Exploring the spectral patterns of chemical elements in soils through proximal VIS-NIR spectroscopy and ML (a pilot study: Armenia).

Andrey Medvedev [1], Gevorg Tepanosyan[1], Shushanik Asmaryan[1], Rima Avetisyan[1], Vahagn Muradyan[1], Igor Sereda[1], Astghik Gevorgyan[1], Karine Davtyan[1], Lilit Sahakyan[1]

[1] Center for Ecological-Noosphere Studies National Academy of Sciences of Armenia

Keywords: proximal sensing, VIS-NIR spectroscopy, spectral signatures, soil geochemistry, spectral patterns, Armenia

Abstract

Armenia has limited agricultural soil resources, which varies in type, vertical spatial distribution and quality and the information on soil chemical, physical and biological attributes are very limited. To establish datasets to cover these gaps and to develop the best combination of the methods of soil qualitative and quantitative assessments and classification of their spatial variability dedicated studies has been conducting. This work is aimed to investigate and reveal the main characteristics of the spectral reflectance of the chemical elements in soils in Armenia using a combination of proximal VIS-NIR and X-ray spectroscopy methods. 30 soil samples were collected from different regions of Armenia characterized with different landscape properties. The content of a number of chemical elements were determined by X-ray spectroscopy and spectral signals of the samples via field spectroradiometer were collected. The comparison of the data of X-ray and VIS-NIR spectroscopy defines three clusters, led by the soil macro components: I cluster - SiO_2, Al_2O_3, Fe_2O_3; II cluster – CaO and III cluster – K_2O. Correlation analysis of the content of these macro components and their spectral reflectance values shows that they vary in spectral patterns. In conclusion it can be stated that the main macro components in the studied soil types have specific spectral patterns. The responsible spectral bands are mainly visible and NIR (400-750nm) and two SWIR bands (800-1750nm and 1950-2400nm). Among the studied elements, the exception was CaO as the most well distinguished macro component over the full spectral range 350nm to 2500nm, except (1750 – 1950nm).

1. Introduction

Soil is valuable resource for agricultural and a primary source of food and energy for people. Therefore, the sustainable soil resources management requires a comprehensive understanding of the soil chemical, mineralogical and other characteristics (Silver et al, 2021, John W, et al, 1994). Armenia has limited agricultural soil resources, which varies in type, vertical spatial distribution and quality (National Atlas of Armenia, 2007). However, the information on soil chemical, physical and biological attributes are very limited. Environmental science community of Armenia works on the establishment of datasets to cover these gaps and seeks to develop the best combination of the methods of soil qualitative and quantitative assessments and classification of their spatial variability (Tepanosyan et al, 2023, Tepanosyan et al, 2024). In line with in situ measurements and lab analysis the application of remote sensing technics is of great importance supporting and ensuring the creation of a national soil databases or spectral library in Armenia (Bellinaso et al 2010, Ma et al, 2023). Compared with the traditional analytical methods proximal sensing such as VIS-NIR and SWIR spectrometry is fast and cost effective (Hong et al, 2018, Yu et al, 2023). The spectral range commonly used to analyse soil spectral patterns includes 400–700 nm (visible, Vis), 700–1100 nm (near-infrared, NIR), and 1100–2500 nm (shortwave infrared, SWIR). These ranges can be captured by field or laboratory sensors and serve as the foundation for optical aerial and satellite remote sensing, including reflectance and imaging spectroscopy (Ma et al, 2023: Dalponte et al., 2023). As not all wavelengths show correlations with the specified soils variable values, the right portions of the spectral wavelengths need to be selected for further estimations of soil properties (Yu et al, 2023). So, at a very first stage of the studies towards the systemising the spectral characteristics of Armenia, a pilot study was conducted and a research goal was stated to investigate and reveal the main characteristics of the spectral reflectance of the chemical elements in soils in Armenia using a combination of proximal VIS-NIR and X-ray spectroscopy methods.

