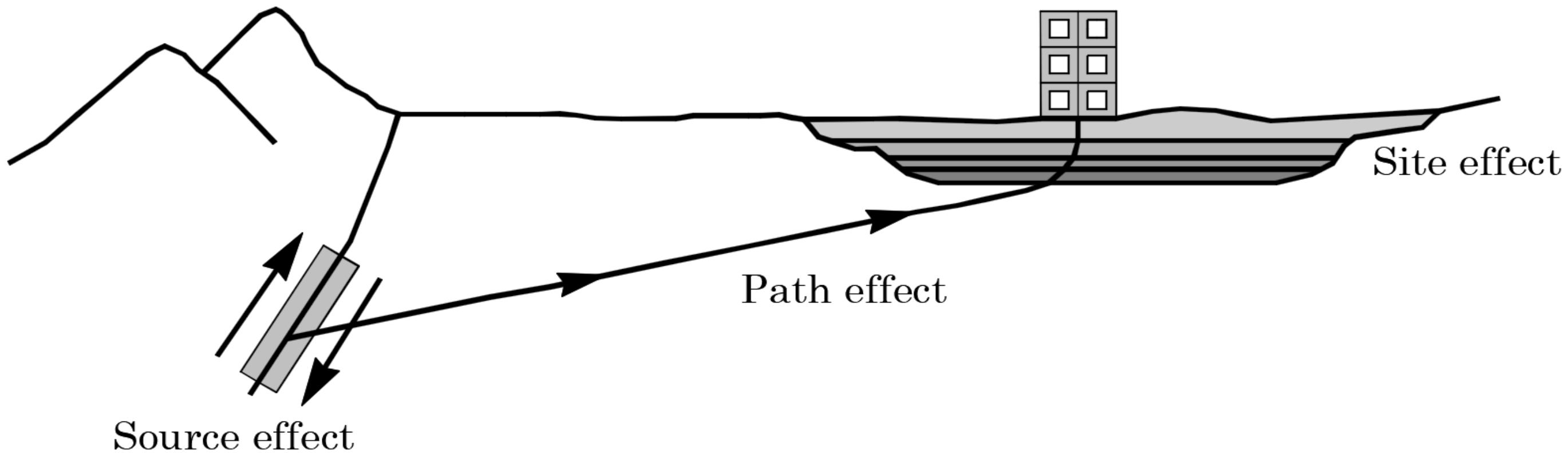
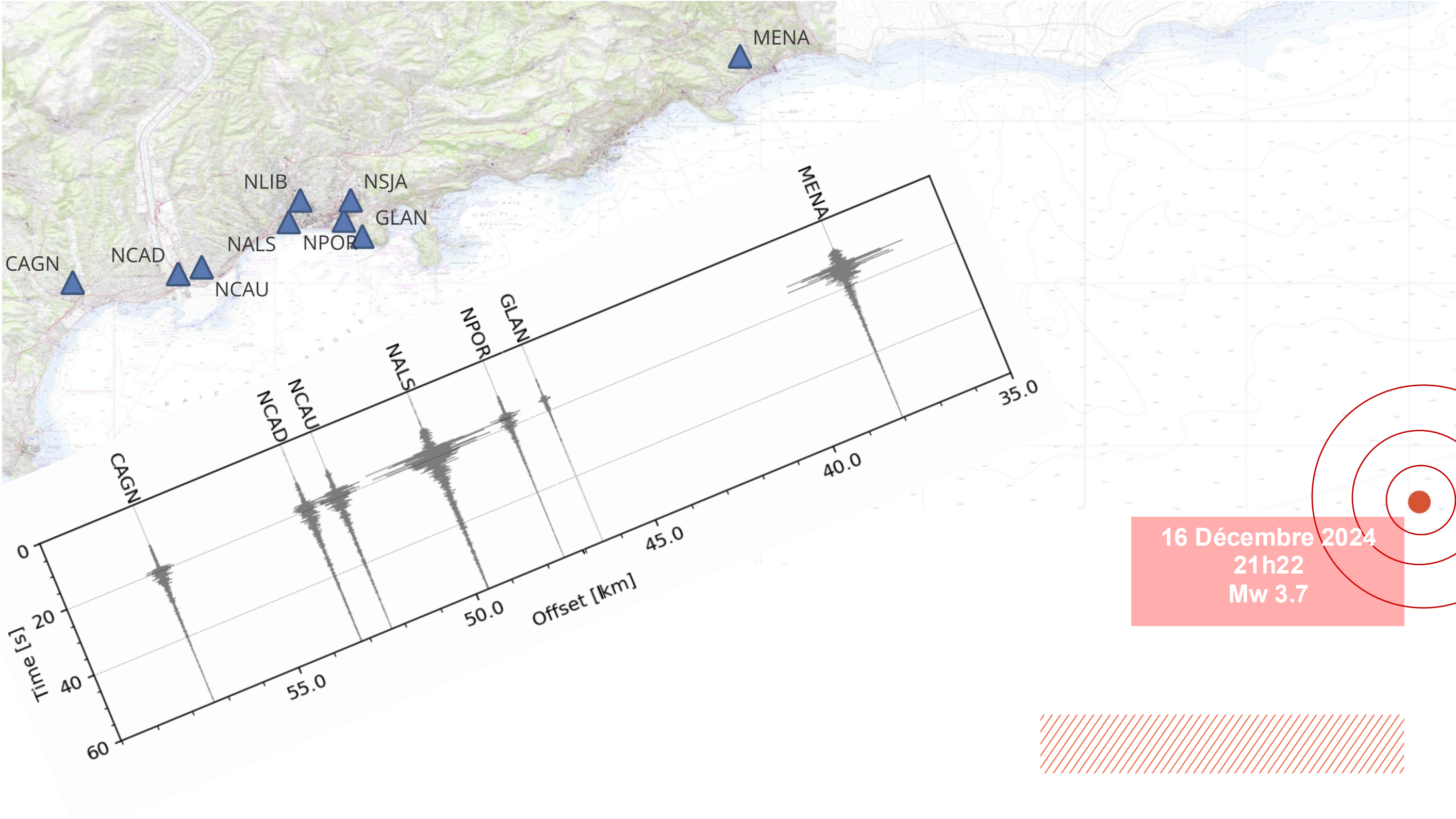


Prediction of shear modulus for clayey soils using ensemble machine learning method

Borderon Julien, Julie Régnier and Nathalie Dufour

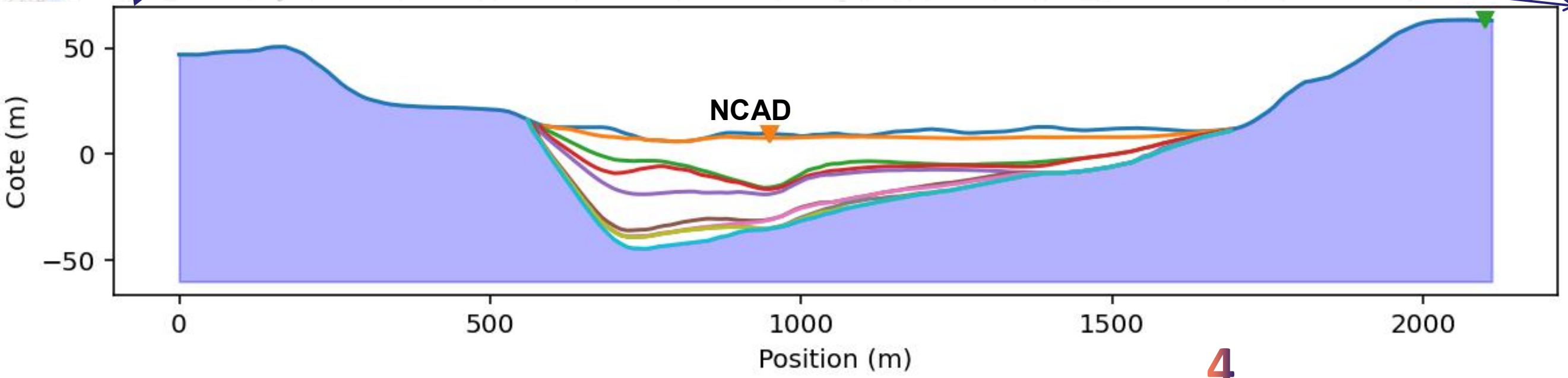
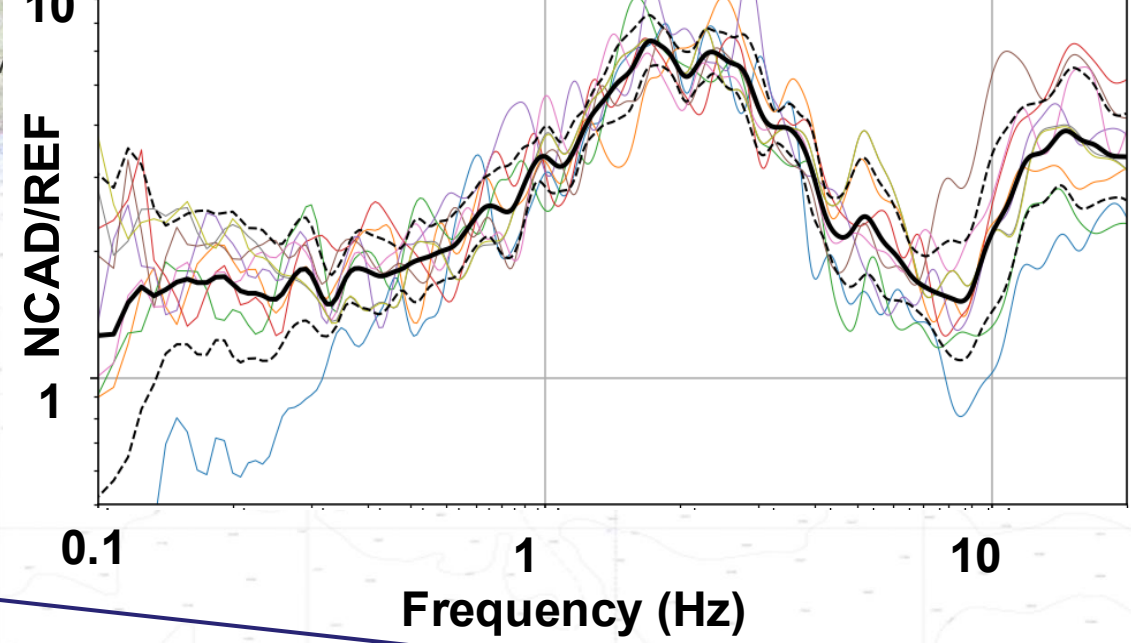
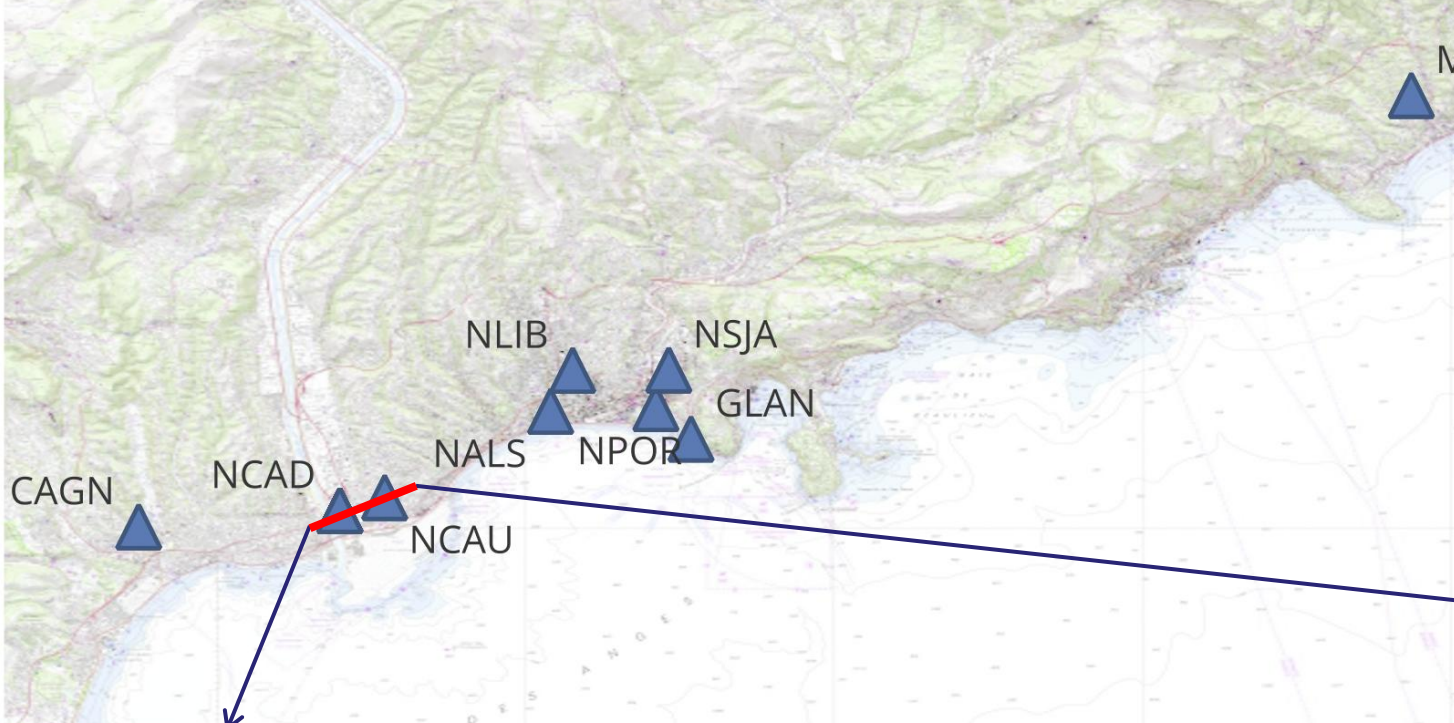
INTRODUCTION





