

杜星, 孙永福, 宋玉鹏, 等. 基于谱聚类算法的海底滑坡危险性评价[J]. 海洋学报, 2021, 43(1): 93–101, doi:10.12284/hyxb2021023

Du Xing, Sun Yongfu, Song Yupeng, et al. Risk assessment of submarine landslide based on spectral clustering[J]. Haiyang Xuebao, 2021, 43(1): 93–101, doi:10.12284/hyxb2021023

基于谱聚类算法的海底滑坡危险性评价

杜星^{1,2}, 孙永福^{3,4}, 宋玉鹏^{2,3}, 修宗祥^{2,3}, 单冶钢^{1*}

(1. 中国电建集团华东勘测设计研究院有限公司, 浙江 杭州 311122; 2. 自然资源部第一海洋研究所, 山东 青岛 266061; 3. 青岛海洋科学与技术试点国家实验室 海洋地质过程与环境功能实验室, 山东 青岛 266237; 4. 国家深海基地管理中心, 山东 青岛 266237)

摘要: 海底滑坡的危险性评价与分区, 对海洋工程设施的选址和危险预防具有指导作用。本文基于无监督机器学习的谱聚类算法对黄河口埕岛海域展开了海底滑坡危险性评价, 构建了输入参数为 9、输出类别为 4、核函数参数为 0.08 的海底滑坡危险性评价模型。使用该模型进行评价, 将研究区分为海底滑坡危险性高、较高、较低和低的区域。评价结果与地质环境因素分布特征对比显示, 最重要的影响因素为海底沉积物类型和水动力作用, 最重要的触发因子为液化。模型参数分析结果显示, 合理简化输入因子可获得精度略低的评价结果, 而核函数参数是影响评价准确性的重要指标。以上研究表明, 谱聚类算法能够较好地用于海底滑坡危险性评价, 数据类别丰富度和精度是影响评价精细程度的重要因素。

关键词: 海底滑坡; 黄河口; 危险性评价; 无监督学习; 谱聚类

中图分类号: P642.22

文献标志码: A

文章编号: 0253-4193(2021)01-0093-09

1 引言

海底滑坡是一种极具破坏力的海洋地质灾害现象。大型海底滑坡甚至能够引起数千立方千米的沉积物进行长距离运移^[1], 对海洋采油平台^[2]、海底光缆^[3]、电缆等各类工程设施造成损害甚至引发海啸, 造成通信中断、平台倒塌等一系列事故, 给人类的生命安全和财产造成巨大的威胁^[4-5]。因此, 研究海底滑坡的特征、机理并进行危险性预测对于海洋工程的安全性具有重要作用。

目前, 海底滑坡的研究方向主要有: 利用高精度地球物理探测进行滑坡形态识别和分类^[6-8]、通过数值分析的方法开展海底滑坡稳定性计算^[9-10]以及使用常规水槽或离心机等物理模型试验模拟滑坡过程^[11-13]。

虽然通过上述常规研究取得了较大的进展, 然而受海底滑坡自身控制条件复杂、触发影响因素众多、监测难度大等原因的影响, 危险性评价分级方面的研究仍然较为不足。基于机器学习进行滑坡危险性预测的方法在陆上滑坡研究^[14-17]中已有运用并成为一种热门的研究手段, 然而该方法用于海底滑坡的研究仍然较少。同时由于陆上滑坡规模相对较小、结果易于监测、影响因素单一等原因, 适用于有监督机器学习, 而海底滑坡大多无法得到明确的影响因素和危险性结果, 所以不太适用于陆地上常用的方法。为此, 本文引入了无监督机器学习中的谱聚类算法构建海底滑坡危险性评价体系, 结合研究区各类地质环境影响因子展开海底滑坡危险性评价分级。

收稿日期: 2019-12-11; 修订日期: 2020-02-23。

基金项目: 中国电建集团华东勘测设计研究院有限公司 201 项目(KY2018-ZD-01); 山东省自然科学基金(ZR2020QD073); 国家重点研发计划项目(2017YFC0307305); 国家自然科学基金(41606084)。

作者简介: 杜星(1991—), 男, 辽宁省大连市人, 工程师, 主要从事海洋工程地质与灾害地质方面研究。E-mail: duxing@fio.org.cn

* 通信作者: 单冶钢(1965—), 男, 高级工程师(教授级), 主要从事水电工程地质、岩土力学方面的研究。E-mail: shan_zg@ecidi.com

2 谱聚类算法

谱聚类算法 (Spectral Clustering) 是人工智能领域的一种无监督机器学习算法, 主要用于处理没有标识信息的数据。通过对样本数据的拉普拉斯矩阵的特征向量进行聚类, 从而达到对样本数据聚类的目的。谱聚类将高维空间的数据映射到低维, 然后在低维空间用其他聚类算法进行聚类。与最常用的 Kmeans 聚类方法相比, 谱聚类算法使用了降维技术, 更适用于高维数据的处理, 同时对于处理稀疏数据更为有效。海底滑坡危险性评价的输入因子类别多、数据不集中, 因此采取谱聚类比传统的聚类方法更为适用。谱聚类通过输入 n 个样本点 $X = \{x_1, x_2, \dots, x_n\}$ 和聚类簇的数目 k , 最终输出聚类簇 A_1, A_2, \dots, A_k 。具体算法步骤为:

