Natural Hazards
and Earth System
Sciences

Open Access

EGU

Discussions

1 **Risk assessment of liquefaction-induced hazards using Bayesian**

2 **network based on standard penetration test data**

3

4 Tang Xiao-Wei[1], Qiu Jiang-Nan[2], Hu Ji-Lei [3,*]

5 [1]State Key Laboratory of Coastal and Offshore Engineering. Dalian University of Technology, Dalian,

6 116024, China. Institute of Geotechnical Engineering, Dalian University of Technology, Dalian,

7 116024, China

8 [2]Faculty of Management and Economics, Dalian University of Technology, Dalian, 116024, China

9 [3]School of Civil Engineering and Mechanics, Huazhong University of Science and Technology, Wuhan,

10 430074, China

11 Correspondence to: Hu Ji-Lei (hujl@hust.edu.cn)

12

13 **Abstract.** Liquefaction-induced hazards are responsible for considerable damages to engineering

14 structures during major earthquakes. Presently, there is not any effective empirical approach that can

15 assess different liquefaction-induced hazards in one model, such as sand boils, ground cracks,

16 settlement, and lateral spreading, due to the uncertainties and complexity of multiple related factors of

17 seismic liquefaction and liquefaction-induced hazards. This study used Bayesian network method to

18 integrate multiple important factors of seismic liquefaction, sand boils, ground cracks, settlement and

19 lateral spreading into a model based on standard penetration test historical data, so that the constructed

20 Bayesian network model can assess the four different liquefaction-induced hazards together for free

21 fields. In the study case, compared with the artificial neural network technology and the Ishihara and

22 Yoshimine simplified method, the Bayesian network method performed a better classification ability,

23 because its prediction probabilities of *Accuracy*, *Brier score*, *Recall*, *Precision*, and area under the curve

24 of receiver operating characteristic (*AUC* of *ROC*) are better, which illustrated that the Bayesian

25 network method is a good alternative tool for risk assessment of liquefaction-induced hazards.

Natural Hazards
and Earth System
Sciences
Open Access
Discussions
EGU

1   Furthermore, the performances of the application of the BN model in estimating liquefaction-induced

2   hazards in the Japan's Northeast Pacific Offshore Earthquake also prove the correctness and reliability

3   of it compared with the liquefaction potential index approach. Except for assessing the severity of

4   hazards induced by soil liquefaction, the proposed Bayesian network model can also predict whether the

5   soil is liquefied or not after an earthquake, and it can deduce the process of a chain reaction of the

6   liquefaction-induced hazards and do backward reasoning, the assessment results from the proposed

7   model could provide informative guidelines for decision-makers to detect damage state of a field

8   induced by liquefaction.

9   **1   Introduction**

10   Prediction of liquefaction potential and assessment of liquefaction-induced hazards are two great and

11   closely related problems. The former problem just needs to answer whether the soil is liquefied or not

12   after an earthquake, whereas the latter problem not only needs to predict whether liquefaction-induced

13   hazards occur or not after soil liquefaction but also needs to assess the severity of different hazards

14   induced by liquefaction. Prediction of liquefaction potential of foundation soils is only the first step of

15   assessment of liquefaction hazards, which was well studied in the recent decades, such as the simplified

16   methods (Seed and Idriss 1971, 1982; Starks and Olsen 1995; Stokoe and Nazarian 1985) based on

17   standard penetration test (SPT), cone penetration test (CPT) and shear wave velocity measurement,

18   laboratory testing, numerical methods, and empirical liquefaction models (Goh 1994; Pal 2006; Toprak

19   et al. 1999) based on historical data. What is more important to the engineers is the effect of

20   liquefaction-induced hazards on foundation or superstructure after seismic liquefaction, but relatively

21   few studies have focused on it (Juang et al. 2005).

22   Field evidence of liquefaction-induced hazards in historical earthquakes mainly consisted of sand boils,

23   ground cracks, settlement and tilting of structures, and lateral spreading failures. Several methods have

24   been proposed to calculate these hazards, including numerical simulation method, laboratory test

25   approach, and field-testing-based method. Although recent advances in physical model experiment and

26   computational modeling of liquefaction-induced ground deformation were quite promising, challenges

27   remain yet on the critical unresolved problems, such as, without a perfect physical numerical model for

28   totally describing the complicated mechanic characteristics of soils, high expense and difficulty

1   associated with obtaining and testing high quality undisturbed samples of loose sandy soils,

2   time-consuming, and large costs of manpower and material resources. Therefore, empirical liquefaction

3   models based on a historical database compiled from historical earthquakes are likely best suited to

4   provide a simple, reliable, and direct way to assess liquefaction-induced hazards for many years to come

5   in the field of geotechnical earthquake engineering (Zhang et al. 2002). As for the empirical liquefaction

6   method, liquefaction potential index (LPI) has been used to characterize liquefaction-induced hazards

7   worldwide that is proposed by Iwasaki et al. (1982). After that several approaches based on the LPI,

8   such as damage severity index (DSI) developed by Juang et al. (2005) that allowed for evaluation of the

9   severity of liquefaction-induced ground damage at or near foundations, the Ishihara inspired LPIISH

10 extended by Maurer et al. (2015) that was found to be consonant with observed surface effects and

11 showed improvement over LPI in mitigating false-positive predictions, and liquefaction severity

12 number (LSN) developed by Tonkin and Taylor (2013) that was developed following the liquefaction

13 damage observations from the Canterbury Earthquake Sequence to reflect the damaging effects of

14 shallow liquefaction on residential land and foundations. In addition, generalized analytical or empirical

15 techniques for estimating a single type of ground failures (e.g. settlement or lateral spreading) induced

16 by liquefaction have been proposed in the decades (Youd and Perkins 1987; Youd et al. 2002; Goh and

17 Zhang 2014; Ishihara and Yoshimine 1992; Zhang et al. 2002; Wu and Seed 2004; Cetin et al. 2009;

18 Juang et al. 2013). With the rapid development of computer technology and mathematics, many new

19 artificial intelligence methods for assessing liquefaction-induced ground deformation (Wang and

20 Rahman 1999; Baziar and Ghorbani 2005; Javadi et al. 2006; Garcia et al. 2008; Rezania et al. 2011)

21 have been developed based on the historical data. However, these methods cannot assess sand boils and

22 ground cracks, and they only can estimate a single type of hazards, e.g. lateral spreading or settlement.

23 Because there is not a generic model for calculating or assessing sand boils, ground cracks, lateral

24 spreading, and settlement simultaneously, and then evaluating the overall severity of hazards induced by

25 liquefaction after an earthquake at present, it is necessary to develop a generic model for assessing all

26 types of liquefaction-induced hazards at a given site subjected to an earthquake. The primary objective

27 of this paper is to use the Bayesian network (BN) technology to integrate soil liquefaction, liquefaction

28 potential index, the four types of hazards (ground cracks, sand boils, lateral spreading, and settlement)

Natural Hazards
and Earth System
Sciences
Open Access
Discussions
EGU

1   induced by liquefaction and severity of liquefaction-induced hazards (describing the overall situation of

2   a site) into one model based on SPT historical data for deduce the process of chain reaction of the

3   hazards from an earthquake event to seismic liquefaction to liquefaction-induced hazards, that cover the

4   shortage of the existing simplified methods only assessing one single liquefaction-induced hazard. The

5   BN model is trained and tested separately using two different datasets of field case histories. The results

6   of the BN model for evaluation of liquefaction-induced hazards are compared with an artificial neural

7   network (ANN) model to verify its effectiveness and robustness using several performance indexes.

8   The remainder of this paper is organized as follows. At first, why used BN technology to assess the

9   hazards induced by seismic liquefaction was explained, and then construction of a BN model for

10   liquefaction-induced hazards was presented. Thirdly, the study case was used to verify the effectiveness

11   and robustness of the BN model compared with an ANN model and the Ishihara and Yoshimine

12   simplified method (Ishihara and Yoshimine, 1992) using several performance indexes, and the BN

13   model was applied for evaluating hazards induced by liquefaction in the 2011 Tohoku earthquake in

14   Japan that illustrated its availability. Finally, advantages and results of the BN model were discussed

15   with other methods, such as the ANN model and the LPI method. The conclusions and future works

16   were presented at the end of this paper.

2   Why Bayesian network?

18   Assessment of liquefaction-induced hazards is a complex engineering problem due to the heterogeneous

19   nature of soils, the participation of a large number of factors involved, and uncertainties of these factors.

20   The existed methods either were developed by statistics or only could assess a type of hazards such as

21   settlement or lateral spreading, and they hardly considered the effects of uncertainties on the

22   performances of the models, especially which purely relied on data-driven whereas ignoring the effects

23   of empirical knowledge or domain knowledge on the assessment of liquefaction-induced hazards.

24   However, the latest development in Bayesian network (BN) technology can combine empirical

25   knowledge and historical data that just right provides new opportunities to develop better performing

26   tools for complex problems in probabilistic terms such as the problem of liquefaction-induced hazards.

27   The BN is one of the most effective theoretical models for knowledge representation and reasoning

28   under the influence of uncertainty and high non-linear relationships among variables (Pearl 1988).

29   Firstly, BN can offer a rationalist and coherent theory for reasoning under the condition of various

Natural Hazards
and Earth System
Sciences

Discussions

EGU

1   uncertainties (e.g. uncertainties of parameters, models, and the domain knowledge) and complexities

2   that are described in terms of subjective beliefs or probabilities to reflect the interdependent relationship

3   between variables. Moreover, it can integrate different kinds of domain knowledge and multi-source

4   information or various quantitative factors and qualitative factors into a consistent system, and facilitate

5   multiple hazards and their interdependencies in a model. Especially, it allows not only sequential

6   inference (from causes to results) but also reverses inference (from results to causes) under conditions

7   of complete data, even incomplete data, and provide efficient frameworks for probabilistic updating and

8   the assessment of component performance when given new evidence.

