Response to Reviewer comments

We are grateful to the reviewers for their thoughtful and constructive feedback, and for the time taken to engage so carefully with the manuscript. We have implemented a number of changes in response to these suggestions, which we feel have significantly improved the paper.

Reviewer 1:

Dawkins et al have presented a portfolio level assessment of expected damages from rainfall induced floods for non-residential built area across Europe. Given the dearth of broadly applicable spatial data and lack of international hydrological modeling, **this is a well written and commendable effort to provide a data driven approach to estimating portfolio risk, especially for insurance and pension plan providers.**

Based on my review, I noted a few major issues like the mapping from rainfall to flood hazard, inaccuracies such as calculation of AEP, and unclear figures. I also noted a few minor corrections. Hopefully the authors and the editor would find these comments insightful and beneficial for improving the manuscript for future readers.

1. One of the primary limitations of this study is the mapping step from rainfall hazard to flood hazard. No validation has been provided for the approach, and this step severely limits the applicability of this study. While the authors have pointed out in the limitations section that their analysis is illustrative, it would be good to emphasize this limitation both in the Limitations section and throughout the manuscript.

Response:
We thank the reviewer for this important comment. We fully agree that the rainfall–flood mapping step is a key limitation of our framework and that its implications for applicability should be more clearly and consistently communicated.

In response, we have revised the manuscript to explicitly emphasise (i) the simplified and unvalidated nature of this mapping, (ii) the resulting structural uncertainty, and (iii) the intended illustrative rather than decision-grade use of the results. Specifically, we have added and revised text in multiple sections as detailed below:

We have added the following sentences towards the end of the abstract:

“A key methodological limitation is the simplified statistical mapping between rainfall extremes and flood depth, which replaces a physically based hydrological modelling chain and is not explicitly validated against observations or hydrological simulations. As such, the results should be interpreted as indicative rather than decision-ready estimates of flood risk.”

In the final paragraph of the introduction, we have revised the description of the framework and added the following clarification:

“While are method is transparent and pragmatic, a key component of this framework is a simplified statistical mapping between rainfall extremes and flood depth, which substitutes for a physically based hydrological modelling chain. This introduces substantial structural uncertainty, and therefore the resulting risk estimates are intended for illustrative and awareness-raising purposes rather than for decision-making at the individual asset level.”

In the second paragraph of the ‘Estimating flood depth hazard information’ subsection, we have added:

“An important caveat of this approach is that the relationship between rainfall and flood depth is represented using a statistical mapping rather than a physically based hydrological model. This mapping is not explicitly validated against observed flood events or hydrological simulations and therefore may not fully capture key processes such as runoff generation, river routing, floodplain storage, and local flood defences. As such, the methodology is suited for demonstration purposes and raising awareness.”

We have strengthened the description of the mapping step in the ‘Relating rainfall and flood depth return-levels’ section by replacing the relevant sentence with:

“The relationships derived here represent a simplified statistical approximation of a highly complex physical system. They do not explicitly account for key hydrological processes and are not validated against independent flood observations or model output.”

We have strengthened the first limitation by replacing it with the following (note this is also added to in response to subsequent reviewer comments):

“Simplified rainfall–flood mapping and lack of hydrological modelling: Limited availability of consistent, open hydrological modelling data at continental scale necessitates the use of a simplified statistical mapping between rainfall extremes and flood depth. In some locations, flood models now exist that can provide flood risk information at very high resolution over limited regional domains (e.g. for the UK, at up to 20-25 m resolution) however, important details of methods and data used for flood risk estimates may not be in the public domain (Bates et al., 2023). Our simplified statistical mapping approach replaces a physically based modelling chain, and therefore cannot represent key processes such as runoff generation, antecedent soil moisture, flow routing, floodplain dynamics, and the influence of flood defences. Critically, our simplified rainfall–flood relationship is not explicitly validated against observed flood events or independent hydrological model output. As a result, there is substantial structural uncertainty in the resulting flood depth estimates, particularly in their absolute magnitude. Consequently, the risk estimates presented in this study should not be interpreted as location-specific or decision-grade assessments, but rather as illustrative estimates intended to demonstrate methodology and explore relative changes and sensitivities.”

We have added the following clarifying sentence to the ‘Societal implications’ section to reinforce interpretation:

“In particular, the simplified and unvalidated mapping between rainfall and flood hazard means that the outputs should be interpreted as indicative rather than precise estimates of asset-level risk.”

We have added the following sentence in the concluding section:

“A key limitation of the approach is the use of a simplified and unvalidated mapping between rainfall extremes and flood hazard, constraining the realism and applicability of the results for decision-making.”

We believe these revisions address the reviewer’s concern by ensuring that the limitations of the rainfall-to-flood mapping are clearly communicated, both as a central caveat and as a defining characteristic of the study’s intended use.

2. Line 324: I am not convinced that there would be 1:1 mapping between the X-year return period rainfall and X-year return period flood. Surely, multiple rainfall events could cause the

same flooding, and in its most accurate form, each flood hazard curve would be a weighted integration of the rainfall's entire spatio-temporal probability distribution. Can the authors provide justification for this simplification?

As the reviewer notes, flood frequency curves are formally derived from integrating the full spatio-temporal rainfall distribution with catchment processes. In this setting, rainfall events with differing temporal and spatial characteristics may produce similar flood magnitudes, and conversely, similar rainfall inputs can lead to different flood responses.

Our approach does not seek to represent this full process chain. Instead, the rainfall–flood relationship in our framework is intended to be viewed as a quantile-to-quantile mapping rather than a direct event-based correspondence. In other words, we relate rainfall and flood distributions through their return levels, without implying that individual rainfall events map directly to specific flood events.

This simplification is adopted in light of the aims and constraints of the study, and is motivated by three main considerations:

1. Consistency with design-event approaches: The assumption is consistent with widely used engineering and risk assessment approaches, where rainfall events of a given return period are used to estimate flood magnitudes of the same nominal return period (e.g. design storm methods). While simplified, this provides a transparent and interpretable link between hazard variables.
2. Tractability and use of open data: A more physically complete representation would require hydrological and hydrodynamic modelling driven by spatially and temporally resolved rainfall fields, ideally calibrated at the catchment scale. However, such datasets and models are not consistently available as open-source products across continental domains. A key aim of this study is to demonstrate a fully open and reproducible framework; the simplified mapping enables this while still providing first-order insights.
3. Approximate monotonic relationship between extremes: Although not exact, there is a physically motivated expectation that higher rainfall quantiles are associated with higher flood quantiles at a given location. The spline-based mapping is therefore intended as a smooth, monotonic approximation of this relationship, rather than a mechanistic representation of flood generation.

We fully agree that this simplification omits important processes and introduces structural uncertainty. As such, the resulting flood hazard estimates should be interpreted as approximate and indicative.

We have added the following paragraph to the Hazard methodology (rainfall–flood mapping section) when introducing the spline-based mapping:

"It is important to emphasise that this approach does not assume a direct event-based 1:1 correspondence between rainfall and flood events. Rather, it represents an idealised quantile-to-quantile mapping between rainfall and flood return-level distributions. In practice, related assumptions are commonly made in engineering 'design event' approaches, where rainfall events of a given return period are used to estimate flood magnitudes of the same nominal return period (e.g. Kourtis et al., 2020). However, in reality, flood frequency curves arise from the integration of the full spatio-temporal rainfall distribution with catchment processes, and multiple rainfall events with differing characteristics can produce similar flood magnitudes. As such, the relationships derived here represent a simplified statistical approximation of a highly complex physical system. They do not explicitly account for key hydrological processes and are not validated against independent flood observations or model output. Consequently, this approach should be interpreted as a monotonic, first-order approximation of the relationship between rainfall and flood extremes, intended to enable large-scale analysis in the absence of consistent open hydrological datasets, rather than a physically complete or decision-ready representation of flood hazard."

We have added the following clarification within the rainfall–flood mapping limitation:

"In particular, the mapping between rainfall and flood hazard is implemented as a quantile-to-quantile relationship. This does not reflect the full convolution of rainfall variability with catchment processes; whereby multiple rainfall events can lead to similar flood responses. As such, the method provides only a first-order approximation to the underlying flood hazard distribution."

3. Line 366: From figure 6, it appears that Gothenburg and Edinburgh were both used for fitting the splines, so it creates a circular logic to both fit the data to the CEMS data points and to state that the fitted curves align well with the CEMS data points.

We thank the reviewer for this helpful comment and agree that the original wording could be interpreted as implying validation of the spline fitting against the same data used to construct it. This was not our intention. The purpose of Figure 7 is to illustrate the interpolation approach and how discrete flood depth return-level points (from the CEMS maps) are extended into a continuous curve and applied consistently to both the historical and future periods.

In response, we have revised the text to clearly state that the agreement between the fitted curves and the CEMS data points is by construction, and to clarify how the fitted relationships are subsequently used to estimate future flood depth return-level curves. The revised paragraph now reads:

"Figure 7 (b) and (d) show the fitted continuous flood depth return-level curves for both the historical and future periods alongside the original CEMS flood map data points used in the spline fitting. By construction, the historical flood depth curves pass through these data points, and therefore the close alignment reflects the successful spline interpolation approach. The same fitted relationships are then applied to the future rainfall return-levels, allowing the estimation of corresponding future flood depth return-level curves."

4. Line 231: Why does only scale vary spatially, and not shape?

We thank the reviewer for raising this important point. We note that this modelling choice and its justification were already discussed in the Supplementary Material. However, we agree that this is a key aspect of the methodology and should be made clear in the main text.

In response, we have moved this explanation from the Supplementary Material into the main manuscript (Section 2.2.3), where we now explicitly describe why the shape parameter is held constant while the scale parameter varies spatially, emphasising considerations of robustness, parameter identifiability, and avoidance of overfitting:

"Holding the shape parameter constant in space is common practice in spatial modelling, because it has high sensitivity to small fluctuations in the data. Granting it too much flexibility can result in unstable and unreliable estimates, ultimately compromising the robustness of the model. This assumption is typically most suitable in applications where the spatial domain is moderate in size and where systematic, large-scale variation in higher-order distributional properties (such as skewness or tail behaviour) is not expected. While there may be genuine spatial variation in the rainfall tail behaviour within our study region (e.g., due to orography), evidence from spatial extremes modelling shows that adding unnecessary spatial flexibility to higher-order parameters risks overfitting and poor tail dependence characterisation; robust practice prioritises parsimonious marginal models over highly parameterised, spatially varying shapes (Davison et al., 2012).

Consistent with this, our model checking does not provide evidence that allowing the shape parameter to vary spatially would materially improve the fit."

5. Line 232: Is the spatial variation of scale from this method parametric or non-parametric? If parametric, it will be helpful for the readers to include its functional form. If non-parametric, can you specify the spatial grid spacing, and how the 12 km dataset is mapped on to this spatial grid?

We thank the reviewer for this helpful question and agree that further clarification of the modelling approach is warranted.

The spatial variation of the scale parameter is modelled using a semi-parametric approach**,** specifically a Generalised Additive Model (GAM) as implemented in the R package *mgcv,* which is integrated into the evgam package stated. In this framework, the dependence of the scale parameter on spatial coordinates (longitude and latitude) and mean rainfall is represented using penalised regression splines, rather than a fully parametric functional form. This allows flexible, smooth spatial variation to be captured while avoiding overfitting using penalisation.

