Utilisation de l'analyse de variance pour évaluer les raisons de la variabilité des mouvements forts du sol

John Douglas, Pierre Gehl

▶ **To cite this version:**

John Douglas, Pierre Gehl. Utilisation de l'analyse de variance pour évaluer les raisons de la variabilité des mouvements forts du sol. 7ème Colloque National AFPS 2007, Jul 2007, Châtenay Malabry, France. 8 p. hal-00714398

HAL Id: hal-00714398
https://brgm.hal.science/hal-00714398

Submitted on 4 Jul 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Utilisation de l'analyse de variance pour évaluer le rôle des effets de source et de site dans la variabilité des mouvements forts du sol

John Douglas — Pierre Gehl

BRGM – ARN/RIS

3, avenue C. Guillemin, BP 36009

45060 Orléans Cedex 2

j.douglas@brgm.fr; p.gehl@brgm.fr

RÉSUMÉ. *Cet article a pour objet de présenter une méthode permettant d'évaluer quantitativement l'importance relative de la modélisation des effets de site et de source dans l'écart entre mouvements du sol prédits et mouvements mesurés. La méthode proposée suit une analyse de variance (ANOVA) sur la différence entre accélérations calculées (à l'aide de modèles empiriques) et accélérations mesurées (plusieurs séismes enregistrés par les mêmes stations) : l'analyse de variance permet alors de projeter la variance globale en composantes dues aux effets de site ou de source non modélisés.*

Afin de tester cette procédure, plusieurs séries d'enregistrements du mouvement fort du sol sont utilisées : pour la séquence sismique d'Ombrie en Italie, cette méthode permet de conclure qu'une grande majorité des variations observées entre mouvements calculés et mouvements mesurés est attribuable à des effets de site non modélisés. A l'inverse, la séquence des Saintes montre que les variations observées sont dues en grande partie à un manque de modélisation de la source.

Cette méthode statistique élémentaire pourrait donc être utilisée pour décider dans quel domaine porter les efforts de modélisation afin de réduire les incertitudes au niveau des modèles empiriques.

ABSTRACT. *A simple method to quantitatively assess the relative importance of unmodelled site and source effects on the observed variation in ground motions is presented. The method consists of analysis of variance (ANOVA) using the computed residuals with respect to an empirical ground-motion model for strong-motion records of various earthquakes recorded at a common set of stations. ANOVA divides the overall variance into the components due to site and source effects not modelled by the ground-motion model.*

To test this procedure, several sets of observed strong-motion records are used. For the data from Umbria, Italy, it is found that the vast majority of the observed variance is attributable to unmodelled site effects. In contrast, the variation in ground motions in the Les Saintes (French Antilles) sequence is largely attributable to source effects not modelled in the ground motion estimation equations used.

This simple technique could be used to decide, for a given set of strong-motion data, where effort should be concerned to reduce the large uncertainties in empirical ground-motion estimates.

MOTS-CLÉS : *données accélérométriques, lois d'atténuation, effets de site, effets de source, régression linéaire à deux composantes*

KEYWORDS: *strong-motion data, ground-motion models, analysis of variance, site effects, source effects, two-way-fit plots*

1. Introduction

L'analyse de variance (ANOVA) est un outil puissant développé par R.A. Fisher (1990), qui permet de séparer la variance globale d'une série d'observations en des composantes associées aux sources possibles de variabilité (Moroney, 1990). Dans une application antérieure de l'analyse de variance au domaine de la sismologie, Douglas (2003, 2004) utilise cette méthode pour mettre en évidence des possibles dépendances régionales des mouvements forts du sol, entre cinq régions européennes, et entre l'Europe, la Californie et la Nouvelle-Zélande.

L'analyse de variance est une méthode simple pour mettre en évidence les contributions relatives des effets de site ou de source sur la variance globale des niveaux d'accélération du sol. Avec cet outil, les variations observées au niveau des mouvements du sol peuvent être séparées en variations causées par la source, celles causées par le site et celles dont l'origine est inexpliquée, probablement due à la propagation. Etant donné que les accélérogrammes sont associés à des magnitudes, des mécanismes de rupture, des distances épicentrales et des classes de site très variés, dont les effets sur les niveaux d'accélération sont déjà connus, nous étudierons ici l'écart entre les valeurs enregistrées et celles obtenues par les lois d'atténuation empiriques. Cette approche permet d'éliminer à peu près les effets primaires de la magnitude, du type de mécanisme, de la distance épicentrale, et des classes de sol : ainsi, seules apparaîtront les variations dont les causes n'ont pas été modélisées par les équations. Nous appliquons ici cette méthode pour les grandeurs suivantes : accélération maximale (PGA) ; et accélération du spectre de réponse élastique à 5% d'amortissement (SA) aux périodes suivantes : 0.2, 0.5, 1.0 et 2.0 s. Les accélérations sont issues de la composante horizontale la plus élevée.

