Answer to reviewer 2

We have added our response below each of your comments (in bold).

The study highlights the role of nature-based solutions (NbS) in mitigating floods and agricultural droughts, particularly through soil-vegetation approaches that improve soil health and ecosystem resilience. The authors found that local soil characteristics significantly influence NbS effectiveness. Using hydrological models, it showed that well-drained soils enhance flood and drought resilience, as hedgerows improve infiltration and hydraulic properties. In contrast, waterlogged soils show limited benefits due to saturation. The study emphasizes the need to consider spatial soil variability when designing and placing NbS for maximum impact. In general, it is a well-elaborated and well-written manuscript, but I think that the value of the study, its pros, cons and system limits should be pointed out a bit better. I have listed my suggestions below. In times of increasing numbers of floods and droughts, such studies are very relevant in order to be able to derive resilient measures to mitigate the negative effects of these events. For me it is a valuable modelling experiment that cleary shows the potential but also the existing gaps in this field of research.

Abstract

Well written. I would already here mention what kind of model you used and what NbS you investigated, makes it more specfic. If you cannot mention all, give the most important ones / examples.

I think you could also highlight a bit more that is an experimental study (in my opinion), because you also investigate in how far the model is able to simulate the effect of the retention measures. I would write this. Point out what is innovative, what was surprising, and what are pros, cons and limits of your general and modelling approach. It should made be clearer what the value of your study is.

We agree with your remarks and will revise the abstract accordingly:

- We will specify the type of model used and the NbS investigated.
- Regarding the comment on the 'experimental study,' we will emphasize this aspect more in the abstract. Additionally, we will introduce a specific first objective: "Present a modelling approach that represents both the natural drainage characteristics of the soil and spatialized NbS scenarios, as well as their interactions." We will also review Sections 3.1 and 3.2 in light of this objective.

<u>See a proposed revised abstract :</u>

Nature-based solutions (NbS) are increasingly being explored as effective strategies for mitigating hydrological extremes, such as floods and agricultural droughts. Among these, soil-vegetation-based approaches may play a key role in improving soil health, enhancing ecosystem services, and restoring the natural hydrological cycle in productive agricultural and forestry catchments, making these landscapes more resilient to climate change. However, the influence of local factors, such as soil characteristics, on the effectiveness of these interventions is often overlooked. This is largely due to the fact that commonly used lumped empirical/conceptual modelling approaches often oversimplify the complex interactions across soil – water processes at multiple scales.

This study presents an innovative approach to modelling the effects of NbS landscape planning scenarios, explicitly simulating soil water fluxes. This approach allows for the investigation of how the spatial variability of soil properties influences NbS effectiveness in mitigating both floods and agricultural droughts.

A fully distributed, physically based hydrological model was used to represent NbS at the catchment scale, simulating water fluxes (e.g., infiltration, evapotranspiration, runoff, soil water content) on a variable-resolution grid. This model relies on measurable local parameters (e.g., topography, soil properties, vegetation) to capture small-scale hydrological processes, enabling NbS scenarios representation through spatial and temporal parameter adjustments. Hydrological simulations were conducted for two catchments: one agricultural and one forest dominated, each integrating a specific landscape planning scenario. In the agricultural catchment, the scenario involved the integration of hedgerows, reduced tillage practices and soil pitting for maize crops. In the forested catchment, NbS focused on forest diversification, practices aimed at limiting soil compaction in forest and restoration of peatlands. NbS effectiveness was assessed not only downstream but also where they are implemented within the catchment using key spatial indicators. Among them, an indicator evaluating the susceptibility of vegetation cover to water stress at a given location was newly developed.

