



Rapid earthquake loss updating of spatially distributed systems via sampling-based bayesian inference

Pierre Gehl, Rosemary Fayjaloun, Li Sun, Enrico Tubaldi, Caterina Negulescu, Ekin Özer, Dina D'ayala

► To cite this version:

Pierre Gehl, Rosemary Fayjaloun, Li Sun, Enrico Tubaldi, Caterina Negulescu, et al.. Rapid earthquake loss updating of spatially distributed systems via sampling-based bayesian inference. Bulletin of Earthquake Engineering, 2022, 10.1007/s10518-022-01349-4 . hal-03660000

HAL Id: hal-03660000

<https://brgm.hal.science/hal-03660000>

Submitted on 8 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Rapid earthquake loss updating of spatially distributed systems via sampling-based bayesian inference

Pierre Gehl¹ · Rosemary Fayjaloun¹ · Li Sun² · Enrico Tubaldi³ · Caterina Negulescu¹ · Ekin Özer³ · Dina D'Ayala²

Received: 28 September 2021 / Accepted: 3 February 2022 / Published online: 4 March 2022
© The Author(s) 2022

Abstract

Within moments following an earthquake event, observations collected from the affected area can be used to define a picture of expected losses and to provide emergency services with accurate information. A Bayesian Network framework could be used to update the prior loss estimates based on ground-motion prediction equations and fragility curves, considering various field observations (i.e., evidence). While very appealing in theory, Bayesian Networks pose many challenges when applied to real-world infrastructure systems, especially in terms of scalability. The present study explores the applicability of approximate Bayesian inference, based on Monte-Carlo Markov-Chain sampling algorithms, to a real-world network of roads and built areas where expected loss metrics pertain to the accessibility between damaged areas and hospitals in the region. Observations are gathered either from free-field stations (for updating the ground-motion field) or from structure-mounted stations (for the updating of the damage states of infrastructure components). It is found that the proposed Bayesian approach is able to process a system comprising hundreds of components with reasonable accuracy, time and computation cost. Emergency managers may readily use the updated loss distributions to make informed decisions.

Keywords Bayesian inference · Critical infrastructure · Seismic risk · Loss updating · Road network

1 Introduction

Rapid loss assessment following an earthquake event is of utmost importance for crisis managers, in order to quantify the extent of the disaster and localize ‘hot-spots’ before planning adequate emergency measures. Therefore, several rapid response systems have been developed worldwide, as detailed in the review by Guérin-Marthe et al. (2021). Such

✉ Pierre Gehl
p.gehl@brgm.fr

¹ Risks and Prevention Division, BRGM, Orléans, France

² Department of Civil, Environmental and Geomatic Engineering, University College London, London, UK

³ Department of Civil & Environmental Engineering, University of Strathclyde, Glasgow, UK

tools usually couple shake-maps (i.e., the updated field of ground-motion parameters, based on earthquake's characteristics and recordings from seismic stations) with damage and loss assessment models (i.e., fragility or vulnerability functions). While such systems are mostly applied to common buildings and the estimation of casualties, the treatment of the performance loss of critical infrastructure and its consequences remains mostly overlooked. Only a few rapid response systems include the damage estimation of infrastructure components (e.g., USGS ShakeCast system; Lin et al. 2018), and none of them evaluate losses at system level so far. The study by Toma-Danila (2018) proposes an extensive risk assessment of the functionality of road networks, with measures of intervention times for ambulances and firefighters; however post-earthquake field observations are not taken into account. The inherent complexity of infrastructure systems, composed of interdependent components of various types, poses extra challenges in terms of modelling and simulation time (Franchin 2014).

Moreover, Guérin-Marthe et al. (2021) introduce a distinction between two types of rapid response systems:

- Procedures that do not account for near-real time observations: the characteristics of the earthquake (e.g., magnitude, location, style-of-faulting) are just used to generate a ground-shaking map from a ground-motion model (GMM). Then, this map is fed into damage and loss estimation models in order to generate a very quick overview of the situation (TELES, Yeh et al. 2006; QLARM, Trendafiloski et al. 2011).
- Procedures that couple shake-maps of the event with damage and loss estimation models in order to account for near-real time observations (e.g., USGS PAGER, Wald et al. 2010; ELER, Zülfikar et al. 2017; Auclair et al. 2014).

In the latter case, integrating strong-motion recordings of the event into the shake-map provides more accurate estimates of the intensity measures (IMs) at the vulnerable sites, thus leading to more reliable prediction of losses. However, it should be noted that, in most cases, only the mean values of the shake-map are used in the loss assessment step, ignoring the fact that GMM outcomes describe a full probabilistic distribution of IMs, even when conditioned on observations. A rigorous rapid response system should ensure the propagation of the uncertainties due to the estimation of ground shaking up to loss predictions. Therefore, this paper investigates a proof-of-concept rapid response procedure that would integrate the following features, in answer to the aforementioned gaps: (1) loss estimation for built areas and infrastructure systems, (2) integration of various types of observations to constrain the predictions, (3) propagation of all sources of uncertainty, from hazard to losses, and (4) ability to treat real-world systems over large areas.

In order to address points (2) and (3), Bayesian Networks (BNs) have emerged as a very promising mathematical tool, well adapted to seismic risk analyses that mobilize a probabilistic framework and dependencies between many variables (Bensi et al. 2011). The inference operations on a BN enable the combination of the initial estimates (i.e., prior distribution provided by predictive models) and of field observations (i.e., providing evidence at some nodes) in order to generate updated posterior distributions of the variables of interest, under Bayes' rule. Therefore, thanks to their probabilistic inference capabilities, BNs constitute an adequate solution for the updating of damage and loss estimates in the crisis phase (Bensi et al. 2015). Regarding the ground-shaking part, a shake-map algorithm based on Gaussian BNs has also been proposed by Gehl et al. (2017). Bensi et al. (2013) have demonstrated the application of BN modelling

techniques to the seismic risk assessment of infrastructure systems, thus ensuring the feasibility of point (1).

However, physical infrastructure components are usually strongly interconnected, and most of them contribute to the performance of the system (e.g., sets of bridges contributing to numerous possible routes in a road network). As a result, evaluation of the system performance measure (S) based on the combination of the states of n components is one typical case of dimensionality curse, as the size of the system node S (i.e., linked to the number of parent nodes) grows exponentially with n (i.e., combinatorial explosion). Such limitations hinder the application of exact inference algorithms to large real-world systems (Cavalieri et al. 2017; Gehl et al. 2018), i.e. point (4).

Bensi et al. (2013) have explored various strategies based on the BN's topology in order to alleviate the scalability issue: they advocate the grouping of components into parallel or series sub-systems through the identification of minimum link sets (MLSs) or minimum cut sets, thus limiting the amount of edges converging towards a given node. These BN formulations are all based on discrete variables: continuous variables (e.g., IM) are discretized into interval sets in order to be combined with variables that have discrete states by default (e.g., damage states of a component). More recent efforts have focused on solving this issue by improving data structures in discrete BNs and reducing memory storage via conditional probability matrices (Byun et al. 2019). This matrix-based BN framework has been further generalized by Byun and Song (2021) to be applied to multi-state systems. Thanks to the decomposition into minimum link sets (MLSs) and the identification of super-components (i.e., sub-sets of components that are only in series or parallel), Applegate and Tien (2019) have also proposed a computationally efficient approach that has the capacity to deal with interdependent infrastructure systems.

Alternative BN frameworks have also been proposed, such as the use of a compression algorithm to reduce the memory storage space of the conditional probability table of the system node S (Tien et al. 2016, 2017). On the other hand, Pozzi and Der Kiureghian (2013) have investigated Gaussian BNs, which consist of continuous variables and have therefore the merit of greatly reducing the sizes of (discrete) conditional probability tables (CPTs). However, this approach requires all variables to be represented by a Gaussian distribution, which is not the case of the components' states (i.e., discrete damage states). This strong limitation prevents the use of exact inference algorithms, and approximate inference engines, such as importance sampling or Gibbs sampling, have to be used instead.

The aforementioned issues have been identified and characterized by Cavalieri et al. (2017), who examined current challenges posed by the BN modelling of real-world infrastructure systems, in terms of both computational complexity and level of system analysis (i.e., formulation of the system performance measure). These considerations have led to the development of an approximate BN approach (Gehl et al. 2018), where an incomplete naïve formulation (i.e., a set of component nodes converging towards the system node) is learned from off-line stochastic loss scenarios: although less straightforward to implement, this method is able to deal with large real-world systems and to estimate elaborate system performance measures (i.e., not limited to connectivity analyses). However, this approach is based on the sampling of numerous earthquake events in the exposed area, based on the regional seismicity, which is not required in a rapid response context where the earthquake parameters (location, magnitude) are usually known with reasonable accuracy within minutes (Cremen and Galasso 2020). Moreover, the corresponding BN uses discretized variables in an exact inference scheme, so that the width of the intervals (discrete states) may influence the accuracy of the results.

Most of the aforementioned studies have mostly focused on the component-to-system links (e.g., Byun et al. 2019; Applegate and Tien 2019; Tien et al. 2016), assuming for instance that components are root nodes in the BN, associated with a marginal probability of failure, or that a single root node IM points towards all component nodes. However, in the context of post-earthquake rapid assessment, the accurate consideration of hazard-related variables leading to component failure probabilities (and their respective correlation structure) is of utmost importance to reach situation awareness. While the modelling of the spatial correlation of seismic IMs has been extensively described by Bensi et al. (2011), its coupling with component and system variables adds yet another layer of complexity. Moreover, seismic responses of components may also be cross-correlated, further complicating the structure of the BN. If a discrete BN formulation is adopted, assuming 20 interval sets for each IM variable (which is not a very accurate description of all possible values) may lead to CPT sizes of up to 20^n for n vulnerable sites. Strategies to simplify the spatial correlation structure, such as Dunnett-Sobel decomposition (Dunnett and Sobel 1955), may be used; however, they also constitute a form of approximation that becomes less accurate for a large number of sites. As an example, the approximate BN approach by Gehl et al. (2018) has partly solved the component-to-system scalability issue; however, the computational bottleneck has been moved to the joint estimation of IM at the locations of all components in the system.

Therefore, pending the development of *ad hoc* BN algorithms that are able to address some of the scalability issues, it is proposed here to adopt a more pragmatic approach based on a sampling inference algorithm (i.e., Monte-Carlo Markov-Chain sampling). This procedure, implemented in the OpenBUGS software (Lunn et al. 2009), has the benefit of avoiding some of the scalability issues and of considering different types of variables (continuous and discrete). The objective here is not to introduce new inference algorithms, but rather to exploit state-of-the-art techniques in order to demonstrate the use of BNs in an operational capacity during the rapid response phase. To this end, the following features are put forward in the proposed loss updating procedure:

- Joint estimation of the probabilistic distribution of damage to residential buildings and connectivity loss of a road network;
- Integration of two types of observations (i.e., evidence), namely the knowledge of the IM at some locations of the exposed area (either via accelerometric stations or via macroseismic testimonies) and the estimation of the damage state (DS) of some physical components (from structural health monitoring techniques);
- Modelling of the spatial correlation of IMs, as well as the correlation between seismic capacities of components;
- Idealization of the road network into a conceptual *Super-Network*, in order to reduce the number of components to process;
- Decomposition of the system into MLSs for connectivity analysis.

