

文章编号: 1000-4653(2023)04-0085-08

基于 NeuralProphet-LSTM 组合模型的 港口货物吞吐量预测

杨宇鸽, 郝杨杨, 王逸文

(上海海事大学 物流研究中心, 上海 201306)

摘要: 为进一步提高货物吞吐量预测准确性, 提出基于 NeuralProphet 时间序列模型与长短期记忆 (LSTM) 神经网络的组合预测模型。首先利用 NeuralProphet 模型对港口货物吞吐量数据进行训练得到预测值并计算残差序列, 然后对残差数据建立 LSTM 神经网络模型进行预报修正, 重构得到最终的预测值。以上海港、厦门港的月度货物吞吐量数据为样本展开试验, 结果表明, 该模型能够有效地解决数据异常波动造成的预测结果误差大、预测效果不稳定等问题; 相比于传统单一模型与 LSTM-支持向量机 (SVM)、Bi-LSTM 等组合模型, NeuralProphet-LSTM 模型预测精度更高, 可帮助港航企业及时调整规划决策与经营策略。

关键词: 港口吞吐量; 组合模型; 预测; NeuralProphet

中图分类号: U695.2

文献标志码: A

DOI: 10.3969/j.issn.1000-4653.2023.04.012

Prediction of port cargo throughput using NeuralProphet-LSTM combination model

YANG Yuge, HAO Yangyang, WANG Yiwen

(Logistics Research Center, Shanghai Maritime University, Shanghai 201306, China)

Abstract: Conventional method has not been successful in predicting cargo throughput of ports due to the great fluctuation caused by complex external factors. In order to improve the accuracy of cargo throughput prediction for ports, the combination of NeuralProphet time series model and LSMT neural network is introduced. NeuralProphet model is trained to produce, according to input cargo throughput data, a preliminary prediction with a residual sequence. The LSMT model is used to correct the preliminary prediction and output final prediction. The combination model is verified through process the monthly throughput data from Shanghai Port and Xiamen Port. The advantage of the combination model in handling abnormal fluctuation over other configurations is demonstrated.

Key words: port throughput; combination model; prediction; NeuralProphet

港口吞吐量是反映港口生产经营活动成果的重要量化指标,也是衡量地区和城市建设发展、经济水平的量化参考依据。但是港口货物吞吐量具有随机性、波动性、非线性等特点,不易预测准确。尤其是“十三五”以来,产业结构面临调整,国际国内经济环境复杂,以及2018年开始的中美贸易摩擦和2020年全球新冠肺炎病毒感染疫情等持续冲击,使

得港口吞吐量走势存在较多变数。因此,面对港口货物吞吐量的不稳定波动,对吞吐量的月度数据预测变得更为重要,科学准确地预测能够及时帮助港航企业调整经营策略或帮助政府合理制定港口发展计划等。

目前,已有不少学者提出港口货物吞吐量预测的模型。常用的预测模型有灰色预测模型、回归模

收稿日期: 2022-05-10

基金项目: 上海市科技计划项目 (23692106900)

作者简介: 杨宇鸽 (1999—), 女, 硕士研究生, 研究方向为港口物流。E-mail: 1124313129@qq.com

通信作者: 郝杨杨 (1984—), 男, 高级工程师, 硕士生导师, 研究方向为港航信息化与仿真优化。E-mail: yyhao@shmtu.edu.cn

引用格式: 杨宇鸽, 郝杨杨, 王逸文. 基于 NeuralProphet-LSTM 组合模型的港口货物吞吐量预测 [J]. 中国航海, 2023, 46(4): 85-92.

YANG Y G, HAO Y Y, WANG Y W. Prediction of port cargo throughput using NeuralProphet-LSTM combination model [J]. Navigation of China, 2023, 46(4): 85-92. (in Chinese)

型、时间序列模型等^[1]。如杨金花^[2]使用灰色模型 GM(1,1) 对上海港集装箱吞吐量进行预测;吴琛^[3]使用基于 ARIMA 时间序列模型用于广州港集装箱吞吐量预测。虽这些模型取得一定的预测效果,但灰色预测模型仅需少量数据即可进行预测,多应用于货物吞吐量年度数据的预测,而传统的时间序列模型对于突变波动的预测效果较差,在非线性数据的处理上有所不足。

