

张加庆.基于大数据的地震损失价值评估模型设计[J].地震工程学报,2018,40(2):356-362.doi:10.3969/j.issn.1000-0844.2018.02.356

ZHANG Jiaqing.Design of Earthquake Loss Value Evaluation Model Based on Big Data[J].China Earthquake Engineering Journal,2018,40(2):356-362.doi:10.3969/j.issn.1000-0844.2018.02.356

基于大数据的地震损失价值评估模型设计^①

张加庆

(青海省地震局,青海 西宁 810001)

摘要:破坏性地震往往导致严重的经济损失及人员伤亡,对地震损失价值评价有助于震前找出抗震弱点,提高抗震能力,实现减轻地震灾害损失的目的。传统空间模型在地震数据处理过程中,无法处理大数据对空间尺度选择的干扰,存在地震损失评估结果偏差大以及波动性高的弊端。因此,在云计算平台下,提出基于大数据的地震损失价值评估模型设计,对模型 HAZ-China 大数据的服务层次、地震应用服务层以及 HAZ-China 大数据体系结构进行设计,为用户提供震前、震时以及震后的地震损失价值评估服务。模型采用 HBase 分布式数据库实现大数据的存储和分析,设计房屋震害数据库以及云计算模型,通过考虑大数据因素的地震灾害损失综合评估过程,实现地震损失价值的准确评估。实验结果说明,所设计模型可实现地震损失价值的准确评估,具有较高的评估精度和稳定性。

关键词:大数据;地震损失;价值;评估模型;云计算平台

中图分类号:P315.9

文献标志码:A

文章编号:1000-0844(2018)02-0356-07

DOI:10.3969/j.issn.1000-0844.2018.02.356

Design of Earthquake Loss Value Evaluation Model Based on Big Data

ZHANG Jiaqing

(Earthquake Administration of Qinghai Province, Xining 81000, Qinghai, China)

Abstract: Destructive earthquakes often lead to serious economic losses and casualties, so it is important to evaluate the value of earthquake loss, for it can help to improve the seismic capacity and reduce the loss of earthquake disasters. In seismic data processing, traditional spatial models cannot deal with the interference of big data on the spatial scale selection, and drawbacks of large deviation and high volatility of seismic loss assessment results must be faced. Therefore, this study designs a seismic loss assessment model based on the cloud computing platform, HAZ-China. The service levels, seismic application service layer, and system structure of the cloud computing platform, HAZ-China, were designed to provide pre-, co-, and post-seismic loss value evaluation services. The model can realize the storage and analysis of big data, then design the building damage database and cloud computing model using the distributed database, HBase. The model can accurately evaluate the seismic loss value through a comprehensive assessment of the seismic disaster loss. The experimental results showed that the designed model can realize an accurate evaluation of the value of earthquake loss, and had high convergence, efficiency, and stability.

Key words: big data; earthquake loss; value; evaluation model; cloud computing platform

① 收稿日期:2017-08-20

基金项目:青海省基础研究项目(2017-ZJ-775)

