Using Machine Learning to Match Striae Pattern on Land Engraved Areas of Bullets

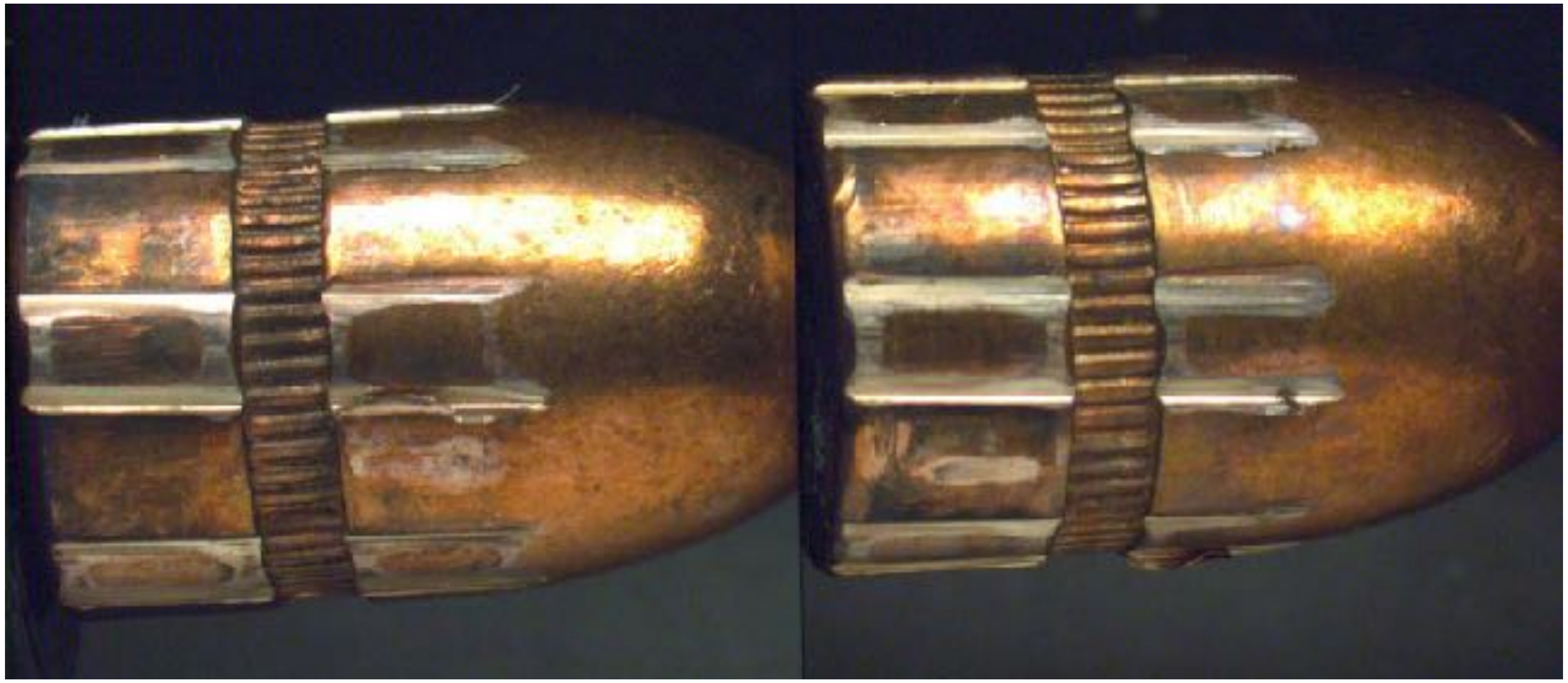
Heike Hofmann (hofmann@iastate.edu, @heike_hh)
ISU CSAFE bullet team

Center for Statistics and Applications in Forensic Evidence
ForensicStats.org

Inaugural Ingram Olkin S3 Forum on Gun Violence Prevention

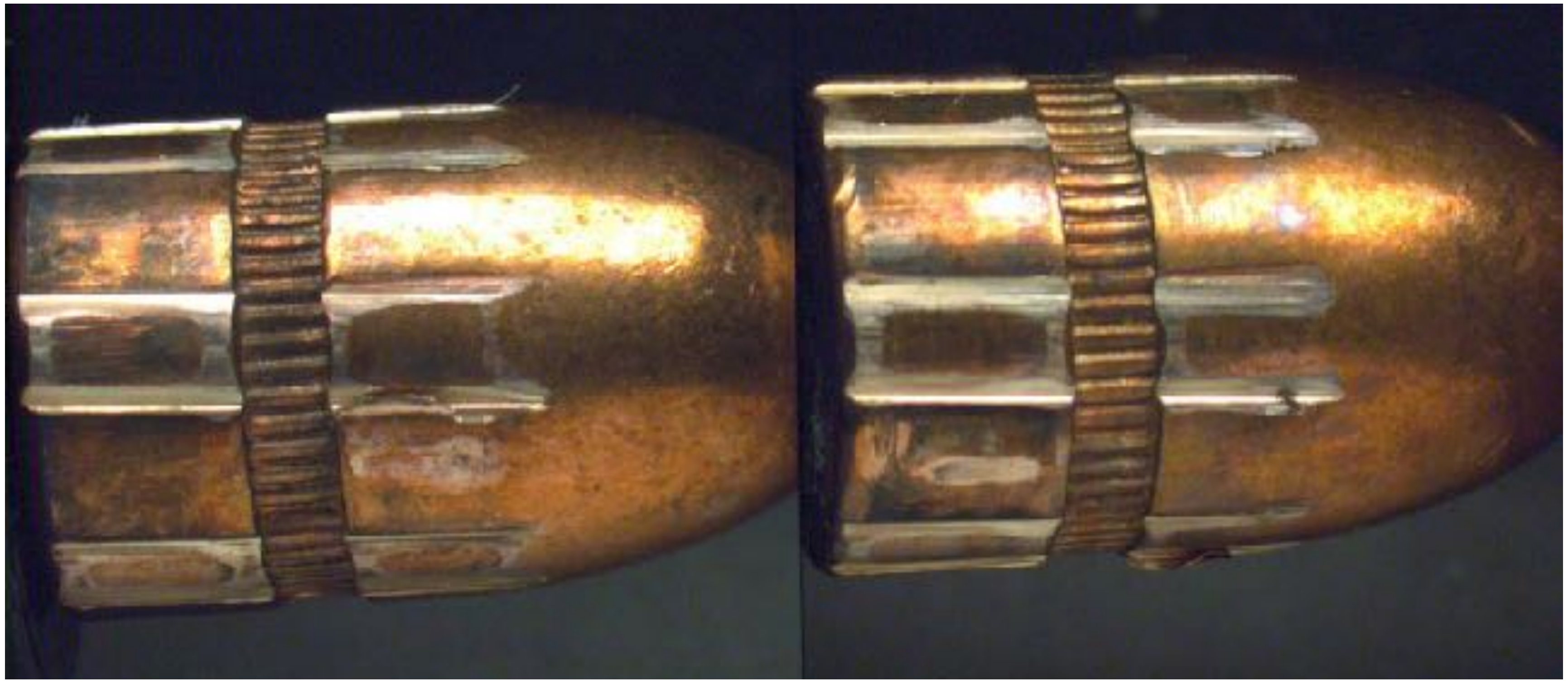
Disclaimer

- ★ This work has nothing directly to do with directly preventing Gun Violence
- ★ Our motive is to establish a scientific foundation for the evaluation of forensic (pattern) evidence



Outline

- ★ Background, Data, and access to it
- ★ Methodology to extract information from raw data
- ★ Results from Matching



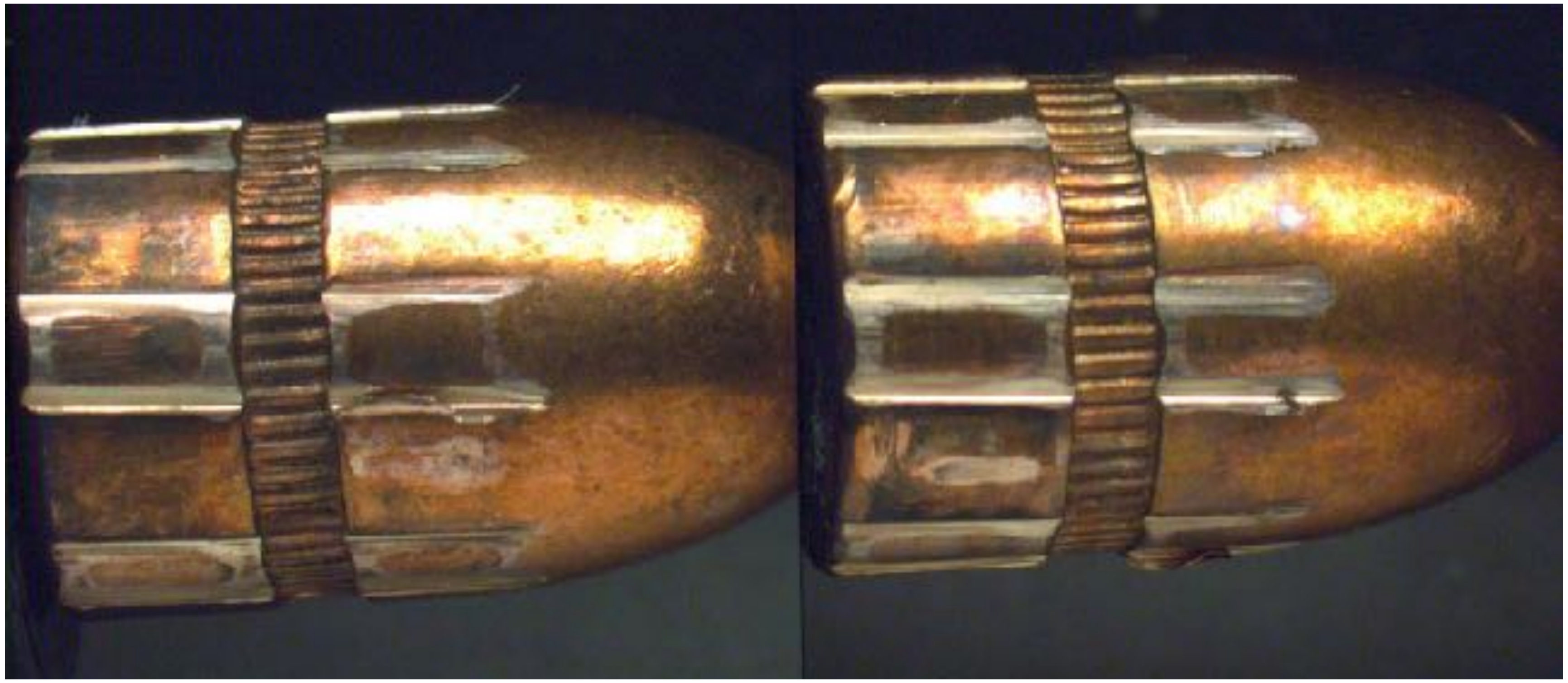
Over-arching Objective

★ **Same Source Problem**: were two bullets fired through the same gun barrel?

★ **Currently**: Firearms and Toolmarks Examiner use visual inspection under a comparison microscope: *subject bias, error rates?*

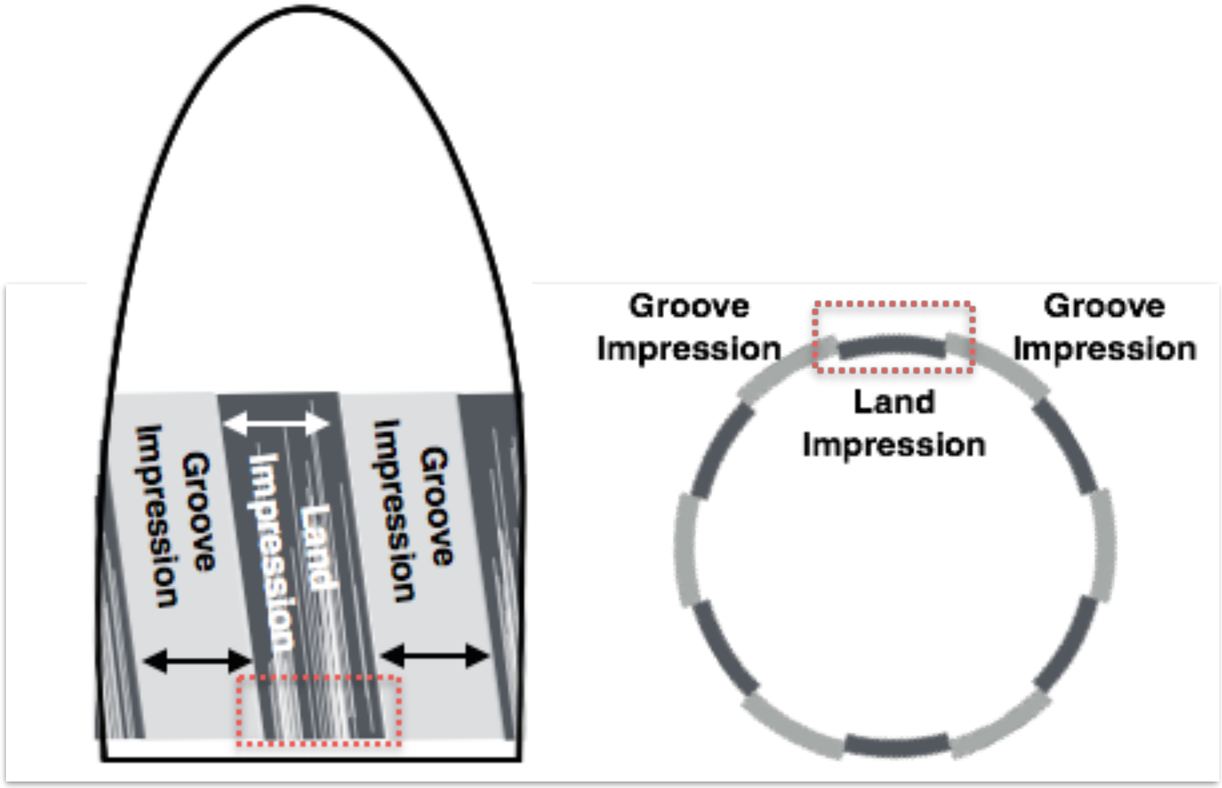
"much forensic evidence – including, for example, bite marks and firearm and toolmark identification is introduced in criminal trials without any meaningful scientific validation, determination of error rates, or reliability testing." (National Research Council 2009)

★ **Goals**: (1) determine *score* as objective measure for the match, (2) establish error rates



Barrel rifling and striae

- ★ Barrel rifling introduces land and groove impressions on bullets
- ★ micro-imperfections introduce striation marks



Data Sources

- ★ NIST Ballistics Toolmarks Research Database: https://tsapps.nist.gov/NRBTD
- ★ 2d images and 3d scans of cartridge cases (firing pin and breech face impressions) and bullets (Land engraved areas)
- ★ Relatively little data on bullets, larger number of cartridge cases

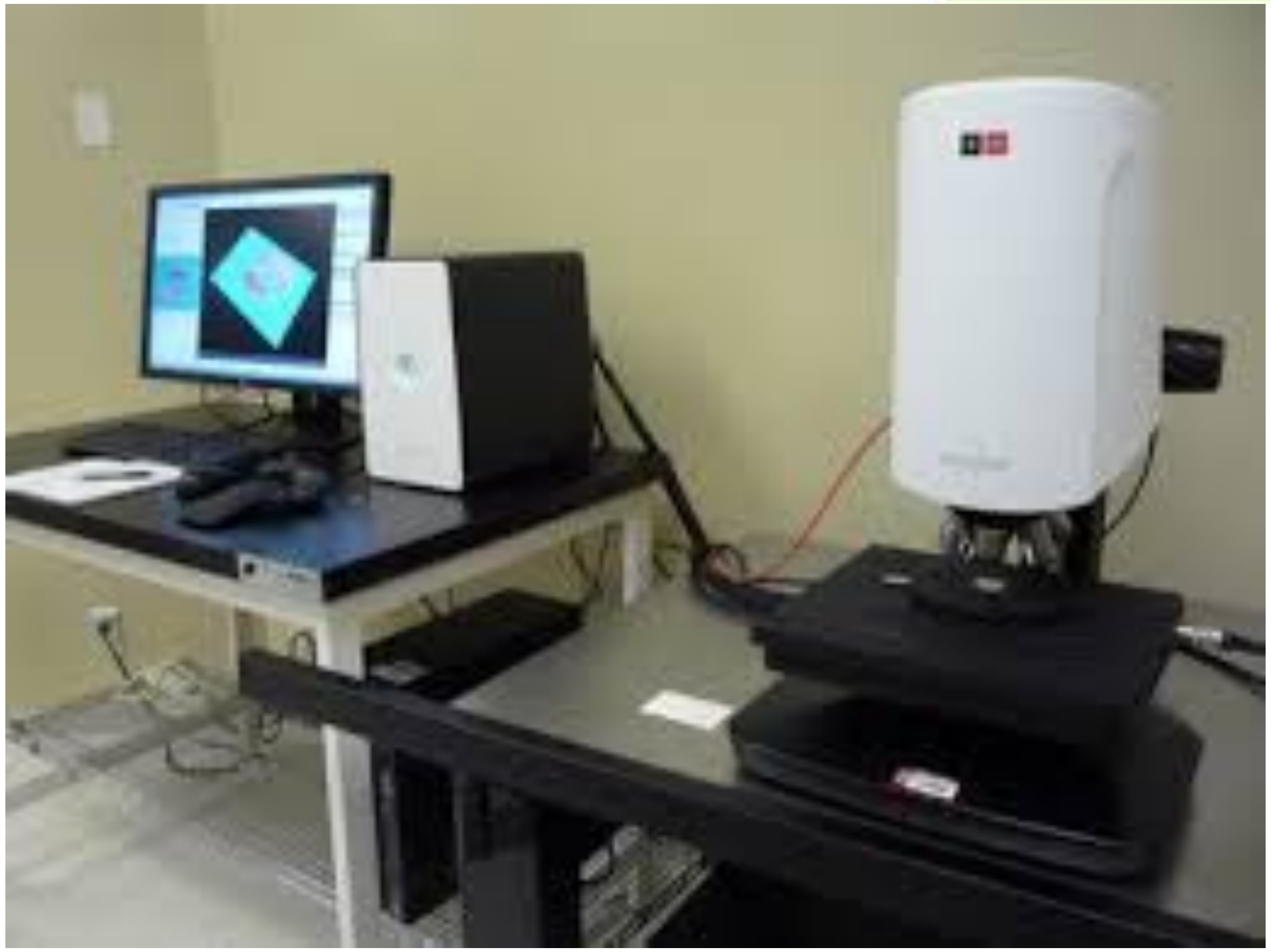
Microscope Facility

Roy J. Carver High Resolution Microscopy Facility

Two Sensofar Confocal Light Microscopes

Four undergraduates scanning bullet lands

3d topographic images: height measurements on x-y grid



csafe
Center for Statistics and Applications in Forensic Evidence
ForensicStats.org

