

灰色系统理论在陕西省农作物秸秆可收集量预测中的 SAS 应用

王涛¹, 宋喜芳^{1*}, 常小箭¹, 赵永锋², 王辉²

(1. 陕西省西安市农业技术推广中心, 陕西西安 710061; 2. 陕西省西安市周至县土壤与肥料工作站, 陕西西安 710400)

摘要 [目的]在 SAS 环境下,运用灰色系统理论对陕西省农作物秸秆可收集量进行预测。[方法]以草谷比和可收集系数估算 2005—2015 年陕西省农作物秸秆可收集量。以农村就业人口、农作物播种面积、农用化肥施用量和农业机械总动力作为影响农作物秸秆可收集量的 4 个主要因素进行灰色关联度分析。在 SAS 环境下,利用 GM(1,1) 灰色模型和基于 GM(1,1) 的多元回归模型对 2016—2020 年的陕西农作物秸秆可收集量进行预测,并对模型精度与误差进行分析比较。[结果]基于 GM(1,1) 的多元回归模型预测精度高于 GM(1,1) 模型的精度,较准确预测了 2016—2020 年陕西农作物秸秆的可收集量。[结论]准确预测农作物秸秆可收集量可为政府开展农业面源污染防治、提高秸秆综合利用提供强有力的数据支撑。

关键词 陕西; 秸秆; 可收集量; 灰色预测; SAS 模型

中图分类号 S-058 文献标识码 A 文章编号 0517-6611(2018)06-0186-04

Application of Gray System Theory in the Prediction of Shaanxi Province' Straw Resource Collectable Amount Based on SAS

WANG Tao, SONG Xi-fang, CHANG Xiao-jian et al (Xi'an Agricultural Technology Extension Center, Xi'an, Shaanxi 710061)

Abstract [Objective] To predict Shaanxi Province's straw resource' collectable amount based on SAS by using gray system theory. [Method] Collectable amount of straw resource from 2005 to 2015 was estimated based on the residue to grain ratio and collection coefficient. To analyze the rural employ population, crop sown area, consumption of chemical fertilizers and total power of agricultural machinery as four major factors affecting collectable amount of crop straw. GM(1,1) gray model and multiple regression model based on GM(1,1) were used to predict the collectable amount of crop straw in Shaanxi from 2016 to 2020, using SAS. [Result] The prediction accuracy of multiple regression model based on GM(1,1) was higher than GM(1,1) prediction. It accurately predicted the collectable amount of crop straw in Shaanxi Province from 2016 to 2020. [Conclusion] The reliable prediction of crop straw' collectable amount in Shaanxi provides a strong data support for carrying out the prevention and control of agricultural surface pollution and improving the comprehensive utilization of straw for the government.

Key words Shaanxi; Crop straw; Collectable amount; Gray prediction; SAS model

能源短缺和环境污染问题已经成为制约人类社会发展的主要瓶颈。生物质能以其可再生、资源丰富、生态环境友好而逐渐成为一种重要的新替代能源,它仅次于煤炭、石油和天然气而居于世界能源消费总量第 4 位^[1]。农作物秸秆作为农业生产的剩余物,是一种重要的生物质资源,其中相当一部分农作物秸秆可作为能源资源进行规模化利用。有效利用这部分生物质资源不但可缓解能源需求压力,消除废弃生物质资源带来的环境污染,而且是振兴农村经济、增加农村就业、提高农民收入、改善农村人居环境的需要^[2]。对陕西省秸秆可收集量进行分析及预测是建立和发展农村生物质能项目的前提和基础,可为打好农业面源污染攻坚战提供数据支持。目前,已采用常规方法对陕西省秸秆资源量进行估算^[3-5],但是尚无陕西省秸秆可收集量的影响因素分析及预测。为此,笔者基于灰色系统理论利用 SAS 统计分析软件,对陕西省秸秆可收集量进行分析和预测。

