

# 基于集合经验模态分解与 Elman 神经网络的线椒株高预测

张 瑜, 汪小昆\*, 孙国祥, 李永博

(1. 南京农业大学工学院, 南京 210031; 2. 江苏省现代设施农业技术与装备工程实验室, 南京 210031)

**摘 要:** 为提高温室环境控制系统的有效性, 针对作物生长量的变化与环境因子的变化存在时间尺度不统一的问题, 该文基于集合经验模态分解 (ensemble empirical mode decomposition, EEMD) 与 Elman 神经网络建模, 提出一种线椒株高生长量预测方法。以 8819 线椒为试验对象, 分别对线椒株高及其环境因子进行 EEMD 分解, 对各尺度下的时间序列建立 EEMD-Elman 预测模型。结果表明: 应用 EEMD-Elman 神经网络建立线椒株高生长量预测模型, 模型预测值与实测值的平均绝对误差为 1.69 cm, 相关决定系数为 0.996, 标准误差为 1.104, 模型预测结果与实测值呈极显著性相关。研究结果可以解决作物生长变化与环境变化时间尺度不统一的问题, 为温室环境控制系统的控制目标的优化提供有效参数。

**关键词:** 神经网络; 模型; 优化; 集合经验模态分解; Elman 神经网络; 线椒株高; 预测

doi: 10.11975/j.issn.1002-6819.2015.18.024

中图分类号: S625.5

文献标志码: A

文章编号: 1002-6819(2015)-18-0169-06

张 瑜, 汪小昆, 孙国祥, 李永博. 基于集合经验模态分解与 Elman 神经网络的线椒株高预测[J]. 农业工程学报, 2015, 31(18): 169—174. doi: 10.11975/j.issn.1002-6819.2015.18.024 <http://www.tcsae.org>

Zhang Yu, Wang Xiaochan, Sun Guoxiang, Li Yongbo. Prediction of cayenne pepper plant height based on ensemble empirical mode decomposition and Elman neural network[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2015, 31(18): 169—174. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.2015.18.024 <http://www.tcsae.org>

## 0 引 言

作物生长量预测为温室环境调控目标优化提供重要依据<sup>[1-2]</sup>。目前, 温室作物生长量预测多通过建立作物外观形态模拟模型的方法实现<sup>[3-6]</sup>。如, calvalho 等<sup>[7]</sup>, Larsen 等<sup>[8]</sup>, Heuvelink 等<sup>[9]</sup>, Nothnagl 等<sup>[10]</sup>分别建立了菊花生长模拟模型, 对菊花的株高、花朵数、花朵大小等生长量进行预测。上述模型研究的最直接意义是及时掌握温室环境控制系统对作物所产生的作用, 为温室环境优化控制提供理论依据。然而, 作物生长模拟模型无法真实地反映作物的实际生长情况, 使得环境控制系统无法根据作物实际生长情况做出有效的响应, 降低了温室环境控制系统的有效性。所以, 根据作物实际生长情况对作物生长量进行预测, 可以更真实地了解环境控制的效果<sup>[11]</sup>, 提高温室环境控制系统的有效性<sup>[12]</sup>。

根据作物实际生长情况对环境进行调控是提高温室环境控制系统有效性的重要方法。但是作物生长量的变化与环境因子的变化存在时间尺度不统一的特点, 作物生长量的变化是缓慢的, 而环境因子的变化是瞬时的。

单纯采用平均值对两个时间尺度不统一的时间序列进行建模、预测, 无法清晰地反映两种时间序列在不同时间尺度下的波动情况, 造成预测结果不准确。因此, 需要对两者进行多时间尺度分解, 利用不同时间尺度下的时间序列特征进行预测与重构, 实现作物实际生长量的预测, 为温室环境控制系统提供实际有效的控制依据。

线椒是温室中常见种植作物, 株高是反映线椒长势的一项重要生长参数, 能够直观地反映出在温室环境控制系统对作物生长的作用效果。因此, 本文基于集合经验模态分解 (ensemble empirical mode decomposition, EEMD) 与 Elman 神经网络, 对线椒株高及其环境参数时间序列进行分解与重构, 分别对分解后的时间序列建立非线性 Elman 神经网络, 根据种植期间, 线椒株高当前的实际生长量, 重构并预测线椒株高在同一生长期内的生长量, 为温室作物生长预测提供精确、实时的预测方法, 为温室环境控制系统控制目标的优化提供实时有效的控制依据, 实现提高温室环境控制系统有效性的目的。

## 1 数据与方法

### 1.1 试验数据采集

本文试验对象为 2014 年种植的 8819 线椒(定植后生长期, 2014 年 5 月 28 日至 2014 年 7 月 8 日)及对应的温室内环境因子(温室内温度、相对湿度、室内总辐射)。试验于江苏省南京市南京农业大学工学院, 博远楼顶 venlo 型玻璃温室内进行。温室环境因子由 Eco-Watch 生态网络监测系统每 15 min 采集一次, 取日平均值用于模

