

# 基于遗传算法的人工神经网络模型在冬小麦根系分布预报中的应用\*

罗长寿 左 强\*\* 李保国

(中国农业大学土壤和水科学系, 北京 100094)

**Application of artificial neural network based on the genetic algorithm in predicting the root distribution of winter wheat.** LUO Changshou, ZUO Qiang, LI Baoguo (*Department of Soil and Water Science, China Agricultural University, Beijing 100094, China*). - *Chin. J. Appl. Ecol.*, 2004, 15(2): 354~356.

In this study, a controlled experiment of winter wheat under water stress at the seedling stage was conducted in soil columns in greenhouse. Based on the data gotten from the experiment, a model to estimate root length density distribution was developed through optimizing the weights of neural network by genetic algorithm. The neural network model was constructed by using forward neural network framework, by applying the strategy of the roulette wheel selection and reserving the most optimizing series of weights, which were composed by real codes. This model was applied to predict the root length density distribution of winter wheat, and the predicted root length density had good agreement with experiment data. The way could save a lot of manpower and material resources for determining the root length density distribution of winter wheat.

**Key words** Artificial neural network, Genetic algorithm, Water stress, Root length density distribution, Prediction.

文章编号 1001-9332(2004)02-0354-03 中图分类号 S126 文献标识码 A

## 1 引 言

我国水资源紧缺,农业用水为全国用水大户.作物是农田节水的重要环节.作物与水分的关系,历来是许多学科致力研究的热点,其中必不可少的是作物根系吸水特性的研究.在根系吸水模型中大都包含根系分布函数.在田间采集根系,得到较为准确的根系分布规律较为困难,常用的方法是利用已有的资料建立经验性的模型或对实测的数据进行拟合<sup>[5]</sup>.然而,根系在土壤剖面中的分布并无标准的模式,会随作物和土壤环境等众多的因素产生极大的变化,因此用前述方法得到的根系数据往往与实际值间差距较大.植物的根系与冠层相互协调、相互适应,二者之间有着紧密的联系.众多的文献研究了不同条件下根、冠间干物质的分配关系<sup>[4]</sup>.土壤水分对根系分布的影响普遍存在<sup>[8]</sup>.据此本研究拟基于土壤水分和冠部参数,对根系分布进行定量预报,即通过间接的、相对容易的途径获得水分胁迫条件下相对准确的根系参数.

人工神经网络是一种模拟人脑信息处理方法的非线性系统,比较适合于一些信息复杂、知识背景不清楚和推理规则不明确问题的建模.多层神经网络的BP模型是最重要的人工神经网络模型之一,在农业等领域有着广泛的应用<sup>[3,6]</sup>.然而BP算法是一种梯度下降搜索方法,存在容易陷入误差函数的局部极小等不足之处.遗传算法是模拟自然界生物进化过程的计算模型,是一种全局优化搜索算法.本文应用遗传算法与人工神经网络模型相结合的方法,对水分胁迫

下的根系分布函数进行了定量预报.

## 2 材料与方 法

### 2.1 试验概况

供试小麦品种为农大189.土壤为砂土,粒径1~0.25 mm的占55.61%,0.25~0.2 mm为20.72%,0.2~0.1 mm为20.77%,0.1~0.02 mm为0.34%,0.02~0.002 mm为0.12%;装填容重 $1.64 \text{ g} \cdot \text{cm}^{-3}$ ,田间持水量 $0.07 \text{ cm}^3 \cdot \text{cm}^{-3}$ .

土柱为内径10 cm的灰色聚乙烯管.按 $1.64 \text{ g} \cdot \text{cm}^{-3}$ 的容重每5 cm 1层分层,装填砂土,深度为40 cm,分别于2、5、10、15、25和35 cm处处理设负压计.

试验在小麦苗期进行,设置1个对照(水分充足, w)、2个处理(水分胁迫1, w<sub>1</sub>;水分胁迫2, w<sub>2</sub>),3次重复.试验过程中,使各处理得到充足养分.

4月27日播种,每个土柱内定株4棵,相当于大田常规密度 $4.50 \times 10^6 \text{ 株} \cdot \text{hm}^{-2}$ ,5月3日出苗.出苗后砂土表面覆盖3 cm的石英砂,防止裸间蒸发.5月10日开始处理,对照每3 d浇1次营养液,保证土壤水分充分供应,称重法控制水分含量;水分胁迫1中,土柱每6 d浇1次水,浇水量为对照的70%;水分胁迫2中,土柱每6 d浇1次水,浇水量为对

\* 国家重点基础研究发展规划项目(G1999011700)及国家863计划资助项目(2001AA242031).

