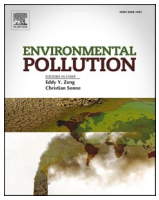
Environmental Pollution 333 (2023) 122066

Contents lists available at ScienceDirect

Environmental Pollution

journal homepage: www.elsevier.com/locate/envpol



Hg and As pollution in the soil-plant system evaluated by combining multispectral UAV-RS, geochemical survey and machine learning☆

L. Salgado [a,b], C.A. López-Sánchez [a], A. Colina [b,c], D. Baragaño [b,d], R. Forján [b,e], J.R. Gallego [b,*]

[a] SMartForest Research Group, Department of Biology of Organisms and Systems Biology, University of Oviedo, 33600 Mieres, Spain
[b] Environmental Biogeochemistry & Raw Materials Group and Institute of Natural Resources and Territorial Planning (INDUROT), University of Oviedo, 33600 Mieres, Spain
[c] Department of Geography, Campus del Milán, University of Oviedo, 33011 Oviedo, Spain
[d] Escuela Politécnica de Ingeniería de Minas y Energía, University of Cantabria, 39316 Torrelavega, Spain
[e] Plant Production Area, Department of Biology of Organisms and Systems Biology, University of Oviedo, 33600 Mieres, Spain

ARTICLE INFO

Keywords:
Arsenic
Mercury
Remote sensing
Random forest
Soil
Plant

ABSTRACT

The combination of a low-density geochemical survey, multispectral data obtained with Unmanned Aerial Vehicle-Remote Sensing (UAV-RS), and a machine learning technique was tested in the search for a statistically robust prediction of contaminant distribution in soil and vegetation, for zones with a highly variable pollutant load. To this end, a novel methodology was devised by means of a limited geochemical study of topsoil and vegetation combined with multispectral data obtained by UAV-RS. The methodology was verified in an area affected by Hg and As contamination that typifies abandoned mining-metallurgy sites in recent decades. A broad selection of spectral indices were calculated to evaluate soil-plant system response, and four machine learning techniques (Multiple Linear Regression, Random Forest, Generalized Boosted Models, and Multivariate Adaptive Regression Spline) were tested to obtain robust statistical models. Random Forest (RF) provided the best non-biased models for As and Hg concentration in soil and vegetation, with R^2 and rRMSE (%) ranging from 0.501 to 0.630 and from 180.72 to 46.31, respectively, and with acceptable values for RPD and RPIQ statistics. The prediction and mapping of contaminant content and distribution in the study area were well enough adjusted to the geochemical data and revealed superior accuracy for As than Hg, and for vegetation than topsoil. The results were more precise than those obtained in comparable studies that applied satellite or spectrometry data. In conclusion, the methodology presented emerges as a powerful tool for studies addressing soil and vegetation pollution and an alternative approach to classical geochemical studies, which are time-consuming and expensive.

1. Introduction

Soil degradation, in particular pollution, affects the supply of ecosystem services such as carbon sequestration, nutrient and water cycles, erosion control, climate regulation, and habitats for plants and animals, as well as opportunities for human development . Industrial and mining areas are often sources of pollutants, causing severe deterioration of environmental compartments, particularly soil . Potentially Toxic Elements (PTEs), such as heavy metals and As, are among the most common soil pollutants, and the evaluation of their abundance, distribution, and mobility is complex (Boente et al., 2022; Forján et al., 2018) and may condition risk assessment and the selection of the most suitable remediation technologies. Wherever potentially polluted soils are used for farming and/or agricultural activities, it is crucial to determine PTE concentrations in order to evaluate health risks (Ballabio et al., 2021; Gil-Díaz et al., 2019). PTEs enter the food chain through crops, pastures and meadows used to rear livestock (Allevato et al., 2019; Manea et al., 2020), and therefore, PTE monitoring not only in soil but also in vegetation is needed (Antonious & Kochhar, 2009; Guerra et al., 2012). Traditional soil and vegetation sampling is carried out on a grid pattern, followed by chemical analysis and subsequent interpretation. However, this approach can be costly, slow, and not well-suited for heterogeneous soils. Indeed, in these cases, multi-element pollution, different sources of pollutants, and external variables (geomorphological, meteorological,

☆ This paper has been recommended for acceptance by Dr Michael Bank.

* Corresponding author.
E-mail address: jgallego@uniovi.es (J.R. Gallego).

https://doi.org/10.1016/j.envpol.2023.122066
Received 14 April 2023; Received in revised form 14 June 2023; Accepted 15 June 2023
Available online 19 June 2023

etc.) should be taken into account (Boente et al., 2022).

To evaluate the health of soil and vegetation, optical remote sensing (ORS) observations provide a reliable alternative approach with a much higher spatial resolution than a sampling grid (Chuvieco & Huete, 2009). However, classical ORS, such as airborne and spaceborne sensors, have spatial and temporal resolution limitations and high operational costs, although images are usually free for users. These limitations can be solved with the use of Unmanned Aerial Vehicle & Remote Sensing (UAV-RS) technologies (Matese et al., 2015; Riveros-Burgos et al., 2021), which are widely used in forestry (Giannetti et al., 2020; Guerra-Hernández et al., 2019), agronomy (Jurado et al., 2020; Rallo et al., 2020; Viera-Torres et al., 2020), climate change research (van Tiggelen et al., 2021), natural risk management (Filkov et al., 2021; Weber et al., 2020), and soil science (Garg et al., 2020; Hout et al., 2020). UAV-RS has been used in soil pollution studies (Choe et al., 2008; Jia et al., 2021a) and, when compared with conventional methodologies, it is faster, less expensive, and non-invasive (Chabrillat et al., 2019). In this context, when present at high concentrations, PTEs interact with natural aggregates in soil (Chakraborty et al., 2017) and alter vegetation growth (Zinnert et al., 2013). Thus, plants are bioindicators of soil health (Gholizadeh & Kopačková, 2019). Moreover, anomalies can be detected with UAV-RS via the following: (i) directly from reflectance information; (ii) from the derivate information provided by spectral indices based on the reflectance properties of coverage; and (iii) from the combination of these data sources (Kureel et al., 2021).

The spectral reflectance measured in remote sensing studies indicates the physical-optical response of plant tissues, thus providing information about plant physiology (Cho et al., 2012; Gholizadeh & Kopačková, 2019) but also the biophysical properties of the canopy, the reflectance of the soil on which the vegetation is growing, the observation geometry, and the characteristics of the light source (Asner, 1998). The presence of high PTE concentrations regulates the amount of energy absorbed/emitted by a surface, thereby making it easier to interpret the influence of these elements on the electromagnetic spectrum (Chakraborty et al., 2014; Jia et al., 2021b). These geochemical alterations can be detected through the chlorophyll dynamics and cell structure of vegetation as there is a correlation between PTE concentrations, vegetation alterations (e.g., stress), and the reflectance in VIS-NIR spectral range (Dunagan et al., 2007; Shi et al., 2016b; Wang et al., 2022).

Irrespective of the methodology used for data acquisition, the mapping of PTE distribution in soil and vegetation is challenging. A series of single-component contamination indices or indicators are commonly mapped (Boente et al., 2022) typically using Geostatistical Interpolation Methods (GIMs), such as kriging, which were developed to calculate the spatial distribution of variables with a relatively homogeneous distribution (Jia et al., 2021a; Leung et al., 2018). However, GIMs are not suitable to study point source pollution. Also, in order to be used properly, these methods may depend on highly correlated predictors or, at least, on the existence of autocorrelation together with a dense sampling to build a valuable spatial model (Demšar, 2006; Zhang et al., 2018). As an alternative, machine learning (ML) techniques have emerged as the most popular approaches to address the limitations of GIMs (Chakraborty et al., 2015; Emadi et al., 2020). Within ML techniques, algorithms that simplify the function to a known form are called parametric ML algorithms. In contrast, non-parametric methods do not make strong assumptions about the form of the mapping function and thus they are free to learn any functional form from the training data. As regards remote sensing, considerable attention has been given to non-parametric techniques (Gao et al., 2022; García-Gutiérrez et al., 2015; Novo-Fernández et al., 2019).

The use of remote sensing technologies can accelerate soil pollution mapping at high resolution and at a lower cost and in less time than soil survey approaches. In this regard, recent studies have addressed the electromagnetic spectrum for modelling the concentrations of chemical elements in soil, both with proximal (e.g., Gholizadeh et al., 2020) and remote (e.g., Boente et al., 2020) sensors. On occasions, ML techniques have been applied to improve the interpretation (e.g., Jia et al., 2021b), although the models obtained are usually complex and the role of the variables selected may be difficult to interpret (Jia et al., 2021a). In fact, these studies reported several shortcomings. In this regard, for instance, Gholizadeh et al. (2020) obtained results that were difficult to extrapolate to continuous spatial information because the spectral data collected were punctual. Jia et al. (2021b), in turn, identified As point sources using spectral information combined with distances to potential sources and to receptors, thus achieving continuous spatial information but not based solely on spectral data. On the whole, challenges remain and further research into combining remote sensing and ML for soil pollution mapping is required.

Given the preceding considerations, here we propose a novel methodology based on the combination of a low-density sampling with the potential of multispectral data obtained by UAV-RS, both followed by the selection of the most appropriate ML technique to interpret the results. We hypothesized that this approach addresses the following specific objectives:

- Effectiveness to cover the above-mentioned research gaps and the issues of mapping point-source pollution.
- Capability to focus simultaneously on soil and vegetation (soil-plant system).
- A site heavily affected by PTE pollution was selected to exemplify the capacity of the approach proposed to attain continuous spatial information about As and Hg concentrations in soil and vegetation.

2. Materials and methods

Given the complexity and variety of the methodologies used in this study, a flow chart is presented in Fig. 1.

2.1. Study area

The study area is located in the surroundings of La Soterraña mine in Asturias (NW Spain) (Fig. 2), which was one of the largest Hg-producing regions in the world during the 20th century (Fernández et al., 2020). Hg exploitations are usually associated with environmental risk due to the abundance of PTEs such as Hg and As (Gallego et al., 2015; Higueras et al., 2006), which are highly toxic to most organisms (D'Aniello et al., 2018; Dadová et al., 2016).

In the past, the study site also hosted extensive metallurgical activities, which produced high volumes of waste. Mining and metallurgical waste were disposed of in two spoil heaps and several stockpiles, which show high concentrations of PTEs (Fernández et al., 2020). These heaps are affected by wind erosion and surface runoff and thus have greatly disturbed the surrounding soils (Boente et al., 2022). In parallel to mine-exploitation dynamics, the land coverage of the study area is associated with small-scale agriculture and livestock farming, thus grass and crops cover intermingle with forestry (Fig. 2).

2.2. Geochemical approach

2.2.1. Sampling design

Given the topography of the area and the results of previous surveys (see Boente et al., 2020, 2022 and references therein), 46 soil sampling stations were established in grasslands, meadows, and ditches in the surroundings of the former mining-metallurgy site (Fig. 2). These stations were georeferenced with a Spectra SP60 GNSS receiver (Trimble, California, USA) in Real-Time Kinematic Position (RTK), with a precision of 8 mm ± 1 ppm horizontal, and 15 mm ± 1 ppm vertical. Each station covered a 50-cm side square in which soil and herbaceous vegetation were sampled.

