

On-Scene Cocaine Detection in Seized Samples

by Handheld Near Infrared Spectrometry and Machine Learning Algorithms

Ruben Kranenburg

Introduction



Ruben Kranenburg
external PhD student

Laboratory coordinator of the Amsterdam Police illicit-drugs laboratory



Arian van Asten
full professor

Professor of Forensic Analytical Chemistry and On-Scene Chemical Analysis

POLITIE

UNIVERSITY OF AMSTERDAM







Outline

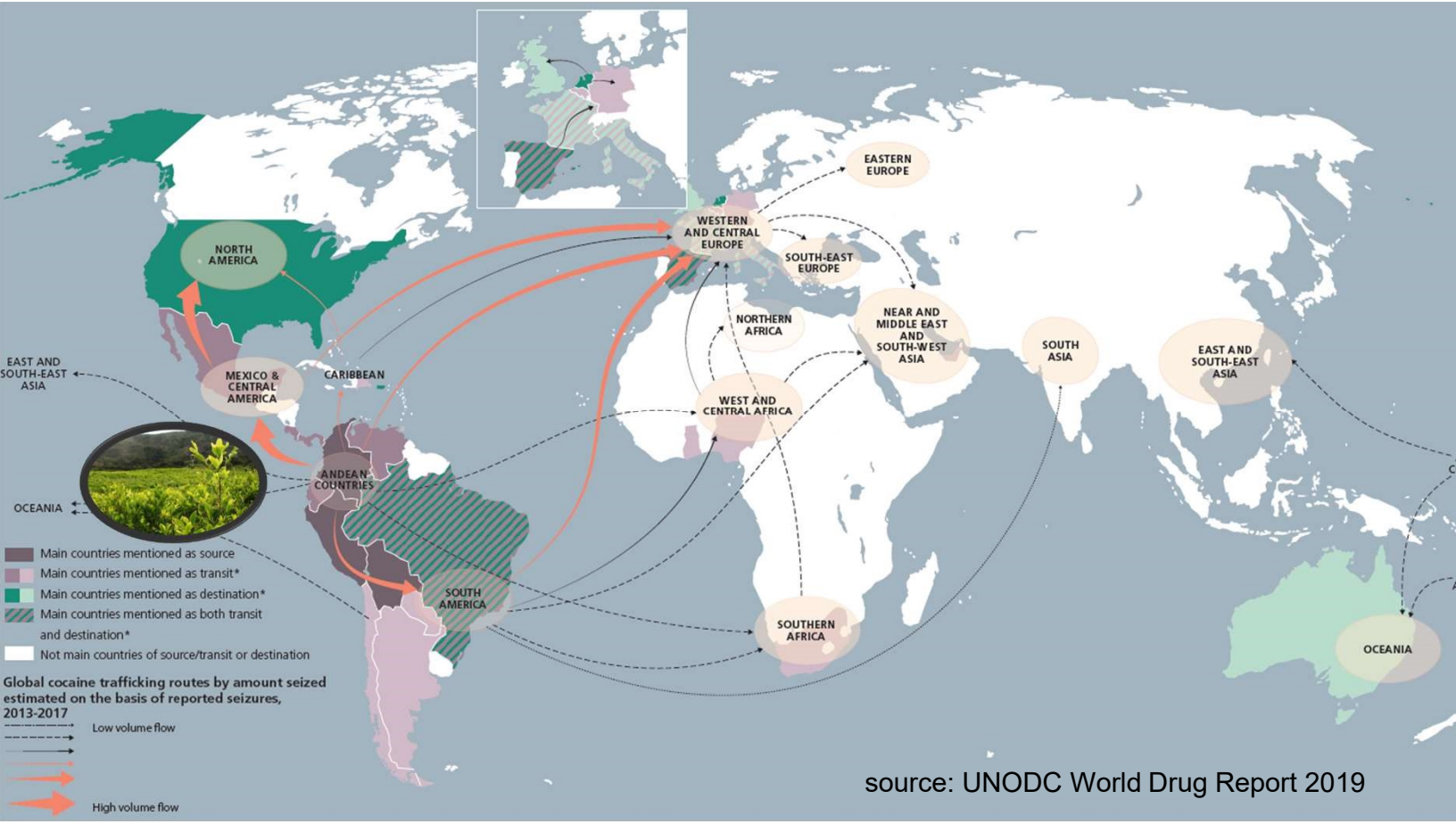
- The cocaine market
- Forensic laboratory testing
- NIR spectroscopy
  1. Spectral data
  2. Model development
  3. Performance
- Applicability
- Conclusions & outlook



Outline

- **The cocaine market**
- Forensic laboratory testing
- NIR spectroscopy
  1. Spectral data
  2. Model development
  3. Performance
- Applicability
- Conclusions & outlook





source: UNODC World Drug Report 2019



Global number of cocaine users
2017

18 million



COCAINE

Global cultivation
2017

Change from previous year
+15%

245,000 ha

343,137 x

Global production
2017

Change from previous year
+25%

pure cocaine

1,976 tons
at 100% purity



sources: UNODC World Drug Report 2019
EU European Drug Report 2018

Outline

- The cocaine market
- **Forensic laboratory testing**
- NIR spectroscopy
  1. Spectral data
  2. Model development
  3. Performance
- Applicability
- Conclusions & outlook



On-scene



Laboratory







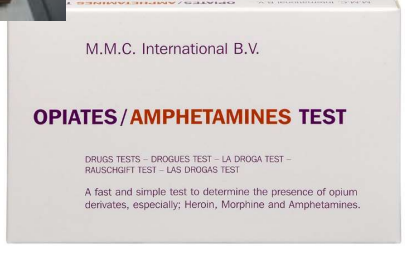
- transport
- manipulation
- paperwork
- time

