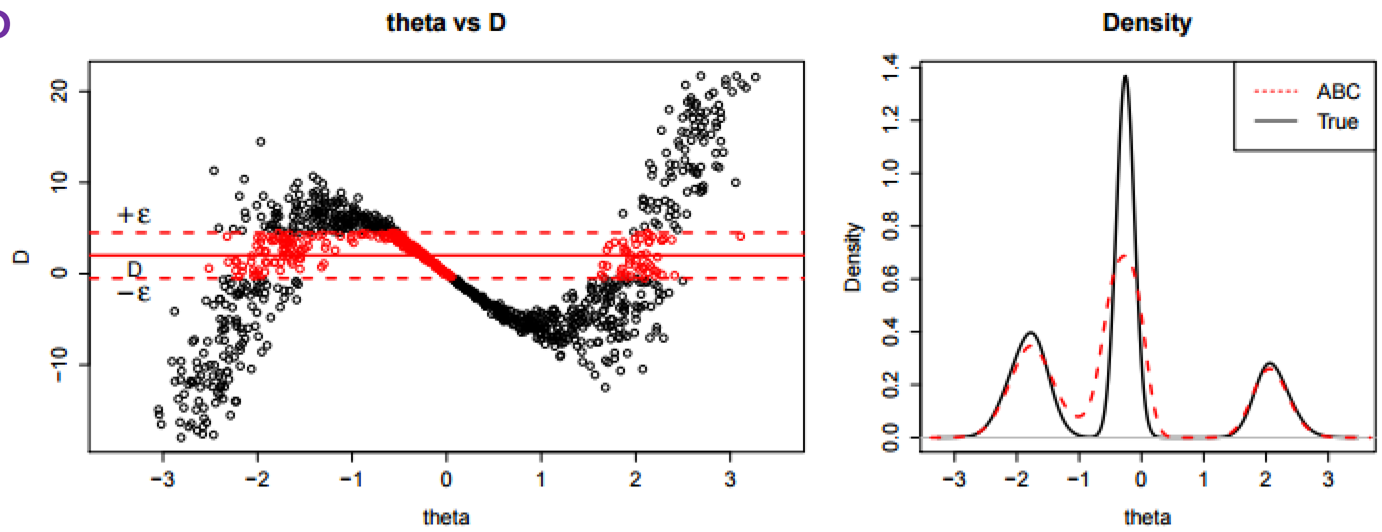
Approximate Bayesian Computation ABC

1. Draw θ from the prior : $\theta \sim p(\theta)$;

2. Simulate a realisation from the model: $H(\theta) \sim p(H|\theta)$;

3. Accept θ if and only if $\rho(S(H_{obs}), S(H(\theta))) \leq \varepsilon$.

ε=2.5



Wilkinson (2015)