Taken separately, none of the above-listed features constitute original developments; however the main novelty resides in their integration under a unique framework in order to facilitate the treatment of large real-world systems. Section 2 of this study details the proposed BN approach for the loss updating of a road network and a built area. A small numerical example is developed in Sect. 3, in order to demonstrate the approach and check the accuracy of the sampling scheme. Then, Sect. 4 applies the BN framework to a real-world area, namely a road network connecting dozens of municipalities in the Pyrenees mountain range (France). Finally, results are discussed in Sect. 5, where a sensitivity study

is also carried out to investigate the influence of parameters that have been seldom studied, such as the correlation between the seismic capacities of the exposed physical components (i.e., bridges and building typologies).

2 Proposed approach to update losses from observations

This section describes the procedure that is put forward to update losses in a rapid response context, by making use of various types of field observations. First, the main principles are outlined, along with the datasets and models that are required for the updating. Then, the process to conceptualize the physical road network into an abstract *Super-Network* is detailed. Finally, the implementation of the BN in the OpenBUGS software (Lunn et al. 2009) is discussed.

2.1 Main principles of the loss updating procedure

The proposed loss updating procedure, summarized in Fig. 1, is articulated around two main types of data:

- ‘Static’ data, which includes the conventional models and datasets that are needed to perform conventional risk assessment: knowing the characteristics of the earthquake

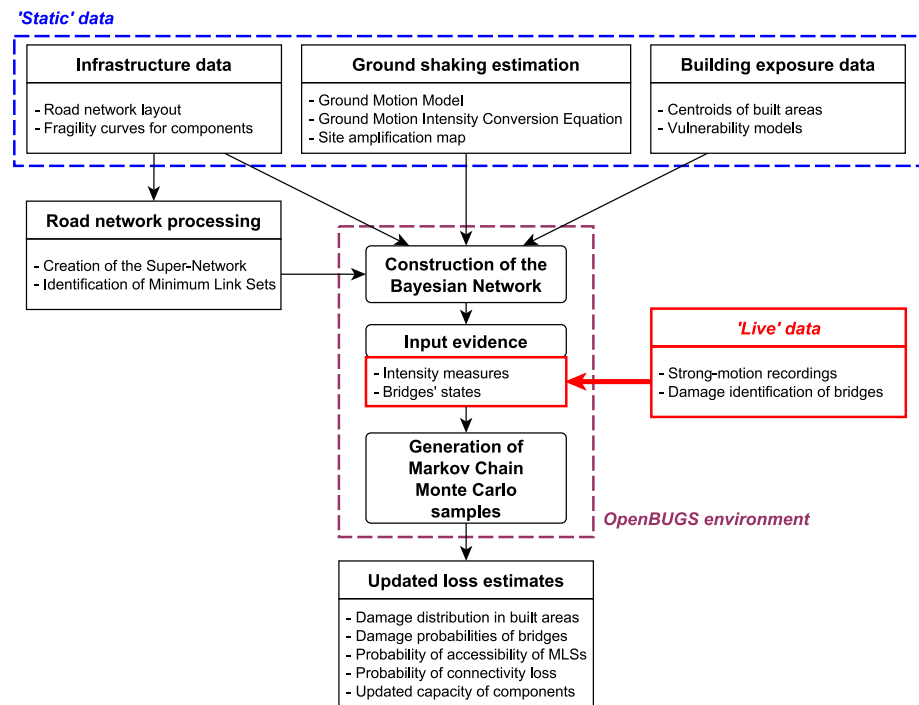


Fig. 1 Proof-of-concept of the procedure for the rapid earthquake loss assessment of built areas and infrastructure systems

(magnitude and location), a GMM is applied to estimate the IM distributions at the locations of the exposed elements, accounting for soil amplification factors. Vulnerability models and fragility curves are also used to predict the damage probabilities of exposed elements, given the IM distributions. Finally, in the case of a road network, the knowledge of the graph topology is also needed in order to access system performance measures such as the network connectivity between two points (Argyroudis et al. 2015). All these models put together result in the prior distributions of the variables of interest (i.e., IM , DS , S).

- ‘Live’ data, which represents the observations (strong-motion recordings and damage identification of bridges) that can be entered as evidence in the variables IM and DS . The objective is to constrain the prior estimates by generating posterior distributions of the variables of interest, given the evidence.

The road network is subjected to a pre-processing step, with the objective of simplifying the initial graph topology into an abstract *Super-Network* that is specifically suited to connectivity analysis. This step has the effect of reducing the network’s size and complexity, thus facilitating the identification of MLSSs. All the data is then entered in a BN structure, which is solved with OpenBUGS: the sampling inference algorithm generates Monte-Carlo Markov Chains (MCMCs), which contain numerous samples of all variables, given the evidence. These samples are finally used to build empirical posterior distributions, with the effect of improving situational awareness for crisis managers. The loss updating procedure is able to generate a wide range of outcomes, such as damage distributions in built areas, damage probabilities of bridges or the probability that the road network connectivity is lost. The seismic response of bridges may also be updated, given the observations of some damaged bridges, so that updated fragility models may be built and applied to subsequent earthquakes.

2.2 Road network processing

The proposed BN model currently relies on the estimation of the connectivity loss of the network, i.e. whether two given locations along the network are still connected or not (binary indicator). Since the analysis is limited to simple connectivity, it is proposed to reduce the complexity of the problem by building a conceptual network (i.e., Super-Network) based on the physical one, via the following steps (see Fig. 2):

- Starting from a physical network with n nodes and m edges (among which, q bridges represent vulnerable edges), only nodes that have a degree equal to 1 (extremities of the network) or strictly above 2 (intersections) are kept as nodes of interest.
- By virtually removing the q vulnerable edges from the adjacency matrix of the network, it is possible to quickly identify all nodes that are accessible from each other without crossing a bridge (i.e., the nodes that stay connected with each other in the altered adjacency matrix).
- These nodes are grouped into a Super-Node: by definition, all nodes belonging to the same Super-Node are always connected with each other. If a bridge is used to connect two nodes within the same Super-Node, it may be removed from the network (such as bridge #4 in Fig. 2). This assumption is only valid when performing a simple connectivity analysis since other performance measures based on travel distance or duration would require the knowledge of alternate paths and so on.

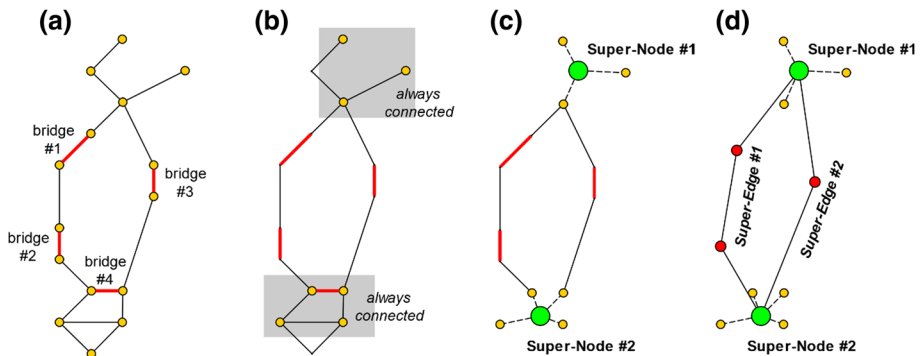


Fig. 2 Illustration of the construction of a Super-Network from a physical network. **a** Initial physical network; **b** Connectivity without the vulnerable edges; **c** Creation of Super-Nodes; **d** Creation of Super-Edges

- (d) The sets of bridges (i.e., vulnerable paths) that are linking two Super-Nodes are identified: this step is facilitated by the fact that only nodes of degree 2 may constitute vulnerable paths (i.e., extremities and intersections are absorbed in Super-Nodes). These paths are referred to as Super-Edges, and they represent sub-systems of bridges in series. Bridges may be then represented by the coordinates of their centroids.

This step leads to a dramatic reduction of the size of the network: in the example in Fig. 2, the initial topology of 15 nodes and 17 edges is simplified into a Super-Network of 2 Super-Nodes and 2 Super-Edges.

In order to identify all MLSS, the recursive algorithm proposed by Cavalieri et al. (2017) has been implemented. It only considers edges (initially, pipelines in a water supply system) as vulnerable components, which is perfectly compatible with the notion of Super-Edge (i.e., representing an in-series system of vulnerable bridges). The recursion starts with the graph G containing all edges and it identifies the shortest path between a predefined origin and destination. Then an edge is removed, and the search for the shortest path continues on the reduced graph G^* . Each time, the shortest path is stored into the list of MLSS, until the whole network has been explored.

2.3 Proposed BN modelling

BNs organize random variables in the form of a directed acyclic graph, where the dependencies between the various nodes are represented by conditional probabilities. An inference is performed on the BN when one or more nodes are observed (i.e., evidence is entered by specifying a given state) and when the probabilities of the other nodes are updated. In the case of a forward analysis, evidence may be entered at the root nodes, and the updated distributions can be estimated at the child nodes (e.g., distribution of infrastructure losses given the occurrence of some natural hazard events). Conversely, a backward analysis consists in the inference of the root nodes based on the observation of a given child node (e.g., updated distribution of the seismic intensity level given the observation of a given loss level).

It is proposed here to build a BN with the OpenBUGS tool (Lunn et al. 2009), which enables the modelling of continuous and discrete variables in the same BN. The OpenBUGS library is freely available and integrated in the R (www.r-project.org)

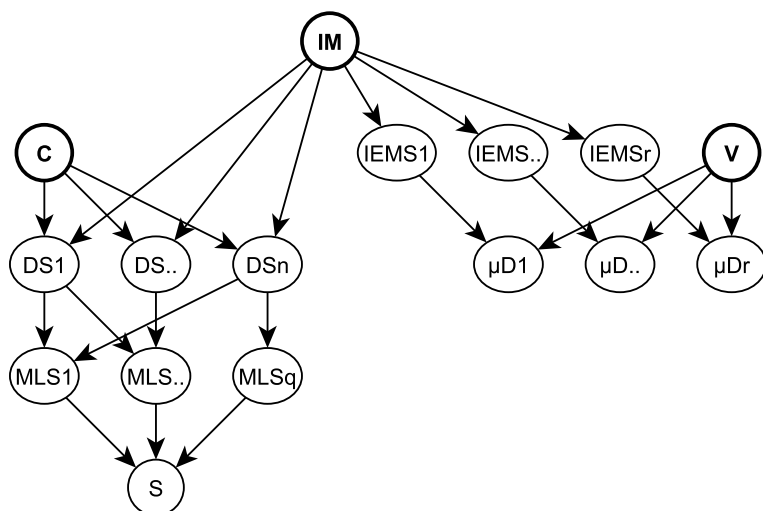


Fig. 3 Proposed BN formulation for a system decomposed into MLSs and for built areas

environment (see demonstration script in the Electronic Supplement of the paper). This tool has been used in various seismic risk applications, such as the derivation of empirical fragility functions for residential buildings (Ioannou et al. 2020), the updating of damage models for water pipelines (Gehl et al. 2021), or the integration of various types of observations for the identification of structural damage (Tubaldi et al. 2021). The trade-off is the use of approximate inference, via MCMC sampling, instead of exact inference such as the junction-tree algorithm (Huang and Darwiche 1996).

