

doi:10.6041/j.issn.1000-1298.2015.04.016

基于基因表达式编程算法的参考作物腾发量模拟计算*

王升 陈洪松 聂云鹏 付智勇 王克林 丁亚丽

(中国科学院亚热带农业生态研究所, 长沙 410125)

摘要: 选取都安气象站5年(2008—2012年)的逐日气象数据,包括日最高气温、最低气温、平均风速、日照时数以及相对湿度5个气象要素的不同组合作为输入,并以FAO-56 Penman-Monteith法(FAO P-M)的计算结果作为标准值,采用基因表达式编程算法(GEP)及径向基函数网络算法(RBFNN)对参考作物腾发量 ET_0 进行模拟计算,并将模拟结果与Hargreaves模型的计算结果进行比较,用决定系数 R^2 和均方根误差 R_{MSE} 作为评价指标。结果表明,GEP模型能够捕捉到 ET_0 的变化,具有较强的适用性,与FAO P-M公式的计算值有很高的一致性。引入关键气象因子(气温和相对湿度)后,模型的决定系数 R^2 达到0.914,均方根误差 R_{MSE} 为0.240 mm/d。在相同输入情况下GEP模型计算精度高于RBFNN模型和Hargreaves模型,并建立了可以替代Hargreaves模型的GEP模型及缺少相对湿度RH时的GEP模型。结果表明,在缺乏相关气象因子时,可以利用GEP模型模拟 ET_0 。

关键词: 参考作物腾发量 基因表达式编程 气象因子组合 Penman-Monteith模型 Hargreaves模型

中图分类号: S161.4 **文献标识码:** A **文章编号:** 1000-1298(2015)04-0106-07

Simulation of Reference Evapotranspiration Based on Gene-expression Programming Method

Wang Sheng Chen Hongsong Nie Yunpeng Fu Zhiyong Wang Kelin Ding Yali

(Institute of Subtropical Agriculture, Chinese Academy of Sciences, Changsha 410125, China)

Abstract: Reference evapotranspiration (ET_0) is a major component of the hydrological cycle. Accurate assessment of evapotranspiration is needed for water resources management and irrigation scheduling. The performance ability of gene-expression programming (GEP) and radical basis function neural network (RBFNN) was investigated for modeling ET_0 in weather station of Du'an for a 5-year period (2008—2012). The data set was comprised of daily maximum temperature, minimum temperature, sunshine duration and relative humidity, which was employed for modeling ET_0 by using FAO-56 Penman-Monteith equation as reference. GEP results were compared with RBFNN and Hargreaves models, and their performances were evaluated through determination coefficient (R^2) and root mean square error R_{MSE} . Based on the comparisons, GEP was found to perform better than RBFNN and Hargreaves models. The GEP model which can replace Hargreaves model and the GEP model without relative humidity were established. Statistically, GEP is an effectual modeling tool for successfully computing reference evapotranspiration.

Key words: Reference evapotranspiration Gene-expression programming Meteorological factors combination Penman-Monteith model Hargreaves model

引言

参考作物腾发量(ET_0)是计算作物蒸发蒸腾量

的关键因子,准确的估算参考作物腾发量对作物需水管理、灌溉制度设计、水资源管理等具有重要意义。目前,有许多采用气象因子计算 ET_0 的数学模

收稿日期:2014-05-18 修回日期:2014-07-25

*国家自然科学基金资助项目(41171187,31100294)和中国科学院西部行动计划资助项目(KZCX2-XB3-10)

作者简介:王升,博士生,主要从事生态水文研究,E-mail:hjdx@foxmail.com

通讯作者:陈洪松,研究员,博士生导师,主要从事土壤物理与水土保持研究,E-mail:hbchs@isa.ac.cn

型^[1],使用较多的是世界粮食和农业组织推荐的 FAO-56 Penman-Monteith 模型(以下简称 FAO P-M 模型),它也常常用于对其他 ET_o 模型的校准^[2]。FAO P-M 模型主要有两大优势:一方面它能够在世界范围内的不同气候区使用,另一方面其结果在世界各地已通过蒸渗仪校准^[3]。其最大的缺陷是需要较为完备的气象数据,然而,即使在发达国家能同时测量这些气象数据的气象站点也是有限的,而且其对数据质量有严格的要求,所以其应用受到一定的限制^[1]。

