Multivariate tiered approach to highlight the link between large-scale integrated pesticide concentrations from POCIS and watershed land-uses

Marion Bernard[a], Sebastien Boutry[b], Robin Guibal[c], Soizic Morin[b], Sophie Lissalde[c], Gilles Guibaud[c], Margaux Saüt[d], Jean-Pierre Rebillard[e], Nicolas Mazzella[b]*

[a] *SGS Life Science Services, 90 Av. des Hauts de la Chaume BP 28, 86281 Saint-Benoît*

[b] *Inrae, UR EABX, 50 Avenue de Verdun, 33612 Cestas, France*

[c] *Université de Limoges, E2Lim, 123 Avenue Albert Thomas, Limoges, Cedex 87060, France*

[d] *DREAL Occitanie, 1 Rue de la Cité Administrative, 31000 Toulouse*

[e] *Agence de l'Eau Adour-Garonne (AEAG), 90 Rue du Feretra, 31078 Toulouse CEDEX 4, France*

*** Corresponding author at: Inrae, UR EABX, 50 Avenue de Verdun, 33612 Cestas, France**

E-mail address: nicolas.mazzella@inrae.fr

ABSTRACT

This paper describes an automatized multi-step methodology in order to identify the relationships between integrative pesticide quantifications and land-use on a given watershed. This methodology contains multivariate statistical analyses such as hierarchical cluster analysis (HCA) and principal component analysis (PCA), which are commonly used for the interpretation of complex geospatial datasets. A large amount of pesticide concentration data were collected along 1-year monitoring in 2016, for 50 sites located on the Adour Garonne basin (South-West France). For those sampling sites, concentrations of 37 selected pesticides were investigated during six periods of 14-days immersion of integrative samplers. Specifically, the sampling devices used were Polar Organic Chemical Integrative Sampler (POCIS), providing time-weighted average concentration estimates. For each studied site, the associated watershed and its land-use repartition were determined based on the Corine Land Cover 2012 and geographical information system (GIS) aggregation of data. The HCA clustered the 50 sites into five groups with similar main land uses. After that, the datasets of pesticide integrated concentration and land use repartition were analyzed in a PCA. The key variables (pesticide distribution and concentrations) responsible for sampling site discrimination showed consistent patterns of distribution with specific land uses. In order to confirm these observations, pesticide fingerprints (based on the waffle method) of sites with contrasted land use relative to the surface areas were compared. These fingerprints confirmed that there was different and specific patterns, visible at a glance, of pesticide occurrence in surface water, in relation with their initial use at the catchment level. This method allowed identifying sources of contamination that could be interesting to prevent or contain pesticide pollutions beyond simply acting on the most at-risk areas.

Keywords: passive sampling, pesticides, multivariate analysis, river catchments, tiered methodology

1. Introduction

The use of organic pesticides in an agricultural or non-agricultural context leads to the contamination of the aquatic environment. This anthropogenic influence has an impact on aquatic ecosystems, and also human beings. Within this frame of reference, monitoring networks are performed in order to evaluate the water quality for pesticides. These networks acquire each year large amounts of data (*e.g.* monthly pesticide concentrations). These data are usually compared to regulatory thresholds such as those from the Water Framework Directive (WFD [1]) in order to evaluate the quality of water bodies, and highlight the vulnerable and degraded areas. In addition, these data could be used to identify the impact of human activities regardless of their intensity. This means finding a way to identify sources of contamination for implementing corrective actions to prevent pesticide pollutions on the most at-risk areas. However, accessing this type of information with a large amount of data collection can be time-consuming and laborious.

A study performed by Macary, et al. [2] used the agricultural pressure (land use, farmers practices) of different size watersheds and soil characteristics relating to the environmental vulnerability of the surrounding surface water environment (slope, pedology of agricultural parcels…) to establish an indicator of pesticides contamination risk (Phytopixal method). With this multi-criteria method, the authors proposed cartographic projections of the pesticides contamination risks for the studied watershed (*i.e.* the Coteaux de Gascogne, South-West of France). Another study performed by Morin, et al. [3] on the same watershed demonstrated that the mapping of pesticide contamination risk with the use of the Phytopixal method allowed obtaining a relevant estimate of pesticide real exposure *in situ* (*i.e.* correlated with pesticide concentration measured), and further assessment of toxic impacts on the diatom communities in the Neste river system. Several other studies using multivariate statistical analyses (hierarchical clustering and/or principal component analysis) were successful in establishing links between surface water chemistry (*e.g.* nutrients, trace elements, pesticides) and land uses [4-8]. In the latter studies, grab sampling was used to describe surface water contamination ny chemicals. However, this sampling strategy provides data with a lack of temporal representativeness since it corresponds to a point-in-time snapshot, and then contamination fluctuations could be missed [9-10]. Thus, with the aim of establishing relationships between global estimates of land use (low temporal resolution) and overall water quality, integrated measurements of water quality would likely improve correlations over spot samplings.

In our study, an alternative sampling approach which corresponded to the use of passive samplers was performed. This sampling strategy which consists of immersing a device in water for a fixed period allows the *in situ* pre-concentration of target compounds, and then provides a time weighted average concentration (TWAC) with integration of contamination fluctuations [11-12]. For 50 sites located on the river of the Adour-Garonne Basin (South-West of France), the concentrations of 37 selected pesticides were investigated during six periods of 14 days spread over the year 2016, by the use of the Polar Organic Chemical Integrative Sampler (POCIS). POCIS are widely used for the sampling of moderate polar pesticides with $0 < \log K_{ow} < 4$ [13]. Due to TWAC estimates obtains, such passive samples allows an increase of both number of quantified pesticides and detection frequencies compared to grab sampling [14-18]. In addition, large-scale spatial and temporal trend studies demonstrated that complex mixtures of pesticides accumulated by passive samplers can be well correlated with land use [19-20].

By using the large amount of pesticide data collected during this one year monitoring, a visual and automatized methodology was proposed in order to link pesticide contaminations of river waters and watershed land uses. To do that, information on land use of each catchment area containing the various sampling sites was determined from Corine Land Cover 2012. Land uses and pesticide datasets were analyzed against each other with multivariate statistical methods, such as hierarchical cluster and principal component analysis. Furthermore, pesticide fingerprints of each sampling sites was established by using "waffles". Finally, this study aimed to propose a tiered methodology in order to highlight the most threatened areas and periods in terms of pesticide contaminations.

2. Materials and methods

2.1. Data acquisition

2.1.1. Study area: the Adour-Garonne basin

The Adour-Garonne basin is located in the southwest of France and covers an area of 117,650 km². It is composed of 116,817 km of rivers and of a coastline strip of 650 km. This basin has a population of *c.a.* 7,000,000 inhabitants with pronounced rural character (30 % of the population), 35 cities with more than 20,000 inhabitants each (28 % of the population) and two metropolises (Toulouse and Bordeaux) with *c.a.* 750,000 inhabitants each. Concerning its land use, and according to the Corine Land Cover 2012 (CLC 2012), 55 % of the basin's surface area corresponds to agricultural areas (*i.e.* crops, vineyards, orchards…) and 40 % to forest areas (Figure 1). The remaining 5 % are divided between artificial areas (*e.g.* urban, industrial areas), wetlands (*e.g.* marshes) and water surfaces (*e.g.* lakes). For this study, 50 sampling sites from the Water Framework Directive network were selected. These sampling sites (Figure 1) were characterized by a diversity of land uses, implying different water pesticide contamination profiles [14].

2.1.2. *Pesticides quantification in surface water with POCIS*

The 50 selected sites within the Adour-Garonne basin were sampled during 6 periods of 14 days evenly distributed between March and December 2016 by the use of passive samplers, in order to obtain pesticide contamination levels. In this study, the Polar Organic Chemical Integrative Sampler (POCIS) was used in his "Pharmaceutical" configuration [13, 21]. With this sampler, 37 neutral and moderately polar pesticides ($0.57 < \log K_{ow} < 4.14$) were investigated. These 37 selected compounds (Table 1) included different chemical families (*e.g.* chloroacetamides, ureas, etc.) and biological activities (*i.e.* herbicides, fungicides, insecticides, as well as their respective known metabolites) in order to cover a large range of treatments. All the sample processing and analyses were described in Bernard, et al. [14]. Briefly, before their field deployment, POCIS were prepared at the laboratory with 200 mg of Oasis HLB sorbent (30µm particle size, 810 m^2g^{-1}, divinylbenzene *N*-vinyl-pyrrolidone, Waters, France) enclosed between two microporous polyethersulfone membranes (PES – 90 mm diameter and 0.1µm pore size, PALL®, VWR, France) and compressed by two holder washers. After the exposure period, POCIS were disassembled and the pesticides were extracted from the sorbent with 3 mL of methanol then 3 mL of methanol:ethyl acetate 75:25 (v/v). The 6 mL extract obtained was evaporated with a Speedvac system (1h30 at 60°C – Thermo Fisher Scientific, France) after adding 10 µL of internal standards. Before analysis, the samples extract was reconstituted with 1 mL of ultrapure water:acetonitrile 90:10 (v/v). The analysis of pesticides from POCIS was performed with two liquid chromatography apparatus coupled with high-resolution mass spectrometry (*i.e.* HPLC-MS/MS and UHPLC-Q-ToF). For the both analytical methods, the compound-dependent instrumental quantification limits (IQL in µg L^{-1}) were determined with method validation (NF T 90-210, AFNOR [22]) and grouped in Table S 1. In addition and according to the method of Poulier, et al. [18], the compound and analytical technique dependent quantification limits for POCIS (QL_P in µg L^{-1}) were calculated and also grouped in Table S 1.

For each sampling site and period, POCIS analysis provided a time-weighted average concentration of each pesticide in water ($\overline{C_w}$ in µg L^{-1}). At last, this 1-year pesticide monitoring provided a large amount of data, which was equal to 11,100 water data on pesticide concentrations acquired (*i.e.* 50 sites × 6 periods × 37 pesticides).

2.1.3. *Preparation of the pesticide datasets*

Because of the large amount of data collected, a rigorous and automatized data preparation methodology was implemented. In this study, the raw data was composed by the pesticide concentrations measured in the POCIS receiving phases (C_{POCIS} in ng per POCIS) after their field deployment. These concentrations were grouped in the same database for all stations, campaigns and pesticides. Then, these raw data followed the processing described in Figure S 1, which was performed with two software programs: Microsoft Excel (version 14.0.7212.5000) and R software (R Core Team, 2017). This preparation step was essential to check the data before their use in the methodology development. Firstly, they were sorted according to the IQL (in µg L^{-1} - Table S 1) and normalized by the mass of the receiving phase recovered inside the POCIS ($M_{Sorbent}$ in g) to obtain C_{POCIS} in ng g^{-1}. Second, the time-weighted average concentrations ($\overline{C_w}$ in µg L^{-1}) of each pesticide in water were calculated as described in Bernard et al. (2019), and then sorted with the QL_P (in µg L^{-1}, Table S 1). Afterwards, quantification frequencies (QF in %) for each studied pesticide (n = 37) and for all stations and periods combined were calculated (*i.e.* annual QF). When annual QF was lower than 10 %, the compound was removed from all the dataset, because contamination levels measured were not significant and these variables provided no discrimination power between water qualities from different locations, in contrast to other compounds with higher QF. This approach also allows ensuring robustness of statistical treatments by removing substantial noise from the analysis. After this step, the number of studied pesticides or metabolites decreased from 37 to 23 (Table 1). Consequently, the total number of data was reduced to 7038 $\overline{C_w}$ values in µg L^{-1}.