For soils, at each sampling station, a 1-kg composite sample comprising four increments of the first 5 cm of soil was collected using a

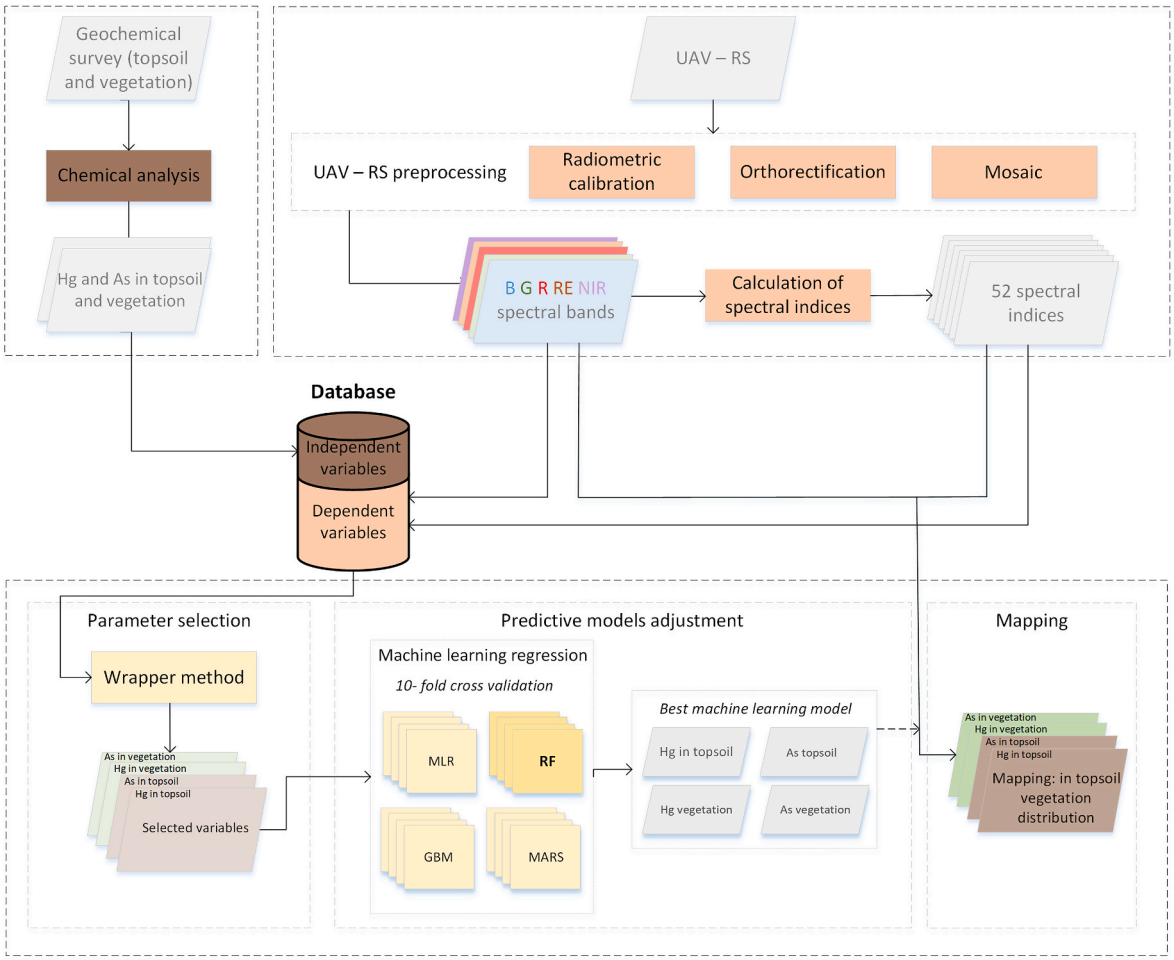


Fig. 1. Process workflow.

Dutch Edelman probe. In turn, vegetation sampling was focused on the aerial parts of all herbaceous plants present in the squares. Both soil and vegetation samples were conserved in sterilized plastic bags and stored at 4 °C until analyses.

2.2.2. Chemical analyses of soil and vegetation

Soil samples were air-dried at room temperature (<25 °C), thus avoiding Hg loss by volatilization. They were then sieved through 2-mm mesh to remove large particles, and material below 2 mm was quartered in a Jones riffle to obtain representative subsamples. Finally, an aliquot of 1 g soil was ground in an agate mortar to achieve a particle size of <100 μm. These samples were sent to an ISO 9002 and ISO-17025 accredited laboratory (Acmelabs, Vancouver) for chemical analysis. The concentrations of As and Hg were determined by ICP-MS (Inductively Coupled Plasma-Mass Spectrometry) after aqua regia extraction, with detection limits of 2 mg kg^{-1} for As and 0.5 mg kg^{-1} for Hg.

Vegetation samples were dried in an oven at 30 °C, ensuring no Hg volatilization. They were then ground using an Ultra Centrifugal Mill ZM 200 (Retsch, Verder Scientific GmbH & Co. KG, Haan, Germany), obtaining a homogeneous subsample with a grain size below 40 μm. Acid digestion using a mixture of H_2O_2 and HNO_3 (1:6 v/v) in a microwave oven (Milestone ETHOS 1) was carried out, and Hg and As were then determined with ICP-OES (Optima 4300 DV; PerkinElmer) using detection limits of 0.036 mg kg^{-1} and 0.04 mg kg^{-1}, respectively.

To address the capacity of vegetation to accumulate As or Hg, the transfer coefficient (TrC) (Busuioc et al., 2011; Forján et al., 2018) was calculated as follows:

$$TrC = \frac{C_V}{C_T}$$

Expression 1. Transfer coefficient of PTE concentrations between vegetation and topsoil. In this case, C_V stands for As or Hg concentration in vegetation and C_T for As or Hg concentration in topsoil.

2. 3Multispectral data approach

2.3.1. Data acquisition

Multispectral data were obtained from a UAV-RS flight. The flight coverage is shown in Fig. 1 and consisted of a surface area of 69 ha downstream of the spoil heaps. A P4 Multispectral UAV-RS from SZ DJI Technology Co. Ltd ® (Nanshan, Shenzhen, China), designed for precision agriculture, was used. This device integrates a multispectral camera composed of six individual CMOS sensors, one RGB, and five narrow-band sensors: blue (B), green (G), red (R), red edge (RE), and near-infrared (NIR) (Table 1).

This UAV-RS system also integrates a sunlight sensor (which provides absolute reflectance measurements without calibration), a GNSS high-precision RTK positioning module, and the TimeSync system, which gives precise and real-time positioning data for each image, thus optimizing photogrammetric results and providing centimeter-level accuracy without Ground Control Points (GCPs).

Due to the complex orography, the photogrammetric survey was carried out with a transversal and longitudinal overlap of 70% and at an average height of 100 m. The flight was conducted during the early summer of 2020 to obtain radiometric information on the herbaceous

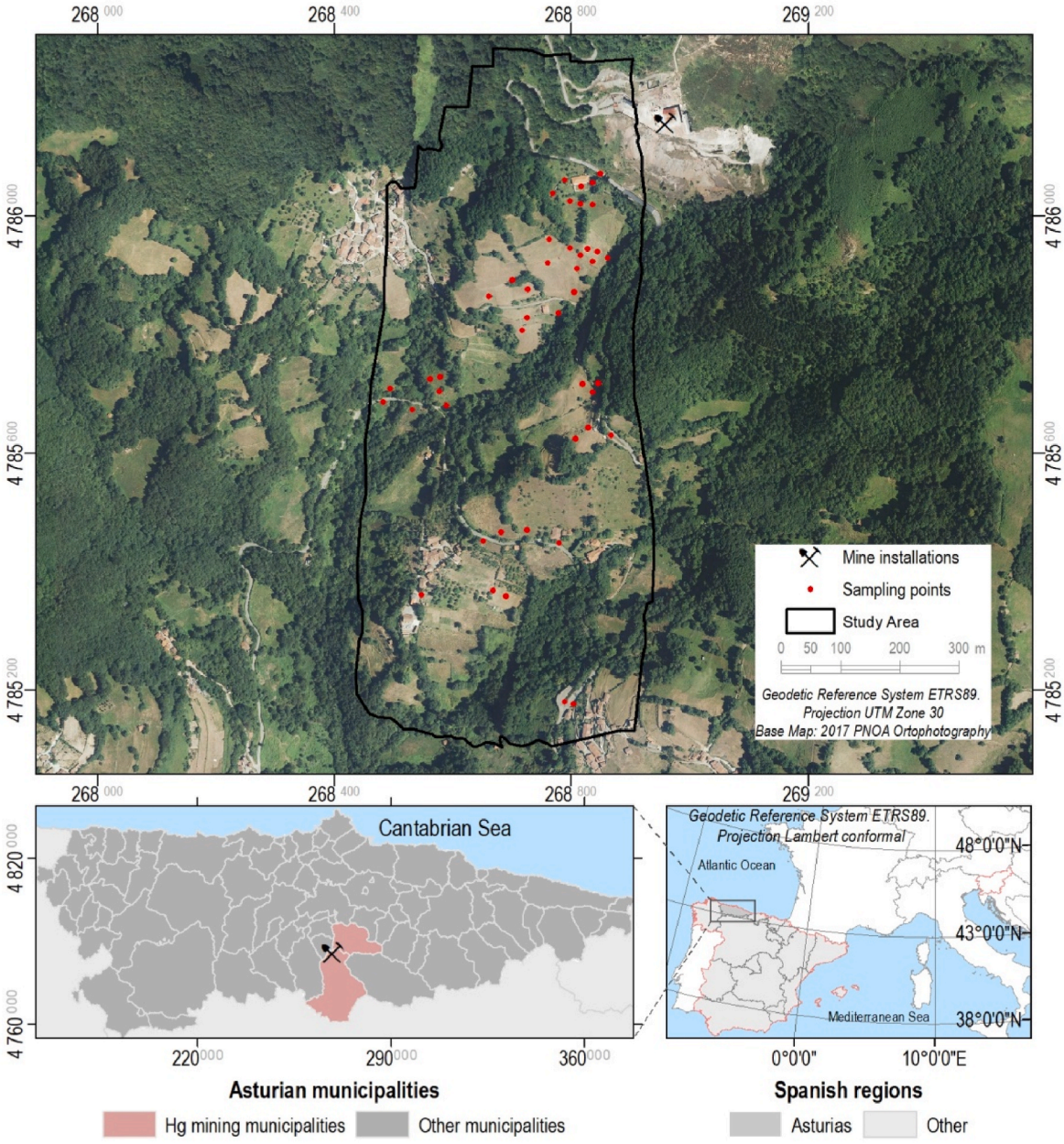


Fig. 2. Location and general view of the study area.

Table 1
Relation between bands and the electromagnetic spectrum.

Band	Central wavelength (nm)	Wavelength width (nm)
B	450	16
G	560	16
R	650	16
RE	760	16
NIR	840	26

vegetation, which was fully grown and absorbed the soil components at this time. In addition, in this period of the year, the soil was not affected by extreme weather conditions such as heavy rains, or alterations caused by agro-livestock activities such as mowing, which started in the weeks after the flight.

Data processing was carried out with the photogrammetry and drone mapping software PIX4D Mapper (Pix4D S.A., Prilly, Switzerland), which is useful to obtain optimal results with multispectral data from UAV-RS (Boente et al., 2020; Handique et al., 2020). The steps followed included point cloud generation, 3D model construction, feature extraction, and multispectral band generation.

2.3.2. Analysis and interpretation

2.3.2.1. a grassland areas. Grassland areas were identified using a non-supervised classification carried out with the minimum distance algorithm included in the SCP package from QGIS 3.4.8 (Zürich, Switzerland) and post-processed manually. The classification distinguished between two coverages: (i) low vegetation, which included grassland, pasture, and crops; and (ii) coverage associated with forestry and unproductive areas, which included bare ground, roads, and population settlements.

2.3.2.2. b Spectral bands used. Fig. 3 shows the spectral location of each band for the P4 Multispectral UAV-RS used in this work. The G band is

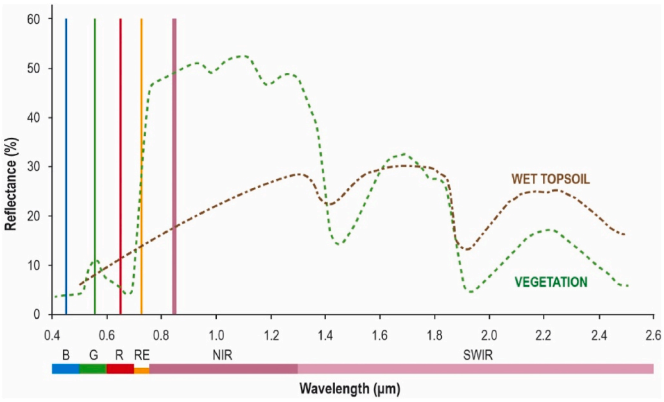


Fig. 3. Typical spectral response specification of wet topsoil and healthy green vegetation (Hoffer, 1978) and distribution of P4 multispectral bands used in this study (vertical lines). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

located in the highest reflectance of the VIS spectral range for vigorous vegetation, while the B and R bands are in the low part. The RE band used is located in the higher slope area of reflectance and the NIR band close to one of the peaks. In this context, healthy plant spectra are characterized by strong absorption in the visible spectrum (450–680 nm) due to chlorophyll, and strong reflectance in the NIR range (700–1200 nm) as a result of the internal scattering of light at cell wall-air interfaces (Dunagan et al., 2007). Also, the RE range (680–740 nm) presents the highest increase of reflectance, being sensitive to variations in leaf chlorophyll absorption (Gitelson & Merzlyak, 1996).

