



THE ROLE OF STATISTICAL METHODS IN ASSESSING THE EVIDENTIAL VALUE OF PHYSICO-CHEMICAL DATA

Grzegorz ZADORA

Institute of Forensic Research, Krakow, Poland

Abstract

The usefulness of statistical approaches, especially those related to the likelihood ratio, in evaluating glass fragments is presented in this paper. The glass fragments were analysed and characterised according to their elemental content as determined by the SEM-EDX method.

Key words

Glass; Classification; Comparison; Forensic statistics; SEM-EDX.

Received 16 April 2006; accepted 15 May 2006

1. Introduction

It is universally known that various types of materials can be subjected to physico-chemical examinations by forensic scientists, e.g. glass fragments, paints, fibres, plastics, petroleum products, organic solvents, gunshot residues. Various analytical methods have been used for these examinations and various kinds of information have been obtained. These include:

- morphological data – e.g. the appearance of a cross-section of fibres, the number and thickness of layers in a cross-section of automotive paint;
- qualitative data – e.g. gas chromatograms of accelerants detected in a sample collected from a suspicious fire scene; chromatograms of pyrolysis products of a sample of paint or fibres; infrared spectrum of an unknown substance;
- quantitative data – e.g. the refractive index of a glass fragment, the concentration of elements in a glass fragment determined by a scanning electron microscope with an X-ray detector (SEM-EDX),

areas under the peaks of a chromatogram of gasoline detected in fire debris.

Questions raised by the police, prosecutors and the courts concern identification of evidence material and/or an association between two or more evidence materials. Therefore, data obtained during casework should be reliable. Data can be confirmed to be reliable when the particular analytical method is fully validated. One of the definitions of validation states that it is a process by which objective data are obtained through carrying out measurements, with the aim of achieving particular aims and meeting certain analytical requirements (i.e. solving specific analytical problems) [7]. Parameters such as repeatability, intermediate precision, reproducibility, and accuracy should be recorded during the validation process for a particular quantitative technique. By the same token, statistical parameters such as standard deviations and relative standard deviation should be determined and regression-correlation analysis as well as an analysis of variance should be carried out during determinations of the above-mentioned parameters. The use of statistical

tools in physico-chemical analysis of evidences should not be viewed as a passing fad, but a contemporary necessity for the validation of results.

The necessity and usefulness of application of statistical methods in the forensic sphere is not only associated with validation. Successful application of statistical methods in the identification-classification and comparison of glass fragments analysed by SEM-EDX is an example [1, 4, 12, 18, 20, 21]. This method has been applied to glass in the Institute of Forensic Research (IFR) since 1997.

2. Glass recovery and elemental analysis

Any glass fragments in clothing are removed by brushing the submitted garments. The collected debris is studied under an SMXX Carl Zeiss optical microscope at 100 magnification with the aim of isolating any glass-like fragments. Fragments recovered for analysis typically have linear dimensions of about 0.1–0.2 mm. Selected glass-like fragments are placed on a 12 mm aluminium SEM stub covered with self-adhesive carbon tape. Once mounted, they are carbon coated using a Swiss Bal-Tech SCD sputterer. The prepared stub is mounted in the sample chamber of a Jeol JSM-5800 scanning electron microscope equipped with a Link ISIS 300 energy dispersive X-ray spectrometer. A minimum of three measurements is taken for each glass fragment. The following measurements conditions were applied: accelerating voltages 20 kV, count time 50 seconds, magnification 1000–2000 times, calibration element, Co. The SEMQuant option was used to calculate the percentages by weight of selected elements in each sample. This is a part of the LINK ISIS software from Oxford Instruments Ltd, of the United Kingdom. This option applied the ZAF correction procedure, which makes corrections for the effects of differences in the atomic number (Z), absorption (A) and X-ray fluorescence (F). The selected analytical conditions allowed all elements to be determined except Li and B. The target elements and their concentrations considered in the research presented in this paper were oxygen, sodium, magnesium, aluminium, silicon, potassium, calcium and iron.

3. Identification of glass fragments

The SEM-EDX method is a useful tool for identification of glass particles among other glassy microtraces recovered from debris. Occasionally, certain sand grains can look like a glass particle, especially

when their linear dimensions are less than 0.2 mm. Therefore, it is necessary to identify glass fragments among other particles collected on an SEM stub. This can be accomplished by means of the analytical results obtained by SEM-EDX analysis. The identification is made on the basis of qualitative results readily discernible in the spectra. The differentiation between true glass and glass-like particles such as quartz grains can be accomplished very quickly due to the obvious and significant differences in elemental content (compare spectra in Figure 1 and 2).

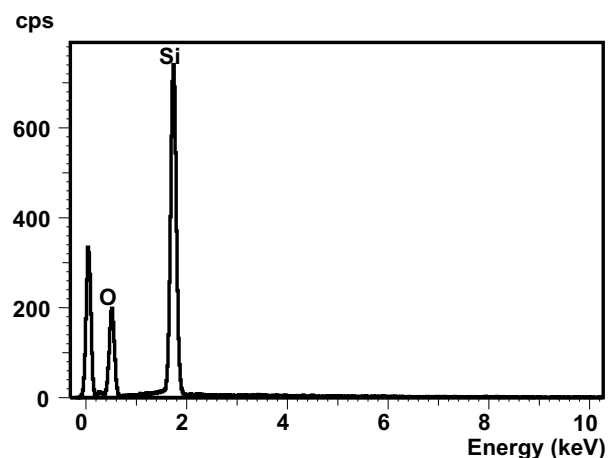


Fig. 1. A SEM-EDX spectrum of a sand particle (SiO_2).

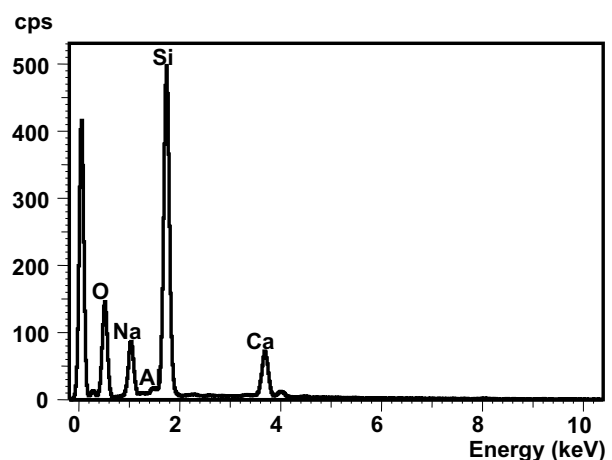


Fig. 2. A SEM-EDX spectrum of a glass fragment (from a colourless pane of float glass – $\text{SiO}_2\cdot\text{CaO}\cdot\text{Na}_2\text{O}\cdot\text{Al}_2\text{O}_3$).

4. Classification of glass fragments

When a glass microtrace is detected among microtraces recovered from debris and mounted on an SEM-stub, a classification should be made as to the type of glass present (e.g., windows, containers, headlamps). This process is especially important when there is

a lack of a comparison sample. Determination of the type of glass could help investigators focus their search for the appropriate comparison materials.

The preliminary classification scheme worked out and applied at the Institute of Forensic Research [18, 21] was made on the basis of analytical results from 210 samples. Inspection of results from these samples allowed identification of specific elements whose ranges of concentration in the considered groups did not overlap. This non-statistical approach only permitted a correct classification of glass samples originating from general types or categories having special physico-chemical features and therefore a special elemental content (See sets 2–5 in Figure 3). An example would be optical glass that contains relatively high concentrations of potassium and barium, present to improve its optical properties.

