EnviroInfo 2005 (Brno)
Informatics for Environmental Protection - Networking Environmental Information
 ISBN: 80-210-3780-6

Environmental Contamination with Endocrine Disruptors and Pharmaceuticals: An Environmetrical Evaluation Approach

Kristina Voigt[1], Rainer Brüggemann[2], Stefan Pudenz[3] and Hagen Scherb[1]

Abstract

Achieving sustainable development in the environmental and health sector it is absolutely necessary to keep the ground- and consequently the drinking water free of contaminants. Unfortunately several chemicals are detected in environmental media. An intensive literature survey was performed for 7 environmental targets and the presence of 12 selected chemicals (pharmaceuticals and endocrine disruptors) therin. 75 articles recently published in international scientific journals performed the basis for the evaluation approach. The seven attributes to be looked upon the presence of at least one of the 12 drugs in those 7 environmental media. The consideration of the environmental media in any publication is coded by 0 (not available) or 1 (available). A 75 (objects) x 7 (attributes) data-matrix results. Several methods of mathematical and statistical background are applied. These are the Hasse Diagram Technique, a method derived from discrete mathematics, the POSAC (Partially Ordered Scalogram Analysis with Coordinates) method and the PCA (Principle Component Analysis), both multivariate statistical methods. Applying the Hasse Diagram Technique Method the most important publications are revealed as well as the least important ones. The PCA data-analysis method shows an exceptional position for the environmental media surface water and waste water.

1. Introduction: Chemicals in the Environment

Chemicals have been detected in environmental media around the world as the result of human activities for quite a number of years. Among them are also pharmaceutically active compounds (PHACs), personal care products (PPCPs) and endocrine disrupting chemicals (EDCs). These chemicals are consumed in the ton-range per year. They are introduced into sewage to a high extent by households. As these compounds are not totally removed during sewage treatment they are discharged in appreciable quantities into receiving waters through municipal sewage treatment plants (STPs) effluents (Ternes 2005). The process of removal of chemicals in wastewater treatment plants in Europe is described by Ternes (2004). This reveals that the topic of pharmaceuticals and endocrine disruptors in the environment is an evolving issue and asks for further data evaluation strategies. That is the reason why we chose the following 12 drugs for an intensive literature study: Bezafibrate (lipid regulator), Carbamazepine (anti epileptic), Clofibric acid (blood lipid regulator), Diclofenac (anti rheumatic), Diazepam (psychiatric drug), Fenofibrate (lipid regulator), Ibuprofen (analgetic), Metoprolol (beta blocker), Phenazone (analgetic), Ethinyl Estradiol (sex hormone, steroid), Roxithomycine (antibiotic), Sulfamethoxazole (antibiotic). The contraceptive Ethinyl Estradiol proved to be an endocrine disruptor. We performed a literature study looking for current, that is to say articles published in the time period 2000-2004. We found 75 articles in international scientific journals which formed the basis for the evaluation approach. Attributes to be looked upon were the availability of

[1] GSF - National Research Center for Environment and Health, Neuherberg, Germany
[2]Leibniz- Institute of Fresh Water Ecology and Inland Fisheries, Berlin, Germany
email: kvoigt@gsf.de, Internet: http://www.gsf.de/ibb/homepages/voigt
[3] Criterion, Berlin, Germany

the pharmaceuticals in the media surface water (SF), groundwater (GW), drinking water (DW), sewage sludge (SS), wastewater (WW), soil (SO) and sediment SE). The consideration of the environmental media in any publication is coded by 0 (not available) or 1 (available). Hence a 75 (objects) x 7 (attributes) data-matrix with the entries c_{ij} (i=1,…75, j=1,..7)) results.

$$c_{ij} = \begin{cases} 1 \; if \; the \; target \; et \; j \; and \; one \; of \; the 12 \; drugs \; is \; discussed \; in \; the \; i^{th} \; publication \\ 0 \; else \end{cases}$$

2. Environmetrical / Chemometrical Methods Used: HDT, POSAC and PCA

As the data situation in environmental sciences in combination with chemical substances becomes more and more complex, it poses a great challenge for establishing new environmetrical and chemometrical data-analysis methods. Einax summarizes in a recent publication important new chemometrical methods (Einax 2003). One of these challenging new chemometrical and environmetrical method is the method of evaluation by order theory, based on the theory of partially ordered sets, and its specific application, known in literature as the Hasse Diagram Technique (HDT). A new decision support tool is based on WHASSE and is called METEOR (METhod of Evaluation by ORder theory) and is explained by an example of evaluating chemical and environmental databases (Voigt 2004b).