收稿日期: 2015-06-04 修订日期: 2015-08-24

基金项目: 国家自然科学基金项目 (61273227)、江苏省自然科学基金 (BK20150686)

作者简介: 张瑜, 女 (汉), 吉林长春人, 讲师, 博士, 农业电气化与自动化、检测技术与自动化装置。南京 南京农业大学工学院, 210031。

Email: zhangyu64@hotmail.com

\*通信作者: 汪小昆, 男 (汉), 教授, 博士, 从事农业生物环境工程、设施农业温室环境模拟与调控方面研究。南京 南京农业大学工学院, 210031。

Email: wangxiaochan@njau.edu.cn

型建立。由于温室内温度湿度等环境因子分布的不均匀性,所以本文选取在温室内东面密集种植的 16 株线椒作为试验样本,以降低温室内不同小气候条件对作物生长形态的影响。作物均生长在适宜的环境条件下,未发生环境胁迫与营养胁迫情况。其中 10 株为测试集,用于建立模型,另外 6 株为验证集,用于模型验证。考虑到作物的个体差异,均取其平均值。网络的输入量为温室内环境因子的日平均值,输出量为测试集中 10 株线椒株高的每日测量值的平均值,测量方法采用文献[13]中线性激光扫描法。线椒采用椰糠基质,独立种植于栽培袋中。灌溉营养液配方采用山东农业大学番茄辣椒配方,  $E_c$  值为 2.0, pH 值为 6.5。

## 1.2 分解方法

在解决时间序列多尺度分解的方法中,小波分解和奇异谱分解等方法均需要依赖于预先给定的基函数,无法根据数据自身的时间尺度特征进行分解,可能会由此分解出虚假或无效的波动分量<sup>[14-20]</sup>。集合经验模态分解方法利用高斯白噪声具有频率均匀分布的统计特性,无需预先设定任何基函数,自适应地将数据分解为有限个本征模函数(intrinsic mode function, IMF),所分解出来的各 IMF 分量包含了原始数据的不同时间尺度的局部特征信号<sup>[21-26]</sup>。所以,本文采用 EEMD 分解方法对试验数据(温室环境因子与线椒株高生长量)进行多尺度分解,分解结果可以由公式(1)求得,重构结果由公式(2)求得。

$$C_j(t) = \frac{1}{N} \sum_{i=1}^N C_{ij}(t) \quad (1)$$

$$X(t) = \sum_{j=1}^n C_j(t) + r_n \quad (2)$$

式中:  $C_{ij}(t)$  表示第  $i$  次加入白噪声后分解所得的第  $j$  个 IMF 分量,  $t$  表示时间;  $C_j(t)$  为分解得到的第  $j$  个 IMF 分量;  $N$  为添加白噪声序列的数目;  $X(t)$  为原始信号;  $r_n$  为分解后的残留余项,  $n$  表示分解次数,当  $r_n$  为单调信号时,分解终止。

## 1.3 预测方法

Elman 神经网络是一种带有反馈的两层 BP 网络结构,其反馈连接是从隐含层的输出到其输入端。这种反馈方式使得 Elman 网络能够探测和识别时变模式。其隐含层又称为反馈层,这种特殊的两层网络可以任意精度逼近任意函数,唯一的要求是其隐含层必须具有足够的神经元数目。隐含层神经元数越多,则逼近复杂函数的精度就越高<sup>[27-28]</sup>。Elman 神经网络的非线性表达原理<sup>[29]</sup>如式(3)所示。

$$\begin{cases} x_c(k) = x(k-1) \\ x(k) = f(w_2(u(k-1)) + w_1 x_c(k)) \\ y(k) = g(w_3 x(k)) \end{cases} \quad (3)$$

式中:  $y$ 、 $x$ 、 $u$ 、 $x_c$  分别表示  $m$  维输出节点向量、 $n$  维中间层节点单元向量、 $l$  维输入向量和  $n$  维反馈状态向量;  $w_3$ 、 $w_2$ 、 $w_1$  分别表示中间层到输出层、输入层到中间层、承接层到中间层的连接权值。  $g$  为输出神经元的传递函数,是

中间层输出的线性组合。  $f$  为中间层神经元的传递函数。

## 2 结果与分析

### 2.1 EEMD 分解与重构

图 1 为原始时间序列。根据式(1)本文分别对线椒株高及其环境因子(温室温度、相对湿度、室内总辐射)进行 EEMD 分解,分别得到了 5 个本征模函数,分别为 IMF1~IMF5,如图 2 所示。本征模函数表征了数据的内在的振动模式。图 2 中,IMF1~IMF4 反应了线椒株高和各个环境因子在不同频率下振荡情况,IMF5 为变化趋势项,反应了线椒株高在生长期内的整体变化趋势。EEMD 分解重构后结果的相关参数如表 1 所示。由表 1 可知,各个环境因子和线椒株高时间序列经过 EEMD 分解重构后,与原始时间序列相关系数在有效数字范围内近似为 1,标准误差接近于 0,平均绝对误差接近于 0,表明重构结果十分有效。

