

朱 垚,宋 欣,何建祥,等. 基于小波降噪和自适应加权法的温室数据融合[J]. 江苏农业科学,2021,49(5):180-186.  
doi:10.15889/j.issn.1002-1302.2021.05.033

# 基于小波降噪和自适应加权法的温室数据融合

朱 垚,宋 欣,何建祥,杨 磊

(天津农学院工程技术学院,天津 300384)

**摘要:**针对温室环境信息多节点数据采集中存在误差大、冲突多和冗余等问题,提出一种基于小波降噪和自适应加权的多传感器数据融合算法。通过小波降噪对采集的数据进行处理,使其具有良好的光滑度和稳定性;利用自适应加权算法对多传感器数据进行融合,得到测量数据的最优估计值。实例验证结果表明,采用所提算法可以有效降低原始数据中存在的噪声和冗余,得到方差较小的数据融合值,可以提高测量数据的准确性和降低数据的传输量;具备良好的稳健性,能够实现对温室环境信息的可靠性和一致性描述。

**关键词:**多传感器;数据融合;小波降噪;自适应加权算法;温室

**中图分类号:**S126      **文献标志码:**A      **文章编号:**1002-1302(2021)05-0180-06

近年来,随着我国温室生产的大面积推广与应用,如何实现对温室环境的优化调控,为农作物生长提供适宜的环境,达到优质、高效和低耗的生产目的<sup>[1]</sup>,已逐渐成为研究热点。温室作为一个复杂的热力学系统,具有多变量、滞后性、时变性、非线性和强耦合等特点<sup>[2]</sup>,传统控制方法难以实现对温室环境的精准调控。目前普遍采用的是智能控制方法,如模糊控制、神经网络、多目标优化等,以及多种方法交叉融合的混合控制。这些方法都是以精确的温室环境模型作为基础<sup>[3]</sup>,而温室环境模型的精确构建则要依赖于环境信息采集的鲁棒性和容错性。

因为温室内各点环境信息分布不均匀,同时也为了克服单点采集存在的不确定性和局限性,对同一个环境参数通常要进行多点采集。但由于外部环境和传感器内部构造等多种因素的影响<sup>[4]</sup>,采集到的多组数据中往往存在异常值、数据缺失、误差大、冲突多、信息冗杂等问题,严重影响了数据的一致性和可靠性<sup>[5]</sup>。因此,有必要通过数据融合技术

对采集到的数据进行处理,得到温室环境信息完整且一致性的描述,从而提高环境信息的鲁棒性和稳健性,以实现温室环境的精准调控。同时,也有助于减少数据传输量,提高控制系统的工作效率,降低无线传感器网络能耗。

目前,数据融合技术在设施农业环境控制方面的应用已取得了一些研究成果。熊迎军等提出在对采集数据进行一致性检测和 3 次指数平滑的基础上,利用一种新型支持度函数进行幂均方融合的方法,能够在得到较高融合精度的同时缩短算法的运行时间;但是该方法在对数据进行预处理时,采用分布图法对异常数据进行了剔除,会导致样本数据的减少,以及异构传感器间样本数量的不一致,从而影响后续的融合精度<sup>[6]</sup>。初洪龙等提出了基于信任度的多传感器数据融合方法,利用自定义的指数信任度函数,分配每个传感器数据的权重,实现多传感器的数据融合<sup>[7]</sup>;但指数运算相对复杂,开销大,不适用于温室无线传感器网络。王振等在利用狄克逊准则对采集数据进行预处理的基础上,采用卡尔曼滤波和基于最小二乘法原理的加权方法实现了数据融合,并通过丢包补偿策略提高了融合进度<sup>[8]</sup>。该方法仍旧存在剔除数据所导致的样本数量减少问题,而且卡尔曼滤波的融合效果过于依赖状态空间模型的精确性。此外,还有将聚类分析、模糊理论与专家系统相结合的融合方法<sup>[9-10]</sup>,但专家系统依赖于一定的先验知识,才能够进行推理。

针对上述问题,本研究提出先采用小波降噪对

收稿日期:2020-04-15

基金项目:天津市科委企业科技特派员项目(编号:18JCTPJC67600、19JCTPJC59200);教育部产学研合作协同育人项目(编号:201702177062);天津农学院研究生科研创新项目(编号:2019XY042)。

