A Bayesian Network for Both Land Subsidence Risk and Soil Contamination Risk Evaluation in Large-Scale Reclaimed Lands of Shanghai, China

Jianxiu Wang[1,2](✉), Linbo Wu[1], Tianliang Yang[3], Tao Cui[1], Zhe Wang[1], Weiqiang Lin[1], Lijie Jin[4], and Tianrong Huang[5]

[1] College of Civil Engineering, Tongji University, Shanghai 200092, China
wang_jianxiu@163.com
[2] Key Laboratory of Geotechnicalal and Underground Engineering of Ministry of Education, Tongji University, Shanghai 200092, China
[3] Key Laboratory of Land Subsidence Monitoring, Prevention and Control of Ministry of Land and Resources, Shanghai Institute of Geological Survey, Shanghai 200072, China
[4] No. 1 Institute of Geology and Mineral Resource Exploration of Shandong Province, Jinan 250014, Shandong, China
[5] Shanghai Urban Construction Design and Research Institute (Group) Co., Ltd., Shanghai 200125, China

Abstract. Shanghai has the largest newly reclaimed area in China, and the associated soil contamination risk exists. To evaluate both land subsidence risks and soil heavy metal contamination (HMC) risks in Shanghai's coastal reclaimed regions, a Bayesian Network (BN) was established taking sixteen variables such as average initial void of underlying strata, land reclamation time, thickness of the reclaimed layer, and soil HMC into consideration. In the BN, the ultimate land subsidence since July 1st 2018 and the land subsidence velocity characterized by the land subsidence from July 1st 2018 to January 1st 2019 analytically evaluated in typical reclaimed regions of Shanghai were used to characterize land subsidence. Seven heavy metal elements, i.e., Zn, Cd, As, Ni, Cu, Pb and Cr were included into the HMC evaluation. Influence strength of the BN arcs and node sensitivity were analyzed. The BN analysis shows that the two study zones hold remarkable differences in field basic characteristics, geotechnical parameters and soil HMC distribution. The three land subsidence risk variables hold little correlation with the HMC distribution, but they are both correlated tightly with the basic characteristics of the study points.

Keywords: Bayesian Network · Land subsidence · Soil contamination · Land reclamation

1 Introduction

Over half of human population in the world lives within 60 km of coast, and the population proportion is increasing [1]. Consequently, large-scale development of land reclamation were conducted to meet the demand for land. Many countries, such as

© Springer Nature Switzerland AG 2020
A. G. Correia et al. (Eds.): ICITG 2019, SSGG, pp. 47–56, 2020.
https://doi.org/10.1007/978-3-030-32029-4_4

United States, South Korea, Netherlands and China reclaimed over 1000 km^2 lands [2], more than 1860 km^2 lands were reclaimed in China during 2000–2016, and a large portion of the lands held significant contamination risks [3]. Field surveys showed that contamination including heavy metal contamination within marine sediments and dredged materials existed in many regions [4], especially in river estuaries [5–7]. Many land reclamation programs using sediments of unknown quality for fills were conducted under poor environmental supervision especially in developing regions, but soil contamination usually was not considered in studies on those new lands.

The city of Shanghai, China obtained over 300 km^2 land using land reclamation [8], and the area is still expanding for the need of city development. Contamination of Zn [9], Cd [10–12], As [11] within surface sediments in some Shanghai's coastal zones were detected to be contaminated. The authors conducted series of field surveys on both geotechnical properties and soil environmental properties during 2016–2018, obtaining numbers of correlated data therein. Worrying surface subsidence risks and heavy metal contamination (HMC) were revealed within Shanghai's typical reclaimed lands.

Bayesian network (BN) is practical in multi-variables analysis [13, 14], and has been widely used in many cases, e.g., earthquake disaster chain analysis [15], marine transportation risk assessment [16], and coastal hazard risk reduction assessment [17].

This work aims to build a comprehensive BN to evaluate surface subsidence risks and HMC in Shanghai's typical reclaimed regions, and analyze correlations among the considered variables.

2 Materials and Methods

2.1 Shanghai's Typical Reclaimed Lands

Detailed borehole surveys and lab tests on both geotechnical data and soil environmental data was done since 2016 in typical reclaimed fields of Hengsha Island and Laogang Town, Shanghai [3]. Depth of the boreholes ranged from 45 m to 75 m, the investigated layers are listed in Table 1.