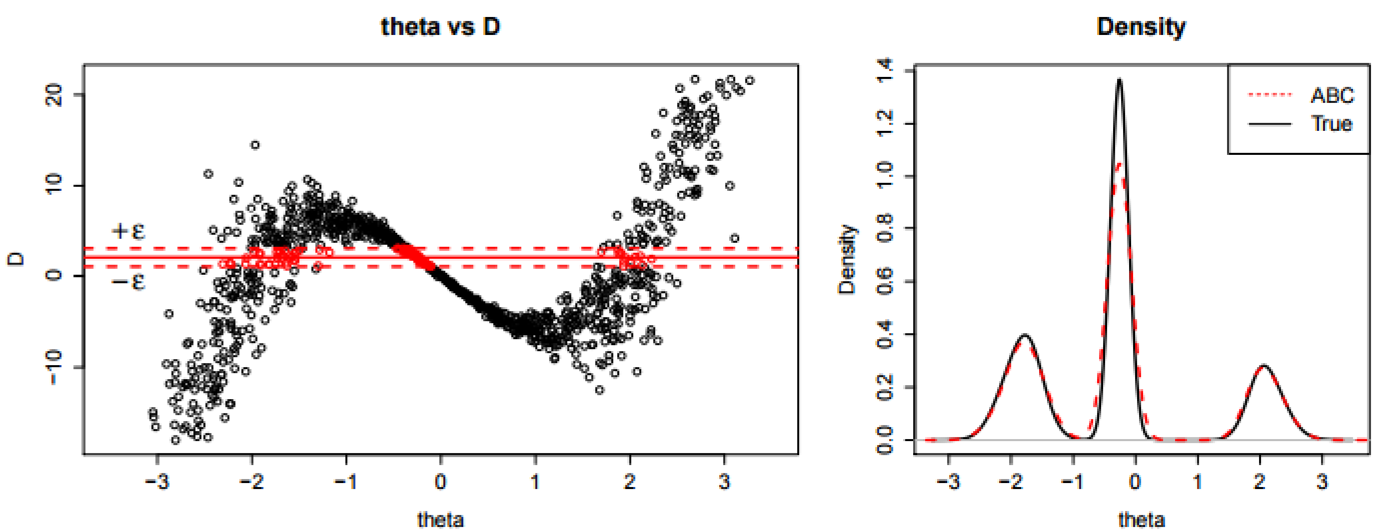
Approximate Bayesian Computation ABC

1. Draw θ from the prior : $\theta \sim p(\theta)$;

2. Simulate a realisation from the model: $H(\theta) \sim p(H|\theta)$;

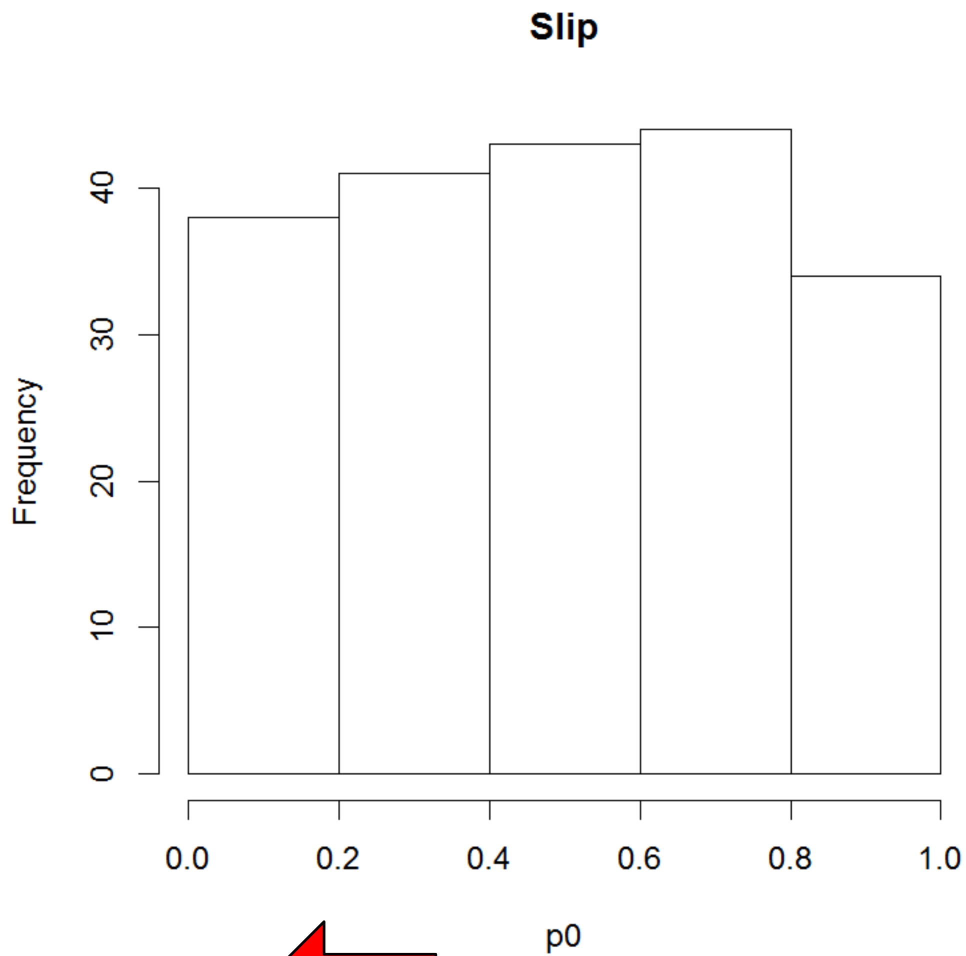
3. Accept θ if and only if $\rho(S(H_{obs}), S(H(\theta))) \leq \varepsilon$.