Dans cet article, la technique proposée est appliquée à deux séries d'enregistrements de mouvements forts. La première provient de cinq stations ayant enregistré les quatre mêmes séismes (20 enregistrements au total) de la séquence sismique d'Ombrie (Italie centrale) en 1997-1998. La seconde série contient des enregistrements de six stations ayant enregistré six séismes (36 enregistrements en tout) durant la séquence sismique des Saintes (Guadeloupe, France) en 2004-2005. Voir Douglas & Gehl (2006) pour des détails et deux autres applications de cette technique.

L'utilisation efficace de l'analyse de variance nécessite de prendre en compte uniquement les stations ayant toutes enregistré tous les séismes considérés, ce qui réduit considérablement la quantité de données exploitables. Les deux parties suivantes détaillent les résultats obtenus avec les séries d'enregistrements des Ombrie et des Saintes : pour visualiser les données, nous introduisons la régression linéaire à deux composantes (Tukey, 1972). La dernière partie est consacrée à la discussion des résultats et à la conclusion.

2. La séquence sismique d'Ombrie (Italie) en 1997-1998

Les données et les paramètres associés (magnitude, mécanisme, distance épicentrale…) utilisés ici sont les mêmes que ceux que Ambraseys (et al. 2005) a employés pour définir les équations de prédiction du PGA et de l'accélération spectrale. Le Tableau 1 résume les données retenues pour l'étude.

De nombreuses données issues de la séquence d'Ombrie ont été utilisées pour l'élaboration du modèle de mouvement du sol d'Ambraseys (et al., 2005), c'est pourquoi nous utilisons ces lois d'atténuation pour calculer les résidus de ces données (écart entre les valeurs observées et les valeurs estimées avec les équations). Le Tableau 2 reprend les résultats de l'analyse de variance pour le PGA et l'accélération spectrale à 0.2 et 0.5 s. En raison de bruits parasites pour les longues périodes, l'analyse de variance n'a pu être menée pour des périodes supérieures à 0.5 s (voir Ambraseys et al. 2005 pour la procédure détaillée du traitement des enregistrements).

Séismes			A	B	C	D
JJ/MM/AA			26/09/97	26/09/97	03/10/97	06/10/97
HH:MM			00:33	09:40	08:55	23:24
Mw			5.7	6.0	5.3	5.5
Station	**Code**	**Classe de sol**		**de**		
Assisi Stallone	AS010	Rock	*21*	*14*	19	20
Colfiorito	CLF	Stiff soil	*0*	*5*	7	7
Gubbio Piana	GBP	Soft soil	*38*	*27*	37	38
Nocera Umbra	NCR	Rock	*11*	*1*	10	11
Rieti	RTI	Very soft soil	*61*	*66*	67	63

Tableau 1. *Données de la séquence d'Ombrie de 1997-1998 utilisées dans cette étude.* ***de*** *représente la distance épicentrale et les distances en italique représentent la distance entre la station et la projection de la faille sur la surface (distance de Joyner-Boore, Joyner et al., 1981).*

Paramètre	**Effet**	**Σ**	**d.l.**	**σ^2**	**Ratio**	
PGA	Stations	0.9934	4	0.2483	13.9	***
	Séismes	0.1492	3	0.0298	1.7	
	Résidu	0.2152	12	0.0179		
	Total	1.3577	19			
SA 0.2 s	Stations	1.3069	4	0.3267	8.6	**
	Séismes	0.1654	3	0.0331	0.9	
	Résidu	0.4585	12	0.0382		
	Total	1.9308	19			
SA 0.5 s	Stations	0.6966	4	0.1741	7.2	**
	Séismes	0.0885	3	0.0177	0.7	
	Résidu	0.2916	12	0.0243		
	Total	1.0767	19			

