Dependence among randomly acquired characteristics on shoeprints and their features.

Naomi K. Damary  
*The Hebrew University of Jerusalem*

Micha Mandel  
*The Hebrew University of Jerusalem*

Sarena Wiesner  
*Israel National Police Division of Identification and Forensic Science (DIFS)*

Yoram Yekutieli  
*Hadassah Academic College*

Yaron Shor  
*Israel National Police Division of Identification and Forensic Science (DIFS)*

See next page for additional authors

Follow this and additional works at: https://lib.dr.iastate.edu/csafe_pubs

Part of the Forensic Science and Technology Commons

**Recommended Citation**

Damary, Naomi K.; Mandel, Micha; Wiesner, Sarena; Yekutieli, Yoram; Shor, Yaron; and Spiegelman, Clifford, "Dependence among randomly acquired characteristics on shoeprints and their features." (2018).  
*CSAFE Publications*. 50.  
https://lib.dr.iastate.edu/csafe_pubs/50

This Article is brought to you for free and open access by the Center for Statistics and Applications in Forensic Evidence at Iowa State University Digital Repository. It has been accepted for inclusion in CSAFE Publications by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Dependence among randomly acquired characteristics on shoeprints and their features.

Abstract
Randomly acquired characteristics (RACs), also known as accidental marks, are random markings on a shoe sole, such as scratches or holes, that are used by forensic experts to compare a suspect's shoe with a print found at the crime scene. This article investigates the relationships among three features of a RAC: its location, shape type and orientation. If these features, as well as the RACs, are independent of each other, a simple probabilistic calculation could be used to evaluate the rarity of a RAC and hence the evidential value of the shoe and print comparison, whereas a correlation among the features would complicate the analysis. Using a data set of about 380 shoes, it is found that RACs and their features are not independent, and moreover, are not independent of the shoe sole pattern. It is argued that some of the dependencies found are caused by the elements of the sole. The results have important implications for the way forensic experts should evaluate the degree of rarity of a combination of RACs.

Keywords
Independence test, Shoe pattern, Randomly acquired characteristics, Frequency tables, Accidentals, Shoeprints, Foot wear comparison

Disciplines
Forensic Science and Technology

Comments

Authors
Naomi K. Damary, Micha Mandel, Sarena Wiesner, Yoram Yekutieli, Yaron Shor, and Clifford Spiegelman
Dependence among randomly acquired characteristics on shoeprints and their features

Naomi Kaplan Damary\textsuperscript{a,*}, Micha Mandel\textsuperscript{b}, Sarena Wiesner\textsuperscript{c}, Yoram Yekutieli\textsuperscript{d}, Yaron Shor\textsuperscript{e}, Clifford Spiegelman\textsuperscript{f}

\textsuperscript{a} Department of Statistics, The Hebrew University of Jerusalem, 91905 Jerusalem, Israel
\textsuperscript{b} Department of Statistics, The Hebrew University of Jerusalem, 91905 Jerusalem, Israel
\textsuperscript{c} Israel National Police Division of Identification and Forensic Science (DIFS), 1 Bar-Lev Road, 91906 Jerusalem, Israel
\textsuperscript{d} Department of Computer Science, Hadassah Academic College, 37 Hanevimim St., 910801 Jerusalem, Israel
\textsuperscript{e} Israel National Police Division of Identification and Forensic Science (DIFS), 1 Bar-Lev Road, 91906 Jerusalem, Israel
\textsuperscript{f} Department of Statistics, Texas A&M University, TX, USA

\section{A R T I C L E  I N F O}

\textbf{Keywords:}
- Independence test
- Shoe pattern
- Randomly acquired characteristics
- Frequency tables
- Accidents
- Shoeprints
- Footwear comparison

\section{A B S T R A C T}

Randomly acquired characteristics (RACs), also known as accidental marks, are random markings on a shoe sole, such as scratches or holes, that are used by forensic experts to compare a suspect’s shoe with a print found at the crime scene. This article investigates the relationships among three features of a RAC: its location, shape type and orientation. If these features, as well as the RACs, are independent of each other, a simple probabilistic calculation could be used to evaluate the rarity of a RAC and hence the evidential value of the shoe and print comparison, whereas a correlation among the features would complicate the analysis. Using a data set of about 380 shoes, it is found that RACs and their features are not independent, and moreover, are not independent of the shoe sole pattern. It is argued that some of the dependencies found are caused by the elements of the sole. The results have important implications for the way forensic experts should evaluate the degree of rarity of a combination of RACs.

