Dynamic Post-Earthquake Updating of Regional Damage Estimates Using Gaussian Processes

Lukas Bodenmann[a,*], Yves Reuland[a], Božidar Stojadinović[a]

[a] Dept. of Civil, Environmental and Geomatic Engineering, ETH Zurich, Stefano-Franscini-Platz 5, Zürich, 8093, Switzerland

Abstract

The widespread earthquake damage to the built environment induces severe short- and long-term societal consequences. Better community resilience may be achieved through well-organized recovery. Decisions to organize the recovery process are taken under intense time pressure using limited, and potentially inaccurate, data on the severity and the spatial distribution of building damage. We propose to use Gaussian Process inference models to fuse the available inspection data with a pre-existing earthquake risk model to dynamically update regional post-earthquake damage estimates and thereby support a well-organized recovery. The proposed method consistently aggregates the gradually incoming building damage inspection data to reduce the uncertainty in ground shaking intensity geographic distribution and to update regional building damage estimates. The performance of the proposed Gaussian Process methodology is demonstrated on one fictitious earthquake scenario and two real earthquake damage datasets. A comparison with purely data-driven methods shows that the proposed method reduces the number of building inspections required to provide reliable and precise damage predictions.

Keywords: Post-earthquake damage assessment, Gaussian Process models, Regional earthquake risk models, Uncertainty reduction

1. Introduction

Damaging earthquakes carry the potential of catastrophic consequences. In addition to direct financial losses, injuries, and fatalities caused by damages to the built environment and short-term disruption in societal functions, earthquakes may give rise to negative long-term consequences, such as increases in home purchase prices and scarcity of low-cost rentals [1]. Housing is an essential infrastructure system of a community, one that provides shelter for its inhabitants, enables higher societal functions and, as shown by the Covid-19 pandemic, provides a safe place to work remotely. The severity and duration of disruptions to such an infrastructure system are not only affected by the seismic resistance of buildings, but also by the efficiency and speed of public and private stakeholder response. Information on the regional severity of the inflicted seismic damage plays a central role in organising the emergency response and recovery efforts [2]. However, such information is scarce, incomplete, and imprecise in the early aftermath of an earthquake. Simultaneously, under

*Corresponding author

intense time pressure, stakeholders and agencies need to take crucial decisions that affect immediate disaster assistance as well as the long-term recovery. Post-earthquake inspection of buildings in the affected region, in addition to enabling safe building re-occupancy, leads to a systematic collection of useful building-specific data [3, 4]. However, visual inspection of all buildings within the affected region is time- and resource-intensive and may take several weeks [5, 6]. Therefore, decision makers rely on damage estimates deduced from incomplete and uncertain data.

Earthquake risk models may provide initial, quantitative, and spatially exhaustive damage estimates. Such risk models aim to quantify the probability of earthquake-induced consequences on the built environment within a region and, in a broader sense, on its users and inhabitants. While this study focuses on the use of risk models for rapid post-earthquake damage estimation [7, 8], risk models are also widely employed in pre-event settings to quantify long-term seismic risk and to conduct scenario-based analyses [9, 10]. Risk models require information on the exposed buildings, in particular on their seismic vulnerability, together with estimates of the spatial distribution of ground motion intensity. In post-event applications of risk models, the latter can be constrained using ground motion recordings [11]. The precision and accuracy of the resulting damage estimates depends, amongst other, on the density of the seismic network, the level of detail of the available exposure information and, finally, how well the seismic vulnerability estimates reflect local conditions of the built inventory. While model-based damage estimates are particularly helpful in the immediate aftermath of an earthquake, when damage observations are not available yet, the inherent multiple sources of uncertainty underline the need for a methodological framework to confirm or correct early damage estimates using empirical evidence, once available.

Permanent seismic instrumentation and monitoring of some buildings may help to detect and quantify earthquake-induced damage. This near-real-time data source has gained increasing attention at building [12, 13, 14] and regional levels [15]. Yet, translating data-driven damage indicators into robust decision-making remains an open research topic. In addition, extensive regional instrumentation of buildings is not realistic in the short-to-medium term and therefore, robust methodologies to link damage of single buildings with regional impact assessment, such as proposed in this paper, will continue to be of interest. Remote-sensing data, such as satellite, aerial, and drone imagery, comprise another data source. However, most automated image-based applications either focus on binary assessment of building collapse, or involve damage proxies that are difficult to link with the actual building damage [16, 17, 18]. In addition, remotely taken images are limited to the building exterior and suffer from occlusion, complex facade ornaments, and changing shadow patterns [19].

Expert-conducted building inspection in the aftermath of an earthquake guarantees a continuous, yet slow, data inflow. Such inspections qualify the induced damage to individual buildings and, as a byproduct, provide important building information that is typically missing in an exposure database (e.g., on the actual lateral load-resisting system). Given this data inflow, machine-learning and statistical inference tools may be leveraged to estimate damage sustained by buildings that have not been inspected yet. Kovačević et al. [20] have applied random forests (RFs) to predict the damage state for individual buildings due to the 2010 Kraljevo (Serbia) earthquake. Stojadinovic et al. [21] have extended the use of RFs and formulated an operational methodology for rapid regional repair cost estimation that requires careful prioritization of inspections. Using inspection data from the 2015 Nepal earthquake, Loos et al. [22] employed geo-statistical techniques to predict the mean damage state aggregated to equally spaced grid cells and, assuming normally-distributed mean damage states, produced variance-based uncertainty estimates for their predictions. However, the proposed method only applies to continuous data, requiring aggregation and averaging of categorical inspection results over grid cells. Apart from the risk of information loss, this also imposes operational constraints on the inspection process, because all buildings in a grid cell have to be inspected before the corresponding results can be used. Sheibani and Ou [23, 24] explored Gaussian Process (GP) models by focusing on continuous quantities that either require sensors within buildings, such as peak floor acceleration, or become available much later in the recovery process (such as repair costs).

We propose to use GP models to fuse the available inspection data with a pre-existing earthquake risk model in order to dynamically update regional post-earthquake damage estimates. The proposed risk model informed GP (RMGP) framework allows processing of observed damage, in the form of multiple ordinal categories (i.e., damage states), to reduce the uncertainty of the geographical distribution of the ground shaking intensity, and to simultaneously update the regional building damage estimates. Thus, instead of learning a new, entirely data-driven, model after the event, individual risk model components are updated using different GPs. This allows for increasingly constrained predictions of building damage that can be consistently aggregated at different spatial scales. The present work is based on theoretical studies of (approximate) Bayesian updating schemes for infrastructure systems focusing on the failure probability of individual components [25, 26, 27] and overall system performance [28].

The structure of this paper is as follows: First, we present the mathematical background of GP models in a general context (Section 2), followed by an introduction to earthquake risk models in Section 3. Then, Section 4 describes the risk model inference and updating process using post-earthquake visual damage inspection data of the proposed RMGP framework. Finally, in Section 5, the RMGP framework is applied to the real data from the 1998 Pollino (Italy) and the 2010 Kraljevo (Serbia) earthquakes, as well as to the simulated data for a fictitious event in Zurich (Switzerland).

2. Statistical Learning with Gaussian Process Models

A Gaussian Process model is a versatile tool to regress and classify data under the Bayesian paradigm [29, 30]. GP models assume a directional dependency between a d-dimensional input vector (covariate) $\mathbf{x}$ from some domain $\mathcal{X}$ and the corresponding observable scalar output (response) y. Assuming this dependency is separable into a systematic and a random component, the systematic dependency is given by a latent function $f : \mathcal{X} \to \mathbb{R}$, such that the likelihood of the output takes the form $p(y|f(\mathbf{x}), \boldsymbol{\vartheta})$. We use $\boldsymbol{\vartheta}$ to denote additional parameters of this likelihood. The objective of GP regression and classification is to produce probabilistic predictions of outputs y_* at the desired target inputs $\mathbf{x}_*$, via inferring knowledge about the function f from training data and prior beliefs.

Consider a training set of empirical pairwise observations $\mathcal{D} = \{(\mathbf{x}_i, y_i)|i = 1, ..., m\}$, where the vector $\mathbf{y} = [y_1, ..., y_m]^\top$ and the matrix $\mathbf{X} = [\mathbf{x}_1, ..., \mathbf{x}_m]^\top$ collect the outputs and inputs, respectively. Besides the training data, GP inference about f also involves formalizing the prior belief about the latent function. Instead of assuming a fixed parametric form of f with a finite number of parameters $\boldsymbol{\psi}$ and making inference about $\boldsymbol{\psi}$ only[1], a GP serves as a prior on the latent function

$$f \sim \mathcal{GP}\left(m(\mathbf{x}, \boldsymbol{\theta}_m), k(\mathbf{x}, \mathbf{x}', \boldsymbol{\theta}_k)\right) \ , \tag{1}$$

[1] An example is the linear model $f(\mathbf{x}, \boldsymbol{\psi}) = \mathbf{x}^\top \boldsymbol{\psi}$, where prior uncertainty about f is usually expressed in terms of a prior distribution on $\boldsymbol{\psi}$.

where $m(\cdot)$ and $k(\cdot)$ denote the mean and positive definite covariance function with parameters $\boldsymbol{\theta}_m$ and $\boldsymbol{\theta}_k$, respectively. While we present mean and covariance functions that are used for regional earthquake risk modeling in Section 3, the reader is referred to [29] for a comprehensive discussion.

Technically, a GP is a collection of random variables, any finite subset of which follows a joint multivariate normal distribution. A GP prior on f means that each input vector $\mathbf{x} \in \mathcal{X}$ has an associated random variable $f(\mathbf{x})$, and that the prior joint distribution $p(\mathbf{f}|\mathbf{X}, \boldsymbol{\theta})$ for a collection of function values $\mathbf{f} = [f(\mathbf{x}_1), ..., f(\mathbf{x}_n)]^\top$ associated with an arbitrary set of n inputs is multivariate normal. The mean vector $\mathbf{m} \in \mathbb{R}^{n\times 1}$ of this distribution has entries $[\mathbf{m}]_i = m(\mathbf{x}_i, \boldsymbol{\theta}_m)$ and the covariance matrix $\mathbf{K} \in \mathbb{R}^{n\times n}$ has entries $[\mathbf{K}]_{ij} = k(\mathbf{x}_i, \mathbf{x}_j, \boldsymbol{\theta}_k)$.

For the inference process, the posterior distribution of the function values $\mathbf{f}$ is computed through Bayes' rule and under the assumption of a factorizing likelihood:

$$p(\mathbf{f}|\mathcal{D}, \boldsymbol{\theta}, \boldsymbol{\vartheta}) = \frac{p(\mathbf{y}|\mathbf{f}, \boldsymbol{\vartheta})p(\mathbf{f}|\mathbf{X}, \boldsymbol{\theta})}{p(\mathcal{D}|\boldsymbol{\theta}, \boldsymbol{\vartheta})} = \frac{p(\mathbf{f}|\mathbf{X}, \boldsymbol{\theta})}{p(\mathcal{D}|\boldsymbol{\theta}, \boldsymbol{\vartheta})} \prod_{i=1}^{m} p(y_i|f(\mathbf{x}_i), \boldsymbol{\vartheta}) \ , \tag{2}$$

where $p(\mathcal{D}|\boldsymbol{\theta}, \boldsymbol{\vartheta}) = \int p(\mathbf{y}|\mathbf{f}, \boldsymbol{\vartheta})p(\mathbf{f}|\boldsymbol{\theta})\mathrm{d}\mathbf{f}$ is the marginal likelihood or model evidence. Then, the posterior predictive distribution of function values $\mathbf{f}_*$ at desired target inputs $\mathbf{X}_*$ is derived by marginalizing the function variables over the available training inputs:

$$p(\mathbf{f}_*|\mathcal{D}, \boldsymbol{\theta}, \boldsymbol{\vartheta}) = \int p(\mathbf{f}_*|\mathbf{f}, \boldsymbol{\theta})p(\mathbf{f}|\mathcal{D}, \boldsymbol{\theta}, \boldsymbol{\vartheta})\mathrm{d}\mathbf{f} \ . \tag{3}$$

The probabilistic predictions of the target outputs are derived by taking the expectation $p(\mathbf{y}_*|\mathcal{D}, \boldsymbol{\theta}, \boldsymbol{\vartheta}) = \int p(\mathbf{y}_*|\mathbf{f}_*, \boldsymbol{\vartheta})p(\mathbf{f}_*|\mathcal{D}, \boldsymbol{\theta}, \boldsymbol{\vartheta})\mathrm{d}\mathbf{f}_*$. The analytical tractability of posterior distributions depends on the type of the output quantity of interest and assumptions about the likelihood $p(y|f(\mathbf{x}), \boldsymbol{\vartheta})$. The following Section 2.1 serves as a an introductory example where we use a Gaussian likelihood. Section 2.2 specifies the likelihood for non-Gaussian cases, i.e., ordinal and nominal categorical data. Those are of special interest for earthquake risk models, because they describe discrete building damage categories and typological building classes (see Section 3). Subsequently, we present approximate inference in cases with such non-Gaussian mappings.

2.1. Introductory Example: Gaussian Likelihood

For continuous outputs $y \in \mathbb{R}$ and under the assumption of independent normally distributed noise, we can assume a Gaussian likelihood $p(y|f(\mathbf{x}), \boldsymbol{\vartheta}) = \mathcal{N}(f(\mathbf{x}), \sigma_0)$. The noise variance σ_0^2 is thus an additional parameter of the likelihood $\boldsymbol{\vartheta} = \sigma_0^2$. Bayesian inference of the latent function f in this model is analytically tractable, while the posterior process $f|\mathcal{D}$ is again a GP. Thus, the posterior of the function values at the input points of the training set $\mathbf{X}$, $p(\mathbf{f}|\mathcal{D})$, as well as the posterior predictive distribution for an arbitrary set of target inputs $\mathbf{X}_*$, $p(\mathbf{f}_*|\mathcal{D})$, follow multivariate normal distributions. By evaluating the covariance matrix between target and training inputs, $\mathbf{K}_*$, and between the target inputs themselves, $\mathbf{K}_{**}$, we can compute the mean and covariance of the posterior predictive as:

$$\boldsymbol{\mu}_* = \mathbf{m}_* + \mathbf{K}_*(\mathbf{K} + \sigma_0^2\mathbf{I})^{-1}(\mathbf{y} - \mathbf{m}) \tag{4a}$$

$$\boldsymbol{\Sigma}_{**} = \mathbf{K}_{**} - \mathbf{K}_*(\mathbf{K} + \sigma_0^2\mathbf{I})^{-1}\mathbf{K}_*^\top \ , \tag{4b}$$

where $\mathbf{I}$ is the identity matrix with the same dimensions as $\mathbf{K}$.

Finally, the predictive distribution of outputs $\mathbf{y}_*$ is also multivariate normal with identical mean, but with the addition of the noise variance σ_0^2 to the diagonal elements of $\mathbf{\Sigma}_{**}$, e.g., $p(\mathbf{y}_*|\mathcal{D}, \boldsymbol{\theta}, \boldsymbol{\vartheta}) = \mathcal{N}(\boldsymbol{\mu}_*, \mathbf{\Sigma}_{**} + \sigma_0^2\mathbf{I})$.

To illustrate GP regression, we use a one-dimensional toy example, where the prior GP is specified using the popular squared exponential covariance function defined as $k(\mathbf{x}, \mathbf{x}', \boldsymbol{\theta}_k) = \sigma^2 \exp(- \parallel \mathbf{x} - \mathbf{x}' \parallel_2^2 /2\ell^2)$. Here σ^2 is the variance, ℓ the length scale, and $\parallel \cdot \parallel_2$ is a Euclidean distance measure. The variance controls the amplitude, while the length scale controls the range of correlation: lower length scale values imply a correlation that decreases faster with the distance between input points. Considering the introductory character of this example, a zero mean function ($m(\mathbf{x}) = 0$) is chosen and thus, the hyper-parameters of the prior GP are limited to $\boldsymbol{\theta} = (\sigma^2, \ell)$.

