European Network on New Sensing Technologies for Air Pollution Control and Environmental Sustainability - *EuNetAir*

COST Action TD1105

INTERNATIONAL WG1-WG4 MEETING on

New Sensing Technologies and Methods for Air-Pollution Monitoring

European Environment Agency - EEA

Copenhagen, Denmark, 3 - 4 October 2013

Action Start date: 01/07/2012 - Action End date: 30/06/2016 - Year 2: 2013-2014 (*Ongoing Action*)

Multivariate modelling of spectroscopic data for tracking sources of particulate air-borne matter

Agricultural University of Iceland

Arngrímur Thorlacius

MC Member

Iceland

COST is supported by the EU Framework Programme

European Science Foundation

ESF provides the COST Office through a European Commission contract

Introduction

- On an island in "the middle" of the Atlantic ocean, with only about three inhabitants per square kilometer, one might expect to be reasonably free from air pollution.

- We do, however, in Iceland have severe problems with air-borne particulate matter in our two largest urban areas i.e. in Reykjavík and Akureyri, where the EU limit values for PM10 are now exceeded on several occasions each winter.

- This lecture describes efforts to develop methodology to trace the origin of this pollution involving a multivariate mathematical model

Methodology in short

- Sources of the pollution are simulated by collecting dust from suspected pure sources (asphalt, soot, soil, volcanic ash etc.) onto filters, in the same or very similar manner as with the collection of ambient dust samples.

- All samples, real and synthesized, undergo a multitude of quantitative measurements yielding concentration values or spectral points. These measurement results are then used in a multivariate model to reveal the amount of each source in the real dust samples.

- Solving such a problem is referred to as ***Source Appointment*** or ***Source Apportionment***

Suspected sources

- Soil (two samples)
- Asphalt (two samples)
- Soot (from a diesel engine)
- Brakes (from a brake repair shop)
- Salt (de-icing salt)
- Volcanic ash (two samples, Eyjafjallajökull and Grímsvötn)

Sample collection

Partisol Plus 2025 air sampler

The arrangement for collecting diesel exhaust

The model – schematic view

VARIABLES	Na	Fe	....	2007nm	2439nm
SOURCES					
Ash EJ	a1	b1	....	y1	z1
As GV	a2	b2	....	y2	z2
Brakes	a3	b3	....	y3	z3
Soot	a4	b4	....	y4	z4
Soil A	a5	b5	....	y5	z5
Soil B	a6	b6	....	y6	z6
Asphalt 1	a7	b7	....	y7	z7
Asphalt 2	a8	b8	....	y8	z8
Salt	a9	b9	....	y9	z9

MODEL

VARIABLES	Na	Fe	....	2007nm	2439nm
SAMPLES					
Sample 1	as 1	bs 1	....	ys 1	zs 1
Sample 2	as 2	bs 2	....	ys 2	zs 2
:	:	:	....	:	:
:	:	:	....	:	:
:	:	:	....	:	:
:	:	:	....	:	:
:	:	:	....	:	:
:	:	:	....	:	:
:	:	:	....	:	:
:	:	:	....	:	:
Sample n	as n9	bsn9	....	ys9	zs9

The model – methodology

- The algorithm for the multivariate calculations is called Partial Least Squares Regression (PLS). It utilizes principal components calculations. It was developed within the field of chemometrics but is now becoming popular in other fields, mainly due to its flexibility

- All calculations were performed with the Sirius chemometric software package (Pattern Recognition Systems AS)

Experimental

- Dust samples collected on glass fibrer filters (Pall) with Partisol Plus 2025 Sequential air sampler (Thermo Scientific) equipped for PM10
- Reflectance scans (NIR-Vis) taken with 2 nm resolution from 400nm to 2500nm (NIRSystems 6500)
- Filters digested in teflon containers (Questron high pressure bombs) under pressure through microwave heating
- Digests analysed by Inductively Coupled Plasma Optical Emission Spectrometry (Jobin Yvon Ultima 2) determining: Ca, Mg, K, Na, S, Fe, Mn, Cu, Zn, Al, Si, Ti, Ba, Sr, Cr, Co, Zr, W

Reflectance measurements

- As with traditional IR, NIR spectra are often dominated by the organic constituents of a sample, but with less resolution. All organic molecules will contribute. Adding the visible region we get a measure of colour differences, which add to the information content

- Here we only want to use the spectral data to trace a compound spectrum of contributions from individual source materials. In other words, we don´t have to know what chemical compounds are contributing to the spectra nor their concentrations. Only that spectral contributions will add up linearly to give a final NIR spectrum.

