Forensic analysis of glass fragments as evidence in a robbery case

Aditya Ranade

December 10, 2021

1 Introduction

A robbery has been reported at a house near Iowa State University. The glass window in the living room was found to be broken from outside to gain entry into the house. On conducting a preliminary search, the police found CCTV surveillance video from nearby houses which indicates a person gained entry into the house by breaking the window.

On searching near the house, police identified a person with characteristics similar to the person seen in the surveillance videos and declared this person (referred to as XYZ) as their prime suspect. Upon close inspection, glass particles were found on the coat of the prime suspect and the police suspect these glass particles to be from the window broken from the house where robbery took place.

The police forensics team collected the glass fragments from the crime scene and from the suspect's coat. Due to the urgent nature of the case and the time crunch, the forensics team could take only three replicate measurements of trace elemental concentrations of 18 elements on both sets of the glass using LA-ICP-MS which can be found in table 2 (glass found at the house) and table 3 (glass found on the suspect's coat)

The prosecutor believes the glass particles found on the suspect XYZ's coat are from the broken window pane from the house where robbery took place. A statistician on the forensics team helps with the analysis.

2 Data

The data used in this project is obtained from the glass element data set available through CSAFE open source data website. The data of trace elemental concentrations of 18 elements on both sets of the glass using LA-ICP-MS can be found in table 2 and table 3 (for the measurements of glass found at the crime scene and the suspect's coat respectively).

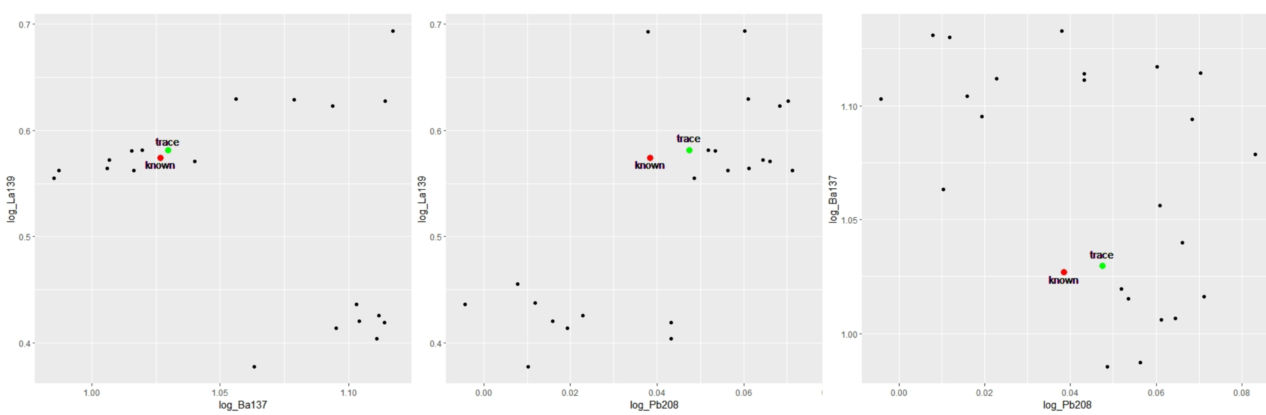

Figure 1: Pairwise plots of the mean elemental concentrations

Figure 1 shows the pairwise plots of the mean elemental concentrations for combination of 3 variables. Black colored points are glass fragments from the background dataset whereas red and green colored points are from the known fragment and trace fragment respectively.

The list of the elements whose measurements were used in this analysis can be found in table 4 in appendix A. Table 4 also indicates if the elemental concentration data (variables) satisfies normality (using Shapiro Wilk Test) . Since the data considered for analysis has 10 variables and we have only 3 measurements each for the glass found at the crime scene and the glass found on the suspect to estimate the 10 parameters, it was decided to employ Principal Component Analysis (PCA) to reduce the dimensions. We also have access to the past historic data of glass particles (measured on various instruments including LA-ICP-MS) collected over time which will be used in one of the methods (ABC Bayes factor approach).