2.3.2.3. c Spectral indices. A spectral index is the combination of multiple single bands through algebraic operations (Boente et al., 2020; Stroppiana et al., 2012). In this regard, for this study, the following classification of four types of spectral indices was established:

(i) Spectral vegetation indices: Due to the interaction of Hg and As with normal vegetation development (Shi et al., 2016a), spectral vegetation indices can provide information about soil degradation as they can detect vegetation stress and vitality by measuring chlorophyll and cellular alterations. Examples of such indices include SIPI (Blackburn, 1998; Penuelas et al., 1995), OSAVI (Rondeaux et al., 1996), GNDVI (Gitelson et al., 1996; Gitelson & Merzlyak, 1994), NDVI (Rouse et al., 1974) and SAVI (Huete, 1988). For more details see Table S1.

(ii) Spectral polluted soil indices: Regarding PTEs, several indices efficiently estimate pollution stages, including DVI (Kooistra et al., 2010), MSAVI (Kooistra et al., 2004), CSVI (Zhang et al., 2017), ADVI (Zhang et al., 2019), HMSSI (Zagajewski et al., 2017), and REP (Clevers et al., 2010). For more details see Table S2.

(iii) Spectral colour indices: Unusual concentrations of PTEs, including As and Hg, impair the function of photosynthetic pigments (Siddiqui et al., 2020; Zhang et al., 2021) and can even cause oxidative stress (Sahu et al., 2011), which directly alters the normal coloration of vegetation (Mei et al., 2021). When studying these alterations, spectral indices are commonly used, for example CI (Escadafal, 1989) and CI2 (Escadafal et al., 1994) as colour indices, BI as a brightness index (Escadafal, 1989), HI as a shape index (Escadafal et al., 1994), RI and SI as redness and saturation indices (Mandal, 2016), and Y and HUE as intensity indices (Escadafal et al., 1994; Liu & Paulsen, 2000). For more details see Table S3.

(iv) Spectral texture indices: These were achieved with the "*r.texture*" tool from Geographic Resources Analysis Support System (GRASS) version 7.6.1., implemented in the software Quantum GIS (QGIS) version 3.4.8. (Zürich, Switzerland). This tool creates spectral texture indices based on other indices and, in our case, NDVI was chosen. Note that to define the texture of the coverage, several variables can be considered, including variations in grey levels, the presence of regular patterns, tone or grey level intensities, and spatial relationships (Haralick et al., 1973). For more details see Table S4.

2.4. Data analysis

2.4.1. ML techniques

The selection of an appropriate ML technique to handle the dataset obtained was one of the main tasks in this study. In fact, ML success is determined by the correct dimensional reduction of independent variables from the original set. Four techniques (one parametric and three non-parametric) were considered. These techniques were selected on the basis of their ease of implementation with our database after several attempts (data not shown) using a similar approach to that described by Biney et al. (2023). The rationale for the selection of the four specific ML approaches is described below. R software (Ihaka & Gentleman, 1996) was used to carry out calculations.

(a) One of the most commonly used parametric techniques is Multiple Linear Regression (MLR) as the models obtained are easier to interpret than other advanced multivariate algorithms (Chung et al., 2020; Ghazali et al., 2010). In remote sensing studies, MLR is a very common approach (Golchoubian et al., 2012; Roces-Díaz et al., 2014; Novo-Fernández et al., 2019). Nevertheless, MLR is based on the fundamental least squares assumption of independence and equal distribution of errors with zero mean and constant variance, which can be breached by non-normality, multicollinearity of variables, and heteroscedasticity of error variance. These limitations can be avoided using non-parametric techniques.

(b) The Random Forest (RF) technique, first proposed by Breiman (2001), is a widely used non-parametric classification and regression approach consisting of a large number of decision trees. Each ensemble tree is developed with distinct independent variables, which are randomly selected from the original dataset. RF algorithms owe their greater stability and better accuracy to the use of randomized sampling with a large number of trees, which allows them to outperform models based on individual trees (Immitzer et al., 2016). The final regression estimate for each sample is obtained as a weighted mean value of the estimates of a large number of individual trees. As the number of trees increases to the maximum number defined by the user, the variance of trees is high while the bias is low (Breiman, 2001). RF techniques have been increasingly used in remote sensing for environmental applications. They have achieved better results and speed of processing than other predictive techniques (Belgiu & Drăgu, 2016; Gao et al., 2022; Novo-Fernández et al., 2019) while presenting results that are relatively insensitive to the number of input number and their multicollinearity (Gislason et al., 2006).

(c) Generalized Boosted Models (GBMs) were developed in the early 2000s (Friedman, 2002) . This algorithm is a flexible non-parametric ML technique for classification. GBMs are a combination of two techniques, decision tree algorithms, with the creation of successive weak tree ensembles, and boosting methods, where each tree learns from the previous ones. GBMs repeatedly fit multiple decision trees to improve accuracy. These models show the effect of each predictor after evaluating the influence of others. In addition, they are characterized by robustness when there are missing values and outliers. The use of

this algorithm is growing in remote sensing research (Henrique Brant Dias et al., 2021; Tavera Busso et al., 2021).

(d) Multivariate Adaptive Regression Splines (MARS) were proposed by Friedman (1991) as a method for flexible regression modelling of high-dimensional data. MARS is a non-parametric regression or classification, based on the fitting of linear regression models in different space intervals of independent variables. The process is also executed between dependent variables and the model can reduce additive effects and interaction between dependent variables. One of the strengths of this method is the processing of missing values and the prevention of overshooting by a self-test. MARS has recently been used in remote sensing for environmental applications (Jafari et al., 2022; Quang et al., 2022).

2.4.2. Feature selection

As mentioned above, an optimal set of variables is crucial to identify and eliminate irrelevant data and thus improve computational performance, which in turn brings about savings in cost and time. In this regard, feature selection is focused on the reduction of the degree of overfitting in the training dataset, thus addressing a major drawback of ML techniques (Novo-Fernández et al., 2019; Xu & Wang, 2010).

In our case, Wrapper methods were used to select the subsample of predictor variables (features) that provided the best results (Zhiwei & Xinghua, 2010). To this end, we used the *CVParameterSelection* method implemented in WEKA software (Hall et al., 2009; Hornik et al., 2009) from Waitako University (New Zealand).

2.4.3. Performance evaluation

Supervised learning algorithms, including train models and test data, were checked with the k-fold cross-validation method. In this widely used approach, the data set is divided into k subsets, in this case 10. Each subset is used as test data and the other nine sets are used as training sets each time the algorithm is executed. Each model runs 100 times, which implies a 10-fold cross-validation of each subset repeated 10 times per training dataset.

To evaluate the quality of model fit, the following statistics were calculated: (i) the pseudo-coefficient of determination (R^2); (ii) relative root mean squared error of cross-validation (rRMSE, %); (iii) the residual prediction deviation (RPD), i.e., the ratio between the standard deviation to the RMSE (Vaudour et al., 2022); and (iv) the ratio of performance to interquartile distance (RPIQ) (Bellon-Maurel et al., 2010). Specifically, RPD and RPIQ were included in the study as robust statistics that are not biased by the high variability of the data (see below Table 2). The relation between the four ML methods considered was evaluated using the paired *t*-test (corrected) ($\alpha = 0.05$), which is based on Student's t-criterion.

3. Results

3.1. Geochemical study

Regarding pollution levels, Table 2 and Figs. S1 and S2 show the total concentrations of Hg and As in soil and vegetation at the 46 sampling points.

The total Hg content in soil ranged from 1.90 to 616.00 mg kg^{-1}. The higher values were located in sampling stations near the spoil heaps (Fig. S1), where herbaceous cover for livestock was observed. The total As content in soil ranged from 86.00 to 3170.00 mg kg^{-1}, again with the higher values in the vicinity of the spoil heaps (Fig. S1). The differences between the mean and median, and the high percentages of coefficient of variation (CV) (Table 2) revealed a notably heterogeneous distribution of As and Hg.

In vegetation, Hg concentration ranged between 0.07 and 0.75 mg kg^{-1}, and again the higher levels corresponded to samples taken close to the spoil heaps. However, a higher dispersion of Hg in vegetation than in soil was observed (Fig. S2) as some relevant values were found in samples far from the spoil heaps. The total As concentration in vegetation ranged from 0.30 to 7.70 mg kg^{-1}, with the highest values being found in a heterogeneous geographical distribution. Like PTE concentrations in the topsoil, the differences between the mean and median, and the high percentages of the CV (Table 2) also revealed an irregular distribution, especially for As.

Soil samples were characterized as neutral acid, with pH values of between 7.62 and 4.71 and a mean of 6.28. The organic matter content (%) in the topsoil ranged from 10 to 46%, the lowest values being found near the mining area. In the soil/plant system, TrC showed slightly higher values for Hg than As in most of the sampling points (Table S5).

As a result of supervised classification and post-processing, the study area selected was finally reduced to 16 ha, which included all low vegetation (herbaceous plants) present in the flight area. These plants were classified into two groups: (i) meadow vegetation; and (ii) gutter, road and heap grass. Meadows comprised species associated with *Arrhenatherion elatioris* or mountain hay meadows, and the major families found were *Peaceae, Primulaceae, Apiaceae, Araceae, Asteraceae, Caprifoliaceae, Caryophyllaceae, Convolvulaceae, Dennstaedtiaceae, Fabaceae, Lamiaceae, Linaceae, Plantaginaceae, Polygalaceae, Polygonaceae, Ranunculaceae, Rosaceae,* and *Rubiaceae.* Gutter, road, and heap grass included species associated with the *Equisetaceae, Rosaceae, Hypericaceae, Lamiaceae, Caprifoliaceae, Poaceae, Plantaginaceae,* and *Rubiaceae* families. A detailed list of the species observed in each type of grass vegetation is described in Table S6.

3.2. Geostatistical analysis and adjustment of predictive models

3.2.1. Feature selection process

Four response (dependent) variables were considered: Hg and As in topsoil, and Hg and As in vegetation (henceforth Hg_s, As_s, and, Hg_v, As_v respectively), whereas 57 predictor variables or features were initially considered (i.e., 5 bands, 23 spectral vegetation indices, 8 spectral polluted soil indices, 8 spectral colour indices, and 13 spectral texture indices). Once feature selection with Weka software (Wrapper methodology) had been completed, 6, 9, 6, and 7 significant features were chosen for Hg_s, Hg_v, As_s, and As_v, respectively from the 57 initially available (see section 3.2.3 for details on the different features selected for each response variable).

3.2.2. Adjustment of predictive models

Predictive models for Hg_s, Hg_v, As_s, and As_v were obtained with R software once the specific features selected had been considered. The goodness-of-fit statistics given by the regression methods used to model the response variables are shown in Table 3.

As indicated in the preceding table, RF showed the best R^2 values in all cases, and therefore it could be initially the regression technique of choice in this study. Higher values of rRMSE were obtained in all cases

Table 2
Descriptive statistics of PTE concentrations (mg kg^{-1}) in topsoil and vegetation.