近年来,人工智能技术(Artificial intelligence, AI),如人工神经网络(ANN)、自适应模糊神经推理系统(ANFIS)和基因编程(GP)等被引入水文学和水资源工程学领域,取得了较大的进展。Kumar 等^[4]、Landeras 等^[3]、霍再林等^[5]和徐俊增等^[6]采用 ANN 对 ET_o 的模拟计算进行了研究,表明 ANN 预测 ET_o 是可行的。Kisi 等^[7-8]发现 ANFIS 能够成功地用于计算 ET_o ;Moghaddamnia 等^[9]在干燥炎热的伊朗分别采用 ANN 和 ANFIS 技术模拟计算 ET_o ,表明 ANFIS 的结果精度更高。顾世祥等^[10]探讨了径向基函数神经网络(RBFNN)模拟和预测 ET_o 的可行性,发现其具有一定的精度。Khu 等^[11]在法国采用 GP 进行实时径流预测,并将结果与实测数据及传统方法的结果进行比较,表明 GP 模型能够获得足够精度的结果。Aytok 等^[12]利用 GP 建立了降雨-产流模型,Shiri 等^[13]比较了 GP 和 ANFIS 对地下水位波动的模拟精度,表明 GP 模拟的效果更好。基因表达式编程(GEP)是葡萄牙学者 Ferreira 于 2001 年提出的革命性新型算法^[14],是一种通用的自适应式随机搜索算法,能够在缺乏先验知识,只有实验数据的情况下挖掘出较为准确的公式,以其较强的普适性和较高的精确度在很多应用领域都取得了很好的实际效果。目前在水文学领域有关 GEP 算法的应用研究很有限^[15],在必要的气象资料较为缺乏的中国,还未开展过应用 GEP 技术模拟计算 ET_o 的研究。本文利用都安县的气象数据,采用不同气象因子的组合,以 FAO P-M 模型的结果为标准,比较 GEP、RBFNN 和 Hargreaves 模型对 ET_o 的模拟效果,并建立采用较少气象因子获得足够精度 ET_o 的 GEP 模型,探索 GEP 算法在 ET_o 模拟计算领域的应用,以期对当地农业用水管理提供理论支持。

1 材料与方法

1.1 数据来源

气象数据来自都安县气象站(北纬 23.93°,东

经 108.10°,海拔高度 170.8 m),包括:平均气温 T_{mean} 、最高气温 T_{max} 、最低气温 T_{min} 、相对湿度 RH 、日照时数 n 和风速 u 的 5 年数据(2008—2012 年)。因为 ET_o 与其他气候因素相比变化程度较小,由 5 年的逐日 ET_o 值对 ET_o 模型进行建立及校正是可以接受的。

1.2 参考作物腾发量模型

1.2.1 FAO P-M 模型

以 FAO P-M 模型^[1]计算的 ET_o 作为标准值,对其他方法计算的结果进行评估,其形式为

$$ET_o = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T + 273} u (e_s - e_a)}{\Delta + \gamma(1 + 0.34u)} \quad (1)$$

式中 R_n ——太阳净辐射通量, MJ/(m²·d),由日照时数结合地理位置信息算得

G ——土壤热通量, MJ/(m²·d)

Δ ——饱和水汽压-温度曲线的斜率, kPa/K

γ ——湿度常数, kPa/K

e_s ——饱和水汽压, kPa

e_a ——实际水汽压, kPa

u ——2 m 高度处的平均风速, m/s

T ——2 m 高度处的平均气温, °C

1.2.2 Hargreaves 模型

模型形式为

$$ET_o = 0.0023R_a \left(\frac{T_{\text{max}} - T_{\text{min}}}{2} + 17.8 \right) \sqrt{T_{\text{max}} - T_{\text{min}}} \quad (2)$$

式中 R_a ——碧空太阳总辐射,其值由气象站点纬度及日序数算得^[16]

1.3 ET_o 敏感性分析

ET_o 的敏感性分析是指在指定时期内一个或几个气象要素发生变化时,对 ET_o 变化影响程度的分析。了解 ET_o 受单要素或多要素变化引起气候变化的敏感性,可以更加深刻的认识气候变化和 ET_o 之间的关系^[17]。用数学式定义,以气象变量的敏感系数来反映其敏感性的分析方法称为敏感系数法。基于 FAO P-M 模型推求的无量纲相对敏感系数(以下简称敏感系数)为

$$S_{V_i} = \lim_{\Delta V_i \rightarrow 0} \frac{\Delta ET_o / ET_o}{\Delta V_i / V_i} = \frac{\partial ET_o}{\partial V_i} \frac{V_i}{ET_o} \quad (3)$$

式中 V_i ——第 i 个变量

不同气象要素有不同的变化范围和单位,为了便于敏感性的比较,敏感系数法采用求偏导数将所研究对象转化为无量纲的形式。敏感系数正值表明 ET_o 随着气候变量的增加而增加,反之亦然;其绝对值越大则气候变量对 ET_o 的影响越大,反之亦然。

本文敏感系数计算过程中偏导数计算通过 Matlab 软件中的符号计算实现,其他运算通过数值计算实现。

1.4 基因表达式编程算法概述

基因表达式编程(Gene expression programming, GEP)起源于遗传算法(Genetic algorithms, GA)和遗传编程(Genetic programming, GP),并继承了二者的优点。在 GEP 进化过程中,既保持了 GA 算法遗传操作的便捷性,又拥有 GP 算法的非线性树结构,实现了利用简单编码解决复杂问题的目的,运算效率比传统的遗传算法和遗传编程高出 100 ~ 10 000 倍^[14]。

