

doi: 10.7621/cjarrp.1005-9121.20171029

· 区域农业 ·

基于 Logistic 及文本分析的华北地下水漏斗区 农户种植马铃薯意愿研究*

马云倩, 郭燕枝, 徐海泉, 王秀丽*

(农业部食物与营养发展研究所, 北京 100081)

摘要 [目的] 对华北地下水漏斗区—河北地区的农户种植马铃薯的意愿及其影响因素, 以及农户生产种植中遇到的问题进行定量与定性分析。[方法] 基于 149 家马铃薯种植户的问卷调查, 运用 Logistic 与文本分析方法进行研究。[结果] 经定量分析表明, 农业人口占家庭人口的比例、单位面积投入、土地使用情况、资金使用情况、马铃薯种植年限、单位面积产量以及销售渠道等这 7 个因素对农户种植马铃薯的意愿产生显著性影响; 经文本分析发现, 农户急切需要得到政府的帮助, 特别是生产种植技术、补贴政策、机械化程度以及销售渠道等方面的援助。[结论] 农业人口占家庭人口的比例越高, 单位面积投入越高, 土地使用以及资金使用越高, 则农户种植马铃薯的意愿就越低; 马铃薯种植年限越高, 单位面积产量越高以及销售畅通, 农户种植马铃薯的意愿就越高; 农户的基本情况如农民的年龄与文化程度、家庭收入特征、用水情况、培训以及补贴状况等因素对马铃薯种植意愿的影响并不明显。

关键词 马铃薯 种植意愿 影响因素 Logistic 模型 文本分析

中图分类号: F327; F323.1 **文献标识码**: A **文章编号**: 1005-9121[2017]10210-06

0 引言

我国水资源匮乏, 特别是近年来随着经济的不断发展, 水资源缺乏形势日益严峻。水资源紧缺已经成为制约我国粮食生产、产业区域布局乃至整个社会经济发展的重要问题, 在这样的形势下, 水资源节约型作物必然成为未来农业的发展趋势^[1-2]。马铃薯俗称土豆, 耐旱、耐寒、耐贫瘠, 正是一种典型的节约水资源环境友好型作物。不仅如此, 马铃薯脂肪含量低, 营养价值丰富, 能够满足人民的营养需求, 对于改善居民的膳食营养结构具有重要意义。我国政府站在保护环境、节约资源、确保粮食安全以及促进居民营养健康的角度, 提出了马铃薯主粮化战略, 将马铃薯作为即水稻、小麦、玉米之后的第四大主粮作物进行推广。而河北地处华北地下水漏斗区, 属于严重的水资源缺乏省份, 正需要这样一种水资源节约型作物。在水资源缺乏与国家马铃薯主粮化战略背景下, 马铃薯具有很大的发展空间, 极有可能成为河北地区的一种主栽作物与主消粮食。而农户作为马铃薯的生产主体, 他们的种植意愿将直接影响马铃薯的推广与生产, 为此, 研究河北地区农户种植马铃薯的意愿及其影响因素也尤为重要。

在目前的研究中, 对某种农作物农户种植意愿的研究有很多, 例如, 周清明^[3]采用 Logit 模型对农户种粮意愿的影响因素开展了定量研究, 研究发现, 粮食价格、种植规模以及农户的个体特征变量对农户种粮的意愿产生正向影响, 而农户的受教育程度及农资价格则对种植意愿产生负向影响; 宋金田和祁春节^[4]则采用 Logit 模型对我国柑橘主产区的农户种植柑橘的意愿及影响因素开展了定量研究, 研究表明, 种植户的年龄、是否参加过培训、种植年限以及是否了解市场信息等对农户种植柑橘的意愿产生显著性影

收稿日期: 2017-01-09

作者简介: 马云倩 (1987—), 女, 山东潍坊人, 助理研究员。研究方向: 食物营养决策支持

*通讯作者: 王秀丽 (1979—), 女, 山西临汾人, 副研究员。研究方向: 食物与营养发展战略。Email: wangxiuli01@caas.cn

*基金项目: 公益性行业 (农业) 科研专项“马铃薯主粮化关键技术体系研究与示范” (201503001)

