

基于 CEEMD 和 SVR 的城市居民用电负荷短期预测

王乙儒

(中央财经大学, 北京 102200)

摘要:城市居民用电负荷短期预测是智能电网的重要组成部分,为了解决居民用电负荷序列因其具有随机性、非平稳性等特征导致的难以准确预测的问题,本文对家庭电力消耗数据进行研究,提出基于互补集合经验模态分解和支持向量回归的混合预测模型,以期为相关人员提供参考。

关键词:用电负荷短期预测;互补集合经验模态分解;支持向量回归;混合预测模型

中图分类号:TM614

文献标识码:A

文章编号:1004-7344(2021)31-0304-03

0 引言

城市居民用电负荷短期预测在智能电网中发挥着重要作用,对居民用电负荷的短期预测能帮助人们合理有效地管理电力系统,减少用电成本,有助于社会的可持续发展,并对未来国家电力行业的发展有重要意义。

用电负荷的短期预测是指对未来几个小时或几天的电负荷数据进行预测。受节假日、天气和居民习惯等因素的影响,居民短期用电负荷序列具有非平稳性和随机性的特点,比较难以预测。常用的短期负荷预测的方法有支持向量机、K 邻近和神经网络^[1]等,这些单一模型预测的方法原理简单、精度低、泛化能力弱,不能作为有效模型使用。为了提高预测精度,混合预测模型被提出,混合预测模型即是在原来预测模型的基础上加上数据预处理的方法。对于时间序列数据,数据预处理阶段使用信号分析的方法可以很好地减弱非平稳性对模型精度的影响^[2,3]。文献[2]使用最小二乘支持向量机与经验模态分解(EMD)建立模型进行预测,并分析了分解的分量和特征之间的相关性,结果与传统单一模型相比预测精度得到提高。文献[3]使用 ARIMA 与集合经验模态分解(EEMD)相结合的方法,对月负荷量用 EEMD 预处理,相比于 EMD, EEMD 更能抑制噪声的影响,随后使用 ARIMA 模型进行预测,结果优于单一的 ARIMA 模型。

本文提出了一种基于互补集合经验模态分解和支持向量回归的 CEEMD-SVR 混合预测模型。CEEMD 算法分解用电负荷时间序列可以有效的抑制噪声引起的误差,SVR 使用一定的方法将数据映射到高维区域,并通过少数的支持向量确定最终结果,具有一定的鲁棒性。通过实例分析验证了模型的有效性。

1 理论基础

1.1 互补集合经验模态分解

EMD 将信号分解为不同的 IMF,把原始信号转化为相对平

稳的信号分量。非平稳信号适合用 EMD 分解,但会出现模态重叠的问题。EEMD 通过加入高斯白噪声来抑制可能出现的模态重叠问题。CEEMD 是 EMD 和 EEMD 的改进,CEEMD 在原始信号中加入一对互补的白噪声,经过多次平均后噪声的影响会逐渐减弱。CEEMD 的过程描述如下:

(1)在原始振动信号 $x(t)$ 中加入一对正负互补的固定强度白噪声 ω_1 。使用 EMD 算法对信号 $x(t) + \mu\omega_1$ 进行分解, μ 为噪声的幅值,并计算一系列模态的均值得到 IMF_1 :

$$IMF_1 = \frac{1}{N} \sum_{i=1}^N E_i[x(t) + \mu\omega_1] \quad (1)$$

式中: E_i -EMD 分解的第 i 个分量。

(2)使用原始振动信号 $x(t)$ 减去 IMF_1 ,会得到一个残差分量,记为 r_1 。 r_1 中加入新的白噪声 ω_1 ,使用 EMD 算法进一步分解,然后得到 IMF_2 。

(3)重复上述过程,直到得到第 $(k-1)$ 个残差分量 r_{k-1} ,然后得到 IMF_k 。

(4)重复步骤 3,提取每个 IMF,直到残差分量中的极值点数不大于 2,最终的残差分量记为 R 。最后,原始信号 $x(t)$ 可写成多个 IMF 分量和 R 之和。

1.2 支持向量回归(SVR)

