Coupling empirical and deterministic models to assess surface water contamination

M.Terrado\textsuperscript{a}, M.P. Lavigne\textsuperscript{b}, D. Barceló\textsuperscript{a}, S. Duchesne\textsuperscript{b}, J.P. Villeneuve\textsuperscript{b}, A.N. Rousseau\textsuperscript{b}, S. Tremblay\textsuperscript{b} and R. Tauler\textsuperscript{a}

\textsuperscript{a} Department of Environmental Chemistry, IIQAB-CSIC. Jordi Girona 18-26, 08034 Barcelona, Catalonia, Spain (mtcqam/dbcqam/rtaqam@iiqab.csic.es)

\textsuperscript{b} Institut National de la Recherche Scientifique, INRS-ETE. 490, rue de la Couronne, Québec, G1K 9A9, Canada (martin-pierre_lavigne/sophie.duchesne/jean-pierre_villeneuve/alain_rousseau/sebastien.tremblay@ete.inrs.ca)

Abstract: Integration of chemometrics and deterministic modelling is proposed as a source for extracting useful information from data sets obtained in environmental monitoring studies. Organic contamination in the Llobregat river basin (Catalonia, NE Spain) was investigated for the period 2003-06. From the application of Multivariate Curve Resolution using Alternating Least Squares (MCR-ALS), 5 organic contamination sources were identified, explaining 87\% of the total data variance. Pollution from chlorinated pesticides and alkylphenols showed a decrease in recent years, while contamination caused by PAHs had a rather constant distribution in space as well as in time. On the other hand, diffuse pollution from a contamination source described by contribution of terbutryn (triazine) and chlorpyrifos (organophosphate insecticide) presented an increase in 2006 with regard to 2005. Dynamics of this increasing identified pollution source was modelled using the system GIBSI. Knowledge about existing cultures, usages of the pesticides defining the contamination source and their application dates and quantities was necessary to be used as input for the model calibration. Therefore, the impact of existing agricultural practices over the basin can be simulated and compared with possible scenarios involving different practices. After watershed modelling, the effects of a proposed measure can be evaluated.

Keywords: chemometrics; deterministic modelling; hydrological modelling; contamination sources; pesticides

1. INTRODUCTION

The big challenge for the European water and environmental managers nowadays is the implementation of the Water Framework Directive (WFD). It is for this reason that new policy instruments have been introduced to protect and improve European waters. An ecological and holistic water status assessment approach, a river basin water planning and a strategy for elimination of pollution by dangerous substances are included among them. To achieve acceptance for environmental policies and successful implementation, trustworthy methods for the identification of contamination sources affecting the quality of surface water in a specific river basin are needed. In the same way, monitoring of the identified sources is also important in order to assess and control contamination distribution, since it allows establishing those parameters which mostly affect pollutants dispersion into the environment as well as calculating the expected effect of a proposed measure.

Chemometric data analysis methods [Massart et al., 1998] assume chemical concentrations measured in a particular sample as affected by multiple contributions coming from independent sources of different origin (industry, agriculture, urbanisation, nature, etc.). These sources are defined describing their chemical composition as well as their geographical and temporal distributions. Multivariate Curve Resolution using Alternating Least Squares (MCR-ALS) [Tauler, 1995] is a multivariate statistical method which allows
the transformation of complex data sets into a new perspective in which the more relevant information is made more obvious, decomposing the environmental data set in components describing the main sources of data variance. While chemometric methods are based on observation and analysis of experimental data, deterministic models are defined as numerical methods based on general laws or theoretical principles [Arheimer and Olsson, 2003]. Traditionally, scientific communities using both types of models have been clearly differentiated. Coupling of these two approaches is proposed so as to merge the information obtained from both methodologies, obtaining a better understanding of contamination problems affecting a river basin.

