

基于最大信息系数和多目标Stacking集成学习的 综合能源系统多元负荷预测

崔树银,汪昕杰

(上海电力大学 经济与管理学院,上海 200090)

摘要:精确的多元负荷预测对于综合能源系统的能源调度与运行规划起到重要的作用。对电、热、冷负荷单独进行预测的传统方法会忽略多元负荷间的耦合关系。针对这一问题,提出一种基于多目标Stacking集成学习的多元负荷协同预测模型。引入最大信息系数对多元负荷及天气因素进行相关性分析,并提出负荷耦合形态指标来深度挖掘多元负荷间的耦合关系;将多目标回归与Stacking集成学习模型相结合,建立多元负荷协同预测模型;通过实际算例验证所提模型的有效性,算例结果表明,与其他预测模型相比,所提模型预测精度更高。

关键词:多目标回归;Stacking集成学习;综合能源系统;最大信息系数;正则化贪心森林算法

中图分类号:TM 715

文献标志码:A

DOI:10.16081/j.epae.202202025

0 引言

能源是人类社会发展的关键因素。据预测,2030年能源需求将增长50%以上^[1]。在传统能源子系统单一运行模式下,电、热、冷等能源之间无法得到充分转化,导致能源利用效率低下以及清洁能源无法被充分消纳等^[2]。因此,综合能源系统IES(Integrated Energy System)和能源互联网EI(Energy Internet)被相继提出。IES作为EI的物理载体,包含电力供应、热力供应、天然气供应等^[3]。多元负荷预测的精度直接关系到IES的能源调度和运行规划是否合理有效。传统负荷预测只需考虑单一能源负荷,而IES负荷预测需综合考虑电、热、冷等多类型能源之间的耦合关联,使得IES多元负荷预测的难度增大。

目前,针对传统单一能源子系统负荷预测研究已取得较好的研究成果,例如:在传统的电力负荷预测方面,文献[4]考虑天气条件、电价、日历信息和历史负荷数据,采用联合概率密度方法进行预测;在热负荷预测方面,文献[5]将灰狼优化算法和差分进化相结合来对支持向量机进行优化,该预测方法同样适用于换热站和热源;在冷负荷预测方面,文献[6]针对冰蓄冷空调中冷负荷预测存在数据关联度低的情况,提出了一种组合神经网络模型来解决这一问题。

相较于单一负荷预测,IES多元负荷预测相关研究仍处于初级阶段。近年来,机器学习等人工智能技术广泛应用于负荷预测领域,针对IES多元负荷

间复杂的耦合关系,无需建立准确的数学模型,仅通过非线性映射分层学习就能得到数据之间的关系。文献[7]利用混沌理论进行变量空间重构,并采用卡尔曼滤波对电、热、冷负荷进行预测;文献[8]将多任务学习和长短时记忆网络相结合,通过共享层来模拟各能源子系统负荷之间的耦合特性;文献[9]通过构建深度结构的多任务学习模型,考虑热、电、气之间的耦合性,建立了“离线+在线”的预测模式;文献[10]基于多层径向基函数神经网络模型实现了电、气、热多元负荷预测;文献[11]利用核主成分分析法对原始数据进行降维,并分别用深度双向长短期记忆神经网络和多元线性回归模型进行多元负荷预测,然后对采用这2个模型预测的结果进行重构得到最终结果。上述文献针对多元负荷预测大多采用单一预测模型或将几个模型相组合的方法。近年来,多模型融合算法在负荷预测领域取得了一定的成就,文献[12]在Stacking集成学习框架下融合多个机器学习及深度学习算法,通过实际算例证明多模型融合算法相较于单一模型或简单的模型组合方法有着更高的预测精度。目前将多模型融合算法用于多元负荷预测领域的研究较少。因此,本文借助Stacking集成学习的思想,同时引入多目标回归来解决传统Stacking集成学习模型无法进行多变量预测的问题。

综上,本文提出一种基于多目标Stacking集成学习模型的多元负荷协同预测方法。首先,引入最大信息系数MIC(Maximal Information Coefficient)理论,对多元负荷及天气因素进行相关性分析;其次,引入负荷耦合形态指标,深度挖掘多元负荷间的耦合关系,并通过计算得到的MIC值进行输入变量筛选;然后,将多目标回归的思想与Stacking集成学习

收稿日期:2021-06-11;修回日期:2021-12-29

在线出版日期:2022-02-28

基金项目:国家自然科学基金资助项目(71972127)

Project supported by the National Natural Science Foundation of China(71972127)

模型相结合,建立多元负荷协同预测模型;最后,通过实际算例验证所提模型的可行性。

1 多元负荷及天气因素相关性分析

IES是一个以电力系统为核心,具有各种信息化设备的能源平衡系统,可以实现电、热、冷、气的多能互补。具体的IES交互图^[13]如附录A图A1所示。从图中可以看出,IES各能源子系统之间可以通过某种能源转换装置进行能源二次利用。不同负荷之间能通过特定形式实现相互转换,表明了多元负荷之间存在复杂的耦合关系。因此,深度挖掘多元负荷之间的耦合特性,同时分析多元负荷及天气因素之间的相关性,这对于提高模型预测精度起到了重要的作用。

1.1 MIC

MIC^[14]于2011年首次被提出,不仅可以衡量2个变量之间是否存在线性关系,还可以观测变量间是否具有正弦性、周期性等非线性关系,而且MIC对于含有噪声的样本具有较好的鲁棒性。MIC值的取值范围在0~1之间,MIC值越大,表明2个变量间的关联程度越高。

MIC的基本原理为:如果2个变量之间存在相关性,则可以在由2个变量组成的散点图上绘制网格,并利用互信息MI(Mutual Information)和网格划分来计算MIC。利用MIC进行相关性分析的具体步骤如下。

1)设天气因素变量 $A=\{a_i\}(i=1,2,\dots,n)$ 和多元负荷 $B=\{b_j\}(j=1,2,\dots,m)$,其中 n 和 m 分别为天气因素和多元负荷变量的数量。即MI值 $f_{\text{MI}}(A,B)$ 的计算公式为:

$$f_{\text{MI}}(A,B)=\sum_{a_i \in A} \sum_{b_j \in B} p(a_i, b_j) \log_2 \frac{p(a_i, b_j)}{p(a_i)p(b_j)} \quad (1)$$

式中: $p(a_i, b_j)$ 为变量 a_i 和 b_j 的联合概率密度; $p(a_i)$ 和 $p(b_j)$ 分别为变量 a_i 和 b_j 的边缘概率密度。

