

Dear reviewers,

We warmly thank the two reviewers for their relevant and useful feedback and comments on our paper titled "Constraining pesticide degradation in conceptual distributed catchment models with compound-specific isotope analysis (CSIA)." We have now carefully addressed and incorporated these comments to emphasize the originality and potential of this study.

Please find below the detailed responses and modifications made (in blue) based on the comments, which have been incorporated into the revised version. The line numbers correspond to the revised manuscript, which already includes the following modifications.

Reviewer 1:

### **General comments**

The study showed that incorporating CSIA reduced uncertainty in PiBEACH model for estimating S-metolachlor concentration in soil. The study is interesting, with a good conclusion and abstract. However, the structure of the article requires MAJOR improvements. Some are mentioned below, but the authors' additional effort is also welcome.

We thank Reviewer 1 for their valuable comments. We have now reorganized the structure of the article based on these suggestions and added a new section to the Materials and Methods, introducing the six metrics.

### **Specific comments**

**Ln 11:** What do you mean by "the calibration of pesticide dissipation"? Which parameters are related to pesticide dissipation?

We have now clarified the sentence pointing the parameters involved in the pesticide dissipation.

*"This limitation complicates the calibration of parameters involved in the pesticide dissipation, e.g., biodegradation and volatilization half times ( $DT_{50}$ ) or partition coefficient ( $K_{oc}$ ) for sorption/desorption, often leading to equifinality. As a result, it hinders the reliable prediction of pesticide persistence in soil and its transportation from agricultural plots to the catchment outlet."*

**Ln 16** This is confusing; previously, you mentioned "calibration of pesticide dissipation." Typically, you calibrate the parameters of the model. Now you are mentioning "predict daily pesticide dissipation." In this case, you are saying that pesticide dissipation is like an output of the model PiBEACH. Can you explain this? It is very important to understand your research.

Thank you for highlighting this confusion. We have now clarified this issue. Pesticide dissipation in soil results from both degradation processes (biodegradation, photolysis, hydrolysis) and transport between compartments (soil to air via volatilization, soil to surface water via runoff, or soil to groundwater via leaching). Therefore, pesticide dissipation is not a direct output of the model, such as the pesticide concentration at the catchment outlet, but can be inferred from the model's internal compartments and fluxes. However, a mass balance

of pesticide fate can be calculated at each time step, allowing for the prediction of both overall pesticide dissipation and the specific fraction that is definitively degraded. The novelty of the PiBEACH model lies in its ability to refine this prediction by distinguishing between degradative and non-degradative processes, thereby reducing uncertainty.

We have now clarified the sentence to point this novelty:

*“The research was conducted in a 47-ha crop catchment, focusing on the widely used pre-emergence herbicide S-metolachlor. A novel conceptual model, named PiBEACH, was developed to predict pesticide dissipation in soil and its transport to rivers, **reducing equifinality and associated uncertainty on the extent of the different dissipation processes by incorporating changes of the carbon isotopic signatures ( $\delta^{13}\text{C}$ ) of the targeted pesticide during degradation.**”*

**Ln 23**, In the phrase “the incorporation of CSIA data into conceptual distributed hydrological models,” you should mention that only for PiBEACH, you did not test more conceptual hydrological models (perhaps the results are different). For this article, it could be beneficial to incorporate one more model.

We have revised the following sentence to express that we focused on the PiBEACH model. We agreed to explore the integration of isotopic fractionation in other models in future work and this study highlights the integration of CSIA data in PiBEACH. Therefore, the primary goal of this paper was to lay the groundwork for applying a similar approach to other pesticide models, highlighting the benefits, limitations, and additional field data required, rather than focusing on comparing the added value of CSIA across different models, which is beyond the scope of the current study. This has now been specified:

*“In summary, the incorporation of CSIA data into **the** conceptual distributed hydrological model **PiBEACH** holds the potential to alleviate parameter equifinality, therewith significantly improving our ability to predict the dynamics of pesticide degradation on the catchment scale. **With concentration of pesticide in topsoil distributed across the catchment and throughout the crop growing season, PiBEACH paves the way for coupling with distributed event-based models, such as OpenLISEM OLP, which are better suited to address the intra-event dynamics of pesticide transport.**”*

**Ln 41**. In the phrase “In addition, existing models frequently fail to accurately quantify the contribution of concomitant pesticide dissipation processes for accurate estimation of off-site pesticide transport (Gassmann et al., 2021).”

Previously, you provided information on the contribution of processes for the pesticide dissipation process, such as degradation, sorption, and volatilization. I think this term (pesticide dissipation) requires a small paragraph for a definition first because, in the phrase of Ln41, it is not clear what you mean by “the contribution of concomitant pesticide dissipation processes.....” Commonly, hydrological models simulate (quantify) “concentrations” over time or “transport”; for example, you can see SWAP-PEARL, HYDRUS, or MACRO models. Therefore, the models do not quantify pesticide dissipation processes as an output; they are mechanisms **IN** the model. When I read about pesticide dissipation, the first thing that came to mind was DT50 (degradation). Therefore,

I suggest including a small explanation of the term, changing it, or citing an author explaining the concept.

We have now revised the sentences to address this comment and clarify how the reactive transport of pesticides, encompassing various processes, leads to pesticide dissipation at the catchment scale, i.e., a reduction in pesticide concentration in the soil compartment.

*While pressure on aquatic ecosystems is increasing, the contribution of each pesticide dissipation processes in soil driving to a mass decrease and pesticide off-site transport toward the catchment outlet remains difficult to evaluate and predict accurately. In this context, reliable and validated hydrological models that predict pesticide dissipation and off-site transport have the potential to address fundamental questions, such as the relationship between hydro-climatic factors and pesticide dissipation processes and the transfer risks of pesticides on the catchment scale. The contribution of degradative processes, including both biotic or abiotic degradation, as well as non-degradative pesticide dissipation processes, such as sorption, leaching, volatilization, off-site transport, must be quantified for this purpose in models (Larsbo and Jarvis, 2005; Steffens et al., 2015; Gassmann, 2021).*

**Ln 45** In the phrase “data limitations to constrain model parameters,” Please include at the end of this paragraph “by calibration”. So, the final phrase should read “data limitations to constrain model parameters by calibration.” Remember that several pesticide parameters can be obtained by laboratory analysis.

We have now revised the sentence as suggested

**Ln 55-57** I think this phrase is key for your purpose. Maybe after this phrase, you should provide your opinion and mention that this is one of the research goals, increasing the understanding of your research.

We have added a final sentence in this section to address the challenge for the modelling community.

*Consequently, reducing the uncertainty of  $DT_{50}$  values remains a challenge at the catchment scale, even though it is a priority for enhancing the predictive capacity of transport models.*

**Ln 72** Is the incorporation of CSIA an “alternative” or “complementary” approach?

We have revised the sentence as suggested:

*In this context, pesticide compound specific isotope analysis (CSIA) offers an complementary approach because information on parent pesticide degradation is derived independently from concentrations and produced transformation products, and relies on stable isotope ratios (Elsner and Imfeld, 2016; Hofstetter et al., 2024).*

**Ln 138** It mentioned soil samples from 13 marked plots. I GUESS you mean the black dots in Figure 1. We still do not know which samples were obtained from weekly measurements (green dots in Fig. 1). The methodology needs to be organized better.

We have now revised the caption of the Figure 1, and the following sentence, to explain the weekly weighted samples of topsoil for each transect from regular sampling (green dots) along transects (see Alvarez-Zaldivar et al., 2018).

Figure 1: The headwater catchment of Alteckendorf (Bas-Rhin, France), illustrating the experimental setup, which includes transects (weighted samples collected at the green dots along the red transects) and plot sampling (black dots), as well as the land use for 2016. The "Other" category encompasses roads, grass strips, and orchards.

Lines 146-148

Soil samples were collected from 13 marked plots (before S-metolachlor applications and on days 1, 50, and 100 post-application), as well as from the north, valley, and south transects across the catchment (weekly weighted samples from the green dots in Fig. 1, spanning from March 19th, i.e., day 0, to July 12th, i.e., day 115). These samples were analyzed for the quantification and CSIA ( $\delta^{13}\text{C}$ ) of S-metolachlor in soil

**Ln 149** I suggest including the total number of observations you have for calibration. You have three types of observations to include in an objective function. Concentrations of S-metolachlor in the topsoil, CSIA, and runoff discharge. It is important to know that number because you are calibrating 25 parameters (which is a lot!) In my experience, you should not calibrate more than five, which still is a lot. Otherwise, you get problems with parameter uncertainties.

Thank you for the suggestion. We have mobilized a total of 103 topsoil samples (i.e., 13 plots  $\times$  4 sampling dates + 3 transects  $\times$  17 weeks) to measure S-metolachlor concentration and  $\delta^{13}\text{C}$  values across the catchment and throughout the crop growing season. This effort was aimed at constraining the topsoil dissipation of S-metolachlor, as well as 115 daily discharge measurements to assess the overall hydrological response of the catchment.

The S-metolachlor concentrations and associated  $\delta^{13}\text{C}$  values monitored at the catchment outlet were not incorporated into the calibration process, as PiBEACH, using daily rainfall data, was not designed to simulate accurate sub-daily intra-event runoff, which accounts for a significant portion of seasonal export (Lefrancq et al., 2012). However, the S-metolachlor concentrations at the outlet, derived from 51 samples, were used for external validation to assess the limitations of PiBEACH, particularly in terms of S-metolachlor export at the outlet (Panel C, Figure 5, and the new section in Materials and Methods). We recognize that calibrating 25 parameters is a considerable undertaking, which justifies the dataset collected across the catchment and throughout the growing season, including the original use of isotopic signatures. We have now addressed the issue of parameter uncertainty with a rigorous methodology, which allowed us to accurately associate our results with the corresponding confidence intervals, as shown in Figures 3, 4, and 5.

We added this information in the manuscript (Line 447):

Calibrating 25 parameters is challenging and justifies the extensive dataset collected across the catchment and throughout the growing season. This unique dataset included the use of isotopic signatures and comprises a total of 103 topsoil samples analyzed for S-metolachlor concentration and  $\delta^{13}\text{C}$ , 115 daily discharge measurements, and 51 river outlet samples with S-metolachlor concentration.

**Ln 170.** OK, PiBEACH needs meteorological records, among other things, but it did not provide details of how those measurements were taken or analyzed in laboratory conditions.