(1) 计算 $n \times n$ 的相似度矩阵 W (包括最小邻近值法、 k 临近法及全连接法), 下式为本文使用的全连接法表达式:

$$s_{ij} = s(x_i, x_j) = \sum_{i=1, j=1}^n \exp \frac{-\|x_i - x_j\|^2}{2\sigma}, \quad (1)$$

式中, W 为 s_{ij} 组成的相似度矩阵; 核函数参数 σ 控制着样本点的邻域宽度, 即 σ 越大表示样本点与距离较远的样本点的相似度越大;

(2) 计算度矩阵 D :

$$d_i = \sum_{j=1}^n \omega_{ij}, \text{ 即相似度矩阵 } W \text{ 的每一行元素之和。}$$

D 为 d_i 组成的 $n \times n$ 对角矩阵;

(3) 计算拉普拉斯矩阵 $L = D - W$;

(4) 计算 L 的特征值, 将特征值从小到大排序, 取前 k 个特征值, 并计算前 k 个特征值的特征向量 u_1, u_2, \dots, u_n ;

(5) 将上面的 k 个列向量组成矩阵 $U = \{u_1, u_2, \dots, u_n\}$, $U \in R^{n \times k}$;

(6) 令 $y^i \in R^k$ 是 U 的第 i 行的向量, 其中 $i = 1, 2, \dots, n$;

(7) 使用 Kmeans 算法将新样本点 $Y = \{y_1, y_2, \dots, y_n\}$ 聚类成簇 C_1, C_2, \dots, C_k ;

(8) 输出簇 A_1, A_2, \dots, A_k , 其中 $A_i = \{j | y_j \in C_i\}$ 。

以上为谱聚类算法的描述。即先根据样本点计算相似度矩阵, 然后计算度矩阵和拉普拉斯矩阵, 接着计算拉普拉斯矩阵前 k 个特征值对应的特征向量, 最后将这 k 个特征值对应的特征向量组成的矩阵 U , U 的每一行成为一个新生成的样本点, 对这些新生成的样本点聚成 k 类, 最后输出聚类的结果。

3 数据和方法

3.1 数据来源

本文研究数据由自然资源部第一海洋研究所在黄河口埕岛海域通过地球物理探测、钻探及监测调查等手段获得^[18], 包含了研究区详细的各类地质特征资料, 并根据地质环境数据特点将水深、波浪、液化深度等各类影响因素分为 4 个级别。黄河口埕岛海域分布着大规模的海底滑坡体系, 具有高含水、浅表多发、广泛分布等特点^[19-20], 水深普遍小于 18 m, 底质类型在 15 m 水深以浅主要为粉土夹杂少量砂土, 大于 15 m 时为粉质黏土和黏土 (图 1)。海底坡度较大的区域大于 0.002°, 主要分布在 6 m 和 12 m 水深等深线周围。易液化区域集中于 3~15 m 水深之间, 从区域中央至外缘液化强度逐渐减小。同时, 中石化胜利油田采油厂的主力区块也位于本研究海域内部, 分布着上百座海上采油平台以及数十条海底输油、输气管道和海底电缆。该区域具有大量详细的物探、钻探以及现场监测获得的各类数据, 能够准确有效地为无监督机器学习提供输入数据。

3.2 研究方法

由于海底滑坡问题的环境影响因素、触发因子众多且难以实时监测, 无法使用地质参数与滑坡危险性之间对应的结果进行有监督学习, 因此本文以无监督学习方法为基础建立黄河口水下三角洲海底滑坡危险性评价模型。基于谱聚类算法的海底滑坡危险性评价主要分为输入输出参数类别确定、原始数据处理、网络参数选取以及评价 4 个部分。

(1) 输入、输出参数类别

海底滑坡危险性影响因素众多, 主要可以分为内因和外因, 其中内因是滑坡体本身的物理力学性质; 外因则受不同地质环境的影响存在多样性。本研究区的内因主要有沉积物类型、海底地形坡度、土体强度, 表征了滑坡体本身; 外因为水深、50 年一遇波高大小、底层最大流速、液化、海底冲刷以及人类工程活动情况。另外, 将区域评价结果分为 4 类是评价分级较为常用的一种方式。因此本研究中输入节点数为 9 个, 对应输入类别的 9 个参数; 输出节点数为 4 个, 分别为危险性高、危险性较高、危险性较低和危险性低 (图 2)。

