

基于贝叶斯网络的朝鲜海峡大气能见度推理模型及实验

单雨龙, 张韧, 毛科峰

(国防科技大学气象海洋学院, 江苏 南京 211101)

摘 要: 针对当前海上贸易航道通航风险评估工作中存在的能见度数据缺失等问题, 提出基于贝叶斯网络的能见度数据推理技术。通过研究海域的确定、节点因子的选取、样本数据集的生成、推理模型的构建及参数学习和推理计算等流程, 构建了基于贝叶斯网络技术的能见度数据推理模型, 并以朝鲜海峡海域为例展开试验分析。结果表明: 能见度具有年变化和年际变化特征规律, 利用多年某月的数据作为训练样本推理该月的能见度等级具有较高的准确性, 且相同样本形式下样本数据数量与推理结果准确性呈正相关。

关键词: 贝叶斯网络; 能见度; 朝鲜海峡

中图分类号: P427.2 **文献标识码:** A **文章编号:** 1003-0239(2019)01-0086-11

1 引言

由于中国九成以上的对外贸易通过海运完成, 故开展海上贸易航线的通航风险评估工作极为重要^[1]。尽管海上交通在便利性和经济性等方面相对陆上和空中交通均具有较大优势^[2], 但其通航风险充满较大的不确定性, 故开展海上航道的通航风险评估工作意义重大。由于能见度是航道通航风险的重要构成因子, 故获取航道所经海域高精度的能见度数据是客观评估航道通航风险的前提。当前大多数专家学者在对航道通航风险的评估工作中所用能见度数据为各个海域气象站点所记录数据, 且通常通过插值技术获取所研究海域的能见度数据。由于部分海区气象站点分布稀疏, 数据量少, 基于插值技术得到的能见度数据较为粗糙, 且缺乏科学性, 故通过其他技术获取海上高精度的能见度数据尤为重要。

针对大气能见度的推理计算, 前人已做如下研究: 马浩等^[3]利用相对湿度和PM 2.5 浓度对能见度进行非线性拟合, 并以杭州、宁波、温州三市为例对模型进行检验, 经检验三市的拟合值与观测值相关

系数均达0.75以上; 李学彬等^[4]利用Mile 散射理论分析了相对湿度对能见度的影响, 并根据实际测量的气溶胶谱分布和相对湿度计算了大气能见度; 傅刚^[5]介绍了3种基于液态水含量的大气消光系数经验计算公式; Gultepe等^[6]提出了一种新的计算雾的能见度的微物理参数化方法, 将能见度作为液态水含量和云滴粒子数密度两者的函数。以上学者所使用的计算能见度的方法主要有非线性拟合、基于消光系数控制方程的数值计算和基于经验公式的数值计算。其中方法一中拟合函数的选取具有较大的主观性; 方法二在计算过程中需获得云雾滴谱和云滴粒子数密度, 但这两种数据不易得到可靠的观测或数值计算资料^[7-8]; 方法三在计算过程中是以无降水为假设前提, 普适性偏低, 且经验公式具有较大不稳定性;

大气能见度的重要影响因子有大气消光系数、气溶胶的化学成分、气象因子等, 其中气象因子对大气能见度的影响机制主要通过大气消光系数和气溶胶的化学成分来完成^[9]。由于气象数据具有易获取性, 故通过气象数据推理能见度不仅保证了推理结果的科学性, 同时保证了所用数据的可靠性。

收稿日期: 2018-06-04; 修回日期: 018-08-13。

作者简介: 单雨龙(1994-), 男, 硕士在读, 主要从事军事海洋环境保障与风险评估研究。E-mail: 1151613926@qq.com

通讯作者: 张韧(1963-), 男, 教授, 博士, 主要从事军事海气相互作用及海洋灾害风险评估研究。E-mail: zren63@126.com

贝叶斯网络技术以概率论和图论知识为基础,不仅具有强大的概率推理能力,同时可将概率推理以可视化的方式进行表达,故贝叶斯网络技术非常适用于能见度的推理工作。

2 贝叶斯网络理论概述

贝叶斯网络(Bayesian Network, BN)又称信念网络、概率网络、因果网络,是综合运用概率论和图论知识进行概率推理的工具。贝叶斯网络由网络结构 G 和网络参数 Θ 两部分组成,其完整的数学表达形式为 $B=(G, \Theta)$, 其中 $G=(V(G), E(G))$, 表示一有向环图, $V(G)$ 是节点集合,节点表示所研究问题域中的变量, $E(G)$ 为有向弧集合,有向弧用于表示节点之间的概率依赖关系。网络参数 Θ 表示每个网络节点的条件概率分布,定量表达了节点之间的概率依赖程度^[10]。

