CLASSIFICATION OF STREET DRUGS USING NMR SPECTROSCOPY AND CHEMOMETRIC ANALYSIS WITH UNCERTAINTY ESTIMATION

National Institute of Standards and Technology
U.S. Department of Commerce

Werickson Fortunato de Carvalho Rocha, Aaron A. Urbas, David A. Sheen, Katrice A. Lippa

Chemical Sciences Division, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA
E-mail: wfr1@nist.gov

Introduction

NMR spectroscopy has become a useful tool for the forensic analysis because it can be used to analyze samples in the solid, liquid, and gas phases. The high reproducibility, sensitivity, and selectivity of this technique facilitates quantitative and qualitative analysis of unknown substances of importance to law enforcement activities. Due to the similarity between NMR spectra of similar substances and the large size of NMR spectra, the use of chemometric tools with uncertainty estimation will provide improved substance identification and recognition.

Objective

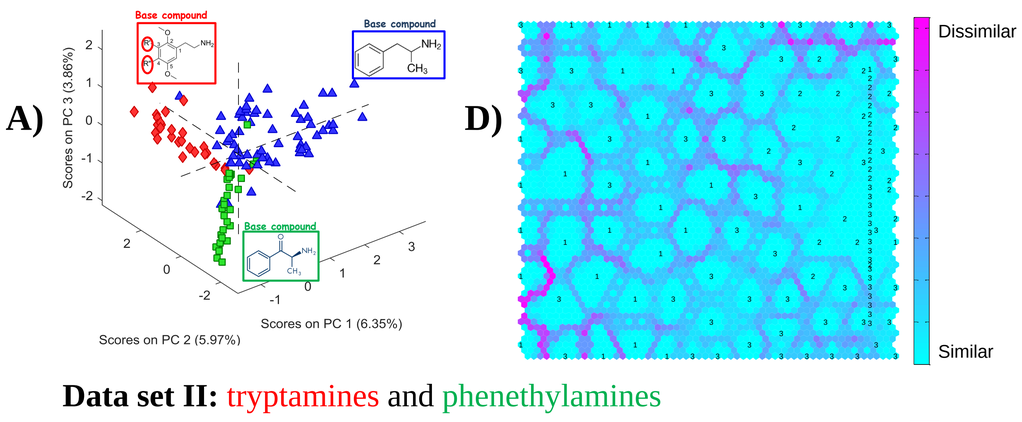
The objective of this study was to classify street drugs using NMR spectroscopy and chemometric analysis with uncertainty estimation. The method proposed is aimed at classifying suspected, unknown drug substances and could be used as a complementary method to classical forensic inspection.

Methodology

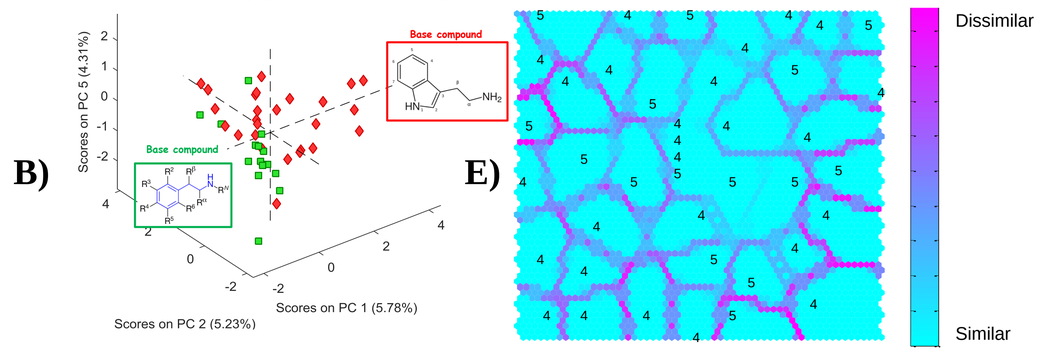
A set of 217 ^{1}H NMR spectra was provided by the German Federal Criminal Police Office which were broadly grouped into seven different chemical classes: 2C-x phenethylamines (n = 28), cathinones (n = 31), amphetamines (n = 63), tryptamines (n = 26), phenethylamines (n = 15), piperazines (n = 14), and methylenedioxy-phenethylamines (n = 20). The phenethylamines class includes substituted phenethylamine compounds that do not fall into one of the other related sub-classes. The raw spectra were binned into uniform width chemical shift windows with either 500 or 2000 bins per spectrum. Exploratory analysis and classification was done using PCA, SOM, and PLS-DA models.