	Concentration (mg kg^{-1}) in topsoil						Concentration (mg kg^{-1}) in vegetation					
	mean	median	min	max	Std.	CV	mean	median	min	max	Std.	CV
Hg	48.41	9.00	1.90	616.00	115.34	238.24	0.27	0.23	0.07	0.75	0.17	62.32
As	514.04	247.00	86.00	3170.00	645.34	125.54	2.14	1.34	0.30	7.70	1.92	89.67

Table 3
Goodness-of-fit statistics given by ML for Hg_s, Hg_v, As_s, and As_v (mg kg^{-1}). Methods used were Multiple Linear Regression (MLR), Random Forest (RF), Generalized Boosted Models (GBMs), and Multivariate Adaptive Regression Splines (MARS). All values represent the mean of 100 model runs (i.e., 10 replicates, each with 10-fold cross-validation). The performance of the regression methods was compared by using different statistics based on the model errors (mean field values were considered as true values): pseudo-coefficient of determination (R^2), relative root mean square error (rRMSE, as a %), the Ratio of Performance to Deviation (RPD) and the Ratio of Performance to Interquartile Distance (RPIQ). The best results are indicated in bold.

Regression methods	Statistics	Hg_s	As_s	Hg_v	As_v
MLR	R^2	0.41	0.31	0.48	0.47
	rRMSE	290.42	149.65	96.83	84.05
	RPD	0.82	0.84	0.64	1.07
	RPIQ	0.17	0.67	0.98	1.67
RF(1)	R^2	**0.50**	**0.53**	**0.57**	**0.63**
	rRMSE	**180.72**	**91.61**	**46.31**	**66.67**
	RPD	**1.32**	**1.37**	**1.35**	**1.35**
	RPIQ	**0.27**	**1.09**	**2.06**	**2.10**
GBMs(2)	R^2	0.40	0.41	0.49	0.48
	rRMSE	183.09	104.35	52.74	73.26
	RPD	1.30	1.20	1.18	1.22
	RPIQ	0.26	0.96	1.81	1.91
MARS(3)	R^2	0.422	0.38	0.51	0.42
	rRMSE	210.37	148.14	615.63	89.89
	RPD	1.13	0.85	1.01	1.00
	RPIQ	0.23	0.67	1.55	1.56

(1) 1500 trees, 1–15 splits.
(2) 1500 trees, shrinkage parameter 0.1.
(3) 3°, nprune = 2,4,6,8,10.

for the topsoil (Hg_s and As_s) as a consequence of the high variability of the input data (Table 2), whereas the lower values observed for vegetation (Hg_v and As_v) are coherent with lower variability of the input data and a better R^2 (Table 3). When comparing the rRMSE values between the different methods, the best results were obtained again with RF in all cases (Table 3). In addition, RPD and RPIQ, both unbiased statistics, presented also the best results for RF. Concretely, RPD showed values close to 1.40 for the four variables whereas RPIQ revealed higher variability, with the minimum value below 1 (Hg_s), and the values associated with vegetation for both elements above 2 (Table 3). It must be also pointed out that, as regards PTE concentrations, and irrespective of the ML technique used, the goodness-of-fit statistics of the models were better for Hg and As concentrations in vegetation than in topsoil (Table 3).

Predicted (estimated with RF) vs. observed (geochemical data) values for Hg_s, Hg_v, As_s, and As_v were compared in a graphical analysis (Fig. 4). A linear model fitted to the scatter plot did not reveal any important problems related to heteroscedasticity or lack of normality.

In general terms, the results were more accurate in vegetation than in topsoil, while for chemical elements they were more precise for As than for Hg. The best R^2 found was for As in vegetation, with a value of 0.630. In brief, the following conclusions were drawn:

(i) The analysis of vegetation models (Fig. 4C and 4D) showed the best fit with the 1:1 line among tested models, i.e., the models were capable of capturing the relationship between variables and thus generating reliable prediction. The Hg_v model tended slightly to underestimate high concentrations and to overestimate low concentrations whereas the As_v model revealed a very slight overestimation at the highest concentrations and almost negligible underestimation at low concentrations.
(ii) As_s model (Fig. 4B) was acceptable for low concentrations but got worser as concentrations were higher (specially above of 1000 mg kg^{-1}).
(iii) Hg_s model (Fig. 4A) was biased in most of the ranges with, both notable, overestimation at high concentrations and underestimation at low concentrations.

3.2.3. Variable importance

As mentioned above, the feature selection process defined a small number of predictor variables from the initial set of 57; thus, once the RF algorithm had been applied, the model obtained in R software revealed the different importance of these predictor variables in the four study cases (Hg_s, Hg_v As_s, and As_v). The percentages of importance of the different features (predictor variables) were graphed (Fig. 5).

For Hg_s, the most important feature group was the spectral polluted soil indices as, remarkably, the ADVI spectral index accounted for approximately 50% of the normalized score in the model (Fig. 5). There were other relevant groups of indices, such as the spectral colour indices (close to 30% of normalized score for CI2), and spectral vegetation indices and the green band (less than 10% in both cases). In the case of As_s, within the nine significant features found, the most important group was the spectral vegetation indices (49%), which included the MTCI index with >40% of the normalized score. Spectral colour indices, spectral texture indices, and spectral polluted soils indices accounted for 32%, 11%, and 7% respectively. As regards Hg_v, the group of spectral vegetation indices was the most important (above 60%) and the variable MTCI spectral index presented the highest normalized score (well above 30%), whereas the spectral colour indices achieved 33% and spectral polluted soils indices only 3%. Finally, as a result of the feature selection process for As_v, spectral vegetation indices (39%) and spectral colour indices (38%) accounted for 38% with SI slightly above 20% of the normalized score, very similar to the. Of note, as it is directly captured by the red channel of the UAV-RS, the red band achieved more than 20% of the normalized score.

3.3. Mapping

As explained above and having considered the goodness-of-fit statistics and the graphical analysis of predictive model adjustment, RF emerged as the ML technique with the best accuracy. The spatial distributions of Hg and As in the study area resulting from the application of RF, both for topsoil and vegetation data, are shown in Fig. 6.

The high spatial resolution of the maps revealed how Hg_s, Hg_v, As_s, and As_v vary depending on the distance to the main pollution source (spoil heaps). Hg and As in topsoil (Fig. 6) presented the highest concentrations near the spoil heaps, but Hg showed a much lower dispersion throughout the study area than As. Concentrations of these two PTEs in vegetation were distributed similarly, revealing the highest concentrations of Hg mostly close to the spoil heap and greater dispersion of high contents for As.

4. Discussion

Given the capacity of UAV-RS to detect geochemical anomalies remotely, these technologies are being increasingly used in mining exploration (Booysen et al., 2020; Tejado-Ramos et al., 2021). With similar premises, remote sensing studies on soil pollution are being established using satellite technologies, alone or combined with spectrometry techniques (Gholizadeh et al., 2018; Khosravi et al., 2022), and also by UAV-RS systems (Jia et al., 2021a).

In this context, soil alterations caused by the presence of metal (loid)s have been related to spectral changes in the visible region (Khosravi et al., 2018; Kooistra et al., 2001) as evidenced in the Hg topsoil model (Hg_s), where the green band was one of the features selected in accordance with Tucker (1979) and ADVI (Zhang et al., 2019) was shown as the most important index. Thus, the triggers of vegetation spectral response are in this study clearly linked to the high damage in vegetation caused by Hg (Baragaño et al., 2022); in addition, the most influent spectral range to Hg_v predictive model was red-red edge, an spectral

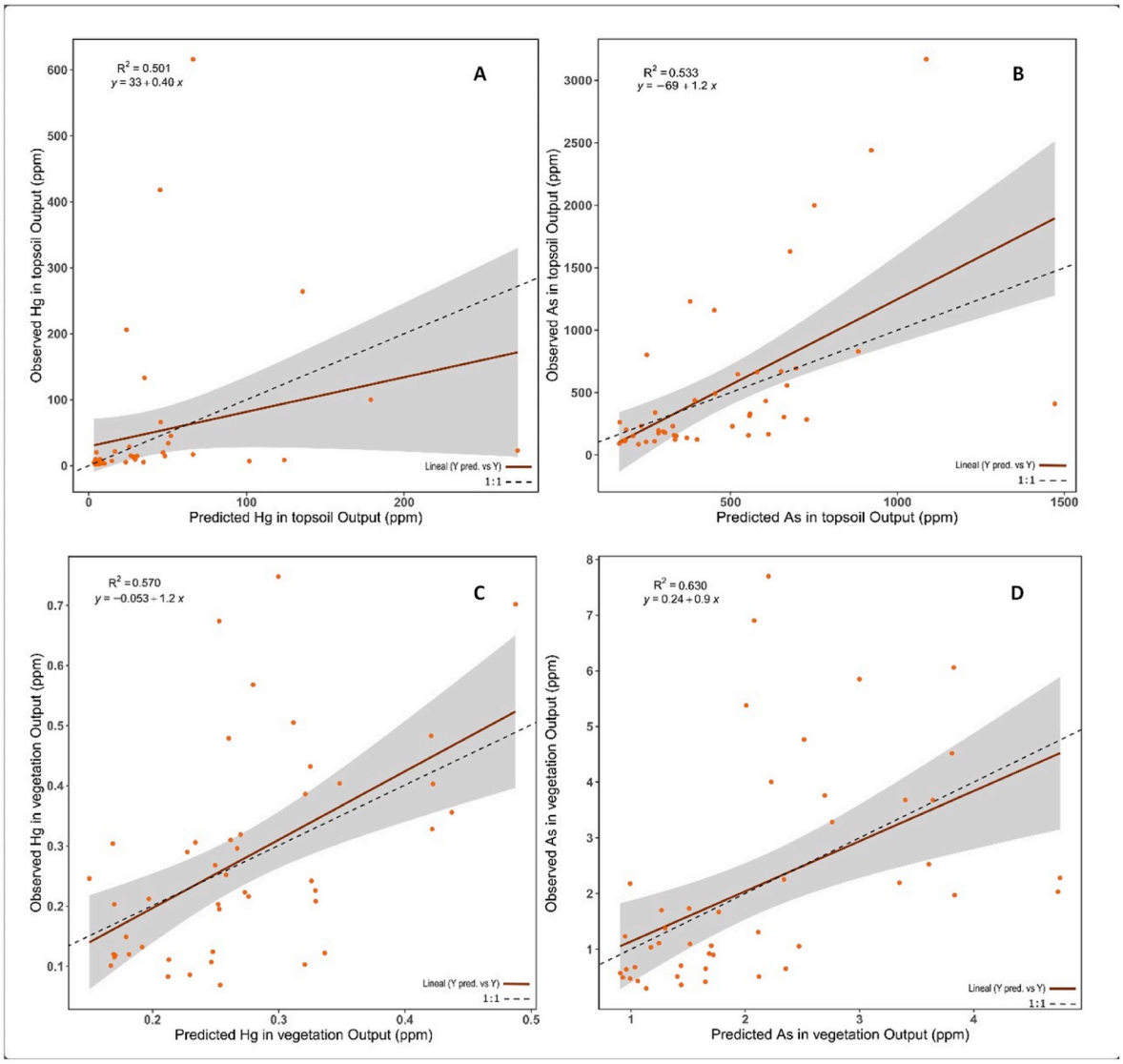


Fig. 4. Graphic representation of predicted vs. observed values (RF model) for Hg_s (A), As_s (B), Hg_v (C), and As_v (D). Values on the X-axis correspond to predicted concentrations (mg kg^{-1}) and those on the Y-axis to observed concentrations (mg kg^{-1}).

area strongly linked to alterations in chlorophyll and vegetation cell structure (Hunt et al., 2012; Mandal, 2016), In turn, soil alterations caused by As have been linked to spectral changes in the Red-NIR region (Wei et al., 2019). This effect was evidenced in the As topsoil model (As_s) as the MTCI (Dash & Curran, 2004) and SI (Mandal, 2016) indices were relevant, whereas the most influent spectral area for As_v predictive model was red, which is also strongly linked to alterations in chlorophyll as mentioned above, a typical symptom produced by As in vegetation (Fernández et al., 2020). On the whole, regarding the spectral index classification performed herein, the spectral vegetation index category was the most important in all cases, including topsoil models, and this corroborated the notable and direct relation between PTEs in topsoil and vegetation observed in the geochemical data (Forján et al., 2017). The meaning of these indices is related to the inhibition of enzyme production and the subsequent destruction of chlorophyll in foliage, and to the damage of the cell structure of the plants (Dash & Curran, 2004; Hunt et al., 2012; Mandal, 2016). Furthermore, these symptoms have been locally identified in the vegetation present in the study area (Baragaño et al., 2022).