3 Methods

We try to apply multiple techniques to the data since we have only three measurement replications for the glass particles found at the crime scene (house where robbery took place) and the glass particles found on the suspect's coat.

3.1 Proposition

We consider the specific source source level proposition. The prosecution proposition (H_p) and the defence proposition (H_d) are as follows

H_p : The fragment found from the suspect's coat is from the smashed window from the house where robbery took place.

H_d : The fragment found from the suspect's coat does not come from the smashed window from the house where robbery took place.

These propositions remain the same in all the methods used in this analysis.

3.2 Method 1: Two Stage Approach using ASTM approach

First approach taken was the two stage approach using the ASTM method. The prosecution proposition (H_p) and defense proposition (H_d) are as described in section 3.1 .

3.2.1 Stage 1

As the name suggests, the Stage 1 is the first stage of comparison where we check if the trace comes from the specific source. If we have an association, then we proceed to the second stage to check the significance of the association.

ASTM approach is performed on the raw elemental concentrations. We calculate a similarity measure C based on the trace and source $C(u, s)$ and compare it to a threshold τ. If $C(u, s) \geq \tau$, we conclude association and then calculate the 3 types of coincidence probabilities (random match probability) which indicates coincidental association.

We considered measurements of 10 elements (variables), and the mean concentrations in fragment recovered from the suspect is compared to the match intervals for the corresponding elements from the known fragment (crime scene). The mean concentrations of all the 10 elements fall within the match intervals hence we can conclude that the two fragments are chemically indistinguishable from each other. We proceed to stage 2 which is checking the significance of comparison.

3.2.2 Stage 2 : Trace Anchored

In stage 2 trace anchored approach, we compare the trace fragment to different fragments (sources) in background data and calculate the probability of trace fragment matching a randomly selected fragment (source) from the background population as

$$p_C(A|u) = \frac{1}{N_a} \sum_{i=1}^{N_a} I[C(u, A_i) \geq \tau]$$

The random match probability of 0.4348 indicates the trace fragment matches 10 out of the 23 glass fragments from the background database. This indicates weak support for the prosecution proposition (H_p).

3.2.3 Stage 2 : Source Anchored

In stage 2 source anchored approach, we compare the known fragment to different fragments (sources) in background data and calculate the probability of source fragment matching a randomly selected fragment (source) from the background population as

$$p_C(A|s) = \frac{1}{N_a} \sum_{i=1}^{N_a} I[C(s, A_i) \geq \tau]$$

The random match probability of 0 indicates the known fragment does not match any of the 23 glass fragments from the background database. This indicates a strong support for the prosecution proposition (H_p).

3.2.4 Stage 2 : General Match

In stage 2 general match approach, we compare two randomly selected fragments in background database and calculate the probability of all such combinations of the fragments matching in the background database as

$$p_C(A) = \frac{1}{N} \sum_{i \neq j} I[C(A_i, A_j) \geq \tau]$$

The random match probability of 0.2411 indicates 122 pairs out of the possible 506 pairs from the 23 glass fragments in the background database . This again indicates a moderate support for the prosecution proposition (H_p).

With this we can say the ASTM two stage approach provides a moderate evidence to prosecution proposition since we found the association with moderate to high coincidence probability.

3.3 Method 2: Two Stage Approach using Hotelling T^2 approach

Second approach considered was the two stage approach using the Hotelling's T^2 approach. The prosecution proposition (H_p) and defense proposition (H_d) are as described in section 3.1

This method is performed on the log transformed elemental concentrations. Here we use the comparison metric C as the Hotelling's T^2 which follows a multivariate normal distribution. Due to the low replications of the measurements for the source and trace fragment, we reduce the number of elements considered and use elements Ba137, La139 and Pb208 (on log scale) for the analysis. A multivariate normality test was performed on the three selected elements which failed to reject non normality and hence we can considerably proceed with the Hotelling's T^2 approach which requires multivariate normality.

