

# 控制与决策

Control and Decision

## 基于TVA-TCN的制造过程关键参数多步预测方法

彭慧, 朱雪靖, 周晓锋, 李帅, 刘舒锐

引用本文:

彭慧, 朱雪靖, 周晓锋, 李帅, 刘舒锐. 基于TVA-TCN的制造过程关键参数多步预测方法[J]. *控制与决策*, 2022, 37(12): 3321-3328.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2021.0922>

---

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### 基于深度卷积神经网络的刀具寿命动态预测研究

Research on dynamic prediction of tool life based on deep convolutional neural network  
*控制与决策*. 2022, 37(8): 2119-2126 <https://doi.org/10.13195/j.kzyjc.2020.1530>

#### 自适应感受野网络的行人重识别

Adaptive receptive network for person re-identification  
*控制与决策*. 2022, 37(1): 119-126 <https://doi.org/10.13195/j.kzyjc.2020.0505>

#### 基于时域扩张残差网络和双分支结构的人体行为识别

Human behavior recognition based on time domain extended residual network and dual branching structure  
*控制与决策*. 2022, 37(11): 2993-3002 <https://doi.org/10.13195/j.kzyjc.2021.0648>

#### 一种并行LSTM-FCN模型在船舶航迹预测中的应用

Parallel LSTM-FCN model applied to vessel trajectory prediction  
*控制与决策*. 2022, 37(8): 1955-1961 <https://doi.org/10.13195/j.kzyjc.2020.1795>

#### 一种基于深度学习的时间序列预测方法

A time series prediction method based on deep learning  
*控制与决策*. 2021, 36(3): 645-652 <https://doi.org/10.13195/j.kzyjc.2019.0809>

# 基于TVA-TCN的制造过程关键参数多步预测方法

彭 慧<sup>1,2,3</sup>, 朱雪靖<sup>1,2,3,4</sup>, 周晓锋<sup>1,2,3†</sup>, 李 帅<sup>1,2,3,4</sup>, 刘舒锐<sup>1,2,3</sup>

(1. 中国科学院 网络化控制系统重点实验室, 沈阳 110016; 2. 中国科学院 沈阳自动化研究所, 沈阳 110016;  
3. 中国科学院 机器人与智能制造创新研究院, 沈阳 110169; 4. 中国科学院大学, 北京 100049)

**摘要:** 制造过程关键参数的准确预测对制造过程的精确控制起关键作用, 现有预测方法通常未考虑时间动态特性, 多步预测性能不佳, 无法满足制造过程实际需求. 对此, 提出一种基于时变注意力时间卷积网络(TVA-TCN)的制造过程关键参数多步预测方法. 首先, 鉴于普通卷积网络感受野的局限性, 利用多通道时间卷积网络提取数据的长期依赖关系, 并使用Softplus激活函数降低对数据异常值的敏感度; 其次, 提出一种时变模型结构, 通过提取上一时间步的隐藏层信息和输出信息, 使得模型不仅能够随时间动态更新, 而且可以缓解梯度消失, 从而提高多步预测性能; 最后, 利用食品加工制造过程的实际数据进行多步预测实验, 结果表明所提出方法与传统的方法相比具有明显的优势.

**关键词:** 多步预测; 制造过程; 时变模型; 时间卷积网络

**中图分类号:** TP391; TP273 **文献标志码:** A

**DOI:** 10.13195/j.kzyjc.2021.0922

**引用格式:** 彭慧, 朱雪靖, 周晓锋, 等. 基于 TVA-TCN 的制造过程关键参数多步预测方法[J]. 控制与决策, 2022, 37(12): 3321-3328.

## Multi-step prediction method for key parameters of manufacturing process based on TVA-TCN

PENG Hui<sup>1,2,3</sup>, ZHU Xue-jing<sup>1,2,3,4</sup>, ZHOU Xiao-feng<sup>1,2,3†</sup>, LI Shuai<sup>1,2,3,4</sup>, LIU Shu-rui<sup>1,2,3</sup>

(1. Key Laboratory of Networked Control Systems, Chinese Academy of Sciences, Shenyang 110016, China; 2. Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China; 3. Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang 110169, China; 4. University of Chinese Academy of Sciences, Beijing 100049, China)

**Abstract:** Accurate prediction of key parameters in tobacco primary processing plays a key role in its precise optimization and control. Existing prediction methods usually do not consider time dynamic characteristics, and the performance of multi-step prediction is not good, which cannot meet the actual index needs of tobacco primary processing. In response to the above problems, a multi-step prediction method for key parameters of tobacco primary processing based on the time-varying attention-temporal convolutional network (TVA-TCN) is proposed. Firstly, for the key information in the input variables, the attention mechanism is introduced to capture the information. Then, a multi-channel temporal convolutional network is used to extract the long-term dependence of the data. Finally, by extracting the hidden layer information and output information of the previous time step, the model can be dynamically updated over time, thereby improving the performance of multi-step prediction. Multi-step prediction experiments are carried out using real data of tobacco primary processing, and the results show that the proposed method has obvious advantages compared with traditional methods.

**Keywords:** multi-step prediction; manufacturing process; time-varying model; temporal convolutional network

## 0 引言

制造业作为国民经济的支柱产业之一, 在我国经济发展中具有重要的地位. 所制造出的产品关键参数是否达到规定的质量指标, 直接影响着制造产品的品质. 考虑到制造过程通常具有非线性和大滞后等

特征, 对关键参数的测量往往耗时长, 难以检测, 因此制造过程关键参数的准确预测就显得尤为重要, 并且制造过程关键参数的准确预测对制造过程的优化控制具有重要意义<sup>[1]</sup>.

近年来, 很多研究者使用机器学习和深度学习技

收稿日期: 2021-05-25; 录用日期: 2021-09-22.

基金项目: 辽宁省重点研发计划项目(2020JH2/10100039).

