

基于灰色马尔科夫模型的吉林省农民收入预测

李晶晶¹, 刘文明^{2*}, 费红梅¹

(1. 吉林农业大学, 长春 130118; 2. 吉林省农业科学院, 长春 130033)

摘要: 本文选取吉林省农民家庭平均每人可支配收入作为研究对象, 运用2005~2018年的宏观统计数据, 首先构建灰色预测模型, 在此基础上利用马尔科夫链对灰色预测值进行修正, 将模拟值与真实值进行对比, 模拟程度较高。因此灰色马尔科夫预测模型能够对吉林省农民家庭平均每人可支配收入进行合理的预测, 最终提出促进农民收入的建议。

关键词: 吉林省; 农民家庭平均每人可支配收入; 灰色马尔科夫模型

中图分类号: F323.8

文献标识码: A

文章编号: 2096-5877(2020)06-0106-04

Prediction of Farmers' Income in Jilin Province Based on Gray Markov Model

LI Jingjing¹, LIU Wenming^{2*}, FEI Hongmei¹

(1. Jilin Agricultural University, Changchun 130118; 2. Jilin Academy of Agricultural Sciences, Changchun 130033, China)

Abstract: In this paper, the average disposable income of farmers' families in Jilin Province was selected as the research object. Based on the macro statistical data from 2005 to 2018, the gray prediction model was first constructed. On this basis, the predicted gray value was modified by using Markov chain, and the simulated value was compared with the real value, with a higher simulation degree. Therefore, the gray Markov prediction model can reasonably forecast the average disposable income of farmers in Jilin Province and finally put forward suggestions to promote farmers' income.

Key words: Jilin Province; Average disposable income of farmers' families; Gray Markov model

吉林省位于中国的黄金玉米带, 是全国重要的商品粮基地, 同时也是农业人口大省, 因而农民增收显得至关重要。2016年吉林省“十三五”规划提出到2020年农民收入比2010年翻一番的目标, 本文对吉林省农民收入进行适时预测, 观察农民收入的变化趋势, 以期为实现农村经济发展的合理目标制定提高农民收入相关政策提供理论支持。国内学者的相关研究主要围绕灰色预测法、非线性退势法、多元线性回归法、主成分分析法、小波变换法、ARMA法和BP神经网络预测法展开^[1-6]。从农民收入水平看, 农民收入绝对数

额在上升, 但增长速度放缓^[7]。从收入结构看, 工资性收入是影响农民收入最主要因素, 需要大力发展非农产业, 转移农村剩余劳动力, 以增加农民收入^[8]。从收入的变化周期看, 全国农民人均纯收入增长与变化存在着7年和40年两个主要波动周期^[9]。吉林省农民家庭平均每人可支配收入在全国没有绝对优势, 但从收入结构的角度看, 2014年其家庭经营纯收入居全国首位, 而家庭经营收入主要是以种粮收入为主, 农民增收的经验对于我国部分省份具有一定的指导意义。本文以吉林省作为研究对象, 选取2005~2018年农民家庭平均每人可支配收入的宏观数据, 运用灰色马尔科夫组合模型测算吉林省农民收入。并对2019~2021年农民家庭平均每人可支配收入做出合理预测, 以实现农民收入快速提高的目标。

1 数据和模型说明

1.1 数据的选取和说明

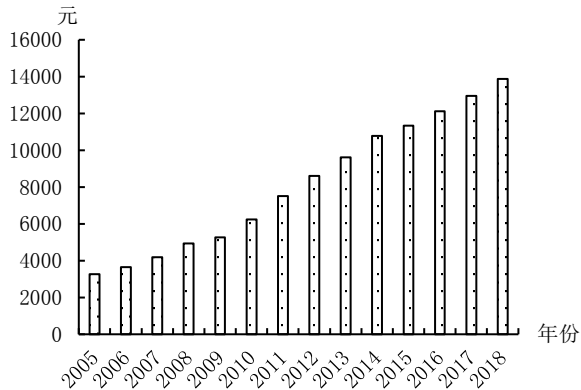
从图1可以看出, 2005~2018年吉林省农民家庭平均每人可支配收入总体呈现上升趋势,

