

报名号#1077

# 2008 年 第一届“数学中国杯”

## 数学建模网络挑战赛

### 承 诺

我们仔细阅读了首届“数学中国杯”数学建模网络挑战赛的竞赛规则。

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。

我们允许数学中国网站([www.madio.net](http://www.madio.net))公布论文，以供网友之间学习交流，数学中国网站以非商业目的的论文交流不需要提前取得我们的同意。

我们的参赛报名号为：1077

参赛队员（签名）：

队员 1：张海龙

队员 2：张伟腾

队员 3：刘锋

参赛队教练员（签名）：

参赛队伍组别：大学组

报名号#1077

# 2008 年 第一届“数学中国杯”

## 数学建模网络挑战赛

### 编 号 专 用 页

参赛队伍的参赛号码：（请各个参赛队提前填写好）：

1077

竞赛统一编号（由竞赛组委会送至评委团前编号）：

竞赛评阅编号（由竞赛评委团评阅前进行编号）：

# 2008 年 第一届“数学中国杯” 数学建模网络挑战赛

## 题 目 沙尘暴的分析与预测

关 键 词 气候影响指数模型、沙尘天气气象等级评估模型、沙尘暴风险预测模型、非线性回归、图像适线法

### 摘 要：

在本文中，我们从数据信息和机理两个角度出发，着重分析了各个气候因素对沙尘暴的发生次数的影响，发现影响沙尘暴的最主要因素是月平均风速、大风日数、气温变化、下垫面风蚀因子、土壤湿润度。

针对问题一，经过综合权衡，将上述五种气候因素对沙尘暴次数的影响综合为一个量——气候影响指数对沙尘暴次数的影响，建立了气候影响指数模型，从而得出以上述五个气候因素为自变量，以气候影响指数为因变量的函数关系式。然后利用 **Curve Expert** 回归软件进行非线性回归，得到沙尘暴次数关于气候影响指数的回归曲线。这样就建立起沙尘暴次数关于各气候因素的函数表达式。这里特别要突出三点，一是类比温度对气候影响指数的影响的量化方法，来量化大风日数对气候影响指数的影响，使得模型更为贴合实际；二是回归分析时通过剔除个别异常点，优化了回归效果；三是考虑到当月的土壤湿润度还受到上月的影响，将 12 个月份分成四类分别刻画土壤湿润度对沙尘暴次数的影响。

针对问题二，参照国家标准，将沙尘天气分成四个等级，即大风、浮沉、扬尘和沙尘暴。通过探究能见度与辐射量之间的函数关系，再结合四个等级的划分依据即可完成对于给定辐射量下沙尘天气的评价。对沙尘暴风险进行预测时，我们引入灾害熵的概念，应用图像适线法，建立了沙尘暴风险预测模型，从而可以实现一定的预测时间段上遭遇的沙尘暴的爆发强度等级的预测，并根据预测结果发出相应的预警信号。

参赛队号 1077

参赛密码             
(由组委会填写)

所选题目 C 题

## 问题重述

沙尘暴是一种灾害性天气，近年来频繁发生，对环境和人们的生产、生活都造成了影响甚至危害。对沙尘暴的发生条件需要加以研究，以便人们进行有效的预测和预防。请为沙尘暴的发生建立数学模型，具体问题如下：

问题一：对沙尘暴的发生次数进行分析，建立相应的数学模型，使人们可以根据相关气象数据的监测结果，估计出沙尘暴在某地区某段时间内发生的次数。

问题二：评估沙尘天气气象等级，预测在近期内观测站所在地发生沙尘暴的风险，并在高风险时发出警报信息。

## 问题假设

1. 沙尘暴的影响因素在本论文中仅考虑与月平均风速、大风日数、气温变化、下垫面风蚀因子、土壤湿润度、太阳辐射等五个因素有关，降水量、蒸发量等其它因素对其影响均可划归到这五类的作用中去。
2. 沙尘天气发生时仅考虑一种沙尘天气，而不考虑发生过程中可能出现的转变情况。
3. 题目中所给的数据是真实反映客观实际的，包括经数据提炼分析得出的大风日数只在同年度不同月份间发生变化，而不随年度变化，我们也认为其是准确可靠的。
4. 假设所给的月平均风速的数据为 10~12m 处的数据。

## 符号说明

由于在模型中我们引用的较多公式出自参考文献，计算公式涉及到一些专业的符号或术语，因而在符号的使用过程中不加更改地对其进行了引用。

符号说明	
$F_i$	第 <i>i</i> 月份沙尘暴发生次数
$A_i$	第 <i>i</i> 月份下垫面风蚀气候因子
$CD_i$	第 <i>i</i> 月份气候影响指数
$B_i$	第 <i>i</i> 月份大风日数

报名号#1077

$D_i$	第i月份土壤湿润指数
$ETP_i$	第i月份月潜在蒸发量
$Q_i$	月降雨量
$p_i$	月平均气压 (mb)
$d_i$	月的天数
$W_i$	在10~12m高度处观测的月平均风速 (m/s)
$t_i$	月平均气温 (°C)
$w_{oi}$	在温度为 $t_i$ 时的饱和水汽压 (mmHg)
$h_i$	月平均相对湿度
$R$	能见度
$\lambda_1$	消光系数
$\lambda_2$	消光率
$\lambda_3$	穿透率
$H$	沙尘天气时空气中尘埃的作用距离
$S$	灾损系列
$S_a$	灾损系列的均值
$C_v$	灾损系列的变差系数；
$C_s$	灾损系列的偏态系数
$\phi$	离均系数

## 问题分析

### (一) 问题一的分析：

由于题目数据是以按月给出的，所以在具体考虑时，以月为单位考虑气候因素对沙尘暴发生次数的影响。以一个观测站（鼎新 52446）为研究对象，评估其某一特定月份实测的各个气候因素（月平均风速、大风日数、平均等）对该月份沙尘暴发生次数的影响，经过综合权衡，得出以任一特定月份的沙尘暴发生次数为因变量，以该月份各个测得指标为自变量的确定函数模型，以达到对该观测站某年某月的沙尘暴发生次数进行估计的目的。

经过权威认证，气候因素对沙尘暴次数的月际变化将起到非常重要的作用。首先从沙尘暴天气发生的4个基本条件出发，经过定性分析，风速和湿润度分别构成了形成沙尘暴天气的动力条件和起沙的可能性，地面温度是形成沙尘暴的热力条件，下垫面风蚀气候因子（沙源）则是沙尘暴天气形成的物质条件。而且，经过数据分析和实际考察，沙尘暴形成还与大风日数存在着密切的联系。且月平均风速、大风日数、地面温差、下垫面风蚀气候因子等4项指标与沙尘暴次数正相关，而下垫面湿润度越大与沙尘暴次数负相关。

针对鼎新的实际情况，我们建立了气候影响指数模型来刻画各气候因素对沙

尘暴次数的月际变化的作用，将各气候因素所起的作用综合反映为一个量对沙尘暴次数的影响，即气候影响指数 $CD_i$ 。之后，针对该模型求出的气候影响指数与沙尘暴天数的图像关系进行回归分析，得出以沙尘暴天数为因变量，以气候影响指数为自变量的回归函数。