责任编委: 牛玉刚.

†通讯作者. E-mail: zhouxf@sia.cn.

术进行制造过程的关键参数预测,并且取得了不错的效果.文献[2]基于协同深度动态特征提取和可变时延估计,提出一种基于数据驱动的制造过程参数预测模型,能够较好地预测出存在时间延迟的过程变量.文献[3]针对制造过程中部分参数难以直接测量的问题,提出基于神经模糊技术和人工神经网络的参数预测方法,实现对关键参数的准确软测量.文献[4]提出一种多重采样率径向基函数自回归模型,用于捕获和量化制造过程中的全局非线性特征.文献[5]选择牛顿模型来描述制造过程,结合多维连续随机变量非线性函数均值和方差的近似计算方法,从而建立描述制造过程中关键参数均值和方差的稳定性预测模型,具有较好的预测效果.

虽然上述方法在关键参数单步预测上效果较好,但是对制造过程的时间动态特征变化的捕捉不够灵敏.考虑到制造过程的滞后性,实际生产过程的关键参数预测往往需要考虑一段较长时间的结果,可以让操作人员根据未来关键参数趋势提前进行其余参数调节,从而提高制造过程的精确性和稳定性.文献[6]针对火力发电厂存煤量预测精度不高的问题,提出一种基于奇异谱分析和长短期记忆神经网络的多变量多步预测模型.文献[7]提出一种基于卷积神经网络-长短期记忆(CNN-LSTM)的预测模型,将其用于多步风电预测.文献[8]考虑每日参考蒸散量预报可以帮助农民进行灌溉计划,使用一维卷积神经网络来进行参考蒸散量的多步预测.文献[9]提出一种对植物茎直径变化的多步预测方法,先将小波分解应用于原始数据,然后提出一个包含LSTM和注意力机制的递归神经网络.文献[10]开发了样本熵来识别原始时间序列中的非线性和不确定性,并且通过多目标正弦余弦优化算法来调整最小二乘支持向量机,进而对电力需求序列进行多步提前预测.

以上方法采用的主要是基于RNN和CNN的方法,其中对于时间序列预测中常用的循环神经网络(recurrent neural network, RNN),随着时间的推移,RNN中的神经元是复制不变的,因此RNN具有时不变性<sup>[11]</sup>.对于CNN,其具有空间上的不变性,将其运用到时间序列预测中,也具有时不变性<sup>[12]</sup>.因此,时间不变性和空间不变性模型对于多步时间序列预测而言,缺乏对输入输出序列长度动态变化的精确捕捉能力.为了解决上述问题,本文提出一种基于TVA-TCN的多步预测方法用于制造过程关键参数预测,利用具有跟随时间变化的动态细胞结构,实现长时间序列的多步预测,并且在食品加工制造过程的实际数据集上进行实验,实验结果表明所提出的方法在性能方面优于传统方法.

本文的贡献如下:1)提出一种基于TVA-TCN的

预测方法来实现制造过程关键参数的多步预测;2)通过提取上一时间步的隐藏层信息和输出信息,使得模型能够随时间动态更新;3)提出的时变模型结构可以很好地适应时间步上的变化,并且可以缓解梯度消失;4)鉴于普通卷积网络感受野的局限性,使用多通道的时间卷积网络进行时间序列特征的提取.

## 1 基本方法

### 1.1 注意力机制

注意力机制由于其直观性、通用性和可解释性,在人工智能领域已经广泛使用<sup>[13]</sup>,其是在每个时间步上,通过参考先前的隐藏状态,自适应地提取相关的外部变量序列,从而实现自适应地选择最相关的输入特征的一种算法<sup>[14]</sup>.

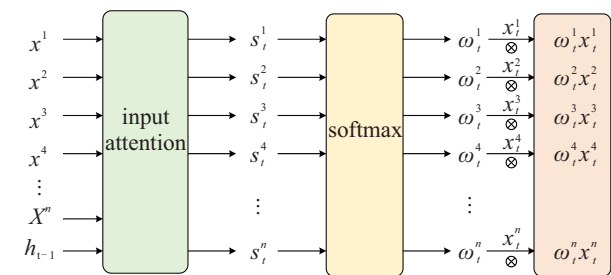


图1 注意力机制结构

注意力机制结构如图1所示, $x^1, x^2, \dots, x^n$ 为 $t$ 时刻的 $n$ 个输入变量,参考上一时刻的隐藏状态 $h_{t-1}$ ,构造注意力机制

$$s_t^i = v_s^T \tanh(W_s h_{t-1} + U_s x_t^i + b), \quad (1)$$

$$w_t^i = \frac{\exp(s_t^i)}{\sum_{i=1}^n \exp(s_t^i)}. \quad (2)$$

其中: $v_s, W_s, U_s, b$ 是需要学习的自适应参数,在求解 $w_t^i$ 时应用Softmax函数来确保所有的注意力权重之和为1,从而可以自适应地提取外部输入变量序列,即

$$\tilde{X}_t = (w_t^1 x_t^1, w_t^2 x_t^2, \dots, w_t^n x_t^n)^T. \quad (3)$$

然后将带有自适应权重的 $\tilde{X}_t$ 作为输入送进下一层网络训练.

### 1.2 时间卷积网络

为了防止产生时间穿越,造成模型使用未来的数据训练而产生过拟合,需要使用因果卷积进行时间上的严格约束<sup>[15]</sup>.时间卷积网络采用膨胀卷积提取历史时间特征<sup>[16]</sup>,网络堆叠少量的隐藏层即可捕获较长的历史时间信息.序列元素 $s$ 的扩张卷积可表示为

$$F(s) = (x_{*df})(s) = \sum_{i=0}^{k-1} f(i) \cdot x_{s-di}. \quad (4)$$

其中: $d$ 是扩张因子, $k$ 是卷积核大小.扩张卷积使得提取的时间特征窗口呈指数型增长.

