

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
Ministère de l'enseignement supérieur et de la recherche scientifique
Université Abou Bekr BELKAID – Tlemcen



Faculté de Technologie
Département de Génie Civil

THESE DE DOCTORAT
Présentée par

Ahmed BOUDGHENE STAMBOULI

Sujet de la thèse

**Estimation de l'effet de site par la
méthode des réseaux de neurones
artificiels**

Soutenue le : 08/05/2018 devant le jury composé de

Mohammed MATALLAH	Président	UAB. Tlemcen
Djawad ZENDAGUI	Directeur	UAB. Tlemcen
Boumediène DERRAS	Co-Directeur	U. Saida
Mohammed FARSI	Examineur	CGS Alger
Mohammed Amine CHIKH	Examineur	UAB. Tlemcen
Pierre_yves BARD	Invité	UJF .Grenoble

Remerciements

Je remercie DIEU tout puissant qui m'a donné vraiment la patience et la volonté de mener et terminer cette thèse.

Le présent travail de thèse de doctorat s'est déroulé au sein du laboratoire d'évaluation et management du risque (Risam)-Département de génie civil, Université Abou Bakr Belkaid Tlemcen en collaboration avec le laboratoire Isterre de l'université Joseph Fourier. Il a été dirigé du côté algérien par le Prof Djawad. Zendagui et le Prof Boumédiène Derras et du côté français par le Prof Pierre-Yves. Bard. Je tiens à les remercier vivement pour m'avoir dirigé dans ce travail et de m'apporter par leurs compétences un support technique et de précieux conseils et encouragements et un soutien lors des différentes étapes de la recherche.

Mes remerciements vont à l'honneur des membres du jury le professeur Mohammed FARSI et le professeur Mohammed Amine CHIKH, d'avoir accepté d'examiner ce travail.

Ainsi je remercie vivement le professeur Mohammed MATALLAH pour avoir bien voulu présider mon jury.

J'aimerais aussi remercier tous ceux qui ont contribué à l'élaboration de ce travail et particulièrement mes chers amis pour l'aide qui m'ont apporté pour réaliser ce travail.

Enfin je remercie du fond du cœur ma famille en particulier ma chère mère, mon père pour son soutien, sa patience, son aide précieuse et surtout son amour qui m'a aidé à surmonter les moments les plus difficiles.

Table des Matières

ملخص.....	xv
Résumé.....	xvii
Abstract.....	xix
Notations.....	xxi
Liste des Tableaux.....	xxiii
Liste des Figures.....	xxv
Chapitre 1.....	1
Introduction.....	1
1.1 Introduction.....	1
1.2 Plan de la thèse.....	3
Chapitre 2.....	5
Revue bibliographique sur les effets de site et identification de la problématique.....	5
2.1 Introduction.....	5
2.2 Cas de séismes caractérisés par les effets de site.....	6
2.2.1 Mexique (Séisme de Michoacan, 1985).....	6
2.2.2 USA (Séisme de Loma Prietta, 1989).....	8
2.2.3 Japon (Séisme de Kobe, 1995).....	9
2.2.4 Algérie (Séisme de Boumerdes, 2003).....	11
2.3 Effets de site : Définition et mesures.....	11
2.3.1 Définition de l'effet de site.....	11
2.3.2 Méthodes de mesure de l'effet de site.....	13
2.3.3 Analyse critique et identification de la problématique.....	22
2.4 Conclusion.....	22
Chapitre 3.....	25
Mise en place de la méthode.....	25
3.1 Introduction.....	25
3.2 Méthode de calcul du facteur d'Amplification (AF).....	25
3.2.1 Définition et hypothèses de base.....	25
3.2.2 Signal sismique.....	26
3.2.3 Réponse sismique d'un profil de sol stratifié horizontalement.....	30
3.2.4 Caractérisation des profils de sol.....	38
3.3 Etablissement de la base de données.....	44
3.4 Méthode d'estimation de l'AF par les réseaux de neurones.....	46
3.5 Conclusion.....	47
Chapitre 4.....	49
Description de l'approche neuronale.....	49
4.1 Introduction.....	49

4.2	Inspiration du neurone formel	49
4.3	Concept de base des réseaux de neurones artificielles	51
4.4	Types de réseaux neurones artificiel (RNA)	52
4.4.1	Réseaux de neurones statiques (réseaux non bouclés)	52
4.4.2	Réseaux de neurones dynamiques ou réseaux bouclés (ou récurrents)	54
4.5	Types d'apprentissage	55
4.5.1	Rétro- propagation du gradient (RPG)	56
4.5.2	Fonction de transfert (ou d'activation)	56
4.5.3	Avantages et inconvénients des différents RNAs statiques et supervisés utilisés pour l'approximation des fonctions	57
4.6	Aspect théorique du GRNN : cas particulier d'une seule sortie	59
4.6.1	Mise en œuvre de GRNN sous Matlab :Aspect pratique	60
4.6.2	Etape d'élaboration d'un modèle avec un réseau GRNN	63
4.6.3	Mesure des performances	64
4.6.4	Détermination et optimisation de la largeur de la gaussienne	65
4.7	Conclusion.....	69
Chapitre 5	71
Résultats et interprétations	71
5.1	Introduction	71
5.2	Description des jeux de données.....	71
5.2.1	Facteur d'amplification en fonction de la période.....	72
5.2.2	Facteur d'amplification en fonction de la fréquence normalisée	75
5.2.3	Facteur d'amplification à courte et moyenne période.....	77
5.3	Description des différents modèles de réseaux de neurones	78
5.4	Résultats.....	79
5.4.1	Nomenclature des résultats.....	79
5.4.2	Dérivation et analyse du terme d'erreur	84
5.4.3	Approche globale	92
5.4.4	Variation des facteurs de Borchardt en utilisant GRNN	98
5.5	Conclusion.....	108
Conclusion	111
Bibliographie	119

ملخص

معظم كتب هندسة الزلزالية تأخذ بعين الاعتبار تأثير الموقع عن طريق عامل التضخيم (AF) عن طريق تعديل أطراف استجابة التسارع من الصخور. ويتم هذا التغيير من خلال معلومات الموقع ذات الصلة بالخصائص الهندسية والميكانيكية لملاح التربة. ولذلك فمن المهم تحديد معالم الموقع التي تتحكم بشكل أفضل في عامل التضخيم التلقائي. والغرض من هذا البحث هو تقديم تقييم كمي لأداء المعلومات الموقع المختلفة (وتسمى أيضا وكيل) للتنبؤ الخصائص الرئيسية لتضخيم (AF).

ولهذه الغاية، وضعنا منهجية تستخدم طريقة الشبكة العصبية الاصطناعية. ومدخلات الشبكة هي البروكسيات التربة التي ينبغي استخدامها، و عامل التضخيم (AF) التي تحددتها نظرية انتشار الموجات الزلزالية، والنتائج هو (AF). من أجل الحصول على عينة كبيرة من البيانات، تم استخدام 858 أعمدة التربة الفعلية مع 14 التسارع الفعلي مع محتوى متغير التردد. يتم تقييم أداء كل تركيبة بروكسي موقع عن طريق تقليل التباين من التباين الأولي لعوامل التضخيم لمحات 858، مع الأخذ في الاعتبار كل من المدى الزمني الكامل والفترة القصيرة والمتوسطة المدى المرتبطة بعوامل بورشردت (borchedt) (Fa, Fv).

النتائج التي تم الحصول عليها تبين أن تضخيم يمكن تقريبه مرضيا مع عدد محدود من بروكسيات الموقع (1-6). بما أن الممارسة البرمجية المعتادة تتضمن عددا محدودا من بروكسيات الموقع (عادة 1، أحيانا 2)، يتم إجراء تحليل لتحديد معالم الموقع "الأفضل أداء". الأفضل هو التباين السرعة الشاملة بين الأساس الأساسي والسرعة الدنيا في عمود التربة. كما أنها واحدة من المعايير الأكثر صعوبة ومكلفة لقياس - وخاصة في حالة الرواسب سميكة - ولذلك يفضل غيرها من المعلومات أكثر ملاءمة، على وجه الخصوص (f_0 , V_{s30}) التي لوحظ أفضل مزيج. من منظور الرمز، يتم تقديم المعادلات الاعتماد على هذه عوامل لتنبأ التضخيم فترة قصيرة ومتوسطة Fa و Fv على هذين المعيارين.

الكلمات المفتاحية: انتشار موجة 1d، معلومات عن الموقع، عامل التضخيم، الشبكة العصبية.

Résumé

La plupart des codes sismiques prennent en compte l'effet du site par l'intermédiaire d'un facteur d'amplification (AF) en modifiant les spectres de réponse d'accélération aux rocher. Cette modification est opérée par le biais de paramètres de site qui sont liés aux propriétés géométriques et mécaniques du profil de sol. Il est donc intéressant d'identifier les paramètres du site qui contrôlent le mieux le facteur d'amplification AF. Le but du présent travail de recherche est de fournir une évaluation quantitative de la performance de différents paramètres de site (appelé aussi proxy) pour prédire les principales caractéristiques de l'amplification 'AF' et par conséquent celles des facteurs de Borchardt F_a et F_v .

A cet effet, nous avons établi une méthodologie basée sur les réseaux de neurones généralisés. Les inputs du réseau sont les proxies à utiliser et les facteurs d'amplifications (ou les F_a et F_v) déterminés par la théorie de propagation d'ondes. Afin d'avoir un large échantillon de données, on a utilisé 858 colonnes de sol réel soumis à 14 accélérogrammes réels ayant un contenu fréquentiel variable. La performance de chaque combinaison de proxy de site est évaluée par la réduction de la variance par rapport à la variabilité initiale des facteurs d'amplification pour les profils choisis, en considérant à la fois la plage de période entière et à des périodes spécifique (à courte et à moyenne) période associées aux facteurs de Borchardt F_a et F_v .

Les résultats obtenus montrent que l'amplification d'un profil de sol peut être approximée de façon satisfaisante avec un nombre limité de proxies de site (1 à 6). Comme la pratique de code habituelle implique un nombre limité de proxies de site (généralement 1, parfois 2), une analyse de sensibilisée est effectuée pour identifier les paramètres du site «les plus performants». Le meilleur est le contraste de vitesse global entre le substratum sous-jacent et la vitesse minimale dans la colonne du sol. Comme il s'agit d'un des paramètres les plus difficiles et les plus onéreux à mesurer- en particulier dans le cas des dépôts épais, on préfère d'autres paramètres plus commodes, en particulier le couple (V_{s30} , f_0). Dans une perspective de code, des équations et des tracés sont fournis décrivant la dépendance des facteurs d'amplification à courte et moyenne période (" F_a " et " F_v ") en fonction de ces deux paramètres.

Mots clés : Réponse 1D linéaire, Paramètre de site, Facteur d'amplification, Réseau de neurone.

Abstract

Most seismic codes take into account the effect of the site via an amplification factor (AF) by modifying the acceleration response spectra of the rock. This change is made through site parameters that are related to the geometric and mechanical properties of the soil profile. It is therefore interesting to identify the site parameters that best control the AF amplification factor. The purpose of this research paper is to provide a quantitative assessment of the performance of various site parameters (also called proxy) to predict the main characteristics of AF and therefore those of Borchardt Fa and Fv factors.

To this end, we have established a methodology that uses the generalized neural network method. The inputs of the network are the proxy(s) to be used and the AF (or Fa and Fv) determined by the seismic wave propagation theory and the output is the AF (or Fa and Fv). In order to have a large sample of data, 858 actual soil columns were used with 14 actual accelerograms with variable frequency content. The performance of each site proxy combination is evaluated by reducing the variance from the initial variability of the amplification factors for the 858 profiles, considering both the entire period range and the short and medium ranges period associated with Borchardt factors Fa and Fv.

Results obtained show that the amplification of a soil profile can be satisfactorily approximated with a limited number of site proxies (1-6). As usual code practice involves a limited number of site proxies (usually 1, sometimes 2), a sensitivity analysis is performed to identify the "best performing" site parameters. The best is the overall velocity contrast between the underlying bedrock and the minimum velocity in the soil column. As it is one of the most difficult and expensive parameters to measure - especially in the case of thick deposits - other more convenient parameters are preferred, in particular (V_{s30} , f_0) which has been observed the best combination. From a code perspective, equations and plots are provided describing the dependence of short and medium period on amplification factors "Fa" and "Fv" with these two site proxies.

Keywords: 1D linear site response – site proxies – amplification factors – neural network.

Principales notations utilisées

Les principales notations utilisées sont présentées ci-dessous. Les autres symboles précisant des détails ou introduits pour les besoins d'une démonstration sont présentés dans le cours du texte.

AF	Amplification Factor
Ω	Quantité sismique
PSA(T)	Pseudo spectre d'accélération
f	Fréquence
T(f)	gain de filtrage de BETTERWORTH
Mw	Magnitude du moment
PGA	peak ground aceleration
w	fréquence d'onde
u	le déplacement
Z	la profondeur
t	le temps
K*	nombre d'onde complexe
Vs	la vitesse de cisaillement de la couche considéré.
ξ :	Facteur d'amortissement
Q	facteur de qualité qui décrit l'énergie de dissipation
A _m , B _m	les amplitudes de la propagation d'onde
α	le rapport d'impédance
FFT	la transformée de Fourier rapide
E ₁ (ω)	Le spectre de Fourier à la surface libre
E _{n+1} (ω)	le spectre de Fourier au niveau du substratum rocheux
T _{1 n+1} (ω)	la fonction de transfert entre l'entrée et la surface.
(Depth)	Profondeur
V _{sm}	la vitesse de cisaillement moyenne sur le profils
(V _{s30})	la vitesse de cisaillement moyenne des ondes sur 30m
(V _{bedrock})	la vitesse de cisaillement du substratum
(Cv)	le contraste de vitesse
f ₀	la fréquence fondamentale du profil de sol
$\sigma_{AF}(P_k, \theta, T_i)$	Ecart type
$\sigma_{0Max}(\theta)$	Variabilité initiale
n _T	le nombre de périodes utilisé égal à 271
yi' :	la valeur prédite en fonction du vecteur d'entrée P comme inconnue
Q	le nombre de vecteur soumis à l'apprentissage
R	le nombre d'éléments du vecteur d'entrée (les paramètres de sites).
IW	les poids entre la première couche et la couche cachée
(P)	les entrées

b	le biais
spread	la largeur de la gaussienne
SSE	la somme carrée des erreurs
y'_{ij}	La sortie prédit (ie : facteur amplification à i periode pour j profiles).
\hat{y}_{ij}	la sortie du calcul exact (ie : facteur amplification à i periode pour j profiles).
(r_{ij})	Résidus
LW	les poids entre la couche cachée et la couche de sommation.
Q	le nombre de vecteur soumis à l'apprentissage.
npf	Nombre d'amplification prise à chaque periode ou frequence normalisée.
R	le nombre d'éléments du vecteur d'entrée(les paramètres de sites).
R_c	coefficient de corrélation.
$\bar{\hat{y}}$	La moyenne de la sortie du calcul exact (modèle linéaire viscoélastique) sur tous les profils de sol pris.
\bar{y}'	La moyenne de la sortie prédit
\hat{y}_j	la moyenne de la sortie du calcul exact (modèle linéaire viscoélastique) pour J^{ieme} profil de sol.
y'_j	La moyenne de la sortie prédit pour j^{ieme} profil de sol.
$\epsilon_{GRNN}(\theta, T_j)$	une erreur dépendante de la période entre prédit et analytique
Fa	facteur de borcherdt à courte période
Fv	facteur de borcherdt à moyenne période
GRNN	Generalized regression neural network
RVFLNN	Réseau fonctionnel de réseau fonctionnel de vecteur aléatoire (Random Vector Functional Link Neural Network)
PMC	Perceptron multicouche
RBF	Radial basis function
PNN	probabiliste neural networks
RP	Profils réel
NP	Profils normalisé
TP	Profils tronqué
dist	distance
radbas	Fonction gaussienne

Liste des Tableaux

Tableau 1 Valeurs de l'accélération maximale des trois composantes enregistrées sur les sites de Treasure Island et Yerba Buena Island.....	8
Tableau 2 Valeur de la vitesse maximale en (m/s) lors du séisme de Kobe.....	10
Tableau 3 Catégories de techniques d'évaluation expérimentales de l'effet de site (Bard et Riepl-Thomas, 2000).....	14
Tableau 4 Informations sur les conditions de site et le nombre de spectre de réponse utilisé pour élaborer le système neuronal.....	19
Tableau 5 Classification de site utilisée dans le choix du RNA.....	20
Tableau 6 Résumé des séismes choisis.....	29
Tableau 7 exemple du profil de sol pris dans le calcul.....	36
Tableau 8 Facteur de corrélation(R^2) entre les paramètres de site.....	41
Tableau 9 facteur de corrélation entre les paramètres normalisés en substratum avec 800m/s.....	42
Tableau 10 facteur de corrélation entre les paramètres des profils tronqués (TP).....	43
Tableau 11 Type de fonction de transfert les plus utilisées.....	57
Tableau 12 Avantage et inconvénient de quelque réseaux de neurones statiques.....	58
Tableau 13 Calcul des variabilités.....	73
Tableau 14 : Calcul des variabilités pour des profils normalisés en fréquence.....	77
Tableau 15 Calcul des variabilités pour F_a et F_v cas des profils (RP).....	77
Tableau 16 Listes des modèles GRNN considérés.....	80
Tableau 17 Paramètres de site pour Sol 1 et Sol 2.....	81
Tableau 18 Écart type des résidus de modèle pour différents modèles GRNN impliquant les facteurs d'amplification pour (RP, NP et TP), application à différents combinaisons de paramètres de site.....	93
Tableau 19 Variation de l'écart type avec le nombre de paramètres du site pour le cas des profils (RP) (profils réels).....	106
Tableau 20 Variation de l'écart type avec le nombre de paramètres du site pour le cas (NP), (profils normalisé).....	107
Tableau 21 Variation de l'écart type avec le nombre de paramètres du site pour le cas (TP) (profils tronqué).....	107

Liste des Figures

Figure 2-1 Enregistrements du tremblement de terre de Michoacan (1985) de la côte Pacifique jusqu'à Mexico (Çelebi, et al, 1988).	6
Figure 2-2 Amplitudes maximales selon les zones de Michoacan-Mexico à gauche et à droite le pseudo spectre d'accélération en fonction de la période (Stewart, 1997).	7
Figure 2-3 Réponses spectrales d'accélération enregistrées au niveau de Treasure Island et YerbaBuena Island (Stewart, 1997).	9
Figure 2-4 Carte du bassin d'Osaka montrant la ceinture des dommages défini comme un secteur où l'intensité en JMA a dépassée VII. (Pitarka et al, 1998).	10
Figure 2-5 Comparaison d'enregistrements sismiques sur la roche et aux sédiments du séisme de Lancy de 1995 (Lebrun, 1997).	12
Figure 2-6 L'effet de site (piégeage des ondes).	13
Figure 2-7 les différentes méthodes adoptées pour déterminer les effets de site. (Cadet et al 2007).où (f_0) représente la fréquence de résonance des sites, (A_0) l'amplification correspondante provoquée par le site à cette fréquence et (FDT) la fonction de transfert.	15
Figure 2-8 Enregistrements en surface et en profondeur (-103m) sur une même station (KIK-NET) (Derras et al, 2004).	16
Figure 2-9 HVSR de l'événement du 16/09/2002 enregistré sur la station TTRH03 (KIK-NET).	17
Figure 2-10 Architecture du réseau de neurone.....	20
Figure 2-11 Spectres de réponse d'accélération estimés par RNA et calculés à partir des enregistrements réels.....	21
Figure 3-1 Exemple de pseudo spectre d'accélération avec un amortissement de 5%.	27
Figure 3-2 Gains de filtres de Butterworth passe-bas d'ordre 4 en fonction de la fréquence.....	28
Figure 3-3. Pseudo spectre d'accélération normalisée (avec 5% d'amortissement) pour un ensemble de 14 accélérogrammes réels récolté en substratum (S_m : représente la moyenne géométrique des 14 pseudos spectre d'accélération normalisée).	30
Figure 3-4 Système de couches de sol stratifié horizontalement soumis à la propagation d'onde S verticale.....	32
Figure 3-5 Terminologie utilisée dans l'analyse de la réponse de site	35
Figure 3-6 Facteur d'amplification fourrier calculer parle modèle linéaire élastiques, en comparaison notre programme avec Deepsoil.	36
Figure 3-7 Accélération en surface par le modèle linéaire élastiques, en comparaison notre programme avec Deepsoil.	37
Figure 3-8. Facteur d'amplification fourrier calculer parle modèle linéaire viscoélastique ou on compare notre programme avec Deepsoil.....	37

Figure 3-9 Accélération en surface par le modèle linéaire viscoélastique, en comparaison notre programme avec Deepsoil.	38
Figure 3-10 : Distribution en histogramme des paramètres de la base de donnée à savoir (Depth , f_0 , C_v , V_{sm} , V_{s30} et $V_{bedrock}$) pour tous les profils (cas profils RP), la courbe rouge représente la distribution normale théorique.	40
Figure 3-11 : Distribution en histogramme des paramètres de la base de donnée à savoir (Depth , f_0 , C_v , V_{sm} et V_{s30}) pour des profils normalisés en substratum (NP), la courbe rouge représente la distribution normale théorique.	42
Figure 3-12 : Distribution en histogramme des paramètres de la base de donnée à savoir (Depth , f_0 , C_v , V_{sm} et V_{s30}) pour des profils tronqués (TP), la courbe rouge représente la distribution normale théorique.	43
Figure 3-13 : Définition des paramètres pris par les réseaux neurones en utilisant 05 paramètres d'entrés pour les profils normalisés (NP) et tronqués(TP).	47
Figure 4-1: Un neurone biologique et ses principaux composants (Alani, 2008).	50
Figure 4-2: Structure générale du neurone formel (Derras et al, 2011).	51
Figure 4-3: Architecture des réseaux neurones.	52
Figure 4-4 Architecture du GRNN avec une seule sortie.	60
Figure 4-5 :Approximation d'une fonction quelconque Y (en noir) par des fonctions gaussiennes(données en différentes couleurs (Alani, 2008).	61
Figure 4-6 : Structure du GRNN (Matlab).	61
Figure 4-7 Variation de la performance en fonction du la largeur de la gaussienne(spread), ((a)la performance est mesuré par Sse), (b par le résidu max)).	67
Figure 4-8 Biais et variance contribuant à l'erreur totale.	68
Figure 5-1 : (a) Variation de l'amplification moyenne $A_{0\theta, Ti}$, (b) Variation de la variabilité $\sigma_{0\theta, Ti}$	72
Figure 5-2 : Facteur d'amplification pour tous les profils en fonction de la période $AFmPk, \theta, Ti$, ainsi la variation de l'amplification moyenne de $A_{0\theta, Ti}$ (courbe rouge),et éventuellement la variation de $AFmPk, \theta, Ti \pm \sigma_{0\theta, Ti}$ application pour les différents type de profils (a : RP , b : NP et c :TP).	74
Figure 5-3 : (a) Variation du nombre de profils à chaque fréquence normalisés (f/f_0), (b) Exemple de facteur d'amplification réelle et interpolé pour le même profil.	75
Figure 5-4: Facteur d'amplification pour tous les profils en fonction de la fréquence normalisées $AFmPk, \theta, fi/f_0$, ainsi la variation de l'amplification moyenne de $A_{0\theta, fi/f_0}$ (courbe rouge) et éventuellement la variation de $AFmPk, \theta, fi/f_0 \pm \sigma_{0\theta, fi/f_0}$ application pour les différents type de profils ((a) RP, (b) NP et (c) TP).	76

Figure 5-5: Comparaison entre $AFmSol1,1, Ti$ et les prédictions du GRNN pour le Sol1.....	82
Figure 5-6: Comparaison entre $AFmSol2,1, Ti$ et les prédictions du GRNN pour le Sol2.....	82
Figure 5-7 : Variation des résultats calculés par la méthode classique (propagation d'onde) (Target Outputs) en fonction des résultats prédit par la méthode GRNN à (0.01s, 0.1s, 0.5s, 1s, 2s et 4s) pour les 858 profils application pour (RP).....	83
Figure 5-8 : Variation de $\varepsilon_{GRNN\theta} = 1, Ti$ pour différentes combinaisons de paramètres de site et de $\sigma_{0\theta} = 1, Ti$	85
Figure 5-9 : Variation de $\varepsilon_{GRNN\theta} = 0, Ti$ pour différentes combinaisons de paramètres de site et de $\sigma_{0\theta} = 0, Ti$	86
Figure 5-10 : Variation de $\varepsilon_{GRNN\theta} = 2, Ti$ pour différentes combinaisons de paramètres de site et de $\sigma_{0\theta} = 2, Ti$	86
Figure 5-11 : Variation de $\varepsilon_{GRNN\theta} = 1, fif0$ pour différentes combinaisons de paramètres de site et de $\sigma_{0\theta} = 1, fif0$	87
Figure 5-12 : Variation de $\varepsilon_{GRNN\theta} = 0, fif0$ pour différentes combinaisons de paramètres de site et de $\sigma_{0\theta} = 0, fif0$	87
Figure 5-13 : Variation de $\varepsilon_{GRNN\theta} = 2, fif0$ pour différentes combinaisons de paramètres de site et de $\sigma_{0\theta} = 2, fif0$	88
Figure 5-14: Illustration de la réduction progressive de l'écart type des résidus entre les modèles GRNN et le AF réelle avec le nombre de proxies de sites considérées pour les profils (RP).....	90
Figure 5-15: Valeurs résiduelles (différence entre le log10 de l'accélération spectrale observée et celle prédite) à 0.01,0.1s,0.5s,1s,2s et 4s en ordonné Les résidus sont en fonction de la profondeur total (Depth), f_0, V_{sm} , contraste de vitesse (Cv) et V_{s30} pour tous les profils.....	91
Figure 5-16: Indicateurs de performance pour les différents proxies de site en terme (reduction of standard deviation) pour le cas des profils réels (RP) en période.....	95
Figure 5-17: Indicateurs de performance pour les différents proxies de site en terme (reduction of standard deviation) pour le cas des profils de sol normalisé(NP) en période.....	95
Figure 5-18: Indicateurs de performance pour les différents proxies de site en terme (reduction of standard deviation) pour le cas des profils de sol tronqué (TP) en période.....	96
Figure 5-19: Indicateurs de performance pour les différents proxies de site en terme (reduction of standard deviation) pour le cas des profils (RP, NP et TP) en fréquence normalisée.....	97
Figure 5-20: Variation des facteurs de Borchardt à courte periode (Fa) avec (f_0, V_{s30}) pour les profils de sol ((a) RP, (b) NP et (c) TP).....	99
Figure 5-21: Variation des facteurs de Borchardt à longue période (Fv) avec (f_0, V_{s30}) pour les profils de sol ((a) RP , (b) NP et (c) TP).....	100
Figure 5-22: Distribution des données initiales dans le plan (f_0, V_{s30}) pour les (a) RP, (b) NP et (c) TP.....	103

Figure 5-23: Indicateurs de performance en termes de réduction de l'écart-type des résidus pour les différents proxies de site (différentes courbes) pour le cas des profils RP, NP et TP ((a1, b1 et c1) : Fa et (a2, b2 et c2) : Fv)..... 105

Chapitre 1

Introduction

1.1 Introduction

Il admet actuellement que les effets de site ont un impact important sur le mouvement sismique du sol et pourraient donc causer des dommages accrus sur les structures. Par exemple, au cours du tremblement de terre du Grand Michoacan au Mexique (Bard et al., 1988), l'amplification induite par les effets du site a été reconnue comme la cause principale de l'effondrement des structures.

Cette amplification est habituellement décrite comme le rapport entre les paramètres sismiques en surface et en substratum rocheux pour un site particulier. Parmi ces paramètres, les spectres de réponse sont les plus utilisés en ingénierie. En effet, la plupart des codes de construction (parasismique) utilisent des spectres de réponse pour définir les charges sismiques de conception sur les structures construites. De plus, les études d'évaluation des risques utilisent des spectres de réponse en accélération pour définir le mouvement sismique et utilisent les équations de prédiction de mouvement (GMPEs) corrélant les ordonnées spectrales à l'amplitude, la distance et les paramètres du site. Pour des raisons de simplicité, dans ces GMPEs, les paramètres de site sont décrits par des "proxies" très brutes. Le plus courant est le paramètre " V_{s30} ", correspondant à la moyenne harmonique de la vitesse de l'onde S sur les 30 premiers mètres, introduit pour la première fois par Borchardt (1994), puis largement utilisé (voir par exemple Martin et Dobry, 1994 ; Dickenson et Seed, 1995 ; Dobry et al., 2000 ; Rodriguez-Marek et

Bray 2001 ; Pitalakis et al., 2001 et tous les GMPEs récents, NGA: Abrahamson et al., 2008;NGAWest2 :(Gregor et al., 2014; Ancheta et al., 2014) et les ensembles RESORCE: Douglas et al., 2014). D'autres auteurs remplacent ou complètent ce paramètre par d'autres paramètres du site tels que la fréquence fondamentale f_0 (Cadet et al., 2012 ; Castellaro et al., 2008 ; Luzi et al.,2001 ; Pitalakis et al.,2012, 2013)ou la profondeur à un niveau de roche dure défini par un seuil de vitesse (de 1 km/s, réf, jusqu'à 2,5 km/s, réf). Les termes associés à ces proxies de site permettent de quantifier le "facteur d'amplification" dépendant de la fréquence de la "roche standard" (habituellement caractérisé par $V_{s30} = 760$ à 800 m/s). Les mêmes proxies sont également utilisés dans les codes réglementaires pour affiner les caractéristiques des spectres de conception (c.-à-d., PGA, facteur à moyenne et courte période) du aux conditions de site: c'est le cas par exemple des codes internationaux du bâtiment, IBC 2012, UBC97, EC8.

Cependant, ces proxies de site capturent distinctement toute la physique de l'amplification du site. Des sites distincts avec des valeurs similaires de proxies de site (par exemple, V_{s30}) pourraient avoir des caractéristiques d'amplification beaucoup différentes. Cela a au moins deux conséquences :

- A) Il influe de façon significative sur la variabilité aléatoire des GMPEs en augmentant la durée de l'événement, ce qui augmente les risques, en particulier à des périodes de retour longues.
- B) Les termes de site correspondants peuvent présenter des variations significatives d'un GMPE à un autre, en fonction des données de mouvement utilisées pour leur dérivation. En effet la relation entre V_{s30} et la structure de vitesse d'un profil à autre n'est pas la même dans le bassin de Los Angeles, Les plaines côtières japonaises ou les bassins intra- bassins intramontagnards des Alpes ou des Appennins.

La recherche d'une meilleure caractérisation de l'effet de site, à travers les facteurs d'amplification, passe par l'identification des meilleurs paramètres de site à utiliser. De cet objectif découle l'ensemble des travaux de la présente thèse. Il s'agit

en autres d'explorer une nouvelle piste de recherche qui va nous permettre de relier le facteur d'amplification à des paramètres de site.

1.2 Plan de la thèse

La présente thèse est organisée en cinq chapitres comme suit :

- Le premier chapitre illustre un aperçu général sur l'organisation de la thèse.
- Le deuxième chapitre dresse une identification des travaux réalisés avec la description de la problématique.
- Le troisième chapitre est consacré à la mise en place de la méthodologie ainsi que la base de données.
- Le quatrième chapitre présente l'outil de la méthode des réseaux de neurones ainsi que son implémentation dans le cas du présent travail.
- Le dernier chapitre donne les résultats obtenus.

En fin, une conclusion clôture le travail avec des perspectives pour de futurs travaux.

Chapitre 2

Revue bibliographique sur les effets de site et identification de la problématique

2.1 Introduction

La compréhension des effets de site revêt d'une importance toute particulière dans la mesure où plusieurs grandes villes à travers le monde sont construites sur des remplissages de vallées alluviales ou sur des bassins sédimentaires. Trifunac a mentionné la phrase suivante "Investigations of the effects of local site conditions on the characteristics of strong earthquake ground motion are probably as old or even older than the discipline of Earthquake Engineering itself" (Trifunac, 1990).

Lors du séisme du Frioul (Italie) en septembre 1976, la ville de Grenoble, située à 500 km de l'épicentre, s'est comportée comme une véritable "caisse de résonance" amplifiant fortement les mouvements sismiques. On sait aujourd'hui que la ville n'est pas à l'abri d'un séisme de magnitude 5,5. Compte tenu de l'étendue actuelle de l'agglomération grenobloise (l'aire urbaine compte aujourd'hui plus de 664 832 habitants, la 10e de France) et de la présence d'installations industrielles, ces phénomènes d'amplification des mouvements sismiques pourraient aujourd'hui avoir un impact désastreux. Cet exemple n'est qu'un exemple parmi tant d'autres.

Le présent chapitre est une étude bibliographique sur les approches d'estimation des effets de site. On commencera par donner un bref aperçu sur quelques cas de séisme mondialement connus, ayant été caractérisé par les effets de site. Les causes et les origines de ces effets sont présentées. Une étude critique sur les méthodes d'évaluation de ces effets est donnée. De cette analyse critique en découlera l'identification de la problématique du présent travail de recherche.

2.2 Cas de séismes caractérisés par les effets de site

2.2.1 Mexique (Séisme de Michoacan, 1985)

Le séisme de Mexico de Michoacan en 1985 a fortement touché la ville de Mexico située à 400 km de l'épicentre (Figure 2-1). Les accélérographes enregistrés montrent une décroissance depuis la côte Pacifique jusqu'à proximité de Mexico. A Mexico même, on observe une amplification très importante des ondes associées à une modification du contenu fréquentiel du signal. Cette modification du contenu fréquentiel du signal est due à la résonance du bassin sédimentaire sur lequel la ville de Mexico est construite.

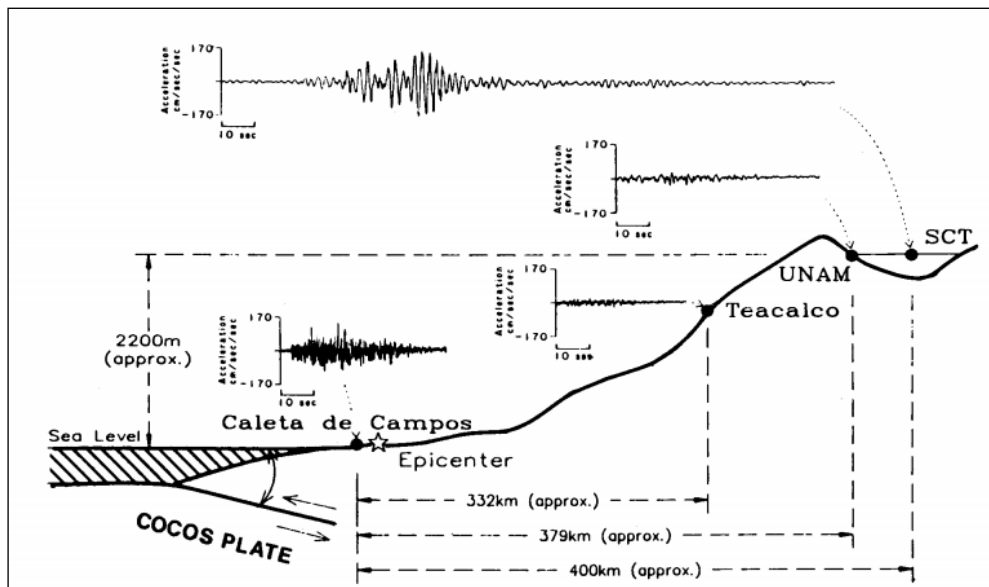


Figure 2-1 Enregistrements du tremblement de terre de Michoacan (1985) de la côte Pacifique jusqu'à Mexico (Çelebi, et al, 1988).

Le séisme de Michoacan ne causa que des dégâts modérés près de son épicentre (par exemple au niveau du site Caleta de Campos), mais provoqua de sérieux dégâts à Mexico City (par exemple au droit de la station SCT) à presque 400 km de l'épicentre. Ce constat donnait à penser que l'effet de site au niveau de Mexico City avait joué un rôle important dans l'amplification des effets du séisme. Le sol de Mexico City peut être divisé en trois sous-parties :

Foothill Zone (FZ) : sol granulaire, basalte, située au sud-ouest de la ville.

Lake Zone (LZ) : d'épais dépôts de sols mous, constitués de limons, argile, cendres provenant des volcans alentour.

Transition Zone (TZ) : située entre les deux précédentes, où les deux types de sols précédents se chevauchent.

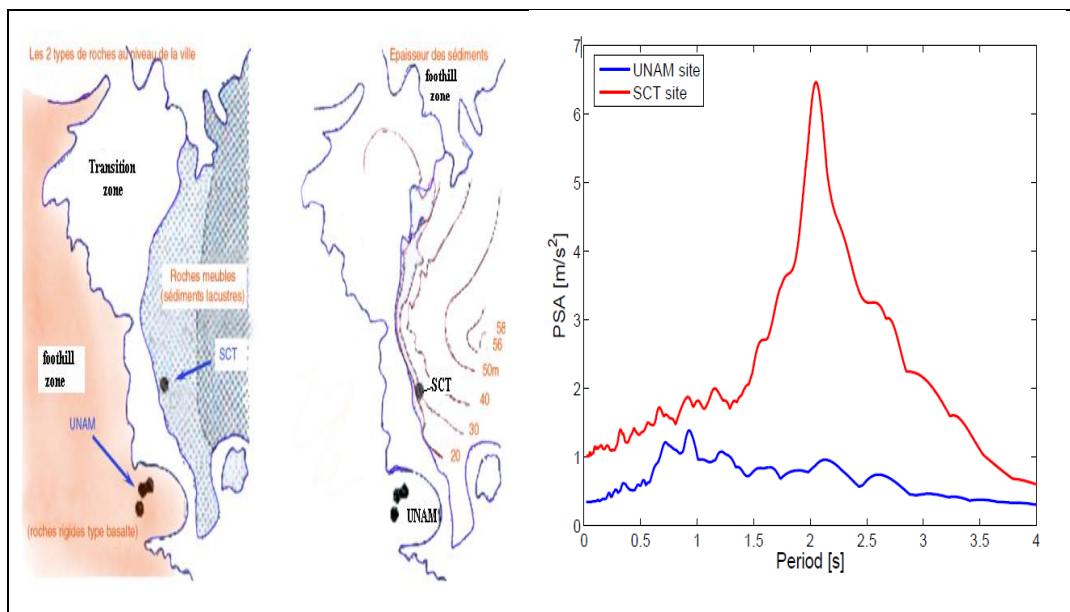


Figure 2-2 Amplitudes maximales selon les zones de Michoacan-Mexico à gauche et à droite le pseudo spectre d'accélération en fonction de la période (Stewart, 1997).

L'épicentre étant situé assez loin de Mexico City, il n'a provoqué que de faibles accélérations au niveau de la FZ (0,03 g à 0,04 g). Dans la TZ, les accélérations ont été sensiblement identiques, mais dans la LZ elles ont été jusqu'à 5 fois supérieures. De plus, le contenu fréquentiel des accélérations montre lui aussi une différence entre les sites de la LZ (Station SCT) et ceux de la FZ (Station

UNAM). Pour une période de 2 s, la pseudo accélération au niveau de SCT est presque 10 fois supérieure à celle obtenue pour la station UNAM.

2.2.2 USA (Séisme de Loma Prieta, 1989)

Le 19 Octobre 1989, un tremblement de terre de magnitude $M_s=7,1$ s'est produit au Mont Loma Prieta, à 100 km au sud de San Francisco, son intensité a été, là encore, plus forte à San Francisco qu'à l'épicentre. De plus le fait que ce séisme ait provoqué des dégâts importants dans certaines régions de la ville, et pas dans d'autres, laissait supposer que l'effet de site avait eu un rôle important.

Dans la baie de San Francisco, le sol est essentiellement composé d'alluvions, et ceux-ci ont, dans certains cas, été sur consolidés et dans d'autres cas ne l'ont pas été. Il en résulte que cette baie peut être divisée en trois régions, l'une de boue de San Francisco (argile limoneuse normalement consolidée), une autre d'alluvions, et la dernière de la roche.

De manière générale, il a été remarqué que l'atténuation des ondes sismiques c'est produit plus rapidement dans la zone rocheuse. Deux instruments ont été installés dans la baie, sur deux îles très proches mais de nature géologique différentes : Yerba Buena Island et Treasure Island. La première est naturelle en affleurement de la roche. La seconde est artificielle et repose sur un banc de sable. Entre ce banc de sable et l'île, on trouve une quinzaine de mètres de boue de San Francisco.

Tableau 1 Valeurs de l'accélération maximale des trois composantes enregistrées sur les sites de Treasure Island et Yerba Buena Island.

Sites	Comp (N-S) _{max}	Comp (E-W) _{max}	Comp (Verticale) _{max}
Treasure Island	0.10g	0.16g	0.02g
Yerba Buena Island	0.03g	0.07g	0.03g

Lors du séisme, les deux instruments, bien que spatialement très proches, ont enregistré des accélérations très différentes : 0.06 g pour Yerba Buena contre 0.16 g pour Treasure, dans la direction Est-Ouest. Ce si est dû à la présence d'une

couche plus " molle " (la boue de San Francisco) qui a provoqué l'amplification du mouvement au niveau de Treasure Island.