在具体确定气候影响指数模型时，需反映不同气候因素对气候影响指数的贡献，文中主要确定了下垫面风蚀气候因子、大风日数和土壤湿润度对气候影响指数的贡献大小：

1、下垫面风蚀气候因子：风蚀气候侵蚀力是气候条件影响风蚀的可能程度的量度，是土地沙化和土壤风蚀评判程度的重要指标，是衡量沙源对沙尘暴次数影响程度的根本指标，国际上通常采用风蚀气候因子表示。在这里我们借用联合国粮农组织[1]给出的公式：

$$A_i = \frac{1}{100} \left( \frac{ETP_i - Q_i}{ETP_i} \right) d_i W_i^3 \quad (1)$$

$$ETP_i = \begin{cases} 0.19(20 + t_i)^2(1 - h_i) & t_i \geq -20 \\ 0 & t_i < -20 \end{cases} \quad (2)$$

$$w_{oi} = \begin{cases} 1.3694 \times 10^9 \exp\left(-\frac{5328.9}{273.2 + t_i}\right) & 0^\circ\text{C} \leq t_i \leq 30^\circ\text{C} \\ 2.6366 \times 10^{10} \exp\left(-\frac{6139.8}{273.2 + t_i}\right) & -40^\circ\text{C} \leq t_i \leq 0^\circ\text{C} \end{cases} \quad (3)$$

其中， $i$ 是月份号； $A_i$ 即为第 $i$ 月份下垫面风蚀气候因子； $ETP_i$ 为月潜在蒸发蒸腾； $Q_i$ 为月降雨量， $d_i$ 是月的天数， $W_i$ 是在10~12m高度处观测的月平均风速（m/s）， $t_i$ 是月平均气温（℃）， $w_{oi}$ 是在温度为 $t_i$ 时的饱和水汽压（mmHg），而 $h_i$ 是月平均相对湿度。

结合所给出的参考数据，可以得出该观测站任一月份的下垫面风蚀气候因子对沙尘暴次数的影响指数。

## 2、大风日数：

经过相关性分析，可以看出大风日数与沙尘暴次数相关系数最大，为0.832，月温度变化与沙尘暴次数相关性次之，相关系数为0.701。而温度变化对气候影响指数的影响程度已由权威资料论证出，可参照式（1）中温度对其的影响程度的刻画类似的将大风日数对气候影响指数的影响作为一个修正项反映到式（1）中去。

## 3、土壤湿润度 $D_i$ 的讨论：

经过查阅相关资料[2]，月土壤湿润度为当月降水量与当月潜在蒸发量之比，

$$\text{即： } D_i = Q_i / ETP_i \quad i=1,2\cdots 12$$

其中 $Q_i$ 为*i*月份降水量(mm)， $ETP_i$ 为潜在蒸散量(mm)。月潜在蒸散量 $ETP_i$ (mm)采用上式(2)计算：

结合上述公式，利用相关数据即可确定出 $D_i$ 的大小。

而此时确定的 $D_i$ 只是一个理论值，事实上，由于土壤湿润度状况对沙尘暴次数的影响有一定的滞后性，所以本月实际土壤湿润度在很大程度上受到上月的土壤湿润度的影响。为此，将相邻月份土壤湿润状况的变化划分为四种情况：一是土壤越来越湿润的变化过程，在这一过程中上一个月份的土壤湿润度相对本月较低，使本月的实际土壤湿润度低于由降水和蒸发所决定的土壤湿润度，从而使本月沙尘暴次数有所增大；第二种情况是土壤越来越干燥的变化过程，在这一过程中土壤的实际湿润度除受本月降水与蒸发的影响外，还受上一个月土壤湿润度的影响，从而使本月沙尘暴次数有所减小；第三种情况是当月的湿润度与上月相比变化不大，此时可认为上月的湿润度对本月不产生影响；最后一种是在每年的11、12、1月份由于北方天气寒冷以致地面封冻，土壤水分不易散失，而使得湿润度对沙尘暴发生次数的影响较小，即对气候影响指数的影响不如上三种情况显著。

所以在对某特定月份考虑时，先根据绘出的土壤湿润度的图像，得出其属于四种变化情况的哪一种，针对其不同类别，对理论土壤湿润度进行修正，从而分别建立函数关系式来刻画不同的相邻月份土壤湿润变化状况对气候影响指数的影响，进一步得出以各气候因素为自变量，以气候影响指数为因变量的各月份的函数表达式。

## (二)、问题二的分析：

1、按照国家气象局对沙尘天气气象等级的划分，依据沙尘天气的能见度和风速大小的差别，将其划分成四级：大风、浮尘、扬尘、沙尘暴[3]。其等级划分评价依据[]为：

	能见度 (km)	风速 (m/s)
大风	10-30	
浮尘	1-10	<4
扬尘	1-10	≥ 4
沙尘暴	< 1	

我们只需针对某次沙尘天气的实测数据求出其能见度，根据该评价依据，即可非常简易的确定出其等级。

**能见度的确定：**空气中的尘埃等会对太阳的辐射量产生很大的影响，而单位体积的尘埃数直接决定了其能见度的大小，这样在能见度与太阳的辐射量之间肯定存在着某种必然的联系，我们现在就要经过机理分析，找出它们之间的这种关系，再根据某次沙尘天气实测数据中给出的太阳的辐射量，则可得出此时的能见度。

## 2、沙尘暴的风险预测：

风险的定义：根据相关资料定义[4]，灾害风险的含义是灾害损失大于某种程度的发生概率大小。由此，可以看出风险不仅与概率有关，还与灾害强度有关。

灾害是一种随机事件，它的发生具有不确定性，因而可以用熵来进行描述。灾害以最无序的方式在各地发生，意味着灾害熵达到了极大值，在给定的约束条件下，当灾害熵取极大值时，灾损系列（各月份的灾害损失程度，用能见度的大小来描述，因为能见度越大，说明受灾程度越小，反之，受灾程度则越大）服从p-III型分布[5]。灾害损失的重现期指未来某一程度的灾害损失出现的平均时间间隔。根据实测的辐射数据得出每月能见度大小，组成一个灾损系列，利用p-III型分布函数，可以计算未来具有一定重现期的灾害损失大小，以达到预测沙尘暴风险的目的。

## 模型准备

### 最大熵原理得出灾损系列的分布：

熵增加原理表明：在孤立或绝热的条件下，系统自发地从非平衡态向平衡态发展的过程是一种熵增加的过程，平衡态与最大熵相对应，在平衡态系统的状态最混乱、最无序。

灾害是一种随机事件，它的发生具有不确定性，因而可以用熵来进行描述，灾害以最无序的方式在各地发生，意味着灾害熵达到了极大值，因此可以用最大熵原理来确定一个地区一定时段内灾损系列(实测的灾害损失资料)的分布。

从灾损系列可知，损失小的灾害出现机会多，损失大的灾害出现机会少，且灾害损失均大于 $a_0$ 。此时约束条件式(2)、(3)成为：

$$\text{灾害损失均大于 } a_0 \Rightarrow \int_{a_0}^{\infty} p(x) dx = 1 \quad (5.1)$$

$$\text{已知函数 } y=x-a_0 \text{ 的平均值给定} \Rightarrow \int_{a_0}^{\infty} (x-a_0) p(x) dx = E(x-a_0) \quad (5.2)$$

$$\text{已知函数 } y=\ln(x-a_0) \text{ 的平均值给定} \Rightarrow \int_{a_0}^{\infty} \ln(x-a_0) \times p(x) dx = E[\ln(x-a_0)] \quad (5.3)$$

