Charles University
Faculty of Natural Sciences
Department of Analytical Chemistry

Summary of Doctoral Thesis

STUDY AND DEVELOPMENT OF SAMPLING METHODS FOR PERSISTENT ORGANIC COMPOUNDS
METHODOLOGY OF DATA ANALYSIS FOR POPS ASSESSMENT

Dipl. Ing. Tomas Ocelka

Prague 2010

TABLE OF CONTENTS

Section Page

1 INTRODUCTION/THEORY

The introductory and theoretical part of this thesis is devoted to Persistent Organic Pollutants (POPs) playing a significant role in environmental protection and human health. The issue of polychlorinated biphenyls (PCBs), polychlorinated dibenzo-p-dioxins and dibenzofurans (PCDD/Fs), pesticides (OCPs) as well as hexachlorobenzene (HCBs), and polybrominated difenylethers (PBDEs) is described in respect to their persistence, bio-accumulation, bio-concentration (with effect of bio-magnification), and long-range transport in the environment[1-7]. Several actions are discussed; mainly the Stockholm Convention as ratified by the Czech Republic that entered into force on 17 May 2004, concerning: pesticides (aldrin, endrin, dieldrin, chlordane, DDT, heptachlor, mirex, toxaphene, HCB), industrial products (PCBs) and unintended by-products (PCDDs and PCDFs); moreover PAHs, PBDEs, lindane, and chlorophenols.

Due to the recent broad use of some pesticides, details about technical mixtures of p,p'-DDT/ o,p'-DDT and their metabolites are revealed[8-13], as well as their use in the Czech Republic[14]. The other special group of pesticides – the HCH is also discussed[15]; among newly assessed POPs, the PBDEs are addressed, despite not being included on the UNEP 12 list of the most frequently targeted chemicals[16-19]. Extended information about their properties are refereed [20-25], as well as toxicity[26-27]. The endocrine disruption chemicals and their effects are briefly mentioned[20-22, 24]. The bioavailability of a particular chemical compound, depending on their chemical properties, is also discussed[7, 28-35]. Among toxicological relevance, tumours initiation, neurodevelopmental toxicity, and thyroid hormone imbalance[25, 36-40] is mentioned, as well as the neurotoxic effects of PBDEs, as similar to those observed for PCBs[41-42].

As a consequence of inventory and regulatory actions to POPs, the importance of data reliability is discussed, mainly focusing on their sampling. The occurrence of POPs at low levels in the environment are connected to methodologies, based on grab and passive sampling in water[43-53], as well as biotic organisms(MBOs)[54-64] or sediments[65-71]. Features of BMOs can help to estimate the lethal dose, approximated by application of the gnostic approach[72-84], as based on recent studies focused on comparison PS and BMOs[85-86]. In general, the *in-situ* passive sampling approach is then broadly discussed, as dosimeters mostly applied to monitor air and water in environment, due to many advantages over the standard (grab) sampling methods[69, 87-96]. Applicability of the semipermeable membrane device (SPMD) as a tool for hydrophobic, lipophilic organic contaminants (and their mixtures) in water of a very low concentration due to bioaccumulation ability is explained together with enhanced applicability to bioassay methods[97-102], as well as methods based on toxicological testing of organic extracts[85-86, 103-104]. Various bioassays and toxicity on extracts were applied within this thesis, too. Some of them were published recently[78, 105-111].

For practical use, general rules and objectives related to passive sampling are discussed, mainly for SPMDs. A tool for lipid reduction is also discussed [112-117]. The grab sampling is illuminated as a good base for wide standardisation[118-119], with a challenging extension to passive sampling. Selection of PS, based on the partition coefficient K_{OW} is described to reveal applicability of given compounds[64, 83, 120-126]. Developmental phases of SPMDs is briefly mentioned[77, 84, 127-128], resulting in 97-99% triolein purity as sequestrant, filled in mostly using layflat LDPE [81, 83], as applied recently[129-131], and subjected to various patents[79-80, 132-138]. Then, basic characteristics of SPMDs and fundamentals factors for its use are given[139-140], as well as a tool that yields time the weighted average (TWA) concentration [65, 67-71, 90, 96, 141-142].

The advantage of PS over grab sampling, arising from an accumulation of contaminants an short-time peaks elimination is emphasized, as well as the applicability for - glacial or mineral water assessment[83, 121]. The methodology attempting to predict the toxicity of compounds based on physicochemical properties and descriptors of compounds, in both aquatic and terrestrial species[80, 143-144] is also introduced. Standardisation and QA/QC aspects (mainly as the Interlaboratory comparison) are mentioned as a future challenge[145-150].

