Using Data Mining to Predict Soil Quality after Application of Biosolids in Agriculture

Jérôme Cortet,* Dragi Kocev, Caroline Ducobu, Sašo Džeroski, Marko Debeljak, and Christophe Schwartz

The amount of biosolids recycled in agriculture has steadily increased during the last decades. However, few models are available to predict the accompanying risks, mainly due to the presence of trace element and organic contaminants, and benefits for soil fertility of their application. This paper deals with using data mining to assess the benefits and risks of biosolids application in agriculture. The analyzed data come from a 10-yr field experiment in northeast France focusing on the effects of biosolid application and mineral fertilization on soil fertility and contamination. Biosolids were applied at agriculturally recommended rates. Biosolids had a significant effect on soil fertility, causing in particular a persistent increase in plant-available phosphorus (P) relative to plots receiving mineral fertilizer. However, soil fertility at seeding and crop management method had greater effects than biosolid application on soil fertility at harvest, especially soil nitrogen (N) content. Levels of trace elements and organic contaminants in soils remained below legal threshold values. Levels of extractable metals correlated more strongly than total metal levels with other factors. Levels of organic contaminants, particularly polycyclic aromatic hydrocarbons, were linked to total metal levels in biosolids and treated soil. This study confirmed that biosolid application at rates recommended for agriculture is a safe option for increasing soil fertility. However, the quality of the biosolids selected has to be taken into account. The results also indicate the power of data mining in examining links between parameters in complex data sets.

J. Environ. Qual. 40:1972–1982 (2011)
doi:10.2134/jeq2011.0155
Posted online 5 Oct. 2011.
Supplemental data file is available online for this article.
Received 19 Apr. 2011.
*Corresponding author (jerome.cortet@ensaia.inpl-nancy.fr).
© ASA, CSSA, SSSA
5585 Guilford Rd., Madison, WI 53711 USA

DESPITE A VAST AMOUNT of scientific literature on biosolids application in agriculture, and successful modeling of related soil fertility aspects (Gabrielle et al., 2005), few models are available to predict the accompanying risks and benefits (mostly urban sludges and compost derivates). Such tools are urgently needed due to the increase in the amount of biosolids recycled in agriculture in recent decades. There has been a gradual strengthening of related legislation in the European Union, particularly in France, in the last 15 yr. During that period, the quantity of sludge has steadily increased; from 800 kt in 2000 to 950 kt (dry wt.) today. In France, the three main methods of sludge disposal are incineration, land-filling, and most important, agricultural land application (60%).

Soil quality has been defined as "the capacity of a specific kind of soil to function, within natural or managed ecosystem boundaries, to sustain plant and animal productivity, maintain or enhance water and air quality, and support human health and habitation" (Karlen et al., 1997). In agricultural contexts, soil fertility and contamination thus appear to be key parameters when surveying soil quality. Numerous case studies have explored the potential positive and/or negative effects of biosolid application on soil quality. Most of the benefits concern soil fertility. Biosolids in general, and sludge in particular, are rich in nutrients, calcium, and organic matter, and their application is known to have more or less long-lasting, favorable effects on both biomass production and chemical and physical soil properties. These effects have been the subject of a considerable number of studies since the late 1970s (e.g., Mitchell et al., 1978; Wei et al., 1985; Logan et al., 1997; Stamatiadis et al., 1999; Aggelides and Londra, 2000; Al-Assiuty et al. (2000)). More recent investigations include those described by Mantovi et al. (2005), Oliver et al. (2005), Oleszczuk (2006), and Brazauskiene et al. (2008). These studies showed that spreading sludge improves the structural properties and fertility of soil and increases its permeability, hydraulic conductivity, and water retention capacity. Moreover, sludge decreases the rate of soil erosion. Sludge spreading also has a favorable effect on soil biological characteristics by stimulating microbial activity and

J. Cortet, C. Ducobu, and C. Schwartz, Nancy Université, UMR INPL/INRA 1120, Laboratoire Sols et Environnement, 2 av. de la Forêt de Haye, 54505 Vandœuvre-lès-Nancy, France; D. Kocev, S. Džeroski, and M. Debeljak, Jozef Stefan Institute, Dep. of Knowledge Technologies, Jamova 39, 1000 Ljubljana, Slovenia. Assigned to Associate Editor César Plaza.

Abbreviations: CA, coal ash; CEC, cation exchange capacity of soil; CMF, control with minimal fertilization; COF, control with optimal fertilization; HWA, household waste ash; LDCSS, lightly dehydrated composted sewage sludge; LDCSSM, lightly dehydrated composted sludge with added metals; LDCSSO, lightly dehydrated composted sludge with added organic pollutants; LDSS, lightly dehydrated sewage sludge; LSS, liquid sewage sludge; MPS, mixed paper sludge; NF, Normes Françaises; PAH, polycyclic aromatic hydrocarbon; PCB, polychlorinated biphenyl.

biomass (Mitchell et al., 1978; Robert, 1996; Banerjee et al., 1997; Criquet et al., 2007).

The main risks concerning biosolids application are associated with soil pollution, particularly the transfer of mineral and organic pollutants within the agrosystem. It has been previously demonstrated that with high amounts of sludge application, possible transfers of pollutants, mainly trace elements, from soil to plants are possible (Juste and Solda, 1977; Morel and Guckert, 1984; Juste and Mench, 1992). Furthermore, concerns about the behavior of organic pollutant, particularly PAHs and polychlorinated biphenyls (PCBs), coming from sewage sludge have been raised (Rogers, 1996; Petersen et al., 2003; Oleszczuk, 2006). These investigations were recently summarized in several reviews showing that sewage sludge used in recommended doses could usually increase fertility and yield, whereas the bioavailability of metals increases only in sludge-amended soil at excessive rates of application over many years (Hargreaves et al., 2008; Singh and Agrawal, 2008; Smith, 2009).

Long-term experiments on the effects of biosolids on soil fertility and contamination are sometimes difficult to interpret, however, because they are, by definition, multifactorial. For instance, it has been shown that the transfer of metals to plants can vary with soil properties, particularly pH or texture (Morel and Guckert, 1984; Juste and Mench, 1992; Brazauskiene et al., 2008; Smith, 2009). Also, classic mineral fertilization is known to be a considerable source of metals. Furthermore, some other factors can modify the results of such studies, such as crop type or agricultural practices (Oleszczuk, 2006). Indeed, even though European and more particularly French regulations have sought to standardize the quality of biosolids for spreading, their nutrient content can change over time (Oliver et al., 2005). Additional factors of differentiation include the application frequency and protocols used. Therefore, the divergence in results obtained from numerous studies underline the complexity of biosolid–soil–plant system relationships. Despite the challenge of obtaining comparable data and given the context of current regulation, stakeholders urgently need methods to assess the effects of biosolid application.

Consequently, we have followed an alternative approach to considering soil responses to the addition of biosolids. Our aim is to propose an integrative method to assess the benefits and risks of biosolid application in agriculture, using data mining on results from a 10-yr field experiment in the northeast of France (1997–2006). The experiment includes four biosolid applications, eight types of biosolids, and several successive crops. We propose here to identify the main factors, in terms of soil and biosolid quality, as well as of agricultural practices, affecting the soil–plant system to be able to predict (i) soil fertility and (ii) soil contamination by heavy metals and organic pollutants. Our main hypothesis is that four applications of biosolids over a 10-yr period, which is very close to the recommended dose permitted by regulations (1.3 times more) (Voynet et al., 1998), can increase fertility without modifying soil contamination levels. For this purpose, we set out to analyze the dataset by using data mining. Data mining is the central step in the process of knowledge discovery in databases and is concerned with applying specific algorithms to find patterns in data. In other words, it takes data as input and outputs knowledge in the form of patterns (i.e., models). Environmental data have previously been successfully modeled using data mining (for an overview, see Debeljak and Dzeroski, 2009, 2011). The application of data mining to soil science covers a wide spectrum of ecological and environmental domains, including habitat modeling (Debeljak et al., 2008), the modeling of population dynamics (Demšar et al., 2006; Debeljak et al., 2011) and the modeling of soil quality (Debeljak et al., 2007; Debeljak et al., 2009). The knowledge obtained from data-mining models can be then combined with the expert knowledge to create a decision support system (e.g., using DEXi [Znidarsic et al., 2006]) for optimal management of soil quality (e.g., Bohanec et al., 2007).

Material and Methods

Experimental Setup and Data Collection

Experiments were performed on the experimental farm at La Bouzule (12 km east of Nancy, France). The climate is oceanic with continental influences. The principal agronomic characteristics of the soil (redoxic neoluvisoil) were described by Martin-Laurent et al. (2004).