(2) 原始数据处理

根据图 1 所示研究点位, 按照研究点位的坐标分别从输入参数中获取相对应的数据。文中所使用的 9 个输入参数的数据已按照不同参数类型及数据特

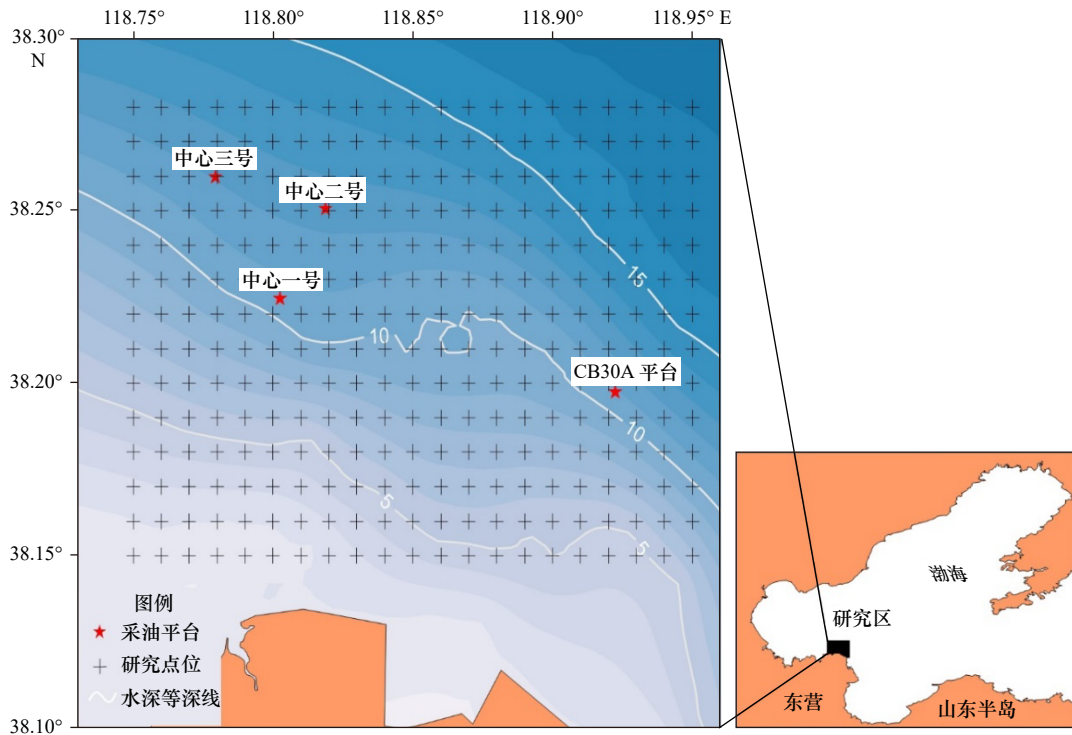


图1 研究区位置及研究点位
Fig. 1 Location of studying area and points

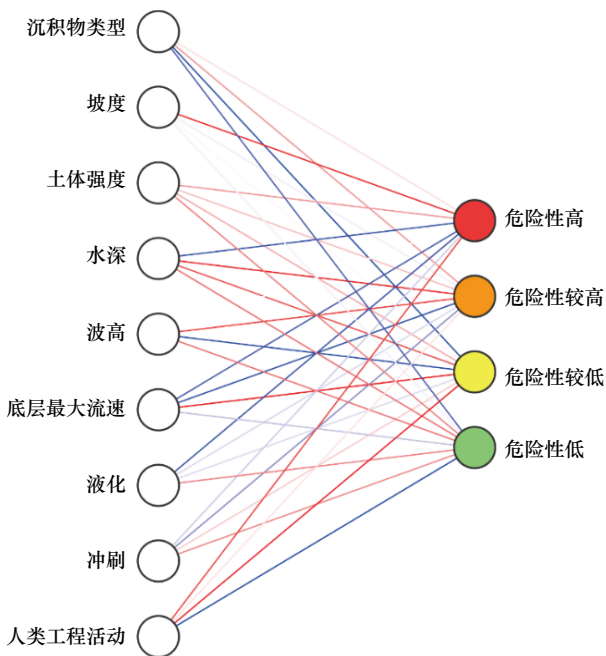


图2 海底滑坡危险性评价网络结构示意图
Fig. 2 Schematic diagram of submarine landslide risk assessment network structure

图中左侧为网络输入因子类别, 右侧为输出因子类别, 连线表示聚类的左右关系示意, 并非一一对应关系
The left side is the network input factor category, and the right side is the output factor category. The connecting line indicates the left-right relationship of clustering, not one-to-one correspondence

点分为了4个级别, 因此每个研究点位拥有9个参数, 每个参数的数值为1~4的不同类别。

(3)核函数参数选择

本文计算相似度矩阵式使用全连接方法, 其中的核函数为高斯核函数(数值大于0)。对核函数参数进行调参时, 保持数据及其他网络参数不变, 令核函数参数为0.01~2.3, 以0.01为间隔进行变化, 针对每个核函数参数均进行聚类并使用 Calinski-Harabasz 评分标准对结果进行打分, 选择评分最高时的核函数参数作为评价网络最终的参数。

(4)海底滑坡危险性评价分区

使用确定了输入输出参数类别、网络参数后可得到最终的海底滑坡无监督学习评价网络, 使用该网络对研究点位数据进行谱聚类训练可获得具有相同特征研究点位的无标签分类结果。

(5)区域标签的赋予

根据谱聚类分析得到的无标签分类结果, 结合研究区地质环境因素分布特征, 对分类结果赋予滑坡危险性评价等级标签, 得到最终的海底滑坡危险性评价结果。

4 结果与分析

4.1 海底滑坡危险性评价结果

通过对不同的核函数参数展开海底滑坡危险性聚类分析后, 基于 Calinski-Harabasz 评分标准的得分

结果见图3。随着核函数参数逐渐增加,聚类效果得分先增加后减小,在核函数参数为0.08时具有最大的得分。得分反映了聚类结果的优化程度,得分越高则说明优化性越好,因此选择核函数参数值为0.08作为滑坡危险性评价网络的最终参数。

