

承 诺 书

我们仔细阅读了第十二届华中地区大学生数学建模邀请赛的竞赛细则。

我们完全明白，在竞赛开始后参赛队员不能以任何方式（包括电话、电子邮件、网上咨询等）与队外的任何人（包括指导教师）研究、讨论与赛题有关的问题。

我们知道，抄袭别人的成果是违反竞赛规则的，如果引用别人的成果或其他公开的资料（包括网上查到的资料），必须按照规定的参考文献的表述方式在正文引用处和参考文献中明确列出。

我们郑重承诺，严格遵守竞赛规则，以保证竞赛的公正、公平性。如有违反竞赛规则的行为，我们将受到严肃处理。

我们的参赛报名号为：163

参赛队员 (签名)

队员 1：尹可汗

队员 2：刘子川

队员 3：王静

湖北省工业与应用数学学会第十二届华中地区大学生数学建模邀请赛组委会

编号专用页

选择的题号：A

参赛的编号：163

（以下内容参赛队伍不需要填写）

竞赛评阅编号：

A 题 超强台风“蝴蝶”的风险评估

摘要

本文针对台风的危险性和各省灾情数据建立了一个关于的台风风险评估的模型，通过“蝴蝶”台风的特征与历史台风拟合，实现路径拟合，并结合龙卷风的形成原理，实现对台风的完整预测，最后提出了对台风多发地带的建议。

针对问题一，借助联合国国际减灾策略(ISDR)2004年提出的风险表达式，危险性 \times 脆弱性，对各省的风险进行评估，我们从题中所给条件，气象官网，各省年鉴中筛选了相关信息，利用台风登陆各省的概率作为危险性的重要指标，把直接经济损失，受灾人数，农田受损面积，房屋受损4个方面的数据规范化，得出灾害程度，分析得出广东省是我国台风影响最为严重的地区，其次是福建省和浙江省，其他沿海地区被遭受的风险相对较低。该结果考虑了较多个有代表性的变量，数据来源可靠，其结果符合我国沿海地区的国情，具有一定的可靠性。

针对问题二，为实现对“蝴蝶”台风的预测，我们查找了该台风的经纬度，台风强度，以及最大风速，利用python语言将该台风与历史台风数据集比较，找出与台风程度最为接近的台风数据，并利用中国台风网的数据库，通过经纬度距离换算，得出模拟的路径，最终得出结论该台风属于西太平洋转向路径，台风从菲律宾以东的海面向西北移动，在 25° N附近转向东北方，向日本方向移动。对我国沿海大陆的影响较小，但是对我台湾的影响较大，这与台风网上所给出的信息与专家的预测一致。该预测方法使用的数据集很大，在一定程度上提高了对台风的预测精准度。

针对问题三，为分析龙卷风对于我国各省的袭击概率，我们结合龙卷风的形成原理，用圆形简化龙卷风模型，查阅文献，找出龙卷风半径和中心最大速度之间的关系。通过echarts软件，实现对经纬度，台风强度，风速的可视化，通过python编程利用朴素贝叶斯分类器的算法计算沿海主要城市与台风中心距离，利用上述公式计算影响范围与各城市的距离，得出结论广东影响最大，依次是海南，台湾，福建，浙江，上海，与第一问的结论基本一致，利用的数据较多，具有一定的准确性，进一步检验了评估模型的正确性，龙卷风的原理分析也为问题四提供了帮助。

针对问题四，结合问题三中关于龙卷风形成原理的分析，查阅文献并参考其他台风多发国家房屋结构的设计，从龙卷风成因，我国房屋结构存在的问题，以及其他国家的房屋设计结构三个方面提出了一些合理的建议，同时为了减小台风对高频省市的受灾损失，我们提出建设一些紧急避难场所来保护受灾群众，减小模型中的脆弱性指标降低风险和经济损失。

本文在风险评估的因素处理上不仅考虑了台风的危害强度，还从我国各省的灾情状况上筛选了有效的数据，利用回归测试来评估分类器，提高了准确率。

关键词： 风险评估 python 概率模型 贝叶斯分类器

目录

1. 问题重述.....	1
1.1. 问题背景.....	1
1.2. 问题提出.....	1
2. 模型假设.....	1
3. 符号说明.....	2
4. 问题一的模型建立与求解.....	2
4.1. 问题描述及分析.....	2
4.2. 模型建立与求解.....	3
4.2.1. 前期准备:	3
4.2.2. 危险性分析.....	7
4.2.3. 脆弱性分析.....	7
4.3. 模型评价与检验.....	9
5. 问题二的模型建立与求解.....	9
5.1. 问题描述及分析.....	9
5.2. 模型建立与求解.....	9
5.3. 模型评价与检验.....	11
6. 问题三的模型建立与求解.....	12
6.1. 问题描述及分析.....	12
6.2. 模型建立与求解.....	12
6.3. 模型评价与检验.....	15
7. 问题四.....	16
8. 模型总结与评价.....	17
参考文献.....	18
附录.....	18

1. 问题重述

1.1. 问题背景

台风是发生在热带洋面上的热带气旋，往往伴有大量的降水和强烈的风暴，在影响范围内的地区往往会遭受到严重的灾害，是十大灾害中造成伤亡人数最多的一种。为了尽量减少台风造成的损失，我国极其重视台风的最佳路径预测和风险评估工作。

1.2. 问题提出

（1）根据中国气象局热带气旋资料中心提供的近 50 年 CMA 热带气旋最佳路径数据集，并收集其他相关资料，对中国各省（以地级市为单位）进行热带气旋的风险评估，评估的主要内容主要包括受灾面积、经济损失、受灾人口和房屋损失四个方面；

