文章编号:1001-2060(2024)04-0140-07

# 基于 STL-Former 的中短期光伏功率预测

李灯熬1,白晓东2,夏鸿伟3,冯 丁4

(1. 山西省能源互联网研究院,山西太原 030000; 2. 太原理工大学 计算机科学与技术学院(大数据学院),山西太原 030024; 3. 新疆石文科技发展有限公司,新疆乌鲁木齐 830011; 4. 太原师范学院 计算机科学与技术学院,山西太原 030619)

摘 要:由于光伏发电具有间歇性和波动性,给电网运行的安全性和稳定性造成危害,对光伏功率进行准确预测可以有效解决这一问题。本文提出一种基于 STL-Former 的中短期光伏功率预测模型,该模型结合了季节趋势局部加权回归分解(STL 分解)与神经网络模型。首先,STL-Former 模型将光伏功率数据通过 STL 分解进行特征扩充,用于提取基于历史序列的周期项、趋势项特征。然后,拼接周期项、趋势项特征和原特征,进行数据预处理和特征编码并使用基于 Informer 模型的神经网络进行功率预测。最后,在真实数据集上进行大量实验。实验结果表明:STL-Former 在中短期光伏功率预测任务中精度较高,其中在 2 h 光伏功率预测任务时,平均绝对值误差为 0.176、均方误差为 0.180;在 28 h 光伏功率预测任务时,平均绝对值误差为 0.170、均方误差为 0.154。

关 键 词:深度学习;光伏功率预测;STL分解

中图分类号:TP391

文献标识码·A

DOI:10.16146/j. cnki. rndlgc. 2024. 04. 017

[引用本文格式]李灯熬,白晓东,夏鸿伟,等. 基于 STL-Former 的中短期光伏功率预测[J]. 热能动力工程,2024,39(4):140 – 146. LI Deng-ao, BAI Xiao-dong, XIA Hong-wei, et al. Short-and medium-term photovoltaic power prediction based on STL-Former[J]. Journal of Engineering for Thermal Energy and Power,2024,39(4):140 – 146.

# Short-and Medium-Term Photovoltaic Power Prediction based on STL-Former

LI Deng-ao<sup>1</sup>, BAI Xiao-dong<sup>2</sup>, XIA Hong-wei<sup>3</sup>, FENG Ding<sup>4</sup>

- (1. Shanxi Province Energy Internet Research Institute, Taiyuan, China, Post Code: 030000; 2. College of Computer Science and Technology (College of Data Science), Taiyuan University of Technology, Taiyuan, China, Post Code: 030024;
  - 3. Xinjiang Shiwen Science and Technology Development Co., Ltd., Urumqi, China, Post Code: 830011;
  - 4. College of Computer Science and Technology, Taiyuan Normal University, Taiyuan, China, Post Code: 030619)

Abstract: The intermittency and volatility of photovoltaic (PV) power generation poses risks to the safety and stability of power grid operations. Accurately predicting PV power can effectively address this issue. Therefore, this paper proposed a short-and medium-term PV power prediction model based on STL-Former, which combined seasonal and trend decomposition using locally weighted regression (STL) and a neural network model. First, the STL-Former model utilized STL decomposition to expand the PV power data, extracting features based on historical sequences such as periodic and trend components. Then, the periodic and trend component features were concatenated with the original features, the data preprocessing and feature encoding were performed, and the power prediction was conducted using a neural network based on the Informer model. Finally, many experiments were carried out on real data sets. The experimental results show that STL-Former has high accuracy in short-and medium-term photovoltaic power forecasting tasks. Among them, the average absolute value error is 0. 176 and the mean square error is 0. 180

收稿日期:2023-07-10: 修订日期:2023-08-08

基金项目:山西省关键核心技术和共性技术研发攻关专项项目(2020XXX007);山西省重点研发计划项目(202102020101006)

in the task of PV power forecasting for 2 hours; the average absolute value error is 0.170 and the mean square error is 0.154 in the task of PV power forecasting for 28 hours.