We acknowledge the reviewer for this useful comment, and we have now added the following sentences.

In section Field site - Lines 148-150

*Daily temperature and potential evapotranspiration were obtained from the MeteoFrance station in Waltenheim sur Zorn located 7 km southwest. Daily rainfall was derived from a tipping bucket rain gauge (Précis Mécanique®, Fig. 1).*

In section PiBEACH development - line 205

*BEACH was fed by daily meteorological records, including rainfall and mean air temperature and potential evaporation, [...].*

**Ln 158** If you mean Supplemental information by "SI", you should put the definition of the acronym one time.

Indeed, thank you. We now use "Supplement" without an acronym, as recommended in the HESS guidelines.

**Ln 178** Can you please mention the mixing topsoil layer length in this phrase? The mixing layer is generally a challenging topic and, in my opinion, an unsolved problem for hydrological models.

This information (first centimeter, from 0 to 1 cm) was originally provided 25 lines after the line 178. We have now included this information as recommended in line 214:

*Compared with BEACH features, PiBEACH includes (i) a mixing topsoil layer (from 0 to 1 cm, McGrath et al., 2008) and additional deepest soil layers to include the shallow groundwater contribution to discharge, (ii) daily changes of topsoil temperature (Neitsch et al., 2011), (iii) daily variations of topsoil hydraulic properties accounting for the impact of agricultural practices with a crop-specific agronomical model (Lefrançois et al., 2018), (iv) pesticides first-order degradation and linear sorption/desorption processes, and (v) pesticide carbon stable isotopic fractionation and transport of isotopologues (e.g., <sup>13</sup>C and <sup>12</sup>C) (Lutz et al., 2017; Alvarez-Zaldivar et al., 2018). PiBEACH requires amounts and dates of pesticide application for each plot.*

As mentioned by the reviewer, the mixing layer concept is probably the most commonly used in pesticide fate models due to its numerical simplicity and its realistic representation, as hortonian runoff processes have repeatedly been shown to be initiated as a near-surface process (Gao et al., 2004; McGrath et al., 2008; Wallender et al., 2008). This concept assumes that the transport is controlled by a mixing layer below the soil surface, in which rainwater, soil solution and runoff water completely and instantaneously mix (Havis et al., 1992). The soil depth interacting with runoff water is variable ranging from 0.25 to 2 cm according to various experimental and modelling studies (Ahuja et al., 1981; Havis et al., 1992; Wallender et al., 2008). Previous studies have shown that the effective depth of interaction is related to the degree of soil aggregation and increases with soil slope, kinetic energy of raindrops and rainfall

intensity (Havis et al., 1992). In fact, the observed mixing-layer depth is often much shallower than the depth required for fitting models to field data (Gao et al., 2004).

Ahuja, L.R., Sharpley, A.N., Yamamoto, M., Menzel, R.G., 1981. *The Depth of Rainfall-Runoff-Soil Interaction as Determined by P-32*. *Water Resources Research*. 17, 969-974.

Gao, B., Walter, M.T., Steenhuis, T.S., Hogarth, W.L., Parlange, J.Y., 2004. *Rainfall induced chemical transport from soil to runoff: theory and experiments*. *Journal of Hydrology*. 295, 291-304.

McGrath, G.S., Hinz, C., Sivapalan, M., 2008. *Modeling the effect of rainfall intermittency on the variability of solute persistence at the soil surface*. *Water Resources Research*. 44.

Wallender, W.W., Joyce, B.A., Ginn, T.R., 2008. *Modeling the Transport of Spray-Applied Pesticides from Fields with Vegetative Cover*. *Transactions of the Asabe*. 51, 1963-1976.

Havis, R.N., Smith, R.E., Adrian, D.D., 1992. *Partitioning Solute Transport between Infiltration and Overland-Flow under Rainfall*. *Water Resources Research*. 28, 2569-2580.

The reference (McGrath et al., 2008) cited at the first mention of the mixing layer in our paper was included to provide a more in-depth understanding of this concept.

Line 214:

*Compared with BEACH features, PiBEACH includes (i) a mixing topsoil layer (assuming that the pesticide transport is controlled by a mixing layer from 0 to 1 cm below the soil surface, in which rainwater, soil solution and runoff water completely and instantaneously mix, McGrath et al., 2008) and additional deepest soil layers to include the shallow groundwater contribution to discharge.*

**Ln 201** The Model only accepts that number of layers. Or is that set by the user?

We relied on the pedological, agronomic, and hydrological knowledge of the studied catchment to develop the conceptual framework for PiBEACH. The number of layers chosen, i.e. five for the Alteckendorf catchment, seems to be the optimal number to simplify vertical heterogeneity based on agronomic criteria (cropping operations), pesticide reactivity, or hydrological functioning with variably saturated and saturated zones. Based on similar expertise, this number of layers could be adjusted for another catchment and implemented in a modified version of PiBEACH.

**Ln 220** In Eq.1, you mentioned that the balance is solved for each layer with soil depth  $Z$ . But the runoff term should not be included in deeper layers. This raises the question of whether you obtain a unique theta per cell or a theta for each layer (5 layers in the profile per cell). It would help if you clarified this, especially for HYDRUS, MACRO, openLISEM, or SWAP-PEARL users.

As noted in line 234 and in Equation 1, PiBEACH calculates the hydrological balance for each cell and layer. The hydrological balance is described in Sheikh et al. (2009), with a detailed implementation provided for each process in Section 6 of the SI. In general, the change in soil water content ( $\theta$ ,  $\text{m}^3 \text{m}^{-3}$ ) for each cell (i) and each layer, defined by soil depth ( $Z$ , mm), is calculated using Equation 1. The resulting soil water content is then available for each cell and layer, facilitating potential coupling with other models such as HYDRUS, MACRO, OpenLISEM OLP, or SWAP-PEARL.

We have moved Equation 1 to follow the description of the layers (line 255) in order to clarify how each term in the equation is applied depending on the specific layer. As noted by the

reviewer, we have added a sentence explaining that runoff (RO) is only computed for the topsoil layer ( $z_0$ ).

Line 257:

*“Runoff (RO) is calculated exclusively for the topsoil layer ( $z_0$ ).”*

**Ln 230:** You mentioned the balance at each cell I was given by Eq. 2. My question is, does this model solve the mass balance per layer (5 in your case), or does it generate a mesh with smaller cells? (Connected with the previous question)

As stated in line 212, equations 1 and 2 are solved for each cell of each layer  $z_0$  to  $z_3$  using a raster approach with a resolution of  $2 \times 2$  m for layers  $z_0$  to  $z_3$  (four layers in total). The deeper layer,  $z_4$ , is represented using a global linear reservoir.

The sentence was revised to clarify the application of the raster approach with a  $2 \times 2$  m resolution for solving the equations.

*To integrate landscape components enhancing or limiting pesticide off-site transport, i.e., grass strips or roads, the Alteckendorf catchment was modelled using a  $2 \times 2$  m resolution raster. The vertical soil profile was represented by five successive layers including, from topsoil to groundwater bottom (Fig. 2):*

**Ln 246.** Where is the mixing layer depth used in Eq. 4? I think you should change  $J$  by  $qx,y$  in this equation.

The equation 4 has now been corrected, and the depth of the mixing layer is now explicitly linked to  $D_{z_0}$ .

Line 269:

*“where RO was runoff (mm) and  $\beta_{RO}$  was a calibration constant ( $1 \geq \beta_{RO} \geq 0$ ) and  $D_{z_0}$  (mm) was the mixing layer ~~topsoil~~ depth (Ahuja and Lehman, 1983).”*

**Ln 296** It is not better to include  $h_{ref}$  than  $\theta_{ref}$ . The water content depends highly on texture;  $0.2 \text{ cm}^3 \text{ cm}^{-3}$  of water content could be either high or low. Instead, the  $h$  does not have that inconvenient.

We acknowledge this suggestion and we agree that water content depends on soil texture. However, we have retained and applied in this study the formulation developed by Walker (1974) and Larsbo and Jarvis (2003), which uses a ratio of soil moisture expressed in  $\theta$  ( $\text{m}^3 \text{ m}^{-3}$ ) rather than  $h$  (mm  $\text{H}_2\text{O}$ ). The reference value,  $\theta_{ref}$ , corresponds to the experimental conditions associated with the  $DT_{50ref}$  (and  $k_{ref}$ ) characterization. In these lab-scale experiments, soil saturation is typically expressed in terms of soil moisture.

**Ln 305** OpenLISEM cannot simulate preferential flow yet, to my knowledge. So, this phrase required modification.

We agree with the reviewer's assessment. OpenLISEM and its pesticide module OLP are not capable of simulating preferential flow and are better suited for catchments with a significant risk of runoff and associated erosion, such as in Alteckendorf during storm events. As a result, we have revised the sentence to focus solely on runoff.

Line 363:

Considering a daily time step, PiBEACH was not designed to accurately simulate ~~fast flow dynamics, i.e., runoff genesis and dynamic across the catchment and preferential water and soil flow at the headwater scale~~ (Sheikh et al., 2009).

**Ln 350** Which threshold did you use to reduce parameters from 43 to 25. Basically, how did you define a parameter as highly sensitive? Additionally, you should mention which outputs were analyzed during sensitivity analysis, i.e., did you include sensitivities to isotope output?

We have calculated and analyzed the mean and standard deviation of the elementary effects (EE) for each parameter, as required by the Morris method. The mean value represents the overall effect of each parameter on the model output, while the standard deviation reflects the interaction between inputs. We excluded 21 parameters where the mean value was zero or close to zero, indicating no significant influence on the output. The outputs considered in this sensitivity analysis included S-metolachlor concentrations at the outlet, composite soil transects (North, Valley, and South), discharge at the outlet, and isotope signatures at both the outlet and composite soil transects. The mean and standard deviation values for the 25 retained parameters are presented in Figures S5 and S6 in the Supplement, for the different output variables.

**Ln 370** Where did you previously mention "daily discharge at the outlet"? You mentioned weekly runoff. The observations are unclear, and you need to structure the methodology better.