For a given earthquake event, the knowledge of epicentre location and magnitude is assumed. Then, as shown in Fig. 3, the variables involved in the loss assessment are the following:

- The vector **IM** represents the distribution of the logarithm of the ground-motion parameter of interest (e.g., peak ground acceleration, PGA) at the locations $i = 1 \dots n$ (for n exposed components) and $j = n+1 \dots n+m$ (for m recordings from seismic stations). It follows a normal distribution of mean $\mu_{\log \mathbf{IM}}$ and covariance $\Sigma_{\mathbf{IM}}$.
- The vector **C** represents the distribution of the logarithm of the seismic response of the n components, also expressed in terms of **IM**. It follows a normal distribution of mean $\mu_{\log \mathbf{C}}$ and covariance $\Sigma_{\mathbf{C}}$.
- The variables DS_i represent the binary damage states (0 = failure, 1 = survival) of the n components. Each DS_i is determined as follows:

$$\begin{cases} DS_i = 0 & \text{if } IM_i > C_i \\ DS_i = 1 & \text{if } IM_i \leq C_i \end{cases} \quad (1)$$

- The variables MLS_k represent the binary states of the q MLSs. In the case of a road network, $MLS_k = 1$ if the MLS route is accessible (i.e., no failed bridges along the itinerary, if we assume that bridges constitute the weak link of the transport infra-

structure), and 0 if not. An MLS_k composed of p components follows the Boolean rule of an in-series system:

$$MLS_k = DS_1 \times \cdots \times DS_p \quad (2)$$

- The variable S represents the performance measure of the system. In the case of a road network, it corresponds to the accessibility between two given points A and B of the network. Thanks to the MLS decomposition, it follows the Boolean rule of an in-parallel system:

$$S = MLS_1 + \cdots + MLS_q \quad (3)$$

With q MLSs, S takes discrete values between 0 and q , thus representing the number of possible routes between A and B.

While the left part of the BN in Fig. 3 is relevant for the performance analysis of an infrastructure system (such as the connectivity between two locations along a road network), it is also possible to expand the study by considering the damage to common buildings. In this case, following the approach detailed in Sedan et al. (2013), built areas are represented as polygons (representing building blocks or city districts) of homogeneous seismic vulnerability. Then, the distribution of damage states within each built area is obtained by applying the semi-empirical vulnerability method by Lagomarsino and Giovinazzi (2006). A vulnerability index V is assigned to each built area, and the mean damage grade μ_D is expressed as a function of the EMS-98 macroseismic intensity I_{EMS} (which is obtained from the distribution of IM via ground-motion intensity conversion equations):

$$\mu_D = 2.5 \left[1 + \tanh \left(\frac{I_{EMS} + 6.25V - 13.1}{2.3} \right) \right] \quad (4)$$

Ultimately, the proportion of damage states in each built area, based on the EMS-98 scale (Grünthal 1998), is obtained by applying a discrete beta distribution based on μ_D (Lagomarsino and Giovinazzi 2006). Then, by considering a vector \mathbf{V} of vulnerability indices, of mean $\mu_{\mathbf{V}}$ and covariance $\Sigma_{\mathbf{V}}$, the BN framework for the damage assessment of common buildings may be defined as in the right part of Fig. 3.

The probabilistic variables (\mathbf{IM} , \mathbf{C} , \mathbf{V}) are entered in the OpenBUGS environment, along with the mathematical expressions (e.g., Eqs. 1–4) that are needed to define deterministic variables (DS , MLS , S , I_{EMS} , μ_D). Evidence may be entered for IM_j (representing ground-motion measurements, as in shake-map procedures), for DS_i (the measurement of the damage state of some components) and μ_D (assuming a field inventory of damages in some built areas).

The Bayesian inference is performed by initiating several MCMC chains, with different combinations of initial conditions (i.e., starting values of the probabilistic variables). Each chain is built with a Gibbs sampling scheme, where variables are successively sampled from the posterior distribution of previous variables: the posterior distribution of a variable is obtained from the product of the prior distribution (e.g., the initial estimate of IM) and the likelihood function (probability of a given observation occurring given the prior distribution).

Within each chain, a large number of samples are generated until convergence is considered to be reached, by running tests as a calibration step as shown in Sect. 5.1. The generated samples are then used to estimate empirical statistics of the variables of interest (i.e., posterior distributions).

2.4 Definition of prior distributions

The prior distributions of the variables that are probabilistically defined (**IM**, **C**, **V**) are assumed to follow normal/lognormal distributions, whose parameters are obtained from predictive uncertain models.

In the case of **IM**, the mean of the logarithm of the ground-motion parameter distribution ($\mu_{\log \text{IM}}$) is given by a ground-motion model (GMM), based on the earthquake characteristics that are assumed to be known with confidence shortly after the event. The covariance matrix Σ_{IM} is assembled as follows:

$$\Sigma_{\text{IM}} = \begin{bmatrix} \sigma_{\eta}^2 + \sigma_{\xi i}^2 & \cdots & \sigma_{\eta}^2 + \rho_{ij} \sigma_{\xi i} \sigma_{\xi j} \\ \vdots & \ddots & \vdots \\ \cdots & \cdots & \sigma_{\eta}^2 + \sigma_{\xi j}^2 \end{bmatrix} \quad (5)$$

where σ_{η} and σ_{ξ} respectively represent the standard deviations of the inter- and intra-event error terms, which are given by the GMM. The term ρ_{ij} represents the spatial correlation of the intra-event error between sites i and j , and it may be defined by available models in the literature (e.g., Jayaram and Baker 2009).

The seismic response **C** of components is provided by fragility curves, where the elements of $\mu_{\log \text{C}}$ correspond to the median fragility, and the standard deviations of **C** correspond to the fragility dispersion β . The dispersion term β may be further decomposed into β_R and β_M , which respectively represent the uncertainty due to record-to-record variability and the uncertainty due to imperfect knowledge or modelling of the component (Crowley et al. 2019). A third type of uncertainty, related to the definition of the damage state threshold, is neglected here for simplification purposes. Therefore, the covariance matrix Σ_{C} is expressed as follows:

$$\Sigma_{\text{C}} = \begin{bmatrix} \beta_{Ri}^2 + \beta_{Mi}^2 & \cdots & \rho_{ij}^R \beta_{Ri} \beta_{Rj} + \rho_{ij}^M \beta_{Mi} \beta_{Mj} \\ \vdots & \ddots & \vdots \\ \cdots & \cdots & \beta_{Rj}^2 + \beta_{Mj}^2 \end{bmatrix} \quad (6)$$

The term ρ_{ij}^R , representing the correlation of the response due to record-to-record variability between components i and j , is very difficult to quantify without the knowledge of the seismic records used in the derivation of the fragility curves. A qualitative rationale may postulate that components i and j – that are spatially very close to each other – may experience ground motion inputs with similar characteristics in terms of duration, frequency content, etc. and therefore their record-to-record variability should be fully correlated. On the other hand, spatially-distant components are likely to be subjected to different ground motions (e.g., different spectral shapes), and therefore their record-to-record variability should be uncorrelated. Such considerations are discussed in Silva (2019), who advocates the use of a correlation structure similar to the one that models the spatial correlation of the intra-event error. Although deserving further investigation, this assumption is also used here for the characterization of ρ_{ij}^R .

The correlation of the component-to-component variability within the same class/typology, represented by ρ_{ij}^M , also requires more knowledge of how the corresponding fragility curves are derived. Therefore, it is proposed to consider the extreme cases in the application (see Sect. 4), namely fully correlated or uncorrelated variability, in order to investigate the impact of these assumptions on the posterior distributions. The decomposition of the dispersion into β_R and β_M is not always detailed in available fragility models (i.e., only

the global dispersion β is specified). Various assumptions regarding this decomposition are also tested in the application example.

Finally, the vulnerability indices \mathbf{V} of built areas are obtained from a field inventory by structural engineers. The standard deviations of the elements of \mathbf{V} may be obtained from the upper and lower bounds of the vulnerability estimates (e.g., see Table 3 of Lagomarsino and Giovinazzi 2006). Regarding the correlation structure, the same issues discussed for the definition of Σ_C hold for Σ_V .

3 Numerical verification of the sampling algorithm

This section investigates the accuracy of the OpenBUGS BN on a synthetic test-case. To this end, a Gaussian BN is introduced as the reference solution for the small studied system.

3.1 Definition of the synthetic system

The accuracy of the sampling-based inference scheme used in OpenBUGS is firstly investigated on a trivial synthetic system. Two components arranged in-series (e.g., two bridges that must be crossed to go from one point to another) are considered: they are respectively 10 and 15 km away from an earthquake epicentre (with assumed magnitude M_w 6.3). A PGA observation is added, located at a site 12.5 km from the epicentre (at an equal distance between the two components). Therefore, the corresponding BN is built for a vector \mathbf{IM} containing 3 elements (the first two representing the sites of the components, and the last one the observation), and a vector \mathbf{C} containing 2 elements (one for each component). The marginal and conditional probabilities of all variables involved in this BN are expressed as follows, using the same convention as in Sect. 2.3 (Eq. 1 to 3):

$$P_{\mathbf{IM}}(\mathbf{im}) = \varphi(\log \mathbf{im}, \boldsymbol{\mu}_{\log \mathbf{IM}}, \boldsymbol{\Sigma}_{\mathbf{IM}}) = \varphi\left(\log \mathbf{im}, \begin{bmatrix} 0.3346 \\ 0.0878 \\ 0.2025 \end{bmatrix}, \begin{bmatrix} 0.1815 & 0.0740 & 0.1132 \\ 0.0740 & 0.1815 & 0.1132 \\ 0.1132 & 0.1132 & 0.1815 \end{bmatrix}\right) \quad (7)$$

$$P_{\mathbf{C}}(\mathbf{c}) = \varphi(\log \mathbf{c}, \boldsymbol{\mu}_{\log \mathbf{C}}, \boldsymbol{\Sigma}_{\mathbf{C}}) = \varphi\left(\log \mathbf{c}, \begin{bmatrix} -0.0083 \\ -0.0083 \end{bmatrix}, \begin{bmatrix} 0.2000 & 0.0400 \\ 0.0400 & 0.2000 \end{bmatrix}\right) \quad (8)$$

where $\varphi(\bullet, \boldsymbol{\mu}, \boldsymbol{\Sigma})$ is the multivariate normal probability density function, of mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$.