**Key words:** deep learning, photovoltaic (PV) power prediction, STL decomposition

# 引言

化石能源(如石油和煤炭)的大量使用导致全球变暖和环境污染问题日益严重,因此新能源(如光伏和风电)成为现代能源系统重要的能源来源形式之一[1]。但是,光伏功率受到太阳辐射、温度、降雨、降雨及气压等因素相互作用影响,较为不稳定,而且其所特有的波动性和间歇性会让电网电力供应不稳定,使得电力调度和管理的难度增加。因此,有必要利用混合预测模型对光伏功率进行高精度预测。

光伏功率的预测模型可以分为物理模型、统计 模型、人工智能模型和混合模型[2]。物理模型主要 通过光照强度、温度等环境因素来预测光伏功率,其 研究主要集中在太阳辐照度的测量上,因为太阳辐 照度和光伏功率直接相关[3]。Frank 等人[4]研究发 现,太阳辐照度是影响光伏功率最重要的因素。 Ameen 等人[5]研究新的特征组合,建立了基于全球 水平辐照度的神经网络模型。常用的统计模型有自 回归模型[6]、指数平滑模型[7]等数学模型。其中, 自回归模型将当前观测值表示为过去时间点观测值 的加权和,通过拟合时间序列数据得到权重。人工 智能模型包含神经网络、支持向量机和遗传算法等, 近年来神经网络的研究较为热门。其中,循环神经 网络(Recurrent Neural Network, RNN)能够通过记 忆单元建立时序数据的长期依赖关系,具有对数据 的动态适应能力;长短期记忆递归神经网络(Long Short Term Memory, LSTM) 在循环神经网络的基础 上进行拓展,提出了遗忘门的结构,使得神经网络能 够学习到较长时间段的序列关系: Transformer 模型 基于传统的编码器 - 解码器模型,使用注意力机制 来代替传统的循环神经网络的结构,取得了较好的 结果:Informer 模型使用概率稀疏自注意力(Prob-Sparse Self-Attention Mechanism)和自注意力蒸馏机 制提取了更长的时间特征[8-14]。混合方法综合使 用统计方法和人工智能技术结合<sup>[15]</sup>。Yang 等人<sup>[16]</sup> 首次将趋势季节局部加权回归分解(STL 分解)、位置编码和卷积神经网络结合用于时间序列预测。ESS-RNN<sup>[17]</sup>使用指数平滑模型提取时间序列中的季节性因素并将其去除。Huang 等人<sup>[18]</sup>结合条件生成对抗网络、卷积神经网络和长短期记忆网络,用于光伏功率的预测。

简单的统计模型难以处理复杂的特征关系,并且建模难以重复使用。人工智能模型能够处理数据中的复杂任务,并且具有复用性,但是效果并不稳定,难以根据不同数据的特性进行建模。混合模型在实践中较为常用,它能够综合统计模型和人工智能模型的优点,提升预测的准确度。但是混合模型涉及到学科交叉,要求使用不同的方法组合使用神经网络。

本文提出了一种新的网络结构,通过 STL 分解与神经网络相结合的方式进行光伏功率预测,研究了这种结构中参数对实验结果的影响,并通过使用真实光伏功率数据集进行统计研究,验证了 STL-Former 模型的有效性。

# 1 STL-Former 模型

## 1.1 光伏功率预测数学模型

光伏功率预测问题是一种回归问题,通常被描述为:

$$\hat{Y} = f(X) \tag{1}$$

式中:X—历史数据的输入矩阵;Y—光伏功率的预测值; $f(\cdot)$ —各种回归方法的抽象函数。

输入矩阵包括有功功率、频率、风速、天气温度、 天气相对湿度、全局水平辐射、漫射水平辐射、风向 等向量。通常根据预测任务的不同而选取不同的时 长,本文选取 2 h 和 28 h 作为预测时长。而抽象函 数可以是统计建模方法、深度学习的卷积神经网络 或是其他各种模型集成。

一般认为光伏功率预测任务所处理的数据是一

种历史时间序列数据,其输入可以描述为,

$$X = x_{t-96}, x_{t-95}, x_{t-94}, \cdots, x_{t-1}, x_t$$
 (2)  
式中: $x_t$ — $t$  时刻的输入数据,由多个特征组成。在本文模型中,采用 96 个数据点(8 h)作为输入进行预测。

## 1.2 相关模型

## 1.2.1 Informer 模型

Informer 模型是基于 Transformer 模型的长时间序列预测模型,通过概率稀疏自注意力将时间复杂度和内存复杂度降低。同时,通过卷积层减半的方式来突出主导注意力,从而有效处理超长序列。模型通过全连接层进行长时间序列的预测,能够提高长序列预测的推理速度并防止误差累积。此外该模型还使用了类似 Transformer 的编码器 – 解码器结构,其中编码器由两部分组成,第1部分由概率稀疏注意力层和标准化层构成,其中注意力层仅对编码器的输入来自注意力;第2部分由前馈层和标准化层构成。并把最后一个编码器的输出作为其中一个概率稀疏注意力机制的解码器输入。解码器由类似的3个结构块组成,分别为两个多头注意力块和1个前馈块。