随着深度学习的不断发展,各类神经网络被运用于港口吞吐量的预测中,包括反向传播(Back Propagation, BP)神经网络、Elman 神经网络、LSTM 神经网络等。刘枚莲等^[4]建立了港口吞吐量预测的 BP 神经网络预测模型,对广西北部湾港口吞吐量进行预测,结果表明 BP 神经网络优于传统回归模型;李广儒等^[5]考虑到 Elman 神经网络具有处理动态信息能力,将其应用于宁波舟山港的货物吞吐量预测,结果证实 Elman 神经网络能较好地适应吞吐量数据随时间变化的特性;王凤武等^[6]构造 LSTM 神经网络的预测模型,并用上海港集装箱吞吐量实际数据验证了其有效性。这些算法对非线性关系预测效果较好,但对网络的初始阈值和权值有较强依赖,容易陷入局部最优解。因此,诸多学者开始提出组合预测方法,通过组合模型能产生具有无偏性的预测,可减少由参数或模型错误识别带来的预测误差,如王振振等^[7]在港口集装箱吞吐量预测中,基于三次指数平滑法得出初始预测值,再用马尔科夫原理得到状态转移矩阵,修正初始预测值,以此提高预测精度;张树奎等^[8]结合灰色模型所需初始数据少和 BP 神经网络非线性拟合能力强的特点,提出灰色神经网络组合模型。这些模型均取得了较好的预测效果,证实了组合模型的有效性,但近年来外部环境变化复杂,吞吐量波动剧烈,传统组合方法未必能够取得最优预测精度,一些学者采用先分解、后集成的方式来提高预测精度,如张丰婷等^[9]基于变分模态分解将吞吐量分解成多个子模态,再分别进行预测和集成,顾子瑜等^[10]提出的 ICEEMDAN-SE-ARIMA&ICAKELM-IKM 预测模型,结果均表明分解集成模型能显著降低原始序列中的复杂性;XIE 等^[11]将数据分解后,通过三种最小二乘支持向量回归模型的预测混合方法应对吞吐量序列的季节性和非线性特点,这类方法确实有助于解决吞吐量中的复杂因素带来的影响,但是计算过程极为复杂、对历史数据要求较高。

对于此问题,本文引入了一种较新颖的时间序列预测模型 NeuralProphet^[12],该模型可以直接分解

出数据中的趋势项、季节性,以及添加节日、特殊事件等外部因素并进行预测,对于少量数据也能有较好的学习效果。此外,考虑到 NeuralProphet 模型终究还是传统的时间序列分析建模思路,高效但是准确性还有待提高,为了进一步捕捉非线性部分的规律,提高模型在港口货物吞吐量预测领域的准确性,结合 LSTM 神经网络对处理非线性部分的优势进行残差修正,增强模型的预测效果。

1 模型分析与设计

1.1 NeuralProphet 模型

NeuralProphet 是一个基于 PyTorch 实现的用户友好型时间序列预测工具。NeuralProphet 模型的核心概念是将其模块化可组合性。它将时间序列分解成六个模块,通过分别拟合这几项,累加起来得到时间序列的预测值 \tilde{y}_t ,基本形式如下:

$$\tilde{y}_t = T(t) + S(t) + E(t) + F(t) + A(t) + L(t) \quad (1)$$

式(1)中, $T(t)$ 是趋势模块。NeuralProphet 采用经典方式对趋势进行建模,表达式如下:

$$T(t_1) = T(t_0) + K \cdot \Delta t = m + K \cdot (t_1 - t_0) \quad (2)$$

其中, K 表示增长率, Δt 表示 t_0 到 t_1 的时间变化, m 代表偏移量。

式(1)中, $S(t)$ 是季节模块。它用傅里叶级数近似表达周期性分量,能够处理高频率数据的多季节性,具体表达式如下:

$$S_p(t) = \sum_{j=1}^k \left(a_j \cdot \cos\left(\frac{2\pi jt}{p}\right) + b_j \cdot \sin\left(\frac{2\pi jt}{p}\right) \right) \quad (3)$$

其中, k 表示傅里叶级数的数量, p 表示固定的周期, a_j, b_j 是被估参数。

式(1)中, $E(t)$ 是事件和节日模块。特殊事件或假日的影响偶尔发生,时间步长 t 上的所有事件对预测值的影响可用以下方程式表示:

$$E(t) = \sum_{e \in E} E_e^*(t) \quad (4)$$

其中, e 表示每个事件都是一个二元变量。 $e \in \{0, 1\}$

式(1)中, $F(t)$ 是未来回归模块,是指在预测期有已知未来值的外部变量, f 表示可选择加法或乘法模式,所有未来回归系数的影响可以用以下方程式表示:

$$F(t) = \sum_{f \in F} F_f^*(t) \quad (5)$$

式(1)中, $A(t)$ 是自回归模块,是指将变量的未来值与其过去值进行回归的过程,自回归的方程表达式如下:

$$A'(t), A'(t+1), \dots, A'(t+h-1) = AR-Net(y_{t-1}, y_{t-2}, \dots, y_{t-p}) \quad (6)$$