灰色系统是指部分信息已知、部分信息未知的系统,它介于一无所知的黑色系统与全部确知的白色系统之间。灰色预测在形式上只运用预测对象自身的时间序列建立模型。与其相关联的因素表面上没有参与运算和建模,并不是说那些因素对预测对象没有影响和作用。灰色系统的“灰”正体现在这里,如影响陕西省秸秆的可收集量与作物品种、耕作制度、收获方式等有密切的关系,同时也受到自然条件、科学发展等方面的制约,是多种因素综合作用的结果,秸秆可收

集量也具有某种灰色性。秸秆可收集量可看成是既有一定已知信息又含有未知信息的灰色系统^[6]。

根据 2005—2015 年陕西省主要农作物产量的统计数据,通过对近年相关研究文献的综合分析,确定主要农作物草谷比、作物秸秆可收集系数等估算参数,对陕西省秸秆可收集量进行估算,并利用灰色系统理论对陕西省农作物秸秆可收集量的影响因素进行灰色关联分析,并在 SAS 环境下利用 GM(1,1) 的多元回归模型进行预测。

1 陕西省农作物秸秆可收集量的估算

陕西省主要农作物产量数据是对秸秆可收集量进行科学准确估算的基础。结合相关文献,采用草谷比的方法对农作物秸秆可收集量进行估算。

1.1 估算方法 陕西省的主要农作物包括小麦、玉米、水稻、棉花、油菜籽等。农作物秸秆可收集量估算公式如下:

$$P_e = \sum_i^n \eta_i (\lambda_i \cdot G_i)$$

式中, P_e 为农作物秸秆可收集量(t); i 为农作物秸秆编号 ($i=1, 2, 3, \dots, n$); η_i 为第 i 种农作物的收集系数; λ_i 为第 i 种农作物的草谷比; G_i 为第 i 种农作物的年产量。

1.2 草谷比及收集系数的整理 综合王亚静等^[7]、李逸辰^[8]的研究成果,确定草谷比。结合陕西实际,选用王亚静等^[7]对农作物秸秆收集系数的研究数据,确定可收集系数。主要农作物产量数据来源于历年《陕西统计年鉴》。陕西主要农作物草谷比及可收集系数见表 1。

1.3 估算结果 经计算得出陕西省 2005—2015 年农作物秸秆的可收集量依次为 1 365.70 万、1 222.99 万、1 247.81 万、1 393.94 万、1317.04 万、1361.56 万、1408.01 万、1 457.98 万、

作者简介 王涛(1982—),男,陕西宝鸡人,农艺师,硕士,从事农业生态与资源保护研究。* 通讯作者,助理农艺师,硕士,从事农业生态与资源保护(生物统计方向)研究。

收稿日期 2017-08-10

表 1 主要农作物草谷比及可收集系数

Table 1 Residue to grain ratio of main crops and coefficient of collectable amount

作物种类 Crops species	草谷比 Residue to grain ratio	可收集系数 Coefficient of collectable amount
小麦 Wheat	1.27	0.83
水稻 Rice	0.94	0.83
玉米 Corn	1.10	0.83
其他谷类 Other grain	2.32	0.83
豆类 Beans	1.36	0.88
棉花 Cotton	2.62	0.90
油菜籽 Rapeseed	2.57	0.85
花生 Peanut	0.86	0.85
其他油类作物 Other oil crops	2.63	0.85
麻类 Bast fibre plants	6.65	0.87
糖类 Sugar crop	0.34	0.88
烟草 Tobacco	0.71	0.90

1 421.05 万、1 416.03 万、1 436.23 万 t。由此可知,从 2005 年以来,陕西省农作物秸秆可收集量总体呈波浪型增长趋势,2015 年较 2005 年增长了 5.16%,较 2006 年增长了 17.44%。

表 2 陕西省农作物秸秆可收集量相关影响因素及关联度

Table 2 Related factors and correlation degree of crop straw' collectable amount in Shaanxi Province

