Forensic Science International 328 (2021) 111035

Contents lists available at ScienceDirect

Forensic Science International

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

A probabilistic approach towards source level inquiries for forensic soil examination based on mineral counts

Yu Chen Lim*, André Marolf, Nicolas Estoppey, Geneviève Massonnet

University of Lausanne, Ecole des sciences criminelles, Batochime, 1015 Lausanne, Switzerland

ARTICLE INFO

Article history:
Received 30 March 2021
Accepted 27 September 2021
Available online 5 October 2021

Keywords:
Forensic geology
Bayesian approach
Heavy mineral fraction
Light mineral fraction
Polarized light microscopy
Scanning electron microscopy
Multivariate statistics
Inter-variability
Intra-variability
Blind tests

ABSTRACT

Forensic soil examination has a well-established foundation in forensic science, this is in part due to the widely varied and complex nature of soil. Within this domain, mineral suite studies are a commonly utilized tool in soil examination. However, statistical or probabilistic approaches towards the interpretation of results from such analysis are lacking and this study aims to fill that gap. Soil samples from four different locations in the city of Lausanne, Switzerland were sampled and their mineral fractions, light and heavy of size between 90 and 180 µm, were studied utilizing microscopical methods. First, the light minerals were identified and counted by employing scanning electron microscopy coupled with energy-dispersive X-ray spectroscopy (SEM-EDS). Second, the heavy minerals were identified and counted manually under a polarized light microscope (PLM). The resulting count data were subjected to various multivariate statistical treatments such as principal components analysis (PCA), hierarchical clustering analysis (HCA), and linear discriminant analysis (LDA). These methods assist in identifying pertinent variables and subsequently in building various classification models. The validities of these models were then tested and evaluated using blind tests. Finally, these methods demonstrate how a probabilistic approach can be taken in the interpretation of the results to answer source level questions.

1. Introduction

The complex nature of soil, combined with its ease of transfer, compels many to consider it to be an ideal type of trace evidence in the domain of forensic science [1–3]. Soil is comprised of several different components (such as minerals, seeds, pollen, anthropogenic particles, etc.) with characteristics related to the source environment as well as parent material. This brings about an exploitable diversity where different components can be studied in a forensic context [1,4]. There are two main utilizations of forensic soil analysis: geolocation and comparison cases. The latter will be the focus of this study and usually involves the analyses of a trace sample of unknown origin, followed by the comparison to various bulk samples of known sources. The goal would be to inform on the potential sources of the trace sample from a closed set of possibilities [3,5,6]. While the analysis of soils should include the employment of multiple independent techniques in order to fully exploit discriminatory power [4], only the analysis of the mineral fraction of a specific size was carried out. This is because the purpose of the study was to explore how such results can be considered in answering source level questions in an informative manner rather than explore the depth of discrimination that can be obtained between these samples. Such interpretations can be later integrated into the wider array of analysis.

The mineral suites make up an important component for comparison due to their variation with the source environment and relative stability [7,8]. Such minerals are well integrated with the other components of soil and its analysis necessitates substantial sample preparation, as described by Palenik [9], to isolate specific fractions. Mineral suites of the sand fraction, specifically between 90 and 180 µm have, according to Palenik [9], a high transfer probability and thus provide a good representation for comparison. In addition, heavy minerals are well concentrated in this size range and yields a representative view of the mineral suites [8].

Minerals of this fraction can be divided into two groups based on their density [8,9]. Light minerals with densities less than 2.89 mg/cm^3 are composed of mainly quartz, carbonates, feldspars, and micas [9]. The distinct classification of phases coupled with the simple varieties found in the light mineral fraction allows for easy identification via elemental analysis such as SEM-EDS [10]. Utilizing such methods, light mineral fraction analysis can be easily automated to

---

* Corresponding author.
E-mail addresses: yuchen.lim@unil.ch (Y.C. Lim), andre.marolf@unil.ch (A. Marolf), nicolas.estoppey@unil.ch (N. Estoppey), genevieve.massonnet@unil.ch (G. Massonnet).

https://doi.org/10.1016/j.forsciint.2021.111035
0379-0738/

provide a proportional count of each type of mineral present. Such an approach mitigates the tediousness of identifying and counting these minerals manually [10,11]. The main drawbacks, however, are the lack of varietal features in this fraction and the fact that these minerals make up the most abundant minerals in the earth's crust [12]. This disadvantage is especially pertinent if one focuses only on chemical compositions and not on physical properties such as morphology [6]. As such, discrimination based on these components may not be immediately apparent, therefore requiring the use of statistical treatments to accentuate their significance.

Heavy minerals with densities greater than 2.89 mg/cm^3 on the other hand possess much more variety and thus present a wider range of compositions [8]. Since minerals possess highly ordered chemical and physical structures, their optical properties are highly characteristic. These properties have been well documented and allow for their identification and counting under polarized light microscopy [8,13,14]. With the correct training and tools, heavy minerals may be identified or classified based on their morphology, color, refractive index, and optical properties under polarized light. This provides a means of obtaining the mineral count of a sample. However, heavy minerals are a small fraction of the soil sample, rarely making up more than 1%. Therefore, in order to study them in appreciable quantities it is necessary to concentrate them since bulk analysis of soil rarely reflects any information originating from this fraction [8,9].

While there exist many methods for the analysis of soils, especially with the advancement in instrumental techniques in recent times [4,15], microscopy, optical or otherwise, remains a preferred tool in the arsenal of many forensic soil examiners. Such methods provide the benefit of allowing the user to physically see and potentially identify the various materials present in a sample. As a result, ambiguity is reduced since individual particles can be characterized and isolated for further analysis [5,16,17]. Furthermore, these methods permit the study of physical characteristics such as morphology and texture, which would otherwise be missed by straight chemical analysis [17].

There are, however, considerations that should be made when using such results to aid in source level inquiries. In general, comparison of soil samples using any method typically leads to one of three conclusions: differentiated, non-differentiated, or inconclusive [5,6,18]. Once this decision has been made, the interpretation of the results often stops there, with no further insight into the significance of this conclusion. It is commonly accepted that the complexity of soil gives a quality of uniqueness to soils from different sources. The claim then follows that it is highly improbable that two soils share the exact same composition, and they will therefore be differentiated in some manner [15]. While this might be the case, natural variability within the sample renders it difficult to draw a definitive line that discriminates two samples, especially if they are closely related in terms of environment and source material. More clearly stated, even though two samples may be similar or dissimilar, it is insufficient to simply draw conclusions based on the ability to differentiate or not between them. Instead, it is important to identify which variables are significant in their discrimination and to what extent. In doing so, the pertinence and strength of the results can be more accurately evaluated. This is related to the intra- and inter-variability of the samples [19]; the variation within the relevant population as compared to the variation between different populations. Understanding this is key to providing useful information towards the answering of the question of common source.

With respect to the comparison of mineral suite compositions, there is a tendency to focus on certain key minerals as a means of discrimination. That is to say that they are present, preferably in abundance, in one sample, but either completely absent or present in an appreciably different quantity in the other sample [18]. While this semi-quantitative approach may be sound, it is limited to samples that are clearly differentiable in terms of composition and offers no means of measuring the significance of a non-differentiation decision. Without a relevant database, there is no way of knowing just how unique such a composition may be and samples with similar composition may be falsely linked due to such a subjective decision. Furthermore, if a key "rare" mineral presents itself in only small concentrations, there would be a high relative standard deviation [20,21] which may result in it being undetected in certain samples when it is in fact present, leading to a false exclusion. There is thus a need, as highlighted by some publications [6,18,22], to take a probabilistic approach towards the interpretation of "non-differentiations" in casework. Such notions are also seconded by the best practices manual published by European Network of Forensic Science Institute's animal, plant and soil working group [23]. In order to carry out such an approach, the use of a case relevant localized database, as well as parametric or non-parametric multivariate statistics to model the inter and intra-variability of the data is required. With such information, evaluations at the source level can be considered under the proposition that the samples share a common source and the proposition that they do not [20]. Hence, the analyst would be able to provide a much more complete interpretation than just simply the ambiguous statement that the two samples "cannot be differentiated", where the significance of this statement is left to the decision maker to interpret.

Such approaches with statistical tools are typical in other aspects of forensic science [19,24] as well as with certain instrumental analysis of soils [6]. Their use has also been explored in the closely related domains of geology and botany in order to classify samples by location based on count data [25]. However, they remain rare when comparing mineral suite compositions for forensic purposes with only a handful of studies found [22,26]. With such methods, the goal is to create a statistical model that would group various samples into established classes, which in this case refers to the source location.