16 Décembre 2024
21h22
Mw 3.7

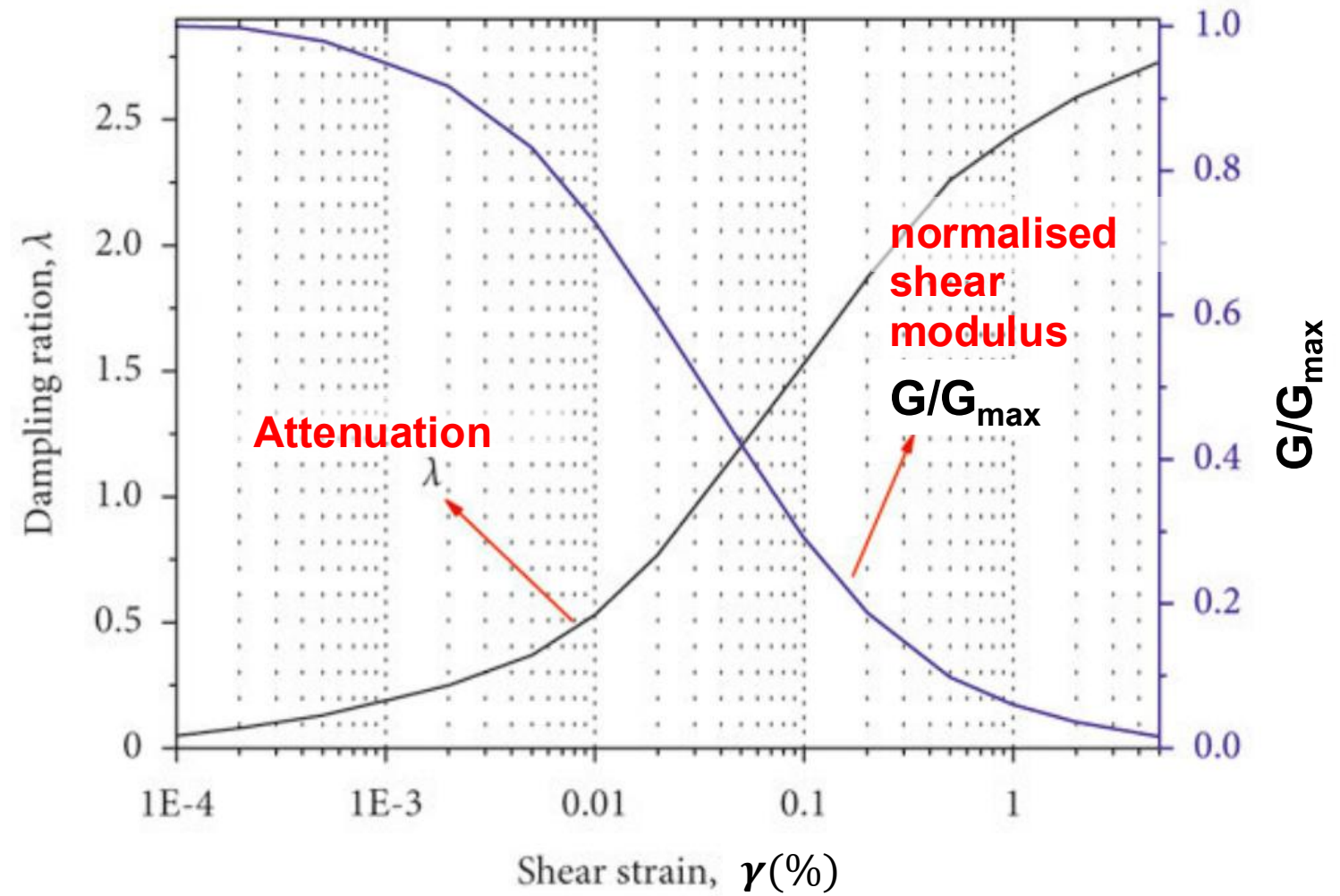
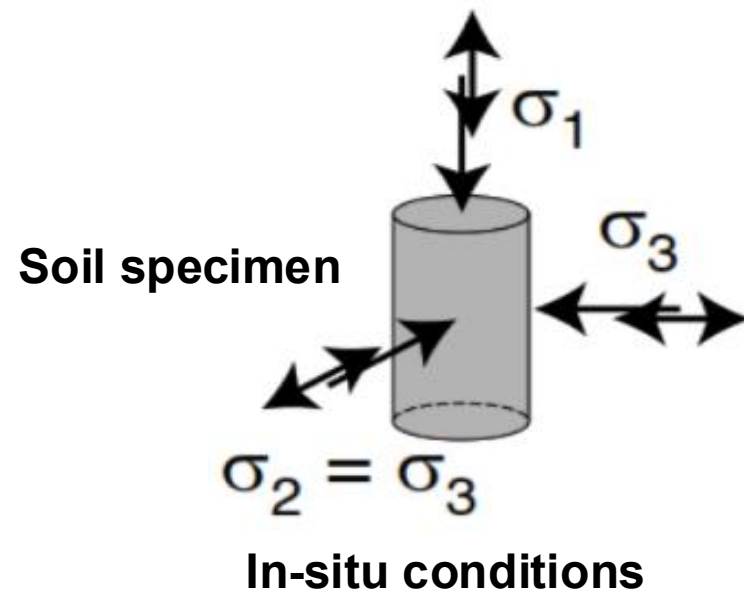
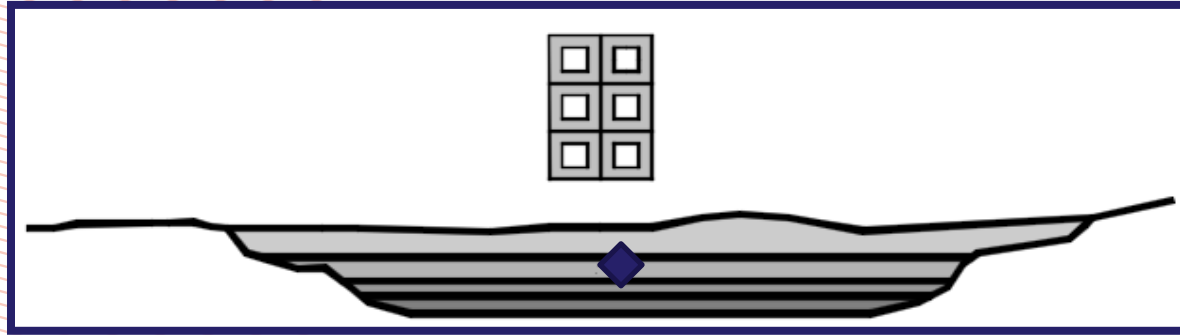




