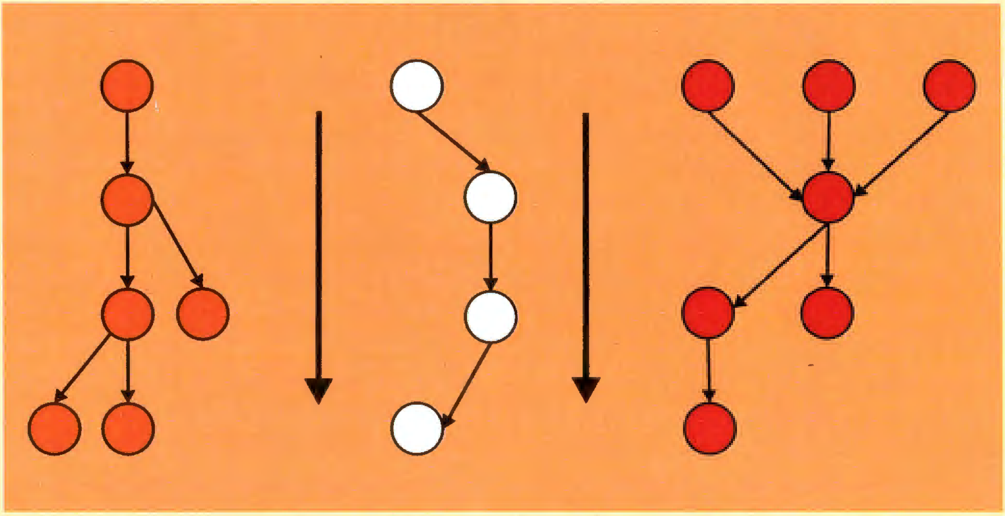
SYSTEMS RESEARCH INSTITUTE
POLISH ACADEMY OF SCIENCES

# MULTICRITERIA ORDERING AND RANKING: PARTIAL ORDERS, AMBIGUITIES AND APPLIED ISSUES



**Jan W. Owsiński and Rainer Brüggemann**
**Editors**

Warsaw 2008

SYSTEMS RESEARCH INSTITUTE
POLISH ACADEMY OF SCIENCES

# MULTICRITERIA ORDERING AND RANKING: PARTIAL ORDERS, AMBIGUITIES AND APPLIED ISSUES

Jan W.Owsiński and Rainer Brüggemann
Editors

Warsaw 2008

This book is the outcome of the international workshop held in Warsaw in October 2008 within the premises of the Systems Research Institute. All papers were refereed and underwent appropriate modification in order to appear in the volume. The views contained in the papers are, however, not necessarily those officially held by the respective institutions involved, especially the Systems Research Institute of the Polish Academy of Sciences.



ISBN 83-894-7521-9
EAN 9788389475213

Technical editing and typesetting:
Jan W.Owsiński, Anna Gostyńska, Aneta M.Pielak

# Applications and Comparisons

*Multicriteria Ordering and Ranking: Partial Orders, Ambiguities and Applied Issues*
*Jan W. Owsiński and Rainer Brüggemann, Editors*

# Hasse Diagrams as Exploratory Tool in Environmental Data Mining: A Case Study

**Stefan Tsakovski* and Vasil Simeonov****

*Chair of Physical Chemistry, Faculty of Chemistry, University of Sofia " St. Kl. Okhridski", 1164 Sofia, 1, J. Bourchier Blvd., Bulgaria, (STsakovski@chem.uni-sofia.bg)

**Chair of Analytical Chemistry, Faculty of Chemistry, University of Sofia "St. Kl. Okhridski", 1164 Sofia, 1, J. Bourchier Blvd., Bulgaria, (VSimeonov@chem.uni-sofia.bg)

The present paper deals with the presentation of a new approach in sediment quality estimation. This approach studies the ranking towards ecotoxicity of lake sediments samples from Turawa lake in Poland with respect to chemical pollutants like polychlorinated biphenyls (PCBs), pesticides, polyaromatic hydrocarbons (PAH) and heavy metals. The sediment quality estimation is done by combining of Hasse diagram technique (HDT) with self-organized maps (SOM). The classification abilities of SOM are used as preprocessing procedure of the initial data matrix before applying HDT. The Hasse diagram analysis shows two sediment pollution "patterns" instead of the four detected by SOM. The Hasse diagram structure is discussed in relation to the environmental management of Turawa lake region.

Keywords: Hasse diagram technique, self-organizing maps, sediment quality, PCB, PAH, heavy metals

## 1. Introduction

Sediments can serve both as reservoirs and as potential sources of contaminants to the water column, and can adversely affect sediment-dwelling organisms, aquatic-dependent wildlife and human health (US EPA 2005).

Effective sediment risk assessment requires finding a relationship between sediment chemistry and toxicity endpoints. In the present state of art one common approach is linking chemical concentrations to toxicity data (MacDonald et al., 2000). This is a typical univariate strategy which produces traditional Sediment

Quality Guidelines (SQG). The problems of the SQG estimation procedure are connected with the bioavailability of sediment contaminants and the effects of covarying chemicals and chemical mixtures. The "mixture paradox" is somehow resolved by "grouping" contaminants, using empirically derived SQG's (Swartz, 1999). However, it is our conviction that the above mentioned problems can be solved, to a large extent, simply with the application of multivariate statistical methods. Moreover, these methods could be applied even to smaller data sets (usually collected during short-term monitoring).

Assessment of the impact of pollution on biological diversity in water bodies requires not only good quality bottom sediment datasets, but also a complete multivariate statistical data analysis. Many studies have been performed using chemometric approaches, like cluster analysis, principal component analysis, receptor modeling, for monitoring datasets as the best way for classification, modeling and interpretation of various environmental compartments, just to cite a limited number of personal case studies (Karadjov et al., 1990; Eddins et al., 1999; Stanimirova et al., 1999; Spanos et al., 2002). It is worth to mention also that in some studies self-organizing maps were used for classification of sediment samples (Chakraborty et al., 2001; Lacassie et al., 2004) and for estimation of the relationship between ecotoxicity and chemical parameters (Tsakovski et al., submitted).

