

利用神经网络-缺省因子法及熵权法 分析气象因素对水面蒸发的影响

邓丽娟

(新疆水利水电科学研究院, 新疆 乌鲁木齐 830049)

摘要: 根据新疆车尔臣流域且末县气象站 2007 年非结冰期(4-9 月)日水面蒸发量及相关常规气象观测资料, 利用神经网络-缺省因子法及熵权法分析了各气象要素对水面蒸发的影响程度。结果表明:水面蒸发对温度与风速最为敏感。希望从气象因素角度出发, 为区域水资源优化调度提供参考。

关键词: 神经网络-缺省因子法; 熵权法; 水面蒸发; 车尔臣河流域

中图分类号: P333

文献标识码: A

文章编号: 1672-643X(2013)03-0043-03

Analysis of impact of meteorological factor on water evaporation based on neural network - default factor method and entropy method

DENG Lijuan

(Xinjiang Research Institute of Water Resources and Hydropower, Urumqi 830049, China)

Abstract: according to daily evaporation and meteorological data during non-glacial period of 2007 from Qiemo meteorological station in Che'erchen River basin of Xinjiang (April to September), the paper used neural network - the default factor method and entropy method to analyze the influence of meteorological factors on water evaporation. Results show that temperature and wind speed is the most sensitive factors on the influence of water surface evaporation. Hope from meteorological factors point of view, to provide reference for regional water resources optimal allocation.

Key words: neural network - default factor method; entropy method; water surface evaporation; Che'erchen river basin

0 前言

新疆车尔臣河流域地处塔里木盆地东南部, 紧邻全国第一大流动性沙漠——塔克拉玛干大沙漠。关于该区域影响日水面蒸发变化的敏感气象因子以及敏感程度的研究少之又少。本文以新疆车尔臣流域且末气象站 2007 年非结冰期(4-9 月)日水面蒸发量观测数据为例, 利用改进的 BP 神经网络技术建立了日水面蒸发模型, 描述了各气象因素对日水面蒸发的影响, 运用缺省因子法分析了各因子对日水面蒸发影响的敏感性, 并采用熵权法对影响因子进行了敏感性综合排序, 以检验神经网络缺省因子法的正确性, 这同时也完成了熵权法的自我检验。

1.1 神经网络模型

1.1.1 神经网络的结构体系和训练算法概述 人工神经网络是一个大规模并行分布式信息处理系

统, 具有一定的性能特点, 类似于人类大脑中的生物神经网络^[1]。神经网络的特点是它的架构, 其代表了节点之间的连接模式、连接权值确定方法和激活功能^[2]。在本研究中, 采用前馈式神经网络结构结合自适应学习速率动量梯度下降反向传播算法(GDX)进行模拟。

(1) 前馈神经网络。随着误差传播学习算法的问世, 前馈神经网络已经成功地应用在不同领域。这种网络架构是最简单的神经网络架构。在前馈神经网络中, 节点一般都安排在层上, 从第一个输入层开始, 并在最后一个输出层结束。它可以有几个隐藏层, 每一层又有一个或多个节点(BP 神经网络结构见图 1)。

(2) 训练算法。本文采用自适应学习速率动量梯度下降反向传播算法(Gradient descent with momentum and adaptive learning rate back propagation,

收稿日期: 2013-01-10; 修回日期: 2013-01-31

基金项目: 水利部公益性行业科研专项经费项目(201201037)

作者简介: 邓丽娟(1980-), 女, 新疆乌鲁木齐人, 工程师, 硕士研究生, 主要从事水利水电工程质量检测。

简称 GDX) 来对模型进行训练。

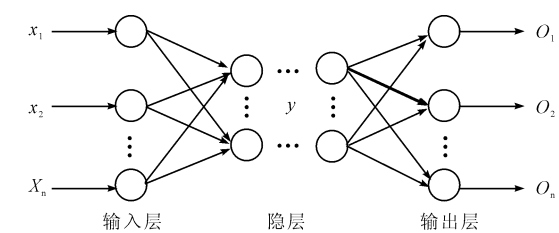


图 1 BP 神经网络结构

自适应学习速率动量梯度下降反向传播算法 (GDX): 在标准 BP 算法中, 学习速率在整个训练期间保持不变, 该算法的结果对于学习速率的正确设

置是非常敏感的。如果学习速率定的太高, 算法可能会振荡和不稳定。如果学习速率定的太小, 算法收敛时间将会太长。为了克服这个问题, GDX 算法结合自适应学习率与动量训练。自适应学习率试图保持尽可能大的学习步长, 同时保持学习的稳定。每个变量根据动量梯度下降进行调整。这就像一个低通滤波器, 动量允许网络忽略错误表面的小功能。该训练算法是最简单、最常见的训练网络方式之一。