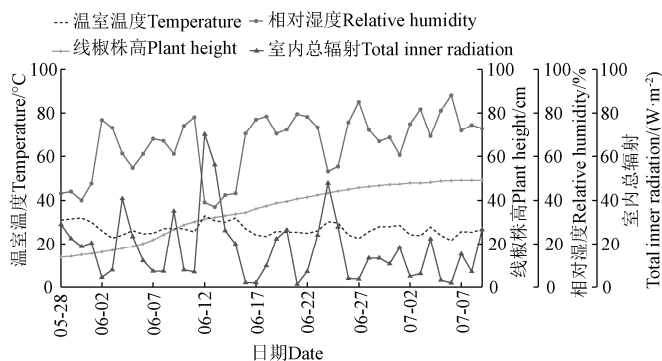


图 1 原始时间序列

Fig.1 Original time series

### 2.2 Elman 神经网络生长量预测与分析

本文 Elman 网络的输入层为温室内环境因子,包括温室温度、相对湿度和室内总辐射,输入节点数为 3;输出层为线椒株高生长量的预测值,节点数为 1。本文采用双层反馈结构,第一层节点数为 10,第二层节点数为 3。反馈层神经元的传递函数为  $\text{tansig}()$ ,线性输出层的传递函数为  $\text{purelin}()$ 。网络训练样本组按照 0.7:0.15:0.15 的比例划分为训练集、验证集和测试集。

为避免神经元饱和,在输入层对输入数据进行归一化处理,将各个数值换算至[0 1]区间内,在输出层对得到的预测结果进行反归一化。根据 Elman 神经网络的非线性空间表达式(公式(3)),本文网络模型的输出结果可以表示为:

$$\begin{cases} x_1(k) = f(w_{12}x_1(k-1) + w_{11}(u(k-1))) \\ x_2(k) = f(w_{22}x_2(k-1) + w_{21}(x_1(k))) \\ y(k) = g(w_3x_2(k)) \end{cases} \quad (4)$$

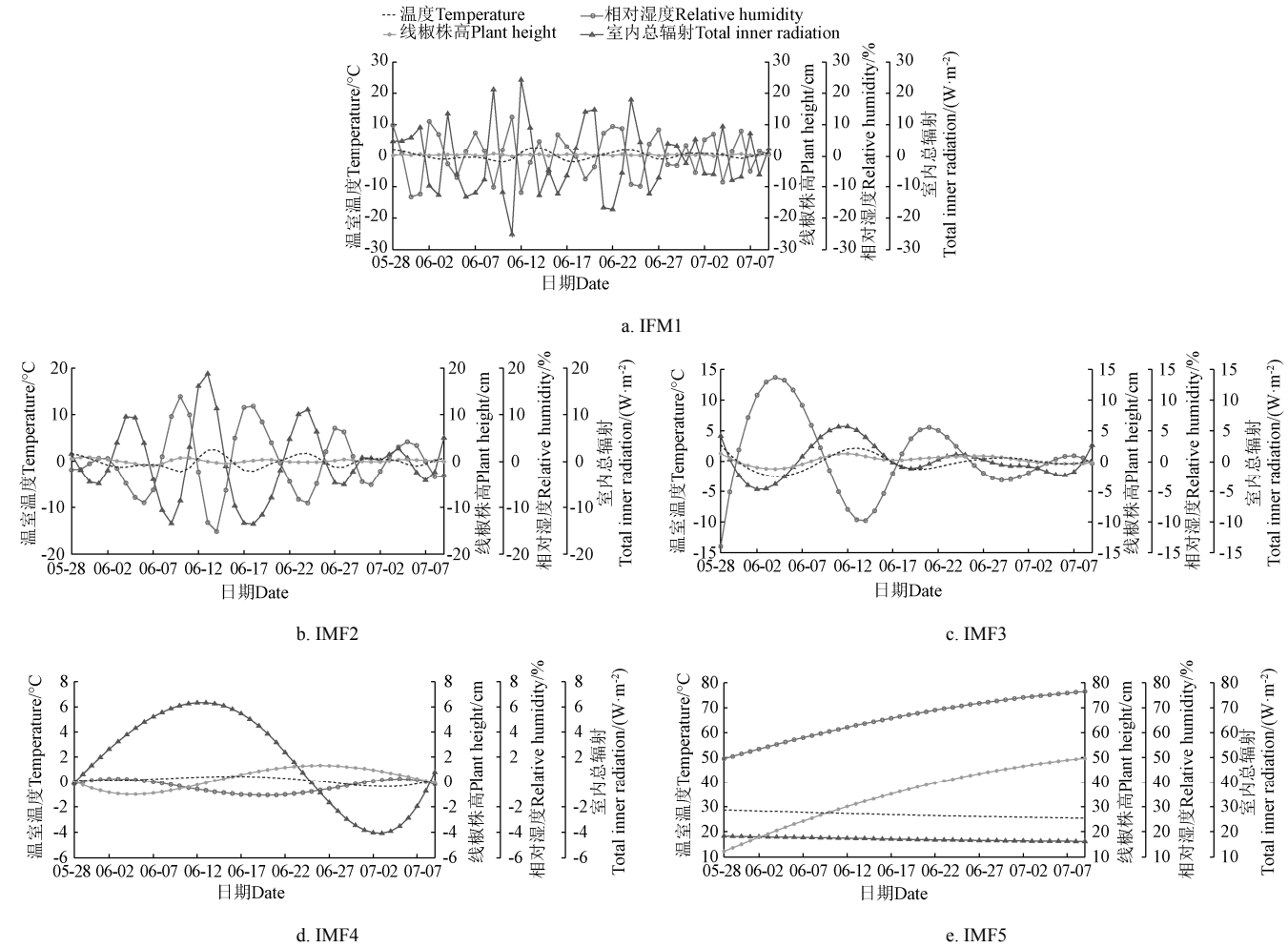
式中:  $u(k-1)$  为输入层的输入;  $x_1(k)$  为第 1 层反馈的输出;  $x_2(k)$  为第 2 层网络的输出,  $y(k)$  为输出结果;  $w_{p1}$  ( $p=1,2$ ;  $p$  为网络层数) 为第 1 层网络和第 2 层网络的承接层到中间层的连接权值;  $w_{p2}$  为第 1 层网络和第 2 层网络的输入层到中间层的连接权值;  $w_3$  为中间层到输出层的连接权值,本文中间层神经元传递函数采用双曲正切传递函数