#### 1.2.2 STL 分解方法

STL 分解是一种常用的时间序列分解方法,通常将数据分解为趋势项、季节项和残差项。该方法多次使用局部多项式回归技术进行平滑估计,以获得每个分量的近似值。其总体公式如下:

$$Y_t = T_t + S_t + R_t \tag{3}$$

通过式(3)将预测数据分解成为了趋势项 $T_t$ 、季节项 $S_t$ 和残差项 $R_t$ 。其中, $Y_t$ 表示在t时刻的观测值(原始数据); $T_t$ 表示t时刻趋势分量,代表时间序列数据中的长期趋势; $S_t$ 表示t时刻季节项分量,捕捉时间序列数据中的周期性重复性的变化; $R_t$ 表示t时刻残差项分量,代表未被趋势项和季节项解释的随机波动。

STL 分解的主要步骤如图 1 所示。首先,将原始数据根据分解周期的大小分成多个子序列,再对子序列进行局部多项式平滑处理,去除掉噪声和快速变化的部分,得到初步的趋势估计。然后,计算每个时间的季节平均量,来估计季节项分量,通过相减

的方式去除趋势项和季节项数据,再次进行平滑处理得到残差项分量。通过重复分解数次,直到趋势项、季节项的估计值收敛为止。经过部分实验发现,STL分解的周期并不会影响最终的实验结果,因此将对照实验中的分解周期 T 设置为 24。

## 1.3 STL-Former 模型

STL-Former 模型首先将原始的时间序列数据进行 STL 分解,拆分为季节项、趋势项和残差 3 部分。把 STL 分解得到的季节项和趋势项作为额外的输入特征,与原始的时间序列数据一起输入到网络模型中。首先,在该模型中使用编码器对输入的时间序列数据和附加的季节性、趋势性特征进行建模,随后采用概率自注意力机制和全局局部信息融合的机制,可以有效捕捉输入序列的全局依赖关系和局部模式。根据编码器的输出,使用解码器生成光伏功率的预测结果。最后,通过均方误差对 STL-Former模型进行优化,使其能够准确预测未来时间点的光伏功率。

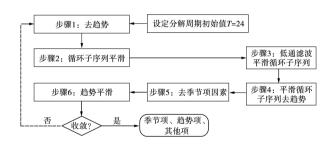


图 1 STL 分解流程图

Fig. 1 STL decomposition flow chart

将 STL 分解得到的季节性和趋势性特征作为额外的输入特征, STL-Former 模型能够更好地利用时间序列数据中的季节性和趋势性信息, 从而提高预测性能。这种结合 STL 分解与 Informer 模型的方法可以在光伏功率预测任务中得到更准确可靠的结果。STL-Former 模型的架构如图 2 所示。

编码器由自注意力层、卷积层和池化层构成,编码器结构可以进行 n 层的堆叠,最后进行向量的拼接,构成编码器的输出。编码器的输出作为自注意力的查询矩阵输入解码器结构中,由于数据仅作为自注意力向量运算的中间向量,因此向量的维数并不会影响自注意力的实现。解码器接收输入数据后

进行遮盖,通过遮盖不同的数据,使网络能够并行进行训练。遮盖数据会串行经过两个注意力层:第1个注意力层会进行一次自注意力,第2个注意力层会将编码器的输出结果作为查询矩阵后再使用注意

力机制。通过全连接网络进行输出,而不是通过递推的方式逐渐生成结果,输出不同长度的结果,避免了推理生成所造成的误差累积,提升了网络预测的准确性。

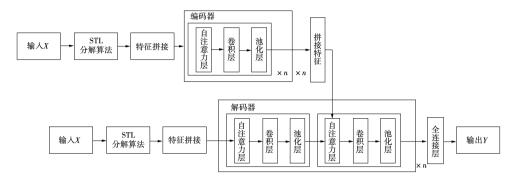


图 2 STL-Former 架构图

Fig. 2 Architecture diagram of STL-Former

# 2 实验结果与分析

#### 2.1 模型输入

神经网络的训练包括训练、验证和测试 3 个阶段,每个阶段都会将过去 96 个时间节点的数据作为输入,每个节点数据包含了湿度、温度和辐照度等特征信息。这些信息作为输入特征输入到神经网络中,t-1,t-2,…,t-96 作为输入,t,t+1,…,t+24 (根据需要预测时间长度)作为输出,根据预测时间长度的不同输出的时间长度也不同。

