

一种基于支持向量机回归的旋转遮光带日射表散射辐照度修正新算法

崇 伟^{1,2} 沙奕卓^{1,3} 行鸿彦^{1,2*} 吕文华^{1,3}

¹ 南京信息工程大学江苏省气象探测与信息处理重点实验室, 江苏 南京 210044
² 南京信息工程大学电子与信息工程学院, 江苏 南京 210044
³ 中国气象局气象探测中心, 北京 100081

摘要 通过比较旋转遮光带日射表(RSP)和参考标准表所测散射辐照度之间的数值差异,分析了太阳总辐照度、环境温度、相对湿度和太阳光谱等气象要素对 RSP 散射辐照度测量误差的影响关系,提出了一种修正 RSP 散射辐照度测量值的新算法。该算法从支持向量机回归预测角度,建立了对 RSP 散射误差修正值的一次预测模型,然后根据误差修正值最优预测模型推导出 RSP 散射辐照度修正算法模型。利用该算法对美国国家太阳辐射研究实验室和劳里观测站采集的 RSP 散射辐照度数据进行修正,修正后两观测站数据的平均偏差和均方根误差分别降低到 -0.2 W/m^2 , 3.3 W/m^2 和 1.9 W/m^2 , 8.5 W/m^2 , 显示算法具有良好的修正性能和适用性。该算法能够有效避免 Vignola 算法中存在的欠修正和 Vignola and Augustyn(VA)算法中存在的过修正现象。

关键词 测量;修正;支持向量机回归;旋转遮光带日射表;散射辐照度

中图分类号 O432; P413.2 文献标识码 A doi: 10.3788/AOS201232.0112001

A New Correction Algorithm for Diffuse Irradiance Measured with Rotating Shadow-Band Pyranometer Based on Support Vector Regression

Chong Wei^{1,2} Sha Yizhuo^{1,3} Xing Hongyan^{1,2} Lü Wenhua^{1,3}

¹ Jiangsu Key Laboratory of Meteorological Observation and Information Processing,
Nanjing University of Information Science and Technology, Nanjing, Jiangsu 210044, China

² College of Electronic and Information Engineering, Nanjing University of Information Science and Technology,
Nanjing, Jiangsu 210044, China

³ Meteorological Observation Center of China Meteorological Administration, Beijing 100081, China

Abstract Through comparison of the numeric difference between diffuse irradiance measured with rotating shadow-band pyranometer (RSP) and standard reference pyranometer, the influence of meteorological elements like global irradiance, ambient temperature, relative humidity and solar spectrum, etc. on measurement errors of diffuse irradiance measured with RSP is analyzed, based on the analysis, a new algorithm of correcting diffuse irradiance measured with RSP is proposed. The algorithm constructs an error modifier prediction model based on support vector regression (SVR) theory first, and then diffuse correction algorithm of RSP can be derived from mentioned prediction model. Applying the correction algorithm to the diffuse irradiance data measured by Solar Radiation Research Laboratory and Lowry Range Solar Station, mean deviation and root mean square error of the two stations' data after correction decrease to -0.2 W/m^2 , 3.3 W/m^2 and 1.9 W/m^2 , 8.5 W/m^2 , which indicates that the algorithm is effective and applicative. Meanwhile, compared with Vignola correction algorithm and Vignola and Augustyn (VA)

收稿日期: 2011-05-09; 收到修改稿日期: 2011-07-31

基金项目: 国家自然科学基金(61072133), 江苏省科技支撑计划(BE2008139)和江苏“六大人才高峰”计划资助课题。

作者简介: 崇 伟(1987—), 男, 硕士研究生, 主要从事计量检定及数据处理等方面的研究。E-mail: ensonic@126.com

导师简介: 沙奕卓(1966—), 男, 硕士, 研究员, 主要从事气象探测技术及仪器标校等方面的研究。

E-mail: syz@cma.gov.cn

* 通信联系人。E-mail: xinghy@nuist.edu.cn

correction algorithm, the algorithm presented can avoid respective undercorrection and overcorrection of similar correction algorithm effectively.

Key words measurement; correction; support vector machine regression; rotating shadow-band pyranometer; diffuse irradiance

OCIS codes 120.3930; 120.5630; 120.5820

1 引 言

太阳辐射是地球-大气系统的根本动力,是气候学研究的重要对象。太阳辐射数据可广泛用于地球-大气系统能量收支变化研究、大气成分特性分析、太阳能资源开发利用^[1]和农业气象服务等领域,具有重要的科研价值和现实意义。传统用于太阳辐照度测量的是高精度热电型日射表,但由于其价格昂贵,维护复杂,不利于大范围商业部署。目前,适宜广泛应用的首选价格低廉的旋转遮光带日射表^[2](RSP),它是基于光电半导体光伏效应制成的复合型日射表,可同时测量太阳总辐照度、散射辐照度,并以此计算出直射辐照度,能够有效应用于太阳辐射测量的实际业务中。

RSP 探测器是光敏半导体器件,其响应度会受到测量环境变化的影响,进而导致辐照度测量误差,其中散射辐照度测量值与本底标准相差最大^[3],最需修正。国内外研究者对改善光电型日射表的测量精度做了大量研究工作,Michalsky 等^[4]为了减小光电型日射表测量太阳总辐照度的误差,提出了一种基于改进余弦响应精度的太阳总辐照度测量方法;李照洲等^[5]对硅陷阱探测器辐射计的温度特性进行了系统研究,提出了加入精密温控系统以减小温度影响、提高精度的方法;King 等^[6,7]研究了光谱和光线入射角对光电模块及太阳辐射传感器的影响,并以此改善了光电型日射表的测量准确度;Vignola^[8]研究发现,光电型日射表测量的散射辐照度受气象要素影响十分严重,为此,他提出了 Vignola 修正算法,针对太阳总辐照度、环境温度和相对湿度修正光电型日射表测量散射的误差,该算法在散射辐照度较低时性能还有待提高;之后 Vignola 等^[3,9]综合已有研究成果,提出了提高 RSP 测量准确度的一整套算法,其中包括针对太阳总辐照度修正 RSP 散射辐照度的修正算法(VA 算法),该算法在实际运用中存在过修正现象。因此,目前的问题在于如何综合多因素影响,针对多影响要素提出 RSP 散射辐照度修正算法,改善已有算法的修正效果。