2)定义一个网格,记作 $G=(x,y)$,将数据集 $D=\{(a_i, b_j)\}$ 中 a_i 和 b_j 的值分别划分到 x,y 这2个格子中,并计算得到 D 在网格 G 划分下的最大MI值为:

$$f_{\text{MI}^*}(D,x,y)=\max f_{\text{MI}}(D|G) \quad (2)$$

式中: $f_{\text{MI}}(D|G)$ 为数据集 D 在 G 下的MI值。

3)对于固定的数据集 D ,将 D 上不同网格 G 的所有最大MI值归一化至(0,1)区间内,归一化的计算公式如下:

$$M(D)_{x,y}=\frac{f_{\text{MI}^*}(D,x,y)}{\log_2 \min\{x,y\}} \quad (3)$$

4)MIC即为归一化后的最大MI值,MIC计算公式如下:

$$f_{\text{MIC}}(D)=\max_{xy < B(\delta)} \{M(D)_{x,y}\} \quad (4)$$

式中: $B(\delta)$ 为关于 δ 的函数,设为 $\delta^{0.6}$ 。

根据上述方法,对实际美国亚利桑那州立大学Tempe校区的全年IES数据及天气数据进行相关性分析。

1.2 多元负荷之间的相关性分析

在IES内部中,不同能源之间可以通过某种特定的形式进行相互转化,这表明多元负荷之间存在一定的相关性。多元负荷曲线如附录A图A2所示。从图中可以看出,电、冷负荷的曲线波动较为接近,都呈现“夏高冬低”的趋势,而热负荷曲线波动正好与之相反,呈现“夏低冬高”的趋势,这表明电、热、冷负荷之间有较强的季节互补性。为了更加直观地量化多元负荷之间的关系,选择采用MIC值来测量电、热、冷负荷间的关联程度。由于电、热、冷负荷的季节互补性,在不同季节电、热、冷负荷存在不同程度的相关性,其量化结果如表1所示。

表1 在不同季节多元负荷之间的MIC值

Table 1 MIC values among multiple loads in different seasons

多元负荷	MIC值			
	春季	夏季	秋季	冬季
电负荷与热负荷	0.4342	0.3959	0.4807	0.7712
电负荷与冷负荷	0.5053	0.6432	0.6051	0.2003
热负荷与冷负荷	0.4647	0.4184	0.6230	0.2951

由表1可知,在不同季节电、热、冷负荷之间的相关程度存在明显差异。本文将MIC值大于0.3的变量作为强相关变量,MIC值小于0.1的变量作为弱相关变量。夏季电负荷与热负荷间的相关程度最低,而电负荷与冷负荷间的相关程度最高,而冬季与夏季正相反;春季和秋季的电、热、冷负荷间的MIC值均大于0.3,表明其具有较强的相关性。上述分析表明在进行多元负荷预测时,不能忽略负荷间的相关性。

1.3 多元负荷与天气因素的相关性分析

电、热、冷负荷除了受到负荷之间的相互耦合影响之外,对于天气因素也较为敏感。不同季节的天气因素对于电、热、冷负荷的影响程度不同。因此,深度挖掘电、热、冷负荷与天气因素之间的相关性,对提升模型的预测精度起到重要的作用,具体结果见附录A表A1。从表中可知,不同季节的天气因素与多元负荷之间的相关程度存在明显差异。如在春季,温度与多元负荷间的MIC均大于0.3,具有较强的相关性,而降雨量与多元负荷间的MIC均小于0.1,相关程度较弱。因此,根据附录A表A1所示的季节差异性分析可知在不同季节应选取不同的天气因素作为输入变量。

1.4 负荷耦合形态指标的相关性分析

由于IES特有的交互结构,其多元负荷间具有

较强的耦合特性,深度挖掘其耦合关系对于提升模型的预测精度起到重要作用。因此,为了反映多元负荷间的耦合特性,提出电冷比、电热比、冷热比这3个负荷耦合形态指标,以充分利用IES多元负荷间的耦合信息来提升多元负荷协同预测的准确性。为了验证负荷耦合形态指标能增强多元负荷的可预测性,利用MIC值来量化负荷耦合形态指标与多元负荷间的相关程度,结果如表2所示。

表2 负荷耦合形态指标与多元负荷间的MIC值

Table 2 MIC values between load coupling morphological indexes and multiple loads

负荷	负荷耦合形态指标	MIC值			
		春季	夏季	秋季	冬季
电负荷	电冷比	0.5397	0.6147	0.4337	0.1728
	电热比	0.6995	0.1992	0.7543	0.3646
	冷热比	0.6026	0.3738	0.6101	0.2521
冷负荷	电冷比	0.8545	0.7002	0.8859	0.4930
	电热比	0.4967	0.3974	0.7178	0.6549
	冷热比	0.8292	0.8297	0.9398	0.8870
热负荷	电冷比	0.3766	0.1580	0.5706	0.5482
	电热比	0.7012	0.3051	0.8027	0.7156
	冷热比	0.5619	0.2422	0.7349	0.7835

由表2可知,负荷耦合形态指标与多元负荷间具有较强的相关性。在夏季,电冷比、冷热比与多元负荷间的相关程度较高,其中冷热比与冷负荷间的MIC超过0.8;在冬季,冷热比、电热比与多元负荷间的相关程度较高,其中电热比与热负荷间的MIC超过0.7。由此可知,负荷耦合形态指标可以很好地反映多元负荷间的耦合关系。因此,选取负荷耦合形态指标作为模型的输入变量来提升多元负荷协同预测的精度,具有一定的可行性。

2 基于多目标Stacking集成学习的多元负荷协同预测模型

2.1 多目标回归

多目标回归^[15]也可以被称为多变量预测和多输出回归。多目标回归是基于多标签分类的思想演化而来的,可以同时预测多个相互关联的连续型变量,通过深度挖掘多个输出变量之间的关联性来提升模型的预测精度。

多目标回归算法目前有堆叠单目标回归SST(Stacked Single-Target)算法和集成回归链ERC(Ensemble of Regressor Chains)算法这2种。SST算法主要分为以下2个阶段:第一阶段,针对每个输出变量,先进行单目标的回归预测;第二阶段,将第一阶段每个输出变量的预测值扩充至原始输入空间,组成新的输入变量,再输入同一个多目标回归模型进行多输出预测。ERC算法为每个输出变量建立一个单独的回归模型,与SST算法不同的是,其在扩

充原始输入空间时,采用的是对先前输出变量的实际值进行扩充,再随机生成多条回归链,每个输出变量基于链的不同位置得到其预测值,最后通过计算平均值得到最终的预测结果。

2.2 双向门控循环单元原理

双向门控循环单元BiGRU(Bidirectional Gated Recurrent Unit)^[16]将具有相反传输方向的2个隐藏层连接到同一输出层,以便于输出层从过去和将来的状态中获取信息。这意味着BiGRU神经网络能够从2个不同的数据方向学习信息,从而进行更准确的预测。BiGRU的思路是将规则的GRU神经单元分为前向传播和后向传播。BiGRU的结构如图1所示。

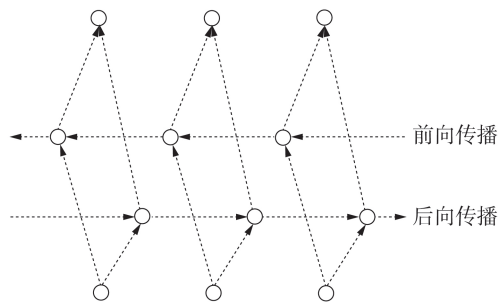


图1 BiGRU结构图

Fig.1 Structure diagram of BiGRU

2.3 正则化贪心森林算法

正则化贪心森林RGF(Regularized Greedy Forest)算法^[17]于2014年被提出,目的是为了应对传统梯度提升决策树GBDT(Gradient Boosting Decision Tree)收敛速度较慢、容易过拟合等不足之处。RGF算法与GBDT不同,在每次迭代时需对整个决策森林进行学习,并在新增树后对全局进行参数优化,同时引入正则项来防止过拟合现象。