时间卷积网络在两个扩张卷积层之间加入跳跃连接,用于解决网络深度增加带来的梯度消失和训练耗时的问题<sup>[17]</sup>,并在每层扩张卷积后加入权重归一化和Dropout层来避免过拟合<sup>[18]</sup>.

## 2 TVA-TCN预测模型

### 2.1 模型描述

模型的输入由两部分组成,一部分是外部变量序列,为  $x = (x_{t-1}, x_{t-2}, \dots, x_{t-T}) \in \mathbf{R}^{n \times T}$ . 其中:  $T$ 是时间窗口的长度,  $n$ 为外部驱动变量个数. 另一部分是目标序列的历史序列,  $y = (y_{t-1}, y_{t-2}, \dots, y_{t-T}), y_i \in \mathbf{R}$ .

注意力机制能够自适应地提取相关的外部变量序列,从而实现自适应地选择最相关的输入特征. 将

经过注意力机制后的带有自适应权重的数据输入时间卷积网络,用于提取多变量时间序列的依赖性.

鉴于需要进行多变量输入的时间特征提取,本文方法中的时间卷积网络采用多通道的扩张卷积核. 多通道时间卷积网络结构如图2所示,使用多通道扩张卷积来进行多变量时间特征的提取,扩张因子  $d$ 为  $[1, 2, 4, 8]$ ,卷积核为2. 同时两个扩张卷积层之间使用多通道卷积进行跳跃连接. 实际生产过程中,会由于传感器采集误差,导致数据存在异常值. 而普通的relu激活函数对异常值特别敏感,会造成relu神经元永久关闭,而softplus激活函数具有相对宽阔的兴奋边界,所以选用softplus作为激活函数.

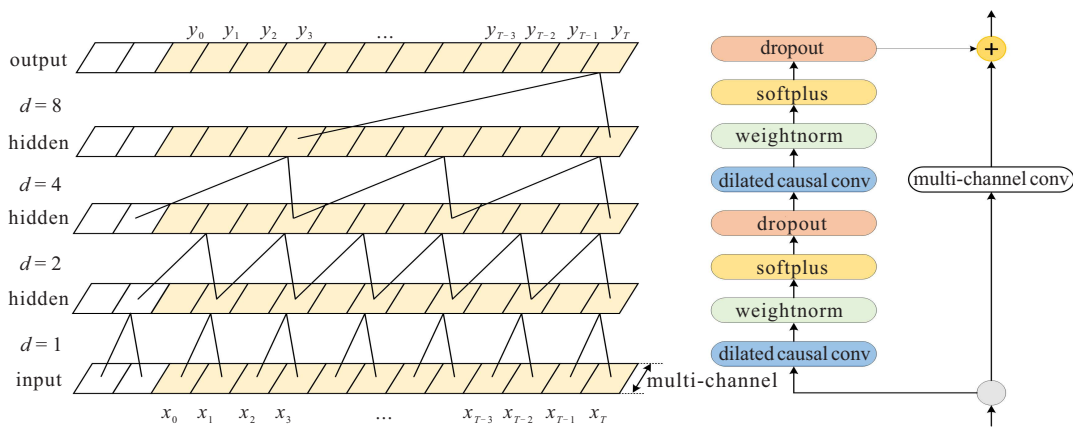


图2 多通道时间卷积网络结构

设经过第  $i$  个残差块的输出为  $y_i$ , 初始输入为  $X_i \in \mathbf{R}^{n+1}$ , 扩张卷积层为  $F_i(s)$ , 多通道卷积为  $C_i^0$ , 则有

$$y_i = C_i^0 X_i + \text{softplus}(F_i(s)(\text{softplus}(F_i(s)X_i))). \quad (5)$$

其中:  $y_n^j$  代表最后一层 (第  $n$  层) 的残差块输出, 总共有  $m$  个隐藏神经元, 最终输出  $Y$  为对其进行全连接的结果.

本文提出一种基于 TVA-TCN 的多步时间预测方法, 其包含注意力机制、时间卷积网络、时变结构. 该结构在  $t$  时刻的输出为

$$Y_t = f_t(h_t), \quad (6)$$

$$h_t = g_t(\text{attn}(X), h_{t-1}, Y_{t-1}). \quad (7)$$

其中:  $X = [x_t^1, x_t^2, x_t^3, \dots, x_t^n]$  为  $t$  时刻外部输入的  $n$  维变量,  $Y_t$  为  $t$  时刻的输出量,  $h_t$  为时间卷积网络之后的输出,  $f_t$  为描述输出层的函数,  $g_t$  为描述时间卷积网络中隐藏层的结构. 其中  $f_t$  和  $g_t$  都是随时间  $t$  变化的函数.

TVA-TCN 结构如图3所示, 原始输入  $X$  经过注意力机制层后产生自适应的权重, 赋予自适应权重的输入送进 TCN 隐藏层中, 经过 TCN 隐藏层后输出最终的  $Y$ .

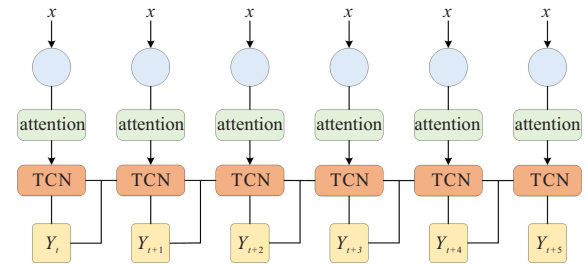


图3 TVA-TCN 结构

LSTM 结构中预测多步时, 只是将上一时刻的隐藏层输出  $h_{t-1}$  与原始输入  $X_t$  输入进下一时刻的隐藏层中. 与 LSTM 网络结构不同的是, TVA-TCN 模型中每个  $t+i$  时刻的隐藏层输入都由原始输入  $X$  经过注意力机制层后的结果, 上一时间步的隐藏层输出  $h_{t-1}$  以及上一时间步的输出  $Y_{t-1}$  构成, 模型可以捕获到上一时间步的信息, 从而使得模型可以根据预测步长进行动态变化.