支持向量机^[10](SVM)是一种基于机构化风险最小原则和统计学 Vapnik-Chervonenks 维(VC)的

机器学习方法,它借助核函数展开和非线性映射,通过在高维空间中求解凸约束条件下的二次回归规划问题,寻找最优回归超平面,这在小样本观测数据情况下寻找规律、拟合预测十分有效。

本文在前人研究成果的基础上,从 RSP 与参考标准表(Eppley 8-48 热电型日射表)各自测得的散射辐照度数值差异入手,分析了太阳总辐照度、环境温度、相对湿度、天顶角和空气质量对 RSP 散射测量误差形成的影响关系,运用支持向量机回归理论(SVR),引入误差修正值作为回归对象,以造成误差的气象要素作为支持因子,构建误差修正值预测模型,根据误差修正值最优预测模型修正 RSP 所测散射辐照度,由此建立起一套基于支持向量机回归的 RSP 散射辐照度修正算法(SVR-DC 算法)。

2 算法理论基础

2.1 SVR 理论

SVR 是 SVM 在函数估计方面的重要应用,它能够整合多因素对预测量的综合影响,拟合出最优的函数模型。设训练集 $T = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_l, y_l)\}$, 其中 $\mathbf{x}_i \in \mathbf{X} = \mathbf{R}^n$ 是输入空间, $y_i \in \mathbf{Y} = \mathbf{R}$ 是输出空间, $i = 1, 2, \dots, l$ 。SVR 问题即要寻找输入空间到输出空间的一个映射 $f: \mathbf{R}^n \rightarrow \mathbf{R}$, 使得 $y = f(\mathbf{x})$ ^[11,12]。

考虑非线性回归问题,根据 Mercer 核展开定理,存在非线性映射 φ , 将输入空间变换到高维特征空间,在高维特征空间中,非线性回归问题转化为线性回归问题,存在回归目标函数^[13]:

$$f(\mathbf{x}) = \mathbf{w}^T \varphi(\mathbf{x}) + b. \quad (1)$$

式中 $\mathbf{w}, \mathbf{x} \in \mathbf{R}^n, b \in \mathbf{R}$ 。由最大间隔回归法,等价于求凸二次规划最优解问题^[14]

$$\min_{\mathbf{w}, b} \left[\frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^l (\xi_i + \xi_i^*) \right], \quad (2)$$

约束条件

$$\begin{cases} [\mathbf{w}^T \varphi(\mathbf{x}_i)] + b - y_i \leq \epsilon^* + \xi_i \\ y_i - [\mathbf{w}^T \varphi(\mathbf{x}_i)] - b \leq \epsilon^* + \xi_i^* \end{cases}, \quad \xi_i, \xi_i^* \geq 0 \quad (3)$$

式中 $i = 1, 2, \dots, l, C > 0$ 为惩罚因子, ϵ 是引入的不敏感损失, ξ_i, ξ_i^* 为松弛系数的上下限。

引入拉格朗日(Lagrange)函数求解该凸二次规划最优解^[14]：

$$L(\mathbf{w}, b, \xi, \xi^*, \mathbf{a}, \mathbf{a}^*, \eta, \eta^*) = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^l (\xi_i + \xi_i^*) - \sum_{i=1}^l (\eta \xi_i + \eta^* \xi_i^*) - \sum_{i=1}^l a_i \{\epsilon^* + \xi_i + y_i - [\mathbf{w}^T \varphi(\mathbf{x}_i)] - b\} - \sum_{i=1}^l a_i^* \{\epsilon^* + \xi_i^* - y_i + [\mathbf{w}^T \varphi(\mathbf{x}_i)] + b\}, \quad (4)$$

式中 $a_i, a_i^*, \eta, \eta^* \geq 0$ 是 Lagrange 乘子, 根据最小二乘原理, 对(4)式 L 分别求 $\mathbf{w}, b, \xi, \xi^*, \eta, \eta^*$ 的偏导, 并令其等于零, 由 Karush-Kuhn-Tucker 最优化条件(KKT 互补条件), 得到回归目标函数

$$f(\mathbf{x}) = \sum_{i=1}^l (a_i^* - a_i) \langle \varphi(\mathbf{x}_i), \varphi(\mathbf{x}) \rangle + b, \quad (5)$$

式中 a_i, a_i^*, b 为待定参数, $a_i, a_i^* \neq 0$ 对应 \mathbf{x}_i 称为支持向量, l 为支持向量数, \mathbf{x} 为待预测量, 最优回归超平面由支持向量样本点决定。

事实上, 非线性变换 φ 的表达式, SVR 能够借助 Mercer 定理将高维特征空间中的内积运算 $\langle \varphi(\mathbf{x}_i), \varphi(\mathbf{x}) \rangle$ 用核函数 $K(\mathbf{x}_i, \mathbf{x})$ 代替, (5) 式可转化为^[15]

$$f(x) = \sum_{i=1}^l (a_i^* - a_i) K(x_i, x) + b. \quad (6)$$

核函数的种类很多, 可以在实际应用中通过试验确定, 而且核函数与向量无关, 可以巧妙地避免“维数灾难”。

2.2 算法物理基础

文献[7]指出, 用于太阳辐射测量的硅半导体探测器响应度 R 受测量环境气象因素的影响, 其计算模型如(7)式所示

$$R = f_{AM} \left(C_1 \frac{G-D}{E_0} f_z + C_2 \frac{D}{E_0} \right) \times [1 - 0.00082(T - 25)]^{-1}, \quad (7)$$