年份 Year	农村就业人员 Rural employed person//万人	播种面积 Sowing area 万 hm ²	农用化肥施用 量折纯量 Consumption of chemical fertilizers//万 t	农业机械总动力 Total power of agricultural machinery//万 kW
2005	949.22	439.12	147.30	1 406.30
2006	948.41	398.35	149.73	1 452.40
2007	925.66	404.47	158.81	1 576.10
2008	902.02	427.45	165.90	1 705.50
2009	871.48	415.41	181.32	1 778.50
2010	848.70	418.56	196.79	1 889.30
2011	816.40	418.10	207.27	2 036.60
2012	789.40	419.03	239.80	2 146.40
2013	774.60	418.35	241.73	2 326.08
2014	774.00	426.21	230.19	2 552.13
2015	778.00	428.42	231.95	2 667.27
关联度 Correlation degree	0.803 6	0.927 2	0.787 3	0.775 2
关联序 Correlation order	2	1	3	4

3 农作物秸秆可收集量的灰色模型预测

灰色预测法是一种对含有不确定因素的系统进行预测的方法。灰色系统分析实质上是将一些已知的数据序列,通过一定的方法处理,使其由散乱状态转向规律化,然后利用微分方程拟合,并由外延进行预测。灰色系统分析方法对于信息不完整或不完整的实际情况具有良好的适用性。

基于少量数据模型 GM(1,1) 的预测被称为灰色预测,是灰色系统应用的重要领域。由于该模型对数据量的要求少、不要求典型的概率分布,因而在许多领域有着广泛的应用。模型在农业生产中得到了较为广泛的运用^[9]。现采用 GM(1,1) 模型预测和基于 GM(1,1) 的多元回归模型预测 2 种方法预测陕西省农作物秸秆可收集量。

3.1 GM(1,1) 模型预测 以 2005—2015 年陕西省农作物

2 农作物秸秆可收集量影响因素的灰色关联度分析

灰色关联度分析是灰色系统理论的一种分析方法,是将因素之间发展趋势的相似或相异程度,亦即“灰色关联度”,作为衡量因素间关联程度的一种方法。灰色关联度分析法是将研究对象及影响因素的因子值视为一条线上的点,与待识别对象及影响因素的因子值所绘制的曲线进行比较,比较它们之间的贴进度,并分别量化,计算出研究对象与待识别对象各影响因素之间的关联度,通过比较各关联度的大小来判断待识别对象对研究对象的影响程度。

参考文献[2]的研究成果,农业生产条件是影响农作物秸秆可收集量的主要因素,依次是农村就业人员、农作物播种面积、农用化肥施用量和农业机械总动力。采用以上 4 个因素作为陕西秸秆可收集量的影响因素指标进行分析,各影响因素历年指标值及均值化关联分析后的关联度结果见表 2。

由表 2 可知,4 个影响因素的关联度都大于 0.700 0,按照关联度由大到小依次为农作物播种面积、农村就业人员、农用化肥施用量和农业机械总动力。

秸秆可收集量为单一变量,运用国际通用标准统计软件 SAS 基于灰色系统理论建模,对 2016—2020 年陕西省秸秆可收集量进行灰色预测,SAS 程序综合了颜杰等^[9]和孔超等^[10]的研究成果。

判断构建的拟合模型是否可信,误差是否很大,必须对其精度进行检验。检验误差有 4 种:相对误差、关联度、均方差比值和小误差概率。模型检验是建模后必不可少的工作。对灰色预测模型的检验一般有残差检验、关联度检验和后验差检验。一般情况下,最常用的是残差检验中的相对误差检验方法^[11]。GM(1,1) 预测的相对误差平均值为 2.34%,预测精度为三级。陕西省农作物秸秆可收集量 GM(1,1) 模型的预测结果见表 3。

表3 陕西省农作物秸秆可收集量 GM(1,1)模型的预测结果

Table 3 Forecasting result of crop straw' collectable amount with GM(1,1) in Shaanxi Province