收稿日期: 2018-12-24

基金项目: 国家自然科学基金(71640039); 吉林省科技发展计划项目(20150418017FG); 吉林省农业科学院创新工程项目(c7208000305、Y81990602); 吉林省社会科学基金(2017BS15)

作者简介: 李晶晶(1991-), 女, 在读博士, 研究方向: 农业经济理论与政策。

通讯作者: 刘文明, 女, 博士, 副研究员, E-mail: 327510910@qq.com



数据来源:2006~2017年《吉林统计年鉴》

图1 吉林省农民家庭平均每人可支配收入

2005年为3263.99元,2018年高达13748元,2004~2018年十四年间年均增加748.86元。以2005年为基期,消除物价影响因素,2005~2018年吉林省农民家庭平均每人可支配收入的名义增长速度和实际增长速度总体呈下降趋势。名义年均增长率达到12.3%,而实际年均增长速率维持在8.8%左右,名义增长速度一直高于实际增长速度。按照实际增长速度划分具体分为四个阶段:一是2005~2008年农民收入增速处于上升阶段。特别是2007~2008年上升幅度最大,达到11.8%,这段时期国家全面取消农业税,发布中央一号文件,加大对“三农”的投入,使得农民收入增速较快。二是2008~2009年农民收入增长速度处于骤降阶段,仅为6.01%,主要是受到2008年金融危机的影响,农产品销售市场不景气,阻碍了农民收入的增长。三是2009~2011年农民收入增长速度处于骤升阶段,增长率从2009年的6.01%上升到2011年的14.23%,这段时期国家加大了粮食直补、农资综合补贴等的支持力度,对重要农产品实施最低收购价政策和临时收储政策,提高了农产品的价格,促使农民的收入增长。四是2011~2018年农民收入增长速度处于“阶梯式”下降阶段,增长率由2011年的14.23%下降到1.1%。在我国经济增长速度由高速增长向中高度增长的新常态背景下,农民自身资源有限,要素配置效率不佳^[9],吉林省农民收入增长乏力,因此农民收入增速放缓。总体来看,随着2016年国家取消东北地区玉米的临储价格政策,吉林省农民收入短期内会受到影响,而未来仍然保持上升态势。

1.2 模型的选择和说明

农民收入的预测包含多种研究方法,但在实际预测中,对研究对象仅选取一种预测方法可能会出现预测精度低等问题,可以将两种及两种以上的预测方法融合起来,充分发挥各自的优势,

达到互惠共利的效果。灰色GM(1,1)预测方法需要少量的样本就可以对事物的长期发展趋势做出较精准的预测,而马尔科夫(Markov)模型需要大量样本,样本需具备平稳特性,这样才能对这种波动性较大的样本做出较为准确的中长期预测,二者结合的灰色马尔科夫(GM-Markov)预测模型融合了各自的优点,弥补了彼此的缺点,能够对吉林省农民收入作出更为精准的预测。

1.2.1 灰色GM(1,1)预测方法

(1)设原始时间序列 $x^{(0)}(k)$ 有n个样本: $x^{(0)}(k)=[x^{(0)}(1), x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(n)]$ 。

(2)由原始时间序列进行一次累加之后得到新的序列: $x^{(1)}(k)=[x^{(1)}(1), x^{(1)}(2), x^{(1)}(3), \dots, x^{(1)}(n)]$ 。

(3)对累加后形成的新序列建立一阶微分方程: $\frac{dx^{(1)}}{dt} + \alpha x^{(1)} = \beta$ 。其中, α 表示发展系数,代表原始序列和累加序列的变化趋势, β 表示灰作用量,反映样本的波动变化。运用最小二乘法进行参数估计,得到上式的矩阵形式:

$$[\alpha\beta]^T = (Z^T Z)^{-1} Z^T Y, \text{ 其中:}$$

$$Z = \begin{pmatrix} -0.5 \times [x^{(1)}(1) + x^{(1)}(2)] & 1 \\ -0.5 \times [x^{(1)}(2) + x^{(1)}(3)] & 1 \\ \dots & \dots \\ -0.5 \times [x^{(1)}(n-1) + x^{(1)}(n)] & 1 \end{pmatrix},$$

$$Y = \begin{pmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \dots \\ x^{(0)}(n) \end{pmatrix}$$