The rainfall data are defined on the native 12 km × 12 km rotated latitude–longitude grid of the UKCP18 regional climate model, and the model is fitted directly on this grid without any reprojection or spatial interpolation. Each grid cell is treated as a single observation location with associated covariates.

The spatial coordinates used in the GAM (grid longitude and latitude) are standardised prior to modelling by centring (subtracting the median) and scaling by their range. This transformation improves numerical stability and comparability of the spline terms in the GAM, while preserving the relative spatial structure of the data. Importantly, this is not a projection or geometric transformation, but a simple rescaling for statistical modelling purposes.

We have added to the text here:

"The spatial variation of the scale parameter is modelled using a GAM framework. In this approach, the dependence on spatial location (longitude and latitude) and mean rainfall is represented using penalised regression splines, forming a semi-parametric model that allows smooth spatial variation without imposing a rigid functional form. The model is fitted directly to rainfall data on the native 12 km x 12 km rotated latitude–longitude grid of the UKCP18 regional climate model, with each grid cell treated as an individual observation location. The spatial coordinates are standardised (centred and scaled) prior to modelling to improve numerical stability, while preserving the relative spatial structure of the data."

6. Line 249: It does not appear that there are any temporal parameters in the fitted distribution. Is the assumption that the distribution parameters are constant over the entirety of each time period? Is that a fair assumption, especially given the authors' statement in the Introduction about 25% of flood losses observed within the most recent 3 years of the last 44?

We thank the reviewer for this comment. Indeed, in the current formulation the parameters of the fitted distribution for each season are assumed to be constant within each time period (i.e. stationary within each ~30-year time slice).

We have adopted this approach for several reasons.

1. This is consistent with the underlying flood hazard modelling framework used to derive the flood depth return level maps (Baugh et al., 2026), which does not include time-varying parameters. Maintaining this consistency avoids introducing methodological discrepancies between the hazard estimates which are mapped to one another in the following step of the method.
2. The use of temporally aggregated “time slices” as approximately stationary representations of a given period or global warming level is a common and well-established practice in climate science, particularly for impact and risk assessments where the focus is on differences between periods rather than continuous temporal evolution (Herger et al., 2015, Schleussner et al., 2016, Kennedy-Asser et al., 2021; Garry et al., 2021; Dawkins et al., 2023).
3. When we explored the inclusion of temporal variation by fitting a model with year included as a smooth term, this did not reveal a clear monotonic trend. Instead, the fitted temporal effect showed oscillatory behaviour, consistent with previous findings that rainfall variability is often dominated by internal climate variability rather than a strong long-term trend (e.g. see page 384 of Wood, 2017). As such, including a time-varying parameter did not materially improve model performance or interpretation in this case.

In response to the reviewer’s comment, we have now clarified this assumption explicitly in the manuscript at several points:

In the Methods (Rainfall Extreme Value Spatial Model section), we have added: *“Distribution parameters are assumed to be stationary within each defined time period (historical/future), primarily to ensure consistency with the underlying flood hazard modelling framework (Baugh et al., 2026), which does not include time-varying parameters. Further, this assumption is consistent with common practice in climate impact assessments where periods are treated as quasi-stationary representations of a given climate state (Herger et al., 2015, Schleussner et al., 2016, Kennedy-Asser et al., 2021; Garry et al., 2021; Dawkins et al., 2023a), and this assumption reflects the absence of a clear monotonic temporal trend in the rainfall distributions analysed here, consistent with previous findings (e.g. see page 384 of Wood, 2017).”*

In the Limitations, we have added: *“Assumption of stationarity within time periods: The assumption of stationarity within each time period (historical and future) in the rainfall EVA model may mask potential within-period non-stationarity, particularly given the concentration of recent losses. However, exploratory analysis including a temporal covariate did not indicate a clear monotonic trend, instead suggesting variability dominated by internal climate variability. Future work could investigate covariate-based approaches to better separate externally forced trends from internal variability.”*

Baugh, Calum; Colonese, Juan; D'Angelo, Claudia; Dottori, Francesco; Neal, Jeffrey; Prudhomme, Christel; Salamon, Peter (2026): River flood hazard maps for Europe and the Mediterranean Basin region. European Commission, Joint Research Centre [Dataset] doi: 10.2905/JRC.WPE5YRR; 10.2905/1D128B6C-A4EE-4858-9E34-6210707F3C81 PID: http://data.europa.eu/89h/1d128b6c-a4ee-4858-9e34-6210707f3c81

Herger, N., Sanderson, B.M. and Knutti, R. (2015) ‘Improved pattern scaling approaches for use in climate impact studies’, *Geophysical Research Letters*, 42(9), pp. 3486–3494. Available at: https://doi.org/10.1002/2015GL063569

Schleussner, C.F., Lissner, T.K., Fischer, E.M., Wohland, J., Perrette, M., Golly, A., Rogelj, J., Childers, K., Schewe, J., Frieler, K., Mengel, M., Hare, W. and Schaeffer, M. (2016) ‘Differential climate impacts for policy-relevant limits to global warming: The case of 1.5 °C and 2 °C’, *Earth System Dynamics*, 7(2), pp. 327–351. Available at: https://doi.org/10.5194/esd-7-327-2016

Garry, F.K., Bernie, D.J., Davie, J.C.S. and Pope, E.C.D. (2021) ‘Future climate risk to UK agriculture from compound events’, *Climate Risk Management*, 32, p. 100282. Available at: https://doi.org/10.1016/j.crm.2021.100282

Kennedy-Asser, A.T., Andrews, O., Mitchell, D.M. and Warren, R.F. (2021) ‘Evaluating heat extremes in the UK Climate Projections (UKCP18)’, *Environmental Research Letters*, 16, p. 014039. Available at: https://doi.org/10.1088/1748-9326/abd4fd

Wood, S.N. (2017) *Generalized Additive Models: An Introduction with R*. 2nd edn. Boca Raton: CRC Press. Available at: https://doi.org/10.1201/9781315370279

7. Paragraph 291: Can the authors include some comments on the QQplots at higher quantiles? The variation is typically largest at the highest quantiles, which is exactly the region that is being intended to be modeled by the EV-GAM. Did the authors notice a systematic under- or over-estimation in this region? What is causing this misfit, and how would adding covariates help reduce this misfit?

Thank you for this comment regarding the behaviour of the Q–Q plots at high quantiles, which are indeed central to the purpose of the extreme value modelling.

As shown in Figure 4 and the equivalent plots in the Supplementary Material, the empirical quantiles lie within the 95% confidence intervals of the fitted model across all locations, seasons, and ensemble members, including at the highest quantiles. We therefore do not observe a systematic over- or under-estimation of the upper tail, outside of the expected levels of uncertainty.

To clarify this interpretation, we have revised the manuscript to more explicitly describe this behaviour. The relevant paragraph now states:

“In some cases in Figure 4, however, such is in the Hamburg grid cell in the summer, many of the quantile points sit close to one end of the 95% confidence interval. While this could potentially indicate systematic over- or under-estimation in some cases, it more plausibly reflects the increased uncertainty associated with representing rare extremes from limited samples. This behaviour is well-known in extreme value analysis and does not indicate systematic bias in the model.”

Regarding the potential role of additional covariates, we agree that incorporating further physically relevant predictors could help explain residual spatial variability and improve the representation of extremes at specific locations. In particular, such covariates may reduce localised misfits where the current model structure does not fully capture spatial heterogeneity in rainfall processes. However, we note that the current model already achieves an adequate statistical fit for the purposes of this study, and the inclusion of additional covariates is identified as a potential area for future refinement rather than a necessary requirement for the present analysis.

8. Paragraph 291: Did the authors explore other probability distributions to model the rainfall data? Since the UKCP18 data includes forecasts, what distributions does that data use, and how does it differ from the authors’ choice?

Thank you for this helpful question. We selected the Generalised Pareto Distribution (GPD) within a peak-over-threshold framework because it is strongly supported by extreme value theory. In particular, for sufficiently high thresholds, exceedances are theoretically expected to follow a GPD

(Pickands–Balkema–de Haan theorem). This makes it a standard and widely adopted choice for modelling rainfall extremes. Alternative approaches within extreme value analysis (e.g. block maxima with a Generalised Extreme Value distribution) rely on the same underlying theory, but the peak-over-threshold GPD approach is often preferred as it makes more efficient use of the data by retaining all threshold exceedances rather than only block maxima.

We have added to the 2nd paragraph of the 'Rainfall Extreme Value Spatial Model' section:

"This GPD distribution choice is consistent with extreme value theory, which shows that exceedances above sufficiently high thresholds converge to a GPD (Pickands–Balkema–de Haan theorem)."

With regard to the UKCP18 data, we believe there may be a misunderstanding in the question. The UKCP18 projections are not generated from an assumed statistical distribution of rainfall. Instead, they are produced from ensembles of physically based climate model simulations that explicitly represent atmospheric processes. The resulting rainfall data therefore reflect the dynamics of the climate model and internal variability within the climate system, rather than a prescribed parametric probability distribution.

9. Line 306: The number of simulated data points seem to be too low for modeling more than ~10 years return period. For example, 500 years would have >180k days, and assuming 10 data points for achieving statistical significance, 1.8M data points would be required. Is there something I am missing from the description of the simulation approach, e.g., that only extreme rainfall events are simulated in which case how is their probability accounted for in the hazard curve generation?

We thank the reviewer for this thoughtful comment. The simulation approach focuses on the upper tail of the rainfall distribution as this is where the model is fitted, specifically exceedances above the 98th percentile of wet days at each location. This corresponds to approximately the top ~2% of wet days, or around ~1.3% of all days (given that approximately 60% of days are wet in many locations, the other 30% are dry).

For a 500-year return period (approximately 180,000 days, given the 360-day climate model years), the relevant portion represented within the peak-over-threshold framework is therefore ~1.3% of this total, corresponding to ~2,340 extreme events/days. Assuming ~10 samples are required to robustly characterise a return level, this implies that ~23,400 simulated exceedances would be sufficient. In this context, we argue that our simulation of 40,000 days (combined across seasons) provides an adequate sample size to estimate high return-period rainfall.

Section S1.3 in the Supplementary Material describes how we calculate the adjusted quantiles to account for this. We have pointed this out more clearly in the manuscript where we reference this part of the Supplementary Material:

"Importantly, appropriate adjustments are applied to these quantiles to account for the peak-over-threshold framework and the fact that only the upper portion of the rainfall distribution is simulated; these adjustments are described in detail in Section S1.3 of the Supplementary Material."

10. Line 303: Given the above comment, it is not completely clear to me why the authors needed to first fit the UKCP18 data to EV-GAM, and then simulate daily rainfall, instead of using the daily rainfall data directly? Over the 30 year time periods, there are already ~11k

data points. Are the hazard values obtained from the simulated approach significantly different than if the authors had used the data directly?

We thank the reviewer for this question. While the dataset contains ~11,000 daily observations, the effective record length remains ~30 years, which limits empirical estimation of return periods to 30 years. Return periods are defined in terms of years rather than data points, meaning the raw data do not provide reliable information about rarer, high-impact events beyond this range (e.g. 100-year events). Fitting the extreme value model is therefore necessary to characterise the tail behaviour of the rainfall distribution and enable extrapolation beyond the observed record.