4



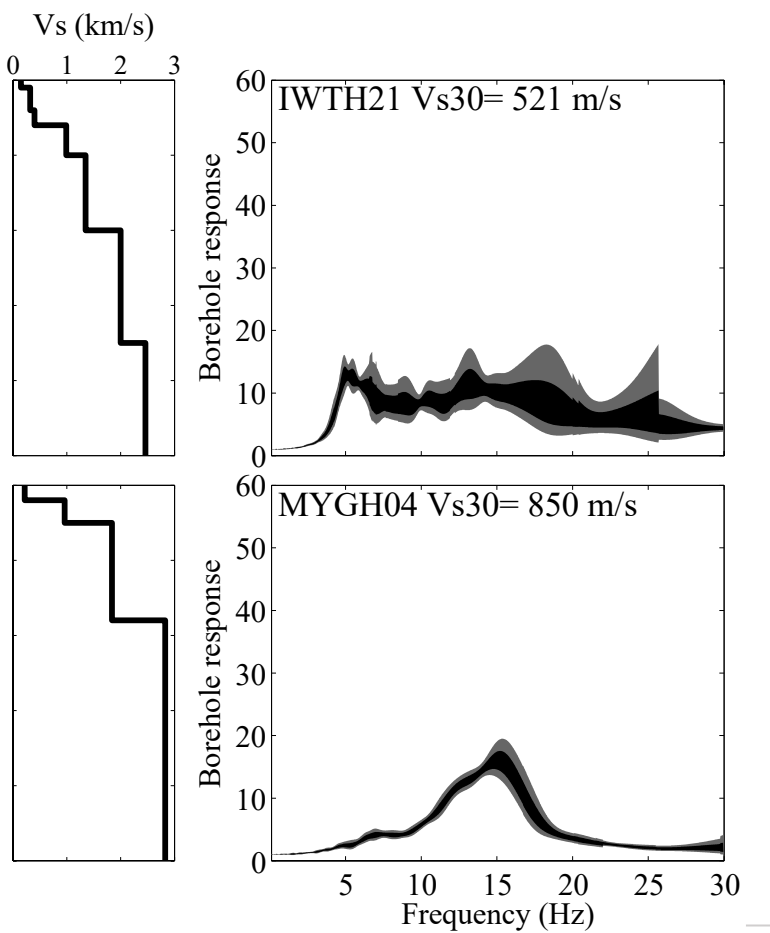
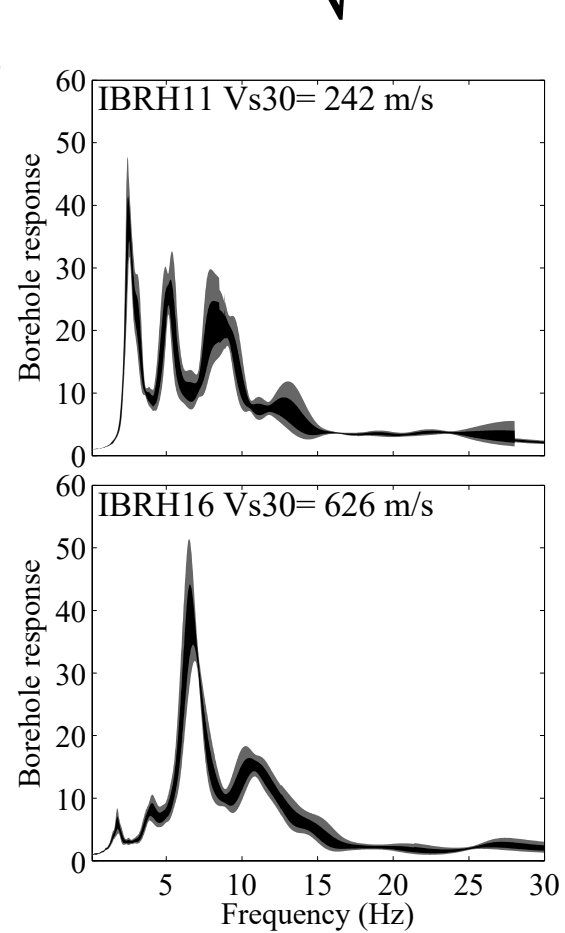
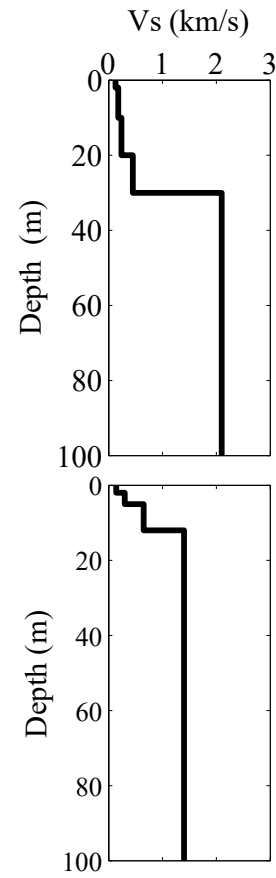
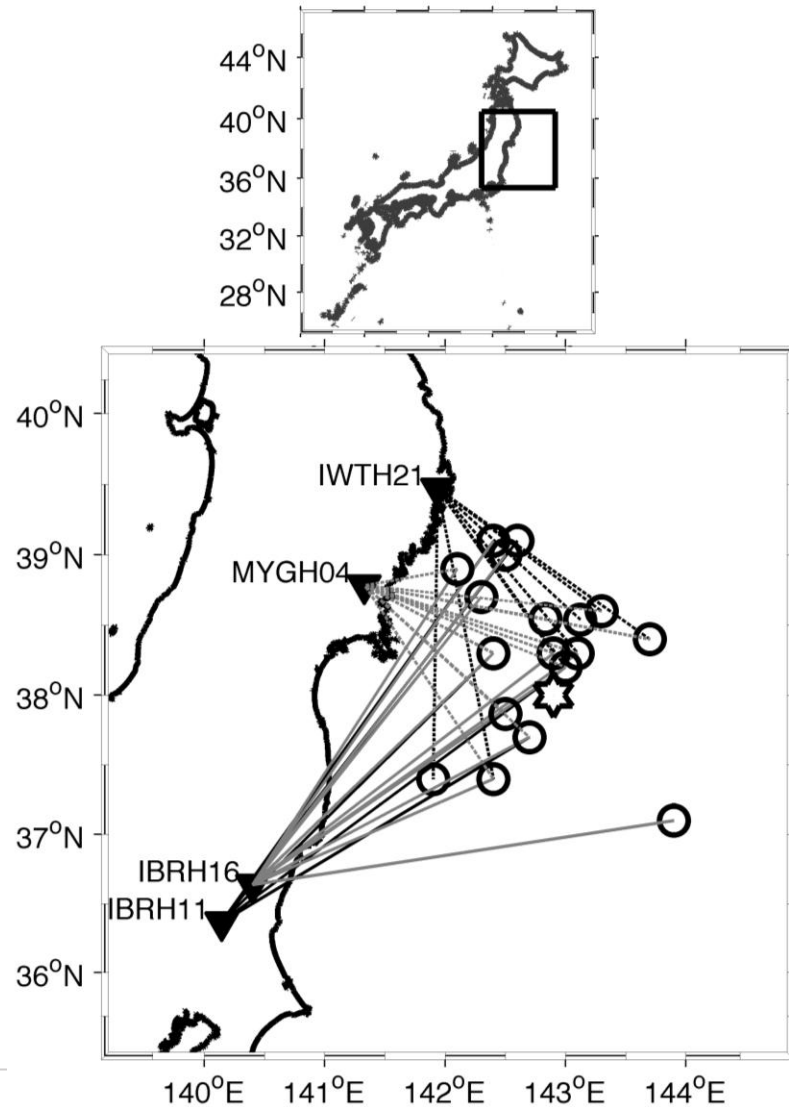
NON-LINEAR SOIL BEHAVIOUR