作者简介:张加庆(1969—),男,河南南阳人,工程师,研究方向:地震安评、地基检测及地震灾害损失评估。E-mail:Yangping258@126.com。

0 引言

当前我国大陆发生的大量灾害性地震对群众的生命和财产安全产生了巨大的威胁。人们开始关注地震应急工作的质量,其中震后经济损失的准确评估是一项重要内容,是地震应急工作有效开展的依据。大数据具有数据规模大、数据高速提升以及数据异构化的特性。采用大数据操作模式可实现地震损失数据的并行操作性能,获取具备洞察力以及决策力的信息资产。传统空间模型在地震数据处理过程中无法处理大数据对空间尺度选择的干扰,存在地震损失评估结果偏差大以及波动性高的弊端。朱永莉等^[3]将强震动记录引入应急性的生命损失快速评估,提出了基于强震动记录的地震生命损失快速评估方法,其建立了考虑建筑地震易损性并与烈度相关的地震生命损失快速评估函数,可较为准确地评估人员损伤情况,但该种评估方法局限性较大,未考虑非人员伤亡的地震损失,且未考察该方法的稳定性。曹彦波等^[4]提出云南地震应急快速评估模型本地化集成研究,实现了震后地震损失速判,并在云南的一次实际地震中,该方法的有效性得到了检验。但是该方法只适用于地震后的损失评估,其应用范围较小,无法为震前准备,降低地震损失提供理论依据。Field 等^[5]提出一种用于加利福尼亚的地震损失评估模型,其有效统计了该地区发生的地震及余震情况,具有一定的实用性。但其受到区域限制,未考虑该方法的大范围适用性。

为了解决该种问题,文章设计基于大数据的地震损失评估模型,在云计算平台下,提高对地震大数据的并行处理效率和精度,获取更为准确的地震损失评估结果。

1 基于大数据的地震损失价值评估模型设计

云计算的关键是虚拟化技术和分布式数据库技术,其运行载体是互联网,可向用户提供海量资源的共享模型,可为大数据环境下的地震灾害数据的处理和分析提供可靠的基础,进而实现地震损失价值的准确评估。因此,利用云计算平台,文章设计基于大数据的地震损失价值评估模型。通过云计算技术,充分分析地震损失大数据,可实现地震应急响应和辅助决策、地震现场损失评估、震后重建资金评估以及地震业务再次开发设置等功能。

1.1 HAZ-China 大数据的服务层次设计

HAZ-China 地震灾害损失评估模型分析的对象

是我国大陆地区,向不同的用户提供震前、震中以及震后的地震信息服务^[2]。HAZ-China 云计算范围分以下 4 个部分:地震应用功能 HAZ-Saa S、地震业务开发功能 HAZ-Paa S、地震数据功能 HAZ-Daa S 以及地震基础管理功能 HAZ-Iaa S,具体如图 1 所示。

地震应用功能层 HAZ-Saa S 通过对用户权限的区别划分,向不同使用者进行地震预警、震中应急指挥、震后扶助和地震科普等。只有在地震相关信息以及工作体系成熟完善的基础之上这些功能才能得以顺利实现,且应尽量避免多部门软硬件的重复设计,才有助于地震业务的正规化发展。地震业务开发功能 HAZ-Paa S 主要负责针对使用者制定相应的具体化细节服务,例如地震统计模型的变化、业务功能的多角度探索等。地震数据功能 HAZ-Daa S 则负责基础数据的获取和分析^[3]以及二次产品信息的获取。基础设施功能层 HAZ-Iaa S 则是以云计算平台的数据长处作为基础,针对用户或者单位进行不同种类的硬件租赁。

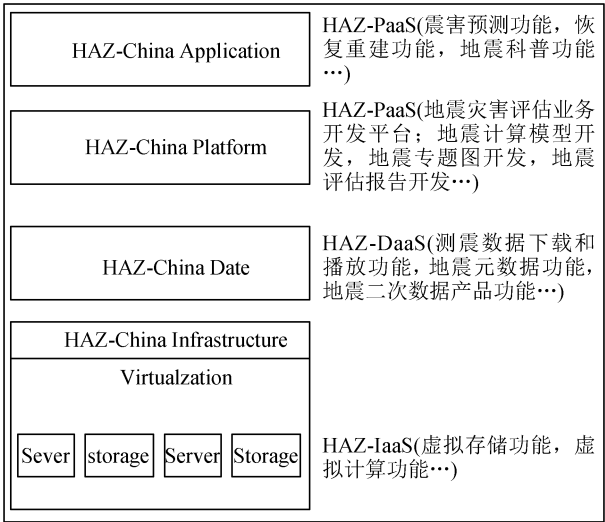


图 1 HAZ-China 的云计算服务层次图

Fig.1 Cloud computing service hierarchical graph of HAZ-China

1.2 地震应用服务层设计

用户通过向软件供货商以支付酬劳的形式借用软件,在 HAZ-China 系统的软件平台上发展所需业务的模式,被称作 HAZ-Saa S 地震应用服务层。分析图 2 可知,地震应用服务层 HAZ-Saa S 为专业用户和非专业用户在不同程度上提供基础数据管理服务等 9 种应用服务。由于综合了我国大部分领域内地震相关数据以及非常规范化的服务模式^[4],使得这一服务层具有两大突出优势,一是节省了相关部

