A158 Current Forensic Models May Make Inappropriate Statistical Assumptions

Max M. Houck, MA, West Virginia University, 1600 University Avenue, 208 Oglebay Hall, Morgantown, WV 26506-6217; Fayola Peters, BS, and Tim Menzies, PhD, West Virginia University, Department of Computer Science, Morgantown, WV 26506; and Simon W. Lewis, PhD, Curtin University of Technology, Department of Applied Chemistry, GPO Box U1987, Perth, 6845, AUSTRALIA*

After attending this presentation, attendees will learn the results of a study designed to show that the simple method of nearest neighbor classification performs comparatively with the current complex statistical models used in the evaluation of trace evidence such as glass and paint.

This presentation will impact the forensic science community by exposing attendees to systematic errors associated with standard statistical models, as they are applied to evidence evaluation.

Forensic interpretation models try to determine the extent to which trace evidence found at a crime scene and on a suspect could have come from a common source. A desired property of such interpretation models is stability; that is, minor changes to the data do not radically change the interpretation. If a model is unstable, then small errors in the data collection can confuse the interpretation. Stability is an important property since forensic data is collected in less-than-ideal situations.

It is shown that several well-known forensic glass fragment interpretation models (Seheult (1978), Grove (1980), Evett (1995), and Walsh (1995) are not stable. More precisely, they exhibit a *zig-zag* behavior: pushing it a little drives it one way and pushing it a little more drives it in the opposite direction. To test if a model zig-zags:

1. Select model inputs at random between their minimum and maximum values,
2. Record the output interpretation,
3. Sort the interpretations by the order of the inputs,

Look for small changes to inputs that lead to large changes in the interpretation. Specifically, look for two small successive increases in any input that leads to a large increase, then decrease in the likelihood interpretation.

Zig-zags were tested for in the Seheult, Grove, Evett, and Walsh models, coding up each model to include not just the crime scene data (for example, refractive index) but also the tuning attributes recommended by the authors of the models. These tuning values come from prior historical or laboratory studies. For example, the Evett and Walsh models derive these values via surveys of different laboratory results. While a useful product of prior research, these tuning parameters can critically and inappropriately affect the interpretation if the tuning values no longer apply in the new situation.

Many zig-zags were found in these models. For example, in the Seheult model, if refractive index of glass at a crime scene is measured at 1.49, a change to 1.50, and then to 1.52 changes interpretation dramatically from 2% to 86% then back to 28%. At four decimal places, the changes are no less dramatic:

RI Ranges	Interpretation%
1.4000 - 1.4236	36
1.4236 - 1.4437	87
1.4437 - 1.4623	22
1.4623 - 1.4803	100
1.4803 - 1.5028	75
1.5028 - 1.5254	36
1.5254 - 1.5441	16
1.5441 - 1.5656	75
1.5656 - 1.5818	22
1.5818 - 1.5999	22

These kinds of changes occur out to 16 decimal places of refractiveindex data. Similar zig-zags can be seen in the Grove, Evett, and Walsh models.

There are two disturbing aspects of these results. Firstly, it has not been previously reported, which raises the issue of whether or not forensic science research checks for model stability. Secondly, it is not true that more recent models (written in the 1990s) are more stable than the older ones (written in the 1970s). In this regard, it is not clear that forensic models are improving their stability over time.

The resolution of this model is the focus of the current research. The authors wish to understand the trade-off between stability and performance: if a model is too stable, it freezes up and cannot respond appropriately to new inputs. New classes of interpretation models that are *stability-aware* are currently being explored. Preliminary results with this approach are encouraging.

Trace, Interpretation, Statistical Models