IMPACT ON SITE RESPONSE

$$V_s = \sqrt{\frac{G}{\rho}}$$

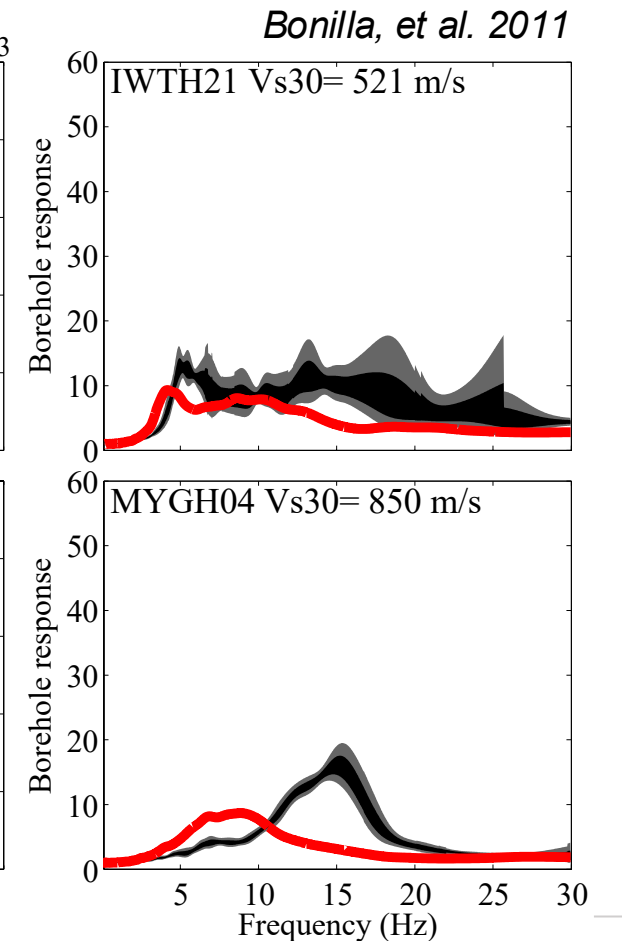
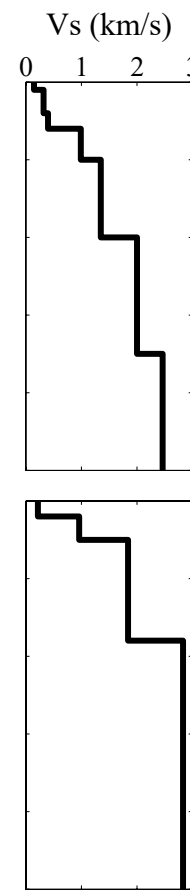
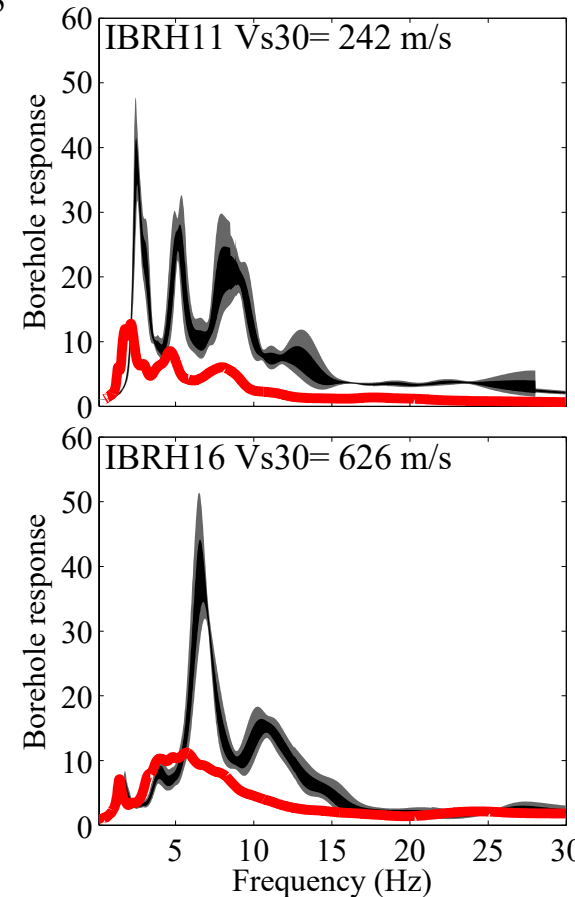
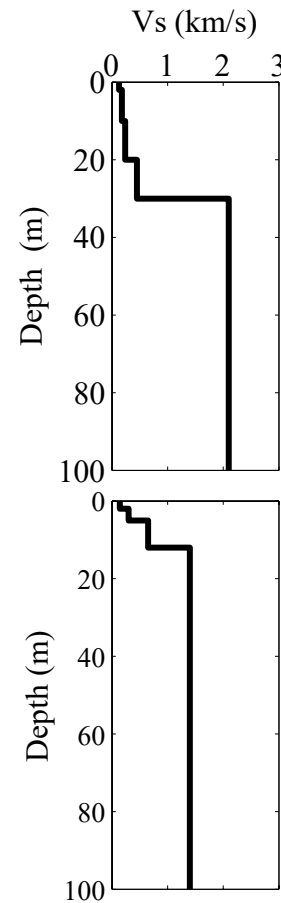
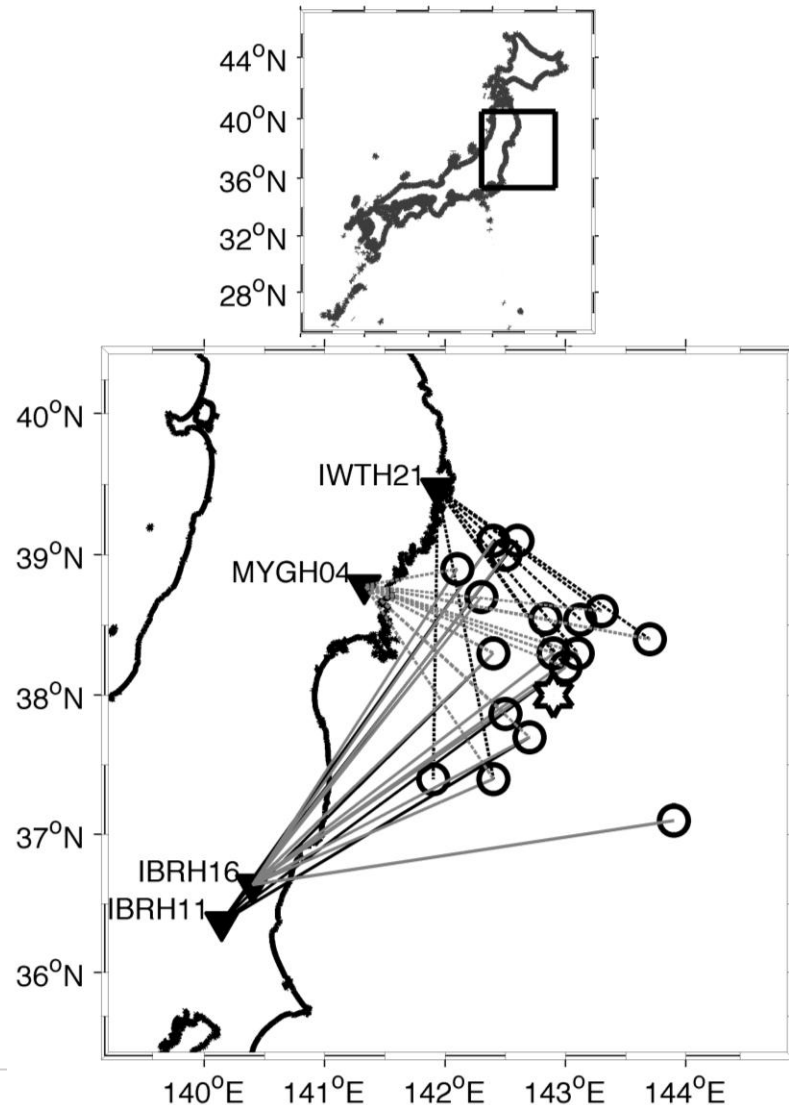
$$f_0 = \frac{V_s}{4H}$$



IMPACT ON SITE RESPONSE

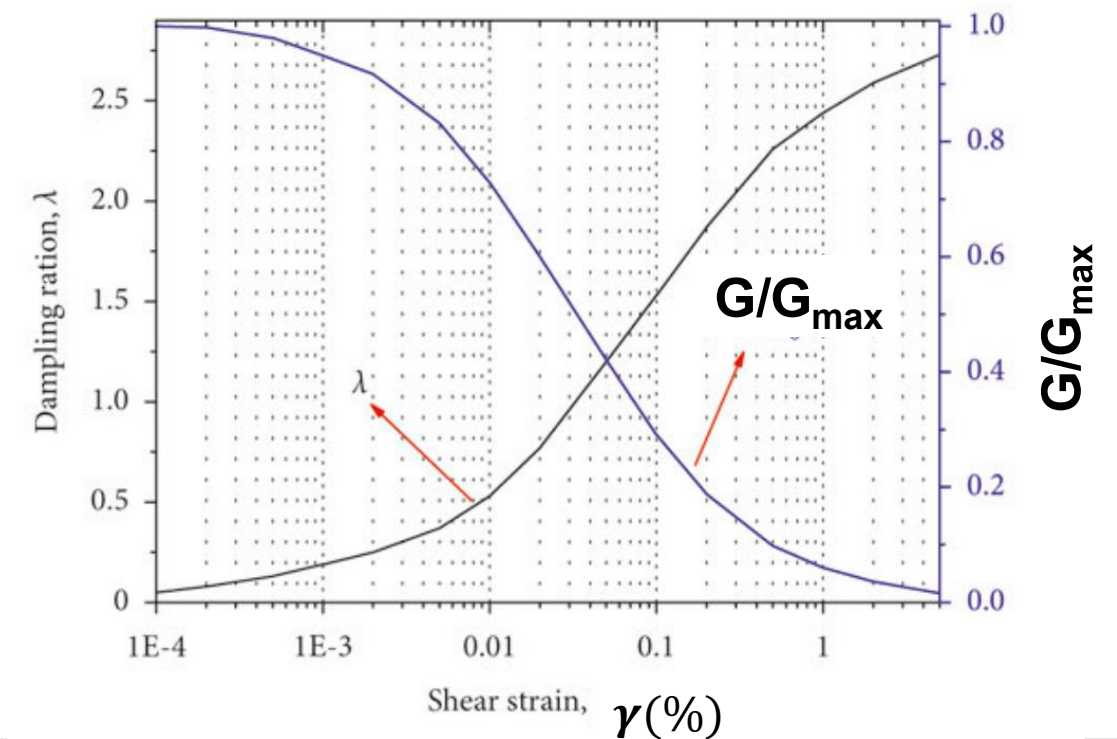
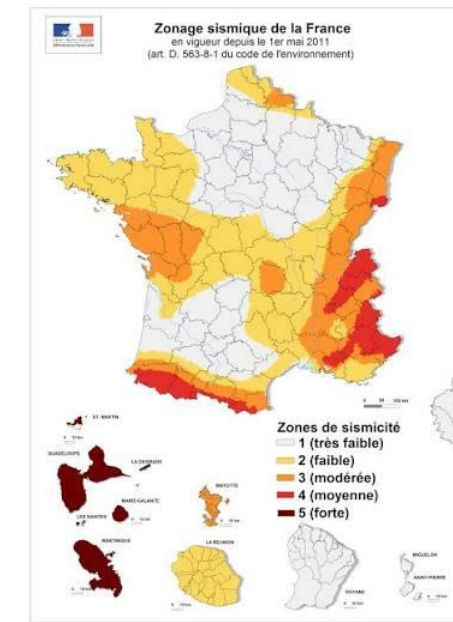
$$V_s = \sqrt{\frac{G}{\rho}}$$

$$f_0 = \frac{V_s}{4H}$$



HOW TO ACCOUNT FOR IT IN FRANCE

- Data: weak motions only
- Strong motions predictions
 - Numerical simulations
 - Non-linear soil properties
- Lab measurements
 - costly
 - Specimen no site behaviour
 - With large uncertainties
- **Need for prediction model of $G/G_{\max}$ and Damping curves**



EXISTING MODELS

- **Parametrics models**

- Most important features : Plastic Index (PI) and Confining Pressure (σ)
- Limitations : too many different equations, don't generalize well

- **Machine learning :**

Marine Clay : Base on Back-propagation Neural Network (Wu et al. 2022)

- No generic for all clay soil
- Not in open access
- Not usable : train on individual values not on entire test

Need for a new non-parametric model fo clayley soil

Functional forms		From
$\frac{G}{G_{max}} = K(\gamma, PI). \sigma_0^{-m(\gamma, PI) - m_0}$ $m(\gamma, PI) - m_0 = 0.272 \left[1 - \tanh \left(\ln \left(\frac{0.000556}{\gamma} \right)^{0.4} \right) \right] . e^{-0.0145 PI^{1.3}}$ $K(\gamma, PI) = 0.5 \left[1 + \tanh \left(\ln \left(\frac{0.000102 + n(PI)}{\gamma} \right)^{0.492} \right) \right]$ $n(PI) = \begin{cases} 0 & PI = 0 \\ 3.37 * 10^{-6} PI^{1.404} & 0 < PI \leq 15 \\ 7 * 10^{-7} PI^{1.976} & 15 < PI \leq 70 \\ 2.7 * 10^{-5} PI^{1.115} & PI > 70 \end{cases}$		Ishibashi & Zhang (1993)
$\frac{G}{G_{max}} = \frac{1}{1 + \left(\frac{\gamma}{\gamma_r} \right)^a}$ with $\gamma_r = (\Phi_1 + \Phi_2 . PI . OCR^{\Phi_3}) . \sigma_0'^{\Phi_4}$ $a = \Phi_5$ Where σ_0' is in atm		
Clayey soil	Mean	Variance
Φ_1	3.52E-02	5.76E-06
Φ_2	7.07E-04	9.12E-09
Φ_3	3.69E-01	7.30E-03
Φ_4	2.97E-01	5.31E-04
Φ_5	9.50E-01	2.25E-04