\*\* 通讯联系人.

2002-04-22收稿,2002-12-10接受.

照的 50%。之后每 6 d 拆除土柱 1 次,获取相关的根系、冠部资料,至 6 月 8 日试验结束,共采样 5 次。

## 2.2 测定项目与方法

根长密度每 6 d 拆除土柱测定 1 次,将土柱纵向剖开,然后横向每 4 cm 一段截取带根的砂土,放于筛上,用水将根冲洗干净;采用德国 AGFA 公司生产的 Snapscan1236 型扫描仪及 Regent Instrument 公司研制的 WinRhizo pro 根系分析软件统计根长。

冠部的测定分析包括测定干物重与叶面积。叶面积测定采用 Snapscan1236 型扫描仪及 WinRhizo pro 根系分析软件中相应的功能来完成。土壤基质势用负压计测定,每天早、晚定时观测。蒸腾量用称重法测定,晚上观测土壤基质势后进行,每 3 d 测定 1 次。

## 2.3 预报根系分布的人工神经网络模型结构

人工神经网络采用前馈式网络<sup>[7]</sup>,包括输入层、输出层、中间层(隐含层)3 层,上下层之间各神经元实现连接,同层之间无连接。地上部干物重、叶面积与根系发育的关系密切<sup>[6]</sup>,这 2 个参数无疑应作为模型的输入项。作物发育时间基本上代表作物的遗传特性,土壤水分条件变化常常是导致根系生长、分布差异的主要原因,故作物发育时间(发芽后天数)与各深度处土壤的基质势也作为模型的输入项。相应各土层深度的根长密度作为输出项。故此,将人工神经网络模型设计为一个 9-9-10 结构的网络,即输入层的神经元数为 9,由发育时间、地上部干物重、叶面积和 2.5、10、15、25 和 35 cm 深度处的基质势构成;中间层神经元数设计为 9;输出层由土柱横向每 4 cm 1 段(共 10 段)的根长密度构成,神经元数为 10;网络共有 190 个权值(包括阈值在内)。

为了满足网络 Sigmoid 函数的条件,数据如果在(0, 1)之外,则作如下归一化处理:

$$x^* = (x - x_{\min}) / (x_{\max} - x_{\min}) \quad (1)$$

式中,  $x^*$  为处理后的数据,  $x$  为处理前的数据,  $x_{\min}$ 、 $x_{\max}$  分别为样本序列的最小值和最大值。

## 2.4 遗传算法优化人工神经网络模型的权值

本研究采用轮盘赌法,并实施最优个体保留策略,以保证收敛到全局最优解<sup>[11]</sup>;同时采用实数进行编码,以提高优化的质量和效率<sup>[9]</sup>。具体过程如下:1)由人工神经网络模型的网络结构可以获得待训网络的权值数,每一权值代表染色体上的一个基因,将各基因的排序固定下来,至此已定义了染色体的长度和结构;2)确定初始种群中染色体的数目  $N$ ,对每条染色体在权值范围内随机产生权值,构成初始群体  $P_{01}$ ;3)用训练样本对染色体种群中的个体所代表的神经网络进行训练,计算每个个体的学习误差( $E$ ):

$$E = \sum_{i=1}^n E_i, \text{ 其中 } E_i = \sum_{l=1}^m (y_l^i - C_l^i)^2 / 2 \quad (2)$$

式中,  $n$  为训练样本个数,  $m$  为输出单元个数,  $y_l^i - C_l^i$  表示用第  $i$  个样本训练时第  $l$  个输出的实际输出与期望输出的差。计算适应度函数值  $f_i = 1/E$ ,它保证误差越小,适应度值越大;4)依据各染色体适应度值,用轮盘赌法进行复制,得到群体  $P_{02}$ ;5)对群体  $P_{02}$  中染色体进行随机配对,并按给定的

需交换的基因数目在染色体上随机选择交换位置,之后依据交叉概率  $P_c$  进行交叉操作,得到群体  $P_{03}$ 。其中交叉操作算子如下<sup>[12]</sup>:

若父代个体  $X_1, X_2$  为区间  $V = [X_{\min}, X_{\max}]$  上均匀分布的随机数,  $Y_1, Y_2, Z_1, Z_2$  由式(3)决定,其中  $\alpha$  为整数, MOD 为取模运算符,则  $Z_1, Z_2$  为交叉后的子代个体。

$$Y_1 = \alpha X_1 + (1 - \alpha) X_2, Y_2 = \alpha X_2 + (1 - \alpha) X_1$$

$$Z_1 = \text{MOD}(Y_1, V), Z_2 = \text{MOD}(Y_2, V) \quad (3)$$

6)依照当前代数的变异概率  $P_m$  及变异量 var,对群体  $P_{03}$  染色体上的每位基因进行变异操作,形成新一代群体  $P_{10}$ 。其中变异率  $P_m$  定义为:

$$P_m = 0.01 + \text{NG} \cdot \text{cof} \quad (4)$$

式中,  $P_m$  表示当前代数的变异率, NG 为自上次进化以来为止连续未进化的代数, cof 为决定染色体强制变异(100% 变异)阈值的系数,通常取值较小。变异量 var 定义为<sup>[9]</sup>:

$$\text{var} = r \cdot q \cdot d \quad (5)$$

式中,  $r$  为随机产生的[0, 1]之间的随机数,  $q$  为权取值范围内的一个固定值,  $d$  为决定变异量 var 的动态参数,初始  $d = 1.0$ ,如果  $c > e$  则  $c = 0$  且  $d = 0.1d$ ;其中,  $c$  为计数器,统计自上次进化以来至当前代为止连续未进化的代数,  $e$  为常量,用于判定是否改变  $d$  的阈值。7)对新群体  $P_{10}$  中各染色体计算其学习误差及适应度值。若适应度值最大的染色体满足学习误差精度要求,则输出该个体,至此学习训练结束,否则转步骤 4)。

## 3 结果与分析

应用前述预报根系分布的神经网络模型结构以及基于实数编码的优化神经网络模型权值的遗传算法,建立了预报水分胁迫条件下根系分布的定量模型,其中初始种群中染色体数目  $N$  为 31,交叉概率  $P_c = 0.9$ ,  $\alpha = 8$ ,每个染色体上有 7 个基因位发生交换;  $\text{cof} = 0.01$ ,  $q = 1$ ,  $e = 10$ ,训练精度设计为学习误差  $E < 0.03$ 。

用 15 个样本中的 14 个样本数据进行学习。学习过程结束后,对未参与学习的实测根长密度分布进行预报。图 1 为模型对 5 月 21 日不同处理实测根长密度分布的预报结果,其中预报对照 w 的学习次数为 4.5 万次,水分胁迫  $w_1$  的学习次数为 5.7 万次,水分胁迫  $w_2$  的学习次数为 6.2 万次。图 1 表明,除局部点预报偏差较大外,整体预报效果较好,完全可以满足实际应用要求。

为进一步验证模型,对 6 月 2 日不同处理下实测根长密度分布进行了预报。图 2 为模型的预报结果,其中预报 6 月 2 日对照 w 的学习次数为 5.5 万次,水分胁迫  $w_1$  的学习次数为 4.7 万次,水分胁迫  $w_2$  的学习次数为 7.5 万次。图 2 同样可表明,整体对根长密度分布的预报效果较好。在预报过程中对不同的样本预报时,学习样本集是由不同的学习样本构成,使得在达到同一训练精度要求时需要不同的学习次数,这可能是由于不同的样本中所蕴涵的信息不同,进而模型学习、提炼其中的知识所需要的学习次数不同所致;模型总体

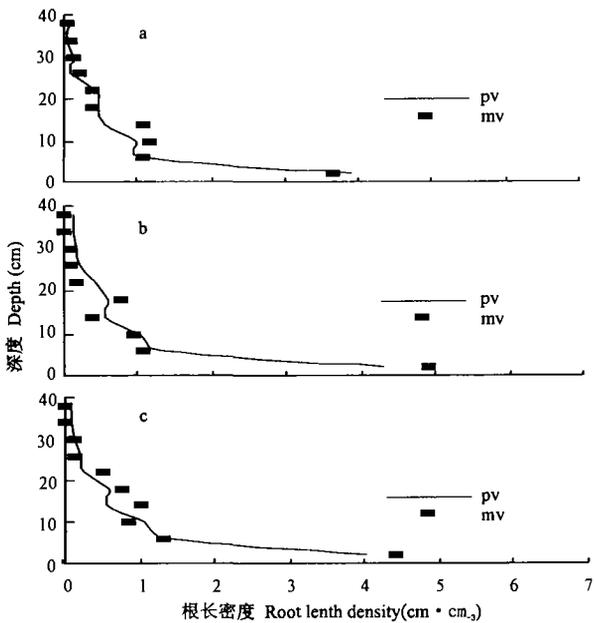


图1 5月21日神经网络根系分布预报模型预报与实测值的比较  
Fig.1 Comparison between data predicted by neural network model of root length density distribution and experimental data on May 21st. a)  $w$ , b)  $w_1$ , c)  $w_2$ , pv) 预报值 Predicted data, mv) 实测值 Measured data. 下同 The same below.