As described in a previous paper (Alvarez-Zaldivar et al., 2018), "discharge at the catchment outlet was continuously measured by a Doppler flow-meter (2150 Isco, Lincoln, Nebraska, USA) with discharge precision of 3% and a time step of 2 minutes. Water was collected by flow proportional sampling using a refrigerated automatic sampler with a total capacity of 3.96 L divided into 12 glass vials each of 330 mL (Isco Avalanche, Lincoln, Nebraska, USA). A predefined discharged volume based on the seasonal rainfall intensity expected for April (50 m<sup>3</sup>), May (100 m<sup>3</sup>) and June (150 m<sup>3</sup>) was chosen allowing for 36 aliquots of 110mL each per week. Water samples were then combined into composite samples according to hydrograph characteristics (base-flow, rising and/or falling limb), yielding one to four samples weekly of volumes >= 990 mL. "

The weekly time step refers only to the field visits for collecting and replacing samples, not to the sampling frequency used to acquire data. We have revised the relevant sentences to clarify this distinction.

Line 175:



Outlet runoff discharge was measured using a Doppler flowmeter (2150 Isco) *with a precision of 3% and a time step of 2 minutes*. Automatic, refrigerated continuous flow proportional sampling (Isco Avalanche with twelve bottles of 330 mL) was conducted at fixed weekly discharge volumes ranging from 50 to 150 m<sup>3</sup> from April to June to capture increasing minimum baseflow discharges observed in 2016 (Alvarez-Zaldivar et al., 2018). To obtain sufficient amount of S-metolachlor for quantification and CSIA, samples were combined *weekly in composite samples based on hydrograph, i.e., base-flow, rising and falling limb, yielding one to four samples weekly of volumes ≥ 990 mL* (Alvarez-Zaldivar et al., 2018).

**Ln 373** OK, you explain that in addition to incorporating CSIA data in the objective function, you will analyze the spatial distribution of this observation. For better understanding, you should use Fig 1 at this point.

*As suggested by the reviewer, we have revised the relevant sentence to clarify this point.*

Lines 436:

*For the latter, KGE metric incorporates topsoil S-metolachlor concentration and  $\delta^{13}\text{C}$  from (i) individual plot observations (13 black dots in Fig. 1), (ii) aggregated plot observation along three transects across the catchment (North, Valley and South transect, Fig. 1) and (iii) transect composite soil (green dots, Fig. 1).*

**Ln 393** If the graph (Fig 3C), the “y” axis is in DT<sub>50</sub>, please use DT<sub>50</sub> instead of  $k_{\text{eg}}$  during the discussion. Additionally, I do not know how this graph and analysis (Fig 3C) provide useful information for the research. All that you mentioned in the graph and discussion can be easily obtained from Eq. 8, derived from another hydrological model, so basically, it is something simple and well-known.

*As suggested by the reviewer, we have now revised the relevant sentence, using DT<sub>50</sub> instead of  $k$  to discuss the temporal variation in S-metolachlor persistence.*

Lines 476:

*Predicted low S-metolachlor degradation rates in the first 10 days of observation ( $DT_{50} = 46 \pm 17$  days) corresponded to an early cold ( $7.1 \pm 1.7$  °C) and dry period (rainfall of 7.4 mm) (Fig. 3A). In contrast, increasing degradation rates ( $DT_{50} = 22 \pm 2$  days) from 50 to 120 days following the first application of S-metolachlor were mainly associated with higher temperatures ( $16.9 \pm 3.6$  T°C).*

**Ln 400** The graph should be improved (.svg). Additionally, you mention the error bars linked to soil bulk density. I think this is important and should be mentioned in the methodology (please improve the organization of methods).

*The figures 3, 4, 5 will be send in HD in the final version.*

As stated in the caption of Figure 3, both observed and predicted soil water content are expressed in  $\text{m}^3 \text{ water m}^{-3} \text{ soil}$ . Additional details have been provided in the Materials and Methods section, explaining how gravimetric measurements obtained in the laboratory (drying samples in an oven set at  $110^\circ\text{C}$ , NF ISO 1146) were converted into volumetric values using soil bulk density. The topsoil bulk density fluctuates throughout the growing season and was estimated daily based on the reference soil bulk density measured below the plow layer. ( $1.5 \pm 0.09 \text{ kg m}^{-3}$ ) and the agronomic module of PiBEACH (Section 7.2 in the Supplement), which was previously validated in Lefrancq et al. (2018, HP).

This information was added in line 170:

*Volumetric topsoil water content ( $\text{m}^3_{\text{water}} \text{ m}^{-3}_{\text{soil}}$ ) was derived from gravimetric water content after drying the samples in an oven set at  $110^\circ\text{C}$  (NF ISO 1146; Lefrancq et al., 2018). This calculation accounted for the seasonal dynamics of topsoil bulk density, as predicted by PiBEACH and detailed in the Supplement (section 7.2).*

In caption of Figure 3 we have revised the targeted sentence:

*The associated error bars are linked to soil bulk density estimations required for gravimetric to volumetric conversions and from the standard error on gravimetric water content measurements.*

**Ln 404** Please indicate if metolachlor concentrations are in liquid or total concentrations (also, this should be explained in the methodology).

We agree with the reviewer's that information on S-metolachlor concentration was missing. Although the filtration methodology was already described (see below and in line 183 of the manuscript), we have now clarified the methodology for river samples and added a sentence (Line 557) in the results section to clarify this point. S-metolachlor concentrations in the particulate phase (i.e.,  $> 0.7 \mu\text{m}$ ) collected in the river were consistently below the quantification limit. Therefore, S-metolachlor concentrations used in the calibration step for the 2016 campaign were based solely on the dissolved phase of the samples collected in the river at the outlet of the catchment.

Line 183:

*To separate S-metolachlor dissolved and particulate phases, river water samples were filtered through  $0.7 \mu\text{m}$  glass fibre filters. Methods of S-metolachlor extraction from soil and water samples with an AutoTrace 280 Solid Phase Extraction (SPE) system (Dionex) with SolEx C18 cartridges (Dionex<sup>®</sup>) and quantification with GC-MS (MS, ISQ™, Thermo Fisher 173 Scientific) were detailed previously (Alvarez-Zaldivar et al., 2018; Gilevska et al., 2022b). Quantification limits for S-metolachlor were  $0.01 \mu\text{g L}^{-1}$  from water and  $0.001 \mu\text{g g}^{-1}$  from soil and total suspended matter samples (d.w.), with a total analytical uncertainty of 8 % and 16 %, respectively.*

New sentence (line 557)

*It should be noted that the S-metolachlor concentrations used to estimate observed S-metolachlor export in river at the outlet of the catchment were based solely on the dissolved phase of river samples. S-metolachlor concentrations in the particulate phase (i.e., > 0.7 μm) of these samples were consistently below the limit of quantification.*

**Ln 420** OK WIC calibration reduced uncertainty, but what happens with the other 24 parameters? In the case of NIC or WIC calibration, do they maintain the uncertainty bounds? Or perhaps it changes, and how? This is important to analyze because that uncertainty may be translated to other parameters. Additionally, is there some field (laboratory) information on the common parameter values for this compound? especially DT50. I agree that DT50 could be one order of magnitude different from the expected one. Still, it is welcome to have the laboratory reference to see if calibration did not modify parameters to “unexplainable” values. That also holds for the 24 remaining parameters that you calibrate. That kind of analysis is expected from this kind of research.

We hypothesized that  $DT_{50, \text{reference}}$  was the primary parameter affected by incorporating stable isotope signatures, as it is the sole parameter governing degradation in PiBEACH. As expected,  $DT_{50, \text{reference}}$  was affected. Specifically, its uncertainty bounds were reduced when  $\delta^{13}\text{C}$  was included in the objective function (WIC calibration). The newly calibrated WIC range for  $DT_{50, \text{reference}}$  ( $18 \pm 4$  days) remained consistent with reported experimental values in the literature. In contrast, the uncertainty bounds for the other 24 parameters were not significantly altered when comparing NIC and WIC calibrations (data not shown in the current version).

In response to the Reviewer’s comment, we have now added a new section in the Supplement (Section 12). This includes a new Figure S8, demonstrating, using  $K_{OC}$  as an example, that CSIA data (WIC) did not reduce the uncertainty of 24 other parameters compared to NIC. This may be due to the pesticide extraction protocols used for soil samples, which account for both the dissolved phase and the extractable sorbed phase. Additionally, sorption generally does not lead to significant isotope fractionation (see also the response to Questions 2 and 3 in the Introduction, addressed by Reviewer 2). In this case, CSIA may not help evaluate sorption and thus constrain  $K_{OC}$  values as isotope signatures for pesticides in the dissolved and particulate phase of soil samples are generally similar.

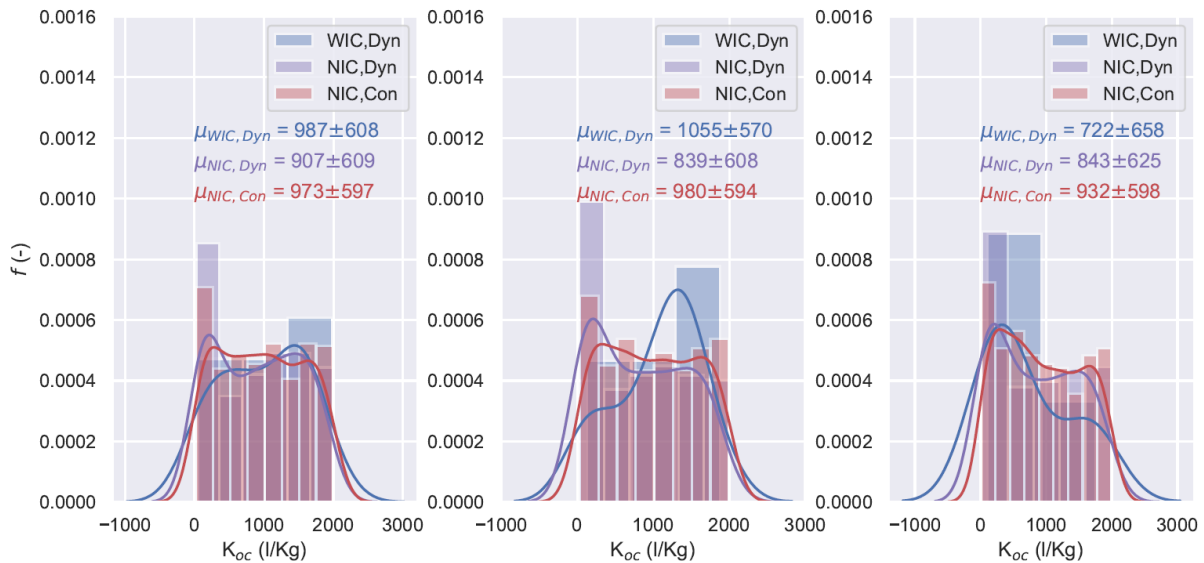


Figure S8: Distribution (out of a total of  $n = 2,500$  runs) of  $K_{oc}$  calibrated with no isotope constraint (NIC,  $n = 672$ ) and with isotope constraint (WIC,  $n = 244$ ) at three sampling resolutions (i.e., composite transect, transect and plot soils). NIC models considered  $KGE_{SM} > 0.5$  and  $KGE_Q > 0.5$ , while WIC models considered  $KGE_{SM} > 0.5$  and  $KGE_Q > 0.5$  and  $KGE_\delta > 0.8$ . Statistics for  $K_{oc}$  distributions are provided as mean (blue for WIC, purple and red for NIC, with  $DT_{50dynamic}$  depending of soil moisture and temperature or  $DT_{50constant}$ , respectively) and standard deviations ( $\mu \pm SD$ ).