tansig(), 输出层神经元采用线性传递函数 purelin()。由双曲正切函数定义, 可得到本文预测网络模型中间层传递函数表达式为:

$$f = \frac{e^{2[w_{p2}u(k-1)+w_{p1}x(k-1)]} - 1}{e^{2[w_{p2}u(k-1)+w_{p1}x(k-1)]} + 1}$$

(5)

式中:  $p=1,2$ ; 表示网络层数。



注: IMF1~IMF4 反应了线椒株高和各个环境因子在不同频率下的振荡情况, IMF5 反应了线椒株高和各个环境因子的变化趋势。  
Note: IMF1-IMF4 reflect the oscillation of variate, including plant height and every environment elements under different frequencies. IMF5 reflects the trend of plant height and every environment elements.

图 2 线椒株高与环境因子 EEMD 分解后各模态分量波动情况  
Fig.2 Oscillation of EEMD results of plant height and environment elements

表 1 EEMD 分解重构结果分析			
Table 1 Summary of EEMD and reconstruction			
时间序列 Time series	决定系数 $R^2$	标准误差 Standard error	平均绝对误差 Mean absolute error
线椒株高 Plant height	1.0000	$2.708 \times 10^{-3}$	$4.58 \times 10^{-5}$ cm
温室温度 Temperature	1.0000	$4.73 \times 10^{-2}$	$3.97 \times 10^{-4}$ °C
相对湿度 Relative humidity	1.0000	$2.516 \times 10^{-3}$	$5.10 \times 10^{-5}$ %
室内总辐射 Total inner radiation	1.0000	$3.445 \times 10^{-3}$	$7.20 \times 10^{-5}$ W·m <sup>-2</sup>

本文中时间序列样本经过 EEMD 分解后, 得到 5 个独立的本征模函数 IMF1~IMF5, 所以需要建立 5 个子 Elman 神经网络, 分别预测生长量的各个时间序列值, 最后经过 EEMD 重构得到最终的线椒株高生长量预测值。具体预测流程如图 3 所示。

各子网络的预测结果统计性描述如表 2 所示, 各个

子网络的预测值与目标值之间相关决定系数分别为 0.71、0.88、0.94、0.95 和 0.98, 平均绝对误差值分别为 0.04、0.09、0.22、0.14 和 1.20 cm。由表 2 可知, 各尺度预测值与实测值的模型拟合效果均表现出极强的相关性, 平均绝对误差均可满足测量精度要求。随着时间尺度的增加, 模型相关性决定系数逐渐增强, 平均绝对误差逐渐增大, IMF5 子网络预测结果的平均绝对误差增大最明显, 可能的原因是 IMF5 为趋势项, 对于变量的波动情况表达不明显。

