

缺失数据下智能温室控制模型的数据插补技术研究

陈秀禹

(大连东软信息学院高等职业技术学院, 辽宁 大连 116023)

✉chenxiuyu@neusoft.edu.cn



摘 要: 针对智能温室控制模型中, 实时温度数据因通信和设备故障等问题造成的数据缺失现象, 提出了在常规神经网络模型基础上, 利用模糊控制进行补偿的插补方法。利用搭建在温室大棚的智能监控系统对人工温室中的温度参数进行采集, 并利用所测数据对上述模型进行插补验证实验。实验结果表明: 与通用模型相比, 所提出的模糊神经网络插补模型改变了传统处理方式中插补数据不精确的现状, 为实时温度数据缺失提供了有效地处理方法, 也为建立智能温室模型提供了数据基础。

关键词: 缺失数据; 智能温室; 数据插补; 模糊神经网络

中图分类号: TP399 **文献标识码:** A

Research on Data Interpolation Technology of Intelligent Greenhouse Control Model under Missing Data

CHEN Xiuyu

(Higher Vocational Technical College, Dalian Neusoft University of Information, Dalian 116023, China)

✉chenxiuyu@neusoft.edu.cn

Abstract: Aiming at real-time temperature data missing caused by communication and equipment failure in intelligent greenhouse control model, this paper proposes an interpolation method based on conventional neural network model and fuzzy control for compensation. An intelligent monitoring system built in the greenhouse is used to collect temperature parameters in the artificial greenhouse, and the measured data is used to perform an interpolation verification experiment on the above model. The experimental results show that: compared with the general model, the proposed fuzzy neural network interpolation model improves the precision of interpolation data in traditional processing methods, provides an effective processing method for real-time temperature data missing, and lays a data foundation for the establishment of intelligent greenhouse model.

Keywords: missing data; intelligent greenhouse; data interpolation; fuzzy neural network

1 引言(Introduction)

准确的温室环境参数对建立智能温室模型, 实施温室种植环境的科学化调控, 防止环境异常变化造成的损失具有十分重要的经济和现实价值^[1-2]。所以, 一般情况下温室内需要长时间监测, 并用传感器将采集信息从温室环境接入互联网, 由于实时温度数据受通信和设备故障等问题影响, 因此数据会不可避免地出现缺失值。通常情况下, 如果缺失数据不多, 可通过直接删除的方式处理, 但更通用的方法是采

用各种方法对所缺失数据进行插补。目前, 针对数据缺失问题, 构建完整数据序列的方法可分为两大类: 单一插补方法和多重插补方法^[3]。这些方式往往都是借助相关统计学方法, 根据缺失数据的规律为缺失值确定一个适合的预测值, 使得数据完整, 在可以利用的现有信息量小、规律模糊时, 常常造成插补结果出现异常。近年来, 神经网络技术也被用来完成相关的数据插补工作^[4], 文中提出一种模糊神经网络温室插补控制模型, 为提高数据的完整性提供了全新的技术支持。

2 数据获取与模型建立(Data acquisition and model building)

准确地监测获取智能温室环境内的温度数据，是建立智能温室控制模型的关键，为收集真实可靠的温度参数进行智能温室控制模型的实验验证。将温室智能监测系统安装于搭建的温室之内，监测温室内的温度参数，考虑到温度系统的一阶特性，设置采样周期为1小时，智能温室控制模型结构如图1所示。系统依照主流物联网架构实现，基本结构分为感知层、网络层、应用层。其中感知层通过温度传感器获取温度信息，并将这些温度数据通过WIFI传给服务器存储，供数据模型建立和插补处理使用。

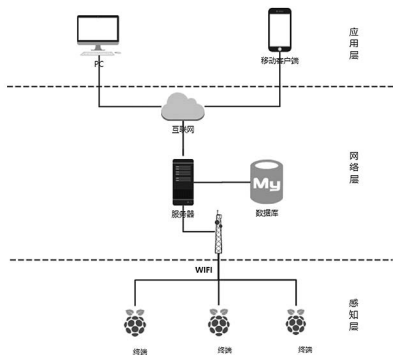


图1 智能温室控制模型结构图

Fig.1 Structure diagram of intelligent greenhouse control model

3 智能温室控制模型数据插补技术研究(Research on data interpolation technology of intelligent greenhouse control model)

