

DOI :10.3880/j. issn. 1006-7647. 2010. 02. 012

冬小麦水肥生产函数最小二乘法回归建模及分析

曹永强^{1 2 3}, 刘琳¹, 姜莉¹, 张伟娜¹

(1. 辽宁师范大学城市与环境学院 辽宁 大连 116029 ; 2. 农业部农业环境与气候变化重点开放实验室 北京 100081 ;
3. 中国科学院水利部水土保持研究所黄土高原土壤侵蚀与旱地农业国家重点实验室 陕西 杨凌 712100)

摘要 : 将最小二乘回归应用于北京市水利科学研究所永乐店试验站冬小麦水肥生产函数建模 , 建立了合理的、具有较强稳定性和预测能力的水肥生产函数模型。由该模型可知 , 水、肥 2 个因素都具有明显的增产效应。水肥耦合效应表现为水肥对产量的协同作用关系。随着作物耗水量的增加 , 化肥的利用效率得到提高 , 而增加化肥的投入量又能提高水分利用效率。

关键词 : 冬小麦种植 ; 最小二乘回归 ; 作物水肥生产函数

中图分类号 : S11+4 文献标识码 : A 文章编号 : 1006-7647(2010)02-0045-04

Modeling and analysis of water-fertilizer production function of winter wheat by use of least-square regression//CAO Yong-qiang^{1,2,3}, LIU Lin¹, JIANG Li¹, ZHANG Wei-na¹(1. School of Urban Planning and Environmental Science, Liaoning Normal University, Dalian 116029, China ; 2. Key Laboratory of Agro-Environment and Climate Change, Ministry of Agriculture, Beijing 100081, China ; 3. State Key Laboratory of Soil Erosion and Dryland Farming on the Loess Plateau, Institute of Water and Soil Conservation, Chinese Academy of Sciences and Ministry of Water Resources, Yangling 712100, China)

Abstract : The least-square regression was applied in the modeling of water-fertilizer production function of winter wheat at Yongledian Irrigation Station of Beijing Hydraulic Research Institute. A rational water-fertilizer production function model with strong stability and prediction performance was established. The modeling results show that the water and the fertilizer have remarkable production increase effect. Their coupling effect is represented by their co-action on the production of winter wheat. With the increase of water consumption of crops, the utilization efficiency of the fertilizer increases, and the increase of the fertilizer input raises the utilization efficiency of the water.

Key words : planting of winter wheat ; least-square regression ; water-fertilizer production function of crops

关于作物的水肥与产量关系研究由来已久 , 但真正的量化关系研究特别是水肥协调关系方面的量化研究则是进入 19 世纪以后^[1]。化肥的大量施用 , 不仅带来农业生产成本的增加 , 还对土壤及地下水造成污染。因此 , 如何提高作物对水肥的利用效率、减少化肥污染、提高农业生产效益 , 是目前及今后长时间内有效地进行农业水肥管理的一个重要方面。解决这一问题的基础是对水-肥-作物产量关系(作物水肥生产函数) 进行研究^[2]。人们从不同的角度入手探讨水分与作物产量的关系 , 建立了多种作物水分生产函数模型。归纳起来 , 这些模型可分为 2 大类 : 一类是以作物耗水(作物蒸腾量或腾发量) 为变量 , 寻找不同生育阶段不同程度的水分亏缺与产量的关系 ; 一类是建立作物生产过程不同农田水分状况与干物质累计量之间的关系^[3]。

Ozsabuncuoglu^[4]通过试验建立了冬小麦的全生育期产量模型 , 该模型为多元线性模型 , 它考虑了生长期和成熟期的降雨、温度和施肥。徐学选等^[5]在黄土丘陵区(固原) 研究了春麦产量与降水和土壤供水的关系 , 定量描述了底墒和生育期降水与产量的关系。笔者在分析北京市水利科学研究所永乐店试验数据的过程中 , 选择最小二乘回归建立了具有较高精度和预测能力的冬小麦水肥生产函数模型 , 并对水肥耦合效应进行了分析。

1 试验和数据

1999 ~ 2000 年在北京市水利科学研究所永乐店试验站的 24 个田间小区进行了冬小麦水肥二因子非充分灌溉试验。试验站位于北京市通县永乐店镇 (地理位置 116.8°E , 39.7°N) , 试验区土质为砂壤土 ,

基金项目 : 国家自然科学基金(50979035) ; 农业部农业环境与气候变化重点开放实验室 2009 年对外开放基金(200901) ; 黄土高原土壤侵蚀与旱地农业国家重点实验室 2008 年对外开放基金主任基金(10501-213)