本文对 IMF1~IMF5 子网络的预测结果进行重构。重构后, 线椒株高预测值与实际测量值之间的拟合关系如图 4 所示。结果表明: 通过 EEMD 多尺度分解, 对线椒株高在不同时间尺度上应用 Elman 神经网络预测再合并, 预测值与测量值达到极显著相关水平, 平均绝对误差值为 1.69 cm, 可以满足温室中作物生长测量的精度要求。

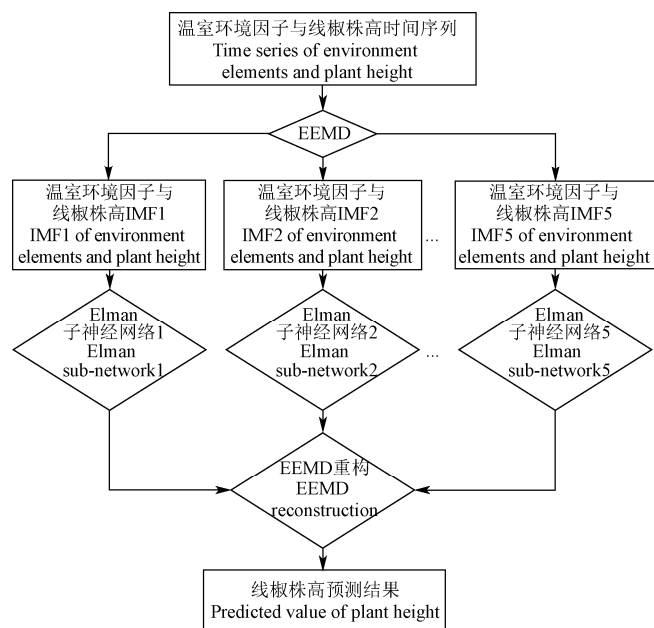


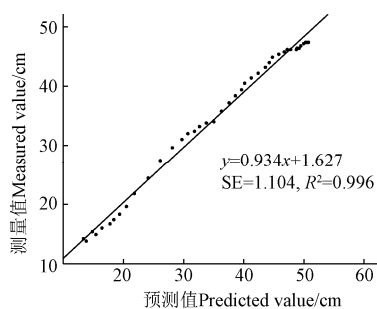
图3 线椒株高预测流程图

Fig.3 Flowchart of pepper plant height prediction

表2 Elman 网络预测线椒株高结果统计分析

Table 2 Summary of Elman neuron network pepper plant height prediction

网络类型 Type of network	决定系数 $R^2$	平均绝对误差 Mean absolute error/cm
Elman 子网络_IMF1 Sub-network_Elman_IMF1	0.71	0.04
Elman 子网络_IMF2 Sub-network_Elman_IMF2	0.88	0.09
Elman 子网络_IMF3 Sub-network_Elman_IMF3	0.94	0.22
Elman 子网络_IMF4 Sub-network_Elman_IMF4	0.95	0.14
Elman 子网络_IMF5 Sub-network_Elman_IMF5	0.98	1.20



注: SE 为标准误差,  $R^2$  为决定系数。

Note: SE is standard error,  $R^2$  is coefficient of determination.

图4 线椒株高分解重构预测结果

Fig.4 Results of pepper plant height prediction

分别应用 EEMD 与 BP 神经网络结合方法<sup>[30]</sup>、EEMD 与 Elman 神经网络结合方法和未经 EEMD 分解的 Elman 神经网络方法进行线椒株高生长量的预测。取试验样本验证集数据, 对所建模型进行验证。验证结果表明, 经过 EEMD 分解的网络模型的预测值和实测值进行比较, 其预测值与实测值均可较好的吻合, 未经过 EEMD 分解的 Elman 网络预测效果稍差。采用决定系数  $R^2$ 、标准误差

SE 和平均绝对误差两个指标分别对 3 个预测模型进行评价, 各指标统计结果见表 3。

表3 3种预测方法线椒株高结果评价

Table 3 Evaluation of three pepper plant height prediction methods

网络类型 Type of network	决定系数 $R^2$	标准误差 Standard error	平均绝对误差 Mean absolute error/cm
EEMD-BP 神经网络 EEMD-BP neural network	0.812	7.012	5.40
EEMD-Elman 神经网络 EEMD-Elman neural network	0.996	1.104	1.69
Elman 神经网络 Elman neural network	0.908	5.032	8.87

由表 3 可知, 经过 EEMD 分解后模型预测结果比未经过分解的模型预测结果准确性高, 经过 EEMD 分解后的 Elman 神经网络的预测结果优于 EEMD-BP 神经网络和未经过 EEMD 分解的 Elman 神经网络预测结果。同时, 由图 4 可知, EEMD-Elman 神经网络的预测结果决定系数  $R^2$  为 0.996, 表明 EEMD-Elman 神经网络输出值与目标值偏差较小, 是十分有效的预测方法。