It should be pointed out that the SEM-EDX method only provides information about major and minor elements present in a glass sample. Trace elements are below the detection limits of this method. It is a common belief that information about trace elements present in a sample is indispensable for effectively comparing and individualising glass evidences. Consequently, the following question is posed: is it possible to solve problems of classification and comparison, on the basis of result of analyses performed by the SEM-EDX method in the absence of information regarding trace elements? The answer is no when a visual inspection of the quantitative data is the only approach. From such an inspection of Figure 3, it can be seen that most samples are located in sets 6 and 7. The classification problem can be solved, however, when cluster analysis (a chemometrics method) is used. This is a common

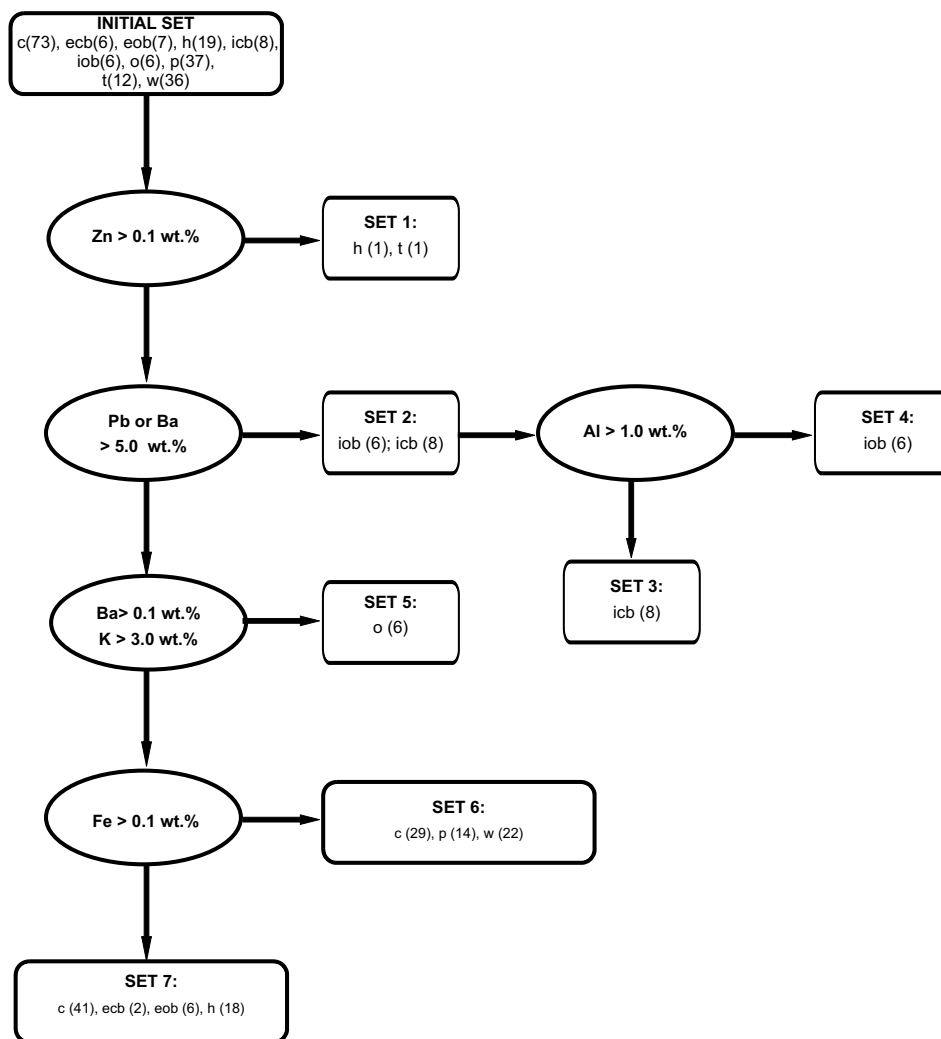


Fig. 3. A classification scheme of glass fragments on the basis of results of SEM-EDX elemental determination (c – car windows, ecb, eob, icb and iob – bulbs, h – headlamps, o – optical glass, p – containers, t – table glass, w – building windows). In brackets – the number of objects originating from a particular class.

technique for data visualisation and classification of materials when samples are described by more than one physico-chemical feature. This method allows us to create clusters out of similar objects. It relies upon a simple rule; the measure of similarity of studied objects is the distance between them. The smaller the distance between objects the more similar they are to each other. The distances between samples were calculated as the squared Euclidean distance and the smallest variance method was applied as a clustering technique. Results of cluster analysis can be presented in the form of dendrograms. A dendrogram resulting from cluster analysis within a set of samples of container glass and glass from automobile and building windows is presented in Figure 4. Two separate clusters are observed for these data.

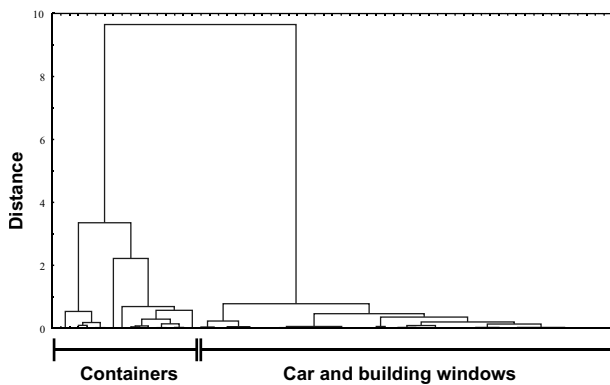


Fig. 4. A dendrogram resulting from cluster analysis within set 6 of container glass, automobile windows and ordinary window panes (similarity measure: squared Euclidean distance, clustering method: the smallest variance; Ward’s method).

Similar results were obtained when glass samples originating from headlamps, bulbs and containers were analysed (Figure 5).

The clusters obtained did not contain glass objects originating from the same type-use class of glass, i.e. they were not homogenous. Therefore, the results for the classification of unknown glass fragments could not be definitive.

It was mentioned that the cluster analysis method is a method of analysis of multivariate data structure. The task of classification of glass fragments for forensic purposes, where glass fragments are analysed by the SEM-EDX method, could be solved by application of methods such as Support Vector Machines or Naive Bayes Classifiers. The application of SVM can be especially successful in the case of classification of glass objects into one of the following glass categories: “car windows” or “building windows” [20].

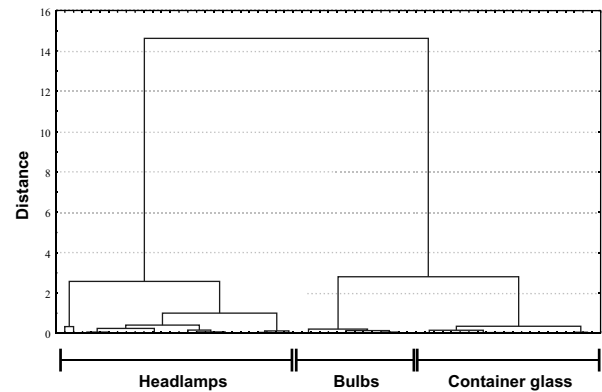


Fig. 5. A dendrogram resulting from cluster analysis within set 7 of car headlamps, bulbs and container glass (similarity measure: Czebyszew distance, clustering method: the smallest variance; Ward’s method).