2.4 多目标Stacking集成学习

传统Stacking集成学习算法通过将不同的机器学习算法融合在一起,并通过K折交叉验证来进行机器学习器训练集的划分,再将机器学习器训练后得到的结果作为元学习器的输入,从而得到最终预测结果。具体预测流程如图2所示。

从图2中可以看出,传统Stacking集成学习模型是对单一任务进行预测的,预测结果也是单一结果输出的。考虑到IES特有的结构特性,电、热、冷负荷之间有较强的耦合关系,将多元负荷单独进行预测往往会忽略多元负荷之间的关联性,从而影响预测精度。针对这类问题,将多目标回归的思想融入Stacking集成学习模型中,充分考虑电、热、冷负荷之间的耦合性,从而进行多元负荷协同预测。多目标Stacking集成学习预测模型主要分为以下2个阶段。

1)第一阶段:基于传统Stacking集成学习模型

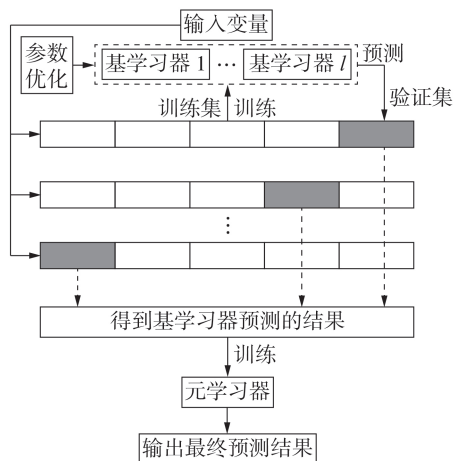


图2 传统Stacking集成学习模型预测流程图

Fig.2 Forecasting flowchart of traditional Stacking ensemble learning model

的多元负荷单独预测。

在第一阶段中,仍使用传统Stacking集成学习模型的基学习器与元学习器训练模式,对电、热、冷负荷分别进行预测。首先,采用5折交叉验证划分训练集,并在第一层基学习器算法选取轻量级梯度提升机LightGBM(Light Gradient Boosting Machine)、极限梯度提升XGBoost(eXtreme Gradient Boosting)、随机森林RF(Random Forest)这类预测性能较好的决策树算法。同时考虑到多元负荷的时序特点,引入BiGRU,在第一层选取不同类型的基学习器,目的是使模型能从不同的空间结构去训练数据,从而达到更高的预测精度;在第二层元学习器模型选取LightGBM模型,以增强模型的泛化能力。

2)第二阶段:扩充原始输入变量。

对原始输入变量进行扩充,将第一阶段单独预测得到的电、热、冷负荷结果扩充至原始输入空间,组成新的输入变量,并且在传统Stacking集成学习模型基础之上再引入一层多目标回归模型。为了防止出现过拟合现象,本文选取RGF算法作为最终多元负荷协同预测模型,并同时输出电、热、冷负荷预测结果。

3 预测流程

本文提出的基于MIC和多目标Stacking集成学习的多元负荷预测方法的具体预测流程图如图3所示。

4 算例分析

本文采用的实际数据为美国亚利桑那州立大学Tempe校区2019年1月1日至2020年2月29日的IES电、热、冷负荷数据,采样间隔为1h,负荷数据来源于该校CAMPUS METABOLISM项目网络数据

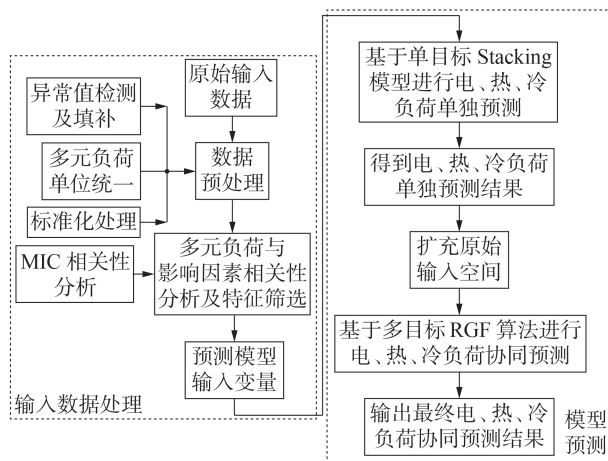


图3 所提预测方法流程图

Fig.3 Flowchart of proposed forecasting method

库,天气数据来源于Dark SKY网站。该校区拥有288座建筑物、52255名师生,同时拥有冷热电联供CCHP(Combined Cooling, Heating and Power)、电锅炉、燃气锅炉以及电转气P2G(Power-to-Gas)等能量转换设备。

为了验证本文模型在一年不同季节下的适用性,根据气象学定义的四季,对美国亚利桑那州立大学Tempe校区的IES负荷数据进行四季划分,并以夏季和冬季为例,选取2019年6月1日至8月27日的电、热、冷负荷数据及天气因素作为夏季预测的训练集样本,8月30日作为夏季典型工作日预测样本,8月31日作为夏季典型周末预测样本;选取2019年12月1日至2020年2月22日作为冬季预测训练集样本,2月25日作为冬季典型工作日预测样本,2月24日作为冬季典型周末预测样本。

本文使用的硬件设施为Inter Core i5-9300H处理器、NVIDIA GTX1660Ti 6 G GDDR6独立显卡,软件平台为Python3.7.1和MATLAB R2020a。

4.1 输入变量与数据预处理

1)考虑到在实际负荷数据记录工程中,人工失误会导致某一时刻负荷数据出现异常情况,因此本文采用 3σ 准则^[18]来进行异常值检测,将检测得到的异常值当作缺失值处理,并采用K-最近邻KNN(K-Nearest Neighbor)算法进行缺失值填补。

2)由于采集到的电、热、冷负荷数据单位不统一,为了不影响模型的鲁棒性,将采用不同单位的负荷进行标准化处理,先根据能源转换计算公式^[19]全部将单位统一为kW,再与其他天气因素数据进行标准化处理,标准化计算公式如下:

$$x^* = \frac{x - \mu}{\sigma} \quad (5)$$

式中: μ 为原始数据样本均值; σ 为原始数据样本的标准差。

3)考虑到多元负荷的时序特性,在进行多元负荷协同预测时,不仅要考虑预测日 $t-1$ 时刻和预测日前一天 $t-1/t/t+1$ 时刻的实际温度、湿度等天气因素,还要考虑预测日前一天 $t-1/t/t+1$ 时刻的电、热、冷负荷数据。因此,结合上文多元负荷及天气因素的相关性分析,本文选取的输入变量如附录A表A2所示。

4.2 评价指标

本文使用平均绝对百分比误差 E_{MAPE} 、平均精度 M_A 和加权平均准确度 W_{MA} 来评价预测结果,在 W_{MA} 的权重分配中,考虑到电力负荷的主导地位,将电、热、冷负荷的权重值分别设为0.6、0.2、0.2,计算公式分别如下:

$$E_{\text{MAPE}} = \frac{1}{n} \sum_{i=1}^n \left| \frac{x(i)-y(i)}{x(i)} \right| \times 100\% \quad (6)$$

$$M_A = 1 - E_{\text{MAPE}} \quad (7)$$

$$W_{\text{MA}} = \alpha_1 (1 - E_{\text{MAPE}_1}) + \alpha_2 (1 - E_{\text{MAPE}_2}) + \dots + \alpha_k (1 - E_{\text{MAPE}_k}) \quad (8)$$