贝叶斯网络的核心理论基础为贝叶斯公式,用于完成贝叶斯网络中概率推理部分。BN网络结构中蕴含条件独立性假设,即若给定根节点的先验概率分布和子节点的条件概率分布,则可依据链式法则推理得到包含所有子节点的联合概率分布,推理公式如下:

$$P(V_1, V_2, \dots, V_k) = \prod_{i=1}^k P(V_i | Pa(V_i)) \quad (1)$$

式中: V_i 为网络节点, $Pa(V_i)$ 为节点 V_i 的父节点集。

贝叶斯网络的结构和参数获取方式有两种,一种是由专家直接给出,二是基于样本数据学习得到。依据数据的完整性划分,贝叶斯网络的结构学习和参学算法如表1所示^[11]。本文所使用的贝叶斯网络结构由业内专家直接给出,参数的确定则基于给定样本和最大似然估计算法(MLE)通过自动学习得到。

3 推理模型构建

3.1 指标的选取与分析

大气能见度,是指人眼在被天气条件限制时能够识别出的最远目标物的距离^[8]。影响大气能见度的因素有很多,包括大气消光系数、气溶胶的化学成分、气象因子等。一方面气象因子对大气能见度的影响机制主要通过大气消光系数和气溶胶的化学成分来完成,另一方面气象因子亦是大气消光系数和气溶胶的化学成分影响大气能见度的外在形式,故仅考虑气象因子对能见度的影响亦能保证最终推理结果的客观性。在海洋区域内,影响大气能见度的气象因子主要有海浪、风速、降水、相对湿度、海雾、海表温度、大气稳定度、浮沉、风沙等,其对大气能见度的影响机理如下:

海浪:由于本文所研究的能见度为航道风险评估服务,故本文所指能见度数据为随船人员在被天气条件限制时能够识别出的最远目标物的距离。由于海浪的大小不仅会直接影响人眼的视线,同时由海浪所引起的船舶的晃动也会干扰随船人员对前方目标物的观测,故海浪的大小与能见度呈负相关。

风速:风不仅会对大气污染物有稀释作用,同时能够改变大气稳定度,对能见度产生较大影响。通常风速越大,对大气污染物的稀释作用越强,大气能见度越好。但由于海洋区域内风会导致海浪的产生,这在一定程度上减弱了大气能见度,故风速与能见度有着复杂的关系。

降水:降水不仅会影响大气消光系数,同时干扰人眼视线,故降水对大气能见度有着重要影响。研究表明,能见度随着降水量的增多呈指数降低,且受降水粒子类型、密度、速度、大小等多种因素影

表1 BN自动学习方法

	数据完整条件下	数据缺失条件下
结构学习算法	基于统计测试的方法(如SGS算法、EP算法)	SEM算法
	基于搜索计分的方法(如K2算法、K3算法)	MCMC算法
参数学习算法	贝叶斯估计方法(Bayesian Estimation)	EM算法
	最大似然估计算法(MLE)	Gibbs抽样算法 高斯逼近算法

响^[12]。

相对湿度:大气相对湿度不仅与降水关系密切,同时相对湿度的变化会引起气溶胶化学组分浓度的变化,此导致颗粒物的散射系数发生改变,从而影响大气能见度^[13];

海雾、浮沉、风沙:此三者对于大气能见度有着直接且重要的影响,其中在海洋区域以海雾为主。

海表温度:一方面海表温度越高,近地层对流越强,污染物浓度越低、大气能见度越好,另一方面海表温度与海雾的发生关系密切^[14]。

大气稳定度:稳定的大气层结可阻止低层大气向上对流发展和水汽向高层扩散,最终使水汽聚集在低层大气中,利于海雾生成^[15]。

由于海面压力、露点温度和海气温差分别对降水和海雾的发生有重要影响^[16],故影响大气能见度的气象因子亦包含海面压力和海气温差。由于大气稳定度的主要影响因子为高低层温差和风速^[17],故可不将低层大气稳定度作为一个独立的影响因子。基于以上考量且受限于资料的可获取性,本文仅考虑海浪、风速、海面压力、降水、相对湿度、海雾、海表温度、海气温差等气象因子对能见度的影响。

基于以上确定的能见度影响因子,构建如下能见度推理BN结构,如图1所示。

3.2 参数学习

3.2.1 样本数据生成

鉴于部分海域能见度资料不易获取,本文选择朝鲜海峡作为研究区域来验证所提技术的可操作

性。研究所用数据为美国国家气候资料中心(National Climate Data Center, NCDC)发布的2000—2017年间的国际海洋大气综合数据集(ICOADS), 由于该数据集是由原始观测资料经后期加工形成, 是目前全世界最大的海洋表面数据集, 故其可靠度较高^[18], 但其并非网格点数据。由于本文在数据的前期处理过程中剔除了部分不完整的数据, 故数据的时间间隔不等, 每一组数据为某船只随船人员某天观测所得。本文首先从数据集中筛选出研究海域的数据, 后将各个节点的数据离散化, 即归类为不同的等级, 各节点等级的划分规则见附录。

