

何南腾,邹嘉南,郭文弟,等.阿克苏地区日蒸散量估算方法研究[J].气象研究与应用,2022,43(2):35–40.

He Nanteng, Zou Jianan, Guo Wendi, et al. Study on estimation method of daily evapotranspiration in Aksu Prefecture [J]. Journal of Meteorological Research and Application, 2022, 43(2): 35–40.

阿克苏地区日蒸散量估算方法研究

何南腾¹, 邹嘉南^{1*}, 郭文弟², 周笑迁¹, 狄迪¹

(1.南京信息工程大学中国气象局气溶胶与云降水重点开放实验室,南京 210044;

2.中国地质调查局西安矿产资源调查中心, 西安 710100)

摘要:采用阿克苏地区2006—2015年日蒸散量数据和气象数据,对几种不同的日蒸散量估算方法进行评估与修正。结果表明,修正后的彭曼公式在阿克苏地区的适用性、准确性有所提高,均方根误差值由11.78减小到8.80;BP神经网络比GRNN神经网络估算效果好,前者Nash-Sutcliffe系数为-0.09,RMSE值为3.27,后者Nash-Sutcliffe系数为-0.44,RMSE值为3.34。研究成果可为极端干旱区的棉花等作物节水种植提供参考。

关键词: 日蒸散量;Penman-Monteith公式;GRNN;BP

中图分类号:P426.2

文献标识码:A

doi: 10.19849/j.cnki.CN45-1356/P.2022.2.06

引言

蒸散发是陆面水循环中重要的水文过程之一,也是联系植物碳交换和水分利用的关键生态过程^[1]。作为大气和地表之间的联系,蒸散量作为水文参量同时出现在地表水量和能量平衡方程中,是陆面过程数值模拟研究中不可缺少的重要边界条件。此外,Rosenberg等^[2]认为70%降落到地球表面的降水是通过蒸发或蒸散作用回到大气中的,在干旱区这个比例可达90%。因此,它对于了解作物蒸散发过程规律、水资源管理以及提高水资源利用效率以及干旱评定等^[3-5]有重要意义。近年来,Santos等结合遥感技术与土壤水量平衡模型,估算了西班牙西南地区棉田2004—2005年蒸散量的季节变化^[6];Nobuhiro等通过能量平衡波文比方法估算了常绿阔叶林蒸散发^[1];Zhao等基于多年涡度相关和小气候梯度系统观测数据库,使用了土壤水量平衡法、波文比法、Penman-Monteith(PM)等6种方法估算并比较绿洲农田蒸散量,发现Penman-Monteith(PM)结果最符合绿洲实际^[7]。由于受不同地区气象和地形、地

貌时空异质性的影响,不同参考作物蒸散量计算公式的适用性各不相同。Penman-Monteith(PM)公式作为FAO专家组唯一推荐使用的参考作物蒸散量计算公式,至今多种参考作物蒸散量计算方法的适用性评价已成为当今的热点问题。然而,参考作物蒸散量计算方法的普适性在极端干旱区却鲜有研究。本文依据阿克苏地区2006—2015年日蒸散量数据与气象资料,对Penman-Monteith(PM)公式进行进一步的优化与适用性分析,以期推荐一个更加精确的参考作物蒸散量计算方法,这对于了解极端干旱区棉田等作物种植以及可持续发展有长远的意义。

此外,由于气象数据具备周期性强、时空性、多维多尺度、各属性相互影响等特征,仅用传统方法对气象数据进行分析和处理会遇到诸多困难^[8]。目前,由于遥感卫星成本低,基于遥感技术和模型估算蒸散发逐渐成为一种经济实用的手段^[9-10]。另外,神经网络作为一种人工智能算法,有别于传统方法,特别是应用在处理非线性问题上有着良好效果。此前,也有专家利用神经网络模型模拟膜下滴灌玉米逐日蒸

收稿日期: 2021-12-14

基金项目: 江苏省高等学校自然科学研究项目(21KJB170002)