（2）对第 2 号超强台风“蝴蝶”在中国各省内的危害进行风险评估，并对所得的结论进行验证；

（3）结合龙卷风形成的原理分析我国各省被龙卷风袭击的概率；

（4）请你们团队写一封信给当地建造房屋的居民和工程师们，针对龙卷风多发地区新建建筑物的结构和外形提出建议。

2. 模型假设

（1）基于东南沿海地区经济发达，科学技术水平高，可假设我国东南沿海地区地级城市的抗灾和防御能力基本相同。

（2）假设台风在移动的过程中，不会发生气旋，气温等多种气象因素的巨大变化。

（3）假设台风的强度与危险性成正比。

（4）假设龙卷风为圆形的，且移动时不改变形状。

3. 符号说明

表 1：符号说明

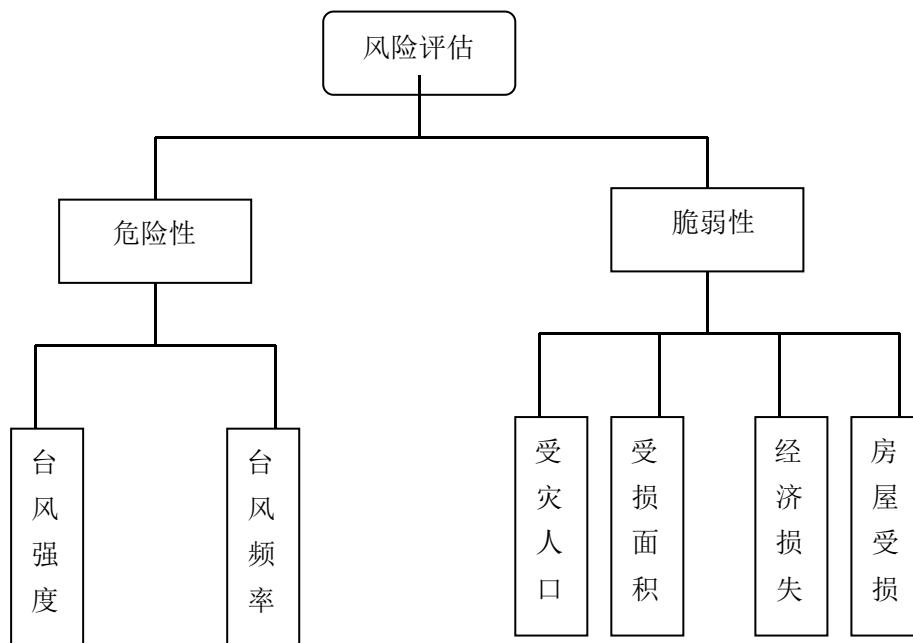
符号	符号描述
ϕ	纬度
λ	经度
$F1F2\cdots Fn$	不同省份
P	台风登陆各省的概率
a, b, c, d	灾害脆弱性因素
I	规范化灾害因素的相对值
G	脆弱性风险指标
d	两点间的距离
θ	圆上的弧度
r	地球的半径
$V(r)$	距离台风为 r 的大风风速
V_m	最大风速
R_{gale}	台风的最大半径
RMW	台风影响范围半径
C	不同省市的集合

4. 问题一的模型建立与求解

4. 1. 问题描述及分析

问题一的主要目的是通过收集资料整理出具有代表性的因素，然后建立出一套有效的各省市台风风险评估模型。风险评估主要由灾害危险性，承灾体脆弱性，以及城市抗灾能力三个方面组成，基于假设 1 的内容，由于东南沿海地区地级城市的抗灾和防御能力基本相同，所以我们主要从灾害危险性，承灾体脆弱性两个方面来分析，以达到简化模型的目的。具体步骤与思路如下：

- 1. 我们搜集和整理中国气象局热带气旋资料中心提供的 CMA 热带气旋最佳路径数据集，统计出建国以来热带气旋的产生数量特征，沿海各省受到各种台风强度的频率，以及台风登陆我国各省的频率，建立危险性分析模型。
- 2. 查找台风灾情的数据，统计和筛选出台风灾害的重要指标，以及各指标在各省的数据情况记录。建立脆弱性分析模型。
- 3. 利用风险运算基本公式，结合上述模型，建立出总风险评估的标准。



4.2. 模型建立与求解

4.2.1. 前期准备

1. 利用中国气象局热带气旋资料中心提供的 CMA 热带气旋最佳路径数据集，利用软件 PyCharm 通过 python 语言将 .txt 文件转化成为 Excel 可识别的 .csv 文件，然后用在线 GIS 分析台风的分布情况，并通过 Excel 对数据进行分析建国以来我国总体上所受台风的影响情况。