#### 2.2 评价指标描述

使用平均绝对误差、均方误差、均方根误差和平均绝对百分比误差、平均平方百分比误差 5 个性能指标作为评价指标来评价模型的优劣。

平均绝对误差:

MAE = 
$$\frac{1}{m} \sum_{i=1}^{m} |(y_i - \hat{y}_i)|$$
 (4)

均方误差:

MSE = 
$$\frac{1}{m} \sum_{i=1}^{m} (Y_i - \hat{Y}_i)^2$$
 (5)

均方根误差:

RMSE = 
$$\sqrt{\frac{1}{m} \sum_{i=1}^{m} (y_i - \hat{y}_i)^2}$$
 (6)

平均绝对百分比误差:

$$MAPE = \frac{1}{m} \sum_{i=1}^{m} \left| \frac{y_i - y_i}{y_i} \right|$$
 (7)

平均平方百分比误差:

$$MSPE = \frac{1}{m} \sum_{i=1}^{m} \left( \frac{y_i - \hat{y}_i}{\gamma_i} \right)^2$$
 (8)

式中 $:\hat{y}_i$ —光伏功率的预测值 $:y_i$ —光伏功率标准化后的值。

#### 2.3 数据分析

使用 Alice Spring 数据中的 DKASC 数据集,选取 2016年4月1日~2016年9月15日的数据,采样间隔为5 min。通过96个时间节点预测不同时间长度的时间节点,一共有48 104 条数据。

从下载的数据中选取实际功率、功率因数、相电压、总谐波失真电压平均值、风速、温度、湿度、全球辐照水平、漫射辐照水平、风向、日降雨量、全局辐照倾斜及漫射辐照倾斜作为分析特征。相关特征的偏相关系数如图 3 所示,颜色深浅表明相关性的强度,深色表明正相关性强,浅色表明负相关性强,介于两者中间的相关性较弱。由图 3 可知,输出实际功率相关性最高的是全球辐照水平和全局辐照倾斜,相关性分别为 0. 93 和 0. 99;温度、漫射辐照水平、漫射辐照倾斜等特征与有功功率的相关性较弱;日降雨量、风向、温度和频率与有功功率相关性较低。

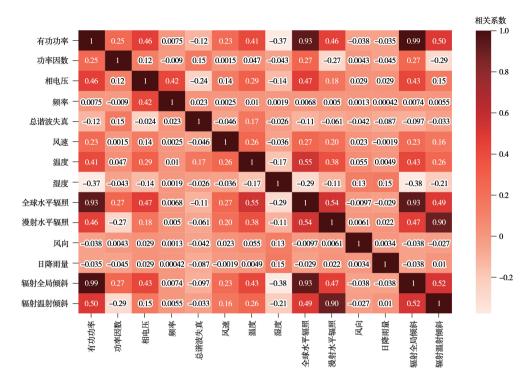


图 3 特征的相关系数图

Fig. 3 Correlation coefficient diagram of features

## 2.4 模型参数设置

为了加速神经网络的收敛,使用规范化的方法,通过每个数据减去均值除以标准差的方式使数据规范化。STL分解周期为24,使用均方误差函数作为训练的损失函数,Dropout设置成为0.05,提前停止训练的耐心设置为3,训练的轮数为5,批次大小设置为32。

#### 2.5 实验结果

光伏功率预测结果如图 4 所示。由图 4 可知, 绝大多数的预测结果都集中在第一象限角平分线周 围,这表明大多数情况下模型能够准确地预测光伏 功率的结果。一部分数据偏离较远,说明出现了预 测功率远小于实际功率的情况。

光伏功率预测误差如图 5(a) 所示。由图 5(a) 可知,共有 100 000 条数据预测进行了误差的评估,均值为 -0.081 6,说明我们的模型能够较为准确地给出无偏的预测结果,同时表明模型会略微倾向于给出较低的功率预测结果。标准差则反映了误差的不确定性,从图 5(a) 看出,误差较为均匀地分布在了横轴的两侧,远离横轴的数据较少,但横轴下侧的数据低谷大于上侧的峰值,说明模型倾向于给出较