(4)解微分方程后得到灰色GM(1,1)预测模型: $\hat{x}^{(1)}(k+1) = \left[x^{(1)}(0) - \frac{\beta}{\alpha} \right] e^{-\alpha k} + \frac{\beta}{\alpha}$ 。

(5)再将 $\hat{x}^{(1)}(k+1)$ 累减还原,得到原始序列的预测值: $\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k)$ 。

但是可能会存在很多研究问题不适合灰色GM(1,1)预测模型,因此要对其进行后验差检验,通过后验差比值C和小误差概率P检验,其中, $C = S_1/S_0$, S_1 为残差的标准差, S_0 为原始序列的标准差, $P = \left\{ \left| x^{(0)}(k) - \hat{x}^{(0)}(k) \right| < 0.6745S_0 \right\}$ 。当 $P > 0.95$, $C < 0.35$ 时,综合预测精度等级为一级,通过了后验差检验。

1.2.2 马尔科夫(Markov)预测模型

马尔科夫(Markov)预测模型是按照事物目前的状态以及呈现的状态转移发展趋势,预测事物在未来的区间可能处于某个状态,并通过取区间

中点进行修正,得出更为准确的预测值。马尔科夫(Markov)预测模型表示为: $x(k) = P(1)^k$ 。

其中, $x(k)$ 为预测样本在时刻k的状态, $P(1)$ 为一步转移概率矩阵。

1.2.3 灰色马尔科夫(GM-Markov)预测模型

(1) 构建灰色 GM(1,1) 预测模型, 求出 $\hat{x}^{(0)}(k+1)$, 得到灰色预测值。

(2) 计算原始值和模拟值的相对误差

$$P_w = \frac{\Delta k}{x^{(0)}(k)} \times 100\%, \Delta k = x^{(0)}(k) - \hat{x}^{(0)}(k)。$$

(3) 按照相对误差划分区间。 $\otimes_i = [\otimes_{1i}, \otimes_{2i}]$, 其中, \otimes_{1i} 是各个区间相对误差的最小限制, \otimes_{2i} 是各个区间相对误差的最大限制。

(4) 计算状态转移矩阵概率。 $p_{ij} = \frac{n_{ij}(k)}{n_i}$, 其

中, $n_{ij}(k)$ 表示从状态 A_i 经过一步转移到状态 A_j 的次数, n_i 表示在样本中 A_i 出现的总次数,具体的转移矩阵如下。由此可以确定矩阵的初始状态,从而预测事物未来的发展状态。

$$P_k = \begin{bmatrix} p_{11}(k)p_{12}(k)\cdots p_{1n}(k) \\ p_{21}(k)p_{22}(k)\cdots p_{2n}(k) \\ \dots\dots\dots \\ p_{n1}(k)p_{n2}(k)\cdots p_{nn}(k) \end{bmatrix}$$

(5) 修正灰色预测值。确定了未来农民收入的变动区间后,需要取区间中点对灰色预测值进行修正,最终得到灰色马尔科夫预测值:

$$Y_{G-M} = \frac{\hat{x}^{(0)}(k+1)}{1 - \frac{1}{2}(\otimes_{1i} + \otimes_{2i})}$$

2 吉林省农民收入预测

2.1 灰色 GM(1,1) 预测

根据图1的2005~2018年吉林省农民收入的数据以及灰色马尔科夫预测模型的原理,运用Matlab程序计算,结果为:

$$[\alpha\beta]^T = (Z^T Z)^{-1} Z^T Y = \begin{bmatrix} -0.10244396 \\ 3806.839784 \end{bmatrix}$$

$$x^{(0)}(1) = 3263.99$$

$$\begin{aligned} \hat{x}^{(1)}(k+1) &= \left[x^{(0)}(1) - \frac{u}{a} \right] e^{-ak} + \frac{u}{a} \\ &= 40424.04884e^{0.10244396k} - 37160.05884 \end{aligned}$$

将各年数据代入上式,对计算结果累减还原,得到原始样本的模拟值、残差和相对误差。由表1可知:后检验差比值C为0.043,小误差概率P为1.0000,建立的灰色GM(1,1)模型预测精度等级