本文共生成朝鲜海峡海域 16 418 组数据,以 2000 年 1 月份为例,生成后的完整训练样本集如表 2 所示。

3.2.2 基于MLE算法的参数学习

参数学习的目的即通过某种算法挖掘样本数据中隐含的各节点之间的概率关系。首先初始化各个节点的概率分布,包括先验概率和条件概率,然后根据推理机制和样本数据集,采用MLE算法进

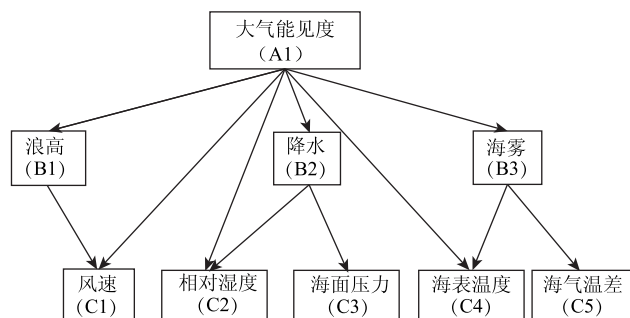


图1 贝叶斯网络结构

表2 完整训练样本集

节点	位置													
	1	2	3	4	5	6	7	8	9	10	11	12	13	...
A1	8	8	9	8	9	9	9	8	7	7	6	4	9	...
B1	2	5	4	4	9	3	1	5	5	5	4	7	4	...
B2	1	1	1	1	1	1	1	1	13	1	1	1	1	...
B3	1	1	1	1	1	1	1	1	1	1	1	1	1	...
C1	4	7	6	6	8	5	4	7	7	11	8	8	4	...
C2	6	7	7	6	7	7	6	9	9	9	8	7	8	...
C3	6	6	8	8	9	9	9	8	7	6	7	6	8	...
C4	30	28	30	30	19	31	31	30	27	29	20	22	25	...
C5	1	1	1	1	1	1	1	1	1	1	10	9	1	...

行参数学习,即对初始化的条件概率分布进行迭代修正以得到与客观训练数据最吻合的概率分布。MLE算法的本质就是试图寻找使得似然函数最大的参数,其工作原理如下:

假定给定独立同分布的观测数据集 $D=\{Y_1, Y_2, \dots, Y_N\}$, 则数据集的似然函数为模型参数的函数,如公式(2)所示:

$$P(D|\theta_s, S^h) = \prod_{i=1}^N P(Y_i|\theta_s, S^h) \quad (2)$$

式中: θ_s 表示未知的参数, S^h 表示网络拓扑结构。由于观测数据集 D 已知,故可通过最大化似然函数来获得参数^[11]。

以2010—2015年1月份数据为训练样本,学习得到的节点A的先验概率分布如表3所示。

4 试验与分析

考虑到能见度具有年变化和年际变化特征规律的可能性,本文分别以朝鲜海峡海域(126°~133°E, 32°~36°N)5 a数据和15 a数据为样本对能见度进行推理,研究海域如图2。鉴于部分海域海雾、降水等数据获取较为困难,本文还将推理模型

进一步简化,并对简化后的模型进行了验证。

4.1 5 a样本数据条件下推理结果分析

以朝鲜海峡海域2010—2014年的数据作为训练样本,推理可得2015—2017年各个月份的能见度等级。为分析能见度是否具有年变化特征规律,将该海域2010—2014年各个月份的多年数据分别作为训练样本,推理得到2015—2017年各个月份的能见度等级,进而可基于公式(3)求得2015—2017年各个月份推理结果的相对误差。

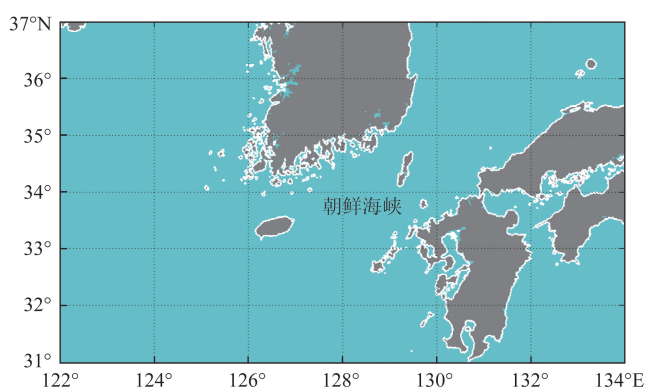


图2 研究海域

表3 节点A先验概率分布

状态	1	2	3	4	5	6	7	8	9	10
先验概率	0.073	0.003	0.001	0.002	0.001	0.013	0.059	0.406	0.420	0.022