1920 - 2015







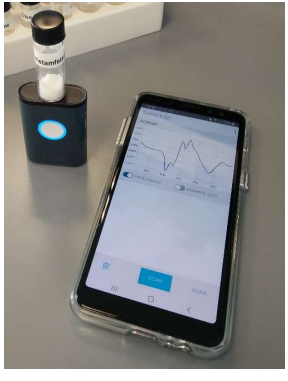


> 2015





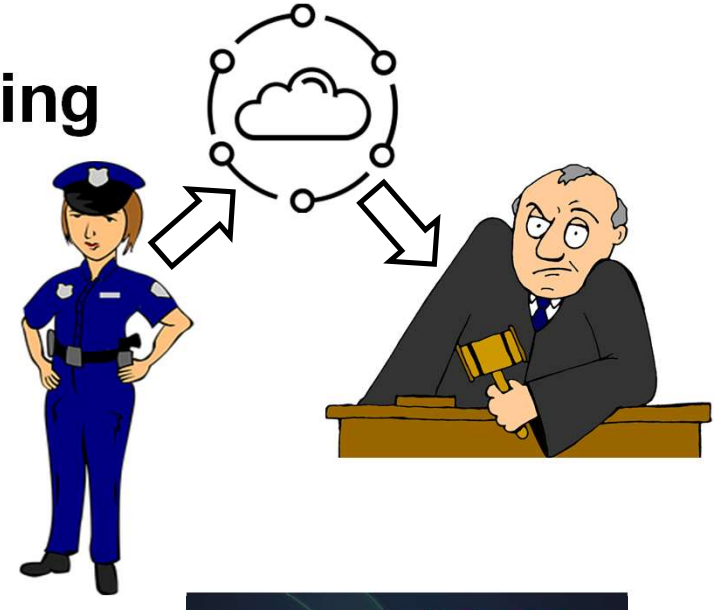


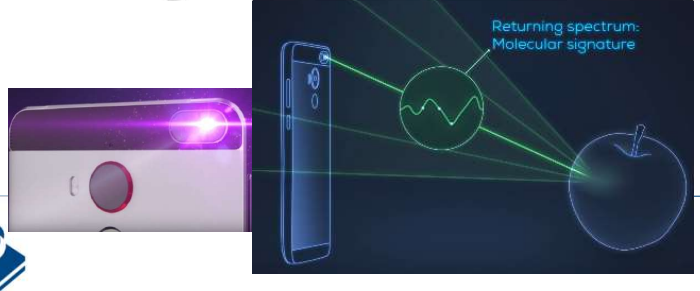


Forensic Laboratory Testing

Questions

- Applicability and evidential value
  - Indication only?
  - Court evidence?
- Selectivity, limit of detection
- Total process efficiency
  - Specialized CSI / border security teams?
  - Every police officer's smartphone?





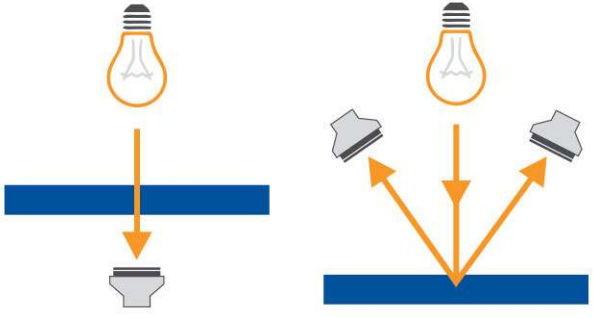
Outline

- The cocaine market
- Forensic laboratory testing
- **NIR spectroscopy**
  1. Spectral data
  2. Model development
  3. Performance
- Applicability
- Conclusions & outlook