其中, $AR-Net$ 可解释为是一个用于时间序列的自动回归前馈神经网络,输入 p 个历史协变量,经过 $AR-Net$ 网络获得 h 步输出,有助于提高拟合的速度与准确性。

式(1)中, $L(t)$ 是滞后回归模块,表达式如下:

$$L'_x(t), L'_x(t+1), \dots, L'_x(t+h-1) = AR-Net(x_{t-1}, x_{t-2}, \dots, x_{t-p}) \quad (7)$$

可以看出,滞后回归器模块在功能上与自回归模块相同,区别在于协变量的最后观测值作为模块的输入。

1.2 LSTM 神经网络模型

除 NeuralProphet 模型能捕捉到的数据关系外,港口货物吞吐量数据还包含了随机非线性部分,这里考虑采用长短时记忆网络(Long Short-Term Memory Network, LSTM)模型来刻画其中的非线性关系。LSTM 是在循环神经网络(Recurrent Neural Network, RNN)的基础上进行的进一步改进。在 RNN 的基础上,LSTM 引入了被称为“控制门”的内部机制,它可以起到调节信息流的作用,其中包括 3 个门的结构:输入门、遗忘门和输出门。这使得 LSTM 具备能选择保留有用信息或遗忘不重要信息的功能,解决 RNN 无法解决的长距离依赖的问题,实现长期记忆。LSTM 的神经元结构如图 1 所示。

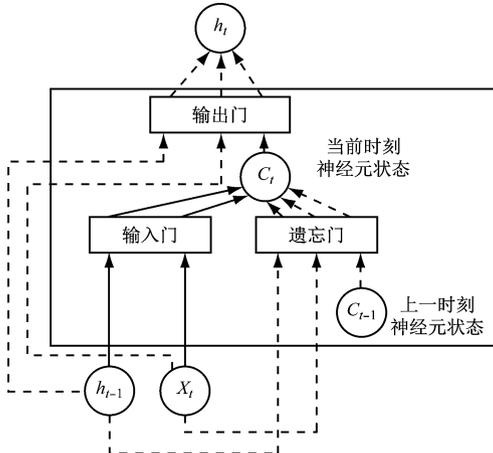


图 1 LSTM 的神经元结构

Fig. 1 Neuron structure of LSTM

X_t 为当前时刻的输入; h_{t-1} 和 h_t 分别表示上一时刻和当前时刻的隐藏层状态; C_{t-1} 和 C_t 分别为上一时刻和当前时刻的神经元状态。LSTM 进行状态

更新和信息输出的主要过程包括以下 3 个部分。

1) 遗忘门

输入上一时刻的状态信息 h_{t-1} 和当前时刻的特征信息 X_t , 两者进行连接,与遗忘门权重矩阵 W_f 相乘,加上偏置项 b_f ,传递到 sigmoid 激活函数中。这个值决定了上一刻神经元状态 C_{t-1} 的值是被记忆还是被遗忘。越接近 0 意味着越应该丢弃,越接近 1 意味着越应该保留。

$$f_t = \sigma(W_f * [h_{t-1}, X_t] + b_f) \quad (8)$$

2) 输入门

i_t 同样具有遗忘功能,表示要更新的数值。通过 tanh 激活函数,创造一个新的候选态向量 \tilde{C}_t 。 i_t 与 \tilde{C}_t 将决定哪些信息是重要且需要保留下来的。最后,需要遗忘的信息与新的要记忆的部分相加,就是要输入下一个节点的神经元状态 C_t 。

$$i_t = \sigma(W_i * [h_{t-1}, X_t] + b_i) \quad (9)$$

$$\tilde{C}_t = \tanh(W_c * [h_{t-1}, X_t] + b_c) \quad (10)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (11)$$

3) 输出阶段

O_t 表示输出控制量,用 tanh 函数对 C_t 进行处理,以确定当前时刻的隐藏状态应携带的信息 h_t

$$O_t = \sigma(W_o * [h_{t-1}, X_t] + b_o) \quad (12)$$

$$h_t = O_t \odot \tanh(C_t) \quad (13)$$

式(8)~式(13)中, W_i 、 W_c 、 W_o 分别是输入门、单元状态、输出门的权重矩阵; b_i 、 b_c 、 b_o 分别为输入门、单元状态、输出门的偏置项; $\sigma(\cdot)$ 是激活函数 sigmoid,其输出值在 $[0, 1]$ 之间; $\tanh(\cdot)$ 是双曲正切激活函数,其输出值在 $[-1, 1]$ 之间; \odot 表示矩阵元素对应相乘的运算; $*$ 表示矩阵的乘法运算。