### 3 结论

应用集合经验模态分解方法对线椒株高及其环境因子(温室温度、相对湿度、室内总辐射)进行多尺度分解与重构。重构后各个环境因子和线椒株高的时间序列与原始时间序列相关系数在有效数字范围内近似为 1, 标准误差接近于 0, 平均绝对误差极接近于 0, 重构结果十分有效。表明集合经验模态分解法(ensemble empirical mode decomposition, EEMD)分解可以解决多尺度时间序列预测问题。应用 EEMD-Elman 神经网络建立线椒株高生长量预测网络, 网络预测值与实测值的平均绝对误差为 1.69 cm, 相关决定系数为 0.996, 标准误差为 1.104, 表明预测结果与实测值呈极显著性相关。3 种预测方法的预测结果进行对比, 表明与未经过 EEMD 分解的预测网络相比, 经过 EEMD 分解的预测网络, 能够反映时间序列自身固有的特性, 更好地诠释时间序列在不同频率的振荡情况; 将经过 EEMD 分解的 EEMD-Elman 神经网络与 EEMD-BP 神经网络的预测结果相比, EEMD-Elman 神经网络预测效果优于 EEMD-BP 神经网络预测效果, 说明 Elman 的双反馈结构可以提高模型的预测精度。

由于本文仅对生长期内的线椒株高进行分析, 并未考虑线椒在幼苗期、开花期及坐果期的生长情况, 所以本文中预测网络仅适用于定植后线椒株高的预测。试验期间, 作物生长未经受任何环境胁迫与营养胁迫, 如外界环境发生较大变化时, 会对网络预测结果造成一定影响。所以, 该预测网络在适应性和稳定性方面存在一定局限性。后续研究应更多关注作物在不同生长阶段, 不同生长参数的预测研究以及环境胁迫与营养胁迫对预测结果精度带来的影响。

综上所述, 集合经验模态分解法可以根据时间序列自身时间尺度特征, 自适应地将时间序列分解到不同时间尺度上, 清晰地表达出原始时间序列在不同时间尺度

上的波动情况, 有效解决温室中环境因子与作物生长量时间尺度不统一的问题。应用 Elman 神经网络双层反馈结构可以提高时间序列的预测精度。EEMD 与 Elman 神经网络结合可以准确预测作物实际生长量, 为温室环境控制系统控制目标优化提供有效参数, 实现有效提高温室环境控制系统的控制效果的目的。