上式右边的值是给定的，如果满足这些约束条件，则熵取极大值时的分布函数为：

$$p(x) = \exp[-\beta_0 - \beta_1(x-a_0) + \beta_2 \ln(x-a_0)] \quad (5.4)$$

由(5.1)与(5.4)可以得到：

$$\exp(\beta_0) = \int_{a_0}^{\infty} \exp[-\beta_0 - \beta_1(x-a_0) + \beta_2 \ln(x-a_0)] dx = \beta_1^{\beta_2-1} \Gamma(\beta_2+1) \quad (5.5)$$

将(5.5)代入(5.4)中得到

$$p(x) = \frac{\beta_1^{\beta_2+1}}{\Gamma(\beta_2+1)} (x-a_0)^{\beta_2} \exp[-\beta_1(x-a_0)] \quad (5.6)$$

可以看出，式(5.6)为p-III型分布函数。



## 模型建立

### 问题一的模型建立与求解

根据问题分析，欲评估各气候因素对沙尘暴次数综合影响，首先确定气候因素对沙尘暴次数的影响指数，即气候影响指数。

### 模型一、气候影响指数模型

模型基本结构为[6]：

$$CD_i = A_i \times (W_i/D_i^2) + \Delta t_i \quad i=1,2,\dots,12 \quad (1)$$

其中*i*为月份， $CD_i$ 是用来描述气候因素对沙尘暴次数的影响指数，即第*i*月份气候影响指数； $A_i$ 为第*i*月份下垫面风蚀气候因子； $W_i$ 为第*i*月份风速月平均值； $D_i$ 是第*i*月份土壤湿润指数； $\Delta t_i$ 为第*i*月份与第*i*-1月份地面平均温度差值。

根据问题分析中对于下垫面风蚀气候因子、大风日数及土壤湿润指数的分析，具体各气候因素在公式（1）中的取值确定如下：

（1）风速平均值 $W_i$ 可以由已知数据直接给出，对于 $\Delta t_i$ 的确定，可以用本月与前月的月平均气温的差值来近似代替地面平均温度差值。

（2）下垫面风蚀气候因子：

在这里我们借用Chepil 公式：

$$A_i = \frac{1}{100} \left( \frac{ETP_i - Q_i}{ETP_i} \right) d_i W_i^3 \quad (2)$$

$$ETP_i = \begin{cases} 0.19(20 + t_i)^2(1 - h_i) & t_i \geq -20 \\ 0 & t_i < -20 \end{cases} \quad (3)$$

公式（2）、（3）中，*i*是月份号； $A_i$ 即为第*i*月份下垫面风蚀气候因子； $ETP_i$ 为月潜在蒸发蒸腾； $Q_i$ 为月降雨量； $P_i$ 是月平均气压（mb）， $d_i$ 是月的天数， $W_i$ 是在10~12m高度处观测的月平均风速（m/s）， $t_i$ 是月平均气温（℃）， $w_{oi}$ 是在温度为 $t_i$ 时的饱和水汽压（mmHg），而 $h_i$ 是月平均相对湿度。

当潜在蒸发量( $ETP_i$ )小于降水 $Q_i$ 时,由于地面不会产生实际的沙尘暴,虽然风蚀的潜在能力仍然存在,但可认为下垫面风蚀气候因子为0。

而饱和水汽压 $w_{oi}$ 又可表示为：

$$w_{oi} = \begin{cases} 1.3694 \times 10^9 \exp\left(-\frac{5328.9}{273.2 + t_i}\right) & 0^\circ\text{C} \leq t_i \leq 30^\circ\text{C} \\ 2.6366 \times 10^{10} \exp\left(-\frac{6139.8}{273.2 + t_i}\right) & -40^\circ\text{C} \leq t_i \leq 0^\circ\text{C} \end{cases} \quad (3)$$

将实测数据代入，得出各月份风蚀气候因子 $A_i$ ，详见表1。

### (3) 大风日数影响程度的确定：

大风日数用符号**B**来表示。

根据问题分析，大风日数与沙尘暴次数正相关，且相关系数 $\text{cov}_1$ 为0.832最大，温度变化与沙尘暴频的相关系数 $\text{cov}_2$ 为0.701。

首先单个对大风日数和沙尘暴次数进行回归，综合各种回归曲线（多项式拟合、指数函数拟合、对数函数拟合）后，所得回归结果的可决系数、剩余标准差的相对大小，取可决系数最大、剩余标准差最小的指数拟合曲线

$$y = 0.45B^{0.912}$$

此式说明了在其他影响因素固定的情况下，大风日数和沙尘暴次数间呈指数关系。此时，我们同样把大风日数对沙尘暴次数的影响综合到气候影响指数中去。那么，依照式（1）中温度变化量对气候影响指数 $CD_i$ 的作用，在考虑大风日数对气候影响指数 $CD_i$ 的作用后，式（1）应修正为：

$$CD_i = A_i \times (W_i/D_i^2) + \Delta t_i + kB^{0.912} \quad i=1,2,\dots,12 \quad (1)$$

其中K值我们取为 $\frac{\text{cov}_1}{\text{cov}_2} = 1.187$ ，以反映与沙尘暴次数间相关系数的不同而导致的其影响程度上的不同。

### (4) 土壤湿润度 $D_i$ 的确定：

根据问题分析，i月份土壤湿润度

$$D_i = Q_i / ETP_i$$

其中 $Q_i$ 为i月份降水量（mm）， $ETP_i$ 为潜在蒸散量（mm）。月潜在蒸散量 $ETP_i$

(mm) 采用上式 (3) 计算：

代入相关参考数据，可确定出理论土壤湿润度  $D_i$  的大小。结果详见表1。其折线图如图1所示。

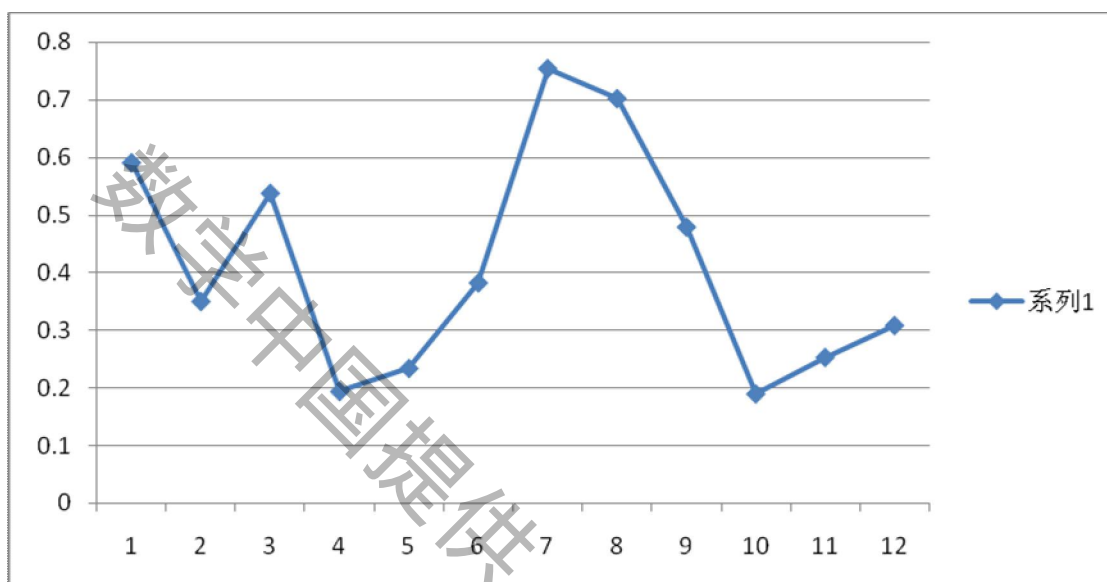


图1 土壤湿润度指数月季变化图

下面基于问题分析中的讨论，结合得出的土壤湿润度指数月季变化图，针对不同月份在相邻月份土壤湿润度变化上的差异，对理论土壤湿润度进行修正，使其更符合实际情况，进一步得出不同月份不同的以各气候因素为自变量，以气候影响指数为因变量的函数表达式。