1.3 NeuralProphet-LSTM 组合预测模型

时间序列数据中蕴含着大量不确定的信息,各种单一预测模型在反映数据信息时存在着一些差异,都有其模型自身的适用条件及局限性,使用单一模型对数据序列进行预测时将不可避免地丢失一些数据信息^[13]。如前文所说,为了能充分捕捉数据中非线性的部分,本文将 NeuralProphet 和 LSTM 两种模型进行组合,提出 NeuralProphet-LSTM 组合模型。在组合模型构建方法的选择上,常见的组合方法有误差倒数法、最小方差法、权重收缩法、优势矩阵法等^[14]。NeuralProphet 模型和 LSTM 模型都在一定程度上能够处理非线性问题,给两者赋权预测意义不大。NeuralProphet 模型的意义在于可以分解复杂模型,并且在外部添加一些重大事件、各种突变状况对吞吐量带来的影响;而残差中的随机非线性部分可

以充分利用 LSTM 神经网络的优势进行补充,因此用了误差值纠正的组合方式,来进一步提高预测的准确性。组合模型预测见图 2。

首先利用 NeuralProphet 算法模型对港口货物吞吐量数据进行训练得到预测值并计算残差序列,然后对残差数据建立 LSTM 神经网络模型进行预报修正,最后重构得到预测值。具体步骤为:

1) 按照时间序列格式对吞吐量原始数据进行处理;

2) 外部输入节假日、突发事件影响,进行参数调优后,利用 NeuralProphet 模型对 t 时刻对应的港口货物吞吐量数据 T_t 进行训练和预测,得到相应的预测数据 L_t ;

3) 将真实数据 T_t 与预测数据 L_t 两者作差,得到误差数据 N_t ,即 $N_t = T_t - L_t$;

4) 用 LSTM 模型对误差进行预测及参数调优,纠正原有误差值,得到新的预测误差;

5) 将 NeuralProphet 模型预测结果 L_t 和误差预测结果 \bar{N}_t 进行合并,最终得到基于 NeuralProphet-LSTM 组合模型的预测结果,即 $\bar{T}_t = L_t + \bar{N}_t$ 。

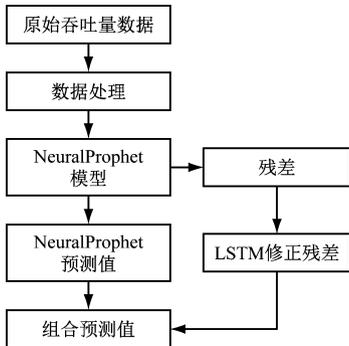


图 2 组合模型预测图

Fig. 2 Combined model prediction graph

2 试验设计

2.1 数据来源及说明

港口货物吞吐量数据具有周期性,且易受外界复杂因素影响。虽不同港口的吞吐量数据有差异,不同港口应对突发事件的能力也有所不同,但吞吐量的影响因素相同,因此总体变化趋势具有一致性。为证明所提出 NeuralProphet-LSTM 模型的普适性和鲁棒性,选择了吞吐量较大的上海港和吞吐量相对较小的厦门港为试验对象进行分析,验证模型处理复杂非线性关系的能力。

根据数据的可获得性,同时考虑近年港口吞吐量发展瓶颈及外界环境影响,分别从上海统计局和厦门统计局选取了上海港和厦门港 2014 年 1 月—

2022 年 3 月期间共 99 个月的货物吞吐量数据作为实例数据,原始时间序列如图 3 和图 4 所示。

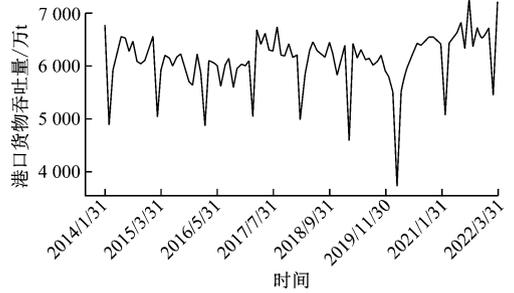


图 3 上海港 2014 年 1 月—2022 年 3 月月度货物吞吐量
Fig. 3 Monthly cargo throughput of Shanghai Port