Consistent with these results, Metternicht (2010) and Verdebout et al. (1994) associated contaminant-induced changes in soil with vegetation health, especially alterations in foliar pigments, chlorophyll activity, and carotenoids, pointing to spectral alterations between 450 and 550 nm (Blue-Green region), as demonstrated in this case by the shifts in GNIR (Bausch & Duke, 1996), MCARI (Daughtry et al., 2000) and SI (Mandal, 2016) in the Hg_v and As_v models. In addition, the efficacy of Red-NIR and Red Edge spectral regions was previously reported (Dunagan et al., 2007; Horler et al., 1983; Sims & Gamon, 2002; Tucker, 1979) in vegetation studies. In this regard, it was observed a relevant role of MTCI (Dash & Curran, 2004) for the Hg_v and the red band and the CI red edge index for As_v models (Hunt et al., 2012).

The estimation of As and Hg spatial distribution in topsoil and vegetation was done by ML techniques, using one parametric method (MLR) and three non-parametric (RF, GBM and MARS) ones, the latter yielding more robust models, thus consistent with previous studies on remote sensing data (Mouazen et al., 2021; Novo-Fernández et al., 2018; Tan et al., 2019). Of note, the RF technique outperformed the other methods tested and therefore emerges as a promising strategy to study the spatial distribution of geochemical elements, both in topsoil and vegetation. In this regard, and consistent with a recent study by Wang et al. (2022), the RF models provided a reliable estimation for distinguishing high and low values, $0.50 < R^2 < 0.65$, with a low-cost and non-invasive methodology. Indeed, the models were unbiased enough, with the exception of Hg_s, and the spectral data had a high spatial

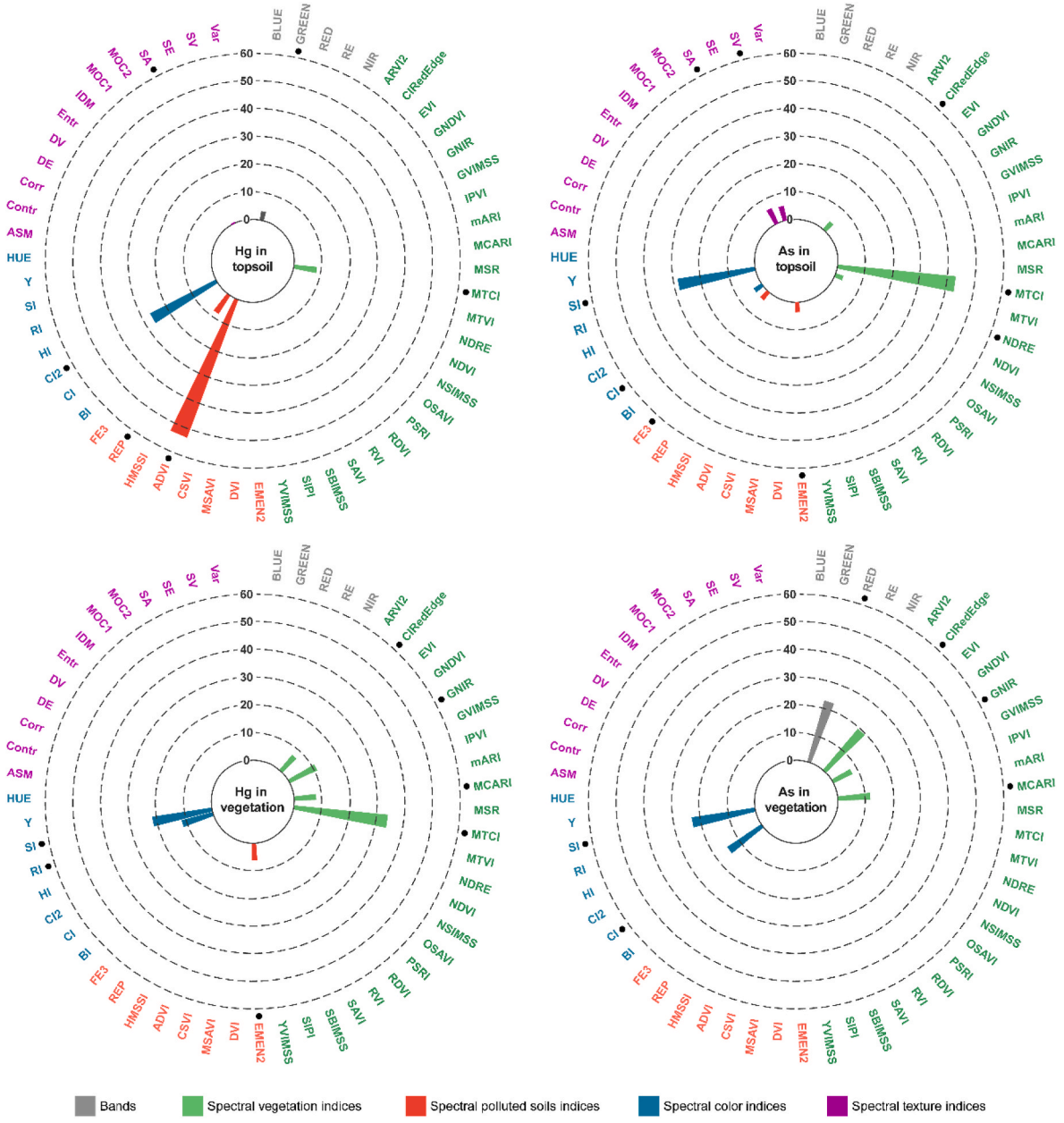


Fig. 5. Type of predictor variable and percentage of importance in the RF models obtained for Hg_s (top left), Hg_v (bottom left), As_s (top right), and As_v (bottom right).

resolution, which is crucial for the mapping of polluted soil (Khosravi et al., 2021). Therefore, our results could be categorized primarily as representative, also given the adjustment between predicted and observed values (Fig. 4); nevertheless an additional evaluation (see below) based on performance figures-of-merit is required for evaluating the quality of model fit.

R^2 values were better for As than Hg, and Hg presented lower error values for vegetation than for soil in terms of rRMSE (Table 3), although these statistics are notably affected by the high variabilitiy of As and, specially, Hg concentrations shown in Table 2. Also, these results may be conditioned by the different transfer ratios of As and Hg from the soil to the plant given that the TrC was higher for Hg than As (Table S5), which implies a higher transfer of Hg to vegetation (Forján et al., 2018). Therefore, we hypothesize that, when As and Hg are translocated to the aerial part, plants are killed earlier in the case of As (McLaughlin et al., 1999). In addition, it should be noted that As and Hg have distinct translocation and accumulation strategies depending on the amount taken up (Zhao et al., 2022). As regards RPD, and according to Casa et al. (2013), values close or above to 1.4 imply prediction ability, as occurs in this study for the four variables studied without any notable diference. However, the values of RPIQ obtained were notably better for vegetation models and in particular, the value of 0.267 for Hg is too low and reveals that, in spite of the acceptable levels found for R^2 and RPD, the Hg_s model may not supply reliable predictions. Note also that the calculation RPIQ for Hg_s was done on a skewed data distribution as also reflected in Fig. 4.

Gholizadeh et al. (2018) and Jia et al. (2021a) highlighted that images with high spatial resolution are frequently insufficient in delivering timely and frequent data with adequate spatial resolution. Nonetheless, this challenge can be effectively tackled by utilizing UAV-RS as a useful tool for monitoring and mapping soil pollution because of its spatial and/or temporal resolutions. In this context, this study revealed equal or superior statistics to other recent remote sensing studies examined although, in our case, cross validation could be a drawback. For instance, Alvyar et al. (2022) carried out a regional study with LANDSAT 8 data and obtained only $R^2 = 0.25$ with RF modelling for As a much

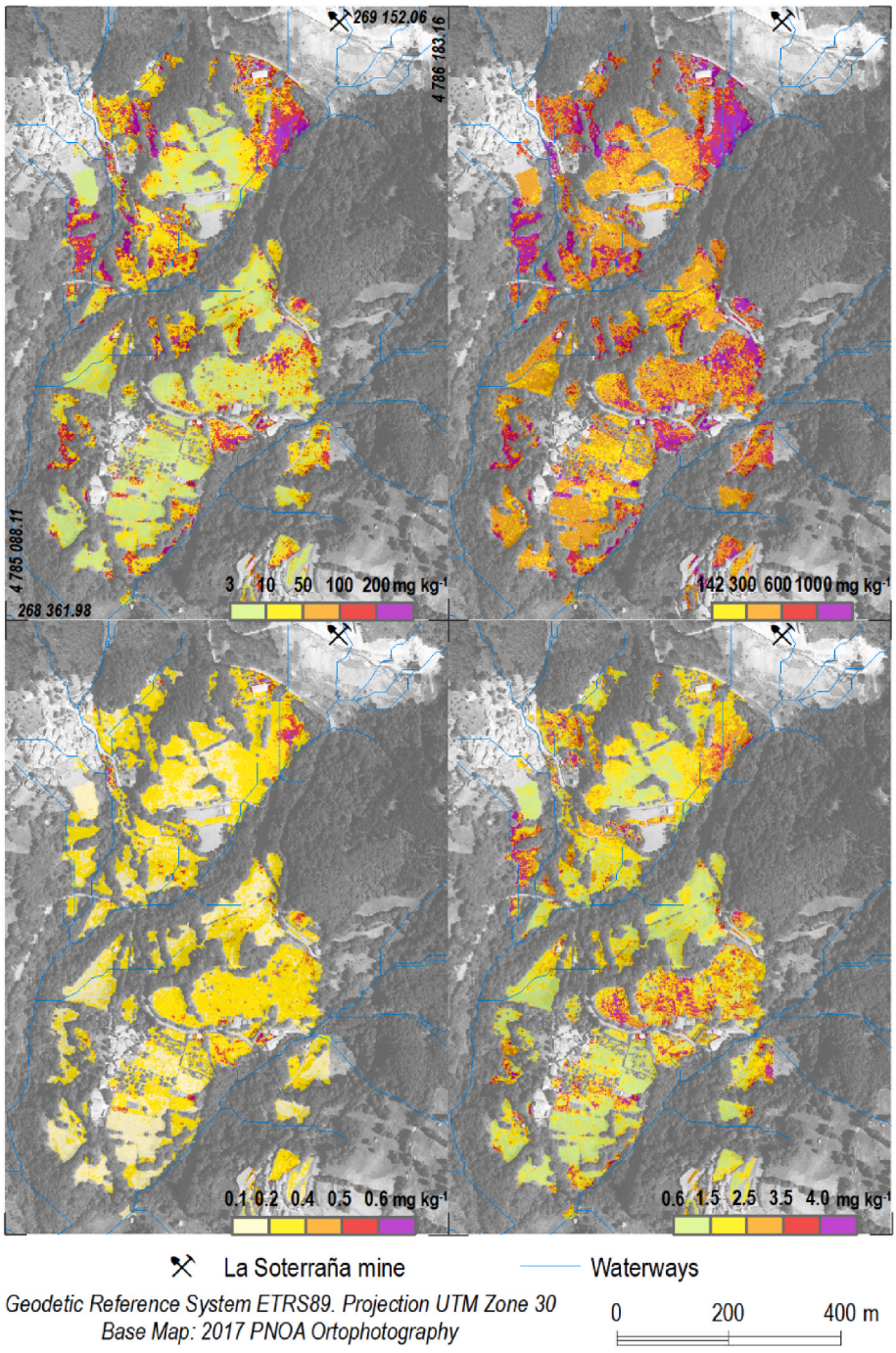


Fig. 6. Spatial distribution of Hg in topsoil (top left), As in topsoil (top right), Hg in vegetation (bottom left), and As in vegetation (bottom right) as obtained after application of the RF algorithm.

lower value than that obtained here, they used an independent validation (only 17 samples as reference) and achieved a limited robustness. Also, in a study in Tai Lake (China), Zhao et al. (2021) obtained R^2 values in soil between 0.218 and 0.283 for As and between 0.02 and 0.03 for Hg using Sentinel-2 with the Back Propagation Neural Network regression technique with an independent validation; it must be pointed out that the concentrations of Hg and As exhibited much lower variability than in the present study. Only Peng et al. (2016) reported similar values to those of this study for As concentration in soil ($R^2 = 0.60$, RPD = 1.53, and RPIQ = 1.48) through RF regression and LANDSAT 8 spectral data after, using 300 sampling points and independent validation. Focusing on vegetation, Dunagan et al. (2007) described a R^2 = 0.670 for Hg with Multiple Stepwise Linear Regression (parametric ML method) using a spectroradiometer, a similar R^2 result to that we obtained through a multispectral UAV but focused only on *Brassica rapa* P. (mustard spinach). In turn, Lv & Liu (2011) did a hyperspectral study in paddy fields and achieved RF values of $R^2 = 0.840$ and RMSE = 3.970 and established a relation between As concentrations in rice plants and visible/near data obtained with a spectrometer.