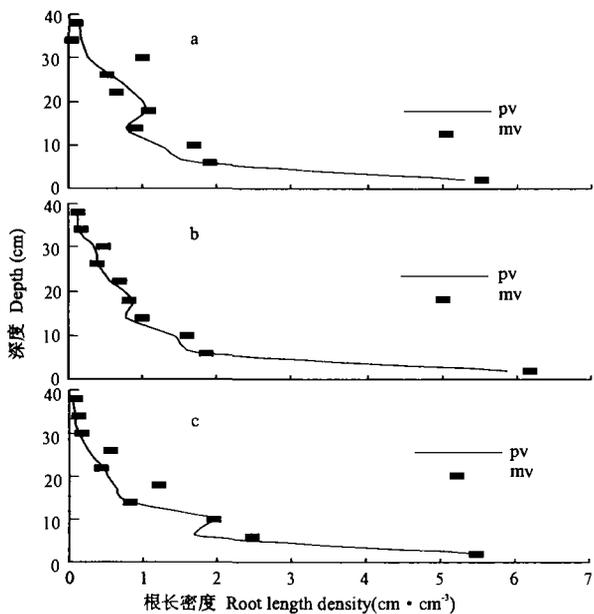


图2 6月2日神经网络根系分布预报模型预报与实测值的比较  
Fig.2 Comparison between data predicted by neural network model of root length density distribution and experimental data on June 2nd.

都达到了一个较好的预报效果,这表明本模型具有良好的收敛性与稳定性.可见,用遗传算法优化神经网络的权值,基于根系分布紧密联系并且易于获取的土壤水分和冠部参数,

建立预报根系分布的人工神经网络模型,对根系分布进行定量预报是可行的.

#### 4 结 语

基于冬小麦苗期室内的土柱实验,用遗传算法训练神经网络的权值,探讨了水分胁迫下应用人工神经网络模型对根系分布的定量预报.结果表明,模型的全局寻优及容错能力强,利用已有的根系分布资料及相应的冠部、土壤水分数据,通过模型学习后,可以对冬小麦的根系分布进行较为准确的预报.

应用该方法代替根系的测定,可节省大量的人力、物力;由此获得的根长密度分布可应用于根系吸水模型中,为农田灌溉及节水提供重要的决策依据.

由于资料的限制,关于苗期以后的情况,有待于进一步研究;随小麦发育期的延长,可以考虑再加入一些输入项,如株高及单株分蘖数等.

#### 参考文献

- 1 Chen G-L(陈国良), Wang X-F(王煦法), Zhuang Z-Q(庄镇泉). 1996. Genetic Algorithm and Application. Beijing: Communication Press. 88~97(in Chinese)
- 2 Huang X-F(黄晓峰), Pan L-D(潘立登), Chen B-H(陈标华), et al. 1999. Estimating reaction kinetics parameters with an improved real-coded genetic algorithm. *Chem Engin Chin Univ(高校化学工程学报)*, 13(1): 50~55
- 3 Koekkoek EJW, Booltink H. 1999. Neural network model to predict soil water retention. *Eur Soil Sci*, 50(3): 489~495
- 4 Thornley JHM. 1998. Modeling shoot: root relations: the only way forward? *Ann Bot*, 81: 165~171
- 5 Wu J, Zhang R, Gui S. 1999. Modeling soil water movement with water uptake by roots. *Plant Soil*, 215: 7~17
- 6 Yang P-L(杨培岭), Feng B(冯斌), Ren S-M(任树梅). 2000. Predicting the growing parameters of crop root under different water conditions by using artificial neural network model. *Trans CSAE(农业工程学报)*, 16(2): 46~49(in Chinese)
- 7 Yuan C-R(袁曾任). 1999. Artificial Neural Network and Application. Beijing: Tinghua Press. 66~70(in Chinese)
- 8 Zhang HP, Liu XN, Zhang HP, et al. 1993. Root system development of winter wheat and its relation to nitrogen, phosphorus and soil moisture in North Plain of China. *Acta Agric Boreali Sin*, 8: 76~82
- 9 Zheng Z-J(郑志军), Zheng S-Q(郑守淇). 2000. Evolving neural network by an adaptive real-coded genetic algorithm. *Com Engin Appl(计算机工程与应用)*, 9: 36~37(in Chinese)

作者简介 罗长寿,男,1974年生,博士,主要从事资源环境信息技术,发表论文6篇. E-mail: luochangshou@etang.com