支持向量机(SVM)是分类领域最有效的工具之一。SVM 区别于一般的分类方法的时,它会将输入映射到高维空间。在处理一些非线性数据时,通过映射的方法可以将问题变为线性,将问题进行简化。SVR 是 SVM 应用在回归问题上的扩展,回归方程为:

$$y = f(x) = \omega \phi(x) + b \quad (2)$$

其中: ω -权值; b -偏差; $\phi(x)$ -可以被核函数替代的非线性函数。SVR 是要找到一个函数,使其误差不大于一个损失 ϵ 。该问题的

目标函数可以被优化为:

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (3)$$

式中: C-惩罚系数, 目标函数的约束条件为:

$$|y_i - w^T x_i - b| \leq \varepsilon + \xi_i \quad (4)$$

核函数的选择会对支持向量回归的结果产生很大的影响。核函数有很多不同的种类, 高斯核函数(RBF)作为用得最多的核函数之一, 适用于大部分场景, 同时训练速度快, 本文也将使用该核函数。

1.3 CEEMD-SVR 短期负荷预测模型

基于负荷时间序列的非平稳性和随机性, 本文提出混合预测模型 CEEMD-SVR。模型对 CEEMD 分解原负荷时间序列得到的不同的 IMF 分别建立 SVR 子模型, 最终的预测结果为子模型预测结果的和。图 1 为模型的基本结构。

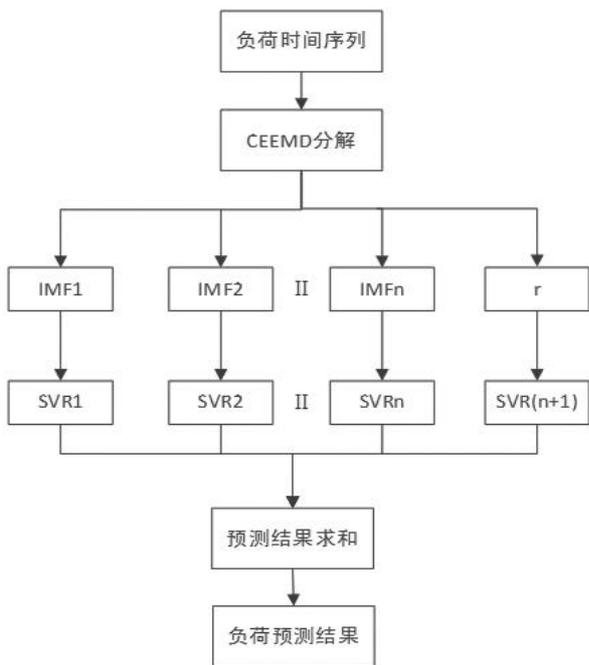


图 1 模型基本结构

2 实验验证

2.1 数据来源

数据集来源于 UCI 数据库中的个人家庭用电数据集。该数据集包含 2006 往后近 4 年的时间里一个家庭的电力消耗, 采样周期为一分钟。选取 2010 年 10 月 1 日到 2010 年 10 月 31 日的的数据, 由于原数据采样频率过高, 对数据进行小时平均化处理, 处理后共有 744 个样本点。选取前 30 天的数据作为训练集, 第 31 天的数据作为预测集。对数据使用 CEEMD 进行分解, 分解的 IMF 频率不同, 高频分量一般代表不确定因素; 一些低频分量具有周期性, 可看作是序列的周期性影响因素。

2.2 评估指标

为了对模型进行评估, 这里采用回归问题常用的三个指标, 分别为 MAE、MSE 和 MAPE, 它们的含义分别为平均绝对误差、均方误差和平均相对误差。这三个指标从不同的方面衡量模型的预测性能, 值越趋于零代表模型的预测效果越好。

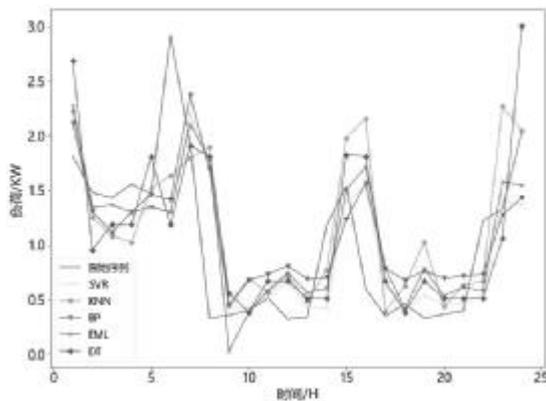
2.3 模型参数

参数的选取对 SVR 有十分重要的影响。在 1.2 节介绍了核函数的选择, 使用的是高斯核函数。除此之外, 还有两个比较重要的参数 C 和 gamma。C 是惩罚系数, 作用于损失函数, C 越大分类越严格。gamma 核函数系数, 特征维度的倒数, 主要控制样本的高维度映射。上述两个重要参数的最佳值通过网格搜索的方法确定, gamma 搜索的范围是 [0.01, 0.1, 1, 10], C 是 [0.1, 1, 10, 100], 并使用 8 折交叉验证。

