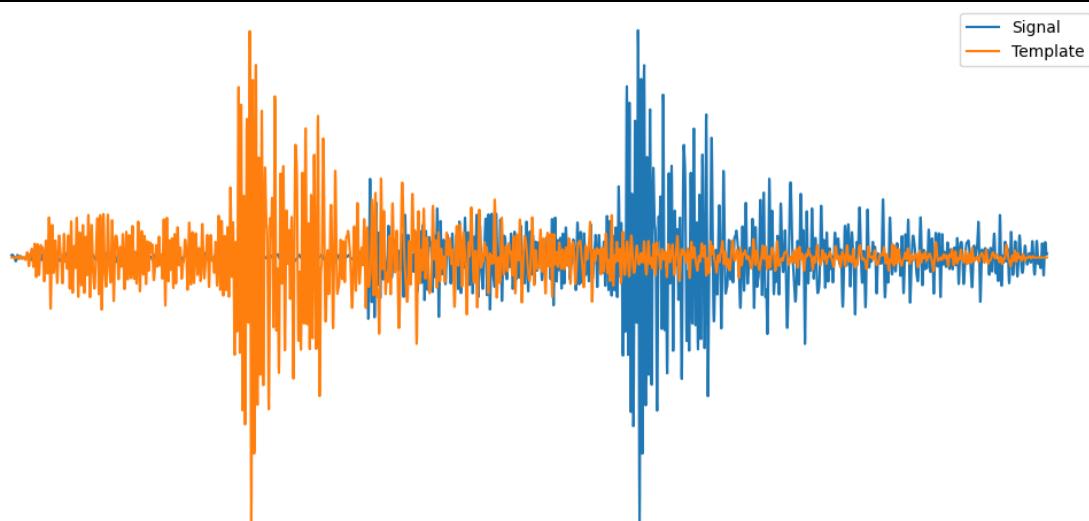


Année Universitaire : 2025–2026

Encadrant(s) : Claudio Satriano, Marie-Paule Bouin

Équipe : Sismologie

# Détection de séismes répétés dans la subduction des Petites Antilles par template-matching



*Rapport de Projet de Synthèse (Semestre 9)  
V.A. Ondes et Signaux*

Élèves :

EV2 LE CAM Louis – [le\\_cam3.eleve@ecole-navale.fr](mailto:le_cam3.eleve@ecole-navale.fr)

EV2 YVARD Marius – [yvard.eleve@ecole-navale.fr](mailto:yvard.eleve@ecole-navale.fr)

# Remerciements

Nous remercions tout d'abord notre maître de stage Claudio Satriano pour son aide, son écoute, et pour tout le temps passé à nous en apprendre plus sur la sismologie dans les Antilles ou ailleurs dans le monde.

Nous remercions ensuite Marie-Paule Bouin, pour le partage de son expérience et pour ses conseils avisés, pour mettre notre sujet en perspective.

De plus, nous remercions Rozenn Desmare, pour sa confiance et son aide dans notre recherche de stage

Enfin, nous remercierons l'EV2 Crepelle, pour ses conseils, sa relecture et sa bonne humeur.



Observatoire volcanologique et sismologique de la Martinique (OVSM), rattaché à l'IPGP

## Présentation de l'organisme d'accueil

L'[IPGP](#) (*Institut de Physique du Globe de Paris*), créé en 1921, est un établissement de recherche et d'enseignement supérieur de 500 chercheurs internationaux et de recherche consacré à l'étude de la Terre et des planètes ou exoplanètes *depuis le noyau jusqu'aux enveloppes fluides les plus superficielles, à travers l'observation, l'expérimentation et la modélisation [sic]*. L'équipe de sismologie compte 38 membres dont 16 chercheurs et enseignants-chercheurs permanents. De nombreuses stations sismiques gérées par l'IPGP partout sur le globe permettent d'approfondir les connaissances dans diverses branches de la sismologie qu'elles soient conventionnelles (étude des sources sismiques, des zones à risque, des dynamiques tectoniques, ou des bruits anthropiques) ou très récentes (sismologie rotationnelle, ou optique), tout en testant de nouveaux capteurs de pointe, améliorés en permanence, et en diffusant des données de qualité en libre accès. Notre travail consistera à éclaircir le comportement de la zone de subduction des Antilles, afin de mieux cerner les risques de cette région de 4 millions d'habitants (2009), en s'appuyant sur des stations de plusieurs générations, implantées sur plusieurs pays.



Site historique et principal de l'IPGP (*Jussieu - Paris V<sup>e</sup>*)

## Fiche de résumé du projet :

Élèves : EV2 LE CAM Louis, EV2 YVARD Marius, EN23  
Filière : Ondes et Signaux  
Organisme d'accueil : IPGP  
Chef de projet : Claudio SATRIANO  
Président de Jury : Jean-Marc LE CAILLEC

## Abstract

The Lesser Antilles subduction mechanisms still remain partially unexplained : GNSS data, coupling models, and geophysical properties modeling do not coincide, nor constitute any consensus concerning current seismic hazard, nor any explanation for 19<sup>th</sup> - century megathrust earthquakes. This study focuses on a seismic swarm located East of Martinique, where *Doucet (2021)* identified families of repeating earthquakes (multiplets) that occurred between 2014 and 2021. We extend this analysis further back in time by taking advantage of the FDF seismic station, which provides continuous recordings since 1998—in contrast to the WI network stations that only began operation in 2014. The use of repeated earthquakes can contribute to improve the understanding of the region, in the vicinity of asperities producing similar seismic time-signals spread over long time-periods. MPOM, BIM and FDF stations are used to reconstitute and then enlarge earthquake families. An adapted template-matching research gives additional events over two decades to existing seismic catalogs, by exploring continuous seismic raw data. Magnitudes are then deduced from several reference families, and tectonic-displacement models are built. Finally, interpretation hypotheses and retrieved data are explored to better understand subduction mechanisms.

**Keywords :** Repeated earthquakes, Lesser Antilles, template-matching, cross-correlation, magnitude, subduction.

## Résumé

La zone de subduction des Petites Antilles est encore partiellement incomprise : Les données GNSS, modèles de couplages ou autres modélisations géophysiques ne forment pas de consensus sur l'état d'aléa sismique, ni sur l'explication de mégaséismes recensés au 19<sup>ème</sup> siècle. Cette étude se concentre sur un essaim de séismes à l'Est de la Martinique, où *Doucet(2021)* a identifié des familles de *repeaters* entre 2014 et 2021. Cette analyse sera prolongée plus tôt dans le temps, en exploitant la station FDF active depuis 1998, à la différence du réseau WI opérant depuis 2014 seulement. L'étude de séismes répétés peut contribuer à éclaircir le comportement local de la subduction, à travers la reconstitution du déplacement tectonique autour de certaines aspérités, dont la rupture successive génère des motifs temporels reconnaissables sur de longues périodes de temps. Les stations MPOM, BIM et FDF sont utilisées pour retrouver puis élargir les familles des Petites Antilles. Un calcul par *template matching* adapté complète les données issues des catalogues sismiques sur deux décennies, en explorant les données brutes. Les magnitudes sont calculées sur plusieurs familles de référence, et des modèles de déplacement tectonique locaux sont reconstitués. Enfin, plusieurs pistes d'interprétation sont explorées selon ces données afin d'éclairer le comportement de la subduction.

**Mots-clefs :** Séismes répétés, Petites Antilles, template-matching, corrélation croisée, magnitude, subduction.

# Table des matières

<b>Introduction</b>	<b>1</b>
<b>Première partie : Contexte et supports théoriques</b>	<b>2</b>
1.1 Contexte des Petites Antilles . . . . .	2
1.2 Utilisation des séismes répétés . . . . .	6
1.3 Le <i>template matching</i> . . . . .	6
1.4 Algorithmes de création des familles . . . . .	8
1.5 Magnitudes, moment sismique, et déplacement de plaque . . . . .	9
1.6 Estimation de la magnitude de moment, et de la <i>corner frequency</i> . . . . .	11
<b>Deuxième partie : Préparation et configuration des calculs</b>	<b>13</b>
2.1 Présentation et fonctionnalités de <i>Requake</i> , <i>SeisCat</i> et <i>SourceSpec</i> . . . . .	13
2.2 Préparation des codes de <i>template matching</i> . . . . .	14
2.3 Chaîne de calcul retenue pour la détection par <i>NCC</i> . . . . .	15
2.4 Paramétrages d'entrée du calcul de Template Matching . . . . .	18
2.5 Utilisation de fichiers <i>.conf</i> . . . . .	21
2.6 Stations et plages de temps utilisées . . . . .	21
<b>Troisième partie : Obtention et exploitation des familles de séismes répétés</b>	<b>23</b>
3.1 Phase I : Débuts avec MPOM . . . . .	23
3.2 Phase II : Premier élargissement avec BIM . . . . .	25
3.3 Phase III : Second élargissement avec FDF . . . . .	27
3.4 Obtention et exploitation des magnitudes . . . . .	31
<b>Discussion</b>	<b>34</b>
4.1 Méthode de constitution des familles et implications . . . . .	34
4.2 Discussion sur les <i>antirepeaters</i> . . . . .	38
<b>Conclusion</b>	<b>39</b>
<b>Glossaire</b>	<b>40</b>
<b>Acronymes</b>	<b>41</b>
<b>Bibliographie</b>	<b>43</b>
<b>Annexes</b>	<b>44</b>

# Table des figures

1	Vue d'ensemble des Antilles . . . . .	2
2	Arc des Petites Antilles . . . . .	2
3	Cartes du réseau W.I. et des stations présentes en Martinique . . . . .	3
4	Schéma d'une coupe de subduction . . . . .	3
5	Activité sismique aux Petites Antilles . . . . .	4
6	Schéma d'une subduction à couplage nul . . . . .	5
7	Schéma d'une subduction à couplage non nul . . . . .	5
8	Schéma introduisant la formation de séismes répétés . . . . .	6
9	Figure d'un pic de cross-corrélation . . . . .	7
10	Schéma d'une hiérarchisation agglomérative . . . . .	8
11	Exemple de dendrogramme . . . . .	9
12	Modèles liant moment sismique et déplacement . . . . .	10
13	Spectres de déplacement : de documentation (gauche), et expérimental (droite)	12
14	Arguments de la fonction <code>correlate_template</code> . . . . .	14
15	Courbe de <i>NCC</i> aux abords d'un pic de détection . . . . .	15
16	Pics de détections successifs de <i>NCC</i> . . . . .	15
17	Égalisation des fréquences d'échantillonnage du template et du signal . . . . .	16
18	Signaux avec un filtre passe-bande [2-10 Hz] . . . . .	16
19	Exemple d'évènements inscrits dans le catalogue . . . . .	17
20	Chaîne de calcul simplifiée . . . . .	17
21	Extrait de WI.MPOM.00.HHZ . . . . .	18
22	Temps de calcul en fonction des durées de chunk . . . . .	19
23	Essai de rognage d'un template . . . . .	19
24	Valeurs de <i>NCC</i> en fonction du taux de décimation . . . . .	20
25	Différences de temps de calcul selon le taux de décimation . . . . .	20
26	Spectres de bruit de FDF, MPOM, DSD, BIM . . . . .	22
27	Période de validité des stations FPF, MPOM, DSD, BIM . . . . .	22
28	Étendue temporelle des familles avec <code>plot_timepans</code> . . . . .	24
29	<code>plot_timepans</code> pour MPOM avec une <i>NCC</i> de 0.8 . . . . .	24
30	<code>plot_timepans</code> du catalogue validé pour BIM sur [2014-2025] . . . . .	26
31	Famille de référence sur BIM à 11 évènements . . . . .	27
32	Histogrammes de <i>NCC</i> <i>a priori/ a posteriori</i> . . . . .	28
33	Comparaison des détections avant et après dégénérescence du template . . . . .	29
34	Evolution des <i>NCC</i> avant/après ciblage sur les ondes S . . . . .	29
35	Évolution des écarts <i>NCC</i> → <i>NCC<sub>s</sub></i> - ordre croissant de <i>NCC</i> . . . . .	30
36	Moyennes pondérées des traces avant et après sélection . . . . .	30
37	Superposition des traces avant et après sélection . . . . .	31
38	Calcul de <i>NCC</i> centré sur les Ondes S . . . . .	31
39	Comparaison des <code>plot_cumulative</code> sur une famille de référence FDF . . . . .	32
40	Régression linéaire de la famille $\mathcal{F}_{inf}$ . . . . .	33
41	Régression linéaire de la famille $\mathcal{F}^{sup}$ . . . . .	33
42	Schéma de la transition sismo-géologique en profondeur sur la subduction.	36
43	Localisation des 68 familles de repeaters sélectionnées sur BIM . . . . .	37
44	Exemple supposé d' <i>antirepeater</i> détecté sur FDF . . . . .	38

# Introduction

L'interface de subduction des Petites Antilles est encore peu connue. Les dynamiques tectoniques mesurées semblent incompatibles avec les séismes majeurs de 1839 et 1843 (Magnitude 8–8.5). Pour tester la possibilité de mégaséismes de type megathrust dans cette région, plusieurs approches sont adoptées. Les données [GNSS](#) sont utilisées pour calculer les vitesses horizontales relatives des plaques, et calculer des champs de contraintes pour proposer des modèles de couplage entre les plaques Caraïbe et Nord-Américaine [1] ; et les propriétés physico-acoustiques de la [lithosphère](#) des Antilles sont évaluées grâce à un réseau de sismomètres marins et de sismographie passive ou active (*airgun shooting*), afin de mieux cerner les dynamiques de la subduction [2].

Des simulations récentes suggèrent un couplage faible de l'interface antillaise, excluant des scénarios à la Tohoku [1]. La plaque Nord-Américaine s'enfoncerait à une vitesse de 20 mm/an sous la plaque Caraïbe sans accumulation de contraintes. Cela entre en contradiction avec les séismes destructeurs de 1839 et 1843, responsables de milliers de morts, documentés par des géologues de l'époque. Des incertitudes demeurent quant à la localisation de leur épicentre et l'absence de tsunami observé reste difficile à concilier avec un scénario de séisme de subduction de type megathrust [3]. Une hypothèse plausible est une rupture au sein de la plaque plongeante (intra-slab) [2] et non au niveau de l'interface de subduction, rendant le scénario compatible avec un couplage interplaqué faible.

Les modèles à grande échelle offrant une vision globale des processus de subduction ([GNSS](#)) peuvent masquer des hétérogénéités locales. Nous adoptons donc une approche à plus petite échelle fondée sur les séismes répétés (repeating earthquakes, ou *repeaters*) : événements distincts réactivant quasi à l'identique une même aspérité, séparés de mois à années, présentant ainsi des formes d'onde très similaires, les conditions de rupture et de propagation étant presque inchangées. La forte corrélation croisée entre les signaux enregistrés permet de les regrouper en familles et par conséquent d'observer indirectement le glissement asismique et la réaccumulation de contraintes dans la zone de subduction, en suivant des aspérités localisées (rayon ~100 m ou moins) (*A.Doucet (2021)* [4]).

Ce rapport de synthèse vise en premier lieu à reprendre et poursuivre les travaux présentés par *A.Doucet*, et à ouvrir de nouvelles réflexions sur les méthodes qui y sont utilisées, et les résultats obtenus. Ce rapport approfondira notamment la recherche des familles, avec des plages de temps allongées, des gestion de bases d'événements plus larges, des seuils de corrélation plus adaptés, et l'ouverture de l'étude d'une même famille sur les données de plusieurs stations. FDF étant assez bruitée, la zone de recherche permise ne s'étendra pas autant que dans *A.Doucet (2021)* [4], cependant les plages temporelles couvertes seront trois fois plus vastes, ce qui améliorera d'autant le recul sur les familles de *repeaters*. Ensuite, la notion même d'appartenance à une famille sera questionnée, notamment sur le plan algorithmique. Enfin, les paramétrages de calcul pour le *template matching* seront analysés en détail, et les magnitudes d'une famille de référence seront calculées et analysées afin d'encadrer les valeurs de déplacement par glissement autour d'une source de séismes répétés. La zone d'intérêt sera fixée au Nord-Est de la Martinique, principalement en raison de la présence de FDF, station active depuis 1998 permettant une étude prolongée de la zone (voir la zone retenue en fig. 3).



# Première partie

## 1.1 Contexte des Petites Antilles

Les Petites Antilles forment un arc volcanique de 850 km à l'Est de la mer des Caraïbes (figures 1, 2). Cet arc, issu de la subduction de la plaque Nord-Américaine sous la plaque Caraïbe, comprend 11 îles volcaniques majeures sur environ 700 km en axe Nord/Sud. La convergence moyenne des plaques est de 19 mm/an [3].

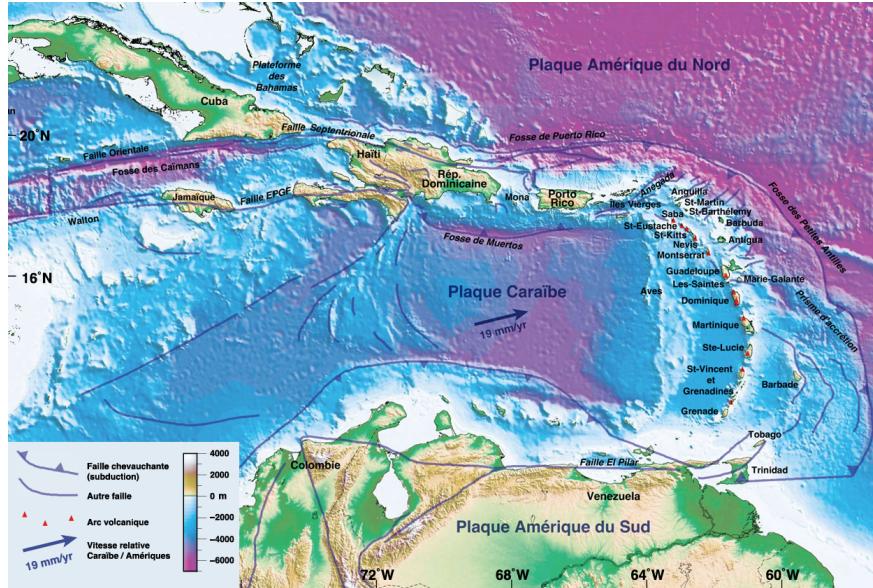


FIGURE 1 – Vue d'ensemble des Antilles et des plaques tectoniques concernées. La flèche bleue au centre indique le mouvement relatif des plaques vers la fosse. Extrait de [1].

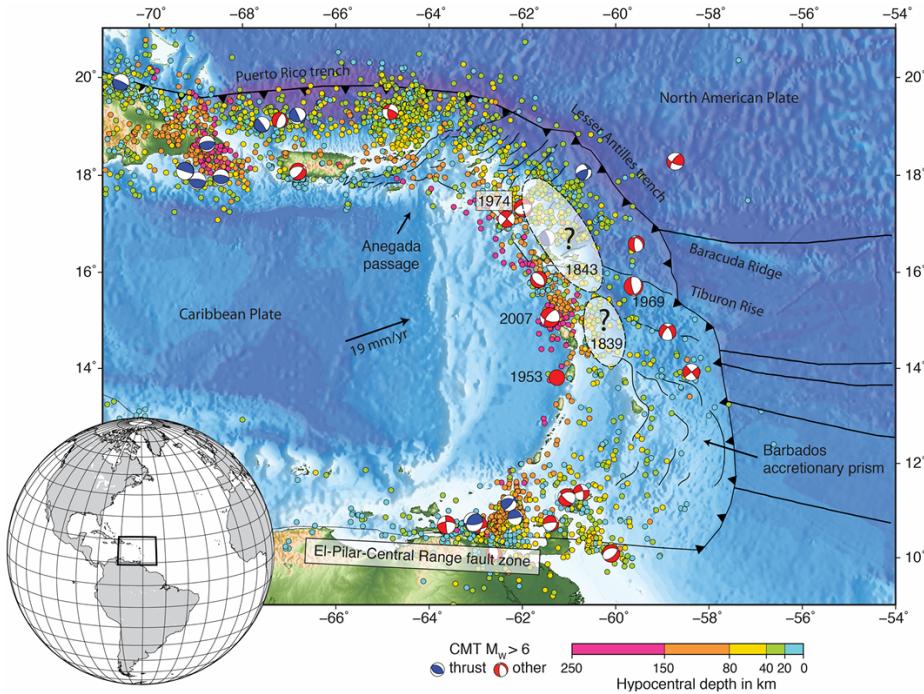
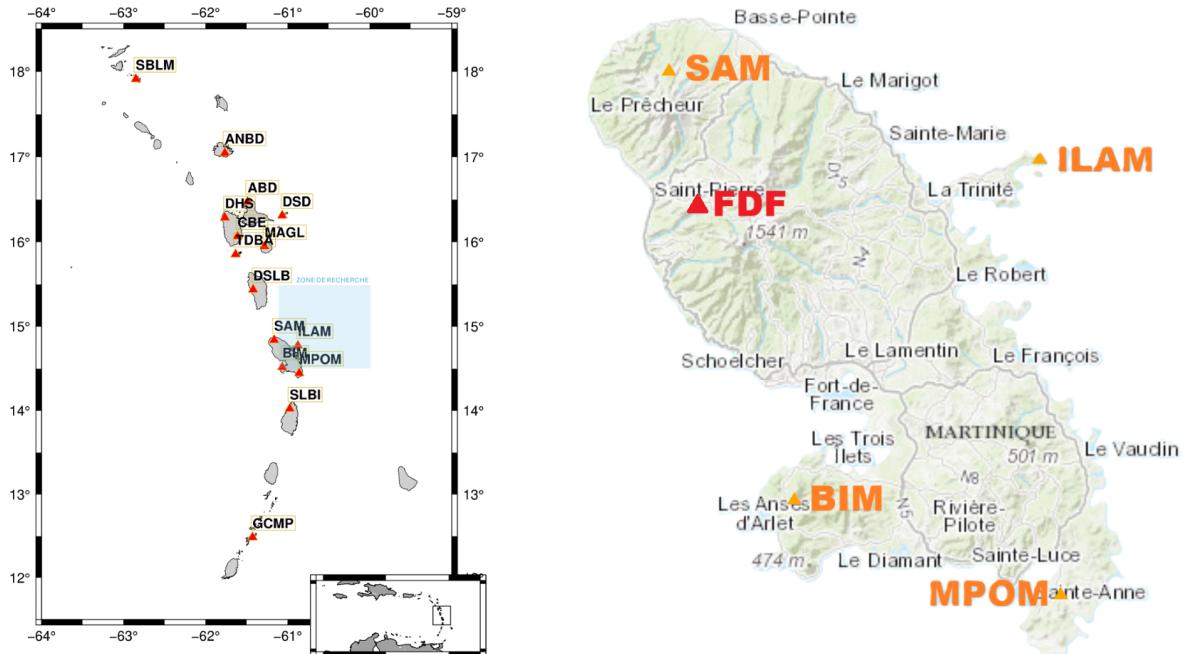


FIGURE 2 – Arc des Petites Antilles, avec les zones supposées des séismes de 1839 et 1843 et les principaux séismes colorés selon la profondeur de l'hypocentre [1]

Cette subduction génère une activité sismique intense : plusieurs milliers de séismes de magnitude inférieure à 5 sont enregistrés chaque année (figure 3). Les foyers se concentrent autour du *slab* — la plaque Nord-Américaine plongeante, et dans la plaque Caraïbe elle-même (séismes intra-plaque). Les figures 4 et 5 illustrent la structure de subduction et les valeurs numériques associées proviennent de [2] : la zone d'interaction principale, comprise entre 20 et 60 km de profondeur (entre le *backstop* et le coin mantellique), correspond aux séismes d'interface (entre les plaques tectoniques). Au-dessus de 20 km et en-dessous de 60 km, les séismes sont intraplaque. Précisément, dans le deuxième cas, en raison des hautes pressions et températures, les roches plongeantes déshydratées peuvent créer ces séismes intraplaque. Au-delà de 200 km, ces roches deviennent ductiles, il n'y a plus de ruptures sismiques.



(a) Carte des stations sismiques du réseau W.I. aux Petites Antilles. (b) Carte des stations sismiques présentes en Martinique (Réseau W.I. et Géoscope).

