第6卷 第4期
1985年11月

华侨大学学报
JOURNAL OF HUACHIAO UNIVERSITY

Vol.6 No.4
Nov. 1985

智能化烟草水份测定仪ZHY—I的数学模型及计算机程序

蔡灿辉

〔计算机科学(电脑)〕

摘要

本文借助理论分析及数理统计方法，巧妙地求出了烟草水份测定仪的数学模型。文中的方法虽然是对烟草湿度测量提出来的，但其实验设计、分析、处理及计算机程序等一整套方法，对许多农作物湿度测量也同样适用。

引言

烟草含水率的测量是卷烟行业中至为关键的课题。没有这种测量，就不可能有现代化的卷烟行业。要进行测量，首先要确定其数学模型，对ZHY—I智能化烟草水份测定仪来说，也就是要求烟草电阻R与含水率X%的关系曲线。但是，电阻R不仅与含水率X%有关，而且还同烟草品种、温度、形状、密度及电极的位置有关。在一般测量中，总是采取一定措施来保证或尽可能保证电极的位置及被测烟草的形状不变，同时尽量争取在品种、含水率、温度不变时密度不变。即使如此，由于烟草品种繁多，不同产地，不同生长期，施用肥料及加工过程中所用的配方、料汁不一样，温度对R的影响也可能不太一样。因此，只用一个数学模型来代表所有的烟丝的R—X关系曲线是不可能的。唯一可行的办法就是通过精密的实验，采集大量的数据，并运用数理统计的方法，求出各种典型烟丝的R—X间的相互关系。才能进而考虑各品种间的关系问题，或者说通过修正，从一品种的模型导出其它品种的近似模型问题。

一、实验的设计

要进行实验就必须先进行精密的设计。实验设计直接关系到求出的数学模型的可用性问题。首先必须考虑的问题是样本的来源，要做一条回归曲线，必须要得到曲线各部分的信

本文1984年10月28日收到。

息，或者说，要使样本点的分布能反应出曲线变化的趋势，这就要求：（1）样本点遍布整个曲线上；（2）在曲线变化陡峭的地方，应使样点分布稠密一些。但在烟厂中，实际可采到的样值一般也就只对应几个点，不可能做出一条曲线，因此必须人工制样。我们采用的方法是回潮法制样。刚回潮后的烟丝所含的水份大多是自由水，而且各部分烟丝的含水量很不均匀，必须经过时效，才能测试。时效时间一般需要2～3天。为了使随机误差尽可能小，同时还考虑到实验条件的限制，每次测量取两个样，采用ZHY—I测烟丝探头，每测量一次可得到四个方位上的值，两个样本可有八个值，由于四个方位上的测量值不呈现相关关系，可近似认为八个独立样本。

其次是测量仪器的选择问题。根据误差公理，有测量就有误差。误差有三大类：系统误差，过失误差及随机误差。后两种误差在数理统计学中有许多克服或削弱的办法，而系统误差则不能依靠统计的方法来消除或减弱。系统误差分两大类：技术误差和理论误差。用数理统计的方法求仪器的数学模型，就是在一定置信度上排除理论误差。而技术误差则取决于作数学模型时跟最终测量时技术条件的相异程度。所谓测量，就是拿未知量去与同类已知量作比较。对于测量仪器，这个已知量也就是它的刻度曲线。当测定数学模型的系统与测量的系统完全一致的话，这就相当于一种代替法测量，误差仅取决于理论误差而没有技术误差。考虑到这一点，所以在测试数学模型时就直接利用ZHY—I本身来测量对应不同水份的电阻值，而用烘箱法求出含水率的实际值。在智能化仪器中，只要把内部软件稍微修改一下，是很容易做到这一点的。表1、2是两种典型烟丝的测量数据。