5. Comparison of glass evidence

The statistical methods were evaluated in a classic comparison problem; could two samples (evidence and control) have originated from the same object? With the aim of solving this problem, the following model was proposed [18]. It was assumed that microtraces could be treated as points in multidimensional space because each sample is described by some elements determined by SEM-EDX. It was also assumed that microtraces originating from separate objects such as two window panes should create two separate sets in multidimensional space; otherwise, they should create one set. The following null hypothesis was considered:

$$H_0: d_i \approx d_e \quad \{1\}$$

where, d_i – the mean of distances calculated between microtraces originating from the same object (Figure 6), d_e – the mean of distances calculated between microtraces originating from different objects (Figure 7).

The following formula was used for distance calculation:

$$d_{A-B} = \sqrt{2 \left(1 - \frac{\sum_{k=1}^n c_{A_k} c_{B_k}}{\sqrt{\sum_{k=1}^n c_{A_k}^2} \sqrt{\sum_{k=1}^n c_{B_k}^2}} \right)} \quad \{2\}$$

where d_{A-B} – the distance between objects A and B , c_{A_k} , c_{B_k} – concentration (wt. %) of the k -th element in compared microtraces A and B , n – number of considered elements.

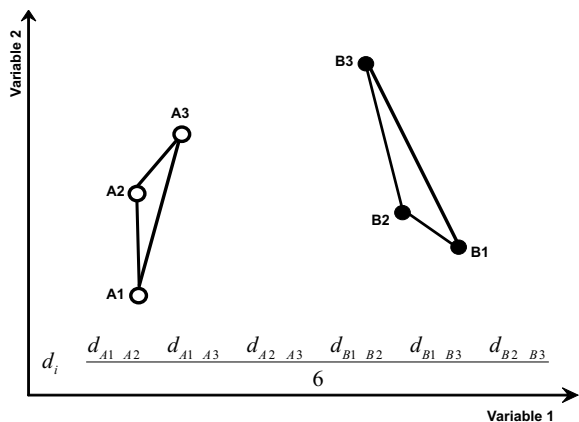


Fig. 6. Illustration of the concept of calculation of mean distance between microtraces originating from the same object; $d_{A1 A2}, d_{A1 A3}, \dots, d_{B2 B3}$ – distances calculated according to equation {2}.

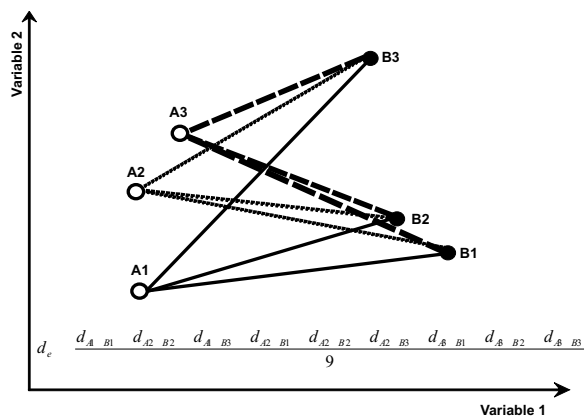


Fig. 7. Illustration of the concept of calculation of mean distance between microtraces originating from different objects; $d_{A1 B1}, d_{A1 B2}, \dots, d_{A3 B3}$ – distances calculated according to equation {2}.

The null hypothesis was tested by the t-Student test (Welch modification). This model worked sufficiently well in that the level of false answers was low when the significance probability p was set as equal to 0.3. This approach was tested through many collaborative tests organised by the Paint and Glass Experts Working Group of the European Network of Forensic Sciences Institutes in which comparisons were correctly answered. Nevertheless, there is a critical issue associated with this approach. This approach controls type I (false negative) error. But it is type II (false positive) errors that must be thoroughly controlled in the forensic field. This critical point relates to all situations where significance tests are used for comparison of evidence and control items.

Moreover, the application of significance tests only allows the analyst to take into account information

about the similarity of compared objects (e.g., differences in the measured elemental content of compared samples) and within-source variation (i.e., variations within the same glass object). Therefore, the answer to the following question could be obtained by the application of a significance test: Do the compared samples have the same physico-chemical features? However, an affirmative answer to this question does not answer the questions: Could the evidence and control sample originate from the same object/What is the strength of the evidence that the compared samples have the same origin?

The answer to these questions requires information regarding sources of error related to between-objects variability of the determined physico-chemical properties as well as information about the commonness or rarity of the determined for the control and evidence samples in the relevant population. Intuition would suggest that it is more probable that the evidence glass samples originated from the same object as the control sample when the matching physico-chemical values are rare in the relevant population than when they are common. This information is not taken into account in significance tests, but it can be through the application of models constructed within likelihood theory, namely the likelihood ratio (LR) approach. LR compares the probability of the measurements on the evidence assuming a common source for the evidence and control glass samples with the probability of the measurements on the evidence assuming different sources for the evidence and control samples. The law of likelihood states that [14]: If A implies that the probability that a random variable X takes the value x is $p_A(x)$, while hypothesis B implies that the probability is $p_B(x)$, then the observation $X = x$ is evidence supporting A over B if and only if $p_A(x) > p_B(x)$, and the likelihood ratio, $p_A(x)/p_B(x)$, measures the strength of that evidence. The letter E describes all data related to evidence, e.g. the elemental content of glass fragments. The likelihood ratio can be expressed in the form:

$$LR = \frac{Pr E | H_p}{Pr E | H_d} \quad \{3\}$$

In the case where the comparison concerns glass fragments for which physico-chemical properties have been determined, the following hypotheses could be considered:

H_p glass fragments recovered from the suspect's clothes are from the same source as the control sample; i.e. the prosecution hypothesis.

H_d glass fragments recovered from the suspect's clothes are not from the same source as the control sample; i.e., the defence hypothesis.

The value of the likelihood ratio is easy to interpret. Values of LR above 1 support the hypothesis considered in the numerator $\{3\}$ and values of LR below 1 support the opposite hypothesis, considered in the denominator $\{3\}$. LR equal to 1 does not support either. Also, a simple rule is applicable: the larger the LR value, the stronger the support for H_p , and vice versa for H_d .

In 1977 Lindley [11] proposed a formula for LR calculations in the case of univariate data when the assumption is that between-source variability is normally distributed. It is, however, unwise from a forensic standpoint to assume a normal distribution of data in relevant populations such as float glass. In the absence of a proven normal distribution of properties, the assumption must be dismissed and in its place, a kernel density estimate for the between-group distribution employed [6].

This univariate LR approach could be theoretically applied to evaluation of glass fragments described by several features. In this approach, LR should be calculated for each variable separately and then all obtained LR_i should be multiplied to obtain a total likelihood ratio LR_t ($LR_t = \prod_{i=1}^n LR_i$, n – number of considered variables). This would be appropriate when the assumption that all considered variables are independent is true. This was a critical point in the application of the LR univariate approach to multi-dimensional data obtained during SEM-EDX analysis because this assumption was not true.

In 2004, the appropriate LR model for multi-variate data was published [12]. This method take into account two sources of errors (variability); within- and between-source variability. This LR approach was ap-

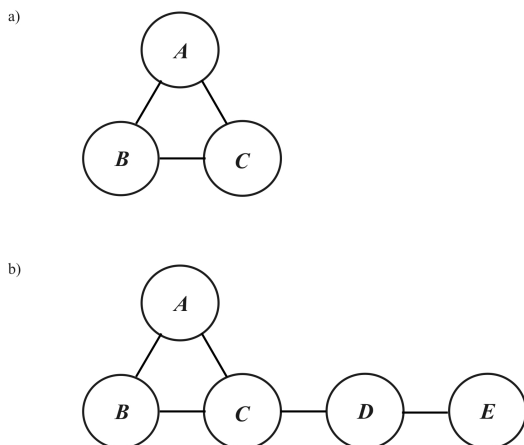


Fig. 8. A graphical model of data structure described by variables a) A, B, C and b) all five variables according to matrix $\{4\}$ when $|r| > 0.25$.

