

基于VMD-TCN的太阳活动指数 $F_{10.7}$ 的短期 预报

王璐瑶

南京信息工程大学数学与统计学院, 江苏 南京

收稿日期: 2024年2月18日; 录用日期: 2024年3月22日; 发布日期: 2024年3月31日

摘要

$F_{10.7}$ 是波长为10.7 cm太阳辐射通量, 它经常作为不同空间天气模型、大气密度经验模型中一个重要的参数输入, 同时也是最广泛使用的衡量太阳活动水平强弱的参数之一, 因此对其进行准确的预报是空间天气研究的一个重点关注的方向。本文利用1957~2019年观测的太阳辐射通量 $F_{10.7}$ 数据建立了基于VMD-TCN的 $F_{10.7}$ 预测模型, 其中将1957~2008年作为模型训练集, 2009~2019年作为模型的测试集, 并使用留一法对数据集进行分组, 进行多次交叉验证。实验结果表明: VMD-TCN模型预报 $F_{10.7}$ 的均方根误差1~4 sfu, 平均绝对误差百分比为0~2 sfu, 相关系数高达0.99。与基于相同数据集建模的其他模型如TCN, VMD-ELM模型相比, VMD-TCN模型预测 $F_{10.7}$ 的均方根误差更低, 相关系数更高, 模型综合预测效果最佳。

关键词

$F_{10.7}$ 预测, VMD分解, TCN模型

Short-Term Prediction of Solar Activity Factor $F_{10.7}$ Based on VMD-TCN Model

Luyao Wang

School of Mathematics and Statistics, Nanjing University of Information Science and Technology, Nanjing
Jiangsu

Received: Feb. 18th, 2024; accepted: Mar. 22nd, 2024; published: Mar. 31st, 2024

Abstract

$F_{10.7}$ is the solar radiation flux with a wavelength of 10.7 cm, which is often used as an important

parameter input in different space weather models, empirical models of atmospheric density, and is also one of the most widely used parameters to measure the strength of solar activity level, so its accurate forecasting is a key concern in space weather research. In this paper, a $F_{10.7}$ prediction model based on VMD-TCN is established using the observed solar radiation flux $F_{10.7}$ data from 1957~2019, in which 1957~2008 is taken as the training set of the model, and 2009~2019 as the test set of the model, and the datasets are grouped by using the leave-one-out method for multiple cross-validations. The experimental results show that the VMD-TCN model predicts $F_{10.7}$ with the root mean square error of 1~4 sfu, the average absolute error percentage of 0~2 sfu, and the correlation coefficient is as high as 0.99. Compared with other models based on the same dataset, such as the TCN, VMD-ELM models, the VMD-TCN model predicts $F_{10.7}$ with a lower root mean square error and a higher correlation coefficient, which gives the best overall prediction effect of the model.

Keywords

$F_{10.7}$, Variational Modal Decomposition, Temporal Convolutional Networks

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

太阳活动对地球上的气候、电磁场和通信系统等都有重要影响, 而 $F_{10.7}$ (2800MHz, 10.7 cm 太阳辐射通量) 是一个较好的衡量太阳活动水平的典型参数, 能够反映太阳活动的周期性变化[1] (Tapping 2013)。 $F_{10.7}$ 指数在航天器轨道预报的大气密度计算及影响通信的电离层预报中, $F_{10.7}$ 指数都是重要的参数。例如, $F_{10.7}$ 被用作电离层模型的控制参数, 用于计算无线电信号特性的变化[2] (Ortikov et al. 2003)。 $F_{10.7}$ 也被广泛应用于卫星、导航、通信和陆地气候[3] [4] (Huang, Liu, and Wang, 2009; Ya Yaetal., 2017 等)。因此, 对 $F_{10.7}$ 进行准确的预报不仅对军事活动的进行有着重要的价值, 而且在空间天气预报这一科学研究方面也具有比较大的意义[5] [6] (Swarup et al., 1963; Tapping and DeTracey, 1990; Henney et al., 2012)。

$F_{10.7}$ 指数预测模型基于时间序列模型。许多研究人员使用不同的方法来建立 $F_{10.7}$ 的预测模型。Mordvinov 等[7]人(1986 年)使用乘法自回归模型来预测 $F_{10.7}$ 的月平均值, 但该模型在预测 $F_{10.7}$ 月平均值时误差较大。Warren 等人[8] (2017 年)为每个预测日期建立了优化的独立模型, 结果表明这种方法的预测效果通常优于自回归方法。Zhong 等人[9] (2005)利用奇异谱分析信号处理技术预测了未来 27 天的太阳活动 $F_{10.7}$ 指数。研究表明, 该方法在预测 $F_{10.7}$ 指数的周期性变化方面表现良好。Henney 等[10] (2012)利用能量传输模型生成的全球太阳磁场预测了 $F_{10.7}$, 提前 1 天的皮尔逊相关系数为 0.97。Liu [11]等人 (2018)应用 Yeates [12] (2007)和 Worden [13] (2000)的两个模型预测了 $F_{10.7}$ 的短期变化。在太阳活动水平较低时, 模型的预测值更接近观测值。

随着机器学习和神经网络的快速发展。许多研究人员对机器学习和神经网络强大的学习能力越来越感兴趣, 并利用它们来研究太阳活动的变化。Huang 等人[3] (2009 年)使用支持向量机回归方法预测了太阳活动 $F_{10.7}$ 的日值。Xiao 等人[14] (2017)使用反向传播神经网络(BP)预测太阳活动日平均指数 $F_{10.7}$, 进行短期预测。结果表明, 利用 BP 神经网络预测太阳活动日指数 $F_{10.7}$ 优于 Huang 等[3] (2009)的结果。Luo 等[15] (2020)提出了一种多步骤预测 10.7 厘米射电通量的方法。该方法结合了经验模式分解(EMD)和反