湿润度指数与沙尘暴次数存在着一定的负相关,并且月际之间存在着显著的变化。由图像可以看出：

(1) 3、6、7月份为土壤比上一月份湿润的月份，即符合问题分析中的第一种情况——土壤越来越湿润，此时我们将 (1) 式修正为：

$$CD_i = A_i \times (2W_i/D_i^2 + 1/D_i^*) + \Delta t_i + 1.187B^{0.912} \quad i=3,6,7 \quad (4)$$

这里  $CD_i$  为土壤越来越湿润过程对应月份的气候影响指数，其中  $D_i^*$  表示前月的土壤湿润指数,其它变量同前。

(2) 2、4、9、10为土壤比上一月份干燥的月份，即符合问题分析中的第二种情况——土壤越来越湿润，此时我们将 (1) 式修正为：

$$CD_i = A_i \times (W_i/D_i^2 - 1/D_i^* \times W_i/D_i) + \Delta t_i + 1.187B^{0.912} \quad i=4,9,10 \quad (5)$$

这里  $CD_i$  为土壤越来越干燥过程对应月份的气候影响指数，变量同前。

(3) 5、8月份为相对前一月份变化不大的月份，属于第三种情况，不需要进行修正，仍为(1)式。

(4) 11、12、1月份处于北方的冷冻季节，其上月的土壤湿润度对当月的土壤湿润度没有影响，属于第四种情况，此时我们将(1)式修正为：

$$CD_i = A_i \times (W_i/D_i) + \Delta t_i + 1.187B^{0.912} \quad i=11,12,1 \quad (6)$$

变量同前。

综合上述对各月沙尘暴次数所建立的气候影响指数模型,得到了应用于所研究地区(鼎新)全年的气候影响指数模型,其表达形式如下:

$$CD_i = \begin{cases} A_i \times (W_i/D_i^2) + \Delta t_i + 1.187B^{0.912} & i=5,8 \\ A_i \times (W_i/D_i^2 - 1/D_i^* \times W_i/D_i) + \Delta t_i + 1.187B^{0.912} & i=2,4,9,10 \\ A_i \times (2W_i/D_i^2 + 1/D_i^*) + \Delta t_i + 1.187B^{0.912} & i=3,6,7 \\ A_i \times (W_i/D_i) + \Delta t_i + 1.187B^{0.912} & i=11,12,1 \end{cases} \quad (7)$$

变量同前。

将鼎新的实测数据代入，可以求得各月份的下垫面风蚀气候因子、土壤湿润度、气候影响指数  $CD_i$ ，如下

月份	土壤湿润度	下垫面风蚀气候因子	气候影响指数	沙尘暴次数
1	0.5902	4.8485	28.2932	1.5
2	0.35	4.003	81.7033	1.311111
3	0.5373	4.4573	129.006	1.466667
4	0.194	5.1592	419.467	2.088889
5	0.234	3.7502	230.235	1.955556
6	0.3821	2.9235	139.28	2.022222
7	0.7531	2.3393	35.1411	1.733333
8	0.7012	1.8958	10.4908	0.844444
9	0.4788	1.5162	5.6842	0.288889
10	0.1899	1.9685	106.954	0.488889
11	0.2531	3.8611	43.2598	1.244444

12	0.3083	9.7837	105.242	1.545455
----	--------	--------	---------	----------

表 1-土壤湿润度、下垫面风蚀气候因子、气候影响指数结果一览表

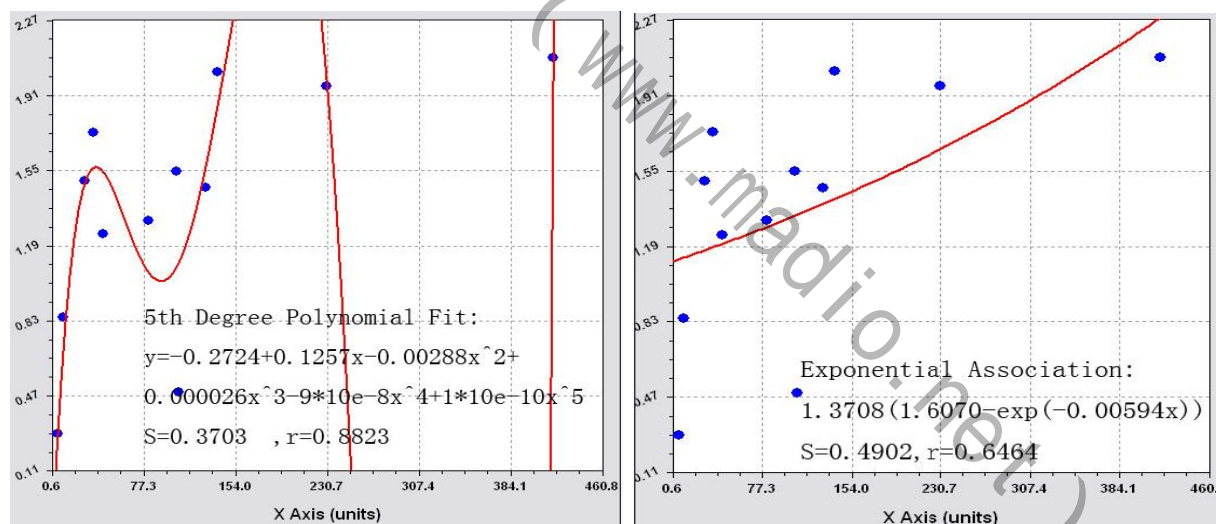
## 2、回归分析

现在我们得出了四种气候因素对于沙尘暴次数  $F_i$  影响的总和度量值，即气候影响指数  $CD_i$ ，之后我们须建立气候影响指数  $CD_i$  与沙尘暴次数  $F_i$  的函数关系式，以进一步将气候因素对沙尘暴次数的影响定量化。

因此，考虑采用非线性回归的方法建立以  $CD_i$  为自变量、以  $F_i$  为因变量的函数关系式。利用 **CurveExpert** 回归软件进行曲线拟合。在确定其回归类型时，我们采用对照法，

分别针对各种回归类型得出其回归方程，每个回归方程对应一组不同的标准偏差值  $s$  和相关系数  $r$ ，通过比较，找出其中标准偏差值最小，相关系数最大的一组值所对应的回归方程。

