Soil can play an important role in forensic cases in linking suspects or objects to a crime scene by comparing samples from the crime scene with samples derived from items. This study uses an adapted ED-XRF analysis (sieving instead of grinding to prevent destruction of microfossils) to produce elemental composition data of 20 elements. Different data processing techniques and statistical distances were evaluated using data from 50 samples and the log-LR cost (C_{llr}). The best performing combination, Canberra distance, relative data, and square root values, is used to construct a discriminative model. Examples of the spatial resolution of the method in crime scenes are shown for three locations, and sampling strategy is discussed. Twelve test cases were analyzed, and results showed that the method is applicable. The study shows how the combination of an analysis technique, a database, and a discriminative model can be used to compare multiple soil samples quickly.

KEYWORDS: forensic science, elemental composition, Canberra distance, trace evidence, energy-dispersive X-ray fluorescence, soil comparison

Soil is often encountered as trace evidence on a wide variety of items and is a complex matrix of various organic and inorganic particles. If the composition of this complex matrix can be properly analyzed and compared to other soil samples, a powerful forensic tool is added to the toolbox of the forensic investigator. Techniques have been developed to analyze a wide range of characteristics or components such as color, pollen, diatoms, bacterial populations, particle size distributions, sand grain shapes, heavy minerals, and elements (1–7).

In casework, the question of common source is complicated by the different criminalistically relevant sources to compare the soil trace with. This source, for example, can be a footprint at a crime scene, the entire crime scene or the forest containing the crime scene. This is further complicated by the limitations of geological resolution where for instance, the soil in the footprint cannot be distinguished from the path it is found at, which in turn can be more extensive than the crime scene delimited by the crime scene officers. In the Bayesian framework, the evidential value depends on the hypotheses. For soil comparisons, this means that when a different definition of the term source is used, the evidential value could change. In this paper, we define source as an area that can be characterized by one sample from one point. This definition implies that when a site can be represented by one sample, site and source mean the same. This also implicates that the size of a site varies with the method used to characterize the soil and the geological variation present at the chosen area.

Combining different techniques to investigate the hypothesis of a possible common source (8–10) will improve the robustness of the comparison and, using independent characteristics, could also improve the evidential value of the comparison. From that perspective, the use of the elemental composition of soil samples can be very informative in combination with, for example, microfossil content (pollen, diatoms, phytoliths) or bacterial populations. As the amount of sample material is often very limited in forensic investigations, a nondestructive technique which allows for other analyses to be subsequently applied to the same soil sample is very valuable. For a forensic application, the results from the analysis should be compared using an objective method (e.g., 11,12). In addition, the results of the objective comparison should fit the Bayesian approach, which is recommended by the European Network of Forensic Science Institutes (ENFSI) guidelines for reporting evaluative forensic evidence. In this paper, we present such a nondestructive and objective method for the forensic comparison of the elemental composition of soil samples, which can potentially be extended to other quantitative measures of soil components. We assess the forensic implications of sample processing, treatment of detection limits, data processing for optimal representation of evidential value, optimal spatial sample resolution, and transport and storage effects.

Adapting to Nondestructive Analysis

For analysis of the elemental composition of soil, different techniques have been described, such as XRD (13), ICP-AES and ICP-MS (8), and ED-XRF (14). For soil comparisons, we chose to analyze the elemental composition of soil samples using ED-XRF. This technique is suitable for different soil types, and the minimum sample size is large enough for a heterogeneous matrix such as soil. A key advantage is that after the ED-XRF

1Netherlands Forensic Institute, Laan van Ypenburg 6, 2497 GB The Hague, The Netherlands.
2Department of Physical Geography, Utrecht University, Heidelberglaan 2 3584 CS, Utrecht, The Netherlands.

Received 17 Mar. 2016; and in revised form 28 July 2016; accepted 30 July 2016.

© 2016 American Academy of Forensic Sciences
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
analysis, the soil sample can be easily removed from the sample cup to be used for other techniques. When analyzing soil samples with ED-XRF for other purposes, for instance to quantify pollution with heavy metals, samples are usually ground to a powder to homogenize samples and increase precision (14). However, grinding is very detrimental to for instance pollen, diatoms, and phytoliths in the sample. To prevent the destruction of these microfossils, we changed the sample processing to sieving over a 250-micron mesh. As the sample is less homogeneous after sieving than after grinding it to a powder, we analyzed the sample three times with a thorough mixing between each analysis. To test whether sieved samples yielded similar results to powdered samples, we analyzed 10 soil samples using both methods (sieving and powdering).

