

DOI: 10.19783/j.cnki.pspc.210573

关键信息缺失下基于相空间重构及机器学习的电力负荷预测

侯慧¹, 王晴¹, 赵波², 章雷其², 吴细秀¹, 谢长君¹

(1. 武汉理工大学自动化学院, 湖北 武汉 430070; 2. 国网浙江省电力有限公司电力科学研究院, 浙江 杭州 310014)

摘要: 随着碳交易系统的发展, 准确预测电力能源消耗对于能源管理是至关重要的。为实现在缺失天气等多种关键信息下的电力负荷预测, 首先采用混沌理论中的相空间重构技术对历史负荷时间序列进行处理, 根据排列熵验证混沌特性。并利用 8 种机器学习模型进行预测与比较, 其中包括 4 种以神经网络为基础的机器学习模型、3 种以统计学习为基础的机器学习模型及 1 种基准模型。其次采用灰色关联度法对预测精度较高的极限学习机(ELM)和极端梯度提升(XGBoost)进行组合, 构建了 ELM-XGBoost 模型。最后将 ELM-XGBoost 应用于一日至一周内不同时间尺度的负荷预测。结果表明, 预测精度随预测时间尺度增加而呈现降低的趋势, 且在日负荷预测中, 所构建的 ELM-XGBoost 模型预测精度得到提升, 应用效果良好。

关键词: 电力负荷预测; 关键信息; 极限学习机; 极端梯度提升; 相空间重构; 排列熵

Power load forecasting without key information based on phase space reconstruction and machine learning

HOU Hui¹, WANG Qing¹, ZHAO Bo², ZHANG Lei², WU Xixiu¹, XIE Changjun¹

(1. School of Automation, Wuhan University of Technology, Wuhan 430070, China;

2. State Grid Zhejiang Electric Power Co., Ltd. Research Institute, Hangzhou 310014, China)

Abstract: With the development of the carbon trading system, accurate forecasting of power consumption is crucial for energy management. For power load forecasting without key information such as weather information, phase space reconstruction technique of chaos theory is adopted first to deal with historical load time series. Permutation entropy is used to validate the chaotic characteristic. 8 kinds of machine learning models are used to forecast and make comparisons. These models are: 4 kinds of neural network, 3 kinds of statistical machine learning and 1 kind of benchmark. Secondly, two optimal models, extreme learning machine (ELM) and extreme gradient boosting (XGBoost), are combined by a grey relational degree method to construct an ELM-XGBoost model. Finally, ELM-XGBoost is applied to forecast with different time scales from one day to one week. Results show that forecasting accuracy decreases with the increase of forecasting time scale. In daily load forecasting, the accuracy of ELM-XGBoost model is improved with a better application effect.

This work is supported by the National Key Research and Development Program of China (No. 2020YFB1506802).

Key words: power load forecasting; key information; extreme learning machine; extreme gradient boosting; phase space reconstruction; permutation entropy

0 引言

准确的电力负荷预测对电网稳定运行有着重要作用。电力负荷大小受多种关键信息影响: 温度、湿度、风速、日期类型、地区 GDP 等。然而, 通常

情况下难以获得所有关键信息的准确数据。因此, 若能在关键信息缺失情况下实施较为准确的电力负荷预测, 则可以省去收集和处理多种数据的繁琐过程, 简化电力负荷预测步骤。

现有短期电力负荷预测的研究主要分为多元时间序列分析和单一时间序列分析。多元时间序列分析考虑天气等多种关键信息数据。如文献[1]选取最大负荷、平均负荷、用电量、负荷密度和土地信息

基金项目: 国家重点研发计划项目资助(2020YFB1506802); 国家自然科学基金项目资助(52177110)

等负荷属性作为多源影响因素,建立了空间负荷预测模型。文献[2]将负荷历史数据、温度数据、日期数据和节假日信息分别输入长短期记忆网络(Long Short Term Memory, LSTM)和轻梯度提升机(Light Gradient Boosting Machine, LightGBM),通过最优加权组合法得出组合模型的预测值。文献[3]将气温、年份、月份、日期和工作日信息等多种关键信息输入梯度提升树,基于预测模型识别重要影响因素。为提高预测精度,文献[4]将天气预报的误差作为负荷预测的特征之一。然而,多种关键信息的输入在提高预测精度的同时会带来繁琐的数据处理及特征处理过程。数据处理包括数据清洗、集成及变换等步骤。特征提取可通过经验模态分解^[5]、卷积神经网络^[6]、关联矩阵^[7]、聚类分析^[8]等方法实现。将所选定特征输入到机器学习模型或深度学习模型^[9-11]中,从而获得预测结果。综上可知,涉及多种关键信息下的负荷预测具有较复杂的数据处理及特征处理过程,且上述研究对天气等关键信息数据依赖性较强,若不能获取准确的关键信息数据,形成多元时间序列,则较难开展精确预测。