We have added to the 1st paragraph of the 'Rainfall Extreme Value Spatial Model' section:

"Due to the limited (~30-year) record length in each time slice, empirical estimation of rare high return-period events beyond the 1-in-30-year event is not reliable, necessitating the use of a statistical model."

11. Line 461: Since the vulnerability function is a CDF, can the authors comment on their choice to use a spline fit instead of fitting a probability distribution, like lognormal?

Our interpretation of the vulnerability function (flood depth–damage curve) is as a mapping from water depth to fraction of damage, rather than a cumulative distribution function (CDF), which maps a variable to a probability. For example 0.6 on the vulnerability function means 60% expected fraction of damage, where as 0.6 on a CDF means 60% probability of being below something. Accordingly, we do not treat the function as a probability distribution to be fitted (e.g. lognormal), but instead as deterministic point-based information. The spline is used purely as a flexible interpolation method to provide a smooth and continuous representation between the tabulated depth–damage values reported in the EU dataset.

12. Line 471: It would help the readers better understand the methodology if the authors included the integration equation to illustrate how the hazard curve is combined with the vulnerability function? Additionally, closed form integration is generally not possible so can the authors also include their discretization step size?

We thank the reviewer for this helpful suggestion. We have revised the manuscript to (i) explicitly include the mathematical definition of the Expected Annual Damage (EAD), (ii) clarify how hazard and vulnerability are combined, and (iii) include additional detail on the discretisation used in the numerical integration. While the trapezoidal integration approach was already described later in the section (Step 3), we agree that introducing this information earlier and with greater clarity improves the transparency of the methodology.

We have edited the beginning of Section 2.5 to:

"For any given location/grid cell that is at risk of flooding, hazard, exposure, and vulnerability information can be combined to estimate flood risk. In this study, flood risk is represented by the Expected Annual Damage (EAD), which is defined as the expectation of flood damage over the full range of event probabilities:

$$EAD = \int D(h)\, dp(h), \quad (1)$$

$$D(h) = V(h) \times E, \quad (2)$$

where h is the flood depth intensity, $D(h)$ denotes the damage cost associated with a flood of depth h, and $p(h)$ is the corresponding annual exceedance probability. The damage cost is expressed as the product of the vulnerability function $V(h)$, which maps flood depth intensity h to fractional damage (see Figure 8 d), and the local financial exposure E (here taken as the total value of assets;

see Figure X(c)). The vulnerability function is applied deterministically to the hazard intensity, such that flood depths from the hazard return-level curve are translated directly into fractional damage (for example, a flood depth of 1 m corresponds to approximately 30% damage at that location).

As this integral for EAD cannot generally be evaluated in closed form, it is approximated numerically. Here, this is achieved using the composite trapezoidal rule, implemented using the Python scipy.integrate.trapezoid function, evaluated at discrete annual exceedance probabilities corresponding to return periods of 1 to 500 years at integer increments."

We also now refer to these equations and the terms in them in the method steps later in this section.

13. Figure 9 and line 484: While it is a reasonably good assumption for quick calculations, the authors have incorrectly stated that AEP is the inverse of the return period (otherwise for a 6-month return period, the annual exceedance probability will be 2, which is impossible). Please include the correct conversion based on the Poisson distribution to convert return periods to AEP.

We thank the reviewer for highlighting this point. We have revised the manuscript to clarify that the relationship AEP=1/T is an approximation that is valid for rare events, and we now include the correct conversion based on a Poisson assumption.

We have added to the beginning of Section 2.5 (just after the text shown for the comment above):

"The annual exceedance probability (AEP), denoted p(h) in Equation (1), is related to the return period T under a Poisson assumption as:

$$AEP = 1 - \exp(-1 / T) \qquad (3)$$

where T is the return period in years. For rare events (i.e. large T), this expression approaches the commonly used approximation $AEP \approx 1 / T$*. We note that the approximation* $AEP \approx 1/T$ *would yield a maximum AEP of 1 at* $T = 1$*; however, the Poisson-based formulation provides a more physically consistent representation of exceedance probability for frequent events. In this study, the lowest return period considered is the 1-year event, corresponding to a maximum AEP of approximately* $1 - exp\ (-1) \approx 0.63$ *under the Poisson assumption."*

We have then modified any reference to this relationship by pointing to this equation. We have also changed Figure 9 (c) and (f) to reflect this change and updated all subsequent figures with the updated EAD values calculated as the area under this corrected AEP curves.

14. Paragraph 496: It will be helpful if the authors can present actual loss damages from any recent flooding events at one of the presented cities, especially if that event's return period is quantified. This would help the authors validate their EAD estimates, and help readers anchor this study's EAD calculations, otherwise it is difficult to verify their order of magnitude.

Thank you for this suggestion. We explain in the manuscript that the EAD estimates presented here are calculated at the scale of individual 100 m × 100 m grid cells and are primarily intended as illustrative examples to demonstrate how risk varies spatially across different locations. As such, they are not directly comparable to reported damages from real-world flood events, which typically affect much larger areas of a city and aggregate losses across many exposed assets.

Furthermore, the EAD estimates reflect long-term average risk derived from hazard information over a 30-year climate period within a climate model, rather than the impact of a single realised event.

Observed damages from real-world floods are strongly influenced by local factors such as detailed hydrodynamics, defence infrastructure, and temporary adaptation measures, which are not explicitly represented in our framework.

We are therefore concerned that a direct comparison with reported event losses could be misleading and may suggest a level of validation or precision that is beyond the scope of this simplified, large-scale methodology. Instead, this study is intended to provide a consistent and scalable framework for comparing relative risk across spatially distributed portfolios, rather than reproducing site-specific loss estimates. We have added to the description of Figure 10:

"These EAD estimates are calculated at the grid-cell scale and are intended to illustrate relative spatial variability in risk; they are not directly comparable to reported losses from individual real-world flood events, which typically reflect aggregated impacts over much larger spatial extents and depend on additional local factors."

15. Line 508: Why are there ensemble members for the historical time period? Wouldn't the historical data be static and known, and ensembles only be present for future forecasted scenarios?

Thank you for this question. While observational data for the historical period are indeed fixed, climate model projection simulations are typically run as ensembles even for the historical period. Each ensemble member represents a physically plausible realisation of the climate system under the same external forcings but with slightly different initial conditions, and they are not tied to the observed internal variability over the historical period.

This is important because the observed record represents only one realisation of internal climate variability, whereas the ensemble captures the range of variability that could have occurred but did not happen by chance. Using ensembles therefore provides a more complete characterisation of extremes and their uncertainty than relying on the single observed trajectory.

In addition, running ensembles for both historical and future periods enables a consistent, like-for-like comparison. Any systematic model biases are present in both periods, so analysing relative changes (rather than absolute values) helps mitigate their impact and provides a more robust assessment of projected changes in risk.

16. Line 537: It would be interesting to see the variability in estimates for simulated portfolios with the same flood risk sampling, e.g., 1000 portfolios with 20% flood-risk regions, and 1000 portfolios with 80% flood-risk regions. This would test the hypothesis whether the percentage of built area in flood risk regions is by itself a good proxy for EAD.

Thank you for this comment, and for the related feedback on Figure 12. We have revised both the figure and the accompanying text accordingly. In particular, we now explicitly address the proposed hypothesis and present an improved approach to decomposing uncertainty arising from portfolio composition and climate ensemble variability. The text and figure have changed to:

"A deeper analysis of how varying the proportion of at-risk locations influences overall risk is presented in Figure 12, which shows how historical and future EAD varies across synthetic portfolios as a function of the proportion of assets at risk. The boxplots summarise variability arising from both portfolio composition and climate ensemble members, with EAD averaged over the alternate dimension in each case.

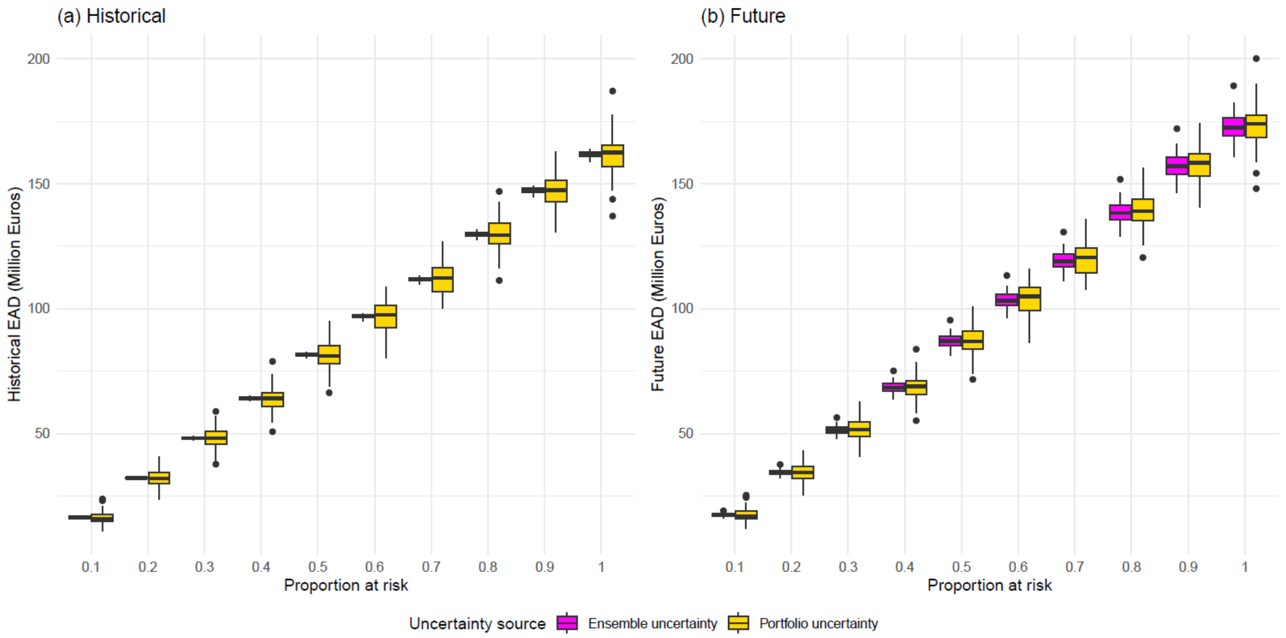


Figure 12: Comparison of uncertainty in estimated flood risk (Expected Annual Damage, EAD) arising from portfolio composition and climate model ensemble members. Boxplots show EAD as a function of the proportion of portfolio locations at risk of flooding (0.1-1 in increments of 0.1). For each proportion at risk, the distribution in yellow represents variability across 100 synthetic portfolios (portfolio uncertainty), where EAD is first averaged over the 12 UKCP18 RCM ensemble members for each portfolio. The distribution in magenta represents variability across the 12 ensemble members (ensemble uncertainty), where EAD is first averaged over the 100 portfolios for each ensemble member. This is shown for (a) Historical climate EAD and (b) Future climate EAD. In both panels, EAD is calculated using the same underlying portfolios and ensemble members, enabling a direct comparison of the relative contributions of portfolio and climate uncertainty.

