A Bayesian Network Model to Assess Seismic Risk of Reinforced Concrete Girder Bridges

Paolo Franchin
Associate Professor, Dept. of Structural & Geotechnical Engineering, Sapienza University of Rome, Rome, Italy

Alessio Lupoi
Assistant Professor, Dept. of Structural & Geotechnical Engineering, Sapienza University of Rome, Rome, Italy

Fabrizio Noto
Professional Engineer, METIS srl, Rome, Italy

Solomon Tesfamariam
Associate Professor, Dept. of Civil Engineering, University of British Columbia, Kelowna, Canada

ABSTRACT: Infrastructure owners or governmental agencies need tools for rapid screening of assets in order to prioritize resources allocation for detailed risk assessment. This paper provides one such tool based on Bayesian Networks and aimed at replacing so-called generic/typological seismic fragility functions for reinforced concrete girder bridges. Resources for detailed assessments should be allocated to bridges with highest consequence of damage, for which site hazard, bridge fragility and traffic data are needed. The presented Bayesian Network predicts the seismic fragility of a bridge at a given site based on data that can be obtained by visual inspection at low cost. Results show that the predicted fragilities are of sufficient accuracy for establishing relative ranking based on risk and assign priorities. While the actual data employed to train the network (establishing conditional probability tables) refer to the Italian bridge stock, the network structure and engineering judgment behind it can be easily transferred to other situations.

The Italian national highway administration (ANAS) owns over 5,000 bridges and less than 500 have been subject to detailed seismic assessment. Rapid and reliable screening tools for highway bridges are of paramount importance for prioritizing detailed assessment and retrofit budget allocation. A risk-based approach, with consideration of site-specific hazard and typological fragility functions (to quantify likelihood of failure), weighted with surveyed traffic data and network importance of the bridge (to quantify consequence of failure), can provide a first estimate of prioritization. The reliability of typological fragility functions, however, is questionable (Borzi et al., 2015). The purpose of this paper is thus to present a Bayesian belief network (BBN) based model developed to predict fragility of reinforced concrete (RC) girder bridges with consideration of limited set of parameters (e.g. profile of the bridge, pier height, hazard) that can be surveyed with minimum effort, by visual inspection or from prevalent design practices.

The BBN captures believed relations between a set of input variables and output variable (attainment or exceedance of a limit state). The model has been developed employing engineering judgment to select the relevant random variables, and then iteratively draw and refine the BBN structure relying numerical simulation of response for a large bridge database (DB). The DB is used to support the training algorithm in

determining the conditional probability tables (CPTs).

1. TRAINING DATA SET

Numerical response data have been produced for 485 bridges in the SOAWE dataset, owned by ANAS. Discarding single-span bridges, data from 390 bridges have been utilized to train the BBN. A description of the properties of these bridges can be found in Borzi et al. (2015): the most frequent characteristics are less than five spans, single-stem piers between 10 m and 20 m high with box-type section, piled foundation on B soil category (Eurocode 8 classification), and simply supported decks on thin non-seismically designed elastomeric pads.

Finite element (FE) response history analyses have been carried out on the 390 bridges in a consistent manner using purpose-made software BR.I.T.NE.Y. The software queries the relational DB for site, geometry and mechanical properties of each bridge, and sets up a 3D inelastic FE model and analysis in OpenSees (McKenna et al. 2010). Detailed description of the modeling choices and the software design is provided in Borzi et al. (2015).

Inelastic response-history analysis is employed to evaluate the seismic fragility curve of each bridge. It is stressed how fragility is not meaningful per se, being a non-transferrable bridge and site property. It is used as a convenient intermediate result to obtain the mean annual frequency (MAF) of exceedance of the limit state (LS) of interest at a lower cost. Herein it is formulated as:

$$p_{LS}(a) = p\left(Y_{LS} \geq 1 \middle| A = a\right) \tag{1}$$

where A denotes the peak ground acceleration (PGA), and Y_{LS} is a global indicator of the attainment or exceedance of the considered LS. The latter is formulated in such a way that the unit value marks the LS, and larger values correspond to non-satisfactory states (Jalayer et al. 2007). Girder bridges, due to their topology, can be described as serial systems (failure in any span implies system failure). Thus the following simple expression is used for Y_{LS}:

$$Y_{LS} = \max_{\text{piers}}\left[\max_{\text{failure modes}}\left(D/C\right)\right] \tag{2}$$

where the maximum of the (peak over time) demand (D) to capacity (C) ratio is taken over all piers and considered failure modes.