使用建立的基于无监督学习模型对黄河口海底滑坡危险性进行了评价(图4),整个研究区从西南至

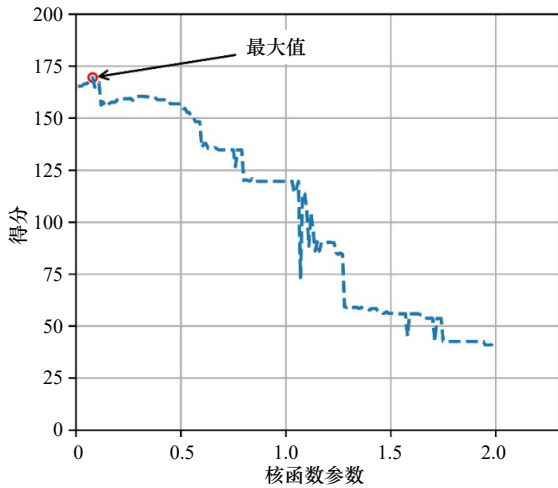


图3 评价网络得分随核函数参数的变化

Fig. 3 Evaluation network score changes with kernel function parameters

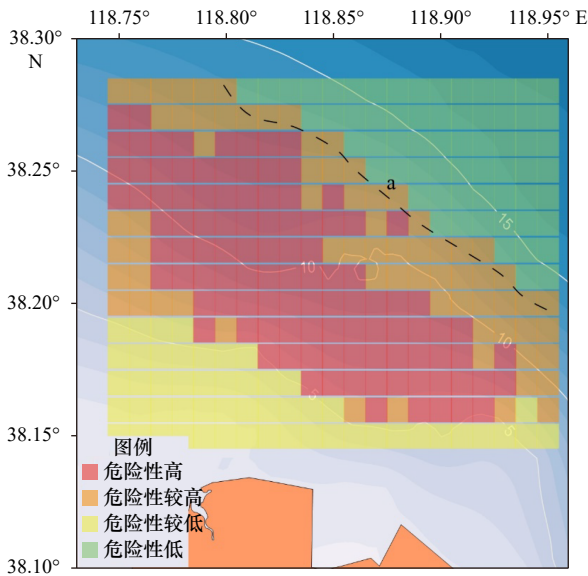


图4 黄河口海底滑坡危险性评价结果

Fig. 4 Risk assessment result of submarine landslides in the Yellow River Estuary

虚线a为危险性较高和危险性低分区间的突变界限,由海底沉积物类型的突变引起

The dotted line a is the mutation boundary between high risk and low risk zones, which is caused by the mutation of seabed sediment types

东北被分为了4个区域。分别为水深5 m以浅、水深14 m以深、水深6~12 m之间及其周边区域。海床受到的水动力作用在波浪、水深及海流共同作用下,随着水深从浅到深变化首先逐渐增强,在9~10 m附近达到最大值,随后逐渐减弱。同时,研究区大致以水深14~15 m等深线为分界线,深水区海底土体为黏土类沉积物,浅水区为粉土夹杂少量粉砂。另外,地球物理探测调查结果显示,研究区内15 m以深基本没有液化的发生,而9~10 m水深等深线附近最易发生液化等现象。因此不难判断,海底滑坡危险性高的区域:大致范围在6~12 m水深等深线内呈条带状分布,长轴方向与水深等深线方向平行,代表了本研究区内最容易出现海底滑坡的区域;危险性较高的区域:分布在危险性高的部分外围,水深范围扩展至5~14 m,代表了本研究区内较为容易出现海底滑坡的区域;危险性较低的区域:分布在研究区的水深小于5 m位置,表示研究区内较不容易出现海底滑坡的区域;危险性低的区域:位于研究区水深14 m以深的位置,表示本研究区内最不容易产生海底滑坡的区域。

本文得到的聚类结果显示,黄河口海底滑坡危险性较高和高的区域位于水深5~14 m区域内,其他水深位置危险性较低。彭俊等^[21]通过高精度地球物理探测手段对黄河三角洲水下岸坡展开了调查,调查结果显示,在水深6~14 m区域主要发育了海底滑坡,6 m以浅和14 m以深未见海底滑坡发育。该物探调查结果与本文利用无监督机器学习得到的海底滑坡危险性分区结果一致性较高,均显示水深5~14 m的区域容易产生海底滑坡。表明本文使用的算法可以取得与真实情况较为接近的结果,能够较好地适用于海底滑坡危险性评价。

4.2 影响因素与触发因子

研究区海底滑坡危险性随着水深逐渐增加总体上呈现较低-较高-高-较高-低的趋势变化。当水深小于5 m时,波浪、海流等水动力作用较小,因此滑坡的触发因子强度不足,导致了水深小的区域危险性较低。随着水深逐渐增加,海床受到的水动力作用越来越强,但水深增加至一定程度时海床受到的水动力作用又会减小。因此滑坡危险性也呈现先增强后减弱的趋势,最严重位置在水深9~10 m附近。若研究区海底沉积物类型一致,则在14 m水深位置应先出现危险性较低的区域再逐渐过渡至危险性低的区域。但由于海底沉积物类型由可液化的粉土突

变为难以液化的黏土类沉积物,因此造成了危险性分区的突变(图4,突变分界线为a)。在水深变深引起的水动力作用减弱和海底沉积物性质的共同影响下,研究区内水深大于14 m的区域成为了最不容易产生海底滑坡的部分。坡度、土体强度和冲刷虽然与最终结果呈现相关关系,但远没有水动力和沉积物类型与结果相关性程度高。而人类工程活动分布对结果影响最小,研究区内海底滑坡的分布几乎未受到人类活动的影响。总的来说,研究区内水深、波高、底层最大流速和海底沉积物类型是最重要的影响因素,坡度、土体强度和冲刷影响程度次之,人类工程活动影响最弱。

