

冬小麦生长条件下改进遗传算法 在根系水盐运移模型中的应用研究

罗长寿^{1,2}, 左 强², 李保国², 王 东²

(1. 北京市农林科学院农业科技信息研究所, 北京 100089; 2 中国农业大学资源与环境学院, 北京 100094)

摘 要: 应用改进遗传算法, 优化人工神经网络模型的权值, 对盐分存在下的冬小麦根系分布进行定量预报, 将获得的根系分布参数与根系吸水模型以及水盐运移模型相结合, 进行了水分、盐分分布的数值模拟。结果表明, 应用改进遗传算法可以为根系吸水模型提供所需的根系参数, 并且可以较好地对土壤中水分、盐分的运移分布情况进行模拟; 该方法建模简单、实用, 模型对于土壤次生盐渍化的防治与微咸水的灌溉利用等具有参考价值。

关键词: 改进遗传算法; 根系分布; 水分; 盐分

中图分类号: S152.7; S153.6; S512.1

文献标识码: A

文章编号: 1002-6819(2005)11-0038-05

罗长寿, 左 强, 李保国, 等. 冬小麦生长条件下改进遗传算法在根系水盐运移模型中的应用研究[J]. 农业工程学报, 2005, 21(11): 38- 42

Luo Changshou, Zuo Qiang, Li Baoguo, et al. Simulating soil water and solute transport in a soil wheat system using a neural network model with an improved genetic algorithm [J]. Transactions of the CSAE, 2005, 21(11): 38- 42 (in Chinese with English abstract)

0 引言

我国水资源紧缺, 关于微咸水灌溉以及作物根系吸水条件下的土壤水盐分布情况, 一直是非常活跃的研究领域^[1]。在应用较为广泛的根系吸水模型中, 大都包含根系分布函数。常用的获取根系分布参数的方法是建立经验性的模型或对实测的数据进行拟合^[2]。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, 因此所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^[3]。本文应用改进遗传算法, 通过间接的、相对容易的途径, 对盐分存在下的根系分布进行定量预报, 通过间接的、相对容易的途径获得盐分存在条件下相对准确的根系分布参数, 并将之与水盐运移模型相结合, 研究根系吸水条件下的水盐运移情况, 这对于盐分存在条件下作物的生长发育、防止土壤次生盐渍化及微咸水的利用等具有重要意义。

1 根系吸水条件下水分、盐分运移试验设置

1.1 试验设置

设置冬小麦室内土柱试验, 供试品种为农大189, 土壤为砂土, 装填容重1.64 g/cm³, 田持0.07 (cm³/cm³)。土柱为内径10 cm 的灰色聚乙烯管。按1.64 g/cm³的容重每5 cm 一层分层装填砂土, 装填深度为40 cm, 分别

于5、10、15、25、35 cm 处理设负压计及盐分传感器。试验在小麦苗期进行, 设置1个对照处理(无盐分胁迫)、2个盐分处理(盐分处理1, 盐分处理2), 3次重复。对照处理的营养液不加NaCl; 盐分处理1中, 营养液中加入NaCl, NaCl浓度为3 g/L; 盐分处理2中, 营养液中加入NaCl, NaCl浓度为6 g/L。每个土柱内定株4棵, 相当于大田常规密度450万株/hm²。出苗后砂土表面覆盖3 cm 的石英砂, 防止裸间蒸发。各处理均保证水分、养分充分供应, 称重法控制水分含量。开始试验处理后, 每6 d 拆除土柱一次, 获取相关的根系、冠部资料。

1.2 根系吸水条件下水分、盐分运移的数学模型

在本试验中, 垂直一维非饱和流情况下, 包含根系吸水的水分运移定解问题为:

$$\begin{aligned} C(h) \frac{\partial h}{\partial t} &= \frac{\partial}{\partial z} \left(K(h) \frac{\partial h}{\partial z} \right) - \frac{\partial K(h)}{\partial z} - S(z, t) \\ h(z, 0) &= h_0(z), \quad 0 \leq z \leq l_z \\ \left[-K(h) \frac{\partial h}{\partial z} + K(h) \right]_{z=0} &= E(t), \quad t > 0 \\ h(l_z, t) &= h_1(t), \quad t > 0 \end{aligned} \quad (1)$$