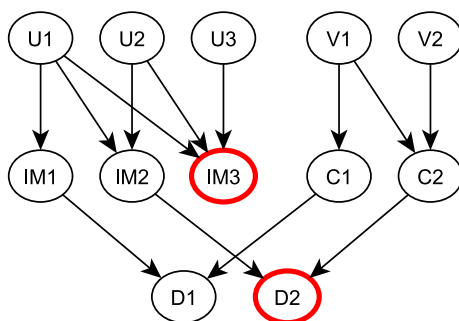
$$\begin{cases} P(DS_1 = ds_1 | im_1, c_1) = (1 - ds_1) \cdot H(im_1 - c_1) + ds_1 \cdot H(c_1 - im_1) \\ P(DS_2 = ds_2 | im_2, c_2) = (1 - ds_2) \cdot H(im_2 - c_2) + ds_2 \cdot H(c_2 - im_2) \end{cases} \quad (9)$$

where $H(\bullet)$ is the Heaviside function, $ds_i = 0$ denotes failure and $ds_i = 1$ denotes survival, for $i = 1, 2$.

$$P(S = s | ds_1, ds_2) = (1 - s) \cdot (1 - ds_1 \cdot ds_2) + s \cdot ds_1 \cdot ds_2 \quad (10)$$

where $S = 0$ denotes disconnection (at least one of the in-series bridges has failed) and $S = 1$ denotes accessibility.

Fig. 4 Gaussian BN used for the numerical verification on the synthetic example. Red circles represent nodes that are evidenced (see Sect. 3.3)



3.2 Verification method

Such a system is small enough to be solved with a BN using an exact inference algorithm such as the junction-tree technique. However, the discretization of all continuous variables (IM , C) also represents a challenge that could lead to inaccuracies in the outcome. Therefore, it is proposed to apply a Gaussian BN to the problem described in Sect. 3.1. A Gaussian BN contains only continuous variables with normal distributions and it also has the benefit of being compatible with exact inference algorithms. The Gaussian BN is presented in Fig. 4, with the following set of variables and their related distributions:

- U_i , $i = 1, 2, 3$: Standard normal variables that are used to model the correlation between the IMs.

$$U_i \sim \mathcal{N}(0, 1) \quad (11)$$

- V_i , $i = 1, 2$: Standard normal variables that are used to model the correlation between the seismic capacities of the components.

$$V_i \sim \mathcal{N}(0, 1) \quad (12)$$

- IM_i , $i = 1, 2, 3$: Logarithm of the IM of interest (here, PGA), which follows a normal distribution where the mean is the linear combination of conditioning variables U_i .

$$IM_i \sim \mathcal{N}\left(IM_{0,i} + \sigma_i \sum_{j=1}^i t_{ij} U_j, \epsilon^2\right) \quad (13)$$

In Equation (13), the terms t_{ij} are elements of the transformation matrix \mathbf{T} , which is used to generate a correlated vector \mathbf{Z} of standard normal variables by transforming the independent variables U_i (assembled into the vector \mathbf{U}): $\mathbf{Z} = \mathbf{T}\mathbf{U}$. The lower triangular matrix \mathbf{T} results from the Cholesky decomposition of the correlation matrix \mathbf{R} , which is obtained from the covariance matrix $\Sigma_{\mathbf{IM}}$ in Eq. (5). $IM_{0,i}$ represents the mean GMM estimation of IM at location i , and σ_i is the standard deviation provided by the GMM.

- C_{ii} , $i = 1, 2$: Logarithm of the seismic capacity of component i , which follows a normal distribution where the mean is the linear combination of conditioning variables V_i .

$$C_i \sim \mathcal{N}\left(C_{0,i} + \beta_i \sum_{j=1}^i t'_{ij} V_j, \epsilon^2\right) \quad (14)$$

In Equation (14), the terms t'_{ij} are elements of the transformation matrix \mathbf{T}' , which is used to generate a correlated vector \mathbf{Z}' of standard normal variables by transforming the independent variables V_i (assembled into the vector \mathbf{V}): $\mathbf{Z}' = \mathbf{T}'\mathbf{V}$. The lower triangular matrix \mathbf{T}' results from the Cholesky decomposition of the correlation matrix \mathbf{R}' , which is obtained from the covariance matrix \sum_C in Eq. (6). $C_{0,i}$ represents the prior mean of the seismic capacity of component i , and β_i is the standard deviation provided by the fragility curve.

- D_i , $i = 1, 2$: Damage indicator of component i , which follows a normal distribution where the mean is the linear combination of conditioning variables IM_i and C_i .

$$D_i \sim \mathcal{N}(C_i - IM_i, \epsilon^2) \quad (15)$$

The damage indicator is a continuous variable, however it may be easily interpreted as a discrete damage state. If D_i is positive, then it means that $C_i > IM_i$, which translates into $DS_i = 1$, as defined in Eq. (1). Conversely, D_i being negative implies that $DS_i = 0$ (i.e., component failure). Therefore, without breaking the Gaussian assumption in the BN, it is possible to access $P_{f,i}$, the probability of failure of component i :

$$P_{f,i} = P(D_i \leq 0) = \Phi\left(-\frac{\mu_{D_i}}{\sigma_{D_i}}\right) \quad (16)$$

where μ_{D_i} and σ_{D_i} are respectively the mean and standard-deviation of variable D_i , which are obtained when solving the Gaussian BN. Φ represents the standard normal cumulative distribution function.

The resulting Gaussian BN in Fig. 4 is implemented in the Bayes Net toolbox (Murphy 2001), which has the ability to apply exact inference (i.e., junction-tree algorithm) to continuous Gaussian variables.

With the proposed Gaussian BN, an issue arises when needing to specify that a given component i has failed or survived as an evidence. Since such an evidence takes the form of an inequality (e.g., $D_i > 0$ for survival), it cannot be entered directly into the BN, which can only treat hard evidence (i.e., a scalar value determining a continuous variable). Therefore it is proposed to decompose the posterior distributions over all possible values of D_i , given the damage state of the bridge. A similar decomposition framework for the treatment of soft evidence has been introduced by Fayjaloun et al. (2021). Let Y be a variable of interest in the BN; then its posterior distribution given the observation of $DS_i = 1$ is decomposed as follows:

$$P(Y|DS_i = 1) = \int_{-\infty}^{+\infty} P(Y|D_i)P(D_i|DS_i = 1)dD_i \quad (17)$$

$P(D_i | DS_i=1)$ represents the conditional probability of observing the value D_i given the damage state of the component: this probability is estimated by using the *a priori* distribution of D_i and by truncating its probability density function, named p_{trunc} (i.e., keeping only the positive part of the support and normalizing the function):

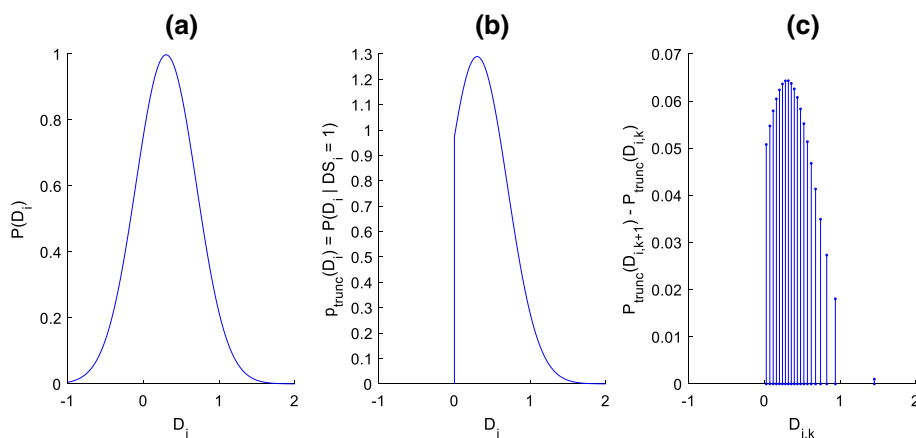


Fig. 5 **a** Prior distribution of D_i ; **b** Truncated and normalized distribution of D_i ; **c** Discretized weights applied to the decomposition

$$P(Y|DS_i = 1) = \int_{-\infty}^{+\infty} P(Y|D_i) p_{trunc}(D_i) dD_i \approx \sum_k P\left(Y|D_i = \frac{D_{i,k+1} + D_{i,k}}{2}\right) [P_{trunc}(D_{i,k+1}) - P_{trunc}(D_{i,k})] \quad (18)$$

where P_{trunc} represents the cumulative distribution function of D_i . If a discretization is applied to the continuous variable D_i , then the integral may be approximated by summing over the discrete intervals $[D_{i,k}; D_{i,k+1}]$, as shown in Fig. 5.

With a very refined discretization, this approach is able to converge quickly towards the exact solution. However, it requires to loop over a large number of BNs (each time defined with a different evidence value), and the computations using such an approach become intractable when several observations need to be entered. Therefore, this method is only used on the trivial synthetic example in order to provide trustworthy results that can be compared to the MCMC sampling-based approach.

3.3 Results

The comparison between the Gaussian BN and the OpenBUGS outcomes is detailed in Table 1, by considering two evidence scenarios: scenario 1 only considers that $\log im_3 = -0.1$; while scenario 2 assumes that an additional type of observation is available, i.e. that component #2 has survived (i.e., $ds_2 = 1$). The R script that builds the corresponding OpenBUGS BN is provided as an Electronic Supplement of the paper, as a demonstrator of the updating procedure. An example of OpenBUGS model is shown in Fig. 6.

Due to the sampling approximation present in both methods, the posterior values cannot be perfectly matched; however, they are in very close agreement, and most discrepancies appear around the 3rd or 4th decimals. Such a level of precision is sufficient for seismic risk applications since it appears that the mismatch is negligible in light of the much greater uncertainties that usually accompany loss assessment models. This numerical test is also an opportunity to verify the soundness of the Bayesian modelling assumptions: all variables are updated given the evidence, and their posterior distribution is shifted in the expected direction (e.g., reduction of the seismic intensity due to

the observation of a low-intensity im_3 , better seismic resistance of the components due to the observation of the survival of component #2, etc.). It is also worth noting that the uncertainty on these variables is gradually reduced as more evidence is entered (i.e., smaller standard-deviations in evidence scenario 2).

A more systematic investigation of the updating capabilities of the BN approach is shown in Fig. 7, where the values of im_3 are gradually increased in evidence scenario 2. The top plots show the evolution of the posterior distributions of variables IM_2 and C_2 , in terms of 16th, 50th and 84th percentiles, obtained with the OpenBUGS approach. The bottom plots show the evolution of the error between the Gaussian BN (GBN) and OpenBUGS (OBN) estimates, which is defined as follows:

$$Err = \frac{E[Y|IM_3 = im_3, DS_2]_{OBN} - E[Y|IM_3 = im_3, DS_2]_{GBN}}{E[Y|IM_3 = im_3, DS_2]_{GBN}} \quad (19)$$

On the right plots of Fig. 7, evidence scenario 2 is modified in order to consider the case where the failure of component #2 is observed.