Finally, from the perspective of the environmental significance of this research for the specific site addressed, the spatial distributions revealed greater geochemical mobility of As than Hg. This behaviour is linked to the geochemical characteristics of these two PTEs (Gallego et al., 2015), with a notable nugget effect of Hg distribution (Boente

et al., 2022), and with a more continuous distribution of As throughout the study area (Matanzas et al., 2021). Regarding the absolute values estimated in the model, the entire study area presents a potential risk for human health given that all the estimated soil concentrations were above 1 mg kg^{-1} of Hg and 40 mg kg^{-1} of As, thus exceeding the Soil Screening Levels in force for soils used for agricultural or livestock activities (Boente et al., 2022). In the case of the spatial distribution of Hg and As in vegetation, according to Markert (1992), the limit for toxicological risks is 0.10 mg kg^{-1}, which was also exceeded for both PTEs throughout the study area. In this regard, the precision obtained in this study strongly suggests that the selected approach (limited geochemical survey, multispectral acquisition, and RF) could be applied to similar cases where contamination by As, Hg, or both is very high. In other cases (different PTEs or more moderate levels of contamination), an adddditional prospection of ML methods would probably be required.

5. Conclusions

The use of UAV-RS technologies to address soil pollution issues has increased in recent years. These technologies are complementary, reliable, and rapid alternatives to classical geochemical methods, and their combination with ML approaches may facilitate the modelling of contaminant distribution and concentrations in both diffuse and point source contamination scenarios. In this context, a novel approach based on the combination of a geochemical study with a limited number of topsoil and vegetation samples, and multispectral data obtained by UAV-RS was presented, and the most appropriate ML technique for interpreting results was identified.

A paradigmatic site affected by Hg and As pollution was selected to exemplify the proposed methodology, and a comprehensive selection of spectral indices that are useful to evaluate soil-plant system health was accomplished. Moreover, four ML techniques (one parametric and three non-parametric) were checked in order to obtain robust statistical models. The performance of the regression methods was compared by using four different statistics that revealed that the RF model achieved the most accurate results. This method enabled the prediction and mapping of contaminant contents in the entire study area, with acceptable adjustment to the geochemical data initially obtained. In general terms, results were slightly more precise for As than Hg, and for vegetation than topsoil. Also, a better accuracy than that obtained in similar studies using satellite or even spectrometry data was achieved.

In conclusion, here it was demonstrated that Hg and As concentrations in soil-plant system can be modelled using a low-density geochemical survey and high-resolution spectral data from UAV-RS, thus obtaining continuous spatial information with statistical robustness. This study also highlights the need for a correct choice of model evaluation statistics in order to minimize the influence of the heterogeneity of the dependent variables. The approach described herein is a powerful tool to study PTEs pollution and arises as an alternative to classical geochemical surveys, which are laborious and expensive.

Author statement

Salgado: Conceptualization, Methodology, Software, Formal Analysis, Investigation, Data Curation, Visualization, Writing – Original Draft. López-Sánchez: Conceptualization, Methodology, Software, Formal Analysis, Data Curation, Writing - Review & Editing. Colina: Formal Analysis, Investigation, Visualization. Baragaño: Investigation, Writing - Review & Editing. Forján: Investigation, Writing - Review & Editing. Gallego: Conceptualization, Writing - Review & Editing, Supervision, Funding Acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This research was partially funded by the project NANOCAREM (AEI/Spain, FEDER/EU, MCI-20-PID2019-106939 GB-I00). Lorena Salgado obtained a grant, "Ayudas para la realización de tesis doctorales. Modalidad A (PAPI-21-PF-27)", funded by the University of Oviedo and Banco Santander. MU-21-UP2021-030 32892642

The authors thank Jesús Valderrábano (INDUROT, University of Oviedo) and Carlos Boente (CIQSO, University of Huelva) for technical support, and also anonymous reviewers who helped to improve the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envpol.2023.122066.

References

Allevato, E., Stazi, S.R., Marabottini, R., D'Annibale, A., 2019. Mechanisms of arsenic assimilation by plants and countermeasures to attenuate its accumulation in crops other than rice. In: Ecotoxicology and Environmental Safety, vol. 185. Academic Press, 109701.

Alvyar, Z., Shahbazi, F., Oustan, S., Dengiz, O., Minasny, B., 2022. Digital mapping of potentially toxic elements enrichment in soils of Urmia Lake due to water level decline. Sci. Total Environ. 808, 152086.

Antonious, G.F., Kochhar, T.S., 2009. Mobility of heavy metals from soil into hot pepper fruits: a field study. Bull. Environ. Contam. Toxicol. 82 (1), 59–63.

Asner, G.P., 1998. Biophysical and biochemical sources of variability in canopy reflectance. Remote Sens. Environ. 64 (3), 234–253.

Ballabio, C., Jiskra, M., Osterwalder, S., Borrelli, P., Montanarella, L., Panagos, P., 2021. A spatial assessment of mercury content in the European Union topsoil. Sci. Total Environ. 769, 144755.

Baragaño, D., Forján, R., Álvarez, N., Gallego, J.R., González, A., 2022. Zero valent iron nanoparticles and organic fertilizer assisted phytoremediation in a mining soil: arsenic and mercury accumulation and effects on the antioxidative system of *medicago sativa* L. J. Hazard Mater. 433 https://doi.org/10.1016/j.jhazmat.2022.128748.

Bausch, W.C., Duke, H.R., 1996. Remote sensing of plant nitrogen status in corn. Trans. ASAE 39 (5), 1869–1875.

Belgiu, M., Drăgu, L., 2016. Random forest in remote sensing: a review of applications and future directions. ISPRS J. Photogrammetry Remote Sens. 114, 24–31.

Bellon-Maurel, V., Fernandez-Ahumada, E., Palagos, B., Roger, J., McBratney, A., 2010. Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. TrAC, Trends Anal. Chem. 29, 1073–1081.

Blackburn, G.A., 1998. Spectral indices for estimating photosynthetic pigment concentrations: a test using senescent tree leaves. Int. J. Rem. Sens. 19 (4), 657–675.

Boente, Carlos, Salgado, L., Romero-Macías, E., Colina, A., López-Sánchez, C.A., Gallego, J.R., 2020. Correlation between geochemical and multispectral patterns in an area severely contaminated by former Hg-as mining. ISPRS Int. J. Geo-Inf. 9 (12), 739.

Boente, C., Baragaño, D., García-González, N., Forján, R., Colina, A., Gallego, J.R., 2022. A holistic methodology to study geochemical and geomorphological control of the distribution of potentially toxic elements in soil. Catena 208, 105730.

Booysen, R., Jackisch, R., Lorenz, S., Zimmermann, R., Kirsch, M., Nex, P.A.M., Gloaguen, R., 2020. Detection of REEs with lightweight UAV-based hyperspectral imaging. Sci. Rep. 10 (1), 1–12.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.

Busuioc, G., Cristina Elekes, C., Stihi, C., Iordache, S., Constantin Ciulei, S., 2011. The bioaccumulation and translocation of Fe, Zn, and Cu in species of mushrooms from Russula genus. Environ. Sci. Pollut. Res. 18, 890–896.

Casa, R., Castaldi, F., Pascucci, S., Basso, B., Pignatti, S., 2013. Geophysical and hyperspectral data fusion techniques for in-field estimation of soil properties. Vadose Zone J. 12 (4) https://doi.org/10.2136/vzj2012.0201.

Chabrillat, S., Ben-Dor, E., Cierniewski, J., Gomez, C., Schmid, T., van Wesemael, B., 2019. Imaging spectroscopy for soil mapping and monitoring. Surv. Geophys. 40, 361–399.

Chakraborty, S., Weindorf, D.C., Li, B., Ali, M.N., Majumdar, K., Ray, D.P., 2014. Analysis of petroleum contaminated soils by spectral modeling and pure response profile recovery of n-hexane. Environ. Pollut. 190, 10–18.