**Ln 439.** OK, it would be nice to see if there are significant differences in the comparison variance. The plot situation showed a smaller reduction in the SD. The “plot” may not be recommended, but you need ANDEVA or something like that for that affirmation.

We agree with the reviewer’s observation. The  $DT_{50}$  distributions calibrated with and without isotopic constraints were similar at the plot resolution but differed at the transect and transect composite sample resolutions. We have now entirely revised this statement to avoid any affirmation.

Line 521:

The mean calibrated  $DT_{50,Ref}$  values for transect composite soil, transects across the catchment, and plot resolutions differed by less than four days (Fig. 4). However, the uncertainties, as indicated by standard deviations, were significantly lower—approximately half—with isotopic constraints (WIC) compared to no isotopic constraints (NIC) for transect composite soil and transects, while they remained unchanged at the plot resolution. This indicates that higher-resolution soil sampling (plots > transects > transect composite samples) did not improve significantly model calibration in this catchment. Therefore, pooling transect samples into a single composite soil sample for CSIA appears sufficient to constrain degradation parameters, while also reducing sampling and analytical efforts. This approach may effectively capture temporal variations in pesticide degradation at the catchment scale.

**Ln 444** “Modeling results”: Do you mean calibration results?  $\epsilon_c$  was calibrated? In this case, did you compare it with previous studies on that parameter? Something like that should be expected for DT50 and perhaps other calibrated parameters. (Similar comment that above)

We have now revised the sentence to clarify that  $\epsilon_c$  was indeed one of the 24 calibrated parameters, and the calibrating value is closed to the experimental reference derived from lab-scale study on the Alteckendorf soil (Alvarez-Zaldivar et al., 2018). In response to the Reviewer's question (see Point 16 in Reviewer 2's section), we have now clarified the different values of the isotope fractionation factor ( $\epsilon_c$ ) used in this study.

Line 536:

*Additionally, the incorporation of CSIA into PiBEACH enabled the calibration of the carbon isotope fractionation value ( $\epsilon_c$ ) for S-metolachlor degradation in topsoil. The calibration yielded  $\epsilon_c = -2.7 \pm 0.6 \text{ ‰}$ , which aligns with the prevailing transformation pathway suggested for this catchment (Alvarez-Zaldivar et al., 2018). Although the  $\epsilon_c$  value used in this previous study ( $\epsilon_c = -1.5 \pm 0.5 \text{ ‰}$ ) may have slightly overestimated degradation extent, it remained within uncertainty ranges obtained in the present study from the model ensemble.*

**Ln 457** Why did you mention that degradation obtained with the model was like observations? In which part of the methodology did you explain that compound degradation was performed under laboratory conditions or other? You need to improve the structure of the article.

We thank the reviewer for their comments. We have now carefully checked and reorganized the structure of the article by adding a new section to the Materials and Methods, which introduces the six metrics. Additionally, we have now provided a detailed explanation of how the observed values, i.e. specifically topsoil concentration,  $\delta^{13}\text{C}$ , and discharge and concentration in river at the outlet, were used to derive the observed values for three of the six metrics.

Line 449:

## **2.7 Metrics of pesticide persistence and transport risks**

*Six metrics were derived from PiBEACH outputs to assess the persistence and transport risk of S-metolachlor at the catchment scale. These metrics were expressed as percentages of the applied S-metolachlor and included: (1) S-metolachlor degradation, (2) the remaining bioavailable mass of S-metolachlor (BAM) in the topsoil (i.e., dissolved and reversibly sorbed S-metolachlor), (3) S-metolachlor export to the catchment outlet via runoff and drainage. The metrics 1 to 3 were derived from observations to evaluate the predictability of PiBEACH. Specifically, the observed remaining bioavailable mass of S-metolachlor (BAM) in the topsoil was calculated through spatial extrapolation of weekly topsoil samples. The observed S-metolachlor degradation was estimated using the mean  $\delta^{13}\text{C}$  derived from the same spatial extrapolation, combined with the enrichment factor ( $\epsilon$ ) obtained from microcosm lab-scale experiments using soil from Alteckendorf. Finally, S-metolachlor export to the catchment outlet was calculated by integrating continuous discharge measurements with sub-weekly S-metolachlor concentration samples from the river. (4) off-site transport of S-metolachlor*

*through leaching, (5) the remaining mass of irreversibly sorbed S-metolachlor (i.e., aged S-metolachlor), and (6) S-metolachlor volatilization, from the first application on March 19 (day 0) to July 12, 2016 (day 115). The metrics 4 to 6 remained difficult to estimate using the catchment-scale sampling campaign and were not computed.*

**Ln461** Here, we explain how the BAM% was obtained, but this is a methodological part. This explanation is now included in the newly added Materials and Methods section.

Line 460:

*Specifically, the observed remaining bioavailable mass of S-metolachlor (BAM) in the topsoil was calculated through spatial extrapolation of weekly topsoil samples*

#### **Technical comments**

**L10.** to the limited availability of field data **capable of distinguishing** between. Please change by “distinguishing”.

As suggested by the reviewer, we have revised the sentence:

Line 11:

*Predicting pesticide dissipation at the catchment scale using hydrological models is often challenging due to the limited availability of field data that can distinguish between degradative and non-degradative processes.*

**L17.** incorporating changes **of** the carbon isotopic signatures. Please change “of” by “in”.  
we modified the sentence as suggested

**L36.** leaching, volatilization, off-site transport, Please included “**and**” before off-site.  
we modified the sentence as suggested

L46. each of which generally **accounting** for several physico-chemical. Please modify by accounts.  
we modified the sentence as suggested

L59. in topsoil and at the catchment outlet **is** currently. Please change “is” by “are”.  
we modified the sentence as suggested

L70. it may also generate unknow and. Please change “unknow” by “unknown”.  
we modified the sentence as suggested

And some more that you can check.

We have thoroughly revised the manuscript to correct any spelling errors.

---

#### **General comment**

Payraudeau et al. present a novel modelling study advancing the field of simulation of pesticide and isotope ratios at catchment scale. Studies as this one are rarely conducted but very important to understand more on pesticide fate at catchment scale. I have two main concerns that I strongly recommend that the authors take into consideration. These are elaborated below.

We thank Reviewer 2 for the relevant comments. In response, we have now revised in details the manuscript, specifically addressing the two main concerns: (i) the added benefits of this distributed model approach compared to the earlier lumped (non-spatially distributed) model study by Lutz et al. (2017), and (ii) the use of pesticide concentrations and isotope ratios at the catchment outlet.

### **Main comments**

1. A key novelty of this study is the application of a spatially distributed catchment model to simulate pesticide concentrations and isotope ratios (only in topsoil unfortunately, see main comment #2), as stated also on line 92. My first main critique to this study is that the added benefit of a spatially distributed model for this purpose is barely demonstrated nor discussed, albeit expectations are raised in lines 87-89. In that sense the added benefit of this distributed model approach compared to the earlier lumped (not spatially-distributed) model study by Lutz et al. (2017) - applied to the same catchment but with an earlier dataset limited to catchment outlet (but no topsoil) data - does not become clear. The focus in the model is more on linking degradation rates to temporal variations in temperature and soil moisture but that can also be simulated with a lumped model approach. There must be spatial heterogeneity in the catchment like in pesticide application at plot level, but also in spatial variation in soil texture and soil moisture content (Figure 1 nicely shows the spatial variations in crop types; and the descriptive text mentions various soil types). The added benefit of simulating such spatial variations much better in a distributed model than in a lumped model is not discussed. I recommend that the authors show and discuss this key advantage better as that seems the main novelty of this study. Or could a lumped model not have resulted in an equally good model fit as page 5 states that the spatial variation in soil parameters was in fact low in this (specific) catchment?

PiBEACH was initially developed as a companion model for distributed event-based pesticide transport models, such as OpenLISEM OLP (Commelin et al., 2024), to enhance understanding and support management decisions and risk assessments at the headwater catchment scale. Event-based models are valuable tools for predicting intra-hourly rainfall-runoff dynamics and the associated off-site transport of pesticides. However, their accuracy depends on the initial conditions, which are affected by hydrological dynamics and pesticide dissipation processes occurring between rainfall events. PiBEACH was designed to predict catchment properties that evolve throughout the growing season and across the catchment, such as plow layer bulk density, soil moisture, and pesticide concentrations in both dissolved and particulate fractions. PiBEACH considers the spatial distribution of (i) agricultural plots with different crops and associated pesticide applications based on farmers' technical practices (e.g., pesticide type, application date, and dosage, as shown in Fig. 1 and Table S2 in the Supplement), (ii) non-target areas, i.e. areas where pesticide applications are not intended,

such as grass strips, hedges, and roads. (Fig. 1), which affected runoff generation and propagation (Lefrancq et al., 2017), (iii) different soil types and slope profiles (Lefrancq et al., 2017), and (iv) soil moisture and temperature dynamics as a function of hillslope position.

While lumped models can capture pesticide dynamics at the catchment scale, as highlighted by Lutz et al. (2017), we hypothesized that only a distributed model can effectively support management decisions related to the crop mosaics within the catchment to mitigate pesticide fluxes at this scale.

In summary, both lumped and distributed approaches can be used to investigate pesticide dynamics throughout the growing season. However, only a distributed model can account for spatial heterogeneity across the catchment and support further management decisions.

As outlined in the following responses to the Reviewer's comments, we have now clarified in the manuscript the complementarity of these two approaches, which were developed and applied to the same catchment.