作者简介: 何南腾(2001—),男,学士,研究方向为颗粒物浓度反演。E-mail: 201983300656@nuist.edu.cn

* 通讯作者: 邹嘉南(1990—),男,博士,讲师,研究方向为气溶胶物理化学。E-mail: Zoujn16@nuist.edu.cn

散量^[11]。本研究利用了反向传播神经网络(Back Propagation Neural Network, 下文统一称为BP神经网络)、广义回归神经网络(General Regression Neural Network, 下文统一称为GRNN神经网络)两种神经网络对日蒸散量数据进行拟合,希望通过此寻找出一种较为准确的日蒸散量估算方法。

1 资料与方法

1.1 研究区概况

阿克苏地区位于新疆中部地区、天山山脉中段南麓、塔里木盆地的北缘,78°03'~78°39'E,39°30'~42°41'N,面积约为1325×10⁴hm²。综合观测场位于阿克苏站区内,地理坐标:80°45'E,40°37'N,海拔1028m,2006年建立,长方形,大小150m×90m,可以满足100a尺度的采样要求。

1.2 数据来源

本研究使用了阿克苏站中综合观测场农田生态系统2006—2015年有测量记录的日蒸散量数据(100个样本,水量平衡法,单位mm·d⁻¹,来自国家生态科学数据中心资源共享服务平台(<http://www.cnern.org.cn/>)),在本研究中用于日蒸散量估算值的对比与验证。另外,使用了相匹配日期的阿克苏站的气象资料,包括日平均气温、日最高气温、最低气温、日照时数、平均相对湿度、平均水汽压、平均风速等气候要素观测数据和经纬度、海拔高度等地理数据。由于国家生态科学数据中心资源共享服务平台中没有阿克苏站以上需求的日尺度气象资料,所以,本研究需要的气象资料引用了美国国家海洋和大气管理局(NOAA)国家环境信息中心(NCEI)相应的日尺度数据。

对于Penman-Monteith(PM)公式和GRNN神经网络、BP神经网络,为方便对估算方法的评估,本研究采用同样的样本进行对比。其中,PM公式包括了所有样本的计算,GRNN神经网络、BP神经网络验证集中的30个样本在总样本中随机选出。

1.3 计算方法

1.3.1 Penman-Monteith(PM)公式

在蒸散量计算中,彭曼法(Penman-Monteith)基于能量平衡和水汽扩散理论,不受作物种类、土壤类型和水汽条件等因素的限制,仅考虑气候要素的影响,其计算得出的潜在蒸散量与实测值最为接近,因此被广泛应用到区域干湿状况的评价研究中^[12]。表达式如下:

$$PE = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T_{mean} + 273} u_2(e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \quad (1)$$

式中,PE为潜在蒸散量,Δ表示空气气温为T时的饱和水气压斜率,计算方式如(2),本次研究中T取日平均气温,T_{mean}表示日平均温度(℃),e_s表示饱和水气压(kPa),e_a表示实际水气压(kPa);R_n为地表净辐射,等于收入短波辐射R_{ns}和支出的净长波辐射R_{nl}之差;G为土壤热通量,一天至十天的时间尺度,其参考草地的土壤热容量相当小,可以忽略不计。γ为干湿表常数,计算方式如(4),其中P为大气压;u₂为2m高风速,u_z表示10m高处风速,z表示高度,取10m。

$$\Delta = \frac{4098 \times \left[0.6108 \times \exp \left(\frac{17.27T}{T+237.3} \right) \right]}{(T+237.3)^2} \quad (2)$$

$$R_n = R_{ns} - R_{nl} \quad (3)$$

$$\gamma = 0.665 \times 10^{-3}P \quad (4)$$

$$u_2 = u_z \frac{4.87}{\ln(67.8z - 5.42)} \quad (5)$$

1.3.2 方法校准

在日蒸散量计算中,主要对地表净辐射项的参数做出调整。目前对于地表净辐射有诸多经验方法已被广泛用于估计世界各地的净辐射,然而,这些经验方法需要特定地点的参数校准。Xiao等^[13]量化了地表反照率的时空变化以更好地估计中国地区的净短波辐射,并基于原位观测重新校准了FAO-56 PM方法的参数以更好地估计净长波辐射。此方法也可用于彭曼法(Penman-Monteith)对日蒸散量的计算。地表净辐射等于收入短波辐射R_{ns}和支出的净长波辐射R_{nl}之差。对于短波辐射有:

$$R_{ns} = (1-\alpha)R_s \quad (6)$$

其中,R_s为太阳辐射,对于阿克苏地区(40°37'N),校准后有:

$$R_{ns}' = (1-\alpha')R_s \quad (7)$$

其中α'随月份变化如表1所示

此外,对于净长波辐射R_{nl}有

$$R_{nl} = \sigma \frac{T_{max}^4 + T_{min}^4}{2} (a + b \sqrt{e_a}) (1.35 \frac{R_s}{R_{s0}} - 0.35) \quad (8)$$