Darendeli (2001)

$$\frac{G}{G_{max}} = \frac{1}{1 + \left(\frac{\gamma}{\gamma_r} \right)^a}$$

$$\gamma_r = \gamma_{r1} \cdot \left(\frac{\sigma'_m}{P_a} \right)^k$$

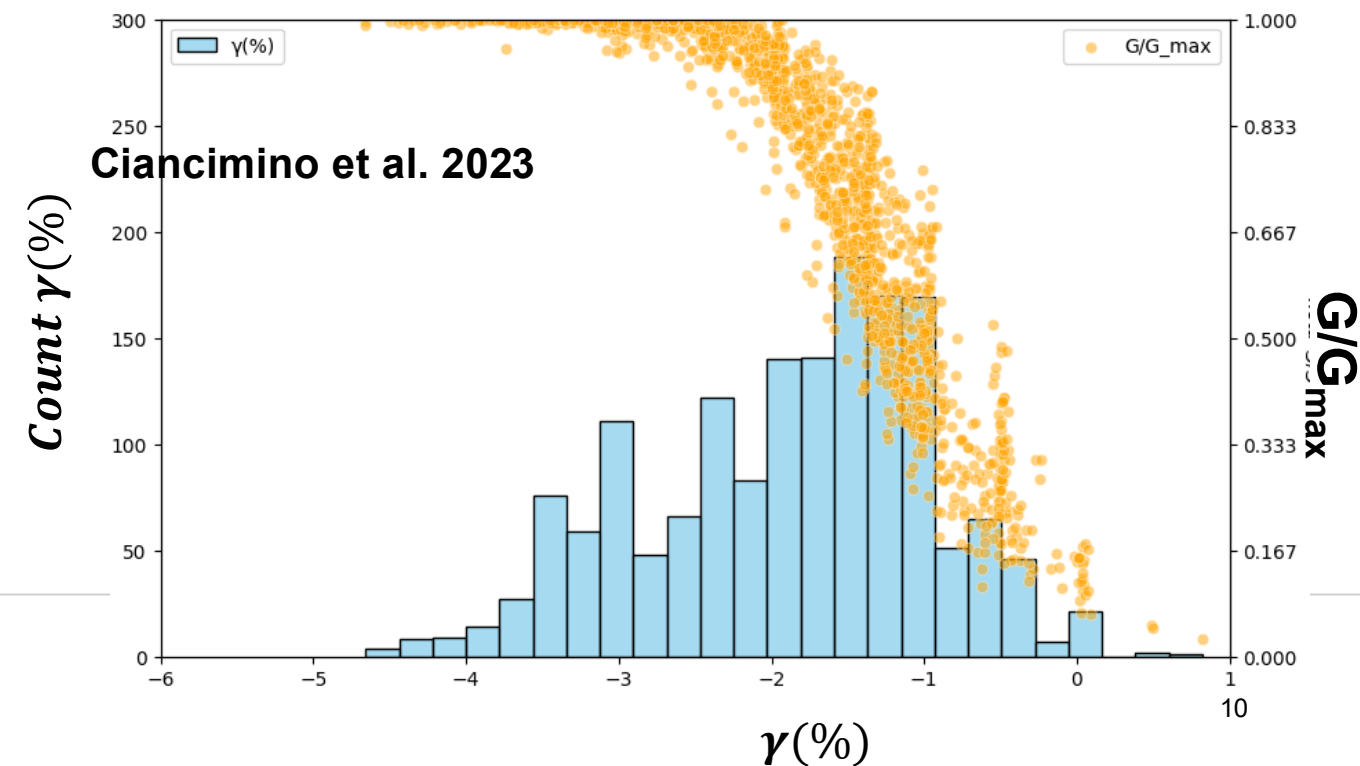
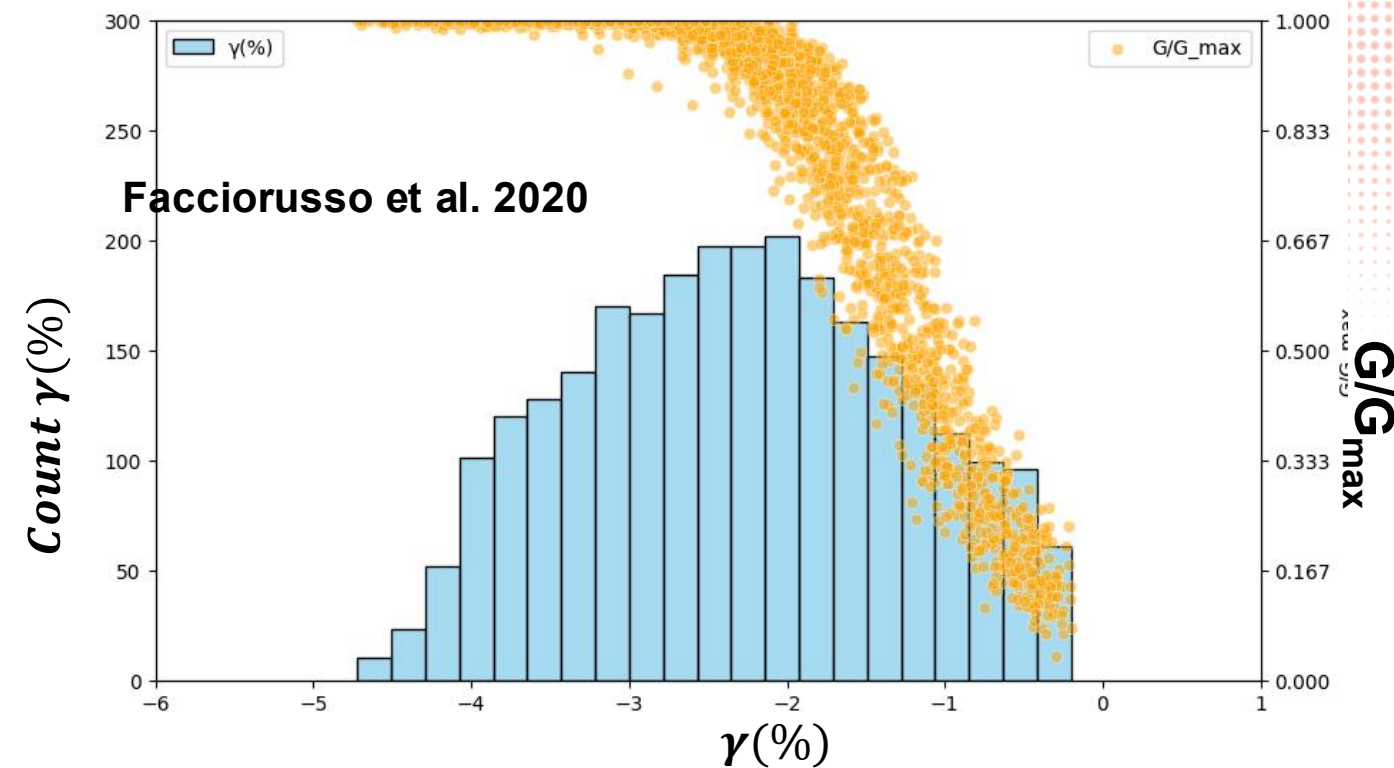
$$\sigma'_m = \frac{(\sigma'_v + \sigma'_h)}{3} = \sigma'_v \frac{(1 + 2 \cdot K'_0)}{3}$$

Zhang, et al. (2005)

AVAILABLE DATABASES

- **FACCIORUSSO : 170** undisturbed specimens (clay) tested (resonant shear tests and cyclic torsional shear test)
- **CIANCIMINO : 187** undisturbed soil specimens (clay, silt) tested (resonant shear tests, cyclic torsional shear test and cyclic double specimen direct simple shear tests)

For each specimen : 6 to 20 values



AVAILABLE DATABASES

Soil physical parameters and condition of lab tests

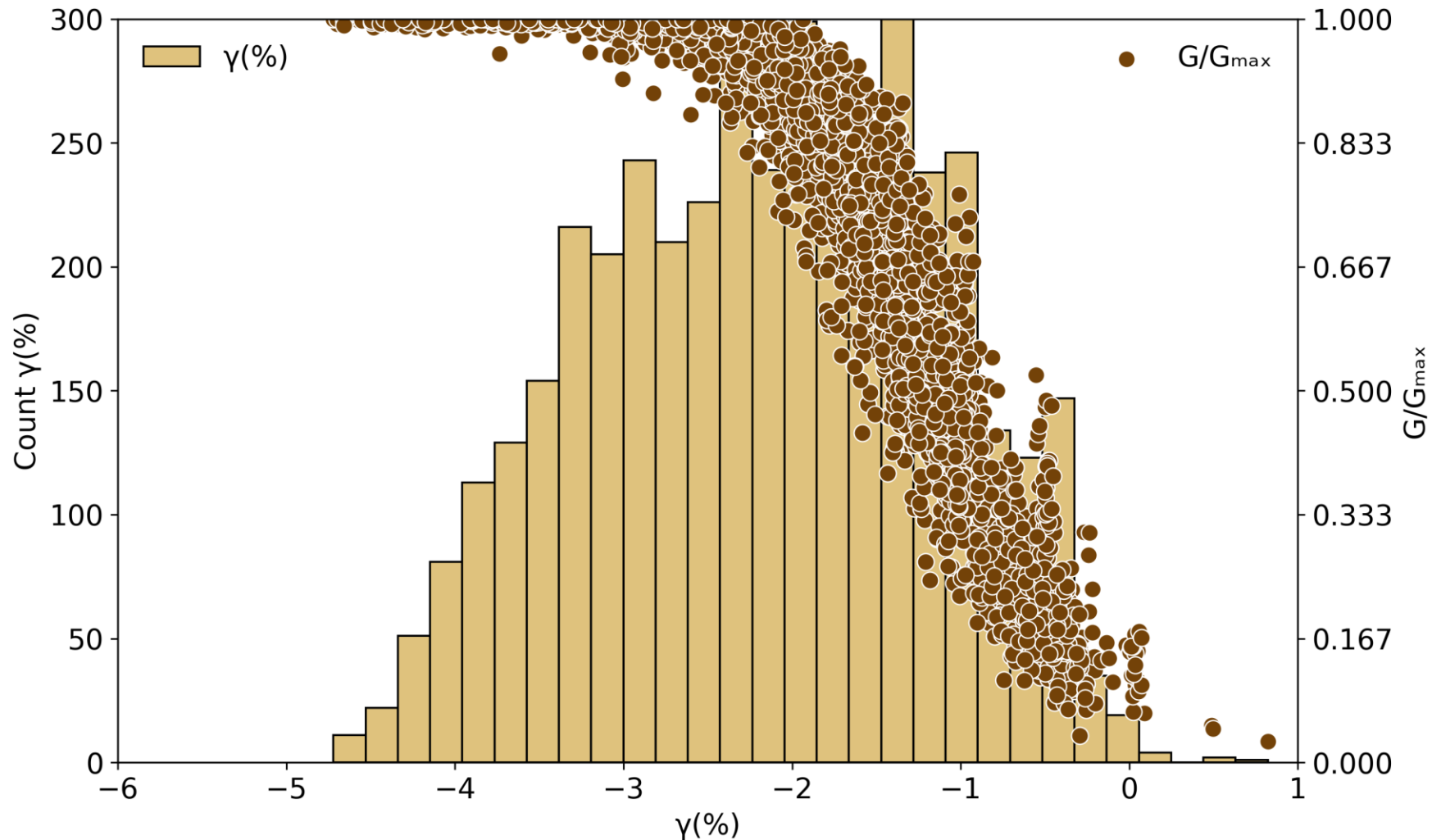
Features		Ciancimino	Facciorusso
UCSC	Geological description	x	x
z	Depth of soil specimen (m)	x	x
σ'_c	Confining pressure (kPa)	x	x
ρ	Density (t/m3)	x	x
PI	Plasticity Index (%)	x	x
LL	Liquid limit (%)	x	x
w	Water content (%)	x	x
e_0	Void ratio (-)	x	x
OCR	Over Consolidation Ratio	x	x