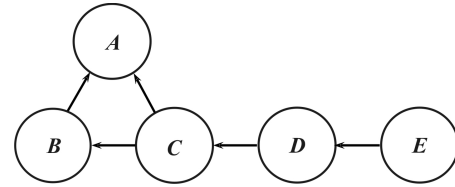


Fig. 9. A direct acyclic graph of data structure according to matrix $\{4\}$ when $|r| > 0.25$.

plied to data [12] where glass fragments were described by three variables. Glass fragments entered in the Institute of Forensic Research’s glass database are described by seven variables. The LR could be calculated by using all variables but this required an assumption that all variables were correlated. This assumption could be accepted, but it must be remembered that during any calculations of LR , an estimation of parameters such as means, variances and covariance should carried out. These estimations require a lot of analytical data, i.e. a large number of analysed glass samples; far more than were accessible in the IFR glass database.

Therefore, it was proposed to use graphical models for a reduction in the number of variables considered for LR calculation. The graphical model is derived from a consideration of a normalised inverse covariance matrix. As an illustration of graphical models, the following theoretical matrix for five variables is presented $\{4\}$.

	A	B	C	D	E	
A	1.00	0.86	0.31	0.19	0.15	
B	0.86	1.00	0.75	0.02	0.01	
C	0.31	0.75	1.00	0.97	0.05	$\{4\}$
D	0.19	0.02	0.97	1.00	0.26	
E	0.15	0.01	0.05	0.26	1.00	

This matrix includes the negatives of the partial correlation coefficients between the corresponding pairs of variables, given the remaining variables. The decision has to be made that a pair of variables may be considered conditionally dependent when the corresponding value of a partial correlation coefficient r is less than certain in absolute value, e.g. $|r| > 0.25$. The variables are represented by nodes. Pairs of variables (nodes) in which the module of partial correlation matrix is greater than r are joined by a straight line. For example, it can be seen in matrix $\{4\}$ that variable A is partially correlated to variables B and C . Variable B is also correlated to variable C as $|r| = 0.75$. The graphical representation of this is presented in Figure 8 a.

The variable C is also partially correlated to variable D ($|r| = 0.97$) and variable D is also partially correlated to variable E ($|r| = 0.26$) (see Figure 8 b). In the

next step arrows are added in such a way that the resulting graph is acyclic (Figure 9).

How can the information presented in the acyclic graph be used to calculate LR ? It is necessary to express a probability density function. This is possible by analysing the dependence of each variable on other variables, i.e. observation of arrows pointing to a particular node. For example, two arrows point to node B (Figure 9). This means that variable B is conditionally dependent on variables A and C . This could be represented by the following probability density function:

$$f_{B|A,C} = \frac{f_{A,B,C}}{f_{A,C}} \quad \{5\}$$

The joint probability function present in the nominator and denominator could be evaluated by the LR approach for multivariate data:

$$f_{B|A,C} = \frac{LR_{A,B,C}}{LR_{A,C}} \quad \{6\}$$

In a similar way for other nodes (Figure 9), the following dependences could be found:

$$f_{A|C} = \frac{LR_{A,C}}{LR_C}; \quad \{7\}$$

$$f_{C|D} = \frac{LR_{C,D}}{LR_D}; \quad \{8\}$$

$$f_{D|E} = \frac{LR_{D,E}}{LR_E}; \quad \{9\}$$

$$f_E = LR_E. \quad \{10\}$$

The multiplying of expressions {6}–{10} gives an expression for the calculation of total value of likelihood ratio (LR_t):

$$LR_t = \frac{LR_{A,B,C} LR_{C,D} LR_{D,E}}{LR_C LR_D} \quad \{11\}$$

One could easily observe that a maximum of three variables are used in the calculation of each partial LR_i , ($i = A, B, C; C, D; D, E; C, D$), but at the same time information about each of five variables is taken into account for LR_t calculations. Therefore, the use of a graphical model allows correct estimation of parameters such as mean, variance and covariance.

This approach was first tested on a set of 130 samples of glass (mainly float glass) described by five variables that had a multivariate normal distribution [1]. Moreover, a third source of variability was also considered in this paper, i.e. measurement error – the

variation of data obtained during analysis of the same glass fragment.

It was previously mentioned that it is nearly impossible to assume a multivariate normal distribution for between-source data. Therefore, a kernel density estimate for between-source data should be applied. A suitable model was worked out. Moreover, a two-level model, after application of a undirected graphical model, was tested on a set of 200 samples analysed by the SEM-EDX method. Details of this can be found in [4].

It should be mentioned that these approaches [1, 2, 4, 12], based on the calculation of a likelihood ratio and the application of graphical models to the evaluation of glass evidence in the form of univariate and multivariate data, can also be applied to the evaluation of other types of evidence when:

- a) they are characterised by continuous variables;
- b) population data are available that make it possible to estimate sources of variability and provide an estimate of the rarity of selected physico-chemical properties.

6. Conclusions

The application of statistical methods to the solving of various forensic problems related to glass evidence presented here provides an example of their potential usefulness in forensic science [2]. Methods based on likelihood theory and the calculation of likelihood ratios are very suitable for solving the problem of comparisons between evidence and control samples.

Mathematical descriptions of these models are generally quite complicated and sophisticated. The complexity of these models is probably one of the limitations on their common usage by forensic scientists. It is also essential that models applied to solving forensic problems should be comprehensible to non-statisticians – judges, prosecutors and members of a jury.

Recently an increasing number of papers have been observed describing approaches based on application of graphical models (e.g. Bayesian networks) for forensic purposes [3, 5, 8, 9, 13, 15, 16]. These models allow non-statisticians to see relationships between variables that describe evidentiary materials without the concurrent loss of their mathematical content.