Ten treatments were applied to experimental plots. The following six different biosolid types were applied: liquid sewage sludge (LSS), lightly dehydrated sewage sludge (LDSS), lightly dehydrated composted sewage sludge (LDCSS), lightly dehydrated composted sludge with added organic pollutants (LDCSSO), lightly dehydrated composted sludge with added metals (LDCSSM), and mixed paper sludge (MPS). Additionally, two other types of wastes were tested: coal ash (CA), and household waste ash (HWA). For convenience, these two other types of wastes will also be called biosolids throughout this paper. These eight biosolids are representative of materials available in France for application on agricultural soils (ADEME and Cabinet Arthur Andersen, 1999). The two control plots did not receive any biosolids but were subject to mineral fertilization. The first control (control with minimal fertilization [CMF]) received a minimal fertilization sufficient only to avoid plant nutrient deficiency, whereas the second control (control with optimal fertilization [COF]) received a higher and optimal quantitative fertilization, which corresponds to the doses usually applied in agriculture and allowing optimal yield (Supplemental Table S1). A split block and plot design was created, including four plots (10 by 4 m) per treatment. Thus, a total of 40 plots, including the 10 treatments, were performed. The principal biosolid characteristics are summarized in Table 1. All types of biosolids, except LDCSSO for benzo(a)pyrene and fluoranthene, met regulatory compliance needs concerning their fertilizing qualities and pollutant concentrations. The mean of LDCSSM trace elements met the regulation, but the cadmium (Cd) concentration was higher than the legal threshold value on two occasions. Each type of biosolid was applied manually to each plot according to regulations (four applications of 10 t ha^{-1} within 10 yr, representing 160 kg of dry matter for each plot). For all the plots, biosolids were complemented with a mineral fertilization corresponding to the level of COF treatment. This was to avoid nutrient deficiency for plants and to conform to current agricultural practices combining biosolid application and mineral fertilization.

A chronological sequence of agricultural practices and soil sampling is described in Supplemental Table S1. Agricultural practices included tillage (0–20 cm) to incorporate biosolids, followed by rotovator cultivation to prepare sowing. The crop rotation included spring rape (*Brassica napus* L. var. *napus*) ('Jaguar', 5 kg ha^{-1}), winter rape ('Amber', 6 kg ha^{-1}), winter wheat (*Triticum aestivum* L.) ('Texel', 190 kg ha^{-1}) and maize (*Zea mays* L.) ('Anjou 258' variety, 100 000 seeds ha^{-1}).

Soil samples were usually collected twice a year, before sowing and biosolid application, and at harvest time. At each sampling date, from T0 to TXIII (Supplemental Table S1), soil agronomic parameters corresponding to the tillage horizon (0–20 cm) were assessed. All the analytical work was done by the national laboratory for soil analysis of the Institut National de la Recherche Agronomique (INRA). The laboratory is certified by the French Ministry of Agriculture for soil characterization. This laboratory is working under AQ ISO 17025 and COFRAC and is a member of the International Soil-analytical Exchange Program (Wageningen, the Netherlands) (16 samples per year). Characterization of soils was performed with reference to Normes Françaises (NF) standards. Quality control for soils was based on the use of certified soil samples (GBW 07401, 07402, 07404, 07405, and 07406), samples from interlaboratory comparisons, internal control samples, and duplicates of the analysis.

Analyses included pH (NF ISO 14254), cation exchange capacity (CEC) (NF ISO 11260), organic carbon (C_{org}) (NF ISO 14235), total organic matter, total nitrogen (NF ISO 11261), available phosphorus (P_{Olsen}) (NF ISO 11263), and C/N (AFNOR, 2004).

Trace elements (Cd, Cu, Ni, Pb, and Zn) were analyzed by ICP–AES (NF EN ISO 11885) after mineralization (NF ISO 1466, 1995 and NF X 31-151, 1993). Total PAHs and total PCBs were analyzed following the standardized protocol XP-X 33-012, 2000 and NF ISO 13877, 1999. Furthermore, for dates from TVII to TXII, extractable $CaCl_2$ trace elements were analyzed to assess the availability and mobility of the contaminants (Lebourg et al., 1996), but measurements for each treatment were obtained from a pooled sample from the 4 plots.

Data Analysis

Descriptive Statistics

Differences between treatments were assessed for each sampling date for each fertility (pH, C_{org}, total N, P_{Olsen}, and C/N) and for each total trace element (total Cd, Cu, Ni, Pb, Zn) parameter, using one-way ANOVA followed by a Tukey HSD pairwise test. No statistics were calculated for organic pollutants and extractible $CaCl_2$ trace elements, as measurements for each treatment were obtained from a pooled sample from the four plots. All statistical calculations were performed using R software (Ihaka and Gentleman, 1996).

Scenarios and Dataset

To examine the effects of biosolid application on soil fertility and to see if these lead to the contamination of the soil with pollutants, we considered four different scenarios (i.e., data analysis setups). The general goal of this data analysis was to predict soil properties at harvest time from variables describing soil properties before sowing, biosolid properties, mineral fertilization, and management practices. Scenario 1 aims to predict soil fertility parameters (C_{org}, total N, C/N, P_{Olsen}), Scenarios 2 and 3 aim respectively to predict total and $CaCl_2$

Table 1. Biosolid characteristics, trace elements and organic pollutants content. Mean of four values for each parameter (one for each spreading during the study). Values in italics are close to or above the threshold values.

	LSS†	LDSS	LDCSS	LDCSSO	LDCSSM	MPS	CA	HWA	Threshold values‡
Dry matter (DM), %	3.70	23.78	39.23	39.18	38.63	47.70	76.15	80.58	–
Organic matter, %	54.55	49.13	61.60	55.20	60.48	48.25	6.70	40.85	–
Organic carbon (C), %	31.60	27.50	33.28	30.65	32.75	27.35	5.63	22.40	–
Total nitrogen (N), %	5.95	3.85	1.95	1.65	1.93	0.83	0.06	1.28	–
C/N	5.30	7.18	17.38	23.23	18.93	34.30	107.4	17.80	–
pH_{water}	7.65	7.73	6.25	6.75	6.05	7.53	9.08	8.25	–
CaO, ‰	84.75	76.75	50.48	79.30	54.30	126.30	22.63	83.18	–
MgO, ‰	7.65	9.48	6.85	11.73	6.95	22.55	32.85	12.83	–
P_2O_5, ‰	72.25	67.33	34.20	26.75	35.20	5.00	2.63	6.25	–
K_2O, ‰	7.90	7.70	8.15	7.10	8.95	4.10	35.30	12.50	–
Cd, mg kg^{-1} DM	1.25	3.28	2.20	3.37	*9.89*	0.23	0.59	1.81	10
Cu, mg kg^{-1} DM	495.75	450.25	232.50	230.75	495.25	167.75	95.50	218.75	1000
Ni, mg kg^{-1} DM	42.75	38.75	25.50	31.75	84.00	11.50	125.50	44.00	200
Pb, mg kg^{-1} DM	284.00	285.75	160.5	162.75	284.00	25.50	104.75	373.25	800
Zn, mg kg^{-1} DM	1271.25	1358.50	682.75	702.00	1455.75	148.25	249.75	1113.25	3000
Benzo(a)pyrene, mg kg^{-1} DM	0.26	0.35	0.77	*10.20*	0.15	0.05	0.01	0.04	2
Fluoranthene, mg kg^{-1} DM	1.78	0.98	2.48	*18.63*	0.44	0.01	0.01	0.08	5
Benzo(b)fluoranthene, mg kg^{-1} DM	0.35	0.35	0.08	0.32	0.18	0.58	0.01	0.16	2.5
7 PCB§ sum, mg kg^{-1} DM	0.16	0.16	0.11	0.07	0.12	0.15	0.01	0.20	0.8

† LSS, liquid sewage sludge; LDSS, lightly dehydrated sewage sludge; LDCSS, lightly dehydrated composted sewage sludge; LDCSSO, lightly dehydrated composted sludge added with organic pollutants; LDCSSM, lightly dehydrated composted sludge added with metals; MPS, mixed paper sludge; CA, coal ashes; HWA, household waste ashes.

‡ French Council decision of 8 Jan. 1998 (Voynet et al., 1998).

§ PCB, polychlorinated biphenyl.

extractible trace elements (Cd, Cu, Ni, Pb, Zn) in the soil, and Scenario 4 predicts organic pollutants concentrations (total PAH, total PCB) in the soil. These four scenarios are described in detail in Table 2.

For 2003, 2004, and 2005, soil data was collected just after the harvest of maize. To make predictions for Scenarios 1 and 4, the total dataset was used ($N = 360$). For Scenarios 2 and 3, extractable $CaCl_2$ trace elements had not been measured throughout the experiment, explaining the smaller dataset ($N = 200$).

Data Mining

The CLUS data-mining system (http://dtai.cs.kuleuven.be/clus/) was used to construct regression tree models (Blockeel and Struyf, 2002). Regression trees predict the value of a numeric target variable (Breiman et al., 1984). They have a hierarchical structure, where the internal nodes contain tests on the input attributes and the leaves predictions for the target variable. Tree construction starts with the complete set of data and recursively splits the data, selecting an attribute test at each step. The heuristic diagram used to select the attribute tests in the internal nodes is an intracluster variation summed over the subsets induced by the test. Intracluster variation is defined as $N \cdot Var(y)$, with N the number of examples in the cluster and $Var(y)$ the variance of target variable y in the cluster. Lower intrasubset variation results in predictions that are more accurate.