土壤密度为 1.4 t/m^3 ,田间持水量为 36.5% 。化肥的施用主要是氮肥 ,为了进行对照 ,试验分为不施肥、少施肥、中施肥和高施肥 4 种情况。各小区的水肥投入量和产量见表 1。用 E_T 表示由田间水量平衡分析求得的作物在整个生育期间的腾发量(腾发量 = 降雨量 + 灌溉量 - 土壤根系层含水的增加量) ,有了腾发量(或灌溉量)和施肥量 ,就可以建立作物全生育期的腾发量、施肥量与最终产量的函数模型。这种模型适用于缺少大量试验数据区域的作物产量与水肥关系的模拟 ,尤其是在干旱地区应用起来很方便。

表 1 1999 ~ 2000 年水肥试验数据

试验样本	E_T/mm	化肥施量/ ($\text{kg}\cdot\text{hm}^{-2}$)	作物产量/ ($\text{kg}\cdot\text{hm}^{-2}$)
A1	211.6	600.0	4011.0
A2	210.1	400.5	4246.5
A3	306.9	600.0	6351.0
A4	190.1	0	3183.0
B1	254.5	600.0	4831.5
B2	250.3	400.5	4737.0
B4	226.9	0	4087.5
C1	244.3	600.0	4257.0
C2	267.5	400.5	5709.0
C4	252.3	0	4410.0
D1	265.8	600.0	5218.5
D2	265.9	400.5	6367.5
D3	266.6	199.5	5443.5
E2	290.0	400.5	5467.5
E3	293.0	199.5	5542.5
E4	273.8	0	5341.5
F1	307.2	600.0	6334.5
F2	307.1	400.5	6088.5
F3	305.3	199.5	6334.5
F4	301.0	0	6444.0
B3	243.0	199.5	4447.5
C3	304.9	400.5	6040.5
D4	255.7	0	5140.5
E1	294.2	600.0	5887.5

注 :B3 ,C3 ,D4 ,E1 为模型检验样本 ,其余为建模样本。 E_T 、化肥施量及作物产量平均值分别为 266.167 mm , 21.671 kg/hm^2 和 349.783 kg/hm^2 ;标准差分别为 38.683 mm , 16.900 kg/hm^2 和 61.631 kg/hm^2 。

2 最小二乘回归的基本原理

普通最小二乘法(ordinary least square ,简称 OLS)是应用最多的参数估计方法。最小二乘估计是一种方差最小的估计 ,有很强的优良性。它通过最小化误差的平方和找到一组数据的最佳函数匹配 ,求得一些未知的真值。

在实际问题中 ,与某随机变量 y 有关系的变量往往有多个 ,而且它们之间的关系是非线性的。研究随机变量 y 与一般变量 $x_1 ,x_2 ,\dots ,x_t$ 之间的相关关系问题便是多元非线性回归问题。为了寻找 y

与一般变量 $x_1 ,x_2 ,\dots ,x_t$ 之间的相关关系 ,首先必须收集 n 组独立观测值^[6] $x_{i1} ,x_{i2} ,\dots ,x_{it} ,y_i$ ($i = 1 ,2 ,\dots ,n$)。

对 n 个有限样本的多元回归模型 :

$$\begin{cases} Y = X\beta + \varepsilon \\ \varepsilon \sim N_n(0, \sigma^2 I_n) \end{cases} \quad (1)$$

其中

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \quad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{(t+3)/2} \end{pmatrix} \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

$$X = \begin{pmatrix} 1 & x_{11} & \dots & x_{1t} & x_{11}^2 & \dots & x_{11} & x_{1t} \\ 1 & x_{21} & \dots & x_{2t} & x_{21}^2 & \dots & x_{21} & x_{2t} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & x_{n1} & \dots & x_{nt} & x_{n1}^2 & \dots & x_{n1} & x_{nt} \end{pmatrix}$$

参数的最小二乘估计 :令

$$Q(\beta_0 ,\beta_1 ,\dots ,\beta_{(t+3)/2}) = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_{i1} - \dots - \beta_{(t+3)/2} x_{it}^2)^2 \quad (2)$$

则各 β_i 的最小二乘估计 $\hat{\beta}_i$ 应该满足

$$Q(\hat{\beta}_0 ,\hat{\beta}_1 ,\dots ,\hat{\beta}_{(t+3)/2}) = \min Q(\beta_0 ,\beta_1 ,\dots ,\beta_{(t+3)/2}) \quad (3)$$