References

1. Aitken C. G. G., Lucy D., Zadora G. [et al.], Evaluation of trace evidence for three-level multivariate data with the

- use of graphical models, *Computational Statistics and Data Analysis* 2006; 50, 2571–2588.
2. Aitken C. G. G., Taroni F., *Statistics and the evaluation of evidence for forensic scientists*, John Wiley and Sons, Chichester 2004.
 3. Aitken C. G. G., Taroni F., Garbolino P., A graphical modelling for the evaluation of cross-transfer evidence in DNA profiles, *Theoretical Population Biology*, 2003, 63, 179–190.
 4. Aitken C. G. G., Zadora G., Lucy D., A two-level model for evidence evaluation, *Journal of Forensic Sciences* 2007, 52, 412–419.
 5. Chan K. P. S., Aitken C. G. G., Estimation of the Bayes factor in a forensic science problem, *Journal of Statistical Computation and Solution* 1989, 33, 249–264.
 6. Curran J. M., Hicks T. N., Buckleton J. S., *Forensic interpretation of glass evidence*, CRC Press, Boca Raton 2000.
 7. Dawid P., Mortera J., Pascali V. L. [et al.], Probabilistic expert systems for forensic inference from genetic markers, *Scandinavian Journal of Statistics* 2002, 29, 577–595.
 8. Evett W., A Bayesian approach to the problem of interpreting glass evidence in forensic science casework, *Journal of the Forensic Science Society*, 1986, 26, 3–18.
 9. Evett W., Gill P. D., Jackson G. [et al.], Interpreting small quantities of DNA; the hierarchy of proposition and the use of Bayesian networks, *Journal of Forensic Sciences* 2002, 47, 520–530.
 10. Garbolino P., Evaluation of evidence using Bayesian networks, *Forensic Science International* 2002, 125, 149–155.
 11. Lindley D. V., A problem in forensic sciences, *Biometrika* 1977, 64, 207–213.
 12. Lucy D., The evaluation of trace evidence in the form of multivariate data, *Applied Statistics* 2004, 53, 109–122.
 13. Mortera J., Dawid A. P., Lauritzen S. L., Probabilistic experts system for DNA mixture profiling, *Theoretical Population Biology* 2003, 63, 191–205.
 14. Ryland R., *Statistical evidence*, Chapman and Hall, London 1997.
 15. Taroni F., Biedermann A., Garbolino P. [et al.], A general approach to Bayesian networks for the interpretation of evidence, *Forensic Science International* 2004, 139, 5–16.
 16. Taroni F., Garbolino P., Bayesian Networks and the evaluation of scientific evidence: a theoretical approach, *Problems of Forensic Sciences* 2001, 46, 173–179.
 17. www.enfsi.org.
 18. Zadora G., Brożek-Mucha Z., Parczewski A., A classification of glass microtraces, *Problems of Forensic Sciences* 2001, 47, 137–143.
 19. Zadora G., Brożek-Mucha Z., Parczewski A., A differentiation of glass microtraces of the same qualitative composition, *Problems of Forensic Sciences* 2001, 47, 144–146.
 20. Zadora G., Glass analysis for forensic purposes – a comparison of classification methods, *Journal of Chemometrics* 2007, 20, 1–13.
 21. Zadora G., Piekoszewski W., Parczewski A., An application of chosen similarity measurements of objects for forensic purposes, *Bulletin of the International Statistical Institute*, 54th session, 2003, volume LX, book 2, 364–367.

Corresponding author

Grzegorz Zadora
 Institute of Forensic Research
 ul. Westerplatte 9
 31-033 Krakow, Poland
 e-mail: gzadora@ies.krakow.pl

ROLA METOD STATYSTYCZNYCH W OSZACOWANIU WARTOŚCI DOWODOWEJ DANYCH FIZYKOCHEMICZNYCH

1. Wstęp

Powszechnie wiadomo, że przedmiotem badań fizykochemicznych wykonywanych przez biegłego sądowego mogą być różnego rodzaju materiały, takie jak okruszki szkła, lakier, włókna, tworzywo sztuczne, produkty ropopochodne, rozpuszczalniki organiczne lub ślady powystrzałowe. Dlatego też podczas wspomnianych badań stosuje się takie metody, które umożliwiają uzyskanie rzetelnych wyników dla obiektów dostępnych do analizy w bardzo małych ilościach. W trakcie badań uzyskuje się liczne informacje o analizowanych obiektach, które można podzielić na:

- cechy morfologiczne np. wygląd przekroju poprzecznego włókna; liczba i grubość powłok lakierowych pochodzących z pojazdu, które ustalono na podstawie obserwacji przy użyciu mikroskopu optycznego;
- dane jakościowe – chromatogramy uzyskane w trakcie analizy chromatograficznej próbki pobranej z miejsca pożaru na obecność lotnych łatwopalnych produktów ropopochodnych; chromatogramy uzyskane w trakcie analizy produktów pirolizy próbki lakieru lub włókna; widmo w podczerwieni uzyskane dla nieznannej substancji;
- dane ilościowe – np. wartość współczynnika załamania światła wyznaczona dla okruszki szkła; zawartość pierwiastków w analizowanym fragmencie szkła wyznaczona na podstawie analizy metodą SEM-EDX; pola powierzchni pików związków obecnych na chromatogramie uzyskanym w trakcie analizy próbek zabezpieczonych na miejscu pożaru.

Pytania zadawane przez zleceniodawców, tj. policję, prokuraturę lub sądy, dotyczą problemu identyfikacji (klasyfikacji) materiału dowodowego lub też ustalenia powiązań pomiędzy materiałem dowodowym a porównawczym. Dlatego też dane uzyskane w trakcie analizy materiału dowodowego powinny być rzetelne, a takie dane mogą być uzyskane wówczas, gdy stosowana metoda analityczna jest w pełni zwalidowana. Jedną z definicji walidacji mówi, że jest to proces, w którym poprzez wykonanie pomiarów uzyskiwane są obiektywne dane mające na celu osiągnięcie konkretnych celów i zaspokojenie wymagań analitycznych (tj. rozwiązanie konkretnych problemów analitycznych) [7]. Parametry takie, jak powtarzalność, odtwarzalność, precyzja i dokładność, powinny być udokumentowane w procesie walidacji danej metody ilościowej. Tym samym parametry statystyczne, np. odchylenie standardowe, względne odchylenie standardowe, jak też analiza regresji i korelacji oraz analiza wariancji, muszą być wykonane w celu spełnienia

powyższych cech metody, które są wymagane w procesie walidacji. Należy dodać, że zastosowanie metod statystycznych powinno być obecnie traktowane nie jak przemijająca moda, ale konieczność, zapewniająca procesowi walidacji rzetelność.

Konieczność i użyteczność zastosowania tych metod w naukach sądowych nie jest związana wyłącznie z procesem walidacji. Na przykład użycie ich w celu rozwiązania problemów klasyfikacji i porównania okruszków szkła dla potrzeb wymiaru sprawiedliwości można znaleźć w licznych publikacjach [1, 4, 12, 18, 20, 21]. Opisane tam metody stosowane są w analizie okruszków szkła w Instytucie Ekspertyz Sądowych w Krakowie.

2. Odzyskiwanie okruszków szkła

Okruszkami szkła obecne na odzieży dostarczonej do badań mogą być odzyskane poprzez zastosowanie procesu szczotkowania. Uzyskane wówczas otrzępiny są analizowane przy stukrotnym powiększeniu za pomocą mikroskopu optycznego SMXX Carl Zeiss w celu wyizolowania okruszków szkła. Większość odzyskanych okruszków szkła ma rozmiar liniowy pomiędzy 0,1 a 0,2 mm. Wyselekcjonowane okruszki szkła są umieszczane na samoprzylepnych krążkach węglowych, które z kolei naklejane są na 12 mm aluminiowe stoliki stosowane w skaningowym mikroskopie elektronowym (SEM). Następnie powierzchnia okruszków szkła pokrywana jest cienką warstwą węgla za pomocą procesu napyłania z zastosowaniem napyłarki firmy Bal-Tech SCD (Lichtenstein). Tak przygotowana próbka montowana jest w komorze mikroskopu skaningowego Jeol JSM-5800 (Japonia) sprzężonego ze spektrometrem promieniowania rentgenowskiego Link ISIS 300 (Wielka Brytania). Wykonywane są minimum trzy pomiary każdego z analizowanych fragmentów szkła w następujących warunkach pomiarowych: napięcie przyspieszające 20 kV, czas pomiaru (dokonywanie zliczeń) 50 sekund, powiększenie 1000 2000 razy, pierwiastek kalibracyjny Co. Opcja SEMQuant, która jest częścią programu LINK ISIS (Oxford Instruments Ltd, Wielka Brytania), zastosowana została w celu określenia zawartości pierwiastków obecnych w widmie analizowanej próbki, a wyrażonej w procentach wagowych. Opcja ta stosuje procedurę korekcyjną ZAF, która uwzględnia poprawki liczby atomowej (Z), absorpcji (A) i fluorescencji promieniowania rentgenowskiego (F). Stosowane warunki pomiarowe pozwalają na oznaczenie wszystkich pierwiastków za wyjątkiem Li i B. Kluczowymi pierwiastkami w przypadku

analizy okruszków szkła są jednak tlen, sód, magnez, aluminium, krzem, potas, wapń i żelazo.