The partial ordering theory approaches have a relatively limited place in environmental risk assessment related mainly to prioritizing of chemicals according to environmental hazards (Brüggemann, 1996; Halfon et al., 1998). Also some studies deals with ranking of environmental media according to their environmental status (Brüggemann, 2001; Brüggemann, 2003) or ranking of monitoring systems (Voigt et al., 2004).

The aim of the present study is to use Hasse diagram as exploratory tool in sediment quality estimation after SOM classification. The expectation is obtain additional information for the relationship between ecotoxicity and chemical pollutants unifying in an original way the advantages of SOM classification as a pre-processing approach and Hasse diagram technique as ranking method.

## 2. Experimental

### 2.1. Sampling region, sampling and sample analysis

In a study dealing with the cluster and principal components analysis of bottom sediments from Turawa Lake (Simeonov, 2007), important information about the sampling region, sampling and analytical procedures is thoroughly described. The Turawa Lake sampling region is located in south-western Poland (Opole Voividship) and the main pollution sources are industry, agricultural activity and urban and domestic sewage.

The bottom sediment dataset was obtained during a sampling campaign in 2004. From the initial collection of a total 34 sediment cores (0.00 – 8.00 m in length), out of 260 samples, only 59 were chosen for the present multivariate statistical analysis, as they comprise all the necessary data without any missing results for both chemical and ecotoxicity parameters (42 variables in total, including exotoxicity parameters, pesticides, congeners, PAH and heavy metals). The samples were representative for the surface layer of the sediment and involved 59 sampling sites from the bottom sediment of Turawa Lake. In Fig. 1, the sampling grid is presented.

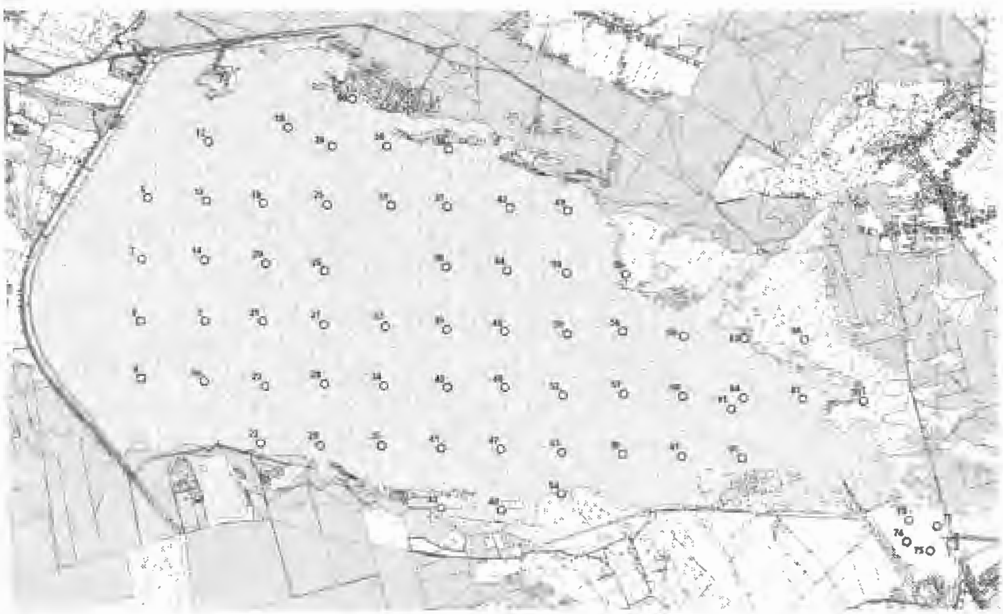


Figure 1. Sampling grid for the bottom sediments (surface layer) from Turawa Lake

Chemical analysis of the samples included various instrumental methods complied with different chemical and ecotoxicity variables. Gas chromatography coupled with mass spectrometry was used for polychlorinated byphenil congeners (PCB), organochlorine pesticides, and polyaromatic hydrocarbons (PAH) determination; inductively coupled plasma – atomic emission spectrometry – for Cr, Cu, Ni, V, Fe, Al, Li; electrothermal atomic absorption spectrometry – for Cd and Pb; hydride generation atomic absorption spectrometry – for As, and cold vapor atomic absorption spectrometry – for Hg.

Table 1. Organic chemical species and their abbreviations used in the study

| **acronym** | **name** |
|---|---|
| PCB28 | 2,4,4'-trichlorobiphenyl |
| PCB52 | 2,2',5,5'-tetrachloro-1,1'-biphenyl |
| PCB101 | 2,2',4,5,5'-pentachlorobiphenyl |
| PCB118 | 2,3',4,4',5-pentachlorobiphenyl |
| PCB138 | 2,2',3,4,4',5'-hexachlorobiphenyl |
| PCB153 | 2,2',4,4',5,5'-hexachloro-1,1'-biphenyl |
| PCB_180 | 2,2',3,4,4',5,5'-heptachlorobiphenyl |
| a_HCH | alpha-1,2,3,4,5,6-hexachlorocyclohexane |
| b_HCH | beta-1,2,3,4,5,6-hexachlorocyclohexane |
| g_HCH | gamma-1,2,3,4,5,6-hexachlorocyclohexane |
| hepta_Cl | heptachlor |
| aldrine | 1,2,3,4,10,10-hexachloro-1,4,4a,5,8,8a-hexahydro-1,4:5,8-dimethanonaphthalene |
| hepta_Cl_B | heptachlor epoxide isomer B |
| pp_DDE | 1,1-dichloro-2,2-bis(p-chlorophenyl)ethylene |
| op_DDD | o,p-dichlorodiphenyl dichloroethane |
| dieldrine | (1aα,2ß,2aα,3ß,6ß,6aα,7ß,7aα)-3,4,5,6-9,9-hexachloro-1a,2,2a,3,6,6a,7,7a-octahydro-2,7:3,6-dimethanonaphth[2,3-b]oxirene |
| endrine | 3,4,5,6,9,9,-hexachloro-1a,2,2a,3,6,6a,7,7a-octahydro-2,7:3,6-dimethanonaphth[2,3-b]oxirene |
| pp_DDD | 1,1-bis(p-chlorophenyl)-2,2-dichloroethane |
| op_DDT | o,p'-dichloro-1,1-diphenyl-2,2,2-trichloroethane |
| pp_DDT | p,p'-dichloro-1,1-diphenyl-2,2,2-trichloroethane |
| HCB | hexachlorobenzene |
| BaA | benzo[a]anthracene |
| BbF | benzo[b]fluoranthene |
| BkF | benzo[k]fluoranthene |
| BaP | benzo[a]pyrene |
| IndP | indeno[123-cd]pyrene |
| DahA | dibenzo[ah]anthracene |
| BPer | benzo[ghi]perylene |

Chronic toxicity was tested in the presence of *Heterocypris incongruens* crustacean (Ostradoxkit, 2005) and mortality (MORT in arbitrary units; a.u.) was chosen for numerical output of the toxicity tests. In Table 1, the coded names of the chemical pollutants are given.