Tableau 2. *Résultats de l'analyse de variance pour les données issues de la séquence d'Ombrie de 1997-1998, où* ***Σ*** *représente la somme des carrés,* ***d.l.*** *le nombre de degrés de liberté,* σ^2 *la variance,* ***Ratio*** *le ratio entre le σ^2 de l'effet (**Stations** ou **Séismes**) et le σ^2 du **Résidu**. Le nombre d'astérisques indique la significativité de l'effet d'après le test F de Fisher (*** correspond à un seuil de probabilité de 0.1%, ** à un seuil de probabilité de 1% et * à un seul de probabilité de 5%).*

Cette analyse montre que les effets de site non modélisés aux stations considérées contribuent en grande partie à la variabilité observée dans les niveaux d'accélération (ceci est révélé par une valeur de σ^2 bien plus importante pour la composante **Stations** que pour **Séismes** ou **Résidu**). Les données utilisées ici ont été enregistrées par les stations de Nocera Umbra (NCR), Gubbio-Piana (GBP) et Rieti (RTI), qui sont sujettes à d'importants effets de site (Ambraseys et al,. 2005), lesquels sont mal intégrés dans les lois d'atténuation empiriques. La station de Nocera Umbra est située près d'une faille verticale avec des roches largement

fracturées qui amplifient les signaux à haute fréquence (Marra et al., 2000) ; les stations de Gubbio-Piana et Rieti sont situées dans des bassins sédimentaires qui génèrent des ondes de surface de grande amplitude (Castro et al., 2004). La conclusion apportée par la méthode ANOVA n'est donc pas surprenante. Il peut être intéressant de noter que cette analyse révèle que les effets de source ne sont pas primordiaux pour ces données, ce qui signifie qu'ajouter des paramètres de source en plus de ceux déjà présents (magnitude, mécanisme) aurait peu de chance de réduire l'écart-type du modèle de prédiction.

La Figure 1 représente une régression linéaire à deux composantes (Tukey, 1972) des résidus calculés (écart entre les accélérations calculée et observée) pour chaque combinaison de séisme (droites de pente négative) et de station (droites de pente positive) : le résidu est indiqué par l'axe des ordonnées (le logarithme de l'accélération mesuré moins le logarithme de l'accélération prédite). La méthode de construction de ce type de diagramme est expliquée par Tukey (1972). Par exemple, l'intersection de la droite «Station NCR » et de la droite «Séisme A » donne la valeur approximative de la différence entre le logarithme de l'accélération mesurée et le logarithme de l'accélération calculée pour l'enregistrement à la station de Nocera Umbra du séisme du 26 septembre 1997 à 00:33. Il faut cependant remarquer que l'axe des ordonnées ne donne pas les résidus exacts, étant donné que cette régression linéaire à deux composantes ne fait qu'indiquer une tendance.

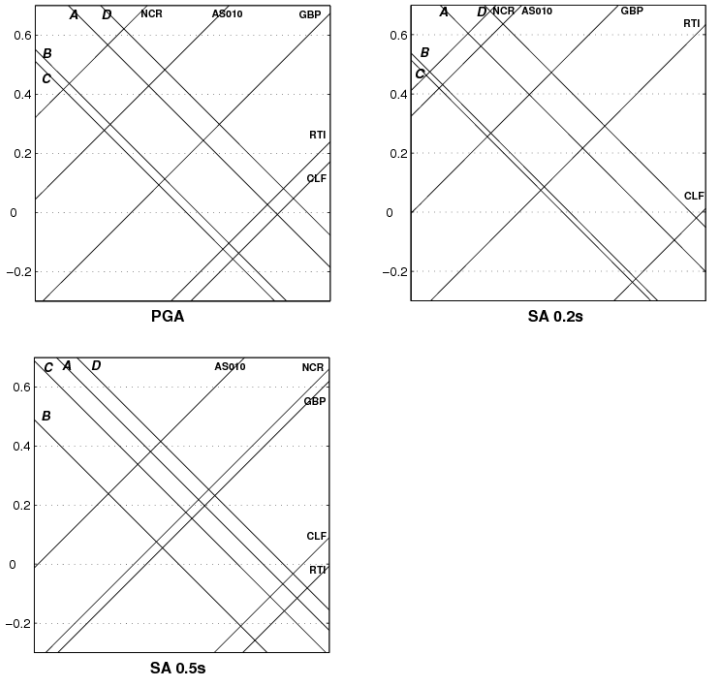

Figure 1. *Régression linéaire à deux composantes pour les données issues de la séquence d'Ombrie de 1997-1998. Les nombres sur l'axe des ordonnées indiquent approximativement les résidus par rapport aux relations de prédiction de Ambraseys et al. (2005).*

3. La séquence sismique des Saintes (Guadeloupe) en 2004-2005

Comme pour la séquence d'Ombrie, la série des Saintes consiste en plusieurs chocs de magnitude similaire et enregistrés à des distances comparables. Les données et les paramètres associés ont été estimés par Douglas et al. (2006). Le Tableau 3 reprend les données utilisées pour l'étude.