My second main critique is that the authors did not make use of the pesticide concentrations and isotope ratios at the catchment outlet (besides the topsoil data) as also reported by Alvarez-Zaldivar et al. (2018). I strongly recommend that the authors add the simulation of the catchment outlet data as reported by Alvarez-Zaldivar et al. (2018) to this work. I see this as a pre-requisite to allow to draw conclusions on pesticide degradation at the catchment scale. The model is now only calibrated to observations on the first cm of soil in the catchment. The (degradation) model is not calibrated to anything deeper than this top one cm. Calibrating the model also to the available catchment outlet data enables to draw conclusions on pesticide degradation at the whole catchment scale, and would make this a really unique study. All conclusions drawn on catchment scale are in the current version based on extrapolation of the topsoil data. Also include a discussion on the pros/cons of calibrating such models with only topsoil data (as done in this current work) vs. outlet data (as done by Lutz et al. 2017), or both. See later "other comments" that elaborate on this main comment #2.

We acknowledge the comment of the Reviewer to clarify this issue in the manuscript. We used river discharge at the outlet in the calibration function and the S-metolachlor fluxes to examine the predictability of PiBEACH (Fig. 5C). As suggested by Reviewer 1, we have now reorganized the article's structure by adding a new section in the Materials and Methods section, which introduces the six metrics simulated after calibration and derived from soil and river samples. Additionally, we have provided a detailed explanation of how the observed values, including discharge and S-metolachlor concentration at the outlet, were used to derive the observed values for three of the six metrics. As expected, given PiBEACH's limited ability to capture intra-daily runoff generation, it is unable to accurately simulate S-metolachlor fluxes at the catchment scale (Fig. 5C), representing less than 2% of the applied amount. Therefore, we opted not to include outlet concentration and isotopic signature, mainly collected during rainfall-runoff events (Alvarez-Zaldivar et al., 2018), in the calibration function to avoid a biased calibration process, which could lead to parameter fitting for the wrong reasons in an attempt to compensate for the model's conceptual limitations.

As emphasized by the reviewer, we prioritized topsoil data during the PiBEACH calibration process, specifically using 103 samples with S-metolachlor concentration and isotopic



signatures. We derived two internal variables, i.e. predicted soil temperature and water content, to extrapolate degradation across time, space, and depth. The degradation dynamics, incorporating soil temperature and water content in PiBEACH, was validated through specific lab-scale experiments on Alteckendorf soil (see Supplement, Section 9). The results enhanced the interpretation of S-metolachlor export, considering the daily limitations of PiBEACH, with predicted and observed export rates of  $2 \pm 6\%$  and 0.3% of the applied S-metolachlor, respectively.

To summarize, we leveraged a substantial portion of the data collected at the catchment outlet for both calibration (discharge in the KGE objective function) and predictability assessment (discharge and S-metolachlor concentrations to estimate export). Due to the conceptual limitations of PiBEACH, particularly its daily time step, the comparison between simulated and observed S-metolachlor reactive transport was only partially conducted, in contrast to previous lumped modelling approaches (Lutz et al., 2017).

The added value of integrating both observed S-metolachlor concentrations and isotopic signatures, at the outlet and in the topsoil, into the calibration process of distributed modelling will be further explored in the upcoming PiBEACH-OpenLISEM OLP modelling study. This study is now in collaboration with Wageningen University, the developer of the OLP pesticide module (Commelin et al., 2024).

This second major concern of the Reviewer, along with the relevant data from the river outlet, has now been incorporated into the revised manuscript, as detailed below.

#### **Additional comments:**

1. Title: why “models”? Plural? It seems only one model has been applied?

As suggested by the reviewer, we have now revised the title to use the singular form of model.

#### 2. Introduction:

1. Line 72: Explain the concept of CSIA better such that the reader can understand how information on degradation can be derived from isotope ratios.

We have revised the following sentence to clarify the connection between isotopic signature and degradation:

*This results in a kinetic isotope effect leaving a chemical imprint in the form of characteristic changes in isotope ratios of reacting molecules (Elsner, 2010). The isotope enrichment factor ( $\epsilon_{lab}$ , e.g for carbon) can be derived from closed microcosm degradation experiments and used to quantify pesticide degradation in field studies (Alvarez-Zaldivar et al., 2018).* In contrast, non-degradative pesticide dissipation processes, such as sorption generally do not result in significant isotope fractionation (Schmidt et al., 2004; Alvarez-Zaldivar et al., 2018; Droz et al., 2021).

2. Line 78: some publications have shown that sorption can lead to significant isotope fractionation effects in the spreading direction of pollution which might

be relevant in catchment studies as well. Why not test this with this model, as this seems straightforward?

Non-destructive processes, such as dispersion and sorption, can induce significant isotope fractionation, but only under specific conditions, including strong sorption behavior, early transient diffusion, and multi-step sorption in aquifers. In such environments, the slightly stronger sorption of light isotopologues can enhance enrichment in the heavier isotopologues at the leading edge of a sorptive contaminant pulse (early stages) while reducing enrichment at the trailing edge (later stages) (Eckert et al., 2013; van Breukelen and Prommer, 2008; van Breukelen and Rolle, 2012). This effect is particularly relevant when biodegradation is slow.

In our case, S-metolachlor is not expected to exhibit strong sorption behavior, as strongly sorbing contaminants typically have  $\log K_{ow} > 2$  and water solubility  $< 1$  ppm, and S-metolachlor degradation in soil was fast. Molecular diffusion, which may lead to non-equilibrium behavior due to concentration gradients between mobile and immobile phases, plays a negligible role in contaminant redistribution when S-metolachlor is first introduced into the field. Furthermore, unlike highly sorptive contaminants in aquifers, S-metolachlor does not form distinct front and tail effects since it is applied uniformly across the catchment and gradually diffuses toward the outlet. Although low but measurable isotope fractionation may occur following multiple sorption–desorption cycles (Kopinke et al., 2017), sorption-induced stable isotope fractionation is generally considered negligible for moderately sorptive pesticides like S-metolachlor in surface agro-ecosystems during seasonal transport.

For these reasons and because sorption-induced isotope fractionation is expected to be very low and negligible in our case, we did not simulate sorption-induced isotope fractionation in our model.

We have now explicitly stated this in the manuscript (line 89), and four additional references have been included:

*Non-destructive processes such as dispersion and sorption can induce isotope fractionation under conditions of strong sorption behavior, early transient diffusion, and multiple sorption–desorption cycles (Eckert et al., 2013; van Breukelen and Prommer, 2008; van Breukelen and Rolle, 2012; Kopinke et al., 2017). However, as sorption-induced isotope fractionation is considered negligible for S-metolachlor during seasonal transport in surface agro-ecosystems and biodegradation in soil was fast, it was not included in the model.*

3. Lines 77-83: I feel that neglecting isotope fractionation effects of non-transformation processes can be more convincingly elaborated.

We agree with the reviewer.

We have now specified in the manuscript (line 89):

*“Given the physico-chemical properties of S-metolachlor, abiotic degradation processes (e.g., photolysis) and their associated isotope fractionation effects were excluded from the modelling. Non-destructive processes such as dispersion and sorption can induce isotope*

*fractionation under conditions of strong sorption behavior, early transient diffusion, and multiple sorption–desorption cycles (Eckert et al., 2013; van Breukelen and Prommer, 2008; van Breukelen and Rolle, 2012; Kopinke et al., 2017). However, as sorption-induced isotope fractionation is considered negligible for S-metolachlor during seasonal transport in surface agro-ecosystems and biodegradation in soil was fast, it was not included in the model."*

### 3. Methods:

#### 1. Line 177: explain what a mixing topsoil layer is. Ploughed?

As also mentioned by the Reviewer 1, the mixing layer concept is probably the most commonly used in pesticide fate models due to its numerical simplicity and its realistic representation, as hortonian runoff processes have repeatedly been shown to be initiated as a near-surface process (Gao et al., 2004; McGrath et al., 2008; Wallender et al., 2008). This concept assumes that the transport is controlled by a mixing layer below the soil surface, in which rainwater, soil solution and runoff water completely and instantaneously mix (Havis et al., 1992). The soil depth interacting with runoff water is variable ranging from 0.25 to 2 cm according to various experimental and modelling studies (Havis et al., 1992; Wallender et al., 2008). Previous studies have shown that the effective depth of interaction is related to the degree of soil aggregation and increases with soil slope, kinetic energy of raindrops and rainfall intensity (Havis et al., 1992). In fact, the observed mixing-layer depth is often much shallower than the depth required for fitting models to field data (Gao et al., 2004).

The ploughed layer extends from 1 to 30 cm below the surface (second layer in PiBEACH). We have revised the sentence to provide more detailed information about the mixing layer directly in the article.

Gao, B., Walter, M.T., Steenhuis, T.S., Hogarth, W.L., Parlange, J.Y., 2004. Rainfall induced chemical transport from soil to runoff: theory and experiments. *Journal of Hydrology*. 295, 291-304.

Havis, R.N., Smith, R.E., Adrian D.D., 1992. Partitioning solute transport between infiltration and overland flow under rainfall. *Water Resour. Res.*, 28, 2569-2580.

Wallender, W.W., Joyce, B.A., Ginn, T.R., 2008. Modeling the Transport of Spray-Applied Pesticides from Fields with Vegetative Cover. *Transactions of the Asabe*. 51, 1963-1976.

Additionally, the reference (McGrath et al., 2008) has been cited at the first mention of the mixing layer to offer a more comprehensive understanding of this concept.

Line 215:

*Compared with BEACH features, PiBEACH includes (i) a mixing topsoil layer (assuming that the pesticide transport is controlled by a mixing layer from 0 to 1 cm below the soil surface, in which rainwater, soil solution and runoff water completely and instantaneously mix, McGrath et al., 2008) and additional deepest soil layers to include the shallow groundwater contribution to discharge,*

#### 2. Some textual errors on lines 168 and 177.

We revised the targeted sentences:

Daily vertical water fluxes across soil layers, followed by lateral fluxes, were considered at the cell scale from upstream to downstream. The surface flow direction was extracted from the digital elevation model, and flow-accumulation was determined using grid-based functions, without relying on a numerical scheme.