2.1 Hasse Diagram Technique (HDT)

The Hasse Diagram Technique is well explained in a variety of different environmental and chemical as well as statistical journals. In a previous issue of Environmetrics the HDT and POSAC methods are explained in detail (Voigt 2004a). A very comprehensive description of the HDT can be found in a recent publication by the second author (Brüggemann 2001, 2002). That is the reason why only some aspects are picked out, which will be useful in the subsequent application. Hasse Diagrams visualize the order relations within objects: Two objects, also called elements (if the aspect of belonging to sets is important) x, y of an object set are considered as being ordered, e.g. $x \leq y$, if all scores of x are less or equal than those of y. Hasse Diagrams are acyclic digraphs and objects are drawn as small circles together with an appropriate identifier. The edges of this graph are the cover-relations; that means, edges which express simply the transitivity are omitted, as they bear redundant information. In our applications the circles near the top of the page (of the Hasse Diagram) indicate objects that are the "better" objects according to the criteria used to rank them: The objects not "covered" by other objects are called maximal objects. Objects which do not cover other objects are called minimal objects. The WHASSE program is developed by Rainer Brüggemann (2001, 2002), for commercial applications it is recommended to contact the company Criterion – Evaluation and Information Management (Criterion 2005). The commercial program is named ProRank.

2.2 Data Reduction Methods: POSAC and PCA

There are many statistical approaches of condensing a data-matrix by the creation of new variables. This process – called ordination – is often used to visualize relationships in two dimensions. These new variables (in general, two variables), which are derived from the original ones, are constructed to optimize some specific criteria. For example, principal component analysis (PCA) creates new axes to explain as much as possible of the variance of the data-matrix. This idea can be applied if order relations (comparability as well as incomparability) are considered as the essential aspect of the data to be preserved in the analysis (Borg 1995). This method – construction of new axes which presents correctly as many as possible of the order relations – is called Partially Ordered Scalogram Analysis with Coordinates (POSAC). POSAC is integrated in the program package SYSTAT (SPSS 2005) under the feature of statistics, data reduction. For a better interpretation of the new axes correlations between old and new variables can be

calculated (Borg 1995). The background of the POSAC method as well as the mathematics in it, is described in a textbook, entitled "Multiple Scaling" by Shye (Shye 1985). Beyond this, the starting point is the same as in the HDT. A more detailed description of the POSAC method in comparison with the Hasse Diagram Technique as well as their application on the subject of the evaluation of air-pollutant monitoring systems in Europe is published by Voigt et al. (Voigt 2004a). The related METEOR method (Method of Evaluation by Order Theory) is explained by an example of evaluating chemical and environmental databases (Voigt 2004b).

PCA is a procedure for analyzing multivariate data which transform the original variables into new ones that are uncorrelated and account for decreasing proportions of the variance in the data. The aim of this method is to reduce the dimensionality of the data. The new variables, the principal components, are defined as linear functions of the original variables. If the first few principal components count for a large percentage of the variance of the observations (say above 70 %) they can be used both to simplify subsequent analyses and to display and summarize the data on a parsimonious manner (Everitt 1998).