Chakraborty, S., Weindorf, D.C., Li, B., Ali Aldabaa, A.A., Ghosh, R.K., Paul, S., Nasim Ali, M., 2015. Development of a hybrid proximal sensing method for rapid identification of petroleum contaminated soils. Sci. Total Environ. 514, 399–408.
Chakraborty, S., Weindorf, D.C., Deb, S., Li, B., Paul, S., Choudhury, A., Ray, D.P., 2017. Rapid assessment of regional soil arsenic pollution risk via diffuse reflectance spectroscopy. Geoderma 289, 72–81.
Cho, M.A., Debbab, P., Mutanga, O., Dudeni-Tlhoneb, N., Magadla, T., Khuluseb, S.A., 2012. Potential utility of the spectral red-edge region of SumbandilaSat imagery for assessing indigenous forest structure and health. Int. J. Appl. Earth Obs. Geoinf. 16 (1), 85–93.
Choe, E., van der Meer, F., van Ruitenbeek, F., van der Werff, H., de Smeth, B., Kim, K. W., 2008. Mapping of heavy metal pollution in stream sediments using combined geochemistry, field spectroscopy, and hyperspectral remote sensing: a case study of the Rodalquilar mining area, SE Spain. Remote Sens. Environ. 112 (7), 3222–3233.
Chung, K.L., Cheng, J.S., Yang, H. Bin, 2020. Effective chroma subsampling and luma modification for RGB full-color images using the multiple linear regression technique. IEEE Access 8, 118315–118323.
Chuvieco, E., Huete, A., 2009. Fundamentals of satellite remote sensing. In: Fundamentals of Satellite Remote Sensing. CRC Press.
Clevers, J.G.P.W., Kooistra, L., Salas, E.A.L., 2010. Study of heavy metal contamination in river floodplains using the red-edge position in spectroscopic data. Int. J. Rem. Sens. 25 (19), 3883–3895.
Dadová, J., Andráš, P., Kupka, J., Krnáč, J., Andráš, P., Hroncová, E., Midula, P., 2016. Mercury contamination from historical mining territory at Malachov Hg-deposit (Central Slovakia). Environ. Sci. Pollut. Control Ser. 23 (3), 2914–2927.
Dash, J., Curran, P.J., 2004. The MERIS terrestrial chlorophyll index. Int. J. Rem. Sens. 25 (23), 5403–5413.
Daughtry, C.S.T., Walthall, C.L., Kim, M.S., De Colstoun, E.B., McMurtrey, J.E., 2000. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. Remote Sens. Environ. 74 (2), 229–239.
Demšar, J., 2006. Statistical comparisons of classifiers over multiple data sets. J. Mach. Learn. Res. 7, 1–30.
Dunagan, S.C., Gilmore, M.S., Varekamp, J.C., 2007. Effects of mercury on visible/near-infrared reflectance spectra of mustard spinach plants (*Brassica rapa P.*). Environ. Pollut. 148 (1), 301–311.
D'Aniello, A., Hartog, N., Sweijen, T., Pianese, D., 2018. Infiltration and distribution of elemental mercury DNAPL in water-saturated porous media: experimental and numerical investigation. Water Air Soil Pollut. 229 (1), 1–17.
Emadi, M., Taghizadeh-Mehrjardi, R., Cherati, A., Danesh, M., Mosavi, A., Scholten, T., 2020. Predicting and mapping of soil organic carbon using machine learning algorithms in Northern Iran. Rem. Sens. 12 (14), 2234.
Escadafal, Richard. Caractérisation de la surface des sols arides par observations de terrain et par télédétection. Applications: Exemple de la Région de Tataouine (Tunisie). PhD thesis.Université Pierre et Marie Curie. https://core.ac.uk/download/pdf/39865262.pdf.
Escadafal, R., Belghith, A., Ben Moussa, H., 1994. Indices spectraux pou la télédetection de la dégradation des milieux naturels en Tunisie aride. In: Communication Présentée Au Sixième SymposiumInternational "Mesure Physique et Signatures En Télédetection". Val d'Isère, France, pp. 17–21.
Fernández, B., Lara, L.M., Menéndez-Aguado, J.M., Ayala, J., García-González, N., Salgado, L., Colina, A., Gallego, J.L.R., 2020. A multi-faceted, environmental forensic characterization of a paradigmatic brownfield polluted by hazardous waste containing Hg, As, PAHs and dioxins. Sci. Total Environ. 726, 138546.
Filkov, A., Cirulis, B., Penman, T., 2021. Quantifying merging fire behaviour phenomena using unmanned aerial vehicle technology. Int. J. Wildland Fire 30 (3), 197.
Forján, R., Rodríguez-Vila, A., Pedrol, N., Covelo, E.F., 2017. Application of compost and biochar with Brassica juncea L. To reduce phytoavailable concentrations in a settling pond mine soil. Waste and Biomass Valorization 9 (5), 821–834.
Forján, R., Rodríguez-Vila, A., Cerqueira, B., Covelo, E.F., 2018. Effects of compost and technosol amendments on metal concentrations in a mine soil planted with Brassica juncea L. Environ. Sci. Pollut. Control Ser. 25, 19713–19727.
Friedman, J.H., 1991. Multivariate adaptive regression Splines. Ann. Stat. 19 (1), 1–67.
Friedman, J.H., 2002. Stochastic gradient boosting. Comput. Stat. Data Anal. 38, 367–378.
Gallego, J.R., Esquinas, N., Rodríguez-Valdés, E., Menéndez-Aguado, J.M., Sierra, C., 2015. Comprehensive waste characterization and organic pollution co-occurrence in a Hg and as mining and metallurgy brownfield. J. Hazard Mater. 300, 561–571.
Gao, B., Stein, A., Wang, J., 2022. A two-point machine learning method for the spatial prediction of soil pollution. Int. J. Appl. Earth Obs. Geoinf. 108, 102742.
García-Gutiérrez, J., Martínez-Álvarez, F., Troncoso, A., Riquelme, J.C., 2015. A comparison of machine learning regression techniques for LiDAR-derived estimation of forest variables. Neurocomputing 167, 24–31.
Garg, P., Nasimi, R., Ozdagli, A., Zhang, S., Mascarenas, D.D.L., Reda Taha, M., Moreu, F., 2020. Measuring transverse displacements using unmanned aerial systems laser Doppler vibrometer (UAS-LDV): development and field validation. Sensors 20 (21), 6051.
Ghazali, N.A., Ramli, N.A., Yahaya, A.S., Yusof, N.F.F.M., Sansuddin, N., Al Madhoun, W. A., 2010. Transformation of nitrogen dioxide into ozone and prediction of ozone concentrations using multiple linear regression techniques. Environ. Monit. Assess. 165 (1–4), 475–489.
Gholizadeh, A., Kopačková, V., 2019. Detecting vegetation stress as a soil contamination proxy_ a review of optical proximal and remote sensing techniques. In: International Journal of Environmental Science and Technology, vol. 16. Center for Environmental and Energy Research and Studies, pp. 2511–2524. Issue 3.
Gholizadeh, Asa, Saberioon, M., Ben-Dor, E., Borůvka, L., 2018. Monitoring of selected soil contaminants using proximal and remote sensing techniques: background, state-of-the-art and future perspectives. Crit. Rev. Environ. Sci. Technol. 48 (3), 243–278.
Gholizadeh, Asa, Saberioon, M., Ben-Dor, E., Viscarra Rossel, R.A., Borůvka, L., 2020. Modelling potentially toxic elements in forest soils with vis–NIR spectra and learning algorithms. Environ. Pollut. 267, 115574.
Giannetti, F., Puletti, N., Puliti, S., Travaglini, D., Chirici, G., 2020. Assessment of UAV photogrammetric DTM-independent variables for modelling and mapping forest structural indices in mixed temperate forests. Ecol. Indicat. 117, 106513.
Gil-Díaz, M., Rodríguez-Valdés, E., Alonso, J., Baragaño, D., Gallego, J.R., Lobo, M.C., 2019. Nanoremediation and long-term monitoring of brownfield soil highly polluted with as and Hg. Sci. Total Environ. 675, 165–175.
Gislason, P.O., Benediktsson, J.A., Sveinsson, J.R., 2006. Random forests for land cover classification. Pattern Recogn. Lett. 27 (4), 294–300.
Gitelson, A., Merzlyak, M.N., 1994. Quantitative estimation of chlorophyll-a using reflectance spectra: experiments with autumn chestnut and maple leaves. J. Photochem. Photobiol., B 22 (3), 247–252.
Gitelson, A.A., Merzlyak, M.N., 1996. Signature analysis of leaf reflectance spectra: algorithm development for remote sensing of chlorophyll. J. Plant Physiol. 148 (3–4), 494–500.
Gitelson, A.A., Merzlyak, M.N., Lichtenthaler, H.K., 1996. Detection of red edge position and chlorophyll content by reflectance measurements near 700 nm. J. Plant Physiol. 148 (3–4), 501–508.
Golchoubian, H., Moayyedi, G., Fazilati, H., 2012. Spectroscopic studies on Solvatochromism of mixed-chelate copper(II) complexes using MLR technique. Spectrochim. Acta, Part A 85 (1), 25–30.
Guerra, F., Trevizam, A.R., Muraoka, T., Marcante, N.C., Canniatti-Brazaca, S.G., 2012. Heavy metals in vegetables and potential risk for human health. Sci. Agric. 69 (1), 54–60.
Guerra-Hernández, Cosenza, Cardil, Silva, Botequim, Soares, Silva, González-Ferreiro, Díaz-Varela, 2019. Predicting growing stock volume of Eucalyptus plantations using 3-D point clouds derived from UAV imagery and ALS data. Forests 10 (10), 905.
Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H., 2009. The WEKA data mining software: an update. SIGKDD Explorations 11 (1), 10–18.
Handique, B.K., Goswami, C., Gupta, C., Pandit, S., Gogoi, S., Jadi, R., Jena, P., Borah, G., Raju, P.L.N., 2020. Hierarchical classification for assessment of horticultural crops in mixed cropping pattern using uav-borne multi-spectral sensor. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives 43 (B3), 67–74.
Haralick, R.M., Dinstein, I., Shanmugam, K., 1973. Textural features for image classification. IEEE Transactions on Systems, Man and Cybernetics, SMC 3 (6), 610–621.
Henrique Brant Dias, S., Filgueiras, R.I., Iná cio Fernandes Filho, E., Santos ArcanjoID, G., Henrique da Silva, G., Chartuni Mantovani, E., Franç da CunhaID, F., 2021. Reference evapotranspiration of Brazil modeled with machine learning techniques and remote sensing. PLoS One 16 (2), e0245834.
Higueras, P., Oyarzun, R., Lillo, J., Sánchez-Hernández, J.C., Molina, J.A., Esbrí, J.M., Lorenzo, S., 2006. The Almadén district (Spain): anatomy of one of the world's largest Hg-contaminated sites. Sci. Total Environ. 356 (1–3), 112–124.
Hoffer, R.M., 1978. Biological and physical considerations in applying computer-aided analysis techniques to remote sensor data. In: Remote Sensing: The Quantitative Approach, pp. 227–289.
Horler, D.N.H., Dockray, M., Barber, J., 1983. The red edge of plant leaf reflectance. Int. J. Rem. Sens. 4 (2), 273–288.
Hornik, K., Buchta, C., Zeileis, A., 2009. Open-source machine learning: R meets Weka. Comput. Stat. 24, 225–232.
Hout, R., Maleval, V., Mahe, G., Rouvellac, E., Crouzevialle, R., Cerbelaud, F., 2020. UAV and LiDAR data in the service of bank gully erosion measurement in rambla de Algeciras lakeshore. Water 12 (10), 2748.
Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). Remote Sens. Environ. 25 (3), 295–309.
Hunt, E.R., Doraiswamy, P.C., McMurtrey, J.E., Daughtry, C.S.T., Perry, E.M., Akhmedov, B., 2012. A visible band index for remote sensing leaf chlorophyll content at the Canopy scale. Int. J. Appl. Earth Obs. Geoinf. 21 (1), 103–112.
Ihaka, R., Gentleman, R., 1996. R: a language for data analysis and graphics. J. Comput. Graph Stat. 5 (3), 299.
Immitzer, M., Vuolo, F., Atzberger, C., 2016. First experience with sentinel-2 data for crop and tree species classifications in central europe. Rem. Sens. 8 (3), 166.
Jafari, R., Amiri, M., Asgari, F., Tarkesh, M., 2022. Dust source susceptibility mapping based on remote sensing and machine learning techniques. Ecol. Inf. 72.
Jia, X., Cao, Y., O'Connor, D., Zhu, J., Tsang, D.C.W., Zou, B., Hou, D., 2021a. Mapping soil pollution by using drone image recognition and machine learning at an arsenic-contaminated agricultural field. Environ. Pollut. 270, 116281.
Jia, X., O'Connor, D., Shi, Z., Hou, D., 2021b. VIRS based detection in combination with machine learning for mapping soil pollution. Environ. Pollut. 268, 115845.
Jurado, J.M., Pádua, L., Feito, F.R., Sousa, J.J., 2020. Automatic grapevine trunk detection on UAV-based point cloud. Rem. Sens. 12 (18), 3043.
Khosravi, V., Doulati Ardejani, F., Yousefi, S., Aryafar, A., 2018. Monitoring soil lead and zinc contents via combination of spectroscopy with extreme learning machine and other data mining methods. Geoderma 318, 29–41.
Khosravi, V., Doulati Ardejani, F., Gholizadeh, A., Saberioon, M., 2021. Satellite imagery for monitoring and mapping soil chromium pollution in a mine waste dump. Rem. Sens. 13 (7), 1277, 2021, Vol. 13, *Page 1277*.
Khosravi, V., Gholizadeh, A., Saberioon, M., 2022. Soil toxic elements determination using integration of Sentinel-2 and Landsat-8 images: effect of fusion techniques on model performance. Environ. Pollut. 310, 119828.