表4 相对误差分布表

样本	月份											
	1	2	3	4	5	6	7	8	9	10	11	12
1	0.079	0.084	0.117	0.106	0.130	0.113	0.146	0.075	0.098	0.097	0.083	0.098
2	0.083	0.087	0.104	0.100	0.129	0.151	0.158	0.085	0.106	0.101	0.081	0.067
3	0.081	0.094	0.093	0.109	0.117	0.125	0.167	0.150	0.134	0.098	0.089	0.081
4	0.081	0.097	0.154	0.129	0.137	0.171	0.200	0.093	0.132	0.098	0.084	0.100
5	0.107	0.113	0.131	0.144	0.186	0.174	0.142	0.141	0.102	0.118	0.106	0.106
6	0.106	0.132	0.154	0.134	0.163	0.205	0.168	0.103	0.136	0.145	0.113	0.112
7	0.139	0.319	0.284	0.237	0.177	0.195	0.142	0.101	0.137	0.172	0.123	0.113
8	0.089	0.123	0.124	0.136	0.176	0.245	0.177	0.115	0.106	0.142	0.153	0.106
9	0.063	0.100	0.092	0.143	0.130	0.153	0.148	0.087	0.096	0.119	0.078	0.070
10	0.071	0.074	0.095	0.102	0.116	0.192	0.135	0.076	0.101	0.100	0.075	0.063
11	0.080	0.077	0.082	0.098	0.132	0.173	0.128	0.080	0.105	0.097	0.078	0.062
12	0.087	0.082	0.087	0.098	0.138	0.158	0.152	0.087	0.121	0.123	0.090	0.071
平均值	0.089	0.115	0.126	0.128	0.144	0.171	0.155	0.099	0.115	0.118	0.096	0.087

$$\delta = \frac{|c-r|}{r} \quad (3)$$

式中: c 表示推理结果, r 表示观测值。

由于 ICOADS 数据集由世界范围内许多船只的随船观测数据汇总所得, 故每月均有多组数据。表 4 中样本 1 表示 2010—2014 年 1 月份共 559 组数据作为训练样本, 样本 12 表示 2010—2014 年 12 月份共 590 组数据作为训练样本, 每个月份的数据量不一致; 月份 1 表示 2015—2017 年 1 月份共 292 组数据的平均相对误差。将每月的推理误差进行年平均, 可得不同样本条件下各个月份的能见度年均误差, 如表 4 中平均值项所示。

将各个样本条件下推理得到的样本所在月份能见度相对误差与该月不同样本条件下的误差均值进行对比, 即表 4 中正对角线上数值与对应月份的平均值项进行对比, 如图 3 所示:

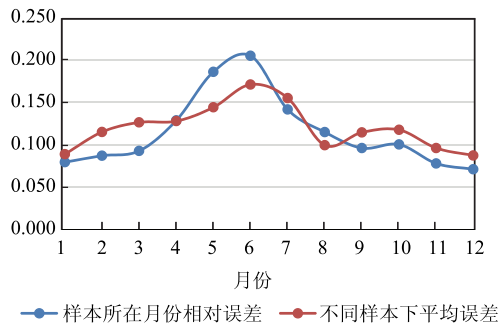


图3 5年样本条件下的误差对比图

可得, 除 4、5、6、8 月份外, 不同样本数据下样本所在月份相对误差均小于对应月份不同样本下的平均误差, 且两者 4 月份的误差值相差极小。由于样本数据时间跨度仅为 5 a, 不同月份下的样本数据并没有充分包含年际变化信息, 故由推理结果可得大气能见度具有年变化特征, 即由多年某一月份的气象数据作为样本推理样本数据所在月份的能见度精度更高。

将 2010—2014 年全部数据作为训练样本进行推理, 得到的能见度误差与不同样本数据下样本所在月份相对误差进行对比, 如图 4 所示。可得除 5、6、8 月份外, 不同样本数据下推理得到的样本所在月份相对误差均小于对应月份全样本条件下推理得到的能见度误差, 由此进一步证实大气能见度具有显著的年变化特征规律。

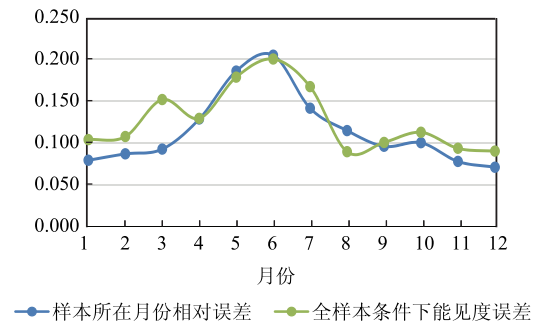


图4 误差对比图(2010—2014年全部数据作为训练样本)