Final data set

Several variables were discarded previous to or during multivariate modelling for various reasons. The following were retained:

Elements: Na, Fe, Mn, Zn, Si, Ti, Ba, Sr, Co, Zr, W

NIR/Vis* (nm) : 407, 679, 855, 983, 1223, 1335, 1415, 1879, 2007, 2439

* Reflectance values ($R = \log[1/ I_R]$) recorded at the stated wavelength

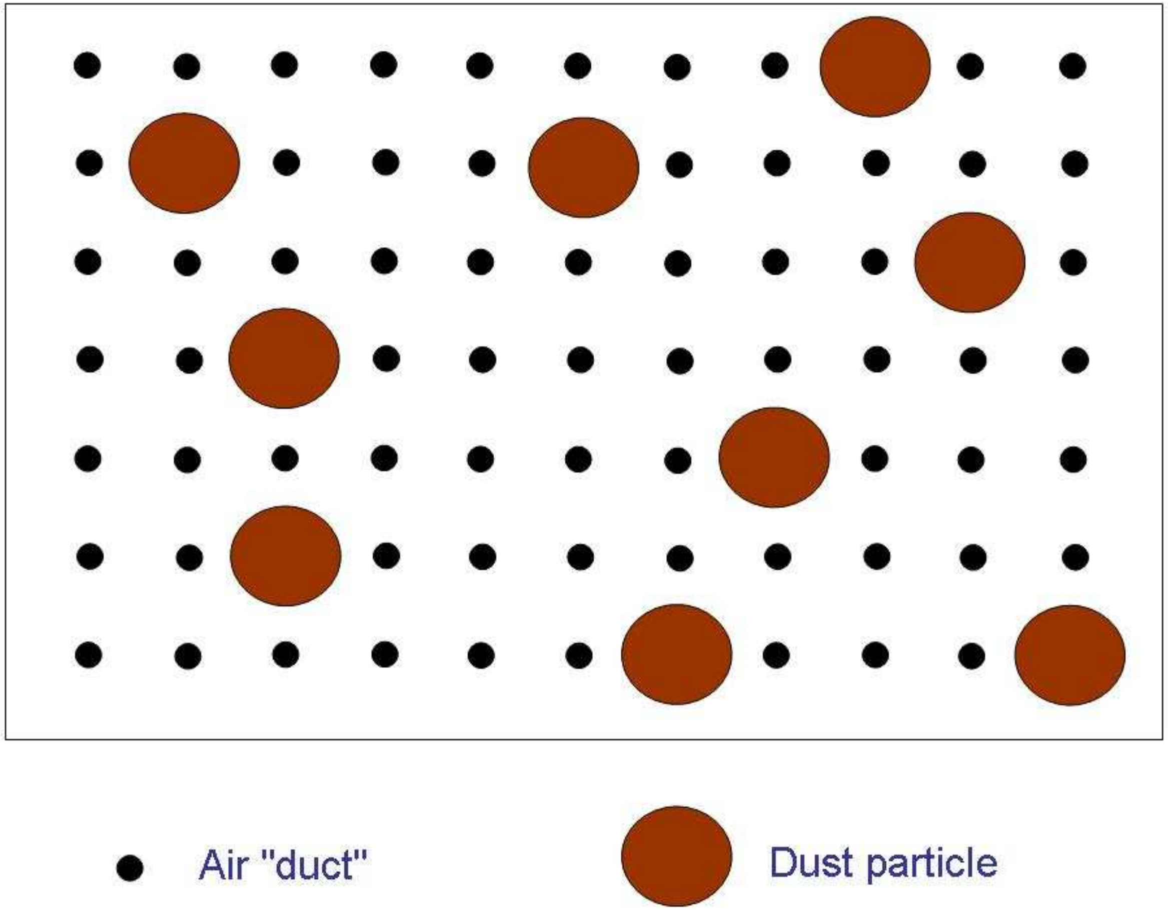
Analytical problems

- The glass fiber filters contain a significant not constant amount of several analytes. Problem can be only partially solved by washing the filters in acid prior to sample collection

- The filters exhibit a significant reflectance over the entire wavelength range. This background can not be easily subtracted as it varies with the amount of dust