1.4.1 GEP 的基因和染色体结构

在 GEP 中,基因组或者染色体组成一个线性的符号串,这个符号串由一个或多个基因组成。GEP 基因由头部(head)和尾部(tail)组成。头部由函数操作符(来自于函数集:运算符和初等函数)或终结符(来自于终结符集:自变量和常数)组成,而尾部仅含有终结符。头部长度 h 根据具体问题选定,尾部长度 l 则是 h 和 m 的函数

$$l = h(m - 1) + 1 \quad (4)$$

式中 m ——所需参数量最多的函数的参数个数

对于开方、正弦和余弦运算, m 取 1; 而加减乘除等四则运算, m 取 2。GEP 中基因的这种结构避免了遗传操作中产生大量的无效编码,提高了算法的执行效率。

GEP 的编码规则为:将表达式根据其语意表示为表达式树(ET),然后从上到下,从左到右按层遍历 ET,这样得到的符号序列即为基因编码的有效部分。该序列又称为 K 表达式(K-expression)。GEP 中编码串的长度是固定的,它由位于前段的 K 表达式和后面的填充部分组成。例如,对于表达式 $\sqrt{(a+b)(c+d)}$,函数集 $\{+, -, *, Q\}$, Q 表示开平方,终止符集 $\{a, b, c, d\}$,表示常数或自变量,图 1 为对应的表达式树(ET)。

层次遍历得到 K 表达式:“ $Q * + - abcd$ ”。显然 $m = 2$,如果取 $h = 5$,则尾部长度 $l = 6$,那么基因长度为 11。由于 K 表达式只有 8 位(编码区域),则尾部剩余部分由终止符中的元素随机填充(非编码区域),所以基因编码可以为 $Q * + - abcdaab$ 。本例中基因表达式长度为 8 个字符,基因候补的非编码区域为程序的进化提供了很大的空间。

在 GEP 中,若干个类似的等长基因按照一定的组合方式构成了 GEP 染色体。程序运行时,基因的数目以及基因的头部长度都是事先选定的。染色体中每个基因片段都可以解码成一个子表达式树

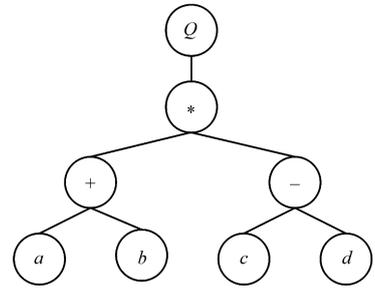


图 1 表达式树

Fig. 1 Expression tree

(Sub-expression tree, Sub-ET),多个子表达式树构成了更复杂的多子树(Multi-subunit ET)表达式树。GEP 染色体的这种特殊结构以及丰富的遗传算子为 GEP 解决复杂问题提供了基本保证。

1.4.2 GEP 的适应度函数

在 GEP 的符号回归和函数发现过程中,所有的进化计算算法都需要对代表个体的问题解答进行评价。由于解答是一段程序,确切的说是一个表达式,对表达式进行评价,就是要评测利用该表达式得到的计算值和目标值的接近程度。Ferreira 提出了 2 种适应度计算函数,一种为基于绝对误差的适应度函数,另一种为基于相对误差的适应度函数。

基于绝对误差的适应度函数

$$f_i = \sum_{j=1}^{C_i} (M - |C_{(i,j)} - T_{(j)}|) \quad (5)$$

基于相对误差的适应度函数

$$f_i = \sum_{j=1}^{C_i} \left(M - \left| 100 \frac{C_{(i,j)} - T_{(j)}}{T_{(j)}} \right| \right) \quad (6)$$

式中 M ——常数,它控制适应度的最大取值

$C_{(i,j)}$ ——由第 i 个染色体对应的函数表达式利用第 j 组数据集得到的函数值

$T_{(j)}$ ——第 j 组数据集的真实值

C_i ——样本的数目

1.4.3 GEP 的遗传操作

GEP 的染色体结构是其解决问题的物质基础,而染色体在遗传过程中丰富多样的操作算子是 GEP 解决复杂问题的重要保证。通常 GEP 包括选择、变异、插串和重组等几大类遗传算子。

选择算子:根据适应度函数值,选择种群中适应度好的个体并复制到下一代产生新的种群,这里选择和复制是同步进行的。

变异算子:在维持种群多样性中起着主要的作用,变异可以发生在染色体的任意位置,但为了保持染色体结构的完整性,发生在头部的变异可以变异成函数符和终结符,而发生在尾部的变异只能变异成终结符。