2.2. Methods

SOM is an algorithm used to visualize and interpret large high-dimensional data sets (Kohonen, 2001). SOM is unsupervised pattern cognition method similarly to Cluster analysis. The main advantage of SOM is the simultaneous classification of variables and objects (sampling locations). Typical applications are visualization of process states or financial results by representing the central dependencies within the data on the map. The map consists of a regular grid of processing units called neurons (see Fig. 2).

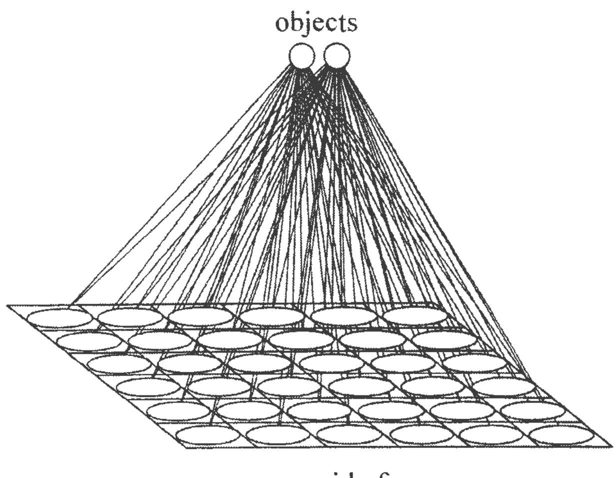



Figure 2. Self-organizing map architecture

A model of some multidimensional observation, eventually a vector consisting of features (variables, attributes), is associated with each unit. The map attempts to represent all available observations with optimal accuracy using a restricted set of models. At the same time the models become ordered on the grid so that similar models are close to each other and dissimilar models far from each other. Fitting of the model vectors is usually carried out by a sequential regression process, where

$t$ = 1,2,... is the step index: For each sample $x(t)$, first the winner index $c$ (best matching unit - BMU) is identified by the condition:

$$\forall i, \|x(t) - m_c(t)\| \leq \|x(t) - m_i(t)\|.$$

After finding the BMU, the weight vectors of the SOM are updated so that the BMU is moved closer to the input vector in the input space. Further, all model vectors or a subset of them that belong to nodes centered around node $c = c(\mathbf{x})$ are updated as:

$$m_i(t+1) = m_i(t) + h_{c(x),i}(x(t) - m_i(t)).$$

Here $h_{c(x),i}$ is the "neighborhood function", a decreasing function of the distance between the $i$-th and $c$-th nodes on the map grid. This regression is usually reiterated over the available objects.

The trained map could be graphically presented by 2D planes for each variable indicating variable distribution values on the different map regions by different colours. To perform SOM-based classification a free Teuvo Kohonen toolbox (SOM Toolbox 2.0) was applied, which can be downloaded together with documentation (Vesanto et al., 2000).

Main advantages of SOM algorithm compared to other pre-processing (classification and dimensionality reduction) techniques are: (1) the projection of variables similarity in the form of features' planes delivers semi-quantitative information about the distribution of a given feature in the space of the cases; (2) SOM visualization enables presentation both similarity between positive as well as negative correlated features; (3) SOM visualization and SOM-supported classification is able to indicate "outliers" i.e. those features or cases which do not belong to a well-organized, homogeneous populations; (4) SOM is noise tolerant (this property is highly desirable when site-measured data are used).

Hasse diagrams visualize order relations between objects described by certain number of attributes. Hasse diagram technique (HDT) is well described elsewhere (Brüggemann et al, 2001) and here only brief description concerning present study will be mention.

In HDT the ranking of objects (elements), E, is done with respect to all variables (attributes), which is called the "information basis" (IB). The processed data matrix **Q** (NxR) contains N elements and R attributes. The entry $y_{ir}$ of **Q** is the numerical value of the *r-th* attribute of the *i*-th element. The two elements $s$ and $t$ are comparable if:

$$s,t \in E; s \leq t \Leftrightarrow y(s) \leq y(t)$$

$$y(s) \leq y(t) \Leftrightarrow y_r(s) \leq y_r(t) \text{ for all } y_r \in IB$$

If there is only one $y_r$ for which $y_r(s) > y_r(t)$ then the objects $s$ and $t$ are incomparable. Partial order set could be easily developed by Hasse matrix which collected relations between each pair of elements. The information stored in Hasse matrix is visualized by Hasse diagram, where the objects are drawn as small circles together with an appropriate identifier. In present study the circles near of Hasse diagram indicate objects that are the 'polluted' 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*. In some diagrams there also exist *isolated objects* which can be considered as maximal and minimal objects at the same time. A *chain* is a set of comparable elements, therefore levels can be defined as the longest chain within the diagram. An antichain is a set of mutually incomparable elements. The height of the Hasse diagram is the longest chain and longest antichain is its width.

The sensitivity analysis of Hasse diagram towards attributes describing objects could be done by dissimilarity matrix (W matrix). The W-matrix represents the influence of the attributes by metric distance between posets, based on different subsets of IB (R-1 attributes). All calculations concerning HDT were performed by the software package WHASSE (Brüggemann et al., 1995).

It seems reasonable to combine the advantages of SOM classification with the main advantage of HDT to reveal in a hierarchical ranking way the links between objects.