ε=1



Wilkinson (2015)

Diagnostic of ABC - example

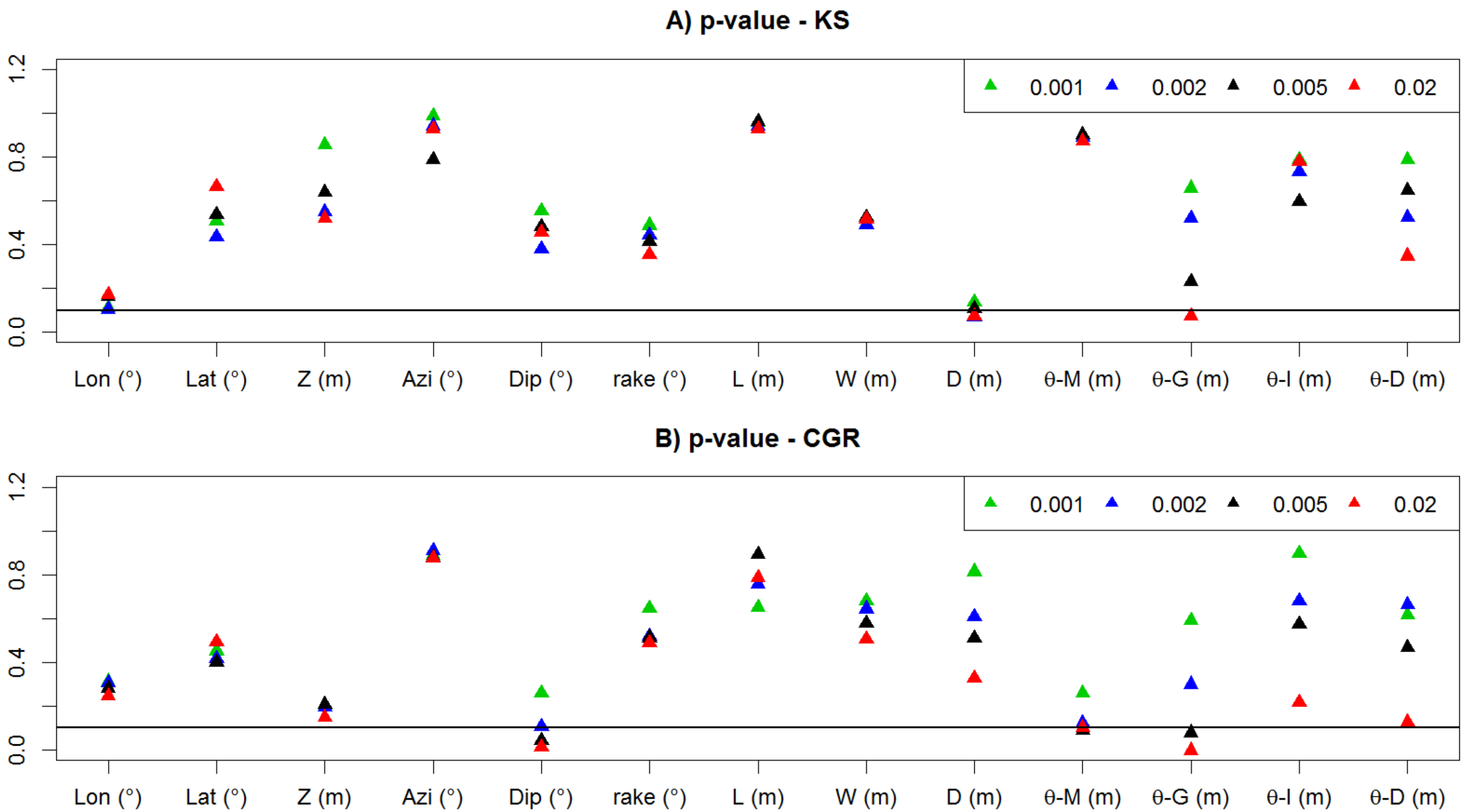


Distribution should be uniform

Prangle et al. (2014)

