

Voigt K., Bruggemann, R., Scherb H., Cok I, Schramm K.W.,
2014. Modelling Similarities between the Contamination with
Endocrine Disruptors of Human Breast Milk Samples and Pine
Needles, in: Ames D.P., Quinn N.W.T., Rizzoli A.E., (Eds.)
Proceedings of the 7th International Congress on
Environmental Modelling and Software, June 15-19, 2014, San
Diego, California, USA. ISBN: 978-88-9035-744-2.



Bold Visions for Environmental Modelling

San Diego, California June 15-19, 2014

[DRAFT] Proceedings of the 7th International Congress on Environmental Modelling and Software (iEMSs) June 15-19, 2014, San Diego, California, USA

How to cite the full proceedings:

Ames, D.P., Quinn, N.W.T., Rizzoli, A.E. (Eds.), 2014. Proceedings of the 7th International Congress on Environmental Modelling and Software, June 15-19, San Diego, California, USA. ISBN: 978-88-9035-744-2

How to cite an individual paper:

Smith, J., Young, B., Taylor, J., 2014. This is the title of our paper. In: Ames, D.P., Quinn, N.W.T., Rizzoli, A.E. (Eds.), Proceedings of the 7th International Congress on Environmental Modelling and Software, June 15-19, San Diego, California, USA. ISBN: 978-88-9035-744-2

Peer Review:

Each paper has been peer reviewed by at least two independent reviewers with possible outcomes of reject, revise, and accept.

Creative Commons License:



The Proceedings of the 7th International Congress on Environmental Modelling and Software by Ames, D.P., Quinn, N.W.T., Rizzoli, A.E. (Eds.) is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/). This is a [Free Culture License](https://creativecommons.org/licenses/by/4.0/).



Modelling Similarities of Endocrine Disruptors in Pine Needles and Human Breast Milk

Kristina Voigt¹, Rainer Bruggemann², Hagen Scherb¹, Ismet Cok³, Karl-Werner Schramm^{4,5}

Helmholtz Zentrum München – German Research Center for Environmental Health

Institute of Computational Biology

Ingolstädter Landstr. 1, D-85764 Neuherberg, Germany, kvoigt@helmholtz-muenchen.de,

scherb@helmholtz-muenchen.de

²*Leibniz-Institute of Fresh Water Ecology and Inland Fisheries, Berlin, Germany brg_home@web.de*

³*Department of Toxicology, Faculty of Pharmacy, Gazi University, 06330, Ankara, Turkey,*

icetin@gazi.edu.tr

⁴*Helmholtz Zentrum München, German Research Center for Environmental Health, Molecular*

EXposomics (MEX), Ingolstädter Landstrasse 1, 85764 Neuherberg, Germany, [\[muenchen.de\]\(mailto:muenchen.de\)*](mailto:schramm@helmholtz-</i></p></div><div data-bbox=)*

⁵*TUM, Wissenschaftszentrum Weihenstephan fuer Ernährung, Landnutzung und Umwelt,*

Department fuer Biowissenschaften, Weihenstephaner Steig 23, 85350 Freising, Germany

Abstract: In a recently performed monitoring project, 18 OCPs (Organochlorine pesticides) in paired samples of pine needles and human breast milk samples were analysed in different regions in the Taurus Mountains in Turkey. The aim of our data evaluation approach is to find out whether there are conformities between the paired human breast milk and pine needle samples. An appropriate data analysis method to identify such conformities and differences is the discrete mathematical method named Hasse diagram technique. The software package used is the PyHasse software. It comprises several modules, which are of great support in the evaluation of environmental data. In this presentation we applied the main Hasse Diagram Technique Module (mHDCI2_7), the Similarity Analysis, and the CombiSimilarity7 for the comparison of two data matrices. The new module CombiSimilarity resulted in evidence that needles' and breast milk pollution are similar if taken here from related locations. This means that although we find visible differences in the Hasse diagrams of breast milk samples and needles, the calculations by the new CombiSimilarity tool reveal conformity of the two data sets.

Keywords: Organochlorine Pesticides; Human Breast Milk; Pine Needles; Hasse Diagram Technique; PyHasse software.