## 3. Results and discussion

In our previous study (Tsakovski et al., submitted) the relationship between ecotoxicity (MORT) and chemical components (polluting species like polychlorinated biphenyls (PCBs), pesticides, polyaromatic hydrocarbons (PAH), heavy metals) of lake sediments samples from Turawa Lake, Poland was investigated by an application of self-organizing maps. From the SOM classification of sampling sites with respect to chemical parameters obtained, it is possible to select four groups of ecotoxicity and to analyze within each one of them the relationship of the chemical parameters to the ecotoxicity (presented by mortality value). The four groups, having enhanced mortality, were compared with respect to their "concentration indices," introduced by the authors. The concentration index for each chemical parameter is formed as a ratio of the average concentration of the parameter in the group, divided by the gross average of all data for the same parameter. The fraction was multiplied by 100 to get the index in percentage. With this data normalisation, it is relatively easy to visualize the connection between mortality and a given chemical parameter. It could be assumed that a high mortality index is achieved if other quality parameters have enhanced indexes. Thus, the overall chronic toxicity could be attributed to certain chemicals for each group of sites indicated on the SOM pattern of mortality. It has to be emphasized that only enhanced mortality values were of interest for this classification mode. So, 26 out of all 59 sites were involved in the assessment procedure. Those which are not taken in consideration possess mortality lower than 30%. In Figs. 3a, 3b, 3c and 3d the values of the indices for the chemical pollutants and mortality are shown.