式中 f_{AM}, f_z 分别为空气质量和天顶角的函数, 其

表 1 数据资料的采集设备、时段及精度信息

Table 1 Information of data collection equipments, time intervals and precision

References	Instrument	Time	Accuracy	
			MBE	RMSE
Diffuse irradiance (standard)	Eppley 8-48	6:00~19:00	±5%	±3 W/m ²
Diffuse irradiance (to be modified)	RSP LI-200	6:00~19:00	-15%~10%	5%~6%
Total irradiance (modified)	RSP LI-200	6:00~19:00	+1.0%~-1.2%	4%~6%
Temperature	Vaisala WXT520	6:00~19:00	±0.3 °C	
Relative humidity	Vaisala WXT520	6:00~19:00	±3%~±5%	
Atmospheric pressure	Vaisala WXT520	6:00~19:00	±0.05~±0.10 kPa	

为研究气象要素对 RSP 测量散射辐照度误差的影响, 需要利用同期的太阳总辐照度、环境温度、相对湿度和气压数据。文中太阳总辐照度由 RSP

形式根据实验确定; G, D, T 分别为太阳总辐射、散射辐射值和环境温度, C_1, C_2 是 $E_0 = 1000 \text{ W/m}^2$ 时直接辐射和散射辐射的检定常数。由此可见, 要测量的散射辐照度自然会因空气质量、天顶角、环境温度和总辐照度的变化出现差异, 这就为本文从气象因素修正 RSP 测量的散射辐照度提供了物理基础。

3 RSP 散射辐照度测量误差分析

3.1 资料来源

由于国内太阳辐射观测业务中 RSP 应用稀少, 获取其测量资料受到限制。本文采用美国国家太阳辐射研究实验室(SRRL)于 2010 年 9 月 1 日至 2010 年 12 月 30 日期间测量的时平均数据进行研究。SRRL 位于科罗拉多州南桌山上(105.18°W, 39.74°N), 海拔 1829 m, 这里地势开阔, 从日出到日落都没有遮挡物, 常年太阳辐射充足。SRRL 是世界上最大的持续运行的太阳辐射观测本底站, 拥有 75 套用于气象要素观测的设备, 其观测数据质量可靠且对外开放。文中待修正的 RSP 是由 Irradiance 公司生产的第二代旋转遮光带日射表, 光电探测器为 LICOR LI-200, 采集散射辐照度, 其精度信息如表 1 所示; 参考标准表是由 Eppley 公司生产的 8-48 热电型总日射表, 加遮光盘, 采集参考标准散射辐照度, 精度信息如表 1 所示, 其平均偏差和标准偏差都优于未经修正的 RSP。

测得, 已经过修正。环境温度、相对湿度和气压由 Vaisala 公司的 WXT520 测量变送器采集, 精度信息列于表 1。同时选用了天顶角和空气质量的计算

值作为误差影响因子,在去除了不满足气候学极值检查规则的误差样本后,最终共取得 121 天共 1300 组样本数据。

3.2 数据预处理

在进行数据分析之前,为消除不同气象要素数据间的量级差异,提高运算效率,对测量数据作极值归一化处理^[16,17]

$$v_{\text{nor}} = \frac{v_i - v_{\text{min}}}{v_{\text{max}} - v_{\text{min}}}, \quad (8)$$

式中 v_{nor} 为要素归一化值, v_i 为该要素实际值, v_{min} 为该要素最小值, v_{max} 为该要素最大值,这样,处理后各因子数据尽落于 $[0, 1]$ 区间内。

需要指出的是,各地的海拔高度不同,直接计算得到的空气质量不能用于相关的数据处理,需要进行气压修正^[7]:

$$M_a = (P/P_0) \cdot M, \quad (9)$$

式中 M_a 为气压修正后的空气质量, P 为观测点气压, $P_0 = 1013.25 \times 10^2$ Pa, 是海平面一个标准大气压, M 为气压修正前观测站点的空气质量。

3.3 误差分析

将参考表 Eppley 8-48 所测散射辐照度视为真值, RSP 所测散射辐照度的绝对误差表示为

$$\Delta r = D(u) - D_s, \quad (10)$$

式中 $D(u)$ 是由 RSP 测得未修正的散射辐照度, D_s 为 Eppley 8-48 测得视为真值的散射辐照度, Δr 即为 RSP 所测散射辐照度的绝对误差。统计数据资料可以发现, RSP 所测 $D(u)$ 比 D_s 平均低 11.9%, 标准差为 0.938, 可见 RSP 对散射辐照度的测量误差是不容忽略的, 需加以修正。

为方便修正模型的建立, 引入误差修正值作为研究对象。误差修正值是与绝对误差 Δr 绝对值相等、符号相反的量^[17], 用 e 表示为

$$e = -\Delta r = D_s - D(u). \quad (11)$$

这样, 影响因子对 RSP 散射辐照度误差的影响关系转化为对其误差修正值 e 的影响关系。记 e_{nor} 为误差修正值 e 归一化处理后的值, 图 1 分别给出了 e_{nor} 随气象影响因子的分布及物理变化趋势, 其中变化趋势由 e_{nor} 随太阳总辐照度、环境温度、天顶角和空气质量(气压修正)的二次、线性和三次拟合表示, 他们的最大残差模分别为 4.2565, 7.1555, 5.9723 和 6.0206。多项式拟合的阶数越高, 其最大残差模越小, 但其表达式形式也越复杂。文中之所以选择二次、线性和三次拟合曲线, 就是因为此处升高拟合阶数, 其最大残差模已没有太大改善, 但拟合形