The modelling approach was validated by accurately reproducing river discharge and saturated zone dynamics. It effectively captures soil natural drainage characteristics and provides a reasonable representation of NbS effectiveness, as indicated by consistency between simulated and literature values. A key finding of this study is that NbS effectiveness in enhancing flood and drought resilience strongly depends on soil natural drainage characteristics. In well drained soils, hedgerows significantly enhanced infiltration by improving soil hydraulic properties and creating additional air space in the soil's porosity through higher rates of evapotranspiration. In contrast, improving hydraulic properties in waterlogged soils had minimal impact on infiltration due to saturation, with anoxic conditions potentially limiting transpiration. Additionally, the study highlights that well-drained soils *offer co-benefits for resilience to agricultural droughts, as they are more likely to experience water deficits that NbS can mitigate. In contrast, such benefits are generally absent in waterlogged soils, which rarely face water scarcity.*

Results highlight that the evaluation of NbS effectiveness should recognize the spatial variability in their performance. This variability should inform the type and location of NbS to increase their overall effectiveness. The study underscores the need to move away from siloed evaluations of NbS and instead adopt a coherent, territory-based approach. Our study encourages decision-makers to envision more hydrologically resilient landscapes—an increasingly urgent necessity in our latitudes. This modelling approach provides a solid foundation for future research, with potential for refining NbS effectiveness assessments. Addressing its data requirements and enhancing uncertainty analysis will further strengthen its application.

1.Introduction

General comments: Well elaborated. For me, however, it is not clear whether the two study areas were already affected by floods and droughts, so perhaps this could be added here (also in the description of the study areas). I hope I did not miss such a statement in the manuscript.

Yes, the study areas have already been affected by floods and droughts. Initially, we did not plan to expand on the specific characteristics of our study areas in the introduction, as we wanted to emphasize that our study is not just a case study but that our approach could also be valuable for other regions and catchments. Therefore, we only mentioned how floods and droughts affected Central Europe in the last years.

However, we will clarify in the description of the study areas that our specific study area, the Vesdre Valley, was the most affected region in Belgium during the July 2021 floods, with 39 fatalities and material damages amounting to several hundred million euros. Some Belgian climate scientists have modeled that, given climate change, such event could occur again two or three times until 2050, meaning the area remains at risk (https://orbi.uliege.be/handle/2268/312886).

Regarding droughts, we would not say that the study area is more affected than other parts of Belgium. However, Belgium experienced unprecedented droughts in the summers of 2018, 2019, and 2020 (https://orbi.uliege.be/handle/2268/300105). It induced *Picea abies* mortality (https://link.springer.com/article/10.1007/s10661-024-12372-0) and loss of crop productivity (https://www.cra.wallonie.be/uploads/2018/08/S%C3%A9cheresse_2018_en_Wallonie_20180813-v2.pdf). Climatologists also predict an increase in drought frequency in the coming decades (especially if global warming exceeds 1.5°C), which could negatively impact the region, especially peatlands (due to a growing summer moisture deficit), *Picea abies* forestry, and agriculture.

Line 47 ff (NbS): I think that the wealth of different "perspectives" and opinions on NbS can be confusing, you can be a bit more specific regarding your measures (some of them seems to be Natural (Small) Water Retention Measures ;o)). Have a look at our paper that tries to bring a bit more order into all these concepts:https://www.mdpi.com/2071-1050/16/3/1308. I was thinking of it when reading the concerning sentences in the manuscript. Just have a look at our paper, if it helps, that's fine, if not, that's fine too. No obligation / need to cite, only if it helps.

Indeed, choosing the right overarching concept to categorize our measures was not straightforward. We had a look at your paper but the question of which term is best suited for our paper remains open, as we believe that each has its pros and cons.

On one hand, NbS might be too broad, as it includes solutions beyond those specifically aimed at water management, which is the main focus of our article. On the other hand, the term Natural (Small) Water Retention Measures (NSWRM) is more specific, as it encompasses only measures primarily intended for water management, even though they may also provide co-benefits such as biodiversity enhancement, climate resilience, etc.

As hydrologists, we might be tempted to classify all our measures as NSWRM, viewing water management as the primary challenge. However, from a forest manager's perspective for example, measures such as forest diversification may serve other primary objectives—forest resilience, biodiversity, aesthetic value, and economic benefits—that are more significant than hydrological regulation.

We will mention other common overarching concepts, including NSWRM, in the introduction and cite your work, which clearly defines these terms. This would give readers a reference for further explanation of these terminologies.

Line 76 ff (models): You might provide a brief concise overview on some of such models either here or in the model framework description (2.2) which also lead to your model selection.

Yes, we propose to overview other commonly used models (SWAT+, SCS, HEC-HMS/RAS, etc.) and cite key studies that have used them to assess NbS efficiency. For pros and cons of each model, we would refer directly to 10.1016/j.scitotenv.2021.147058, which provides a comprehensive synthesis.

