Proceedings of the 5th ECPMG 2024
5th European Conference on Physical Modelling in Geotechnics
2-4 October 2024, Delft, the Netherlands

Predicting moisture levels in response to climatic variables: An approach based on spectral analysis.

Y.F. Vargas-Alzate
Polytechnic University of Catalonia, Barcelona, Spain, yeudy.felipe.vargas@upc.edu

J. Vaunat, A.M. Macias-Gutierrez, A. Di Mariano, A.M. Zapata-Franco
Polytechnic University of Catalonia, Barcelona, Spain, {jean.vaunat, andres.macias, alessandra.dimariano, ana.maria.zapata}@upc.edu

ABSTRACT: Changes in water content under climatic actions is one important process in several geotechnical issues, such as earth-structures deformation, foundation heave/settlement or slope failures seated in the vadose zone. Under certain conditions, these changes may act as a forcing condition for the underlying saturated water column and control cyclic pore pressure at higher depths. The deterministic modelling of these effects requires considering soil-atmosphere thermo-hydro-mechanical interactions under long time high-frequency series. Associated computational cost is a limitation for this type of simulations, particularly when going to large geometry or regional analyses. This article shows two strategies to estimate the soil hydraulic response in time under climatic series handled by Fourier decomposition. These strategies present similarities with techniques used to analyse soil response under seismic actions and propose to define soil transfer functions relating the input (climatic time series) to the output (water content or suction/pore pressure variation at several depths).

1 INTRODUCTION

Expected scenarios of climate change point to a modification of the Intensity-Duration-Frequency curves (IDR) of precipitations, in which long periods of dryness will be suddenly interrupted by very intensive rainfalls (Heidari et al., 2020). This extreme phenomenon would rapidly increase the water content within a dry soil mass, leading to major failure probabilities. Modification of IDR may also play a significant role in earthquake-induced soil phenomena such as liquefaction and seismic slope stability (Keefer, 2002; Roback et al., 2018). Accounting for such effects requires estimation of water pressures or water contents within soil masses, under the whole climatic histogram.

Variations of pore pressures or water contents are governed by the mass balance of water, which results mathematically in a convection-diffusion equation. Under climatic actions, it is moreover coupled with the heat equation to model water phase changes and evaporation. These equations describe various diffusion processes of water, vapour and heat in porous media (Stępniewski et al., 2011), which translate into an attenuation and shift of phase at the depth of surface fluctuations. Under simplifying assumptions, they can be reduced to a system of two diffusion equations that can be solved using classical techniques, avoiding direct-integration schemes (Canuto et al., 2007). For instance, spectral methods by mode decomposition, potentially involving the use of the Fast Fourier Transform (FFT, Cooley & Tukey, 1965). These techniques can be used to reduce the order of the problem by assessing, in the frequency domain, problems unmanageable within the limit of the current computational capacity (long time-periods – predictions under climate change – and large areas – city, hydrological basins). They also present the advantage of handling input defined as a frequency spectrum, avoiding working with time series of climate change scenarios.

Reduction of the computational cost by suppression of the time dimension allows to develop a computational framework to estimate the probability of occurrence of water pressure/water content above thresholds related to soil failure, for a prescribed spectrum of climatic variables. Moreover, computational effort reduction allows considering uncertainties involved, for instance, in the soil mechanical properties, acting loads, climate-related variables, etc.

This article examines two approaches to develop reduced order models capable to predict the evolution of water content in soils, denoted as θ, with time. The first approach is based on the development of transfer functions between climatic variables and θ. These functions are obtained as the ratio between the Fourier amplitude spectra of both θ and the climatic variables.

In addition, it has been explored the development of multi-regression models to enhance the predictability of the first approach. The second approach has been oriented to study the capabilities to predict θ by means of a parametric transfer function. The latter has been developed by considering specific boundary conditions in terms of flux.

2 CLIMATIC VARIABLES AND WATER CONTENT IN THE SOIL

Records of a monitoring campaign carried out in an upper horizontal soil layer during seven years (max. depth 1 m, Calvet et al., 1999) have been analysed. They contain the time-history evolution of the soil water content at several depths. In addition, a series of climatic variables of the area have also been recorded.

2.1 Water content in the soil

Figure 1 shows the time evolution of θ with depth, in the analysed soil profile. Specifically, θ has been recorded at ten depths, the shallowest being at 5 cm from the surface whilst the deepest is at 95 cm.