In this work, integration of the proposed methodologies was tested over a pilot study catchment: the Llobregat river basin (Catalonia, NE Spain). Draining an area of around 5000 km², the river is approximately 160 km long. It is born in Castellar de n’Hug (Berguedà) and goes by until reaching the Mediterranean Sea, in el Prat de Llobregat (Baix Llobregat) (see Figure 1). La Baells dam is located not very far from its source and regulates the river flow. Llobregat river has two main tributaries, Cardener and Anoia. The headwaters of Llobregat and Cardener are in a rather unpolluted area. However, the two river midwaters flow through a densely populated and industrialized region where salt and potash mines (in Cardona and Súria, respectively) exist. Agriculture is significant in this zone, mainly devoted to cereal farming, such as wheat, barley and fodder. Anoia headwaters are in an agricultural area. Downwaters flow through an industrialized area, and near the confluence with the Llobregat river the main activity is agriculture again (mainly vineyards), while industrial and urban contribution attains a lower importance. At its lower course, the river flows through one of the most populated areas in the Mediterranean region, waters receiving high inputs from industry and urbanisation. An intensive agricultural activity exists near the delta, with predominance of orchard and garden vegetables, such as lettuce, cabbage, green onions and artichokes.

2. METHODS

This section is divided in two parts. Part 2.1 describes empirical models applied in this work. Data pretreatment and analysis were performed under the Matlab 7.0 (2004; The Mathworks, MA, USA) computer and visualisation environment, using PLS Toolbox 3.5

Figure 1. Location map of the Llobregat river basin with identification of quality stations: LL005, Llobregat river at Guardiola de Berguedà; LL015, La Baells reservoir at Berga; LL020, Llobregat river at Olvan; LL030, Merlès stream at Santa Maria de Merlès; LL035, Llobregat river at Balsareny; LL040, Gavarr slack at Sant Feliu de Llobregat; LL045, Calders river at Navarcles; LL050, Or river at Sant Fruíts de Bages; LL055, Llobregat river at El Pont de Vilomara; LL065, Cardener river at Olius; LL075, Negre river at Clarana de Cardener; LL080, Cardener river at Cardona; LL085, Cardener river at Castellgalí; LL090, Llobregat river at Castellbell i el Vilar; LL095, Llobregat river at Olesa de Montserrat; LL100, Llobregat river at Abona; LL105, Anoia river at Jorba; LL110, Anoia river at Vilanova del Camí; LL120, Riudebitlles stream at Sant Sadurní d’Anoia; LL125, Avernó river at Subirats; LL130, Anoia river at Martorell; LL135, Rubí stream at Castellbisbal; LL140, Llobregat river at Sant Joan Despí; LL145, Llobregat river at Barcelona.
functions (Eigenvector research Ltd., Manson, WA, USA). Part 2.2 describes deterministic models applied to the case of study. The physiographic database of the distributed hydrological model was prepared using the software PHYSITEL 3.0 (2006; INRS-ETE, Québec). The software HYDROTTEL 2.6 (2002; INRS-ETE, Québec) allowed the hydrological calibration of the whole watershed. Eventually, GIBSI (Gestion Intégrée des Bassins versants à l’aide d’un Système Informatisé) (2007; INRS-ETE, Québec) was the integrated modelling system used for watershed management.