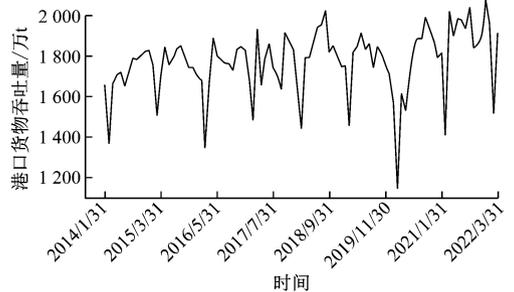


图 4 厦门港 2014 年 1 月—2022 年 3 月月度货物吞吐量
Fig. 4 Monthly cargo throughput of Xiamen Port

从图 3 和图 4 可以看出,近年来上海港和厦门港的港口货物吞吐量在总体上趋势一致,小浮动上升,受到不同尺度事件影响,如中美贸易摩擦^[15]、新冠肺炎病毒感染疫情等,呈现出复杂多变的非线性特征。另外,在每年的 2 月会出现明显的下降,该时段是我国的农历新年,进出口贸易受到节假日的影响。

2.2 NeuralProphet 时间序列模型拟合试验

NeuralProphet 模型进行拟合时所需要的两列数据分别是“ds”和“y”,其中“ds”表示时间戳,“y”表示时间序列的值,因此,在处理数据时,需要修改数据的列名,另外按照此前观察到的吞吐量在每年 2 月下降明显的规律,设置外部事件,命名为“chunjie”以及设置 2020 年新冠肺炎病毒感染疫情突发事件,命名为“epidemic”,进而进行模型拟合。

图 5 和图 6 分别是模型对上海港货物吞吐量数据及厦门港货物吞吐量数据成分的单分析。从上至下分别是 2014—2022 年港口货物吞吐量数据被 NeuralProphet 加法模型分解出的增长趋势、突变点的变化率、年周期趋势和外部事件,可以看出:近年来,港口货物吞吐量到达了一定程度上的发展“瓶颈期”,无论是上海港还是厦门港都面临发展增速放缓,尤其新冠肺炎病毒感染疫情暴发,更是极大程度上改变了吞吐量的走向,此外,从周期趋势中可以