For me, it is in general still a problem that the measures we simulate in the model show an effect almost immediately, which does not correspond to reality. Depending on the landscape conditions, there is a delay in these effects (such as retention). If you see this similar, you could discuss that somewhere (intro or discussion, maybe better there).

Yes we see it similar and, indeed, this aspect deserves to be considered.

The POST scenarios that we modeled were scenarios where measures were considered "mature". We did not model the transition between the BASELINE and POST scenario.

We will clarify this in the 2.2.2 Landscape planning scenario section, and discuss the implication of this hypothesis in the study limitations section.

Line 86 ff (objectives) For me, as mentioned, it is also an experimental study that tests the capabiities of your (modelling) procedure for your task. I mean you clearly show the pros and cons and gaps (what still has to be done).

Yes, we will add an objective on that as mentionned above.

2.Materials and Methods
2.1 Study area

General comment: See my comment before, add (maybe here) info if the study areas already faced floods and / or droughts..

Figure 1: The soil map is a bit hard to read.

Yes, we will separate the soil map from the figure 1, make it bigger and change the symbology to beter represent the stone content, which is currently hard to read.

Line 106: Sentence "Apart from peatlands,.." . I think you can delete this sentence, it is already written in line 100 ("Soils are mostly silty.")(except of the stone content) and does not provide additional information).

Right. In response to Reviewer 1's comment, we planned to rewrite a dedicated section in the Material and Methods on soil description. The soil map (discussed above) will be included in this section.

2.2 Modelling framework

General comment: See my comment in the intro - provide a brief concise description on other models with similar capabilities (or weaknesses) such as MIKE SHE / MIKE 1D.

2.2.1 Hydrological model

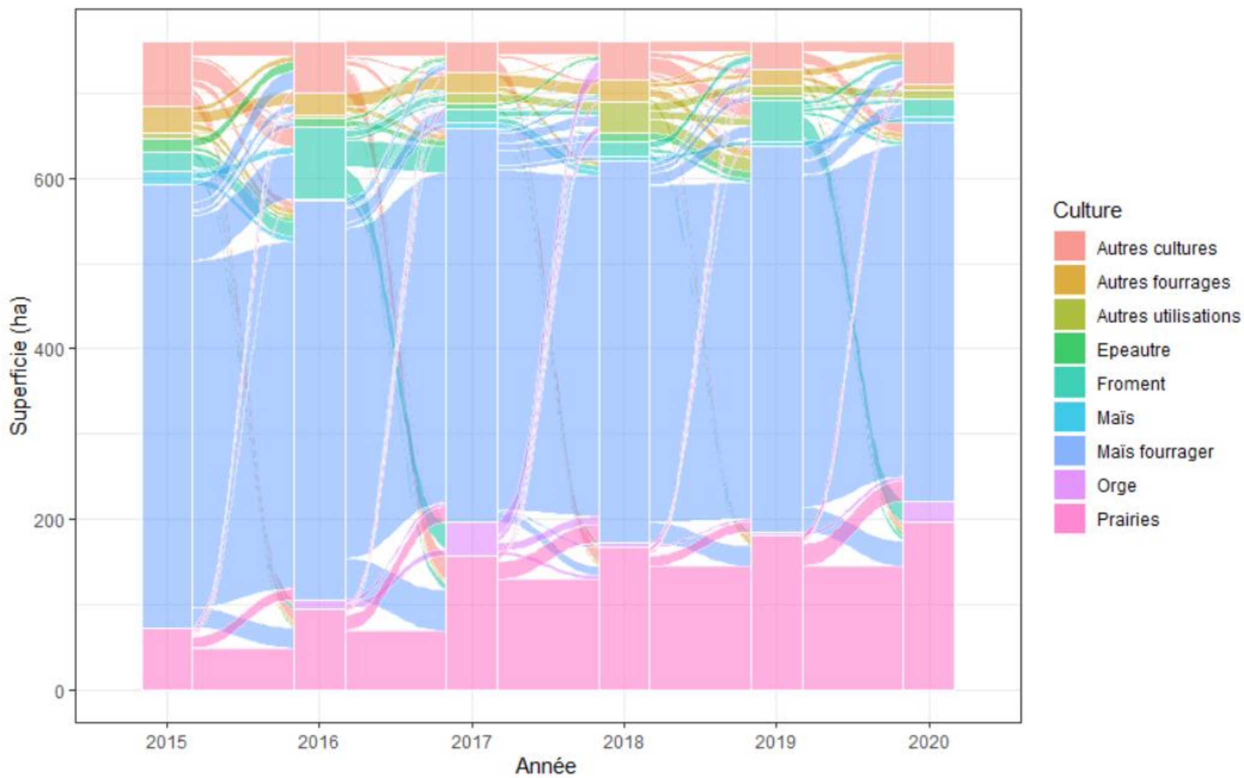
Line 124 ff (data description): I would add a table listing the data and describe the most relevant characterists / information it. See this example (section 2.3 Model inputs) https://www.sciencedirect.com/science/article/pii/S1470160X1931012X?via%3Dihub