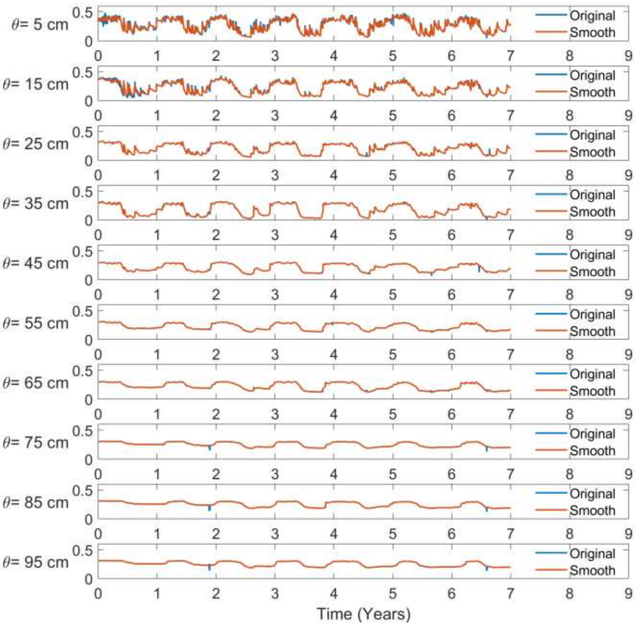

Figure 1. Water content in soils

Due to some problems with the recording devices, some gaps in the data have been observed. They have been interpolated using a smooth function around a given data gap.

2.2 Climatic variables

Data on solar and atmospheric radiation (S.R. and A.R., respectively), rainfall intensity, temperature (T), wind speed (W.S.) and relative humidity (R.H.) have been collected in the aforementioned campaign (Climate-related variables); Figure 2 shows the time-history evolution of them. This figure also shows data smoothing to analyse how the mean evolves over time. Data smoothing has not been applied to the time-history climate data.

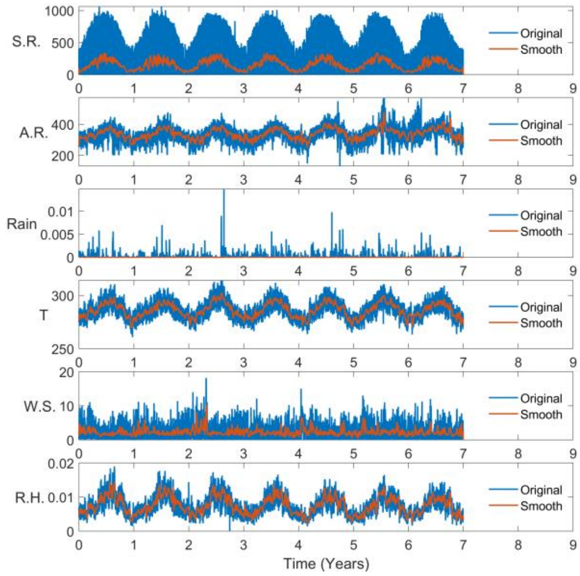

Figure 2. Climatic variables

3 SPECTRAL ANALYSIS

The previous records can be seen as input (Climatic variables) and output variables (water contents at depth), namely I/O variables. Based on this perspective, these records have been used to develop TFs capable of making forecasts considering climatic variables as input. These transfer functions have been obtained as ratios between the Fourier's transform of the recorded I/O variables. Specifically, this approach seeks to characterize θ from a transfer function, TF. This function is obtained by considering the Fast Fourier transform (FFT) of θ and the Fast Fourier transform of a single climatic variable. The following equation shows the mathematical form that characterizes a TF:

$$TF_{i,j}(\omega) = \frac{f_{i,j}(\omega)}{g_j(\omega)} \quad (1)$$

where $f_{i,j}(\omega)$ is the Fourier amplitude spectrum of θ_i; $g_j(\omega)$ is the Fourier amplitude spectrum of the climatic variable g; subscripts i and j stand for a specific depth and a given climatic variable, respectively. The main hypothesis is that, given the time history of θ and a specific climatic variable, g_j, the TF obtained through Equation (1) can be used to approximate the Fast Fourier Transform of the Soil Response Parameter (SPR) to another record, g_k:

$$f_{i,k}(\omega) \approx TF_{i,j}(\omega) * g_k(\omega) \tag{2}$$

To validate this approach, the recorded data have been divided into two subsets: ($f_{i,j}$,g_j) and ($f_{i,k}$,g_k). The first one is used to develop the *TFs* whilst the second to evaluate the consistency of the results. It is worth mentioning that both subsets have data for 3 years. The seventh year has not been considered in the analysis to ensure balance between learning and validation time periods while maintaining signal stationarity.