2. Material and Methods

2.1 Study area

The pilot study was conducted in 2020. 30 soil samples were collected from 5 administrative regions (marzes) of Armenia (Ararat, Armavir, Gegharqunik, Vayots Dzor, Aragatsotn) characterized with different soil types and landscape properties (Figure 1). The studied area exceptionally reflects the vertical zoning of the country, which starts from semi-desert to subalpine and nival landscape zones distributed on the altitudes 850-4092m (Ghazaryan, 2013). The difference in altitude of about 3000m meters, the average annual temperature variation in the country is >170C, and the average aerothermal gradient is 0.6 for every 100 meters of altitude (Soil resources of Mediterranean and Caucasus countries, 2013). Armenia is notable for an exceptional abundance of solar irradiation: the reported duration of sunshine varies from 1,900 to 2,800 hours per year, while the annual number of days when the sun does not shine is from 18 to 64. The tension of solar radiation is also high. The great diversity of bioclimatic and lithological-geometric conditions as well as the long and diverse economic activity have led to formation of different soil types (Ghazaryan, 2013).

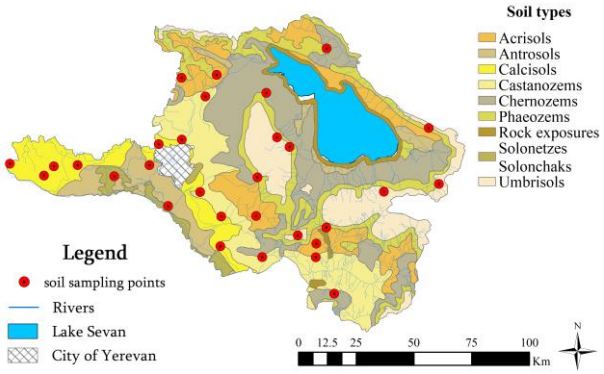


Figure 1. Spatial distribution of sampling sites.

2.2 In situ measurements

The soil samples of the studied area are mainly represented by calcisols, chernozems, acrisols, castanozems, phaeozems etc. (National Atlas of Armenia, 2007) (Figure 1).

The samples were taken from the above 20 cm of soil layer after removing grass cover is such was presented. Each sample consisted of 5 subsamples which were placed in the plastic bag and transported to the laboratory. At the laboratory, the samples were dried in a drying oven (SMO28-2) at 40ºC. The dry samples were then crushed using a ceramic mortar and pestle and sieved through a 2 mm sieve. The samples were pulverized in vibratory grinder (IV6, Vibrotechnik) and the contents of Co, Zn, Fe_2O_3, Y, Mn, V, Ti, Ga, Al_2O_3, Nb, SiO_2, Rb, Br, S, Sr, Zr, Ba, P, K_2O, Pb, As, Ni, Cu, Cr, CaO elements were determined by X-ray spectroscopy (Rigaku NEX DE).

2.3 Proximal Sensing

Afterwards, the spectral signals of the 2 mm sieved soil samples were collected via a passive NaturaSpec spectroradiometer over the full spectral range 350nm to 2500nm.

N	***Specifications***	***Features***
1	Spectral Range	350-2500nm
2	Spectral Resolution	2.7nm@700nm
		5.5nm@1500nm
		5.8nm@2100nm

Table 1. The spectral specifications of NaturaSpec spectroradiometer.

Spectrometric analysis, each soil sample, sieved to 2 mm, underwent passive, non-contact spectrometry. The samples were evenly spread in a 5 mm layer within Petri dishes and placed in an open area under bright, stable natural sunlight. Spectral measurements were carried out using the NaturaSpec spectroradiometer across the full spectral range of 350 nm to 2500 nm. Measurements were taken at a height of 5–7 cm above the surface in a strictly orthogonal orientation. A total of 10 measurements were performed, with slight sensor shifts relative to the previous measurement area to minimize potential anomalies. Before and after scanning each sample, a calibration measurement was conducted using a white reference panel with known reflectance values, following the same procedure. To determine the absolute spectral reflectance of the soil samples, the average value of the two extreme white panel measurements was calculated, multiplied by its spectral reflectance coefficient at the given wavelength, and divided by the mean soil measurement value. These calculations were automated using a Python scripting.

2.4 Comparison between UAV-derived and in-situ-measured LST.

To unveil the link between the studies elements, their groups and spectral signals the generated data was preprocessed using the normality test (Shapiro-Wilk), Spearman correlation analysis, PCA, k-means and hierarchical clustering. Fig 2 shows the workflow of the study.