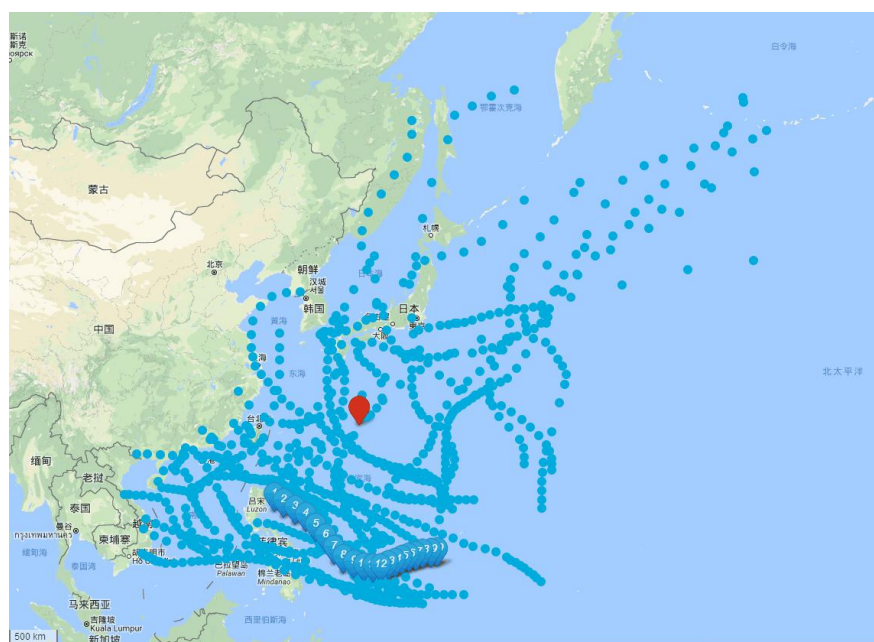


图 1: GIS 热带气旋最佳路径数据集分布图

由图 1 的信息，我们可以看出我国东南沿海地区受到台风的影响十分频繁，尤其是广东，台湾，海南这些低纬度的地区受台风的影响最为突出。

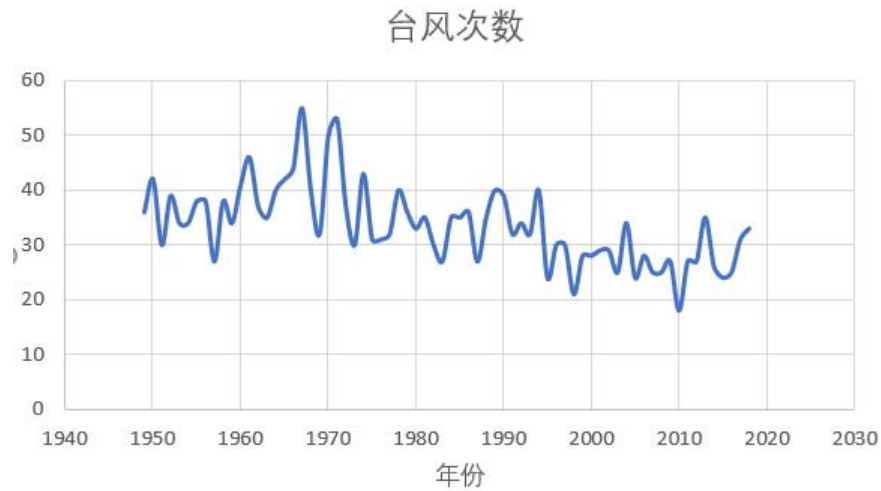


图 2：1949~2018 年间西北太平洋热带气旋出现次数统计

建国以来，西太平洋地区共有 2348 次热带气旋，平均每年约有 34 次，频率较高，在图 2 中，我们可以看出每年的台风变化复杂，但是始终围绕在 30 次附近，对整个亚洲东南部造成的影响十分巨大，所以台风问题也值得我们去重视。

2. 利用中国气象局热带气旋资料中心，搜集了登陆我国的气象的热带气旋数据集，整理出我登陆我国各地区的台风频率，以及台风强度，并且筛选出来台风影响的主要地区。

表 1：1949-2018 热带气旋在我国登录情况

省份	登陆次数
广东	263
海南	162
台湾	148
福建	127
浙江	44
广西	35
山东	20
香港	14
辽宁	13
上海	8
江苏	7
天津	1

根据表 1 的情况我们可以看出来，排名前五位的省份中，广东省，海南省，台湾省，福建省登录次数均超过 100 次，受到热带气旋的影响十分严重。

省份	TD	TS	STS	TY	STY	SuperTY	登陆次数
广东	51	55	73	55	8	2	244
海南	40	27	31	37	3	2	140
台湾	13	13	27	60	23	6	142
福建	18	28	34	38	3	1	122
浙江	5	4	13	16	5	1	44
江苏	3	1	5	4	1	1	15
香港	3	2	1	4	3	1	14
天津	1	0	0	0	0	0	1
上海	8	3	15	13	5	1	37

表 3：热带气旋等级

名称	强度	风速
TD	0	低于热带低压
TD（热带低压）	1	10.8-17.1m/s
TS（热带风暴）	2	17.2-24.4 m/s
STS（强热带风暴）	3	24.5-32.6 m/s
TY（台风）	4	32.7-41.4 m/s
STY（强台风）	5	41.5-50.9 m/s
SuperTY（超强台风）	6	51.0 m/s

通过表 2，表 3 的分析，我们基本可以确定这些省份属于台风的多发地区，经过资料检索，三级以上的台风皆可以对这些地区造成不同程度的损坏，而且 3 级以上的台风频率均已经超过一半，故我们将对该区域的地级市进行研究分析。

3. 收集关于台风灾情的各种资料与各省的地理基础信息的年鉴资料，选取重要的指标，对模型建立奠定基础。^[1, 2]

表 4：2009-2017 广东，福建，浙江省市的台风受灾情况列表

年份	省市	受灾人数（万人）	受灾面积（千公顷）	直接经济损失（亿元）	房屋受损（间）
2017	广东	171.46	41.88	55.77	34
2016	广东	194.6	23.04	9.26	12
2015	广东	376.4	24.71	28.76	78
2014	广东	514.99	46.03	60.41	11717
2013	广东	509.44	50.64	74.2	5382
2012	广东	204.02	114.68	17.47	1391
2011	广东	77.92	123.67	12.63	600
2010	广东	236.3	28.81	30.62	12000
2009	广东	361.24	289.95	38.99	6300
2017	福建	53.01	2.61	1.21	1606.75
2016	福建	3.59	20.48	16.06	571
2015	福建	34.18	30.68	30.79	3858
2014	福建	8.65	16.68	4.29	53
2013	福建	38.07	55.96	45.06	6066
2012	福建	1.42	10.59	2.64	6
2011	福建	56.76	1.62	5.32	200
2010	福建	116.42	14.79	33.1	700
2009	福建	165	73.52	19.83	1400
2017	浙江	318.5	1.08	0.87	1606.75
2016	浙江	116.44	5.33	2.4	2
2015	浙江	327.85	27.44	11.2	116
2014	浙江	93.07	9.55	4.33	60
2013	浙江	736.62	378.1	28.17	2124
2012	浙江	318.5	68.36	42.57	781.75
2011	浙江	318.5	68.36	1.92	781.75
2010	浙江	318.5	14.79	33.1	781.75
2009	浙江	318.5	42.21	11.85	781.75