Finally, the necessity of evaluation of POPs using complex statistical and alternative methods is discussed: exploratory and marginal statistical analysis[85-86, 103-104, 151-152], transformation, followed by Principal Component Analysis (PCA), FA (Factor Analysis), Cluster Analysis (CLU) [153-154], revealing relations between variables and observations. The concept of statistical approach is theoretically and briefly described[152, 155-156]. For a complex evaluation of multivariate data, distinctive for POPs, the usage of robust methodology is expressed, if classical statistics do not pass data normality. The central limit theorem, assuming randomness, independence and stationary conditions[22, 157-166] with an a priori required data model, is mentioned. The necessity for handling and interpreting small scale and non-normal data is the reason for adoption of complementary method based on the gnostic approach[167-170], with its own theoretical background, procedures, and recent applications[139-140, 171-176]. A rough description of gnostic theory is also reflected in this work, whereas details can be found in literature[153-154, 167-170, 177]; realized applications attracted its use for environmental data[139, 172].

2 EXPERIMENTAL PART

A. Semipermeable membrane device (SPMDs)

The following standard design of SPMDs was used for all experimental work: the lay flat thin-walled tube of nonporous (with transient cavities) material LDPE, filled with 1 ml of synthetic lipid – triolein, neutral triglyceride (1,2,3-tri-[cis-9-octacenoyl]glycerol) of high purity (>97%), which makes a thin film in membrane. The parameters: width 2.5 cm (lay-flat), overall length 91 cm and wall thickness 75 -90μm, overall sampling area is about 460 cm^2, total about 4.5 g. An exploded view of overall SPMDs is given in Figure 1. Transports of contaminants were through transient pores, with specific diameter approx. 10^{-9}m (similar to postulated size of transient cavities in biomembranes as 9.8 10^{-9}m)[153-154, 177]. The use is confronted with non-standard PS despite some attempts at their calibration[121, 124-125, 178-179]. On the other hand, the use as a tool for effective method of separation of organic contaminants from lipids[180-181], based on SPMDs, is demonstrated. Figure 2 brings used sampling arrangements. For underground water, the protective shroud was developed within this thesis. This new design (to be soon patented) allows installation two (instead one) SPMDs (e.g. 1 for toxicity and 1 for chemicals).

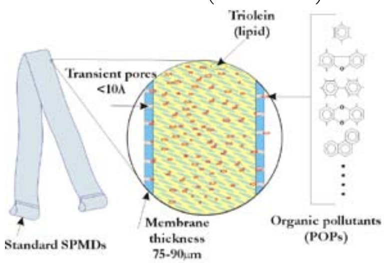


Figure 1 Exploded view of standard SPMDs

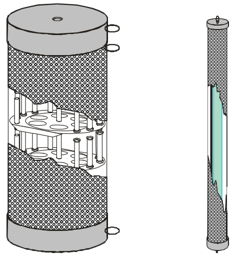


Figure 2 Sampling arrangements – protective shrouds for SPMDs for surface (left) and underground water (right)

Based on detailed insight of chemical uptake, the ambient concentration was estimated from:

$$C_{W,i} = C_{S,i} V_S / R_{S,i} t \qquad \text{Eq. 1}$$

The term $R_{S,i}$ is the *sampling rate* parameter of contaminant *i* and provides a conceptual link between classical extraction techniques and passive sampling with SPMDs. Sampling rate is in $L.d^{-1}$, which can be interpreted as the volume of water (L), dialysed per day, for a particular contaminant (*i*) and time(*t*); $C_{S,I}$ is concentration found in SPMDs, V_S denotes total volume of SPMDs. Sampling rates were used to calculate ambient concentration of contaminant $C_{W,i}$. They were used for a large number of compounds: PAHs[120, 122], PCDDs/PCDFs[132, 137, 182-195], PCBs[73, 121, 137, 193, 195-201] and a number of pesticides[137, 195-196, 201-203]. The calculations were adopted from recent publications[121, 199, 204-205]. In the latter phase, the PRCs (Performance Reference Compounds) approach was used; based on the principle of use the kind of internal standards

(PRCs), added prior to the exposure into SPMDs devices considered SPMDs as calibrated *in-situ*. Evaluation is then based on the amount of dissipated PRCs and those found at the beginning of the sampling[127, 137, 183, 190, 206-209]. Adopted general requirements for deployment were recently described[81, 121, 189, 210-211].