2.2 TVA-TCN随时间变化的优点

2.2.1 LSTM的时不变性

为了与TVA-TCN随时间动态变化的特点进行对比,本文分析了时间序列预测中常用的LSTM模型.

LSTM的细胞结构如图4所示,有

$$h_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tanh(\sigma(W_f[h_{t-1}, x_t] + b_f) \times C_{t-1} + \sigma(W_i[h_{t-1}, x_t] + b_i) \sigma(W_C[h_{t-1}, x_t] + b_C)). \quad (8)$$

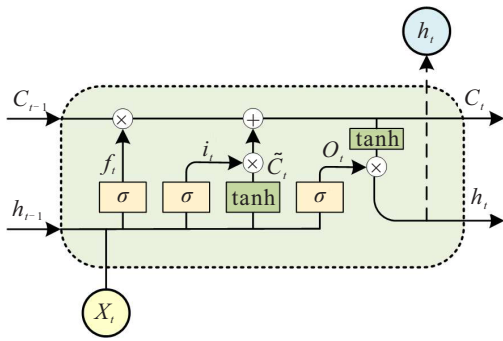


图4 LSTM细胞结构

LSTM整体结构可以简化为如图5所示,有

$$h_t = g(wh_{t-1}, uX_t), \quad (9)$$

$$Y_t = f(h_t). \quad (10)$$

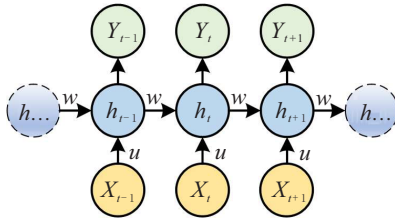


图5 LSTM整体结构

给定两个输入  $X_t, X_{t-\tau}$ , 两个输出  $Y_t, Y_{t-\tau}$ , 两个隐层状态  $h_t, h_{t-\tau}$ . 为了保持输入相同, 令  $X_t = X_{t-\tau}$ , 相似的,  $h_t = h_{t-\tau}$ . 其中,  $f$  和  $g$  都是不随时间改变的函数. 要证明模型是具有时不变性等价于证明  $Y_t = Y_{t-\tau}$ .

$$Y_t = f(g(wh_{t-1}, uX_t)), \quad (11)$$

$$Y_t = f(g(uX_t, [wg(uX_{t-1}, [\dots [wg(uX_1, wh_0)]])])), \quad (12)$$

$$Y_{t-\tau} = f(g(wh_{t-\tau-1}, uX_{t-\tau})), \quad (13)$$

$$Y_{t-\tau} = f(g(uX_{t-\tau}, [wg(uX_{t-\tau-1}, [\dots [wg(uX_{1-\tau}, wh_{0-\tau})]]])). \quad (14)$$

因为  $\forall t \leq \tau, X_{t-\tau} = 0$  和  $\forall t \leq \tau, h_{t-\tau} = 0$ , 所以

由式(14)可得

$$Y_{t-\tau} = f(g(uX_{t-\tau}, [wg(uX_{t-\tau-1}, [\dots [wg(uX_1, wh_0)]])])). \quad (15)$$

因为  $X_t = X_{t-\tau}, h_t = h_{t-\tau}$ , 由式(15)可得

$$Y_{t-\tau} = f(g(uX_t, [wg(uX_{t-1}, [\dots [wg(uX_1, wh_0)]])])). \quad (16)$$

所以  $Y_t = Y_{t-\tau}$ , 即LSTM模型是时不变的.

2.2.2 TVA-TCN的时变性

给定两个输入  $X_t, X_{t-\tau}$ , 两个输出  $Y_t, Y_{t-\tau}$ , 两个隐藏层状态  $h_t, h_{t-\tau}$ . 为了保持输入相同, 令  $X_t = X_{t-\tau}$ , 同时这里假设输入经过注意力机制操作后的值也相等, 即  $\text{Attn}(X_t) = \text{Attn}(X_{t-\tau})$ , 相似的,  $h_t = h_{t-\tau}$ . 要证明模型是具有时不变性等价于证明  $Y_t = Y_{t-\tau}$ .

$$Y_t = f_t(g_t(\text{Attn}(X_t), h_t Y_t)), \quad (17)$$

$$Y_{t-\tau} = f_{t-\tau}(g_{t-\tau}(\text{Attn}(X_{t-\tau}), h_{t-\tau-1}, Y_{t-\tau-1})). \quad (18)$$

由于  $\text{Attn}(X_t) = \text{Attn}(X_{t-\tau}), h_t = h_{t-\tau}$ , 由式(18)可得

$$Y_t = f_t(g_t(\text{Attn}(X_{t-\tau}), h_{t-\tau}, Y_{t-1})). \quad (19)$$

因为  $f_t \neq f_{t-\tau}$  且  $g_t \neq g_{t-\tau}$  且  $Y_{t-1} \neq Y_{t-\tau-1}$ , 所以  $Y_t \neq Y_{t-\tau}$ , 即模型是随时间动态变化的. 可以看出, 传统的LSTM模型具有时不变性, 下一时刻的预测模型没有很好地使用到上一时刻预测结果的信息. 而本文提出的TVA-TCN模型是随时间动态变化的, 模型可以随着预测步长的变化而产生自适应变化.