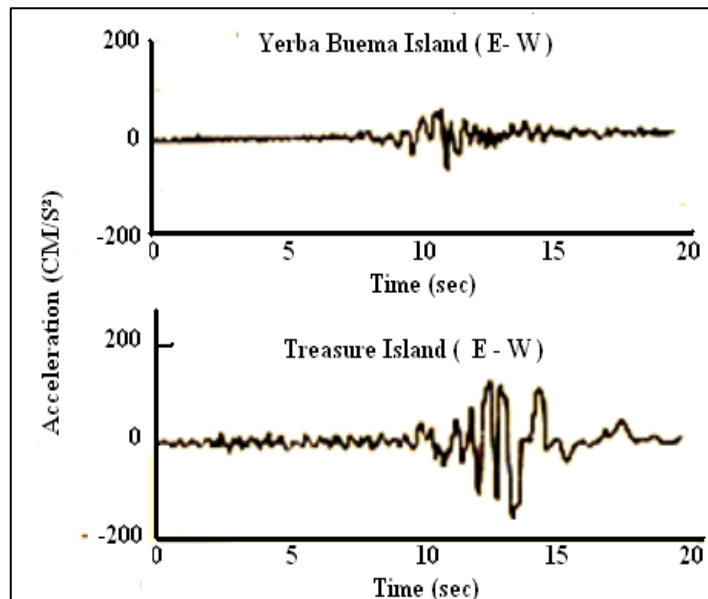


Figure 2-3 Réponses spectrales d'accélération enregistrées au niveau de Treasure Island et Yerba Buena Island (Stewart, 1997).

C'est le même effet qui a provoqué la destruction du Cypress Viaduct : une partie de cet ouvrage reposait sur un sol mou, la boue de San Francisco, tandis que l'autre partie, qui elle a résisté, ne l'était pas.

2.2.3 Japon (Séisme de Kobe, 1995)

Le tremblement de terre de Hyogo-Ken Nanbu de Kobe du 17 janvier 1995 de magnitude 7.2 sur l'échelle JMA (ou 6.9 sur l'échelle de Richter) a causé des dégâts considérables dans la région de Kobe, en particulier dans une zone localisée désignée sous le nom « ceinture des dommages », où la plupart des bâtiments se sont effondrés.

La ceinture de dommages a 1km de largeur et 30 kilomètre de longueur, située à environ 1 kilomètre de la faille de Rokko (Figure 2-4).

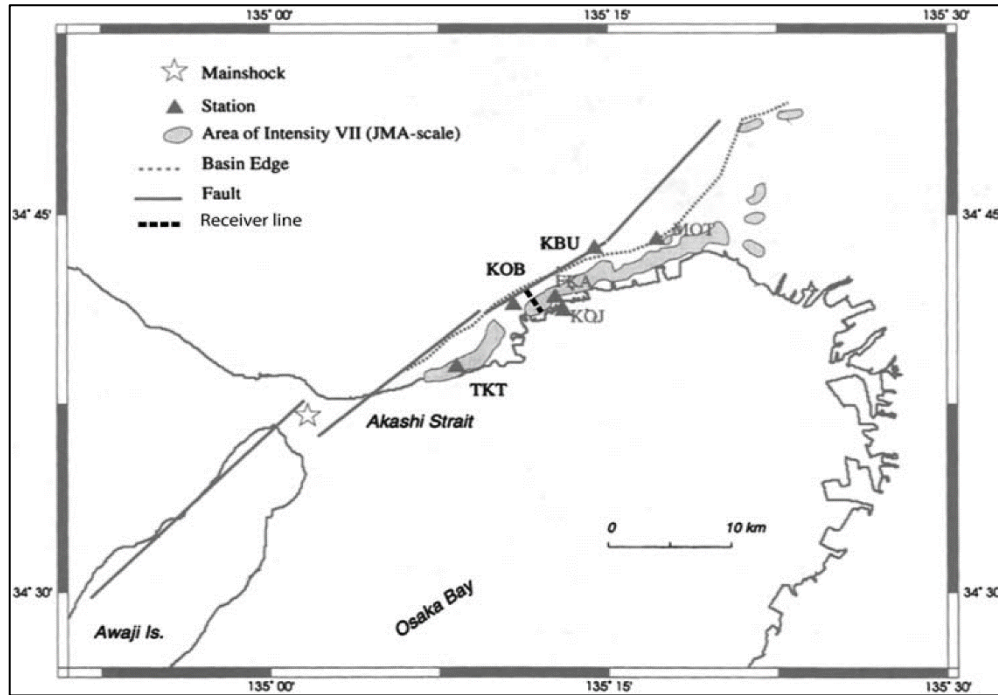


Figure 2-4 Carte du bassin d'Osaka montrant la ceinture des dommages défini comme un secteur où l'intensité en JMA a dépassée VII. (Pitarka et al, 1998).

Le tableau (Tableau 2) montre les valeurs de la vitesse maximale en (m/s) enregistrée sur certaines stations ; le radial correspond à la composante perpendiculaire de la faille et la transversale correspond à la composante parallèle à la faille. Notez la grande amplification de la composante radiale à la station de TKT, situé dans la ceinture des dommages (Pitarka et al., 1998).

Tableau 2 Valeur de la vitesse maximale en (m/s) lors du séisme de Kobe.

Station	Radial	Transversale	Verticale
KBU	0.36	0.39	0.18
KOJ	0.86	0.38	0.34
KOB	1.02	0.56	0.49
TKT	1.57	0.63	0.13

Afin d'expliquer cette répartition des dommages particuliers, Kawase en 1996 a invoqué un effet de propagation des ondes du à l'effet de bord du bassin. Ce dernier est causé par l'interférence entre le signal directe des ondes S incidentes sur le bassin et les ondes de surface diffractées au bord du bassin.

2.2.4 Algérie (Séisme de Boumerdes, 2003)

Le tremblement de terre de Boumerdes du 21 mai 2003 de magnitude 6.9 sur l'échelle de Richter a causé des dégâts considérables :

- ✓ Des milliers de mort et plus de 5 Milliards de dollar de dégâts
- ✓ Des milliers d'immeubles étaient touchés à Alger, à Bordj Ménaïel alors que l'épicentre du séisme était à plus de de 60 km (Yelles-Chaouche. et al, 2003).

Le séisme a été même ressenti jusqu'à 250 Km de l'épicentre où l'accélération enregistrée a atteint 0.02g. Les stations Kaddera 1 et 2 situées à 20 Km de l'épicentre ont enregistré le mouvement. Le plus étonnant dans l'enregistrement de ces deux stations c'est la différence importante voir énorme (Station 1 un PGA de 0.34g, Station 2 un PGA de 0.58g dans la direction est-ouest). De plus l'analyse montre que les signaux enregistrés en champ proche (Keddara) sont assez riche en haute fréquence qui préconise la forte présence de l'effet de site c'est-à-dire que l'onde sismique reste piégée.

2.3 Effets de site : Définition et mesures

2.3.1 Définition de l'effet de site

L'effet de site est défini comme étant l'amplification d'une quantité sismique (Accélération, vitesse, etc....) enregistrée à la surface libre. Il se mesure par le ratio de cette quantité en surface et en profondeur (ou à un affleurement rocheux). Cette amplification se produit principalement dans deux cas : au sommet des collines et des montagnes (effet de site causé par la topographie), et/ou lorsque une vallée a fait l'objet d'une importante sédimentation peu consolidée (par exemple vallée remplie par les alluvions récentes, les marais ou les tourbières). Il est à noter que le terme "Effet de site" fait référence dans la plupart des travaux à ce dernier cas.

Dans les deux cas, les ondes sismiques sont piégées dans la structure (dans la montagne, elles se réfléchissent d'un côté à un autre ; dans la vallée, une fois elles ont pénétré dans la zone d'alluvions, meubles elles ne peuvent pas s'échapper et se

réfléchissent sur les côtés et sur le fond de la vallée). Les études théoriques montrent que l'amplitude des ondes peut être amplifiée d'un facteur de presque 10 par de tels effets. Par ailleurs, dans le cas de la vallée, l'effet disparaît dès que l'on quitte la zone d'alluvions et que l'on gagne le flanc rocheux de la vallée.

La figure (Figure 2-5) illustre ce phénomène. Il s'agit du séisme de Lancy ($M = 5.2$). L'examen de cette figure montre qu'il vaut mieux construire sur le rocher qu'en plein milieu d'une plaine alluviale si l'on veut éviter des dommages liés aux effets de site. En revanche, monté trop haut sur le flanc de la vallée donne naissance à un effet de site topographique, car on se rapprocherait alors de l'autre cas (sommet de montagne) qui donne lui aussi des amplifications du mouvement du sol.

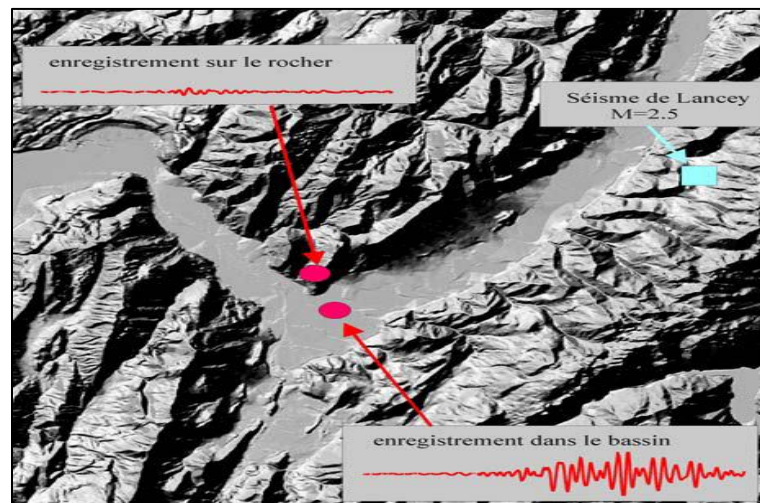


Figure 2-5 Comparaison d'enregistrements sismiques sur la roche et aux sédiments du séisme de Lancy de 1995 (Lebrun, 1997).

Les effets de stratigraphie se manifestent lorsque les variations latérales des propriétés mécaniques et géométriques du sous-sol sont négligeables comme dans le cas des vallées peu encaissée et des bassins sédimentaires. Le modèle de sol est constitué d'une superposition de couche homogènes, horizontales d'épaisseur constante surmontant un substratum rocheux.

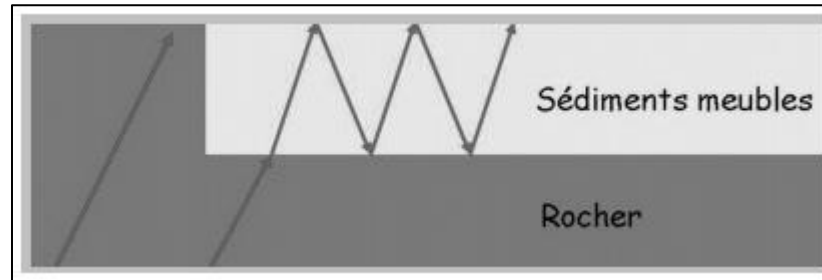


Figure 2-6 L'effet de site (piégeage des ondes).

Les causes des effets de site dans ce cas sont dues aux manifestations de deux phénomènes : le premier lorsque le contraste d'impédance est marqué entre les deux milieux, l'amplitude de l'onde réfractée est amplifiée par rapport à celle de l'onde incidente et Le deuxième cause quand les ondes sismiques arrivant à la surface se réfléchissent aux interfaces entre différentes couches et restent piégées aux abords de la surface ; l'énergie est donc piégée à l'intérieur de la couche sédimentaire superficielle.

2.3.2 Méthodes de mesure de l'effet de site

2.3.2.1 Cadre général

Les effets de site que l'on peut observer présentent l'intérêt d'être prévisibles, cependant, ce n'est que depuis une trentaine d'année avec le développement des techniques de simulations numériques et le nombre croissant des stations sur le terrain que sa quantification s'est révélée possible. La mesure de l'effet de site se fait par le biais des méthodes expérimentales et les numériques.

Les méthodes expérimentales se basent sur l'analyse des enregistrements dans un site particulier. Le principe général de l'utilisation des enregistrements de séismes consiste à comparer le signal enregistré sur le rocher à celui enregistré dans le bassin. Bard et Riepl-Thomas (2000) ont répertorié quatre techniques expérimentales pour évaluer les effets de site (*Tableau 3*).

Tableau 3 Catégories de techniques d'évaluation expérimentales de l'effet de site (Bard et Riepl-Thomas, 2000).

Données sources	Méthodes	Informations obtenues	Limites	Coût
Observations Macrosismiques		-Corrélation avec la nature du sous-sol -Micro zonage	Informations uniquement Qualitatives	Faible
Enregistrement de micro-tremblements provenant des vibrations ambiantes : « bruit » du vent, des vagues, du trafic, des machines industrielles	Spectres de micro-Tremblements	-Corrélations avec l'épaisseur des couches molles sur le substratum -Estimation de la période prédominante du sol et du filtrage des fréquences -Pour les périodes longues, estimation des périodes propres du site et des amplifications correspondantes	-« Bruit » caractéristique d'un site particulier ; effets de source et de cheminement -Nécessité d'un suivi très rigoureux en raison de la différence de « bruit » entre le jour et la nuit -Utilisation qualitative Uniquement	Faible
	Coefficients spectraux de bruit			Faible
	Rapport H/V (technique de Nagoshi-Nakamura)			Une des techniques moins chères
	Méthode SASW : enregistrements en série (Aki)			Calculs nécessitant des ordinateurs puissants
Enregistrement de séismes d'intensité faible à modérée	Spectres avec site de référence	Estimation de la réponse du Site	-Informations locales -Pas d'information sur la réponse d'un site au comportement anélastique	-Utilisation d'instruments de mesure sensibles -Besoin d'au moins 20 enregistrements -Plusieurs mois nécessaires pour l'instrumentation, le recueil des données et l'analyse -Amortissement de l'investissement avec le nombre d'investigations
	Spectres sans site de référence			
Informations sur les grands mouvements sismiques		-Estimation de la réponse, même si le comportement du sol est non linéaire -Micro zonage	-Possibilités d'utilisation rares.	-Installation et maintenance de l'instrumentation à l'échelle d'une grande ville

Les méthodes numériques se basent sur une représentation précise de la vallée et sur les paramètres géophysiques et géotechniques du site à étudié.

Pour simuler correctement la propagation des ondes dans une vallée, il faut connaître le profil de vitesse des ondes, la nature des matériaux, la profondeur des sédiments ou encore la taille de la vallée mais cela demande des investigations géophysiques très coûteuses en particulier pour des forages profonds selon la

complexité de la vallée, la précision souhaitée, et la capacité de calcul dont on dispose, on va utiliser des modélisations 1D, 2D ou 3D.

L'avantage de ces méthodes numériques est de pouvoir faire varier les différents paramètres caractéristiques du modèle et de comprendre physiquement les phénomènes qui rentrent en jeu. Par contre, leur limitation provient de nécessité de connaître précisément toutes les caractéristiques géométriques et dynamiques des formations ainsi que le champ d'ondes incident. De plus, les hypothèses portant sur le comportement du sol sont aussi très contraignante. La figure ci-dessous illustre les différentes approches possibles dans l'étude de l'effet de site.

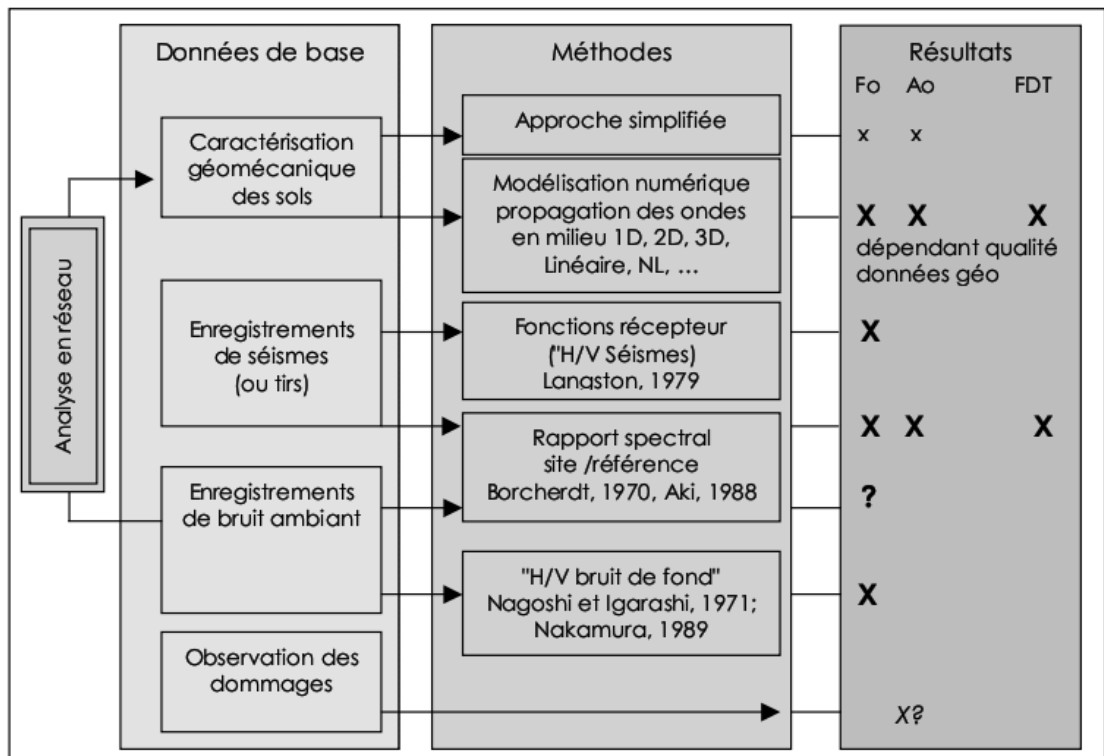


Figure 2-7 les différentes méthodes adoptées pour déterminer les effets de site. (Cadet et al 2007). où (f₀) représente la fréquence de résonance des sites, (A₀) l'amplification correspondante provoquée par le site à cette fréquence et (FDT) la fonction de transfert.

2.3.2.2 Mesures expérimentales <<in situ et observations sur site>> (Réseau en puits) :

Cette méthode est basée sur la comparaison des enregistrements en surface et en profondeur d'une sollicitation donnée sur le même site (Figure 2-8). Celle-ci

donne un aperçu sur le degré d'amplification surface/profondeur (par exemple le rapport spectral entre les deux signaux) qui permet de caractériser l'effet de site. Cette méthode n'est pas largement utilisée car elle requiert un dispositif expérimental très coûteux.

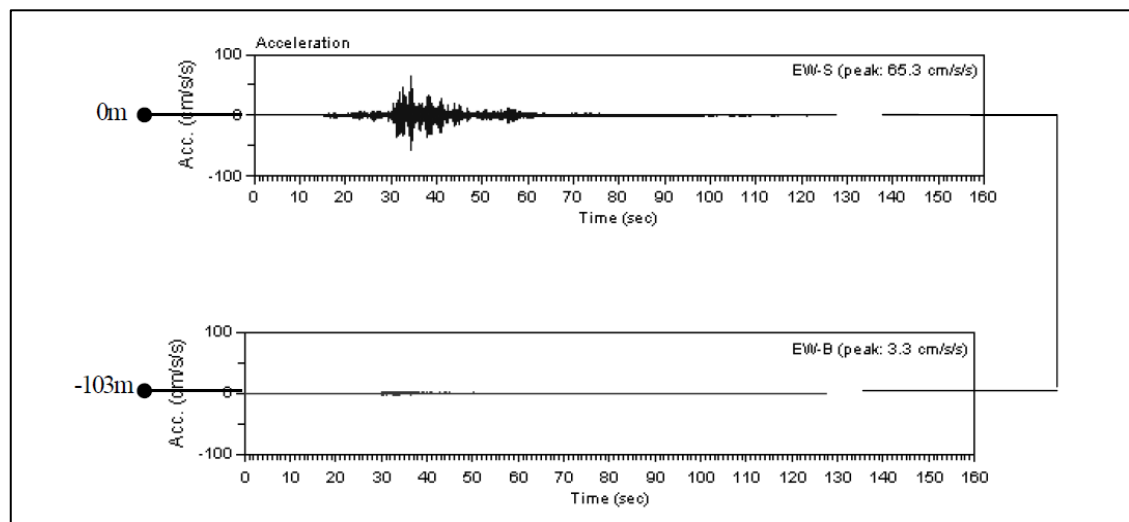


Figure 2-8 Enregistrements en surface et en profondeur (-103m) sur une même station (KIK-NET) (Derras et al, 2004).

2.3.2.3 Méthodes H/V :

Cette méthode est appelée aussi HVSR (horizontal vertical spectral ratio), elle est très intéressante du fait de sa simplicité et son économie, elle est basée sur le rapport spectral entre la composante horizontale et la composante verticale des enregistrements en surface (Duval et al, 1994; 1998). La pratique montre pourtant qu'elle est assez efficace pour évaluer la fréquence fondamentale des effets de site (Bard et al, 1998). Il est signalé que les spectres utilisés proviennent uniquement de l'enregistrement de bruit de fond.

Cette méthode a été introduite pour la première fois par Nogoshi (Nogoshi, 1971) et complétée plus tard par Nakamura en 1989 (Nakamura, 1989)), la méthode est basée sur les enregistrements des bruits ambiant générés par le trafic, les activités urbaines ...le vent et autre. Nakamura a montré que l'effet de site peut être exprimé par le rapport spectral de la composante horizontal et vertical du bruit ambiant à la surface du fait que ce rapport est lié à la fréquence fondamentale du

sol et par conséquent au facteur d'amplification. Plusieurs études ont été effectuées avec cette méthode montrant que cette dernière a une bonne stabilité expérimentale et qui peut être utilisée pour obtenir la fréquence fondamentale de résonance d'un site à l'aide de la formule développée par Theodulidis en 1995, indiquée sur la figure (Figure 2-9) où S_{EW} , S_{NS} et S_{UP} sont les spectres de Fourier des accélérogrammes enregistrés à la surface de la terre.

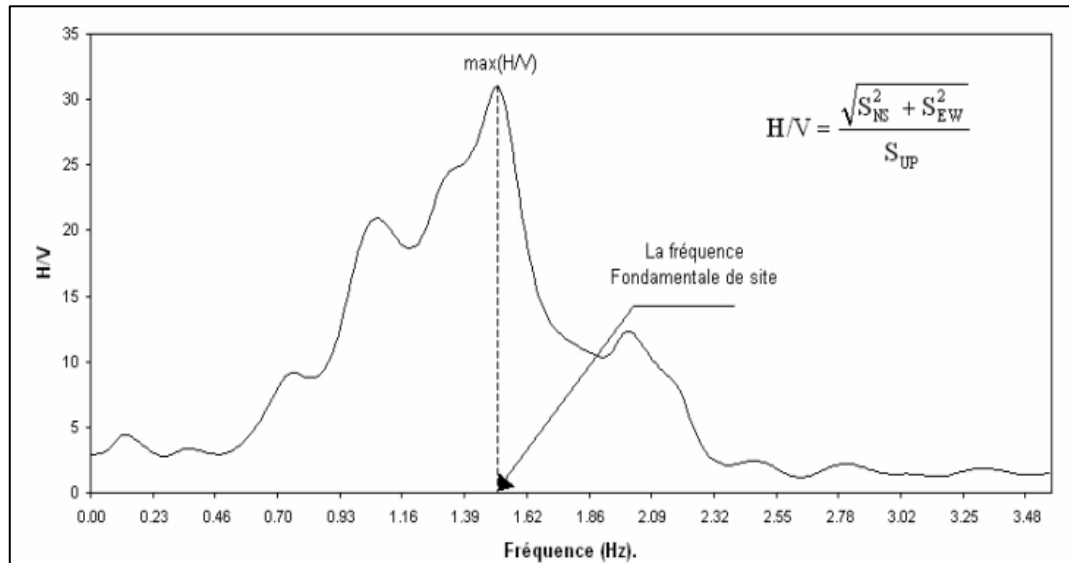


Figure 2-9 HVSR de l'événement du 16/09/2002 enregistré sur la station TTRH03 (KIK-NET).

En 1998 Riepl et al ont appliqué cette méthode aux ensembles de données de mouvement faibles et forts. En 1999 Bard considère que le pic H/V a essentiellement pour origine les propriétés des ondes de surface. En 2004 Bonnefoy-Claudet et al, à partir de simulations numériques de bruit de fond sismique, ont étudié en détail l'origine du pic H/V : dans le cas de milieux stratifiés et quel que soit la contraste d'impédance (forte, modéré, faible), on conclut que la fréquence du pic H/V est très proche de la fréquence de résonance du site.

Le principal problème de cette méthode c'est qu'elle est utilisable dans des régions à faible sismicité car l'utilisation dans les mouvements faibles (petits séismes, bruit de fond sismique) et le comportement non linéaire des sols, c'est-à-dire que la réaction du sol due aux mouvements forts (dont on souhaite déterminer les effets) n'est pas toujours proportionnelle à celle aux mouvements faibles. En

particulier, les sols très mous amplifient moins les mouvements forts car ils se déforment de façon irréversible et amortissent l'onde qui s'y propage.

2.3.2.4 Méthodes numérique

Les modèles unidimensionnels (1D) intègrent une représentation stratigraphique du sous-sol et vue leur simplicité de mise en œuvre et d'application, ils sont très utilisés en génie parasismique (interaction sol/structure). Certains logiciels (SHAKE, EERA, Deepsoil,...) permettent même d'intégrer des lois de comportement linéaire équivalent ou non-linéaire dans les différentes couches. L'analyse unidimensionnelle est également à la base des normes parasismiques pour la prise en compte de l'influence de l'hétérogénéité verticale du sol sur le mouvement sismique de surface.

L'approximation 1D montre cependant ses limites quand le sous-sol présente des hétérogénéités latérales et cela se traduit par la réponse sismique qui n'est plus négligeable (Pitilakis, 1999). Du coup, cette méthode n'est valable qu'à l'échelle très locale d'un bâtiment isolé ou dans le cas des vallées alluvionnaires peu encaissées (Semblat, 2000b).

2.3.2.5 Réseau de neurones artificiels :

Derras et al, (2010) ont élaboré un système de plusieurs RNA. Le but est de tester la capacité de la méthode neuronale à tracer le spectre de réponse à la surface libre de la terre à partir de celui enregistré en profondeur est présentée. Pour ce faire, les auteurs ont utilisé un sous ensemble de la base de donnée KiK-net (tableau ci-dessous).

Tableau 4 Informations sur les conditions de site et le nombre de spectre de réponse utilisé pour élaborer le système neuronal

Classes	Code des Profils	V_{s30} (m/s)	H_T (m)	H_S (m)	F_0 (Hz)	Nombre total des enregistrements utilisés	Le nombre des enregistrements utilisés dans la phase d'apprentissage	Le nombre des enregistrements gardés pour le test
A1	KGWH02	180	200	54	1.09	22	21	1
B2	EHMNH04	254	200	64	1.41	14	13	1
	SMNH07	318	209	60	1.74	12	11	1
	EHMNH07	391	200	20	3.2	15	14	1
C3	SMNH03	445	101	34	3.35	12	11	1
	SMNH02	503	101	25	4.6	12	11	1
	KOCH03	668	100	32	5.30	16	15	1
	SMNH05	711	101	14	7.5	12	11	1
D4	EHMNH03	802	101	10	11.42	13	12	1
	KOCH05	1072	100	6	15.42	12	11	1

Les spectres de réponses en accélération (S_a) adoptés sont de l'ordre de 140 couples entrés/sorties [les deux composantes NS (nord-sud) et EW (est-ouest) sont utilisées] c'est-à-dire 140 (S_a) à la surface libre de la terre et 140 (S_a) enregistrés à une profondeur variant entre 100m à 200m. Le nombre des (S_a) réservés à l'apprentissage est de 130 et 10 (S_a) sont conservé à la phase test.

Le système neuronal est un ensemble de RNA. Chaque RNA est caractérisé par une classe de site (tableau ci-dessous) et par la valeur maximale spectrale de

référence. L'apprentissage est effectué avec la technique de la rétro propagation avec un pas constant.

Tableau 5 Classification de site utilisée dans le choix du RNA

Type de sol	Vitesse de cisaillement V_{s30} (m/s)	La fréquence de résonance f_0 (Hz)
A1 : très meuble	<200	<1.67
B2 : meuble	$200 \leq V_{s30} < 400$	$1.67 \leq F_0 < 3.33$
C3 : Ferme	$400 \leq V_{s30} < 800$	$3.33 \leq F_0 < 6.67$
D4 : rocheux	≥ 800	≥ 6.67

Trois paramètres de site ont été utilisés, il s'agit de H_{800} qui représente l'épaisseur jusqu'à une vitesse de cisaillement égale à 800 m/sec et la fréquence de résonance f_0 ainsi que le V_{sm} (vitesse moyenne sur tout le profil). Les inputs de chaque RNA sont l'ordonnée spectrale en accélération et la période correspondante. La seule sortie est donnée par l'ordonnée spectrale en accélération à la surface libre. L'architecture du RNA est donnée sur la figure ci-dessous.

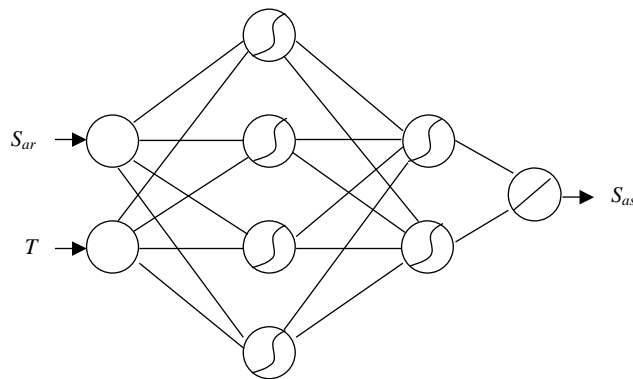


Figure 2-10 Architecture du réseau de neurone

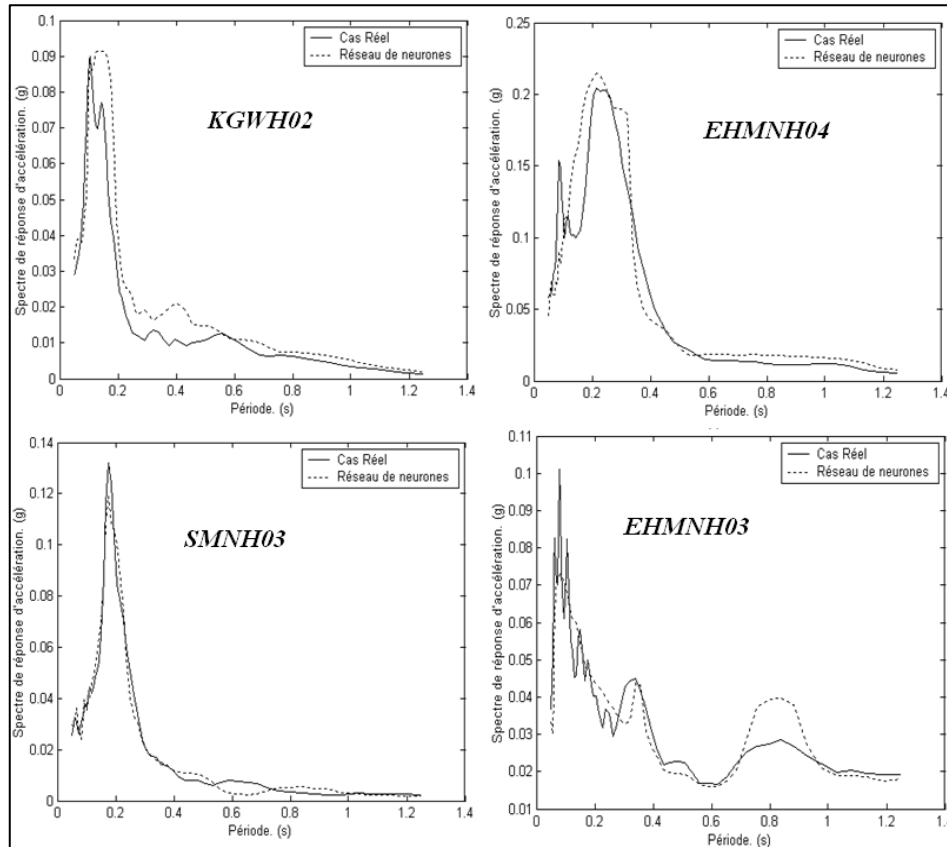


Figure 2-11 Spectres de réponse d'accélération estimés par RNA et calculés à partir des enregistrements réels.

Les 10 spectres de réponse gardés pour la phase test ont été représentés conjointement avec ceux simulés par le système neuronal. On donne ici quatre exemples (figure ci-dessous). Une lecture rapide de ces tracés montre la ressemblance entre les cas réels et les simulations. Le problème de l'estimation des pics spectraux persiste toujours : dépendance de la base de données initiale.

Dans cette étude les auteurs ont utilisé les mêmes sites que ceux de la phase apprentissage, et cela explique la ressemblance existante entre les enregistrements et les simulations. Les manques observés dans le présent article sont aux niveaux de la méthode d'apprentissage, du non prise en compte du problème de sur-apprentissage et de la limitation dans la base de données utilisée pour l'apprentissage.

2.3.3 Analyse critique et identification de la problématique

Il apparaît de la section précédente que les méthodes de mesures de l'effet de site sont multiples pour au moins deux raisons : la méthodologie de la mesure (Expérimentale vs numérique) le type de la quantité sismique utilisée pour caractériser le mouvement sismique. Ainsi pour un site donné et suite à un séisme, l'effet de site peut être observé ou non suivant l'une des approches citées dans la section précédente. Toutefois, sa quantification par des approches différentes ne donnera pas la même valeur. La raison étant que le facteur d'amplification est relié à deux entités : les caractéristiques sismiques et les caractéristiques de site. Ces deux entités sont complexes et ne peuvent pas être facilement caractérisés. A titre d'exemple, considérons les caractéristiques de site. Un sol est caractérisé par ses propriétés mécaniques et géométriques qui peuvent être par exemple la vitesse de cisaillement et les épaisseurs de couche. Un modèle qui permet d'obtenir le facteur d'amplification en utilisant ces propriétés est compliqué à mettre en place du fait de leur nombre important.

Il serait donc intéressant de mettre en place une approche qui permet d'obtenir le facteur d'amplification en fonction d'un certain nombre de paramètres de site qui peuvent caractériser le site. Cette tâche est en théorie simple à mettre en place mais en pratique reste assez compliqué du fait de la complexité de l'introduction des caractéristiques sismiques.

L'objectif de la présente thèse trouve son fondement dans la recherche d'une relation entre le facteur d'amplification et un certain nombre de paramètres de site. Alternativement, cet objectif en appelle un autre quel est le nombre minimum de paramètres de site pour estimer "correctement" un facteur d'amplification.

2.4 Conclusion

Après cette brève revue bibliographique sur l'importance des dégâts causés par l'effet de site ou il est impérativement claire que l'effet de site doit être pris dans les code parasismique afin de bien dimensionner nos ouvrages. La prise de compte de

cette effet se traduit pratiquement par le facteur d'amplification qui le rapport entre le spectre en surface et en profondeur.

L'une de ces approches qui se situe, en point de vue méthodologique, à mi-chemin entre les approches expérimentales et numériques est l'approche neuronale. Cette dernière se base essentiellement sur les données expérimentales en utilisant le concept numérique. Cette approche peut être intéressante du moment où les données d'enregistrements et de sites sont de plus en plus disponibles. Si on suit cette logique, les réseaux de neurones artificiels peuvent être utilisés pour estimer les paramètres caractérisant le phénomène d'effet de site et ce à la place des méthodes conventionnelles. L'utilisation de la technique neuronale s'avère intéressante quand le nombre des variables prises en compte est élevé (Ghaboussi et al.,1998 ; Hashash et al.,2004 ; Derras et al.,2012 ; Boudghene Stambouli et al.,2017) et quand la base de données est petite par rapport au domaine d'investigation et c'est le cas de la plupart des base de données par tout à travers le monde.

L'idée principale de la présente thèse est d'identifier les effets des paramètres de site sur le facteur d'amplification. Le chapitre suivant est consacré à la mise en place de la méthodologie.

Chapitre 3

Mise en place de la méthode

3.1 Introduction

L'objectif principal du présent travail de recherche consiste à élaborer une relation entre le facteur d'amplification et un certain nombre de paramètres de site. Alternativement, cet objectif en appelle un autre celui du nombre minimum de paramètres de site à utiliser pour estimer "correctement" un facteur d'amplification. Afin d'atteindre ce double objectif, on se propose dans cette partie de mettre en place la méthode pour la construction de la base de donnée.

3.2 Méthode de calcul du facteur d'Amplification (AF)

3.2.1 Définition et hypothèses de base

Considérons une quantité sismique ' Ω '. Cette quantité peut être l'accélération maximale du signal sismique (PGA), sa vitesse maximale (PGV) ou toute autre quantité. Le facteur d'amplification (Amplification Factor, AF) est défini par l'expression suivante :

$$AF = \frac{\Omega_{Surface}}{\Omega_{Substratum} \text{ ou } \Omega_{Roche \text{ Affleurante}}} \quad (3-1)$$

Une telle façon de mesurer l'amplification est très pratique car le facteur AF est adimensionnel. La quantification de l'effet de site par le facteur d'amplification

dépend donc du type de la quantité sismique. Il est donc préférable d'utiliser des quantités couramment utilisées dans l'ingénierie. Dans le présent travail on utilisera le pseudo spectre en accélération comme quantité sismique du fait de sa large utilisation par les ingénieurs. Toutefois, la méthode qui sera mise en place reste valable quel que soit la nature de la quantité sismique. Par ailleurs la mesure de cette quantité au niveau du substratum est en soit assez contraignante et on lui préfère souvent la mesure au niveau de la roche affleurant.

L'évaluation du AF s'effectue en considérant un certain nombre d'hypothèses :

a) Signal sismique

Le choix du signal sismique joue un rôle important surtout que la quantité sismique choisie est le spectre de réponse. Pour ce faire un choix judicieux de ces signaux est très important. On aura donc à choisir des signaux dont le contenu couvre l'ensemble des fréquences.

b) Type de modélisation du comportement de sol :

On a considéré que le sol a un comportement viscoélastique linéaire.

c) Profil de sol :

On considère que le sol, stratifié horizontalement, repose sur un substratum

3.2.2 Signal sismique

Le signal sismique peut être caractérisé par plusieurs paramètres avec toutefois deux paramètres qui sont les plus en vue : la valeur maximale et le contenu fréquentiel. L'allure du pseudo spectre d'accélération notée $PSA(T)$ ou $PSA(f)$ avec T la période de la structure et f la fréquence de la structure permet d'identifier le contenu fréquentiel du signal. Sa valeur maximale peut être par contre facilement identifiée.

Initialement 73 enregistrements ont été récoltés. L'idée consiste à sélectionner des événements cohérents en effectuant les opérations suivantes :

- a) Filtrer le signal : fréquence de filtrage $>0.15\text{Hz}$, afin d'éliminer le bruit à basse fréquence (fréquences parasites dans le signal).

- b) Calculer les 73 PSA(T)
- c) Normaliser les 73 PSA par rapport au PGA correspondant.

Cette étape permet de réduire la variation dans l'allure par rapport aux spectres non normalisés.

- d) Localiser la gamme de fréquence où la valeur maximale du PSA(T) est préconisée.

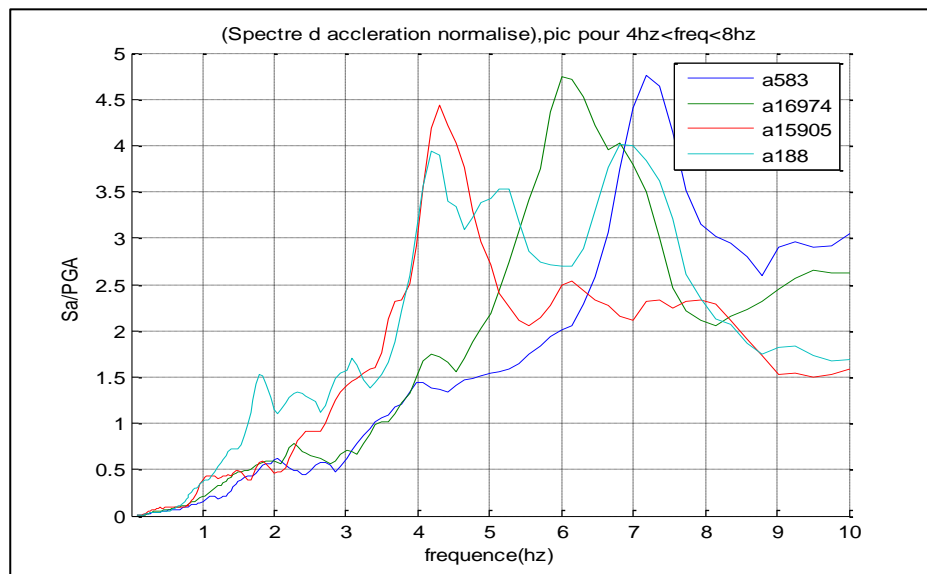


Figure 3-1 Exemple de pseudo spectre d'accélération avec un amortissement de 5%.