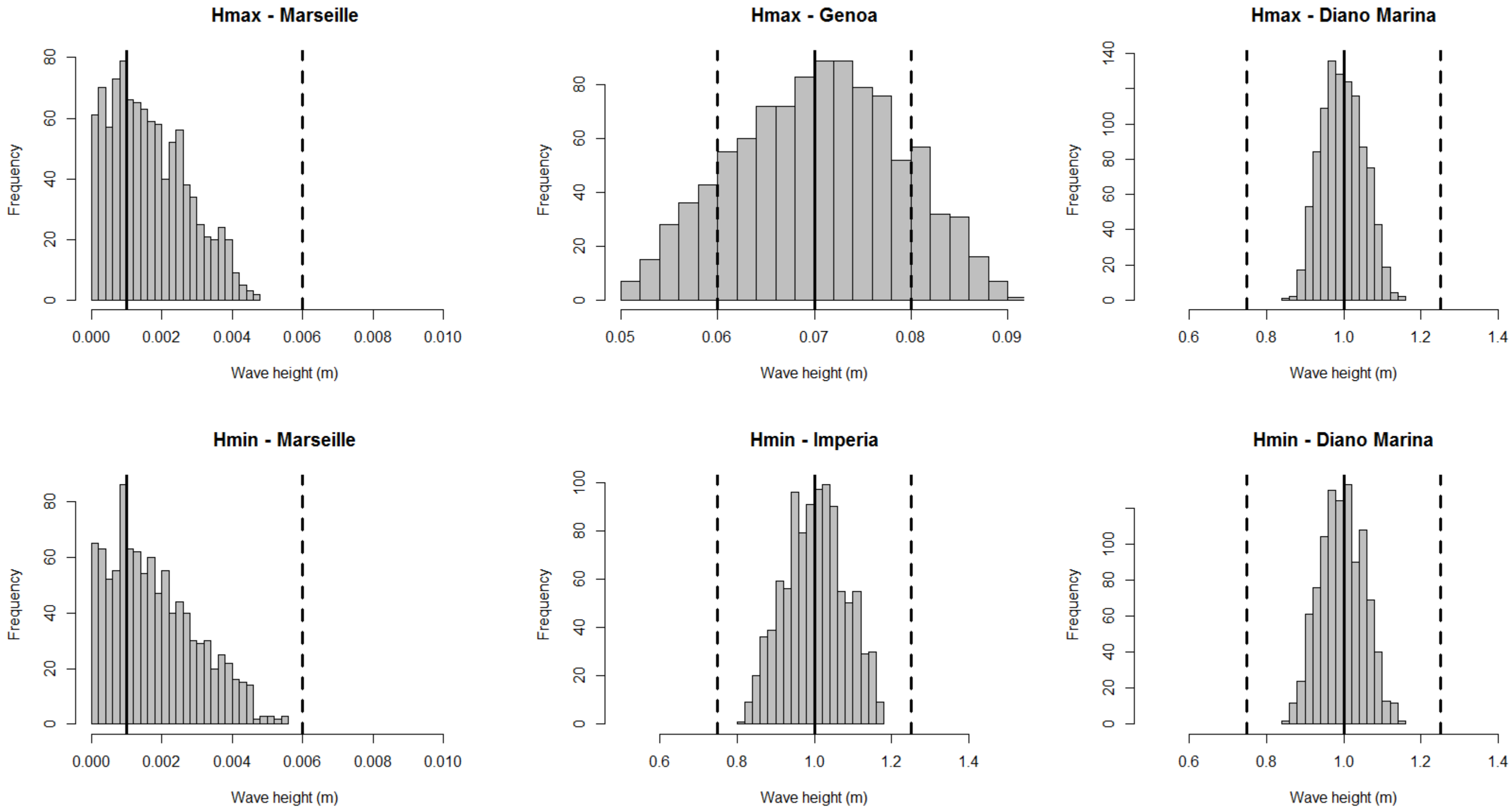
Choosing the proportion of accepted samples