Figure 12 shows that EAD increases systematically with the proportion of at-risk locations in both the historical and future periods, highlighting the strong dependence of risk on portfolio exposure. In particular, portfolios with low (e.g., 20%) and high (e.g., 80%) proportions of at-risk assets exhibit clearly separated EAD distributions, indicating that the proportion at risk provides a useful first-order proxy for EAD. However, there remains a substantial spread in EAD values within each exposure level, such that portfolios with the same proportion of at-risk assets can still experience markedly different levels of damage. This demonstrates that, while the proportion at risk captures a large component of flood risk, it is not sufficient on its own to fully explain differences in EAD, and additional factors such as asset value distribution and spatial configuration also play an important role.

Both panels in Figure 12 clearly show that the variability in EAD associated with portfolio composition is substantially greater than that arising from the UKCP18 RCM ensemble. At each proportion of assets at risk, the spread of the yellow boxplots (portfolio uncertainty, where EAD is averaged across ensemble members) is consistently much larger than that of the magenta boxplots (ensemble uncertainty, where EAD is averaged across portfolios). This demonstrates that the spatial distribution and value of assets within a portfolio exert a stronger influence on EAD than the specific climate realisation used to represent extreme rainfall.

Nevertheless, the magenta boxplots indicate a non-negligible spread across ensemble members, particularly at higher proportions of at-risk assets, showing that climate uncertainty still contributes to variation in EAD. Furthermore, the increase in EAD with proportion at risk is evident in both panels, with higher-risk portfolios exhibiting larger absolute variability. Comparing panels (a) and (b) in Figure 12, the overall spread across ensemble members increases in the future period, suggesting that EAD *estimates become increasingly sensitive to the choice of ensemble member as the analysis period approaches the middle of the century.*

Taken together, these results highlight that while portfolio composition is currently the dominant driver of variability in estimated flood risk, uncertainty due to climate projections becomes increasingly important as risk levels increase and into the future. This reinforces the need for portfolio-specific risk assessments, as both the proportion and spatial distribution of exposed assets strongly influence estimated impacts and their associated uncertainties."

17. Figure 12: It is not clear why the authors chose the particular distribution of the portfolio samples in panel b. In fact the variability in panel b appears less than that of panel a, which appears counter-intuitive. Separate plots with 1000 samples at the same flood risk percentage might be more helpful to compare the effect of flood risk area. A plot showing the flood risk area percent on x-axis with EAD distribution on y-axis would also be informative.

Please see our response to comment 16.

18. Paragraph 586: In order to illustrate the recoupment of costs, a discount factor should be added to the EAD.

We agree that the inclusion of discounting is important when evaluating the recoupment of adaptation costs over time. However, in this study the EAD is used as a measure of long-term average annual risk, rather than as part of a full cost–benefit or investment appraisal framework.

Incorporating discounting would require additional assumptions regarding time horizons, discount rates, and the temporal profile of both damages and adaptation costs, which are context-specific and beyond the scope of this simplified, large-scale analysis.

Our intention here is to provide a consistent and transparent framework for comparing relative flood risk across locations and portfolios, rather than undertaking detailed economic appraisal. We have clarified this limitation in the manuscript and note that discounting would be an important extension for future work focused on adaptation decision-making.

We have added to the end of Section 4:

"We note that the EAD represents an average annualised risk metric and does not incorporate discounting. Incorporation of discount rates would be required for formal cost–benefit analysis of adaptation measures, but this is beyond the scope of the present study and would require context-specific assumptions regarding time horizons and economic parameters."

19. Line 48: The authors highlighted variability in vendor estimates for flood risk. However, their manuscript does not address this limitation, nor does it provide a consistent framework that can be used to reduce this variability, especially given their rainfall to flood hazard mapping step. I would encourage the authors to highlight this limitation.

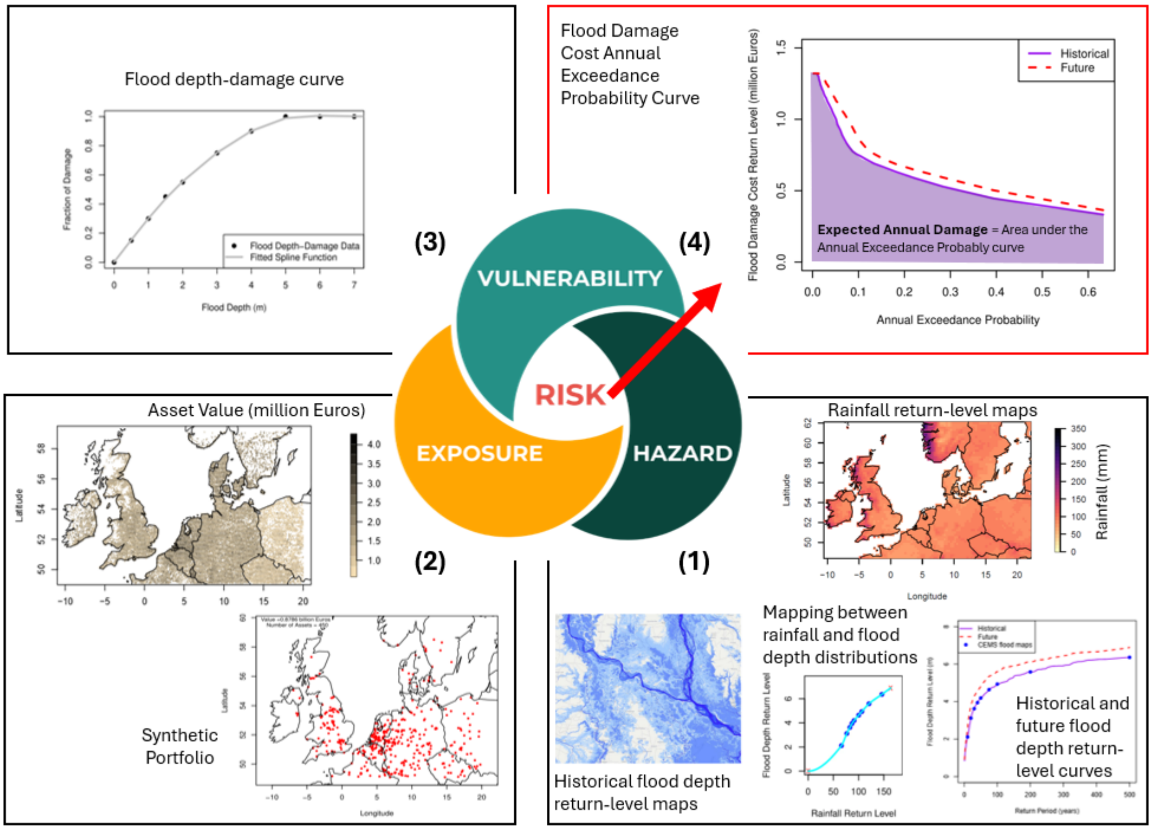
We agree that substantial variability exists across vendor flood risk estimates, often due to differences in data sources, modelling assumptions, and the use of proprietary or "black-box" methodologies. While our framework does not seek to define a single "best" approach to flood risk estimation, it is explicitly designed to improve transparency and comparability by making each modelling step (from rainfall to flood hazard, vulnerability, and exposure) explicit and reproducible.

We've added to the last paragraph of the introduction:

"In common with existing commercial flood risk products, there is no single ` best' method for translating rainfall extremes into flood hazard, and different modelling choices can lead to substantially different risk estimates. Rather than eliminating this variability, the framework presented here seeks to make key assumptions explicit by providing a transparent and reproducible method. This transparency enables clearer comparison between alternative model configurations, such as different portfolio constructions, time periods, or modelling choices, supporting more informed interpretation of differences in risk estimates, in contrast to proprietary or black-box methods where such differences may be difficult to diagnose."

20. Figure 1: While the sub-figures represent the various components of the methodology, it is difficult to interpret their flow, as mentioned in paragraph 105. A flowchart or similar step-by-step schematic will make it easier to understand how the figure represents the calculation flow.

We thank the reviewer for this helpful suggestion. We have revised Figure 1 to improve clarity of the methodological flow by explicitly numbering the key stages of the framework (hazard, exposure, vulnerability, and risk). We have also expanded the figure caption to include a clear, step-by-step description of how these components are combined, with references to the relevant sections of the manuscript. These changes are intended to provide a more intuitive representation of the calculation flow while retaining the original schematic structure, as a Venn diagram is often used for this type of risk schematic.



"Figure 1. A schematic of the approach used to combine information about the hazard, exposure and vulnerability to estimate idealised flood risk (here Expected Annual Damage) for any given time period and at any given location. The stages of the approach are numbered: (1) Hazard: Estimate the flood hazard by constructing a flood-depth return-level curve for each grid cell, based on rainfall

return-level curves and a statistical mapping from rainfall extremes to flood depth (see Section 2.2). (2) Exposure: Quantify the exposed asset value within each grid cell using built area and country-specific asset values per unit area (see Section 2.3). (3) Vulnerability: Apply a depth–damage (vulnerability) function to the flood-depth return-level curve to map flood intensity to the fraction of asset damage, producing a damage-fraction return-level curve (see Section 2.4). (4) Risk: Combine vulnerability and exposure to estimate flood damage costs and quantify risk by calculating the Expected Annual Damage (EAD) as the numerical integral (area under the curve) of the flood damage cost annual exceedance probability (AEP) curve (see Section 2.5)."

21. Figure 7: It would be helpful to include the return levels for all 12 ensemble members in the figure, in order to better understand their variability within the main manuscript.

Thank you for this helpful suggestion. We agree that visualising the spread across all 12 ensemble members is important for understanding the variability in return levels. However, the primary objective of Figure 7 is to illustrate the rainfall to flood depth return level curve mapping, rather than to fully characterise ensemble uncertainty within the main manuscript.

To maintain clarity and readability in the main text, we have therefore retained a simplified representation in Figure 7. The full spread of return levels across all 12 ensemble members is presented in the Supplementary Material (Fig S2), where this variability can be more clearly and comprehensively illustrated, without overcomplicating the main figure.

22. Figure 8b: Since the value is constant per country, it will be better represented by a table, than a gridded map which gives the impression of more spatial variability.

We appreciate the reviewer's observation. While the asset value per square metre is indeed constant at the country level, we have chosen to retain the gridded map representation for consistency with panels (a) and (c), and to allow direct spatial comparison across exposure and vulnerability components within the same visual framework. A tabulated version of the underlying country-level asset values is available in the database accompanying the Global Flood Depth-Damage Functions: Methodology and Database with Guidelines (referenced in the manuscript as Huizinga et al., 2017).

We have added a note on this in the figure caption: *"Note that asset values per square metre in subplot (b) are applied uniformly at the country level; spatial patterns therefore reflect national differences rather than within-country variability; a tabulated version of these values is available in the Global Flood Depth-Damage Functions database (Huizinga et al., 2017)."*

23. Figure 10: Due to the narrow dispersion, the panel (a) is difficult to read. Suggest removing or lightening the horizontal grid, and perhaps splitting the y-axis as Odense is primarily responsible for the extended y-scale. A log y-axis would be another option, given the losses in millions €.

Thank you for these suggestions. We have split Figure 10 into two to improve readability and lightened the horizontal gridlines in panel (a). Please see our response to comment 25 for further details.