一般情况下，温度数据的缺失属于随机缺失模式，缺乏相应的统计规律进行单一插补操作。

3.1 BP神经网络

BP是一种误差反向传播神经网络，有很强的学习及处理非线性问题的能力，并可利用样本学习，对信息间的内在关系进行仿真，以逼近实际数据系统^[5]。

传统的三层BP神经网络结构图如图2所示。

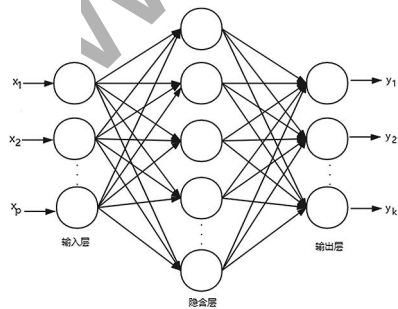


图2 BP神经网络结构图

Fig.2 Structure diagram of BP neural network

其基本学习过程主要分为以下四部分：模式顺序传播、误差逆传播、记忆训练和学习收敛。标准BP神经网络学习算

法流程图如图3所示。

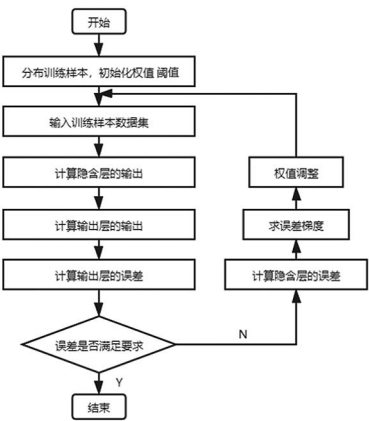


图3 标准BP神经网络学习算法流程图

Fig.3 Flow chart of standard BP neural network learning algorithm

3.2 BP神经网络时间识别序列插补模型

神经网络用于数据插补的主要方法是用完整的温度数据集训练网络，将缺失变量Y之外的其他变量的温度数据作为网络的输入，将缺失温度Y作为网络输出。待网络训练满足要求后，把其他测试温度数据作为网络的输入，网络输出值即为真正缺失数据的预测值。

本文为了增强神经网络对时间序列的识别性能，构造了双向时间识别序列^[4]，也就是采用缺失时间段前后已有温度数据共同对缺失数据进行预测，提高神经网络拟合准确性。神经网络计算过程如下：

- (1)取本日24小时温度数据，随机缺失数据6个，缺失率为25%。
- (2)假设本日时间窗口存在 $2n+1$ 个时间段， $n=1$ ，温度数据的输入量为缺失数据的前后 $2n$ 个时间段，共四个数据。输出量为缺失数据。
- (3)根据Kolmogorov定理，若BP神经网络构建中间层数目 n_2 输入层数目 n_1 ，则 $n_1=\log_2 n_2$ ，故隐含层节点数为16。
- (4)选取输入层到中间层神经元的传递函数为tansig函数，选取输出层神经元传递函数也为tansig函数。

采用如表1所示一天的温度数据对BP神经网络进行训练。

表1 训练数据表(以一天24小时为例)

Tab.1 Training data table(Take 24 hours a day as an example)

IN1	IN2	IN3	IN4	01
20	21	21	20	21
20	19	18	20	19
20	17	14	20	17
20	20	20	20	20
17	19	18	20	19
22	23	24	21	23

完成已有样本训练后，以另一天的24小时温度数据作为输入送入测试网络，输出为缺失值，设检验误差指标 $E=(\text{温度实际值}-\text{神经网络输出值})$ 。如果 E 值不高于1，则认为训练结果可靠，从而实现对缺失值的预测，完成温度数据的插补。