9   In recent decades, the BN has been widely applied in risk analysis of the engineering field, such as

10 catastrophic risk (Li et al. 2010a; Li et al. 2010b; Li et al. 2012), earthquake risk damage (Bayraktarli et

11 al. 2005; Bayraktarli and Faber 2011; Bensi et al. 2009 and 2014), embankment dam risk (Zhang et al.

12 2011; Xu et al. 2011; Peng and Zhang 2012), landslide hazards (Song et al. 2012; Liang et al. 2012),

13 and soil liquefaction (Bayraktarli 2006; Hu et al. 2015), in which application of the BN in assessing

14 liquefaction-induced damages have never been involved before. An important sign is given by the fact

15 that, the numbers of relevant publications in this field from 2001 to 2015 were obtained by retrieving

16 with "BN" and "risk analysis" in Web of Science database annually increased from about 3 to 50 (as

17 shown in Fig. 1). Especially in recent five years, BN technology is generally popular with engineers and

18 researchers for assessment of risk analysis. The BN technology has been proven to be a robust method

19 for risk analysis.

20 **3   A BN model for liquefaction-induced hazards**

21 **3.1 Probabilistic reasoning of BN**

22 A Bayesian network combines the knowledge of graph theory and statistics theory that consists of

23 models and arcs or links with conditional probabilities. Its inference algorithms are based on the

24 Bayesian rule, the chain rule, and the conditional independence rule as follows

25 $$P(X \mid Y) = \frac{P(Y \mid X) \cdot P(X)}{P(Y)} ,$$ (1)

26 $$P(x_1, \cdots, x_n) = P(x_1)P(x_2 \mid x_1) \cdots P(x_n \mid x_1, x_2, \cdots, x_{n-1}) ,$$ (2)

27 $$P(x_1, \cdots, x_n) = \prod_{i=1}^{n} P(x_i \mid \pi(x_i)) ,$$ (3)

Natural Hazards
and Earth System
Sciences
Discussions

1   Where: P(Y) is called prior probability, P(X|Y) is one's belief for hypothesis X upon observing evidence

2   Y which is known as posterior probability, P(Y|X) is the likelihood that Y is observed if X is true.

3    $\pi(x_i)$ is a set of values for the parents of X_i.

4   A generic BN model for liquefaction-induced hazards (as shown in Fig. 2) is constructed with domain

5   knowledge to illustrate how to reason in the assessment of liquefaction-induced hazards. There are three

6   types of nodes in the BN model: (1) input nodes, they are soil parameters (SP), earthquake parameters

7   (EP) and field conditions (FC) that are factors of seismic liquefaction; (2) state nodes, they are

8   liquefaction potential (LP) and liquefaction potential index (LPI) that separately shows whether soil is

9   liquefied or not, and expresses how much the degree of soil liquefaction is; and (3) output nodes, they

10   are liquefaction-induced hazards (LH), such as lateral spreading, settlement, ground cracks, and sand

11   boils, which express the severity of liquefaction-induced hazards. The nodes are connected by twelve

12   arcs or links. In the risk assessment of liquefaction-induced hazards, if evidence comes from input

13   nodes, a posteriori probability or belief of the target variable (LH) being in a certain state (e.g. severe)

14   can be derived by the three above formulas of the BN method as follows:

15
$$P(\text{LH} = severe \mid SP, EP, FC) = \frac{P(\text{LH} = severe, SP, EP, FC)}{P(SP, EP, FC)}$$
$$= \frac{P(SP, EP, FC \mid \text{LH} = severe)P(\text{LH} = severe)}{P(SP, EP, FC)}$$
$$= \frac{\sum P(SP, EP, FC, LP, LPI \mid \text{LH} = severe)\sum P(\text{LH} = severe)}{\sum P(\text{LH}, LP, LPI, SP, EP, FC)}$$

16
$$P(\text{LH}, LP, LPI, SP, EP, FC) = P(SP) \cdot P(FC) \cdot P(EP \mid SP, FC) \cdot P(LP \mid SP, EP, FC)$$
$$\cdot P(LPI \mid \text{LH}, FC) \cdot P(\text{LH} \mid SP, EP, FC, LP, LPI)$$

17   **3.2 Construction of a BN model for liquefaction-induced hazards**

18   Strong earthquakes can cause liquefaction and therewith ground failures in the form of four types of

19   hazard, such as sand boils, ground cracks, settlement-induced tilting of structures, and lateral spreading.

20   Table 1 lists some factors of liquefaction potential (LP), liquefaction potential index (LPI), four types of

21   hazards induced by liquefaction and severity of liquefaction-induced hazards (SLH), in which LP and

22   LPI are used to describe the state of soil liquefaction, four types of liquefaction-induced hazards are

23   used to identify different damages and their severity after seismic liquefaction, and SLH is a

24   comprehensive index intergrading the four types of hazards indexes together that is used to describe the

25   overall severity of disasters after liquefaction. Also, Table 1 lists some empirical modeling methods that

26   can be used as domain knowledge to construct a BN model of liquefaction-induced hazards. Hu et al.

Natural Hazards
and Earth System
Sciences

Open Access

Discussions

EGU

1   (2016) constructed a BN model for liquefaction potential (as shown in Fig. 3) that considered twelve

2   factors: the magnitude of the earthquake (ME), epicentral distance (ED), duration of the earthquake

3   (DE), peak ground acceleration (PGA), fines content (FC), soil type (ST), average particle size (D_{50}),

4   SPT number (SPTN), vertical effective stress (σ_v'), groundwater table (GT), depth of soil deposit (DSD)

5   and the thickness of the soil layer (TSL). In the seismic parameters, the bigger the magnitude of the

6   earthquake is and the nearer epicentral distance is, the longer duration of the earthquake is and the

7   bigger PGA is, the greater liquefaction potential is. In the soil parameters, anti-liquefaction behavior of

8   soil will largely change with different fines content, when the fines content increases within 30%, the

9   liquefaction strength decreases, but when the fines content is beyond 30%, the liquefaction strength

10   increases as fines content increases, and when the fines content is more than 50% (silt and sandy silt),

11   the soil is hardly liquefied. In addition, fines content can change the type of soil. Normally, purified clay

12   and silt cannot be liquefied, whereas poorly graded sand and silty sand are easily liquefied. The bigger

13   average particle size is, and the bigger SPT number is, the smaller the probability of soil liquefaction is.

14   In the field condition, the deeper depth of soil deposit is, the bigger vertical effective stress is, the

15   increase of pore water pressure is hard to overcome vertical effective stress, so soil liquefaction is not

16   easy to occur. In addition, shallow groundwater table and thin soil can partly reduce the probability of

17   soil liquefaction. So a state node (LPI) and output nodes (sand boils, ground cracks, lateral spreading,

18   settlement, and SLH) just need to be added into the exist BN model of liquefaction potential (as shown

19   in Fig. 3) based on the generic BN model in Fig. 2. A new BN model for liquefaction-induced hazards

20   (as shown in Fig. 4) was constructed according to domain knowledge of the hazards in Table 1. The

21   factor, ground slope that effects GC and LS, was not considered in the BN model of

22   liquefaction-induced hazards because the collected datum in this paper did not contain it.

23   Earthquake liquefaction-induced hazards are a chain reaction process originated from the earthquake

24   event to soil liquefaction to its pertinent hazards, in which different input values deduce different

25   liquefaction states and different the degrees of liquefaction, and then the outputs of the former system

26   (e.g. LP) are used as input information of the latter system to result in different hazard events (e.g. sand

27   boils, lateral spreading, et al.). The whole process of earthquake liquefaction-induced hazards can be

28   described as follow: at beginning of explosion earthquake, earthquake parameters, soil characteristics

29   and field conditions are considered as control variables, and their prior probabilities are calculated by

30   parameter learning. And then a posterior probability of the output variable (e.g. LP) can be reasoned for

31   estimating whether the event can be triggered or not. If the event occurs, its conditional probability is

32   replaced by the posterior probability that will be considered as evidence variable for input. Finally, a

Natural Hazards
and Earth System
Sciences

Open Access

EGU

Discussions

1   posterior probability of the latter event (e.g. LP) will be calculated using the new conditional probability

2   of the former event for estimating its grade. Repeat the above process until grades of all hazards events

3   are identified at last.

4   **4   The study case**

5   **4.1 Dataset**

6   In this study, the data set including 442 SPT borings from post-earthquake in-situ tests at liquefied (245

7   SPT borings) and non-liquefied (197 SPT borings) sites in Taiwan, Japan, and the USA were collected.

8   332 SPT borings containing 184 liquefied sites and 148 non-liquefied sites were used to train the BN

9   model, and the rest 110 SPT borings were used to test the effectiveness and robustness of the BN model.

