

基于改进双重压缩和激励与多头特征注意力机制的电-热负荷协同预测

余强¹, 韩静娴¹, 杨子梁¹, 宋济东¹, 杨德昌¹, 齐海杰², 于芑³

(1. 中国农业大学信息与电气工程学院, 北京 100083; 2. 国网智能电网研究院有限公司, 北京 102200;
3. 国网山东省电力公司电力科学研究院, 山东 济南 250003)

摘要:综合能源系统中负荷多样且存在耦合,为提升负荷预测精度,提出一种基于改进双重注意力机制的分组卷积神经网络-门控循环单元短期电-热负荷协同预测模型。通过改进的压缩和激励注意力为各输入通道加权,再对其进行分组卷积;利用多头特征注意力对卷积结果进行赋权,并利用输入门控循环单元模型对负荷进行预测。算例仿真结果表明,所提模型的平均绝对百分比误差均低于3%。

关键词:综合能源系统;负荷预测;分组卷积神经网络;门控循环单元;改进的压缩和激励注意力机制;多头特征注意力机制

中图分类号:TK01*9

文献标志码:A

DOI:10.16081/j.epae.202409030

0 引言

在系统规划和运行中,综合能源系统可实现不同能源系统的优势互补,但是多种能源的统筹规划必须建立在对供给能力与负荷需求的全面调查与梳理之上。与传统电力负荷不同,综合能源系统中负荷的多样性、不同种类负荷之间的耦合特性以及用户行为的影响,使得负荷消耗特性愈发复杂,精确的负荷预测结果尤为重要。

早期人们采用传统的时间序列法^[1]、统计学法^[2-3]等来对负荷进行预测,随着研究手段的进步,学者们尝试采用传统预测方法以外的技术(例如人工神经网络^[4-5]、专家系统技术^[6]、支持向量机^[7-8]等)来对负荷进行预测,也取得了良好的预测效果。近年来,随着长短期记忆神经网络^[9]等方法的提出,学者们使用这些方法进行负荷预测,结果表明,这些方法的预测效果基本上优于传统的时间序列法,但通常单一模型存在缺陷,为了提高模型预测精度,不少学者尝试运用组合模型。文献^[10]采用卷积神经网络(convolutional neural network, CNN)与长短期记忆神经网络相结合的融合模型,建立电负荷预测模型,结果表明,该模型精度高于单一模型;文献^[11]提出一种基于特征筛选的CNN-双向长短期记忆(bi-

directional long short-term memory, BiLSTM)网络的短期电力负荷预测方法,利用CNN提取特征向量,在高维空间中形成高维预测特征向量值,并采用BiLSTM网络进行双向循环训练,结果表明,该组合方法能提升多维电力负荷预测精度;文献^[12]将历史负荷数据分解成多个子序列负荷,将天气、日期、工作日类型等因素作为输入特征,采用CNN-BiLSTM预测后得到各个子序列负荷的预测值,并将其相加重构实际负荷预测曲线,结果表明,连续周电负荷平均绝对百分比误差(mean absolute percentage error, MAPE)低于4%;文献^[13-15]利用模态分解方法将历史负荷数据分解成多个负荷序列,对每个子序列运用不同的模型分别进行预测,并将结果相加获得实际值;文献^[16]将数据输入CNN,利用特征向量构造时间序列形式,并将其输入双向门控循环单元(gated recurrent unit, GRU)网络中,完成短期负荷预测;文献^[17]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。上述运用组合模型进行预测的研究,利用对输入变量进行筛选或者分解重构的方法提取输入变量的重要信息,并将其输入神经网络进行训练,但神经网络对输入变量的无差别对待可能会引起预测误差。

为了减小上述误差,部分关于负荷预测的研究更注重挖掘特征变量的关键信息。文献^[18]在模型中添加注意力机制,对住宅区电负荷进行预测;文献^[19]进行基于时间模式注意力机制的电力负荷预测;文献^[20]引入注意力机制来调节捕获的信息向量权重,赋予各历史状态对应的注意力权重;文献^[21]在改进长短期记忆神经网络预测模型中加入注意力层,注意力机制可赋予模型隐含层不同的权重;

收稿日期:2024-01-09;修回日期:2024-07-30

在线出版日期:2024-10-08

基金项目:国家电网有限公司科技项目(SGSDDK00PDJS2250114);国家自然科学基金资助项目(52377127)

Project supported by the Science and Technology Project of State Grid Corporation of China(SGSDDK00PDJS2250114) and the National Natural Science Foundation of China(52377127)

