Example Introductions

The following examples, lightly edited for clarity of illustration, include notes indicating the topic of each paragraph. Both introductions contain the same key topics in the same order, and neither mixes topics within paragraphs. The second is longer than the first, but follows the same structure. Add labels like this to your own draft to check that you are staying on topic.

Example 1

Baker J.W. (2007). "Quantitative classification of near-fault ground motions using wavelet analysis," *Bulletin of the Seismological Society of America*, 97 (5), 1486–1501.

[Establish the territory—pulse-like ground motions are important] Near-fault ground motions containing strong velocity pulses are of interest in the fields of seismology and earthquake engineering. A quantitative approach for identifying these ground motions is proposed here and used to perform a variety of basic studies of their properties. These ground motions, which are here referred to as "pulse-like ground motions," have been identified as imposing extreme demands on structures to an extent not predicted by typical measures such as response spectra [*e.g.*, 1–10]. Theoretical considerations also provide an indication of seismological conditions that may result in occurrence of velocity pulses due to, for example, directivity effects [11–13]. While the effect is relatively well studied, a hindrance to incorporating these effects in probabilistic seismic hazard analysis and engineering building codes is that a quantitative method for identifying these velocity pulses does not yet exist. This means that a variety of researchers have assembled sets of pulse-like or near-fault ground motions, but these classifications are not easily reproducible [*e.g.*, 14–17].

[Establish the niche—current pulse-like classifications depend on judgment] The ground motions identified in past studies are typically selected because the velocity time history of the ground motion is dominated by a large pulse, as seen, for example, in Figure 1a, and/or because source–site geometry suggests that a directivity pulse might be likely to occur at the site where the motion was recorded. Selection of pulse-like ground motions using these approaches requires some level of judgment, and for many ground motions, such as those shown in Figure 1b and c, the classification may not be obvious. Identification of non-pulse-like motions at near-fault locations is also challenging for the same reasons, although it has not received as much attention.

[Establish the niche—lack of classification causes problems] The lack of a quantitative classification scheme for recorded ground motions has hindered progress toward obtaining results such as the probability that a ground motion with a given earthquake magnitude, distance, and source–site geometry will contain a velocity pulse. Knowledge of this probability is useful for applications such as probabilistic seismic hazard analysis [18,19]. The lack of quantitative classifications also means that electronic libraries of recorded ground motions do not list any statistics indicating whether a given ground motion contains a velocity pulse, and this limits the ability of the science and engineering communities to access these ground motions and study their effects for research or practical applications.

[Occupy the niche] In this article, an approach for detecting pulses in ground motions is proposed and investigated. The procedure uses wavelet-based signal processing to identify and extract the largest velocity pulse from a ground motion; if the extracted pulse is "large" relative to the remaining features in the ground motion, the ground motion is classified as pulse-like. The period of a detected velocity pulse, a parameter of interest to engineers, is also easily determined. The classification algorithm is computationally inexpensive, so large libraries of recorded ground motions can be (and have been) analyzed. Although some of the identified pulses are likely not caused by directivity effects, the approach is useful for identifying a set of records potentially exhibiting directivity effects, which can then be manually considered more carefully. Alternatively, the ground motions could be used (without further classification) for structural response calculations, under the assumption that pulses will cause similar effects regardless of their causal mechanism.

Example 2

Baker, J.W., and Chen, Y. (2020). "Ground motion spatial correlation fitting methods and estimation uncertainty." *Earthquake Engineering & Structural Dynamics*, 49 (15), 1662–1681.

[Establish the territory—spatial correlations in ground motions are important] Spatial correlations in ground motion intensities have been studied for nearly two decades [*e.g.*, 1-7]. These studies calibrate models using statistical analysis of recorded data from past earthquakes, in a manner similar to ground motion prediction models. The resulting models are important for estimating risks to distributed systems such as portfolios of insured properties and distributed infrastructure systems [*e.g.*, 8-12].