To improve the predictive performance and/or interpretability of the trees, trees can be pruned. In this work, we use three different pruning techniques: minimal number of instances in a leaf, maximal tree depth, and maximal size. The first two techniques are utilized during the construction of the tree, and the third technique is used after the tree is built. More details about these techniques can be found in Struyf and Džeroski (2006).

The construction of regression trees is a widely used modeling approach (Tan et al., 2006); regression trees enable easy understanding of results, where interpretation of the studied phenomena from a systemic point of view is possible. The method for the construction of regression trees is nonparametric (which does not require prior assumptions concerning the distribution probability of the predicted and other attributes). In addition, it is not computationally expensive even in the case of large datasets. Also, the process of tree construction can handle redundant attributes and noise effectively.

During the modeling process, we explored the aforementioned pruning settings to control the size and complexity of the constructed regression trees:

- minimum instances in a leaf (i.e., minimum number of samples from which the predicted value is calculated): 16, 24 or 32;
- maximum tree depth (i.e., number of hierarchical levels of the tree): 2, 3, or 4; and
- maximum size (i.e., maximum number of predictions [leaves]): 7, 9, or 11.

Additionally, these setting values were combined (i.e., minimum number of instances in a leaf was set to 32 and the maximal tree depth to 3) to find the regression trees with the best performance and the most interpretable structure. In this way, we obtained a set of models (i.e., regression trees) for each scenario.

The predictive performance of the models on unseen data was assessed according to the following three quantitative criteria estimated by 10-fold cross-validation: correlation coefficient, root mean squared error (RMSE) and relative root mean squared error (RRMSE). One model from each scenario, i.e., the best according to the criteria of size, predictive performance, and interpretability, was selected for further presentation and interpretation in this paper.

Results

For each soil characteristic (attribute), the results present (i) the significant differences between treatments for each sampling date and (ii) the main points describing the decision trees resulting from data mining.

Table 2. Predicted and independent attributes in each of the scenarios.

	Predicted attributes†	Independent attributes			
		Soil properties before sowing†	Biosolid properties‡	Mineral fertilizers§	Management and temporal aspects of farming practices¶
Scenario 1 Soil fertility	C_{org}, N_{tot}, C/N, P_2O_5, in soil at harvest time	pH_{H2O} [-log10 H+], CEC, C_{org}, N_{tot}, C/N, P_2O_5	pH_{H2O} [-log10 H+], C_{org}, N_{NH4+}, N_{tot}, C/N, P_2O_5, CaO, MgO, K_2O	N, P, K	Time since last biosolid spreading; time since last mineral fertilization; number of previous biosolid spreadings; type of the current crop, previous crop (type of the crop 1 yr before and type of the crop 2 yr before)
Scenario 2 Trace elements	Total Cd, Cu, Ni, Pb, Zn in soil at harvest time	pH_{H2O} [-log10 H+], CEC, C_{org}, N_{tot}, C/N, P_2O_5, Total and $CaCl_2$ extractible Cd, Cu, Ni, Pb, Zn	pH_{H2O} [-log10 H+], C_{org}, dry matter, organic matter, total Cd, Cu, Ni, Pb, Zn, PAH, PCB	N, P, K, Total Cd, Cr, Cu, Zn, Ni	
Scenario 3 Extractible trace elements	$CaCl_2$ extractible Cd, Cu, Ni, Pb, Zn in soil at harvest time	pH_{H2O} [-log10 H+], CEC, C_{org}, N_{tot}, C/N, P_2O_5, Total and $CaCl_2$ extractible Cd, Cu, Ni, Pb, Zn			
Scenario 4 Organic pollutants	PAH and PCB in soil at harvest time	pH_{H2O} [-log10 H+], CEC, C_{org}, N_{tot}, C/N, P_2O_5			

† C_{org} (organic carbon), N_{tot} (total nitrogen), P_2O_5, CaO, MgO, and K_2O in soils are in ‰; cation exchange capacity (CEC) is in cmol kg^{-1}; Cd, Cu, Ni, Pb, Zn, PAH (polycyclic aromatic hydrocarbon), and PCB (polychlorinated biphenyl) are in mg kg^{-1} of dry matter.

‡ C_{org}, N_{tot}, dry matter, and organic matter in biosolids are in %; CEC is in cmol kg^{-1}; P_2O_5, CaO, MgO, K_2O are in ‰; Cd, Cu, Ni, Pb, Zn are in mg kg^{-1}.

§ N, P, K are in unit ha^{-1}; Cd, Cr, Cu, Zn, Ni are in g ha^{-1}.

¶ Time is in months.

Soil Fertility (Scenario 1)

The data for C_{org}, total N, and P_{Olsen} are presented in Supplemental Table S2. At several sampling dates, C_{org} was significantly lower in COF and CMF compared with LDCSSM, LDCSSO, LDSS, LSS, and MPS ($p < 0.05$), yet at the end of the study (date TXIII), no significant difference could be observed between treatments. Significant increases in total N and available P were observed after the first three LSS and LDSS applications. When applications ceased, N concentrations tended to decrease, but not available P, which had accumulated for these two kinds of biosolids. Nitrogen concentration was significantly greater in LSS than in COF for dates TIII, TV, TVII, TVIII, TIX, TX, TXI, and TXII ($p < 0.05$), yet at the end of the experiment, the same level of N was observed for both treatments. In the case of $P_{Olsen,}$ significant differences occurred between treatments throughout the study, usually with greater LSS and LDSS concentrations than for COF. However, at the end of the experiment, P concentrations were only greater in LDSS than in CA.

The results of the data-mining models (Fig. 1) indicate that soil fertility parameters (C_{org}, total N, C/N, and P_{Olsen}) after harvest depend mostly on three types of variables: (i) soil fertility before sowing, (ii) characteristics of the biosolids, the number and period of biosolid applications, and (iii) previous crops (Fig. 1). Organic C in the soil after harvest is generally positively correlated with the initial C_{org} before sowing. To a lesser extent, an increase in P_{Olsen} in soil before sowing is also correlated with an increase in soil C_{org} after harvest (Fig. 1, top left, $r = 0.65$). Soil N after harvest is positively correlated to the initial N content in soil before sowing but also to the date of last mineral fertilization. In cases where fertilization occurred 6 mo previously, the N content in soil after harvest was lower than when fertilization occurred four and 5 mo before harvest (Fig. 1, top-right, $r = 0.71$). Soil C/N model predictions showed a shift corresponding to an initial modification of land use (i.e., a shift from grassland to arable crop). Thus, the C/N ratio was negatively affected when the soil 2 yr prior to harvest was under grass. Then, the C/N ratio was positively affected (i) by the number of biosolid applications, reaching its maximum after four applications, and (ii) by the quality of biosolids, as the C/N ratio was positively correlated with the C/N ratio in the biosolids (Fig. 1, bottom-left, $r = 0.81$). Phosphorus content after harvest was clearly positively correlated with the P_{Olsen} content before sowing. Furthermore, soil P_{Olsen} was slightly positively correlated with the biosolid CaO content. It also tended to decrease when the last biosolid spreading occurred at least 12 mo before (Fig. 1, bottom-right, $r = 0.89$).

Trace Elements

Total Metals (Scenario 2)

The data on the total metal contents measured in the soil are presented in Supplemental Table S3. At the end of the experiment, all the metals except Cd and Zn were at the same level as in the initial samples. A slight increase in Zn (20%) was observed, whereas Cd more than tripled during the period.

Significant differences between treatments appeared during the field experiment. For Cd, these differences fluctuated throughout the study, but at the end, COF, CMF, and MPS con-