The same three failure modes have been considered for girder bridges as in Borzi et al. (2015): the ductile flexural mode, checked in terms of chord rotation in the piers, and the brittle shear and unseating modes, checked in terms of pier shear force and bearings' displacements, respectively. Abutments and foundation failure modes have not been considered, based on the traditionally conservative design of bridge foundation in Italy (Calvi et al., 2013).

From a functional point of view, two LSs suffice: light (LD) and severe damage (SD). When damage is light, the bridge can be open, at least to emergency traffic. It is thus available in network connectivity analysis (emergency operations), but excluded from traffic flow analysis (indirect loss). When damage is severe the risk of collapse in the event of an aftershock is high (Franchin and Pinto, 2009) and the bridge must be closed. In this respect SD and higher damage LSs, up to collapse, are equivalent in terms of traffic restriction decision-making. In this work, we focus on LD, due to its observed prevalence in past Italian earthquakes (Calvi et al., 2013). In terms of capacities, LD occurs when rotation θ exceeds the yield value θ_y, the displacement is such that the deck loses support from the bearing seat (but remains on the pier cap), $u>u_{seat}$, or shear strength is exceeded, $V>V_R$.

Fragility curves are constructed by multiple-stripe analysis (MSA) (Jalayer and Cornell, 2009), i.e. parametrically increasing a chosen intensity measure (IM, the PGA in this study) and estimating the probability of exceedance at each level, based on the sample of responses obtained from motions selected to match the current IM level. This procedure represents the most rigorous application of total probability theorem

for the MAF evaluation, since it allows for changes in the motions used at each IM level.

Recent availability of efficient algorithms (Bradley 2013; Lin et al. 2013) has greatly facilitated the selection of appropriate ground motions at each IM level. The generalized conditional intensity measure method (Bradley, 2013) has been used to select motions at each site and IM level. The method postulates a joint lognormal distribution $f(\mathbf{IM})$ of a vector of IMs, based on their proven marginal lognormality and availability of cross-IM correlation models. It then derives, at each intensity level IM*=im*, the conditional distributions $f(\mathbf{IM}_{n-1}|IM^*=im^*)$ of the subvector from which IM* employed for the hazard curve is removed (A in the case at hand). Motions are finally selected in the required number as those of minimum distance with those sampled from each $f(\mathbf{IM}_{n-1}|IM^*=im^*)$.

The chosen vector of IMs includes, beside PGA, the spectral accelerations at ten periods: 0.10s, 0.15s, 0.20s, 0.30s, 0.40s, 0.50s, 0.75s, 1.00s, 1.50s and 2.00s. Marginal distributions (including PGA hazard) are obtained by the ground motion prediction equation by Boore and Atkinson (2008), and the Italian seismic source model ZS9 (INGV, 2004). Correlation between spectral ordinates is described with the model by Baker and Jayaram (2008). Sampling of motions has been carried out in an automatic manner, but caution has been put in not sampling different motions for bridges at the same site (such as e.g. parallel viaducts on the same highway). In such cases, only one set of motions has been sampled and used for the analysis of both bridges.

In order to strike a balance between accuracy and computational cost, nine intensity levels and ten motions per level (totaling 90 inelastic response history analyses for each of the 390 bridges, i.e. almost 40,000 runs) were kept as in (Borzi et al., 2015). Even with n = 10 natural motions, the empirical distribution falls within the Kolmogorov-Smirnov bounds of the target conditional ones, $f(\mathbf{IM}_{n-1}|IM^*=im^*)$, for all IMs and intensity levels, and resulted in generally stable results in most cases.

