

Feasibility study on the establishment of local economic early warning and prediction mechanism—Take the annual corn output in Changzhi area as an example

Huijuan Liu

Changzhi City Statistical Monitoring Center, Changzhi, Shanxi, 046000, China

Abstract

The deepening reform of statistical work in the new era should be anchored in serving the high-quality development of China's modern economy. Advanced analysis by the statistical system can help local governments maintain rapid economic growth while avoiding potential risks and hidden dangers. This paper uses the "Changzhi Statistical Yearbook" from 1997 to 2022 as data sources, employing WPS Excel, a commonly used software tool, to establish a Cobb-Douglas prediction model for annual corn production in Changzhi region. It also delves into the feasibility of establishing a long-term mechanism for early warning and predictive work in local economic statistics. Tests show that the model established in this paper can effectively predict annual corn production, with an average error rate below 5%.

Keywords

prediction and early warning; statistical yearbook; prediction model; time series

地方经济预警预测机制建立的可行性研究——以长治地区玉米年产量为例

刘惠娟

长治市统计监测中心, 中国·山西 长治 046000

摘要

新时代统计工作的深化改革应锚定服务于中国式现代化经济的高质量发展, 统计系统的超前研判可以帮助地方政府在保持经济快速发展的同时, 规避潜在的风险隐患。本文以1997—2022年《长治统计年鉴》为数据来源, 采用WPS Excel这一常用工作软件作为数据分析工具, 建立了长治地区玉米年生产量的柯布-道格拉斯预测模型, 并深入探讨地方经济统计预警预测工作长效机制建立的可行性。测试表明, 本文建立的模型可以有效预测玉米年产量, 平均差错率低于5%。

关键词

预测预警; 统计年鉴; 预测模型; 时间序列

1 引言

党的二十届三中全会对进一步全面深化改革、推进中国式现代化做出了科学擘画和战略部署。政府决策往往希望经济快速发展的同时, 可能避免潜在的风险, 这就要求统计系统做到信息的超前预判。利用统计工作的优势, 建立统计预测预警机制, 全力化解经济社会发展中可能遇到的问题, 为及时布置、防患于未然奠定基础^[1]。大数据背景下, 年鉴资料作为统计工作者最为方便快捷, 翔实可靠的资料性用书, 信息量巨大, 本文以长治地区1997—2022年《长治统计年鉴》为数据来源, 利用WPS Excel为数据分析工具,

通过对影响玉米生产的生产资料和投入情况做分析, 并拟合出符合玉米生产的C-D函数模型, 为扩大长治地区玉米生产、提高农业总产值具有指导意义, 并将探讨如何利用地方统计数据, 通过科学建模的手段, 达到对经济运行预警预测的目的, 为地方预警预测提供一条新思路, 讨论现阶段地方经济预警预测机制建立的可行性。

2 案例选择和建模过程

2.1 相关数据整理

玉米的投入产出数据来自《长治市统计年鉴》^[2](1997—2022), 本研究假定玉米产量是土地(播种面积)、投入因素(农业劳动力人口、施肥量)的函数。生产因素主要有土地; 投入因素主要有劳动力、肥料。忽略气候、管理因素以及玉米品种改良等因素。

【作者简介】刘惠娟(1982—), 女, 中国山西长治人, 硕士, 副高级统计师, 从事财务统计学研究。

2.2 确定时间序列的成分

1997—2022年间，长治地区玉米总产量呈增长态势；播种面积自2003年起开始稳步增加，到2011年后基本稳定在200千公顷以上略有增长；农业劳动力人口虽然波动较大，但仍保持增长的大趋势；施肥量在1997~2012年持续增长，2012年施肥量达379108吨，后逐年下降。总之，玉米的年产量及其影响因素均为带有趋势性的时间序列。

2.3 数据处理及预测模型的建立

C-D函数与回归函数间是对数与反对数的关系，使用WPS Excel对数据进行分析。将欲建立的玉米产量函数 $Y_{in} = K \times X_{in}$ ，对弹性系数 n 进行对数转换，即 $X'_i = \log X_i$ ，可得回归方程：

$$Y' = \log Y = \log K + \sum_{n=1}^3 \log X_n$$

其中， $i=1,2,3,\dots,20$ 代表1997—2022年的各年的观测次序； Y_i 表示长治地区 i 年的玉米总产量； X_{1i} 表示长治地区 i 年的玉米播种面积； X_{2i} 表示长治地区 i 年的农业劳动力人口； X_{3i} 表示长治地区 i 年的施肥量。 n 为1、2、3分别表示玉米播种面积、农业劳动力人口、施肥量对玉米总产量弹性系数， K 为转化效率系数。

用WPS Excel对影响玉米生长的三个主要因素玉米播种面积(X_1)、农业劳动力人口(X_2)、施肥量(X_3)；进行对数转换后，再进行多元回归分析，可得多元回归方程为：

$$Y' = \log Y = 0.546881846X'_1 + 0.967464911X'_2 + 1.004120268X'_3 - 6.332758726$$