其中,σ=4.903×10⁻⁹为斯蒂芬-玻尔兹曼常数(MK·K⁻⁴·m⁻²·d⁻¹)。T_{max}⁴和T_{min}⁴分别为最高和最低绝对气温(K=℃+273.16);e_a为实际蒸汽压;R_s为太阳辐射,R_{s0}为晴空辐射;a、b为经验回归系数(无量纲)。

表 1 阿克苏地区地表反照率随月份变化

月份	反照率
1月	$\alpha=0.324$
2月	$\alpha=0.287$
3月	$\alpha=0.198$
4月	$\alpha=0.186$
5月	$\alpha=0.182$
6月	$\alpha=0.182$
7月	$\alpha=0.183$
8月	$\alpha=0.185$
9月	$\alpha=0.186$
10月	$\alpha=0.2195$
11月	$\alpha=0.272$
12月	$\alpha=0.318$
年平均	$\alpha=0.224$

对于阿克苏地区 $a=0.50-0.002\times lat=0.419$, $b=0.25+0.002\times lat=-0.169$ 。所以对于修正后的地表净辐射项可列为:

$$R_n' = R_{ns}' - R_{nl}' = (1-\alpha') R_s - \sigma \frac{T_{max}^4 + T_{min}^4}{2} (0.419 - 0.169\sqrt{e_a}) (1.35 \frac{R_s}{R_{s0}} - 0.35) \quad (9)$$

1.3.3 神经网络

(1) GRNN 神经网络

GRNN 具备很强的非线性映射能力、柔性结构以及较好的容错性, 适用于解决非线性问题。GRNN 在逼近能力和学习速度上也有很强的优势, 网络最后收敛于样本量积聚较多的优化回归面, 另外在样本数据较少时, 也有一个较好的预测效果。

GRNN 神经网络由四层构成^[14]。如图 1 所示, 分别为输入层、模式层、求和层和输出层。假设样本特征集为: $\{trx_1, trx_2, \dots, trx_m\}$, 其中 $trx_i=[x_1, x_2, \dots, x_n]$ 。标签集为: $\{try_1, try_2, \dots, try_m\}$, 其中 $try_i=[y_1, y_2, \dots, y_k]$ 。

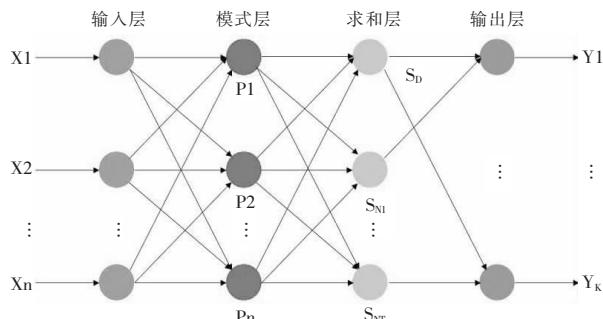


图 1 GRNN 神经网络结构图

输入测试样本, 节点个数等于样本的特征维度。模式层中节点个数等于训练样本的个数, 第 i 个测试样本 trx_i 与第 j 个测试样本 trx_j 之间的 Gauss 函数取值为:

$$Gauss(trx_i - trx_j) = e^{-\frac{\|trx_i - trx_j\|}{2\delta^2}} \quad (10)$$

求和层节点个数等于输出样本维度加 $1 \times (k+1)$, 求和层的输出为两部分, 第一个节点输出模式层输出的算术和, 其余 k 个节点的输出为模式输出的加权和。假设对于测试样本 tex , 模式层的输出为 $\{g_1, g_2, \dots, g_m\}$ 。