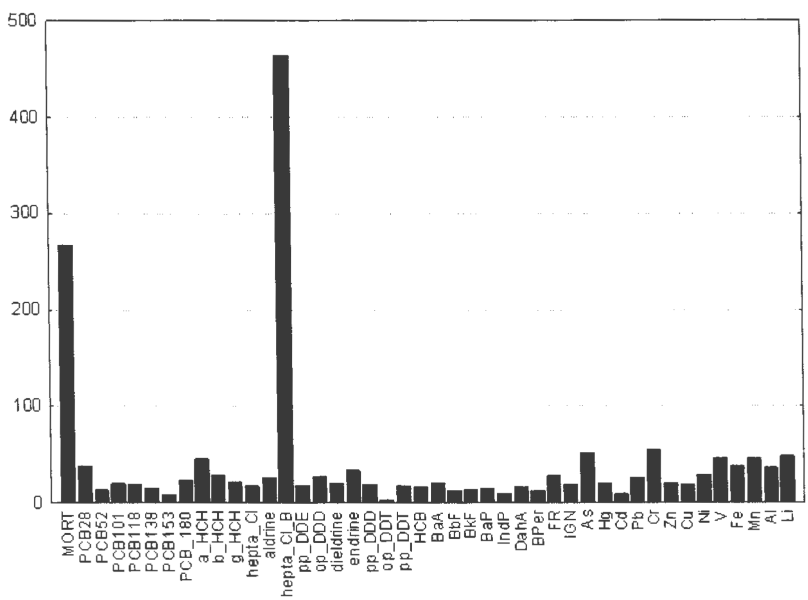



Figure 3a. Indices distribution for all parameters in group 1

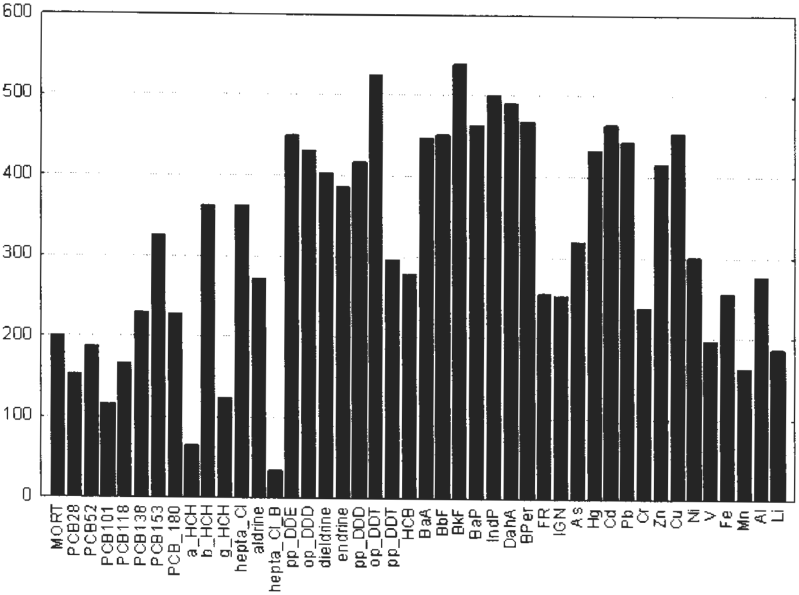



Figure 3b. Indices distribution for all parameters in group 2

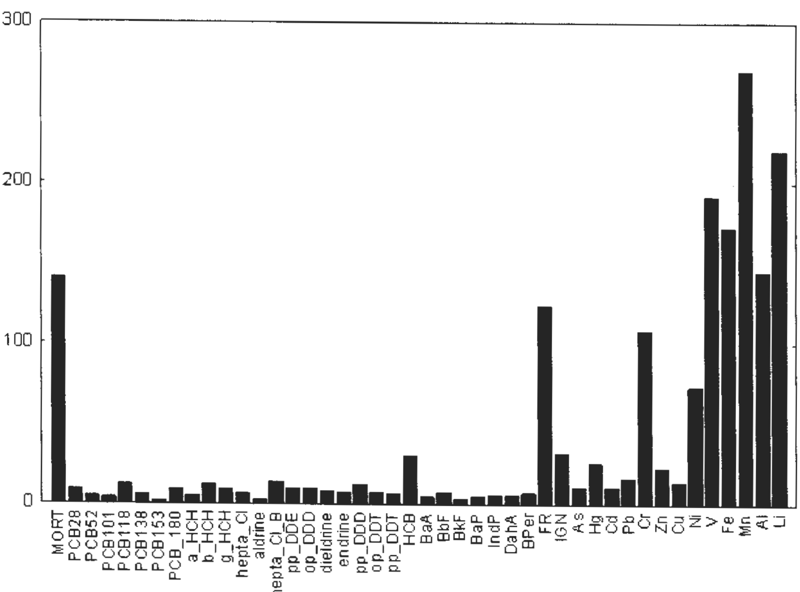



Figure 3c. Indices distribution for all parameters in group 3

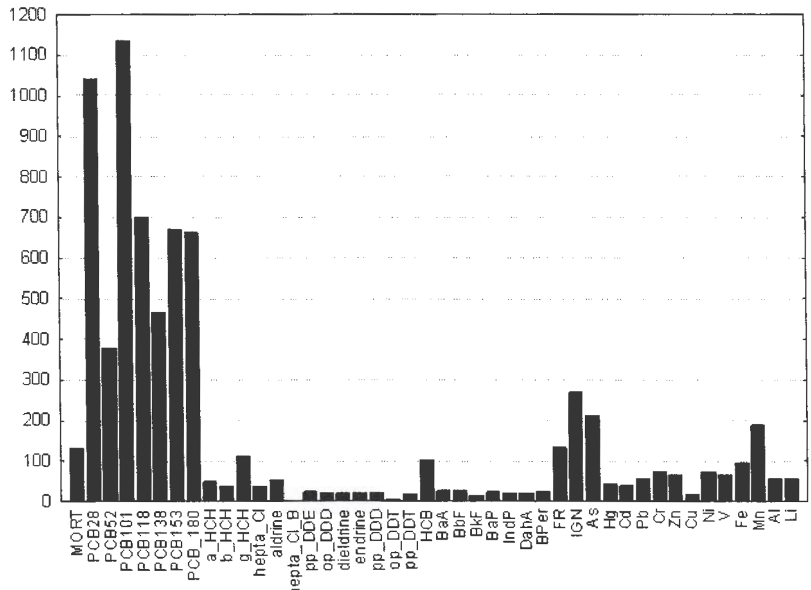



Figure 3d. Indices distribution for all parameters in group 4

The four individual high-mortality groups formed on the SOM could be interpreted as follows.

***Group 1***: Nine sites are included in this pattern, having the numbers 18, 34, 55, 64, 70, 73, 75, 76, 80. The mortality index for the group is 268 %, having the highest value compared to all other values. As seen in Fig. 3a, the only chemical parameter with a very high index (nearly 500 %) is heptachlor B. It may be assumed that for this particular group (and location) of sites, chronic toxicity is due mainly to the toxic effect of specific accumulation of this pesticide. Conditionally, this pattern of chronic toxicity could be named "***heptachlor B***" produced toxicity.

***Group 2:*** Another 9 sites are included in the second classification group (sites 6, 7, 8, 9, 13, 14, 19, 20, 25). They are characterized by a relatively lower chronic toxicity index (nearly 200 % or two and a half times lower than that of group 1). In this case, the contribution of heptachlor B or HCH to the mortality is obviously negligible (Fig. 3b). All other chemicals, however, show high indices: between 100 and 200 % are many PCB chemicals (PCB 28, PCB 52, PCB 101, PCB 118), gHCH, V, Mn, Li; all other quality parameters are over 200 %, as extremely high indices are found for ppDDE, opDDD, dieldrine, endrine, ppDDT, op DDT, all PAHs, Hg, Cd, Pb, Zn, Cu. Thus, this cluster of sites (located almost without exception in the left end of Turawa Lake) could be conditionally related to a ***"pesticide and PAH"*** caused chronic toxicity, but without any specific effect of a certain chemical as in the previous case. Probably, the sedimentation of different chemicals is facilitated exactly in this part of the lake due to stream dynamics reasons.

***Group 3***: Five sites are included in the next level of the chronic toxicity pattern. The mortality index is again lower than before (nearly 140 %) and it is related to relatively high indices of several metals – V, Fe, Mn, Al, Li (Fig. 3c). It may be assumed that chronic mortality is linked directly in this situation with the sediment structure (all metals are mainly major components of the sediments, the fraction size is also a structural characteristics). Therefore, this pattern of chronic toxicity is ***"heavy metal"*** caused. The sites' identification numbers are 27, 40, 53, 58, 65 forming a relatively compact group of sites near the right bench of the Lake, close to the sites of the first group.

***Group 4***: The last obvious pattern of sites grouped in the mortality SOM includes only 3 sites (16, 51, 57). The group has an index of 130 % and the calculations indicate that this chronic mortality is probably related to the very high concentrations (indices) of all PCBs (between 600 and 1100 %). The conditional name of this group could be ***"PCB congeners"*** caused mortality. The location of the three sites in this case is quite occasional. The indices are presented in Fig. 3d.

This classification, aiming to find a relationship between mortality parameters of the sediment samples and their chemical composition, proves unambiguously that this relationship is quite complex and multivariate.

The appropriate next step in sediment quality estimation could be ranking of sediments according to chemical pollutants significantly affecting chronic toxicity and comparison between poset obtained with measured mortality values.

Environmental data often are associated with significant degree of uncertainty inherent in ranking analysis. This general limitation of HDT could be solved by using of multivariate statistical techniques like clustering, principal component analysis or multidimensional scaling as pre-processing tools. The aim of abovementioned techniques is to reduce dimensionality of the input data matrix. In this study SOM are used as preprocessing tool since they have opportunity for simultaneous classification of attributes (chemical pollutants) and objects (sediment samples).

The ordering of 2D SOM planes (see Fig. 4) presents classification of chemical pollutants and mortality with respect to surface sediments.

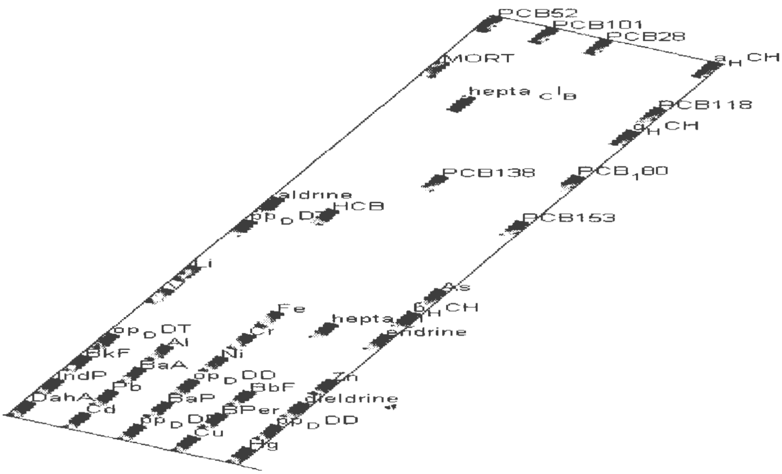


Figure 4. Classification of chemical pollutants and mortality

The presented classification scheme gives opportunity, by detected similarities between chemical pollutants and mortality, for reducing the number of chemical pollutants which describe sediments. The "new" attributes are obtained by summing of concentrations of similar chemical pollutants (see Table 2).