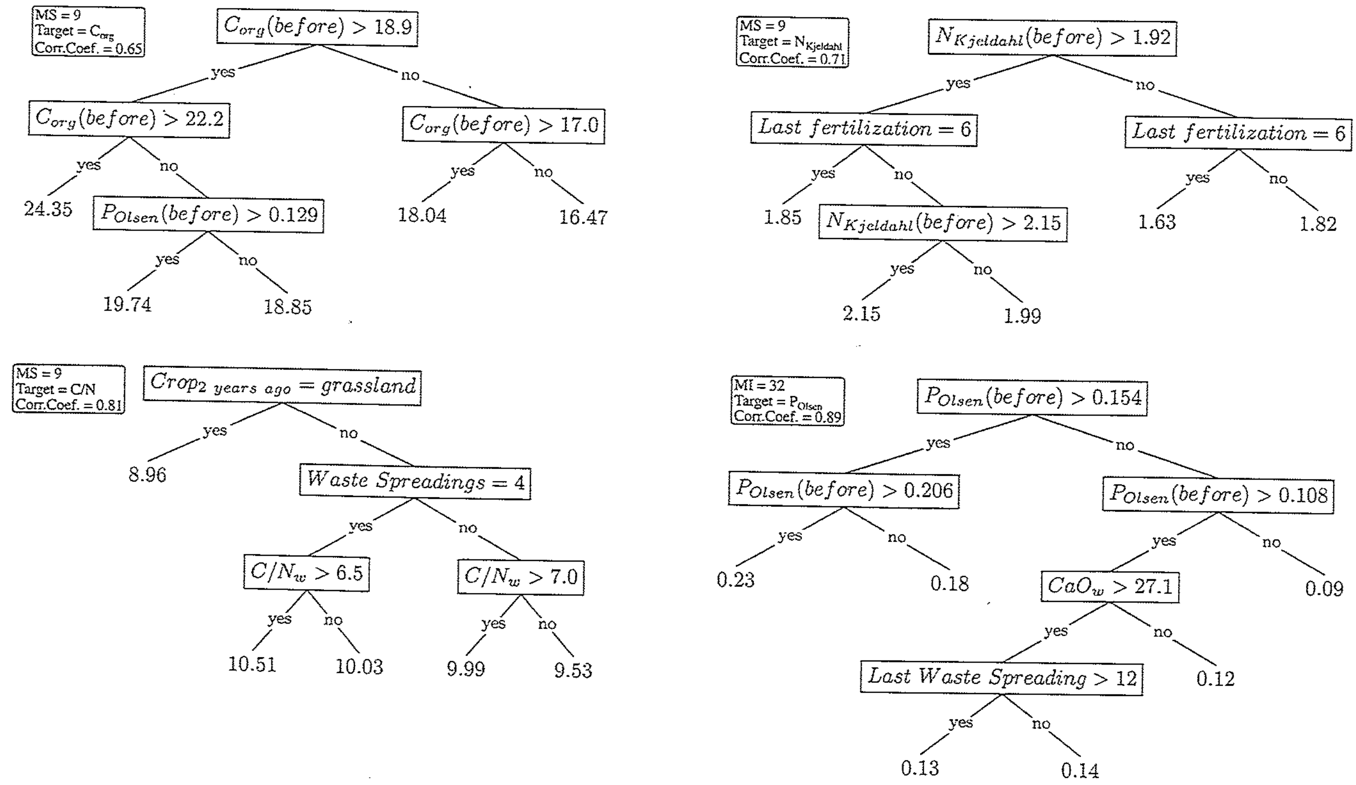

Fig. 1. Regression trees predicting values for soil fertility parameters at harvest time (Scenario 1) MS: maximal size of the constructed regression trees (i.e., maximum number of predictions [leaves]). MI: minimum instances in a leaf (i.e., minimum number of samples from which the predicted value is calculated). "(before)" = content in soil at sowing time; "w" = content in biosolids. Time is expressed in months. See Table 2 for further explanation of attributes. Trees should be read like "what-if" questions from the root of each tree. The closer the attribute is to the root of the tree, the more it is statistically influent for the considered predicted attribute.

tained significantly greater Cd than did LDSS and LDCSSO. Copper contents appeared significantly lower in control plots than in the other plots receiving treatments at dates TII, TV, TVI, TVII, TX, and TXI. At the end of the experiment, however, no significant difference could be detected between treatments. Nickel did not show any significant differences between treatments throughout the study. Lead showed significant differences during the study and remained greater in HWA than in LDCSSO, MPS, CA, CMF, and COF at the end of the study. Zinc appeared to be significantly lower in COF than after biosolid treatments at dates TIII, TVI, TXI, and TXII, but no significant difference could be detected between treatments at the end of the experiment.

An examination of the data-mining models (Fig. 2) shows that "total Cd" was mostly influenced by the attributes "previous crop," "time since last biosolid spreading," and the extractible Cd and Pb in the soil before sowing (Fig. 2, top left, r = 0.88). The highest values were obtained when "previous crop" was "not winter wheat" (e.g., when the crop was maize), when "last biosolid spreading" occurred at least 18 mo before harvest, and when "extractable Pb" in the soil before sowing was greater than 50 μg kg^{-1} of dry matter. The lowest values were observed if "previous crop" was "winter wheat," and "extractable Cd" in the soil was lower than 22.2 μg kg^{-1} of dry matter before sowing.

"Total Cu" was mainly influenced by the total Cu and Zn content in soil before sowing, and the attribute "previous crops" (Fig. 2, top right, r = 0.79). In fact, a positive correlation was found with "previous Cu content," but a negative one with "previous Zn content." Finally, the highest values were obtained with high Cu and moderate Zn contents in the soil before sowing and when "previous crop" was "not winter wheat." The lowest values of total Cu after harvest were obtained with low initial Cu content but high initial Zn contents in a soil before sowing.

"Total Ni" was mostly influenced by "previous crop," with the highest values occurring when maize was cultivated the previous year (Fig. 2, center left, r = 0.85). In this case, a higher pH of the biosolid tended to increase the total Ni in the soil. In the other case, a negative correlation was found with total Zn in the soil at the time of sowing.

"Total Pb" in soil after harvest positively increased when the "previous crop" was "maize" (Fig. 2, center right, r = 0.88). In this case, a negative correlation was found with "total Ni" in the soil before sowing. When the "previous crop" was "not maize," a negative correlation was observed with "total Cd" in the soil before sowing, but a positive correlation with "total Pb" in the biosolids. Finally the highest values were observed when "maize" was "previous crop" combined with a soil with low Ni content. The lowest values were obtained when the "previous crop" was "not maize", with a high Cd content soil.

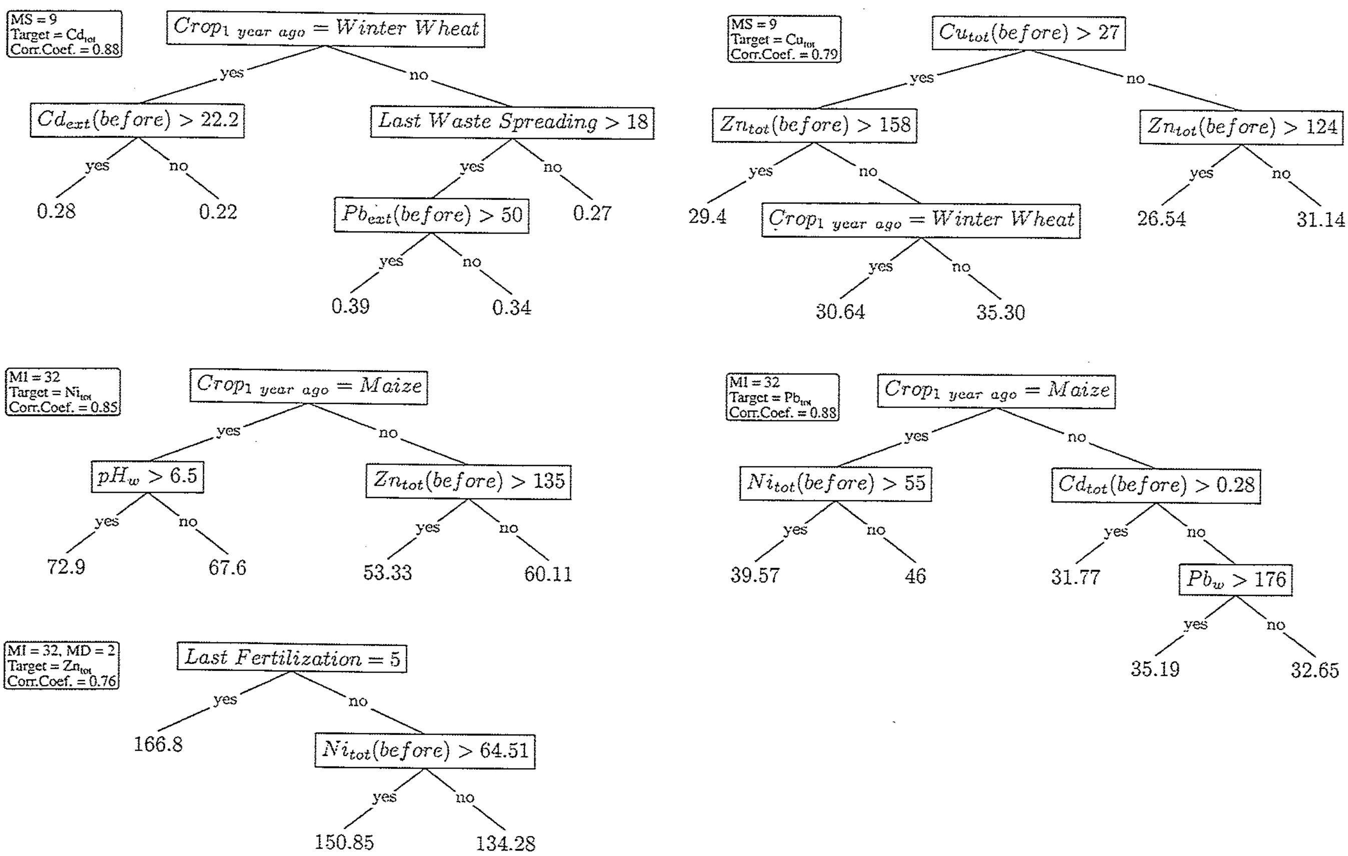

Fig. 2. Regression trees predicting values for total metals in soil (mg kg^{-1} dry matter [DM]) at harvest time (Scenario 2). MS: maximum size of the constructed regression trees (i.e., maximum number of predictions [leaves]); MI = minimum instances in a leaf (i.e., minimum number of samples from which the predicted value is calculated); MD = maximum tree depth (i.e., number of hierarchical levels of the tree); "(before)" = content in soil at sowing time; "w" = content in biosolids; "tot" = total metal content (mg kg^{-1} DM); "ext" = $CaCl_2$–extractible metal content (μg kg^{-1} DM). Time is expressed in months. See Table 2 for further explanation concerning attributes. Trees should be read like "what-if" questions from the root of each tree. The closer the attribute is to the root of the tree, the more it is statistically influent for the considered predicted attribute

"Total Zn" at harvest seemed to be mainly influenced by the "last mineral fertilization." The highest values were observed when this fertilization occurred 5 mo before harvest (Fig. 2, bottom left, $r = 0.76$). Where this was not the case, a positive correlation was observed with "total Ni before sowing."