→ only consider parameters well represented in the databases (ex : too many NaN values for OCR)

-> only the 7 first features

DATA CLEANING

- Remove test with z-score >3
- Remove test smaller than 6 points
- Remove test with a rise greater than 0.075 between 2 points
- **dataset of 6635 rows**
 - 2886 from Facciorusso
 - 1628 from Ciancimino



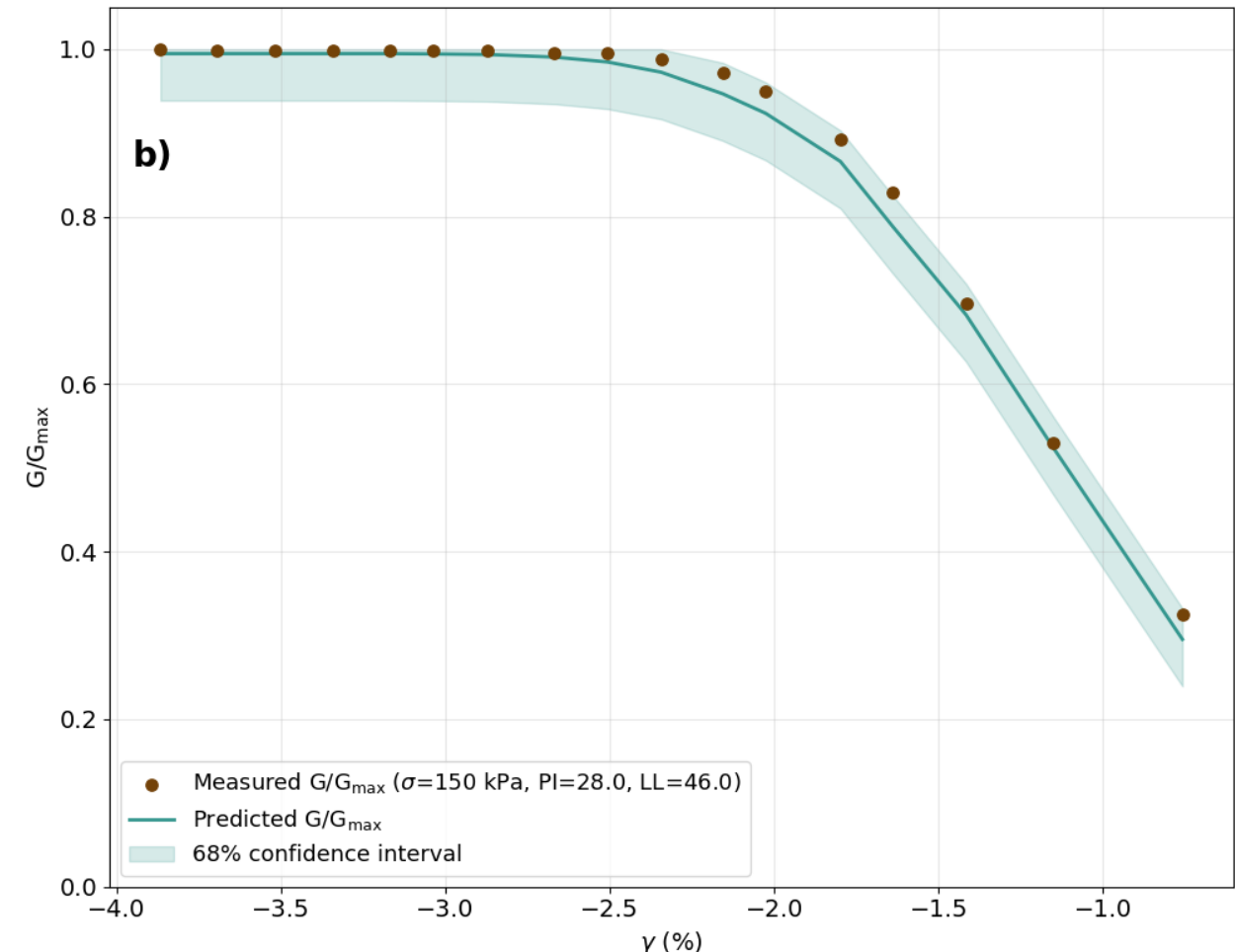
from Bordon et al. 2026

Machine learning Model

- 5 ML families tested → **XGBoost selected** ($R^2 = 0.97$, MAE = 0.03)
- Boosted trees: robust to correlated features
- 9 inputs: Z , σ , ρ , PI, LL, w , e_0 , γ + USCS class → G/Gmax
- Interpretability via **SHAP**
- 80/20 split **by laboratory test** (no leakage between train and test)
- 3-fold CV for hyperparameter tuning (+0.3% only)

A accurate model on unseen laboratory tests

- **$R^2 = 0.97$** on the test set (row-by-row = averaged per test)
- Captures the reverse-sigmoidal, monotonic $G/G_{\max}$ decrease
- Accuracy depends on data density:
 - $G/G_{\max} > 0.9$: $R^2 = 0.996$
 - $[0.8-0.9]$: $R^2 = 0.973$
 - $[0.4-0.5]$: $R^2 = 0.916$
- Lower R^2 at strong nonlinearity = scarce data and lower intrinsic variance, not larger absolute errors



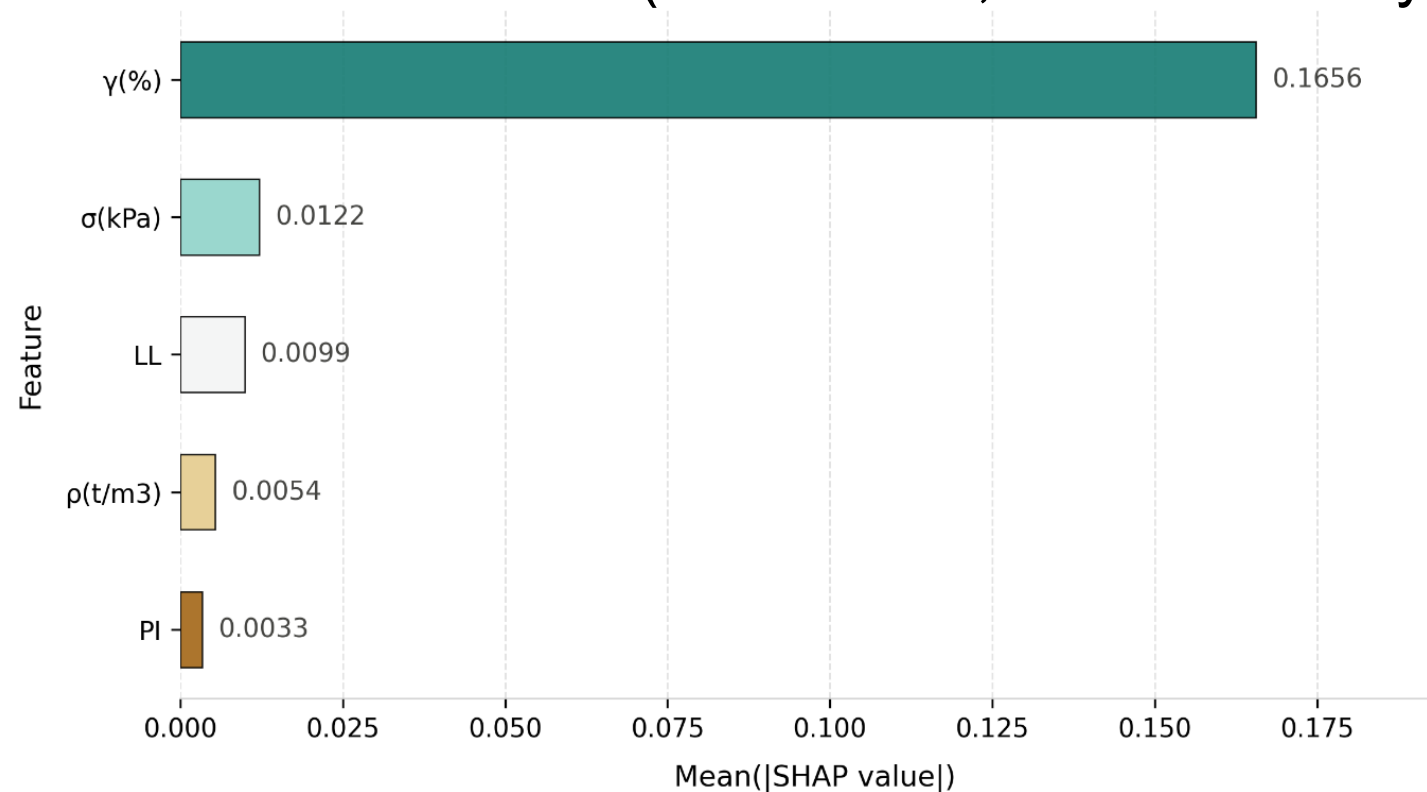
from Bordon et al. 2026

Feature importance and physical consistency

SHAP ranking:

shear strain $\gg$ confining pressure $>$ liquid limit

- Confining pressure is 2nd (absent from several earlier studies)
- LL consistent with the literature (linked to PI, but needs only one lab test)