在实际情况下,部分研究难以获得与某地区电力负荷历史数据相匹配的关键信息数据,如负荷密度、土地信息、温度信息等,或所获取关键信息数据类型有限,限制了预测精度的提升。因此,为实现在缺失天气等多种关键信息下的电力负荷预测及简化预测步骤,有必要对单一时间序列的负荷预测进行研究。针对单一时间序列下的研究,文献[12-14]指出,一般电网的电力负荷时间序列具有混沌特性,利用负荷的混沌特性进行负荷预测,可以不直接考虑影响负荷的因素,而是对受诸多关键信息影响的负荷历史数据进行分析。据此,文献[15]提出基于相空间重构理论和神经网络融合的改进混沌神经网络,通过对天津配电网和地区网的历史负荷进行训练和测试,证实了所提方法的有效性。文献[16]提出基于相空间重构理论和递归神经网络相结合的预测模型。文献[17-18]采用相空间重构网络对时间序列进行预测。虽然上述研究不涉及影响负荷的关键信息数据,仅使用负荷历史数据完成预测,但其所采用的预测模型以神经网络为主,模型较为单一。现有研究中对于基于单一时间序列下的机器学习算法对比研究较少,且预测精度有待提高。

综上,本文首先采用混沌理论中的相空间重构技术与机器学习模型相结合的方法,通过对比8种机器学习模型的预测精度,寻找单一时间序列下合适的电力负荷预测模型。其次,对于较优的模型,在不同时间长度下进一步验证方法的正确性。最后

采用灰色关联度法进行模型组合,提出了由ELM和XGBoost组合而成的ELM-XGBoost模型,算例结果表明,在日负荷预测中,本文组合模型能够在缺失关键信息下取得较其他组合模型更优的预测效果。

1 理论介绍

1.1 相空间重构原理

相空间重构理论(Phase Space Reconstruction Theory, PSRT)认为混沌系统中各分量的演化都是由与之相互作用的其他分量(如气象因素等分量)所决定,即每个分量的演化过程中都隐含着系统的全部信息^[19]。

设混沌时间序列为 $x_1, x_2, \dots, x_{N-1}, x_N$, 对 $x_i (i = 1, 2, \dots, N - (m-1)\tau)$, 作如下变换^[20]。

$$x_i = (x_i, x_{i+\tau}, x_{i+2\tau}, \dots, x_{i+(m-1)\tau})^T \quad (1)$$

式中: τ 为延迟时间; m 为嵌入维数。

根据相空间重构法,将混沌时间序列 $x_1, x_2, \dots, x_{N-1}, x_N$ 转换成延时为 τ 、维数为 m 的新数据空间^[21],即

$$\mathbf{X} = \begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_{N-(m-1)\tau} \\ x_{1+\tau} & x_{2+\tau} & x_{3+\tau} & \cdots & x_{N-(m-2)\tau} \\ x_{1+2\tau} & x_{2+2\tau} & x_{3+2\tau} & \cdots & x_{N-(m-3)\tau} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{1+(m-1)\tau} & x_{2+(m-1)\tau} & x_{3+(m-1)\tau} & \cdots & x_N \end{bmatrix} \quad (2)$$

式中,每一列表示一个向量或相点。另外,从式(1)可以看出,对于 N 个数据点,赋予延迟时间 τ ,嵌入维数 m ,可以重构出 $N - (m-1)\tau$ 个向量或相点。因此,若预测时间序列长度为 L ,则所需数据点个数为 $L + (m-1)\tau$ 。

1.1.1 延迟时间计算

为求取延迟时间,本文选取互信息法。式(3)为混沌时间序列 $x_1, x_2, \dots, x_{N-1}, x_N$ 的互信息与延迟时间的函数关系^[22],取互信息函数的第一个极小值点为延迟时间。

$$I(\tau) = \sum_{i=1}^N P(x_i, x_{i+\tau}) \ln \left[\frac{P(x_i, x_{i+\tau})}{P(x_i)P(x_{i+\tau})} \right] \quad (3)$$

式中: $P(x_i)$ 为 x_i 在时间序列 $x_1, x_2, \dots, x_{N-1}, x_N$ 中出现的概率; $P(x_{i+\tau})$ 为 $x_{i+\tau}$ 在时间序列 $x_{1+\tau}, x_{2+\tau}, \dots, x_{N-1+\tau}, x_{N+\tau}$ 中出现的概率; $P(x_i, x_{i+\tau})$ 为 x_i 和 $x_{i+\tau}$ 分别在时间序列 $x_1, x_2, \dots, x_{N-1}, x_N$ 和 $x_{1+\tau}, x_{2+\tau}, \dots, x_{N-1+\tau}, x_{N+\tau}$ 中同时出现的联合概率。