A cet effet, nous avons choisi 4 plages de fréquence, à savoir ($f < 2\text{Hz}$), ($2\text{ Hz} < f < 4\text{ Hz}$), ($4\text{ Hz} < f < 8\text{ Hz}$) et ($8\text{ Hz} < f < 16\text{ Hz}$). A titre d'exemple, nous présentons sur la figure (Figure 3-1) les PSA normalisés qui correspondent à la gamme de fréquence [4 et 8] Hz.

A l'issue de cette étape qui consiste à choisir des signaux qui couvrent une large gamme de fréquence en pic. Cependant, il ne reste que 14 signaux au lieu de 73 répondant à ce critère (Tableau 6). Pour éviter la variabilité liée à échantillonnage, un pas de temps unique de 0.01 sec a été choisi. Si le pas du signal est inférieur à 0.01 sec, un filtrage passe bas à 30 Hz est adopté afin d'éviter des coupures brutales. De ce fait, nous avons utilisé la méthode Butterworth qui présente une coupure s'atténuant linéairement. Cette dernière a été décrite pour la première fois Stephen Butterworth $T(f)$ (Figure 3-2) et est défini par:

$$T(f) = \frac{1}{\sqrt{1+(f/f_c)^{2n}}} \quad (3-2)$$

Avec n : ordre généralement pris égal à 4.

$T(f)$: Le gain du filtre.

f_c : Fréquence de coupure du filtre.

f : Fréquence angulaire du signal en rad/s.

La figure ci-dessous présente la fonction de transfert du filtrage passe bas avec une fréquence de coupure de 30 Hz et n pris égal à 4.

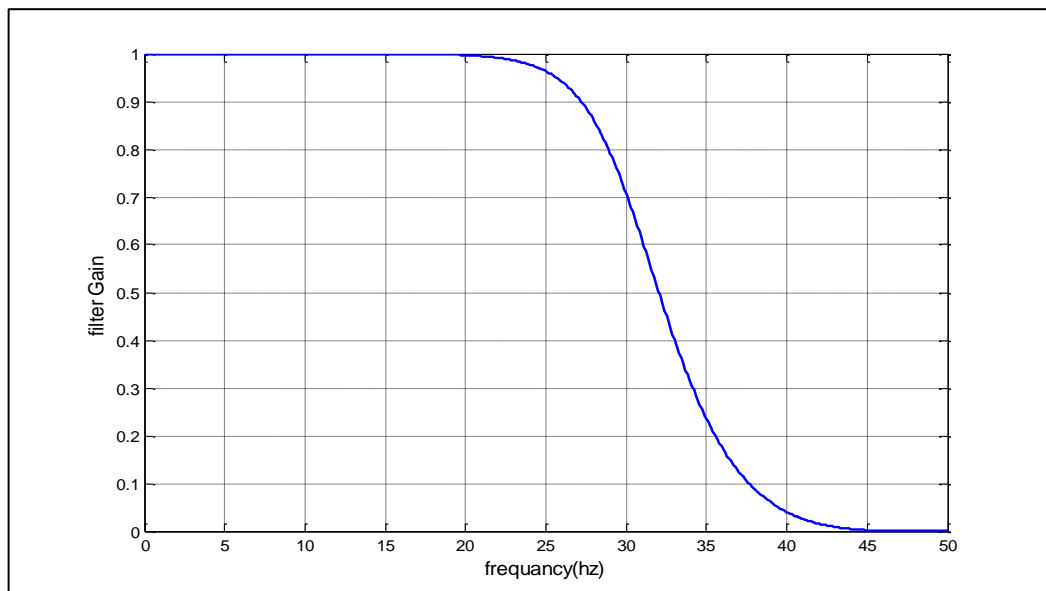


Figure 3-2 Gains de filtres de Butterworth passe-bas d'ordre 4 en fonction de la fréquence.

En ce qui concerne les signaux qui ont un pas de temps supérieur à 0.01 sec, une interpolation linéaire a été utilisée, le tableau ci-dessous résume les étapes déjà évoquées.

Tableau 6 Résumé des séismes choisis.

Identification	RESO RCE	PGA (m/s ²)	Earthquake	Station	Mw	Distance	Frequency content of the highest PSA
S1	16149	1.269	Manjil (Iran) 20/06/1990	Zanjan	7.7	56	f<2Hz
S2	16351	1.288	Olfus (Iceland) 29/05/2008	Ljosafoss-Hydroelectric Power Station	6.1	11	f<2Hz
S3	16352	3.190	Olfus (Iceland) 29/05/2008	Selfoss City Hall	6.1	3	f<2Hz
S4	15537	2.074	South Iceland 17/06/2000	Thjorsarbru	5.0	10	2Hz<f<4Hz
S5	6756	3.106	South Iceland 17/06/2000	Flagbjarnarholt	6.6	15	2Hz<f<4Hz
S6	6765	3.140	South Iceland 17/06/2000	Thjorsarbrun	6.6	4	2Hz<f<4Hz
S7	583	1.0030	Lazio-Abruzzo 11/05/1984	Villetta Barrea	5.3	2	4Hz<f<8Hz
S8	16974	1.4740	L'Aquila 07/04/2009 (aftershock)	Pescomaggiore	5.3	10	4Hz<f<8Hz
S9	15905	3.0060	Firouzabad (Iran) 20/06/1994	Zarrat	5.8	11	4Hz<f<8Hz
S10	188	1.4960	Basso Tirenno 15/04/19878	Naso	5.5	16	4Hz<f<8Hz
S11	17116	1.023	L'Aquila aftershock 09/04/2009	Montereale	4.9	10	8Hz<f<16Hz
S12	6802	4.260	South Iceland aftershock 21/06/2000	Thjorsartun	6.4	3	8Hz<f<16Hz
S13	16996	1.393	L'Aquila aftershock 07/04/2009	L'Aquila Via Aterna – Il Moror	4.2	2	8Hz<f<16Hz
S14	6789	0.803	South Iceland aftershock 21/06/2000	Hveragerdi-Church	6.4	23	8Hz<f<16Hz

Afin de faciliter le calcul, les signaux sont rendus à la même durée qui représente la plus grande durée 'T_{max}' parmi l'ensemble des signaux (Tableau 6). Les accélérations qui sont comprises entre leur durée du signal en question et 'T_{max}' sont transformées à une valeur nulle.

L'examen de la figure (Figure 3-3) montre que les 14 signaux ont des contenus fréquentiels couvrant toutes les gammes fréquences. Du moment que l'utilisation

de ces 14 spectres de réponses va donner 14 AF. De ce fait on a introduit un spectre de réponse défini comme étant la moyenne géométrique des 14 spectres de réponse. Cette valeur moyenne est nommée ‘Sm’.

L’amplification ‘AF’ est définie comme étant le rapport entre le spectre de réponse en surface et en profondeur pour un profil de sol. Pour ce faire on doit de définir comment estimer ce spectre de réponse en surface connaissant le spectre de réponse en profondeur ainsi que les caractéristiques du sol. Dans le présent travail on a eu recours à la théorie de propagation des ondes sismiques.

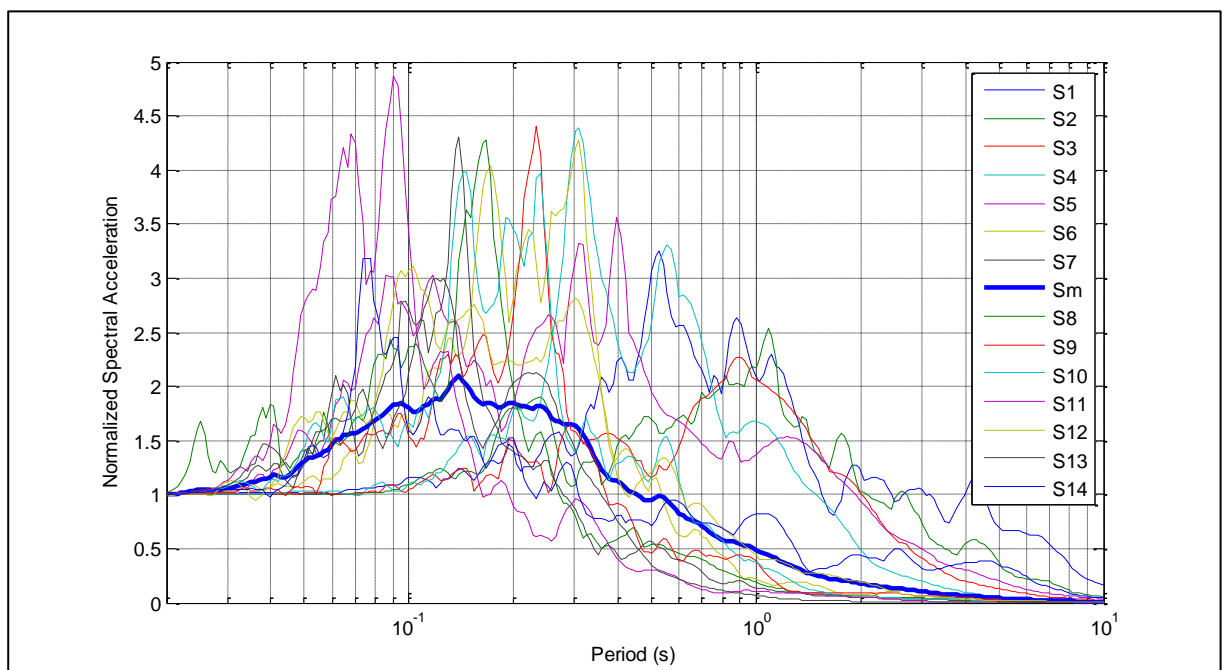


Figure 3-3. Pseudo spectre d’accélération normalisée (avec 5% d’amortissement) pour un ensemble de 14 accélérogrammes réels récolté en substratum (Sm : représente la moyenne géométrique des 14 pseudos spectre d’accélération normalisée).

3.2.3 Réponse sismique d’un profil de sol stratifié horizontalement

3.2.3.1 Introduction

Le problème consiste à trouver la réponse d’un profil de sol à une onde de volume, en utilisant la théorie de propagation d’ondes dans un milieu continu. C’est un problème tridimensionnel puisque l’onde se propage dans toutes les directions.

La source est considérée comme constituée d'une ligne (faille) et le site est supposé être placé à une distance suffisamment éloignée de la source, le problème devient donc bidimensionnel.

Dans cette analyse le profil du sol est modélisé en milieu continu. À cet effet, les conditions suivantes sont adoptées :

- En génie parasismique l'hypothèse admise est que le mouvement horizontal résulte de la propagation d'onde de cisaillement, et le mouvement vertical de la propagation de l'onde de compression.
- Chaque couche du profil est affectée d'un système d'axes (u, z_m) , étant également supposée homogène, isotrope.
- Le profil de sol est supposé infini dans le sens u .
- Pour la modélisation, le profil de sol est constitué de n couches horizontales d'épaisseur h_m , de masse volumique ρ_m , de vitesse de cisaillement complexe ν_{s_m} et de coefficient d'amortissement ξ_m .
- Le profil de sol est supposé soumis à la propagation verticale d'ondes de cisaillement, par conséquent, le vecteur déplacement en tout point présente une seule composante non nulle selon l'axe x : la propagation verticale d'onde de cisaillement ne peut induire qu'une déformation horizontale.

L'approche utilisée pour cette dernière est celle de 'Thompson-Haskell', la forme d'équation utilisée est décrite comme suit :

$$G \frac{\partial^2 u}{\partial z^2} + \eta \frac{\partial^3 u}{\partial z^2 \partial t} = \rho \frac{\partial^2 u}{\partial t^2} \quad (3-3)$$

G: Module de cisaillement de la couche considéré.

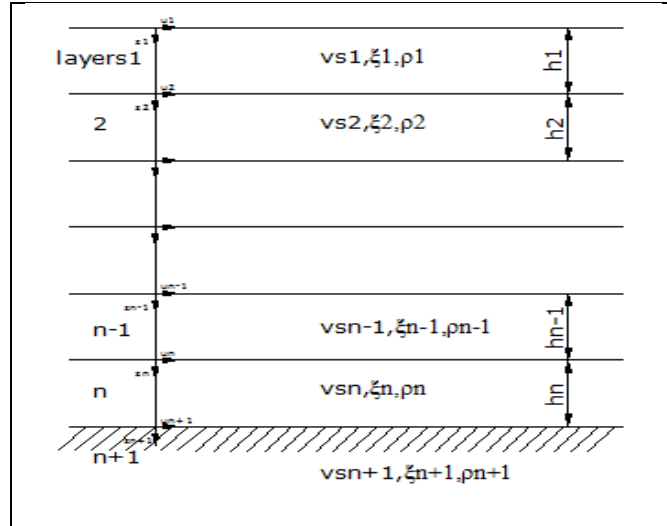


Figure 3-4 Système de couches de sol stratifié horizontalement soumis à la propagation d'onde S verticale.

La solution de l'équation d'onde (Equation(3-3)) unidimensionnelle, pour une fréquence d'onde unique (ω) fournit le déplacement (u) en fonction de la profondeur (z) et du temps (t) pour chaque couche de sol m est (Kramer, 1996) :

$$u_m(z) = A_m e^{(i\omega t)} e^{(ik_m^* z)} + B_m e^{(i\omega t)} e^{(-ik_m^* z)} \quad (3-4)$$

Où :

$$k^* = \frac{\omega}{vs^*} \quad (3-5)$$

k^* : nombre d'onde complexe.

$$vs^* = vs(1+i\xi) \quad (3-6)$$

vs : la vitesse de cisaillement de la couche considéré.

ξ : Facteur d'amortissement.

D'après Hashash et al (2004), le facteur d'amortissement minimum est défini comme suit :

$$\xi = \frac{1}{2Q} \quad (3-7)$$

Q : facteur de qualité qui décrit l'énergie de dissipation.

$$Q = \frac{vs}{10} \quad (3-8)$$

$$\text{D'où : } \xi = \frac{5}{vs} \quad (3-9)$$

A_m et B_m représentent les amplitudes de propagation d'onde vers le haut et vers le bas respectivement, on peut les déterminer à partir des conditions aux limites suivantes :

- Forme modale normalisée à l'unité en surface :

$$\lim_{z \rightarrow 0} u_1(z) = 1 \Rightarrow A_1 = 1 \quad (3-10)$$

- Contrainte de cisaillement nulle à la surface libre :

$$\lim_{z \rightarrow 0} G_1^* \frac{du_1}{dz} = 0 \Rightarrow A_1 = B_1 \quad (3-11)$$

- Continuité de la contrainte de cisaillement entre la couche 'm' et la couche 'm+1' :

$$\tau_m(z_m = h_m, t) = \tau_{m+1}(z_{m+1} = 0, t) \quad (3-12)$$

(m+1) c'est le numéro de la couche qui se trouve au-dessous la couche (m) (comme présenter sur la figure (Figure 3-4)). On sait que :

$$\tau(z, t) = G^* \frac{\partial u}{\partial z} = iG^* k^* (Ae^{(i\omega t)} e^{(ik^*z)} - Be^{(i\omega t)} e^{(-ik^*z)}) \quad (3-13)$$

D'où

$$A_{m+1} - B_{m+1} = \frac{k_m^* G_m^*}{k_{m+1}^* G_{m+1}^*} (A_m e^{(i\omega t)} e^{(ik_m^*z)} - B_m e^{(i\omega t)} e^{(-ik_m^*z)}) \quad (3-14)$$

On peut définir un autre paramètre α qu'est le rapport d'impédance entre la couche 'm' et 'm+1' :

$$\alpha_m = \frac{k_m^* G_m^*}{k_{m+1}^* G_{m+1}^*} \quad (3-15)$$

D'où l'équation devient $A_{m+1} - B_{m+1} = \alpha_m (A_m e^{(i\omega t)} e^{(ik_m^*z)} - B_m e^{(i\omega t)} e^{(-ik_m^*z)})$

- Continuité du déplacement entre la couche 'm' et 'm+1'

$$u_m(z_m = h_m, t) = u_{m+1}(z_{m+1} = 0, t) \quad (3-16)$$

$$A_{m+1} + B_{m+1} = (A_m e^{i\omega t} e^{i\kappa_m^* z} + B_m e^{i\omega t} e^{-i\kappa_m^* z}) \quad (3-17)$$

Les amplitudes A_{m+1} et B_{m+1} en fonction de A_m et B_m sont déterminés à partir des équations précédentes. D'où on peut écrire :

$$A_{m+1} = \frac{1}{2} A_m (1 + \alpha_m) e^{i\kappa_m^* h_m} + \frac{1}{2} B_m (1 - \alpha_m) e^{-i\kappa_m^* h_m} \quad (3-18)$$

$$B_{m+1} = \frac{1}{2} A_m (1 - \alpha_m) e^{i\kappa_m^* h_m} + \frac{1}{2} B_m (1 + \alpha_m) e^{-i\kappa_m^* h_m} \quad (3-19)$$

3.2.3.2 Différentes étapes de calcul de la réponse du sol

L'amplitude à la surface libre d'un sol peut être définie comme le produit entre l'amplitude d'entrée et la fonction de transfert qui relie la surface et la couche qui comporte le mouvement d'entrée. Pour cela on a développé un programme de calcul de la réponse d'un profil du sol sous un chargement sismique, qui peut être obtenu par les étapes suivantes (Kramer, 1996 ; Yezli-Khati, 2003) :

1. Exprimer le mouvement d'entrée dans le domaine fréquentiel en utilisant la transformée de Fourier rapide (FFT). Cette transformé de Fourier aura une partie réelle et une partie imaginaire.

La transformée de Fourier d'une fonction continue dans le temps $f(t)$ est donnée par :

$$F(\omega) = \int_{-\infty}^{+\infty} f(t) e^{-i\omega t} dt \quad (3-20)$$

qui peut également s'écrire sous une forme complexe :

$$F(\omega) = E(\omega) e^{-i\phi(\omega)t} = C(\omega) - iS(\omega) \quad (3-21)$$

Avec :

$$E(\omega) = \sqrt{C(\omega)^2 + S(\omega)^2} \quad (3-22)$$

Où $E(\omega)$ représente le spectre de l'amplitude Fourier.

2. Définir la fonction de transfert entre la couche d'entrée et la couche de sortie, qui aura également une partie réelle et une partie imaginaire.

3. Faire le produit entre le spectre de Fourier au niveau du mouvement d'entrée est la fonction de transfert reliant la couche d'entrée et de sortie, ce produit donne le spectre de Fourier à la surface libre, qui aura également une partie réelle et une partie imaginaire, soit :

$$E_1(\omega) = E_{n+1}(\omega) \times |T_{1\ n+1}(\omega)| \tag{3-23}$$

$E_1(\omega)$: Le spectre de Fourier à la surface libre.

$E_{n+1}(\omega)$: le spectre de Fourier au niveau du substratum rocheux

$|T_{1\ n+1}(\omega)|$: la fonction de transfert entre l'entrée et la surface.

4. Une fois, le spectre de Fourier à la surface libre étant connu. Il est possible de déterminer l'accélérogramme à la surface libre par la transformée de Fourier inverse.

D'autre part, la fonction de transfert T_{m+1} est le rapport de l'amplitude de déplacement au toit de la couche m à l'amplitude de déplacement au toit de la couche $m+1$:

$$T_{m+1}(\omega) = \frac{A_m + B_m}{A_{m+1} + B_{m+1}} \tag{3-24}$$

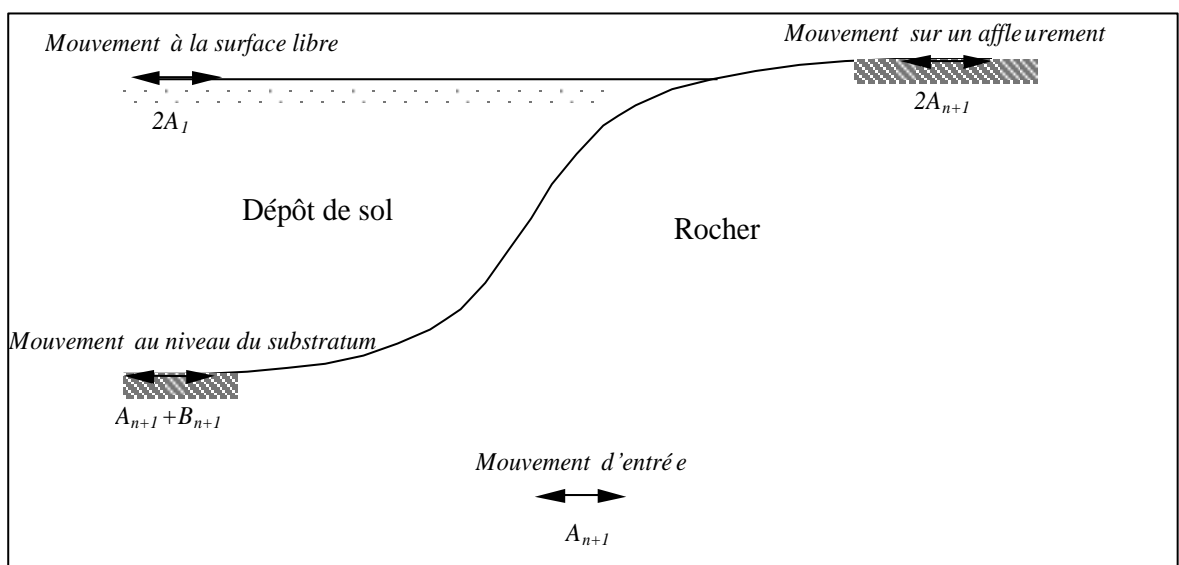


Figure 3-5 Terminologie utilisée dans l'analyse de la réponse de site .

Nous avons confronté le programme de calcul que nous avons développé avec DEEP-soil (www.illinois.edu) et EERA (geoinfo.usc.edu/gees) en utilisant le même profil de sol (voir ci-dessous). Le séisme de Kobe1995 a été choisi dans les calculs pour montrer que les résultats obtenus par notre propre code et celui de Deepsoil. Les figures (Figure 3-6 à Figure 3-9) illustrent le cas étudié.

La nécessité de développer ce programme résulte du fait que l'introduction des données pour les autres logiciels (Deepsoil ou EERA) se fait profil par profil alors que dans notre cas on sera amené à faire des milliers de simulations d'un seul cout.

Tableau 7 exemple du profil de sol pris dans le calcul.

	Epaisseur(m)	Vs(m/s)	ρ (kN/m ³)	L'amortissement
Couche1	30	200	20	2.5%
substratum	0	800	20	0.625%

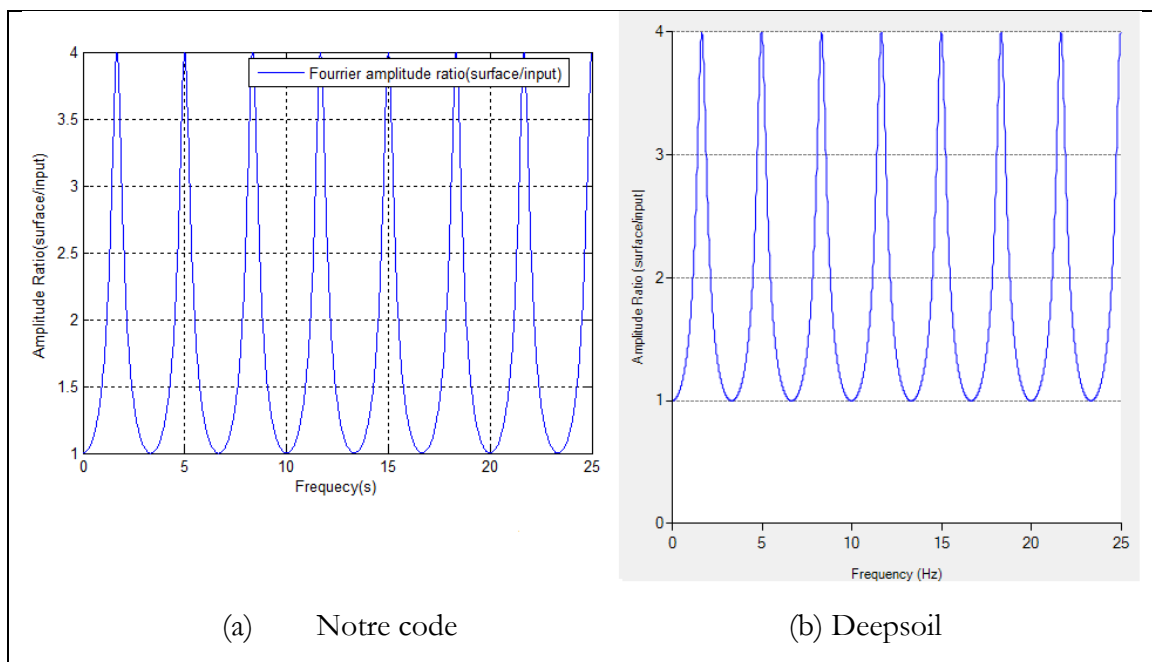


Figure 3-6 Facteur d'amplification fourrier calculer parle modèle linéaire élastiques, en comparaison notre programme avec Deepsoil.

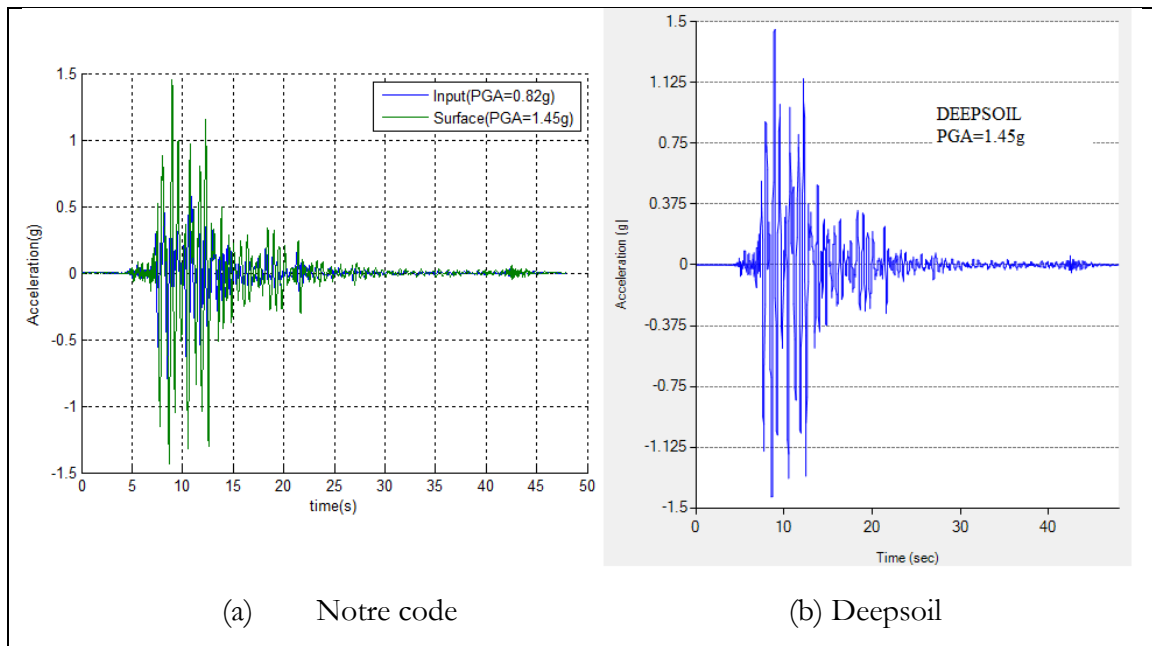


Figure 3-7 Accélération en surface par le modèle linéaire élastiques, en comparaison notre programme avec Deepsoil.

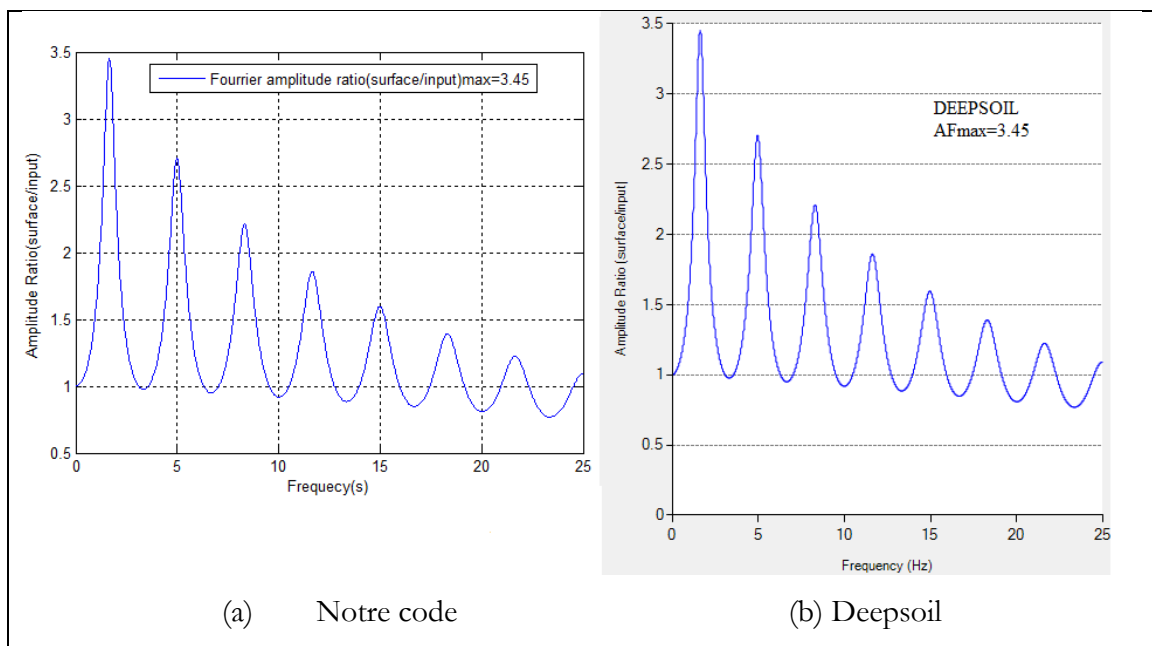


Figure 3-8. Facteur d'amplification fourrier calculer parle modèle linéaire viscoélastique ou on compare notre programme avec Deepsoil.

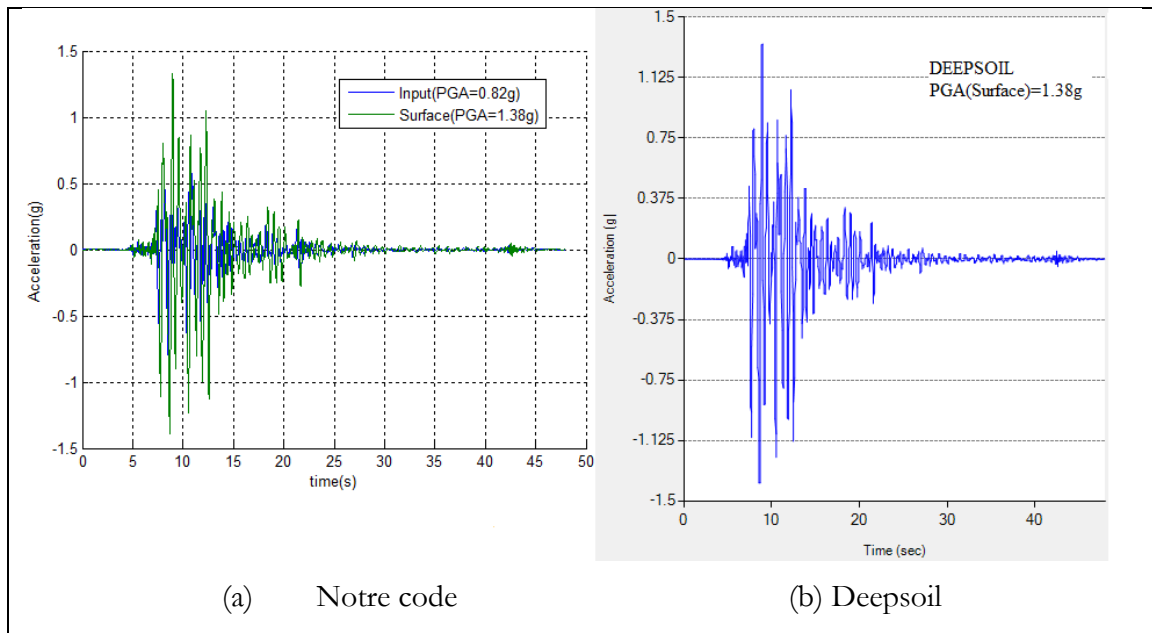


Figure 3-9 Accélération en surface par le modèle linéaire viscoélastique, en comparaison notre programme avec Deepsoil.

3.2.4 Caractérisation des profils de sol

Dans la présente étude on a repris la compilation de 858 profils de sol mises en place par C. Cornou (Salameh 2016). Cet ensemble de profil de sol est composé d'environ 600 sites KiK-net japonais, plus de 200 sites des États-Unis, mis à disposition par D. Boore (<http://quake.usgs.gov/~boore>) et 22 Sites européens mesurés au cours du projet NERIES (Di Giulio et al., 2012). Toutefois, on note que la méthodologie mise en place peut être appliquée pour d'autres bases de données de profils de sol.

Les principales caractéristiques de l'ensemble des profils résultants sont présentées dans Salameh et al. (2016) et Almakari et al. (2016).

Ces profils de sol sont dénommées ci-après Profils de sol initiaux ou Real Profile (RP). On peut caractériser chaque profil par une série de paramètres de site qui est en quelque sorte une carte d'identité de chaque profil. Ces paramètres sont au nombre de six (06) : profondeur (Depth), la vitesse de cisaillement moyenne sur le profils (V_{sm}), la vitesse de cisaillement moyenne des ondes sur 30m supérieure (V_{s30}), la vitesse de cisaillement du substratum ($V_{bedrock}$), le contraste de vitesse

(Cv) et de la fréquence fondamentale du profil de sol. Elles sont définies comme suit:

$$Depth = \sum_{i=1}^n h_i \quad (3-25)$$

$$V_{sm} = \sum_{i=1}^n h_i / \sum_{i=1}^n \frac{h_i}{V_i} \quad (3-26)$$

Avec $V_m = \sqrt{G_m / \rho_m}$ est la vitesse de cisaillement de la couche (i)

$$V_{s30} = 30 / \sum_{i=1}^{l30} \frac{h_i}{V_i} \quad (3-27)$$

Ou l30 :est le numéro de la couche ou ne dépasse pas 30m de profondeur.

$$Cv = \frac{V_{bedrock}}{V_1} \quad (3-28)$$

Dans cette étude, la fréquence fondamentale (f_0) du profil du sol est déterminée par la méthode Rayleigh simplifiée (Dobry et al, 1976), (Equations ci-dessous).

$$x_{i-1} = x_i + \frac{z_i + z_{i-1}}{v_i^2} H_i \quad (3-29)$$

avec $x_n = 0$.

$$w_5^2 = \frac{4 \sum_{i=1}^{i=n} \frac{(z_i + z_{i-1})^2}{v_i^2} H_i}{\sum_{i=1}^{i=n} (x_i + x_{i-1})^2 H_i} \quad (3-30)$$

$$T_0 \approx T_5 = \frac{2\Pi}{w_5} \quad (3-31)$$

$$f_0 = \frac{1}{T_0} \quad (3-32)$$

Les distributions du logarithme de ces paramètres suivent une distribution normale. De plus, les coefficients de corrélation (R^2) des paramètres ont été calculées (voir *Tableau 8*).

Plusieurs conclusions peuvent être tirées de l'examen de la figure (*Figure 3-10*) et le tableau correspondant.

- Les distributions suivent une loi log normale.
- Les paramètres de site sont parfois bien corrélés entre eux (exemple V_{sm} et V_{s30}) et pour d'autres ils ne le sont pas (C_v et V_{s30})
- La vitesse au niveau du substratum peut dépasser la valeur de 2000m/s.

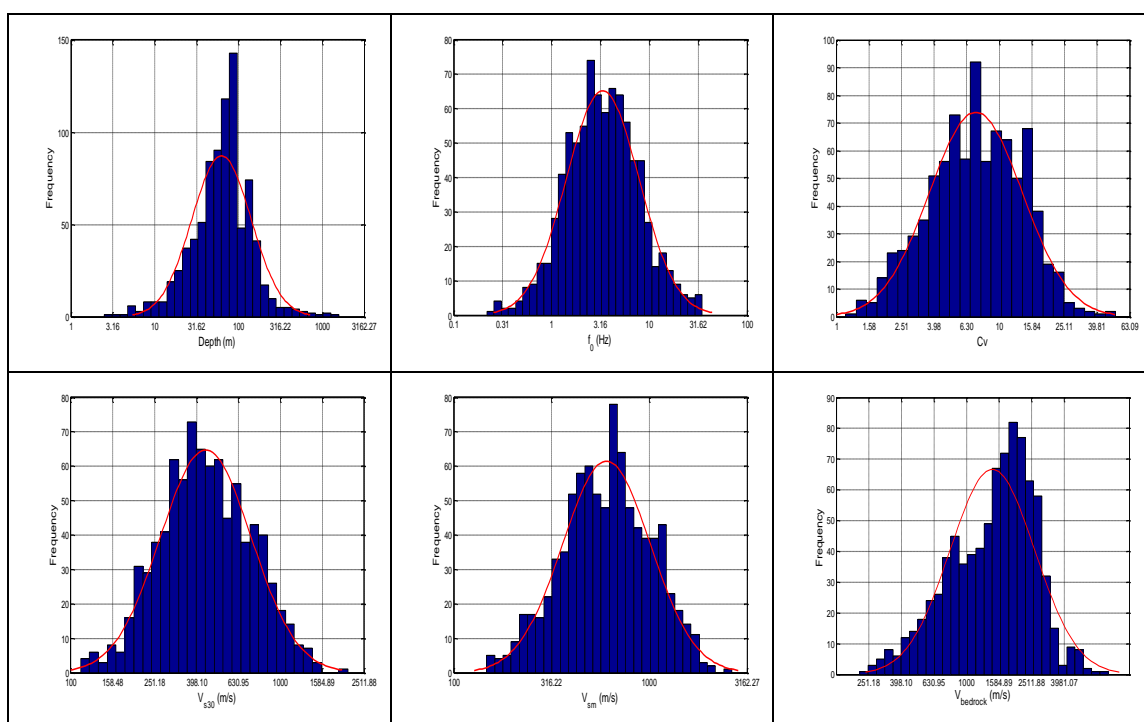


Figure 3-10 : Distribution en histogramme des paramètres de la base de donnée à savoir (Depth , f_0 , C_v , V_{sm} , V_{s30} et $V_{bedrock}$) pour tous les profils (cas profils RP), la courbe rouge représente la distribution normale théorique.

Tableau 8 Facteur de corrélation(R^2) entre les paramètres de site.

	Depth	f_0	C_v	V_{sm}	V_{s30}	$V_{bedrock}$
Depth	1	0.3441	0.0903	0.2434	0.0388	0.0962
f_0		1	0.0196	0.1858	0.4050	0.3102
C_v			1	0.1596	0.0559	0.5345
V_{sm}				1	0.8515	0.6784
V_{s30}					1	0.5531
$V_{bedrock}$						1

Du moment que les vitesses de cisaillement au niveau du substratum atteignent des valeurs importantes et du fait que la majorité des codes sismiques considèrent que $V_{bedrock}$ est égale à 800 m/s, nous avons constitué deux autres nouvelles bases de données de profils de sol issues de la base de données initiale, en premier lieu avec des profils normalisés en profondeur ou pour chaque profil de sol normalisé parmi les 858 profils, est modifié comme suit :

$$V'_i = V_i \cdot 800 / V_{bedrock} \quad (3-33)$$

$$h'_i = h_i \cdot 800 / V_{bedrock} \quad (3-34)$$

$$Depth = \left(\frac{800}{V_{bedrock}} \right) * Depth \quad (3-35)$$

$$V_{sm} = \left(\frac{800}{V_{bedrock}} \right) * V_{sm} \quad (3-36)$$

$$V_{s30} = 30 / \sum_{j=1}^{l'_{30}} \frac{h'_j}{V'_j} \quad (3-37)$$

Où l'_{30} est le numéro de la couche, ou ne dépasse pas 30m de profondeur modifiée. Il est nécessaire de dire que la vitesse du bedrock est égale à 800m/s, le C_v et f_0 restent constants. Les nouveaux profils de sol modifiés sont nommés Profil normalisé en substratum ou Normalized Profile (NP). Pour respecter la réalité, on a éliminé tous les profils qui ont une vitesse de cisaillement de l'une des couches

inferieur à 80m/s. cela induit un changement du nombre de profils à 570 aux lieux de 858.