Based on analysis of the mineral suites of the sand fraction, this study explores how the results can be considered in answering source level questions in an informative manner by using statistical tools. Two phases were executed in succession, the first being a mostly exploratory phase to understand the data, find patterns of interest, and identify, if any, key potentially discriminating variables. Secondly, with the information obtained during the first phase, a more directed approach was then taken in building a predictive model that would be able to class an unknown sample into the various categories with known or at least appreciable uncertainty. Not only can such a model be evaluated for its precision and accuracy, but it also furnishes its classifications with transparency. This provides a more complete picture for the interpretation of the data.

Soils from four different parks in Lausanne, Switzerland were collected and analyzed for this study. These samples have been known to be difficult to differentiate with other methods such as pH, elementary composition or particle size distribution, etc. [27]. These samples share a similar composition given the similar geological and biological setting of the area. Lausanne lies on the Molasses basin along with the other major cities and cantons in Switzerland, soils originating from the basin thus share many similarities due to the parallel conditions of weather, elevation, parent materials, etc. [28]. As such it was expected that the raw mineral counts would easily produce a subjective "non-differentiated" result, despite being from different locations. Multivariate statistics were applied to better understand and clarify such a result, with the hope that a more informative interpretation could be obtained for answering source level questions.

2. Materials and method

2.1. Sampling and sample preparation

Samples in this study were obtained from four different parks in the Lausanne area: Valency (VA), Mont-Repos (RE), Denantou (DE), and Bourget (BO) as shown in the to-scale map in Fig. 1. These locations are a mix of urban and natural settings, although the soils are most likely of natural origin and have not been imported. The exact sampling points within these parks were chosen based on their seclusion and accessibility from a main track. Respectively in these locations, samples of 20 g were collected from the surface from a central point (CC), followed by two points found 2 m to the north (NN) and south (SS). Soils from each sampling point were collected with inert tools, packaged in plastic seal bags, and dried at ambient temperature.

To prepare the respective samples for analysis, 3–4 g of subsample was first obtained by coning and quartering to ensure homogeneity. The process described by Palenik [9] was then carried out to separate both the light and heavy mineral fraction of size between 90 and 180 µm. Three replicates of all the center points were obtained for analysis with two additional replicates of the North and South locations respectively for the heavy mineral fraction. Each sample was labelled by their location, the cardinal direction, and the replicate number. For example: VACC1 refers to the first replicate obtained from the central point of Valency park.

2.2. Light mineral fraction

A subsample of the light mineral fraction was obtained once again by quartering. This subsample was then mounted in a ClaroFast hot mounting acrylic resin using a Struers Cito-Press mounting press. Parameters of the mounting process can be found in the Annex. These resin mounts were then polished using silicon carbide and subsequently diamond polishers of sizes, 6, 3 and 1 µm progressively. These mounts were then carbon coated with a Cressington Carbon Coater at 4.0 V which gives a coat of around 15 nm. They were then observed under a Tescan Mira II LMU SEM with a Penta-FET EDS detector. The AZtec software was utilized to carry out the automatic counts. The exact parameters for imaging and quantifying X-ray counts can be found in the Annex. For each mount, the image of an area was first taken to establish the thresholds of contrast in order for the software to identify the minerals as target areas to analyze. An erosion function was utilized to minimize issues due to adjacent touching particles being counted as one particle, and a hole-fill function was added to tackle the issues of particles with gaps. Once these parameters were optimized, an automation process was set in place to count and analyze these mineral particles across the mount. These particles were then classified based on their elemental composition by percent weight according to the scheme found in the Annex Table 1. Where possible, up to approximately 1000 particles were counted per preparation. 12 samples consisting of 3 replicates of each center point were analyzed using this method.

2.3. Heavy mineral fraction

The heavy mineral fractions were respectively mounted on microscope slides using Cargille immersion oil with refractive index of $n_D = 1.660$. The resulting yield of minerals was sufficient to produce 1–3 microscopic slides for each replicate with approximately three to five hundred particles on each slide. All slides were utilized in counting to ensure the results were representative. The slides were then observed utilizing a Leica Leitz DMRXP polarized light microscope. Minerals were then classified based on their optical characteristics and morphology under 100 x or 200 x magnification, with references to the literature for consistency [8]. The method described by Graves [14] was adapted; the rotating stage of the microscope was fixed and the slide was allowed to traverse horizontally and minerals that came into the field of view that intersected the ocular crosshairs were identified and counted. Once the end of a row was reached, the slide was then displaced vertically by a sufficient distance to avoid double counting. This process ensures consistent counting and up to around 300 particles were counted for each replicate. In total, 28 samples were observed, comprising 3 replicates for the center point (CC) and two replicates for the North (NN) and South (SS) locations.

2.4. Data treatment

The numerical data obtained for each sample was first normalized to a hundred percent based on total mineral count in order to compare fairly across the samples. R statistical software was then used in all the data treatment. First, to explore and visualize the data, each mineral type was plotted against their proportion and separated by location to evaluate their discriminating potential. The light mineral fraction was eliminated as an effective means of discrimination on this basis. Variables, if any, were then preliminarily selected for evaluation based on their perceived initial potential and confidence in identification. Multinomial logistic regression on each of these selected variables was carried out to evaluate and select suitably discriminating minerals based on their p-values, using VA as a reference location. These p-values were calculated from the Z-score for a normal distribution under the null hypothesis that the counts of each variable are not different for distinct locations. It thus refers to the probability of observing the results given that the null hypothesis is true [21].

Three different pretreatment methods were applied: total proportion normalization, square root, and a combination of both. These were evaluated using the same visual and regression methods, which resulted in selecting both pretreatment methods. Using multinomial logistic regression as a guide, the proportions of minerals were compared between locations to evaluate if their proportions were different. The minerals with p-values less than 0.1 for at least two of the locations, as well as dolomite were selected for the next phase.

Pretreated values of selected variables were then used to obtain their principal component scores utilizing their covariance. After which, the selection of variables was further refined based on their loadings, clustering ability, and percent variation explained. The refined list of variables was then applied to HCA methods where four distance measures were utilized: Euclidean, Manhattan, Minkowski, and Bray-Curtis [24,25,29]. For each distance measure, four clustering methods: simple, complete, average, and ward were tested to obtain sixteen different dendrograms. The dendrogram with the best separation and most accurate groupings was selected.

A linear discriminant model was then built with the same parameters and evaluated for its performance by sampling one-thousand times from the dataset. Each repeated sampling splits the data into a training and test set, which was used to evaluate the false positive rate for each classification and overall average performance. The complete dataset was then used to construct a final model. The performance was calculated by the average percent of correct classifications, and the false positive rate for each class was calculated by dividing the number of false positives by the sum of false positives and true negatives.

Finally, two sets of blind tests were carried out with two samples from the center points in each instance. The first set of blinds labelled A and B, were analyzed for their heavy mineral fraction as described with the known samples. The resulting numerical counts were then treated in the same manner and used to predict their PCA and LDA scores, as well as to calculate their distance relative to the

Fig. 1. Map of the Lausanne area with the four locations labelled (VA = Valency, RE=Mont-Repos, DE Denantou, and BO=Bourget).[1]

Table 1
Light mineral fraction count for respective samples and replicate.

Sample	Quartz	Carbonate	K-spar	Albite	Anorthoclase	Oligoclase	Anorthite	Muscovite	Biotite	Not-classified	Total
RECC1	589	40	16	51	22	49	1	0	0	105	873
RECC2	472	21	10	5	5	14	0	0	0	82	609
RECC3	716	25	35	1	45	48	0	0	0	69	939
BOCC1	667	46	21	100	32	57	0	2	3	101	1029
BOCC2	717	32	24	69	34	65	0	0	6	101	1048
BOCC3	697	36	41	61	52	67	0	0	1	66	1021
DECC1	726	1	7	86	30	53	0	0	2	110	1015
DECC2	287	1	3	1	6	10	0	0	0	42	350
DECC3	714	0	24	62	64	55	0	0	1	46	966
VACC1	409	1	10	2	9	19	0	0	0	50	500
VACC2	692	54	31	15	54	77	0	0	0	77	1000
VACC3	676	29	22	66	60	48	0	0	4	56	961

known dataset. The models were then used to classify the possible locations of these blind samples using all three models. The LDA model provided a posterior probability for its classification using an equally distributed prior probability of 0.25 for each location.

After the termination of the first set of blind tests, refinements in variable selection were carried out to improve the models. Minerals with counts of less than 5% in all locations were eliminated and the models were rebuilt using the remaining variables. The process was then repeated with the second set of blind samples labelled C and D using the refined models for classification.