式中: n 为样本数; $x(i)$ 为 i 时刻负荷实际值; $y(i)$ 为 i 时刻负荷预测值; $\alpha_j (j=1, 2, \dots, k)$ 为第 j 类负荷的权重值; $E_{\text{MAPE}_j} (j=1, 2, \dots, k)$ 为第 j 类负荷的平均绝对百分比误差值; k 为负荷类型的数量。 E_{MAPE} 值在0~100%之间, E_{MAPE} 值越小,预测误差越低; M_A 和 W_{MA} 值在0~100%之间, M_A 和 W_{MA} 值越大,预测精度越高。

4.3 参数选取

本文采用贝叶斯优化算法^[20]对各模型进行参数寻优,各模型参数选取如表3所示。

4.4 预测结果对比分析

4.4.1 考虑负荷耦合形态指标前后的对比分析

为了验证本文所提出的负荷耦合形态指标能够提升模型的可预测性,设置以下2种案例进行对比分析:案例1考虑负荷耦合形态指标,采用本文模型进行电、热、冷负荷协同预测;案例2不考虑负荷耦合形态指标,采用本文模型进行电、热、冷负荷协同预测。夏季和冬季考虑负荷耦合形态指标前后的电、热、冷负荷预测精度分别如表4和表5所示。

表4 夏季考虑负荷耦合形态指标前后的预测精度结果

Table 4 Forecasting accuracy results before and after considering load coupling morphological indexes in summer

案例	指标	工作日			周末		
		电负荷	热负荷	冷负荷	电负荷	热负荷	冷负荷
1	$E_{\text{MAPE}} / \%$	0.81	1.24	0.68	0.33	0.38	0.62
2		2.75	2.46	3.48	2.03	1.80	2.62
1	$W_{\text{MA}} / \%$	99.13			99.60		
2		97.16			97.90		

表5 冬季考虑负荷耦合形态指标前后的预测精度结果

Table 5 Forecasting accuracy results before and after considering load coupling morphological indexes in winter

案例	指标	工作日			周末		
		电负荷	热负荷	冷负荷	电负荷	热负荷	冷负荷
1	$E_{\text{MAPE}} / \%$	1.00	0.58	1.03	0.93	1.04	0.71
2		3.51	4.31	6.40	1.99	3.68	3.03
1	$W_{\text{MA}} / \%$	99.08			99.09		
2		95.75			97.46		

从表4和表5中可知,案例1的预测精度优于案例2。在夏季典型工作日场景下,案例1的电、热、冷负荷预测误差 E_{MAPE} 相较于案例2分别减少了1.94%、1.22%、2.80%,预测精度 W_{MA} 提升了1.97%;在冬季典型工作日场景下,案例1的电、热、冷负荷预测误差 E_{MAPE} 相较于案例2分别减少了2.51%、3.73%、5.37%,预测精度 W_{MA} 提升了3.33%。这表明本文所提出的负荷耦合形态指标增强了模型的可预测性。

4.4.2 多元负荷协同预测与单一负荷预测对比分析

为了验证多元负荷协同预测在不同时间步长下的普适性和优越性,将预测步长分别取1、6、12 h,并与单一负荷预测进行详细对比分析,以夏季为例,对比结果如表6所示。从表中可知,随着预测步长的增大,多元负荷协同预测和单一负荷预测的预测误差 E_{MAPE} 都逐渐变大,但在不同预测步长下,多元负荷协同预测模型都始终优于单一负荷预测模型。在典型工作日场景下,当预测步长为12 h时,多元负荷预测误差区间为[1.73%, 3.74%],而单一负荷预

表3 参数选取

Table 3 Selection of parameters

算法	参数取值
RGF	最大叶子数为400,算法为最小惩罚、同层级零和约束的RGF算法,损失函数为平方损失函数
LightGBM	学习率为0.078,最大树深度为7,最大叶子数为15,提升算法为GBDT,分裂的最小损失减少值为0.37,最小样本权重和为0.0016,L1正则项为0,L2正则项为0.48,提升次数为200
XGBoost	学习率为0.0756,最大树深度为4,树的数目为100,训练样本采样比例为1,L2正则项为0.5,提升算法采用树模型,特征采样的比例为1
RF	树的数目为100,分支所需的最小样本数为2,叶子节点存在所需最小样本数为1
BiGRU	神经元数量为64,全连接层层数为3层,优化器为Adam算法,学习率采用回调函数自主调整,迭代次数为100,批量数量为72

表6 不同预测步长下的多元负荷协同预测与单一负荷预测精度对比结果

Table 6 Comparison results of forecasting accuracy between multivariate load collaborative forecasting and single-load forecasting in summer with different forecasting step sizes

预测模式	日类型	预测步长 / h	$E_{MAPE} / \%$		
			电负荷	热负荷	冷负荷
多元负荷 协同预测	工作日	1	0.81	0.68	1.24
		6	1.65	2.15	2.97
		12	1.73	2.38	3.74
	周末	1	0.33	0.38	0.62
		6	0.91	1.21	1.49
		12	1.39	1.76	2.79
单一负荷预测	工作日	1	2.84	2.37	2.87
		6	3.69	3.63	3.68
		12	5.33	6.04	7.24
	周末	1	2.50	2.24	2.61
		6	3.58	3.01	3.20
		12	6.75	6.76	8.75

测误差区间为 $[5.33\%, 7.24\%]$,误差明显增大。这是因为多元负荷间具有较强的耦合关系,而单一负荷预测忽略了多元负荷间的耦合关系。

在同一预测步长下多元负荷协同预测所花费的时间为35.38 s,而各负荷单独预测所花费的时间分别为电负荷32.92 s、热负荷33.03 s、冷负荷32.00 s,多元负荷协同预测的训练时间仅为单一负荷预测训练时间总和的1/3左右。以上分析充分验证了多元负荷协同预测模型的优越性。

4.4.3 多模型融合与单一模型对比分析

为了验证本文采用的多模型融合的协同预测模型在预测精度方面的优势,将本文模型与集成学习中各单一模型算法进行对比,进行对比的单一模型算法包括RGF、XGBoost、BiGRU算法,为了保证实验的公平性,仍采用相同数据进行分析,并以夏季为例,具体实验结果如附录A图A3、A4和表7所示。

表7 多模型融合与单一模型预测对比结果

Table 7 Comparison of forecasting results between multi-model fusion and single model

指标	模型	工作日			周末		
		电负荷	热负荷	冷负荷	电负荷	热负荷	冷负荷
$E_{MAPE} / \%$	RGF	1.86	2.04	1.99	1.32	1.13	1.56
	XGBoost	1.95	2.10	1.90	1.48	1.29	1.56
	BiGRU	1.39	1.32	1.40	1.20	1.25	1.20
	本文模型	0.81	0.68	1.24	0.33	0.38	0.62
$W_{MA} / \%$	RGF	98.08	98.08	98.08	98.67	98.67	98.67
	XGBoost	98.03	98.03	98.03	98.54	98.54	98.54
	BiGRU	98.62	98.62	98.62	98.79	98.79	98.79
	本文模型	99.13	99.13	99.13	99.60	99.60	99.60