作者简介:朱 垚(1995—),男,浙江绍兴人,硕士研究生,主要从事智能农机装备关键技术研究。E-mail:2833465528@qq.com。

通信作者:宋 欣,博士,副教授,主要从事智能农机装备关键技术研究。E-mail:songxin-tju@163.com。

采集的数据进行处理,使数据序列具有良好的光滑度和稳定性;然后利用自适应加权算法对多传感器数据进行融合,从而实现对温室环境信息的可靠性和一致性描述。

## 1 温室环境采集数据的小波降噪

由于小波分析适用于由短时高频成分和长时低频成分组成且无太多先验知识的数据进行降噪处理<sup>[11]</sup>,故可用于对温室环境信息采集数据的处理。此外,由于采集温室环境参数会受限传感器的布置以及使用成本,采集节点相对较少,所获得的监测数据样本有限,故采用小波分析法还能充分利用采集的数据,避免剔除数据所造成的样本数据减少和不一致问题。小波降噪的处理过程如下<sup>[11-12]</sup>:

(1)小波分解:选择适当的小波基函数,确定分解层数,并利用 Mallat 算法对原始数据进行分解;设原始数据为  $f(x)$ ,则在尺度  $j$  下的低频部分为  $A_j^d f$ ,高频部分为  $D_j f$ ,以 0 级为最高分辨率方向,则原始数据最终可被分解为

$$f(x) = A_{-J}^d f + \sum_{j=-J}^{-1} D_j f. \quad (1)$$

式中, $J$  为分散层数,一般取值为 5 或 6。

(2)高频系数的阈值量化:选定合适的阈值对第 1 层至第  $N$  层的高频系数进行量化处理;

(3)小波重构:根据小波分解的第  $N$  层低频系数和经过量化处理后的第 1 层至第  $N$  层高频系数进行小波重构,得到降噪后的离散数据。

## 2 自适应加权融合算法

### 2.1 自适应加权融合算法的基本原理

自适应加权融合算法的基本原理<sup>[13]</sup>如图 1 所示,以总均方误差最小的为最优条件,依据每个传感器的测量值  $x_i (i=1,2,3,\dots,m)$ ,以自适应方式获得其最优权重  $w_i (i=1,2,3,\dots,m)$ ,使得最终的融合结果  $\hat{x}$  达到最优。该算法能够有效综合多传感器的测量信息,并具有算法简洁、无需先验知识、实时性好等优点,适用于温室环境信息的数据融合<sup>[14]</sup>。

融合结果  $\hat{x}$  和各传感器权重  $w_i$  满足如下关系:

$$\begin{cases} \hat{x} = \sum_{i=1}^m w_i x_i \\ \sum_{i=1}^m w_i = 1 \end{cases}. \quad (2)$$

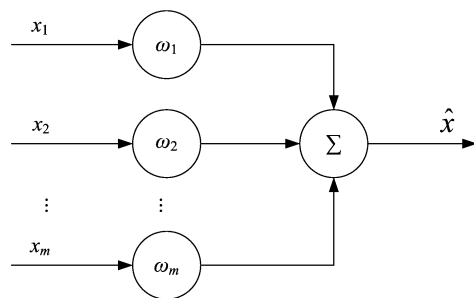


图1 自适应加权算法的基本原理

总均方误差( $\sigma^2$ )为

$$\sigma^2 = E[(x - \hat{x})^2] = \sum_{i=1}^m w_i^2 \sigma_i^2. \quad (3)$$

式中: $\sigma_i^2 (i=1,2,3,\dots,m)$  为各传感器的方差; $x$  为温室环境参数的真值,各传感器的测量值为  $x_i$  彼此相互独立且为  $x$  的无偏估计。

根据自适应加权算法的基本原理可知,当总均方误差  $\sigma^2$  为最小值时,其对应的各传感器权重  $w_i$  为最优,因此,采用拉格朗日乘数法对公式(3)求极值,得到极点:

$$w_i = \frac{1}{\sigma_i^2 \sum_{i=1}^m \frac{1}{\sigma_i^2}}. \quad (4)$$

此时,对应的最小均方误差为

$$\sigma_{min}^2 = \frac{1}{\sum_{i=1}^m \frac{1}{\sigma_i^2}}. \quad (5)$$

最终数据融合结果为

$$\hat{x} = \sum_{i=1}^m \frac{x_i}{\sigma_i^2 \sum_{i=1}^m \frac{1}{\sigma_i^2}}. \quad (6)$$