3. Results

3.1. Light mineral fraction

Table 1 shows the different counts for the established light minerals as classified automatically by the Aztec software based on their EDS spectra. In total, 12 samples were analyzed and 18,544 particles classified.

A visual representation of the mineral proportions can be seen in Fig. 2. As seen, quartz accounts for the large majority of the minerals present. Other mineral types typically make up for less than 10% of the light mineral suite. Visually speaking, the minerals do not appear to have much discriminating power aside from a few candidates that can potentially differentiate between DE and BO locations. These are: quartz, carbonates, oligoclase, and potassium rich feldspars (K-spar).

Scaled principal components obtained using this data set without biotite, muscovite, or anorthite show no inherent potential to group or separate these samples by location (see Fig. 3). Pretreatments applied show no improvement in separating these samples and do not appear to mitigate the heavy weighting of the quartz fraction. An argument could however be made for DE samples being isolated in the bottom right quadrant mainly due to the loadings of the carbonates. In addition, the BO samples were also tightly grouped due to the loadings of quartz and oligoclase, however, these samples have huge overlaps with other location samples.

3.2. Heavy mineral fraction

Ten main types of minerals were identified to be present consistently in the heavy mineral fractions: apatite, chlorite, chloritoid,

---

[1] Map from the Guichet Cartographique Cantonal, https://www.geo.vd.ch/

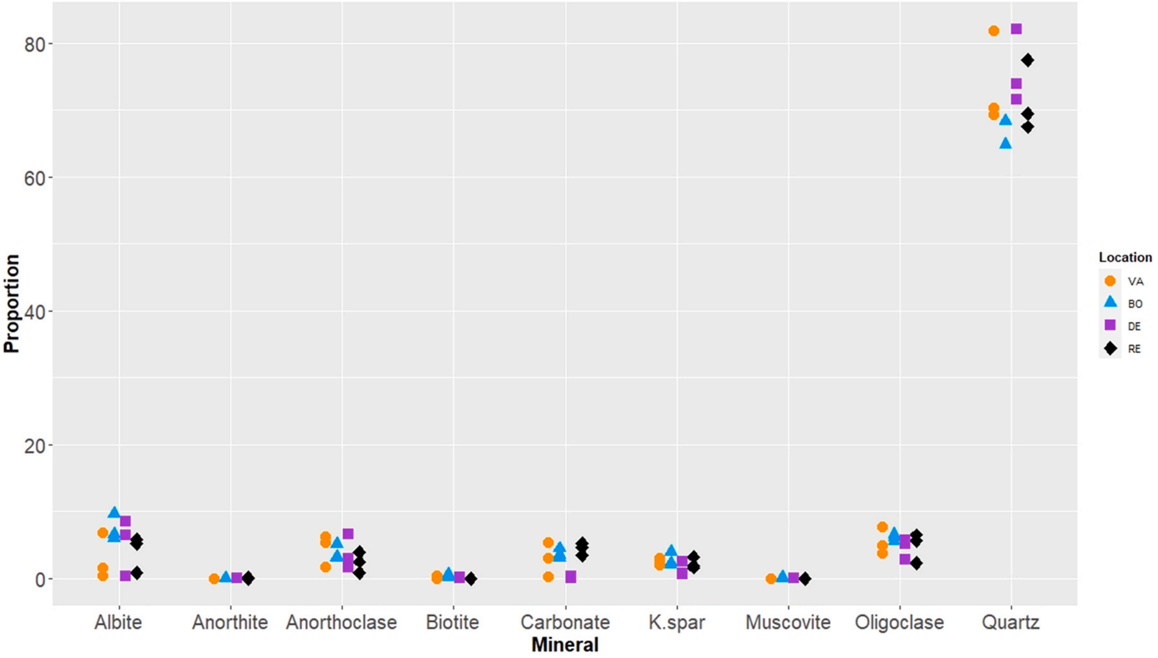

Fig. 2. Percent proportions of light minerals from different locations (n = 12).

Fig. 3. First two principal component scores and variable loadings of light minerals (n = 12). Small symbols represent the scores of the actual sample and the big symbols represent the center of each location group.

Table 2
Heavy mineral counts of respective samples and replicates.

Sample	Garnet	Glaucophane	Tourmaline	Hornblende	Chloritoid	Dolomite	Apatite	Epidote	Chlorite	Opaques	Others	Total
BOCC1	2	3	2	7	2	41	8	21	49	38	130	303
BOCC2	1	1	3	6	1	43	7	19	53	34	152	320
BOCC3	7	4	3	8	1	48	20	29	49	27	120	316
BONN1	5	3	5	8	3	44	9	25	36	33	158	329
BONN2	4	3	2	10	3	54	7	21	44	40	170	358
BOSS1	6	3	3	9	4	39	19	35	60	46	132	356
BOSS2	13	3	2	10	2	32	6	32	49	37	145	331
DECC1	4	3	2	17	3	1	17	51	29	34	145	306
DECC2	10	4	4	9	4	0	22	43	24	59	127	306
DECC3	7	3	3	16	6	1	18	58	25	57	144	338
DENN1	14	3	5	14	5	1	14	38	25	44	148	311
DENN2	10	6	7	14	4	2	14	53	36	58	146	350
DESS1	11	4	5	15	2	1	7	30	22	48	152	297
DESS2	11	3	5	10	7	0	7	44	27	47	139	300
RECC1	5	2	3	3	0	26	16	20	29	45	149	298
RECC2	4	1	2	2	2	30	7	25	33	58	145	309
RECC3	14	3	2	7	2	24	14	36	40	64	125	331
RENN1	4	1	7	13	1	33	10	22	29	50	185	355
RENN2	7	2	3	7	5	50	10	31	45	50	136	346
RESS1	7	2	4	5	2	50	9	34	43	66	167	389
RESS2	5	4	3	9	2	48	9	28	34	65	119	326
VACC1	5	3	3	12	4	25	4	28	37	30	151	302
VACC2	4	5	3	14	1	17	9	43	40	45	144	325
VACC3	5	3	5	15	5	24	20	34	48	45	136	340
VANN1	6	3	4	11	10	36	12	34	34	39	135	324
VANN2	9	5	5	16	3	45	10	34	38	40	131	336
VASS1	4	4	4	12	7	30	9	32	25	45	146	318
VASS2	2	5	3	18	9	33	19	23	45	41	142	340
Blind A	6	4	4	17	7	1	21	46	23	46	131	306
Blind B	8	3	3	16	4	42	10	26	30	48	145	335
Blind C	8	4	8	7	4	48	12	22	39	26	175	353
Blind D	4	2	2	5	1	38	8	20	22	39	156	297

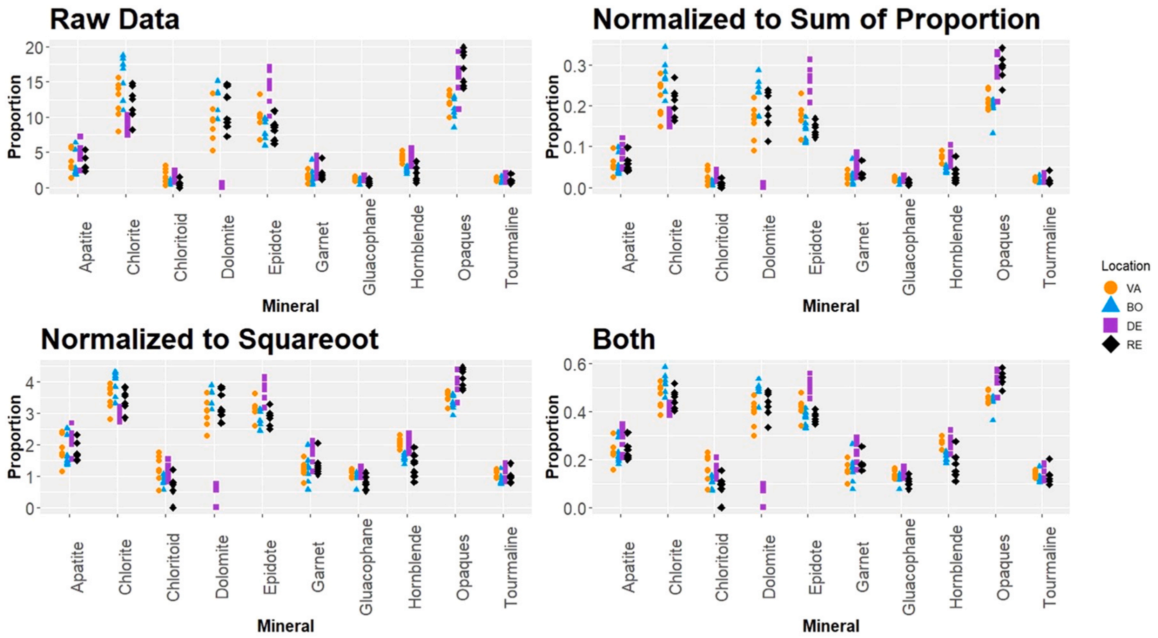

Fig. 4. Percent proportion of selected minerals in the heavy mineral fraction by location and pretreatment (n = 28).

dolomite, epidote, garnet, glaucophane, hornblende, tourmaline, and opaque minerals. In total, 32 samples, including the blinds, were counted representing a total of 10,451 particles visually identified and counted.