看出港口货物吞吐量存在着明显的周期性变化。

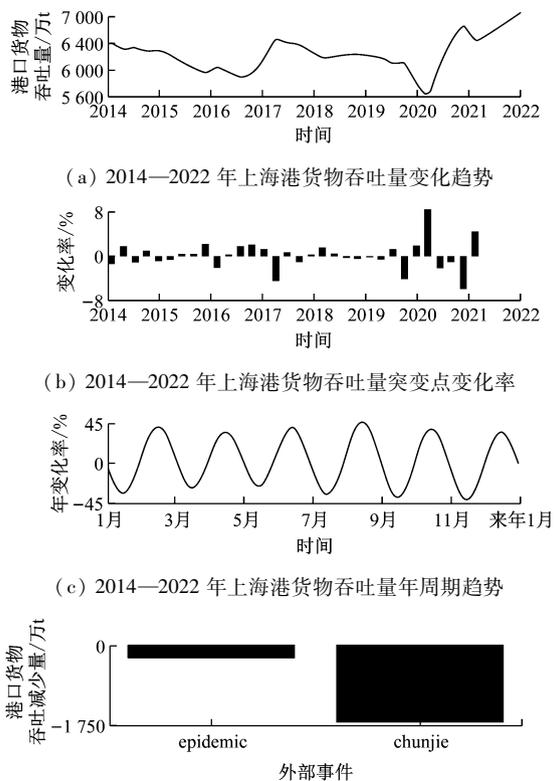


图 5 上海港成分分析图

Fig. 5 Component analysis diagram of Shanghai Port

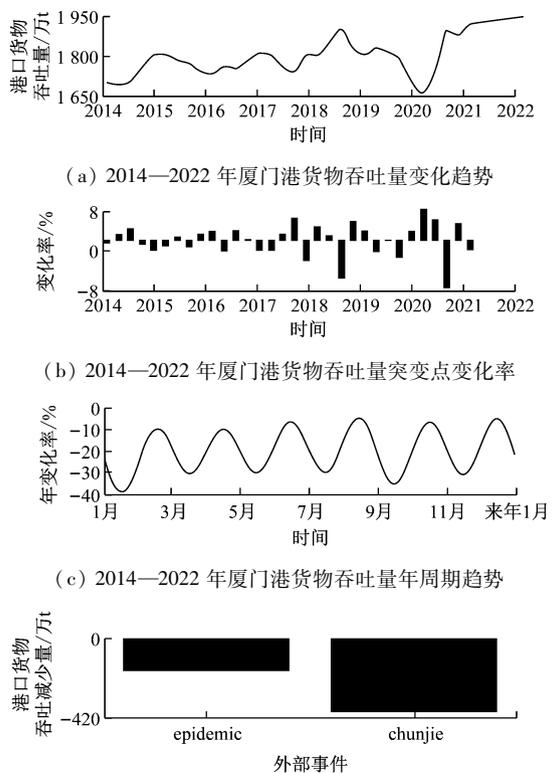


图 6 厦门港成分分析图

Fig. 6 Component analysis diagram of Xiamen Port

图 7 显示的是 NeuralProphet 模型在上海港数据集上的拟合效果,图 8 则为厦门港数据拟合效果。从图 7 和图 8 可以看到,整体走势是基本符合的,既符合季节性,又符合线性趋势的缓慢增长,并且在外部干预下也能成功拟合突变点对数据的影响。这说明 NeuralProphet 模型的拟合度较好,可以用于港口货物吞吐量的预测,但是存在不少该模型无法捕获到的无规律离散部分以及一些与真实值的偏差数据,因此通过残差修正来提高预测精度。

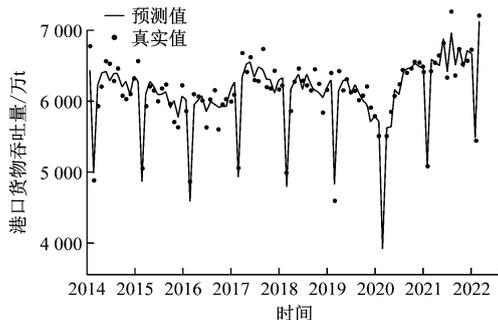


图 7 上海港数据拟合效果

Fig. 7 Data fitting effect of Shanghai Port

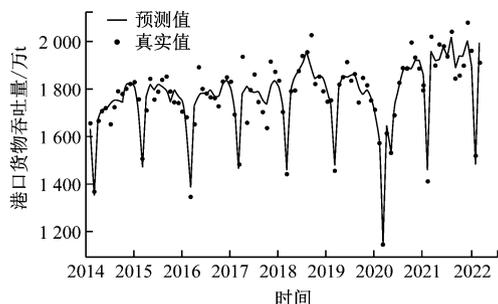


图 8 厦门港数据拟合效果

Fig. 8 Data fitting effect of Xiamen Port

2.3 组合模型预测试验

基于上述提出的 NeuralProphet-LSTM 预测模型,将数据集分为两类,其中 2014 年 1 月—2020 年 9 月期间 81 个月的数据作为训练集用于训练模型,2020 年 10 月—2022 年 3 月期间 18 个月的数据作为测试集对模型拟合效果进行验证。在 Python3.7 软件的基础上,经过多次试验,将具体参数设置为如表 1 所示。

图 9 和图 10 分别为上海港和厦门港通过 NeuralProphet-LSTM 组合模型吞吐量的预测值与真实值的对比。从图中可以看出,针对不同港口,模型预测的结果与吞吐量真实值都有较高吻合度,即使经历了 2020 年初新冠肺炎病毒感染疫情的重大打击,也能通过历史数据以及 2020 年后半年的回升趋势对之后的吞吐量数据做出较为准确的预测,证实该模型对港口货物吞吐量预测具有普适性与有效性。

表1 参数设置
Tab.1 Parameter setting

模型名称	参数名	取值
NeuralProphet	增长趋势类型	linear
NeuralProphet	突变点数量/个	31
NeuralProphet	估计趋势变化点的历史比例	0.9
NeuralProphet	季节性模式	加法模式
LSTM	输入层神经元数量/个	125
LSTM	隐藏层数量/层	2
LSTM	隐藏层神经元数量/个	15
LSTM	输出层神经元数量/个	1
LSTM	迭代数量/次	100
LSTM	一次训练所选取的样本数/个	6

3 结果分析

3.1 评价指标选取

本文采用了均方根误差(Root Mean Squared Error, R_{MSE})、平均绝对误差(Mean Absolute Error, M_{AE})、平均绝对百分比误差(Mean Absolute Percentage Error, M_{APE})这三个指标作为检验模型预测结果的标准。其计算分别如下所示:

$$R_{MSE} = \sqrt{\frac{1}{N} \sum_{n=1}^N (y_n - \hat{y}_n)^2} \quad (14)$$

$$M_{AE} = \frac{1}{N} \sum_{n=1}^N |y_n - \hat{y}_n| \quad (15)$$

$$M_{APE} = \frac{100}{N} \sum_{n=1}^N \left| \frac{y_n - \hat{y}_n}{\hat{y}_n} \right| \quad (16)$$