1.1.2 嵌入维数计算

为求取嵌入维数, 本文选取伪近邻法。从几何的角度来看, 所观测到的混沌时间序列实际上是高维相空间中系统运动的轨迹在低维空间上的投影。在投影过程中, 轨迹会发生一定的扭曲, 原先在高维相空间中不临近的点, 投影到低维空间中, 就有可能成为相邻的两个点, 这些点被称为伪最近邻点。相空间的重构实际上还原出混沌的运动轨迹。随着嵌入维数的增大, 原系统运动的轨迹会不断的展开, 伪最近邻点也会逐渐的消失。混沌时间序列相空间重构以后, 对于空间中的相点 $\mathbf{X}(t) = (x_t, x_{t+\tau}, x_{t+2\tau}, \dots, x_{t+(m-1)\tau})$ 都存在一个最近的邻点, 两个相点之间的距离为 $R_m(t) = \|\mathbf{X}(t) - \mathbf{X}_j(t)\|$, 当维数增加到 $m+1$ 时, 这两点之间的距离会产生变化, 变化后的距离为^[23]

$$R_{m+1} = \sqrt{R_m^2(t) + \|\mathbf{X}(t+m\tau) - \mathbf{X}_j(t+m\tau)\|^2} \quad (4)$$

如果 R_{m+1} 比 R_m 大很多, 可以认为这是由于在高维相空间中原本并不相邻的两个点投影到低维空间中变为相邻点所造成的。于是有

$$S_m = \frac{R_{m+1}(t)}{R_m(t)} \quad (5)$$

若 $S_m > S$, 那么 $\mathbf{X}(t)$ 与 $\mathbf{X}_j(t)$ 就是伪最近邻点, 其中 S 为阈值。

对混沌时间序列数据, 从嵌入维数 $m=2$ 开始, 计算伪最近邻点的比值或者伪最近邻点的个数。增加嵌入维数, 重新计算直到伪最近邻点的比值小于 5% 或者伪最近邻点个数不再减少, 此时的嵌入维数可以认为能够使得混沌的运动轨迹被全部展开, 也即为最合适的嵌入维数。

1.1.3 混沌特性验证

本文选取排列熵作为混沌特性验证的依据。排列熵算法可分析混沌时间序列的复杂性, 其值介于 0 和 1 之间, 越接近 1, 表示时间序列的混沌特性越强, 具体计算方法可参见文献[24]。

1.2 机器学习模型

本文所选用机器学习模型如图 1 所示。

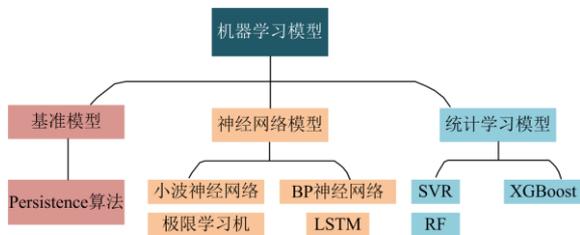


图 1 机器学习模型概览图

Fig. 1 An overview diagram of the data-driven models

基准模型是衡量其他技术模型好坏的参考基准。若模型的预测效果等于或弱于基准模型, 则该模型应被修正或是舍弃。Persistence 模型为常见的有监督机器学习基准模型。Persistence 算法的原理是将 $t-1$ 时刻的负荷值作为 t 时刻的负荷预测值。BP 神经网络(Back Propagation Neural Network, BPNN)是一种按误差反向传播训练的多层前馈网络, 它的基本思想是梯度下降法, 利用梯度搜索技术, 使网络的实际输出值和期望输出值的误差均方差为最小^[25]。小波神经网络(Wavelet Neural Network, WNN)在 BP 神经网络的基础上, 将小波基函数作为隐含层神经元激活函数的神经网络^[26]。支持向量回归(Support Vector Regression, SVR)是支持向量机的回归推广形式, SVR 通过构造一个超平面来解决回归预测问题^[27]。极限学习机(Extreme Learning Machine, ELM)为求解单隐层神经网络的算法, 该算法的提出使得神经网络中的参数不通过多次迭代调整, 而是根据求解公式得到^[28]。随机森林(Random Forest, RF)为采用 Bagging 思想对决策树进行集成的一种算法^[29]。XGBoost 在随机森林的基础上, 引入 Boosting 集成思想, 进一步提高了算法的精确度^[30]。LSTM 是一种特殊的循环神经网络, 解决了循环神经网络在长序列训练过程中的梯度消失和梯度爆炸等问题^[30]。

1.3 基于灰色关联度的加权组合

基于灰色关联度的加权组合法是通过求灰色关联度最大化来确定各单一预测模型所占的比重。

假定某一长度为 N 的时间序列在 M 种预测模型下的预测值为 x_{nm} , 其中, $m=1, 2, 3, \dots, M$, $n=1, 2, 3, \dots, N$ 。则其组合模型为

$$x_n = \sum_{m=1}^M l_m x_{nm} \quad (6)$$

依据文献[31], 基于灰色关联度的组合预测模型求解权重 l_m 等价于求解式(7)。

$$\min \left(\sum_{n=1}^N \left(\sum_{m=1}^M l_m e_{nm} \right)^2 \right)^{\frac{1}{2}} \quad (7)$$