The raw heavy mineral counts of these ten minerals can be found in Table 2. Remaining minerals under “others” include rock fragments that were polycrystalline, and quartz that have hematite inclusions that render them heavy. Also present in minute amounts were glauconite, zircon, and some pyroxenes. The raw data revealed some potential for differentiating between certain locations. This is especially the case for chlorite, dolomite, epidote, hornblende, and the opaque minerals as seen in Fig. 4. There is also some indication that Hornblende counts are distinct between VA and DE locations from BO and RE locations, a trend that is also observed to a weaker extent in glaucophane. Both chlorite and epidote proportions show strong potential to separate between DE and BO samples, as well as

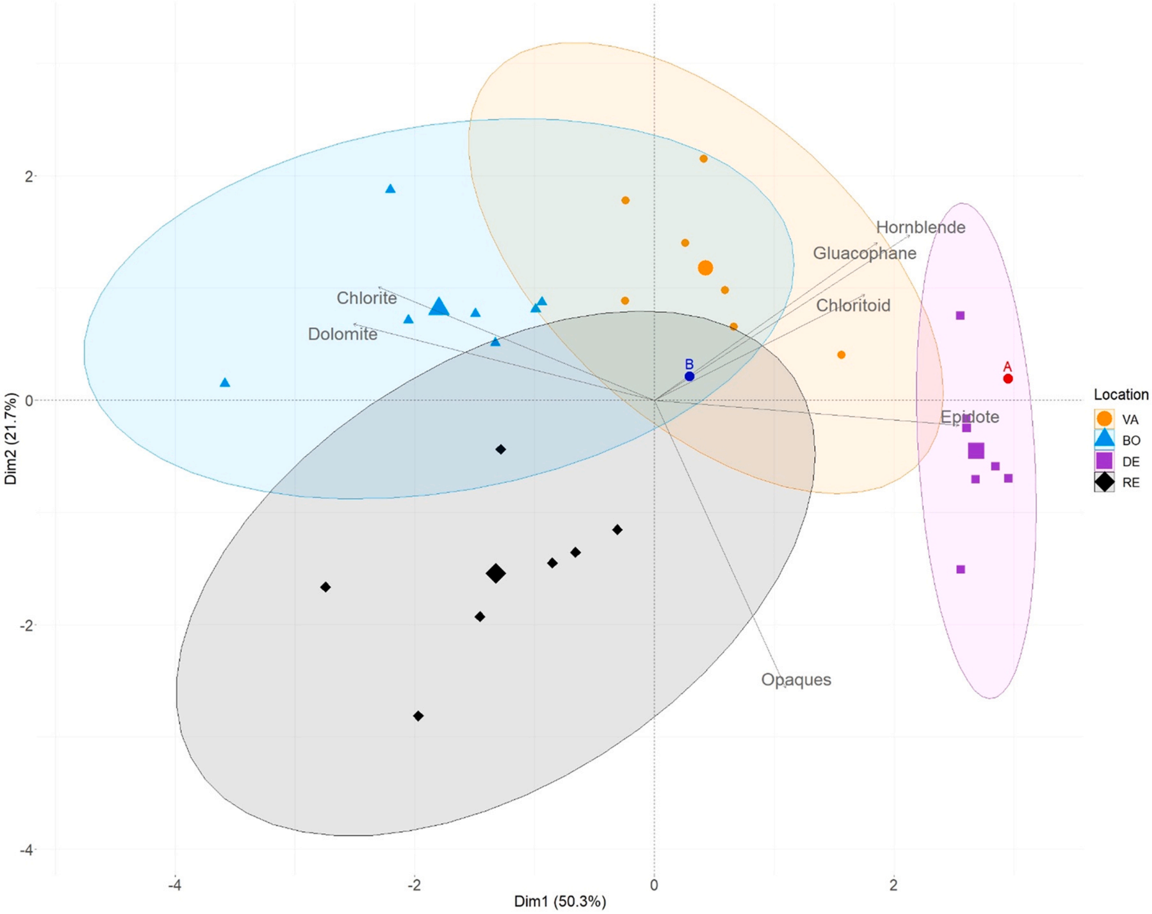

Fig. 5. First two principal component scores and variable loadings of heavy minerals with prediction of blinds (n = 28) with the predictive scores of two blind samples. Small symbols represent the scores of the actual sample and the big symbols represent the center of each location group.

to a lesser extent from both VA and RE samples. Dolomite strongly separates DE samples from the rest of the samples given that there are close to none. Finally, the proportions of opaque minerals indicate a potential difference between counts in VA and BO samples from DE and RE samples. When the samples are grouped by their cardinal directions (NN, SS, and CC) there are no obvious trends that could be observed with respect to their mineral counts.

Applications of data pretreatment methods generally aid in spreading out the variation in low proportional count data relative to the higher proportions, resulting in clearer separations in some cases. As seen in Fig. 4, the spread of the proportions within some locations were reduced for higher proportion minerals such as in BO and DE for chlorite, RE for epidote, and BO, as well as VA for the opaque. The combination of increased inter-variation and decreased intra-variation appears to aid in the separation of certain groups based on specific minerals whose discriminating potential was already observed in the raw data.

Multinomial logistic regression for each mineral by pretreatment also highlights these effects with a general decrease in p-values in certain important minerals as shown in the Annex Table 2. Overall, the utilization of both pretreatment methods seems to alter the variations in the most favorable way in terms of separating by location and was therefore chosen. The following six minerals were selected based on the criterion mentioned involving their p-values (< 0.1) from the multinomial logistic regression: hornblende, glaucophane, chloritoid, epidote, chlorite, and the opaque minerals. Dolomite was also selected despite not meeting the p-value criteria, as it was a clear discriminating variable based on the visual plots. Its high p-value in the multinomial logistic regression will be discussed in the next section.

3.3. First attempt at a model

Principal component scores derived from these minerals show a strong clustering and separation for DE samples based on the first two principal components. While the other three locations do show separation, there is a more uncertainty in their groupings, with their confidence ellipses exhibiting overlap (see Fig. 5). Based on the loadings in Fig. 5, DE samples are primarily isolated from the other location samples based on the first principal component, which is heavily weighted by the proportion of dolomite, epidote, and to a lesser extent chlorite. DE samples have a much higher variation in the second principal component mostly influenced by the opaque minerals. This first principal component is also responsible for the weak separation between the BO and VA groups. Between these two groups, the opaque mineral proportion does not vary significantly, and they are separated from the RE samples by the second principal component based on the opaque minerals. Finally, RE and VA are also separated in the diagonal direction based on glaucophane, hornblende, and chloritoid counts.

Distance calculation as well as the dendrograms reveal that utilizing the Manhattan distance with a ward clustering yields the most accurate groupings. In which all of the DE and VA samples were correctly classed, two BO samples were incorrectly classed as VA, and one RE sample was incorrectly classed as BO shown in Fig. 6.

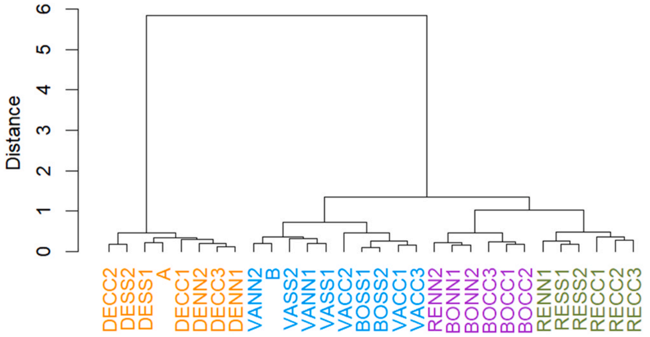

Fig. 6. HCA of selected minerals using manhattan distance and ward clustering (n = 28) with the inclusion of two blind samples.

Table 3
False Positive Rates (n = 1000) and Posterior Probabilities for Blind Samples A and B. Classification Using LDA With Selected Minerals from the first model. (n = 28).