2.3 TVA-TCN缓解梯度消失的优点

当前深度学习是根据损失函数计算误差, 并通过反向传播的方式来优化神经网络, 而将误差从最后一层传输到第1层的过程需要通过链式法则来处理. 链式法则是一种连续乘法的形式, 因此当网络层数较深时, 梯度消失和梯度爆炸问题变得越来越明显. 根据损失函数计算出来的误差通过反向传播对权重进行更新时, 其梯度会接近0或者无穷大, 即产生了梯度消失或者梯度爆炸. 而在TVA-TCN中, 通过在隐藏层和输出层之间进行交错输出, 链式被分解为多个项的总和, 而多个项的和比多个项的乘积更加稳定, 因此能够缓解梯度消失.

TVA-TCN的交叉输出结构如图6所示, 对于每一层有1个隐藏层(时间卷积网络)和1个输出层, 对于每一层有

$$g^l = a(h^l), \tag{20}$$

$$h^{l+1} = w_h^l g^l + b_h^l, \tag{21}$$

$$f^l = w_f^l g^l + b_f^l, \tag{22}$$

$$Y^l = a(f^l). \tag{23}$$

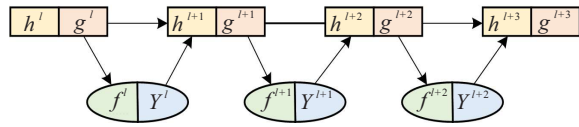


图6 交叉输出结构

利用微积分的链式法则, 损失函数 \$L\$ 对第 \$l\$ 层的权重 \$w\_h^l\$ 的导数为

$$\frac{\partial L}{\partial w_h^l} = \sum_{i=1}^n \frac{\partial L}{\partial f^{l+i-1}} \frac{\partial f^{l+i-1}}{\partial g^{l+i-1}} \cdots \frac{\partial h^{l+2}}{\partial g^{l+1}} \frac{\partial h^{l+1}}{\partial g^l} \frac{\partial g^l}{\partial w_h^l} + \frac{\partial L}{\partial h^{l+i}} \frac{\partial h^{l+i}}{\partial g^{l+i-1}} \cdots \frac{\partial h^{l+2}}{\partial g^{l+1}} \frac{\partial h^{l+1}}{\partial g^l} \frac{\partial g^l}{\partial w_h^l}. \tag{24}$$

由式(24)可得, 损失函数 \$L\$ 对第 \$l\$ 层的权重 \$w\_h^l\$ 的导数为多个项的和, 相较于使用多项式的乘积, 值小于 1 的多项和比值小于 1 的多项积趋向于 0 的速度更慢, 从而缓解了梯度消失.

### 3 实验验证

#### 3.1 制造过程案例

以某食品加工制造过程为例, 加工制造过程主要可以分为真空回潮、松散回潮、发酵、干燥等环节. 松散回潮是指对原料进行松散和增温增湿, 使原料的含水率和温度增至一定水平. 但松散回潮环节的原材料含水率波动较大, 会导致出口参数控制效果较差, 准确的长时间预测结果可以让操作人员根据关键参数未来趋势提前进行其余参数调节, 从而提高出口参数的稳定性. 因此, 本文主要关注松散回潮环节, 如图 7 所示. 经过切片工序的原料, 被铺开均匀倒入传送带, 随传送带到达入口位置, 入口传感器检测入口原料参数并记录. 原料进入滚筒后, 滚筒入口处的蒸汽喷射装置对其喷射蒸汽, 充分润湿后, 在滚筒内向前翻滚均匀. 经过一段时间后, 原料到达出口传感器位置, 出口传感器检测出口参数并记录, 最终的产品沿着传送带转移至下一工序.

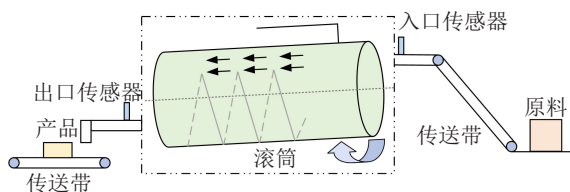


图7 松散回潮环节

在此食品加工过程中, 出口水分为此过程需要预测的关键参数. 出口水分直接影响产品的颜色、光泽、口感, 也影响食品加工过程中原料的损耗程度. 因此,

出口水分在此食品加工过程中占有重要的地位.

#### 3.2 数据分析与处理

本文使用某食品加工制造过程的实际生产数据, 从某食品加工厂的制造执行系统中提取数据. 数据采集时间约一个月, 采样时间间隔为 6s, 总计 84 个批次、37 212 个样本. 数据的特征包括温度仪(出口)温度实际值(%), 加水实际值(kg/h), 实际加水累计量(kg), 热风温度(°C), 入口水分(%), 热交换蒸汽阀开度(%)等 21 个相关性强的变量.

原料从入口传感器到出口传感器的过程为 4 min, 在实际生产中出口含水率的工艺评价指标为 ±1.5(含水率单位:%). 利用随机选择的方式, 以批次比例为 59:25 划分训练集测试集, 并进行多次交叉验证.

#### 3.3 基于 TVA-TCN 的预测结果

TVA-TCN 模型的结构主要由注意力机制层和时间卷积网络堆叠而成, 并且在下一预测时间步的隐藏层中加入上一时间步的隐藏层信息和输出信息, 预测目标是未来 30s 的制造过程关键参数出口含水率的百分比值.

TVA-TCN 模型中多通道卷积的通道为 21, 训练时隐藏神经元为 48 个, 输入的时间窗口为 20, 预测输出的时间窗口为 5, Epoch 的值设为 100, Batchsize 的值设为 16, 学习率设置为 0.001, 训练过程使用的优化器为 Adam, 训练所使用的机器为 GPU2080Ti.

与本文所提出的 TVA-TCN 进行实验对比的多步预测方法包括 LSTM<sup>[6]</sup>、CNN-LSTM<sup>[7]</sup>、TCN<sup>[15]</sup>、基于双边注意力机制的模型 DA-LSTM<sup>[19]</sup> 以及基于注意力机制的 TCN 模型 A-TCN. 其中 A-TCN 为 TVA-TCN 去掉时变结构后的模型. 为了衡量各种时间序列预测方法的有效性, 采用平均绝对误差(MAE)、均方根误差(RMSE)、平均绝对百分比误差(MAPE)作为评价指标.