\section{1. Introduction}

The 2009 NRC report, \textit{Strengthening Forensic Science in the United States: A Path Forward} \cite{4}, and the recent PCAST report to the president, \textit{Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods} \cite{3}, call for improving the scientific basis of forensic procedures. The current paper examines basic assumptions in the field of shoe comparison, currently used in defining the degree of rarity (DOR) for a given shoe sole \cite{2,6,8}.

Footwear examination starts with a comparison of the size, pattern and wear of the shoe in question to the print found in the crime scene, the sole’s DOR of the shoe is then calculated on the basis of Randomly Acquired Characteristics (RACs) that scar its contact surface. These marks are caused by abrasion on the outsole. Unlike manufacturing flaws, the creation of RACs depends on the owner’s walking patterns, the material of shoe sole, the surface with which it comes in contact, the pattern of the shoe and its wear and tear. RACs are characterized by several features \cite{6,5} such as their location, orientation (angle) and shape (defined below), which are used for comparison of shoes.

In contrast to the clear definitions of location and orientation (presented in Section 2), the definition of shape presents a special challenge. \cite{5} define four categories of shapes: lines/curves, circles, triangles, and irregular shaped features. \cite{6} refers to Standardized Individual Characteristics: point, line, curve, enclosure, and three dimensional. In this study we adopt the 7 categories of shape used by the Israeli Police Division of Identification and Forensic Science (DIFS), as described in Section 2 and Fig. 3 \cite{8}. All of these definitions are influenced by the design of the shoe sole, as only certain parts actually come into contact with the ground. Thus, the shapes used in this study are not really shapes, but types of RACs in the context of shape. For the sake of simplicity we have chosen to call them “shape types”.

\* Corresponding author.
E-mail addresses: naomi.kaplan@mail.huji.ac.il (N. Kaplan Damary), micha.mandel@mail.huji.ac.il (M. Mandel), sarenawiez@gmail.com (S. Wiesner), yoramye@hadassah.ac.il (Y. Yekutieli), yaronshor@gmail.com (Y. Shor), cliff@stat.tamu.edu (C. Spiegelman).

\url{https://doi.org/10.1016/j.forsciint.2017.11.038}
Previous studies assume that the features of a given RAC are independent of each other, as are the RACs themselves [6,8,7] and therefore the DOR of a single RAC is calculated by multiplying the probabilities of its features and the DOR of the entire shoe is calculated by multiplying the DOR of all RACs. Let \( \ell_i, s_i, o_i \) be the location, shape and orientation of RAC \( i \), and \( \Pr(\ell_i), \Pr(s_i) \) and \( \Pr(o_i) \) be the corresponding probabilities, then, under the independence assumption, RAC \( i \)'s DOR is calculated as

\[
\text{DOR}_i = \Pr(\ell_i) \times \Pr(s_i) \times \Pr(o_i),
\]

and the overall DOR is the product of the DORs of all RACs found on the shoe: \( \text{DOR} = \prod_i \text{DOR}_i \). However, if the independence assumption does not hold, the calculation is incorrect, and the rarity of a certain shoe may be overestimated. The goal of the current paper is to statistically test the independence assumption.

Conventional practice focuses on the attempt of forensic experts to match crime scene impressions with a suspect’s shoe. The current study focuses on the shoe’s rarity on the assumption that linkage of the suspect to a rare print would be much more damning than connection to a common print that could potentially have been made by others. Thus, an understanding of the independence among RACs and the subsequent effect on shoe DOR is of considerable significance.

