A Bayesian network approach for the identification of relationships between drivers of chlordecone bioaccumulation in plants

Yohan Liber, Denis Cornet, Régis Tournebize, Cyril Feidt, Maurice Mahieu, Francois Laurent, Jean-Philippe Bedell

A Bayesian network approach for the identification of relationships between drivers of chlordecone bioaccumulation in plants

Yohan Liber [b], Denis Cornet [c,d], Régis Tournebize [e], Cyril Feidt [f], Maurice Mahieu [g], François Laurent [b], Jean-Philippe Bedell [a]

a : LEHNA, UMR 5023, CNRS, ENTPE, University of Lyon, F-69518 Vaulx-en-Velin, France
b : INRA, UMR1331, Toxalim, F-31027 Toulouse Cedex 3, France
c : CIRAD, UMR AGAP, F-34398 Montpellier, France
d : University of Montpellier, CIRAD, INRA, Montpellier SupAgro, Montpellier, France
e : INRA. UR1321, ASTRO, F-97170 Petit-Bourg, Guadeloupe, France
f : URAFPA, Université de Lorraine, INRA USC340, F-54500 Vandoeuvre-lès-Nancy, France
g : INRA, URZ, UR 143, F-97170 Petit-Bourg, Guadeloupe, France

Published in Environmental Science and Pollution Research. Available online: DOI 10.1007/s11356-019-07449-0

Keywords

Bayesian network modeling, Bioaccumulation, Chlordecone, Land management, Pesticide

Abstract

Plants were sampled from four different types of chlordecone-contaminated land in Guadeloupe (West Indies). The objective was to investigate the importance of biological and agri-environmental parameters in the ability of plants to bioaccumulate chlordecone. Among the plant traits studied, only the growth habit significantly affected chlordecone transfer, since prostrate plants concentrated more chlordecone than erect plants. In addition, intensification of land use has led to a significant increase in the amount of chlordecone absorbed by plants. The use of Bayesian networks uncovers some hypothesis and identifies paths for reflection and possible studies to identify and quantify relationships that explain our data.

1. Introduction

The French West Indies are facing an environmental catastrophe due to soil and water contamination by chlordecone (CLD). This organochlorine pesticide was used to control the banana weevil (Cosmopolites sordidus) in Guadeloupe and Martinique from 1972 to 1993. Cabidoche et al. (2009) reported that chlordecone pollution is expected to last several decades in nitisols and several centuries in ferralsols and andosols. In Guadeloupe, 15% of the island's agricultural land is contaminated, at least 2100 ha of agricultural soil with between 0.25 and 1 mg kg–1 of dry soil, and 3100 ha > 1 mg kg–1 (Cabidoche et al. 2006). Such diffuse contamination cannot be treated by ex situ depollution methods, although progress has been made recently (Chaussonnerie et al. 2016; Mouvet et al. 2016). To clean up the soil, or to avoid contamination of neighboring agricultural plants, realistic solutions are needed at plot scale. Phytoremediation is a potential solution. Transfer of chlordecone to the aboveground plant parts has been demonstrated in many studies (Cabidoche and Lesueur-Jannoyer 2012; Lorber-Pascal et al. 2016; Liber et al. 2018). Of the 3200-plant species in Guadeloupe (Fournet 2002), only a few cultivated species have been analyzed so far, as part of specific studies on dietary risk (Cabidoche and Lesueur-Jannoyer 2012; Clostre et al. 2015). Cultivated species represent less than 6% of the island's plant diversity (Lazzeri and Mouhoud 2010). Despite this important floristic potential, we do not have enough knowledge to identify the plant characteristics that enable accumulation of chlordecone. Such knowledge would make it possible to select species or identify key traits to consider in phytoremediation of these contaminated areas. The aim of the present study was thus to identify, in vivo, the relationships between physical-anatomical and environmental drivers that could promote extraction of chlordecone from the soil and its accumulation in plants.

2. Material and Methods

2.2 Sampling and preparation

Sampling was conducted in the watersheds of the Pères and Pérou rivers south of the island of Basse-Terre, Guadeloupe. Due to heterogeneous spreading, chlordecone contamination in soils varies considerably (Clostre et al. 2014). To ensure sampling from contaminated sites, plots were chosen based on past use of chlordecone. Three types of sites were selected. First, banana and sugar cane plantations. As these plots are still in use, sampling focused on weeds and shrubs growing in the inter-rows. Second, Creole gardens, a traditional horticultural system based on vegetatively multiplied crops and permanent soil cover, aself-sustaining system used by private individuals. Third, former banana plantations, now abandoned and covered by spontaneous vegetation, which host a wide variety of plant species.