年份 Year	可收集量 Collectable amount//万 t	预测量 Predicted amount//万 t	相对误差 Relative error//%
2005	1 365.70	1 365.70	0
2006	1 222.99	1 271.18	3.94
2007	1 247.81	1 291.84	3.53
2008	1 393.94	1 312.83	-5.82
2009	1 317.04	1 334.17	1.30
2010	1 361.56	1 355.85	-0.42
2011	1 408.01	1 377.88	-2.14
2012	1 457.98	1 400.27	-3.96
2013	1 421.05	1 423.03	0.14
2014	1 416.03	1 446.16	2.13
2015	1 436.23	1 469.66	2.33
2016	—	1 493.54	—
2017	—	1 517.81	—
2018	—	1 542.48	—
2019	—	1 567.54	—
2020	—	1 593.02	—

3.2 基于 GM(1,1)的多元回归模型预测

3.2.1 多元回归方程。根据灰色关联分析的结果,选取农

村劳动力(X_1)、农作物播种面积(X_2)、农业机械总动力(X_3)、化肥施用量(X_4)这4个影响因素作为自变量,秸秆可收集量(Y)为因变量,进行多元回归分析。基于2005—2015年陕西省秸秆可收集量及各影响因素的数据,应用SAS软件建立多元回归方程,回归模型拟合如下:

$$Y = 38.53501 - 0.52667X_1 + 4.03361X_2 + 1.34451X_3 - 0.08995X_4$$

在 $\alpha = 0.01$ 的置信度水平, F 比值的概率为0.0007,模型达到极显著水平,表明回归模型拟合效果优良。从回归方程可见,农作物播种面积(X_2)、农业机械总动力(X_3)与秸秆可收集量呈正相关,增大农作物播种面积、提高农业机械总动力可以提高秸秆可收集量。

3.2.2 4个因素的GM(1,1)模型预测。以2005—2015年农村劳动力、农作物播种面积、农用化肥施用量和农业机械总动力4个影响因素的历史数据为基础,运用GM(1,1)模型分别对各影响因素2016—2020年的发展预测,结果见表4。

表4 秸秆可收集量影响因素的GM(1,1)模型预测

Table 4 Forecast of influencing factors of crop straw' collectable amount with GM(1,1) in Shaanxi Province

年份 Year	乡村就业人员 Rural employed person//万人		播种面积 Sowing area 万 hm^2		农用化肥施用量折纯量 Consumption of chemical fertilizers//万 t		农业机械总动力 Total power of agricultural machinery//万 kW	
	真实值 Truth value	模拟值 Simulation value	真实值 Truth value	模拟值 Simulation value	真实值 Truth value	模拟值 Simulation value	真实值 Truth value	模拟值 Simulation value
2005	949.22	949.22	439.12	439.12	147.30	147.30	1 406.30	1 406.30
2006	948.41	942.87	398.35	406.99	149.73	156.28	1 452.40	1 458.52
2007	925.66	919.10	404.47	409.27	158.81	164.76	1 576.10	1 560.08
2008	902.02	895.93	427.45	411.57	165.90	173.70	1 705.50	1 668.70
2009	871.48	873.34	415.41	413.89	181.32	183.12	1 778.50	1 784.89
2010	848.70	851.32	418.56	416.21	196.79	193.06	1 889.30	1 909.16
2011	816.40	829.86	418.10	418.55	207.27	203.53	2 036.60	2 042.09
2012	789.40	808.94	419.03	420.90	239.80	214.58	2 146.40	2 184.28
2013	774.60	788.54	418.35	423.27	241.73	226.22	2 326.08	2 336.36
2014	774.00	768.66	426.21	425.65	230.19	238.50	2 552.13	2 499.04
2015	778.00	749.28	428.42	428.04	231.95	251.44	2 667.27	2 673.04
平均相对误差 Average relative error	—	0.011 6	—	0.009 0	—	0.042 2	—	0.008 9
2016	—	730.39	—	430.45	—	265.08	—	2 859.15
2017	—	711.98	—	432.87	—	279.47	—	3 058.23
2018	—	694.03	—	435.30	—	294.63	—	3 271.16
2019	—	676.53	—	437.75	—	310.62	—	3 498.92
2020	—	659.47	—	440.21	—	327.47	—	3 742.54