求和层第一个节点的输出为:

$$S_D = \sum_{i=1}^m g_i \quad (11)$$

其余 k 个节点的输出为:

$$S_{Nj} = \sum_{i=1}^m y_{ij} g_i, j=1, 2, \dots, k \quad (12)$$

其中加权系数 y_{ij} 为第 j 个模式层节点对应的训练样本的标签的第 j 个元素。

输出层节点个数等于标签向量的维度, 每个节点的输出等于对应的求和层输出与求和层第一个节点输出相除。

(2) BP 神经网络

本次研究还采用了 BP 神经网络(如图 2), 其含有输入层、隐含层和输出层三个部分。BP 神经网络使用梯度下降算法^[15]。假设输入层神经元为 $P=[p_1, p_2, \dots, p_i]$, 隐含层神经元为 $S=[s_1, s_2, \dots, s_k]$, 输出层神经元为 $Y=[y_1, y_2, \dots, y_j]$ 。另外, $w_{k,i}^1$ 表示输入层第 i 个神经元连接隐含层第 k 个神经元的权值, $w_{j,k}^2$ 表示隐含层第 k 个神经元连接输出层第 j 个神经元的权值。隐含层的激发函数为 f^1 , 输出层的激发函数为 f^2 , $b_{s_k}^1$ 表示隐含层神经元阈值, $b_{y_j}^2$ 表示输出层神经元阈值, 其中, $w_{k,i}^1, w_{j,k}^2, b_{s_k}^1, b_{y_j}^2 \in (-1, 1)$ 。

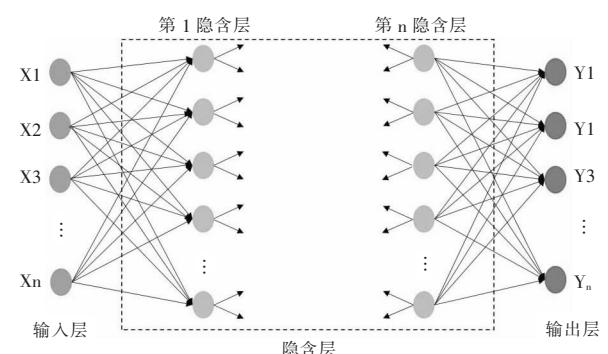


图 2 BP 神经网络结构图

输入层各神经元与隐含层各神经由相应的权值连接, 隐含层的第一个神经元 x_1 , 加权求和从输入层每一个神经元处得到输出值 $p_1 \times w_{1,1}^1 + p_2 \times w_{1,2}^1 + \dots + p_i \times w_{1,i}^1$, 加上阈值 $b_{s_i}^1$, 通过激发函数 f_1 , 得到该神经元的输出值:

$$x_1^1 = f^1((p_1 \times w_{1,1}^1 + p_2 \times w_{1,2}^1 + \dots + p_i \times w_{1,i}^1) + b_{s_i}^1) \quad (13)$$

然后, 输出层第一个神经元 y_1 接收隐含层每一个神经元输出值, 并加权求和得 $x_1 \times w_{1,1}^1 + x_2 \times w_{1,2}^1 + \dots + x_k \times w_{1,k}^1$, 加上阈值 $b_{y_1}^2$, 通过激发函数 f^2 , 得到输出层该神经元的输出值:

$$y_1^1 = f^2((x_1 \times w_{1,1}^1 + x_2 \times w_{1,2}^1 + \dots + x_k \times w_{1,k}^1) + b_{y_1}^2) \quad (14)$$

(3) 评估方法

拟合优度指标通常用于评价改进方法的性能, 通常对于此类研究采用 Nash–Sutcliffe 系数(NSE)、平均偏差误差(MBE)和均方根误差(RMSE)。表达式如下^[13]:

$$NSE = 1 - \frac{\sum_{i=1}^n (R_e^i - R_0^i)^2}{\sum_{i=1}^n (R_e^i - R_{ave})^2} \quad (15)$$

$$MBE = \frac{\sum_{i=1}^n (R_e^i - R_0^i)}{n} \quad (16)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (R_e^i - R_0^i)^2}{n}} \quad (17)$$