表5 误差分布表

样本	月份											
	1	2	3	4	5	6	7	8	9	10	11	12
1	0.076	0.088	0.111	0.100	0.131	0.122	0.145	0.074	0.095	0.087	0.080	0.076
2	0.075	0.074	0.099	0.096	0.133	0.132	0.153	0.109	0.113	0.114	0.073	0.063
3	0.087	0.105	0.145	0.113	0.131	0.134	0.180	0.171	0.121	0.106	0.103	0.083
4	0.073	0.104	0.122	0.121	0.142	0.186	0.136	0.070	0.100	0.082	0.075	0.091
5	0.106	0.122	0.144	0.123	0.180	0.175	0.166	0.152	0.098	0.124	0.122	0.106
6	0.118	0.130	0.190	0.127	0.176	0.190	0.168	0.104	0.129	0.137	0.113	0.102
7	0.127	0.217	0.252	0.200	0.183	0.176	0.153	0.097	0.118	0.146	0.104	0.103
8	0.116	0.125	0.145	0.153	0.178	0.222	0.171	0.094	0.114	0.140	0.127	0.101
9	0.081	0.108	0.083	0.116	0.173	0.180	0.153	0.082	0.100	0.114	0.072	0.062
10	0.088	0.084	0.079	0.106	0.129	0.195	0.146	0.072	0.099	0.115	0.082	0.063
11	0.086	0.067	0.080	0.096	0.140	0.166	0.119	0.072	0.099	0.091	0.079	0.063
12	0.087	0.078	0.088	0.094	0.144	0.172	0.145	0.079	0.102	0.106	0.083	0.069
平均值	0.093	0.109	0.128	0.120	0.153	0.171	0.153	0.098	0.107	0.113	0.093	0.082

鉴于部分资料不易获取,本模型并未考虑沙尘暴、雾霾等气象因子以及非气象因子等要素对大气能见度的影响,故推理结果存在一定误差。由图3、4可得,相对误差总体保持在0.21级以内,准确性较高,故基于本模型推理得到的能见度结果具有精度较高。

4.2 15 a样本数据条件下推理结果分析

为了进一步验证大气能见度的年变化特征,同时检验大气能见度变化是否具有年际变化特征规律,取2000—2014年的数据作为训练样本,推理可得2015—2017年的能见度等级。分析步骤同5年样本数据条件下的推理结果分析,可得不同样本数据条件下的各个月份的年均相对误差,如表5所示:

将各个样本条件下推理得到的样本所在月份能见度相对误差与该月不同样本条件下的误差均值进行对比,如图5所示。

可得,3、4、5、6、10月份中样本所在月份相对误差均大于对应月份不同样本下平均误差,7月份两种误差值相等,能见度年变化特征减弱。由于样本数据时间跨度较长,单个月份的样本不仅包含年变化特征信息,同时将年际变化特征隐含在内,故相对于5 a样本数据下的推理结果,15 a样本数据下的推理结果所展现出来的大气能见度年变化特征稍弱。由图5可得样本所在月份相对误差值最高为0.190,小于5 a样本条件下的最高误差0.205,可得尽管能见度年变化特征减弱,但由于数据量的增多使得推理结果精度更高。

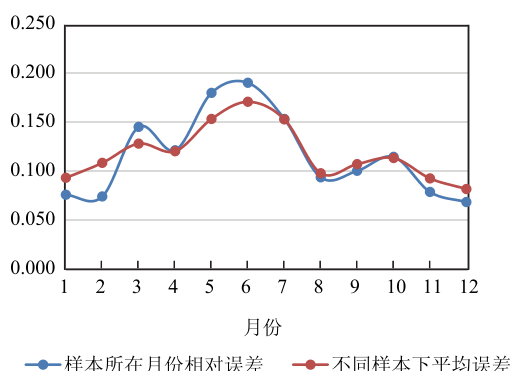


图5 15年样本条件下的误差对比图

以15 a所有数据作为样本,将推理结果与不同样本数据下推理得到的样本所在月份能见度误差进行对比,如图6所示。

可得除5、8月份外,样本所在月份相对误差均小于对应月份在全样本条件下推理得到相对误差,此结果进一步验证了大气能见度的年变化特征规律。为进一步验证大气能见度的年际特征规律,取1970—2017年间朝鲜海峡海域各个月份的能见度等级,其变化趋势如图7所示:

图8为各月能见度等级的48 a平均值变化情况,由图7、8可得,该海域每年能见度最小值均出现在5、6、7月份,最大值均出现在10、11、12、1月份,能见度变化具有显著的年变化特征,且每年变化趋势基本一致。从图7可得,2008—2013年间能见度异常减小,经查阅原始数据,该时间段内朝鲜海峡海域许多地点的能见度等级数据记录为1,即明确说明该时间下朝鲜海峡的能见度较差。具体原因各时间各地点不尽相同,但主要有三,一是该时间下该海域受海雾影响严重,二是受降水影响严重,三是悬浮在空气中的尘埃对能见度的影响;由于部分数据记录不完整,仅有能见度数据而无降水和海雾数据,故无法给出所有时间地点能见度较差的具体原因,但由于数据为实测数据,故能见度数据本身的科学性具有保证。

图9曲线部分为1970—2017年间朝鲜海峡海域各年最低能见度等级变化趋势,直线部分为历史最小值且为极小值出现年份的连线,可得从1970—2006年能见度的极小值每隔4~10 a有所降低,具

