Dear recommender and reviewers,

Many thanks for the detailed reviews regarding our paper titled "Chemical effects on ecological interactions within a model-experiment loop". The comments and suggestions were all largely relevant and helpful. We have addressed all of the queries for the two referees for whom you forwarded responses.
As suggested by the recommender, we included the entire ODD model description in the main text, so that it is complete by itself and does not require reading other papers, which improves the overall clarity of the paper. Otherwise, we now discuss more in details the choices we made when discarding data to test their relevance. We also elaborated on how transferable our suggestions on the experimental design are, especially for other species or chemicals. A table in the SI to summarize the different experiments was also added to help reading the experiment section.
You will find our detailed answers to each of the reviewers' comments below, in italic blue. Changes in the manuscript are highlighted in blue. We hope that you now find the paper suitable for publication. If you have any further queries please do not hesitate to get in contact with me.

Yours sincerely,

The authors.

Comments from Reviewer 1:

The manuscript reports an interesting study whereby different combinations of experimental data are used for parameterisation of a 3-species microcosm model in an effort to identify appropriate testing strategies for generating ecotoxicological data that avoids unnecessary testing. Furthermore, this particular experimental and modelling setup aims to tease apart direct and indirect effects of chemical exposure. The principle of using models and experiments in an integrative, iterative process for optimisation is not new, but this manuscript presents a nice illustration of how this approach could be used in small, multi-species microcosms. The authors finish by recommending a testing strategy for future microcosms studies. Below are a few comments:

Experimental setup: P6 L114 Experiment 1 is referred to as "Experiment without sediment". This suggests that all other experiments were performed with sediment, however, it is not clear whether this is the case when reading through the design of the other experiments in the manuscript. Clarification is required. This is especially pertinent to the experiments with cadmium, as the fate of cadmium ions will depend upon processes such as absorption, precipitation etc and result in both suspended and dissolved forms partitioned across a number of chemical species. Due to the low solubility of cadmium, much of it may precipitate and bind to the sediment which may alter the bioavailability to (and the effect on) the organisms. Whilst the authors state that analytical confirmation of cadmium was performed throughout the experiments, presenting mean measured values, depending on the presence of sediment (or not) may necessitate some additional discussion on the fate of cadmium in the system and any implications of its toxic effect and modelling outcomes (e.g. this may further inform the discussion on the NEC for Daphnia survival when compared with that in the wider literature as discussed on P17 L361-377).

We agree that including sediment makes a massive change in cadmium fate and bioavailibility in the microcosm. In order to simplify this aspect, we chose to include in the present study only experiments conducted without sediment, which is now specified l.109. The term "without sediment" corresponds to the denomination in Lamonica et al. 2016a, where some experiments were conducted with sediment (those ones were not included in the present study) and some without.

Table 2: Why are there no 2-species experiments including Daphnia and algae in the cadmium exposed studies?

We did not perform 2-species experiments with daphnids and algae because we showed in Lamonica et al 2026b that the competition between duckweeds and algae has not effect on algae dynamics (with and without cadmium), and we made the assumption that there is no interaction between duckweeds and daphnids. Hence, we supposed that the 2-species experiments with daphnids and

algae are equivalent to the 3-species experiments, so we decided not to perform the 2-species experiments to reduce experimental effort.

P15 L 311 states that algae-daphnid interaction (grazing) parameter wasn't well supported in the model fit. This is quite a critical parameter if exploring multi-species interactions and could the 2-species study have helped to address this? Inclusion of these 2-species experiments would also have provided more interesting results in the modelling using only partial experimental results (as the Daphnia results would also have changed i.e. the "super-imposed" look of the daphnia results in figure 4 between the partial datasets and reference data would potentially have differed).

As mentioned above, we supposed that the 2-species experiments with daphnids and algae are equivalent to the 3-species experiments, so we decided not to perform the 2-species experiments to reduce experimental effort.

Modelling assumptions:
P11 L218 "we make the assumption that the contaminant toxicokinetics is fast" Given the potential importance of this assumption on the modelling results and effects observed, can the authors provide further justification for the validity of this assumption.

We added references (Gestin et al., 2021, Ratier and Charles, 2022) l.299 to support this assumption.

P13 L269 The reference data were considered the "best possible estimates in the present case study in view of the model and all available data". The authors have been clear that what they are comparing is one estimate against another (better) estimate (the best possible estimate given the limitations of the data available). It would be interesting to know whether the additional 2-species tests mentioned above would have improved this reference estimate and if so, what implications that would have on the results and conclusions (i.e. is the experimental plan incomplete thereby potentially undermining the reference estimate).

That is true, nevertheless our answer above explains the fact that the experimental plan was incomplete.

Results and discussion:
I generally found these sections an appropriate reflection of the study outcomes. How to integrate indirect effects as seen in this study into regulatory chemical risk assessment globally may be a challenge ahead of us, as well as how to deal with other species and chemicals in a microcosm-modelling approach. However, models are likely to only become more important in the future for application

in risk assessment so studies such as this one which optimise testing in line with modelling will only become more essential. Given the results obtained in the study, the authors' recommendation for future testing strategies is fair, even if the aim of reduced testing may not always be realised in practice given the remaining difficulty of knowing in advance what data is most essential.