FIGURE 3 – Cartes des stations sismiques du réseau W.I. aux Petites Antilles et en Martinique. Les noms sont des acronymes des sites d'installation. Le carré bleu (voir fig.3a) indique la zone de recherche des *repeater*. La station historique de FDF est en rouge fig.3b

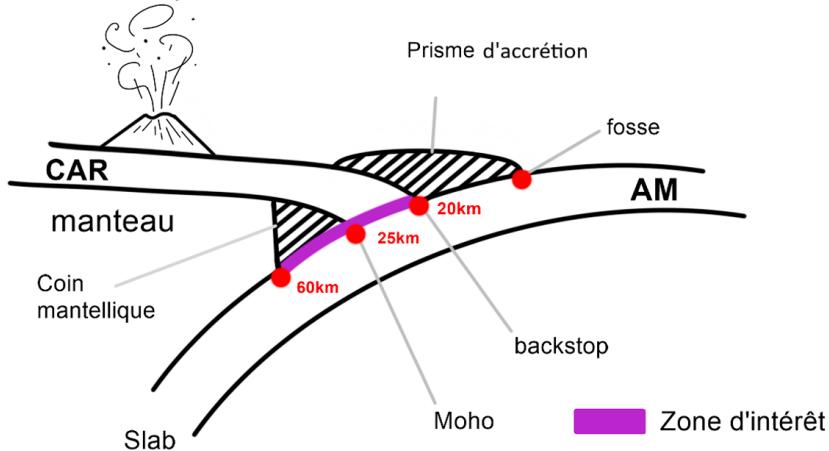


FIGURE 4 – Schéma de subduction et terminologie associée. Zone violette : région potentielle de génération des séismes répétés. Points rouges : frontières entre interfaces en mouvement. Extrait de [4]

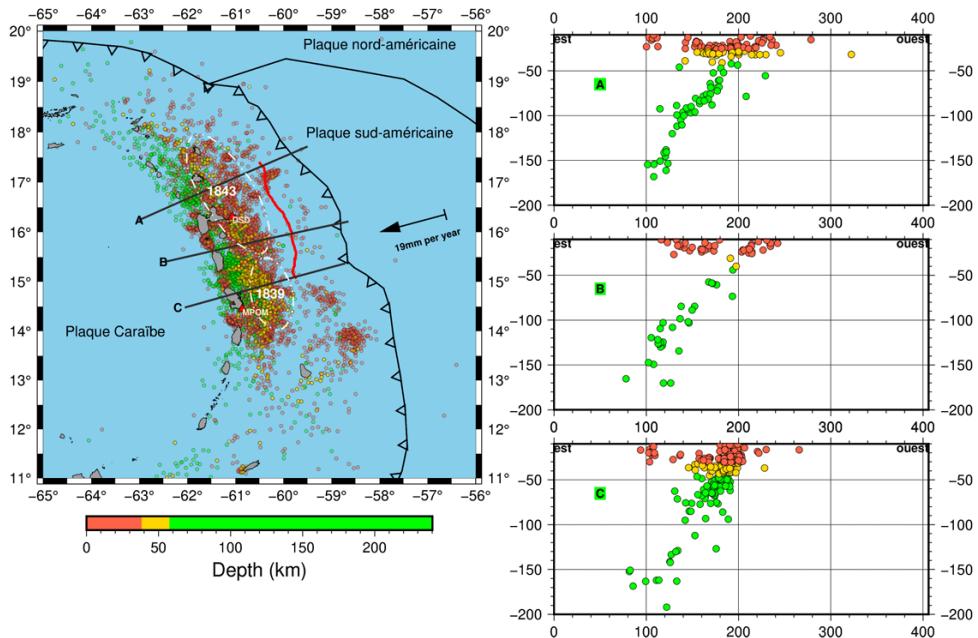


FIGURE 5 – Activité sismique aux Petites Antilles entre 2014 et 2021 selon les données de l’IPGP. La ligne rouge correspond au *backstop* [2]. Les séismes sont colorés selon leur profondeur : rouge pour [0-40 km], orange pour [40-60 km], vert pour 60 km et plus. Les coupes A,B,C sont signalés par des segments noirs sur la carte de gauche. Extrait de [4]

Le degré de **couplage** entre les plaques est un indicateur clé du risque de séismes de type **megathrust**. Il traduit la proportion de **glissement** bloqué à l’interface de contact interplaque :

- **Couplage nul (0%)** : La plaque plongeante est en fluage (**glissement** continu) et n’entraîne pas la plaque subductée (il n’y a pas d’accumulation de contraintes) ;
- **Couplage total (100%)** : La plaque plongeante est comme soudée par friction (**blocage complet**) et la plaque subductée est poussée comme un ressort qui peut se libérer violemment sous forme de séisme (accumulation d’énergie élastique) ;
- **Couplage partiel (X%)** : Correspond au rapport des vitesses de déplacement entre la plaque subductée et la plaque plongeante.

Les figures 6 et 7 illustrent ces différentes situations de couplage [5].

L’étude des Petites Antilles fait face à un problème profond et récurrent : l’explication de l’occurrence de deux mégaséismes en 1839 et 1843 au large de la Guadeloupe et de la Martinique (voir les zones supposées des épicentres en fig.2), ayant causé plusieurs milliers de morts. Dans les deux cas, le **couplage** fait justement partie des éléments étudiés pour mieux décrire les comportements de la subduction et tenter d’étudier la possibilité de mégaséismes aux Antilles. Un certain nombre d’éléments se contredisent : les modèles de couplage sont incompatibles avec les observations **GNSS**. Les analyses de sédiments mettent en évidence les traces de séismes majeurs, mais aucun tsunami n’a été observé en 1839 ou 1843, ce qui semble improbable. Cette étude vise donc à éclairer les motifs de cette contradiction.

Dans les Petites Antilles, l’étude de couplage par suivi **GNSS** des vitesses de déplacement suggère un couplage faible, voire très faible. En faisant l’hypothèse que le couplage a pu être élevé dans le passé, on s’attendrait à observer plus de 175 ans après les séismes, au moins une reprise du couplage entre les plaques, ce qui n’est pas le cas [1].



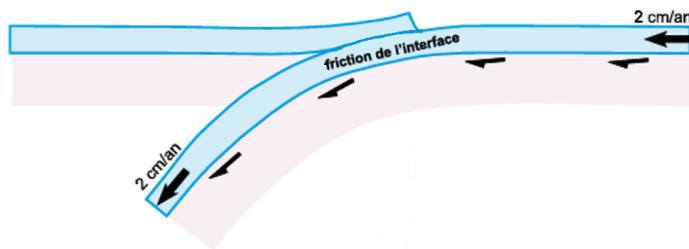


FIGURE 6 – Schéma synthétique de subduction avec un couplage nul (0%). La friction à l'interface n'est pas suffisante et la plaque subductée n'accumule par d'énergie, ni ne se contracte. La vitesse relative du mouvement entre les plaques est égale à la vitesse de la plaque plongeante. Adapté de [5].

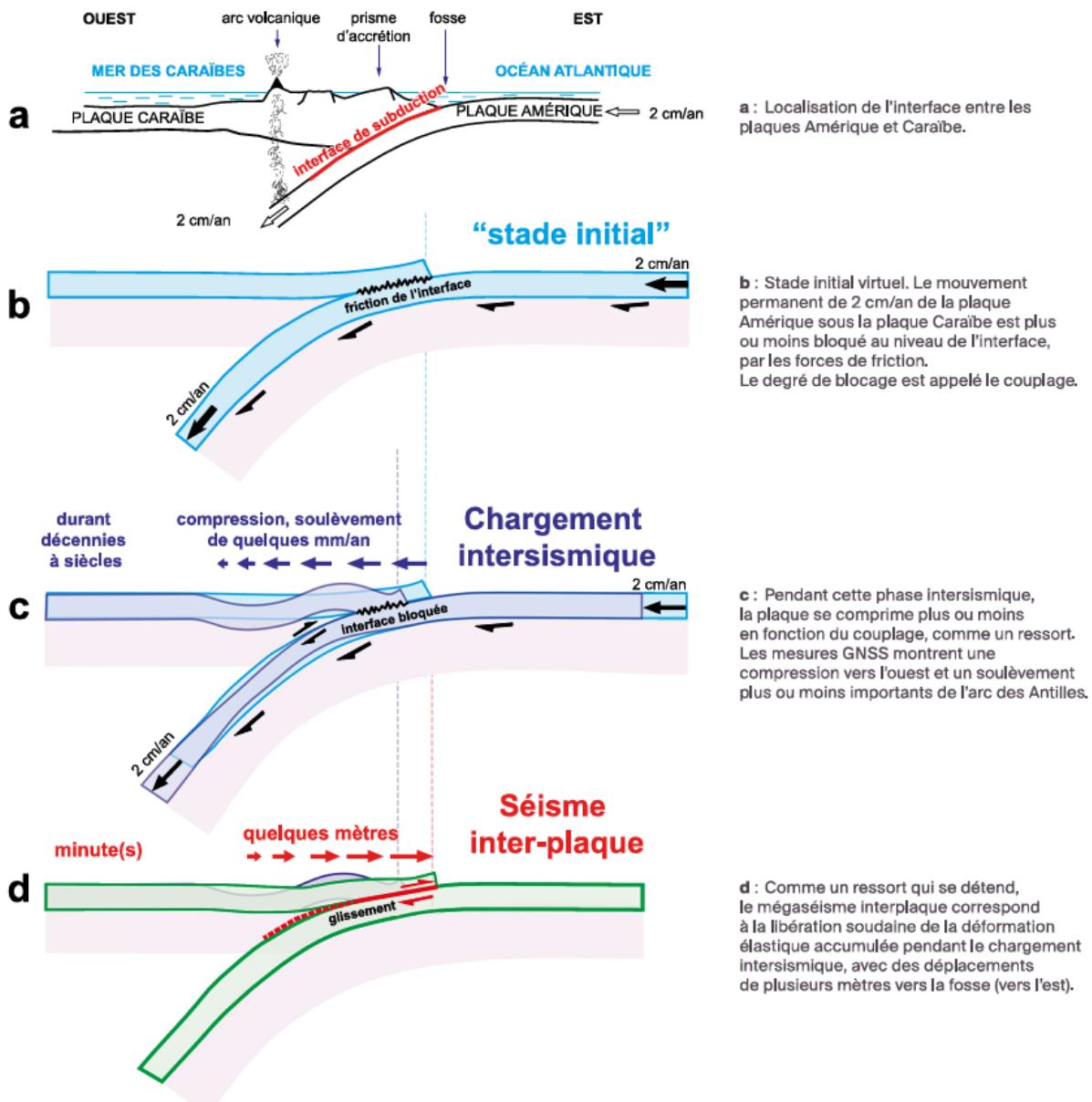


FIGURE 7 – Description d'une fosse de subduction à couplage non nul. La friction générée à l'interface entre les plaques (b) crée une contraction longitudinale et transversale sur la plaque subductée qui recule et accumule de l'énergie (c). Le relâchement des contraintes engendre un séisme violent (d). Extrait de [5].



## 1.2 Utilisation des séismes répétés

Le couplage renseigne sur l'état mécanique de l'interface entre les plaques. Lorsqu'il est estimé à partir des vitesses mesurées par [GNSS](#), il correspond à une vision intégrée à grande échelle, qui ne permet pas de décrire les variations locales du comportement de la faille. Pour accéder à ces échelles plus fines, on utilise l'étude des séismes répétés : ces événements, qui rompent successivement sur une même [aspérité](#), reflètent le [glissement](#) local sur la zone de subduction. Le glissement ainsi estimé peut ensuite être comparé aux vitesses de plaque déduites des mesures GNSS globales. La figure 8 extraite de [6] schématise les courbes de glissement pour une plaque plongeante, avec des zones "coulissantes" (glissement proportionnel au temps) et des zones d'irrégularités (glissement en escalier avec des paliers bien définis).

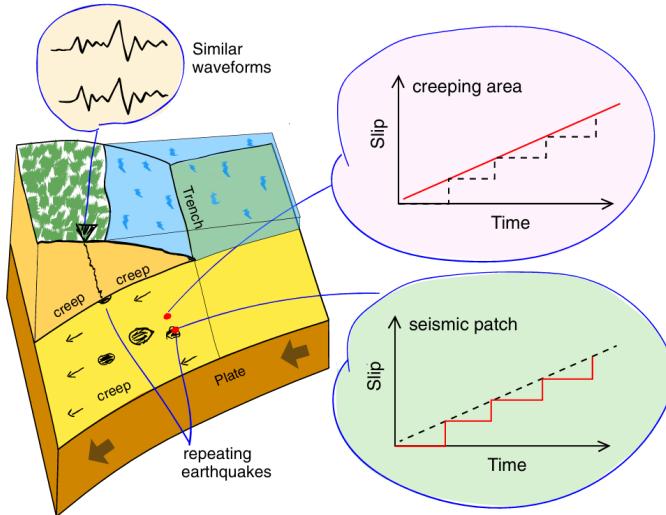


FIGURE 8 – Schéma introduisant la formation des séismes répétés (*repeaters*). Ils apparaissent au niveau d'[aspérités](#) de l'interface de subduction. La figure illustre deux modes de [glissement](#) : un glissement continu (en haut) produisant un déplacement linéaire, et un glissement par à-coups (en bas) correspondant à des ruptures successives au même endroit, générant une courbe de déplacement en paliers. Extrait de [6]

Les [aspérités](#) de l'interface entraînent un glissement discontinu, marqué par des ruptures successives. Leur détection sismique permet d'estimer le glissement associé à chaque rupture, puis, sur le long terme, de reconstituer l'évolution de leur profil de glissement et la vitesse moyenne des zones voisines. Ce rapport s'attache donc à l'analyse d'enregistrements sismiques sur plusieurs décennies afin de mieux comprendre le couplage des Petites Antilles.

## 1.3 Le *template matching*

Le *template matching* (recherche de motif) est une méthode permettant la recherche d'un motif temporel donné, appelé *template*, dans un signal beaucoup plus long. La similitude entre le template et le signal est déterminée par un calcul de *cross correlation* (corrélation croisée) normalisée, ou [NCC](#). La corrélation croisée entre deux signaux continus consiste en un calcul d'intégration entre deux séries temporelles, décalées d'un facteur  $\tau$ . Pour chaque valeur de  $\tau$ , on calcule la *NCC*, dont le maximum indique le décalage optimal pour lequel le template correspond le mieux au signal. La normalisation adoptée est celle du *facteur de corrélation de Pearson*, consistant à diviser le produit scalaire des deux signaux par le produit de leurs écarts types respectifs  $\sigma_T \sigma_S$ .



La formule utilisée pour des séries temporelles discrètes est rappelée par [7] sous sa forme générale :

$$NCC(t) = \frac{1}{N_s N_c} \sum_{s,c} \frac{\sum_{n=1}^N T_{s,c}(t_n) S_{s,c}(t + t_n + \tau_{s,c})}{\sqrt{\sum_{n=1}^N T_{s,c}^2(t_n) \sum_{n=1}^N S_{s,c}^2(t + t_n + \tau_{s,c})}} \quad (1)$$

Dans l'équation 1,  $s$ ,  $c$  et  $n$  sont les indices de la station, du canal (*ex* : HHZ), et du temps, et  $N_s$ ,  $N_c$ ,  $N$  désignent le nombre de stations, le nombre de composantes, et la longueur des séries temporelles discrètes.  $T$ ,  $S$  sont les signaux temporels du template utilisé et des données continues de la station.  $\tau_{s,c}$  est le décalage temporel pour une station donnée  $s$ , et une composante donnée  $c$  (il permet de synchroniser les signaux dont on calcule la corrélation). Cela revient à calculer, pour une station et un canal donné, le rapport d'une corrélation sous la forme (Produit Scalaire / Produit des Normes), car  $\|X_{s,c}\| = \sqrt{\sum_{n=1}^N X_{s,c}^2(t_n)} \propto \sigma_X$ .

En pratique, le calcul de  $NCC$  sera mené sur une seule station et un seul canal, car une seule station de référence sera utilisée à la fois (MPOM, BIM, DSD ou FDF, voir la figure 3), soit  $N_s = N_c = 1$ , et  $\tau_{s,c}$  négligé, soit :

$$NCC(t) = \frac{\sum_{n=1}^N T(t_n) S(t + t_n)}{\sqrt{\sum_{n=1}^N T^2(t_n) \sum_{n=1}^N S^2(t + t_n)}} \quad (2)$$

Ou encore :

$$NCC(t) = \frac{\sum_{n=1}^N T(t_n) S(t + t_n)}{N \sigma_T \sigma_S} \quad (3)$$

Ce qui fait apparaître explicitement la normalisation et le moyennage suivant le produit scalaire. Un calcul de  $NCC$  glissant, en utilisant un même template sur tout un signal permet, à chaque emplacement du signal d'obtenir une valeur maximale de similitude locale. Sur l'ensemble du signal, les similitudes maximales entre template et motif cherché se distinguent par des valeurs élevées de  $NCC$ , idéalement des pics. De tels pics seront considérés comme des détections s'ils dépassent un certain seuil (Voir en troisième partie : 2.6).

La figure 9 montre un exemple de pic de  $NCC$  lorsque la fenêtre du template se superpose avec un motif d'intérêt.

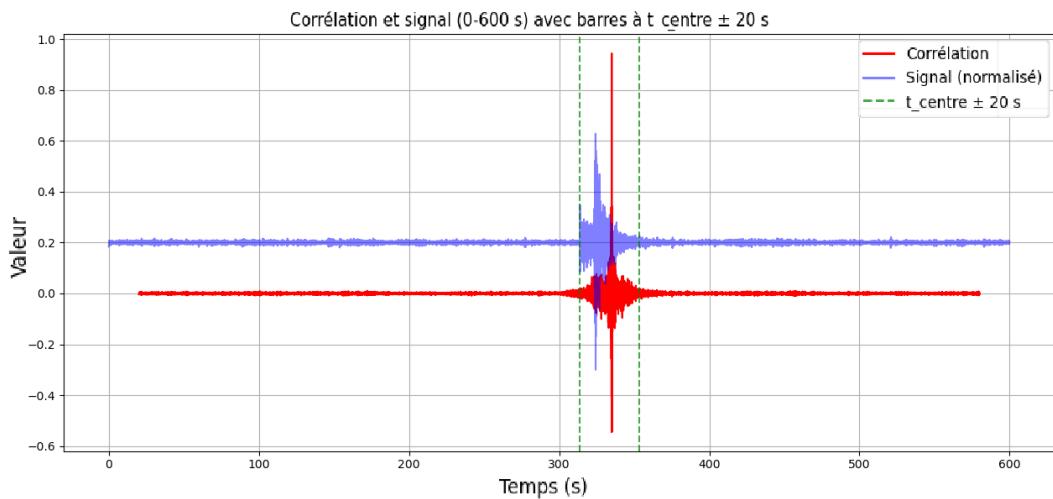


FIGURE 9 – Test de calcul de  $NCC$  sur un séisme, en utilisant comme template la fenêtre découpée de ce séisme. La section du signal entre les pointillés verts est donc égale au template. On observe l'apparition d'un pic de  $NCC$  (rouge) lorsque la fenêtre délimitée par les rebords en pointillés verts se superpose avec un séisme d'intérêt (bleu). Par convention, la valeur de  $NCC$  est attribuée au milieu de la fenêtre. Le template est de 40 secondes, et débute dès l'apparition des ondes P.



## 1.4 Algorithmes de création des familles

La création de familles à partir des catalogues de séismes est fondamentale afin d'identifier différents clusters d'évènements caractérisés par un motif de template. Cette étape est importante : en effet, la pertinence des familles a une influence directe sur la qualité des templates qui leur sont associés, et donc sur les recherches d'éventuels nouveaux évènements non répertoriés. Or chaque algorithme implique un certain type de familles, les deux principaux étant :

- **Première méthode** : Recherche par composantes connexes, ou agglomération hiérarchique. Cette méthode permet de construire des familles "de proche en proche", selon un seuil de  $NCC$  donné. Si deux évènements  $A$  et  $B$  ont une  $NCC > 0.95$  par exemple, on les place dans une seule famille. Si un évènement  $C$  a une  $NCC$  de 0.45 avec  $A$  mais de 0.97 avec  $B$ , alors il est inclus dans la famille. Cette méthode consiste à assembler des familles par proximités successives. Un exemple synthétique en est donné en figure 10. Son avantage est sa simplicité algorithmique, et son approche visuelle et intuitive. Sa complexité en calcul est linéaire par rapport au nombre d'évènements ( $\mathcal{O}(n)$ ), après avoir calculé les  $NCC$  au préalable. Cependant, elle présente le désavantage de relier dans une famille certaines paires d'évènements ayant mutuellement une valeur de  $NCC$  basse, typiquement inférieur au seuil de  $NCC$  requis pour ajouter un évènement à la famille.
- **Seconde méthode** : UPGMA (*Unweighted Pair Group Method with Arithmetic mean*). Cette méthode héritée de la biologie crée des familles itérativement de façon à créer un arbre phylogénétique des évènements sismiques. Un haut degré de similarité (une haute  $NCC$ ), implique une position basse dans l'arbre. L'algorithme se base sur la matrice des distances associée à l'ensemble des évènements disponible. À chaque étape, les deux évènements partageant une distance minimale (reliee à la hauteur de la corrélation), sont fusionnés (on crée un embranchement), et on recommence. La distance calculée à chaque étape entre deux clusters est une moyenne pondérée par la taille de ces clusters. [formule] Ce fonctionnement permet de favoriser l'émergence de clusters fidèles à la réalité, bien qu'il désavantage les petites familles. **N.B.** : la mention "Unweighted" du nom de la méthode UPGMA ne vient donc pas du calcul des distances, mais du fait que chaque évènement est considéré avec un poids égal, puisque le poids d'une famille est proportionnel au nombre de ses évènements. À la fin de la création de l'arbre phylogénétique (ou dendrogramme), on "coupe" l'arbre au niveau du seuil de corrélation souhaité, ce qui fait apparaître les différentes familles (figure 11). Cet algorithme est plus robuste mais plus coûteux également, avec une complexité en  $\mathcal{O}(n^2)$ , où  $n$  est le nombre de séismes à clusteriser.

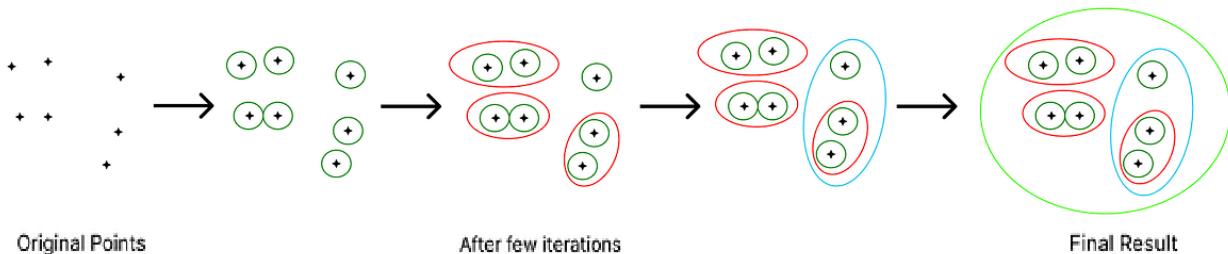


FIGURE 10 – Exemple schématique d'une hiérarchisation agglomérative. Les clusters sont formés selon un critère de moindre distance. Tous les points de la figure appartiennent à une unique famille. Ils sont agglomérés au fur et à mesure, jusqu'à constituer la famille entière.



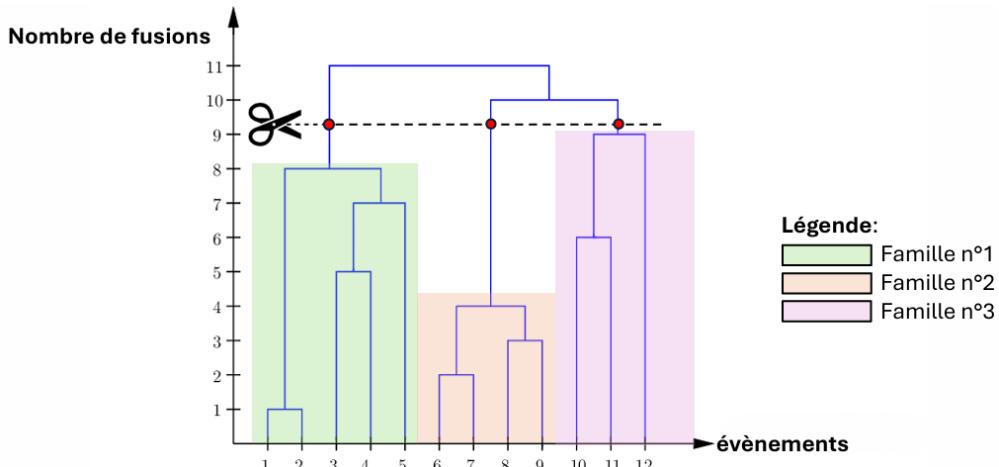


FIGURE 11 – Exemple de dendrogramme (ou arbre phylogénétique). Les événements sont sur l'axe des abscisses. Plus deux événements ont un embranchement commun bas, plus ils sont proches, car les algorithmes utilisent une métrique de distance minimale. Si à la fin de l'algorithme de classification on coupe avec un seuil de *NCC* exigeant, cela équivaut à couper le dendrogramme vers les branches les plus basses, donc à créer de nombreuses familles très homogènes. Si ce seuil est moins fort, les familles seront moins nombreuses et plus diversifiées.

D'autres méthodes existent, citons notamment [WPGMA](#), la méthode complémentaire d'[UPGMA](#), dont le calcul de distances n'est pas pondéré entre familles, ce qui permet de mieux valoriser les familles moins nombreuses. Dans ce cas, chaque famille a un poids égal, donc les événements ont un poids inégal. La dualité UPGMA-WPGMA illustre le dilemme entre la cohérence structurelle des familles et leur nombre. Un compromis pourrait être trouvé dans un calcul de distance avec une pondération intermédiaire entre les deux cas, et dont les paramètres de pondération pourraient être ajustés selon les jeux de données à traiter (qualité des données, absence de bruit...).

## 1.5 Magnitudes, moment sismique, et déplacement de plaque

La magnitude est une grandeur associée à la puissance d'un séisme, plus précisément la représentation logarithmique de la quantité d'énergie libérée lors de ce séisme. Il existe plusieurs définitions de la magnitude. La plus célèbre est probablement la magnitude de Richter, ou magnitude locale ( $M_L$ , sans unité), mais cette magnitude est peu utilisée car elle est imprécise aux hautes valeurs de magnitude, (typiquement  $> 6$ ), et nécessite d'être étalonnée pour une région donnée. Une autre magnitude, plus courante, est la magnitude de moment ( $M_W$ , sans unité), cette fois directement reliée au moment sismique ( $M_0$ , en N.m, homogène à une énergie). C'est la magnitude de moment qui sera considérée dans ce projet. Cette grandeur est explicitement reliée au moment sismique, et donc à la quantité d'énergie libérée, et elle reste consistante aux hautes magnitudes, sans étalonnage nécessaire. La relation entre  $M_W$  et  $M_0$  est :

$$M_W = \frac{2}{3}(\log_{10}(M_0) - 9,1) \quad (4)$$

Ou de façon équivalente,

$$M_0 = 10^{\frac{3}{2}M_W + 9,1} \quad (5)$$



Le moment sismique, est une grandeur dont le nom rappelle que l'énergie libérée lors d'un séisme peut se modéliser à un couple de forces appliqué au niveau d'une faille. On définit en général ce moment comme :

$$M_0 = \mu S \Delta u \quad (6)$$

Où :

- $\mu$  : module de rigidité (ou de cisaillement) de la roche, en  $Pa$  ;
- $S$  : surface de la roche rompue lors du séisme ;
- $\Delta u$  : déplacement de la roche le long de la faille associé au séisme.