Compared with BEACH, PiBEACH includes (i) a mixing topsoil layer (assuming that the pesticide transport is controlled by a mixing layer from 0.25 to 2 cm below the soil surface, in which rainwater, soil solution and runoff water completely and instantaneously mix, McGrath et al., 2008) and additional deepest soil layers to account for the contribution of the shallow groundwater contribution to discharge, ...

3. Line 181: I think that not the various isotopologues are simulated in the model but the bulk heavy and light carbon in the pesticide?

As suggested by the reviewer, we have revised the sentence.

[...] and (v) pesticide carbon stable isotopic fractionation and transport of bulk heavy and light carbon in the pesticide (e.g.,  $^{13}\text{C}$  and  $^{12}\text{C}$ ) (Lutz et al., 2017; Alvarez-Zaldivar et al., 2018).

4. Line 187: "numerical diffusion"? Do they authors mean numerical dispersion?

We agree with the reviewer that 'numerical dispersion' refers to a type of truncation error that occurs when solving the diffusion-advection equation using finite differences, and have revised the relevant sentence accordingly.

5. Line 188: explain what the export coefficient is.

We provided the definition of the export coefficient:

This is because the numerical *dispersion diffusion* (Gatel et al., 2020) is generally of the same order of magnitude as pesticide export coefficient (i.e., the ratio of the mass transported at the outlet to the applied mass at the catchment scale), i.e., from 1‰ to 1% of the applied amount of pesticides.

6. Line 192-193: does this then not prohibit the simulation of hot-moments?

We assumed that the aspect targeted by this comment is the daily time step, which constrains its ability to predict fast flow dynamics, such as runoff generation and preferential water flow. We fully acknowledge that a daily model, using daily rainfall data, cannot capture intra-hourly rainfall intensity, which drives runoff intensity and surface off-site pesticide transport. However, this limitation, which was one of the key motivations for developing PiBEACH in conjunction with an event-based model like OpenLISEM OLP, does not, in our opinion, preclude the simulation of hot moments for degradation. Considering "soil moisture and temperature"- dependent degradation, we can account for periods of slower or faster degradation in the topsoil across the catchment, primarily driven by temperature. For example, during the first application in April, the  $DT_{50}$  was close to 100 days, with an average temperature of  $9.1 \pm 2.9^\circ\text{C}$  (Fig. 3C and Table S2). As topsoil temperatures progressively increased (Fig. 3A), the degradation rate accelerated, reaching a  $DT_{50}$  of approximately 25 days

in May, four times faster than in April. The impact of topsoil soil moisture on the  $DT_{50}$  was secondary (Fig. 3A and C). Between March and July, the topsoil temperature increased, displaying a more pronounced temperature gradient with depth, which resulted in reduced reactivity at greater depths. For example, in July, the  $DT_{50}$  values were 25 days in the topsoil (Fig. 3C) and 34 days at a 2-meter depth (data not shown).

In summary, the "soil moisture and temperature" dependence, as a proxy for microbial activity, provided a standardized approach for depicting degradation kinetics across the catchment and throughout the year.

**Model approach general:** I do not understand why this model has not been applied to simulate data at the catchment outlet as it seems to be designed to this end. Could not a simpler model been applied to simulate only the data in the topsoil?

We thank the reviewer for this comment. We agree that a lumped and parsimonious modelling approach, such as that developed by Lutz et al. (2017), effectively captures the dynamics of pesticide degradation and transport at the catchment scale. However, this approach was not designed to provide a distributed representation of pesticide concentrations in the topsoil throughout the growing season to assess off-site pesticide transport resulting from runoff and erosion. Such information is critical for informing management decisions and conducting risk assessments. PiBEACH was initially developed for this purpose as a companion model for distributed, event-based pesticide transport models, such as OpenLISEM OLP (Commelin et al., 2024), and cannot be used independently to simulate data at the catchment outlet. We agree that a lumped and parsimonious modelling approach such as developed by Lutz et al., (2017) was able to capture the dynamic of pesticide degradation and transport at the catchment scale. Therefore, this lumped approach was not designed to provide distributed extent of pesticide amount in topsoil when a rainfall event causes runoff and erosion with associated pesticide transport to support management decisions and risk assessments. PiBEACH was initially developed for this purpose as a companion model for distributed event-based pesticide transport models, such as OpenLISEM OLP (Commelin et al., 2024) and cannot be directly applied to simulate data at the catchment outlet. PiBEACH was not originally designed to simulate discharge and the associated pesticide concentrations in headwater catchments, where surface runoff and off-site transport can play a significant role in pesticide fluxes in rivers, such as in the Alteckendorf catchment with loamy soils (Lefrancq et al., 2018). Additionally, PiBEACH does not account for erosion and the transport of pesticides sorbed in the suspended matter, which is another limitation for simulating pesticide concentrations in rivers. The particulate fraction of some pesticides can be substantial at the outlet following intense rainfall events (Lefrancq et al., 2018).

As discussed previously, PiBEACH has a dual objective:

- To simulate both degradation and non-degradative processes across the catchment and throughout the crop growing season, providing key inputs for distributed, event-based models such as OpenLISEM OLP, which simulates the transport of both dissolved and particulate pesticides, based on the remaining pesticides in the mixing layer. This contributes significantly to pesticide modelling at the catchment scale.
- To model based-flow and the associated pesticide concentrations, enabling the integration of discharge and pesticide concentrations between rainfall events.

We thank the reviewer for highlighting the importance of considering the complexity or simplicity of a model in providing targeted answers. We have not verified in the section on model description in the Materials and Methods and this issue has been addressed: “the development of the PiBEACH conceptual model relied on field knowledge of hydrological dynamics and associated S-metolachlor flows in the Alteckendorf catchment, translating our experimentalist’s understanding (perceptual model) into conceptual model”, with an appropriate level of complexity to meet the two objectives.

To clarify the application domain of PiBEACH and prevent any misinterpretation of its capabilities and uses, we have now provided additional details in the title:

*Constraining topsoil pesticide degradation in a conceptual distributed catchment models with compound-specific isotope analysis (CSIA)*

In the Abstract (Line 28):

*Based on pesticide concentrations in topsoil distributed across the catchment and throughout the crop growing season, PiBEACH paves the way for coupling with a distributed event-based models, such as OpenLISEM OLP, which are better suited to address the intra-event dynamics of pesticide transport towards the catchment outlet.*

As detailed below, this information is also provided in the Materials and Methods section (see the new section on metric calculation, “2.7 Metrics of Pesticide Persistence and Transport Risks,” Line 449, page 13 of this document) and in the Results and Discussion section (Line 589, with further details on the last page of this document).

Lines 205-206: what is meant with the depth of the groundwater layers varying constantly from upstream to downstream?

As described, the Alteckendorf catchment was not equipped with piezometers or wells to monitor the dynamics of the water table. Therefore, we adopted a conceptual approach based on our understanding of soil structure (Lefrancq et al., 2017, 2018), capable of simulating the observed seasonal fluctuations in the water table when monitoring data is available (Molenat et al., 2005). A two-layer shallow aquifer (z3 and z4) was defined, with its thickness decreasing from the upper hillslopes (maximum of 23.2 m) to the riverbanks. The shallow aquifer is encased by unsaturated layers (Z0 to Z2, with a total thickness of 80 cm), which include a mixing layer, the plough layer, and a layer affected by artificial drainage, as determined by our soil structure investigation (Lefrancq et al., 2017, 2018). The shallow aquifer was divided into an upper variably saturated layer (z3), influenced by aquifer recharge, and a lower saturated layer that drains towards the river, functioning as a global linear reservoir (Manfreda et al., 2005). A calibration parameter (Zf) was used to distinguish the upper and lower portions of the shallow aquifer, accounting for the gradual decrease in thickness of both the saturated and variably saturated layers. Consequently, we revised the targeted sentence by replacing 'constantly' with 'gradually' to more accurately reflect our approach.

*The depth of groundwater layers (z3 + z4) varied gradually from upstream to downstream, exhibiting seasonal fluctuations driven by aquifer recharge, as observed in the hillslopes of monitored catchments (Molénat et al., 2005).*

With a new reference:

Molénat, J., Gascuel-Oudou, C., Davy, P., Durand P.: How to model shallow water-table depth variations: the case of the Kervidy-Naizin catchment, France, *Hydrol. Process.*, 19, 901-920, <https://doi.org/10.1002/hyp.5546>, 2005.

7. Page 9: as sorption was simulated in the model, why not test / show that sorption isotope fractionation has limited/negligible effects on the simulated isotope ratios in topsoil (and especially at the outlet?)?

As discussed previously in the point 2 raised by the Reviewer, S-metolachlor is not expected to exhibit strong sorption behavior, as strongly sorbing contaminants typically have  $\log K_{ow} > 2$  and water solubility  $< 1$  ppm, and S-metolachlor degradation in soil was fast. Molecular diffusion plays a negligible role in contaminant redistribution when S-metolachlor is first introduced into the field, potentially leading to non-equilibrium behavior due to concentration gradients between mobile and immobile phases. Furthermore, unlike highly sorptive contaminants in aquifers, S-metolachlor does not form distinct front and tail effects since it is applied uniformly to corn and sugar beet plots across the catchment and gradually diffuses toward the outlet. Although low but measurable isotope fractionation may occur following multiple sorption–desorption cycles (Kopinke et al., 2017), sorption-induced stable isotope fractionation is generally considered negligible for moderately sorptive pesticides like S-metolachlor in surface agro-ecosystems during seasonal transport. Thus, we did not simulate sorption-induced isotope fractionation in our model.

8. Page 10: add bit of explanation to argue why it was chosen for to simulate only the bioavailable fraction? How would it influence model outcomes compared to the more common model formulation that simulates biodegradation of the total dissolved fraction? Is this in line with earlier catchment degradation model studies?

As noted by the reviewer regarding common model formulations, we also accounted for the biodegradation of the total dissolved fraction. Additionally, we assumed that the reversibly sorbed fraction of pesticides, subject to daily sorption/desorption equilibrium, undergoes biodegradation at a similar rate, expressed by the  $DT_{50,reference}$  value. However, as in several pesticide transport models, e.g. PEARL, PELMO, PRZM or MACRO, we included an aged fraction that is more strongly sorbed and undergoes limited degradation.

In summary, consistent with the following models, we considered three pesticide pools: the bioavailable fraction, comprising both the (1) total dissolved fraction and (2) the reversibly sorbed fraction, and (3) a non-extractable fraction.