For each sampled species, about 250 g of fresh aboveground biomass was collected and stored in individual plastic bags with three to five sub-samples according to the patch size sample of vegetal. The smallest plants were pooled to obtain the required biomass for analysis. For details and list of species as site typology see Supplementary data N°1.

The corresponding bulk soil (up to 20 cm of depth) was hand-picked and stored in 1-L aluminum trays. To avoid contamination, nitrile gloves were used and changed between each soil sampling. Distilled water tanks were used to clean the plant samples in three successive steps: (1) in the first tank large particles were removed with a brush, (2) a second tank was used to hand wash and decant the finest dust particles, and (3) a last tank was used to rinse the samples. Only the water in the last tank was changed between samples. All the samples were dried in an oven at 70 °C for 72 h before being weighed. Thereafter, the soil samples were crushed and sieved to 2 Øand the plant samples were ground to a powder using a jaw crusher with a 0.2-Øsieve. Fifty grams of dry soil was put in glass jars, and about 20 g of dry plant powder was packed in plastic bags for shipment to the laboratory. Analyses were performed at the Drôme departmental laboratory (LDA 26), France. Chlordecone content was quantified by liquid chromatography coupled with mass spectrometry (LCMS) according to standard ANSES LSA-INS-0161.

2.3 Observations

Plant features were categorized in four binary and three continuous variables: the photosynthetic pathway (C3 or C4 metabolism), the root type (fasciculate or storage root systems), the phenological stage (flowering or vegetative), growth habit (erect or prostrate growth), chlordecone concentration in shoot tissue (µg kg–1 DV), total biomass harvested (g DV), and the total amount of accumulated chlordecone (µg) (see SD N°1).

Each plot was geolocalized and overlain on a detailed soil map (ORSTOM 1981) to identify their specific soil classification (see SD N°1). Based on this classification and on the literature (Dorel et al. 2000), three derived variables were built: (1) soil type (i.e., (i) Nitisols grouping ferralitic soils with halloysite and low organic matter content, (ii) andosols including andosols with allophane, brown andic soils, and perhydrated andosols); (2) soil organic matter content (%); and (3) soil allophane content (%). Soil characteristics were completed by measuring the chlordecone concentration in the soil samples (µg kg–1 DV).

Based on the current use of the plot, a land use binary variable was derived, describing the current land intensification level (intensive for banana or sugar cane and extensive for fallows and Creole gardens).

Finally, for each pair of soil-plant observations, the bioconcentration factor in the shoot (BCF) was estimated as the ratio between the CLD concentrations in the aboveground plant parts and the CLD concentration in the associated soil.

2.4 Statistical analysis

Bayesian networks (BNs) have emerged as a suitable approach to complex systems when prior knowledge is lacking. BNs are increasingly used as an integrative tool in agricultural (Cornet et al. 2016; Drury et al. 2017) and environmental sciences (Aguilera et al. 2011). Unlike multivariable regression that seeks to identify the covariates associated with certain variables of interest, Bayesian network analysis goes much further by empirically separating these variables into those that depend directly and indirectly on the concentration variable. Bayesian network modeling has the potential to reveal much more about the key features of complex biological systems than currently available approaches (Lewis and McCormick 2012). Unlike the factorial approach, Bayesian network modeling does not attempt to reduce the dimensionality of the dataset, allowing for more straightforward biological interpretation of the results.

Bayesian networks are multivariate models with two main components: first, a qualitative component illustrated by a directed acyclic graph (DAG), which is a graphical representation of the joint probability distribution of all the random variables. Nodes of the DAG account for variables and the arrows between them represent direct and indirect relationships based on knowledge or statistical associations. Second, a quantitative component of conditional probabilities quantifying the strength and the uncertainty of the relationships between variables (Jensen and Nielsen 2007). We used an additive Bayesian network (ABN) that extends the usual generalized linear model to multiple dependent variables (Lewis and McCormick 2012; Pittavino et al. 2017). All analyses were conducted using R software, version 3.4.4 (R Development Core Team 2017) and the R “abn” package (Pittavino et al. 2016). In the absence of any prior knowledge or data, all DAG structures were equally supported by a uniform prior. A three-step procedure (Pittavino et al. 2017) was used:

Identifying an optimal model that best supported the observed data using exact search and log marginal likelihood as model score;

Adjusting the model for overfitting using Markov chain Monte Carlo (MCMC) simulations implemented in JAGS (Babyak 2004). Arcs present in less than 50% of the MCMC simulations (2500/5000) were considered not to be robust and removed from the DAG generated in the first step (Friedman et al. 1999, Pittavino et al. 2017);

Estimating the marginal posterior log odds ratio and 95% credible intervals for each parameter from the posterior distribution

3. Results and Discussion

ABN allowed us to identify the best dependency structure among our observed variables (Fig 1). MCMC adjustment led us to discard all three relationships explaining soil chlordecone concentration. However, with 2463 recoveries over 5000 MCMC simulations, trimming the relationship of soil organic matter with soil chlordecone maybe questionable. Gathering more data or using informative prior based on the literature could have led us to keep this dependency.

DAG revealed no significant relationship between BCF and other drivers. Most of the information carried by the calculated BCF variable is included in the measured plant chlordecone content which is privileged in the analysis. No robust relationship was found between the CLD concentration in soil and in the plant, suggesting that transfers of CLD to the plant are more related to the bioavailability of CLD in soil than to its relative concentration. Moreover, at field scale, past pesticide application practices account for the potential stock of pollutant in the soil (Woignier et al. 2014). The lack of relationship explaining soil chlordecone concentration in our study may first be caused by the lack of information on past applications. This was confirmed by the direct relationship between soil organic matter contents and chlordecone concentrations in the plant. Surprisingly, the dataset showed an increase in CLD concentration in the plant with an increase in the percentage of organic matter in the soil. In previous studies, organic matter was shown to play a role in pesticide sequestration in soil (Barriuso et al. 1997) including CLD (Woignier et al. 2016). Similarly, we observed that the land use variable revealed an increase in the amount of chlordecone absorbed by plants in intensive cropping systems (Fig. 1, positive and significant relationship between Land use and Plant CLD amount).

To understand how land use influences CLD accumulation in plants despite high organic matter contents, Fig. 2 examines the significance of the relationships between the land use variable and several other parameters. No relationship between land use and plant biomass appeared to justify the increase in CLD levels in plants growing in intensive land use conditions (Fig. 2a). Soils with the highest organic matter content were also those under the most intensive agriculture but their CLD concentrations did not differ significantly from soils with low agricultural intensification (Fig. 2b). It has been shown that soils under conventional management (conventional tillage, N-rich amendment, associated with intensive cropping systems) exhibit fewer aggregates formation that can be linked to a better soil structuration and then to a potential better soil stability for rain or hydrological events (Blanco-Canqui and Lal 2008; Trivedi et al. 2018; Rowley et al. 2018). From the results reported by Sierra et al. (2015) for several cropping systems in Guadeloupe, we calculated that the rate of soil organic carbon (SOC) mineralization in andosols under intensive banana cultivation is 4–5 times higher than in andosols or nitisols in no till systems. The release of organic matter accelerated by intensive land use may have increased the desorption of chlordecone into the aqueous phase of the soil. This phenomenon enhances the bioavailability of chlordecone for the plant, which explains the higher bioaccumulation in samples harvested on intensively cultivated andosols (Fig. 2c). Another relationship expressed by the DAG concerns the possible influence of plant growth habit on CLD accumulation. Prostrate plants concentrate more chlordecone than erect plants (Figs. 1 and 2d). The difference in statistical significance effect of growth habit between Figs. 1 and 2d is first linked to different dependent variables, respectively plant CLD concentration and BCF. While BCF is dependent on soil CLD concentration, plant CLD concentration is not leading to different results (e.g., a unique low BCF value can have multiple causes such as low plant CLD concentration or high soil CLD concentration). These results also highlight the advantage of Bayesian networks over Wilcoxon tests. Indeed, Bayesian network, by modeling efficiently the joint probability distribution of all variables at once, allows for solving common issues related to mediated effect and/or multicollinearity by using conditional dependencies (Darwiche 2009). Because there was no significant difference in soil CLD concentrations between these two growth modes (Fig. 2e), this pattern of contamination can be explained by deposits of particle-bound chlordecone on the aboveground parts of prostrate plants through rain splash or direct soaking in the soil solution (Dreicer et al. 1984; Collins et al. 2006). Once in contact with the shoots, organic chemicals bound to the particles may diffuse through the lipophilic cuticle and become sorbed or permeate the plant (Collins et al. 2006).

