Manuscript PCI Ecotox Env Chem

Dear Pierre Labadie,

Thank you for the extremely helpful feedback you and the reviewers provided on our manuscript. We really appreciate the time you have taken through this and for pointing further pertinent references from the literature.

We have now implemented the suggested changes in a revised version of the manuscript. Among the most important changes suggested by the reviewers: we have run the analyses with the appropriate year for the spatial data set (2017) and incorporated minor changes in the results following this update. We also addressed the general comment on wording, by replacing the use of “cocktails” with “mixtures” where appropriate. We tested the stability of clustering across years and explained our choice of comparing 2017 to all years (including 2017) instead of presenting the year by year analysis. Finally, we addressed all minors’ comments by both reviewers. Note, that all responses below use the new results presented in the manuscript.

We detail below our point-by-point responses to the reviewers’ comments. The corresponding revisions are highlighted in the revised manuscript by using the track changes mode; other changes resulting from updated analyses are also highlighted.

Identifying pesticide cocktails at country-wide scale

Review by Clémentine FRITSCH

General comments

This study presents an original and well-designed method to characterize the mixtures of currently used pesticides that wildlife may be exposed to in agrosystems. This approach was developed at country scale, based on available public data about sales of pesticides and cropping, and seems reliable to identify specific groups of pesticide mixtures, identify the spatial distribution of such groups, and identify both pesticides of broad use that are applied nationally and on different crops and “discriminating” compounds which are of specific use geographically and/or agronomically.

Within the current context of dramatic biodiversity loss in agricultural landscapes, providing knowledge to help understanding and predicting wildlife exposure to and unintentional impacts of pesticides on ecosystems is of crucial importance, making this manuscript of high scientific interest. The methods are scientifically sound, and the MS is well written.

As a consequence, I have a few minor remarks listed hereafter, but also I point out several drawbacks. The first one is about the recurrent use of the word “cocktail” even in cases where the related concepts and paradigms would justify using “mixture” instead.

The second concern is about a validation of the approach, a way to corroborate the findings and interpretation, which is partly lacking. This may be achieved by considering further comparison with one of the references already cited, and by adding data and results from environmental surveys in France which report the mixtures actually found in environmental matrices (soil, air, water) and biota.

The test of the temporal robustness should be refined by comparing the output from the “core” year 2017 to the output obtained from the dataset on the other years without including 2017. These issues are detailed in specific comments.

Thank you for the positive comments and for pointing out the issue regarding the use of “cocktail” and “mixture”. We checked their use throughout the manuscript and replaced “cocktail” with “mixture” where appropriate. See below for a more detailed response.

We agree that a comparison of the purchase data to environmental contamination data would strengthen our results. To test whether data on pesticide purchase can be representative of environmental contamination we used a dataset on water contamination in French rivers and water bodies (Naïades; http://www.naiades.eaufrance.fr/acces-donnee). We generally found positive, albeit not extremely strong, correlations between purchase and contamination data, suggesting that the pesticide purchase data we used are indeed related to environmental contamination. A full description of the analysis and correlation results is detailed below.

Regarding the test of the temporal robustness of our results, it seemed to us more appropriate to compare each available year to 2017, than to compare the period 2014-2018 excluding 2017 to the year 2017 as the meaning of a time period with a missing year is unclear. Results of this new robustness analysis are presented below.

I thought in some cases within the manuscript that English language may be seen as of unequal quality. But I am not a native English speaker, I can neither perform a fully relevant review of English writing nor assert whether a review of English language is really needed or not.

We have proofread the manuscript carefully.

Detailed comments

Title and throughout the manuscript – usually the word “cocktail” is used when dealing with (toxic) responses of organisms, since it refers to concepts and paradigms related to “cocktail effects” with underlying mechanisms and modelling of antagonism, additivity and synergy between compounds in inducing effects in organisms. Within the context of environmental chemistry and issues related to exposure, which is the case here in this study, the word “mixture” of pesticides or exposure/co-exposure to multiple compounds is used. This is the case in scientific literature as well as in regulatory context (see for instance (Beronius et al., 2020)). I recommend in the title and when relevant in the manuscript to use “mixture” instead of “cocktail”, and keep the latest only when referring to toxicological issues.