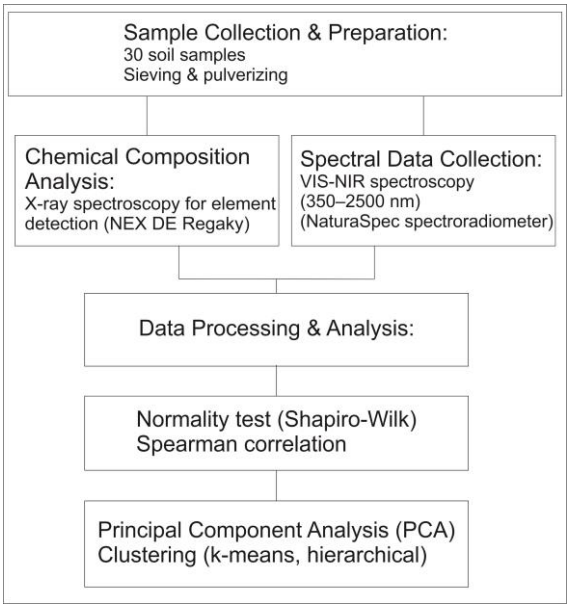


Figure 2. The flowchart of the study

3. Results and discussions

The descriptive statistics of the chemical elements are shown in the table 2. Among the studied elements As, S and Br show high values (>100%) of the coefficient of variations (CV) 246%, 163%, 122% correspondingly. Relatively low (<60%) level of CV was observed for Ti, V, Mn, Ni, Zn, Rb, Y, Zr, Al_2O_3, SiO_2, Fe_2O_3, CaO, K_2O, Nb, P, Ga, Co with values of 25.6, 29.5, 30.7, 43.1, 50.4, 44.0, 22.2, 28.7, 16.2, 19.8, 26.5, 29.8, 41.9, 46.5, 26.3, 39.1 respectively, suggesting the homogeneous distribution of these elements, which are typical for the ground areas (Beus et al, 1976).

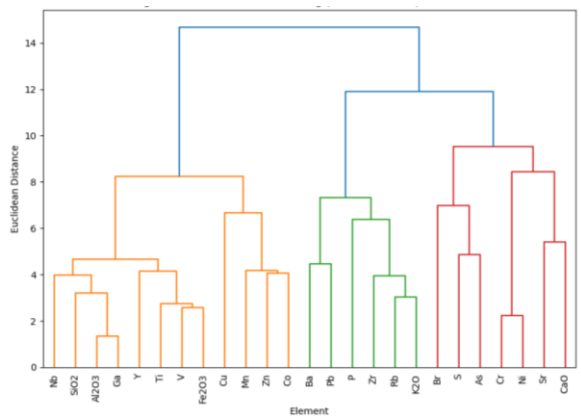
Shapiro-Wilk normality test shows that the studied elements were mainly non-normally distributed. The Figure 1 shows that three clusters of the elements are formed. In each cluster, the soil macro components were observed: I cluster - SiO_2, Al_2O_3, Fe_2O_3; II cluster – CaO and III cluster – K_2O (fig 1).

Elem.	Mean	Med.	SD	Skew.	Min	Max.	CV,%
S	1462	871	2383	4.82	110	13539	5.0
Ti	4904	4954	1257	-.999	1600	6998	25.6
V	237	230.50	69.9	.258	99	400	29.5
Cr	224	176.00	191	1.94	19	869	85.4

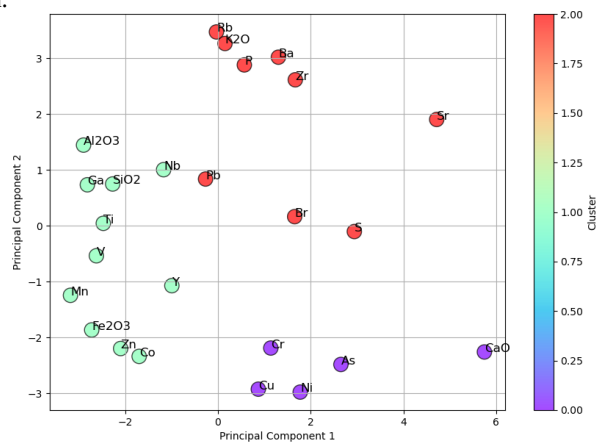
Mn	1022	1013.00	314	.461	392	1926	30.7
Ni	112	97.250	48.3	.995	25.9	243	43.1
Cu	53.9	37.600	34.6	1.72	17.6	154	64.2
Zn	82.6	73.250	41.6	3.39	23.1	271	50.4
As	18.6	6.785	45.8	4.09	.96	229	245.8
Br	6.08	4.6350	7.43	4.16	.349	42	122.1
Rb	66.0	63.850	29.0	.824	14.0	154	44.0
Sr	393	328.50	244	1.60	98.4	1126	62.0
Y	20.1	21.650	4.5	-1.23	7.02	26.1	22.2
Zr	204	203.50	58.5	.112	66.0	343	28.7
SiO_2	488400	505029	79240	-1.23	267568	605367	16.2
Al2O3	128648	131059	25500	-1.05	59962	167710	19.8
Fe2O3	60953	61141	16155	0.481	24625	111942	26.5
CaO	84539	61055	83592	1.76	15374	360095	98.9
K2O	17285	16681	5150	0.670	7232	32677	29.8
Nb	12.6	12.6	5.30	1.41	1.510	31.800	41.9
P	1348	1370	627	0.003	436	2291	46.5
Ba	288	240	285	4.08	51.0	1641	98.9
Pb	16.9	14.5	16	4.56	3.14	97	95.3
Ga	11.4	11.9	3.0	-1.23	3.73	16.4	26.3
Co	52.9	50.6	21	1.24	11.8	124	39.1