由于 $Q(\beta_0 ,\beta_1 ,\dots ,\beta_{(t+3)/2})$ 是 $\beta_0 ,\beta_1 ,\dots ,\beta_{(t+3)/2}$ 的非负二次式 ,且关于 $\beta_0 ,\beta_1 ,\dots ,\beta_{(t+3)/2}$ 均可微 ,因而根据微积分原理 , $\hat{\beta}_0 ,\hat{\beta}_1 ,\dots ,\hat{\beta}_{(t+3)/2}$ 是下列方程组的解 :

$$\begin{cases} \frac{\partial Q}{\partial \beta_0} = 0 \\ \frac{\partial Q}{\partial \beta_j} = 0 \quad j = 1, 2, \dots, (t+3)/2 \end{cases} \quad (4)$$

通过整理可将其用矩阵表示为

$$(X'X)\hat{\beta} = X'Y \quad (5)$$

当 $(X'X)^{-1}$ 存在时 ,其最小估计为

$$\hat{\beta} = (X'X)^{-1}X'Y \quad (6)$$

3 最小二乘回归建模

关于冬小麦产量与农田耗水量的关系众说不一 ,有直线、抛物线及对数型等关系 ,多数研究者采用抛物线(二次多项式)来描述两者的关系 :当耗水量超过需水量时 ,产量不但不增加 ,而且有可能降低。然而由于影响产量的因子不止是水分 ,肥料也是影响作物产量的重要因子 ,不妨将冬小麦产量与水肥 2 个因子的关系用二元二次多项式进行拟合。

设 E_T 为 x_1 ,化肥投入量为 x_2 ,产量为 y ,通过多元回归 ,可以建立包含交叉项的水肥生产函数

模型：

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1^2 + \beta_4x_2^2 + \beta_5x_1x_2 \quad (7)$$

利用最小二乘回归原理 ,借助 SPSS 软件可以求得 1999 ~ 2000 年的水肥生产函数模型各项系数如下 : $\beta_0 = - 78.159$, $\beta_1 = 1.388$, $\beta_2 = 2.831$, $\beta_3 = 0.000\,737$, $\beta_4 = - 0.023\,5$, $\beta_5 = 0.000\,28$ 。复相关检验 : $R^2 = 0.866\,3$, $R = 0.931$;显著性检验 : $F = 23.65 > F_{0.95}(5,14) = 3.66$,可认为上述回归方程 (式 (7))是显著的。

产量模型计算值与实测值的拟合结果如图 1 所示。另假设化肥量为 150 kg/hm^2 , $200\text{ mm} < E_T < 320\text{ mm}$ 的水肥组合系列 ,采用 1999 ~ 2000 年的水肥生产函数模型系数计算相应的产量。从表 2 可知计算结果较接近实际值 ,说明最小二乘回归可求得正确的水肥生产函数模型。

表 2 最小二乘回归模型模拟得到的冬小麦产量

E_T/mm	产量/ ($\text{kg}\cdot\text{hm}^{-2}$)	E_T/mm	产量/ ($\text{kg}\cdot\text{hm}^{-2}$)
200	3 750	266	5 325
220	4 215	280	5 670
240	4 695	300	6 165
260	5 157	320	6 675

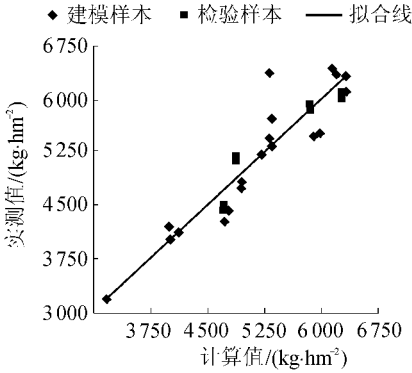


图 1 最小二乘回归产量拟合

多元回归模型的使用有一个前提 ,就是自变量之间不能存在严重的多重相关性。当 X 中的变量高度相关时 ,行列式 $|X'X|$ 几乎接近于零 (条件数趋于无穷大) ,求 $X'X$ 的逆阵会含有严重的舍入误差 ,因此回归系数容易受较大舍入误差的影响 ,从而增加估计值的抽样变异性。可以用简单相关系数粗略地分析自变量之间的多重相关性。表 3 列出了 1999 ~ 2000