Beronius, A., Zilliacus, J., Hanberg, A., Luijten, M., van der Voet, H., van Klaveren, J., 2020. Methodology for health risk assessment of combined exposures to multiple chemicals. Food Chem. Toxicol. 143, 111520. https://doi.org/10.1016/j.fct.2020.111520

We initially avoided the word “mixtures” to avoid confusion with the statistical tool “mixture model”. However, the reviewer is right and we have modified the text throughout the manuscript to replace “cocktail” with “mixture” where appropriate. We have hence renamed “mixture models” to “model-based clustering” to avoid the confusion (p7, l.179).

Line 88. The reference of the database and/or URL must be added.
Indeed the reference was missing, we added it (P4; l.94)

Line 93. "[…] a well-suited case country to try and identify pesticide […]" something weird in this sentence. "[…] a well-suited case country to try *identifying* pesticide […]"?
We changed P4 line 99 to "[…] a well-suited case country to identify pesticide […]"

Line 96-101. The end of the introduction looks like a conclusion or summary rather than a classical end of introduction where expectations and hypotheses are usually described. I suggest to modify.
The end of the introduction was changed to underline the specific objectives of the study. We specify three objectives and describe the hypotheses and expectations (P5 lines 101-110)

Line 106. The amount of what? Please provide details.
We reformulated P5 line 116 "[…] the amount of each active substance purchased […]"

Lines 103-112. To make it clear for the readers, it would be necessary to explain the administrative and spatial meaning of postcode in France. (what is "township" versus "district", the average size of township surfaces and their potential variability).
The range of postcode areas was added and we now explain the different divisions of France Region>Department> Postcode, with a reference to European NUTS (ecostat) levels (P. 5, lines 120).

Lines 124-125. The probability that compounds are not used over the same postcode than bought is high looking at the organization of farms nowadays. It may be worth mentioning that however they are likely to be used closely, in the vicinity of the administrative location of the farmer, so in surrounding postcodes (which allow studying the spatial patterns at the national scale).
This comment has been addressed by specifying both the approximation ("pesticides may not always be spread exactly in the postcode where farmers are domiciled ") and explaining the spatial proximity ("but are unlikely to be used beyond the neighbouring postcodes"); see P6 lines 138-141

Lines 139-140. Is this related or similar to the "Registre parcellaire graphique (RPG)"? On official websites it is indicated that the annual RPG versions before 2013 (2010, 2011 et 2012) are available on the webpages of the "Agence de service et de paiement" while the most recent RPG versions since are available at "data.gouv.fr". Since the data of pesticide sales in 2017 were studied, why not using the RPG data from 2017? The authors should explain why not using the RPG and what was the year of concern.
LPIS is exactly the "Registre parcellaire graphique (RPG): we used the generic name, related to the Common Agricultural Policy in the European Union, Land Parcel Identification System (LPIS) to name it in English, but we have added the French name to avoid ambiguities (P6, line 155).
Regarding the focal year, we started the analyses with the data we had at hand, but you are correct that it would make a lot more sense to match pesticide data with land use data. The manuscript is now fully updated with year 2017 for LPIS. This change of data has changed some of the results, but only slightly, as expected due to the relative stability of cropland area from one year to the other: the number of

postcode groups is now 19 instead of 18, and the number of core substances is 12 instead of 13. These minor changes are highlighted in the manuscript.

Lines 144-146. As far as I well understand, the total cropping area was summed up, thus merging both conventional and organic farming? Could this introduce a bias in further analyses and interpretation of the data since most of, if not all in the list studied, synthetic pesticides are not used in organic farming and the surfaces of organic farming might not be homogenously distributed in space at the township and/or the national scale. The plots cultivated under organic farming are available in the RPG.

We thank the reviewer for this comment. In fact, the proportion of organic farming does vary across postcodes and postcode groups (see Figure A), with mostly two postcode groups (*n* and *o*) having large fractions of organic farming. These postcodes cover chiefly mountain areas and have relatively low cropland areas (see Figure 2 in the manuscript). The proportion of organic farming varies much less across the other postcode groups. Therefore, organic farming should not affect our results much. For postcode groups *n* and *o*, the higher proportion of organic farming might lead to higher local heterogeneity, with conventional fields receiving more pesticides than estimated by the mixture model. However, it remains true that the higher proportion of organic fields in these postcode groups lead to a lower overall registered pesticide load. Organic farming also raises the issue of other substances without mandatory reporting that may be used in organically farmed areas. This is now explained P. 23, lines 676-679.