From Fig. 7, the following points may be noted:

Table 1 Posterior distributions of the variables of interest, for the two evidence scenarios. Observed values (evidence) are in bold font. C_i and IM_i are expressed in $\log m/s^2$

	Prior	Evidence scenario 1 ($im_3 = -0.1$)		Evidence scenario 2 ($im_3 = -0.1, ds_2 = 1$)	
		Gaussian BN	OpenBUGS BN	Gaussian BN	OpenBUGS BN
$P(s = 0)$	0.8320	0.7576	0.7592	0.5717	0.5706
$\overline{im_1} ; \sigma_{im1}$	0.3346 ; 0.4260	0.1459 ; 0.3330	0.1469 ; 0.3334	0.1420 ; 0.3332	0.1408 ; 0.3326
$\overline{im_2} ; \sigma_{im2}$	0.0878 ; 0.4260	- 0.1009 ; 0.3330	- 0.0994 ; 0.3334	- 0.2391 ; 0.2954	- 0.2387 ; 0.2957
$\overline{im_3} ; \sigma_{im3}$	0.2025 ; 0.4260	- 0.1 ; /	- 0.1 ; /	- 0.1 ; /	- 0.1 ; /
$\overline{c_1} ; \sigma_{c1}$	- 0.0083 ; 0.4472	- 0.0083 ; 0.4317	- 0.0083 ; 0.4469	0.0416 ; 0.4433	0.0415 ; 0.4437
$\overline{c_2} ; \sigma_{c2}$	- 0.0083 ; 0.4472	- 0.0083 ; 0.4317	- 0.0081 ; 0.4473	0.2411 ; 0.3510	0.2419 ; 0.3508
$P(ds_1 = 0)$	0.7106	0.6090	0.6101	0.5717	0.5706
$P(ds_2 = 0)$	0.5618	0.4341	0.4350	0	0

- In the case that $DS_2 = 1$ is observed, C_2 is well above IM_2 for low values of im_3 and the gap narrows as im_3 increases. IM_2 is influenced by the evidence of IM_3 and tends to increase, which has the effect of forcing C_2 to increase even more in order to respect the condition $DS_2 = 1$. It has also been found that the correlation between IM_2 and C_2 increases, reflecting the higher constraints that are applied by large values of im_3 .
- In the case that $DS_2 = 0$ is observed, low values of im_3 induce low values of IM_2 as well, forcing C_2 to be reduced dramatically. As im_3 increases, the constraints become more relaxed, and IM_2 increases enough so that C_2 deviates less from its prior distribution.


```

model
{
  IM[1:3] ~ dmnorm(muIM[, ], tau[, ])
  tau[1:3, 1:3] <- inverse(cov_IM[, ])
  C[1:2] ~ dmnorm(muC[, ], tau2[, ])
  tau2[1:2, 1:2] <- inverse(cov_C[, ])

  for (i in 1:2) {
    frag[i] <- phi((C[i] - IM[i])/1.00000E-06)
  }
  DS[1, 1] ~ dbern(frag[1])
  Dobs[1, 1] ~ dbern(frag[2])

  PfSystem <- 1 - DS[1, 1] * Dobs[1, 1]

  prec.IM <- 1/1.00000E-06
  IMobs[1, 1] ~ dnorm(IM[3], prec.IM)
}

```

Fig. 6 OpenBUGS model for scenario 2. Observations (evidence) are the quantities Dobs and IMobs

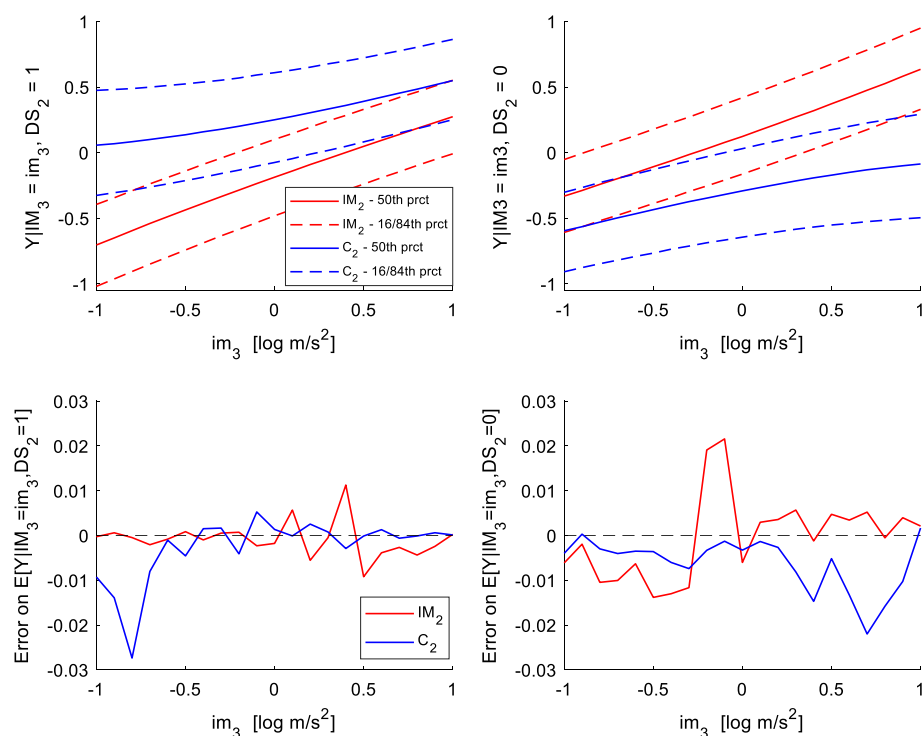


Fig. 7 Top left: posterior distributions of variables IM_2 and C_2 , in terms of 16th, 50th and 84th percentiles, with evidence $IM_3 = im_3$ and $DS_2 = 1$; Top right: posterior distributions of variables IM_2 and C_2 , in terms of 16th, 50th and 84th percentiles, with evidence $IM_3 = im_3$ and $DS_2 = 0$; Bottom: Error rate between the Gaussian BN and OpenBUGS estimates

- Regarding the error rates, no significant biases or trends are observed. The errors seem to be well constrained under 0.03 (i.e., around 3% of deviation at the maximum), confirming the suitability of the sampling-based BN for the problem at hand.

4 Application to a real road network

This section describes the case-study area in the French Pyrenees mountain range and the construction of the corresponding BN model.

4.1 Description of the case-study area

The proposed BN approach for rapid response is applied to a road network composed of 118 bridges (i.e., vulnerable components), which connects 53 municipalities (i.e., built areas) located in a valley around Bagnères-de-Luchon (Pyrenees, France). The case-study area is detailed in Fig. 8.

The typologies of the 118 bridges are identified based on photographs and aerial pictures, and their conditional probability of failure is defined by fragility functions, some of which are taken from the SYNER-G database (Crowley et al. 2011). The most common bridge type consists of short single-span bridges (length < 50 m), followed by masonry arch bridges. In total, 18 different fragility curves have been assigned (3 models for the 83 single-span bridges, 3 for the 7 continuous multi-span bridges, and 12 for the 28 arch bridges). The fragility curve corresponding to the first limit state is considered as the

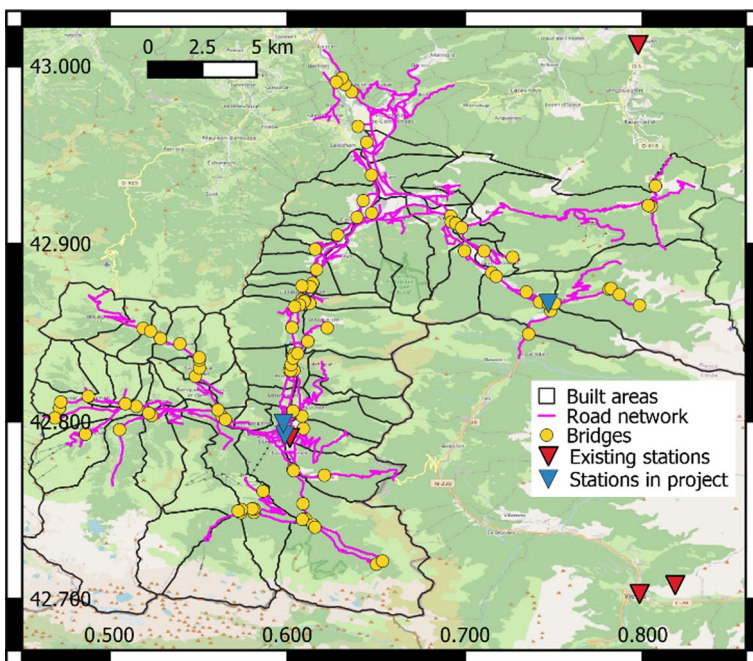


Fig. 8 Situation map of the Luchon case-study area

threshold of the loss of functionality of the bridge (i.e., failure of the component), assuming that even small structural damage might be enough to prevent safe passage.

For the 53 municipalities, the information concerning the characteristics of the residential buildings is extracted from the building census data freely provided by the French National Institute of Statistics and Economic Studies (INSEE). From INSEE data, we are using mainly three items: the building type (e.g. individual building, collective building ...), the number of stories and the construction age period. We associate the age period range of construction to the evolution of the construction techniques and the evolution of seismic design codes. With these data, we classify the different buildings types into the EMS-98 vulnerability classes (Grünthal 1998). For each building type in each municipality, we estimate the associated number of buildings and the associated population. Then, for the building types distributed into the EMS-98 classes, we estimate the RISK-UE vulnerability indices based on the method developed by Lagomarsino and Giovinazzi (2006).

The studied area is surrounded by several seismic stations (see Fig. 8), which are used as sources of observations to constrain estimates of the strong-motion field. All selected fragility curves use PGA as IM; therefore, the regional GMM by Tapia (2006) is applied here for the estimation of the prior distribution of IM. For PGA, the spatial correlation model by Jayaram and Baker (2009) is expressed as:

$$\rho_{ij} = \exp\left(-\frac{3h}{8.5}\right) \quad (20)$$

where h is the inter-site distance in km.

Finally, site effects are modelled via soil amplification factors, which were estimated from local investigations and soil measurements (Roullé et al. 2012).

4.2 Construction of the BN model

The assembled Super-Network is displayed in Fig. 9, where the position of the abstract Super-Nodes is estimated as the mean of the coordinates of the nodes belonging to each Super-Node. This preliminary step leads to a dramatic simplification of the network since the initial 607 nodes and 689 edges are now replaced by 52 Super-Nodes and 67 Super-Edges. Out of the initial 118 bridges, only 96 are kept in the Super-Network, since some of them are not useful to ensure the connectivity between Super-Nodes.

For demonstration purposes, the connectivity between a location A (downtown of Bag-nères-de-Luchon) and a location B (Northern end of the road network) is studied here, in order to simulate the possibility to send rescue teams from the North to the affected area after a potential earthquake (see Fig. 9). Thanks to the Super-Network conceptualisation, the decomposition of all possible paths between A and B becomes computationally manageable: 60 MLSs are identified, mobilising 25 bridges in total (i.e., most bridges are involved in multiple MLSs), as shown in Fig. 10.