式却将大大复杂。因此, 在保证合理描述 e_{nor} 变化趋势的前提下, 选用了形式简单且最大残差模较小的拟合方式。

图 1(a) 给出 e_{nor} 在 $0 \sim 950$ W/m^2 总辐照度范围内的分布图, 其中曲线为 e_{nor} 相对于太阳总辐照度的二次拟合曲线, 它反映了 RSP 测量散射辐照度的误差随太阳总辐照度攀升而增大。这是由于太阳总辐照度强弱引起太阳辐射光谱特性变化, 对 LICOR 光谱响应度产生影响造成的。在晴朗天气下, RSP 测得的 $D(u)$ 比在多云天气下测得的 $D(u)$ 小 20%~30%^[8]。根据数据资料分析, 太阳总辐照度每增加 100 W/m^2 , RSP 测量散射辐照度的误差平均增加 2.907 W/m^2 。

图 1(b) 给出 e_{nor} 随环境温度的分布情况及其相对于环境温度的线性变化曲线, e_{nor} 随环境温度逐步上升, 其与环境温度的相关系数为 0.2975。45 $^{\circ}\text{C}$ 的温度变化范围内, e_{nor} 变化率为 0.007, 说明 RSP 散射辐照度测量误差随环境温度升高而有所增加。造成这种现象的原因在于, LICOR LI-200 响应度受温度变化的影响大于热电型日射表响应度受温度变化的影响^[18]。

太阳辐射光谱对 RSP 测量散射辐照度的准确度有影响, 本文通过天顶角和空气质量加以体现。图 1(c) 和图 1(d) 分别给出 e_{nor} 随天顶角和空气质量(气压修正) M_a 的分布规律。图 1(c) 中曲线表示 e_{nor} 随天顶角变化的三次拟合曲线, 随着天顶角的增加, e_{nor} 呈减小趋势, 说明 RSP 测量散射辐照度的误差随天顶角增大而减小。图 1(d) 中 e_{nor} 样本点集中分布在 M_a 介于 $0 \sim 5$ 和 $20 \sim 25$ 之间的区域, 位于 $20 \sim 25$ 之间的 e_{nor} 比落于 $0 \sim 5$ 之间的 e_{nor} 平均减少了 53.7%, e_{nor} 相对于 M_a 的三次拟合曲线显示 e_{nor} 随 M_a 增大总体上趋于减小, 即 RSP 测量散射辐照度的误差随 M_a 增大而减小。造成这种现象的原因之一, 天顶角和空气质量在清晨或傍晚时段较大, 此时大气中介于 $640 \sim 750$ nm 的红光所占比重增加, 而天顶角和空气质量较小的中午时段, 大气中介于 $450 \sim 480$ nm 的蓝光占据主导, 由于 LICOR LI-200 对红光的光谱响应度高于蓝光的光谱响应度, 所以, 天顶角和空气质量较大时, RSP 测得的散射辐照度要相对准确; 另一个原因是一天中, 热电型辐射表光谱响应度变化不大, 受余弦响应影响较小^[19], Eppley 8-48 的测量值不会因光谱变化而起伏太大。如此双重作用影响下, 造成了 RSP 散射辐照度测量误差随天顶角和 M_a 增大而减小。

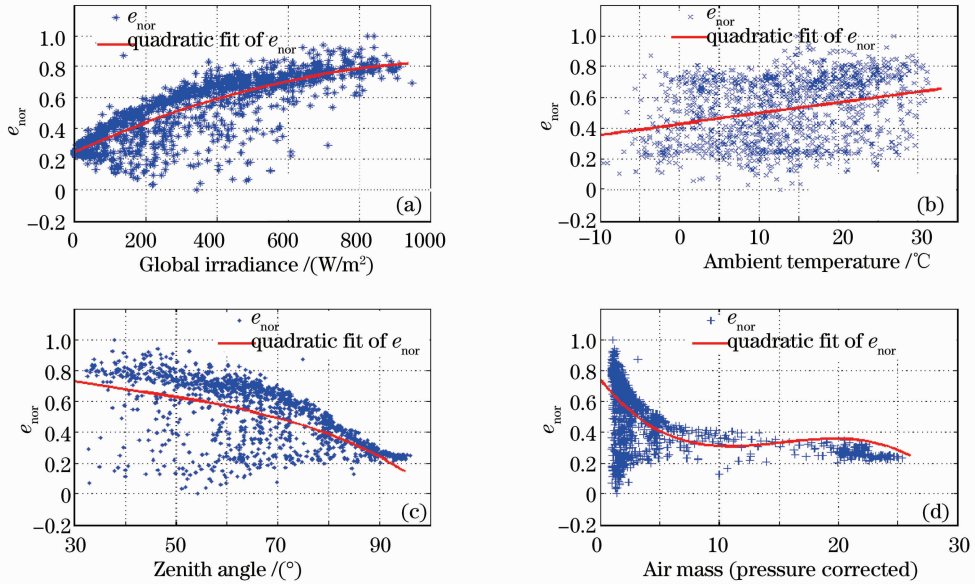


图1 e_{nor} 随气象影响因子的分布情况及变化趋势。(a)随太阳总辐照度;(b)随环境温度;(c)随天顶角;(d)随空气质量(气压修正)

Fig. 1 Plot of distribution and trend of e_{nor} versus the meteorological influence factors. (a) Versus global irradiance of sun; (b) versus ambient temperature; (c) versus zenith angle; (d) versus air mass (pressure corrected)

另外,相对湿度对RSP散射辐照度测量准确度有着不可忽略的影响。相对湿度表征着空气中的水汽含量,研究表明^[20],空气中水汽含量增加,水汽对近红外波段700~1000 nm的太阳辐射将产生不可忽略的吸收作用,其最大红外吸收窗口介于900~980 nm之间^[21]。然而,在300~950 nm之间,LICOR LI-200的光谱响应度是随辐射波长增大而增加的,在950 nm左右达到最大^[6]。即LICOR LI-200光谱响应度最大对应的谱段恰是水汽吸收最严重的谱段,因此,随着相对湿度的增大,RSP测量到的近红外波段散射辐照度相对减少,但是Eppley 8-48测量近红外波段散射时受影响程度较小,因而造成误差修正值 e_{nor} 随相对湿度增大而增加。