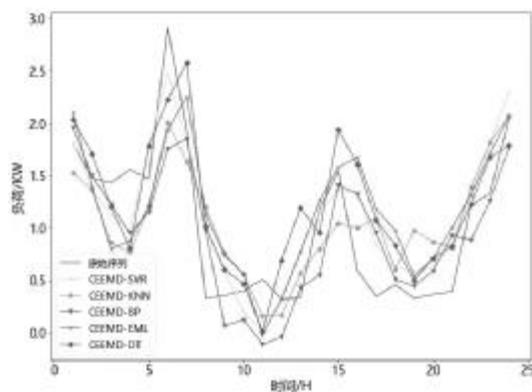
模型的输入长度和超前预测的步长也是需要确定的。部分学者的研究表明, 输入长度相同的情况下超前预测的步长增加会降低预测精度^[9], 因此这里选择超前一步预测。模型输入的长度通过实验确定, 考虑到模型的精度和训练时长, 选择输入长度为 3。

2.4 预测结果分析

为了验证本文提出的模型是否有效, 引入其他的单一和混合预测模型进行对比。单一模型分别为 SVR、KNN、BP、极限学习机(ELM)和决策树(DT); 混合模型为 CEEMD 和单一模型的组合。模型中出现的所有相同的方法参数也都是一样的。其中图 2(a) 和图 2(b) 分别是原始序列和单一模型或混合模型的预测值拟合图。表 1 是所有模型的评价指标展示。



(a) 单一模型结果



(b) 混合模型结果

图 2 不同模型的预测结果

从图 2(a) 中看出, 各单一模型的预测结果序列图跟原序列形状基本相似, 但有部分预测值存在较大的偏差, 预测的总体效果并不好。而且由图可知, 预测有滞后现象, 这表明单一模型不能很好地学习到负荷序列的内部规律。从图 2(b) 中看出, 各混合

(下转第 308 页)

方和根求相对合成标准不确定度:

$$u_{\text{crel}} = \sqrt{u_{1\text{rel}}^2 + u_{2\text{rel}}^2 + u_{3\text{rel}}^2 + u_{4\text{rel}}^2}$$

$$= \sqrt{(3.12 \times 10^{-3})^2 + (5.73 \times 10^{-3})^2 + (4.13 \times 10^{-3})^2 + (8.16 \times 10^{-4})^2}$$

$$= 0.0078 \quad (20)$$

$$u_c = u_{\text{crel}} \times \omega = 0.0078 \times 0.0213\% = 1.66 \times 10^{-6} \quad (21)$$

2.3.6 扩展不确定度

通常取 95% 置信水平, 包含因子 $k=2$, 扩展不确定度为:

$$U = u_c \times k = 1.66 \times 10^{-6} \times 2 = 3.32 \times 10^{-6} \quad (22)$$

2.3.7 不确定度报告

根据扩展不确定度来确定电感耦合等离子体发射光谱法测定镁及镁合金中锰含量为: $0.0213\% \pm 0.000332\%$, $k=2$ 。

3 结论

通过对电感耦合等离子体发射光谱法测定镁及镁合金中锰含量的不确定度进行系统的分析, 计算出合成不确定度, 给出电感耦合等离子体发射光谱法测定镁及镁合金中锰含量为 $0.0213\% \pm 0.000332\%$, $k=2$ 。锰含量的置信区间为: $[0.020968\%, 0.021632\%]$ 。

从各分量评定的量值看, 在诸多分量中, 工作曲线变动性和锰标准溶液对合成不确定度的贡献较大, 因此规范分析操作, 提高标准溶液配制的精确度, 不断优化仪器参数, 都可以减小工作曲线变动性和锰标准溶液对不确定度的影响, 对分析测试准确度和可靠性的提高非常重要。同时本文介绍的 ICP 光谱法测定