Séismes			A	B	C	D	E	F
JJ/MM/AA			21/11/04	21/11/04	21/11/04	27/11/04	02/12/04	14/02/05
HH:MM			11:41	13:37	18:53	23:44	14:47	18:05
Mw			6.3	5.3	5.4	4.9	5.0	5.8
Station	**Code**	**Classe**			**de**			
Saint-Claude Belfond	GJYA	Rock	34	30	24	39	35	26
Ecole Pigeon	PIGA	Rock	50	46	41	55	51	42
Institut Pasteur Abymes	IPTA	Rock	51	48	45	58	53	47
Stade Morne à l'Eau	MESA	Soft soil	62	59	57	69	65	59
Radar Meteo France	MOLA	Rock	62	60	59	70	65	61
Saint François	SFGA	Rock	62	61	62	69	65	63

Tableau 3. *Données issues de la séquence des Saintes en 2004-2005. Voir la légende du Tableau 1 pour la signification des abréviations.*

Douglas et al. (2006) examinent quantitativement la capacité de neuf séries d'équations récentes de prédiction du mouvement du sol, à partir de séismes crustaux superficiels enregistrés aux Antilles (comme ceux du Tableau 3). Il en ressort qu'aucun des modèles testés ne prédit avec précision les niveaux d'accélération observés, qui sont généralement plus faibles et plus variables que ceux prédits par les modèles. Les lois d'atténuation de Ambraseys et al. (2005) sont cependant celles qui sont les plus appropriées parmi tous les modèles examinés : ce modèle de prédiction a donc été utilisé ici aussi. Le Tableau 4 donne les résultats de l'analyse de variance pour les données de la séquence des Saintes.

Contrairement aux données de la séquence d'Ombrie, les résultats sur les données des Saintes montrent que les effets de source non modélisés contribuent en majorité à la variance globale et que cet effet est très significatif. Douglas et al. (2006) étudient les niveaux d'accélération enregistrés suite à deux paires de répliques du séisme des Saintes : cette analyse est réalisée en calculant les ratios entre les spectres de réponse de plusieurs enregistrements communs à deux répliques, après avoir corrigé des différences mineures au niveau de la magnitude ou de la distance. Les ratios calculés mettent donc en évidence l'effet de la source sur la variabilité du niveau d'accélération, étant donné que les effets de site ont été retirés à travers le calcul du ratio. Il a été remarqué que, pour une paire d'événements, la source causait des différences d'un facteur 10 au niveau du mouvement enregistré pour certaines périodes et certaines stations. Les effets de source très significatifs mis en évidence dans la présente étude confirment donc ce résultat. Le Tableau 4 montre aussi que les effets de site ont peu d'influence à haute fréquence, alors qu'à basse fréquence, les effets de site non modélisés contribuent de manière significative à la variance globale des niveaux d'accélération. Cinq des six stations sélectionnées

sont classées comme des stations au rocher, donc on pouvait effectivement s'attendre à des effets de site moins significatifs que pour les données d'Ombrie, où les stations sont situées sur des sites avec des formations géologiques plus hétérogènes. Sur la Figure 2, toutes les droites représentant les stations semblent très proches, sauf une (GJYA), ce qui montre que les effets de site sont globalement bien modélisés pour ces données. A l'inverse, la dispersion importante entre les droites représentant les séismes confirme les importants effets de source mis en évidence par l'analyse de variance.

