EnviroInfo 2011: Innovations in Sharing Environmental Observations and Information
Copyright 2011 Shaker Verlag Aachen, ISBN: 978-3-8440-0451-9

Similarities of Environmental Health Data of Persistent Organic Pollutants in three Countries Analyzed by the PyHasse Software

Kristina Voigt [1], Rainer Bruggemann [2] Hagen Scherb[1], ısmet Cok[3] Birgül Mazmnci[4], Ali Mazmanci[4], Cafer Turgut[5], Karl-Werner Schramm[67]

Abstract

Chemicals, such as Persistent Organic Pollutants (POPs), introduced into the environment by human activity may disrupt the endocrine system of animals including fish and wildlife as well as humans and produce adverse effects because of the crucial role hormones play in controlling development. POPs are detected worldwide. In order to evaluate the data on numerous studies on POPs sound mathematical and statistical data evaluation methods are needed. The data analysis method demonstrated in this paper, is based on the theory of partially ordered sets and provides a generalized ranking. Partial order is a discipline of Discrete Mathematics and one may consider partial order as an example of mathematics without arithmetic. The data analysis is performed with the free available software package PyHasse, written by the second author, which provides apart from the calculation of Hasse diagrams many features, such as for example the similarity analysis applied in this paper.

Studies on POPs were performed in Denmark (1997-2001) and Finland (1997-1999) as well as in Turkey in 2010. In our data analysis approach we investigated data sets of breast milk samples of women in Denmark and Finland which contained detectable levels of 20 Persistent Organic Pollutants (POPs). These results have already been published by Voigt et al, 2010. In a study performed in the Taurus Mountains area in Turkey the same 20 POPs were detected in breast milk samples. The question arises whether the above mentioned methodology of partial orders can find differences or similarities among these countries.

Applying the sub-routine Similarity of the PyHasse software the similarities between data sets can be identified. The combinations are for the similarity analyses: Turkey with Denmark, Turkey with Finland, and Denmark with Finland.

In the similarity analysis different types of relations are distinguished which quantify the similarities of the two compared data sets. The highest degree of similarity can be found comparing Denmark with Finland. However, there is also some similarity regarding Turkey – Denmark and Turkey – Finland. This means that the breast milk samples in all three countries are similarly contaminated with respect to their quality the 20 POPs looked upon.

[1] Helmholtz Zentrum Muenchen, German Research Center for Environmental Health,
Institute of Biomathematics and Biometry, Ingolstaedter Landstr.1, 85764 Neuherberg, Germany, kvoigt@helmholtz-muenchen.de
[2] Leibniz-Institute of Fresh Water Ecology and Inland Fisheries, Berlin, Germany,
[3] Department of Toxicology, Faculty of Pharmacy, Gazi University, 06330, Ankara, Turkey
4 Department of Biology, Faculty of Sciences and Letters, University of Mersin, Turkey
[5] Faculty of Agriculture, Adnan Menderes University, 09100 Aydin, Turkey
[6] Helmholtz Zentrum München, German Research Center for Environmental Health, Institute of Ecological Chemistry, Ingolstädter Landstrasse 1, 85764 Neuherberg, Germany
[7] TUM, Wissenschaftszentrum Weihenstephan fuer Ernaehrung und Landnutzung, Department fuer Biowissenschaften, Weihenstephaner Steig 23, 85350 Freising, Germany

1. Introduction

The most intractable pollutants are nuclear wastes, hazardous wastes, and wastes that threaten global biogeochemical processes, such as greenhouse gases. They are chemically the hardest to sequester or detoxify, physiologically the hardest for our senses to detect, and economically and politically the most difficult to regulate [Meadows et al., 2004, p. 112]. In a recently published paper regarding a safe operating space for humanity, nine processes are stated for which it is absolutely necessary to define boundaries. One of these processes is chemical pollution for which the boundaries have – unfortunately - yet to be determined [Rockström et al., 2009]. We focus in this paper on human-synthesized chemicals, called Persistent Organic Pollutants (POPs). They have never before existed on the planet, and therefore no organisms have evolved in nature to break them down and render them harmless.

