

## **Příloha 1/B**

### **Vyhodnocení experimentálních dat – Ostravsko**

## 1 Úvod

Cílem této zprávy je vyhodnocení dat, která byla pořízena v roce 2003 (jaro, podzim) v oblasti města Ostravy. Jednalo se jedno stanoviště (Ostrava-Přívoz), na kterém byly provedeny dvě odběrové kampaně. Jednalo se o 24 hodinové odběry v průběhu jednoho měsíce. Datové soubory, získané analýzou odebraných vzorků, obsahovaly údaje o koncentraci polyaromatických uhlovodíků (PAU) a některých prvků. Dále byly provedeny podpisy 6 významných energetických zdrojů z hlediska obsahů PAU a některých prvků.

Pro matematicko-statistické zpracování těchto souborů byly využity multivariační statistické metody a program US EPA CMB 8.2.

## 2 Teoretická část

### 2.1 Multivariační průzkumová analýza

Prvotní informaci o struktuře dat může poskytnout **metoda hlavních komponent (PCA)**.

#### **Metoda hlavních komponent**

Metoda hlavních komponent neboli PCA je metodou vícerozměrné popisné statistiky. Matematickým způsobem snižuje počet proměnných charakteristik tím, že vytváří teoretické ("umělé") proměnné neboli parametry, které v sobě však zahrnují informaci o všech původních proměnných. Tyto nové proměnné, které vytvářejí nový souřadnicový prostor, nazýváme hlavními komponentami. Umístění bodů v tomto novém prostoru nazýváme *score*, zastoupení původních proměnných v hlavních komponentách nazýváme zátěžemi nebo komponentními váhami (*loadings*). Podaří-li se touto metodou vyjádřit původní parametry souboru pomocí menšího počtu přímo neměřitelných teoretických proměnných, které je možno racionálně interpretovat, získáme úspornější a efektivnější popis souboru studovaných objektů i struktury závislostí mezi jejich parametry. Zároveň si můžeme utvořit představu o druhu a intenzitě působení procesů, které podmiňují pozorovanou variabilitu datových souborů. PCA analýza může tedy sloužit k několika účelům:

- a) k identifikaci teoretických proměnných,
- b) k transformaci dat do systému ortogonálně proměnných,
- c) ke klasifikaci objektů nebo proměnných
- d) k redukci rozměrů zkoumaného problému, tj. redukci rozměru prostoru nahrazením většího počtu původních proměnných čili parametrů novým souborem o menším počtu teoretických proměnných.

Při výpočtu komponentní analýzy jsou důležité výše uvedené komponentní zátěže (*loadings*). Jsou to hodnoty, které se získávají průmětem vektorů z prostoru na ortogonální osy hlavních komponent, čili projekční koeficienty jednotlivých vektorů. Bývají také někdy označovány jako korelační koeficienty původních charakteristik s jednotlivými komponentami. Při interpretaci je nutné vzít v úvahu hlavní komponenty s největšími vlastními čísly a zjišťovat, které charakteristiky s jednotlivými komponentami korelují. Vysoké kladné a vysoké

záporné váhy pro prvky téže komponenty jednoduše odrážejí antagonistické chování těchto prvků. Tímto způsobem si můžeme učinit představu o tom, které konkrétní vlivy nebo procesy jsou příčinou pozorované variability charakteristik a závislosti mezi jejich hodnotami.

## 2.2 Sdružovací (klastrová) analýza

### **Základní principy sdružovací analýzy**

Výchozím bodem většiny sdružovacích metod je stanovení míry podobnosti objektů. Většinou požadujeme, aby byla nulou vyjádřena maximální rozdílnost a jedničkou totožnost. Z praktických důvodů však často pracujeme s mírami nepodobnosti neboli se vzdálenostmi, které nabývají nulových hodnot v případě totožnosti a jejichž hodnoty rostou se zvětšující se rozdílností objektů. Hodnota "vzdálenosti" představuje spíše míru nepodobnosti a její doplněk je výrazem podobnosti. Nejznámější je euklidovská vzdálenost mezi dvěma body (jedincem A a B), určeným ortogonálními vektory:

$$d_{AB} = \sqrt{\sum_{i=1}^n (X_{iA} - X_{iB})^2}, \text{ kde}$$

$X_{iA}$  = hodnota znaku  $i$  příslušná objektu A

$X_{iB}$  = hodnota znaku  $i$  příslušná objektu B

Společnou nevýhodou všech vzdáleností je závislost na jednotkách měření (vzdálenost nejvíce ovlivní charakteristiky s největšími hodnotami směrodatných odchylek). Nepříznivě se projeví i vliv korelovatelnosti charakteristik.

Všechny techniky sdružovací analýzy vycházejí z matice pozorování, která není předem rozdělena ani podle jedinců, ani podle proměnných. Na základě matice pozorování příslušných jedinců se formuluje postup jak sdružit jedince (způsob Q) nebo proměnné (způsob R) do navzájem oddělených skupin na podkladě některého z koeficientů podobnosti.

## 2.3. Lineární diskriminační analýza

Metody se používá pro řešení otázek třídění jedinců do určitých, předem definovaných skupin. Lze to doložit na příkladu dvou nebo více souborů, které jsou charakterizovány systémem znaků (proměnných, parametrů). O daném jedinci máme rozhodnout, do kterého ze souborů patří. Diskriminační funkce vymezuje kritéria pro oddělení jednotlivých souborů a tím i zařazení jedince.

Diskriminační analýza se liší od sdružovacích metod tím, že určuje příslušnost jedinců do předem vymezených skupin, zatímco sdružovací analýzy hledají skupiny přítomných jedinců a vytvářejí jejich systém, aniž by zpočátku byl určen charakter nebo počet skupin.

Diskriminační analýzou lze řešit v podstatě dvě skupiny klasifikačních problémů:

1. Otázky samostatnosti skupin jedinců vydělených na základě souboru znaků (proměnných). Tento problém lze stručně formulovat takto: Máme-li dva nebo více souborů představovaných jednoznačně rozříděnými jedinci, vyjádříme pomocí souboru znaků (proměnných) „vzdálenost“ mezi těmito soubory a tak získáme představu míry nezávislosti mezi jednotlivými soubory navzájem. Je-li vzdálenost mezi určitou dvojicí souborů menší než zvolené kritérium významnosti rozdílu mezi skutečně samostatnými skupinami, jsme oprávněni revidovat správnost původní zvolené klasifikace a hledat příčiny a chyby.
2. Otázky zařazení sporných jedinců do předem definovaných skupin na základě typických souborů. Každý z problematických souborů je zařazen do určité skupiny podle afinity svých znaků k té či oné skupině. Lze určit pravděpodobnost příslušnosti jedince k určeným skupinám.

Významnou součástí diskriminační analýzy je výpočet diskriminační funkce a testování „vzdáleností“ mezi soubory.

## 2.4 Receptorové modelování

Receptorové modelování používá chemické a fyzikální vlastnosti plynů a aerosolových částic, které jsou změřeny pro emisní zdroje a receptory (v českém prostředí používáme slovo imisní charakteristiky místa), pro identifikaci a kvantifikaci příspěvku jednotlivých zdrojů na jednotlivých imisních (odběrových) místech. Z matematického hlediska zpracování emisních a imisních dat zahrnuje:

- a) chemickou hmotnostní bilanci
- b) faktorovou analýzu datové matice
- c) vícenásobnou lineární regresi

Základním požadavkem pro receptorové modelování je:

- a) stabilita emisních zdrojů v monitorovaném období
- b) nalezení všech významných emisních zdrojů, které přispívají k charakterizaci imisních míst

V práci byl použit programový soubor CMB 8.2, který byl poskytnut pracovníky US EPA v North Carolina.

### 3 Experimentální část

Odběrové vzorkovací plány, analýza odebraných vzorků a vytvoření datových matic pro oblast města Ostravy jsou předmětem dílčí zprávy za rok 2004. Pro vyhodnocení experimentálních dat byl použit programový soubor XLSTAT (Addinsoft, 2005, [www.xlstat.com](http://www.xlstat.com)).

### 4 Vyhodnocení experimentálních dat a diskuse

Byly zpracovány 4 datové soubory imisních dat:

1. PAU, jaro, 16. dubna až 17. května 2003
2. PAU, podzim, 20. října až 22. listopadu 2003
3. kovy, jaro, 16. dubna až 17. května 2003
4. kovy, podzim, 20. října až 22. listopadu 2003

Dále byly zpracovány dva soubory emisních dat, pro PAU a kovy z těchto významných energetických zdrojů:

Název zdroje	Označení
Spalovna průmyslových odpadů, 17.8.04	SPOVO
Energetika Vítkovice, 19.8.04	ENVI
Elektrárna Třebovice, 22.8.04	ETRE
Třinecké železářny, 25.8.04	TRIZ
OKD, koksovna Svoboda, 26.8.04	OKDk
Elektrárna Dětmorovice, 28.8.04	DETM

V souborech je použito pro označování jednotlivých PAU těchto zkratek:

Phe	phenanthren	CAS: 000085-01-800
Ant	anthracen	CAS: 000120-12-700
Flt	fluoranthen	CAS: 000206-44-000
Pyr	pyren	CAS: 000129-00-000
B[a]A	benzo[a]anthracen	CAS: 000056-55-300
Chr	chrysen	CAS: 000218-01-900
B[e]P	benzo[e]pyren	CAS: 000192-97-200
B[b]F	benzo[b]fluoranthen	CAS: 000205-99-200
B[k]F	benzo[k]fluoranthen	CAS: 000207-08-900
B[a]P	benzo[a]pyren	CAS: 000050-32-800
DB[ah]A	dibenzo[a,h]anthracen	CAS: 000053-70-300
B[ghi]P	benzo[g,h,i]perylene	CAS: 000191-24-200
I[cd]P	indeno[1,2,3-cd]pyren	CAS: 000193-39-500

## 4.1 Průzkumová analýza dat

Soubor PAU – jaro

Soubor obsahoval 32 datových vektorů. Každý z vektorů je charakterizován 13 koncentracemi PAU (ng/m<sup>3</sup>).  $\chi^2$  test potvrdil vysoký stupeň korelace mezi jednotlivými proměnnými ( $\chi^2 = 1080$ ,  $\chi^2(\text{krit}) = 99$ ). Základní statistické charakteristiky souboru jsou uvedeny v Tabulce I.