2.1.4. *Data analysis / methodology developed*

The methodology developed was adapted from several studies that used multivariate analyses to better characterize the sources of water chemicals contamination [4-5, 7-8]. Because of the large amount of data collected, their processing and interpretation are often difficult. In this context, multivariate statistics are useful approaches since they decrease the number of components in a complex dataset by identifying key variables responsible for patterns of sampling sites [23].

2.1.5. *Site classification based on watershed land use.*

For each site, the associated watershed was determined using the GRASS plugging and QGIS software (version Las Palmas 2.18). To do that, the "r.water.oulet" function was used to delimit the watershed using the outflow coordinates and the digital elevation model (DEM). Then, the area of each watershed was calculated. At last, the percentages of land uses of each watershed were evaluated by making an intersection between the CLC 2012 layers (pixel size: 500 m * 500 m) and the determined watershed. All the sampling sites information is available in Table 2.

A hierarchical cluster analysis (HCA) was performed on the site data in order to identify groups of sites that had similar land use profiles. A classification scheme using the Euclidean distance for similarity measures between percentages of land use was performed. The Ward's method was used for the establishment of the links between sites to improve distinctive power of the classification [8, 24].

2.1.6. Relationships between water pesticides contamination and watershed land use.

In an effort to show the relationships between quantified pesticides and land use at the sampling site, principal component analyses (PCA) were performed on the pesticides concentrations measured for each site and period, after standardization of the POCIS data ($\overline{C_w}$ in µg L^{-1}) by a Yeo-Johnson transformation [25]. This transformation accounted for the high variability of the contamination levels between sites, periods and compounds [14, 18, 26] and reduced the influence of extreme values on variance [23]. Moreover, the data transformed this way achieved a normal distribution, which is required to conduct this multivariate statistical analysis [27]. In this analysis, the 23 studied pesticides were considered as variables, and individuals corresponded to each sampled site at each period (*i.e.* period-station code).

The axes extracted from the PCA performed with the contamination levels measured ($\overline{C_w}$ in µg L^{-1}) for the 23 studied pesticides and all periods (n=6) and sites (n=50) were then rotated with the Kaiser Varimax criterion[28] in order to enhance the interpretation. This Varimax rotation makes it possible to bring the groups of variables closer to the axes that allow obtaining a better evaluation of their contribution, by associating a limited number of factors to each variable. The correlation between the variables (*i.e.* pesticides) and the defined axes are expressed by the loadings (or eigenvalues). For each axis and variable, these loadings were grouped in Table S 2 and allowed highlighting the most discriminating variables.

At last, site groups obtained with HCA based on land use were projected on the PCA two dimensional plots in order to picture rough correlations between quantified pesticides and sampling sites (land use).

2.1.7. Specific pesticide fingerprint and relationship with land use.

To go more in depth in understanding and visualizing the relationships between land use and pesticides contamination, sites exhibiting specific profiles based on the PCA were focused on. Specifically, sites with contrasted profiles (forest, agriculture, and vineyard) were selected. The seasonality and magnitude of pesticide contamination profiles were illustrated using the "waffle" method [29]. These multivariate analyses and the pesticides fingerprint were performed with R Software (R Core Team, 2017), packages "ade 4"[30], "adegraphics" [31] and "cluster" [32] ; the waffles were drawn using "ggplot2" [33] and "treemapify"[34] packages.

3. Results and discussion

3.1. Site classification based on the watershed land use

Hierarchical cluster analysis (HCA) was initially performed on the sampling site dataset, based on the percentage of land use associated with the respective watershed of each studied site (Table 2). The optimal clustering (Figure 2) discriminated five major groups of sites. Group I gathers sites from watersheds dominated by agricultural areas like corn, sunflower and winter or spring wheat. Group II includes sites consist in a combination of forests, agricultural areas and pastures with equivalent surface area percentages (Table 2). Groups III, IV and V represent sites from forest, pasture and vineyard dominated watersheds, respectively.

The Group I dominance (23/50 sites) was expected because the Adour-Garonne basin has a large part of field crops (*i.e.* 55%, Figure 2 and Table 2). The Group II (15/50 sites) and Group III (6/50 sites) are consistent because forests represent the second most important part of this territory (*i.e.* 40%). Lastly, there are only three sites in either group IV or V, that was also consistent with overall land use of the basin, with the two main vineyard areas around Cognac and Bordeaux cities, and some small areas of cattle breeding in the medium altitude mountains of Massif Central. Moreover, it would be expected some sites to be locally dominated by artificial territories (*e.g.* Bordeaux and Toulouse metropolises), but in fact, it is never reached high percentages of land use of such watersheds. For each group obtained with the HCA, the inter-group variability was at least 54 % (group IV). This variability, due to the absence of buffer zone determination for each watershed, may lead to some discrepancies between the dominant portion of land use in the watershed and that near the sampled site.

3.2. Linking pesticide contamination with watershed land use

A principal component analysis (PCA) was performed with the contamination levels measured for the 23 studied pesticides for all periods (n=6) and sites (n=50). This PCA projection (Figure 3 and Figure 4) allowed to visualize local trends in pesticide concentrations and to establish correlations between pesticide contamination patterns and the land use groups defined before (Table 2). Six axes with eigenvalues greater than 1 (Kaiser Criterion) were selected, they explained 73.2 % of the total variance of the dataset. The Varimax-rotated axes loading matrices of pesticides are grouped in Table S 2. The first axis (A1) explained 18.2 % of the total variance and had strong loadings (≥ 0.75) related to flurtamone (FTM), epoxyconazol (EPX), cyproconazole (CYP), metolachlor (MTC), and moderate loadings (between 0.50 and 0.75) with tebuconazole (TBZ), atrazine (ATZ) and desethylatrazine (DEA). The second axis (A2) explaining 17.5 % of the total variance had strong loadings with desethylterbuthylazine (DET), desisopropylatrazine (DIA), simazine (SMZ), terbuthylazine (TUZ), and moderate loadings of desethylatrazine (DEA). A two-dimensional plot of A1 against A2 (Figure 3) was performed in order to show the correlation between variables and individuals.

Figure 3 b allowed the discrimination of two groups (previously obtained with a HCA, Table 2 and Figure 2): I and V, which mainly correspond to agricultural areas and vineyards, respectively. The other groups (II, III and IV) were not clearly separated from Group I and did not seem correlated with any of the pesticides having strong or moderate loadings on these first two axes. The first axis structured Group I which includes sites with agricultural areas dominated watersheds (Table 2) and was correlated with CYP, EPX, FTM, and MTC (Figure 3 a). The two fungicides (CYP and EPX) and the two herbicides (MTC and FTM) are often used in agricultural context in France, especially for the treatment of cereal, corn and sunflower crops (https://ephy.anses.fr). The moderate loadings of TBZ can be explained by the fact that this fungicide is used for both the treatment of cereal crops and vineyards (less specific use than CYP, for example). Then, the moderate loadings of the prohibited ATZ (banned in 2003 in France) and its main metabolite DEA illustrate their likely high persistence in soils [35-37] and recurrent release in waters, since decades after its ban it is still quantified in surface waters [14, 26]. Globally, the correlations between these pesticides and agricultural areas were consistent.

A gradient of concentration was observed along A1, with the individuals located on the positive values (Figure 3b). This gradient was explained by the seasonal trends of the contamination levels measured, which are linked to specific treatment periods. In agreement with Bernard, et al. [14], metolachlor concentrations were the highest in May, as the result of the treatment of corn crops at this period (*i.e.* March to June). Group V sites (*i.e.* vineyards dominated watersheds, Table 2) were characterized by SMZ and its metabolite DIA, as well as TUZ and its metabolite DET, along Axis 2. These two pesticides which had been used for the treatment of vineyards are banned since 2003 for SMZ, and 2008 for TUZ. However, they are still quantified in water together with their metabolites [16], as also shown by our data. In a lesser extent, the loading with A2 of norflurazon (NFZ), an herbicide typicaly used in vineyards during the same period, appears to be consistent in terms of residulal contamination. As previously mentionned for atrazine, the presence of these compounds in surface water (and also groundwater) despite their prohibition for several years can be explained by their persistence in soil after their application [38-39]. For TUZ, a study performed by Carretta, et al. [40] demonstrated that its dissipation from soil is strongly influenced by the granulometry, organic carbon content and depth. Consequently, quantification of SMZ and TUZ in surface waters can be explained by their remobilization from soil which can depends on soil texture, physico-chemical properties of the compound (including log K_{oc} and log K_{ow}, Table 1), as well as climatic conditions such as rainfall. A study performed by Hildebrandt, et al. [5] in three sampling sites of North Spain showed the impact of intensive vineyard cultivation on surface and groundwater quality. In this study, they also demonstrated correlations between these four same pesticides (SMZ, DIA, TUZ and DET) and the presence of vineyards, in agreement with our observations. Complementary information was provided by the other axes. Axis 3 (10.8% of the variance) had strong loadings related to chlortoluron (CTU), isoproturon (IPU), and moderate loadings with metazachlor (MTZ) and imidacloprid (IMI) (Table 2), also correlated with group I. This observation is consistent because they are generally all used for the treatment of cereals crops, while Axis 4 (Figure 4) showing moderate loading with dimetomorph (DMM), carbendazime (CBZ) and NFZ correlated with vineyard sites (Group V). NFZ and CBZ were banned in 2004 and 2008, respectively, but as previously mentioned for atrazine they are still quantified in surface waters due to their likely remobilization from the soil [41]. DMM was also quantified in a vineyard watershed by [42]. Globally, the pesticides correlated with Axis 1 and agricultural areas are mainly herbicides (CTU, IPU and MTZ) as found by Van Metre, et al. [19], while those correlated with Axis 3 are mainly fungicides (DMM and CBZ). These observations are in agreement with specific pesticides use because vineyards are more sensitive to fungal development than corn crops, for example.

To conclude, this PCA revealed that the pesticides quantified in the rivers with POCIS are regularly linked with the main land use of the corresponding watersheds, especially in the case of agricultural areas and vineyards (groups I and V, respectively – Table 2). Conversely, the groups II, III and IV were not clearly associated with the contamination profile found. Such result can be explained by the fact that these sites are assumed either to be slightly or not contaminated by pesticides (*i.e.* areas dominated by forest or pastures), or phytosanitary treatments with various compounds in the case of watershed characterized by mixed land uses. To confirm these statistical analyses and go more in depth in contamination profiles for selected sites identified by the PCA, site fingerprints were further addressed using the waffle method.