**Processing the Elemental Data**

The next step after producing reliable elemental data is processing the data so they can be used for a statistical analysis. As part of the data is commonly reported as “below the detection limit” and quantified data is less reliable below the limit of quantification (LoQ, theoretically 10/3 the detection limit), a common question is what to do with values below these two limits. For the presented method, we removed elements with >90% of measurements below the LoQ. Remaining elements with values below the LoQ had these values replaced by the value of this limit. This prevents artificial differences between samples with a value just below the LoQ when compared with samples with a value just above LoQ. After this, processing elements with a high concentration (bulk elements) are still in the same dataset as elements with a low concentration (trace elements). The differences between samples in trace elements can be lost in the much larger differences between samples in bulk elements. To compensate for this, multiple data processing techniques were tested in combination with different statistical methods (see next section).

**Determining Optimal Statistical Method for Objective Comparisons**

To compare elemental data, multivariate statistical methods or statistical distance measures can be used. As multivariate statistical methods are less suited to compositional data (data that sums to one or 100% for each sample) (15), we chose to use a statistical distance measure. This quantifies the difference (or similarity) between two sets of data, such as the elemental data from two soil samples. An elemental concentration dataset with 50 soil samples was tested with five statistical distance measures.

To determine the best performing combination of data processing method and statistical distance measure, we chose to divide all the distances calculated between the 50 samples into two groups: a group with the distances calculated between the elemental data from the triplicate analyses of the same sample (hereafter abbreviated as the “same” group) and a group with the distances calculated between the elemental data from samples from different sites (“different” group) (4). The further apart these two groups are, the better the discrimination of the model in assigning an unknown sample to one of these groups. The number of false positives and false negatives should also be low. The log-LR cost \(C_{\text{LR}}\) (16)) is a measure for the number and weight of false LR results (both positive and negative); an ideal model would have a \(C_{\text{LR}}\) of zero. The combination of data processing method and statistical distance measure with the lowest \(C_{\text{LR}}\) was selected as the best performing.

**Spatial Resolution**

To interpret the results of a soil comparison, spatial resolution, and effects of transfer, storage, and sampling should be taken into account. The spatial resolution must be high enough to be of forensic use, meaning that it should consistently discriminate between two questioned sites, and not so high that sampling the correct point within the site is difficult. A high resolution is especially relevant for forensic casework in the Netherlands as the variability of the topsoil is generally lower than other European countries, due to the lower diversity in source material for the topsoil and the large amounts of soil that are transported and mixed for construction and agriculture. The degree of separation between the “same” group and the “different” group is also related to the spatial resolution of the method.

A high spatial resolution also means the method is sensitive to small-scale variation within a site. This has an effect on the sampling strategy of a crime scene; how many samples should be taken and where. When an unknown sample is matched to a sample from a crime scene, it is important for the forensic conclusion to delimit the matching area to a larger area including the crime scene, just the crime scene or a small portion within the crime scene. Ideally, almost no difference in elemental composition should be found within the crime scene and a large difference should be found between the crime scene and other relevant sites. The within-site variability is shown for three different crime sites with samples from different points within these sites using the statistical distances calculated between the sampling points.

**Effects of Transfer, Sampling, and Storage**

During the transfer from the soil surface to an item (such as a shoe), subsequent sampling, and varying storage conditions, soil samples can change (17,18). Due to the heterogeneity of soil, these changes can have a larger effect when smaller soil samples are analyzed and compared to the soil from which the sample originated. To test how the comparison method is affected by this, we collected test cases in cooperation with the Police Academy of the Netherlands. Students of the Police Academy, the training center for the national police, were asked to send in test cases with the following assignment: Make a shoe print in soil and collect a soil sample from the shoe sole, a soil sample from the upper centimeter of the resulting shoe print, and an unrelated soil sample from another site. The soil samples were collected using normal casework procedures and sent to the laboratory. No further instructions were given on how to sample or how to store the samples. Twelve complete test cases were received. After analysis, the results were discussed with the students of the Police Academy and additional information about the test cases was gathered. The information send along with the test case and the additional information given later was used to establish the physical distance between the site of the shoe print and the unrelated site and their description.