TVA-TCN 与经典的时间序列预测模型在预测时长为 6s~30s 的预测结果如表 1 所示. 可以看出, CNN-LSTM 以两层结构进行特征抽取, 第 1 层使用 CNN 进行数据特征的预抽取, 第 2 层使用 LSTM 对预抽取的数据特征进行叠加抽取, 因此 CNN-LSTM 的预测效果优于 LSTM 模型. TCN 由于使用了扩张卷积和残差连接, 拥有更大的感受野, 预测未来时可以使用到更多的历史信息, 因此预测效果优于 LSTM 模型. DA-LSTM 是使用驱动序列注意力和时间步长注意力机制的 LSTM 模型<sup>[19]</sup>, 能够提取相关性高的驱动序列和时间步长序列, 因此在实际数据集上的预测性

表1 提前多步预测的出口含水率结果

预测时长/s	预测模型	MAE	RMSE	MAPE/%
6	LSTM	0.1603	0.2010	0.877
	CNN-LSTM	0.1531	0.2002	0.836
	TCN	0.1499	0.2017	0.821
	DA-LSTM	0.1496	0.1957	0.818
	A-TCN	0.1454	0.1902	0.798
	TVA-TCN	<b>0.1204</b>	<b>0.1609</b>	<b>0.653</b>
12	LSTM	0.1851	0.2326	1.012
	CNN-LSTM	0.1781	0.2312	0.968
	TCN	0.1724	0.2302	0.943
	DA-LSTM	0.1707	0.2234	0.934
	A-TCN	0.1689	0.2204	0.919
	TVA-TCN	<b>0.1480</b>	<b>0.1972</b>	<b>0.802</b>
18	LSTM	0.2091	0.2629	1.144
	CNN-LSTM	0.1925	0.2518	1.047
	TCN	0.1875	0.2504	1.025
	DA-LSTM	0.1873	0.2458	1.025
	A-TCN	0.1851	0.2401	1.012
	TVA-TCN	<b>0.1688</b>	<b>0.2223</b>	<b>0.917</b>
24	LSTM	0.2246	0.2846	1.230
	CNN-LSTM	0.2148	0.2747	1.170
	TCN	0.2033	0.2666	1.112
	DA-LSTM	0.2013	0.2672	1.101
	A-TCN	0.1956	0.2612	1.081
	TVA-TCN	<b>0.1841</b>	<b>0.2439</b>	<b>0.999</b>
30	LSTM	0.2390	0.3027	1.309
	CNN-LSTM	0.2264	0.2865	1.236
	TCN	0.2173	0.2846	1.189
	DA-LSTM	0.2117	0.2814	1.158
	A-TCN	0.2102	0.2794	1.142
	TVA-TCN	<b>0.1992</b>	<b>0.2619</b>	<b>1.082</b>

能上优于LSTM模型. A-TCN为TVA-TCN去掉时变结构后的模型,即使用驱动序列注意力的TCN模型,其能够提取相关性高的驱动序列,因此预测效果上优于TCN模型.由表1可以看出,TVA-TCN在预测时长为6s~30s的预测结果中都取得了较好的结果,因为TVA-TCN是一个时变模型,对于下一时刻的预测都会自适应地使用到上一时刻的输出信息和隐藏层信息,因此能够适应多步预测的动态变化,并且TVA-TCN中使用了注意力机制对输入变量进行特征的自适应捕获,将提取到的相关性高的输入特征送进多通道扩张卷积中进行特征提取,从而达到了较优的预测效果.

原料在松散回潮过程的时间为4min,因此随机截取4min的时间进行预测结果的对比. CNN-LSTM、DA-LSTM、TCN与A-TCN是4个预测效果较优的时间序列预测模型,TVA-TCN与其提前6s预测出口含水率的结果如图8所示.可以看出,TVA-TCN在大多数时间内,预测结果非常接近实际出口含水率,并且预测误差的绝对值低于其他两个模型.结合表1和图8可以明显发现TVA-TCN在MAE、MSE、MAPE这3项指标上均表现更好.在提前6s预测出口含水率时,TVA-TCN方法的MAE较其余最优方法降低17.2%,RMSE降低15.4%,MAPE降低18.2%.