空气气压对RSP散射测量误差的影响较小,它与 e_{nor} 的相关系数只有0.1884,所以本文不将气压列入影响因子序列,而是用作参考,修正空气质量。

4 构建支持向量机回归修正模型

为修正太阳总辐照度、环境温度、相对湿度、天顶角和空气质量(气压修正)对RSP散射辐照度测量造成的误差,以归一化误差修正值 e_{nor} 值为回归对象,太阳总辐照度、环境温度、相对湿度、天顶角和空气质量(气压修正)为支持因子,构建支持向量机回归模型。考虑所建模型的优化,将1300组数据分为三个数据集:训练数据集、优化数据集和检验数据

集,各占总样本的75%,20%和5%,分别用于支持向量机回归模型的建立、所建模型参数优化和性能检验。

本文选择高斯型径向基函数作为核函数

$$K(\mathbf{x}_i, \mathbf{x}) = \exp(-\|\mathbf{x} - \mathbf{x}_i\|/\sigma^2), \quad (12)$$

式中 \mathbf{x}_i 为支持向量, \mathbf{x} 为待预测量, σ 为径向基核参数。构建支持向量机回归模型涉及到惩罚因子 C 、不敏感损失系数 ϵ 和径向基核参数 σ 的选择,本文实验中采用逐步筛选法加以确定,最终分别取 $C=140.06$, $\epsilon=16.17343$, $\sigma=6.88754$,模型择优标准为相关系数和均方根误差(RMSE)^[22]。

对检验数据集样本进行预测,效果如图2所示, e_{nor} 的预测值与实况值变化趋势基本一致,相关系数为0.921,RMSE为0.0728,表明 e_{nor} 预测模型的效果良好。

根据极值归一化(8)式,误差修正值 e 实际预测值可利用 e_{nor} 的最优预测值,还原出来,即

$$e_p = e_{\text{nor}_p} \times (e_{\text{max}} - e_{\text{min}}) + e_{\text{min}}, \quad (13)$$

式中 e_{nor_p} 为 e_{nor} 的最优预测值, e_p 即是 e 实际最优预测值, e_{max} 和 e_{min} 分别是误差修正值 e 的最大值和最小值。用 e 实际预测值 e_p 修正RSP散射辐照度,推导出RSP散射辐照度测量误差修正算法SVR-DC,

$$D(c) = D(u) + e_p, \quad (14)$$

式中 $D(c)$ 为修正后的散射辐照度。

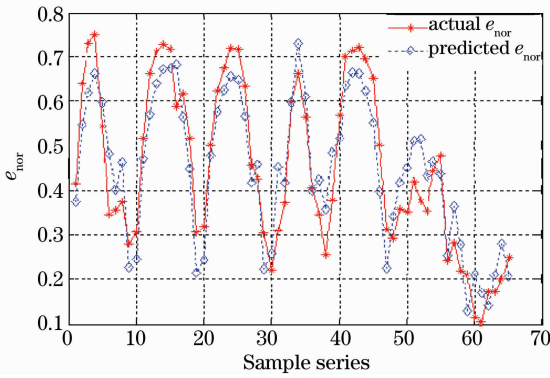


图 2 e_{nor} 的实况值与预测值对比图

Fig. 2 Plot of comparison between actual e_{nor} and predicted e_{nor}

5 SVR-DC 修正性能评价

修正算法模型 SVR-DC 构建完成后,要对其修正效果进行评估,首先对其自身的修正性能进行考察,以检验其修正有效性和适用性;然后再与同类修正算法的修正效果进行对比,以检验其优越性。

5.1 SVR-DC 性能评估

为证明 SVR-DC 修正算法自身的有效性,引入测量准确度作为评价标准。测量准确度表示测量结果与被测量真值之间的一致程度^[23],它由平均偏差(MBE)和均方根误差(RMSE)综合体现。MBE 反映测量结果中系统误差的大小,而 RMSE 反映测量结果中随机误差的大小。以 D_s 作为散射辐照度真值,修正前 RSP 散射辐照度 $D(u)$ 和修正后散射辐照度 $D(c)$ 的平均偏差和均方根误差可以分别用(15)和(16)式计算:

$$e_{mb}(x) = \frac{\sum_{j=1}^n D(x)_j - D_{s_j}}{n}, \quad (15)$$

$$e_{rms}(x) = \sqrt{\frac{\sum_{j=1}^n (D(x)_j - D_{s_j})^2}{n}}, \quad (16)$$

式中 $e_{mb}(x)$ 和 $e_{rms}(x)$ 分别表示 RSP 所测散射辐照度的平均偏差和均方根误差, $D(x)_j$ 为样本序列 j 对应的散射辐照度值,此处 x 取 u 或 c , 分别对应修正前和修正后的情况, D_{s_j} 为样本序列 j 对应的散射辐照度真值, n 为测量样本数。

$e_{mb}(x)$ 越趋于 0, 说明测量结果的系统误差越小, 其中若 $e_{mb}(c) > 0$, 则存在过修正, 若 $e_{mb}(c) < 0$, 则存在欠修正; $e_{rms}(x)$ 越接近于 0, 说明测量结果的随机误差越小, 综合起来, 测量准确度就越高。计算检验数据集的 MBE 和 RMSE 如表 2 所示, 修正

后散射辐照度的 MBE 和 RMSE 较修正前减小了一个数量级, 系统误差接近于 0, 随机误差也大大减小, 修正效果显著。

表 2 SVR-DC 修正前后的 MBE 和 RMSE/(W/m²)

Table 2 MBE and RMSE before and after SVR-DC correction/(W/m²)