2. BN MODEL DEVELOPMENT

2.1. BNs and general modeling issues

BBN is a graphical model that describes a probabilistic relationship among a set of variables (Pearl 1988). A BBN is represented with a directed acyclic graph (DAG), where the nodes represent variables of interest (e.g., concrete strength, pier height, etc.), and the links between them indicate informational or causal dependencies among the variables. The BBN B over a set of variables $X = \{x_1,\ldots, x_k\}$, $k \geq 1$ is a network structure B_s, which denotes a DAG over X and a set of CPTs B_p (Figure 1). The relations between the variables in a BBN, as depicted in Figure 1, show variables $x_{1,2}$ are said to be the parent of x_3 (x_3 is the child of $x_{1,2}$ if the link goes from $x_{1,2}$ to x_3). The CPT, $B_p = \{p(x_i|pa(x_i)|x_i \in X)\}$, gives the causal relation between parents pa(x_i) and child x_i. The joint distribution of the random variables in the BBN is given as the product of the conditional probability distributions (Bensi et al. 2014):

$$p(x_1,\ldots,x_n) = \prod_{i=1}^{n} p\left(x_i \middle| pa(x_i)\right) \qquad (3)$$

With increasing number of parents and corresponding states, two problems arise. The first has to do with Bayesian inference, since the child CPT size increases exponentially with the associated computational problems (Bensi et al. 2014). The second has to do with the development of the CPTs. Having data to support establishment of probabilities for all parents' states combinations becomes increasingly unlikely. One modeling technique utilized in this paper is that of inverting arrows direction (Norsys Software Corp., 2006), which preserves the statistical dependence (e.g. Figure 2).

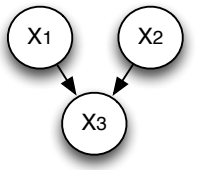
2.2. Expert-judgment leading to BN structure

Tesfamariam and Liu (2013) have shown that a BBN structure generated through expert knowledge can furnish comparable results to those generated through learning algorithms, with the advantage, however, that the causal rela-

tion between different parameters of engineering significance can be maintained. Thus, in this paper an expert-driven BBN structure is developed.

X_1	PROB.
L	$P(X_1=L)$
M	$P(X_1=M)$
H	$P(X_1=H)$

X_2	PROB.
L	$P(X_2=L)$
H	$P(X_2=H)$

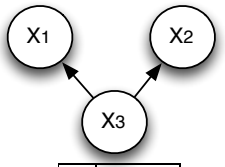


X_1	X_2	PROBABILITY	
		L	H
L	L	$P(X_3=L\|X_1=X_2=L)$	$P(X_3=H\|X_1=X_2=L)$
L	H	$P(X_3=L\|X_1=L,X_2=H)$	$P(X_3=H\|X_1=L,X_2=H)$
M	L	$P(X_3=L\|X_1=M,X_2=L)$	$P(X_3=H\|X_1=M,X_2=L)$
M	H	$P(X_3=L\|X_1=M,X_2=H)$	$P(X_3=H\|X_1=M,X_2=H)$
H	L	$P(X_3=L\|X_1=H,X_2=L)$	$P(X_3=H\|X_1=H,X_2=L)$
H	H	$P(X_3=L\|X_1=X_2=H)$	$P(X_3=H\|X_1=X_2=H)$

Figure 1: Basic BBN with probability tables.

X_3	Probability		
	L	M	H
L	$p(X_1=L\|X_3=L)$	$p(X_1=M\|X_3=L)$	$p(X_1=H\|X_3=L)$
H	$p(X_1=L\|X_3=H)$	$p(X_1=M\|X_3=H)$	$p(X_1=H\|X_3=H)$

X_3	Probability		
	L	M	H
L	$p(X_2=L\|X_3=L)$	$p(X_2=M\|X_3=L)$	$p(X_2=H\|X_3=L)$
H	$p(X_2=L\|X_3=H)$	$p(X_2=M\|X_3=H)$	$p(X_2=H\|X_3=H)$



X_3	Prob.
L	$p(X_3=L)$
H	$p(X_3=H)$

Figure 2: The BBN in Figure 1, with reversed arcs, and updated probability tables.

Preliminary analysis of the numerical response results (from BR.I.T.N.E.Y.) has shown that the three failure modes do not contribute in the same proportions to the exceedance of the LD LS (this is not the case for the SD LS, not dealt with here). Figure 3 shows the frequency histograms of failure modes over the numerical runs (each run being one bridge, at one intensity level A=a, with one of the chosen motions) where Y_{LD}>1, i.e. failure occurred. These results show that LD is dominated by the flexural mode. Thus the light damage indicator, Y_{LD}, is quantified only in terms of flexural failure mode.