表1 GM(1,1)误差检验表

年份	原始值	模拟值	残差	相对误差(%)
2005	3263.99	3263.99	0	0
2006	3641.13	4360.77	-719.64	-19.76
2007	4189.89	4831.19	-641.30	-15.31
2008	4932.74	5352.36	-419.62	-8.51
2009	5265.91	5929.75	-663.84	-12.61
2010	6237.40	6569.43	-332.03	-5.32
2011	7510.00	7278.11	231.89	3.09
2012	8598.20	8063.24	534.96	6.22
2013	9621.20	8933.07	688.13	7.15
2014	10780.10	9896.73	883.37	8.19
2015	11326.00	10964.35	361.65	3.19
2016	12122.94	12147.13	-24.19	-0.20
2017	12590.00	13457.51	-507.51	-3.92
2018	13748.00	14609.25	-861.25	-6.26

表2 GM(1,1)2019~2021年吉林省农民人均纯收入

年份	预测结果(元)
2019	16517.60
2020	17899.50
2021	18873.51

为一级,据此可以预测2019~2021年吉林省农民家庭平均每人可支配收入(表2)。

2.2 马尔科夫链修正灰色预测值

为了使吉林省农民收入的预测值更接近真实值,利用马尔科夫链对灰色预测值进行修正,根据表2计算的相对误差,划分为6个区间,最小区间限制为-20%,最大区间限制为10%,区间划分上含下不含(表3)。

表3 吉林省农民收入状态区间划分

标签	区间	年份	数量
A ₁	(-20%, -10%)	2006, 2007, 2009	3
A ₂	(-10%, -5%)	2008, 2010, 2018	3
A ₃	(-5%, -2%)	2017	1
A ₄	(-2%, 2%)	2005, 2016	2
A ₅	(2%, 5%)	2011, 2015	2
A ₆	(5%, 10%)	2012~2014	3

由此得到一步状态转移概率矩阵如下:

$$P(1) = \begin{bmatrix} 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \\ 2/3 & 0 & 0 & 0 & 1/3 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 1/3 & 2/3 \end{bmatrix}$$

由此计算出2005~2018年灰色马尔科夫预测值,并与GM(1,1)预测值进行比较(表4)。

表4 GM(1,1)和GM-Markov预测结果比较

年份	原始值(元)	GM(1,1)		GM-Markov	
		模拟值(元)	相对误差(%)	模拟值(元)	相对误差(%)
2005	3 263.99	3 263.99	0	3 263.99	0
2006	3 641.13	4 360.77	-19.76	3 791.98	-4.14
2007	4 189.89	4 831.19	-15.31	4 201.04	-0.27
2008	4 932.74	5 352.36	-8.51	4 978.94	-0.94
2009	5 265.91	5 929.75	-12.61	5 156.31	2.08
2010	6 237.40	6 569.43	-5.32	6 111.10	2.02
2011	7 510.00	7 278.11	3.09	7 542.08	-0.42
2012	8 598.20	8 063.24	6.22	8 717.02	-1.38
2013	9 621.20	8 933.07	7.15	9 657.37	-0.38
2014	10 780.10	9 896.73	8.19	10 699.17	0.75
2015	11 326.00	10 964.35	3.19	11 362.02	-0.32
2016	12 122.94	12 147.13	-0.20	12 122.94	0
2017	12 590.00	13 457.51	-3.92	13 002.43	-0.41
2018	13 748.00	14 609.25	-6.26	14 405.07	-3.83

由表4可以看出,经过马尔科夫修正的预测值更接近真实值,总体平均误差率仅为-0.56%,而灰色预测值的平均误差率相对较高,达到-3.23%,这说明灰色马尔科夫预测模型可以较为准确地预测吉林省农民收入,最后对吉林省2019~2021年农民收入进行预测。由于2018年划分为第二个状态区间,取 $P(1)$ 矩阵中第二行中最大值处于的状态区间,因此2019年吉林省农民家庭平均每人可支配收入划分为第一个状态区间,取其状态区间中点进行马尔科夫修正,结果为14 363.13元。

再由同样方法预测2020和2021年吉林省农民家庭平均每人可支配收入,结果分别为15 564.78元、16 411.75元。

2.3 后验差检验

灰色马尔科夫预测值较为接近真实值,但是还要通过后验差比值 C 和小误差概率 P 对其进行后验差检验,这样才能保证模型可靠。 S_1 代表残差的标准差, S_0 代表原始值的标准差, $C = S_1/S_0 = \frac{152.87}{3 521.14} = 0.043$ 。