向传播神经网络(BP), 构建了预测 $F_{10.7}$ 值的 EMD-BP 模型。Zhang 等人[16] (2020)提出了一种利用长短期记忆网络(LSTM)方法对太阳活动日平均指数 $F_{10.7}$ 进行短期预报的方法。该预报具有 0.98 的高相关系数和 6.20~6.35 sfu 的低均方根误差。尽管上述基于递归神经网络(RNN)的结构及其变体实现了良好的 $F_{10.7}$ 预测精度, 但模型的训练过程往往耗费大量时间和计算内存, 而且在网络训练过程中还经常遇到梯度爆炸或梯度消失等问题[17] [18] (Zachary et al., 2015, Yang et al., 2021)。同时, 太阳辐射通量变化非常剧烈, 其序列值本身的非线性非平稳程度就很强, 所以关于太阳辐射通量的中短期预测问题仍然有很多更适合非线性系统建模的机器学习模型或组合模型可以尝试, 并且极端事件预测能力有待于进一步提高。针对这些问题, 本文尝试引入了变分模式分解算法(VMD), 对 $F_{10.7}$ 指数进行分解。VMD 算法能有效处理时间序列数据的周期性和强非线性。其次引入 TCN 神经网络对 $F_{10.7}$ 指数预测。TCN 神经网络对低复杂度和低非线性的时间序列具有良好的预测能力。因此, 将 VMD 方法和 TCN 模型相结合, 构建了 $F_{10.7}$ 指数的 VMD-TCN 预测新模型, 并试图提高 $F_{10.7}$ 指数的预测精度。并将预测结果与基于相同数据集建模的其他模型如 TCN, VMD-ELM 模型相比, 验证 VMD-TCN 预测 $F_{10.7}$ 的有效性和可靠性。

2. 数据来源与处理

2.1. 数据来源

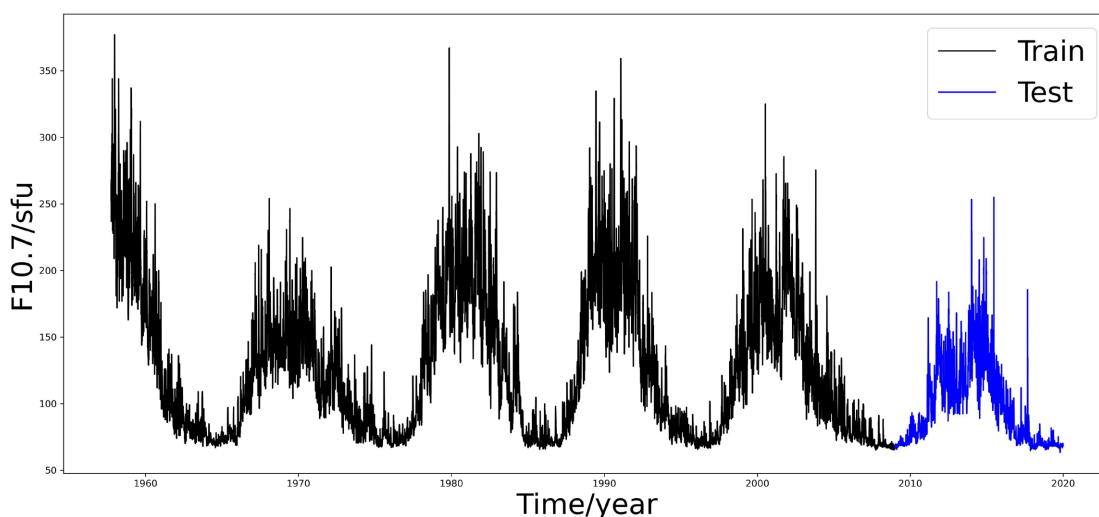


Figure 1. 1957 to 2019 data for $F_{10.7}$, where black indicates the training set and blue indicates the test set
图 1. $F_{10.7}$ 的 1957 到 2019 的数据, 其中黑色表示训练集, 蓝色表示测试集

$F_{10.7}$ 代表波长为 10.7 cm 的太阳辐射通量, 该指数大小反映太阳活动性强弱。 $F_{10.7}$ 是记录太阳活动水平时间最长的指数之一, 10.7 厘米太阳通量以太阳通量单位($1 \text{ sfu} = 10^{-22} \text{ W} \cdot \text{m}^{-2} \text{ Hz}^{-1}$)。本文收集了 1957 年至 2019 年的太阳辐射通量($F_{10.7}$)日值数该数据来源于美国国家海洋和大气管理局网站 (<http://www.celestrak.com/SpaceData/>)。 $F_{10.7}$ 代表波长为 10.7 厘米的太阳辐射通量, 该指数的大小描述了太阳活动的强度。10.7 厘米的每日太阳通量数据来自美国国家海洋和大气管理局的网站。每天进行三次通量测定。每个 10.7 厘米太阳通量测量都以三个值表示: 观测值、调整值和 URSI 系列 D 值(绝对值)。观测值是太阳射电望远镜测量的数字。它被两个量调制: 太阳活动水平和地球与太阳之间的变化距离。由于它是太阳活动撞击地球所导致的排放量的测量方法, 因此它是研究地球现象时使用的量。在研究太阳时, 不希望有地球与太阳之间距离变化所引起的 10.7 厘米太阳通量的年度调制。然而, 在太阳通量监测器需要准确获取和跟踪太阳的星历计算过程中, 得到的副产品之一是太阳和地球之间的距离。因此,

产生了一个额外的值，称为调整值，它考虑了地球 - 太阳距离的变化，并代表平均距离。通量密度的绝对测量是相当困难的。天文学家试图通过应用比例因子来匹配不同频率的太阳通量密度数据与频谱。通过将每个波长与校准的频谱相结合，可以获得一系列 D 通量，其中 D 通量等于 0.9 乘以调整通量[19] (Tanaka et al., 1973)。

三月至十月的测量时间分别为世界时 1700 时、2000 时(当地正午)和 2300 时。然而，由于地处山谷和相对较高的纬度，在一年的其他时间里不可能保持这些时间。因此，从 11 月到 2 月，通量测定时间在 1800、2000 和 2200 改变，以使太阳在地平线以上足够高，可以很好地测量。因此，我们选择调整后的通量值 $F_{10.7}$ 在晚上 8 点测量的值，我们使用 1957 年到 2019 年的 $F_{10.7}$ 数据，并选择 1957 年 11 月 10 日至 2008 年 12 月 31 日 51 年的数据作为训练集，2009 年 1 月 1 日至 2019 年 12 月 31 日 11 年的数据作为测试集。图 1 显示了 $F_{10.7}$ 数据随时间变化的曲线图，黑色代表训练数据，蓝色代表测试数据。它们将在下文中被分别使用。