The efficiency of a variable to predict the dynamic response of a system, using a transfer function, depends on the correlation between both (the variable and the response of the system). Therefore, at a first stage, the causality between I/O variables was evaluated. Specifically, a series of correlation analyses were performed to identify the variables that most influence the prediction of θ (Wagener & Pianosi, 2019). First, the bivariate distribution between pairs of I/O variables has been studied. The efficiency (predictive capacity) of the input variables has been measured in terms of the coefficient of determination, R^2. As a product of this analysis, a ranking based on the efficiency of the input variables has been performed. Table 1 presents the R^2 values between I/O variables. The data show that T and R.H. have a strong correlation with θ, which emphasizes the main role of evaporation in water content changes at that site.

Table 1. R^2 between climatic variables and water content

Z (cm)	S.R.	A.R.	Rain	T	W.S.	R.H.
5	-0.18	-0.44	0.07	-0.58	0.07	-0.55
15	-0.15	-0.45	0.05	-0.55	0.11	-0.54
25	-0.14	-0.45	0.04	-0.55	0.10	-0.55
35	-0.12	-0.45	0.03	-0.53	0.11	-0.55
45	-0.11	-0.44	0.03	-0.51	0.11	-0.53
55	-0.07	-0.39	0.02	-0.43	0.11	-0.44
65	-0.04	-0.36	0.03	-0.37	0.11	-0.37
75	0.02	-0.27	0.02	-0.25	0.11	-0.25
85	0.01	-0.29	0.02	-0.27	0.09	-0.26
95	0.04	-0.23	0.02	-0.18	0.11	-0.18
Mean	**-0.07**	**-0.38**	**0.03**	**-0.42**	**0.10**	**-0.42**

Based on the results presented in Table 1 and employing the first approach, a prediction of θ was performed and T was used to develop the transfer functions. Figure 3 shows the obtained results. There is good agreement between predicted and recorded data in terms of maximum values for all the analysed depths. Simulated data of θ show high frequency fluctuations which are not observed in recorded data.

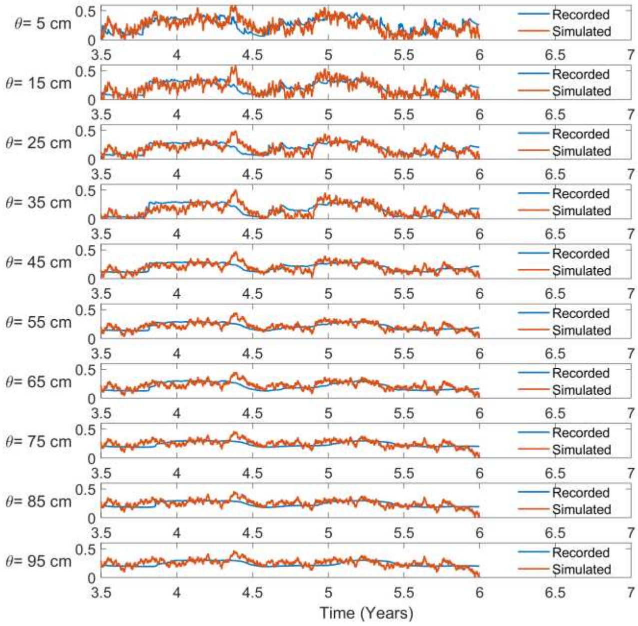

Figure 3. SRP's prediction by using TFs

Predictions presented in Figure 3 only extract information from temperature to develop the transfer functions. As it has been observed in Table 1, this variable is the most correlated with the water content in the soil. However, by using the concept of multi-regression analysis, it is possible to identify pairs of variables that combined increase the correlation with the soil water content. To do so, the following multi-regression model has been employed:

$$y = \alpha_0 + \sum_{i=1}^{N_{IV}} \alpha_i z_1 + \varepsilon \tag{3}$$

where N_{IV} represents the number of information variables, IV, considered in the regression model; α_n represents the coefficients providing the best fit between model and data; z_1 is a basic function; ε represents the residuals. Based on this model, Table 2 presents the R^2 values between I/O variables by considering that the input, I, is the summation of two IVs. In this table only the mean R^2 has been presented (Last row in Table 1). In addition, the α_n coefficients providing the best fit are also presented.