The dataset $\mathcal{D}$ consists of $n = 40$ datapoints, at which y is sampled from the function $\sin(0.5x)+ 0.5\cos(0.2x)$ polluted with a white noise $\epsilon \sim \mathcal{N}(0, \sqrt{0.1})$. Input positions x are uniformly sampled in two separate intervals. A first model mimics the situation where the hyper-parameters are fixed to 1, 7 and 0.15 for the GP variance σ^2, length scale ℓ, and noise variance σ_0^2, respectively. Figure 1a shows the posterior predictive distribution over the input space $x_* \in [-20, 20]$. The resulting posterior function, whose density is shown as the shaded area, gives a poor fit to the generating function, represented by the solid line. The fit is especially poor in regions far from available data points. Figure 1c plots the posterior predictive in the output space $p(\mathbf{y}_*|\mathcal{D}, \boldsymbol{\theta}, \boldsymbol{\vartheta})$, with the mean function identical to Figure 1a above, but with the variance being increased by the noise σ_0^2.

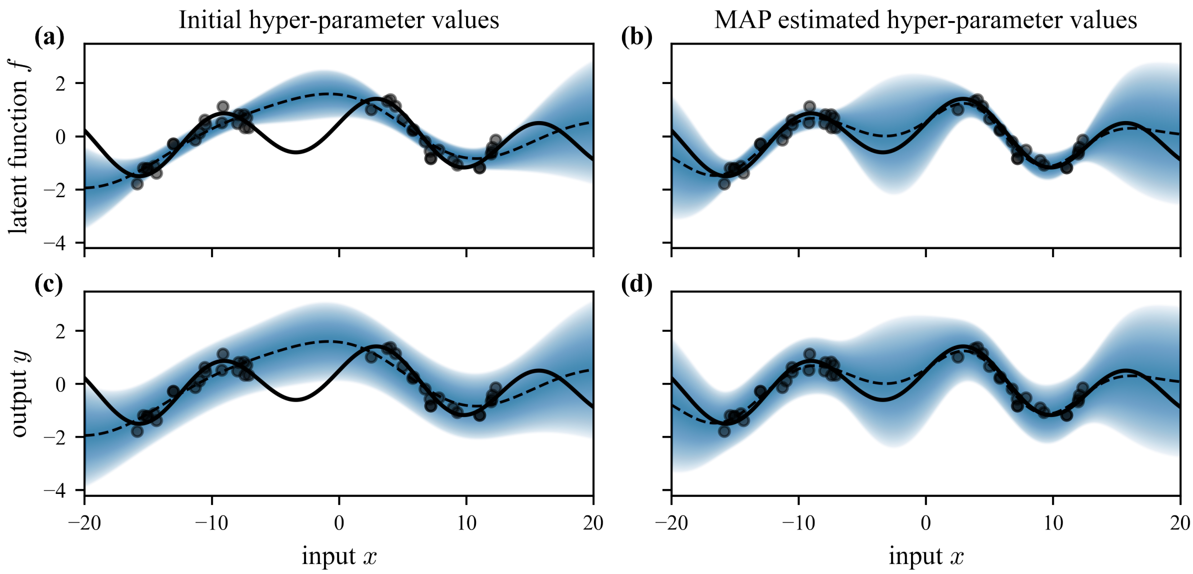

Figure 1: GP regression with independent normal noise. The input data $\mathcal{D}$, indicated using dots, was generated from the function shown as a solid line with additive normal noise. The top row of plots shows the posterior of the latent functions, where the shaded area shows the density $p(f|\mathcal{D}, \boldsymbol{\theta}, \boldsymbol{\vartheta})$ and the dashed lines indicate the corresponding mean. The bottom row of plots shows the posterior in the output space $p(y|\mathcal{D}, \boldsymbol{\theta}, \boldsymbol{\vartheta})$. In the left column the parameters θ and ϑ are fixed, whereas in the right column the parameters are estimated as their MAP values.

The choice of appropriate hyper-parameters $\boldsymbol{\theta}$ and $\boldsymbol{\vartheta}$ prior to the inference process is often challenging, especially in situations with limited prior domain knowledge. In a full Bayesian setting hyper-prior distributions $p(\boldsymbol{\theta})$ and $p(\boldsymbol{\vartheta})$ are assigned and inference is performed over these parameters to update the joint posterior of the latent function and the hyper-parameters simultaneously.

This approach is analytically intractable and thus, requires an approximation via sampling, such as Markov-chain Monte Carlo. In this study, we use the maximum a-posteriori (MAP) estimates of the hyper-parameters, for example, the mode of their posterior distributions $p(\boldsymbol{\theta}, \boldsymbol{\vartheta}|\mathcal{D})$, via solving an optimization problem

$$(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\vartheta}}) = \underset{\boldsymbol{\theta}, \boldsymbol{\vartheta}}{\operatorname{argmax}} \log p(\mathcal{D}|\boldsymbol{\theta}, \boldsymbol{\vartheta}) + \log p(\boldsymbol{\theta}) + \log p(\boldsymbol{\vartheta}) \ , \tag{5}$$

which is analytically tractable for a Gaussian likelihood, and factorizing hyper-prior distributions. In this example we chose log-normal hyper-priors, with the median values identical to the initial values above, and the standard deviations set to 0.8 for ℓ and σ^2 and to 0.2 for σ_0^2.

The posterior function derived using the MAP-estimated hyper-parameters provides a better fit to the true data generating function (see Figure 1b and d). Such improved fit results from the reduction of the estimated length scale to a value of 3 from the initially set value of 7. The smaller length scale ℓ reduces the overconfidence in regions away from known data points, evidenced by the larger spread in the posterior density of the function f.

GP regression, as presented above, is also widely applied in the geo-statistics field, where it is known as kriging [31] and naturally focused mostly on two-dimensional input spaces. The method for rapid post-earthquake damage mapping proposed by Loos et al. [22], first uses the non-spatial inputs in least-squares regression to remove a linear trend in the data and secondly uses the geo-coordinates to update GP models with the remaining residuals. The method proposed by Sheibani and Ou [23] also relies on GP regression but applies it to the entire input space and not only on the geo-coordinates. Both methods, however, have in common that they rely on a Gaussian likelihood, which explains why they are not directly applicable to categorical data, such as the damage states assigned during post-earthquake building inspections.

2.2. Non-Gaussian Likelihoods for Categorical Data

If the output quantity of interest is categorical, meaning it only takes values from a finite set of discrete categories, likelihood functions that differ from the Gaussian likelihood used in Section 2.1, are required. Although the inference process loses its analytical tractability for categorical data, GP models also apply in these contexts [32, 33]. The next two paragraphs contain commonly used likelihoods for ordered categorical output data (ordinal GP regression) and nominal categorical outputs (multi-class GP classification). Finally, a method for approximate Bayesian inference in these non-conjugate cases is presented.

Ordinal GP regression. Ordinal data is defined by scalar outputs y_i being elements of a finite set $\mathcal{Y}$ of $c+1$ ordered categories, which are denoted as integers, $y_i \in \mathcal{Y} = \{0, 1, ..., c\}$, with preserved ascending ordering information. According to [33], GP regression can be rewritten for such ordinal data by employing c parameters $\eta_1 < ... < \eta_c$ to define the class membership probabilities of label y_i conditional on (latent) function value $f_i = f(\mathbf{x}_i)$:

$$p(y_i|f_i, \boldsymbol{\vartheta}) = \begin{cases} 1 - \Phi\left[\left(f_i - \eta_{y_i+1}\right)/\beta\right], & \text{if } y_i = 0 \\ \Phi\left[\left(f_i - \eta_{y_i}\right)/\beta\right], & \text{if } y_i = c \\ \Phi\left[\left(f_i - \eta_{y_i}\right)/\beta\right] - \Phi\left[\left(f_i - \eta_{y_i+1}\right)/\beta\right], & \text{otherwise} \end{cases} \tag{6}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function (CDF). The threshold parameters, $\boldsymbol{\eta} = (\eta_1, ..., \eta_c)$, partition the real line into contiguous intervals, mapping the continuous function f_i into the discrete variable y_i. The dispersion of these threshold parameters is denoted as β. The parameters of the likelihood function thus become $\boldsymbol{\vartheta} = (\boldsymbol{\eta}, \beta)$.

Multi-class GP classification. Nominal categorical data, similar to ordinal data, is discrete and finite with possible class outputs $y_i \in \mathcal{C} = \{0, 1, ..., c\}$, where c is the number of classes. Unlike ordinal labels, nominal class labels do not provide any ranking information and, instead of having one function f as is the case for ordinal data, c is independent latent GP functions, i.e., one exists for each class. The function values for data point $\mathbf{x}_i$ are denoted as $\mathbf{g}_i = (g_i^0, ..., g_i^c)$, where we denote functions as g to avoid confusion with preceding sections. The conditional class membership probabilities of label y_i are calculated using the softmax function as [29]:

$$p(y_i|\mathbf{g}_i) = \frac{\exp g_i^{y_i}}{\sum_{k=1}^{c} \exp g_i^k} \ . \tag{7}$$

If for class k we write $\mathbf{g}^k = (g_1^k, ..., g_n^k)$, then $p(\mathbf{g}^k|\mathbf{X})$ follows a multivariate normal distribution. Furthermore, because the c latent functions are independent, so is $p(\mathbf{g}|\mathbf{X})$, where $\mathbf{g} = (\mathbf{g}^1, ..., \mathbf{g}^c)$. In this case, the likelihood function has no specific parameters $\boldsymbol{\vartheta} = \emptyset$.

Variational Gaussian approximation. Because of the non-Gaussian likelihoods defined in Eq. 6 and Eq. 7, the posterior distribution defined in Eq. 2, is not analytically tractable and is thus approximated using the variational Gaussian approximation scheme [34]. The true, but intractable, posterior is approximated with a multivariate normal distribution $q(\mathbf{f})$ that is close to the true posterior $q(\mathbf{f}) = \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \approx p(\mathbf{f}|\mathcal{D}, \boldsymbol{\theta}, \boldsymbol{\vartheta})$. Specifically, we search for parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ that minimize the Kullback-Leibler (KL) divergence between the variational distribution $q(\mathbf{f})$ and the true posterior. We can write this KL divergence as

$$\mathrm{KL}\left(q(\mathbf{f}) \ \| \ p(\mathbf{f}|\mathcal{D}, \boldsymbol{\theta}, \boldsymbol{\vartheta})\right) = \log p(\mathcal{D}|\boldsymbol{\theta}, \boldsymbol{\vartheta}) + \mathcal{F}_{\mathrm{vfe}}(q, \boldsymbol{\theta}, \boldsymbol{\vartheta}) \ , \tag{8}$$

where $\mathcal{F}_{\mathrm{vfe}}(\cdot)$ denotes the variational free energy. Because of the non-Gaussian likelihood, the evidence $p(\mathcal{D}|\boldsymbol{\theta}, \boldsymbol{\vartheta})$ is not analytically tractable undermining a direct evaluation of Eq. 8. However, the evidence being a constant term, minimizing Eq. 8 is equivalent to minimizing the variational free energy. The latter is expressed as

$$\mathcal{F}_{\mathrm{vfe}}(q, \boldsymbol{\theta}, \boldsymbol{\vartheta}) = \mathrm{KL}\left(q(\mathbf{f}) \ \| \ p(\mathbf{f}|\boldsymbol{\theta})\right) - \sum_{i=1}^{n} \left[\int \log(p(y_i|f_i, \boldsymbol{\vartheta})) q(f_i) df_i \right] \ . \tag{9}$$

The first term in Eq. 9 is analytically tractable because it is the KL-divergence between two multivariate normal distributions, the variational distribution and the prior of f. The one-dimensional integrals in the second term, also called the expected log likelihood of the data with respect to $q(\mathbf{f})$, can, for example, be approximated using Gauss-Hermite quadrature. In order to keep the variational free energy small, a model should provide a good explanation of the data (via large values for the second term) while not deviating too far from the prior (by keeping the first term small). Negative variational free energy is also called the evidence lower bound (ELBO), because it provides a lower bound to the log of evidence. This follows from Eq. 8 and the fact that the KL-divergence measure is always positive. Therefore, $\mathcal{F}_{\mathrm{vfe}}(\cdot)$ is used to learn the parameters of the variational distribution $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, while it simultaneously provides approximate MAP estimates for the hyper-parameters $\boldsymbol{\theta}$ and $\boldsymbol{\vartheta}$, using the following objective function:

$$(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\vartheta}}) = \underset{\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\theta}, \boldsymbol{\vartheta}}{\operatorname{argmax}} \ -\mathcal{F}_{\mathrm{vfe}}(q, \boldsymbol{\theta}, \boldsymbol{\vartheta}) + \log p(\boldsymbol{\theta}) + \log p(\boldsymbol{\vartheta}) \ . \tag{10}$$

The application of this objective function to derive MAP estimates for the parameters $\boldsymbol{\vartheta}$ of the ordinal likelihood require a two-fold re-parameterization: First, to enable a parallel estimation of the threshold parameters and their dispersion, we perform an up-front scaling $\tilde{\boldsymbol{\eta}} = \boldsymbol{\eta}/\beta$. Second, to fix the order of the threshold parameters, positive parameters $\Delta_y = \tilde{\eta}_y - \tilde{\eta}_{y-1}$ are introduced for categories $y > 0$. Thus, a MAP estimation is performed for parameters $\boldsymbol{\vartheta} = \{\beta, \tilde{\eta}_0, \Delta_1, ..., \Delta_c\}$.

Because the approximated posterior distribution, $q(\mathbf{f})$, is multivariate normal, the approximate posterior predictive $q(\mathbf{f}_*|\hat{\boldsymbol{\theta}}) = \int p(\mathbf{f}_*|\mathbf{f}, \hat{\boldsymbol{\theta}})q(\mathbf{f})\mathrm{d}\mathbf{f}$ is also multivariate normal. Consequently, we can evaluate its parameters as

$$\boldsymbol{\mu}_* = \mathbf{m}_* + \mathbf{A}\boldsymbol{\mu} \tag{11a}$$

$$\boldsymbol{\Sigma}_{**} = \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^\top + \mathbf{B} \ , \tag{11b}$$

where $\mathbf{A} = \mathbf{K}_*\mathbf{K}^{-1}$ and $\mathbf{B} = \mathbf{K}_{**} - \mathbf{K}_*\mathbf{K}^{-1}\mathbf{K}_*^\top$. In the following section we link the outlined GP theory to regional earthquake risk models and we specify well-suited mean and covariance functions of the different GPs described above.

3. Regional Earthquake Risk Models

Regional risk models, when employed in a post-earthquake context, provide rapid, yet uncertain, damage predictions at regional scale because their components contain uncertainties. The geographical distribution of shaking intensity measures remains uncertain despite seismic network station measurements. Detailed information about the buildings in the affected region is often not available, making modeling and earthquake response simulation for each buildings within a region practically impossible. Thus, buildings in the region are clustered into several predefined types per their seismic vulnerability. Given the lack of data and knowledge, such average vulnerability often represents buildings at a national, if not continental, scale [35]. Finally, building type attribution models, or exposure mapping models, correlate building types with socio-econonomic indicators and building attributes from public databases, such as building height, value and year of construction [36, 37], in order to overcome the lack of building-specific information. Such typological attribution models may have limited applicability in the region hit by an earthquake, further adding to the uncertainties arising from the shaking intensity and vulnerability risk model components.

In this study, we assume existence of public databases to obtain the basic building-specific information. For a building, this information is gathered in a d-dimensional vector $\mathbf{x}$, whose individual entries are, for example, the geographic coordinates, the construction year, and the number of stories. Building damage is then modelled as a consequence of earthquake ground motion intensity at its location, using vulnerability functions expressing the probability of reaching a given damage category as a function of ground motion intensity. Because of the lack of knowledge and data, we employ vulnerability functions defined for building types expected to behave similarly under earthquake loading. Then, we use typological attribution models to map individual buildings to these pre-defined building types.