2.2. 数据处理

本文对数据进行标准化处理。在建模之前对输入数据进行预处理操作，确保模型预测结果的适用性、可靠性和对标准的统一性。我们通过将原始数据转换为无量纲的标准化值，从而避免不同属性值对不相同的指标的影响，因此实验结果更具备比较性，不会影响实验结果的精确率。一般常用的数据标准化方法包括线性函数标准化(Min-Max 缩放)，Z-score 标准化和线性刻度标准化。

本文模型数据主要使用 Min-Max 归一化方法，首先对原始流量数据做预先处理，并根据式 1 将时间序列输入 X 转换为 X'

$$X' = \frac{X - X_{\max}}{X_{\max} - X_{\min}} \quad (1)$$

式中 X_{\max} 是数据样本中的最大值，而 X_{\min} 是数据样本中的最小值。归一化的样本数据 X 通过式 2 被缩放到 -1 和 1 之间的值。

$$X = 2 * X' - 1 \quad (2)$$

3. 方法

3.1. VMD 分解方法

VMD 算法常用于处理非平稳信号，可有效地提取出数据特征[18]。对于输入信号，首先构造变分问题，假设原始信号 f 被分解为 k 个分量，保证分解序列为具有中心频率的有限带宽的模态分量，同时各模态的估计带宽之和最小，约束条件为所有模态之和与原始信号相等，则 VMD 算法由此产生约束的变分问题为：

$$\begin{cases} \min_{\{u_k\}, \{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right)^* u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \\ s.t. \sum_{k=1}^K u_k = f \end{cases} \quad (3)$$

式中: K 为需要分解的模态个数(正整数)， $\{u_k\}$ ， $\{\omega_k\}$ 分别对应分解后第 K 个模态分量和中心频率， $\delta(t)$ 为狄拉克函数， $*$ 为卷积运算符。

引入 Lagrange 乘法算子 λ ，将约束变分问题转变为非约束变分问题，得到增广 Lagrange 表达式为：

$$L(\{u_k\}, \{\omega_k\}, \lambda) = \alpha \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) u_k(t) \right] e^{-j\omega_k z} \right\|_2^2 + \left\| f(t) - \sum_k u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_k u_k(t) \right\rangle \quad (4)$$

式中： α 为二次惩罚因子，作用是降低高斯噪声的干扰。利用交替方向乘子(ADMM)迭代算法结合 Parseval/Plancherel、傅里叶等距变换，优化得到各模态分量和中心频率，并搜寻增广 Lagrange 函数的鞍点，交替寻优迭代后的 $\{u_k\}$ ， $\{\omega_k\}$ 和 λ 的表达式如下，详细过程如下所示：

$$\begin{cases} \hat{u}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)^2} \\ \omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_k^{n+1}(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k^{n+1}(\omega)|^2 d\omega} \\ \hat{\lambda}^{n+1}(\omega) = \hat{\lambda}^n(\omega) + \gamma \left(\hat{f}(\omega) - \sum_k \hat{u}_k^{n+1}(\omega) \right) \end{cases} \quad (5)$$

式中， γ 为信号的噪声容忍度。通过反复迭代直到满足收敛条件或达到最大迭代次数即停止，收敛条件为：

$$\sum_k \frac{\|\hat{u}_k^{n+1} - \hat{u}_k^n\|}{\|\hat{u}_k^n\|} < \varepsilon \quad (6)$$

3.2. TCN 网络

时序卷积网络(Temporal Convolutional Network, TCN)在 2018 年由 Bai 等[20]提出，它是一种新型的时间序列预测模型，是在卷积神经网络(CNN)的基础上进行改进的，其模型架构包括因果卷积(Causal Convolution)、空洞卷积(Dilated Convolution)和残差模块(Residual Block)。TCN 的梯度稳定使其能够有效避免模型训练过程中发生梯度消失或梯度爆炸导致模型训练失败的问题，同时具有灵活的感受野等优点。因此，我们将 TCN 应用于 $F_{10.7}$ 的预测中。

3.2.1. $F_{10.7}$ 序列建模

对于单变量时间的预测，TCN 模型的输入是时间序列的滞后观测值，输出是未来的序列值，其中每一组输入模式都是由在时间序列中移动一个固定长度的窗口组成[21]。预测的原理如图 2 表示：

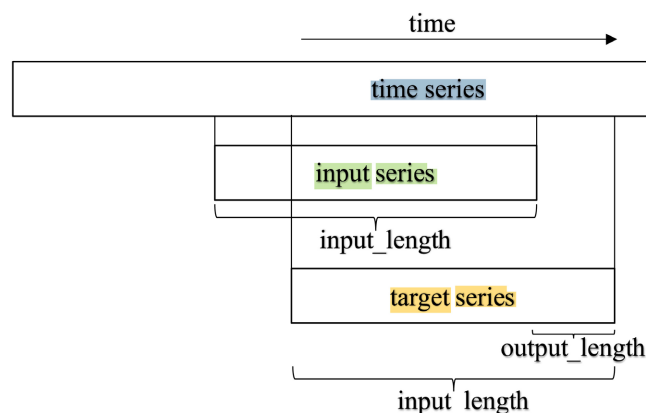


Figure 2. Schematic of sequence data prediction

图 2. 序列数据预测示意图

假设输入数据为 $x = (x_0, x_1, \dots, x_T)$ ，对应期望的输出序列为 $y = (y_0, y_1, \dots, y_T)$ ，其中 x, y 两个序列要满足因果关系，即须使用先前时刻观测得到的输入 x_0, x_1, \dots, x_{t-1} 来预测 t 时刻的输出 y_t ，TCN 网络的建模目标是产生映射的任何隐函数 f ，即本文序列的预测问题可表示为：

$$\hat{y}_1, \dots, \hat{y}_{T+1} = f(x_0, x_1, \dots, x_t, \dots, x_T) \tag{7}$$