响；刘月华等^[5]利用 Logit 模型在恩施山区开展调研，对该地区农户种烟的意愿进行了实证分析，经研究表明，烟叶种植年限和对收购政策的满意度对烟农种植意愿产生正向显著影响；闫志明等^[6]运用 Logistic 模型对新疆南部农户种植棉花的影响因素开展了定量研究，研究表明，政府对棉花补贴、土地肥力、保险以及专家的技术指导对棉花生产产生显著性影响；宋长鸣^[7]利用 Logistic 模型对蔬菜价格波动背景下生产者种植意愿的变化情况开展了研究，研究表明，以蔬菜种植收入为主的农户，由于收入来源单一，在价格蔬菜价格波动时并不会放弃对蔬菜的种植；王亚坤等^[1]利用 Logit 模型对农户种植谷子意愿的影响因素开展了研究，研究表明，农户种植业收入占家庭总收入的比重、家庭耕地面积以及谷子种植补贴是影响农户种植意愿的显著性因素，且都是正向影响。综上所述，目前学者对某种作物的种植意愿研究已有很多，但是对农户种植马铃薯意愿及其影响因素的研究极少或者说尚且没有，在国家马铃薯主粮化背景下，亟需对农户种植马铃薯意愿开展调查研究。文章调查了河北地区 149 家马铃薯种植户，并对他们种植马铃薯的意愿、影响因素，及其反映的问题开展定量与定性研究，以期为当地马铃薯的推广种植及产业发展，以及国家马铃薯主粮化战略，提供合理的科学依据。

1 研究假设及种植影响因素分析

1.1 农户行为的理论假设

在经济学中存在经济人假说的理论，农户被假设为经济人，那么他们的种植行为也就是为了追求自身利益的最大化。农户种植行为是指农户为了追求个人利益最大化，在自身条件及现有自然、社会环境下，进行的生产选择活动^[3-5]。但是在经济学中，农户的种植行为是否是理性的，也存在诸多争议^[4-5]。以俄国经济学家恰亚诺夫^[8]为代表的学派认为农户的生产行为是非理性的，他们的生产行为并不是追求利益最大化，而是为了满足消费需要和劳动辛苦程度上的平衡；以美国经济学家舒尔茨^[9]为代表的学派则认为农户的生产行为是理性的，农户会主动了解市场信息，并且根据信息会对自己拥有的资源进行生产选择以获取最大利益。在文中采用舒尔茨的观点，认为农户为理性的经济人。

1.2 农户种植马铃薯意愿的影响因素

农户作为理性的经济人，他们种植马铃薯的意愿受多种因素影响。目前关于农户种植马铃薯意愿的研究极少或者没有，但是关于农户种植某种农作物意愿的研究却有很多，在前人研究的基础上^[3-7]，本文将从 5 个方面对农户种植马铃薯的意愿及其影响因素进行探讨：农户基本特征、马铃薯生产特征、农户家庭收入特征、资源禀赋约束、外部环境等。

(1) 农户基本特征。主要包括农户的年龄、文化程度、农业人口占家庭人口的比例等 3 个变量。种植户的年龄对马铃薯种植意愿能够产生影响，具体而言，种植马铃薯是体力活，需要年轻力壮的人从事，但是目前农村的现状是很多年轻人外出打工，从事农业种植的人年龄趋于老龄化。文化程度对马铃薯种植意愿也会产生影响，受教育程度与农户了解的信息和掌握资源的能力、对生产决策的判断有关，因此对农户种植意愿会产生一定的影响。农业人口占家庭人口的比例即一个家庭中从事农业人口的人数多少，会对种植某种作物的意愿产生影响。

(2) 马铃薯生产特征。包括种植马铃薯的年限、单位面积投入、单位面积收入以及单位面积产量这 4 个变量。马铃薯的种植年限会对农户的种植意愿产生影响，原因在与种植年限的长短会对农户的种植经验产生直接影响。单位面积投入代表着生产马铃薯的成本，成本的高低直接影响这农户种植马铃薯的意愿。单位面积收入与单位面积产量是农户种植马铃薯的动力，必定会对农户的种植意愿产生直接的影响。

(3) 农户家庭收入特征。采用马铃薯收入占家庭收入的比重这个变量来衡量，该变量能够说明农户家庭的收入来源问题以及该家庭对马铃薯种植的依赖程度。

(4) 资源禀赋约束。从事某种生产活动需要有劳动要素的投入。在生产马铃薯的过程中，使用水、土地是否需要付费，是否受到限制，以及使用的资金是够能够自理，有无相关的种植培训等，这 4 个变量会直接影响农户种植马铃薯的积极性和意愿。

(5) 外部环境。外部环境对农户种植马铃薯的行为影响也是不容忽视的。比如补贴政策, 政府是否对马铃薯的种植有扶持, 有无种植相关的补贴, 直接影响农户的种植积极性。再者, 销售渠道的影响也是至关重要的。农户生产的马铃薯, 能够顺利的销售出去, 有一个畅通的销售渠道, 将极大地增强农户种植马铃薯的信心, 提高种植意愿。