This study focuses on the physical damage to residential buildings as the output of the GP model. Two facts motivate this choice: First, most models of other earthquake consequences, such as recovery time and repair costs, use building damage as the starting point [38]. Second, empirical observations of building damage are the main building-specific data source to constrain estimates of other impacts after an earthquake occurs.

3.1. Ground Motion Intensity Measures

Predicting earthquake-induced damage to a spatially distributed set of buildings in this study involves the following assumptions: (i) Conditional on a specific scenario event, the logarithm of the ground motion intensity measure (IM) at dispersed building sites follow a multivariate normal distribution; (ii) Conditional on IM, damage to individual buildings is independent; (iii) The capacity of a building is given in terms of an IM threshold at which the building transitions into the next damage state; (iv) The capacities of all buildings in the region are expressed with respect to the same ground motion IM.

Let im_i denote the ground motion IM at location $\mathbf{x}_i$. The characteristics of the corresponding event, such as its magnitude, source geometry and style of faulting, are collected in vector $\mathbf{e}$. To estimate im_i, we employ empirical and ergodic ground motion models (GMMs) of the form

$$\ln im_i = m(\mathbf{x}_i, \mathbf{e}, \boldsymbol{\theta}_m) + \delta_{\mathrm{W},i} + \delta_{\mathrm{B}} \ , \tag{12}$$

where $m(\cdot)$ denotes the GMM trend function with parameters $\boldsymbol{\theta}_m$ that predicts the median IM conditional on the magnitude, the style of faulting, the source-to-site distance, and the local soil conditions. The terms $\delta_{\mathrm{W},i}$ and δ_{B} denote the within-event and between-event residuals that are both normally distributed with zero mean and variances σ^2_{W} and σ^2_{B}, respectively. Standard GMMs provide the parameters of Eq. 12 based on mixed-effects regression on ground motion recordings from historic earthquakes. The between-event residual indicates a systematic deviation from the trend function for a given event, whereas the remaining deviations are summarized in the within-event residual. Thus, for a specific event, δ_{B} is identical for two sites $\mathbf{x}_i$ and $\mathbf{x}_j$, whereas within-event residuals, $\delta_{\mathrm{W},i}$ and $\delta_{\mathrm{W},j}$ differ.

Past studies suggest that the latter are spatially dependent and that their joint distribution is multivariate normal [39, 40]. Most of these empirical evaluation studies conclude that the spatial correlation, ρ, decreases exponentially with the euclidean distance between two sites:

$$\rho(\mathbf{x}_i, \mathbf{x}_j, \ell) = \exp\left(\frac{- \parallel \mathbf{x}_i^{(1:2)} - \mathbf{x}_j^{(1:2)} \parallel_2}{\ell} \right) \ , \tag{13}$$

where ℓ denotes the length scale. For illustration, the distance at which ρ is lower than 5% is approximately 3ℓ. The superscript $(1:2)$ indicates that the correlation function acts on the geographical coordinates (i.e. Easting and Northing) of the input vectors, here assumed to be stored in dimensions one and two.

All random variables in Eq. 12 are Gaussian and the correlation function in Eq. 13 leads to positive definite covariances. Therefore, the random field of logarithmic ground motion IMs $f = \ln im$ follows a GP as defined in Eq. 1. The mean GP function is the same as the trend function of the GMM $m(\mathbf{x}_i, \mathbf{e}, \boldsymbol{\theta}_m)$, while the GP covariance function is $k(\mathbf{x}, \mathbf{x}', \boldsymbol{\theta}_k) = \sigma^2_{\mathrm{W}} \rho(\mathbf{x}, \mathbf{x}', \ell) + \sigma^2_{\mathrm{B}}$. For a finite set of buildings with inputs $\mathbf{X}$ and event characteristics $\mathbf{e}$, the joint distribution $p(\mathbf{f}|\mathbf{X}, \mathbf{e}, \boldsymbol{\theta}_f)$ is multivariate normal. The hyper-parameters $\boldsymbol{\theta}_f$ of this GP model are the GMM parameters $(\boldsymbol{\theta}_m, \sigma^2_{\mathrm{W}}, \sigma^2_{\mathrm{B}})$ and the length scale ℓ of the spatial correlation model.

The above formulation can be extended to cases involving multiple IMs. In such a case, each IM is represented as a separate GP with different mean and covariance functions. The correlation between the GPs can then be modelled using, for example, a linear model of co-regionalization [41].

3.2. Building Damage

We denote the damage state of building i as a discrete random variable Y_i with sequential, collectively exhaustive and mutually exclusive categories $y \in \mathcal{Y} = \{0, 1, ..., c_y\}$, where $y = 0$ indicates no damage. A building enters a certain damage category $y > 0$, if the ground motion IM at the building location exceeds its corresponding capacity threshold. In compliance with the state of the art, we assume these thresholds to be log-normally distributed with location parameters $\eta_1 < \eta_2 < ... < \eta_{c_y}$ and a common dispersion parameter β. Based on the logarithm of IM, here denoted as f_i, the probability of entering a damage state $y > 0$ is derived as

$$P(Y_i \geq y | f_i, \boldsymbol{\vartheta}) = \Phi\left(\frac{f_i - \eta_y}{\beta}\right) , \tag{14}$$

where $\boldsymbol{\vartheta}$ collects the parameters of the log-normal distributions. The expression in Eq. 14 is called a vulnerability function. The dispersion parameter β is chosen to be the same for all damage categories to prevent negative probability masses or, in other words, a crossing of the vulnerability functions. Given the similarity to the likelihood employed in case of ordinal GP regression defined in Eq. 6, inference using damage data from a small number of inspected buildings can be performed.

The approaches to estimate the parameters $\boldsymbol{\vartheta}$ for existing building stocks differ in whether they are calibrated using numerical models of buildings, damage data from past earthquake events, or a combination of these two data sources [42, 43, 44]. Here we assume that $\boldsymbol{\vartheta}$ have been derived for a finite set of building types $\mathcal{B} = \{1, 2, ..., c_b\}$, such as low-rise unreinforced masonry (URM) buildings with flexible floor diaphragms or mid-rise reinforced concrete (RC) shear wall buildings. Ideally, the inputs $\mathbf{x}_i$ would contain the corresponding vulnerability type $b_i \in \mathcal{B}$ for each building. In reality, however, this is seldom the case. For example, the number of floors could be available from municipal databases or estimated from digital elevation models to separate low- and mid-rise buildings. Furthermore, some building attributes are uncertain, if not unknown, for large subsets of buildings within the region. For example, acquiring input data on the lateral load resisting system often requires detailed and time-consuming field surveys. Typological attribution models are used solve this missing data problem.

3.3. Typological Attribution Model

Typological attribution, in the present context, describes the process of deriving unknown taxonomic attributes of buildings from known information. As discussed in Section 3.2, the required attributes depend on the building vulnerability types $\mathcal{B}$. The models discussed in this section, assume that the input data $\mathbf{x}_i$ for each building i contains the construction period or year, and the number of floors or an approximate height indicator. The lateral load resisting system (LLRS) is assumed unavailable, as in most real-world scenarios.

We model the randomness pertaining to the LLRS using a discrete random variable A_i with c_a possible attribute combinations $\mathcal{A} = \{1, 2, ..., c_a\}$. For example, $A_i = 1$ indicates an URM building with flexible floors and $A_i = 2$ indicates an URM building with stiff floors. A typological attribution model provides class membership probabilities for A_i as a function of known input data $\mathbf{x}_i$. This makes it possible to generate type samples a_i for all buildings first, and then attribute a vulnerability type $b = h(\mathbf{x}, a)$ using a deterministic function $h : \mathcal{X} \cup \mathcal{A} \rightarrow \mathcal{B}$, as well as gather the corresponding parameters of the vulnerability function as described above. As an example, the deterministic function $h(\cdot)$ may simply add the height class (i.e. low-rise) of a building, stored in $\mathbf{x}_i$, to the sampled unknown attribute combination a_i.

In the absence of region-specific data, a typology attribution model based on expert opinion presents the best alternative. Experts may provide insights into the proportions of building types in the region, for instance, by specifying the percentage of reinforced concrete shear wall buildings built in the region in certain periods. In addition, unknown attribute combinations for buildings in geographic proximity and/or of similar age and geometry might be correlated.

In Bodenmann et al. [45], we proposed to use latent functions to account for such variability and to incorporate typological information gathered during post-earthquake inspections. By analogy to the multi-class GP classification scheme presented in Section 2.2, we employ c_a independent latent functions $\mathbf{g} = [g^{(1)}, ..., g^{(c_a)}]$ and impose a GP prior on them. The class membership probabilities conditional on these functions are given by Eq. 7. The prior mean function $u(\cdot)$ is calibrated using an expert opinion based proportion estimate. For the covariance function, we combine multiple squared-exponential covariance functions. These account for spatially correlated deviations over longer length scales, due to regional differences in construction practices, v_{LS}, and short-scale correlated deviations for classes of nearby buildings that share similar age and number of stories v_{SS}. The resulting covariance function is denoted as

$$v(\cdot) = v_{LS} + v_{SS,1} \cdot v_{SS,2} \cdot v_{SS,3} \ , \tag{15}$$

where the individual functions of $v_{SS}(\cdot)$ measure similarity with respect to the geo-coordinates, the number of floors and the construction year. The adopted multiplicative structure of the covariance function limits high correlation to buildings whose three individual functions are all very similar. Detailed examples of typological attribution models are shown in the three case studies presented in Section 5 and the corresponding appendices.

3.4. Risk Model Workflow

The damage to residential buildings $\mathbf{X}$ resulting from a specified earthquake scenario $\mathbf{e}$ is estimated using the following framework: First, we draw r samples from the multivariate normal distributions $p(\mathbf{f}|\mathbf{e}, \mathbf{X})$ (ground motion intensity) and $p(\mathbf{g}|\mathbf{X})$ (typological attribution). Then, for each sample and for each building i, we first sample unknown attributes a_i from the categorical distribution $p(A_i|\mathbf{g}_i)$, with probabilities given by Eq. 7, and subsequently sample a damage state y_i from the categorical distribution $p(Y_i|f_i, a_i)$, with probabilities given by Eq. 6. Finally, for each sample j, buildings with identical state of damage are counted and the values stored in a vector $\mathbf{n}_j = \{n_{yj}|y \in \mathcal{Y}\}$. Thus, we obtain r samples from $p(\mathbf{N}|\mathbf{e})$, which denotes the joint predictive distribution of how the exposed buildings are partitioned amongst the different damage states. The marginal predictive distributions $p(N_y|\mathbf{e})$ of the entries of random vector $\mathbf{N}$ denote the predicted number of buildings in a certain damage state y.

4. Model Updating with Post-Earthquake Data

In the immediate aftermath of an earthquake it is imperative to use the early-arriving ground motion intensity and building damage data as soon as possible to constrain the uncertainty of regional damage estimates and to organize recovery. The following sections explain the inference steps that use the ground motion intensity data measured by seismic networks, and the building damage and typological data obtained by inspection. Figure 2 schematically illustrates the proposed risk model informed GP (RMGP) framework. The first rapid damage estimates are obtained using a regional risk model, where estimates of ground motion IMs are constrained by seismic recordings, analogous to the current shake map systems. Then, data from the first inspected buildings is used to further constrain the shake map and, in parallel, update the building vulnerability function parameters as well as the typological attribution model. The Python implementation of the RMGP framework is publicly available[2] and builds upon the software library GPFlow [46].

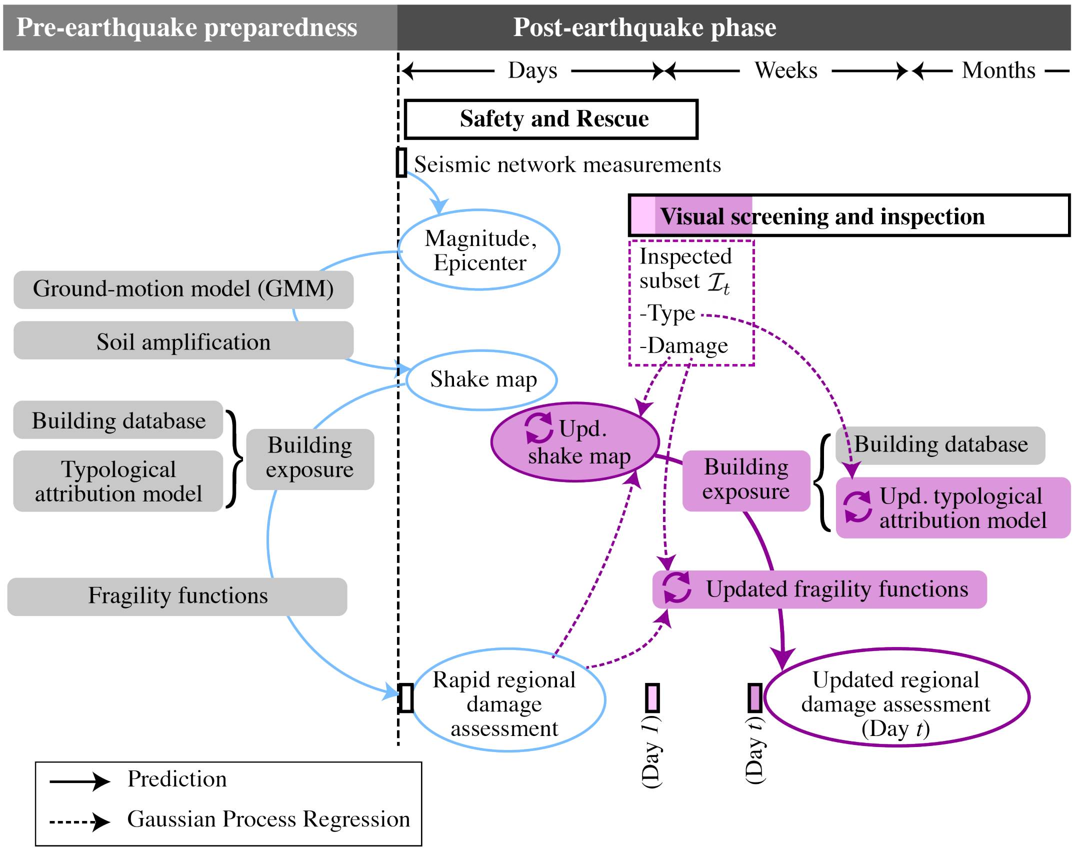

Figure 2: Schematic overview of the post-earthquake implementation of regional risk models and their GP dynamic updating using inspection data.

We use the subscripts $\mathcal{S}, \mathcal{I}$ and $\mathcal{T}$ to denote index sets related to the seismic stations, the inspected buildings, and the target buildings for which we aim to predict the damage category. In the immediate aftermath of an event, the model predicts the damage to buildings with indices $\mathcal{T} = \{1, 2, ..., n\}$. Indices of buildings inspected after t time steps are denoted as $\mathcal{I}_t \subset \mathcal{T}$, while the indices of the remaining not (yet) inspected buildings are denoted as $\mathcal{T}_t = \mathcal{T} \setminus \mathcal{I}_t$.