式中 h ——土壤水基质势, cm; z ——空间坐标, cm, 向下为正; $C(h)$ ——容水度, cm⁻¹; $K(h)$ ——非饱和导水率, cm · d⁻¹; t ——时间坐标, d; $h_0(z)$, $h_1(t)$ ——分别表示已知函数(或离散点); l_z ——模拟区域垂直总深度, cm; $S(z, t)$ ——根系吸水速率, d⁻¹; $E(t)$ ——蒸发、灌水强度(蒸发时取“-”号, 灌水时取“+”号), cm · d⁻¹。

本试验研究中, 不考虑土壤不动水体作用及吸附作用, 忽略土壤温度势的影响, 垂直一维非饱和土壤盐分运移方程如下

收稿日期: 2005-04-07 修订日期: 2005-07-04

基金项目: “中国—以色列科学与战略研究开发专项资金合作研究项目”; “国家重点基础研究发展规划项目”(G1999011709)部分研究内容

作者简介: 罗长寿(1974-), 男, 山西怀仁人, 博士, 主要从事农业信息技术研究。北京市农林科学院农业科技信息研究所, 100089。

Email: luocs@agri.ac.cn



$$\frac{\partial C(z,t)}{\partial t} = \frac{\partial}{\partial z} \left(\Theta_{sh} \frac{\partial C}{\partial z} \right) - \frac{\partial (qC)}{\partial z}$$

$$C(z, 0) = C_0(z) \quad 0 \leq z \leq L_z, t = 0 \quad (2)$$

$$\left[-\Theta_{sh} \frac{\partial C}{\partial z} + qC \right]_{z=0} = q_0 C_0(t) \quad z = 0, t > 0$$

$$C(L_z, t) = C(t) \quad z = L_z, t > 0$$

式中 C ——土壤盐分浓度, $g \cdot L^{-1}$; D_{sh} ——水动力弥散系数, $cm^2 \cdot d^{-1}$; q ——非饱和达西流速, $cm \cdot d^{-1}$; $C_0(z)$ ——已知剖面浓度, $g \cdot L^{-1}$; $C(t)$ ——已知浓度, $g \cdot L^{-1}$; $C_0(t)$ ——灌溉水的浓度, $g \cdot L^{-1}$, 蒸发时为 0。

2 基于改进遗传算法的根系分布预报模型

作物存在条件下, 在水盐运移模型模拟过程中, 根系吸水源汇项的确定对于模拟土壤水分与盐分的分布准确与否起着关键的作用。但是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。本研究针对根系与冠层紧密联系、相互协调、相互适应的特性, 利用人工神经网络适合于信息复杂问题建模的特性, 构建预报根系分布参数的人工神经网络模型; 并利用遗传算法的简单易用、鲁棒性强、全局寻优的基本特性, 应用改进遗传算法, 对人工神经网络模型的权值进行优化, 建立预报根系分布参数的人工神经网络模型。建立的基于改进遗传算法的人工神经网络模型可提供根系分布参数, 进而可与根系吸水模型和水盐运移模型相结合, 对根系吸水条件下的水盐运移分布情况进行模拟分析。

2.1 预报根系分布参数的人工神经网络模型的构建

地上部干物重、叶面积与根系发育的关系密切, 这两个参数作为模型的输入项; 作物发育时间基本上代表作物的遗传特性, 土壤水分、盐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、各深度处土壤的基质势与盐分也作为模型的输入项。相应的各土层深度的根长密度作为输出项。在本研究中使用前馈式人工神经网络, 该网络包括三层^[4]: 输入层, 输出层, 中间层(隐含层), 上下层之间各神经元实现连接, 同层之间无连接。

2.2 遗传算法优化人工神经网络模型权值的过程

遗传算法是一种模拟大自然生物进化过程的计算模型, 它以其简单、鲁棒性强、全局寻找以及不受搜索空间限制性条件约束等特点而日益受到人们的关注。遗传算法模拟了自然选择和自然遗传过程中发生的繁殖、交配和突变现象, 将问题的求解表示成染色体, 构成初始染色体群体, 并将它们置于问题的环境中去, 根据适者生存的原则, 从中选择出适应环境的染色体, 对之进行复制、交叉和变异操作, 进而产生更加适应环境的新一代染色体群, 如此不断循环进化, 直至收敛到一个最适应环境的个体上, 求得问题的最优解^[5]。本文中遗传算法优化人工神经网络模型权值的主要过程如下: 应用实数编码策略对人工神经网络模型的权值进行染色体基因编码; 生成初始的染色体群体; 计算群体中个体适应度值; 进行遗传操作, 包括采用轮盘赌法进行