表 3 回归变量之间的简单相关系数

变量	回归系数				
	x_1	x_2	x_1^2	x_2^2	x_1x_2
x_1	1.000	- 0.294	- 0.993	0.026	0.289
x_2		1.000	0.364	- 0.254	- 0.900
x_1^2			1.000	0.004	- 0.375
x_2^2				1.000	- 0.176
x_1x_2					1.000

年水肥生产函数回归模型自变量之间的简单相关系数。从表 3 中可知 ,回归变量之间的相关系数比较小 ,所以认为回归变量之间不存在多重相关性。

最小二乘回归的具体计算通过 SPSS 软件实现 ,具体参见文献 [7]。

4 对最小二乘回归模型的分析 and 讨论

由最小二乘法求得的水肥生产函数可以拟合 1999 ~ 2000 年冬小麦实测产量与计算产量。从图 1 来看 ,拟合结果具有较高的精度。但由于试验数据是小样本容量 ,还不足以据此判断所建立回归模型的正确性、稳健性及预测能力。为了进行判断 ,可以假设代表各个试验处理水平的数据 ,利用模型模拟 ,根据产量的模拟值来验证模型是否合理 [8]。图 2 反映了化肥投入量不同时的 E_T 和产量关系。从图 2 中可以断定最小二乘回归得到的水肥生产函数模型是稳定可靠的 ,可以据此分析水肥耦合效应 [9]。

- 化肥投入量为 0 时
- 化肥投入量为 0 时拟合线
- 化肥投入量为 5 kg 时
- 化肥投入量为 5 kg 时拟合线
- 化肥投入量为 15 kg 时
- 化肥投入量为 15 kg 时拟合线
- 化肥投入量为 25 kg 时
- 化肥投入量为 25 kg 时拟合线

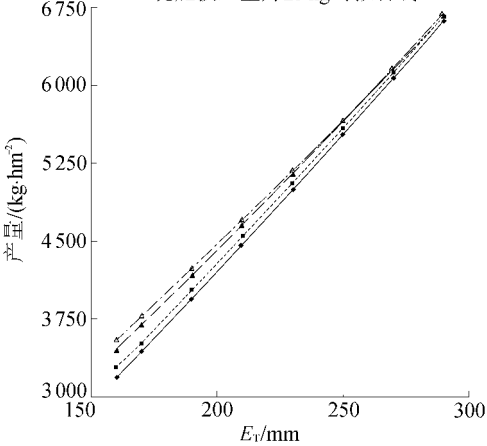


图 2 最小二乘回归模型模拟的 1999 ~ 2000 年冬小麦产量
由式 (7) 和表 2 对 1999 ~ 2000 年的水肥生产函数求一阶偏导数：

$$\frac{\partial y}{\partial x_1} = 1.388 + 0.001\,47x_1 + 0.000\,28x_2 \quad (8)$$

$$\frac{\partial y}{\partial x_2} = 2.831 - 0.047x_2 + 0.000\,28x_1 \quad (9)$$

根据式 (8) 式 (9) 可以进行如下分析。

a. 水肥的单因子效应。水肥生产函数一阶偏导数的常数项都大于零 ,且一次项系数远小于常数项 ,所以水肥投入量处于一定区域时两者都有明显的增产效应 [10]。由式 (8) 和式 (9) 可知 ,当 E_T 不变时 ,随着化肥投入量的增加 ,边际效应递减 ,甚至成

为负效应。例如,若 E_T 固定为 250 mm,当化肥投入量由 45 kg/hm² 增加到 375 kg/hm² 时,化肥投入量的边际效应由 2.66 下降到 1.73。 $\frac{\partial y}{\partial x_1}$ 值的大小表示了 E_T 的边际效应。在化肥投入量一定的情况下 $\frac{\partial y}{\partial x_1}$ 随着 E_T 的增加而增加。但由于一次项系数很小,所以可以认为 $\frac{\partial y}{\partial x_1}$ 基本保持不变。也就是说,在化肥投入量一定的情况下,产量与 E_T 的关系可以近似为直线关系。

b. 水肥耦合效应。水肥生产函数的交叉项化肥的系数为正表明,水肥耦合效应表现为对产量的协同作用,该作用能促进产量的提高。增加化肥投入量有利于提高作物水分利用效率,反之亦然。

5 结 语

用多元非线性回归方法分析水分、肥料对冬小麦产量的耦合影响,结果表明,只有当水分和肥料合理搭配时才能充分发挥它们的作用。多元回归分析时,最小二乘回归可正确地估计模型参数。利用该方法建立了北京市水利科学研究所永乐店灌溉试验站 1999~2000 年的水肥生产函数回归模型。