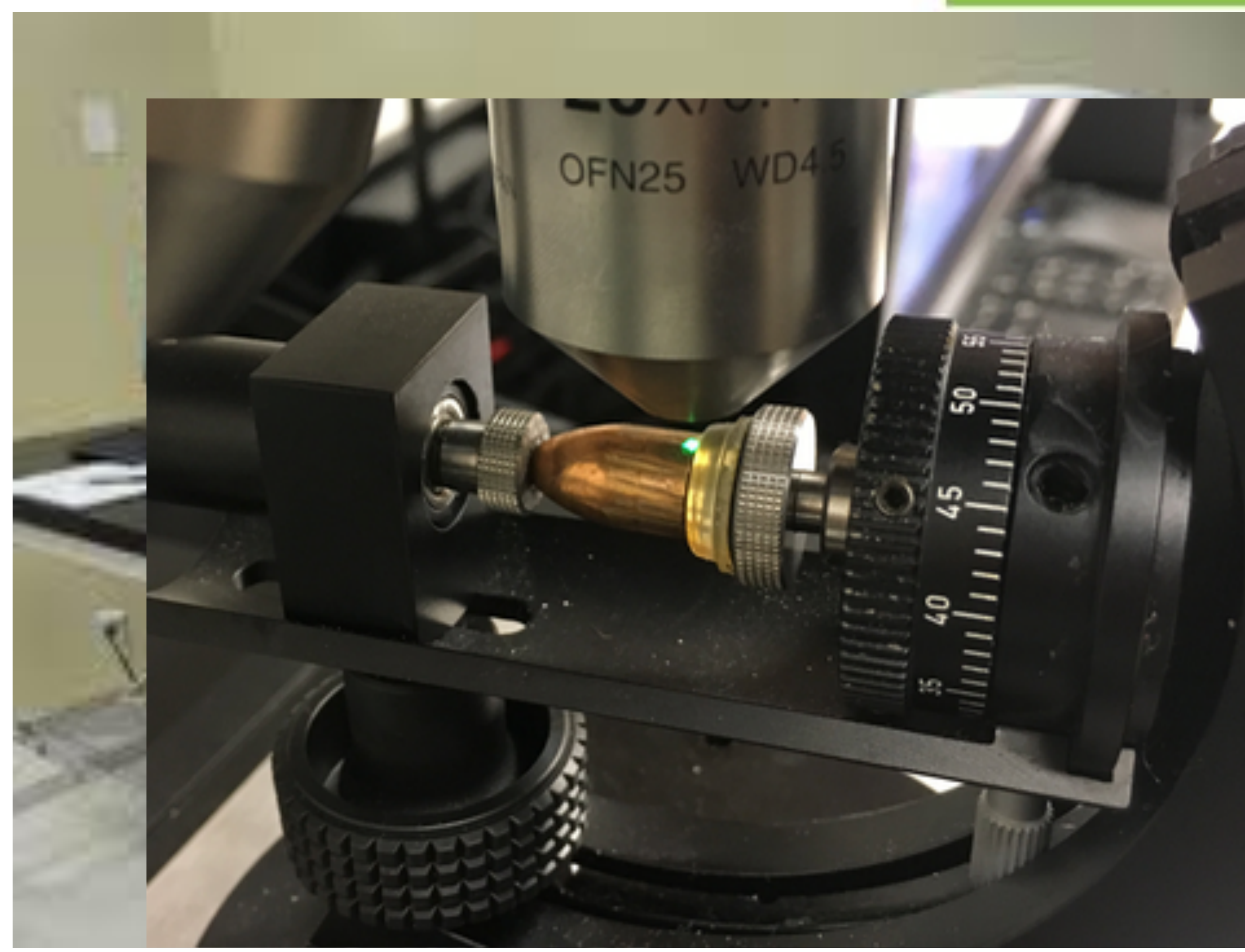
Microscope Facility

Roy J. Carver High Resolution Microscopy Facility

Two Sensofar Confocal Light Microscopes

Four undergraduates scanning bullet lands

3d topographic images: height measurements on x-y grid



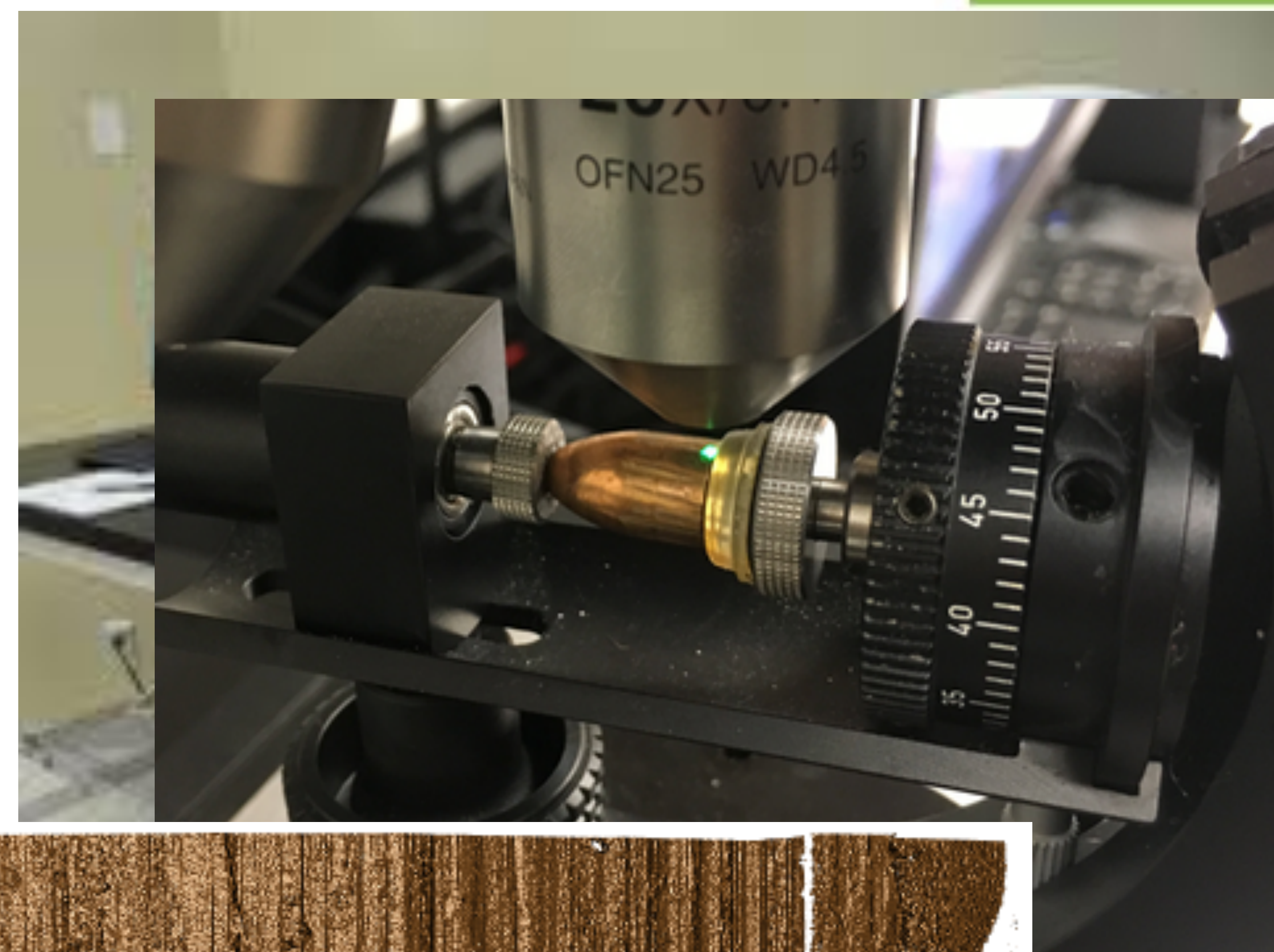
Microscope Facility

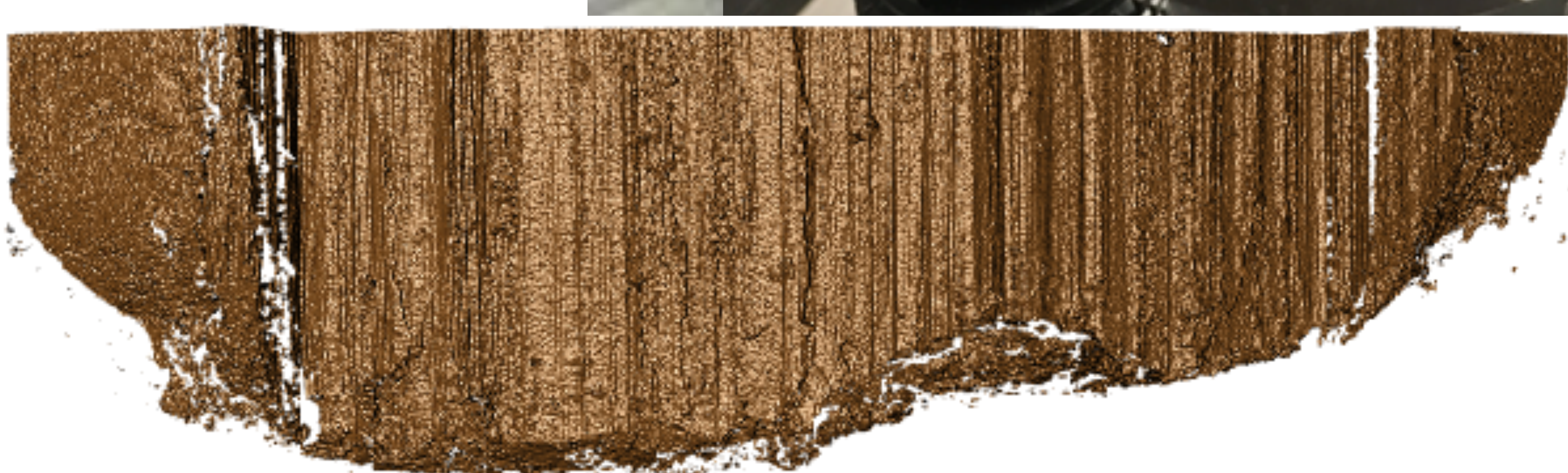
Roy J. Carver High Resolution Microscopy Facility

Two Sensofar Confocal Light Microscopes

Four undergraduates scanning bullet lands

3d topographic images: height measurements on x-y grid





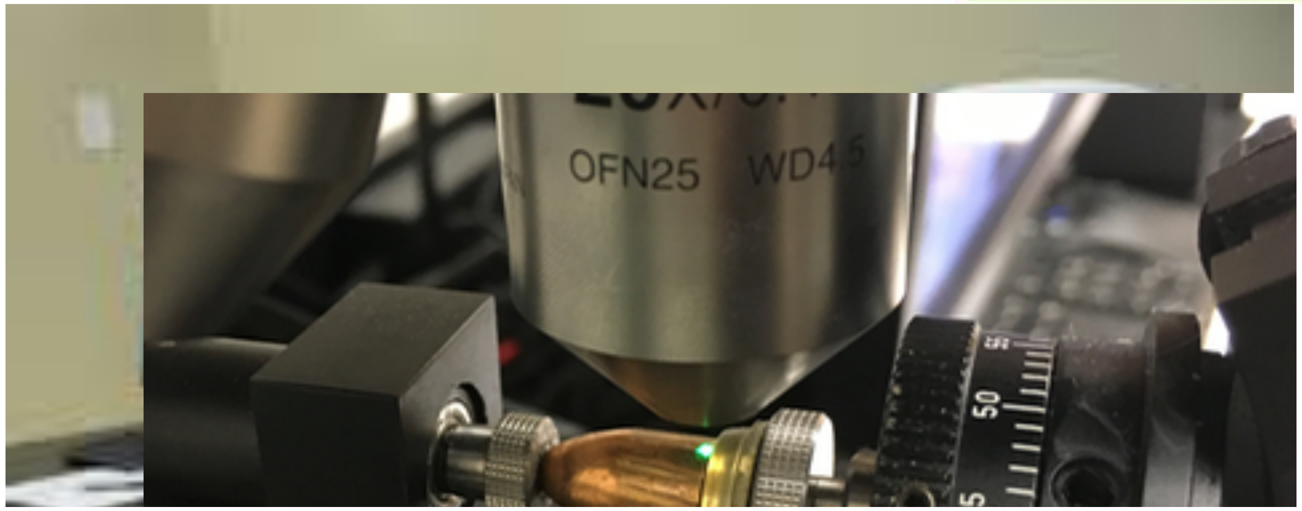
csafe
Center for Statistics and Applications in Forensic Evidence
ForensicStats.org

Microscope Facility

Roy J. Carver High Resolution Microscopy Facility

Two Sensofar Confocal Light Microscopes

Four undergraduates scanning bullet lands



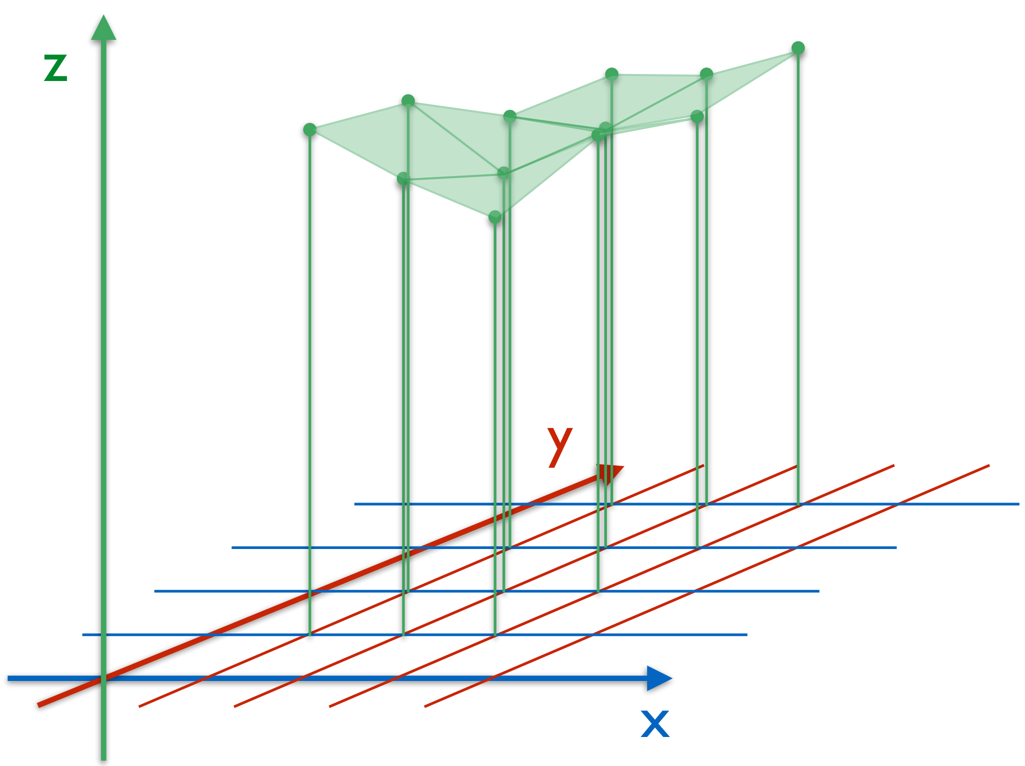
Data from CL Microscope

x-y-z files

Data captured on grid of $0.645\ \mu m \times 0.645\ \mu m$

Total captured area for each land

~ 2.2 mm x 0.6 mm



x - y - z file

x	y	z
18.705	0.000	-25.221138
19.350	0.000	-25.253155
19.995	0.000	-25.335022
20.640	0.000	-25.418171
21.285	0.000	-25.477917
21.930	0.000	-25.541687
22.575	0.000	-25.673903
23.220	0.000	-25.966341
23.865	0.000	-40.070286
24.510	0.000	-40.407612
25.155	0.000	-40.587063
25.800	0.000	-33.437973
26.445	0.000	-33.691895
27.090	0.000	-39.690674
27.735	0.000	-40.317741