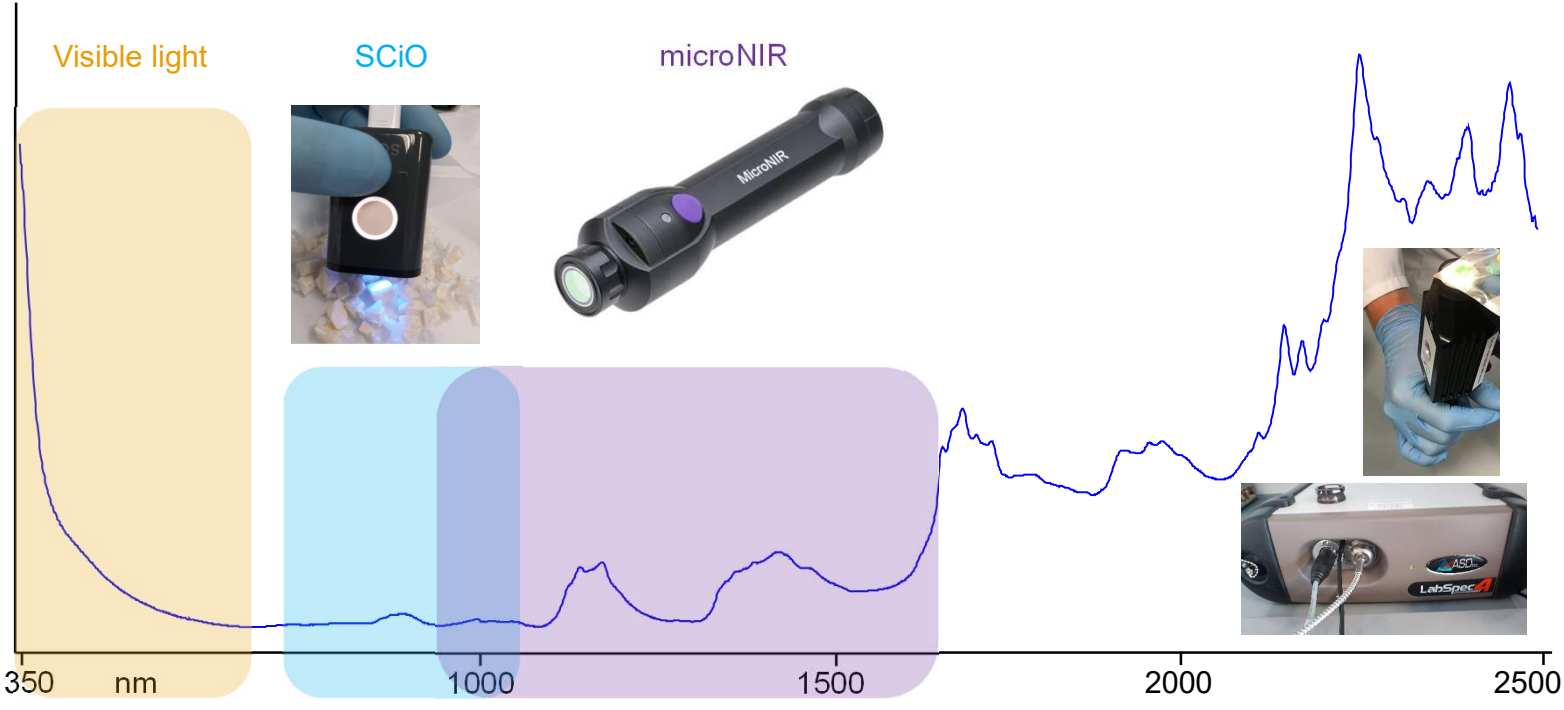
Near Infrared Spectroscopy

- ~700 – 2500 nm
- diffuse reflection analysis



absorbance reflection

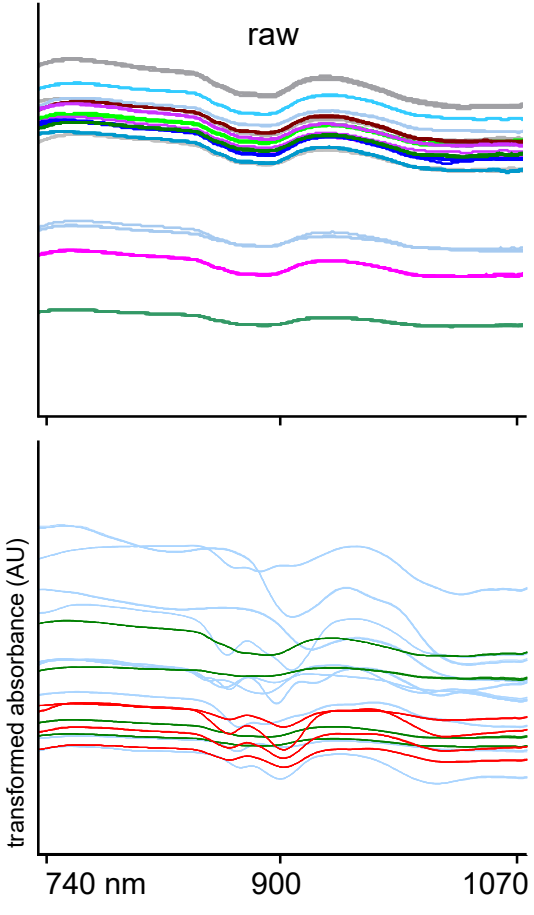
source: FOSS analytics

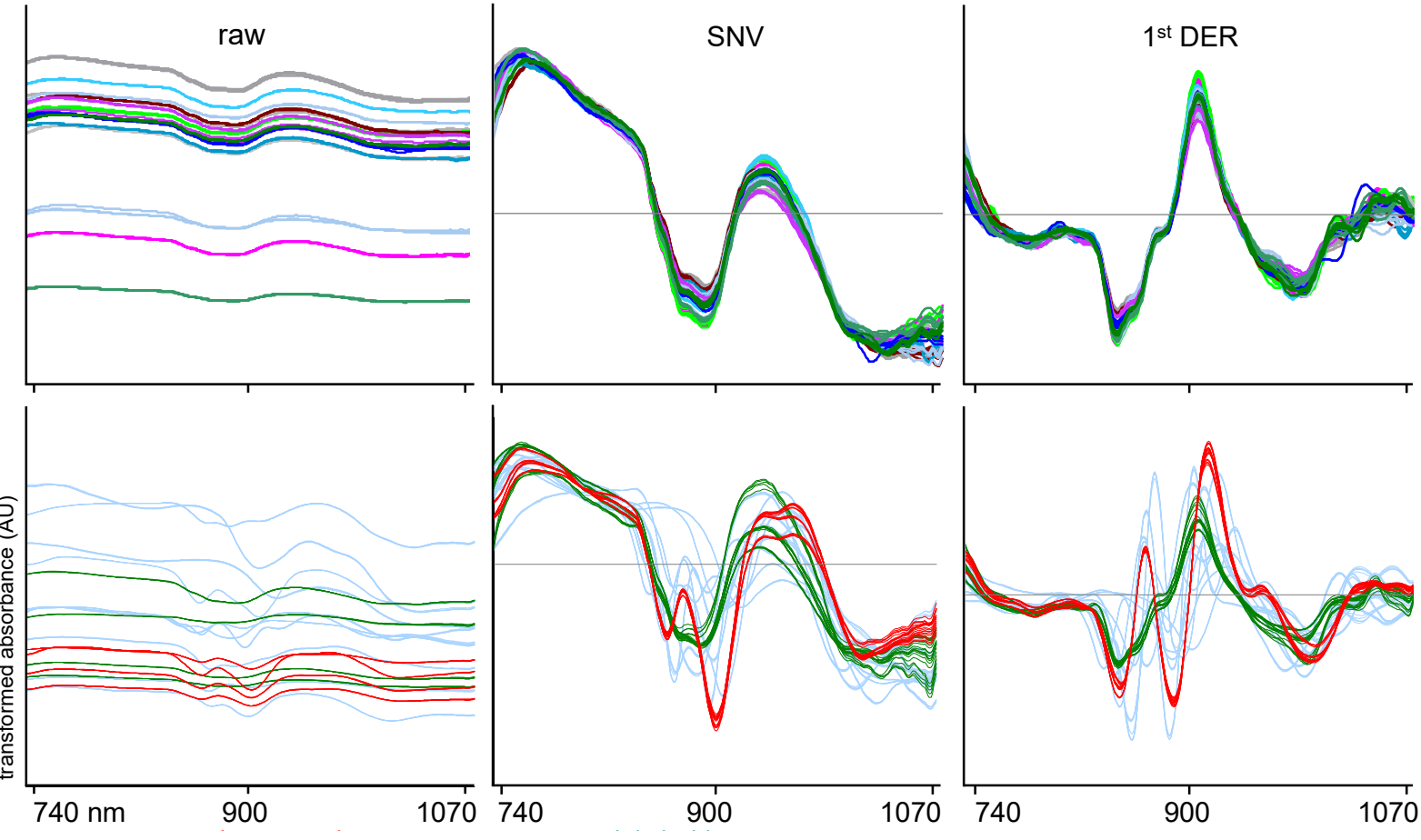


	SCiO	microNIR	ASD LabSpec4
	740 – 1070 nm	950 – 1650 nm	350 – 2500 nm
Cocaine HCl			

Near Infrared Spectroscopy

- ~700 – 2500 nm
- diffuse reflection analysis
  - Scattering effects
  - Signal intensity differences
- spectral preprocessing




raw
SNV
1st DER
transformed absorbance (AU)
740 nm
900
1070
740
900
1070
740
900
1070

Model design

goal	***to do***
robustness	multiple scanners
	harmonized method
representative	actual case samples
rapid	automated processing

Model design

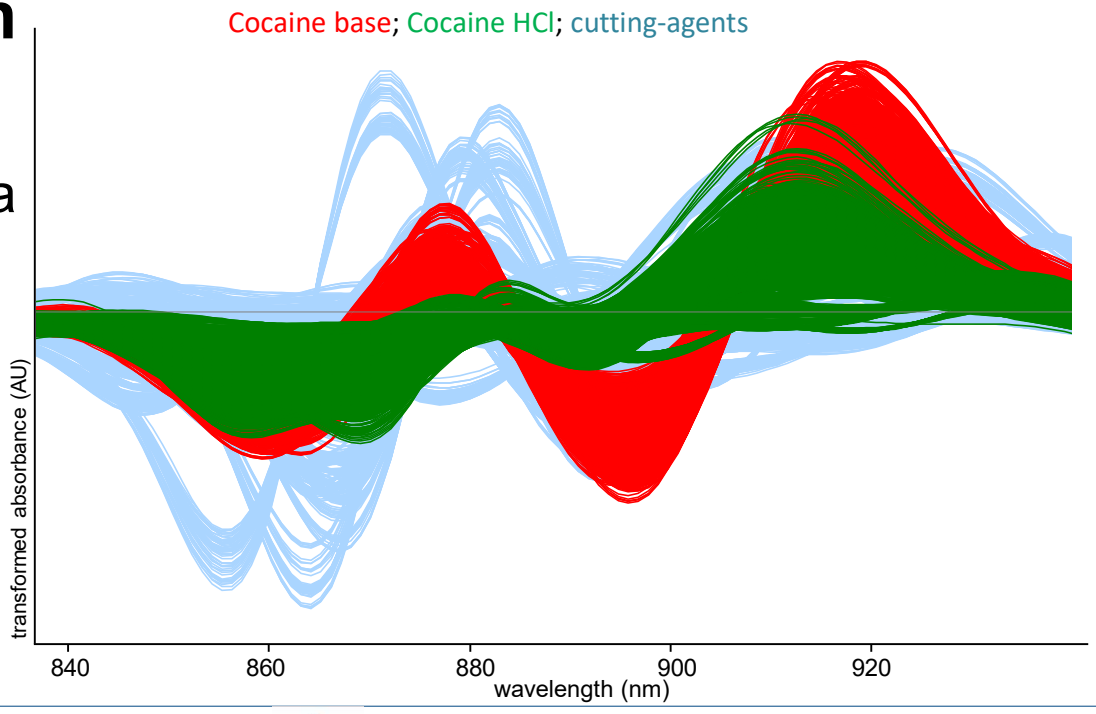
- 15 instruments
- > 10 operators
- 520 samples
- 10059 scans
- glass vials