10   Because of incomplete data, e.g. the proportion of missing data for D_{50} is about 15.2%, the proportion

11   of missing data for vertical effective stress is about 29.4%, and the proportion of missing data for the

12   thickness of soil layer is about 38.9%, expectation-maximization (EM) algorithm (Lauritzen 1995) was

13   used for training the 332 SPT borings data to obtain conditional probability table for the BN model due

14   to its more robustness compared with other algorithms and its suitableness for the data contained large

15   missing values. Briefly, the EM method is an iterative algorithm to find maximum likelihood estimation

16   or maximum a posterior estimation of parameters that repeatedly takes a Bayesian net and uses it to

17   obtain a better one by doing an expectation (E) step followed by a maximization (M) step until

18   convergence of the algorithm. In the E step, it uses regular Bayesian net inference with the existing

19   Bayesian net to compute the expected value of all the missing data, and then the M step finds the

20   maximum likelihood Bayesian net given the now extended data (e.g. original data plus expected values

21   of missing data). The data from the 1999 Chi-Chi earthquake in Taiwan (M_w=7.6) were downloaded

22   from   http://www.ces.clemson.edu/chichi/TW-LIQ/In-situ-Test.htm   and

23   http://peer.berkeley.edu/lifelines/research_projects/3A02/. Information of special "small magnitude"

24   was from the 1957 Daly City (California) earthquake (M_w=5. 3) of the USA and the 1987 Whittier

25   Narrows earthquake (M_w=5. 9) of the USA that refers to Cetin et al. (2000). The data from the 2011

26   Tohoku earthquake in Japan (M_w=9. 0) were provided by the research center for the management of

27   disaster and the environment, Tokushima University in Japan. Effects of liquefaction induced by these

28   earthquakes observed, including sand boils, settlement, ground cracks, and lateral spreading (as shown

29   in Fig. 5), resulted in destroying cropland, blocking the channels, and severe damages or collapses of

30   many buildings, highways, bridges, harbor facilities and other infrastructure components.

31   The grading standard of liquefaction and liquefaction-induced hazards according to domain knowledge

Natural Hazards
and Earth System
Sciences
Discussions
Open Access

1   is as seen in Table 2, e.g. LPI is divided into four grades according to Iwasaki (1982), non-liquefaction

2   (LPI=0), slight liquefaction (0<LPI≤5), moderate liquefaction (5<LPI≤15), and serious liquefaction

3   (LPI>15). SLH is divided into four grades according to the experience of disaster in the field of

4   engineering that is described in detail as seen in Table 3. After that, according to the descriptions of

5   SLH, a statistic summary of liquefaction-induced hazards data is presented in Fig. 6. It can be seen that

6   (1) liquefaction doesn't have to induce hazards, but the occurrence of the liquefaction-induced hazards

7   is based on liquefaction, (2) LPI is not a good index for a description of the severity of

8   liquefaction-induced hazards because the efficacy of the LPI framework and accuracy of derivative

9   liquefaction hazards are uncertain, e.g. serious liquefaction of the LPI occurs in the none SLH in Fig. 6

10   (1) and slight liquefaction of the LPI occur in the severe SLH in Fig. 6 (4), but there is a rule that the

11   bigger the LPI is, the corresponding severity of the liquefaction-induced hazards is, (3) SB, S, GC, and

12   LS are macroscopical phenomena of hazards induced by liquefaction, and there is also a trend that the

13   bigger the values of these indexes are, the more severity of the SLH is, (4) the classifications for the

14   four different types of hazards in Fig. 6 almost accord with the descriptions of the field ground damage

15   status in Table 3.

16   Fig. 7 shows ratios of all influence factors for the severe status of the SLH. It can be easily seen that

17   most severe damage sites suffered from a big or super earthquakes with long loading, their epicentral

18   distance were near to the earthquake sources, and their PGA should be higher enough. As for soil

19   characteristics, pure sand or silty sand with moderate fines particles and moderate D_{50} easily result in

20   severe damage, but few sites with gravel soil and sandy silt were also suffered from severe damage. The

21   damaging phenomena showed that even though gravel and sandy silt were not easy to be liquefied, if

22   they were liquefied in a big earthquake, they also resulted in severe damage. The small SPTN means

23   that the sand soil is so loose that it more easily results in settlement and lateral spreading after

24   liquefaction because loose sand is more easily compressed and flowing during seismic liquefaction. As

25   for field conditions, the shallowly deposited sandy soil layer is thin so that its effective stress is small,

26   and groundwater table is near to the ground surface, this kind of field is easily suffered from severe

27   damage. The above laws fit well with engineering practical experience. The sum of ratios of three

28   variables, such as D_{50}, σ_v' and the thickness of soil layer, is not 1 due to data missing mentioned in

29   Section of Dataset.

4.2 Performance indexes

31   In this section, in order to comprehensively evaluate the performances of the two probabilistic models

1   for liquefaction-induced hazards, several performance indexes are used, which are *Accuracy*, *Prediction*,

2   *Recall*, area under the curve of the receiver operating characteristic (*AUC* of *ROC*), and *Brier score*. The

3   details of these indexes are briefly introduced as follows.

4   The *Accuracy* is a measure of the percentage of correctly classified instances for each class, which is a

5   widely used metric for measuring the overall performance of a classifier. For instance, the *Accuracy* is

6   equal to 0.9 that means 90% of data can be correctly classified. However, it does not mean that the

7   accuracies of each class are all 90%, perhaps the accuracy of one class is high but others are very low.

8   Therefore, the evaluation of the predictive capability based on the *Accuracy* alone can be misleading

9   when a class imbalance exists in a data set. Other indexes, *Precision*, *Recall*, and *AUC* of *ROC*, should

10 be used to further measure the performance of each class for a model or a classifier.

11 The *Recall* refers to the probability of detection of a class, and it measures the proportion of correctly

12 predicted positive instances among the actual positive ones. If a classifier can achieve a higher *Recall*

13 for a class that means it can detect more positive instances of the class. The *Precision* refers to the

14 proportion of true positives among the instances predicted as positive for a single class, but it cannot

15 measure how the classifier detects the actual positive instances. A classifier with high *Precision* but with

16 a lower *Recall* is less useful because it cannot detect significant positive instances, especially in the risk

17 assessment, security and warning are major concerns, a good classifier should detect more positive

18 instances and relatively high accuracy of prediction, namely, have a high Recall and acceptable

19 *Precision*.

20 The ROC curve is a graphical plot developed by the false positive rate (the proportion of all negatives

21 that still yield positive test outcomes) on the x-axis and the true positive rate or *Recall* on the y-axis that

22 can present an overly optimistic view of an algorithm's performance. The *AUC* of *ROC* is the area

23 between the horizontal axis and the *ROC* curve that is a comprehensive scalar value representing a

24 classifier's expected performance. The *AUC* of *ROC* is between 0.5 and 1, where a value close to 0.5 is

25 less precise, while a value close to 1.0 is more precise. Therefore, the bigger the AUC of *ROC* is, the

26 better prediction the performance of the classifier can get.

27 *Brier score* is used to measure the quality of probabilistic forecasts for discrete events, which was

28 proposed by Brier (1950). Suppose that on each of n occasions an event can occur in only one of r

29 possible classes and on one such occasion, i , the forecast probabilities are $f_{i1}, f_{i2}, \cdots, f_{ir}$, that the event

30 will occur in classes $1, 2, 3, \cdots, r$, respectively. The *Brier score* (B) can be defined by

Natural Hazards
and Earth System
Sciences
Discussions

1 $$B = \frac{1}{n}\sum_{j=1}^{r}\sum_{i=1}^{n}\left(f_{ij} - E_{ij}\right)^2, \qquad\qquad (4)$$

2 Where: $\sum_{i=1}^{r} f_{ij} = 1, i = 1, 2, 3, \cdots, n$. E_{ij} takes the value 1 or 0 according to whether the event occurred in

3 class j or not. For instance, in the study case of this paper, 110 SPT borings used for testing (n=110),

4 SLH has four classes, none, minor, medium, and severe (r=4), in which a probability or confidence

5 statement (f_{ij}) was made for each SPT boring instance. The *Brier score* is between 0 and 2, where the

6 minimum *Brier score* is 0 for perfect prediction and the maximum value is 2 for the worst possible

7 prediction.

8 **5   Results**

9 **5.1 Comparison of predictive results**

10 The comparisons of probability predictive results of the two models, the BN model (in Fig. 4) and the

11 ANN model used same parameters, are shown in Table 4. Comparing the probability of *Accuracy*,

12 except for LS, the *Accuracy* values of the BN model for the other types of hazards and SLH are

13 obviously more than the ANN's, and comparing the *Brier score*, its values of the BN model except for

14 LS and SLH are less than ANN's, which both illustrate that the overall performances of the BN model

15 are better than the ANN model. As for each type hazard induced by liquefaction and SLH, *Recall*,

16 *Precision*, and *AUC* of *ROC* of BN model for each class are almost more than the ANN's, which also

17 represents that BN model is better than the ANN model. Therefore, the proposed BN approach is better

18 than the ANN technology, and its performances are acceptable for monitoring and forecast of seismic

19 liquefaction-induced hazards. In addition, comprising the computation time, the BN model using the

20 EM algorithm outperforms the ANN model contained 20 hidden layers using radial basis function,

21 requiring 36 iteration steps (about 19.8 CPU seconds) to converge to a stable state, therefore the

22 convergence rate of the BN model is faster than the ANN model in this paper.

23 Furthermore there are not any effective simplified methods for estimating ground cracks and sand boils,

24 and simplified methods for calculating lateral spreading, e.g. Bartlett and Youd (1995), Wang and

25 Rahman (1999), Goh et al. (2014), et al. that need the free face ratio or the ground slope, however, these

26 two factors were not contained in collected datum in this paper. Therefore, ground cracks, sand boils

27 and lateral spreading cannot be estimated by simplified methods, while settlement can be calculated by

28 the I&Y simplified method proposed by Ishihara and Yoshimine (1992) as shown in Table 4. It can be

Natural Hazards
and Earth System
Sciences

Open Access

Discussions

EGU

1   clearly seen that predictive results of data-driven methods, such as BN and ANN, are better than I&Y

2   simplified method, but the simplified method can get a certain value (as shown in Fig. 8) rather than an

3   interval value or a probability. In addition, the simplified method was only constructed by a relationship

4   among relative density, the factor of safety against liquefaction (F_L) and volumetric strain (ε_v) to

5   estimate the settlement of a site, in which the factor of safety against liquefaction was obtained by

6   integrating earthquake intensity and SPTN by empirical formulas or empirical coefficient so that it

7   might cause calculation error and result in considerable prediction error, such as the predictive result of

8   small settlement as shown in Table 4, the precision of the simplified method is only 0.069. However, the

9   data-driven methods can integrate multiple factors of liquefaction-induced hazards into a model so that

10   they can get better predictive performances than simplified methods.