Extractable Metals (Scenario 3)

The data on extractable metals are shown in Supplemental Table S4. The extractable metal concentrations were always much lower than those observed for total metal content (50 times), except for Cd, whose extractable fraction represented approximately 10% of the total metal content. No significant difference could be observed between treatments. However, soils with MPS tended to contain lower concentrations of Cd, Zn, and Ni, than those treated with other biosolids.

Examining the data-mining results (Fig. 3), "extractable soil Cd" seemed to be strongly positively correlated with "Cu biosolid content" (Fig. 3, top left, $r = 0.83$). In the case of high Cu content in biosolids, the highest values were obtained when the "previous crop" was "not wheat." In the case of low Cu content in biosolids, a negative correlation was observed with "initial total Zn" soil content before sowing and a positive correlation with "initial total Pb" content in soil.

"Extractable soil Cu" was negatively correlated with the "initial total soil Cd" before sowing (Fig. 3, center left, $r = 0.84$). In the case of high total Cd content in soil, the lowest extractable Cu concentrations were obtained when the "previous crop" was "winter wheat." In the case of low total Cd content in the soil, the highest values were observed with high organic C in biosolids.

"Extractable Ni" in the soil was strongly influenced by the attribute "previous crop," with lowest values when "wheat" was cultivated 1 yr before (Fig. 3, bottom left, $r = 0.80$). In this case, a positive correlation was observed with "extractable Ni in soil before sowing." In the case of "maize," the quality of biosolids influenced "extractable Ni," with high extractable soil Ni values correlated with low dry matter content in biosolids.

"Extractable Pb" in the soil at harvest was most increased when the "last mineral fertilization" was recent (4 mo) (Fig. 3, top right, $r = 0.88$). In this case, "extractable Pb" was negatively correlated with the "total Cd in the soil before sowing." Where this was not the case, "extractable Pb content" was negatively correlated with "initial extractable Ni before sowing" and positively correlated with "initial extractable Cu before sowing."

"Extractable Zn" in soil at harvest was mainly influenced by "Cu content in biosolids," with the highest levels obtained with high Cu contents when the "previous crop" was "maize" (Fig. 3, center right, $r = 0.81$). When "Cu content in biosolids" was lower, "extractible Zn in soil" was negatively correlated to the "dry matter of biosolids" but positively to "organic C in the soil."

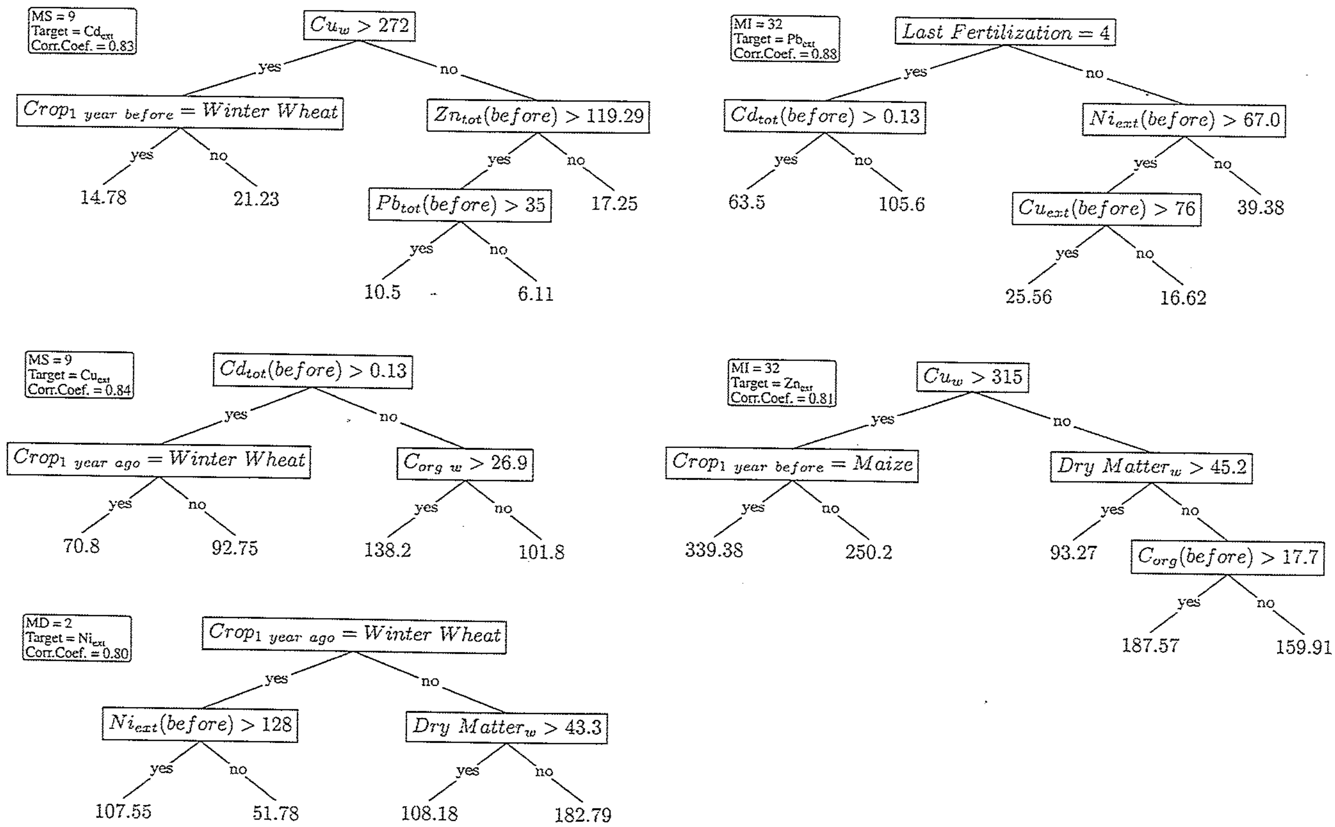

Fig. 3. Regression trees predicting values for $CaCl_2$ extractable metals in soil (μg kg^{-1} dry matter [DM]) at harvest time (Scenario 3). MS = maximum size of the constructed regression trees (i.e., maximum number of predictions [leaves]); MI = minimum instances in a leaf (i.e., minimum number of samples from which the predicted value is calculated); MD: maximum tree depth (i.e., number of hierarchical levels of the tree); "(before)" = content in soil at sowing time; "w" = content in biosolids; "tot" = total metal content (mg kg^{-1} DM); "ext" = $CaCl_2$-extractible metal content (μg kg^{-1} DM). Time is expressed in months. See Table 2 for further explanation concerning attributes. Trees should be read like "what-if" questions from the root of each tree. The closer the attribute is to the root of the tree, the more it is statistically influent for the considered predicted attribute.

Organic Pollutants (Scenario 4)

The data on total PAHs and PCBs in the soils are presented in Supplemental Table S5. Concerning PAHs, the values were usually under or very close to quantification limits. They never exceeded 2.6 mg kg^{-1} of soil, obtained for the treatment LDSS at date TIX. At the end of the study (date TXIII), no PAH could be detected in the soil.

Examining the data-mining results (Fig. 4), "total PAHs concentrations" were mainly correlated to metal contents in soil and biosolids (Fig. 4, top, r = 0.64). They were mainly negatively correlated with "total soil Ni in soil before sowing." Furthermore, in the case of low Ni content in soil, "total PAHs concentrations" were positively correlated to "Cu content in biosolids" and negatively correlated to "Ni content in biosolids" and to "total soil Pb in soil before sowing."

Total PCB values were very low, close to, or below quantification limits. The maximum value was obtained in the LDCSSO treatment at date TIII, with 40 $\mu g \cdot kg^{-1}$ of soil. At the end of the study, no PCB could be detected in the soil samples for any treatment. The regression trees in Fig. 4 show that "total PCBs" mostly increased when "total Cd in soil before sowing" was low (Fig. 4, bottom, r = 0.92). If "total Cd in soil before sowing" was high, it was positively correlated to "total PCB in biosolids" and "phosphorus fertilization."