	$x=u$	$x=c$
$e_{mb}(x)$	-8.8	-0.2
$e_{rms}(x)$	12.2	3.3

为检验 SVR-DC 修正算法及建模所选参数是否具有普遍意义, 选用异于 SRRL 测量站点及其测量时间的数据进行修正。本文随机选择了科罗拉多州立劳里观测站(LRSS: Lowry Range Solar Station, 39.60701°N, 104.58017°W, 1860 m) 2011 年 6 月 16 至 2011 年 6 月 19 日三天早晨 6 点至晚上 19 点之间 RSP 测量的散射辐照度数据, 同时利用了同期的气象影响要素数据, 其经 SVR-DC 算法修正前后的对比情况如 3 所示, 相关系数高达 0.997。修正前 MBE 等于 -10.3 W/m², RMSE 等于 11.7 W/m², 修正后 MBE 和 RMSE 分别等于 1.9 和 8.5 W/m², 较修正前至少减小了 27%。可见, SVR-DC 修正算法对不同测量站点、不同测量时间 RSP 测量的散射辐照度也具有不错的修正效果。

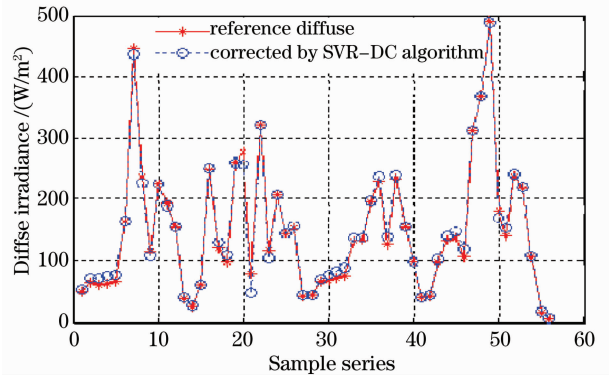


图 3 LRSS 观测站 RSP 所测散射辐照度经 SVR-DC 修正效果图

Fig. 3 Plot of SVR-DC correcting result for diffuse irradiance measured by LRSS

5.2 同类算法性能比较

为证明 SVR-DC 修正算法在同类修正算法中具有优势, 将 SVR-DC 修正算法与同类 Vignola 修正算法和 VA 修正算法的修正效果进行对比。图 4 所示为检验数据集 RSP 散射辐照度经过 Vignola 算法、VA 算法和 SVR-DC 算法修正后与参考标准散射辐照度的对比效果。

1) Vignola 修正算法是 Vignola 在前人实验研

究的基础上,从单因素影响入手,针对太阳总辐照度、环境温度和相对湿度三个要素,分别提出一个修正 RSP 散射辐照度的表达式,然后将三者修正结果叠加,进而实现修正 RSP 所测散射辐照度的效果。这存在割裂因素影响的不足,且未考虑空气质量和天顶角所造成的误差,容易造成低散射辐照度时的欠修正现象。SVR-DC 改进了这种割裂因素影响的缺陷,综合考虑总辐照度、环境温度、相对湿度、天顶角和空气质量的影响,从整体上修正气象要素造成的误差,效果改善明显。

从图 4(a)可以看出,经 Vignola 算法修正后, RSP 散射辐照度 $D(c)$ 与参考标准散射辐照度 D_s 的变化趋势趋于一致,相关系数达 0.9961,说明 Vignola 算法能够较好修正 RSP 散射辐照度的测量误差。但不可否认,在散射辐照度低于 50 W/m^2 时,经 Vignola 算法修正得到的 $D(c)$ 多处低于真值 D_s ,表 3 给出其 MBE 等于 $-3.3 \text{ W/m}^2 < 0$,RMSE

等于 6.6 W/m^2 ,欠修正现象明显。图 4(c)所示为经 SVR-DC 算法修正后 RSP 散射辐照度 $D(c)$ 与参考真值 D_s 间的效果对比图, $D(c)$ 与 D_s 的相关系数高达 0.9992,其 MBE 和 RMSE 在表 2 中已经给出,分别为 -0.2 W/m^2 和 3.3 W/m^2 ,分别较 Vignola 算法减小了 50% 以上。因此,在避免欠修正和修正空气质量造成的误差方面,SVR-DC 性能有所改进,效果优于 Vignola 算法。

2) VA 修正算法是在 Vignola 算法基础上提出的,它仅以太阳总辐照度作为影响因素,修正 RSP 散射辐照度的测量值,环境温度、相对湿度、天顶角和空气质量造成的误差都由总辐照度代为体现,这就导致高散射时过修正现象的存在。SVR-DC 算法综合考虑了总辐照度、环境温度、相对湿度、天顶角和空气质量每一个因素造成的误差,从整体上修正其不利影响,避免了高散射时 VA 修正算法存在的过修正现象。