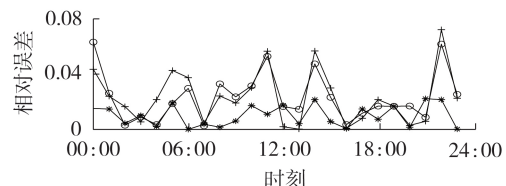
从表7中可知,在不同日类型场景下,本文所提出的多模型融合的协同预测模型在预测精度 W_{MA} 上

均优于各单一预测模型。在典型工作日场景下,本文模型的电、热、冷负荷的预测误差 E_{MAPE} 分别为0.81%、0.68%、1.24%,而基于XGBoost的模型预测结果最差, E_{MAPE} 分别达到了1.95%、2.10%、1.90%,分别为本文模型预测误差的2.41倍、3.09倍、1.53倍;在典型周末场景下,本文模型的电、热、冷负荷的预测误差 E_{MAPE} 分别为0.33%、0.38%、0.62%,而基于XGBoost的模型预测结果仍为最差, E_{MAPE} 分别达到了1.48%、1.29%、1.56%,分别为本文模型预测误差的4.48倍、3.39倍、2.52倍。通过与各单一预测模型进行对比,再次验证了多模型融合的协同预测的优越性。

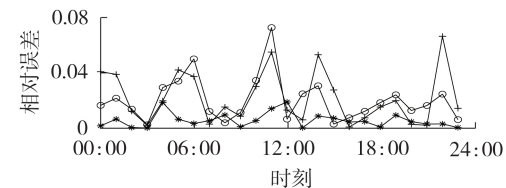
4.4.4 不同多元负荷预测模型对比分析

为了进一步突出本文所提出的多元负荷协同预测模型的优势,选取多目标回归和多任务学习算法来进行对比分析。其中多目标回归选取的学习器为LightGBM算法,多任务学习算法的共享层选取BP神经网络来进行共享学习,以冬季为例,对比实验结果如图4、5和表8所示。

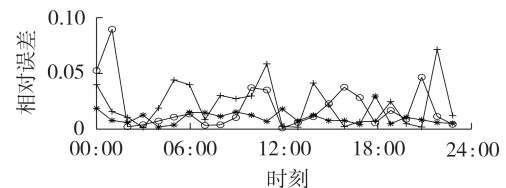
从图4和图5中可知,在冬季典型工作日场景下,电、热、冷负荷预测相对误差最大值主要集中在在工作时段,即08:00—19:00时段;在冬季典型周末场景下,电、热、冷负荷预测相对误差最大值主要出现



(a) 电负荷



(b) 热负荷



(c) 冷负荷

— 本文模型, - - 多目标回归, ... 多任务学习算法

图4 冬季典型工作日多元负荷预测模型相对误差对比图

Fig.4 Relative error comparison diagram of multivariate load forecasting model on typical working days in winter

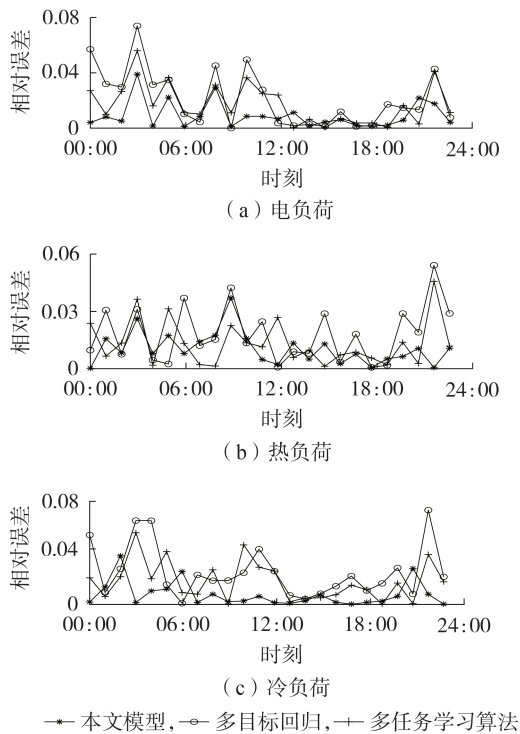


图5 冬季典型周末多元负荷预测模型相对误差对比图

Fig.5 Relative error comparison diagram of multivariate load forecasting model on typical weekends in winter

表8 不同多元负荷预测模型的预测结果
Table 8 Forecasting results of different multivariate load forecasting models

指标	模型	工作日			周末		
		电负荷	热负荷	冷负荷	电负荷	热负荷	冷负荷
$E_{MAPE} / \%$	多目标回归	2.33	2.04	1.96	2.15	1.80	2.34
	多任务学习算法	2.36	2.23	2.18	1.73	1.33	1.67
	本文模型	1.00	0.58	1.03	0.93	1.04	0.71
$W_{MA} / \%$	多目标回归	97.80	97.80	97.80	97.88	97.88	97.88
	多任务学习算法	97.70	97.70	97.70	98.36	98.36	98.36
	本文模型	99.08	99.08	99.08	99.09	99.09	99.09

在00:00—05:00、20:00—24:00时段。本文所提出的多元负荷协同预测模型总体误差波动较为平稳,而其他2种多元负荷预测模型误差波动则较为剧烈,主要有以下3点原因。

1) 本文采用的多模型融合下的多元负荷协同预测模型,通过选取不同类型的算法来达到互补的作用,其中树模型算法可以很好地学习多元负荷及天气因素之间的非线性关系;BiGRU可以从不同方向全面地对数据进行学习;RGF算法可以在单一负荷误差波动较为剧烈时,能从不同数据结构空间去学习多元负荷间的耦合信息,防止发生过拟合现象。

2) 多目标LightGBM算法不具有处理时间序列的功能,因此在处理时序特点较为明显的数据时具有明显的劣势。

3) 多任务学习算法虽然在共享层学习了多元负荷间的耦合关系,但BP神经网络的网络层较为简单,无法更深层次地挖掘多元负荷间的耦合信息。

从表8中可知,在预测精度方面,本文采用的协同预测模型的预测误差 E_{MAPE} 在 $[0.5\%, 1.1\%]$ 区间之内,而其他2种多元负荷预测模型预测误差 E_{MAPE} 在 $[1.3\%, 2.4\%]$ 范围内,可见本文模型明显优于其他2种模型。对于预测精度 W_{MA} 值,本文模型也优于其他2种模型。

综上所述,本文所采用的多元负荷协同预测模型,在Stacking集成学习框架下进行多模型融合,同时引入BiGRU神经网络和RGF算法作为学习器模型,前者可以从不同方向对数据进行全面的学习,后者能防止过拟合现象的发生。为了充分利用多元负荷间的耦合关系,将多目标回归思想与多模型融合相结合来实现多元负荷协同预测。

5 结论

本文考虑IES中电、热、冷负荷之间的耦合关系,提出负荷耦合形态指标来深度挖掘多元负荷间的耦合性,并引入MIC的相关性分析来进行特征筛选,再将多目标回归与Stacking集成学习相结合,建立电、热、冷负荷协同预测模型,最后通过实际算例得出以下结论:

1) 在考虑负荷耦合形态指标的情况下,协同预测模型较好地利用了多元负荷间的耦合信息,本文模型的预测精度得到了明显提升;

2) 本文模型解决了传统Stacking集成学习模型只能进行单一目标预测的问题,并在不同预测步长下证明了多元负荷协同预测相较于单一负荷预测受到预测步长的影响较小,预测精度和训练时间明显优于单一负荷预测;

3) 在Stacking集成学习模型中,分别引入RGF算法和BiGRU,前者能够防止过拟合现象的发生,后者能够从不同的数据方向学习信息,并且与其他多输出预测模型相比,其具有更高的预测精度且相对误差曲线波动较为平稳,模型的稳定性更强。

在之后的研究中,笔者将进一步考虑多元负荷时序耦合特性,挖掘更深层次的耦合特征,从而进一步提升预测精度。

附录见本刊网络版(<http://www.epae.cn>)。

参考文献:

- [1] SUGANTHI L, SAMUEL A A. Energy models for demand forecasting-a review[J]. Renewable and Sustainable Energy Reviews, 2012, 16(2): 1223-1240.
- [2] 程浩忠, 胡泉, 王莉, 等. 区域综合能源系统规划研究综述[J]. 电力系统自动化, 2019, 43(7): 2-13.

- CHENG Haozhong, HU Xiao, WANG Li, et al. Review on research of regional integrated energy system planning[J]. Automation of Electric Power Systems, 2019, 43(7): 2-13.
- [3] 熊文, 刘育权, 苏万煌, 等. 考虑多能互补的区域综合能源系统多种储能优化配置[J]. 电力自动化设备, 2019, 39(1): 118-126.
- XIONG Wen, LIU Yuquan, SU Wanhua, et al. Optimal configuration of multi-energy storage in regional integrated energy system considering multi-energy complementation[J]. Electric Power Automation Equipment, 2019, 39(1): 118-126.
- [4] YAN Y D, LI S F, LI W Q, et al. Power load probability density forecasting using Gaussian process quantile regression[J]. Applied Energy, 2018, 213: 499-509.
- [5] 于晓娟, 齐先硕, 顾吉浩, 等. 基于混合算法优化支持向量机的供热负荷预测模型[J]. 河北工业大学学报, 2019, 48(5): 39-46.
- YU Xiaojuan, QI Xianshuo, GU Jihao, et al. Heat load prediction model based on support vector machine optimized by hybrid algorithm[J]. Journal of Hebei University of Technology, 2019, 48(5): 39-46.
- [6] 于军琪, 井文强, 赵安军, 等. 基于改进 PSO-BP 算法的冷负荷预测模型[J]. 系统仿真实报, 2021, 33(1): 54-61.
- YU Junqi, JING Wenqiang, ZHAO Anjun, et al. Cold load prediction model based on improved PSO-BP algorithm[J]. Journal of System Simulation, 2021, 33(1): 54-61.
- [7] 赵峰, 孙波, 张承慧. 基于多变量相空间重构和卡尔曼滤波的冷热电联供系统负荷预测方法[J]. 中国电机工程学报, 2016, 36(2): 399-406.
- ZHAO Feng, SUN Bo, ZHANG Chenghui. Cooling, heating and electrical load forecasting method for CCHP system based on multivariate phase space reconstruction and Kalman filter[J]. Proceedings of the CSEE, 2016, 36(2): 399-406.
- [8] 孙庆凯, 王小君, 张义志, 等. 基于 LSTM 和多任务学习的综合能源系统多元负荷预测[J]. 电力系统自动化, 2021, 45(5): 63-70.
- SUN Qingkai, WANG Xiaojun, ZHANG Yizhi, et al. Multiple load prediction of integrated energy system based on long short-term memory and multi-task learning[J]. Automation of Electric Power Systems, 2021, 45(5): 63-70.
- [9] 史佳琪, 谭涛, 郭经, 等. 基于深度结构多任务学习的园区型综合能源系统多元负荷预测[J]. 电网技术, 2018, 42(3): 698-707.
- SHI Jiaqi, TAN Tao, GUO Jing, et al. Multi-task learning based on deep architecture for various types of load forecasting in regional energy system integration[J]. Power System Technology, 2018, 42(3): 698-707.
- [10] 翟晶晶, 吴晓蓓, 王力立. 基于径向基函数神经网络的综合能源系统多元负荷短期预测[J]. 电力需求侧管理, 2019, 21(4): 23-27, 34.
- ZHAI Jingjing, WU Xiaobei, WANG Lili. Multiple short-term load forecasting in integrated energy system based on RBF-NN model[J]. Power Demand Side Management, 2019, 21(4): 23-27, 34.
- [11] 陈锦鹏, 胡志坚, 陈伟楠, 等. 二次模态分解组合 DBiLSTM-MLR 的综合能源系统负荷预测[J]. 电力系统自动化, 2021, 45(13): 85-94.
- CHEN Jinpeng, HU Zhijian, CHEN Weinan, et al. Load prediction of integrated energy system based on combination of quadratic modal decomposition and deep bidirectional long short-term memory and multiple linear regression[J]. Automation of Electric Power Systems, 2021, 45(13): 85-94.
- [12] 杨国清, 张凯, 王德意, 等. 基于包络线聚类的多模融合超短期光伏功率预测算法[J]. 电力自动化设备, 2021, 41(2): 39-46.
- YANG Guoqing, ZHANG Kai, WANG Deyi, et al. Multi-mode fusion ultra-short-term photovoltaic power prediction algorithm based on envelope clustering[J]. Electric Power Automation Equipment, 2021, 41(2): 39-46.
- [13] 黄子硕, 何桂雄, 闫华光, 等. 园区级综合能源系统优化模型功能综述及展望[J]. 电力自动化设备, 2020, 40(1): 10-18.
- HUANG Zishuo, HE Guixiong, YAN Huaguang, et al. Overview and prospect of optimization model function for community-scale integrated energy system[J]. Electric Power Automation Equipment, 2020, 40(1): 10-18.
- [14] 赵雅雪, 王旭, 蒋传文, 等. 基于最大信息系数相关性分析和改进多层级门控 LSTM 的短期电价预测方法[J]. 中国电机工程学报, 2021, 41(1): 135-146, 404.
- ZHAO Yaxue, WANG Xu, JIANG Chuanwen, et al. A novel short-term electricity price forecasting method based on correlation analysis with the maximal information coefficient and modified multi-hierarchy gated LSTM[J]. Proceedings of the CSEE, 2021, 41(1): 135-146, 404.
- [15] 王进, 高选人, 张睿, 等. 结合目标特定特征和目标相关性的多目标回归[J]. 电子学报, 2020, 48(11): 2092-2100.
- WANG Jin, GAO Xuanren, ZHANG Rui, et al. Multi-target regression via specific features and inter-target correlations[J]. Acta Electronica Sinica, 2020, 48(11): 2092-2100.
- [16] 张立鹏, 毕凤荣, 程建刚, 等. 基于注意力 BiGRU 的机械故障诊断方法研究[J]. 振动与冲击, 2021, 40(5): 113-118.
- ZHANG Lipeng, BI Fengrong, CHENG Jiangang, et al. Mechanical fault diagnosis method based on attention BiGRU[J]. Journal of Vibration and Shock, 2021, 40(5): 113-118.
- [17] 黄明增, 文云峰, 汪荣华, 等. 基于正则化贪心森林的多维频率指标智能化预测方法[J]. 电力建设, 2020, 41(9): 124-131.
- HUANG Mingzeng, WEN Yunfeng, WANG Ronghua, et al. Intelligent prediction for multi-dimensional frequency indicators based on regularized greedy forests[J]. Electric Power Construction, 2020, 41(9): 124-131.
- [18] 张振慧, 张正江, 胡桂廷, 等. 基于拉依达准则与线性拟合的改进型无迹卡尔曼滤波粗大误差补偿算法[J]. 计算机测量与控制, 2019, 27(11): 153-156, 162.
- ZHANG Zhenhui, ZHANG Zhengjiang, HU Guiting, et al. Improved unscented Kalman filter based on gross error compensation algorithm with pauta criterion and linear fitting[J]. Computer Measurement & Control, 2019, 27(11): 153-156, 162.
- [19] 张涌新, 沈弘, 马静. 综合能源系统负荷特性分析及应用研究[J]. 电力建设, 2018, 39(9): 18-29.
- ZHANG Yongxin, SHEN Hong, MA Jing. Load characteristic analysis and application research on integrated energy system[J]. Electric Power Construction, 2018, 39(9): 18-29.
- [20] 邓帅. 基于改进贝叶斯优化算法的 CNN 超参数优化方法[J]. 计算机应用研究, 2019, 36(7): 1984-1987.
- DENG Shuai. Hyper-parameter optimization of CNN based on improved Bayesian optimization algorithm[J]. Application Research of Computers, 2019, 36(7): 1984-1987.

