

基于神经网络的宏观农业生产预测模型的研究

褚庆全 李立军 张志鹏

(中国农业大学 农学与生物技术学院,北京 100094)

摘要 为探索宏观农业生产系统预测的新方法,构建了基于人工神经网络的预测模型,利用1994—2003年的气象、经济、生产、投入、技术、价格各方面的数据对我国粮食生产进行了拟合分析,并预测了2004年粮食总产,预测的结果为46125.46万t。结果表明,与灰色系统相比,本文建立的模型具有90%以上的拟合精度,模型具有容错能力、联想能力和学习能力,可以用来尝试解决农业生产系统预测问题。

关键词 人工神经网络;BP模型;农业生产系统;预测

中图分类号 S117

文章编号 1007-4333(2004)05-0001-05

文献标识码 A

Prediction model of agricultural production based on artificial neural network

Chu Qingquan, Li Lijun, Zhang Zhipeng

(College of Agronomy and Biotechnology, China Agricultural University, Beijing 100094, China)

Abstract To explore new prediction methods of macro agricultural production system, one method of modeling the prediction of agricultural production based on BP (back propagation) model is established. Fitness analysis of foodstuff production was made and the total production of 461.2546 million ton in 2004 was forecasted using all the data of meteorology, economy, production, inputs, technology and prices. Compared with gray system method, the model has over 90% of precise fitness, tolerance on errors, association and studying capacity. This model can be applied to solve issue of prediction of agricultural production system.

Key words artificial neural network; back propagation model; agricultural production system; forecast

农业生产系统是以区域自然资源和社会经济条件为基础的复杂系统,具有生物学特点和经济学特点,从播种到收获,要受到栽培技术措施、作物生育规律、土壤资源、气候资源、生物环境以及粮食价格、粮食进出口数量、物质投入等的影响。要使我国的农业生产系统保持稳定健康的发展,进行比较准确的预测以有效地宏观调控是非常重要的。我国的农业职能部门每年要做3次生产情况预测,分别在产前、产中、产后,其中以产前和产中最为重视,已能根据预测作出正确的调控措施决策,实现从宏观上控制粮食生产。农业生产的决策和预测是一个复杂的过程,决策者要充分考虑社会经济条件与自然条件,包括投入、人口、土地、自然条件、农产品的供求

和政策信息等,因此需要对各种综合数据信息进行加工处理,选用合适的模型和分析方法的系统,支持和辅助决策者解决各种决策问题或制定相关发展计划。

目前国际上对农业生产系统的预测常用方法是灰色系统法、统计动力学方法和组合预测法^[1]。近年来,我国一些学者采用多元回归、灰色系统、模糊数学、时间序列分析、组合预测等方法对宏观农业生产的情况进行了不同角度的探讨,取得了一些的进展^[2~4]。据文献报道,其预测误差通常为5%~10%。这些方法预测结果不够理想的主要原因是:第一,没有用系统科学的方法进行预测;第二,没有把作物生产系统看作一个复杂系统,而是只针对气

收稿日期:2004-04-26

基金项目:科技部十五攻关资助项目(2001BA513B3)

作者简介:褚庆全,讲师,在职博士研究生,主要从事农业生产系统研究。

象等少数因素或某一类重要因素进行研究^[5]。有些方法中所考虑的因子过于简单,没能综合考虑影响农业生产系统的众多因素,或者没能挖掘出这些因素间的关系,预测的结果很难有说服力,或者有些模式实际上很难付诸实践^[6]。