4.1. Seismic Recordings

This section explains how to constrain the predictions of ground motion intensity f using seismic recordings, where we closely follow Worden et al. [47]. We denote the dataset of recorded ground motion intensity measures from a seismic network as $\mathcal{D}_\mathcal{S} = \{(\mathbf{x}_i, z_i) | i \in \mathcal{S}\}$, where z_i are noise-free measurements of the ground motion intensity ($\ln z_i = f_i$). In a first step, we condition

[2] `https://gitlab.ethz.ch/bodlukas/earthquake-rmgp`

the distribution of the (constant) inter-event residual on the seismic recordings by deriving its posterior distribution $p(\delta_{\mathrm{B}}|\mathcal{D}_{\mathcal{S}}) = \mathcal{N}(\xi_{\mathrm{B}}, \psi_{\mathrm{B}})$, where the parameters are calculated as

$$\psi_{\mathrm{B}}^2 = \left(\frac{1}{\sigma_{\mathrm{B}}^2} + \frac{\mathbf{1}^\top \mathbf{C}_{\mathcal{S}\mathcal{S}}^{-1}\mathbf{1}}{\sigma_{\mathrm{W}}^2}\right)^{-1} , \tag{16a}$$

$$\xi_{\mathrm{B}} = \frac{\psi_{\mathrm{B}}^2}{\sigma_{\mathrm{W}}^2} \left(\mathbf{1}^\top \mathbf{C}_{\mathcal{S}\mathcal{S}}^{-1}(\ln \mathbf{z} - \mathbf{m}_{\mathcal{S}})\right) , \tag{16b}$$

with $\mathbf{C}_{\mathcal{S}\mathcal{S}}$ denoting the correlation matrix of the within-event residuals with entries $[\mathbf{C}_{\mathcal{S}\mathcal{S}}]_{ij} = \rho(\mathbf{x}_{\mathcal{S}i}, \mathbf{x}_{\mathcal{S}j}, \ell)$. Then we estimate the posterior predictive distribution of f evaluated at the target inputs $\mathbf{X}_{\mathcal{T}} = \{\mathbf{x}_i | i \in \mathcal{T}\}$, which is a multivariate normal distribution $p(\mathbf{f}_{\mathcal{T}}|\mathcal{D}_{\mathcal{S}}) = \mathcal{N}(\boldsymbol{\nu}_{\mathcal{T}}, \boldsymbol{\Psi}_{\mathcal{T}\mathcal{T}})$. For this calculation the posterior mean of the between-event residual ξ_{B} is added to the prior GP mean function as a constant term, and the remaining variance of the between-event residual ψ_{B}^2 is added to the prior variance of the within-event residual σ_{W}^2. The mean vector $\boldsymbol{\nu}_{\mathcal{T}}$ and covariance matrix $\boldsymbol{\Psi}_{\mathcal{T}\mathcal{T}}$ of the posterior predictive distribution are then computed as

$$\boldsymbol{\nu}_{\mathcal{T}} = \mathbf{m}_{\mathcal{T}} + \xi_{\mathrm{B}} + \mathbf{C}_{\mathcal{T}\mathcal{S}}\mathbf{C}_{\mathcal{S}\mathcal{S}}^{-1}(\mathbf{z} - \mathbf{m}_{\mathcal{S}} - \xi_{\mathrm{B}}) , \tag{17a}$$

$$\boldsymbol{\Psi}_{\mathcal{T}\mathcal{T}} = (\sigma_{\mathrm{W}}^2 + \psi_{\mathrm{B}}^2)\left(\mathbf{C}_{\mathcal{T}\mathcal{T}} - \mathbf{C}_{\mathcal{T}\mathcal{S}}\mathbf{C}_{\mathcal{S}\mathcal{S}}^{-1}\mathbf{C}_{\mathcal{T}\mathcal{S}}^\top\right) , \tag{17b}$$

where matrices $\mathbf{C}_{\mathcal{T}\mathcal{S}}$ and $\mathbf{C}_{\mathcal{T}\mathcal{T}}$ denote the correlation matrices evaluated between the target points and seismic stations and between the target points themselves. To generate correlated realizations of damage states for the target buildings $\mathbf{X}_{\mathcal{T}}$, we follow the workflow described in Section 3.4, where we sample from the posterior $p(\mathbf{f}_{\mathcal{T}}|\mathcal{D}_{\mathcal{S}})$ instead of the prior $p(\mathbf{f}_{\mathcal{T}})$.

4.2. Inspection Data

During inspection of a certain building i experts usually provide an estimate of the inflicted damage y_i via attributing a certain damage category. Here we assume that those are consistent with the damage state $\mathcal{Y}$ employed in the fragility functions of the risk model $y_i \in \mathcal{Y}$. Besides this damage description, experts are also asked to provide some basic building attributes, e.g., the dominating material of the lateral load-resisting system and its type. Again we assume that this information allows attribution of a typological combination $a_i \in \mathcal{A}$ employed in the risk model. We denote the dataset from building inspections conducted up to a given time step t after the event as $\mathcal{D}_{\mathcal{I}t} = \{(\mathbf{x}_i, y_i, a_i) | i \in \mathcal{I}_t\}$, where we omit the subscript t in the following for better readability. This data is used to perform two separate inference steps: First, we perform ordinal GP regression using the function f, where we also account for the seismic recordings $\mathcal{D}_{\mathcal{S}}$. Second, we perform multi-class GP classification using the functions $\mathbf{g}$ related to the unknown typological attributes.

In the first step, we use a_i to allocate a certain fragility class $b_i = h(\mathbf{x}_i, a_i)$ for each of the inspected buildings. Then we use the data y_i and the likelihood for ordinal data (see Eq. 6) to perform variational inference on f. Specifically, we approximate the true posterior $p(\mathbf{f}_{\mathcal{I}}|\mathcal{D}_{\mathcal{S}}, \mathcal{D}_{\mathcal{I}}) \propto p(\mathbf{y}|\mathbf{f}_{\mathcal{I}})p(\mathbf{f}_{\mathcal{I}}|\mathcal{D}_{\mathcal{S}})$ with a variational multivariate normal distribution $q(\mathbf{f}_{\mathcal{I}})$. Note that the posterior is conditioned on both the inspection data $\mathcal{D}_{\mathcal{I}}$ and seismic recordings $\mathcal{D}_{\mathcal{S}}$. To evaluate the variational free energy in Eq. 9, we replace the prior $p(\mathbf{f}_{\mathcal{I}})$ with the posterior predictive $p(\mathbf{f}_{\mathcal{I}}|\mathcal{D}_{\mathcal{S}})$ from Section 4.1. As outlined in Section 2.2, the objective function of this variational inference step (given by Eq. 10) allows for a simultaneous MAP estimation of the model hyper-parameters. We exploit this capacity in order to update the vulnerability functions. Specifically, we calculate

MAP estimates of the damage threshold parameters $\hat{\boldsymbol{\eta}}_b$ for all vulnerability types $b \in \mathcal{B}$. Given the limited amount of available building inspections in the early aftermath of an earthquake event, we keep the dispersion parameters (β in Eq. 14) and other hyper-parameters, such as the empirical GMM parameters, fixed at the values specified in the risk model.

In a second step we perform approximate inference on the latent functions $\mathbf{g}$ using the data a_i via standard variational inference as described in Section 2.2. The true posterior $p(\mathbf{g}_\mathcal{I}|\mathcal{D}_\mathcal{I})$ is approximated by a multivariate normal distribution $q(\mathbf{g}_\mathcal{I})$. We keep the parameters of the mean functions fixed to the initial expert-judgment based estimates and perform MAP estimation for the variances and length scales of the covariance function specified in Eq. 15.

To predict damage for a set of not (yet) inspected target buildings $\mathbf{X}_{\mathcal{T}t}$, we follow the procedure outlined in Section 3.4 with two exceptions: First, we sample from the (approximate) posterior predictive distributions $q(\mathbf{f}_\mathcal{T})$ and $q(\mathbf{g}_\mathcal{T})$. The parameters of these multivariate normal distributions are calculated using Eq. 11. To take into account the seismic recordings in case of function f, we use the mean vector $\boldsymbol{\nu}_\mathcal{T}$ and covariance matrices $\boldsymbol{\Psi}_{\mathcal{T}\mathcal{T}}$ and $\boldsymbol{\Psi}_{\mathcal{T}\mathcal{I}}$ to evaluate matrices $\mathbf{A}$ and $\mathbf{B}$ in Eq. 11. Second, to sample damage categories we employ the updated fragility functions $p(Y|f, a, \boldsymbol{\vartheta})$, using the inferred MAP-estimates of the threshold parameters.

5. Application to three case-studies

Three case studies, one simulated earthquake scenario in the Swiss Canton of Zurich and two real-world case-studies, namely the 1998 Pollino (Italy) and the 2010 Kraljevo (Serbia) earthquake events, are used to apply and evaluate the proposed RMGP framework. Specifically, we seek to predict the spatial pattern of earthquake damage as the number of buildings assigned to be in a given damage state within sub-regions (i.e. zip-codes or municipalities). Table 1 summarizes key characteristics of the three examples, while Figure 3 shows the location of the epicenters, the boundaries to the considered sub-regions, and the locations of seismic network stations, if available. Whereas the earthquake events share similar magnitudes, the three case studies strongly differ in terms of the size of the considered region, the number of exposed buildings, and the coverage of the seismic network. The Zurich case represents a densely populated region that is equipped with four seismic network stations. The Pollino case study covers the largest area but has a relatively low building density. Finally, the Kraljevo case study covers a small region with few buildings and no seismic network stations.

Table 1: Characteristics of the earthquake events and the considered regions in three case studies.

	Zurich (simulated)	Pollino (1998)	Kraljevo (2010)
Magnitude $\mathrm{M_w}$	5.8	5.6	5.5
Number of buildings	33594	20528	1959
Area size [km^2]	97.1	931.1	25.4
Number of subregions	22	14	3

Four levels indicate the severity of earthquake damage to residential buildings: no damage, slight, moderate, and extensive damage. These four damage states are encoded as increasing integers $\mathcal{Y} = \{0, 1, 2, 3\}$. We mimic a post-event application by providing first predictions in the immediate aftermath ($t = 0$) using shake maps obtained following the process described in Section 4.1, which are constrained using only ground motion recordings $\mathcal{D}_\mathcal{S}$. Subsequently, gradually

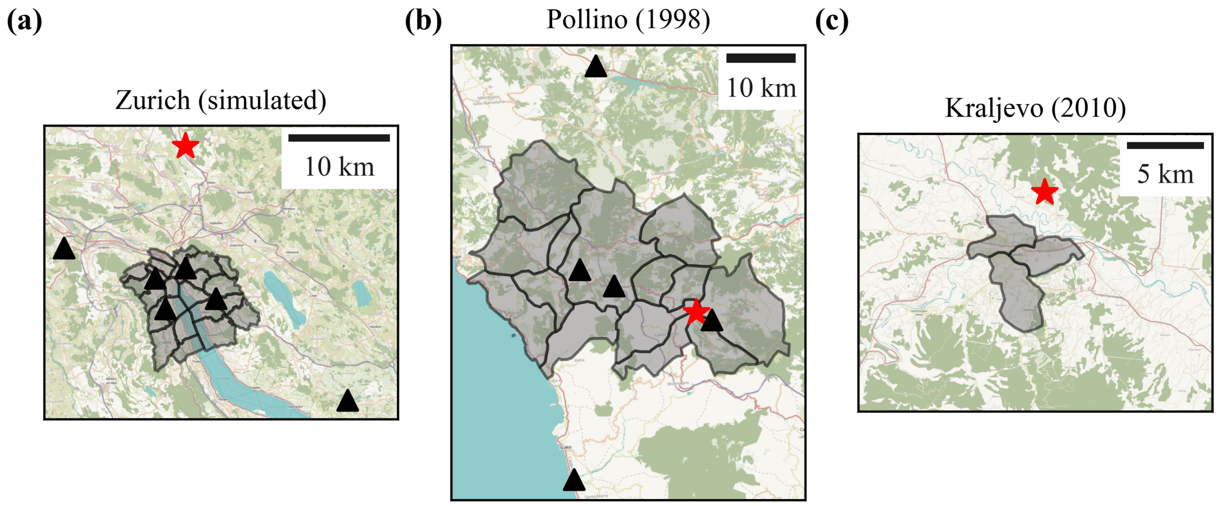

Figure 3: Overview of the three case studies: (a) Simulated M6 scenario in the canton of Zurich (Switzerland), (b) Pollino (Italy) earthquake from 1998 and (c) Kraljevo (Serbia) earthquake from 2010. The grey shaded areas delimit the sub-regions of interest, the star indicate the epicenter locations, and the triangles indicate the stations of seismic networks, if available.

increasing amounts of data from building inspections $\mathcal{D}_{\mathcal{I}t}$ enable dynamic updating the risk model components[3] (see Section 4.2), and thereby lead to improved impact predictions in subsequent time steps $t > 0$. To assess the performance of the proposed RMGP framework in each time step t and for each subregion k, we compare its predictions $\mathbf{N}_{tk}$ of the actual number of buildings in each damage state to predictions obtained using two entirely data-driven methods (presented in Section 5.2).

5.1. Data and Models

For all three case-studies, the available building-specific input data $\mathbf{X}_{\mathcal{T}}$ consists of the geographical coordinates, the Eurocode 8 soil class [48], the construction year (or period), and the number of stories (either exact or as a range). The event characteristics $\mathbf{e}$ consist of the earthquake (moment-) magnitude, the coordinates of the epicenter and the style of faulting.

The historic inspection data of the Pollino and Kraljevo events do not contain information regarding the time when a building was inspected. Therefore, the inspection process is simulated in the same manner as for the simulated Zurich case study. Namely, at the beginning of each time step (inspection day), an un-inspected building is randomly assigned to each available inspection team. Each inspection team then continues by additionally inspecting the four geographically closest (non-inspected) buildings to complete this time step. Whereas this procedure successfully produces a spatial clustering of inspection data, the time trajectory of the entire inspection process is simplified. Therefore, the results are presented as a function of the number of inspected buildings available for inference rather than in terms of a specific time since the event.

In the following, the three case studies are briefly introduced. In addition, the supplementary material contains further information about the respective prior risk models, the inspection result, the building databases and, for the Pollino case study, the required data pre-processing steps.

[3]The inference process at each time step starts from the initial risk model and shake map, and considers data from all buildings that were inspected until this time step.

5.1.1. *Zurich*

This case study examines damage to approximately 34000 residential buildings within the Canton of Zurich (Switzerland) inflicted by a $M_w = 5.8$ scenario earthquake (see Figure3). The following paragraphs introduce the available input data, the risk model components and the models used to generate the 'true' building damage.

Available input data. Building-specific information $\mathbf{X}_{\mathcal{T}}$ is retrieved from the office for spatial development of the Canton of Zurich (2020). Approximately 75% of the building stock was built before 1980, with 65% of all buildings having between 3 and 5 stories. The locations of the seismic network stations $\mathbf{X}_{\mathcal{T}}$ are fictitious and deliberately chosen to present a dense network within the region of interest. We simplify by assuming soil class B in the entire region.

Risk model components. The earthquake risk model, assumed to be available before the scenario earthquake, builds upon the empirical ground motion model of Akkar and Bommer [50] and the spatial correlation model of Esposito and Iervolino [51]. The employed vulnerability functions are defined for the peak ground acceleration (PGA) IM, and specified for 12 building types $b \in \mathcal{B}$, encompassing buildings from older low-rise (1-2 stories) URM with flexible floor diaphragms to high-rise (6 stories or higher) RC shear wall buildings designed according to modern seismic guidelines. Because of the available input information $\mathbf{x}$, the randomness pertains to four unknown typological attribute combinations $a \in \mathcal{A}$, which are URM with either flexible or stiff floor diaphragms, and RC buildings with either shear walls or infill frames forming the structural system.

Data-generating models. The same input information, $\mathbf{X}_{\mathcal{T}}$, specified above is used to generate the '*true*' data. Four cross-correlated random fields of PGA and elastic, 5%-damped spectral acceleration at periods of 0.3, 0.6 and 1.0 seconds are generated using the GMM of Chiou and Youngs [52]. The spatial cross-correlation of the within-event residuals is accounted for using the model of Markhvida et al. [53] and the cross-correlation of the inter-event residuals follows the relations provided by Baker and Cornell [54]. Finally, the vulnerability functions developed by Martins and Silva [35] are used to simulate damage for each building. These functions are defined for one of the four aforementioned IMs, depending on the building type and the number of stories. The data-generating model is deliberately chosen to differ from the risk model used for updating within RMGP to reflect real conditions in which models never fully capture the real processes leading to regional damage distributions.