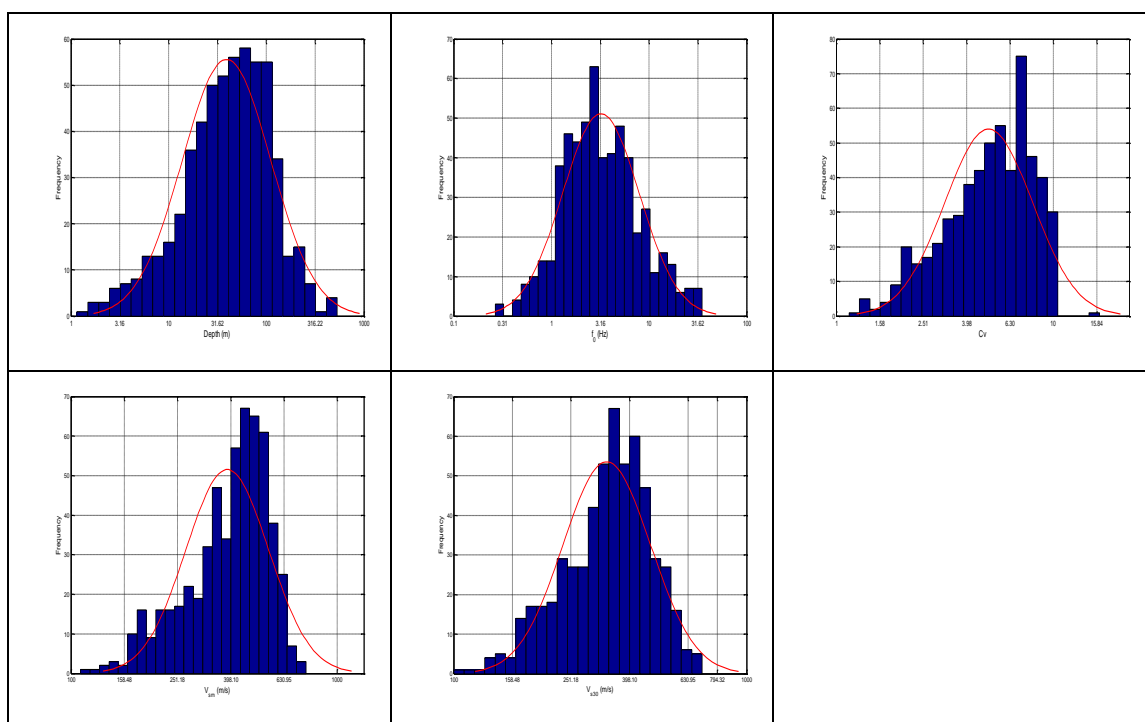


Figure 3-11 : Distribution en histogramme des paramètres de la base de donnée à savoir (Depth , f_0 , C_v , V_{sm} et V_{s30}) pour des profils normalisés en substratum (NP), la courbe rouge représente la distribution normale théorique.

Tableau 9 facteur de corrélation entre les paramètres normalisés en substratum avec 800m/s.

Correlation factor (R^2)	Depth	f_0	C_v	V_{sm}	V_{s30}
Depth	1	0.4814	0.1935	0.4245	0.0713
f_0		1	0.0196	0.2542	0.0356
C_v			1	0.5124	0.4900
V_{sm}				1	0.8744
V_{s30}					1

En deuxième lieux une autre philosophie est adoptée qui consiste à prendre des profils appelé tronqués ou Turnicated Profile (TP). Dans ce contexte, chaque profil de sol est maintenu inchangé jusqu'à la profondeur Z_{800} où la vitesse de cisaillement de la couche considéré dépasse 800 m/s et au-delà, la vitesse est fixée à 800 m/s, chaque fois que la vitesse du substratum rocheux réel est inférieure à 800 m/s, la vitesse du substratum rocheux est

portée à 800 m/s. Cependant les mêmes formules utilisées pour les profils réelles restent inchangées, leur distribution se présente comme suit :

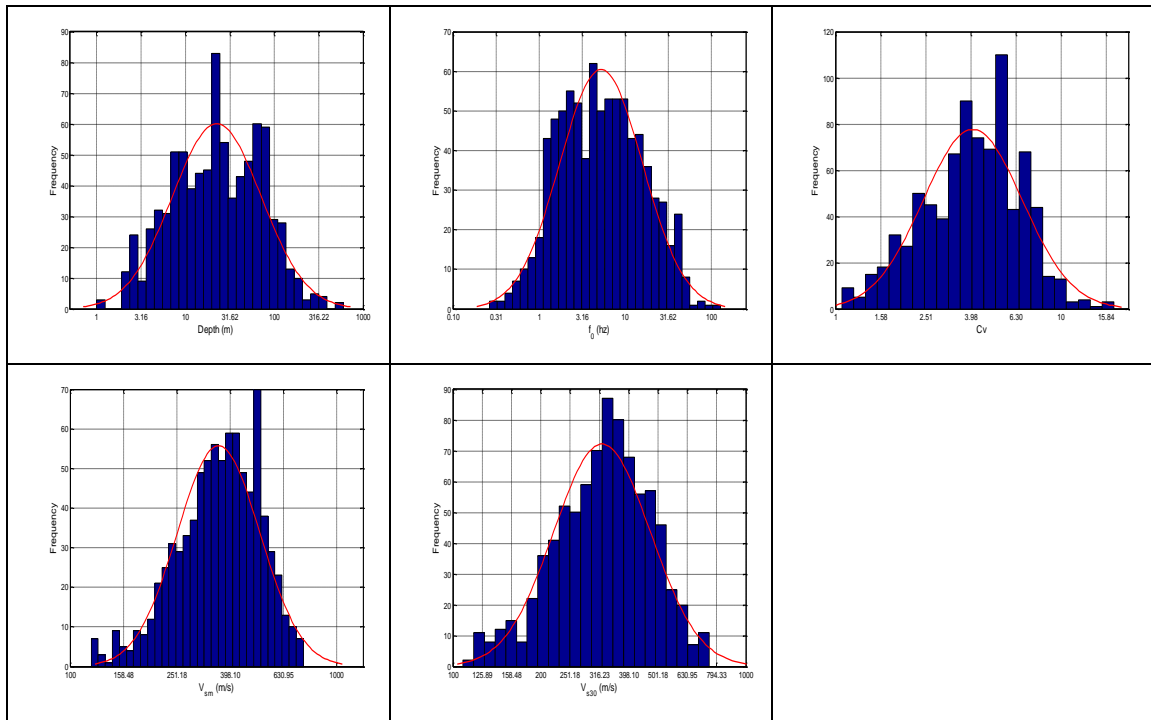


Figure 3-12 : Distribution en histogramme des paramètres de la base de donnée à savoir (Depth , f_0 , Cv , V_{sm} et V_{s30}) pour des profils tronqués (TP), la courbe rouge représente la distribution normale théorique.

Tableau 10 facteur de corrélation entre les paramètres des profils tronqués (TP)

Correlation factor (R^2)	Depth	f_0	Cv	V_{sm}	V_{s30}
Depth	1	0.4176	0.1973	0.2430	0.0664
f_0		1	0.3864	0.0629	0.2529
Cv			1	0.3852	0.5436
V_{sm}				1	0.8811
V_{s30}					1

Les tableaux (Tableau 8, Tableau 9 et Tableau 10) indiquent la variation des coefficients de détermination (R^2) entre chaque paire de paramètres d'entrés pour les trois types de profils (RP, NP et TP). Il y a clairement une tendance générale à une certaine corrélation entre les paramètres de vitesse, en particulier V_{sm} et V_{s30} , $V_{bedrock}$ et V_{sm} , V_{s30} et Cv. En outre les tableaux indiquent que les corrélations sont plus faibles parfois autour de 0 pour les autres cas. A ce titre on note une faible corrélation pour les paires (Cv, f_0), (Depth, V_{s30}), (Depth, Cv) et (Depth, $V_{Bedrock}$).

Ces indicateurs de corrélation sont très utiles pour sélectionner les paramètres de site qui sont indépendants afin de déterminer les modèles reliant l'amplification du site aux caractéristiques local du site.

En outre la fonction de distribution cumulative (cdf) est fourni en annexe pour les différents paramètres d'entrés (Depth, $f_0, V_{sm}, V_{s30}, C_v, V_{bedrock}$) en fonction des différents types de profils de sol (RP, NP et TP).

3.3 Etablissement de la base de données

Le facteur d'amplification est calculé comme le rapport des spectres de réponse à la surface du sol et la roche de référence (substratum rocheux).

$$AF(T) = \frac{SA(T)_s}{SA(T)_b} \quad (3-38)$$

Cela est calculé pour les trois ensembles de base de données (RP, NP et TP) soumis à 14 excitations sismiques. L'ensemble résultant de facteurs d'amplification est donc constitué de $(858+570+855) \times 14 = 31962$ amplification 'AF'. Le facteur d'amplification est noté $AF(P_k, \theta, S_l, T_i)$, avec :

- P_k , $k = 1, \dots, 858$ pour identifier le profil de sol.
- $\theta = 1$ pour les profils RP, $\theta = 0$ pour les profils NP et $\theta = 2$ pour les profils TP.
- S_l , $l = 1, 14$ indique l'excitation (voir Section 3.2.2), $S_m : l \neq m$ indique la moyenne géométrique des 14 spectre de réponse pour chaque profil.
- T_i , ($i = 1, \dots, 271$) est la $i^{\text{ème}}$ période. En fait, les facteurs d'amplification ont été systématiquement calculés pour 271 valeurs, également espacées entre 0,01 s et 10 s sur un axe logarithmique équidistant (c'est-à-dire également espacées de façon égale entre 0,1 Hz et 100 Hz sur un axe de fréquence).

Une fois le AF calculée pour un profil particulier k, et pour 14 excitations sismiques, le facteur d'amplification moyen du site noté $AF_m(P_k, \theta, T_i)$ ou par abréviation AF_m est calculé comme la moyenne géométrique des 14 facteurs d'amplification individuels.

$$\log[AF_m(P_k, \theta, T_i)] = \left(\frac{1}{14}\right) \sum_{l=1}^{14} \log[AF(P_k, \theta, S_l, T_i)] \quad (3-39)$$

L'écart-type correspondant a également été estimé

$$\sigma(P_k, \theta, T_i) = \sqrt{\frac{\sum_{l=1}^{14} [\log(AF(P_k, \theta, S_l, T_i)) - \log(AF_m(P_k, \theta, T_i))]^2}{14}} \quad (3-40)$$

Par la suite, la notation abrégée AF correspondra en fait à la moyenne géométrique 'AF_m(P_k, θ, T_i)'.

Entre autre des paramètres supplémentaires sont introduits pour mesurer la variabilité des résultats.

- La moyenne des AF sur tous les profils est notée A₀(θ, T_i) et elle est définie par la moyenne géométrique des 858 profils

$$\log(A_0(\theta, T_i)) = \frac{1}{n_p} \sum_{k=1}^{n_p} \log[AF_m(P_k, \theta, T_i)] \quad (3-41)$$

Avec : n_p est le nombre total de profils égal à 858 pour 'RP', 570 pour 'NP' et 855 pour 'TP'.

La variabilité initiale, définie comme l'écart-type du facteur d'amplification moyen sur tous les profils.

$$\sigma_0(\theta, T_i) = \sqrt{\frac{1}{n_p} \sum_{k=1}^{n_p} [\log(AF_m(P_k, \theta, T_i)) - \log(A_0(\theta, T_i))]^2} \quad (3-42)$$

- La variabilité initiale maximum est définie comme suite :

$$\sigma_{0Max}(\theta) = \text{Max}_{T_i} [\sigma_0(\theta, T_i)] \quad (3-43)$$

- La variabilité initiale sur toute la période est

$$\sigma_{t0}(\theta) = \frac{1}{n_T} \sum_{i=1}^{n_T} \sigma_0(\theta, T_i) \quad (3-44)$$

Ou n_T est le nombre de périodes utilisé égal à 271.

3.4 Méthode d'estimation de l'AF par les réseaux de neurones

L'objectif principal de la présente thèse, trouve son fondement dans la recherche d'une relation entre le facteur d'amplification et un certain nombre de paramètres de site. Alternativement, cet objectif en appelle un autre quel est le nombre minimum de paramètres de site pour estimer "correctement" un facteur d'amplification. Pour ce faire on dispose maintenant d'une base de données constitué des éléments suivants :

- 03 bases de données pour le profil de sol : Profil de sol initial (RP), Profil normalisé en substratum (NP) et profils tronqué (TP) avec pour chacun ses paramètres de site soit $2 \times (858 + 570 + 855) = 4566$ valeurs pour le profil de sol initial.
- Pour chaque profil de sol, on a le AF c'est-à-dire $AF_m(P_k, \theta, T_i)$ soit au total $2283 \times 271 = 618693$ valeurs de AF.

Les très grandes quantités de valeurs de AF nous ont amené à utiliser les techniques d'intelligence artificielle. La figure (Figure 3-13) illustre parfaitement cette situation où in-fine on aura à introduire les paramètres de site dans un réseau RNA pour obtenir le AF (appelé aussi Spectral amplification factor). On aura donc d'abord à constituer ce réseau de neurones avec comme données pour chaque profil de sol le AF théorique (Equation (3-39)) et les valeurs des paramètres de site, la sortie est le AF généré. Il faut aussi s'assurer que la variabilité des valeurs dans la base de données ne soit pas inférieure à celle qui sera observée dans les valeurs générés par le réseau de neurones.

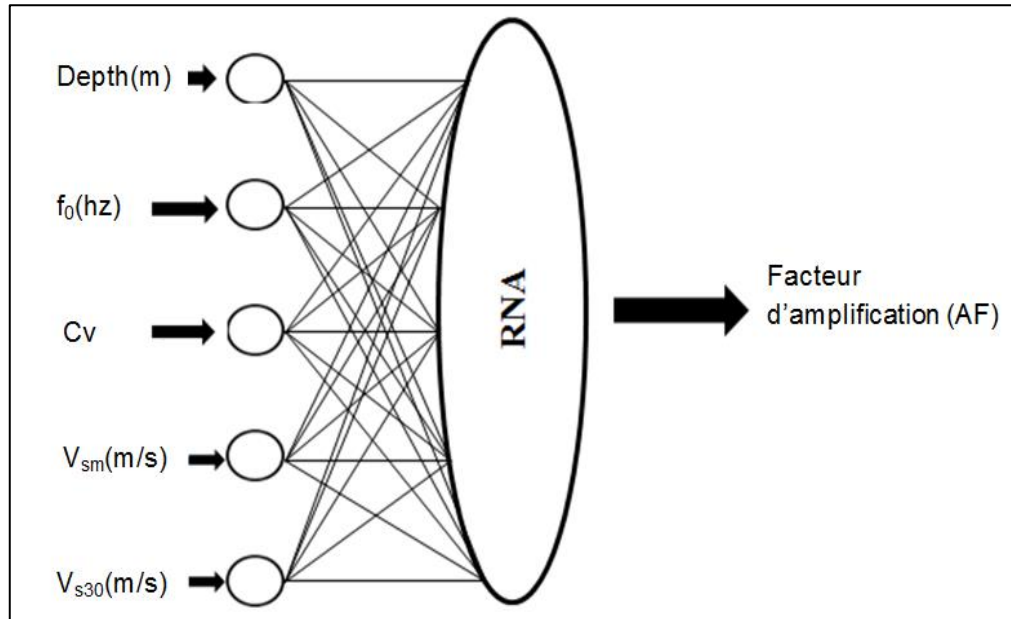


Figure 3-13 : Définition des paramètres pris par les réseaux neurones en utilisant 05 paramètres d'entrés pour les profils normalisés (NP) et tronqués(TP).

Il est nécessaire de rappeler que pour les profils réels, on a la vitesse de substratum ($V_{bedrock}$) en plus des 5 paramètres définis dans la figure (Figure 3-13).

Dans la stratégie d'estimation de l'AF on utilise le AF tel que défini par l'Equation (3-39) c'est-à-dire en utilisant le terme $AF_m(P_k, \theta, T_i)$. Cependant on peut utiliser le terme $AF_m\left(P_k, \theta, \frac{f_i}{f_0}\right)$ c'est à dire que l'on va procéder à une description de l'AF en fonction de la fréquence au lieu de la période. On a pensé à normaliser la fréquence par la fréquence de résonance de chaque profils.

3.5 Conclusion

Le présent chapitre nous a permis de mettre en place la méthode ainsi que ses inputs. Il s'agit d'utiliser la méthode des réseaux de neurones pour déduire des facteurs d'amplification à partir d'une base de données composée de facteur d'amplification tirée de l'analyse 1D viscoélastique linéaire et de paramètre de site. Le chapitre suivant est justement consacré à la présente de la méthode des réseaux de neurones et de l'implémentation de la méthodologie décrite plus haut.

Chapitre 4

Description de l'approche neuronale

4.1 Introduction

Ce chapitre a pour objectif la description de la méthode de réseaux de neurones artificiels. On présentera en premier lieu le cadre général et les avantages de cette méthode qui dans son fondement est capable d'effectuer des opérations semblables à celle du cerveau humain. Ensuite on décrira la méthode de réseaux de neurones de régression généralisée (GRNN). De par ses avantages, le GRNN sera utilisé pour estimer le facteur d'amplification spectral et déduire la variabilité aléatoire associé à chaque proxy de site. Enfin l'implémentation du GRNN sous « Matlab » sera mise en évidence.

4.2 Inspiration du neurone formel

Les cellules nerveuses de l'être humain, souvent appelées neurones sont des éléments de base du système nerveux central qui se compose de mille milliards neurones. Ces dernières sont des cellules composées de trois parties :

1-Un corps cellulaire qui contient le noyau et se charge d'effectuer les transformations biochimiques essentielles à la synthèse des éléments assurant la vie du neurone.

2-Les dendrites sont des ramifications du corps cellulaire, sa fonction principale permettent au neurone de capter les signaux qui lui provenant de l'extérieur.

3-L'axone généralement plus long que les dendrites, Il sert de moyen de transport pour les signaux émis par le neurone.

Les connexions entre neurones sont réalisées au niveau des synapses, lieu de proximité d'axone émetteur et dendrites réceptrices.

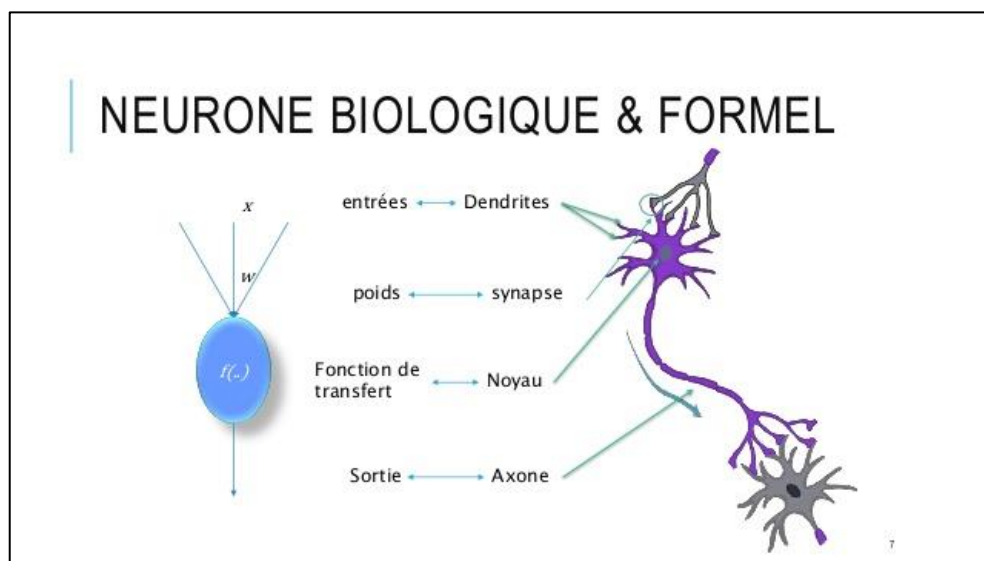


Figure 4-1: Un neurone biologique et ses principaux composants (Alani, 2008).

A partir des principes de fonctionnement du neurone biologique, le neurone formel (ou artificiel) est créé qu'est tout simplement une modélisation mathématique du neurone biologique (Figure 4-1).

Un neurone formel se caractérise par trois concepts : son état, ses connexions avec d'autres neurones et sa fonction d'activation. L'architecture d'un neurone artificiel est illustrée sur la figure (Figure 4-2).

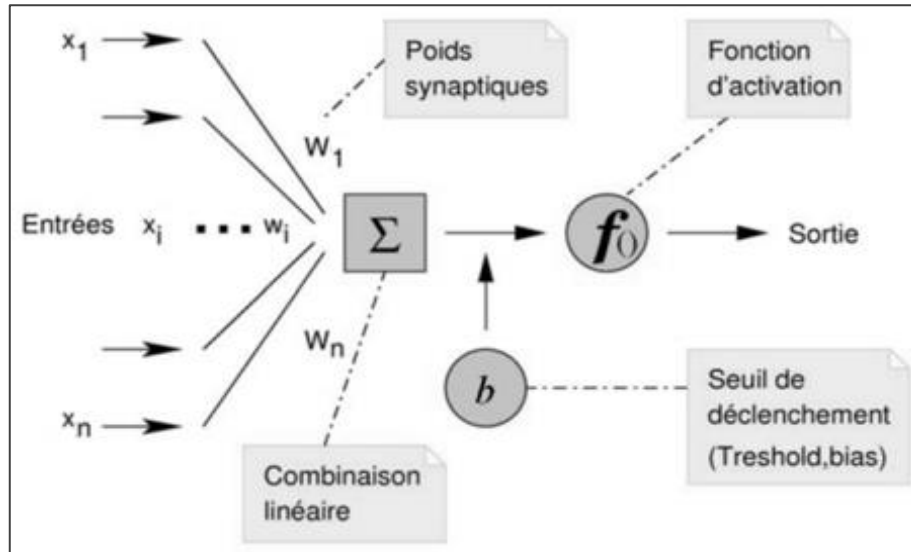


Figure 4-2: Structure générale du neurone formel (Derras et al, 2011).

4.3 Concept de base des réseaux de neurones artificielles

Les réseaux de neurones sont des opérateurs mathématiques, possédant des entrées qui peuvent être les sorties d'autres neurones, des entrées de signaux extérieurs ou la sortie du réseau. La valeur de la sortie d'un neurone résulte du calcul de la somme des entrées, pondérées par des coefficients (dits poids de connexions ou poids synaptiques) et du calcul d'une fonction non linéaire (dis fonction d'activation) de cette somme pondérée. Si cette somme dépasse un certain seuil (appelé le biais), le neurone est activé et transmet une réponse (sous forme de potentiel d'action) de manière unidirectionnelle dont la valeur est celle de son activation. Si le neurone n'est pas activé, il ne transmet rien.

La capacité de traitement des réseaux est stockée sous forme de poids d'interconnexions, qu'est obtenus soit par un processus d'apprentissage ou/et d'adaptation à partir d'un ensemble d'exemples d'apprentissage. L'interconnexion peut être totale ou partielle. Dans le premier cas, on parle d'un RNA totalement connecté (Figure 4-3) et dans le deuxième cas partiellement connecté.

Ces neurones formels sont connectés entre eux de diverses manières pour créer les réseaux de neurones artificiels.

L'architecture du réseau est décrite par l'ensemble de neurones interconnectés entre eux. Cette architecture peut varier d'une connectivité totale ou réseaux entièrement connectés (réseau à connexions complètes) à une connectivité locale ou réseaux à couches (réseau à sens unique), la figure ci-dessous illustre une architecture générale d'un réseau neurone.

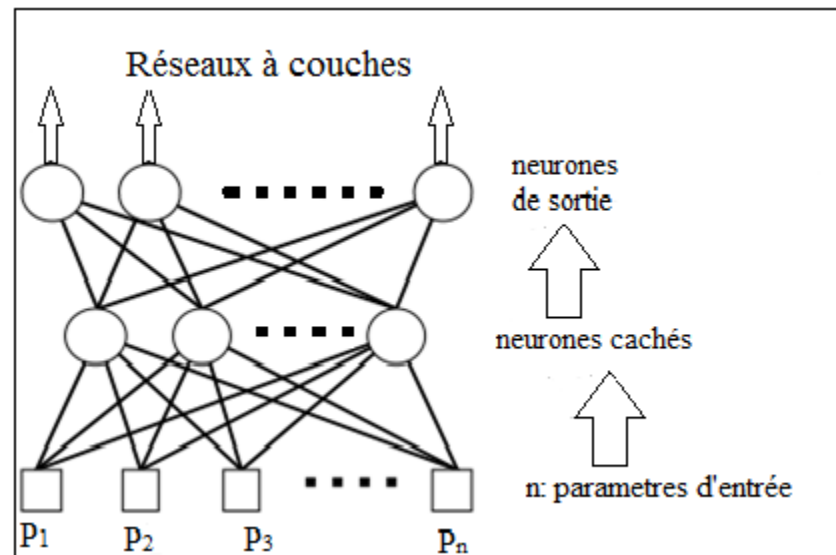


Figure 4-3: Architecture des réseaux neurones.

4.4 Types de réseaux neurones artificiel (RNA)

Les RNAs peuvent être classés en deux catégories selon leur dépendance avec le temps : les réseaux de neurones statiques (ou non bouclés) qui sont indépendants du temps et les réseaux de neurones dynamiques (ou bouclés) qui dépendent du temps.

4.4.1 Réseaux de neurones statiques (réseaux non bouclés)

Dans le cas des réseaux de neurones statiques, la sortie actuelle d'un neurone n'a aucune influence sur les sorties futures des autres neurones, Dans ce cas, le temps n'est pas significatif.

1. Perceptron multicouche (PMC)

Parmi les types les plus répondus des réseaux de neurones statiques on note le perceptron multicouche. Ce type de réseau est composé généralement de trois couches, le rôle de chaque couche est défini comme suite :

- La première couche est appelée aussi la couche d'entrée. Elle recevra les données source, la taille de cette couche est déterminée par le nombre de variables d'entrées.
- La deuxième couche est la couche cachée (voir les couches cachées). Les fonctions d'activation sont généralement non linéaires, Le choix de sa taille (nombre de neurones) n'est pas automatique et doit être ajusté. Il sera souvent préférable pour obtenir la taille optimale, de varier le nombre de neurones.
- La troisième couche est appelée couche de sortie. Elle donne le résultat obtenu après compilation par le réseau des données d'entrée dans la première couche. Sa taille est déterminée par le nombre de variables dont on a besoin en sortie.

2. Le réseau (Random Vector Functional Link Neural Network) (RVFLNN)

Le réseau (RVFLNN) (Réseau fonctionnel de réseau de vecteur aléatoire fonctionnel) est un modèle du perceptron développé par Pao et al.1989. Son architecture est similaire à celle du PMC avec des connexions supplémentaires entre la couche d'entrée et la couche de sortie. L'ajout de cette dernière montre que la connexion supplémentaire entre la couche d'entrée et de sortie joue un rôle important dans l'amélioration de la performance du RNA.

3. Réseau de type fonction à base radiale (RBF)

Dans la première couche du réseau, les distances entre le vecteur d'entrée et le centre de la fonction à base radiale sont à calculer ainsi que la largeur de la gaussienne. La somme du produit entre les fonctions RBF et les poids entre la couche cachée et la couche de sortie plus le biais associé à la couche de sortie représente le résultat du neurone de sortie.

La fonction à base radiale (exemple une gaussienne) est utilisée aussi, comment une fonction d'activation, dans les réseaux probabilistes (probabiliste neural networks (PNN)) et dans les réseaux à régression généralisée (Generalized régression neural network (GRNN)) et les autres réseaux.

4. Réseaux probabiliste (probabiliste neural networks (PNN))

Ce réseau peut être employé pour des problèmes de classification. Quand une entrée est présentée, la première couche calcule les distances entre le vecteur d'entrée et tous les vecteurs d'apprentissage et produit un vecteur dont les éléments indiquent comment ce vecteur d'entrée est proche de chaque vecteur d'apprentissage. La deuxième couche additionne ces contributions pour chaque classe des entrées pour produire à la sortie du réseau un vecteur des probabilités. Enfin, une fonction de transfert de concurrence sur la sortie de la deuxième couche sélectionne le maximum de ces probabilités, et produit un 1 pour cette classe et 0 pour les autres classes (Alani, 2008).

5. Réseaux à régression généralisée (generalized regression neural network (GRNN))

Généralement le GRNN est utilisé pour approximer une fonction continue quelconque où cette dernière est approchée par une série de fonction gaussienne construit en combinaison linéaires. Chacune de ces fonctions gaussiennes peuvent nous fournir un point de la fonction désiré.

4.4.2 Réseaux de neurones dynamiques ou réseaux bouclés (ou récurrents)

Les réseaux de neurones dynamiques appelés aussi réseaux récurrents ont une évolution dépendante du temps. La sortie d'un neurone du réseau peut donc être fonction d'elle-même ; cela n'est évidemment concevable que si la notion de temps est explicitement prise en considération. Ainsi, à chaque connexion d'un réseau de neurones bouclé est attaché, outre un poids comme pour les réseaux non bouclés, un retard, multiple entier (éventuellement nul) de l'unité de temps choisie. Une grandeur, à un instant donné, ne pouvant pas être fonction de sa propre valeur au

même instant, tout cycle du graphe du réseau doit avoir un retard non nul (Dreyfus, 2008).

Il est à noter que, dans la présente thèse nous utilisons un réseau de neurones statique, du fait que la fonction à prédire (facteur d'amplification de site) ne dépend pas du temps.

4.5 Types d'apprentissage

Le but d'apprentissage est de se rapprocher le plus possible de l'optimum, en ajustant les paramètres du modèle tel que les poids synaptiques. Dans la littérature, nous pouvons trouver plusieurs algorithmes d'apprentissage ; les plus connues sont la règle de Hebb (Parizeau, 2004), la règle de Windrow-Holff (1960), et la règle de Delta généralisé appelée également Rétro propagation (Parizeau, 2004).

L'apprentissage peut être supervisé ou non supervisé. Cette distinction repose sur la forme des exemples d'apprentissages, couples entrée/sortie associés à des réseaux multicouches dans le cas de l'apprentissage supervisé, alors que l'on ne dispose que des valeurs d'entrée pour l'apprentissage non supervisé.

On trouve l'apprentissage supervisé, entre autres, dans le perceptron où le superviseur fournit au réseau des couples d'entrée/sortie, il fait apprendre au réseau l'ensemble de ces couples, par une méthode d'apprentissage en comparant pour chacun d'entre eux la sortie effective du réseau et la sortie désirée, L'apprentissage est terminé lorsque tous les couples entrée/sortie sont jugés reconnus par le réseau.

Le type de l'apprentissage non supervisé consiste à détecter automatiquement des régularités qui figurent dans les exemples présentés et à modifier les poids des connexions pour que les exemples ayant les mêmes caractéristiques de régularité provoquent la même sortie.

Un autre type d'apprentissage commence d'être utilisé se décrit par l'apprentissage semi supervisé. Le principale avantage de cette méthode c'est qu'elle vise les problèmes avec peu de données étiquetées et une grande quantité de

données non étiquetées où ce type de situation est assez fréquent quand la base est assez pauvre en donnée ou coûteuse .

Il est nécessaire de signaler que le type d'apprentissage le plus utilisé c'est l'apprentissage supervisé car la plupart des algorithmes d'apprentissage des réseaux de neurones sont des algorithmes d'optimisation dans ce cas la méthode d'optimisation la plus célèbre et la plus utilisée en analyse neuronale est la rétro-propagation du gradient de l'erreur.

4.5.1 Rétro- propagation du gradient (RPG)

La RPG a été créé en généralisant les règles d'apprentissage de Widrow-Hoff (1960). Le but est d'ajuster les poids synaptiques et les biais pour lesquels la fonction de coût « E » (méthode qui calcul des erreurs entre la prédiction et l'initial), calculée sur les points de l'ensemble d'apprentissage, soit minimale.

A chaque couple entrée/sortie, une erreur est calculée, ainsi le gradient de l'erreur est déterminé. Ce gradient est la dérivée partielle de l'erreur par rapport au poids synaptique ($\delta E/\delta W$). Par la RPG, nous pouvons optimiser le modèle neuronal en modifiant les poids et les biais. Deux familles se présentent pour la mise à jour des paramètres.

- 1) Une première famille dite offline traite le choix des paramètres avant le démarrage de l'algorithme. Généralement ce choix se fait manuellement, et dépend de la habileté de l'auteur à choisir le bon jeu de paramètres par rapport au problème traité.
- 2) La deuxième famille de pré-réglages est dite online, la configuration choisie change au fur et à mesure que la recherche algorithmique continue.

4.5.2 Fonction de transfert (ou d'activation)

Les formes et les formules des fonctions de transferts les plus usuelles sont représentées sur le tableau suivant :

Tableau 11 Type de fonction de transfert les plus utilisées.

Nom de la fonction	Relation d'entrée/sortie	Icone	Nom Matlab
Seuil	$a=0$ si $n<0$ $a=1$ si $n\geq 0$		hardlim
Seuil symétrique	$a=-1$ si $n<0$ $a=1$ si $n\geq 0$		hardlims
linéaire	$a=n$		purelin
Linéaire saturée	$a=0$ si $n<0$ $a=n$ si $0\leq n\leq 1$ $a=1$ si $n>1$		satlin
Linéaire saturée symétrique	$a=-1$ si $n<-1$ $a=n$ si $-1\leq n\leq 1$ $a=1$ si $n>1$		satlins
Linéaire positive	$a=0$ si $n<0$ $a=n$ si $n\geq 0$		poslin
sigmoïde	$a = \frac{1}{1 + e^{-n}}$		logsig
Tangente hyperbolique	$a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$		tansig
compétitive	$a=0$ autrement $a=1$ si n maximum		compet
gaussienne	$a = e^{-n^2}$		radbas

4.5.3 Avantages et inconvénients des différents RNAs statiques et supervisés utilisés pour l'approximation des fonctions

Le modèle neuronal est un approximateur universel. Cybenko (1989) et Funahashi (1989). Il a la capacité de capturer les phénomènes physiques sous-jacents (Derras et al, 2012). Cependant, il est préférable, avant chaque étude, utilisant l'approche neuronale comme outil de modélisation, de bien choisir le type de RNA le mieux adapté à la problématique en question. Pour cette raison nous comparons plusieurs variantes de RNA.

Koivo (1994), Park et al (2002), Fatima et hamid (2009), Mssaf et Belmajdoub (2015) ont comparé différents réseaux neurones. Le tableau (*Tableau 12*) montre une synthèse de cette comparaison.

Après cette étude critique, notre choix s'est fixé sur le réseau de neurones de régression généralisé GRNN. Le principal avantage du GRNN, en plus de ceux citée dans le tableau (*Tableau 12*), est la non existence de l'apprentissage non supervisé comme dans le réseau RBF ; car le nombre de gaussienne dans la couche cachée est fixe. Donc ce type de réseau peut être une bonne alternative au RBF et au PMC. En outre, le GRNN et comme son nom l'indique peut, potentiellement, remplacer la méthode classique de la régression multiple. Avec le GRNN nous voulons simultanément prédire le facteur d'amplification spectrale de site et analysé leurs incertitudes

Dans la section suivante, une description théorique du GRNN est donnée.

Tableau 12 Avantages et inconvénient de quelque réseaux de neurones statiques

RNA	Avantages	Inconvénients
PMC	<ul style="list-style-type: none"> • Accepte les données bruitées et la classification non-linéaire. • Possibilité d'ajouter plusieurs couches cachées • Plus maniable • Représentation globale de l'espace • Architecture simple 	<ul style="list-style-type: none"> • Le nombre de couches cachées et de neurones par couche cachée sont à définir. • Pas d'architecture unique • Risque de tomber dans un minimum local • Demande un temps de calcul non négligeable
RVFLNN	<ul style="list-style-type: none"> • Accepte les données bruitées et la classification non-linéaire. • Représentation globale de l'espace • Des entrées optimisées 	<ul style="list-style-type: none"> • Architecture plus complexe que dans le PMC • Le nombre de couches cachées et de neurones par couche cachée sont à définir. • Pas d'architecture unique • Risque de tomber dans un minimum local • Demande un temps de calcul non négligeable
RBF	<ul style="list-style-type: none"> • Accepte les données bruitées et la classification non-linéaire. • Représentation locale de l'espace • Convergence rapide 	<ul style="list-style-type: none"> • Nombre de neurones dans la couche cachée égale aux nombre des exemples utilisés : Nécessite une grande capacité de calcul.
GRNN	<ul style="list-style-type: none"> • Dédié pour l'approximation des fonctions • Architecture unique et simple • Apprentissage simple et convergence rapide • Peu de paramètres à optimiser 	<ul style="list-style-type: none"> • A utiliser seulement avec les bases de données non volumineuses.

4.6 Aspect théorique du GRNN : cas particulier d'une seule sortie

Le réseau de neurones GRNN a une architecture assez semblable à celle des réseaux de neurones à base radiales (radial basis networks ou RBF) (Wasserman, 1993). Le GRNN est composé de 4 couches : la couche d'entrée, couche de motifs (pattern), la couche de sommation et la couche de sortie (Figure 4-4). (voir : Specht, 1991 ; Wasserman, 1993 ; Hannan et al., 2010 ; Boudghene Stambouli et al., 2017).

La couche d'entrée transporte simplement les données à la couche suivante par une Architecture parallèle. En ce qui concerne la présente étude, les entrées du réseau de neurones sont le ou les proxies des conditions locales de site. Il s'agit de : la profondeur total du sédiment (Depth), la fréquence de résonance (f_0), la contraste de vitesse (Cv), V_{s30} , la vitesse moyenne sur tout le profil (V_{sm}), et la vitesse au niveau du substratum rocheux. Cette vitesse est prise égale à 800m/s pour les profils normalisés et les profils tronqués.

Dans la couche de motifs la fonction d'activation est gaussienne (voir *Tableau 11*), le nombre de neurones est déterminé automatiquement et représente le nombre d'échantillon (exemples) d'apprentissage.

La couche de sommation c'est la troisième couche du réseau. Elle se constitue de deux neurones : SN (le numérateur) et SD (le dénominateur). Les deux sont régis par une fonction d'activation linéaire. Dans cette couche s'effectue le calcul de la somme des réponses pondérées des motifs. Les poids de liaisons entre le neurone SN et le $j^{\text{ème}}$ neurone de la couche des motifs est représenté par LW_j qu'est la sortie elle-même de la couche des motifs. SD (sommation des neurones) est utilisée pour calculer les sorties non pondérée des neurones de la couche des motifs. Les poids de liaisons entre le neurone SD et la couche des motifs sont égales à l'unité (1) (voir Figure 4-4).

Enfin, la couche de sortie, représente les prédictions $y=SN/SD$ (Hannan et al., 2010). Dans l'actuelle étude il s'agit des facteurs d'amplifications spectrales.

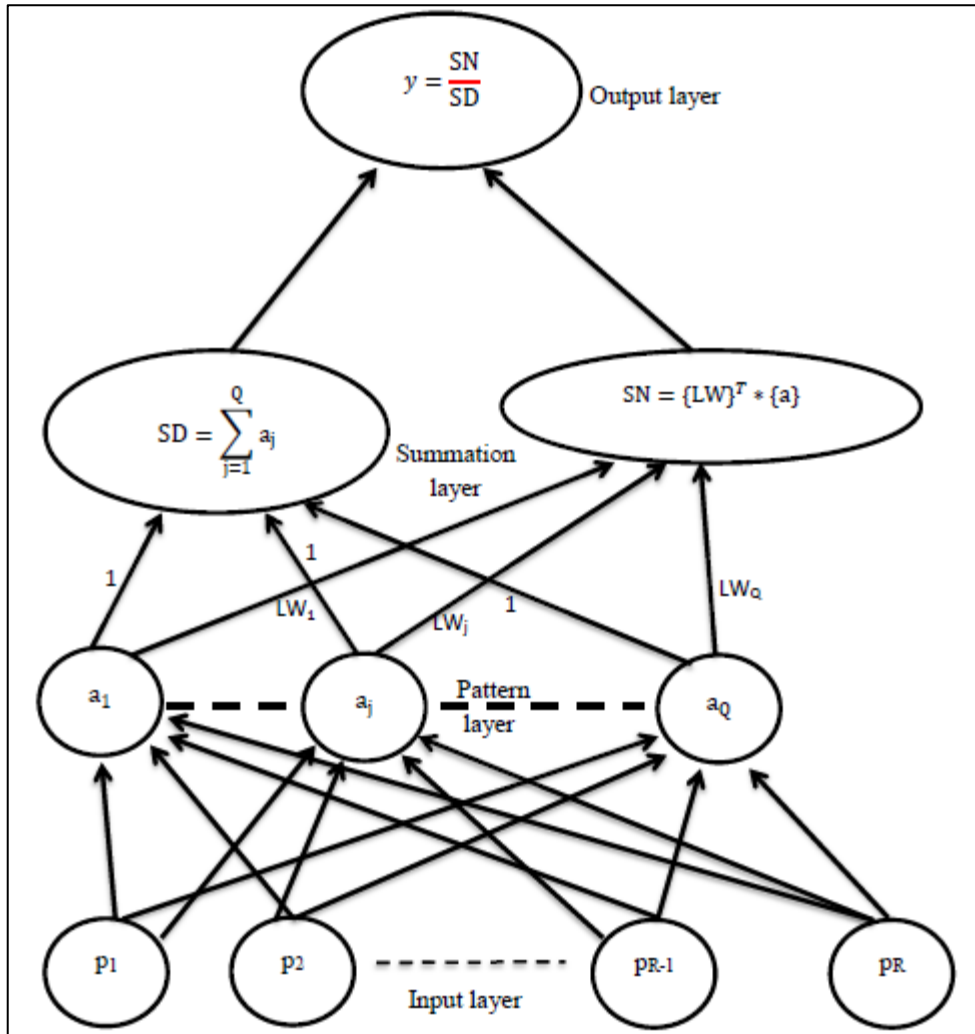


Figure 4-4 Architecture du GRNN avec une seule sortie.