其中， x_i 和 y_i 分别为 i 时刻的观测值与预测值， f 为 TCN 网络训练出来的函数映射。

时间卷积网络(Time Convolutional Network, TCN)是在卷积神经网络(Convolutional Neural Network, CNN)基础上发展起来的算法之一[20]。TCN 采用一维卷积网络，由膨胀因果卷积和残差模块组成。下文将分别介绍。

3.2.2. 膨胀因果卷积原理

一维卷积可对时间序列进行运算并提取各种特征，但随着时间序列长度的增长，规则的卷积网络则需要更多的卷积层以接收更长的序列。扩展卷积则在卷积基础上进行了改进，允许卷积时的输入存在间隔采样，层数为 L 且卷积核大小为 k 的接受域为：

$$r = 2^{(L-1)}k \tag{8}$$

时间序列中元素 s 的因果扩展卷积运算 F 定义为：

$$F(s) = (\bar{x} * f)(s) = \sum_{i=0}^{k-1} f(i) \cdot \bar{x}_{s-d-i} \tag{9}$$

式中： $x = (x_0, x_1, \dots, x_T)$ 为输入向量， d 为扩展因子， $*$ 为因果扩展卷积运算符， f 为卷积核向量， k 为卷积核大小， $s-d-i$ 表示输入过去的方向。

膨胀因果卷积结构如图 3 所示，可以通过调节感受野尺寸以满足不同长度序列感知需求，从而解决 CNN 中时间建模长度受卷积核尺寸限制这一难题。

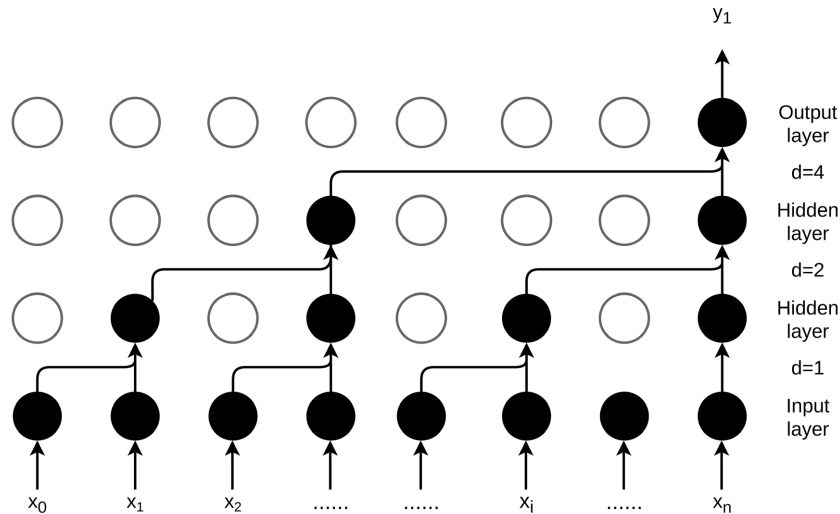


Figure 3. Inflated causal convolution
图 3. 膨胀因果卷积

3.2.3. 残差模块

TCN 中残差模块的结构如图 4 中残差链接使网络能够跨层数进行信息传递，从而避免层数过多而造成信息损失。引入残差卷积进行跳层连接，并进行 1×1 卷积操作，保证输入与输出保持一致。