插串算子: GEP 中特有的遗传算子, 包括 IS 插串、RIS 插串(根插串)和基因插串。

重组算子: 2 个同源染色体通过交配重组形成新的染色体, 从而产生新的个体, 包括单点重组、两点重组和基因重组。

1.5 基于 GEP 的参考作物腾发量建模

基于基因表达式编程算法的气象变量模拟计算 ET_0 一共有 5 个步骤:

(1) 确定输入变量组合。影响 ET_0 的 5 个直接气象因素中, 相对湿度 RH 、日照时数 n 和风速 u 的观测数据是许多地区缺乏的, 而最高气温 T_{max} 及最低气温 T_{min} 较易观测和预报。本研究选择温度项与 RH 、 n 、 u 组合作为输入因子。由于碧空太阳总辐射 R_a 反映了太阳理论辐射, 仅与观测点的地理位置和日期相关而无需观测, 引入模型可提高模拟的精度^[18], 故 R_a 也可作为一个输入因子。

(2) 选择函数集。包括 $x^2, x^3, \ln x, e^x, x^y, +, -, \times, \div$ 。

(3) 选择适应度函数。本研究选用基于绝对误差的适应度函数式(4)。

(4) 选择所需的遗传算子及相应概率, 如表 1 所示。

(5) 选择遗传终止的标准。当 R_{MSE} 接近于 0 且基本稳定时即停止, 接受产生的模型。

以都安站 2008—2010 年的日气象资料及 FAO P-M 模型计算的 ET_0 为训练样本, 用 2011—2012 年的日气象资料及 FAO P-M 模型计算的 ET_0 为检验样本, 得到不同气象因子组合下的 ET_0 模型。

表 1 GEP 模型中遗传算子的取值

Tab.1 Genetic operators employed in GEP models

参数	数值	参数	数值
染色体个数	30	单点重组算子	0.3
头部长度	8	两点重组算子	0.3
基因个数	3	基因重组算子	0.1
连接函数	+	基因插串算子	0.1
适应度函数	R_{MSE}	IS 插串算子	0.1
变异算子	0.044	RIS 插串算子	0.1
倒串算子	0.1		

1.6 径向基函数神经网络模型法 (RBFNN)

径向基函数神经网络模型(Radial basis function-artificial neural network, RBF-ANN)是一种 3 层前向式网络, 其模型结构详见文献[10]。

1.7 评价指标

本研究采用 2 个评价指标, 即决定系数 R^2 和均方根误差 R_{MSE} , 其计算式为

$$R^2 = \frac{\left[\sum_{i=1}^N (y_i - \bar{y})(y'_i - \bar{y}') \right]^2}{\sum_{i=1}^N (y_i - \bar{y})^2 \sum_{i=1}^N (y'_i - \bar{y}')^2} \quad (7)$$

$$R_{MSE} = \sqrt{\frac{\sum_{i=1}^N (y_i - y'_i)^2}{N}} \quad (8)$$

式中 y_i, y'_i ——FAO P-M 和其他模型计算的 ET_0 值
 \bar{y}, \bar{y}' —— y 和 y'_i 的平均值
 N ——检验模型时的天数, 取值 731 d

2 结果与分析

2.1 都安站 ET_0 敏感性分析

以都安站近 60 年(1951—2010 年)的气温、风速、净辐射和相对湿度的月平均值计算月敏感系数, S_{Rn}, S_T, S_u, S_{RH} 分别是太阳辐射、平均温度、风速和相对湿度的敏感系数, 其中 R_n 是由日照时数根据 Penman-Monteith 公式计算出来的太阳净短波辐射, 由于 S_{RH} 为负值, 为便于比较, 图中 S_{RH} 为原 S_{RH} 乘以 -1 得到。

图 2 为都安站气象因子敏感系数曲线。结果表明 1—12 月份太阳净辐射(日照时数)、气温和风速对 ET_0 的敏感系数均为正值, 而相对湿度对 ET_0 的敏感系数则为负值, 说明净辐射、气温和风速与 ET_0 正相关, 相对湿度与 ET_0 负相关。太阳辐射(日照时数)对 ET_0 的敏感系数呈单峰型分布, 夏季较大, 风速对 ET_0 的敏感系数呈单谷型分布, 夏季较小, 相对湿度对 ET_0 的敏感系数在雨季(6、7、8 月份)较小; 气温对 ET_0 的敏感系数在年内变化较小。对都安站一年的 ET_0 影响较大的因素依次为相对湿度、太阳净辐射(日照时数)、气温、风速。