.
.
.

x3p format

ISO standard ISO5436 – 2000

- x3p is a container format, consisting of
  - a binary surface matrix
  - an xml file with meta information (specifications of the capturing device, operator information, data specific records)
- Tools for working with x3p files: OpenFMC (C, Matlab) Suite of R packages developed at CSAFE (x3ptools, bulletxtrctr)

Data collected at CSAFE

- In collaboration with Forensic Labs and Police Departments
  - Srinivasan Rathinam, LAPD:
    4 bullets per barrel for 626 Beretta firearms
  - Steve Kramer, St Louis PD:
    2 SigSauer barrels with 192 fired bullets each
  - Melissa McNally, Houston FSI:
    test sets (6 kits with 25 bullets each),
    persistence data shots 11-50 for eight Ruger barrels
  - Hamby Sets 10, 36, 44, 224, and a clone (35 bullets each)
- Total of > 20k scans of Land engraved areas

csafe
Center for Statistics and Applications in Forensic Evidence
ForensicStats.org

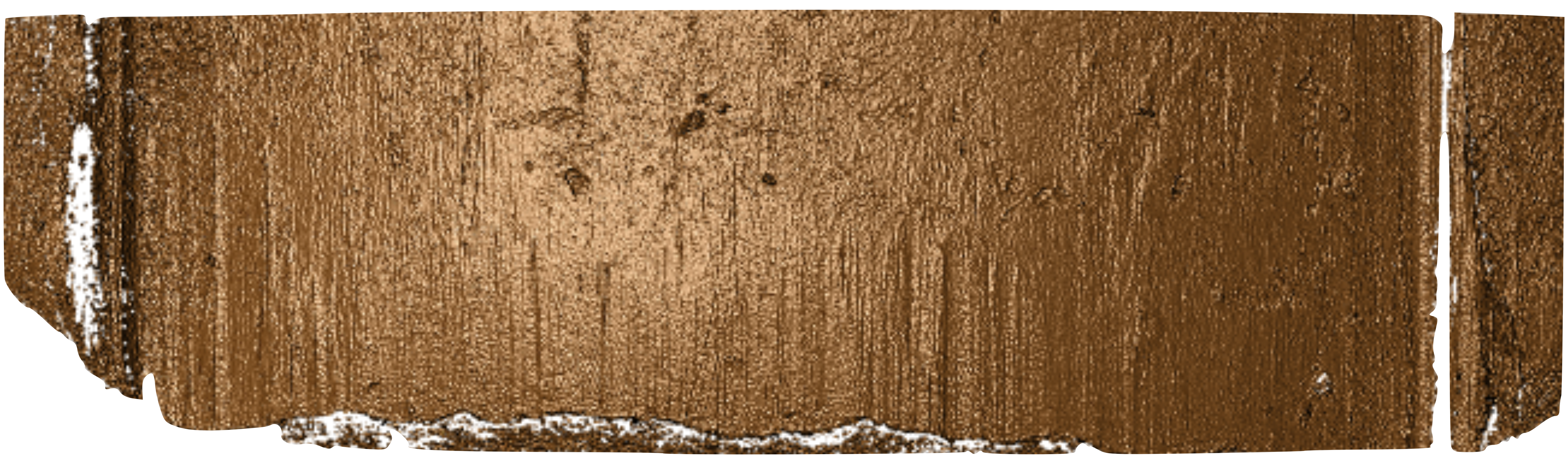
From raw scans to data for analysis

csafe
Center for Statistics and Applications in Forensic Evidence
ForensicStats.org

Statistical Analysis

Automatic matching score

Step 1: identify region suitable for matching



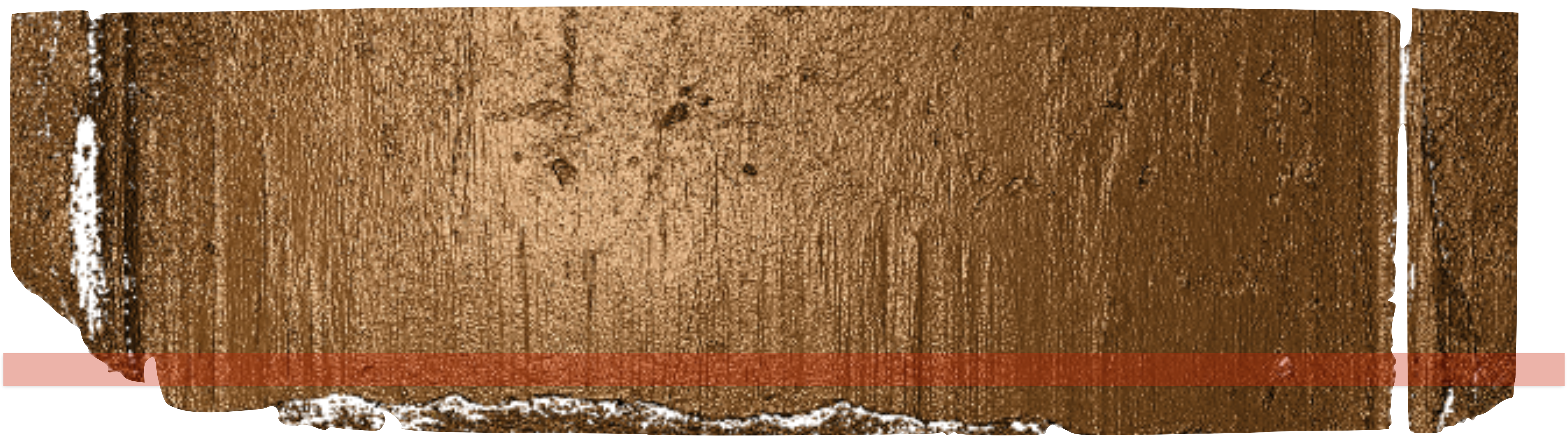
Region close to heel of bullet

Avoid break-off

Statistical Analysis

Automatic matching score

Step 1: identify region suitable for matching



Region close to heel of bullet

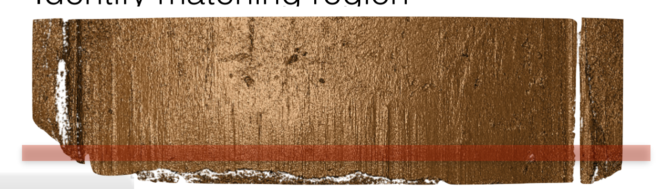
Avoid break-off

Statistical Analysis

Automatic matching score

Step 1b: from scan to crosscut

Identify matching region



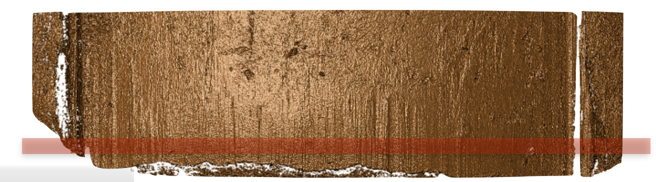
csafe
Center for Statistics and Applications in Forensic Evidence
ForensicStats.org