Results and Discussion

Data set I: 2C-x phenethylamines, cathinones, and amphetamines



Data set II: tryptamines and phenethylamines



Data set III: piperazines and methylenedioxy-phenethylamines

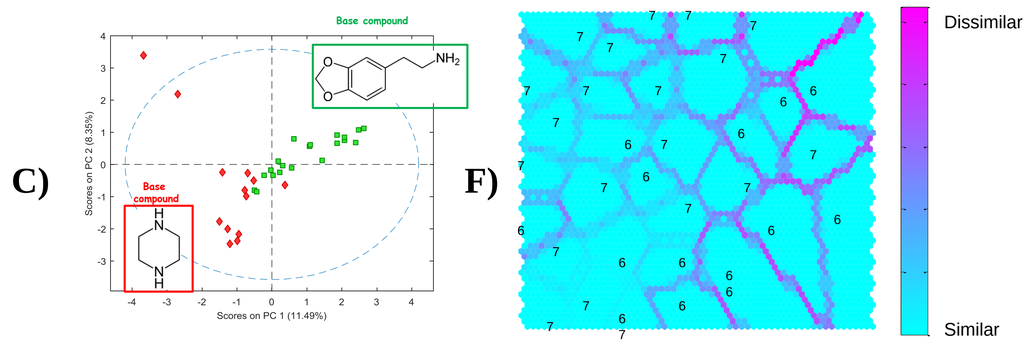


Fig. 1 Results of the exploratory analysis for the spectra 2000 bins per spectrum. A), B), and C) PCA models; D), E), and F) SOM models.

Table 1: Classification of parameters obtained for PLS-DA models built from **Data set I** for the spectra with 2000 bins per spectrum (blue column) and 500 bins per spectrum (gray column).

	CALIBRATION						VALIDATION					
CLASS	Class1		Class2		Class3		Class1		Class2		Class3	
N[A]	18	13	18	21	49	51	10	15	13	10	14	12
N[B]	0	0	0	0	0	1	1	1	3	3	2	0
ME (%)[C]	0	0	0	0	0	2	10	6.7	23	30	14	0
TP	1	1	1	1	1	0.98	0.9	1	0.69	0.7	0.93	1
FP	0	0	0	0	0	0	0.1	0	0	0	0.04	0
TN	1	1	1	1	1	1	1	1	1	1	1	1
FN	0	0	0	0	0	0.03	0	0.05	0.16	0.11	0	0
SENS	1	1	1	1	1	0.98	0.9	0.93	0.69	0.7	0.93	1
SPEC	1	1	1	1	1	1	1	1	1	1	1	1
R	0.97	0.91	0.98	0.94	0.97	0.93	0.86	0.88	0.86	0.86	0.83	0.82
RMSE	0.10	0.14	0.08	0.14	0.11	0.18	0.23	0.24	0.28	0.24	0.28	0.27

[a]N: number of samples in each class; [b]N: number of misclassified samples classes. [c]ME (%): misclassification error; TP: true positive; FP: false positive; TN: true negative; FN: false negative; Sens: sensitivity; Spec: specificity. R: Pearson's correlation coefficient for calibration and validation. RMSE: root mean square error for calibration and validation. **Class1:** 2C-x phenethylamines. **Class2:** cathinones. **Class3:** amphetamines.