**Tabulka I Základní statistické charakteristiky souboru PAU – jaro**

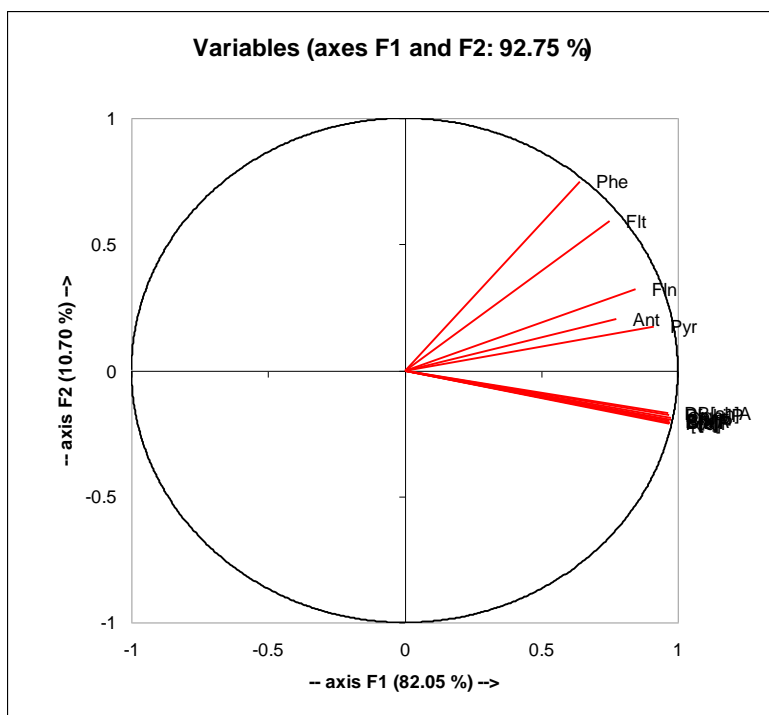
PAU	průměr	SD	RSD
Fln	64.4	31.9	49.5%
Phe	103.0	36.2	35.1%
Ant	2.9	1.6	54.0%
Flt	28.5	11.5	40.4%
Pyr	11.0	4.5	40.4%
B[a]A	2.9	2.5	85.8%
Chr	3.7	2.7	72.8%
B[b]F	3.6	2.7	74.8%
B[k]F	1.5	1.1	73.6%
B[a]P	2.6	2.0	74.9%
DB[ah]A	0.3	0.3	86.3%
B[ghi]P	2.6	1.8	67.7%
I[cd]P	2.7	2.0	73.2%

SD - výběrová směrodatná odchylka, RSD – rel. SD

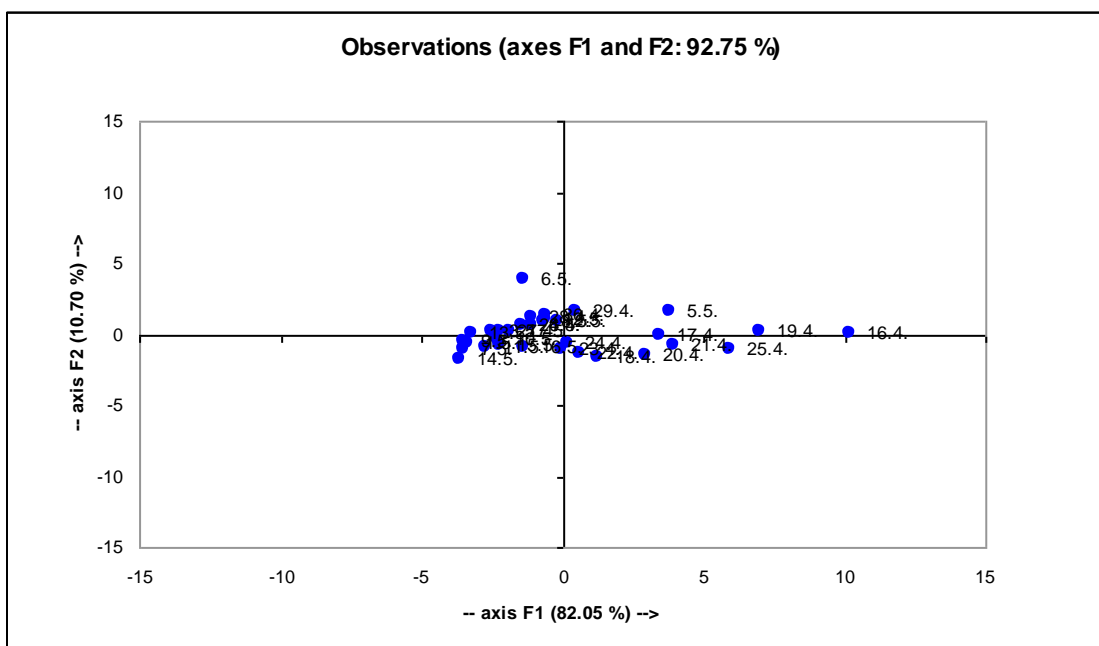
Vysoké hodnoty RSD svědčí o tom, že data nejsou homogenní, tj. nepocházejí z jednoho zdroje.

Na obrázku 1 je graf korelační kružnice, znázorňující korelaci původních proměnných (PAH) s prvními dvěma hlavními komponentami. První dvě hlavní komponenty nesou 92,8% veškeré informace o tomto souboru, jsou tedy postačující pro dokonalý popis původních proměnných. Všechny proměnné vykazují vysokou vzájemnou korelaci (na korelační kružnici jsou umístěny blízko sebe), i když tvoří dvě skupiny proměnných. V jedné jsou Fln, Phe, Ant, Flt a Pyr (lehčí PAU), ve druhé ostatní, těžší polyaromatické uhlovodíky.

Na obrázku 2 je graf skóre, který ukazuje rozdělení všech „vzorků“ (vektory pozorování) v prostoru prvních dvou hlavních komponent. Z obrázku je patrné, že vzorky netvoří pouze jeden shluk, pocházejí tedy z více emisních zdrojů.



**Obr. 1**  
Korelační  
kružnice pro  
soubor PAU -  
jaro



**Obr. 2** Graf skóre pro soubor PAU – jaro

### Soubor PAU – podzim

Soubor obsahoval 34 datových vektorů. Každý z vektorů je charakterizován 13 koncentracemi PAU (ng/m<sup>3</sup>).  $\chi^2$  test potvrdil vysoký stupeň korelace mezi jednotlivými proměnnými ( $\chi^2 = 1544$ ,  $\chi^2(\text{krit}) = 100$ ). Základní statistické charakteristiky souboru jsou uvedeny v Tabulce II.

**Tabulka II Základní statistické charakteristiky souboru PAU – podzim**

PAU	průměr	SD	RSD
Fln	66.1	46.4	70.1%
Phe	118.0	66.7	56.5%
Ant	16.4	15.1	92.0%
Flt	48.9	36.4	74.5%
Pyr	31.5	23.1	73.4%
B[a]A	10.0	8.9	88.5%
Chr	8.9	7.5	83.7%
B[b]F	7.1	6.2	88.3%
B[k]F	3.6	3.1	87.1%
B[a]P	6.5	5.9	90.2%
DB[ah]A	0.6	0.5	95.1%
B[ghi]P	4.5	3.8	86.1%
I[cd]P	6.0	5.3	87.9%

SD - výběrová směrodatná odchylka, RSD – rel. SD

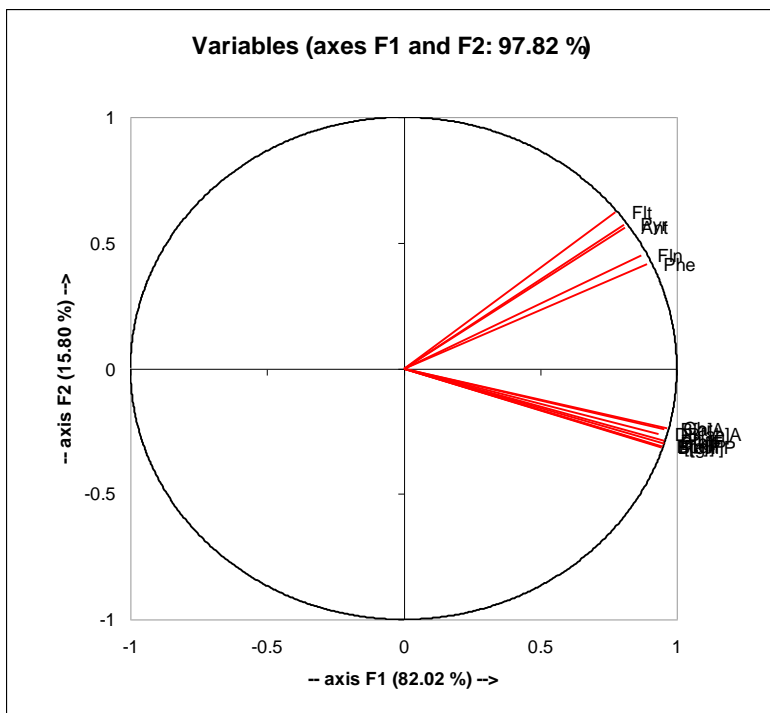
Vysoké hodnoty RSD opět, jako v předchozím případě, svědčí o značné nehomogenitě datové matice. Srovnáme-li průměrné hodnoty jednotlivých PAU v jarním a podzimním období, vidíme, že na podzim jsou hodnoty vyšší. Je to zřejmě způsobeno jinými zdroji emisí, než jsou velké energetické zdroje.