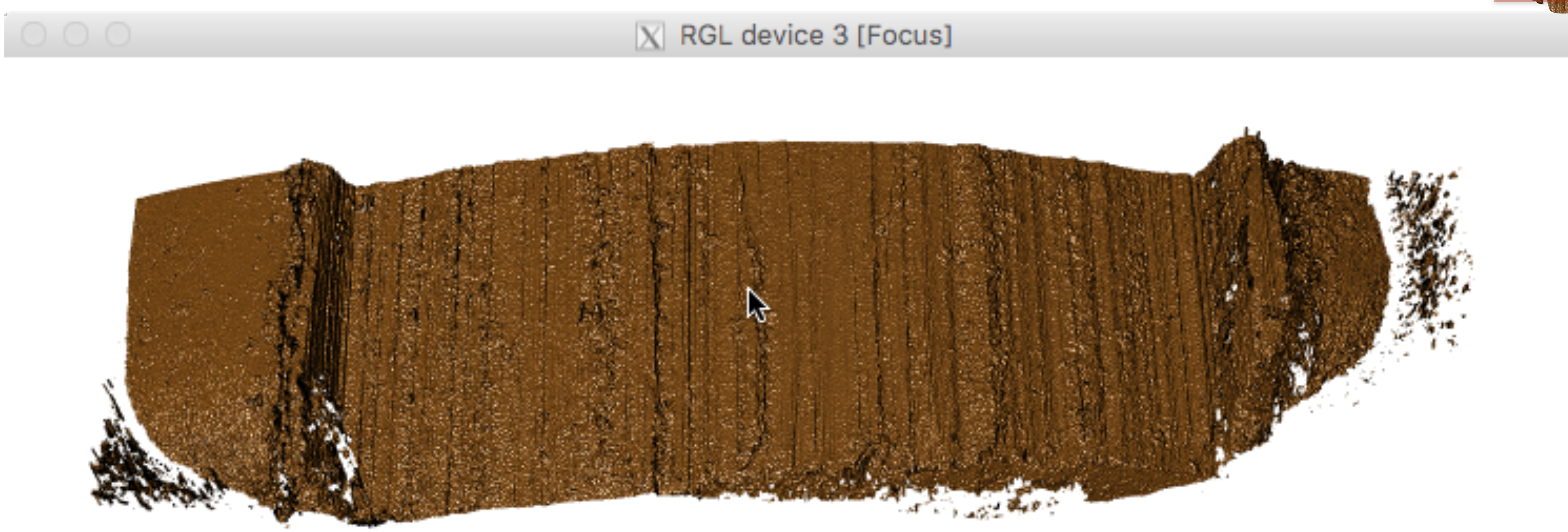
Statistical Analysis

Automatic matching score

Step 1b: from scan to crosscut

Identify matching region



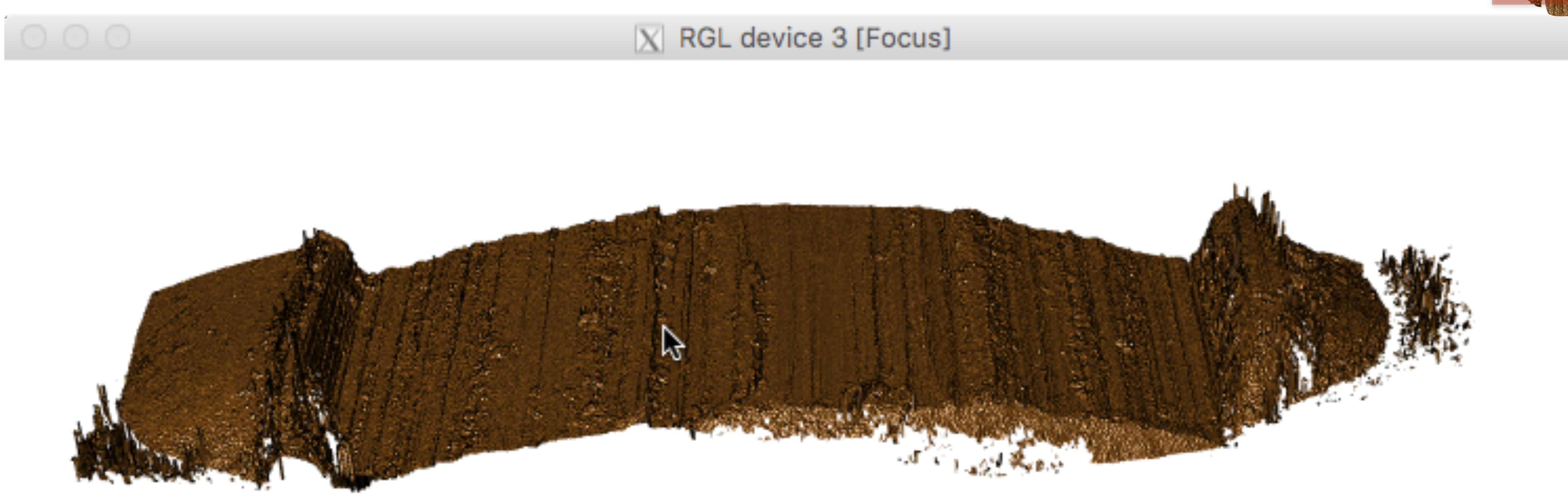


Statistical Analysis

Automatic matching score

Step 1b: from scan to crosscut

Identify matching region

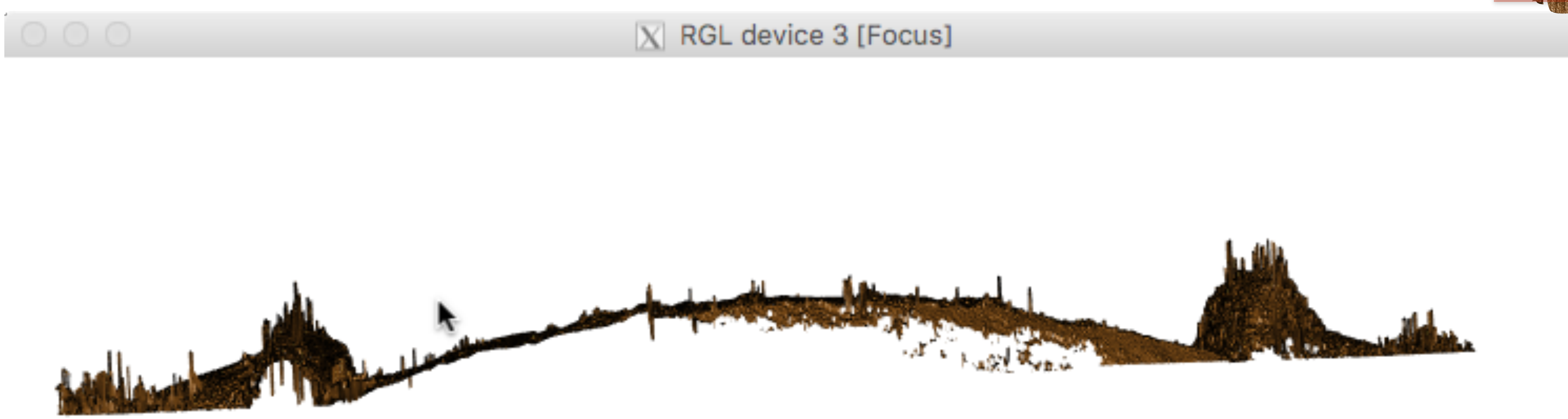


Statistical Analysis

Automatic matching score

Step 1b: from scan to crosscut

Identify matching region

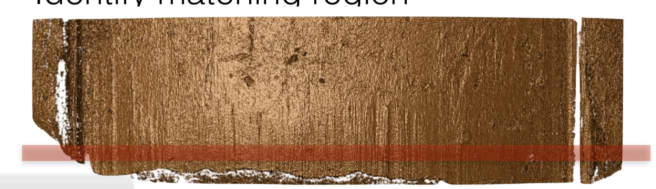


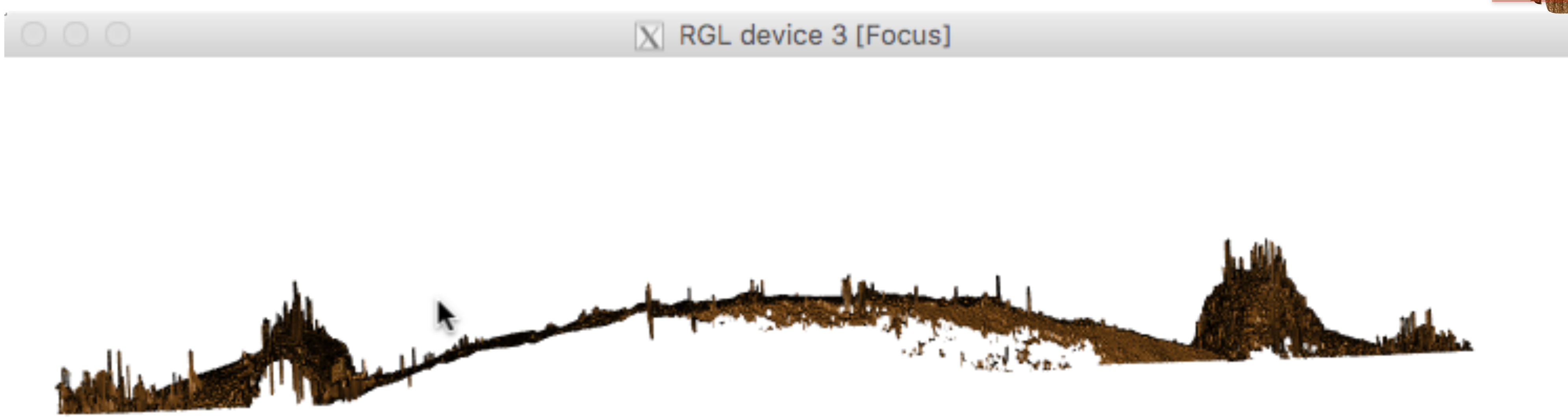
Statistical Analysis

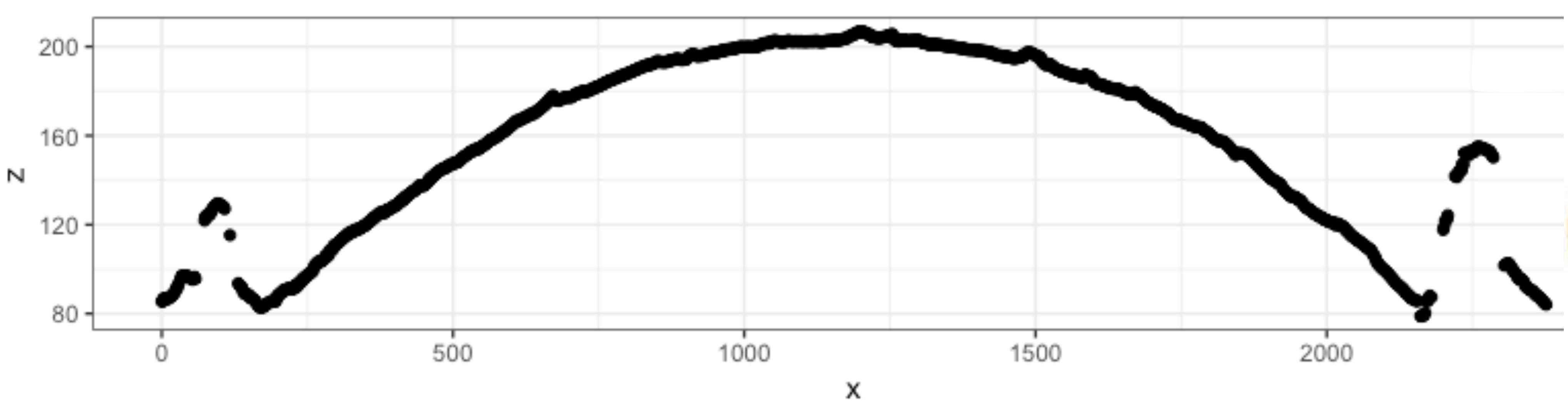
Automatic matching score

Step 1b: from scan to crosscut

Identify matching region





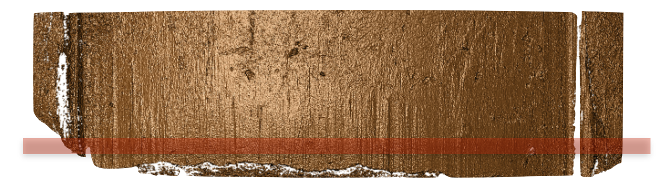


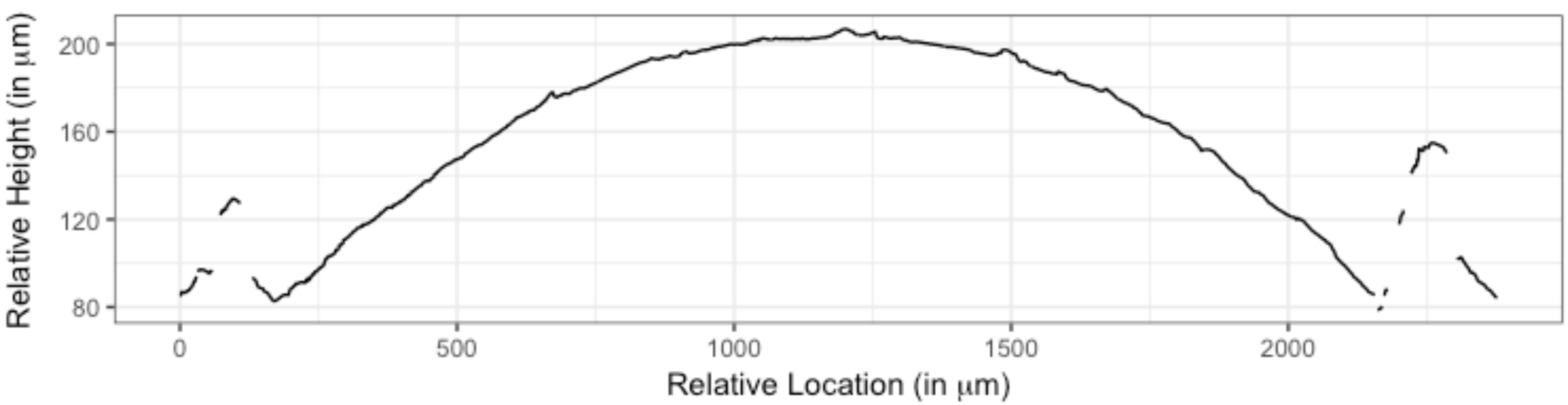
Statistical Analysis

Automatic matching score

Step 2: Identify groove locations

Identify matching region





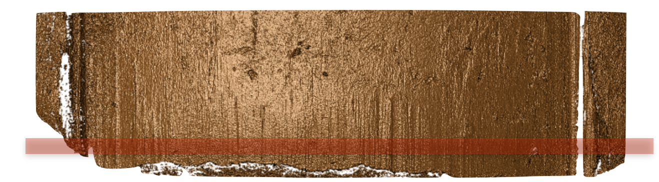
Shoulders (locations outside the grooves) are removed

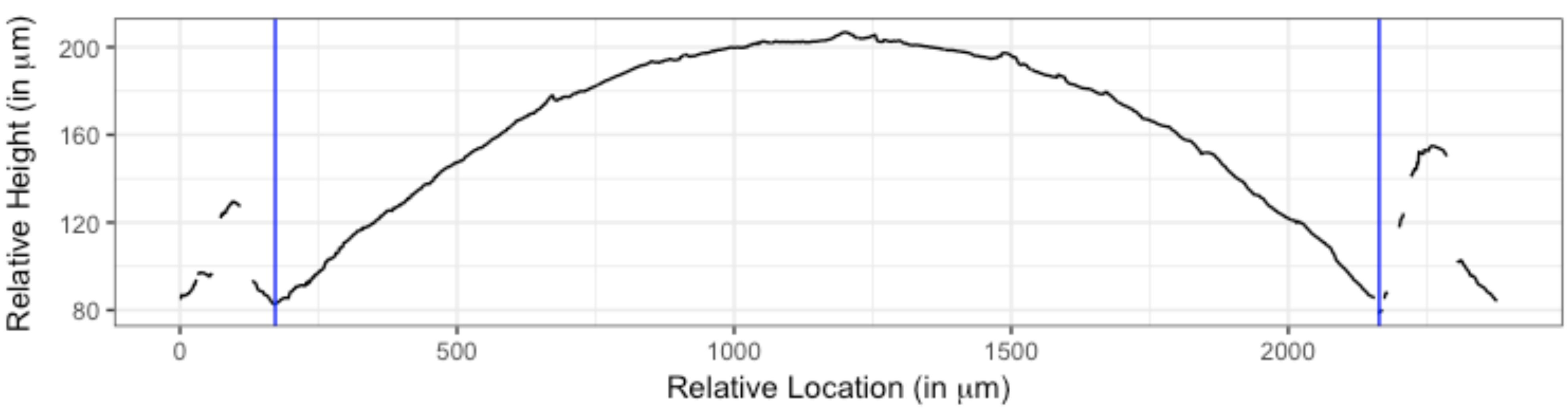
Statistical Analysis

Automatic matching score

Step 2: Identify groove locations

Identify matching region



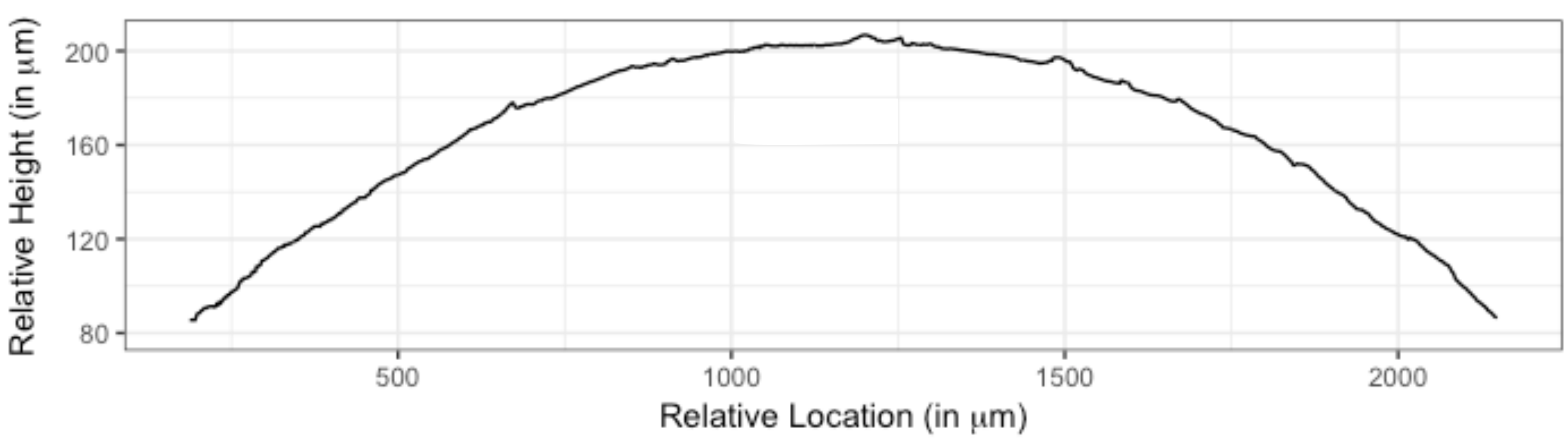


Shoulders (locations outside the grooves) are removed

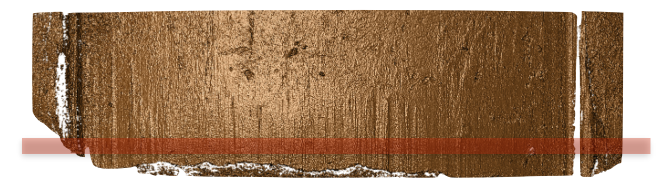
Statistical Analysis

Automatic matching score

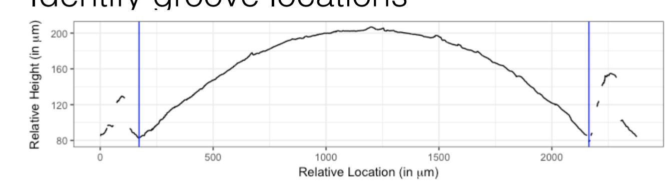
Step 3: Fit curvature



Identify matching region



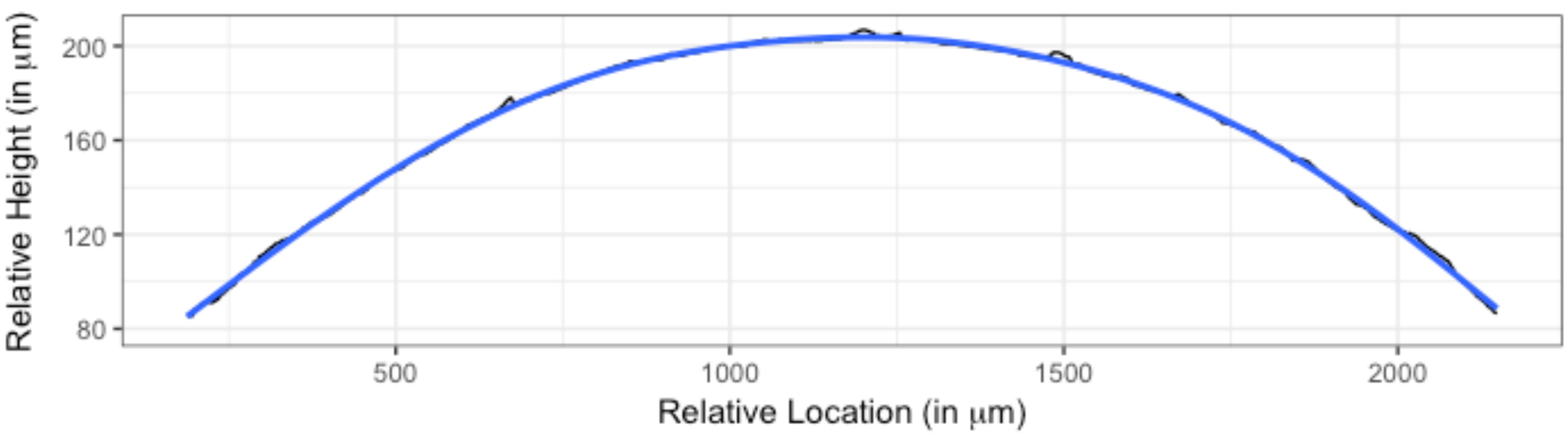
Identify groove locations



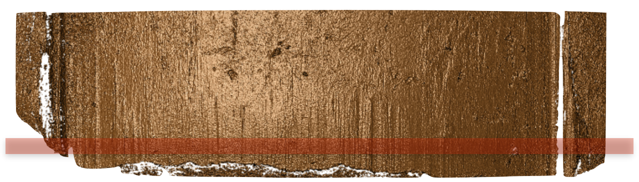
Statistical Analysis

Automatic matching score

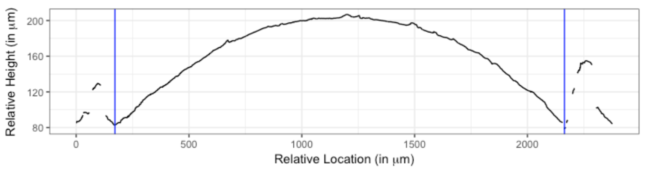
Step 3: Fit curvature



Identify matching region



Identify groove locations

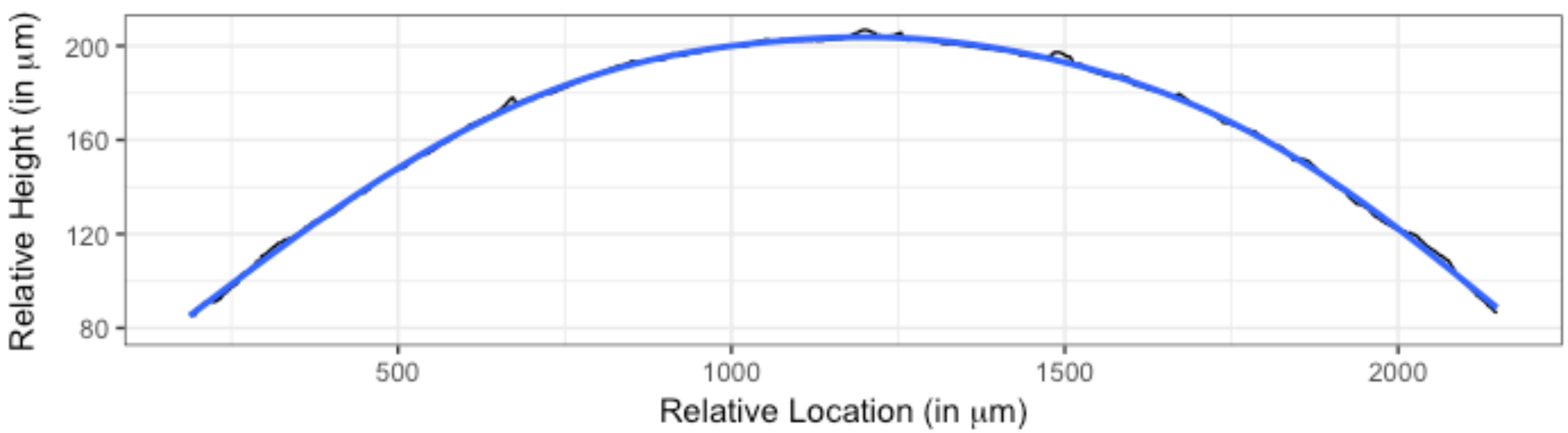


csafe
Center for Statistics and Applications in Forensic Evidence
ForensicStats.org

