Response to Reviewer #2

Author's remarks

We thank the reviewer for their valuable comments and appreciate the time and effort they put into reviewing our manuscript.

In response to feedback received from Reviewer #1, we have reframed the analysis around flood-related credit constraints and reduced the emphasis placed on mortgage default as the main outcome of interest. Our model framework estimates the degree to which uninsured flood damage increased the potential for financial states (e.g., negative equity and cashflow problems) that can impair a mortgage borrower's ability to fund repairs to their property through additional borrowing; however, given the lack of data on mortgage performance, the degree to which these states are predictive of default remains unclear. Whether negative equity or cashflow problems trigger a borrower to default is also likely to depend on the availability of alternative sources of funding for home repairs that were not explicitly modeled, including disaster assistance grants, personal savings, and informal transfers from family and friends. As such, we have moved away from describing these borrowers as "at risk of default" and instead describe them as "credit constrained," reflecting their diminished capacity to fund repairs by taking on debt while acknowledging that their long-term recovery outcomes are uncertain and depend on several factors not captured by our analysis. We believe that this more modest framing aligns our study objectives and research questions more closely with our methodological approach.

As part of this reframing, we have made substantial changes to the introduction and background sections of the manuscript, and have modified our terminology as follows:

Original terminology	**New terminology**
"at risk of default"	"credit constrained"
"at risk of strategic default"	"collateral constrained"
"at risk of cashflow default"	"income constrained"
"at risk of double-trigger default"	"constrained by both income and collateral"

Please be aware that our responses to Reviewer #2's comments utilize the new terminology, even when their comments use the original terminology. In addition, line numbers referenced in our response correspond to the revised manuscript and may differ from those referenced in the reviewer's comments. A point-by-point response to their review is included below, with reviewer comments shown in **black** and our replies in **blue**.

Review of “Flood risks to the financial stability of residential mortgage borrowers: An integrated modeling approach”

General Comment Statement

The paper presents a significant and well documented contribution to the field of climate financial risk. The integrated, "bottom-up" modelling framework, which links property-level flood damage to household financial distress, is a commendable and ambitious effort to advance the understanding of this critical issue.

Our reply: We thank the reviewer for their thoughtful evaluation of our analysis.

As I am not an expert in US insurance market, I focused the review mainly on the modelling part of the work. While the framework is conceptually sound, the analysis concludes that its current implementation contains a cascade of methodological limitations that is likely going to lead to a systematic underestimation of the true financial risk.

Our reply: We thank the reviewer for highlighting this important concern. We agree that our estimates of flood damage exposure and (by extension) post-flood credit constraints are likely to be conservative due to the low sensitivity of our flood damage model, which fails to detect many properties that were damaged. We have revised the manuscript in several places to clarify for readers that our findings should be interpreted in light of this limitation (see responses below).

Line 18: The finding that the evaluated floods "generated $4.0 billion in property damage" is a key quantitative output. However, this figure should be interpreted as a conservative bound. As mentioned below (see comments on Line 271), the damage detection model fails to identify a majority of properties that actually sustained damage, meaning the true total damage might be higher.

Our reply: We have revised this sentence in the abstract to clarify that $4.0 billion is a conservative bound.

Revised text from lines 18-19: Conservative projections suggest that the floods evaluated generated $4.0 billion in property damage across the study area, of which 66% was uninsured.

Line 20: The statement that 32% of affected borrowers lacked sufficient income or collateral, "placing them at an elevated risk of default," is based on the underwriting criteria used. Is there any risk that the criteria used could lead to an underestimation of the number of borrowers who would be denied credit and thus be at risk?

Our reply: The underwriting criteria used to classify borrowers as income constrained or collateral constrained reflect the maximum allowable DTI and CLTV ratios under a government program that insures mortgages made by lenders to disaster-affected property owners; however, borrowers below these limits can still be denied a loan based on their credit history. While we believe that existing mortgage borrowers are likely to have high credit scores relative to the general population, those with a history of missed payments may face additional challenges in accessing repair loans that are not captured by our modeling framework. We have included the following text in the manuscript to highlight how the omission of factors related to credit history are likely to influence our findings.

Revised text from lines 547-551: It should be noted that borrowers meeting these ratio-based criteria can still be denied a loan due to unsatisfactory credit history—a process that is not represented in our modeling framework. While existing mortgage borrowers have (by definition) previously met lending standards and likely possess higher credit scores than the general population, the omission of factors related to credit history may cause us to underestimate the share of flood-exposed borrowers who would be denied a loan.

Lines 234-240: The generation of "pseudo-absence" points is a pragmatic solution to incomplete data but introduces noise. The authors' own validation (Line 276) shows that model precision increases significantly when these points are excluded, suggesting that a number of these randomly generated "undamaged" points likely distorted the model's training (actually damaged?).