Table 1. Strata in the study regions.

Strata	Soil type	Strata	Soil type
①$_{3\text{-}1}$	Landfill	⑤$_{1\text{-}2}$	Silt interbedded with clay
①$_{3\text{-}2}$	Hydraulic fill	⑤$_{3}$	Silty clay
①$_{3\text{-}3}$	Mud	⑥	Silty clay
②$_{2}$	Muddy and silty clay	⑦$_{1}$	Sandy silt
②$_{3}$	Silt and sand	⑦$_{2}$	Silt and sand
③$_{1}$	Muddy and silty clay	⑧	Clay
④	Muddy clay	⑨	Sand with silt
⑤$_{1\text{-}1}$	Clay		

Using the data detected in 2016–2018, the authors evaluated surface subsidence risks on basis of one dimensional Terzaghi's soil consolidation theory [18]. The evaluated land subsidence caused by land reclamation was comprised of self-weight consolidation subsidence of reclaimed layers and compression consolidation subsidence of underlying sedimentary beddings. Land subsidence since July 1st 2018 (s_{18}) was evaluated to characterize the total land subsidence risk. The land subsidence from July 1st 2018 to January 1st 2019 (v_{18}) was evaluated to act as the land subsidence velocity in the analyzed points. And comprehensive land subsidence risk (R_c) simultaneously considering s_{18} and v_{18} was also evaluated. Heavy metal contamination within soils considering concentration of seven heavy metal elements, i.e., Zn, Cd, As, Ni, Cu, Pb and Cr was evaluated in thirty two boreholes among which eleven boreholes were drilled in Laogang Town and twenty one in Hengsha Island. Considering that the contaminated layers including the reclaimed layer ①$_{3\text{-}2}$ to layer ⑤$_{1\text{-}1}$, and that layer ②$_2$–③$_1$ were dredged as the major fill source for land reclamation in the study regions, only the top layers downward to layer ⑤$_{1\text{-}1}$ were analyzed in the HMC analysis, and they were divided into three groups, i.e., the fill layers, shallow sediments (layer ②$_2$–③$_1$) and deep sediments (layer ④–⑤$_{1\text{-}1}$).

2.2 BN Analysis

BNs are acyclic directed graphs with nodes that represent random variables and arcs that represent direct probabilistic dependences among them [13, 14, 16]. BN's structure is a qualitative illustration of the interactions among the modeled variables. Based on the conditional independence and the chain rule [13], the joint probability distribution of a set of random variables U = {B1, …, Bn} can be defined in the BN as follows:

$$P(U) = \prod_{i=1}^{n} P\{B_i|Pa(B_i)\} \tag{1}$$

where P(U) is the joint probability distribution of U, and Pa(Bi) is the parent set of variable Bi. Information flow in BN obeys the following rules: (1) information may flow through a serial (B1→B2→B3)/diverging (B1←B2→B3) connection unless the evidence for the intermediate variable B2 is given; (2) information may flow through a converging (B1→B2←B3) connection whenever the state of the intermediate variable B2 or one of its descendants is given [19]. BN determines the relevant variables to a given target variable by using these rules. According to Bayes theorem, posterior probability distribution of events can be yielded by updating their prior probability given new observations, i.e. evidence E [20]:

$$P(U|E) = P(U,E) \Big/ P(E) = P(U,E) \Big/ \sum_{U} P(U,E) \tag{2}$$

In the BN analysis, subject-matter-expert judgments [16] from correlated sources, such as geotechnical field surveys, laboratory tests, data collection, and standards learning, were firstly executed to determine concrete BN variables and global BN

Table 2. Categorization of analyzed variables in the BN.