**Materials and Methods**

**Collecting and Preparing the Soil Samples**

Soil samples were collected in small plastic jars from the upper centimeter of topsoil (at least 10 g) from approximately 30 by 30 cm and dried and stored at room temperature until analysis.
For analysis, all soil samples were dried at 40°C for at least 16 h and then gently crushed using a pestle and mortar. Large botanical remains (roots, twigs and leaves) or anthropogenic particles (such as screws or pieces of brick) were removed by hand.

**ED-XRF Analysis**

Samples (one to three grams) were transferred to XRF sample cups (Chemplex, USA) with an internal diameter of 24 mm and a four-µm prolene film (Chemplex, USA).

For the adaption to a nondestructive analysis, 20 samples were analyzed once. For the other research questions, the samples were analyzed three times and thoroughly mixed in between. To determine the optimal combination of data processing and statistical distance measure, 50 sieved samples were analyzed resulting in elemental data of 150 analyses. For the examples of within-site variation, 26 sieved samples were analyzed resulting in elemental data of 78 analyses, and to test the effects of transfer, sampling, and storage, 36 sieved samples were analyzed resulting in elemental data of 108 analyses.

All samples were analyzed with a Xepos bench top ED-XRF spectrometer (Spectro Analytical Systems (SAS), Germany), equipped with a 50 W X-ray tube with Pd anode and Be exit window and a Peltier cooled Si drift detector (energy resolution <160 eV). Each sample was analyzed in a sample chamber flushed with helium, using four targets: Al$_2$O$_3$ as Barkla polarizer, an HOPG crystal as Bragg polarizer and a Mo target, and a Co target as secondary targets. The quantification was performed with the X-LabPro software using the TurboQuant method (SAS, Germany). This software is based on the fundamental parameters approach (FPA) with an automatic correction for the matrix used and is calibrated by the manufacturer on pure chemicals and certified reference materials.

Due to the limits of the instrument and low X-ray yields, the lighter elements showed an increased sensitivity to matrix and grain size effects. Therefore, the elements Si and Na were excluded. The element Fe (iron) was also excluded as in coursework, iron was found to be a common contaminant, especially when comparing soil traces from rusty tools with soil samples from a site. Elements which were not measurable above the limit of quantification (LoQ) in 90% or more of the samples of the largest dataset (50 samples) were also excluded. The LoQ was defined as $10/3$ of the limit of detection as given by the software. The elemental concentration data of the 20 remaining elements reported by the X-LabPro software was used: Mg, Al, P, S, Cl, K, Ca, Ti, Cr, Mn, Ni, Cu, Zn, Br, Rb, Sr, Y, Zr, Ba, and Pb (ordered by atomic number). Values below LoQ were replaced by their LoQ.

**Adapting to Nondestructive Analysis**

Ten soil samples with different soil types (three clay, four sand, and three organic-rich soil samples) were selected. Each of the 10 samples was homogenized and divided into two equal parts. One part was sieved through a sieve with a 250 micron mesh (sieved samples), and the other part was ground to a powder using a Vibratory Micro Mill Pulverisette-0 (Fritsch GmbH, Germany) for 5 min (powdered samples). The data were plotted against each other to an indication whether the measurements on any of the used elements were particularly sensitive to this change.

**Data Processing and Calculating Statistical Distance Measures**

To determine the optimal combination of data processing and statistical distance measure, soil samples were sampled from 50 different sites throughout the Netherlands. These sites included among others home gardens, public parks, fields, woodlands, and edges of roads, covering a variety of soil types composed mainly of sand, clay/silt, organic material or a mixture of these. The geographical distribution of the 50 samples is given in Fig. 1.