Paramètre	**Effet**	**Σ**	**d.l.**	**σ^2**	**Ratio**	
PGA	Stations	0.3725	5	0.0621	1.5	
	Séismes	4.4501	5	0.7417	18.5	***
	Résidu	1.0037	25	0.0401		
	Total	5.8263	35			
SA 0.2 s	Stations	0.7439	5	0.1240	3.1	*
	Séismes	4.9627	5	0.8271	20.8	***
	Résidu	0.9922	25	0.0397		
	Total	6.6988	35			
SA 0.5 s	Stations	0.5026	5	0.0838	2.4	
	Séismes	5.0288	5	0.8381	24.1	***
	Résidu	0.8703	25	0.0348		
	Total	6.4017	35			
SA 1.0 s	Stations	0.8659	5	0.1443	5.9	**
	Séismes	4.3639	5	0.7273	29.6	***
	Résidu	0.6150	25	0.0246		
	Total	5.8447	35			
SA 2.0 s	Stations	2.3950	5	0.3992	19.9	***
	Séismes	6.2325	5	1.0388	51.7	***
	Résidu	0.5017	25	0.0201		
	Total	9.1292	35			

Tableau 4. *Résultats de l'analyse de variance pour les données issues de la séquence des Saintes de 2004-2005. Voir la légende du Tableau 2 pour la signification des abréviations.*

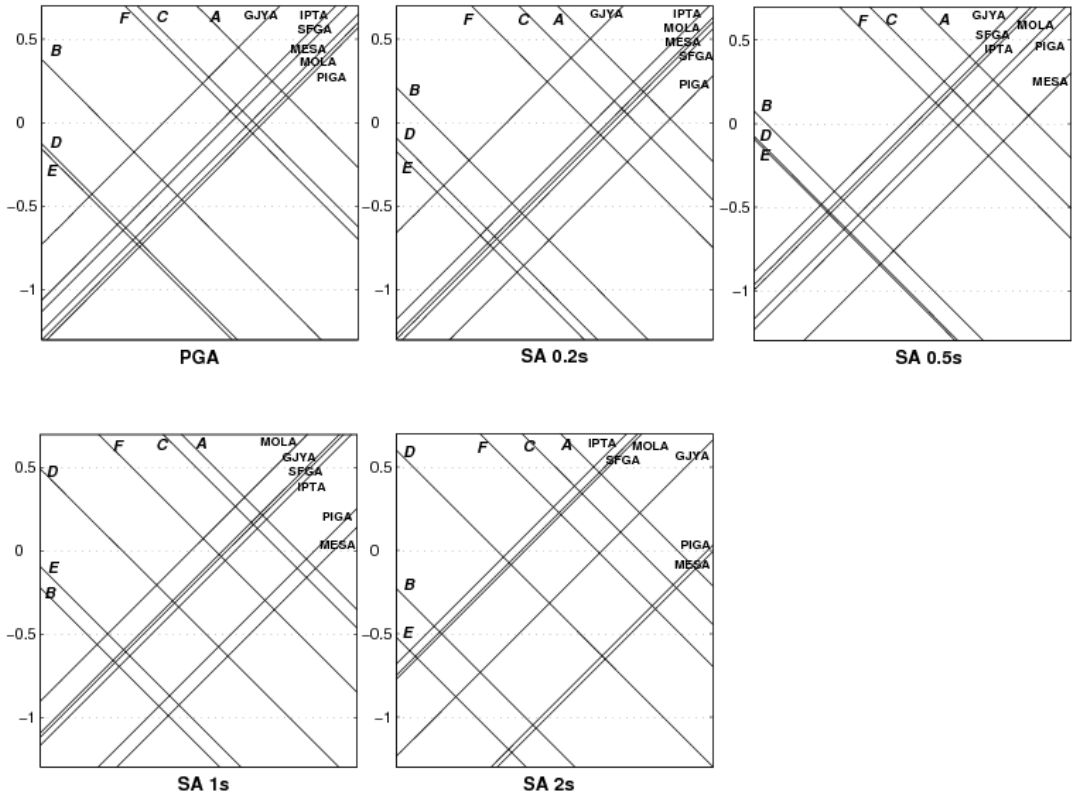

Figure 2. *Régression linéaire à deux composantes pour les données issues de la séquence des Saintes de 2004-2005. Les nombres sur l'axe des ordonnées indiquent approximativement les résidus par rapport aux relations de prédiction de Ambraseys et al. (2005).*

5. Conclusion

Ce bref article a proposé une méthode quantitative simple pour décomposer la variabilité dans les niveaux d'accélération en celle attribuable aux effets de site et celle imputable aux effets de source. Cette méthode a été appliquée à quatre séries de données. Nous avons remarqué que pour deux de ces séquences (Ombrie 1997-1998 et Molise 2002-2003), les effets de site non modélisés sont plus significatifs que les effets de source pour expliquer la variabilité observée dans les résidus. Néanmoins, pour les autres séries étudiées (Kocaeli 1999 et Les Saintes 2004-2005) les effets de source non modélisés contribuent le plus à la variance globale. Par ailleurs, les diagrammes de régression à deux composantes, introduits par Tukey (1972), ont permis de démontrer graphiquement les deux effets à l'origine des incertitudes dans l'évaluation des niveaux d'accélération.