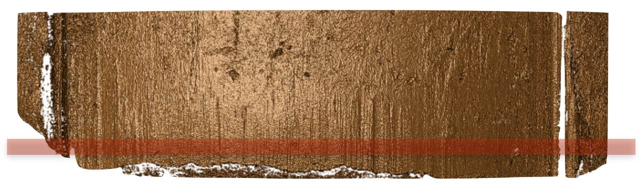
Statistical Analysis

Automatic matching score

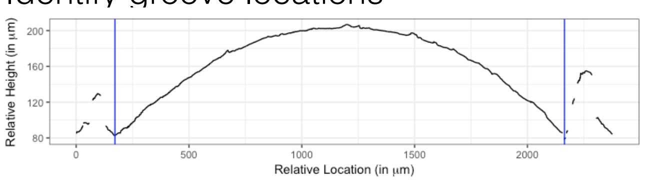
Step 3: Fit curvature & get signature



Identify matching region



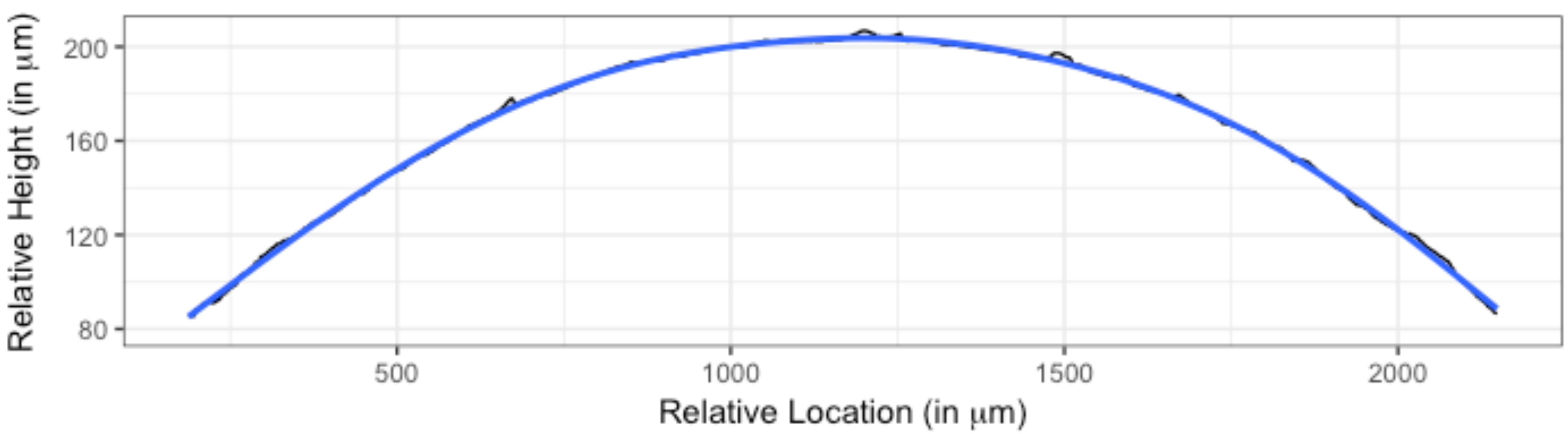
Identify groove locations



Statistical Analysis

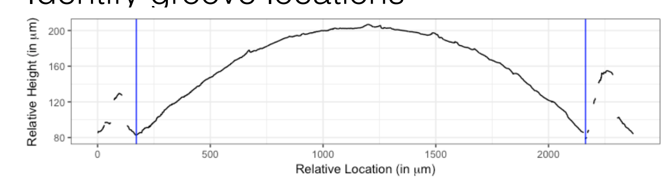
Automatic matching score

Step 3: Fit curvature & get signature



Identify matching region

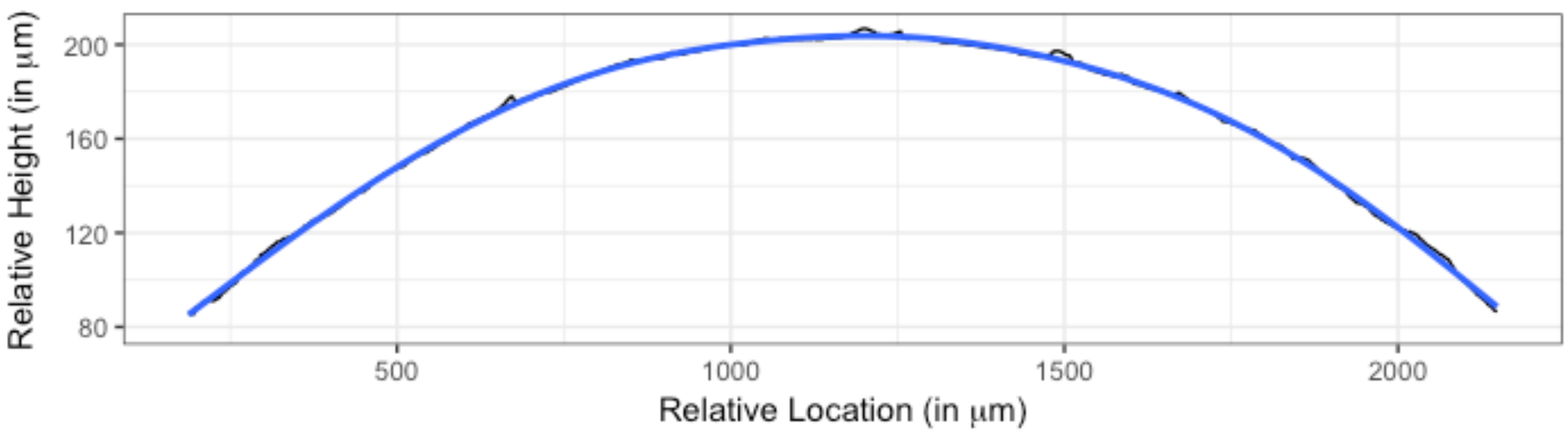
Identify groove locations

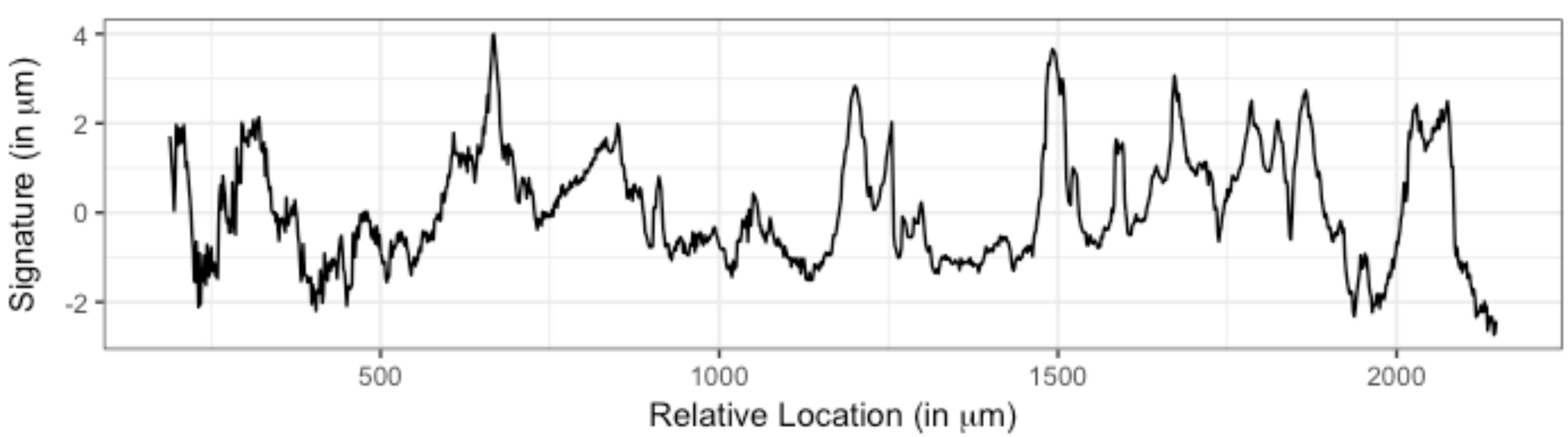


Statistical Analysis

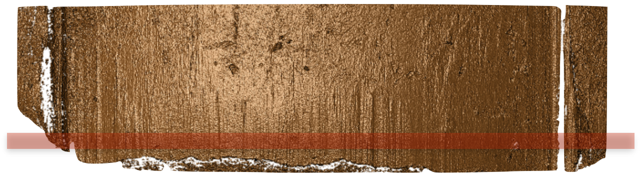
Automatic matching score

Step 3: Fit curvature & get signature

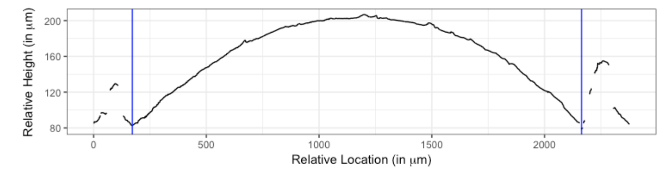




Identify matching region



Identify groove locations



csafe

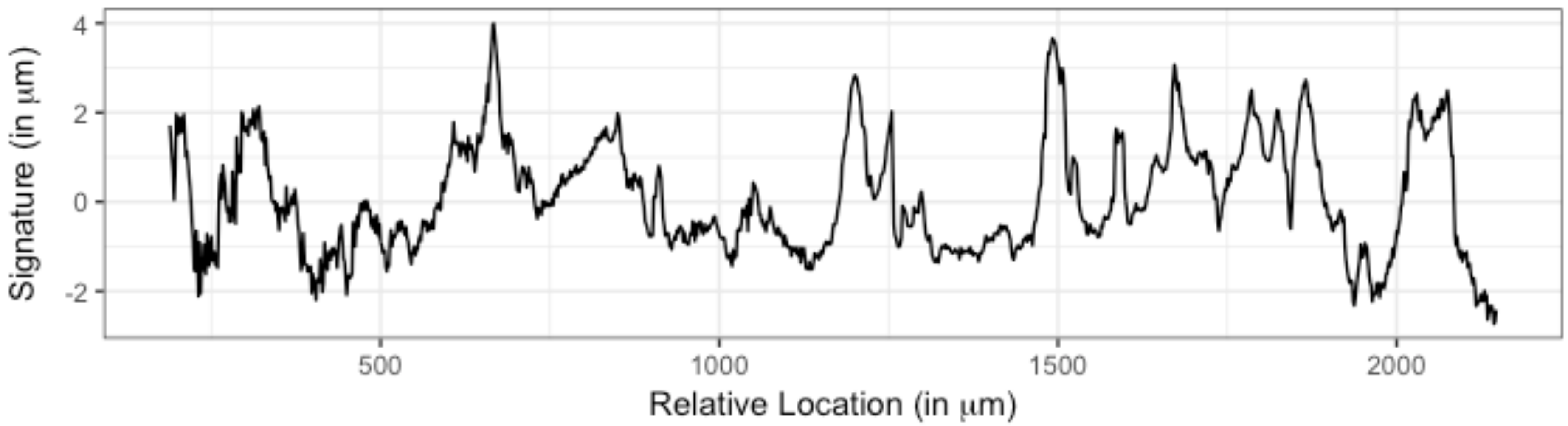
Center for Statistics and Applications in Forensic Evidence