Using LR-GUI software (19), different data processing options and statistical distance measures were tested on the dataset with data from 50 different sites. The following eight data processing options were tested: no processing, converting the data to relative data (sum equals 100%, norm), dividing the concentration value of each element by the standard deviation of that element in the dataset (sd), using the normal logarithm of the concentration values (log10), using the square root of the concentration values (sqrt), and the following combinations of these options: sd and norm, log10 and norm, sqrt and norm. The data processing options were combined with five statistical distance measures: Canberra, Chebyshev, Minkowski, Bray–Curtis, and Euclidean.

The Canberra distance takes the absolute difference per element between samples, weighs it to the total sums of the element in question, and those weighed differences are then summed. The upper limit is the same as the number of elements compared (in this case 20). Chebyshev distance sums all the absolute differences per element between samples with no upper limit and considers all differences for all elements equally. Bray–Curtis distance sums all the differences per element between samples and then divides this by the sum of all element values. This distance considers all

**FIG. 1—Geographical distribution of the 50 soil samples from the Netherlands contained in the dataset to determine the optimal combination of data processing and statistical distance measure. In total, 10 clay/silt, 26 sand, and 14 organic-rich soil samples were collected.**
differences equally and has an upper limit of one. As the sums of both samples are used, the sums of both samples have to be equal, so usually percentage data is used. Minkowski distance takes the absolute difference between samples for an element, raises it to the power N, sums these differences to power N for all the elements, and then takes the Nth root of this sum. The Minkowski distance used has $N = \frac{1}{2}$ and the Euclidean distance (a type of Minkowski distance) has $N = 2$.

Chebyshev, Bray–Curtis, and Euclidean distance measures are mainly influenced by larger differences between elements (mainly the bulk elements with higher values) and are less sensitive to trace elements when these are combined with bulk elements. The Minkowski (with $N = \frac{1}{2}$) distance measure is influenced less by larger differences between elements, but is still more sensitive to larger differences than small ones. The Canberra distance measure is mainly influenced by differences between elements with a low concentration (trace elements) and is very sensitive to trace elements when combined with bulk elements.

The 40 possible combinations of data processing options and statistical distance measures were evaluated using the log-LR cost ($C_{llr}$) between the “same” and the “different” groups using the gamma distribution model fitted to the data. The best performing ($C_{llr}$) combination of data processing method and statistical distance measure was recalculated using PopTools (http://www.cse.csiro.au/poptools) in MS Excel.

Spatial Resolution

Three additional sites were sampled to show examples of the different spatial resolution of the method within different sites. The sites were chosen as examples of the broad variety in crime scenes. Site 1 was a back garden roughly 10 by 10 m with three garden beds. Each garden bed was sampled twice yielding six samples in total for this site. Site 2 was a caravan camping field with hedges and dirt roads. Ten samples were taken in a 20 m circle around one camping site. Site 3 was a ditch next to a rural asphalt concrete road with agricultural fields next to it. Ten samples were taken over a 50 m transect: along the ditch and along both sides of the road. Canberra (CNB) distances between the elemental data were calculated for the samples from each site. For site 1, 153 distances were calculated, and for sites 2 and 3, 435 distances were calculated for each. The CNB distances were categorized as: same sampling spot, up to 2 m, 3–5 m, 6–10 m, 11–15 m, 16–20 m, 21–30 m, and 31–50 m apart. The CNB distances in the categories were then evaluated using the groups as defined by the database constructed as described in the materials and methods section on data processing.

Effects of Transfer, Sampling, and Storage

Twelve complete test cases were received. On arrival, samples were directly processed, each sample cup was analyzed three times, and data were processed and CNB distances calculated (as described in materials and methods). Calculating CNB distances between the elemental data of two soil samples both measured three times therefore resulted in nine CNB distances. The average of these CNB distances was used as a consensus result for the two samples.

Results and Discussion

Adapting to Nondestructive Analysis

In Fig. 2, the square root of the concentration of 20 elements in the 10 sieved samples is plotted against the same 20 elements in the 10 powdered samples. The square root of the concentration values was used to compensate in part for the wide distribution of the data. Fig. 2 shows that many of the concentration values fall in the area between 90% and 110% between treatments, but there are some outliers.