Table 2. Grouping of chemical pollutants

| Attributes | Chemical pollutants |
|---|---|
| SUM PCB 1 | PCB 28, PCB 52, PCB 101 |
| SUM PCB 2 | PCB 118, PCB 138, PCB 153, PCB 180 |
| SUM HM | Hg, Cd, Pb, Cr, Cu, Ni, Li, V, Fe, Al |
| SUM PAH | BaA, BbF, BkF, BaP, IndP, DahA, BPer |
| SUM PEC 1 | b_HCH, hepta_Cl, pp_DDE, op_DDD, dieldrine, endrine, pp_DDD, op_DDT |
| SUM PEC 2 | pp_DDT, HCB, aldrine |

Additionally the concentrations of hepta_Cl_B and As are included as attributes. The a_HCH, g_HCH are omitted because they possess high concentrations only for one sediment location.

The new set of attributes was used for SOM classification of sediment locations. The 59 sediment samples are projected on 2D map with dimensionality 5x8. On Fig. 5 the number of plane units and number of objects related to each unit (hexagon) are shown. The 59 sediments are grouped in 26 plane units.

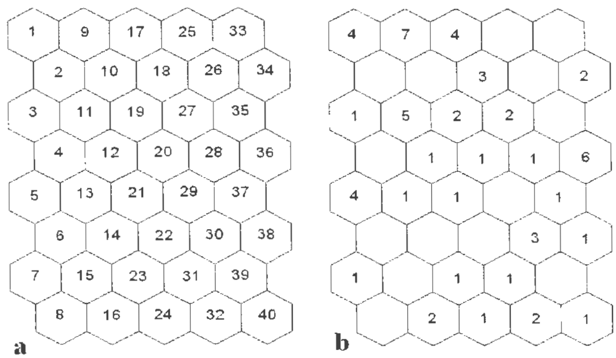



Figure 5. 2D SOM plane (a) unit number, (b) number of objects in each unit.

Each populated unit could be used as equivalence class, which includes the matching objects (Table 3). For attribute values the map trained vector of corresponding unit will be used.

Table 3. Equivalence classes

| Equivalence class (unit number) | Objects (sediment grid location) | Equivalence class (unit number) | Objects (sediment grid location) |
|---|---|---|---|
| 1 | 23, 35, 49, 54 | 21 | 45 |
| 3 | 29 | 23 | 9 |
| 5 | 22, 24, 30, 56 | 24 | 13 |
| 7 | 28 | 27 | 55, 70 |
| 9 | 15, 42, 50, 59A, 60, 65, 67 | 28 | 58 |
| 11 | 33, 38, 52, 59, 74 | 30 | 14, 20, 21 |
| 12 | 36 | 31 | 6 |
| 13 | 16 | 32 | 7, 8 |
| 16 | 19, 46 | 34 | 18, 75 |
| 17 | 47, 48, 61, 63 | 36 | 34, 53, 64, 73, 76, 81 |
| 18 | 40, 41, 66 | 37 | 57 |
| 19 | 39, 46 | 38 | 25 |
| 20 | 27 | 40 | 51 |

Ranking of equivalence classes by HDT was performed. The order of attributes in Hasse diagrams presented in Fig. 6 is as follows: SUM PCB 1, SUM PCB 2, hepta_Cl_B, SUM PEC 1, SUM PEC 2, SUM PAH, As, SUM HM.

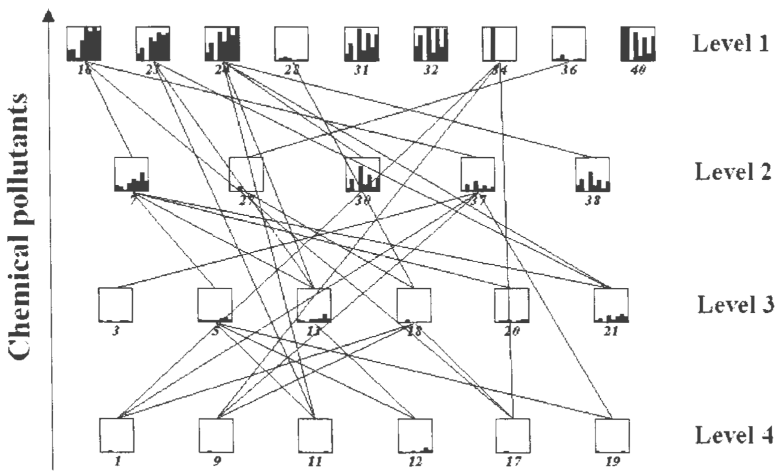



Figure 6. Hasse diagram for the 26x8 reduced data set

It can be observed that Turawa lake sediments have been ranked in four priority levels. The diagram points out 9 classes (16, 23, 24, 28, 31, 32, 34, 36, 40) in Level 1 that represent the most polluted locations (17 out of 59) in Turawa lake. Three of the classes are isolated elements: 31, 32 and 40. Class 40 possesses the highest values for SUM PCB 1 and SUM PCB 2, till classes 31 and 32 for SUM PCB 1 and SUM PCB 2. The other elements could be divided in two groups. The first one includes 28, 34 and 36, where hepta_Cl_B is dominant pollutant. In the second group (16, 23, 24) all the other pollutants (pesticides, PAH's, As and heavy metals) exhibit elevated values. The mortality values of the classes in Level 1 are in the range 42-55 a.u.

Five classes (7, 27, 30, 37, 38) are located at the second level. They again could be divided into two groups: class 27 belongs to "hepta_Cl_B" pattern till the other to "mixed" one. It is worth to mention that sequences $16 \geq 37$ and $24 \geq 38$ do not correspond to chronic toxicity of the classes. The possible reasons for the higher mortality values of classes 37 and 38 could be the presence of other pollutants or specific pollutant transformations.

The level 3 and 4 consolidate 12 classes representing 34 out of 59 sediment locations. All objects could be defined as less polluted and have mortality values less than 30 a.u.

The sensitivity analysis of Hasse diagram was performed by calculation of dissimilarity W matrix for different combinations of R-1 attributes and the calculated sensitivities are presented in Table 4.

