Predicting Site-to-Site Variability in Ground Motion across Italy

POLITECNICO MILANO 1863

Authors :
Mattéo Bevilacqua, Camilla Bocciolone, Martina Maria Calandro, Mattia De Bartolomeis, Pınar Nur Özkaplan

Politecnico di Milano
Academic Year 2024–2025
Applied Statistics
Group 11

Academic Supervisor:
Prof. Alessandra Menafoglio
With advice from INGV researchers :
Dr. Sara Sgobba, Dr. Giulio Brunelli, Dr. Giovanni Lanzano

INTRODUCTION

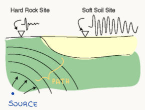
ABSTRACT

Understanding and predicting the spatial variability of earthquake ground motion is a critical challenge in seismic hazard assessment. This study focuses on modeling the *site-to-site variability term* (δS2S) of Ground Motion Models using records from more than 900 Italian stations. δS2S was measured across multiple spectral periods (ranging from 0.01s 10s) capturing frequency-dependent amplification effects. We tested several statistical techniques to identify the primary factors influencing δS2S and to build a nationwide prediction model.

BACKGROUND

Recent studies show that recorded ground-motion amplitudes vary systematically with local site conditions. In geophysical Ground Motion Models (GMMs), this behaviour is captured by a specific term, called δS2S :

$$\text{Log}\,(IM_{es}(T)) = \mu_{es}(T) + \overbrace{\delta L2L\ (T)}^{\text{SOURCE}} + \overbrace{\delta P2P_s}^{\text{PATH}} + \delta S2S_s(T) + \varepsilon_s\ (T)$$



Proxies such as Vs30 lithostratigraphic class (rock/soil layers), and surface slope have already been used to predict δS2S on regional scales. Recent research leverages nationwide datasets and integrates classical statistical approaches with spatial modeling techniques to investigate the geophysical controls on δS2S and assess their predictive potential.

The physical meaning of our target and covariates is reported in the following table :

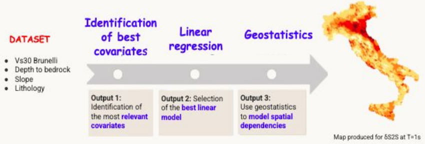
δS2S SA: site-to-site residual	It represents the deviation of the observed ground motion at the site from the median values empirically predicted by the Ground Motion Model. δS2S can refer to different measures of ground motion severity. In this work, it specifically refers to Spectral Acceleration (SA), a key period-dependent parameter in Earthquake Engineering.
Vs30	Average shear-wave velocity in the top 30 m of soil, representative of site stiffness.
Slope	Surface topographic slope, which can influence seismic wave amplification through scattering and basin edge effects.
Depth to bedrock	Depth of the layer of hard rock (bedrock) under the soft surface layers. Greater depth to bedrock can lead to stronger amplification of ground motion.
Lithology	Type and physical characteristics of rocks or sediments that make up the Earth's surface.

OBJECTIVES

This project aimed to develop a statistical model to predict δS2S across Italy. The main objectives were:

- ❑ To achieve accurate prediction of δS2S;
- ❑ To quantify the relative contribution of various geological factors to its spatial variability;
- ❑ To evaluate the potential benefits of incorporating spatial interpolation using Kriging models;
- ❑ To assess the added value of Functional Principal Component Analysis (FPCA) compared to models based on conventional covariates.

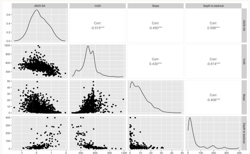
WORKFLOW



METHODOLOGY & RESULTS

EXPLORATORY DATA ANALYSIS

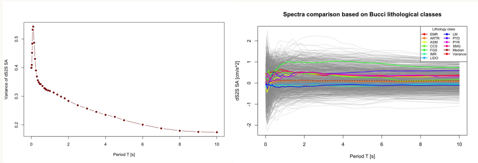
Marginal distributions and scatter plots for predictors and δS2S highlighted the correlations we expected to have:



As Vs30 ↓, softness of the soil ↑ and Amplification ↑
As slope ↓ Amplification ↓
As depth to bedrock ↑ Amplification ↑

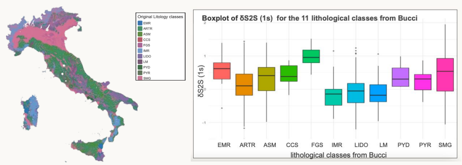
Focusing on δS2S:

- The variability differs across periods, showing higher values at short ones, implying potential heterogeneous challenges in prediction.

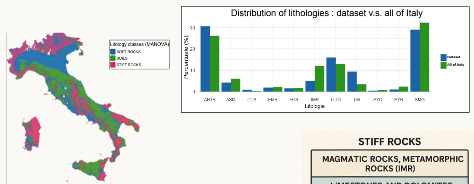


MANOVA

- Used in order to assess the impact of lithology classes on δS2S.
- The results show that δS2S present statistical significant dependence on lithology when it is tested jointly over the period (p-value=0 in MANOVA one-way) and singularly with each period (p-value=0 in ANOVA one-way)



- By analysing Bonferroni confidence intervals and observing some classes were under represented in Italy we conducted further analysis to reduce the number of classes.



- Statistical and lithological analysis led us to 3 different regroupings. The relevance of these groupings was confirmed by a final MANOVA.