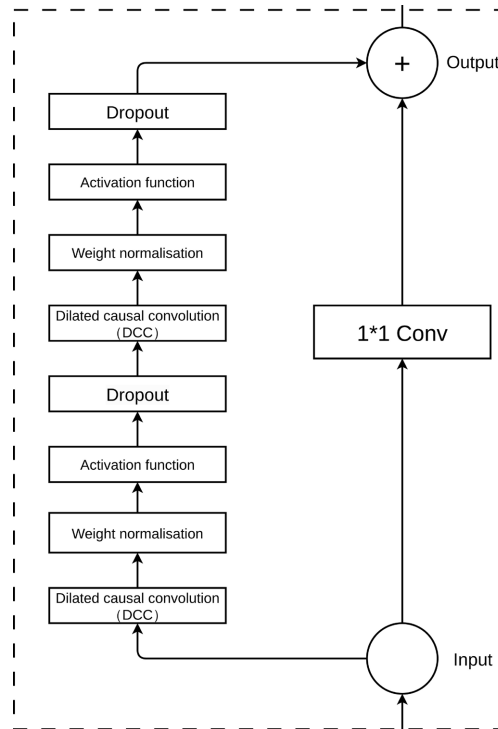


Figure 4. TCN residual module structure
图 4. TCN 残差模块结构

3.3. VMD-TCN 模型

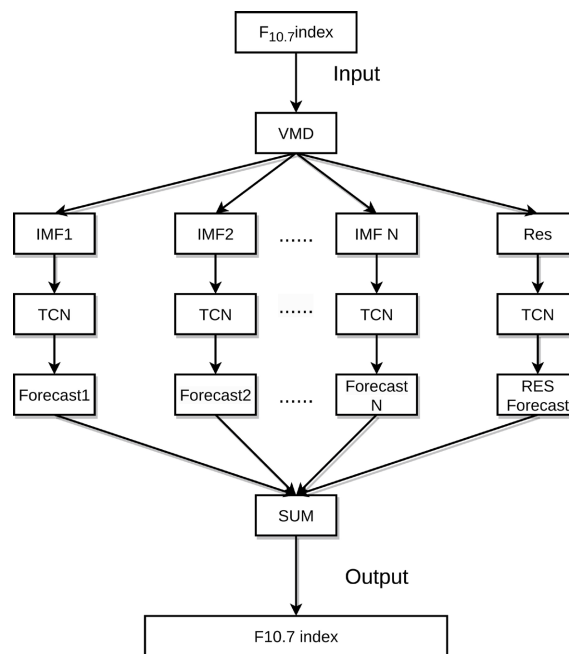


Figure 5. Technology roadmap
图 5. 技术路线图

VMD 可以减少具有高复杂性和强非线性的时间序列数据。作为一种改进的递归神经网络，TCN 网

络不仅能解决 RNN 无法处理长距离依赖性的问题,还能解决神经网络中常见的梯度爆炸或梯度消失等问题。它在处理时间序列数据方面非常有效。因此,我们尝试将 TCN 神经网络模型与 VMD 算法相结合,提前 1~3 天预测 $F_{10.7}$ 指数。将这种组合方法被称为 VMD-TCN 模型。具体步骤如图 5 所示。图 5 首先将原始数据通过 VMD 分解方法分解得到的多个 IMF 分量和一个残差分量,接下来使用 TCN 模型分别对各分量进行训练预测,最后将各分量的预测结果相加得到最终预测结果。VMD 可以减少具有高复杂性和强非线性的时间序列数据,而本文的数据具有非线性,非平稳的特点,通过 VMD 方法将数据分解成中高低频信号,可以使得 TCN 模型更好的捕捉数据内部的信息,从而达到提高预报效果的结果。

1. 我们使用 1957 年至 2019 年的 $F_{10.7}$ 数据,并使用 VMD 算法对 $F_{10.7}$ 数据进行分解,分解得到 5 个 IMF 分量和一个残差分量。

2. 使用 TCN 模型分别训练这些分解后的分量,在完成 TCN 模型的训练后,我们可以预测得到提前 1~3 天每个分量的预测结果。

3. 将各部分提前 1~3 天的预测结果相加,就可以得到 $F_{10.7}$ 提前 1~3 天的预测值。

由图 4 可知,时序卷积网络模型主要采用 Adam 梯度下降优化算法对其进行优化,并自动学习记忆时间序列的特征生成最优模型,最后通过全连接层后进行结果预测。另外本文在预测时使用 L1 范数进行正则化[22](Zhao et al., 2017),使用 Dropout 防止模型过拟合。为了量化本模型的预报性能,我们选取了 5 个评估量来衡量模型的性能,分别为平均绝对误差、均方根误差、平均绝对百分比误差、相关系数,这五个常见的模型评价指标去评价预测的预测效果。

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\hat{f}_i - F_i| \quad (10)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (f_i - F_i)^2} \quad (11)$$

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \frac{|f_i - F_i|}{f_i} \quad (12)$$

$$\text{MRE} = \frac{1}{N} \sum_{i=1}^N \frac{|f_i - F_i|}{f_i} * 100\% \quad (13)$$

$$R = \frac{\sum_{i=1}^N (f_i - \bar{f})(F_i - \bar{F})}{\sqrt{\sum_{i=1}^N (f_i - \bar{f})^2} \sqrt{\sum_{i=1}^N (F_i - \bar{F})^2}} \quad (14)$$

其中, MAE 表示平均绝对误差, MAPE 表示平均绝对百分比误差, RMSE 表示均方根误差, MRE 表示平均相对误差, R 表示线性相关系数, N 表示样本个数, F_i 表示预报值, f_i 表示观测值, \bar{f} 表示观测值期望, \bar{F} 表示预报值期望。每个指标都从不同角度去评估模型, MAE 表示预测值和实际值之间的平均绝对误差, MAPE 表示预测值和实际值之间的平均绝对百分比误差, RMSE 代表预测值和实际值之间的误差, MRE 表示表示预测值和实际值之间的平均相对误差, R 代表预测值与实际值之间的趋势拟合程度。因此 MRE, MAE 和 RMSE 和 MAPE 越小, R 越大, 模型预测越佳。

3. 结果与讨论

3.1. VMD 分解结果分析

本研究采用 VMD 算法对 1957 年至 2019 年的 $F_{10.7}$ 数据进行分解。分解后得到 5 个 IMF 分量和残差。

原始 $F_{10.7}$ 指数序列的值是分解后的 6 个子分量的和，残差可通过原始数据减去 5 个 IMF 分量表示。图 6 显示了 2015 年 $F_{10.7}$ 指数的分解结果、其中，IMF1 和 IMF2 为高频信号，IMF3 为中频信号，IMF4 和 IMF5 为低频信号。分解结果显示，IMF3, IMF4 和 IMF5 分量的曲线更平滑，拐点更少，复杂度更低，与原始 $F_{10.7}$ 数据相比，曲线更平滑、拐点更少、复杂度更低，这有利于神经网络 TCN 模型的建立。虽然 IMF1 和 IMF2 波动很大，但从定量的角度来看，IMF1 和 IMF2 所占的比例很小，因此对预测结果的影响也很小。所以使用神经网络模型可以有效地预测它们。

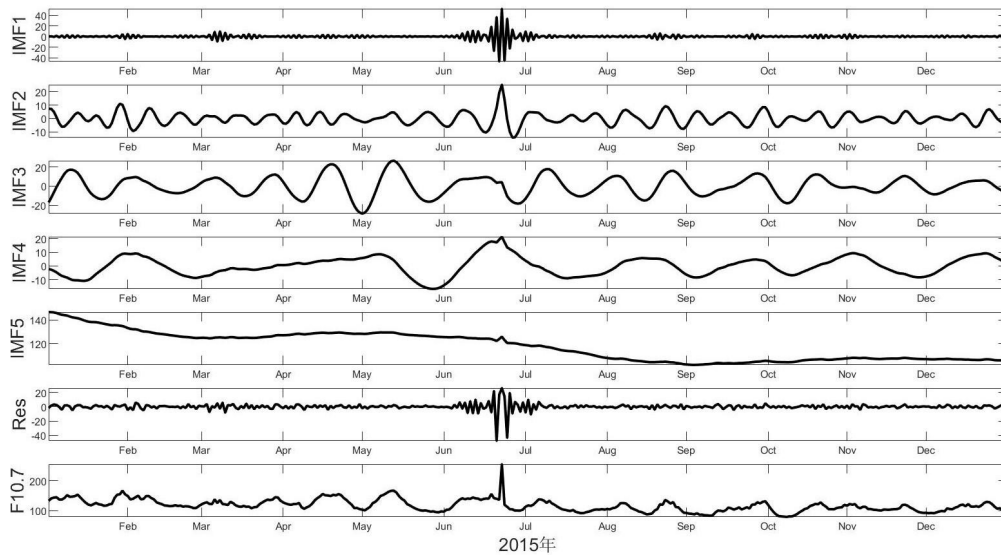


Figure 6. $F_{10.7}$ index after using VMD decomposition methodology during 2015 year
图 6. 2015 年 $F_{10.7}$ 指数使用 VMD 分解方法后的结果