选择操作, 以及按照一定的规则进行交叉和变异操作; 染色体群体中适应度值的再次计算; 如果满足停止搜索判据, 叠代停止, 输出问题的最优解; 否则, 转向步骤。为保证最优个体不被遗传操作所破坏, 在遗传操作过程中实施最优个体保留策略。

2.3 模型中适应度函数的确定

适应度函数是用来评判染色体群体中个体的优劣程度的指标, 遗传算法利用适应度值这一信息来指导搜索方向, 我们采用以下形式的适应度函数:

用输入和输出构成的训练样本, 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 , 如

$$E = \sum_{i=1}^n E_i, \text{ 其中 } E_i = \sum_{l=1}^m (y_l^i - C_l^i)^2 / 2 \quad (3)$$

式中 n ——训练样本个数; m ——输出单元个数; $y_l^i - C_l^i$ ——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。适应度函数由下式确定

$$f_s = 1/E$$

计算适应度函数值, 适应度函数保证误差越小, 适应度值越大。

2.4 对遗传算子的改进

应用遗传算法来训练人工神经网络模型的权值时, 交叉算子和变异算子在优化过程中占据着极其重要的地位。由于简单遗传算法(SGA)在优化人工神经网络模型的权值时, 染色体有可能在搜索空间内收缩或发散, 导致在搜索的快速性、全局收敛性方面往往达不到预期的效果。通过对交叉算子和变异算子的改进^[6], 可减少遗传算法优化神经网络模型权值的时间, 并提高遗传算法优化神经网络模型的收敛性和稳定性。具体如下:

对于交叉算子:

若 x_1, x_2 为父代个体, 为区间 $V = [x_{min}, x_{max}]$ 上均匀分布的随机数, 交叉后的子代个体 z_1, z_2 由下式产生^[7],

$$y_1 = \alpha x_1 + (1 - \alpha)x_2, y_2 = \alpha x_2 + (1 - \alpha)x_1 \quad (4)$$

$$z_1 = MOD(y_1, V), z_2 = MOD(y_2, V)$$

其中 α ——整数; MOD ——取模运算符。

对于变异算子, 采用如下规则^[8]:

$$P_m = 0.001 + NG \cdot cof \quad (5)$$

式中 P_m ——当前代数的变异率; NG ——自上次进化以来为止连续未进化的代数; cof ——决定染色体强制变异(100% 变异) 阈值的系数。

$$var = rand \cdot w \cdot dyna \quad (6)$$

式中 var ——变异量; $rand$ ——随机产生的[0, 1]之间的随机数; w ——权的取值范围内的一个固定值; $dyna$ ——决定变异量 var 的动态参数, 初始 $dyna = 1.0$; 如果 $counter > nochange$, 则 $dyna = dyna \times 0.1$ 且 $counter = 0$; 其中, $counter$ 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; $nochange$ ——常量, 是判定是否改变 $dyna$ 的阈值。

通过上述交叉算子与变异算子的组合应用, 一方面

产生的子代个体在搜索空间内是均匀分布的, 同时又可形成一个自适应的变异机制, 故可使遗传算法在优化神经网络模型权值的全过程中始终保持着较高的效率, 其性能检验见文献[9]。采用上述改进后的遗传算法, 可使遗传算法优化神经网络模型的收敛性得到明显改善, 并且减少求解时间, 从而达到提高优化根系分布预报模型效率的目的。

2.5 模型参数的获取

在前述试验设置中, 按下述方法取得相关的模型参数:

冠部参数: 每 6 d 测定一次, 测定项目包括干物重、叶面积; 叶面积的测定采用德国 AGFA 公司生产的 Snapscan 1236 型扫描仪进行叶面积扫描, 之后用 Regent Instrument 公司研制的 WinRhizo pro 分析软件进行叶面积计算。

根长密度的测定: 每 6 d 拆除土柱一次, 将土柱纵向剖开, 然后横每 4 cm 一段截取带根的砂土, 放于筛上, 用水将根冲洗干净; 用 Snapscan 1236 型扫描仪扫描根系, 用 WinRhizo pro 根系分析软件统计根长。