9. Line 265: “depth-dependent degradation”. The argument not to include this in the model seems rather weak (because there are no data available from deeper

soil layers; but there are in fact catchment outlet data available). If that reasoning is followed the authors should also only draw conclusions on what happens in the top soil layer and not in the catchment as a whole. Lutz et al. 2017 included “depth-dependent degradation” to enable fitting the concentration and isotope ratios of pesticides at the catchment outlet. But the current model is not constrained with such outlet data and therefore there will be high uncertainty on what happens at the catchment scale. In fact as the model is only calibrated with topsoil data (and discharge at the outlet but that is not that relevant for catchment scale degradation extent), the model is therefore only informative on what happens in the upper 1 cm of the catchment. See also main comment #2.

We fully agree with the Reviewer that the apparent depth-dependent degradation may be a parsimonious approach to represent complexity of depth-dependent microorganism activity as shown in different recent papers. We have chosen in this paper to explicitly link the persistence of pesticide with soil moisture and temperature as key drivers of microorganism activity integrating both temporal, spatial and depth dynamics such as in MACRO or PEARL models. We tested at the lab scale on the Alteckendorf soil, the predictability of this approach (Figure S4) before to implement it in PiBEACH. With this "soil moisture and temperature"-dependent degradation, we can account for periods of slower or faster degradation in the topsoil across the catchment. Figure 3C highlights cold and hot moments, i.e., slower and then faster degradation, primarily driven by temperature. The  $DT_{50}$  was close to 100 days during the first application in April at  $9.1 \pm 2.9^\circ\text{C}$ , before decreasing to 25 days in May. As topsoil temperatures progressively increased (Fig. 3A), the degradation rate stabilized in May, with a  $DT_{50}$  of approximately 25 days—four times faster than in April. The impact of topsoil soil moisture on the predicted  $DT_{50}$  was secondary (Fig. 3A and C). Between March and July, the topsoil temperature increased, displaying a more pronounced temperature gradient with depth, which resulted in reduced reactivity at greater depths. For example, in July, the  $DT_{50}$  values were 25 days in the topsoil (Fig. 3C) and 34 days at a 2-meter depth (data not shown). The adopted “soil moisture and temperature” dependence as a proxy for microorganism activity allow an uniformized method to depict degradation kinetic across the catchment and throughout the year.

This decrease resulted in a moderate, depth-dependent degradation of PiBEACH, in line with predictions from the calibration of the lumped model developed by Lutz et al. (2017) for this catchment. In both cases, i.e., the depth-dependent degradation (Lutz et al., 2017) and soil moisture and temperature-dependent degradation, the adopted approach can only tackle tendencies of microbial activity which remains largely unknown and complex to monitor along soil depth and at the catchment scale.

The next step, opened with the PiBEACH development, will be the combination with an event-based model such as OpenLISEM OLP which is expected to significantly improve the intra-daily runoff and associated dissolved and particulate transport from the plots to the outlet. This combined approach opens the way to open the grey box of soil reactivity by introducing new dependence and calibrate associated parameters.



We have now clarified the assumptions of soil moisture and temperature dependence in the Materials and Methods section and completed the discussion (line 316), as suggested by the Reviewer.

Line 316:

*Degradation rates generally decrease over soil depth, e.g., due to lower microbial activity (Cruz et al., 2008; Lutz et al., 2017) or sorption (Arias-Estevez et al., 2008) in deeper soil layer. The lack of concentration and isotope data for S-metolachlor in deeper soil layers prevented direct consideration of depth-dependent degradation.*

Line 353:

Between March and July, the topsoil temperature increased, displaying a more pronounced temperature gradient with depth, which resulted in reduced reactivity at greater depths. For example, in July, the DT<sub>50</sub> values were 25 days in the topsoil (Fig. 3C) and 34 days at a 2-meter depth (data not shown), as predicted by the lumped model developed by Lutz et al. (2017) for this catchment. In both cases, the depth-dependent degradation (Lutz et al., 2017) and the soil moisture and temperature-dependent degradation, the adopted framework can only capture general trends in microbial activity, which remains largely unknown and challenging to monitor across soil depth and at the catchment scale.

10. The three applications as shown in Fig. 3 are not described in the methods section (or I could not find it): how much pesticide was applied when and where? Exactly this heterogeneity in source zone variation calls for a model like this one but the advantage of using a distributed model has not been illustrated.

The application details were provided in the initial version, in the Supplement (Table S2).

Table S2. Applied mass (Kg) of active ingredient (S-metolachlor) per transect by date and days since 1st application. Ranges indicates uncertainty of exact application date (Alvarez-Zaldivar et al., 2018).

App. No.	Date	Days	North	Valley	South
A1	March 20 - 25 <sup>th</sup>	0 - 5	5.1	1.6	11.1
A2	April 13 - 14 <sup>th</sup>	25 - 26	8.0	1.8	2.9
A3	May 25 - 31 <sup>st</sup>	67 - 73	7.2	2.4	0.0
Total			20.2	5.9	14.0

We have now revised the sentence in the Materials and Methods to summarize Table S2, emphasizing the spatial and temporal variability of S-metolachlor applications at the plot scale. In our view, this variability highlights the necessity of a distributed modelling approach such as PiBEACH.

Line 161:

*The three application dates with, associated doses and formulation of S-metolachlor were collected for each plot through farmer surveys, providing spatially distributed data for the modelling. In summary, applications took place from late March to late May, with total amounts*

of 20.2 kg, 5.9 kg, and 14 kg applied to plots in the North, Valley, and South transects, respectively (Table S2).

## Results & Discussion:

11. Generally: explain why concentrations and isotope ratios of pesticide at the catchment outlet were not used. I assumed this was done when starting reading the paper but only in the R&D section it became clear that only data from the upper 1cm of the catchment skin were measured and simulated. How then can conclusions be drawn on overall degradation at the catchment body when these simulations are not constrained with outlet data? Thus relatedly, what is the relevance of the simulations of the upper 1 cm in the soil as presented in the key result Figure 3 on what happens at catchment scale? See also main comment #2.

We acknowledge Reviewer 2's comments regarding the relevance of the catchment results (Figure 5). As suggested by Reviewer 1, we have now reorganized the article's structure by adding a new section to the Materials and Methods, which introduces the six metrics simulated after calibration and derived from soil and river samples. Additionally, we have provided a detailed explanation of how the observed values, i.e., topsoil S-metolachlor concentration and  $\delta^{13}\text{C}$ , but also river discharge, and S-metolachlor concentration at the outlet, were used to derive the observed values for three of the six metrics.

The model was calibrated using topsoil concentration ( $\text{KGESM} > 0.5$ ) and isotopic signature ( $\text{KGE}\delta > 0.8$ ), along with the flow rate at the outlet ( $\text{KGEQ} > 0.5$ ), to represent hydrological functioning at the catchment scale. As mentioned in the Results section (lines 438–440), “*The tendency of PiBEACH to overestimate runoff generation is related to the basic assumptions of the SCS-CN approach (see section 6.2, SI), and to the daily time-step concealing daily distribution of rainfall intensity*”. This result (Figure 3B) confirms the limitations of PiBEACH in accurately representing intra-daily and intra-hourly runoff generation and the associated fast-flow pesticide transport.

The S-metolachlor concentrations measured at the outlet throughout the growing season were also used to derive the observed export metric (Figure 5C). As expected, given PiBEACH's limited ability to capture intra-daily runoff generation, it is unable to accurately simulate S-metolachlor fluxes at the catchment scale representing less than 2% of the applied amount. Therefore, we opted not to include outlet concentration and isotopic signature in the calibration function to avoid a biased calibration process, which could lead to parameter fitting for the wrong reasons in an attempt to compensate for the model's conceptual limitations.

The added value of incorporating observed S-metolachlor concentrations and isotopic signatures at the outlet in the calibration process will be explored in the upcoming PiBEACH-OpenLISEM OLP modelling study.

It is important to emphasize that the primary goal of PiBEACH, which has been now clarified in the manuscript, is to simulate the degradation and persistence of pesticides in topsoil layer

for coupling with event-based models such as OpenLISEM OLP. Figures 5A (isotopic dynamics) and 5B (remaining mass in topsoil) demonstrate that PiBEACH performs well in achieving this objective.

12. The three levels of sampling resolution as presented in Figure 4: composite, transects, plots are unclear, and not clearly described and explained in the method section. This would help also on improving on main comment #1.

Following the suggestion of both reviewers, we have now revised the caption of the Figure 1, and the following sentence, to explain the weekly weighted samples of topsoil for each transect from regular sampling (green dots) along transects (see Alvarez-Zaldivar et al., 2018).

*Figure 1: The headwater catchment of Alteckendorf (Bas-Rhin, France), illustrating the experimental setup, which includes transects (weighted samples collected at the green dots along the red transects) and plot sampling (black dots), as well as the land use for 2016. The "Other" category encompasses roads, grass strips, and orchards.*

Lines 146-148

*Soil samples were collected from 13 marked plots (before S-metolachlor applications and on days 1, 50, and 100 post-application), as well as from the north, valley, and south transects across the catchment (weekly weighted samples from the green dots in Fig. 1, spanning from March 19th, i.e., day 0, to July 12th, i.e., day 115). These samples were analyzed for the quantification and CSIA ( $\delta^{13}\text{C}$ ) of S-metolachlor in soil*

13. Lines 440-441: in my words: "more detailed spatial soil sampling did not help to better constraint the model". Why then was a spatially distributed model needed?

Mobilizing data of isotopic signatures through higher-resolution soil sampling (plots > transects > transect composite samples) did not significantly improve model calibration in this catchment. Therefore, pooling weekly transect samples into a single composite soil sample for CSIA appears sufficient to constrain the degradation parameter, i.e.  $DT_{50 \text{ reference}}$ , while also reducing sampling and analytical efforts. This composite approach may effectively capture temporal variations in pesticide degradation at the catchment scale.

We hypothesized that the transect composite approach, with a single averaged topsoil  $\delta^{13}\text{C}$  value per week, could also enhance the calibration of lumped models such as the one developed by Lutz et al. (2017). However, this result does not disqualify distributed modelling approaches such as PiBEACH, which was primarily designed to spatially represent pesticide concentrations, hydraulic properties, and topsoil moisture as initial conditions for distributed event-based models.

14. Lines 443-447: the difference in fractionation factor -2.7 vs. -1.5 is quite large and needs further explanation. Also compare and discuss the outcome with the fractionation factor calibrated by Lutz et al (2017): - 1.3 permil.