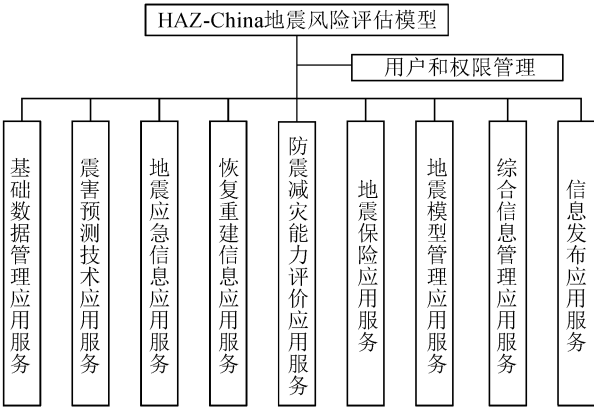


图 2 地震应用服务模块

Fig.2 Seismic application service module

门购买和开发关于地震风险评估软件的开支；二是完善了地震风险相关的评估体系。若该层应用服务无法为用户提供服务，则用户需要通过地震业务开发平台服务层 HAZPaaS 实施二次规划，对自身的应用进行设置^[5]，同时在地震数据服务层 HAZ-DaaS 塑造私有数据库。

1.3 HAZ-China 大数据体系结构设计

SOA 构建层、管理中间件层、资源池层以及物理资源层组成了 HAZ-China 云计算平台体系（图 3）。这四层之间的关联度小，一层状况坍塌并不会对其他层造成致命伤害，同时又各司其职，形成系统、科学的运转状态。



图 3 HAZ-China 云计算平台体系结构

Fig.3 System structure of HAZ-China cloud computing platform

1.4 HBase 分布式数据库运用

HBase 即非关系型数据库，其具有高效率、可伸缩等特点，是 Google 的 BigTable 的开源实现，可聚集众多低成本的电脑通过 HDFS 分布式文件系统进行文件保存，形成巨大的无线数据保存体系。本文模型采用 HBase 分布式数据库实现地震大数据的分析和存储，该数据库的优点是：（1）列的可扩展性。大部分情况下，应对列簇、行键实现进行设定，基于无数的列构成了列簇，所以之后对其实施动态分布即可。（2）实现高性能的并发读写操作。仅

凭 HDFS 文件模式难以实现以上功能。（3）能够对数据实现切分。HBase 能够智能地切分到达上线的数据，使得 HBase 的运行效率明显增强。（4）高可靠性。多而复杂的数据被反复复制会留下庞大的数据节点，使得存储的难度大大增加，然而 HBase 避免了因节点的中断导致数据丢失现象^[6]。

1.4.1 房屋震害数据库设计

大量数据根据键-值的模式排列在表格之内，表格又保存于 HBase 内，因此数据以这种形式保存在 HBase 内。基于 GBT18208.4-2011，采用以下公式

可计算出地震房屋破损的程度^[7]:

$$L = \sum_{i=1}^m \sum_{s=1}^n \sum_{j=1}^5 (S_{is} \times R_{isj} \times D_{isj} \times P_{is}) \quad (1)$$

其中: i 评估区 s 类房屋建筑面积、房屋 j 级破坏的破坏比、 j 级破坏的损失比以及房屋重置单价分别用 S_{is} 、 R_{isj} 、 D_{isj} 以及 P_{is} 来描述。