Node		Discretizetion		
No.	Name(unit)	Range	Levels	Level label
N01	Thickness of reclaimed layer (m)	[0,2]	Thin	TN
		(2,5]	Thick	TK
		(5,+∞)	Very Thick	VTK
N02	Predominant soil type of fill		Clayer soils	CS
			Silty-Sandy soils	SSS
N03	Land reclamation time (Year)	[0,6]	short term	ST
		(6,13]	Moderate term	MT
		(13,20]	Long term	LT
		(20,+∞)	Very long term	VLT
N04	Point location		Yangtze estuary island zone	YEIZ
			Yangtze estuary coastal plain zone	YECPZ
N05	Average initial void ratio of reclaimed layer	[0,0.6)	Extremely low void ratio	RELV
		[0.6,0.75)	Very low void ratio	RVLV
		[0.75,0.9)	Low void ratio	RLV
		[0.9,1.0)	High void ratio	RHV
		[1.0,1.5)	Very high void ratio	RVHV
N06	Average hydraulic conductivity of reclaimed layer (cm/s)	[0,1E-7)	Practically impermeable	RPI
		[1E-7,5E-7)	Very weak permeability	RVWP
		[5E-7,3E-6)	Weak permeability	RWP
		[3E-6,1E-5)	High permeability	RHP
		[1E-5,1E-4)	Very high permeability	RVHP
		[1E-4,+∞)	Extremely high permeability	REHP
N07	Average compressibility coefficient of reclaimed layer (Mpa^{-1})	(0,0.1]	Very low compressibility	RVLC
		(0.1,0.2]	Low compressibility	RLC
		(0.2,0.4]	Moderate compressibility	RMC
		(0.4,0.5]	High compressibility	RHC
		(0.5,+∞)	Very high compressibility	RVHC
N08	Average initial void ratio of underlying strata	[0,0.6)	Extremely low void ratio	UELV
		[0.6,0.75)	Very low void ratio	UVLV
		[0.75,0.9)	Low void ratio	ULV

Node		Discretizetion		
No.	Name(unit)	Range	Levels	Level label
N08	Average initial void ratio of underlying strata	[0.9,1.0)	High void ratio	UHV
		[1.0,1.5)	Very high void ratio	UVHV
N09	Average hydraulic conductivity of underlying strata (cm/s)	[0,1E-7)	Practically impermeable	UPI
		[1E-7,5E-7)	Very weak permeability	UVWP
		[5E-7,3E-6)	Weak permeability	UWP
		[3E-6,1E-5)	High permeability	UHP
		[1E-5,1E-4)	Very high permeability	UVHP
N10	Average compressibility coefficient of underlying strata (Mpa^{-1})	(0,0.1]	Very low compressibility	UVLC
		(0.1,0.2]	Low compressibility	ULC
		(0.2,0.4]	Moderate compressibility	UMC
		(0.4,0.5]	High compressibility	UHC
		(0.5,+∞)	Very high compressibility	UVHC
N11	s_{18} (mm)	[0,200)	No risk	SNR
		[200,1000)	Low Risk	SLR
		[1000,+∞)	High risk	SHR
N12	v_{18} (mm/six months)	[0,1.83]	No risk	VNR
		(1.83,2.00]	Low risk	VLR
		(2.00,10.00]	Moderate risk	VMR
		(10.00,+∞)	High risk	VHR
N13	R_c	SNR and VNR	No risk	TNR
		VMR but not SHR	Moderate risk	TMR
		SHR or VHR	High risk	THR
		Othe conditions	Low risk	TLR
N14	Fill HMC contamination		Yes	FY
			No	FN
N15	Shallow sediment HMC pollution		Yes	SY
			No	SN
N16	Deep sediment HMC pollution		Yes	DY
			No	DN

structure and to build a BN for risk evaluation in the study regions. The BN was constituted of sixteen nodes among which four nodes (N01-N04) are basic characteristics of the data points, six nodes (N05-N10) showing geotechnical parameters of analyzed layers, three nodes (N011-N013) used for evaluating subsidence risk levels, and three (N14-N16) used to assess soil contamination (Table 2). Secondly, data preparation based on the analysis in the first step was conducted carefully. The continuous variables were transformed into discrete variables to make BN a discrete network that can be applicable to discrete information in practice (Table 2). Variable discretization was conducted on the basis of subject matter experts' judgments. Thirdly, primary structure learning was performed with the prepared data and outlined global BN structure. The primarily learned BN structure was further adjusted to rectify the arcs among the BN nodes that go against common theories and logic. Finally, data learning was conducted using the established BN structure, thus producing the final BN model.