B. Used biota; bioassays

Biotic organisms, *Dreissena polymorpha* were used for comparison to the SPMDs; they were collected in a sand-pit with very clean water. Small shells were placed in polypropylene cages (100 pieces/cage) and fixed at the same float as the SPMDs.

To assess toxicity after exposure, the SPMDs were also used in various unique studies within this thesis[79]: WWTP in Ostrava, and, on the contrary, in a clean environment[144, 212-213]. Their selection was based on the particular sensitivity of SPMDs to exposed POPs. The toxicological response was based on EC_{50}, expressed as the V_{TOX} parameter that was developed during our research recently[107]. The following tests were used[79, 101, 104, 107, 110-111, 144, 213-215]: *Vibrio-fischeri*, acute toxicity test with cladophora *Daphnia pulex*, a test with rainbow trout *Oncorhynchus mykiss* and a test of inhibition of algal growth of *Scenedesmus subspicatus* and/or *Selenastrum capricornutum*. A test with bioluminescent bacterium was carried out following the standard procedures (ISO 11348)[79, 216].

C. Sampling protocol

Standard protocol for sampling, handling and analysis was considered as the main prerequisite for producing comparable results of appropriate quality[217], following QA/QC protocol developed within this thesis, and used for accreditation. It can be seen as a recommendation, verified in practical applications.

Ref	Operation	**Description/ Requirements /Blank (frequency)** site/project/lot/sample set *)= if applicable
1	Fabrication of SPMDs	Defined QA/QC conditions / FaBl (0/1*/1*/0)
1	Storage	--/Air-tight can, about -10°C/PrBl (0/1*/1*/0)
2	Transport	--/Air-tight can, ≈0°C/ TrBl (1/1*/0/0)
2	Deployment/Retrieval	SPMDs into or from the rack or holder/short time/FBl (1/1*/0/0)
2	Exposure	Up to 30 days/security of sampling site, temperature recording/No blanks required
3	Exterior cleaning	Removal of all external impurities/rinsing by water, fast operation/ReBl (0/1*/0/0)
4a	Dialysis	250 ml solvent, 2 exchange/24 hour per one solvent/FaBl*, PrBl (0/1*/1*/0)
5	Dialysate (chemicals) Dialysate (bioassay)	Collection of all solvent/spike by isotopic standards/see 4a Collection of all solvent/no spike, taking of aliquot/ see 4a
6	Clean-up	Oligomers, impurities removal/by analytical method/see 4a + ReBl (0/0/0/1)
7	Fractionation (chemicals) Fract. (immunoassay)	Isolation of analytical fractions/by analytical method/see 6 Isolation of analytical fractions/by analytical method/see 6
8	Pre-concentration	10μl/nonane solvent/see 6

Ref	Operation	**Description/ Requirements /Blank (frequency)** site/project/lot/sample set *)= if applicable
9	Chemical analysis	Quantification of chemicals/by analytical method/see 6
10	Immunoassay/bioassay	Toxicological response/by analytical method/see 6
11	Data analysis, reporting	Ambient concentration by Rs/PRCs; PCA, FA, GA/data pre-treatment/Evaluation of QA/QC, including blanks
SP	Sampling plan	New definition of sampling plan based on new foundations/given from 11 and new project purposes/given from 11

Table 1 Abbreviated protocol for handling, sampling and analysis of SPMDs

D. Analytical methods

Determination of PAHs was carried out by HPLC-FLD method, in combination with method GC/MS/MS. PCDD/Fs, PCBs and OCPs were analyzed by GC/MS/MS on GCQ or PolarisQ (Thermoquest). The clean-up method and optimisation of MS/MS detection were described elsewhere[218]. Used methods were accredited according to ČSN EN ISO/IEC 17025.

E. Statistical and gnostic methods

Within this thesis, some data exhibited distinctive features: some measured values were very *low* or *not detected*, requiring treatment as low-censored; their ranges were either narrow or extremely broad, *very low size* (rarely repeated measurements on one site), distribution of data was seldom *normal* or *log-normal*, data tended to *polymodality (non-homogeneity)*, and were *skewed*. Therefore data analysis required exact, robust and reliable procedures – from data transformation to evaluation of relevant quantitative characteristics and interpretation, as also theoretically described[103, 139-140, 153-154]. Unlike statistic approaches, recently broadly used, the gnostic approach, based on novel fundamentals[103, 139-140] was also used[219-221]; due to rigid fundamentals it hides strong potential for treatment and evaluation of the environmental data [85, 103-104, 139].

