

doi:10.6041/j.issn.1000-1298.2014.10.023

基于交叉验证的农田土壤饱和导水率传递函数研究^{*}

孙 美¹ 张晓琳¹ 冯绍元² 霍再林¹

(1. 中国农业大学水利与土木工程学院, 北京 100083; 2. 扬州大学水利科学与工程学院, 扬州 225009)

摘要: 基于北京大兴区农田土壤剖面采样数据,采用多元回归及多折交叉验证方法,建立并验证以不同土壤质地、有机质含量为自变量的农田土壤饱和导水率(K_s)土壤传递函数(PTF),对比了其他3种现有的 K_s 土壤传递函数预测结果,并对0~80 cm农田土壤 K_s 进行预测。结果表明:以土壤粉粒质量分数、粘粒质量分数、有机质质量比的组合项及常数项为自变量建立的多元回归方程,对所采剖面样本0~100 cm内土壤饱和导水率预测的均方根误差为40.525 cm/d,平均相对误差为204.738%,决定系数为0.544,相关系数为0.742,均好于其他3种预测结果。大兴区域农田土壤 K_s 表层变化范围较大,北部及西南部永定河一带 K_s 较小。所建立的PTF模型对采育镇60~80 cm深度处的粘壤土夹层也有较好表征,可以用作大兴区域土壤饱和导水率的评估。

关键词: 农田土壤 饱和导水率 土壤传递函数 多元回归 多折交叉验证

中图分类号: S152.3; S152.7 **文献标识码:** A **文章编号:** 1000-1298(2014)10-0147-06

引言

土壤水分运动是影响田间作物生长和地下水环境的重要影响因子,是农田水肥管理、区域水文循环及水资源转换等研究的基础。土壤饱和导水率(K_s)反映了不同环境条件对土壤水分入渗性能的影响,是描述土壤水分运动的最基本的量化参数之一,它与土壤质地、土壤容积密度、土壤有机质等土壤基本物理性质有着密切关系。大部分模拟土壤水分、溶质运移的农业系统模型均需要输入较精确的 K_s 值。由于传统方法对大面积直接测量获取土壤饱和导水率存在困难,因而根据土壤基本物理性质参数建立估计预测饱和导水率的传递函数(Pedotransfer function, PTF)的间接方法研究一直备受关注。多元回归法是建立土壤饱和导水率传递函数的简单且运用最广的方法,其他建立PTF的方法还有人工神经网络法(ANN)^[1-3]、数据分类分组方法^[4]、分类与回归树法等^[5-6]。国内外学者们分别针对不同地区^[2, 7-8]、不同土壤质地(如砂土^[9]、壤土和粘土^[10]等)或不同土壤类型(如稻田土^[11-12]、热带土^[13]等)建立了土壤饱和导水率PTFs。通过对土壤剖面的土壤水分特征曲线及土壤饱和导水率样本进行多元回归分析, Li等^[14]建立了封丘地区土壤水分特征参数PTFs模型,并用于玉米、小麦产量

的区域尺度预测。王相平^[15]建立了北京通州区土壤水分特征参数PTFs模型,与Li等^[14]建立的PTFs模型对比发现,通州地区模型对饱和导水率 K_s 的预测结果优于Li等^[14]建立的模型。李晓鹏等^[16]研究认为,土壤传递函数与同一地区试验数据统计资料的土壤适用性较高。正是由于PTFs本身具有较强的地域性,使得土壤传递函数在不同地区上的应用适用性具有局限性。北京市大兴区是华北地区较典型的集约化农业种植区,为北京重要的农副产品生产供应基地。大兴区土壤饱和导水率的分布特征直接关系到区域水肥利用效率及农作物生产等研究,从而影响农业面源污染负荷及地下水安全等问题。为更好地了解大兴区域土壤渗透性情况,本文通过分析北京市大兴区土壤质地、有机质含量的实测数据,采用多元回归方法建立大兴区农田土壤饱和导水率 K_s 传递函数模型,并通过与多折交叉验证方法相结合,在模型建立同时对模型进行验证,以优化建立模型的参数,提高模型精度及稳定性,同时与文献[15]、文献[14]、文献[17]3种方法进行比较,探讨PTF在研究区土壤饱和导水率预测的适应性。