这里分别针对四种回归类型——多项式回归、指数回归、线性回归等进行曲线拟合，其拟合曲线如下：



报名号#1077

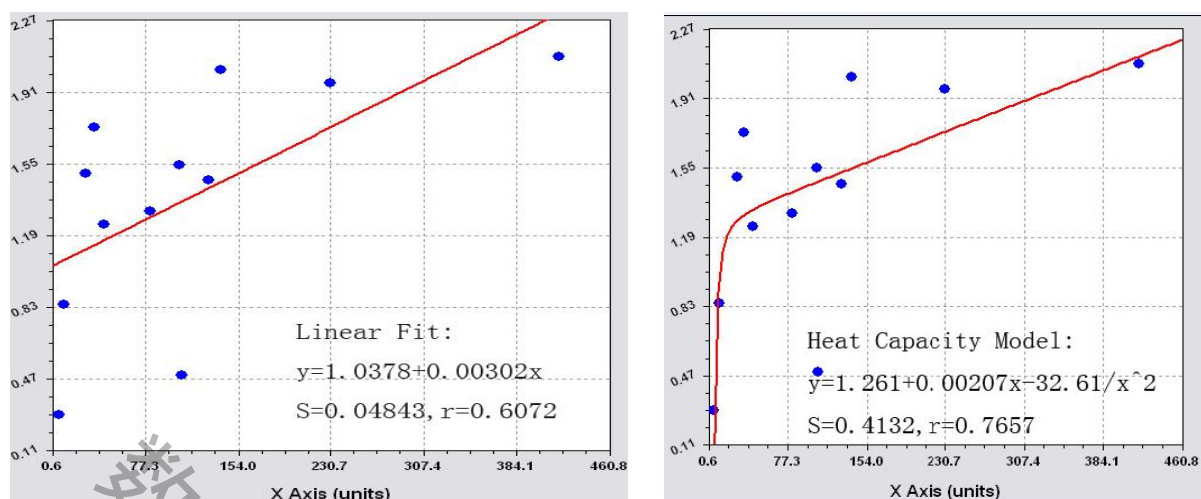


图 2 拟合曲线对比图

通过标准偏差值  $s$  和相关系数  $r$  大小的比较，可以看出，五次多形式拟合后，其标准偏差最小，相关系数最大，但由于其抖动很大，所以我们采用标准偏差较小，相关系数较大，同时抖动较小的 **heat capacity model** 对其进行拟合。

### 剔除异常点

在拟合前首先对原始数据点进行残差分析（见下图），发现有一个异常点（106.954, 0.488889），对曲线拟合影响很大。由于该回归参数估计是残差平方和最小的估计，当异常值远大于其它观测值时，会导致拟合值偏向异常值的拟合错误，使回归模型失真。因此，我们将该异常点剔除。

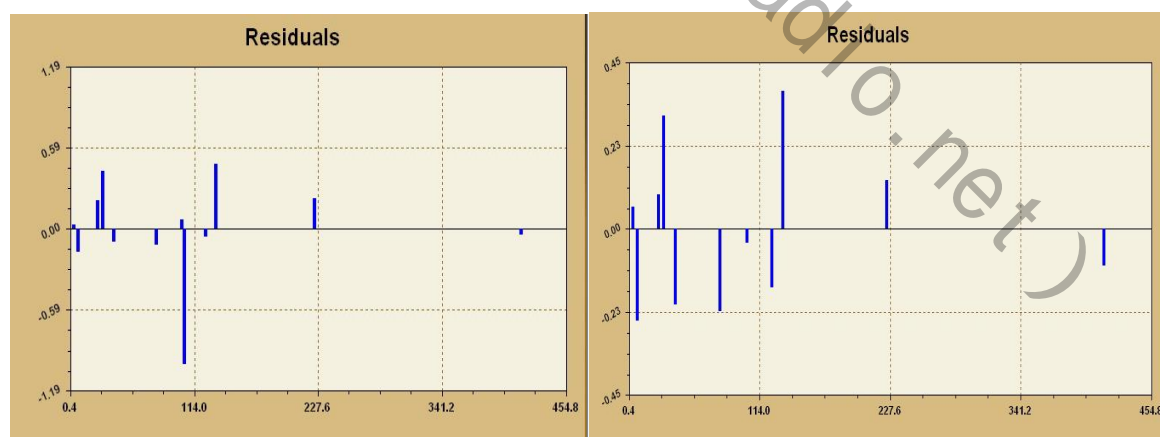


图 3 剔除前残差图

剔除后残差图

将异常点剔除后，经过曲线拟合得出其回归方程为

$$F_i = 1.392 + 0.00188 \times CD_i + \frac{-37.8763}{CD_i^2} \quad i=1,2,3 \dots 12 \quad (8)$$

相应可以得出拟合曲线的标准偏差  $S=0.236$ ，相关系数  $r=0.919$ 。不难发现其标准偏差  $S$  相对较小，相关系数很大，可以达到 0.9 以上，说明该曲线的回归效果比较好，回归是显著的。

## 问题二的模型建立与求解

### 1、天气气象等级评价模型

首先，经过查阅资料得出天气气象等级评价依据表如下：

	能见度 (km)	风速 (m/s)
大风	10-30	
沙尘	1-10	<4
扬尘	1-10	>4
沙尘暴	<1	

表 2-天气气象等级评价依据表

根据问题分析：我们现在所要做的就是如何根据参考数据得出所要评估的沙尘天气的可见度。具体求解如下：

引入参数能见度  $R$  消光系数  $\lambda_1$ 、消光率  $\lambda_2$ 、穿透率  $\lambda_3$ ，其定义参见符号说明。而经查阅资料可知，它们三者之间有着密切的联系[7]：

$$R = 3.912 / \lambda_1 \quad (9)$$

$$\lambda_2 = 1 - \lambda_3 \quad (10)$$

消光系数是指单位传播距离下的相对消光量，而对于本题所要讨论的是空气中的沙尘消光大小，因此这时所要考虑的传播距离是沙尘天气时空气中尘埃的作用距离，相对消光量即为其消光率。由此可以确定出消光系数与消光率之间的关系：

$$\lambda_1 = \lambda_2 / H \quad (11)$$

其中  $H$  为沙尘天气时空气中尘埃的作用距离，一般取一常量，经借鉴相关权威文献，在此取设定  $H$  为 210m。

$$\text{而穿透率 } \lambda_3 = \frac{\text{穿透强度}}{\text{入射强度}} * 100\% \quad (12), [8]$$

又因光强度与光能量成正比，则 (12) 式可等效为

$$\lambda_3 = \frac{\text{直射辐射量}}{\text{总辐射量}} * 100\% = \frac{w_1}{w_2} * 100\% \quad (13)$$

这样将 (9)、(10)、(11)、(12) 四式相结合，得出能见度关于辐射量的函数关系式：

$$R = 3.912 \times H / (1 - w_1 / w_2) \quad (14)$$

根据该公式，若某次沙尘天气的太阳辐射量实测数据已知，则可直接计算

出该沙尘天气的能见度，结合给定的风速，参考天气气象等级评价依据表则可评价出其天气气象等级。

## 2 沙尘暴风险预测模型

结合问题分析，我们引入灾害熵[5]建立沙尘暴风险预测模型，对其风险进行预测。

基于模型准备中所阐释的理论，如果灾损系列用  $s$  来表示，那么灾害熵取极大值时， $s$  服从  $p$ -III 型分布(经过整理)：

$$p(s) = \frac{\beta^\alpha}{\Gamma(\alpha)} (s - a_0)^{\alpha-1} \exp[-\beta(s - a_0)] \quad (15)$$

式中  $\alpha$ 、 $\beta$ 、 $a_0$  是三个待定的未知参数，它们与灾损系列统计参数  $\bar{S}$ 、 $C_v$ 、 $C_s$  的关系为：

$$\alpha = \frac{4}{C_s^2}; \quad \beta = \frac{2}{\bar{S} C_v C_s}; \quad a_0 = \bar{S} \left(1 - \frac{2C_v}{C_s}\right) \quad (16)$$