式中 n 为观测次数; R_e 为估计数据; R_0 为观测

数据; R_{ave} 为观测数据的平均值。Nash–Sutcliffe 系数取值范围为 $-\infty \sim 1$, 其值越接近 1, 均方根误差值越小, 模型性能越好。平均偏差误差为正值表示高估了结果, 其为负值则表示低估了结果。

2 结果与分析

2.1 自相关性分析

表 2 为日蒸散量与日最高温度、日最低温度、露点温度、风速、日照时长、降雨量等六个气象要素相关系数的统计。其中, p 值表示原假设为真时得到的样本观察结果或者更极端结果出现的概率。* 表示有统计学差异($p < 0.05$), ** 表示差异显著($p < 0.01$), 显著性差异。它是统计学上对数据差异性的评价。通常情况下, 实验结果达到 0.05 水平或 0.01 水平, 才可以说数据间具有差异显著或是极显著。

从表 2 中可以看出, 日蒸散量与其他六个气象要素量均呈正相关, 与日最高温度、日最低温度和日平均露点温度相关系数分别为 0.3142、0.4338、0.4624, 且均达到 0.01 显著性水平。神经网络对结果的预测如果是建立在具有高度相互影响的因素上, 其预测模型和结果会相对可靠, 综合彭曼公式和相关系数考虑, 本研究利用日最高温度、日最低温度、日平均露点温度、风速、日照时长、纬度、海拔、日序作为自变量, 日蒸散量作为因变量, 建立神经网络模型。

表 2 自相关性分析

	PE	Tmax	Tmin	Tdew	WS	n	P
PE	1	0.3142**	0.4338**	0.4624**	0.0710	0.0314	0.1224
Tmax		1	0.7034**	0.6294**	0.0242	0.4869**	-0.1598
Tmin			1	0.8086**	0.3676**	0.0011	0.0257
Tdew				1	-0.0158	-0.0245	0.1394
WS					1	-0.1380	0.0799
n						1	-0.2679**
P							1

(表中 PE 表示日蒸散量, Tmax 表示日最高气温, Tmin 表示日最低气温, Tdew 表示日平均露点温度, WS 表示日平均风速, n 表示日照时长, P 表示日降水量)

2.2 修正后彭曼法(Penman–Monteith)

如图 3 所示, 修正过后彭曼公式对于日蒸散量的估算比修正前要小(MBE 值由 8.4428 减小到 5.6030), 且更接近于真实测量值。从拟合效果来看, Nash–Sutcliffe 系数修正前与修正后分别为 0.0328

与 0.0132, 能说明结果接近观测值的平均值水平, 总体结果可信。同时, 修正前后的均方根误差分别为 11.7830 和 8.7981。所以总的来说, 修正后彭曼公式效果优于修正前, 对阿克苏地区日蒸散量计算的适用性和准确性相比修正前有所增加。

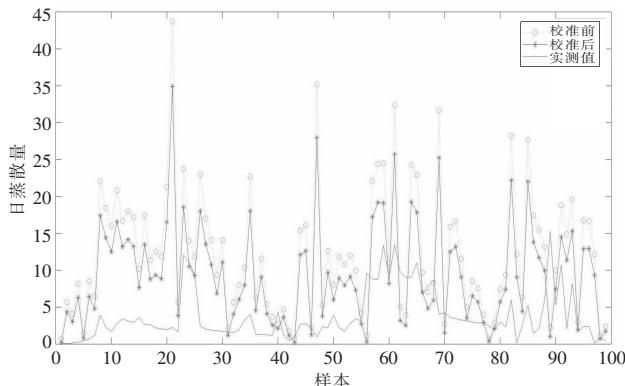


图3 修正前后彭曼公式拟合结果对比

2.3 GRNN 与 BP 神经网络估算

此次研究使用了一百个样本,分成 70 个训练集与 30 个测试集,使用了日最高温度、日最低温度、露点温度、风速、日照时长、纬度、海拔、日序等八个变量作为输入值,日蒸散量作为输出值。将 GRNN 神经网络拟合结果与实际测量值进行评估,其 Nash-Sutcliffe 系数为 -0.4410, 均方根误差为 3.3431, 说明结果接近观测值的平均值水平, 总体结果可信。另外, 测试集中仍有少部分预测值(4 个)与测量值相比有较大误差。这涉及到神经网络的泛化性问题, 此次研究的训练样本数量较少, 导致有少量较大误差的预测值产生。