Table 2. The descriptive statistics of the chemical elements

Correlation analysis of the content of these macro components and their spectral reflectance values shows that they vary in spectral patterns. In the I cluster the macro components shows a significant negative correlation in near infrared and two short-wave infrared (SWIR) bands (800-1750nm and 1950-2400nm) as follows (SiO_2` r=-0.5, p<0.05; $Al_2O_3$ r=-0.7, p<0.01; $Fe_2O_3$ r=-0.7, p<0.01). II group represented by CaO macro component shows a significant positive correlation in visible band 450-650nm` r=0.75; p<0.01 and in near-infrared and two short-wave infrared bands (750-1750 and 1950-2400nm) varies between r = {0.7-0.75} with a different level of significance p<0,01 and p<0,001.



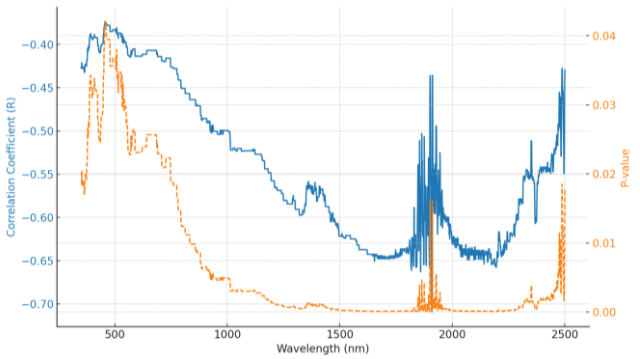
a.



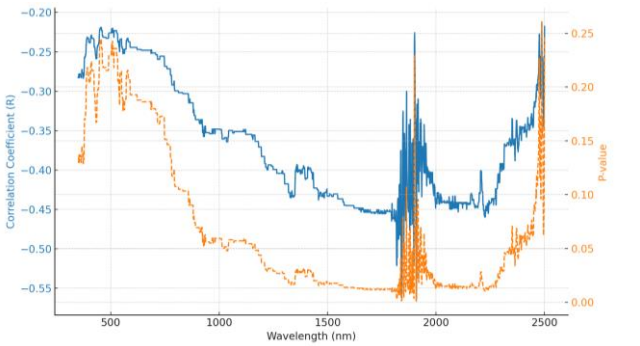
b.

Figure 3. Clustering of elements: a. hierarchical; b. K-Means with PCA

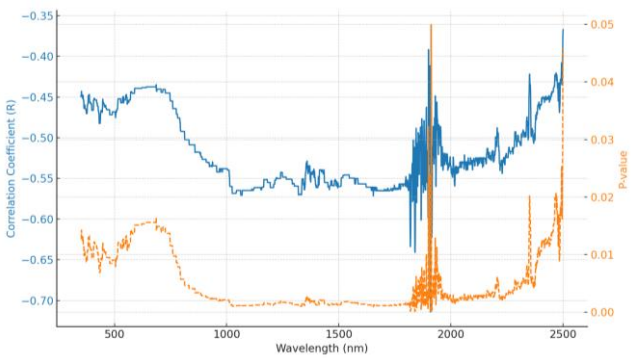
Correlation analysis of the content of these macro components and their spectral reflectance values shows that they vary in spectral patterns. In the I cluster the macro components shows a significant negative correlation in near infrared and two short-wave infrared (SWIR) bands (800-1750nm and 1950-2400nm) as follows (SiO_2` r=-0.5, p<0.05; $Al_2O_3$ r=-0.7, p<0.01; $Fe_2O_3$ r=-0.7, p<0.01). II group represented by CaO macro component shows a significant positive correlation in visible band 450-650nm` r=0.75; p<0.01 and in near-infrared and two short-wave infrared bands (750-1750 and 1950-2400nm) varies between r = {0.7-0.75} with a different level of significance p<0,01 and p<0,001. Finally, K_2O in the third group shows no significant correlation over the full spectral range (Figure 3).