式中  $\bar{S}$ ——灾损系列的均值； $C_v$ ——灾损系列的变差系数； $C_s$ ——灾损系列的偏态系数。

在未来灾害分析中，需要求出指定概率  $P$  所对应的灾害损失  $SP$  即

$$p = p(s > SP) = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_{SP}^{\infty} (s - a_0)^{\alpha-1} \exp[-\beta(s - a_0)] ds \quad (17)$$

式中  $P$  由决策部门根据风险水平或抗灾能力的大小选定， $P$  一经确定， $SP$  即为估计的具有一定概率(或重现期  $T$ )的灾害损失。这里  $T = 1/P$ ，如要求  $P = 1\%$  的灾害损失  $S_{1\%}$ ，那么重现期  $T = 1/1\% = 100$  年， $S_{1\%}$  即为百年一遇的灾害损失。因此，重现期指未来某一程度的灾害损失出现的平均时间间隔。

为了根据分布函数式(17)计算  $SP$ ，通过推导，可令  $\Phi = (SP - \bar{S})/(\bar{S} \times C_v)$ ，式中  $\Phi$  称为离均系数，有表可查[]，那么

$$SP = (\Phi C_v + 1) \bar{S} \quad (18)$$

因此，在  $\bar{S}$ 、 $C_v$ 、 $C_s$  和  $P$  已知的条件下，即可求得  $SP$ 。这里  $\bar{S}$ 、 $C_v$ 、 $C_s$  可根

$$\text{据灾损系列来计算: } \bar{S} = \frac{1}{n} \sum_{i=1}^n S_i, \quad C_v = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (S_i - \bar{S})^2} \quad (19)$$

而  $C_s$  用  $C_s = m C_v$  来估计( $m$  为倍数)。

### 图解适线法

由于用上述方法计算的  $\bar{S}$ 、 $C_v$ 、 $C_s$  误差较大，因此一般用图解适线的方法



来确定  $\bar{S}$ 、 $C_v$ 、 $C_s$ ，而把原先计算的  $\bar{S}_0$ 、 $C_{v0}$ 、 $C_{s0}$  作为图解适线时的参数初试值。

图解适线时，首先把灾损系列按大到小排列  $S_1 \geq S_2 \geq S_3 \geq \dots S_n$ ，那么，它们各项的概率为：

$$p_i = i/(n+1) \quad (20)$$

然后把灾损系列和对应的  $p_i$  点绘在概率格纸上，得到离散的灾损点子，再根据  $\bar{S}_0$ 、 $C_{v0}$ 、 $C_{s0}$  和式(18)计算 SP，把 SP 和  $p_i$  也点绘在概率格纸上，得到一条连续的理论曲线。如果理论曲线和灾损点子配合不好，那么改变  $\bar{S}$ 、 $C_v$ 、 $C_s$  重新适线，直至两者吻合为止，此时的  $\bar{S}$ 、 $C_v$ 、 $C_s$  即为待求的统计参数。

最后根据指定概率 P 和最佳拟合的理论曲线，即可求得具有一定重现期的灾害损失 SP。

在本题中用能见度 R 来衡量灾损系列的大小。因为能见度越大，沙尘天气等级越低，受灾程度越小，反之则受灾程度越大。说明能见度 R 与受灾程度呈负相关的关系，并且能见度 R 直接决定了受灾程度的大小。

这样针对本题目给出所想要预测的时间段长度 T（即模型中的重现期），即可得出与之相对应的概率  $P=1/T$ ，结合式（17），即可求出灾害损失的大小，即能见度 R 的大小。从而可以在一定的概率下对沙尘暴的爆发强度等级进行预测。

根据能见度的大小，将沙尘暴的爆发等级及其相应的预警信号[9]分成三类：

沙尘暴的爆发等级	预警信号	能见度 ( km )
强	红	$\leq 0.05$
中	橙	0.05—0.5
弱	黄	0.5—1

表 3 沙尘暴的爆发等级及其相应的预警信号划分表

结合沙尘天气等级评价模型中对于能见度的计算，可以得出 12 个月份平均能见度，按照平均能见度由大到小的顺序排序，并确定出与之相对应的概率值，借助图像适线法的一般原理，最终得出理论曲线与点子最佳拟合时，

$\bar{S}=1.026$ ， $C_s=1.2$ ， $C_v=0.167$ 。结合式（16），求出  $\alpha=2.78$ ， $\beta=9.73$ ， $a_0=0.7404$ ，

借助式（17）即可得出一定预期时间段内会遭遇沙尘暴的能见度大小，此时可根据沙尘暴的爆发等级及其相应的预警信号划分表，确定出爆发等级及不发或者发

出何种颜色的预警信号。

基于以上理论，及求解出的求出的  $\alpha = 2.78, \beta = 9.73, a_0 = 0.7404$ ，可以得出在 12 个不同的预期时间段分别会遭遇的沙尘暴的能见度大小和相应的沙尘暴的爆发等级及其预测信号。详见下表

其表示的具体含义是：在一个月内会遭遇能见度为 1.32 的沙尘天气，在两个月内会遭遇能见度为 0.1245 的中等沙尘暴天气，此时发出橙色信号，依次类似，在一年内会遭遇能见度为 0.0452 的强沙尘暴天气，此时发出红色信号。

概率值	能见度 (km)	沙尘暴的爆发等级	预警信号	概率值	能见度 (km)	沙尘暴的爆发等级	预警信号
1	1.32	未	未	1/7	0.0586	中	橙
1/2	0.1245	中	橙	1/8	0.0548	中	橙
1/3	0.0938	中	橙	1/9	0.0518	中	橙
1/4	0.0789	中	橙	1/10	0.0493	强	红
1/5	0.0698	中	橙	1/11	0.0471	强	红
1/6	0.0634	中	橙	1/12	0.0452	强	红

表 4 沙尘暴的爆发等级及其相应的预警信号预测表

## 结果分析

### 1、对模型一拟合结果的分析

在下图中，将拟合后所得的方程式 (8) 所决定的么 45 年内各月的平均沙尘暴次数与实际的沙尘暴次数进行比较，我们发现：

(1) 除 2、5、6、10 月份外其它月份的平均沙尘暴次数与实际吻合的较好，且整体吻合较好，说明该模型能较好地反映出诸多的气候影响因素对沙尘暴发生次数的内在作用机理，对于沙尘暴爆发可能性的预测有一定的指导意义。