1 材料与方法

1.1 研究区概况

试验区在北京市大兴区,全区总面积为1 036 km²,

收稿日期: 2013-11-26 修回日期: 2013-12-23

^{*} 公益性行业(农业)科研专项经费资助项目(201203077)

作者简介: 孙美, 博士生, 主要从事农田水循环研究, E-mail: sunmei-0451@163.com

通讯作者: 霍再林, 副教授, 主要从事农田水循环与水土环境、作物水盐响应及高效用水研究, E-mail: huozl@cau.edu.cn

地势自西向东南缓倾,海拔 13.4 ~ 52.0 m 之间,东经 116°13′ ~ 116°43′,北纬 39°26′ ~ 39°51′,属暖温带半湿润大陆季风气候,年平均气温为 11.6℃,年平均降水量 556.4 mm。全区土壤以褐土为主,占全区土壤面积 95% 以上;其次为潮土,主要分布于西南平原地区及少量北部地区。大兴区土壤质地类型以砂质、轻壤质和砂壤质为主,分别占总面积的 33.62%、31.36% 和 26.68%,其他质地类型总和不到 10%。

1.2 样品采集与测定

为了得到较准确的土壤属性数据,根据北京市 1:1 000 000 土壤类型图、北京市大兴区土地利用类型图(2005 年)及取样点分布均匀的原则,于 2010 年 7 月在大兴区 14 个乡镇的农田开挖 20 个土壤剖面,同时布置 94 个土钻点^[18]。每个土壤剖面取样深度为 200 cm,依据土壤剖面分层属性,分 5 ~ 6 层用环刀取原状土,共 84 个样本点,每点测定土壤含水率、饱和导水率、颗粒组成、土壤容积密度、有机质含量;每个土钻点取样深度 0 ~ 20、20 ~ 40、40 ~ 60、60 ~ 80 cm,每层测定土壤含水率、颗粒组成、有机质含量。大兴区行政区划图、20 个剖面点及 94 个土钻点分布如图 1 所示。

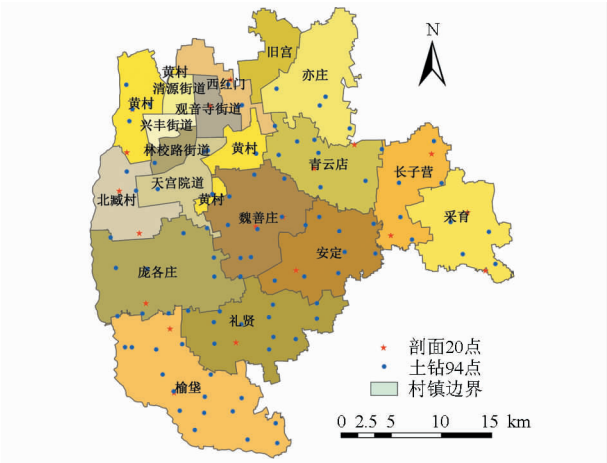


图 1 研究区采样点分布示意图

Fig. 1 Spatial distribution diagram of sample points in Daxing district

土壤含水率测定采用干燥法;土壤饱和导水率采用常水头法;粒径组成测定采用 Mastersizer 2000 A 型激光粒度分析仪;土壤容积密度测定采用环刀法;土壤有机质测定采用重铬酸钾容量法-稀释热法。

1.3 土壤饱和和导水率传递函数及多折交叉验证方法

在建立大兴区土壤饱和和导水率 PTF 之前,首先分析研究区域开挖 20 剖面中 0 ~ 100 cm 的 42 个土壤样品基本理化性状数据及其相应的数据变换形式

与土壤饱和和导水率 K_s 之间的相关关系,确定传递函数中的输入变量 $X(x_1, x_2, \dots, x_n)$ 。采用多元回归方法建立 K_s 与 X 之间的回归模型,为使最终获得的回归模型有较高的精度及稳定性,运用多折交叉验证的方法对模型参数逐步寻优。寻优过程主要分为两部分,一部分为模型可选择参数组合初选,以保证模型对局部数据模拟具有较高稳定性;另一部分为模型参数终选,以保证模型对整体数据模拟具有较高精度。具体步骤为:

(1)模型由于输入变量较多,由 SPSS 逐步回归法得到纳入回归方程中的自变量 X 。

(2)利用 Matlab 实现随机分组 1 000 次,每次分组 42 个样品以 9:1 的比例随机分成训练组及验证组。

(3)分别用训练组进行多元回归建模,用验证组验证模型对剩余数据的预测精度,如此可获得共 1 000 个土壤饱和和导水率的 PTF 模型,同时分别计算这 1 000 个模型训练组与验证组对 K_s 的均方根误差(式(1))及平均相对误差(式(2))(R_{Ti} 、 A_{Ti} 、 R_{Vi} 、 A_{Vi} ,其中 $i = 1, 2, \dots, 1\,000$),并计算训练组与验证组 R_{MSE} 及 A_{PE} 差值: R_{Di} 、 A_{Di} 。

(4)根据绘出的 R_D 、 A_D 的频率累积图(图 2)的分布规律,初步选取 R_D 、 A_D 两误差统计参数均较小的参数组(R_{Di} 、 A_{Di} 在两累积分布图同时小于 70% 的参数组),以作模型参数终选的备选组。

(5)对 42 个样品重新进行 1 000 次随机分组抽样成新的训练组和验证组,分别计算备选组模型对新训练组及验证组 1 000 次 R_{Di} 、 A_{Di} 的平均值 $\bar{R}_D$ 、 $\bar{A}_D$ 及对整体数据模拟的决定系数(R^2),选择几组模型中 $\bar{R}_D$ 、 $\bar{A}_D$ 及 R^2 较小的一组确定为最终模型,以保证所建立的模型对整体数据模型的准确性。

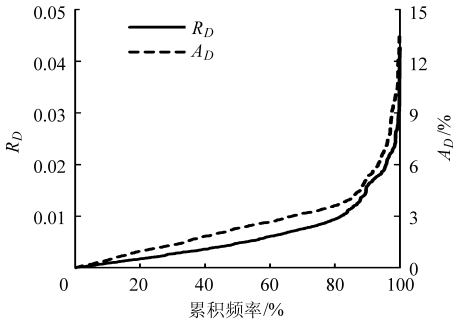


图 2 R_D 、 A_D 频率累积分布图

Fig. 2 Cumulative frequency distribution map for R_D and A_D

1.4 不同土壤饱和和导水率传递函数

本文选取文献[15]、文献[14]、文献[17]共 3 种表征土壤饱和和导水率的 PTF 与本文所建立的大兴区土壤传递函数做对比。其中,文献[15]、文献[14]分别根据北京通州区、河南封丘地区土壤特性,采用

多元回归方法建立 PTF;文献[17]为美国盐土实验室根据 1 913 个不同岩性土壤,采用神经网络技术建立的 PTF 程序软件。

1.5 模型评价指标

利用均方根误差 (R_{MSE})、平均相对误差 (A_{PE})、决定系数 (R^2) 和相关系数 4 个统计量评价各个模型模拟值和观测值的吻合程度。用标准差及变异系数 (C_v) 评价土壤基础物理性质的空间离散程度。

$$R_{MSE} = \sqrt{\frac{1}{n} \sum (O - P)^2}$$
 (1)

$$A_{PE} = \frac{1}{n} \sum |(O - P)/O| \times 100\%$$
 (2)

$$R^2 = 1 - \frac{\sum (O - P)^2}{\sum (O - \bar{O})^2}$$
 (3)

$$C_v = \left(\sqrt{\frac{1}{n} \sum (O - \bar{O})^2} / \bar{O} \right) \times 100\%$$
 (4)

式中 O ——土壤饱和和导水率的实测值
 $\bar{O}$ ——各实测值平均值
 P ——土壤饱和和导水率传递函数的预测值
 n ——样本个数

2 结果与分析

2.1 大兴区野外采样土壤物理特性统计特征

大兴区土壤的颗粒组成根据美国制土壤分类,主要是壤质砂土、砂质壤土、壤土、砂质粘土为主。其土壤的砂粒含量整体偏高。野外采样开挖剖面的 20 个采样点 0 ~ 100 cm 农田土壤饱和和导水率及土钻点 94 个采样点的土壤基本物理特性的统计特征如表 1 所示。可以看出,研究区域内土壤 K_s 及有机质质量比的取值范围较大,其中 K_s 的变化范围为 1.32 ~ 205.03 cm/d,变异系数 C_v 值最大,达到 139.0%。土壤有机质质量比的 C_v 值次之,达到 75.42%,最大值出现在土壤以粘壤土为主的东部采育地区,达到 22.07 g/kg,属土壤养分含量划分等级中高级。研究区土壤不同粒径含量的变化系数相差不大,均在 30% ~ 40% 之间,变化范围从大到小依次为砂粒、粉粒、粘粒含量。这主要是由于大兴区属永定河冲积平原,西部地区壤质砂土较多,东部采育地区多壤土或粘土,土壤颗粒组分不均造成;土壤容积密度的变异性最小, C_v 值仅为 8.7%,变化范围为 1.134 ~ 1.668 g/cm³。