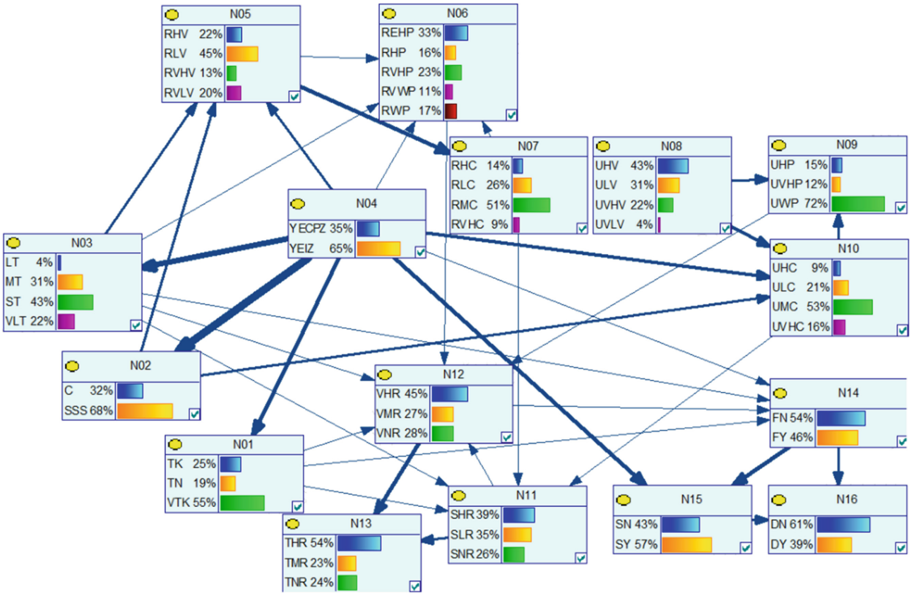

Fig. 1. Land subsidence risk and soil HMC contamination BN showing arc influence strength.

3 Results

Influence strength between the Node N04 with its connected nodes generally shows stronger than other arcs, indicating the point location holds robust correlation with other variables, especially the soil type, land reclamation time, s_{18} and shallow sediment HMC. Influence between the three geotechnical parameters of fill layers, average initial soil void ratio (N05), average hydraulic conductivity (N06) and average compressibility coefficient (N07), only appears strong between Node N05 and Node N07. Contrastingly, influence among the three geotechnical parameters all shows similarly strong. Influence among the three HMC nodes is significantly strong, especially that between the fill HMC (N14) and the shallow sediment HMC (N15) (Fig. 1).

Sensitivity analysis on both the land subsidence risk levels and soil contamination indicates the factors which can provide a high probability of correct evaluation on the target node (Figs. 2, 3, 4, 5 and 6). The s_{18} risk is remarkably sensitive to the four basic characteristics of the data points (N01-N04), the void ratio nodes and the compressibility coefficient nodes (Fig. 2a). The v_{18} risk is sensitive to all the basic characteristics of the data points and the geotechnical parameter nodes among which the point location, land reclamation time and the fill layer thickness hold the highest sensitivity (Fig. 2b). Sensitivity of the R_c risk reflects a combination of that of s_{18} risk and v_{18} risk (Fig. 3). The fill HMC shows sensitive to point location, land reclamation time and fill layer thickness (Fig. 4). The shallow sediment HMC acts sensitive to the fill HMC, but only slightly sensitive to point location (Fig. 5). The deep sediment HMC shows