3.3. Confirmation of statistical observations with pesticide contamination fingerprints

The PCA showed that the group III (*i.e.* forests dominated watershed) was not correlated with any peculiar pesticides, suggesting that the corresponding stations are slightly, or even not contaminated. One site among this Group III was chosen in order to illustarte this observation (*i.e.* site n° 5216210, Table 2 and Figure 5 a). For all sampling periods over the year 2016, the pesticide fingerprint of this site showed that it was almost not contaminated by the 23 searched pesticides (Figure 5 b). Only acetochlor (ATC) and alachlor (ALA) were detected in May, but with very low contamination levels, near the QL_P (Table S 1), which can be considered as residual ultra-traces (compounds banned since 2013 and 2008, respectively). The proximity of some agricultural and artificial areas only at the outlet of the watershed does not seem affecting the quality of this watercourse. This observation probably reflects the ability of the river to progressively dilute the pesticides mainly used upstream of the monitoring area.

The previous PCA discriminated some sites among the group I (*i.e.* agricultural dominated watershed) according to the axes 1 and 3. The pesticide fingerprints of two sites from the group I with different land use repartition were considered hereafter. For instance, the catchment associated to the site n°5080960 (Figure 6 a) is composed of 81 % of agricultural areas, 15 % of forests and 2 % of pastures and vineyards. Adversely, the watershed of site n°51156950 (Figure 7 a) is composed by 82 % of agricultural areas, 12 % of artificial areas (*e.g.* urban areas), 3 % of forest and 2 % of pastures. The pesticide fingerprints (parts b of Figure 6 and Figure 7) can be expected to exhibit high similarities because of agricultural areas dominating within the group I. Nonetheless, some features due to the minor part of their other land use can be noticed, with higher artificial surfaces for the Hers Mort station (Figure 7 a). In these two cases, fingerprints highlighted complex mixtures of pesticides, with the frequent occurrence (≥ 50 % of the 6 sampling periods, for a selected chemical) of ALA, ATZ, CTU, FTM, MTC and SMZ as herbicides, or DMM, EPX and TBZ as fongicides. These results showed that the watercourses were contaminated, at least once along the year, by almost all the 23 studied pesticides, in contrast with results obtained previously for a representative site of the Group III (Figure 5 b). In each site, pesticides used specifically in agricultural context were quantified (*e.g.* MTC, TBZ, EPX,…) with varying contamination levels over the year 2016. The month of May provided the highest number of quantified pesticides, with scores of 20/23 and 18/23 pesticides, and the highest contamination levels with average concentrations of 6.44 µg L^{-1} and 4.20 µg L^{-1} (*i.e.*

sums of each quantified contaminant), for the site n°5080960 and the site n°5156950, respectively.

These two fingerprints also demonstrated some differences related to the minority land use. Indeed, site n°5156950 (Figure 7) exhibited IMI with high quantification frequency and stable contamination levels, in contrast with site n°5080960 (Figure 6). Until 2016, this insecticide was be used either in agricultural contexts, like cereal crops, or for non-agricultural purposes, like the treatment of private gardens or pets. For site n°5156950, the IMI residual and stable contamination levels can be associated with the non-agricultural context because of the proximity of artificial areas near the sampling site (Figure 7 a) [14]. In contrast, IMI contamination of site n°5080960 was rather due to the use of this insecticide in an agricultural context. Actually, it was quantified during Spring and November only, which corresponds to its typical application periods. The same observation can be made for DIU, banned since 2008 in agricultural treatments, which is now used only as a biocide for fouling treatment on house materials, and was observed all through the year at site n°5156950. The site n°5080960 also exhibits some trace levels of residues that were less, or even not quantified on site n°5156950, such as SMZ and NFZ. Such herbicides were characteristic of vineyard treatments in the early 2000s in France, and probably correspond here to remnant background contaminations. These results are supported by some vineyards still occurring on the Gupie at Sainte Bazeille station (*i.e.* 2 % of land use, Table 2).

This kind of observation is more pronounced with the pesticide fingerprint obtained for group V sites, which were representative of catchments characterized by higher surface of vineyards (Table 2, Figure 8 a). The waffles of the site n°5075900 (Figure 8 b) exhibited a complex mixture of pesticides and metabolites containing NFZ, DMM, SMZ and DIA for each sampling period. But this time, contamination levels were higher than those previously observed for site n°5080960. This observation can be explained by a likely dilution effect occurring at sampling sites located far from the source of contamination (case of site n°5080960) [42]. In addition, TUZ and its degradation product DET were also quantified during all periods [5], which was not the case for the sites of group I. Conversely, CYP, EPX and FTM were barely or even not quantified on the site n°5075900, which involves their specific uses in vineyard agricultural context. On the other hand, some pesticides typically used on field crops were also quantified on this site, such as MTC, but contamination levels were less marked in this case, with a maximal value of 0.5582 µg L^{-1} (Figure 8 b). TBZ was quantified as well during each period, but with contamination levels closer to agricultural sites of Group I. This observation can be explained by the non-specific uses of this fungicide, as previously mentioned. In addition, other studies performed on either an agricultural [18] or a vineyard [42] watershed quantified this compound.

These fingerprints obtained with the waffle method combined with POCIS data confirmed that there is a different and specific pattern to the pesticide impacts of land uses on the surface water quality. These simple graphical representations of TWACs allow seeing quickly which contaminant was mainly quantified and when it occurred. In addition, the waffles allow a quick comparison of the pesticide fingerprints between sites.

4. Conclusions and perspectives

The purpose of this paper was to develop and propose a methodology for optimizing the interpretation of the links between land use and pesticides quantified in surface water, using a large dataset. Data used in this study were collected during a 1-year pesticide monitoring performed with the deployment of POCIS during 6 periods of 14 days, over 50 sampling sites. Each site was identified and delimited in function of its watershed, with the associated land-use aggregates. These two datasets were confronted with multivariate analyses such as hierarchical cluster analysis (HCA) and principal component analysis (PCA), which allowed identifying key variables responsible for patterns of sampling sites.

The HCA allowed grouping the 50 sites into 5 groups with similar main land uses. The PCA discriminated two groups of sites and showed that some pesticides were specific to a peculiar phytosanitary uses. To confirm these observations, pesticides fingerprints of four sites with opposing land use were compared. These fingerprints confirmed that there is a different and specific pattern related to the pesticide detections with POCIS and the land uses for the concerned catchments. For instance, these fingerprints quickly showed that forest-dominated watersheds display a slight contamination by selected pesticides. In addition, it appears that a minor share of the watershed land use may contribute to the water quality, and some dilution effects would occur when the source was far from the sampling site.

However, this methodology could be further improved to demonstrate even finer links between land-use and quantified contaminants with passive samplers. One option would be to apply a buffer of X km around the sampling site or around the linear of the watercourse, in order to reduce the size and spatial variability of the study area [2]. Another alternative could be to project the graphical land parcel register (RPG – https://www.data.gouv.fr/fr/datasets/registre-parcellaire-graphique-rpg-contours-des-parcelles-et-ilots-culturaux-et-leur-groupe-de-cultures-majoritaire/) in order to identify more precisely the crops near the sampling site, and to achieve a higher level of detail regarding contamination sources.

To conclude, this methodology could be considered to establish similar links for other type of contaminants, such as pharmaceutical residues. To do this, the TWACs of these contaminants in the water would be necessary, and then coupled with indicators or characteristic uses on territories.

Acknowledgments

This study was financed by the Adour-Garonne French Water Agency. Authors would like to thank Aurelie Moreira, Gwilherm Jan and Sylvia Moreira from Inrae, Julie Leblanc, Patrice Fondanèche, Karine Cleries and Robin Guibal from University of Limoges, and Jessica Partaud and Lomig Leborgne from the Adour-Garonne Water Agency for technical, field and laboratory assistance.

Conflict of interest

The authors declared no competing interest.

Funding

This work was supported by the Adour-Garonne Water Agency.

References

1. EU., Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a Framework for Community action in the field of water policy. Official Journal of European commission (Vol. 1999). **2000**.
2. Macary, F.; Morin, S.; Probst, J. L.; Saudubray, F., A multi-scale method to assess pesticide contamination risks in agricultural watersheds. . *Ecological Indicators* **2014,** *36*, 624–639.
3. Morin, S.; Bottin, M.; Mazzella, N.; Macary, F.; Delmas, F.; Winterton, P.; Coste, M., Linking diatom community structure to pesticide input as evaluated through a spatial contamination potential (Phytopixal): A case study in the Neste river system (South-West France). *Aquatic Toxicology* **2009,** *94*, 28–39.
4. Fitzpatrick, M. L.; Long, D. T.; Pijanowski, B. C., Exploring the effects of urban and agricultural land use on surface water chemistry, across a regional watershed, using multivariate statistics. *Applied Geochemistry* **2007,** *22*, 1825–1840.
5. Hildebrandt, A.; Guillamón, M.; Lacorte, S.; Tauler, R.; Barceló, D., Impact of pesticides used in agriculture and vineyards to surface and groundwater quality (North Spain). *Water Research* **2008,** *42*, 3315–3326.
6. Metcalfe, C. D.; Helm, P.; Paterson, G.; Kaltenecker, G.; Murray, C.; Nowierski, M.; Sultana, T., Pesticides related to land use in watersheds of the Great Lakes basin. . *Science of the Total Environment* **2009,** *648*, 681–692.
7. Tariq, S. R.; Shah, M. H.; Shaheen, N.; Khalique, A.; Manzoor, S.; Jaffar, M., Multivariate analysis of trace metal levels in tannery effluents in relation to soil and water: A case study from Peshawar, Pakistan. . *Journal of Environmental Management* **2006,** *79*, 20-29.
8. Yidana, S. M.; Ophori, D.; Banoeng-Yakubo, B., A multivariate statistical analysis of surface water chemistry data-The Ankobra Basin, Ghana. *Journal of Environmental Management* **2008,** *86*, 80–87.
9. Allan, I. J.; Vrana, B.; Greenwood, R.; Mills, G. A.; Knutsson, J.; Holmberg, A.; Guigues, N.; Fouillac, A. M.; Laschi, S., Strategic monitoring for the European Water Framework Directive. *Trends in Analytical Chemistry* **2006,** *25*, 704–715.
10. Novic, A. J.; Brien, D. S. O.; Kaserzon, S. L.; Hawker, D. W.; Lewis, S. E.; Mueller, J. F., Monitoring Herbicide Concentrations and Loads during a Flood Event: A Comparison of Grab Sampling with Passive Sampling. . *Environmental Science and Technology* **2017,** *51*, 3880–3891.
11. Bernard, M.; Boutry, S.; Tapie, N.; Budzinski, H.; Mazzella, N., Lab-scale investigation of the ability of Polar Organic Chemical Integrative Sampler to catch short pesticide contamination peaks. . *Environmental Science and Pollution Research* **2022,** *29*, 40-50.
12. Vrana, B.; Mills, G. A.; Allan, I. J.; Dominiak, E.; Svensson, K.; Knutsson, J.; Morrison, G.; Greenwood, R., Passive sampling techniques for monitoring pollutants in water. *Trends in Analytical Chemistry* **2005,** *24*, 845-868.
13. Alvarez, D. A.; Petty, J. D.; Huckins, J. N.; Jones-Lepp, T. L.; Getting, D. T.; Goddard, J. P.; Manahan, S. E., Development of a passive, in situ, integrative sampler for hydrophilic organic contaminants in aquatic environments. *Environmental Toxicology and Chemistry* **2004,** *23* (7), 1640-1648.
14. Bernard, M.; Boutry, S.; Lissalde, S.; Guibaud, G.; Saüt, M.; Rebillard, J.-P.; Mazzella, N., Combination of passive and grab sampling strategies improves the assessment of pesticide occurrence and contamination levels in a large-scale watershed. . *Science of The Total Environment* **2019,** *651*, 684–695.
15. Criquet, J.; Dumoulin, D.; Howsam, M.; Mondamert, L.; Goossens, J.-F.; Prygiel, J.; Billon, G., Comparison of POCIS passive samplers vs. composite water sampling: A case study. *Science of The Total Environment* **2017,** *609*, 982-991.