3.3.1 Stage 1

We considered measurements of the 3 selected elements (variables), and check if the difference in the mean concentrations of the 3 elements is significant. The Hotelling's T^2 test statistic is 15.48 with a corresponding p-value of 0.0931 which indicates we have have no evidence to reject the null hypothesis of equal means at 5% level of significance and hence can considerably conclude that the glass fragments found on suspect cannot be chemically distinguished from the glass found at the crime scene (house where the robbery took place). We proceed to stage 2 which is checking the significance of comparison.

3.3.2 Stage 2 : Trace Anchored

In stage 2 trace anchored approach, we compare the trace fragment to different fragments (sources) in background data and calculate the probability of trace fragment matching a randomly selected fragment from the background population. The formula for coincidence probability is same as in section 3.2.2 . The random match probability of 0.1739 indicates the trace fragment matches 4 out of the 23 glass fragments from the background database. This indicates moderate support for the prosecution proposition (H_p).

3.3.3 Stage 2 : Source Anchored

In stage 2 source anchored approach, we compare the known fragment to different fragments (sources) in background data and calculate the probability of source fragment matching a randomly selected fragment from the background population. The formula for coincidence probability is same as in section 3.2.3 . The random match probability of 0.3043 indicates the known fragment matches 7 out of the 23 glass fragments from the background database. This indicates a moderate support for the prosecution proposition (H_p).

3.3.4 Stage 2 : General Match

In stage 2 general match approach, we compare two randomly selected fragments in background database and calculate the probability of all such combinations of the fragments matching in the background database. The formula for coincidence probability is same as in section 3.2.4 .The random match probability of 0.2292 indicates 58 pairs out of the possible 253 pairs from the 23 glass fragments in the background database . This again indicates a moderate support for the prosecution proposition (H_p).

With this we can say the Hotelling's T^2 two stage approach provides a moderate evidence to prosecution proposition since we found the association with considerably moderate coincidence probability.

3.4 Method 3: Neyman-Pearson Likelihood ratio Approach

Third approach considered was Neyman-Pearson Likelihood ratio Approach. The prosecution proposition (H_p) and defense proposition (H_d) are the same as described in section 3.1

In the likelihood ratio approach, calculate the ratio of likelihood of evidence assuming prosecution proposition is true and the likelihood of evidence assuming defense proposition is true. The likelihood ratio gives us the strength of support to the prosecution proposition over the defense proposition. Due to the low replications of the measurements for the source and trace fragment, we reduce the number of elements considered and use elements Ba137, La139 and Pb208 (on log scale) for the analysis.

$$\hat{\lambda_{ss}}(E) = \lambda_{ss}(\hat{\theta_p}, \hat{\theta_d}|E) \text{ where}$$
$$\lambda_{ss}(\hat{\lambda_p}, \hat{\lambda_d}|E) = \frac{L(\hat{\theta_s^*}|s)L(\hat{\theta_s^*}|u)L(\hat{\theta_a}|A)}{L(\hat{\theta_s}|s)L(\hat{\theta_a^*}|u)L(\hat{\theta_a^*}|A)}$$
$$\hat{\theta_p} = \{\hat{\theta_s^*} = T(s,u), \hat{\theta_a} = T(A)\} \text{ under prosecution model}$$
$$\hat{\theta_d} = \{\hat{\theta_s} = T(s), \hat{\theta_a^*} = T(u,A)\} \text{ under defence model}$$

The specific source Neyman-Pearson LR is 0.0656 which is less than 1 (likelihood under defence model is higher than the likelihood under prosecution model) which indicates we are 15 times more likely to observe the evidence if H_d (defence proposition) is true than if H_p (prosecution proposition) is true. This is interesting result since it gives contrasting results compared to the two stage (ASTM and Hotelling's T^2) approach.