Environmental and human exposure to Persistent Organic Pollutants (POPs) has been the subject of scientific investigation and political regulation for almost 40 years. The signature of the Stockholm Convention on Persistent Organic Pollutants in May 2001, and its amendment in 2010 [UNEP, 2011] is one of the most obvious and important results of the scientific and political discussion on global pollutants. The Stockholm Convention on Persistent Organic Pollutants is a global treaty to protect human health and the environment from chemicals that remain intact in the environment for long periods, become widely distributed geographically, accumulate in the fatty tissue of humans and wildlife, and have adverse effects to human health or to the environment. Exposure to POPs can lead to serious health effects including certain cancers, birth defects (like cryptorchidism), dysfunctional immune and reproductive systems, greater susceptibility to disease and even diminished intelligence. Within the last few years, questions and concerns have also focused on the hypothesis that EDC (Endocrine Disrupting Chemicals) may be involved in the dramatic rise in the incidence of metabolic disorders such as obesity and diabetes observed worldwide the last 40 years [Lyche et al., 2011].Given their long range transport, no one government acting alone can protect is citizens or its environment from POPs.

The evaluation of the data is a major step in the elucidation of the danger of these chemicals. Currently 22 POPs are listed in the convention mentioned above. 20 of these chemicals are in the focus of our data evaluation approach.

2. Data Analysis Method: Hasse Diagram Technique

2.1 Basic Principles

The data analysis method demonstrated here, is based on the theory of partially ordered sets and provides a generalized ranking. Partial order is a discipline of Discrete Mathematics and will only be briefly explained in this paper in order to follow the data-analysis steps. A good overview can be found in Bruggemann and Carlsen [2006] and Bruggemann and Patil [2010].

First step: We need a set of objects. We call this set of objects the ground set, and denote it as G. Second step: We need an operation between any two objects. As an evaluation is our aim, we must compare the objects. Is object "a" better than object "b"? If objects a and b are comparable we write $a \perp b$, albeit this in general important symbol is not needed in our actual analysis.

Third step: We do not only want that two objects are comparable, but we also would like to know the orientation: Is "a" better or worse than "b"? Therefore the signs ≤ and ≥ are introduced: $a \leq b$ "may" denote that a is better than b, $a \geq b$ "may" indicate that a is worse b.

Fourth step: Why "may"? The essential point is that we have to define, when we will consider object a as better than b, i.e. the signs "≤" and "≥" alone do not help in an evaluation procedure, we must give them an appropriate sense, i.e. an appropriate contextual orientation. If objects are not comparable this is indi-

cated by the symbol ||. A subset G' ⊆ G in which for all object pairs (x,y) ∈ G'^2 x || y holds, is called an antichain.
Fifth step: Independently how we define ≤ and ≥ the ground set equipped with e.g. "≤" must obey three axioms, if we want to speak from a partially ordered set (poset):
Reflexivity: An object a can be compared with itself: a ≤ a
Antisymmetry: If a is better b, and at the same time b is better than a, then a = b. We write:
a ≤ b and b ≤ a ⇒ a ≅ b and consider a and b as equivalent.
Transitivity: If a is better than b and at the same time b is better than c, then a is better than c.
a ≤ b, b ≤ c ⇒ a ≤ c.
If the ≤-relation is defined properly; the ground set G equipped with ≤ is a partially ordered set. A widespread notation is: (G, ≤).

The transitivity axiom is a very important one and filters out many situations, which could also be considered as a matter of evaluation: For example: competitions in sports: Crack "a" wins about crack "b", in turn crack "b" wins about "c", but unexpectedly crack "a" does not win about "c"! The mathematical analysis of this kind of generalized comparisons is done in tournament theory.
Sixth step: Why can a partially ordered set be represented by a directed graph? Consider the objects of a ground set as vertices. Then in any case, where for (a,b) ∈ G^2 is valid a ≤ b we draw an arrow starting from b and ending in a. Because of the transitivity we can omit such arrows, which are represented by a sequence of ≤ - relations. Hence, most of the advices how to construct a Hasse diagram are using the concept of a "cover relation" $\prec$: Two objects a,b for which is valid that a ≤ b are in a cover relation, if there is no third element in between. Then a Hasse diagram is a graph of cover relations with additional conventions how to locate the objects in the drawing plane.

2.2 PyHasse Software

The data analysis is performed by the free available software package PyHasse, written by the second author [Bruggemann and Patil, 2010], which provides apart from the calculation of Hasse diagrams many features, such as for example the similarity analysis applied in this paper.