Cette dernière relation permet de faire le lien entre les propriétés mécaniques de la roche et la quantité d'énergie libérée pendant le séisme.

Cependant, les paramètres  $\mu$ ,  $S$  et  $\Delta u$  étant très complexes, voire impossible à estimer pour un séisme donné, une autre relation sera employée : un modèle empirique, dit de *Nadeau et Johnson* [6][8] permettant, dans le cadre des séismes répétés sur une faille donnée, de relier  $M_0$ , exprimé en  $dyne.cm$ , ou  $10^{-7} N.m$ , au déplacement *local* de la roche, noté  $d$ , exprimé en cm. On a alors :

$$d = 10^{-2.36} \cdot M_0^{0.17} \quad (7)$$

D'autres modèles existent comme ceux de *Beeler*, ou *Eshelby-Crack* [4]. Ces modèles sont présentés sur la figure 12 ci-dessous. Il est donc possible, connaissant la magnitude d'un évènement sismique de type *repeater*, de lui associer un déplacement, en mm ou cm, de la faille au niveau de la zone de rupture. Ceci constitue l'un des objectifs de ce projet de synthèse.

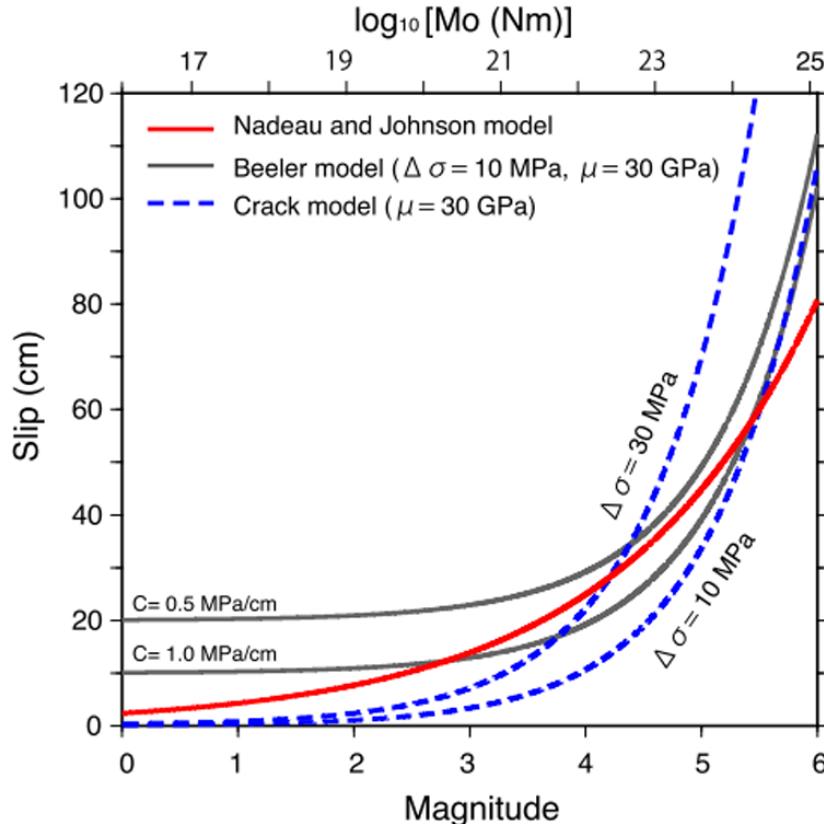


FIGURE 12 – Modèles liant moment sismique et déplacement : Nadeau-Johnson, Beeler, Crack. Ces modèles dépendent aussi des paramètres géophysiques de la zone de rupture, par exemple de  $\mu$  pour le modèle de Crack (bleu). Extrait de [6].

## 1.6 Estimation de la magnitude de moment, et de la *corner frequency*

Le paragraphe suivant vise à décrire le socle théorique utilisé pour estimer la magnitude de moment  $M_W$  d'un évènement sismique. Dans un second temps la grandeur nommée *corner frequency*, ou fréquence coin  $f_c$  sera également introduite.

La première étape est de calculer un spectre de déplacement des données sismométriques pour un évènement. Ce spectre de déplacement peut se calculer sur les [ondes P](#), ou les [ondes S](#) d'un séisme. Le calcul sur les ondes P ne sera pas exploité ici. En pratique, les ondes S seront majoritairement utilisés, car plus énergétiques et visibles. On note  $f : \rightarrow S^{p|s}(f)$  la fonction associée au spectre de l'évènement pour les ondes P ou S. On utilisera donc souvent  $S^s(f)$ . Ce spectre correspond à la norme Euclidienne des 3 composantes spectrales correspondant aux 3 composantes spatiales du sismomètre, par exemple les composantes E,N,Z, pour une base orthonormée d'axes ( $\vec{E}$ st,  $\vec{N}$ ord,  $\vec{E}$ st  $\wedge$   $\vec{N}$ ord). Les composantes spectrales sont les modules de leur transformée de Fourier respectives.

$$\text{On a donc : } S^{p|s}(f) = \sqrt{\left(S_E^{p|s}(f)\right)^2 + \left(S_N^{p|s}(f)\right)^2 + \left(S_Z^{p|s}(f)\right)^2} \quad (8)$$

Ce spectre permet d'étudier la signature spectrale des signaux sismiques indépendamment de l'orientation spatiale des capteurs. En effet, les propriétés de magnitude d'un séisme sont perçues de façon isotrope, et seule importe l'amplitude résultante du séisme telle qu'elle est perçue par un sismomètre.

La relation permettant d'estimer la magnitude de moment à partir des données sismométriques est la suivante [9] :

$$S^{p|s}(f) = \frac{1}{\mathcal{G}(r)} \times \frac{FR_{\Theta\Phi}}{4\pi\rho_h^{1/2}\rho_r^{1/2}c_h^{5/2}c_r^{1/2}} \times M_0 \times \frac{1}{1 + \left(\frac{f}{f_c^{p|s}}\right)^2} \times e^{-\pi ft^*} \quad (9)$$

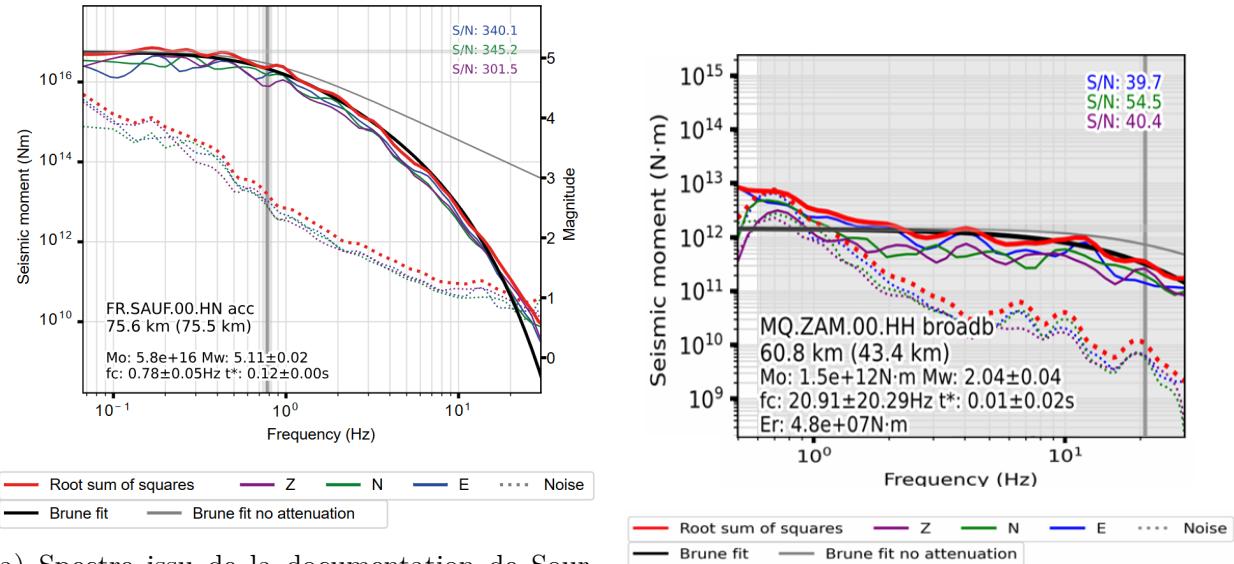
où :

- $S^{p|s}$  est le spectre d'amplitude calculé précédemment ;
- $\mathcal{G}(r)$  est une fonction décrivant la dispersion géométrique de l'énergie libérée, pour une distance radiale  $r$ .  $\mathcal{G}(r)$  est souvent de la forme  $r : \rightarrow r^n$  avec  $n = 1$  par défaut (propagation sphérique) ;
- $F$  est un facteur influencé par la réflexion de l'énergie. Souvent on a  $F = 2$  pour l'énergie incidente et l'énergie réfléchie. ;
- $R_{\Theta\Phi}$  traduit l'anisotropie de la libération d'énergie depuis l'[hypocentre](#). Il y a souvent des lobes privilégiés pour les ondes P et S ;
- $\rho_h$  and  $\rho_r$  sont les densités du milieu à l'hypocentre et au niveau du récepteur (le sismomètre) ;
- $c_h$  and  $c_r$  sont les vitesses de propagation des ondes P ou S à l'hypocentre et au niveau du récepteur ;
- $M_0$  est le moment sismique ;
- $f$  est la fréquence de travail ;
- $f_c^{p|s}$  est la fréquence coin (*corner frequency*) ;
- $t^*$  est un paramètre d'atténuation, en secondes.

Ceci est un résultat du modèle de *Brune* (1970) [10][11]. Dans l'équation ci-dessus, le terme de gauche est donc accessible directement en manipulant les données expérimentales. Le terme de droite, quant à lui est partiellement connu. L'ensemble des constantes liées à la propagation sont supposées connues. Le terme de droite se ré-écrit donc :

$$C \times \frac{M_0}{1 + \left(\frac{f}{f_c^{ps}}\right)^2} \times e^{-\pi f t^*} \quad (10)$$

C étant connue, les trois paramètres  $M_0$ ,  $f_c$ , et  $t^*$  sont ajustés par régression non-linéaire pour correspondre au mieux au spectre  $S^s(f)$ . Les trois valeurs en sont alors déduites. Tous les calculs de régression pour estimer les paramètres évoqués ci-dessus sont réalisés numériquement. Un exemple de spectre produit par l'outil SourceSpec décrit à la partie 2.1 est donné en figure 13a.



(a) Spectre issu de la documentation de SourceSpec. **En rouge** : terme expérimental issu des données sismiques, exprimé dans les relations 8 et 9. **En noir** : ajustement par le modèle de *Brune* (plateau basse fréquence puis pente  $\propto f^{-2}$ ).  $f_c$  est indiqué par la barre grise. **En pointillés** : niveaux de bruit utilisés pour calculer le *SNR*, afin de s'assurer de la qualité du signal traité. Tiré de la [documentation de SourceSpec](#).

(b) Spectre expérimental obtenu sur FDF. Ce type de spectre permet directement d'obtenir une magnitude exploitable pour étudier le glissement au niveau des *aspérités*, via des modèles comme celui de *Nadeau-Johnson*. Les hautes valeurs de *SNR* permettent ici une bonne qualité de spectre, contrairement à d'autres événements où les faibles *SNR* contraignent voire empêchent le calcul de magnitudes.

FIGURE 13 – Spectres de déplacement : de documentation (gauche), et expérimental (droite)

La fréquence coin  $f_c$  est l'un des trois paramètres ajustés lors de la régression. Sachant que  $t^* \ll 1\text{s}$ , on a  $S^s(f) \propto \frac{1}{1+(f/f_c^{ps})^2}$ , soit un comportement de *passe-bas* : plateau aux basses fréquences puis décroissance  $\propto f^{-2}$  aux hautes fréquences. Entre ces deux régimes, la fréquence  $f_c$  marque la transition entre le plateau et la pente d'atténuation progressive, d'où son nom de fréquence *coin*. Cette fréquence, outre son caractère important sur la propagation de l'énergie sismique, joue un rôle central pour l'analyse des sources des *repeater* et des chutes de contrainte associées aux événements sismiques des familles de *repeater*.



# Deuxième partie

## 2.1 Présentation et fonctionnalités de *Requake*, *SeisCat* et *SourceSpec*

Requake [12] est un outil *open source* développé en Python pour l'analyse des séismes répétés, exécutable depuis un environnement type [Anaconda](#) (ex. VS Code), il fonctionne sous Windows, Linux et MacOS. Son développement collaboratif sur GitHub permet de soumettre des propositions d'amélioration de code. Les différentes fonctions proposées par Requake permettent notamment de :

- Parcourir les catalogues de sismicité pour identifier d'éventuels séismes répétés ;
- Construire des catalogues de similarité avec des contraintes de localisation ;
- Former des familles de séismes répétés à partir de ces catalogues de similarité ;
- Compléter les familles en détectant, par *template matching*, des séismes non catalogués. Cette fonction n'était pas fonctionnelle auparavant. Nous avons donc complété son algorithme de *template matching*, dont la réalisation est détaillé par la suite ;
- Visualiser les évènements d'une famille dans le temps ;
- Tracer les signaux d'une famille et leur corrélation avec la moyenne.

Ces fonctions reposent sur le module ObsPy [13], possédant de nombreux outils de traitement sismologique. Certains fichiers de code de Requake seront adaptés pour optimiser la recherche de nouvelles familles sur deux décennies de données. **N.B** : La description des principales fonctions de Requake est donnée dans le [glossaire](#).

La version utilisée pour ce projet est : `requake 0.6+74.gd3c4318` [12]

Le second outil open-source, SeisCat [14], génère des catalogues d'évènements sismiques à partir de critères géographiques, temporels ou de magnitude, en interrogeant les bases de données IPGP, IRIS, etc. Il peut aussi concaténer ces données aux catalogues existants. Les catalogues sont directement formatés pour être utilisés par Requake. L'outil est également employé pour lancer les calculs de magnitude via la fonction *run* de SourceSpec (voir section [3.4](#)).

La version utilisée pour ce rapport est : `seiscat 0.8+71.gea89e63` [14]

Le dernier outil GitHub manipulé pour ce projet, SourceSpec, [9] est utilisé après la constitution des familles pour estimer la magnitude de moment ( $M_w$ ) et la fréquence coin ( $f_c$ ). Pour chaque évènement, il récupère les traces des stations voisines, calcule leurs spectres de déplacement, et applique le modèle de *Brune* avec quelques hypothèses physiques supplémentaires pour en déduire l'estimation de  $M_w$  et  $f_c$ . Plus le nombre de traces est élevé, plus les estimations peuvent être précises. Ainsi, la recherche de données inclut les accéléromètres : bien que moins sensibles aux faibles magnitudes, l'intégration temporelle de leurs signaux fournit des vitesses ( $m.s^{-1}$ ) exploitables comme celles des sismomètres. L'outil enregistre les stations utilisées et génère un rapport HTML synthétisant les spectres et la qualité des données.

La version utilisée pour ce rapport est : `source_spec 1.8+167.ga768e63` [9]

## 2.2 Préparation des codes de *template matching*

L’élaboration d’un algorithme de template matching s’est faite en plusieurs étapes. Tout d’abord, nous avons réalisé des codes *standalone* afin de lancer les premiers essais de détections. Les premiers signaux testés pour la détection par seuil de *NCC* sont des signaux synthétiques de type ondelette (gaussienne  $\times$  sinus) et permettent d’identifier clairement les étapes de calcul et d’affichage nécessaires pour mettre en avant les pics de *NCC* correspondants. Nous avons ensuite basculé sur des signaux réels, en intégrant des fonctions Obspy déjà optimisées ([13], dont la fonction `correlate_template` qui calcule une corrélation en utilisant la *FFT* (*Fast Fourier Transform*)), puis nous avons enfin perfectionné notre chaîne de calcul (voir en section suivante). Les premiers signaux réels testés faisaient 1 à 2 heures au maximum, déjà centrés sur des séismes de dates UTC connues d’avance. Dès lors que la rapidité et la fiabilité des détections a été assurée, les calculs de *template matching* ont été lancés sur plusieurs semaines, puis mois, puis années. L’algorithme résultant a été intégré progressivement dans le corps de Requake, dans la fonction `scan_templates`, disponible en annexe GitHub [ici](#).

Les arguments de `correlate_template` sont précisés dans la figure 14 :

```
correlate_template(data, template, mode='valid', normalize='full', demean=True, method='auto') [source]
```

FIGURE 14 – Arguments employés dans la fonction `correlate_template` de Obspy, pour les calculs de `scan_templates` (voir [glossaire](#)) (voir la [documentation Obspy](#)).

- `mode = "valid"` : La dimension du tableau plus petit est toujours incluse dans celle du plus grand (pas de dépassement) ;
- `normalize = "full"` : normalisation indépendante à chaque position du template (car les propriétés statistiques du signal ne sont pas réellement stationnaires, donc il faut normaliser avec des écarts-types locaux) ;
- `demean = True` : Recentrage à chaque nouvelle position du template (moyenne mise à zéro) ;
- `method = "auto"` : Usage de la *FFT* si nécessaire/possible.

Sur signaux réels, la corrélation d’un motif similaire au template produit un pic net de *NCC* proche de 1. On observe souvent un léger aplatissement avant le pic maximal (figure 15), proportionnel à la durée précédant les ondes P dans le template. Un template plus court réduit ce plateau, mais augmente la sensibilité au bruit : les valeurs de *NCC* deviennent instables et sujettes à des erreurs numériques. À l’inverse, un template plus long lisse les irrégularités. L’aplatissement n’affecte pas la détection ; il pourrait même servir à distinguer les vrais pics de ceux dus au bruit, si une baisse locale d’énergie ( $E_{[t_a, t_b]}(NCC) = \sum_{n=0}^N |NCC[t_a + n \times \Delta t]|^2$ ), où  $N = \left\lfloor \frac{t_b - t_a}{\Delta t} \right\rfloor$  est observée avant le maximum. La figure 16 montre cet effet sur deux détections successives.

Les codes finaux de calcul de *NCC* sont ensuite implémentés dans Requake, dans le fichier `scan_templates.py`. Les paramètres y sont ajustés conjointement au fichier de configuration `config.conf`, qui centralise les variables globales.



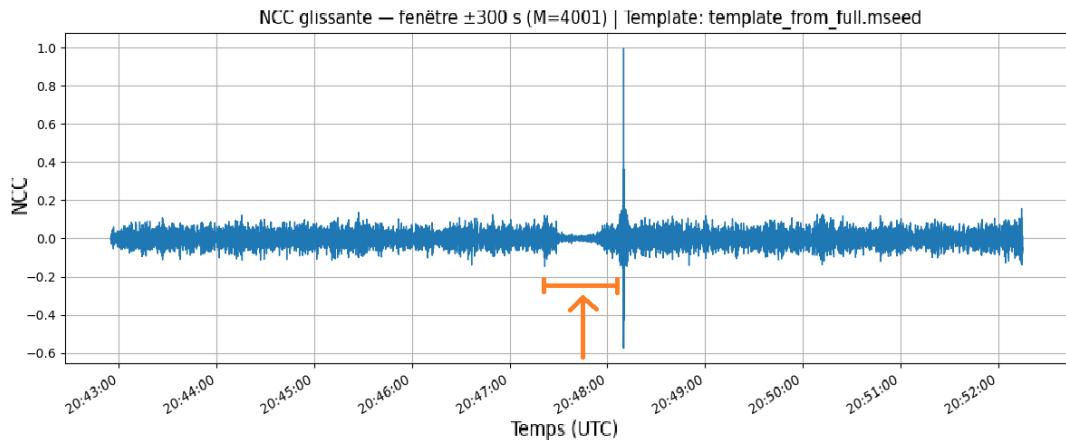


FIGURE 15 – Courbe de *NCC* aux abords d'un pic de détection. Le marquage orange signale la présence d'un aplatissement caractéristique, dont la largeur dépend de la largeur du template.

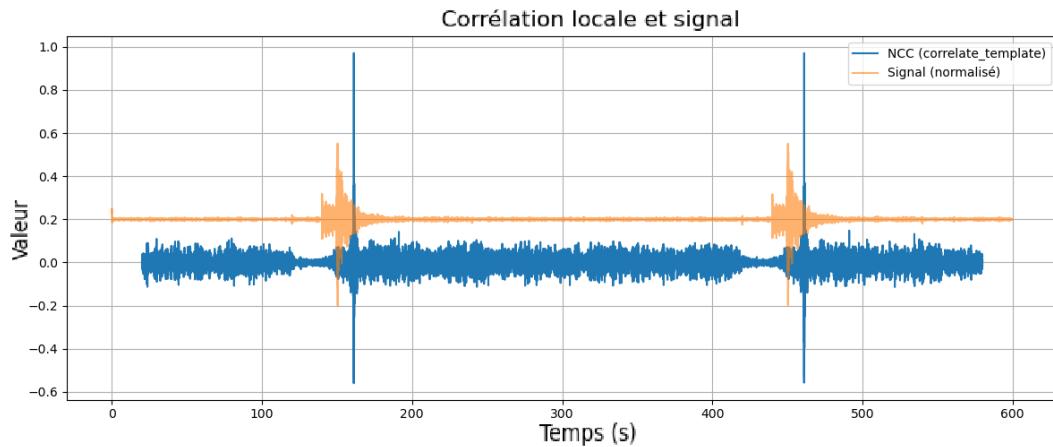


FIGURE 16 – Pics de détections successifs de *NCC* (bleu) autour de deux séismes similaires au template (orange). Le signal sismique a été généré et normalisé synthétiquement à des fins de clarté, mais les signaux réels sont souvent bien plus denses et bruités.

## 2.3 Chaîne de calcul retenue pour la détection par *NCC*

Cette section décrit les principales étapes de l'algorithme de calcul de la *NCC* dans la fonction `scan_family_templates` du fichier `scan_templates.py` (voir le [glossaire](#)). La chaîne de calcul résumée est exposée en figure 20 :

- Importation des données sismiques entre deux dates [UTC](#) via le [webservice](#) IPGP, en utilisant le format NET.STA.LOC.CHAN (ex. : WI.MPOM.00.HHZ) ;
- Copie explicite des signaux de trace et de template pour ne pas altérer les données initiales ;
- Décimation des signaux (facteur 2) pour optimiser le calcul, avec égalisation des fréquences d'échantillonnage du template et du signal. Si la décimation ne respecte pas la condition de [Nyquist](#) (20 Hz ici), le signal le plus lent est interpolé ; sinon, celui qui le permet est décimé. Un filtrage passe-bas précède toute décimation pour éviter l'aliasing. La figure 17 résume l'arbre de décision associé à cette étape ;



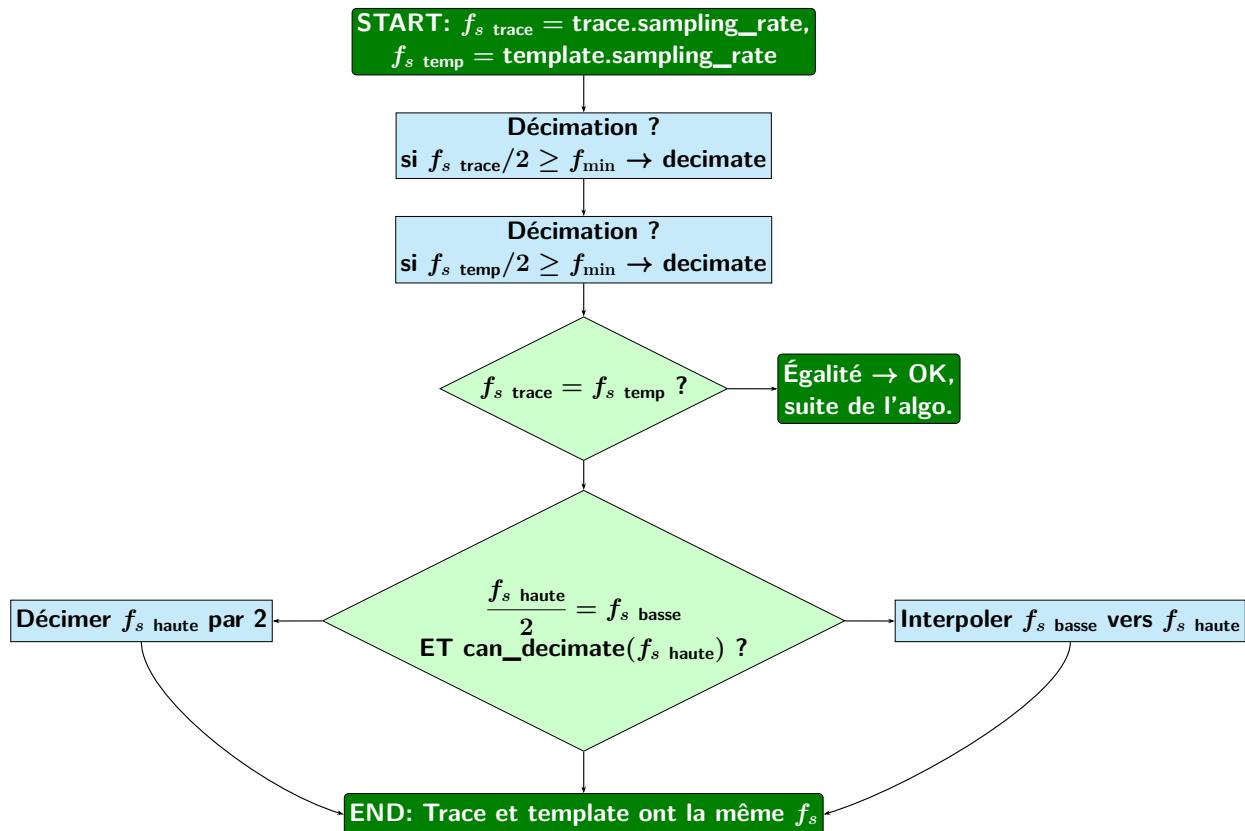


FIGURE 17 – Arbre de décision d'égalisation des fréquences d'échantillonnage du template et du signal sismique.  $f_{basse}$  est définie comme  $\min(f_s_{trace}, f_s_{temp})$ , et inversement pour  $f_{haute}$

- D) Filtrage de la trace et du template dans la bande [2–10 Hz] (cf. [4]) afin d'isoler les fréquences sismiques pour atténuer les bruits anthropiques et de la houle. La figure 18 donne un aperçu de l'influence du filtrage sur un signal sismique brut ;
- E) Conversion en tableaux NumPy de float64 pour optimiser les calculs ;
- F) Découpage des signaux par slice (isolement de la fenêtre utile au calcul) ;
- G) Calcul de la corrélation croisée normalisée (NCC), puis bornage des valeurs dans [-1 ; 1]. Ce traitement évite des arrondis de floats64 sans signification physique ;

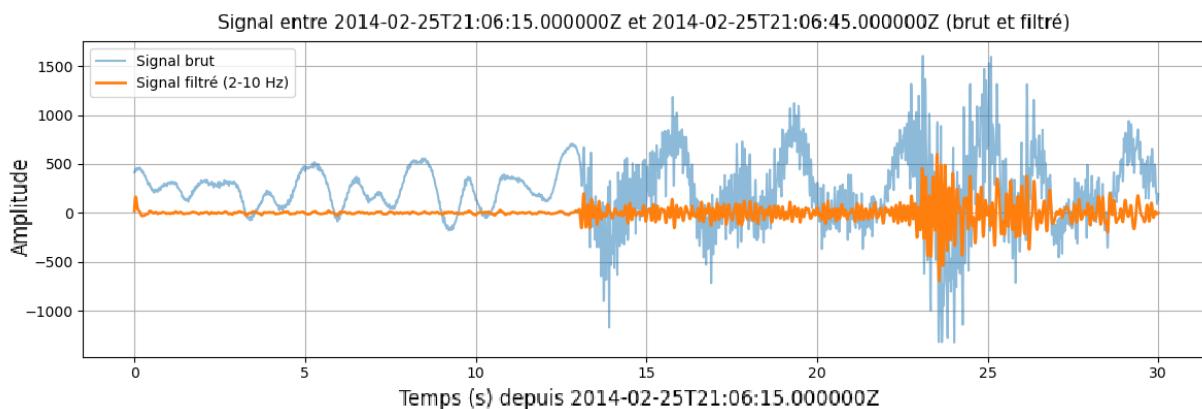


FIGURE 18 – Exemple de filtrage réalisé pour la chaîne de calcul de scan\_templates. Un filtre passe-bande [2-10 Hz] permet de s'affranchir des tendances de basses fréquences gêantes (par ex. les marées terrestres), et de bruits divers (houle, bruit anthropique, etc.).