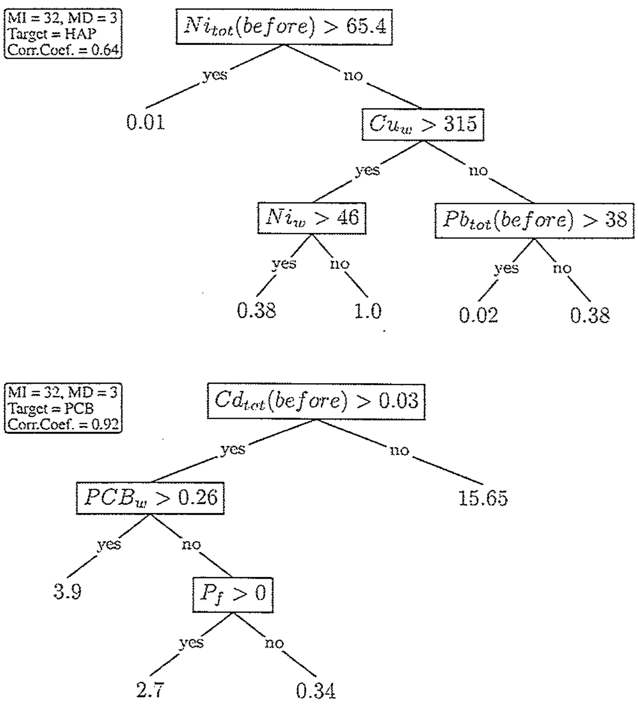

Fig. 4. Regression trees predicting values for total polycyclic aromatic hydrocarbons (PAH) and polychlorinated biphenyl (PCB) (mg kg^{-1} dry matter [DM]) at harvest time (Scenario 4). MI = minimum instances in a leaf (i.e., minimum number of samples from which the predicted value is calculated); MD = maximum tree depth (i.e., number of hierarchical levels of the tree); "(before)" = content in soil at sowing time; "w" = content in biosolids; "tot" = total metal content (mg kg^{-1} DM); "f" = content in fertilizers. See Table 2 for further explanation concerning attributes. Trees should be read like "what-if" questions from the root of each tree. The closer the attribute is to the root of the tree, the more it is statistically influent for the considered predicted attribute.

Discussion

These results demonstrate that data mining is a potent tool for extracting patterns from a large database, here including results of a 10-yr field survey after the application of eight biosolids representative of biosolids potentially used in agriculture.

Soil Fertility

The regression trees for predicting soil fertility parameters show that the initial fertility of soil and crop management are of greater importance than are biosolid applications, especially when biosolids are complemented with mineral fertilizers, which is the most probable case in crop production. Thus, organic matter and N content at harvest are mainly predicted by their contents in soil before sowing: the greater they are at sowing, the greater they will be at harvest. Even if significant differences in N concentrations between biosolid treatments have been shown in the soil, mineral fertilization applied each year is still more predictive of soil N content at harvest than the biosolid N content or the time elapsed since its last application. This result confirms previous reports showing very contrasting results pertaining to the effects on soil N content of biosolid applications. Indeed, these results are highly dependent on the type of crop, the quality of biosolids, or the dose of biosolid applied. Moreover, biosolids are often reported to be less effective in supplying N than inorganic mineral fertilizers (Hargreaves et al., 2008).

In contrast, that the soil C/N ratio can be influenced by biosolids has been confirmed, with values increasing with the number of biosolid applications and higher biosolid C/N ratio. However, even for C/N, crop rotation is still the main driving force. Thus, a change from pasture to a crop production system leads to a decrease in C/N, which has been already been demonstrated in long-term experiments (Blair et al., 2006).

Furthermore, this study has demonstrated that biosolids, particularly biosolids presenting a high CaO content, can slightly increase the available P content in soil at harvest, even though the initial soil P content at sowing is still the best predictor. In addition, it is not available P in biosolids that influences the available P in soil, but other characteristics of biosolids, such as their CaO content. These results confirm that the physicochemical characteristics of biosolids need to be taken into account when assessing the impact of their application in agriculture. They lead us to the conclusion that available P and C/N ratio of a soil is probably the best indicator—at least better than organic matter and total N—to be included in decision support systems for stakeholders wishing to assess agricultural practice management involving the use of biosolids.

Concerning available P, the above conclusion is supported by recent results obtained under controlled conditions, showing strong but transient phosphatase activity and available P increases following sewage sludge application. These effects vary with the initial sludge chemical characteristics but not with soil initial characteristics (Criquet et al., 2007). Field studies have also confirmed that soil P is strongly modified, usually increased, by sewage sludge compost application (Korboulewsky et al., 2002; Mantovi et al., 2005).

Trace Elements

Total trace element concentrations in soils were systematically lower than threshold values accepted by French legislation (Voynet et al., 1998). In addition to the chemical forms in which each trace metal occurs, metal mobility and behavior into the soil are known to be governed by soil properties, such as the pH, or the abundance of constituents that can easily retain metals, such as iron and manganese oxides or organic matter (Baize, 2009).

Cadmium

An increase in a soil's total Cd is observed over a period of time, from low values at the beginning of the experiment to comparatively higher values at the end. This result is reflected by the regression tree models: low values are observed when the previous crop was winter wheat, which is the case at the beginning of the experiment, and higher values are observed later. Similarly, the situation with the "last biosolids spread occurring more than 18 months previously," which is correlated to the highest total Cd values, corresponds to the end of the experiment. No treatment effect is shown by the model, and the increase in time is probably due to another parameter independent of any treatment in the system. Mineral fertilization cannot explain this global increase, as no difference occurs between the two controls CMF and COF. Aerial deposit increases might be suspected but would require more investigation.

When comparing $CaCl_2$–extractable Cd with total Cd, the opposite situation is observed, with a decrease in concentration levels with time following the last biosolid application. This result has been commonly observed and indicates that the metal evolves over time toward less mobile forms (Morel and Guckert, 1984). Furthermore, due to the significant effect of biosolids on soil pH (Supplemental Table S2), soil extractable Cd is less important after MPS application compared with other biosolids. This result is also reflected by our models, showing that the main factor able to predict extractable Cd is linked to the biosolid quality, particularly biosolid Cu content. Thus, the different sewage sludges (from liquid to composted sludge: LSS, LDSS, LDCSS, LDCSSO, and LDCSSM) contain more Cu and are clearly discriminated by the model from the other biosolids (MPS, CA, and HWA).

For liquid and composted sludge, the $CaCl_2$–extractable Cd fraction is dependent on time (expressed here by "crop the year before"), with a decrease in concentrations over periods of years following sludge application. For MPS, CA, and HWA, extractable Cd is positively correlated with Zn and negatively with Pb, which agrees with the literature relating to the behavior of metals following biosolids application (Smith, 2009). This leads us to the conclusions that for extractable Cd prediction, (i) the initial composition of biosolids is the strongest indicator; (ii) the maturation processes occurring during sludge composting are not enough to discriminate liquid and composted sludge; and (iii) no correct prediction on Cd bioavailability can be performed from total Cd content, as no clear relation can be established between available and total Cd. Finally, our long-term in situ results confirm a recent review showing that biosolids applied in accordance with the 1998 French regulation had no significant impact on total soil Cd concentrations (Baize, 2009).

Copper

Contrary to Cd, total Cu in soil at harvest mostly depends on total Cu in soil before and, to a lesser extent, on total Zn in the soil before harvest, showing some antagonism between the two metals. This result of the model confirms the dynamics of total Cu in soil, with no real increase with time in situ. Total Cu is thus rather more soil-property dependent than time dependent, being partly influenced by the previous crop. No biosolid effect was identified by the model, even though significant differences were noticed during the study.

The picture looks different when considering "extractable Cu," which is strongly correlated to "total Cd in soil before sowing." This is due to the coincidence in time of decreasing extractable Cu and increasing total Cd. Indeed, "$CaCl_2$ extractable Cu" is time dependent. However, at the beginning of the experiment, when the extractable Cu in the soil was the highest, the amount of extractable Cu also depended on the biosolid quality, particularly its C_{org} content: high extractable Cu in soil is linked to high C_{org} in biosolids. This result agrees with recent studies showing that the addition to soil of organic matter from composted biosolids can raise the extractable concentrations of Cu compared with unamended soil receiving only mineral fertilizers (Herencia et al., 2008).

Nickel

Just as with total Cd content, total Ni is time dependent, increasing during the study for all the treatments, when "previous crop" was "maize." However, at the end of the experiment, the biosolid quality, and particularly the biosolid pH, also seemed to influence total Ni in soil. Lower values were predicted when composted sludge (pH < 6.5) was used than with the other biosolids. We suspect that a lower pH in biosolids could increase the mobility of Ni in soil and thus reduce the total amount in the soil. This result is supported by the significantly lower soil pH when using sewage sludge (whether composted or not) compared with MPS, CA, or HWA (Supplemental Table S2). This hypothesis is confirmed by the literature on the subject of sequential extraction during sewage sludge composting (Amir et al., 2005). Indeed, Ni was observed to behave differently from other metals. This behavior could be attributed to the high proportion of Ni in the raw material that is present in an organic, readily biodegradable form mobilizing Ni during the composting process (Smith, 2009). As with total Cu, some antagonism of total Ni is noticed with total Zn, Ni increasing when Zn decreases.