Na obrázku 3 je graf korelační kružnice, znázorňující korelaci původních proměnných (PAU) s prvními dvěma hlavními komponentami. První dvě hlavní komponenty nesou téměř 98% informace o tomto souboru, jsou tedy postačující pro dokonalý popis původních proměnných. Všechny proměnné vykazují vysokou vzájemnou korelaci, i když tvoří dvě skupiny proměnných. V jedné jsou Fln, Phe, Ant, Flt a Pyr (lehčí PAU), ve druhé ostatní, těžší polyaromatické uhlovodíky. Situace je tedy stejná jako na jaře.

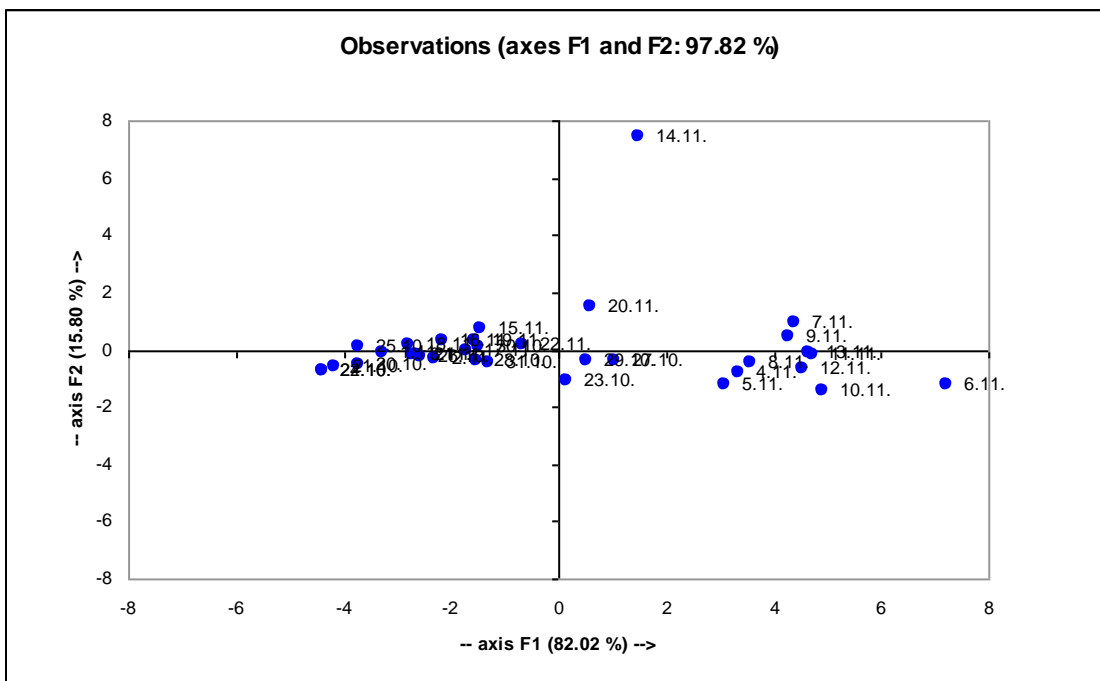
Na obrázku 4 je graf skóre, který ukazuje rozdělení všech „vzorků“ (vektory pozorování) v prostoru prvních dvou hlavních komponent. Z obrázku je patrné, že vzorky netvoří pouze jeden shluk, pocházejí tedy z více emisních zdrojů.

Je velký rozdíl v rozptýlení dat v jarním a podzimním období. Je to zřejmě způsobeno působením jiných emisních zdrojů, než tomu je v případě jarního období. Na obrázku jsou patrné dva body, které tvoří samostatné shluky.





Obr. 3 Korelační kružnice pro soubor PAU – podzim



Obr. 4 Graf skóre pro soubor PAU –podzim

Soubor kovy – jaro

Soubor obsahoval 28 datových vektorů. Každý z vektorů je charakterizován 13 koncentracemi kovů ( $\text{ng/m}^3$ ).  $\chi^2$  test potvrdil vysoký stupeň korelace mezi jednotlivými proměnnými ( $\chi^2 = 573$ ,  $\chi^2(\text{krit}) = 99$ ). Základní statistické charakteristiky souboru jsou uvedeny v Tabulce III.

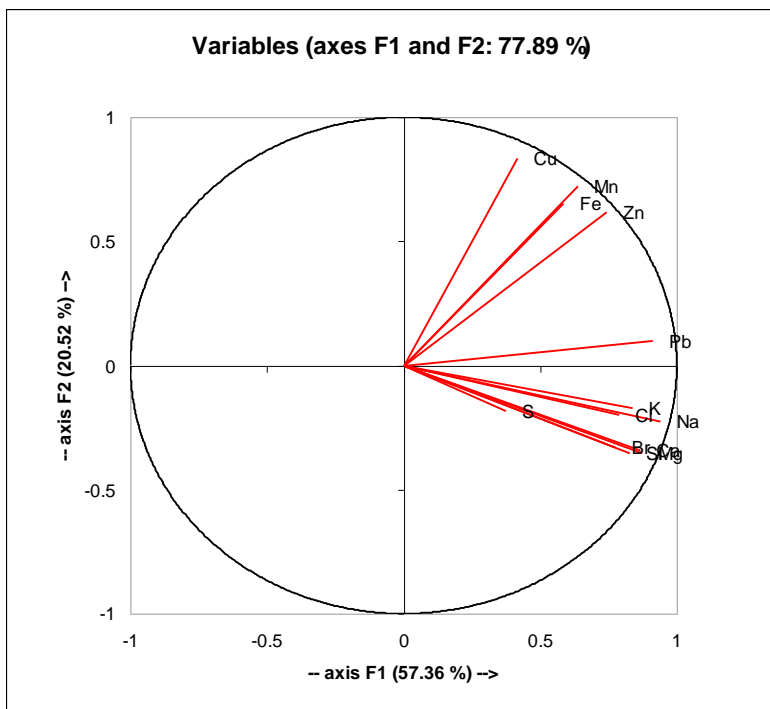
**Tabulka III Základní statistické charakteristiky souboru kovy – jaro**

Prvek	průměr	SD	RSD
Na	166	221	133%
Mg	69	98	142%
Si	256	310	121%
S	1545	881	57%
Cl	213	295	138%
K	271	198	73%
Ca	213	392	184%
Mn	15	11	76%
Fe	404	330	82%
Cu	7	5	72%
Zn	118	76	64%
Br	7	5	73%
Pb	35	23	66%

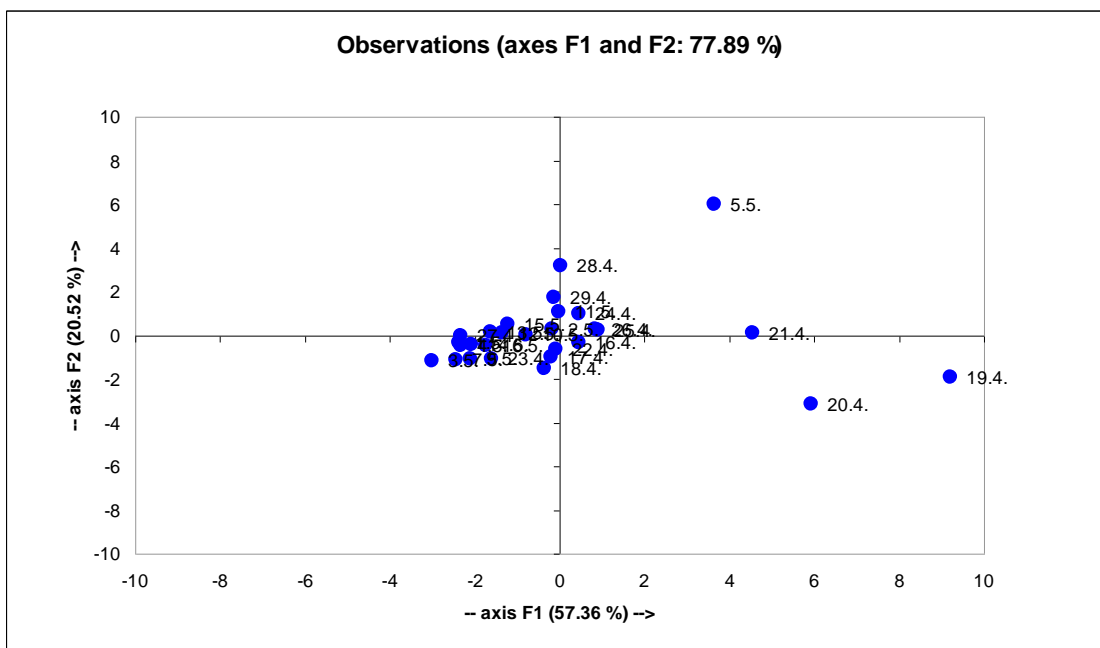
Soubor je charakterizován vysokými obsahy síry a železa, rovněž obsahy křemíku, draslíku a vápníku patří k vysokým. Veliká variabilita jednotlivých prvků svědčí o multizdrojové kontaminaci.