作者简介:



崔树银

崔树银(1968—),男,副教授,博士,主要研究方向为综合能源系统优化(E-mail: shiyincui@126.com);

汪昕杰(1996—),男,硕士研究生,主要研究方向为综合能源系统多元负荷预测(E-mail: maochoi00@126.com)。

(编辑 李玮)

(下转第81页 continued on page 81)

构及其应用[J]. 电工技术学报, 2014, 29(8): 173-179.

ZHANG Jianpo, ZHAO Chengyong, SUN Haifeng, et al. Improved topology of modular multilevel converter and application[J]. Transactions of China Electrotechnical Society, 2014, 29(8): 173-179.

作者简介:

束洪春(1961—),男,教授,博士研究生导师,博士,从事电力系统新型继电保护与故障测距、故障录波、数字信号处



束洪春

理等方面的教研工作(E-mail: kmshe@sina.com.cn);

王文韬(1997—),男,硕士研究生,主要研究方向为柔性直流输电技术(E-mail: 1574261522@qq.com);

江耀曦(1975—),女,讲师,博士研究生,通信作者,主要研究方向为柔性直流输电技术(E-mail: 1765540810@qq.com)。

(编辑 王欣竹)

Novel MMC sub-module topology with DC fault clearing capability

SHU Hongchun¹, WANG Wentao¹, JIANG Yaoxi², SHAO Zongxue¹, WANG Rui¹, LIAO Mengli¹

(1. Faculty of Electric Power Engineering, Kunming University of Science and Technology, Kunming 650500, China;

2. Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China)

Abstract: When short circuit fault occurs on DC line of MMC-HVDC (High Voltage Direct Circuit based on Modular Multilevel Converter) system, the DC fault current cannot be blocked by traditional HBSM (Half-Bridge Sub-Module). In this regard, a FLHBSM (novel Five-Level Half-Bridge Sub-Module) topology is proposed. FLHBSM has the same control structure as HBSM with the addition of DC fault clearing capability. Compared with other sub-module topologies with DC fault clearing capability, FLHBSM reduces the number of switching devices and the area occupied by the converter station, reduces the loss and provides high control flexibility, and not only improves the economy of MMC, but also enhances its operation reliability. Meanwhile, FLHBSM can be compatible with the capacitor voltage balancing strategies suitable for HBSM, and can choose better capacitor voltage balancing strategy according to the actual needs. A simulation model of bi-terminal ± 100 kV 21-level MMC-HVDC system based on FLHBSM is built on MATLAB/Simulink platform, which verifies the effectiveness of DC fault clearing capacity by using FLHBSM.

Key words: MMC; sub-module topology; DC fault clearing capability; voltage balancing strategy; economy

(上接第39页 continued from page 39)

Multivariate load forecasting in integrated energy system based on maximal information coefficient and multi-objective Stacking ensemble learning

CUI Shuyin, WANG Xinjie

(School of Economics and Management, Shanghai University of Electric Power, Shanghai 200090, China)

Abstract: The accurate multivariate load forecasting plays an important role in energy dispatching and operation planning of integrated energy system. Traditional methods for forecasting electrical, thermal, and cooling loads separately ignore the coupling relationship between multivariate loads. In order to solve this problem, a multivariate load collaborative forecasting model based on multi-objective Stacking ensemble learning is proposed. The maximal information coefficient is introduced to analyze the correlation between multivariate load and weather factors, and the load coupling morphological index is proposed to deeply explore the coupling relationship between multivariate loads. Then, the multi-objective regression and Stacking ensemble learning model are combined to establish the multivariate load cooperative forecasting model. A practical example is given to verify the effectiveness of the proposed model, and the results of numerical examples show that the forecasting accuracy of the proposed model is higher than that of other forecasting models.

Key words: multi-objective regression; Stacking ensemble learning; integrated energy system; maximal information coefficient; regularized greedy forest algorithm