For statistical analysis, the exploratory data analysis was first carried out to evaluate the actual sample distribution. A curve of the probability density function for each analyzed parameter was examined and the analysis was based on histogram, symmetry plot, halfsum plot, Box-and-whisker plot, quantile plot and rankit Q-Q plot, while the coefficient of skewness measures the asymmetry of the observations. The Jarque-Berra test ($\alpha = 0.05$) was preliminarily used to test the normality of concentration distribution within each element. For transformation, the Box-Cox[153-154, 166, 222] was used as the most frequently applied statistical method. After transformation, a series of complex multivariate methods were used.

Principal component analysis (PCA), applied to analyses for simplifying data matrix of a large number of observed variables to a set of linear combinations of the original variables, especially for POPs, where a number of variables exceeded 100 (including congeners of PCDD/Fs and PCBs). The notion of *explained variance* ranged from 0 (no explanation) to 1 (complete explanation of variable) provides information, how a combination of particular variables fits the analyzed data. For simplicity, a graphical method of projection was used to reveal “important”

PC, called Cattell's scree graph.

Factor analysis (FA) was carried out to determine the basic latent data structure using the following settings: the analysis was based on the correlation matrix and the obtained *factors* were rotated using a Varimax normalized algorithm, which allows for an easier interpretation of the factors loadings and maximization of the variance explained by the extracted factors.

Cluster analysis (CA) was considered as an exploratory data analysis technique for solving the classification problem. It comprises an unsupervised classification procedure that involves measuring either the distance (or the similarity) or the correlation coefficients between objects to be clustered. The information obtained from the measured variables is used to reveal the natural clusters existing between the studied samples. The similarities in this case were quantified through the Euclidean distance measurement. The CA was considered as a complementary method to PCA and FA.

Gnostic analysis (GA) – by the original mathematical background[139-140] it was first applied for pilot environmental studies within this thesis[85-86, 103-104], further expanded and completed into the practical guide[223]. The main goal of gnostic development was to make an expansion of alternative methodology for practical applications to small data sets. Nowadays this method is supported by the Web-based Open-Source system (R-project based on the R-language[224]).

The differences between statistical and gnostic approaches can be summarised in Table 2.

Problem/aspect	Approach	
	Statistics	**Gnostics**
Quantity of data to be treated	Mass data, large data samples	Individual data and small samples
A priory given statistical model of data	Required	Not used. Data assumed to satisfy algebraic requirements of measurement theory. Model to be applied results from data only.
Main theoretical tool	Additive measure over a sigma-algebra	Non-additive measure over two bi-algebras
Axiomatic	Formal, based an the Central Limit Theorem	Based on laws of Nature
Notion of probability	Formally defined	Derived mathematically from the Clausius' deterministic data entropy
Notion of information	Formally introduced	Derived mathematically from the Clausius' deterministic data entropy
Inherent geometry	Euclidean geometry	Riemannian geometry determined by data
Optimality criterions of estimation	Formal (e.g. least squares or max. likelihood)	Minimum information loss or entropy increase determined by variation principle
Variation features	Non-existent	Variation theorems for data errors, information and entropy
Bounds of a data sample	Ambiguous, dependent on a subjective decision	Uniquely and objectively determined by data
Robustness of an estimate	In classical statistics – not available In robust statistics achievable by means of an artificial superstructure over the basic theory	Resulting from the basic theory as its inherent and natural feature
Convergence of the two theories	Unknown	Proved for the case of high quality data
Ties with existing theories of Nature	Not defined	Proved close relations to classical thermodynamics, relativistic mechanics and both classical and robust statistics.

Table 2 General comparison of statistical and gnostic approach

F. Applications

Experimental work within this thesis falls into the following applications: fat reduction to assist analytical methods for determining POPs[225], sampling in a clean environment, revealing geogenic background (glacial water, mineral water), as well as identification of key pollutants in drinking water[80, 136, 212, 226-228], industrial applications focused on the assessment on contamination from runoff waster, flooding episode around hot-spots and assessment of waste water treatment plants[107, 135, 143, 213, 226, 229], and monitoring long-term studies, having a transnational effect for adoption on an international scale[230-232], revealing interactions chemicals and ecotoxicological effects. As an application within remediation, the identification of underground water sources was performed in concert with Integral Pumping Test[233]. Within all studies, data were subjected to appropriate data analysis: statistical (EDA, PCA, FA, CLU) and a gnostic one where censoring and interpretation of LOD values are addressed[234]. From all findings within applications, selected conclusions can be accepted.