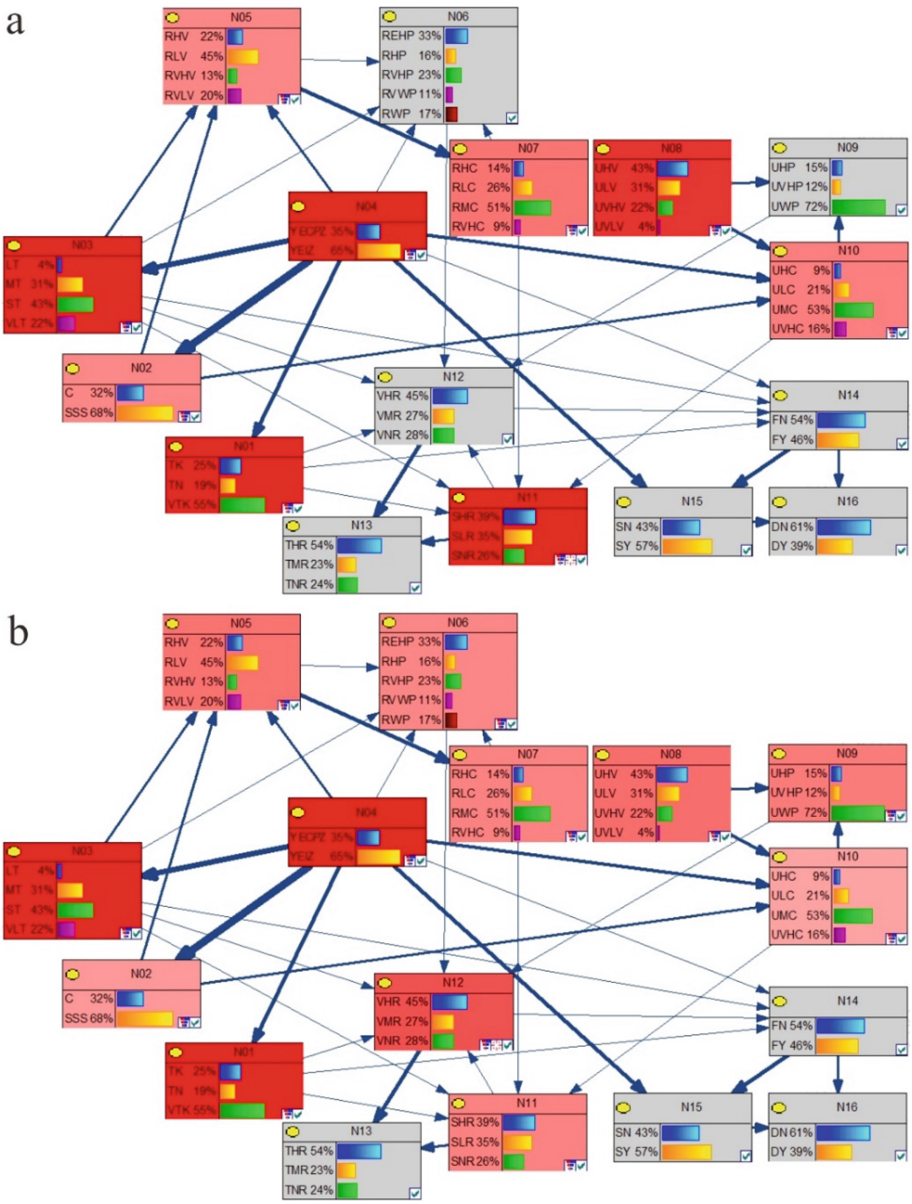

Fig. 2. Sensitivity analysis respectively setting N11 (a), N12 (b) as the target variable (Bar chart color: the more red indicating the sensitivity more high, gray showing no sensitivity).

sensitive to the other two HMC nodes, land reclamation time and point location (Fig. 6). Sensitivity between the three subsidence risk nodes and the three HMC nodes appears slight.

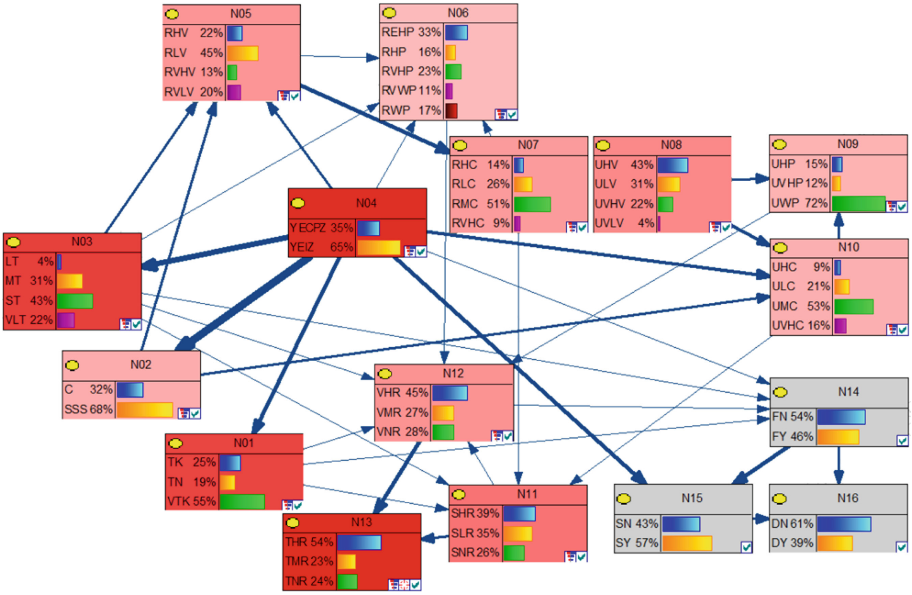

Fig. 3. Sensitivity analysis setting N13 as the target variable (Bar chart color: the more red indicating the sensitivity more high, gray showing no sensitivity).