11   **5.2 Casual reasoning using the BN model**

12   Based on the developed BN model, probabilities of the liquefaction-induced hazards were inference

13   through causal reasoning in this section. The third column in Table 5 shows the posterior probabilities of

14   all grades of LP, LPI, and its induced hazards. It can be seen that when the input information about

15   earthquake parameters, soil characteristics, and field conditions are not known, the probabilities of all

16   grades of in each output variable are no big different except for LS and SB due to its serious imbalance

17   data in different grades. However, when a site is determined to be liquefied and the probability of yes

18   status of LP is changed into 100%, it can be seen in the fourth column that probabilities of none status

19   of LPI and all hazards decrease in some extent but probabilities of other states in LPI and all hazards

20   increase largely. Furthermore, if the site is serious liquefied, the probability of yes status of LP and

21   serious status of LPI are changed into 100% as seen in the fifth column in Table 5, the probabilities of

22   grades (except none status) of all hazards continue increasing, in which GC occurs with 66.1%

23   probability, serious sand boil occurs with 69.6% probability, the probability of Big LS is 9.5%, the

24   probability of big settlement is 49.8%, and the probability of severe of SLH is 64.1%. It shows that

25   liquefaction-induced hazards are much more severe in the serious liquefied site. And then

26   macro-liquefaction phenomena, such as GC and serious SB, are observed, probabilities of the big status

27   of other hazards continue increasing slightly as seen in the sixth column in Table 5 so that the predictive

28   results are more close to the actual situation. Therefore, according to the above deduce process, the BN

29   model can calculate the posterior probability of LP based on conditional probabilities of input variables

30   for estimating whether a site is liquefied or not, if it is liquefied, its posterior probability will be

31   considered as input information for the latter variable's prediction. Reasoning step by step like that can

32   give all predictive results of liquefaction-induced hazards. In addition, when the prior probabilities of all

1   input variables, such earthquake parameters, soil characteristics and field conditions, can be determined

2   actually at first, the predictive performances of all hazards will be improved to a large content. For

3   instance, a site suffered from a super earthquake with long duration, survey shows that the SLH is

4   severe with big settlement, no lateral spreading, serious sand boils and ground crack, the input variables

5   of the site are that the epicentral distance is near, PGA is higher, soil type is sand with some fine particle,

6   its D_{50} is medium, SPT number shows the sand is loose, $\sigma_{v'}$ is small, groundwater table is shallow, both

7   depth and thickness of the sand layer are moderate. The reasoning probability value of LP is 99.9%, LPI

8   is identified as serious with 43.8% probability, GC doesn't occur with 51.4% probability that doesn't

9   match with the survey result, SB is identified as many with 76.5% probability, LS is identified as none

10   with 85.0% probability, settlement is identified as big with 53.1% probability, and SLH is identified as

11   severe with 52.6% probability according to the input information. After that the site can be determined

12   as a liquefied area with serious liquefaction degree, so LP is intended to be 100% probability, and the

13   probability of serious state of LPI is also intended to be 100%, then the probabilities of all hazards are

14   changed, GC occurs with probability 100% that matches with the survey result, LS is identified as none

15   with 100% probability (increased by 15%), settlement is identified as big with 100% probability

16   (increased by 46.9%), and SLH is identified as severe with 100% probability (increased by 47.4%).

17 **5.3 Diagnostic reasoning using the BN model**

18   In order to detect what kind of situations more easily result in severe damage of a site, the most

19   probable explanations of LP (Yes), LPI (Serious), GC (Yes), SB (Many), LS (Big) and S (Big) are

20   inference using the diagnostic reasoning capabilities of the BN model, the results are shown in Table 6.

21   It can be seen that the loose silty sand (medium D_{50}) contained moderated fines particles deposits

22   shallowly (small $\sigma_{v'}$) on a site with a low underground water level that is suffered from a super

23   earthquake with moderate duration, and the epicentral distance is moderate, which is more easily

24   subjected to be liquefied. The most probable explanations of GC and many SB are same as LP's under

25   conditions of serious or moderate soil liquefaction, but the most probable explanations of big LS and S

26   are tiny different from LP's, such as PGA and sandy soil, the reason is that big LS and big settlement

27   need more seismic intensity than occurrences of sand boils and ground cracks, and sand is more easily

28   flowing and compressed after liquefaction than the sand containing fines particles. In addition,

29   occurrences of big LS and big settlement often accompany by many sand boils, while ground cracks

30   may or may not occur. The above results are according with the analysis results in Fig. 7. In addition, If

31   soil characteristics, field conditions, and hazards of the field are known, how strong the earthquake

32   intensity (status of magnitude of the earthquake, duration of earthquake, PGA and epicentral distance)

Natural Hazards
and Earth System
Sciences
Discussions

1   resulting in the liquefaction-induced hazards of the field can be estimated using backward inference

2   ability of the BN method, that can provide some references for aseismatic design.

3   **5.4 Sensitivity analysis of liquefaction-induced hazards**

4   Sensitivity analysis can detect how much each factor has an impact on the target variable. In this section,

5   mutual information is used to assess sensibility that is a measure of the mutual dependence between the

6   two variables. The results of mutual information for different liquefaction-induced hazards are

7   computed separately in the BN model as shown in Table 7. It is shown that the thickness of soil layer is

8   the most sensitive variable for GC, and relative important factors are the depth of the soil layer, D_{50}, and

9   the duration of earthquake compared to the rest other factors. As for SB, groundwater table is the most

10 sensitive variable, and relative important factors are the thickness of the soil layer, SPTN, the duration

11 of the earthquake, PGA, depth of soil layer, and σ_v' compared to the rest other factors. As for S, PGA is

12 the most sensitive variable, and relative important factors are SPTN, the duration of the earthquake, and

13 depth of soil layer compare to the rest other factors. As for LS, PGA is also the most sensitive variable,

14 and relative important factors are D_{50}, the thickness of the soil layer, depth of soil layer, and soil type

15 compared with the rest other factors. These results are highly consistent with the domain knowledge in

16 Table 1. When comprehensive comparisons with the most sensitive factors and relative important

17 factors of the four types of liquefaction-induced hazards and SLH are conducted, the duration of

18 earthquake, PGA, SPTN, depth of soil deposit, and the thickness of soil layer are more important than

19 other factors, because their overlap sections are more than three times. In these five factors, a

20 combination of SPTN and the earthquake intensity described by the duration of earthquake and PGA

21 can detect the degree of soil liquefaction. And then the depth of soil deposit and the thickness of soil

22 layer combine with relative density (determined by SPTN) based on the degree of soil liquefaction to

23 result in the soil volumetric strain. Consequently, liquefaction-induced hazards, e.g. settlement and

24 lateral spreading, can be estimated. Therefore, in order to mitigate seismic liquefaction-induced hazards

25 of a filed, dispose of relative density of sandy soil, the depth of sandy soil deposit and the thickness of

26 sandy soil layer are the crucial factors.

27 **6   Application of the BN model**

28 The above BN model is applied to assess the liquefaction-induced hazards in the Japan's Northeast

29 Pacific Offshore Earthquake on March 11, 2011. The research regions are Ibaraki prefecture, Chiba

30 prefecture, Saitama prefecture, Kanagawa prefecture and Tokyo city that contain 196 investigation sites.

31 The assessment results of the SLH are shown in Fig. 9, in which the blue circle means little to none

Natural Hazards
and Earth System
Sciences
Discussions

liquefaction-induced hazards, the green circle means minor liquefaction-induced hazards, the yellow circle means medium liquefaction-induced hazards, the orange circle means severe liquefaction-induced hazards and the red circle means prediction error. In the 196 real fields, the prediction accuracies of four types of liquefaction-induced hazards, respectively, are 99.50% for lateral spreading, 81.63% for sand boils, 80.61% for settlement, 89.8% for ground cracks, and 84.1% for SLH. In addition, the prediction accuracies of four different levels of SLH (Little to none, Minor, Medium and Severe) separately are 79.83%, 84.62%, 81.25% and 79.83%, which prove the validity of the BN model in general. However, the prediction accuracies of the LPI approach (Iwasaki et al. 1982) for the four different levels of SLH are 36.96%, 8.82%, 68%, and 42.22% separately, which are much worse than the prediction results of the BN model.

In this earthquake, the areas with greater losses and a larger number of liquefaction sites are in Ibaraki prefecture and Tokyo city that are near to the sea compared with other places, which contained 78 sites with different degrees of hazards, and in which disasters of about 50 sites are medium or severe. It can be known from Table 3 that free site with medium disasters or severe disasters is easy to cause sand boils and ground cracks phenomenon, lateral spreading and settlement, which is the result of foundation failure, further to cause building damage and bridge collapse. Therefore the BN model of assessing liquefaction-induced hazards not only accurately assess range values of lateral spreading and settlement, quantity of sand boils, and ground cracks phenomenon, but also accurately predict the severity of hazards induced by liquefaction for foundation on the free site, and then qualitatively assess disasters of a building or other structures according to the engineering experiences relationship between disasters of the foundation and losses of upper structural, which provide guidelines for prevention and mitigation engineering of natural disasters.

7 Discussion

This paper constructed a probability model for liquefaction-induced hazards using BN technology. As a means of probability inference, BN technology offers several specific advantages over other methods in the evaluation of catastrophic, which primarily can support a good platform for integrating different kinds of hazards and their interdependencies into a consistent system (Li et al. 2010b). However, existing empirical methods for estimating hazards induced by seismic liquefaction, such as the empirical formulas constructed by Youd et al. (1987, 2002) and the multiple linear regression (MLR) model constructed by Goh and Zhang (2014) for estimating lateral spreading, the different simplified procedures for estimating the settlement proposed by Ishihara and Yoshimine (1992), Zhang et al.