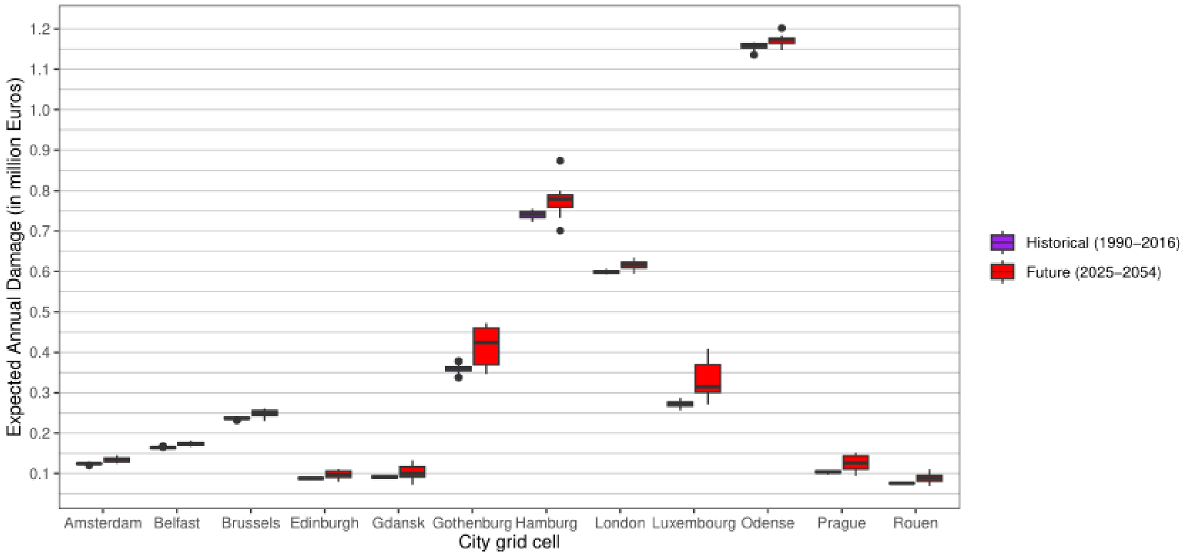
24. Figure 10: Can a statement be added in the caption describing the boxplot: whether the boxes represent 25-75 interquantiles, whiskers, etc.?

We have added to this caption: *"Boxes represent the interquartile range (25th–75th percentiles) of ensemble EAD, with the central line indicating the median; whiskers extend to 1.5 times the interquartile range, and points beyond this range are shown as outliers."*

25. Figure 10: Panel (b) map and the histogram are difficult to read, and the differentiation of colors is not apparent.

In response to comment 15 and this comment, we have modified panel (b) to show the relative % change in EAD (calculated as (future EAD - historical EAD)/historical EAD). This simplifies interpretation of the colour scale, and the accompanying histogram has therefore been removed as it was no longer informative. In addition, the colour scale has been revised, and, following comment 23, the figure has been split into two to improve overall readability. The revised version of Figure 10 is shown below (the other two subplots are now in a separate figure).

(a) Ensemble spread of historical and future Expected Annual Damage for 12 city grid cells



(b) Ensemble mean % Change in Expected Annual Damage (future – historical)/historical

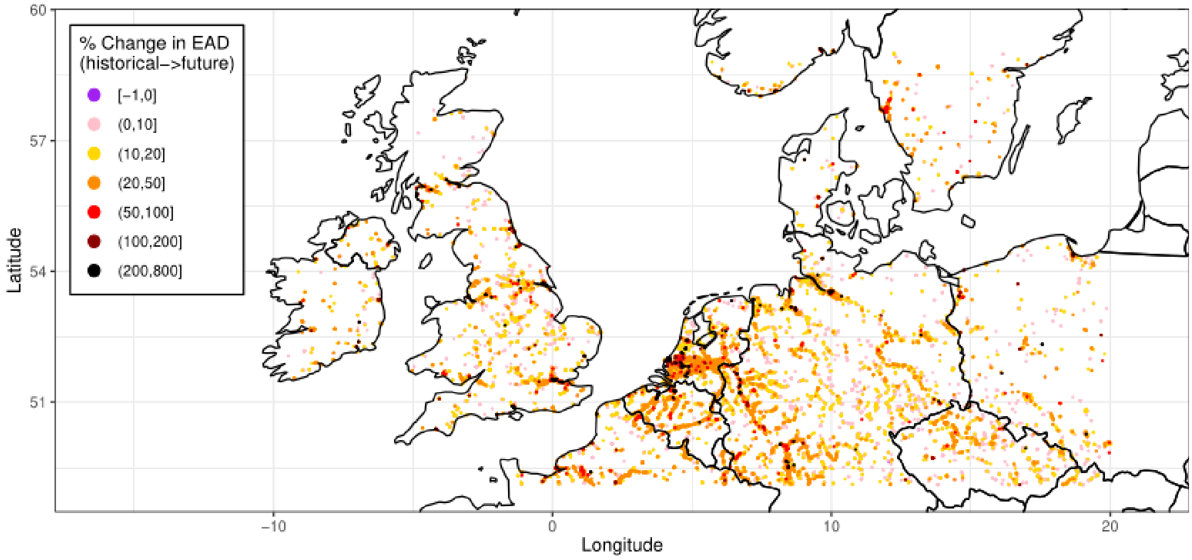


Figure 10. A comparison of historical and future flood risk, expressed as Expected Annual Damage (EAD) derived using the methodology described in Section 2: (a) UKCP18 RCM ensemble spread in EAD for 12 100 m x 100 m grid cells representative of the central regions of 12 European cities (see

Table 1 for more detail) in historical (1990–2016) and future (2025–2054) periods, Boxes represent the interquartile range (25th–75th percentiles) of ensemble EAD, with the central line indicating the median; whiskers extend to 1.5 times the interquartile range, and points beyond this range are shown as outliers; (b) UKCP18 RCM ensemble mean percentage change in EAD between historical and future periods, calculates as (future EAD - historical EAD)/historical EAD in each location. A value of this metric between 50-100 means the future EAD is 50-100% greater than the historical EAD.

26. Figure 12: Due to significant overlap in the dots, the overlaid ensembles are not informative. I would suggest changing the figure to make it clearer, e.g., showing the mean and standard deviation from the ensembles.

Please see our response to comment 16.

27. Figure 12: Why do some sampled portfolios have 0 EAD in panel b?

In the original version of Figure 12, the sampling strategy was not clearly described. In fact, portfolios were generated across proportions of at-risk assets ranging from 0.01 to 1.00 in increments of 0.01, with 10 portfolios sampled at each increment. As a result, some portfolios contained as little as 1% of assets in flood-risk regions, which led to very low EAD (~0) values when those assets were located in lower-risk areas. In the revised Figure, we instead consider proportions starting from 10%, removing this artefact (see our response to comment 16).

28. Figure 12: The numbering on the ensemble members seem to indicate that there are more than 12 ensemble members(15?). Can the authors add a statement if they selected a subset of ensembles and their selection criteria?

Thank you for this observation. The UKCP18 Regional Climate Model (RCM) ensemble comprises 12 active members, which are internally labelled 01, 04, 05, 06, 07, 08, 09, 10, 11, 12, 13, and 15. The missing indices correspond to global model simulations that were excluded due to physical instability or unrealistic climate behaviour, as part of the Perturbed Parameter Ensemble (PPE) design. Our updated version of this figure (see our response to comment 16) doesn't show ensemble member numbering, so we have not included this information in the manuscript.

29. Figure 13: At AEP = 1, the EAD must be zero as only 0 damage can occur with certainty. The figure does not represent this.

Thank you for this comment. We agree this should be the case in principle. In the revised analysis, AEP is defined using the Poisson representation (as per comment 13), for which the maximum AEP corresponding to a 1-year return period is approximately 0.63. We do not extrapolate to higher AEP values (i.e., return periods shorter than 1 year), and the x-axis in Figure 13 reflects this. This has now been clarified in the figure caption: *"AEP is defined using the Poisson representation (see Section 2.5), such that the maximum AEP corresponding to a 1-year return period is approximately 0.63. Values beyond this (i.e. return periods shorter than 1 year) are not shown."*

30. Figure 13: Can the authors clarify the axes of panel (a). The x-axis seems to represent the Flood hazard AEP, and the y-axis would be the *Expected* flood damage cost. The y-axis scale is in € so it does not seem to represent the "return level".

Thank you for flagging this issue in axis naming. We have changed the y axis *to 'Expected Flood Damage Cost'*. We note that the same issue is seen in Figure 9 (c) and (f). We have changed these axes labels too.

31. Line 45: What are transition risk tools?

Thank you for highlighting this. We agree that we can better define these tools here. We therefore propose an adjustment of the sentence to: *“For example, Bingler and Colesanti Senni (2022) found that transition risk tools (designed for understanding risk associated with the transition to a lower carbon economy) often suffered from poor transparency and communication around uncertainties and assumptions made.”*

32. Line 91: What does “idealised” refer to?

Thank you for this question. We have clarified our meaning in the text:

“The aim of this study is to produce idealised flood risk information (maps, summary plots, statistics) that demonstrate both current and plausible future physical climate risks to property portfolios. Here, “idealised” refers to the use of simplified and synthetic representations of hazard, exposure and vulnerability, designed to isolate key drivers of risk and to illustrate the methodological framework, rather than to provide detailed, site-specific predictions.”

33. Line 118: Nit: *Plausible* hazard scenarios is a more representative terminology, but “possible” is also acceptable.

We have changed the word to ‘plausible’ here.

34. Line 118: ...probabilities *of occurrence*.

We have added this.

35. Line 120: Can the authors add references to some existing hazard products, perhaps in the Introduction section, to illustrate that the authors researched for available event sets among widely available hazard products, bolstering their statement that no such event sets exist? This would also support their next paragraph since hazard products at different return levels are available.

Thank you for this comment. We stated *“Currently, to our knowledge, no open-source flood depth hazard event sets exist that cover Europe in a consistent way, which constrains the application of this risk assessment approach”*. We realise that this phrasing could be confusing as we then use hazard products at specific return levels which are available across Europe (i.e. https://data.jrc.ec.europa.eu/collection/id-0054). We have therefore adjusted the sentence and add a sentence to acknowledge what data products are available, as suggested by the reviewer:

“Currently, to our knowledge, no open-source flood depth hazard event sets exist for all return periods covering Europe in a consistent way for past and future climate length (20-30 year) periods, which constrains the application of this risk assessment approach. Although not an official flood hazard map, open-source river flood hazard data are available across Europe for nine return levels (Baugh et al., 2024, 2026) alongside recently released satellite derived flood maps (Betterle and Salamon, 2026). Historical Analysis of Natural Hazards in Europe (naturalhazards.eu) also provide historical information on European floods as well as modelled potential floods over the years 1950-2020 (Tilloy et al., 2024 and Paprotny et al., 2024).”

Baugh, Calum; Colonese, Juan; D'Angelo, Claudia; Dottori, Francesco; Neal, Jeffrey; Prudhomme, Christel; Salamon, Peter (2026): River flood hazard maps for Europe and the Mediterranean Basin region. European Commission, Joint Research Centre [Dataset] doi: 10.2905/JRC.WPE5YRR; 10.2905/1D128B6C-A4EE-4858-9E34-6210707F3C81 PID: http://data.europa.eu/89h/1d128b6c-a4ee-4858-9e34-6210707f3c81

Betterle, A., Salamon, P. (2026): Satellite-derived flood depth maps for Europe. European Commission, Joint Research Centre [Dataset] doi: 10.2905/JRC.NNQVVER; 10.2905/0bc96690-b89c-4909-9166-c2c322a20130 PID: http://data.europa.eu/89h/0bc96690-b89c-4909-9166-c2c322a20130

Tilloy A., Paprotny D., Grimaldi S., Gomes G., Bianchi A., Lange S., Beck H., Mazzetti C., Feyen L. (2024) HERA: a high-resolution pan-European hydrological reanalysis (1950-2020). Earth System Science Data, 17:293-316 https://doi.org/10.5194/essd-17-293-2025.