The $CaCl_2$–extractable Ni shows another pattern and looks to be primarily time dependent, with, as observed for Cd and Cu, lower values when "winter wheat" was the "previous crop." It is interesting to observe that extractable Ni is also dependent on the amount of "extractable Ni in the soil before sowing," but also on the "dry matter content of biosolid." This observation helps to distinguish between MPS, CA, and HWA and other biosolids.

Lead

Total Pb seems to depend mostly on time, but some antagonism with total Ni and Cd is also confirmed. Furthermore, an increase in total Pb content in biosolids is shown to induce an increase in Pb in a soil. However, unlike other metals, "extractable Pb" follows more or less the same rules as "total Pb," as some antagonism is found between "total Cd" and "extractable Ni." This result leads us to the conclusion that the mobility of Pb is weak, as has been noted many times in the literature (Smith, 2009). Furthermore, the positive correlation found between extractable Pb and Cu confirms their similar behavior.

Zinc

The slight increase in total Zn over time is reflected by the model, as the maximum Zn content coincided with a mineral fertilization occurring 5 mo previously. Only a positive correlation with total Ni was noticeable. No treatment effect was predicted by the model, suggesting the relative inefficiency of biosolid effects as predictors of total Zn.

A completely different situation is observed for $CaCl_2$–extractable Zn, which is influenced, like extractable Cd, by the biosolids quality and particularly Cu content. Furthermore, as for extractable Ni, when the dry matter of biosolids increased, the Zn extractable fraction decreased. The hypothesis suggested for Cd and Ni can be also argued here. Finally, to a lesser extent, the $CaCl_2$–extractable Zn fraction is classically influenced by C_{org} content in the soil.

Organic Pollutants

The PAH values obtained were lower than commonly considered to be hazardous in several countries: 40 mg kg^{-1} in the Netherlands, 20 mg kg^{-1} in Canada (Quebec province) and Switzerland, and 50 mg kg^{-1} in the United Kingdom (Costes and Druelle, 1997; Conseil Fédéral Suisse, 1998). The same conclusion can be drawn concerning PCBs with regards to the 200 mg kg^{-1} threshold accepted in Switzerland.

The results concerning organic pollutants are more difficult to interpret than those relating to metals because the evolution of these pollutants, after the addition of biosolids, is the combined result of multiple processes, including adsorption, desorption, bioformation, volatilization, photodegradation, leaching, and incorporation into humic substances (Oleszczuk, 2006). In our study, PAHs seemed to be linked to the presence of metals in soils and biosolids, particularly Cu, Ni, and Pb. In fact, this result is not surprising, as some studies have demonstrated that metals could increase the sorption by the soil of phenanthrene, which is used as a representative PAH (Gao et al., 2006).

Similarly, using soil collected from the same site (La Bouzule), it has been previously demonstrated that the adsorption of phenanthrene was higher in the presence of metals, which may be related to a change in the structural conformation of organic molecules in soils. Furthermore, metals would increase the extractability of phenanthrene (Saison et al., 2004). However, in our case, it remains unclear why correlations are sometimes positive and sometimes negative with Cu in biosolids, with Ni in soil and biosolids, and with Pb in biosolids.

On the other hand, it is interesting to observe that PCBs are also correlated to the presence of metals, particularly Cd in soil. However, unlike PAHs, PCBs in soil are also correlated to PCBs in biosolids and to the P provided by fertilization. Thus, PCBs measured in the soil are probably a combination of PCBs originating in biosolids and mineral fertilizers in interaction with metals in the soil.

Finally, our results show that (i) biosolid spreading at recommended doses offers a good alternative for agriculture to increase soil fertility, as this can provide at least the same amounts of N and C, and more P as mineral fertilization; (ii) a long-term effect of biosolids is observed on available P and C/N in soils; and (iii) significant differences occur between the different types of biosolids. Concerning metallic and organic trace elements, the measured values were far lower than the commonly accepted values in Europe. This means that biosolid spreading at recommended doses is safe in terms of contaminant behavior and accumulation in soils.

Consequently, these results show that data mining is a powerful tool for extracting and linking various attributes into patterns consistent with existing knowledge of the biogeochemical functioning of the soil–biosolid–plant system. As an example, our approach has confirmed that depending on the biosolid quality, the P content provided could vary greatly and should thus be included, together with N, in any decision support system used to decide on the application level of these biosolids. More generally, the results obtained by this study enable us to identify and choose relevant indicators to enhance and refine a decision support system for efficient biosolid application in agriculture.

Acknowledgments

This work was supported by the ADEME (Isabelle Feix and Antonio Bispo), via the work program VADETOX (contract no. 0375C0006), and the INRA, via the bilateral collaborative program WAS between France and Slovenia. Thanks to all the scientific and technical partners of VADETOX for fruitful discussions. We are also particularly grateful to Antonio Bispo (ADEME), Anne-Marie Charissou (IPL EST), Louis Florentin (LSE INPL INRA), and all the technical staff (particularly Bernard Colin and Stéphane Colin) for their help during the collection of samples and analysis of results.

References

ADEME and Cabinet Arthur Andersen. 1999. Examen de la situation de la filière de recyclage agricole des boues d'épuration urbaines en Europe et dans divers autres pays du monde, no. 3358. ADEME Editions, Paris.

AFNOR. 2004. Evaluation de la Qualité des Sols. Vol. 1. Méthodes d'Analyse chimique. AFNOR, Paris.

Aggelides, S.M., and P.A. Londra. 2000. Effects of compost produced from town wastes and sewage sludge on the physical properties of a loamy and clay soil. Bioresour. Technol. 71:253–259. doi:10.1016/S0960-8524(99)00074-7

Al-Assiuty, A.I., M.A. Khalil, and H.M. Abdel-Lateif. 2000. Effects of dry sludge application on soil microarthropod communities in a reclaimed desert ecosystem. Pedobiologia 44:567–578. doi:10.1078/S0031-4056(04)70072-4

Amir, S., N. Hafidi, G. Merlina, and J. Revel. 2005. Sequential extraction of heavy metals during composting of sewage sludge. Chemosphere 59:801–810. doi:10.1016/j.chemosphere.2004.11.016

Baize, D. 2009. Cadmium in soils and cereal grains after sewage-sludge application on French soils: A review. Agron. Sustain. Dev. 29:175–184. doi:10.1051/agro:2008031

Banerjee, M.R., D.L. Burton, and S. Depoe. 1997. Impact of sewage sludge application on soil biological characteristics. Agric. Ecosyst. Environ. 66:241–249. doi:10.1016/S0167-8809(97)00129-1

Blair, N., R.D. Faulkner, A.R. Till, and P.R. Poulton. 2006. Long-term management impacts on soil C, N and physical fertility: Part I. Broadbalk experiment. Soil Tillage Res. 91:30–38. doi:10.1016/j.still.2005.11.002

Blockeel, H., and J. Struyf. 2002. Efficient algorithms for decision tree cross-validation. J. Mach. Learn. Res. 3:621–650.

Bohanec, M., J. Cortet, B. Griffiths, M. Žnidaršič, M. Debeljak, S. Caul, J. Thompson, and P.H. Krogh. 2007. A qualitative multi-attribute model for assessing the impact of cropping systems on soil quality. Pedobiologia 51:239–250. doi:10.1016/j.pedobi.2007.03.006

Brazauskiene, D.-M., V. Paulauskas, and N. Sabiene. 2008. Speciation of Zn, Cu, and Pb in the soil depending on soil texture and fertilization with sewage sludge compost. J. Soils Sediments 8:184–192. doi:10.1007/s11368-008-0004-6

Breiman, L., J.H. Friedman, R.A. Olshen, and C.J. Stone. 1984. Classification and regression trees. Wadsworth, Belmont, CA.

Conseil Fédéral Suisse. 1998. Ordonnance sur les atteintes portées aux sols (OSol). 814.12. Conseil Fédéral Suisse, Bern.

Costes, J.M., and V. Druelle. 1997. Les hydrocarbures aromatiques polycycliques dans l'environnement: La réhabilitation des anciens sites industriels. Rev. I.F.R. Petrol. 52:425–440.

Criquet, S., A. Braud, and S. Nèble. 2007. Short-term effects of sewage sludge application on phosphatase activities and available P fractions in mediterranean soils. Soil Biol. Biochem. 39:921–929. doi:10.1016/j.soilbio.2006.11.002

Debeljak, M., J. Cortet, D. Demsar, P.H. Krogh, and S. Dzeroski. 2007. Hierarchical classification of environmental factors and agricultural practices affecting soil fauna under cropping systems using Bt maize. Pedobiologia 51:229–238. doi:10.1016/j.pedobi.2007.04.009

Debeljak, M., and S. Dzeroski. 2009. Applications of data mining in ecological modelling. p. 409–423. *In* S.E. Jørgensen, T-S. Chon, and F. Recknagel (ed.) Handbook of ecological modelling and informatics. WIT Press, Southampton, UK.