Natural Hazards
and Earth System
Sciences
Open Access
Discussions
EGU

1   (2002), Wu and Seed (2004), and Juang et al. (2013), which only can assess a single type of ground

2   failures, and cannot estimate ground cracks and sand boils. Another method, the LPI approach, can

3   quantify liquefaction severity of a site by providing a unique value for the entire soil column instead of

4   several factors of safety per layer. However, calibration of LPI to observe liquefaction severity is

5   limited, and the efficacy of the LPI framework and accuracy of derivative liquefaction-induced hazards

6   are uncertain (Maurer et al. 2014). When the value of LPI is big (LPI>15), whereas there may be not

7   phenomenon of settlement or ground cracks, but when the value of LPI is small (LPI<5), whereas a

8   serious sand boils with a long duration and wide scale, lateral spreading and severe subsidence occur, so

9   the real SLH are largely not consistent with the prediction results of the LPI approach, which can be

10 found in Fig 6 and in the prediction results of the LPI approach in the section of Application of the BN

11 Model. In fact, LPI only can reflect the degree of liquefaction in a site, but cannot detect real situations

12 of ground damage, the relation between LPI and the types of liquefaction-induced hazards was not

13 being examined systematically that might be a qualitative relation in some extent.

14 Comparing the two methods, the BN method and the ANN method, although both of them are

15 supervised learning methods, the BN method is a generative model, whereas the ANN method is a

16 discriminative model, therefore, the BN method can obtain the joint probability distribution of the

17 parameters so that it can describe distributions of data in statistical terms and combine with a strong

18 probabilistic theory enabling it to give an objective interpretation, and its computation time is faster

19 than discriminative models such as the ANN method, even when sample size increases quickly, the BN

20 method can quickly converge to the true model, especially, when the data contains hidden parameters,

21 the BN method can still develop a robust model, but the ANN method cannot do it (Correa et al. 2009).

22 As for the models, in the BN model, each node means a random variable that has actual meaning, and

23 the arrow between two nodes means causation, but each node in the ANN model is not random variable

24 that has not actual meaning, and the arrow between two nodes just means weighting functional

25 relationship that has not other meanings, such as causation or logistic relationship, making it difficult to

26 explain the results. In addition, except predicting the different hazards induced by liquefaction, the

27 constructed BN model in this paper also can predict liquefaction potential, the *Accuracy* of liquefaction

28 potential is 0.80 using the testing data in the dataset. If using the ANN technology, a new model should

29 be reconstructed by studying the training data to predict the *Accuracy* of liquefaction potential, whereas

30 the BN model can predict directly without retraining. Especially, the BN method can reason forward

31 and reason backward, such as assessing hazards induced by liquefaction with given earthquake

32 parameters, soil parameters, and field conditions, or predicting what kind of soil properties and field

Natural Hazards
and Earth System
Sciences
Discussions

1   conditions are when hazards are known after an earthquake, whereas the ANN method just only can do

2   forward reasoning.

8   Conclusions and future work

4   Given uncertainty and complexity of liquefaction-induced hazards, this paper developed a generic BN

5   model for estimating risk assessment of different hazards induced by seismic liquefaction based on

6   historical disaster data that provides a platform integrating a variety of information sources from

7   different fields and facilitates different hazards induced by liquefaction into a model. The findings in

8   this paper are as follows:

9   (1) After comparing with ANN technology using several performance indexes, the BN model achieved

10 a better *Accuracy* and a better *Brier score* for overall performance, and a better *Recall*, a better

11 *Precision*, and a better *AUC* of *ROC* for each damage state (e.g. sand boils, settlement et al.) and its

12 computation time is faster than the ANN method. This illustrates that the BN method is suitable for risk

13 assessment of liquefaction-induced hazards influenced by complex multiple factors. Comparing with

14 the I&Y simplified method for estimating settlement, the data-driven methods (BN and ANN) are

15 superior to it. Furthermore, the performances of the application of the BN model in estimating

16 liquefaction-induced hazards in the Japan's Northeast Pacific Offshore Earthquake also prove the

17 correctness and reliability of it compared with the LPI approach.

18 (2) The BN model can deduce the process of a chain reaction of the liquefaction-induced hazards and do

19 backward reasoning, such as inference from input variables (earthquake parameters, soil characteristics

20 and field conditions) to soil liquefaction to different hazards events or from soil liquefaction to different

21 hazards events to input variables. In addition, the most probable explanations of LP, Serious LPI, GC,

22 many SB, Big LS and Big S in the BN model show that the loose silty sand or sandy soil (medium D_{50})

23 contained moderated fines particles deposits shallowly (small σ_v') on a site with low underground water

24 level that is suffered a super earthquake with moderate duration, and the epicentral distance is moderate,

25 it is more easily suffered to be liquefied and severe hazards induced by liquefaction.

26 (3) After sensitivity analysis for liquefaction-induced hazards separately, the most sensitive factors of

27 different hazards are different. When comprehensively comparing with these important factors of the

28 four types of liquefaction-induced hazards, the duration of the earthquake, PGA, SPTN, depth of soil

29 deposit, and the thickness of soil layer are more important than other factors, which contribute to the

30 soil volumetric strain.

1   Because the occurrence of liquefaction may cause no damage, little or severe damage to the ground

2   surface or infrastructure, the constructed BN model in this paper can accurately assess severity of

3   hazards after seismic liquefaction, and provides guidelines to engineers with which fields need to be

4   dealt with, rather than dealing with all fields with liquefaction, thus, large of costs will be saved. In this

5   study, more historical data will be collected to update the conditional probability table to improve the

6   BN model, especially collecting historical data contained statues of small and medium of lateral

7   spreading due to a lack of them in this paper's data, and then the nodes of utility and decision actions

8   will be trying to add into the BN model to extend it, so that the new model can test how different actions

9   will result in different hazards and different expected utilities of loss. The results can serve as significant

10   information for decision making in the earthquake resistance and hazard reduction.

11   **Acknowledgments**

17

Natural Hazards
and Earth System
Sciences
Discussions

Open Access

EGU

1   **References**

2   Bardet J.P., Kapuskar M. Liquefaction sand boils in San Francisco during 1989 Loma

3   Prieta earthquake. Journal of Geotechnical Engineering, 119(3), 543-562, 1993.

4   Bartlett S.F., Youd T.L. Empirical prediction of liquefaction-induce lateral spread.

5   Journal of Geotechnical Engineering, 121(4), 316-329, 1995.

6   Bayraktarli Y.Y. Application of Bayesian probabilistic networks for liquefaction of

7   soil. 6th International Ph.D. Symposium in Civil Engineering, Zurich, 8, 23-26, 2006.

8   Bayraktarli Y.Y., Faber Michael H. Bayesian probabilistic network approach for

9   managing earthquake risks of cities. Georisk: Assessment and Management of Risk for

10   Engineered Systems and Geohazards, 5(1), 2-24, 2011.

11   Bayraktarli Y.Y., Ulfkjaer J.P., Yazgan Ufuk, et al. On the Application of Bayesian

12   Probabilistic Networks for Earthquake risk management. Proceedings of the Ninth

13   International Conference on Structural Safety and Reliability (ICOSSAR 05), Rome,

14   June, 20-23, 2005.

15   Baziar M.H., Ghorbani A. Evaluation of lateral spreading using artificial neural

16   networks. Soil Dynamics and Earthquake Engineering, 25(1), 1-9, 2005.

17   Bensi M.T., Kiureghian A.D., and Straub D. A Bayesian network framework for

18   post-earthquake infrastructure system performance assessment. Proceedings of TCLEE

19   2009, Lifeline Earthquake Engineering in a Multihazard Environment, 28 June-1 July.

20   Oakland, CA, 2009.

21   Bensi Michelle, Kiureghian A.D., and Straub D. Framework for post-earthquake risk

22   assessment and decision making for infrastructure Systems. ASCE-ASME Journal of

23   risk uncertainty engineering systems, Part A, Civil Engineering, 1(1), 04014003, 2014.

24   Brier G.W. Verification of forecasts expressed in terms of probability. Monthly

25   Weather Review, 78(1), 1-3, 1950.

1 Cetin K.O., Bilge H.T., Wu Jiaer, et al. Probabilistic model for the assessment of

2 cyclically induced reconsolidation (volumetric) settlements. Journal of Geotechnical

3 and Geoenvironmental Engineering, 135(3), 387-398, 2009.

4 Cetin K.O., Seed R.B., Kiureghian A.D., et al. SPT-based probabilistic and

5 deterministic assessment of seismic soil liquefaction initiation hazard. Pacific

6 Earthquake Engineering Research Repot No. PEER-2000/05, 2000.

7 Correa M., Bielza C., Teixeira Pamies J. Comparison of Bayesian networks and

8 artificial neural networks for quality detection in a machining process. Expert Systems

9 with Applications, 36(3), 7270-7279, 2009.

10 Garcia S.R., Romo M.P., Botero E. A neurofuzzy system to analyze

11 liquefaction-induced lateral spread. Soil Dynamics and Earthquake Engineering, 28(3),

12 169-180, 2008.

13 Goh Anthony T.C. Seismic liquefaction potential assessed by neural networks. Journal

14 Geotechnical Engineering, 120(9), 1467-1480, 1994.

15 Goh Anthony T.C., and Zhang W.G. An improvement to MLR model for predicting

16 liquefaction-induced lateral spread using multivariate adaptive regression splines.

17 Engineering Geology, 170(20), 1-10, 2014.

18 Hu Jilei, Tang Xiaowei, and Qiu Jiangnan. A Bayesian network approach for

19 predicting seismic liquefaction based on interpretive structural modeling. Georisk:

20 Assessment and Management of Risk for Engineered Systems and Geohazards, 9(3),

21 200-217, 2015.