	BO	DE	RE	VA
False Positive Rate (%)	11.64	0.00	19.00	9.98
Posterior Probability (A)	<0.001	>0.999	<0.001	<0.001
Posterior Probability (B)	0.0659	<0.001	0.150	0.784

LDA using the same variables and resampling for a thousand times yielded an average performance of about 85%, the false positives rates are shown in Table 3. The model is thus apparently effective in distinguishing samples of DE origin, with much more uncertainty in classifying samples of other locations, especially RE, which performed the worst.

3.4. First set of blind tests, samples A and B

Utilizing these three models, the identities of the blind tests were approximated. Predicted principal components scores in Fig. 5 show blind Sample A to be distinctly clustered with the DE samples and blind sample B to be positioned close to the center of the other three locations, although noticeably closer to VA. Hierarchical clustering using the Manhattan distance measurement and ward clustering also classified the blind sample A into the DE group and the blind sample B into the VA group in Fig. 6.

The LDA model provided the posterior probabilities shown in Table 3. This model classified blind sample A into the DE group with a probability of close to 1, with the other classes being negligible in their probability. As for blind sample B, the model gave the highest posterior probability for VA but non-negligible values for RE and BO, in that order. The true identities of the blind samples for A and B are DE and RE respectively. Hence, blind sample A was classified accurately, while there were certain inaccuracies for blind sample B. It was concluded that there were issues with this attempt at a model and further refinements were carried out along with an addition of a second set of blind tests to test the new models.

3.5. Model refinement, evaluation, and all blind tests (Samples A to D)

In this second attempt of a model, only data from minerals that had more than 5% proportions were used. These minerals represent the high-count variables and are chlorite, dolomite, epidote, and the opaque minerals. The first two principal components scores derived from these minerals resulted in a larger overlap of the confidence ellipses between locations and a larger grouping size, although separation is still visible. Sample B is located in the overlap of the VA and RE samples but much closer to the center of the RE samples this time around. Sample A is still distinctly clustered among the DE group. With respect to the new blind samples, scores for sample C and D landed close to the center of the BO and RE clusters respectively. These scores are shown in Fig. 7.

False positive rates increased for the classification of BO, RE, and VA classes to approximately 20% for each, and overall model performance also deteriorated to around 80% in the reworked LDA model. Posterior probabilities for sample A remained the same, with a close to 1 probability for DE and negligible probabilities for the other classifications. Concerning sample B, the two highest posterior probabilities were for VA and RE with approximately an even split between the two. As for the second set of blind samples, the highest posterior probability for sample C and D are BO and RE respectively, with other locations presenting significantly lower to negligible posterior probabilities. Table 4.

The identities for the second set of blind tests were BO and RE for samples C and D respectively. Therefore, they appear to be correctly classed using the refined model.

4. Discussion

4.1. Light mineral fraction

At a preliminary glance, none of the minerals in the light mineral appear to have strong discrimination potential between locations. Quartz and potassium rich feldspars appear to have some potential between only two locations (DE and BO) and carbonates have a strong potential in discriminating DE samples from other locations. This trend in the carbonates can also be observed in the heavy mineral fraction with dolomite, a carbonate variety. Overall, the light minerals do not appear to have a strong classification potential for all four groups as compared to the heavy minerals. Moreover, the proportions are heavily weighted by quartz which resulted in the other minerals representing around 10% of the proportion at best. Data pretreatments were unable to reduce this heavy weighting from quartz. As a result, focus was drawn towards the heavy mineral fraction to differentiate between these locations and no further work or data treatment was attempted on the light mineral fractions.

4.2. Heavy mineral fraction

The heavy mineral fraction was, as expected, a lot more varied in terms of mineral types. With ten mineral types consistently detected in all samples. Other minerals were detected in certain samples, but they were at an alarmingly low count and were inconsistent between replicates such that it would not be appropriate to use them as classifiers. These were thus disregarded, but kept as part of the total count for normalization.

As seen in the proportion plots (Fig. 4), there is evidence to suggest that several of these minerals could aid in the classification. Most of these are higher count minerals (more than 10% on average) with visible separations in at least one of the locations. A strong example is that of dolomite for location DE where little to none were present in this location. DE has a large population of pine trees with acidic leaves. It is thus possible that the carbonates in the soil were removed from this location due to reactions with the acids [12]. This is supported by these locations having a lower pH from previous tests [27]. Whatever the explanation may be, dolomite is a strong classification parameter for the DE samples. However, the lower count minerals proportions appear to be more random, mirroring the trends seen in the light mineral fraction.

Data pretreatment was employed with the intention of reducing the weight of these higher count variables on the classification and evaluating if some of the lower count minerals had classification potential which could be enhanced. The plot of proportions indicate that the pretreatments appear to alter the inter- and intra-variability in a favorable way that improves the separation power in the minerals. Namely, by reducing the intra-variability within a location

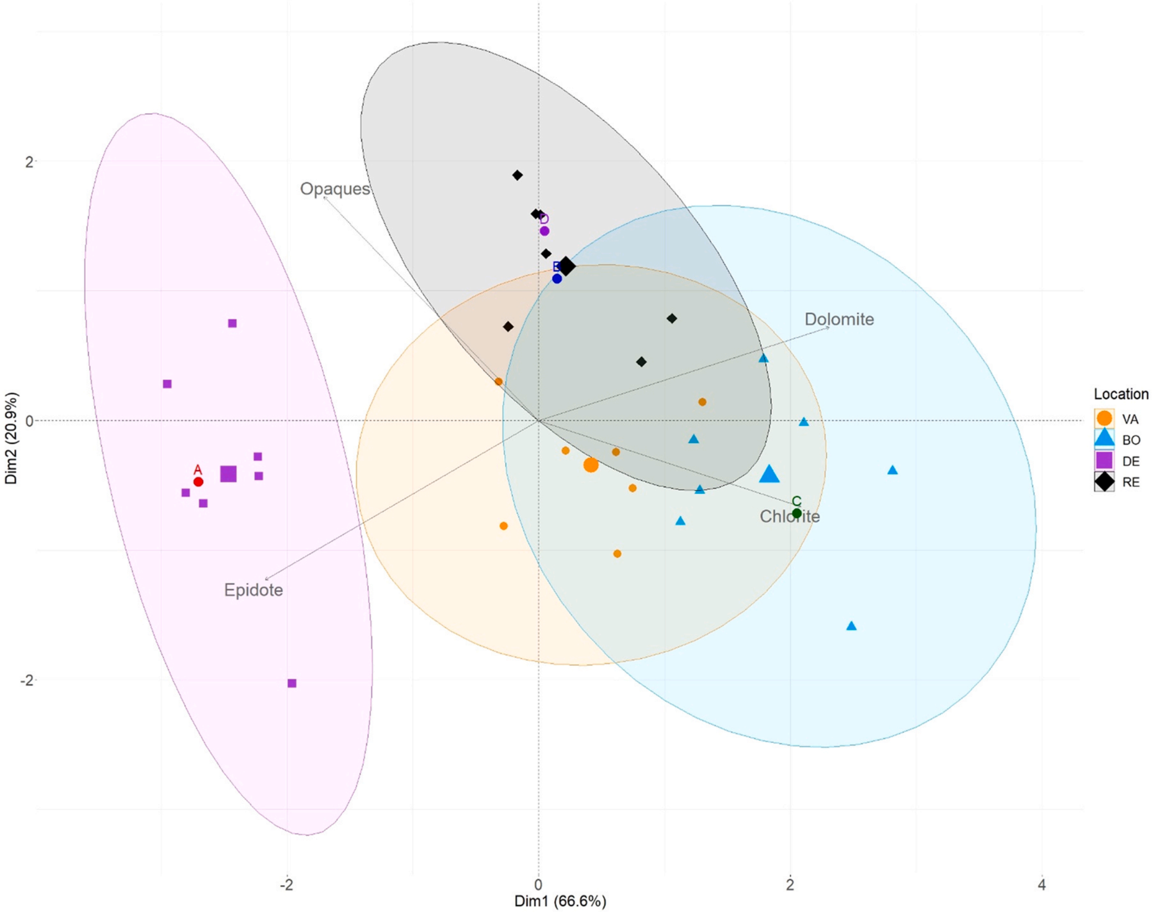

Fig. 7. Principal component scores from high count minerals and predicted scores of blinds (n = 28) with the predictive scores of four blind samples. Small symbols represent the scores of the actual sample and the big symbols represent the center of each location group.

Table 4
False positive rates (n = 1000) and posterior probabilities for blind sample classification using high count minerals.