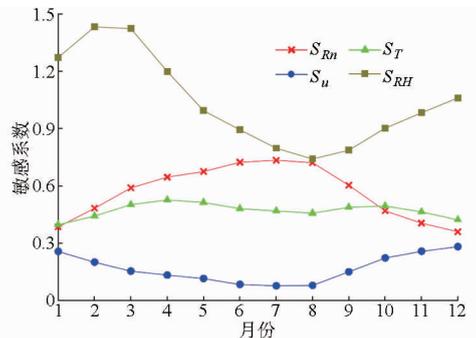


图 2 都安站气象变量敏感系数

Fig.2 Sensitivity coefficients of climatic variables of Du'an station

2.2 不同气象因子组合下 GEP 和 RBFNN 模型结果对比

分别采用 GEP、RBFNN 算法在不同气象因子组合下, 建立的 ET_0 模型计算结果与 FAO P-M 模型

结果的决定系数(R^2)和均方根误差(R_{MSE})统计结果如表2所示。可见GEP模型在不同的输入组合下 R^2 的变化范围为0.830~0.942, R_{MSE} 的变化范围为0.375~0.646 mm/d。

表2 不同输入因子的组合及模拟结果

Tab.2 Error statistics of variable input factors for applied models

方案	输入因子	GEP		RBFNN	
		R^2	R_{MSE} /(mm·d ⁻¹)	R^2	R_{MSE} /(mm·d ⁻¹)
1	T_{min}, T_{max}	0.830	0.646	0.829	0.662
2	T_{min}, T_{max}, RH	0.906	0.485	0.904	0.545
3	T_{min}, T_{max}, u	0.848	0.605	0.825	0.664
4	T_{min}, T_{max}, n	0.895	0.510	0.897	0.500
5	T_{min}, T_{max}, R_a	0.835	0.633	0.828	0.673
6	$T_{min}, T_{max}, R_a, RH$	0.914	0.457	0.903	0.542
7	T_{min}, T_{max}, R_a, u	0.865	0.579	0.835	0.642
8	T_{min}, T_{max}, R_a, n	0.901	0.491	0.833	0.666
9	$T_{min}, T_{max}, R_a, RH, u$	0.859	0.606	0.889	0.520
10	$T_{min}, T_{max}, R_a, u, n$	0.920	0.449	0.834	0.643
11	$T_{min}, T_{max}, R_a, RH, n$	0.919	0.442	0.903	0.542
12	$T_{min}, T_{max}, R_a, RH, u, n$	0.942	0.375	0.888	0.520
Hargreaves	T_{min}, T_{max}, R_a	0.813	0.693		

从表2可看出,当缺少或增加某些关键的气象因子时,会显著影响模型的精度。方案1为气温模型,方案2、3、4分别在方案1的基础上添加RH、u和n,可见与方案1相比方案2、3、4的精度都有所提高,其中方案2气象因子组合精度提高最显著, R^2 由0.830增加到0.906, R_{MSE} 由0.646 mm/d降低到0.485 mm/d。其次为方案4,方案3。这可由 ET_o 对气象因子的敏感性来解释,2.1节的分析表明,都安站 ET_o 敏感的气象因子由大到小依次为相对湿度、日照时数、气温和风速。

方案5、6、7、8分别在方案1、2、3、4的基础上添加了 R_a ,精度均得到了提高。这是由于蒸散发的能

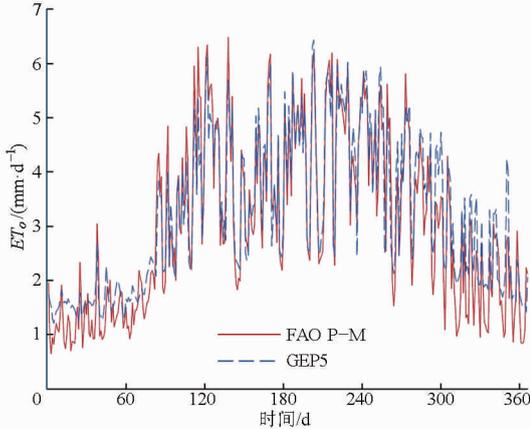


图3 GEP5和GEP10与FAO P-M结果比较

Fig.3 Comparison of results for GEP5 and GEP10 with ET_o from FAO P-M

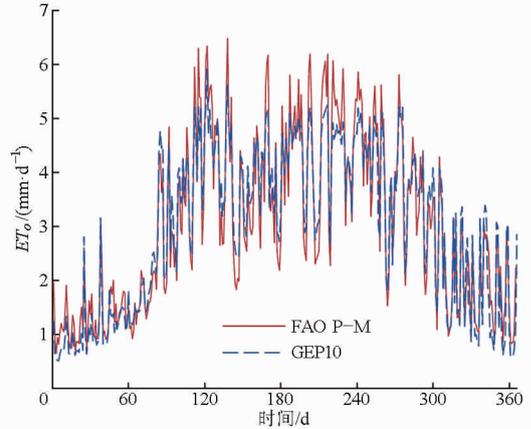
量来源为太阳辐射,一年中理论太阳辐射主要受日地相对距离和太阳高度角的控制,日地相对距离与日序数有关,太阳高度角与地理位置有关, R_a 综合反映了这两个因素,因此引入 R_a 可提高 ET_o 模型的精度,而且 R_a 不需要观测,可由计算得到,避免了观测导致的误差。另外可以发现,方案5的输入因子和传统方法Hargreaves模型相同,但精度得到了提高。由于GEP算法能够提供算术表达式,与其他神经网络算法的非线性黑箱过程不同,因此在仅有气温缺乏其他气象资料,且灌溉人员未掌握相关神经网络技术时,方案5得到的算术模型比较适用。方案5中3个输入因子 T_{min}, T_{max}, R_a 的GEP模型为