1) TThe SPMDs as a *tool for fat reduction* - can be applicable for fat content higher than 1%(m/m), which corresponds to *feeding stuffs* (premixes), *food* (feeding oil, raw animal (pig) fat, butter, natural and dried eggs, sausages, salami, chocolate, whipped cream, cheese, beef, pork, fish, poultry, kidney, liver, dried milk), *human adipose tissue*. For those samples, recoveries of added isotopic standards can be expected within a 70-95% interval. This is in a good agreement with recently published results[120, 122].
2) TThe assessment of *anthropogenic background in Wieliczka* salt mine - the profile of PAHs and quantitative content reflect the geogenic background. Results of exposed SPMDs in parallel exhibit RSD within the interval 1 to 13%; this RSD includes analytical procedure uncertainty. Results show excellent repeatability. It is apparent that only a few of the monitored PAHs were detected. Mainly those having 3 aromatic rings; only one having 4-ring structure was identified (pyrene). No carcinogenic PAHs were detected, like benzo(a)pyrene. The concentration level is very low (for the cleanest profile F-II): phenanthrene (0.25 ng/l), anthracene (0.03 ng/l), fluoranthene (0,18 ng/l) and pyrene (0.18 ng/l). On the contrary, measurable concentrations of PCBs were found, with profiles slightly different to measured profiles of technical mixtures, frequently originating from recent use (Z-25 and FII). For each sample, OctaCB were the latest isomer group detected, having finger-prints feature for Delor 106 (or alternative, with domination of Hexa-CB) appear. Detectable concentrations were at the levels of 5- 80 pg/l of the sum of PCB congeners that respond to the high sensitivity of used SPMDs.

For the *Assessment of POPs in Poděbradka mineral waters* the level of contamination of groundwater is very low, if compared to similar studies and limit values for drinking water. The following table summarizes obtained data for PAHs, PCBs and selected OCPs. Table 3 shows approximate sums of main contaminant groups. These values are rough since the concentrations of a high number of assessed compounds were below the detection limit.

The most frequent contaminant of all analyzed substances was phenanthrene with the highest concentration of 13.9 $ng.l^{-1}$ in the Buda bore hole. Chemical parameters like PAHs, PCBs and OCPs concentration together with toxicity to *Desmodesmus subspicatus*, *Daphnia magna* and *Vibrio fischeri* were determined. Although very low level of contamination was observed, secondary contamination of PCBs through bedrock was observed. Used bioassays showed no dependence on varying concentrations of PAHs under 16 $ng.l^{-1}$. Changes in PCBs concentration from 22 to 168 $pg.l^{-1}$ induced homogenous changes in the toxicity response of all used organisms. HCH Concentration 470 $pg.l^{-1}$ together with the DDT concentration 26.7

pg.l^{-1} caused a significant response of bacterium, although crustaceans and algae showed no sensitivity to this level of pesticides.

pg/L	Buda	Plazy	Malý Vestec	Poděbrady
Sum of PAHs	16159	4380	2807	179
Sum of PCBs	95	168	72	22
Sum of HCH	220	140	220	470
HCB	7.9	8.3	8.3	2.2
Sum of DDT	4.2	4.1	3.9	26.7

Table 3 Poděbrady: concentrations of selected contaminants (pg.l^{-1})

3) The *Comparison of SPMDs and biotic sampling systems*,

The PCBs there were broadly used in 1959 to 1984 in former Czechoslovakia[15, 235-237]. Contamination of PCB[238-240] reveals contamination of 2,3,7,8-TCDD in the presence of mono- and non-ortho PCBs[53] that is the result of recent monitoring[15, 241-242]. The assessment of contamination of rivers, soils, suspended particulate matters, sediments and fluvial soils were published [15, 226, 241-244], however, water contamination is not too known.

The results showed that the river Elbe is contaminated primarily by Delor 103, whereas the river Morava by Delor 106. Total contamination of the river Elbe, as compared with the river Morava, is approx. 5 times higher. Concentrations of non-ortho and mono-ortho-chlorinated PCBs in *Dreissena polymorpha* are considerably lower when compared with the content in rivers, while the 2.2´-di-ortho-chlorinated PCBs are practically the same. Dominating accumulation of tri-ortho chlorinated biphenyls as well as the pronounced accumulation of 2.6-di-ortho- and tetra-ortho-chlorinated PCBs has been found. This is most probably caused by high thermodynamic stability and rigid structure of mentioned compounds (dihedral angel, aromatic rings perpendicular to each other, internal barriers of rotation). Fish have relatively low metabolism ability for high chlorine content of PCB congeners. Biotransformation appears to be limited to congeners with tri- and tetra-ortho chlorinated positions. Lesser chlorinated PCBs are more easily biotransformed. It is generally accepted that the bioconcentration of PCBs in aquatic organisms correlates with the degree of chlorination, the stereochemistry and also with lipophilicity.