表1　　友谊牌烟丝测量数据

X% \ R(MΩ) \ N	1	2	3	4	5	6	7	8	$\bar{R}$
12.6	170	208	192	132	154	250	181	186	181
13.0	147	102	147	104	143	165	125	113	131
14.5	38	38	35.4	38	43.6	37.1	46.1	46.1	40.3
14.7	31.8	31.8	30.2	27.9	30.2	29.4	32.7	24.9	29.9
15.4	23.4	24.2	20.6	19.5	20.0	24.9	20.6	29.8	22.8
16.5	8.38	9.29	8.16	11.2	14.8	20.6	17.8	19.5	13.7
16.6	7.93	10.5	13.9	10.9	12.8	9.29	8.16	9.29	10.3
17.1	8.07	9.97	8.84	11.5	10.5	9.07	11.2	10.5	10.1
18.2	2.5	3.9	2.7	2.79	3.12	3.12	2.79	2.96	2.99
19.4	1.82	1.87	1.59	1.54	1.73	2.13	1.73	1.78	1.77
19.8	1.28	1.43	1.92	1.63	1.87	1.63	1.28	1.36	1.55
19.9	0.983	1.037	1.21	0.815	0.903	1.24	1.01	0.837	1.00
20.4	1.68	1.73	1.43	1.62	1.43	1.28	1.5	1.5	1.52
22.6	0.448**	0.517	0.522	0.519	0.52	0.517	0.517	0.518	0.524
23.5	0.31	0.31	0.378	0.369	0.293	0.339	0.257	0.339	0.324
25.4	0.19	0.195	0.184	0.211*	0.179	0.184	0.179	0.184	0.187
25.5	0.242**	0.179**	0.223	0.217	0.217	0.217	0.223	0.223	0.220
30.7	0.0817	0.0741	0.0791	0.0648	0.0704	0.0741	0.0704	0.0741	0.0736

表2　　鹭江牌烟丝测量数据

$R(M\Omega)$ \ N X%	1	2	3	4	5	6	7	8	$\bar{R}$
12.1	400	650	715	567	457	861	740	740	641
12.8	400	457	345	310	400	517	414	471	415
13.0	165	197	236	257	271	197	154	286	220
14.4	96.3	104	122	93.6	82.1	57.4	60.7	125	92.6
15.0	51.1	36.2	39.8	30.9	67.3	86.5	38.9	34.5	48.2
15.1	49.9	16.9	27.2	29.4	27.9	30.2	35.4	15.2	29.0
15.8	30.9	65.7	46.1	22.8	38.9	27.2	41.1	35.4	38.5
16.1	14.4	19.5	29.4	27.2	29.4	19.5	20.6	20.6	22.6
18.6	5.0	4.85	5.43	5.72	5.58	7.74	6.01	7.55	5.99
19.2	5.29	3.46	3.29	3.79	3.9	3.57	4.85	3.79	3.99
19.3	2.7	1.92	2.38	1.63	4.01	2.44	2.01	3.39	2.56
20.2	0.956	1.09	1.50	1.47	1.94	1.17	1.24	1.09	1.31
20.3	1.78	2.01	1.92	3.46	1.54	1.36	1.73	2.44	2.03
22.6	0.586	0.554	0.538	0.448	0.497	0.460	0.602	0.635	0.540
23.1	0.837	0.586	0.859	0.554	0.509	0.30	0.586	0.485	0.59
26.4	0.228	0.217	0.19	0.19	0.206	0.20	0.184	0.257	0.209

二、数据处理与数学模型的建立

数据处理的目的在于建立数学模型，其最简单的方法是多项式回归。但是，并不是每个函数在其所关心的范围内都可以用多项式来逼近的。从数学分析中我们知道$(1+X)^{\alpha}$当α不是非负整数时，在$X>1$处其台劳展开是发散的。因此如何选择逼近函数，对数学模型的建立，是至关重要的。而对所研究的对象的物理性质稍加探讨，对逼近函数的选取，是相当有帮助的。

从烟草的化学成份看来，烟草的导电性能主要取决于内部电解质的离子浓度。当含水率X趋于零时，电解质的电离度也趋于零，只有少数几个自由电子能在电场作用下作定向运动，因而电阻很大。随着X增大电解质开始电离，这时只要少数几个离子产生就会对导电率产生很大的影响，因而电阻下降很快。但是，随着X的增大，可电离的离子数目越来越少，而且产生同样数目的离子对离子浓度变化率影响也越来越小，因此电阻R随X增大而下降的趋势也越来越缓慢，最后趋于某R_0亦即X—R的关系近似服从负幂函数式：

$$R=R_o+D/(X+q)^n$$
$$X=C/(R-R_o)^n-a \qquad (1)$$

或把表1、2的数据画成曲线(图1)可以看出这些曲线都基本符合式(1)。因此，问题就在于如何定出常数R_o，C，n，a。如果直接用最小二乘法对(1)式进行回归，即令：

$$\sum_i\left[X_i-C/(R_i-R_o)^n+a\right]^2=min$$

则求解R_0，C，n，a时势必要求解超越方程组。而即使采用数值解法，这还是一个头痛的，有时甚至是无法解决的问题。考虑到X不太大时，总有$R \gg R_0$，这时式(1)可化为：

