

基于贝叶斯网络的河流突发性水质污染事故风险评估

孙鹏程, 陈吉宁

(清华大学环境科学与工程系, 北京 100084)

摘要: 事故造成的水质突发性风险评估对于河流水质安全管理具有重要意义. 通过贝叶斯网络直观地表示事故风险源和河流水质之间的相关性, 并用时序蒙特卡罗算法将风险源状态模拟、水质模拟和贝叶斯网络推理过程结合, 可以对多个风险源共同影响下的河流突发性水质污染事故的超标风险进行量化评估. 案例研究表明, 多风险源对同一受纳水体的水质突发性污染事故风险的耦合影响十分显著, 在进行流域水安全管理时必须对多风险源进行综合管理. 同时, 模型的诊断推理功能可为流域关键风险源识别和管理提供决策支持.

关键词: 风险评估; 贝叶斯网络; 时序蒙特卡罗模拟; 水质模型

中图分类号: X820.4 文献标识码: A 文章编号: 0250-3301(2009)01-0047-05

Risk Assessment of River Water Quality Under Accidental Pollution Based on Bayesian Networks

SUN Peng-cheng, CHEN Ji-ning

(Department of Environmental Science and Engineering, Tsinghua University, Beijing 100084, China)

Abstract: Risk assessment for accidental pollution plays an important role in river water quality management. Bayesian networks can be applied to represent the relationship between pollution sources and river water quality intuitively. A time sequential Monte Carlo algorithm, integrated with pollution sources model, water quality model and Bayesian reasoning, is developed to quantify river water quality risk under the collective influence of multiple pollution sources. A case study shows that multiple pollution sources have obvious effect on water quality risk of the receiving water body, which means that integrated management should be developed for multiple risk sources. The model could also support the decision-making process of river basin management through identification of critical pollution sources.

Key words: risk assessment; Bayesian networks; time sequential Monte Carlo; water quality model

近年来, 以沱江氨氮排放事件、松花江苯类物质泄漏事件、珠江北江含镉污水排放事件等为代表的一系列重大水环境突发性污染事件频繁发生, 据国家环境保护总局统计, 2005 年接到的水污染突发事件报告为 41 起, 2006 年接报处理水污染突发事件增加到 95 起. 水污染事故发生突然, 发生后危害大且难以控制, 因此在事故发生前对河流水质进行风险评估, 并针对性地采取预防管理措施是减小事故危害的重要手段.

目前, 国内外对于突发性水质污染事故的风险评估研究比较少, 已有研究多集中在单个风险源的风险评估及事故泄漏后污染物的行为模拟和危害后果分析方面^[1~5]. 我国的河流系统存在着多风险源区域性集中分布的特征, 而对于多风险源河流系统的综合风险评价目前尚缺乏足够的研究. 本文从保证河流水质安全角度出发, 以贝叶斯网络 (Bayesian networks) 技术表达河流系统中各要素的不确定性及各要素之间的相关性, 集成风险源状态模型和水质模型后应用时序蒙特卡罗随机模拟技术进行贝叶斯网络推理, 实现多风险源河流系统突发性水质污染

事故风险的量化评估.

1 贝叶斯网络及贝叶斯网络推理概述

贝叶斯网络是以图形化的方式表达不确定性信息的概率模型, 能够直观地表达系统内各元素之间的相互影响关系. 一个贝叶斯网络是一个有向无环图, 节点表示随机变量, 节点之间的有向弧表示相互关系, 每个节点均具有其自身的条件概率表 (conditional probability table, CPT)^[6,7].

在贝叶斯网络模型中, 当部分变量 (通常是输入变量) 的概率确定后, 就能够通过基础的概率运算和贝叶斯理论计算所有节点或者部分节点的概率, 实现贝叶斯网络推理. 现有的贝叶斯网络推理算法可分为精确推理算法和近似推理算法两大类^[8,9]. 精确推理算法基于贝叶斯概率公式进行推理, 适用于结构简单、网络规模小的贝叶斯网络; 随机模拟方法 (stochastic sampling) 为主的近似推理算法在不改变

收稿日期: 2008-01-14; 修订日期: 2008-04-15

作者简介: 孙鹏程 (1981 ~), 男, 博士生, 主要研究方向为河流水质模拟, E-mail: spc03@mails.tsinghua.edu.cn

计算结果正确性的前提下,降低了计算精度,从而简化计算复杂性,在网络结构复杂、规模较大的贝叶斯网络中应用较多.