Assessment of POPs in Rivers in the Czech Republic

Czech rivers have been monitored since 2003 by group of PS[1], and biota was initiated with 19 profiles; since 2007, there were new 2 profiles added. Configuration for sampling was realised as depicted above. Protective shrouds were fitted recently to floats used in biomonitoring studies.

Profiles for each group are presented. For statistical evaluation, non-normal distribution and outliers were evident. Therefore, average and median data are presented. The period 2003-2007

[1] SPMDs (for non-polar compounds) and DGTs (for metals), POCIS (for pharmaceuticals and pesticides)

is discussed. As for PCBs, the highest concentrations were found in the Bílina River in Ústí n/L. ($C_{average/median}$= 4155 / 3263 pg/l), the lowest one in Elbe-Debrné ($C_{average/median}$= 426 / 330 pg/l) that can be considered as a background value for monitoring.

As for probable sources, the elevated contamination by PCBs most probably comes from industrial sources, despite banned production of PCBs in 1980s in the former Czechoslovakia. Since beginning production, over 21kt were manufactured in Chemko Strážské and over 11kt was used inside Czechoslovakia[15, 235-237]. With respect to Delor 106 and Delor 103 recently used in technical mixtures, there is an apparent changing of ratio of isomeric patterns per year. Roughly, based on summary concentration of PCBs isomers and its development from 2003 to 2007, there is a rise in Tri- and Tetra-CB isomers (3-16%), whereas, Hexa-CB is decreasing proportionally (1-15%).

The contamination of particular profiles are discussed in detail , e.g. Elbe-Valy most probably originates from Pardubice-Semtín, with dominant content of Delor 103[15]. However, composition of PCBs varies up to Elbe-Obříství, beyond Spolana Neratovice. This means that Spolana Neratovice can also contribute to PCBs contamination by Delor 106, despite the recently found dominant contamination by OCPs and PCDD/Fs[15, 51, 245].

As for OCPs, major contamination was identified in Bílina-Ústí ($C_{average/median}$ = 3939 / 2991 pg/l) and Elbe-Obříství ($C_{average/median}$ = 3892 / 3231pg/l), whereas the lowest concentration (about three times) was found in Jizera-Předměřice ($C_{average/median}$= 663 pg/l).

Variations among years are apparent from the profiles as well as sampling places, making difficulties in formulating serious conclusions about exact causalities. The results from published data[244] show that there is a relative good correspondence with burdens of soil and sediments, found in the hundreds ng/g[44, 47, 49], or at biotic organisms[15, 241], even though the SPMDs exhibit only dissolved phase, without metabolic mass exchange. As for PCDD/Fs, the most remarkable profile is Odra-Bohumín ($C_{average/median}$ = 3428 / 2152 fg/l), which corresponds with the highest contamination of PAHs for this profile ($C_{average/median}$ = 223 / 185 ng/l) that is not surprising due to the chemistry production, coal industry and metallurgy - steelwork and WWTP effluents as well as landfills of Chemical plant Ostrava-Hrušov[242]. On the other hand, as seen from the list of measured isomers by SPMDs, the majority was below the LOD (number of observations with LOD within the range 84-100%), meaning that contamination of PCDD/Fs falls into the range of low environmental contamination, even after accumulation by SPMDs (only sequestering the dissolved phase). There are apparent exclusions for isomers measured above LODs, mainly TCDD, TCDF, PeCDF (number of observations with LOD within the range of 0-16%). From the temporal trends viewpoint, as evaluated from the most contaminated profiles, the overall trend is not uprising.

They differ from gnostic distributions not only in mean levels, but also in the variability, reflected in the shape of DF. Apparently, distributions are not normal (Gaussian), domains are finite, densities are asymmetric, mainly for PBDEs, PAHs, PCBs. Special attention is to be put to PBDEs. Due to the high variability of data, a robust (gnostic) approach was used in data evaluation. The shape of PBDE's DF can be roughly found as similar to PCDDs, reflecting their similar bioaccumulation property with respect to SPMDs, as well as similarity of molecular structures.

Similarly, statistical EDA has revealed non-normal distribution, densities asymmetric, parameters like skewness and kurtosis has revealed necessity data transformation. The Box-Cox transformation was performed on non-normal data, prior to the multivariate data analysis.