对于 BP 神经网络, 为取消样本数量对拟合效果的误差, 同样分成 70 个训练集与 30 个测试集, 使用了日最高温度、日最低温度、露点温度、风速、日照时长、纬度、海拔、日序等八个变量作为输入值, 日蒸散量作为输出值。估算结果与实际值的 Nash-Sutcliffe 系数为 -0.0883, 均方根误差为 3.2659, 总体结果可信。其中仍有少量的预测值与测量值相比有较大误差。另外可以看出, BP 神经网络较 GRNN 神经网络对日蒸散量估算有更好的效果。

经过对比, BP 神经网络对于日蒸散量的估算效果优于 GRNN 神经网络。所以对于 BP 神经网络, 本研究做了进一步的评估。图 4 显示神经网络输出对训练、验证和测试集的目标回归图。四幅图的 $y=x$ 线表示目标输出, 不同颜色的实线表示神经网络输出。为了完美的配合, 数据应该沿着 45 度的线下降。可以看出训练集与验证集输出值贴近目标值, 且回归系数分别达到 0.94244 和 0.94606。同时, 测试集回归系数为 0.6113, 效果不如训练集与验证集, 本研究认为是由数据量偏少造成。另外, 总体的输出值与目标值也贴近且回归系数达到了 0.90871, 证明 BP 神经网络对于日蒸散量的估算问题上, 有一个很好的拟合效果。

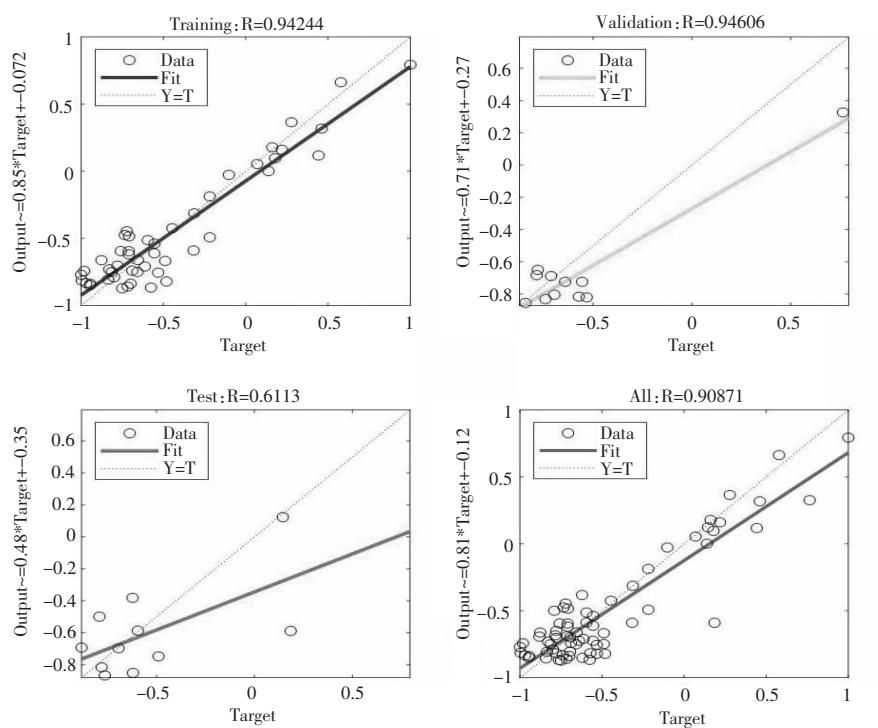


图4 BP 神经网络回归各集预测与实际值回归图

3 结论

本次研究对彭曼公式中地表净辐射项做了一个修正,使其更好的服务于阿克苏地区,此外对于日蒸散量的计算方法本身,采用了GRNN和BP两种机器学习方法进行了拟合计算,发现BP神经网络在解决此类问题上有较高的准确性。这一研究对阿克苏地区的日蒸散量估算、了解极端干旱区的棉田等作物种植有长远的意义。