低的功率预测结果。光伏功率预测值与真实值差值的统计结果如图 5(b)中所示,其中纵轴代表的是出现的差值数量,横轴代表的差值的大小,曲线是对统计结果进行了连线的可视化处理。从图中可以看出,预测误差近似成正态分布,误差值具有向左的偏度,说明预测结果比真实结果低的值多于预测结果比真实结果高的值,模型在预测的过程中偏向于预测较低的结果。原因可能是数据源当中大量的微小负值影响了训练。

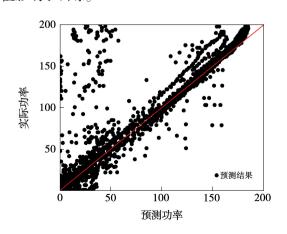


图 4 光伏功率预测值和实际值的散点图 Fig. 4 Scatter chart of predicted value and actual value of PV power

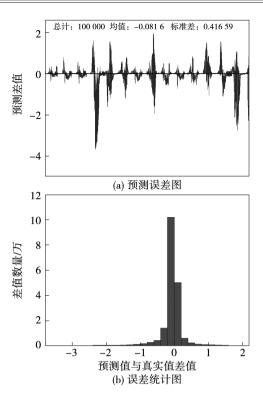


图 5 光伏功率预测值和实际值误差统计 Fig. 5 Error statistics of predicted value and actual value of PV power

## 2.6 不同算法对照

本文提出的 STL-Former 模型的各项数据都取得了较好的成果。预测时间长度为2h和28h模型的误差如表1和表2所示。由表1可知,在预测时间长度为2h的任务中,平均绝对误差相较于原模型降低了33.1%,均方误差降低了30.6%。其他各项评价指标均有所提升,在与传统的 LSTM 比较时,平均绝对误差降低了44.4%,均方误差降低了28.4%。由表2可知,28h光伏功率预测任务中,STL-Former 模型各损失函数均较其他模型降低了1倍以上。

表 1 2 h 光伏功率预测模型误差比较

Tab. 1 Error comparison of 2 h PV power prediction models

评价指标	STL-Former	Informer	LSTM	RNN
MAE	0.176	0.263	0.317	0.409
MSE	0.180	0.259	0.252	0.271
RMSE	0.425	0.509	0.502	0.521
MAPE	0.762	1.091	1.122	1.021
MSPE	55.758	84.723	94. 531	50.641

表 2 28 h 光伏功率预测模型误差比较

Tab. 2 Error comparison of 28 h PV power prediction models

评价指标	STL-Former	Informer	LSTM	RNN
MAE	0.170	0.402	0.598	0.558
MSE	0.154	0.433	0.612	0.575
RMSE	0.392	0.658	0.782	0.758
MAPE	0.748	1.673	1.484	1.516
MSPE	83.506	237.468	106.454	121.742

单变量的预测结果如图 6 所示。可以看到, STL-Former 模型的损失值最低,效果最好,说明该模型能够更好地完成短、中期光伏功率序列的预测。

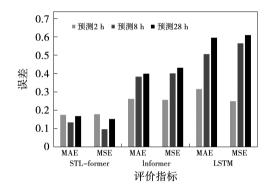


图 6 评价指标单变量预测结果图

Fig. 6 Univariate prediction results of evaluation indexes

#### 2.7 STL 分解周期对实验结果的影响分析

不同 STL 分解周期对实验结果的影响如图 7 所示。由图 7 可知,不同时间长度的任务目标和分解周期的预测结果不同。对于光伏功率为 2 h 的预测任务来说,选取分解周期为 96 较优。对于光伏功率为 28 h 的预测任务,分解周期选取 720 较优。

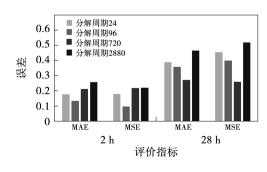


图 7 不同 STL 分解周期预测结果 Fig. 7 Prediction results of different STL decomposition periods