贝叶斯网络(BNs)的结构和推理过程具有适合于风险评价建模和分析的特点^[10],BNs 结构可以表达变量间的不确定性关系和多态性特征,BNs 的前向推理(因果推理)可以计算各种故障条件下系统发生风险的联合概率,进行风险评价;BNs 的诊断推理可以计算系统发生风险时各个组件状态的条件概率,进行系统诊断,开展针对性的风险管理^[11].

在河流水质风险评价和管理工作中应用贝叶斯网络,可以方便地表征多变量之间的相互影响关系,处理各种不同因素不确定性的量化问题,而贝叶斯网络的推理过程也易于同水质模型集成以评价水质风险和进行风险源诊断,在水质安全管理中有着广泛的应用前景.

贝叶斯网络构建和推理的一般步骤为^[12,13]:① 识别影响和描述单元状态的变量;② 识别变量之间的关系;③ 将各变量状态离散化并填充 CPT 表,变量较多时,将变量分类分组表示;④ 选择合适的算法进行推理,求出各种概率值.

2 基于贝叶斯网络的河流突发性水质污染事故风险评估模型

2.1 河流系统突发性水质污染事故的贝叶斯网络模型

构建河流系统突发性水质污染事故的贝叶斯网络模型,首先通过河流水质评价确定需要进行风险管理的水质指标(也可根据管理的要求进行确定),然后对流域内污染物的风险来源进行排查,根据风险源与水体及风险源之间的相互影响关系确定贝叶斯网络结构.

图 1 所示为南水北调东线工程中的梁济运河-东平湖段.该河段规划了 8 座污水处理厂以实现対入河污染物量的削减,考虑影响断面水质的主要因素为 8 个污水处理厂(风险源)、上游水源以及水体对污染物的降解能力以及水文状况,构建该河段的贝叶斯网络见图 2.

确定网络结构后,应该进一步确定各个节点变量的条件概率表(CPT).每个节点的条件概率表中需包括该节点所有父节点不同状况组合情况下该节点状态的条件概率.河流系统中,断面水质往往受多个风险源共同影响.以图 1 所示系统为例, O_3 断面水质受 8 个风险源影响,即使每个风险源只有 2 种

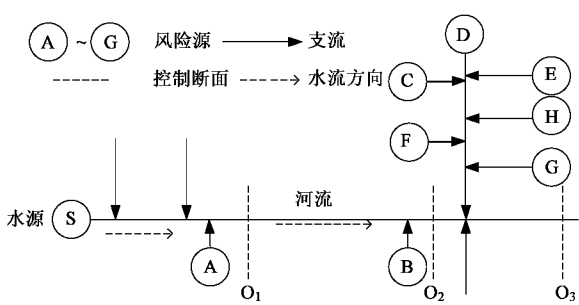


图 1 示例河流(梁济运河-东平湖)
Fig.1 Sketch map of a river system(Liangji canal-Dongping Lake)

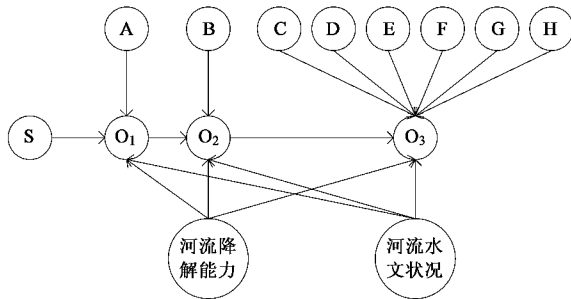


图 2 河流突发性水质污染事故风险贝叶斯网络
Fig.2 Bayesian networks for accidental pollution risk of river water quality

状态变化,其静态组合个数也达到 $2^8 = 256$ 种,而各个风险源状态变化的动态特征又会使得它们之间的组合方式更为复杂,比如 2 个风险源之间既可能同时发生事故,也可能在一个事故开始一段时间后另一个发生事故等.因此,河流系统的复杂性、河流水质污染事故的动态性使得 CPT 的确定非常困难,而直接应用 CPT 进行精确推理也是很难实现的.