- H) Extraction des maxima locaux puis sélection des pics principaux. Si plusieurs pics sont espacés de moins de deux jours, seul le plus fort est conservé. En cas de plateau continu de *NCC* au-delà du seuil, une détection unique est créée au centre (valeur moyenne) du plateau ;
- I) Rédaction de chaque détection dans un catalogue .txt formatée conformément à la figure 19.

```

quake_out > template_catalogs > catalog15.G.FDF.00.BHZ.txt > data
1 reqk1999giknlr|1999-03-31T00:17:09|None|None|None|reqake0.6+74.gd3c4318.dirty|None|None|None|None|None|None|0.54
2 reqk1999mfuffq|1999-06-22T01:13:12|None|None|None|reqake0.6+74.gd3c4318.dirty|None|None|None|None|None|None|0.54
3 reqk2000hfbonn|2000-04-11T06:49:21|None|None|None|reqake0.6+74.gd3c4318.dirty|None|None|None|None|None|None|0.54
4 reqk2000rqmcej|2000-09-05T05:19:44|None|None|None|reqake0.6+74.gd3c4318.dirty|None|None|None|None|None|None|0.52
5 reqk2000sqefvd|2000-09-19T03:14:49|None|None|None|reqake0.6+74.gd3c4318.dirty|None|None|None|None|None|None|0.50
6 reqk2001bimmlc|2001-01-19T16:03:15|None|None|None|reqake0.6+74.gd3c4318.dirty|None|None|None|None|None|None|0.53
7 reqk2001iajigf|2001-04-23T19:25:32|None|None|None|reqake0.6+74.gd3c4318.dirty|None|None|None|None|None|None|0.51
    
```

FIGURE 19 – Exemple d'évènements inscrits dans le catalogue. Certains évènements sont inconnus des catalogues. D'autres sont répertoriés. Les deux premiers champs sont le *event ID* et la date *UTC*. Le dernier champ est la *NCC* avec le template

Le taper (paramètre 0.05, soit 5% de la longueur des bords, de type "cosine") est utile pour éviter les lobes secondaires parasites suite à un filtrage. Idéalement appliqué avant le filtre passe-bande et après la décimation, le taper n'a pas été intégré ici car il n'améliorait pas la détection. La gestion des maxima suffisait à supprimer les pics parasites. Son utilité reste toutefois théoriquement fondée.

Il en va de même pour la fonction `detrend` de Obspy. Les données sismiques corrompues peuvent présenter un décentrage ou une dérive lente (*offset*). Bien que le filtre passe-bande et `correlate_template` corrigent en partie ce problème, appliquer un `detrend` systématique reste prudent pour éviter des détections irréalistes.

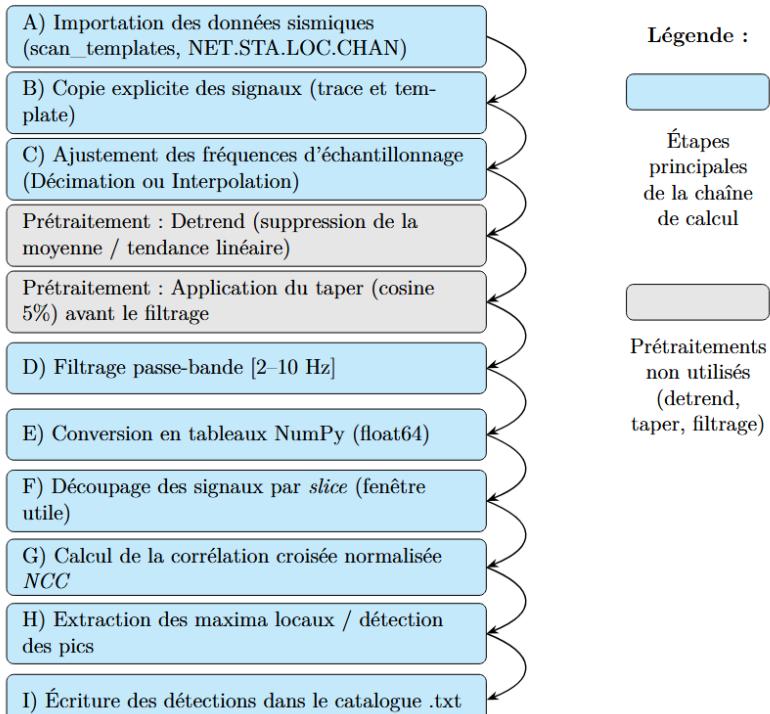


FIGURE 20 – Chaîne de calcul simplifiée, avec en gris les étapes théoriquement nécessaires, mais contournées par le présent code (voir en annexe GitHub : [ici](#)), sans pertes en performances de détection.

La figure 21 montre une section des données de MPOM partiellement corrompues par un décentrage de la masse du sismographe. Ce type d'anomalie rend les détections partiellement inexploitables car certaines étapes de la chaîne de calcul ne résistent pas à certaines irrégularités de très haute fréquence. Bien que rares, elles soulignent l'importance de pré-traiter systématiquement les données.

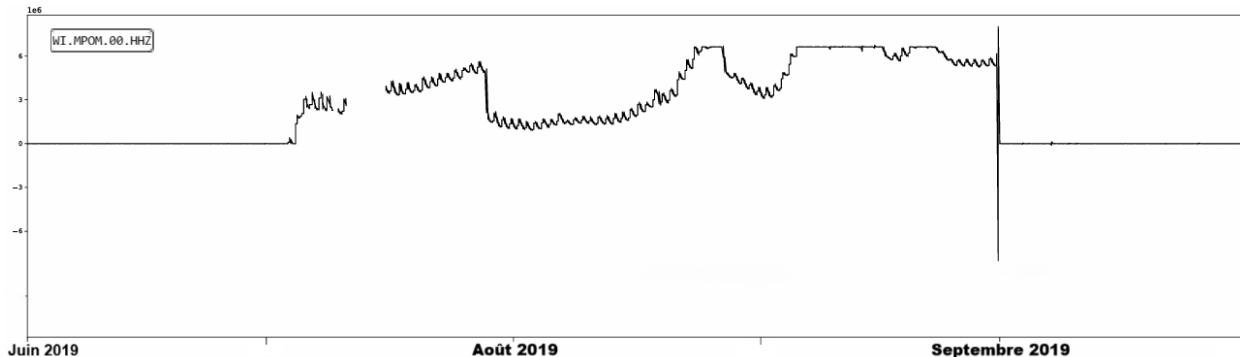


FIGURE 21 – Signal extrait de WI.MPOM.00.HHZ (juillet–octobre 2019). La moyenne, censée être nulle, montre une anomalie durable appelée décentrage, due à un défaut de position de l'élément mobile du sismomètre. Ce décentrage corrompt les détections par *NCC* : certaines zones sont saturées (plateaux) et des discontinuités perturbent la chaîne de calcul (impulsion fin octobre 2019).

Si le type de données est connu à l'avance, certaines vérifications (gestion d'erreurs, égalisation de fréquence, respect de Nyquist) peuvent être omises, ce qui accélère sensiblement le calcul. Les paramètres du `config.conf` doivent alors rester cohérents avec les données traitées.

Le code de `scan_templates` adapté et utilisé pour ce projet est disponible en annexe. Le fichier original est disponible sur le dossier GitHub : [12].

## 2.4 Paramétrages d'entrée du calcul de Template Matching

Les calculs de corrélation étant destinés à couvrir des périodes de plusieurs années, leur exécution peut durer plusieurs jours. Il est donc nécessaire d'optimiser les paramètres dans `config.conf` (fichier de configuration). Les paramètres ajustables détaillés dans les paragraphes suivants sont le *chunk*, la durée du template, le taux de décimation, et le parallélisme :

-Le chunk (fr. fragment) : Le paramètre *chunk* définit le découpage de la période de calcul en sous-sections successives, optimisant la gestion mémoire en ne chargeant que les données utiles à chaque segment. Bien qu'un découpage plus fin augmente le nombre de requêtes serveur, il améliore l'exploitation de la **RAM**, les données étant immédiatement recyclées pour le chunk suivant. Afin d'estimer la durée optimale du chunk, on effectue un même calcul en faisant varier sa taille (1h, 1j, etc. jusqu'à 1 mois). On fera ensuite varier la longueur de la plage temporelle de calcul pour vérifier que la valeur optimale du chunk est indépendante de la longueur de cette plage. La figure 22 présente les résultats en échelle logarithmique. On constate globalement que la valeur optimale d'un chunk est de 24 heures. Sur un calcul d'un an, on observe quasiment un rapport de 50 en temps de calcul entre une valeur de 24 heures et d'un mois. La valeur de 24 heures sera donc adoptée par la suite.



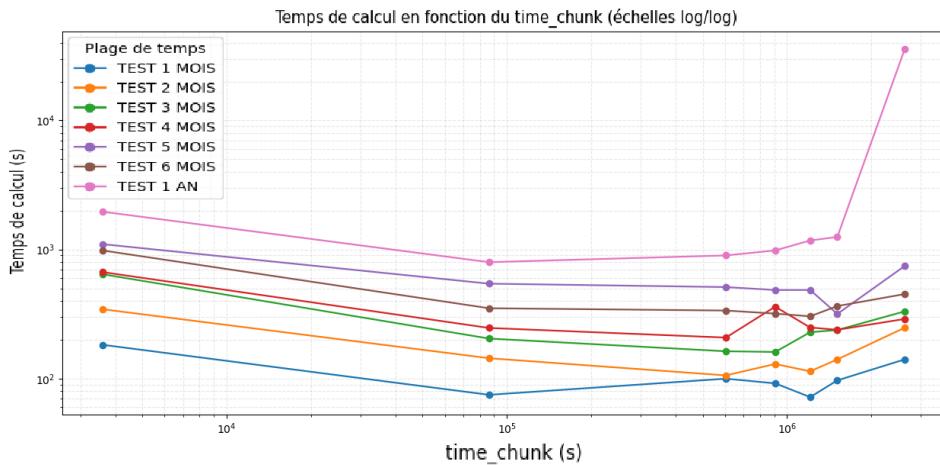


FIGURE 22 – Comparaison des temps de calcul pour plusieurs durées de chunk et plages temporelles (échelles log-log). L’optimum apparaît pour  $86400\text{ s} = 24\text{ h}$ , soit environ  $10^5\text{ s}$ .

-La durée du template : La durée d’un template produit par défaut sur Requake est de 2 minutes. La production d’un fichier d’extension .sac contenant le template est faite de telle sorte que les **ondes P** soient par défaut placées 20 secondes après le début du signal. 40-50s après les ondes P, le bruit est prépondérant. La fin du template consiste en une diminution des **ondes S** qui disparaissent dans le bruit. On pourrait donc chercher à rogner le signal afin de simplifier le fichier contenant le signal du template, sans perte d’informations (voir la figure 23). Le rognage n’apporte aucune amélioration notable : le gain de temps est nul et la sensibilité au bruit augmente. Finalement, la durée du template, éventuellement rognable via la fonction `trim` de Obspy, ne sera pas modifiée et reste donc à deux minutes.

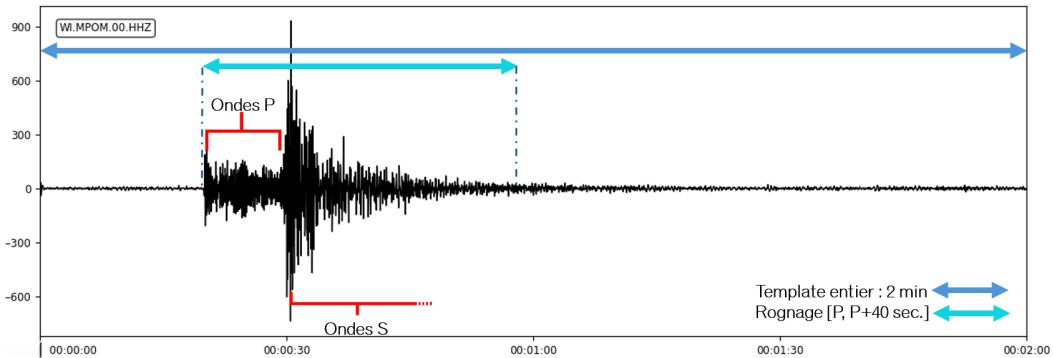


FIGURE 23 – Essai de rognage d’un template de 40 s débutant sur les ondes P. Les durées des templates complet et rogné (fonction `trim` de Obspy) figurent en bas à droite.

-Le taux de décimation : Ce paramètre réduit la complexité du calcul en sous-échantillonnant le signal. Un taux de décimation de  $N$  consiste à ne garder que  $1/N$  des valeurs, ce qui divise la fréquence effective d’échantillonnage par  $N$ . Étant donné que les courbes sismométriques manipulées sont filtrées en passe-bande sur [2 - 10 Hz], Avec une fréquence initiale de 100 Hz,  $N$  doit rester inférieur ou égal à 5 pour respecter la fréquence de Nyquist de 20 Hz (*cf.* Théorème d’échantillonnage de Nyquist-Shannon). On effectue la décimation avec la fonction Obspy `decimate`, qui permet si nécessaire de filtrer le signal en passe-bas, pour éviter tout risque d’aliasing (repliement de spectre). L’augmentation de  $N$  fait chuter les pics de  $NCC$ . Dans le cas  $N = 5$  (cas limite) les pics sont effondrés avec des  $NCC$  de 0.2 ou 0.3 typiquement, ce qui est inexploitable pour la détection de templates. Elle supprime les hautes fréquences nécessaires à une superposition précise du template. Suite à différents



essais, et afin de conserver l'avantage de la décimation lors du calcul sans détériorer la qualité de la détection par  $NCC$ , une valeur de  $N = 2$  est retenue. La figure 24 expose la variation des valeurs de  $NCC$  lorsqu'on fait varier le taux de décimation, et met en évidence la diminution négligeable des valeurs de  $NCC$  quand  $N$  passe de 1 à 2, avant de chuter lorsque  $N$  augmente encore. Finalement, la figure 25 montre un gain de temps de calcul d'un facteur 10 pour  $N = 2$  sur un mois de données.

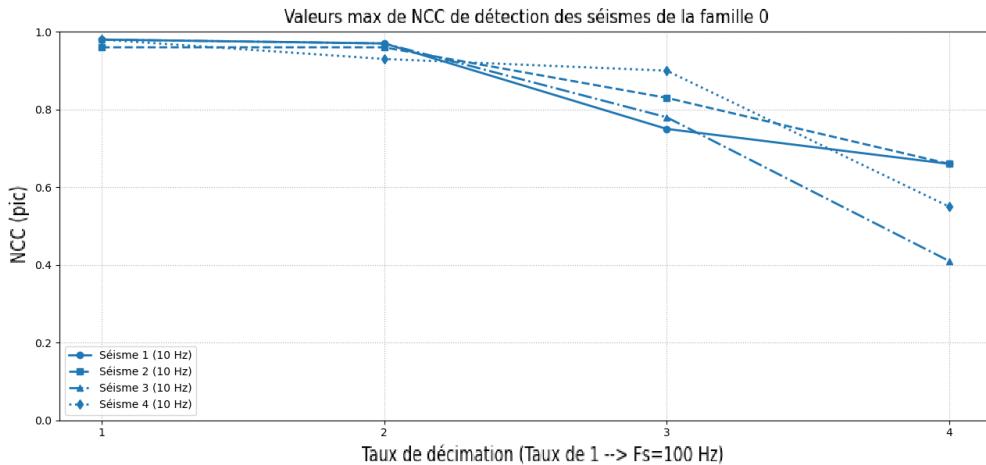


FIGURE 24 – Diminution des valeurs de  $NCC$  lorsque le taux de décimation augmente. Quatre séismes de référence ont été étudiés ici. Leur  $NCC$  a été calculée successivement, après décimation des séismes et du template. Les valeurs atteintes pour un taux de décimation  $\geq 3$  sont inexploitables.

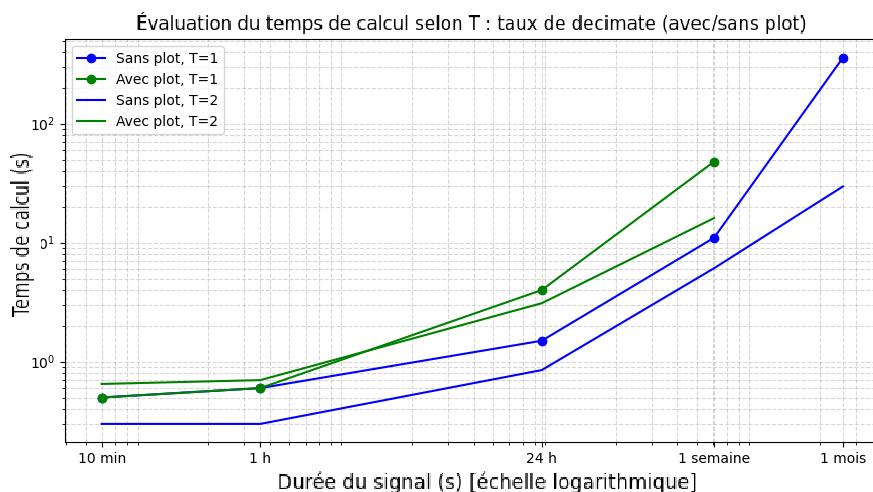


FIGURE 25 – Temps de calcul de la  $NCC$  par `scan_templates` avec ou sans décimation (taux de 1 ou 2). Les courbes sans *plot* sont naturellement plus basses. Sans *plot*, le passage d'un taux de 1 (décimation nulle) à 2 réduit d'un facteur proche de 10 le temps de calcul. Ainsi, bien qu'une valeur de décimation  $\leq 2$  soit adoptée, le gain de temps est fort.

-Le mode de lancement du calcul sur Requake : Requake donne la possibilité d'effectuer une recherche de template sur plusieurs templates en même temps. L'augmentation du nombre de templates a soulevé la question du mode de calcul : *lancer un template à la fois ou plusieurs en parallèle* ? D'après les essais effectués, la recherche en profondeur (avec plusieurs templates simultanément) s'avère de loin la plus efficace. Sur une plage d'une décennie, et pour un calcul avec 40 templates en parallèle, le gain estimé se compte en dizaines d'heures par rapport à l'autre approche. Malgré des variations non précisément chiffrables liées aux requêtes serveur, le gain est suffisamment net pour être retenu.



## 2.5 Utilisation de fichiers .conf

Le fichier de configuration de requake (`requake.conf`) est un fichier rassemblant tous les paramètres numériques nécessaires aux calculs effectués par Requake. Il est le pendant des données sismiques d'entrée, exploitées par Requake et permet leur exploitation en paramétrant les différentes fonctions (`scan_templates`, `scan_catalog`, `build_families` etc., voir le [glossaire](#)). Le formatage de ce fichier est contrôlé par un second fichier `configspec.conf`, faisant partie du code source, qui définit le type d'arguments à définir pour lancer le calcul, et la structure de données attendues dans le `requake.conf`. Si besoin, il est possible de rajouter des paramètres dans le `configspec` afin de mieux jouer sur les données. Par exemple, le taux de décimation a été implémenté, donc nous avons ajouté ce paramètre aux fichiers `.conf`. Cela permet de modifier l'exécution du code, sans toucher au code source. La répartition des variables entre les fichiers `.conf` et le code source doit permettre une bonne ergonomie d'utilisation des outils comme Requake.

## 2.6 Stations et plages de temps utilisées

Dans le cadre de ce PDS, l'étude est portée particulièrement sur le comportement de la subduction des Antilles à l'Est de la Martinique, car des stations opérées par l'IPGP y produisent des données continues et accessibles, permettant un suivi pluriannuel de la faille. Ces stations sont également peu bruitées, leurs signaux restent propres sur la bande fréquentielle utile. La carte des stations du réseau large bande [VI](#), dont certaines sont opérées par l'IPGP, est donnée en figure 3, avec un encadré délimitant la zone de recherche des événements sismiques de ce projet. Bien que certaines stations soient situées hors de la zone d'étude, elles peuvent néanmoins détecter les séismes qui s'y produisent.

La qualité des données sismiques est évaluée à l'aide de l'outil `SeedPSD` du RéSiF (Réseau Sismologique et Géodésique Français : <https://ws.resif.fr/resifws/seedpsd/1/histogram>), qui calcule les spectres de densité de probabilité de bruit d'une station, ceux de la figure 26 en sont issus. Ces spectres indiquent la puissance spectrale attendue sur une plage fréquentielle donnée et sont généralement calculés sur des longues périodes, afin d'être représentatif de l'environnement spectral à l'année sur une station. Pour calculer un spectre sur plusieurs années, on découpe la plage de temps en une succession de fenêtres (typiquement de 1 ou 2 heures) avec un *overlap* (débordement). Sur chacune de ces fenêtres, un calcul de PSD (densité spectrale de puissance) est mené, suivi d'un lissage spécifique. Enfin, tous les spectres `PSD` sont superposés. Sur les graphes, pour une fréquence  $f$  donnée, les niveaux de bruits les plus fréquents sont associés à une intensité spectrale plus haute, et à une couleur associée. Les maxima de puissance spectrale pour chaque fréquence sont positionnés entre deux courbes appelées *HNM* (*High Noise Model*) et *LNM* (*Low Noise Model*) (en noir sur les spectres fig.26). Ces courbes sont les limites acceptables des valeurs de niveaux de bruit. Ainsi le `LNM` correspond au bruit le plus faible qu'il est possible d'avoir sur les meilleures stations existantes, et `HNM` correspond aux pires stations. Le bruit recherché doit rester entre ces deux courbes, ce critère permettant de discriminer des stations au profit d'autres.

La bande des séismes [2–10 Hz] sera utilisée pour cette étude, pour minimiser la pollution par le bruit (anthropique, océanique, etc.). En comparant les performances spectrales des stations, MPOM sera préférée à BIM sur les données d'après 2014, jusqu'à la disparition des données MPOM de 2022, date où BIM sera de nouveau considérée. DSD (Désirade - Guadeloupe) sera écartée pour des raisons de qualité de bruit. Les spectres de FDF, MPOM, DSD et BIM (les stations principalement utilisées *in fine*) sont fournis en figure 26.



**N.B :** Par convention, les spectres de bruits sont calculés à partir d'accélérations dérivées (au sens de  $\partial/\partial t$ ) des signaux sismiques, d'où l'unité  $(m/s^2)^2/Hz$ . Cette grandeur est convertie en dB pour la lisibilité des spectres.

