

吴凡, 阙志萍. 基于多分类 logistic 模型的铁路水害分级警戒概率预报研究[J]. 灾害学, 2017, 32(1): 17-21. [WU Fan and Que Zhiping. Graded alerting probability forecast of railway water disaster based on the multinomial logistic regression model [J]. Journal of Catastrophology, 2017, 32(1): 17-21. doi: 10.3969/j.issn.1000-811X.2017.01.004.]

# 基于多分类 logistic 模型的铁路水害分级警戒概率预报研究\*

吴凡, 阙志萍

(江西省气象服务中心, 江西 南昌 330046)

**摘要:** 利用 2004-2014 年鹰厦线的铁路水害资料和降水资料, 分析了鹰厦铁路水害的时空分布特征, 研究了两种地形的铁路水害分级警戒概率预报模型。结果表明: 鹰厦铁路水害的年际变化差异大且分布不均, 发生时间集中在汛期, 局地性较强; 山脉地形铁路水害分级警戒概率预报模型检验的出巡警戒、限速警戒以及封锁警戒的准确率分别为 78.3%、58.8% 以及 74.4%; 相对平坦地形铁路水害分级警戒概率预报模型检验的出巡警戒、限速警戒以及封锁警戒的准确率分别为 85.0%、75.8% 以及 87.7%; 在铁路防洪调度中结合水害概率预报模型以及 3 个因子的降水量值共同判断警戒级别的效果更优, 可为铁路安全运营和高效调度提供一定的指导。

**关键词:** 鹰厦铁路; 铁路水害; 分级警戒; 降水; 多分类 logistic 回归

**中图分类号:** X43; P49 **文献标志码:** A **文章编号:** 1000-811X(2017)01-0017-05

doi: 10.3969/j.issn.1000-811X.2017.01.004

随着铁路的高速发展, 各类灾害性天气对铁路运输安全造成的威胁日益凸显<sup>[1-3]</sup>。由于山洪或排水不良汇水水流能量, 降雨入渗或降雨与周围地质环境作用, 直接或由其诱发的次生灾害作用于铁路路基轨道、桥梁、隧道、涵洞等基础设施, 并造成段道、限速事故的灾害, 称为铁路水害<sup>[4]</sup>。有关铁路水害的研究较多, 李晓霞等<sup>[5]</sup>统计分析了三兰铁路局铁路水害时空分布特征及与气象条件的关系, 建立了铁路水害预警模型; 接小峰等<sup>[6]</sup>提出了基于 E-1 阈值曲线的防洪区间阈值预测方法; 杨思全等<sup>[7-8]</sup>在分析铁路成灾机理和致灾因子的基础上构建了铁路环境水害危险度计算模型。现阶段铁路部门针对水害影响运输安全的警戒工作主要是出巡警戒、限速警戒、封锁警戒, 然而水害分级警戒概率预报的相关研究较少。本文基于铁路沿线的水害信息资料和降水量资料, 利用多分类 logistic 回归方法, 研究铁路水害分级警戒的概率预报模型, 以期能为铁路部门高效调度、安全运营提供一定的参考。

## 1 资料与方法

### 1.1 研究区概况

鹰厦铁路是中国东南沿海重要的铁路干线之

一, 北起江西省鹰潭市, 南至福建省厦门市, 全长 697.7 km, 沿线包括 93 个铁路工务段编组站(图 1)。根据南昌铁路局辖区内路段的气候特征和地理环境, 将鹰厦线进一步细化为 9 个区段, 分别为鹰潭-肖家、肖家-资溪、资溪-邵武、邵武-吉舟、吉舟-青州、青州-卓宅、卓宅-梅水坑、梅水坑-龙海和龙海-厦门区段, 其中大部分区段受武夷山脉地形影响, 地形地貌复杂, 多为土质埕坡, 在遇强降水过程时雨水冲刷坡面, 容易导致水害的发生, 少数区段地势平坦, 且靠近城市, 排水系统更好, 水害频率相对较少。考虑到样本量的影响, 本文分别选取青州-卓宅区段、卓宅-龙海区段为代表研究山脉地形、相对平坦地形的铁路水害分级警戒概率预报模型。