While the available building data (coordinates, construction year, number of stories) is deemed realistic, the building vulnerability functions and soil conditions are not. Therefore, neither the 'true' simulated results nor the prior risk model predictions reflect the real earthquake risk conditions in Zurich.

5.1.2. *Pollino*

On September 9, 1998, the Pollino region in Italy was struck by a $M_w = 5.6$ earthquake that caused widespread building damage in the region's municipalities. This case study examines approximately 20000 residential buildings spread over 14 subregions (see Figure 3).

Available input data. The event characteristics and the seismic recordings are taken from the Engineering Strong-Motion (ESM) database [55]. The main data source for this case study is the publicly available building damage database Da.D.O. [56], consisting of inspection results for about 13000 residential buildings in the affected region. This database provides the geographic coordinates, the construction period (from pre-1919 to post-1981) and the number of stories, with the latter categorized into three height classes: Low-rise (1-2 stories), mid-rise (3-4 stories) and high-rise (¿5 stories). In cases where multiple entries were lumped at the same coordinates, the European Settlement Map [57] serves as reference to attribute unique coordinates to each building.

The available data does not cover the entire building stock. The Italian National Institute of Statistics [58] census data from 2001 provides information on the number of residential buildings from a certain construction period and with a certain number of stories, at municipality level. Based on this information, we augment the damage database with the missing buildings, for which spatial coordinates are generated using the European Settlement Map. The 2001 census data is the temporarily closest to the 1998 event, yet, changes that occurred in the three years may induce errors that are deemed small and thus, acceptable. Finally, soil classes are attributed according to Forte et al. [59].

Risk model components. The earthquake risk model, assumed to be available prior to the event, builds on the empirical GMM of Bindi et al. [60] and the spatial correlation model of Esposito and Iervolino [51]. In accordance with the Italian risk model [10], buildings are first classified according to their material: URM or RC. For URM buildings, Rosti et al. [61] empirically derived vulnerability functions for three building types (A, B, and C1, with decreasing susceptibility to earthquake-induced damage) further separated into low-rise and mid-/high-rise buildings. For RC buildings, we employ vulnerability functions from Rosti et al. [44] that differentiate two building types (C2, D) if the building was built before and after the introduction of seismic design guidelines. RC buildings are further separated into low-, mid- and high-rise buildings. Thus, a total of 12 classes $b \in \mathcal{B}$ are defined.

Processing of inspection results. The available inspection results provide damage descriptions for individual groups of structural elements, such as the vertical and horizontal elements of the lateral load resisting system. Following the approach by Dolce et al. [56], we transform these descriptions into overall levels of building damage and link them to one of the four typological attribute combinations $a \in \mathcal{A}$. Given the absence of documented damage, the buildings, which are added to augment the database, are assumed to be undamaged. This assumption, despite not reflecting small and/or undocumented damage, is assumed reasonable.

5.1.3. Kraljevo

On November 3, 2010, an earthquake of $M_w = 5.4$ struck the region of Kraljevo (Serbia), claiming two fatalities and leading to damage reported on about 16000 structures. This case study examines damage to approximately 2000 residential buildings, for which detailed information are publicly available [21]. We group these 2000 buildings artificially into three subregions (see Figure3).

Available input data. The event characteristics are taken from the earthquake catalogue published by the Seismological Survey of Serbia [62]. No seismic recordings from stations sufficiently close to the Kraljevo region are available. The aforementioned building dataset covers single-family houses, 90% of which have no more than two stories. No additional pre-processing of this dataset was performed; thus we consider following building information is available before the event: geographical coordinates, number of stories, construction year, and soil class.

Risk model components. The earthquake risk model, builds on the empirical ground motion model of Akkar and Bommer [50] and the spatial correlation model of Esposito and Iervolino [51]. Stojadinovic et al. [21] have classified buildings into four main categories, Adobe, URM with two different types of brick, and URM with tie-beams. The latter are further divided into three subcategories based on the construction year, giving 6 distinct building types, where additionally low- and mid-rise buildings are separated leading to a total of 12 vulnerability types $b \in \mathcal{B}$. Due to the lack of region-specific fragility functions, we employ vulnerability functions provided in Martins and Silva [35]. In our prior risk model we consider the four main categories as the random typological attribute combinations $a \in \mathcal{A}$. Using the available input information, we establish a deterministic link to the 12 vulnerability types $b \in \mathcal{B}$.

5.2. Data-Driven Methods for Comparative Analysis

To assess the quality of the predictions obtained using the proposed RMGP framework, two purely data-driven regression methods are used. A simple linear model is selected as basic option, while a random forest model reflects a common regression approach from previous literature studies [63, 20]. The following building-specific features are used for the data-driven regression: geo-coordinates, epicentral distance, construction year, number of stories, and soil class.

Ordered Linear Probit (OLP). regression is similar to the ordinal GP regression outlined in Section 2.2, with the exception of the latent function f not being a GP but a linear combination of the features $\mathbf{x}^\top \boldsymbol{\psi}$. The form of the likelihood is identical to Eq6 and the threshold parameters $\boldsymbol{\eta}$, together with the feature weights $\boldsymbol{\psi}$, are estimated by minimizing the negative log-likelihood.

Random Forest (RF). classifiers are ensembles of decision trees that are trained using bootstrap samples of the data [64]. For each set of bootstrap samples, a decision tree is grown by recursively partitioning the input space until all terminal nodes contain a pre-defined minimum number of samples. During training, a random subset of the d input variables is selected at each node and the best split is chosen as the split that leads to the largest decrease in Gini-impurity [65]. For prediction, the class membership probabilities are derived via averaging the probabilities predicted by the individual trees[4]. The number of trees in the forest, the minimal required number of samples in a node and the size of the feature subset to split the nodes are hyperparameters of this model. We fixed the number of trees in the forest to 1000 and perform a grid-search amongst the pre-specified values for the other two hyper-parameters, keeping the combination that returns the highest out-of-bag performance.

A GP Regression method for model specification following Sheibani and Ou [23] was also implemented. However, because the damage data is ordinal (and not continuous), this regression method showed the worse performance and thus, has been omitted from the presentation of the results. Indeed, the application of a regression model to ordinal (or categorical) data would require an additional training step that includes a model to translate the continuous output into the discrete categories.

[4]Based on the features of a given test input, a trained decision tree assigns the corresponding terminal node and calculates the class membership probability as the proportion of class labels in this terminal node that were seen during training.

5.3. Results

Probabilistic predictions $p(\mathbf{N}_k)$ of the distribution of damage states $y \in \mathcal{Y}$ sustained by the exposed buildings in subregion k are the principal metric used to compare the proposed RMGP framework to the purely data-driven models. We introduce two error metrics: i) the marginal prediction error (MPE) that measures the performance on predicting the number of buildings in any single damage state; and (ii) the joint prediction error (JPE) that measures performance jointly over all damage states.

The predictive multivariate distribution $p(\mathbf{N}_k)$ is approximated via r Monte-Carlo samples $\mathbf{n}_{kr}$, as described in Section 3.4. The vector $\check{\mathbf{n}}_k$ denotes for a subregion k the 'true' number of buildings in each damage state. We employ the continuous ranked probability score (CRPS) as the error metric to quantify the MPE. The CRPS metric is often used to compare the performance of probabilistic forecasts, for instance from weather forecasting systems [66, 67, 68]. Thus, for the amount of buildings in any subregion k being in damage state y, the MPE for predictive distribution $p(N_{ky})$ with respect to the true value $\check{n}_{ky}$ is approximated via r samples as

$$\mathrm{MPE}_{ky} \hat{=} \mathrm{CRPS}(p(N_{ky}); \check{n}_{ky}) \approx \frac{1}{r} \sum_{j=1}^{r} \left(\parallel n_{kyj} - \check{n}_{ky} \parallel_2 - \frac{1}{2r} \sum_{i=1}^{r} \parallel n_{kyj} - n_{kyi} \parallel_2 \right). \tag{18}$$

The first term of the MPE formulation measures the bias - or lack of accuracy - of the predictions, while the second term of the MPE accounts for the dispersion - or lack of precision. For deterministic predictions, the MPE simplifies to the absolute error metric, which renders the interpretation of the MPE more intuitive. Also, using the definition of Eq. 18, the MPE is given in the same units than the predicted quantity, which is, in this case, the number of buildings. Finally, the CRPS can be readily extended to the multivariate case, as discussed in the following. The energy score (ES) is used to quantify the JPE. The ES generalizes the CRPS towards multivariate predictions. The ES is calculated in a similar way to Eq. 18, by replacing the scalar quantities n_{kyj} and $\check{n}_{ky}$ with vectors $\mathbf{n}_{kr}$ and $\check{\mathbf{n}}_k$, which include all of the damage states. Hence, for the ES, the first term in Eq. 18 measures the Euclidean distance between the sampled damage state partitioning and the true partitioning of the exposed buildings in a four-dimensional space (because we consider four damage states in the case studies). Again, the ES as JPE has the same units as the predicted quantity.

5.3.1. Zurich

The Zurich case study, consisting a simulated earthquake scenario, contains the most exhaustive information of the geographical distribution of the ground motion intensity and the 'true' damage. Therefore, the local performance of impact estimation is reported for the Zurich case study in Figure 4, which contains the predicted number of moderately damaged buildings in a single subregion after the M5.8 scenario earthquake (see Figure 3 for the earthquake details). When increasing amounts of buildings are inspected and thus, available to train the GP model, the kernel density of the predictions made with the proposed RMGP framework increases, meaning that the prediction uncertainty decreases. The initial prediction - relying exclusively on the risk model and the shake map that is established using seismic-network measurements - suffers from large uncertainties: the amount of buildings with moderate damage ranges from below 100 to over 800 buildings, which corresponds to a range from 7% to 54% of the total amount of 1471 buildings in this subregion. Such large uncertainties undermine good decision making and the organisation of recovery activities. The predictions stemming from the RF model, which do not have a physics-based component, require a minimum amount of initial building inspections to enable predictions and therefore, are represented only from the second time step, $t = 1$. When comparing the proposed RMGP framework and the RF model, the latter returns more precise yet less accurate results as the 'true' value is not contained within the 90%-confidence interval. The over-confident predictions with almost no uncertainty produced by the RF model stem from the implicit model assumptions. Conditional on the input values $\mathbf{x}$ of some buildings, the RF model considers damage to these buildings as independent, regardless of how close the buildings are. The RMGP framework, on the other hand, generates spatially correlated samples of ground motion intensity, while building damage is then assumed to be independent, conditioned on these samples.

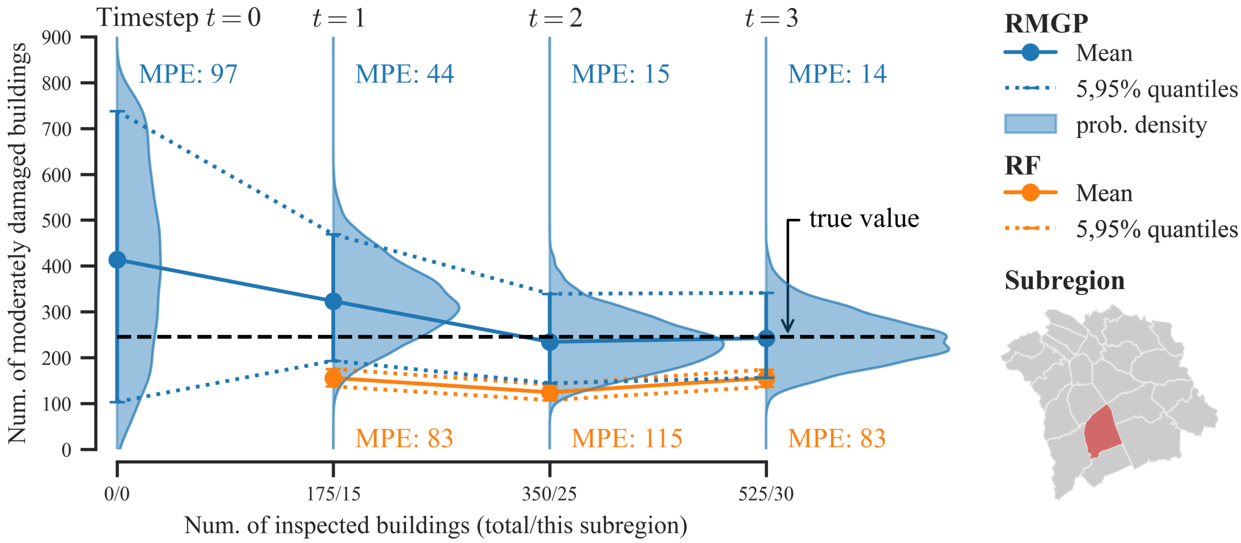

Figure 4: Zurich case study: Evolution of the predictions of the number of moderately damaged buildings ($y = 2$) in the indicated subregion over four time steps with increasing amounts of inspection data. Predictions made with the proposed RMGP framework (blue) and an entirely data-driven random forest (RF, orange) model are compared with the true value. The numerical values indicate the marginal prediction error (MPE) at the four time steps for both methods.

In addition to the probabilistic distribution of the predicted number of moderately damaged buildings, the MPE values, provided in Figure 4 and derived using Equation 18, quantify the reduction of the predictive uncertainty and the convergence of the median value towards the 'true' amount of moderately damaged buildings. Comparing the kernel density distributions with the MPE values highlights the joint contribution of precision and accuracy of the MPE. Providing the MPE for all four damage states (from no damage, $y = 0$, to extensive damage, $y = 3$) of the considered subregion, Table 2 complements Figure 4. The four time steps coincide with the time steps reported in Figure 4, where inspection data is available for 0, 0.5, 1 and 1.5 % of all 34000 buildings. Finally, Table 2 also provides the JPE, which measures the error in the joint distribution of building damage states assigned in this subregion.

While Figure 4 represents the results for a single subregion, impact estimates generally cover the entire affected region. The error in damage predictions for each subregion, represented using the JPE, is therefore reported in Figure 5. The geographical distribution of prediction errors confirms that the RMGP framework rapidly converges for all subregions, while the JPE remains high for the RF model, especially in some subregions. Also, the cumulative joint prediction error (CJPE),

Table 2: Marginal and joint prediction error for damage state counts without inspection data ($t = 0$) and using data gathered in three inspection time steps for the subregion indicated in Figure 4.

			RMGP				**RF**		
			$t=0$	$t=1$	$t=2$	$t=3$	$t=1$	$t=2$	$t=3$
Marginal Prediction Error[1]	MPE_{ky}	$y=0$	118	65	22	26	36	87	27
		$y=1$	363	15	9	26	58	36	61
		$y=2$	97	44	15	14	83	115	83
		$y=3$	14	4	8	4	23	21	16
Joint Prediction Error	JPE_k		365	80	31	41	113	149	107

[1] The true damage state counts in this subregion are [519, 676, 245, 31] for damage states $y \in [0, 1, 2, 3]$, respectively.

obtained by summing the JPE over all subregions, shows the convergence of the prediction results towards the 'true' building damage. Note that the results shown so far consider the outcome of a single inspection process, which leads to the singularity that the predictions at time step $t = 2$ are slightly more accurate than at time step $t = 3$. This singularity may originate from a particularly efficient distribution of buildings inspected until time step $t = 2$ or a slight overfit to a badly balanced inspection set at time step $t = 3$. Given the influence of a single random inspection process, multiple inspection runs are considered in the following.