条件概率表(CPT)的本质是描述节点之间的相互关系,而河流水质模型的基本功能是描述风险源和水体水质之间的响应关系.因此,本研究应用贝叶斯网络表达河流系统中风险源与水体水质之间的关系后,通过随机采样获得各风险源排放行为的时间序列,利用水质模型进行仿真模拟,借助近似算法实现贝叶斯网络的各种推理功能,以解决贝叶斯网络在河流系统水质安全管理中应用时确定 CPT 的困难,充分利用贝叶斯网络在风险评价和管理方面的优势.

2.2 贝叶斯网络推理的时序蒙特卡罗算法

以概率基本公式为基础的精确推理算法很难适用于突发性水质污染事故的风险评估,本研究借鉴电力系统可靠性评估的时序蒙特卡罗算法^[14,15],构建河流突发性水质污染事故风险的时序蒙特卡罗模

拟算法,进行贝叶斯网络近似推理,算法包括风险源状态模拟和河流水质蒙特卡罗模拟两部分。

2.2.1 风险源状态模型

将风险源的运行过程离散为不同的状态,如正常运行状态和故障状态,确定各状态发生概率、持续时间及排放强度。

(1)状态发生概率 风险源运行时在不同状态之间进行转化,其转化过程可以用马尔可夫过程来表示^[16,17],其特点是从一个状态到另一个状态转移的概率仅仅取决于转移前后的 2 个状态,转移的快慢或强度由转移率来表示。

以 2 个状态模型为例,假设风险源有正常运行和发生事故 2 种状态,处于正常运行状态的平均时间为 MTTF,风险源处于故障状态的平均时间为 MTTR,则风险源由正常运行状态向故障状态转移的转移率为:

$$\lambda = \frac{1}{\text{MTTF}} \quad (1)$$

由故障状态向正常运行状态转移的转移率为:

$$\mu = \frac{1}{\text{MTTR}} \quad (2)$$

通过计算可以获得系统运行稳定后处于正常和故障状态的概率:

$$P_1 = \frac{\mu}{\lambda + \mu} \quad (3)$$

$$P_2 = \frac{\lambda}{\lambda + \mu} \quad (4)$$

(2)事故状态持续时间 对于不同的风险源,事故状态持续时间 T_{ik} 的确定方式也不尽相同。在很短时间内发生的泄漏事故,可以按照瞬时排放进行计算;对某些确定的事故情景,持续时间按照确定的时间进行计算;其余情景的持续时间可以将 T_{ik} 处理为服从指数分布或者威布尔分布的变量进行计算。

当 T_{ik} 服从指数分布时,用逆变换法得到:

$$T_{ik} = -T_{\text{MTTR}} \ln r \quad (5)$$

式中, r 为(0, 1)之间均匀分布的随机数, T_{MTTR} 为平均事故状态时间, T_{ik} 也可用威布尔分布等进行处理。

(3)事故排放强度 发生事故后的污染物排放强度往往与风险源的规模、安全设施条件、地理特征、人口经济等自然因素和社会因素有关,其表达方式可以分为污染物排放量、流量、浓度等指标。对于瞬时排放的风险源,可以用污染物排放总量来表征排放强度;对于具有一定的时间持续特征的风险源可以用流量和浓度来表征其排放强度。在风险评估中,一般选

择可能发生的危害最大情景来确定排放强度。

2.2.2 时序蒙特卡罗模拟算法

突发性风险评估需要考虑风险源的时变特征,将电力系统可靠性评估中的时序蒙特卡罗算法(Time Sequential Monte Carlo Simulation Technique)和河流水质模型进行集成,应用到河流突发性水质污染事故风险评估之中,算法计算步骤如下:

① $t = 0$ 时,设系统从正常运行状态开始,即此时所有风险源均处在正常运行状态,无突发性污染物排放;

② 为系统中的每个风险源产生一个随机数(本研究中利用 fortran90 自带的随机数函数产生随机数),通过已经确定的风险源状态模型中风险源的状态概率确定下一时刻风险源的状态,记录所有处于故障状态的风险源;

③ 为处于故障状态的风险源产生随机数(产生方式同步骤②),根据已建立的风险源状态模型确定风险源的故障状态持续时间,并确定相应的污染物排放强度;