Kooistra, L., Wehrens, R., Leuven, R.S.E.W., Buydens, L.M.C., 2001. Possibilities of visible-near-infrared spectroscopy for the assessment of soil contamination in river floodplains. Anal. Chim. Acta 446 (1–2), 97–105.
Kooistra, L., Salas, E.A.L., Clevers, J.G.P.W., Wehrens, R., Leuven, R.S.E.W., Nienhuis, P. H., Buydens, L.M.C., 2004. Exploring field vegetation reflectance as an indicator of soil contamination in river floodplains. Environ. Pollut. 127 (2), 281–290.
Kooistra, L., Leuven, R.S.E.W., Wehrens, R., Nienhuis, P.H., Buydens, L.M.C., 2010. A comparison of methods to relate grass reflectance to soil metal contamination. Int. J. Rem. Sens. 24 (24), 4995–5010.
Kureel, N., Sarup, J., Matin, S., Goswami, S., Kureel, K., 2021. Modelling vegetation health and stress using hypersepctral remote sensing data. Modeling Earth Systems and Environment 8 (1), 733–748.
Leung, Y.F., Liu, W., Li, J.S., Wang, L., Tsang, D.C.W., Lo, C.Y., Leung, M.T., Poon, C.S., 2018. Three-dimensional spatial variability of arsenic-containing soil from geogenic source in Hong Kong: implications on sampling strategies. Sci. Total Environ. 633, 836–847.
Liu, J., Paulsen, M.R., 2000. Corn whiteness measurement and classification using machine vision. Trans. ASAE 43 (3), 757–763.
Lv, J., Liu, X., 2011. Predicting arsenic concentration in rice plants from hyperspectral data using random forests. Adv. Intell. Syst. Comput. 128, 601–606.
Mandal, U.K., 2016. Spectral color indices based geospatial modeling of soil organic matter in Chitwan district, Nepal. Int. Arch. Photogram. Rem. Sens. Spatial Inf. Sci.
Manea, D.N., Ienciu, A.A., Ştef, R., Şmuleac, I.L., Gergen, I.I., Nica, D.V., 2020. Health risk assessment of dietary heavy metals intake from fruits and vegetables grown in selected old mining areas—a case study: the banat area of southern carpathians. Int. J. Environ. Res. Publ. Health 17 (14), 1–19.
Markert, B., 1992. Establishing of "Reference Plant" for inorganic characterization of different plant species by chemical fingerprinting. Water, Air, Soil Pollut. 64 (3–4), 533–538.
Matanzas, N., Afif, E., Díaz, T.E., Luis, J., Gallego, R., Hodge, V., Petruzzelli, G., Quirós, G., 2021. Screening of pioneer metallophyte plant species with phytoremediation potential at a severely contaminated Hg and as mining site. Environments 8 (7), 63.
Matese, A., Toscano, P., Di Gennaro, S., Genesio, L., Vaccari, F., Primicerio, J., Belli, C., Zaldei, A., Bianconi, R., Gioli, B., 2015. Intercomparison of UAV, aircraft and satellite remote sensing platforms for precision viticulture. Rem. Sens. 7 (3), 2971–2990.
McLaughlin, M.J., Parker, D.R., Clarke, J.M., 1999. Metals and micronutrients - food safety issues. Field Crop. Res. 60 (1–2), 143–163.
Mei, L., Zhu, Y., Zhang, X., Zhou, X., Zhong, Z., Li, H., Li, Y., Li, X., Daud, M.K., Chen, J., Zhu, S., 2021. Mercury-induced phytotoxicity and responses in upland cotton (gossypium hirsutum L.) seedlings. Plants 10 (8), 1494, 2021, Vol. 10, Page 1494.
Metternicht, G., 2010. Vegetation indices derived from high-resolution airborne videography for precision crop management. Int. J. Rem. Sens. 24 (14), 2855–2877.
Mouazen, A.M., Nyarko, F., Qaswar, M., Tóth, G., Gobin, A., Moshou, D., 2021. Spatiotemporal prediction and mapping of heavy metals at regional scale using regression methods and landsat 7. Rem. Sens. 13 (22), 4615, 2021, Vol. 13, *Page 4615*.
Novo-Fernández, A., Franks, S., Wehenkel, C., López-Serrano, P.M., Molinier, M., López-Sánchez, C.A., 2018. Landsat time series analysis for temperate forest cover change detection in the Sierra Madre Occidental, Durango, Mexico. Int. J. Appl. Earth Obs. Geoinf. 73, 230–244.
Novo-Fernández, A., Barrio-Anta, M., Recondo, C., Cámara-Obregón, A., López-Sánchez, C.A., 2019. Integration of national forest inventory and nationwide airborne laser scanning data to improve forest yield predictions in north-western Spain. Rem. Sens. 11 (14), 1693.
Peng, Y., Kheir, R.B., Adhikari, K., Malinowski, R., Greve, M.B., Knadel, M., Greve, M.H., 2016. Digital mapping of toxic metals in Qatari soils using remote sensing and ancillary data. Rem. Sens. 8 (12), 1003, 2016, Vol. 8, Page 1003.
Penuelas, J., Frederic, B., Filella, I., 1995. Semi-empirical indices to assess carotenoids/Chlorophyll a ratio from leaf spectral reflectance. Photosynthetica 31 (2), 221–230.
Quang, N.H., Quinn, C.H., Carrie, R., Stringer, L.C., Hue, L. T. Van, Hackney, C.R., Tan, D. Van, 2022. Comparisons of regression and machine learning methods for estimating mangrove above-ground biomass using multiple remote sensing data in the red River Estuaries of Vietnam. Remote Sens. Appl.: Soc. Environ. 26.
Rallo, P., de Castro, A.I., López-Granados, F., Morales-Sillero, A., Torres-Sánchez, J., Jiménez, M.R., Jiménez-Brenes, F.M., Casanova, L., Suárez, M.P., 2020. Exploring UAV-imagery to support genotype selection in olive breeding programs. Sci. Hortic. 273, 109615.
Riveros-Burgos, C., Ortega-Farías, S., Morales-Salinas, L., Fuentes-Peñailillo, F., Tian, F., 2021. Assessment of the clumped model to estimate olive orchard evapotranspiration using meteorological data and UAV-based thermal infrared imagery. Irrigat. Sci. 39 (1), 63–80.
Roces-Díaz, J.V., Jiménez-Alfaro, B., Álvarez-Álvarez, P., Álvarez-García, M.A., 2014. Environmental niche and distribution of six deciduous tree species in the Spanish atlantic region. IForest 8 (2), 214–221.
Rondeaux, G., Steven, M., Baret, F., 1996. Optimization of soil-adjusted vegetation indices. Remote Sens. Environ. 55 (2), 95–107.
Rouse, J., Haas, R.H., Schell, J.A., Deering, D., 1974. Monitoring Vegetation Systems in the Great Plains with ERTS, vol. 1. NASA Special Publication, pp. 309–317.
Sahu, G.K., Upadhyay, S., Sahoo, B.B., 2011. Mercury induced phytotoxicity and oxidative stress in wheat (Triticum aestivum L.) plants. Physiol. Mol. Biol. Plants 18 (1), 21–31, 2011 18:1.
Shi, T., Liu, H., Chen, Y., Wang, J., Wu, G., 2016a. Estimation of arsenic in agricultural soils using hyperspectral vegetation indices of rice. J. Hazard Mater. 308, 243–252.
Shi, T., Wang, J., Chen, Y., Wu, G., 2016b. Improving the prediction of arsenic contents in agricultural soils by combining the reflectance spectroscopy of soils and rice plants. Int. J. Appl. Earth Obs. Geoinf. 52, 95–103.
Siddiqui, M.H., Alamri, S., Nasir Khan, M., Corpas, F.J., Al-Amri, A.A., Alsubaie, Q.D., Ali, H.M., Kalaji, H.M., Ahmad, P., 2020. Melatonin and calcium function synergistically to promote the resilience through ROS metabolism under arsenic-induced stress. J. Hazard Mater. 398, 122882.
Sims, D.A., Gamon, J.A., 2002. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. Remote Sens. Environ. 81 (2–3), 337–354.
Stroppiana, D., Bordogna, G., Carrara, P., Boschetti, M., Boschetti, L., Brivio, P.A., 2012. A method for extracting burned areas from Landsat TM/ETM+ images by soft aggregation of multiple Spectral Indices and a region growing algorithm. ISPRS J. Photogrammetry Remote Sens. 69, 88–102.
Tan, K., Ma, W., Wu, F., Du, Q., 2019. Random forest–based estimation of heavy metal concentration in agricultural soils with hyperspectral sensor data. Environ. Monit. Assess. 191 (7), 1–14.
Tavera Busso, I., Rodríguez Núñez, M., Amarillo, A.C., Mettan, F., Carreras, H.A., 2021. Modeling air pollution-related hospital admissions employing remote sensing and geographical information systems. Atmos. Environ. vol. 261.
Tejado-Ramos, J.-J., Chocarro-León, M., Barrero-Béjar, I., Valverde-Calvo, A., Ferreras-Moreno, M., Giraldo-Pavón, F., Tarragona-Pérez, C., 2021. Enhancement of the sustainability of wolfram mining using drone remote sensing technology. Remote Sens. Appl.: Society and Environment 23, 100542.
Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens. Environ. 8 (2), 127–150.
van Tiggelen, M., Smeets, P., Reijmer, C., Wouters, B., Steiner, J., Nieuwstraten, E., Immerzeel, W., van den Broeke, M., 2021. Mapping the aerodynamic roughness of the Greenland ice sheet surface using ICESat-2: evaluation over the K-transect. Cryosphere Discuss. 1–28.
Vaudour, E., Gholizadeh, A., Castaldi, F., Saberioon, M., Borůvka, L., Urbina-Salazar, D., Van Wesemael, B., 2022. Satellite imagery to map topsoil organic carbon content over cultivated areas: an overview. Rem. Sens. 14 (12) https://doi.org/10.3390/rs14122917.
Verdebout, J., Jacquemoud, S., Schmuck, G., 1994. Optical properties of leaves: modelling and experimental studies. Imaging Spectrometry — a Tool for Environmental Observations 169–191.
Viera-Torres, M., Sinde-González, I., Gil-Docampo, M., Bravo-Yandún, V., Toulkeridis, T., 2020. Generating the baseline in the early detection of bud rot and red ring disease in oil palms by geospatial technologies. Rem. Sens. 12 (19), 3229.
Wang, Y., Zhang, X., Sun, W., Wang, J., Ding, S., Liu, S., 2022. Effects of hyperspectral data with different spectral resolutions on the estimation of soil heavy metal content: From ground-based and airborne data to satellite-simulated data. Science of The Total Environment 838, 156129.
Weber, C., von Eichel-Streiber, J., Rodrigo-Comino, J., Altenburg, J., Udelhoven, T., 2020. Automotive radar in a UAV to assess earth surface processes and land responses. Sensors 20 (16), 4463.
Wei, L., Yuan, Z., Yu, M., Huang, C., Cao, L., 2019. Estimation of arsenic content in soil based on laboratory and field reflectance spectroscopy. Sensors 19 (18), 3904.
Xu, Z., Wang, X., 2010. Research for Information Extraction Based on Wrapper Model Algorithm, vol. 2010. 2nd International Conference on Computer Research and Development, ICCRD, pp. 652–655.
Zagajewski, B., Tømmervik, H., Bjerke, J.W., Raczko, E., Bochenek, Z., Kłos, A., Jarocińska, A., Lavender, S., Ziółkowski, D., 2017. Intraspecific differences in spectral reflectance curves as indicators of reduced vitality in high-arctic plants. Rem. Sens. 9, 1289.
Zhang, Chengye, Ren, H., Qin, Q., Ersoy, O.K., 2017. A new narrow band vegetation index for characterizing the degree of vegetation stress due to copper: the copper stress vegetation index (CSVI). Remote Sensing Letters 8 (6), 576–585.
Zhang, Jinhao, Xiao, M., Gao, L., Fu, J., 2018. A novel projection outline based active learning method and its combination with Kriging metamodel for hybrid reliability analysis with random and interval variables. Comput. Methods Appl. Mech. Eng. 341, 32–52.
Zhang, Chao, Yang, K., Li, Y., Cheng, F., Rong, K., 2019. Spectral characteristics and the study of pollution degree of maize leaves under copper and lead stress. J. Indian Soc. Remote Sens. 48 (1), 21–33, 2019 48:1.
Zhang, Jie, Hamza, A., Xie, Z., Hussain, S., Brestic, M., Tahir, M.A., Ulhassan, Z., Yu, M., Allakhverdiev, S.I., Shabala, S., 2021. Arsenic transport and interaction with plant metabolism: clues for improving agricultural productivity and food safety. Environ. Pollut. 290, 117987.
Zhao, H., Liu, P., Qiao, B., Wu, K., 2021. The spatial distribution and prediction of soil heavy metals based on measured samples and multi-spectral images in Tai Lake of China. Land 10 (11), 1227, 2021, Vol. 10, *Page 1227*.
Zhao, F.J., Tang, Z., Song, J.J., Huang, X.Y., Wang, P., 2022. Toxic metals and metalloids: uptake, transport, detoxification, phytoremediation, and crop improvement for safer food. Mol. Plant 15 (1), 27–44.
Zhiwei, X., Xinghua, W., 2010. Research for Information Extraction Based on Wrapper Model Algorithm. Second International Conference on Computer Research and Development, Kuala Lumpur, Malaysia, pp. 652–655.
Zinnert, J.C., Via, S.M., Young, D.R., 2013. Distinguishing natural from anthropogenic stress in plants: physiology, fluorescence and hyperspectral reflectance. Plant Soil 366 (1–2), 133–141.
Biney, J.K.M., Houška, J., Volánek, J., Abebrese, D.K., Cervenka, J., 2023. Examining the influence of bare soil UAV imagery combined with auxiliary datasets to estimate and map soil organic carbon distribution in an erosion-prone agricultural field. Sci Total Environ. 870,161973.