土壤盐分的测定: 用 TYC-II 型土壤盐分传感器获取, 每天按时观测。

土壤基质势的测定: 由负压计测定, 每天定时观测。
蒸腾量: 称重法获取, 每 3 d 测定一次。

3 模型应用与分析

3.1 根系分布预报结果分析

依据前述对构建预报根系分布参数的人工神经网络模型描述以及本试验设置, 本文预报盐胁迫下冬小麦根系分布的模型为一个 13-9-10 结构的网络, 即: 输入层由发育时间、地上部干物重、叶面积和 5、10、15、25、35 cm 深度处的基质势与盐分值构成, 神经元数为 13。输出层由土柱横每 4 cm 一段的根长密度构成, 神经元数为 10。中间层神经元数的设计在理论上并没有一个明确的规定, 一般依据经验而定。当中间层神经元数较少时, 模型全局优化能力减弱; 当中间层神经元数增多时, 模型的泛化能力减弱, 即对未经学习的样本的预报能力减弱。在本试验中中间层神经元数定为 9。应用改进的遗传算法优化神经网络模型的权值, 其中在遗传操作过程中实施最优个体保留策略, 群体规模为 31, 训练精度设计为学习误差 $E < 0.05$ 。学习结束后, 对未参与学习的样本进行预报。

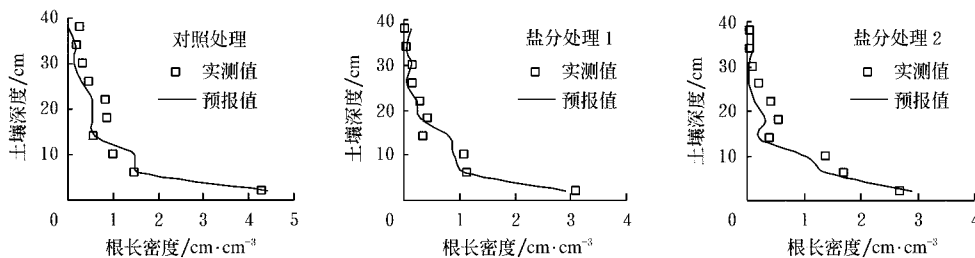


图 1 模型预报与实测结果比较

Fig. 1 Comparison between measured and estimated root length densities

图 1 是预报与实测结果的比较。图示表明, 除个别点外, 整体预报效果较好。同时, 在上述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, 但是模型总体都达到了一个较好的预报效果, 表明了模型具有良好的收敛性与稳定性。由图中可看出, 由于根系主要分布在土表以下 25cm 以内, 所以我们对此范围内的根长密度预报与实测的结果进行了平均相对误差分析及相关系数的计算, 其结果见表 1。在计算平均相对误差的过程中, 由于个别点的预报效果较差, 导致预报数据平均相对误差的增大; 但是其整体相关系数为 0.97, 表明模型总体的预报效果较好。同时, 由表 1 还可看出, 模型对于对照处理的预报效果要好于盐分处理 1, 而盐分处理 1 的预报效果又好于盐分处理 2, 这也从一个侧面反映了在盐分存在条件下, 作物根系生长出现更为复杂多变的局面。考虑到作物与环境的相互作用的复杂性与多变性, 认为本模型的预报结果是可以接受的, 可以满足实际应用的要求。也就是说, 基于与根系分布紧密联系并且易于获取的土壤水分、盐分和冠部参数, 用改进的遗传算法优化神经网络模型的权值, 对根系分

布进行定量预报是可行的, 可以用相对较易获取的参数取得较难测定的根系分布资料。

表 1 根长密度预报与实测结果误差分析

Table 1 Error analysis of results between measured and estimated root length densities

| 处 理 | 平均相对误差 | 相关系数 ($n = 18$) |
|--------|--------|-------------------|
| 对照处理 | 0.24 | 0.97 |
| 盐分处理 1 | 0.26 | |
| 盐分处理 2 | 0.27 | |

3.2 盐胁迫条件下的根系吸水模型的构建

有关盐胁迫条件(无水分胁迫)下根系吸水的规律, 应用较为广泛的当属 Feddes 等(1978)提出的根系吸水模型, 即

$$S = \alpha(h_0)S_{\max} \quad (7)$$

式中 S —— 根系吸水速率, 表示单位时间单位土体的根系吸水量, $\text{cm}^3 \cdot \text{cm}^{-3} \cdot \text{d}^{-1}$; h_0 —— 土壤水渗透势, cm ; S_{\max} —— 根系的 最大吸水速率, d^{-1} , 按下式计算^[2](Wu 等, 1999), 即:

$$S_{\max} = \frac{T_p(t)L_d(z,t)}{L_r(t) \int_0^{L_r(t)} L_d(z,t) dz} \quad (8)$$