The present study focuses on the relationship among the features of a given RAC using data collected by the DIFS [8,7]. Three propositions are tested; each contends that the RACs are created randomly and independently on the shoe sole, but each includes in addition:

1. The shape type of the RAC is independent of its location.
2. The orientation of the RAC is independent of its shape type.
3. The orientation of the RAC is independent of its location.

The data and methods used in this article are described in Sections 2 and 3. In Section 4 the propositions above are tested and rejected, showing that there is in fact an association among the features. In Section 5 two analyses are carried out in an attempt to control the effect of the sole pattern. The questions from Section 4 and similar issues regarding the possible reasons for the dependencies are investigated in Section 6. Section 7 concludes the paper with a discussion.

2. Data

The DIFS has constructed one of the largest databases of RACs to date [8,7], including about 13,500 RACs from 380 lab impressions (shoe prints). In order to define the RACs’ locations and orientations, a crucial preprocessing step was to normalize all shoe impressions on a standardized \( x-y \) axis with identical length and orientation. For each test impression, the top and bottom of the shoeprint were marked to indicate the direction of the major axis and to determine the length of the shoe. The origin of axes was set at the middle point between the two marked extremities. The minor axis was defined as the line perpendicular to the major axis that passes through the origin of axes. Finally, all axes were standardized to the same length. More details of this normalization process can be found in Appendix A1.

The three features of a RAC are defined as follows:

1. Location: the normalized shoe sole was divided into 14 sub-areas, on the basis of experts’ knowledge; see Fig. 1.
2. Orientation: determined by the angle of the RAC with respect to the \( x \) axis of the shoe (see Fig. 2 for an example). In addition, the orientations were divided into 9 groups (20° each). As some shapes have potentially larger errors than others, a grade measure was used to express the degree of the shape elongation [8]. Orientation error was found to be a function of the orientation grade. When analysing orientation, the analysis was
conducted only on RACs with a low degree of error in their orientation (grade>5, meaning shapes that are roughly 5 times longer than their width).

3 Shape type 1 is based on the definitions determined by the Israeli Police DIFS for classification purposes (see Fig. 3):

(1) Scratch – a long and narrow tear of the sole.
(2) Hole – a cut or tear that does not fit any of the other categories.
(3) Cut-off corner – a corner of a pattern element that is torn.
(4) Rift – a tear that crosses a narrow pattern element.
(5) Foreign object – a stone, pebble etc. stuck in between pattern elements.
(6) Schalamach – micro-tear of the border of an element.
(7) Missing part – a significant section of the sole pattern is torn off.

Since the number of RACs of type 5 (Foreign Object) and 7 (Missing Part) is small, these were omitted when analysing shape type.²

3. Methods

The three features (location, orientation and shape type) are categorical variables with different number of categories. In order to test independence between a pair of features, the data is arranged in a two-way contingency tables in which \( O_{ij} \) is the number of RACs in category \( i \) in the first feature and category \( j \) in the second. Let \( E_{ij} \) be the expected number of observations under the proposition that the features are independent. Specifically, \( E_{ij} = n_i \times m_j/N \), where \( n_i \) is the total number of RACs in category \( i \) of the first feature, \( m_j \) is the total number of RACs in category \( j \) of the second feature, and \( N \) is the total number of RACs. The test is conducted based on the standard \( \chi^2 \) statistic:

\[
\chi^2 = \sum_{i} \sum_{j} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}.
\]

The null hypothesis of independence includes two parts. First, it assumes that RACs are independent, meaning that the occurrence of a particularly configured RAC does not predict the occurrence of another with a certain configuration. Second, the hypothesis assumes that RAC features are also independent, i.e. the shape type distribution does not depend on the RAC’s location or orientation. Under the null proposition, this statistic has approximately a Chi-squared distribution with \((I-1) \times (J-1)\) degrees of freedom, where \( I \) and \( J \) are the total numbers of categories of the two variables from which the table is constructed. The larger the value of \( \chi^2 \), the more evidence exists against independence [1]. For contingency tables having small cell counts, the \( p \)-value is computed using repeated simulations. This is done by random sampling from the set of all contingency tables with the same given margins.