Initial number of random samples = 1 million



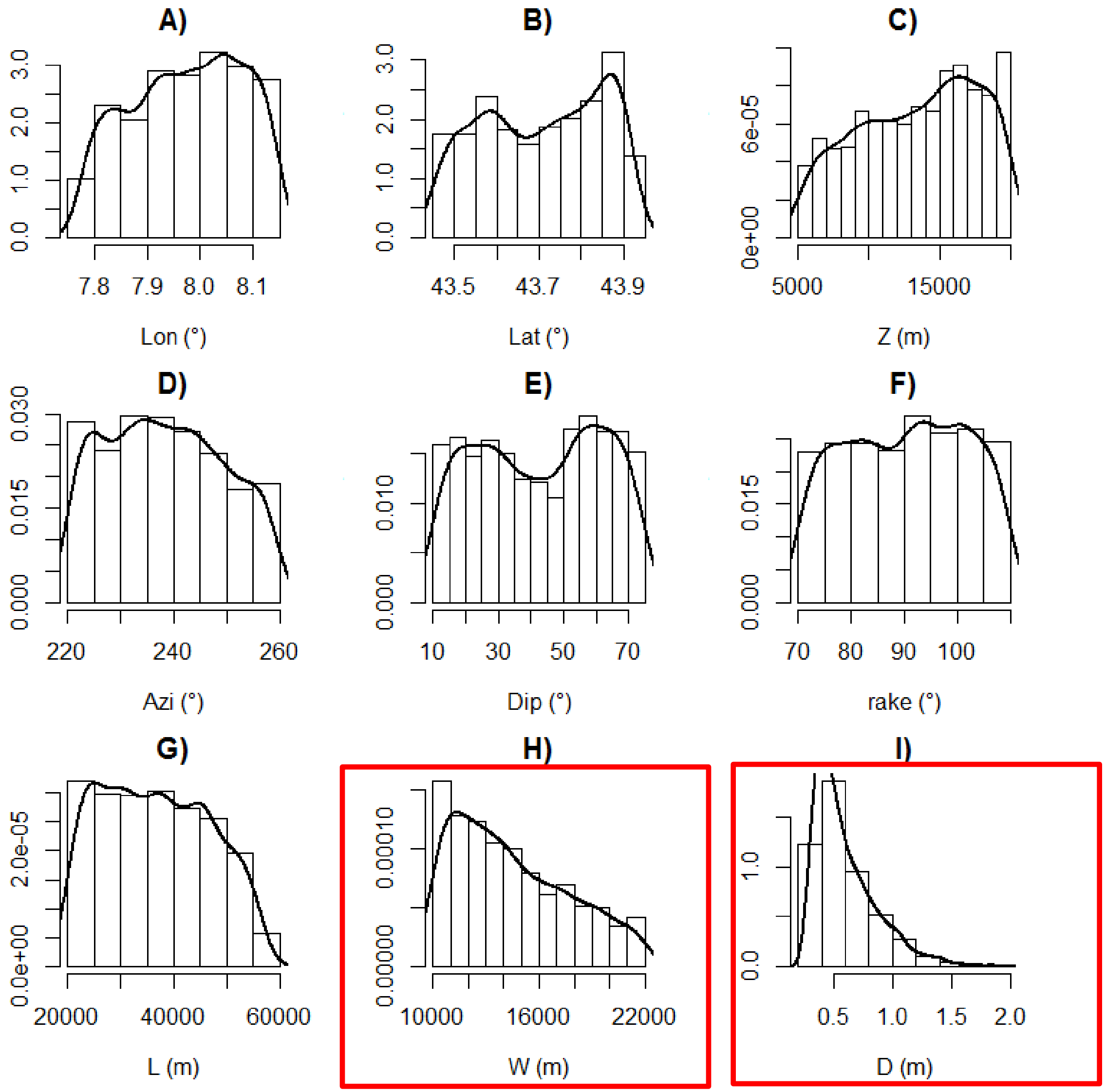
Verifying the coverage property (Prangle et al. 2014)