### 2.2 传感器方差的求解

利用各传感器所测数据间的相关性对方差进行求解。设 2 个独立的传感器  $p$  和  $q$ ,其待估计真值为  $x$ ,采样数据分别为  $x_p, x_q$ ,其对应的测量误差分别为  $e_p, e_q$ ,且为零均值的平稳噪声,则有

$$\begin{cases} x_p = x + e_p \\ x_q = x + e_q \end{cases}. \quad (7)$$

传感器  $p$  的方差可表示为

$$\sigma_p^2 = E[e_p^2]. \quad (8)$$

$x_p, x_q$  的互协方差( $R_{pq}$ )为

$$R_{pq} = E[x_p x_q] = E[x^2]. \quad (9)$$

$x_p$  的自协方差( $R_{pp}$ )为

$$R_{pp} = E[x_p^2] = E[x^2] + E[e_p^2]. \quad (10)$$

由公式(8)、公式(9)、公式(10)可得:

$$\sigma_p^2 = R_{pp} - R_{pq} \quad (11)$$

设传感器在  $k$  时刻时,  $R_{pp}$  和  $R_{pq}$  的时间域估计值分别为  $R_{pp}(k)$ 、 $R_{pq}(k)$ , 则有:

$$R_{pp}(k) = \frac{1}{k} \sum_{j=1}^k x_p(j) x_p(j) = \frac{k-1}{k} R_{pp}(k-1) + \frac{1}{k} x_p(k) x_p(k) \quad (12)$$

$$R_{pq}(k) = \frac{1}{k} \sum_{j=1}^k x_p(j) x_q(j) = \frac{k-1}{k} R_{pq}(k-1) + \frac{1}{k} x_p(k) x_q(k) \quad (13)$$

$R_{pq}$  可用  $R_{pq}(k)$  的均值  $\bar{R}_{pq}(k)$  进行估计:

$$R_{pq} = \bar{R}_{pq}(k) = \frac{1}{m-1} \sum_{\substack{p=1 \\ q \neq p}}^m R_{pq}(k) \quad (14)$$

由此可知, 各传感器方差可通过自协方差  $R_{pp}$  和

互协方差  $R_{pq}$  的时间域估计值进行估计。

### 3 实例分析

试验数据采集自辽宁省某温室试验基地, 以 2019 年 3 月 1 日采集到的温度、湿度数据为样本, 进行算法验证。温室中按照一定距离间隔分别布置 2 个传感器节点, 每隔 30 min 采集 1 次数据, 共采集 48 组样本。

#### 3.1 小波降噪预处理

利用 Matlab 软件对样本数据进行小波降噪处理。通过多次试验结果对比, 最终选择 db5 小波基函数进行 3 层分解, 各层阈值选用软阈值函数, 阈值估计采用 RigrSure 阈值法, 降噪重构后的数据与原始样本数据的对比结果如图 2 所示。