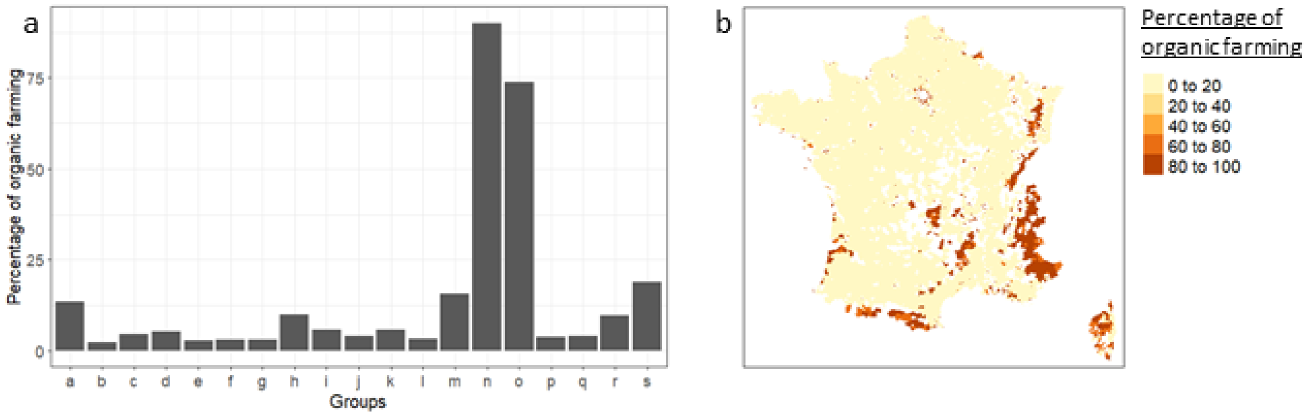


Figure A: (a) Histogram of the percentages of organic farming per group of postcodes. (b) Map of the percentage of organic farming per group of postcode.

Line 260. Please check the writing of “k”.

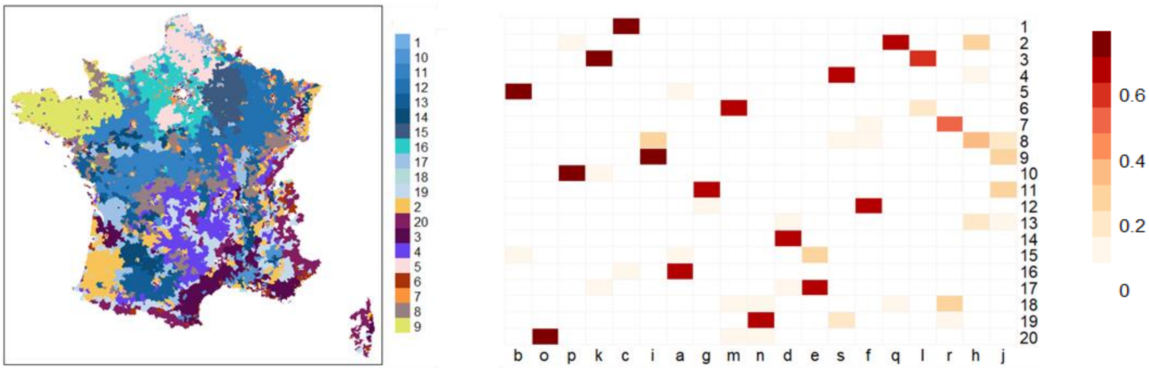
The writing of “k” was changed see lines 196,272,290,341,355

Line 301-312. It is a very good point that authors checked the reliability of the method over time. However, the fact that the data from the year 2017 are present in the two dataset is a flaw, since this of course artificially increases the probability of correlation, creates an absence of statistical independence between the measurements. It may be recommended to perform the same approach but comparing the output from 2017 to the output from the dataset ***2015-2018 without 2017***.