It can be observed that the pair (S.R., T) is the most correlated. Based on this information, it has been developed a new variable, which takes information from the solar radiation and temperature, to predict water content. This variable employs the regression coefficients shown in Table 2. Figure 4 shows the new prediction based on this enhanced variable. It can be observed that the consistency between recorded and simulated data is higher compared to the use of a single IV (see Figure 3).

Table 2 R^2 for pairs of variables

IV_1	IV_2	R^2	α_0	α_1	α_2
S.R.	**A.R.**	0.38	0.329	0.000	0.000
S.R.	**Rain**	0.09	0.250	0.000	7.145
S.R.	**T**	0.47	0.691	0.000	-0.002
S.R.	**W.S.**	0.15	0.245	0.000	0.002
S.R.	**R.H.**	0.43	0.271	0.000	-2.969
A.R.	**Rain**	0.38	0.330	0.000	10.603
A.R.	**T**	0.45	0.401	0.000	0.000
A.R.	**W.S.**	0.39	0.324	0.000	0.002
A.R.	**R.H.**	0.44	0.324	0.000	-0.549
Rain	**T**	0.42	0.528	5.360	-0.001
Rain	**W.S.**	0.11	0.246	5.262	0.002
Rain	**R.H.**	0.42	0.272	8.009	-2.723
T	**W.S.**	0.46	0.558	-0.001	0.003
T	**R.H.**	0.45	0.416	-0.001	-1.528
W.S.	**R.H.**	0.44	0.266	0.003	-2.793

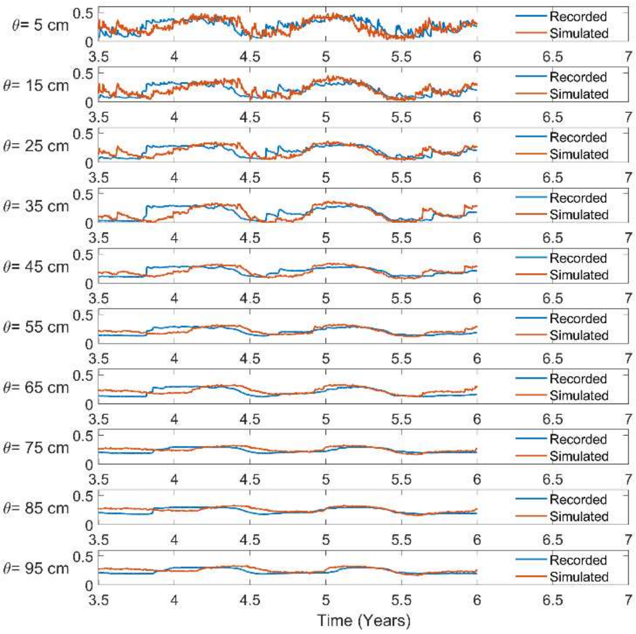

Figure 4 SRP's prediction based on two IV

4 PARAMETRIC TRANSFER FUNCTION

Another approach to predict θ by means of spectral methods is to develop a parametric transfer function. To do so, special boundary conditions need to be considered. In this case, the flux ($\frac{d\theta}{dt}$) at a given depth, L'_z, was set at zero and the θ value recorded at a depth of 5 cm was imposed at the surface (See Figure 1, top). In this manner, the analytical solution of the diffusion equation provides the expression of θ at depth z:

$$\theta = \delta_i * A_i \sum_{i=1}^{n} \sin(\omega_i t + \alpha_i * (z - L_z) + \varphi_i + \frac{\pi}{2}) \quad (4)$$

where $\delta_i = \frac{e^{\alpha_i z} - e^{-\alpha_i z}}{e^{\alpha_i L'_z} - e^{-\alpha_i L'_z}}$; $\alpha_i = \sqrt{\frac{\omega_i}{2\kappa}}$; κ is the diffusion coefficient, in a mathematical sense; A_i and φ_i are the amplitude and the phase of the harmonic with frequency ω_i, extracted from the Fourier's amplitude spectrum, respectively; $\bar{\theta}$ is the mean of the water content during the period taken into account.

Equation (4) has allowed estimating, at all the depths, the pair (κ, L'_z) that minimizes the mean square error (MSE) with respect to the recorded data. This was achieved by means of the Monte Carlo method. It is worth mentioning that the estimation provided by Equation 3 depends on the number of harmonics. Thus, the Monte Carlo minimization was performed using 100 harmonics. Figure 5 shows the evolution of the MSE as a function of κ and L'_z.