6. Remerciements

Cette étude a été financée par des projets de recherche du BRGM. Nous avons utilisé des données provenant de réseaux accélérométriques maintenus par ENEL (maintenant par SSN), ENEA, l'Observatoire de Kandilli

(Université de Bogazici), le BRGM et le RAP-IPGP. Les réseaux accélérométriques de Guadeloupe sont opérés par le BRGM et l'Institut de Physique du Globe de Paris (IPGP), sous l'égide du Réseau Accélérométrique Permanent (RAP) de France. Le centre de données du RAP est basé au Laboratoire de Géophysique Interne et de Tectonophysique, à Grenoble. Nous sommes très reconnaissants envers le personnel de ces organismes qui gèrent les réseaux accélérométriques qui ont fourni les données, sans lesquelles cette étude aurait été impossible.

7. Références

Ambraseys N.N., Douglas J., Sarma S.K., Smit P.M., "Equations for the estimation of strong ground motions from shallow crustal earthquakes using data from Europe and the Middle East: Horizontal peak ground acceleration and spectral acceleration", *Bulletin of Earthquake Engineering* 3 (1), 2005, p. 1-53.

Ambraseys N.N., Free M.W., "Surface-wave magnitude calibration for European region earthquakes", *Journal of Earthquake Engineering* 1 (1), 1997, p. 1-22.

Baturay M.B., Stewart J.P., "Uncertainty and bias in ground-motion estimates from ground response analyses", *Bulletin of the Seismological Society of America* 93 (5), octobre 2003, p. 2025-2042.

Boore D.M., Joyner W.B., Fumal T.E., "Equations for estimating horizontal response spectra and peak acceleration from western North American earthquakes: A summary of recent work", *Seismological Research Letters* 68 (1), janvier-février 1997, p. 128-153.

Castellaro S., Mulargia F., Kagan Y.Y., "Regression problems for magnitudes", *Geophysical Journal International* 165, 2006, p. 913-930.

Castro R.R., Pacor F., Bindi D., Franceshina G., Luzi L., "Site response of strong ground motion stations in the Umbria, central Italy, region", *Bulletin of the Seismological Society of America* 94 (2), 2004, p. 576-590.

Douglas J., "Earthquake ground motion estimation using strong-motion records: A review of equations for the estimation of peak ground acceleration and response spectral ordinates", *Earth-Science Reviews* 61 (1-2), 2003, p. 43-104.

Douglas J., "Use of analysis of variance for the investigation of regional dependence of strong ground motions", *Proceedings of Thirteenth World Conference on Earth Engineering*, 2004, Paper n°29.

Douglas J., Bertil D., Roullé A., Dominique P., Jousset P., "A preliminary investigation of strong-motion data from the French Antilles", *Journal of Seismology* 10 (3), 2006, p. 271-299.

Douglas J., Gehl P., "Investigating strong ground-motion variability using analysis of variance", *Soil Dynamics and Earthquake Engineering*, 2006, soumis.

Fisher R.A., *"Statistical methods, Experimental Designs, and Scientific Inference"*, Oxford, Oxford University Press, 1990.

Joyner W.B., Boore D.M., "Peak horizontal acceleration and velocity from strong-motion records including records from the 1979 Imperial Valley, California, earthquake", *Bulletin of the Seismological Society of America* 71 (6), décembre 1981, p. 2011-2038.

Marra F., Azzara R., Belluci F., Caserta A., Cultrera G., Mele G., Palombo B., Rovelli A., Boschi E., "Large amplification of ground motion at rock sites within a fault zone in Nocera Umbra (central Italy)", *Journal of Seismology* 4 (4), 2000, p. 543-554.

Moroney M.J., *"Facts from figures"*, Penguin Books, 2nd Edition, 1990.

Tukey J.W., "Some graphic and semigraphic displays", Bancroft T.A. (Ed.), *Statistical Papers in Honor of George W. Snedecor*, The Iowa State University Press, Ames, Iowa, 1972, p. 293-316.