表5 预测精度等级划分

预测等级	后验差比值 C	小误差概率 P	预测等级	后验差比值 C	小误差概率 P
一级	< 0.35	> 0.95	三级	> 0.7	< 0.45
二级	> 0.8	< 0.5	四级	≤ 0.7	≥ 0.65

$P = \left\{ \left| x^{(0)}(k) - \hat{x}^{(0)}(k) \right| < 0.674 5S_0 \right\}$, 最终求得 $P = 1$,按照表5的评定,当 $C < 0.35$, $P > 0.95$,模型等级为一级,预测精度高。据此灰色预测模型通过了后验差检验,可以对吉林省农民收入作出可靠预测。

3 结论及建议

3.1 结论

灰色马尔科夫预测模型预测结果和实测值的拟合度比灰色GM(1,1)预测模型更高,能客观反映吉林省农民家庭平均每人可支配收入的变化趋势,预测的结果较为精准,2005~2018年马尔科夫修正值平均相对误差率只有-0.56%,最大相对

误差率为-4.14%,后验差比值 C 为0.043,小误差概率为1,符合精度要求。从预测结果来看,2018~2020年吉林省农民家庭平均每人可支配收入名义年均增长5.76%,2011~2018年为9.16%,农民收入增长速度趋缓。吉林省需要及时作出相关政策调整,以增加农民收入。吉林省农民家庭平均每人可支配收入具有以家庭经营收入为主,工资性收入严重滞后的特性,吉林省需要继续保持农业的优势,同时兼顾非农产业的发展。

3.2 建议

3.2.1 发展适度规模经营

吉林省拥有松辽平原的天然地理条件,黑土资源丰富。但是吉林省土地细碎化严重,(下转第139页)

- 春:长春工业大学,2013.
- [10] 贾楠,赵二劳,范建凤,等.洋葱黄酮超声波辅助提取工艺优化及其抗氧化活性研究[J].河南工业大学学报(自然科学版),2015,36(2):62-65,70.
- [11] 刘超,高炯杰.超声波辅助法提取毛葱黄酮的工艺研究[J].食品研究与开发,2015,36(18):50-54.
- [12] 雷红松.大孔吸附树脂分离纯化荷叶黄酮与生物碱研究[D].长沙:湖南农业大学,2008.
- [13] 杨玲娥,康永锋,马晨晨,等.超声波辅助提取番木瓜提取液抗氧化活性研究[J].吉林农业科学,2013,38(1):75-79.
- [14] Wang P F, Mu X P, Du J J, et al.Flavonoid content and radical scavenging activity in fruits of Chinese dwarf cherry (*Cerasus humilis*) genotypes [J]. Journal of Forestry Research, 2018, 29(1): 55-63.
- [15] 方玉梅,张春生,谭萍,等.金针菇黄酮类化合物的抗氧化性作用[J].食品研究与开发,2012,33(3):15-18.
- [16] 张黎明,李瑞超,郝利民,等.响应面优化玛咖叶总黄酮提取工艺及其抗氧化活性研究[J].现代食品科技,2014,30(4):233-239.
- [17] 段筱杉,张朝辉,应锐,等.海芦笋黄酮的抗氧化作用及对CCl₄致小鼠急性肝损伤的保护作用[J].水产学报,2017,41(12):1946-1955.
- [18] 钟方丽,王文姣,王晓林,等.猴腿蹄盖蕨、玉竹总黄酮的体外抗氧化活性研究[J].中国食品添加剂,2016(6):65-72.
- [19] 周宁,赵晓璐,屈潇毅,等.超声辅助提取甘蔗叶总黄酮与还原力的测定[J].应用化工,2016,45(10):1883-1886.
- [20] 包骥林,王丹,陆葵青,等.桂花黄酮类化合物体内抗氧化活性的研究[J].食品研究与开发,2018,39(21):14-20.
- [21] Xican Li. 2-Phenyl-4,4,5,5-tetramethylimidazoline-1-oxyl 3-oxide (PTIO·) Radical scavenging: A New and Simple Antioxidant Assay *In Vitro*[J]. Journal of Agricultural & Food Chemistry, 2017, 65: 6288-6297.
- [22] 谢晓凤,童莲花,童德胜,等.马齿苋总黄酮的提取及其浓缩汁抗氧化性研究[J].食品工业,2013,38(2):192-197.
- [23] Lijun Sun, Jianbao Zhang, Xiaoyun Lu, et al.Evaluation to the antioxidant activity of total flavonoids extract from persimmon (*Diospyros kaki* L.)leaves[J]. Food and Chemical Toxicology, 2011, 49(10): 2689-2696.
- [24] Qiuhui Hu, Jie Yu, Wenjian Yang, et al.Identification of flavonoids from *Flammulina velutipes* and its neuroprotective effect on pheochromocytoma-12 cells[J]. Food Chemistry, 2016, 204: 274-282.
- [25] 李胜华,臧汝瑛,刘秀河,等.香辛料对蔬菜中亚硝酸盐的清除能力研究[J].山东食品发酵,2012(2):3-7.
- [26] 胡利,贾冬英,姚开,等.桑叶黄酮对亚硝酸盐的体外清除作用研究[J].中国调味品,2015,40(3):1-5.
- [27] Narumol P, Weerasak S, Patcharawee N, et al.In vitro antioxidant activity study of novel chromone derivatives[J].Chemical Biology&Drug Design, 2012, 79(6): 981-989.