2.1 Empirical models

Structure of data and data pretreatment

From its extensive annual environmental monitoring program, the Catalan Water Agency (Agència Catalana de l’Aigua, ACA) is periodically obtaining large data sets with different parameters related to water quality. For the present study, different data matrices (one per year) were arranged with the concentration of 17 organic compounds in 24 sampling sites distributed all over the Llobregat river system for the period 2003-06 (see Figure 1 for sampling sites location). The four obtained data matrices had identical dimensions, variables and sampling sites coinciding every year. Analysed organic compounds (concentrations in ng/L) were triazines (atrazine, terbutryn and simazine), organophosphate plaguicides (chlorpyrifos and diazinon), chlorinated plaguicides (lindane, endosulphan I, endosulphan II, endosulphan sulphate and ΣHCHs), polyaromatic hydrocarbons (pyrene, fluoranthene and fenanthrene) and alklyphenols (nonylphenol etoxylate (e=0), nonylphenol etoxylate (e=1), nonylphenol etoxylate (e=2) and t-octylphenol). Compounds were selected on account of periodical data availability. Some of them, such as atrazine, chlorpyrifos, endosulfan, fluoranthene, HCHs, nonylphenol, octylphenol and simazine are substances proposed for regulation in the proposal of the prioritary substances list from the Water Framework Directive. Variables having too many values below the detection limit were excluded from the data set. For this reason, fenitrothion and fluorene, initially considered, were finally removed to perform the analysis. The four data matrices were concatenated one on the top of the other, keeping columns in common (variables) and giving an augmented data matrix $D_{aug}$, of dimensions 96 rows x 17 columns. Because variable magnitudes differed considerably, scaling of $D_{aug}$ was considered to be necessary. Analysis without scaling was also performed to allow comparison, although results will not be displayed here for brevity. Scaling of individual matrices is neither presented in this work since it has been shown to result in a loss of information about the temporal component when identifying contamination sources distribution [Terrado et al., 2007].

Chemometric analysis

The multivariate data analysis method MCR-ALS was applied to data. It has the advantage of decomposing the data matrix by more natural constraints than those used in other frequently applied methods like Principal Component Analysis (PCA) [Joliffe, 1986], and environmental results can be interpreted straightforwardly. The drawback of MCR-ALS method is that solutions may not be unique and caution should be taken to check the reliability of the obtained results [Tauler, 2001]. MCR-ALS decomposes the augmented data matrix $D_{aug}$ by a constrained ALS algorithm using two alternating least-squares steps. Matrix decomposition gives a matrix of loadings ($V^T$), with the chemical composition of the main patterns of data variance, a matrix of scores ($U_{aug}$), describing the geographical and temporal distributions of the identified patterns and finally, a matrix of error ($E_{aug}$) containing the percentage of data variance not explained by the model. Information about geographical and temporal distribution of contamination sources appears intermixed in the columns of the resolved augmented scores matrix, being therefore, not directly available. However, MCR-ALS solutions can be additionally constrained to follow a trilinear model or to recover the information in the two modes, geographical and temporal [Tauler et al., 1998]. When trilinearity is applied, the profiles in the three different modes - $U_{geo}$ (scores matrix containing geographical information), $U_{temp}$ (scores matrix containing temporal information) and $V^T$ - are directly recovered and can be compared to obtained profiles.
when using other methods based on a trilinear model, such as PARAFAC [Bro, 1997]. Constraints used for all data matrices in this study were non-negativity for the two modes (augmented scores mode and loadings in the second mode) and normalisation of the loadings matrix to equal length. Singular value decomposition was applied over data matrices when row-wise as well as column-wise augmentation was performed before applying any pretreatment on data. In row-wise augmentation, matrices were concatenated having rows in common (samples) while in column-wise augmentation, matrices had columns (variables) in common. This gave an estimation about how trilinear data matrices were. Since obtained results were quite similar, the application of the trilinearity constraint was considered to be suitable to resolve the main lines of data variation in a more simple way [Tauler et al., 1995, Felipe-Sotelo et al., 2006 and Peré-Trepat et al., 2007].

2.2 Deterministic models

Determination of the drainage structure of the basin

The complete drainage structure of the basin was obtained using PHYSITEL [Turcotte et al., 2001], a software program designed specifically to prepare the watershed database for HYDROTEL (the hydrological model used as a second step for the watershed modelling). The most important data file for the determination of the drainage structure was a digital elevation model (DEM) based on a square grid with a specified spatial resolution (in this study, 200 m). Using the DEM, the drainage path from each cell to one of the cells surrounding it was defined. It was necessary to identify the cell corresponding to the outlet of the basin in order to define the complete area delimitation. After this, definition of the river network was possible, considering cells draining an area greater than a specified threshold. Finally, the determination of relatively homogeneous hydrological units (RHHU) was achieved, consisting on the creation of sub-watersheds associated to each river section of the modelled hydrological network. The software also allowed calculation of the percentage of land uses and soil types existing over each RHHU. Thus, a land use layer in raster format was obtained from the Catalan Government for year 2002. As no suitable soil map was available for the Llobregat basin, it was digitized for the present work taking an existing map of Catalonia and the Balear Islands form the Institute of Catalan Studies (IEC) as a reference.