3.5 Method 4: ABC Bayes Factor Approach

Fourth approach considered was Approximate Bayesian Computation (ABC) Bayes Factor Approach. The prosecution proposition (H_p) and defense proposition (H_d) are the same as described in section 3.1

The Approximate Bayesian Computation Bayes factor approach works when the likelihood is not tractable, and only retains the samples which are very close to the observed sample. We try to sample from the target posterior distribution to evaluate the (intractable) likelihood function.

$$\beta_{ss}(E) = \frac{f(u|\theta_s)d\pi(\theta_s|s)}{f(u|\theta_a)d\pi(\theta_a|A)}$$

The numerator will be estimated by the prosecution model (M_p) and the denominator will be estimated by the defence model (M_d) where the models are defined as below

$M_p : u_j \sim \mathcal{N}_{10}(\mu_s, \Sigma_s)$ where $j = 1, \ldots, N_u = 3$

$M_d : B_u \sim \mathcal{N}_{10}(\mu_a, \Sigma_a)$; $u_j|b_u \sim \mathcal{N}_{10}(b_u, \Sigma_w)$ where $j = 1, \ldots, N_u = 3$

and the priors will be

$\mu_s \sim \mathcal{N}_{10}(m, B)$

$\mu_a \sim \mathcal{N}_{10}(m, KB)$ where K=10 constant

$\Sigma_s \sim \mathcal{W}_{10}^{-1}(W, \eta_e)$ where $\eta_e = 55$ is degrees of freedom

$\Sigma_a \sim \mathcal{W}_{10}^{-1}(B, \eta_b)$ where $\eta_b = 55$ is degrees of freedom

$\Sigma_w \sim \mathcal{W}_{10}^{-1}(W, \eta_e)$ where $\eta_e = 55$ is degrees of freedom

We use the sample mean as the summary statistic (η) with Euclidean distance function (ρ), 0.4 tolerance level (ϵ) and equal prior model probabilities (0.5) which results in odds ratio 1. The ABC Bayes factor comes out to be 2.4322 which indicates we are approximately 2.5 times more likely to observe the evidence if H_p is true than if H_d is true. This method gives results which similar to the two stage approach in the sense that it indicates support to the prosecution proposition. The tolerance level was set from analysis of results where data on known matches was run.

4 Results

The results of the four methods used for analysis are summarized in table 1

Method	Result	Interpretation	Correct
Two stage ASTM	Association (moderate RM probability)	Support for H_p	Yes
Two stage Hotelling's T^2	Association (moderate RM probability)	Support for H_p	Yes
NP Likelihood approach	LR = 0.06	Support for H_d	No
ABC Bayes factor approach	BF = 2.43	Support for H_p	Yes

Table 1: Summary of four methods

While choosing the data for this project, the glass fragment found on suspect and from the crime scene was taken from the same window pane fragment. Hence ideally we would like the methods to indicate evidence in favor of the prosecution proposition H_p. Two stage approach (both ASTM and Hotelling's T^2) as well as the ABC bayes factor approach indicates support for H_p whereas the NP likelihood approach indicates support in favor of H_d

We had to employ dimension reduction technique in two stage Hotelling's T^2 approach and Neyman Pearson Likelihood approach due to the limited replicate measurements on the glass fragments found on the suspect and at the crime scene. Initially a principal component analysis (PCA) was considered but the principal components fail the multivariate normality assumption. Hence we select the 3 variables such that they are univariately normal as well as combined multivariate normal.

One of the possible explanation for the NP likelihood approach showing opposite result as compared to the other 3 methods is due to the use of subset of the data. We considered measurements of only 3 elements for the method to be used (namely Ba137, La139 and Pb208 on the log scale) due to the limited replications of measurements on the glass fragments found at the crime scene and on the suspect. The 3 elements considered were all univariate normal as well as combined multivariate normal. If we consider some other combination of 3 variables, we might get different results. It might be that the selected variables do not have sufficiant discriminatory power but are well behaved on their own.

5 Discussion

Looking at the results obtained from the 4 methods, we can conclude that the glass fragment found on the suspect XYZ seems to have originated from the broken glass window in the house where the burglary took place.

One of the shortcoming of this analysis is the less number of replicate measurements on the glass fragments found at the crime scene and those found on the suspect. This limits the use of certain techniques (in our case the Two stage approach using Hotelling's T^2 approach and Neyman Pearson Likelihood Ratio approach). We had to use dimension reduction (reduce the trace elements) to use the specified methods. It would have been better if we had more replicates as that would have allowed us to use more techniques as well as yielded more accurate results.