Thank you for the very encouraging comment.

Whilst generally well written, the meaning of a few sentences are unclear and should be addressed.

We carefully reviewed the manuscript.

Note. I have not evaluated the R script and model code, though these are available for readers as a pre-requisite for submission to PCI.

Comments from Reviewer 2:
I enjoyed reading this paper because it attempts to make a concrete link between experimental study design and the use of data in modelling. Additionally, the model itself, considering the population dynamics and interactions between 3 species under cadmium exposure is interesting.

Thank you for the very positive comment.

There are some general issues which remain unclear to me after reading the manuscript.

Why chose the 4 reduced datasets, what is the rationale for them. I can imagine that it is easier to maintain single species lab experiments than a microcosm. Hence, omitting experiments that require additional microcosms might be a practical (perhaps also financial) benefit. Likewise, leaving out some of the lower exposure concentrations could be a way to reduce animal testing without losing the signal of cadmium effects (which might be assumed to manifest itself at higher concentrations?). Adding these kinds of "rationales" of why the omitted data were chose would be helpful.

We now discuss the choice of the datasets in the last section of the Discussion (l.557-570), addressing the benefit of omitting some of the experiments.

It is not clear how the final recommendation came to be formulated L 440 – 451, about which datasets are best needed to inform modelling work. In fact, how transferable is this approach to inform study design to other compounds or species sets? For example, if an endocrine disrupting contaminant was tested which perhaps has a strong non-linear effect, including effects at low concentrations, would you still get to the same conclusions about omitting low concentration data? Discussing a bit more the context in which the results and recommendations are to be placed would be very helpful.

We have elaborated on the recommendation in the last paragraph of the Discussion (l.585-): we now mention the change in the selected species, and also in the type of contaminant.

The study seems well conducted and scientifically sound. Below I provide a number of comments and suggestions that are mainly aimed at improving the clarity of the paper. None of them, however, are major flaws.

Introduction
L. 48 – 52: This section is a bit vague, not clearly linked to the previous text, and misses some details. I suggest to extend this into a separate paragraph showing concrete examples (or references) of where models have been able to link

(extrapolate) between levels of biological organisation, and specifically (related to the text above) how models have been able to explicitly account for species interactions.

We developed this paragraph and added some references to illustrate extrapolation between levels of biological organisation thanks to models l.52-59.

L. 61 – 62: Why do more complex experimental designs resist formal optimisation?

We now mention that the increase in complexity is due to species interaction and indirect effects of the contaminant (l.69-70).

L. 67 – 70: This sounds strange: you first need to collect data, then model these data, and then you can improve how to collect the data in the first place (after you have already collected them). Perhaps this just needs a few words at the end of L.70 ... improve the experimental design for studies with microcosm experiments with similar species and compounds (or do you think it could also be useful to give guidance on experiments with more species or under compound mixtures?)

Thank you for this suggestion, we added this at the end of the sentence l.78-79.

L. 73: I find it a bit difficult to understand "direct" and "indirect" effects. One could argue that direct effects are the interactions of a pollutant with a specific target molecule. Please explain in a bit more detail, e.g. "direct effects of the contaminant on a species in isolation", and "indirect effects via contaminant effects on species interactions". After reading the discussion I understand it a bit better, it seems direct relates to effects on specific, modelled processes, and indirect relates to effects on state variables which then cascade to affect processes where these state variables are inputs. A clear definition at the start of the paper would be useful.

We added explanations on what we mean by direct and indirect effects l.82-83.

L. 71 – 76: it seems to me that there is a 3rd aim: to develop critical effect concentrations for key population regulating processes (i.e. EC50 in stress functions). In fact, on L 85-86 this is mentioned as an explicit step in the project (and a discussion is given L 362 – 389).

This is indeed right, this third aim is now mentioned l.84-85.

L. 73: "how to get back from modelling to experimentation" sounds a bit bulky. How about: "how model outcomes can inform experimental design"

Thank you for the suggestion, we changed the formulation l.85-86.

L. 83: It is not clear to me how "this" permits to identify direct and indirect effects. Do you mean that you used all data, including data where species occur in isolation and where they occur as a community of 3 species, to estimate model parameters, which then allows you to identify direct and indirect effects, respectively. (or did I understand wrongly, see also previous comment L 73).

We now explicitly state that using data where species occur in isolation and where they occur as a community of species allowed us to discriminate direct and indirect effects l.95-96.

L 86. I understand that you cannot say everything at once, but it would be useful to specify which processes in order to make the text less vague and easier to follow "different processes (growth of the 3 species, survival of Dapnia, and strength of interspecies interaction)"

We now name the general processes l.99-100.

Experiments and observed data L 109: does measuring the size of the daphnids affect their survival?

Based on the extensive experience of the authors in experimenting with daphnids, it does not.

L 112-147: The description is nice. However, I found it helpful to put this into a table to get a clearer overview. I easily get lost reading this. I would recommend adding such a table (if not in the main manuscript, then in the appendix). In fact, while making the overview table, I noticed some missing info regarding the duration and replication of some experiments (see separate review file)

Thank you for this suggestion, we have added the table in the supplementary information (SI Table S1). We also added the missing information in the text.