2 马铃薯种植意愿的实证研究

2.1 数据来源

该文数据来源于 2016 年对河北地区农户种植马铃薯意愿的调查, 样本区域涉及河北的 11 个地级市, 调查采用随机入户调查的方式, 收集到了 149 家马铃薯种植户的数据。主要调查内容涉及农户的基本特征、马铃薯生产特征、家庭收入特征、资源禀赋约束、外部环境等 5 方面的问题。

2.2 模型选择和变量说明

该文研究的是农户种植马铃薯的意愿问题, 调查中农户就“是否愿意继续种植马铃薯”一问做出回答, 回答结果有两种: 愿意和不愿意。采用研究农户种植某种农作物意愿的常用模型 Logistic 模型^[4], 以农户是否愿意继续种植马铃薯为因变量(愿意: 1, 不愿意: 0)以农户的基本特征、马铃薯生产特征、家庭收入特征、资源禀赋约束、外部环境等 5 个方面的 14 个影响因素为自变量来进行建模分析。各变量的赋值情况如表 1 所示。

表 1 变量的选择与赋值

变量类别	变量名称	代码	赋值
农户种植意愿	因变量	Y	愿意 = 1, 不愿意 = 0
农户基本情况	年龄	X1	60 岁以上 = 1, 50 ~ 59 岁 = 2, 40 ~ 49 岁 = 3, 30 ~ 39 岁 = 4, 18 ~ 29 岁 = 5
	文化水平	X2	小学 = 1, 初中 = 2, 高中 = 3, 大学 = 4, 研究生 = 5
	农业人口占家庭人口比例	X3	按实际比例
马铃薯生产特征	种植年限	X4	1 年及以下 = 1, 2 ~ 5 年 = 2, 5 年以上 = 3
	单位面积投入	X5	按实际投入
	单位面积收入	X6	按实际收入
	单位面积产量	X7	按实际产量
家庭收入特征	马铃薯收入占家庭收入的比重	X8	0% = 1, 1% ~ 30% = 2, 36% ~ 65% = 3, 66% ~ 90% = 4, 91% 以上 = 5
资源禀赋约束	用水	X9	灌溉用水需要付费 = 1, 不需要 = 0
	用地	X10	使用土地需要付费 = 1, 不需要 = 0
	资金使用	X11	种植马铃薯的资金能够自理 = 1, 不能 = 0
	培训	X12	接受过马铃薯生产知识相关培训 = 1, 没有 = 0
外部环境	种植相关补贴	X13	有补贴 = 1, 没有 = 0
	销售渠道	X14	自己食用不出售 = 1, 自己销售 = 2, 中间商收购 = 3, 与企业或组织签订合同 = 4

2.3 运行结果与影响因素分析

运用 R 软件编写代码进行回归分析。由于影响因素较多, 无法认为确定哪些是真正的影响因素, 因此, 文中采用逐步回归方法进行变量筛选^[10], 运行结果分类结果如表 2 和表 3 所示。

逐步回归筛选变量之后的 Logistic 回归输出中 AIC = 68.212。表 3 中, 行代表正确的类, 列代表模型判断的类, 对角线外为错判个数, 该模型的误判率为 0.1544, 正确的判断率在 80% 以上, 说明模型拟合度较好。

在回归过程中, 变量 X1 (农户的年龄)、X2 (农户的文化水平)、X6 (单位面积收入)、X9 (种植马铃薯用水情况) 以及 X12 (是否接受过相关培训) 等 5 个变量被淘汰掉, 这说明这个 5 个影响因素对农户种植马铃薯的意愿影响很小或者没有。从表 2 中可以看出, 当显著性水平 α 为 0.05 时, X4 (种植年限)、X5 (单位面积投入)、X7 (单位面积产量)、X10 (用地) 以及 X11 (资金使用) 这 5 个变量能够通过显著性检验。但在该文中设定的显著性水平 α 为 0.1, 因此通过显著性检验的变量则为 X3 (农业人口占家庭人口比例), X4 (种植年限)、X5 (单位面积投入)、X7 (单位面积产量)、X10 (用地)、X11 (资金