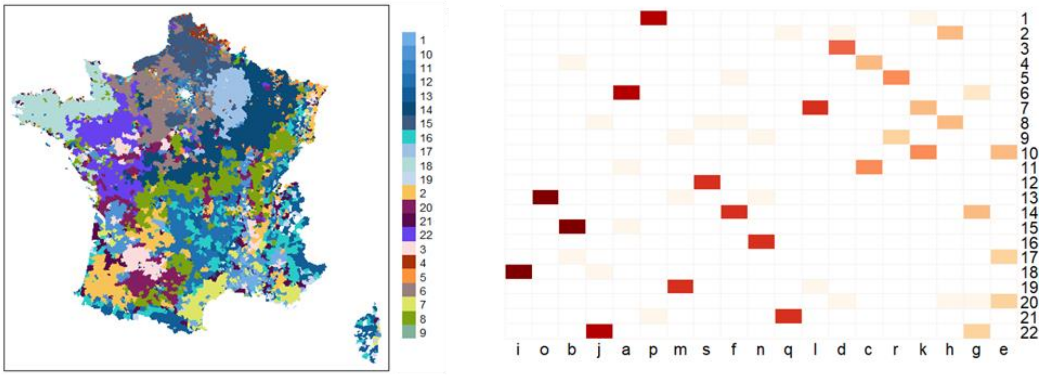
We agree that including year 2017 in the two analyses contributes to obtaining correlated results. However, the rationale behind the comparison of the 2017 outputs with the 2015-2018 outputs was to evaluate how sensitive the results were with respect to the choice of a specific year vs. a longer time period. We could compare 2017 vs. 2015, 2016 and 2018 pooled together, but it seems to us that the latter grouping with a missing year in the middle of a time series would be a bit weird. Instead, we performed the analysis per year, comparing, 2017 to 2015, 2016 and 2018 separately. The results, shown below, indicate that the clustering is relatively stable through time. First, the number of clusters varies little, between 19 in 2017 and 25 in 2018. Second, all the main clusters are found in all years, albeit sometimes with splitting (e.g. Brittany can be split into a western and a southern part).

We hence chose to leave the results of the 2015-2018 clustering in the manuscript, as a test of the sensitivity to the time range chosen for analysis, because an analysis of the temporal variation of the clustering is beyond the scope of the manuscript, but we could replace it with the year-to-year comparison if the reviewers and/or editors think it relevant.