#### 参 考 文 献

- [1] Straten G, van Challa H, Buwalda F. Towards user accepted optimal control of greenhouse climate[J]. Computers and Electronics in Agriculture, 2000, 26(3): 221—238.
- [2] Pucheta J A, Schugurensky C, Fullana R, et al. Optimal greenhouse control of tomato seedling crops[J]. Computers and Electronics in Agriculture, 2006, 50(1): 70—82.
- [3] 倪纪恒, 罗卫红, 李永秀, 等. 温室番茄干物质分配与产量的模拟分析[J]. 应用生态学报, 2006, 17(5): 811—816. Ni Jiheng, Luo Weihong, Li Yongxiu, et al. Simulation of greenhouse tomato dry matter partitioning and yield prediction[J]. Chinese Journal of Applied Ecology, 2006, 17(5): 811—816. (in Chinese with English abstract)
- [4] 李永秀, 罗卫红, 倪纪恒, 等. 温室黄瓜干物质分配与产量预测模拟模型初步研究[J]. 农业工程学报, 2006, 22(2): 116—121. Li Yongxiu, Luo Weihong, Ni Jiheng, et al. Simulation of dry matter partitioning and yield prediction in greenhouse cucumber[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2006, 22(2): 116—121. (in Chinese with English abstract)
- [5] Ma Yuntao, Wen Meiping, Guo Yan, et al. Parameter optimization and field validation of the functional-structural model GREENLAB for maize at different population densities[J]. Annals of Botany, 2008, 101(8): 1185—1194.
- [6] Yan Hongping, Kang Mengzheng, de Reffye P, et al. A dynamic, architectural plant model simulating resource-dependent growth[J]. Annals of Botany, 2004, 93(5): 591—602.
- [7] Carvalho S M P, Heuvelink E, Cascals R, et al. Effect of day and night temperature on internode and stem length in chrysanthemum: Is everything explained by DIF? [J]. Annals of Botany, 2002, 90(1): 111—118.
- [8] Larsen R U, Cecilia H. Predicting leaf unfolding in flower induced shoots of greenhouse grown chrysanthemum[J]. Scientia Horticulturae, 1995, 63(3/4): 225—239.
- [9] Carvalho S M P, Heuvelink E, van Kooten O. Effect of light intensity, plant density and flower bud removal on flower size and number in cut chrysanthemum[J]. Acta Horticulturae, 2002, 593(3): 33—38.
- [10] Nothnag I M, Kosiba A, Larsen R U. Predicting the effect of irradiance and temperature on the flower diameter of greenhouse grown chrysanthemum[J]. Scientia Horticulturae, 2004, 99(3/4): 319—329.
- [11] 王定成, 乔晓军, 汪春秀, 等. 基于作物响应的温室环境 SVMR 控制仿真[J]. 农业工程学报, 2010, 26(增刊 2): 290—293. Wang Dingcheng, Qiao Xiaojun, Wang Chunxiu, et al. Simulation of greenhouse environment SVMR control based on plant response[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2010, 26(Supp.2): 290—293. (in Chinese with English abstract)
- [12] 李永博, 周伟, 汪小岳, 等. 基于奇异摄动理论的综合温室控制系统设计[J]. 农业机械学报, 2012, 43(5): 184—189. Li Yongbo, Zhou Wei, Wang Xiaochan, et al. Greenhouse control system design based on singular perturbation theory[J]. Transactions of the Chinese Society for Agricultural Machinery, 2012, 43(5): 184—189. (in Chinese with English abstract)
- [13] 张瑜, 汪小岳, 孙国祥, 等. 基于激光视觉的温室作物茎叶量测方法研究[J]. 农业机械学报, 2014, 45(9): 254—259. Zhang Yu, Wang Xiaochan, Sun Guoxiang, et al. Leaves and stems measurement of plants based on laser vision in greenhouses[J]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(9): 254—259. (in Chinese with English abstract)
- [14] Field R. Spatial species-richness gradients across scales: A meta-analysis[J]. Journal of biogeography, 2009, 36(1): 132—147.
- [15] Crawford J. Multi-scale investigations of alpine vascular plant species in the San Juan Mountains of Colorado, USA GLORIA target region. Scientifica Acta, 2008, 2(2): 65—69.
- [16] Foody G M. Spatial non-stationarity and scale-dependency in the relationship between species richness and environmental determinants for the sub-Saharan endemic avifauna[J]. Global Ecology and Biogeography, 2004, 13(4): 315—320.
- [17] Keitt T H, Urban D L. Scale-specific inference using wavelets[J]. Ecology, 2005, 86(9): 2497—2504.
- [18] Grinsted A, Moore J C, Jevrejeva S. Application of cross wavelet transform and wavelet coherence to geophysical time series[J]. Nonlinear Processes in Geophysics, 2004, 11: 561—566.
- [19] Zhou Y, Tao T, Mei X, et al. Feed-axis gearbox condition monitoring using built-in position sensors and EEMD method[J]. Robotics and Computer-Integrated Manufacturing, 2011, 27(4): 785—793.
- [20] Ai-Ho L, Che-Chou S, Pai-Chi L. Potential contrast improvement in ultrasound pulse inversion imaging using EMD and EEMD[J]. Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on, 2010, 57(2): 317—326.
- [21] Huang N E, Zheng S, Long S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis[J]. Proc R Soc Lond, 1998, 454: 903—995.
- [22] Huang Norden E, Attoh-Okine N O. The Hilbert-Huang transform in engineering[M]. England, Taylor & Francis, 2005.
- [23] Huang Yongxiang, François G., Schmitt B., Time dependent intrinsic correlation analysis of temperature and dissolved oxygen time series using empirical mode decomposition[J]. Journal of Marine Systems, 2014, 130: 90—100.
- [24] Huang Norden E, Wu Zhaohua. A review on Hilbert-Huang transform: Method and its applications to geophysical studies[J]. Reviews of Geophysics, 2008, 46(2): 1—23.
- [25] Huang N E, Wu Manli C, Long S R, et al. A confidence limit for the empirical mode decomposition and Hilbert spectral analysis[J]. Proceedings Royal Society of London, 2003, A(459): 2317—2345.
- [26] Wu Zhaohua, Huang Norden E. Hilbert-Huang Transform and its applications[C]//Huang Norden E, Shen Samuel S P, Hilbert-Huang Transform and its applications, Singapore: World Scientific Publishing Company, 2005: 107—125.
- [27] 刘荣, 方鹏飞. 改进 Elman 神经网络的综合气象短期负荷预测[J]. 电力系统保护与控制, 2012, 40(22): 113—117. Liu Rong, Fang Gefei. Short-term load forecasting with comprehensive weather factors based on improved Elman neural network[J]. Power System Protection and Control, 2012, 40(22): 113—117. (in Chinese with English abstract)
- [28] 周云龙, 王强, 孙斌, 等. 基于希尔伯特—黄变换与 Elman 神经网络的气液两相流型识别方法[J]. 中国电机工程学报, 2007, 27(11): 50—56. Zhou Yunlong, Wang Qiang, Sun Bin, et al. Applied study of Hilbert-Huang transform and Elman neural network on flow