3. Identyfikacja szkła wśród innych próbek

Metoda SEM-EDX jest użytecznym narzędziem identyfikacji okruszków szkła spośród innych okruszków wyglądających jak szkło. Zdarza się bowiem, że ziarna piasku mogą przypominać okruszki szkła, szczególnie, gdy ich rozmiar liniowy jest mniejszy niż 0,2 mm. Dlatego też konieczne jest, aby odróżnić okruszki szkła od ziaren piasku. Identyfikacja taka jest dokonywana na podstawie rezultatów analizy jakościowej, tj. analizy widma SEM-EDX. Na pytanie, czy analizowany okruszek to fragment szkła, czy też okruszek wyglądający jak szkło, można odpowiedzieć bardzo szybko, ponieważ widmo szkła znacząco różni się od widma piasku (porównaj widma na rycinach 1 i 2).

4. Klasyfikacja okruszków szkła

W przypadku, gdy okruszki szkła znajdują się wśród tych okruszków, które ujawniono w otrzepinach i następnie umieszczono na stoliku SEM, to wówczas często konieczne jest przeprowadzenie ich klasyfikacji, tj. przypisanie ich do jakiejś kategorii użytkowej szkła (np. szkło okienne, pochodzące z opakowania, klosz reflektora samochodowego). Ten proces jest szczególnie istotny w przypadku braku materiału porównawczego, ponieważ określenie rodzaju szkła może pomóc osobie prowadzącej dochodzenie w ustaleniu, jakiego materiału porównawczego powinna szukać.

Pierwszy schemat klasyfikacji szkła opracowany w Instytucie Ekspertyz Sądowych [18, 21] powstał na podstawie rezultatów analizy próbek szkła z 210 różnych obiektów. Analiza wizualna danych pozwoliła wyróżnić te pierwiastki, których przedziały zawartości w rozpatrywanych kategoriach użytkowych szkła nie pokrywały się. Takie niestatystyczne podejście pozwoliło na poprawną klasyfikację próbek szkła pochodzących z tych kategorii szkła, które posiadają specjalne właściwości fizykochemiczne i z tego powodu specjalny skład chemiczny (zbiory 2-5 na rycinie 3). Jako przykład można podać szkło optyczne zawierające relatywnie dużą zawartość potasu i baru, czyli pierwiastków poprawiających właściwości optyczne szkła.

Należy nadmienić, że metoda SEM-EDX dostarcza informacji tylko o pierwiastkach występujących w stosunkowo dużej ilości w szkłe (tzw. główne i średnie składniki). Pierwiastki obecne w ilościach śladowych znajdują się poniżej poziomu detekcji tej metody. Istnieje powszechne przekonanie, że informacja o zawartości pierwiastków śladowych w próbce szkła jest niezbędna w celu

efektywnego rozwiązania wspomnianych wcześniej problemów analizy szkła dla potrzeb wymiaru sprawiedliwości. Czy jest jednak możliwe rozwiązanie problemu klasyfikacji i porównania próbek szkła na podstawie rezultatów analizy SEM-EDX, gdy brak jest informacji o pierwiastkach śladowych? Odpowiedź jest negatywna, gdy przeprowadzana jest tylko wizualna analiza danych. Takie postępowanie pozwala wyłącznie ułożyć większość próbek w zbiorach 6 i 7. W rozwiązaniu problemu klasyfikacji pomocne okazało się zastosowanie metody chemometrycznej, jaką jest analiza skupień. Metoda ta jest często stosowaną metodą do wizualizacji danych i klasyfikacji obiektów w przypadku, gdy opisane są przez więcej niż jedną cechę fizykochemiczną. Metoda ta pozwala grupować próbki szkła w zbiory obiektów podobnych. Opiera się ona na prostej zasadzie, że miarą podobieństwa dwóch próbek jest odległość pomiędzy nimi. Im mniejsza odległość, tym dwa obiekty bardziej są do siebie podobne. W przypadku przeprowadzonych badań odległość pomiędzy nimi wyrażona była jako kwadrat odległości euklidesowej, a jako metodę budowania skupień zastosowano metodę najmniejszej wariancji. Rezultaty analizy skupień najczęściej prezentowane są w formie dendrogramów. Na rycinie 4 prezentowany jest dendrogram uzyskany w trakcie wykonanej analizy skupień dla próbek pochodzących z opakowań szklanych, szyb samochodowych i szyb okiennych, na którym widoczne są dwa oddzielne skupienia. Podobne rezultaty uzyskano, kiedy analizowano próbki pochodzące z kloszy reflektorów samochodowych, żarówek i opakowań szklanych.

Jak łatwo zauważyć, poszczególne skupienia nie zawierają, tylko i wyłącznie, obiektów z jednej klasy użytkowej szkła, tzn. nie są one jednorodne. Dlatego też zaklasyfikowanie nieznannej próbki szkła na podstawie tej metody nie może być katagoryczne.

Jak wspomniano, metoda analizy skupień jest metodą chemometryczną służącą analizie struktury zbiorów danych ustalonych dla obiektów opisanych przez więcej niż jedną zmienną. W celu rozwiązania problemu klasyfikacji okruszków szkła na podstawie danych uzyskanych za pomocą metody SEM-EDX, można zastosować metody klasyfikacji obiektów, takie jak klasyfikatory maksymalnodległościowe lub też naiwne klasyfikatory bayesowskie. Zastosowanie szczególnie tej pierwszej metody w przypadku klasyfikowania obiektów do kategorii „szyby samochodowe” lub „szyby okienne” jest bardzo skuteczne [20].

5. Porównywanie okruszków szkła

Metody statystyczne są bardzo użyteczne w procesie porównywania próbek szkła i ustalenia, czy dwie próbki (dowodowa i kontrolna) mogą pochodzić z tego samego obiektu. W celu udzielenia odpowiedzi na to pytanie,

zaproprowany został opisany niżej tok postępowania [18]. Jak już wspomniano, okruchy szkła można traktować jako punkty w przestrzeni wielowymiarowej, ponieważ każda z próbek może być opisana przez zawartość kilku pierwiastków, które oznaczono techniką SEM-EDX. Ponadto mikroślady pochodzące z dwóch różnych obiektów szkła powinny tworzyć dwa oddzielne zbiory w przestrzeni wielowymiarowej, zaś w przeciwnym przypadku jeden zbiór. Dlatego też została rozważona następująca hipoteza zerowa:

$$H_0: d_i = d_e, \quad \{1\}$$

gdzie d_i to obliczona średnia odległość w przestrzeni wielowymiarowej pomiędzy próbkami pochodzącymi z tego samego obiektu (rycina 6), a d_e to średnia odległość w przestrzeni wielowymiarowej obliczona pomiędzy próbkami pochodzącymi z różnych obiektów (rycina 7). W celu obliczania odległości została zastosowana następująca zależność:

$$d_{A B} = \sqrt{2 \frac{1}{n} \frac{\sum_{k=1}^n c_{A_k} c_{B_k}}{\sqrt{\frac{\sum_{k=1}^n c_{A_k}^2}{n} \frac{\sum_{k=1}^n c_{B_k}^2}{n}}}}, \quad \{2\}$$

gdzie $d_{A B}$ to odległość pomiędzy obiektami A i B , c_{A_k} , c_{B_k} – zawartość (% wag.) k -tego pierwiastka w porównywanych próbkach szkła A i B , n – liczba rozważanych pierwiastków.