模型的模拟结果表明该模型参数估计合理。对模型的分析表明, E_T 和化肥 2 个因素都具有明显的增产效应。在化肥投入量不变的情况下,作物的水分利用效率随着耗水量的增加而有所增加;在作物

耗水量保持不变的情况下,作物肥料利用效率随着肥料投入量的增加而下降。在一定范围内,增加化肥投入量能够提高作物产量,但在一定范围之外增加化肥投入会引起作物减产。

参考文献:

[1] 翟丙年,李生秀.冬小麦产量的水肥耦合模型[J].中国工程科学,2002,4(9):69-74.

[2] 张凤翔,周明耀等.郭凯泉.水肥耦合对冬小麦生长和产量的影响[J].水利与建筑工程学报,2005(2):24-26.

[3] 沈荣开,张瑜芳,黄冠华.作物水分生产函数与农田非充分灌溉研究述评[J].水科学进展,1995,4(3):249.

[4] OZSABUNCUOGLU I H. Production functions for wheat : a case study of Southwestern Anatolian Project (SAP) Region [J]. Agricultural Economics ,1998(18):75-87.

[5] 徐学选,穆兴民.小麦水肥产量效应研究进展[J].干旱地区农业研究,1999,17(3):6-12.

[6] 周纪芾.回归分析[M].上海:华东师范大学出版社,1993:33-50.

[7] 薛薇.SPSS 统计分析方法及应用[M].北京:电子工业出版社,2004.

[8] 康绍忠,胡笑涛,蔡焕杰,等.现代农业与生态节水的理论创新及研究重点[J].水利学报,2004(12):1-7.

[9] 胡庆芳,尚松浩,温守光.萧河冬小麦水肥生产函数偏最小二乘回归建模及分析[J].节水灌溉,2006(1):1-4.

[10] 薛亮.夏玉米分根交替灌溉施肥的水肥耦合效应研究[D].杨凌:西北农林科技大学,2008.

(收稿日期 2009-04-08 编辑:高建群)

(上接第 31 页)

河道地形等条件的影响。为了进一步完善纵向一维过饱和 TDG 释放预测模型,需积累更多的过饱和 TDG 原型观测资料,并对 k_T 及 φ_T 的取值范围作进一步率定。同时,可以对泄水的非恒定流特征进行研究,并开展二维或三维天然河道释放过程预测,进一步完善过饱和 TDG 沿程释放模型。

参考文献:

[1] US Environment Protection Agency. Quality criteria for water [R]. Washington , D. C. :Environment Protection Agency , 1986.

[2] 谭德彩.三峡工程致气体过饱和和对鱼类致死效应的研究[D].重庆:西南大学,2006.

[3] US Army Corps of Engineers. Technical analysis of TDG processes[R]. Jacksonville District :US Army Corps of Engineers-Northwest Division , Environmental Resources and Fish Planning Offices 2005.

[4] POLITANO M ,CARRICA P ,TURAN C ,et al. Prediction of the total dissolved gas downstream of spillways using a two-

phase flow mode[C]. World Water Congress.[s.l.].2004.

[5] WEBER L ,HUANG H ,LAI Y ,et al. Modeling total dissolved gas production and transport downstream of spillway : three-dimensional development and application[C]. IAHR&INBO [s.l.] 2004.

[6] LI Ran ,LI Jia ,LI Ke-feng ,et al. Field observation of total dissolved gas supersaturation of high dams[C]. 33IAHR Congress.[s.l.] 2008.

[7] 蒋亮,李嘉,李然,等.紫坪铺下游过饱和和溶解气体原型观测研究[J].水科学进展,2008,19(3):367-371.

[8] 蒋亮,李然,李嘉,等.高坝下游水体中溶解气体过饱和和问题研究[J].四川大学学报:工程科学版,2008,40(5):69-73.

[9] 程香菊,陈永灿.大坝泄洪下游水体溶解气体超饱和和理论分析及应用[J].水科学进展,2007,18(3):346-350.

[10] 冯镜洁,李然,李克锋,等.高坝下游过饱和 TDG 释放过程研究[J].水力发电学报,2010,29(1):7-12.

[11] 李然,李嘉,李克锋,等.高坝工程总溶解气体过饱和和预测研究[J].中国科学,2009,39(12):1887-2006.

(收稿日期 2009-10-28 编辑:高建群)