模型检验,一般要求平均相对误差越小越好,从表4的平均相对误差可见,乡村就业人数、农用化肥施用量预测精度均为二级,农作物播种面积、农业机械总动力预测精度均为一级。

3.2.3 基于GM(1,1)的多元回归模型预测。将4个因素的真实值和未来5年的预测值代入上述多元回归方程,可得出2005—2015年陕西省农作物秸秆可收集量的多元回归模

拟值和2016—2020年基于GM(1,1)的多元回归组合预测值。

由表5可知,基于GM(1,1)模型的多元回归模型对历史数据的预测精度为1.11%,预测精度为二级,较GM(1,1)模型预测精度有很大的提高。可见,基于GM(1,1)的多元回归模型集成了回归分析及灰色理论的优点,因此获得了较好的预测效果。

表 5 陕西省秸秆可收集量 GM(1,1) 预测和基于 GM(1,1) 的多元回归组合预测比较

Table 5 GM(1,1) prediction of crop straw' collectable amount and multiple regression model based on GM(1,1) in Shaanxi Province

年份 Year	真实值 Truth value	GM(1,1) 预测 GM(1,1) prediction	基于 GM(1,1) 的多元回归组合预测 Multiple regression model based on GM(1,1)
2005	1 365.70	1 365.70	1 381.40
2006	1 222.99	1 271.18	1 216.49
2007	1 247.81	1 291.84	1 254.24
2008	1 393.94	1 312.83	1 357.28
2009	1 317.04	1 334.17	1 338.97
2010	1 361.56	1 355.85	1 374.50
2011	1 408.01	1 377.88	1 390.50
2012	1 457.98	1 400.27	1 442.33
2013	1 421.05	1 423.03	1 433.81
2014	1 416.03	1 446.16	1 429.99
2015	1 436.23	1 469.66	1 428.80
平均相对误差 Average relative error	—	0.023 3	0.011 1
2016	—	1 493.54	1 489.35
2017	—	1 517.81	1 510.25
2018	—	1 542.48	1 530.73
2019	—	1 567.54	1 550.84
2020	—	1 593.02	1 570.49

4 讨论与结论

GM(1,1) 模型本身运用灰色系统理论,利用 SAS 程序实现模型分析与预测,仅从陕西省农作物秸秆可收集量的数据序列本身去寻找数据变化规律,并没有考虑其他相关影响因

素对可收集量的影响,因此相对误差略高,精度较低。

基于 GM(1,1) 的多元回归预测模型从灰色关联度的视角,选取 4 个影响因素进入回归分析模型,分析事物发展变化的影响因素及影响程度,通过 GM(1,1) 模型预测影响因素的变化趋势,然后按照回归分析的建模思想再构建预测模型。该预测方法集成了回归分析及灰色理论的优点,因此相对误差降低,预测效果较好。

加大农作物播种面积和提高农业机械总动力,将提高陕西省农作物秸秆可收集量。通过基于 GM(1,1) 模型的多元回归模型预测结果可见,陕西省农作物秸秆可收集量在 2010—2015 年呈不断增长趋势,到 2020 年将达 1 570.49 万 t,该研究可为政府开展农业面源污染防治、做出秸秆综合利用决策提供数据支撑。