预测时间长度为2h和28h的不同STL分解周期比较结果分别如表3和表4所示。由表3可知,使用分解周期为96(8h)的STL分解方法可以获得较好的预测精度,并且预测结果在各项指标上普遍低于其他长度。说明STL分解提取出的光伏功率序列特征是有效的,并可以提升神经网络的训练结果。由表4可知,在预测时间长度为28h的序列光伏功率预测任务中,当采用720(144h)分解周期时,所得到的分解特征较为优秀,并且4个指标的结果均低于其他周期长度。这进一步证明了STL分解提取的特征是有效的,能够为神经网络的训练提供帮助。

表 3 不同 STL 分解周期比较结果(预测时间长度 2 h)

Tab. 3 Comparison of different STL decomposition periods

(Predicted time is 2 h)

评价指标	24	48	96	336	720	1 440
MAE	0.176	0.140	0.135	0.170	0.211	0.220
MSE	0.180	0.132	0.098	0.154	0.218	0. 192
RMSE	0.425	0.363	0.312	0.392	0.467	0.438
MAPE	0.762	0.567	0.512	0.748	0.986	1.112
MSPE	55.75	34.02	21.74	83.50	98.90	127.45

表 4 不同 STL 分解周期比较结果(预测时间长度 28 h)

Tab. 4 Comparison of different STL decomposition

periods (Predicted time is 28 h)

评价指标	24	48	96	336	720	1 440
MAE	0.388	0.382	0.357	0.287	0.270	0.358
MSE	0.453	0.448	0.397	0.290	0.258	0.370
RMSE	0.673	0.669	0.630	0.539	0.508	0.609
MAPE	1.649	1.580	1.522	1.461	1.413	1.601
MSPE	257.1	231.8	217.8	234.8	223.9	249.7

通过以上实验结果可以得出,STL 分解方法能够成功提取有效的光伏功率序列特征,提升神经网络训练精度,但在不同任务长度下分解周期的最优值不同。这为光伏功率预测任务中的神经网络训练提供了一种有效的特征提取方法。

# 3 结 论

采用改变模型参数和统计的方式,研究不同参数对神经网络最终预测结果的影响,并与其他网络

进行对比,得出结论:

- (1) 所提出的 STL-Former 模型能够更好地应用于中、短期光伏功率预测任务,具有更高的预测精度。相较于 Informer 模型,平均绝对误差降低了33.1%、均方误差降低了30.6%
- (2)针对不同时间长度的预测任务,最佳的STL分解周期不同。在2h的光伏功率预测任务中取STL分解的周期为96(8h)效果较好,在28h的光伏功率预测任务中取分解周期为720(60h)效果较好。

# 参考文献:

- [1] IRENA. Renewable capacity statistics 2023 [R]. 978-92-9260-525-4, Abu Dhabi; International Renewable Energy Agency, 2023.
- [2] PAZIKADIN A R, RIFAI D, ALI K, et al. Solar irradiance measurement instrumentation and power solar generation forecasting based on artificial neural networks (ANN): A review of five years research trend [J]. Science of the Total Environment, 2020, 715:136848.
- [3] MOHANTY S, PATRA P K, SAHOO S, et al. Forecasting of solar energy with application for a growing economy like India: Survey and implication[J]. Renewable and Sustainable Energy Reviews, 2017,78:539-553.
- [4] FRANK V,ZACHARY D, JOSH P, et al. Effects of changing spectral radiation distribution on the performance of photodiode pyranometers [J]. Solar Energy, 2016, 129;224 235.
- [5] AMEEN B, BALZTER H, JARVIS C, et al. Modelling hourly global horizontal irradiance from satellite-derived datasets and climate variables as new inputs with artificial neural networks [J]. Energies, 2019, 12:148.
- [6] BOX G E P, JENKINS G M, REINSEL G C. Time series analysis: Forecasting and control[M]. London: Palgrave Macmillan, 2013.
- [7] GARDNER E S. Exponential smoothing: The state of the art[J]. Journal of Forecasting, 1985, 4(1):1-28.
- [8] ELMAN J L. Finding structure in time [J]. Cognitive Science, 1990,14:179-211.
- [9] LIM B, ZOHREN S, ROBERTS S. Recurrent neural filters: Learning independent Bayesian filtering steps for time series prediction
  [C]// International Joint Conference on Neural Network: IEEE,
  Glasgow, UK, July 19 24, 2020.
- [10] TOM Y, DEVAMANYU H, SOUJANYA P, et al. Recent trends in deep learning based natural language processing [J]. Journal of Land Use Science, 2018, 13 (1/3):55-75.

(下转第153页)