Then, PCA was performed, to explain the variance–covariance structure of a complex multivariate data set through a few linear combinations of the original variables (in this case, isomers PCB, selected congeners PCDD/F, ∑PAHs, ∑HCHs, ∑DDTs (consisting of DDTs, + metabolites DDDs and DDEs), selected PBDEs, and toxicological parameters).

The principal component analysis was performed as a pattern recognition method to compare the bioconcentration of mentioned parameters in SPMDs. There was major clustering ∑PCBs – ∑Tri-∑TetraCB, ∑PBDEs, ∑PAHs, together with ∑TCDF and ∑PeCDF. The second cluster is formed by higher chlorinated ∑PCBs (∑Penta-∑OctaCB), ∑DDTs (and their metabolites) and ∑HCHs.

Robustly estimated correlation (gnostic) coefficients[103, 139-140] were tested for statistical significance, giving similar results, however, without necessity of data transformations. It was shown that is useful in combination of methods, especially if data do not fits normal distribution and there are in small data sets. In such cases for correlation, the best results are expected if used robust correlation, ig. Spearman, Kendall or gnostic.

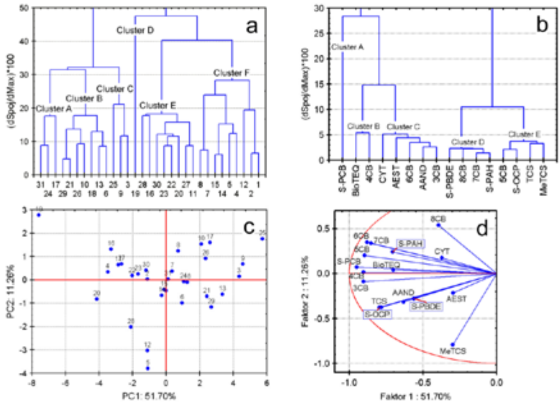
Identification of endocrine disrupting chemicals by SPMDs

Endocrine disrupting chemicals (EDCs) represent a variety of chemical classes, such as persistent organic compounds like PCBs, PAHs, OCPs, PBDEs, hormones, and other industrially formed compounds, like dioxins. Those compounds can cause many adverse effects in aquatic environment[246-250]; endocrine disruptors act via nuclear receptors, nonnuclear steroid hormone receptors, nonsteroid receptors, orphan receptors, enzymatic pathways involved in steroid biosynthesis and/or metabolism, and numerous other mechanisms that converge upon endocrine and reproductive systems[251-252].

The semipermeable membrane devices (SPMDs) of standard arrangement was used to identify non-polar organics: polycyclic aromatic hydrocarbons (PAHs), polychlorinated biphenyls (PCBs), organochlorinated pesticides (OCPs), polybrominated difenyl ethers (PBDEs), triclosane (TCS) and methyltriclosane (MeTCS) within the framework of assessment of endocrine disrupting chemicals (EDCs) and bioassays in parallel in vitro responses to reveal cytotoxicity (Cyt 1/IC50), antiestrogenicity (AES 1/IC50), antiandrogenicity (AAND 1/IC50) and AhR-mediated activity (Bio-TEQ)[253].

The study focused on the impact of Brno agglomeration and municipal waste water treatment plant effluent on upper (lower) streams of Svitava and Svratka rivers.

After EDA diagnosis, indicating some outliers at high values and asymmetric, skewed distribution, based on combined sample skewness and kurtosis test $21.13 > \chi^2(0.95, 2) = 5.992$, the normality of data distribution was rejected and Box-Cox transformation was performed. Then Cluster analysis CLU was performed. The dendrogram of water samples resulting from CLU (Euclidean distance and Ward´s agglomerative method) of Box-Cox normalized data set is represented, together with PCA, on the Graph 1a-d.



Graph 1 EDCs in Brno: dendrograms and PCA of sampling sites and parameters

Cluster B being close to cluster C contains BioTEQ and 4CB, while cluster C contains CYT, AEST, 6CB, AAND and 3CB. Another two clusters D and E seem to be similar but cluster D contains ∑PBDE, 8CB, 7CB and ∑PAH while cluster E contains 5CB, ∑OCP, TCS and MeTCS. This outlying sample is attributed to MeTCS occurrence that can be accounted to biodegradation product of TCS due to biological process of WWTP. As for toxicity responses, cytotoxicity was higher for samples from locations downstream from Brno city. Treatment processes in the WWTP decreased the cytotoxicity of the waste waters. SPMD samples elicited significant AhR-mediated activity. The antiestrogenic activity appeared at most sampling sites. Results were confirmed by FA, with Varimax rotation of the axes defined by PCs (FA) explains again the 63% of the variability of the Box-Cox normalized data but it partially modifies the weight of the normalized variables in three varifactors. From the point of view biotests, there is apparent AhR-mediated activity (dioxin-like activity) that corresponds to low chlorinated PCBs.