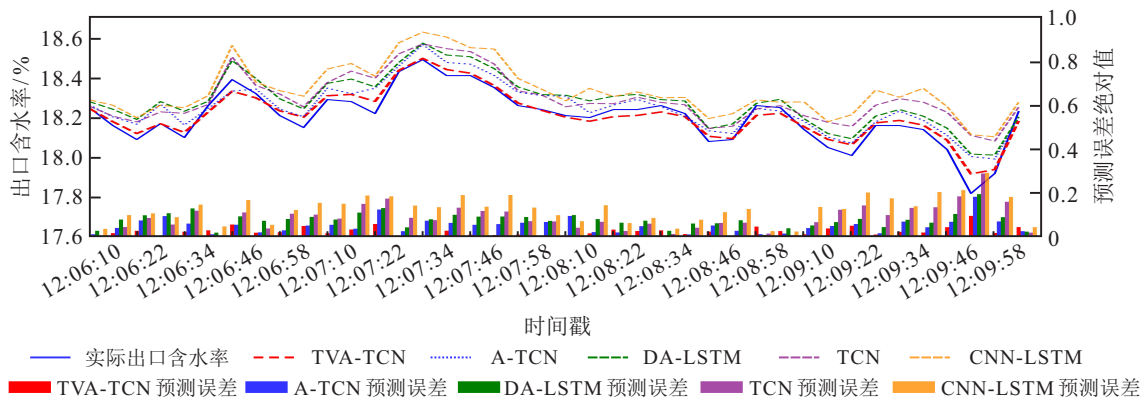


图8 提前6s的出口含水率预测结果

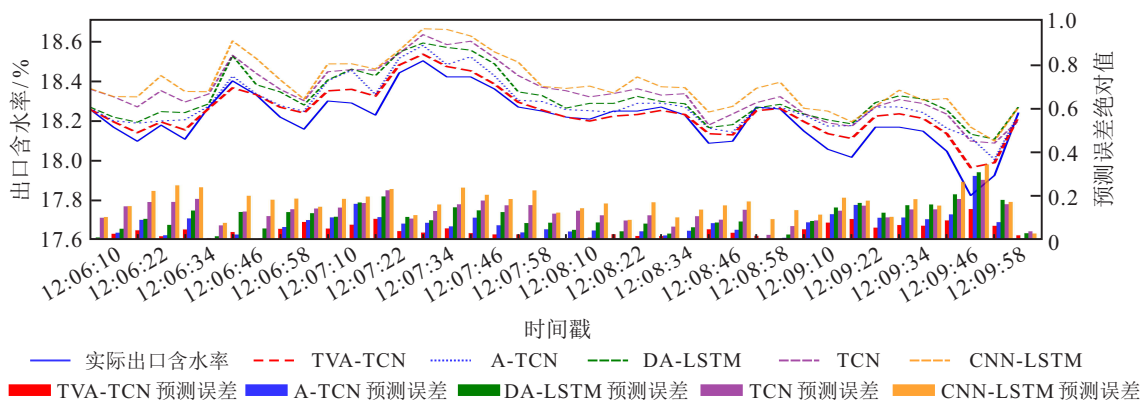


图9 提前30s的出口含水率预测结果

TVA-TCN与CNN-LSTM、TCN、DA-LSTM、A-TCN进行提前30s预测出口含水率的结果如图9所示。可以看出,TVA-TCN均优于其他4个模型,并且每个时间戳的预测误差绝对值大部分处于最小值的状态。在提前30s预测出口含水率时,TVA-TCN方法的MAE较其余最优方法降低5.2%,RMSE降低6.3%,MAPE降低5.3%。由于预测步长的增长使得预测误差的基数变大,导致相对比值变小,进而相对于提前6s预测出口含水率TVA-TCN提升的性能有所下降,但预测性能较好。

TVA-TCN与LSTM、CNN-LSTM、TCN、DA-LSTM、A-TCN在提前预测6s~30s的平均MAE如图10所示。进一步考虑参数指标考察要求,选用平均MAE在0~0.2区间中的占比来评估方法的整体预测性能。

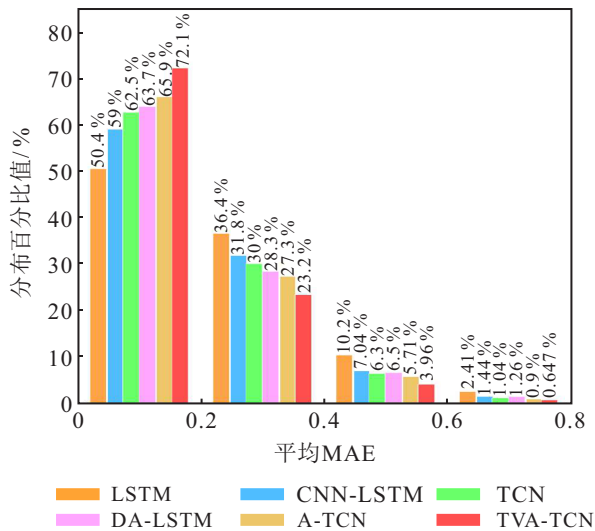


图10 出口含水率预测平均MAE

由图10可以看出,TVA-TCN的多步预测平均MAE有72.1%分布在0~0.2区间内,而次优模型A-TCN有65.9%分布在0~0.2区间内,TVA-TCN相对于A-TCN在0~0.2区间内明显增加了6.2%,这说明TVA-TCN在多步预测中的预测误差大部分是比较小的。并且,平均MAE在0.4~0.8区间中,TVA-TCN所占的百分比最小,可以说明TVA-TCN预测误差偏大的情况较少。可以发现,本文所提出的TVA-TCN在提前多步预测中的预测精度上相较其他时间序列预测模型具有一定优势,在实际应用中具有重要参考意义。

## 4 结论

本文面向制造过程关键参数的多步预测任务,研究了基于深度学习的时间序列多步预测方法。针对现有多步预测方法无法随时间动态变化的问题,提

出基于TVA-TCN的多步预测方法。该方法通过提取上一时间步的输出值和隐藏层值,通过捕获上一时间步的信息,实现模型能够跟随预测步长动态变化的特点。此外,使用多通道时间卷积网络对历史信息进行特征提取。最后在某食品加工制造过程实际数据上开展实验研究,结果验证了本文方法具有更为优秀的多步预测性能。本文方法主要用于单变量的多步预测,对于多变量的多步预测并未涉及,因此实现多变量的多步预测是后续的一项重要研究工作。