	BO	DE	RE	VA
False Positive Rate (%)	18.21	0.00	18.34	15.56
Posterior Probability (A)	<0.001	>0.999	<0.001	<0.001
Posterior Probability (B)	0.0348	<0.001	0.493	0.472
Posterior Probability (C)	0.872	<0.001	<0.001	0.128
Posterior Probability (D)	0.008	>0.001	0.981	0.011

and increasing the inter-variability between locations. The multinomial logistic regression on these minerals also supports this notion, with several p-values lowered after the pretreatment was applied (see Annex Table 2). It is important to note however that multinomial logistic regression operates under the assumption that classes cannot be perfectly distinguished. As such they do not yield accurate p-values for minerals when classes are clearly distinct such as dolomite for DE. Post pretreatment, several minerals appear to remain unexploitable for classification purposes, and were thus eliminated based on their apparent lack of separation in the proportion plot and higher p-value for more than one location class. These were apatite, tourmaline, and garnet. The threshold was set at 0.1 for the p-values, while an admittedly arbitrary choice, it allowed for the selection of variables that could still be adjusted moving forward. The sole exception to this rule was dolomite who's low p-value was due to its strong discriminatory power [21]. Overall, pretreatment was able to improve the classification ability of some of the lower count minerals such as chloritoid, hornblende, and glaucophane.

The selected list of pretreated variables contributed to principal components that were able to cluster DE samples strongly and the other three locations with less certainty. Nevertheless, there was an apparent grouping and tendency for each sample to be clustered by their location. Based on the observations of the loadings (Fig. 5), the locations were grouped into their respective locations not based on all the selected variables but between one to three key mineral proportions. The fact that RE and VA samples were separated in the first model most strongly by hornblende, glaucophane, and chloritoid proportions became crucial when the first set of blind tests were evaluated. The key discriminant variables may differ for completely different locations and should thus be evaluated based on a case-by-case basis.

Hierarchical clustering, as well as LDA demonstrated more of the same ideas with the DE samples being distinctly classed and a higher uncertainty in classing the other three samples by location. The additional benefit of the LDA model is the ability to provide a quantification of these uncertainties, thus evaluating the performance of the model. False positives rates for RE classification appear to be the worst with BO and VA not far behind. Such quantification of the uncertainty can be applied to case-work scenarios to clarify the uncertainties in the results, as well as in model selection. Additionally, posterior probabilities calculated in the classification of unknown samples allow for concrete information to be provided on the potential sources rather than just an exclusion or inclusion result [20]. Based purely on the counts of the heavy mineral fractions (see Table 2 and Fig. 4) it would be difficult to determine at which point these samples can be classified as different, especially for the BO, VA, and RE samples. Assuming a valid construction, these models give

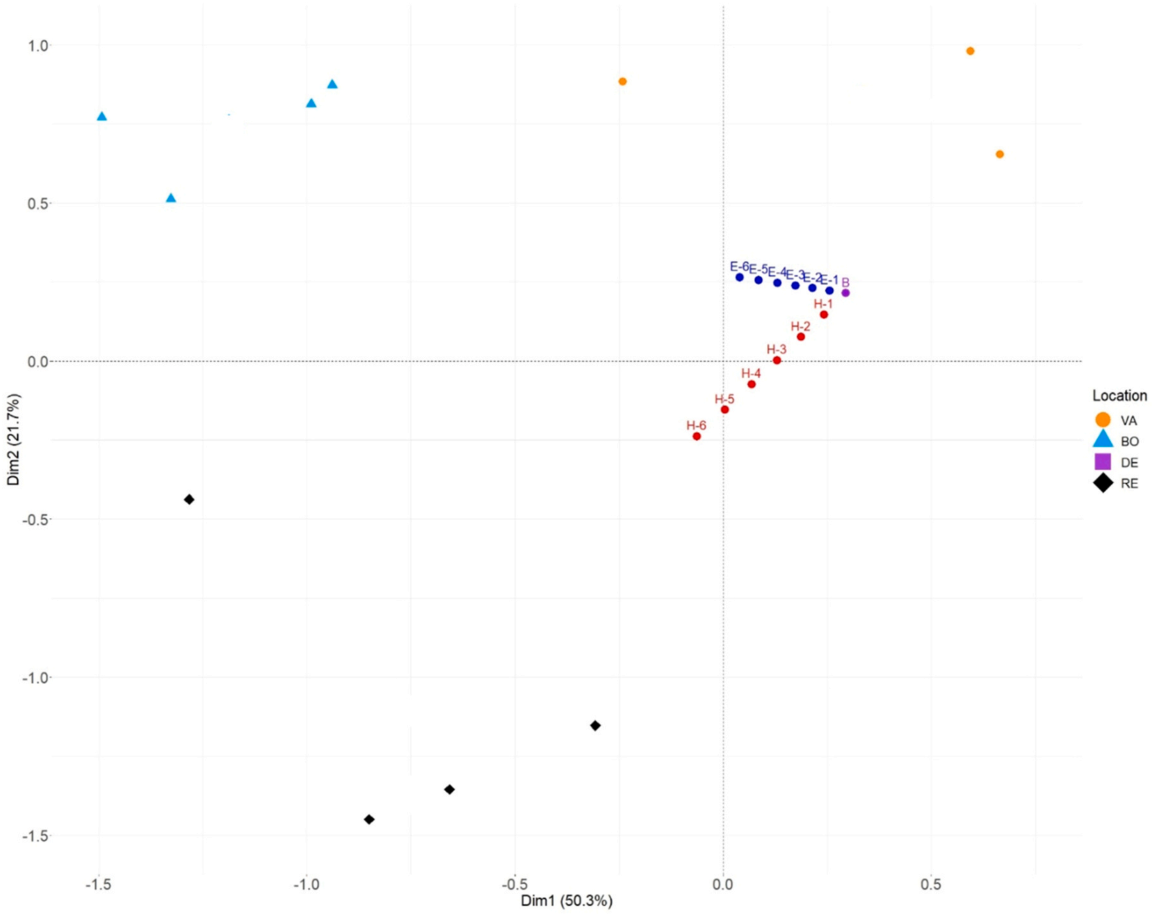

Fig. 8. Zoomed in plot of principal components of the first model with sample B and manually altered counts of epidote (E) or hornblende (H) for sample B. The point labelled B is the original location of the blind sample B, while each point labelled E and H are the alteration in count for epidote and hornblende respectively followed by the number of counts subtracted.

Table 5
Posterior Probabilities from the first LDA model for Sample B and Sample B with Altered Counts of minus 6 for Hornblende and Epidote Respectively.

	BO	DE	RE	VA
Posterior Probability (B)	0.0659	<0.001	0.150	0.784
Posterior Probability (B, Hornblende - 6)	0.122	<0.001	0.832	0.447
Posterior Probability (B, Epidote - 6)	0.036	>0.001	0.166	0.797

insight into the tendencies of the mineral counts as well as the limitations of the analysis in a meaningful way. Although it may be clear that certain mineral count proportions vary in such a way that suggest a difference based on location, without having an idea of how the variation affects these discriminatory properties, it would be easy to falsely exclude or include an unknown sample based solely on the apparent differences or similarities in the counts. The information provided from the models serves to give a more accurate interpretation of an unknown sample. In the case of this study, these models would be able to classify an unknown into the DE group with an appreciably high certainty and with less certainty in the other three groups, depending on the mineral counts. As such, certain samples could be more confidently classified than others depending on their variation. This was explored in the blind testing phase.

The incorrect classification of blind sample B into the VA class accentuates the uncertainties associated with the first model. Additionally, it is worth noting that the models acknowledge this uncertainty instead of providing a concrete statement of inclusion or exclusion. Without developing these uncertainties, the analyst would be stuck making a decision on the samples' classification and could wrongly classify the sample with unfounded confidence. Nevertheless, the posterior probabilities provided for the classification for blind sample B in Table 3 were still determined to be higher than an acceptable margin of error. There was thus clearly a need to refine the models.

As observed in the loadings plot, there are several redundancies in terms of the variables, with glaucophane, hornblende, and chloritoid heavily weighing the separation between VA and RE samples, the two locations in contest for sample B. These variables are all low count minerals of less than 5% proportion. As previously mentioned, there is a risk in the employment of such data because the estimate of relative standard error for these parameters are higher. In other words, a small change in counts affects the normalized proportion to a more significant extent as compared to the same changes in a higher count mineral. Fig. 8 illustrates this phenomenon with the hornblende (H) and epidote (E) count in sample B being manipulated by a subtraction of six in increments of one. The shift was much more drastic for each subtraction in hornblende count as compared to epidote. Overall, the shift for hornblende was enough to move sample B much closer to the RE cluster whereas the shift for epidote kept sample B in the same cluster of VA. Numerically, in principal component space, a difference of six in the hornblende counts shifted sample B by 0.577 units. This is contrasted to 0.283 units for epidote, a difference of approximately half. The changes to the posterior probability due to these alterations also show a much more significant effect on the classification with changes in hornblende counts as compared to epidote (see Table 5). When the hornblende count was reduced by six, posterior probability shifted in favor of RE instead of VA. Whereas for epidote, a change in value of six barely affected the posterior probabilities.