Na obr. 5 je znázorněna korelační kružnice pro první dvě hlavní komponenty, které nesou 78% veškeré informace. Zajímavé jsou dvě skupiny prvků, které spolu vysoce korelují. V první skupině to jsou Mn, Fe, Cu a Zn, ve druhé pak Na, Mg, Si, Cl, K, Ca a Br. Poněkud osamoceně stojí Pb a S. Síra vykazuje vysoké koncentrace ve všech vzorcích, o čemž svědčí relativně nízká variabilita (RSD = 57%).

Na obrázku 6 je ukázán graf skóre pro první dvě hlavní komponenty. Z grafu je patrné, že data tvoří jeden velký shluk, vnitřně členěný, a 4 body jsou jasně odlišné. To může svědčit i o náhlých změnách poměrů při monitorování. Další analýza bude provedena při zpracování dat lineární diskriminační analýzou.



Obr. 5 Korelační kružnice pro soubor kovy - jaro



Obr. 6 Graf skóre pro soubor kovy - jaro

Soubor kovy – podzim

Soubor obsahoval 29 datových vektorů. Každý z vektorů je charakterizován 11 koncentracemi kovů ( $\text{ng/m}^3$ ).  $\chi^2$  test potvrdil vysoký stupeň korelace mezi jednotlivými proměnnými ( $\chi^2 = 460$ ,  $\chi^2(\text{krit}) = 73$ ). Základní statistické charakteristiky souboru jsou uvedeny v Tabulce IV.

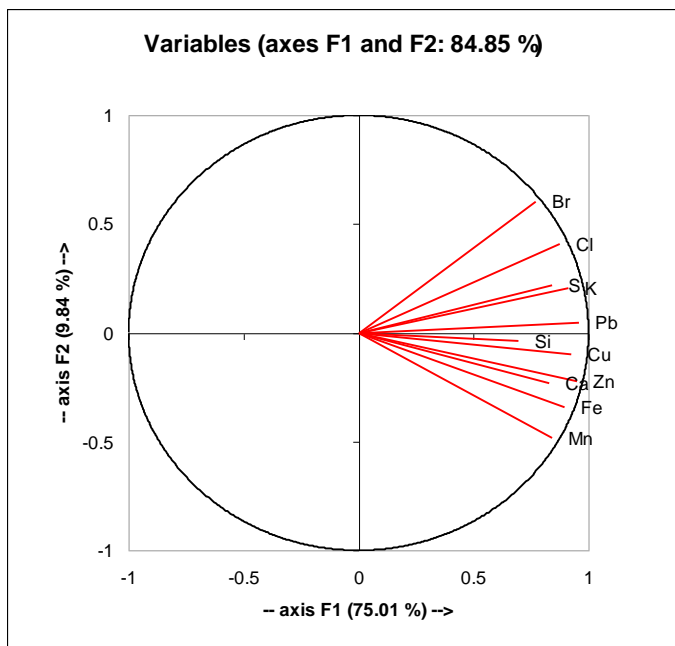
**Tabulka IV Základní statistické charakteristiky souboru kovy – podzim**

Prvek	Průměr	SD	RSD
Si	214	123	57%
S	1630	1145	70%
Cl	1214	1600	132%
K	406	483	119%
Ca	56	36	64%
Mn	25	29	112%
Fe	409	475	116%
Cu	17	15	86%
Zn	303	269	89%
Br	11	12	109%
Pb	72	70	98%

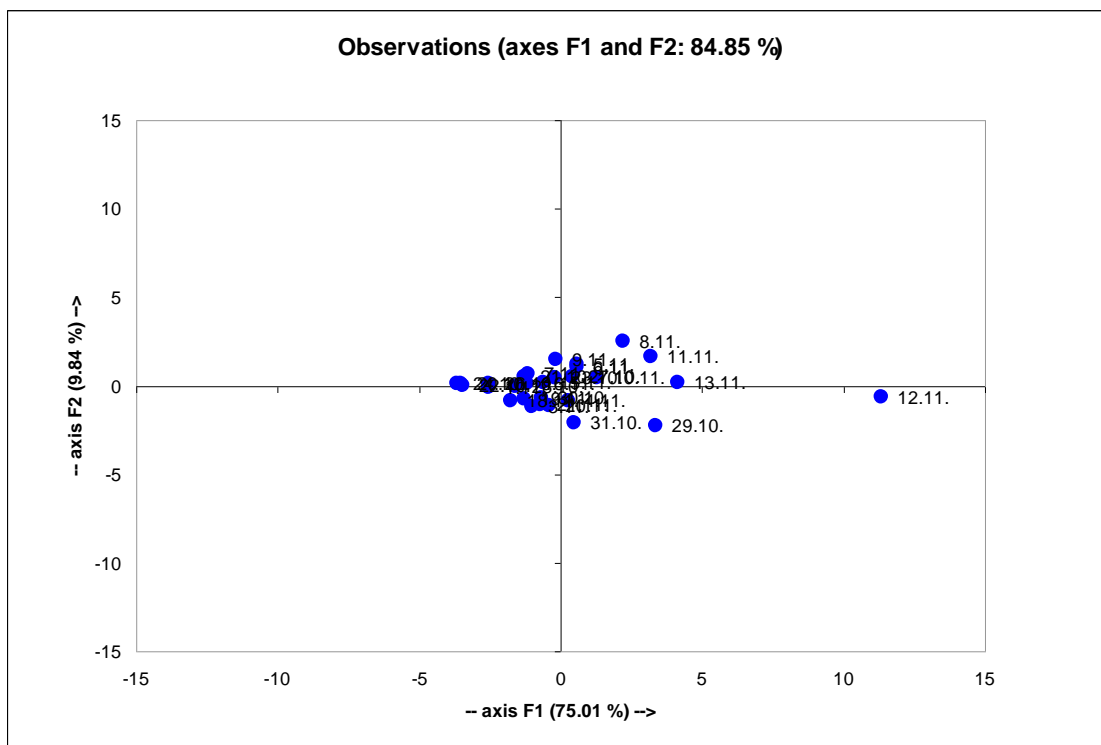
Soubor je charakterizován vysokými obsahy síry a chloru, rovněž obsahy železa, draslíku a zinku patří k vysokým. Profil kovů je zcela odlišný od jarního období. Chybí Na a Mg, jejichž obsahy byly ve většině vzorů pod mezí detekce. Veliká variabilita jednotlivých prvků svědčí o multizdrojové kontaminaci.

Na obr. 7 je znázorněna korelační kružnice pro první dvě hlavní komponenty, které nesou 85% veškeré informace. Zajímavé jsou dvě skupiny prvků, které spolu vysoce korelují. V první skupině to jsou S, Cl, K a Br, ve druhé pak Ca, Mn, Fe, Cu a Zn. Poněkud osamoceně stojí Pb a Si. Křemík vykazuje nízké koncentrace ve všech vzorcích, o čemž svědčí relativně nízká variabilita (RSD = 57%).

Na obr. 8 je ukázán graf skóre pro první dvě hlavní komponenty. Z grafu je patrné, že data tvoří jeden velký shluk, vnitřně členěný a 1 bod je jasně odlišný. Další analýza bude provedena při zpracování dat lineární diskriminační analýzou.



Obr. 7 Korelační kružnice pro soubor kovů - podzim



Obr. 8 Graf skóre pro soubor kovů - podzim

## 4.2 Klastrová analýza a lineární diskriminační analýza

Klastrová analýza spolu se znázorněním pomocí dendrogramů byly použity pro zařazení jednotlivých vektorů pozorování do příslušných skupin. Lineární diskriminační analýza (LDA) totiž potřebuje, jako vstupní údaj, znalost předběžného rozdělení do jednotlivých skupin. Tyto vstupní informace byly v LDA dále iterovány, až se dospělo k nejlepšímu rozdělení. Charakteristiky jednotlivých skupin sloužili k odhadu charakteristik emisních zdrojů.

### Soubor PAU – podzim

Soubor byl klastrovou analýzou rozdělen do 5 shluků. V tabulce V je ukázáno nejlepší rozdělení, které poskytla klastrová analýza nejbližšího souseda (k-means). V další tabulce VI jsou uvedeny charakteristiky jednotlivých shluků (těžiště shluků) neboli prototypy shluků.

### **Tabulka V Rozdělení souboru PAU – jaro do shluků**

SHLUK	1	2	3	4	5
VELIKOST	12	10	6	5	1
	26.10.	4.11.	20.10.	23.10.	14.11.
	28.10.	5.11.	21.10.	27.10.	
	30.10.	6.11.	22.10.	29.10.	
	31.10.	7.11.	24.10.	20.11.	
	2.11.	8.11.	25.10.	22.11.	
	3.11.	9.11.	1.11.		
	15.11.	10.11.			
	16.11.	11.11.			
	17.11.	12.11.			
	18.11.	13.11.			
	19.11.				
	21.11.				