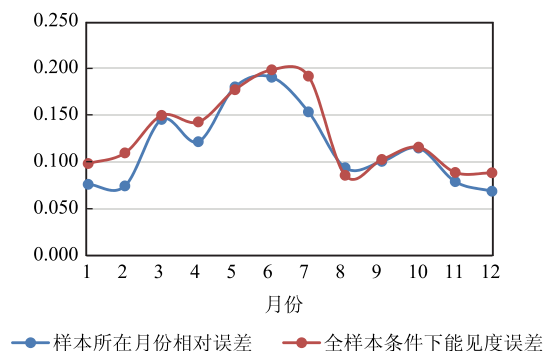


图6 当月能见度误差与全样本条件下能见度误差对比图

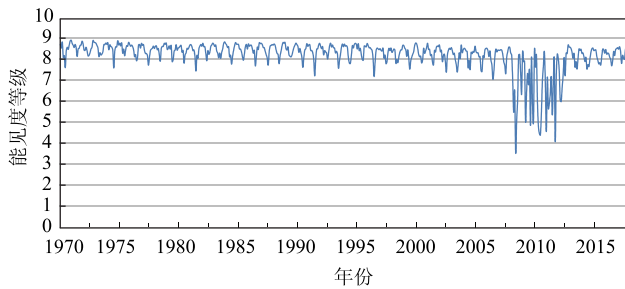


图7 1970—2017年间朝鲜海峡海域各个月份能见度等级变化图

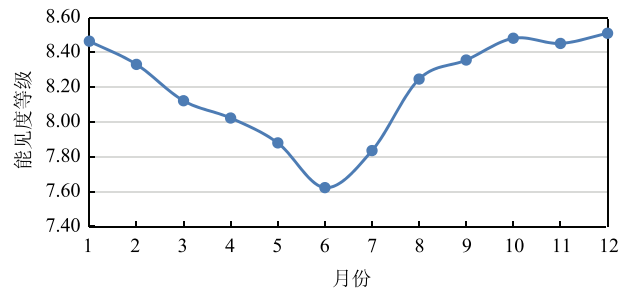


图8 各月能见度等级的多年平均值变化图

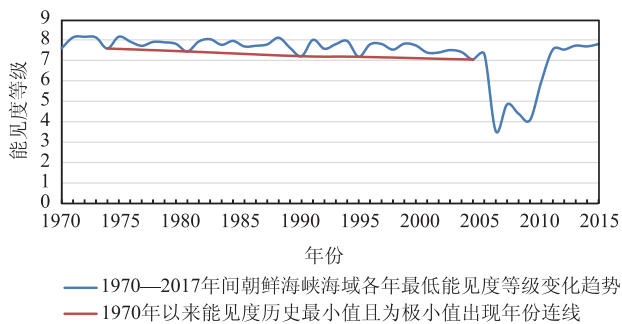


图9 1970—2017年间朝鲜海峡海域各年最低能见度等级变化图

表6 历史能见度最小值与极小值出现年份及其能见度等级

年份	1974	1981	1991	1996	2006
能见度等级	7.58	7.434	7.203	7.173	7.038

体如表6。由于2008年该海域受降水、海雾等天气现象影响严重,能见度达到48 a中最小值,归为异常情况,故不将其纳入规律考察范围。由此可判定能见度具有年际变化规律,变化周期4~10 a不等。

综合以上分析可得,在朝鲜海峡海域能见度的变化具有年变化、年际变化特征。利用贝叶斯网络技术对能见度进行推理时,样本数据不应盲目求多,应以各个月份的多年历史数据分别作为样本来推理样本所在月份的能见度。

4.3 基于简化的贝叶斯网络能见度推理

以上工作验证了贝叶斯网络技术在能见度推理上的可操作性,鉴于以北极海域为代表的许多海域中降水、海雾、海浪等资料的不易获取,需将推理模型进一步简化以适应该海域能见度的推理工作,

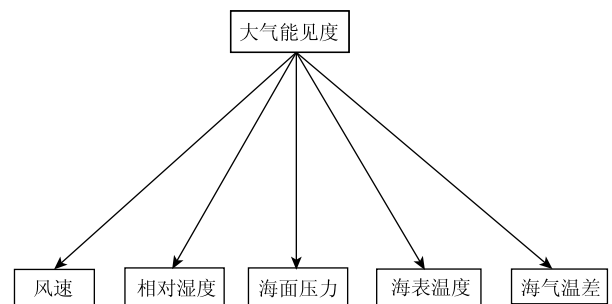


图10 简化后的BN网络

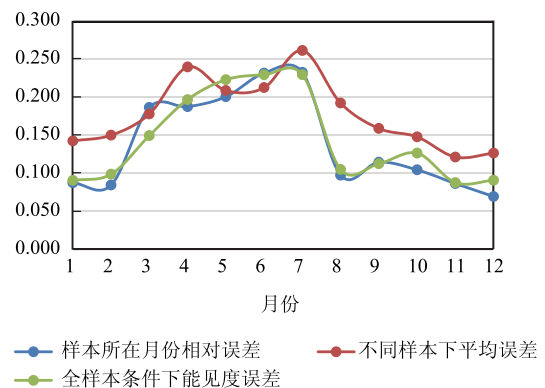


图11 误差对比图

简化后的贝叶斯网络如图10所示:

基于简化后的BN网络,在15 a样本数据条件下以同样的方法对能见度进行推理,结果如图11所示:

样本所在月份的相对误差的年均值为0.140,全样本条件下能见度误差年均值为0.144,且从图中可得样本所在月份相对误差值明显小于不同样本下平均误差值,符合前文所揭示的规律。其中,

样本所在月份相对误差最大值为0.232,略大于完整BN网络下的推理最大误差0.19,符合客观现实。同时此误差表明实际能见度等级与推理得到的能见度等级之间的最大差值不高于3级,准确度较高。

5 结论与不足

基于试验结果,可得出以下结论:

(1)能见度具有年变化、年际变化特征。由实验可得,朝鲜海峡海域每年能见度最小值出现在5、6、7月份,最大值出现在11、12、1月份,且在1970—2006年间每隔4~10 a能见度的极小值有所下降;

(2)基于BN技术对能见度进行推理的准确性较高。由于能见度具有年变化特征,故所用样本数据不应盲目求多,经实验得利用多年某月份的数据作为样本推理该月的能见度等级具有最高的准确性。5 a数据为样本条件下各月能见度推理误差均不高于0.21级,15 a数据为样本条件下各月能见度推理误差均不高于0.19级,可得相同样本形式下样本数据较多者准确性较高;

(3)降水、海雾等天气现象对能见度具有较大影响。简化后的BN并未将降水、海雾等天气现象直接纳入子节点,致使在相同样本条件下基于简化后的BN推理误差较基于完整意义上的BN推理误差高出0.048。由于简化BN考虑到了相对湿度、海面压力、海表温度、海汽温差等因素对能见度的影响,且相对湿度和海面压力会直接影响降水,海表温度和海汽温差会直接影响海雾,故两种BN网络下的推理误差均保持在较小的区间内。

由于本文所推理出的能见度为等级数据,无法计算出能见度确切数值,故基于本文所提技术得到的能见度数据应用范围有限。为进一步解决此问题,可将能见度等级进一步细化,将各个能见度等级对应的数值区间进一步缩小。

参考文献:

- [1] 中国外贸海上运输占九成 港口成经济发展命脉[N]. 人民日报, 2016-06-20.
- [2] 北极航道的优与劣[N]. 航运交易公报, 2017-06-06.
- [3] 樊高峰, 马浩, 张小伟, 等. 相对湿度和PM_{2.5}浓度对大气能见度的影响研究: 基于小时资料的多站对比分析[J]. 气象学报, 2016, 74(6): 959-973.
- [4] 李学彬, 徐青山, 魏合理, 等. 相对湿度与能见度的相关性研究[C]//2007年中国气象学会年会论文集. 广州: 中国气象学会, 2007: 5.
- [5] 傅刚, 李晓岚, 魏娜. 大气能见度研究[J]. 中国海洋大学学报, 2009, 39(5): 855-862.
- [6] Gültepe I, Müller M D, Boybeyi Z. A new visibility parameterization for warm-fog applications in numerical weather prediction models[J]. Journal of Applied Meteorology and Climatology, 2006, 45(11): 1469-1480.
- [7] 胡波, 杜惠良, 郝世峰, 等. 一种统计技术结合动力释用的沿海海雾预报方法[J]. 海洋预报, 2014, 31(5): 82-86.
- [8] 盛裴轩, 毛节泰, 李建国, 等. 大气物理学[M]. 北京: 北京大学出版社, 2003.
- [9] 陈丽, 翟崇治, 蒋佳凌, 等. 大气能见度重要影响因子分析[J]. 山东化工, 2015, 44(9): 169-171.
- [10] 李明, 张韧, 洪梅, 等. 基于信息流改进的贝叶斯网络结构学习算法[J]. 系统工程与电子技术, 2018, 40(6): 1385-1390.
- [11] 史志富, 张安. 贝叶斯网络理论及其在军事系统中的应用[M]. 北京: 国防工业出版社, 2012.
- [12] 刘西川, 高太长, 刘磊, 等. 降水现象对大气消光系数和能见度的影响[J]. 应用气象学报, 2010, 21(4): 433-441.
- [13] Hyslop N P. Impaired visibility: the air pollution people see[J]. Atmospheric Environment, 2009, 43(1): 182-195.
- [14] 王鑫. 黄海海雾形成的气候特征以及能见度分析[D]. 青岛: 中国海洋大学, 2004.
- [15] 卞若玢, 杨波, 赵世昌. 大连市旅顺口区海雾气候特征及形成机制分析[J]. 现代农业科技, 2017, (12): 211-212.
- [16] 曲平, 解以扬, 刘丽丽, 等. 1988—2010年渤海湾海雾特征分析[J]. 高原气象, 2014, 33(1): 285-293.
- [17] 李祥余. 大气稳定度分类方法及判据比较研究[J]. 环境与可持续发展, 2015, 40(6): 93-95.
- [18] 李晓婷, 郑沛楠, 王建丰, 等. 常用海洋数据资料简介[J]. 海洋预报, 2010, 27(5): 81-89.