Good idea. We will include the following table :

Data type	Data source	Resolution		Descripition
		Spatial	Temporal	
Rainfall	Gridded observational data	5 km	Daily	Gridded and point observations are combined to obtain hourly 5 km gridded rainfall
	https://hydrometrie.wallonie.be	Point observations	Hourly	
Reference evapotranspiration	Gridded observational data	5 km	Daily	Estimated with the Penman-Monteith equation (Allan et al., 1998)
River discharge	https://hydrometrie.wallonie.be	Point observations	Hourly	To calibrate and validate the models
DEM	LIDAXES (version 2) - MNT, 2023	2 m	2013-2014	It is a hydrologically conditioned DEM
LULC	WALOUS 2018 - Série, 2024; 140 Bassine et al., 2020	5 m	2018	This data served as model input to describe the spatial distribution of LAI, Kc, Rd and M value of Manning.
Soil	« Carte Numérique des Sols de Wallonie »	1/20.000	1947 - 1991	The soil map of Belgium was used as input to identify homogenious soil unit, stone content and soil depth. It is also used to validate the model with respect to the natural drainage classification.
	Textures et fractions granulométriques de référence des sols de Wallonie - Série	50 m 3 depths (0-40 cm, 40-80 cm, 80-120 cm)	1947 - 2020	Map of textural fractions: clay (0–2 µm), silt (2–50 µm), and sand (50–2000 µm). To each layer of each homogenious soil unit is assigned mean values of textural fractions.
	Updated European hydraulic pedotransfer functions (euptfv2)	/	/	Retention and hydraulic conductivity curves each layer of each homogenious soil unit were derived from the pedotransfer function 1 using depth and mean values of textural fractions as predictors
Geology	Sohier, 2011	1 km	/	This data originate from a Bibliographic study of the hydrogeological context of the Vesdre catchment. It gives the spatial repartition of homogenious hydrogeologic units and an initial guess of hydraulic conductivities, which was refined during the modle calibration.

Line 137 ff: What about land management (tillage crop rotation, fertilization, etc.) data, for instance from agricultural statistics, data from Integrated Administration and Control System (IACS) or interviews?

To our knowledge, systematic statistics on agricultural practices (tillage or fertilization, ...) are almost non-existent or not publicly available in our study areas. In Wallonia, we have a public database that is anonymized and updated annually based on farmers' declarations, providing information only on the main crop of the current year. In our study areas, apart from permanent grasslands, most main crops were maize, as shown in this figure illustrating crop rotation dynamics in the Vesdre catchment. (https://orbi.uliege.be/handle/2268/296474). Apologies, the figure is only available in French.



The LULC map we used incorporates data from this database (2018), which enables us to distinguish different types of crops. Therefore, we decided to simplify croplands by considering two categories: Open production surfaces and croplands (see table 1), assuming maize as a representative crop for parameterizing vegetation development (LAI, Kc and root depth) in crops.

Line 201 (Moriasi et al., 2007): I know Daniel Moriasi and I highly appreciate his work, and I know that many people use this performance guideline, but I think it is not always the best method to evaluate model performance. In some cases it might not say very much

about how well the hydrological dynamics are represented by the model (you describe the dynamics later in the text, so all fine (just a comment)).

We agree with your point. The Moriasi guidelines are widely used, and we included them as a reference to provide readers with a familiar benchmark. However, we acknowledge their limitations and discuss them in the text (line 534 - 540).