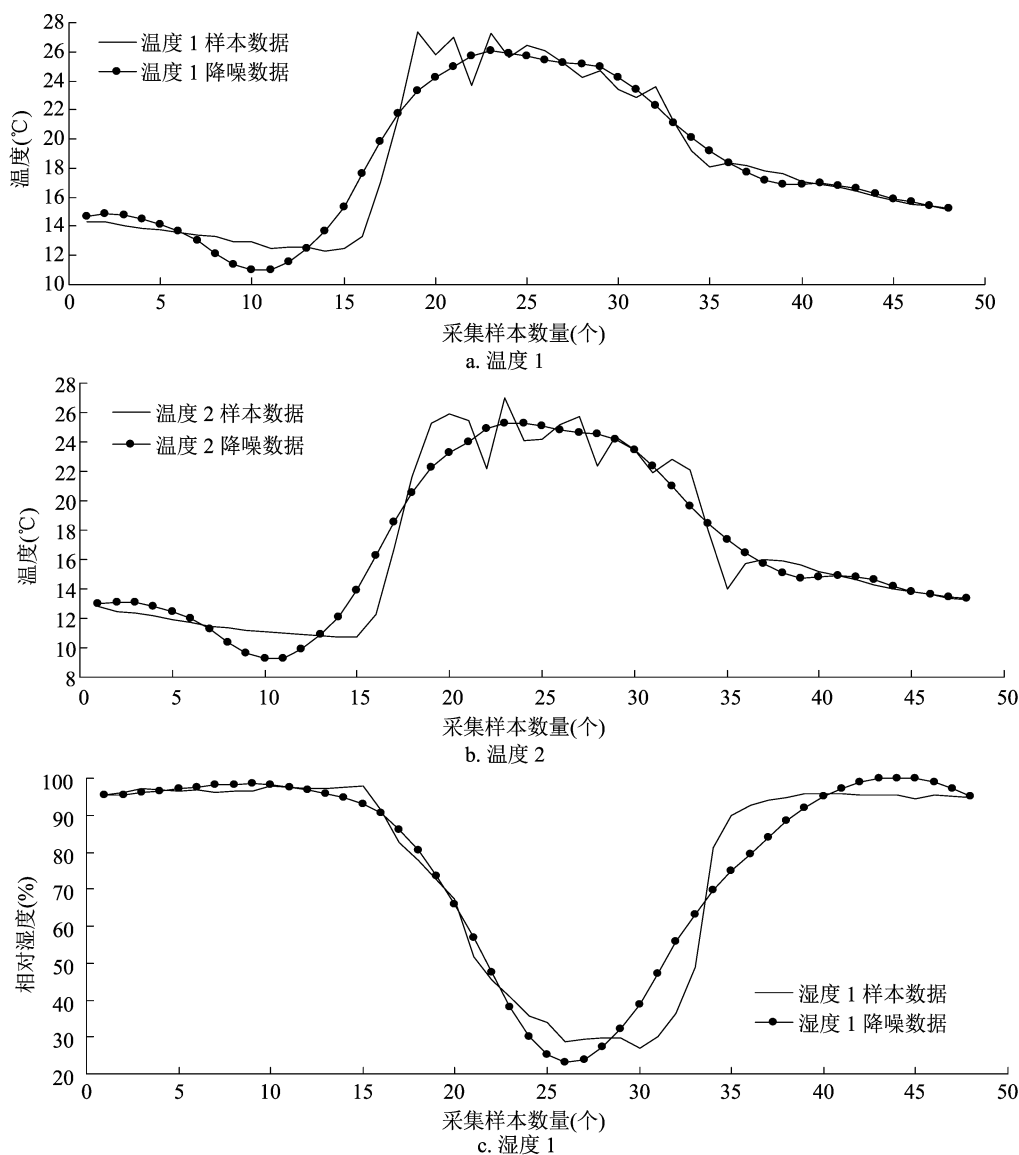
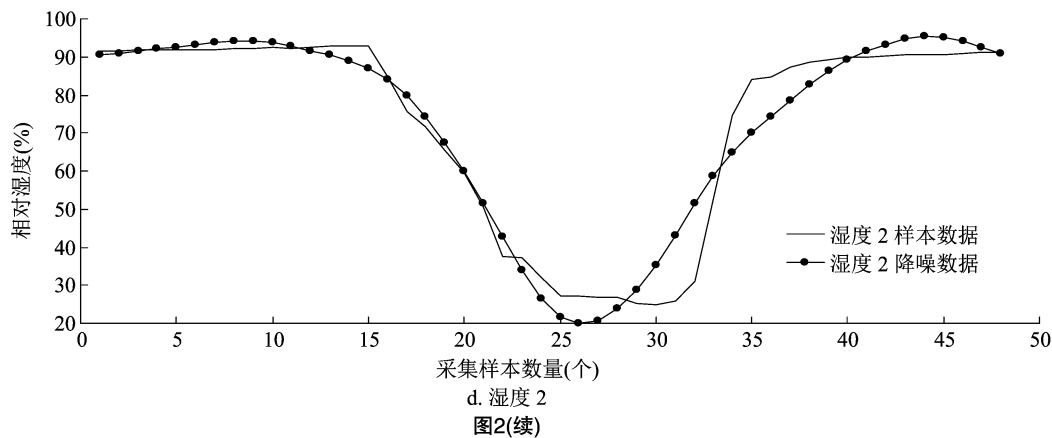


图2 温湿度采集数据降噪效果



从图 2 - a、图 2 - b 可以看出,温度曲线在中间时段存在较大波动,经过小波降噪处理后的数据曲线无奇异值出现,具有良好的光滑度和稳定性,起到了很好的降噪效果。

3.2 多传感器数据融合

根据“2.1”节中所述的自适应加权融合算法的基本原理,利用 Matlab 软件实现其求解过程。在权重求最优解时,各传感器权重会随着采样次数的增加而变化,当样本数量达到一定数值时,权重会趋于稳定,并最终收敛于某一数值。如图 3 - a 所示,温度权重大约在样本数量为 40 时趋于稳定,最终分

别收敛于 0.53、0.47。如图 3 - b 所示,湿度权重大约在样本数量为 35 时趋于稳定,最终分别收敛于 0.52、0.48。表 1 所示为收敛后,满足总均方误差最小条件的各温度、湿度传感器权重的最优解,及其对应的最小均方误差。可以看出,融合后的总均方误差要低于各个传感器的方差,证明经过数据融合可以得到各测量值的最优估计,从而提高传感器数据的准确性。