16. Guibal, R.; Lissalde, S.; Charriau, A.; Poulier, G.; Mazzella, N.; Guibaud, G., Coupling passive sampling and time of flight mass spectrometry for a better estimation of polar pesticide freshwater contamination: Simultaneous target quantification and screening analysis. *Journal of Chromatography A* **2015,** *1387*, 75-85.
17. Lissalde, S.; Mazzella, N.; Mazellier, P., Polar organic chemical integrative samplers for pesticides monitoring: Impacts of field exposure conditions. *Science of The Total Environment* **2014,** *488-489*, 188-196.
18. Poulier, G.; Lissalde, S.; Charriau, A.; Buzier, R.; Delmas, F.; Gery, K.; Moreira, A.; Guibaud, G.; Mazzella, N., Can POCIS be used in Water Framework Directive (2000 / 60 / EC) monitoring networks? A study focusing on pesticides in a French agricultural watershed. *Science of the Total Environment* **2014,** *497-498*, 282-292.
19. Van Metre, P. C.; Alvarez, D. A.; Mahler, B. J.; Nowell, L.; Sandstrom, M.; Moran, P., Complex mixtures of Pesticides in Midwest U.S. streams indicated by POCIS time-integrating samplers. *Environmental Pollution* **2017,** *220*, 431-440.
20. O'Brien, D.; Lewis, S.; Davis, A.; Gallen, C.; Smith, R.; Turner, R.; Warne, M.; Turner, S.; Caswell, S.; Mueller, J. F.; Brodie, J., Spatial and Temporal Variability in Pesticide Exposure Downstream of a Heavily Irrigated Cropping Area: Application of Different Monitoring Techniques. *Journal of Agricultural and Food Chemistry* **2016,** *64* (20), 3975-3989.
21. Mazzella, N.; Dubernet, J.-F.; Delmas, F., Determination of kinetic and equilibrium regimes in the operation of polar organic chemical integrative samplers: Application to the passive sampling of the polar herbicides in aquatic environments. *Journal of Chromatography A* **2007,** *1154* (1), 42-51.
22. AFNOR, NF T90-210, - Qualité de l'eau - Protocole d'évaluation initiale des performances d'une méthode dans un laboratoire. **2018**.
23. Newman, M. C.; Clements, W. H., Multivariate approaches. In Ecotoxicology: a comprehensive treatment. Press, B. R. C., Ed. 2007; pp 479-495.
24. Güler, C.; Thyne, G. D.; McCray, J. E.; Turner, K. A., Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. *Hydrogeology Journal* **2002,** *10* (4), 455-474.
25. Yeo, I.-K.; Johnson, R. A., A New Family of Power Transformations to Improve Normality or Symmetry. *Biometrika* **2000,** *87* (4), 954-959.
26. Guibal, R.; Lissalde, S.; Leblanc, J.; Cleries, K.; Charriau, A.; Poulier, G.; Mazzella, N.; Rebillard, J.-P.; Brizard, Y.; Guibaud, G., Two sampling strategies for an overview of pesticide contamination in an agriculture-extensive headwater stream. *Environmental Science and Pollution Research* **2018,** *25* (15), 14280-14293.
27. McCune, B.; Grace, J. B.; Urban, D. L., Analysis of Ecological Communities. *MjM Software Design* **2002**, 300.
28. Kaiser, H. F., The varimax criterion for analytic rotation in factor analysis. . *Psychometrika* **1958,** *23*, 187–200.
29. Lagnel, J.-M., Manuel de data visualisation (Dunod). **2017**.
30. Dray, S.; Dufour, A.-B., The ade4 Package: Implementing the Duality Diagram for Ecologists. *Journal of Statistical Software* **2007,** *22* (4), 1 - 20.
31. Siberchicot, A.; Julien-Laferrière, A.; Dufour, A.-B.; Thioulouse, J.; Dray, S., adegraphics: An S4 Lattice-Based Package for the Representation of Multivariate Data. *R Journal* **2017,** *9*, 198-212.
32. Maechler M; Rousseeuw P; Struyf A; Hubert M; K, H., cluster: Cluster Analysis Basics and Extensions. R package version 2.1.4 — For new features, see the 'Changelog' file (in the package source), https://CRAN.R-project.org/package=cluster. **2022**.
33. Wickham, H., Ggplot2: Elegant Graphics for Data Analysis. 2nd Edition, Springer, New York. https://doi.org/10.1007/978-0-387-98141-3. **2009**.
34. Wilkins, D., https://CRAN.R-project.org/package=treemapify. **2018**.

35. Capriel, P.; Haisch, A.; Khan, S. U., Distribution and nature of bound (nonextractable) residues of atrazine in a mineral soil nine years after the herbicide application. *Journal of Agricultural and Food Chemistry* **1985,** *33* (4), 567-569.
36. Jablonowski, N. D.; Köppchen, S.; Hofmann, D.; Schäffer, A.; Burauel, P., Persistence of 14C-labeled atrazine and its residues in a field lysimeter soil after 22years. *Environmental Pollution* **2009,** *157* (7), 2126-2131.
37. Loiseau, L.; Barriuso, E., Characterization of the Atrazine's Bound (Nonextractable) Residues Using Fractionation Techniques for Soil Organic Matter. *Environmental Science & Technology* **2002,** *36* (4), 683-689.
38. Louchart, X.; Voltz, M.; Andrieux, P.; Moussa, R., Herbicide Transport to Surface Waters at Field and Watershed Scales in a Mediterranean Vineyard Area. *Journal of Environmental Quality* **2001,** *30* (3), 982-991.
39. Mudhoo, A.; Garg, V. K., Sorption, Transport and Transformation of Atrazine in Soils, Minerals and Composts: A Review. *Pedosphere* **2011,** *21* (1), 11-25.
40. Carretta, L.; Cardinali, A.; Marotta, E.; Zanin, G.; Masin, R., Dissipation of terbuthylazine, metolachlor, and mesotrione in soils with contrasting texture. *Journal of Environmental Science and Health, Part B* **2018,** *53* (10), 661-668.
41. Assoumani, A.; Coquery, M.; Liger, L.; Mazzella, N.; Margoum, C., Field application of passive SBSE for the monitoring of pesticides in surface waters. *Environmental Science and Pollution Research* **2015,** *22* (6), 3997-4008.
42. Rabiet, M.; Margoum, C.; Gouy, V.; Carluer, N.; Coquery, M., Assessing pesticide concentrations and fluxes in the stream of a small vineyard catchment - Effect of sampling frequency. *Environmental Pollution* **2010,** *158*, 737–748.

Table captions

Table 1. Abbreviations, log K_{ow} (octanol-water partition coefficient), log K_{oc} (organic carbon-water partition coefficient), families, biological activities and regulatory status of the 37 studied pesticides; pesticides written in bold presented annual quantification frequencies higher than 10%.

Table 2. The 50 sampling stations with their codes, names, HCA groups obtained, area of the associated watershed and land use repartitions.

Figure captions

Figure 1. Map projection of the Adour-Garonne basin with the 50 sampling stations selected (black dots); land use according to Corine Land Cover 2012; yellow: agricultural areas (i.e. cereal and vegetable crops); orange: vineyards; green: forests and semi-natural areas; light green: pastures; red: artificial surfaces; blue lines; major river networks.

Figure 2. Hierarchical cluster diagram of the 50 stations classified based on the percentage of land use; the numbers at the bottom of this graphical representation correspond to the station's code (as listed in Table 2); y-axis corresponds to the Euclidian distance. The roman numbers in the tree's branches correspond to a posteriori established groups stations with dominant land use: I: agricultural areas, II: Combination of forests, agricultural areas and pastures; III: forest; IV: pasture; V: vineyard.

Figure 3. Two-dimensional plot of PCA computed from the concentration of pesticides measured at each sampling station and period. The first two axes (A1 and A2) explain 35.7 % of the variability; (a) explanatory variables (23 studied pesticides identified with their abbreviations, listed in Table 1); (b) individuals (samples) and projection of the HCA groups identified with their respective color, the ellipse is drawn around the barycenter of the group, darker areas correspond to 50 % of the group's data.

Figure 4. Two-dimensional plot of PCA computed from the concentration of pesticides measured at each sampling station and period. The two axes (A3 and A4) explain 20.2 % of the variability; (a) explanatory variables (23 studied pesticides identified with their abbreviations, listed in listed in Table 1); (b) individuals (samples) and projection of the HCA groups identified with their respective color, the ellipse is drawn around the barycenter of the group, darker areas correspond to 50 % of the group's data.

Figure 5. Station n°5216210 located on the Gave de Pau River at Rieulhes (France). (a) Map projection of the watershed associated with this station (black dot); land use according to Corine Land Cover 2012; yellow: agricultural areas (i.e. cereal and vegetable crops); green: forests and semi-natural areas; light green: pastures; red: artificial surfaces; blue lines; major river networks. (b) Pesticides fingerprints (waffles diagrams of measured concentration in $\mu g\ L^{-1}$) obtained at this station for each sampling period (on the X-axis) and each studied pesticide (on the Y-axis, pesticides abbreviation are listed in Table 1); pesticides are grouped by biological activity where green corresponds to herbicides, red to insecticides, blue to fungicides and purple to metabolites.