报名号#1077

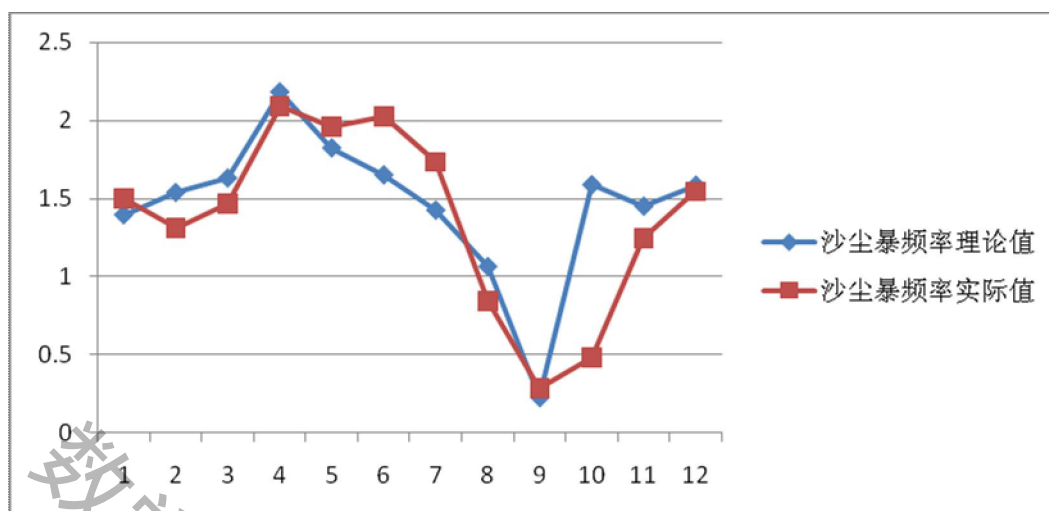


图4 沙尘暴次数理论（拟合后数据）与实际数值对照图

（2）由表-各指标与沙尘暴次数对照一览表可以看出，在所有影响沙尘暴次数的指标中，

大风日数、温度差值及月均风速等三者同沙尘暴次数相关系数相对较大，从一定程度上，反映出这三者对于沙尘暴次数的作用较为明显。各指标都是通过影响气候影响指数进而对沙尘暴次数产生影响的，在我们得到的气候影响指数方程（1）中，也不难看出各个指标对沙尘暴次数影响的影响程度：

由于众气候影响因素中只有土壤湿润度同沙尘暴次数负相关，所以尽管它与沙尘暴次数相关系数相对较小，但它对于沙尘暴次数影响较大，因而是一个平方项。

月份	土壤湿润度	月平均风速	大风日数	温度差值	气候风蚀因子	气候影响指数	沙尘暴次数
1	0.5902	3.370455	2	-1.64091	4.8485	26.0492	1.5
2	0.35	3.288889	1.9	3.93429	4.003	79.5621	1.311111
3	0.5373	3.386667	4	7.475557	4.4573	124.775	1.466667
4	0.194	3.608889	5.5	8.386663	5.1592	413.804	2.088889
5	0.234	3.184444	5.1	6.78889	3.7502	224.95	1.955556
6	0.3821	2.924444	4.9	4.57111	2.9235	134.185	2.022222
7	0.7531	2.691111	4.4	2.20223	2.3393	30.5239	1.733333
8	0.7012	2.513333	2.3	-1.74889	1.8958	7.9406	0.844444
9	0.478	2.3511	1.4	-6.175	1.5162	4.0651	0.288889

	8	11		56			
10	0.189 9	2.52	1.3	-8.051 11	1.9685	105.44 1	0.488889
11	0.253 1	3.24	2.2	-8.615 56	3.8611	40.811 2	1.244444
12	0.308 3	3.4704 55	2	-7.126 71	9.7837	102.99 8	1.545455
各指标与沙尘暴次数的相关系数	-0.11 55	0.6877 75	0.8320 27	0.7012 91	0.4335 78	0.6016 57	1

表 5 各气候指标与沙尘暴次数对照一览表

## 2、异常点剔除后拟合效果差异分析

通过剔除前曲线拟合效果与剔除后曲线拟合效果对比分析剔除的合理性和必要性。

剔除前后其拟合曲线分别为：

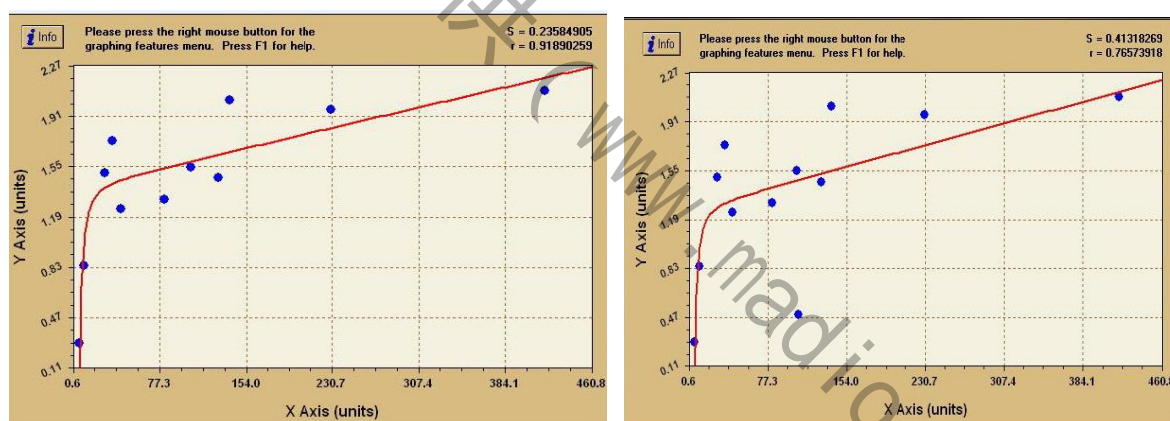


图 5 剔除后回归曲线图

剔除前回归曲线图

相应可以得出两拟合曲线的标准偏差与  $S$  相关系数  $r$  的对比，详见回归效果对比表不难发现剔除异常点后，回归方程的标准偏差  $S$  减小了近一倍，相关系数也有很大幅度的增加，达到了 0.9 以上，说明剔除异常点后拟合曲线的回归效果远较剔除前回归效果优，回归更为显著。从而进一步说明了剔除异常点是非常合理的也是很有必要的。

	$S$	$r$
剔除前	0.413	0.765
剔除后	0.236	0.918

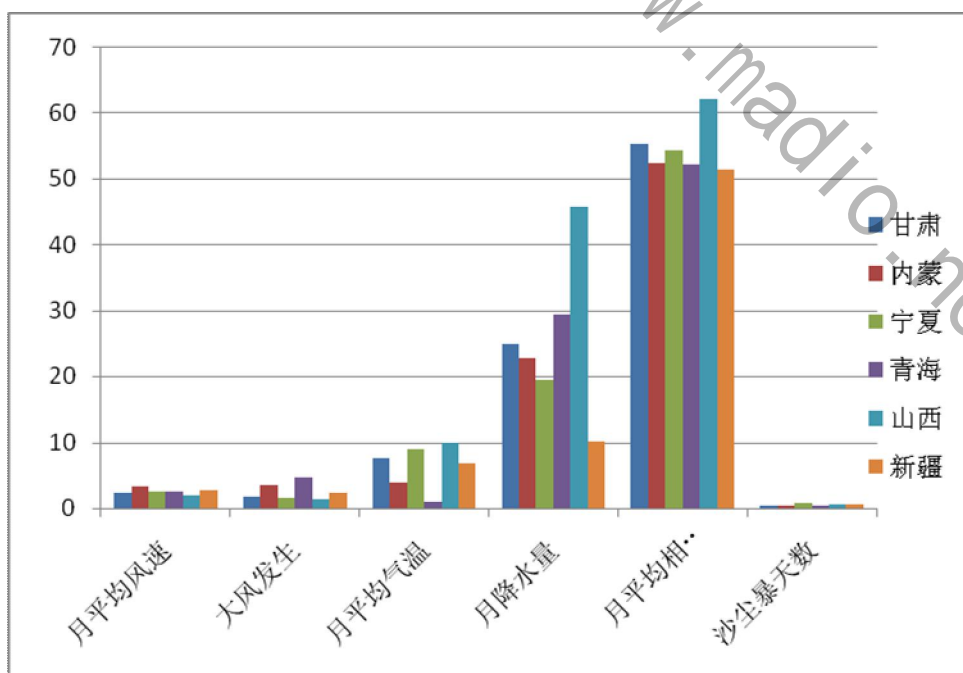
表 6 回归效果对比表

## 模型的进一步讨论

1、在模型二中，在引入式（11） $\lambda_1=\lambda_2/H$  来确定消光系数 $\lambda_1$ 与消光率 $\lambda_2$ 之间的关系时，综合相关的资料并经过理论分析，确定出沙尘天气影响空气层的厚度H为210m，但此值是在210m以上的高空沙尘粒子浓度较小，忽略其对消光率的影响的基础上提出的，但事实上，当沙尘等级较高时，还是或多或少存在着一定影响的，因此H值该应该具体情况具体设定。