PyHasse is written in the modern interpreter language Python (vs 2.7). Python is used as 'rapid prototyping' programming language, is considered as a'Very High Level Language' (VHLL) and allows a high level of abstraction [Müller and Schwarzer, 2007]. PHYTHON can freely be downloaded from http://www.python.org/download, is platform independent and it runs on many operating systems. It also has access to many packages either in algebra, in statistics or in graphics or even in preparing sound effects, game developments and image processing, allows object-oriented programming and is itself modularized to a very high degree (http://www.pythonology.com/home). In contrast to the older but professionally programmed WHASSE, PyHasse is considered "experimental" software, which should bridge the gap between professional software and software, exclusively developed in laboratories and in general only applicable for the programming scientist.

2.3 PyHasse Module: Similarity

In complex data-sets it is often necessary to compare different sets of criteria which are quantified by the attributes or (synonym) the indicators. In the similarity analysis we intend to calculate the similarity of different posets (partially ordered sets) based on the same ground set. This similarity analysis is an important feature of the newly developed software package called PyHasse.

The outcome of a partial order for two objects a, b may be $a < b$, $a > b$, $a \parallel b$, $a \cong b$. When two partial orders $(G, \leq_1)$ and $(G, \leq_2)$ are to be compared, then the combinations are count such as: $a <_1 b$, $a <_2 b$, or $a <_1 b$, $a >_2 b$, or $a \|_1 b$, $a <_2 b$ for which we use the shorthand notation >> , >< , || etc.
Most important are the entries like >> or <<, which are counting the 'isotone' character of both partial orders (ISO in Table 2) and the entries like ><, <> which contribute to the 'antitone' character, i.e. to the conflicts between the two partial orders (ANTI in Table 2). Following Rademaker et al. [2008]: two posets are in conflict or "contradict each other" (are antitone) on two objects $x, y \in P$,"if we have $x <_1 y$ and $y <_2 x$ or $y <_1 x$ and $x <_2 y$". There are still more combinations to look upon: <||, >||, ||>,||< , =||, ||= or || || are considered as indifferent (IND in Table 2), combinations like > =, < =, = <, = > are called weak isotone (WISO in Table 2). Finally the entry of type = = contributes to equivalence relations (IDE in Table 2).
A recent example of the application of the similarity on an environmental health data set is given by Voigt et al. [2010].

3. Chemicals to be analyzed

Studies on POPs were performed in Denmark (1997-2001) and Finland (1997-1999) as well as in Turkey in 2010.

A possible human association between maternal exposure to organochlorine compounds used as pesticides and cryptorchidism among male children has been investigated in Denmark and Finland. These investigations have shown striking differences in semen quality and testicular cancer rate between Denmark and Finland. Since malformation of the testis is a shared risk factor for these conditions, a joint prospective study for the prevalence of congenital cryptorchidism was executed in Denmark (1997-2001) and Finland (1997-1999) [Boisen et al., 2004], [Damgaard et al., 2006], [Shen et al., 2010]. The contamination situation on the occurrence and effects of Persistent Organic Pollutants (POPs) in the Taurus Mountains in Turkey was studied recently [Turgut et al., 2011]. Taurus Mountains were suggested for this study because of their potential to act as a sink for organic pollutants; hence this region can reflect the atmospheric pollution in Turkey as well as neighboring countries e.g Arabia, Africa and Russia.

Here also breast milk samples were analyzed containing measurable levels of POPs [Schramm, personal communication, 2011].

The chemicals found and which we want to analyze are listed in Table 1 together with their used acronym and their CAS-Number. In our data analysis approach we investigated data sets of breast milk samples of women in Denmark and Finland which contained measurable levels of 20 Persistent Organic Pollutants (POPs). These results have already been published by Voigt et al., [2010]. In the study performed in the Taurus Mountains area in Turkey the same 20 POPs were detected in breast milk samples. The 20 chemicals are listed together with their acronyms used in the further evaluation and their CAS-No in Table 1.

Table 1:
20 Persistent organic Pollutants in Breast Milk Samples in Denmark, Finland and Turkey

Nr.	Acronym	Name	CAS-Number
01	AHCH	alpha-Hexachlorcyclohexane	319-84-6
02	BHCH	beta-Hexachlorcyclohexane	319-85-7
03	GHCH	gamma-Hexachlorcyclohexane	58-89-9
04	DHCH	delta-Hexachlorcyclohexane	319-86-8
05	PECB	Pentachlorobenzene	608-93-5
06	HCBE	Hexachlorobenzene	118-74-1
07	PPDT	p, p'-Dichlordiphenyltrichlorethane	50-29-3