式中, e_{nm} 为第 m 种预测模型在 n 时刻的预测误差。

1.4 评价指标

本文选取两个无量纲评价指标和两个有量纲评价指标。无量纲评价指标: 均方根误差的变异系数 E_{CVRMSE} 和平均绝对百分比误差 E_{MAPE} 。有量纲评价指标: 均方根误差 E_{RMSE} 和平均误差 E_{ME} 。表达式如式(8)一式(11)所示。

$$E_{CVRMSE} = \frac{\sqrt{\sum_i (\hat{y}_i - y_i)^2 / N}}{\sum_i y_i / N} \quad (8)$$

$$E_{MAPE} = \frac{100\%}{N} \sum_i \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (9)$$

$$E_{RMSE} = \frac{1}{N} \sqrt{\sum_i (\hat{y}_i - y_i)^2} \quad (10)$$

$$E_{ME} = \frac{1}{N} \sum_i |\hat{y}_i - y_i| \quad (11)$$

式中： y_i 和 \hat{y}_i 分别为第 i 个样本的实际值和预测值； N 为样本总数。

2 基于相空间重构及机器学习的电力负荷预测

相空间重构作为数据处理的一种方式将历史负荷数据重构为高维相空间，并认为该高维相空间包含了影响历史负荷数据的多种关键信息特征。因此，将相空间重构与机器学习结合时，其结合方式是将相空间重构后的高维相空间作为机器学习模型的训练输入数据。

基于相空间重构及机器学习的电力负荷预测方法训练流程如下：

(1) 计算历史负荷数据的延迟时间与嵌入维数；

(2) 划分历史负荷数据为训练集和测试集，根据延迟时间和嵌入维数对训练集数据进行相空间重构，即为训练输入数据；

(3) 将原始训练集后移一个预测时间长度，即为训练输出数据；

(4) 将训练输入数据与训练输出数据输入机器学习模型，即可进行训练。

本文算例流程如图 2 所示。

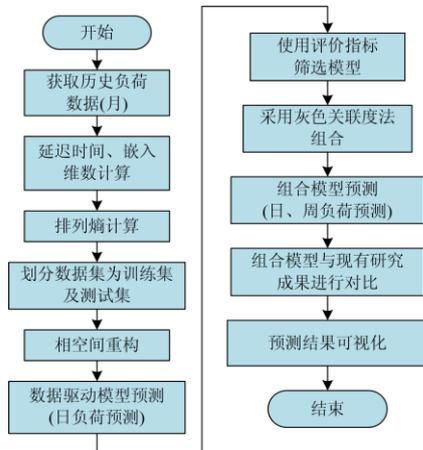


图 2 算例流程图

Fig. 2 Flow chart of case study

3 算例分析

3.1 数据来源

本文负荷数据来源为美国 PJM 电网的数据管理工具 Data Miner 2 网站。本文所指“关键信息缺失”表示缺失与电力负荷相关的天气信息、地理信息、人口信息等，并仅使用历史负荷数据实施预测。对历史负荷数据进行相空间重构，重构后的二维时间序列即为机器学习模型的输入特征。

3.2 相空间重构结果

采用互信息法求得历史负荷数据的互信息函数随延迟时间变换的图像如图 3 所示。从图 3 中可以看出，互信息函数在延迟时间 τ 为 4 时达到极小值点，但该极小值点不太明显，当延迟时间从 2 增加至 10 时，互信息函数值并无较大波动。因此，延迟时间还需要根据排列熵计算进一步确定。采用伪近邻法求得历史负荷数据的伪近邻率随嵌入维数变换的图像如图 4 所示。当嵌入维数增加到 3 时，伪近邻率下降至 5% 以下。因此，取嵌入维数 m 为 3。图 5 为延迟时间和嵌入维数多种取值组合下的排列熵计算结果。从图 5 中可以看出，嵌入维数为 3 时，当且仅当延迟时间取 8 时排列熵取到最大值，最大值为 1。因此，取延迟时间 τ 为 8。重构后的三维相空间如图 6 所示。