![FIG. 2—Elemental data of sieved samples plotted against powdered samples. Square root of average elemental concentration (µg/g) data of 10 soil samples after sieving (vertical axis) plotted against the same samples when powdered (horizontal axis). Every dot represents one element from one sample. The inset (lower right) shows the elements with lower values in more detail. The dotted line in the middle represents the point where both treatments would give exactly the same results, and the dotted lines on both sides represent the results plus and minus 10%. [Color figure can be viewed at wileyonlinelibrary.com].]
Although there are some differences between the sieved and the powdered samples, the values of most of the 20 elements are similar. The differences make a direct comparison of data from sieved with powdered samples less reliable, but the similarity enables the comparison between sieved samples using these 20 elements. As sieving is less destructive than grinding to a powder, we chose to use sieving (250 micron mesh) for the other samples in this study.

Data Processing and Calculating Statistical Distance Measures

The 40 possible combinations of data processing and statistical distance measures were evaluated using the log-LR cost ($C_{llr}$) between the “same” and the “different” groups in the database with 50 different sites. In Table 1, the $C_{llr}$ per combination is given. Note that the Bray–Curtis distance measure was not suitable for combinations where the sum of compared elemental compositions was not equal.

The best performing (lowest $C_{llr}$, 0.027) combination of data processing method and statistical distance measure is taking the square root of the concentration values followed by converting into relative data combined with the Canberra statistical distance measure.

The distances calculated between elemental data from the “same” group and the distances calculated between elemental data from the “different” group were plotted as a frequency histogram for the Canberra distance measure (Fig. 3).

The Canberra frequency histogram (Fig. 3) can be used as a discriminative model in soil comparisons to assign two samples to the “same” group when the calculated CNB distance is lower than 0.6 or to the “different” group when the CNB distance is higher than 1.1. There is an area of overlap of the two groups between the CNB distances of 0.6 and 1.1. In the area of overlap, the ratio between the bars can be used as an indication of the odds of the two samples having the same or a different

<table>
<thead>
<tr>
<th>Data Processing</th>
<th>Canberra</th>
<th>Chebyshev</th>
<th>Minkowski</th>
<th>Bray–Curtis</th>
<th>Euclidean</th>
</tr>
</thead>
<tbody>
<tr>
<td>No processing</td>
<td>0.109</td>
<td>0.264</td>
<td>0.162</td>
<td>–</td>
<td>0.221</td>
</tr>
<tr>
<td>/sd</td>
<td>0.109</td>
<td>0.167</td>
<td>0.074</td>
<td>0.107</td>
<td>0.147</td>
</tr>
<tr>
<td>log10</td>
<td>0.117</td>
<td>0.110</td>
<td>0.123</td>
<td>–</td>
<td>0.211</td>
</tr>
<tr>
<td>sqrt</td>
<td>0.106</td>
<td>0.140</td>
<td>0.134</td>
<td>–</td>
<td>0.130</td>
</tr>
<tr>
<td>/sd - norm</td>
<td>0.034</td>
<td>0.328</td>
<td>0.083</td>
<td>0.099</td>
<td>0.179</td>
</tr>
<tr>
<td>log10 - norm</td>
<td>0.062</td>
<td>0.116</td>
<td>0.057</td>
<td>0.045</td>
<td>0.043</td>
</tr>
<tr>
<td>sqrt - norm</td>
<td>0.027</td>
<td>0.071</td>
<td>0.040</td>
<td>0.040</td>
<td>0.051</td>
</tr>
</tbody>
</table>

In each cell, the $C_{llr}$ value is given for the combination of data processing (rows) and statistical distance (columns) using the same dataset, unless the data were not suitable for that combination (+). Data processing abbreviations: converting the data to relative data (norm), dividing the concentration value of each element by the standard deviation of that element in the dataset (/sd), the normal logarithm of the concentration values (log10), the square root of the concentration values (sqrt). Lowest value of $C_{llr}$ in bold.

Although there are some differences between the sieved and the powdered samples, the values of most of the 20 elements are similar. The differences make a direct comparison of data from sieved with powdered samples less reliable, but the similarity enables the comparison between sieved samples using these 20 elements. As sieving is less destructive than grinding to a powder, we chose to use sieving (250 micron mesh) for the other samples in this study.

The best performing (lowest $C_{llr}$, 0.027) combination of data processing method and statistical distance measure is taking the square root of the concentration values followed by converting into relative data combined with the Canberra statistical distance measure.