$$ET_o = 1.0013 \left(\frac{R_a}{T_{min} + T_{max}} \right)^2 + \ln(\ln(T_{min} + T_{min} R_a)) - \sqrt[3]{\frac{T_{min}}{T_{max} - T_{min}}} - e^{\sqrt{T_{min}}} \quad (9)$$

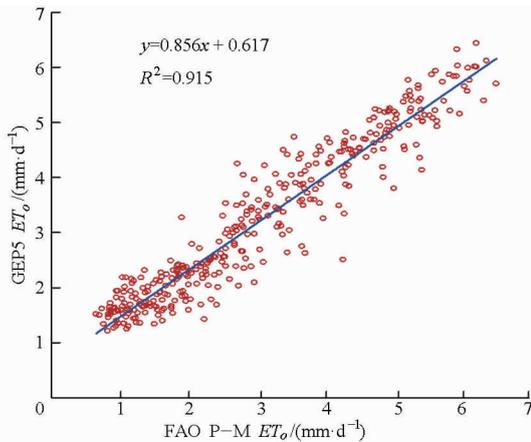
方案9、10、11为5因子模型,分别在方案6、7、8的基础上引入了u、n和RH,其中方案10的精度提高最大。方案12的输入因子和FAO P-M模型的输入因子相同,因此其精度也是最高的,但方案12与FAO P-M模型的结果仍然有差异,并不能替代FAO P-M模型,因此在气象数据完备的地区,FAO P-M模型是计算 ET_o 最好的方法。在相对湿度数据不易获取或容易缺失的情况下,方案10($R^2 = 0.92, R_{MSE} = 0.449$ mm/d)是最好的选择。方案10的输入因子为 $T_{min}, T_{max}, R_a, u, n$,在都安站其 ET_o 表达式为

$$ET_o = (R_a + u + T_{min})^{0.125 \ln T_{max}} + \ln(n + 671.734 \frac{1}{e^{T_{min} + T_{min}}}) + \sqrt{u} - 3.799 \quad (10)$$

在检验期2012年的方案5和方案10的模拟值与FAO P-M模型的 ET_o 对比如图3所示。由图3可见,GEP模型能够捕捉到逐日 ET_o 在年内的变化趋势,且方案10(GEP10)的结果比方案5(GEP5)更



接近 FAO P-M 模型的值。图 4 为线性回归分析, 检验期 2012 年方案 5 和方案 10 的模拟值与 FAO P-M 模型标准值的散点图也表明, GEP 模型与



FAO P-M 模型的计算值具有很高的 consistency。方案 10 的斜率更接近于 1 (0.921), 截距接近于 0 (0.198)。

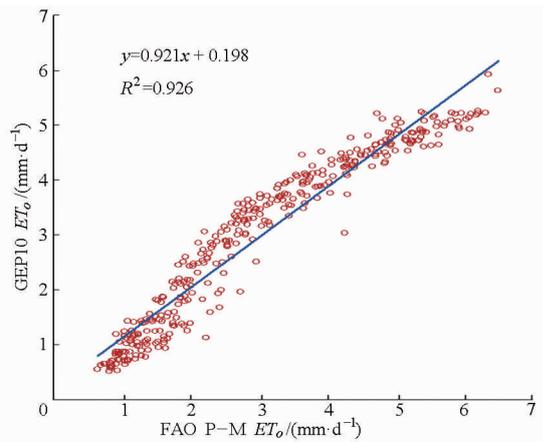


图 4 GEP5 和 GEP10 与 FAO P-M 值的线性回归分析

Fig. 4 Comparative scatter plots between results of GEP5 and GEP10 with ET_0 from FAO P-M

表 2 还给出了不同气象因子组合下, 采用径向基函数神经网络 (RBFNN) 模型获得的 ET_0 统计结果。结果表明, 除方案 4 和方案 9, GEP 模型的计算精度均高于 RBFNN 模型。与 RBFNN 及其他神经网络模型相比, GEP 模型最主要的优点是它可以明确地给出自变量和因变量之间的算术表达式, 这一方面便于挖掘自变量和因变量之间的理论关系, 另一方面即使不了解神经网络的灌溉工作人员也可以使用, 而 RBFNN 及其他神经网络模型是一个黑箱过程, 使用难度较大。

3 结论

(1) 研究 ET_0 模型时, 在缺少太阳净辐射而用日照时数进行计算的情况下, 引入反映日序数变化和地理位置信息的碧空太阳总辐射值 R_0 , 可以明显提高模型精度, 而且该参数不需要观测, 可以通过简单