2.2.2 Landscape planning scenarios
General comment: Are these scenarios and suggestions just your ideas or are they based on some River basin or landscape management plans or something like that in your region that suggest some of them?
Would be good to know.

The measures we tested are based on a strategic territorial development framework called the *Schéma stratégique multidisciplinaire du bassin versant de la Vesdre* (Strategic Multidisciplinary Plan for the Vesdre River Basin). This study was conducted in response to the July 2021 floods and aims to rethink territorial development in the Vesdre basin in all its aspects, including hydrography, agriculture, forestry, natural areas, public spaces, mobility, and urbanization. The overarching goal is to make the region less vulnerable to future floods and droughts.

Rather than providing a concrete action plan, the study outlines a long-term vision for sustainable territorial development, proposing a range of potential measures. Our role was to assess the potential effectiveness of such measures proposed in this vision that may have an effect on the catchment's hydrology at mid-long term using hydrological modeling. To do so, we tested three scenarios, each implementing as many feasible measures as possible within a specific context (agriculture, forests, and peatlands), to estimate their maximum potential for flood and drought mitigation. A report has been produced :

https://orbi.uliege.be/handle/2268/314437

This was briefly mentioned in the acknowledgements, but we will cite the Strategic Multidisciplinary Plan for the Vesdre River Basin directly in the 2.2.2 Landscape planning scenarios section, and elabore about (like explained above).

Line 229 ff (Agricutural practices): These are partly very small areas (if I understand this correctly). Is it relevant?

Yes, these crop areas cover only 2.6% of the total catchment 1 surface. Consequently, their impact on the hydrograph at the catchment scale (or outlet) is probably limited due to their small extent.

We should, however, mention that the agricultural practice we implemented (soil pitting for maize crops) is an incremental rather than a transformational measure. Unlike no-till farming, it does not require farmers to change their agricultural system.

Our scenarios were aimed at implementing as many feasible measures as possible in our study site to demonstrate their potential combined maximal effecteveness while staying true to the catchment's existing conditions (LULC). We sought to demonstrate a range of options, which we also evaluate directly where they are implemented. This is one reason we have chosen a fully distributed hydrological model: it allows us to assess the effectiveness of measures not only at the outlet but also within the watershed, directly where they are implemented and effective.

Is it relevant? We believe that when planning flood risk reduction through hydrological measures, "*every square meter counts*".

Table 3: The table is titled "summary of hypotheses.." - are there other papers that used these or similar parameter modifications to "describe" such measures? How to confirm or reject the hypotheses? By measurements? These are these experiments? I guess there is no kind of a parameter database for such measures in the model (similar to SWAT) that can be extended?

Yes, these are indeed hypotheses, and we acknowledge that some values were chosen somewhat arbitrarily and sometimes even based on "expert judgment". MIKE SHE does not have a database of widely accepted parameter values for these measures, unlike SWAT model with the "Conservation Practice Modeling Guide for SWAT and APEX" from Waidler et al., 2011.

Some of the values we used can be found in the literature. For example, to modify the saturated hydraulic conductivity (Ksat) of the soil beneath hedgerows, we based our approach on the experimental results of Holden et al. (https://www.sciencedirect.com/science/article/pii/S0167880918304894. Other values are derived from measurements conducted in our laboratory, such as the assessment of roughness at the base of hedgerows using hydraulic resistance measurements (https://intelleau.wixsite.com/projet), and also found in literature (https://theses.fr/2018AMIE0031). In certain cases, parameterization follows a logical approach. For instance, when modifying the vegetation type, parameters such as LAI, crop coefficient (Kc), and root depth are adjusted based on pre-existing values in the BASELINE model (see Appendix 1).

Although reference values from previous studies could be used, we cannot assume they apply directly to our specific case. Claiming otherwise would overlook the uncertainty associated with these values. Therefore, we consider them as hypotheses, and our results depend on these assumptions.