22 Hu Jilei, Tang Xiaowei, and Qiu Jiangnan. Assessment of Seismic liquefaction

23 potential based on Bayesian network constructed from domain knowledge and history

24 data. Soil Dynamic and Earthquake Engineering, 89, 49-60, 2016.

25 Ishihara Kenji and Yoshimine Mitsutoshi. Evaluation of settlements in sand deposits

26 following liquefaction during earthquakes. Soils and Foundations, 32(1), 173-188,

27 1992.

1   Iwasaki T., Tokida K., Tatsuoka F., et al. Microzonation for soil liquefaction potential

2   using simplified methods. Proc. 3rd International Earthquake Microzonation

3   Conference, Seattle, 1319-1330, 1982.

4   Javadi A.A, Rezania M., and Nezhad M.M. Evaluation of liquefaction-induced lateral

5   displacements using genetic programming. Computers and Geotechnics, 33(4-5),

6   222-233, 2006.

7   Juang C. Hsein, Ching Jianye, Wang Lei, et al. Simplified procedure for estimation of

8   liquefaction-induced settlement and site-specific probabilistic settlement exceedance

9   curve using cone penetration test (CPT). Canadian Geotechnical Journal, 50(10),

10 1055-1066, 2013.

11 Juang C. Hsein, Yuan Haiming, Li David Kun, et al. Estimating severity of

12 liquefaction-induced damage near foundation. Soil Dynamics and Earthquake

13 Engineering, 25(5), 403-411, 2005.

14 Lauritzen S.L. The EM algorithm for graphical association models with missing data.

15 Computational Statistics and Data Analysis, 19(2), 191-201, 1995.

16 Li L.F., Wang J.F., Leung H., et al. Assessment of catastrophic risk using Bayesian

17 network constructed from domain knowledge and spatial data. Risk Analysis, 30(7),

18 1157-1175, 2010a.

19 Li L.F., Wang J.F., and Leung H. Using spatial analysis and Bayesian network to

20 model the vulnerability and make insurance pricing of catastrophic risk. International

21 Journal of Geographical Information Science, 24(12), 1759-1784, 2010b.

22 Li L.F., Wang J.F., Leung H., et al. A Bayesian method to mine spatial data sets to

23 evaluate the vulnerability of human beings to catastrophic risk. Risk Analysis, 32(6),

24 1072-1092, 2012.

25 Liang W.J., Zhuang D.F., Jiang D., et al. Assessment of debris flow hazards using a

26 Bayesian Network. Geomorphology, 171-172, 94-100, 2012.

Maurer B.W., Green R.A., Cubrinovski M., et al. Evaluation of the liquefaction potential index for assessing liquefaction hazard in Christchurch, New Zealand. Journal of Geotechnical and Geoenvironmental Engineering, 140(7), 04014032, 2014.

Maurer B.W., Green R.A., Taylor O-D.S. Moving towards an improved index for assessing liquefaction hazard: lessons from historical data. Soils and Foundations, 55(4), 778-787, 2015.

Pal Mahesh. Support vector machines-based modeling of seismic liquefaction potential. International Journal for Numerical and Analytical Methods in Geomechanics, 30(10), 983-996, 2006.

Pearl, J. Probabilistic Reasoning in Intelligent Systems. Morgan Kaufmann Publishers, San Mateo, California, 1988.

Peng M., Zhang L.M. Analysis of human risks due to dam break floods - part 1: A new model based on Bayesian networks. Natural Hazards, 64(1), 903-933, 2012.

Rezania M., Faramarzi A., and Javadi A.A. An evolutionary based approach for assessment of earthquake-induced soil liquefaction and lateral displacement. Engineering Applications of Artificial Intelligence, 24(1), 142-153, 2011.

Seed H.B., Idriss I.M. Simplified procedure for evaluating soil liquefaction potential. Journal of the Soil Mechanics and Foundation Engineering Division, 97(9), 1249-1273, 1971.

Seed H.B., Idriss I.M. Ground motions and soil liquefaction during earthquakes. Monograph series. Earthquake Engineering Research Institute, Berkeley, Calif., 1982.

Song Yiquan, Gong Jianhua, Gao Sheng, et al. Susceptibility assessment of earthquake-induced landslides using Bayesian network: A case study in Beichuan, China. Computers & Geosciences, 42, 189-199, 2011.

Starks T.D., Olsen S.M. Liquefaction resistance using CPT and field case histories. Journal of Geotechnical Engineering, 121(12), 856–869, 1995.

Natural Hazards
and Earth System
Sciences
Discussions

EGU

Open Access

1   Stokoe K.H., Nazarian S. Use of Rayleigh waves in liquefaction studies. In: Woods R

2   D, ed. Measurement and Use of Shear Wave Velocity for Evaluating Dynamic Soil

3   Properties. ASCE, New York, 1-17, 1985.

4   Tonkin & Taylor Ltd. Liquefaction vulnerability study. Report to Earthquake

5   Commission, Tonkin & Taylor ref. 52020.0200/v1.0 [prepared by S. van Ballegooy

6   and P. Malan], 2013.

7   Wang J., Rahman M.S. A neural network model for liquefaction-induced horizontal

8   ground displacement. Soil Dynamics and Earthquake Engineering, 18(8), 555-568,

9   1999.

10   Weber, P., Medina-Oliva, G., Simon, C., et al. Overview on Bayesian Networks

11   application for dependability, risk analysis and maintenance areas. Engineering

12   Application of Artificial Intelligence, 25(4), 671-682, 2012.

13   Wu Jiaer, and Seed R.B. Estimation of liquefaction-induced ground settlement (case

14   studies). Proc. 5th International Conference on Case Histories in Geotechnical

15   Engineering, New York, April 13, 13-17, 2004.

16   Toprak S., Holzer T.L., Bennett M.J., et al. CPT- and SPT-based probabilistic

17   assessment of liquefaction potential. Proc. 7th U.S.-Japan Workshop on Earthquake

18   Resistant Design of Lifeline Facilities and Countermeasures Against Soil Liquefaction,

19   MCEER, Seattle, 69-86, 1999.

20   Xu Y., Zhang L.M., and Jia J.S. Diagnosis of embankment dam distresses using

21   Bayesian networks. Part II. Diagnosis of a specific distressed dam. Canadian

22   Geotechnical Journal, 48(11), 1645-1657, 2011.

23   Youd T.L. Geologic effects-liquefaction and associated ground failure. Proceedings of

24   the Geologic and Hydrologic Hazards Training Program, U.S. Geological Survey

25   Open-File Report, 210-232, 1984.

1 Youd T.L., Hansen C.M., and Bartlett S.F. Revised multilinear regression equations

2 for prediction of lateral spread displacement. Journal of Geotechnical and

3 Geoenvironmental Engineering, 128(12), 1007-1017, 2002.

4 Youd T.L., Perkins D.M. Mapping of liquefaction severity index. Journal of

5 Geotechnical Engineering, 113(11), 1374-1392, 1987.

6 Zhang G., Robertson P.K., and Brachman R.W.I. Estimating liquefaction-induced

7 ground settlements from CPT for level ground. Canadian Geotechnical Journal, 39(5),

8 1168-1180, 2002.

9 Zhang L.M., Xu Y., Jia J.S., et al. Diagnosis of embankment dam distresses using

10 Bayesian networks. Part I. Global-level characteristics based on a dam distress

11 database. Canadian Geotechnical Journal, 48(11), 1630-1644, 2011.

12

Natural Hazards
and Earth System
Sciences

Discussions

1   Table 1. Factors of liquefaction and its induced hazards and empirical modeling

2   methods.

Category	Liquefaction and its induced hazards	Factors	Empirical methods
Liquefaction state	Liquefaction potential (LP)	The magnitude of earthquake, epicentral distance, the duration of earthquake, peak ground acceleration (PGA), fines content, soil type, average particle size (D_{50}), SPT number (SPTN), vertical effective stress (σ_v'), groundwater table, depth of soil deposit, and the thickness of soil layer	Hu et al. (2016)
	Liquefaction potential index (LPI)	LP, depth of soil deposit, and the thickness of soil layer	Iwasaki et al. (1982); Maurer et al. (2015)
Liquefaction -induced hazards	Sand boils (SB)	LP, LPI, depth of soil deposit, the thickness of soil layer, and groundwater table	Bardet and Kapuskar (1993)
	Ground cracks (GC)	LP, LPI, D50, depth of soil deposit, the thickness of soil layer, and ground slope (θ)	Youd (1984)
	Lateral spreading (LS)	LP, LPI, PGA, The magnitude of earthquake, epicentral distance, depth of soil deposit, the thickness of soil layer, D50, and θ	Bartlett and Youd (1995); Wang and Rahman (1999); Goh et al. (2014)
	Settlement (S)	LP, LPI, PGA, depth of soil deposit, the thickness of soil layer, soil type, LS, SB	Zhang, Robertson, and Rrachamn (2002); Cetin et al. (2009); Juang et al. (2013)
Comprehensive index	Severity of liquefaction-induced hazards (SLH)	LP, LPI, SB, GC, LS, S	-

Natural Hazards
and Earth System
Sciences
Discussions

1   Table 2. Grading standard for liquefaction and liquefaction-induced hazards.

Factor	No. of grade	Grade	Data number	Range
Liquefaction potential	2	None	245	-
		Yes	197	-
Liquefaction potential index	4	Non-liquefaction	145	0
		Slight liquefaction	97	$0 < LPI \leq 5$
		Moderate liquefaction	106	$5 < LPI \leq 15$
		Serious liquefaction	94	$15 < LPI$
Settlement (m)	4	None	238	0
		Small	23	$0 < S \leq 0.1$
		Medium	54	$0.1 < S \leq 0.3$
		Big	127	$0.3 < S$
Sand boils	4	None	275	-
		Less	21	-
		Medium	11	-
		Many	135	-
Ground crack	2	None	106	-
		Yes	336	-
Lateral Spreading (m)	4	None	437	0
		Small	0	$0 < LS \leq 0.1$
		Medium	0	$0.1 < LS \leq 0.3$
		Big	5	$0.3 < LS$
Severity of liquefaction-induced hazards	4	Little to None	238	-
		Minor	28	-
		Medium	46	-
		Severe	130	-