文献[22]使用注意力机制进行优化,通过赋予长短期记忆神经网络隐含层不同的权重增强重要信息的作用。上述研究在利用组合模型的基础上,重构特征变量,引入注意力机制对模型输入赋予权重比,使模型赋予重要特征变量更多的注意力,但组合模型结构复杂,具有双重输入,上述研究仅对隐含层赋予时间注意力机制,而忽略了变量之间的相互关联作用以及神经网络无差别对待输入变量的特点,因此预测效果还有进一步提升的空间。

综上,本文提出一种基于改进双重注意力机制的短期电-热负荷协同预测模型。改进的压缩和激励(squeeze and excitation, SE)注意力机制为各输入通道特征加权,对输入变量进行分组卷积(group convolution, GC),并对关键信息赋予高权重,改进多头特征注意力对卷积结果进行加权后,输入GRU对负荷进行预测,从而利用双重注意力提高模型的泛化能力,解决神经网络无差别对待输入特征的问题。此外,在模型中采用多输出通道来进一步提高模型的灵活性和适应性。

1 基于改进双重注意力机制的负荷预测模型

1.1 分组卷积神经网络

CNN模型可对高维数据提取数据特征,降低模型复杂度,其网络结构包括输入层、卷积层、池化层、全连接层和输出层。传统的CNN输入特征图尺寸为 $H \times W \times C$ (H 、 W 、 C 分别为特征图的高、宽和通道数),对所有输入变量进行卷积,其卷积核大小为 $h_1 \times w_1 \times c_1$ (h_1 、 w_1 、 c_1 分别为卷积核的高、宽和通道数)。而分组卷积神经网络(group convolutional neural network, GCNN)将输入特征变量分成 g 组,卷积核大小为 $h_1 \times w_1 \times (c_1/g)$,再分别采用 g 个卷积核进行卷积,分组后卷积核的尺寸减小至原来的 $1/g$,参数量也相对减少,生成 C_2 张特征图,提升模型的计算效率。CNN与GCNN的结构原理图如附录A图A1所示。

本文选用GCNN对输入的各时序特征变量进行分组卷积,GCNN结构包含3个卷积层和1个展平操作,每个卷积层包含1个卷积操作和1个批量归一化操作。相较于CNN,GCNN以较小的计算量降低数据维度,提取不同时刻变量的重要信息,有利于提高模型的泛化能力。

1.2 GRU网络

为了解决循环神经网络(recurrent neural network, RNN)在处理长时间序列时出现的长期记忆和梯度消失的问题^[23],学者们提出GRU网络。GRU网络内部的门控结构不仅能解决梯度消失的问题,而且能够完整地考虑时间序列中的长短期依赖关系,从而学习到负荷变化的情况,有效提升模型精度。

GRU网络内部结构较简单,缩短了训练时长,且训练效果较好。GRU网络内部结构如附录A图A2所示。

GRU通过当前时刻 t 的输入 x_t 和上一节点传递的隐状态 h_{t-1} 得到下一节点隐状态 h_t ,其内部含有重置门 r_t 和更新门 Z_t ,重置门用于控制前一时刻隐状态在当前时刻的保留程度,其值越小说明保留的信息越少,更新门控制前一时刻的隐状态和加入当前节点的部分信息,其值越大说明前一时刻保留的信息越多。计算过程如下:

$$r_t = \sigma(W_r x_t + U_r h_{t-1}) \quad (1)$$

$$Z_t = \sigma(W_z x_t + U_z h_{t-1}) \quad (2)$$

$$\tilde{h}_t = \tanh(W_c x_t + U_c (r_t h_{t-1})) \quad (3)$$

$$h_t = Z_t \tilde{h}_t + (1 - Z_t) h_{t-1} \quad (4)$$

式中: σ 为激活函数Sigmoid; h_t 为 t 时刻的隐层输出; \tilde{h}_t 为 h_{t-1} 和 x_t 拼接后的矩阵,并通过激活函数 \tanh 进行映射; U_r 、 U_z 、 U_c 、 W_r 、 W_z 、 W_c 、 W_r 为参数矩阵; $1 - Z_t$ 表示对 h_{t-1} 的遗忘程度。

1.3 改进的SE注意力机制

SE注意力机制是通过自动学习的方式在通道维度增加注意力机制,为每个通道赋予权重值,使神经网络更加关注某些通道,从而获取更重要的信息,提升模型精度。本文采用GCNN对输入时间序列特征变量进行分组再卷积,其分组相当于通道,对每组赋予相应权重,突出关键时序变量,从而提升预测精度。

传统的SE注意力机制(如附录A图A3所示)通过Squeeze将每个通道的二维特征($H \times W \times C$)压缩为1个实数($1 \times 1 \times C$),通过Excitation给每个特征通道赋予一个权重值,对特征权重归一化后将其加到各通道上(Scale)。传统的SE注意力机制的Squeeze仅采用最大池化来提取整张特征图的空间像素信息,并经共享隐层后压缩特征,这都容易造成信息丢失问题。