4. Conclusion

Analysis of the dependency relationships of in situ parameters that may affect plant chlordecone uptake revealed that land management is of paramount importance in plant contamination. The most important bioconcentration factors were found in soils intensively managed, highlighting that we need to improve the carbon cycles in such agricultural systems. Bayesian networks permit us to think as a first step that the role of contact between the soil and the plant in the contamination of prostrate species can be one explanation. The intake of contaminated soil is the exposure route regularly put forward to explain the contamination of livestock by chlordecone. For example, grazing cattle can intake up to 4% of their ration as soil (Jurianz et al. 2017; Collas et al. 2019). However, drinking water and food are other potential routes of contamination. Our finding has clear implications for healthy crop or animal production management, notably by giving information for plant species that can be grazed by cattle reared on polluted areas.

Itis difficult to identify biological drivers that allow plants to bioaccumulate chlordecone in the natural environment. The interference of many uncontrolled variables introduces confounding factors in the analysis that complicate interpretation of the results. The use of Bayesian networks showed some hypothesis and identifies paths for reflection and possible studies to identify and quantify relationships that explain our data. To further understand and identify key plant traits affecting chlordecone bioaccumulation in such a multivariate and complex system, it would be necessary either to increase the number of samples, or to conduct comparative experiments with matched samples differing only in one or a few key characteristics.

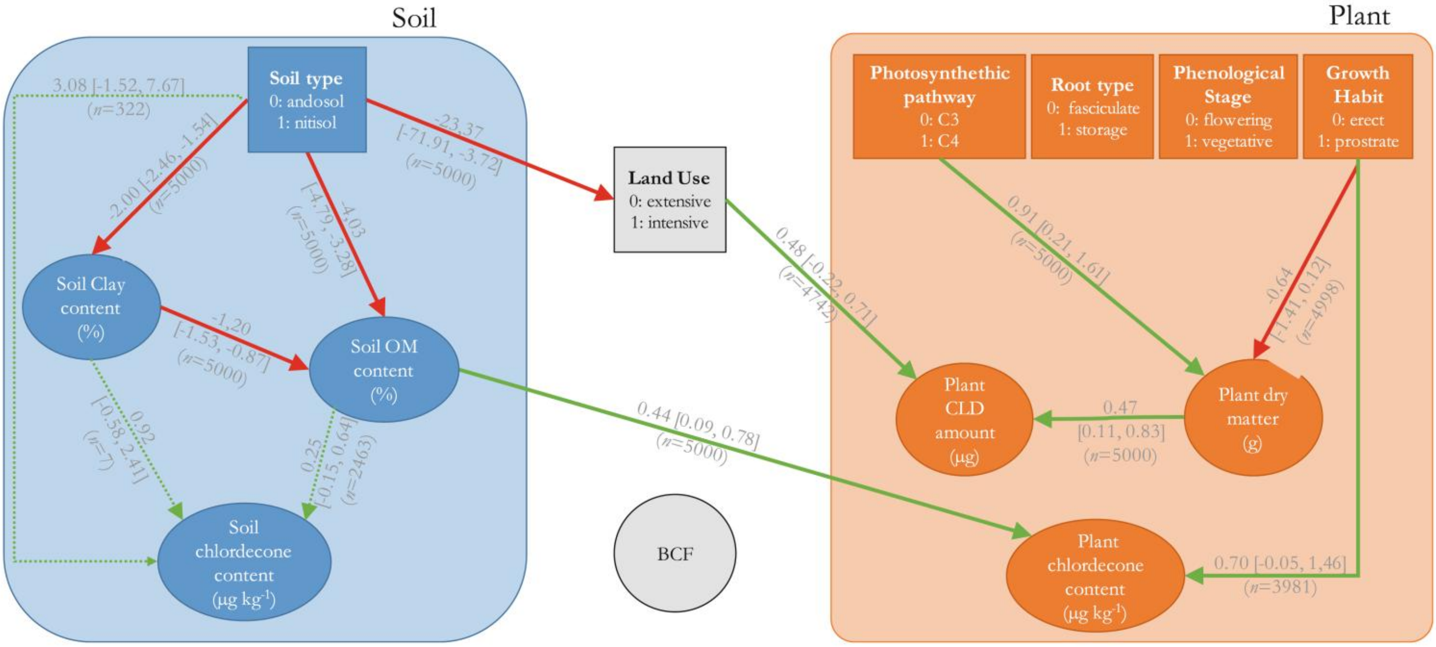

Figure 1. Final optimal additive Bayesian network (ABN) model, afteradjustment for overfitting (trimmed relationships are illustrated withdotted arcs), evaluating factors linked with the accumulation ofchlordecone in the plant in Guadeloupe (n = 25).