由于海南省和台湾省的数据缺失过多，则将 F2 和 F3 的数据剔除。为保证结果的精确性，所以没有采用他们的数据，虽然某些省份个别数据的缺失，但这些数据来源中国海洋公报，准确可靠，且这些数据都是近年来的统计，通过假设各省的灾害遭遇情况变化不大，故我们用平均值来代替个别缺失的数据，以保证预测的效果准确。

4.2.2. 危险性分析

在前提准备的工作下，我们基本确定了我们的模型区域——广东省(F1)，浙江省(F)，福建省，我们利用表2所给的数据计算出了各省市3级强度台风下的频率列表。

表5：模型区域的3级台风登陆的频率

省份	3级以上
广东	0.56
福建	0.53
浙江	0.79

然后我们根据假设，3级台风以下的所有风险可以抗击，实现不损失，而台风强度可以看做与危险性成正比关系，那么我们用期望值来实现危险性的评估值。

表6：各级台风的频率情况

省份	STS=3	TY=4	STY=5	superTY=6	低于3级=0
广东	0.30	0.23	0.03	0.01	0.43
福建	0.28	0.31	0.02	0.01	0.38
浙江	0.30	0.36	0.11	0.02	0.20

可算得危险性期望评估：

$$E(F) = P(STS) \times STS + P(TY) \times TY + P(STY) \times STY + P(superTY) \times superTY + 0$$

$$\text{广东: } E(F1) = 2.03 \quad \text{福建: } E(F4) = 2.24 \quad \text{浙江: } E(F5) = 3.01$$

4.2.3. 脆弱性分析

脆弱性是从人文和自然意义的两个方面分析灾害对于整个社会的损失程度的反映，在受伤人口，农田受灾面积，经济损失，房屋受损，地表破坏，基础设施与重点工程破坏，生态环境稳定性与恢复力指标，这8个方面，我们选取了具有代表性和普遍性的4个指标来进行模型的预测，即受灾人口，农田受损面积，经济损失，房屋损坏。经过分析，由于经济损失中，间接经济损失难以估算，故我们只研究直接经济损失，本文引用了灾害损失定量计算模型作为损失程度的标准，来实现对于脆弱性标准的估计。

对于数据我们先求取其平均值：

$$S = \sum_{n=1}^9 s_i / 9$$

表 7：受灾指标的平均值

省市	a=受灾人数 (万人)	b=受灾面积（千 公顷）	c=直接经济损 失（亿元）	d=房屋受损（间）
广东	294.04	82.60	36.46	4168.22
福建	53.01	25.21	17.59	1606.75
浙江	318.50	68.36	15.16	781.75

为了实现对灾害的定量化运算，首先必须把受灾人口，农田受损面积，直接经济损失，房屋损坏，折算成规范化指数：^[3]

$$\begin{aligned} I_a &= \lg(a) - 1; \\ I_b &= \lg(b) - 2; \\ I_c &= \lg(c) - 2; \\ I_d &= \lg(d) - 2; \\ G &= I_a + I_b + I_c + I_d; \end{aligned}$$

然后将各因子的规范化指数相加，作为灾害的损失定量指标，得到我们的脆弱性分析结果，如表 8

表 8：脆弱性灾害指数结果

最后利用自然灾害的风险概念基本公式：

省份	灾害指 数 G	受灾人数 I _a	受灾面积 I _b	直接经济损 I _c	房屋 I _d
广东	8.346	3.46	2.91	1.56	0.4168
福建	7.526	3.72	2.401	1.245	0.160675
浙江	7.58	3.5	2.83	1.18	0.0781

$$R = E(F_i) \times G(F_i)$$

得到如下结果

表 9：各省风险指标

省份	R
广东	22.81
福建	16.24
浙江	16.94

4.3. 模型评价与检验

优点：该模型在分析风险评估搜集了全面而且可靠的数据，模型的运算时间复杂度较低，模型考虑的因素有 6 个且具有代表性，使得预测的模型评估具有普遍性。

缺点：该模型在预测风险评估时使用的方法过于简便，在精准度上还不够精准，同时每个城市的抗灾能力也是有差异的，而且抗灾能力对于抗击台风还是有很大的作用的，这会使得模型建立更加精确。

5. 问题二的模型建立与求解

5.1. 问题描述及分析

2019 年的台风“蝴蝶”于 2 月 20 日凌晨形成，从起始点（155.1° E，5.1° N）开始爆发向西北方向前进，并逐步加强。至 2 月 23 日已经成为风速为 58 的 superTY 级别台风，在经纬度为（142.7° E，12.1° N）附近。台风 2 号“蝴蝶”还在持续奔跑，中国大陆各省上都是不安分。尽管最近南支槽减弱，副高东退，但是南北方还是不平静。是华南一带降雨没停，我国两广江西福建等地出现中到大雨，其中江西南部 and 广东北部等地部分地区更是出现了 50~86 毫米的暴雨。为此，必须要对第 2 号超强台风“蝴蝶”在中国各省内的危害进行风险评估。