为弥补上述SE注意力的缺陷,本文提出一种改进的双重SE注意力机制,如图1所示,即在传统的SE注意力机制上,单独另设一条SE注意力机制,其中采用全局平均池化(Averagepool)操作来压缩特征图信息,将最大池化(Maxpool)和平均池化压缩的信息串联(Contact)后输入4个全连接层中,不进行维度压缩,对4个全连接层的输出取平均值,再经Relu与Sigmoid函数压缩到0.5~1,为各通道赋予0.5~1的权重。计算公式为:

$$X = \text{Contact}(\text{Averagepool}, \text{Maxpool}) \quad (5)$$

$$M = \sigma(QX + b) \quad (6)$$

式中: X 为压缩后的信息串联得到的矩阵; M 为通过

线性变换获得的权值矩阵; Q 为权重矩阵; b 为偏置矩阵。

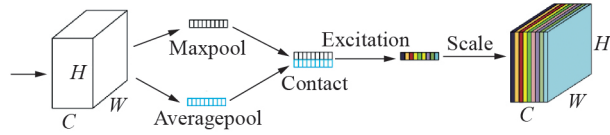


图 1 改进的SE注意力机制

Fig.1 Improved SE attention mechanism

通过 4 个全连接层,即进行 4 次线性变换,输出 4 次线性变换矩阵 $M^1 \sim M^4$,对其取平均值得到 M' ,再通过修正线性单元 Relu 函数与激活函数 Sigmoid 获得权重值 α_m ,如式(7)、(8)所示。

$$M' = \frac{1}{4} \sum_{s=1}^4 M^s \quad (7)$$

$$\alpha_m = \sigma(\text{Relu}(M')) \quad (8)$$

对各通道采用上述公式进行计算,即可得到各通道的权重值,再为各通道进行赋权。赋权矩阵 α_m 为:

$$\alpha_m = [\alpha_1, \alpha_2, \dots, \alpha_c] \quad (9)$$

式中: $\alpha_i (i=1, 2, \dots, C)$ 为第 i 个通道的权重值。

将历史 C 小时的气象特征、历史负荷数据构建为输入特征矩阵 X_A (每组数据含有 C 个通道,每个通道含有该时刻经过筛选的气象和负荷数据),如式(10)所示。

$$X_A = [x'_1, x'_2, \dots, x'_c] \quad (10)$$

式中: $x'_a (a=1, 2, \dots, C)$ 为预测时刻历史滞后 a 小时的特征变量矩阵。

通过改进的 SE 注意力机制计算赋权矩阵 α_m ,对输入特征变量进行赋权得到初步赋权矩阵 X'_A ,如式(11)所示。

$$X'_A = [\alpha_1 x'_1, \alpha_2 x'_2, \dots, \alpha_c x'_c] \quad (11)$$

由于本文直接的输入特征是各基本能耗单元的负荷形态在多时间断面下的不同表现形式,特征规模较大,因此需要优化的特征注意力权重系数较多。若采用 Softmax 函数,则会导致最终得到的各特征注意力权重系数的数量级较小,即使对于重要特征,其注意力权重也会因受特征规模的影响而维持在一个较小的数值。此外,Softmax 的影响是持续性的,若

在初始时刻没有学习好,则可能导致注意力权重对应的整幅特征图在整个学习过程中的特征提取效果较差。

综上所述,本文先将向量 M' 经 Relu 函数单侧抑制到正区间,再利用 Sigmoid 进行归一化,目的是将各特征的注意力权重系数限制在 0.5~1,即使是最不重要的特征,也对其赋予 0.5 的权重,重要特征与不重要特征之间的权重系数保持一定的差距。此外,即使模型的学习效果不佳,各特征的注意力系数也相差不大,对模型的泛化性能影响较小,起到了一种保护作用。

1.4 基于软注意力的多头特征注意力机制

注意力机制可分为硬注意力机制与软注意力机制,硬注意力机制以非 0 即 1 的方式限制无效信息,从而达到聚焦有效信息的目的。与硬注意力机制不同,软注意力机制通过对各特征赋予权重参数实现模型对不同特征有差异性的关注。