④ 根据风险源排放量确定河流水质模型的输入参数;

⑤ 利用水质模型模拟河流水质变化,记录河流随风险源排放的水质变化,记录重要观测断面的超标时间和污染物浓度值;

⑥ 计算河流水质超标风险概率: $f = \text{超标时间} / \text{模拟时间}$;

⑦ 重复②~⑥,直到规定的模拟时长或者 f 值收敛。

当模拟时间足够长时,通过水质模拟结果可以获得河流系统水质的时空概率分布,进一步将水质模拟结果同已经确定的水质风险评价标准作比较,即可以得到突发性事件造成的河流系统的水质超标概率,同时风险源状态变化的模拟样本集和河流水质变化的模拟结果样本集也能为进行其他贝叶斯网络推理提供基本信息。

3 案例计算

以图 1 中的河段为例进行突发性水质风险评估,评估的水质指标为高锰酸盐指数和氨氮,系统的贝叶斯网络模型见图 2。8 个风险源为规划建设的污水处理厂,用两状态模型模拟污水处理厂的状态变化,污水处理厂故障状态时排放强度按照污水处理厂来水水质和设计处理能力确定,故障时间假设服从指数分布。经过对一些污水处理厂的调研表明,

污水处理厂在建成后运行的一段时间内事故发生率较高,运行稳定后逐渐好转,事故发生率降低,参考调研结果本研究中事故发生时间比例取 3%. 计算中,上游来水满足地表水Ⅲ类水体标准,风险评估标准为地表水Ⅲ类标准. 风险源各项参数见表 1.

表 1 风险源计算参数
Table 1 Parameters of risk sources

风险源	MTR MTTF + MTR /%	T _{MTR} /h	事故排放强度/g·s ⁻¹	
			高锰酸盐指数	氨氮
A	3	24	46.3	13.9
B	3	24	46.3	13.9
C	3	24	81.0	25.7
D	3	24	67.3	17.6
E	3	24	92.6	27.8
F	3	24	92.6	27.8
G	3	24	93.7	33.0
H	3	24	57.9	17.4

选择水质模型对河流水质进行时序蒙特卡罗模拟,模拟时长为 50 万 h,并进行 10 次平行计算实验测试随机算法的采样随机性和计算结果的收敛性.

表 2 各断面高锰酸盐指数和氨氮超标概率评估结果

Table 2 Risk assessment results of permanganate index and NH₄⁺-N at different sections

实验序号	O ₁		O ₂		O ₃	
	高锰酸盐指数	氨氮	高锰酸盐指数	氨氮	高锰酸盐指数	氨氮
1	0.026 8	0.027 8	0.000 373	0.048 8	0.175	0.216
2	0.029 5	0.030 6	0.000 173	0.051 6	0.171	0.210
3	0.031 4	0.032 5	0.000 736	0.051 5	0.179	0.221
4	0.030 6	0.031 7	0.000 564	0.052 0	0.176	0.217
5	0.028 6	0.029 7	0.000 358	0.051 2	0.179	0.222
6	0.027 5	0.028 6	0.000 540	0.047 1	0.176	0.216
7	0.026 7	0.027 7	0.000 610	0.047 0	0.171	0.210
8	0.026 2	0.027 2	0.000 461	0.048 2	0.176	0.215
9	0.024 6	0.025 7	0.000 622	0.050 5	0.177	0.216
10	0.029 2	0.030 3	0.000 704	0.051 3	0.188	0.224