Les détails de mise en œuvre sous Matlab avec la toolbox ANN (www.mathworks.com) ainsi que les équations associées sont décrites dans la section suivant.

4.6.1 Mise en œuvre de GRNN sous Matlab : Aspect pratique

Le GRNN est souvent utilisé pour avoir une approximation d'une fonction continue quelconque, cette dernière est approchée par une combinaison linéaire des fonctions gaussiennes bien choisies, où chaque fonction gaussienne fournis un point de la fonction désirée 'y' (Figure 4-5).

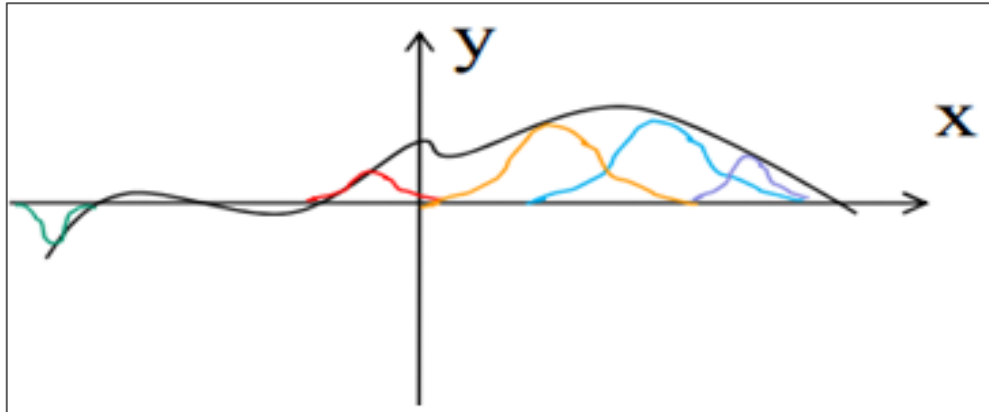


Figure 4-5 :Approximation d’une fonction quelconque Y (en noir) par des fonctions gaussiennes(données en différentes couleurs (Alani, 2008).

La figure ci-dessous représente la structure du réseau GRNN telle qu’elle est donnée sous Matlab (www. mathworks.com).

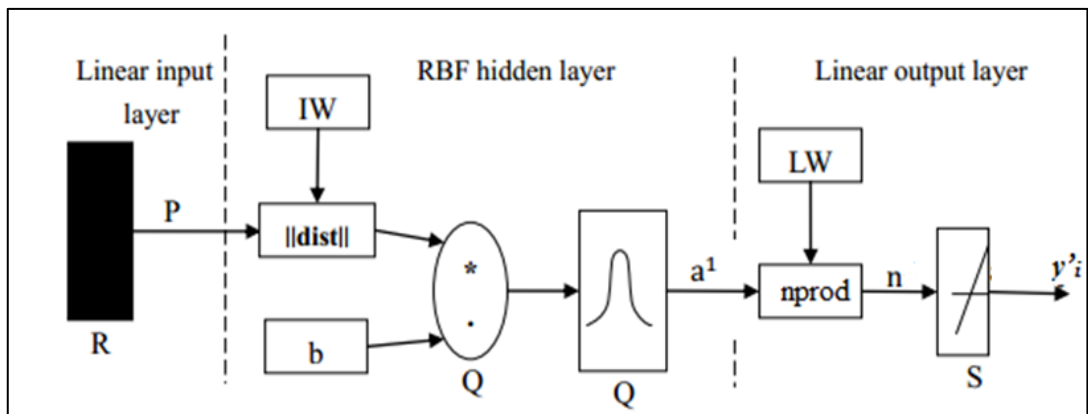


Figure 4-6 : Structure du GRNN (Matlab).

Le GRNN est un modèle qui se compose, comme c’est indiqué dans la section précédente, d’une couche d’entrée linéaire, la couche qui abrite la fonction d’activation (e.g. gaussienne), la couche de sommation et la couche de sortie linéaire (voir Figure 4-6).

La première couche transporte les données à la couche cachée. Dans cette dernière, le nombre de neurones est égal au nombre d’exemples l’apprentissage. La différence entre les poids (IW : centre de la gaussienne) et la valeur d’entrée (choisis) représente la distance euclidienne ($dist$) (équation 4.1). Pour une distance ($dist$) grande on aura une valeur petite (faible importance), inversement, avec une

petite (dist) la sortie est importante (valeur grande). Cela donne une bonne capacité d'approximation locale, ce qui est important pour améliorer l'efficacité de l'apprentissage GRNN (Liang et al, 2014) (voir Figure 4-6). Dans cette même figure, le b représente le seuil (le biais). La fonction de transfert de cette couche est une fonction gaussienne axisymétrique et est caractérisée par son centre et sa largeur.

$$\|dist\| = \|IW - P\| = \sqrt{\sum_{k=1}^R (IW_{jk} - P_k)^2} \quad (4-1)$$

k=1,2,...,R ; (R est le nombre de paramètres d'entrée (les proxies de site)).

j=1,2,3,...,Q ; (Q est le nombre de neurone dans la couche cachée ou le nombre de profiles pris dans l'apprentissage).

P:les entrées (pour notre cas se sont les proxies définis au paravent).

IW : les poids entre la couche d'entrée et la couche cachée :centre de la gaussienne.

$$.a = radbas(\|IW - P\|b) \quad (4-2)$$

La fonction de transfert (radbas) de chaque neurone est une fonction à base radial qu'est présenté par:

$$.radbas(n) = e^{-n^2} \quad (4-3)$$

$$b = \frac{\sqrt{\log(5)}}{spread} \quad (4-4)$$

avec :

b :biais ; spread : la largeur de la gaussienne.

La dernière couche représente la couche de sortie représentée dans Matlab par la fonction 'normprod'. La i^{ème} sortie est donnée par n_i :

$$n_i = \text{normprod}(LW, a) = \frac{LW * [a]}{\sum_{j=1}^Q a_j} \quad (4-5)$$

$j=1,2,3,\dots,Q$; (Q est le nombre de neurone dans la couche cachée ou le nombre de profiles pris dans l'apprentissage).

$i=1,2,3,\dots,\text{npf}$; npf : le nombre de vecteur de sortie.

LW : la matrice des poids entre la couche cachée et la couche de sommation, sa taille est ($Q*\text{npf}$).

$$y'_i = \text{purelin}(n_i) = n_i \quad (4-6)$$

purelin : Fonction d'activation linéaire.

y'_i : dans cette étude, il s'agit du facteur d'amplification à prédire à i période ou à i fréquence normalisée.

4.6.2 Etape d'élaboration d'un modèle avec un réseau GRNN

La procédure d'apprentissage du GRNN peut être divisée en six étapes :

1) La subdivision du jeu de données en deux sous-ensembles : le premier pour la phase d'apprentissage et le deuxième pour la phase test.

2) La création des réseaux GRNN, tel qu'il est évoqué dans la section précédente

3) Simulation du modèle : Le paramètre le plus délicat à simuler, est la largeur de la gaussienne (spread) ; puisque le nombre de neurones est fixé automatiquement (voir Benrcia et al., 2010; Hannan et al., 2010; Liang et al., 2014). La largeur de la gaussienne est obtenue par une façon approximative où deux approches sont utilisées et mentionnées à la fin de ce chapitre.

4) Détermination des performances du réseau GRNN : les performances sont déterminées par l'analyse des résidus (r_{ij}), l'écart type et le coefficient de corrélation. Et surtout par la vérification des tendances physiques.

5) Spécification des paramètres du réseau GRNN : A l'issue de l'étape 3) et 4) (analyse statistique et physique), nous pouvons choisir la largeur de la gaussienne du modèle final.

6) Obtention du modèle final: Avec le GRNN nous pouvons écrire l'équation de prédiction qui est en fonction des paramètres d'entrées et des différents paramètres du réseau (biais(b), spread, (poids :IW,LW...)). Dans la présente thèse il s'agit de l'équation qui donne les deux facteurs d'amplification spectrale de Borchardt (F_a et F_v) en fonction des proxies de site.

4.6.3 Mesure des performances

La comparaison des résultats obtenus par le modèle neuronal et ceux mesurés sur site doivent être confrontés. Généralement on utilise des paramètres statistiques pour mesurer la fiabilité du modèle et par la suite le valider. On fournit des paramètres qui peuvent être employés pour mesurer les performances du réseau de neurones pour un ensemble de données particulières.

✓ R_c : coefficient de corrélation.

$$R_c = \frac{\frac{\sum_j (y'_j - \bar{y}')(\hat{y}_j - \bar{\hat{y}})}{Q}}{\sqrt{\frac{\sum_j (\hat{y}_j - \bar{\hat{y}})^2}{Q}} \sqrt{\frac{\sum_j (y'_j - \bar{y}')^2}{Q}}} \quad (4-7)$$

J : indique le nombre de profils soumis à l'apprentissage (J=1 :Q).

$\bar{\hat{y}}$: La moyenne de la sortie du calcul exact (modèle linéaire viscoélastique) sur tous les profils de sol pris.

\bar{y}' : La moyenne de la sortie prédit

\hat{y}_J : la moyenne de la sortie du calcul exact (modèle linéaire viscoélastique) pour J^{ieme} profil de sol.

y'_j : La moyenne de la sortie prédit pour j^{ieme} profil de sol.

- L'écart type du modèle (σ_{model})

$$\sigma_{model} = \sqrt{\frac{\sum_{j=1}^Q \left(\sum_{i=1}^{npf} (\log(y'_{i,j}) - \log(\widehat{y}_{i,j}))^2 \right)}{Q \cdot npf}} \quad (4-8)$$

- L'écart type max (σ_{max})

$$\sigma_{max} = \max_{(i=1:npf)} \left(\sqrt{\frac{\sum_{j=1}^Q \left((\log(y'_{i,j}) - \log(\widehat{y}_{i,j}))^2 \right)}{Q}} \right) \quad (4-9)$$

Où

y'_{ij} : La sortie prédit (ie : facteur amplification à i periode pour j profiles).

$\widehat{y}_{i,j}$: La sortie prix de la base de donnée (ie : facteur amplification à i periode pour j profiles).

- Somme carrée des erreurs (Sse)

$$Sse = \sum_{j=1}^Q \left(\sum_{i=1}^{npf} (\log(y'_{i,j}) - \log(\widehat{y}_{i,j}))^2 \right) \quad (4-10)$$

- Résidus max (r_{ij})

$$r_{ij} = \frac{1}{Q} \sum_{j=1}^Q (\max_{(i=1:npf)} (\log(y'_{i,j}) - \log(\widehat{y}_{i,j}))) \quad (4-11)$$

4.6.4 Détermination et optimisation de la largeur de la gaussienne

Il est à noter que, dans le GRNN, la largeur de la gaussienne représente le seul paramètre à configurer : Pour optimiser cette largeur nous avons utilisé les deux approches complémentaires suivantes :

1 Dans la première approche, la largeur peut être optimisée par la validation croisée (voir annexes), en calculant la performance du réseau par la somme carrée des erreurs (Sse) : donnée par l'équation (4-10). La valeur de la largeur avec laquelle nous avons obtenu le Sse le plus faible représente la valeur de la largeur optimale (un exemple est donné sur la figure (Figure 4-7a). Dans cet exemple nous avons utilisé 50% de la base de données (phase test). Dans cette figure nous remarquons que la largeur optimale est comprise entre 0 et 0.5.

À savoir que les entrées et les sorties ont été normalisées par la fonction (mapstd) (voir : www.mathworks.com). Dans « mapstd » la moyenne et l'écart type sont égales à « 0 » et « 1 » respectivement.

La formule générale de mapstd d'un vecteur P avec des sorties y peut s'exprimer comme suite :

$$y = (p - pmean) \left(\frac{ystd}{pstd} \right) + ymean \quad (4-12)$$

Où :

Pmean et ymean sont les moyennes des vecteurs p et y respectivement.

ystd et pstd sont l'écart type des vecteurs y et p par rapport leur moyennes respectivement. (Pour toute information : voir : www.mathworks.com)

2 Dans la deuxième approche, la largeur de la gaussienne est optimisée en calculant le résidu max (r_{ij}) (équation 4-11). Par la suite nous calculons l'écart type et le coefficient de corrélation ainsi les résidus. Il est à noter que dans cette approche, toute la base de données est utilisée pour l'apprentissage. A savoir que les entrées et les sorties n'ont pas été normalisés.

- Un exemple de calcul est donné sur la figure (Figure 4-7b). Cette figure montre que la performance (r_{ij}) est bonne lorsque la largeur est comprise entre 0.01 et 0.2. A cette dernière valeur (r_{ij}) des résidus est égale à 0.33.

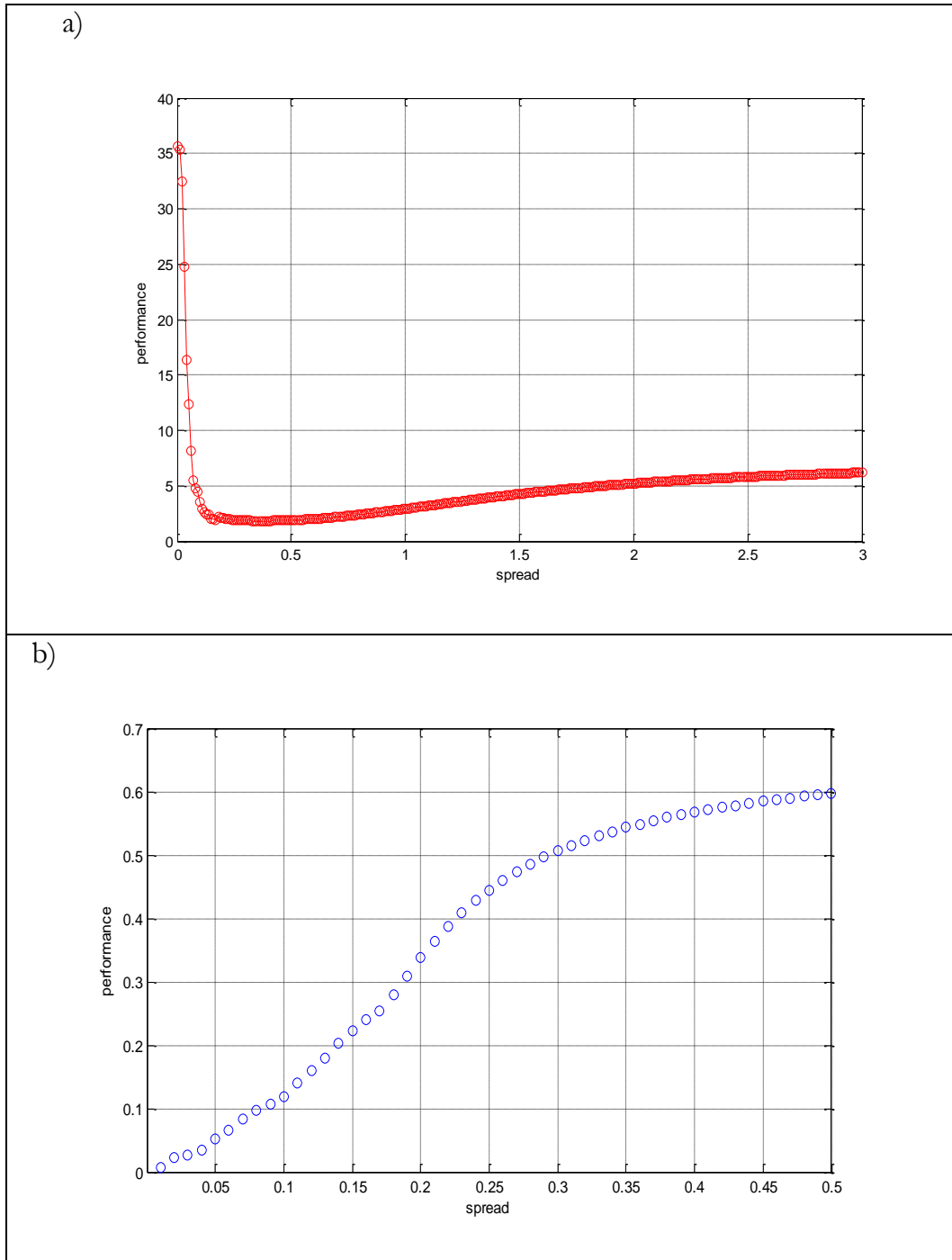


Figure 4-7 Variation de la performance en fonction du la largeur de la gaussienne(spread), ((a)la performance est mesuré par Sse), (b par le résidu max)).

Le but est de trouver une valeur exacte de la largeur de la gaussienne car des valeurs trop petites provoquent un sur-apprentissage, tandis que des valeurs de trop grandes provoquent un sous-apprentissage. La figure ci-dessous illustre ce problème.

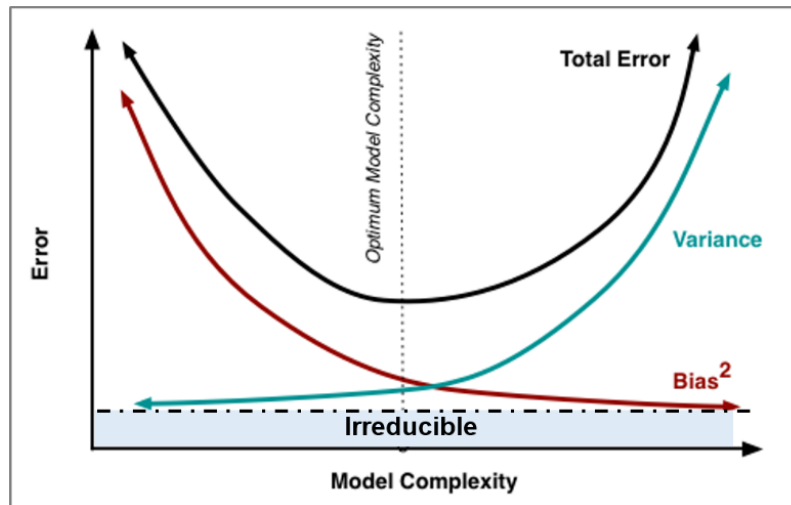


Figure 4-8 Biais et variance contribuant à l'erreur totale.
(<http://scott.fortmann-roe.com/docs/BiasVariance.html>).

Donc, il faut avoir un compromis entre le biais et la variance des résidus. Pour ce faire, nous avons pris plusieurs jeux de données (50% de la base de données totale) triés aléatoirement.

Avec chaque jeu de données nous avons calculé la largeur optimale (en utilisant les deux approches déjà citées). A la fin, nous avons pris la moyenne arithmétique de la largeur de chaque échantillon soumis à l'apprentissage.

$$spread = \frac{1}{k} \sum_{u=1}^k spread(k) \quad (4-13)$$

Spread(k) : la largeur trouver pour un échantillons k soumis à l'apprentissage.

Spread est la largeur finale utilisée pour le modèle final.

Il est à noter que, avec cette procédure de calcul, la valeur de la largeur reste préliminaire. La capacité du modèle à prédire des facteurs d'amplification qui respectent l'aspect physique du phénomène est décisif au choix de la largeur de la gaussienne.

4.7 Conclusion

Dans ce chapitre nous avons donné un aperçu sur la méthode des réseaux de neurones artificiels. Nous avons montré aussi que le réseau de neurones de régression généralisée (GRNN) représente une bonne alternative à la fois par rapport aux modèles classiques de régression et aux perceptrons multicouches (PMC). Le plus grand avantage des GRNN est la simplicité de sa mise en œuvre. Il converge rapidement vers un minimum global (Specht, 1991), en étant économique en ressources mémoire et en temps de calcul. D'autre part, son architecture unique présente un atout non négligeable puisque le seul paramètre à optimiser est la largeur de la gaussienne.

Chapitre 5

Résultats et interprétations

5.1 Introduction

La question abordée dans cette étude est d'identifier les meilleurs paramètres du site qui expliquent de manière optimale et donc prédisent le facteur d'amplification propre au site. L'accent est mis ici sur la réponse 1D des colonnes de sols stratifiés horizontalement et les relations entre les facteurs d'amplification correspondants sur les spectres de réponse et un nombre limité de «paramètres récapitulatifs» décrivant les caractéristiques globales du profil de sol.

5.2 Description des jeux de données

Les codes parasismiques actuels prennent l'effet de site en les considérant par l'intermédiaire de l'introduction de certaines catégories de sol basées sur certains paramètres, parmi lesquelles : la vitesse de cisaillement moyenne sur toute la profondeur utilisée dans le RPA99 « Règlement parasismique algérien, 1999 » (DTR, 2003), la vitesse de cisaillement moyenne sur trente mètres de profondeur est considérée dans le code de bâtiments UBC97 « Uniform Building Code, 1997 » (Klimis et al ,1998), est utilisé également dans « Eurocode8,1994 » (Bard.et al ,2000) et la fréquence caractéristique de sol, donnée par Arias (Saragoni. et al, 1998).

Les inputs des réseaux de neurones sont constitués des paramètres de site et des facteurs d'amplification spectrale en fonction de la période, des facteurs d'amplifications pour des fréquences normalisés. Ce choix initial des paramètres de site va être consolidé avec des tests sous forme d'exemples toute en calculant certains paramètres statistiques comme (l'écart-type, coefficient de corrélation R^2 , ...) et ce là à chaque paramètre d'entrée . De plus on va évaluer les tendances pour mieux comprendre la physique du modèle.

5.2.1 Facteur d'amplification en fonction de la période

La base de donnée décrite dans le chapitre III est composée des amplifications spectrales $AF_m(P_k, \theta, T_i)$ pour tous les types de profils (RP, NP et TP). A partir de cette base de données on calcule l'amplification moyenne $A_0(\theta, T_i)$ et ainsi la variabilité $\sigma_0(\theta, T_i)$ associé à chaque jeux de profils (RP, NP et TP). La figure (Figure 5-1) donne la variation de ces deux fonctions pour les profils (RP). Il est à noter que la figure (Figure 5-1a) donne l' amplification moyenne pour les profils de sol type (RP), la figure (Figure 5-1b) donne la variabilité au sein des données de la base de données pour les profils (RP).

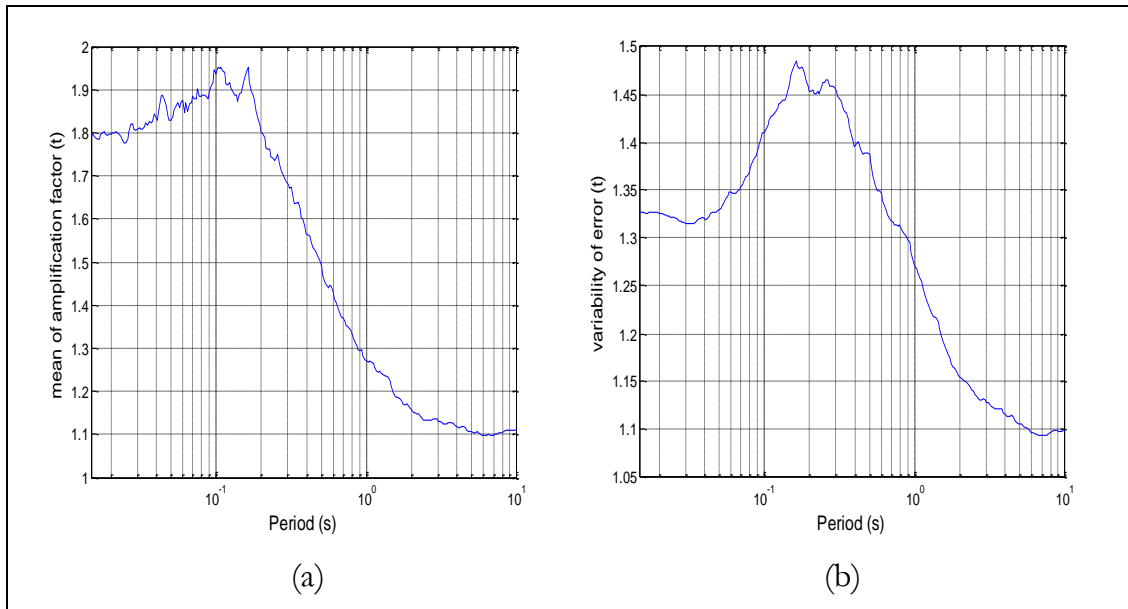


Figure 5-1 : (a) Variation de l'amplification moyenne $A_0(\theta, T_i)$, (b) Variation de la variabilité $\sigma_0(\theta, T_i)$.

Le tableau ci-dessous présente les valeurs de la variabilité (total, max ainsi à des périodes spécifiés) au sein de la base de donnée pour tous les types de profils (RP, NP et TP).

Tableau 13 Calcul des variabilités.

Total initial variability $\sigma_m(\theta = 1)$ (RP)	0.1178	Total initial variability $\sigma_m(\theta = 0)$ (NP)	0.0846	Total initial variability $\sigma_m(\theta = 2)$ (TP)	0.0896
Maximum initial variability $\sigma_{max}(\theta = 1)$	0.1717	Maximum initial variability $\sigma_{max}(\theta = 0)$	0.1232	Maximum initial variability $\sigma_{max}(\theta = 2)$	0.1317
$\sigma(\theta = 1, T = 0.01 \text{ s})$	0.1227	$\sigma(\theta = 0, T = 0.01 \text{ s})$	0.0811	$\sigma(\theta = 2, T = 0.01 \text{ s})$	0.0866
$\sigma(\theta = 1, T = 0.02 \text{ s})$	0.1226	$\sigma(\theta = 0, T = 0.02 \text{ s})$	0.0809	$\sigma(\theta = 2, T = 0.02 \text{ s})$	0.0861
$\sigma(\theta = 1, T = 0.04 \text{ s})$	0.1206	$\sigma(\theta = 0, T = 0.04 \text{ s})$	0.0756	$\sigma(\theta = 2, T = 0.04 \text{ s})$	0.0759
$\sigma(\theta = 1, T = 0.07 \text{ s})$	0.1314	$\sigma(\theta = 0, T = 0.07 \text{ s})$	0.0883	$\sigma(\theta = 2, T = 0.07 \text{ s})$	0.0861
$\sigma(\theta = 1, T = 0.1 \text{ s})$	0.1494	$\sigma(\theta = 0, T = 0.1 \text{ s})$	0.1062	$\sigma(\theta = 2, T = 0.1 \text{ s})$	0.106
$\sigma(\theta = 1, T = 0.2 \text{ s})$	0.1623	$\sigma(\theta = 0, T = 0.2 \text{ s})$	0.1169	$\sigma(\theta = 2, T = 0.2 \text{ s})$	0.1242
$\sigma(\theta = 1, T = 0.4 \text{ s})$	0.1446	$\sigma(\theta = 0, T = 0.4 \text{ s})$	0.1089	$\sigma(\theta = 2, T = 0.4 \text{ s})$	0.1188
$\sigma(\theta = 1, T = 0.7 \text{ s})$	0.1200	$\sigma(\theta = 0, T = 0.7 \text{ s})$	0.098	$\sigma(\theta = 2, T = 0.7 \text{ s})$	0.1093
$\sigma(\theta = 1, T = 1.0 \text{ s})$	0.1040	$\sigma(\theta = 0, T = 1.0 \text{ s})$	0.0873	$\sigma(\theta = 2, T = 1.0 \text{ s})$	0.0982
$\sigma(\theta = 1, T = 2.0 \text{ s})$	0.0626	$\sigma(\theta = 0, T = 2.0 \text{ s})$	0.0503	$\sigma(\theta = 2, T = 2.0 \text{ s})$	0.0552
$\sigma(\theta = 1, T = 4.0 \text{ s})$	0.0477	$\sigma(\theta = 1, T = 4.0 \text{ s})$	0.033	$\sigma(\theta = 1, T = 4.0 \text{ s})$	0.036
$\sigma(\theta = 1, T = 7.0 \text{ s})$	0.0388	$\sigma(\theta = 1, T = 7.0 \text{ s})$	0.027	$\sigma(\theta = 1, T = 7.0 \text{ s})$	0.0296
$\sigma(\theta = 1, T = 10.0 \text{ s})$	0.0412	$\sigma(\theta = 0, T = 10.0 \text{ s})$	0.0297	$\sigma(\theta = 2, T = 10.0 \text{ s})$	0.0329

Il est important de noter que les AF qui seront simulés par la GRNN doivent avoir de préférence une variabilité inférieure à celle de la base de données.

La figure ci-dessous présente le calcul de l'amplification pour n profils de sol à différents type de profils de sol (RP, NP et TP) ainsi leur correspondant amplification moyenne plus au moins la variabilité associée à chaque cas.

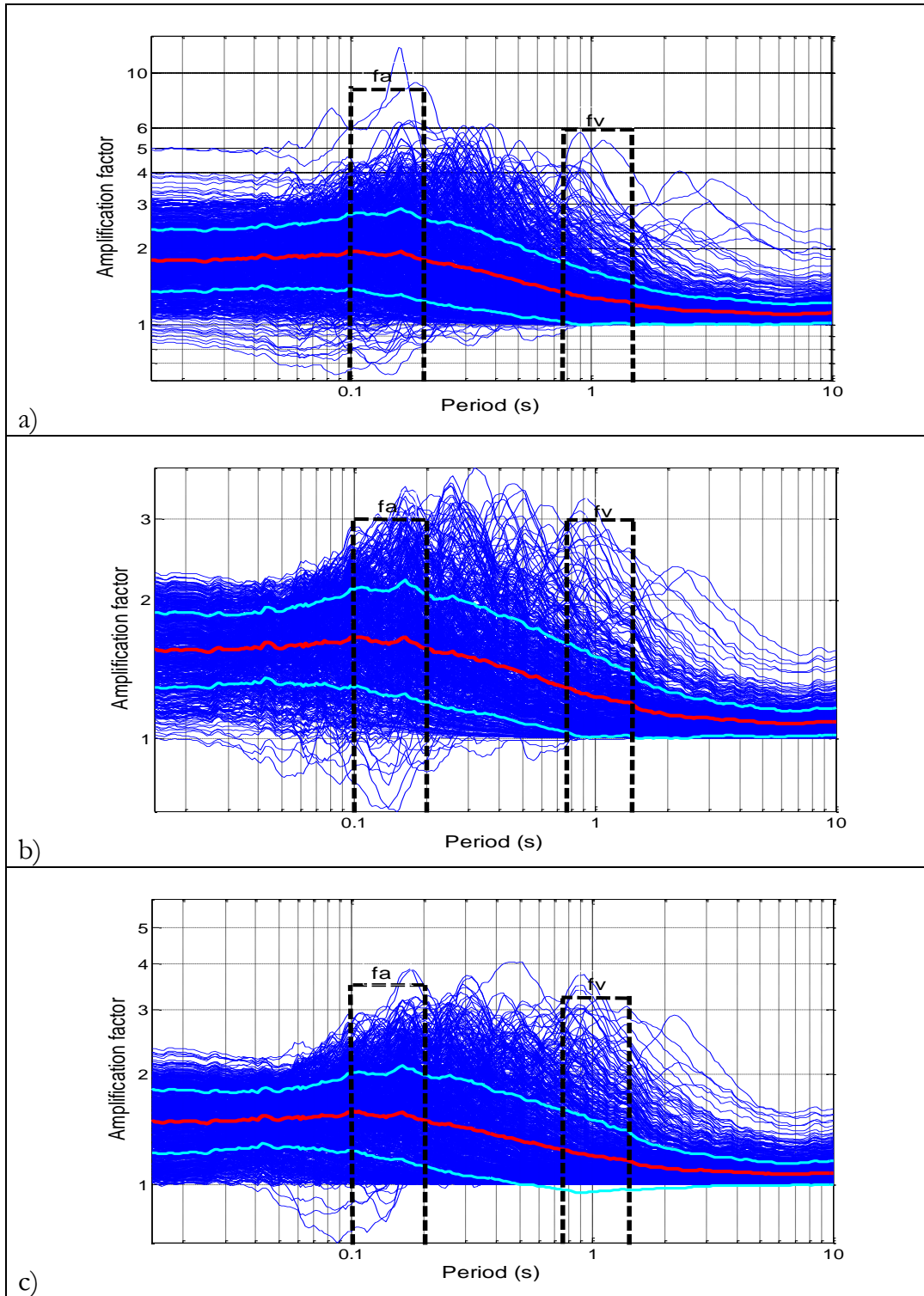


Figure 5-2 : Facteur d'amplification pour tous les profils en fonction de la période $AF_m(P_k, \theta, T_i)$, ainsi la variation de l'amplification moyenne de $A_0(\theta, T_i)$ (courbe rouge), et éventuellement la variation de $AF_m(P_k, \theta, T_i) \pm \sigma_0(\theta, T_i)$ application pour les différents type de profils (a : RP , b : NP et c : TP).

5.2.2 Facteur d'amplification en fonction de la fréquence normalisée

Une deuxième méthode consiste à prendre tous les facteurs d'amplifications spectrales $AF_m \left(P_k, \theta, \frac{f_i}{f_0} \right)$ normalisées en fréquence par la fréquence caractéristique pour chaque site (f_0) calculé par la méthode de de Rayleigh simplifié (défini dans le chapitre III).

Lors de la normalisation des fréquences pour chaque site, il s'est avéré qu'on n'a pas le même nombre de profils à chaque fréquence. La courbe en cloche ci-dessous montre la variation du nombre de profils en fonction de la fréquence normalisée pour les différents type jeux de données (RP, NP et TP). Pour ne pas avoir un nombre de profils très réduit par cette normalisation on a choisi de considérer dans l'intervalle de fréquence normalisée (f/f_0) [0.03, 30]. Une fois ce choix a été fait, on a procédé à une interpolation linéaire de tous les profils pour avoir les mêmes valeurs des fréquences normalisés à chaque colonne pour tous les profils et éventuellement pour les trois jeux de données (RP, NP et TP). La figure ci-dessous celle de droite montre un exemple d'un profil réelle (RP) et un autre interpolé.

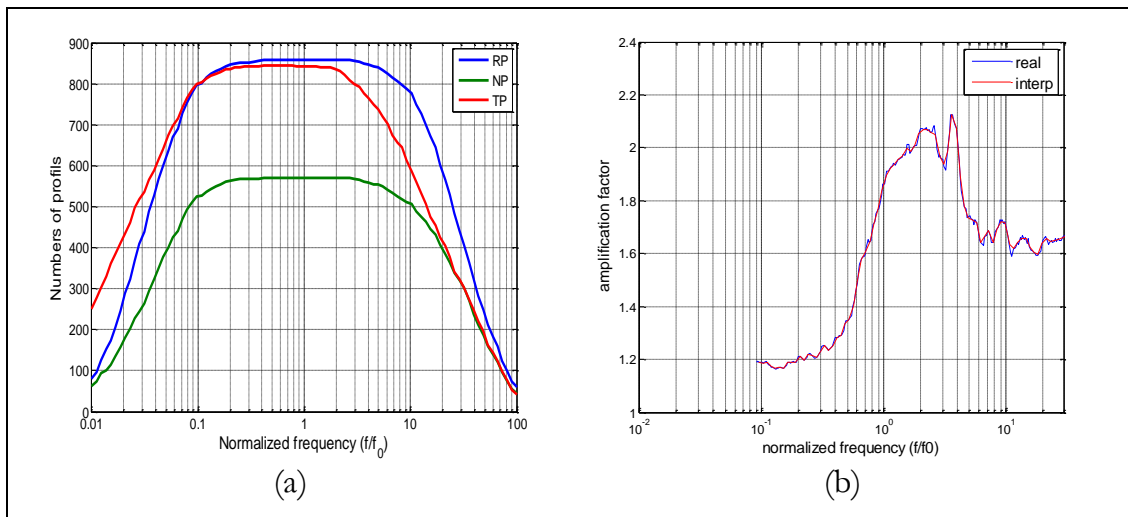


Figure 5-3 : (a) Variation du nombre de profils à chaque fréquence normalisée (f/f_0), (b) Exemple de facteur d'amplification réelle et interpolé pour le même profil.

La figure (Figure 5-4) donne la variation de l'amplification en fonction de la fréquence normalisée et ce là pour tous les types profils (RP, NP et TP). Ainsi leur amplification moyenne correspondant plus au moins à la variabilité associée à chaque cas.

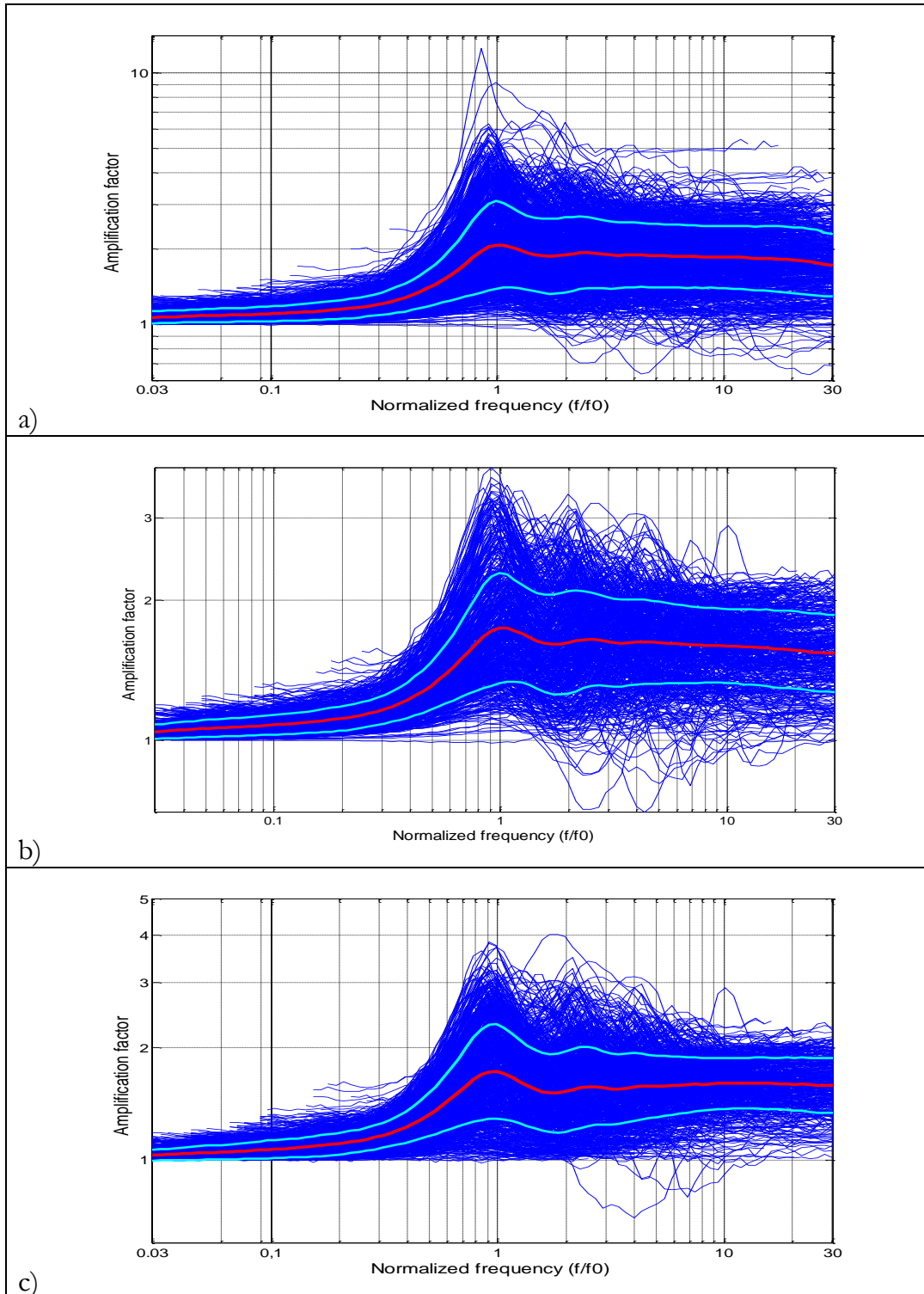


Figure 5-4: Facteur d'amplification pour tous les profils en fonction de la fréquence normalisées $AF_m(P_k, \theta, f_i/f_0)$, ainsi la variation de l'amplification moyenne de $A_0(\theta, f_i/f_0)$ (courbe rouge) et éventuellement la variation de $AF_m(P_k, \theta, f_i/f_0) \pm \sigma_0(\theta, f_i/f_0)$ application pour les différents type de profils ((a) RP, (b) NP et (c) TP).

Le tableau ci-dessous montre les valeurs de la variabilité (total, max ainsi à des fréquences normalisées spécifiés) au sein de la base de donnée pour tous les types profils (RP, NP et TP).

Tableau 14 : Calcul des variabilités pour des profils normalisés en fréquence.