表 1 采样点土壤饱和和导水率 K_s 及物理特性统计特性
Tab.1 Statistical characteristics of K_s and physical property for sample points

采样点	统计指标	砂粒质量 分数/%	粉粒质量 分数/%	粘粒质量 分数/%	有机质质量比/ (g·kg ⁻¹)	土壤容积密度/ (g·cm ⁻³)	土壤饱和和导水率/ (cm·d ⁻¹)
剖面 20 点	最大值	74.10	86.02	14.07	22.07	1.668	205.03
	最小值	2.17	22.74	2.56	0.90	1.134	1.32
	标准差	17.45	14.95	3.23	5.10	0.124	47.52
	平均值	43.63	49.31	7.06	6.76	1.421	34.18
	C_v /%	39.99	30.33	45.73	75.42	8.70	139.01
土钻 94 点	最大值	90.26	83.89	14.83	26.01		
	最小值	3.99	8.85	0.60	0.36		
	标准差	20.12	17.20	3.24	4.58		
	平均值	50.64	43.43	5.93	6.72		
	C_v /%	39.72	39.61	54.65	68.16		

2.2 土壤饱和和导水率 PTFs

2.2.1 大兴区土壤饱和和导水率 PTF 模型

本文在建立饱和和导水率的三次多元回归模型 (PTF) 时,涉及的变量有土壤砂粒质量分数 (S_a)、粉粒质量分数 (S_i)、粘粒质量分数 (C_l)、有机质质量比 (O_m)、容积密度 (B_d)。由于三次多项式涉及的全部可能自变量较多,本文仅给出与 K_s 在 $p < 0.05$ 水平下土壤基本物理性质等变量的相关分析结果,如表 2 所示。其中若以“*”代表上述某一变量,则*²、*³分别为相应自变量的二次幂、三次幂;“×”为对不同变量做乘积。可以看出,在 $p < 0.05$ 水平上 $S_a \times O_m$ 、 $S_a \times S_i \times O_m$ 、 O_m 等均与 K_s 存在较好的相关性;在 $p < 0.01$ 水平下, O_m^3 、

$C_l \times O_m^2$ 、 $O_m^2 \times B_d$ 、 $S_i \times O_m^2$ 与 K_s 的相关性系数最高,其中 $S_i \times O_m^2$ 与 K_s 呈明显负相关。此外, K_s 与土壤质地一阶变量之间没有明显的相关性,这与文献[14]的研究结果相似。以上相关分析的结果可以揭示各土壤基本性质变量与 K_s 之间的线性关系,为更准确、便捷地建立土壤饱和和导水率 PTF 模型提供参考。

按照 1.3 节中步骤(1),最终进入优化过程的模型自变量为 $S_i \times O_m^2$ 、 $C_l \times O_m^2$ 、 O_m^3 及常数项。经过模型初选阶段,共获得 4 组符合要求的模型参数组, $\bar{R}_D$ 、 $\bar{A}_D$ 及各参数组对全部 42 点数据模拟的 R^2 如表 3 所示。可以看出,第 301 组参数的 $\bar{R}_D$ 、 $\bar{A}_D$ 相对较小,且 $R^2 = 0.544$ 达到最大,相比其他组参数组合更

合理。若传递函数的变量中加入容积密度 B_d 输入项,回归后模型决定系数 R^2 为 0.556,提高幅度并不显著,且土壤容积密度的测定需要挖取土壤剖面,并伴有取样及测定工作,明显不利于模型在区域上的

推广。因此本文将较易获取的土壤粒径组成及有机质质量比作为传递函数的输入项,以方便准确地估算区域尺度的土壤水力学参数。最终建立的土壤饱和导水率 PTFs 模型如表 4 所示。

表 2 土壤饱和导水率 K_s 与基本物理性质相关矩阵
Tab.2 Correlation matrix of K_s via basic soil properties