通常情况下海底滑坡的外界影响触发因子有地震、天然气水合物、波浪作用、火山活动、海啸等一系列条件,不同的海域受周边地质环境条件的影响存在不同的因子。本文研究区内并不存在地震、海啸、火山等剧烈地质现象,主要外界影响因子为波浪引起的海底土体液化。黄河口埕岛海域在50年一遇波浪工况下的液化深度分布如图5所示^[22],随着水深逐渐加深,液化深度也呈现先增加后减小的分布情况,在

水深6~12 m之间具有较大的液化深度分布。对比图4、图5可知,研究区液化深度分布的特征与本文得到的海底滑坡危险性评价结果有着密切的相关性和一致性,均在水深10 m等深线附近分布着地质灾害最严重的区域,当水深减小或继续增加时地质灾害的严重性越来越弱。足以充分表明波浪作用下海底沉积物液化是研究区内海底滑坡最重要的触发因子。

5 算法参数讨论

5.1 输入因子简化

本文采用了沉积物类型、海底地形坡度、土体强度、水深、波高、底层最大流速、液化、海底冲刷以及人类工程活动情况9种参数作为网络输入因子,进行海底滑坡危险性分析。因某些海域输入因子的难获取性,也可以选择沉积物类型、水深、波高和底层最大流速作为输入因子进行谱聚类。沉积物类型能够表述海底土体的性质。水深、波高、底层最大流速能够代表水动力作用强度,也可换算求得海底地形坡度,并影响液化、冲刷。因此简化的模型选择沉积物类型、水深、波高和底层最大流速作为输入因子进行

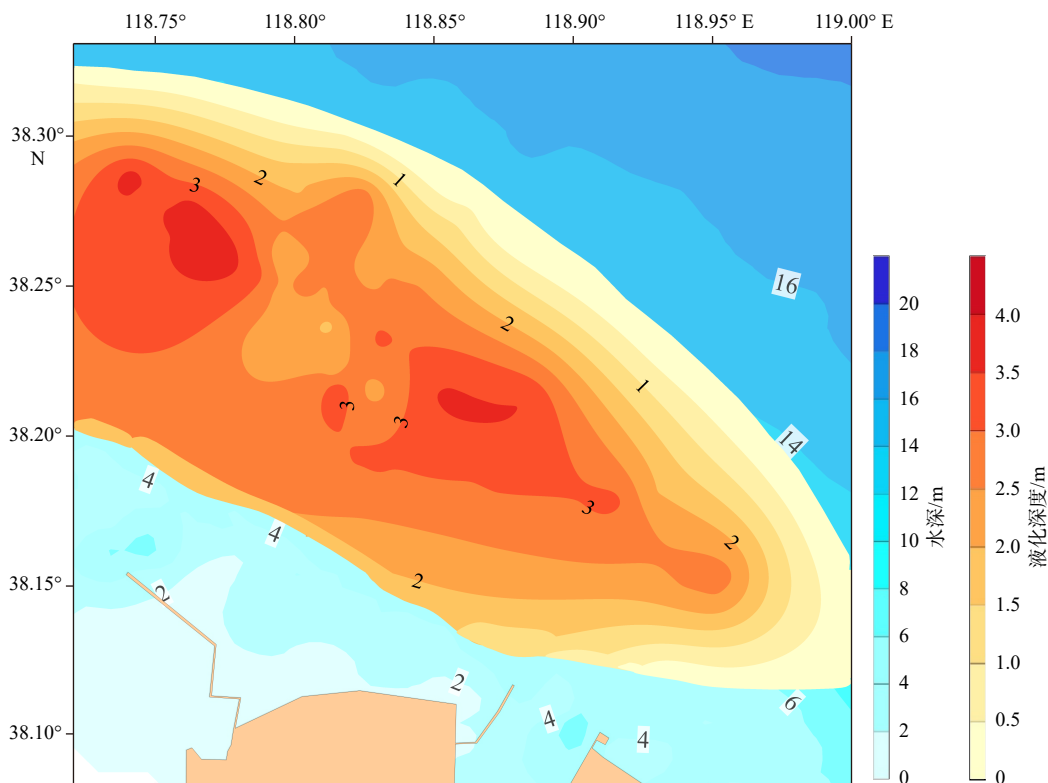


图5 黄河口埕岛海域液化深度分布^[22]

Fig. 5 Distribution of liquefaction depth in Chengdao sea area of the Yellow River Estuary^[22]

液化深度分布使用了50年一遇风浪条件下数据进行计算,并与物探测得的液化扰动层分布进行了对比验证

The liquefaction depth distribution is calculated by using the data of 50 years return period wind and wave, and compared with the distribution of liquefaction disturbance layer measured by geophysical prospecting

谱聚类。评价结果如图 6 所示,与未简化评价结果对比,危险性低和较低的区域基本一致,但是危险性高的区域范围减小、危险性较高区域范围扩大且二者未呈现包含关系。由此可见,适当地对输入因子进行简化能够较为粗略地对研究区展开海底滑坡危险性评价,但评价结果的准确度相对降低。