## 参考文献(References)

- [1] 陈龙, 刘全利, 王霖青, 等. 基于数据的流程工业生产指标预测方法综述[J]. 自动化学报, 2017, 43(6): 944-954.  
(Chen L, Liu Q L, Wang L Q, et al. Data-driven prediction on performance indicators in process industry: A survey[J]. Acta Automatica Sinica, 2017, 43(6): 944-954.)
- [2] Yao L, Ge Z Q. Cooperative deep dynamic feature extraction and variable time-delay estimation for industrial quality prediction[J]. IEEE Transactions on Industrial Informatics, 2021, 17(6): 3782-3792.
- [3] Fatima S A, Ramli N, Taqvi S A A, et al. Prediction of industrial debutanizer column compositions using data-driven ANFIS- and ANN-based approaches[J]. Neural Computing and Applications, 2021, 33(14): 8375-8387.
- [4] Zhou F, Peng H, Ruan W J, et al. Cubic-RBF-ARX modeling and model-based optimal setting control in head and tail stages of cut tobacco drying process[J]. Neural Computing and Applications, 2018, 30(4): 1039-1053.
- [5] Zhang M J, Huang F, Chen Q, et al. The prediction model of moisture content's stabilization during tobacco strip drying process[C]. Proceedings of 21th International Drying Symposium. Valencia: Universitat Politècnica València, 2018: 1511-1517.
- [6] 孔雯, 车权, 赵慧荣, 等. 基于奇异谱分析与长短时记忆神经网络的电厂存煤量短期预测[J]. 信息与控制, 2020, 49(6): 742-751.  
(Kong W, Che Q, Zhao H R, et al. Short-term prediction of coal stock in power plant based on singular spectrum analysis and long short-term memory neural network[J]. Information and Control, 2020, 49(6): 742-751.)
- [7] Wu Q Y, Guan F, Lv C, et al. Ultra-short-term multi-step wind power forecasting based on CNN-LSTM[J]. IET Renewable Power Generation, 2021, 15(5): 1019-1029.
- [8] Ferreira L B, da Cunha F F. Multi-step ahead forecasting of daily reference evapotranspiration using deep learning[J]. Computers and Electronics in Agriculture, 2020, 178: 105728.

- [9] Alhnaity B, Kollias S, Leontidis G, et al. An autoencoder wavelet based deep neural network with attention mechanism for multi-step prediction of plant growth[J]. *Information Sciences*, 2021, 560: 35-50.
- [10] Li R R, Chen X L, Balezentis T, et al. Multi-step least squares support vector machine modeling approach for forecasting short-term electricity demand with application[J]. *Neural Computing and Applications*, 2021, 33(1): 301-320.
- [11] Hewamalage H, Bergmeir C, Bandara K. Recurrent Neural Networks for Time Series Forecasting: Current status and future directions[J]. *International Journal of Forecasting*, 2021, 37(1): 388-427.
- [12] 胡美玉, 张云洲, 秦操, 等. 基于深度卷积神经网络的语义地图构建[J]. *机器人*, 2019, 41(4): 452-463. (Hu M Y, Zhang Y Z, Qin C, et al. Semantic map construction based on deep convolutional neural network[J]. *Robot*, 2019, 41(4): 452-463.)
- [13] 鹿天柱, 钱晓超, 何舒, 等. 一种基于深度学习的时间序列预测方法[J]. *控制与决策*, 2021, 36(3): 645-652. (Lu T Z, Qian X C, He S, et al. A time series prediction method based on deep learning[J]. *Control and Decision*, 2021, 36(3): 645-652.)
- [14] Cho K, Courville A, Bengio Y. Describing multimedia content using attention-based encoder-decoder networks[J]. *IEEE Transactions on Multimedia*, 2015, 17(11): 1875-1886.
- [15] Zhai N J, Zhou X F. Temperature prediction of heating furnace based on deep transfer learning[J]. *Sensors: Basel, Switzerland*, 2020, 20(17): 4676.
- [16] Wan R Z, Mei S P, Wang J, et al. Multivariate temporal convolutional network: A deep neural networks approach for multivariate time series forecasting[J]. *Electronics*, 2019, 8(8): 876.
- [17] 翟乃举, 周晓锋, 李帅, 等. 基于迁移学习和知识蒸馏的加热炉温度预测[J]. *计算机集成制造系统*, 2022, 28(6): 1860-1869. (Zhai N J, Zhou X F, Li S, et al. Prediction method of furnace temperature based on transfer learning and knowledge distillation[J]. *Computer Integrated Manufacturing Systems*, 2022, 28(6): 1860-1869.)
- [18] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(4): 640-651.
- [19] Yoon N, Kim J, Lim J L, et al. Dual-stage attention-based LSTM for simulating performance of brackish water treatment plant[J]. *Desalination*, 2021, 512: 115107.

### 作者简介

彭慧(1963—), 男, 研究员, 硕士, 从事执行制造系统、应用系统集成等研究, E-mail: penghui@sia.cn;

朱雪靖(1997—), 男, 硕士生, 从事时间序列预测与机器学习的研究, E-mail: zhuxuejing@sia.cn;

周晓锋(1978—), 女, 副研究员, 博士, 从事过程监测与故障诊断、机器学习与大数据分析处理等研究, E-mail: zhouxuf@sia.cn;

李帅(1988—), 男, 副研究员, 博士, 从事过程监测与故障诊断、数据挖掘等研究, E-mail: lishuai@sia.cn;

刘舒锐(1993—), 男, 助理研究员, 硕士, 从事工业过程建模、机器学习的研究, E-mail: liushurui@sia.cn.

(责任编辑: 孙艺红)

## 下 期 要 目

- 基于机器视觉的喷涂机器人轨迹规划与涂装质量检测研究综述..... 訾 斌, 等
- 6G通信感知一体化理论与技术..... 彭木根, 等
- 基于自适应大规模邻域搜索算法的多车辆与多无人机协同配送方法..... 伍国华, 等
- 电梯群控系统调度模型及其改进ADMM分解算法..... 张家瑞, 等
- 基于浅层定位的动态细化目标检测网络..... 郑荣元, 等
- 基于刺激-响应机制的改进鸡群算法..... 王英聪, 等
- 基于朴素贝叶斯的室内VLC网络天线选择方法..... 冷亭亭, 等
- 一种交互演化改进鲸鱼算法及其收敛性分析..... 刘景森, 等
- 基于两级筛选机制及深度学习组合模型实现短时交通流预测..... 徐先峰, 等
- 基于辅助输入信号和冗余执行器的主动故障诊断..... 吕雪燕, 等