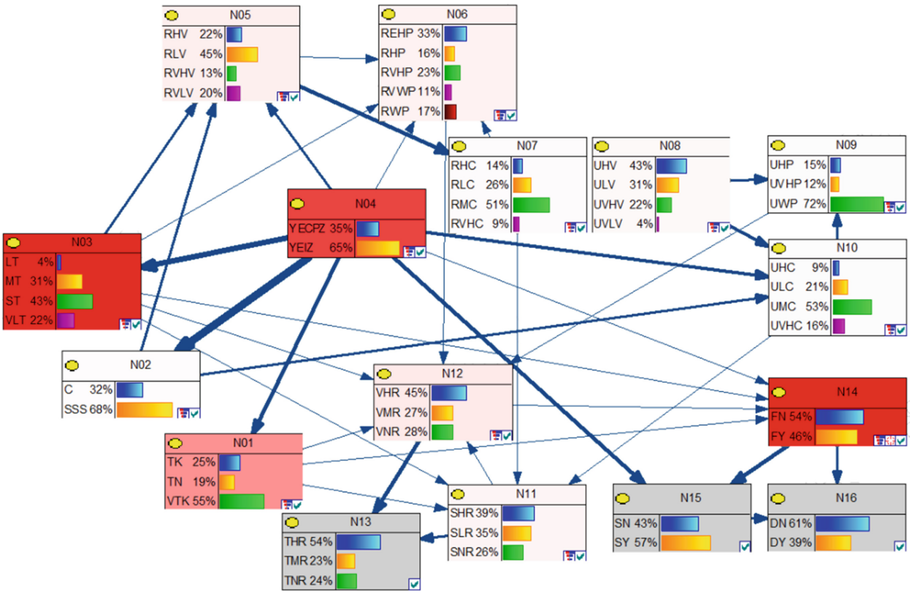

Fig. 4. Sensitivity analysis setting N14 as the target variable (Bar chart color: the more red indicating the sensitivity more high, gray showing no sensitivity).

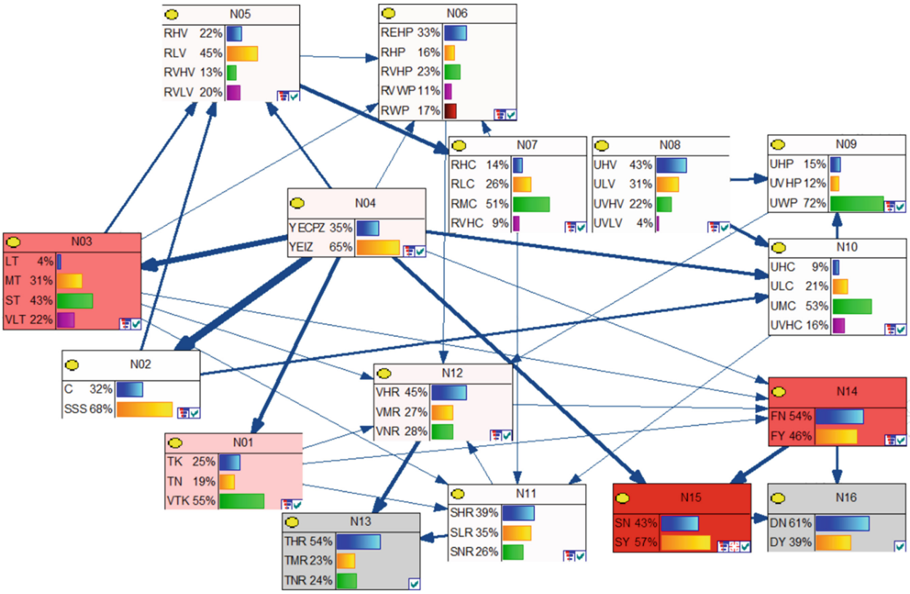

Fig. 5. Sensitivity analysis setting N15 as the target variable (Bar chart color: the more red indicating the sensitivity more high, gray showing no sensitivity).