### 1.2 资料来源

2004-2014 年的铁路水害的灾情信息来自南昌铁路局, 主要包括了 2004-2014 年的铁路水害发生日期、发生时间、水害类型、是否影响行车等; 2008-2014 年的降水量数据来自铁路工务段雨量日报表及气象站, 主要包括了逐日降水量、日 10 min 最大降水量和日小时最大降水量。降水量检测仪器置于铁路工务段观测点附近, 能较好地反映铁路沿线的降水情况。

\* 收稿日期: 2016-07-21 修回日期: 2016-09-03

基金项目: 江西省科技厅项目(20142BBG70035); 南昌铁路局科委项目“局管内水害临界雨量及评估模型研究”(201411)  
第一作者简介: 吴凡(1988-), 男, 江西南城人, 硕士, 主要从事铁路气象服务研究. E-mail: wufan\_1988@163.com

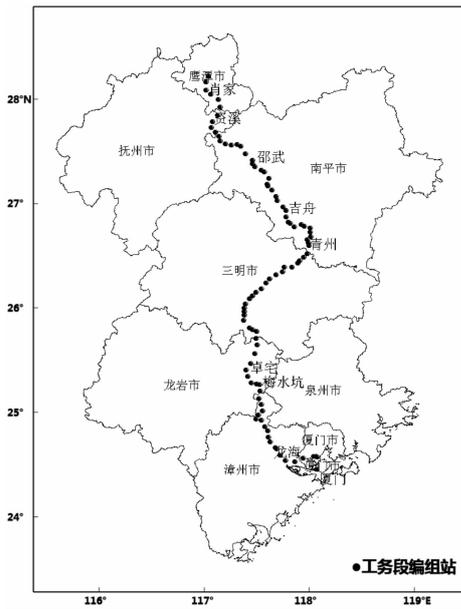


图1 鹰厦铁路沿线区段划分示意图

### 1.3 研究方法

通常意义上的 Logistic 回归要求因变量  $y$  只有两种取值, 即常用的二分类 Logistic 模型, 当因变量  $y$  取值有两种以上时, 就要用多分类 Logistic 回归分析<sup>[9-10]</sup>。如果因变量为  $k$  个有序分类变量, 可设  $y$  的  $k$  个取值按等级顺序为 1、2、...、 $k$ 。对  $y$  取值小于等于等级 ( $j=1、2、\dots、k$ ) 的概率即取前  $j$  个值得累积概率用来表示, 则有:

$$p(y=j) = \left[ \begin{array}{l} p_1 = p(y \leq 1); \\ p_j - p_{j-1} = p(y \leq j) - p(y \leq j-1); \\ p_k = 1 - p_{k-1} = 1 - p(y \leq k-1)。 \end{array} \right] \quad (1)$$

$$\ln \frac{p_j}{1-p_j} = \beta_{0j} + \sum_{i=1}^m \beta_i x_i \quad (2)$$

$$p_j = \frac{\exp(\beta_{0j} + \sum_{i=1}^m \beta_i x_i)}{1 + \exp(\beta_{0j} + \sum_{i=1}^m \beta_i x_i)} \quad (3)$$

## 2 结果分析

### 2.1 铁路水害时空分布特征

2004 - 2014 年鹰厦线共出现铁路水害 1 008 例, 水害发生频次分布不均, 差异较大, 其年际变化如图 2 所示, 可以看出 2006 年、2010 年和 2012 年鹰厦线的铁路水害发生较严重, 尤其是 2010 年铁路水害发生频次最多, 达 360 次。铁路水害的发生除了明显的年际变化还有明显的年内变化, 从图 3 还可以发现铁路水害发生的时间段与铁路沿线区域的汛期时间相对应, 主要出现在 4 - 8 月, 其中 6 月铁路水害发生频次最多。利用 2010

年青州 - 卓宅区段内的 3 个铁路气象观测站的降水资料, 分析其月降水量的分布如图 3 所示, 发现降水主要集中在 4 - 6 月, 且 6 月降水量最大, 与铁路水害分布特征一致, 可见降水是诱发铁路沿线水害的主要原因之一。