Dynamic modelling L. 162 and 167: Please double check that these are the correct references, I have the feeling that Lamonica et al 2016 a and b have been interchanged.

We carefully checked the references throughout the manuscript.

L. 162. When I check Lamonica et al 2016b (or a?) it says exactly the same: "the interaction is modelled with a Lotka-Volterra type 1 model". So, referring to the Lamonica et al 2016 paper does not assist me with further details. Perhaps you can just add a few words for clarity (e.g. something like this? "... Lotka-Volterra type I model both algal and duckweed growth rate are directly proportional to ?food?")

We now include the entire ODD protocol in the main text, thus the reader will find all model assumptions section 3.4 Design concepts, in particular sections 3.4.1 to 3.4.4, and equations for the complete model explicitly written Eq. (7).

L. 166 Since the interaction is an important part of this paper (i.e. the aim to disentangle direct vs. indirect effects), I would find it useful to have a bit more info (and equations) about the grazing and ingestion, rather than just a reference to another paper. (I checked the Lamonica et al 2016b (not a) paper and found a good description there, but it gets a bit much to check the supplements and two other papers to find the info needed to understand a relatively important part of the study).

The process description and the equations have been added l.307-319, Eqs. (3,4,5).

L. 178: Does this sentence contradict the sentence above (L. 175: "...algae and duckweeds are competing..."), as well as the use of the Lotka-Volterra type I model for competitive interaction?

We now specify that the competition is unilateral l.305.

L 181 – 193: It would be good to make explicit reference to the supplements (section 1.7.6 / eqn 6). Alternatively (but this might be a personal preference), it seems a shame that the actual equations are buried in the supplements, considering that this paper is developing a model.

We agree on that, and also following the recommender advice we have moved the complete set of equations in the main text Eq. (7)).

L. 194: "We use stochasticity..." This is quite vague, and could be done in different ways. Does this only apply to the binomial distribution for the number of daphnids? I can see in Fig S1 that there are some variances added but it is a bit hard to decipher exactly where. (This sentence about stochasticity is perhaps also a bit out of place here, as you continue to describe the deterministic model equations related to Cadmium stress in the following paragraphs. Perhaps the stochasticity deserves a section on its own in the paper (e.g. just before statistical inference)?)

Following the ODD protocol, how we include stochasticity in the model is now described in a specific section 3.4.5 (l.242-250).

L. 197-200: Are there any references to support these assumptions?

Actually, some studies suggest that carrying capacities might be affected by chemicals (Hendriks et al. ETC 2005). Although, from what we previously observed and modeled in this particular microcosm and experiment, it is fair to assume that only growth rates are affected by cadmium.

Statistical inference L. 255: Was only the Gelman Rubin diagnostic used, or did you also do a visual check. Was there a certain criterium or cut-off used to decide if the Gelman Rubin diagnostic was sufficient?

We now specify that we used a cut-off of 1.01 l.373.

Look-back on the experimental design Table 2: I did not immediately understand what the Cadmium concentrations C0 – C12 were referring to, but after making the table of the datasets (see comment L 112-147), I now see that there are indeed 4 groups of experimental concentrations. So, perhaps adding the table of datasets to the paper would help others as well.

As previously mentioned, we have added the table in the supplementary information (SI Table S1).

Results
L 302. Something is wrong with this sentence, the English does not make sense.

We corrected the sentence (l.419).

L. 331: Is this a typo, I counted 5 stress functions (5 plots per dataset)

We rephrased the sentence (l.449).

Figure 4: Why not also display the median prediction (or mode as you mention in L. 265)?

We find that the figure is already quite full, so we would rather leave it without the median prediction.

Discussion
L. 386: Wouldn't this be more due to the low value of bk rather than the narrowness of the posterior distribution?

This is indeed right, the sentence has been corrected (l.503-504).

L 393: It would be nice to have some quantification of how big the negative effect was, and what is considered "slight".

The overall quantification can be found looking at the slopes of the growth and survival rates according to cadmium concentrations and at the values of EC_{50} for those processes, thus we now refer to Figure 4 (l.510). We agree that the term "slight" is ambiguous, so we modified the sentence to highlight that the effect of cadmium on growth was lighter than on survival (l.509).

L. 430: similar stress functions were obtained with dataset B and the reference dataset, but looking at the appendix, some of the parameters do differ when using dataset B. Perhaps this needs a mention and some discussion on whether or not changes in individual parameters are relevant to make recommendations regarding the design of lab experiments.

We now mention in the discussion (l.548-549) that there were changes in the estimates of parameters, which need to be taken into account if the purpose is to estimate EC_{50} for a given process, for instance.

Conclusion
L. 464 – 475: I do not see a clear link between this paragraph and the paper.

We removed this paragraph, which was indeed unrelated to the rest of the text.

Figures
Reference to the figure numbers seem off in a number of places throughout the text. The labels in figure 4 are quite small.

We verified the references to the tables and figures. The small labels are mainly an issue of the overall figure size, it should be better at the final format.