Simulation of hydrological processes over the basin

For the hydrological simulation, the physically-based, distributed, hydrological model HYDROTEL was used [Fortin et al., 1995]. It consists of 6 computational models which are run in a cascade at each time step: 1) estimation of meteorological variables, 2) snow accumulation and melt, 3) potential evapotranspiration, 4) vertical water budget, 5) flow on RHHUs and 6) channel routing. For every model, the software offers the choice of different algorithms. Hydrological, meteorological and physiographical data were needed. Daily flow measures at different stations over the river network were furnished by the Water Catalan Agency for the different hydrological years of the studied period. Measures of flow at the outlet of the basin (or close to it) were essential. Meteorological data were obtained from the Catalan Meteorological Service (Servei Meteorològic de Catalunya, SMC) also at a daily frequency. Parameters used in the model calibration were established adopting the works from Lavigne, 2006 and Turcotte et al., 2007 as reference, and taking into account the regional differences existing between them and the region considered in this work. Parameters were adapted as much as possible to the Llobregat river basin [Bathurst et al., 2006]. Although the model was able to simulate the effect of dams, they were not considered in this work for reasons of data requirements.

Model calibration for the Llobregat river basin was performed in two different stages. In the first stage, a calibration of the whole basin was done focussing in the final part of the river. Abrera hydrological station (see Figure 2.a) was selected to start the calibration for i) being located sufficiently away from dam effects (as said, these effects were not modelled in the present study); ii) having a complete meteorological and hydrological data register for the studied period of time; iii) being placed in an area affected by the different contamination
sources existing in the basin (industrial, agricultural and urban); and iv) being the closer station to the outlet over whom no flow diminution was observed. Martorell and Sant Joan Despí stations were found Abrera station downstream. Despite being closer to the outlet, they were not chosen for calibration due to an observed flow decrease probably caused by water collecting at these sites, which made hydrological simulation more difficult. Once the first stage was finished and calibration values were obtained for Abrera station, a second simulation was performed dividing the basin in 6 different sub-basins (groups of RHHUs) (see Figure 2.a) and using results from the previous calibration as initial calibration parameters. In this way, the basin is not considered as a whole, and values of the different parameters can be adjusted independently for each group of RHHUs. This offers a higher flexibility to the model which sometimes can be necessary in order to obtain a more accurate calibration. Data about 14 hydrological stations were used, each station being associated to a river section (see Figure 2.b).

![Figure 2. HYDROTEL calibration steps: a) Division of the Llobregat basin in UHHRs and sub-basins (groups of UHHRs). Spatial distribution of hydrological stations; b) Association of the different hydrological stations to a specific river section.](image)

HYDROTEL calibration for the Llobregat basin was based in the strategy proposed by Turcotte et al., 2003. Parameters being more often modified during calibration were the recession coefficient (RC), the 2nd and 3rd soil layers depth (Z2 and Z3 respectively) and the potential evapotranspiration multiplication factor (FETP). RC, Z2 and Z3 are calibration parameters controlling the vertical water budget module of stream flow forecasting over the basin while FETP controls potential evapotranspiration. RC regulates contribution of the 3rd soil layer to flow. It is associated with prolonged summer drought recession objectives. Z3 depth affects the water quantity of this layer, horizontal flow and vertical flow between this layer and Z2. FETP is related to variables such as soil humidity content and meteorological conditions. As its relative sensitivity is high for all the periods of the year, flow is rapidly affected by FETP modification. That makes FETP and Z3 suitable to regulate annual and summer flow volumes. Z2 has effects over flow simulation during the summer and fall periods (usually, the rainy season) [Lavigne, 2006].