农业生产系统的预测均属于非可实验系统预测,与可实验系统的预测相比主要有以下不同之处:1)这类系统的预测是用历史数据进行的,使用历史数据,则会有时(事)过境迁的变化;2)这类系统的影响因素众多,而且其中多数因素是不能进行人为控制的,这为预测造成了很多困难^[7,8]。要做到准确的宏观农业生产预测是有一定难度的。这是因为在预测中许多不可控因素,因素众多且复杂,主要包括:生产技术因素,如种籽、化肥、农家肥、农机、作物耕作及栽培技术等;社会经济因素,如农业政策、价格、农业管理等;气象因素,如作物各个生长阶段的气温、日照、降水、风雷等;自然因素,如耕地数量、地形、地势、土壤、病虫害等;还有各种随机因素,因此用一般的预测方法难以全部涵盖。目前,用于宏观农业生产系统预测的成熟方法也不多。模糊理论和计算机技术的快速发展为农业生产系统的预测提供了可靠的工具和可能。人工神经网络方法已经用于农业生产的各个方面,如精确农业中的模糊控制、技术管理以及作物和杂草的识别,谷物、水果的分类和分级,病虫害的预测等方面^[9~11]。但人工神经网络用于宏观农业生产的预测和决策报道的却很少。本文尝试利用人工神经网络理论,综合考虑多种因素,如气候、经济技术、自然、价格和政策等因素来预测全国粮食总产量,建立用于宏观农业生产预测和决策的神经网络模型,并实现计算机软件。

1 人工神经网络的原理和方法

人工神经网络(artificial neural network)简称神经网络(ANN),是基于现代生物学研究人脑组织的成果基础上,用大量简单的处理单元广泛连接组成的复杂网络,来模拟人类大脑神经网络的学习、记忆、推理、归纳等功能。ANN是一个大规模自组织、自适应的非线性动力系统。它能较好地模拟人的形象思维,具有大规模并行协同处理能力,具有较强的容错能力、联想能力和学习能力,能根据外界环境的变化修改自己的行为^[13]。

反向传播 BP(back propagation)模型是一种监督训练多层神经网络,由 Rumelhart 和 McClelland

1985 年提出^[5]。它具有很强的自组织、自适应能力。网络一般有 3 层:输入层、隐含层和输出层(图 1, x_i 为输入因子, w_{ij} 为权重, y_k 为输出因子)。本文采用具有多输入节点和多输出节点的 3 层 BP 神经网络,含一层隐层。BP 模型拓扑结构见图 1。

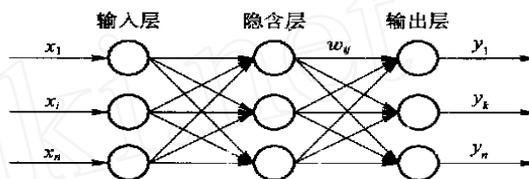


图 1 BP 网络模型

Fig. 1 BP network model

当有信息向网络输入时,信息首先由输入层传至隐层节点,经特性函数作用后,再传至输出层输出,其间每经过一层都要由相应的特性函数进行变换。节点的特性函数要求是可微的,通常选用 S 型函数,本文选用 $f(x) = 1/(1 + e^{-x})$ 作为特性函数。

其运行方式见图 2。信息流层进入网络后,传到隐含层结点,经过响应传递函数(一般取 Sigmoid 函数)的作用,再传到输出结点,计算模型输出值与理想输出值之间的误差,如不符合要求,则沿误差最大方向(误差梯度方向)反传回去,调整各层之间的权值,如此反复,当输出信号和教师信号间的误差之和 E 小于某一初始设定的精度时,学习过程即告结束,此时保存最终调整好的权值 w_{ij} ,并给出输出结果^[13]。

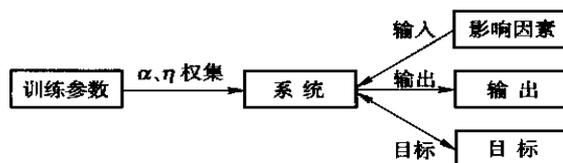


图 2 BP 网络模型的运行方式

Fig. 2 Operating way of BP network model

2 利用 ANN 模型对农业生产系统进行分析 and 预测

本文利用建立的模型对 1994—2003 年粮食生产影响因子的权重进行了分析,并对 2004 年的粮食生产情况进行了预测。

2.1 粮食产量影响因素的权重分析

影响粮食生产的因素很多,辨别哪些因素对产量的影响大,哪些因素的影响小,对于分析和调控粮食生产很重要。在进行预测前,首先利用 ANN 模

型分析影响因子对粮食产量的权重。选取的因素包括耕地面积、粮食作物播种面积、化肥施用量、有效灌溉面积、农机总动力、乡村劳动力、农村用电量、受灾面积、降雨不足年际变化率、良种百分率、粮食进口量、粮食出口量等因子。选取 1994—2003 年时间序列样本进行训练和学习。首先进行学习率的优化,得到优化后的学习率,取值为 0.85,即此时网络的学习过程更具快速性和平稳性,网络也不会出现“过学习”的现象。误差 = 0.001,得到最终训练结果的权重值(表 1)。