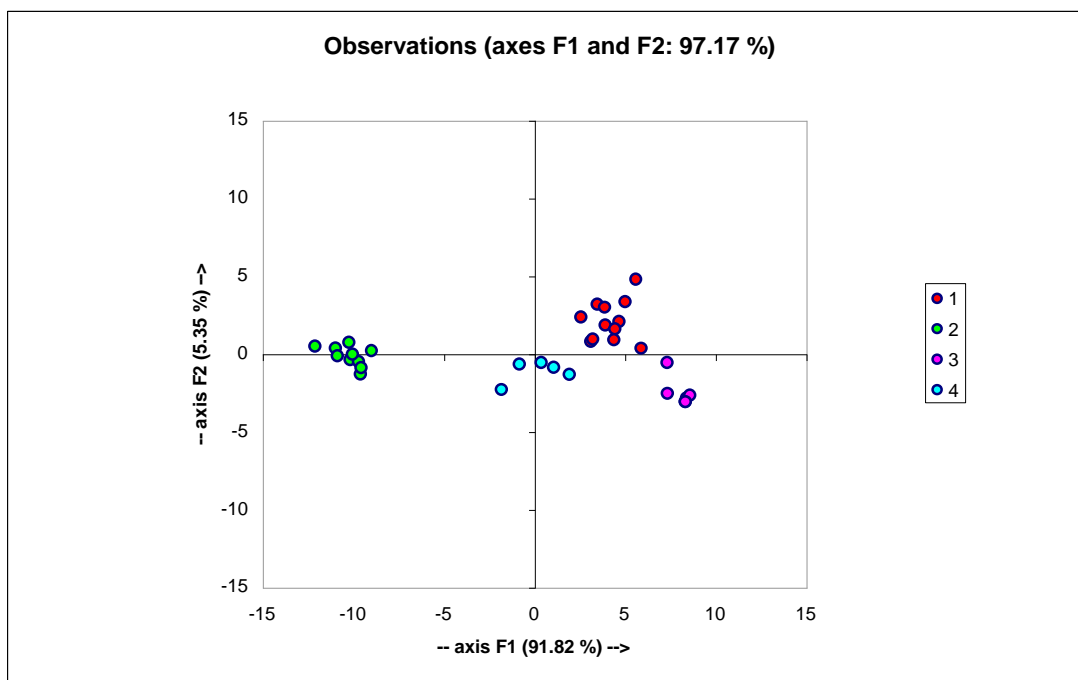
Označení vzorků: datum odběru

**Tabulka VI Charakteristiky prototypů**

PROTOTYP	Fln	Phe	Ant	Flt	Pyr	B[a]a	Chr
1	40.4	84.1	7.4	33.0	20.8	4.3	4.5
2	115.2	186.4	31.2	79.7	52.8	22.3	19.3
3	12.2	27.3	1.8	6.9	4.1	0.5	0.6
4	69.7	140.1	16.1	49.2	31.4	11.8	10.1
5	188.7	275.0	66.7	181.9	112.4	2.7	2.2

PROTOTYP	B[b]F	B[k]F	B[a]P	DB[ah]A	B[ghi]P	I[cd]P
1	3.5	1.8	3.0	0.2	2.4	3.0
2	15.8	8.0	14.8	1.3	9.8	13.5
3	0.5	0.2	0.4	0.1	0.3	0.4
4	7.6	3.9	7.1	0.6	4.5	6.2
5	0.1	0.1	0.1	0.1	0.1	0.1

Na obrázku 9 je ukázán výsledek LDA, tj. rozdělení do jednotlivých shluků v prostoru prvních dvou hlavních komponent.



Obr. 9 Výsledek LDA pro soubor PAU – jaro

Z výsledku LDA vyplývá, že soubor PAU – jaro (imísí data) je tvořen čtyřmi shluky, jasně vychýlený bod ze 14.11. nebyl vzat v úvahu. Dá se tedy předpokládat, že situace na monitorovacím místě je ovlivňována minimálně třemi emisními zdroji.

Soubor PAU – jaro

Soubor byl klastrovou analýzou rozdělen do 5 shluků. V tabulce VII je ukázáno nejlepší rozdělení, v tabulce VIII jsou uvedeny charakteristiky jednotlivých shluků.

Tabulka VII Rozdělení souboru PAU – jaro do shluků

SHLUK	1	2	3	4	5
VELIKOST	12	9	5	4	2
	27.4.	26.4.	17.4.	18.4.	16.4.
	1.5.	28.4.	20.4.	22.4.	19.4.
	3.5.	29.4.	21.4.	23.4.	
	7.5.	30.4.	25.4.	24.4.	
	9.5.	2.5.	5.5.		
	10.5.	4.5.			
	11.5.	6.5.			
	13.5.	8.5.			
	14.5.	12.5.			
	15.5.				
	16.5.				
	17.5.				

Označení vzorků: datum odběru

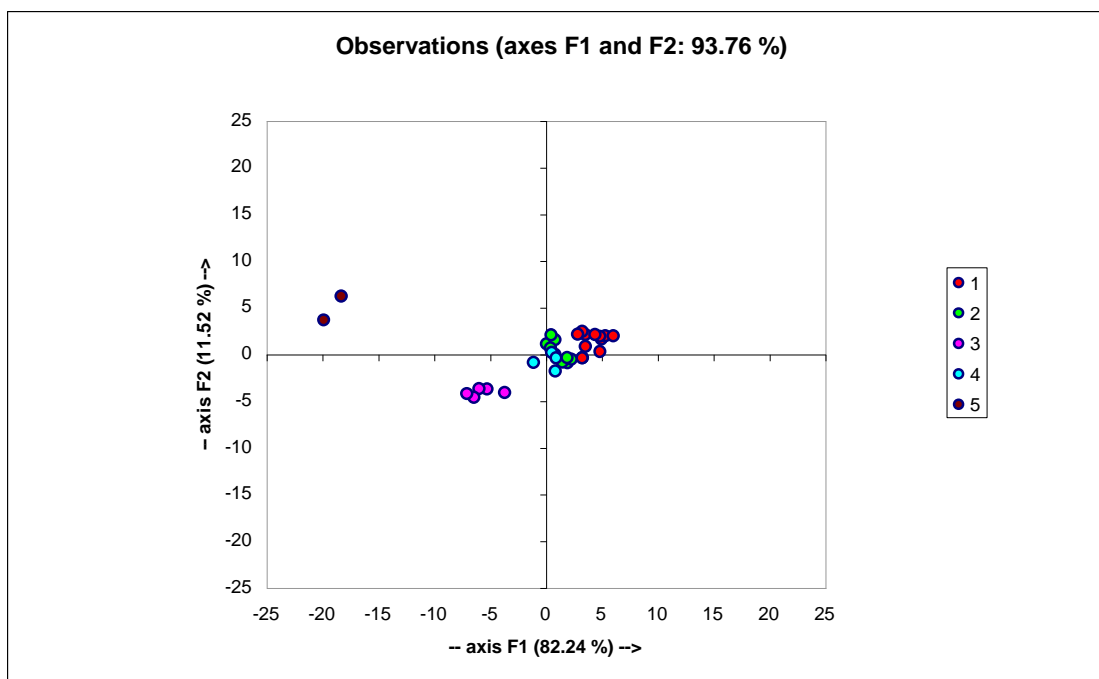
**Tabulka VIII Charakteristiky prototypů**

PROTOTYP	Fln	Phe	Ant	Flt	Pyr	B[a]A	Chr
1	39.0	71.8	1.8	18.3	7.0	1.2	1.8
2	72.5	127.7	3.1	33.7	11.2	1.7	2.5
3	86.7	127.5	3.5	39.5	16.3	6.1	7.1
4	50.4	81.7	2.7	23.8	11.8	3.6	4.6
5	152.9	161.2	7.1	48.5	20.1	9.5	10.7

PROTOTYP	B[b]F	B[k]F	B[a]P	DB[ah]A	B[ghi]P	I[cd]P
1	1.7	0.7	1.2	0.1	1.3	1.2
2	2.3	1.0	1.7	0.2	1.8	1.9
3	7.2	3.1	5.4	0.6	5.1	5.1
4	4.6	2.0	3.3	0.4	3.2	3.7
5	10.4	4.3	7.3	1.0	6.5	7.6

Na obrázku 10 je ukázáno rozdělení do čtyř shluků jako výsledek LDA.





Obr. 10 Výsledek LDA pro soubor PAU – jaro

Z obrázku zřetelně vyplývá, že imisní data jsou rozdělena opět do čtyř shluků, tentokrát s dvěma vychýlenými body.

#### Soubor kovy – podzim

Soubor byl klastrovou analýzou rozdělen do 5 shluků. V tabulce IX je ukázáno nejlepší rozdělení, které vzniklo v procesu výpočtu metodou nejbližšího souseda. V tabulce X jsou uvedeny charakteristiky jednotlivých shluků.