Aspects of analysis of multivariate data

Besides common statistical approaches, gnostic approach was also used. Conclusions from the applications are especially relevant for POPs and their levels, as found above using SPMDs. It is noteworthy that data transformation (e.g. Box Cox), despite its rigorousness, tends data to interdependence, as was found in practical applications within this thesis.

Due to some consequences that are relevant for small data samples[254], mainly: (i) statistical tests have relatively low power, (ii) probability of selection of most appropriate parametric distribution will be low, (iii) use of empirical distribution may be problematic, (iv) parameter estimates will have high statistical error; approximation of the sampling distribution may be relatively poor, the use of both non-robust statistical correlation (Pearson), robust statistical correlation (Spearman, Kendal) and robust gnostic seems to be practical, for further PCA application.

The *non-parametric* methods were used from the available correlations. An overview of used methods is given in the Table 4.

ID	Approach description	Method source
SP	Common statistics	Pearson
SK	Robust statistics	Kendall
SS	Robust statistics	Spearman
GN	Mathematical gnostic	Kovanic

Table 4 Czech Rivers: overview of used approaches and methods of correlation

Using these methods, various correlation coefficients were estimated for given groups of contaminants. Evaluation of correlations are given on the Table 5. The importance of used correlation approaches is apparent from the results.

Method	SP	SK	SS	GN
SP	0	0.136	0.147	0.161
SK	0.136	0	0.084	0.087
SS	0.147	0.084	0	0.047
GN	0.161	0.087	0.047	0

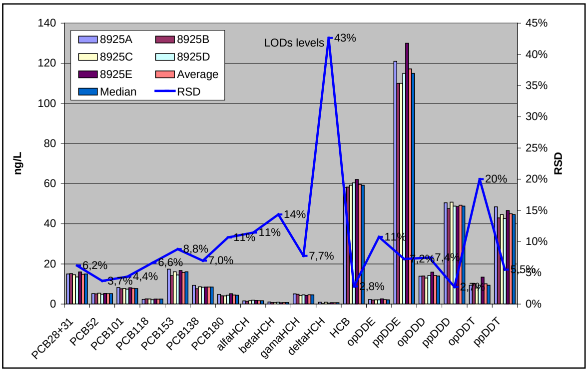
Table 5 Czech Rivers: comparison of various correlation coefficients and methods

From this overview and the results compiled above, the smallest difference between vectors is between Spearman's and gnostic correlations, which proves the best fit between those methods. Both methods are the most suitable for further evaluation of the relationship of data. However, Kendal's method seems to be acceptable too, but an important feature should be taken into account: gnostic correlation coefficient is estimated by the gnostic distribution functions, which enable usage of censored data. The classical Pearson's correlation, despite its wide use, is less robust.

Specific conclusions were made for regression models resulting from the use of gnostic regressions. All methods were realised as Weighted Least Squares. The best results was reached by the method ("gnostic") with R=0.986 and STD = 0.0044, together with ("fair") with R = 0.979 and STD = 0.0063. Another appropriate method was ("Andrews") with R = 0.704 and STD = 0.0063. In all results, there is an apparent opposite effect of PCBs against PAHs. This fact is explainable by grouping effect and the loss of information distinctive for both PCBs and PAHs. However, regression methodology plays an important complementary methodology, especially when a robust method is used. The use of the most robust methods is advisable, even if data are of a low scale and/or normality criteria are not met.

Evaluation of repeatability

Within the framework of all studies realised recently as well as within this thesis, there was much data taken for QA/QC aspect. On the graph below there are results from parallel exposure of PCBs, OCPs, as one example from the series.



Graph 2 QA/QC: parallel results of individual PCBs and OCPs, surface water, N=5
PCBs - within indicator congeners

Detailed results can be found among source data, as supplied in digital form only to this thesis. As from validation protocols, uncertainty was set up within 20-30%. From all data there is an apparent RSD within 5-15% involving both sampling and analytical uncertainty. For exceptional cases, where values are approaching LODs, the RSD is approx. within 20-40%. Similar results were found for toxicity testing, within RSD 5-25%.

All aforementioned results underline the suitability of SPMDs for various applications as a tool with powerful reproducibility within broad concentration scales.