K_s	O_m	O_m^2	$S_a \times O_m$	$C_l \times O_m$	$O_m \times B_d$	O_m^3	$S_a^2 \times O_m$	$S_a \times S_i \times O_m$
相关性	0.366 ^a	0.464 ^b	0.381 ^a	0.346 ^a	0.318 ^a	0.507 ^b	0.304 ^a	0.378 ^a
显著性	0.017	0.002	0.013	0.025	0.040	0.001	0.050	0.014
n	42	42	42	42	42	42	42	42

K_s	$S_a \times C_l \times O_m$	$S_a \times O_m^2$	$S_a \times O_m \times B_d$	$S_i \times O_m^2$	$C_l^2 \times O_m$	$C_l \times O_m^2$	$C_l \times O_m \times B_d$	$O_m^2 \times B_d$
相关性	0.449 ^b	0.498 ^b	0.336 ^a	-0.402 ^b	0.311 ^a	0.460 ^b	0.307 ^a	0.437 ^b
显著性	0.003	0.001	0.030	0.008	0.045	0.002	0.048	0.004
n	42	42	42	42	42	42	42	42

注:a 为在 0.05 水平(双侧)下显著相关;b 为在 0.01 水平(双侧)下显著相关。

表 3 K_s 初选参数组合及评价误差值
Tab.3 Parameter sets and evaluation errors for K_s in primary election

编号	$\overline{R_D}$	$\overline{A_D}/\%$	R^2	常数项	$S_i \times O_m^2$	$C_l \times O_m^2$	O_m^3
301	0.310	6.837	0.544	32.098	-1.813	4.432	39.508
356	0.022	11.476	0.510	32.061	-1.812	4.439	39.435
527	0.599	14.096	0.500	30.023	-2.014	6.585	36.018
758	0.561	28.916	0.508	32.092	-2.268	7.651	39.835

表 4 4 种土壤传递函数的输入变量、函数形式及预测 K_s 评价误差值

Tab.4 Input variables, specific forms, and the evaluation error for K_s of four pedo-transfer functions

函数名称	输入变量	函数形式	$R_{MSE}/(\text{cm} \cdot \text{d}^{-1})$	$A_{PE}/\%$	R^2	r
大兴区 PTF	$S_a、S_i、C_l、O_m$	$K_s = 32.098 - 1.813S_i \times O_m^2 + 4.432C_l \times O_m^2 + 39.508O_m^3$	40.525	204.738	0.544	0.742
文献[15]	$S_i、C_l、\ln S_a$	$K_s = 98.702 - 0.586S_i + 1.819C_l - 15.66\ln S_a$	37.518	499.220	0	-0.010
文献[14]	$S_a、C_l、B_d、O_m$	$\ln K_s = 13.262 - 1.914\ln S_a - 0.974\ln S_i - 0.058C_l - 1.709\ln O_m + 2.885O_m - 8.026\ln B_d$	53.808	255.307	0.090	0.481
文献[17]	$S_a、S_i、C_l、B_d$	—	26.570	958.675	0	0.251

注:大兴区 PTF 采用美国制粒径分级,文献[15]采用美国制粒径分级,文献[14]采用国际制粒径分级,文献[17]采用美国制粒径分级;3 种传递函数计算所得 K_s 单位均为 cm/d 。

2.2.2 不同土壤饱和导水率传递函数对比

图 3 为大兴区土壤 PTF 与其他 3 组传递函数实测值与预测值对比图,由于 K_s 变化范围较大且服从对数分布的随机变量^[19],为清楚地反映不同传递函数的模拟效果,本文中采用对数坐标。由图 3 及不同 PTF 对饱和导水率预测的模拟评价指标 R_{MSE} 、 A_{PE} 、 R^2 及相关系数 r (表 4)可以看出,文献[14]函数的预测结果大部分偏小,相关系数为 0.481, R^2 仅为 0.09;文献[15]、文献[17]的预测结果多数偏大,精度较差, R_{MSE} 、 A_{PE} 值均较大, R^2 分别为零,认为不能用来预测大兴区域土壤导水率。文献[15]建立的模型虽基于北京市通州区土壤,距离本文研究区域非常近,但由于通州区土壤主要为轻壤质土,大兴区主要为砂质及砂壤土,两区土壤的质地有差别,使文献[15]建立的模型对大兴区土壤饱和导水率预测结果也较差。本文所