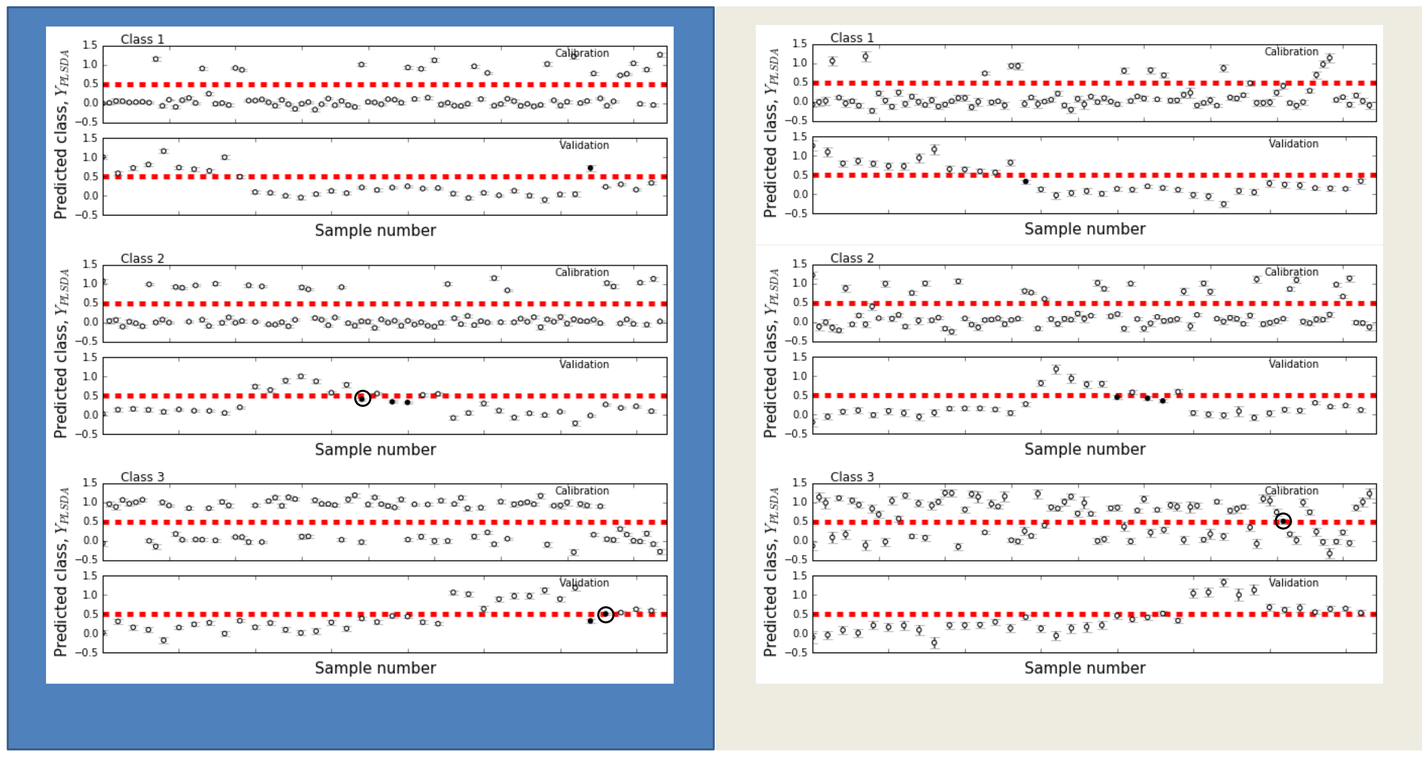


Fig. 2 Predicted classes for the calibration and validation sets of the PLS-DA models from Table 1, showing the members of each of the three classes (above dashed line) with prediction intervals for all samples analyzed by the residual bootstrap method. Some samples, highlighted by black circles, have prediction intervals that include the class boundary and so cannot be confidently classified.

Conclusions

Exploratory analysis was used to identify different chemical classes. PCA can identify general structural classes of molecules while SOM can discriminate among the derivatives of a general class of molecules. Estimating the uncertainty by PLS-DA model provides a conclusion that is more reliable and complete for a classification in forensic analysis.

The NMR spectra were provided by the German Federal Criminal Police Office (BKA) under the supervision of Torsten Schoenberger.