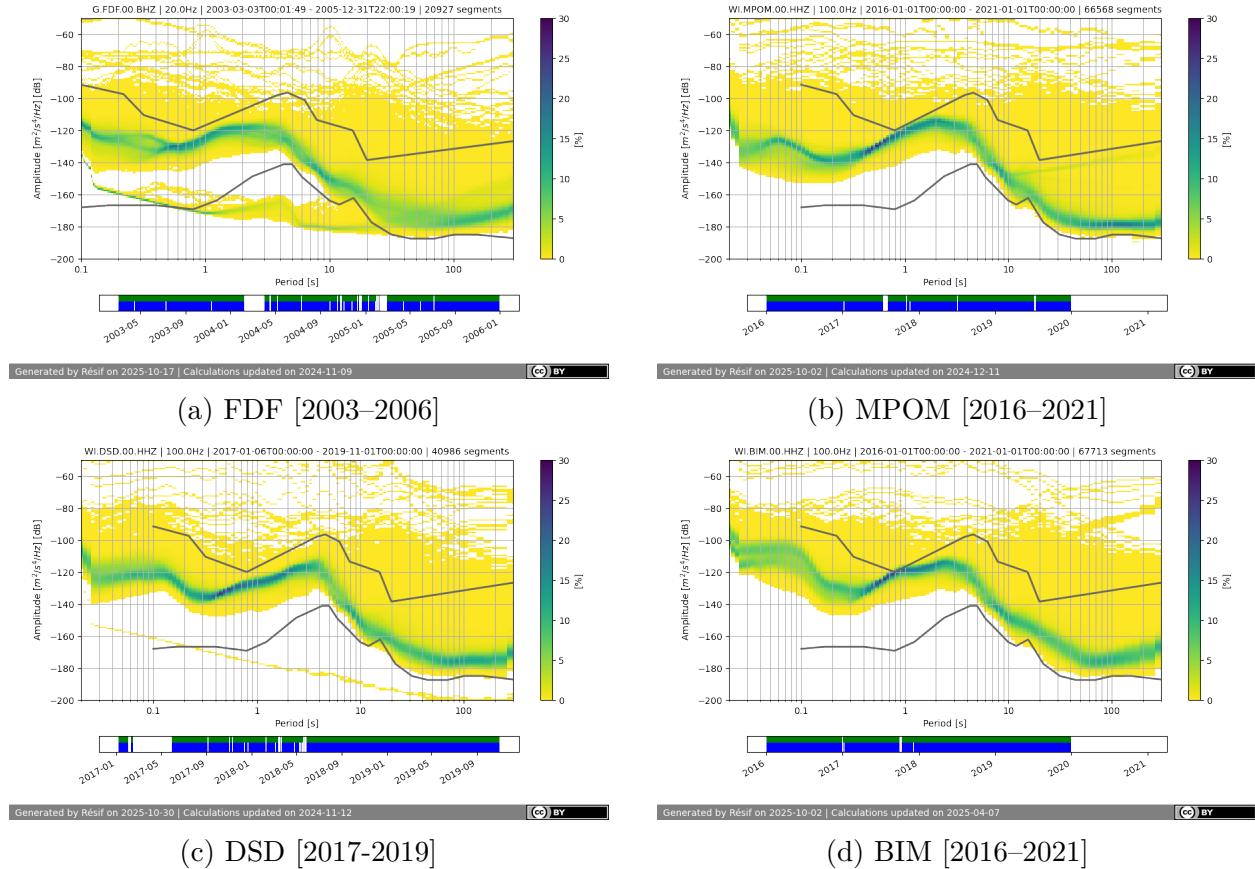


FIGURE 26 – Spectres des stations FDF, MPOM, DSD, BIM générés sur RéSiF. Les spectres sont calculés sur plusieurs années. L'abscisse correspond à la période. La couleur correspond, pour  $T$  donnée, à une densité de probabilité de bruit, dont les valeurs maximales sont en bleu. Les courbes en noir sont associées aux modèles HNM (en haut), et LNM (en bas).

Pour retrouver les familles de séismes de [4]], les stations MPOM (Morne Pois) et DSD (Désirade) sont exploitées pour la période 2014–2022. Les données récentes de MPOM n'étant plus disponibles (station vandalisée), BIM (Morne Bigot – Martinique) est utilisée pour la période 2014–2025. Enfin, afin d'analyser les anciens séismes au large de la Martinique, la station FDF (Fort-De-France - Martinique) est exploitée à partir de fin de décembre 1998. Le résumé des plages d'utilisation des stations est donné ci-après en figure 27.

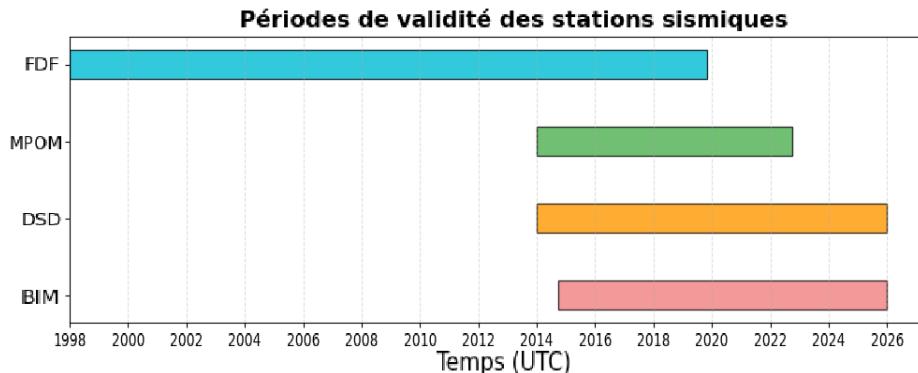


FIGURE 27 – Plages de temps exploitées sur les stations FDF, MPOM, DSD, BIM

# Troisième partie

Cette partie constitue le centre de l'exploitation et l'analyse des données de cette étude. Dans les anciennes versions de Requake, il n'était pas possible de lancer des recherches par *template matching* sur des séries de données réelles prolongées. L'analyse de familles de *repeaters* se faisait exclusivement à partir des séismes répertoriés dans les catalogues sismiques. Dans son étude, *A.Doucet (2021)* étudie les familles de *repeaters* de MPOM sur [2014-2021] en se basant uniquement sur les séismes déjà inventoriés. Il ne reste plus qu'à constituer les familles avec les fonctions Requake dédiées pour calculer la *NCC* entre deux événements déjà ciblés, puis à former les familles selon un algorithme de clusterisation donné, comme **UPGMA** (via [build\\_families](#), [scan\\_catalog](#)).

Or l'implémentation d'un nouvel algorithme de recherche par template dans le code de Requake démultiplie les chances de découvrir de nouveaux séismes inconnus, bien que présents dans les données, et qui pourront être intégrés aux futures familles. Ainsi, les recherches par catalogues sont précisées par le *template matching* : La section [3.1](#) s'attachera à effectuer la recherche de famille sur MPOM en utilisant pour la première fois la recherche par le dipptyque *catalogue + template matching*, en discutant le seuil de *NCC* appliqué par *A.Doucet (2021)*. Par la suite, afin d'obtenir des familles plus fournies et plus longues, la section [3.2](#) prendra comme référence la station BIM, en redéfinissant l'utilisation des seuils par *NCC* pour sélectionner des événements. Enfin, en section [3.3](#), une famille de BIM de référence sera retrouvée dans les données de FDF (transposition de famille), ce qui permettra *in fine* d'agrandir la famille en cherchant des nouveaux événements directement depuis FDF sur près de 2 décennies de données ([1998-2019]). La famille élargie sur FDF permettra enfin d'estimer le profil de déplacement au niveau de son [aspérité](#) et de donner une valeur estimée des niveaux de couplage associés.

## 3.1 Phase I : Débuts avec MPOM

La première station à avoir été utilisée sur le long terme est MPOM, en raison de la qualité de son spectre de bruit par rapport à DSD notamment. Elle présente un bruit moins important sur la zone d'intérêt. **N.B :** La figure [26](#) ne rend pas compte de cette comparaison, car le spectre de DSD n'a pas pu être calculé sur la plage où il est de moins bonne qualité. Pour MPOM, les données seront utilisées sur la période [2014-2022] (figure [27](#)).

Dans un premier temps, afin de retrouver les séismes identifiés dans le rapport de *A.Doucet [4]*, les paramètres de recherche ont été inchangés. Les premiers calculs par [scan\\_templates](#) ont permis de retrouver les familles précédemment identifiées. Ces familles sont rappelées en figure [28](#).

Afin de trouver des séismes non détectés par *A.Doucet(2021) [4]* à cause du seuil initial ( $NCC=0.95$ ), celui-ci a été abaissé à 0.9. Cette opération a permis de révéler plusieurs signaux similaires au template, auparavant ignorés. Cependant, cet abaissement du seuil a suscité une nouvelle interrogation : en effet, certaines valeurs de *NCC* après la décimation étaient inférieures à d'autres *NCC* d'événements qui n'avaient pas été détectés précédemment. Cela impliquait que si le critère de constitution des familles était uniquement basé sur le seuil, alors il fallait nécessairement inclure dans les familles déjà connues des événements qui n'avaient pas été considérés jusqu'alors.

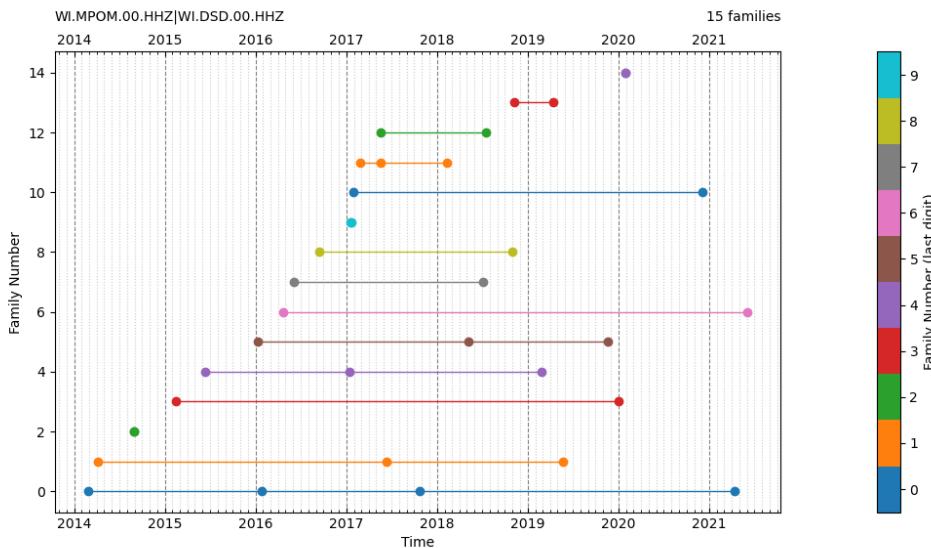


FIGURE 28 – Figure montrant l’étendue temporelle des familles ([plot\\_timespans](#), voir [glossaire](#)), avec les séismes *repeaters* marqués par des points. La figure montre les résultats bruts, en amont du tri, où les trois familles-singulet seront supprimées.

Pour ces raisons, un nouveau calcul de constitution des familles a été mené, mais cette fois-ci en abaissant le seuil à 0.8 (puis 0.7). Ces modifications ont mécaniquement fait augmenter le nombre de familles à 60 familles (voir en figure 29), uniquement sur le changement du seuil de *NCC*. Cet assouplissement du seuil de *NCC* a pour objectif de faire apparaître plus d’événements dans les familles et d’apporter *in fine* plus d’informations sur les ruptures correspondant à une [aspérité](#) donnée. Les nouvelles familles ont ensuite été triées et contrôlées visuellement, afin de s’assurer de l’absence d’événements non pertinents. De plus, elles ont été comparées aux familles trouvées avec des seuils de *NCC* plus hauts, afin de s’assurer de la conservation de tous les événements d’une analyse à l’autre.

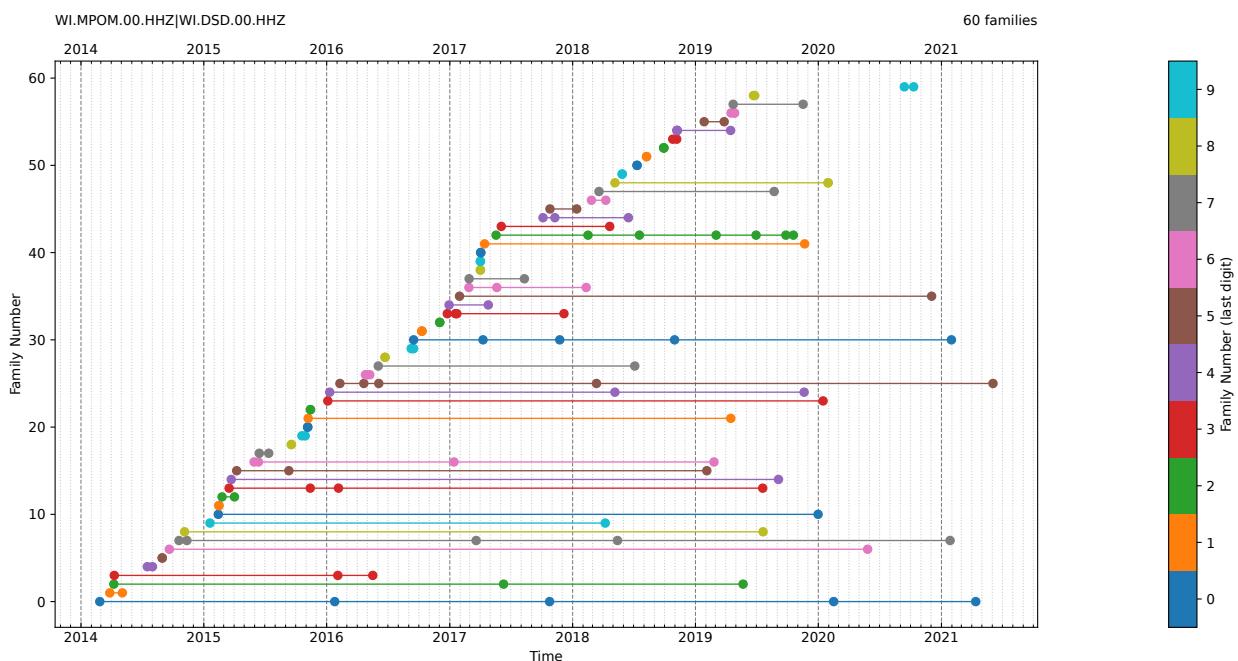


FIGURE 29 – Figure du [plot\\_timespans](#) pour MPOM avec un seuil de *NCC* de 0.8, illustrant la nette augmentation des effectifs de famille lors de l’abaissement sur seuil de corrélation. 60 familles sont présentes. Figure brute (en amont du tri des familles).



Pour finir, une comparaison explicite entre les évènements des familles détectés dans les catalogues ou par `scan_templates`, sur les 60 familles de la figure 29 a été effectuée une fois les familles créées. L'analyse par *template matching* est effectuée sur une période de temps identique à celle du catalogue utilisé pour la création des familles. En théorie, les évènements sont déjà connus. Mais outre les évènements déjà connus et répertoriés que l'on peut s'attendre à retrouver par *template matching*, il serait possible de détecter des séismes de la famille qui n'ont pas été répertoriés. Ainsi, il fallait pour chaque famille :

- A- Comparer le nombre d'évènements détectés à celui prévu par l'analyse du catalogue ;
- B- En cas d'écart : vérifier si les évènements supplémentaires figurent ou non dans les catalogues ;
- C- Identifier les séismes absents et rechercher les causes de non-détection.

Au total, sur les 60 familles (166 évènements) de MPOM identifiées, le calcul de template matching a été effectué avec une décimation de 1 puis de 2 pour s'assurer, en comparant les sorties du calcul, de la non-perte d'informations de détection (en plus d'un gain de 32 heures entre les deux calculs). Le bilan est le suivant : sur 166 évènements attendus d'après les catalogues,

- 168 ont été détectés par `scan_templates` (n'incluant pas exactement les 166 évènements des catalogues, mentionnés ci-dessus).
- On compte au total :
  - 7 séismes retenus par `build_families` mais non retenus par `scan_templates` dont 6 récupérés en supprimant la décimation (décimation = 1) et 1 non récupéré ;
  - 7 séismes non attendus par `build_families` mais détectés par `scan_templates` ;
  - 1 séisme non attendu par `build_families` et non présent dans le catalogue. Il s'agit d'un nouveau séisme qui n'avait pas été répertorié par l'**OVSM** .

Cette analyse permet de valider malgré tout une forte concordance entre les données des catalogues et l'analyse des données continues.

En raison de la destruction de MPOM, la période de [2021-présent (2025)] est devenue inexploitable. Or cette plage de temps représente un peu moins de 40% de la période de 2014 à nos jours, ce qui est un manque à gagner pour la précision des familles de séismes. La largeur de la fenêtre de temps est d'autant plus importante que des séismes orphelins peuvent se voir associer un nouveau séisme à 10 ans d'intervalle. En effet, des familles telles qu'elles sont identifiées ne sont qu'un aperçu instantané de ruptures existantes sur des échelles de temps géologiques. Plus le recul temporel est fort, et plus la construction de familles prend du sens. C'est la raison pour laquelle la station MPOM a été abandonnée au profit de BIM (Morne Bigot - Martinique), active sur la période [2014-présent (2025)]. L'exploitation des données de BIM fait l'objet de la section suivante.

## 3.2 Phase II : Premier élargissement avec BIM

La station BIM (Morne Bigot - Martinique) a été utilisée de façon similaire à la station MPOM. Lors des premières recherches, les paramètres identiques à ceux utilisés pour MPOM ont été appliqués, dans le but de maintenir la cohérence de la méthode et de confirmer la reproductibilité des résultats. Les premières observations ont permis de retrouver les mêmes groupes de séismes qu'auparavant. Cependant, il a été constaté que les données de MPOM manquaient à partir de 2021. Cette lacune est due à l'interruption permanente de la station durant cette période, rendant la tranche temporelle [2021-présent] inutilisable.

Pour continuer les études sur une période prolongée et ininterrompue, une autre station devait être choisie, alliant faible bruit et stabilité instrumentale. En service continu depuis 2012, la station BIM a prouvé être le meilleur choix : elle offre non seulement des données s'étendant sur une période plus longue que celles de MPOM, mais elle renforce aussi la solidité statistique de l'identification des familles grâce à la persistance des enregistrements jusqu'en 2025.

Dès les premières analyses effectuées sur BIM, une amélioration significative de la durée des familles et du nombre d'événements par famille a été constatée. L'incorporation de données plus actuelles (2021-2025) a permis d'élargir les périodes temporelles observées, certaines familles s'étendant sur une durée supérieure à dix ans. Les seuils de *NCC* ont été progressivement abaissés (de 0.9 à 0.6) pour révéler de nouvelles familles et augmenter le nombre d'événements par famille.

Cependant, la seule identification automatique ne permettait pas d'assurer la qualité des résultats. Un processus itératif de sélection manuelle a donc été mis en place : pour chaque groupe identifié, les événements ont été scrutés individuellement pour valider leur authenticité. Cette vérification s'appuyait sur l'examen des formes d'ondes, des valeurs de *NCC*, ainsi que sur la concordance temporelle et fréquentielle entre les événements. Les familles caractérisées par des signaux incohérents, excessivement bruités ou présentant des valeurs de *NCC* trop faibles ont été éliminées.

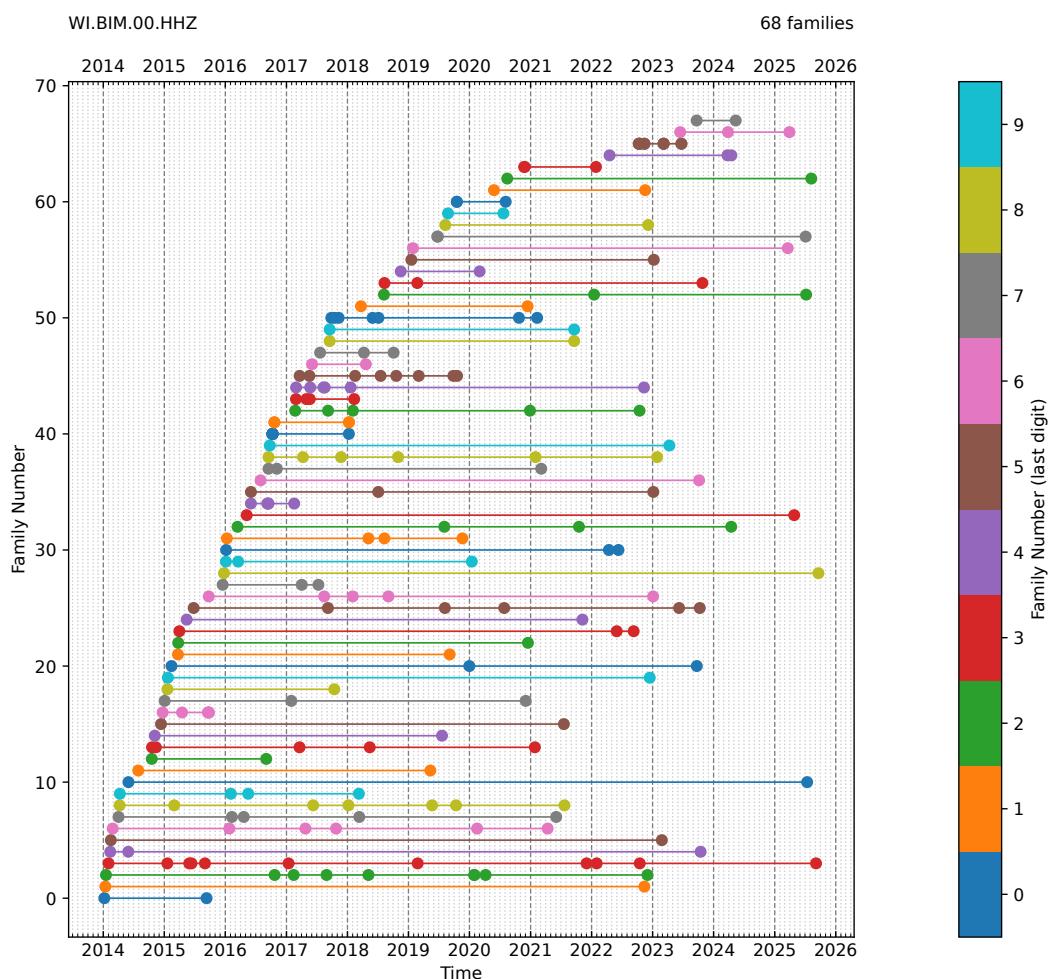


FIGURE 30 – Figure du ([plot\\_timespans](#)) pour BIM sur [2014-2025]. 68 familles sont présentes. La quantité de familles couvrant des longues plages de temps a pu être augmentée significativement par rapport à MPOM. L'une des familles atteint 11.6 ans d'envergure pour une dizaine d'événements.

Le tri a été réalisé en plusieurs itérations successives, en commençant par un seuil de  $NCC=0.6$ , avant d'augmenter graduellement à 0.7, 0.8 et finalement 0.9. Un tableau de suivi a été continuellement mis à jour (au format Excel), listant les familles validées et les événements liés. Cette méthode graduelle a facilité l'élimination des détections accidentelles tout en préservant les familles véritablement pertinentes.

Une fois cette étape de sélection finalisée, un nouveau fichier de familles au format *Requake* a été généré. Celui-ci ne contient que les familles considérées comme robustes, c'est-à-dire validées à la fois par le calcul automatique et par le contrôle visuel. Ce fichier sert désormais de base fiable pour les analyses ultérieures et pour les comparaisons avec d'autres stations.

Le résultat final de cette étape sur BIM présente un ensemble cohérent de 68 familles. Ces 68 familles, couvrant [2014–2025], serviront de base pour les recherches sur FDF, toujours dans l'objectif de prolonger encore la profondeur temporelle des analyses. Le passage, pour une famille d'événements, d'une station à une autre sera nommé *transposition* par la suite. L'ensemble des familles validées, et leur étendue temporelle, sont présentées en figure 30.

### 3.3 Phase III : Second élargissement avec FDF

À partir d'une base de 68 familles fiables (jusqu'à 11,6 ans ; fig. 30), l'étude vise à prolonger la profondeur temporelle. La station FDF (Fort-de-France, [1998–fin 2019], fig. 27) est donc retenue. Les familles de BIM sont transposées sur FDF par des recherches *scan\_templates* ciblées autour des événements déjà connus.

Cependant, la sélection des familles de BIM n'étant à ce stade pas encore achevée, et par souci de clarté, il a été décidé de commencer par se concentrer sur une unique famille. Pour une première transposition sur FDF, la famille BIM n°07 (11 événements sur ~12 ans, espacements réguliers) est choisie. Cette dernière est présentée en figure 31.