Our reply: We appreciate the reviewer's point and agree that the use of pseudo-absences introduces some label noise into the training data. However, this step was necessary to correct for the selection bias inherent in the address-level insurance data, which disproportionately captured damaged (claim) locations. Without pseudo-absences, flood presence locations are overrepresented in the training data, leading to systematic overprediction of flood damage across the study area, particularly for pre-2009 events where the number of missing address-level policies was high. While it is true that precision improves when pseudo-absences are excluded from the validation data, this result mainly reflects a change in how the model is being evaluated rather than a true increase in predictive performance. Since our goal is to generalize predictions to the broader set of properties (including those without insurance), including pseudo-absences in the training data provides a more representative sample, even if it introduces some label noise.

Revised text from lines 292-294: While the inclusion of pseudo-absences likely introduces some label noise into the training data, this step was necessary to correct for the bias inherent in the address-level insurance data, which disproportionately captured damaged (claim) locations.

Line 271: A very low sensitivity of just 12% to 42% means the model fails to identify between 58% and 88% of properties that were actually damaged. This is a foundational error that guarantees a systematic underestimation of the total number of impacted households and the total damage costs. All subsequent risk estimates are therefore performed on a small fraction of the true at-risk population.

Our reply: We thank the reviewer for prompting us to address this issue more explicitly in the manuscript. We agree that the model's low sensitivity indicates that many damaged properties were not detected, leading to an underestimation of true flood exposure. We have revised the manuscript to more clearly acknowledge this limitation and to explain its cause and implications. Specifically, we now note that low sensitivity is characteristic of models trained on class-imbalanced data, and we clarify that our results should be interpreted as a conservative lower bound on total flood exposure rather than a central estimate.

Revised text from lines 324-333: When identifying damaged properties, the model exhibited high accuracy (≥92%) and specificity (≥98%) but low sensitivity, with true positive rates of between 12% and 42% across events. This behavior is characteristic of machine learning classifiers trained on class imbalanced data where the positive class (e.g., presence of flood damage) is rare compared to the negative class (Haixiang et al., 2017; He and Cheng, 2021). Among properties that were misclassified by our model in cross-validation, false positive and false negative predictions respectively accounted for 12% and 88% of model errors across the seven evaluated events (Table S4). Collectively, these results suggest that our model often fails to detect properties that were damaged, which is likely to lead to a systematic underestimation of the true level of flood exposure within the study area. As such, our projections of flood damage exposure (and, by extension, flood-related credit constraints) should be interpreted as a conservative bound as opposed to a central estimate.

Lines 272-276: The authors' framing of this result is misleading. The model's high precision is emphasized while downplaying the severe consequence of the high false-negative rate. In risk assessment, particularly for disaster aid, the cost of a false negative (failing to identify a household in need) is high.

Our reply: We have removed the text related to the model's high precision from the paragraph describing our cross-validation results, which now places greater emphasis on the consequences of our model's high false-negative rate. The revised paragraph is shown in our response to the reviewer's previous comment.

Lines 343-345: The reported accuracy is a major concern. Only 54% of the model's value predictions fall within ±20% of the actual sale price. The authors later note this is the largest source of uncertainty in their final results (Lines 674).

Our reply: We acknowledge the substantial uncertainty in our property value estimates. To address this, we have added text in the revised manuscript highlighting potential sources of error in our property valuation model. In response to comments from Reviewer #1, we have also included a variance-based sensitivity analysis that quantifies the contribution of uncertain model inputs (including property values) to variation in our projections of post-flood credit constraints. This analysis confirms that property values are the largest source of uncertainty in our results.

Revised text from lines 406-414: The substantial uncertainty in our property value estimates likely arises from a combination of factors, including: (1) the limited number of property-specific details in NCEM's statewide building inventory, which describes basic structural attributes but lacks information on other price-relevant characteristics such as recent improvements or deferred maintenance; (2) the presence of sales that do not reflect fair market values (e.g., intrafamily transfers) in the training and validation data, which can bias model predictions; and (3) geolocation errors that may result in mismatches between recorded sales and parcel geometries. Future work could potentially enhance the performance of the property valuation model by introducing filters to identify arms-length sales and by adding predictors that capture property-specific attributes related to structural defense and prior flood exposure (Nolte et al., 2024; Pollack and Kaufmann, 2022).

Lines 357-360: Is the use of GSE data to model the entire market, creating a bias of the "typical" borrowing population? If yes, it should be stipulated to keep the modelling results in perspective.

Our reply: We thank the reviewer for highlighting this important modeling assumption. We have added the following text to the manuscript to clarify the degree to which the GSE loans are representative of the broader U.S. mortgage market.

Revised text from lines 480-490: It is important to note that mortgages acquired by the GSEs—which account for approximately half of all U.S. mortgage originations (GAO, 2019)—consist of "conforming" loans that meet standardized requirements related to loan size, borrower credit quality, and documentation. Mortgages that are not represented in the GSE data include "jumbo" loans whose amounts exceed the conforming loan limit, which are typically associated with very expensive properties; "subprime" loans made to borrowers with questionable credit history or unverifiable income, which peaked at 15% of the U.S. mortgage market in the years leading up to the 2007 subprime mortgage crisis (Agarwal and Ho, 2007); and loans insured by government programs targeting specific groups such as first-time homebuyers, veterans, and active-duty military personnel (Jones, 2022; Perl, 2018). As such, borrower attributes that were simulated based on GSE data primarily reflect the characteristics of middle-income, creditworthy borrowers, and may underrepresent the characteristics of households at both the upper and lower ends of the wealth distribution and of communities in North Carolina with a large military presence such as Cumberland, Onslow, and Craven counties (N.C. Department of Military and Veterans Affairs, 2025).