Total initial variability $\sigma_m(\theta = 1)$	0.1060	Total initial variability $\sigma_m(\theta = 0)$	0.0724	Total initial variability $\sigma_m(\theta = 2)$	0.0714
Maximum initial variability $\sigma_{max}(\theta = 1)$	0.1736	Maximum initial variability $\sigma_{max}(\theta = 0)$	0.1187	Maximum initial variability $\sigma_{max}(\theta = 2)$	0.1268
$\sigma(\theta = 1, f/f_0 = 0.05)$	0.0261	$\sigma(\theta = 0, f/f_0 = 0.05)$	0.0183	$\sigma(\theta = 2, f/f_0 = 0.05)$	0.018
$\sigma(\theta = 1, f/f_0 = 0.1)$	0.0301	$\sigma(\theta = 0, f/f_0 = 0.1)$	0.0215	$\sigma(\theta = 2, f/f_0 = 0.1)$	0.024
$\sigma(\theta = 1, f/f_0 = 0.2)$	0.0373	$\sigma(\theta = 0, f/f_0 = 0.2)$	0.0269	$\sigma(\theta = 2, f/f_0 = 0.2)$	0.029
$\sigma(\theta = 1, f/f_0 = 0.4)$	0.0600	$\sigma(\theta = 0, f/f_0 = 0.4)$	0.0378	$\sigma(\theta = 2, f/f_0 = 0.4)$	0.0418
$\sigma(\theta = 1, f/f_0 = 0.7)$	0.1305	$\sigma(\theta = 0, f/f_0 = 0.7)$	0.0904	$\sigma(\theta = 2, f/f_0 = 0.7)$	0.102
$\sigma(\theta = 1, f/f_0 = 1.0)$	0.1673	$\sigma(\theta = 0, f/f_0 = 1.0)$	0.1149	$\sigma(\theta = 2, f/f_0 = 1.0)$	0.1228
$\sigma(\theta = 1, f/f_0 = 2.0)$	0.1473	$\sigma(\theta = 0, f/f_0 = 2.0)$	0.1101	$\sigma(\theta = 2, f/f_0 = 2.0)$	0.1074
$\sigma(\theta = 1, f/f_0 = 4.0)$	0.1281	$\sigma(\theta = 0, f/f_0 = 4.0)$	0.0912	$\sigma(\theta = 2, f/f_0 = 4.0)$	0.0908
$\sigma(\theta = 1, f/f_0 = 7.0)$	0.1248	$\sigma(\theta = 0, f/f_0 = 7.0)$	0.0829	$\sigma(\theta = 2, f/f_0 = 7.0)$	0.0741
$\sigma(\theta = 1, f/f_0 = 10.0)$	0.1239	$\sigma(\theta = 0, f/f_0 = 10.0)$	0.0794	$\sigma(\theta = 2, f/f_0 = 10.0)$	0.0682
$\sigma(\theta = 1, f/f_0 = 20.0)$	0.1312	$\sigma(\theta = 0, f/f_0 = 20.0)$	0.0833	$\sigma(\theta = 2, f/f_0 = 20.0)$	0.0712

5.2.3 Facteur d'amplification à courte et moyenne période

La troisième méthode consiste à calculer deux facteurs, un à courte période (0.1Sec) et l'autre à moyenne période (1Sec) et ceci à partir des résultats initiaux tels :

-Fa est le facteur d'amplification à courte période, dans ce cadre on a adopté une moyenne géométrique des facteurs d'amplification entre [0.1s 0.2s] pour chaque profils (voir Figure 5-2).

-Fv est le facteur d'amplification à moyenne période, dans ce cadre on a adopté une moyenne géométrique des facteurs d'amplification entre [0.75s 1.5s] pour chaque profils (voir Figure 5-2).

Tableau 15 Calcul des variabilités pour Fa et Fv cas des profils (RP).

Variability total of error (in log10)(Fa)	0.1543
Variability total of error (in log10)(Fv)	0.0995
Moyenne de Fa	1.8978
Moyenne de Fv	1.2699

5.3 Description des différents modèles de réseaux de neurones

Comme présenté dans les chapitres précédents, une série de 858 profils de sol réel a été considérée et leur réponse viscoélastique linéaire à des ondes S incidentes verticalement est calculée pour 14 accelerogrammes réels distincts couvrant une large plage de fréquences. On a aussi indiqué que ces profils de sol seront normalisés (NP).

Pour chaque site P_k , on a calculé le facteur d'amplification $AF(P_k, \theta, S_l, T_i)$

et ensuite le facteur d'amplification géométrique moyen $AF_m(P_k, \theta, T_i)$ a été dérivé des 14 facteurs d'amplification (voir Section 3.3) pour chaque profil. Il est nécessaire de rappeler que le même travail a été effectué pour des profils tronqués à 800m/s.

Les inputs des multiples modèles de réseaux sont en fait les facteurs d'amplification calculés par la réponse viscoélastique $AF_m(P_k, \theta, T_i)$ ainsi que les paramètres du site P_k . Ces paramètres sont au nombre de six :

$$\{x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad x_6\}^T = \{Depth \quad V_{sm} \quad V_{s30} \quad C_v \quad f_0 \quad V_{bedrock}\}^T \quad (5-1)$$

Plusieurs modèles de réseaux de neurones ont été développés car notre input dépend du nombre de paramètres de site utilisés (proxy). Ainsi considérons par exemple le cas où on a introduit un seul paramètre de site, on peut avoir six modèles de réseaux de neurones car on a six paramètres de site. Par exemple, si on choisit d'introduire deux paramètres de site, on aura donc plusieurs combinaisons. Par contre lorsqu'on décide de prendre six paramètres de site, on aura uniquement un seul cas. Donc le facteur d'amplification obtenu par le réseau de neurone dépendra donc du facteur d'amplification $AF_m(P_k, \theta, T_i)$ mais aussi du nombre de paramètres de site. En fait ce facteur sera noté: $AF(P_k, \theta, T_i, x_1 \dots x_R)$.

Il a été aussi indiqué que l'input pourra être $AF_m \left(P_k, \theta, \frac{f_i}{f_0} \right)$ au lieu de $AF_m(P_k, \theta, T_i)$ et par conséquent l'output du réseau de neurone sera $AF \left(P_k, \theta, \frac{f_i}{f_0}, x_1 \dots x_R \right)$.

Enfin, on aura des modèles de réseaux de neurones où l'input sera le paramètre ou les paramètres de site et l'output sera le F_a et F_v .

La section suivante a pour objectif de présenter les résultats obtenus et de mener des études de sensibilité pour identifier la performance relative de plusieurs proxies de site, en vue de proposer des ensembles de combinaisons optimales offrant un bon compromis entre la pertinence physique et la pratique.

5.4 Résultats

5.4.1 Nomenclature des résultats

Le premier objectif à atteindre est de tester la capacité de cette approche en utilisant seulement un nombre limité de sites pour prédire de manière satisfaisante les facteurs d'amplification réels.

Dans ce cadre, nous avons dérivé un grand nombre de modèles GRNN utilisant toutes les combinaisons possibles de paramètres d'entrée et analysé l'évolution de l'écart-type des résidus («prédit» - «valeurs réelles»), en la comparant également aux valeurs initiales de variabilité pour chaque période ($\sigma_0(\theta, T_i)$), et la variabilité globale $\sigma_{t0}(\theta)$ définie dans les chapitres précédents (Equation (3-42)-(3-43)).

Ces modèles GRNN ont été dérivés pour chacun des deux ensembles des 858 facteurs d'amplification (c'est-à-dire sur l'axe de fréquence réelle et normalisée), en considérant les divers ensembles de variables explicatives énumérés dans le tableau (Tableau 16). Un total de 186 modèles GRNN ont été obtenu (chacun pour toutes les fréquences / périodes considérées, ainsi à courte et longue période pour les facteurs de Borchardt F_a et F_v), leurs performances sont analysés de manière systématique à travers les valeurs des résidus, le coefficient de corrélation R^2 , ainsi

les tendances physiques. L'output s'écrit donc $AF(P_k, \theta, T_i, x_1 \dots x_R)$. Pour des raisons de simplification d'écriture nous l'avons écrit $AF_{GRNN}(P_k, \theta, T_i)$ ou simplement $AF(T)$ au lieu de $AF(P_k, \theta, T_i, x_1 \dots x_R)$ (voir *Tableau 16*).

Tableau 16 Listes des modèles GRNN considérés.

GRNN Output	Profil de sol	Paramètres de site	Nombre de Combinaison
$AF(T), Fa, Fv$	Réel	06 Depth, f_0 , Cv $V_{sm}, V_{s30}, V_{bedrock}$	06 pour 01 paramètre 15 pour 02 paramètres 20 pour 03 paramètres 15 pour 04 paramètres 06 pour 05 paramètres 01 pour 06 paramètres
$AF(f/f_0)$	Réel	5 Depth, Cv, V_{sm} $V_{s30}, V_{bedrock}$	05 pour 01 paramètre 10 pour 02 paramètres 10 pour 03 paramètres 05 pour 04 paramètres 01 pour 05 paramètres
$AF(T), Fa, Fv$	Normalisé, Tronqué	5 Depth, f_0, Cv V_{sm}, V_{s30}	05 pour 01 paramètre 10 pour 02 paramètres 10 pour 03 paramètres 05 pour 04 paramètres 01 pour 05 paramètres
$AF(f/f_0)$	Normalisé, Tronqué	4 Depth, Cv V_{sm}, V_{s30}	04 pour 01 paramètre 06 pour 02 paramètres 04 pour 03 paramètres 01 pour 04 paramètres

Avant de discuter de la performance des différents proxies de site, nous comparons dans la figure (Figure 5-5), l'AF prédit par quelques modèles GRNN au AF réelle pour deux profils de sol (*Tableau 17*):

-Sol 1: Les paramètres du site appartiennent à la base de données.

-Sol 2: Paramètres de site n'appartiennent pas à la base de données.

Le facteur d'amplification de la base de données s'écrit $AF_m(P_k, 1, T_i)$.

Les profils de sol Sol1 ou Sol2 n'appartiennent pas à la base données c'est à dire qu'ils ne correspondent pas à l'un des $P_k, k = 1, \dots, 858$ profils mais les paramètres de site de Sol1 ont des valeurs qui appartiennent à la base de données. Les facteurs d'amplification seront notés $AF_m(Sol1, 1, T_i)$ et $AF_m(Sol2, 1, T_i)$.

Ils sont obtenus par l'analyse 1D viscoélastique avec notre code et seront appelés Target Output. Le facteur d'amplification obtenu par le GRNN sera note GRNN output (Combinaison des paramètres de sites). Par exemple si l'on introduit tous les paramètres de site alors le facteur d'amplification obtenu par la GRNN sera GRNN output (All parameters) que l'on notera $AF_{GRNN}(P_k, \theta, T_i)$. Comme on a 186 combinaisons possibles on va ci-après montrer que quelques cas.

Tableau 17 Paramètres de site pour Sol 1 et Sol 2

Sol	Depth (m)	f_0 (Hz)	V_{sm} (m/s)	C_v	V_{s30} (m/s)	$V_{bedrock}$ (m/s)
Sol 1	72	3.69	603	12.33	333	1850
Sol 2	163	1.44	746	8.33	472	1000

L'examen des figures (Figure 5-5 et Figure 5-6) montrent clairement que le AF prédit (GRNN output) est évidemment identique à celui calculé (Target output) pour les deux sols. On remarque que lorsque le nombre de paramètres de site utilisé diminue alors les deux AF ne sont pas totalement identiques, ils sont même largement différents pour certaines périodes de la structures.

De toute évidence, les différences entre les prédictions et les facteurs d'amplification réels varient d'un profil de sol à l'autre. C'est pourquoi il est important d'analyser l'écart-type des résidus pour obtenir un aperçu statistiquement significatif des performances relatives de chacun des proxies de site considérés qui contrôlent nos facteurs d'amplification.

La figure (Figure 5-7) donne la variation des résultats simulés par la méthode classique (propagation d'onde modèle linéaire viscoélastique) en fonction des résultats prédits par le GRNN application pour une combinaison constituée de cinq paramètres de site ($V_{s30}, f_0, C_v, V_{sm}$ et Depth).

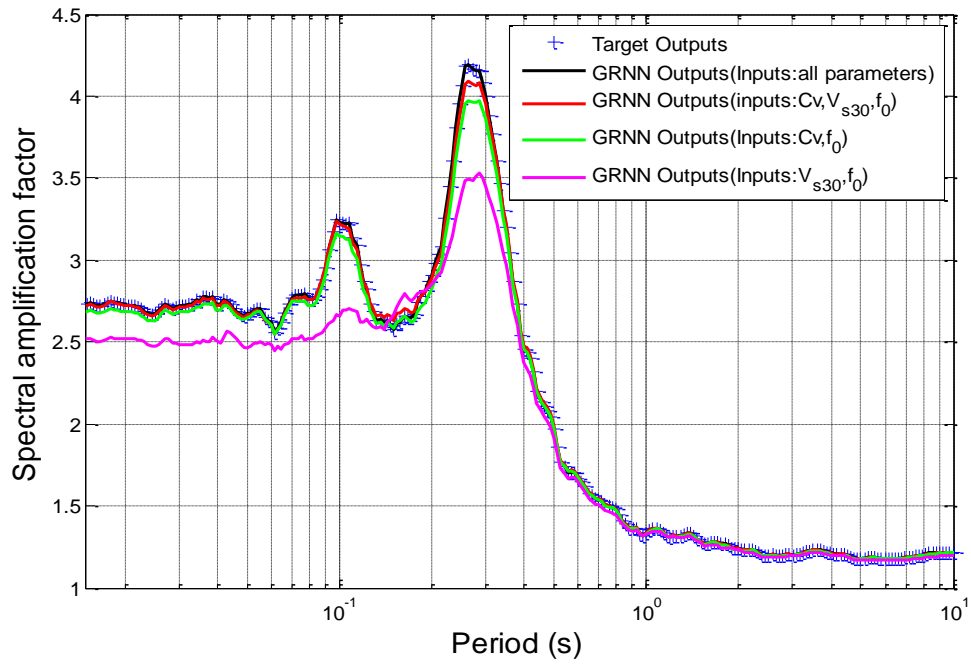


Figure 5-5: Comparaison entre $AF_m(Sol1,1,T_i)$ et les prédictions du GRNN pour le Sol1.

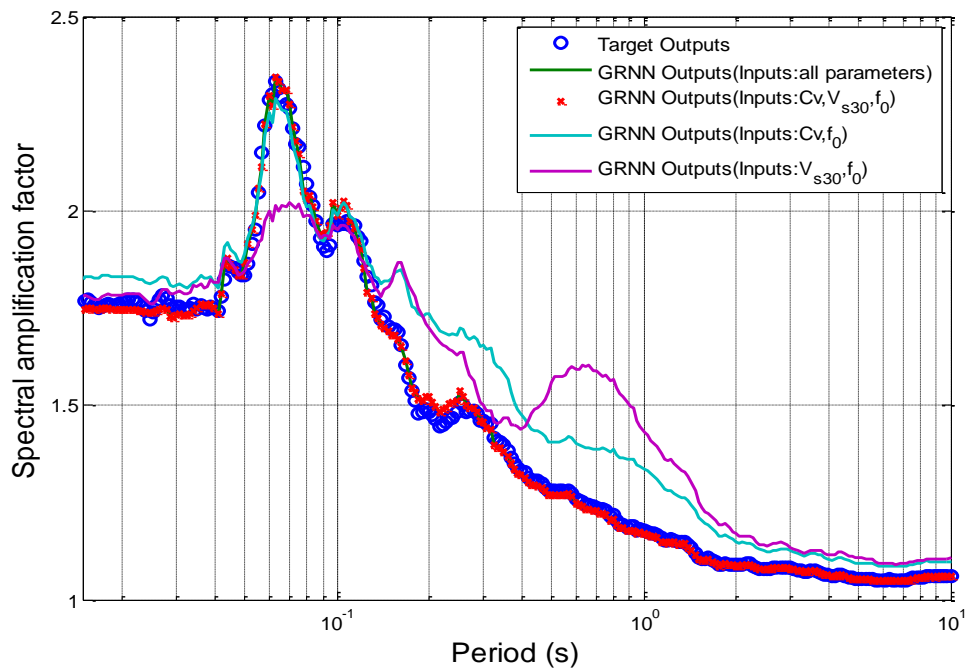


Figure 5-6: Comparaison entre $AF_m(Sol2,1,T_i)$ et les prédictions du GRNN pour le Sol2.

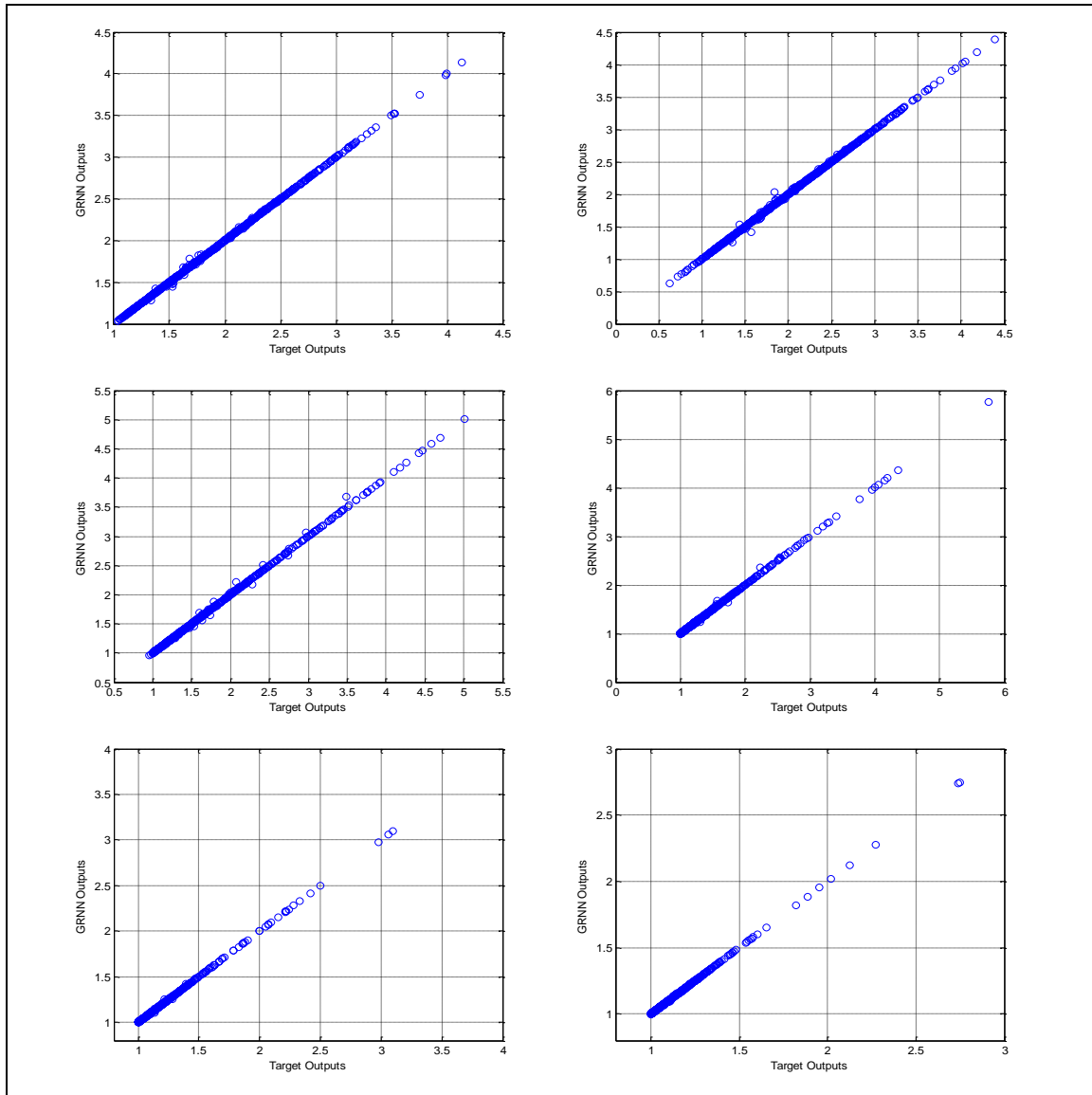


Figure 5-7 : Variation des résultats calculés par la méthode classique (propagation d’onde) (Target Outputs) en fonction des résultats prédit par la méthode GRNN à (0.01s, 0.1s, 0.5s, 1s, 2s et 4s) pour les 858 profils application pour (RP).

D’après la courbe ci-dessus on peut conclure que la GRNN fonctionne de manière fiable, vu que les valeurs simulées par la GRNN et par la méthode classique (linéaire viscoélastique) sont du même ordre de grandeur. Ainsi on obtient un facteur de corrélation supérieur à 0.9. En outre ces courbes illustrent la stationnarité du modèle puisque tous les points appartiennent à la même bissectrice de tous les 858 profils de sols pris pour l’apprentissage à différentes périodes.

La comparaison des facteurs d'amplifications Targets et GRNN outputs telle que présentées ci-dessous ne fournit qu'une description visuelle des différences qui peuvent exister, il serait plus intéressant d'effectuer cette comparaison en utilisant la notion d'écart type. La section suivante est dédiée à cette fin.

5.4.2 Dérivation et analyse du terme d'erreur

L'erreur entre la prédiction et les valeurs réelles a été estimée avec les quantités suivantes, en comparaison avec les variabilités initiales des facteurs d'amplification tels que définis par les équations ((3-42) et (3-43)).

- Pour chaque période et chaque modèle GRNN, une erreur dépendante de la période (appelée Ecart type du résidu ou standard deviation) peut être définie en relation avec le terme de variabilité initiale (Equivalent au $\sigma_0(\theta, T_i)$, Equation(3-42)). La notation de cette erreur doit s'écrire $\varepsilon(\theta, T_i, x_1 \dots x_R)$. Toutefois pour ne pas alourdir le texte elle sera écrite dès fois $\varepsilon_{GRNN}(\theta, T_i)$ et on indiquera les paramètres qui seront introduits

$$\varepsilon(\theta, T_i, x_1 \dots x_R) = \sqrt{\frac{1}{n_p} \sum_{k=1}^{n_p} [\log(AF(P_k, \theta, T_i, x_1 \dots x_R)) - \log(AF_m(P_k, \theta, T_i))]^2}$$
(5-2)

- De même, en relation avec la variabilité initiale maximale $\sigma_{0Max}(\theta)$, une "erreur maximale" est définie comme étant le maximum sur toutes les périodes / fréquences de ce terme d'erreur dépendante de la période

$$\varepsilon_{max}(\theta, x_1 \dots x_R) = \max[\varepsilon(\theta, T_i, x_1 \dots x_R)]$$
(5-3)

- Enfin, en relation avec le terme global de la variabilité initiale $\sigma_{t0}(\theta)$, (Equation (3-44)), une erreur globale est définie comme la moyenne sur toutes les erreurs à chaque période.

$$\varepsilon_{t,GRNN} \equiv \varepsilon_m(\theta, x_1 \dots x_R) = \frac{1}{n_p} \sum_{i=1}^{n_p} \varepsilon(\theta, T_i, x_1 \dots x_R)$$
(5-4)

Des exemples d'erreur dépendante de la période $\varepsilon(\theta, T_i, x_1 \dots x_R)$ sont affichés dans les figures (Figure 5-8 , Figure 5-9 et Figure 5-10) respectivement pour les profils (RP, NP et TP) en fonction de la période réelle et les figures (Figure 5-11, Figure 5-12 et Figure 5-13) pour les mêmes type de profils en fonction de la fréquence normalisée. Au niveau de ces figures on a aussi tracé les variabilités initiales (initial standard déviation) des ensembles de facteurs d'amplification c'est-à-dire $\sigma_0(\theta, T_i)$ et $\sigma_0\left(\theta, \frac{f_i}{f_0}\right)$.

Par exemple la courbe qui correspond à all Parameters c'est la fonction $\varepsilon(\theta, T_i, Depth, V_{sm}, V_{s30}, C_v, f_0, V_{bedrock})$.

La courbe qui correspond à Input : f_0, V_{s30}, C_v c'est la fonction $\varepsilon(\theta, T_i, f_0, V_{s30}, C_v)$. Ces deux fonctions peuvent être écrites $\varepsilon_{GRNN}(\theta, T_i)$.

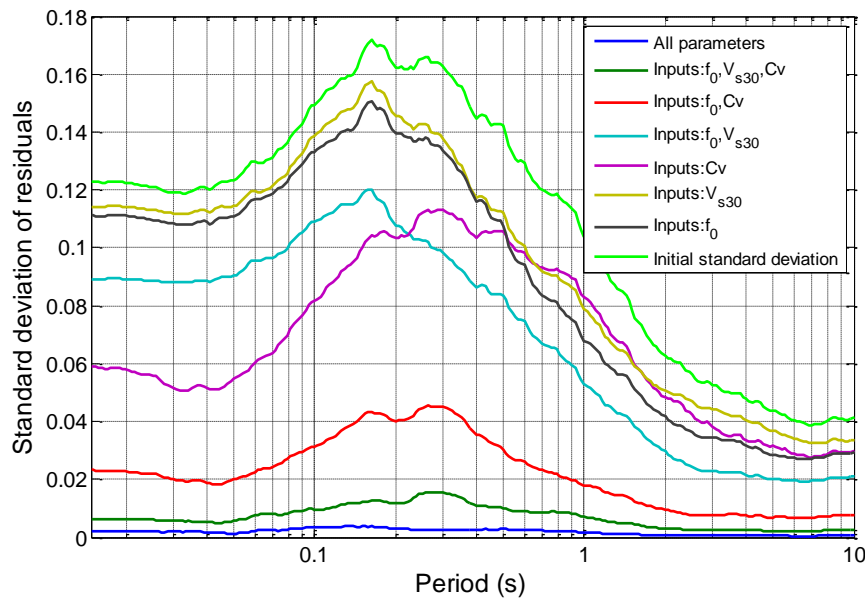


Figure 5-8 : Variation de $\varepsilon_{GRNN}(\theta = 1, T_i)$ pour différentes combinaisons de paramètres de site et de $\sigma_0(\theta = 1, T_i)$

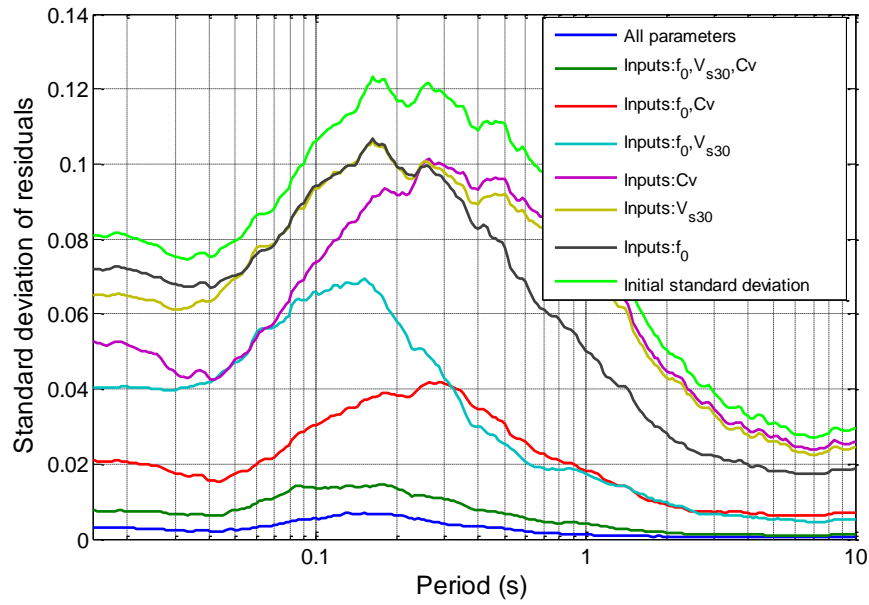


Figure 5-9 : Variation de $\varepsilon_{GRNN}(\theta = 0, T_i)$ pour différentes combinaisons de paramètres de site et de $\sigma_0(\theta = 0, T_i)$.

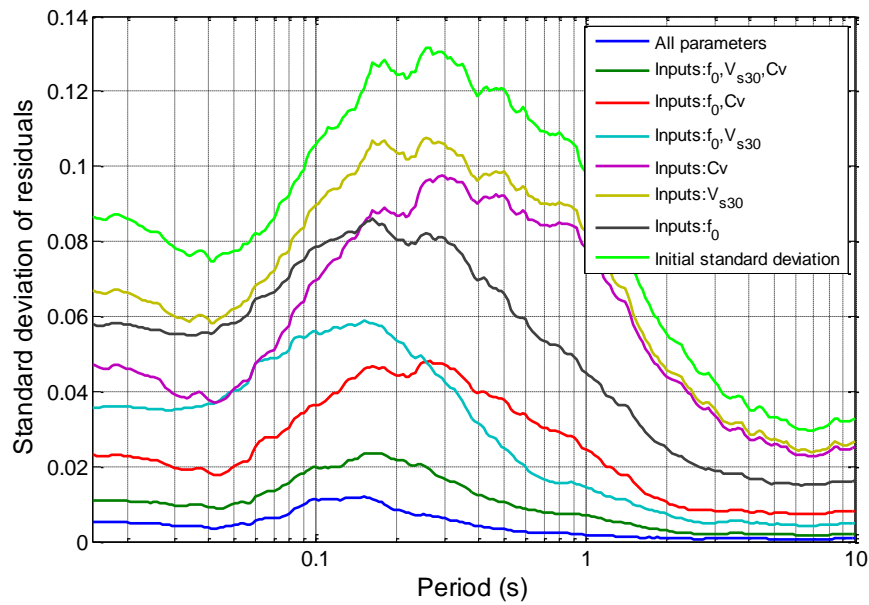


Figure 5-10 : Variation de $\varepsilon_{GRNN}(\theta = 2, T_i)$ pour différentes combinaisons de paramètres de site et de $\sigma_0(\theta = 2, T_i)$.

Dans la deuxième étude, en considérons des fréquences normalisées, cependant le paramètre "f₀" a été remplacé par le paramètre "Depth", qui est totalement indépendant des paramètres de vitesse.

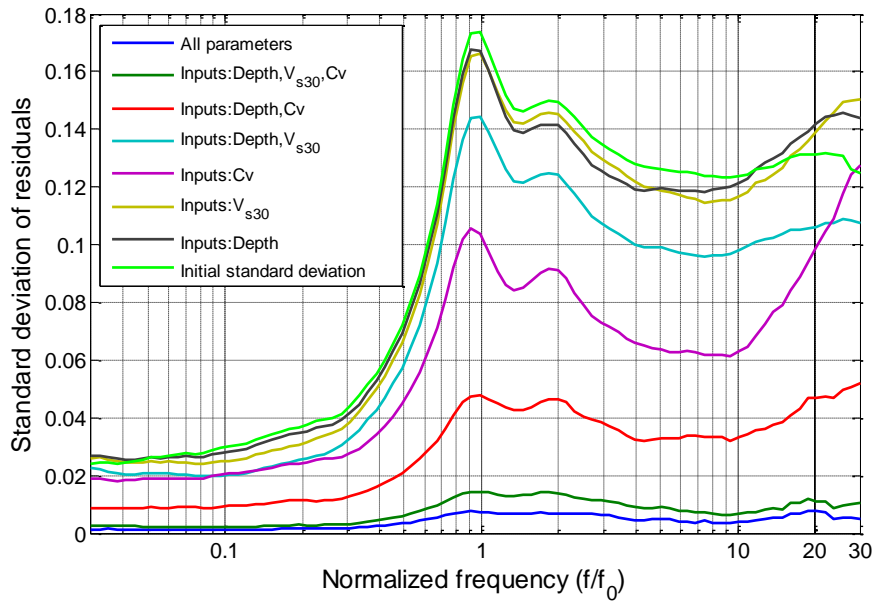


Figure 5-11 : Variation de $\epsilon_{GRNN}(\theta = 1, f_i/f_0)$ pour différentes combinaisons de paramètres de site et de $\sigma_0(\theta = 1, f_i/f_0)$.

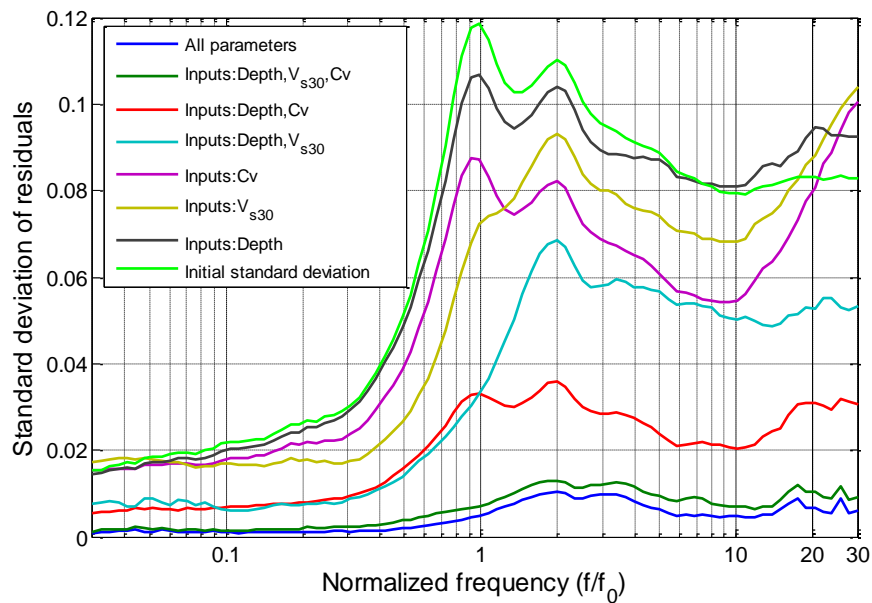


Figure 5-12 : Variation de $\epsilon_{GRNN}(\theta = 0, f_i/f_0)$ pour différentes combinaisons de paramètres de site et de $\sigma_0(\theta = 0, f_i/f_0)$

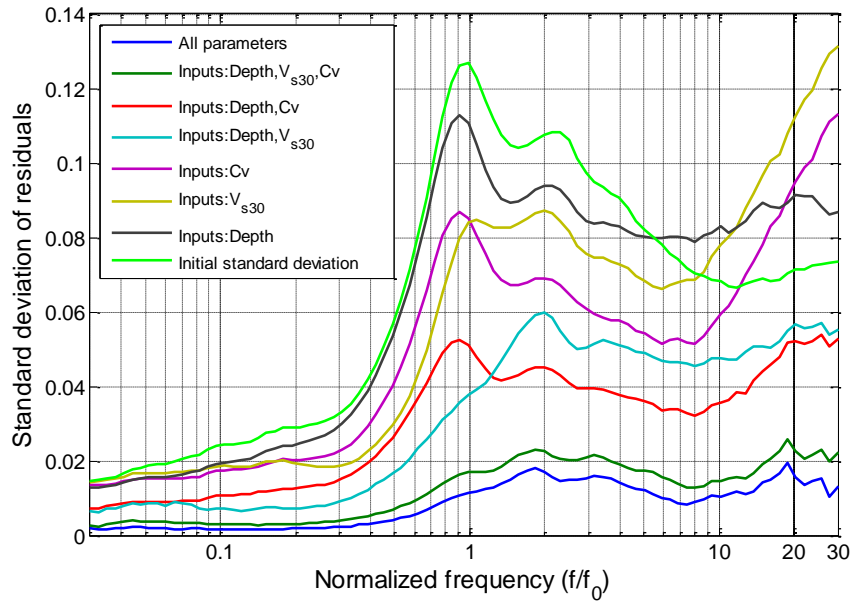


Figure 5-13 : Variation de $\varepsilon_{\text{GRNN}}(\theta = 2, f_i/f_0)$ pour différentes combinaisons de paramètres de site et de $\sigma_0(\theta = 2, f_i/f_0)$

L'examen des figures (Figure 5-8 à Figure 5-13) permet de dresser les conclusions suivantes :

- Pour l'ensemble des cas présentés, la variabilité initiale ($\sigma_0(\theta, T_i)$, où $\sigma_0(\theta, \frac{f_i}{f_0})$) est plus importante que la variation de n'importe quel modèle GRNN même à un seul paramètre de site en entrés.
- Le paramètre de site Cv seul permet une explication significative de la AF (c'est-à-dire que $\varepsilon_{\text{GRNN}}(\theta, T_i)$ est significativement plus petit que $\sigma_0(\theta, T_i)$), parfois même mieux que si l'on considère deux autres proxies de site tels que (f_0, V_{s30}). En revanche, ce résultat n'est toutefois pas valable pour les jeux de profils NP et TP, en raison de l'uniformité de la vitesse du substratum rocheux à 800 m/s, qui réduit considérablement l'importance relative du contraste de vitesse (Cv) par rapport à V_{s30} .
- AF prédit par GRNN avec 3 paramètres (Cv, f_0, V_{s30}) est très puissant pour prédire l'AF réel, tandis que le «de proxies avec tous les paramètres» est plus efficace: $\varepsilon_{\text{GRNN}}(\theta, T_i)$ est inférieur à 5% du

terme initial $\sigma_0(\theta, T_i)$ et ce résultat est valable pour les trois jeux de données

Ces résidus sont maximaux dans la plage, courte à intermédiaire période (réelle) et autour de la fréquence fondamentale f_0 en considérons la fréquence normalisée (Figure 5-11, Figure 5-12 et Figure 5-13) : ces maxima correspondent à la plage de fréquence de la plus grande variable initiale.

- Le paramètre V_{s30} qui est largement utilisé est plus efficace pour quand elle est associée avec la fréquence fondamentale (f_0) et quand la vitesse du substratum rocheux est uniforme (Figure 5-9 et Figure 5-10) c'est-à-dire pour les profils (NP et TP). Mais pour tous les autres cas il est nettement moins performant que les paramètres C_v ou f_0 .

Cependant, ces résultats ne sont que partiels puisque seulement 6 des nombreux modèles possibles sont considérés.

La figure (Figure 5-14) affiche l'évolution de l'erreur globale de prédiction $\epsilon_{t,GRNN}$ en fonction du nombre de proxies (ou paramètre de site). Bien qu'il diminue nettement avec un nombre croissant de proxies de site explicatif, il présente également une dispersion significative pour un nombre donné de proxies de site explicatif : cela indique que certains proxies de site ont un meilleur rendement que d'autres dans le contrôle du facteur d'amplification.

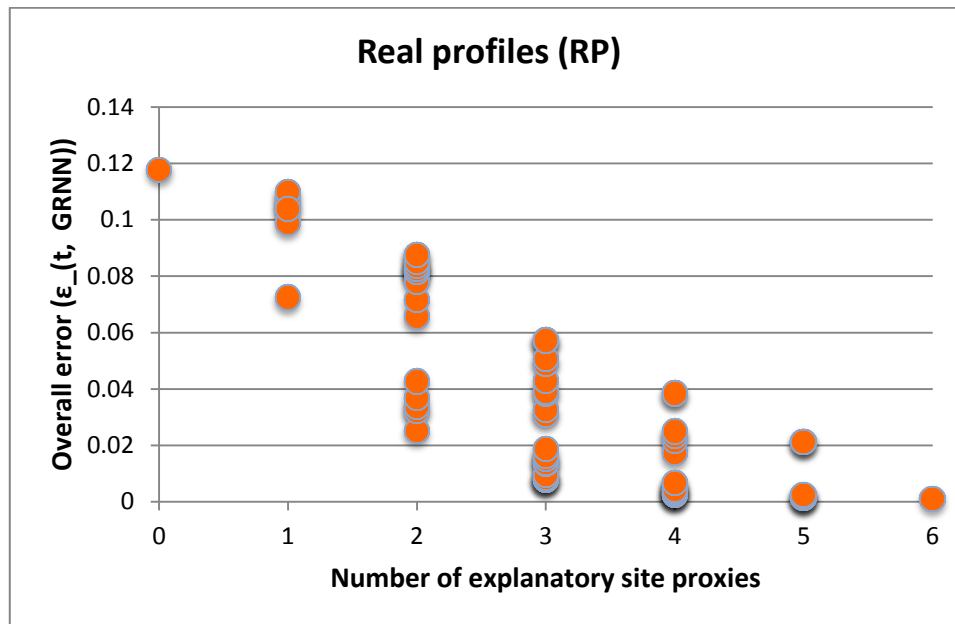


Figure 5-14: Illustration de la réduction progressive de l'écart type des résidus entre les modèles GRNN et le AF réelle avec le nombre de proxies de sites considérées pour les profils (RP).

En annexe, une illustration de la réduction progressive de l'écart type des résidus entre les modèles GRNN et le AF réelle pour les jeux de données (NP et TP) est donnée.

Une autre façon de voir, consiste à déterminer de la distribution et la variation des erreurs (résidus) en fonction des paramètres de sites. La figure (Figure 5-15) illustre ce concept où l'erreur est calculé pour les 858 profils de sol en fonction des paramètres pris en entrés. À cet effet le modèle GRNN considéré est construit à partir de seulement 5 paramètres décrit par (V_{s30} , Depth, f_0 , V_{sm} et le contraste de vitesse (Cv)).