The Chi-square test assumes that both the features of a RAC and the RACs themselves are independent. If the null independence assumption is rejected, the test cannot indicate whether the features of a RAC are dependent, the RACs are dependent or both. Therefore, a test that is less sensitive to possible dependence among RACs has been used in order to examine the hypothesis that RAC features are independent, while controlling the dependence among RACs. This was conducted to test independence between the location and shape features by sampling one RAC from each shoe, thus creating a database of 386 independent RACs with their location and shape features. These RACs are independent since they were sampled from different shoes. The Chi-square statistic was calculated based on the sampled data. Since the sample size is small, only the 5 horizontal rows of the original sub areas shown in Fig. 1 are used. To minimize the effect of the random sampling, the process was repeated 100 times and the average of the chi-square statistics was used. The null distribution was calculated by sampling one shape and one location independently from each shoe. This statistic, which is based on one RAC from each shoe, reduces considerably the number of observations and hence is expected to have small power. A second test uses the Fisher method. Instead of averaging the 100 Chi-square statistics as described above, it calculates the \( p \)-value for each sample and combines the 100 \( p \)-values by \(-\sum \ln(p_i)\), where \( p_i \) is the \( p \)-value obtained for the \( i \)th simulated data set. This was compared to the null distribution, Gamma(100, 1). This approximates the global null distribution of independence between features of a RAC and independence of the RACs themselves, assuming the \( p \)-values are independent.

4. Data analysis

We start by testing independence between the shape type and the location of the RACs. The frequency of shape types in different sub-areas is presented in Table 1.

---

1 Shape type was defined in a different way than the shape feature in Yekutieli et al. [8], Wiener et al. [7], Stone [6] and Speir et al. [5] and therefore the conclusions here cannot be applied to their research without further examination.

2 Analyses including these two types were also conducted with \( p \)-values calculated using simulations as described in Section 3. All results were similar to those without these two shape types.
The independent and Shalamach residuals in areas 3, 8, and 11 and shape type is the nature of the shoe sole element, since certain shape types are defined by these very elements. For example, a cut-off corner can only appear in an element that contains a corner. Schalamach RACs are micro tears of the boarders of elements (resulting from wear). Thus, it is reasonable that in areas where there is less pressure caused by the foot (2, 9) there will be less RACs of that type. A possible explanation for the abundance of Rifts in areas 3 and 8 is that most shoes do not have a contact surface in these areas, and those shoes which do, have patterns that contain lines, the only elements in which the Rift type can appear. Holes can occur on almost every element which may explain the small absolute value of its residuals.

The second proposition concerns independence between orientation and shape type. The results of the Chi-square independence test reveals that association does exist between the shape type and the orientation (p-value < 0.001).

The extreme residuals are presented in Table 2. Extreme positive residuals mean that there are more observations than would be expected under the independence assumption in that category, and extreme negative ones mean there are less. The residual analysis suggests that Scratch shapes tend to have an angle that is proximate to the y (long) axis of the shoe i.e. tend to have a high degree of orientation (in absolute value). Holes and Shalamach tend to have an orientation that is in proximity to the x axis of the shoe i.e. a lower orientation. Cut-off corner shapes tend to have an orientation of (−50, −30) and (30, 50), that is an angle of ±45°. This is caused by the orientation of the square elements, which are indeed usually squares and not diamonds. Rifts do not have significant residuals since they can occur in different thicknesses, which influence the orientation.

In addition, use of the two statistics which test the hypothesis of independence between the location and shape features, while controlling the dependence among RACs, produced results that led to the rejection of the independence hypothesis (p-values < 0.001).