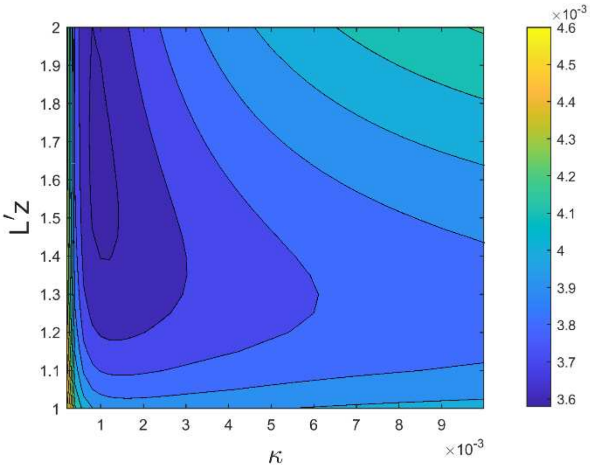

Figure 5. MSE as a function of κ and L'_z

From Figure 5, it can be inferred that the optimal pair (κ, L'_z) is κ= 0.001 and L'_z= 1.6 m. By using these values, θ has been estimated for all depths; Figure 6 shows a comparison of both simulated and recorded data, using 100 harmonics in Equation 3. Simulated and recorded data are in good agreement for all depths. Simulated results do reproduce well both the attenuation with depth of the water change amplitudes and the disappearance of the higher frequency modes. This is consistent with the characteristics of the diffusion equation underlying the infiltration/ evaporation process and indicates that the sensitivity of water content changes to atmospheric fluctuations decays with depth.

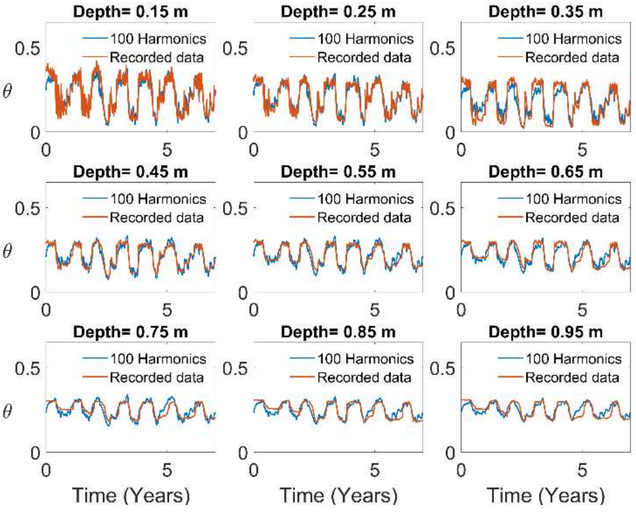

Figure 6. θ prediction considering 100 harmonics

To study possible reduction of the computational cost, another analysis was conducted using the same pair (κ, L'_z) that minimizes the MSE and considering only the 10 harmonics (Figure 7) provided with the higher amplitudes. The agreement between simulated and recorded data is similar to the one observed for the analysis with 100 harmonics. This provides information about the most relevant modes to be used in future predictions of water content under climatic actions, at this site. The number of harmonics to be used decreases with depth as the result of the attenuation of the fluctuations of higher frequencies with the distance to atmospheric boundary conditions.

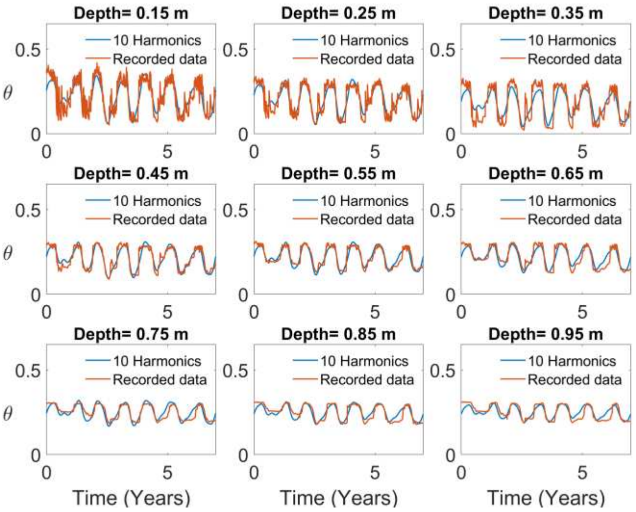

Figure 7. θ prediction considering 10 harmonics

5 CONCLUSIONS

This article presents two approaches based on spectral methods, for estimating soil water contents at different depths. The attractiveness of both methods relies mostly on the reduction of computational times. Data from a monitoring campaign carried out in an upper horizontal soil layer (max. depth 1 m, Calvet et al., 1999) were considered. These records implicitly contain information on the interaction of complex physical processes that trigger highly random variability. They provide an opportunity to conduct causality studies aimed at analysing which climate variables (input) have the highest explanatory capacity to predict soil water content, θ (output variable).