Line 583: The conclusion must be interpreted as a conservative floor, not a central estimate, due to the cascading methodological issues outlined above and should be mentioned as such.

Our reply: We have added the following text to the manuscript to underscore that our projections of flood-related credit constraints are likely to be conservative.

Revised text from lines 691-695: Over the study period, 7,180 mortgage borrowers were projected to face flood-related credit constraints as indicated by ACLTV > 100% or ADTI > 45%. This number represents a small share (0.15%) of all mortgages originated during the study period but a substantial fraction (32%) of those exposed to flooding. Given the relatively low sensitivity of our flood damage model observed in cross-validation (Section 3.3), our projections may underestimate the true number of borrowers exposed to flooding over the study period and prevalence of flood-related credit constraints.

Revised text from lines 894-897: In addition, cross-validation results suggest that our machine learning-based approach often failed to detect properties that were damaged, which is likely to contribute to a systematic underestimation of the true level of flood exposure within the study area. For these reasons, our projections of flood damage exposure and flood-related credit constraints should be interpreted as conservative bounds as opposed to central estimates.

Line 494-675: While the paper's focus is on flood risk, its analysis spans a period in which the U.S. housing market underwent its most significant shock in generations. From 2008, the model might overlook a critical variable that shaped housing values, credit availability, and the underlying financial health of borrowers.

Our reply: We thank the reviewer for encouraging us to consider how the 2008 global financial crisis (GFC) may have impacted the underlying financial health and credit access of mortgage borrowers during the study period. We have included the following text in the revised manuscript to highlight how these factors (which were not explicitly modeled) are likely to influence our findings. It is worth noting that our property value model, which uses observed property sales as an input, reflects the effects of the GFC on home prices and (by extension) borrower equity.

Revised text from lines 905-914: Fourth, our model framework does not account for how factors related to the 2008 global financial crisis (GFC) may have impacted the financial health and credit access of mortgage borrowers during the study period. These include elevated rates of unemployment that persisted for several years following the GFC and a tightening of mortgage lending standards that reduced the availability of credit to property owners. Mortgage lending standards in the U.S. underwent a gradual loosening during the early 2000s leading up to the crisis, followed by a sharp tightening during the 2007-2009 period that led to increases in loan denial rates and more stringent LTV and DTI requirements (Vojtech et al., 2020). Of the seven flood events evaluated in this study, the effects of the GFC would be most relevant for Hurricane Irene, which occurred in 2011 when the economy was still recovering from the crisis. If the elevated rate of unemployment and reduced credit supply during this period were incorporated into our model, projections of credit constraints among borrowers exposed to flooding from Hurricane Irene would likely be higher.

References

Agarwal, S. and Ho, C. T.: Comparing the Prime and Subprime Mortgage Markets, Federal Reserve Bank of Chicago, 2007.

GAO: Prolonged Conservatorships of Fannie Mae and Freddie Mac Prompt Need for Reform, Government Accountability Office, Washington, D.C., 2019.

Haixiang, G., Yijing, L., Shang, J., Mingyun, G., Yuanyue, H., and Bing, G.: Learning from class-imbalanced data: Review of methods and applications, Expert Systems with Applications, 73, 220–239, https://doi.org/10.1016/j.eswa.2016.12.035, 2017.

He, J. and Cheng, M. X.: Weighting Methods for Rare Event Identification From Imbalanced Datasets, Front Big Data, 4, 715320, https://doi.org/10.3389/fdata.2021.715320, 2021.

Jones, K.: FHA-Insured Home Loans: An Overview, Congressional Research Service, 2022.

N.C. Department of Military and Veterans Affairs: Military Bases in North Carolina, https://www.milvets.nc.gov/benefits-services/military-bases-north-carolina#Fort_Bragg, 2025.

Nolte, C., Boyle, K. J., Chaudhry, A. M., Clapp, C., Guignet, D., Hennighausen, H., Kushner, I., Liao, Y., Mamun, S., Pollack, A., Richardson, J., Sundquist, S., Swedberg, K., and Uhl, J. H.: Data Practices for Studying the Impacts of Environmental Amenities and Hazards with Nationwide Property Data, Land Economics, 100, 200–221, https://doi.org/10.3368/le.100.1.102122-0090R, 2024.

Perl, L.: VA Housing: Guaranteed Loans, Direct Loans, and Specially Adapted Housing Grants, Congressional Research Service, 2018.

Pollack, A. B. and Kaufmann, R. K.: Increasing storm risk, structural defense, and house prices in the Florida Keys, Ecological Economics, 194, 107350, https://doi.org/10.1016/j.ecolecon.2022.107350, 2022.

Vojtech, C. M., Kay, B. S., and Driscoll, J. C.: The real consequences of bank mortgage lending standards, Journal of Financial Intermediation, 44, 100846, https://doi.org/10.1016/j.jfi.2019.100846, 2020.