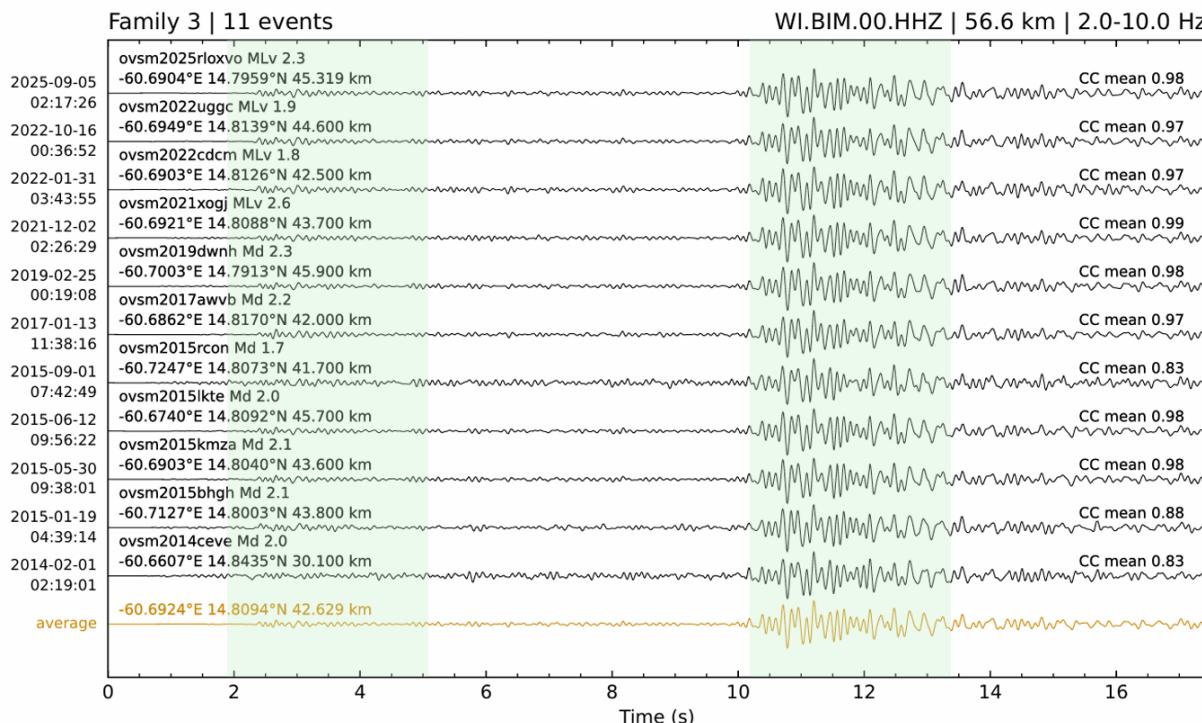
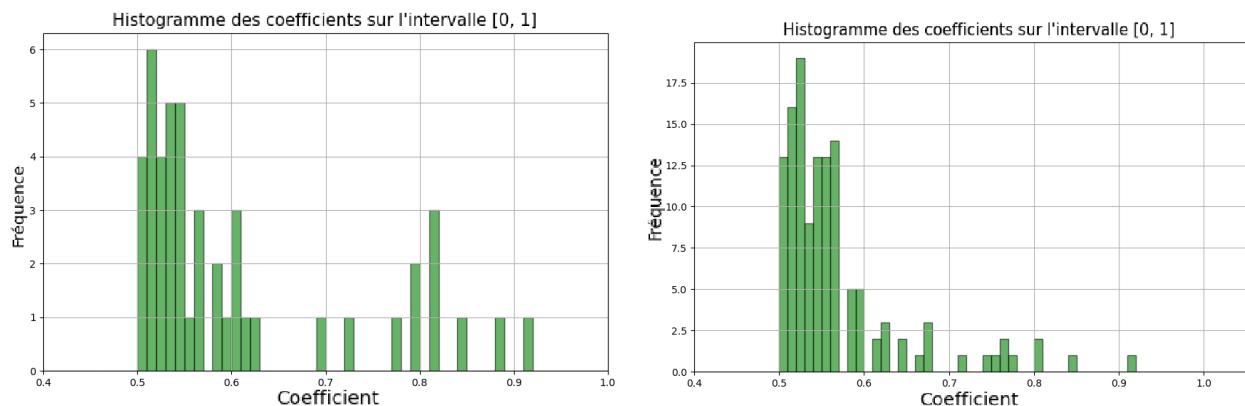


FIGURE 31 – Famille n°07 de la station BIM retenue pour référence. Les événements retenus ont été alignés par un critère de  $NCC$ . Les zones de surbrillance (en vert) indiquent l'emplacement des *ondes P* (gauche) et *ondes S* (droite) alignées. Leurs  $NCC$  avec le template moyen (en bas) est affiché sur la droite. Ces événements sont uniquement issus des catalogues.

Voici les étapes de sa transposition de BIM vers FDF :

- 1) *Template matching* sur les données continues de BIM : les 11 évènements de BIM sont retrouvés (qualité variable) ;
- 2) En plus de ceux retrouvés et issus des catalogues, 8 évènements supplémentaires sont détectés. Trois sont écartées (erreurs de calcul, courbes non réalistes, ou  $NCC$  incohérentes). Cinq sont conservées ( $NCC \in [0.68; 0.98]$ ). La famille BIM finale fait alors 16 évènements avec le template 7 en référence ;
- 3) Lors de la recherche sur FDF, les évènements de BIM postérieurs à fin 2019 étaient inaccessibles, la station ayant cessé son activité. Les autres retrouvent une correspondance : 11 évènements composent la première famille FDF par analogie. Il nous faut alors un template propre à FDF, l'ancien étant inadapté car la **fonction de Green** du trajet source - station change. Un nouveau template 07 de référence pour FDF est défini comme la moyenne des 10 premiers évènements (un d'eux étant trop bruité) ;
- 4) Avant d'entamer la recherche par *template* de FDF, il faut savoir que de 1998 à 2002, le champ "LOC" (*Location Code*, voir la documentation de référence, le SEED Reference Manual [ici](#)) n'existe pas, les requêtes FDSN utilisent donc G.FDF..BHZ (vs G.FDF.00.BHZ). De plus, La fréquence d'échantillonnage (20 Hz) a imposé d'adapter la chaîne de calcul à cause de la décimation que l'on ne pouvait plus utiliser. Le *Template matching* sur FDF de 1998 à 2019 avec un seuil de  $NCC = 0.50$  (volontairement bas) abouti à 47 détections, en vue d'un tri ultérieur.

En testant le *template* moyen des 47 évènements comme nouvelle référence, `scan_templates` retourne 128 détections. Le tracé est hétérogène et les  $NCC$  chutent : le *template* "dégénère" par moyenne de signaux dissemblables. Les histogrammes illustrent ce basculement : deux amas nets autour de 0,65 pour les 47 évènements initiaux dits "*a priori*" (fig. 32a) contre un profil quasi-exponentiel "*a posteriori*" (fig. 32b). Surtout, le *template* dégénéré ne retrouve plus certains évènements pourtant identifiés *a priori* (fig. 33).



(a) Histogramme de  $NCC$  sur les 47 évènements *a priori*. Deux clusters autour du seuil de  $NCC = 0.65$  sont distinguables, et pertinents pour séparer les évènements pertinents des non-pertinents.

(b) Histogramme des  $NCC$  sur les 128 évènements *a posteriori* : les deux clusters disparaissent. Le profil "exponentiel" suggère une perte d'information, le bruit ressemblant plus facilement au template dégénéré.

FIGURE 32 – Histogrammes de  $NCC$  (a) *a priori* (47 ev.) et (b) *a posteriori* (128 ev.).

Afin de passer outre cette « dégénérescence », seules les détections *a priori* (associés au template fondé sur les familles certifiées sur BIM, transposées sur FDF) sont retenues. La famille finale de FDF n° 07 sera donc bornée aux 47 évènements sélectionnés initialement.

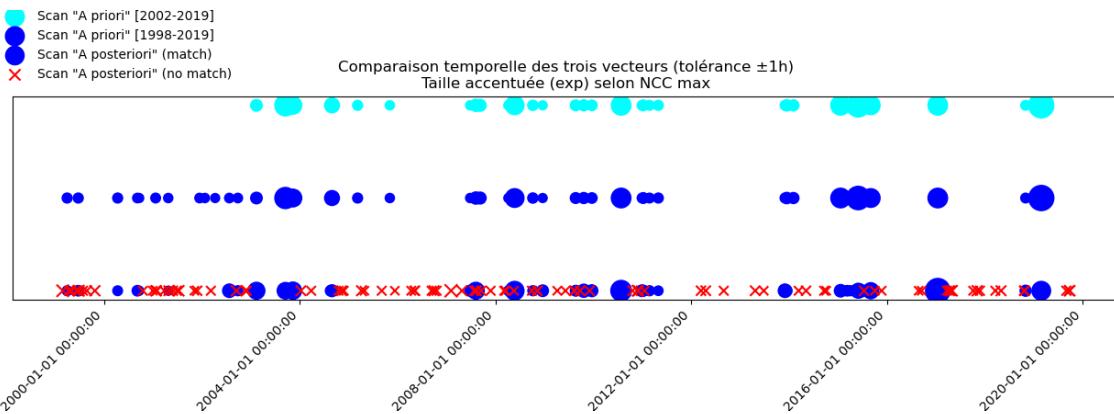


FIGURE 33 – Bleu clair - haut : détections *a priori* [2002–2019] avant correction du problème d’importation des données FDSN. Bleu marine - milieu : détections *a priori* complètes [1998–2019] ; les 47 événements de la famille de référence y figurent. Bleu marine et rouge - bas : détections *a posteriori*; croix rouges = fausses, ronds bleus = vraies. La taille des marqueurs est proportionnelle à leurs *NCC*. La plupart des fortes *NCC* sont retrouvées, mais de nombreuses fausses détections ont aussi une *NCC* non négligeable.

Les estimations de déplacement de la plaque modélisées par une loi physique de type Nadeau et Johnson [8], et implémentées par `plot_cumulative` (exemple en figure 39) dépendent du nombre d’événements et de leurs magnitudes. Afin d’encadrer le déplacement réel, deux ensembles sont définis à partir des 47 événements FDF :

$\mathcal{F}_{inf}$  (Famille stricte) : Un seuil initial de  $NCC > 0.65$  puis un contrôle visuel aboutissent à 20 événements (voir annexes). Les fig. 36a, 36b, et 37 illustrent la sélectivité par rapport aux 47 *a priori*.

$\mathcal{F}^{sup}$  (Famille moins stricte, contenant la précédente) : Avec comme critère ( $NCC > 0.6$ )  $\cup ((NCC > 0.4) \cap (NCC_s > 0.7))$  (la  $NCC_s$  étant détaillée fig. 38) on obtient 30 événements. Le ciblage des ondes S donne systématiquement une  $NCC_s$  plus élevée que la  $NCC$  mais sans rapport de proportionnalité (fig. 34, 35). D’où le choix d’un seuil combiné plutôt que sur  $NCC_s$  seule, qui aurait conduit à accepter des événements aux  $NCC$  extrêmement basses. **N.B** : familles *inf/sup* visualisables [ici](#) (annexes GitHub).

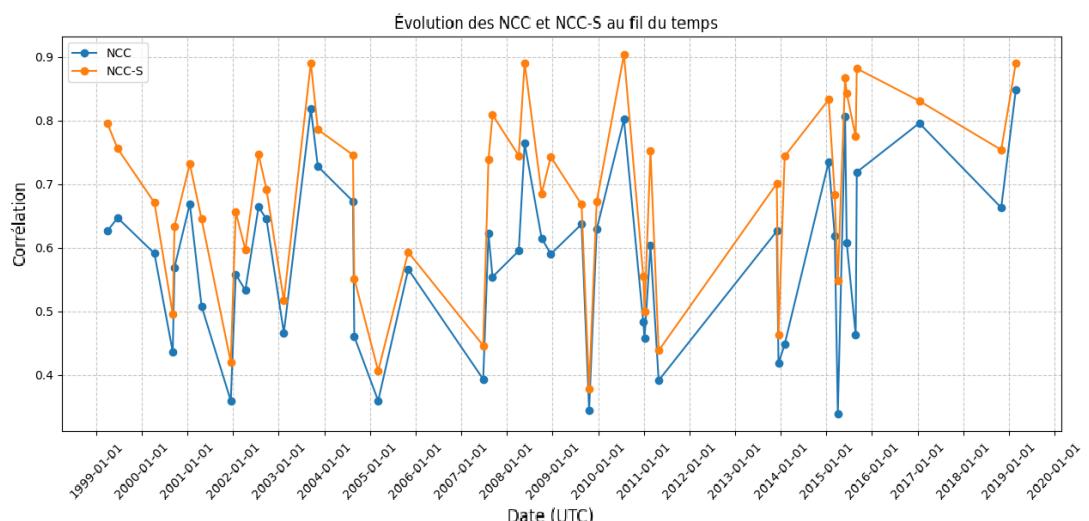


FIGURE 34 – Evolution des *NCC* des 47 événements, avant (bleu) et après (orange) le ciblage du calcul de la *NCC* sur les ondes S. L’amélioration des valeurs de *NCC* est omniprésente. On voudrait néanmoins comprendre la logique de cette amélioration, d’où la figure 35.



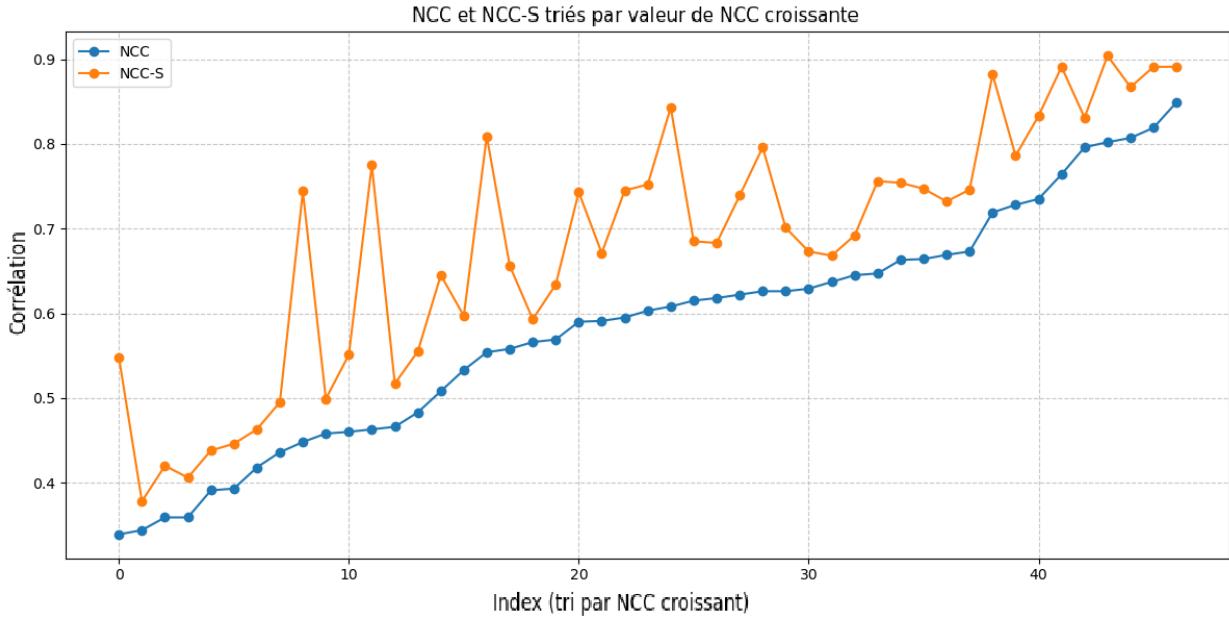
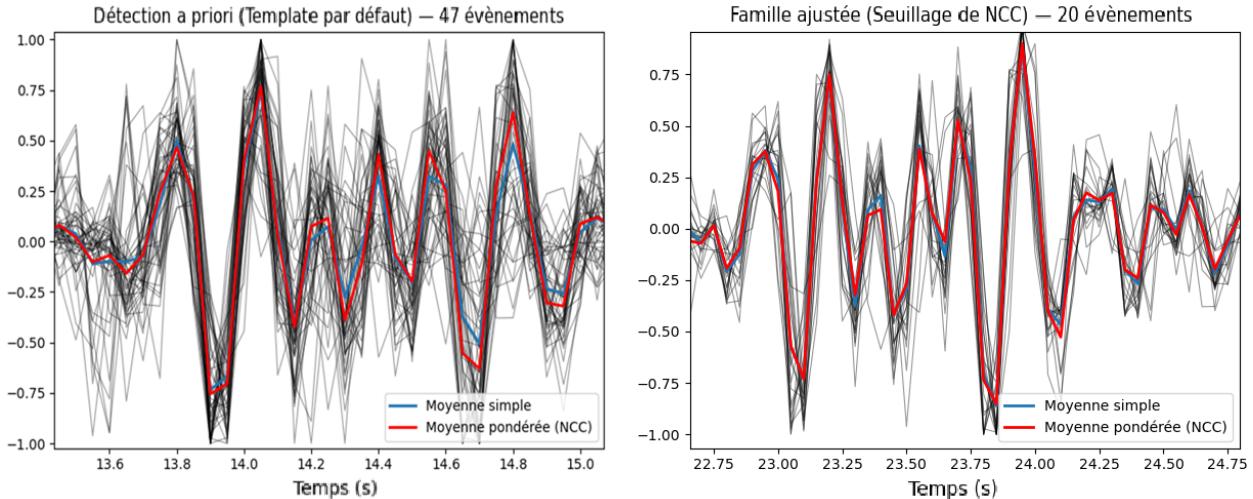


FIGURE 35 – Réorganisation des évènements de la figure 34 par ordre de  $NCC$  croissante. Les écarts entre  $NCC$  et  $NCC_s$  ne sont pas liés à la valeur de la  $NCC$ . Un seuillage sur la  $NCC_s$  seul n'est donc pas pertinent car une valeur de  $NCC_s$  donnée peut correspondre à des valeurs de  $NCC$  très disparates. D'où l'idée d'un critère combiné pour la famille  $\mathcal{F}^{sup}$ .



(a) Superposition des enveloppes alignées pour les 47 évènements *a priori* autour des ondes S. Le template moyen (moyenne arithmétique) associé est affiché en rouge, et la moyenne pondérée est affichée en bleu. Le "bruit" de l'ensemble est notable

(b) Superposition des enveloppes alignées pour les 20 évènements sélectionnés par seuillage sur la  $NCC$  + vérification visuelle, (ç.à.d la famille  $\mathcal{F}_{inf}$ ) autour des ondes S. Le template moyen (moyenne arithmétique) associé est affiché en rouge, et la moyenne pondérée est affichée en bleu.

FIGURE 36 – Superposition des traces alignées avant/après sélection pour la famille  $\mathcal{F}_{inf}$ . La variance (dispersion) de l'ensemble en 36b est bien moindre qu'en 36a. Le calcul de la moyenne pondérée augmente modestement les détails de haute fréquence du signal moyenné. La variation par rapport à la moyenne arithmétique se distingue surtout aux extrema.



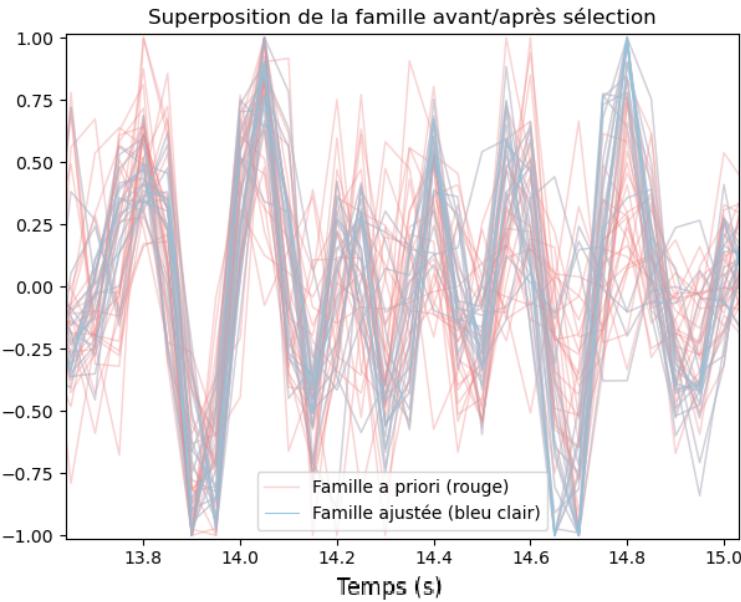


FIGURE 37 – Coloration des traces selon leur appartenance : traces *a priori* (en rouge), et traces sélectionnées pour la famille  $\mathcal{F}_{inf}$  (en bleu clair). Le sous-ensemble sélectionné en bleu possède un bruit (au sens de variance intra-familial) bien plus faible que le bruit total associé aux 47 événements. Autrement dit, pour une abscisse de temps donné, l’ensemble des ordonnées des traces sélectionnées est bien moins dispersé que l’ensemble des ordonnées sur les 47 traces *a priori*.

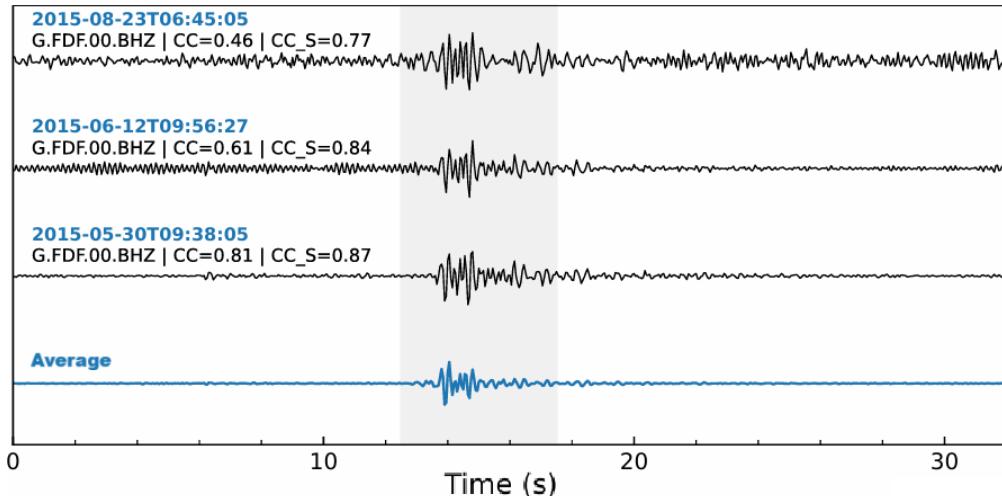


FIGURE 38 – Exemple de calcul de NCC sur plusieurs traces de la famille *a priori* de 47 événements sur FDF. On compare la *NCC* entière à la *NCC<sub>s</sub>* (centrée sur les ondes S). Les deux *NCC* sont affichées sur la gauche des traces. Toutes les valeurs sont en fig. 34.

### 3.4 Obtention et exploitation des magnitudes

Les magnitudes des familles  $\mathcal{F}_{inf}$  et  $\mathcal{F}^{sup}$  sont calculées par SourceSpec avec un seuil de  $SNR \geq 5$  en raison des faibles magnitudes. Certaines traces incohérentes de stations lointaines ou bruitées sont exclues ; toutes les magnitudes ne sont donc pas obtenues avec SourceSpec. Les valeurs de magnitudes de SourceSpec étant peu similaires aux magnitudes des catalogues, la priorité a été donnée aux magnitudes de SourceSpec (c.à.d. aux valeurs expérimentales). Ainsi, une régression semi-empirique a été construite entre les magnitudes  $M_W$  issues de Sour-



ceSpec, et les amplitudes instrumentales des séismes. L'amplitude instrumentale (en  $m.s^{-1}$ ) est obtenu à partir de l'amplitude numérique, via la fonction `remove_response` d'Obspy. La régression par moindres carrés de la forme  $M_W = a.log(A_{instrumentale}) + b$  est ajustée pour chaque famille :  $r^2 = 0.798$  pour la famille  $\mathcal{F}_{inf}$  et  $r^2 = 0.830$  pour la famille  $\mathcal{F}^{sup}$ . Ces régressions (fig.40 et 41) permettent d'estimer les  $M_W$  manquants.

La dernière étape consiste en la construction d'un catalogue des événements des deux familles, et de leurs magnitudes associées. La figure 39 superpose les `plot_cumulative` (graphe par paliers) des deux familles dont l'écart est notable. Outre l'ajout de nouveaux événements aux familles, le graphe de déplacement couvre une plage de temps trois fois plus importante que pour *A.Doucet (2021)* [4]. La pente moyenne de la famille  $\mathcal{F}^{sup}$  est proche de  $2 \text{ cm} \cdot \text{an}^{-1}$ , cohérente avec la vitesse de la plaque américaine (borne supérieure plausible). Cette démonstration repose sur une famille ; or 68 familles robustes issues de BIM restent transposables sur FDF, ouvrant la voie à un échantillonnage régional plus fin des vitesses de *glissement*. Les *couplages* pour les deux familles sont de 4% et 51% pour les familles  $\mathcal{F}_{inf}$  et  $\mathcal{F}^{sup}$  respectivement, ce qui exclut d'emblée les hauts couplages, également rejettés par *van Rijsingen* [1], qui évoque plutôt un *couplage* entre 0 et 10% pour le Nord-Est de la Martinique.