In the first approach, transfer functions connecting variables that influence climate (Solar and atmospheric radiation, rainfall intensity, temperature, wind speed, atmospheric pressure and relative humidity) and θ at several depths were operated. Data show that temperature and relative humidity are the atmospheric variables with the highest predictive capacity of θ. The correlation between these climatic variables and θ increases if a multi-regression model composed of sub-groups of climatic variables is employed (See Vargas-Alzate et al., 2022). One of the main shortcomings of the first approach, when employing only one *IV*, is that simulated data tend to fluctuate more than recorded series, over short time periods. However, these fluctuations can be mitigated if filters that act on the resulting signals are considered. Future research should be conducted to this end. Anyhow, using two *IV*s significantly increases the coherence between simulated and recorded data.

With regard to the second approach, a deterministic transfer function was employed, based on specific boundary conditions. Predictions based on this method are better than those based on the first approach, provided that parameters are previously estimated by back-analysis on sub-sets of data series. As depth increases, it is possible to consider fewer harmonics than close to the soil surface. Estimating the optimal number of harmonics as a function of depth would help improving the model efficiency. The use of a specific parameter that varies with depth would help addressing the intrinsic non-linearity of the diffusion coefficient.

6 ACKNOWLEDGEMENTS

This research has been funded by the Spanish Research Agency (AEI) through the project with reference: TED2021-132559B-I00 – J-02970.

7 REFERENCES

Calvet, J.-C., Bessemoulin, P., Noilhan, J., Berne, C., Braud, I., Courault, D., Fritz, N., Gonzalez-Sosa, E., Goutorbe, J.-P., Haverkamp, R., Jaubert, G., Kergoat, L., Lachaud, G., Laurent, J.-P., Mordelet, P., Olioso, A., Péris, P., Roujean, J.-L., Thony, J.-L., … Vignes, D. (1999). MUREX: a land-surface field experiment to study the annual cycle of the energy and water budgets. *Annales Geophysicae*, *17*(6), 838–854. https://doi.org/10.1007/s00585-999-0838-2

Canuto, C., Quarteroni, A., Hussaini, M. Y., & Zang, T. A. (2007). *Spectral Methods*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-30728-0

Cooley, J. W., & Tukey, J. W. (1965). An Algorithm for the Machine Calculation of Complex Fourier Series. *Mathematics of Computation*, *19*(90), 297. https://doi.org/10.2307/2003354

Heidari, H., Arabi, M., Ghanbari, M., & Warziniack, T. (2020). A Probabilistic Approach for Characterization of Sub-Annual Socioeconomic Drought Intensity-Duration-Frequency (IDF) Relationships in a Changing Environment. *Water*, *12*(6), 1522. https://doi.org/10.3390/w12061522

Keefer, D. K. (2002). Investigating Landslides Caused by Earthquakes – A Historical Review. *Surveys in Geophysics*, *23*(6), 473–510. https://doi.org/10.1023/A:1021274710840

Roback, K., Clark, M. K., West, A. J., Zekkos, D., Li, G., Gallen, S. F., Chamlagain, D., & Godt, J. W. (2018). The size, distribution, and mobility of landslides caused by the 2015 Mw7.8 Gorkha earthquake, Nepal. *Geomorphology*, *301*, 121–138. https://doi.org/10.1016/J.GEOMORPH.2017.01.030

Stępniewski, W., Sobczuk, H., & Widomski, M. (2011). *Diffusion in Soils* (pp. 214–220). https://doi.org/10.1007/978-90-481-3585-1_273

Vargas-Alzate, Y. F., Hurtado, J. E., & Pujades, L. G. (2022). New insights into the relationship between seismic intensity measures and nonlinear structural response. *Bulletin of Earthquake Engineering*, *20*(5), 2329–2365. https://doi.org/10.1007/s10518-021-01283-x

Wagener, T., & Pianosi, F. (2019). What has Global Sensitivity Analysis ever done for us? A systematic review to support scientific advancement and to inform policy-making in earth system modelling. *Earth-Science Reviews*, *194*, 1–18. https://doi.org/10.1016/j.earscirev.2019.04.006