Binary variables are shown as squares and continuous variables as ovals. Green and red arrows stand for respectively positive and negative. Numbers represent log odds ratios (for binary variables) and mean effects (for continuous variables) of significant directly dependent variables in the ABN model with credible intervals in brackets and the number of arc recovery during MCMC simulations in parentheses. Arcs in the ABN model show only significant statistical dependency.

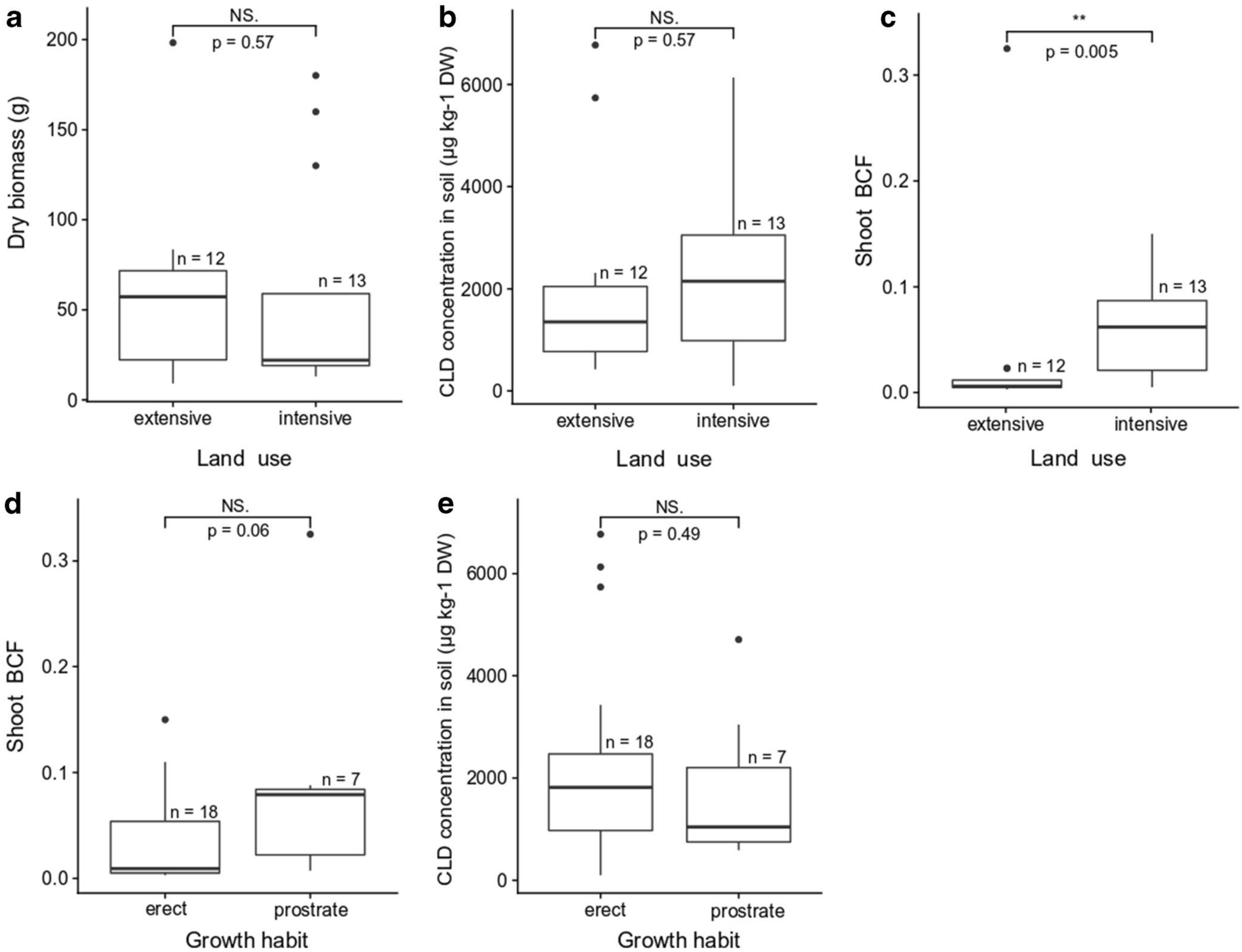

Figure 2. Graphical comparison of the two drivers (land use and growthhabit) highlighted by the directed acyclic graph (DAG) with the otherparameters tested in the study.