08	OPDT	o, p’-Dichlordiphenyltrichlorethane	789-02-6
09	PPDD	p, p’-Dichlordiphenyldichlorethane	72-54-8
10	OPDD	o, p’-Dichlordiphenyldichlorethane	53-19-0
11	PPDE	p, p’-Dichlordiphenyldichlorethene	72-55-9
12	OPDE	o, p’-Dichlordiphenyldichlorethene	3424-82-6
13	TCHL	trans-Chlordane	5103-74-2
14	CCHL	cis-Chlordane	5103-71-9
15	OXYC	Oxychlordane	27304-13-8
16	CHCE	cis-Heptachloroepoxide	1024-57-3
17	DIEL	Dieldrin	60-57-1
18	END1	Endosulfan-1	959-98-8
19	MECH	Methoxychlor	72-43-5
20	MIRE	Mirex	2385-85-5

The question arises whether the above mentioned methodology of partial orders can find differences or similarities among these countries.

We will now evaluate the following data sets:

20 chemicals x 65 breast milk samples in Denmark (DK)
20 chemicals x 65 breast milk samples in Finland (FL)
20 chemicals x 44 breast milk samples in Turkey (TK).

4. Application of Partial Order Technique (Hasse Diagram Technique)

4.1 Basic Hasse Diagram Technique

In the first evaluation step we perform the data analyses on the three data sets Denmark, Finland and Turkey. The results of the corresponding 3 Hasse diagrams are given in Figure 1.

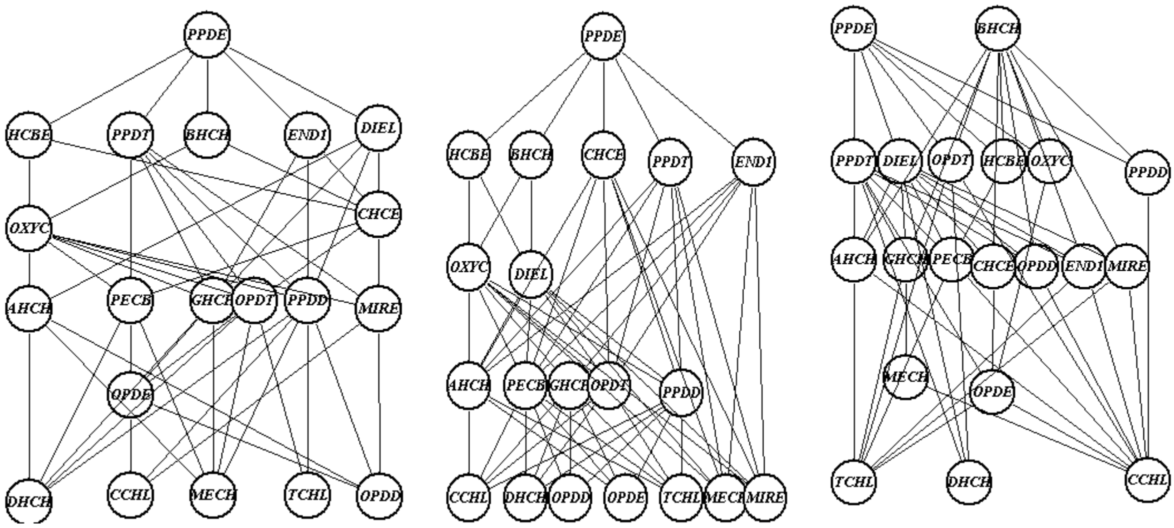


Figure 1: Hasse diagrams for 3 countries: Denmark (left), Finland (middle), Turkey (right)