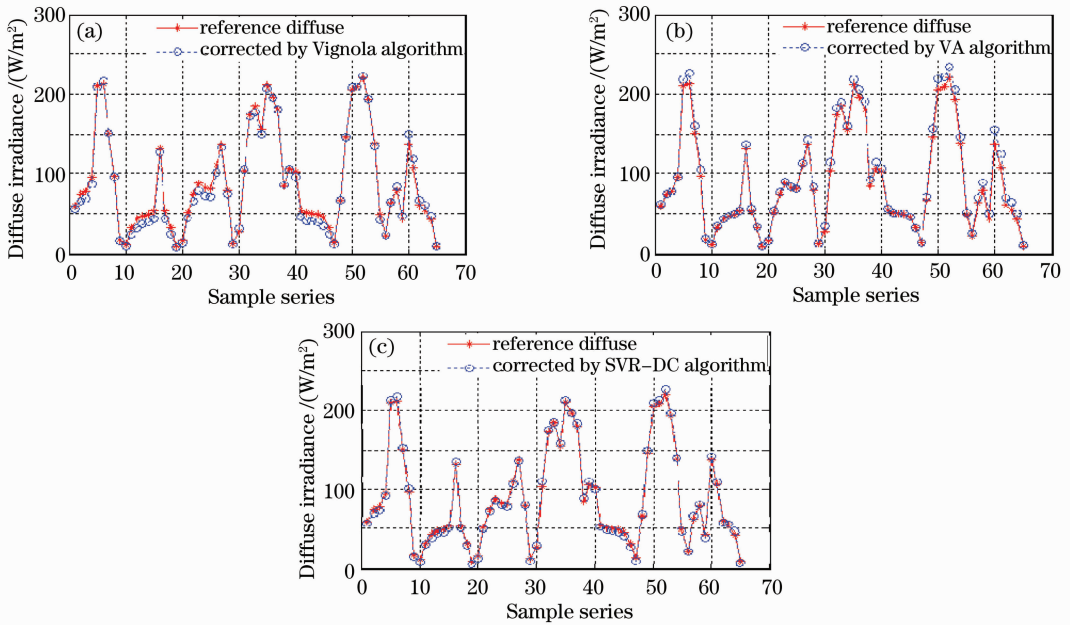


图 4 经三种修正算法修正后 RSP 散射辐照度与参考散射辐照度 D_s 的对比效果。(a)经 Vignola 算法修正后; (b)经 VA 算法修正后; (c)经 SVR-DC 算法修正后

Fig. 4 Plot of comparison between RSP diffuse irradiance corrected by three correction algorithms and reference diffuse irradiance. (a) Corrected by Vignola algorithm; (b) corrected by VA algorithm; (c) corrected by SVR-DC algorithm

图 4(b)所示为经 VA 算法修正后 RSP 散射辐照度 $D(c)$ 与参考散射辐照度 D_s 间的效果对比图,它们的相关系数为 0.9985,比 Vignola 算法有所提高,且不存在欠修正。但是,当散射辐照度超出 100 W/m^2 时, $D(c)$ 在多个样本点上明显高于 D_s ,出现了过修正。统计显示,此时 $D(c)$ 比 D_s 平均高出 5.5%,最大时高出 18.9 W/m^2 。经 VA 算法修正后 RSP 散射

辐照度的 MBE 和 RMSE 如表 3 所示,其 MBE 等于 $4.9 \text{ W/m}^2 > 0$,RMSE 等于 6.9 W/m^2 ,显示过修正存在。同样,与图 4(c)相比,经 SVR-DC 修正后 RSP 散射辐照度的 MBE 仅为 -0.2 W/m^2 ,RMSE 为 3.3 W/m^2 ,较 VA 算法同样有 50% 以上的改善。可见,SVR-DC 算法能较为准确地计算出散射误差修正值,减小测量结果的系统误差和随机误差,弥补

VA 算法在过修正方面的不足。

表 3 Vignola 算法、VA 算法修正后散射辐照度 $D(c)$ 的 MBE 与 RMSE/(W/m²)

Table 3 MBE and RMSE of diffuse irradiance $D(c)$ corrected by Vignola algorithm and VA algorithm/(W/m²)

Type of correction algorithm	$e_{mb}(c)$	$e_{rms}(c)$
Vignola algorithm	-3.3	6.6
VA algorithm	4.9	6.9

6 结 论

分析了太阳总辐照度、环境温度、相对湿度、天顶角和空气质量(气压修正)对 RSP 散射辐照度测量误差的影响关系,从支持向量机回归预测角度,针对误差修正值构建预测模型,提出了以误差修正值的最优预测修正 RSP 所测散射辐照度的 SVR-DC 算法。对 SRRL 检验数据和 LRSS 测量数据的修正结果显示,修正后两个观测站 RSP 散射辐照度的 MBE 和 RMSE 值分别为 -0.2 W/m^2 , 3.3 W/m^2 和 1.9 W/m^2 , 8.5 W/m^2 , 较修正前减小至少 27%, 系统误差和随机误差大大减小,修正效果明显。而且,在与同类 Vignola 修正算法和 VA 修正算法的效果对比中,SVR-DC 算法能够弥补散射辐照度低于 50 W/m^2 时 Vignola 算法的欠修正现象,避免散射辐照度高于 100 W/m^2 时 VA 算法的过修正现象,MBE 和 RMSE 比经 Vignola 算法和 VA 算法修正后减小 50% 以上。综合来看,SVR-DC 算法修正效果有效,适用性较好,且在同类算法中具有优势,具有良好的实际应用前景。

参 考 文 献

- 1 Sha Yizhuo, Bian Zeqiang, Lü Wenhua. Design of solar energy resources observation station in China [J]. *Meteorological, Hydrological and Marine Instruments*, 2010, (1): 6~9
沙奕卓, 边泽强, 吕文华. 我国太阳能资源观测站的设计[J]. *气象水文海洋仪器*, 2010, (1): 6~9
- 2 Lee Harrison, Joseph Michalsky, Jerry Berndt. Automated multifactor rotating shadowband radiometer; an instrument for optical depth and radiation measurements[J]. *Appl. Opt.*, 1994, **33**(22): 5118~5125
- 3 J. Augustyn, T. Geer, F. Vignola *et al.*. Improving the accuracy of low cost measurement of direct normal solar irradiance [C]. Proc. Solar 2002, American Solar Energy Society, 2002
- 4 J. J. Michalsky, L. C. Harrison, W. E. Berkheiser. Cosine response characteristics of some radiometric and photometric sensors[J]. *Solar Energy*, 1995, **54**(6): 397~402
- 5 Li Zhaozhou, Zheng Xiaobing, Wu Haoyu *et al.*. Temperature characteristics research of high-precision spectral radiometer standard detector [J]. *Acta Optica Sinica*, 2004, **24**(3): 401~407
李照洲, 郑小兵, 吴浩宇等. 高精度光谱辐射标准探测器的温度特性研究[J]. *光学学报*, 2004, **24**(3): 401~407