训练结果如图4所示，由图可见，单纯神经网络训练的结果在几点上有明显偏差，误差较大，难于直接用于数据插补。

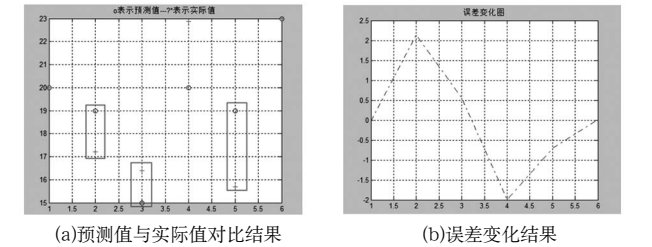


图4 神经网络测试结果

Fig.4 Neural network test results

3.3 改进的模糊神经网络模型

神经网络模型虽然能够进行非线性的数据插补，但在处理极端无规律信息上，整个网络难于控制，所以往往在温度发生特殊变化时难于达到最优插补效果。

模糊控制依靠专家经验建立模糊集合、隶属函数并通过模糊推理得到控制结果，其推理规则主要由专家从实践中获得，难于复制^[6]。

目前神经网络和模糊控制结合的方式，一般分三种：利用神经网络的运行结果驱动模糊控制器；利用神经网络来完成模糊规则的记忆；或利用神经网络优化模糊控制器参数^[7-8]。

根据智能温室数据插补基本的需求，采用神经网络运行结果驱动模糊控制器方式建立改进的模糊神经网络插补模型，步骤如下。

(1)模糊规则制定：对确定的温室历史同期数据温度规律进行模糊化，对输出模糊量进行解模糊判决(即通过各种解模糊方法，完成由模糊量到精确量的转化)，形成对神经网络温度模型的增量，对神经网络插补模型进行补偿。

(2)结果比较：分别采用未加入模糊补偿量的神经网络和加入模糊补偿量的神经网络进行同期数据的训练和测试，分析结果。

下面介绍模糊规则的制定。

(1)输入量 X_1 ：历史同期温度误差

为了分析历史同期温度数据的规律，结合温度的历史极值，建立历史同期温度误差的基本论域 X_1 为 $[-6, 6]$ ，设定温室系统温度误差模糊集合的论域为 $\{-6, -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, 6\}$ ，可得温度的量化因子 $K_1=1$ ，把温度误差的模糊量分为七个等级：负大、负中、负小、零、正小、正中、正大，用英文缩写为：NB、NM、NS、ZO、PS、PM、PB，可建立温度隶属表如表2所示。

PB，可建立温度隶属表如表2所示。

表2 历史同期温度隶属表

Tab.2 Historical temperature membership table for the same period

模糊量	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6
NB	1	0.5											
NM		0.5	1	0.5									
NS				0.5	1	0.5							
ZO						0.5	1	0.5					
PS									0.5	1	0.5		
PM											0.5	1	0.5
PB												1	0.5

(2)输入量 X_2 ：本期温度误差

为了分析本采样周期温度是否存在突变，确定本测试窗口期温度误差的基本论域 X_2 为 $[-4, 4]$ ，设定控制系统温度误差模糊集合的论域为 $\{-4, -3, -2, -1, 0, 1, 2, 3, 4\}$ ，可得温度的量化因子 $K_1=1$ ，把温度误差的模糊量分为五个等级：负大、负小、零、正小、正大，用英文缩写为：NB、NS、ZO、PS、PB。同理也可建立本期温度隶属表。

(3)输出量

利用历史同期温度得出本年度温度变化基本趋势，利用本期数据得出本日温度变化程度，得出本期已存在数据是否存在极端温度的变化。选取输出量为插补数据增幅补偿值，将增量分成五个等级：减少大量、减少少量、不增加、增加少量、增加大量。所以其对应的模糊语言为NB、NS、ZO、PS、PB。设定论域 $\{-2, -1, 0, 1, 2\}$ ，可建控制规律隶属表，如表3所示。

表3 控制规律隶属表

Tab.3 Membership table of control law

X_1	X_2	Y
NB	NB	-2
NB	NS	-1
NB	ZO	0
NB	PS	1
NB	PB	2

此模糊控制的输出量涵盖了神经网络无法说明的人类专家经验，利用这部分输出作为神经网络输出变量的补偿，就可以在现有数据基础上增加专家经验，避免插补数据出现极值。

4 实验数据分析(Experimental data analysis)

选择日温度作为输入，将随机缺失温度数据作为输出，分别采用两种技术实现智能温室数据的插补处理，两种插补技术效果对比如表4所示(结果1对应神经网络，结果2对应模糊神经网络)。可以看出模糊神经网络插值结果较单纯的神经网络