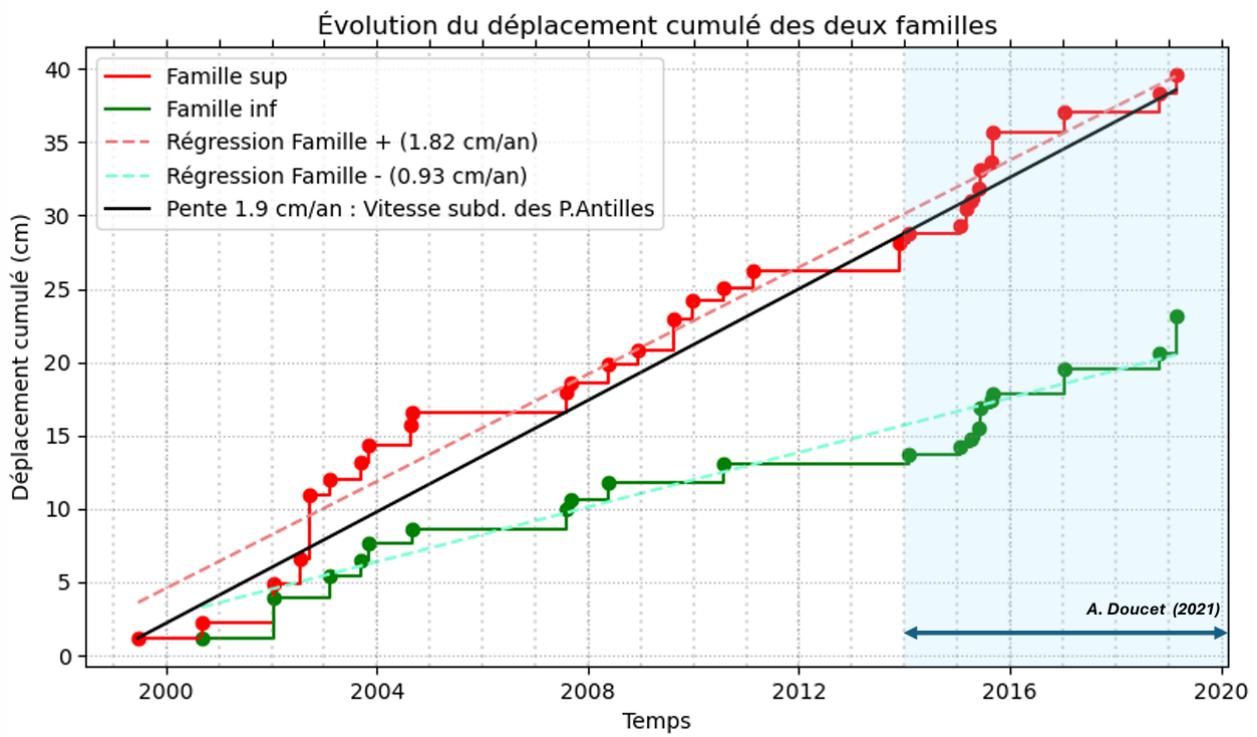


FIGURE 39 – Superposition des `plot_cumulative` des familles  $\mathcal{F}^{sup}$  (rouge) et  $\mathcal{F}_{inf}$  (vert foncé). Sur chacune des familles, des droites de régression sont affichées en pointillés pour estimer la vitesse moyenne de déplacement. La droite en noir correspond à une vitesse de déplacement de  $19 \text{ mm/an}$ , c'est-à-dire la vitesse de la subduction des Petites Antilles. Sur la plage considérée de [1999-2019], la vitesse moyenne de la  $\mathcal{F}^{sup}$  est très proche de celle de la subduction. On voit nettement que sur l'intervalle [2014-2021] (en bleu) étudié par *A.Doucet* dans [4], le coefficient directeur local est bien plus élevé, d'où l'intérêt de cette étude sur une période prolongée. Pour une vitesse de subduction de  $19 \text{ mm/an}$ ,  $\mathcal{F}^{sup}$  est associée à un couplage de 4% et  $\mathcal{F}_{inf}$  est associé à un couplage de 51%.

$$\text{Régression linéaire : } M_w = a \cdot \log_{10}(A) + b \\ r^2 = 0.798$$

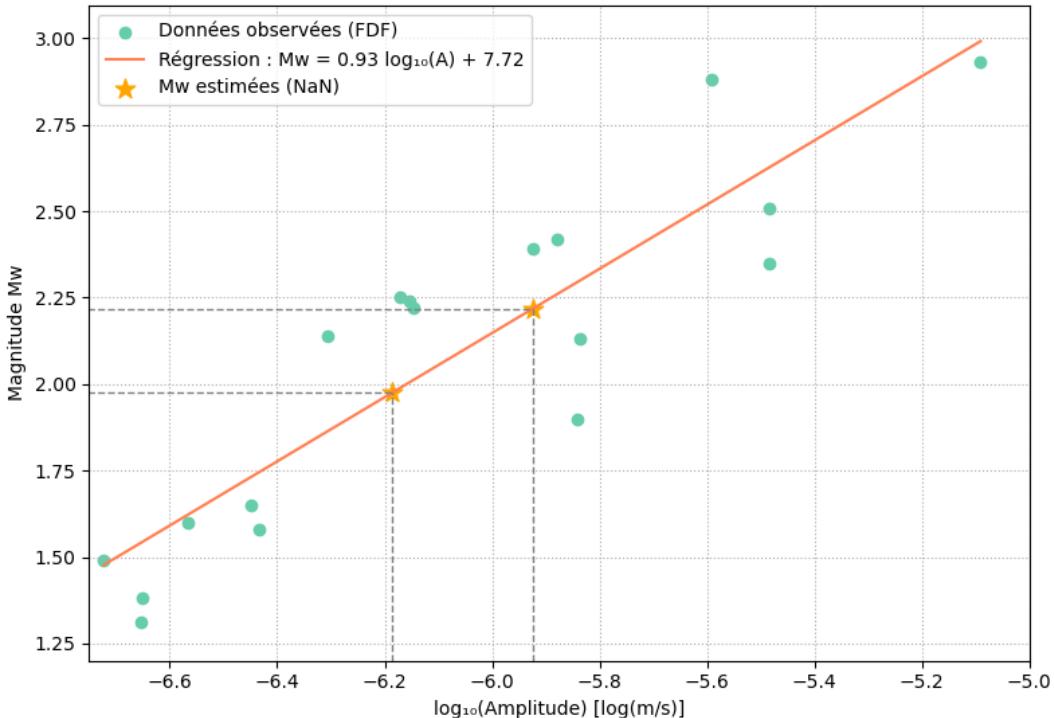


FIGURE 40 – Régression linéaire entre  $M_w$  et  $\log(A)$  par moindres carrés, sur les événements de la famille  $\mathcal{F}_{inf}$ .  $r^2 = 0.798$ . Les segments en pointillés désignent la lecture sur la droite de régression des  $M_w$  à partir de l'amplitude d'évènements dont la magnitude n'avait pas été calculée sur SourceSpec (2 évènements concernés).

$$\text{Régression linéaire : } M_w = a \cdot \log_{10}(A) + b \\ r^2 = 0.830$$

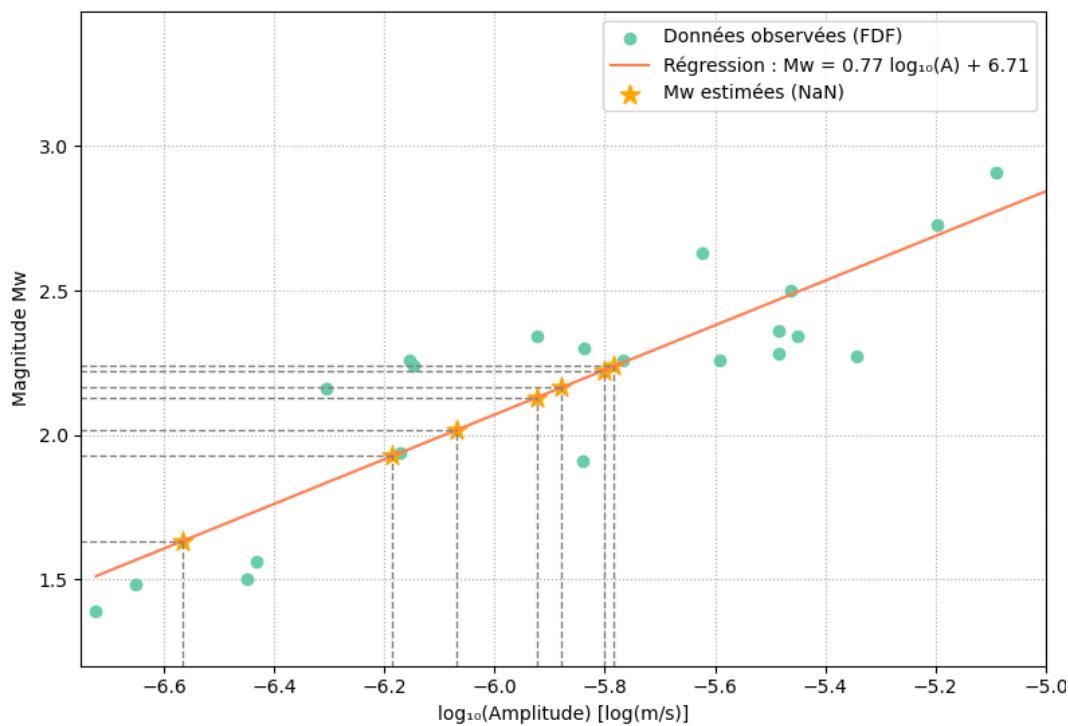


FIGURE 41 – Régression linéaire entre  $M_w$  et  $\log(A)$  par moindres carrés, sur les événements de la famille  $\mathcal{F}^{sup}$ .  $r^2 = 0.830$ . Les segments en pointillés désignent la lecture sur la droite de régression des  $M_w$  à partir de l'amplitude d'évènements dont la magnitude n'avait pas été calculée sur SourceSpec (7 évènements concernés).



# Discussion

## 4.1 Méthode de constitution des familles et implications

La constitution des familles est une étape fondamentale sur laquelle reposent toutes les analyses ultérieures de ce projet (Magnitudes de moment, déplacements de la plaque tectonique, et in fine, la compréhension du [couplage](#)). Or cette étape est soumise à plusieurs facteurs d'incertitude : les familles ne sont pas toujours identiquement restituées d'une station à l'autre, les évènements sont attribués aux familles avec une certaine probabilité d'erreur, les données sont parfois non-fiables, voire indisponibles, et les templates de recherche ne sont pas toujours optimaux. Ces différents éléments d'incertitude doivent être appréhendés afin d'optimiser ou de relativiser la qualité des familles analysées.

Premièrement, les stations étant positionnées sur des sites parfois très éloignés ou de nature différente et n'ayant pas les mêmes performances spectrales, certains évènements visibles sur une station ne le sont pas sur une autre station. Ainsi, l'étude d'une famille sur une unique station introduit un biais de sélection dont on peut tenter de s'affranchir en regardant les données des stations alentour (si elles existent). Une approche visant à fusionner les données de plusieurs stations améliorera la précision des familles. Cet enjeu a été rencontré lors de la transposition des familles *MPOM* → *BIM*, puis *BIM* → *FDF*. Il peut être vu, au contraire, comme un critère pour juger de la pertinence d'un évènement : plus sa visibilité est avérée d'une station à l'autre, plus sa contribution à la famille est plausible.

En outre, les familles obtenues et étudiées dans le cadre de ce rapport ont été obtenues principalement par seuillage de leur valeur de corrélation croisée normalisée (*NCC*), mais la valeur de ce seuil a été sensiblement abaissée au fil des réflexions. Cette valeur ne peut être fixée sans prendre en compte la méthode de constitution des familles de Requake (via [build\\_families](#), par exemple [UPGMA](#), voir la section 1.4), ou la variabilité du bruit au sein des évènements d'une famille supposée. L'analyse de la *NCC* sur les [ondes S](#) uniquement permet d'affiner l'analyse par rapport au template entier, mais ce ne peut être une fin en soi en raison de la variabilité possible des [ondes P](#), souvent moins visibles en raison de leur faible énergie. L'analyse "à l'œil" a souvent permis de trancher sur certains évènements ambigus, mais cette vérification au cas par cas n'a de sens que si aucun de ces cas ambigus n'est rejeté d'avance par un seuil de *NCC* trop haut. En effet, une *NCC* basse est parfois accidentelle, et certains signaux bruités observés de près sont bien plus similaires à un template que d'autres évènements moins bruités mais intrinsèquement moins similaires au *template*. Afin d'éviter les procédures de sélection *ad hoc* et peu facilement généralisables, peut-être une approche par extraction de features et d'algorithmes de clusterization pourrait-elle permettre de systématiser les détections et la qualité des familles générées ? Il faudrait pour cela décider quels *features* significatifs extraire, sachant que cela ne pourrait-être ni l'amplitude des ondes P ou S, ni le niveau de bruit, ni le temps de détection. Si une telle liste de *features* est établie, la méthode de classification par *random forests* (forêts aléatoires) pourrait offrir une possibilité de classification résiliente au bruit et à la diversité des signaux au sein d'une même famille.

Troisièmement, les données disponibles ne sont pas toujours exploitables, pour plusieurs raisons. Les données sont transmises des stations sismiques vers les serveurs via des réseaux de satellites. Or ces satellites sont parfois saturés ou momentanément inopérants, ce qui rompt la continuité des données (les données ne sont pas nécessairement perdues, mais plus complexes à extraire). Les stations sismiques sont aussi parfois placées en maintenance, afin d'en améliorer les capteurs, ce qui perturbe momentanément les données, ou change leur format. Enfin, des anomalies ponctuelles (décentrage de masse, défauts divers), ou l'arrêt simple de la production de données (remplacement par une station plus récente, ou cas accidentel comme MPOM). Ces irrégularités incitent à combiner les données des différentes stations des Caraïbes afin d'éclaircir les trous de données tout en rendant les détections plus fiables, comme évoqué en début de section. Elles incitent aussi à une résilience des algorithmes de template matching face aux changements des formats de données ou de requêtes (typiquement la fréquence d'échantillonnage, ou le champ "LOC", *Location Code*), d'autant plus probables pour des calculs sur plusieurs décennies. L'analyse technique des performances du réseau de stations WI pourra être consultée plus avant dans la publication suivante : *Anglade et al.*, <https://doi.org/10.5194/adgeo-40-43-2015>.

Enfin, la constitution d'une famille de séismes répétés est une opération délicate car elle repose sur un template dont on fait l'hypothèse qu'il est parfaitement représentatif de la famille. Cette confiance a priori est nécessaire car les catalogues fournissent parfois peu d'événements, et lors de la recherche sur les données continues, il faut faire l'hypothèse que le template est représentatif de la famille. Cette représentativité peut être partiellement contrôlée via le calcul d'obtention du template : par défaut obtenu comme une moyenne arithmétique des événements répertoriés, il pourrait être obtenu par une pondération des événements selon leur plausibilité (qui pourrait être déterminée manuellement, et non par la valeur de *NCC* avec le template, ce qui introduirait un biais d'auto-réalisation). Cette étape est importante car un template, censé refléter toute la famille, est utilisé alors même que la famille est inconnue. Il faut une correspondance de la famille vers le template, et non l'inverse ! De plus, lorsque la famille s'agrandit, le template ne doit pas perdre en finesse, et le phénomène de lissage par moyennage ne doit pas faire dégénérer la détection, en raison d'une perte d'information (voir les résultats de 3.3). Un nouvel algorithme de détection pourrait être proposé afin de trouver un compromis entre la précision du template restreignant les détections, et sa représentativité d'une famille et sa propension à détecter plus d'événements qu'il n'en faut.

Les diagrammes en paliers générés par requake `plot_cumulative` associent à chaque événement un déplacement de la plaque en *mm* ou *cm*. Ces diagrammes associent donc à une plage de temps, le déplacement estimé de la plaque sur cette plage de temps, par somme de tous les déplacements associés aux événements successifs. Or les facteurs d'incertitudes tels que mentionnés dans les paragraphes précédents mènent à retirer ou à ajouter des événements selon la rigidité des critères de sélection. Cela amène, suivant les procédés, à faire varier le nombre d'événements dans les familles, et donc les estimations de déplacement des plaques. Par principe de précaution, et afin d'éviter de proposer des valeurs incohérentes, les déplacements tectoniques sur une période donnée pourraient être analysés par une logique de borne inférieure et supérieure : Tous les événements les plus plausibles, apparaissant quoiqu'il arrive dans une famille seront comptabilisés et interviennent dans un déplacement minimal de la plaque. Inversement, les familles les plus optimistes pourraient avoir tendance à surévaluer le déplacement de la plaque. Plutôt que de considérer un déplacement unique associé à une période, on pourrait proposer une plage de déplacement représentant l'incertitude sur les événements, avec un déplacement limite minimal et un déplacement limite maximal.

Nous avons donc mis en évidence la possibilité de construire, pour chaque famille de *repeaters* un intervalle de confiance sur le déplacement, et donc sur le **couplage**. Tandis que *A. Doucet (2021)* [4] se focalisait sur les familles issues des catalogues, nous avons enrichi les familles, et précisé les profils de déplacements cumulés, tout en augmentant par 3 la plage de temps d'étude. Cela permet de prendre un recul plus grand sur le comportement de **glissement** au niveau des **aspérités**, d'autant plus que les intervalles de glissement cumulés comme la figure 39 peuvent être réalisés pour chacune des 68 familles recueillies sur BIM et de bonne qualité, ou sur toute autre famille jugée pertinente aux Petites Antilles à l'avenir. Il est également à noter que la famille de référence utilisée pour aboutir à  $\mathcal{F}_{inf}$  et  $\mathcal{F}^{sup}$  dans la troisième partie du rapport (BIM\_07), est située à une profondeur relativement grande (42,6 km, voir l'emplacement en 43, et le schéma de la subduction sur la transition entre *repeaters* et *slow-slip* en figure 42), c'est-à-dire à la base de la zone couplée, sur l'interface entre le coin mantellique et la plaque plongeante (voir le schéma en 4). En conséquence de quoi, il est attendu que le couplage soit très faible, voire quasi-nul, comme c'est le cas pour  $\mathcal{F}^{sup}$  en figure 39. Pour cette raison, il est serait intéressant de poursuivre l'étude des profils de glissement à des profondeurs plus superficielles, notamment pour les familles de BIM situées au Nord-Est de la famille de référence de la figure 43, afin de mieux comprendre la relation entre couplage et profondeur. De même, afin de mieux reconstituer les dynamiques de couplage pour différentes profondeurs sur la subduction, il serait judicieux de renforcer la fiabilité des localisations de famille de *repeaters*, dans la mesure où elles impactent directement l'interprétation des valeurs de couplage.

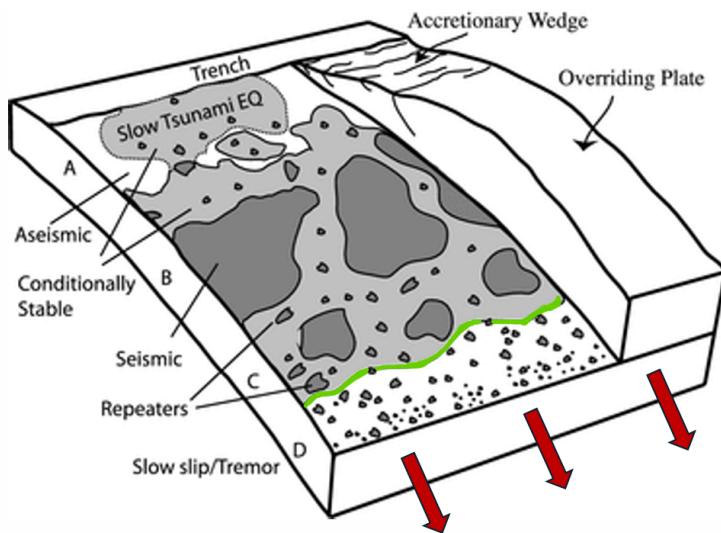


FIGURE 42 – Schéma de la transition sismo-géologique en profondeur au niveau de la subduction. La ligne verte sépare la zone d'apparition de séismes ou *repeaters* (en haut) de la zone où on peut observer du *slow-slip* (en bas). BIM\_07, à une profondeur de 42,6 km se situe proche de cette délimitation. L'étude des autres familles devra donc s'interpréter suivant cette profondeur de façon à mieux comprendre les profils de glissement. Adapté de <https://doi.org/10.1029/2011JB009133>.

Le modèle reliant déplacement et magnitude pourra également être adapté en fonction de l'estimation des chutes de contraintes ( $\Delta\sigma$ , en MPa). Alors que le modèle de *Beeler* aura tendance à augmenter la hauteur des "marches" sur les profils de glissement, le modèle d'*Eshelby* pourra être approfondi si les régressions de SourceSpec permettent de mieux estimer les paramètres du modèle. La vitesse de glissement moyenne de la famille  $\mathcal{F}^{sup}$  converge de façon nette vers la vitesse relative de la subduction de 19 mm/an, ce qui constitue une

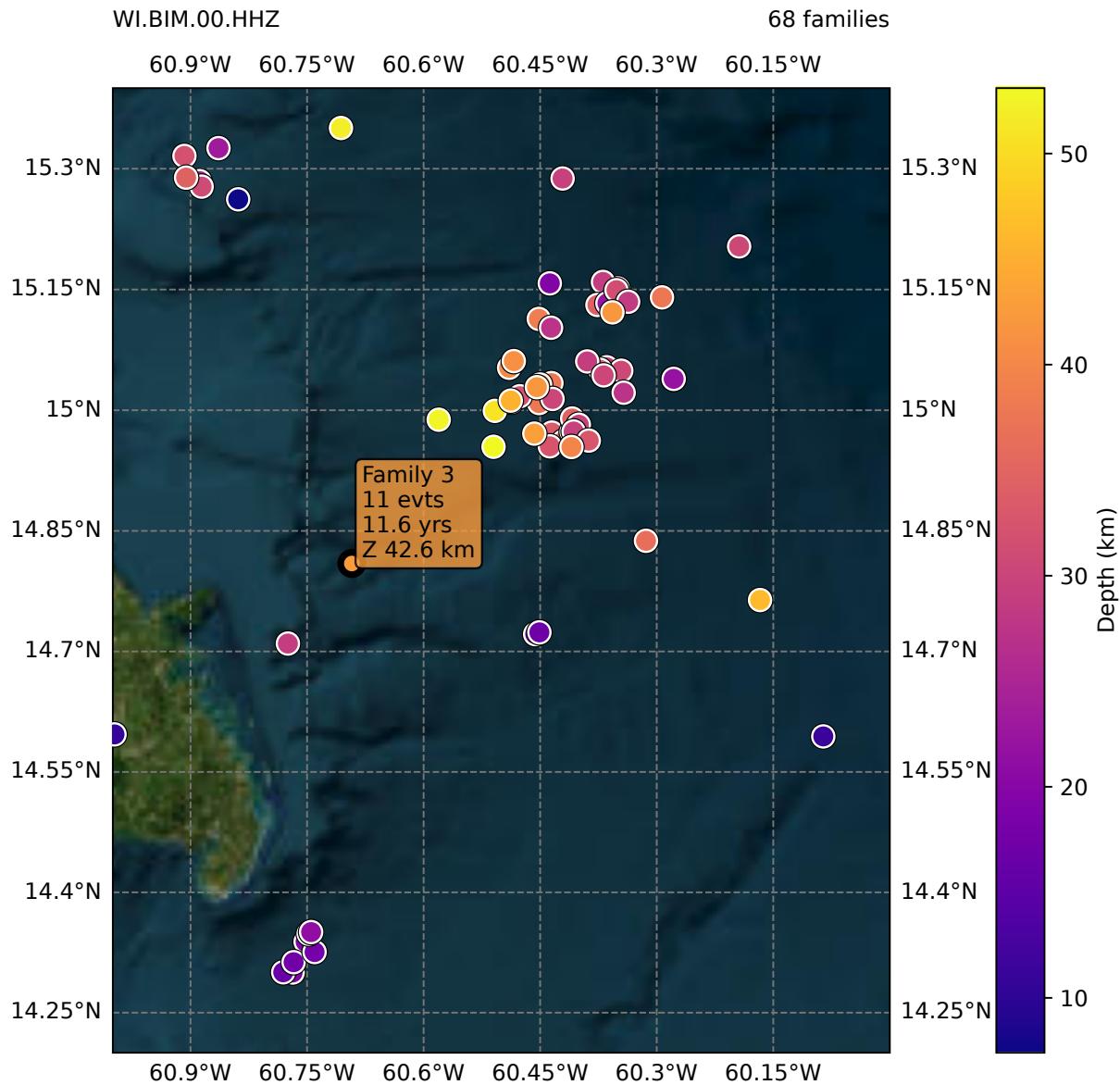


FIGURE 43 – Position des barycentres (arithmétiques) des évènements des 68 familles retenues sur BIM, après extraction/sélection depuis les catalogues, puis un tri après un `scan_templates`. Le Sud-Est de la Martinique apparaît en bas à gauche. La famille mise en valeur dans l’encadré est la famille prise comme référence dans la troisième partie, pour arriver à la constitution de  $\mathcal{F}_{inf}$  et  $\mathcal{F}^{sup}$ , soit BIM n°07 (numérotée par un 3 après le tri des familles de BIM effectué ultérieurement).

avancée par rapport à *A. Doucet(2021)*, où l’étude étant trop restreinte dans le temps, la vitesse locale de déplacement n’était pas représentative de la tendance générale mieux visible sur plusieurs décennies. Ces résultats suggèrent l’importance que pourrait avoir un suivi **GNSS** au niveau de ces sources de repeaters. Enfin, le profil des graphes de glissement comporte plusieurs motifs de clusters de *repeaters* séparés de plusieurs années : ainsi par exemple, quasiment aucun séisme n’a lieu sur la figure 39 entre 2004-2008 et 2011-2015. Cette distribution temporelle spécifique évoque les travaux de *Naoki Ushida* sur le *slow slip*, ou glissement asismique (sans génération de séismes), et la répartition des *repeaters* est associée à un taux de glissement corrélé temporellement à des séismes de haute magnitude qui peuvent les précéder (<https://www.researchgate.net/publications/292212226>).