Natural Hazards
and Earth System
Sciences
Discussions

1   Table 3. Description of the severity of liquefaction-induced hazards.

Severity of liquefaction-induced hazard	Description of field ground status
Little to None	Non-liquefaction. There is no sand boils phenomenon and no ground failure.
Minor	Slight liquefaction. The phenomenon of the sand boil is sporadic, but there is no ground failure.
Medium	Moderate liquefaction. There is a medium sand boil phenomenon, which has a short duration, small gushing quantity and small scale, the quantity of surface subsidence is less than 3% of the sand layer thickness that can cause structural damage, and the tiny crack of ground occurs, but there is no lateral spreading.
Severe	Serious liquefaction. There is a serious sand boil phenomenon, which has a long duration, large gushing quantity and wide scale, surface largely crazes, and lateral spreading and severe subsidence affect structures' services. The quantity of surface subsidence is more than 3% of the sand layer thickness.

1 Table 4. Comparison of predictive performances of liquefaction-induced hazards.

Category	Method	Accuracy	Brier score	Damage state	Recall	Precision	AUC of ROC
Ground cracks	BN	0.909	0.070	Yes	0.742	0.920	0.780
				None	0.975	0.920	0.962
	ANN	0.873	0.091	Yes	0.581	0.947	0.641
				None	0.987	0.857	0.949
Sand boils	BN	0.918	0.106	Many	0.932	0.911	0.558
				Medium	-	-	-
				Less	0.857	0.857	0.667
				None	0.932	0.948	0.982
	ANN	0.736	0.130	Many	0.591	0.813	0.652
				Medium	-	-	-
				Less	0.000	0.000	0.000
				None	0.932	0.733	0.973
Settlement	BN	0.836	0.110	Big	0.867	0.703	0.845
				Medium	0.815	0.957	0.745
				Small	1.000	0.600	1.000
				None	0.840	0.933	1.000
	ANN	0.745	0.130	Big	0.667	0.741	0.815
				Medium	0.444	0.857	0.542
				Small	0.000	0.000	0.000
				None	1.000	0.735	1.000
	I&Y Simplified Method	0.727	-	Big	0.862	1.000	-
				Medium	0.778	0.840	-
				Small	0.667	0.069	-
				None	0.600	1.000	-
Lateral spreading	BN	0.955	0.024	Big	1.000	0.286	1.000
				Medium	-	-	-
				Small	-	-	-
				None	0.954	1.000	1.000
	ANN	0.982	0.018	Big	0.000	-	0.000
				Medium	-	-	-
				Small	-	-	-
				None	1.000	0.982	1.000
Severity of liquefaction-induced hazards	BN	0.936	0.124	Severe	0.935	0.967	0.879
				Medium	0.857	0.900	0.626
				Minor	0.875	0.700	1.000
				None	0.980	0.980	0.980
	ANN	0.718	0.117	Severe	0.710	0.710	0.785
				Medium	0.333	0.636	0.776
				Minor	0.000	0.000	0.000
				None	1.000	0.746	1.000

Natural Hazards
and Earth System
Sciences
Discussions

Open Access

EGU

1   Table 5. The posterior probabilities of partial output variables

| Output variable	Grade	Risk probability 1	Risk probability 2	Risk probability 3	Risk probability 4
Liquefaction potential	Yes	0.572	**1**	**1**	**1**
	None	0.428	**0**	**0**	**0**
Liquefaction potential index	Serious	0.220	0.385	**1**	**1**
	Moderate	0.257	0.450	**0**	**0**
	Slight	0.207	0.136	**0**	**0**
	None	0.316	0.286	**0**	**0**
Ground cracks	Yes	0.239	0.408	0.661	**1**
	None	0.761	0.592	0.339	**0**
Sand boils	Many	0.304	0.515	0.696	**1**
	Medium	0.0554	0.0809	0.0676	**0**
	Less	0.0506	0.0725	0.0355	**0**
	None	0.590	0.331	0.201	**0**
Lateral Spreading	Big	0.076	0.0811	0.095	0.0927
	Medium	0.0698	0.0702	0.0764	0.0782
	Small	0.0698	0.0702	0.0764	0.0782
	None	0.784	0.778	0.748	0.751
Settlement	Big	0.255	0.362	0.498	0.523
	Medium	0.179	0.229	0.140	0.120
	Small	0.162	0.198	0.212	0.240
	None	0.404	0.212	0.150	0.116
SLH	Severe	0.277	0.416	0.641	0.746
	Medium	0.168	0.225	0.0966	0.697
	Minor	0.140	0.174	0.147	0.114
	None	0.415	0.185	0.116	0.697

1   Table 6. The most probable explanation of LP, serious LPI, GC, many SB, big LS and

2   big S in the BN model

Factor	LP	LPI	GC	SB	LS	S
Earthquake magnitude	Super	Super	Super	Super	Super	Super
Epicentral distance	Medium	Medium	Medium	Medium	Medium	Medium
Duration of earthquake	Medium	Medium	Medium	Medium	Medium	Medium
PGA	Medium	Medium	Medium	Medium	Higher	Higher
Fines content	Medium	Medium	Medium	Medium	Medium	Medium
Soil type	SM	SM	SM	SM	SP	SP
D_{50}	Medium	Medium	Medium	Medium	Medium	Medium
SPT No.	Loose	Loose	Loose	Loose	Loose	Loose
σ'	Small	Small	Small	Small	Small	Small
Groundwater table	Shallow	Shallow	Shallow	Shallow	Shallow	Shallow
Depth of soil layer	Shallow	Shallow	Shallow	Shallow	Shallow	Shallow
Thickness of soil layer	Thin	Medium	Medium	Thin	Medium	Thin
LP	-	Yes	Yes	Yes	Yes	Yes
LPI	-	-	Serious	Moderate	Serious	Moderate
GC	-	-	-	None	Yes	None
SB	-	-	-	-	Many	Many
LS	-	-	-	-	-	None

Natural Hazards
and Earth System
Sciences
Discussions

1   Table 7. Sensitivity analysis of seismic liquefaction-induced hazards.

Factor	Mutual information				
	GC	SB	S	LS	SLH
The magnitude of earthquake	0.002	0.003	0.002	0.001	0.002
Epicentral distance	0.004	0.007	0.007	0.002	0.008
The duration of earthquake	**0.008**	**0.016**	**0.013**	0.002	**0.015**
PGA	0.004	**0.011**	**0.029**	**0.123**	**0.026**
Fines content	0.001	0.001	0.001	0.003	0.001
Soil type	0.001	0.002	0.006	**0.013**	0.005
D_{50}	**0.009**	0.001	0.002	**0.029**	0.003
SPTN	0.004	**0.017**	**0.017**	0.003	**0.019**
σ_v'	0.001	**0.010**	0.007	0.006	0.008
Groundwater table	0.000	**0.054**	0.003	0.002	0.004
Depth of soil deposit	**0.013**	**0.010**	**0.009**	**0.014**	**0.010**
The thickness of soil layer	**0.035**	**0.023**	0.006	**0.028**	0.005

Natural Hazards
and Earth System
Sciences
Discussions
Open Access
EGU
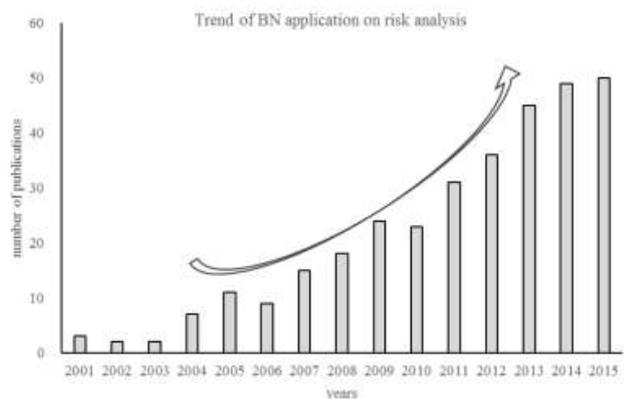
1

2
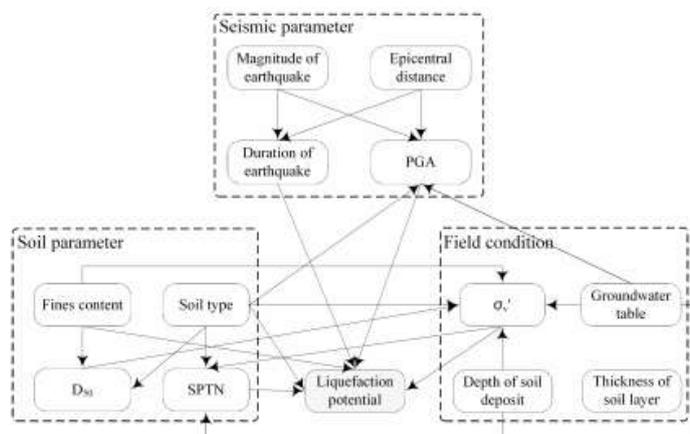
3   Figure 1. Increasing application of BN in risk analysis (update Weber et al. 2012).
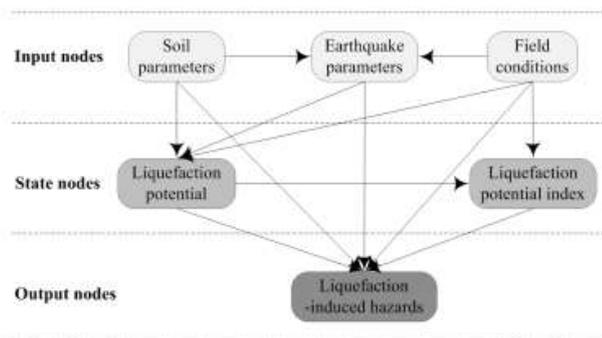
1