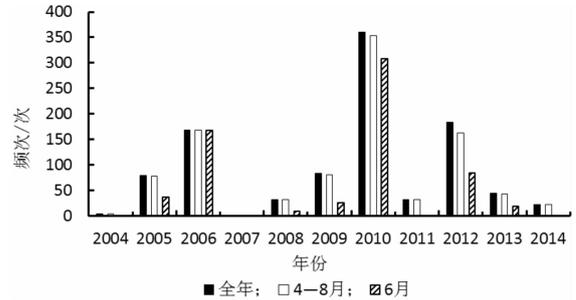


图2 2004 - 2014 年鹰厦铁路沿线水害发生频次分布

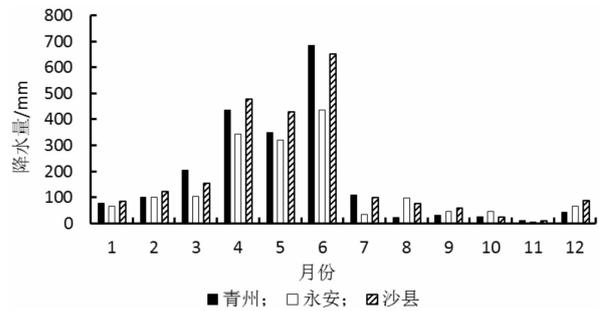


图3 2010 年青州 - 卓宅区段 3 个气象站月降水量分布

鹰厦线的铁路水害除了时间分布上的明显差异, 还有明显的空间分布特征。2004 - 2014 年鹰厦铁路沿线 9 个区段内的水害发生频次如图 4, 可以看出鹰潭 - 肖家、肖家 - 资溪和龙海 - 厦门 3 个区段的铁路水害发生频次最少, 分别为 3 次、15 次和 23 次; 吉舟 - 青州、青州 - 卓宅和梅水坑 - 龙海区段的铁路水害发生频次最多, 分别为 155 次、446 次和 216 次; 这一方面与区段划分的距离有关系, 另一方面是由于受武夷山脉地形的影响, 地形陡峭, 地表坡度较大且多为土质陡坡, 遇到强降水过程时雨水冲刷坡面, 坡面未及时做好防护加固, 导致水害发生, 从而影响铁路正常的运营。总体而言, 鹰厦线的铁路水害年际变化差异大且分布不均, 发生时间集中在汛期且局地性较强。

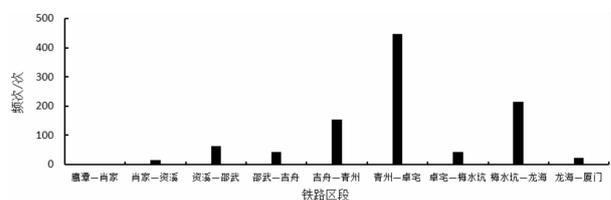


图4 2004 - 2014 年鹰厦铁路沿线各区段铁路水害发生频次

## 2.2 铁路水害分级警戒概率预报模型

### 2.2.1 铁路水害与降水因子的相关分析

收集 2008 - 2014 年鹰厦铁路沿线的水害资料, 在对同期 1 h 最大降水量、24 h 降水量、连续降水量资料筛选和剔除等质量控制的基础上, 分析铁路水害影响行车次数与同期降水因子的相关性(表 1), 由表 1 可见, 铁路水害发生前 1 h 最大降水量、水害发生前 24 h 降水量、连续降水量与铁路水害影响列车运行均呈正相关关系, 相关系数分别为 0.516、0.575 和 0.543, 相关性较好且均通过了显著性检验, 说明这几个气象因子是铁路水害影响行车的主要降水因子。铁路水害发生前 1 h 最大降水量主要是考虑短时降水强度较大, 易造成土体被雨水冲空流去, 造成路肩坍塌, 诱发泥石流灾害等; 铁路水害发生前 24 h 降水量主要是考虑强降水对床道的冲刷, 影响路基稳定, 或因岩土松动造成倒树侵线, 甚至汇水冲刷造成泥石流灾害等; 连续降水量主要是考虑连续降水造成的降水量积累, 导致铁路水害点的路基长期受雨水冲刷、浸泡, 使土质松软, 土体含水饱和失衡, 抗滑力下降, 发生溜坍或坍塌等。

表 1 2008 - 2014 年鹰厦铁路线降水因子与铁路水害相关系数

	1h 最大降水量	24h 降水量	连续降水量
铁路水害影响行车	0.516*	0.575*	0.543*

注: \*  $p \leq 0.01$

### 2.2.2 青州 - 卓宅区段水害警戒 logistic 回归模型

以青州 - 卓宅区段作为山脉地形铁路的代表, 选取 2008 - 2014 年该区段铁路水害对应时期的 1 h 最大降水量( $x_1$ )、24 h 降水量( $x_2$ )以及连续降水量( $x_3$ )作为自变量因子构建铁路水害分级警戒概率模型如下所示, 其中铁路水害的警戒级别分为 I、II、III 三个等级, 分别对应封锁警戒、限速警戒、出巡警戒。

