Response to Reviewer Comments

Jack W. Baker

March 8, 2025

This file illustrates good practices in responding to reviewer comments. It is based on a real response, but has been shortened and edited for clarity. Some commentary is provided in right-margin notes.

We thank the reviewers for their helpful feedback, which zeroed in on some of the tricky issues in this topic, and noted some important areas where we could clarify the manuscript.

Some preamble text to explain the general approach to some themes in the review comments.

As a general comment, we note that the reviewers raise several questions about the broader FEMA P-58 fragility functions, in addition to the new procedure proposed here. These are reasonable questions, though fully considering them here would put us in the difficult position of proposing deviations from a consensus document that has been developed over 15 years of deliberations. As our stated scope is to improve the conceptual fragility procedure to incorporate damage dependencies, we feel that the broader reconsideration of the FEMA P-58 fragility function parameters is reasonably considered as out of scope. That said, several of the review questions are reasonable so we have added additional discussion to the manuscript to acknowledge these potential issues and point to paths forward where possible.

The specific reviewer comments have been listed in order below, followed by our responses in italic text and quotations from the manuscript in grey boxes.

Reviewer 1

Q 1.1 In a parenthesis in Line 433 the authors vaguely allude to the fact that one may simply give correlation coefficients in order to impose statistical dependence between the damage in different components. I am not necessarily advocating for that approach, although it is quite transparent, but I wonder if it should be acknowledged as an alternative, if only in a passage in the paper? Suitable probability transformations exist for the purpose.

Thank you for this suggestion. We have added the following text at line 276 to explain this issue:

Note the very short response to indicate we agree, and the focus on what changed in the text.

> Third, dependence in component capacities does not need to be specified via the above decomposition of ϵ into multiple contributing sources. Instead, dependence could be introduced by sampling capacity terms for each component from

> a multivariate distribution and specifying a covariance matrix for that distribution. It can be shown that the above decomposition weights can be converted into equivalent covariance values. However, the authors and project reviewers found the above weighting approach more intuitive for specifying judgment-based dependencies.

Q 1.2 In this manuscript, the authors' pedagogical instincts are admirable. That effort is also reflected in Figure 2, but I have a question about the labels on the abscissa and ordinate axis of Figure 2b. Perhaps remove the ordinate label, keep the Demand (D) label, and introduce a Capacity (C) label on the abscissa axis of the capacity PDF insert? Also, I am not sure that the vertical blue dashed lines should extend below the capacity PDF. The main point, I suppose, is to link the gray-shaded area in the PDF, i.e., p*, to the ordinate value, p*, in Figure 2a.

Thank you for these suggestions. We have revised figure to to update the labels as suggested, and to rearrange some figure elements to improve clarity.

An occasional 'thank you' can keep the tone positive, especially when the reviewer provides a helpful pointer.

Q 1.3 The authors acknowledge in several instances, including in Line 430, that it is still challenging to know exactly what level of dependence to introduce. That is fair, and does not retract from the value of the paper. However, any need for specifying correlation, or statistical dependence more generally, is a sign that our physical engineering models can be improved. Would there be an opening somewhere to suggest what modeling efforts would be needed in order to let the physics lead to the manifestation of dependence? No need to go far here, because this comment is easily interpreted as a criticism of the present use of fragility functions.

This is a fair point, and we agree that some additional discussion could be valuable. We have expanded a previously short remark in the Conclusions to a full paragraph on this topic (line 454):

If you agree with the comment, say it, so the reviewer can quickly tell whether they are reading a rebuttal or are reading and edit to confirm that it addresses the point.

> While a parameterization has been proposed here based on current judgement, further study of damage data offers the opportunity to refine the parameterization. One data analysis approach that would be informative would be to perform component damage tests that hold some conditions fixed (e.g., construction method) and vary another (e.g., loading protocol), in order to quantify how much component capacity variability comes from each source of uncertainty. Another approach would be to use random effects models to study data sets like that in Figure 1 and attribute observed variability to the tests' varying component configurations, loading conditions, and other factors. From a modeling perspective, some components' damage could be studied by using high-fidelity numerical models (e.g., for steel connection fractures) and varying component and material characteristics, as well as loading time series.

Q 1.4 Line 70: "components" should be "component"

Thank you–we have fixed this typo.

Reviewer 2

Q 2.1 Equation 12: Does this indicate that 20% of the simulations will have full dependent components, 60% of the simulations will have system dependent components and the remaining 20% of realizations will have fully independent components? If yes, the authors should mention this here for the readers understanding. If no, the readers would benefit from further explanation of how these partial dependencies are incorporated in the analysis.

Thank you for this comment. To clarify, we have added a sentence to the discussion of Equation 12:

> This model assumes that some of the uncertainty is shared by all components (20%), and some is shared by all components of the same system (60%), with the remainder unique to each component.