5.2 核函数参数

为检验该参数对聚类结果的影响程度,令核函数参数分别为 0.1、1、5、10,并构建 4 个不同参数的评价网络进行海底滑坡危险性评价。评价结果如图 7 所示,随着核函数参数的逐渐增加,危险性高的区域呈现一种扩大的趋势。 σ 为 0.01 和 0.1 时评价结果与上文得到的结果相似,只是危险性高的区域范围稍有变化。当 σ 为 1 时,危险性高的区域已完全与危险性较低区域相邻,未能体现出二者间危险性较高区域的过渡。当 σ 为 10 时,该模型已经无法进行有效的评价。以上研究结果表明,核函数参数评分相近时都可以较好地与研究区进行聚类分析,但相差较大时则无法准确对数据进行聚类分析。

5.3 算法适用性

海底滑坡、海底地震等多种海底地质灾害难以有效地监测到产生的过程和结果,因此并没有初始因素—最终结果——对应的数据用于有监督机器学习的训练。然而无监督机器学习算法以其不需要结果、可以自行学习的功能可以有效克服这一问题,所以无监督机器学习算法成为了能够解决此类问题的关键。

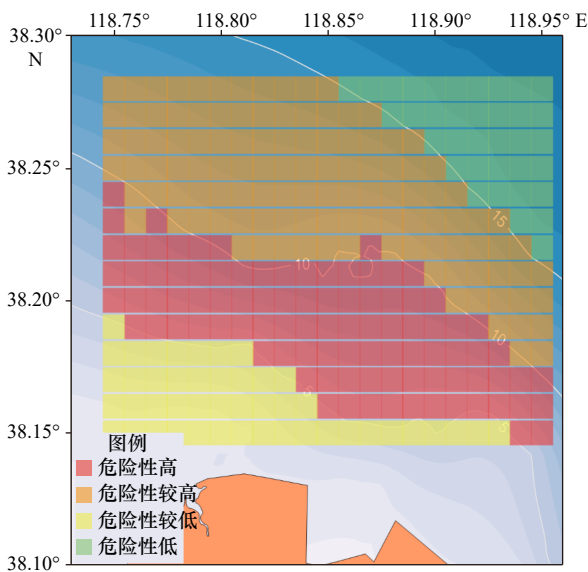


图 6 简化输入因子得到的海底滑坡危险性评价结果

Fig. 6 Risk assessment results of submarine landslides from simplified input factors

结果表明,谱聚类方法得到的海底滑坡危险性分类与实际海底滑坡区域分布较为一致,具有较高的相似性和可靠性。研究区内产生的滑坡与传统海底滑坡相比具有规模小、坡度缓、滑动距离短等特点,二者的滑坡危害性显然无法相提并论。因此,聚类的结果并不表示海底滑坡危险性的绝对值,而表示了与研究区内危险性的相对值,不同危险性高低代表了在研究区内产生海底滑坡现象的可能性和严重程度。使用谱聚类方法进行分类后,可以良好地估计研究区域内海底滑坡危险性的相对高低,有利于进行海洋工程建设时的选址。

现场实测数据的缺乏仍然是限制海底滑坡危险性分析发展的最重要因素,5.1 节结果表明只有获得更丰富、更精确、更详细的现场地质环境数据,才能够更加准确的评价结果。进行其他海域的海底滑坡危险性评价时,需要根据研究区地质环境因素特点,选择适宜的输入因子和输出因子。受深海调查数据的影响,本文未选在东海、南海等滑坡发育更为明显的区域,接下来将进一步搜集资料与现场探测相结合,本文提出的算法可用于更加广泛的研究区域。

6 结论

本文使用了无监督机器学习的技术手段,构建了海底滑坡危险性评价方法,通过对黄河口埕岛海域的多种地质参数分析并进行了计算,分析了海底滑坡危险性评价结果并对算法参数展开了讨论,主要得到以下结论:

(1) 使用基于无监督机器学习的谱聚类算法能够对黄河口海底滑坡进行准确的评价分类,表示研究区内不同分组的海底滑坡相对危险性,当网络中含有 9 个输入节点、4 个输出节点、核函数参数为 0.08 时具有最好的聚类效果。

(2) 黄河口埕岛海域海底滑坡危险性与水深分布相关性较大,危险区域分布于 6~12 m 水深,较危险区域分布于危险区域外围的 5~14 m 水深,较安全区域分布于水深小于 5 m 的位置,安全区域位于水深大于 14 m 的位置。

(3) 基于谱聚类方法得到的黄河口埕岛海域海底滑坡危险性分区结果与各类地质环境因素都存在相关关系,其中,沉积物类型和水动力作用是最根本的影响因素,液化是本研究区海底滑坡最重要的触发因子。