此外,神经网络自身拟合准确性还与数据量有关,阿克苏站观测数据量偏少是本研究一大局限性,所以后续仍需有大量的观测数据投入到神经网络模型中。

致谢: 本文数据来源于国家生态科学数据中心资源共享服务平台(<http://www.cnern.org.cn/>),感谢此网站提供的阿克苏站日蒸散量观测数据。

参考文献:

- [1] 赵文智,吉喜斌,刘鹤.蒸散发观测研究进展及绿洲蒸散研究展望[J].干旱区研究,2011,28(3):463–470.
- [2] Rosenberg N J, Blad B L, Verma S B. Microclimate –the Biological Environment of Plants[M]. New York: John Wiley&Sons, 1983.
- [3] 吴初梅,何鹏,戴平风.蒸发量季节变化特点与干旱发生关系的初步分析[J].气象研究与应用,2008(3):8–11.
- [4] 莫新,李英梅.干旱指标的探讨[J].广西气象,1991(1):25–27.
- [5] 卢海新,朱天赐.厦门干旱成因和防旱抗旱对策[J].气象研究与应用,2010,31(3):24–26.
- [6] C. Santos et al. Integrating satellite-based evapotranspiration with simulation models for irrigation management at the scheme level [J]. Irrigation Science, 2008, 26 (3): 277–288.(重复出现)
- [7] Zhao W Z, Liu B, Zhang Z H. Water Requirements of Maize in the Middle Heihe River Basin, China [J]. Agricultural Water Management, 2009, 97(2): 215–223
- [8] 彭昱忠,王谦,元昌安,等.数据挖掘技术在气象预报研究中的应用[J].干旱气象,2015,33(1):19–27.
- [9] 郭二旺,郭乙霏,张凌杰,等.基于SEBAL模型和Landsat-8遥感数据的农田蒸散发估算[J].灌溉排水学报,2019,38(1):83–89.
- [10] 吴凡,陈植华,胡成.基于METRIC模型的江汉平原蒸散量估算适应性分析[J].安全与环境工程,2021,28(2):149–158.
- [11] 陈志君,朱振闯,孙仕军,等.Stacking集成模型模拟膜下滴灌玉米逐日蒸散量和作物系数[J].农业工程学报,2021,37(5):95–104.
- [12] 陈超,庞艳梅,潘学标,等.四川地区参考作物蒸散量的变化特征及气候影响因素分析[J].中国农业气象,2011,32(1):35–40.
- [13] Xiao M Z, Kong D D. Improvement in the Estimation of Daily Net Surface Radiation in China [J]. Journal of Irrigation and Drainage Engineering, 2021, 147(3) .
- [14] Chasset P O. GRNN: General regression neural network [J]. Revue De Physique Appliquée, 2013(6):1321–1325.
- [15] McClelland J L, Kawamoto A H. Mechanisms of Sentence Processing: Assigning Roles to Constituents of Sentences [M]. Cambridge: MITPress, 1986.

Study on estimation method of daily evapotranspiration in Aksu Prefecture

He Nanteng¹, Zou Jianan^{1*}, Guo Wendi², Zhou Xiaoqian¹, Di Di¹

(1.Key Laboratory for Aerosol–Cloud–Precipitation of China Meteorological Administration, Nanjing University of Information Science and Technology, Nanjing 210044, China; 2. Xi'an Center of Mineral Resources Survey, China Geological Survey, Xi'an 710100, China)

Abstract: Based on the daily evapotranspiration data and meteorological data of Aksu from 2006 to 2015, several different estimation methods of daily evapotranspiration were evaluated and modified. The applicability and accuracy of the modified Penman–Monteith (PM) formula in Aksu are improved, and the root mean square error is reduced from 11.78 to 8.80. The BP neural network (Back Propagation Neural Network) has better estimation effect than the GRNN (General Regression Neural Network). The former has a Nash–Sutcliffe coefficient of -0.09 and an RMSE value of 3.27, while the latter has a Nash–Sutcliffe coefficient of -0.44 and an RMSE value of 3.34. The research results can provide references for water-saving planting of cotton and other crops in extremely arid regions.

Key words: daily evapotranspiration; Penman–Monteith formula; GRNN; BP