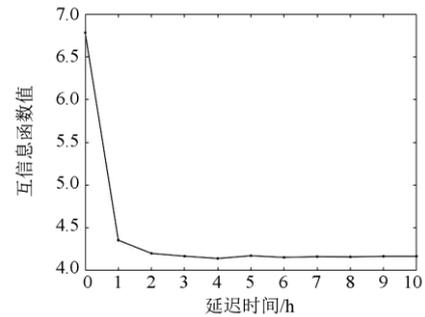


图 3 历史负荷数据的互信息函数

Fig. 3 Mutual information function of historical data

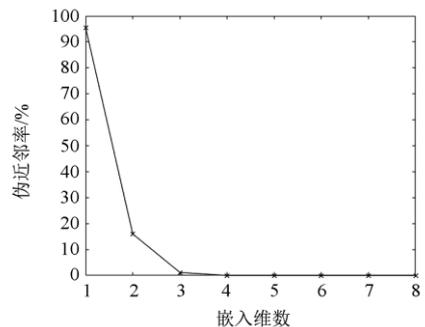


图 4 历史负荷数据的伪近邻率

Fig. 4 Pseudo nearest neighbor rate function of historical data

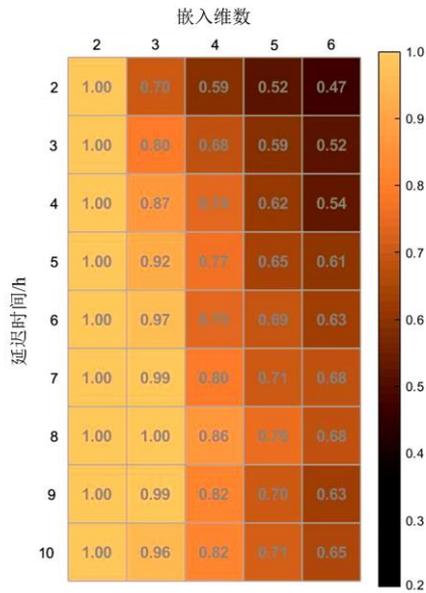


图 5 排列熵计算

Fig. 5 Calculation of permutation entropy

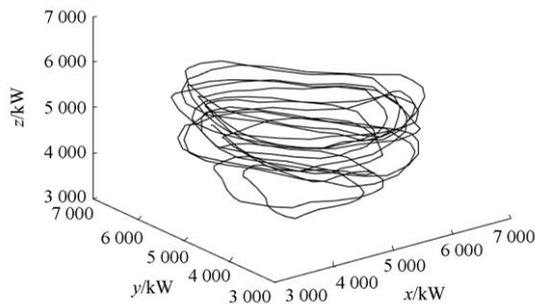


图 6 重构后的三维相空间

Fig. 6 Reconstructed three-dimensional phase space

3.3 超参数设置

表 1 为所使用机器学习模型的超参数设置。

表 1 超参数设置

Table 1 Hyperparameter setting

模型名称	超参数设置
Persistence	—
BPNN	net.trainparam.epochs=500, net.trainparam.goal=0.0001, net.trainParam.lr=0.01
WNN	maxgen=50, net.trainParam.lr=0.01
ELM	—
LSTM	units=1, loss='mae', optimizer='sgd', epochs=50, batch_size=1, verbose=2
SVR	kernel='linear', C=0.001
XGBoost	objective='reg:squarederror', n_estimators=20
RF	n_estimators=2000, random_state=0

3.4 日负荷预测结果对比

将相空间重构后的训练集输入输出、测试集输入输出分别输入各个机器学习模型中。根据 1.2 节所述：若预测时间序列长度为 L ，则所需数据点个数为 $L + (m - 1)\tau$ 。以日负荷预测为例，其输入输出设置如下：训练输入为相空间重构后的 2019 年 7 月 2 日 12 时至 2019 年 7 月 28 日 24 时，训练输出为 2019 年 7 月 3 日 0 时至 2019 年 7 月 29 日 24 时，测试输入为相空间重构后的 2019 年 7 月 28 日 12 时至 2019 年 7 月 29 日 24 时，测试输出为 2019 年 7 月 30 日 0 时至 24 时，数据间隔为 1 h。

采用表 1 中的超参数设置，首先对月历史数据进行相似日负荷预测。除基准模型外，其余模型均通过自主学习训练集数据特征输出测试集预测结果，各个模型在测试集上的预测效果如表 2 所示。

表 2 日负荷预测结果对比

Table 2 Comparison of daily load forecasting results

模型名称	$E_{CVRMSE}/\%$	$E_{MAPE}/\%$	E_{RMSE}/kW	E_{ME}/kW
Persistence	4.75	4.21	226.27	196.46
BPNN	6.19	7.93	295.38	568.52
WNN	5.15	5.79	215.37	370.33
ELM	0.20	0.03	9.63	105.65
LSTM	3.64	3.31	173.92	158.48
SVR	2.84	2.12	136.00	105.86
XGBoost	2.64	1.91	126.08	96.54
RF	2.70	1.92	129.23	96.67

从表 2 中可以看出，BP 神经网络和小波神经网络在测试集上的预测精度低于基准模型，而在实际应用中应舍弃。在神经网络模型中，若以 E_{CVRMSE} 、 E_{MAPE} 或 E_{RMSE} 为标准，则 ELM 误差最低；若以 E_{ME} 为标准，则 XGBoost 表现最优。因此，为达到更高的预测精度，本文选用 ELM 和 XGBoost 两种模型进行组合预测。

3.5 多种时间尺度下预测精度对比

采用 ELM 和 XGBoost 分别进行 1 天至 7 天内不同时间尺度下的负荷预测，并采用灰色关联度法在不同预测时间长度下进行权重分配，结果如表 3 所示。所构建组合模型 ELM-XGBoost 在不同时间长度下的预测精度如图 7 所示。