式(1) 直观地显示出, S_{is} 、 R_{isj} 、 D_{isj} 以及 P_{is} 是计算区域中房屋财产损失的必备因素。基于 HBase 按行输出的规律,本次规划一定要为日后的查找与修改提供极大便利。定义行键 Num 来描述房屋的代码,定义列簇 Info,这里保存部分重要内容, S 、 R 、 D 以及 P 是前面四种数据的缩写,将其划分在 Info 的下属行,出于对这些原因的考虑,绘制 HBase 数据表。通过输送大量数据到 HBase 表格的方式,使得逾越上限时被智能地分散成众多个 splits,形成 regions,为将来 Mapreduce 程序提供完善的数据储备^[8]。

1.4.2 云计算平台下基于大数据的操作系统运模

为使程序可在数以千计的计算机内同时操作,需应用 Mapreduce 编程模型进行编程。函数 Mapper 以及 Reducer 构成了 Mapreduce 编程模型的中心主体,由图 4 进行描述。通过拆散重型文件使其变成无数个小数据,并在各个节点分析这些小数据,即通过这些节点的 Map 模块把待处理的数据转换成键-值的模式^[9]。Reduce 模块得到这些键-值形式的数据之后再对其进行处理,把具备同样 Key 值特征的 Value 值进行合并,进而输出最后结果。这种多台计算机同时协作运行的做法远远超过以往单纯在一台计算机上操作数据的形式,即云计算的优点所在,使得数据运算的速度发生了翻天覆地的变化。

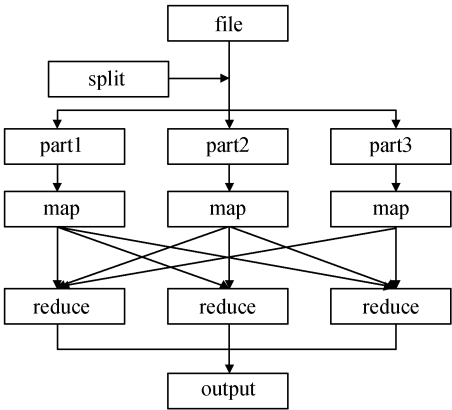


图 4 Mapreduce 工作原理图

Fig.4 Mapreduce working principle diagram

1.4.3 地震灾害损失综合评估

本文模型通过考虑大数据因素下的地震灾害损失综合评估过程实现地震损失价值的准确评估。时间、位置以及地震强度分别用 t 、 $E(x_e, y_e)$ 以及 M_s 来表示。 $D(x_d, y_d)$ 表示待分析的灾区, $ds = dx dy$ 是其中一个具体区域,经历地震后极易造成灾害损失。下列公式是椭圆形衰减规律,假设地震的各项数据符合此公式:

$$Y_a(M_s, R_a) = f_a(M_s, R_a) + \epsilon_a \quad (2)$$

$$Y_b(M_s, R_b) = f_b(M_s, R_b) + \epsilon_b \quad (3)$$

式中:椭圆的长、短半轴分别用 R_a 和 R_b 来描述, ϵ_a 和 ϵ_b 表示符合上升排列的任意值。在地震 Y 的作用下, D 点遭受的灾害用 $L(D)$ 表示,那么 L 表示震中 E 点造成的财产损失,用下式进行详细描述:

$$L_E = \int L_E(D) ds \quad (4)$$

地震的严重程度决定了其灾害程度,因为存在一部分不确定因素,导致对 D 点的震动程度也无法做出准确判断^[10],用 $P(y=Y)$ 来表示发生震动值 Y 的概率,这种情况下地震受到的灾害可用下式来描述:

$$L_E = \iint L_E(D | Y) P(y=Y) ds dy \quad (5)$$

地震损失评估模型应在地震形成后的短时间内进行,而受到大数据环境的不确定性干扰,地震速报数据同设备定位的参数存在一定的偏差^[11],同地震的宏观震中也存在较高的偏差,导致地震损失评估结果的准确度降低。为提高地震损失评估精度,应全面分析该种大数据产生的不确定性。