表 1 影响因子对粮食总产的权重

Table 1 Weights of total foodstuff production influence factors

因子	权重	因子	权重
耕地面积	8.0	受灾面积	4.2
粮食作物播种面积	5.3	降雨不足年际变化率	3.7
化肥施用量	6.9	良种百分率	2.5
有效灌溉面积	7.6	粮食进口量	0.4
农机总动力	1.5	粮食出口量	0.2
乡村劳动力	0.3	粮食收购价格指数	-1.8
农村用电量	0.8	农田建设支出	4.4

从全国总体情况来看,在 1994—2003 年期间,

耕地面积是影响粮食总产的重要因素之一,1997—2003 年的 6 年中,耕地面积大幅度减少,导致 1999 年之后的粮食总产连续 4 年滑坡。在物质投入因素方面,有效灌溉面积和化肥施用量的权重较大,分别为 7.6 和 6.9,说明在此期间灌溉面积的不断增加是粮食总产提高的最重要的因素,其次是化肥施用量的影响;农机动力和农村用电量对总产的影响相对较小。在气象因子中,受灾面积和降雨不足年际变化率的权重分别为 4.2 和 3.7,对粮食总产的影响也较大,这和生产实际情况相符合,2000—2003 年北方出现了较大范围的比较严重旱灾,2000 年粮食总产比 1999 年下降了 0.46 亿 t。在政策因素中,粮食收购价格指数的权重较小,说明它对粮食总产增加或减少的作用不很明显。而农田基建支出的贡献率在政策因素中较大。

2.2 ANN 模型的预测分析

以全国粮食总产量 (Y) 为输出神经元,以包括耕地面积、粮食作物播种面积、化肥施用量、有效灌溉面积、农机总动力、农村用电量、受灾面积、降雨不足年际变化率、农田建设支出等几个因子为输入神经元进行分析。网络训练回检结果与灰色系统回检结果相比较见表 2。

表 2 神经网络训练后回检结果与灰色系统回检结果比较

Table 2 Comparison of results tested past data between trained neural network and gray system

年份	粮食总产量/万 t	ANN 回检值/万 t	相对误差/ %	灰色系统回检值/万 t	相对误差/ %
1994	44 510.10	43 830.50	- 1.53	41 299.43	- 7.21
1995	46 661.80	44 976.56	- 3.61	48 692.34	4.35
1996	50 452.80	48 245.23	- 4.38	53 674.66	6.39
1997	49 417.70	51 890.56	5.00	53 143.79	7.54
1998	51 229.30	53 956.58	5.32	53 953.06	5.32
1999	50 838.60	51 962.45	2.21	48 989.07	- 3.64
2000	46 217.54	49 345.76	6.77	51 256.35	10.90
2001	45 263.80	47 537.58	5.02	49 125.66	8.53
2002	45 706.00	44 045.56	- 3.63	48 753.63	6.67
2003	43 067.00	45 202.56	4.96	46 046.56	6.92

注:资料来源于《中国统计年鉴》(1984—2002 年)

由表 2 可以看出,与灰色系统相比,经过训练的 ANN 模型与实际值拟合度较好,具有较高的预测精度,ANN 模型预测的最大误差 6.77%,平均误差为 4.24%;而灰色系统的预测的最大误差 10.90%,平均为 6.74%。将灰色系统和神经网络回检结果和预测结果与实际值相比较(图 3)可以看出,与灰色

系统相比,神经网络模型的回检结果与实际值比较接近。灰色系统预测是从系统的一个或几个离散数列中找出系统的变化关系,建立系统的连续变化模型。灰色系统所建立的是一个指数模型,当粮食总产发生零增长或负增长时,系统误差就会加大,预测周期越多误差越严重,并且灰色系统本身没有很强