1.1.2 BP 神经网络设计 在 Matlab R2011a 中, BP 神经网络设计包括网络的生成、训练及仿真可以通过调用一系列函数来实现, 见表 1。

表 1 BP 神经网络设计常用函数

函数名称	功能	常用格式
newff	生成一个 BP 神经网络	$\text{net} = \text{newff}(PR, [S_1, S_2, \dots, S_N], \{TF_1, TF_2, \dots, TF_N\}, BTF, BLF, PF)$
train	对 BP 网络进行训练	$[\text{net}, \text{tr}] = \text{train}(\text{net}, P, T)$
sim	对 BP 网络进行仿真	$a = \text{sim}(\text{net}, P)$

表 1 中, PR 为 $R \times 2$ 维矩阵, 表示输入层 R 个神经元的取值范围; 若神经网络为 N 层, 则 $[S_1, S_2, \dots, S_N]$ 中各元素分别表示各层神经元 (节点) 的数量; $\{TF_1, TF_2, \dots, TF_N\}$ 中各元素分别表示各层神经元所采用的传递函数, 本文中 BTF 为神经网络训练时所选定的训练函数; BLF 为选定的学习规则函数; PF 为误差函数; train 函数常用格式中的 tr 存储训练过程中的步数信息和误差信息; net 表示前面生成的神经网络对象; P 为网络输入样本矢量集; T 为对应的目标样本矢量集; sim 函数常用格式中的 net 表示前面生成并经过训练的神经网络^[3]。

1.2 熵权法

熵本身是一个热力学概念, 最先由 Shannon 引入信息论, 熵值反映了信息无序化程度, 其值越小, 系统无序度越小, 因此可用信息熵评价所获系统信息的有序度及其效用, 其计算步骤如下^[4]:

(1) 假定针对评价目标 (在本文中为水面蒸发量), 共采集了 m 个样本, 每个样本有 n 个影响因素, 构建判断矩阵 R :

$R = (x_{ij})_{m \times n}, (i = 1, 2, \dots, m; j = 1, 2, \dots, n)$ (1)

式中: x_{ij} 为第 i 个样本的第 j 个影响因素所对应的数值。

(2) 影响因素指标性质的判定。这里所指的指标性质是指影响因素与评价目标的正负相关性, 若影响因素与评价目标正相关, 则表明该影响因素为正指标, 反之则为负指标。

(3) 影响因素指标性质确定后, 将评价矩阵 R 进行归一化处理:

$$b_{ij} = \frac{x_{ij} - x_{\min}}{x_{\max} - x_{\min}} \quad (2)$$

式中: $x_{\max}, x_{\min}$ 分别为同一因素下不同样本中数值最大者和数值最小者。

(4) 根据熵值的定义, m 个样本 n 个影响因素, 则影响因素的熵值为:

$$H_j = \frac{(1 + b_{ij})}{-\ln m} \left(\sum_{i=1}^m f_{ij} \ln f_{ij} \right) \quad (i = 1, 2, \dots, m; j = 1, 2, \dots, n) \quad (3)$$

$$f_{ij} = \frac{(1 + b_{ij})}{\sum_{i=1}^m (1 + b_{ij})} \quad (4)$$

(5) 根据熵值的大小, 对影响因素的敏感性进行排序。熵值越大, 表明敏感性越大, 反之则越小。

2 模型计算

2.1 数据来源

本文以新疆车尔臣河流域且末气象站 2007 年 4 月至 9 月的非冰期日气象数据为基础, 采用 BP 神经网络模型与熵权法研究各气象因素对水面蒸发量的敏感程度。选取日均气温 ($x_1, ^\circ\text{C}$)、日最高气温 ($x_2, ^\circ\text{C}$)、日最低气温 ($x_3, ^\circ\text{C}$)、日均风速 ($x_4, \text{m/s}$)、日均水汽压 (x_5, hPa)、日均相对湿度 ($x_6, \%$) 与日照时数 (x_7, h) 这 7 因素为子因素, 即影响因素; 日水面蒸发量 (E, mm) 为母因素。在数据处理过程中, 为了保证模型计算的可靠性, 本文剔除了有日降水的气象数据, 最终有效建模样本数为 147 组。

2.2 输入、输出因子选取及训练样本预处理

针对研究区实际情况,建立了日水面蒸发 BP 网络模型,选取日均气温、日最高气温、日最低气温、日均风速、日均水汽压、日均相对湿度与日照时数作为输入因子,日水面蒸发量作为输出变量。气象数据来源于 2007 年 4 月至 9 月的日气象观测资料。所有相关数据均参与网络的训练与检验。考虑到前述各气象因子在数值上变化较大,且量纲也不同,故训练时首先采用归一方法将这些数据统一处理到 [0,1] 区间^[5]。