To conclude, the null proposition is rejected in favour of dependence between the RACs or their features. The implication of this finding is that probabilities cannot be multiplied in order to

---

**Table 2**
The residuals of the independence test between shape types and orientation. Significant residuals with absolute value >3 are colored in orange (dark grey in the print version) and those with absolute value >2 are colored in yellow (light grey in the print version).

<table>
<thead>
<tr>
<th>x</th>
<th>y</th>
<th>(-90, -70]</th>
<th>(-70, -50]</th>
<th>(-50, -30]</th>
<th>(-30, -10]</th>
<th>(-10, 10]</th>
<th>(10, 30]</th>
<th>(30, 50]</th>
<th>(50, 70]</th>
<th>(70, 90]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>6.28</td>
<td>3.75</td>
<td>1.07</td>
<td>-4.55</td>
<td>-10.53</td>
<td>-3.01</td>
<td>0.74</td>
<td>5.35</td>
<td>3.30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2.18</td>
<td>-0.39</td>
<td>-1.50</td>
<td>3.53</td>
<td>2.22</td>
<td>2.05</td>
<td>-0.93</td>
<td>-2.47</td>
<td>-0.66</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.86</td>
<td>-0.35</td>
<td>3.14</td>
<td>0.83</td>
<td>-0.72</td>
<td>-0.76</td>
<td>2.94</td>
<td>-1.44</td>
<td>-2.54</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.23</td>
<td>-0.39</td>
<td>0.57</td>
<td>0.29</td>
<td>-0.07</td>
<td>0.25</td>
<td>0.40</td>
<td>-0.68</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.28</td>
<td>-5.79</td>
<td>-1.56</td>
<td>1.20</td>
<td>10.02</td>
<td>3.78</td>
<td>-1.78</td>
<td>-4.02</td>
<td>-3.94</td>
</tr>
</tbody>
</table>
calculate the DOR of a certain RAC or of a certain shoe sole. It is therefore very important to understand the dependence among the features of a single RAC as well as among the RACs themselves. The next section further investigates several possible causes for these dependencies.

5. Further investigation of the shape type feature

Seven shape types are presented in this study: scratch, hole, cut-off corner, rift, foreign object, Schalamach, and missing part. The definition of these shape types is convenient for expert practitioners but it confuses shape types and shoe-sole elements, thereby creating dependence between them. These are not really “shapes” but rather “RAC types” as they indicate how the RAC was created, and certain shape types are indeed defined by the shoe-sole elements. It is possible that a definition of shape types that is unrelated to shoe-sole elements would lead to different results. To examine this possibility, two further analyses were conducted.

The first repeats the analysis above using only scratches and holes, as these are shape types that do not involve elements of the shoe. The second analysis ignores the shape type and instead tests dependence between the size of the RAC and its location and orientation.

The first analysis did not reject the independence assumption between the two shapes and location (p-value = 0.9), but found association between the two shapes and orientation (p-value < 0.001). This may be a result of the previous finding of scratches tending to have an orientation that is parallel to the walking direction (y axis).

For the second analysis, the size of each RAC was calculated as the area in square pixels which is the product of the length and width of the tight bounding box that encloses the RAC. This does not measure the actual area of the RAC but serves as an approximation of its size, since the actual size is difficult to calculate, especially for open shapes that do not contain a well-defined area. The size was divided into 20 groups, each containing 5% of the observations. As in the previous analyses, a Chi-square test was used.

An association was found between the size of the RAC and each of the features (location, shape and orientation) with p-value < 0.001 in each test. These analyses indicate that the independence assumption among the RAC features does not hold. In future studies it is planned to use a better definition of shape and retest the assumptions.

6. Further investigation of the dependence among the features

Differences among elements of the shoe soles, walking patterns, walking environments and so on, affect the number of RACs and their features. Therefore, the dependence among the features of a RAC may be caused by the differences among shoe soles. We indeed found, using a Chi-square test, that different shoes were prone to develop RACs with different features (see Section 6.1). We believe that the elements of the sole have tremendous impact on the location, shape and orientation of the RAC and thus we analyse subgroups of shoes having similar patterns.