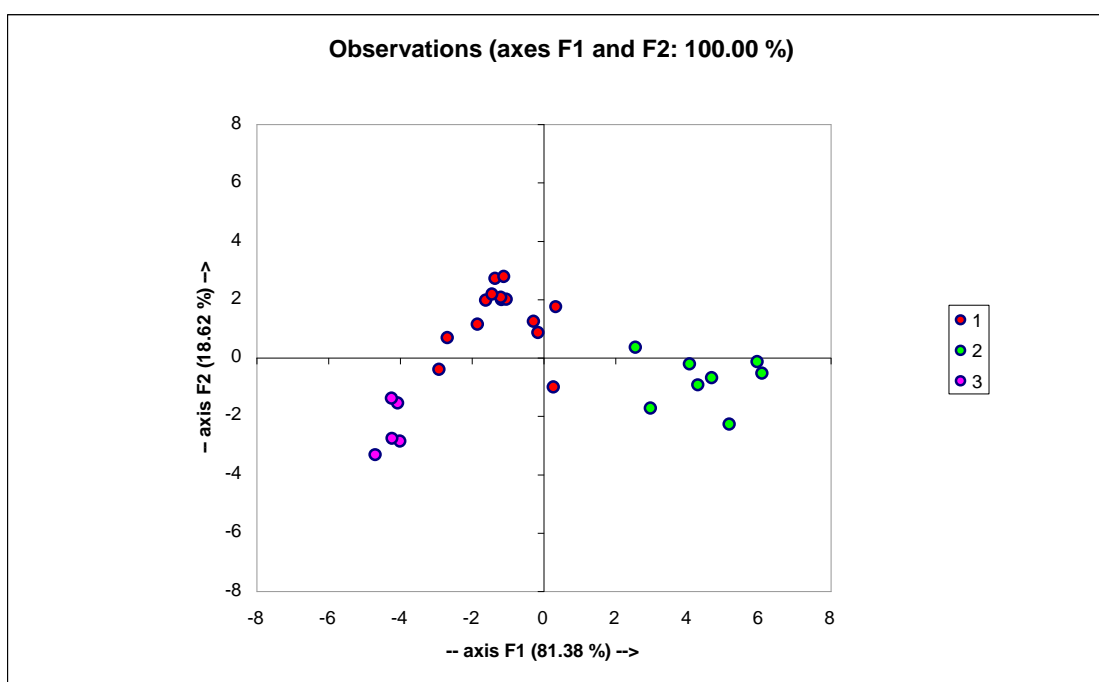
#### **Tabulka IX Rozdělení souboru kovy – podzim do shluků**

SHLUK	1	2	3	4	5
VELIKOST	14	8	5	1	1
	21.10.	27.10.	20.10.	12.11.	29.10.
	23.10.	5.11.	22.10.		
	28.10.	6.11.	24.10.		
	30.10.	8.11.	26.10.		
	31.10.	9.11.	1.11.		
	3.11.	10.11.			
	4.11.	11.11.			
	7.11.	13.11.			
	15.11.				
	16.11.				
	18.11.				
	19.11.				
	20.11.				
	21.11.				

**Tabulka X Charakteristiky prototypů**

PROTOTYP	Si	S	Cl	K	Ca	Mn	Fe	Cu	Zn	Br	Pb
1	187	1442	419	241	55	24	367	13	260	5	49
2	291	2145	2738	665	71	20	380	24	384	25	107
3	63	404	29	63	13	1	34	1	31	0	8
4	375	5583	6291	2355	144	135	2493	69	1371	40	375
5	573	2298	993	399	92	98	1036	38	558	9	141

Na obrázku 11 je ukázán výsledek LDA a rozdělení do tří shluků.



Obr. 11 Výsledek LDA pro soubor kovy – podzim

Z obrázku vyplývá jasné rozdělení imisních dat do tří shluků, data ze 29.10. a 12.11. nebyla vzata při výpočtu v úvahu, neboť tvoří jasné vychýlené body.

Soubor kovy – jaro

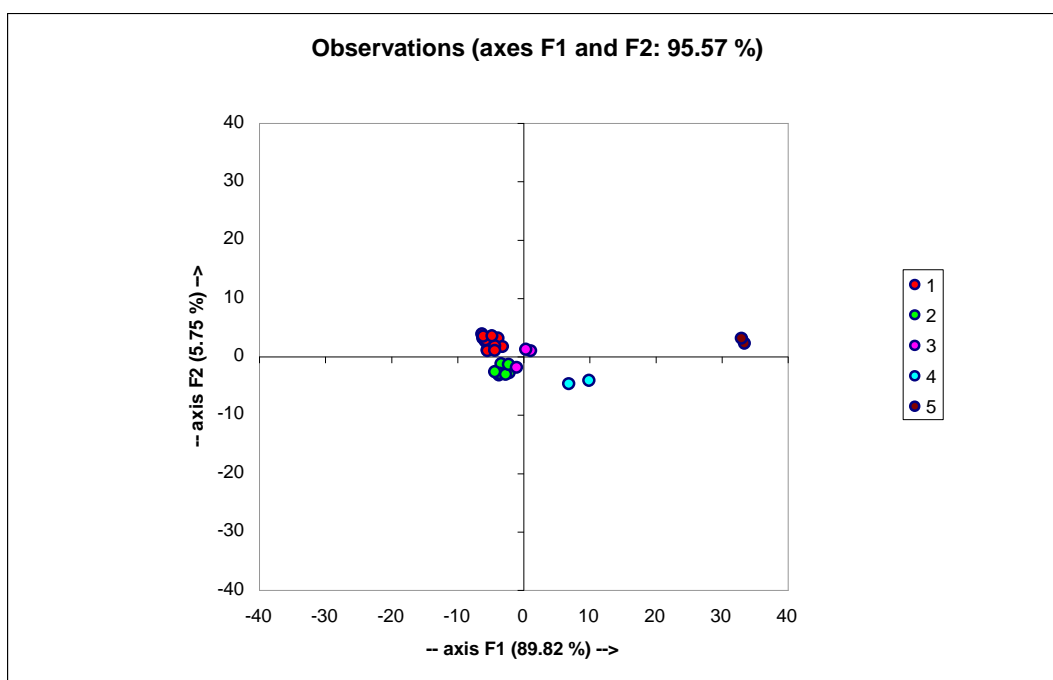
Soubor byl klastrovou analýzou rozdělen do 5 shluků. V tabulce XI je ukázáno nejlepší rozdělení do shluků pomocí metody nejbližšího souseda. V tabulce XII jsou uvedeny charakteristiky jednotlivých shluků. Na obrázku 12 je výsledek LDA a rozdělení do 5 shluků.

**Tabulka XI Rozdělení souboru kovy – jaro do shluků**

SHLUK	1	2	3	4	5
VELIKOST	13	8	3	2	2
	23.4.	16.4.	28.4.	21.4.	19.4.
	27.4.	17.4.	29.4.	5.5.	20.4.
	1.5.	18.4.	11.5.		
	3.5.	22.4.			
	4.5.	24.4.			
	7.5.	25.4.			
	9.5.	26.4.			
	10.5.	2.5.			
	12.5.				
	13.5.				
	14.5.				
	15.5.				
	16.5.				

**Tabulka XII Charakteristiky prototypů**

PROTOTYP	Na	Mg	Si	S	Cl	K	Ca	Mn	Fe	Cu	Zn	Br	Pb
1	61	24	123	1079	67	130	40	7	200	5	66	4	21
2	126	50	169	2424	277	384	101	14	298	7	128	9	39
3	99	52	286	903	117	151	233	27	943	11	148	3	23
4	384	137	466	1503	316	493	458	41	1077	17	300	9	80
5	887	396	1209	2067	946	695	1511	21	668	7	194	18	83



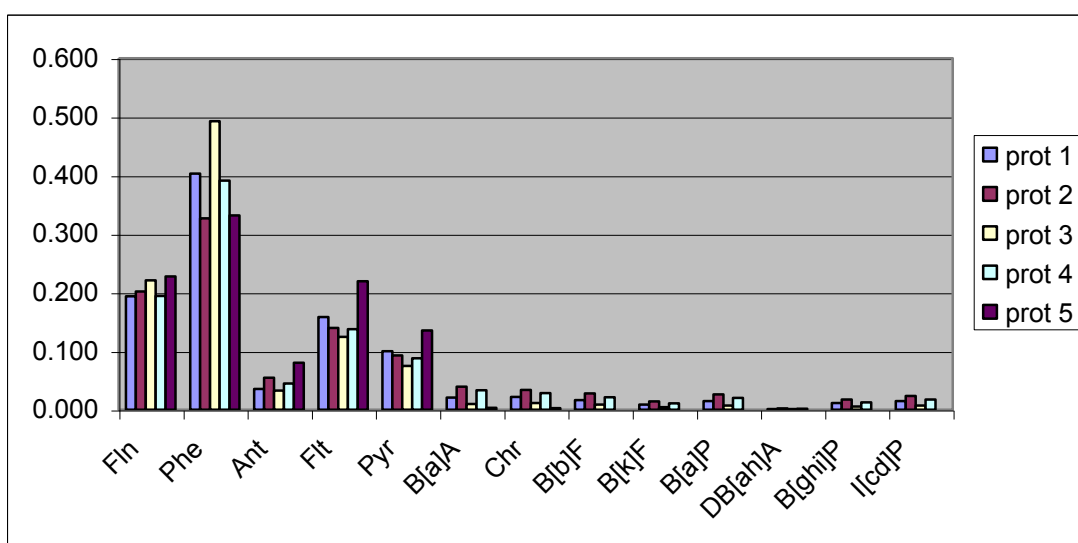
Obr. 12 Výsledek LDA pro soubor kovy - jaro

Z obrázku vyplývá zajímavá skutečnost: tři největší klastry tvoří poměrně kompaktní celek, jsou zřejmé i dvě dvojice odlehlých bodů.