Figure 6. Station n°5080960 located on the Gupie River at Sainte Bazeille (France). (a) Map projection of the watershed associated with this station (black dot); land use according to Corine Land Cover 2012; yellow: agricultural areas (i.e. cereal and vegetable crops); orange: vineyards; green: forests and semi-natural areas; light green: pastures; red: artificial surfaces; blue lines; major river networks. (b) Pesticides fingerprints (waffles diagrams of measured concentration in $\mu g\ L^{-1}$*) obtained at this station for each sampling period (on the X-axis) and each studied pesticide (on the Y-axis, pesticides abbreviation are listed in Table 1); pesticides are grouped by biological activity where green corresponds to herbicides, red to insecticides, blue to fungicides and purple to metabolites.*

Figure 7. Station n°5156950 located on the Hers Mort River at Saint Sauveur (France). (a) Map projection of the watershed associated with this station (black dot); land use according to Corine Land Cover 2012; yellow: agricultural areas (i.e. cereal and vegetable crops); green: forests and semi-natural areas; light green: pastures; red: artificial surfaces; blue lines; major river networks. (b) Pesticides fingerprints (waffles diagrams of measured concentration in $\mu g\ L^{-1}$*) obtained at this station for each sampling period (on the X-axis) and each studied pesticide (on the Y-axis, pesticides abbreviation are listed in Table 1); pesticides are grouped by biological activity where green corresponds to herbicides, red to insecticides, blue to fungicides and purple to metabolites.*

Figure 8. Station n°5075900 located on the Euille River at Laroque (France). (a) Map projection of the watershed associated with this station (black dot); land use according to Corine Land Cover 2012; orange: vineyards; yellow: agricultural areas (i.e. cereal and vegetable crops); green: forests and semi-natural areas; light green: pastures; red: artificial surfaces; blue lines; major river networks. (b) Pesticides fingerprints (waffles diagrams of measured concentration in $\mu g\ L^{-1}$*) obtained at this station for each sampling period (on the X-axis) and each studied pesticide (on the Y-axis, pesticides abbreviation are listed in Table 1); pesticides are grouped by biological activity where green corresponds to herbicides, red to insecticides, blue to fungicides and purple to metabolites.*

Supporting information captions

Table S 1. The 37 studied compounds with their uptake rate values (k_u), sampling rate values (R_s), instrumental quantification limit IQL and limits of quantification for POCIS (QL_p) associated with the two analytical techniques.

Table S 2. Varimax rotated axes loading matrices from PCA of pesticide concentrations (n=284); Bold value indicate strong loadings (≥ 0.75); Italic value indicate moderate loadings (≥ 0.5 and < 0.75).

Figure S 1. Data processing applied on the pesticide concentrations obtained after the field deployment of POCIS; P: number of studied pesticides; red box: incoming dataset; orange box: outgoing dataset.

Table 1.

Compound	Abbreviation	Log K_{ow}[a]	Log K_{oc}[b]	Family	Biological activity	Authorized/banned as pesticide (in France, 12/2016) [c,d]
Acetochlor	**ATC**	4.14	2.31	Chloroacetamide	Herbicide	Banned (2013)
Alachlor	**ALA**	2.97	3.09	Chloroacetamide	Herbicide	Banned (2008)
Atrazine	**ATZ**	2.50	2	Triazine	Herbicide	Banned (2003)
Azoxystrobin	AZS	2.50	2.63	Strobilurine	Fungicide	Authorized
Carbaryl	CBY	2.36	2.32	Carbamate	Insecticide	Banned (2008)
Carbendazim	**CBZ**	1.50	2.6	Carbamate	Fungicide	Banned (2008)
Carbofuran	**CBF**	1.62	1.34	Carbamate	Insecticide	Banned (2008)
Chlortoluron	**CTU**	2.50	2.15	Urea	Herbicide	Authorized
Cyproconazole	**CYP**	3.09	2.59	Triazole	Fungicide	Authorized
Desethylatrazine	**DEA**	1.51	1.26	Triazine	Metabolite	-
Desethylterbuthylazine	**DET**	2.23	2.17	Triazine	Metabolite	-
Desisopropylatrazine	**DIA**	1.15	1.84	Triazine	Metabolite	-
1-(3,4-dichlorophenyl)-3-methyl urea	DCPMU	2.94	2.06	Urea	Metabolite	-
1-(3,4-Dichlorophenyl) urea	DCPU	2.65	1.99	Urea	Metabolite	-
Dimetachlore	DMC	2.17	1.8	Chloroacetamide	Herbicide	Authorized
Dimetomorph	**DMM**	2.68	2.54	Morpholine	Fungicide	Authorized
Diuron	**DIU**	2.80	3.03	Urea	Herbicide	Banned for agriculture use (2008)/authorized as biocide
Epoxyconazol	**EPX**	3.30	3.26	Triazole	Fungicide	Authorized
Flurtamone	**FTM**	3.20	2.52	Furanone	Herbicide	Authorized

Flusilazol	FSZ	3.75	3.22	Triazole	Fungicide	Banned (2008)
Hexazinone	HXZ	1.17	1.73	Triazinone	Herbicide	Banned (2007)
Imidacloprid	**IMI**	0.57	2.28	Neonicotinoid	Insecticide	Authorized (Banned 2018)
1-(4-isopropylphenyl)-3-methyl urea	IPPMU	2.63	2.17	Urea	Metabolite	-
1-(4-isopropylphenyl) urea	IPPU	2.16	2.25	Urea	Metabolite	-
Isoproturon	**IPU**	2.50	2.14	Urea	Herbicide	Authorized (Banned 2017)
Linuron	LINU	3.00	2.79	Urea	Herbicide	Authorized
Metazachlor	**MTZ**	2.49	1.88	Chloroacetamide	Herbicide	Authorized
Methomyl	MTY	1.24	1.4	Carbamate	Insecticide	Banned (2009)
Metolachlor	**MTC**	3.40	2.3	Chloroacetamide	Herbicide	Authorized
Metoxuron	MTX	1.60	2.08	Urea	Herbicide	Banned (2007)
Norflurazon	**NFZ**	2.30	2.85	Pyridazinone	Herbicide	Banned (2004)
Norflurazon-desmethyl	NFZD	1.72	3.43	Pyridazinone	Metabolite	-
Pirimicarb	PYC	1.70	2.59	Carbamate	Insecticide	Authorized
Simazine	**SMZ**	2.18	2.11	Triazine	Herbicide	Banned (2003)
Tebuconazole	**TBZ**	3.70	3.19	Triazole	Fungicide	Authorized
Terbuthylazine	**TUZ**	3.40	2.34	Triazole	Herbicide	Banned (2003)
Thiodicarbe	TIC	1.62	2.62	Carbamate	Insecticide	Banned (2008)

[a] Data from Ineris (https://substances.ineris.fr/fr/)

[b] Information from the Pesticide Properties DataBase – PPDB (http://sitem.herts.ac.uk/aeru/ppdb/en/atoz.htm)

[c] Information from the international office of water (OIEau - https://dev.oieau.fr/ag-pesticides/substances)

[d] Information from E-Phy-Anses (https://ephy.anses.fr/)

Table 2.

Station	Station name	HCA groups	Area of the watershed (km²)	Artificial surfaces	Agricultural areas	Vineyards	Pastures	Forests and semi-natural areas	Comments
5008000	The Seugne at St-Germain de Lusignan	I	369	2%	76%	8%	2%	13%	Agriculture
5011600	The Beau at Saint-Médard	I	219	3%	74%	5%	1%	17%	
5012000	The Antenne at Javrezac	I	338	3%	65%	21%	2%	9%	
5013150	The Tourtrat at Reparsac	I	28	3%	64%	24%	0%	8%	
5023100	The Lizonne downstream Bioussac	I	132	0%	67%	0%	7%	26%	
5030000	The Dronne in Coutras	I	5857	2%	57%	0%	6%	34%	
5079100	The Dropt in Loubens	I	2345	1%	74%	5%	6%	13%	
5079200	The Andouille at Roquebrune	I	67	0%	69%	12%	2%	17%	
5080960	The Gupie in Sainte Bazeille	I	219	1%	81%	2%	2%	15%	
5083300	The Trec at Longueville	I	334	2%	87%	0%	2%	7%	
5104000	The Garonne upstream Lot	I	71646	3%	52%	1%	9%	33%	
5106850	The Gélise upstream Rimbez	I	338	2%	56%	17%	9%	15%	
5107000	The Grande Baise at Bapaume	I	2523	2%	81%	0%	6%	10%	
5115550	The Gèze at Castelnau Magnoac	I	28	1%	71%	0%	6%	16%	
5129150	The Rieu Tort at Labastide St Pierre	I	132	5%	49%	35%	1%	8%	
5155000	The Save at Grenade	I	2188	2%	81%	0%	8%	10%	
5156950	The Hers Mort at Saint-Sauveur	I	1824	12%	82%	0%	2%	3%	
5157100	The Sausse at Toulouse	I	224	12%	84%	0%	0%	4%	

5158700	The Aussonnelle at Seilh	I	357	21%	60%	0%	1%	18%	
5219000	The Luy at Saint-Pandélon	I	2166	3%	70%	0%	8%	19%	
5228280	The Douze at Mauvezin d'Armagnac	I	604	0%	63%	10%	9%	17%	
5229100	The Midour at Lannemaignan	I	924	1%	65%	7%	11%	16%	
5231370	The Adour at Borderes	I	5857	4%	55%	0%	10%	31%	
5015320	The Eaux-Claires at Puymerle	II	473	0%	33%	0%	13%	**54%**	Mix Forests + agriculture + pastures
5042000	The Auvézère at Pont Rognac	II	4885	1%	**41%**	0%	22%	35%	
5049000	The Vézère at Le Bugue	II	1824	3%	26%	0%	27%	**43%**	
5054600	The Solane downstream Naves	II	224	6%	**33%**	0%	31%	31%	
5064000	The Cère at Sansac	II	357	6%	22%	0%	33%	**39%**	
5068640	The Sumène upstream Valette	II	3831	0%	9%	0%	27%	**62%**	
5068890	The Rhue at St-Thomas	II	5	1%	6%	0%	40%	**52%**	
5093550	The Riou Mort downstream Viviez	II	369	7%	28%	0%	32%	**33%**	
5099170	The Boralde Flaujaguèse downstream Espalion	II	219	0%	3%	0%	45%	**52%**	
5128000	The Aveyron at Lugans	II	473	2%	29%	0%	34%	**35%**	
5134000	The Agout at Ambrès	II	4885	3%	36%	0%	17%	**44%**	
5167008	The Grand Hers upstream Vixiège	II	1850	2%	27%	0%	12%	**58%**	
5211550	The Luzoué at Monein	II	5	0%	22%	0%	20%	**58%**	
5225100	The Midouze at Tartas	II	5711	2%	34%	2%	3%	**58%**	
5226000	The Midouze at Campagne	II	4770	2%	36%	3%	3%	**55%**	
5065500	The Jordanne upstream Mandailles-St-Julien	III	1850	0%	1%	0%	19%	80%	Forests
5181000	The Garonne at Labarthe Inard	III	3831	3%	11%	0%	8%	78%	
5191000	The Leyre at Lamothe	III	3583	2%	14%	0%	0%	84%	
5191900	The Leyre at Belin Beliet	III	2741	1%	17%	0%	0%	82%	