STIFF ROCKS

- MAGMATIC ROCKS, METAMORPHIC ROCKS (IMR)
- LIMESTONES AND DOLOMITES (LIDO)
- LIMESTONE-MARLY (LM)
- EFFUSIVE MAGMATIC ROCKS (EMR)

SOFT ROCKS

- VOLCANIC PYROCLASTIC DEPOSITS (PYD)
- PYROCLASTIC ROCKS (PYR)
- SOILS WITH MIXED GRANULOMETRY (SMG)
- FINE GRAINED SOILS (B1 class) (FGS)

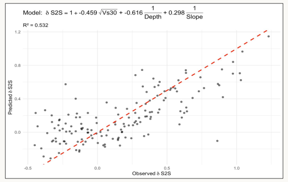
SOILS

- SANDSTONES, PELITIC-ARENACEOUS COMPLEXES, CALCAREOUS ARENACEOUS COMPLEXES, QUNGLOMERATES, TRAVERTINE, QUARTZARENESTR)
- ARGILLITES, SILTSTONES, MARLS (ASM)
- CEMENTED COARSE SEDIMENTS (CCS)

LINEAR REGRESSION

We performed feature selection by:

- Retaining only variables with correlation |r| > 0.30 with the target
- Generating transformations (log, sqrt, inverse, square, exponential) and pairwise interactions
- Applying manual forward selection based on p-values, AIC and BIC
- Final model includes 3 predictors out of 22 candidates, all with VIF < 10



GEOSTATISTICS

To account for the spatial structure of our data, kriging was employed:

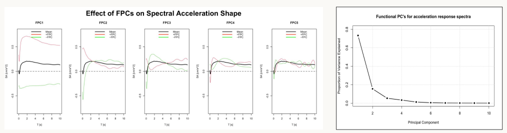
- Ordinary kriging was initially tested, but due to non-stationarity, results were unsatisfactory.
- Universal kriging was then applied, including covariates significantly correlated with dS2S (based on linear regression).
- For model validation, Leave-One-Out Cross-Validation (LOOCV) with Mean Squared Error (MSE) was performed. The results are summarized in the following table:

Variable to model	Covariates	LOO MSE	LOO MSE standard deviation
δS2S (PGA)	i. Vs30-Brunelli	0.2535	0.5038
	ii. (1/slope) + sqrt(Vs30_Brunelli)	0.2482	0.4997
	iii. (1/slope) + sqrt(Vs30_Brunelli) + (1/depth_to_bedrock)	0.2482	0.4997
	iv. (1/slope) + sqrt(Vs30_Brunelli) + lito classes	0.2538	0.5001
	v. (1/slope) + sqrt(Vs30_Brunelli) + geographical location	0.2519	0.5034
δS2S (1s)	i. Vs30-Brunelli	0.3576	0.5549
	ii. (1/slope) + sqrt(Vs30_Brunelli)	0.3487	0.5915
	iii. (1/slope) + sqrt(Vs30_Brunelli) + (1/depth_to_bedrock)	0.3487	0.5915
	iv. (1/slope) + sqrt(Vs30_Brunelli) + litoclasses	0.3059	0.5578
	v. (1/slope) + sqrt(Vs30_Brunelli) + geographical location	0.3537	0.5965

- The above table shows that best results are obtained when including as covariates vs30_Brunelli, slope and the lito classification ! Adding other covariates did not strike significantly better results ; hence, we followed the principle of Ockam's razor and chose to keep our final model as is.

FUNCTIONAL PRINCIPAL COMPONENT ANALYSIS

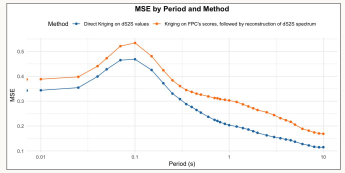
Next question : can we take advantage of correlation between time periods ?
⇒ compute FPCs



It seems we can !
The first four Functional Principal Components explain more than 97.5% of variability, we'll therefore try to predict FPC's and reconstruct the spectrum from these !

GEOSTATISTICS on FPC's

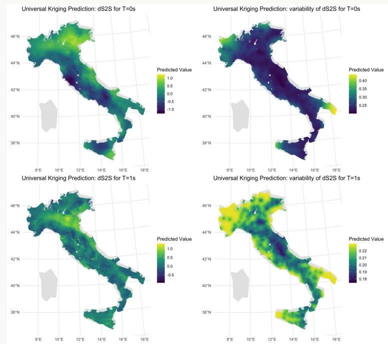
We then tried taking advantage of the time-period correlation and applied Kriging on functional Principal Components:



Key Takeaway :
Using FPC's unfortunately yielded less precise answers. Here are some possible explanations :

- Error amplification from conversion FPC's → time periods
- Bias with using a limited number of FPC's

MAPS OF δS2S



1. δS2S patterns → local amplification effects are frequency-dependent, which aligns with known site effects in seismology.
2. Spatial prediction is more reliable in data-rich areas (northern and central Italy), and less so in southern/extreme regions — typical in geostatistical modeling.
3. T = 1s predictions are more variable and perhaps more sensitive to local site conditions, which is expected as longer-period motions are more affected by deep soil and basin effects.

CONCLUSION

- ➢ Vs30 Brunelli turned out to be the most significant predictor for the model, as expected, considering that it is an estimated parameter that incorporates various factors, including lithological information.
- ➢ We observed that δS2S at longer periods shows a much stronger correlation with the main predictors compared to δS2S at shorter periods, which makes the latter inherently more difficult to predict.
- ➢ Geostatistical analysis confirmed a strong spatial dependence in δS2S. The use of Universal Kriging allowed for reliable spatial interpolation of δS2S, improving predictions.
- ➢ Functional Principal Component Analysis (FPCA) effectively summarized the δS2S spectral curves using a few dominant components. Nevertheless, kriging applied to FPCs yielded less satisfactory results, likely due to error amplification during the spectral reconstruction phase and the intrinsic limitations of the approach.