Integrated modelling of the basin

The integrated modelling system for watershed management GIBSI (Gestion Intégrée des Bassins versants à l’aide d’un Système Informatisé) was used for the Llobregat river basin modelling. The system can either be used as a data management system or as an impact assessment tool to study the effect of management scenarios on surface water quality using
mathematical models [Villeneuve et al., 1998 and Quilbé et al., 2007]. It is comprised of a relational database management system (RDMS), physically-based simulation models (hydrology, soil erosion, chemical pollutants transport and water quality), management models (point source, diffuse source and reservoir management models) and a geographic information system (GIS). Evaluation is performed via scenarios comparison. Two types of data, spatial and alphanumerical, are needed to be integrated in GIBSI. Spatial data integration was done by means of the GIS GRASSLAND. In this study, raster and vector data layers with a 200 m resolution and UTM projection European Datum 1950 Zone 31N were used. Required layers were a DEM of the whole area, the basin contours, municipalities, land uses, pedology, slope, sub-basins, SSUs (which corresponded to UHHRs in HYDROTEL) and river sections. Digital cartography was obtained from the Catalan Government (Generalitat de Catalunya) web site. Alphanumerical information needed as input depended on the simulation models to run. The aim of the present work was the implementation of the pesticides transport model, for which turning of the hydrological and soil erosion models was required, since outputs of one model were used as inputs for the next one. The water quality model was not used in this work owing to a lack of data availability. The hydrological model integrated in GIBSI was HYDROTEL, which has already been described in the previous section. Using data calculated by the hydrological model for each specific SSU, the erosion model derived from the algorithms of the Universal Soil Loss Equation (USLE), calculates the sediment yield transported to the river system corresponding to every land use over a SSU. Values of the different parameters used for calibration were extracted from the literature [Bathurst et al., 2006, Rawls and Brakensiek, 1989, Wischmeier, 1978]. The chemical transport model is based on the nitrogen, phosphorus and pesticide transport algorithms of the Soil and Water Assessment Tool (SWAT) [Arnold et al., 1998]. Transport model is used to calculate the agricultural pollutant loads exported to the river flow. Nitrogen and phosphorus algorithms (PO module) were not used in this study since attention was only focussed on pesticides transport (PES module). Pesticides algorithms simulate application, degradation and transport of pesticides by water and eroded soil particles over each agricultural land use in every SSU of the basin. Information about products cultivated in the region of study was obtained from a Vegetal Defence Association (Associació de Defensa Vegetal del Baix Llobregat). As information about agricultural practices was not directly available, it was extracted from the fitosanitary warnings published by the Catalan Department of Agriculture, Alimentation and Rural Action (DARP). Pesticides application taxes were obtained from the vademecum of fitosanitary products, document collecting recommendations of suitable doses to use at each pesticide application.