式中, y_n 为吞吐量的真实值, \hat{y}_n 为吞吐量的预测值, 取值越小, 说明模型的预测性能就越好。

3.2 对比分析

为进一步判断 NeuralProphet-LSTM 模型的性能, 先与单一模型(支持向量机(Support Vector Machine, SVM)、单一 NeuralProphet 模型、LSTM 神经网络模型)进行对比。将这些模型分别对 2014 年 1 月至 2020 年 9 月的数据进行训练, 对 2020 年 10 月至 2022 年 3 月的数据进行预测, 并计算各个模型的 R_{MAE} 、 M_{AE} 和 M_{APE} 值, 计算结果见表 2。从表 2 可以看出, NeuralProphet-LSTM 模型的预测效果显然优于单一模型, 同时证实了模型通过组合后在原有的基础上提高了预测精度。

除了对比单一模型外, 本文还将 NeuralProphet-LSTM 组合模型与 GM(1,1)-BP 模型、LSTM-SVM 模型^[16]、Bi-LSTM 模型^[17]的预测性能进行对比, 评价指标的计算结果见表 3。从表 3 可知, 各类组合模型在预测港口货物吞吐量方面, 都实现了不同程度的预测效果, 但是可能受到新冠肺炎病毒感染疫情对吞吐量的波动影响导致之后的预测值出现较大偏差, 效果不如 NeuralProphet-LSTM 组合模型。

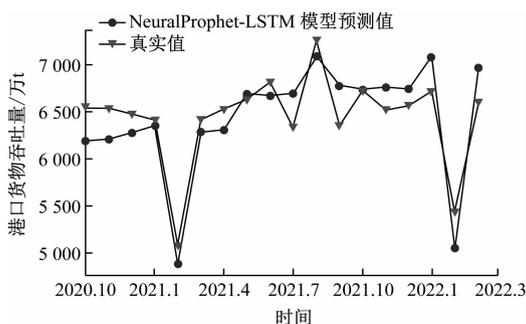


图9 上海港 NeuralProphet-LSTM 模型预测值与真实值对比

Fig.9 Comparison between the predicted value and the actual value of NeuralProphet-LSTM model in Shanghai Port

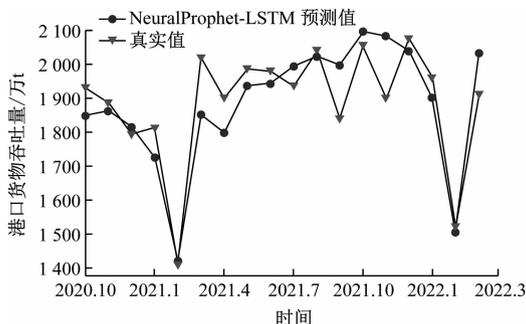


图10 厦门港 NeuralProphet-LSTM 模型预测值与真实值对比

Fig.10 Comparison between the predicted value and the actual value of NeuralProphet-LSTM model in Xiamen Port

表2 与单一模型评价指标对比

Tab.2 Compared with the evaluation indicators of a single mode

模型	RMSE		MAE		MAPE/%	
	上海港	厦门港	上海港	厦门港	上海港	厦门港
SVM	558.93	194.42	449.33	168.89	7.89	9.19
NeuralProphet	319.03	125.83	252.92	89.97	3.70	4.50
LSTM	453.82	191.15	367.18	138.44	6.53	6.79
NeuralProphet-LSTM	170.70	78.63	145.15	62.05	2.27	3.33

表 3 与组合模型评价指标对比

Tab.3 Compares the evaluation indexes with those of the combined model

模型	RMSE		MAE		MAPE/%	
	上海港	厦门港	上海港	厦门港	上海港	厦门港
GM(1,1)-BP	466.61	173.35	310.33	117.94	5.18	6.74
Bi-LSTM	490.47	178.24	400.38	132.76	6.43	7.39
LSTM-SVM	313.45	104.73	216.84	75.88	3.60	4.37
NeuralProphet-LSTM	170.70	78.63	145.15	62.05	2.27	3.33