Hipotezę zerową testowano za pomocą testu t-Studenta (modyfikacja Welch). Model ten działał wystarczająco poprawnie, tzn. uzyskano w nim niewielki procent odpowiedzi fałszywie pozytywnych i fałszywie negatywnych przy założonym poziomie istotności = 0,3. Poprawność modelu testowano i potwierdzono w oparciu o różnego rodzaju testy przeprowadzone przez Grupę Roboczą ds. Badania Lakierów i Szkła, która działa w ramach Europejskiej Sieci Instytutów Nauk Sądowych. Niemniej jednak model ten ma przynajmniej jedną wadę. Podejście to kontroluje bowiem poziom błędów I rodzaju (odpowiedzi fałszywie negatywne), a jak wiadomo, w naukach sądowych bardziej istotne jest to, aby był kontrolowany poziom błędów II rodzaju (odpowiedzi fałszywie pozytywnych). Ta niedogodność dotyczy zastosowania wszystkich testów istotności dla potrzeb wymiaru sprawiedliwości.

Ponadto zastosowanie testu istotności pozwala jedynie na uwzględnienie informacji o podobieństwie porównywanych obiektów poprzez określenie różnic zawartości poszczególnych pierwiastków w próbkach oraz zmienności analizowanych cech fizykochemicznych w obrębie pojedynczego obiektu szklanego. Dlatego też wyłącznie odpowiedź na następujące pytanie: „czy próbki mają takie same właściwości fizykochemiczne?” może być udzielona po zastosowaniu testów istotności. Niemniej jednak

nawet odpowiedź pozytywna nie jest rozwiązaniem problemu z punktu widzenia wymiaru sprawiedliwości najważniejszego, tj. „jaka jest siła dowodu, że porównywane próbki mają to samo pochodzenie?”

Odpowiedź wymaga uwzględnienia informacji odnoszących się do źródła błędów związanych ze zmiennością wartości fizykochemicznych w różnych obiektach szklanych, jak też informacji odnośnie do rzadkości oznaczanych właściwości fizykochemicznych dla próbki porównawczej i kontrolnej w populacji odniesienia. Intuicja sugeruje, że jest bardziej wiarygodne, że próbki dowodowe szkła pochodzą z tego samego obiektu co próbka kontrolna w przypadku, gdy właściwości fizykochemiczne oznaczone dla porównywanych obiektów są rzadkie w populacji generalnej niż wtedy, gdy są one często spotykane. Takie informacje nie są uwzględniane w testach istotności, ale mogą być wzięte pod uwagę w modelu opartym na ideach pochodzących z teorii wiarygodności, a konkretnie w oparciu o iloraz wiarygodności (LR). Iloraz wiarygodności LR porównuje prawdopodobieństwo warunkowe, że zaobserwowane podobieństwo wartości cech fizykochemicznych oznaczane dla próbki dowodowej i próbki porównawczej szkła powstało w przypadku, gdy próbki te pochodzą z tego samego obiektu (tzw. hipoteza prokuratury) oraz prawdopodobieństwo, że taka sytuacja miała miejsce w przypadku, gdy próbki te pochodzą z dwóch różnych obiektów (tzw. hipoteza obrony). Definicja ilorazu wiarygodności jest następująca: Jeżeli prawdopodobieństwo, że zmienna losowa X przyjmie wartość x wynosi $p_A(x)$ przy założonej hipotezie A , a $p_B(x)$ przy założonej hipotezie B , to zaobserwowana wartość $X = x$ wspiera zdarzenie A w przeciwieństwie do zdarzenia B tylko i tylko wtedy, gdy $p_A(x) > p_B(x)$, a iloraz wiarygodności $p_A(x)/p_B(x)$ jest miarą siły takiego dowodu. Iloraz wiarygodności można wyrazić w postaci:

$$LR = \frac{Pr \{E | H_p\}}{Pr \{E | H_d\}}. \quad \{3\}$$

W przypadku, gdy problem porównawczy dotyczy fragmentów szkła, dla których oznaczono ich właściwości fizykochemiczne, to rozważane są następujące hipotezy:

H_p fragmenty szkła ujawnione na odzieży podejrzanego i próbka kontrolna pochodzą z tego samego obiektu; tzw. hipoteza prokuratora;

H_d fragmenty szkła ujawnione na odzieży podejrzanego i próbka kontrolna pochodzą z różnych obiektów; tzw. hipoteza obrony;

E dowód; informacje odstępne o materiale dowodowym, np. skład pierwiastkowy szkła.

Wartość ilorazu wiarygodności jest łatwa do interpretacji, ponieważ wartość LR większa niż 1 wspiera hipotezę rozważaną w liczniku $\{3\}$, a wartość LR mniejsza niż 1 wspiera hipotezę rozważaną w mianowniku $\{3\}$.

LR równy 1 nie wspiera żadnej z nich. Ma tutaj również zastosowanie inna prosta zasada: im większa wartość LR , tym silniejsze wsparcie dla H_p i *vice versa* dla H_d .

W 1977 roku Lindley [11] zaproponował model do obliczania wartości LR w przypadku, gdy porównywane próbki opisane są przez jedną zmienną i gdy można założyć, że rozkład tej cechy w populacji generalnej (np. w różnych obiektach szklanych) jest rozkładem normalnym, chociaż, jak wiadomo, założenie to nie jest poprawne w przypadku obiektów szklanych. Dlatego też bardziej wskazane jest zastosowanie innego modelu. Do modelowania rozkładu danej cechy w populacji generalnej zastosować można bowiem podejście, w którym rozkład prawdopodobieństwa analizowanej cechy wyznaczony jest poprzez zastosowanie metody estymacji jądrowej [5].

Taki jednowymiarowy model do obliczania wartości LR teoretycznie może zostać zastosowany w celu oszacowania wartości dowodowej okruszków szkła opisanych przez kilka cech. W takim przypadku LR należy obliczyć oddzielnie dla każdej zmiennej, a następnie tak uzyskane wartości LR_i pomnożyć w celu uzyskania ostatecznej wartości ilorazu wiarygodności (LR_i) ($LR_i = \prod_{i=1}^n LR_i$,

n – liczba rozważanych zmiennych). Takie podejście możliwe jest do zastosowania wówczas, gdy wszystkie rozważane zmienne są od siebie niezależne. To założenie nie jest z reguły prawdziwe w przypadku oszacowywania wartości dowodowej próbek szkła opisanych przez wiele zmiennych oznaczonych za pomocą metody SEM-EDX.

W 2004 roku opublikowana została informacja o modelu LR służącym do analizy danych wielowymiarowych [12]. Model ten uwzględnia dwa źródła zmienności, tj. zmienność wewnątrz- i pomiędzyobiektoową i pierwotnie był stosowany w przypadku, gdy próbki szkła opisywały 3 zmienne [12]. Próbki szkła znajdujące się w bazie danych Instytutu Ekspertyz Sądowych są opisane przez siedem zmiennych. Wspomniany model LR może zostać zastosowany i w tym przypadku, ale gdy założymy się, że wszystkie zmienne są zależne. Założenie to można zaakceptować, jakkolwiek należy pamiętać, że podczas obliczania wartości LR w którymkolwiek ze wspomnianych modeli konieczne jest rzetelne oszacowanie parametrów występujących w modelu, tzn. średnich, wariancji i kowariancji. Prawidłowe określenie tych wartości wymaga dużej liczby danych analitycznych, więcej, niż zebrano w bazie danych Instytutu Ekspertyz Sądowych.