从表 2 中可以看出一共有两组关于 Logistic 模型的回归系数, 24 h 降水量和连续降水量都通过显著性检验, 1 h 最大降水量的显著性检验值为 0.068 1, 通过显著性水平 0.1 的检验。表 3 是用于检验模型的有效性, 由于显著性水平值  $< 0.000 1$ , 所以该模型整体是有效的。所得到的 logistic 回归模型为:

$$p_1 = p(y = 1 \text{ 等级}) = p(\text{封锁}) = \frac{\exp(-5.4454 + 0.0263x_1 + 0.0209x_2 + 0.015x_3)}{1 + \exp(-5.4454 + 0.0263x_1 + 0.0209x_2 + 0.015x_3)}; \quad (4)$$

$$p_2 = p(y \leq 2 \text{ 等级}) = p(\text{封锁或限速}) =$$

$$\frac{\exp(-3.1185 + 0.0263x_1 + 0.0209x_2 + 0.015x_3)}{1 + \exp(-3.1185 + 0.0263x_1 + 0.0209x_2 + 0.015x_3)} \quad (5)$$

即可以获得不同级别警戒概率预报模型:

$$\text{封锁警戒: } p(\text{I}) = p_1. \quad (6)$$

$$\text{限速警戒: } p(\text{II}) = p_2 - p_1. \quad (7)$$

$$\text{出巡警戒: } p(\text{III}) = 1 - p_2. \quad (8)$$

表 2 模型参数部分的参数估计

参数	df	参数估计值	标准差	卡方	显著水平
$\beta_{01}$	1	-5.4454	0.4781	129.707	$< 0.0001$
$\beta_{02}$	1	-3.1185	0.3667	72.3142	$< 0.0001$
$x_1$	1	0.0263	0.0144	3.3272	0.0681
$x_2$	1	0.0209	0.00516	16.4284	$< 0.0001$
$x_3$	1	0.0150	0.00190	62.2685	$< 0.0001$

表 3 模型的有效性检验

模型	模型拟合标准		似然比检验	
	-2 倍对数似然值	卡方	df	显著水平
仅截距	650.489			
最终	449.388	57.8656	3	$< 0.0001$

根据上述方程进行回报检验, 检验结果如表 4 所示, 可以看出出巡警戒预报与封锁警戒预报的效果都较好, 准确率分别为 78.3% 与 74.4%, 限速警戒预报的效果较差, 准确率为 58.8%。分析这三种警戒预报的错报、漏报相对应的降水量发现, 大部分是由于 1 h 最大降水量、24 h 降水量以及连续降水量三者中的 1 个或 2 个因子的值很小, 但是另外 1 个因子的值较大, 因此采用 Logistic 回归模型预测警戒级别的概率结合 3 个因子的降水量值共同判断警戒级别的效果更优。

表 4 青州 - 卓宅区段水害警戒概率预报模型检验结果

出巡预报/%		限速预报/%		封锁预报/%	
准确率	78.3	准确率	58.8	准确率	74.4
错报限速	18.6	限速报封锁	17.1	封锁报限速	25.6
错报封锁	3.1	限速报出巡	24.1	封锁报出巡	0.0

### 2.2.3 卓宅 - 龙海区段水害警戒 logistic 回归模型

以卓宅 - 龙海段铁路作为相对平坦地形铁路的代表, 选取 2008 - 2014 年该区段铁路水害对应时期的 1 h 最大降水量( $x_1$ )、24 h 降水量( $x_2$ )以及连续降水量( $x_3$ )作为自变量因子构建铁路水害分级警戒概率模型如下所示, 其中铁路水害的警戒级别分为 I、II、III 三个等级, 分别对应封锁警戒、限速警戒、出巡警戒。

从表 5 中可以看出一共有两组关于 Logistic 模型的回归系数, 1 h 最大降水量、24 h 降水量以及连续降水量都通过显著性水平 0.01 的检验, 表 6 是用于检验最终方程的有效性, 由于显著性水平