从图 7 中可以看出，随着预测时间尺度增加，预测误差逐渐增大。可见，相空间重构技术在应用于电力负荷预测时，对于短期负荷预测应用效果更优。

表 3 ELM-XGBoost 模型权重分配

Table 3 Weight allocation of ELM-XGBoost model

模型	1 天	2 天	4 天	7 天
ELM	0.330	0.698	0.321	0.075
XGBoost	0.670	0.302	0.679	0.925

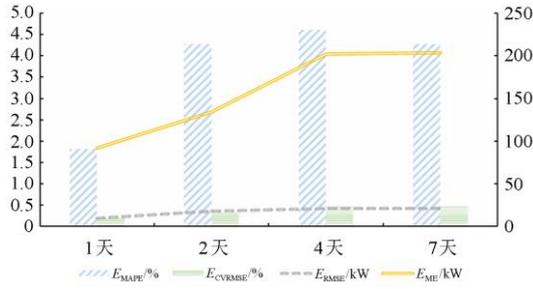


图 7 不同时间尺度下预测精度指标对比

Fig. 7 Comparison of forecasting accuracy in different time scales

3.6 ELM-XGBoost 模型预测结果

分别采用不同组合模型对 1 天及 7 天的数据进行负荷预测，将 ELM-XGBoost 模型与现有文献[2]中所提模型进行对比，包括 BP-SVM 和 GBM-GP，并对将 XGBoost 换成 SVR 得到的 ELM-SVR 模型及将 ELM 换成 LSTM 得到的 LSTM-XGBoost 模型进行对比，得到结果如表 4 所示。

表 4 不同模型结果对比

Table 4 Comparison of different models

模型	$E_{MAPE}/\%$		E_{RMSE}/kW	
	1 天	7 天	1 天	7 天
ELM-XGBoost	1.81	4.28	117.89	274.78
BP-SVM	2.93	4.26	5 497.83	7 675.91
GBM-GP	2.03	2.72	3 648.09	5 435.05
ELM-SVR	3.59	4.35	193.52	278.38
LSTM-XGBoost	2.05	4.32	143.64	278.56

表 4 中的 BP-SVM 模型和 GBM-GP 模型为考虑多种关键信息条件下的组合模型负荷预测，ELM-SVR 模型及 LSTM-XGBoost 模型为关键信息缺失下的组合模型负荷预测。由表 4 可知，与考虑多种关键信息条件下的组合模型相比，在日负荷预测中，本文所构建 ELM-XGBoost 模型的 E_{MAPE} 达到最小，为 1.81%，优于 BP-SVM 模型及 GBM-GP 模型，而周负荷预测误差则稍大，进一步验证了关键信息缺失下相空间重构技术在长时间尺度下的不适应性。综上，在短期负荷预测中，基于相空间重构的电力负荷预测不仅能省去收集和处理多种因素数据的繁琐工作，还能提高预测精度，具有一定的实用价值。与关键信息缺失下的组合模型相比，不论在日负荷预测还是周负荷预测情况下，对比 E_{MAPE} 和 E_{RMSE} 两种指标，本文所构建 ELM-XGBoost 模型均优于 ELM-SVR 模型和 LSTM-XGBoost 模型，显示了本文所构建 ELM-XGBoost 模型的优越性。

图 8、图 9 分别为两种单一模型和组合模型的

日负荷预测曲线和周负荷预测曲线。

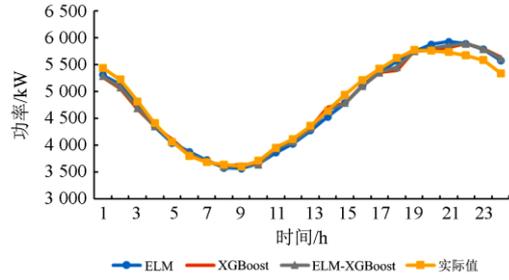


图 8 日负荷预测曲线

Fig. 8 Forecasting curve of daily load forecasting

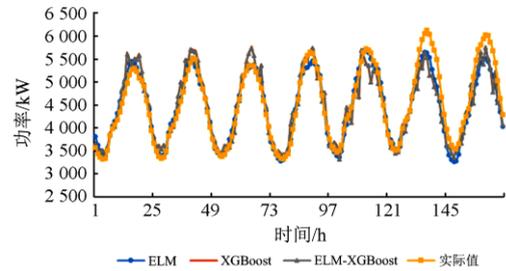


图 9 周负荷预测曲线

Fig. 9 Forecasting curve of weekly load forecasting

从图 8 和图 9 可以看出，在日负荷预测和周负荷预测中，ELM-XGBoost 模型对于日负荷与周负荷均具有良好的跟随性能，显示了在缺失关键信息的情况下相空间重构技术对于电力负荷的适用性。另外，ELM-XGBoost 模型在日负荷预测中的预测精度高于周负荷预测，说明混沌理论中的相空间重构技术更适用于短期(日)负荷预测。

4 结论

为实现在缺失天气等关键信息数据下进行电力负荷预测，本文采用混沌理论中相空间重构技术与机器学习结合的形式进行预测，为提高预测精度，从 8 种机器学习模型中择优组合，通过算例仿真得出以下结论。

(1) 在相空间重构条件下，ELM 和 XGBoost 为所选模型中的最优预测模型。

(2) 采用灰色关联度法对 ELM 和 XGBoost 进行组合，构建了 ELM-XGBoost 模型。在日负荷预测中，与现有部分文献中考虑多种关键信息的负荷预测模型相比，ELM-XGBoost 模型表现更优。