为了快速估测出地震发生时所带来的损失,考虑重点应放在速报地震参数和宏观地震参数的不同中,针对可以判定的速报地震震中,现实宏观震中的地点可设置为到达概率模型的任意变量^[12]。相对于速报地震震中设为 $E(x_e, y_e)$,宏观地震震中发生在空间 $G(x_m, y_m)$ 点的概率密度函数是 $g(x_{me}, y_{me})$,此间 $x_{me} = x_m - x_e, y_{me} = y_m - y_e$ 。那么地震损失价值为 L_e 代表宏观地震震中以及速报震中全部相同情况下的损失价值,所形成的概率密度是 $g(0,0)$ 。则针对灾区全部损失可写为:

$$L = \iint L_{Gg}(x_{me}, y_{me}) dx_{me} dy_{me} \quad (6)$$

综合式(4)(G 点等于 E 点),得出:

$$L = \iiint L_G(D | Y) P(y=Y) g(X_{me}, y_{me}) ds dy dx_{me} dy_{me} \quad (7)$$

因为大数据环境下的地震震级以及地震动减缓具有未知性,应根据设置假设范围内的置信度,通过极限状态实施预判^[13]。针对易损性判断和判断估计的未知性,延伸到详细的判断目标(包括直接损失、次生灾害损失等),做到有目的的纠正预判。所以式(7)是一个在大数据环境下针对地震灾害未知性的损害综合预判模型^[14],其可实现地震损失价值的准确评估。

2 实验分析

2.1 本文模型的训练和仿真分析

实验分析 2012 年 6 月 24 日宁蒗—盐源 5.7 级地震,采用本文模型对该地震直接经济损失实施评估。该地震的震源深度是 11 km,受灾人口和面积分别是 11.62 万人以及 2 218 km²,包括永登、盐源以及当雄地区。该地区的设防加速度是 0.15g,2011 年地震地区的人均 GDP 是 7 336 元,产业结构比是 57.6%,受灾面积是 1 278 km²。基于这些因素,通过本文设计的地震损失价值评估模型实施仿真,获取本文模型的实验结果并列于表 1。

表 1 实验检测数据统计结果

Table 1 Statistical results of experimental data	永登	盐津	当雄
实际经济损失/万元	6 668	22 818	40 025
模型运算结果/万元	6 457	22 503	39 865

为了更加直观地分析本文模型的预测结果同真实地震经济损失的对比情况,基于表 2 绘制不同地区的对比柱形图(图 5)。分析该图能够直观看出本文模型对不同地区地震直接损失的运算结果同该地

表 2 测试样本输出的相对误差

Table 2 Relative error of test sample output	永登	盐津	当雄
相对误差	9.65%	8.62%	0.29%
测试样本实际输出	10 125.66	2 874.36	5 151.63
测试样本目标结果	9 185	2 638	5 176

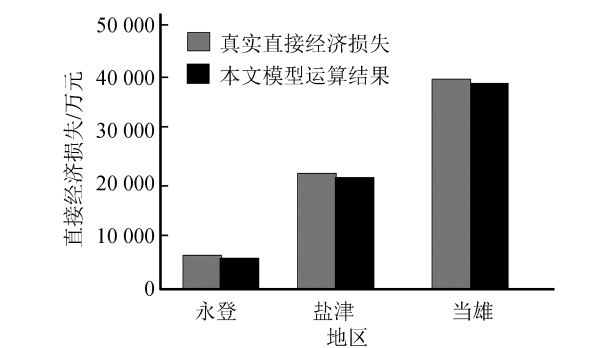


图 5 检测数据对比图

Fig.5 Contrast diagram of test data

真实的直接经济损失基本一致,说明本文设计的基于大数据的地震损失价值评估模型是有效的,实现了地震损失的准确预测。

本文模型的训练样本、测试样本以及确认样本的误差曲线用图 6 描述。分析该图能够得出,训练开始阶段 3 种样本的误差曲线都不断降低,说明本文模型的泛化性能较高。实验设置样本误差曲线连续 5 步提升则不再进行仿真,避免本文模型出现过拟合问题。