We believe that a key advantage of our modelling framework is its reliance on theoretically measurable parameters that represent a physical reality. In contrast, more empirical and conceptual model-based approaches modify and calibrate "conceptual" parameters—such as the CN value (used in SWAT?) to represent a measure's effect on infiltration—based on expected impacts rather than direct measurements. This advantage could be enhanced through synergies with post-implementation monitoring, as parameter values could be refined using direct measurements after implementation, further improving the credibility and reliability of model outputs. Finally, a sensitivity analysis of final results concerning the choice of these values (covering a plausible range) would add significant value to our approach. However, it is story that, we believe, extend beyond the scope our paper. Nevertheless, this could lead to potential future research.

This limitation of our approach and the associated knowledge gap will be discussed further in the "3.7 Study limitations and gaps" section.

Is your model able to simulate the effect of each of these measures in equal quality? Or are some "better simulated" than others?

We believe that the model adequately represents the key hydrological processes affected by these measures, such as infiltration, evapotranspiration, and runoff. However, as with any model, it remains a simplification of reality, meaning that certain processes are not explicitly accounted for. For example, the formation of cracks during soil desiccation, which can lead to preferential flow pathways (line 508), the recycling of atmospheric moisture (line 574) are not included or the temporal variability of soil hydraulic propeties (line 542 : https://doi.org/10.1111/ejss.13558)

As for whether the model simulates the effect of each measure with equal accuracy, we are unfortunately unable to provide an answer at this stage. To do so, we would need direct field measurements of key hydrological variables—such as infiltration rates, evaporation, and soil moisture content—at locations where these measures have been implemented. Such data would allow for a direct comparison with model outputs and provide concrete validation of the model's ability to represent these measures accurately.

2.2.3 Integrated hydrological analysis of model outputs

Line 248 (Hydrological indicators): So, you developed these indicators? I think there several indicators out like this, I think you should provide an overview. Infiltration and

agricultural drought, etc., are not new "indicators". I would better point out what is new here (because - again - you write that you developed them) and what is better compared to the exiting ones. Do you see it as as a kind of "enhancement" to other hydrological indicators, such as the "Indicators for Hydrological Alteration (IHA)", introduced mainly by Richter et al. and Poff et al. and many more?

Yes, the names we assigned to the indicators we calculated or developed may not have been well chosen. We will rename them. We also will clarify which indicators were simply calculated (already existing) and which were newly developed.

The indicators we referred to as “Outlet hydrograph” and “Infiltration” can be found in the literature. For the first indicator (“Outlet hydrograph”), we will rename it as “peak flow percent reduction” and “flood volume percent reduction”, which are commonly found in literature. For the second indicator (“Infiltration”), a more precise term would be “spatial variation of infiltration rates.” While this type of indicator is not new, to our knowledge, it is rarely used to assess the effectiveness of NBS from a spatially explicit perspective at the subplot scale on slopes. This aspect, we believe, adds some originality to our study.

The last indicator, which we called “Agricultural drought,” is more novel (we developed it), as we have not seen it in other studies. It simultaneously considers soil water status (characterized by soil matric potential at different depths) and the vegetation’s ability to extract water (characterized by rooting depth). This indicator reflects the susceptibility of the vegetation cover to water stress at a given location, characterized by duration as a function of stress intensity (defined by different matric potential thresholds) and return periods. A more appropiate name for this indicator would be “spatial Vegetation Drought Stress Duration Frequency (sVDSDF)”.

Our study aims not only to assess hydrological impacts at the river scale (on river flow regime) but also to consider the entire watershed, including slopes and other relevant processes such as infiltration or susceptibility of the vegetation cover to water stress. We selected and developed these indicators to effectively represent the impact of NbS measures (even if we could be more exaustive) while ensuring they are easily interpretable by decision-makers, who are often non-specialists. This aspect, we believe, differentiate our approach from the "Indicators for Hydrological Alteration (IHA)," which were primarily designed to assess hydrological impacts on aquatic, wetland and riparian ecosystems with an analisys of the river flow regimes, with less emphasis on catchments slopes/drainages areas.

3.Results and discussion

General comment: Well elaborated and written.

3.4 Spatial variability of effectiveness of NbS against flood

Line 426 ff (These findings..). The question is again) if your results - considering all uncertainties - justify such a statement (second part of the sentence)?

Maybe we should adopt more cautious terms regarding this sentence. We plan to reformulate the sentence with :

“These findings suggest that well-drained soils may offer greater potential for improving infiltration, implying that, natural soil drainage characteristics should be considered when prioritizing areas of NbS implementation."