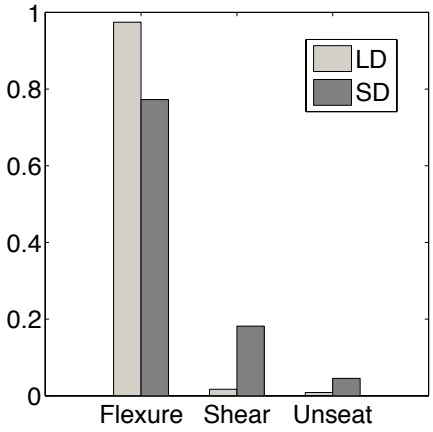


Figure 3: Contribution of each failure mode to the LD and SD limit states.

The variables to be selected belong to the general categories of material properties, geometry including details, and site characteristics. The requirement to be satisfied is, to the extent possible, limit the selected variables to properties that can be easily obtained by visual inspection or from prevalent design practices. This requirement is dictated by the intended use of the models to predict bridge fragility without a detailed seismic assessment, i.e. in the absence of the associated tests and inspections data. Thus, data on materials and reinforcement are kept to a minimum (concrete strength f_c, steel yield strength f_y and longitudinal steel area A_s).

As far as geometry is concerned, girder bridges are made of one or more decks, supported over a variable number of piers. A first problem to be solved is the characterization of bridges with different number of piers with the same number of variables. Initial assessment of the flexural failure results indicated that in over 70% of cases system damage, as expressed by Eq.(2), is associated with either the SP, TP or the SP=TP (Figure 4). Consequently, beside global geometrical variables, only variables pertaining the SP and TP have been considered in the BBN model.

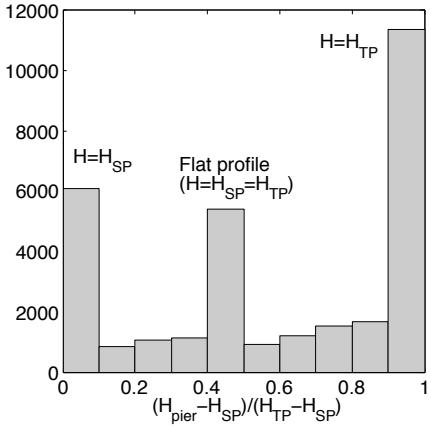


Figure 4: Relative frequency of the height of the pier leading to failure, as a proportion of SP and TP heights: zero means the pier leading to failure is the shortest one, one means it is the tallest.

Based on the previous considerations, the BBN shown in Figure 5 is formulated. Variables related to the TP and SP are shown on the left and right, respectively, while the central part of the BBN contains global variables that take on a single value for the entire bridge (material properties, which are spatially variable, have been modeled as constant in the numerical analyses in BRI.T.N.E.Y. and thus kept constant also in the BBN model).

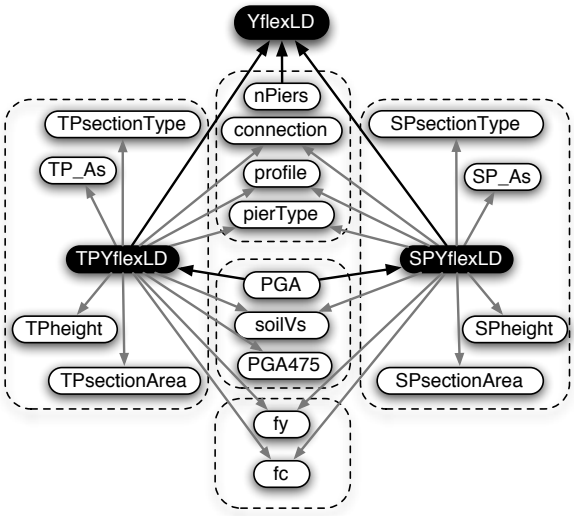


Figure 5: BBN structure for the LD limit state assessment (input and output variables in white and black shading, respectively, grey arrows indicate reversal of causal relation, see Figure 2).

The (flexural) light damage indicator *YflexLD* is related to *TPYflexD* and *SPYflexD* only. The observable engineering properties that affect piers flexural damage are: *pierType* and height (*TPheight* and *SPheight*), section area (*TPsectionArea* and *SPsectionArea*) and type (*TPsectionType* and *SPsectionType*), area of longitudinal steel (*TP_As* and *SP_As*) and material properties (concrete *fc* and steel *fy*). Pier height and section area are good proxies that correlate well with pier stiffness and yield deformation capacity, especially when complemented with pier and section types. The advantage over use of aspect ratio, which is also related to stiffness and θ_y, is that aspect ratio must be defined in both plan directions, unless the pier is circular or square.