式中 z ——土壤深度, cm; $T_p(t)$ ——潜在蒸腾强度, $\text{cm} \cdot \text{d}^{-1}$; $L_d(z,t)$ ——根长密度分布, $\text{cm} \cdot \text{cm}^{-3}$; $L_r(t)$ ——最大扎根深度, cm; $\alpha(h_0)$ ——无水分限制时渗透势的修正系数, 无量纲; 对于 $\alpha(h_0)$, 认为在无水分、养分胁迫条件下, 其较为实用的形式如下^[10] (Hom aee (1999)):

$$\alpha(h_0) = 1 - \frac{\alpha}{360}(h_0^* - h_0) \quad (9)$$

式中 h_0^* ——渗透势临界值, cm; α ——每增加单位电导率 ($\text{mS} \cdot \text{cm}^{-1}$) 降低吸水量的斜率值, cm^{-1} , $\alpha = 0.073$ 。由式(7)、(9) 可得盐分胁迫条件(无水分胁迫)下的根系吸水模型为

$$S = [1 - \frac{\alpha}{360}(h_0^* - h_0)] \frac{T_p L_d(z)}{L_r \int_0^{L_r} L_d(z) dz} \quad (10)$$

在上述根系吸水模型中, 根长密度分布由基于改进遗传算法的根系分布预报模型, 通过利用地上部干物重、叶面积、发育时间、基质势与盐分参数来提供。其中 h_0^* 应用实测值进行反求, 方法如下: 已知 t_1 、 t_2 两个时刻的实测土壤剖面含水量分布、盐分浓度分布、根长密度分布函数 $L_d(z)$ 及潜在蒸腾强度 T_p , 给定 h_0^* 一个初值, 从 t_1 时刻起, 按照式(10) 计算根系吸水速率, 作为源汇项求解水分、盐分运移方程, 计算出 t_2 时刻土壤含水量的模拟值, 获得该时刻土壤剖面含水量模拟值与实测值

之间的误差平方和; 然后再给 h_0^* 赋一个值, 用同样的方法计算该误差平方和, 直到误差平方和最小, 此时的 h_0^* 值即为所求参数。本试验中计算结果 $h_0^* = -560 \text{ cm}$ 。

3.3 水分、盐分运移模拟结果分析

用改进的遗传算法优化神经网络模型的权值, 对根系分布进行定量预报, 将预报所得的根系分布参数与根系吸水模型相结合, 进行水分、盐分运移分布模拟。在盐分胁迫下, 考虑根系吸水的水分运移方程采用 Crank-Nicolson 差分格式求解, 盐分求解方程采用 Bresler (1973) 二阶差分法求解上述定解问题, 对土壤水分、盐分进行数值模拟。模拟过程中, 空间步长 $\Delta z = 1 \text{ cm}$, 时间步长 $\Delta t_{j+1} = 1.25\Delta t_j$, 下标 $j, j+1$ 表示时间步长的序数, 迭代水势控制标准 $\epsilon = 0.01$, 上边界土面蒸发强度 $E(t) = -0.038 \text{ cm} \cdot \text{d}^{-1}$, 由于实测数据的限制, 地表处的含水率由线性外推给出, 下边界取在 40 cm 处, 水分通量、盐分通量均为 0, 最大扎根深度 $L_r = 40 \text{ cm}$ 。图 2 是其中一个处理土壤剖面的水分、盐分的运移模拟结果, 模拟时间为 3 d, 其中平均潜在蒸腾强度 T_p 为 $0.38 (\text{cm} \cdot \text{d}^{-1})$ 。图 2 的结果表明, 应用该方法可以较好地预报盐分胁迫下土壤中水分的分布情况, 同时, 对于盐分的分布情况, 除个别点外, 整体模拟效果较好。整体上, 作为一种方法上的尝试, 我们认为, 应用改进的遗传算法对根系分布进行定量预报, 将预报所得的结果与根系吸水模型相结合, 进行水盐运移分布模拟, 可以较好地模拟盐分胁迫、作物生长条件下土壤中水分、盐分的分布情况。