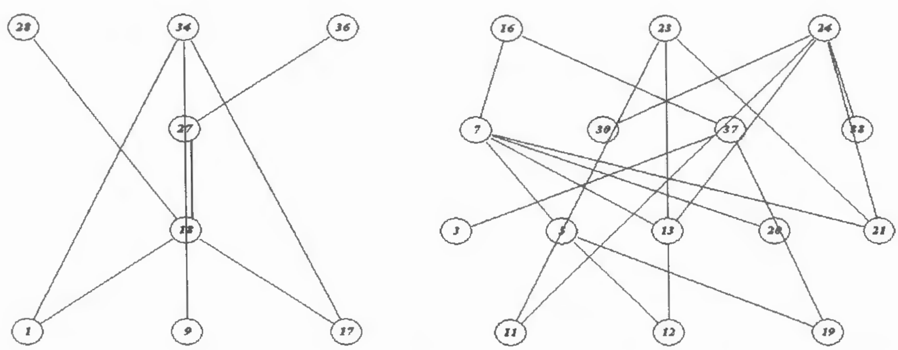
Table 4. Attributes sensitivities

| Attribute | Sensitivity |
|---|---|
| SUM PCB 1 | 4 |
| SUM PCB 2 | 1 |
| hepta_Cl_B | 189 |
| SUM PEC 1 | 1 |

| Attribute | Sensitivity |
|---|---|
| SUM PEC 2 | 0 |
| SUM PAH | 0 |
| As | 3 |
| SUM HM | 0 |

Therefore, the pollutant hepta_Cl_B is the most important one within the attribute set. The sensitivity of SUM PEC 2, SUM PAH and SUM HM is 0.

It could be concluded that Hasse diagram has two "pollution" branches: "hepta_Cl_B" and "mixed" one. The branches including respective maxima are presented on Fig. 7.




a b

Figure 7. The hepta_Cl_B (a) and mixed branch (b) of Hasse diagram presented on Fig. 6

The hepta_Cl_B branch (Fig. 7a) includes 29 sediment locations. All sequences in the branch correspond to measured mortality values. It could be concluded that quality status of a half of the investigated sediments in Turawa lake depends on hepta_Cl_B concentrations. The "mixed" branch (Fig. 7b) of Hasse diagram includes the other 30 sediment locations. The rest of attributes arrange sediment locations in two big groups. First group (level 1 and 2) has mortality values higher than 30 a.u. and the second one lower than 30 a.u. Again, some contradictions between pollutants sequences and chronic toxicity values discussed above were observed.

It is interesting to have a look on the position of the chains in the two branches on the 2D plane of SOM (see Fig. 8). Two groups of chains are located in different regions of SOM plane. The chains in hepta_Cl_B patterns (Fig. 8a) start from the most toxic class 36 and end to classes 1, 9 and 17 which collected the sediment locations with lowest pollution level. The chains in mixed branch (Fig. 8b) starting from class 16 represent the second direction from toxic to non-toxic sediment locations.

The information obtained by HDT is very useful concerning chronic toxicity origin with respect to sediment locations. Further monitoring procedures could be restricted to one sediment location from each class. It is reasonable that sampling should include the isolated elements representing unique sediment locations and chains in the diagram where the changes in pollutant concentrations could be monitored. Future monitoring results could be placed in already trained SOM and in Hasse diagram obtained, respectively. This opportunity is of high priority for possible remediation activities.

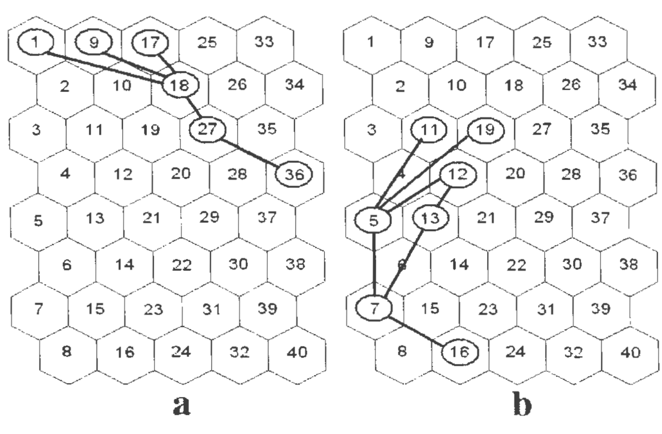



Figure 8 The location of the chains in hepta_Cl_B branch (a) and in "mixed" branch (b) on the SOM plane

The last step in the analysis of sediment status is to compare total ranking of sediment location with their mortality index. For this purpose the sediment poset was represent by a set of linear extensions. A summary of the linear extension analysis is presented in Table 5.

Table 5. Linear extension analysis

| Class | Min | $R_{kav}$ | Max | U(x) |
|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 0 |
| 3 | 2 | 2 | 2 | 0 |
| 5 | 10 | 10 | 10 | 0 |
| 7 | 13 | 13 | 13 | 0 |
| 9 | 3 | 3 | 3 | 0 |
| 11 | 4 | 4 | 4 | 0 |
| 12 | 5 | 5 | 5 | 0 |
| 13 | 6 | 6 | 6 | 0 |
| 16 | 16 | 22.71 | 26 | 10 |
| 17 | 7 | 7 | 7 | 0 |
| 18 | 8 | 8 | 8 | 0 |
| 19 | 9 | 9 | 9 | 0 |
| 20 | 11 | 11 | 11 | 0 |
| 21 | 12 | 12 | 12 | 0 |
| 23 | 14 | 17.56 | 26 | 12 |
| 24 | 17 | 23.8 | 26 | 9 |
| 27 | 14 | 15.84 | 25 | 11 |
| 28 | 15 | 20.31 | 26 | 11 |
| 30 | 15 | 18.75 | 25 | 10 |
| 31 | 15 | 20.38 | 26 | 11 |
| 32 | 15 | 20.39 | 26 | 11 |
| 34 | 15 | 20.39 | 26 | 11 |
| 36 | 15 | 21.31 | 26 | 11 |
| 37 | 15 | 18.53 | 25 | 10 |
| 38 | 15 | 19.21 | 25 | 10 |
| 40 | 15 | 20.82 | 26 | 11 |

Note: Min – minimum rank, Max – maximum rank, $R_{kav}$ – averaged rank; U(x) - uncertainty

Comparison between $R_{kav}$ and mortality index of the classes in Hasse diagram is shown on Figure 9. The good correlation coefficient (0.900) between $R_{kav}$ and mortality index shows that the selected set of attributes could be used for estimation of sediment status with respect to chronic toxicity. Again two distinct separate groups involving level 1 and 2 in the one and level 3 and 4 in the other are observed.