This occurrence is likely due to the fact that the hornblende minerals have a lower average count in the samples, which results in any deviation having a greater proportional difference with respect to the mean [21]. As an example, for a sample with a mean count of 10, a difference of 1 in the count corresponds to a 10% change whereas for a mean of 30, it is 3%. The estimate from a lower mean count is thus more prone to error, especially over a small number of trials. Therefore, it seems apparent that caution should be exercised when using a low count mineral and the threshold should be dependent on the circumstance. If lower count minerals were desired to be employed, either a higher number of total counts or more replicates should be applied to get a more accurate depiction of the mean and standard deviation. These corrections may not be practical due to the time-consuming nature of this analysis, as well as small sample sizes in case work. It might thus be pragmatic to stick to more prevalent minerals. The models were then refined to drop all variables that accounted for less than 5% in more than two locations. Four variables remained: dolomite, epidote, chlorite, and the opaque minerals.

The refined PCA and LDA models provided a much more reasonable evaluation of the uncertainty towards the classification of unknown sample B, with the disadvantage of a loss in performance. Although locational groups are no longer as clearly separated from before, and false positive rates have increased, these newer models better reflect the reality of the similarities between the groups. Additionally, the classification for the two sets of blinds were also much more accurate. The higher discrimination with the previous model was likely a result of over parameterization, a model fitted to the training data specifically, but does not in fact yield accurate results for data outside the training set. A conservative model such as the latter type should be favored for application when the true identities of unknowns cannot be verified. Finally, it should be noted that there are several factors that could influence the sample counts outside of random variation for both fractions. Most of these issues arise from the sample preparation phase, where identical treatment of each sample is not possible due to random and human error. An interesting aspect to note was that the first set of blinds were not homogenized before sample preparation, whereas the blinds in the second set were. It is possible that this resulted in the lower confidence in classification for blind sample B as compared to C and D.

As previously mentioned, such a probabilistic approach is highly applicable to casework. Firstly, it allows for the advancement of the interpretation of the results beyond “non-differentiation”. Based solely on their mineral counts, samples of VA, RE, and BO may have been described as “cannot be differentiated”. The location DE is a classic example of an ideal situation where one single mineral provides the smoking gun for differentiation. Yet it is still important to quantify this conclusion, the result that sample A “non-differentiated” from the location DE clearly does not have the same strength as the three other blind samples being classified as “non-differentiated”. By studying the variation and tendencies of these counts, useful patterns and parameters could be identified that begin to differentiate between the three questionable locations. There is no doubt that just as with any method, an uncertainty comes with this approach, but these uncertainties are identifiable and quantifiable. This thus permits the dissemination of increasingly accurate information for the decision maker. For example, instead of describing blind sample B as simply being non-differentiated from the three locations in question, one can provide the conclusion that the mineral count does not support the proposition that it is of BO origin and provides approximately equal support for both VA and RE origins. This is much more revealing than stating simply that the samples cannot be differentiated. It explains to what extent the samples are similar or dissimilar in a way that is useful in answering a case relevant question [20]. It is also important to note that only one component of the soils were evaluated, and there is a plethora of other means to analyze the soil [4,15] that could serve to decrease the uncertainty of the interpretation and provide a more robust classification of the various samples. Previous methods (XRF, pH measures, particle size distributions, and GC-MS) applied to attempt to differentiate these samples yielded similar results, in that there was a high overlap between the intra and inter-variability. In order to combine these analyses in a way that correctly weights the discrimination of the different methods, a probabilistic approach has to be employed [20].

Finally, the samples worked with in this study (3–4 g) are quite large by forensic science standards. Questioned or trace samples in case work are often miniscule in nature [3]. Replicates are thus most likely impossible for the such samples. In addition, the minute samples may mean that a critical number of minerals to be counted may not be reached. This study is limited in this regard and future work could investigate the effects of smaller total counts on interpretation. Another issue that is not evaluated in this study is that of differential transfer. Questioned samples are typically found on a secondary substrate after being in contact with the primary source. This creates another layer of uncertainty as to what proportion of minerals are transferred and if there are differences between different mineral types. Since comparisons would be made to reference samples collected from the source locations, they would not have gone through the same unknown interactions with the substrate or be diluted by background particles already present on the target surface. As such, it is very unlikely that a one-to-one comparison would be appropriate [6,18]. It would be interesting and pertinent to explore the effects of transfer, persistence, and background of the mineral suite composition in soil onto different target substrates for future studies [30]. Such a study would also benefit the interpretations at the activity level [20].

5. Conclusion

Four parks in the Lausanne area were sampled for their soils. These samples were treated and analyzed for their mineral composition at the size between 90 and 180 μm and separated into light and heavy mineral fractions. The resulting mineral proportions were then treated with multivariate statistical methods to provide a probabilistic approach in answering source level questions.

The multivariate statistical evaluation of the light mineral counts in this mineral suite study revealed a limited utility of the light mineral fractions to discriminate between samples from the four locations. Due to the low potential, efforts concerning this fraction were abandoned to focus on the more promising fraction.

On the other hand, the evaluation of heavy mineral counts demonstrates greater potential. The pretreatment of the heavy mineral fraction improved discrimination for certain variables and resulted in the successful creation of three separate models that were able to group the samples based on their locations. The uncertainty of such models was abundantly clear either visually or quantifiably. Blind testing successfully attributed three out of four samples to their actual sources but also revealed certain weaknesses of these models. For the last sample, classification was not as straightforward, but the models were still able to provide a correct classification with lower confidence. It was also established that low count variables were not suitable for classification due to the higher relative standard error. Care should thus be taken with small sample sizes when choosing variables to use as classifiers. These models also highlight the prevalence of uncertainty in classification and demonstrate that the interpretations do not have to be confined to strict categories of differentiated or not differentiated. Posterior probabilities, as well as error rates, can be calculated to quantify the uncertainty in order to provide more accurate information towards answering source level questions. Studies looking into probabilistic approaches towards forensic soil examination should be further explored to increase the aid towards the interpretation of results.

Such a probabilistic approach is highly applicable to complex casework as it allows for the advancement of the interpretation of the results beyond “non-differentiation” and enables to quantify this conclusion. There is no doubt that just as with any method, an uncertainty comes with this approach, but these uncertainties are identifiable and quantifiable. This avoids the so-called fall off the cliff effect of using the generic terms relating to “differentiation” to describe samples that have varying probative values. Such an approach thus permits the dissemination of increasingly accurate information for the decision maker. This information can also then be combined in a sensible way with other analysis of soil samples that may also be done in parallel.

Future work could explore such approaches in other components in the soil or single mineral studies. Furthermore, mineral suite studies such as this one can also be expanded towards activity level problems by exploring issues of transfer, persistence, and background on the mineral counts.

Declarations of interest

None.

Annex

See Table A1 and A2.

Table A1
Light mineral fraction identification scheme.

Mineral	Sub-type	Classification Criteria
Quartz	–	$Si > 79\%$, and $P < 15\%$
Carbonate	–	$Ca > 80\%$
Feldspar	Orthoclase (K)	$45\% < Si < 72\%$, $P < 12\%$, $12\% < Al < 35\%$, $9\% < K < 35\%$, $Na < K$, $Ca < 6\%$, and $Mg < 5\%$
	Albite (Na)	$65\% < Si < 85\%$, $P < 12\%$, $12\% < Al < 28\%$, $2\% < Na < 15\%$, $Ca = 0\%$, $K = 0\%$, and $Fe < 6\%$
	Anorthoclase (Na-K)	$55\% < Si < 85\%$, $P < 12\%$, $12\% < Al < 28\%$, $K > 0\%$, $2\% < Na < 15\%$, $Ca = 0\%$, and $Fe < 6\%$
	Oligoclase (Na-Ca)	$53\% < Si < 76\%$, $P < 12\%$, $12\% < Al < 28\%$, $Na > 2\%$, $Ca > 0\%$, and $(Na + Ca) > K$, and $Fe < 6\%$
	Anorthite (Ca)	$48\% < Si < 66\%$, $P < 12\%$, $12\% < Al < 30\%$, $Na = 0\%$ K, $Ca > 13\%$, and $Fe < 6\%$
Mica	Biotite	$45\% < Si < 72\%$, $P < 12\%$, $12\% < Al < 35$, $9\% < K < 35\%$, $1.5\% < Mg < 40\%$, and $1.5\% < Fe < 50\%$
	Muscovite	$45\% < Si < 72\%$, $P < 12\%$, $12\% < Al < 35$, $9\% < K < 35\%$, and $Mg < 10\%$,

Table A2
P-values for respective minerals at each location compared to VA with multinomial logistic regression ($n = 28$).