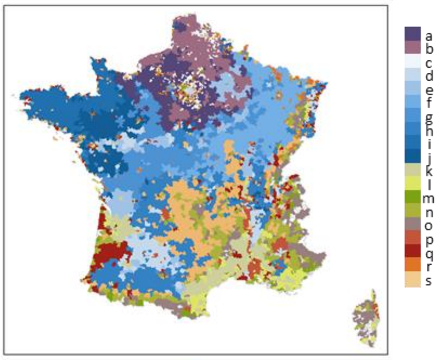
2015



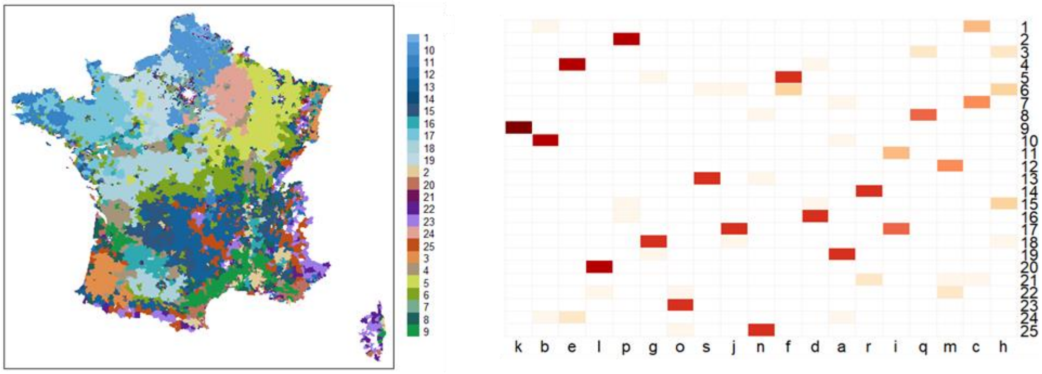
2016



2017



2018



All

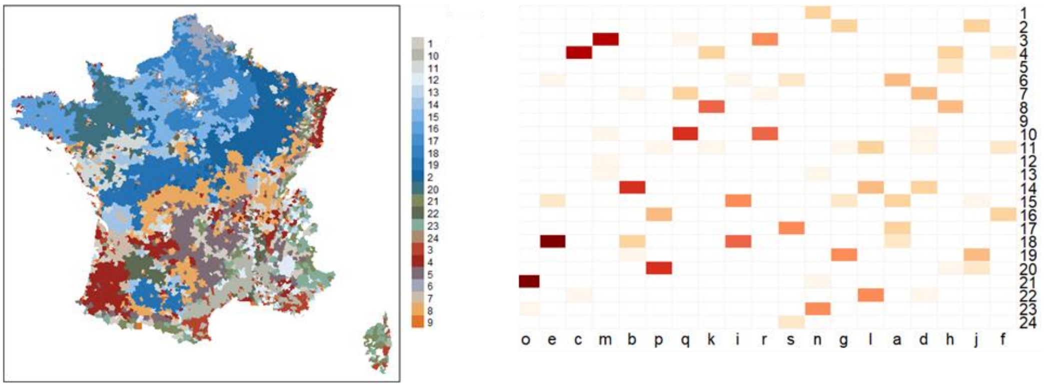


Figure B: Differences and similarities in the clustering of postcodes produced by model-based clustering run on data from 2017 against each other year and all years pooled. Maps of each year from 2015 to 2018 are shown. Next to each map, a heatmap corresponding to the similarity of the 2017 groups with the groups of the corresponding year, expressed as the percentage postcodes from 2017 groups that were found in the various groups of the corresponding year. "ALL" correspond to years 2015 to 2018 pooled together. The heatmaps should be read vertically: for example, when comparing results of the mixture model performed on 2017 data to the one performed on the 2015-2018 data (last row of the figure), 2017 group i is split mostly into 2015-2018 groups 16(53%) and 20 (40%). In contrast, 79% postcodes of 2017 group e are found in 2015-2018 group 14.

592 entirely surprising because of the presence of the 2017 data in both analyses.
We did not understand this comment. We suspect this might be a leftover from a copy-paste of our text.

Line 334. The end of the sentence in the Figure' legend may be missing.
Corrected. "The dendrogram was obtained using an agglomerative hierarchical clustering" (P13 line 370)

Figure 3. I get the point that authors focused on readability of the graph, choosing not to show the pesticide names, that is understandable. However, given the key issue of the study, i.e. "identify the number and composition of pesticide cocktails potentially occurring in French farmland" and provide lists of compounds that could of concern for regulation and for combined tests of toxicity, it would definitely be better to have the names of the pesticides at least in one principal figure of the manuscript. Moreover, large parts of the results / discussion mention pesticides that cannot be seen on information within the text but only as Supplementary Information. I recommend modifying the Figure, adding the names of the compounds in Figure 3 as it is done in Figure S8.
We understand the importance of showing substances name, but to balance between something readable, while keeping a one-page figure allowing an overview of purchase probability of all substances, we only put core substance names, as well as the names of substances referred to in the text, on this figure (P16).

Lines 458-462. An interaction between the factors "type of crops" and "geographic location" is likely to be expected because plant pathogens and organisms considered as pests are heterogeneously distributed over space at the national scale due to climate and to distribution range. This means that for a given type of crops, the composition of pesticide mixture may be expected to vary between north/south and east/west for instance. While for a given geographic region the mixture of pesticides may be expected to differ between crops. Except for broad-spectrum pesticides that are not specific to one target plant protection issue only but may be used under various contexts.
This was indeed expected and we now mention this expectation in the introduction (see P3, lines 106-110). Going into more details into the mechanisms responsible for the relationship (or lack thereof) between crop type and pesticide purchases would be necessary, but would require the development of new statistical tools, which is beyond the scope of this manuscript.

Another important issue is about the "spatially closer postcodes groups": what is the extent, the spatial scale considered as "close"?

We agree that "close" is somewhat vague, and we changed it to adjacent/non-adjacent / neighbour (P19, l. 563 and P25 l.740)

Could we expect a role of farmer's cooperatives, regulatory of administrative bodies (e.g. "chambre régionale d'agriculture") and advisers/dealers of pesticides sellers on influencing the geographic patterns and thus the correlation between the geographic distance and active substance compositions of groups? If so, what is the spatial scale of correlation expected? Regional more than local or the other way round?

Although we did not find a significant correlation between the geographic distance and the active substance composition of groups (P19 l.561), we discussed the potential causal effects of pesticide market/distribution on our results (P25, l. 744). However, going further into such potential effect would require an in-depth analysis of the pesticide market and farming advice that we think is beyond the scope of study.

How to disentangle from the effects of climate/spatial distribution of pests/spatial distribution of crops?