建立的土壤传递函数相关系数为 0.742, R^2 为 0.544,预测结果优于其他 2 种模型预测结果,证明了土壤传递函数具有明显的区域性。对于部分 K_s 较小 ($K_s < 10 \text{ cm}/\text{d}$) 的实测值,大兴区 PTF 模型有过高估计;但大部分 K_s 实测值与大兴区土壤 PTF 的模拟值分布在 1:1 线附近。可以认为本文所建立的大兴区土壤传递函数模型基本上反映了研究区域土壤水力特征性质,可以用作大兴区域土壤饱和导水率的评估。

2.3 大兴区域农田土壤饱和导水率预测

根据前文所建立的大兴区土壤饱和导水率传递函数,估算其他 94 土钻采样点 0~80 cm 的土壤饱和导水率,并利用 ArcGIS 中的克里金插值法,对预测的点源数据进行空间区域尺度扩展,得到大兴区 0~80 cm 农田土壤饱和导水率空间分布图(图 4)。可以看出,不同深度的土壤饱和导水率取值范围不

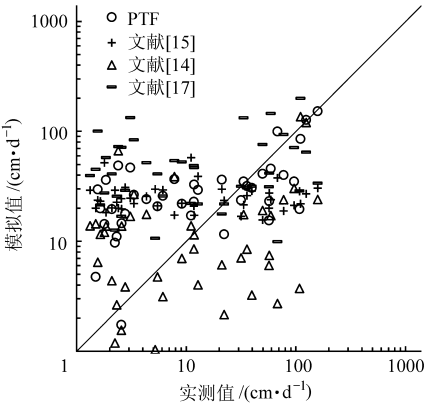


图 3 大兴区土壤 K_s 实测值与预测值对比

Fig.3 Comparison of K_s between measured data and simulation

同,表层土壤 K_s 空间变异性最大,且数据普遍大于其他深度。这主要是因为表层土壤由于受人为影响大,存在较多植物根系或虫洞,土壤较疏松且孔隙度

相对大,通透性好,水分较易通过入渗进入土体,从而使得 K_s 较大。另一方面,从土壤 PTF 函数归一化灵敏度^[20]分析结果来看(图 5),3 个参数对大兴区土壤饱和导水率的敏感程度从大到小依次为土壤有机质质量比、土壤粘粒质量分数、土壤粉粒质量分数。表层土壤由于灌水、施肥等田间管理措施,有机质质量比较深层土壤大,使得大兴区土壤 K_s 在 0 ~ 20 cm 深度较其他深度略大, K_s 取值从 10 ~ 280 cm/d 不等,以旧宫地区、西红门地区 K_s 值最大,其他 3 层土壤导水率变化范围较小。在 40 ~ 60 cm 深度土壤中,黄村地区及北臧村镇砂粒含量较多,土壤饱和导水率较大, K_s 大部分在 25 ~ 40 cm/d 内。采育镇 60 ~ 80 cm 深度 K_s 为 15 ~ 20 cm/d 之间,这主要是因为该地区此层内粘粒含量较多,土壤容积密度较大,粘壤土居多,饱和导水率较小。