Paprotny D., Rhein B., Vousdoukas M. I., Teferenko P., Dottori F., Śledziowski J., Treu S., Feyen L., Kreibich H., Mengel M. (2024) Merging modelled and reported flood impacts in Europe in a combined flood event catalogue, 1950-2020. Hydrology and Earth System Sciences 28:3983-4010 https://hess.copernicus.org/articles/28/3983/2024

36. Line 126: Similar to above comment, Figure 1 does not adequately illustrate how the probability distributions will be combined.

As per comment 20, we have updated this figure and the caption, and in this part of the text added some extra detail: *"By combining this curve with vulnerability and exposure information, EAD can be approximated by integrating over the annual exceedance probability curve of the resulting impact intensities (as in the red box in Figure 1, where each step of the method is described in the sections referenced in the figure caption)"*

37. Line 128: Can the authors add a citation for the identified dataset on its first mention?

Yes, good idea - we have added this.

38. Line 272: Figure 3c -> Figure 3d

Thank you for spotting this, we have corrected the typo.

39. Figure 6: Regarding the two additional points, what is the associated y-axis point corresponding to the highest rainfall x-point?

Thank you for this question, and we apologise for the lack of clarity in the manuscript. The y-axis value corresponding to the highest rainfall point is obtained through a controlled extrapolation: the rainfall range is extended to a common maximum to ensure stable interpolation, and the associated flood depth is estimated by linearly extrapolating from the gradient of the two highest historical points. This avoids artefacts associated with unconstrained spline extrapolation and ensures a physically consistent, monotonic relationship.

We have added to the figure caption: *"The red crosses indicate two additional points included in the spline fitting to enable extrapolation beyond the nine quantiles: one at the origin (0,0), i.e. no rain = no flood, and one at the maximum rainfall across the combined historical and future periods. The associated maximum flood depth is estimated by linearly extrapolating from the gradient of the two highest historical points, ensuring stable, monotonic, and physically consistent behaviour."*

40. Line 351: Can the authors elaborate on "several modelling caveats"?

In response to your comment 2, Lines 351-353 have been rewritten and this phrase is no longer in the text. This has been rewritten as: *"It is important to emphasise that this approach does not assume a direct event-based 1:1 correspondence between rainfall and flood events. Rather, it represents an idealised quantile-to-quantile mapping between rainfall and flood return-level distributions. In practice, related assumptions are commonly made in engineering 'design event' approaches, where rainfall events of a given return period are used to estimate flood magnitudes of the same nominal return period (e.g. Kourtis et al., 2020). However, in reality, flood frequency curves arise from the integration of the full spatio-temporal rainfall distribution with catchment processes, and multiple rainfall events with differing characteristics can produce similar flood magnitudes. As such, the relationships derived here represent a simplified statistical approximation of a highly complex physical system. They do not explicitly account for key hydrological processes and are not validated against independent flood observations or model output. Consequently, this approach should be interpreted as a monotonic, first-order approximation of the relationship between rainfall and flood extremes, intended to enable large-scale analysis in the absence of consistent open hydrological datasets, rather than a physically complete or decision-ready representation of flood hazard."*

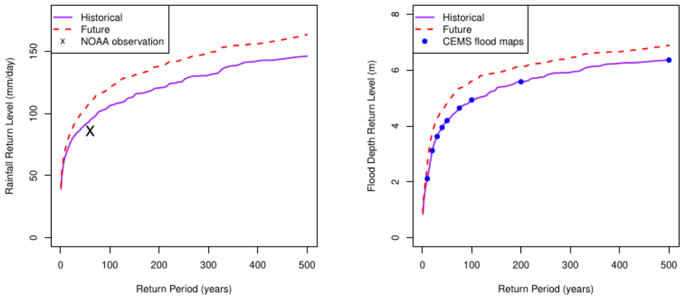
We have also added an additional consideration of the caveat associated with the mismatch between the spatial resolutions of the rainfall and flood depth data – see our response to Reviewer 2's first comment.

41. Line 352: Can the authors provide supporting evidence for the relationships being informative, and an associated citation for their next statement?

As above, this phrase is no longer in the manuscript text, replaced with more detailed caveats – see our response to Reviewer 2's first comment.

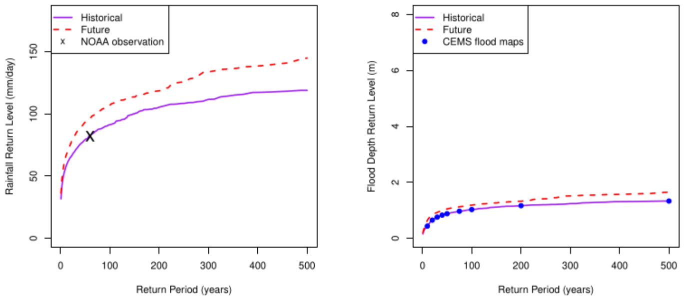
42. Line 359: Can the authors mark the 60-year return period values on panels a and c?

This is a great idea, thank you for this suggestion. Please see the updated figure below:



(a) Gothenburg: Rainfall return level curve

(b) Gothenburg: Flood depth return level curve



(c) Edinburgh: Rainfall return level curve

(d) Edinburgh: Flood depth return level curve

We have also referred to these points being on the figure and added this to the figure caption.

43. Line 368: Can the authors qualify the return-period, e.g., return-level estimates for ≥ 100 years…?

Thank for this suggestion. We have added this here.

44. Line 369: It is unclear to me from the figure why the estimates are notably high if they align with the CEMS flood map values, which are also ≥5m for ~100-yr return periods.

We agree that the original wording was unclear. Our intention was not to suggest that the spline-derived estimates are inconsistent with the CEMS flood map values, but rather that the CEMS values themselves appear high for this particular location. In this case, the high return-level estimates (>5 m) arise because the fitted curves closely follow the CEMS flood depth data points, which already exhibit relatively large values at higher return periods.

We have revised the text to clarify that the apparent overestimation originates from the underlying CEMS flood hazard dataset rather than from the interpolation or modelling approach itself: *“We note that for Gothenburg, the historical and future flood depth return-level estimates for return periods greater than 100 years appear notably high (reaching in excess of 5 metres). These values closely reflect the corresponding CEMS flood map estimates, which themselves exhibit relatively large flood depths for high return periods in this grid cell. Although previous flood modelling work for Gothenburg has reported water depths in excess of 2 metres (Filipova and Rana, 2012), this discrepancy suggests that the known caveats associated with the CEMS dataset (discussed in Section 2.2.1) may be contributing to an overestimation of flood depths at this location.”*

45. Line 497: Would it be possible to include the total non-residential property value of the representative 100m cells in Table 1?

Yes, this is a good idea. We have added these values to Table 1.

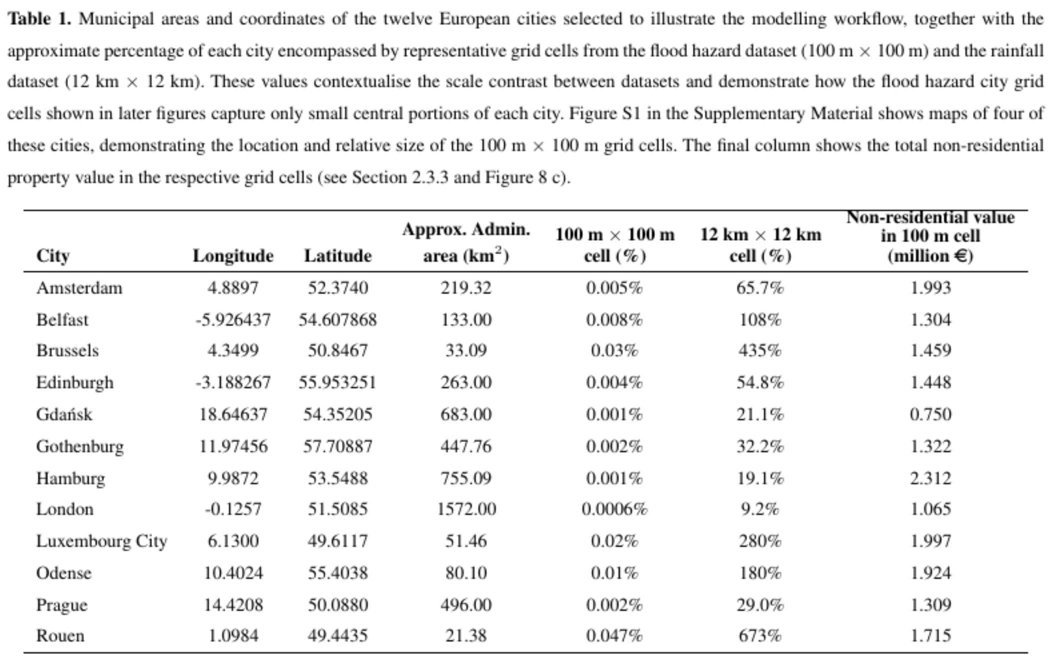


Table 1. Municipal areas and coordinates of the twelve European cities selected to illustrate the modelling workflow, together with the approximate percentage of each city encompassed by representative grid cells from the flood hazard dataset (100 m × 100 m) and the rainfall dataset (12 km × 12 km). These values contextualise the scale contrast between datasets and demonstrate how the flood hazard city grid cells shown in later figures capture only small central portions of each city. Figure S1 in the Supplementary Material shows maps of four of these cities, demonstrating the location and relative size of the 100 m × 100 m grid cells. The final column shows the total non-residential property value in the respective grid cells (see Section 2.3.3 and Figure 8 c).

City	Longitude	Latitude	Approx. Admin. area (km^2)	100 m × 100 m cell (%)	12 km × 12 km cell (%)	Non-residential value in 100 m cell (million €)
Amsterdam	4.8897	52.3740	219.32	0.005%	65.7%	1.993
Belfast	-5.926437	54.607868	133.00	0.008%	108%	1.304
Brussels	4.3499	50.8467	33.09	0.03%	435%	1.459
Edinburgh	-3.188267	55.953251	263.00	0.004%	54.8%	1.448
Gdańsk	18.64637	54.35205	683.00	0.001%	21.1%	0.750
Gothenburg	11.97456	57.70887	447.76	0.002%	32.2%	1.322
Hamburg	9.9872	53.5488	755.09	0.001%	19.1%	2.312
London	-0.1257	51.5085	1572.00	0.0006%	9.2%	1.065
Luxembourg City	6.1300	49.6117	51.46	0.02%	280%	1.997
Odense	10.4024	55.4038	80.10	0.01%	180%	1.924
Prague	14.4208	50.0880	496.00	0.002%	29.0%	1.309
Rouen	1.0984	49.4435	21.38	0.047%	673%	1.715

46. Paragraph 524: Can the differences between the portfolios be described in this paragraph instead of two paragraphs later? This would help readers anchor the differences in EAD.