1. INTRODUCTION

Organochlorine pesticides (OCPs) have been reported in different environmental and human media worldwide for ages. The ubiquitous distribution of persistent organic pollutants (POPs) in air, including results of air monitoring by passive air sampling in five continents was performed and published. The authors demonstrated that these chemicals were found everywhere and demanded action to reduce environmental contamination by, and human exposure to, hazardous chemicals (Bogdal et al., 2013). Most of these chemicals have endocrine potential. Endocrine disruptors are exogenous agents that interfere with the production, release, transport, metabolism, binding, action or elimination of the natural hormones in the body responsible for the maintenance of homeostasis and the regulation of developmental processes (Birnbaum, 2013). A large number of recently published studies support the association between chemical exposure and obesity as well as the risk of type 2 diabetes (Mannetje et al., 2012, Thayer et al., 2012). The public health implications of such an association could be severe, and could be a heavy loading of national economies. A great effort would be necessary to reduce the external and internal exposure to OCPs to decrease the social burden of such impacts (Lee, 2012).

The occurrence of OCPs in the Taurus Mountains in Turkey motivated to an international study on these chemicals in environmental and human media. Whereas soil, pine needles and air samples from the same location should allow comparisons anyhow, human samples might be slightly different due to their mobility. Thus the study design targeted only humans which are living and feeding mostly from their habitat for a long time and had been less mobile than individuals in cities or other areas.

The contamination of soils has already been evaluated by Turgut et al. (2012). The occurrence of OCPs in breast milk samples was the objective of a recently published study (Voigt et al., 2012). Furthermore, we analysed the relationship between the contamination of soil and breast milk samples in the region (Bruggemann et al., 2014). The current study should outline the translational power between locally paired needle and human breast milk samples for organochlorine pesticides employing PyHasse software.

2. DATA ANALYSIS METHOD AND SOFTWARE PYHASSE

2.1 Partial Order Technique and PyHasse Software

The theory of partial order ranking is presented as an overview in many publications. We only mention two recent ones here (Bruggemann & Patil, 2010, 2011).

The analysis of partial order relations as a discipline of discrete mathematics can only be briefly explained in order to follow the data analysis steps. A detailed description of the procedure can be found in Voigt et al., (2013).

The partial order technique, also known as Hasse diagram technique is a ranking method. It is frequently applied when objects should be evaluated by a variety of criteria also called attributes. A partially ordered set or poset can be presented in a graphical form, called a Hasse diagram. The main frame of the Hasse diagram technique is the so-called four point program:

1. Selecting a set of objects of interest which are to be compared, E . The so-called ground set.
2. Selecting a set of attributes, by which the comparison is performed, called the information base IB .
3. Find a common orientation for all attributes.
4. Analysing $x, y \in E$ where one of the following relations is valid: Equivalence, comparability, incomparability.

Very often it is of interest to identify the maximal, minimal, and isolated objects:

- Maximal objects have no upper neighbour in the Hasse diagram.
- Minimal objects have no lower neighbour in the Hasse diagram.
- Isolated objects have neither upper nor lower neighbours.

When the indicators are oriented in such a way that large values indicate a risky situation, then the maximal and isolated objects ("extremal" objects) are those which are of special concern. When there are several extremal objects, the reason of "being risky" is to be traced back to the values of the different attributes. Minimal objects are relatively harmless; however, isolated objects are of concern because their indicator values (attributes) are very particular.

The software package PyHasse can be obtained from the second author Dr. Rainer Bruggemann (brg_home@web.de). The software tools have recently been described (Voigt et al., 2010). PyHasse is based on the interpreter programming language Python. PyHasse is in a steady development and currently comprises more than 100 different modules, together with four libraries (two of them developed by the second author). In this paper objects are the selected chemicals, attributes are the measurements in human breast milk and pine needles.

2.2 Similarity

In complex data sets it is often necessary to compare different sets of criteria (attributes). In the similarity analysis we intend to calculate the similarity of different posets (partially ordered sets). This similarity analysis is an important feature of PyHasse, which has been described in detail by Bruggemann and Patil (2010, 2011).

Summing up, in order to describe the behavior of two partial orders in a compact way we use the wording:

- ISO: isotone: matchings ($<, <$) or ($>, >$)
- ANTI: antitone: the matchings ($>, <$) or ($<, >$)

- WISO: weak isotone: the following matchings: (< , ≃) , (> , ≃), (≃, <), (≃, >)
- IND: indifferent: all matchings where || is part of the pair.
- IDE: equivalent: matching (≃, ≃).