值  $< 0.0001$ ，所以该模型整体是有效的。所得到的 logistic 回归模型为：

$$p_1 = p(y=1 \text{ 等级}) = p(\text{封锁}) = \frac{\exp(-10.0242 + 0.1594x_1 + 0.0226x_2 + 0.0099x_3)}{1 + \exp(-10.0242 + 0.1594x_1 + 0.0226x_2 + 0.0099x_3)}; \quad (9)$$

$$p_2 = p(y \leq 2 \text{ 等级}) = p(\text{封锁或限速}) = \frac{\exp(-6.0922 + 0.1594x_1 + 0.0226x_2 + 0.0099x_3)}{1 + \exp(-6.0922 + 0.1594x_1 + 0.0226x_2 + 0.0099x_3)}. \quad (10)$$

即可以获得不同级别警戒概率预报模型：

$$\text{封锁警戒：} p(\text{I}) = p_1. \quad (11)$$

$$\text{限速警戒：} p(\text{II}) = p_2 - p_1. \quad (12)$$

$$\text{出巡警戒：} p(\text{III}) = 1 - p_2. \quad (13)$$

表 5 模型参数部分的参数估计

参数	df	参数估计值	标准差	卡方	显著水平
$\beta_{01}$	1	-10.0242	0.9824	104.1215	$< 0.0001$
$\beta_{02}$	1	-6.0922	0.7146	72.6890	$< 0.0001$
$x_1$	1	0.1594	0.0180	78.0117	$< 0.0001$
$x_2$	1	0.0226	0.00711	10.0717	0.0015
$x_3$	1	0.0099	0.00360	7.5411	0.0060

表 6 模型的有效性检验

模型	模型拟合标准		似然比检验		
	-2 倍对数似然值	卡方	df	显著水平	
仅截距	482.036				
最终	185.140	23.2079	3	$< 0.0001$	

根据上述方程进行回报检验，检验结果如表 7 所示，可以看出相对平坦地形铁路的分级警戒概率模型效果明显优于山脉地形的警戒模型，其出巡警戒预报、封锁警戒预报的效果都较好，准确率分别为 85.0% 与 87.7%，限速警戒预报的效果较差，准确率仅 75.8%，其中限速报出巡为 18.1%。分析这三种警戒预报的错报、漏报相对应的降水因子发现，错报和漏报的情况与山脉地形基本一致，大部分是由于 1 h 最大降水量、24 h 降水量以及连续降水量三者中的 1 个或 2 个因子的值很小，但是另外 1 个因子的值较大，因此采用 Logistic 回归模型预测相对平坦地形铁路警戒级别的概率结合 3 个因子的降水量值共同判断警戒级别的效果最优。

表 7 概率预报检验结果

出巡预报/%		限速预报/%		封锁预报/%	
准确率	85.0	准确率	75.8	准确率	87.8
错报限速	15.0	限速报封锁	6.1	封锁报限速	9.5
错报封锁	0.0	限速报出巡	18.1	封锁报出巡	2.7

### 3 结论与讨论

(1) 鹰厦线的铁路水害年际变化差异大且分布不均，发生时间集中在汛期且局地性较强。

(2) 以青州 - 卓宅区段研究的山脉型铁路水害分级警戒概率预报模型的总体结果较好，出巡警戒预报与封锁警戒预报的准确率分别为 78.3% 与 74.4%，限速警戒预报的效果较差，准确率为 58.8%。

(3) 以卓宅 - 龙海区段研究的相对平坦地形铁路水害分级警戒概率预报模型的效果很好，出巡警戒预报、封锁警戒预报的准确率分别为 85.0% 与 87.7%，限速警戒预报的准确率为 75.8%。

(4) 分析两种不同地形的铁路水害概率预报模型的错报、漏报例子发现，三种警戒的错报或者漏报大部分是由于 1h 最大降水量、24h 降水量以及连续降水量三者中的 1 个或 2 个因子的值很小，但是另 1 个因子的降水值很大，因此在铁路防洪调度中结合水害概率预报模型以及降水量因子共同判断警戒级别的效果更优。