We rewrote the introduction to make it clear that our approach primarily describes a pattern, but does not aim at identifying the mechanisms. We also added a paragraph in the discussion (P25, l. 744) to mention the possible drivers of this clustering: climate and pest outbreaks, and provide perspectives on how to discriminate them. For example, we propose a model-based approach with environmental covariates in the model and test whether and how clusters are modified (P25 l.726-728). Again, addressing the drivers of the spatial patterns of pesticide purchases would require the development of new statistical tools, which is beyond the scope of this manuscript.

Line 501. Thiram… tebuconazole

This has been changed in the main text, lines 22,452,607

Line 512. "substances are used with in the buying area", with?

Corrected. "substances are used within the buying area" (P21 line 618)

Line 531. "effects have already been studied. but mostly on pairs of substances". Please check, the dot likely not should be there.

This has been changed in the main text (P21 line 638)

Line 533. A few recent studies under controlled conditions addressed broader mixtures, for instance using mixtures of herbicide/insecticide/fungicide or even larger using soils sampled in *natura*. See for instance:

Glinski, D.A., Purucker, S.T., Van Meter, R.J., Black, M.C., Henderson, W.M., 2019. Endogenous and exogenous biomarker analysis in terrestrial phase amphibians (Lithobates sphenocephala) following dermal exposure to pesticide mixtures. Environ. Chem. 16, 55–67. https://doi.org/10.1071/en18163

Panico, S.C., van Gestel, C.A.M., Verweij, R.A., Rault, M., Bertrand, C., Menacho Barriga, C.A., Coeurdassier, M., Fritsch, C., Gimbert, F., Pelosi, C., 2022. Field mixtures of currently used pesticides in agricultural soil pose a risk to soil invertebrates. Environ. Pollut. 305, 119290. https://doi.org/10.1016/j.envpol.2022.119290

Van Meter, R.J., Glinski, D.A., Purucker, S.T., Henderson, W.M., 2018. Influence of exposure to pesticide mixtures on the metabolomic profile in post-metamorphic green frogs (Lithobates clamitans). Sci. Total Environ. 624, 1348–1359. https://doi.org/10.1016/j.scitotenv.2017.12.175

Thank you for these references, we have included some of them in the text (P22 line 639).

Line 550. It would be necessary to address further comparisons with environmental surveys to check whether the core and discriminating substances identified in this study in group of postcodes have indeed been also detected in the environment or biota over the given postcodes. And also to check for the "crop effect" among other factors. The reference to the results of Silva et al (2019) is indeed relevant, and could be even more used to compare the results on the influence of crop type and geographical gradients (N-S, E-W). However, one might consider it of marginal significance to support the findings about the compounds characterizing the groups in the present French study and where they are used at national level. The fact that the compounds were indeed found (or not found) in field surveys in France could support the conclusions of the study about the relevance of the approach to identify mixtures of concern for environmental risks. The main trouble is of course to find data about current pesticide screening in soil, water, air or biota that could be used to compare with the data on purchase in France. Although incomplete in terms of selected screened compounds, year of sampling, location etc… some references/dataset may be useful for comparisons with the sales of pesticides presented here to ensure the reliance of the study. The authors may consider for instance the following:

https://data.eaufrance.fr/ - screenings of banned and currently used pesticides in surface water are performed regularly at the national level and the data are available.

https://www.atmo-france.org/article/phytatmo - in some years, measurements are performed in air samples using a multi-residue analytical menu screening many compounds and results are delivered by region or district.

Décuq, C., Bourdat-Deschamps, M., Benoit, P., Bertrand, C., Benabdallah, R., Esnault, B., Durand, B., Loubet, B., Fritsch, C., Pelosi, C., Gaba, S., Bretagnolle, V., Bedos, C., 2022. A multiresidue analytical method on air and rainwater for assessing pesticide atmospheric contamination in untreated areas. Sci. Total Environ. 823, 153582. https://doi.org/10.1016/j.scitotenv.2022.153582

Fritsch, C., Appenzeller, B., Burkart, L., Coeurdassier, M., Scheifler, R., Raoul, F., Driget, V., Powolny, T., Gagnaison, C., Rieffel, D., Afonso, E., Goydadin, A.-C., Hardy, E.M., Palazzi, P., Schaeffer, C., Gaba, S., Bretagnolle, V., Bertrand, C., Pelosi, C., 2022. Pervasive exposure of wild small mammals to legacy and currently used pesticide mixtures in arable landscapes. Sci. Rep. 12, 15904. https://doi.org/10.1038/s41598-022-19959-y