Frequently, data matrices to be compared are quite different in size. Then it is not clear whether the similarity relations are due to either the different numbers of attributes or to the different values of the attributes.

2.3 CombiSimilarity

That is the reason why the similarity analysis tool was extended by further algorithms taking the different number of attributes into account. The extended tool is named CombiSimilarity. Given two data matrices $dm1$ and $dm2$. Let $m1$ be the number of columns of $dm1$ and $m2$ those of $dm2$. Without loss of generality we assume $m2 > m1$. Two variants are under discussion:

1) From $dm2$ a set of $\binom{m2}{m1}$ data matrices $dm2'$ is selected, where now the number of attributes of $dm2'$ is the same as that of $dm1$.

2) Each single column of $dm1$ is combined with each single column of $dm2$. Thus we obtain $m1 \cdot m2$ new matrices, all with only two columns.

Both variants lead to a distribution of quantities ISO, ANTI, etc., which can be further examined by appropriate statistical tests, such as the "t-test" allowing for unequal variances. The idea is, to eliminate the bare effect of a greater number of indicators, describing nevertheless the same context.

The focus of this CombiSimilarity approach is on the matching of Isotone (ISO) relations. A detailed description of the CombiSimilarity tool is given by Bruggemann et al. (2014).

3. APPLICATION OF PYHASSE METHODS ON ENVIRONMENT AND HEALTH DATA SETS

3.1 Data sets to be analysed

The occurrence of POPs/OCPs in the Taurus Mountains in Turkey was studied recently (Turgut et al., 2012, Voigt et al., 2012). Five locations were selected in Taurus Mountains to collect pine needles and breast milk of stationary women. Pine needles are effective bio monitors due to the excellent uptake properties of their waxy layer and breast milk as well-known as the state of the art biomonitor for females who in our study mainly feed on localized food items surrogated from pine needles. The needles receive their OCP content from air pollution.

In this approach we want to examine the occurrence of 18 POPs listed in Table 1 in a human medium, namely breast milk, as well as in an environmental medium, namely pine needles.

Table 1. 18 OCPs detected in breast milk and pine needle samples in the Taurus Mountains, 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	PECB	Pentachlorobenzene	608-93-5
05	HCBE	Hexachlorobenzene	118-74-1
06	PPDT	p, p'-Dichlordiphenyltrichlorethane	50-29-3
07	OPDT	o, p'-Dichlordiphenyltrichlorethane	789-02-6
08	PPDD	p, p'-Dichlordiphenyldichlorethane	72-54-8
09	OPDD	o, p'-Dichlordiphenyldichlorethane	53-19-0
10	PPDE	p, p'-Dichlordiphenyldichlorethene	72-55-9
11	OPDE	o, p'-Dichlordiphenyldichlorethene	3424-82-6
12	OXYC	Oxychlorane	27304-13-8
13	CHCE	cis-Heptachloroepoxide	1024-57-3
14	DIEL	Dieldrin	60-57-1
15	END1	Endosulfan-1	959-98-8
16	END2	Endosulfan-2	33213-65-9

17	MECH	Methoxychlor	72-43-5
18	MIRE	Mirex	2385-85-5

3.2 Main Partial Order Model: Hasse Diagrams for 2 Data matrices

First we calculate the Hasse diagram for the 18 chemicals in 7 pine needle samples (Figure 1, lhs) and for 18 chemicals in 44 human breast milk samples (Figure 1, rhs). We apply the main Hasse diagram technique module of the PyHasse software (mHDCI2_7).