collected (dust particles cover a portion of the filter surface and thus shield the background reflectance. This an be handled with an iterative procedure during modeling

Reflectance from filter surface varies with the amount of collected particles

A collected particle will cover a part of the filter surface and thus reduce the background reflectance. This reduction will at first be linear with number (or mass) of collected particles and then the effect will level off with increased overlap



A simple shielding model

The following relationship can be used to mimic the reduced background reflectance caused by particles from one source, S_i :

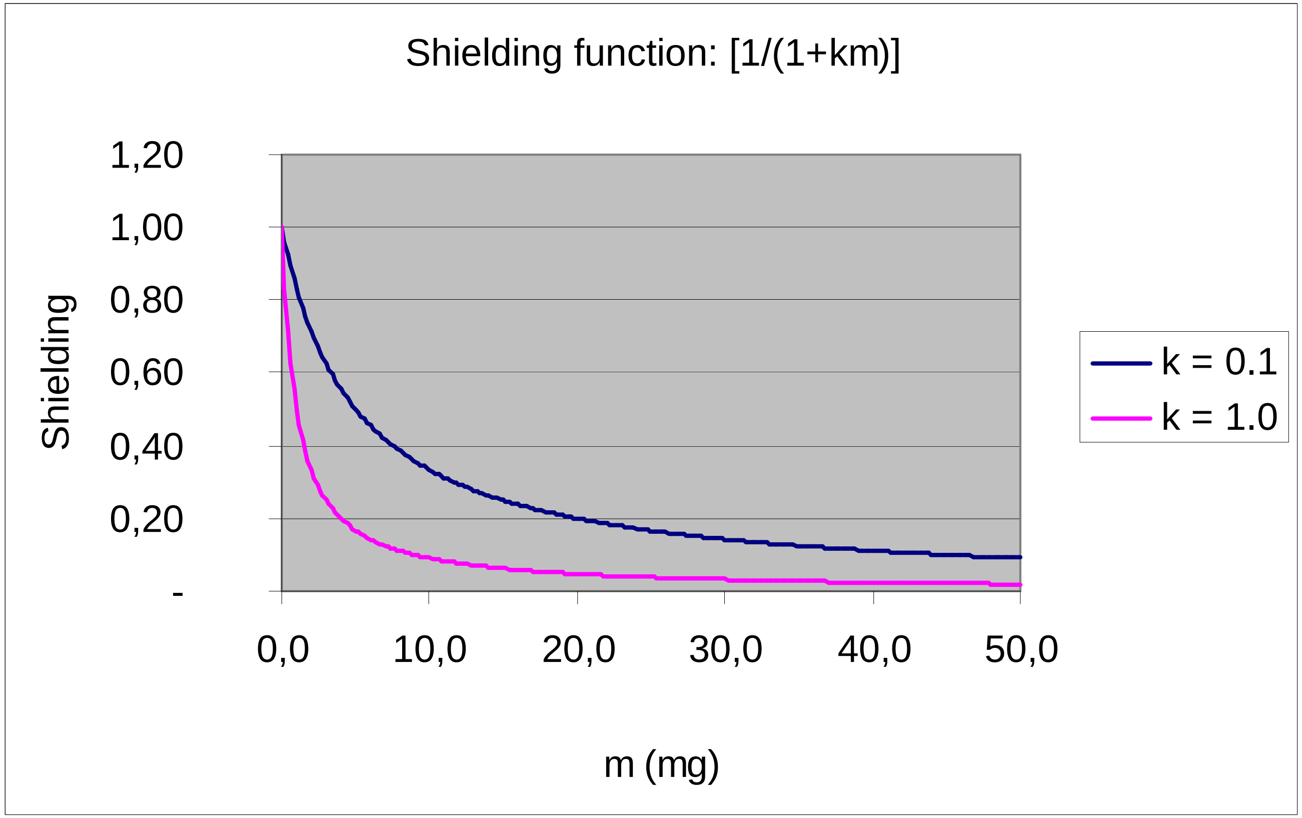
$R_{background} = R_{blank} \cdot [1/(1 + k_i \cdot m_i)]$