ForensicStats.org

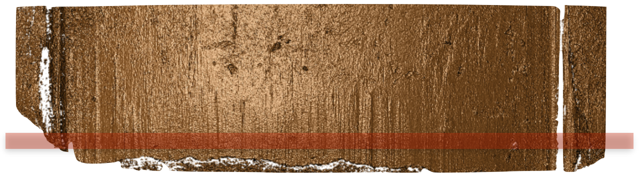
Statistical Analysis

Automatic matching score

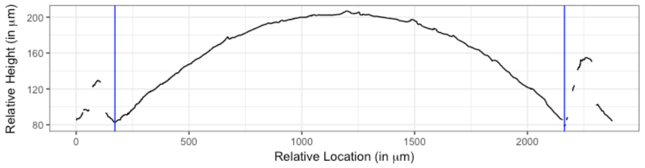
Step 4: Align signatures



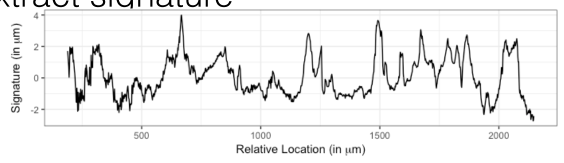
Identify matching region



Identify groove locations



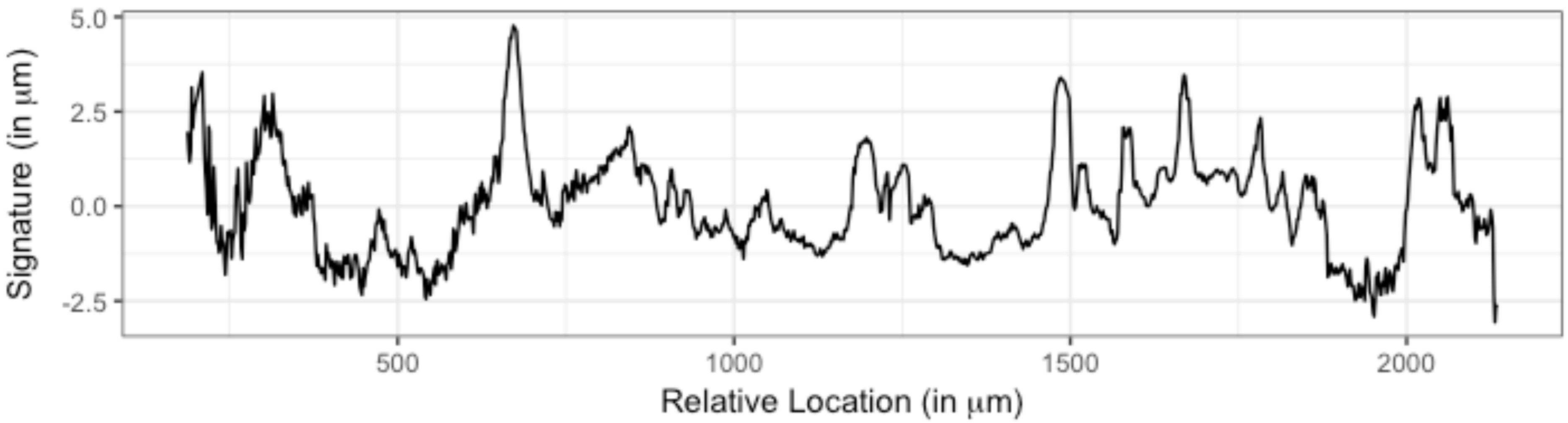
Extract signature

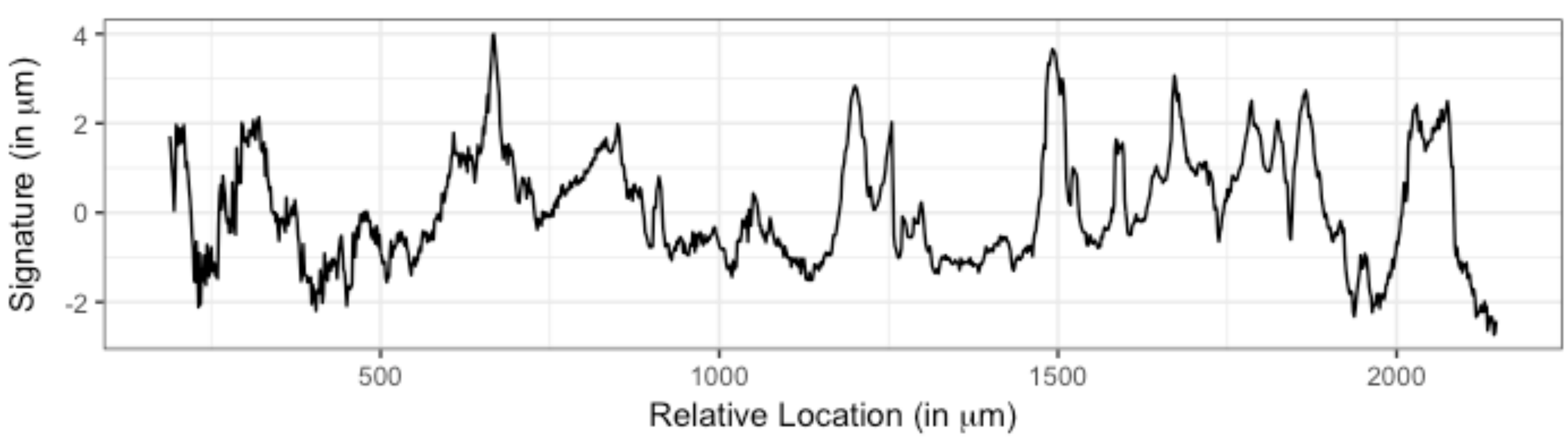


Statistical Analysis

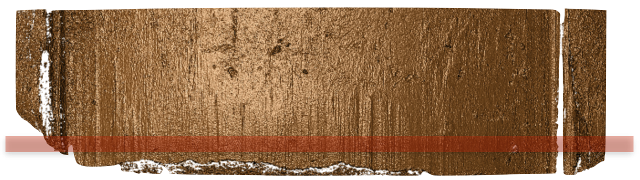
Automatic matching score

Step 4: Align signatures

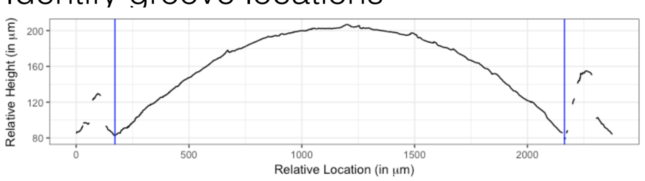




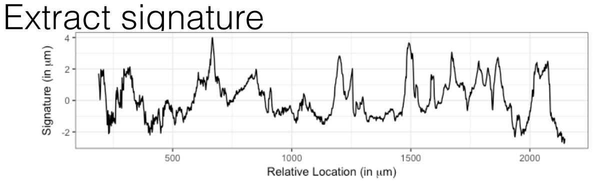
Identify matching region



Identify groove locations



Extract signature

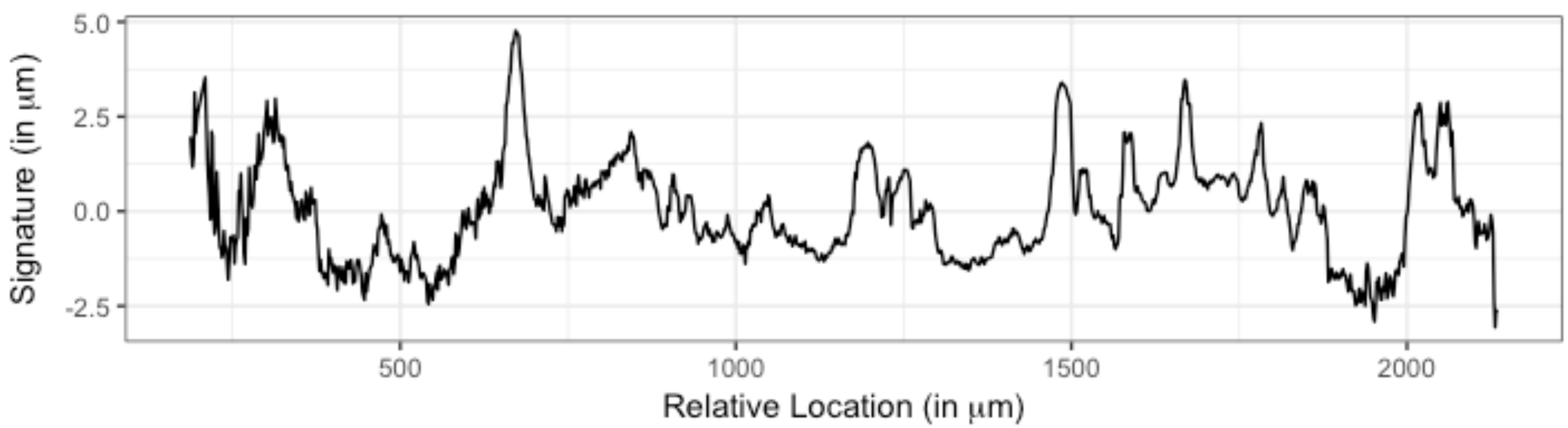


csafe
Center for Statistics and Applications in Forensic Evidence
ForensicStats.org

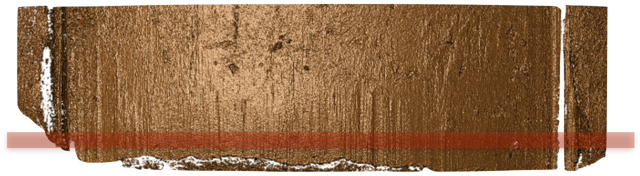
Statistical Analysis

Automatic matching score

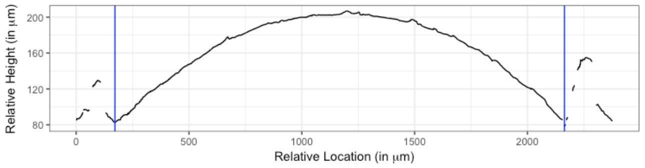
Step 4: Align signatures



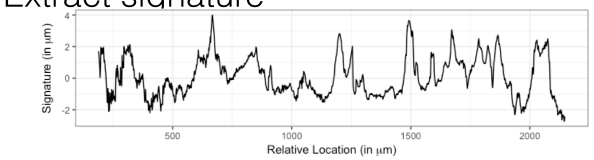
Identify matching region



Identify groove locations



Extract signature



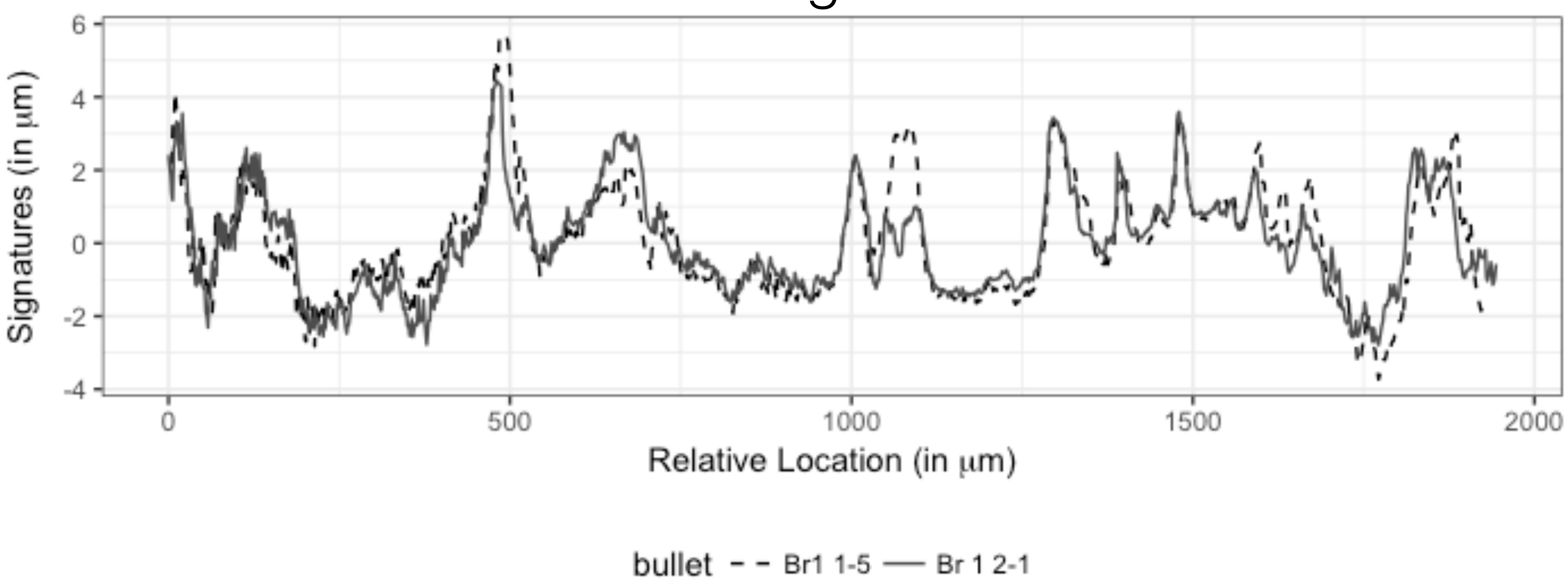
csafe
Center for Statistics and Applications in Forensic Evidence
ForensicStats.org

Statistical Analysis

Automatic matching score

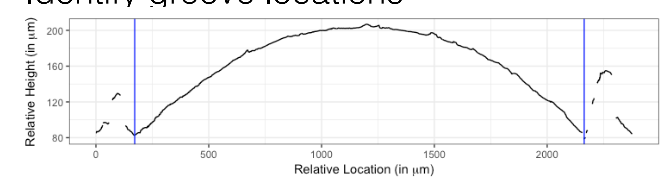
Step 4: Align signatures

Horizontal shifts to find best alignment

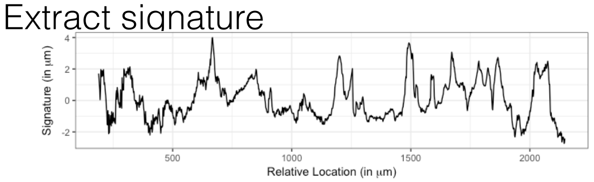


Identify matching region

Identify groove locations



Extract signature



csafe
Center for Statistics and Applications in Forensic Evidence
ForensicStats.org

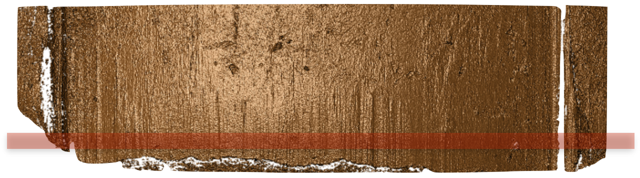
Statistical Analysis

Automatic matching score

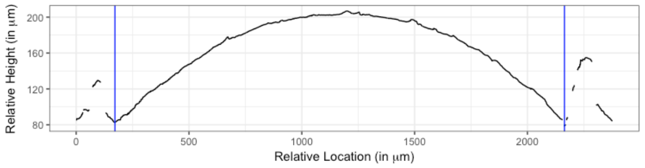
Step 5: Extract features

Feature should distinguish between a match and a non-match

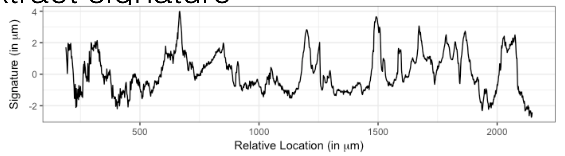
Identify matching region

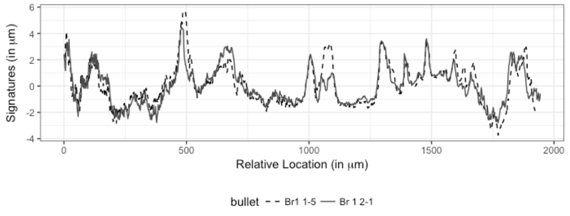


Identify groove locations



Extract signature





Statistical Analysis

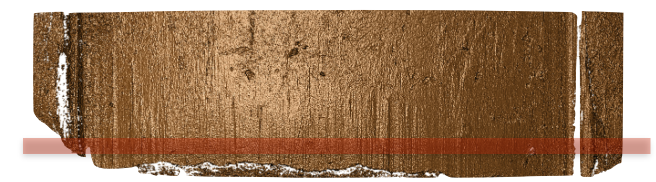
Automatic matching score

Step 5: Extract features

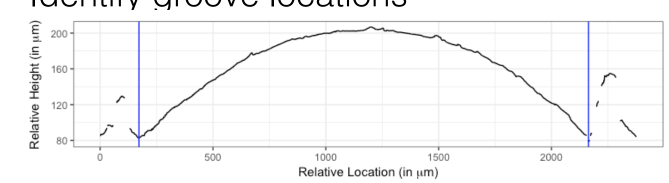
Feature should distinguish between a match and a non-match

★ # matches/mis-matches of peaks & valleys

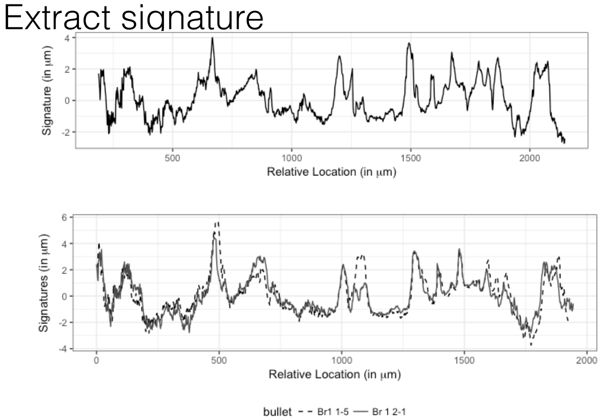
Identify matching region



Identify groove locations



Extract signature



Statistical Analysis

Automatic matching score

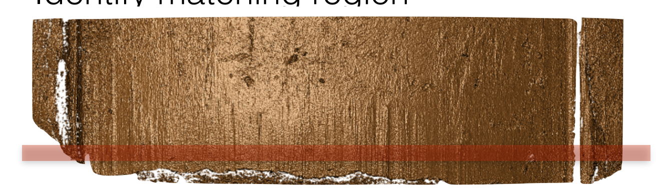
Step 5: Extract features

Feature should distinguish between a match and a non-match

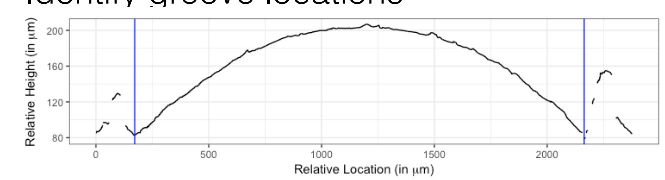
★ # matches/mis-matches of peaks & valleys

★ # consecutive matches/mis-matches(cms)

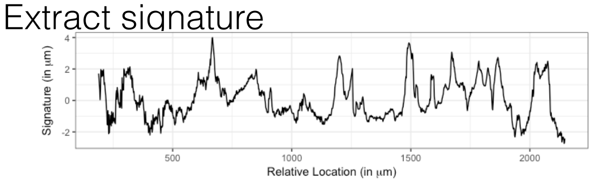
Identify matching region

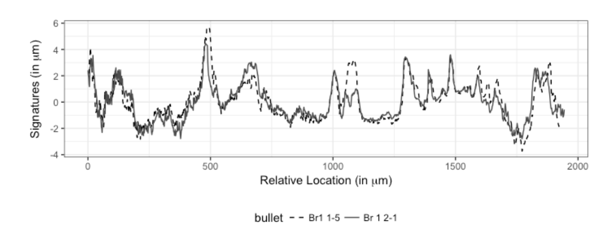


Identify groove locations



Extract signature





Statistical Analysis

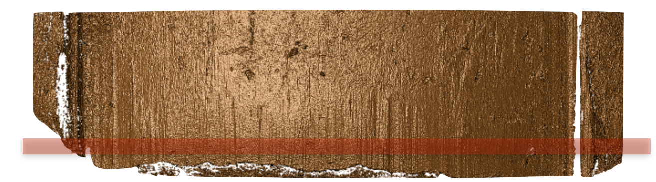
Automatic matching score

Step 5: Extract features

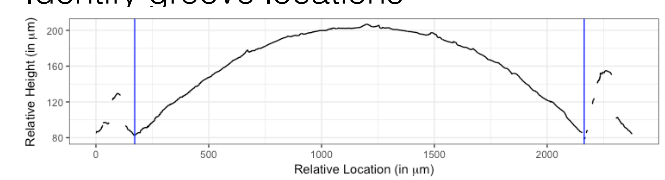
Feature should distinguish between a match and a non-match

- ★ # matches/mis-matches of peaks & valleys
- ★ # consecutive matches/mis-matches(cms)
- ★ depth of peaks/valleys

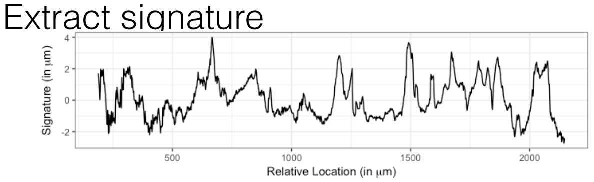
Identify matching region

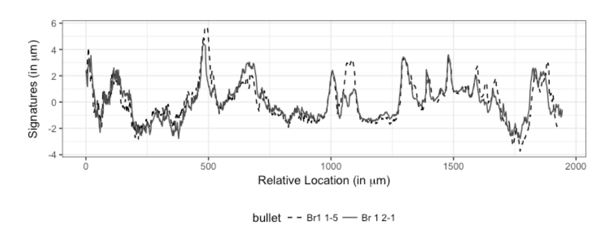


Identify groove locations



Extract signature





Statistical Analysis

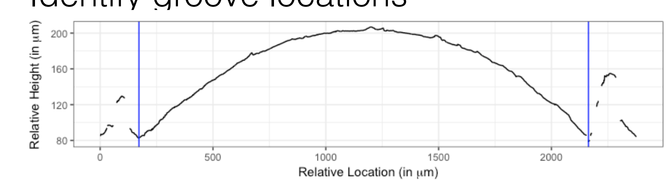
Automatic matching score

Step 5: Extract features

Identify matching region

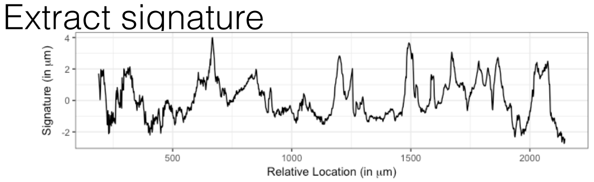
Feature should distinguish between a match and a non-match

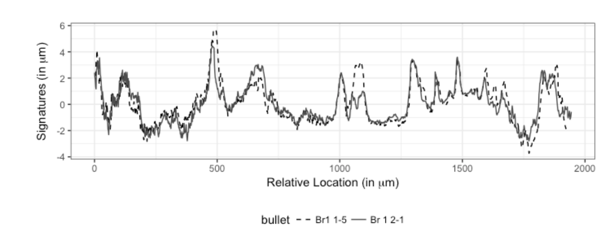
Identify groove locations



- ★ # matches/mis-matches of peaks & valleys
- ★ # consecutive matches/mis-matches(cms)
- ★ depth of peaks/valleys
- ★ area between the signatures