3.2. 预报结果分析

Table 1. VMD-TCN model prediction errors (MAE, RMSE, MAPE) for $F_{10.7}$ data for the period 1996~2019 and R
表 1. VMD-TCN 模型对 1996~2019 年期间 $F_{10.7}$ 数据的预测误差(MAE, RMSE, MAPE)和 R

Year	1 Day in Advance				2 Day in Advance				3 Day in Advance			
	MAE (sfu)	RMSE (sfu)	MAPE (sfu)	R	MAE (sfu)	RMSE (sfu)	MAPE (sfu)	R	MAE (sfu)	RMSE (sfu)	MAPE (sfu)	R
2009	0.38	0.47	0.53	0.9917	0.68	0.84	0.96	0.9717	0.57	0.77	0.79	0.9609
2010	0.53	0.67	0.66	0.9929	1.02	1.32	1.27	0.9708	1.16	1.54	1.44	0.9551
2011	1.1	1.53	0.98	0.9980	2.22	3.21	1.89	0.9911	3.26	4.69	2.65	0.9675
2012	1.45	2.03	1.18	0.9940	2.74	3.99	2.22	0.9768	2.76	3.64	2.23	0.9833
2013	1.12	1.53	0.91	0.9970	2.24	3.01	1.80	0.9881	4.00	5.8	2.72	0.9764
2014	1.88	2.75	1.25	0.9948	3.47	5.01	2.33	0.9827	4.00	5.84	2.72	0.9764
2015	2.06	4.42	1.68	0.9750	3.18	5.79	2.63	0.9575	3.88	8.13	3.21	0.9145
2016	0.74	0.98	0.82	0.9965	1.37	1.86	1.50	0.9870	1.67	2.22	1.83	0.9809
2017	0.91	1.80	1.07	0.9869	1.42	2.79	1.65	0.9672	1.58	3.52	1.82	0.9467
2018	0.3	0.46	0.53	0.9936	0.71	0.89	1.02	0.9758	0.61	0.81	0.87	0.9662
2019	0.38	0.48	0.55	0.9909	0.72	0.90	1.03	0.9661	0.61	0.85	0.86	0.9503
Total	1.00	1.94	0.93	0.9979	1.79	3.15	1.67	0.9946	2.09	3.97	1.90	0.9912

表 1 显示了不同年份 VMD-TCN 的预测值和观测值之间的统计参数, 这反映了 VMD-TCN 模型在不同年份的性能表现。从表 1 中可以得出, 提前 1~3 天 $F_{10.7}$ 预测结果的均方根误差是 1~4 sfu, 平均绝对值误差是 1~3 sfu, 平均绝对百分比误差是 1~2 sfu, 相关性是 0.99, 这意味着模型的性能不会因为提前预测的时间的不同而出现显著性的差异。这体现了 VMD-TCN 模型的稳定性。

由表 1 可见, 在不同年份, 本文的预报精度存在一定差异。为评估本文方法在不同年份的精度状况, 以太阳光度强度为标准, 将完整太阳周期划分为太阳活动平静期与太阳活动活跃期, 并分析预报精度与太阳活动强度的关系, 如图 7 所示。在图 7(a)中, 红色点线图表示第 24 太阳活动周内各年年均 $F_{10.7}$ 变化情况, 其纵轴位于图 7(a)左边, 图中标注皆为红色。本文结合第 24 周实际状况, 将年均 $F_{10.7}$ 高于 110 sfu 的年份视为太阳活动活跃期, 其余年份为太阳活动平静期, 图 7(a)中绿色划线为太阳活动低年与高年的划分标准线。由图可见, 2011 年至 2015 年的 5 年间, $F_{10.7}$ 年均值显著高于其他年份, 视其为太阳活动活跃期; 其余年份 $F_{10.7}$ 年均值低且变化较缓慢, 视为太阳活动平静期。图 7(a)中柱状图表示本文方法预报第 24 太阳活动周 $F_{10.7}$ 的各年年均相对误差, 其纵轴位于图 7(a)右, 图中标注皆为蓝色。结合图 7(a)中线图和柱状图可以发现, 本文方法预报平均相对误差与太阳活动强度存在较强的负相关。图 7(b)显示了第 24 太阳活动周预测值与真实值, 红线表示 $F_{10.7}$ 的观测值, 蓝色线表示 VMD-TCN 模型的预测值。由图 7(b)可知, VMD-TCN 模型很好地预测出了 $F_{10.7}$ 的走势, 并且在大部分时间内, 真实值和预测值的数值大小吻合较好。尤其在 $F_{10.7}$ 峰值时, VMD-TCN 模型的预测效果也与真实值相符合, 并且在太阳活动高年期间, VMD-TCN 模型的预测结果也表现优异。