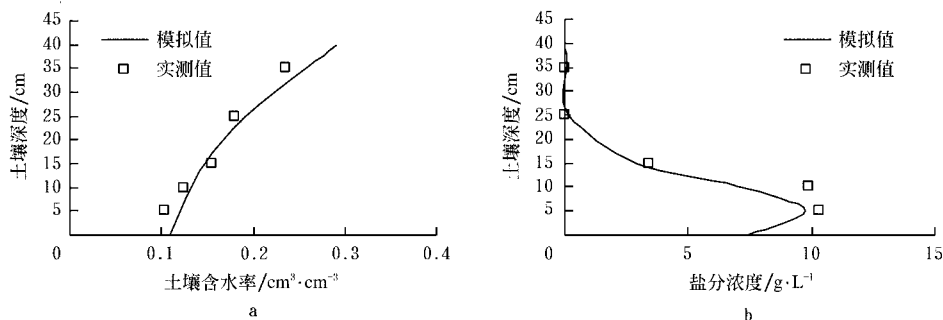


图 2 土壤含水量、盐分浓度模拟值与实测值的比较

Fig 2 Comparison between measured and estimated soil water content, salinity

4 结 语

作物根系吸水是土壤-植物-大气连续体水分传输问题研究中的一个重要部分, 同时又是根区土壤水分动态模拟必不可少的资料。通过应用改进的遗传算法, 为盐分存在下的根系分布参数获取提供一种相对较为简便、实用的方法。将该方法获得的根系参数与根系吸水模型及水盐运移模型相结合, 可以较好地模拟土壤中的水分、盐分的分布。该方法的应用对于研究盐分胁迫下的根系分布、盐分与作物生长的关系, 以及土壤水分、盐分运移分布规律等具有重要的作用, 并且该方法的应用可为微咸水的利用及防止土壤次生盐渍化等提供一

定的决策依据。

[参 考 文 献]

- [1] Khosla B K, Gupta R K. Response of wheat to saline irrigation and drainage [J]. Agricultural Water Management, 1997, 32: 285- 291.
- [2] Wu J R, Zhang S G. Modeling soil water movement with water uptake by roots[J]. Plant and Soil, 1999, 215: 7- 17.
- [3] Thornley J H M. Modelling Shoot: Root Relations: the Only Way Forward? [J]. Annal of Botany, 1998, 81: 165 - 171
- [4] 袁曾任. 人工神经网络及其应用[M]. 清华大学出版

- 社, 1999: 66- 70
- [5] 陈国良, 王煦法, 庄镇泉, 等. 遗传算法及其应用[M]. 北京: 人民邮电出版社, 1996
- [6] 罗长寿. 应用基于遗传算法的神经网络模型估算水盐胁迫下的冬小麦根系分布[D]. 中国农业大学, 2002
- [7] 黄晓峰, 潘立登, 陈标华, 等. 用改进的实数编码遗传算法估计反应动力学参数[J]. 高等化学工程学报, 1999, 13(1): 50- 55
- [8] 郑志军, 郑守淇. 用基于实数编码的自适应遗传算法进化神经网络[J]. 计算机工程与应用, 2000, (9): 36- 37.
- [9] 罗长寿, 周丽英. 改进遗传算法的神经网络模型研究[J]. 情报杂志, 2005, 24(5): 65- 66
- [10] Homae M. Root water uptake under non-uniform transient salinity and water stress [D]. Wageningen Agricultural University, the Netherlands 1999: 41- 54, 93- 97.

Simulating soil water and solute transport in a soil-wheat system using a neural network model with an improved genetic algorithm

Luo Changshou^{1,2}, Zuo Qiang², Li Baoguo², Wang Dong²

(1. Institute of Information on Science and Technology of Agriculture, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100089, China; 2. College of Resources and Environment, China Agricultural University, Beijing 100094, China)

Abstract: An improved genetic algorithm was applied and examined to optimize the weights of a neural network model for estimating root length density (RLD) distributions of winter wheat under salinity stress. Thereafter, soil water and solute transport with root-water-uptake in a soil-wheat system was simulated numerically, in which the estimated RLD distributions were incorporated. The results showed that the estimated RLD distributions of winter wheat using the neural network model combined with the improved genetic algorithm, as well as the simulated soil water content and salinity distributions, were comparably in agreement with the experimental data. The method can be used in modeling flow and transport under salinity or saline water irrigated areas.

Key word: improved genetic algorithm; root length density distribution; soil water content; salinity