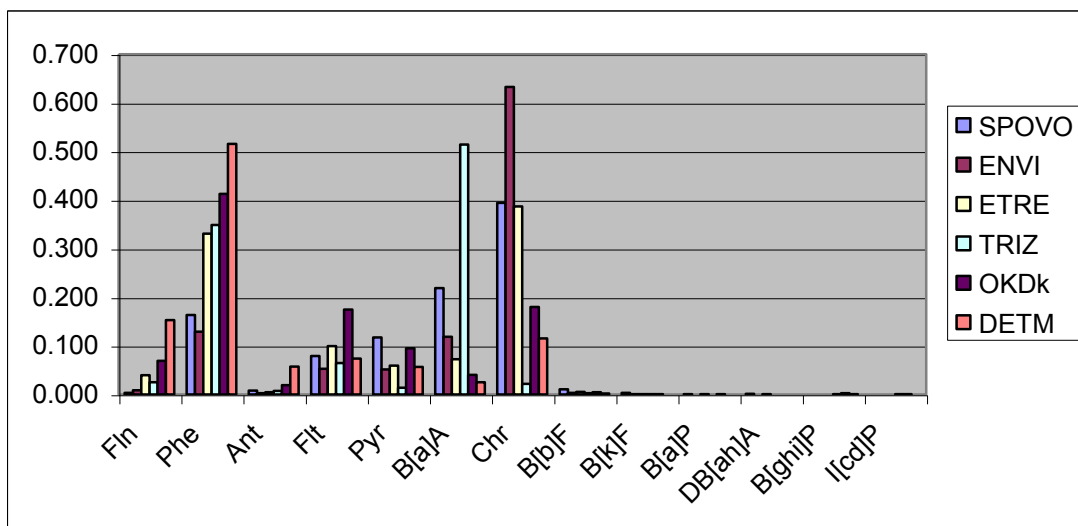
## Srovnání prototypů jednotlivých souborů s podpisem zdroje

Na obrázcích 13 – 18 je uvedeno srovnání prototypů vypočtených LDA s podpisem zdrojů pro PAU i kovy. Ve všech případech bylo použito poměrné zastoupení jednotlivých složek jak v prototypyech, tak v podpisech zdrojů. Pro lepší přehlednost uvádíme podpisy zdrojů vždy pod každým souborem imisních dat.

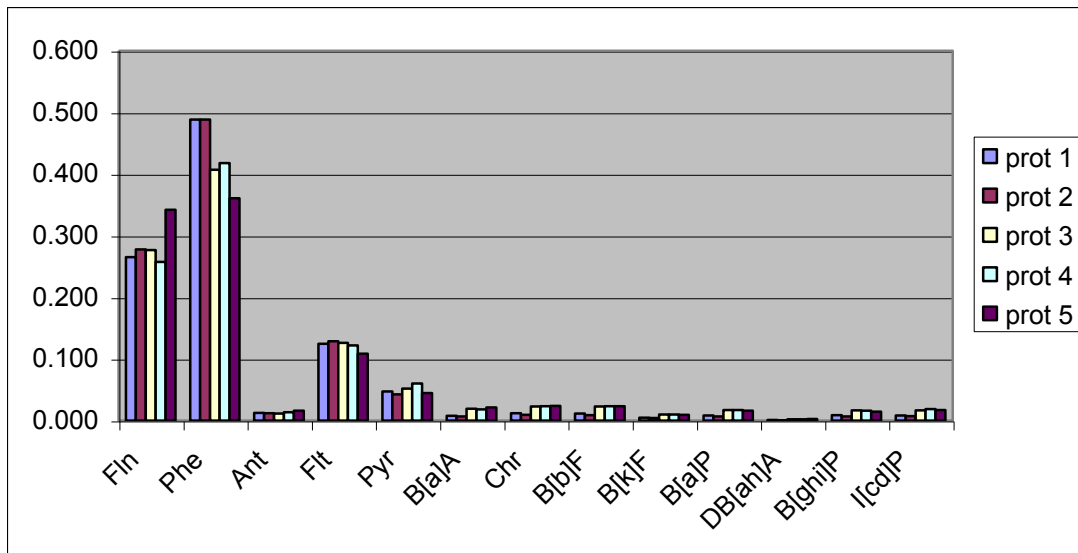
Soubory pro PAU



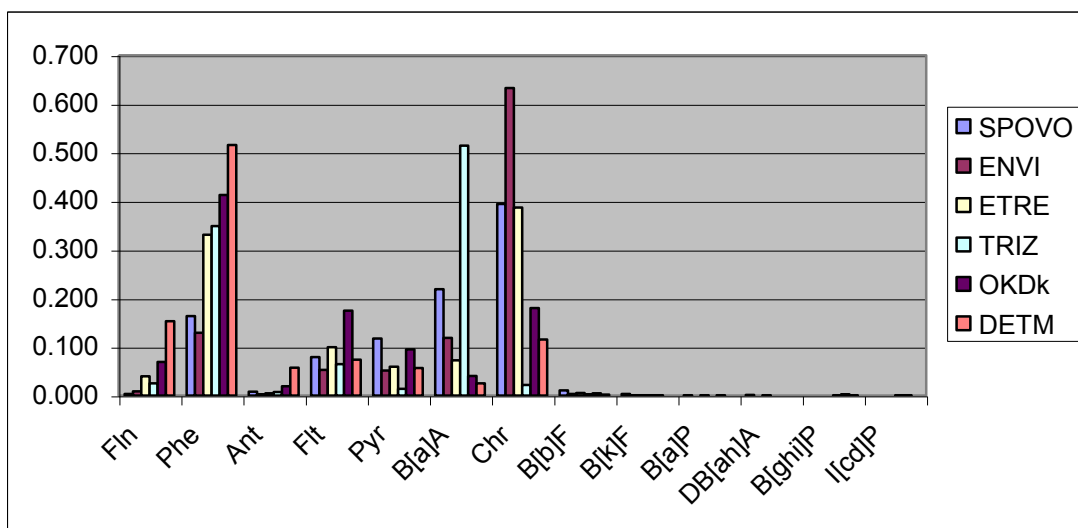
Obr. 13 Poměrné zastoupení PAU v soubor PAU – podzim



Obr. 14 Poměrné zastoupení PAU ve zdrojích emisí



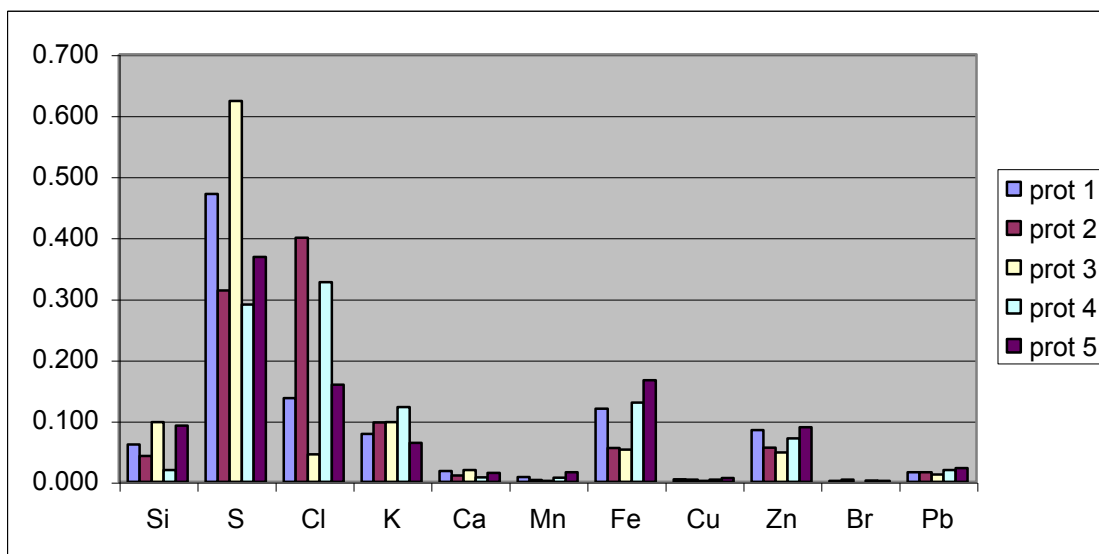
Obr. 15 Poměrné zastoupení PAU v souboru PAU – jaro