公式1

以上回归方程显著，将公式1进行反对数转换，可以转换为C-D函数。

$$Y = 4.6477341X_1^{0.546881846}X_2^{0.967464911}X_3^{1.004120268}$$

公式2

2.4 模型精度的评价

将《年鉴》资料中的实测值与模型拟合的回测值进行比较，计算实测值与回测值之间的平均误差为：-7274吨，与实测值平均数1138564.91吨之间的差错率为-0.064%，可见本模型的拟合精度总体良好。由于平均差的计算简便，受极端数据影响较大，故差错率区间适当放宽至±5%之间。

3 结论与讨论

3.1 长治市玉米产量预测方程的分析与讨论

3.1.1 模型分析结论

通过对1997—2022年玉米产量与各因素间的回归方程分析可知，通过了显著水平等于0.001的F检验，说明该模型可信度较高。修正样本可决系数 $R^2=0.71$ ，表明玉米播种面积(X_1)、农业劳动力人口(X_2)、施肥量(X_3)对玉米总产量(Y)增长的解释能力为70.7%。该解释能力表明了该模型在70.7%的程度上拟合了样本数据。

3.1.2 影响因素分析

玉米播种面积、农业劳动力人口、施肥量的偏回归系数均大于0，表明三者与玉米产品存在显著的正相关。近年来，长治市农用耕地面积增长已临近极限，农业人口的保持也面临挑战。据陈建梅等研究，我国粮食增长在2005年以前主要靠化肥的投入获得，增加化肥投入量对粮食作物增长的作用程度高达139.745%^[3]。而2005年以后，玉米增产对化肥的依赖性逐渐减小。从长治地区的施肥量调查来看，2012年前，单位面积化肥投入量逐年增加而产量却在增长的总态势下出现上下波动，因此，指导农民合理施肥，提高化肥利用率，是农业生产中亟待解决的问题。

3.2 建模过程中遇到的问题

3.2.1 确定所预测变量的时间序列成分

本例是分析预测全市的玉米总产量。1997—2022年间，随着优良品种的开发、先进农业机械的应用，玉米种植的规模化、管理的科学化等；农田水利设施的建设，玉米财政补贴，农业税费改革政策等国家农业政策的实施^[4]，使得玉米产量应呈现一个上升趋势，这是符合现实规律的。

3.2.2 确定主要的影响因子

自变量一般选择与因变量明显相关的指标作为自变量，且控制在3个左右。本例中，《长治统计年鉴》收录的农业统计指标非常多，摘出1997—2022年的玉米总产量数据资料为因变量(Y)，再选择最主要的影响因素玉米播种面积(X_1)、农业劳动力人口(X_2)、施肥量(X_3)为自变量。来分析生产资料和投入对玉米总产量的影响。其他指标如农用机械总量、水电站个数等等对玉米生产也有一定的影响，但不是最直接最主要的影响因素，按照随机因素处理，不做过多讨论。另外，即便是主要因素之间，也有可能存在多重共线性关系。本例中，用t检验来检验回归系数，根据WPS Excel的分析结果， $t_1=3.129$ ， $t_2=1.735$ ， $t_3=3.624$ ，相对应的P值为0.0049、0.0968、0.0015，给定显著性水平 $\alpha=0.05$ ，可知 t_2 相应的 $P>0.05$ ，由此可知，劳动人口对玉米产量的影响不显著。回归方程中 $F=21.17$ ，相应的P值为1.1057。可见，回归方程通过F检验而回归系数未通过t检验，暗示着模型中有可能存在多重共线性的关系，需要进一步判断，或者采用向前选择、向后选择、逐步回归等方法来剔除多重共线性关系的影响，本文不再赘述。

3.2.3 预测模型的选择

玉米总产量是一个多因素影响的趋势性时间序列，指数曲线较为适用于此类序列的预测。柯布-道格拉斯函数是分析资源“投入”和产品“产出”间经济数量关系最常用的一种生产函数，也称为C-D生产函数。该函数将生产规模报酬划分为递增、递减和不变三种类型，成为经济学中技术进步测算和政策制定的理论基础，使经济增长理论从定性描述转向数理建模，开创了定量分析生产要素的先河。C-D函数比较广泛地适用于农业技术经济中，也可应用于政府制定

产业政策时,预测不同要素投入策略的经济效果。

3.2.4 预测模型的有效性分析

预测是对未来的估计和认识,不可能与未来实际完全相符,所以有误差是不可避免的,如果发现预测误差超出了容许的范围,那就应查明原因,必要时需要对模型进行修正^[5]。本例中回归方程的总方差为0.0672。它蕴含了有关模型基本假设的重要信息。如果回归模型正确的话,可以将残差看作误差的观测值。本文采用了最简单的平均误差来作为评价标准,主要是考虑到便于计算和理解。