3. Evaluation of the 75 Articles (Objects) by Seven Environmental Media (Attributes)

3.1 Evaluation by Hasse Diagram Technique

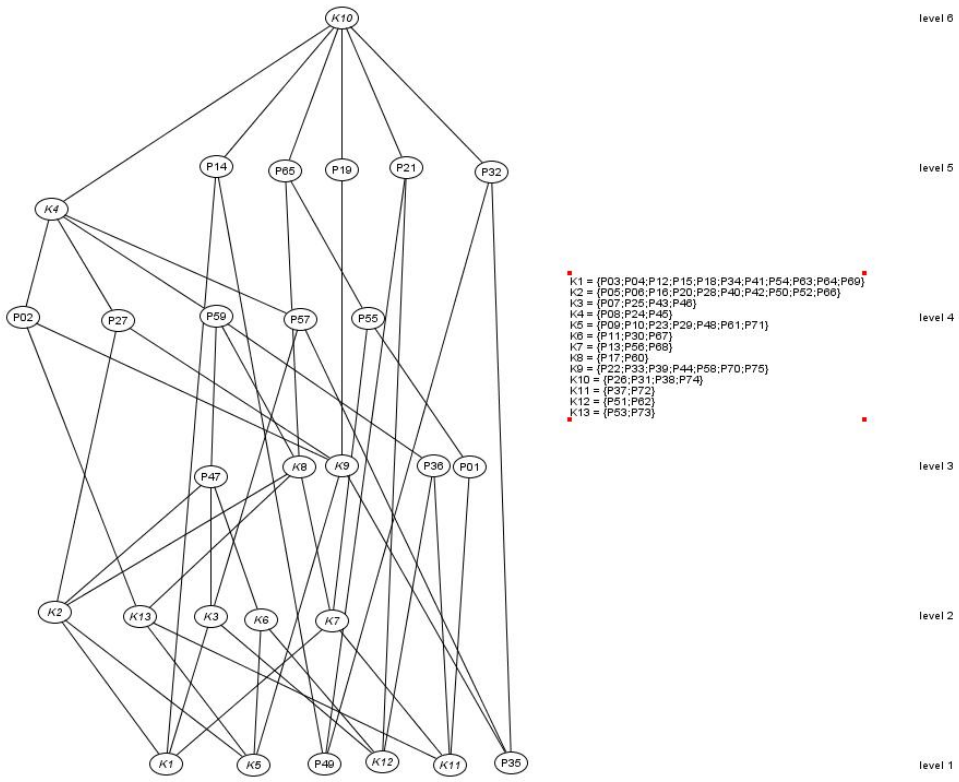

Fig. 1: Hasse Diagram for 75x7 Data-matrix (75 Publications, 7 Environmental Media) For this data analysis we used the ProRank program (Criterion 2005).

The Hasse Diagram in Figure 1 is structured into six levels. Some objects are denoted by K which means they comprise several objects (so-called equivalence classes) and other objects are called P which means

publication number x. The diagram shows one maximal object K10 which means that the publications P26, P31, P38, P74 are equivalent. As K10 is connected with all other objects, this implies that these four publications have more information than all other objects. These articles are review articles which cover many environmental media. We can detect six minimal objects, four of them form equivalence classes. The information in these articles is very small and/or specialized.
Taking a look at all objects (object set) 1457 incomparabilities, 1158 comparabilities and 167 equalities are counted. Although some important ranking information can be drawn out of the diagram, the figure calls for data reduction procedures for more comprehensive conclusions.

3.2 Evaluation by POSAC and PCA

The POSAC - Partially Ordered Scalogram Analysis with Coordinates reduces the data-matrix into a two-dimensional space. The two new variables are named latent order variables (LOV1, LOV2). The equivalence classes are named by the first object of that class. 79 % of the profile pairs are correctly presented, this means that 21 % of the information is lost. The maximal object in the Hasse diagram K10 (P26, P31, P38, P74) shows high values in LOV1 and LOV2. A differentiation between the Hasse diagrams is demonstrated for the minimal objects. They have high values for one latent order variable and low values for the other variable. Figure 2 left hand side shows the POSAC plot.