使用)以及 X14 (销售渠道)这 7 个影响因素。

(1) 农户基本情况。农户基本情况的 3 个影响因素,仅农业人口占家庭人口的比例 (X3) 这一因素通过了显著性检验。变量 X3 的回归系数为负值,这表明农业人口占家庭人口比例越高,农户种植马铃薯的意愿就越低。这是因为,一个家庭中从事农业人口的人数越多,对种植某种作物的选择也就越多,相应的种植马铃薯的意愿也就没那么强烈。X3 这一变量的取值主要集中在 0.3~0.4 之间的 41 户与 0.5 到 0.7 之间的 85 户,2 个区间共有 126 户,占总样本的 85%,该变量的均值为 0.51。这说明该地区农业人口占家庭人口的比例这一因素在家庭间存在一定的差异,一半以上的家庭中从事农业的人口超过 50%,而从事农业的人口越多,他们选择种植其他农作物的可能性就越大,种植马铃薯的意愿也就越小。

(2) 马铃薯生产特征。在这方面的影响因素中,马铃薯的种植年限 (X4)、单位面积投入 (X5) 以及单位面积产量 (X7) 这 3 个影响因素通过了显著性检验。从表 2 的回归结果中可以看出,变量 X4 的回归系数为正,这说明马铃薯的种植年限对农户种植马铃薯的意愿影响为正向影响。也就是说,种植年限越长,农户积累的种植经验也就越多,对马铃薯种植的依赖性也就越高,农户的种植意愿也就越强烈。单位面积投入 (X5) 这一变量的回归系数为负值,这表明单位面积投入对农户种植的意愿产生负向影响,即单位面积投入越高,农户越不愿意种植马铃薯。这是合乎常理的,单位面积投入代表生产成本,成本越高,种植的积极性就会受到影响,种植意愿也就变弱。单位面积产量 (X7) 的回归系数为正值,该变量对农户种植意愿的影响为正向的,即马铃薯的单位面积产量越高,越能提高农户种植马铃薯信心与积极性,进而种植意愿也就越高。无论是用来出售还是自己食用,马铃薯的单位面积产量越高越能提高农户的种植意愿。

(3) 农户家庭收入特征。虽然在建模结果中显示变量 X8 (马铃薯收入占家庭收入的比重)并没有通过显著性检验,但是在建模过程中该变量并没有被淘汰,这说明该变量对农户种植意愿也存在一定的影响,只是影响并没有像其他变量一样那么显著。

(4) 资源禀赋约束。在这方面,土地使用情况 (X10) 以及资金使用情况 (X11) 通过了显著性检验。变量 X10 与 X11 这两个变量在回归中的系数均为负值,这说明,土地和资金的使用限制越多,农户对于种植马铃薯的意愿也就越低。这与现实是相符的,生产要素的使用受到约束,自然生产就会受到不利的影 响,相应的农户种植马铃薯的意愿也就下降。

(5) 外部环境。在这一方面的影响因素中,销售渠道 (X14) 通过了显著性检验,而且该变量在回归方程中的系数为正值,这说明销售渠道越畅通,农户种植马铃薯的意愿也就越高。这是显而易见的问题,农户种植的马铃薯有了很好的销售保障,他们便可以没有后顾之忧的去生产种植,自然种植意愿也就很高。而变量种植相关补贴 (X13) 虽然没有通过显著性检验,但该变量在逐步回归过程中被保留下来,这说明该变量对农户种植意愿也存在一定的影响。

3 基于文本分析的农户种植问题分析

为了更全面地研究影响农户种植马铃薯意愿的因素,在问卷设计中,不仅设置了客观性的问题,也设置了主观题“您在马铃薯种植、加工以及销售环节还存在什么问题?需要哪些方面的政府政策支持及补贴?”。这一问题有半数以上的农户做出了回答,认为在生产种植、加工及销售方面存在问题,需要政府

表 2 逐步回归筛选变量之后的 Logistic 回归输出结果

	估计值	标准差	Z 值	P 值
解释变量	1.111 6	2.685 3	0.414 0	0.678 9
X3	-4.502 4	2.404 4	-1.873 0	0.061 1
X4	2.402 4	0.846 0	2.840 0	0.004 5**
X5	-0.005 4	0.001 8	-3.016 0	0.002 6**
X7	0.003 2	0.001 0	3.132 0	0.001 7**
X8	-1.154 7	0.745 4	-1.549 0	0.121 4
X10	-2.850 2	1.098 6	-2.594 0	0.009 5**
X11	-2.807 1	1.432 1	-1.960 0	0.049 9*
X13	1.734 5	1.132 8	1.531 0	0.125 7
X14	1.284 7	0.718 9	1.787 0	0.073 9

注:显著性代码 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’