2、在模型三中，由于我们采用的方法是目估适线法，即通过观察理论曲线与灾损点子配合得优异度程度来不断地修正 $\bar{S}$ 、 $C_v$ 、 $C_s$ 的取值，以获得配合程度最好的一组值，但这难免会产生人为的误差。为了克服目估适线时的人为误差，可以在计算机上，利用多参数模型搜索的方法和离差绝对值之和最小的适线准则来确定这三个统计参数。

3.结合我们的模型，我们从实际防沙治沙的角度出发，我们作出了各个省份的气候情况，为沙尘暴的实际预测，防治及研究提供一定的指导及借鉴意义。



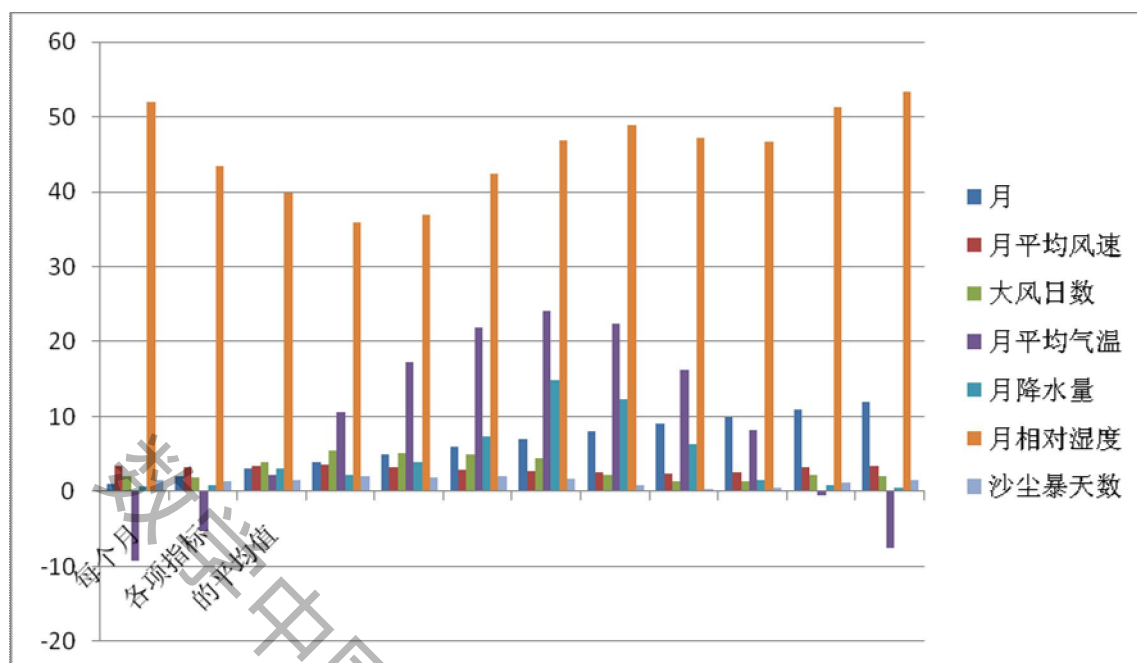


图 6 各省气候因子对比图

## 模型优缺点

### 优点：

- 1、针对某一特定地区建立了适用于该地区的三类模型：气候影响指数模型、沙尘天气等级评价模型及沙尘暴风险预测模型，较好解决了问题要求，且模型的针对性较强。
- 2、模型一中，经过综合权衡，将月平均风速、大风日数、土壤湿润度、地面温差、下垫面风蚀气候因子等气候因素对沙尘暴次数的影响综合为一个量——气候影响指数对沙尘暴次数的影响，通过广泛查阅相关资料，借助机理分析，深入探究了各因素与沙尘暴次数的内在作用机理，模型可操作性很强，可将该研究方法推广到一般领域，模型适用性和实用性较强。
- 3、模型一中，引入大风日数对沙尘暴次数的影响，从而增加了模型的合理性，使模型更贴合实际。
- 3、在模型二中，从理论层面深入讨论了能见度与辐射度之间的内在联系，说服力较强，且针对沙尘天气等级的评价方法简单易行。
- 4、在模型三中，巧妙地引入灾害熵对未来的沙尘暴风险进行预测，从而极大地减少了未来诸多的不确定性因素给风险预测带来的困难。

### 缺点：

- 1、由于各气候影响因素量纲不同，其数值的相对大小差别比较大，因而将他们合并讨论时，有可能会出现相对数值较大的因素弱化其它因素对沙尘暴次数的作用的情况，而使得结果产生变异。

报名号#1077

2、文中仅针对一个地区建立相应的模型，而不同的地区由于地理因素上的差异其具体模型显然不完全一样，但我们并没有考虑地理位置的不同对模型产生的影响，一方面造成了很大的信息浪费，另一方面给模型的适用性带来了一定限制。但是，此种研究方法同样可以应用到其它地区进行分地区分析，即虽然各地区具体情况不同，但模型仍可以推广到一般领域。

## 参考文献

- [1] FAO. A provisional Methodology for Soil Degradation Assessment [M]. Rome : Food and Agriculture Organization of the United Nations ,1979.
- [2] 石慧兰, 王新堂, 邵志勇, 陈成国, 张建海 德州生态质量气象评价中湿润指数定量分析 《山东气象》第 27 卷 , 2007 年 3 月
- [3] 张广兴, 李霞 沙尘暴观测及分级标准研究 《中国沙漠》第 23 卷第 5 期, 2003 年 9 月
- [4] 任鲁川 灾害熵: 概念引入与应用案例 《自然灾害学报》第 9 卷 2 期, 2000 年 5 月
- [5] 冯利华, 李凤全 基于最大熵原理的灾害损失分析 《数学的实践与认识》第 35 卷第 8 期, 2005 年 8 月
- [6] 李艳春 宁夏中北部地区沙尘暴频率气候影响指数模型的建立 《干旱气象》Vol. 22 NO. 1 Mar , 2004
- [7] 王毅, 饶瑞中 空间斜程能见度的影响因素分析 《强激光与离子束》第 15 卷第 10 期 2003 年 10 月
- [8] 辛金元, 张文煜, 袁九毅, 刘立超 沙尘气溶胶对直接太阳辐射的衰减研究 《中国沙漠》 Vol. 23 No. 3 May 2003
- [9] 姚润丰 中央气象台发布沙尘暴橙色警报 <http://www.sina.com.cn> 08. . 4. 12