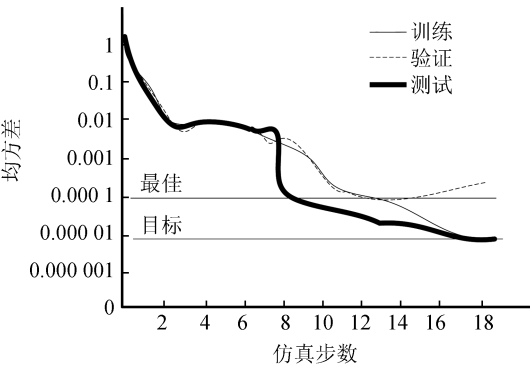


图 6 不同样本的误差曲线

Fig.6 Error curves of different samples

2.2 对比分析结果

实验基于 Matlab 2010 仿真软件对不同地震损失价值评估模型进行验证,采用的数据来自于国家地震网中的大数据集,其中 50 组训练数据以及 20 组预测数据集构成本次实验的两类数据。本次实验根据训练时间、训练精度以及预测精度三要素将本文模型和传统空间模型的实验数据进行对比,其精度用均方根误差来衡量^[15],得出本文模型的实用性更胜一筹。实验统计了不同数据量情况下两种模型进行地震损失价值评估的结果(表 3)。

分析表 3 可得,本文模型花费的训练时间不到传统空间模型的 1%,其效率远远超过空间模型,完成了大数据环境下地震灾害损失的高效率预测。

实验对比分析了本文模型和空间模型在不同数据量下的训练时间、训练精度以及预测精度情况(图 7)。

图 7(a)中两种模型花费的训练时间差距更为明显,本文模型的训练时间远远低于空间模型,说明本文模型具有较好的大数据处理性能。分析图 7(b)、(c)能够看出,随着数据量的不断提升,本文模型的训练精度以及预测精度都高于空间模型,本文模型具有较高评估精度;且本文模型的训练精度和预测精度变化更为平稳,空间模型的训练精度和预

表 3 不同模型的实验结果

Table 3 Experimental results of different models

数据集数据 数量/kB	本文模型			空间模型		
	训练时间/s	训练精度/%	预测精度/%	训练时间/s	训练精度/%	预测精度/%
5	0.285 3	0.110 4	0.077 5	48.486 1	0.812	0.538
10	0.347 7	0.130 2	0.070 8	44.176 1	0.112	0.48
15	0.285 3	0.308 6	0.062 1	46.087 2	0.537	0.137
20	0.254 1	0.102 1	0.067 2	45.087 2	0.335	0.713
25	0.238 5	0.101 8	0.064 6	43.121 6	0.137	0.42
30	0.207 3	0.154 2	0.061 3	45.673 6	0.28	0.137
35	0.301	0.320 3	0.056 8	46.172 8	0.168	0.183
40	0.270 7	0.170 7	0.053 6	44.404 4	0.35	0.405
45	0.316 5	0.181	0.050 2	36.703 5	0.118	0.15
50	0.316 5	0.107 8	0.034 8	46.126	0.763	0.873
55	0.435	0.163 6	0.011	44.660 6	0.537	0.375