将融合算法与算术平均法的融合效果相比(图 4)可以看出,应用所提算法得到的融合曲线比算术平均法融合曲线更加光滑,数据波动小,具有很好

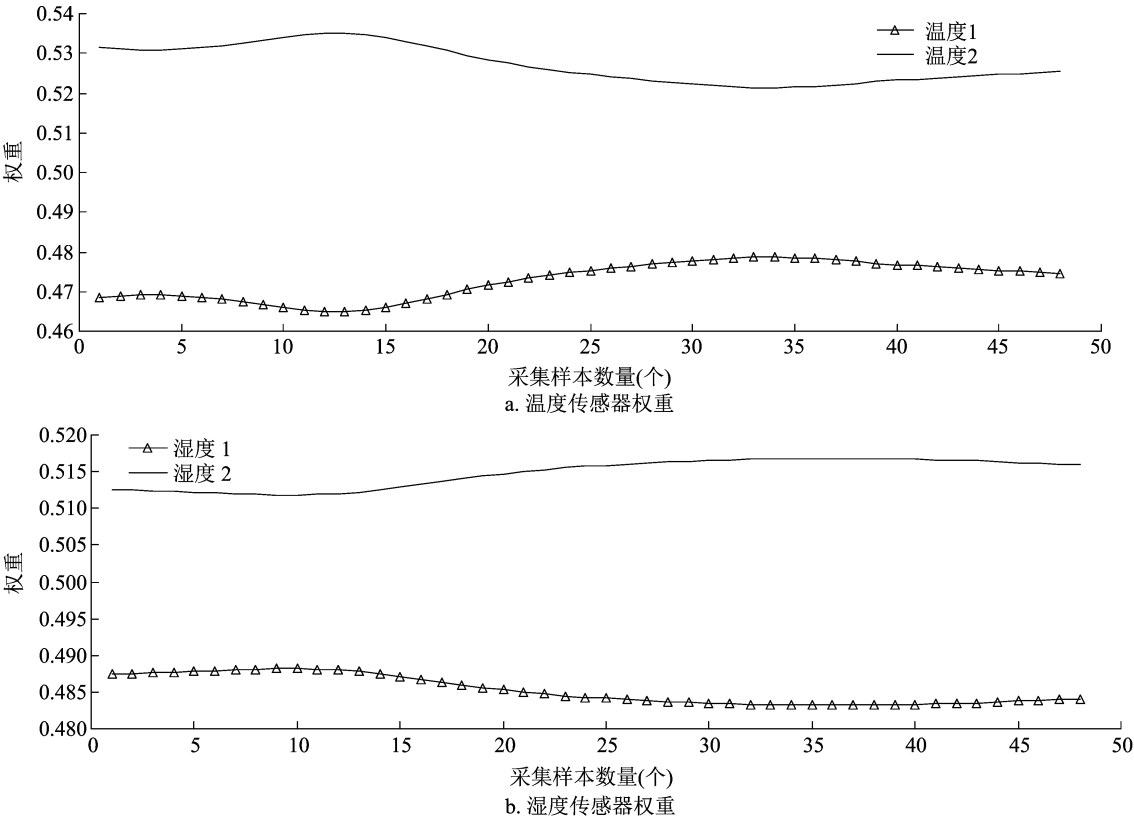


图3 温湿度传感器权值变化曲线

表 1 温湿度传感器最小均方误差、最优权重及总均方误差

项目	$\sigma_1^2$	$\sigma_2^2$	$w_1$	$w_2$	$\sigma^2$
温度	0.25	0.23	0.47	0.53	0.12
湿度	0.38	0.36	0.48	0.52	0.18

的稳定性。为了进一步对比融合效果,分别计算了 2 种融合结果的标准偏差进行比较,如图 5 所示,采用本研究所提算法得到的融合结果,其标准偏差曲线更加平滑和稳定,从而证明其融合结果更加可靠。