### 参考文献：

- [1] 郑度, 张懿铨. 青藏铁路沿线生态与环境安全[M]. 杭州: 浙江科学技术出版社, 2009: 19-88.
- [2] 刘峰贵, 张海峰, 陈琼, 等. 青藏铁路沿线自然灾害地理组合特征分析[J]. 地理科学, 2010, 30(3): 384-390.
- [3] 安志山, 张克存, 屈建军, 等. 青藏铁路沿线风沙灾害特点及成因分析[J]. 水土保持研究, 2014, 21(2): 285-289.
- [4] 接小峰, 徐林荣, 曹禄来, 等. 铁路水害特征与致灾概率预测模型研究[J]. 灾害学, 2015, 30(1): 187-192.
- [5] 李晓霞, 王勇, 关向宁, 等. 铁路水害及预警模型[J]. 干旱区研究, 2011, 28(1): 181-186.
- [6] 接小峰, 徐林荣, 曹禄来, 等. 铁路水害预警雨量阈值预测方法及其应用[J]. 自然灾害学报, 2015, 24(5): 177-183.
- [7] 杨思全, 陈亚宁, 王昂生, 等. 基于灰色理论的铁路环境水害危险度计算模型构建研究[J]. 中国科学院研究生院学报, 2002, 19(4): 366-371.
- [8] 杨思全, 陈亚宁, 王昂生, 等. 基于灰色理论的铁路水害危险度计算模型[J]. 自然灾害学报, 2002, 11(2): 119-124.
- [9] 常振海, 刘薇. Logistic 回归模型及其应用[J]. 延边大学学报(自然科学版), 2012, 38(1): 28-32.
- [10] 高歌, 张明芝. 多分类有序反应变量 Logistic 回归及应用[J]. 同济大学学报, 2003, 31(10): 1237-1241.

## Graded Alerting Probability Forecast of Railway Water Disaster Based on the Multinomial Logistic Regression Model

WU Fan and QUE Zhiping

(*Meteorological Service Center of Jiangxi, Nanchang 330046, China*)

**Abstract:** Based on the railway water disaster data and precipitation data of Yingxia Railway from 2004 to 2014, the temporal and spatial distribution characteristics of the railway water disaster are analyzed, and the graded alerting probability forecast model of railway water disaster by two different terrain are researched. The results show that the annual variation and distribution of flood disaster are different, railway water disaster always occurred in flood season. The test results of probability forecast model by mountain terrain show that the graded alerting accuracy rates are 78.3%, 58.8% and 74.4%. The test results of probability forecast model by relatively flat terrain show that the graded alerting accuracy rates were 85.0%, 75.8%, and 87.7%. Alerting effect could be better when combined with the probability forecast model and the precipitation values of 3 factors, the result have an important guidance for the safe operation and efficient operation of the railway department.

**Key words:** Yingxia Railway; railway water disaster; graded alertness; Precipitation Multinomial Logistic Regression

(上接第 16 页)

## Engineering Disease Classification and Control in Aeolian Soil Area

ZHANG Yu<sup>1,2</sup>, ZHANG Xiangdong<sup>3</sup>, CHEN Tielin<sup>1,2</sup> and LIU Jiashun<sup>3</sup>  
(1. *School of Civil Engineering, Beijing Jiaotong University, Beijing 100044, China*;  
2. *Key Laboratory of Urban Underground Engineering (Ministry of Education), Beijing Jiaotong University, Beijing 100044, China*; 3. *Institute of Civil Engineering and Transport, Liaoning Technical University, Fuxin 123000, China*)

**Abstract:** As for the disease classification, influencing factors and controlling measurements for Aeolian Soil engineering in western Liaoning Province of China, methods of field investigation, indoor experiment and construction requirements analysis are used. The disease classifications are obtained in terms of disease phenomenon and disease mechanism respectively, then analyzing the four-level influencing factors under the Engineering-Geosphere System and Analytic Hierarchy Process, and giving the controlling measurements, thus forming the systematical and simple frame for engineering disease identification and control. The results show: ①Aeolian Soil is silty clay with coarse granule, good gradation, medium compressibility, low shear strength, easy liquefaction, apparent structure and non-obvious collapsibility; ②Engineering disease in Aeolian Soil area often develops under the reason of compressible settlement, shearing damage, vibration-subsidence and liquefaction, collapsibility, frost-heave and thaw-settlement, weathering, current scour, etc.; ③Based on the Engineering-Geosphere System, the main influencing factors for engineering disease in Aeolian Soil area are physicommechanical property of Aeolian Soil, human activity and natural environment; ④It's necessary to establish systematical frame of disease identification, adjusting and control, thus making the work normative, processive, standardizing and informatized.

**Key words:** Aeolian soil; engineering; disease; classification; geosphere system; control