我们对历史台风进行拟合，找出与“蝴蝶”相似的台风模型，再对“蝴蝶”进行预测路径和与以往台风对大陆各省的风险评估进行模拟，通过第一问中的模型对“蝴蝶”的风险评估进行验证从而得出结论。

5.2. 模型建立与求解

为了找到“蝴蝶”的轨迹与历史数据中相拟合的台风，先找出蝴蝶中的特别参考的坐标经纬度：起点（155.1° E，5.1° N），最大中心风速点（140.1° E，13.7° N），终点（135.6° E|17.5° N）。为此简化模型，在这里我们假设地球为均匀球体（ $r=6371.393$ 公里），

在历史数据中找到与这三个点连成的直线段距离最短的台风，通过 python 计算经纬度转化距离的公式：

$$\text{hav}\left(\frac{d}{r}\right) = \text{hav}(\psi_2 - \psi_1) + \cos(\psi_1)\cos(\psi_2)\text{hav}(\lambda_2 - \lambda_1)$$

$$\text{hav}(\theta) = \sin^2\left(\frac{\theta}{2}\right) = \frac{1 - \cos(\theta)}{2}$$

先看第一个式，等号的右边，输入参数有 ϕ 、 λ ，等号的右边有 d 、 r ，其中 ϕ 表示纬度， λ 表示经度， d 是我们要的两点的距离。

r 是地球半径， d/r 表示两点在圆上的弧度 θ 。通过整理两个式子可得：

$$d = 2r * \arcsin \left(\sqrt{\text{hav}(\psi_2 - \psi_1) + \cos(\psi_1)\cos(\psi_2)\text{hav}(\lambda_2 - \lambda_1)} \right)$$

$$= 2r * \arcsin \left(\sqrt{\sin^2\left(\frac{\varphi_2 - \varphi_1}{2}\right) + \cos(\psi_2)\cos(\psi_1)\sin^2\left(\frac{\lambda_2 - \lambda_1}{2}\right)} \right)$$

也就是说，只要根据上面最后这行式子，带入相应的经纬度值，以及地球半径，就可以得到我们想要的两点的球面距离，其我们想要的结果如下：

距离： 318.64	日期： 2003041600
距离： 335.3	日期： 2004061518
距离： 375.08	日期： 2004061600
距离： 391.99	日期： 2004082306
距离： 382.09	日期： 2004082312
距离： 399.37	日期： 2004082318
距离： 392.99	日期： 2009100406
距离： 272.79	日期： 2009112512
距离： 162.12	日期： 2009112518
距离： 132.15	日期： 2009112600
距离： 178.24	日期： 2009112606
距离： 245.58	日期： 2009112612
距离： 307.07	日期： 2009112618
距离： 374.55	日期： 2009112700
距离： 249.47	日期： 2013101800
距离： 236.29	日期： 2013101806
距离： 271.16	日期： 2013101812
距离： 286.46	日期： 2013101818
距离： 294.06	日期： 2013101900
距离： 345.38	日期： 2013101906
距离： 398.53	日期： 2013101912

图 3：拟合出的历史台风

通过“中国气象局”提供的历史台风网站，找出上述的与“蝴蝶”拟合的路径轨迹图：

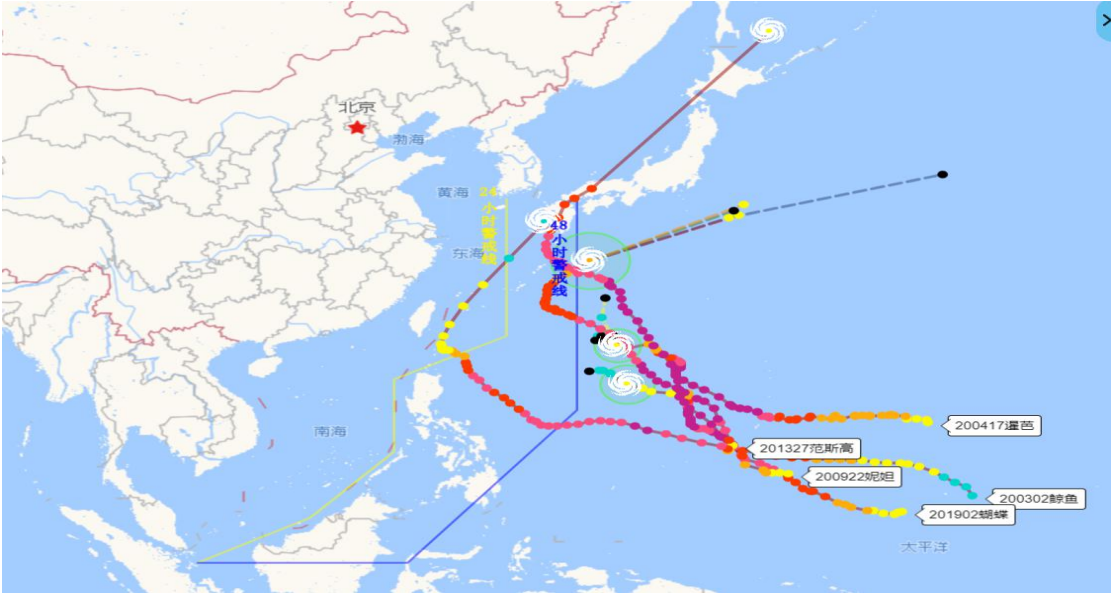


图 4：与“蝴蝶”拟合的历史台风

结合第一问对各省的风险评估进行预测，发现台风“蝴蝶”对各省的风险指标都不算太大。根据和表 9 的数据可以看出，最终得出结论该台风属于西太平洋转向路径，台风从菲律宾以东的海面向西北移动，在 25° N 附近转向东北方，向日本方向移动。对我国沿海大陆的影响较小，但是对我台湾稍微有点影响 。