(责任编辑:王昱)

(上接第109页)

小农户分散式生产阻碍了规模化经营进程^[10],从职业农民角度来讲,发展适度规模经营是其增收的有效方式。因此需要积极发展家庭农场和职业农民,有利于提高粮食生产效率,促进农民收入快速增长。

3.2.2 调整农业经营结构

从家庭经营收入内部的比重变化来看,牧业收入比重始终次于种植业,居于第二位。特别是2000年开始吉林省畜牧业数量和质量均有飞跃,2000~2009年牧业收入所占比重均保持在10%以上;而从牧业产值在农林牧渔总产值所占比重来看,2016年已达到38.35%,且种植业本身具有周期性和季节性,因此发展畜牧业既可以提高劳动效率,又可以延长农副产品的产业链,“农牧结合”的生产模式有利于农业的可持续发展。

3.2.3 支持非农产业发展

从兼业农民的角度来讲,需要提高非农产业的就业能力,以增加工资性收入。2017年吉林省的城镇化率上升到56.65%,这主要是人口数量的城镇化,更重要的是注重发展人口质量的城镇化,即实现人民生活水平和幸福感的城镇化。应提升城镇化的发展质量,繁荣二三产业,为农民提供更多的就业机会和初来城镇的生活补贴,保

证兼业农民转移的持久性和稳定性,实现农民的工资性收入稳步增长。

参考文献:

- [1] 王浩.基于非线性退势法的农民收入预测[J].江苏农业科学,2012,40(6):376-379.
- [2] 卢小丽,何光.基于ARMA模型的四川省农村居民收入趋势预测[J].中国农学通报,2012,28(5):110-114.
- [3] 彭婵娟.福建省农民人均纯收入的组合预测分析[J].中国农学通报,2013,29(35):152-157.
- [4] 张海燕.基于多元线性回归模型的四川农村居民收入增长分析[J].统计与决策,2010(13):88-90.
- [5] 佟光霁,张琳.基于灰色马尔科夫模型的农民收入质与量预测[J].商业研究,2012(9):135-140.
- [6] 曹跃群,周加斌,吴颖,等.基于小波变换的农民收入增长波动关系预测分析[J].华东经济管理,2009,23(5):44-48.
- [7] 吴迪,刘文明,舒坤良.吉林省自然灾害对种植业及农民收入的影响研究[J].东北农业科学,2016,41(3):104-108.
- [8] 余晓洋,田帅,刘帅.杭州市农民工工资性收入变动及其影响因素分析[J].东北农业科学,2017,42(3):50-54.
- [9] 刘文明.基于农户角度的吉林省耕地可持续利用分析[J].东北农业科学,2018,43(3):39-42.
- [10] 费红梅,刘文明,王立,等.农户土地流出处意及其影响因素分析[J].东北农业科学,2017,42(6):69-72.

(责任编辑:王丝语)