Finally, this structure is implemented in OpenBUGS in order to build the corresponding BN, which contains the following nodes:

- The vector **IM**, containing 156 elements (96 bridge sites + 53 built area centroids + 7 observations);
- The vector **C**, containing 96 elements;
- The vector **V**, containing 53 elements;
- 96 *DS* nodes;

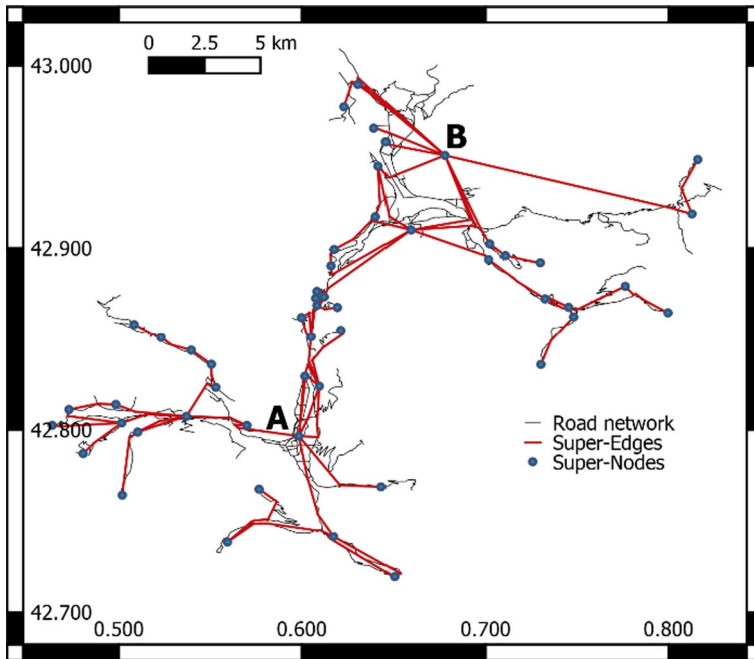


Fig. 9 Construction of the Super-Network from the physical network

- 53 μ_D nodes;
- 60 MLS nodes;
- 1 S node.

The updating capabilities of the BN are tested with a hypothetical Mw 6.3 earthquake scenario, located south of the road network (0.60° lon, 42.65° lat). Hypothetical observations from 7 seismic stations and damage measures on 5 bridges are set as evidence in the BN (2 survivals and 3 failures). It is assumed that one of the monitored bridges belongs to the identified MLS s, and it has been observed as intact (see Fig. 10).

5 Results and discussion

This section details the updated loss distributions obtained from the BN while discussing the impact of various factors, such as the number of MCMC samples or the correlation between the seismic response of the bridges.

5.1 Checking the accuracy of the updated ground-motion field

As the numerical test in Sect. 3 could only be performed on a small example with a reduced number of variables, another test is conducted here on all IM variables involved in the case-study. In the BN, the **IM** vector represents a spatially-correlated Gaussian field, for which analytical solutions have been developed (Vanmarcke 1983; Stafford

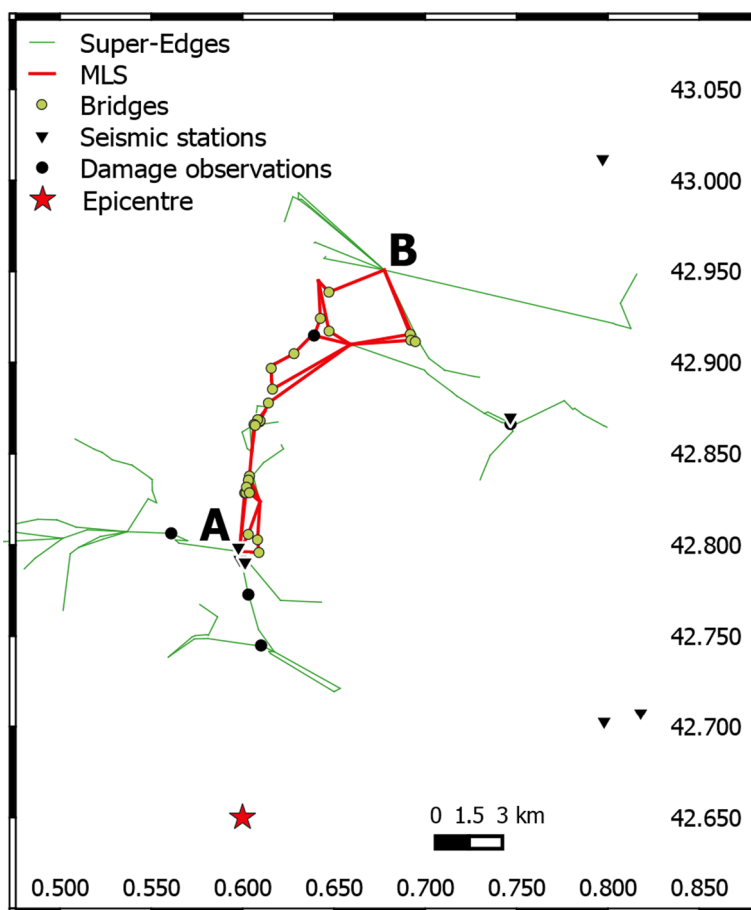


Fig. 10 Summary of the identified MLSs and location of the observations

2012): such an updating procedure is actually adopted in current systems for the generation of shake-maps (Worden et al. 2018), thus constituting the reference solution. Since this solution is only available for the ground-motion field, the IM variables are updated here using only the 7 seismic observations (i.e., no observations of the damage states of bridges). IMs are updated at the locations of the 96 bridges of interest and the centroids of the 53 built areas: the updated distributions are compared to the analytical solution for different sizes of MCMC samples (respectively three chains of 3 000, 5 000, 15 000 and 25 000 samples each, with the removal of a burn-in phase of 2 000 samples in each chain). The errors rates $\varepsilon_{E(IM)}$ and $\varepsilon_{\sigma \log IM}$ of the posterior distributions estimated with the BN with respect to the exact analytical solution, at all vulnerable sites, are summarized in Fig. 11:

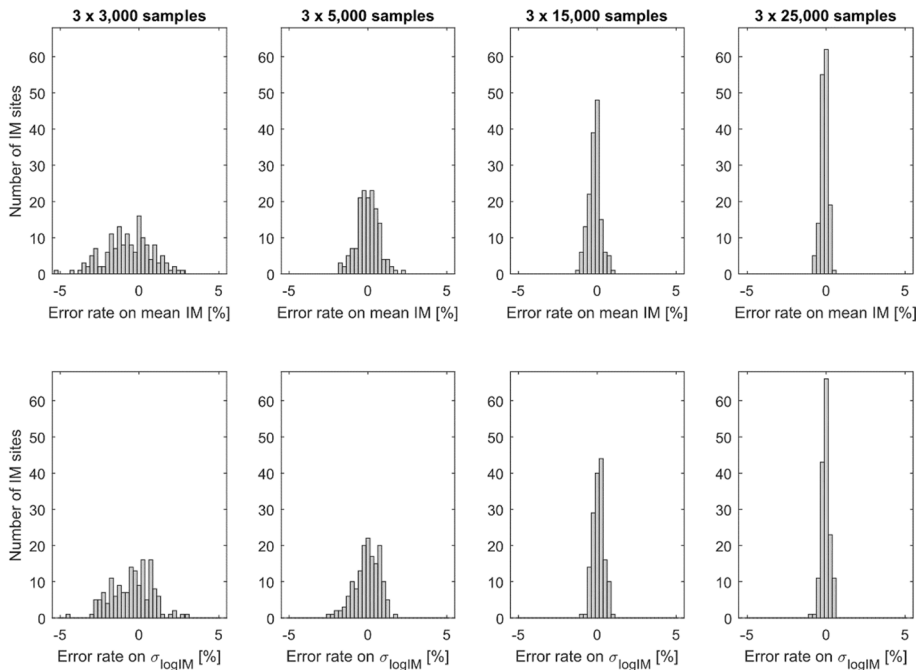


Fig. 11 Histograms of the error rate of the mean and standard deviation of the posterior IM distribution at the various sites, with respect to the analytical solution, for various lengths of the MCMC chains

$$\begin{cases} \epsilon_{E(IM)} = \frac{E(IM_{sampled}) - E(IM_{exact})}{E(IM_{exact})} \times 100 \\ \epsilon_{\sigma_{logIM}} = \frac{\sigma_{logIM_{sampled}} - \sigma_{logIM_{exact}}}{\sigma_{logIM_{exact}}} \times 100 \end{cases} \quad (21)$$

where $E(IM)$ is the expectation of the posterior IM distribution and σ_{logIM} is the standard deviation of the logarithm of IM.

The estimated posterior distributions of IM are globally in very good agreement with the analytical solution, both in terms of mean and standard deviation, thus providing further confidence with respect to the application of the proposed sampling-based inference algorithm. This comparison exercise serves as a preliminary step to calibrate the size of samples and ensure proper convergence of the posterior distributions. In the current application, Fig. 11 shows that satisfying accuracy levels are quickly reached as the size of samples grows. Therefore, it is decided to opt for a chain size of 15 000 samples, in order to keep the error rate well below 1%. With this choice, the complete BN (including also *C*, *DS*, *MLS* and *S* nodes) is solved in less than half an hour with a medium-performance personal computer (8 GB memory): this timeframe is in line with what should be expected from elaborate rapid response systems, although it could be further improved with the use of dedicated computing servers.

5.2 Updated loss estimates

Following the discussion of Sect. 2.4 on the assumptions to be made for the covariance models, three different correlation hypotheses are tested for the response \mathbf{C} of bridges:

- *Corr1*: no correlation is introduced, so that $\Sigma_{\mathbf{C}}$ is simply a diagonal matrix.
- *Corr2*: only the correlation of the record-to-record variability is introduced, with a correlation model decreasing with inter-bridge distance (i.e., Eq. 6 with $\rho_{ij}^M = 0$).
- *Corr3*: in addition to the correlation of the record-to-record variability, a full correlation between bridges of the same type (i.e., using the same fragility model) is assumed, i.e. $\rho_{ij}^M = 1$ if bridges i and j are in the same typology.

Global results for the system connectivity are detailed in Table 2, where it is shown that the proposed approach is also able to identify which MLS is the most likely to remain accessible (i.e., “best” MLS).

A significant difference is observed between the prior and posterior probabilities of disconnection, due to the assumption that the damage states of 5 bridges were entered as evidence (see Fig. 10). The evidence of a bridge’s state has an impact on the system at various levels:

- The observation of a bridge failure directly modifies the accessibility of the MLS(s) to which it belongs, and in turn system connectivity.
- If the seismic response \mathbf{C} is modelled with a constrained correlation model (i.e., *Corr2* or *Corr3*), then the observation of a bridge’s state may modify the seismic response of other bridges, in turn modifying their probability of failure and ultimately the accessibility of the MLSs to which they belong.
- Finally, as seen in Fig. 7, the observation of a bridge’s state may also modify the distribution of the IM at the base of the bridge, in turn modifying the ground-shaking field in the vicinity. In the BN model, the state of a bridge i is dependent on both the seismic response C_i and on the shaking level IM_i ; therefore, both parent nodes \mathbf{C} and \mathbf{IM} are affected by the evidence on the bridge’s state.