What about the delay of measures effects?

See our response above.

Regarding flood protection measures - when are floods are too strong, that measures have no effect anymore?

This paragraph concerns results of effectiveness of measures in term amount of infiltrated water focussing on the rainfall event between 13 and 18 July 2021 which was the most severe ever recorded in the region, with an estimated return period of 300–400 years. These results are therefore relevant to extreme flood events. But yes, with climate change such an event, or even more severe ones, occurring again cannot be ruled out.

Line 468 ff ("These areas, we believe.." ff): See my comments above. Are the results reliable enough to state this - would you tell that a stakeholder / water manager, etc.? OK, you eaken your statement with last sentence, but think of reformulate it.

Yes. We propose to rephrase by :

“Further research would be valuable in determining whether these areas should be prioritized for implementing NbS aimed at retaining overland runoff and enhancing reinfiltration.”

3.5 Spatial variability of effectiveness of NbS against agricultural drought

General comment: You could perform allocation change experiments" (searching for the most effective measures (or measure combinations) at specific locations)? We use multi-criteria optimization for that (see https://www.optain.eu; see also project deliverables)). Could be a point for discussion?

3.6 Synergies and trade-offs between flood and drought mitigation with NbS

General comment: See my previous comment regarding "allocation experiments" / combination of measures.

That is something we had not considered, mainly because we relied on the "*Schéma stratégique multidisciplinaire du bassin versant de la Vesdre"* (Strategic Multidisciplinary Plan for the Vesdre River Basin) which served as a reference for our scenarios. But yes, we had a look at your "D5.1: Common optimisation protocol" from the OPTAIN project and your paper on CoMOLA (https://doi.org/10.1016/j.envsoft.2019.05.003). We found your approach very interesting and it deserves to be discussed in our article.

If we understand correctly, a major limitation to applying this approach to our particular modelling framework is the number of simulations/scenarios required to explore the solution space and converge towards a set of optimal solutions. This would require a significant simplification of our models in order to reduce computation times. We believe that SWAT+ would be a better tool to perform such multi-criteria optimization procedure. We propose, though, to extend our last paragraph in section 3.6, mentioning your work:

"Finally, it is important to note that the balance between synergies and trade-offs depends not only on the type of NbS but also on their spatial implementation. As discussed in previous sections, NbS were more effective in soils without waterlogging, both in mitigating flood risks and addressing agricultural droughts. This opens the door to allocation change experiments and multi-objective optimization.

Since different NbS or combinations of NbS can create both benefits and trade-offs across multiple (often conflicting) objectives, it is essential to identify one or several optimal implementation strategies that maximize benefits while minimizing trade-offs. Among other, one approach is based on Pareto-optimality (Deb, 2014), where a set of optimal solutions is generated such that no objective can be improved without worsening another. From this set of optimal alternatives, decision-makers can explore and the select appropriate solutions based on their priorities.

Among the available tools for such analyses, the Constrained Multi-Objective Optimization of Land Use Allocation (CoMOLA) software (Strauch et al., 2019) has been developed to facilitate these assessments and has been applied, notably, within the OPTAIN project (www.optain.eu)."

3.7 Study limitations and knowledge gaps

General comments: In general, I would emphasize the value of the study a little better.

Indeed, in this section, we are—as we believe rightly so—as critical as possible of our approach. We aim to be transparent by acknowledging that our approach is not perfect and that many questions remain unanswered. However, we believe, this does not diminish the value of our study in any way.

From our perspective, the value of our study lies in its process-based approach, making full use of the available data. We demonstrate that our modeling approach effectively captures the fine-scale interactions between soil and water processes across multiple spatial scales. This enables us to assess the effectiveness of measures in detail, considering the interactions occurring at different scales.

The measures we consider are "nice to have" solutions, offering multifunctionallity and being easier to implement than hard engineering measures. Rather than focusing solely on the riverbed, they target catchment slopes, aiming to boost vertical water fluxes such as infiltration, recharge, and evapotranspiration. At the same time, we evaluate their effects on both floods and droughts.

This underscores the need to move away from siloed evaluations of NbS, which are still common among land managers whose governance structures remain highly compartmentalized. Instead, a coherent, territory approach is essential. Our study encourages decision-makers to envision more hydrologically resilient landscapes—an increasingly urgent necessity in our latitudes.