Among global variables, *profile* and *connection* are the geometrical parameters that contribute most to increase/decrease the demand/capacity ratio in the piers. Further, given the serial system assumption, at parity of other factors the vulnerability of the bridge will increase with the number of piers. As a result, the number of piers (*nPiers*) is directly linked to the global damage indicator, *YflexLD*. Global variables include also site-dependent properties: the average 30-meters shear wave velocity, *soilVs* and *PGA475*, the PGA with mean return period of 475 years. The latter, which has almost unit correlation with the area under the hazard curve, is meant to represent the effect of the site seismicity (whatever is not captured by IM*=*A*) on the fragility.

Finally, all variables are either discrete in nature or discretized. The ranges are determined by sensitivity analysis and engineering judgment in order to maximize the effect on dependent variables (and of course *TPYflexLD*, *SPYflexLD* and *YflexLD* have the last interval corresponding to values larger than one). In the software package used for this study (NETICA by Norsys Software Corp., 2006), three data training algorithms (Cooper and Herskovits, 1992) are provided: counting, expectation-maximization and gradient descent. The simple counting algorithm is used

herein, since results are available for all nodes/variables. During the counting process, where data is not available for a specific parents' combination, equal probabilities are usually assigned to the corresponding child states in the CPT. In this work, however, the possibility of missing data is minimized by changing the direction of the arrows (grey arrows in Figure 5).

3. MODEL PERFORMANCE

The model performance is assessed by comparing results, in terms of both fragility curves and associated risk values, obtained from numerical analysis with those produced by the model. Risk is computed with the closed-form proposed by Vamvatsikos (2013):

$$\lambda_{LS} = \int_0^\infty p_{LS}(a)\left|d\lambda_A(a)\right| = \\ = \sqrt{p}k_0^{1-p}\left[\lambda_A\left(\exp\left(\mu_{\ln A}\right)\right)\right]^p \exp\left(\frac{1}{2}pk_1^2\sigma_{\ln A}^2\right) \quad (4)$$

based on a quadratic fit of the PGA hazard in log-log space, $\lambda_A = k_0(-k_1 \ln a - k_2 \ln^2 a)$, and on a lognormal fit (by maximum likelihood estimation) to the fragility curve parameters μ_{lnA} and σ_{lnA}. The latter is available at nine discrete points (up to 0.3g), in the form of Y_{LD} values for the numerical analyses, and probability of $Y_{LD}>1$ for the BBN model (the value for state $\{0.99,\infty\}$ from the CPT of *YflexLD*, updated after observed values have been assigned as evidence to all other observable variables with the exception of PGA, which is varied parametrically). Parameter $p=(1+2k_2\sigma^2_{\ln A})^{-1}$ in Eq.(4) is a measure of hazard curvature and when $p=1$ Eq.(4) collapses on the well-known risk closed-form based on linear hazard. Figure 6 shows the results for one sample bridge (for which the model performs quite well).

Actually, the agreement between the BBN-based and the numerical fragilities is remarkable for most bridges. When this is not the case, differences can have both signs (underestimation or overestimation). BBN-based probabilities are monotonically increasing with PGA in a smooth manner, whereas the numerical results are showing larger variability, due to the number of ground motions per IM* level used being limited with respect to the total variability (n=10). In many cases the BBN-based results are more stable and reflect the fact that two bridges are very similar (also in terms of site) and should therefore have almost the same fragility. Other cases, however, may point at the possibility that the model is missing some important parameter. This is the subject of ongoing further research on the BBN structure.

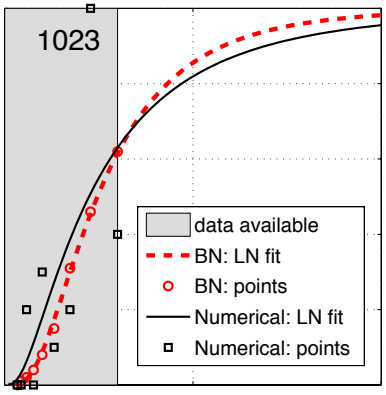


Figure 6: BBN-based and numerical fragility curves for sample bridge 1023. Gray shade denotes intensities where FE analyses support the curves: everything above 0.3g is an extrapolation of the LN fit.