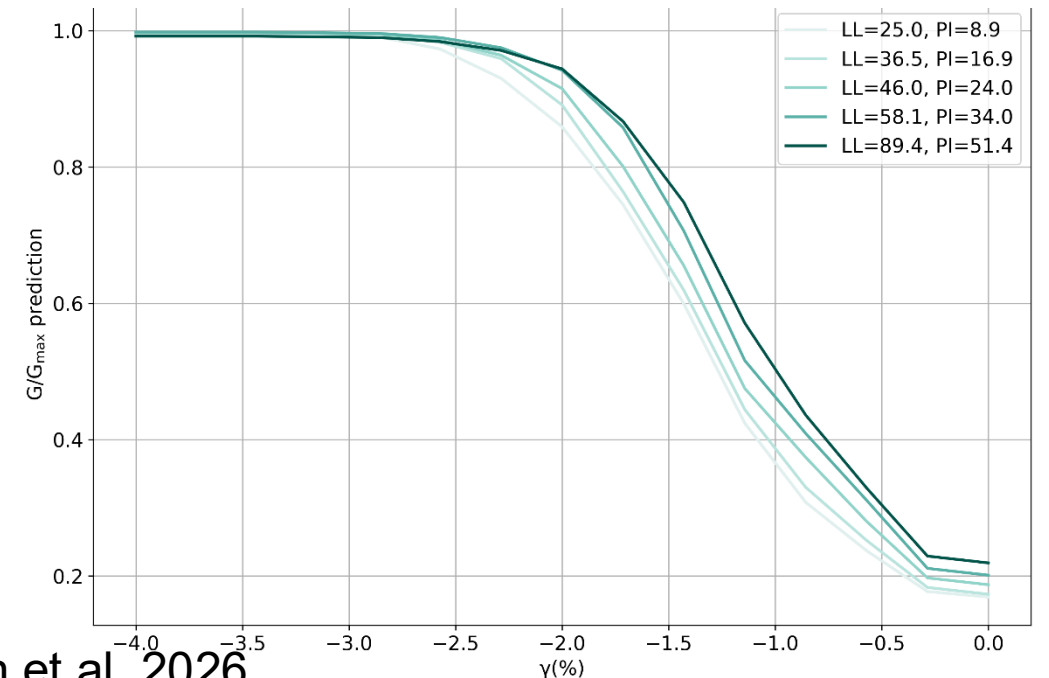
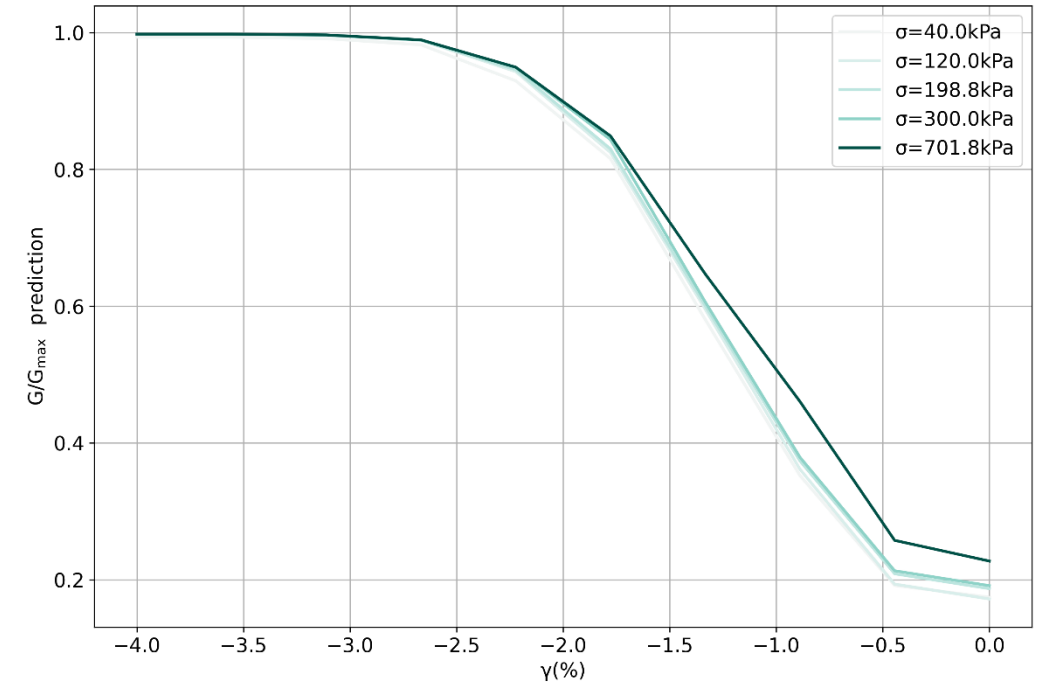


from Bordon et al. 2026

Feature importance and physical consistency

Confining pressure $\uparrow \rightarrow G/G_{\max} \uparrow$, effect negligible at low strain and stronger at high strain (not captured by earlier equations)

LL and PI $\uparrow \rightarrow$ slower stiffness degradation (consistent with Vucetic et al., 1991)
Model is interpretable and consistent with physical expectations



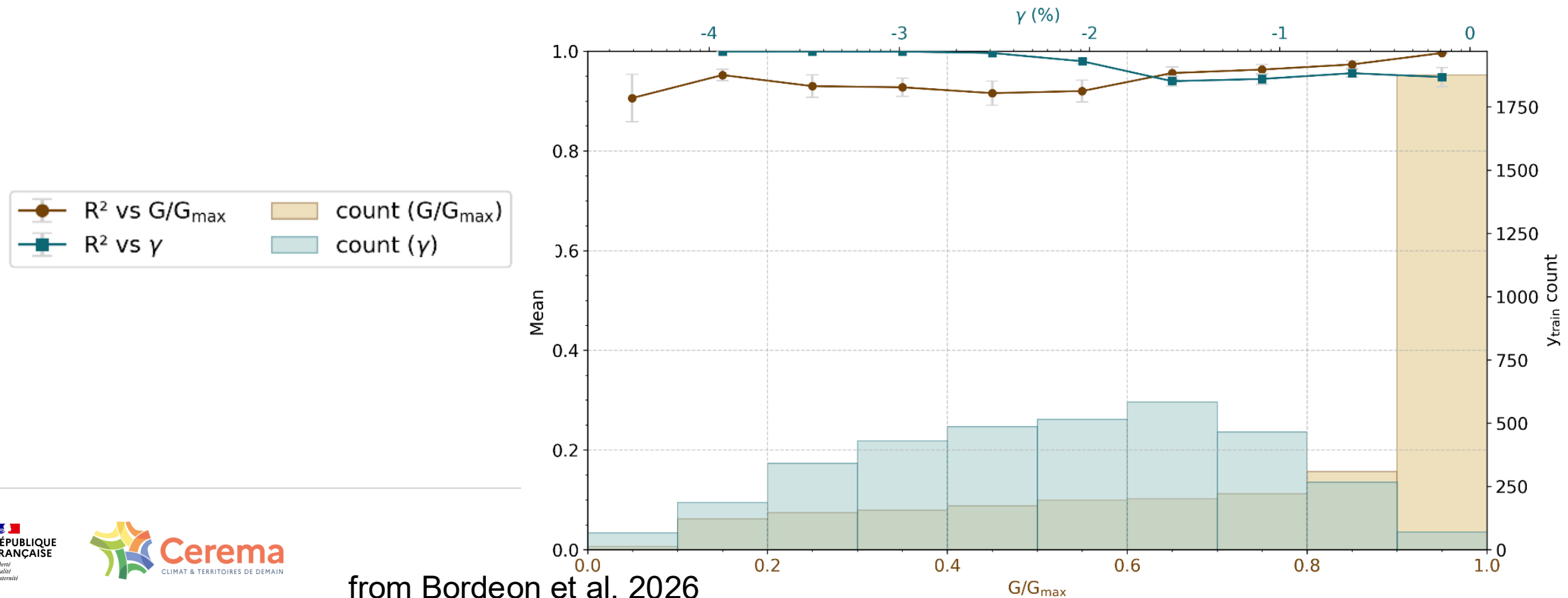
Quantifying model performance

Data concentrated in $[0.9-1.0]$ ($\gamma < 10^{-2} \%$, quasi-elastic range)

$G/G_{\max} > 0.9$: $R^2 = 0.996 \pm 0.001$

$[0.8-0.9]$: $R^2 = 0.973 \pm 0.007$

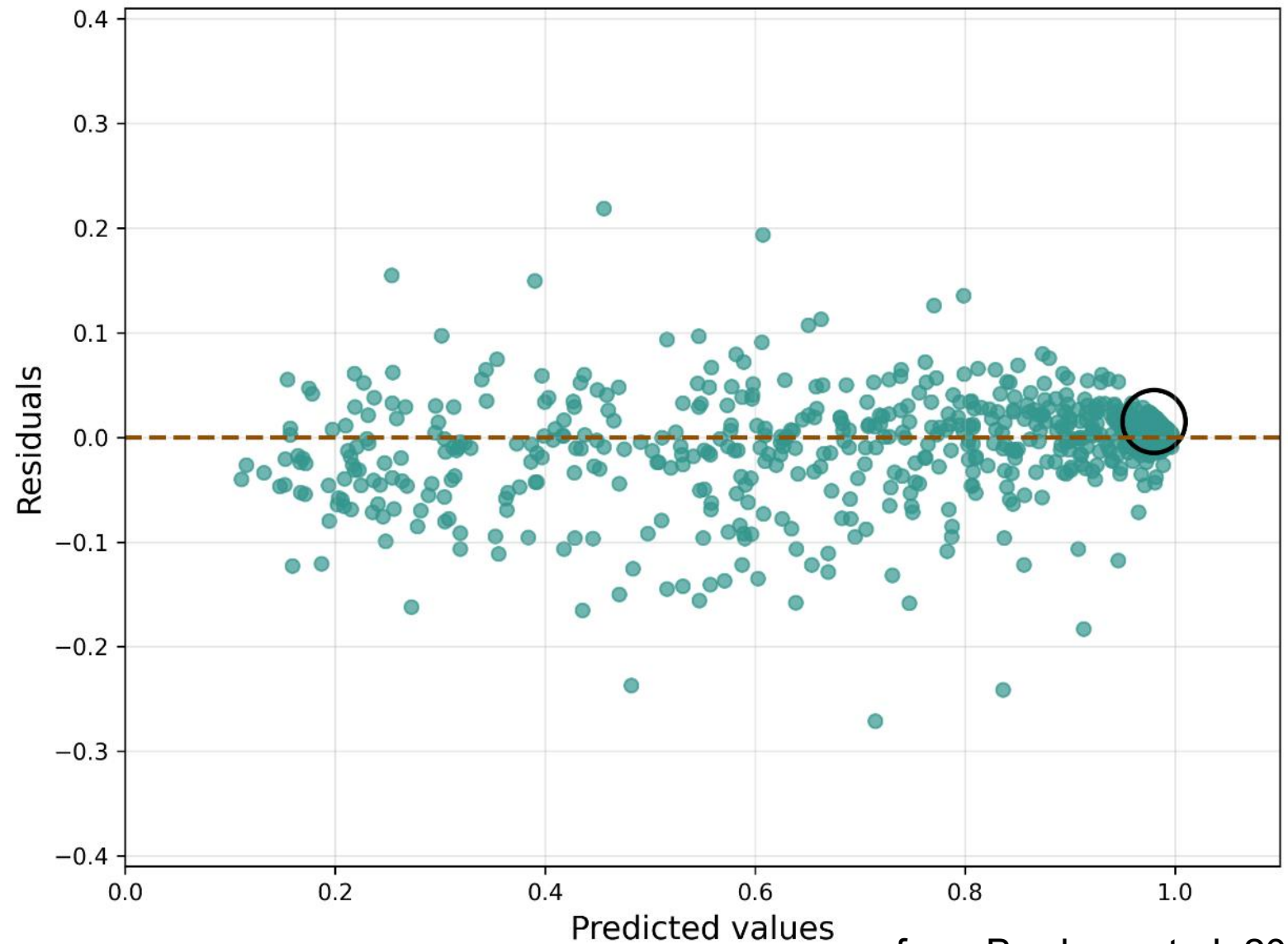
$[0.4-0.5]$: $R^2 = 0.916 \pm 0.024$ (scarcer data)