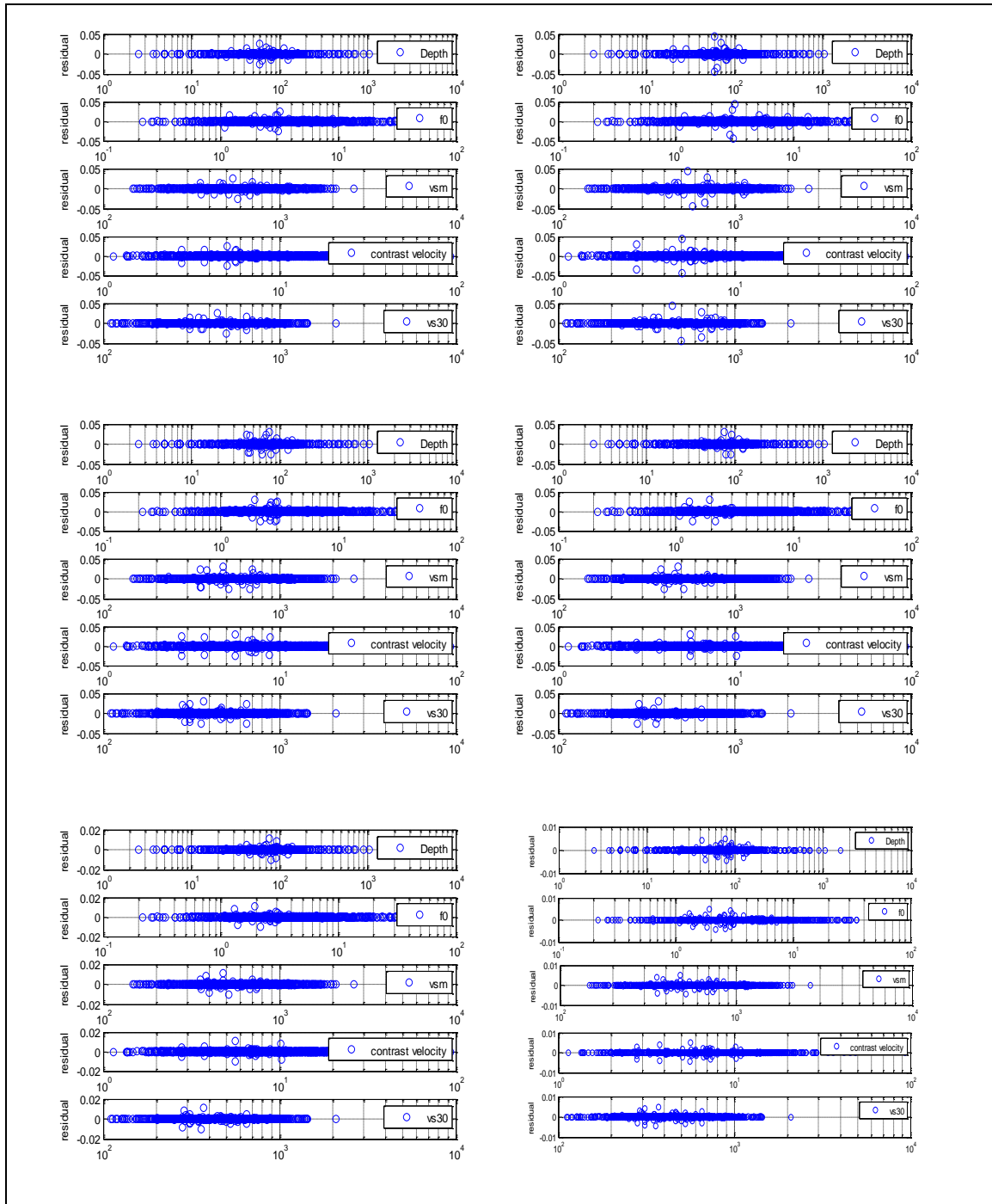


Figure 5-15: Valeurs résiduelles (différence entre le log10 de l'accélération spectrale observée et celle prédite) à 0.01, 0.1s, 0.5s, 1s, 2s et 4s en ordonné. Les résidus sont en fonction de la profondeur total (Depth), f_0 , V_{sm} , contraste de vitesse (Cv) et V_{s30} pour tous les profils.

Les figures (Figure 5-15) montrent d'une façon générale que l'incertitude est faible ainsi que l'absence de biais. Elles montrent aussi que l'ensemble de ces résidus sont compris entre -0.05 et +0.05 au maximum.

Le nombre de combinaison est très important et les résultats obtenus ne sont valables que pour quelques combinaisons. La section suivante est consacrée à la mise en place d'une stratégie pour prendre en compte l'ensemble des combinaisons

5.4.3 Approche globale

En considérant un grand nombre de combinaisons possibles (indiqué à titre d'exemple sur le tableau ci-dessous), nous avons analysé les performances respectives de chaque proxy en évaluant, pour un nombre donné de proxy de site, la valeur moyenne de $\varepsilon_{t,GRNN}$ pour toutes les combinaisons afin de déterminer le paramètre le plus pertinent.

A cet effet on a introduit la notion de réduction de la variance. Considérons par exemple que pour notre input que deux paramètres de site : x_1 et un autre paramètre de site. Les combinaisons qui incluent x_1 sont égales à 5 : $\{(x_1, x_2); \dots (x_1, x_6)\}$. Par conséquent les standard deviation sont décrits comme suit: $\{\varepsilon_m(\theta, x_1, x_2); \dots; \varepsilon_m(\theta, x_1, x_6)\}$. Donc on peut calculer la valeur moyenne $\varphi_m(\theta, x_1, 2)$ qui permet d'obtenir la variabilité du paramètre x_1 lorsqu'on considère uniquement deux paramètres de site comme entrée.

$$\varphi_m(\theta, x_1, 2) = \frac{1}{5} \sum_{l=2}^{l=6} \varepsilon_m(\theta, x_1, x_l) \quad (5-5)$$

Cette valeur est ensuite comparée à l'erreur initiale à travers la relation suivante qui donne une appréciation sur la réduction de l'écart type:

$$RS_m(\theta, x_1, 2) = 1 - \frac{\varphi_m(\theta, x_1, 2)}{\sigma_{t0}(\theta)} \quad (5-6)$$

Dans le cas où on s'intéresse à la performance du paramètre x_1 lorsqu'on considère uniquement trois paramètres de site comme entrée alors cette réduction s'écrit maintenant :

$$RS_m(\theta, x_1, 3) = 1 - \frac{\varphi_m(\theta, x_1, 3)}{\sigma_{t0}(\theta)} \quad (5-7)$$

Dans le cas général cette réduction s'écrit :

$$RS_m(\theta, x_1, NPR) = 1 - \frac{\varphi_m(\theta, x_1, NPR)}{\sigma_{t0}(\theta)} \quad (5-8)$$

NPR est le nombre de paramètre de site à considérer $\{NPR = 1, \dots, R\}$

Aussi il y a une autre entité utilisé qui définit par la réduction de variance, son équation peut être exprimé comme suite :

$$RV_m(\theta, x_1, NPR) = 1 - \left(\frac{\varphi_m(\theta, x_1, NPR)}{\sigma_{t0}(\theta)} \right)^2 \quad (5-9)$$

Tableau 18 Écart type des résidus de modèle pour différents modèles GRNN impliquant les facteurs d'amplification pour (RP, NP et TP), application à différents combinaisons de paramètres de site.

Number of parameters	Description of the parameters	Standard deviation (RP)	Variance reduction (RP)	Standard deviation (NP)	Variance reduction (NP)	Standard deviation (TP)	Variance reduction (TP)
All (6) parameters	Depth + f_0 + V_{sm} + C_v + V_{s30} + $V_{bedrock}$	0.0011	99.99%	0.0032	96.21%	0.0053	94.08%
Three parameters (best triplet)	f_0 + C_v + V_{s30}	0.0079	99.55%	0.0079	90.66%	0.0118	86.83%
Two parameters (best pair)	f_0 + C_v	0.0251	95.46%	0.0233	72.45%	0.0278	68.97%
Two parameters (convenient pair)	f_0 + V_{s30}	0.0782	55.9%	0.0382	54.84%	0.0339	62.16%
One parameter (best)	C_v	0.0725	62.1%	0.0652	22.93%	0.0622	30.58%
One parameter (usual)	V_{s30}	0.1038	22.4%	0.0715	15.84%	0.0733	18.19%
One parameter	f_0	0.099	29.4%	0.0678	19.85%	0.0563	37.16%
Overall Initial variability term $\sigma_m(\theta)$		0.1178		0.0846		0.0896	

Toutes les possibilités des combinaisons en termes de proxies de site ont été considérées, la figure (Figure 5-16) illustre ce concept.

L'examen attentif de la figure (Figure 5-16) montre qu'il existe un paramètre qui affecte mieux le facteur d'amplification qui est décrit par le contraste de vitesse C_v . Il est facilement compréhensible que le contraste de vitesse domine le contraste d'impédance qui contrôle l'amplification réelle (pour le cas simple, à couche unique). Tous les autres paramètres sont pratiquement identiques, avec toutefois une performance légèrement meilleure pour la fréquence fondamentale, et un peu plus mauvais pour les paramètres «épaisseur totale» 'Depth' et ' V_{sm} '. En ce qui concerne le proxy V_{s30} qui est largement utilisé, il s'avère qu'il n'est pas aussi fiable que la fréquence fondamentale f_0 mais lorsque ils sont associés tous les deux donnent des résultats meilleurs par rapport au contraste de vitesse tout seul même parfois meilleur ou similaire que le couple (f_0, C_v) surtout pour les profils type (NP ou TP) (voir Figure 5-17 et Figure 5-18).

Donc, Il serait souhaitable autant que possible, de mesurer pour tout profil qu'on veut étudier, le contraste de vitesse entre la profondeur et la surface. Malheureusement, ces mesures sont plutôt rares et /ou coûteuses, et ce proxy de site "optimal" n'est presque jamais disponible.

Donc on se pose la question suivante : quel est le «second choix» optimal, étant donné que C_v n'est pas disponible ?

Un examen attentif des tableaux d'erreurs indique que la paire (f_0, V_{s30}) fournit des erreurs de prédiction presque similaires à C_v seul pour les profils (RP) et beaucoup plus meilleur que C_v seul pour les profils (NP ou TP).

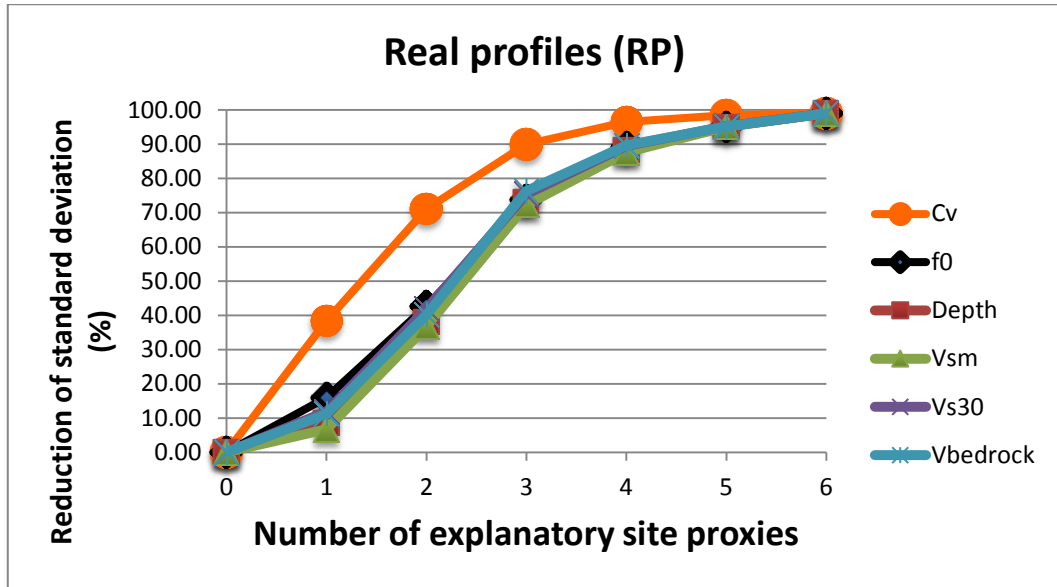


Figure 5-16: Indicateurs de performance pour les différents proxies de site en terme (reduction of standard deviation) pour le cas des profils réels (RP) en période.

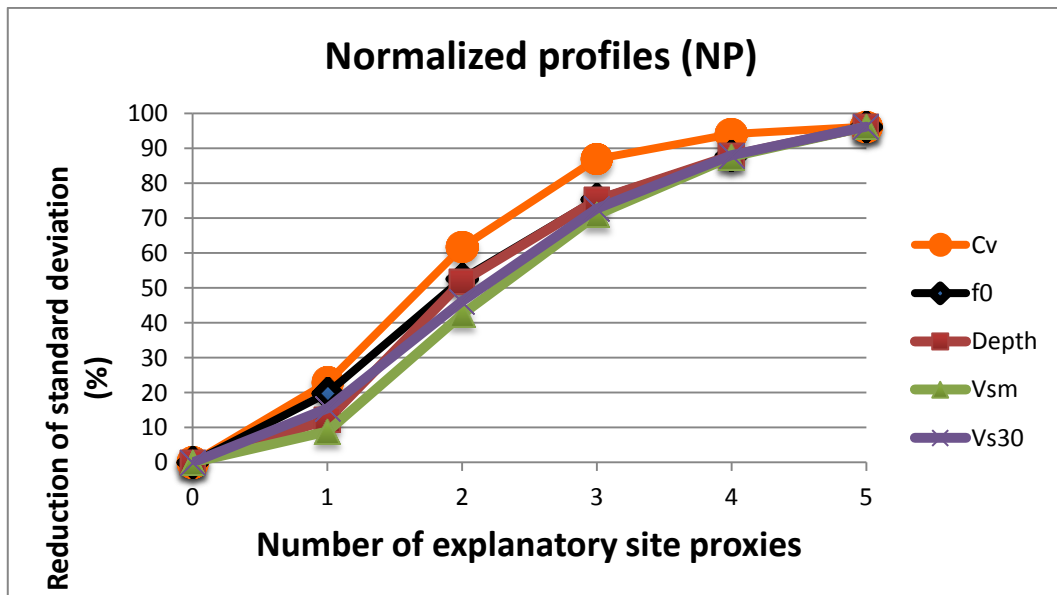


Figure 5-17: Indicateurs de performance pour les différents proxies de site en terme (reduction of standard deviation) pour le cas des profils de sol normalisé(NP) en période.

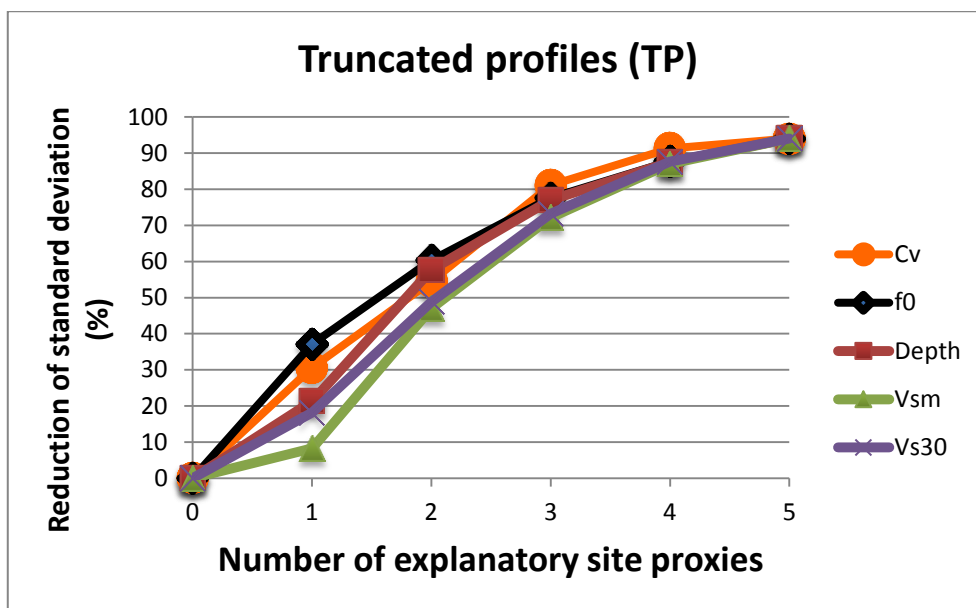


Figure 5-18: Indicateurs de performance pour les différents proxys de site en terme (reduction of standard deviation) pour le cas des profils de sol tronqué (TP) en période.

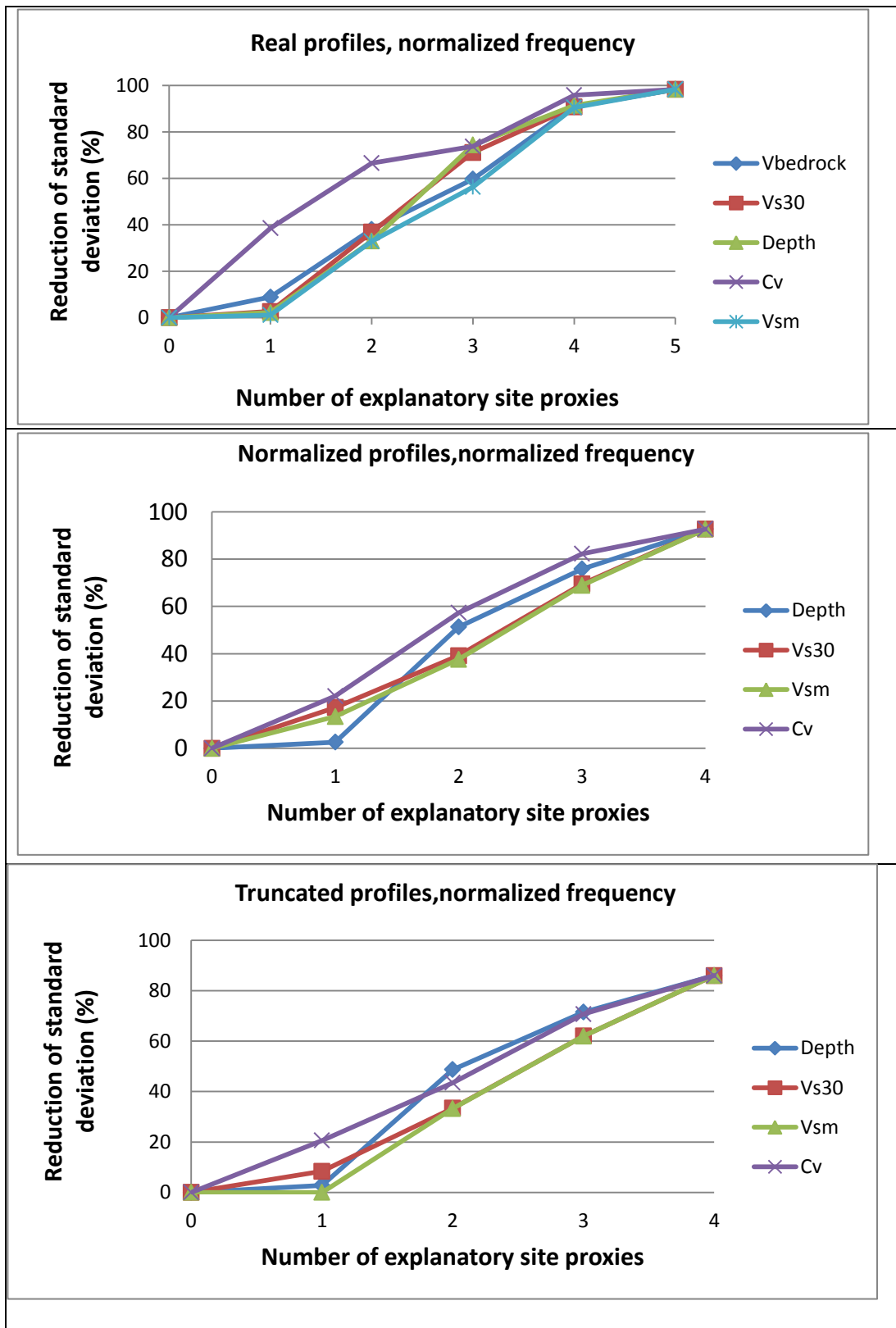


Figure 5-19: Indicateurs de performance pour les différents proxies de site en terme (reduction of standard deviation) pour le cas des profils (RP, NP et TP) en fréquence normalisée.

Enfin, un autre résultat intéressant concerne l'utilité - ou non - de la normalisation de fréquence par f_0 . Une comparaison entre les figures (Figure 5-16 et Figure 5-19) ou bien (Figure 5-17 et Figure 5-19) indique clairement que l'erreur de prédiction est légèrement réduite en considérant f_0 directement comme une entrée plutôt que simplement en normalisant l'axe de fréquence. Par exemple, la réduction sigma est de 79% avec la paire de paramètres (C_v , f_0) et 93% pour le triplet de paramètres (C_v , f_0 , V_{s30}) pour le cas de RP en période, alors qu'elle n'est que de 38% avec le paramètre f_0 et 68% pour la paire de paramètres (C_v , V_{s30}) pour le cas Profil réel (RP) normalisé en fréquence. Cette dernière procédure est cependant simple, et permet d'avoir une formule de prévision moins complexe.

5.4.4 Variation des facteurs de Borchardt en utilisant GRNN

Comme indiqué ci-dessus, les effets du site peuvent être simplement caractérisés par les deux facteurs de Borchardt F_a et F_v , en particulier dans une perspective réglementaire. Nous avons donc calculé les facteurs de Borchardt pour le modèle GRNN en fonction de la paire de proxies de site (f_0 , V_{s30}) qui s'avère plutôt efficace. Les figures (Figure 5-20 et Figure 5-21) montrent la dépendance de ces deux facteurs en fonction de V_{s30} et f_0 respectivement pour les cas (RP, NP et TP). Dans chaque cas, cette dépendance est considérée sur la plage de 5% à 95% de chaque paramètre explicatif considéré, c'est-à-dire [0.8, 14 Hz] pour f_0 , [200, 1000 m/s] pour V_{s30} cas profils (RP) et [150, 550 m/s] pour V_{s30} cas profils (NP et TP).

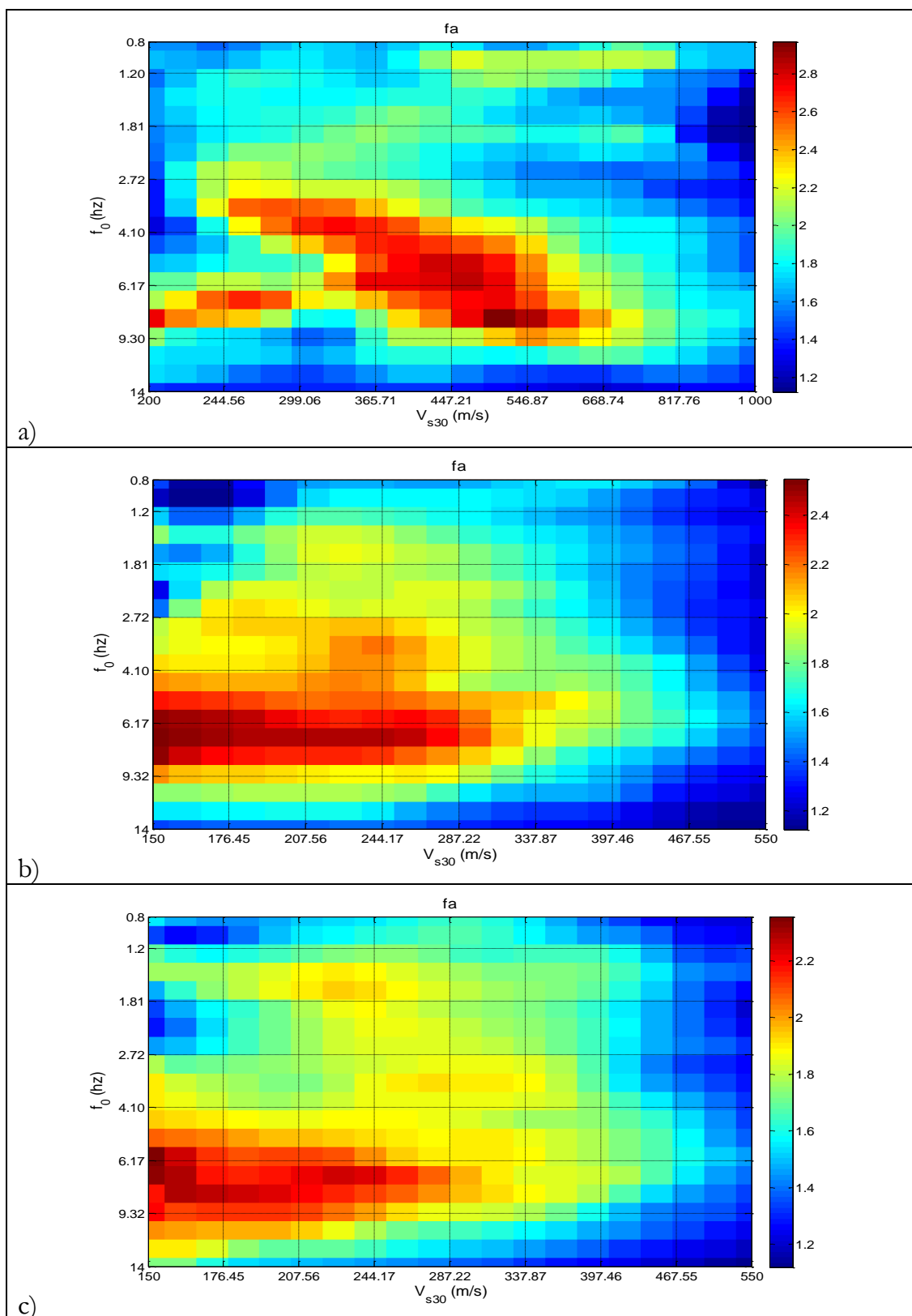


Figure 5-20: Variation des facteurs de Borcherdt à courte période (F_a) avec (f_0 , V_{s30}) pour les profils de sol ((a) RP, (b) NP et (c) TP).

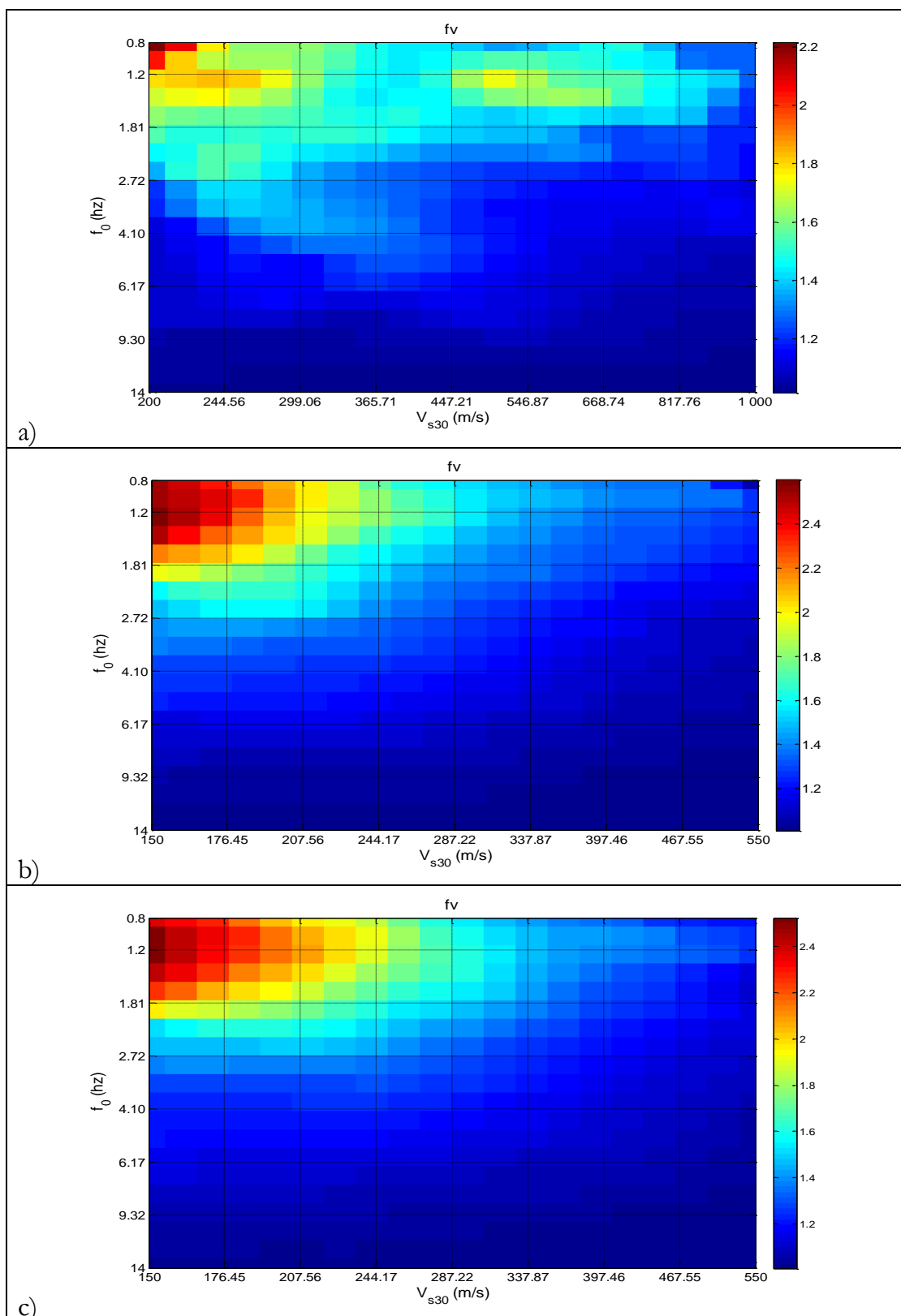


Figure 5-21: Variation des facteurs de Borcherdt à longue période (F_v) avec (f_0, V_{s30}) pour les profils de sol ((a) RP , (b) NP et (c) TP).

Les dépendances correspondantes sont exprimées avec l'équation explicite suivante associée aux modèles GRNN:

$$Fa = 10^{\left(\{W_3\}^T * \{H\} / \sum H_i\right)} \quad (5-10)$$

$$Fv = 10^{\left(\{W_4\}^T * \{H\} / \sum H_i\right)} \quad (5-11)$$

$$\{H\} = \exp\left(-\left(\sqrt{\left(\{W1\} - \log_{10}(f0)\right)^2 + \left(\{W2\} - \log_{10}(V_{s30})\right)^2}\right) \cdot 16.6511^2\right) \quad (5-12)$$

Avec

$$\{H\} = \begin{bmatrix} H_1 \\ \vdots \\ H_i \\ \vdots \\ H_{\text{final}} \end{bmatrix} \quad (5-13)$$

Où $\{W1\}$ et $\{W2\}$ sont des vecteurs de taille 858 constitués des poids entre la couche d'entrée et la couche cachée des 858 profils d'échantillons, alors que $\{W3\}$ et $\{W4\}$ sont les pondérations associées aux connexions entre les couches cachées et la couche de sortie. Un fichier Excel supplémentaire est fourni en annexes pour une utilisation pratique de ces équations.

Le facteur d'amplification à courte période Fa se révèle atteindre des valeurs plus élevées, essentiellement pour les sites à haute fréquence jusqu'à intermédiaire et à des vitesses faibles en surface.

A partir des figures (Figure 5-20 et Figure 5-21) on note que les valeurs maximales sont supérieures à 2,5 dans les cas RP, NP et TP mais correspondent à des combinaisons légèrement différentes du couple (f_0, V_{s30}) :

- Le premier cas (RP), fa est important correspond à des valeurs de V_{s30} autour de 550 m/s (et également à des fréquences fondamentales autour de 6 à 9 Hz) mais pour les profils (NP) et (TP) correspond à des valeurs de V_{s30} inférieures à 300 m/s.

De telles différences sont liées à la possibilité d'une résonance à grande amplitude quand une couche mince de sol raide (faible caractéristique géométrique) se trouve au-dessus d'une roche très dure, situation assez fréquente dans les profils réels mais impossible dans des profils normalisés ou tronqué, en raison de la réduction de la vitesse imposée par le socle de 800 m/s.

En parallèle, le facteur d'amplification à longue période ou intermédiaire F_v atteint des valeurs les plus élevées (même au-dessus de 2) pour les sites à basse fréquence (f_0 inférieur à 1,5 à 2 Hz), et à faible vitesse: (V_{s30} inférieur à 200 m/s).

A l'inverse pour les profils (RP), F_v reste faible (et souvent inférieur à 1,4) pour les sites à haute fréquence (f_0 au-delà de 4 Hz). Il peut rester significatif pour des valeurs de f_0 (entre 1,4 et 1,6)Hz et pour des sites rigides ou ($V_{s30} > 400$ m/s). Ce cas est plus souvent rencontré lorsque le substrat rocheux est assez profond et assez dur correspondant à des fréquences fondamentale inférieure à 2Hz, alors que pour les profils NP ou TP, F_v est supérieur à 1,4 lorsque V_{s30} dépasse 350m/s.

Quelle relation (RP, NP ou TP) doit être utilisée dans un cas pratique ?

Il faut tout d'abord garder à l'esprit que la présente étude a été réalisée uniquement dans le cadre linéaire (comme une première étape de faisabilité), et cela peut expliquer pourquoi les valeurs de F_v restent relativement limitées et souvent plus petites que les valeurs de F_a . Cependant, afin d'obtenir une estimation du premier ordre des valeurs F_a et F_v pour la réponse linéaire d'un site donné, la première étape consiste à obtenir une idée approximative du substratum sous-jacent: s'il y a des chances qu'il soit un substrat rocheux très dur avec des vitesses d'onde S dépassant 1,2 - 1,5 km/s, il est préférable de choisir une des relations (RP).

En revanche, si la vitesse des ondes S du substrat rocheux est proche d'un substrat rocheux "standard" avec une vitesse d'onde S d'environ 800 m/s (par exemple entre 600 et 1000 m/s) et si la valeur V_{s30} est inférieure à 550 m/s, alors il est probablement préférable de sélectionner la relation (NP) ou (TP).

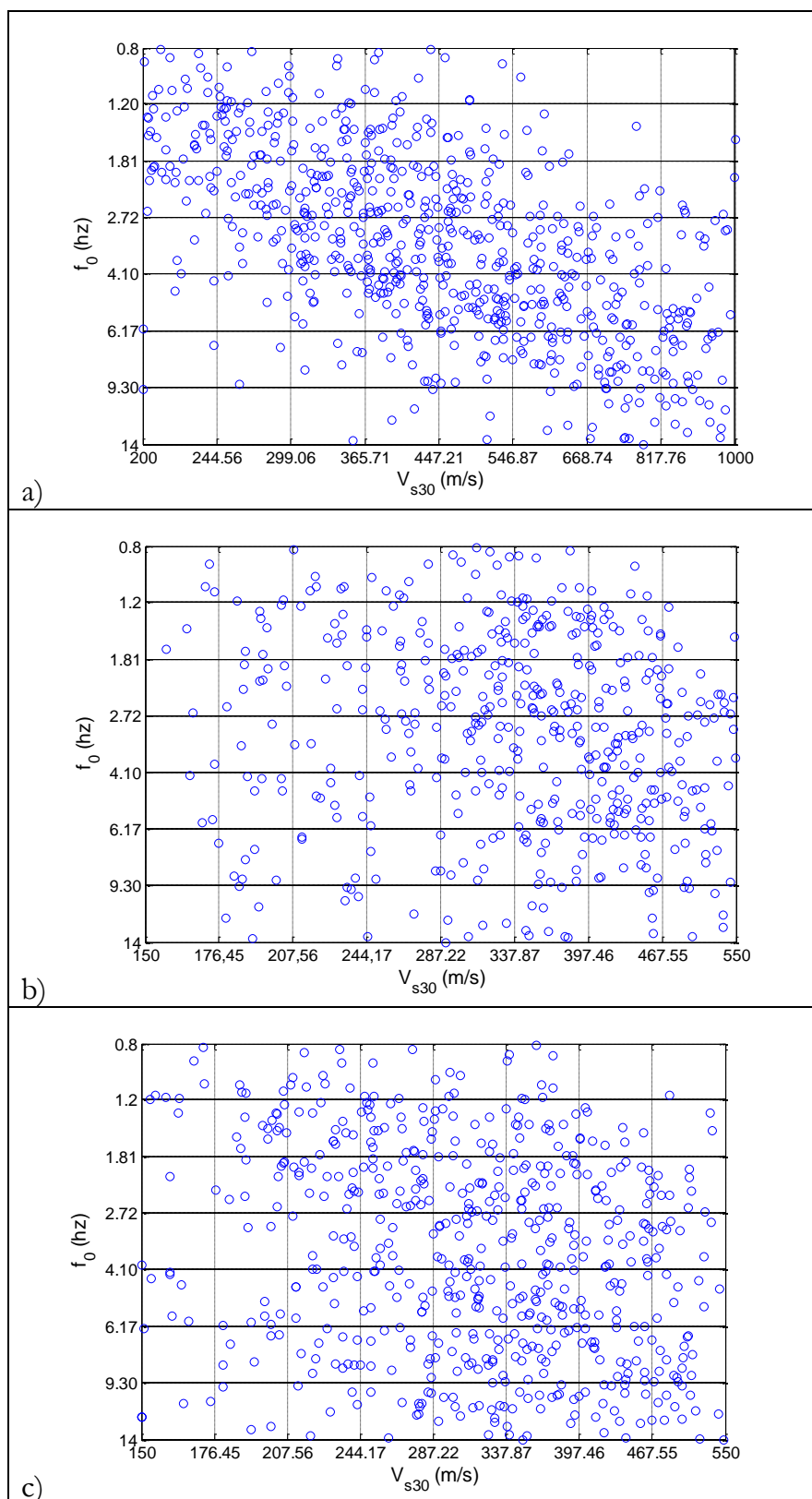


Figure 5-22: Distribution des données initiales dans le plan (f_0 , V_{s30}) pour les (a)

RP, (b) NP et (c) TP.

Les figures (Figure 5-22 a, b et c) donnent la distribution des données dans les plans (f_0 , V_{s30}) pour les profils (RP, NP et TP) .

La relation trouvée en terme de (F_a et F_v) peut être considérée comme fiable pour l'ensemble de la zone rectangulaire décrite par la gamme de fractile 5-95% des deux paramètres (V_{s30} et f_0) pour les cas (NP) et (TP) tandis que pour le modèle (RP) est très faiblement contraint pour les sites à haute fréquence et à faible vitesse (typiquement $f_0 > 5$ Hz et $V_{s30} < 350$ m/s) ainsi pour les sites à basse fréquence et à grande vitesse (typiquement $f_0 < 2$ Hz et $V_{s30} > 500$ m/s).

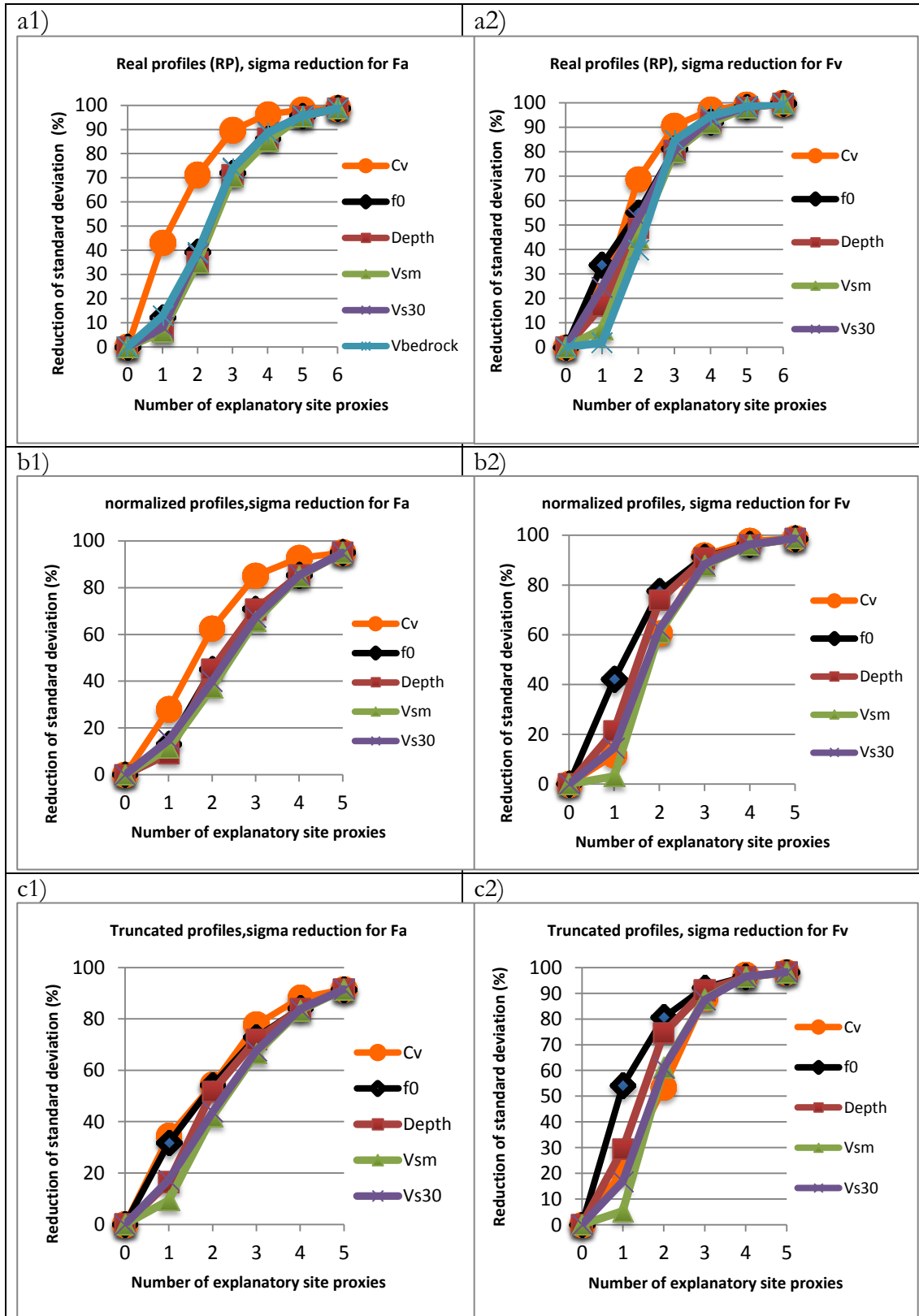


Figure 5-23: Indicateurs de performance en termes de réduction de l'écart-type des résidus pour les différents proxies de site (différentes courbes) pour le cas des profils RP, NP et TP ((a1, b1 et c1) : Fa et (a2, b2 et c2) : Fv).