Debeljak, M., and S. Dzeroski. 2011. Decision trees in ecological modelling. p. 197–211. *In* F. Jopp, H. Reuter, and B. Breckling (ed.) Modelling complex ecological dynamics. Springer-Verlag, Berlin.

Debeljak, M., D. Kocev, W.J.M. Towers, B. Griffiths, and P. Hallett. 2009. Potential of multi-objective models for risk-based mapping of the resilience characteristics of soils: Demonstration at a national level. Soil Use Manage. 25:66–77. doi:10.1111/j.1475-2743.2009.00196.x

Debeljak, M., G: Squire, D. Demšar, M.W. Young, and S. Dzeroski. 2008. Relations between the oilseed rape volunteer seedbank, and soil factors, weed functional groups, and geographical location in the UK. Ecol. Modell. 212:138–146. doi:10.1016/j.ecolmodel.2007.10.019

Debeljak, M., G. Squire, D. Kocev, C. Hawes, M.W. Young, and S. Džeroski. 2011. Analysis of time series data on agroecosystem vegetation using predictive clustering trees. Ecol. Modell. 222:2524–2529. doi:10.1016/j.ecolmodel.2010.10.021

Demšar, D., S. Džeroski, T. Larsen, J. Struyf, J. Axelsen, M. Bruus-Pedersen, and P.H. Krogh. 2006. Using multi-objective classification to model communities of soil. Ecol. Modell. 191:131–143. doi:10.1016/j.ecolmodel.2005.08.017

Gabrielle, B., J. Da-Silveira, S. Houot, and J. Michelin. 2005. Field-scale modelling of carbon and nitrogen dynamics in soils amended with urban biosolid composts. Agric. Ecosyst. Environ. 110:289–299. doi:10.1016/j.agee.2005.04.015

Gao, Y., W. Xiong, W. Ling, and J. Xu. 2006. Sorption of phenanthrene by soils contaminated with heavy metals. Chemosphere 65:1355–1361. doi:10.1016/j.chemosphere.2006.04.030

Hargreaves, J.C., M.S. Adl, and P.R. Warman. 2008. A review of the use of composted municipal solid waste in agriculture. Agric. Ecosyst. Environ. 123:1–14. doi:10.1016/j.agee.2007.07.004

Herencia, J.F., J.C. Ruiz, E. Morillo, S. Melero, J. Villaverde, and C. Maqueda. 2008. The effect of organic and mineral fertilization on micronutrient availability in soil. Soil Sci. 173:69–80. doi:10.1097/ss.0b013e31815a6676

Ihaka, R., and R. Gentleman. 1996. A language for data analysis and graphics. J. Comput. Graph. Statist. 5:299–314. doi:10.2307/1390807

Juste, C., and M. Mench. 1992. Long-term application of sewage sludge and its effects on metal uptake by crops. p. 159–193. *In* D.C. Adriano (ed.) Biochemistry of trace elements. Lewis Publ., Boca Raton, FL.

Juste, C., and P. Solda. 1977. Effets d'applications massives de boues de stations d'épuration urbaines en monoculture de maïs: Action sur le rendement et la composition des plantes et sur quelques caractéristiques du sol. Sci. Sol 3:147–155.

Karlen, D.L., M.J. Mausbach, J.W. Doran, R.G. Cline, R.F. Harris, and G.E. Schuman. 1997. Soil quality: A concept, definition and framework for evaluation. Soil Sci. Soc. Am. J. 61:4–10. doi:10.2136/sssaj1997.03615995006100010001x

Korboulewsky, N., S. Dupouyet, and G. Bonin. 2002. Environmental risks of applying sewage sludge compost to vineyards: Carbon, heavy metals, nitrogen and phosphorus accumulation. J. Environ. Qual. 31:1522–1527. doi:10.2134/jeq2002.1522

Lebourg, A., T. Sterckeman, H. Ciesielski, and N. Proix. 1996. Intérêt de différents réactifs d'extraction chimique pour l'évaluation de la biodisponibilité des métaux en traces du sol. Agronomie 16:201–215. doi:10.1051/agro:19960401

Logan, T.J., B.J. Lindsay, L.E. Goins, and J.A. Ryan. 1997. Field assessment of sludge metal bioavailability to crops: Sludge rate response. J. Environ. Qual. 26:534–550. doi:10.2134/jeq1997.00472425002600020027x

Mantovi, P., G. Baldoni, and G. Toderi. 2005. Reuse of liquid, dewatered, and composted sewage sludge on agricultural land: Effects of long-term application on soil and crop. Water Res. 39:289–296. doi:10.1016/j.watres.2004.10.003

Martin-Laurent, F., L. Cornet, L. Ranjard, J.-C. Lopez-Gutierrez, L. Philippot, C. Schwartz, R. Chaussod, G. Catroux, and G. Soulas. 2004. Estimation of atrazine-degrading genetic potential and activity in three French agricultural soils. FEMS Microbiol. Ecol. 48:425–435. doi:10.1016/j.femsec.2004.03.008

Mitchell, M.J., R. Hartenstein, B.L. Swift, E.F. Neuhauser, B.I. Abrams, R.M. Mulligan, B.A. Brown, D. Craig, and D. Kaplan. 1978. Effects of different sewage sludges on some chemical and biological characteristics of soil. J. Environ. Qual. 7:5591–5598.

Morel, J.L., and A. Guckert. 1984. Evolution en plein champ de la solubilité dans DTPA des métaux lourds du sol introduits par des épandages de boues urbaines chaulées. Agronomie 4:377–386. doi:10.1051/agro:19840409

Oleszczuk, P. 2006. Persistence of polycyclic aromatic hydrocarbons (PAHs) in sewage sludge-amended soil. Chemosphere 65:1616–1626. doi:10.1016/j.chemosphere.2006.03.007

Oliver, I.W., M.J. McLaughlin, and G. Merrington. 2005. Temporal trends of total and potentially available element concentrations in sewage biosolids: A comparison of biosolid surveys conducted 18 years apart. Sci. Total Environ. 337:139–145. doi:10.1016/j.scitotenv.2004.07.003

Petersen, S.O., K. Henriksen, G.K. Mortensen, P.H. Krogh, K.K. Brandt, J. Sørensen, T. Madsen, J. Petersen, and C. Grøn. 2003. Recycling of sewage sludge and household compost to arable land: Fate and effects of organic contaminants, and impact on soil fertility. Soil Tillage Res. 72:139–152. doi:10.1016/S0167-1987(03)00084-9

Robert, M. 1996. Le sol: Interface dans l'environnement, ressource pour le développement. Masson, Paris.

Rogers, H.R. 1996. Sources, behaviour and fate of organic contaminants during sewage treatment and in sewage sludges. Sci. Total Environ. 185:3–26. doi:10.1016/0048-9697(96)05039-5

Saison, C., C. Perrin-Ganier, S. Amellal, J.L. Morel, and M. Schiavon. 2004. Effect of metals on the adsorption and extractibility of ^{14}C-phenanthrene in soils. Chemosphere 55:477–485. doi:10.1016/j.chemosphere.2003.10.059

Singh, R.P., and M. Agrawal. 2008. Potential benefits and risks of land application of sewage sludge. Waste Manag. 28:347–358. doi:10.1016/j.wasman.2006.12.010

Smith, S.R. 2009. A critical review of the bioavailability and impacts of heavy metals in municipal solid waste composts compared to sewage sludge. Environ. Int. 35:142–156. doi:10.1016/j.envint.2008.06.009

Stamatiadis, S., J.W. Doran, and T. Kettler. 1999. Field and laboratory evaluation of soil quality changes resulting from injection of liquid sewage sludge. Appl. Soil Ecol. 12:263–272. doi:10.1016/S0929-1393(99)00007-4

Struyf, J., and S. Džeroski. 2006. Constraint based induction of multi-objective regression trees. p. 222–233. *In* F. Bonchi and J. Boulicaut (ed.) Knowledge Discovery in Inductive Databases. 4th Int. Workshop, KDID'05, Revised, Selected and Invited Papers. LNCS, Vol. 3933. Springer, Berlin.

Tan, P., M. Steinbach, and V. Kumar. 2006. Introduction to data mining. Addison-Wesley, Boston, MA.

Voynet, D., J.-P. Chevènement, L. Le Pensec, E. Zuccarelli, and B. Kouchner. 1998. Arrêté du 8 janvier 1998 fixant les prescriptions techniques applicables aux épandages de boues sur les sols agricoles pris en application du décret n° 97–1133 du 8 décembre 1997 relatif à l'épandage des boues issues du traitement des eaux usées. Journal officiel de la République Française, 31 Jan., p. 1563–1571.

Wei, Q.F., B. Lowery, and A.E. Peterson. 1985. Effect of sludge application on physical properties of silty clay loam soil. J. Environ. Qual. 14:178–180. doi:10.2134/jeq1985.00472425001400020005x

Znidarsic, M., M. Bohanec, and B. Zupan. 2006. proDEX: A DSS tool for environmental decision-making. Environ. Model. Softw. 21:1514–1516. doi:10.1016/j.envsoft.2006.04.003