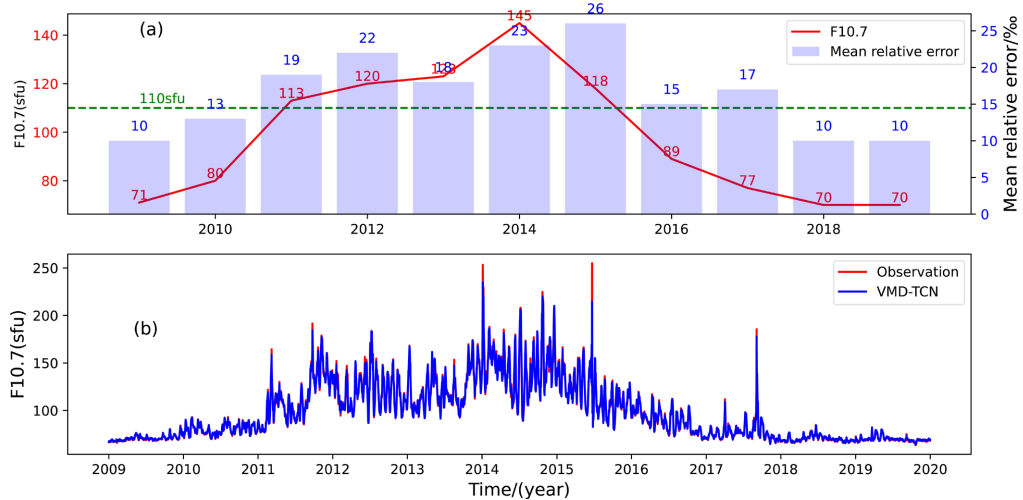


Figure 7. (a) Mean relative error versus $F_{10.7}$ annual mean; (b) Comparison of $F_{10.7}$ forecast values with observed values
图 7. (a) 平均相对误差与 $F_{10.7}$ 年均值关系; (b) $F_{10.7}$ 预报值与观测值的对比

为了更好的评估模型的性能, 我们基于同一数据集进行了对比实验。将 TCN 模型预报结果, 以及 VMD 分解方法与 ELM 模型相结合的预报结果, 与本文 VMD-TCN 模型提前 1~3 天的预报结果进行对比。预报对比结果如表 2 所示。表 2 展示了 TCN、VMD-ELM、VMD-TCN 三种模型在第 24 个太阳活动周期期间对太阳辐射通量 $F_{10.7}$ 提前 1~3 天的预测结果。从表 2 中可以看出, VMD-TCN 模型在第 24 个太阳活动周的提前 1~3 天的均方根误差分别为 1.94、3.15 和 3.97, TCN 模型在第 24 个太阳活动周的提前 1~3 天的均方根误差分别为 5.15、5.16 和 5.66。VMD-ELM 模型在第 24 个太阳活动周的提前 1~3 天的均方根误差分别为 2.08、3.45 和 4.19。VMD-TCN 模型的预测精度优于 TCN 模型和 VMD-ELM 模型, 并且 VMD-TCN 模型的相关系数, 平均绝对误差以及平均绝对百分比误差也相对更好。VMD-TCN 模型预报太阳辐射通

量 $F_{10.7}$ 可行的原因。首先, TCN 网络可以捕捉数据序列的历史信息来处理非线性问题, 并且相比 LSTM、GRU 等递归神经网络, 它不需要使用门控机制, 所以 TCN 更准确, 有更长的记忆, 且 TCN 不存在梯度消失、梯度爆炸问题等优点, 并从训练集中捕捉有用的信息来预测提前 1~3 天的太阳辐射通量 $F_{10.7}$ 。其次, VMD 算法可以减少非平稳数据序列的影响, 所以即使太阳辐射通量 $F_{10.7}$ 突然变化, VMD-TCN 模型也能很好地拟合 $F_{10.7}$ 。这些都表明 VMD-TCN 模型在处理太阳辐射通量 $F_{10.7}$ 数据的非平稳和非线性问题方面具有明显的优势。

Table 2. Comparison of the results of the models for the 24-week forecasts of solar activity

表 2. 各模型在太阳活动 24 周预报结果的对比结果

	Model	MAE	RMSE	MAPE	R
1-day ahead	TCN	2.77	5.15	2.51	0.9852
	VMD-ELM	1.01	2.08	0.92	0.9976
	VMD-TCN	1.00	1.94	0.93	0.9979
2-day ahead	TCN	2.80	5.16	3.50	0.9852
	VMD-ELM	1.83	3.45	1.66	0.9934
	VMD-TCN	1.79	3.15	1.67	0.9946
3-day ahead	TCN	3.99	5.66	3.89	0.9798
	VMD-ELM	2.19	4.19	1.99	0.9902
	VMD-TCN	2.09	3.97	1.90	0.9912

4. 结论

10.7 厘米($F_{10.7}$)的太阳辐射通量是太阳活动的重要指标。它在太阳物理学中的应用包括太阳活动水平的指标、太阳周期特征的预测等。针对 $F_{10.7}$ 观测时间长、具有一定周期的特点, 本文首次将基于机器学习的 TCN 相关理论和技术和信号处理中的 VMD 分解方法引入到空间天气的 $F_{10.7}$ 序列预测中, 通过太阳活动周数据建立了 VMD-TCN 预测模型。并基于同一数据集建立了其他模型 TCN, VMD-ELM 建立作比较, 发现 VMD-TCN 模型的整体预测精度略优于单独的模型 TCN 模型和 VMD-ELM 模型, VMD-TCN 模型预测 $F_{10.7}$ 的均方根误差更低, 相关系数更高, 模型综合预测效果最佳, 这也验证了该模型预测太阳辐射通量的有效性和可靠性。同时, 与前人所建立模型的 $F_{10.7}$ 模型作比较, 发现 TCN 模型的整体预测精度略优于那些广泛使用的自回归(AR)模型[23]和长短期记忆网络[16] (LSTM)模型。特别是对于提前 2、3 天的预测, VMD-TCN 模型略胜一筹。因此, VMD-TCN 模型在 $F_{10.7}$ 的预测中具有重要的意义, 可以帮助更好地理解 and 预测太阳活动的变化。

VMD-TCN 的高精度归功于 VMD 算法有效降低了 $F_{10.7}$ 数据序列的非线性和复杂性, 这有利于 TCN 网络捕捉序列变化的内部规律。尽管 VMD-TCN 方法被证明是预测 $F_{10.7}$ 的可行方法, 但其预测能力仍有进一步提高的空间。未来的工作可以尝试在模型或数据处理中加入物理机制, 并使用更科学的方法来提高模型的泛化能力。

参考文献

- [1] Tapping, K.F. (2013) The 10.7 cm Solar Radio Flux ($F_{10.7}$). *Space Weather*, **11**, 394-406. <https://doi.org/10.1002/swe.20064>
- [2] Ortikov, M.Y., Shemelov, V.A., Shishigin, I.V. and Troitsky, B.V. (2003) Ionospheric Index of Solar Activity Based