$$X = C/R^n - a \qquad (2)$$

若a为已知时，(2)式即可化为线性回归的方程。因此式(2)可以采用对a的区间搜索法来求回归，但计算量较大。考虑到在所用曲线段总有$X >> a$，略去a的作用，对(2)两边取对数，有：

$$l_nX = l_nC - nl_nR \qquad (3)$$

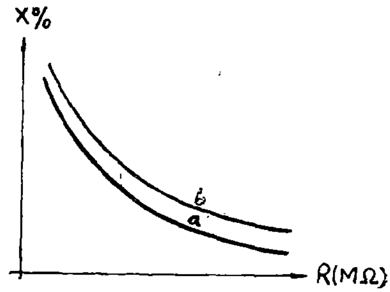

图 1 两种典型烟丝的实验曲线
a—友谊牌 b—鹭江牌

用式(3)所求得的回归方程在R的几何中值附近是基本准确的。但在R偏大或偏小时，就会产生一定偏差，因此需要修正。考虑到a对低端数据影响较大而R_0对高端的回归精度影响较大。故在求回归方程(2)时采用办法如下：

1.用曲线中段的数据求出回归方程(3)；

2.用低端数据求出a′，使标准差最小；

3.对高端数据回归出R_0'。

图2是这种作法的示意。下列为所求的OFRTRAN程序(以表1的数据为基础)。

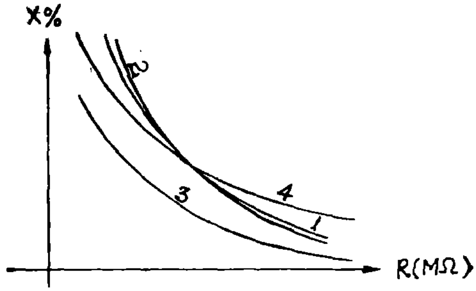

图 2 求曲归曲线示意图
曲线 1—$X = C/(R - R_0)^n - a$
2—$X = C'/(R - R_0')^{n'} - a'$
3—$X = C'/R^{n'} - a'$
4—$X = C'/R^{n'}$

```
C     THIS IS A PROGRAM FOR RELATIONSHIP BETWEEN
C     TOBACCO'S MOISTURE AND ITSRESISTANT DIMENSION X(139),
      R(139), A(139), B(139)
      DO 100 I=1,139
      READ(2,10) X(I), R(I)
10    FORMAT(2F8.4)
      B(I)=ALOG(X(I))
100   A(I)=ALOG(R(I))
      SA=0
      SB=0
      SAA=0
      SBA=0
      DO 200 I=1,118
      SA=SA+A(I)
```

```
      SB = SB + B(I)
      SAA = SAA + A(I)*A(I)
200   SBA = SBA + A(I)*B(I)
      D = -(SBA - SB*SA/118)/(SAA - SA*SA/118)
      C = EXP((SB + D*SA)/118)
      A0 = 0
      DO 300  I = 1,96
300   A0 = A0 + C/R(I)**D - X(I)
      A0 = A0/96
      R0 = 0
      DO 400  I = 132,139
400   R0 = R0 + R(I) - (C/(X(I) + A0))**(1/D)
      R0 = R0/8
      A1 = -A0
      WRITE(3,30)A1, C, R0, D
 30   FORMAT(5X,'X = ',F8.4,'+ ',F8.4,'/R- ',F8.4,'**',F8.5)
      WRITE(3,40)
 40   FORMAT(6X,'R(I)',8X,'X(I)',8X,'XREG(I)',8X,'DX(I)')
      S = 0
      DO 500  I = 1,139
      XREG = C/(R(I) - R0)**D - A0
      DX = X(I) - XREG
      S = S + DX*DX
500   WRITE(3,50) R(I), X(I), XREG, DX
 50   FORMAT(1X,4F12.4)
      S = SQRT(S/137)
      WRITE(3,60)S
 60   FORMAT(/5X,'S = ',F8.4/)
      STOP
      END
```

三、数据处理实例

为了求出表 1 中的实验数据的回归方程，实行如下处理：

1.奇异数据的排除：

在实验中，由于各种原因，可能产生过失误差，使数据偏离真值太远，这种数据必需排除，以保证回归方程的可靠性. 奇异数据判别方法有多种，考虑到数据太多要计算标准差也很麻烦，故采用狄克松准则进行判断，对显著性为 0.01(标有**)及显著性为 0.05(标有*)的数