5216210	The Gave de Pau at Rieulhes	III	2076	2%	3%	0%	7%	**87%**	
5224100	The Retjons at Tartas	III	257	3%	18%	0%	0%	**79%**	
5071300	The Mortagne upstream Tauves	IV	2076	1%	6%	0%	**84%**	10%	
5097000	The Lander downstream St-Flour	IV	604	3%	26%	0%	**60%**	10%	Pastures
5042085	The Arnac stream at Arnac Pompadour	IV	2188	8%	14%	0%	**65%**	13%	
5028110	The Barbanne at Libourne	V	924	1%	5%	**82%**	10%	3%	
5075900	The Euille at Laroque	V	197	2%	10%	**62%**	4%	22%	Vineyards
5078900	The Vignague at Morizès	V	195	1%	15%	**58%**	9%	16%	

Figures

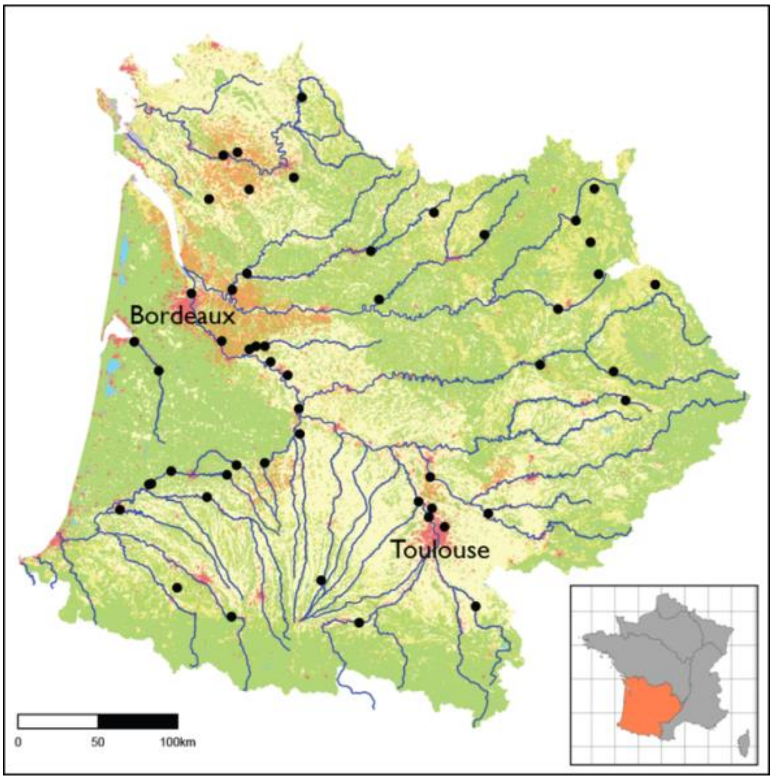

Figure 1. Map projection of the Adour-Garonne basin with the 50 sampling stations selected (black dots); land use according to Corine Land Cover 2012; yellow: agricultural areas (i.e. cereal and vegetable crops); orange: vineyards; green: forests and semi-natural areas; light green: pastures; red: artificial surfaces; blue lines; major river networks.

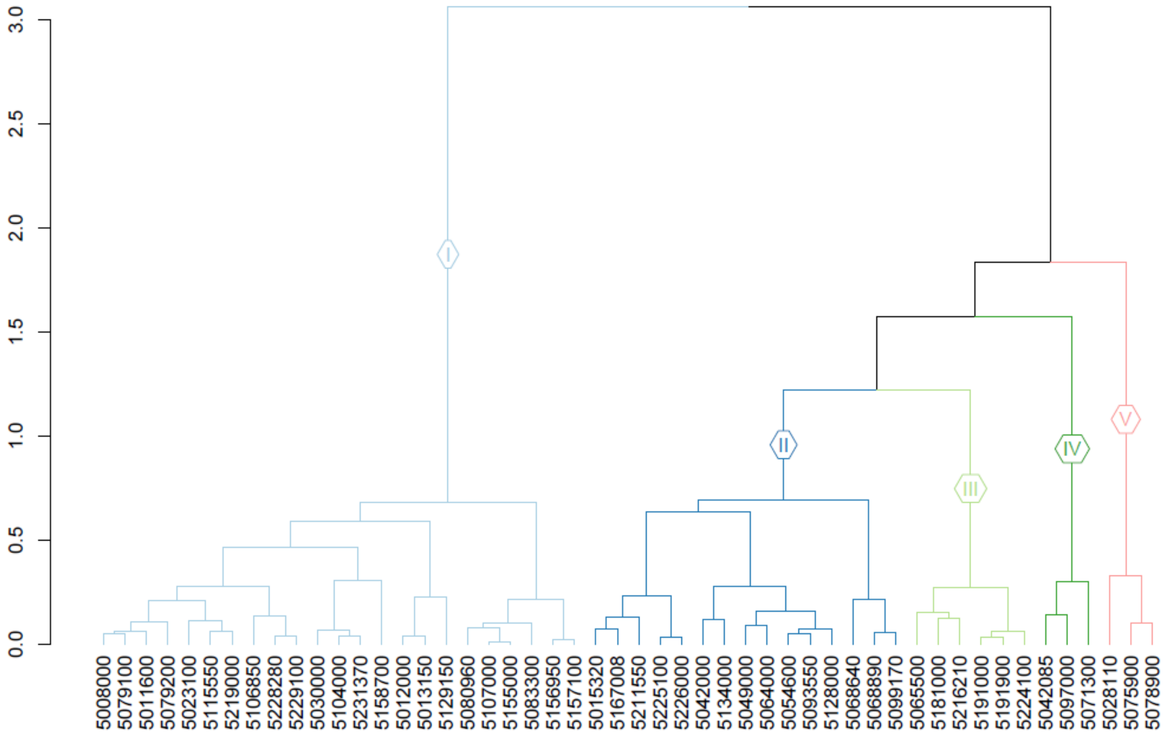

Figure 2. Hierarchical cluster diagram of the 50 stations classified based on the percentage of land use; the numbers at the bottom of this graphical representation correspond to the station's code (as listed in ***Erreur ! Source du renvoi introuvable.****); y-axis corresponds to the Euclidian distance. The roman numbers in the tree's branches correspond to a posteriori established groups stations with dominant land use: I: agricultural areas, II: Combination of forests, agricultural areas and pastures; III: forest; IV: pasture; V: vineyard.*

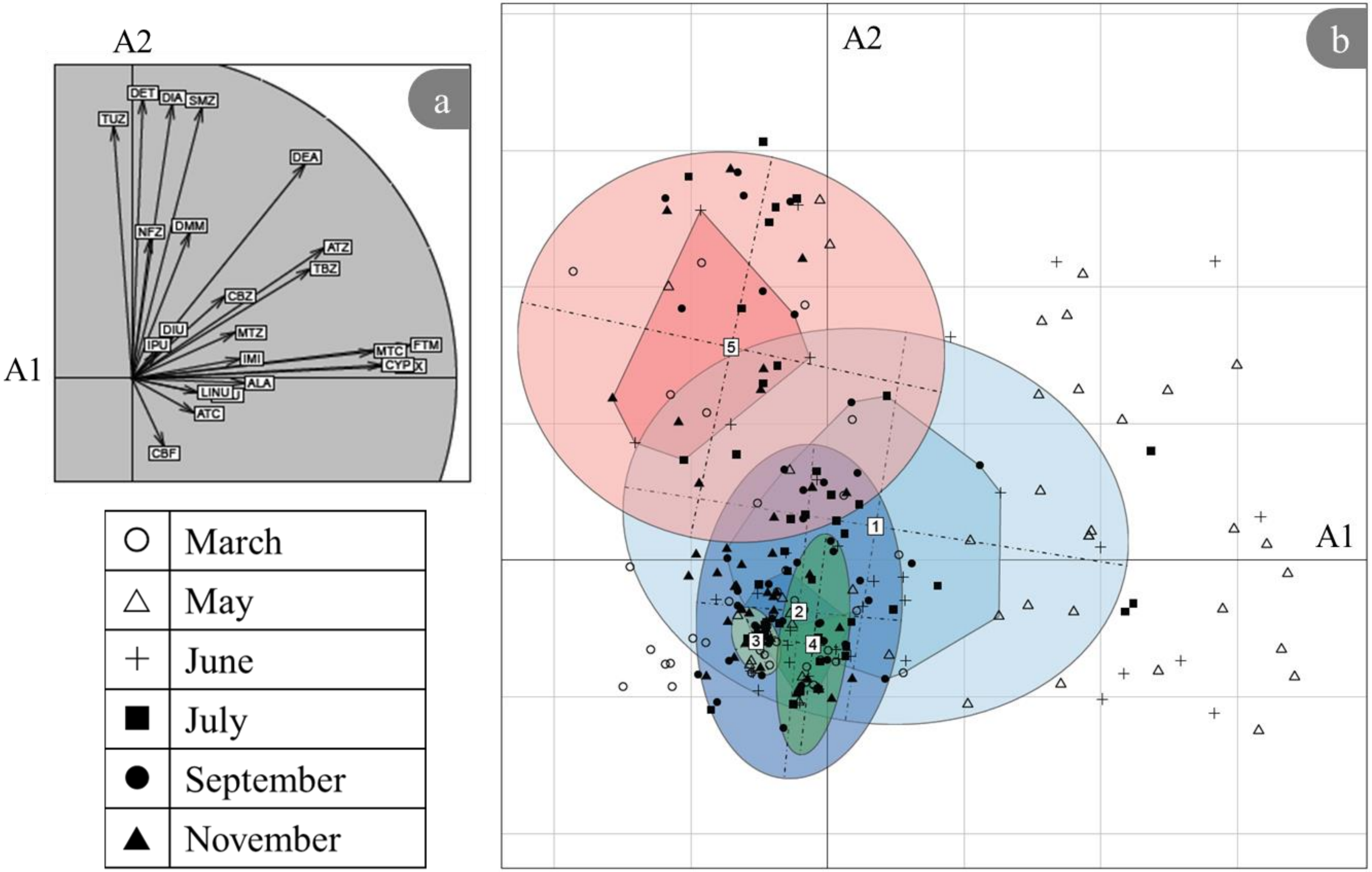

Figure 3. Two-dimensional plot of PCA computed from the concentration of pesticides measured at each sampling station and period. The first two axes (A1 and A2) explain 35.7 % of the variability; (a) explanatory variables (23 studied pesticides identified with their abbreviations, listed in ***Erreur ! Source du renvoi introuvable.****); (b) individuals*

(samples) and projection of the HCA groups identified with their respective color, the ellipse is drawn around the barycenter of the group, darker areas correspond to 50 % of the group's data.