3. RESULTS

3.1. Chemometric results

MCR-ALS application on the organic contamination data matrix identified 5 principal contamination sources existing in the Llobregat river basin for the period 2003-06. Together, these 5 identified sources explained 87.08% of the total variance of data (see Figure 3). The major part of data variance (66.40%) was explained by a contamination source with industrial origin describing PAHs contributions. Distribution of this contamination source was rather constant in space as well as in time (see Figure 5). Spatially, the outstanding sampling site in comparison to the rest of locations was LL085, emplaced in the Cardener river at Castellgalí (see Figure 1). It is located at the central part of the basin, downstream of Manresa, which is one of the most populated municipalities in the basin, having an industrial activity especially important in the sectors of metallurgy, chemistry and leather. Manresa has also the condition of being a radial centre for road links, a key factor for its important commercial activity. The second identified contamination source explained 15.77% of data variance (a significant difference with respect to the first described source). It was loaded by lindane and total HCHs (a sum of different compounds, in which lindane isomer has the major proportion) as well as some contribution of PAHs, a group of compounds which appeared already in the previous identified source of
contamination (see Figure 3). This pattern corresponded to a mixture of pollution with agricultural (for lindane presence) as well as industrial origin (due to PAHs presence). It was a contamination source showing a clear decrease in time, for which its contribution to samples in year 2006 was half the contribution presented in 2003 (see Figure 4). However, for all the studied years, the more affected sampling sites were LL135, located in Rubí stream at the down part of the basin, followed by LL100, in the Llobregat river at Abrera station (see Figure 1). In the region near the confluence of the Llobregat river with the Anoia river (the second important tributary by the left side), there exists an important industrial activity located in Martorell. A third contamination source was identified explaining 8.19% of the total data variance. It was a contamination originated from agricultural practices and described by contribution of the group of analysed endosulphants (endosulphan I, endosulphan II and endosulphan sulphate) (see Figure 3). The highest levels of this contamination in samples were detected in year 2004. It presented a reduction during recent years, attaining its lower contribution in 2006 (see Figure 4). For all the studied period, LL135 followed by LL105, were the locations showing the highest levels of this contamination pattern. LL105 was placed at the upper part of the Anoia river, in an agricultural area mainly devoted to cereal farming. After reaching this point, the Anoia river flows through a zone of vineyards before intersecting the Llobregat river. As site LL135 is found downstream the intersection (see Figure 1), pollutants transported by the Anoia river when passing through this vast agricultural zone, are transported reaching the Llobregat river and downstream. Fourth and fifth identified contamination sources explained a very similar amount of data variance (see Figure 3). A clearly agricultural source defined by terbutryn and chlorpyrifos contribution was identified, explaining 5.30% of data variance. This source showed a decrease in its contribution to samples from 2003 to 2005, but an important increase was observed in 2006 (see Figure 4). LL135 and LL145 sampling sites (close to the river mouth) were identified as the most affected locations by this source. Cultures having a major predominance in the final part of the basin are orchard and fruit trees. Fifth pollution source explained 5.04% of the variance of data and it was loaded by the group of alkylphenols. These compounds can act as endocrine disruptors and are indicators of a type of contamination coming from industry and urbanisation. A clear decrease of this contamination source was identified in recent years for the analysed sites in the Llobregat river basin (see Figure 4).

From the identified contamination sources obtained by chemometric analysis of data matrices, two were found to have an agricultural origin: one describing a contamination by endosulphants and the other due to terbutryn and chlorpyrifos application. This last contamination source was selected for subsequent modelling since terbutryn and chlorpyrifos were found at concentrations above the detection limit in a 20% of the analysed samples, while endosulphan concentrations were below the detection limit in the major part of the analysed samples. Besides, the contamination source defined by terbutryn and chlorpyrifos was the only one showing an increase in the recent years (from 2004 to 2006), so a major attention should be put on it.
Figure 3. Contamination sources of organic compounds identified in the Llobregat river basin. Analysed compounds: 1, atrazine; 2, terbutryn; 3, simazine; 4, chlorpyrifos; 5, diazinon; 6, lindane; 7, endosulphan I; 8, endosulphan II; 9, endosulphan sulphate; 10, sum of hexachlorocyclohexane; 11, pyrene; 12, fluoranthene; 13, fenanthrene; 14, nonylphenol etoxylate (e=0); 15, nonylphenol etoxylate (e=1); 16, nonylphenol etoxylate (e=2); 17, t-octylphenol

Figure 4. Spatial and temporal distribution of organic contamination sources identified in the Llobregat river basin. The 24 sampling sites are ordered in the X axis from north to south for every sampled year (2003 to 2006).
3.2. Deterministic results

As explained in the previous section, Abrera station was focussed for calibration of the hydrology model HYDROTEL. Total observed flow in Abrera for all the analysed period, \(1.52 \times 10^9\) m\(^3\), was very close to the simulated flow by the software HYDROTEL, \(1.55 \times 10^9\) m\(^3\). Obtained values of Nash-Sutcliffe coefficient and Root Mean Square Error (RMSE) were 0.61 and 0.125, respectively. These coefficients gave an idea of how good the simulation was. Nash-Sutcliffe coefficient can vary between one and minus infinite. In general, the closer it is to one, the better the simulation, since simulated flow is more similar to measured flow. For RMSE, simulation is better when its value tends to zero.