参考文献

- [1] 石金明,范敏,席细平,等. 江西省秸秆资源量分析[J]. 能源研究与管理,2012(1):1-4.
- [2] 崔胜先,谢光辉,董仁杰. 灰色系统理论在黑龙江省农作物秸秆可收集量预测中的运用[J]. 东北农业大学学报,2011,42(8):123-130.
- [3] 朱建春,张增强,李荣华. 陕西关中地区作物秸秆资源的综合利用现状及其影响因素模型[J]. 农业环境与发展,2011,28(2):12-17.
- [4] 王宁堂,王军利,李建国. 农作物秸秆综合利用现状、途径及对策[J]. 陕西农业科学,2007(2):112-115.
- [5] 朱建春,李荣华,张增强,等. 陕西作物秸秆的时空分布、综合利用现状与机制[J]. 农业工程学报,2013,29(S1):1-7.
- [6] 冯蕾,王效华. 江苏省农作物秸秆资源量的灰色预测[J]. 能源研究与利用,2010(4):1-3.
- [7] 王亚静,毕于运,高春雨. 中国秸秆资源可收集利用量及其适宜性评价[J]. 中国农业科学,2010,43(9):1852-1859.
- [8] 李逸辰. 陕西省农作物秸秆资源量及其经济价值评估[D]. 长沙:中南林业科技大学,2014.
- [9] 颜杰,相丽驰,方积乾. 灰色预测模型及 SAS 实现[J]. 中国卫生统计,2006,23(1):75-85.
- [10] 孔超,刘元凤,沈续雷. 灰色预测模型的 SAS 程序改进[J]. 中国卫生统计,2008,25(6):640-641.
- [11] 刘树,王燕,胡凤阁. 对灰色预测模型残差问题的探讨[J]. 统计与决策,2008,3(1):9-11.

(上接第 148 页)

- [6] 斯特拉勒 A N. 现代自然地理学[M]. 北京:科学出版社,1983.
- [7] 周淑贞. 气象学与气候学[M]. 3 版. 北京:高等教育出版社,1997.
- [8] 黄秉维. 现代自然地理[M]. 北京:科学出版社,1999.
- [9] 郑度,李炳元. 青藏高原自然环境的演化与分异[J]. 地理研究,1990,9(2):1-10.
- [10] 符淙斌,王强. 气候突变的定义和检测方法[J]. 大气科学,1992,16(4):482-493.
- [11] SIEPIELSKI A M, MORRISSEY M B, BUORO M, et al. Precipitation drives global variation in natural selection[J]. Science,2017,355(6328):959.
- [12] 伍光和田连恕,胡双熙,等. 自然地理学[M]. 3 版. 北京:高等教育出版社,2000.
- [13] STOCKER T F, QIN D, PLATTNER G K, et al. IPCC,2013: Technical Summary[M]// Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Interg-

overnmental Panel on Climate Change. Cambridge: Cambridge University Press,2013:533-535.

- [14] METZ B, DAVIDSON O R, BOSCH P R, et al. Climate change 2007: Mitigation of climate change. Contribution of Working Group III to the fourth assessment report of the Intergovernmental Panel on Climate Change[M]. Cambridge: Cambridge University Press,2007.
- [15] 史培军,唐迪,方伟华,等. 从应对 2008 年低温雨雪冰冻巨灾看我国巨灾风险防范对策[J]. 保险研究,2008,3(5):9-12.
- [16] MITTON S. 2003 年欧洲热浪如何影响生态系统[J]. 科学观察,2008,3(1):61-62.
- [17] 王和. 对建立我国巨灾保险制度的思考[J]. 中国金融,2005(7):50-52.
- [18] SHI P J, YE Q, HAN G Y, et al. Living with global climate diversity—suggestions on international governance for coping with climate change risk[J]. International journal of disaster risk science,2012,3(4):177-184.

本刊提示 来稿请用国家统一的法定计量单位的名称和符号,不要使用国家已废除了的单位。如面积用 hm^2 (公顷)、 m^2 (平方米),不用亩、 尺^2 等;质量用 t (吨)、kg (千克)、mg (毫克),不用再担等;表示浓度的 ppm 一律改用 mg/kg、mg/L 或 $\mu\text{L}/\text{L}$ 。