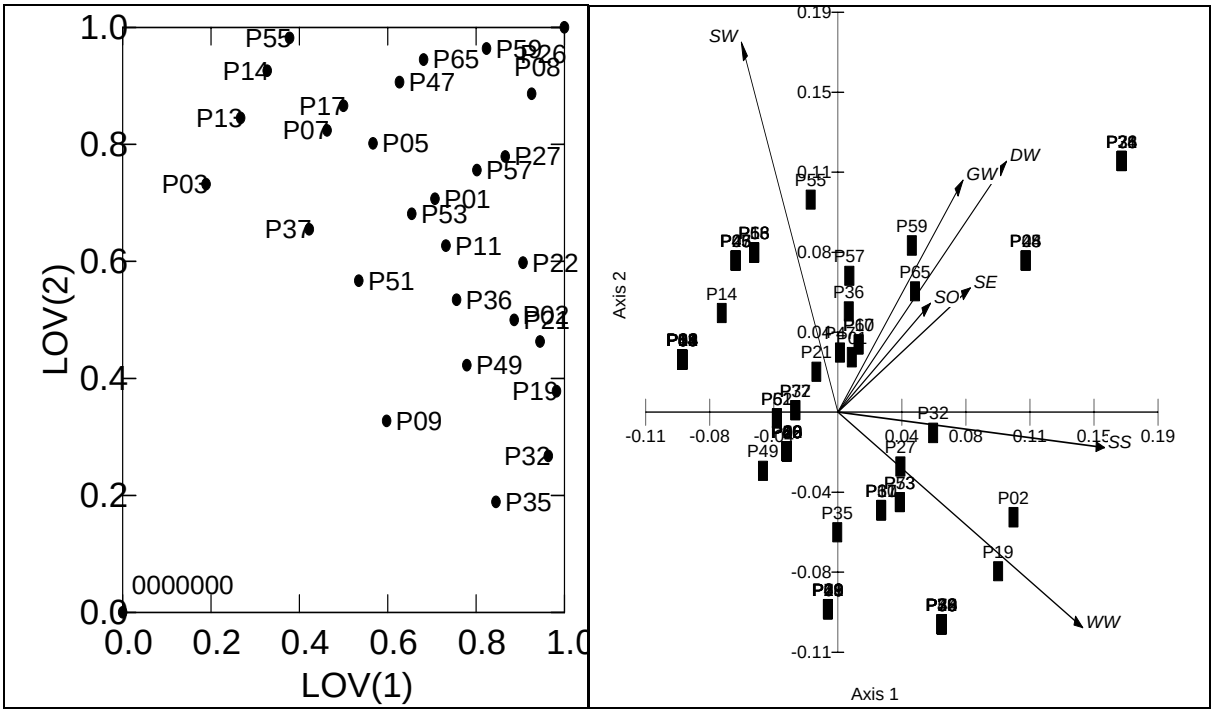

Fig. 2: POSAC Plot (left) (only representative elements of the equivalence classes) and PCA Bi-Plot (right) of 75x7 Data-matrix

3.3 Application of PCA

The PCA was carried out applying the MVSP Multi-Variate Statistical Package (Kovach 2005). The Bi-Plat of the data-matrix is given in Figure 2 right hand side. The declared variance of the first three components is 66 %. The application of the PCA (Principal Component Analysis) reveals an exceptional position of the media SW = surface water" and WW = waste water. The situation of the objects will not be discussed in this short paper.

4. Discussion and Outlook

The short paper reveals the visualisation and detailed data analysis concerning recently published articles about the emerging subject of pharmaceuticals / endocrine disruptors in environmental media. The data-analyses methods derive from the multi-variate statistical aspects "variance" (PCA) and "order relation" (HDT, POSAC). The partial order ranking method Hasse Diagram Technique reveals among other aspects the maximal and minimal objects, that is to say the most and least important publications for the occurrence and detection of 12 pharmaceuticals in different environmental media. The POSAC method differentiates between the minimal objects. In the PCA method the environmental media (attributes) are looked upon. The two media "waste water" and "surface water" have a special position. Further research steps should be performed, e.g. leaving out review articles, enlarging the publication list, including further aspects (attributes), e.g, number of pharmaceuticals, applying the elaborated methodology on collected data (cooperation), scrutinizing the data availability in Internet databases in order to enlarge the knowledge about the important subject of pharmaceuticals and endocrine disruptors in environmental media. We already began with the evaluation of Internet databases concerning their availability on environmental data on the chosen 12 pharmaceuticals and endocrine disruptors.

Bibliography

Borg I, Shye S. (1995): Facet Theory, Form and Content. Sage Publications, Thousand Oaks, pp. 111.112.

Brüggemann R, Welzl G. (2002): Order Theory Meets Statistics – Hasse Diagram Technique, in: Voigt, K, Welzl, G., Order Theoretical Tools in Environmental Sciences, Shaker-Verlag, Aachen, 9-40.

Brüggemann, R., S. Pudenz, et al. (2001). The Use of Hasse Diagrams as a Potential Approach for Inverse QSAR I. f. G. u. Binnenfischerei, SAR and QSAR in Environmental Research. 11: 473-487.

Criterion (2005): WHASSE Software, http://www.criteri-on.de/StartseiteEng.html.

Einax, J. W. (2003): Chemometrics in Environmental Sciences - Challenges and New Methods The Information Society and Enlargement of the European Union, in: Gnauck, A., Heinrich, R. (Eds.), , 2, Metropolis Verlag, Marburg, pp. 51-57.

Everitt B.S. (1998): The Cambridge Dictionary of Statistics, Cambridge University Press, Cambridge,

Kovach Computing Services (2005): Multivariate Statistical Package, http://www.kovcomp.co.uk/mvsp/.

Shye S. (1985): Multiple Scaling; Elsevier Publishers, Amsterdam.

SPSS Science (2005): Systat, http://www.spssscience.com/SYSTAT/index.html.

Ternes, T., Bonerz, M. (et al) (2005): Determination of pharmaceuticals, iodinated contrast media and musk fragrances in sludge by LC/tandem MS and GS/MS, Journal of Chromatography A. 1067, pp. 213-223.

Ternes, T., Joss, A., Siegrist, H. (2004): Scrutinizing Pharmaceuticals and Personal Care Products in Wastewater Treatment, Environmental Science & Technology, 15, pp. 392A-399A.

Voigt, K, Welzl, G, Brüggemann R. (2004a): Data Analysis of Environmental Air Pollutant Monitoring Systems in Europe. Environmetrics 15: 577-596.

Voigt, K., R. Brüggemann, Welzl, G. (2004): Chemical Databases Evaluated by Order Theoretical Tools, Anal Bioanal Chem, 380, pp. 467-474.