络明显更优，这主要是由于神经网络插补只借助了实测样本中的片段数据，无法体现温度数据的整体特点及趋势，不如模糊神经网络插补技术更能保证插值效果的可靠性。

表4 两种插补效果对比表

Tab.4 Comparison table of two insertion effects

实际值	插值结果1	插值结果2	相对误差1	相对误差2
20	20	20	0	0
19	17.2095	18.2095	-1.7905	0.07905
19	16.3927	18.3927	-2.6073	-0.6073
21	22.8733	21.8733	1.8733	0.8733
18	15.6910	17.6910	-2.309	-0.309
23	22.9866	22.9866	-0.0134	-0.0134

5 结论(Conclusion)

利用模糊控制技术和神经网络建立的智能温室缺失数据补偿模型，很好地结合了人类专家经验和神经网络的超强拟合能力，有效解决了温度缺失下智能温室模型中的数据插补问题，此模型在温室控制中的运用得到了较好的控制结果。

参考文献(References)

[1] 孙占鹏,李佳,欧文.多传感器室内环境监测系统[J].传感器与

(上接第36页)

检测速率，满足了目标检测中对检测速度和检测精度的一般要求，证明本文提出的算法在交通图像识别领域具有实用和参考价值。

参考文献(References)

[1] WANG Y Q. An analysis of the Viola-Jones face detection algorithm[J]. Image Processing On Line, 2014, 104(4):128-148.
[2] TORRIONE P A, MORTON K D, RAYN S, et al. Histograms of oriented gradients for landmine detection in ground-penetrating radar data[J]. IEEE Transactions on Geoscience and Remote Sensing, 2013, 52(3):1539-1550.
[3] OUYANG W, ZENG X, WANG X, et al. DeepID-Net: Object detection with deformable part based convolutional neural networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(7):1320-1334.
[4] CHEN L L, WEN L F, BIAN X, et al. Fast single shot multibox detector and its application on vehicle counting system[J]. IET Intelligent Transport Systems, 2018, 12(10):1406-1413.
[5] GE Z, JIE Z, HUANG X, et al. Delving deep into the imbalance of positive proposals in two-stage object detection[J]. Neurocomputing, 2021(425):107-116.

微系统,2017,36(1):87-90.

[2] 张静,石煜,杨继森,等.室内环境智能控制系统设计[J].实验室研究与探索,2016,35(7):65-69.
[3] 王志良,黄珊,陈海涛.黄河流域水文数据插补方法比较及应用[J].人民黄河,2020,42(7):14-18.
[4] 蒋丽丽,姜大庆.基于BP神经网络的农资库存数据插补技术[J].江苏农业科学,2018,46(20):268-271.
[5] JIA W K, ZHAO D, SHEN T, et al. An optimized classification algorithm by BP neural network based on PLS and HCA[J]. Applied Intelligence, 2015, 43(1):176-191.
[6] 缪志农.基于状态空间方法的自适应模糊控制研究[D].四川:西南交通大学,2006.
[7] 周红标,张钰,柏小颖,等.基于自适应模糊神经网络的非线性系统模型预测控制[J].化工学报,2020,71(7):3201-3212.
[8] 刘帅,盛金保,王昭升,等.基于模糊神经网络的水电施工安全隐患评价[J].水利水运工程学报,2020(1):105-111.

作者简介:

陈秀寓(1977-),女,硕士,副教授.研究领域:多种教学模式及其应用,信息教学.

[6] HAN X F. Modified cascade RCNN based on contextual information for vehicle detection[J]. Sensing and Imaging, 2021, 22(1):1-10.
[7] 闫新庆,杨喻涵,陆桂.基于改进Faster-RCNN目标检测算法研究[J].中国新通信,2021,23(01):46-48.
[8] 孙乾宇,张振东.基于YOLOv3增强模型融合的人流密度估计[J].计算机系统应用,2021,30(04):271-276.
[9] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[J]. Computer Vision, 2016, 9905(12):21-37.
[10] 王冠,耿明洋,马勃檀,等.基于孪生区域候选网络的目标跟踪模型[J].小型微型计算机系统,2021,42(04):755-760.
[11] GEIGER A, LENZ S, STILLER C, et al. Vision meets robotics: The KITTI dataset[J]. The International Journal of Robotics Research, 2013, 32(11):1231-1237.

作者简介:

章康明(1998-),男,本科生.研究领域:软件开发,计算机视觉.
曹新(1980-),女,硕士,教授.研究领域:车联网,无线传感器网络.