Panico, S.C., van Gestel, C.A.M., Verweij, R.A., Rault, M., Bertrand, C., Menacho Barriga, C.A., Coeurdassier, M., Fritsch, C., Gimbert, F., Pelosi, C., 2022. Field mixtures of currently used pesticides in agricultural soil pose a risk to soil invertebrates. Environ. Pollut. 305, 119290. https://doi.org/10.1016/j.envpol.2022.119290

Pelosi, C., Bertrand, C., Bretagnolle, V., Coeurdassier, M., Delhomme, O., Deschamps, M., Gaba, S., Millet, M., Nélieu, S., Fritsch, C., 2022. Glyphosate, AMPA and glufosinate in soils and earthworms in a French arable landscape. Chemosphere 301, 134672. https://doi.org/10.1016/j.chemosphere.2022.134672

Pelosi, C., Bertrand, C., Daniele, G., Coeurdassier, M., Benoit, P., Nélieu, S., Lafay, F., Bretagnolle, V., Gaba, S., Vulliet, E., Fritsch, C., 2021. Residues of currently used pesticides in soils and earthworms: A silent threat? Agric. Ecosyst. Environ. 305, 107167. https://doi.org/10.1016/j.agee.2020.107167

Prouteau, L., 2021. Caractérisation de la contamination en pesticides azoles et néonicotinoïdes chez les espèces d'intérêt localisées en région Nouvelle-Aquitaine : développement de méthodes analytiques et applications (Thèse de doctorat). La Rochelle Université, La Rochelle, France.

We agree that a comparison of the purchase data to environmental contamination data would strengthen our results. To test whether data on pesticide purchase can be representative of environmental contamination we used a dataset on water contamination in French rivers and water bodies (Naïades; http://www.naiades.eaufrance.fr/acces-donnee). This database contains the results of analyses performed by water agencies and environmental consultancies targeting selected active substances, 28 of which were present in the database on pesticide purchases. For each of these substances and for each location of water analysis, we averaged the quantity of the active substance measured in the water over all samples collected in 2017. For each substance and postcode, we averaged measured quantities (mg/L) across all sampling sites located in the postcode. As the pesticide residues sampled in the water of one postcode could originate from neighbouring postcodes, we weighted the quantity of substances sampled in a postcode based on the samplings in the neighbour postcodes in a buffer of 10km. For the 28 active substances common to both databases, we calculated the Pearson correlation between the amounts purchased and the quantity measured in water bodies in the postcode and the 10km-buffer neighbourhood. Of these 28 substances, we found a mean correlation of 0.28 (range: -0.05 ; 0.68 ; Figure C and Table A) between the Naïade database and the BNV-d database. Although not very strong, such correlations suggest that the pesticide purchase data we used are indeed related to environmental contamination. We agree that further investigation are needed to ascertain the links between pesticide purchase and environmental contamination but this might need further data to be fully informative. For example, depending on substance characteristics, such as degradability, and the way they are used, such as weather conditions during spraying, we could expect variations in the level of environmental contamination.
A such, we propose not to include this preliminary analysis in the manuscript but stress the need for further detailed analysis (P22 l.656-660)