A Posteriori distributions

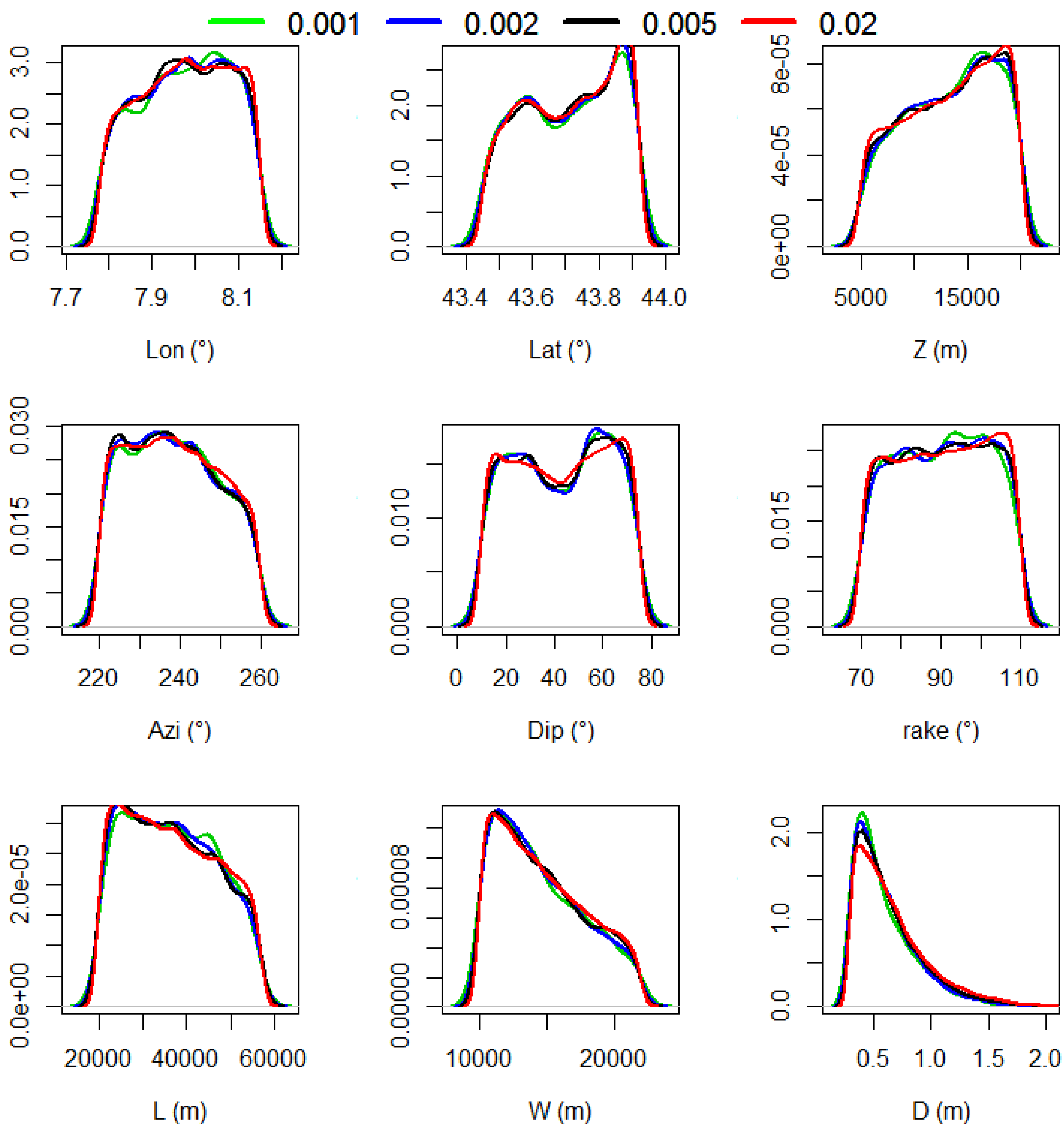


1,000 accepted random samples

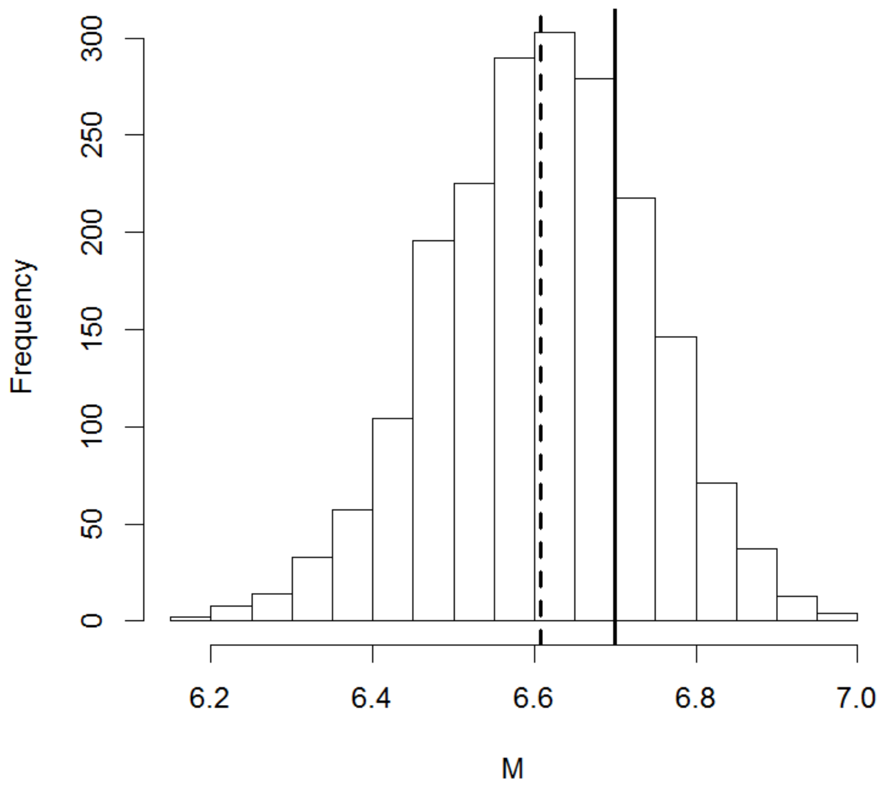
Inference of source parameters



Influence of ABC tuning

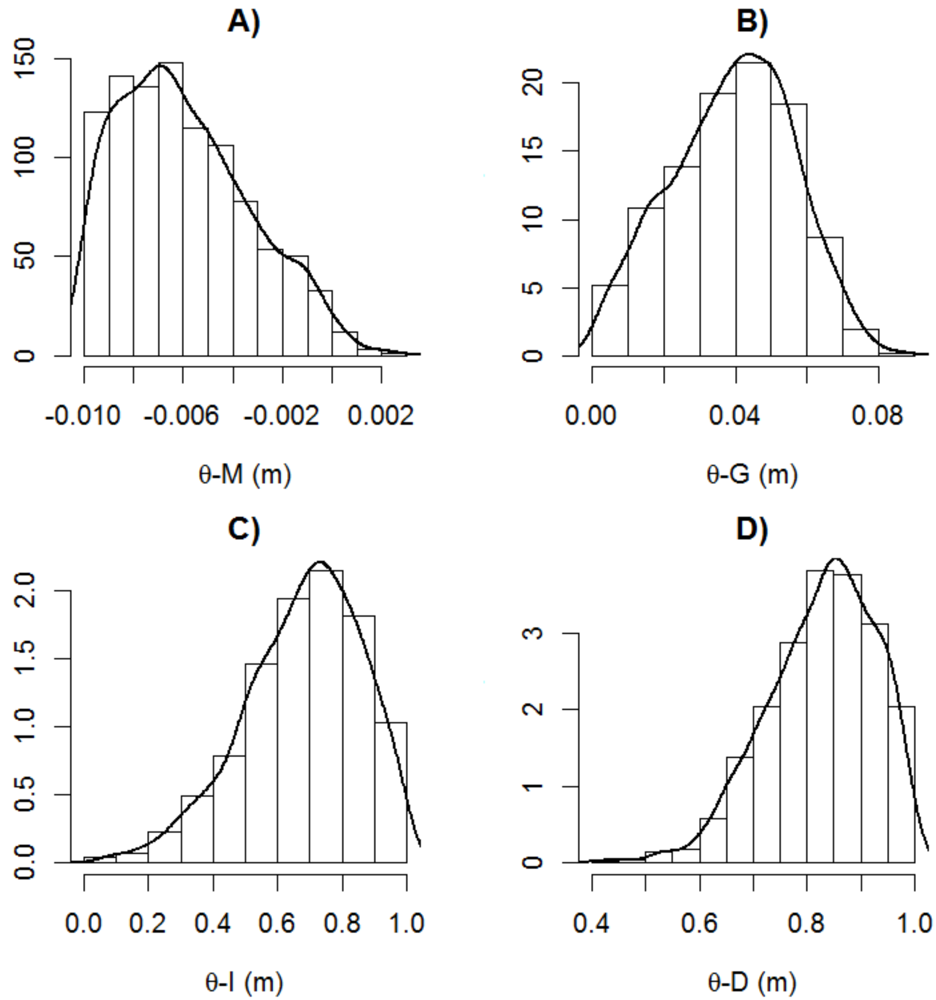


Inference of magnitude - North Dipping model



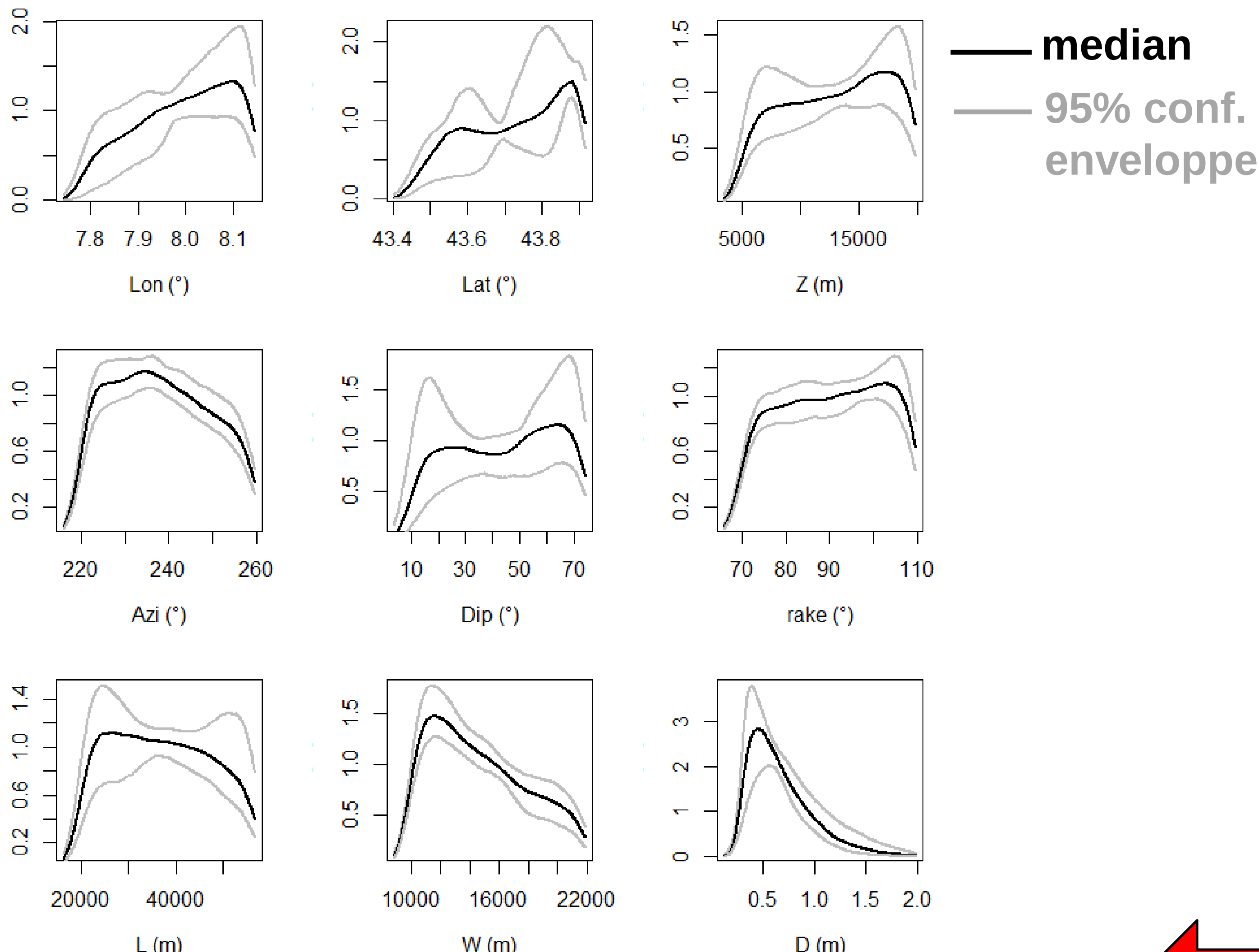