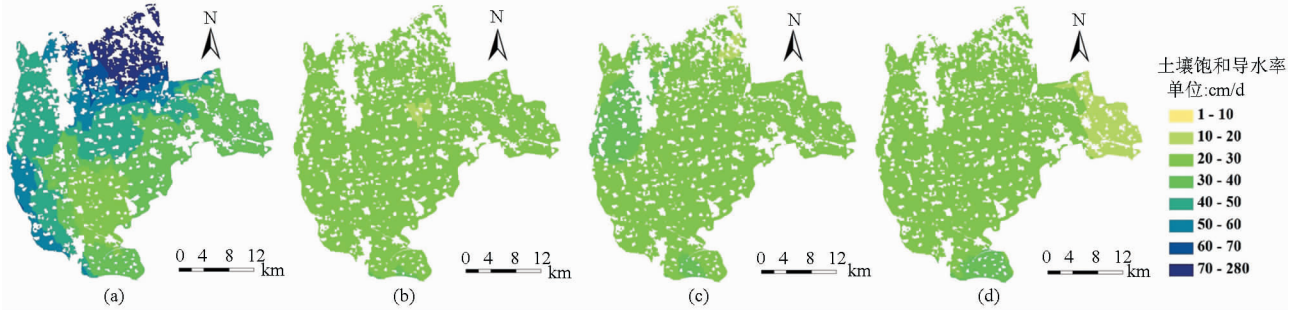


图 4 大兴区 0 ~ 80 cm 土壤饱和导水率 K_s 分布图

Fig.4 Distribution map of K_s in the depth of 0 ~ 80 cm in Daxing district

(a) 0 ~ 20 cm (b) 20 ~ 40 cm (c) 40 ~ 60 cm (d) 60 ~ 80 cm

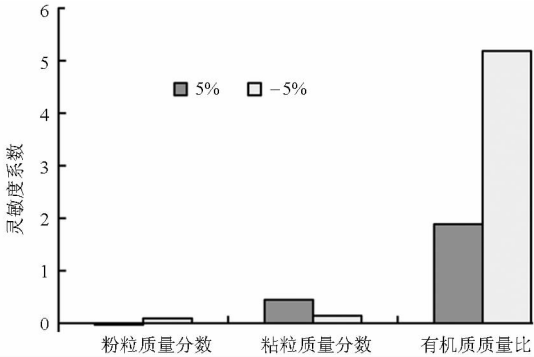


图 5 大兴区土壤 PTF 中各参数的归一化灵敏度系数

Fig.5 Normalized sensitivity coefficient of each parameter in PTF of Daxing district

3 结论

(1)大兴区土壤 K_s 与 O_m 的变异性较大,土壤容

积密度变异性最小, C_v 值分别达到 139%、75.4% 及 8.7%。

(2)以多折交叉验证方法优选后 PTF 模型的 R_{MSE} 为 40.525 cm/d、 A_{PE} 为 204.738%、 R^2 为 0.544,模型效果优于 3 种现有 PTF 模型的预测结果,证明了土壤传递函数在不同地区的应用存在明显的限制。

(3)大兴区域表层 0 ~ 20 cm 土壤饱和导水率的空间变异性较大,取值在 10 ~ 280 cm/d 不等。大兴区北部亦庄地区及西南部庞各庄镇、榆垓镇沿永定河一带 K_s 较小。在 60 ~ 80 cm 深度处,东部采育镇有粘壤土夹层, K_s 预测值较小,在 10 ~ 20 cm/d 之间。

参 考 文 献

1 Schaap M G, Leij F J, van Genuchten M T. Neural network analysis for hierarchical prediction of soil hydraulic properties[J]. Soil Science Society of America Journal, 1998, 62(4): 847 - 855.

2 李慧霞,刘建立,朱安宁,等. 预测天然文岩流域土壤饱和导水率的土壤转换函数方法比较研究[J]. 土壤,2010,42(3):

438 – 445.

Li Huixia, Liu Jianli, Zhu Anning, et al. Comparison study of soil pedo-transfer functions in estimating saturated soil hydraulic conductivity at Tianranwenyanqu Basin[J]. Soils, 2010, 42(3) : 438 – 445. (in Chinese)

3 Mei D, Hui H, Christoph N, et al. 3-D mapping of regional soil hydraulic characteristics through establishment of pedo-transfer functions[J]. Disaster Advances, 2012, 5(4) : 788 – 795.

4 Minasny B, Mcbratney A. The neuro-m method for fitting neural network parametric pedotransfer functions[J]. Soil Science Society of America Journal, 2002, 66(2) : 352 – 361.

5 Coppola A, Comegna A, Dragonetti G, et al. A stochastic texture-based approach for evaluating solute travel times to groundwater at regional scale by coupling gis and transfer function[J]. Procedia Environmental Sciences, 2013, 19: 711 – 722.

6 Friedman J H. Multivariate adaptive regression splines[J]. The Annals of Statistics, 1991, 19(1) : 1 – 67.

7 邹刚华. 亚热带典型流域土壤水文性质模型模拟及氮素分布特征研究[D]. 长沙:中南林业科技大学,2012.

8 张均华,刘建立,张佳宝,等. 常熟地区水稻土饱和导水率的间接方法研究[J]. 土壤通报,2010,41(4) :778 – 782. Zhang Junhua, Liu Jianli, Zhang Jiabao, et al. Indirect methods to estimate saturated soil hydraulic conductivity in paddy soil of Changshu[J]. Chinese Journal of Soil Science, 2010, 41(4) : 778 – 782. (in Chinese)

9 Manyame C, Morgan C L, Heilman J L, et al. Modeling hydraulic properties of sandy soils of Niger using pedotransfer functions[J]. Geoderma, 2007, 141(3 – 4) : 407 – 415.