set 1 'case'		set 2 'negatives'		set 3 '2018 standards'	
samples		**samples**		**samples**	
Cocaine base (n)	38	negatives (n)	39	standards (n)	305
Cocaine HCl (n)	93			negatives (n)	45
mean (%)	*66.1*	*all unique*		*max (%)*	*64.5*
median (%)	*69.2*	- *cutting agents*		*min (%)*	*4.5*
max (%)	*99.4*	- *other drugs*			
min (%)	*17.3*	- *common powers*			
data		**data**		**data**	
SCiO scans	6136	SCiO scans	2420	SCiO scans	1503
Unique SCiO's	13	Unique SCiO's	13	Unique SCiO's	6
replicates	5	replicates	5	replicates	5



Model design

- Different spectra cocaine base cocaine HCl
- Differences for mixtures

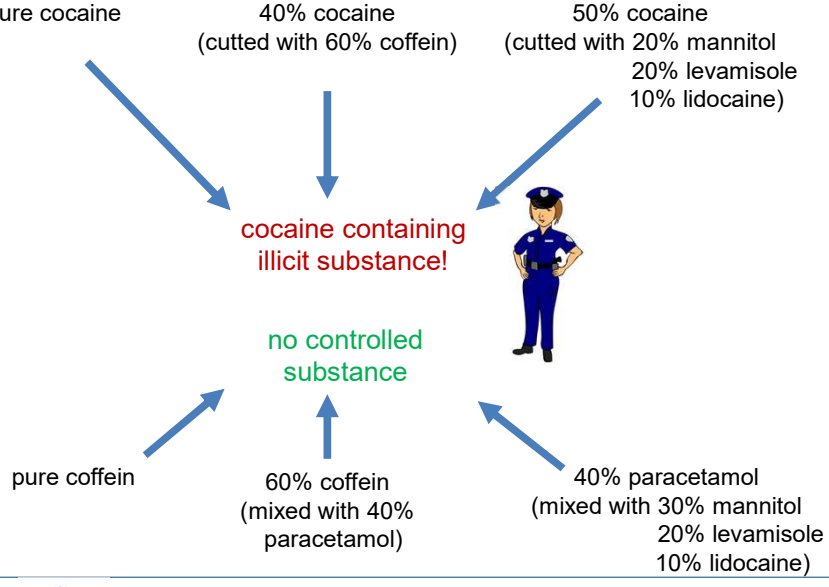


Model design



- Chemometric models
  - PCA
  - SIMCA – PLS
  - PLS-DA

  Too many false positives
- Multi stage models
  - k-Nearest Neighbors
  - Machine learning