的并行计算能力,这是灰色系统产生误差较大的主要原因。而神经网络模型能够对数据进行并行处理,它不像传统方法的建模过程,它是把大量的数据交给按一定结构形式和激励函数构建的人工神经网络。

粮食生产系统是一个非常复杂的系统,它的影响因素比较多,如气候因素、农业科技因素、社会经济因素等,一方面有些因素难以量化(如技术进步因素),另一方面影响因素多而不易穷尽,所以本文只是选取一些主要的、量化的影响因素,用这些因素来预测粮食产量,而不是全部因素,因此这也是预测误差产生的原因之一。利用神经网络模型进行预测时,预测的精度除了和选择的因素有关外,还与样本的训练次数、学习率以及所选择的误差值相关,因此,为了提高精度,减少误差,应加大样本的训练和学习次数,选择更小的值。

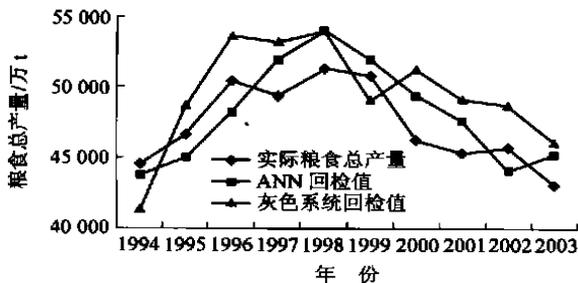


图3 ANN预测值、灰色系统预测值与实际值比较

Fig. 3 Comparison of ANN forecasting value, gray system forecasting value and actual value

由于训练过的网络已“模拟”并“记忆”了输入变量和输出变量之间的“函数”关系,因而可以用它来进行预测,并且由于是用上一年的预测因子作为输入变量,因此采用逐年加入新的信号、重新训练和学习、并向后预测的方法。用上述经过训练的神经网络对2004年我国粮食总产进行预测,预测结果为46 125.46万t。同样,利用上述方法,以1994年到2003年全国粮食总产(Y)为输入神经元,以经济作物播种面积、粮食作物播种面积、单位面积化肥用量等几个因子为输出神经元进行网络学习和训练,然后进行分析。计算后输出结果见表3。

从表3可以看出,到2004年,欲实现46 125.46万t的粮食产量,粮食作物和经济作物的播种面积分别为12 151.12万 hm^2 和2 315.16万 hm^2 ,农机总动力和化肥施用量分别为48 685.32万kW和4 152.25万t。

表3 ANN模型决策变量输出值

Table 3 Output value of ANN modal decision variable

年份	经济作物 播种面积/ 万 hm^2	粮食作物 播种面积/ 万 hm^2	农机 总动力/ 万kW	化肥 施用量/ 万t
2004	2 315.16	12 151.12	48 685.32	4 152.25

3 结论

从以上分析可以看出,选取适当的预测指标,在神经网络中经训练后,可以得到影响粮食总产的主要因子和粮食总产之间的非线性映射关系,回检精度比灰色系统预测精度高。这说明,可以利用人工神经网络模型进行宏观农业生产的预测和决策。同时,针对农业生产系统的预测,初步建立了基于ANN的预测模型,这种模型可以对目标变量的影响因子的权重进行分析,可以用于农业生产系统的预测和决策;并开发了相应的功能软件,只要给定时间历史序列的数据,软件即可进行分析预测。这是神经网络模型应用于农业生产系统预测的新尝试,如果把它与其他计算方法(如遗传算法)相结合,那么可望提高预测的精度。随着计算机技术在农业领域的不断渗透,神经网络必将在解决农业生产系统的非线性问题中的分类、预测等做出更大的贡献。

4 讨论

人工神经网络BP模型可以作为宏观农业生产系统的预测新尝试。和灰色系统相比,ANN模型可以对因子的贡献率进行分析,而且可以同时处理多个输出值的情况。但神经网络BP模型算法是一种数学方法,有其一定的局限性,如学习时间较长,收敛速度慢等,这些都有待于在实践中进一步探索,不断完善。另外,计算误差的大小,一方面取决于样本的训练和学习次数,还和预测因子的选择相关。因此,为了使预测结果更为可靠,还要对预测的指标进行选择,选择相关程度高的预测因子,这样才能是预测的结果更准确。