The distances calculated between elemental data from the “same” group and the distances calculated between elemental data from the “different” group were plotted as a frequency histogram for the Canberra distance measure (Fig. 3).

The Canberra frequency histogram (Fig. 3) can be used as a discriminative model in soil comparisons to assign two samples to the “same” group when the calculated CNB distance is lower than 0.6 or to the “different” group when the CNB distance is higher than 1.1. There is an area of overlap of the two groups between the CNB distances of 0.6 and 1.1. In the area of overlap, the ratio between the bars can be used as an indication of the odds of the two samples having the same or a different

![FIG. 3—Histogram with the frequency distribution of the Canberra distance measure. The percentage of observations (vertical axis) per interval of calculated distance measures (horizontal axis, only upper limit shown per category) with the “same” group in shaded bars ($N = 150$) and the “different” group in black bars ($N = 11,025$).

TABLE 1—Evaluating the combinations of data processing and statistical distance measures with $C_{llr}$ values.
source. This type of discriminative model fits the Bayesian approach, which is recommended by the European Network of Forensic Science Institutes (ENFSI) guidelines for reporting evaluative forensic evidence.

The actual area of overlap is limited, which is partly caused by the size of the database. Increasing the size of the database would add more “rare” outliers on both sides of the distribution curves which would increase the overlap. Alternatively, a model can be fitted over the data to model the edges of the distributions. Both are topics for further research.

Spatial Resolution

For the three sites, the averages, standard deviations, and minimum and maximum values of the CNB distances between the elemental data were calculated. The results are shown in Fig. 4.

At site 1, the back garden with three garden beds, potting soil, or compost was added to the soil and not well mixed. The CNB distances between the samples from the same spot are relatively large when compared to sites 2 and 3 and the database (Fig. 4). The added and unmixed potting soil increased the difference in elemental composition and the resulting CNB distance between samples from the exact same sampling point.

At site 2, a camping field with a similar soil on the entire site, there is a clear increase of CNB distance with increasing physical distance.

At site 3, the samples were taken alongside a road with similar soil and no samples were taken more than a few meters perpendicular to the road to where the agricultural fields started. CNB distances between samples close to each other, but perpendicular to the road, were larger than CNB distances between samples from a longer distance along the road. For instance, a sample from the bottom of the ditch was very different from the samples of the side of the ditch exceeding the “same” group distance limit of 1.10 of the database (Fig. 3). On the other hand, samples 50 m apart along the road still had a CNB distance of 0.71 falling within the area of overlap of the “same” and “different” groups (0.60 < CNB < 1.10). This shows greater differences in elemental composition of the soil due to depth, land use, or construction work than due to physical distance.

Unknown soil samples can be assigned to the “same” group of the database when the CNB distance is lower than 0.60. For site 1, this would include samples within 5 m. For site 2, these values can still be calculated for some samples taken up to 20 m away. For site 3, it would include samples within 30 m.

It is clear from the results of these three example sites that the optimal sampling strategy relates to the local soil variation, both on the surface and below, and is different for each site. Based on the results, for a site with similar soil, one sample per four square meters (2 by 2 m) is likely to be sufficient. In addition, specific spots in a site where either the soil is notably different (for instance because of added potting soil), or has been disturbed (for instance by digging) or where specific crime-related traces (for instance shoe prints) are visible, should also be sampled.

When sampling a site for a case, the variation of the soil within the site is unknown. Multiple samples must be taken from the crime scene, as defined by the crime scene investigators, to be able to determine the within-site variation afterward. This is especially important on sites with a large variation to prevent false negatives. In addition, samples must be taken outside of the crime scene to determine whether the crime scene can be distinguished from the neighboring area and if not, where the spatial limits of the method actually are. These spatial limits are defined by the spatial resolution of the method applied, and the spatial resolution could be improved if multiple methods are combined.

Effects of Transfer, Sampling, and Storage

The unrelated soil samples in the 12 test cases were taken from 30 cm up to 11 km from the shoe print. In all test cases,
the average CNB distance between the soil from the shoe and the shoe print was lower than between the unrelated soil sample and the shoe print (Table 2).