The incorporation of CSIA data into PiBEACH yielded an isotope fractionation factor of  $\epsilon_C = -2.7 \pm 0.6 \text{ ‰}$  for S-metolachlor degradation in topsoil. This value is notably larger than those reported in previous studies, including Alvarez-Zaldivar et al. (2018) ( $\epsilon_C = -1.5 \pm 0.5 \text{ ‰}$ ) and Lutz et al. (2017) ( $\epsilon_C = -1.3 \pm 0.5 \text{ ‰}$ ). While this discrepancy suggested differences in degradation dynamics or environmental conditions, it remained within the uncertainty range obtained from the model ensemble in this study.

The observed stronger fractionation effect ( $\epsilon_C = -2.7 \pm 0.6 \text{ ‰}$ ) may result from variability in environmental conditions, microbial activity, and transformation kinetics, supporting a more refined assessment of S-metolachlor degradation. The fractionation factor reported by Alvarez-Zaldivar et al. (2018) ( $\epsilon_C = -1.5 \text{ ‰}$ ) was obtained under laboratory conditions, where experimental constraints may have reduced isotope fractionation compared to field-scale observations. Similarly, the  $\epsilon_C = -1.3 \pm 0.5 \text{ ‰}$  reported by Lutz et al. (2017) was derived from a different dataset and modelling approach.

The present study incorporates a more extensive dataset and employs model ensembles, capturing a broader range of degradation scenarios and yielding a stronger fractionation effect. This suggests that enzymatic degradation may have been more pronounced, potentially due to higher microbial activity or the presence of a dominant microbial consortia with different metabolic capabilities.

While the previously reported  $\epsilon_C$  values ( $-1.5 \text{ ‰}$  and  $-1.3 \pm 0.5 \text{ ‰}$ ) may have slightly overestimated degradation extent, the discrepancy remains within the uncertainty bounds determined in this study. These findings emphasize the need for site-specific calibration when applying CSIA-based degradation assessments and highlight the importance of model ensemble approaches in accounting for uncertainties in isotope fractionation factors under field conditions.

The difference in isotope fractionation in previous and this study is now discussed in the manuscript (line 536):

*“The integration of CSIA data into PiBEACH yielded an isotope fractionation factor of  $\epsilon_C = -2.7 \pm 0.6 \text{ ‰}$  for S-metolachlor degradation in topsoil, exceeding previously reported values (Alvarez-Zaldivar et al., 2018:  $-1.5 \pm 0.5 \text{ ‰}$ ; Lutz et al., 2017:  $-1.3 \pm 0.5 \text{ ‰}$ ). This discrepancy likely reflects differences in environmental conditions, microbial activity, and transformation pathways, yet remains within the uncertainty range derived from the model ensemble. Compared to laboratory-derived estimates, the stronger fractionation observed in this study suggests alternative degradation pathways or the influence of distinct microbial consortia at the catchment scale. The broader dataset and model ensemble approach provide a more comprehensive assessment of degradation processes, capturing a wider range of field-scale conditions. While previously reported  $\epsilon_C$  values may have slightly overestimated degradation extent, these findings emphasize the need for site-specific calibration in CSIA-based assessments. Furthermore, they highlight the importance of model ensembles in addressing uncertainties associated with isotope fractionation factors in heterogeneous environmental settings, such as agro-ecosystems.”*

15. Line 458+459: “measure”, change into “measurements”

We have corrected the sentence as suggested.

16. Figure 5: what is the relevance if these results at catchment scale when the model is only calibrated on topsoil data but not on outlet pesticide data?

We acknowledge the Reviewer's comments regarding the relevance of the catchment results (Figure 5). As mentioned in the previous point n°12, this has now been clarified by specifying how the soil and river samples were used in a new section to the Materials and Methods. As discussed previously, it should be mentioned that the primary goal of PiBEACH, is to simulate the degradation and persistence of pesticides in topsoil layer for coupling with event-based models such as OpenLISEM OLP. This has now been clarified in the manuscript. Figures 5A (isotopic dynamics) and 5B (remaining mass in topsoil) demonstrate that PiBEACH performs well in achieving this objective.

4. Discussion is largely lacking. Add a larger discussion section to include for example:

We acknowledge the Reviewer's feedback and we have clarified the discussion by incorporating elements related to the three highlighted points and previously discussed.

1. Comparison with findings Lutz et al. (2017) who applied a similar model to the same catchment but then calibrated to earlier data at the catchment outlet instead of the topsoil. Add some discussion to discuss why this study did find that having isotope data let to better model results, whereas this effect was not clearly found by Lutz et al. (2017). Note that Lutz et al (2027) had to include depth-dependent degradation to enable to simulate the outlet pesticide and isotope ratio data.

We have now added the following section, discussing the main contributions of each modelling study (Lutz et al., 2017 and this one), in the results section (line 600), before the conclusion:

*In summary,  $\delta^{13}\text{C}$  data from both soil samples, used to constrain degradation, and the river samples, used to capture additional degradation processes along pesticide transport, can strengthen the robustness of the calibrated parameters. Both lumped (Lutz et al., 2017) and distributed approaches, such as PiBEACH, can be used to investigate pesticide dynamics throughout the growing season. However, only a distributed model can account for spatial heterogeneity across the catchment to support further management decisions. Additionally, PiBEACH provides spatially explicit data on pesticide concentrations, hydraulic properties, and topsoil moisture, which are crucial as initial conditions for distributed event-based models. We leveraged a substantial portion of the data collected at the catchment outlet for both calibration (discharge in the KGE objective function) and predictability assessment (discharge and S-metolachlor concentrations to estimate export). Due to the conceptual limitations of PiBEACH, particularly its daily time step, the comparison between simulated and observed S-metolachlor reactive transport was only partially conducted, in contrast to previous lumped modelling approaches (Lutz et al., 2017). The added value of incorporating observed S-*

*metolachlor concentrations and isotopic signatures at the outlet in the calibration process will be explored in the upcoming PiBEACH-OpenLISEM OLP modelling study.*

## 2. Pros/cons topsoil vs. catchment outlet data.

We have now added the following section, discussing the strengths and weaknesses of using soil and river  $\delta^{13}\text{C}$  data to constrain reactive transport models, in the results section (line 594), before the conclusion:

*This new modelling approach, applied to a similar compound and catchment as in Lutz et al. (2017), allows for a discussion of the strengths and weaknesses of using soil and river  $\delta^{13}\text{C}$  data to constrain reactive transport models. The lumped modelling approach (Lutz et al., 2017) focuses on  $\delta^{13}\text{C}$  of S-metolachlor in the river at the catchment outlet, representing an integrated signature of ongoing degradation at the catchment scale. This model was developed as a standalone approach to evaluate the utility of additional S-metolachlor  $\delta^{13}\text{C}$  data in river systems during calibration. In the PiBEACH framework, we further investigated the added value of the topsoil  $\delta^{13}\text{C}$  data to reduce uncertainty in the degradation kinetics. For three components of the S-metolachlor mass balance, namely, degraded mass, remaining mass in the topsoil, and export at the catchment outlet, their predicted cumulative contributions against observed data could be validated. We did not explore the combined use of both topsoil and river S-metolachlor  $\delta^{13}\text{C}$  in PiBEACH due to conceptual limitations in accurately accounting for rapid surface off-site transport associated with runoff and erosion. However, with the planned coupling of PiBEACH and an event-based model, such as OpenLISEM OLP, integrating both topsoil and river  $\delta^{13}\text{C}$  data may substantially improve the constraints on ongoing degradation at the catchment scale and enhance the representation of surface fast-flow transport during intense rainfall-runoff events.*

3. Line 493: I thought the conclusion was that there are no “hot-spots” in this catchment, as there is limited spatial variation among the topsoil samples? Therefore, add some discussion what the added benefit of a spatial distributed model was in this case.
5. Hot-spot and hot-moments. The paper does not discuss hot-spots in the catchment. The advantage of this model is that it enables to account for spatial heterogeneity but it seems that all model parameters were taken spatially homogeneously except with depth. Therefore, the added benefit to the model applied by Lutz et al. 2017 remains unexplored.

We have now added the following section discussing the points 3 and 5 as suggested by the reviewer 2

*The degradation process in PiBEACH, dependent on soil temperature and moisture, was primarily governed by temperature, exhibiting a distinct seasonal and vertical pattern. Consequently, the predicted degradation hotspot was located in the topsoil layer. Soil*

*moisture, which showed greater spatial variability than temperature, played a secondary role in degradation, with no significant spatial variability observed between plots treated with S-metolachlor throughout the growing season. The limited predicted spatial variation of  $DT_{50}$  across the catchment does not diminish the relevance of PiBEACH's distributed nature, particularly in the context of its coupling with distributed event-based models such as OpenLISEM OLP.*

6. Line 510-511: see earlier comments: as pesticide concentrations and isotope ratios were only measured on the topsoil one cannot simply extrapolate to the catchment scale. Degradation rates are likely much higher in the topsoil than in the deeper soil layers leading to overestimation of a catchment to degrade pesticide. The data to constrain the model also at catchment scale are available. Why not use them?

We have now addressed the second major critique by clarifying in the manuscript (new section to the Materials and Methods, introducing the six metrics and in the new element of discussion) how discharge and S-metolachlor concentration in the river at the catchment outlet were used to calibrate the parameters and assess the simulated S-metolachlor export from the catchment. As previously discussed, we opted not to include outlet concentration and isotopic signature, mainly collected during rainfall-runoff events (Alvarez-Zaldivar et al., 2018), in the calibration function to avoid a biased calibration process, which could lead to parameter fitting for the wrong reasons in an attempt to compensate for the model's conceptual limitations.

*We leveraged a substantial portion of the data collected at the catchment outlet for both calibration (discharge in the KGE objective function) and predictability assessment (discharge and S-metolachlor concentrations to estimate export). Due to the conceptual limitations of PiBEACH, particularly its daily time step, the comparison between simulated and observed S-metolachlor reactive transport was only partially conducted, in contrast to previous lumped modelling approaches (Lutz et al., 2017). The added value of incorporating observed S-metolachlor concentrations and isotopic signatures at the outlet in the calibration process will be explored in the upcoming PiBEACH-OpenLISEM OLP modelling study.*

*Sylvain Payraudeau, on behalf of the co-authors*

*Strasbourg, February 17<sup>th</sup> 2025*