2.3 模型的训练及检验

为使所建模型具有一定的泛化能力,选取 2007 年 4 月至 8 月的 121 组样本作为日水面蒸发的训练样本。

研究表明,3 层 BP 网络可以逼近任何有理函数,本文选取 3 层 BP 网络,隐含层采取双曲正切 S 型函数 $f(x) = 1/(1 + e^{-x})$ 作为 BP 神经网络的传递函数,输出层采取线性函数,利用 MATLAB 软件编写程序,建立日水面蒸发模型。通过对模型输入、输出样本的训练,最终确定网络拓扑结构为 7: 15: 1。采用未参与训练的 26 组(9 月数据)输入、输出样本数据对所建立的模型进行检验(检验结果见图 2)。结果表明,该网络模型具有较高的精度,模型数据点均匀的分布在 $y = x$ 这条直线的两侧;日水面蒸发量平均检验绝对误差为 0.23 mm,相对误差绝对值为 3.7%,能有效表征综合气象因素条件下日水面蒸发的变化情况,满足日水面蒸发预报精度要求。

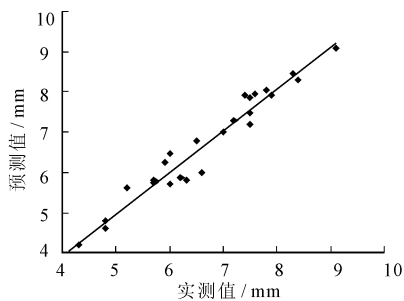


图2 日水面蒸发量神经网络模型检验

2.4 缺省因子法

本文将缺省因子法引入到 BP 神经网络模型计算中,通过对各影响因素的逐一缺省,进行 BP 神经网络模型的建立及检验,根据其检验误差与全因素模型检验误差的比值 R_i 大小来确定缺省因子的敏感程度^[6]。

$$R_i = RMSE_i / RMS_E \tag{5}$$

式中: R_i 为敏感性指数; $RMSE_i$ 为缺省第 i 个因子时模型检验误差; RMS_E 为全因子模型检验误差。若 $RMSE_i > RMSE_j$,表示日水面蒸发对第 i 个因子较第 j 个因子敏感。

2.5 敏感因子分析结果

根据前述方法,分别缺省因子 x_1 、 x_2 、 x_3 、 x_4 、 x_5 、 x_6 、 x_7 ,建立 6 因子的日水面蒸发 BP 神经网络模型。为使 6 因子模型与 7 因子模型具有可比性,在建模过程中采用的训练样本、检验样本均等同于 7 因子模型,同样采用改进的 BP 网络学习算法进行模拟。

通过缺省因子检验法分析比较,结果表明:6 因子模型的检验误差与 7 因子模型相比均有不同程度的增大($R > 1$),说明输入因子对日水面蒸发均有不同程度的影响。利用 BP 神经网络模型建立 6 因子的模型拟合方程。通过缺省因子法分析比较,得到各因素的敏感性指数(见表 2)。

表 2 缺省因子法检验结果

影响因子	平均相对误差	R	敏感性排序
全部因子	3.7		
x_1	10.4	2.81	1
x_2	9.8	2.65	2
x_3	9.5	2.57	3
x_4	8.4	2.27	5
x_5	4.3	1.15	7
x_6	4.5	1.22	6
x_7	9.1	2.46	4

由表 2 可知,温度要素(包括日均温度、日最高温度与日最低温度)对日水面蒸发的敏感性指数较大,其 R 值分别为 2.81、2.65 与 2.57;其次为日照时间和日均风速这两种因素,其 R 值分别为 2.46 和 2.27;再次为日均水汽压与日均相对湿度两种因素,其 R 值分别为 1.15 和 1.22。各气象因素对日水面蒸发的敏感性由大到小排序为: $x_1 > x_2 > x_3 > x_7 > x_4 > x_6 > x_5$ 。

2.6 熵权法的计算及排序

根据式(1~2)至式(1~4),求得非冰期各影响因素的熵值及排序(见表 3)。

表 3 熵值的计算及排序

气象条件	影响因子	x_1	x_2	x_3	x_4	x_5	x_6	x_7
非冰期	熵值	2.374	2.042	2.101	1.882	1.521	1.679	2.227
	排序	1	4	3	5	7	6	2

从图2中可以看,模型验证期径流模拟值与实测值总体变化趋势相近,在量上存在一定的偏差。验证期的 Nash - sutcliffe 效率系数为 0.80,径流相对误差为 11.6%。由率定期和验证期可以看出,模拟效率系数较高,误差较小,表明 SWAT 模型在天山西部山区是可以适用的,能够较好的反映其产流特性。