In the 12 test cases, more than one gram was sampled from the shoes. During preliminary testing, most soil samples of less than one gram could not be accurately measured. Although this minimum sample size is a limit for forensic application in casework where soil traces are often small, we expect that this technical issue can be improved upon, for example, using sample cups with a smaller diameter for the entire database.

In seven of the twelve test cases, the average CNB distance calculated between the soil from the shoe print and the soil from the shoe was less than 0.60 which coincides with the “same” group in the CNB database. In four test cases, the average CNB distance calculated between shoe and shoe print was in the area of overlap (0.60 < CNB < 1.10) and one was higher than 1.10.

None of the average CNB distances between the soil from the shoe and the unrelated sample were less than 0.60, two of 12 test cases were in the area of overlap, and the remaining 10 test cases had values higher than 1.10.

In a Bayesian framework, the likelihood ratio (LR or evidential value) is equal to the odds of the results under assumption of one hypothesis versus the odds of the results under another (nonoverlapping) hypothesis. It is interesting to note that most of the CNB distances have no data for both groups, and LRs can therefore not be calculated (Fig. 3). The apparent contrast between the desired good separation and the resulting small overlap for LR calculations is mainly caused by the sample size of both groups. As noted in the results section on data processing and calculating statistical distance measures, extending the overlap while keeping the majority of the data separated is a topic for further research.

The discriminative model does not take into account the rarity of the specific combination and quantity of the 20 selected elements and of the other elements. How to combine the presented quantitative method with a frequentist approach is also a topic for further research. The presented method can be used to compare a large number of samples quickly and can give an estimate of the evidential value. In casework, additional interpretation using the case-specific circumstances is still necessary.

Based on the results of the validation, it can be seen that, even without further instructions on sampling procedure, most test cases could be resolved at the laboratory. In all test cases, the results of the analysis weighed in favor of the scenario where the soil from the shoe was from the shoe print and not from the alternative (unrelated) site. Seven of the test cases have the alternative site within 30 m of the shoe print. This shows the method can also distinguish between smaller distances, provided proper reference samples are available and different soil types are present.

**Conclusions**

We present a method to compare soil samples using elemental concentration data. The preparation of the samples was adapted by sieving the sample instead of grinding it to a powder, which is much less destructive to characteristics of the sample, such as microfossils.

Several distance measures and data processing methods were tested. Taking the square root of the concentration values followed by a conversion into relative data and using the Canberra statistical distance measure was found to be the best performing combination.

As shown by the separation of the two groups, “same” sampling point and “different” sites, in the dataset and differences between related and unrelated shoeprints in the test cases, the method can distinguish between soil samples from a same sample point or from different sites with a high degree of accuracy. The test cases also indicate that the method is not very sensitive to sampling by different people, which is important in a forensic setting where many different people are involved in sample collection and transfer to the laboratory.

The examples of the spatial resolution within a site indicate that the optimal sampling strategy is different for each site, but one sample per four square meters (for sites with similar soil) is likely to be sufficient with additional sampling at forensically relevant places.

The frequency distribution histogram of the CNB distances can be used when comparing soil samples as a discriminative model to assign two samples to either the “same” or the “different” group. In this discriminative model, CNB distances lower than 0.60 are more likely to be calculated between samples from
the same sampling point than between samples from different sites (with a site defined as an area of 30 m across with the same soil type and land use). CNB distances higher than 1.10 on the other hand are more likely to be calculated between samples from different sites (more than 30 m apart or with different soil type or land use) than between samples from the same sampling point. These two CNB boundaries are influenced by the definition of “site” used for the two groups in the database, but could also be influenced by the laboratory equipment used to measure the samples for the database.

These results show that the presented method, including the database and the discriminative model, is a useful tool to compare multiple soil samples quickly for forensic casework.

Acknowledgments

The authors wish to thank M. Schrader for assisting in the ED-XRF measurements, J. de Zoete and M. Lopatka for the use of their LR-GUI (19) to test different statistical methods, and F. Quaak for scientific discussion and proofreading.

References


Additional information and reprint requests:
Stefan Uitdehaag, M.S.
Netherlands Forensic Institute
Laan van Ypenburg 6
2497 GB The Hague
The Netherlands
E-mail: s.uitdehaag@nfi.minvenj.nl