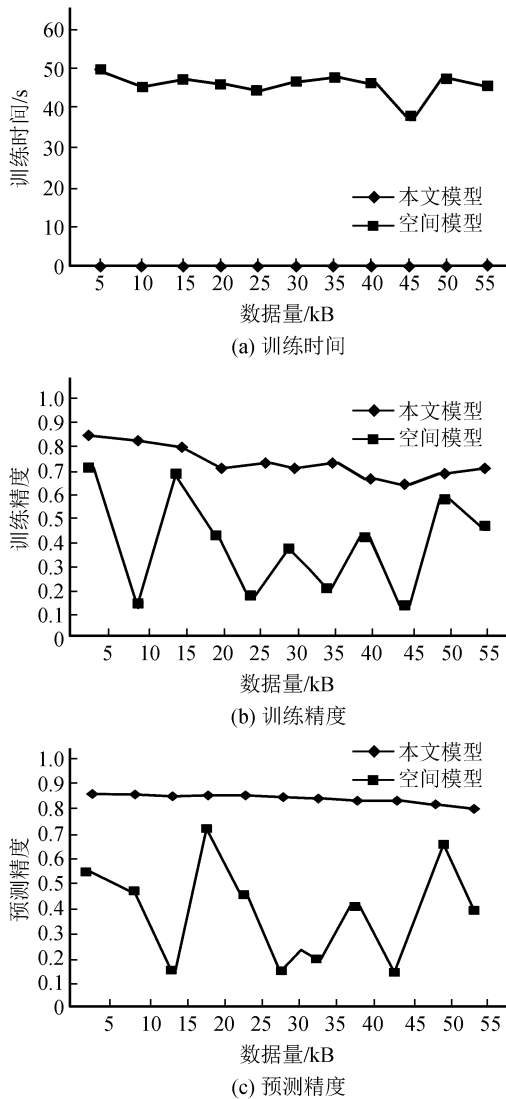


图 7 不同模型的训练时间、训练精度及预测精度对比结果

Fig.7 Comparison between training time,training accuracy, and prediction accuracy different models

测精度随着数据量的逐渐增加,出现了大幅度的波动趋势,说明本文模型具有较高的稳定性。

综合分析上述实验结果可得,相对于空间模型,本文模型对大数据的处理性能更强,能够实现大数据环境下地震损失价值的快速、准确、稳定评估,具有较高的应用价值。

3 结论

本文提出云计算平台下基于大数据的地震损失价值评估模型设计,研究过程中,为了提高地震损失评估精度,全面分析地震大数据产生的不确定性,实现充分分析地震损失数据的基础上,高效率处理地震损失数据,获取更为精准的地震损失评估结果,并且提高了评估结果的稳定性,有效解决了传统方法中存在的评估结果偏差较大且结果稳定性差、波动性高的缺点。

参考文献(References)

[1] 张文路,蒋欢军.地震人员伤亡评估方法与模型研究综述[J].结构工程师,2016,32(3):181-191.
ZHANG Wenlu,JIANG Huanjun.A Review of Methods and Models on Seismic Casualty Estimation[J].Structural Engineers,2016,32(3):181-191.

[2] RONG M,SU M,WANG H.Calculation Method of Quantitative Evaluation Weight Value ω of Railway Tunnel Earthquake Damage Loss Based on the Analytic Hierarchy Process[J].Electronic Journal of Geotechnical Engineering,2015,20(16):6893-6902.

[3] 朱永莉,黎大虎,朱建钢.基于强震动记录的地震生命损失快速评估方法[J].地震学报,2017,39(1):143-154.
ZHU Yongli,LI Dahu,ZHU Jiangang.Rapid Evaluation Method of Life Loss in Earthquakes Based on Strong Ground Motion Records[J].Acta Seismologica Sinica,2017,39(1):143-