的计算获得。

(2) GEP 模型有能力捕捉到 ET_0 的变化, 采用不同气象因子组合作为输入因子时, 随着气象因子的增加, GEP 模型模拟的精度随之提高。对于都安站, 引入关键的气象因子 (气温和相对湿度) 后, 模型的 R^2 达到 0.914, R_{MSE} 达到 0.457 mm/d, 这与不同地区 ET_0 对气象因子的敏感程度有关。

(3) 在仅有气温数据时, 采用 GEP 模型的计算结果优于 Hargreaves 模型, 因此可用 GEP 模型替代 Hargreaves 模型。由于相对湿度 RH 不易获取, 本研究建立了缺少相对湿度 RH 时的 GEP 模型, 具有较高的计算精度 ($R^2 = 0.920$, $R_{MSE} = 0.449$ mm/d)。

(4) 通过与 RBFNN 模型结果对比, GEP 模型计算的 ET_0 精度更高。而且 GEP 模型可以得到具体的数学表达式, 降低了模型使用难度, 使得对计算机技术并不精通的灌溉工作人员也可运用。

参 考 文 献

- Allen R G, Pereira L S, Raes D, et al. Crop evapotranspiration—guidelines for computing crop water requirements—FAO irrigation and drainage paper 56 [R]. Rome: FAO, 1998.
- 王昊, 许士国, 孙砾石. 扎龙湿地参照作物蒸散发估算的经验模型[J]. 水科学进展, 2007, 18(2): 246-251. Wang Hao, Xu Shiguo, Sun Leshi. Empirical models for reference evapotranspiration estimation in Zhalong wetland[J]. Advances in Water Science, 2007, 18(2): 246-251. (in Chinese)
- Landeras G, Ortiz-Barredo A, Lopez J J. Comparison of artificial neural network models and empirical and semi-empirical equations for daily reference evapotranspiration estimation in the Basque Country (Northern Spain) [J]. Agricultural Water Management, 2008, 95(5): 553-565.
- Kumar M, Raghuwanshi N, Singh R, et al. Estimating evapotranspiration using artificial neural network[J]. Journal of Irrigation and Drainage Engineering, 2002, 128(4): 224-233.
- 霍再林, 史海滨, 陈亚新, 等. 参考作物潜在蒸散量的人工神经网络模型研究[J]. 农业工程学报, 2004, 20(1): 40-43.
- 徐俊增, 彭世彰, 张瑞美, 等. 基于气象预报的参考作物蒸发蒸腾量的神经网络预测模型 [J]. 水利学报, 2006, 37(3): 376-379. Xu Junzeng, Peng Shizhang, Zhang Ruimei, et al. Neural network model for reference crop evapotranspiration prediction based on weather forecast[J]. Journal of Hydraulic Engineering, 2006, 37(3): 376-379. (in Chinese)