k_i : a shielding constant for S_i

m_i : the collected mass of S_i

(Note that the relation within the brackets returns a value ≤ 1)

The overall shielding effect is obtained by summing over all sources



If a source material is effective in shielding the background reflectance (large k-value), relatively less mass is needed to produce an effect.

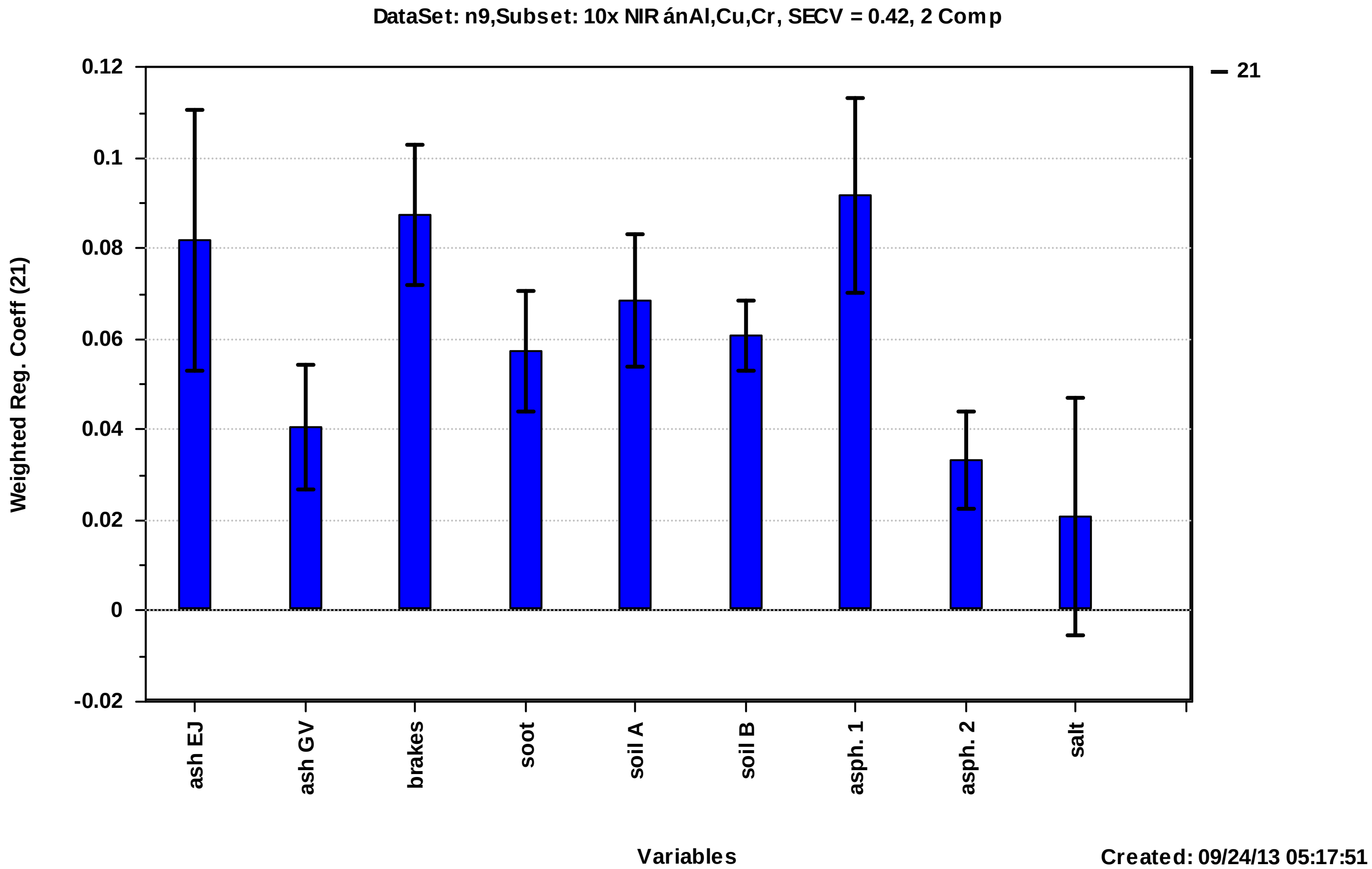
Iterative procedure

1. An initial PLS-model, based on element concentrations only, is calculated
2. Source composition from this first model is used to calculate shielding for the background correction of the reflectance spectra
3. Thus corrected spectra enter an extended PLS model
4. PLS results are used to get more correct reflectance spectra
5. Points 2 to 4 are repeated until convergence is obtained (seen by no further change in source content of unknown samples

Novelty of this approach

- To the best of my knowledge, the quantitative use of such reflectance spectra for this type of work has not been reported elsewhere
- The preparation of filter samples from pure suspected dust sources is also peculiar to the present approach
- So is the combination of the above within the frame of a PLS-model

Model results



- Regression coefficients yield relative contributions from individual sources

Normalized results

Sample no.	48h avg µg/m³	Ash %	Brakes %	Exhaust %	Soil %	Asphalt %	Salt %
21	67	22	17	11	24	22	4
15	43	23	16	9	23	23	5
31	42	18	15	23	20	20	3
24	40	14	13	41	16	14	2
16	40	17	13	33	18	16	2
AVG		**19**	**15**	**24**	**20**	**19**	**3**
MAX	67	23	17	41	24	23	5

- Results for the five "largest" samples (greatest collected mass)

PMF – an alternative approach

- Positive Matrix Factorization (PMF) is another principal component based technique that is widely used for source appointment.