We have revised the manuscript accordingly and moved the description of the key difference between the portfolios earlier in the paragraph - stating that Portfolio 1 and Portfolio 2 contain 20% and 80% of locations at risk of flooding, respectively. We agree that introducing this information at this stage improves clarity and helps the reader better interpret the differences in EAD between the portfolios.

47. Line 526: It would be better to highlight the differences in the spatial distribution rather than the apparent similarity to help readers focus on the relevant differences, e.g., Portfolio 2 appears to have a higher concentration of built area along the coasts.

We thank the reviewer for this suggestion. We agree that highlighting spatial differences can be useful; however, the intention of this example is specifically to demonstrate that portfolios which appear similar in terms of overall spatial distribution and aggregate characteristics can nevertheless exhibit substantially different levels of risk. This reflects a realistic challenge faced by asset managers, where portfolios may not exhibit obvious visual differences despite having materially different exposure to flood hazard.

For this reason, we have retained the emphasis on the apparent similarity between the portfolios in the main text, as this underpins the key message that quantitative risk assessment is necessary to distinguish between them.

We have slightly tweaked the wording here to emphasise this: *“This example highlights the critical need for pension sector asset managers to quantify the physical climate risks embedded within their portfolios. Portfolios that appear broadly similar based on aggregate characteristics may nevertheless have materially different risk profiles, and these differences are not always evident without explicit risk analysis. Accurately assessing this risk is therefore essential, and proactive measures should be taken to reduce exposure where possible.”*

48. Line 536: How is risk of flooding determined: Is it based on flooding depth exceeding some value for a certain X-year hazard?

This is explained in the ‘Relating rainfall and flood depth return-levels’ section in Section 2.2.3 (line 340), and in the caption of Figure 6 where the ‘at-risk’ locations are shown in blue in the top panel. We have added reference to this in this section of the manuscript for completeness: *“This arises primarily because Portfolio 1 and Portfolio 2 are sampled such that 20% and 80% of locations are at risk of flooding respectively (where locations at risk of flooding are indicated by non-zero flood depth for events up to at least the 1-in-75-year return-period in the CEMS flood hazard maps, see the left panel of Figure 6).”*

49. Line 541: It is unclear how to compare the variability of the ensemble members and that of EAD.

See our response to your comment 16 – we have updated this figure to more clearly show this. Thank you for your input in improving this figure.

50. Line 554: It is not clear how to conclude from Figure 12b that portfolios with a higher flood risk area have higher EAD. Similarly for the following statement.

As above, we have updated this Figure – see response to comment 16.

51. Line 573: Does the conversion from 10-year flood to 6-yr flood apply to all ensemble members?

Thank you for this question. Yes, the same methodology is applied consistently across all portfolio locations and all ensemble members. However, the specific conversion from a historical 1-in-10-year event to a future return period (e.g. 6-year in the example shown) is not fixed. Instead, it is calculated independently for each location and each ensemble member based on the corresponding flood depth return-level curves. As a result, the equivalent future return period varies across both space and ensemble members.

We have clarified this point in the manuscript to avoid any ambiguity: *"This approach is applied to all locations in Portfolio 2. It is important to note that this conversion is calculated separately for each location and ensemble member based on their respective return-level curves. As a result, the equivalent future return period corresponding to the historical 1-in-10-year event varies across both space and ensemble members."*

We have also added clarification in the adaptation figure caption that: *"The uncertainty in the box-plots quantifies the UKCP18 RCM ensemble spread, where box, whiskers and outliers are as defined in Figure 10."*

52. Line 587: Can the authors add a citation or justification for their assumed adaptation cost?

Thank you for highlighting this. In the original version, the value of €1000 per m^2 was intended purely as a simplified assumption for the toy model and was not based on an explicit evidence base. Following the reviewer's comment, we revisited the literature and found that this value sits toward the upper end of plausible estimates. In response, we have revised the assumption to €500 per m^2, which is more consistent with published ranges while still representing an upper-bound estimate for non-residential buildings. We have also added an appropriate citation and supporting justification in the manuscript:

"To illustrate this, we use the same toy example and assume that adapting Portfolio 2 assets to withstand the historical 1-in-10-year flood depth incurs a cost of €500 per m^2 for 1 m of flood depth. This assumption is informed by European cost estimates synthesised by Aerts (2018), which suggests that wet and dry flood proofing measures for approximately 1 m of flood depth correspond to costs of order €100-350 per m^2 for residential buildings. Given the focus of this study is on non-residential assets and recognising the additional service complexity and specification typically required for such buildings, we adopt an upper bound value of €500 per m^2 for 1 m of flood depth. This choice is intended to provide a precautionary assumption suitable for illustrative adaptation analysis, rather than site-specific appraisal."

Aerts, J.C.J.H. (2018) A review of cost estimates for flood adaptation. Water, 10(11), p. 1646. doi: https://doi.org/10.3390/w10111646.

Reviewer 2:

The paper addresses a highly relevant topic—financial risk associated with extreme rainfall and flooding—and proposes a transparent framework based on open data. The proposal is useful as a demonstrative and educational framework. However, there are substantial methodological limitations, some acknowledged by the authors, that reduce the quantitative robustness of the conclusions.

We thank the reviewer for their thorough and constructive feedback. Many of the points raised closely overlap with issues highlighted by Reviewer 1, particularly regarding the rainfall–flood mapping, model assumptions, and interpretation of results. In revising the manuscript, we have taken a consistent approach to addressing these concerns, strengthening the clarity of assumptions, limitations, and intended use of the framework. Below we respond to each comment in turn and describe the specific changes made to the manuscript, referring back to the responses to Reviewer 1 where relevant.

Below are some points that, in my opinion, should be clarified and explored in more depth in the paper:

1. Discuss the inconsistency between the different spatial resolutions used (12 km for rainfall versus 100 m for flooding and financial assets);

Thank you for this suggestion. We have added (at around line 350):

"A further caveat arises from the mismatch in spatial resolution between the rainfall and flood depth datasets used (as described above). As a result, multiple flood exposed locations may share identical rainfall return levels despite experiencing different flood dynamics, adding further uncertainty to the derived relationships. The quantile-to-quantile mapping therefore implicitly assumes that rainfall statistics at the coarser climate model scale are representative of the local hydrological forcing relevant for each flood prone grid cell. In reality, sub grid scale variability in rainfall, land cover, drainage, and topography can lead to substantial heterogeneity in local flood response that cannot be resolved at the climate model resolution. This mismatch may also introduce systematic bias, for example where convective rainfall extremes are smoothed at 12 km resolution, potentially leading to underestimation of localised flood-generating intensities. Despite this limitation, the approach remains suitable for capturing large-scale patterns of exposure and relative risk, as the quantile-mapping framework preserves the relative ordering of extreme events across space. This enables consistent comparison of hazard severity and exposure between regions, even if absolute local flood magnitudes are uncertain. Future work could reduce this limitation through higher-resolution climate data or explicit hydrological modelling. This scale inconsistency reinforces the interpretation of the approach as a large-scale approximation rather than a site-specific flood hazard assessment."

And we have added to the 'Simplified rainfall-flood mapping and lack of hydrological modelling' limitation: *"A further limitation arises from the mismatch in spatial resolution between datasets. Rainfall extremes are modelled at 12 km resolution, while flood depths are represented at 100 m resolution. The mapping therefore assumes that rainfall statistics at the climate model scale are representative of local flood-generating conditions."*

2. Lines 320-325: Explain under what physical conditions the rainfall → flood depth approximation can be considered valid.

We thank the reviewer for this helpful comment. This point is closely related to Reviewer 1 Comment 2, where we discuss in detail the assumptions and limitations of the simplified rainfall–flood mapping approach, and we refer the reviewer to that response for a fuller explanation.

In particular, we have added to this section: "It is important to emphasise that this approach does not assume a direct event-based 1:1 correspondence between rainfall and flood events. Rather, it represents an idealised quantile-to-quantile mapping between rainfall and flood return-level distributions. In practice, related assumptions are commonly made in engineering 'design event' *approaches, where rainfall events of a given return period are used to estimate flood magnitudes of the same nominal return period (e.g. Kourtis et al., 2020). However, in reality, flood frequency curves arise from the integration of the full spatio-temporal rainfall distribution with catchment processes, and multiple rainfall events with differing characteristics can produce similar flood magnitudes. As such, the relationships derived here represent a simplified statistical approximation of a highly complex physical system. They do not explicitly account for key hydrological processes and are not validated against independent flood observations or model output. Consequently, this approach should be interpreted as a monotonic, first-order approximation of the relationship between rainfall and flood extremes, intended to enable large-scale analysis in the absence of consistent open hydrological datasets, rather than a physically complete or decision-ready representation of flood hazard."*

3. The entire methodology relies on statistical mapping and does not present explicit observational validation; especially Section 2.2.3 and Section 5. Is it possible to include independent validation using observed flood data, actual economic losses, or official risk maps?

We thank the reviewer for this important comment. This concern closely aligns with Reviewer 1 Comment 1 and Comment 2, where the lack of validation of the rainfall–flood mapping approach and its implications have been discussed in detail. In particular, we have added text in the Abstract, Introduction, Methods, Limitations, and Conclusions emphasising that (i) the rainfall–flood relationship is not validated against observations or independent hydrological modelling, (ii) this introduces substantial structural uncertainty, and (iii) the resulting estimates are intended to be illustrative rather than decision-grade.

We also refer the reviewer to our response to Reviewer 1 Comment 14, where we explain the challenges associated with comparing model outputs to observed economic losses.

To further clarify this point, we propose adding the following sentence to the 'Assumptions around synthetic portfolio design and financial context' Limitations section:

"In addition, the methodology is not validated against observed economic losses, insurance claims, or official flood risk products; given the differences in spatial scale, temporal aggregation, and modelling assumptions, such comparisons are not directly comparable, and the resulting damage estimates should therefore be interpreted as indicative rather than calibrated values. In future potential revision to this framework, where more physically based hydrological and hydrodynamic modelling is used, such validation against observed impacts and official risk products would be essential to ensure the reliability and applicability of the results."

4. Section 2.2.3: Justify the implicit assumption of stationarity of the rainfall-flood relationship under future climate change.

Thank you for raising this important point. We acknowledge that, under climate change, the physical relationship between rainfall extremes and flooding may evolve due to factors such as changes in rainfall intensity structure, soil moisture regimes, land use, or flood management. Explicitly modelling such non-stationarity would require physically based hydrological and hydrodynamic models driven by high-resolution climate data, which are not consistently available as open datasets at continental scale.

We have revised the manuscript to (i) explicitly state the assumption of stationarity of the rainfall–flood relationship within each time period, (ii) clarify that projected changes in flood hazard arise from changes in rainfall extremes rather than changes in the rainfall–flood mapping itself, and (iii) emphasise that this represents a first-order approximation appropriate for large-scale, illustrative analysis.