抵达 O₁ 断面前只受到风险源 A 的影响,因此 O₁ 断面处水质超标风险概率同风险源 A 的事故概率基本一致,约为 3%;O₂ 断面超标风险概率较小是由于

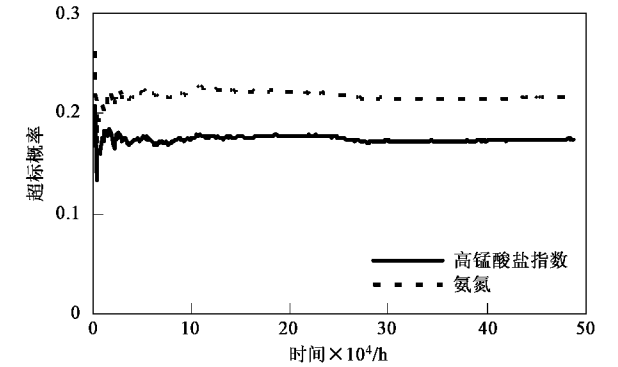


图 3 O₃ 断面水质风险评估

Fig.3 Risk assessment of permanganate index and NH₄⁺-N at section O₃

4 结果与讨论

表 2 给出了 3 个控制断面的超标风险概率计算结果. O₁ 断面高锰酸盐指数超标概率的均值为 0.028 1,标准差为 0.002 14,氨氮超标概率的均值为 0.029 2,标准差为 0.002 15;O₂ 断面高锰酸盐指数超标概率的均值为 0.000 5,标准差为 0.000 17,氨氮超标概率的均值为 0.049 9,标准差为 0.001 97;O₃ 断面高锰酸盐指数超标概率的均值为 0.176 9,标准差为 0.004 86,氨氮超标概率的均值为 0.217,标准差为 0.004 58. 10 次平行模拟实验的计算结果基本一致,证明所构建的算法数值上是稳定的,并且随着断面超标概率的增大精度也有所提高.图 3 是 O₃ 断面超标风险概率计算过程的时间曲线,模拟初始时概率计算值波动较大,大约 30 万 h 后逐渐平缓,这一结果能为确定合理的模拟时间长度提供支持.

评估结果表明,满足地表水Ⅲ类标准的源水在

河流在抵达 O₂ 断面前经过了一段较长的自净过程,水质好转从而使得水体对事故污染负荷的稀释能力增强;而在受多个风险源影响的 O₃ 断面,虽然单个风险源仅有 3%的时间发生事故,其高锰酸盐指数和氨氮的超标风险概率均远远高于 3%. 可见,断面的水质风险受风险源、源水、河流自净能力等多个因素的影响,且多个风险源对同一受纳水体水质风险的耦合影响十分显著. 因此,在进行河流水质风险管理过程中,必须对所有的风险源进行综合考虑,才能真正降低系统水质风险.

应用贝叶斯网络的诊断推理功能,可以识别各个风险源对系统风险的贡献大小,为风险源管理提供依据. 诊断推理通过计算超标事件发生后风险源状态的条件概率判断风险源对系统风险的贡献. 设

超标事故次数为 N , $risk$ 表示河流系统发生水质超标事故, x_i 为风险源状态, 且 $x_i = 1$ 为事故状态, 系统发生超标事故时风险源 i 处于事故状态的次数为 M , 则进行诊断推理的条件概率为:

$$P(x_i = 1 | risk) = \frac{M}{N} \tag{6}$$

图 4 给出了控制断面 O_3 处的高锰酸盐指数超标风险的诊断推理结果. 图中显示, 模拟过程中, O_3 断面高锰酸盐指数超过 6 mg/L 的事件, 各风险源贡献大小相近, 其中距离 O_3 断面最近的污水处理厂 C、D、E、F、G、H 贡献较大; 超过 7 mg/L 的事件中, E、F、G 3 个污水处理厂贡献较大; 超过 8 mg/L 的事件中, 污水处理厂 E、F、G、D、C 贡献较大; 超过 9 mg/L 的事件中, 污水处理厂 G、C 贡献最大, 污水处理厂 A 的贡献超过污水处理厂 H. 由于较高浓度的超标事件多为不同风险源发生事故后叠加造成的, 其各风险源的条件概率值相加可能 > 1 . 污染物超标程度直接影响事故的环境危害后果, 管理中应该综合考虑各风险源对突发性水质污染事故发生概率和危害程度的影响, 制订管理策略.