- on the Data of Measurements of the Spacecraft Signals Characteristics. *Journal of Atmospheric and Solar-Terrestrial Physics*, **65**, 1425-1430. <https://doi.org/10.1016/j.jastp.2003.09.005>
- [3] Huang, C. and Liu, D. (2009) Forecast Daily Indices of Solar Activity, $F_{10.7}$, Using Support Vector Regression Method. *Research in Astronomy and Astrophysics*, **9**, 694-702. <https://doi.org/10.1088/1674-4527/9/6/008>
- [4] Yaya, P., Hecker, L., Dudok, T., Fèvre, C. and Bruinsma, S. (2017) Solar Radio Proxies for Improved Satellite Orbit Prediction. *Journal of Space Weather and Space Climate*, **7**, Article No. A35. <https://doi.org/10.1051/swsc/2017032>
- [5] Tapping, K.F. and DeTracey, B. (1990) The Origin of the 10.7 cm Flux. *Solar Physics*, **127**, 321-332. <https://doi.org/10.1007/BF00152171>
- [6] Swarup, G., Kakinuma, T., Covington, A.E., Harvey, G.A., Mullaly, R.F. and Rome, J. (1963) High Resolution Studies of Ten Solar Active Regions at Wavelengths of 3 - 21 cm. *Astrophysical Journal*, **137**, Article No. 1251. <https://doi.org/10.1086/147601>
- [7] Mordvinov, A.V. (1986) Prediction of Monthly Indices of Solar Activity $F_{10.7}$ on the Basis of a Multiplicative Autoregression Model. 67-73. <https://ui.adsabs.harvard.edu/abs/1986BSolD1985...67M/abstract>
- [8] Warren, H.P., Emmert, J.T. and Crump, N.A. (2017) Linear Forecasting of the $F_{10.7}$ Proxy for Solar Activity. *Space Weather*, **15**, 1039-1051. <https://doi.org/10.1002/2017SW001637>
- [9] Zhong, Q.Z., Liu, S.Q., He, J.X., et al. (2005) Application of Singular Spectrum Analysis to Solar 10.7 cm Radio Flux 27-day Forecast. *Chinese Journal of Space Science*, **25**, 199-203. <https://doi.org/10.11728/cjss2005.03.199>
- [10] Henney, C.J., Toussaint, W.A., White, S.M. and Arge, C.N. (2012) Forecasting $F_{10.7}$ with Solar Magnetic Flux Transport Modeling. *Space Weather*, **10**, S02011. <https://doi.org/10.1029/2011SW000748>
- [11] Liu, C., Zhao, X., Chen, T. and Li, H. (2018) Predicting Short-Term $F_{10.7}$ with Transport Models. *Astrophysics and Space Science*, **363**, Article No. 266. <https://doi.org/10.1007/s10509-018-3476-x>
- [12] Yeates, A.R., Mackay, D.H. and van Ballegooijen, A.A. (2007) Modelling the Global Solar Corona: Filament Chirality Observations and Surface Simulations. *Solar Physics*, **245**, 87-107. <https://doi.org/10.1007/s11207-007-9013-7>
- [13] Worden, J. and Harvey, J. (2000) An Evolving Synoptic Magnetic Flux map and Implications for the Distribution of Photospheric Magnetic Flux. *Solar Physics*, **195**, 247-268. <https://doi.org/10.1023/A:1005272502885>
- [14] Xiao, C., Cheng, G., Zhang, H., Rong, Z., Shen, C., Zhang, B. and Hu, H. (2017) Using Back Propagation Neural Network Method to Forecast Daily Indices of Solar Activity $F_{10.7}$. *Chinese Journal of Space Science*, **37**, 1-7. <https://doi.org/10.11728/cjss2017.01.001>
- [15] Luo, J., Zhu, H., Jiang, Y., Yang, J. and Huang, Y. (2020). The 10.7-cm Radio Flux Multistep Forecasting Based on Empirical Mode Decomposition and Back Propagation Neural Network. *IEEJ Transactions on Electrical and Electronic Engineering*, **15**, 584-592. <https://doi.org/10.1002/tee.23092>
- [16] Zhang, W., Zhao, X., Feng, X., Liu, C., Xiang, N., Li, Z. and Lu, W. (2022) Predicting the Daily 10.7-cm Solar Radio Flux Using the Long Short-Term Memory Method. *Universe*, **8**, Article No. 30. <https://doi.org/10.3390/universe8010030>
- [17] Lipton, Z.C., Berkowitz, J. and Elkan, C. (2015) A Critical Review of Recurrent Neural Networks for Sequence Learning. <https://arxiv.org/abs/1506.00019>
- [18] Yang, H.J., Sun, Y.Q., Zhu, W., et al. (2021) Prediction Method of Dissolved Gas Concentration in Transformer Oil Based on CEEMD-TCN Model. *Electronic Devices*, **44**, 887-892.
- [19] Tanaka, H., Castelli, J.P., Covington, A.E., et al. (1973) Absolute Calibration of Solar Radio Flux Density in the Microwave Region. *Solar Physics*, **29**, 243-262. <https://doi.org/10.1007/BF00153452>
- [20] Bai, S.J., Kolter, J.Z., Koltun, V., et al. (2018) An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling. <https://arxiv.org/abs/1803.01271>
- [21] Wu, X.W. (2009) Time Series Analysis: Univariate and Multivariate Methods. People's University of China Press, Beijing.
- [22] Zhao, H., Gallo, O., Frosio, I. and Kautz, J. (2017) Loss Functions for Image Restoration with Neural Networks. *IEEE Transactions on Computational Imaging*, **3**, 47-57. <https://doi.org/10.1109/TCI.2016.2644865>
- [23] Du, Z.L. (2022) Forecasting the Daily 10.7 cm Solar Radio Flux Using an Autoregressive Model. *Solar Physics*, **295**, Article No. 125. <https://doi.org/10.1007/s11207-020-01689-x>