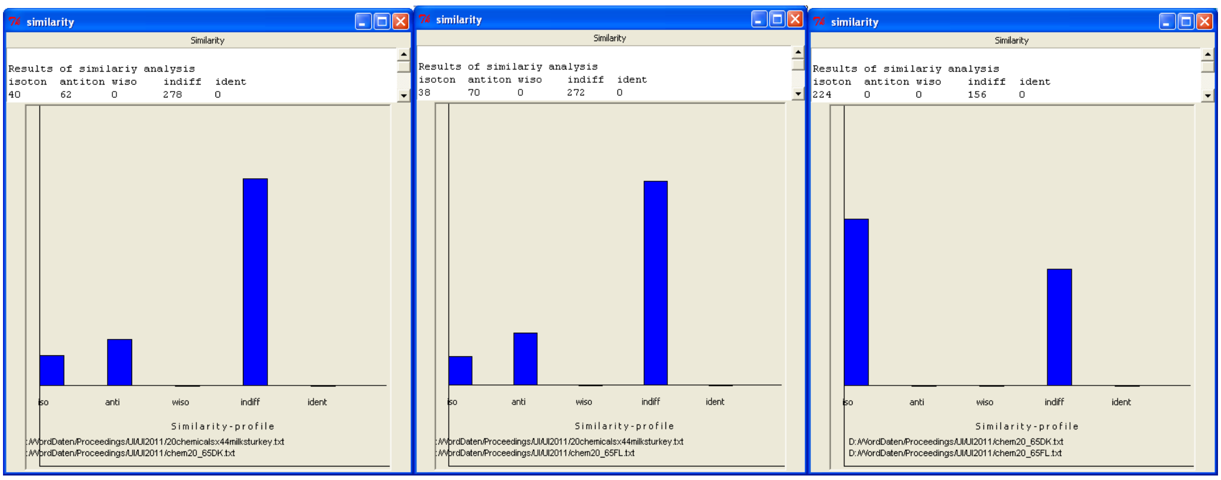
It can clearly be seen that that the chemical PPDE (DDE) is always a maximal object. This means that the chemical is the worst in comparison to all other objects (chemicals) with which it is connected in the downward position. In the Hasse diagram representing the Turkish breast milk sample results (rhs) the chemical BHCH (beta-Hexachlorocyclohexane) is also a maximal object. This chemical is found in the second highest level in the Danish and Finnish Hasse diagrams. The number of minimal objects (those chemicals which show the least pollution in comparison to those objects with which they are connected in the upward position) varies in the three Hasse diagrams. The chemicals CCHL (cis-Chlordane), DHCD (delta-Hexachlorocyclohexane), and TCHL (trans-Chlordane) are always in the minimal position. The Danish diagram (lhs) shows 5 minimal objects; the Finnish (middle position) has 6 objects and the Turkish (rhs) only the above mentioned 3 chemicals.

This demonstrates that the three diagrams have on the one side a lot of similarities in the maximal and minimal positions and differences on the other side. Some of these differences can easily be detected from the diagrams, e.g. the positions of the chemicals CHCE and MECD. Although the different positions and <-relations for single chemicals in the three partial orders are of high interest, we want to find a measure for the similarity or dissimilarity among the three partial orders, whose diagrams are shown in Figure 1.
It is now of importance to calculate the similarities respectively differences of the data sets. A PyHasse module was established by the second author Dr. Rainer Bruggemann which quantifies the similarities and differences between any two data sets respectively any two partial orders.

4.2 PyHasse Module: Similarity

As mentioned above, the module Simi5.py was established to check two data sets for their similarities. For further reading we recommend the recent publications of Bruggemann and Patil [2010], Voigt et al. [2010] and Voigt et al. [2011].

The combinations are for the similarity analyses are:
Turkey with Denmark, Turkey with Finland, Denmark with Finland



Turkey / Denmark   Turkey / Finland   Denmark / Finland

Figure 2: Similarity Analyses of the 3 countries: Turkey, Finland, Denmark

In Figure 2 the bar diagrams of the 3 different similarity analyses are given. The results of the different relations are listed in Table 2.

Table 2:
20 Persistent Organic Pollutants in Breast Milk Samples in Denmark, Finland and Turkey

Data sets / Relations	ISO	ANTI	WISO	IND	IDE
TK/DK	40	62	0	278	0
TK/FL	38	70	0	272	0
DK/FL	224	0	0	156	0

Isotone relations demonstrate a high degree of similarity. Comparing the Danish with the Finnish data they are the most similar data sets. Antitone relations which are found in the comparisons of Turkey with Denmark and with Finland indicate no similarity with respect to some chemicals and may be caused by different uptake mechanisms. The comparison of the Danish with the Finish data does not give a single antitone relation.

The results of the calculation show that Danish and Finnish data sets are similar to each other to a higher degree than the similarities of Turkey to Denmark and Turkey to Finland, resp..

5. Summary and Outlook

The 20 chemicals which we focus in our study are detected in breast milk samples in different countries and regions. The outcome of this simple and preliminary study already reveals that chemicals having high concentration levels in the breast milk samples (i.e. chemicals in the Hasse diagram at or near the top) are nearly the same in all three different countries. This gives a strong indication that we find these Persistent Organic Pollutants in human beings worldwide. Hence our findings corroborate the impairment not only of our environment but also of the human beings with POPs. However we also find differences especially in comparison to data of Turkey, which should be more closely analyzed in order to clarify potential different uptake mechanisms.