The boxplots at the top show the relationships between land use and dry biomass (a), soil concentration of CLD (b), and shoot BCF (c). The boxplots at the bottom show the relationship between growth habit and shoot BCF (d) and soil CLD concentrations (e). Wilcoxon tests were performed to check the significance of the results.

5. References

Aguilera PA, Fernández A, Fernández R, Rumí R, Salmerón A (2011) Bayesian networks in environmental modelling. Environ Model Softw 26(12):1376–1388

Babyak MA (2004) What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. Psychosom Med 66(3):411–421

Barriuso E, Houot S, Serra-Wittling C (1997) Influence of compost addition to soil on the behaviour of herbicides. Pestic Sci 49(1):65–75

Blanco-Canqui H, Lal R (2008). Principles of soil conservation and management. Springer Netherland. ISBN 978-1-4020-8708-0, 620 pp.

Cabidoche YM, Clermont-Dauphin C, Lafont A, Sansoulet J, Cattan P, Achard R, Caron A, Chabrier C (2006) Stockage dans les sols à charges variables et dissipation dans les eaux de zoocides organochlorés autrefois appliqués en bananeraies aux Antilles: relation avec les systèmes de culture. Ministère de l'Ecologie et du Développement Durable, p. 100

Cabidoche YM, Achard R, Cattan P, Clermont-Dauphin, C, Massat F, Sansoulet J (2009) Long-term pollution by chlordecone of tropical volcanic soils in the French West Indies: A simple leaching model accounts for current residue. Environ Pollut 157:1697–1705

Cabidoche YM, Lesueur-Jannoyer M (2012) Contamination of harvested organs in root crops grown on chlordecone polluted soils. Pedosphere 22(4):562–571

Chaussonnerie S, Saaidi PL, Ugarte E et al (2016) Microbial degradation of a recalcitrant pesticide : chlordecone. Front Microbiol 7:2025– 2037

Clostre F, Lesueur-Jannoyer M, Achard R, Letourmy P, Cabidoche YM, Cattan P (2014) Decision support tool for soil sampling of heterogeneous pesticide (chlordecone) pollution. Environ Sci Pollut Res 21(3):1980–1992

Clostre F, Letourmy P, Lesueur-Jannoyer M (2015) Organochlorine (chlordecone) uptake by root vegetables. Chemosphere 118:96–102

Collas C, Mahieu M, Tricheur A, Crini N, Badot P-M, Archimède H, Rychen G, Feidt C, Jurjanz S (2019) Cattle exposure to chlordecone through soil intake. The case-study of tropical grazing practices in the French West Indies. Sci Total Environ 668:161–170. https://doi.org/10.1016/j.scitotenv.2019.02.384

Collins C, Fryer M, Grosso A (2006) Plant uptake of non-ionic organic chemicals. Environ Sci Technol 40(1):45–52. https://doi.org/10. 1021/es0508166

Cornet D, Sierra J, Tournebize R, Gabrielle B, Lewis FI (2016) Bayesian network modeling of early growth stages explains yam interplant yield variability and allows for agronomic improvements in West Africa. Eur J Agron 75:80–88

Darwiche A (2009) Modelling and reasoning with Bayesian networks. Cambridge University Press, Cambridge

Friedman N, Goldszmidt M, Wyner A (1999) Data analysis with Bayesian networks: A bootstrap approach. In UAI, 1999

R Development Core Team (2017) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. http://www.R-project.org

Dorel M, Roger-Estrade J, Manichon H, Delvaux B (2000) Porosity and soil water properties of Caribbean volcanic ash soils. Soil Use Manag 16:133–140

Dreicer M, Hakonson TE, White GC, Whicker FW (1984) Rainsplash as a mechanism for soil contamination of plant surfaces. Health Phys 46:177–187