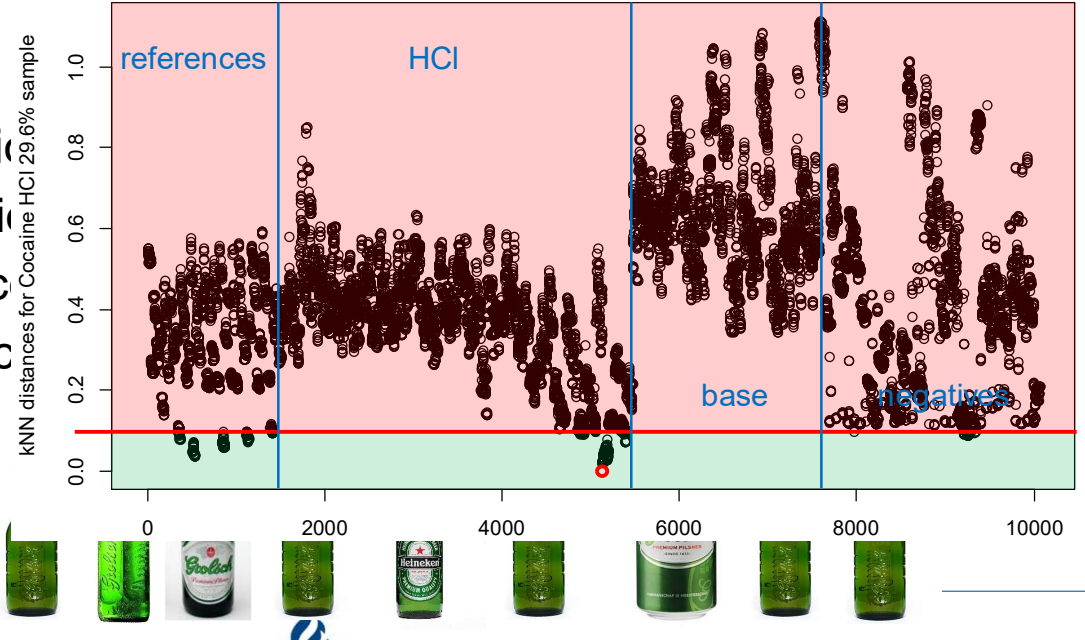
K-Nearest Neighbors

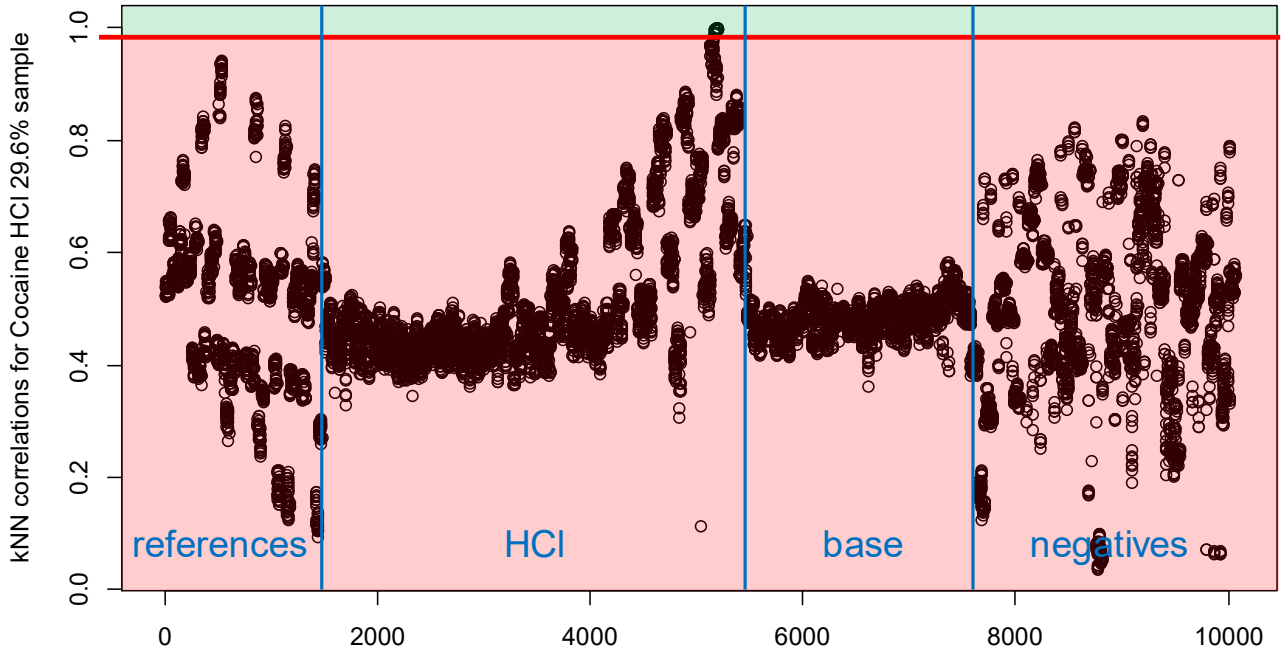
Two types:

- Pearson Correlati
  (focus on similariti
- Euclidean Distanc
  (focus on differenc





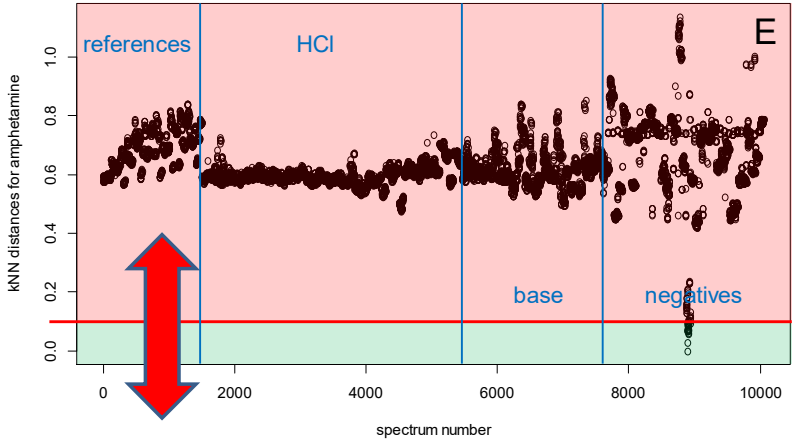




kNN correlations

K-Nearest Neighbors

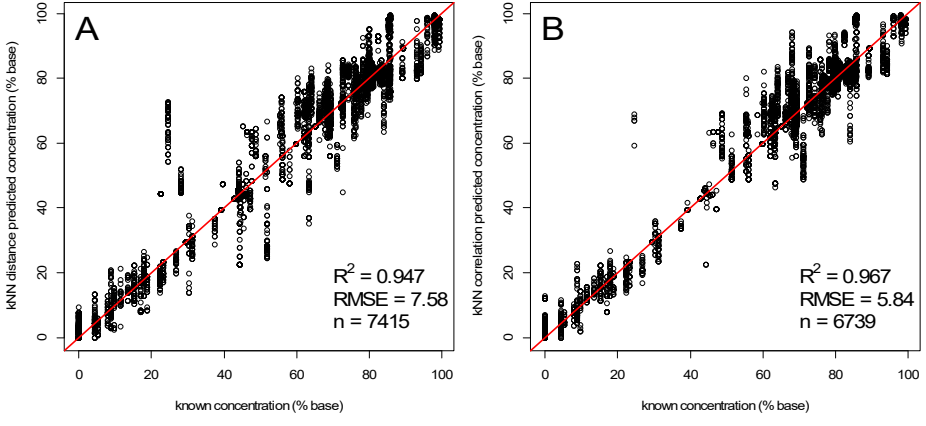
- Optimization:
  - Thresholds for ID
  - Nr. of ‘best’ neighbors
- Our model:
  - Correlations: Top 9
  - Distances: Top 9
- Result: Top 18 (or less…)
- Voting on ID of neighbors