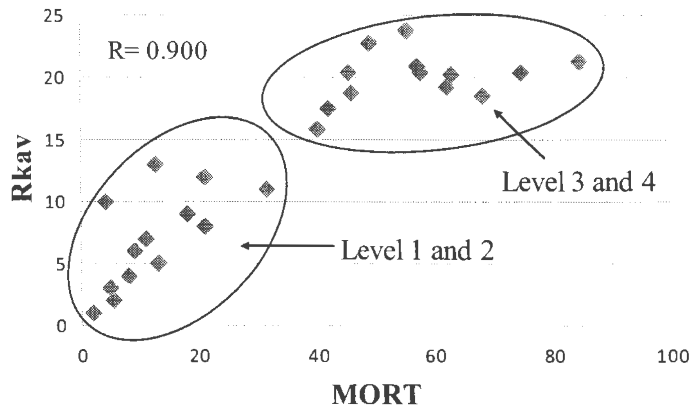



Figure 9. Comparison between $R_{kav}$ and mortality index

## 3. Conclusions

The present study has demonstrated a new efficient approach for estimation of sediment chronic toxicity status by four groups of chemical pollutants: PCB congeners, pesticides, PAHs, and heavy metals. The Hasse diagram obtained after SOM classification ensures additional information about polluting priority of the chemicals, relationships between ecotoxicity and chemical pollutants, which gives formation of two polluting patterns and sequence of sediment locations links.

Finally, the approach suggested could be used for investigated region management with respect to the future monitoring and remediation activities.

## References

Brüggemann R., Halfon E., Bücherl C. (1995) Theoretical base of the program "Hasse", *GSF-Bericht 20/95, Neuherberg.* The software may be obtained by contacting Dr. R. Brüggemann.

Brüggemann R., Voigt K, Halfon, E. (1996) Partially ordered sets - A computerized tool to compare environmental databases. In: Environmental Software Systems, R. Denzer, G. Schimak and D. Russell (Eds.) *Proceedings of the International Symposium on Environmental Software Systems*, Chapman & Hall, 185-195.

Brüggemann R., Halfon E., Welzl G., Voigt K., Steinberg C. (2001) Evaluation of test battery results for the toxicity of chemicals in sediments of Lake Ontario. *J. Chemical Information and Computer Science*, 41, 918-925.

Brüggemann R., Welzl G., Voigt K. (2003) Order theoretical tools for the evaluation of complex regional pollution patterns. *J. Chemical Information and Computer Science*. 43, 1771-1779.

Chakraborty B., Kaustubha R., Hedge A. and Pereira A. (2001) Acoustic seafloor sediment classification using self-organizing feature maps. *IEEE Trans. Geosci. Rem. Sens.*, 39, 2722-2725.

Eddins S., Tsakovski S., Simeonov V., Williams D.F. (1999) Evaluation of bottom sediment data from the Winyah Bay Estuarine system. *Tox. Envir. Chem.,* 70, 461-472.

Halfon E. and Brüggemann R. (1998) On ranking chemicals for environmental hazard. Comparison of methodologies. *Proceedings of the workshop on Order Theoretical Tools in Environmental Sciences*, Berlin, 11-48.

Karadjov M. and Simeonov V. (1990) Interpretation of geochemical data by the use of multivariate data analysis. *Mikrochim. Acta,* III, 191-199.

Kohonen T. (2001) *Self-Organizing Maps*. Third ed. Springer, Berlin.

Lacassie J.P., Roser B., Del Solar J.R., Herve F. (2004) Discovery geochemical patterns using self-organizing neural networks: a new perspective for sedimentary provenance analysis. *Sedimentary Geology*, 165, 175-191.

MacDonald D., Ingersoll C., Berger T. (2000) Development and Evaluation of Consensus-Based Sediment Quality Guidelines for Freshwater Ecosystems, *Arch. Environ. Contam. Toxicol.*, 39, 20-39.

Ostracodtoxkit F. (2005) Chronic direct contact Toxicity Test for Freshwater Sediments. *Standard Operational Procedure*, MicroBioTests Inc., Belgium.

Simeonov V., Wolska L., Kuczynska A., Gurwin J., Tsakovski S., Protasowicki M., Namiesnik J. (2007) Sediment quality assessment by intelligent data analysis. *Trends Anal. Chem.*, 26, 323-331.

Spanos Th., Simeonov V., Andreev G. (2002) Environmetric modeling of emission sources for dry and wet precipitation from an Urban Area. *Talanta*, 58, 367-375.

Stanimirova I., Tsakovski S., Simeonov V. (1999) Multivariate statistical analysis of coastal sediment data. *Fresenius J. Anal.Chem.*, 365, 480-488.

Swartz R. (1999) Consensus Sediment Quality Guidelines for Polycyclic Aromatic Hydrocarbon Mixtures. *Environ. Toxicol. Chem.* 18, 780-787.

Tsakovski S., Kudlak B, Simeonov V., Wolska L., Namiesnik J. (submitted). *Anal. Chim. Acta.*

U.S. EPA (United States Environmental Protection Agency). (2005) Predicting toxicity to amphipods from sediment chemistry. National Center for Environmental Assessment, Washington, DC; EPA/600/R-04/030. Available from: National Technical Information Service, Springfield, VA, and online at http://www.epa.gov/ncea.

Vesanto J., Himberg J., Alhoniemi E., Parhankagas J. (2000) SOM Toolbox for Matlab 5, *Report A57*, http://www.cis.hut.fi/projects/somtoolbox/.

Voigt K., Welzl G., Brüggemann R. (2004) Data analysis of environmental air pollutant monitoring systems in Europe. *Environmetrics*, 15, 577-596.

This book is a collection of papers, prepared in connection with the 8th International Workshop on partial orders, their theoretical and applied developments, which took place in Warsaw, at the Systems Research Institute, in October 2008. The papers deal with software developments (PYHASSE and other existing software), theoretical problems of ranking and ordering under various assumed analytic and decision-making-oriented conditions, as well as experimental studies and down-to-earth pragmatic questions.

ISBN 83-894-7521-9
EAN 9788389475213