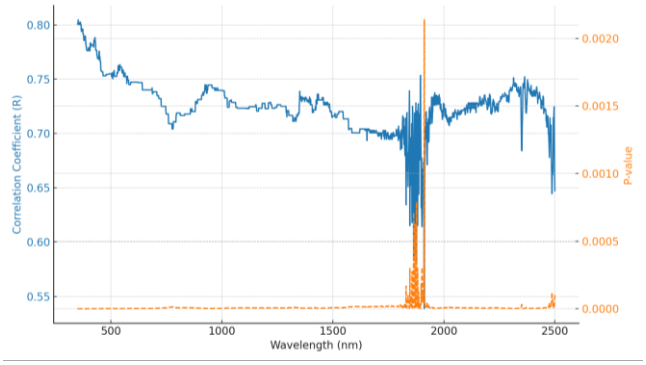
a.



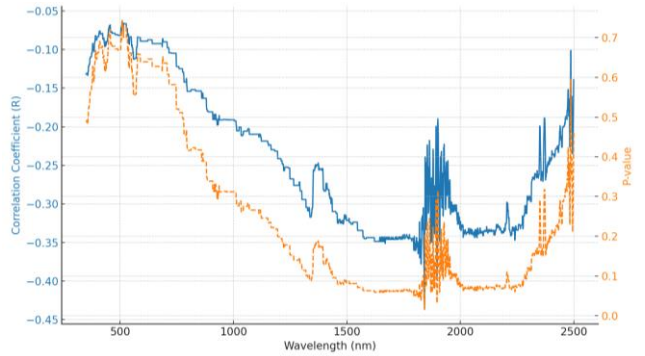
b.



c.



d.



e.

Figure 4. Correlations between spectral reflectance of soil macro components: a. Al_2O_3, b. SiO_2, c. Fe_2O_3, d. CaO, e. K_2O

Spatial distribution of the K-means clustering revealed three spatial groups of the soil samples (Figure 5).

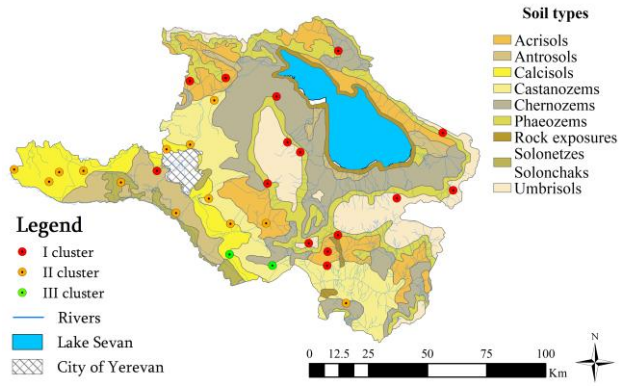


Figure 5. The spatial distribution of soil samples.

It is evident that the identified groups have distinct spatial separation. Particularly, the first cluster is mainly distributed in comparatively high mountainous areas presented by chernozems and castanozems. The second cluster is distributed in the areas where calcisols are dominating. The formation of this cluster is conditioned by the presence of CaO. Moreover, this cluster also included Cu and As. This finding is in line with the previous study by Tepanosyan et al, (2022) where the same combination of element was found on the same area.

It is noteworthy, that in the small spectral window 1750-1950 nm all three groups shows well defined spectral noise (Figure 4), which is another case for the study. However, as stated by Baumann et al, VIS-NIR (400–2500 nm) measurement can be used to estimate multiple soil properties (Baumann et al, 2021). Hence, the next steps suggest the studies of the other soil properties to reveal the spectral response of the other elements associated with the studied macro components.