## 4.2 Discussion sur les *antirepeaters*

Tandis que les séismes répétés, ou *repeaters*, sont associés à un même template, et à la rupture d'une même faille, on appelle *antirepeaters* des séismes dont la corrélation avec le template serait proche de -1. Autrement dit, les *antirepeaters* seraient des séismes totalement inversés par rapport au template initial. Au-delà de la particularité mathématique, ces *antirepeaters* possèdraient une signification théorique précise. Localisés sur la même source que, ou dans sa proximité, ils seraient associés à la même hétérogénéité locale et correspondraient à un déplacement local des roches en sens inverse du sens de déplacement de la faille, par rétractation en raison de l'entraînement par les roches environnantes ayant une vitesse de déplacement différente [15]. Bien qu'encore peu considéré, ce phénomène pourrait éclaircir les dynamiques tectoniques telles que celle étudiée dans les Petites Antilles. Les algorithmes utilisés pour la détection d'événements dans une version modifiée de `scan_templates` n'ont pas considéré la potentielle détection d'*antirepeaters*, et bloquent les détections pour les valeurs de *NCC* négatives. En effet, elles étaient trop souvent attenantes de vraies détections positives, et n'étaient que des artifices de calculs propres à des lobes secondaires de *NCC* autour d'un pic central très proche de 1. Cependant, sur la dernière famille analysée pour FDF, un événement dont le profil semblait quasiment inversé par rapport au template a été identifié. Afin de s'assurer qu'il n'y en avait pas d'autres, un `scan_templates` a été lancé sur la même période, mais avec un template inversé (multiplié par -1). La valeur de *NCC* était plus forte en valeur absolue avec l'anti-template, qu'avec le template initial, comme on peut l'observer en figure 44. Le souhait d'ajouter aux algorithmes existant de Requake la possibilité de détecter des *antirepeaters* pourrait être envisagée à terme, bien qu'une recherche par `scan_templates` avec un template inversée soit déjà tout à fait possible et équivalente à une détection par valeurs de *NCC* négatives. La détection et l'interprétation des *antirepeaters* comme des déplacements locaux opposés à la direction de la plaque pourraient partiellement compléter la logique des représentations de déplacement par paliers, en incluant des marches "négatives", et affineraient les graphes générés par `plot_cumulative`, en les faisant plus encore correspondre au déplacement réel de la plaque Américaine au niveau des irrégularités à l'origine des *repeaters*.

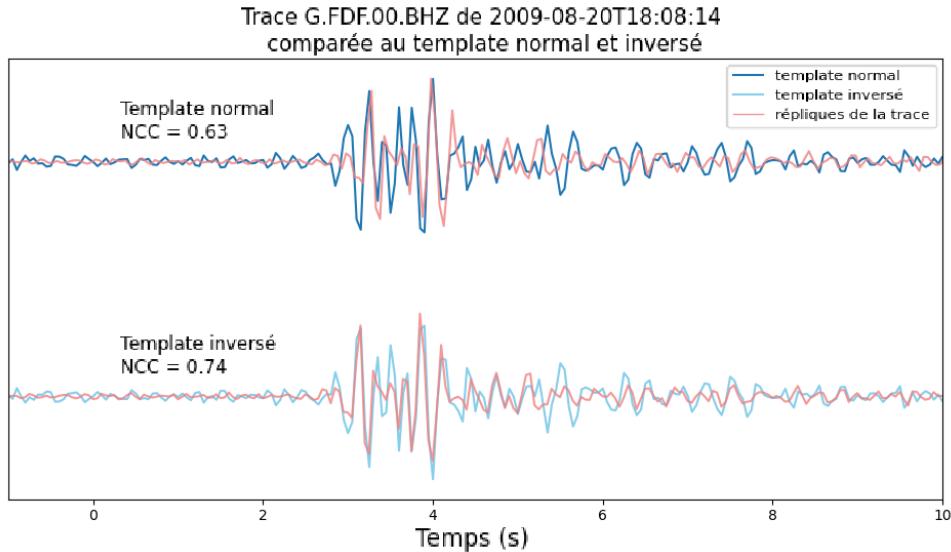


FIGURE 44 – **En haut :** Template (bleu foncé) superposé avec la trace d'un événement sismique (rose) ressemblant à un *antirepeater*, après alignement. **En bas :** Template inversé (bleu clair) superposé avec la trace (rose) du même événement sismique, après alignement. La *NCC* est plus élevée en valeur absolue avec l'*antitemplate* qu'avec le template *normal*.



# Conclusion

Alors que le comportement de la subduction des Petites Antilles demeure partiellement incompris, l'étude des familles de *repeaters* permet d'estimer le comportement de la plaque plongeante au niveau de certaines **aspérités** localisées. Après avoir construit une chaîne de calcul de détection de *repeaters* par *template matching*, et l'avoir intégrée dans Requake, l'exploration des données continues sur plusieurs stations de références (MPOM, BIM et FDF notamment) a permis d'augmenter considérablement le nombre de *repeaters* considérés, et de compléter ce faisant les données des catalogues sismiques. Une nouvelle base de familles a pu être étudiée sur MPOM et BIM, où l'ajustement des seuils de détection par *NCC* a permis d'affiner la sélection des événements sismiques. Dans la mesure où la qualité d'une famille de *repeaters* dépend directement de son envergure temporelle, l'étude des anciennes stations comme FDF a été privilégiée. Une famille de référence sur BIM a été ainsi *retrouvée* sur la station FDF, en découvrant au passage d'autres événements sismiques. Les profils de **glissement** cumulés ont pu être établis, et le **couplage** a pu être estimé. Cette démarche appelle donc à généraliser la méthode d'encadrement du glissement pour un grand nombre de familles, puis de raisonner à grande échelle sur les intervalles obtenus pour le couplage, afin les comparer aux couplages proposés par *van Rijsingen(2021)* [1] notamment.

Cependant, la compréhension du comportement sismo-tectonique de la subduction doit continuer à s'appréhender de façon multidisciplinaire : Les données locales obtenues à l'aide des *repeaters* complètent des modélisation globales, des études géomorphologiques régionales, ou des observations de terrain, et des approches nouvelles récentes viennent encore compléter les questionnements. L'étude de sédiments , ou de la croissance des coraux notamment, a émergé, pour lier l'afaissement annuel moyen de quelques *mm* de certains atolls à une accumulation de contraintes dans l'arc des Petites Antilles (*Feuillet* <https://doi.org/10.1016/j.epsl.2021.117343>). De plus, un article récent de *Haipeng Luo* (<https://doi.org/10.1038/s41561-025-01778-1>) place aussi l'étude des composantes verticales au centre de l'étude des accumulations de contraintes en zone de subduction, suggérant une réinterprétation du cas antillais, d'autant plus que la notion de couplage est de plus en plus considérée comme un paramètre instationnaire, pouvant varier sur plusieurs décennies.

Aussi l'étude des mégathrusts de 1839 et 1843 fait-elle appel à de nombreuses méthodes d'investigation, pour lesquelles l'étude des *repeaters* peut être une plus-value importante. C'est pourquoi l'originalité de cette méthode d'analyse locale et indirecte, permise par la disponibilité de réseaux de stations, est à continuer à exploiter, car nombreuses encore sont les zones à éclairer dans la compréhension de la subduction des Petites Antilles.

---

*Fin du rapport*

---

## Glossaire

***megathrust*** Séisme extrêmement puissant en magnitude, mégaséisme. Exemple : Tohoku (Japon), 2011. [4](#)

***repeater*** Séisme appartenant à une famille de séismes répétés, issus d'une même source. [3](#), [10](#), [12](#), [24](#)

***standalone*** code autonome pouvant s'exécuter indépendamment, sans dépendre d'un autre fichier ou programme. [14](#)

***webservice*** Ensemble de serveurs contenant des données sismiques continues sur plusieurs stations. Les requêtes FDSN formatées permettent de cibler les données précises souhaitées. [15](#)

***build\_families*** Fonction `requake` construisant des familles en se basant uniquement sur des catalogues de séismes répertoriés (sans exploration des données continues). La constitution des familles se fait – actuellement – par composantes connexes ou UPGMA. [21](#), [23](#), [25](#), [34](#), [40](#)

***float64*** Type de nombre à virgule flottante en Python, codé sur 64 bits. Son utilisation permet une meilleure précision et efficacité dans les longs calculs (`scan_template`), sans erreurs d'arrondi. [16](#)

***plot\_cumulative*** Fonction `requake` affichant une figure avec les graphes en paliers associés aux familles. Les changements de paliers sont associés à des séismes, et la hauteur du palier permet de reconstituer le glissement local de la plaque en déplacement. [29](#), [32](#), [35](#), [38](#), [45](#)

***plot\_timestraps*** Fonction `requake` affichant les familles par leur étendue temporelle. Les points sont les évènements de la famille. Cela permet de visualiser l'écartement entre les *repeaters*. [24](#), [26](#)

***remove\_response*** Fonction `Obspy` pour convertir les amplitudes sismiques numériques, en amplitudes physiques (typiquement en  $m.s^{-1}$ ). [32](#)

***scan\_catalog*** Fonction `requake` permettant d'explorer un catalogue d'évènements et de calculer les [\*NCC\*](#) entre les séismes selon un critère de proximité géographique. C'est un préalable de la fonction `build_families`. [21](#), [23](#)

***scan\_templates*** Fonction `requake` permettant la recherche par *template matching* en explorant les données sismiques continues d'une station. La plage de temps du calcul et la dimension des *chunks* et *overlaps* sont gérées dans le fichier `requake.conf`. [14](#), [16](#), [18](#), [20](#), [21](#), [23](#), [25](#), [27](#), [28](#), [37](#), [38](#)

***slice*** Fonction d'`Obspy` utilisée pour extraire une fenêtre temporelle d'un signal sismique. La fonction renvoie une copie extraite, sans modifier le signal original. [16](#)

***trim*** Fonction d'`Obspy` utilisée pour extraire une fenêtre temporelle d'un signal sismique. La fonction renvoie le signal modifié, sans effectuer de copie (modification de l'objet-source). [19](#)

**Anaconda** Distribution Python pour manipuler des environnements Python liés aux calculs scientifiques. [13](#)

**aspérité** Irrégularité des roches, localisée au niveau de la faille de subduction, où la roche résiste au glissement, ce qui induit une succession de ruptures à l'origine de *repeaters* [6](#), [12](#), [23](#), [24](#), [36](#), [39](#)

**couplage** Degré d'interaction entre les deux plaques. Si le couplage est de 100%, alors les deux plaques sont solidaires. Si il est de 0%, alors la subduction se fait sans frictions, par fluage. [4](#), [32](#), [34](#), [36](#), [39](#), [44](#)

**fonction de Green** En sismologie, une fonction de Green correspond à la réponse impulsionale d'un milieu géologique donné. L'impulsion étant le séisme en question. Cette réponse impulsionale dépend directement de la nature géologique et physique du milieu de propagation des ondes entre foyer et station. [28](#)

**glissement** déplacement spatial de la plaque plongeante le long de la faille de subduction. Ce glissement peut être continu, ou discontinu (cas des *repeaters*). [4](#), [6](#), [32](#), [36](#), [39](#)

**hypocentre** Synonyme de *foyer*. Emplacement-source où une rupture de roche libère de l'énergie créant un séisme. Sa projection à la surface est l'*épicentre*. [2](#), [11](#)

**lithosphère** Enveloppe rigide de la surface de la Terre comprenant la croûte terrestre et une partie du manteau supérieur. Elle est divisée en plaques tectoniques [[voir sur Wikipédia](#)]. [1](#)

**NumPy** Format de données sous forme de tableaux numériques ou *arrays*, pour optimiser les calculs scientifiques. [16](#)

**Nyquist** Fréquence limite supérieure que doit contenir un signal pour une fréquence d'échantillonnage donnée. Le théorème de Nyquist-Shannon affirme que  $f_{Nyquist} = f_s/2$  [15](#)

**ondes P** Ondes primaires d'un séisme [7](#), [11](#), [19](#), [27](#), [34](#)

**ondes S** Ondes secondaires d'un séisme [11](#), [19](#), [27](#), [31](#), [34](#)

## Acronymes

**FFT** *Fast Fourier Transform*. Algorithme de calcul de transformée de Fourier discrète, de complexité  $\mathcal{O}(n \log n)$ . [14](#)

**NCC** *Normalized Cross Correlation*, corrélation croisée normalisée. [6](#), [40](#)

**SNR** *Signal to Noise Ratio*, rapport signal à bruit [12](#)

**UTC** *Coordinated Universal Time*. Temps universel civil basé sur une généralisation du temps *GMT*, dit de *Greenwich*. [15](#), [17](#)

**FDSN** *International Federation of Digital Seismograph Networks*. Cet organisme gère l'accès aux données sismiques via des formats de requêtes standardisés aux serveurs. [28](#), [29](#)

**GNSS** Géolocalisation et Navigation par un Système de Satellites [1](#), [4](#), [6](#), [37](#), [45](#)

**HNM** High Noise Model (voir la description des spectres de bruit en [2.6](#)) [21](#), [22](#)

**IPGP** Institut de Physique du Globe de Paris [i](#), [ii](#), [4](#)

**LNM** Low Noise Model (voir la description des spectres de bruit en [2.6](#)) [21](#), [22](#)

**OVSM** Observatoire Volcanologique et Sismologique de la Martinique [25](#)

**PSD** *Power Spectral Density*, densité spectrale de puissance [21](#)

**RAM** *Random Access Memory*. Mémoire vive d'un ordinateur. [18](#)

**RéSiF** Réseau sismologique et géodésique Français. [22](#)

**UPGMA** *Unweighted pair group method with arithmetic mean* [9](#), [23](#), [34](#), [40](#)

**WI** *West Indies*. Réseau de stations sismiques opéré dans la zone des Caraïbes par l'IPGP. [21](#), [35](#)

**WPGMA** *Weighted pair group method with arithmetic mean* [9](#)

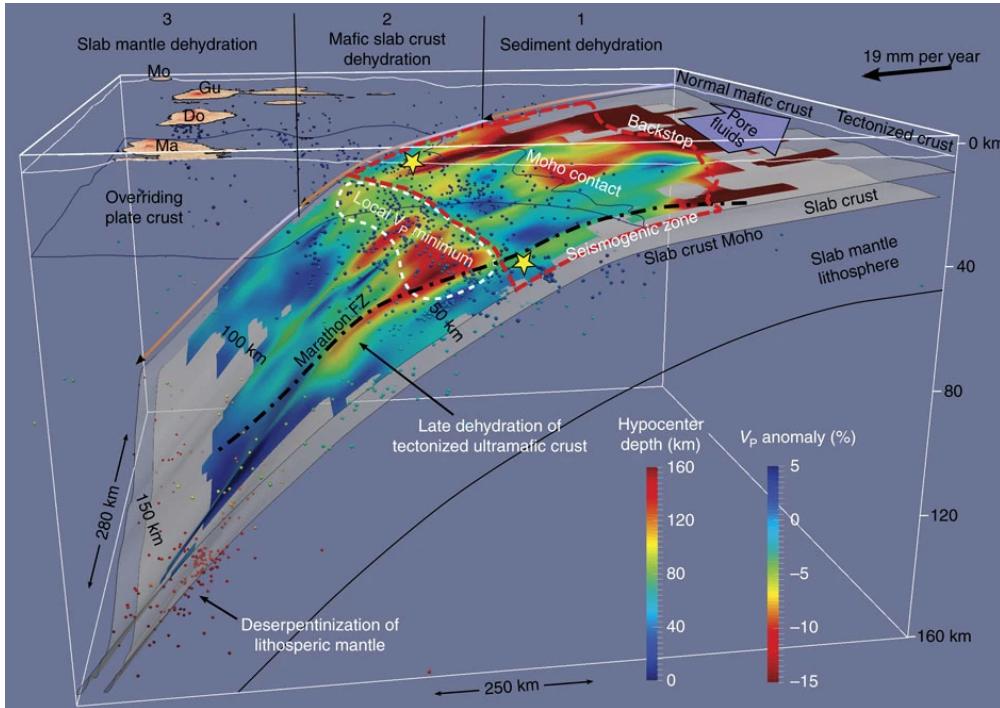


## Références

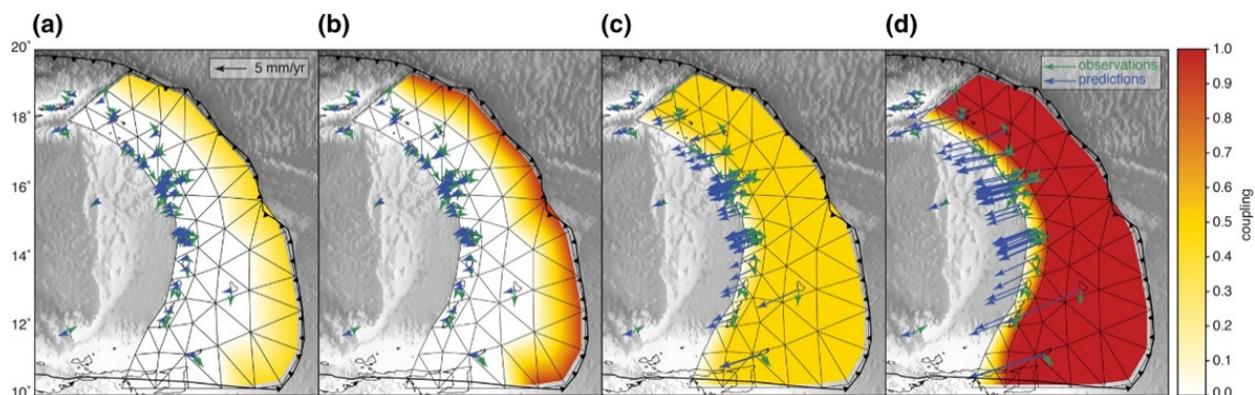
- [1] Eleonora van Rijsingen et al. Inferring Interseismic Coupling Along the Lesser Antilles Arc : A Bayesian Approach. *Journal of Geophysical Research : Solid Earth*, 2021. doi : 10.1029/2020JB020677.
- [2] Michele Paulatto. Dehydration of subducting slow-spread oceanic lithosphere in the Lesser Antilles. *Nature Communications*, 2016. doi : 10.1038/ncomms15980.
- [3] Nathalie Feuillet, François Beauducel, and Paul Tapponnier. Tectonic context of moderate to large historical earthquakes in the Lesser Antilles and mechanical coupling with volcanoes. *Journal of Geophysical Research*, 116, 2011.
- [4] Adèle Doucet. Recherche de séismes répétés dans la zone de subduction des Antilles. Rapport de m1 2020–2021, IPGP, Université de Paris Risques Naturels, Paris, France, mai 2021. Encadrants : C. Satriano et M.-P. Bouin.
- [5] Damien Guillou. PREST 2017-2023 ; Volcans, séismes et tsunamis dans la Caraïbe. Technical report, Interrreg Caraïbes, 2024.
- [6] Naoki Uchida. Detection of repeating earthquakes and their application in characterizing slow fault slip. *Progress in Earth and Planetary Science*, 2019. URL <https://doi.org/10.1186/s40645-019-0284-z>.
- [7] Eric Beaucé. Fast matched-filter (FMF) : an efficient seismic matched-filter search for both CPU and GPU architectures. *Seismological Research Letters*, 2017. URL <https://www.researchgate.net/publication/321487962>.
- [8] Johnson Lane R Nadeau Robert M. Moment release rates and estimates of source parameters for small repeating earthquakes. *B Seismol Soc Am*, 88 :790–814, 1998.
- [9] Claudio Satriano. SourceSpec – Earthquake source parameters from P- or S-wave displacement spectra, 2020. URL <https://github.com/SeismicSource/sourcespec>.
- [10] James N. Brune. Tectonic stress and the spectra of seismic shear waves from earthquakes. *Journal of Geophysical Research*, 75 :4997–5009, 1970. URL <https://doi.org/10.1038/s41561-025-01778-1>.
- [11] Eric Beaucé. Microseismic Constraints on the Mechanical State of the North Anatolian Fault Zone 13 Years After the 1999 M7.4 Izmit Earthquake. *Journal of Geophysical Research : Solid Earth*, 2022. doi : 10.1029/2022JB024416.
- [12] Claudio Satriano. Requake – Repeating earthquakes search and analysis, 2024. URL <https://github.com/SeismicSource/requake>.
- [13] Python framework for seismology, 2017. URL <https://github.com/obspy/obspy>.
- [14] Claudio Satriano. SeisCat – Keep a local seismic catalog, 2024. URL <https://github.com/SeismicSource/seiscat>.
- [15] Simone Cesca, Peter Niemz, and Torsten Dahm. Anti-repeating earthquakes and how to explain them. *Communications Earth & Environment*, 5 :158, 2024. doi : 10.1038/s43247-024-01290-1. URL <https://www.nature.com/articles/s43247-024-01290-1>.

# Annexes

Par souci de clarté/lisibilité, les annexes contenant du code, ou des très grandes images sont déposées sur le lien GitHub suivant : <https://github.com/PDSAntilles/Annexes-au-rapport-de-PDS-Petites-Antilles>. Les présentes annexes contiennent uniquement des figures complétant le présent rapport.

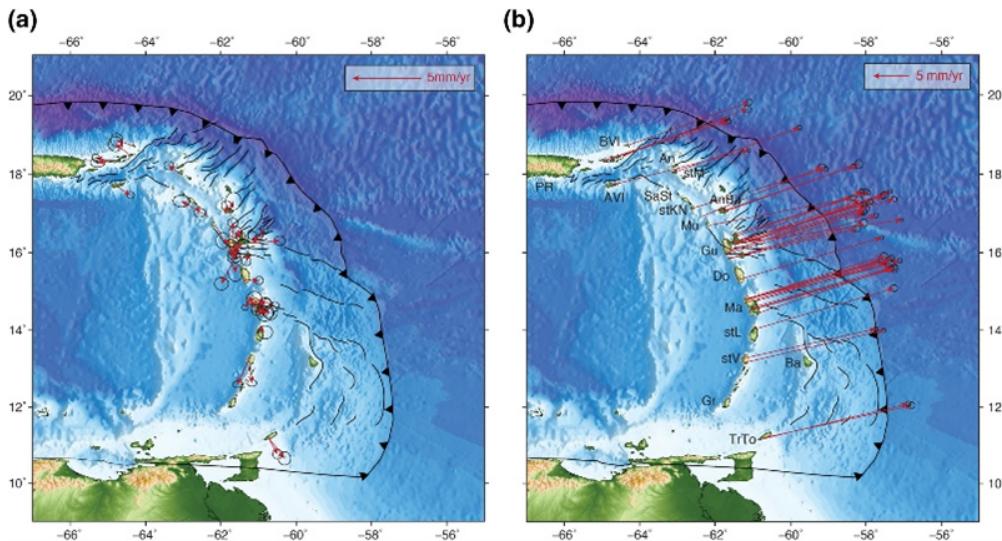


Vue en 3D de la subduction au niveau des Petites Antilles. Les îles volcaniques apparaissent en haut à gauche et la zone sismogène apparaît en pointillés rouges. Extrait de [2].

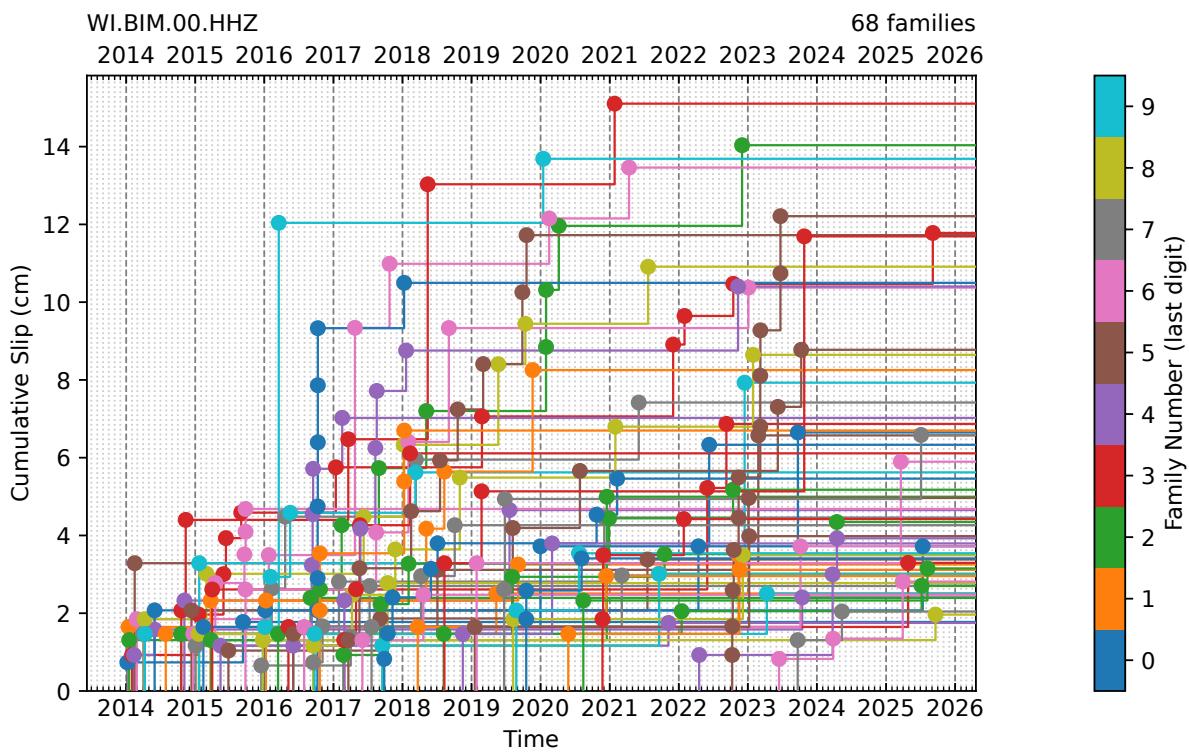


Comparaison entre les déplacements observés (flèches vertes) et supposés (flèches bleues) pour différents modèles de **couplage**. Les couleurs correspondent au modèle de couplage appliqué (voir échelle sur la droite). Les prédictions ont été calculées pour un bloquage à 20 km ou 65 km de profondeur. Il apparaît que les couplages élevés (homogènes ou hétérogènes) sont très improbables, alors que l'hypothèse d'un couplage faible n'est pas à exclure. Extrait de [1].





Observation des déplacements relatifs par **GNSS**, dans le référentiel de la plaque Caraïbe (à gauche), et de la plaque Américaine (à droite). Le déplacement relatif de  $20 \text{ mm/an}$  au niveau de la zone de subduction est mis en valeur par la sous-figure de droite. Extrait de [1].



Courbes en paliers ([plot\\_cumulative](#)) pour les 68 familles de BIM retenues après sélection, couvrant la plage de temps [2014 - présent (2025)]. On constate la variété des profils de glissement parmi les familles.