Extract signature





Statistical Analysis

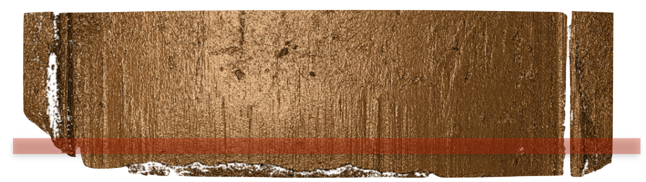
Automatic matching score

Step 5: Extract features

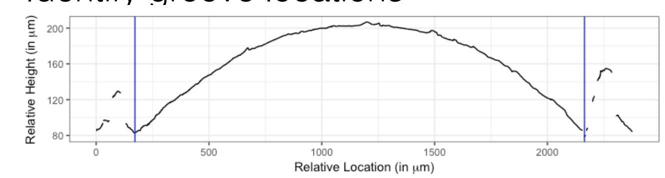
Feature should distinguish between a match and a non-match

- ★ # matches/mis-matches of peaks & valleys
- ★ # consecutive matches/mis-matches(cms)
- ★ depth of peaks/valleys
- ★ area between the signatures
- ★ cross-correlation function

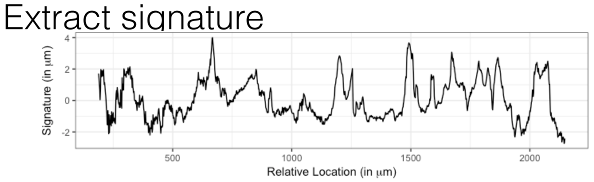
Identify matching region

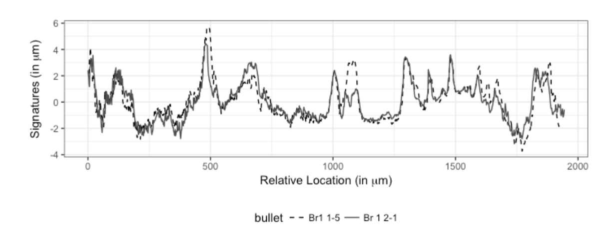


Identify groove locations



Extract signature





Features & comparisons

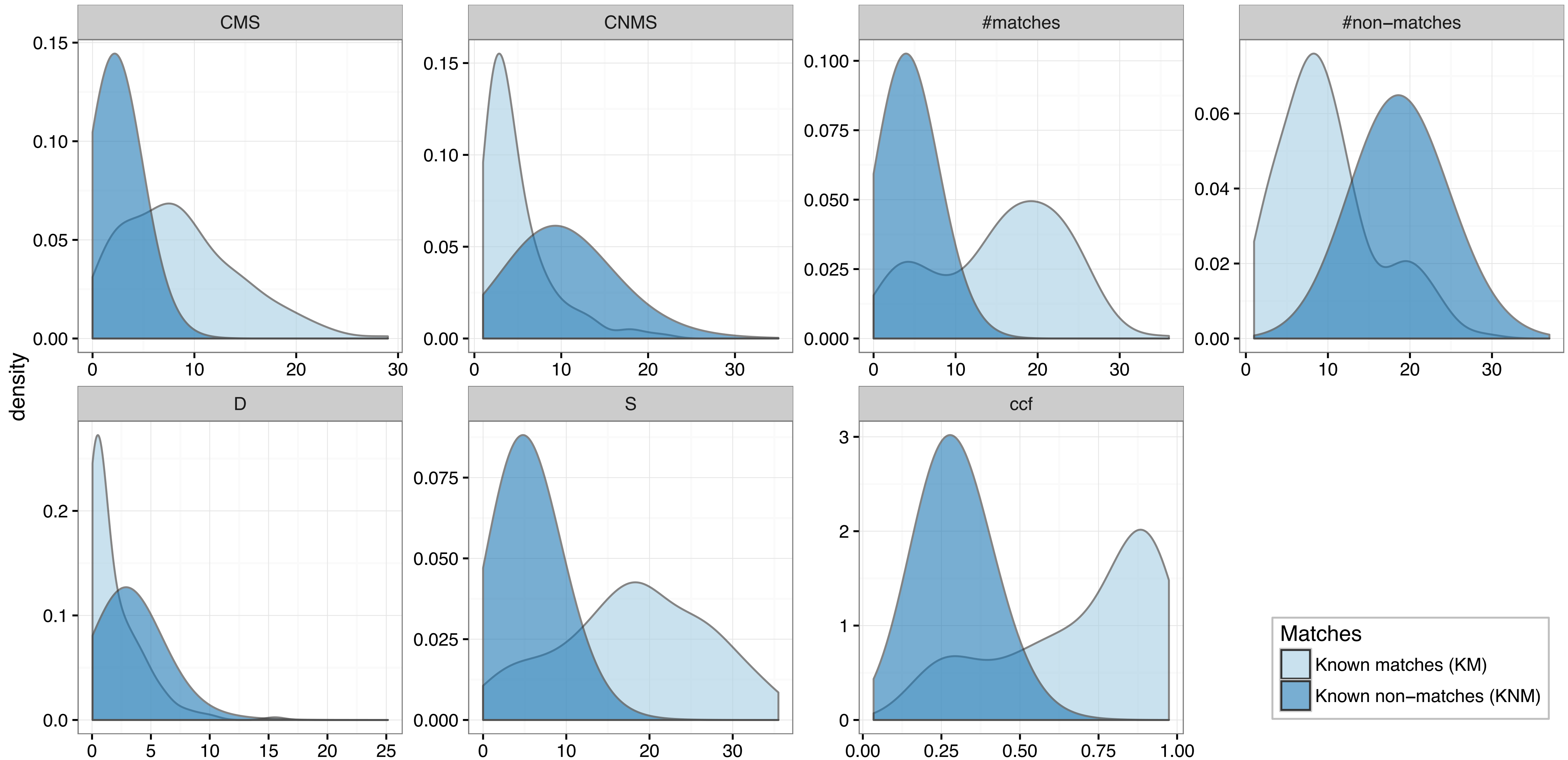
barrel1	bullet1	land1	barrel2	bullet2	land2	ccf	D	matches	mismatches	cms	non_cms
4	2	1	10	1	3	0.26	0.00	2.16	20.49	0.54	8.19
4	2	1	2	1	2	0.30	0.00	2.31	19.41	0.58	8.25
Unk	G	1	4	2	1	0.58	0.00	3.48	18.94	1.74	8.42
4	**1**	**3**	**4**	**2**	**1**	**0.85**	**0.00**	**6.14**	**16.41**	**2.23**	**4.24**
4	2	1	10	2	5	0.38	0.00	2.37	18.61	1.18	6.86
4	2	1	6	2	6	0.32	0.00	4.01	16.43	2.29	4.98
4	2	1	5	2	2	0.24	0.00	2.24	18.02	0.56	5.00

Features & comparisons

barrel1	bullet1	land1	barrel2	bullet2	land2	ccf	D	matches	mismatches	cms	non_cms	Known match
4	2	1	10	1	3	0.26	0.00	2.16	20.49	0.54	8.19	FALSE
4	2	1	2	1	2	0.30	0.00	2.31	19.41	0.58	8.25	FALSE
Unk	G	1	4	2	1	0.58	0.00	3.48	18.94	1.74	8.42	FALSE
4	**1**	**3**	**4**	**2**	**1**	**0.85**	**0.00**	**6.14**	**16.41**	**2.23**	**4.24**	**TRUE**
4	2	1	10	2	5	0.38	0.00	2.37	18.61	1.18	6.86	FALSE
4	2	1	6	2	6	0.32	0.00	4.01	16.43	2.29	4.98	FALSE
4	2	1	5	2	2	0.24	0.00	2.24	18.02	0.56	5.00	FALSE

Features & comparisons

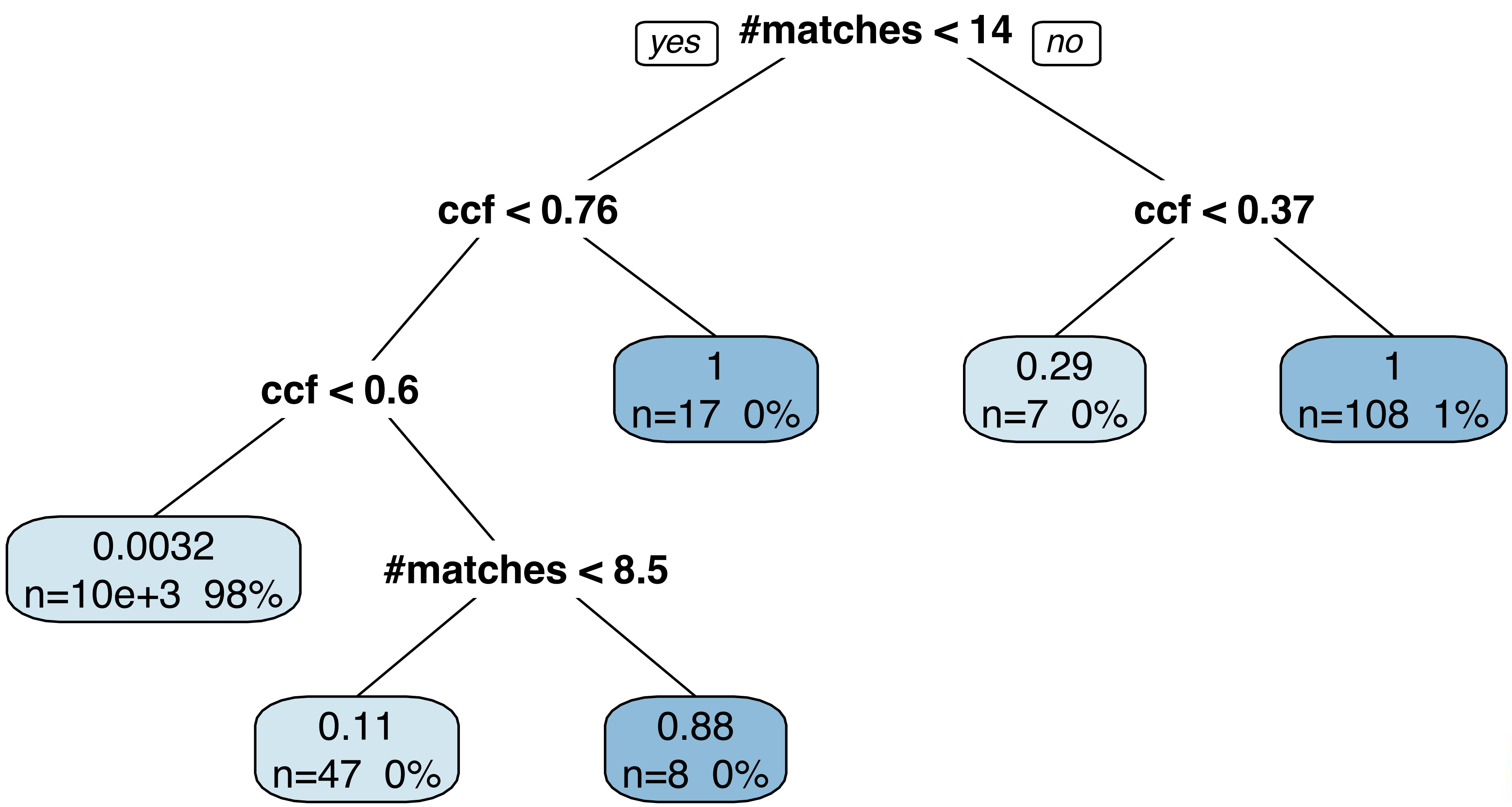
All features show distinction between known matches and known non-matches



Combining Features

Decision Tree

★ Decision Tree (1984 Breiman)

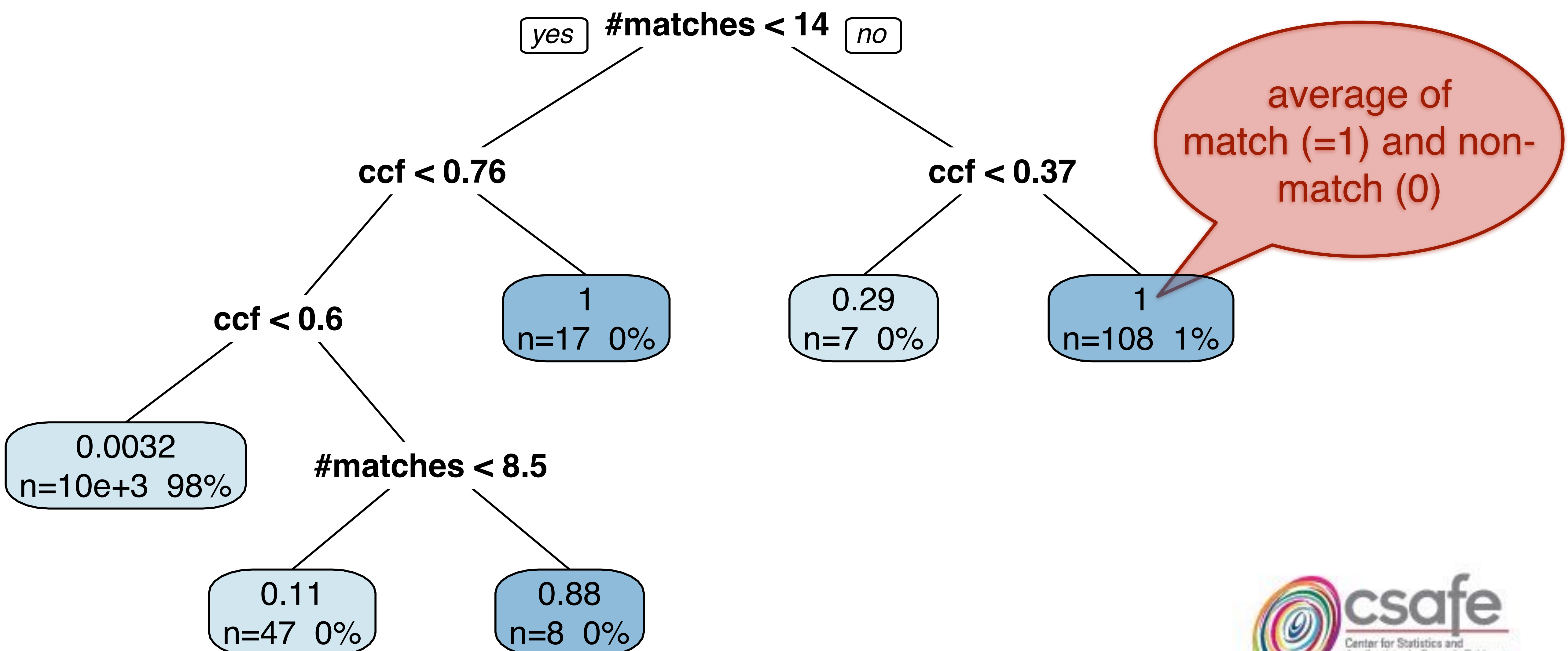


csafe
Center for Statistics and Applications in Forensic Evidence
ForensicStats.org

Combining Features

Decision Tree

- ★ Decision Tree (1984 Breiman)

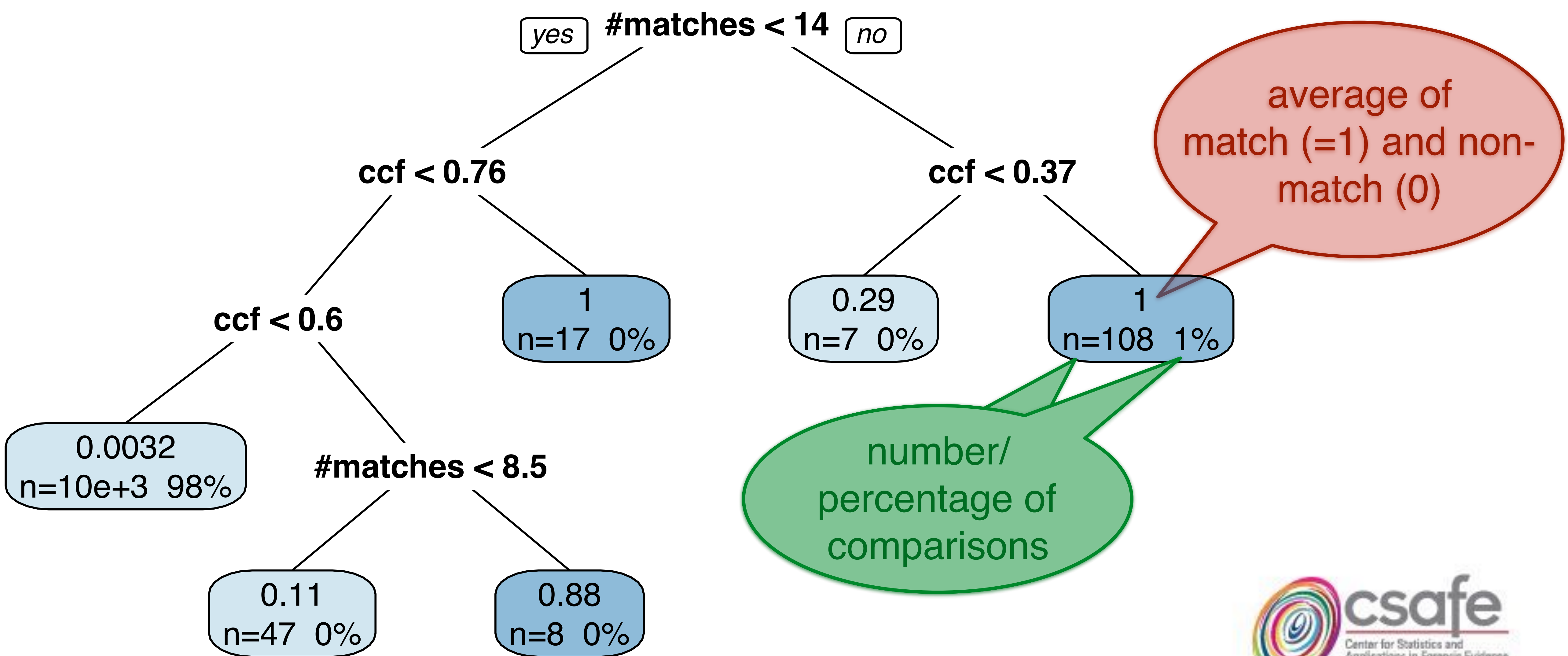


csafe
Center for Statistics and Applications in Forensic Evidence
ForensicStats.org