Figure 7 provides a first global assessment of model performance, in terms of statistics of the fragilities obtained from the BBN (red) and the numerical analyses (black). Mean and 16th – 84th fractiles of all curves are computed at each PGA (i.e. 16th is the lowest one).

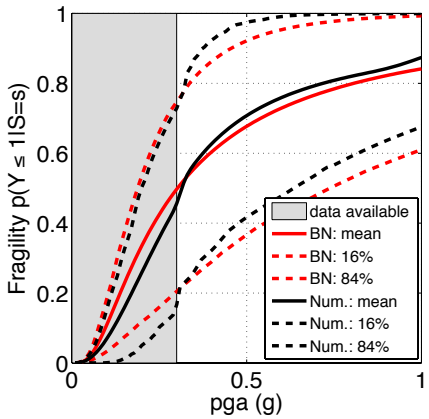


Figure 7: Global comparison of BBN-based and numerical fragilities for LD.

The model provides an excellent representation of the global variability of fragility curves over the considered DB (which is not minor, since the median PGA (PGA corresponding to 0.5 fragility value) varies between the 16th and the 84th fractiles by more than 200%.

Further proof of overall good performance (for the intended purpose) is given in Figure 8, which shows, for all bridges, the MAF of exceeding the LD limit state. The cloud of points gathers around the one-to-one line corresponding to the perfect match between the BBN-based and the numerical fragilities.

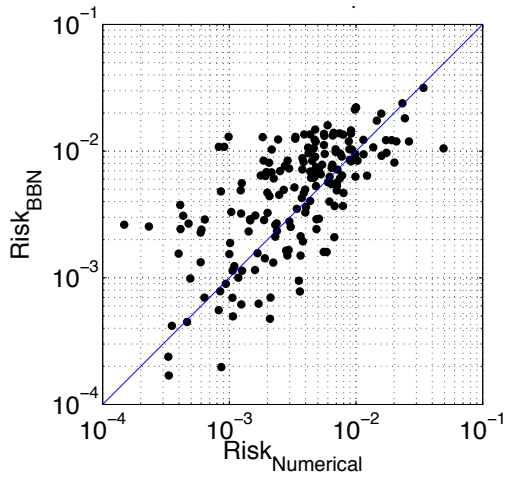


Figure 8: Global comparison of BBN-based and numerical fragilities in terms of normalized risk of LD.

Finally, a major advantage of the BBN model is realized in its capability to deal with cases of incomplete information. Typically, amongst the selected input parameters for the BBN, the most expensive ones to acquire, not readily available, are f_c, f_y and A_s. If one or more of these parameters are unknown, the BBN model can be still used. One approach is to take minimum and maximum estimated values for each unknown parameter and compute the corresponding fragility. Alternatively, evidence for the parameter is not provided to the model, and the uncertainty on the parameter, described by its unconditional distribution from the DB, will be propagated through the network to quantify Y_{LD}. Figure 9 shows an example, for bridge 1019, of the variability to be expected from lack of information on one or combinations of the parameters f_c, f_y and A_s. Sensitivity analysis highlighted how the most influential parameters are TP_As and SP_As, and indeed this is reflected in the fragilities of Figure 9.

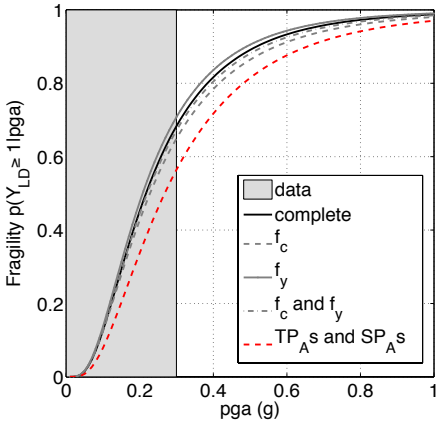


Figure 9. Sensitivity of BBN-based fragility to missing information.

4. CONCLUSIONS

This paper introduced a BBN model for the prediction of the light damage seismic fragility curve of RC girder bridges. The network structure is expert-driven, while the CPTs are trained based on the seismic response of existing 390 bridges, obtained by inelastic analysis with recorded ground motions.

The model is meant for use in prioritization of performance assessment and upgrade/retrofit interventions. Quality of prediction is considered already satisfactory for the intended purpose, even though improvements can be sought both in the BBN structure, but most importantly in the supporting response database (with increased number of motions per intensity level, leading to more stable targets, and with a more balanced database of bridges).