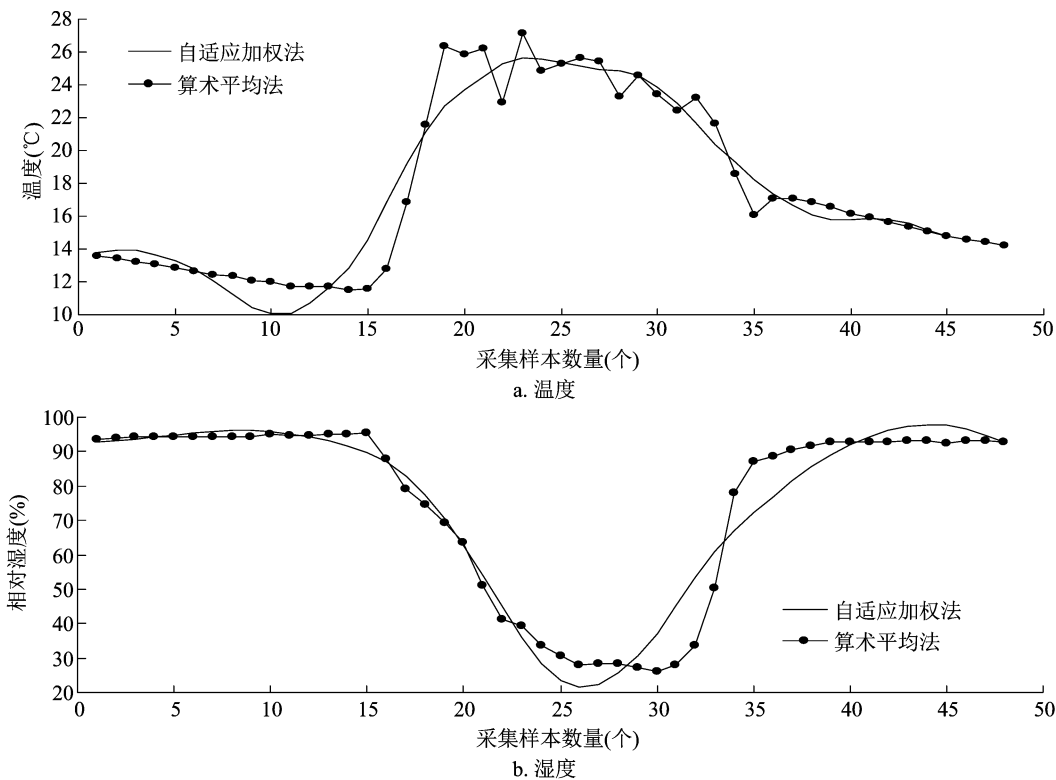


图4 温度、湿度数据融合曲线

3.3 算法稳健性验证

为验证本研究算法的稳健性,增加 1 组 2019 年 3 月 3 日的湿度数据作为干扰数据,模拟采集数据存在异常值的情况。从图 6 可以看出,干扰数据相比正常样本数据最大偏差为 35.7%,存在一定偏离,能够起到干扰效果。

从图 7 可以看出,存在干扰数据的融合曲线依然具有较好的光滑度,并且与无干扰数据融合曲线趋势相一致,由此证明采用本研究所提算法在遇到异常值情况下具有较好的稳健性。

4 结论

针对温室环境信息多节点数据采集中存在误差大、冲突多和冗余等问题,本研究提出了一种基于小波降噪和自适应加权相结合的多传感器数据融合算法。通过实例对所提算法的可行性进行了验证,并且与算术平均法的融合效果,以及加入干扰

数据后的融合效果逐一进行了对比,结果表明采用本研究所提算法可以有效地降低原始数据中存在的噪声和冗余,得到方差较小的数据融合值,提高测量数据的准确性和降低数据的传输量;具备较好的稳健性,能够实现对温室环境信息的可靠性和一致性描述,为后续温室环境建模提供可靠的数据。

参考文献:

[1]毛罕平,晋 春,陈 勇. 温室环境控制方法研究进展分析与展望[J]. 农业机械学报,2018,49(2):1-13.  
[2]赵春江. 精准农业研究与实践[M]. 北京:科学出版社,2009.  
[3]张雪花,张 武,杨 旭,等. 农业温室环境控制方法研究综述[J]. 控制工程,2017,24(1):8-15.  
[4]孙力帆,张雅媛,郑国强,等. 基于 D-S 证据理论的智能温室环境控制决策融合方法[J]. 农业机械学报,2018,49(1):268-275.  
[5]Chen Y K, Si X C, Li Z G. Research on Kalman - filter based multisensor data fusion[J]. Journal of Systems Engineering and Electronics,2007,18(3):497-502.