表 10：各省风险指标

省份	R
广东	4.24
福建	6.78
台湾	9.42
上海	3.15
浙江	5.94

5.3. 模型评价与检验

这与新闻报道的预计的符合：“蝴蝶”将以每小时 20 公里左右的速度向西北方向移动，强度逐渐加强，最强可达强台风级或超强台风级（15-16 级，48-55m/s）。25 日起，“蝴蝶”将在 140E 附近转向偏北方向移动，强度逐渐减弱。未来，“蝴蝶”对我国海区无影响。

6. 问题三的模型建立与求解

6.1. 问题描述及分析

按照问题三要求结合龙卷风形成的原理分析我国各省被龙卷风袭击的概率，我们查阅相关资料了解到龙卷风的成因：

一是在炎热的夏季，强烈的太阳光照射，气温升高非常剧烈，大量的水分从海面蒸发到空气中，水蒸汽到达高空后遇冷变成区域性的积雨云。

二是积雨云在受到强烈的阳光照射后，积雨云中的水分子或空气中的其它分子电子具有较高的能量是因为受到强烈的阳光照射，使云团成为被极化的带电云团。

三是在强烈的阳光照射下，大气压强发生区域性变化产生大气对流形成风，风推动云团定向流动。

四是当这些相对孤立的带电云团在天空中运动时，在大风力的作用，这些高速运动的带电云团在洛伦兹力的作用下发生旋转，因此形成了龙卷风。

根据龙卷风形成的原因，为了确定每个省被龙卷风袭击的概率，对主要受袭击的沿海城市进行分类，利用朴素贝叶斯分类器来算出沿海城市受袭击的概率。

6.2. 模型建立与求解

假设龙卷风为圆形的，且移动时不改变形状。根据龙卷风的形成结构建立龙卷风半径和中心最大速度之间的关系：^[4]

$$V(r) = V_m \cdot \left(\frac{R_{gale}}{RMW} \right)^{-X}$$

其中 X 为：

$$X = \frac{\ln\left(\frac{V_m}{62}\right)}{\ln\left(\frac{R_{gale}}{RMW}\right)}$$

根据数据画出龙卷风半径和中心最大速度之间的关系图：^[5]

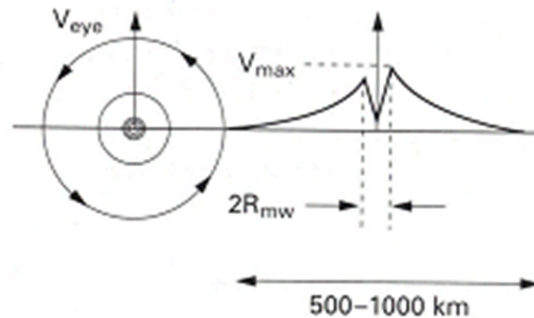


图 5：台风影响半径与速率之间的关系

假设某时刻的台风有 5 项特征 F1、F2、F3、F4、F5、F6，分别为经度、纬度、中心最低气压，台风影响半径，中心最大速度和台风强度等级。现有 m 个类别 C（Category），分别为不同的省份（主要包括了沿海地区的省市）。贝叶斯分类器就是计算出概率最大的那个分类，也就是求下面这个算式的最大值：

$$P(C|F1F2...F6) = P(F1F2...F6|C)P(C) / P(F1F2...F6)$$

将每一时刻的台风出现概率按经纬度进行数据可视化，并按台风强度的颜色进行划分，颜色越深，某地台风出现频率越高，台风强度越大。得到以下散点分布图：



图 6：台风经纬度与强度的关系

由于 $P(F1F2...F6)$ 对于所有的类别都是相同的，可以省略，问题就变成了求 $P(F1F2...F6|C)P(C)$ 的最大值。朴素贝叶斯分类器则是更进一步，假设所有特征都彼此独立，因此

$$P(F1F2...F6|C)P(C) = P(F1|C)P(F2|C) \dots P(F6|C)P(C)$$

上式等号右边的每一项，都可以从统计资料中得到，由此就可以计算出每个类别对应的概率，从而找出最大概率的那个类的所在城市，其结果如图。



图 7：我国各省被台风袭击频率

虽然“所有特征彼此独立”这个假设，在现实中不太可能成立，但是它可以大大简化计算，而且有研究表明对分类结果的准确性影响不大，通过 python 编程实现其分类器验证的实验结果和用 excel 对各省登陆的统计数据如下：

表 10：近 20 年不同强度 TC 袭击中国各省的概率

省份	TD	TS	STS	TY	STY	superTY	概率
广东	51	55	73	55	8	2	0.332%
海南	40	27	31	37	3	2	0.191%
台湾	13	13	27	60	23	6	0.193%
福建	18	28	34	38	3	1	0.169%
浙江	5	4	14	16	5	1	0.059%
上海	9	6	12	11	2	2	0.057%