镁及镁合金中锰含量的不确定度的评定方法可以推广到其他元素的不确定度的评定。

参考文献

- [1] 唐语, 王亚鹏, 张鲁宁. 电感耦合等离子体原子发射光谱法测高碳铬铁中锰钛磷含量的不确定度评定[J]. 工业计量, 2020, 30(1): 55-57.
- [2] IOS/IEC 17025: 2017, General requirements for the competence of calibration and testing laboratories[S]. 2017.
- [3] 镁及镁合金化学分析方法 第 20 部分: ICP-AES 测定元素含量 (GB/T 13748.20—2009) [S]. 北京: 中国标准出版社, 2009.
- [4] 电感耦合等离子体发射光谱法测量结果不确定度评定规范: CSM 01 01 01 04—2006[S]. 北京: 中国标准出版社, 2006.
- [5] 测量不确定度评定与表示: JJF 1059.1—2012[S]. 北京: 中国标准出版社, 2012.
- [6] 国家标准物质研究中心. 化学分析测量不确定度评定: JJF 1135—2005[S]. 2005.
- [7] 河南省计量科学研究院, 上海市计量测试技术研究院, 北京市计量检测科学院研究院. 常用玻璃量器检定规程: JJG 196—2006[S]. 北京: 中国计量出版社, 2006.
- [8] 化学分析中不确定度的评估指南: CNAS-GL06—2006[S]. 2006.

收稿日期: 2021-07-07

作者简介: 李珍妮(1990—), 女, 汉族, 陕西横山人, 硕士研究生, 工程师, 主要从事产品检验检测工作。

(上接第 305 页)

表 1 不同模型的预测误差

预测模型	MAE	MSE	MAPE	
单一模型	SVR	0.377	0.317	52.3
	BP	0.385	0.294	60.3
	ELM	0.400	0.309	67.9
	KNN	0.444	0.398	70.4
	DT	0.483	0.438	66.6
混合模型	CEEMD-SVR	0.256	0.080	38.9
	CEEMD-DT	0.287	0.121	48.8
	CEEMD-BP	0.291	0.139	47.6
	CEEMD-KNN	0.317	0.184	49.1
	CEEMD-ELM	0.332	0.155	58.5

模型的预测结果曲线跟原曲线基本拟合, 预测值有偏差但偏差较小, 而且没有滞后的问题。结合图 2(a) 和 (b) 来看, 使用 CEEMD 对负荷数据进行分解可以有效挖掘负荷数据的内在特性, 提高预测的精准度。

由表 1 可知, 单一模型中, SVR 表现最好, 进一步推测相比于其他模型, SVR 可能更适合此类问题的预测; 混合模型中, CEEMD-SVR 表现最好, 它同时也是所有模型中表现最好的, MAE、MSE 和 MAPE 值都低于其他所有模型。同时, 表 1 中的数据也进一步验证了混合模型比其对应的单一模型的表现更好。

3 结论

对于非平稳的城市居民用电负荷时间序列的预测, 本文使用

互补集合经验模态分解对序列做平稳化处理并建立不同的支持向量回归子模型, 把各子模型预测结果的叠加求和作为最终预测结果。为了验证模型的优越性, 本文引入单一模型和不同混合模型进行对比实验。实验证明了 CEEMD 分解可充分挖掘具有非平稳性及随机性特征的负荷序列的内在规律, 提高模型预测的精度; 并且相比于其他的混合预测模型, CEEMD-SVR 模型的预测效果更佳。

参考文献

- [1] 邹治锐, 高坤, 朱伟, 等. 基于神经网络 BP 算法的局部电网短期负荷预测系统[J]. 湖南电力, 2020, 40(2): 74-77.
- [2] 孔祥玉, 李闯, 郑锋, 等. 基于经验模态分解与特征相关分析的短期负荷预测方法[J]. 电力系统自动化, 2019, 43(5): 46-52.
- [3] 郭建鹏, 余颖铃, 温步瀛. 基于 EEMD-ARIMA 模型的地区月负荷量预测[J]. 电力工程技术, 2018, 37(6): 28-32, 74.
- [4] WANG C, ZHANG H, MA P. Wind power forecasting based on singular spectrum analysis and a new hybrid Laguerre neural network[J]. Applied Energy, 2020, 259(9): 114139.

收稿日期: 2021-07-04

作者简介: 王乙儒(2000—), 男, 汉族, 河南周口人, 本科在读, 研究方向为人工智能。