(3) 相空间重构技术适用于短期负荷预测，其预测精度随预测时间尺度增加而呈现减小的趋势。

综上，本文所提方法不仅能省去搜集和处理天气等多种关键信息数据的繁琐，还能在日负荷预测

中进一步提高预测精度。

参考文献

- [1] 马临超, 齐山成, 牛赛, 等. 考虑小区发展不均衡性和不确定性的多阶段空间负荷预测[J]. 电力系统保护与控制, 2021, 49(1): 91-97.
MA Linchao, QI Shancheng, NIU Sai, et al. Multi-stage spatial load forecasting considering the imbalance and uncertainty of the development of the sub-area[J]. Power System Protection and Control, 2021, 49(1): 91-97.
- [2] 陈纬楠, 胡志坚, 岳菁鹏, 等. 长短期记忆网络和 LightGBM 组合模型的短期负荷预测[J]. 电力系统自动化, 2021, 45(4): 91-97.
CHEN Weinan, HU Zhijian, YUE Jingpeng, et al. Short-term load forecasting based on long-term memory network and LightGBM combination model[J]. Automation of Electric Power Systems, 2021, 45(4): 91-97.
- [3] 庞传军, 余建明, 张波, 等. 基于梯度提升树计及非线性的电力负荷影响因素分析[J]. 电力系统保护与控制, 2020, 48(24): 71-78.
PANG Chuanjun, YU Jianming, ZHANG Bo, et al. Nonlinear correlation analysis of influence factors of a power load based on a gradient boosting decision tree[J]. Power System Protection and Control, 2020, 48(24): 71-78.
- [4] 李天云, 刘自发. 电力系统负荷的混沌特性及预测[J]. 中国电机工程学报, 2000, 20(11): 36-40.
LI Tianyun, LIU Zifa. The chaotic property of power load and its forecasting[J]. Proceedings of the CSEE, 2000, 20(11): 36-40.
- [5] FAY D, RINGWOOD J. On the influence of weather forecast errors in short-term load forecasting models[J]. IEEE Transactions on Power Systems, 2010, 25(3): 1751-1758.
- [6] 陆继翔, 张琪培, 杨志宏, 等. 基于 CNN-LSTM 混合神经网络模型的短期负荷预测方法[J]. 电力系统自动化, 2019, 43(8): 131-137.
LU Jixiang, ZHANG Qipei, YANG Zhihong, et al. Short-term load forecasting method based on CNN-LSTM hybrid neural network model[J]. Automation of Electric Power Systems, 2019, 43(8): 131-137.
- [7] 张志, 杜延菱, 崔慧军, 等. 考虑关联因素的智能化中长期电力负荷预测方法[J]. 电力系统保护与控制, 2019, 47(2): 24-30.
ZHANG Zhi, DU Yanling, CUI Huijun, et al. Intelligent mid-long electricity load forecast method considering associated factors[J]. Power System Protection and Control, 2019, 47(2): 24-30.
- [8] 唐俊熙, 曹华珍, 高崇, 等. 一种基于时间序列数据挖掘的用户负荷曲线分析方法[J]. 电力系统保护与控制, 2021, 49(5): 140-148.
TANG Junxi, CAO Huazhen, GAO Chong, et al. A new user load curve analysis method based on time series data mining[J]. Power System Protection and Control, 2021, 49(5): 140-148.
- [9] ALIPOUR M, AGHAEI J, NOROUZI M, et al. A novel electrical net-load forecasting model based on deep neural networks and wavelet transform integration[J]. Energy, 2020, 205.
- [10] 陆继翔, 张琪培, 杨志宏, 等. 基于 CNN-LSTM 混合神经网络模型的短期负荷预测方法[J]. 电力系统自动化, 2019, 43(8): 131-137.
LU Jixiang, ZHANG Qipei, YANG Zhihong, et al. Short-term load forecasting method based on CNN-LSTM hybrid neural network model[J]. Automation of Electric Power Systems, 2019, 43(8): 131-137.
- [11] 吴潇雨, 和敬涵, 张沛, 等. 灰色投影改进随机森林算法的电力系统短期负荷预测[J]. 电力系统自动化, 2019, 39(12): 50-55.
WU Xiaoyu, HE Jinghan, ZHANG Pei, et al. Power system short-term load forecasting based on improved random forest with grey relation projection[J]. Automation of Electric Power Systems, 2019, 39(12): 50-55.
- [12] O'NEILL-CARRILLO E, HEYDT G T, KOSTELICH E J. Chaotic phenomena in power systems: detection and applications[J]. Electric Machines and Power Systems, 1999, 27(1): 79-91.
- [13] 方仍存, 周建中, 彭兵, 等. 电力负荷混沌动力特性及其短期预测[J]. 电网技术, 2008, 32(4): 61-66.
FANG Rengcun, ZHOU Jianzhong, PENG Bing, et al. Chaotic dynamics of power load and its short-term forecasting[J]. Power System Technology, 2008, 32(4): 61-66.
- [14] 杨正瓴, 林孔元. 电力系统负荷记录混沌特性成因的探讨[J]. 电力系统自动化, 2002, 26(10): 18-22.
YANG Zhengling, LIN Kongyuan. Discussion on causes of chaos in load records of power systems[J]. Automation of Electric Power Systems, 2002, 26(10): 18-22.
- [15] 孙雅明, 张智晟. 相空间重构和混沌神经网络融合的短期负荷预测研究[J]. 中国电机工程学报, 2004, 24(1): 44-48.
SUN Yaming, SUN Zhisheng. A new model of STLF based on the fusion of PSRT and chaotic neural networks[J]. Proceedings of the CSEE, 2004, 24(1): 44-48.
- [16] 张智晟, 孙雅明, 王兆峰, 等. 相空间近邻点与递归神经网络融合的短期负荷预测[J]. 中国电机工程学报,