4. Conclusion

In conclusion, it can be stated that the main macro components in the studied soil types have specific spectral patterns. The responsible spectral bands are mainly visible and NIR (400-750nm) and two SWIR bands (800-1750nm and 1950-2400nm). In the list of the studied elements, the exception was CaO as the most well distinguished macro component over the full spectral range 350nm to 2500nm, except (1750 – 1950nm). The results of this pilot study are crucial for the future more detailed investigations ensuring the appropriate basic for the building of complex multiparameter model for estimation and validation of different soil aspects.

Acknowledgements

The work was supported by the Science Committee of RA, in the frames of the research projects № 22IRF-04 and № 23LCG1E016, the state target program № 1-12/TB-23 and the state budget funding 2025.

References

Baumann, P., Helfenstein, A., Gubler, A., Keller, A., Meuli, R. G., Wächter, D., Lee, J., Viscarra Rossel, R., and Six, J. 2021. Developing the Swiss mid-infrared soil spectral library for local estimation and monitoring, *SOIL*, 7, 525–546, https://doi.org/10.5194/soil-7-525-2021,

Bellinaso Henrique, Demattê José Alexandre Melo, Araújo Romeiro Suzana. 2010. Soil spectral library and its use in soil classification, Rev. Bras. Ciênc. Solo 34 (3), https://doi.org/10.1590/S0100-06832010000300027

Beus, A.A., Grabovskaya, L.I., Tikhonova, N.V., 1976. Environmental geochemistry. Moscow, p. 248

Dalponte, M., ... Frizzera, L., Gianelle, D., 2023. Spectral separability of bark beetle infestation stages: A single-tree time-series analysis using Planet imagery. Ecological Indicators 153. https://doi.org/10.1016/j.ecolind.2023.110349

Ghazaryan H., 2013. Brief Outline of Soils In Armenia. Proceeding the Economic Dimension of Land Degradation, Desertification and Increasing the Resilience of Affected Areas in the Region of Central and Eastern Europe (EDLDIR-2013). Mendel university in Brno press, Czech Republic 2013, ISBN 978-80-7375- 715-1.

Hong, Y., Yu, L., Chen, Y., Liu, Y., Liu, Y., Liu, Y., Cheng, H. 2018. Prediction of Soil Organic Matter by VIS–NIR Spectroscopy Using Normalized Soil Moisture Index as a Proxy of Soil Moisture. *Remote Sens.*, 10, 28. https://doi.org/10.3390/rs10010028

John W. Doran, Timothy B. Parkin, (1994) Defining Soil Quality for a Sustainable Environment, SSSA Special Publications, https://doi.org/10.2136/sssaspecpub35.c1

Joint Research Centre: Institute for Environment and Sustainability, Montanarella, L., Panagos, P. and Yigini, Y., 2013. Soil resources of Mediterranean and Caucasus countries – Extension of the European soil database, Montanarella, L.(editor), Panagos, P.(editor) and Yigini, Y.(editor), Publications Office, , https://data.europa.eu/doi/10.2788/91322

Ma Y., Roudier P., Kumar K., Palmada T., Grealish G., Carrick S., Lilburne L., Triantafilis J. 2023. A soil spectral library of New Zealand, Geoderma Regional, Volume 35, e00726, https://doi.org/10.1016/j.geodrs.2023.e00726

National Atlas of Armenia 2007, Yerevan, Publisher “Tigran Mets”, Yerevan, 232 p.

Silver WL, Perez T, Mayer A,Jones AR. 2021. The role of soil in thecontribution of food and feed. *Phil.Trans. R. Soc. B* 376: 20200181. https://doi.org/10.1098/rstb.2020.0181

Tepanosyan G., Muradyan V., Tepanosyan G., Avetisyan R., Asmaryan S., Sahakyan L., Denk M., Gläßer C. 2023. Exploring relationship of soil PTE geochemical and “VIS-NIR spectroscopy” patterns near Cu–Mo mine (Armenia). Environmental Pollution, Volume 323, 121180, https://doi.org/10.1016/j.envpol.2023.121180

Tepanosyan G., Poghosyan Z., Sahakyan L. 2024. Identification of spatial clusters of potentially toxic elements in different soil types using unsupervised machine learning and compositional data analysis. *Soil & Environmental Health*, Volume 2, Issue 3, 100085, https://doi.org/10.1016/j.seh.2024.100085

Yu B., Yan Ch., Yuan J., Ding N., Chen Z., 2023. Prediction of soil properties based on characteristic wavelengths with optimal spectral resolution by using Vis-NIR spectroscopy, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, Volume 293, 122452, https://doi.org/10.1016/j.saa.2023.122452