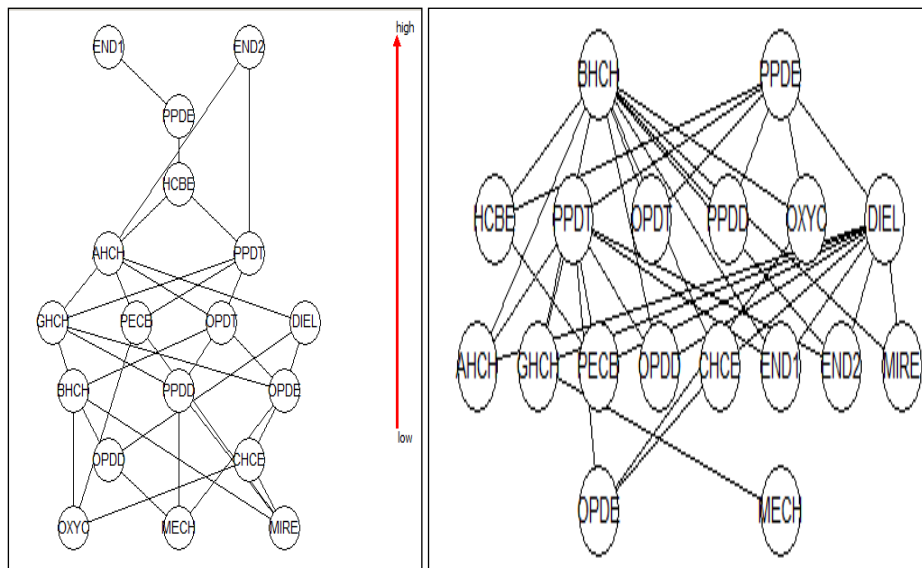


Figure 1. Main Hasse Diagram Technique Module applied on 18x7 Pine Needle (lhs) and 18x44 Breast Milk Samples (rhs).

Endosulfan1 and 2 (END1/END2) are the major pollutants in needle samples whereas DDT (PPDT) and beta-Hexachlorocyclohexane (BHCH) are the maximal objects in human breast milk samples. Concerning Figure 1 lhs (needles) PPDE and PPDT are situated in the second highest level. Methoxychlor (MECH) is a minimal object in both Hasse diagrams. OXY and MIRE are minimal objects in the needles data set and found in the second lowest/highest level in the breast milk samples' data set. This means we can detect several compliances as well as several differences when comparing the two data sets needles versus breast milk pollution. A more precise and sophisticated method is the calculation of similarities applying the similarity module of the PyHasse program.

3.3 Similarity Analysis of needles versus breast milk samples

The Similarity Analysis quantifies the order theoretical differences of two data matrices. We apply this feature on the 18x7 (needle samples) versus 18x44 (breast milk samples) data matrices. The analysis reveals only isotone (60), antitone (28) and indifferent (218) relations. Whereas isotone relations demonstrate a high degree of similarities, antitone relations no similarity and indifferent relations reveal all relations with incomparabilities. In this approach, the relation isotone/indifferent is nearly 1:3.5. This result indicates that for the two data sets of the contamination of needles' samples versus human breast milk samples a statement concerning similarity is hampered by the very large amount of indifferences. Should we take care of the fraction of antitone, which is only around a fifth of the isotone relations?

In this similarity approach we compare two data matrices with different number of attributes, namely, 7 needles' samples with 44 breast milk samples. We want to see whether the CombiSimilarity analyses applied to different size data matrices like in our example here, also provide more precise results.

3.4 CombiSimilarity Analysis of needles' samples versus breast milk samples

The CombiSimilarity is currently only applicable for the isotone relations.

In Table 2 the statistical quantities obtained from the PyHasse module "Combisimilarity7.py" are summarized:

Table 2. Expectation and standard deviation of the theoretical and empirical distributions

	Theoretical distribution (statistical analysis)	Empirical distribution (att_att analysis)
Mean	76.99	90.73
Stdev	13.44	6.42

Empirical data matrices have a mean of 91 whereas the stochastic mean (based on independent permutations) shows 77. Hence, the mean(empir) is larger than the mean of the stochastic distribution and hence there is evidence that the two data matrices (needle and breast milk) are more similar than the stochastically generated distribution. Furthermore, the probability for a wrong decision is calculated and the error probability is less 0.1 %. Therefore, not only the mean(empir) > mean(theor), but also this result is statistically safe. Checking Table 2, it can be verified that even mean(theor) + Stdev(theor) is less than mean(empir)-Stdev(empir).

4. RESULTS, CONCLUSIONS AND OUTLOOK

In this paper, we focus on the possible association between the contamination of pine needles' and human breast milk samples with the same environmental organochlorine pesticides (OCPs). All samples are taken from the Taurus Mountains in Turkey from nearby sites for breast milk donors and needle couples/pairs as their residents. A quantification of the differences can be calculated in the similarity analysis module of the PyHasse program. As a further evaluation step, we applied the newly developed CombiSimilarity tool. This module is still in a testing phase: It must be applied not only to several other and different data matrices, but also the role of the matches other than isotone relations must be checked. Basically the variant 1 with a set of data matrices having the same size must be tested too in order to find out, whether this variant can add more insight or not.