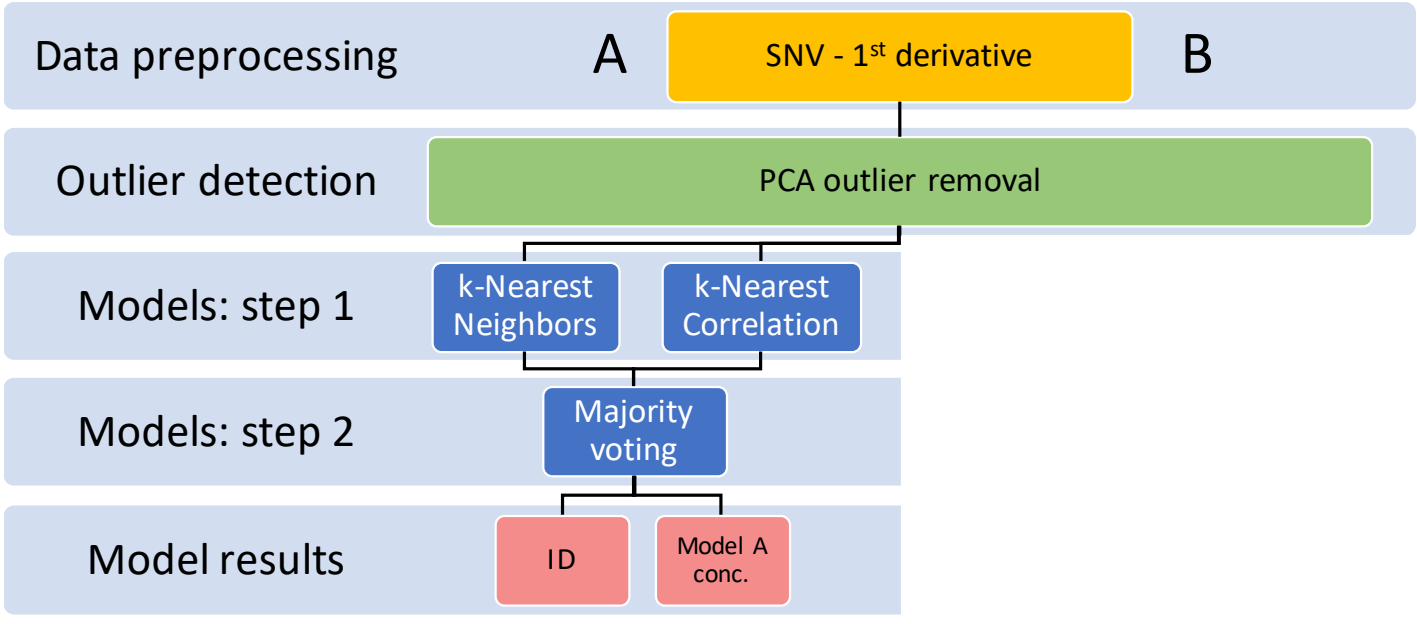


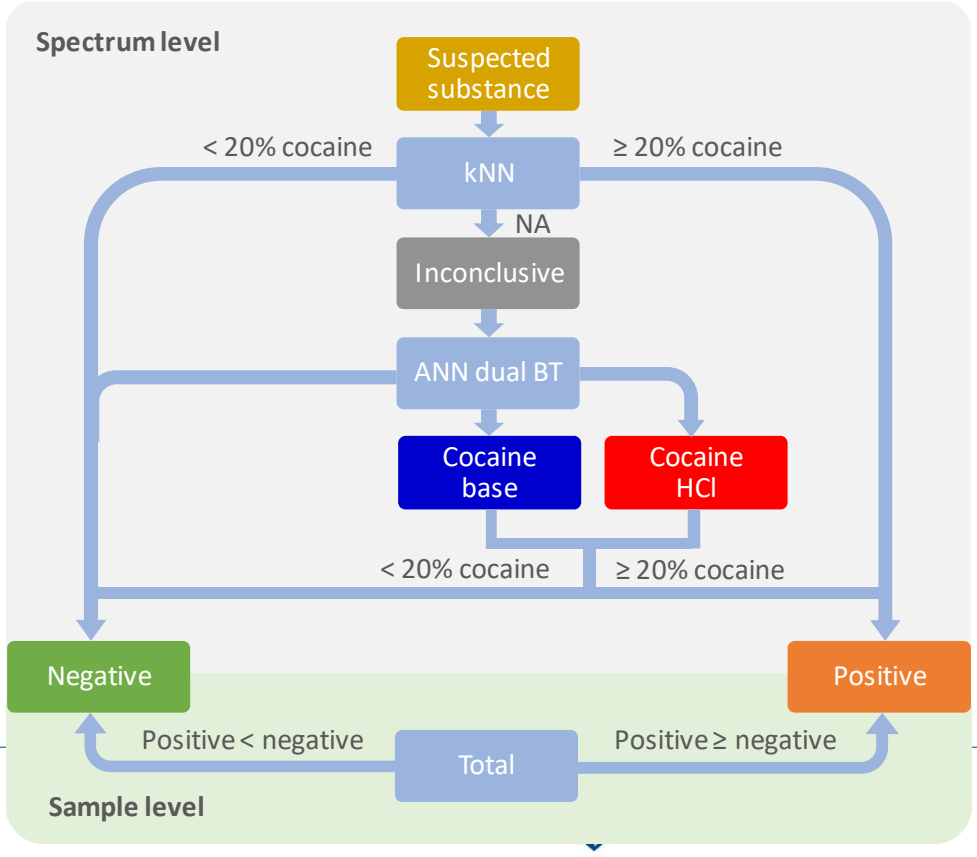
K-Nearest Neighbors performance

- 99% correct classification; acurate concentration prediction
- However, **no result** for new substances *or* unique mixtures (31% of dataset, mostly 'negatives')



Additional step needed…


Data preprocessing
A
SNV - 1st derivative
B
Outlier detection
PCA outlier removal
Models: step 1
k-Nearest Neighbors
k-Nearest Correlation
Models: step 2
Majority voting
Model results
ID
Model A conc.


Spectrum level
Suspected substance
< 20% cocaine
kNN
≥ 20% cocaine
NA
Inconclusive
ANN dual BT
Cocaine base
Cocaine HCl
< 20% cocaine
≥ 20% cocaine
Negative
Positive
Positive < negative
Total
Positive ≥ negative
Sample level

Outline

- The cocaine market
- Forensic laboratory testing
- NIR spectroscopy
  1. Spectral data
  2. Model development
  3. **Performance**
- Applicability
- Conclusions & outlook



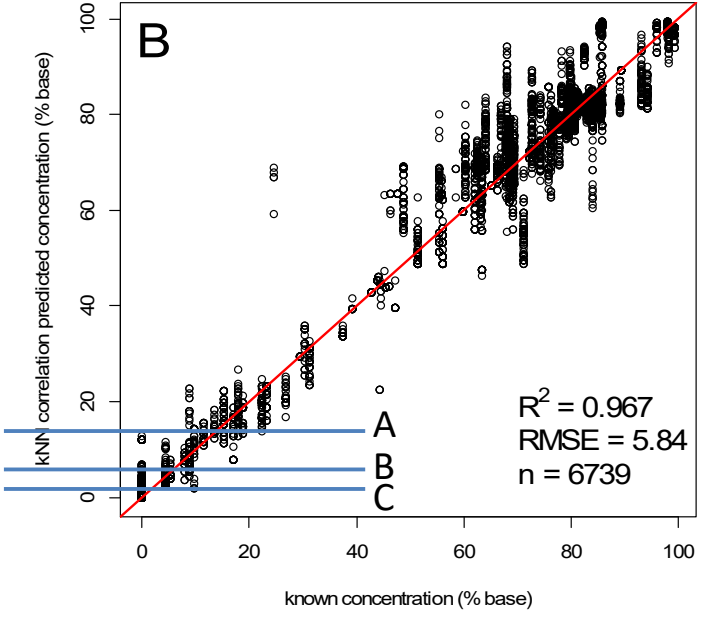
Model performance

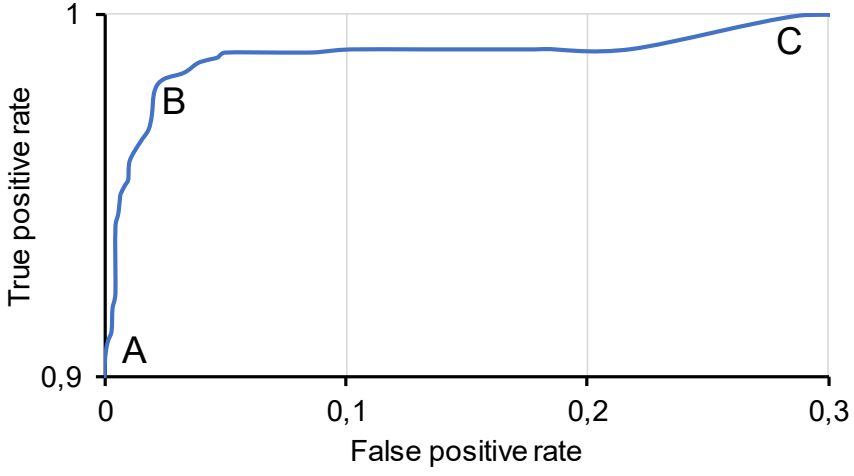
First step: what criteria do we need?