It would also be worthwhile to consider a technique to select variables (trace elements) such that most of the variation in the data is captured without having to consider too many variables. It is certainly a challenge to keep a balance between accuracy of the techniques and keeping the dimensions low. This is a topic I would like to explore further.

6 References

- Dataset used for the analysis `https://github.com/CSAFE-ISU/AOAS-2018-glass-manuscript`
- Danica M. Ommen, Christopher P. Saunders & Cedric Neumann (2017) The characterization of Monte Carlo errors for the quantification of the value of forensic evidence, Journal of Statistical Computation and Simulation, 87:8, 1608-1643, DOI: 10.1080/00949655.2017.1280036
- Andrew van Es, Wim Wiarda, Maarten Hordijk, Ivo Alberink, Peter Vergeer Implementation and assessment of a likelihood ratio approach for the evaluation of LA-ICP-MS evidence in forensic glass analysis
- Statistics and the Evaluation of Evidence for Forensic Scientists (Third Edition) Colin Aitken, Franco Taroni and Silvia Bozza

Appendix A

The 18 elemental concentration of the glass fragments from the broken window found at the house where robbery took place (crime scene) are as follows

Element	Element	Rep 1	Rep 2	Rep 3
Li7	Lithium	1.7	1.68	1.63
Na23	Sodium	104730	105550	105880
Mg25	Magnesium	22750	22690	22610
Al27	Aluminium	2572	2591	2541
K39	Potassium	1303.1	1305.8	1337
Ca42	Calcium	59460	58910	58470
Ti49	Titanium	154.2	152.7	157.9
Mn55	Manganese	17.98	17.87	17.82
Fe57	Iron	670	655	664
Rb85	Rubidium	2.27	2.3	2.35
Sr88	Strontium	28.58	28.72	29.14
Zr90	Zirconium	27.19	27.13	26.19
Ba137	Barium	11.21	10.61	10.12
La139	Lanthanum	3.86	3.67	3.72
Ce140	Cerium	9.08	9	8.99
Nd146	Neodymium	3.22	3.06	3.04
Hf178	Hafnium	0.726	0.682	0.75
Pb208	Lead	1.035	1.132	1.113

Table 2: Trace elemental concentrations using LA-ICP-MS for glass found at the crime scene

The 18 elemental concentration of the glass fragments found on suspect's coat are as follows

Element	Element	Rep 1	Rep 2	Rep 3
Li7	Lithium	2.06	1.92	1.57
Na23	Sodium	106560	105830	103500
Mg25	Magnesium	23020	22580	22730
Al27	Aluminium	2605	2545	2595
K39	Potassium	1319	1325	1306.6
Ca42	Calcium	59830	58140	59630
Ti49	Titanium	156.8	157.1	158.6
Mn55	Manganese	18.13	17.81	17.73
Fe57	Iron	670	677	658
Rb85	Rubidium	2.28	2.29	2.22
Sr88	Strontium	29.37	28.73	29.07
Zr90	Zirconium	27.92	26.61	28.07
Ba137	Barium	11.21	10.56	10.38
La139	Lanthanum	3.83	3.83	3.78
Ce140	Cerium	9.16	9.32	8.9
Nd146	Neodymium	3.23	3.18	3.1
Hf178	Hafnium	0.682	0.661	0.7
Pb208	Lead	1.078	1.129	1.14

Table 3: Trace elemental concentrations using LA-ICP-MS for glass found on suspect's coat

The table indicating the elements used in the analysis and if the elements are Normally distributed (raw data and data on log scale)

Element	Element	Normality	Normality (log scale)
K39	Potassium	No	No
Ti49	Titanium	No	No
Mn55	Manganese	No	No
Rb85	Rubidium	Yes	Yes
Sr88	Strontium	No	No
Zr90	Zirconium	No	No
Ba137	Barium	Yes	Yes
La139	Lanthanum	Yes	Yes
Ce140	Cerium	Yes	Yes
Pb208	Pb	Yes	Yes

Table 4: Trace elements used in the analysis