Finalement, il convient d'indiquer que, pour ce qui est des facteurs d'amplification, toutes les combinaisons possibles de paramètres de site ont été prises en compte ainsi les modèles GRNN associés. Les résultats obtenus ont été analysés. Les performances de certaines d'entre elles sont énumérées dans les tableaux (Tableau 19, Tableau 20 et Tableau 21) respectivement pour les profils ((RP) (NP) et (TP)).

Tableau 19 Variation de l'écart type avec le nombre de paramètres du site pour le cas des profils (RP) (profils réels).

Number of parameters	Explanatory parameters	Fa		Fv	
		Standard deviation of residuals	Variance reduction	Standard deviation of residuals	Variance reduction
All parameters	Depth+ $f_0 + V_{sm}$ + $C_v + V_{s30} + V_{bedrock}$	0.0020	99.98%	2.324e-4	99.999%
Three parameters (best triplet)	$f_0 + C_v + V_{s30}$	0.0094	99.6%	0.0067	99.5%
Two parameters (best pair)	$f_0 + C_v$	0.0349	94.9%	0.0173	97.0%
Two parameters (convenient pair)	$f_0 + V_{s30}$	0.1083	50.7%	0.0515	73.2%
One parameter (best, Fa)	C_v	0.0878	67.6%	0.0786	37.6%
One parameter (best, Fv)	f_0	0.1355	22.9%	0.0661	55.9%
One parameter (usual)	V_{s30}	0.1419	15.4%	0.0752	42.9%
Initial σ		0.1543		0.0995	

Tableau 20 Variation de l'écart type avec le nombre de paramètres du site pour le cas (NP),
(profils normalisé).

(NP)		Fa		Fv	
Number of parameters	Explanatory parameters	Standard deviation of residuals	Variance reduction	Standard deviation of residuals	Variance reduction
All parameters	Depth + f_0 + V_{sm} + Cv + V_{s30}	0.0020	99.97%	2.324e-4	99.996%
Three parameters (best triplet)	f_0 + Cv + V_{s30}	0.0103	99.5%	0.0042	99.8%
Two parameters (best pair)	f_0 + Cv	0.0349	94.9%	0.0173	97.0%
Two parameters (convenient pair)	f_0 + V_{s30}	0.0809	72.5%	0.0280	92.1%
One parameter (best, Fa)	Cv	0.0878	67.6%	0.0786	37.6%
One parameter (best, Fv)	f_0	0.1355	22.9%	0.0661	55.9%
One parameter (usual)	V_{s30}	0.1224	37.1%	0.0932	12.3%
Initial σ		0.1543		0.0995	

Tableau 21 Variation de l'écart type avec le nombre de paramètres du site pour le cas (TP)
(profils tronqué).

(TP)		Fa		Fv	
Number of parameters	Explanatory parameters	Standard deviation of residuals	Variance reduction	Standard deviation of residuals	Variance reduction
All parameters	Depth + f_0 + V_{sm} + Cv + V_{s30}	0.0095	99.28%	0.0017	99.97%
Three parameters (best triplet)	f_0 + Cv + V_{s30}	0.0185	97.2%	0.0061	99.57%
Two parameters (best pair for Fa)	f_0 + Cv	0.0375	88.8%	0.0229	93.9%
Two parameters (convenient pair)	f_0 + V_{s30}	0.0511	79.1%	0.0127	98.1%
One parameter (best, Fa)	Cv	0.0734	57.05%	0.0736	37.9%
One parameter (best, Fv)	f_0	0.0764	53.47%	0.0428	79%
One parameter (usual)	V_{s30}	0.0925	31.4%	0.0777	30.8%
Initial σ		0.112		0.0934	

5.5 Conclusion

La présente étude a permis de bien viser et identifier les paramètres clés contrôlant la réponse du site 1D, en commençant ici par le domaine linéaire comme première étape et cela pour 858 colonnes de sol mesurées sur des sites réels en provenance du Japon, des États-Unis et d'Europe. La réponse linéaire 1D (viscoélastique) a été calculée pour les ondes incidentes planes (type SH) avec une propagation verticale pour un ensemble représentatif d'accélérogrammes d'entrée réels qui couvre une large gamme de fréquences de pics.

La moyenne géométrique de l'amplification correspondante a été obtenue à partir du rapport entre les spectres de réponse d'accélération en surface et d'entrée (en substratum), à la fois en fonction des facteurs d'amplification AF (f) dépendants de la fréquence et des facteurs d'amplification Fa et Fv "récapitulatifs" [0,1 s, 0,2 s] et [0,75 s, 1,5 s], respectivement.

L'approche GRNN a été utilisée (Réseau de Neurone à Régression Généralisée) pour étudier la relation entre ces facteurs d'amplification et les proxies de site "usuelles", c'est-à-dire V_{s30} , f_0 , l'épaisseur total des sédiments 'Depth', la vitesse moyenne des sédiments ' V_{sm} ', le contraste de vitesse et la vitesse du substratum ' C_v '.

Comme les profils réels (RP) présentent une grande variabilité d'un site à un autre, du point de vue : vitesse de rocher. La vitesse et l'épaisseur de chaque couche pour chaque profils a été également multiplier par un coefficient afin de normaliser les profils à une vitesse de roche-mère uniforme à 800 m/s cependant les fonctions de transfert sont à nouveau calculées et les modèles GRNN également étaient développés pour ces profils normalisés (NP), en outre une autre procédure est de considéré pour chaque profil de sol, qui est maintenu inchangé jusqu'à la profondeur Z_{800} où la vitesse dépasse 800 m/s et au-delà la vitesse est fixée à 800 m/s, Chaque fois que la vitesse du substrat rocheux réel est inférieure à 800 m/s, la vitesse du substratum rocheux est alors portée à 800 m/s. Ce troisième ensemble appelé (TP) (profils tronqués) comprend donc également 858 profils de sols.

Similaire avec les profils réels et normalisés en substratum, tout le travail a été refait pour ce nouveau type de profils.

De nombreux modèles GRNN ont été considérés pour chaque cas, avec toutes les combinaisons possibles de proxies de site, d'un seul proxy à tous sous plusieurs combinaisons, ce qui permet de comparer les performances respectives de chaque proxy pour expliquer (et prédire) l'amplification du site.

Les résultats montrent que les caractéristiques clés des facteurs d'amplification dépendants de la fréquence peuvent être reproduites de manière satisfaisante avec un nombre limité de proxies de site. Le paramètre du site le plus performant est le contraste de vitesse global entre la vitesse de la roche mer et la vitesse minimale dans la colonne de sol (qui est généralement à la surface). Ce résultat est valable pour les profils 'RP' en revanche pour les profils 'NP' ou 'TP' ce résultat n'est pas valable de plus, on commence à sentir le rôle de la fréquence de résonance.

Comme le C_v est l'un des paramètres difficiles parfois onéreux à mesurer- en particulier dans le cas des dépôts épais -, on préfère d'autres paramètres plus commodes, parmi lesquels le couple (V_{s30}, f_0) qui permet de réduire la variance des résidus par au moins 60%.

Dans une perspective de code, des équations et des figures sont fournis décrivant la dépendance des facteurs d'amplification "Fa" et "Fv" à courte et moyenne période sur ces deux paramètres $(V_{s30}$ et f_0).

Les résultats montrent que Fa peut atteindre des valeurs élevées pour les sites avec faibles vitesses et forte f_0 (C'est-à-dire des sites minces et mous), tandis que les valeurs les plus élevées de Fv correspondent à des vitesses faibles et à des valeurs f_0 faibles.

Ces résultats ouvrent la voie à des améliorations dans la classification des sites avec une relation physique entre les proxies de site et l'amplification du site. Ce n'est plus cependant qu'une première étape, et les résultats actuels devraient être complétés par plusieurs autres investigations.

D'abord, l'ensemble des profils de sol considérés est dominé par les sites KiK-net, ils sont plutôt rigides bien que ce biais a été quelque peu corrigé avec l'ensemble des profils normalisés 'NP' ou profils tronqués 'TP', il n'est pas entièrement satisfaisant car la procédure de normalisation incluait également une mise à l'échelle de la profondeur pour maintenir la fréquence inchangée. L'ajout de sites plus mous étendrait la plage d'applicabilité des résultats à des sites plus mous et plus épais.

Deuxièmement, ces résultats sont limités dans le cadre linéaire. Une autre prochaine importante étape consistera à considérer les réponses de sites non linéaires. Ou on assigne des caractéristiques non linéaires à différentes couches de chaque profil de sol (information qui est actuellement indisponible) et ajouter au moins une variable explicative dans la couche d'entrée, liée au niveau de chargement (PGA ou autre), sera nécessaire.

Conclusion

La plupart des codes sismiques modernes prennent en compte l'effet du site par l'intermédiaire d'un facteur d'amplification (AF) en modifiant les spectres de réponse d'accélération de la roche par rapport à un «proxy de condition de site», c'est-à-dire que ce sont des paramètres relatifs aux propriétés géométriques et mécaniques du profil de sol .

Pour des raisons pratiques, il est donc intéressant d'identifier les paramètres du site qui contrôlent le mieux le facteur d'amplification AF. Le but de cette étude est de fournir une évaluation quantitative de la performance de différents proxies condition du site pour prédire les principales caractéristiques de la AF, y compris les facteurs d'amplification F_a et F_v souvent utilisés à courte et moyenne période proposés par Borchardt en 1994.

Dans ce but, la réponse viscoélastique linéaire d'un ensemble de 858 colonnes de sol réel est calculée pour un ensemble de 14 accélérogrammes réels avec des pics en spectre de réponse avec des fréquences variables, ainsi la corrélation entre les facteurs d'amplification moyens spécifiques au site correspondants et les proxies du site correspondant seul ou par l'intermédiaire de combinaisons multiples .ce pendant une analyse est réalisée à l'aide du réseau neuronal de régression généralisée (GRNN). La performance de chaque combinaison de proxy de site est évaluée par la réduction de la variance par rapport à la variabilité initiale des facteurs d'amplification pour les 858 profils, en considérant à la fois la plage de période entière et les plages à courte et moyenne période associées aux facteurs de Borchardt F_a et F_v .

Il est montré que l'amplification particulière d'un profil de sol peut être approximée de façon satisfaisante avec un nombre limité de proxies de site (1 à 6) c'est La principale contribution de cette thèse. Comme la pratique des codes

habituelles implique un nombre de proxy de site limité (généralement 1, parfois 2 au max).

Une analyse de sensibilité est effectuée pour identifier les paramètres du site «des plus performants». Le meilleur est le contraste de vitesse global entre le substratum sous-jacent et la vitesse minimale dans la colonne du sol pour les profils originaux 'RP', en revanche ce résultat n'est pas applicable pour les autres bases de données : normalisés 'NP' ou tronqués 'TP' où la fréquence de résonance f_0 devienne prédominante.

Comme le contraste de vitesse est l'un des paramètres le plus difficile parfois plus onéreux à mesurer - en particulier dans le cas des dépôts épais -, on préfère d'autres paramètres plus commodes, en particulier le couple (V_{s30}, f_0) qui permet de réduire d'au moins 60% .

Aussi un autre résultat est fourni décrivant la dépendance des facteurs d'amplification "Fa" et "Fv" à courte et moyenne période sur ces deux paramètres $(V_{s30}$ et $f_0)$.

Les résultats montrent que Fa peut atteindre des valeurs élevées pour les sites avec faibles vitesses et forte fréquence (f_0) (C'est-à-dire des sites minces et mous), tandis que les valeurs les plus élevées de Fv correspondent à des vitesses faibles et à des valeurs f_0 faibles.

Ces résultats ouvrent la voie à des améliorations dans la classification des sites avec une relation physique entre les proxies de site et l'amplification du site. Ce n'est plus cependant qu'une première étape, et les résultats actuels devraient être complétés par plusieurs autres investigations.

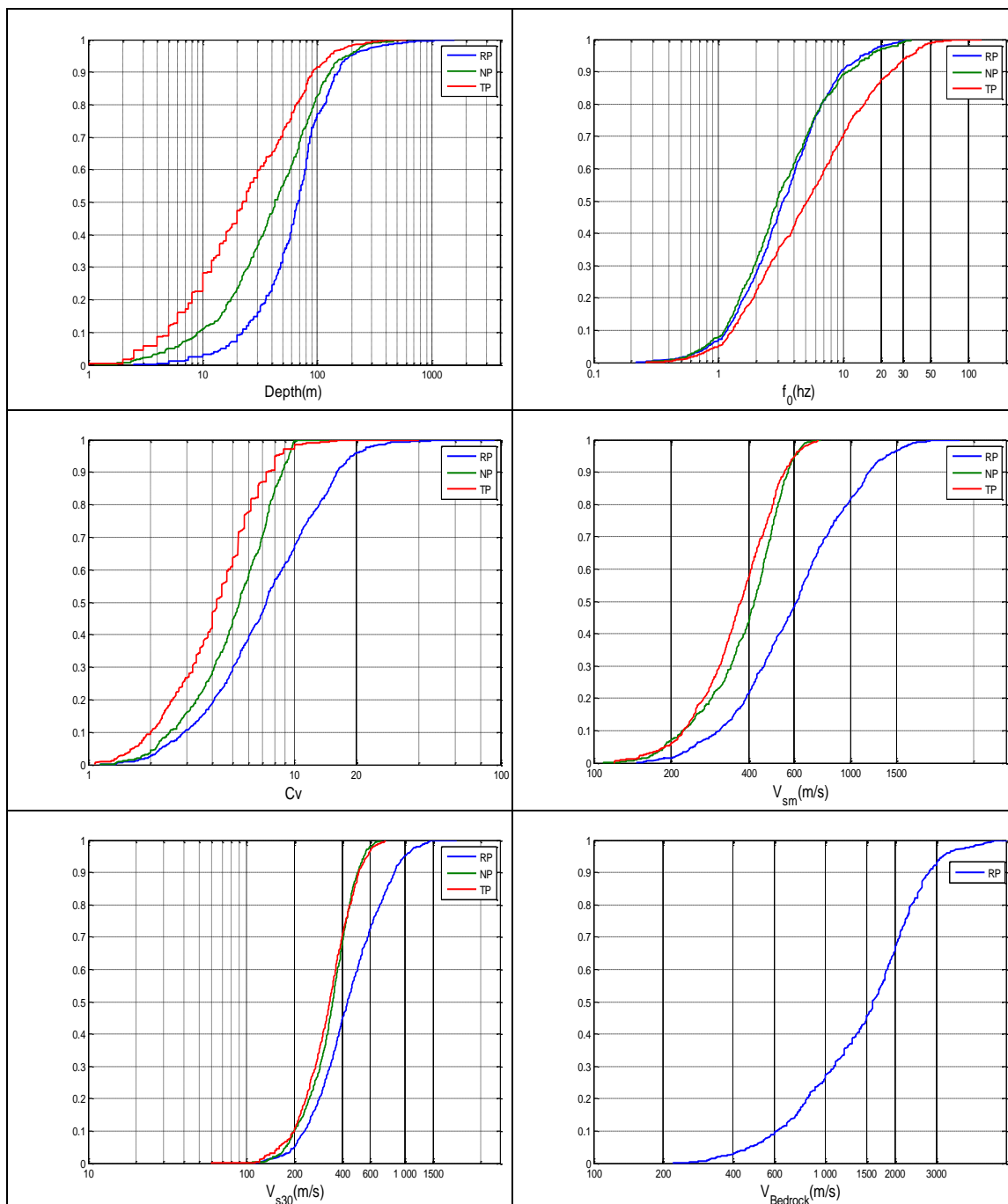
A la fin, il faut d'abord signaler que l'ensemble des profils de sol considérés est dominé par les sites KiK-net, ils sont plutôt rigides bien que ce biais a été quelque peu corrigé avec l'ensemble des profils normalisés ou profils tronqués, il n'est pas entièrement satisfaisant car la procédure de normalisation incluait également une mise à l'échelle de la profondeur pour maintenir la fréquence inchangée. L'ajout de

sites plus moux étendrait la plage d'applicabilité des résultats à des sites plus moux et plus épais.

Deuxièmement, ces résultats sont limités dans le cadre linéaire. Une autre prochaine importante étape consistera à considérer les réponses de sites non linéaires. Ou on Assigne des caractéristiques non linéaires à différentes couches de chaque profil de sol (information qui est actuellement indisponible) et ajouter au moins une variable explicative dans la couche d'entrée, liée au niveau de chargement (PGA et autre), sera nécessaire.

Annexes

Annexe 1 : Fonction de distribution cumulative (cdf) pour les différents paramètres d'entrés (Depth, f_0 , V_{sm} , V_{s30} , C_v , $V_{bedrock}$) application pour les différents profils types (RP, NP et TP).



Annexe 2 : Définition de la Validation croisée

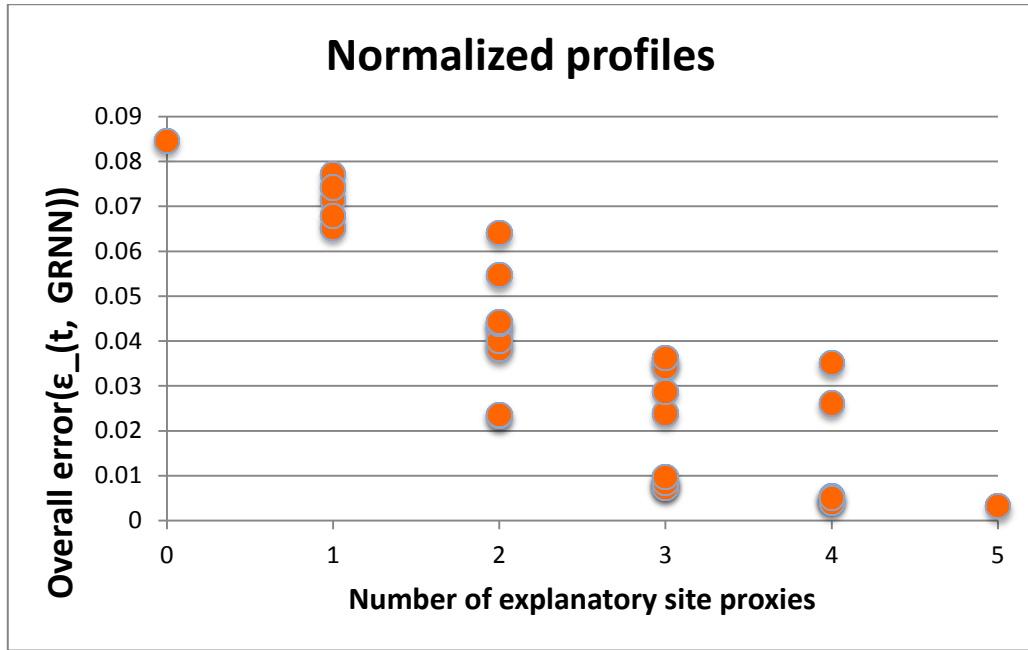
La validation croisée (« cross-validation ») est une méthode d'estimation de fiabilité d'un modèle fondé sur une technique d'échantillonnage. En fait, il y a au moins trois techniques de validation croisée : « tests et validation » ou « holdout method », « k-fold cross-validation » et « leave-one-out cross-validation » (LOOCV).

-La première méthode est très simple, il suffit de diviser l'échantillon de taille n en deux sous échantillons, le premier d'apprentissage (communément supérieur à 60 % de l'échantillon) et le second de test. Le modèle est bâti sur l'échantillon d'apprentissage et validé sur l'échantillon de test. L'erreur est estimée en calculant un test, une mesure ou un score de performance du modèle sur l'échantillon de test, par exemple l'erreur quadratique moyenne.

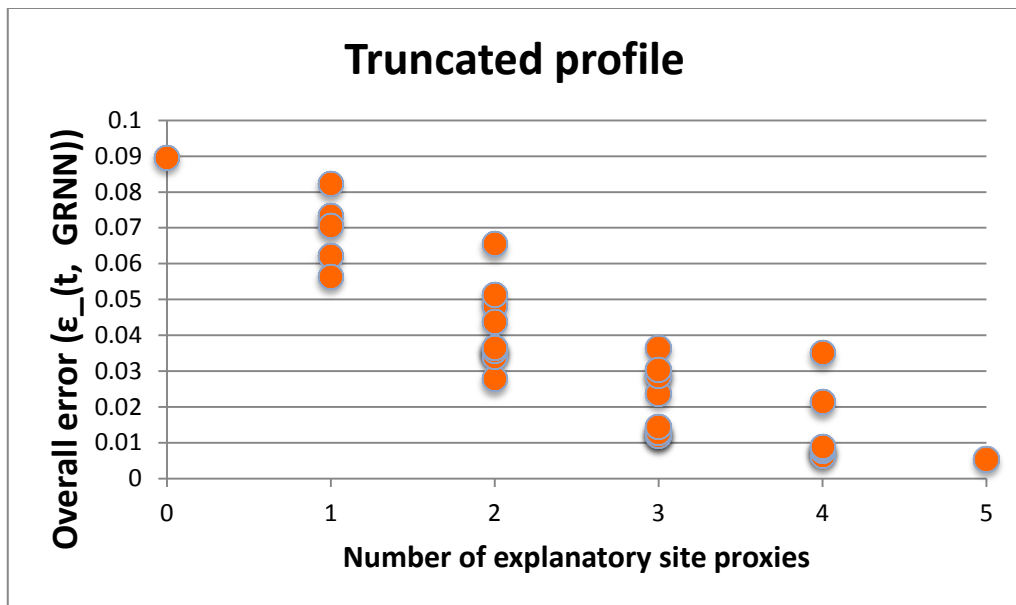
- Dans la seconde, on divise l'échantillon original en k échantillons, puis on sélectionne un des k échantillons comme ensemble de validation et les $(k-1)$ autres échantillons constitueront l'ensemble d'apprentissage. On calcule comme dans la première méthode le score de performance. Puis on répète l'opération en sélectionnant un autre échantillon de validation parmi les $(k-1)$ échantillons qui n'ont pas encore été utilisés pour la validation du modèle. L'opération se répète ainsi k fois pour qu'en fin de compte chaque sous-échantillon ait été utilisé exactement une fois comme ensemble de validation. La moyenne des k erreurs quadratiques moyennes est enfin calculée pour estimer l'erreur de prédiction.

- La troisième méthode est un cas particulier de la deuxième méthode où $k=n$, c'est-à-dire que l'on apprend sur $(n-1)$ observations puis on valide le modèle sur la n ème observation et l'on répète cette opération n fois.

Annexe 3 : Illustration de la réduction progressive de l'écart type des résidus entre les modèles GRNN et le AF réelle avec le nombre de proxies de sites considérées pour les profils (NP).



Annexes 4 : Illustration de la réduction progressive de l'écart type des résidus entre les modèles GRNN et le AF réelle avec le nombre de proxies de sites considérées pour les profils (TP).



Annexes 5 : fichier Excel pour calculer les facteurs d'amplification à partir de V_{s30} et f_0 .

Bibliographie

- Abrahamson, N., G. Atkinson, D. Boore, Y. Bozorgnia, K. Campbell, B. Chiou, I. Idriss, W. Silva, and R. Youngs,. (2008), Comparisons of the NGA ground-motion relations, *Earthquake Spectra*, 24(1), 45-66. doi: 10.1193/1.2924363.
- Aki K, Richards PG (1980) *Quantitative seismology, theory and methods*, vol 1. WH Freeman & Co., New York.
- Alani, T. (2008). Réseaux de Neurones Tutorial en Matlab. Département Informatique ESIEE-Paris, Paris.
- Almakari M, Régnier J, Salameh C, Cadet H, Bard PY, Lopez-Caballero F, Cornou C (2016) Modulation of weak motion site transfer functions by non-linear behavior: a statistical comparison of 1D numerical simulation with KiKnet data. In: *Proceedings of the 5th IASPEI/IAEE international symposium: effects of surface geology on seismic motion*, Taipei, August 15–17 Paper P101C, 14 pp.
- Ancheta, T. D., R. B. Darragh, J. P. Stewart, E. Seyhan, W. J. Silva, B. S.-J. Chiou, K. E. Wooddell, R. W. Graves, A. R. Kottke, D. M. Boore, T. Kishida and J. L. Donahue (2014). NGA-West 2 database, *Earthquake Spectra* 30, 989–1005.
- Bard P.Y., M. Campillo, F.J. Chavez-Garcia & F.J. Sanchez-Sesma (1988). A theoretical investigation of large-and small-scale amplification effects in the Mexico City valley, *Earthquake Spectra* 4, pp. 608-633.
- Bard, P. Y. (1999). Microtremor measurements: a tool for site effect estimation. *The effects of surface geology on seismic motion*, 3, 1251-1279.
- Bard, P.-Y., & Thomas, J.R. (2000). *Wave preparation in complex geological structures and their effects on strong ground motion*. Wave propagation in earthquake engineering. WIT Press Great Britain, SO40 7AA, UK.
- Bard, P. Y., Zare, M., & Ghafory-Ashtiany, M. (1998). The Iranian Accelerometric Data Bank: A Revision and Data Correction. *Journal of Seismology and Earthquake Engineering*, 1(1), 1-22.
- Bardet, J., Ichii, K., & Lin, C. (2000). EERA: a computer program for equivalent-linear earthquake site response analyses of layered soil deposits: University of Southern California, Department of Civil Engineering, <http://geoinfo.usc.edu/gees>.
- Biro Y., and P. Renault (2012), Importance and Impact of Host-to-Target Conversions for Ground Motion Prediction Equations in PSHA, 15 WCEE, Lisboa 2012.
- Bonilla, L. F., Steidl, J. H., Lindley, G. T., Tumarkin, A. G., & Archuleta, R. J. (1997). Site amplification in the San Fernando Valley, California: Variability of site-effect estimation using the S-wave, coda, and H/V methods. *Bulletin of the Seismological Society of America*, 87(3), 710-730.
- Bonnefoy-Claudet, S. (2004). *Nature du bruit de fond sismique: implications pour les études des effets de site* (Doctoral dissertation, Université Joseph-Fourier-Grenoble I).
- Borcherdt, R. D. (1994). Estimates of site-dependent response spectra for design (methodology and justification). *Earthquake spectra*, 10(4), 617-653.

- Borcherdt, R. D. (2002). Empirical evidence for acceleration-dependent amplification factors. *Bulletin of the Seismological Society of America*, 92(2), 761-782.
- Butterworth, S. (1930). On the theory of filter amplifiers. *Wireless Engineer*, 7(6), 536-541.
- Cadet, H., Duval, A. M., Bard, P. Y., Bertrand, E., & Bertrand, (2012). Site effect assessment using KiK-net data: part 2—site amplification prediction equation based on f_0 and V_{sz} . *Bulletin of Earthquake Engineering*, 10(2), 451-489. doi:10.1007/s10518-011-9298-7.
- Cadet, H., Duval, A. M., Bertrand, E., & Bard, P. Y. (2007, July). Case study of noise array measurements in soft clay at l'Ebron, Trièves, Isère, France. In 7ème colloque national du génie parasismique AFPS'07 (p. 9p).
- . Castellaro, S., Mulargia, F., & Rossi, P. L. (2008). VS30: Proxy for seismic amplification?. *Seismological Research Letters*, 79(4), 540-543.
- Cigizoglu HK, Alp M (2005) Generalized regression neural network in modelling river sediment yield. *Adv Eng Softw* 37:63–68.
- Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems (MCCS)*, 2(4), 303-314.
- Demartines, P (1994). Analyse de données par réseaux de neurones auto-organisés, thèse de l'Institut National Polytechnique de Grenoble.
- Derras B., (2004). Estimation du risque lié à l'effet de site et génération d'un spectre de réponse à la surface libre, Mémoire de Magister, Université Abou Baker Belkaid - Tlemcen.
- Derras, B., A. Bekkouche, and D. Zendagui, (2010). Neuronal approach and the use of kik-net network to generate response spectrum on the surface. *Jordan Journal of Civil Engineering*, 4 12-21.
- Derras, B., Bard, P. Y., Cotton, F., & Bekkouche, A. (2012). Adapting the neural network approach to PGA prediction: an example based on the KiK-net data. *Bulletin of the Seismological Society of America*, 102(4), 1446-1461. doi: 10.1785/0120110088.
- Di Giulio, G., A. Savvaidis, M. Ohrnberger, M. Wathelet, C. Cornou, B. Knapmeyer-Endrun, F. Renalier, N. Theodoulidis and P.-Y. Bard, 2012. Exploring the model space and ranking a best class of models in surface-wave dispersion inversion: application at European strong motion sites", *Geophysics* 77, B147 (2012), <http://dx.doi.org/10.1190/geo2011-0116.1>.
- Dickenson, S. E., and R.B. Seed, (1995). Nonlinear dynamic response of soft and deep cohesive soil deposits", *Proc. of Int. Workshop on Site Response*, 1995, Yokosuka, Japan, 2: 67-81.
- Dobry, R., Oweis, I., & Urzua, A. (1976). Simplified procedures for estimating the fundamental period of a soil profile. *Bulletin of the Seismological Society of America*, 66(4), 1293-1321.
- Dobry, R., R. D. Borcherdt, C. B. Crouse, I. M. Idriss, W. N. Joyner, G. R. Martin, M. S. Power, E. E. Rinne and R. B. Seed (2000). New Site coefficients and site classification system used in recent building seismic code provisions". *Earthquake Spectra*, 2000, 16(1): 41-67.

- Douglas, J., Akkar, S., Ameri, G., Bard, P. Y., Bindi, D., Bommer, J. J., ... & Kuehn, N. M. (2014). Comparisons among the five ground-motion models developed using RESORCE for the prediction of response spectral accelerations due to earthquakes in Europe and the Middle East. *Bulletin of earthquake engineering*, 12(1), 341-358. doi 10.1007/s10518-013-9522
- Dreyfus, G., J-M. Martinez, M. Samuelides, M.B. Gordon, F. Badran, and S. Thiria (2008). *Apprentissage statistique : Réseaux de neurones : Cartes topologiques Machines à vecteurs supports*, eyrolles.
- DTR B C 2 48. (2003). Règles parasismiques algériennes- R.P.A.99 version 2003”, document technique réglementaire, centre national de recherche appliquée en génie parasismique, Ministère de l’habitat, Algérie.
- Duval, A. M. (1994). Détermination de la réponse d'un site aux séismes à l'aide du bruit de fond: évaluation expérimentale (Doctoral dissertation, Paris 6).
- Duval, A. M., Méneroud, J. P., Vidal, S., & Singer, A. (1998). Relation between curves obtained from microtremor and site effects observed after Caracas 1967 earthquake. In 11th European Conference on Earthquake Engineering.
- EC8 Eurocode 8. (2004). Design of structures for earthquake resistance—part 1: General rules, seismic actions and rules for buildings. European Committee for Standardization (CEN), EN 1998-1, eurocodes.jrc.eceuropa.eu/ (last accessed February 2016).
- Funahashi, K. I. (1989). On the approximate realization of continuous mappings by neural networks. *Neural networks*, 2(3), 183-192.
- Gelebi, M. (1988). Topographical and geological amplification : Case studies and engineering implications”. International Workshop on Spatial Variation on earthquake Ground Motion, USA.
- Ghaboussi, J., and C.C.J. Lin (1998). New method of generating spectrum compatible accelerograms using neural networks, *Earthquake Engineering and Structural Dynamics*. 27, 377-396.
- Gregor, N., Abrahamson, N. A., Atkinson, G. M., Boore, D. M., Bozorgnia, Y., Campbell, K. W., ... & Silva, W. (2014). Comparison of NGA-West2 GMPEs. *Earthquake Spectra*, 30(3), 1179-1197.
- Koivo, H. N. (1994). Artificial neural networks in fault diagnosis and control. *Control engineering practice*, 2(1), 89-101.
- Hannan SA, Manza RR, Ramteke RJ (2010) Generalized regression neural network and radial basis function for heart disease diagnosis. *Int J CompApp* 7(13):7–13.
- Hashash, Y. M., Groholski, D., Phillips, C., Park, D., & Musgrove, M. (2012), DEEPSOIL 5.1. User Manual and Tutorial, 107, [https:// www.illinois.edu](https://www.illinois.edu) .
- Hashash, Y., Jung, S., & Ghaboussi, J. (2004). Numerical implementation of a neural network based material model in finite element analysis. *International Journal for numerical methods in engineering*, 59(7), 989-1005.
- Haskell, N. A. (1953). The dispersion of surface waves on multilayered media. *Bulletin of the seismological Society of America*, 43(1), 17-34.

- IBC2012. International Building Code 2012 Edition, ISBN 978-1-60983-039-7, International Code Council, <https://archive.org/details/gov.law.icc.ibc.2012>, last accessed November 2016.
- Park, J. W., Harley, R. G., & Venayagamoorthy, G. K. (2002). Comparison of MLP and RBF neural networks using deviation signals for on-line identification of a synchronous generator. In Power Engineering Society Winter Meeting, 2002. IEEE (Vol. 1, pp. 274-279). IEEE.
- JSR, J. Sun CT, Mizutani E (1997). Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence.
- Kawase, H. (1996). The cause of the damage belt in Kobe: 'The basin-edge effect,' constructive interference of the direct S-wave with the basin-induced diffracted/Rayleigh waves. *Seismological Research Letters*, 67(5), 25-34.
- . Kim, B., Lee, D. W., Park, K. Y., Choi, S. R., & Choi, S. (2004). Prediction of plasma etching using a randomized generalized regression neural network. *Vacuum*, 76(1), 37-43.
- Klimis, N.S., B.N. Margaris, and P.K. Koliopoulos (1998). Response spectra estimation according to the EC8 and NEHRP soil classification provisions: a comparison study based on Hellenic data. Proceeding of the 11th European Conference on Earthquake Engineering T8 Site effects, spatial variability of seismic motion.
- Kramer S.L., (1996). *Geotechnical Earthquake Engineering. Civil Engineering and Engineering Mechanics.* Prentice Hall, Upper Saddle River, New Jersey 07458.
- LeBrun, B. (1997). *Les effets de site: étude expérimentale et simulation de trois configurations* (Doctoral dissertation, Université Joseph-Fourier-Grenoble I).
- Liang, C., & Mu, D. (2014). Evaluation Research of Urban Distribution System's Resource Integration based on Generalized Regression Neural Network.
- Luzi, L, R Puglia, F Pacor, MR Gallipoli, D Bindi and M Mucciarelli, (2011). Proposal for a soil classification based on parameters alternative or complementary to Vs, 30. *Bulletin of Earthquake Engineering* 9 (6), 1877-1898.
- Fatima, M., & Hamid, S. (2009). Comparaison de méthodes de classification réseau RBF, MLP et RVFLNN1. *Damascus University Journal* Vol,(25)-No.(2).
- . Martin, G. R., & Dobry, R. (1994). Earthquake site response and seismic code provisions. *NCEER Bulletin*, 8(4), 1-6.
- Msaaf, M., & Belmajdoub, F. (2015, December). L'application des réseaux de neurone de type «feedforward» dans le diagnostic statique. In *Xème Conférence Internationale: Conception et Production Intégrées*.
- Nakamura Y., A method for dynamic characteristics estimation of subsurface using microtremor on the ground surface, 1989, *Quarterly Report Railway Tech. Res. Inst.*, 30-1, pp 25-30.
- Nogoshi, M., & Igarashi, T. (1971). On the amplitude characteristics of microtremor— Part 2: *Journal of the Seismological Society of Japan*, 24, 26–40.
- Parizeau, M. (2004). Réseaux de neurones. GIF-21140 et GIF-64326, 124.

- Pitarka, A., Irikura, K., Iwata, T., & Sekiguchi, H. (1998). Three-dimensional simulation of the near-fault ground motion for the 1995 Hyogo-ken Nanbu (Kobe), Japan, earthquake. *Bulletin of the Seismological Society of America*, 88(2), 428-440.
- Pitilakis, K. D., Raptakis, D. G., & Makra, K. A. (1999, June). Site effects: recent considerations and design provisions. In 2nd Int. Conf. on Earthquake Geotechnical Eng., Lisbon, Balkema ed (pp. 901-912).
- Pitilakis, K. D., Makra, K. A., & Raptakis, D. G. (2001). 2D vs 3D site effects with potential applications to seismic norms: the case of EUROSEISTEST and Thessaloniki. *Proceedings of the XVth ICSMGE, Istanbul*, 123-33.
- Pitilakis, K., Riga, E., & Anastasiadis, A. (2012). Design spectra and amplification factors for Eurocode 8. *Bulletin of Earthquake Engineering*, 10(5), 1377-1400.
- Pitilakis, K., Riga, E., & Anastasiadis, A. (2013). New code site classification, amplification factors and normalized response spectra based on a worldwide ground-motion database. *Bulletin of Earthquake Engineering*, 11(4), 925-966.
- Renault, P.L.A., N.A. Abrahamson, P.-Y. Bard, D. Fäh, A. Pecker, and J. Studer, (2014). PEGASOS Refinement Project, Volume 5, SP3 - Site Response Characterization, 672 pages, available from <http://www.swissnuclear.ch/de/downloads.html>. ©2013-2015 Swissnuclear, Olten, Switzerland.
- Riepl, J., Gariel, J., & Bard, P. (1998). "About the Coda wave signals to estimate local site amplification factors". second international symposium on the effects of surface geology on seismic motion, Yokohama, Japan, 523-530.
- Rodriguez-Marek, A., Bray, J. D., & Abrahamson, N. A. (2001). An empirical geotechnical seismic site response procedure. *Earthquake Spectra*, 17(1), 65-87.
- Salameh, C., P.-Y. Bard, B. Guillier, J. Harb, C. Cornou and M. Almakari, (2016). Using ambient vibration measurements for risk assessment at an urban scale: from numerical proof of concept to a case study IN BEIRUT (LEBANON). Invited keynote lecture, 5th IASPEI / IAEE International Symposium: Effects of Surface Geology on Seismic Motion, August 15-17, 2016, paper #K102A, 17 pages.
- Saragoni, G.R., Lobos, C., & Gómez-Bernal, A. (1998). Site and Earthquake Mechanism Effect on Design Response Spectra". 11th European Conference on Earthquake Engineering, Balkema, Rotterdam, ISBN 90 5410 982 3.
- Schnabel, P. B., J. Lysmer, and H. B. Seed (1973). SHAKE--A Computer Program for Earthquake Response Analysis of Horizontally Layered Sites, Report No, EERC 72 12, Earthquake Engineering Research Center, University of California, Berk'eley.
- Semblat, J. F., Duval, A. M., & Dangla, P. (2000). Numerical analysis of seismic wave amplification in Nice (France) and comparisons with experiments. *Soil Dynamics and Earthquake Engineering*, 19(5), 347-362.
- Specht D. F., (1991). A general regression neural network. *IEEE Trans. Neural Netw* 2 (6), pp. 568–576.
- Stambouli, A. B., Zendagui, D., Bard, P. Y., & Derras, B. (2017). Deriving amplification factors from simple site parameters using generalized regression neural networks: implications for relevant site proxies. *Earth, Planets and Space*, 69(1), 99.
- Stewart, J. (1997). "Key Geotechnical Aspects of the 1989 Loma Prieta Earthquake". National Information Service for Earthquake Engineering :

- Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological cybernetics*, 43(1), 59-69.
- Bentrcia, T., & Tamrabet, Y. (2010). A LOW COMPLEXITY FINITE DIFFERENCES-NEURAL NETWORK COMBINED APPROACH FOR THE MODELING OF PN JUNCTION ELECTRICAL PROPERTIES.
- Theodulidis, N. P., & Bard, P. Y. (1995). Horizontal to vertical spectral ratio and geological conditions: an analysis of strong motion data from Greece and Taiwan (SMART-1). *Soil dynamics and earthquake engineering*, 14(3), 177-197.
- Thomson, W. T. (1950). Transmission of elastic waves through a stratified solid medium. *Journal of applied Physics*, 21(2), 89-93.
- Trifunac, M. D. (1990). How to model amplification of strong earthquake motions by local soil and geologic site conditions. *Earthquake Engineering & Structural Dynamics*, 19(6), 833-846.
- UBC Uniform building code (1997). Structural engineering design provisions (Vol. 2). International Conference of building officials.
- Wasserman P.D., (1993). *Advanced Methods in Neural Computing*. John Wiley & Sons, Inc. New York, NY, USA.
- Widrow, B., & Hoff, M. E. (1960). Adaptive switching circuits (No. TR-1553-1). STANFORD UNIV CA STANFORD ELECTRONICS LABS.
- Pao, Y. (1989). *Adaptive pattern recognition and neural networks*.
- Yelles-Chaouche, A.K., & Djellit, H.(2003). Le séisme de Boumerdès (21 MAI 2003,Mw:6.8)”. Colloque International, Alger, Algérie.
- Yezli-Khati, F., & Afra, H. (2003). Réponse sismique d’un profil de sol a une onde de cisaillement SH”. Séminaire international de géomatériaux.