3.2.5 预测模型的使用实践分析

利用以上C-D函数模型预测出2023年长治地区玉米总产量误差为36873.783吨,差错率为3.3%。分析原因主要是由于农业机械化和无人机广泛应用于农业生产,农业劳动力人口对作物产量的影响越来越小,2022年统计制度改革,农业劳动力人口不再作为农业统计指标进行统计,变量数据缺失对预测结果造成一定影响。这一实践结果也表明,变量与产量的相关性降低时,可以考虑将变量农业劳动力人口更改为其他相关性较强的指标,或考虑使用其他预测模型。

4 关于建立地方性经济预测预警长效机制的讨论

4.1 选择权威数据来源

我国经济社会指标数据涵盖广泛,发布机构多样,政府统计部门作为国家经济运行的专职记录和服务部门,是最具权威性,数据产出量最大,可开发性最强。作为大数据时代的统计人员,在做基层经济预警预测时,最方便快捷,最全面可靠的数据是地方统计年鉴或其他统计数据资料,例如《长治市统计年鉴》中所收录的数据资料分门别类、内容丰富。利用政府统计资料为数据来源,为地方经济增长做全方位、多角度的预测,具有广阔前景。

4.2 保证数据质量

数据质量是统计工作的生命线,如果源头数据误差太大,后面的一切工作都将无从谈起。提高基层统计数据质量,是一切统计工作的基础。我国统计数据质量保障体系已较为完善,通过“法律+技术+监督”的方式多管齐下,保障数据真实可靠。首先,统计法等一系列配套法律体系的建立,保证了统计数据的质量。其次,随着数字技术的发展,政府统计工作应用区块链技术,确保政府统计数据安全可靠、详实有效^[6]。

4.3 提高分析能力

如果说数据质量是统计的生命,那么统计分析就是统计的灵魂。大数据时代,大量的信息充斥着人类的生产生活,信息服务网格化、精细化、智能化要求越来越高。首先,提高统计队伍的业务水平。为提高基层统计分析能力做充分的人才储备。其次,借助分析工具的计算能力可减少基层工作者的工作量,为基层经济的预警预测提供了便捷。最后,多

样的培训和各种统计组织繁荣的大环境也有利于统计分析能力的提高。

4.4 可预测性的判定

地方经济指标的可预测性取决于数据本身的特性、可用信息的完整性和所选模型的适用性。统计年鉴资料已提供了大量带有趋势性和季节性的数据,在可用信息的完整性方面也表现优秀,选用适当的统计模型进行进一步开发,即可达到统计预警预测的目的。预警预测模型是否一劳永逸的适用本地区经济运行,需要对预测模型的有效性进行长期观察和测试。

4.5 其他考虑因素

预测模型的使用除考虑所预测指标本身特性,需分解预测误差、控制过度拟合风险以外,还需要考虑时间和费用成本、是否能够更好地解释问题本质、是否便于预测和优化维护、是否便于计算机处理等一系列问题。

5 结语

党的十八大以来,习近平总书记多次就统计工作做出重要讲话指示批示,明确了统计改革方向和根本遵循。党的十九届五中全会明确提出推进统计现代化改革,广大统计工作者要积极投身于统计现代化改革洪流,深化统计分析服务、有效发挥统计监督作用、完善统计体制机制。国家统计局在关于更加有效发挥统计监督职能作用的意见中明确指出,更加注重统计监测的客观性、预警性、有效性、结合政策要求、区域特点、数据规律,深度监测评价政策措施落实情况 and 实际效果,揭示问题风险,提出政策建议。大力推动大数据在政府统计工作中的应用,加强对非传统数据在经济决策中作用的研究,更好地支撑宏观经济决策。大力开展统计预测预警,充分运用统计数据信息,是加强统计工作的需要,是提高新质生产力的需要,是现代化经济建设的需要。统计预测有完整的理论体系,有独立的方法,统计预测分析应作为一个独立的工作阶段加以强化。抓好统计预测,一定能使整个统计工作更有生气,发挥更大的作用。

参考文献

- [1] 李勃.建立统计预测预警机制的探索与思考.中国信息报(N).2014,11,1第007版.
- [2] 长治市统计局,国家统计局长治调查队.长治统计年鉴[M](1997~2022).北京:中国统计出版社.长治市统计局,国家统计局长治调查队.
- [3] 陈建梅.农业生产资料投入对粮食作物产出影响因素的相关验证分析.经济研究导刊[J].2009(23):32~34.
- [4] 王资.农业政策对中国玉米生产的影响研究(D).浙江:浙江大学,2016.5:9~17.
- [5] 杨荫洲.统计预测方法(一).统计研究[J].1986,6:44~56.
- [6] 郭蓬丽,马美玲.区块链下的政府统计数据溯源机制初探.中国统计[J].2021.01:76~78.