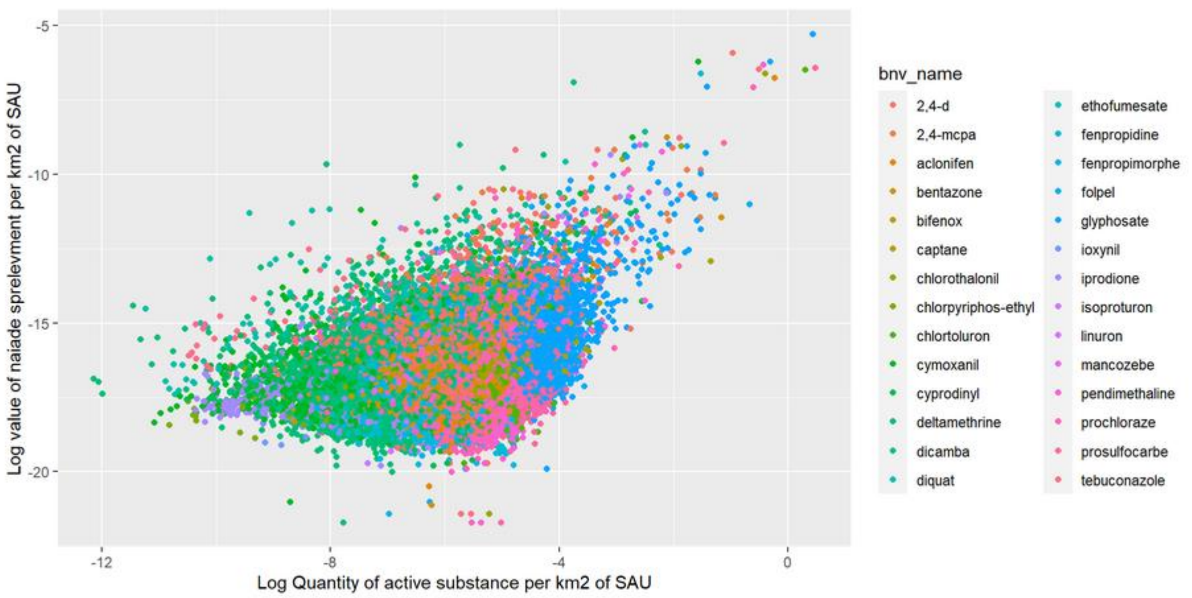


Figure C. Relationships between the quantity of active substances found in surface water from the Naïade database and the quantity of active substances purchased in a 10km buffer around the Naïade sites calculated from the BNV-d database.

Table A. correlations between the quantity of active substance found in surface water from the Naïade database and the quantity of active substance purchased in a 10km buffer around the Naïade sites calculated from the BNV-d database.

substance name	correlation
2,4-d	0.2702
2,4-mcpa	0.2852

aclonifen	0.2569
bentazone	0.3398
chlorothalonil	0.1393
chlortoluron	0.3238
cymoxanil	0.4019
cyprodinyl	0.0925
deltamethrine	0.2543
dicamba	0.2061
diquat	0.2043
folpel	0.4324
glyphosate	0.4719
iprodione	0.5279
mancozebe	0.3172
pendimethaline	0.158
prochloraze	0.1099
prosulfocarbe	0.196
tebuconazole	-0.0564
chlorpyriphos-ethyl	0.3845
ethofumesate	0.389
fenpropidine	0.2246
fenpropimorphe	0.2269
isoproturon	0.3763
bifenox	0.4347
linuron	0.689
captane	0.4788
ioxynil	0.4516

Review by Patrice Couture

This is an important paper, I would even say a much awaited paper, that paves the way for environmentally-relevant ecotoxicological studies on the effects of pesticide mixtures on wildlife. As the authors correctly point out, although several ecotoxicological studies have examined pesticide mixtures, the pesticides in the mixtures are usually selected based on general use over large territories or countries, or targeting some crop types, and may not correspond to relevant mixtures to which particular species of wildlife are exposed. The approach that the authors take is highly innovative and useful. It makes use of the postcodes of pesticide buyers for 279 different molecules in France. After making a number of well explained and reasonable assumptions, they group the 5,361 postcodes in 18 groups that are discriminated by a small number of pesticides. Some pesticides, like glyphosate, are found in most or all 18 groups and called core pesticides, but others are associated to just one or a few of these groups and termed discriminant pesticides. The approach allows associating each of the 18 groups to a specific crop, or assemblage of crops, hence, for example, differenciating Brittany from the Southern regions, where vineyards are concentrated.

The limitations of the study are well explained and dealt with in the cautious interpretation of the data. For example, it may be that a pesticide purchased in a given postcode area at a given date is not used in the year or in the postcode in which it was purchased. The authors eliminated chemicals purchased by the SNCF that centralizes its purchases for use across the French territory. Overall, the data are credible and the conclusions prudent.

The power of this study is that this approach will allow ecotoxicologists to study mixtures of pesticides to which a given species is likely exposed in a given part of its distribution range. The approach taken by the authors is sufficiently well described to be applied by others in other regions and countries where a database of pesticide sales is kept.

From an ecotoxicological perspective, this study is exciting and opens new avenues to research on the impacts of pesticide use at the landscape level. The approach could even be coupled with georeferenced data on biodiversity to further orient studies of the effects of these mixtures.

I do not have the competence to make a judgment on the statistical approach, and that part will need to be assessed by someone with the relevant expertise.

Finally, the style is clear and the English is excellent. A pleasure to read. I attach a pdf to indicate a few corrections.

Thank you very much for the supportive comments. We have incorporated all suggested changes.