- 7 Kisi Ö. Generalized regression neural networks for evapotranspiration modelling[J]. *Hydrological Sciences Journal*, 2006, 51(6): 1092 – 1105.
- 8 Kisi Ö, Öztürk Ö. Adaptive neurofuzzy computing technique for evapotranspiration estimation[J]. *Journal of Irrigation and Drainage Engineering*, 2007, 133(4): 368 – 379.
- 9 Moghaddamnia A, Ghafari Gousheh M, Piri J, et al. Evaporation estimation using artificial neural networks and adaptive neuro-fuzzy inference system techniques[J]. *Advances in Water Resources*, 2009, 32(1): 88 – 97.
- 10 顾世祥, 王士武, 袁宏源. 参考作物腾发量预测的径向基函数法[J]. *水科学进展*, 1999, 10(2): 123 – 128.
Gu Shixiang, Wang Shiwu, Yuan Hongyuan. A radial basis function model predicting daily reference evapotranspiration[J]. *Advances in Water Science*, 1999, 10(2): 123 – 128. (in Chinese)
- 11 Khu S T, Liong S Y, Babovic V, et al. Genetic programming and its application in real-time runoff forecasting[J]. *Journal of the American Water Resources Association*, 2001, 37(2): 439 – 451.
- 12 AYTEK A, ALP M. An application of artificial intelligence for rainfall-runoff modeling[J]. *Journal of Earth System Science*, 2008, 117(2): 145 – 155.
- 13 Shiri J, Kisi Ö. Comparison of genetic programming with neuro-fuzzy systems for predicting short-term water table depth fluctuations[J]. *Computers & Geosciences*, 2011, 37(10): 1692 – 1701.
- 14 元昌安, 彭显忠, 覃晓, 等. 基因表达式编程算法原理与应用[M]. 北京: 科学出版社, 2010.
- 15 Guven A, Kisi Ö. Estimation of suspended sediment yield in natural rivers using machine-coded linear genetic programming[J]. *Water Resources Management*, 2011, 25(2): 691 – 704.
- 16 刘钰, Pereira L S, Teixeira J L, 等. 参照腾发量的新定义及计算方法对比[J]. *水利学报*, 1997, 28(6): 27 – 33.
Liu Yu, Pereira L S, Teixeira J L, et al. Update definition and computation of reference evapotranspiration comparison with former method[J]. *Journal of Hydraulic Engineering*, 1997, 28(6): 27 – 33. (in Chinese)
- 17 Yin Y, Wu S, Chen G, et al. Attribution analyses of potential evapotranspiration changes in China since the 1960s[J]. *Theoretical and Applied Climatology*, 2010, 101(1–2): 19 – 28.
- 18 胡庆芳, 尚松浩, 温守光, 等. 基于投影寻踪的径向基函数网络在参考腾发量预测中的应用[J]. *水利学报*, 2006, 37(9): 1151 – 1154.
Hu Qingfang, Shang Songhao, Wen Shouguang, et al. Using radial basis function network based on project pursuit to forecast reference evapotranspiration[J]. *Journal of Hydraulic Engineering*, 2006, 37(9): 1151 – 1154. (in Chinese)

~~~~~

(上接第 76 页)

- 8 杨晓斌. 基于 SPSS 软件的浦东新区长江引水量分析和预测[J]. *南水北调与水利科技*, 2012, 10(4): 70 – 73.  
Yang Xiaobin. Analysis and prediction of water diversion from the Yangtze River in the Pudong new district based on SPSS[J]. *South-to-North Water Diversion and Water Science & Technology*, 2012, 10(4): 70 – 73. (in Chinese)
- 9 旦木仁加甫. 中长期水文预报与 SPSS 应用[M]. 郑州: 黄河水利出版社, 2011.
- 10 门登霍尔 W, 辛塞奇 T. 统计学[M]. 5 版. 北京: 机械工业出版社, 2011.
- 11 何晓群. 多元统计分析[M]. 3 版. 北京: 中国人民大学出版社, 2012: 1 – 87.
- 12 陆军, 张红涛, 魏德云, 等. 基于时间序列分析的雾滴页面动态接触角预测与建模[J]. *农业机械学报*, 2013, 44(2): 80 – 86.  
Lu Jun, Zhang Hongtao, Wei Deyun, et al. Dynamic contact angle forecasting and modeling for drop leaf interface based on time series analysis[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2013, 44(2): 80 – 86. (in Chinese)
- 13 蒋智, 况明生. 重庆地区近 57 年降雨量变化特征及其影响分析[J]. *亚热带水土保持*, 2009, 21(2): 9 – 13.  
Jiang Zhi, Kuang Mingsheng. Analysis on the characteristic of precipitation variability & effects in recent fifty-seven years in Chongqing area[J]. *Subtropical Soil and Water Conservation*, 2009, 21(2): 9 – 13. (in Chinese)
- 14 曾小凡, 翟建青, 姜彤, 等. 长江流域年降水量的空间特性和演变规律分析[J]. *河海大学学报: 自然科学版*, 2008, 36(6): 727 – 732.  
Zeng Xiaofan, Zhai Jianqing, Jiang Tong, et al. Spatial characteristics and evolutionary trends of annual precipitation in the Yangtze River basin[J]. *Journal of Hohai University: Natural Sciences*, 2008, 36(6): 727 – 732. (in Chinese)
- 15 张焱, 韩军青, 郭刚. 西晋黄土高原地区近 47 年降雨量的统计分析[J]. *干旱区资源与环境*, 2008, 22(1): 89 – 91.  
Zhang Yan, Han Junqing, Guo Gang. Statistical analysis on precipitation in Jinxibei region during the last 47 years[J]. *Journal of Arid Land Resources and Environment*, 2008, 22(1): 89 – 91. (in Chinese)
- 16 方积乾. 生存分析的概念与方法学[J]. *自然杂志*, 1988, 11(11): 826 – 831.
- 17 徐英, 骆福添. 生存分析中几种模型的研究概况[J]. *中国卫生统计*, 2006, 23(4): 364 – 366.
- 18 张娜, 屈忠义, 杨晓, 等. 贝叶斯模型在土壤转换函数中的应用与适应性评价[J]. *农业机械学报*, 2014, 45(2): 149 – 155.  
Zhang Na, Qu Zhongyi, Yang Xiao, et al. Application and adaptability evaluation of Bayesian model in soil transfer functions[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2014, 45(2): 149 – 155. (in Chinese)
- 19 Oakes D. Biometrika centenary: survival analysis[J]. *Biometrika*, 2001, 88(1): 99 – 142.
- 20 Tiwari G, Bang diwala S, Saraswat A, et al. Survival analysis: pedestrian risk exposure at signalized intersections[J]. *Transportation Research Part F: Traffic Psychology and Behaviour*, 2007, 10(2): 77 – 89.