综上所述,无论是单一模型还是组合模型,NeuralProphet-LSTM 模型都有着更好的预测效果,进一步证实了该模型处理变化复杂数据的有效性。

4 结 论

近年来的港口货物吞吐量呈现较为复杂的趋势,针对其非线性、非平稳的特点,本文提出一种基于 NeuralProphet-LSTM 的组合预测模型,利用 NeuralProphet 模型模块化特点,分解时间序列中的复杂数据结构,并进行有效预测,同时结合 LSTM 处理非线性问题优势提高预测精度,得出研究结论如下。

1) 本文所提出的 NeuralProphet-LSTM 模型可用于处理不规则数据,能够有效解决数据异常波动造成的预测结果误差大、预测效果不稳定等问题。面对受新冠肺炎病毒感染疫情、中美贸易摩擦等外部事件冲击下依旧能做出有效预测。另外,适用于不同港口,具有普适性。

2) 相较于传统的 SVM 模型、NeuralProphet、LSTM 神经网络,通过残差修正方法构建的 NeuralProphet-LSTM 组合模型在预测精度上得到了提高,相较于 GM(1,1)-BP、LSTM-SVM、Bi-LSTM 等组合模型,NeuralProphet-LSTM 模型更能充分提取吞吐量中的复杂信息。该模型不仅可以通过内置的加法模型分解组成成分,直观再现吞吐量数据中不同时间尺度的波动特点,并且能够高效拟合数据中的异常波动并自行做出有效预测,比人为的分解-集成更为高效。

未来的研究中,由于人工难以寻找最优参数组合,因此需关注 LSTM 神经网络的权值和阈值,进一步优化提高模型准确性。同时,为了更深入地分析港口货物吞吐量时间序列数据,可有针对性地纳入港口吞吐量的影响因素做进一步的研究。

参 考 文 献

[1] 王苗,高璐. 港口集装箱吞吐量预测方法研究[J]. 上海工程技术大学学报,2020,34(2):201-206.
WANG Z, GAO L. Research on forecasting method of container throughput[J]. Journal of Shanghai University of Engineering Science, 2020, 34(2): 201-206. (in Chinese)

[2] 杨金花,杨艺. 基于灰色模型的上海港集装箱吞吐量预测[J]. 上海海事大学学报,2014,35(2):28-32.
YANG J H, YANG Y. Forecast of Shanghai Port container throughput based on grey model[J]. Journal of Shanghai Maritime University, 2014, 35(2): 28-32. (in Chinese)

[3] 吴琛. 基于时间序列模型的港口集装箱吞吐量预测[J]. 珠江水运,2019(5):73-74.
WU C. Based on the port container throughput forecast of time series model[J]. Transportation Research, 2019(5): 73-74. (in Chinese)

[4] 刘枚莲,朱美华. 基于 BP 神经网络的港口吞吐量预测模型[J]. 系统科学学报,2012,20(4):88-91.
LIU M L, ZHU M H. Port throughput prediction model based on BP neural network[J]. Journal of systems science,2012,20(4):88-91. (in Chinese)

[5] 李广儒,朱庆辉. 基于 Elman 神经网络的港口货物吞吐量预测[J]. 重庆交通大学学报(自然科学版),2020,39(6):8-12.
LI G R, ZHU Q H. Prediction of port cargo throughput based on Elman neural network[J]. Journal of Chongqing Jiaotong University (Natural Science Edition), 2020, 39(6):8-12. (in Chinese)

[6] 王凤武,张晓博,阎际驰,等. 基于 LSTM 的上海港集装箱吞吐量预测[J]. 中国航海,2022,45(2):109-114.
WANG F W, ZHANG X B, YAN J C, et al. Container throughput forecast of Shanghai Port based on LSTM[J]. Navigation of China,2022,45(2):109-114.

[7] 王振振,裴道方,朱宗良,罗天. 基于 ES-Markov 模型的港口集装箱季度吞吐量分析与预测[J]. 中国航海,2019,42(4):125-130.
WANG Z Z, CHANG D F, ZHU Z L, et al. Analysis and prediction of quarterly container throughput based on ES-Markov model[J]. China navigation,2019,42(4): 125-130. (in Chinese)

[8] 张树奎,肖英杰,鲁子爱. 基于灰色神经网络的港口集装箱吞吐量预测模型研究[J]. 重庆交通大学学报(自然科学版),2015,34(5):135-138.
ZHANG S K, XIAO Y J, LU Z A. Research on prediction model of port container throughput based on grey neural network[J]. Journal of Chongqing Jiaotong University (Natural Science Edition), 2015, 34(5): 135-138. (in Chinese)