	Garnet	Glaucophane	Tourmaline	Hornblende	Chloritoid	Dolomite	Apatite	Epidote	Chlorite	Opaques
Raw Data										
VA vs BO	0.814	0.0560	0.134	0.0306	0.0268	0.262	0.705	0.0565	0.610	0.137
VA vs DE	0.0323	0.736	0.330	0.944	0.457	0.868	0.318	0.0354	0.0331	0.0462
VA vs RE	0.394	0.0107	0.407	0.0108	0.0166	0.222	0.657	0.114	0.775	0.0190
Normalization to Total Proportion										
VA vs BO	0.833	0.0611	0.204	0.0233	0.0312	0.0243	0.643	0.375	0.0567	0.220
VA vs DE	0.0318	0.917	0.279	0.564	0.558	0.196	0.198	0.058	0.0399	0.0203
VA vs RE	0.443	0.00996	0.606	0.00826	0.916	0.213	0.686	0.0776	0.726	0.0131
Square Root										
VA vs BO	0.956	0.0547	0.107	0.0285	0.0269	0.0269	0.670	0.0573	0.0624	0.142
VA vs DE	0.0257	0.732	0.389	0.993	0.608	0.695	0.308	0.0324	0.0321	0.0450
VA vs RE	0.408	0.0126	0.330	0.0101	0.0111	0.224	0.729	0.118	0.795	0.0176
Both										
VA vs BO	0.954	0.0564	0.148	0.0239	0.0287	0.0271	0.635	0.037	0.0545	0.225
VA vs DE	0.0212	0.866	0.325	0.633	0.708	0.680	0.188	0.058	0.0376	0.0109
VA vs RE	0.413	0.0130	0.430	0.00686	0.0109	0.229	0.735	0.0807	0.765	0.0131

The following are the parameters used in the sample preparation and SED-EDX analysis of the light mineral fraction:

Cito-Press Parameters

Temperature: 180 °C.
Press time: 5.0 min
Pressure: 150 bar.
Cooldown time:6.0 min

SEM-EDX Parameters

SEM

Accelerating voltage:20 KeV.
Probe current:6.
Process time:5 s
Image scan resolution:1024 px.
Dwell Time:15 μs.

EDS

Acquisition mode:Spot.
Scan time:6 s
Pulse pileup correction: Checked.
Auto-Id:All possible elements checked except Oxygen and Carbon.
The following table describes the classification scheme used for mineral identification adapted from McVicar and Graves [10]:

References

[1] R.W. Fitzpatrick, Soil: forensic analysis, Wiley Encyclopedia Forensic Sci, John Wiley and Sons Limited, New Jersey, 2004, pp. 1–14, https://doi.org/10.1002/9780470061589.fsa096.pub2
[2] D. Pirrie, L. Dawson, G. Graham, Predictive geolocation: forensic soil analysis for provenance determination, Episodes 40 (2017) 141–147, https://doi.org/10.18814/epiiugs/2017/v40i2/017016
[3] D.A. Stoney, A.M. Bowen, P.L. Stoney, Inferential source attribution from dust: review and analysis, Forensic Sci. Rev. 25 (2013) 107–10742.
[4] K. Pye, Geological and Soil Evidence, CRC Press, Florida, 2007, https://doi.org/10.1201/9781420004755
[5] A. Bowen, Individualizing minerals: a proposed approach for forensic soil comparison, Microscope 55 (2007) 59–73.
[6] R.M. Morgan, P.A. Bull, Data interpretation in forensic sediment and soil geochemistry, Environ. Forensics 7 (2006) 325–334, https://doi.org/10.1080/15275920600996248
[7] H. Jenny, Factors of Soil Formation: A System of Quantitative Pedology, Dover, New York, 1994.
[8] M. Mange, H. Maurer, Heavy Minerals in Colour, Springer Science and Buisness Media, Berlin, 2012.
[9] S. Palenik, Heavy minerals in forensic science, Dev. Sediment. 58 (2007) 937–961.
[10] M.J. McVicar, W.J. Graves, The forensic comparison of soils by automated scanning electron microscopy, J. Can. Soc. Forensic Sci. 30 (1997) 241–261, https://doi.org/10.1080/00085030.1997.10757104
[11] D. Pirrie, G.K. Rollinson, M.R. Power, J. Webb, Automated forensic soil mineral analysis; testing the potential of lithotyping, Geol. Soc. Spec. Publ. 384 (2013) 47–64, https://doi.org/10.1144/SP384.17
[12] P. Duchaufour, Introduction à la science du sol, Dunod, Paris, 2014.
[13] D. Bloss, Opt. Crystallogr. (1960), https://doi.org/10.1063/1.3056750
[14] W.J. Graves, A mineralogical soil classification technique for the forensic scientist, J. Forensic Sci. 24 (1979) 10839J, https://doi.org/10.1520/jfs10839j
[15] R. Murray, Evidence of the Earth, second ed., Mountain Press Publishing Company, Montana, 2011.
[16] A.M. Bowen, E.A. Caven, Forensic provenance investigations of soil and sediment samples, Geol. Soc. Spec. Publ. 384 (2013) 9–25, https://doi.org/10.1144/SP384.4
[17] W.C. McCrone, J.G. Delly, The Particle Atlas: An Encyclopedia of Techniques for Small Particle Identification, Ann Arbor Science Publishers, Ann Arbor, 1973.
[18] R.M. Morgan, P.A. Bull, The philosophy, nature and practice of forensic sediment analysis, Prog. Phys. Geogr. 31 (2007) 43–58, https://doi.org/10.1177/0309133307073881
[19] C. Muehlethaler, G. Massonnet, T. Hicks, Evaluation of infrared spectra analyses using a likelihood ratio approach: a practical example of spray paint examination, Sci. Justice 56 (2016) 61–72, https://doi.org/10.1016/j.scijus.2015.12.001
[20] C.G.G. Aitken, F. Taroni, Statistics and the Evaluation of Evidence for Forensic Scientists, second ed., John Wiley & Sons, Ltd, 2005, https://doi.org/10.1002/0470011238
[21] J.M. Curran, Introduction to Data Analysis with R for Forensic Scientists, CRC Press, Florida, 2010, https://doi.org/10.1201/9781420088274
[22] D.A. Stoney, A.M. Bowen, P.L. Stoney, Loss and replacement of small particles on the contact surfaces of footwear during successive exposures, Forensic Sci. Int. 269 (2016) 78–88, https://doi.org/10.1016/j.forsciint.2016.11.015
[23] Best Practice Man. Forensic Comp. Soil Traces 2019.
[24] O. Guéniat, P. Esseiva, Le profilage de l′héroïne et de la cocaïne, Presses polytechniques et universitaires romandes, Lausanne, 2005.
[25] J.R. Bray, J.T. Curtis, An ordination of the upland forest communities of Southern Wisconsin, Ecol. Monogr. 27 (1957) 325–349, https://doi.org/10.2307/1942268
[26] D.A. Stoney, A.M. Bowen, M. Ausdemore, P.L. Stoney, C. Neumann, F.P. Stoney, Differential analysis of very small particles (VSP) from the contact surfaces and recessed areas of footwear, Forensic Sci. Int (2019) 106–114, https://doi.org/10.1016/j.forsciint.2019.02.032
[27] A.R. Marolf, L. Dawson, J. Ross, G. Massonnet, L. Gueissaz, V. T.W, Forensic Geology: analysis of plant-wax compounds on molassic soils in public sites in the greater Lausanne area Pap. Present. ENFSI 2016 APST Meet. Prague.
[28] A. Sommaruga, From the central Jura Mountains to the Molasse Basin (France and Switzerland), Bull. Fuer Angew. Geol. 16 (2011) 63–75.
[29] P. Esseiva, L. Gaste, D. Alvarez, F. Anglada, Illicit drug profiling, reflection on statistical comparisons, Forensic Sci. Int. 207 (2011) 27–34, https://doi.org/10.1016/j.forsciint.2010.08.015
[30] D. Werner, C. Burnier, Y. Yu, A.R. Marolf, Y. Wang, G. Massonnet, Identification of some factors influencing soil transfer on shoes, Sci. Justice 59 (2019) 643–653, https://doi.org/10.1016/j.scijus.2019.07.004