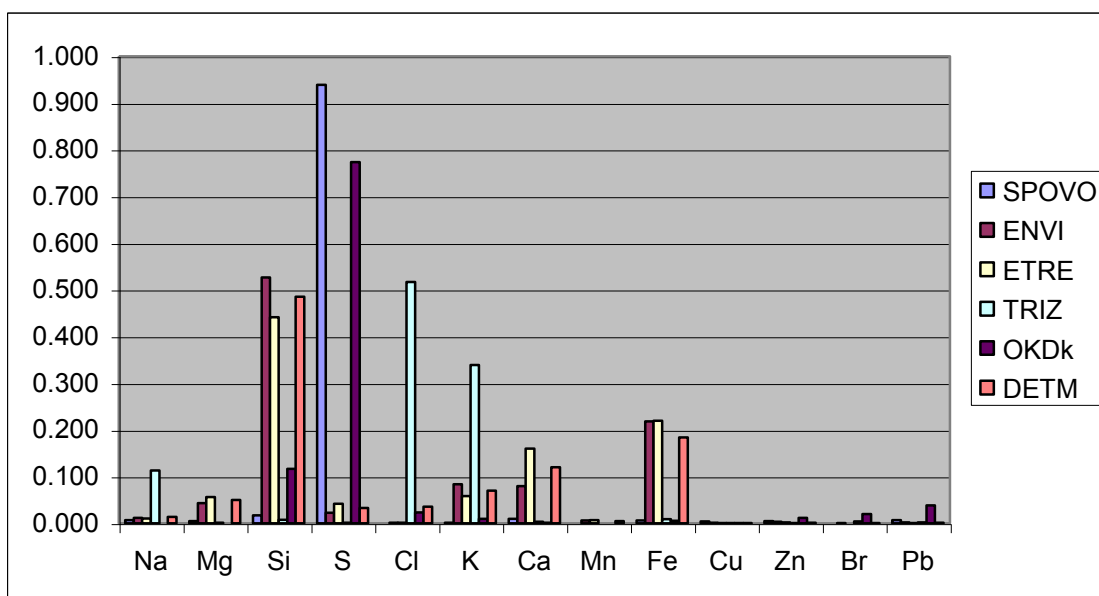
Obr. 14 Poměrné zastoupení PAU ve zdrojích emisí

Z obrázků je patrné, že profily imisních dat jsou zcela rozdílné než profily zdrojů. Lze se tedy oprávněně domnívat, že imisní data v případě PAU nepocházejí z velkých energetických zdrojů.

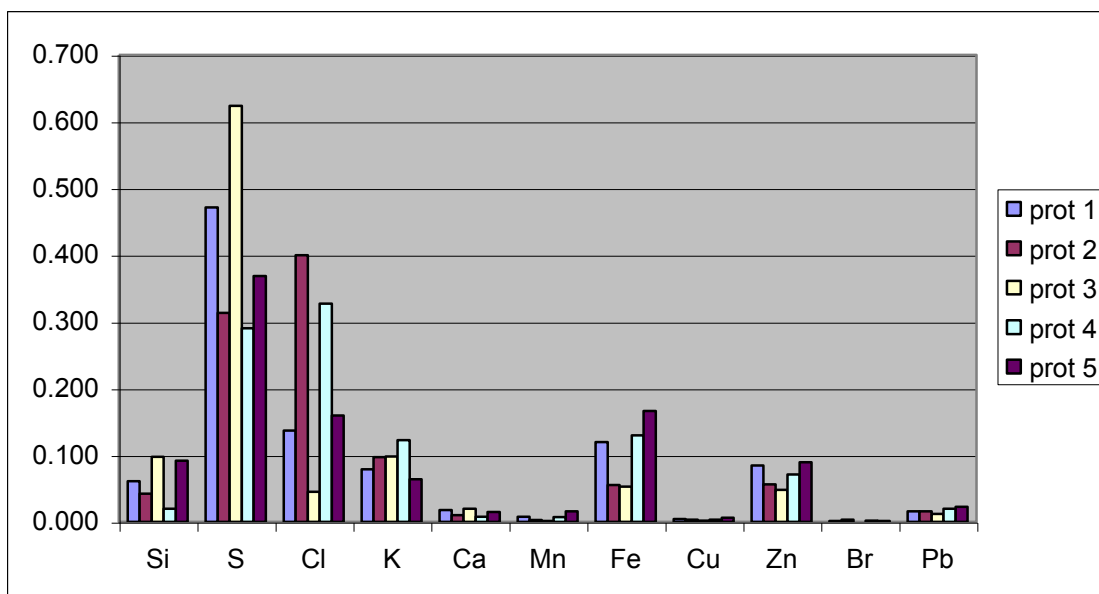
## Soubory pro kovy



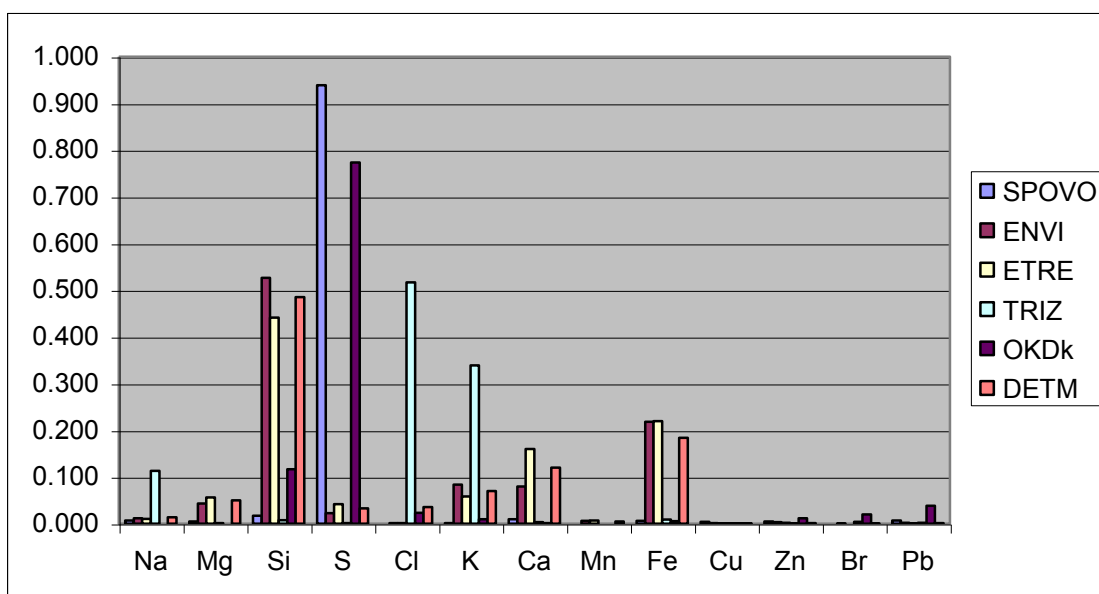
Obr. 16 Poměrné zastoupení kovů v souboru kovy – podzim



Obr. 17 Poměrné zastoupení kovů v emisních zdrojích



Obr. 18 Poměrné zastoupení kovů v souboru kovy – jaro



Obr. 17 Poměrné zastoupení kovů v emisních zdrojích

Z obrázků je patrné, stejně jako v případě PAU, že profily imisních dat jsou rozdílné od profilů emisních zdrojů. Tak např. v souboru kovy – jaro úplně chybí Na a Mg, obsahy křemíků jsou v emisních zdrojích podstatně vyšší než v imisích. Zajímavý je i rozdíl v obsazích Zn: Zn není v emisních zdrojích prakticky přítomen.

Charakteristické indexy

Pro odhad původu zdroje emisí PAU se někdy používají tzv. indexy. Jsou to vlastně poměry koncentrací některých polyaromatických uhlovodíků. Nejčastěji se používají poměry **BaP/BghiP** a **PyR/BaP**. Z dřívějších měření je známo, že hodnoty poměru **BaP/BghiP** pro zdroje z dopravy nabývají hodnot 0,2 – 0,4, kdežto pro zdroje z lokálních topenišť jsou hodnoty větší než 0,8. Pro poměr **PyR/BaP** se hodnoty indexu pro dopravu pohybují v rozmezí 2 – 12, pro lokální topeniště 0,2 – 0,8. Vedle toho se dá použít index **Phe/Ant**, který indikuje např. zpracování dehtových frakcí. V následující tabulce XIII jsou vypočteny indexy pro jednotlivé prototypy v podzimním a jarním období.

**Tabulka XIII Indexy pro jednotlivé prototypy souborů PAU a pro emisní zdroje**

## Emisní zdroje

ZDROJE	BaP/BghiP	PyR/BaP	Phe/Ant
SPOVO	4.7	140	20.1
ENVI	0.1	2535	46.1
ETRE	7.0	93	75.7
TRIZ	0.7	55	46.8
OKDk	0.2	179	21.9
DETM	0.7	209	9.0

## PAU – podzim

PROTOTYP	BaP/BghiP	PyR/BaP	Phe/Ant
1	1.3	6.9	11.3
2	1.5	3.6	6.0
3	1.2	10.9	15.0
4	1.6	4.4	8.7
5	0.0	2248	4.1

## PAU – jaro

PROTOTYP	BaP/BghiP	PyR/BaP	Phe/Ant
1	0.9	5.9	39.6
2	1.0	6.4	40.6
3	1.1	3.0	36.2
4	1.0	3.6	30.7
5	1.1	2.8	22.6

Ze srovnání jednoznačně nevyplývá, zda imisní data nějakým způsobem korelují s daty emisními.



#### 4.3 Receptorové modelování

Výsledky zpracování imisních a emisních dat receptorovým modelem CMB 8.2 jsou uvedeny ve zvláštních přílohách.

### 5 Souhrn

Soubory pocházející z oblasti Ostravy byly podrobeny zpracování multivariační analýzou. Bylo zjištěno, že imisní data pocházejí minimálně ze tří emisních zdrojů, z nichž ani jeden nepatří mezi velké energetické zdroje. Další zpracování těchto dat musí proto zahrnovat jiné zdroje, např. lokální topeniště, dopravu, apod.