We have added to this section: *"It should be noted that this approach assumes the rainfall–flood mapping is stationary across the two time periods. That is, while rainfall extremes are allowed to change between historical and future periods, the statistical relationship used to map rainfall return levels to flood depth return levels is held fixed. As such, the projected changes in flood hazard arise from changes in rainfall extremes rather than changes in the rainfall–flood mapping itself. Explicitly modelling such non-stationarity would require physically based hydrological and hydrodynamic models driven by high-resolution climate data, which are not consistently available as open datasets at continental scale, as discussed in Section 2.2.1."*

And we have added another section in the Limitations: *"Assumption of stationarity in the rainfall–flood relationship: The methodology assumes that the statistical relationship between rainfall extremes and flood depth remains stationary across historical and future climate periods. In reality, climate change may alter hydrological processes such as runoff generation, soil moisture dynamics, drainage efficiency, and flood routing, potentially modifying the rainfall–flood relationship over time. These effects are not explicitly represented in the present framework. As a result, projected changes in flood hazard reflect changes in rainfall extremes rather than potential changes in catchment response. This assumption is appropriate for the large-scale, illustrative analysis undertaken here, but would need to be revisited in applications intended for local-scale or decision-grade flood risk assessment."*

5. Figure 8: Discuss the limitations associated with the use of national average property values, ignoring intra-urban and regional heterogeneity.

Thank you for highlighting this. The use of national average property values is a simplifying assumption that was adopted to enable a consistent, pan-European analysis using fully open and harmonised datasets. We fully agree that this approach does not capture substantial intra-urban and regional heterogeneity in asset values, particularly within large metropolitan areas where property values can vary strongly with location, land use, and building characteristics.

We have revised the manuscript to explicitly acknowledge this limitation and clarify its implications for interpretation, adding to the text in Section 2.3.2 Property value data: *"Asset values per unit area are applied uniformly at the national level. This simplification does not capture intra-urban or regional heterogeneity in property values, particularly within large metropolitan areas, and may therefore misrepresent absolute damages at specific locations."*

And we have added to the limitations section: *"Use of national average exposure values: Exposure is represented using national average non-residential property values per unit area. This approach ignores potentially large intra-urban and regional variations in asset values driven by land use, economic activity, and urban structure. As a result, estimated damages at individual locations may be over- or under-estimated in absolute terms. Incorporating higher-resolution exposure data would be an important extension where suitable open-source datasets become available."*

6. Lines 599-605: Better justify the simplified relationship between precipitation extremes and flood depth, considering that the method ignores fundamental hydrological processes.

Thank you for this comment. This point relates to several previous comments, and we have made a number of revisions to the manuscript to more clearly highlight, critique, and justify the simplified rainfall–flood mapping approach. In particular, we have expanded the discussion in this part of the Limitations section (see our responses to Reviewer 1 Comments 1 and 2).

To better justify this approach in this limitation section we have now also added: *"Despite these limitations, the approach retains value by providing a transparent and internally consistent means of linking changes in rainfall extremes to corresponding changes in flood hazard at large spatial scales, where use of high-resolution physically based hydrological modelling is not feasible using currently available open data."*

7. Lines 620-645: Better justify the exclusive use of a single climate scenario and discuss differences under intermediate scenarios.

Thank you, we agree this needs more clarification in the manuscript. We discuss this initially in Section 2.2.2. We have added further justification to this section: *"Differences between emissions scenarios remain relatively small prior to mid-century (Figure 1 of Met Office, 2018) with stronger divergence emerging after around 2050, largely beyond the period considered in this study. Consequently, while RCP8.5 accelerates the timing of warming relative to present-day policy expectations, the magnitude of warming during the modelled period in this study (2025-2054) is not outside the envelope of plausible 21st-century global temperatures and is largely consistent with other scenarios in this near-future period. The use of a single scenario therefore enables a clear and internally consistent illustration of the methodology without materially affecting the qualitative conclusions. Future work could extend this analysis by incorporating multiple emissions scenarios and higher-resolution climate model ensembles as such datasets become more widely available at continental scale, allowing a more explicit exploration of scenario-dependent uncertainty."*

We have added to this limitation section: *"Regarding the use of a single climate scenario (here RCP8.5), for the near- to mid-century period considered in this study (2025–2054), differences in projected warming between emissions scenarios remain relatively small, with more pronounced divergence emerging primarily after around 2050 (see Figure 1 of Met Office, 2018). Under intermediate emissions scenarios (e.g. RCP4.5 or RCP6.0), projected changes during this period would be expected to follow broadly similar spatial patterns to those under RCP8.5, with differences in magnitude becoming increasingly pronounced later in the century. As a result, the use of a single high-emissions scenario enables a clear and internally consistent illustration of the methodology without materially affecting the qualitative conclusions regarding relative patterns of rainfall extremes and flood risk. Future work, particularly for more distant time horizons, could extend this analysis to multiple emissions scenarios as suitable high-resolution, pan-European climate datasets become available, enabling a more explicit assessment of scenario-dependent uncertainty."*

8. Is it possible to present a sensitivity or stability analysis of the extrapolation performed by the monotonic spline?

We thank the reviewer for this suggestion. The extrapolation in the monotonic spline arises from the need to extend the rainfall–flood relationship beyond the range of available flood return-level data. As described in the manuscript (caption for Figure 6) and in our response to Reviewer 1 Comment 39, this is controlled by the inclusion of a point at (0,0) and an additional upper-bound point: where the maximum rainfall value is mapped to a corresponding flood depth based on linear extrapolation from the gradient of the highest available return-level points. This ensures a smooth, monotonic, and physically consistent continuation of the relationship while avoiding artefacts associated with unconstrained spline extrapolation. The extrapolation is fully constrained by the monotonic spline formulation and the imposed upper-bound control point, and is therefore not expected to exhibit unstable behaviour beyond the assumptions used to define this point.

Based on this, the primary source of uncertainty in the extrapolation is the specification of this upper-bound flood depth. We have chosen this approach explained above as it provides a stable and transparent extension of the relationship that is consistent with the behaviour of the highest observed quantiles, while avoiding unrealistic curvature in the tail.

While we agree that a formal sensitivity analysis could be informative, systematically exploring alternative extrapolation assumptions would require additional modelling choices and is beyond the scope of this illustrative, large-scale framework.

In response to this comment and Reviewer 1 Comment 39, we have added to the Figure 6 caption: *"The red crosses indicate two additional points included in the spline fitting to enable extrapolation beyond the nine quantiles: one at the origin (0,0), i.e. no rain = no flood, and one at the maximum rainfall across the combined historical and future periods. The associated maximum flood depth is estimated by linearly extrapolating from the gradient of the two highest historical points, ensuring stable, monotonic, and physically consistent behaviour. The extrapolated portion of the spline is tightly constrained by the imposed bounding points. While sensitivity to these assumptions could be explored, this lies beyond the scope of the present idealised analysis, which focuses on a single, physically consistent specification."*

As an aside we have exported Figure 6 as a high-resolution PNG rather than a vector PDF to reduce file size and improve rendering performance. This significantly improves responsiveness when viewing the document, without affecting the visual quality of the figure.

9. Discuss in more depth the impacts of the absence of bias correction on the final monetary results.

Thank you for this comment. We agree that the absence of bias correction is an important consideration when interpreting the resulting monetary estimates.

We have added to the first paragraph of Section 3.1 of the results: *"As described in Section 2.2.2, no bias correction is applied to the rainfall data used in this analysis. This, and a number of other modelling caveats (discussed in detail in Section 5), may affect the absolute magnitude of simulated rainfall extremes, and consequently the resulting flood depths and associated monetary estimates. As such, the absolute values of expected annual damage (EAD) should be interpreted with caution, with greater emphasis placed on the relative behaviour and spatial patterns of risk."*

10. Lines 628-635: Explain how random synthetic portfolios adequately represent real financial portfolios.

We thank the reviewer for this comment. The synthetic portfolios are used in the absence of open source real financial information and to provide a neutral and controlled illustration of the methodology. While the exact values may differ for a real financial portfolio, the scientific conclusions and relative patterns of risk we identify are not expected to depend strongly on the portfolios being synthetic or real. Application to more realistic portfolios is, however, an important direction for future work.

We have added to Section 2.3.4: *"In the absence of detailed, openly available data on real financial portfolios, this synthetic approach provides a neutral and controlled framework for exploring the behaviour of the methodology. By avoiding underlying characteristics of real financial portfolios, the use of synthetic portfolios ensures that the resulting patterns of risk reflect the underlying hazard–exposure relationships rather than features specific to any individual portfolio. While the exact values may differ for real financial portfolios, the scientific conclusions and relative patterns of risk are not expected to depend strongly on whether portfolios are synthetic or real. Application to more* *realistic portfolios, where suitable data are available, represents an important direction for future work."*

11. Figure 13: Discuss the limitations of the simplified modeling of adaptation and flood protection measures.

Thank you for this suggestion. We agree that this simplification should be explicitly noted. We have added to the end of Section 4: *"We emphasise that this representation of adaptation is intentionally simplified and serves as an illustrative example within the modelling framework. In particular, the approach assumes uniform protection of assets up to a specified return level, does not account for defence failure, spatial heterogeneity in vulnerability or protection standards, or the phased implementation of adaptation over time. In addition, the adaptation cost model is highly idealised and does not capture site-specific engineering, economic, or policy constraints. As such, the results should be interpreted as indicative of the potential scale and implications of adaptation, rather than as a detailed or decision-ready assessment."*

12. Caution: Moderate the language of the conclusions, as the paper itself acknowledges that the framework is exploratory and demonstrative in nature

We thank the reviewer for this helpful comment. We agree that the conclusions should reflect the exploratory and illustrative nature of the framework. We have therefore moderated the language throughout the conclusions to avoid overstatement and to better align with the underlying assumptions and limitations of the analysis, while retaining the key insights regarding relative patterns of risk and the potential role of adaptation.

The conclusion now reads as follows:

"7 Conclusions

This study presents a transparent and pragmatic framework for assessing physical climate risk from extreme rainfall for European pension portfolios using openly available data and methods. By integrating hazard, exposure, and vulnerability components, we provide illustrative estimates of Expected Annual Damage for both historical and near-future climates.

The analysis highlights four key insights:

- ***Current and future risk is substantial:*** *Pension portfolios may already face notable flood-related financial risk, with this risk likely to increase across Europe under near-term climate change.*
- ***Portfolio composition matters:*** *Differences in asset location and hence value at risk of flooding drive greater variability in risk than climate model uncertainty within this framework.*
- ***There is a need for improved risk information in decision-making:*** *Financial actors (such as asset managers of pension funds) would benefit from incorporating physical climate risk considerations to support long-term financial resilience.*
- ***Climate adaptation reduces risk:*** *Adaptation measures have the potential to substantially reduce expected annual damage (EAD) and improve financial outcomes, although the magnitude of these benefits depends on the simplified assumptions adopted here.*

While the approach is designed for demonstration and awareness-raising, it highlights the value of developing transparent and interpretable tools for assessing physical climate risk in the financial sector. A key limitation of the approach is the use of a simplified and unvalidated mapping between rainfall extremes and flood hazard, which constrains the realism and applicability of the results for *decision-making. Future work should refine the hazard modelling, expand climate scenarios, and incorporate asset-specific vulnerability profiles to enable more comprehensive and decision-relevant risk assessments. Ultimately, integrating climate-informed perspectives into investment planning is likely to be important for improving the resilience of financial systems in a changing climate."*

Final response:

We have added an acknowledgement to the reviewers within the manuscript to thank them for their valuable input.

"Finally, we thank the reviewers for their constructive and thoughtful comments, which have significantly improved the clarity, rigour, and presentation of this manuscript."