Combining Features

Decision Tree

★ Decision Tree (1984 Breiman)

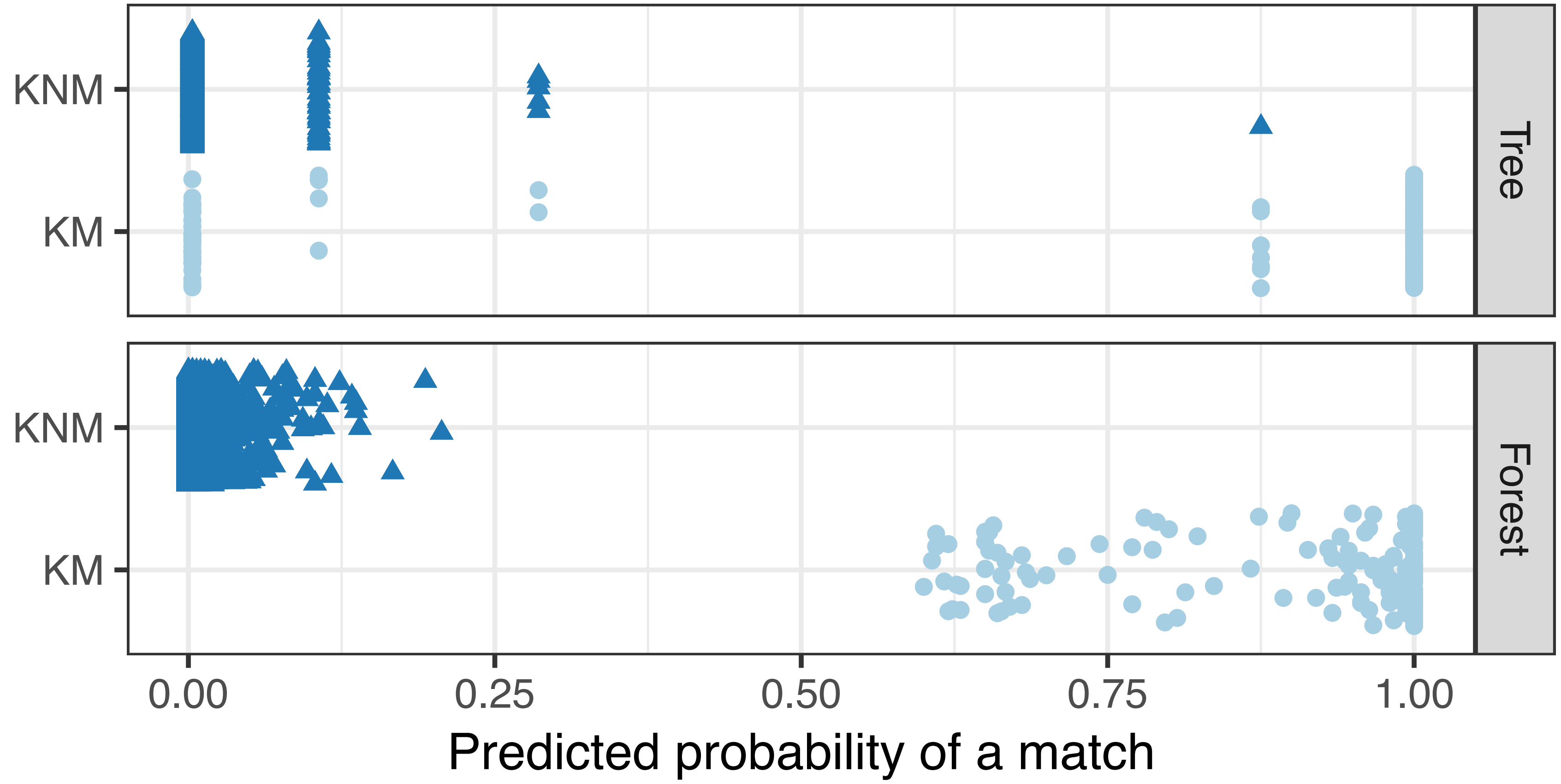


csafe
Center for Statistics and Applications in Forensic Evidence
ForensicStats.org

Combining Features

Random Forest

★ Combination of 500 Decision Trees



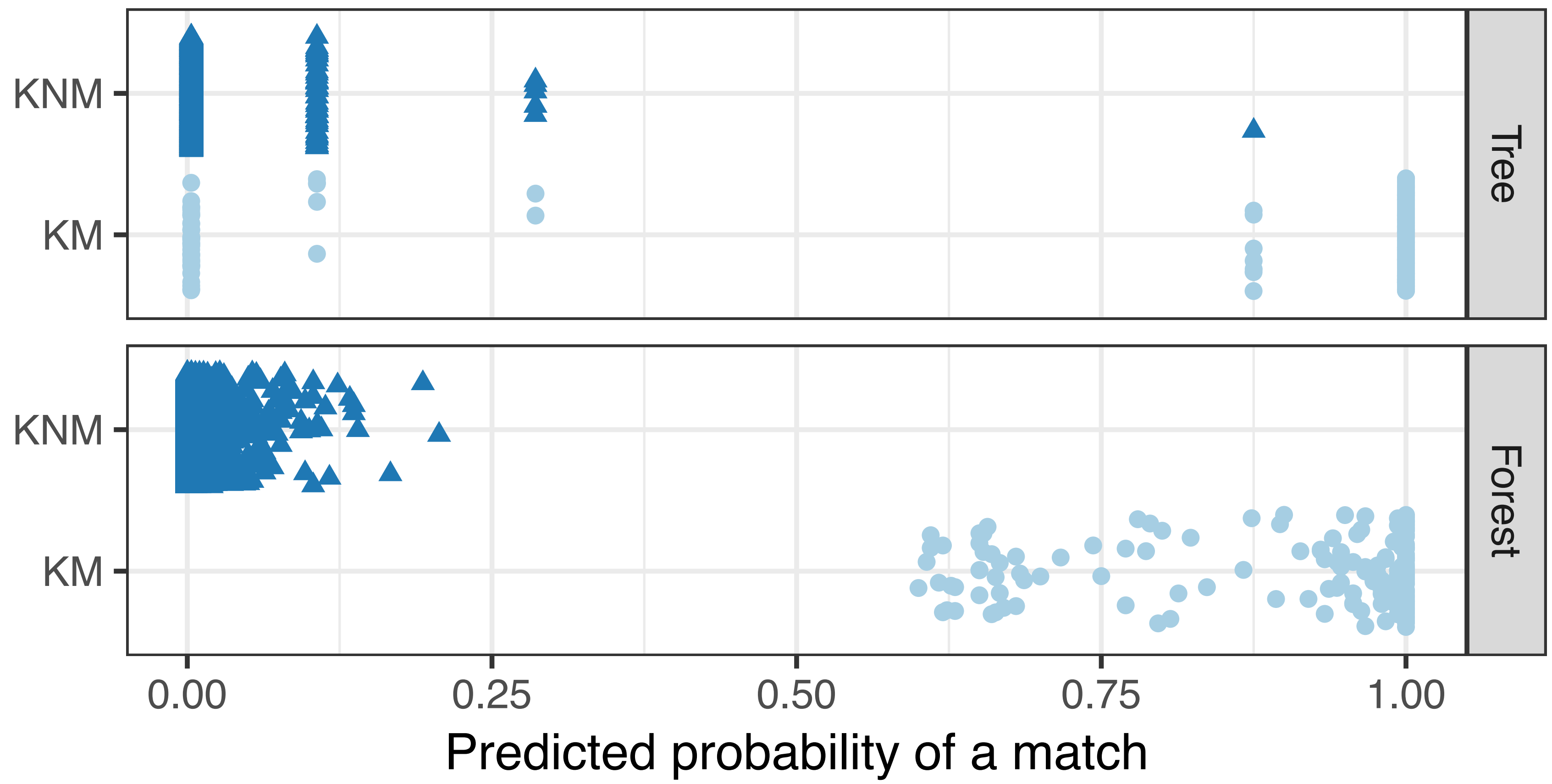
★ one false match (score too high) for tree, several false non-matches (scores too low)

★ no errors for Random Forest score, good separation

Combining Features

Random Forest

- Combination of 500 Decision Trees



Automatic matching of bullet land impressions, Annals of Applied Statistics, Eric Riemer Hare, Heike Hofmann, and Alicia Carriquiry

Algorithmic approaches to match degraded land impressions
Eric Hare; Heike Hofmann; Alicia Carriquiry
Law, Probability and Risk, Volume 16, Issue 4, 1 December 2017, 203–221, https://doi.org/10.1093/lpr/mgx018

- one false match (score too high) for tree, several false non-matches (scores too low)
- no errors for Random Forest score, good separation

csafe
Center for Statistics and Applications in Forensic Evidence
ForensicStats.org

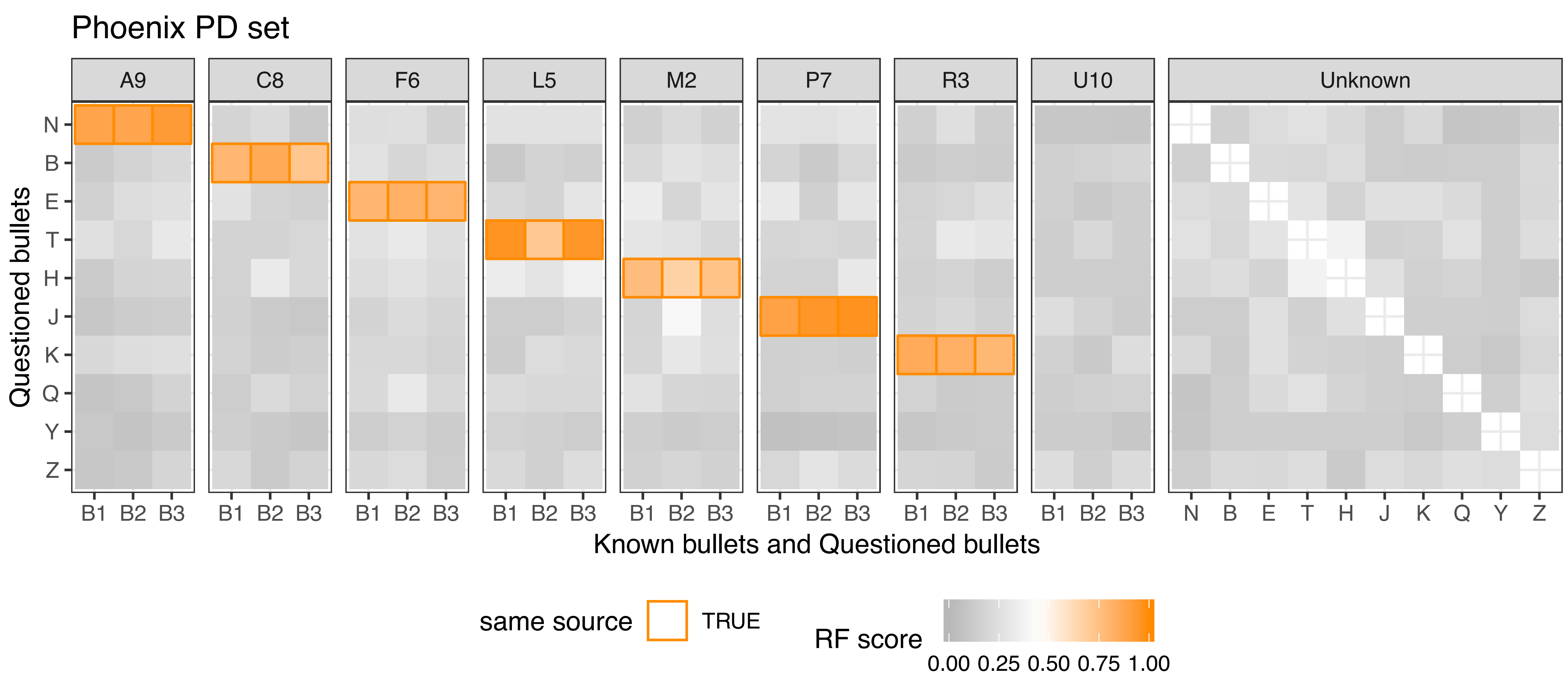
Case validation

csafe
Center for Statistics and Applications in Forensic Evidence
ForensicStats.org

Validating the RF score

Random Forest

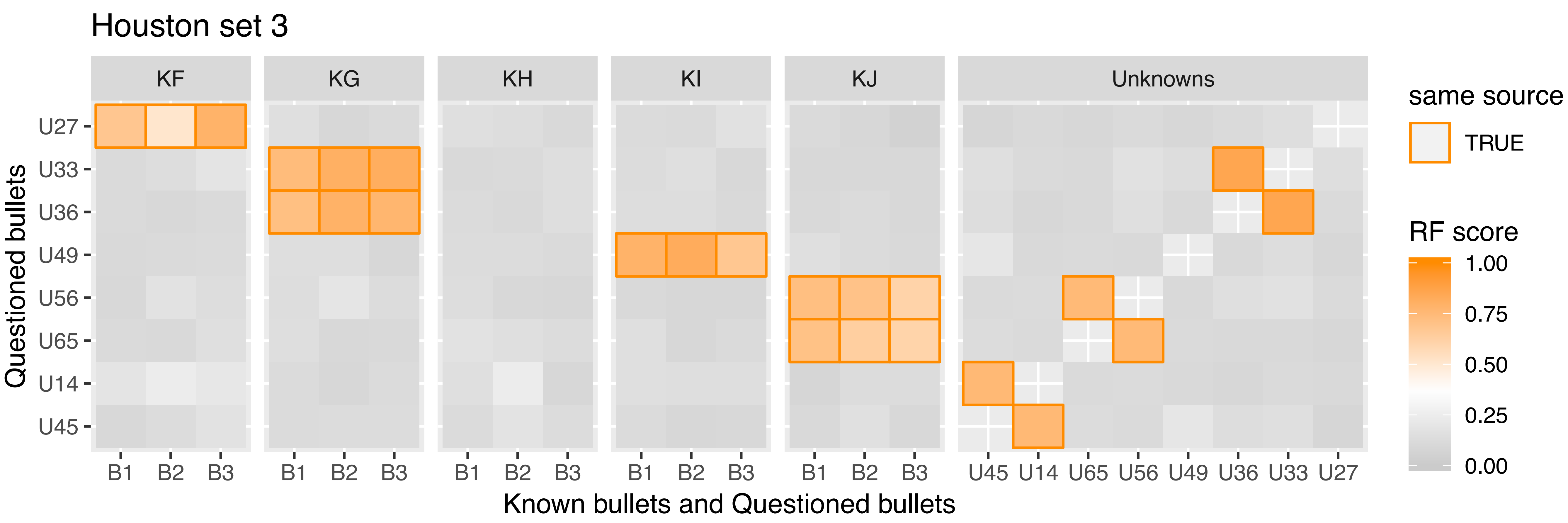
- ★ Phoenix PD Study (Tyler Klep)
  - ★ known matches: eight barrels with three test fires each
  - ★ ten questioned bullets



Validating the RF score

Random Forest

- ★ Houston Test 1 (Melissa McNally)
  - ★ known matches: five barrels with three test fires each
  - ★ eight questioned bullets



Conclusions

- Preliminary results are promising
- Rewarding to work on project with obvious high impact
- Challenges at every step:
  - data collection, data wrangling, feature extraction, modeling
  - theoretical foundations, knowledge transfer to labs

Why is this Data Science?

Result is combination of

- ★ Data management: scans are large (~15 MB each), organizational structure for quality checks, re-scans, …
- ★ Computationally intensive - data processing, supervised learning methods
- ★ Applied Statistics: exploratory analysis, feature extraction, distributional properties, error analysis

csafe
Center for Statistics and Applications in Forensic Evidence
ForensicStats.org

Thank You!

Questions?

Heike Hofmann (hofmann@iastate.edu, @heike_hh)
ISU CSAFE bullet team

csafe
Center for Statistics and Applications in Forensic Evidence
ForensicStats.org