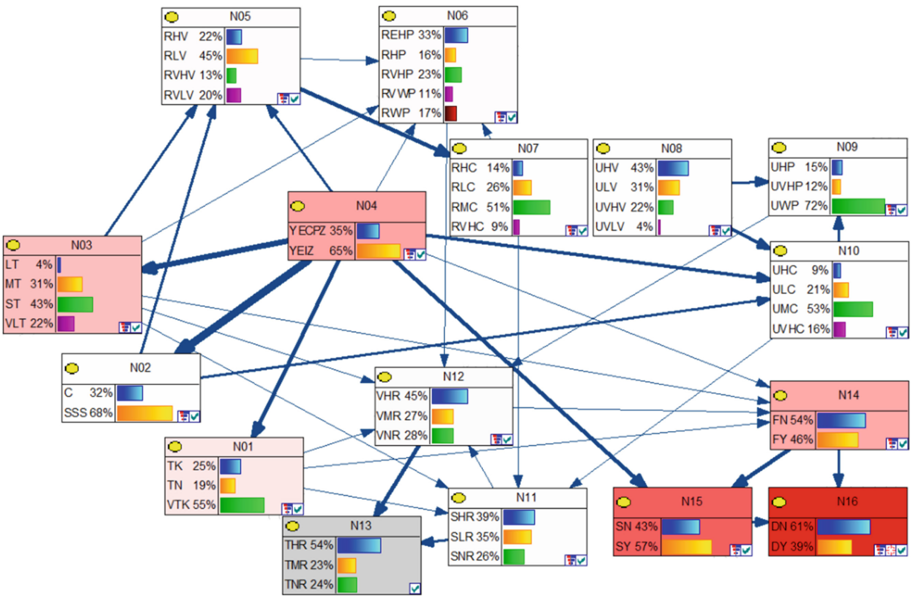

Fig. 6. Sensitivity analysis setting N16 as the target variable (Bar chart color: the more red indicating the sensitivity more high, gray showing no sensitivity).

4 Discussion and Conclusions

Point location holds strong influence on other three basic characteristics, geotechnical parameters and shallow sediment HMC, meanwhile it shows high sensitivity to the three subsidence risk nodes, fill HMC and deep sediment HMC, this essentially reflects significant differences in both land reclamation construction and sedimentary environments between the two study zones, i.e., YEIZ and YECPZ. The different influence strength of arcs among geotechnical parameters of the fill layers and the underlying strata, probably was caused by different texture between the fill soils and sedimentary soils. Rationally, the three land subsidence risk nodes appear sensitive to the six geotechnical parameter nodes since the land subsidence was evaluated using those parameters. But the three HMC nodes develop low sensitivity to the six geotechnical parameters, s_{18} risk and v_{18} risk, indicating that the soil HMC in the study regions hold no significant correlation with the soil geotechnical parameters. The HMC distribution was controlled more by the basic characteristics such as the point location, land reclamation time and fill layer thickness. Noticeably, the three HMC nodes show almost no sensitivity to fill soil types, meaning the HMC distribution held no correlation with the fill soil types in the study regions. The BN can be used as a basis for land subsidence controlling and field remediation planning in the study regions, further field survey data on soil geotechnical properties and HMC distribution would enhance reliability of the BN.

Acknowledgments. This work is supported by China's National Key R&D Program (2017YFC0806000), China's National Key Basic Research Program (2014CB046901); Shanghai Pujiang Program (15PJD039); Shanghai Municipal Science and technology project (18DZ1201301); Opening fund of State Key Laboratory of Geohazard Prevention and Geoenvironment Protection (SKLGP2018K019); Fund of Shanghai Institute of Geological Survey.