The differences between the three correlation models are noticeable: the extreme configurations *Corr1* and *Corr3* should be used as upper and lower bounds of the model outcome, pending an in-depth investigation of appropriate correlation models for the seismic response. Furthermore, each assumption leads to the identification of a different MLS as the least affected route, which could have a large impact on emergency operations. An example of the changes in the rapid estimate of the post-earthquake condition of the network is provided in Fig. 12, both in terms of failure probability of bridges and

Table 2 Results of the Bayesian updating, in terms of probability of disconnection between points A and B (i.e., $S = 0$) and of identification of the “best” MLS, for the various correlation assumptions

	Prior	Posterior		
		<i>Corr1</i>	<i>Corr2</i>	<i>Corr3</i>
Pr(Disconnection)	0.095	0.342	0.353	0.435
“Best” MLS	#45	#49	#21	#3

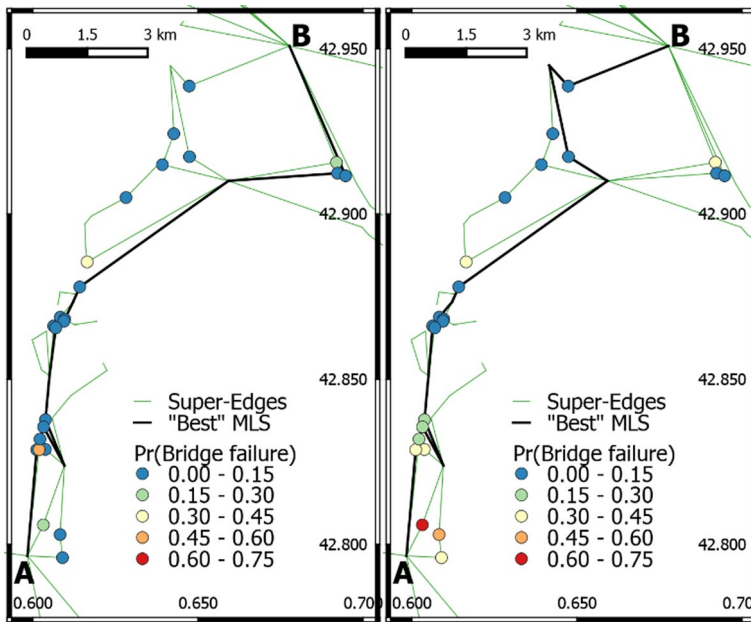


Fig. 12 Left: prior distribution using only the characteristics of the earthquake event; Right: posterior distribution using field observations (with *Corr3* model)

identifying the most accessible MLS. Similar updating may also be performed for the damage assessment of built areas (see Fig. 13): the results are shown for *Corr3* model, which only affects the seismic response **C** of bridges. However, in a BN, the model adopted for the nodes related to a type of component can impact the posterior distributions of all nodes, including those related to other types of components in the system.

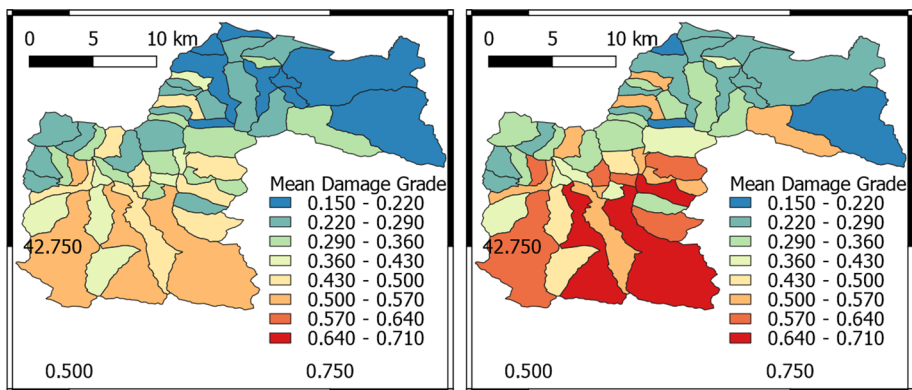


Fig. 13 Left: prior mean of μ_D (Mean Damage Grade) of the built areas; Right: posterior distribution using field observations (with *Corr3* model)

Therefore, the choice of a correlation model **C** has consequences on the distribution of **IM**, which in turn affects the distribution of the mean damage grade for buildings.

Apart from the changes in the mean damage and loss values, the Bayesian updating is especially useful for reducing the related uncertainties; and therefore, gaining more confidence in rapid damage assessment. In Fig. 14, the updated standard-deviations of $\log\text{-PGA}$, used as **IM**, are reduced by half (from around 0.42 in the prior distribution down to the 0.20–0.25 range) in the immediate vicinity of seismic stations. On the other hand, the survival/failure observations on bridges have a reduced impact on the distribution of **IM**, since the probability updating enters here in competition with the updating of **C**, which have an equal contribution in the damage sampling (see Eq. 1).

The updating process of the response **C** is further detailed in Fig. 15, where the posterior distribution of a given C_i is used to assemble a fragility function, through the following expression:

$$P_f(PGA) = \Phi\left(\frac{im - \bar{C}_i}{\sigma_{C_i}}\right) \quad (22)$$

Two opposite configurations are shown in Fig. 15:

- In the top plots, bridge #533 is observed as intact, and its updated response distribution is therefore moving towards higher PGA levels. The response of bridge #546, a neighbouring bridge of the same type (i.e., Type 1), is also updated, but to a lesser extent. As expected,

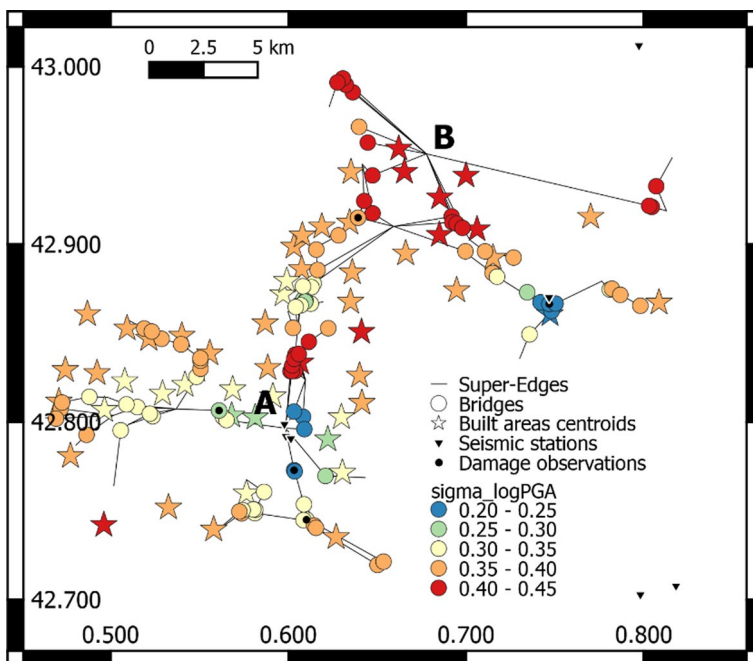


Fig. 14 Posterior standard-deviation $\sigma_{\log\text{PGA}}$ of the IMs at the locations of interest (with *Corr3* model)

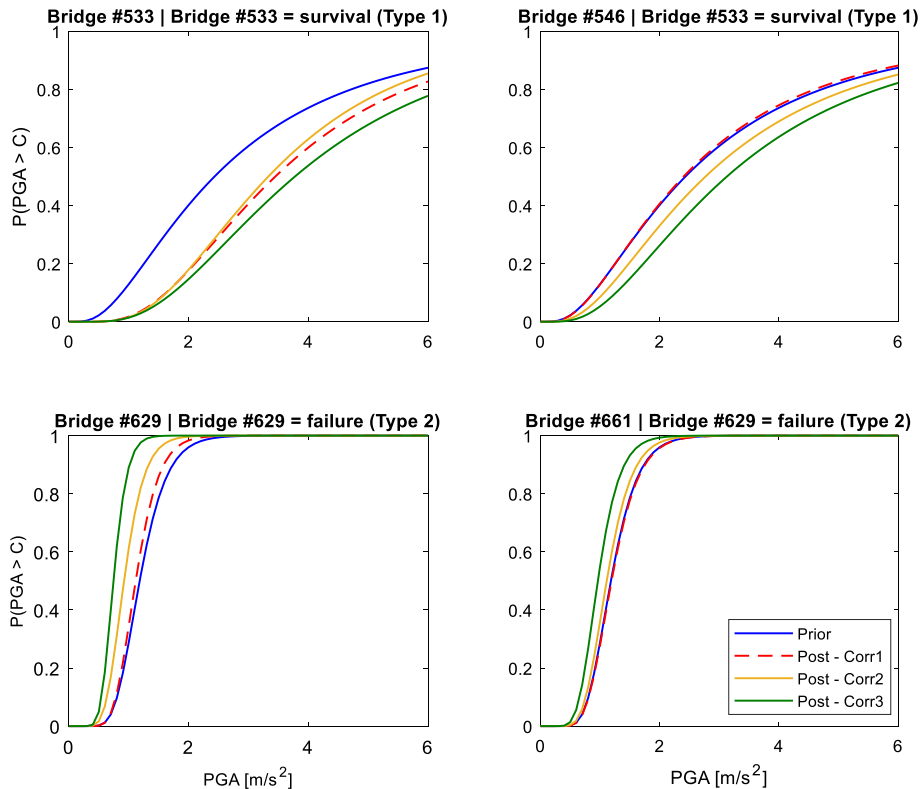


Fig. 15 Updating of the seismic response of bridges, based on damage observations

the correlation model *Corr1* (no correlation at all) has no impact on the response. Since the two bridges are only about 2 km apart, the application of model *Corr2* (only distance-dependent correlation) has a noticeable influence on the response distribution. Model *Corr3* has the most influence since it corresponds to the strongest correlation structure.

- In the bottom plots, bridge #629 is observed as failed, and its updated response distribution is therefore moving towards lower PGA levels. Similar trends are observed for bridge #661 of the same type (i.e., Type 2): in this case, the effect of model *Corr2* is less noticeable, since the two bridges are further apart (10 km).

The updated seismic response may then be used as a prior distribution in subsequent risk analyses: this feature is especially useful in the earthquake aftershock phase, since the updated fragility models may be directly applied to the components that have survived, in order to refine the aftershock risk assessment.

6 Conclusions

This study has investigated the implementation of a Bayesian Network-based framework for improving situational awareness during the rapid response phase following an earthquake event. It has been demonstrated that a BN built in OpenBUGS environment is able to provide updated losses for a real-world road network and built areas based on available observations. The BN is solved with a MCMC sampling scheme, which delivers approximate posterior distributions: this approximate solution, as opposed to exact inference algorithms, is a necessary trade-off in order to treat systems of hundreds of components exposed to spatially distributed hazards. Moreover, the sampling inference scheme used in OpenBUGS can combine continuous (e.g., intensity measures, seismic capacities) and discrete variables (e.g., damage states), which has the benefit of introducing exact distributions instead of discretizing continuous variables. Regarding road networks, when connectivity loss is used as a simple system indicator, the decomposition of the system into MLSs is also essential in reducing the complexity of the BN. As a result, in the studied example, posterior distributions are generated within 20 or 30 min. Moreover, the decomposition into MLSs has the benefit of identifying specific routes associated with probabilities of accessibility: this information has the potential to be used by emergency managers to set up dedicated evacuation routes or safe itineraries to hospitals. Each MLS can also be associated with a travel distance or travel duration, in order to develop more elaborate performance indicators than connectivity loss.