3 结 语

为能够更好研究气候及下垫面的变化对流域产流规律的影响,为其他相关方面的研究提供一定的指导依据,借助研究区 DEM 数据、土地利用数据、土壤数据、水文气象要素等基础数据,以 SWAT2009 模型为平台研究了 SWAT 模型在天山西部山区径流模拟过程中的适用性。

通过对比看出 SWAT 模型在天山西部山区是可以适用的,能够较好的反映其产流特性,率定期和验证期效率系数均为 0.80,径流相对误差分别为 8.6% 和 11.6%,径流模拟值与实测值变化趋势相近。产生误差的主要原因是山区降水与气温在垂直方向上分布的差异性及研究区实测水文气象站点较少,不能较好的反映降水与气温的时空分布规律,并且由于积雪、冰川融水的存在加大的径流模拟的难度,有待于借助精度更高、更好的数据源进行进一步研究。

(上接第 45 页)

由熵值的大小可以判定,非冰期水面蒸发影响因子的敏感性排序为: $x_1 > x_7 > x_3 > x_2 > x_4 > x_6 > x_5$ 。

3 结 语

(1) 基于神经网络 - 缺省因子法与熵权法的水面蒸发敏感因素研究结果表明:非冰期情况下,水面蒸发对温度与风速这二因素最为敏感。在今后该区域的水资源优化调度中,应当充分考虑这二种气象因素对水资源配置的影响,合理优化水资源调度方案。

(2) 神经网络 - 缺省因子法与熵权法分析的结果基本一致,尤其在主要敏感因素的判别上是完全一致的,这充分说明上述两种方法在水面蒸发敏感性分析中的可行性。同时,该方法也为类似问题的解决提供了参考。

参考文献:

- [1] 张明旭,赵海英. 晋江西流域降雨径流的 SWAT 模型模拟[J]. 通化师范学院学报,2008,29(4):74-76.
- [2] 冯国章. 新疆内陆河流水文干旱特征浅析[J]. 干旱地区农业研究,1994,12(2):70-78.
- [3] 姜卉芳,姜毅,陈亮. 新疆河流径流模拟[J]. 新疆农业大学学报,1998,21(3):176-183.
- [4] Philip W G, Manuel R R, Colleen H G, et al. The soil and water assessment tool: Historical, development, applications, and future research directions [J]. Trans ASAE, 2007,50(4):1211-1250.
- [5] 李峰,胡铁松,黄金金. SWAT 模型的原理、结构及其应用研究[J]. 中国农村水利水电,2008(3):24-28.
- [6] 施雅风,刘潮海,王宗太,等. 简明中国冰川目录[M]. 上海:上海科学普及出版社,2005:41-48.
- [7] 黄清华,张万昌. SWAT 模型参数敏感性分析及应用[J]. 干旱区地理,2010,33(1):8-15.
- [8] Nash J E, Sutcliffe J V. Rive flow forecasting through conceptual models, Part1 - a discussion of principles [J]. Journal of Hydrology, 1970,10(3):282-290.
- [9] 姚苏红,朱仲元,张圣微. 基于 SWAT 模型的内蒙古闪电河流域径流模拟研究[J]. 干旱区资源与环境,2013,27(1):175-180.
- [10] 齐贞,杜丽平,刘晓冰,等. SWAT 模型中气象数据库和土壤数据库的构建方法[J]. 河南科学,2011,29(12):1458-1463.
- [11] 张小咏,李佳,杨艳昭,等. 基于 SWAT 模型的长江源区径流模拟[J]. 西北林学院学报,2010,27(5):38-44.

参考文献:

- [1] 陈玉芳. BP 神经网络的算法改进及应用研究[D]. 成都:电子科技大学,2004.
- [2] 杨旭,冯国章. 基于 BP 网络的中长期水文预报精度影响分析[J]. 东北水利水电,2003,21(2):1-3.
- [3] 魏铁军,王永成,刘东. BP 网络在三江平原井灌区地下水埋深预测中的应用[J]. 黑龙江水专学报,2009,36(1):26-30.
- [4] 韩丽伟,付强,刘东,等. 三江平原地下水承载能力综合评价模型的构建及其应用[J]. 水土保持研究,2010,17(2):182-185.
- [5] 张德二. 台安实验区“三水”转化分析及水资源计算研究[J]. 水文,1992,12(5):26-31.
- [6] 于国强,李占斌,张霞,等. 土壤水盐动态的 BP 神经网络模型及灰色关联分析[J]. 农业工程学报,2009,25(11):141-147.