In agreement with past studies: 6.7-6.9 (Ioualalen et al. (2014)

Inference of bias parameters



$$H_{observed} = H_{simulated} + \theta_{bias} \approx \hat{H}(\theta_{Source}) + \theta_{bias}$$

Impact of observation imprecision



ABC-kriging procedure reconducted 1,000 times; observations are randomly varied

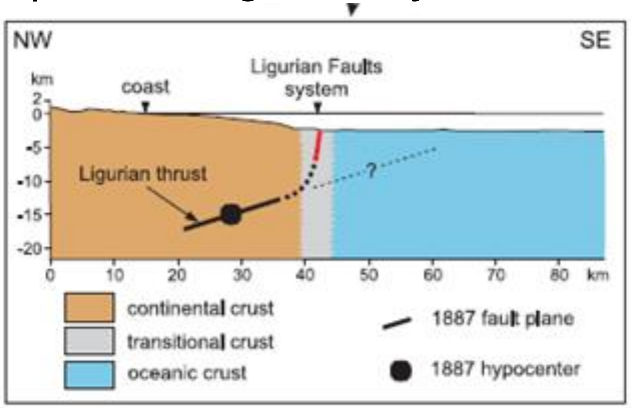
Summary

> **Methodology for combining imprecise observations and numerical simulations**

> **Emulation of the long running numerical simulator (meta-model)**

> **Further work**

- Inferring assumption scenarios i.e. complex fault geometry
- Impact of metamodel error



Ioualalen et al. (2014

THANK YOU

Géosciences pour une Terre durable

brgm

Influence of meta-model error