- 2003, 23(8): 45-50.
ZHANG Zhisheng, SUN Yaming, WANG Zhaofeng, et al. A new STLF approach based on the fusion of optimal neighbor points in phase space and the recursive neural network[J]. Proceedings of the CSEE, 2003, 23(8): 45-50.
- [17] XU Z, ZHONG L, ZHANG A. Phase space reconstruction-based concept or network for time series prediction[J]. IEEE Access, 2019, 7: 163172-163179.
- [18] LI J, ZHANG Q, ZHANG Y, et al. Hidden phase space reconstruction: a novel chaotic time series prediction method for speech signals[J]. Chinese Journal of Electronics, 2018, 27(6): 1221-1228.
- [19] TAKENS F. Detecting strange attractors in turbulence[M]. Heidelberg: Springer Berlin Heidelberg, 1981.
- [20] 王德意, 杨卓, 杨国清. 基于负荷混沌特性和最小二乘支持向量机的短期负荷预测[J]. 电网技术, 2008, 32(7): 66-71.
WANG Deyi, YANG Zhuo, YANG Guoqing. Short-term load forecasting based on chaotic characteristic of loads and least squares support vector machine[J]. Power System Technology, 2008, 32(7): 66-71.
- [21] GUO Fengfan, LI Lingpeng, WEI Chiangong. Short term load forecasting based on phase space reconstruction algorithm and bi-square kernel regression model[J]. Applied Energy, 2018, 224: 13-33.
- [22] ALBERS D, HRIPCSAK G. Estimation of time-delayed mutual information and bias for irregularly and sparsely sampled time-series[J]. Chaos, Solitons & Fractals, 2012, 45(6): 853-860.
- [23] 韩亚军, 李太福. 相空间重构短期风速与发电功率在线预测[J]. 控制工程, 2019, 26(8): 1503-1508.
HAN Yajun, LI Taifu. Online prediction of short-term wind speed and power generation based on phase space reconstruction[J]. Control Engineering of China, 2019, 26(8): 1503-1508.
- [24] LI G, LIT Y, ROOZITALAB F. Midterm load forecasting: a multistep approach based on phase space reconstruction and support vector machine[J]. IEEE Systems Journal, 2020, 14(4): 4967-4977.
- [25] 左远龙, 黄玉水, 杨晓辉, 等. 基于 PFA-MBAS-BP 神经网络模型的光伏发电短期预测[J]. 电力系统保护与控制, 2020, 48(15): 84-91.
ZUO Yuanlong, HUANG Yushui, YANG Xiaohui, et al. Short-term prediction of photovoltaic power generation based on a PFA-MBAS-BP neural network model[J]. Power System Protection and Control, 2020, 48(15): 84-91.
- [26] LIU Y, YANG Q, AN D, et al. An improved fault diagnosis method based on deep wavelet neural network[C] // 2018 Chinese Control and Decision Conference (CCDC), June 9-11, 2018, Shenyang, China: 1048-1053.
- [27] CHEN Y, XU P, CHU Y, et al. Short-term electrical load forecasting using the support vector regression (SVR) model to calculate the demand response baseline for office buildings[J]. Applied Energy, 2017, 195: 659-670.
- [28] ERTUGRUL Ö. Forecasting electricity load by a novel recurrent extreme learning machines approach[J]. International Journal of Electrical Power & Energy Systems, 2016, 78: 429-435.
- [29] KIRANMAI S, LAXMI A. Data mining for classification of power quality problems using WEKA and the effect of attributes on classification accuracy[J]. Protection and Control of Modern Power Systems, 2018, 3(1): 1-12.
- [30] WANG Y, SUN S, CHEN X, et al. Short-term load forecasting of industrial customers based on SVM and XGBoost[J]. International Journal of Electrical Power & Energy Systems, 2021, 129.
- [31] 张和平, 解晓龙. 基于灰色关联度的组合优化模型研究[J]. 统计与决策, 2019, 35(9): 19-23.
ZHANG Heping, XIE Xiaolong. Research on combinatory optimization model based on grey relational degree[J]. Statistics and Decision, 2019, 35(9): 19-23.

收稿日期: 2021-05-17; 修回日期: 2021-09-29

作者简介:

侯慧(1981—), 女, 博士, 副教授, 博士生导师, 研究方向为能源互联网、电动汽车智能充电策略等; E-mail: houhui@whut.edu.cn

王晴(1998—), 女, 硕士研究生, 研究方向为能源互联网、负荷预测等; E-mail: wangqing1342@whut.edu.cn

吴细秀(1975—), 女, 通信作者, 博士, 副教授, 硕士生导师, 研究方向为高电压绝缘、高压电器电磁仿真等。E-mail: wuxixiu@whut.edu.cn

(编辑 许威)