2

3   Figure 2. A generic BN for liquefaction-induced hazards.

Natural Hazards
and Earth System
Sciences
Open Access
Discussions

1

2

3   Figure 3. A BN model of seismic liquefaction (Hu et al. 2016).
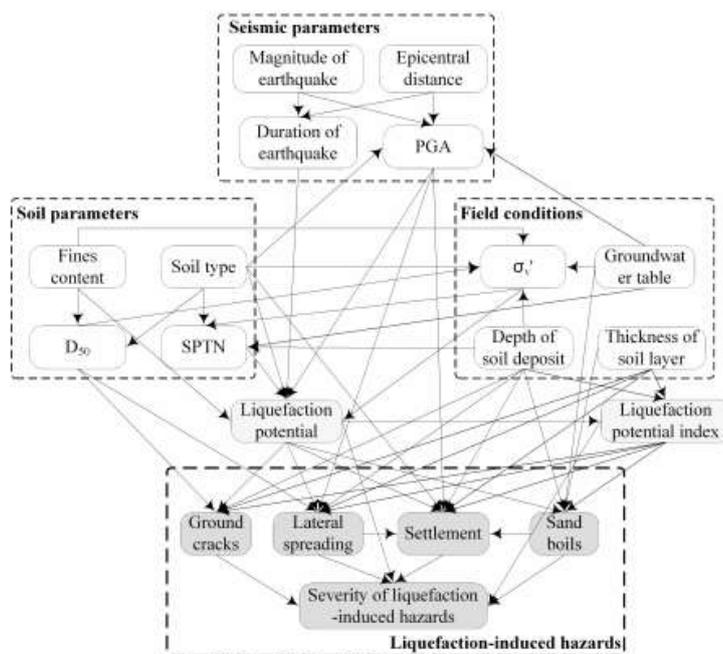
1
2
3   Figure 4. A BN model of seismic liquefaction-induced hazards.

Natural Hazards
and Earth System
Sciences

Discussions

1

2
3

4   Figure 5. Photos showing liquefaction-induced hazards during the Chi-chi earthquake

5   and the 2011 Tohoku earthquake: (1) Sand boils in Chikusei city; (2) Ground cracks at

6   Arakawa   River   in   Toda   city;   (3)   Settlement   at   Taichung   Port

7   (http://www.ces.clemson.edu/chichi/TW-LIQ/Liq-Album/Settlement-7.htm); (4) Lateral

8   spread   induced   failure   of   a   dike   in   Nantou

9   (http://www.ces.clemson.edu/chichi/TW-LIQ/Liq-Album/LatSpread-3.htm).

1
2   (1) Little to none SLH   (2) Minor SLH

3
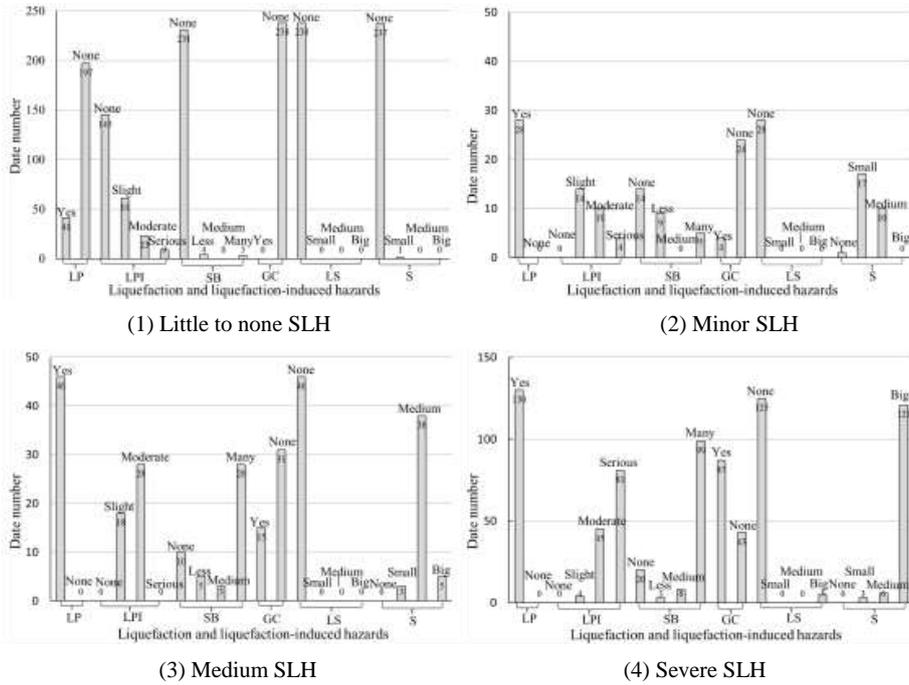4   (3) Medium SLH   (4) Severe SLH
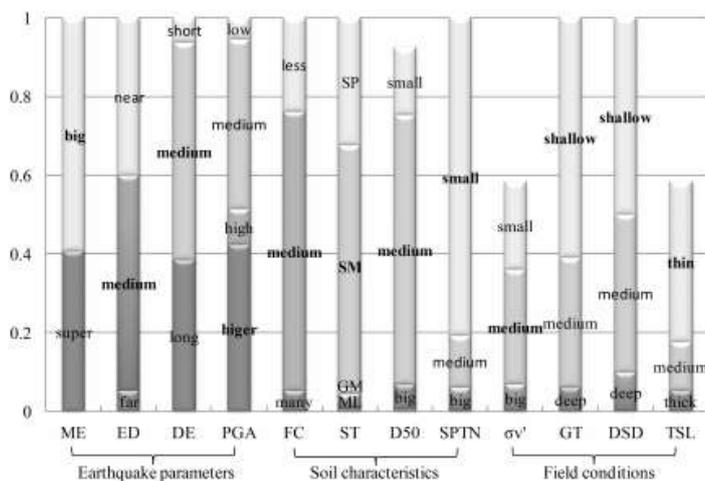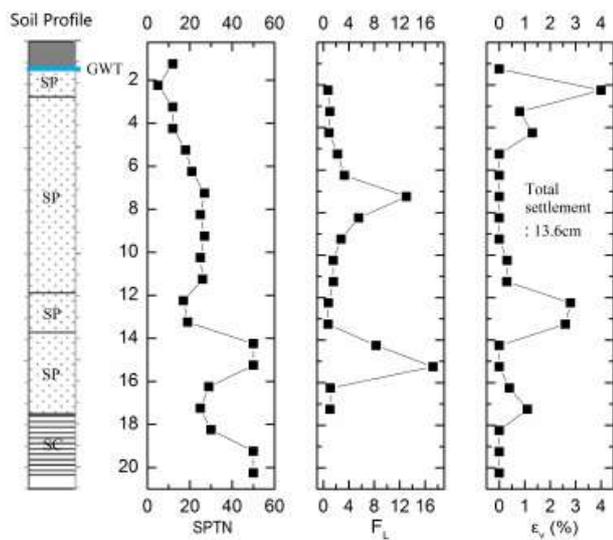
5

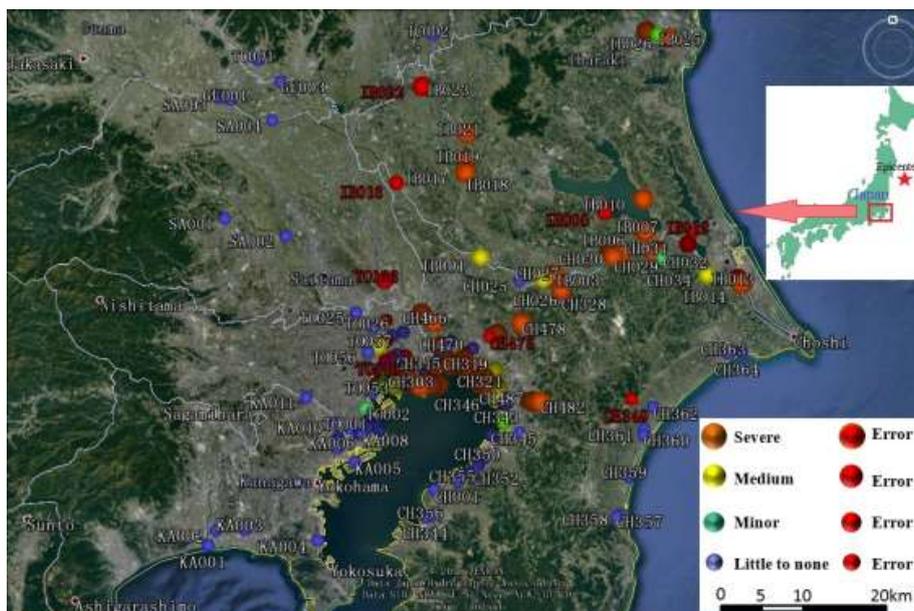6   Figure 6. Statistic summary of seismic liquefaction-induced hazards data.

1

2

3   Figure 7. Ratios of all influence factors for the severe status of SLH

1

2

3   Figure 8. Soil profile and estimate of settlement

Natural Hazards
and Earth System
Sciences
Open Access
Discussions

1
2

3   Figure 9. Assessment results of the severity of hazards induced by seismic liquefaction

4   in the northeast area of Japan in the 2011 Tohoku earthquake.