It is an urgent need to first insist on the exact compliance with the outlaw of these chemicals and furthermore have more and faster development of sustainable chemicals. Environmetrical and chemometrical methods like the applied partial order technique and the developed PyHasse program package can be applied for finding out more sustainable chemicals. We conclude that the data analysis method presented here can well be applied for distinguishing more or less dangerous chemicals. In addition, it should be used in sustainable chemistry in the same manner for detecting more and less sustainable chemicals.

Bibliography

Boisen K.A., Kaleva M., Main K.M., Virtanen H.E., Haavisto A.M., Schmidt I.M., Chellakooty M., Damgaard I.N., Mau C., Reunanen M., Skakkebaek N.E., Toppari J. (2004) Difference in prevalence of congenital cryptorchidism in infants between two Nordic countries. Lancet 363 (9417):1264-1269.

Bruggemann R., Carlsen L. (2006) Partial Order in Environmental Sciences and Chemistry. Springer, Berlin.

Bruggemann R., Patil G.P. (2010) Multicriteria prioritization and partial order in environmental sciences. Environ Ecol Stat 17:383-410.

Damgaard I.N., Skakkebaek N.E., Toppari J., Virtanen H.E., Shen H.Q., Schramm K.-W., Petersen J.H., Jensen T.K., Main K.M., Grp NCS (2006) Persistent pesticides in human breast milk and cryptorchidism. Environmental Health Perspectives 114 (7):1133-1138.

Lyche J.L., Nourizadeh-Lillabadi R., Karlsson C., Stavik B., Berg V., Skare J.U., Alestrom P., Ropstad E. (2011) Natural mixtures of POPs affected body weight gain and induced transcription of genes involved in weight regulation and insulin signaling. Aquat Toxicol 102 (3-4):197-204.

Meadows D., Randers J., Meadows D. (2004) Limits to Growth, The 30-Year Update. Chelsea Green Publishing Company, White River Junction, Vermont, USA.

Muller M., Schwarzer S. (eds) (2007) Python im deutschsprachigen Raum. Tagungsband zum Workshop am 28. September 2006 in Leipzig. Lehmanns Media, Berlin.

Rademaker M., De Baets B., De Meyer H. (2008) New Operations for Informative Combination of Two Partial Order Relations with Illustrations on Pollution Data. Combinatorial Chemistry & High Throughput Screening 11 (9):745-755.

Rockstrom J., Steffen W., Noone K., Persson A., Chapin F.S., 3rd, Lambin E.F., Lenton T.M., Scheffer M., Folke C., Schellnhuber H.J., Nykvist B., de Wit C.A., Hughes T., van der Leeuw S., Rodhe H., Sorlin S., Snyder P.K., Costanza R., Svedin U., Falkenmark M., Karlberg L., Corell R.W., Fabry V.J., Hansen J., Walker B., Liverman D., Richardson K., Crutzen P., Foley J.A. (2009) A safe operating space for humanity. Nature 461 (7263):472-475.

Schramm K.- W. (2011) personal communication on breast milk samples.

Shen H., Main K.M., Andersson A.M., Damgaard I.N., Virtanen H.E., Skakkebaek N.E., Toppari J., Schramm K.-W. (2008) Concentrations of persistent organochlorine compounds in human milk and placenta are higher in Denmark than in Finland. Hum Reprod 23 (1):201-210.

Turgut, C.., Atatanir L., Mazmanci B., Mazmanci M.A., Henkelmann B., Schramm K.W., The occurrence and environmental effect of Persistent Organic Pollutants (POP) in Taurus Mountains soils, submitted to Environmental Science and Pollution Research, accepted, 2011. DOI 10.1007/s11356-011-0561-x.

UNEP (2011) Stockholm Concention on persistent organic pollutants (POPs), http://chem.pops.int/Convention. UNEP, Stockholm Convention Secretariat. http://chem.pops.int/Convention. Accessed 11.05.2011.

Voigt K., Bruggemann R., Scherb H., Shen H.Q., Schramm K.-W. (2010) Evaluating the relationship between chemical exposure and cryptorchidism. Environmental Modelling & Software 25 (12):1801-1812.

Voigt K., Scherb H., Bruggemann R., Schramm K.-W. (2011) Application of the PyHasse Program Features: Sensitivity, Similarity, and Separability for environmental health data. Statistica and Applicazione accepted.