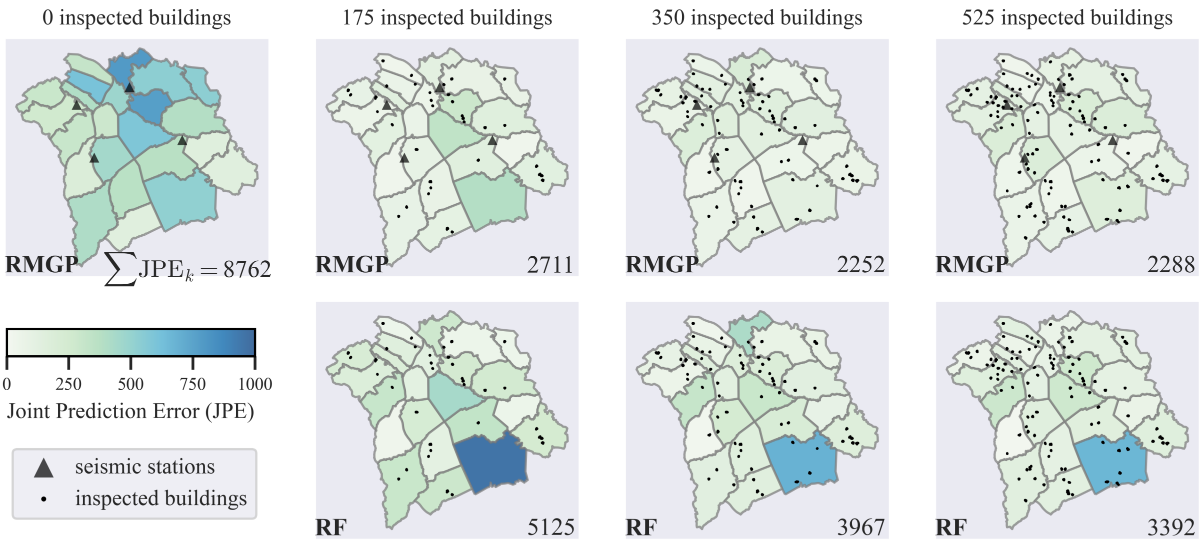

Figure 5: Zurich case study: Joint prediction error (JPE) for the proposed RMGP framework (top row) and a RF model (bottom row) in all considered subregions at four time steps (columns) with increasing amounts of inspection data. The numerical values indicate the cumulative JPE (over all subregions) achieved by both methods in different time steps.

The CJPE is subsequently used to compare the prediction outcomes of RMGP, RF, and OLP methods in Figure 6. The violin outlines presented in this figure provide a visual representation of the distribution of the kernel probability density, i.e., the width of the shaded area corresponds to the proportion of the data at this vertical ordinate. The probability distributions is based on

50 random inspection runs, each leading to one instance of CJPE. The dashed and dotted lines within the violins indicate the median and the inter-quartile range. Statistically, all three methods converge with additional data. For the proposed RMGP framework, the information from the first 175 inspected buildings provides in all cases a significant improvement in the prediction results, with the median CJPE decreasing by over 50% with respect to the initial risk model predictions.

The RMGP framework produces the most reliable predictions, especially with only few available building inspections. The median CJPE of the RMGP framework with 175 inspected buildings is lower than the median CJPE of the RF model after 525 buildings were inspected. Hence, the additional information provided by the underlying risk model reduces the need for building inspections in absence of optimised inspection schemes. The influence of the inspection scheme is larger for the purely data-driven methods, RF and OLP, than for the RMGP, which includes physics-based information. Finally, recall that all buildings are modelled as being on soil class B, meaning that additional influence of varying soil classes is likely to further increase the number of inspected buildings required for data-driven methods, such as RF and OLP, to converge.

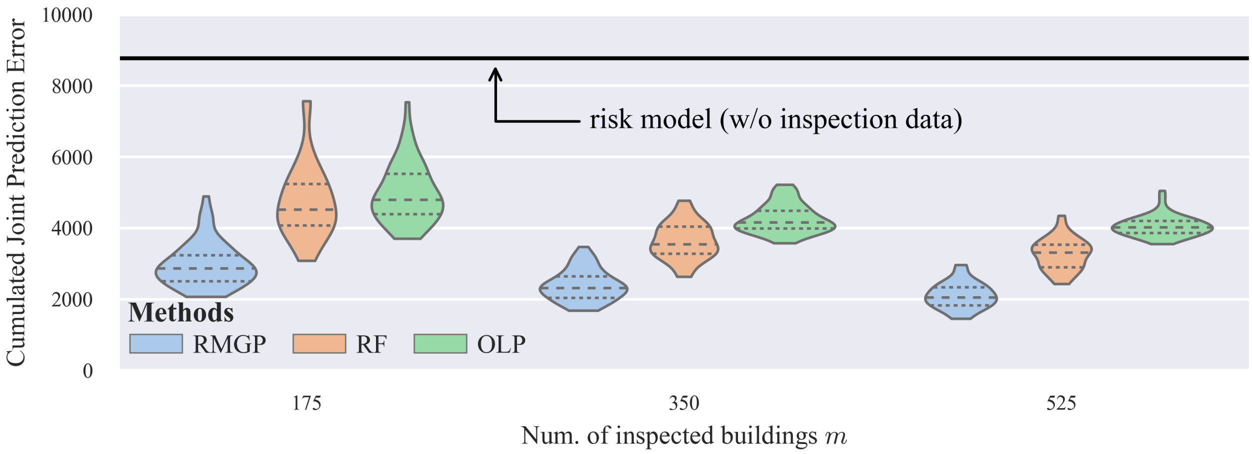

Figure 6: Zurich case study: Violin plots for the cumulative joint prediction error (CJPE) over 50 random inspection processes using the proposed RMGP framework, a random forest (RF) and an ordered linear probit regression (OLP) model for three time steps corresponding to increasing amounts of inspection data available for training. The solid black line illustrates the CJPE obtained with the prior risk model and recordings from seismic stations.

In real case studies, a shake map cannot be known with precision, except at the seismic network stations used to derive the shake map. Therefore, we use this simulated case study to analyze the performance of the proposed RMGP framework in updating function f, which is the logarithmic ground motion IM, in this case expressed as PGA. The top of Figure 7 shows the evolution of the inferred median PGA over the entire region when increasing amount of buildings are inspected[5], while the bottom of the figure illustrates the probability density of the posterior distributions at three specific locations. The uncertainty at specific locations is reduced, especially when some buildings in the vicinity of the location get inspected, as evidenced by the uncertainty reduction after 175 inspections for locations S1 and S2, while at location S3 the variance reduced only after 525 buildings are inspected, some of which are in the same subregion as location S3. This comparison is particularly interesting given that the 'true' damage for the vast majority of the

[5]The median PGA corresponds to the exponential mean of the posterior distribution of function f, which is, for these maps, evaluated at a regularly spaced grid.

inspected buildings originates from IMs other than PGA, namely spectral accelerations at periods of 0.3, 0.6 and 1.0 seconds. The physical meaning of the latent function provides information on the shaking intensity that may prove useful in assessing other earthquake consequences or the risk of cascading events.

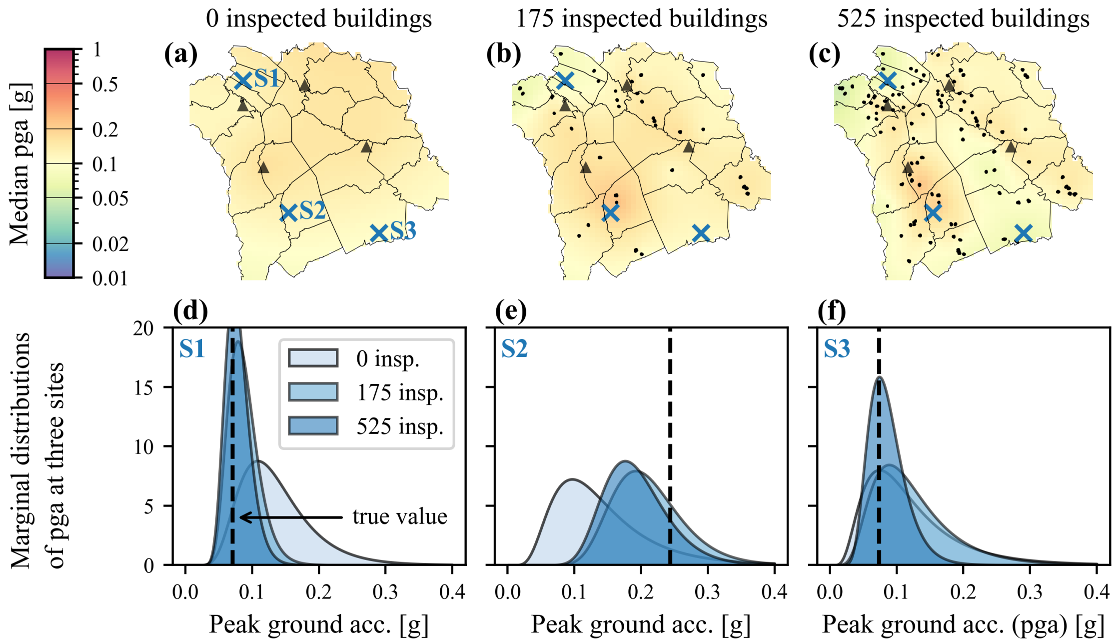

Figure 7: Zurich case study: Inferred estimates of peak ground acceleration (PGA). The top row (a-c) shows maps of the median PGA inferred from increasing amounts of inspection data, while the bottom row (d-f) illustrates the posterior distributions of PGA for three building sites S1, S2 and S3.

5.3.2. Pollino

The rapid impact assessment with the risk model for Pollino, shown in Figure 8 (top left), produces large prediction errors for three subregions. When inspection outcomes are used to update the risk model, the quality of the estimates increases rapidly in all subregions. The CJPE, representing the total number of mis-predicted buildings, obtained by aggregating subregion data, decreases by over 60%, from 3480 to 1397, by fusing the risk model with inspection data of 600 buildings, corresponding to 3% of the entire building stock in the affected region. For this specific inspection sequence, the RF model (shown in the bottom row of Figure 8) performs better than the initial risk model. In addition, with increasing amounts of data, the prediction performance of the RF model, measured through the CJPE, approaches the one of the RMGP framework. This underlines the importance of the additional information provided by the risk model in the RMGP framework, especially in the early stages of inspection, when data is scarce.

As outlined in Section 4, the proposed RMGP framework leverages inspection data not only to constrain the distribution of the ground motion IMs, in this case PGA, but also to update the vulnerability function parameters. Figure 9 shows the evolution of the inferred median PGA in the top row, and the evolution of the estimated vulnerability curves for low-rise URM buildings of class A in the bottom row. Compared to the Zurich case study, the Pollino region is almost ten

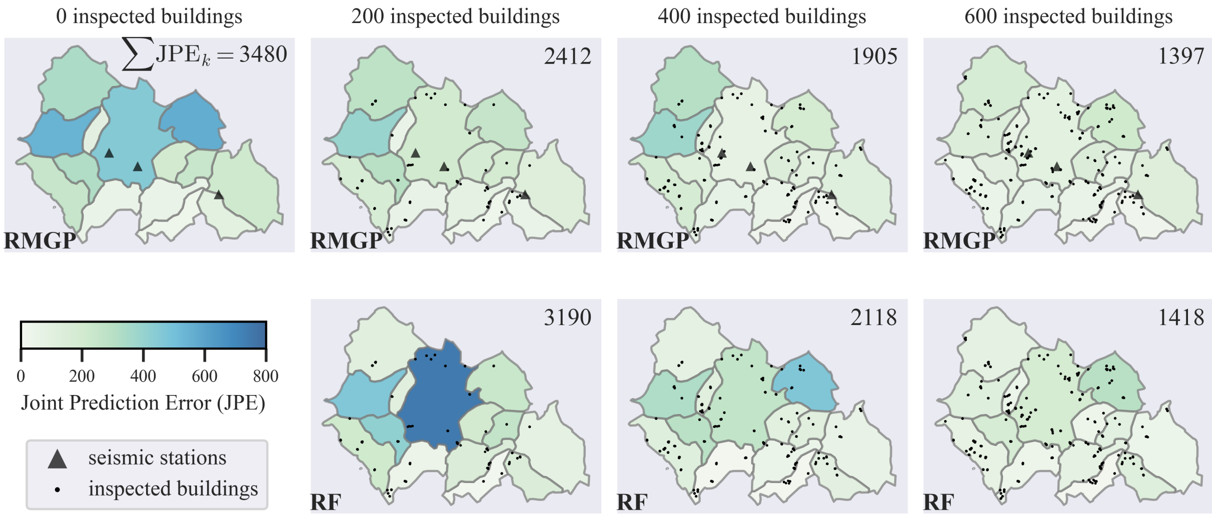

Figure 8: Pollino case study: Joint prediction error (JPE) for the proposed RMGP framework and the RF mdoel (bottom row) in all considered subregions at four time steps (columns) with increasing amounts of inspection data. The numerical values indicate the cumulative (over all subregions) JPE achieved by both methods at the different time steps.

times larger, but has significantly sparser seismic network, which results in larger heterogeneity of the inferred median PGAs. Whereas the initial shake map varies from below 0.02g to 0.25g, fusing inspection data reveals a more fine-grained spatial pattern. While the updating does not change the vulnerability curve for the slight damage state (delimiting $y = 0$ from $y = 1$), the initial vulnerability curve for the extensive damage state shifts towards less conservative predictions with the addition of inspection outcomes. Such updated vulnerability curves may be useful in case of aftershocks in the same region, or to develop risk models in other regions with similar building inventories.

As for the Zurich case study, the violin outlines of kernel densities of the CJPE corresponding to 50 random simulation inspection process outcomes for the Pollino case study are reported in Figure 10. Compared to the simulated Zurich earthquake, the initial Pollino risk model is contained within the range of possible outcomes for the RF and the OLP models. Updating the risk model with inspection outcomes in the RMGP framework leads to a median reduction of approximately 50%, yet, the dispersion of CJPE results due to inspection processes is large for the first 200 inspected buildings. In addition, the median and 75-percentile CJPE values for the RMGP framework with 200 inspected buildings are similar to the corresponding RF model values based on 600 building inspections. Again, this underlines that the fusion of the risk model with inspection data outperforms the risk model without inspection data and the purely data-driven predictions that do not have a physics-based foundation.

5.3.3. Kraljevo

The Kraljevo case study is characterized by a much smaller impacted area (see Figure 3), for which no seismic network stations are available. Nevertheless, as seen in Figure 11, the predictions originating from the proposed RMGP framework outperform the predictions obtained from the RF and the OLP models, as well as the original risk model predictions. The predictions from RF and

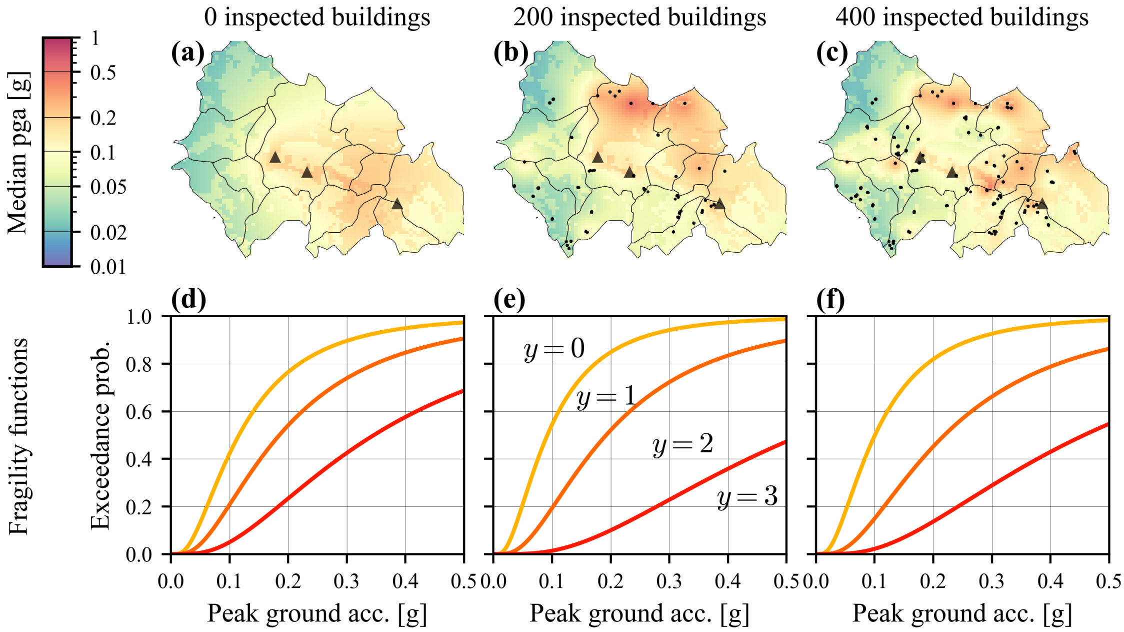

Figure 9: Pollino case study: Top row shows the predicted median PGA inferred from increasing amounts of inspection data. The bottom row shows the evolution of the vulnerability functions for low-rise URM buildings of class A (the most vulnerable ones) with accumulation of the inspection data.