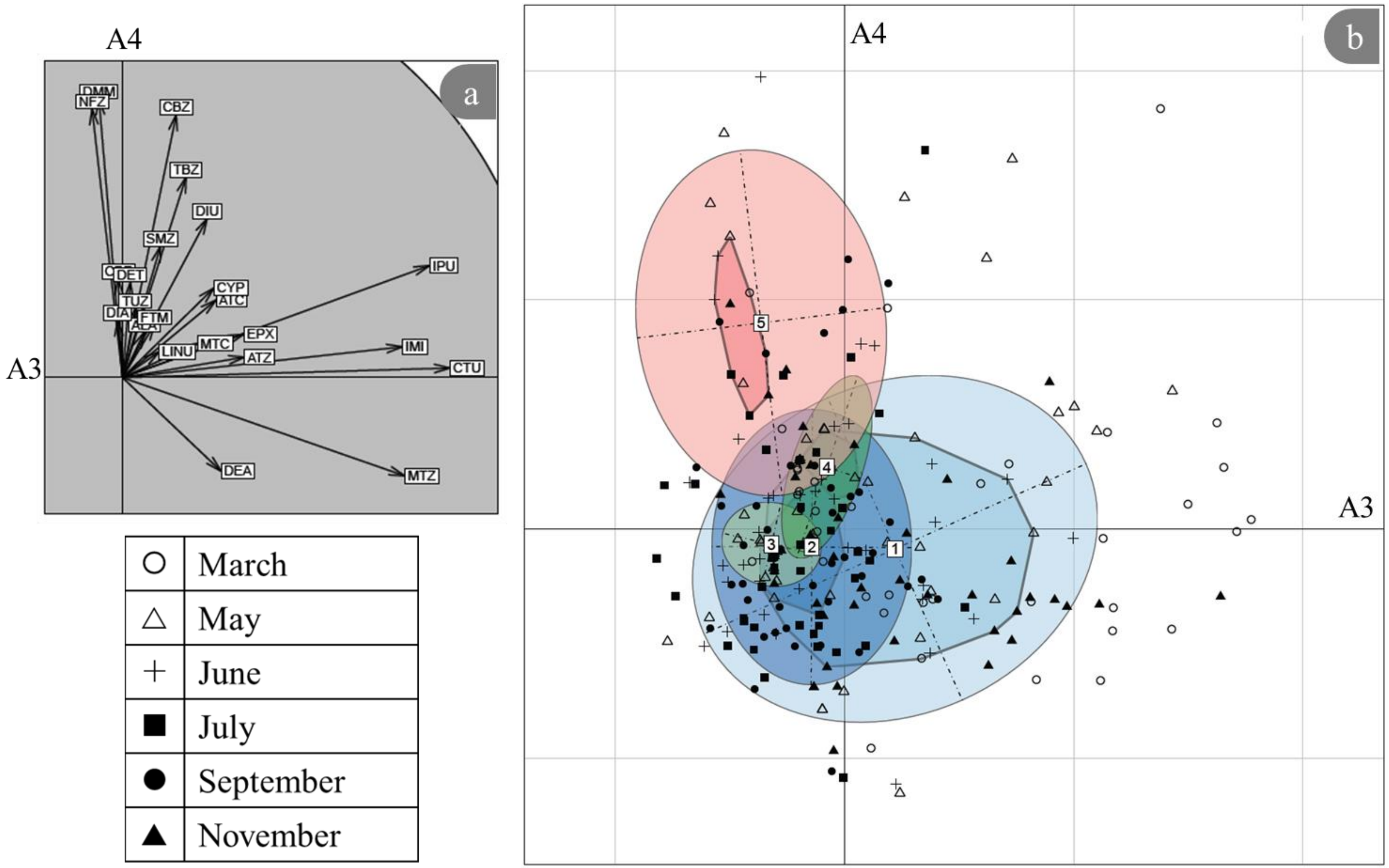

Figure 4. Two-dimensional plot of PCA computed from the concentration of pesticides measured at each sampling station and period. The two axes (A3 and A4) explain 20.2 % of the variability; (a) explanatory variables (23 studied pesticides identified with their abbreviations, listed in ***Erreur ! Source du renvoi introuvable.****); (b) individuals*

(samples) and projection of the HCA groups identified with their respective color, the ellipse is drawn around the barycenter of the group, darker areas correspond to 50 % of the group's data.

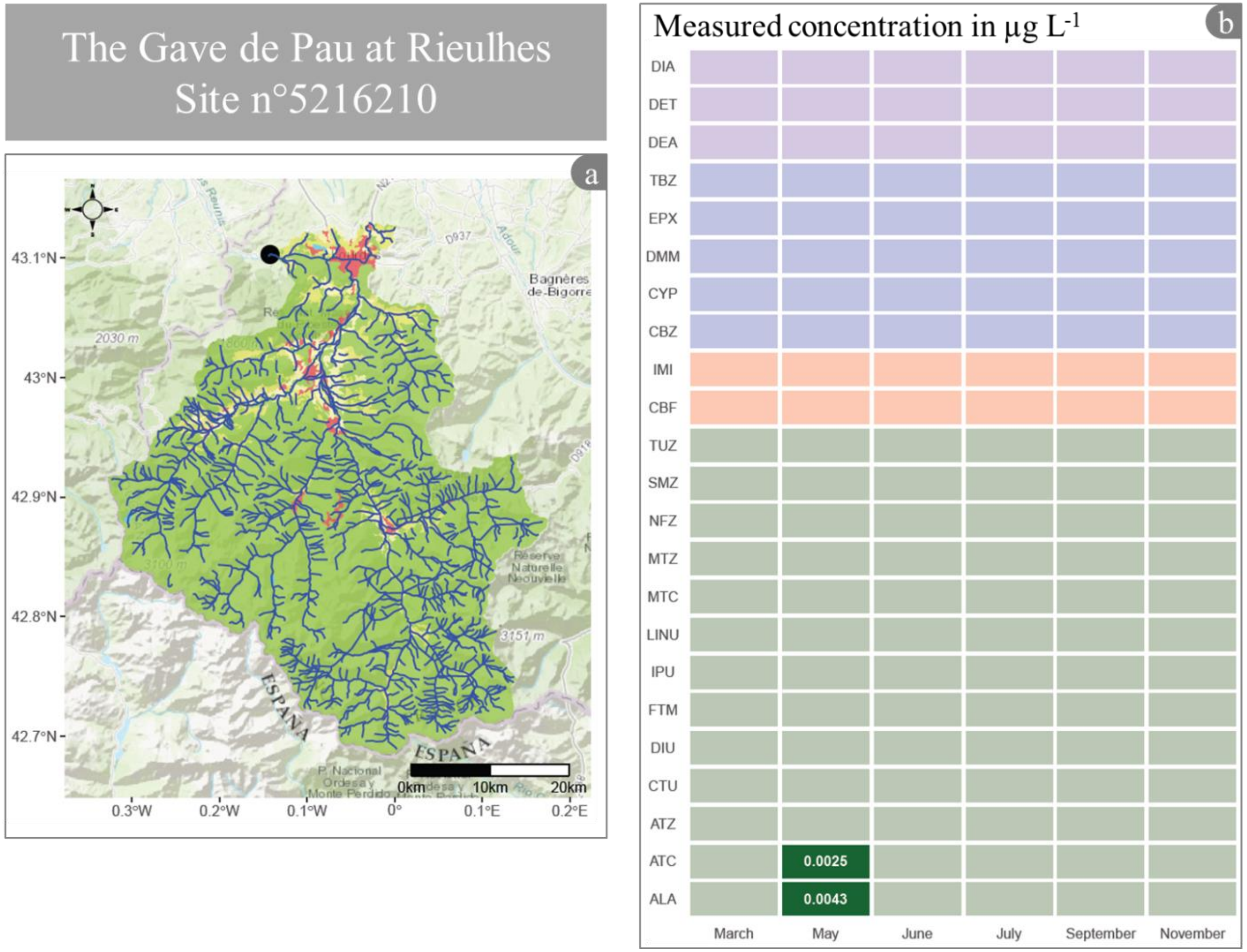

Figure 5. Station n°5216210 located on the Gave de Pau River at Rieulhes (France). (a) Map projection of the watershed associated with this station (black dot); land use according to Corine Land Cover 2012; yellow: agricultural areas (i.e. cereal and vegetable crops); green: forests and semi-natural areas; light green: pastures; red: artificial surfaces; blue lines; major river networks. (b) Pesticides fingerprints (waffles diagrams of measured concentration in µg L^{-1}) obtained at this station for each sampling period (on the X-axis) and each studied pesticide (on the Y-axis, pesticides abbreviation are listed in ***Erreur ! Source du renvoi introuvable.****); pesticides are grouped by biological activity where green corresponds to herbicides, red to insecticides, blue to fungicides and purple to metabolites.*