Dlatego też zaproponowano zastosowanie modelowania graficznego w celu redukcji liczby zmiennych w stosowanym modelu LR . Model graficzny (graf) uzyskiwany jest na podstawie analizy znormalizowanej odwrotnej macierzy kowariancji. W celu ilustracji tego podejścia rozważmy następującą teoretyczną macierz odnoszącą się do 5 zmiennych {4}.

	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	
<i>A</i>	1.00	0.86	0.31	0.19	0.15	
<i>B</i>	0.86	1.00	0.75	0.02	0.01	
<i>C</i>	0.31	0.75	1.00	0.97	0.05	{4}
<i>D</i>	0.19	0.02	0.97	1.00	0.26	
<i>E</i>	0.15	0.01	0.05	0.26	1.00	

Macierz ta zawiera ujemne wartości współczynników korelacji cząstkowej. Na podstawie analizy należy podjąć decyzję, powyżej jakiej wartości współczynnika korelacji cząstkowej r możemy uważać, że dwie zmienne są ze sobą istotnie skorelowane, np. $|r| > 0.25$. W modelu graficznym zmienne są reprezentowane przez węzły. Pary zmiennych (węzłów), dla których wartość absolutna współczynnika korelacji cząstkowej jest większa niż r , są ze sobą łączone linią prostą. Na przykład w macierzy {4} można zaobserwować, że zmienna A jest powiązana ze zmiennymi B i C . Zmienna B jest również powiązana ze zmienną C , ponieważ dla tej pary $|r| = 0.75$. Graf uzyskany na podstawie tak poczynionych obserwacji jest prezentowany na rycinie 8 a.

Zmienna C jest również skorelowana ze zmienną D ($|r| = 0.97$) a zmienna D jest natomiast skorelowana ze zmienną E ($|r| = 0.26$) (patrz rycina 8 b). W kolejnym etapie analizy strzałki są dodawane w taki sposób, aby ostateczny graf nie zawierał cykli, a zmienne (węzły) powiązane były co najwyżej w trójkąty (rycina 9).

Jak zastosować informację zawartą w acyklicznym grafie do obliczenia wartości LR ? W tym celu na podstawie analizy uzyskanego grafu konieczne jest wyrażenie odpowiednich funkcji gęstości prawdopodobieństwa. Jest to możliwe dzięki analizie zależności występujących pomiędzy zmiennymi, tj. obserwacji strzałek na grafie. Na przykład dwie strzałki wskazują węzeł B (rycina 9). Oznacza to, że zmienna B jest warunkowo zależna od zmiennej A i C , co za pomocą funkcji gęstości prawdopodobieństwa można wyrazić w następujący sposób:

$$f_{B|A,C} = \frac{f_{A,B,C}}{f_{A,C}}. \quad \{5\}$$

Funkcje gęstości prawdopodobieństwa obecne w mianowniku i liczniku mogą być z kolei określone przez odpowiednie wyrażenia stosowane do obliczania LR :

$$f_{B|A,C} = \frac{LR_{A,B,C}}{LR_{AC}}. \quad \{6\}$$

Podobnie można postąpić w przypadku innych węzłów (rycina 9):

$$f_{A|C} = \frac{LR_{AC}}{LR_C}; \quad \{7\}$$

$$f C | D \frac{LR_{C,D}}{LR_D}; \quad \{8\}$$

$$f D | E \frac{LR_{D,E}}{LR_E}; \quad \{9\}$$

$$f E \quad LR_E. \quad \{10\}$$

Mnożąc wyrażenia {6} {10}, uzyskamy wyrażenie pozwalające na obliczenie końcowej wartości LR oszacowującej wartość dowodową porównywanych okruchów szkła (LR_t):

$$LR_t \frac{LR_{A,B,C} LR_{C,D} LR_{D,E}}{LR_C LR_D}. \quad \{11\}$$

Jak łatwo zaobserwować, w celu obliczenia każdego z cząstkowych LR_i są stosowane maksymalnie trzy zmienne ($i = A, B, C; C, D; D, E; C, D$), ale jednocześnie w obliczaniu LR_t jest uwzględniana informacja o każdej z pięciu zmiennych. Tym samym, zastosowanie modeli graficznych pozwala na rzetelne oszacowanie parametrów takich jak średnia, wariancja oraz kowariancja.

Metoda ta to została najpierw przetestowana na zbiorze 130 próbek szkła (głównie typu „float”) opisanych przez pięć zmiennych, dla których założono, że mają wielowymiarowy rozkład normalny [1] i w którym były rozważane trzy źródła zmienności, tj. błąd pomiaru pojedynczego okruchu szkła, zmienność pięciu cech fizykochemicznych wewnątrz pojedynczego obiektu szklanego, a także pomiędzy obiektami szklanymi zgromadzonymi w bazie danych.

Jak już wcześniej wspomniano, jest prawie niemożliwe, aby założenie, że zmienność międzyobiektoowa posiada wielowymiarowy rozkład normalny, była prawdziwa. Dlatego też metoda estymacji jądrowej do modelowania wielowymiarowego rozkładu gęstości prawdopodobieństwa została zastosowana w innym modelu uwzględniającym tylko dwa ostatnie źródła zmienności. W tym przypadku model został opracowany również w oparciu o analizę graficzną zależności pomiędzy zmiennymi, ale z zastosowaniem nieukierunkowanych modeli graficznych. Był on z powodzeniem testowany na zbiorze 200 próbek szkła badanych przy pomocy metody SEM-EDX [4].

Należy nadmienić, że proponowane podejścia [1, 2, 4, 12] opracowane do obliczania ilorazu wiarygodności i w oparciu o zastosowanie modeli graficznych, a służące do oszacowania dowodów w postaci okruchów szklanych, które opisane są przez jedną bądź więcej zmiennych, mogą być również zastosowane do innego rodzaju materiału dowodowego w przypadku, gdy:

- są one charakteryzowane przez zmienne ciągłe;
- dostępne są dane populacyjne, które umożliwiają oszacowanie zmienności wewnątrz- i pomiędzy-

obiektowej, jak też oszacowanie rzadkości wybranych do analizy cech fizykochemicznych.

6. Podsumowanie

Przedstawione tutaj zastosowania różnych metod statystycznych do rozwiązania problemów analizy szkła dla potrzeb wymiaru sprawiedliwości dostarczają przykładów ich przydatności w naukach sądowych [2]. Metody oparte na obliczaniu ilorazu wiarygodności są bardzo użyteczne w rozwiązywaniu problemów związanych z porównywaniem próbek dowodowych i kontrolnych.

Formalizm matematyczny tych modeli jest raczej skomplikowany i jest to zapewne największym ograniczeniem sprawiającym, że nie są one powszechnie stosowane przez biegłych sądowych. Należy zaznaczyć, że wskazane jest, by model stosowany w celu rozwiązywania problemów w naukach sądowych był stosunkowo łatwo zrozumiały dla osób nie będących statystykami, a więc sędziów, prokuratorów i ławników.

Obecnie można zaobserwować wzrastającą liczbę publikacji opisujących zastosowanie w naukach sądowych technik analizy danych wielowymiarowych w oparciu o metody graficzne (np. sieci bayesowskie) [3, 5, 8, 9, 13, 15, 16]. Modele te pozwalają osobom nie będącym statystykami zaobserwować zależności występujące pomiędzy zmiennymi opisującymi próbki dowodowe bez jednoczesnej utraty kontekstu matematycznego opracowanego modelu.