We conclude from the simplified similarity analysis that the fraction isotone is approximately one third of the fraction indifferent. However, there is a small antitone fraction. Therefore there is an indication that other processes are possible which favour dissimilarity between the two data matrices.

Now with the refined methodological approach based on the new module CombiSimilarity, the result shows that there is high evidence that needles' and breast milk pollution are very similar if taken from related residential locations. This means that although we find visible differences in the two Hasse diagrams (see Figure 1) the calculations by the new CombiSimilarity tool reveal a conformity of the two data sets, which is justified by an inferential statistical argument.

In a general context this result on the similarity of needles' and breast milk samples shows strong evidence that organochlorine pesticides once distributed into the environment find their path into both environmental and human samples. This implies to our mind that not only more research on the influence of chemicals especially those which have endocrine potential (as this is the case for many OCPs) should be performed in order to strengthen this portability of environmental chemicals' pollution effects. Furthermore, lessons must be learnt from the environmental contamination leading to the contamination of plants (trees, needles) as well as of human beings.

We will extend our research to other types of chemicals, e.g. Polycyclic Aromatic Compounds (PAHs), Polychlorinated Biphenyls (PCBs). For these chemical groups data sets of the pollution of the Turkish Mountains already exist. By applying our software tools to these data sets we will hopefully be able to develop further refinements of the CombiSimilarity module.

5. REFERENCES

- Birnbaum, L.S. (2013). When environmental chemicals act like uncontrolled medicine. *Trends Endocrinol Metab* 24(7) 321-323.
- Bogdal, C., Abad, E., Abalos, M., van Bavel, B., Hagberg, J., Scheringer, M., Fiedler, H. (2013).

- Worldwide distribution of persistent organic pollutants in air, including results of air monitoring by passive air sampling in five continents. *Trac-Trend Anal Chem* 46, 150-161.
- Bruggemann, R., Patil, G.P. (2010). Multicriteria prioritization and partial order in environmental sciences. *Environmental and Ecological Statistics* 17, 383-410.
- Bruggemann, R., Patil, G.P. (2011). *Ranking and Prioritization for Multi-indicator Systems*. Springer, Berlin.
- Bruggemann, R., Scherb, H., Voigt, K. (2014). CombiSimilarity an Innovative Method to Compare Environmental and Health Data Sets with different Attribute Sizes, Example: 18 Organochlorine Pesticides in Soil and Human Breast Milk Samples. In print *Ecotoxicology and Environmental Safety*, DOI: 10.1016/j.ecoenv.2014.03.031.
- Lee, D.H., 2012. Persistent organic pollutants and obesity-related metabolic dysfunction: focusing on type 2 diabetes. *Epidemiology and health* 34, e2012002.
- Mannetje, A., Coakley, J., Mueller, J.F., Harden, F., Toms, L.M., Douwes, J. (2012). Partitioning of persistent organic pollutants (POPs) between human serum and breast milk: a literature review. *Chemosphere* 89, 911-918.
- Thayer, K.A., Heindel, J.J., Bucher, J.R., Gallo, M.A. (2012). Role of environmental chemicals in diabetes and obesity: a National Toxicology Program workshop review. *Environ Health Perspect* 120, 779-789.
- Turgut, C., Atatanir, L., Mazmanci, B., Mazmanci, M.A., Henkelmann, B., Schramm, K.W. (2012). The occurrence and environmental effect of persistent organic pollutants (POPs) in Taurus Mountains soils. *Environ Sci Pollut Res Int* 19, 325-334.
- 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, 1801-1812.
- Voigt, K., Scherb, H., Bruggemann, R., Schramm, K.-W. (2012). Discrete mathematical data analysis method in support of sustainable chemistry. *Fresenius Environmental Bulletin* 21, 2510-2515.
- Voigt, K., Scherb, H., Bruggemann, R., Schramm, K.W. (2013). Discrete mathematical data analysis approach: A valuable assessment method for sustainable chemistry. *Sci Total Environ* 454-455C, 149-153.