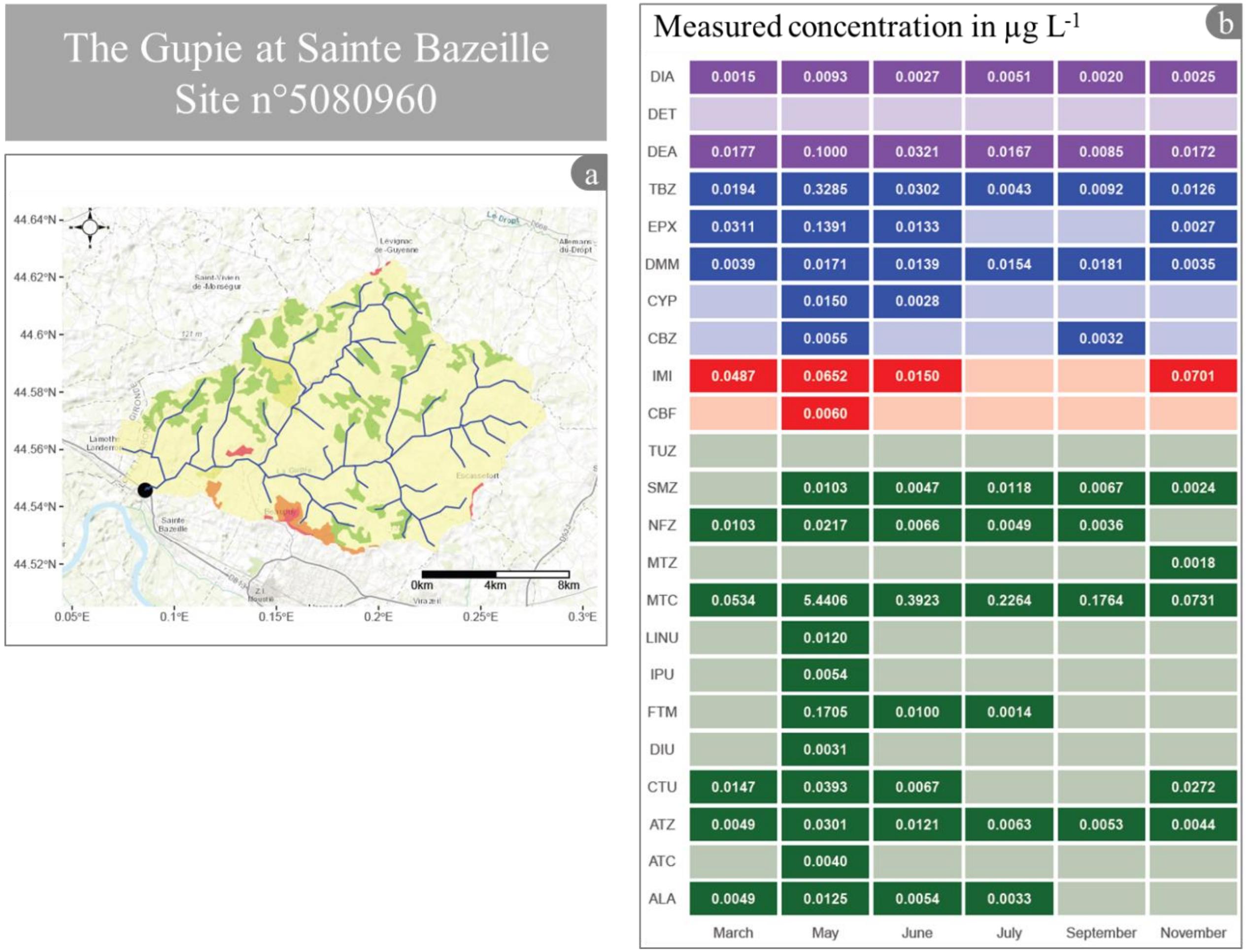

Figure 6. Station n°5080960 located on the Gupie River at Sainte Bazeille (France). (a) Map projection of the watershed associated with this station (black dot); land use according to Corine Land Cover 2012; yellow: agricultural areas (i.e. cereal and vegetable crops); orange: vineyards; green: forests and semi-natural areas; light green: pastures; red: artificial surfaces; blue lines; major river networks. (b) Pesticides fingerprints (waffles diagrams of measured concentration in µg L^{-1}*) obtained at this station for each sampling period (on the X-axis) and each studied pesticide (on the Y-axis, pesticides abbreviation are listed in* ***Erreur ! Source du renvoi introuvable.****); pesticides are grouped by biological activity where green corresponds to herbicides, red to insecticides, blue to fungicides and purple to metabolites.*

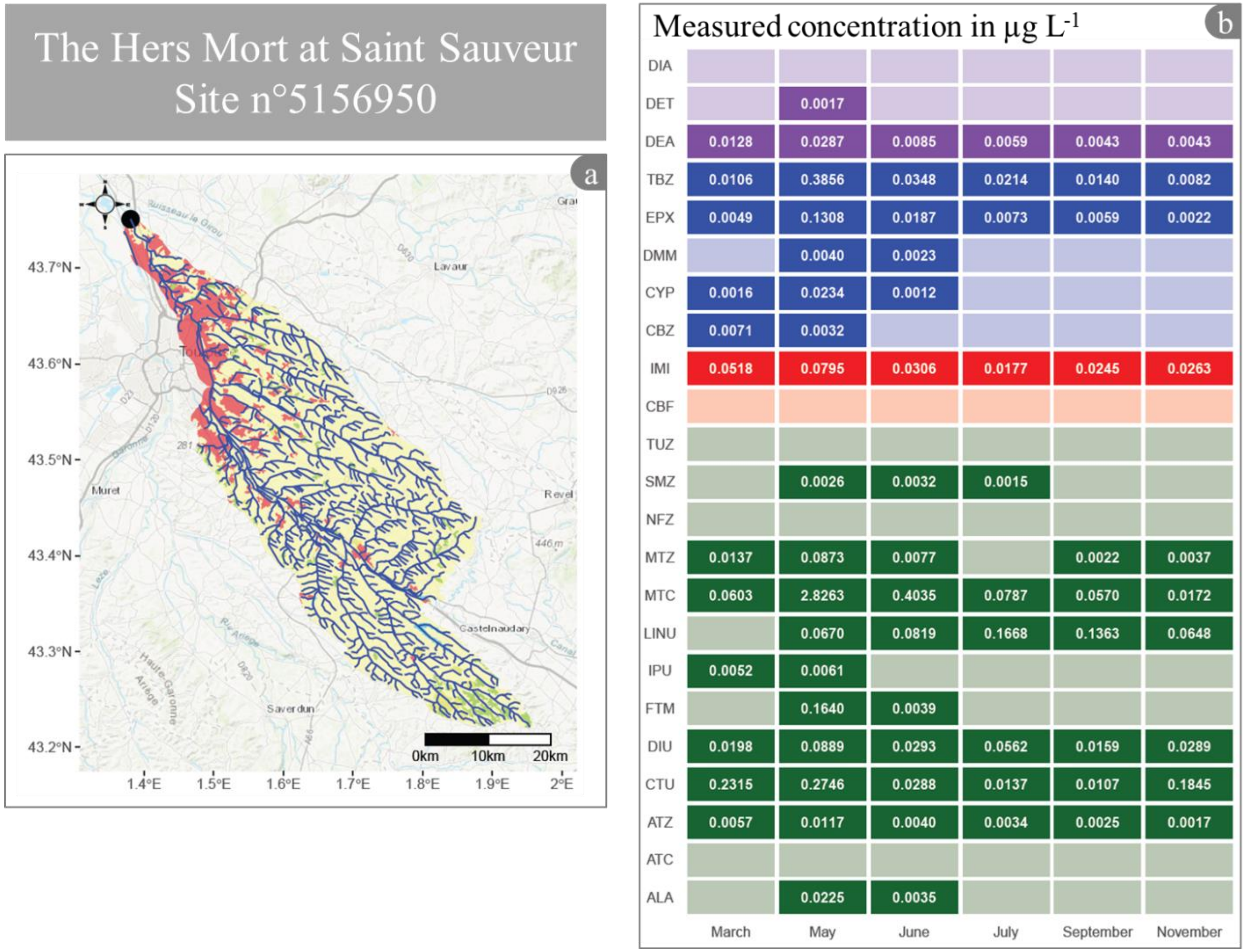

Figure 7. Station n°5156950 located on the Hers Mort River at Saint Sauveur (France). (a) Map projection of the watershed associated with this station (black dot); land use according to Corine Land Cover 2012; yellow: agricultural areas (i.e. cereal and vegetable crops); green: forests and semi-natural areas; light green: pastures; red: artificial surfaces; blue lines; major river networks. (b) Pesticides fingerprints (waffles diagrams of measured concentration in µg L^{-1}) obtained at this station for each sampling period (on the X-axis) and each studied pesticide (on the Y-axis, pesticides abbreviation are listed in ***Erreur ! Source du renvoi introuvable.****); pesticides are grouped by biological activity where green corresponds to herbicides, red to insecticides, blue to fungicides and purple to metabolites.*

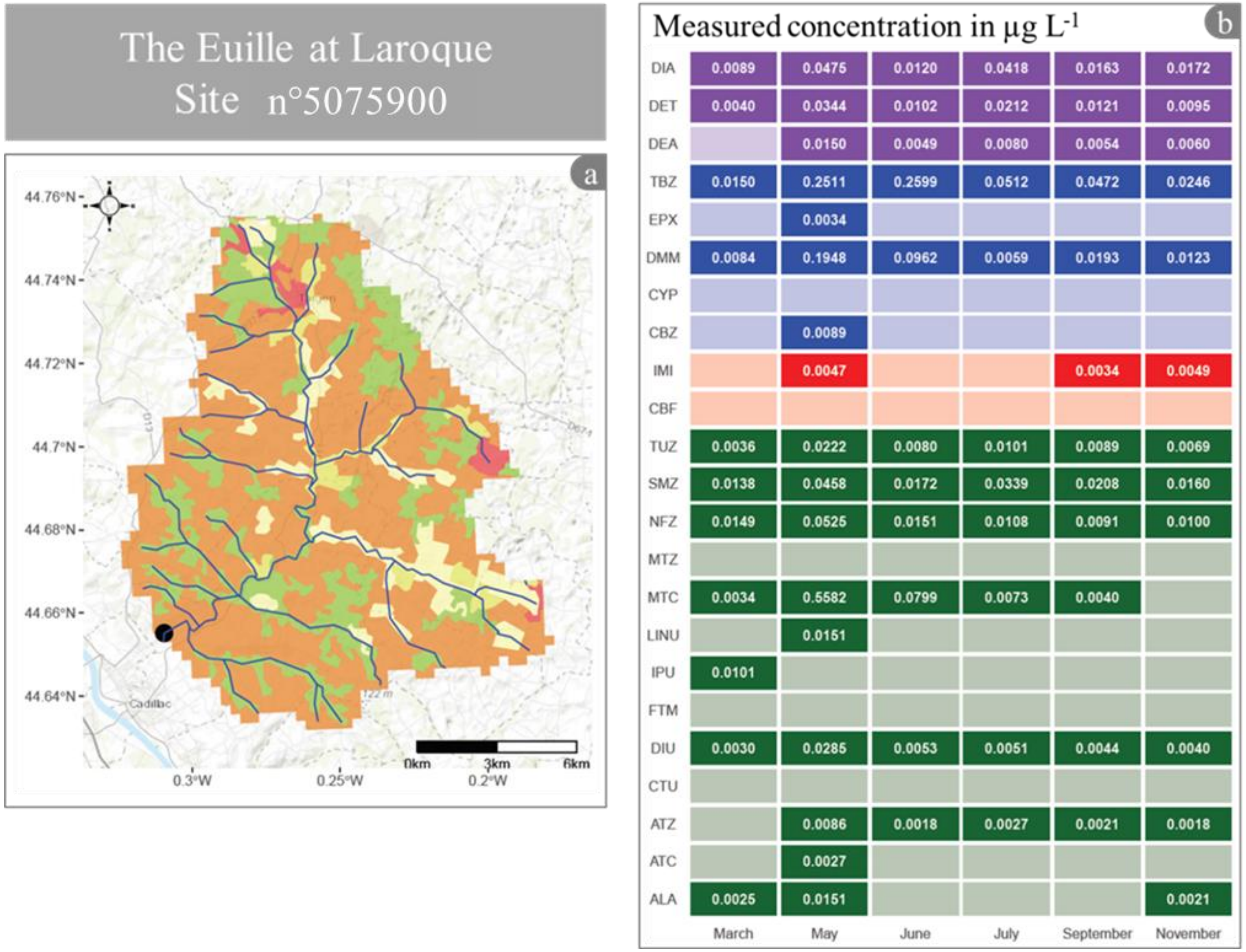

Figure 8. Station n°5075900 located on the Euille River at Laroque (France). (a) Map projection of the watershed associated with this station (black dot); land use according to Corine Land Cover 2012; orange: vineyards; yellow: agricultural areas (i.e. cereal and vegetable crops); green: forests and semi-natural areas; light green: pastures; red: artificial surfaces; blue lines; major river networks. (b) Pesticides fingerprints (waffles diagrams of measured concentration in µg L^{-1}) obtained at this station for each sampling period (on the X-axis) and each studied pesticide (on the Y-axis, pesticides abbreviation are listed in ***Erreur ! Source du renvoi introuvable.****); pesticides are grouped by biological activity where green corresponds to herbicides, red to insecticides, blue to fungicides and purple to metabolites.*