- PMF has been adopted by the US EPA as one of a few methods recommended for source appointment
- PMF requires greater computing power than the present technique (not of great importance)
- PMF has its weakness in resolving similar sources

Future developments

- More trace analytes (elements) will be added to the data set in order to improve predictability.
- I´m working on how to make dust from rubber from tires
- It would be interesting to test the present methodology on other data sets (in other cities).
- Comparison with the PMF method would be very interesting.
- To include information from the synthesized source samples in the PMF may simplify the interpretation needed to complete PMF predictions and it may help in handling similar source materials which is reportedly a problem for the technique.

CONCLUSIONS

- PLS modelling of spectroscopic data, seems to be a worthwhile exercise to establish the origin and content of the somewhat complex mixtures of dust that are diminishing the live quality of urban Icelanders.

- Adding non-destructive reflectance data to the conventional wet-chemical results improves the quality of the data set.

- The use of synthetic pure source samples may be of use in other source appointment methods

Acknowledgments

The Icelandic Road Authority is thanked for financial support for this project

Thanks to COST Action TD1105 for funding the author's Short Term Scientific Misson visiting the chemometrics group of the University in Bergen Norway.

Thanks to Þorsteinn Jóhannsson and the Icelandic Environmental Agency for allowing the use of the dust sampler for this project and to Páll Höskuldsson at Efla Engineering for the planning and implementation of the dust sampling program.

Special thanks to Olav Kvalheim and Egil Nodland at the University of Bergen for valuable guidance with the modeling work.

THANK YOU !!