The database contains shoes of various makes and sole patterns. Three relatively frequent patterns Nike Shox R4 (NSR4, n = 36), Nike Shox NZ (NSNZ, n = 27) and Classic Timberland (CT, n = 22) were identified and are presented in Fig. 5.

Although the database is limited in the number of shoes of the same make, we perform a preliminary analysis in order to test independence between RACs on similar shoes. The same analysis should be conducted with larger databases (at this time, a larger database of shoes with identical elements is being compiled).

6.1. Association between sole patterns and RACs’ features

The first analysis examines whether certain patterns are more likely to develop certain types of RACs than others. To answer this question, Chi-square independence tests were conducted between each of the three features of the RAC and the pattern of the shoe. The p-values of shape type versus pattern, and location versus pattern were calculated using 100,000 simulations, since the numbers of observations are small.

Fig. 5. Three frequent shoe patterns: from left to right NSR4, NSNZ, and CT.
It was found that the pattern is associated with the shape type (p-value < 0.001). Scratches are more likely to appear on CT shoes than on NSR4. This may be a result of the CT pattern having larger elements which are more susceptible to Scratches. In addition, Shcallamach shapers appear more on NSR4 shoes and less on CT. Being heavy duty shoes, the CT pattern is made of stronger materials and hence less prone to Shcallamach, which are micro tears of the border of elements. In addition, there are no shapes of type Rift in CT shoes, as would be expected since this pattern contains no thin lines.

The location of the RACs was also found to be associated with the shoe pattern (p-value < 0.001). Pattern NSR4 has less RACs in area 3 than would be expected under the independence assumption, and pattern NSNZ has more. In addition, pattern CT tends to have less RACs in area 9 and more in areas 11 and 12.

Orientation and pattern are found to be independent, that is, the distribution of orientation in these three patterns is similar.

6.2. Association between features within pattern

Section 4 shows that dependence exists using all of the shoes in the database. The data contain various patterns, and, as shown above, different patterns are prone to develop RACs having different features. It is therefore interesting to check whether the latter association explains the dependence found in Section 4, or whether dependence exists within a given shoe pattern. In order to investigate this, the analysis of Section 4 is repeated for each of the three shoe patterns separately. All analyses were based on Chi square tests with p-values calculated using simulations. The results are given in Table 3.

Dependence exists between shape type and location and between shape type and orientation in the two Nike patterns, but does not exist between orientation and location. The number of CT shoes is relatively small and hence the power of the test is low, but still association exists between orientation and shape type. Although the independence found between orientation and location may be due to small sample sizes, it may also be a real phenomena in these patterns; this should be further explored.

7. Discussion and conclusions

This article examines the relationship among RAC features. It has been shown that association exists between shape type and location and between shape type and orientation even within a specific shoe pattern. The calculation of the DOR as a product of probabilities as suggested by Stone [6] and discussed in Section 1 is therefore invalid, and the dependence should be modelled for a well founded evaluation of the DOR.

An important feature that affects the creation of a RAC is the element of the shoe sole, on which the RACs appear. The analysis takes into account the shoe sole patterns, but since they consist of a large number of different elements, a more detailed analysis is required. This study demonstrated through the use of 3 sample patterns, how dependence between RACs’ features and shoe sole patterns can be tested. Clearly, broader conclusions can be drawn only after examining more patterns, but this study serves as an example of what might be done when more data become available. In future analyses, the relationship between elements and the RACs’ features should be investigated. It is our hope that this will lead to a deeper understanding of the relationships between the features and the elements.

Besides the elements of the shoe sole, other factors specific to the shoe, such as the owner’s walking patterns and the ground with which it comes in contact, may affect the creation of RACs. Thus, a further analysis which takes these factors into account is required.

Our study has several limitations. First, the definition of shape type needs revision. Here, the shape type is defined in a way that confuses shape types and elements, thus causing dependence between them. It is possible that a different definition of shape that does not depend on the element will lead to different results.