- 6 David L. King, Jay A. Kratochvil, William E. Boyson. Measuring solar spectral and angle-of-incidence effects on photovoltaic modules and solar irradiance sensors[C]. 26th IEEE Photovoltaic Specialists Conference, 1997, Conference Record of the Twenty-Sixth IEEE: 1113~1116
- 7 David L. King, Daryl R. Myers. Silicon photodiode pyranometers: operational characteristics, historical experiences, and new calibration procedures [C]. 26th IEEE Photovoltaic Specialists Conference, 1997, Conference Record of the Twenty-Sixth IEEE: 1285~1288
- 8 Frank Vignola. Solar cell based pyranometers: evaluation of the diffuse response [C]. Proceedings of the 1999 Annual Conference, American Solar Energy Society, 1999: 260
- 9 Frank Vignola. Removing systematic errors from rotating shadowband pyranometer data [C]. Proceedings of the 35th ASES Annual Conference July 9-13, 2006, Denver, Colorado, 2006
- 10 Zhang Xuegong. Introduction to statistical learning theory and support vector machines [J]. *Acta Automation Sinica*, 2000, **126**(1): 32~42
张学工. 关于统计学习理论和支持向量机[J]. *自动化学报*, 2000, **126**(1): 32~42
- 11 Xing Hongyan, Jin Tianli. Weak signal estimation in chaotic clutter using wavelet analysis and symmetric LS-SVM regression [J]. *Acta Physica Sinica*, 2010, **59**(1): 140~145
行鸿彦, 金天力. 基于对偶约束最小二乘支持向量机的混沌海杂波背景下中的微弱信号检测[J]. *物理学报*, 2010, **59**(1): 140~145
- 12 Zhu Jia, Wang Zhenhui, Jin Tianli *et al.*. Combination of wavelet decomposition and least square support vector machine to forecast atmospheric ozone content time series [J]. *Climatic and Environmental Research*, 2010, **15**(3): 295~302
朱 佳, 王振会, 金天力等. 基于小波分解和最小二乘支持向量机的大气臭氧含量时间序列预测[J]. *气候与环境研究*, 2010, **15**(3): 295~302
- 13 Zhang Hua, Zeng Jie. Wind speed forecasting model study based on support vector machine [J]. *Acta Energetica Sinica*, 2010, **31**(7): 928~932
张 华, 曾 杰. 基于支持向量机的风速预测模型研究[J]. *太阳能学报*, 2010, **31**(7): 928~932
- 14 Yang Zhixia, Deng Naiyang. The analysis to the solution set of the primal problem in linear support vector ordinal regression[J]. *OR Transactions*, 2007, **11**(3): 105~112
杨志霞, 邓乃扬. 线性支持向量顺序回归机的原始问题的解集分析[J]. *运筹学学报*, 2007, **11**(3): 105~112
- 15 Yu Zhongdang, Wang Longshan. Salt & pepper noise switching filter based on LS-SVR convolution mask [J]. *Acta Optica Sinica*, 2009, **29**(1): 163~168
于忠党, 王龙山. 基于回归型最小二乘支持向量机卷积模板的椒盐噪声开关滤波器[J]. *光学学报*, 2009, **29**(1): 163~168
- 16 Gu Wanlong, Zhu Yeyu, Pan Pan *et al.*. Application of SVM methods to calculation of solar radiation [J]. *Acta Energetica Sinica*, 2010, **31**(1): 56~59
顾万龙, 朱业玉, 潘 攀. 支持向量机方法在太阳辐射计算中的应用[J]. *太阳能学报*, 2010, **31**(1): 56~59
- 17 Sha Dingguo. Error Analysis and Assessment of Uncertainty in Measurement[M]. China Metrology Publishing House, 2003. 10~18
沙定国. 误差分析与测量不确定度评定[M]. 中国计量出版社, 2003. 10~18
- 18 Andrew L. Rosenthal, Jeffrey M. Roberg. Twelve Month Performance Evaluation for the Rotating Shadowband Radiometer [D]. New Mexico: New Mexico State University, Southwest Technology Development Institute, 1994. 18~19
- 19 L. Alados-Arboledas, F. J. Batlles, F. J. Olmo. Solar radiation

- resource assessment by means of silicon cells[J]. *Solar Energy*, 1995, **54**(3): 183~191
- 20 Zhou Liangming, Liu Yuguang. Study on water vapor's absorption to sun's radiation [J]. *Ocean Technology*, 2004, **23**(4): 109~112
周良明, 刘玉光. 利用地物光谱资料研究水汽对太阳辐射的吸收 [J]. *海洋技术*, 2004, **23**(4): 109~112
- 21 Xiong Wei, Shi Hailiang, Wang Yuanjun *et al.*. Study on near-infrared spatial heterodyne spectrometer and detection of water vapor[J]. *Acta Optica Sinica*, 2010, **30**(5): 1512~1515
熊 伟, 施海亮, 汪元钧 等. 近红外空间外差光谱仪及水汽探测研究[J]. *光学学报*, 2010, **30**(5): 1512~1515
- 22 Li Gang, Zhao Jing, Li Jiaying *et al.*. Visible-infrared reflectance spectroscopy applied in rapid screen of diseases[J]. *Acta Optica Sinica*, 2011, **31**(3): 0317001
李 刚, 赵 静, 李家星 等. 可见-近红外反射光谱用于疾病快速筛查[J]. *光学学报*, 2011, **31**(3): 0317001
- 23 Xu Guirong, Chen Bo, Wan Rong *et al.*. Error analysis on precipitable water derived from ground-based GPS with various methods [J]. *Torrential Rain and Disasters*, 2008, **27**(4): 346~350
徐桂荣, 陈 波, 万 蓉 等. 地基 GPS 不同水汽反演方法的误差分析[J]. *暴雨灾害*, 2008, **27**(4): 346~350

栏目编辑: 谢 婧