[Establish the niche—current studies draw differing conclusions about causes of correlations] As our library of recorded ground motions grows over time, spatial correlation studies have grown in refinement, with several studies exploring factors that may cause spatial correlation to vary from one earthquake to another. Goda and Hong [13] report differences in correlations between California and Taiwan ground motions, but no effect of earthquake magnitude. Jayaram and Baker [14] and Sokolov et al. [15] speculate that soil condition heterogeneity may influence spatial correlation. Goda [5] and Heresi and Miranda [16] report that spatial correlations vary from individual earthquake to earthquake, but find no earthquake characteristic that clearly predicts this variation. Schiappapietra and Douglas [17] report high variability in correlations amongst a sequence of earthquakes in Central Italy, and list local site effects or path and azimuthal effects as possible causes, noting Sokolov et al.'s [15] similar speculation. Other studies note a possible trend with earthquake magnitude [18,19] or variation regionally [20]. Other studies group data from multiple earthquakes into a single data set, making the assumption that correlations from the earthquakes are equivalent [4,7,6,21,22].

[Establish the niche—we have limited data and no way to quantify uncertainty in results] Several issues make it difficult to definitively identify important factors. First, because the spatial correlation estimation is empirical and requires many observed ground motions from an earthquake, it is difficult to obtain sufficient data under the conditions of interest. Second, there are no closed-form results from spatial correlation estimation that allow for a quantitative assessment of the uncertainty in a given estimate. Bootstrap estimation is another popular technique to quantify estimation uncertainty. However, it is difficult to apply to spatial data because of challenges with maintaining spatial dependence structure in the replicates while avoiding resampling the same location within a replicate (which provides no information about spatial structure). Some studies have proposed bootstrap techniques that resample from an estimated nonparametric distribution [23] or resampling transformed data [24], but the methods are somewhat complex and have not been adopted widely. For the above reasons, no results have been presented in previous studies to quantify the estimation uncertainty in spatial correlations computed from individual earthquakes.

[Establish the niche—model fitting techniques differ in previous studies] Another issue that has not been evaluated in this literature is the role of the method used to fit parameters for the models, and relative performance of alternative methods. Fitting methods used in prior ground motion studies include manual visual fitting [14], least squares regression on transformed data [16,7], and least squares regression with weighting according to distance or number of data [25]. One general study evaluated several fitting methods using Monte Carlo simulation of spatial data with a known correlation structure [26], and found that the above fitting methods produce systematic differences in results. But that study considered a small number of observed data (16 or 36 stations) on a regular grid—a situation very different than ground motion data coming from greater numbers of irregularly spaced stations. A recent study examined this issue for realistic numbers of stations in earthquake ground motion studies, but the locations in that study were randomly simulated [20]. There are no general statistical results for estimators under arbitrary station configurations [27].

[Establish the niche—unknown relationship between data sample size and estimation uncertainty] Exact estimation of a correlation model from data (*i.e.*, consistency in statistical estimation language) requires a large number of observations at closely spaced distances, with the dense locations not too concentrated at a single location [28]. The above-cited studies of ground motions consider well-recorded earthquakes, but none systematically study the impact of well-recorded versus more-poorly-recorded earthquakes on resulting spatial correlation estimates. Further, the above-cited studies use different methods to fit models to data, and it is unknown what impact those fitting methods have on estimation uncertainty. The baseline estimation uncertainty is important as it establishes a threshold at which variability in observed correlations can be credibly linked to some causal source rather than being due to estimation variability.

[Occupy the niche] To address the above issues, this paper proposes a framework to quantify uncertainty in spatial correlation models, and uses the framework to evaluate estimation uncertainty associated with individual earthquakes and model fitting methods. Section 2 introduces the basic framework for characterizing ground motion amplitudes using ground motion models, and introduces the semivariogram as a tool for quantifying spatial correlations. Methods for fitting semivariogram models are also introduced. Section 3 then introduces a model to describe the various components of apparent uncertainty in semivariogram parameters. A method is proposed for quantifying estimation uncertainty, by synthetically simulating ground motion amplitudes with a known spatial correlation model but observed only at locations corresponding to those of past earthquakes. This method is then applied to the considered earthquakes and fitting methods. The results are then discussed, along with the limitations and broader implications of the work.