参 考 文 献

- [1] 崔振洋,李晓亮. 灰色GM(1,1)模型群及其粮食产量的预测[J]. 山西农业大学学报,1994,14(3):295~297
- [2] 肖冬根,黄璜,陈光,等. 农业专家系统中产量预测模型的研究[J]. 作物研究,2001(3):11~13

- [3] 王建林,王宪彬,太华杰. 中国粮食总产量预测方法研究[J]. 气象学报,2000(6):738~744
- [4] 吴春霞,何勇,蔡建平. 组合预测方法及其在粮食产量预测中的应用[J]. 农业系统科学与综合研究,2002(1):17~19
- [5] 陈锡康,潘晓明. 从农作物产量预测看发展交叉科学研究的重要性[J]. 中国科学院院刊,2000(1):47~48
- [6] 邓聚龙. 灰色预测与决策[M]. 武昌:华中理工大学出版社,1986. 20~30
- [7] 孙明玺. 预测和评价[M]. 杭州:浙江教育出版社,1986. 6~15
- [8] 郁明谏. 多维时空全息协调理论专集[M]. 北京:北京农业大学出版社,1995. 5~18
- [9] 蔡昆争. 神经网络在农业中的应用[J]. 农业系统科学与综合研究,2001(1):54~56
- [10] 谢名洋,黄永春. 基于神经网络的甘蔗产量预测系统[J]. 中国糖料,2000(1):14~19
- [11] 张志鹏. 以全息协调为基础的神经网络方法在农业生产中应用的研究[D]. 北京:中国农业大学,1999.
- [12] 胡守仁. 神经网络导论[M]. 长沙:国防科技大学出版社,1993. 20~35
- [13] 张志鹏,曹庆波,文新亚. 应用人工神经网络方法评定作物生产系统中指标的权重[J]. 农业系统科学与综合研究,2003(2):1~4

科技动态 ·

“中国粮食与食物安全研究中心”在我校成立

2004年7月18日,“中国粮食与食物安全研究中心”在我校成立。该中心聘任国家科技部中国生物技术发展中心王宏广教授为主任、我校农学与生物技术学院李召虎教授和食品科学与营养工程学院罗云波教授为副主任,聘任国际著名农学家、诺贝尔奖获得者布劳格博士及何康、陈锡文、袁隆平等22位著名科学家和有关领导为顾问,聘任全国27所院校和科研机构的51位专家教授作为首批研究员。

该中心将以保障我国粮食与食物安全为最终目标,联合国内外粮食、食物安全领域专家,组建研究队伍和协作网络进行研究,为政府提供决策建议、为粮食生产提供关键技术、为消费者提供科学服务。该中心的成立,将有利于进一步凝聚国内外优势科研力量,加快粮食与食物安全的科技创新步伐,为我国未来粮食与食物安全做出更大的贡献。

国家重大科技专项“粮食主产区保护性耕作制与关键技术研究”课题2004年8月在我校启动

该课题属于国家重大科技专项“粮食丰产科技工程”的第14课题,由我校农学与生物技术学院高旺盛教授主持。课题将按照“农艺技术为主,农机农艺配套;高产粮田为主,突出节本增效;技术集成为主,研究示范结合”的总体思路,在2004—2006年间,集中全国优势力量,重点突破不同农耕类型区周年保护性土壤耕作关键技术、覆盖作物与保护性种植方式、保护性耕作农田高产栽培及水肥高效利用等共性关键技术。该课题立足东北、华北及长江中下游三大平原,主攻水稻、玉米、小麦三大作物,以粮食产量占全国95%以上的11省市为依托,实现粮食高产、节本增效、资源环境保护三者同步发展的最终目标。

将由中国农业大学牵头,成立“全国保护性耕作制度研究协作网”(CCFSR),挂靠在中国农学会耕作制度分会,编辑出版《保护性耕作研究动态》,推动我国保护性耕作技术研究和学术交流。

(科学技术处供稿)