据均予排除。

2.利用上述FORTRAN程序求回归可得回归方程及标准差如下：

$$X=0.0977+20.7255/(R-0.0577)^{0.09569}$$

$$S=0.5255 \tag{4}$$

如果直接采用多项式回归(多项式回归程序参见[3]，但变量及数组均为双精度)可求得回归方程及标准差如下：

(1) 4阶多项式：

$$X_4=7.196\times10^{-8}R^4-3.878\times10^{-5}R^3+6.94\times10^{-3}R^2-0.471R+22.68$$

$$S_4^*=2.64621$$

(2) 5阶多项式：

$$X_5=-1.528\times10^{-9}R^5+9.618\times10^{-7}R^4-2.14\times10^{-4}R^3+0.02R^2-0.743R+23.18$$

$$S_5^*=2.47783$$

(3) 6阶多项式：

$$X_6=-1.9245\times10^{-11}R^6+1.199\times10^{-8}R^5-2.547\times10^{-6}R^4+1.966\times10^{-4}R^3-7.357\times10^{-4}R^2-0.404R+22.65$$

$$S_6^*=2.88748.$$

由上可见，采用本文所述的方法拟合出来的曲线要比多项式拟合精确得多，也简单明了得多。

3.精度估计：

在我们进行测烟草含水率及电阻关系的实验中，对于同一盘烟丝，由于不均匀性等因素的影响，测出的电阻值通常是不一样的，一般可表为 $R=f(X)+\xi$ (5)

其中X为一盘烟丝的平均含水率

ξ为随机变量，可近似认为服从$N(0,\sigma^2)$分布。但实际上我们是用其反函数

$$X=f^{-1}(R)+\eta \tag{6}$$

来求回归的，η不服从$N(0,\sigma^2)$分布，如果只是一次性测量，精度是难于估计的。但考虑到$ZHY-I$每完成一次动作是16次测量的平均，根据中心极限定理，测量平均值的分布是趋于正态分布的，而在工程上当测量次数不小于8时，即可认为$\bar{X}$是正态分布的了。因此$ZHY-1$的测量精度在置信概率为95%时应为$\bar{X}\pm2S\bar{x}=\bar{X}\pm2S/16=\bar{X}\pm S/2$.对表1的数据有

$$S\bar{x}^*=0.1314$$

四、数学模型的适应性——品种补偿问题

能否通过普通校正方法用一台仪器测量不同品种的烟丝，决定仪器存在价值。然而，由于烟丝数学模型中的几个参数a、C、R_0、n、都随品种因素S的变化而变化，但具体形式又很难找出，因而只能靠实验的方法在一定的精度范围内进行补偿。根据循分中值定理，对函数$X=g(R,S)$，当S从S_0变到$S_0+\triangle S$时，函数X偏差为：

$$\triangle X=g(R,S_0+\triangle S)-g(R,S_0)=g'_s(R,S_0+\theta\triangle S)\triangle S$$

通常 g_s' 不会是一个常数。但是由于烟草间的物化性质相差不大，而且含水率变化范围也比较小，实验表明，在一定精度范围内 g_s' 可以认为是一个常数。以表2的数据为例，按最小均方误差原则把品种 a 的相关模型往上平移来作其数学模型可求得回归方程及标准差如下：

$$X_b = 0.7341 + 20.7255/(R - 0.0571)^{0.00569}$$

$$S_b^* = 0.6037 \qquad S\bar{x}_b^* = 0.1509$$

这个数学模型的精度还是令人满意的。

结　束　语

借助本文所建立的数学模型[式(4)]，我们成功地研制了“智能化烟草水份测定仪 ZHY－I”。在实践中证明本文提出的方法是正确的。对于非电量电测来说，刻度曲线的标定是至为关键的，而本文的一些方法，无疑对一些生资材料的含水率的测量也有适用性。

谨向指导该论文撰写的方志成副教授表示深切的谢意。

参　考　文　献

〔1〕丁士晟，多元分析方法及其应用，吉林人民出版社，(1981)。

〔2〕费业泰等，误差理论与数据处理，机械工业出版社，(1983)。

〔3〕方志成等，裂纹深度刻度曲线的数学模型与计算机程序，华侨大学学报，4.1(1983)。

The Mathmatical Model of Micro—processor BasedToqacco Moisture Meter ZHY—I and Computer program

Cai Canhui

Abstract

By means of theoretical analysis and mathmatical statistics, the paper found the mathmatical model of the tobacco moisture meter in a clever way. Though it is only about tobacco′s moisture testing, the ways of its experimental designing, analysing, processing and computer programming can de extensively used iu many other crops′ moisture testing.