(4) 模型核函数参数应选择聚类结果评分高的,

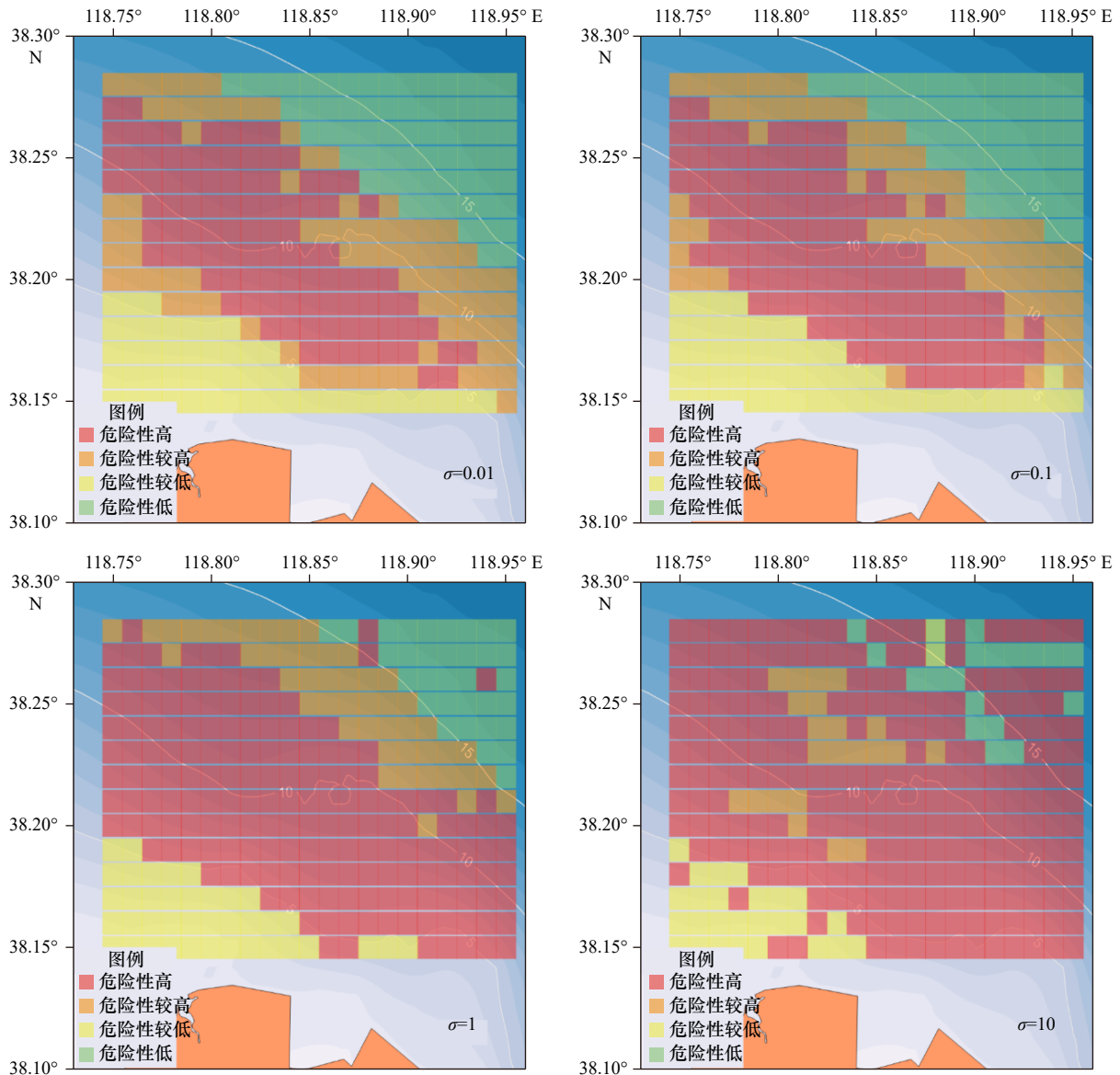


图7 不同核函数参数海底滑坡危险性评价结果

Fig. 7 Risk assessment results of submarine landslides with different kernel function parameters

输入因子类别在数据不足时可适当简化,需包含海底滑坡自身条件和触发因子两方面,评价结果与充足输

入因子结果相比精度降低。

参考文献:

- [1] Solheim A, Bryn P, Berg K, et al. Ormen Lange—An Integrated Study for Safe Field Development in the Storegga Submarine Slide Area[M]. San Diego: Elsevier, 2005.
- [2] Bea R G. How sea floor slides affect offshore structures[J]. Oil & Gas Journal, 1971, 69(48): 88–92.
- [3] Hsu S, Sibuet J. Flow of turbidity current viewed from failures of telecommunication cables[C]//International Conference on Seafloor Mapping for Geohazard Assessment. Ischia, Italy: American Geophysical Union, 2009.
- [4] Kawamura K, Laberg J S, Kanamatsu T. Potential tsunamigenic submarine landslides in active margins[J]. Marine Geology, 2014, 356: 44–49.
- [5] 刘锋. 南海北部陆坡天然气水合物分解引起的海底滑坡与环境风险评价[D]. 青岛: 中国科学院海洋研究所, 2010.
Liu Feng. A safety evaluation for submarine slope instability of the northern South China Sea due to gas hydrate dissociation[D]. Qingdao: Institute of Oceanology, Chinese Academy of Sciences, 2010.
- [6] 胡光海. 东海陆坡海底滑坡识别及致滑因素影响研究[D]. 青岛: 中国海洋大学, 2010.
Hu Guanghai. Identification of submarine landslides along the continental slope of the East China Sea and analysis of factors causing sub-