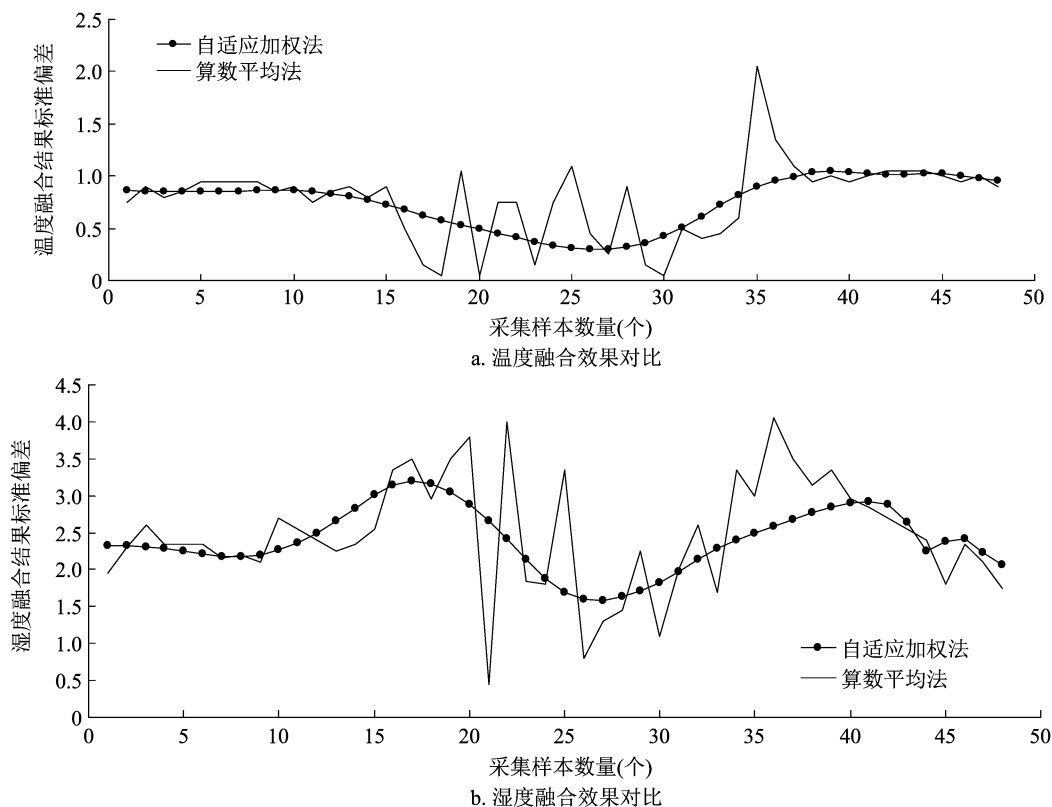


图5 温度、湿度数据融合结果标准偏差对比

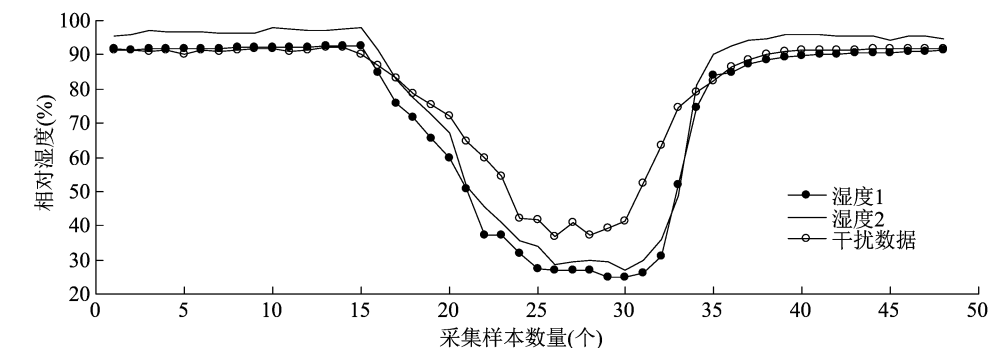


图6 正常数据与干扰数据曲线

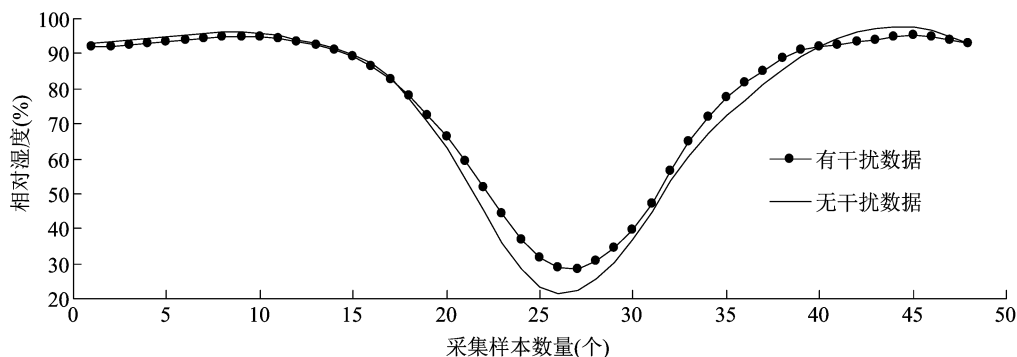


图7 有干扰数据的融合结果对比

宋林鹏,刘世洪,王 翠. 基于词向量 + BiLSTM + CRF 的农业技术需求文本实体提取[J]. 江苏农业科学,2021,49(5):186-193.  
doi:10.15889/j.issn.1002-1302.2021.05.034