- regime identification for gas-liquid two-phase flow[J]. Proceedings of the CSEE, 2007, 27(11): 50–56. (in Chinese with English abstract)
- [29] 丁世飞, 贾伟宽, 许新征, 等. 基于 PLS 的 Elman 神经网络算法研究[J]. 电子学报, 2010, 38(2A): 71–75.  
Ding Shifei, Jia Weikuan, Xu Xinzheng, et al. Elman neural network algorithm based on PLS[J]. ACTA Electronic SINICA, 2010, 38(2A): 71–75. (in Chinese with English abstract)
- [30] 张宏, 马岩, 李勇, 等. 基于遗传 BP 神经网络的核桃破裂功预测模型[J]. 农业工程学报, 2014, 30(18): 78–84.  
Zhang Hong, Ma Yan, Li Yong, et al. Rupture energy prediction mode for walnut shell breaking based on genetic BP neural network[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of CSAE), 2014, 30(18): 78–84. (in Chinese with English abstract)

## Prediction of cayenne pepper plant height based on ensemble empirical mode decomposition and Elman neural network

Zhang Yu, Wang Xiaochan<sup>\*</sup>, Sun Guoxiang, Li Yongbo

(1. College of Engineering, Nanjing Agricultural University, Nanjing 210031, China;

2. Jiangsu Modern Facilities Agricultural Technology and Equipment Engineering Laboratory, Nanjing 210031, China)

**Abstract:** In order to improve the performance of greenhouse control system and figure out the multi-time scale variable problem between plant growth and environment elements in greenhouses, the ensemble empirical mode decomposition (EEMD) and Elman neural network were used to predict cayenne pepper height in this paper. Taking the cayenne pepper 8819 and its plant height and environment elements (temperature, relative humidity, total inner radiation) as research object, plant height and environment elements were decomposed by the method of EEMD. The cayenne pepper plants sampled were irrigated with the nutrient solution with electricity conductivity of 2.0 and pH value of 6.5 from Shandong Agricultural University. Five intrinsic mode functions (imfs) were obtained by the EEMD, named imf1, imf2, imf3, imf4 and imf5. Oscillations of cayenne pepper plant height and environment elements at different frequency were shown by imf1, imf2, imf3 and imf4, while variation trend of cayenne pepper plant height and environment elements were shown in imf5. All imfs were reconstructed by reconstruction of the EEMD with approximate values of 1, 0 and 0 respectively for correlation coefficient, standard error and mean absolute error to original time series. The imfs obtained by EEMD decomposition were used to build plant height prediction model at different time-scale frequencies based on EEMD method and Elman neural network. All the plants of cayenne pepper 8819 sampled were divided into 2 sets, which were training set and testing set. Average values of sampled plants and environment elements were used in neural network building. In this paper, the environment elements (temperature, relative humidity and inner total radiation) formed the input layer of EEMD-Elman network, and the prediction of cayenne pepper plant height was the output layer of EEMD-Elman network. Double-layer feedback structure method was used in this paper, in which there were 10 nodes in the first layer and 3 nodes in the second layer. The function ‘transig’ was used as transfer function in the feedback layer, while the function ‘purelin’ was used in the output layer. Sample data were divided into training set, validation set and testing set in proportion of 0.7:0.15:0.15. Five sub-neural networks were established by imf1, imf2, imf3, imf4 and imf5 of samples’ plant height and environment elements. Final predicted value of sampled plant height was reconstructed by EEMD reconstruction, using the results of 5 sub-neural networks. Results of EEMD-Elman prediction model showed the mean absolute error for plant height was 1.69 cm, with the correlation coefficient of 0.996 and the standard error of 1.104, which meant the prediction value was significantly correlated to the real value. Otherwise, 2 different prediction models were built by the method of EEMD-BP (back propagation) neural network and Elman neural network. Results of EEMD-BP prediction model showed the mean absolute error of 5.40 cm for plant height, with the correlation coefficient of 0.812 and the standard error of 7.012, while results of Elman model showed the mean absolute error of 8.87 cm, with the correlation coefficient of 0.908 and the standard error of 5.032. The results of 3 different prediction models were compared. The prediction of EEMD-Elman neural network was the best among the 3 neural networks. The results of the models with EEMD were better than that of the Elman model without EEMD. So, EEMD could decompose time series into different time scales according to its own features without the disturbance of noises and singular wave signals. Details of original time series were demonstrated decently by EEMD. Fluctuation of original time series could be explained better at different time scales by EEMD. Precision of prediction could be improved by Elman neural network with double-layer feedback structure construction. In conclusion, the combination of EEMD and Elman neural network can be used to figure out the issue of the prediction of multi-time scale between plant growth and environment variation in greenhouses, in order to provide effective references for control objectives optimization in greenhouse control system.

**Key words:** neural network; models; optimization; ensemble empirical mode decomposition; elman neural network; cayenne pepper plant height; prediction