表 3 Logistic 回归分类结果

	Yes	No
Yes	123	5
No	18	3

的帮助。根据农户的回答,文中采用文本分析的方法,利用 R 软件中的“wordcloud”包以及“Rwordseg”包编写代码绘制了农户反映问题的词云及聚类分析图,如图 1 与图 2 所示。



图 1 农户反映问题词云

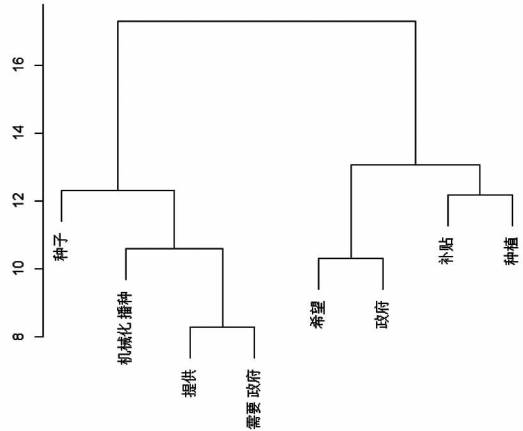


图 2 农户反映问题聚类分析

从图 1 和图 2 中可以看出,农户迫切需要政府的帮助,特别需要政府在种植方面的帮助,需要政府提供相应的补贴,提高机械化播种的程度。与此同时,农户反映他们非常需要政府提供优良的马铃薯种子以及生产种植的技术,做好销售渠道畅通无阻的保障。该文本分析结果与建模结果相互印证相互补充,在建模过程中,变量 X13(种植相关补贴)没有通过显著性检验,但在建模过程中没有被淘汰掉,这说明该变量对农户种植意愿存在一定影响,而文本分析的结果恰好印证了这一结论,这也是对定量模型的一种补充。

4 结论

通过对华北地下水漏斗区——河北地区 149 家农户的马铃薯种植意愿影响因素研究,发现农业人口占家庭人口的比例、马铃薯种植年限、单位面积投入、单位面积产量、土地、资金以及销售渠道等 7 个因素会对马铃薯的种植意愿产生显著性影响,其中农业人口占家庭人口的比例、单位面积投入、土地与资金使用情况这 4 个因素的影响是负向影响,而马铃薯种植年限、单位面积产量以及销售渠道等这 3 个因素产生的影响是正向效果。农户的基本情况如年龄和文化程度、马铃薯收入占家庭收入的比重、用水情况、培训以及补贴等 6 个因素并没有通过显著性检验,也就是说这 6 个因素对农户种植马铃薯意愿的影响不明显。通过文本分析绘制农户反映问题词云图与聚类分析图发现,绝大多数农户需要政府的帮助,特别是种植方面,需要政府提供优良的马铃薯种子,生产种植技术的培训,提供种植相关的补贴,提高播种机械化程度,以及销售方面的信息,做好销售渠道的保障。

参考文献

- [1] 王亚坤,王慧军,杨振立. 我国谷子种植户持续种植意愿的影响因素研究. 中国农业资源与区划, 2016, 37 (2): 96~102
- [2] 罗其友. 农业区域发展学科建设战略思考. 中国农业资源与区划, 2012, 33 (4): 1~4
- [3] 周清明. 农户种粮意愿的影响因素分析. 农业技术经济, 2009, (5): 25~30
- [4] 宋金田,祁春节. 农户柑橘种植意愿及影响因素实证分析——基于我国柑橘主产区 152 个农户的调查. 华中农业大学学报(社会科学版), 2012, (4): 17~21
- [5] 刘月华,王竹青,刘焰,等. 烟农种植意愿的实证分析——基于恩施山区烟农的调查数据. 中国农学通报, 2014, 30 (11): 116~120
- [6] 闫志明,蒲春玲,胡赛,等. 基于新疆南部地区实证的棉农生产行为影响因素分析. 中国农业资源与区划, 2015, 36 (6): 139~145
- [7] 宋长鸣. 蔬菜价格波动背景下生产者种植意愿变化研究——兼论对 Logistic 模型的重新解读. 中国农业大学学报, 2016, 21 (1): 147~156
- [8] 恰亚诺夫. 农民的经济组织. 萧正洪译. 北京: 中央编译出版社, 1996. 9~11
- [9] 舒尔茨. 改造传统农业. 梁小民译. 北京: 商务印书馆, 1987. 90~95
- [10] 吴喜之. 复杂数据统计方法——基于 R 的应用. 北京: 中国人民大学出版社, 2012