Q 2.2 Figure 4: The authors should explain what 'Type 1' and 'Type2' components represent in Figures 4-b/d/f

We have revised the text on line 306 to be more clearer about this notation:

Reference a line number whenever possible, so that the reviewer can quickly locate the passage in the manuscript.

> Figure 4b shows a scatter plot of how many components of "type 1" and "type 2" are damaged in each simulation (where the numbering is used to distinguish between types, but otherwise has no meaning as all five component types have the same fragility parameters).

Q 2.3 Line 311: This statement indicates that a large ϵ_{all} will cause all the components in the building to have higher capacity. Is this correct? The authors should consider rephrasing to avoid confusion.

Yes, that is correct–a particular simulation having a large ϵ_{all} will tend to cause all components to have a larger capacity, because that term appears in the capacity equation (Equation 5). Since that is the intended message, we assume the statement does not need rephrasing.

If you don't change anything, it can be helpful to explain why.

Reviewer 3

Q 3.1 The presented procedure assumes that the component fragility functions in FEMA P-58 already capture all relevant sources of uncertainty discussed in the manuscript (see Equation 7). Several uncertainty sources are listed in the introduction to illustrate why a portion of the variance in capacities shall be shared among components. For example, uncertainty due to using a scalar EDP metric to represent a complex time history of multiple demands is suggested to be shared among all components. Capacity differences due to varying quality of work among different contractors are shared across components in the same system since the same contractor installs them. The variation in performance of various components considered under the same archetype presents variability shared across the specific components installed in the building. These are all reasonable arguments, and the above sources contribute to the overall uncertainty in component vulnerabilities. However, following Equation 7 and the parameters in Equation 12, the authors assume that a substantial part of the variance characterized in the existing FEMA P-58 component database stems from the above sources. For this assumption to hold, component fragility curves must be calibrated to experimental or empirical data that faithfully captures all the above phenomena. Such calibration requires dynamic tests with a representative set of demand time histories applied on a wide range of different specimens that fall into the group modeled by the component archetype and installed by multiple contractors that represent the range of installation quality expected in construction. This type of empirical data is not available for most fragility functions used in FEMA P-58 assessments. Often quasi-static tests are used to observe the damage at certain levels of an engineering demand parameter which does not allow the fragility function to capture the uncertainty due to differences in demand time histories. Similarly, when the same contractor installs all specimens in an experiment, the uncertainty referred to as β_{sys} in the manuscript is not included in the calibrated fragility function. Rather than disaggregating the variance of existing vulnerability models in FEMA P-58, it seems more appropriate to use the existing variance as β_j in Equation 6 and propose additional beta parameters (or weights) that quantify the other sources of uncertainty shared within systems (β_{sys}) and across all components (β_{all}). Such an approach increases the uncertainty in the results compared to the existing FEMA P-58, but that is realistic because β_{sys} and β_{all} are not considered in the current methodology. Such an approach would also incentivize future work to measure these uncertainties and experimental work that clearly defines which sources of uncertainty are included in the experimental data.

This is an interesting comment, and we partially agree with it. Many of the data sets used for fragility calibration do at least partially reflect the β_{sys} and β_{all} effects, and when judgement is used to supplement the data it is often to consider the effect of conditions other than those seen in tests. For example, Miranda and Mosqueda (2011) provide the following description of the data shown in Figure 1:

"Fragility functions developed herein are based entirely on experimental results of racking tests. These tests are primarily static racking tests although results from a few dynamic tests are also included. [The report also goes on to describe differing static loading protocols that were used in various tests, which can partially represent the effect of different loading time histories associated with a single peak displacement.] Most specimens are 8 ft by 8 ft or

Sometimes you will get a really long comment with multiple arguments included, or a comment that you partially but don't fully agree with, and you need to explain why you believe your revisions are sufficient. In such cases, a longer response may be helpful. But try to make this the exception rather than the rule.

11.5 ft by 12 ft, however some of the experimental tests include specimens with return walls and one investigation included two full room specimens with corners and other conditions commonly found in buildings. Many of the specimens include doors... Data from a total of 74 experimental tests from six different research investigations spanning over 40 years were considered for developing the fragility functions."

On the other hand, the variation in conditions across tests is not fully representative of the variation in conditions in real buildings (and it could not be, since there is no central administrator of the profession's test results).

To raise this issue, and provide some comments to this effect, we have added a paragraph starting at line 461 in the Conclusions:

> While a parameterization has been proposed here based on current judgement, further study of damage data offers the opportunity to refine the parameterization. One data analysis approach that would be informative would be to perform component damage tests that hold some conditions fixed (e.g., construction method) and vary another (e.g., loading protocol), in order to quantify how much component capacity variability comes from each source of uncertainty. Another approach would be to use random effects models to study data sets like that in Figure 1 and attribute observed variability to the tests' varying component configurations, loading conditions, and other factors. From a modeling perspective, some components' damage could be studied by using high-fidelity numerical models (e.g., for steel connection fractures) and varying component and material characteristics, as well as loading time series.