- 154.
- [4] 曹彦波,李兆隆,李永强,等.云南地震应急快速评估模型本地化集成研究[J].地震研究,2015,38(1):148-154.
CAO Yanbo,LI Zhaolong,LI Yongqiang,et al.Research on the Localization Integration of Yunnan Earthquake Emergency Rapid Assessment Model [J]. Journal of Seismological Research,2015,38(1):148-154.
- [5] FIELD E, PORTER K, MILNER K,et al.A Prototype Operational Earthquake Loss Model for California Based on UCERF3-ETAS-A First Look at Valuation [J]. Earthquake Spectra,2017,33(2):117-122.
- [6] WALKER B B, TAYLOR-NOONAN C, TABBERNOR A, et al.A Multi-criteria Evaluation Model of Earthquake Vulnerability in Victoria, British Columbia[J]. Natural Hazards, 2014, 74(2):1209-1222.
- [7] 孙君,谭清美.基于 GF-AHP 组合模型的应急物流能力评估——以地震灾害为例[J].北京工业大学学报,2014,40(9):1354-1360.
SUN Jun, TAN Qingmei. Evaluation of Emergency Logistics Capabilities Based on GF-AHP Combined Models——Taking the Earthquake Disaster as an Example[J]. Journal of Beijing University of Technology, 2014, 40(9):1354-1360.
- [8] 宋平,聂高众,邓视,等.基于现场调查的地震灾害损失预评估——以云南省德宏傣族景颇族自治州为例[J].地震地质,2016,38(4):1148-1159.
SONG Ping, NIE Gaozhong, DENG Yan, et al. Pre-assessment of Earthquake-induced Losses Based on Field Survey and Kilometer Grid Dataset: A Case Study from Dehong Dai-Jingpo Autonomous Prefecture[J]. Seismology and Geology, 2016, 38(4):1148-1159.
- [9] YANG X, XIE Z, LING F, et al. Post-Earthquake People Loss Evaluation Based on Seismic Multi-level Hybrid Grid: A Case Study on Yushu $M_s7.1$ Earthquake in China[J]. Open Geosciences, 2016, 8(1):639-649.
- [10] 于萍萍,陈建平,柴福山,等.基于地质大数据理念的模型驱动矿产资源定量预测[J].地质通报,2015,34(7):1333-1343.
YU Pingping, CHEN Jianping, CHAI Fushan, et al. Research on Model-driven Quantitative Prediction and Evaluation of Mineral Resources Based on Geological Big Data Concept[J]. Geological Bulletin of China, 2015, 34(7):1333-1343.
- [11] 郭凯,黄金刚,彭克银,等.大数据技术在海量测震数据中的应用[J].地震研究,2017,40(2):317-323.
GUO Kai, HUANG Jingang, PENG Keyin, et al. Research and Application of Big Data Technology in Mass Seismic Data[J]. Journal of Seismological Research, 2017, 40(2):317-323.
- [12] 王草,包琦玮,李全旺.简支混凝土梁桥 Pushover 分析简化方法[J].震灾防御技术,2014,9(3):439-446.
WANG Cao, BAO Qiwei, LI Quanwang. Simplified Pushover Analysis Method for Simply Supported Concrete Bridges[J]. Technology for Earthquake Disaster Prevention, 2014, 9(3):439-446.
- [13] BERNARDINI G, QUAGLIARINI E, D'Orazio M. Towards Creating a Combined Database for Earthquake Pedestrians' Evacuation Models[J]. Safety Science, 2016, 82:77-94.
- [14] 李芬花,卢东海,初文婷,等.基于多层次模糊综合评价的重力坝风险评估模型研究[J].水利水电技术,2016,47(2):92-96.
LI Fenhua, LU Donghai, CHU Wenting, et al. Multi-hierarchy Fuzzy Comprehensive Evaluation Method-based Study on Risk Assessment Model For Gravity Dam [J]. Water Resources and Hydropower Engineering, 2016, 47(2):92-96.
- [15] 李文倩,佟大威,王振,等.大型地下洞室群相邻块体地震响应[J].河海大学学报(自然科学版),2016,44(3):196-202.
LI Wenqian, TONG Dawei, WANG Zhen, et al. Seismic Responses of Adjacent Blocks in Large-scale Underground Cavern Group [J]. Journal of Hohai University (Natural Sciences), 2016, 44(3):196-202.