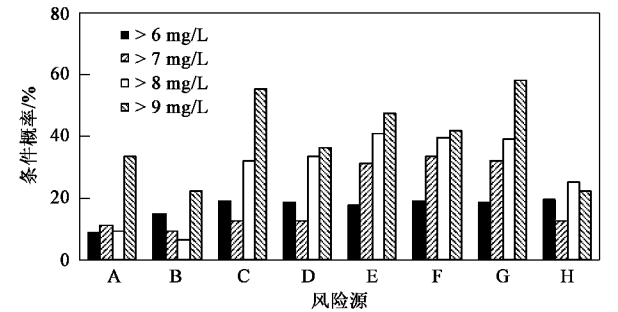


图 4 高锰酸盐指数贝叶斯网络诊断推理(O_3 断面)

Fig.4 Bottom-up reasoning result of Bayesian networks for permanganate index management at section O_3

5 结论

(1)构建了基于贝叶斯网络的多风险源河流突发性水质污染事故的风险评估模型. 对梁济运河段受 8 个污水处理厂影响下的河流系统水质污染事故进行了评价, 结果表明单个污水处理厂事故时间比例为 3% 时, 8 个污水处理厂共同影响下的断面高锰酸盐指数和氨氮突发性污染事故概率分别达到了 17% 和 21%, 可见多风险源对单一受纳水体的突发性水质污染事故风险具有明显的耦合影响, 在河流水质安全管理中应综合考虑.

(2)利用贝叶斯网络的逆向诊断推理, 可以识别

多风险源系统中的关键风险源, 为风险管理提供决策支持, 案例研究中识别了梁济运河段 8 个污水处理厂对控制断面不同超标程度的突发性水质污染事故风险影响的大小, 可以帮助确定各污水处理厂在风险管理中的排序, 为进一步的管理策略制订提供帮助.

参考文献:

[1] Norreys R, Cluckie I. Real time assessment of transient spills (rats) [J]. Water Science and Technology, 1996, 33(2): 187-198.

[2] Schimmer M, Butler B J. Transport behaviour and natural attenuation of organic contaminants at spill sites[J]. Toxicology, 2004, 205(3): 173-179.

[3] 刘国东, 宋国平, 丁晶. 高速公路交通污染事故对河流水质影响的风险评价方法探讨[J]. 环境科学学报, 1999, 19(5): 572-575.

[4] 任红娟. 公路毒性化学品运输事故的的概率计算及环境风险评价方法[J]. 苏州大学学报(工科版), 2006, 26(6): 70-71.

[5] 曾光明, 卓利, 钟政林, 等. 突发性水环境风险评价模型事故泄漏行为的模拟分析[J]. 中国环境科学, 1998, 18(5): 403-406.

[6] Uusitalo L. Advantages and challenges of Bayesian networks in environmental modelling[J]. Ecological Modelling, 2007, 203(3-4): 312-318.

[7] Castelletti A, Soncini-Sessa R. Bayesian networks in water resource modelling and management[J]. Environmental Modelling & Software, 2007, 22(8): 1073-1074.

[8] 黄友平. 贝叶斯网络研究[D]. 北京: 中国科学院计算研究所, 2005.

[9] 霍利民. 基于贝叶斯网络的电力系统可靠性评估[D]. 北京: 华北电力大学, 2004.

[10] Nathan O S, Dana L K. Bayesian parameter estimation in probabilistic risk assessment[J]. Reliability Engineering & System Safety, 1998, 62(1-2): 89-116.

[11] Castelletti A, Soncini-Sessa R. Bayesian Networks and participatory modelling in water resource management [J]. Environmental Modelling & Software, 2007, 22(8): 1075-1088.

[12] Santa Olalla F M, Dominguez A, Ortega F, et al. Bayesian networks in planning a large aquifer in Eastern Mancha, Spain [J]. Environmental Modelling & Software, 2007, 22(8): 1089-1100.

[13] Henriksen H J, Rasmussen P, Brandt G, et al. Public participation modelling using Bayesian networks in management of groundwater contamination[J]. Environmental Modelling & Software, 2007, 22(8): 1101-1113.

[14] Billinton R, Wang P. Distribution system reliability cost/worth analysis using analytical and sequential simulation techniques[J]. IEEE Transactions on Power Systems, 1998, 13(4): 1245-1250.

[15] Wang P, Billinton R. Time-sequential simulation technique for rural distribution system reliability cost/worth evaluation including wind generation as alternative supply[J]. IEEE Proceedings-Generation Transmission and Distribution, 2001, 148(4): 355-360.

[16] 田思进, 孟安. 状态空间法在安全工程中的应用[J]. 工业安全与防尘, 1994, (11): 3-9.

[17] 许开立, 陈宝智. 易燃易爆重大危险源事故概率计算[J]. 中国安全科学学报, 1997, 7(5): 1-5.