根据模型画出我国沿海地区被龙卷风袭击的频率热图如下：

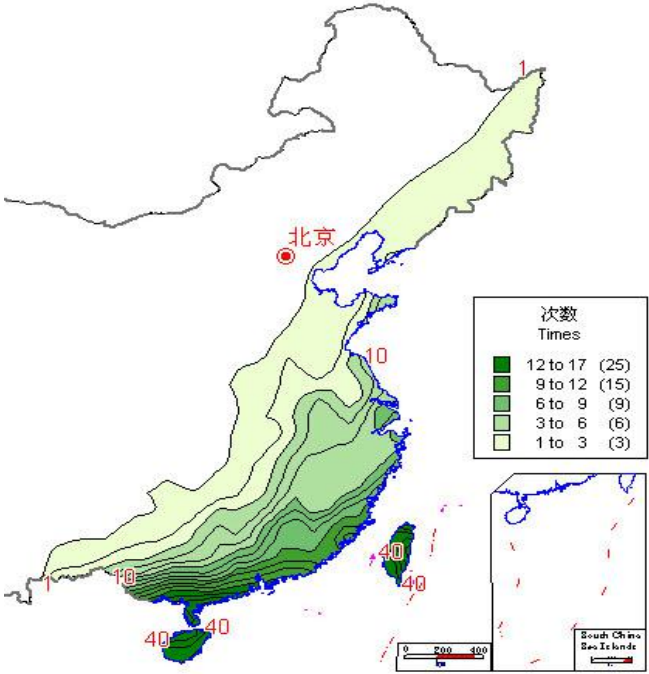


图 8：龙卷风对沿海各省袭击的频率热图

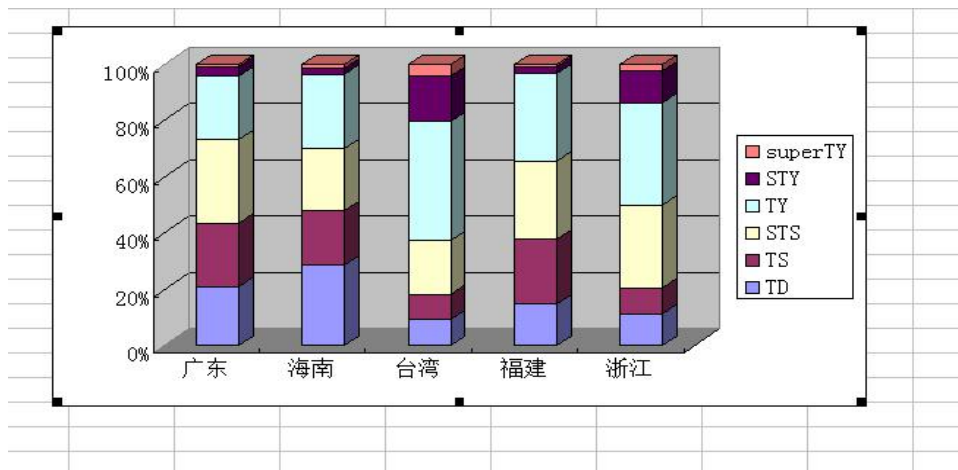


图 9：各省风险评估概率

6.3. 模型评价与检验

通常使用回归测试来评估分类器的准确率，最简单的方法是用构造完成的分类器对训练数据进行分类，然后根据结果给出正确率评估。但这不是一个好方法，因为使用训练数据作为检测数据有可能因为过分拟合而导致结果过于乐观，所以一种更好的方法是在构造初期将训练数据一分为二，用一部分构造分类器，然后用另一部分检测分类器的准确率大概在 83%左右，说明是十分客观的。

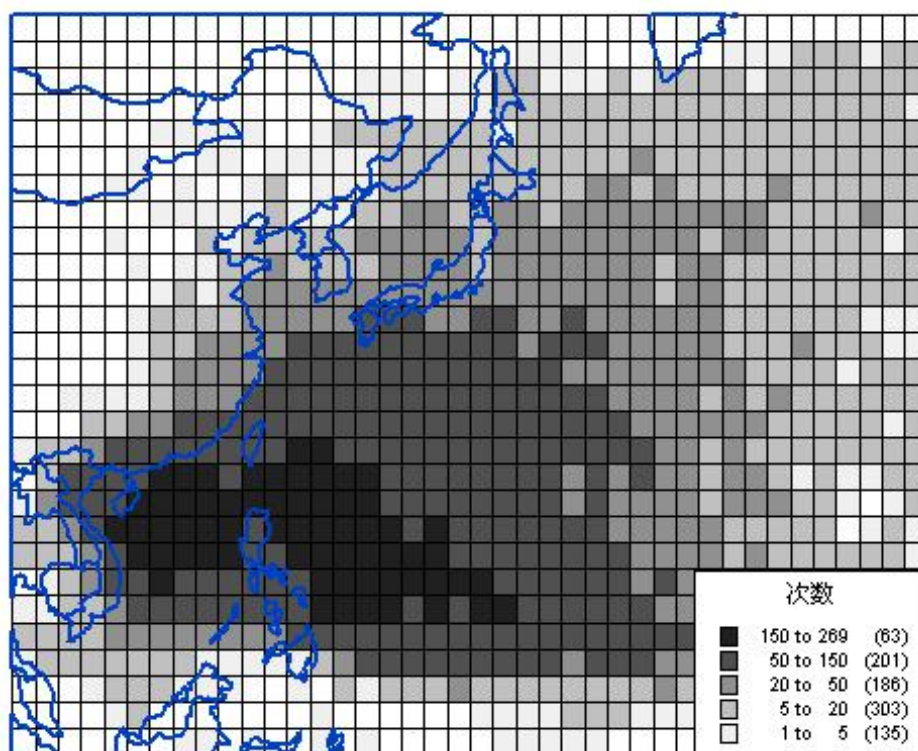


图 10：热带气旋中心位置次数分布

7. 问题四

给居民和工程师们的一封信

亲爱的居民和工程师们：

你们好！我们是武汉理工大学的学生，现在就龙卷风多发地区新建建筑物的结构和外形提出一些建议。

龙卷风是一种破坏力极大的局地强风，是一种小范围的灾害性天气现象。所经之处经常会发成大树连根拔起、建筑物倒塌、车辆被掀翻等现象，甚至可以瞬间摧毁大片的庄稼田地，人畜的生命受到威胁，经济遭受巨大的损失，给人们留下难以抹去的恐惧记忆。