References

1. UN Altas, Coasts and coral reefs. http://www.oceansatlas.org/subtopic/en/c/304/. Accessed 12 Apr 2019
2. Wikipedia, Land reclamation. https://en.wikipedia.org/wiki/Land_reclamation. Accessed 12 Apr 2019
3. Wang, J.X., Wu, L.B., Deng, Y.S., Song, D.S., Liu, W.J., Hu, M.Z., Liu, X.T., Zhou, J.: Investigation and evaluation of contamination in dredged reclaimed land in China. Mar. Georesour. Geotechnol. **36**(5), 603–616 (2018)
4. Ho, K.T., Burgess, R.M., Pelletier, M.C., Serbst, J.R., Ryba, S.A., Cantwell, M.G.: An overview of toxicant identification in sediments and dredged materials. Mar. Pollut. Bull. **44** (4), 286–293 (2002)
5. Caeiro, S., Costa, M.H., Ramos, T.B., Fernandes, F., Silveira, N., Coimbra, A.: Assessing heavy metal contamination in sado estuary sediment: an index analysis approach. Ecol. Indic. **5**(2), 151–169 (2005)
6. EPA (United States Environmental Protection Agency): Assessment of Contaminated Sediments in the Kinnickinnic River Mooring Basin in the Milwaukee Estuary Area of

Concern, Milwaukee, Wisconsin (2016). https://dnr.wi.gov/topic/greatlakes/documents/TurningBasinSiteCharacterizationReport.pdf. Accessed 12 Apr 2019
7. Mukesh, M.V., Chandrasekaran, A., Premkumar, R., Keerthi, B.N.: Metal enrichment and contamination in river and estuary sediments of Tamirabarani, South India. J. Earth Sci. Cli. Change **9**(10), 497–503 (2018)
8. Wang, W., Liu, H., Li, Y., Su, J.: Development and management of land reclamation in China. Ocean Coast Manag. **102**, 415–425 (2014)
9. Liu, L., Xu, S.Y., Chen, Z.L., Yu, J.: Spatial distribution and environmental quality assessment on heavy metals in tidal flat sediments of Shanghai coastal zone. Shanghai Geo **1**, 1–5 (2000)
10. Fang, M., Wu, Y.J., Liu, H., Jia, Y., Zhang, Y., Wang, X.T., Wu, M.H., Zhang, C.L.: Distribution, sources and ecological risk assessment of heavy metals in sediments of the Yangtze River estuary. Acta Sci. Circumstant. **33**(2), 563–569 (2013)
11. Li, L., Ping, X.Y., Wang, Y.L., Jiang, M., Yuan, Q., Shen, X.Q.: Study on the spatial and temporal distribution of the heavy metals in the surface sediments based on multivariate statistical technique from the Changjiang Estuary and its adjacent areas. Environ. Chem. **32**(3), 438–445 (2013)
12. Zhang, W., Feng, H., Chang, J., Qu, J., Xie, H., Yu, L.: Heavy metal contamination in surface sediments of Yangtze River intertidal zone: an assessment from different indexes. Environ. Pollut. **157**(5), 1533–1543 (2009)
13. Pearl, J.: Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann Publishers Inc., San Francisco (1988)
14. Haddawy, P., Hasan, A.H.M.I., Kasantikul, R., Lawpoolsri, S., Sa-Angchai, P., Kaewkungwal, J., Singhasivanonb, P.: Spatiotemporal Bayesian Networks for malaria prediction. Artif. Intell. Med. **84**, 127–138 (2018)
15. Wang, J.X., Gu, X., Huang, T.: Using Bayesian Networks in analyzing powerful earthquake disaster chains. Nat. Hazard **68**(2), 509–527 (2013)
16. Baksh, A.A., Abbassi, R., Garaniya, V., Khan, F.: Marine transportation risk assessment using Bayesian Network: application to Arctic waters. Ocean Eng. **159**, 422–436 (2018)
17. Plomaritis, T.A., Costas, S., Ferreira, O.: Use of a Bayesian Network for coastal hazards, impact and disaster risk reduction assessment at a coastal barrier (Ria Formosa, Portugal). Coast. Eng. **134**, 134–147 (2018)
18. Terzaghi, K.: Erdbaumechanik auf Bodenphysikalisher Grundlage. Deuticke, Vienna (1925)
19. Sagrado, J.D., Rodríguez, F., Berenguel, M., Mena, R.: Bayesian Networks for greenhouse temperature control. J. Appl. Logic **17**(C), 25–35 (2016)
20. Bayes, R.T.: An essay toward solving a problem in the doctrine of chances (reprint). Biometrica **45**(3–4), 293–315 (1702–1761)