An important advantage of the BBN model is that it can be used even with partial inputs, even though all parameters have been chosen to be easy and relatively cheap to obtain. Any input will constrain the predicted fragility bringing it closer to the fragility of the considered bridge. In the worst case of no information about a bridge, the model will attribute to it the mean fragility over the DB.

Future research will focus on the extension of this model to include higher states of damage, and the effects of deterioration. Furthermore, this will be integrated in transportation network analysis for seismic risk management. The model, proposed for bridges, is believed to provide an important stimulus for extension to other structural typologies.

5. ACKNOWLEDGMENTS

Partial funding from the Italian Department of Civil Protection (DPC-Reluis project 2014-2017, project RS6), is gratefully acknowledged.

6. REFERENCES

Baker, J.W., Jayaram, N. (2008). "Correlation of spectral acceleration values from NGA ground motion models." *Earthq. Spectra*, 24(1), 299-317.

Bensi, M., Der Kiureghian, A. Straub, D. (2014). Framework for Post-Earthquake Risk Assessment and Decision Making for Infrastructure Systems. *ASCE-ASME Journal of Risk and Uncertainty in Eng. Systems, Part A: Civil Engineering*.

Boore, D.M., Atkinson, G.M. (2008). "Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01 s and 10.0 s." *Earthq. Spectra*, 24(1), 99-138.

Borzi, B., Ceresa, P., Franchin, P., Noto, F., Calvi, G.M., Pinto P.E. (2015). "Seismic Vulnerability of the Italian Roadway Bridge Stock." *Earthq. Spectra* (10.1193/070413EQS190M)

Bradley, B. (2013). "Ground motion selection for seismic risk analysis of civil infrastructure", Handbook of seismic risk analysis and management of civil infrastructure systems, Chapter 4 (Tesfamariam and Goda eds), Woodhead Publ. Ltd, UK, ISBN 978-0-85709-268-7.

Calvi G.M, Pinto PE, Franchin P (2013) "Chapter 17: Seismic Design Practice in Italy", Bridge Engineering Handbook – 2nd Edition, Chen W-F and Duan L (eds), CRC Press, Boca Raton, FL, ISBN 9781439852187

Cooper, G.F., Herskovits, E. (1992). "A Bayesian method for the induction of probabilistic networks from data." *Machine Learning*, 9, 309-347.

Franchin, P., Pinto, P.E. (2009). "Allowing Traffic over Mainshock-Damaged Bridges." *Journal of Earthq. Eng.*, 13(5), 585–599.

INGV [2004] "Redazione della mappa di pericolosità sismica prevista dall'Ordinanza PCM 3274 del 20 marzo 2003. Rapporto conclusivo per il Dip. di Prot. Civile, INGV, Milano-Roma" (in Italian).

Jalayer, F., Cornell, C. A. (2009). "Alternative non-linear demand estimation methods for probability-based seismic assessments." *Earthq. Eng. & Struct. Dyn.*, 38(8), 951-972.

Jalayer, F., Franchin, P., Pinto, P. E. (2007). "A scalar damage measure for seismic reliability analysis of RC frames." *Earthq. Eng. & Struct. Dyn.*, 36(13), 2059-2079.

Lin, T., Haselton, C.B., Baker, J.W. (2013). "Conditional spectrum-based ground motion selection. Part I: Hazard consistency for risk-based assessments." *Earthq. Eng. & Struct. Dyn.*, 42(12), 1847-1865.

McKenna, F., Scott, M.H., Fenves, G.L. (2010). "Nonlinear Finite-Element Analysis Software Architecture Using Object Composition." *Journal of Computing in Civil Engineering*, 24, pp.95–107.

Norsys Software Corp. (2006). Netica TM Application. <http://www.norsys.com>, downloaded , July 20.

Tesfamariam, S., Liu, Z. (2013). "Seismic risk analysis using Bayesian belief networks", Ch. 7. Handbook of seismic risk analysis and management of civil infrastructure systems (Tesfamariam and Goda eds), Woodhead Publ. Ltd, UK, ISBN 978-0-85709-268-7.

Vamvatsikos, D. (2013). Derivation of new SAC/FEMA performance evaluation solutions with second-order hazard approximation. *Earthq. Eng. & Struct. Dyn.*, 42(8), 1171-1188.