Drury B, Valverde-Rebaza J, Moura MF, de Andrade LA (2017) A survey of the applications of Bayesian networks in agriculture. Eng Appl Artif Intell 65:29–42

Fournet J (2002) Flore illustrée des phanérogames de Guadeloupe et de Martinique. Cirad and Gondwana Editions, p 2538

Jensen FV, Nielsen TD (2007) Bayesian networks and decision graphs. Springer-Verlag, New York

Jurianz S, Collas C, Lastel ML, Godard X, Archimède H, Rychen G, Mahieu M, Feidt C (2017) Evaluation of soil intake by growing creole young bulls in common grazing systems in humid tropical conditions. Animal: 1–9. https://doi.org/10.1017/S1751731116002755

Lazzeri Y, Mouhoud ME (2010). Prospective territoriale en terre de développement durable :une application à la région Guadeloupe. Diasporas protéiformes. Études caribéennes, n°16. URL : http:// journals.openedition.org/etudescaribeennes/10998 ;Accessed 15 august 2010 :https://doi.org/10.4000/etudescaribeennes.10998

Lewis FI, McCormick BJJ (2012) Revealing the complexity of health determinants in resource-poor settings. Am J Epidemiol 176(11): 1051–1059

Liber Y, Letondor C, Pascal-Lorber S, Laurent F (2018) Growth parameters influencing uptake of chlordecone by Miscanthus species. Sci Total Environ 624:831–837

Lorber-Pascal S, Létondor C, Liber Y, Jamin EL, Laurent F (2016) Chlordecone transfer and distribution in maize shoots. J Agric Food Chem 64:409–415

Mouvet C, Dictor M-C, Bristeau S, Breeze D, Mercier A (2016) Remediation by chemical reduction in laboratory mesocosms of three chlordecone-contaminated tropical soils. Environ Sci Pollut Res 24:25500–25512

ORSTOM (Office de la recherche scientifique et technique outre-mer) (1981) Carte des sols n°2838. (IN) Les ressources en eau de surface de la Guadeloupe : figures - annexes /dressé par P. Chaperon ; d'après carte des sols de F. Colmet Daage http://sphaera. cartographie.ird.fr/carte.php?num=2838

Pittavino M, Lewis F, Furrer R (2016). abn: an R package for modelling multivariate data using additive Bayesian networks. The Comprehensive R Archive Network (CRAN), pp. 1–37

Pittavino M, Dreyfus A, Heuer C, Benschop J, Wilson P, Collins-Emerson J, Torgerson PR, Furrer R (2017) Comparison between generalized linear modelling and additive Bayesian network; identification of factors associated with the incidence of antibodies against Leptospira interrogans sv Pomona in meat workers in New Zealand. Acta Trop 173:191–199

Rowley MC, Grand S, Verrecchia EP (2018) Calcium-mediated stabilisation of soil organic carbon. Biogeochemistry 137(1–2):27– 49

Sierra J, Causeret F, Diman JL, Publicol M, Desfontaines L, Cavalier A, Chopin P (2015) Observed and predicted changes in soil carbon stocks under export and diversified agriculture in the Caribbean. The case study of Guadeloupe. Agric Ecosyst Environ 213:252–264

Trivedi P, Wallenstein MD, Delgado-Baquerizo M, Singh BK (2018) Chapter 3 - Microbial modulators and mechanisms of soil carbon storage. Soil Carbon Storage, pp:73–115. https://doi.org/10.1016/ B978-0-12-812766-7.00003-2

Woignier, T., Clostre, F., Cattan, P., Levillain, J., Cabidoche, Y.-M., Lesueur Jannoyer, M. (2014). Diagnosis and management of field pollution in the case of an organochlorine pesticide, the chlordecone. In: Hernandez Soriano, Maria C., dir., Environmental risk assessment of soil contamination (p. 615-636). Rijeka, HRV : InTech. 918 p., https://doi.org/10.5772/57263

Woignier T, Clostre F, Fernandes P, Rangon L, Soler A, Lesueur-Jannoyer M (2016) Compost addition reduces porosity and chlordecone transfer in soil microstructure. Environ Sci Pollut Res Int 23:98–108

6. Acknowledgements

We thank Dr. Jorge Sierra for constructive discussions about soil carbon dynamics associated with land management practices in Guadeloupe.