Quantifying prediction uncertainty

Residuals: **Laplacian,**
homoscedastic

- Constant 68% interval
- **0.056 / +0.037** on $G/G_{\max}$ [0–0.963]
- Upper bound clipped at **1**
($G/G_{\max} \leq 1$)

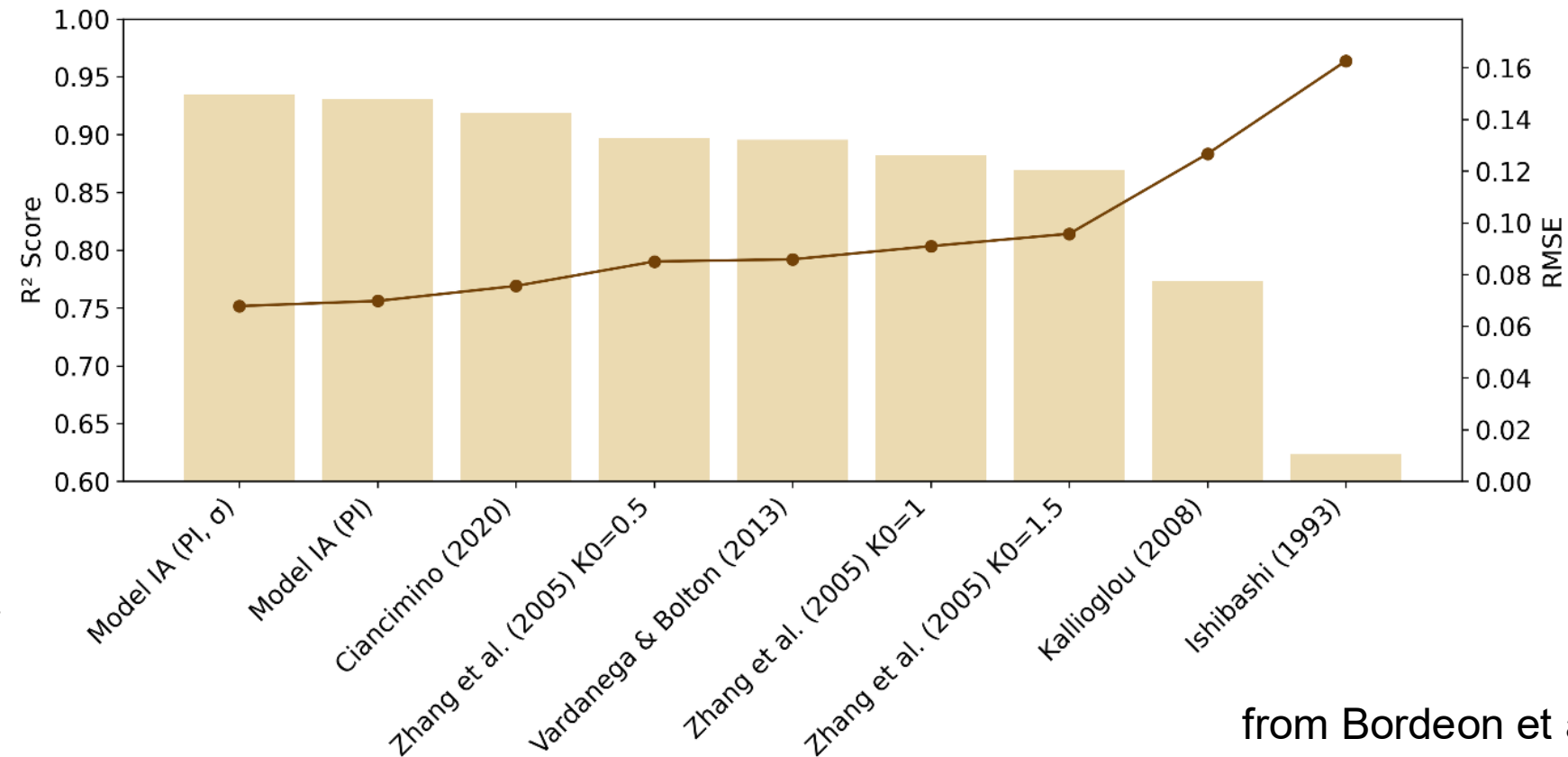


from Bordeon et al. 2026

BENCHMARKING WITH EXISTING MODELS

Test set : Gaudiosi database: **180 laboratory tests** (RC, TS, CDSOSS)

- Fully independent: not used for training
- Two subsets for fair comparison with empirical models:
 - **Model IA (PI):** points with PI
 - **Model IA (PI, σ):** points with PI and confining pressure



CONCLUSIONS

- **Open-access tool:** G/Gmax curves + uncertainty + empirical comparison
- Higher performance than empirical models on independent data, physically consistent
- Limits: Italian CL/CH/ML/MH soils, few high-strain tests
- Next: broader database, more inputs, strain-dependent uncertainty
- **AI complements, not replaces, empirical understanding**

G/Gmax curve as a function of γ (%)

Enter your soil parameters to generate the G/Gmax curve with uncertainties and empirical equations.

PI	Z (m)	e_s	WI
20,00 - +	20,00 - +	0,50 - +	20,00 - +
W	σ (kPa)	ρ (t/m ³)	K_0 (for Zhang)
20,00 - +	200,00 - +	1,50 - +	0,50 - +

Number of points

15

Select USCS class:

☒ CH ☐ CH-CL ☐ CL ☐ CL-CH ☐ CL-ML ☐ MH ☐ MH-OH ☐ ML ☐ ML-OL

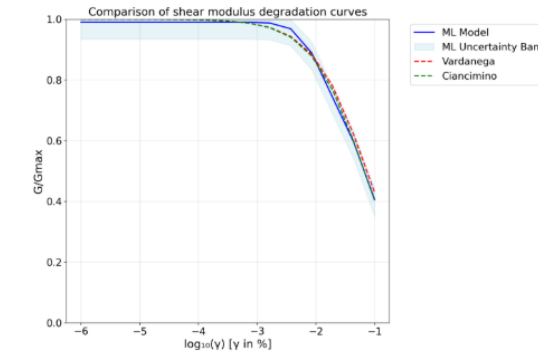
Display options

☒ Show uncertainty bands ☒ Show ML model

Empirical equations to compare

☐ Kollioglou ☐ Ishibashi ☒ Vardanega ☒ Ciamicino ☐ Zhang

Generate curve



Download results as CSV

Parameters used

PI = 20.0, W = 20.0, WI = 20.0
Z = 20.0 m, σ = 200.0 kPa, e_s = 0.5
 ρ = 1.5 t/m³, USCS = CH, K_0 = 0.5

Thanks to Carnot Clim'adapt Institute of Cerema (Accept-IA project)

REFERENCES

- **Borderon Julien, Régnier Julie, Dufour Nathalie. Predicting Normalized Shear Modulus Reduction Curves and Uncertainty for Italian Clayey Soils: A Data-Driven Approach ». Accepted in Soil Dynamic and Earthquake Engineering. 2026**
- Borderon J, Dufour N, Régnier J. Exploring Database Quality Through Shapley Values: Application to Dynamic Soil Parameters Databases. Geotechnics 2025;5:61. <https://doi.org/10.3390/geotechnics5030061>.
- Ciancimino A, Cosentini RM, Foti S, Lanzo G, Pagliaroli A, Pallara O. The PoliTO–UniRoma1 database of cyclic and dynamic laboratory tests: assessment of empirical predictive models. Bull Earthq Eng 2023;21:2569–601. <https://doi.org/10.1007/s10518-022-01573-y>.
- Facciorusso J. An archive of data from resonant column and cyclic torsional shear tests performed on Italian clays. Earthq Spectra 2021;37:545–62. <https://doi.org/10.1177/8755293020936692>.
- Gaudiosi I, Romagnoli G, Albarello D, Fortunato C, Imprescia P, Stigliano F, et al. Shear modulus reduction and damping ratios curves joined with engineering geological units in Italy. Sci Data 2023;10:625.
- Vucetic M, Dobry R. Effect of Soil Plasticity on Cyclic Response. J Geotech Eng 1991;117:89–107. [https://doi.org/10.1061/\(ASCE\)0733-9410\(1991\)117:1\(89\)](https://doi.org/10.1061/(ASCE)0733-9410(1991)117:1(89)).

Thanks

TOUTES NOS ACTUALITÉS, SOLUTIONS,
ACTIVITÉS ET SERVICES

cerema.fr

NOS RESSOURCES DOCUMENTAIRES

doc.cerema.fr

NOS ÉVÈNEMENTS, CONFÉRENCES
ET WÉBINAIRES

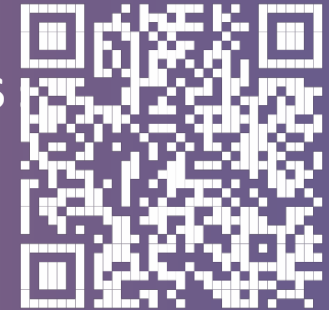
cerema.fr/evenements

NOS FORMATIONS CERTIFIÉES

cerema.fr/formation

NOTRE OFFRE
AUX COLLECTIVITÉS TERRITORIALES
REJOIGNEZ-NOUS !

cerema.fr/collectivites-territoriales



NOUS CONTACTER

cerema.fr/contact

  @cerema