Furthermore, right shoe data were superimposed on the left shoe data as a mirror image, and right and left identifiers were not included in the analysis. It should be examined whether this factor is important in analysing the dependencies.

Moreover, the human factor cannot be ignored. Some experts may have identified more RACs than others. It is possible that some are more sensitive to various types of RACs; one may identify more Holes while another finds more Scratches. These questions should be studied in the future.

Acknowledgements

This material was based upon work partially supported by the National Science Foundation under Grant DMS-1127914 to the Statistical and Applied Mathematical Sciences Institute. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

This work was supported by the Israel Science Foundation (grant no. 519/14).

The authors would like to express their gratitude to the editor, the associate editor and the referees for their valuable comments and suggestions.

Appendix A.

A.1 The normalization process

This normalization was done by first marking a shoe-aligned coordinate system on each print and then standardizing the shoe according to this system. The standardization was done by transforming all measurements from image coordinates to the shoe aligned coordinate system. The process includes the following steps:

(a) Marking of each shoe as described in the body of the text.
(b) Translation of the marked origin of axes to (0,0) using a special Matlab program.
(c) Rotation by the direction of the shoe aligned coordinate system.
(d) Scaling by the length of the shoeprint.
(e) Multiplying the horizontal axis by –1 or 1, to mirror if needed such that all shoeprints will be turned to left shoes.

A.2 The residuals table

Table 4.
Table 4
The residuals $\sqrt{\text{observed expected}}$. Significant residuals with absolute value $>3$ are colored in orange (dark grey in the print version) and those with absolute value $>2$ are colored in yellow (light grey in the print version).

<table>
<thead>
<tr>
<th>Shape type</th>
<th>Sub-area</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scratch</td>
<td>1</td>
<td>2.08</td>
<td>0.86</td>
<td>-2.95</td>
<td>-2.90</td>
<td>-1.66</td>
</tr>
<tr>
<td>Holo</td>
<td>2</td>
<td>-0.49</td>
<td>1.18</td>
<td>-1.52</td>
<td>1.42</td>
<td>-3.00</td>
</tr>
<tr>
<td>Cut-off corner</td>
<td>3</td>
<td>-3.60</td>
<td>0.58</td>
<td>-0.88</td>
<td>6.43</td>
<td>-0.79</td>
</tr>
<tr>
<td>Rift</td>
<td>4</td>
<td>1.09</td>
<td>0.55</td>
<td>1.17</td>
<td>-1.76</td>
<td>-0.78</td>
</tr>
<tr>
<td>Shoallannach</td>
<td>5</td>
<td>0.77</td>
<td>-0.12</td>
<td>0.03</td>
<td>-3.02</td>
<td>1.88</td>
</tr>
<tr>
<td>6</td>
<td>-0.30</td>
<td>-1.50</td>
<td>2.47</td>
<td>-3.13</td>
<td>6.50</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>-1.07</td>
<td>-0.92</td>
<td>1.06</td>
<td>1.32</td>
<td>2.72</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>-2.16</td>
<td>-0.58</td>
<td>2.89</td>
<td>4.34</td>
<td>-0.42</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.74</td>
<td>1.42</td>
<td>-2.64</td>
<td>0.54</td>
<td>-4.42</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.95</td>
<td>1.19</td>
<td>-2.16</td>
<td>-1.93</td>
<td>-1.97</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>-0.59</td>
<td>-1.48</td>
<td>1.70</td>
<td>5.19</td>
<td>-0.80</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.54</td>
<td>-0.09</td>
<td>1.01</td>
<td>-1.88</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0.43</td>
<td>-0.37</td>
<td>-0.47</td>
<td>-1.30</td>
<td>1.89</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>-1.22</td>
<td>-1.31</td>
<td>2.23</td>
<td>5.16</td>
<td>-0.43</td>
<td></td>
</tr>
</tbody>
</table>

References