The input of field observations has the expected effect of refining the loss estimates, thus providing more confidence in the information delivered to emergency responders. The main mechanism behind the updating process is due to the formulation of the covariance matrices Σ_{IM} and Σ_C : while the correlation between intensity measures is well studied, the correlation between the seismic capacities of components is less obvious. Various correlation assumptions have been tested here, in an attempt to quantify this effect: further investigations will be required in order to propose a robust model for the covariance of the seismic response. It is shown that correlated seismic capacities also lead to updated fragility models, depending on the observed damage on components. Such a feature of the BN model would be especially useful to update the risk assessment models in the aftershock phase, for instance. Finally, the results of such a BN application can constitute the starting point of rapid repair strategies for bridges, in order to improve the seismic resilience of transportation systems.

While the present study has focused on methodological developments and on the demonstration of a simple yet realistic case-study, further work should be devoted to the verification of this framework on a real earthquake case, where collected post-earthquake data could be used to replay the sequence of events and evaluate the ability of the BN to reduce uncertainties on loss estimates. Although the theory behind the BN model has been checked (e.g., validation of the ground-motion updating part in Sect. 5.1), an application to a real earthquake case would have the merit of confronting the method to actual conditions (e.g., possibility of missing data, contradicting observations, etc.). Moreover, the BN model is solely based on a connectivity analysis: if each MLS is associated with a travel distance or duration, post-earthquake travel times may be generated. However, such a performance indicator would remain based on free-flow travel conditions (i.e., each edge of the road network is associated with a constant travel speed, whatever the traffic conditions). While the BN model in itself is not able to account for traffic flows, the approach introduced by Gehl et al. (2018) may be applied: it consists of generating off-line stochastic loss

scenarios, computing travel times with a traffic flow model (i.e., with an origin-destination matrix and user-equilibrium equations) and learning an *ad-hoc* BN structure from the simulation outcomes.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10518-022-01349-4>.

Author contributions *Conceptualization and Methodology*: Pierre Gehl, Rosemary Fayjaloun; *Formal analysis and investigation*: Rosemary Fayjaloun, Pierre Gehl, Li Sun, Caterina Negulescu, Enrico Tubaldi, Ekin Özer; *Writing - original draft preparation*: Pierre Gehl; *Writing - review and editing*: All; *Funding acquisition & Supervision*: Pierre Gehl, Enrico Tubaldi, Dina D'Ayala.

Funding This research was funded by the European Union's Horizon 2020 research and innovation program, Grant Number 821046, project TURNkey.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Applegate CJ, Tien I (2019) Framework for probabilistic vulnerability analysis of interdependent infrastructure systems. *J Comput Civil Eng* 33(1):04018058
- Argyroudis S, Selva J, Gehl P, Pitilakis K (2015) Systemic seismic risk assessment of road networks considering interactions with the built environment. *Computer-Aided Civ Infrastruct Eng* 30(7):524–540
- Auclair S, Monfort D, Colas B, Langer T, Bertil D (2014) Outils de réponse rapide pour la gestion opérationnelle de crises sismiques. In *Colloque SAGEO*
- Bensi M, Der Kiureghian A, Straub D (2011) A bayesian network methodology for infrastructure seismic risk assessment and decision support. PEER Report 2011/02. University of California, Berkeley
- Bensi M, Kiureghian D, Straub D (2013) Efficient Bayesian network modeling of systems. *Reliab Eng Syst Saf* 112:200–213
- Bensi M, Kiureghian A, Straub D (2015) Framework for post-earthquake risk assessment and decision making for infrastructure systems. *ASCE-ASME J Risk Uncertain Eng Syst Part A: Civil Eng* 1(1):04014003
- Byun JE, Song J (2021) Generalized matrix-based Bayesian network for multi-state systems. *Reliab Eng Syst Safety*, 107468
- Byun JE, Zwirgmaier K, Straub D, Song J (2019) Matrix-based Bayesian Network for efficient memory storage and flexible inference. *Reliab Eng Syst Saf* 185:533–545
- Cavaleri F, Franchin P, Gehl P, D'Ayala D (2017) Bayesian networks and infrastructure systems: Computational and methodological challenges. In: Gardoni P (ed) *Risk and reliability analysis: Theory and applications*. Springer, Berlin
- Cremen G, Galasso C (2020) Earthquake early warning: Recent advances and perspectives. *Earth-Sci Rev* 205:103184
- Crowley H, Colombi M, Silva V, Monteiro R, Ozcebe S, Fardis M, Tsionis G, Askouni P (2011) Fragility functions for roadway bridges. In *SYNER-G Deliverable Report D3.6*

- Crowley H, Pinho R, van Elk J, Uilenreef J (2019) Probabilistic damage assessment of buildings due to induced seismicity. *Bull Earthq Eng* 17(8):4495–4516
- Dunnett CW, Sobel M (1955) Approximations to the probability integral and certain percentage points of a multivariate analogue of Student's t-distribution. *Biometrika* 42(1/2):258–260
- Fayjaloun R, Gehl P, Auclair S, Boulahya F, Guérin-Marthe S, Roullé A (2021) Integrating strong-motion recordings and Twitter data for a rapid shakemap of macroseismic intensity. *Int J Disaster Risk Reduct* 52:101927
- Franchin P (2014) A computational framework for systemic seismic risk analysis of civil infrastructural systems. In: Pitilakis K, Franchin P, Khazai B, Wenzel H (eds) *SYNER-G: Systemic seismic vulnerability and risk assessment of complex urban, utility, lifeline systems and critical facilities*. Springer, Dordrecht, pp 23–56
- Gehl P, Douglas J, d'Ayala D (2017) Inferring earthquake ground-motion fields with bayesian networks. *Bull Seismologic Soc Am* 107(6):2792–2808
- Gehl P, Cavalieri F, Franchin P (2018) Approximate Bayesian network formulation for the rapid loss assessment of real-world infrastructure systems. *Reliab Eng Syst Saf* 177:80–93
- Gehl P, Matsushima S, Masuda S (2021) Investigation of damage to the water network of Uki City from the 2016 Kumamoto earthquake: derivation of damage functions and construction of infrastructure loss scenarios. *Bull Earthq Eng* 19:685–711
- Grünthal G (1998) European Macroseismic Scale. *Cahiers du Centre Européen de Géodynamique et de Séismologie*, 15, Luxembourg
- Guérin-Marthe S, Gehl P, Fayjaloun R, Negulescu C, Auclair S (2021) Rapid earthquake response: the state-of-the art and recommendations with a focus on European systems. *Int J Disaster Risk Reduct* 52:101958
- Huang C, Darwiche A (1996) Inference in belief networks: A procedural guide. *Int J Approximate Reasoning* 15(3):225–263
- Ioannou I, Chandler RE, Rossetto T (2020) Empirical fragility curves: the effect of uncertainty in ground motion intensity. *Soil Dyn Earthq Eng* 129:105908
- Jayaram N, Baker JW (2009) Correlation model for spatially distributed ground-motion intensities. *Earthq Eng Struct Dynamics* 38(15):1687–1708
- Lagomarsino S, Giovinazzi S (2006) Macroseismic and mechanical models for the vulnerability and damage assessment of current buildings. *Bull Earthq Eng* 4(4):415–443
- Lin K, Wald DJ, Kircher CA, Slosky D, Jaiswal K, Luco N (2018) USGS ShakeCast system advancements. In: *Proceedings of the 11th national conference on earthquake engineering 2018, NCEE 2018: integrating science, engineering, and policy*, 6, pp 3458–3468
- Lunn D, Spiegelhalter D, Thomas A, Best N (2009) The BUGS project: Evolution, critique and future directions (with discussion). *Stat Med* 28:3049–3082
- Murphy K (2001) The Bayes net toolbox for Matlab. *Comput Sci Stat* 33:1024–1034
- Pozzi M, Der Kiureghian A (2013) Gaussian Bayesian network for reliability analysis of a system of bridges. In: *Proceedings of the 11th international conf. on structural safety and reliability*, New York, United States
- Roullé A, Macau A, Figueras S, Monfort D, Lantada N, Susagna T, Irizarry J (2012) Performing seismic scenario in the Luchon - Val d'Aran area, Central Pyrenees. In: *Proceedings of the EUROGeo 2012, Bologna, Italy*
- Sedan O, Negulescu C, Terrier M, Roullé A, Winter T, Bertil D (2013) Armageddon—a tool for seismic risk assessment illustrated with applications. *J Earthquake Eng* 17(2):253–281
- Silva V (2019) Uncertainty and correlation in seismic vulnerability functions of building classes. *Earthq Spectra* 35(4):1515–1539
- Stafford PJ (2012) Evaluation of structural performance in the immediate aftermath of an earthquake: A case study of the 2011 Christchurch earthquake. *Int J Forensic Eng* 1(1):58–77
- Tapia M (2006) Desarrollo Y Aplicación de Métodos Avanzados Para La Caracterización de La Respuesta Sísmica Del Suelo a Escala Regional Y Local. PhD thesis, Universidad Politécnica de Catalunya, Barcelona, Spain
- Tien I, Der Kiureghian A (2016) Algorithms for Bayesian network modeling and reliability assessment of infrastructure systems. *Reliab Eng Syst Saf* 156:134–147
- Tien I, Der Kiureghian A (2017) Reliability assessment of critical infrastructure using Bayesian networks. *J Infrastruct Syst* 23(4):04017025
- Toma-Danila D (2018) A GIS framework for evaluating the implications of urban road network failure due to earthquakes: Bucharest (Romania) case study. *Nat Hazards* 93(1):97–111
- Trendafiloski G, Wyss M, Rosset P (2011) Loss estimation module in the second generation software QLARM. *Human Casualties in Earthquakes*. Springer, Dordrecht, pp 95–106

- Tubaldi E, Özer E, Douglas J, Gehl P (2021) Examining the contribution of near real-time data for rapid seismic loss assessment of structures. *Structural Health Monitoring*, In press
- Vanmarcke E (1983) *Random Fields, Analysis and Synthesis*. The MIT Press, Cambridge, Massachusetts
- Wald D, Jaiswal K, Marano KD, Bausch DB, Hearne MG (2010) *PAGER — Rapid Assessment of an Earthquake's Impact*. USGS Fact Sheet 2010-3036
- Worden CB, Thompson EM, Baker JW, Bradley BA, Luco N, Wald DJ (2018) Spatial and spectral interpolation of ground-motion intensity measure observations. *Bull Seismol Soc Am* 108(2):866–875
- Yeh CH, Loh CH, Tsai KC (2006) Overview of Taiwan earthquake loss estimation system. *Nat Hazards* 37(1–2):23–37
- Zülfikar c, Fercan NOZ, Tunc S, Erdik M (2017) Real-time earthquake shake, damage, and loss mapping for Istanbul metropolitan area. *Earth Planets Space* 69:1–5

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.