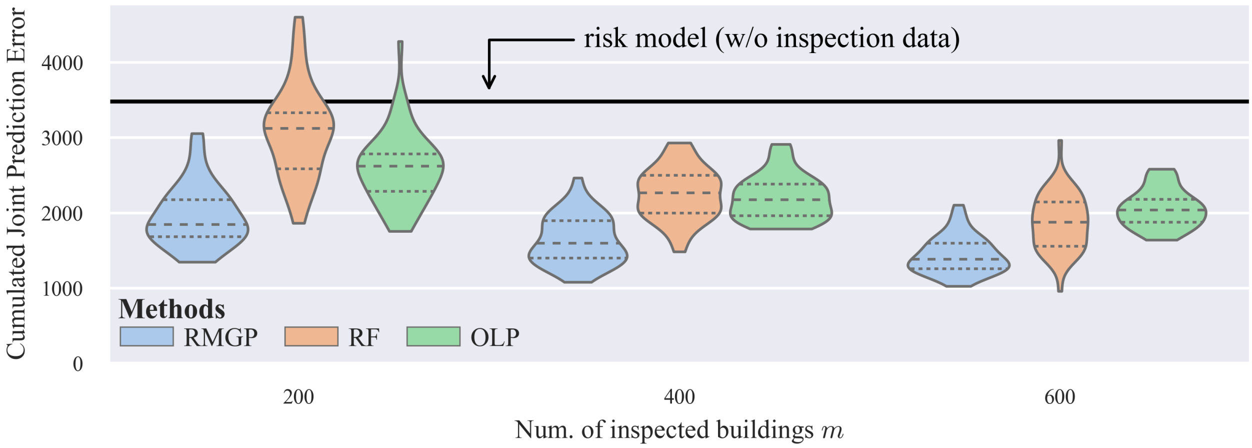

Figure 10: Pollino case study: Violin plots for the cumulative joint prediction error (CJPE) over 50 random inspection processes using the proposed RMGP framework, and the RF and OLP models for three time steps with increasing amounts of inspection data available for training. The solid black line illustrates the CJPE obtained with the prior risk model and recordings from seismic stations.

OLP models are very sensitive to changes in the inspection sequence, particularly when only 40 building have been inspected. Despite the limited geographic range of the case study, 40 buildings are very few considering the amount of information and correlations to be inferred, especially in absence of an underlying risk model.

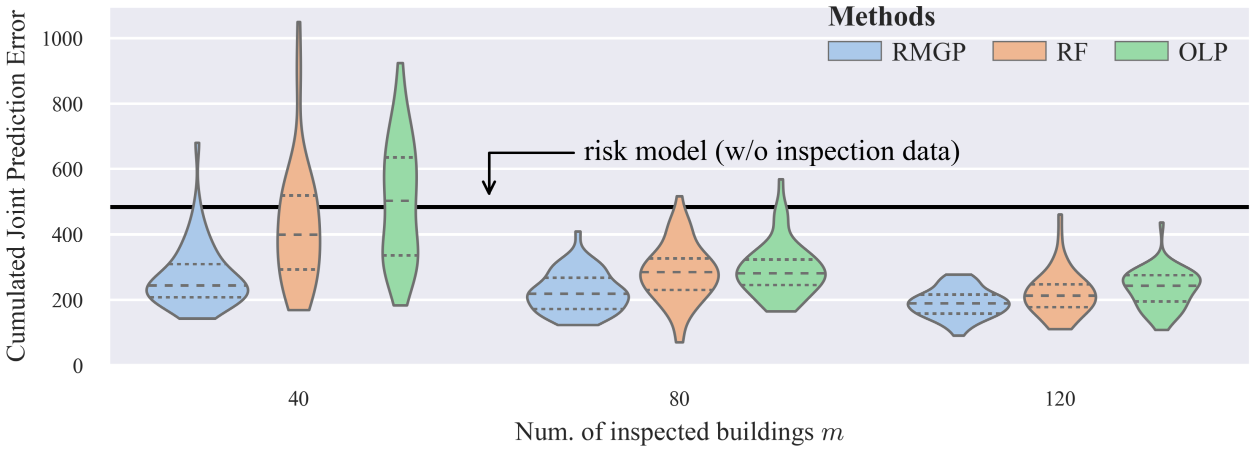

Figure 11: Kraljevo case study: Violin plots for the cumulative joint prediction error (CJPE) over 50 random inspection processes using the proposed RMGP framework, and the RF and OLP models for three time steps with increasing amounts of inspection data available for training. The solid black line illustrates the CJPE obtained with the prior risk model.

5.4. Summary and Limitations

Comparing the results of all three case studies, Table 3 summarizes Figures 6, 10, and 11 by stating the mean, minimum and maximum CJPE values over 50 randomly sampled inspection processes. For a more transparent comparison, the CJPE is normalized by the number of exposed buildings in the respective region and the best-performing methodology is highlighted in bold for each time step. The potential of the proposed RMGP framework to reduce errors and uncertainty in the geographical distribution of earthquake-induced building damage is confirmed in all three case studies.

The first, simulated, Zurich case study shows that despite a relatively dense seismic network, the initial predictions from the assumed prior risk model may be inaccurate. Updating the components of the risk model using early-arriving building damage inspection data rapidly captures the actual trend in the data and thus, provides reliable regional damage estimates after a fraction of the time that would be required to inspect the entire city. This is also confirmed by the case of Pollino, where the initial risk-model predictions are more accurate despite the less dense seismic network. While the numbers of inspected buildings at each time step are similar for the Zurich and Pollino case studies, we used substantially less inspection data in the case of Kraljevo. The first time step is limited to 40 inspected buildings, which translates to a larger span in CJPEs from different inspection sequences.

In all case studies, the proposed RMGP framework performs better than purely data-driven models, such as RF. In addition, for the first time step ($t = 1$), with very limited data, the RMGP framework performs significantly better than the data-driven models. In general, we focus on the early aftermath, with very few inspected buildings (no more than 600), which undermines data-driven methods that typically require large training sets. In addition, GP models account for dependencies in the predictions of multiple, spatially distributed outputs. This is particularly important in cases where aggregated damage statistics are of interest.

Table 3: Summary of the Cumulative Joint Prediction Errors (CJPEs) for all three case studies, using the proposed RMGP framework and the RF and OLP data-drive model, without inspection data ($t = 0$) and using data gathered in three inspection time-steps t after the event. The indicated values are the mean (min., max.) over 50 random inspection processes.

Cumulative Joint Predictive Error (CJPE) in % of the number of exposed buildings					
Case study	Model	$t = 0$	$t = 1$	$t = 2$	$t = 3$
Zurich[1]	RMGP	26.1	**8.8** (**6.1**, **14.6**)	**7.1** (**5.0**, **10.3**)	**6.3** (**4.3**, **8.8**)
	RF		14.1 (9.2, 22.5)	10.8 (7.8, 14.2)	9.7 (7.2, 12.9)
	OLP		14.8 (10.6, 22.4)	12.7 (10.6, 15.5)	12.1 (10.6, 15.0)
Pollino[2]	RMGP	17.0	**9.5** (**6.6**, **14.9**)	**7.9** (**5.3**, **12.0**)	**7.0** (5.0, **10.2**)
	RF		15.0 (9.1, 22.4)	11.0 (7.2, 14.3)	9.1 (**4.7**, 14.4)
	OLP		12.7 (8.6, 20.8)	10.7 (8.7, 14.2)	10.0 (8.0, 12.6)
Kraljevo[3]	RMGP	24.7	**13.9** (**7.3**, **34.7**)	**11.4** (6.3, **20.9**)	**9.7** (**4.7**, **14.2**)
	RF		21.7 (8.6, 53.5)	14.5 (**3.6**, 26.4)	11.2 (5.7, 23.5)
	OLP		25.5 (9.4, 47.1)	15.0 (8.4, 29.0)	11.9 (5.5, 22.3)

[1] Available inspection data for the three time steps: 175 (0.5%), 350 (1.0%), 525 (1.5%).
[2] Available inspection data for the three time steps: 200 (1.0%), 400 (2.0%), 600 (3.0%).
[3] Available inspection data for the three time steps: 40 (2.0%), 80 (4.0%), 120 (6.0%).

It should be noted that the RFs, used for comparison in this paper, are found in several post-earthquake damage prediction studies found in the literature. Considering the rapid evolution of machine-learning tools, better RF models or models based on other approaches, such as neural networks, may outperform the herein adopted versions of data-driven models.

The indicated performance of the initial rapid damage estimates ($t = 0$) naturally depend on the assumed prior risk models and do not allow for any conclusions on how official national risk models might perform in future earthquake events. Such models might profit from more detailed information on the exposed building stock and can incorporate better knowledge from local experts.

The RMGP framework involves a bottom-up earthquake risk model and thus, may prove to have a limited applicability in low-income countries, where geo-coordinates of buildings are not always complete, or even available. However, grass-root exposure modelling [69] may contribute overcoming this limitation.

The case studies idealize the visual inspection outcome as accurate and unbiased. Thus, subjectivity in visual inspections and damage state definitions is not accounted for. However, the good performance of the proposed RMGP framework in the real cases, Pollino and Kraljevo, underlines its applicability to datasets stemming from human inspectors.

The dataset of buildings for which inspection data becomes available in the early aftermath is based on random starting points for inspection teams and subsequent choice of the nearest buildings. However, especially in subregions less affected by the earthquake, inspections may be performed only on building owner demand and are thus limited to buildings that have a higher a-priori probability of being damaged than the randomly chosen buildings. This introduces a sampling bias in the dataset and may undermine the use of machine-learning tools. To overcome this issue, either random buildings need to be inspected in addition to those with an open inspection demand, or special model training schemes need to be implemented. In the case of the RMGP framework, this could be achieved by conditioning the vulnerability functions on the event that the owner reported building damage. Such conditional functions would be used for inference and damage predictions for buildings that requested an inspection, whereas the unconditional functions would be used to predict damage to buildings for which an inspection is not yet requested. We leave this topic for our future work.

6. Conclusion

This paper presents a dynamic updating framework for regional post-earthquake damage estimation, RMGP, that leverages early-arriving observations of ground motion intensity and building damage using Gaussian processes. The framework relies on Gaussian process models to update the ground motion intensity, building type attribution and building type vulnerability components of a regional earthquake risk model that is established prior to the event. Three case studies, focused on regional building damage estimation and updating, illustrate the proposed framework. Our conclusions are:

- RMGP framework provides a powerful tool to combine inspection evidence with prior estimates of traditional risk models, thus enabling precise and accurate predictions of earthquake-inflicted impacts on residential buildings in a fraction of the time required to inspect the entire building stock.
- The parallel updating of ground motion intensity estimates, building vulnerability functions, and building typological attribution models enable a local calibration of prior risk models with broader geographical scopes. The produced byproducts, such as the updated building vulnerability functions, provide useful information for risk modellers beyond rapid damage assessment.
- Using GP models to fuse the early-arriving inspection data with prediction models for probabilistic regional risk analysis requires less data compared to the purely data-driven machine-learning approaches, such as random forest models, to provide these estimates.
- Compared to purely data-driven approaches, predictions obtained using the proposed RMGP framework are more robust with respect to changing inspection sequences and do not require prior optimization or allocation of buildings for priority inspection.

The promising results from the application of the proposed GP dynamic damage updating framework to three case studies highlight its potential for operational use after future earthquakes. Yet, this requires close collaboration between risk modellers and agencies responsible for planning the inspection process. Future work may involve understanding the minimum amount of inspection data that is required to yield stable and reliable updating results, especially in presence of possibly biased inspection results. A further challenge lies in effectively combining the dynamically improving damage estimates with damage-to-loss and recovery models to examine the effect of rapid data collection on financial loss predictions and recovery forecasts. This challenge can be extended to include infrastructure systems that provide essential community services, such as electric power, potable water or communication. Finally, applying the proposed RMGP framework to more complete data, such as that available at later stages of the inspection process, may offer a robust and automated methodology to derive empirical region-specific building vulnerability curves.

Supplementary Material

Supplementary material to reproduce the presented case studies is available in `https://gitlab.ethz.ch/bodlukas/earthquake-rmgp`. It includes detailed descriptions of the prior risk models employed in the three case studies, a thorough explanation on the data pre-processing for the Pollino case study, as well as an illustration of the GP inference scheme.

CRediT authorship contribution statement

Lukas Bodenmann: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Visualization, Writing – original draft preparation, Writing – review & editing. Yves Reuland: Conceptualization, Visualization, Writing – original draft preparation, Writing – review & editing. Božidar Stojadinović: Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Acknowledgements

The work presented in this paper was financially supported by the Real-time Earthquake Risk Reduction for a Resilient Europe 'RISE' project, financed by the European Union's Horizon 2020 research and innovation program under grant agreement No 821115, as well as the ETH Risk Center project 'DynaRisk', financed under grant agreement ETH-11 18-1.

References

[1] S. H. Potter, J. S. Becker, D. M. Johnston, K. P. Rossiter, An overview of the impacts of the 2010-2011 Canterbury earthquakes, International Journal of Disaster Risk Reduction 14 (2015) 6–14. doi:10.1016/J.IJDRR.2015.01.014.

[2] S. Amin, M. Goldstein, Data Against Natural Disasters : Establishing Effective Systems for Relief, Recovery, and Reconstruction, World Bank, Washington D.C., 2008. URL: http://hdl.handle.net/10986/6511.

[3] D. A. Mcentire, J. Cope, Damage Assessment After the Paso Robles (San Simeon, California) Earthquake: Lessons for Emergency Management, 2004.

[4] D. Lallemant, R. Soden, S. Rubinyi, S. Loos, K. Barns, G. Bhattacharjee, Post-Disaster Damage Assessments as Catalysts for Recovery: A Look at Assessments Conducted in the Wake of the 2015 Gorkha, Nepal, Earthquake:, Earthquake Spectra 33 (2017) S435–S451. doi:10.1193/120316EQS222M.

[5] F. Marquis, J. J. Kim, K. J. Elwood, S. E. Chang, Understanding post-earthquake decisions on multi-storey concrete buildings in Christchurch, New Zealand, Bulletin of Earthquake Engineering 15 (2017) 731–758. doi:10.1007/S10518-015-9772-8/FIGURES/5.

[6] K. Kusunoki, Damage assessment in japan and potential use of new technologies in damage assessment., 2021. doi:10.1007/97830306881342.

[7] M. Erdik, K. Şeşetyan, M. B. Demircioĝlu, U. Hancilar, C. Zülfikar, Rapid earthquake loss assessment after damaging earthquakes, Soil Dynamics and Earthquake Engineering 31 (2011) 247–266. doi:10.1016/j.soildyn.2010.03.009.

[8] S. Guérin-Marthe, P. Gehl, C. Negulescu, S. Auclair, R. Fayjaloun, Rapid earthquake response: The state-of-the art and recommendations with a focus on European systems, International Journal of Disaster Risk Reduction 52 (2021) 101958. doi:10.1016/J.IJDRR.2020.101958.