- Zero false positive? (festivals..)
- Zero false negative? (bulk goods testing…)
- Some optimum?

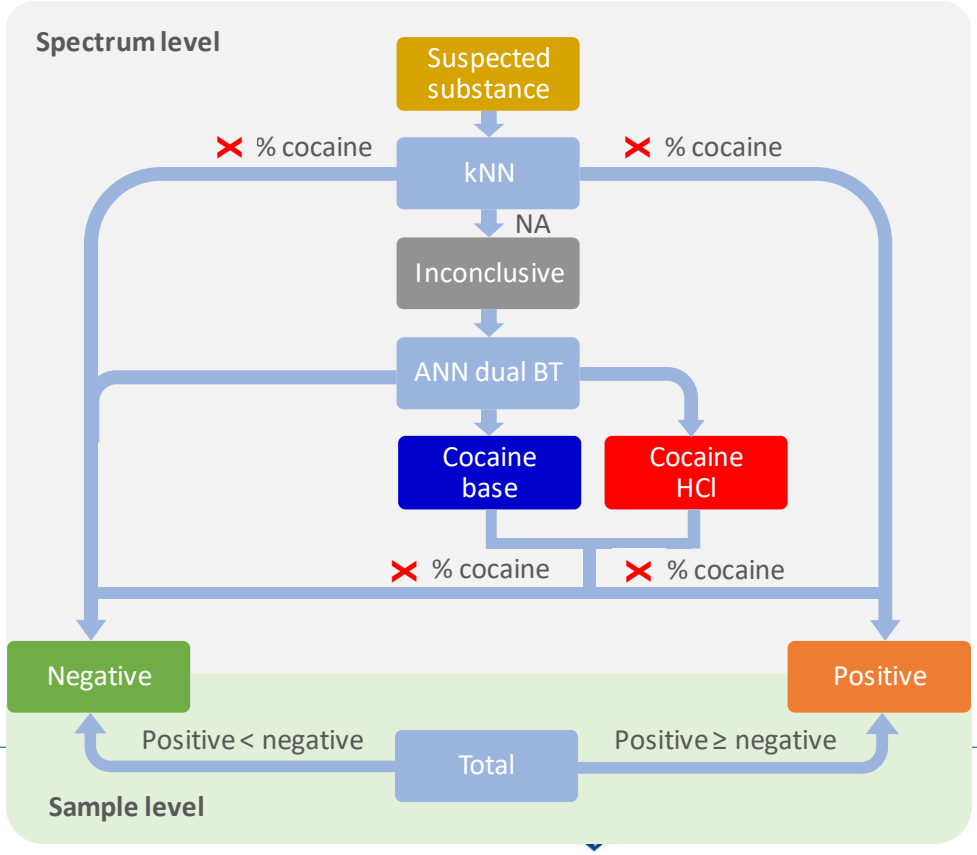








A: 0% FP 10% FN (@ 36 wt% cocaine)
B: ~1% FP ~1% FN (@ 19 wt% cocaine)
C: 30% FP 0% FN (@ 0 wt% cocaine)



✕ = adjustable threshold

Model performance

- Model fully cross-validated
- Leave ‘all from same sample’ out
- Threshold 19 wt%

k-Nearest Neighbors (sub-model A) performance						
pre-processing	True Positives	False Negatives	True Negatives	False Positives	n pos	n neg
SNV - 1st D	98.9%	1.1%	99.3%	0.7%	6184	1011
ANN - dual tree bagging (sub-model B) performance						
pre-processing	True Positives	False Negatives	True Negatives	False Positives	n pos	n neg
SNV - 1st D	96.1%	3.9%	97.3%	2.7%	6821	2638
ANN - dual tree bagging (sub-model B) performance on sub-model A inconclusive spectra						
pre-processing	True Positives	False Negatives	True Negatives	False Positives	n pos	n neg
SNV - 1st D	81.0%	19.0%	97.5%	2.5%	637	1627

note:
637 positives
1627 negatives
Inconclusive! (31%)

Outline

- The cocaine market
- Forensic laboratory testing
- NIR spectroscopy
  1. Spectral data
  2. Model development
  3. Performance
- **Applicability**
- Conclusions & outlook



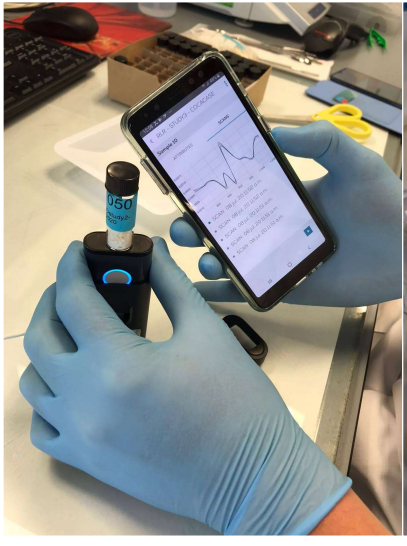
External case samples

- 54 case samples, 625 scans
- 91% TP; 96% TN

False positives:
amphetamine:caffein
diltiazem

False negatives:
< 17 wt% cocaine samples

Inconclusives (kNN part):
new substances or compositions





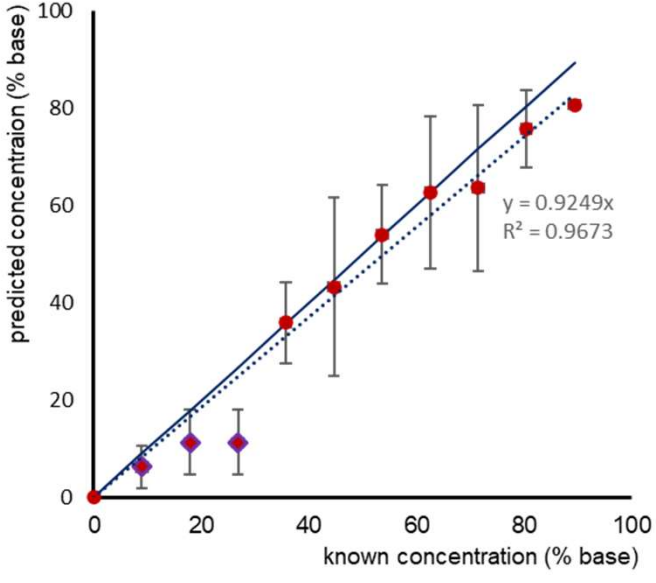
External case samples

Manually cutted cocaine samples

- 10 – 100 wt%
- cutted with 3 common substances

Sub 30 wt% samples:

- false negative
- kNN model inconclusive -> BT-ANN



Outline

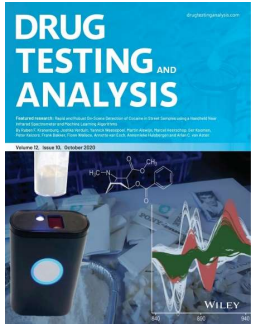
- The cocaine market
- Forensic laboratory testing
- NIR spectroscopy
  1. Spectral data
  2. Model development
  3. Performance
- Applicability
- **Conclusions & outlook**



Conclusions

- Cocaine yields limited spectral signal in 740 – 1070 nm NIR range
- Spectral preprocessing and multi-stage models exploit these features
- Cocaine detection with short wavelength NIR is feasible

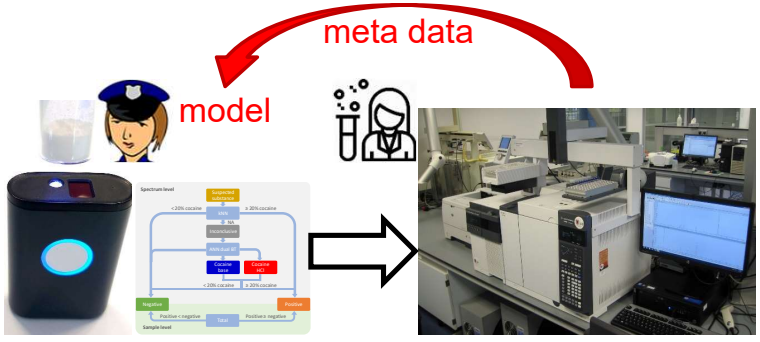
  97% TP and 98% TN at 19 wt% threshold
- Limitations arise at low concentrations & unknown spectra
- Optimization by adjustable threshold is possible



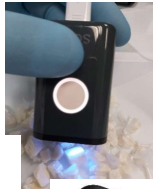
Kranenburg *et al.* **2020**
https://doi.org/10.1002/dta.2895

Outlook

- Machine learning models can adapt to actual market developments



- Limitations:
  - Only demonstrated in glass bottles (NIR transparant)
  - Packaging may influence spectrum



- Other NIR wavelength range



- Data safety & integrity

Acknowledgements

Dutch governmental laboratory partners

Joshka Verduin

Renee de Ridder

Yannick Weesepoel

Martin Alewijn

Arian van Asten

Kevin van Bree

Peter Keizers

Frank Bakker

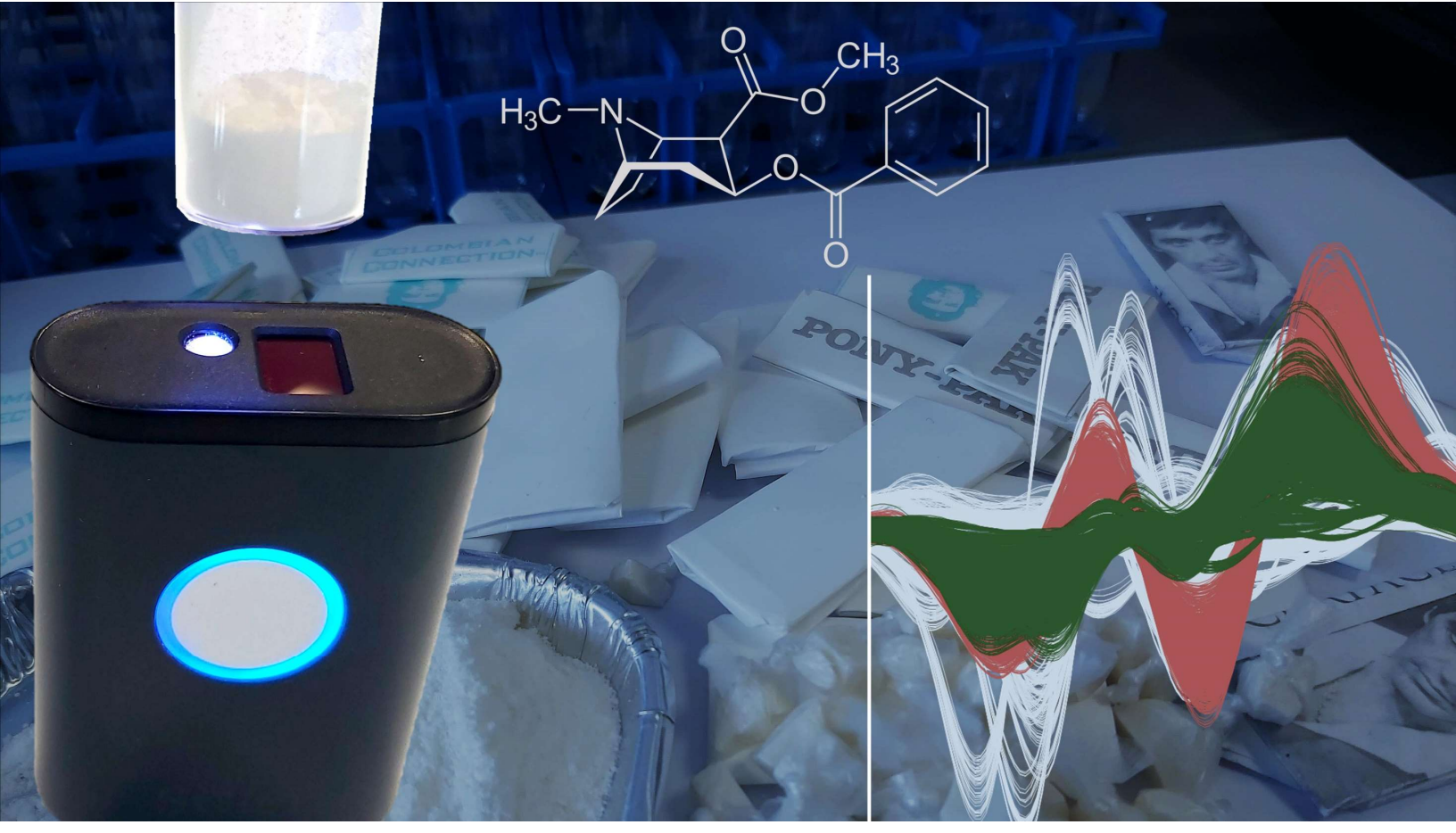
Marcel Heerschop

Ger Koomen

Fionn Wallace

Annette van Esch

Annemieke Hulsbergen


O
CH3
H3C—N
O
O
O
COLOMBIAN CONNECTION
POWY-