10 Wösten J H M. Pedotransfer functions to evaluate soil quality[M]//Gregorich E G, Carter M R. Soil Quality for Crop Production and Ecosystem Health, Developments in Soil Science Amsterdam: Elsevier, 1997,25: 221 – 245.

11 Aimrun W, Amin M S M, Eltaib S M. Effective porosity of paddy soils as an estimation of its saturated hydraulic conductivity[J]. Geoderma, 2004, 121(3) : 197 – 203.

12 Aimrun W, Amin M S M. Pedo-transfer function for saturated hydraulic conductivity of lowland paddy soils[J]. Paddy and Water Environment, 2009, 7(3) : 217 – 225.

13 Hodnett M G, Tomasella J. Marked differences between van Genuchten soil water-retention parameters for temperate and tropical soils; a new water-retention pedo-transfer functions developed for tropical soils[J]. Geoderma, 2002, 108(3) : 155 – 180.

14 Li Y, Chen D, White R E, et al. Estimating soil hydraulic properties of Fengqiu County soils in the North China Plain using pedo-transfer functions[J]. Geoderma, 2007, 138(3 – 4) : 261 – 271.

15 王相平. 区域农田水氮利用效率及氮素淋失风险研究[D]. 北京:中国农业大学,2010.

16 李晓鹏,张佳宝,吉丽青,等. 土壤传递函数在计算土壤饱和导水率中的应用[J]. 灌溉排水学报,2009,28(2) :70 – 73.

17 Schaap M G. Rosetta: Pedotransfer function software[CP/OL][1999]. <http://www.ars.usda.gov/Services/docs.htm?docid=8953>.

18 孙美,蒙格平,张晓琳,等. 集约化种植区硝态氮在土壤剖面中的分布与累积特征[J]. 环境科学学报,2012,32(4) :902 – 908.

19 Tietje O, Hennings V. Accuracy of the saturated hydraulic conductivity prediction by pedo-transfer functions compared to the variability within FAO textural classes[J]. Geoderma, 1996, 69(1) : 71 – 84.

20 朱国威,任理. 根系带水质量模型参数灵敏度分析与标定的研究[J]. 灌溉排水学报,2011,30(2) :5 – 9. Zhu Guowei, Ren Li. Parameters sensitivity analysis and scaling of RZWQM[J]. Journal of Irrigation and Drainage, 2011, 30(2) : 5 – 9. (in Chinese)

Pedo-transfer Function for Saturated Hydraulic Conductivity
of Agricultural Soil Based on Cross-validation

Sun Mei¹ Zhang Xiaolin¹ Feng Shaoyuan² Huo Zailin¹

(1. College of Water Resources and Civil Engineering, China Agricultural University, Beijing 100083, China
2. College of Hydraulic Science and Engineering, Yangzhou University, Yangzhou 225009, China)

Abstract: Based on the sample data of agricultural soil profile, the locally-developed pedo-transfer function (PTF) of agricultural soil saturated hydraulic conductivity (K_s) in Daxing district was established using multiple regression with the independent variables, including constant and the monomials with variables of S_i content, C_i content, and organic matter content. The PTF was validated simultaneously using multi-fold cross-validation method. The performance of Daxing District PTF built in this paper was compared with other three existing PTFs in predicting K_s . The PTF from this study was used in regional prediction of K_s in depth of 0 ~ 80 cm. The results showed that the root-mean-square error, average absolute relative error, and coefficient of determination of multiple regression function for predicting K_s in depth of 0 ~ 100 cm were 40.525 cm/d, 204.738%, 0.544, respectively, which were much lower than that of other three PTFs. The K_s was of great spatial variability in the top layer soil of 0 ~ 20 cm, and was relative small along the Yongding River in the north and southwest of Daxing district. The PTF from this study also performed well in the characterization of layered C_i loam in depth of 60 ~ 80 cm in Caiyu. It can be used in regional evaluation of K_s in Daxing district.

Key words: Agricultural soil Saturated hydraulic conductivity Pedo-transfer function Multiple regression Multi-fold cross-validation