[9] V. Silva, D. Amo-Oduro, A. Calderon, C. Costa, J. Dabbeek, V. Despotaki, L. Martins, M. Pagani, A. Rao, M. Simionato, D. Viganò, C. Yepes-Estrada, A. Acevedo, H. Crowley, N. Horspool, K. Jaiswal, M. Journeay, M. Pittore, Development of a global seismic risk model, Earthquake Spectra 36 (2020) 372–394. doi:10.1177/8755293019899953.

[10] M. Dolce, A. Prota, B. Borzi, F. da Porto, S. Lagomarsino, G. Magenes, C. Moroni, A. Penna, M. Polese, E. Speranza, G. M. Verderame, G. Zuccaro, Seismic risk assessment of residential buildings in Italy, Bulletin of Earthquake Engineering 19 (2021) 2999–3032. doi:10.1007/s10518-020-01009-5.

[11] D. J. Wald, V. Quitoriano, T. H. Heaton, H. Kanamori, C. W. Scrivner, C. B. Worden, TriNet "ShakeMaps": Rapid Generation of Peak Ground Motion and Intensity Maps for Earthquakes in Southern California, Earthquake Spectra 15 (1999) 537–555. URL: http://journals.sagepub.com/doi/10.1193/1.1586057. doi:10.1193/1.1586057.

[12] H. Y. Noh, D. G. Lignos, K. K. Nair, A. S. Kiremidjian, Development of fragility functions as a damage classification/prediction method for steel moment-resisting frames using a wavelet-based damage sensitive feature, Earthquake Engineering & Structural Dynamics 41 (2012) 681–696. doi:10.1002/EQE.1151.

[13] Y. Reuland, P. Lestuzzi, I. F. C. Smith, A model-based data-interpretation framework for post-earthquake building assessment with scarce measurement data, Soil Dynamics and Earthquake Engineering 116 (2019) 253–263. doi:10.1016/J.SOILDYN.2018.10.008.

[14] P. F. Giordano, M. P. Limongelli, The value of structural health monitoring in seismic emergency management of bridges, Structure and Infrastructure Engineering (2020). doi:10.1080/15732479.2020.1862251.

[15] J. A. Goulet, C. Michel, A. D. Kiureghian, Data-driven post-earthquake rapid structural safety assessment, Earthquake Engineering & Structural Dynamics 44 (2015) 549–562. doi:10.1002/EQE.2541.

[16] F. Dell'Acqua, P. Gamba, Remote Sensing and Earthquake Damage Assessment: Experiences, Limits, and Perspectives, Proceedings of the IEEE 100 (2012) 2876–2890. doi:10.1109/JPROC.2012.2196404.

[17] E. Booth, K. Saito, R. Spence, G. Madabhushi, R. T. Eguchi, Validating Assessments of Seismic Damage Made from Remote Sensing:, Earthquake Spectra 27 (2011). doi:10.1193/1.3632109.

[18] S. G. Paal, M. Asce, J.-S. Jeon, I. Brilakis, A. M. Asce, R. Desroches, Automated Damage Index Estimation of Reinforced Concrete Columns for Post-Earthquake Evaluations, Journal of Structural Engineering 141 (2014) 4014228. doi:10.1061/(ASCE)ST.1943-541X.0001200.

[19] N. Kerle, F. Nex, M. Gerke, D. Duarte, A. Vetrivel, UAV-Based Structural Damage Mapping: A Review, ISPRS International Journal of Geo-Information 9 (2019) 14. doi:10.3390/ijgi9010014.

[20] M. Kovačević, Z. Stojadinović, D. Marinković, B. Stojadinović, Sampling and machine learning methods for a rapid earthquake loss assessment system, in: 11th U.S. National Conference on Earthquake Engineering, Earthquake Engineering Research Institute, Los Angeles, CA, 2018.

[21] Z. Stojadinovic, M. Kovacevic, D. Marinkovic, B. Stojadinovic, Rapid earthquake loss assessment based on machine learning and representative sampling:, Earthquake Spectra (2021). doi:10.1177/87552930211042393.

[22] S. Loos, D. Lallemant, J. Baker, J. McCaughey, S. H. Yun, N. Budhathoki, F. Khan, R. Singh, G-DIF: A geospatial data integration framework to rapidly estimate post-earthquake damage, Earthquake Spectra 36 (2020) 1695–1718. doi:10.1177/8755293020926190.

[23] M. Sheibani, G. Ou, The development of Gaussian process regression for effective regional post-earthquake building damage inference, Computer-Aided Civil and Infrastructure Engineering 36 (2021) 264–288. doi:10.1111/MICE.12630.

[24] M. Sheibani, G. Ou, Adaptive local kernels formulation of mutual information with application to active post-seismic building damage inference, Reliability Engineering & System Safety 215 (2021). doi:https://doi.org/10.1016/j.ress.2021.107915.

[25] M. Pozzi, Q. Wang, Gaussian Process Regression and Classification for Probabilistic Damage Assessment of Spatially Distributed Systems, KSCE Journal of Civil Engineering 22 (2018) 1016–1026. doi:10.1007/s12205-018-0014-x.

[26] D. Straub, A. Der Kiureghian, Improved seismic fragility modeling from empirical data, Structural Safety 30 (2008) 320–336. doi:https://doi.org/10.1016/j.strusafe.2007.05.004.

[27] P. Gehl, F. Cavalieri, P. Franchin, Approximate bayesian network formulation for the rapid loss assessment of real-world infrastructure systems, Reliability Engineering & System Safety 177 (2018) 80–93. doi:https://doi.org/10.1016/j.ress.2018.04.022.

[28] M. Bensi, A. Der Kiureghian, D. Straub, Bayesian network modeling of correlated random variables drawn from a gaussian random field, Structural Safety 33 (2011) 317–332. doi:https://doi.org/10.1016/j.strusafe.2011.05.001.

[29] C. Rasmussen, K. Williams, Gaussian processes for machine learning, The MIT Press, Cambridge, MA, 2006.

[30] M. Kuss, Gaussian Process Models for Robust Regression, Classification, and Reinforcement Learning, Ph.D. thesis, Technische Universität Darmstadt, Darmstadt, Germany, 2006.

[31] G. Matheron, The intrinsic random functions and their applications, Advances in Applied Probability 5 (1973) 439–468. doi:10.2307/1425829.

[32] J. Hensman, A. Matthews, Z. Ghahramani, Scalable Variational Gaussian Process Classification, in: Proceedings of the Eigtheenth International Conference on Artificial Intelligence and Statistics, Proceedings of Machine Learning Research, 2015, pp. 351–360. URL: https://proceedings.mlr.press/v38/hensman15.html.

[33] W. Chu, Z. Ghahramani, Gaussian Processes for Ordinal Regression, Journal of Machine Learning Research 6 (2005) 1019–1041. URL: http://www.jmlr.org/papers/volume6/chu05a/chu05a.pdf.
[34] M. Opper, C. Archambeau, The variational gaussian approximation revisited, Neural Computation 21 (2009) 786–792. doi:10.1162/neco.2008.08-07-592.
[35] L. Martins, V. Silva, Development of a fragility and vulnerability model for global seismic risk analyses, Bulletin of Earthquake Engineering (2020). doi:10.1007/s10518-020-00885-1.
[36] L. Diana, J. Thiriot, Y. Reuland, P. Lestuzzi, Application of Association Rules to Determine Building Typological Classes for Seismic Damage Predictions at Regional Scale: The Case Study of Basel, Frontiers in Built Environment 5 (2019). doi:10.3389/fbuil.2019.00051.
[37] H. Crowley, V. Despotaki, D. Rodrigues, V. Silva, D. Toma-Danila, E. Riga, A. Karatzetzou, S. Fotopoulou, Z. Zugic, L. Sousa, S. Ozcebe, P. Gamba, Exposure model for European seismic risk assessment, Earthquake Spectra 36 (2020) 252–273. doi:10.1177/8755293020919429.
[38] H. V. Burton, S. B. Miles, H. Kang, Integrating Performance-Based Engineering and Urban Simulation to Model Post-Earthquake Housing Recovery:, Earthquake Spectra 34 (2019) 1763–1785. doi:10.1193/041017EQS067M.
[39] K. Goda, H. P. Hong, Spatial Correlation of Peak Ground Motions and Response Spectra, Bulletin of the Seismological Society of America 98 (2008) 354–365. doi:10.1785/0120070078.
[40] N. Jayaram, J. W. Baker, Statistical Tests of the Joint Distribution of Spectral Acceleration Values, Bulletin of the Seismological Society of America 98 (2008) 2231–2243. doi:10.1785/0120070208.
[41] Loth, Christophe and Baker, Jack W., A spatial cross-correlation model of spectral accelerations at multiple periods, Earthquake Engineering & Structural Dynamics 42 (2013) 397–417. doi:10.1002/eqe.2212.
[42] D. D'Ayala, A. Meslem, D. Vamvatsikos, K. Porter, T. Rossetto, V. Silva, Guidelines for Analytical Vulnerability Assessment - Low/Mid-Rise, Technical Report, Vulnerability Global Component Project, 2015. doi:10.13117/GEM.VULN-MOD.TR2014.12.
[43] S. Lagomarsino, S. Giovinazzi, Macroseismic and mechanical models for the vulnerability and damage assessment of current buildings, Bulletin of Earthquake Engineering 4 (2006) 415–443. doi:10.1007/s105180069024z.
[44] A. Rosti, C. Del Gaudio, M. Rota, P. Ricci, M. Di Ludovico, A. Penna, G. M. Verderame, Empirical fragility curves for Italian residential RC buildings, Bulletin of Earthquake Engineering 19 (2021) 3165–3183. doi:10.1007/s10518-020-00971-4.
[45] L. Bodenmann, Y. Reuland, B. Stojadinovic, Dynamic Updating of Building Loss Predictions Using Regional Risk Models and Conventional Post-Earthquake Data Sources, in: Proceedings of the 31st European Safety and Reliability Conference (ESREL 2021), 2021, pp. 1411–1418. doi:10.3850/978-981-18-2016-8{_}559-cd.
[46] A. Matthews, M. van der Wilk, T. Nickson, K. Fujii, A. Boukouvalas, P. León-Villagrá, Z. Ghahramani, J. Hensman, GPflow: A Gaussian Process Library using TensorFlow, Journal of Machine Learning Research 18 (2017) 1–6. URL: http://jmlr.org/papers/v18/16-537.html.
[47] C. Worden, E. M. Thompson, J. W. Baker, B. A. Bradley, N. Luco, D. J. Wald, Spatial and spectral interpolation of ground-motion intensity measure observations, Bulletin of the Seismological Society of America 108 (2018) 866–875. doi:10.1785/0120170201.
[48] Eurocode 8, Design of structures for earthquake resistance - Part 1: General rules, seismic actions and rules for buildings, 2004.
[49] Amt für Raumentwicklung, Geografisches Informationssystem des Kantons Zürich (GIS-ZH), 2020. URL: http://maps.zh.ch.
[50] S. Akkar, J. J. Bommer, Empirical Equations for the Prediction of PGA, PGV, and Spectral Accelerations in Europe, the Mediterranean Region, and the Middle East, Seismological Research Letters 81 (2010) 195–206. doi:10.1785/gssrl.81.2.195.
[51] S. Esposito, I. Iervolino, PGA and PGV Spatial Correlation Models Based on European Multievent Datasets, Bulletin of the Seismological Society of America 101 (2011) 2532–2541. doi:10.1785/0120110117.
[52] B. S.-J. Chiou, R. R. Youngs, Update of the Chiou and Youngs NGA Model for the Average Horizontal Component of Peak Ground Motion and Response Spectra, Earthquake Spectra 30 (2014) 1117–1153. doi:10.1193/072813EQS219M.
[53] M. Markhvida, L. Ceferino, J. W. Baker, Modeling spatially correlated spectral accelerations at multiple periods using principal component analysis and geostatistics, Earthquake Engineering & Structural Dynamics 47 (2018) 1107–1123. doi:10.1002/eqe.3007.
[54] J. W. Baker, A. Cornell, Correlation of Response Spectral Values for Multicomponent Ground Motions, Bulletin of the Seismological Society of America 96 (2006) 215–227. doi:10.1785/0120050060.
[55] L. Luzi, G. Lanzano, C. Felicetta, M. D'Amico, E. Russo, S. Sgobba, F. Pacor, ORFEUS Working Group 5, Engineering Strong Motion Database (ESM) (Version 2.0), 2020. doi:10.13127/ESM.2.

[56] M. Dolce, E. Speranza, F. Giordano, B. Borzi, F. Bocchi, C. Conte, A. D. Meo, M. Faravelli, V. Pascale, Observed damage database of past Italian earthquakes: The Da.D.O. WebGIS, Bollettino di Geofisica Teorica ed Applicata 60 (2019) 141–164. doi:10.4430/bgta0254.

[57] C. Corbane, F. Sabo, European Settlement Map from Copernicus Very High Resolution data for reference year 2015, Public Release 2019, 2019. doi:10.2905/8BD2B792-CC33-4C11-AFD1-B8DD60B44F3B.

[58] ISTAT, Censimento generale della popolazione e delle abitazioni, 2001. URL: http://dawinci.istat.it/jsp/MD/.

[59] G. Forte, E. Chioccarelli, M. De Falco, P. Cito, A. Santo, I. Iervolino, Seismic soil classification of Italy based on surface geology and shear-wave velocity measurements, Soil Dynamics and Earthquake Engineering 122 (2019) 79–93. doi:10.1016/j.soildyn.2019.04.002.

[60] D. Bindi, F. Pacor, L. Luzi, R. Puglia, M. Massa, G. Ameri, R. Paolucci, Ground motion prediction equations derived from the Italian strong motion database, Bulletin of Earthquake Engineering 9 (2011) 1899–1920. doi:10.1007/s10518-011-9313-z.

[61] A. Rosti, M. Rota, A. Penna, Empirical fragility curves for Italian URM buildings, Bulletin of Earthquake Engineering 19 (2021) 3057–3076. doi:10.1007/s10518-020-00845-9.

[62] Seismological Survey of Serbia, National Earthquake Catalogue, Technical Report, 2017. URL: http://www.seismo.gov.rs/Seizmicnost/Katalog-zemljotresa.pdf.

[63] S. Roeslin, Q. Ma, H. Juárez-Garcia, A. Gómez-Bernal, J. Wicker, L. Wotherspoon, A machine learning damage prediction model for the 2017 Puebla-Morelos, Mexico, earthquake:, Earthquake Spectra 36 (2020) 314–339. doi:10.1177/8755293020936714.

[64] L. Breiman, Random Forests, Machine Learning 45 (2001) 5–32. doi:10.1023/A:1010933404324.

[65] L. Breiman, J. H. Friedman, R. A. Olshen, C. J. Stone, Classification And Regression Trees, 1984. doi:10.1201/9781315139470.

[66] H. Hersbach, Decomposition of the Continuous Ranked Probability Score for Ensemble Prediction Systems, Weather and Forecasting 15 (2000) 559–570. doi:10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2.

[67] T. Gneiting, A. E. Raftery, Strictly Proper Scoring Rules, Prediction, and Estimation, Journal of the American Statistical Association 102 (2007) 359–378. doi:10.1198/016214506000001437.

[68] T. Gneiting, M. Katzfuss, Probabilistic Forecasting, Annual Review of Statistics and Its Application 1 (2014) 125–151. doi:10.1146/annurev-statistics-062713-085831.

[69] C. Scaini, A. Peresan, A. Tamaro, V. Poggi, C. Barnaba, Can high-school students contribute to seismic risk mitigation? Lessons learned from the development of a crowd-sourced exposure database, International Journal of Disaster Risk Reduction 69 (2022) 102755. doi:10.1016/J.IJDRR.2021.102755.