附录 A

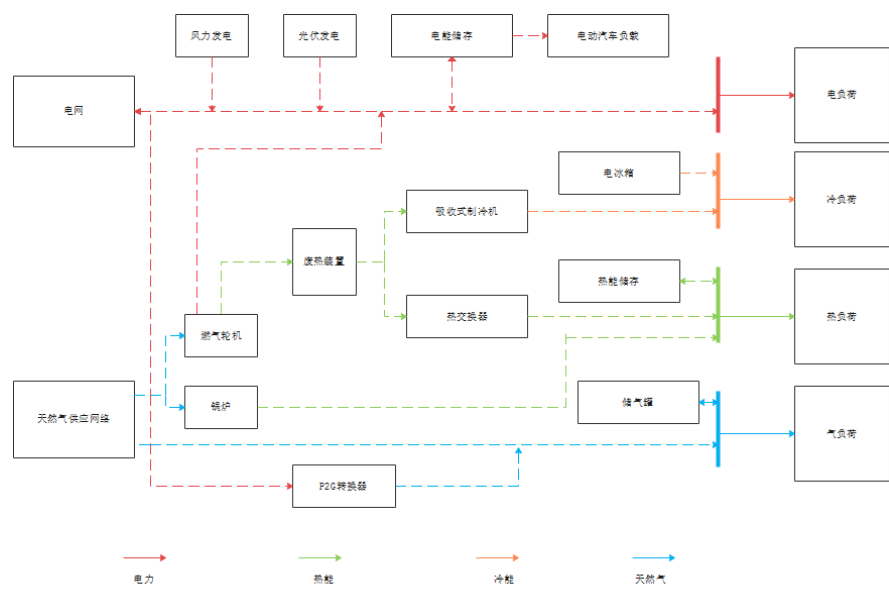


图 A1 IES 交互图
Fig.A1 Interactive diagram of IES

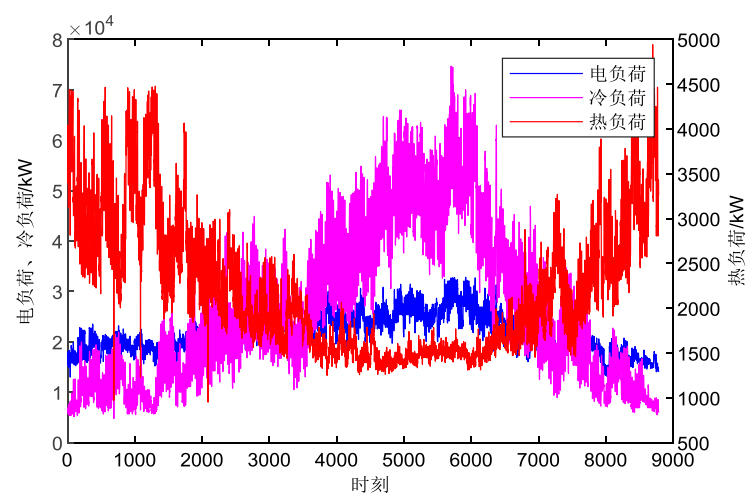


图 A2 多元负荷曲线
Fig.A2 Multivariate load curves

表 A1 不同季节下的多元负荷与天气因素间的 MIC 值
Table A1 MIC values between multivariate loads and weather factors in different seasons

影响因素	多元负荷	春	夏	秋	冬
温度	电	0.3313	0.6447	0.4929	0.6270
	热	0.4471	0.2869	0.6212	0.8889
	冷	0.7071	0.8559	0.7270	0.3383
湿度	电	0.2108	0.2989	0.2160	0.2262
	热	0.1666	0.1163	0.1504	0.2885
	冷	0.4576	0.2460	0.2900	0.2007
大气压	电	0.1274	0.2030	0.2653	0.2073
	热	0.2340	0.0917	0.2606	0.1804
	冷	0.2011	0.1989	0.3253	0.2115
风速	电	0.1095	0.2100	0.2198	0.2029
	热	0.0911	0.1539	0.1641	0.1543
	冷	0.1189	0.2353	0.2312	0.2089
降雨量	电	0.0891	0.0487	0.2328	0.2863
	热	0.0614	0.0176	0.1396	0.1528
	冷	0.0942	0.0487	0.2301	0.1932

表 A2 输入变量设置
Table A2 Input variable settings

输入变量（夏季）	输入变量（冬季）
预测日前一天 $t-1/t/t+1$ 时刻电负荷	预测日前一天 $t-1/t/t+1$ 时刻电负荷
预测日前一天 $t-1/t/t+1$ 时刻热负荷	预测日前一天 $t-1/t/t+1$ 时刻热负荷
预测日前一天 $t-1/t/t+1$ 时刻冷负荷	预测日前一天 $t-1/t/t+1$ 时刻冷负荷
预测日 $t-1$ 时刻温度、湿度、风速	预测日 $t-1$ 时刻温度、湿度、风速、大气压、降雨量
预测日前一天 $t-1/t/t+1$ 时刻温度、湿度、风速	预测日前一天 $t-1/t/t+1$ 时刻温度、湿度、风速、大气压、降雨量
负荷耦合形态指标	负荷耦合形态指标

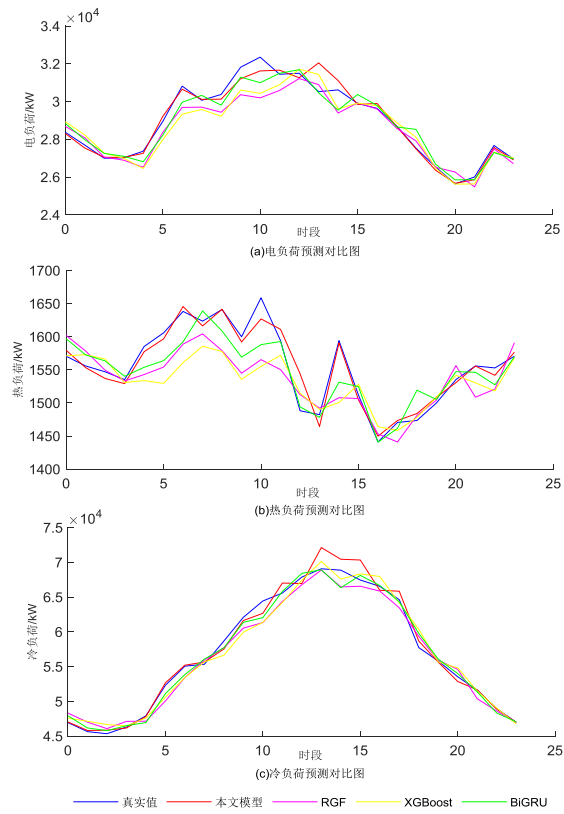


图 A3 典型工作日场景下的多模型融合与单一模型预测对比
 Fig.A3 Comparison of multi-model fusion and single-model prediction in a typical workday scenario

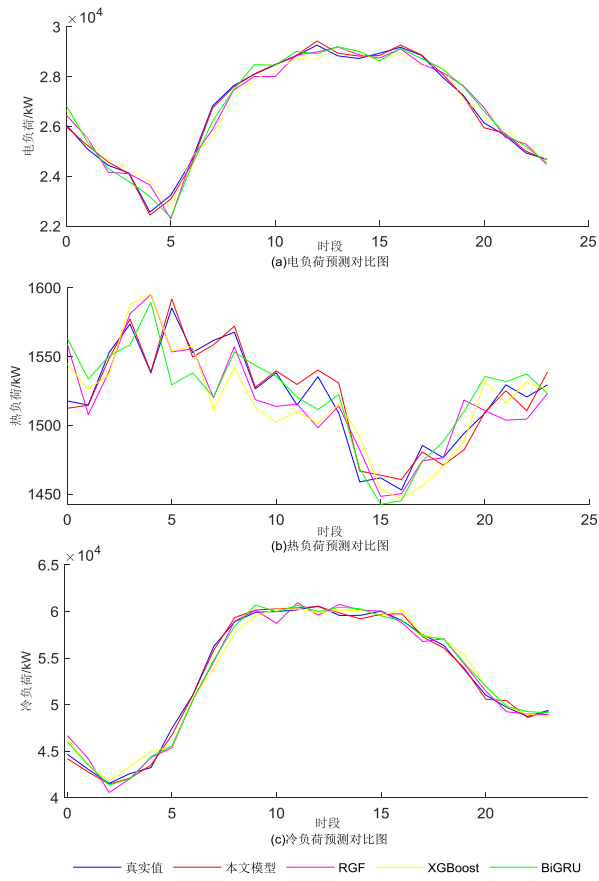


图 A4 典型周末场景下的多模型融合与单一模型预测对比
 Fig.A4 Comparison of multi-model fusion and single-model prediction in a typical weekend scenario