- marine landslides[D]. Qingdao: Ocean University of China, 2010.
- [7] Wang Weiwei, Wang Dawei, Wu Shiguo, et al. Submarine landslides on the north continental slope of the South China Sea[J]. *Journal of Ocean University of China*, 2018, 17(1): 83–100.
- [8] Ilstad T, De Blasio F V, Elverhøi A, et al. On the frontal dynamics and morphology of submarine debris flows[J]. *Marine Geology*, 2004, 213(1/4): 481–497.
- [9] El-Ramly H, Morgenstern N R, Cruden D M. Probabilistic slope stability analysis for practice[J]. *Canadian Geotechnical Journal*, 2002, 39(3): 665–683.
- [10] Griffiths D V, Lane P A. Slope stability analysis by finite elements[J]. *Géotechnique*, 1999, 49(3): 387–403.
- [11] Schwarz H U. Subaqueous Slope Failures: Experiments and Modern Occurrences[M]. Stuttgart: Schweizerbart'sche Verlagsbuchhandlung, 1982.
- [12] Bradshaw A S, Tappin D R, Rugg D. The kinematics of a debris avalanche on the Sumatra margin[M]//Submarine Mass Movements and Their Consequences. *Advances in Natural and Technological Hazards Research*, Vol. 28. Dordrecht: Springer, 2010: 117–125.
- [13] Schofield A N. Use of centrifugal model testing to assess slope stability[J]. *Canadian Geotechnical Journal*, 1978, 15(1): 14–31.
- [14] 吴益平, 滕伟福, 李亚伟. 灰色-神经网络模型在滑坡变形预测中的应用[J]. *岩石力学与工程学报*, 2007, 26(3): 632–636.
Wu Yiping, Teng Weifu, Li Yawei. Application of grey-neural network model to landslide deformation prediction[J]. *Chinese Journal of Rock Mechanics and Engineering*, 2007, 26(3): 632–636.
- [15] Marjanović M, Bajat B, Abolmasov B, et al. Machine learning and landslide assessment in a GIS environment[M]//GeoComputational Analysis and Modeling of Regional Systems. New York: Springer, 2018: 191–213.
- [16] Lun N K, Liew M S, Matori A N, et al. Recent developments in machine learning applications in landslide susceptibility mapping[J]. *AIP Conference Proceedings*, 2017, 1905: 040022.
- [17] Pham B T, Prakash I, Bui D T. Spatial prediction of landslides using a hybrid machine learning approach based on random subspace and classification and regression Trees[J]. *Geomorphology*, 2018, 303: 256–270.
- [18] 孙永福, 胡光海, 宋玉鹏, 等. 近海海底地质灾害预测评价及防控关键技术研究[R]. 青岛: 自然资源部第一海洋研究所, 2016.
Sun Yongfu, Hu Guanghai, Song Yupeng, et al. Study on the key technology of prediction, evaluation and prevention of offshore submarine geological hazards[R]. Qingdao: First Institute of Oceanography, Ministry of Natural Resources, 2016.
- [19] 杨作升, 陈卫民, 陈彰榕, 等. 黄河口水下滑坡体系[J]. *海洋与湖沼*, 1994, 25(6): 573–581.
Yang Zuosheng, Chen Weimin, Chen Zhangrong, et al. Subaqueous landslide system in the Huanghe River (Yellow River) delta[J]. *Oceanologia et Limnologia Sinica*, 1994, 25(6): 573–581.
- [20] Prior D B, Yang Z S, Bornhold B D, et al. The subaqueous delta of the modern Huanghe (Yellow River)[J]. *Geo-Marine Letters*, 1986, 6(2): 67–75.
- [21] 彭俊, 陈沈良, 陈一强, 等. 黄河三角洲侵蚀性岸段水下岸坡地质灾害及其空间分布[J]. *海洋通报*, 2014, 33(1): 1–6.
Peng Jun, Chen Shenliang, Chen Yiqiang, et al. Geological hazards and their spatial distribution in the subaqueous slope at the erosive coast of the Yellow River delta[J]. *Marine Science Bulletin*, 2014, 33(1): 1–6.
- [22] 杜星. 黄河口海底粉土波致孔压精细观测及液化评判方法[D]. 青岛: 国家海洋局第一海洋研究所, 2016.
Du Xing. Fine observation of wave-induced pore water pressure and liquefaction evaluation method of seabed silt in Yellow River Estuary[D]. Qingdao: The First Institute of Oceanography, State Oceanic Administration, 2016.

Risk assessment of submarine landslide based on spectral clustering

Du Xing^{1,2}, Sun Yongfu^{3,4}, Song Yupeng^{2,3}, Xiu Zongxiang^{2,3}, Shan Zhigang¹

(1. POWERCHINA Huadong Engineering Corporation Limited, Hangzhou 311122, China; 2. First Institute of Oceanography, Ministry of Natural Resources, Qingdao 266061, China; 3. Laboratory for Marine Geology, Pilot National Laboratory for Marine Science and Technology (Qingdao), Qingdao 266237, China; 4. National Deep Sea Center, Qingdao 266237, China)

Abstract: The risk assessment and zoning of submarine landslides can guide the site selection and risk prevention of offshore engineering facilities. In this paper, an unsupervised machine learning spectral analysis algorithm was used to evaluate the risk of submarine landslides in the Chengdao sea area of the Yellow River Estuary. A model of submarine landslides risk assessment with 9 input parameters, 4 output parameters and 0.08 as kernel function parameters is constructed. By using this model, the study area can be divided into 4 parts: high, quite high, quite low and low risk of submarine landslide. The comparison between the evaluation results and the distribution characterist-

ics of geological environment factors show that the most important factors are the type of seafloor sediment and hydrodynamic action, and the most important trigger factor is liquefaction. The analysis results of model parameters present that the evaluation results with slightly lower accuracy can be obtained by reasonably simplifying the input factors, and the kernel function parameter is important index affecting the evaluation accuracy. The above research shows that the unsupervised machine learning algorithm can be well used in the risk assessment of submarine landslides, and the richness and accuracy of data categories are important factors affecting the assessment accuracy.

Key words: submarine landslide; Yellow River Estuary; risk assessment; unsupervised machine learning; spectral clustering