We will rework on this section with this, to emphazise a bit more the value of the study. We will also addapt our conclusion and abstract accordingly.

So far, the gaps are highlighted quite well in this chapter (although the word "however" is used a little too often ;o)), but it should be worked out a little better. What about the uncertainties (capability of the model(s) (what are the system limits (scale, describing the efficincy measure under the given circumstances, data, etc.? Can they (uncetainties) be quantified? What does that do to the reliability of the results? Is it sufficient so far to be able to derive reliable recommendations?

Yes we will rework this section following your recommandations.

<u>Model uncertainties</u>: We used all available data, including all gauging stations in the study area and drainage classes from the Belgian soil map, to validate our model results as thoroughly as possible. However, this was likely not sufficient to ensure that we accurately captured the hydrological functioning across the entire study area. We acknowledge this limitation and discuss it in the first paragraph of Section 3.7: Study Limitations and Knowledge Gaps. Additionally, we conducted tests on the uncertainty introduced by pedotransfer functions (using uncertainty estimates from EU-PTF V2) by simulating hydrographs with an alternative model in a smaller catchment. Such analyses, we believe,

extend well beyond the scope of this article, and we plan to address them separately in a future study.

Model capability : As previously discussed, the model remains a simplification of reality, omitting certain subtle processes—such as the temporal variability of soil hydrodynamic properties. However, we believe that it adequately represents key hydrological processes influenced by measures, including infiltration, evapotranspiration, and runoff.

Can uncertainties be quantified ? : Partially, through our model validation procedure. Regarding the representation of measures, our model results align with expectations based on existing literature on NbS effectiveness. However, as noted above, quantifying uncertainties would require direct field measurements of key hydrological variables—such as infiltration rates, evaporation, and soil moisture content—at locations where these measures have been implemented. A sensitivity analysis of measure parameterization could provide valuable insights but would require a dedicated, comprehensive study.

What does that do to the reliability of the results? Is it sufficient so far to be able to derive reliable recommendations? :

As you suggested in lines 426 and 468, we will review the manuscript to take a more cautious approach when drawing recommendations from model outputs with unverifiable uncertainty. Nevertheless, since we used a process-based modeling approach, even if the exact efficiency values remain uncertain, we can still identify which processes are affected by which measures. This allows us to formulate logically sound recommendations. Our key message in this article remains that local soil drainage characteristics should be considered when implementing NbS as it influence their effectiveness.

For me, this is a very good modeling experiment that could be used as a basis for further discussion (scientists, stakeholders), also to specify what needs to be tackled urgently next. To do this, you could involve farmers, authorities, economists, and first ask them for their opinion on the methods, measures and results.

Yes, we actually used this study to convice to stakeholders from the Vesdre Valley that these measures could be effective. This study enabled us to secure a new project aimed at creating “laboratory” catchments of around 2 km^2 in the Vesdre Valley, where such measures would be implemented in practice. The first phase involves iteratively developing a coordinated management plan with various stakeholders in the catchment—authorities, farmers, businesses, and local residents. As hydrologists, our role is to propose an “optimized hydrological management plan”, including a quantified assessment of its effectiveness. This plan would then be reviewed by stakeholders to ensure its feasibility and acceptability. Iteratively, we would model the revised plan to assess its hydrological effectiveness.

Ultimately, we hope that such a plan will be implemented, giving us the opportunity to monitor changes in the catchment's hydrological functioning, evaluate the effectiveness of the measures in practice, the changes on local hydrological parameters and addapt the models (parameter values, etc) as function of.

In this section, we will emphasize that while such a (our) study, although essential to demonstrate the potential effectiveness of these measures in order to convince, is not sufficient to guarantee their concrete implementation. To move forward, a collaborative territorial approach with all stakeholders is needed, one that also integrates human sciences.

I have already mentioned other points (taking into account the delayed effect of the implemented measures, experiments on the spatial allocation of measures (and the combination of measures), flood protection measures – when are floods too strong to be mitigated by these measures? etc.). You basically already have many of the points in this chapter, but I would discuss it a bit more clearly and in a more structured way and – as I said – emphasize the value of the study more clearly.

See our responses above.