Figure 5 displays the results from the hydrologic simulation for the station located in Abrera. The area graphic in the first plot (right axis) represented daily temperatures. In this case, temperatures followed a similar pattern (shape) during the 4 years of study. The 0ºC temperature was indicated in the graphic by a horizontal dark line. On the left, daily precipitation (mm) as well as snow depth on soil were represented by a graphic of bars. Snow was only calculated for those days in which temperature fell below zero and a rain episode took place. However, in the case of the Llobregat basin, few days with this characteristics existed, being therefore, a parameter having a more limited importance. Temperature and precipitation corresponded to average values calculated over the set of RHHUs being placed upstream the simulated station, once interpolation between stations was performed. The central plot in Figure 5 showed the quantity of water available for runoff in the area over the station where flow was simulated. Finally, the plot at the bottom of Figure 5 was a hydrogram representing observed and simulated flows (using a dark area and a light line, respectively) in the river section corresponding to Abrera station. In this work, high flows originated by rain episodes occurring in spring and fall were clearly observed. There was a good synchronisation of high flows and a correct base flow simulation. On the other side, in the case of consecutive rain episodes, simulation of the retarded flow (water abandoning the soil slowly after rain) was not accurate enough. Retarded flow is related to the recession coefficient, a parameter which was difficult to calibrate for the case of study (sometimes simulated flow was higher than observed flow and other times it was lower). Thus, finally, a balance had to be achieved.

Application of the transport model was done for the case of a diffuse agricultural contamination source identified in the Llobregat river basin. As the source was loaded by a similar contribution of terbutryn and chlorpyrifos (see Figure 3), it was introduced to the data base of the GIBSI system as a compound X, integrating the properties of the different compounds defining the source. Thus, compound X had a sorption and a half-life time being the average of the two compounds.
ACKNOWLEDGEMENTS

This work is supported by the European Union FP6 Integrated Project MODELKEY (SSPI-CT-2003-511237-2), the Spanish Ministry of Education and Science (CTQ2006-15052) and the Catalan Government (2007BE-100162).

REFERENCES

- Arheimer, B., and J. Olsson, Integration and coupling of hydrological models with water quality models: applications in Europe, Swedish Meteorological and Hydrological Institute, Sweden, 2003
- Bro, R., PARAFAC. Tutorial and applications, Chemometrics and Intelligent Laboratory Systems, 38, 141-171, 1997
- Fortin, J.P., R. Moussa, C. Bocquillon, J.P. Villeneuve, Hydrotel, a distributed hydrological model compatible with remote sensing and geographical information systems, Revue des Sciences de l’Eau, 8, 97-124, 1995
- Joliffe, I.T., Principal Component Analysis, Springer, New York, 1986
- Peré-Trepat E., A. Ginebreda, R. Tauler, Comparison of different multiway methods for the analysis of geographical metal distributions in fish, sediments and river waters in Catalonia, Chemometrics and Intelligent Laboratory Systems 88, 69-83, 2007
- Tauler, R., Multivariate Curve Resolution applied to second order data, Chemometrics and Intelligent Laboratory Systems, 30, 133-146, 1995.
- Tauler, R., A. Smilde, B. Kowalski, Selectivity, local rank, 3-way data-analysis and ambiguity in multivariate curve resolution, Journal of Chemometrics, 9, 31-58, 1995
- Tauler, R., I. Marques, E. Casassas, Multivariate curve resolution applied to three-way trilinear data: study of spectrofluorimetric acid-base titration of salicylic acid at three excitation wavelengths, Journal of Chemometrics, 12, 55-75, 1998
- Tauler, R., Calculation of maximum and minimum band boundaries of feasible solutions for species profiles obtained by multivariate curve resolution, Journal of Chemometrics, 15, 627-646, 2001
- Terrado, M., M. Kuster, D. Raldúa, M.L. de Alda, D. Barceló, and R. Tauler, Use of chemometric and geostatistical methods to evaluate pesticide pollution in the irrigation and drainage channels of the Ebro river delta during the rice-growing season, Analytical and Bioanalytical Chemistry, 387, 1479-1488, 2007