龙卷风破坏力如此之大的原因有三个：第一，龙卷风具有极高的风速，若直接受到龙卷风的袭击，一般的建筑物很难抵抗如此之高的风速所带来的巨大冲击力，就很有可能倒塌；第二，龙卷风内部与外围有明显的气压差，称为“气压降”，这个特性使龙卷风具有强大的吸附、扭转效果，造成房屋内外压差，会直接吸走一些建筑的屋顶，并且导致四周的墙壁坍塌；第三，因为龙卷风会携带着之前经过的建筑物的碎片、树木的枝叶以及其他杂物前行，形成速度和威力都很大的飞射物，对后袭的建筑物造成更剧烈的二次破坏。

但是，我国部分居民建筑普遍存在一些问题，如设计房屋时构造时考虑较少，导致一些房屋的屋顶与四周的墙面之间的连接件少或强度不高，且屋面较轻，房屋抗风升力的能力几乎没有；存在很多违章建筑，房屋建筑质量差，安全性低，且稳固性和抗风承载力都不达标；建筑之间的距离过近等，都增大了龙卷风对人民生命安全的威胁。

由于不同地区龙卷风发生的时间不定，解决居民的房屋建筑的安全问题便是当务之急。针对这个问题，我们参考日本、越南及菲律宾等台风多发国家建筑结构特点，提出以下建议：

（1）重视房屋结构：经过大量数据统计，在龙卷风袭击后大多数建筑屋面系统受损最为严重，所以应该改善屋面的抗风性和房屋整体的牢固性。对于钢筋混凝土屋面重量略轻，则应强化板端锚固钢筋与屋面圈梁的连接构造；对木屋等柔性屋面，自身刚度小，整体性能差，加上重量较轻，应加强房屋整体的联接，强化屋面覆盖层与檩条、屋架及檩条、屋架与其承重墙、柱间的连接构造，确保房屋建筑形成一个有效的传力路径，增加其牢固性；^[6]

（2）利用好结构的延性：延性好的结构可以吸收和耗散龙卷风的能量且形变能力大，可以减少龙卷风对房屋的破坏，有效地防止建筑倒塌的发生。多使用造柱-圈梁体系，增加墙体接拉钢筋、钢筋砖带，减小构造柱间的距离等，都能有效地增加房屋结构的延性；^[6]

（3）选择合适的建筑地形：软弱的地基不但承载能力差，并且在龙卷风极强的吸附力下，对房屋建筑的嵌固力和基础抗拔能力弱，所以在选择建筑区域时应该避开龙卷风故道、风口、河滩等地，选取平坦地面、稳定基岩且周边有防护林或房屋较密集的地段建造房屋；^[7]

(4) 加强监督管理：违章建筑安全性极低，且当建筑之间距离过小时，会使其坐落于龙卷风尾流区，造成更加严重的破坏^[4]。国家应该严格实施相关政策，如《土地管理法》、《城乡规划法》、《村庄和集镇规划建设管理条例》等，加强对违章建筑、以及建筑之间的距离的管理，尽量减少人民的损失；

(5) 设置紧急避难所：为了使人民的生命安全能够多一层保障，工程师们可以设置避难所，将局部加强的内部小房间、楼梯间、地下室设置为安全岛，也是一种值得提倡的工程应急预案。^[6]

我们应该未雨绸缪，提前做好准备工作，提高建筑和房屋的安全性、稳固性和抗风承载力，切实保障人民的生命安全。

谢谢大家。

8. 模型总结与评价

- (1) 优点：该模型在分析风险评估搜集了全面而且可靠的数据，对每个省份进行概率分析，模型的运算时间复杂度较低，考虑的因素从台风的危险性和我国沿海城市的脆弱性进行评估，有 6 个指标且都具有代表性，使得预测的模型评估具有普遍性。
- (2) 缺点：该模型在预测风险评估时使用的方法过于简便，在精准度上还不够精准，同时每个城市的抗灾能力也是有差异的，而且抗灾能力对于抗击台风还是有很大的作用的，这会使得模型建立更加精确。在对“蝴蝶”进行评估的时候，没有用所有的历史数据，而是只用了最近 20 年作为评估标准，未考虑年代对台风变化与大陆板块的影响。
- (3) 改进方法：对模型考虑的时候可以加上台风对各省的降雨量和抗灾能力对于抗击台风能力。在第三问中用朴素贝叶斯分类器进行对各省风险评估的时候可以改变 Laplace 平滑系数进行调优。整体的概率模型在求出期望的时候应可以加上对各省概率的损失估计。并且可以考虑台风年代对各个省市的影响。

参考文献

- [1] 殷洁, 戴尔阜, 吴绍洪. 中国台风强度等级与可能灾害损失标准研究 [J]. 地理研究, 2013.
- [2] 牛海燕. 中国沿海台风灾害风险评估研究 [D]; 华东师范大学, 2012.
- [3] 魏章进, 隋广军, 唐丹玲. 基于聚类与回归方法的台风灾情统计评估 [J]. 数理统计与管理, 2014.
- [4] 高榕. 中国龙卷风特性统计分析及灾后建构物调查研究 [D]; 北京交通大学, 2018.
- [5] 殷洁, 戴尔阜, 吴绍洪. 中国台风灾害综合风险评估与区划 [J]. 地理科学, 2013.
- [6] 甘文举, 童小龙, 宋彬彬. 龙卷风多发区低层房屋抗风概念设计探讨 [J]. 小城镇建设, 2011.
- [7] 甘文举. 低层房屋龙卷风荷载分析及抗风设计研究 [D]; 湖南大学, 2009.

附录

见附录文件