附录

表 1 能见度等级划分

ICOADS 能见度数据	90	91	92	93	94	95	96	97	98	99
等级	1	2	3	4	5	6	7	8	9	10
大小/km	<0.05	0.05	0.2	0.5	1	2	4	10	20	≥50

表 2 浪高等级划分

浪高大小/m	浪高等级
0	1
0.5	2
1.0	3
1.5	4
2.0	5
2.5	6
3.0	7
3.5	8
4.0	9
4.5	10
5.0	11
5.5	12
6.0	13
6.5	14
7.0	15
7.5	16
8.0	17
8.5	18
9.0	19
9.5	20
10.0	21
>10.0	22

表 4 海雾等级划分

ICOADS 天气状况数据	海雾等级
其他	1
40	2
41	3
42	4
43	5
44	6
45	7
46	8
47	9
48	10
49	11

表 3 降水等级划分

ICOADS 天气状况数据	降水等级
其他	1
50	2
51	3
52	4
53	5
54	6
55	7
56	8
57	9
58	10
59	11
60	12
61	13
62	14
63	15
64	16
65	17
66	18
67	19
68	20
69	21
70	22
71	23
72	24
73	25
74	26
75	27
76	28
77	29
78	30
79	31
80	32
81	33
82	34
83	35
84	36

ICOADS天气状况数据	降水等级
85	37
86	38
87	39
88	40
89	41
90	42
91	43
92	44
93	45
94	46
95	47
96	48
97	49
98	50
99	51

表6 风速等级划分

风速/(m/s)	风力等级
0.0~0.2	1
0.3~1.5	2
1.6~3.3	3
3.4~5.4	4
5.5~7.9	5
8.0~10.7	6
10.8~13.8	7
13.9~17.1	8
17.2~20.7	9
20.8~24.4	10
24.5~28.4	11
28.5~32.6	12
32.7~36.9	13
37.0~41.4	14
41.5~46.1	15
46.2~50.9	16
51.0~56.0	17
56.1~61.2	18
≥61.3	19

表5 海表温度等级划分

海表温度/℃	等级
<-10	1
-10~-9	2
-9~-8	3
-8~-7	4
-7~-6	5
-6~-5	6
-5~-4	7
-4~-3	8
-3~-2	9
-2~-1	10
-1~0	11
0~1	12
1~2	13
2~3	14
3~4	15
4~5	16
5~6	17
6~7	18
7~8	19
8~9	20
9~10	21
10~11	22
11~12	23
12~13	24
13~14	25
14~15	26
15~16	27
16~17	28
17~18	29
18~19	30
19~20	31
20~21	32
21~22	33
22~23	34
23~24	35
24~25	36
25~26	37
26~27	38
27~28	39
28~29	40
29~30	41
≥30	42

表7 相对湿度等级划分

相对湿度/%	0~10	10~20	20~30	30~40	40~50	50~60	60~70	70~80	80~90	90~100
等级	1	2	3	4	5	6	7	8	9	10

表 8 海平面压力等级划分

压力/hPa	等级
<990	1
990~995	2
995~1 000	3
1 000~1 005	4
1 005~1 010	5
1 010~1 015	6
1 015~1 020	7
1 020~1 025	8
1 025~1 030	9
1 030~1 035	10
1 035~1 040	11
>1 040	12

表 9 海气温差等级划分

海气温差/℃	等级
其他	1
-10	2
-9	3
-8	4
-7	5
-5	6
-4	7
-3	8
-2	9
-1	10

Inference model and experiment for visibility in the Korean Strait based on Bayesian Network

SHAN Yu-long, ZHANG Ren, MAO Ke-feng

(Institute of Meteorology, National University of Defense Technology, Nanjing 211101 China)

Abstract: In this paper, considering the lack of visibility data in assessing the risk along marine trade routes, we present a new approach in inferring visibility data based on Bayesian Network. The inference model of visibility data based on Bayesian Network is built through the determination of study area, choice of nodal index, generation of sample datasets, configuration of inference model, parameter learning and inference calculation. The inference model is applied to Korean Strait as an experiment. It is found that visibility reveals annual and inter-annual features. It is of high accuracy to infer the visibility level of a specific month using multiple-year data of that month as training samples. Furthermore, we also demonstrate that the amount of sample data has a positive effect on the inference accuracy.

Key words: Bayesian Network; visibility; Korean Strait