# 基于词向量 + BiLSTM + CRF 的农业技术需求文本实体提取

宋林鹏,刘世洪,王 翠

(中国农业科学院农业信息研究所/农业农村部农业信息服务技术重点实验室,北京 100081)

**摘要:**农业技术需求文本实体提取是农业技术转移工作中需求分析能够量化、特征化的关键技术。基于主流农业技术转移网站上的农业技术需求文本进行分析,提出农业技术需求文本中需求主体对象、需求意图、需求程度 3 种关键实体成分;分别使用传统的 CRF 和基于神经网络的词向量 + BiLSTM + CRF 2 种模型对关键实体进行识别提取。试验结果表明,词向量 + BiLSTM + CRF 在正确率、召回率指标上分别能达到 88.51% 和 82.28% 以上,均优于 CRF 模型。

**关键词:**农业技术需求挖掘;命名实体识别;条件随机场;双向长短时记忆网络

**中图分类号:** S126 **文献标志码:** A **文章编号:** 1002-1302(2021)05-0186-08

农业创新链是一个以农业研发活动为基础,以农业科技成果转化为渠道,以实现农业科技成果产业化应用、市场化推广为目的的产业链<sup>[1]</sup>。在这个产业链中,农业科技成果转化起到了枢纽的作用,其上游是农业研发活动的主体,包括各大高校和科研院所,下游是进行直接农业生产的农民或农业企业,他们希望应用农业技术来改进生产活动,提高生产效益,但由于经济等方面的限制,只能借助于上游研发机构的成果,而农业科技成果转化正是起到了“中间人”的作用,其效果发挥的质量影响着整个产业链的通达程度。农业科技转移平台正是农业科技成果转化的承载实体,完善农业科技转移平

台建设是破解农业科技成果转化难的关键。

当前,我国农业技术转移主要依赖政府主导,农机推广人员作为技术供需双方的中间人,进行线下推广。这种方式存在成本高、效率低、时效性差等问题。一方面,农民或农企(特别是中小农企)对农业技术获取渠道、技术应用状况、收益等方面的信息掌握十分薄弱,许多农企当前依然依靠传统技术甚至过度的人力资源投入进行生产经营,导致农业生产过程成本高昂、效益低下;另一方面,许多高校和科研机构的成果应用不到实际的生产上,形成了技术拥有者和技术需求者脱节的现象。农业技术本质上也是商品,会受到市场规律的影响,市场的需求是主导市场资源配置和生产的主导力量,农业技术的市场需求同样也是农业技术转移市场的主导因素。许多学者也对当前农业技术转移机制脱离市场需求的问题进行了深入研究,并给出了相应的对策,但这些研究大多是从政策、理论方面给出建议<sup>[2-4]</sup>,很少有在具体技术层面提出优化农业技术需求信息挖掘的。

收稿日期:2020-05-11

基金项目:中国农业科学院科技创新工程项目(编号:CAAS-ASTIP-201X-A11)。

作者简介:宋林鹏(1994—),男,云南曲靖人,硕士研究生,主要从事信息技术与数字农业相关研究。E-mail:kuafu1994@163.com。

通信作者:刘世洪,博士,研究员,主要从事农业农村信息化相关研究。E-mail:liushihong@caas.cn。

[8]王 振,白星振,马梦白,等. 一种基于数据预处理和卡尔曼滤波的温室监测数据融合算法[J]. 传感技术学报,2017,30(10):1525-1530.

[9]温 静. 多传感器数据融合技术在温室控制中的应用[J]. 河北农业大学学报,2015,38(4):111-115.

[10]王宪磊,裴玖玲,刘润泽,等. 新疆地区智能温室控制系统研究——基于多传感器融合机理[J]. 农机化研究,2017,39(7):39-42,57.

[11]田胜利,周拥军,葛修润,等. 基于小波分解的建筑物变形监测数据处理[J]. 岩石力学与工程学报,2004,23(15):2639-2642.

[12]邸继征. 小波分析原理[M]. 北京:科学出版社,2010.

[13]翟翌立,戴逸松. 多传感器数据自适应加权融合估计算法的研究[J]. 计量学报,1998,19(1):69-75.

[14]孙国祥. 基于多信息融合的温室黄瓜肥水一体化灌溉系统研究[D]. 南京:南京农业大学,2016.