本文提出一种基于软注意力的多头特征注意力机制,如图 2 所示。图中: $x_1 \sim x_m$ 为输入特征, m 为输入特征总数; $y_1^b \sim y_m^b (b=1, 2, \dots, s)$ 分别为第 b 次线性变换得到的 $x_1 \sim x_m$ 的初步权重值, s 为线性变换次数; $\bar{y}_1 \sim \bar{y}_m$ 分别为 $x_1 \sim x_m$ 的初步均值权重值; $\beta_1 \sim \beta_m$ 分别为 $x_1 \sim x_m$ 的最终注意力权重系数。该注意力模块可以即插即用,为输入特征赋予权重系数,从而使模型具有更好的稳定性与泛化能力。本文模型经过卷积后展平成一维输入变量,对各特征赋予相应权重后再将其输入 GRU 网络,突出关键输入变量,弱化对模型影响较小的输入变量,提升模型精度。

将卷积后的结果经过展平层进行平铺展平,在将其输入 GRU 网络进行训练前,输入特征向量 $X_b = [x_1, x_2, \dots, x_m]$,利用一个共享线性层得到初步权重向量 Y ,如式(12)所示。

$$Y = \sigma(QX_b + b) \quad (12)$$

在此基础上,由于本文采用多头特征注意力机制,即总共进行 s 次线性变换,输出 s 次线性变换矩阵 $Y^1 \sim Y^s$,第 s 次线性变换得到的初步权重向量 $Y^s = [y_1^s, y_2^s, \dots, y_m^s]$,进而对得到的 s 个初步权重向量取均值,得到初步均值权重向量 \bar{Y} ,如式(13)所示。

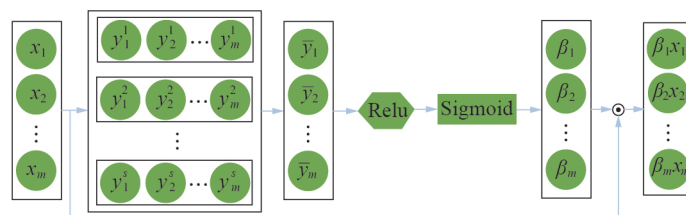


图 2 多头特征注意力机制

Fig.2 Multi-head feature attention mechanism

$$\bar{Y} = \frac{1}{s} \sum_{n=1}^s Y^n = [\bar{y}_1, \bar{y}_2, \dots, \bar{y}_m] \quad (13)$$

利用Relu激活函数将得到的初步均值权值向量 \bar{Y} 限制到非负区间,最后利用Sigmoid函数对权重系数进行归一化处理,得到各输入特征的最终注意力权重系数向量 β ,如式(14)所示。

$$\beta = \text{Sigmoid}(\text{Relu}[\bar{y}_1, \bar{y}_2, \dots, \bar{y}_m]) = [\beta_1, \beta_2, \dots, \beta_m] \quad (14)$$

将最终的注意力权重系数向量与原输入特征向量按元素相乘,得到加权后的输入特征向量,如式(15)所示。

$$X'_B = \beta \odot X_B \quad (15)$$

式中: X'_B 为加权后的输入特征向量;“ \odot ”表示哈达玛积。

本文所提出的2种注意力机制均采用Relu和Sigmoid函数对权值进行归一化,但2种机制的赋权意义不同。改进的SE注意力机制通过对通道(即分组)进行赋权后将其输入GCNN以提取重要信息,而改进的多头注意力机制则是对所有特征进行赋权,使GRU网络对重要的输入变量给予更多的关注。

2 电-热负荷预测模型

2.1 输入特征的处理

本文选择美国亚利桑那州立大学校园综合能源系统负荷数据来进行验证,气象数据来源于美国国家气象局官网^[24-25],时间尺度为1h,一天共8760个样本。本文选用随机森林对特征变量的重要程度进行分析,对输入特征变量进行筛选,重要程度用贡献度来表示,其原理是,计算每个特征在随机森林的每棵决策树的贡献度,并对其取平均值,贡献度越高,则表明相应的特征变量越重要。

为了提升模型的拟合能力,本文选取典型气象数据(风速、可降水量、风向、湿度、温度、气压、云类型、露点、地表反照率)以及电-热负荷数据进行特征选择,其贡献度如图3所示。

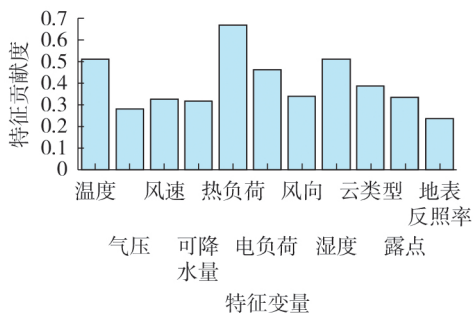


图3 气象特征和历史负荷的贡献度

Fig.3 Contribution of meteorological feature and historical load

由图3可知,可降水量、风速、风向、气压、云类型、露点和地表反照率的贡献度普遍较低,其值均小于0.4,而电负荷、热负荷、温度和湿度的贡献度较高,其值均大于0.4。本文选用GCNN对输入变量进行卷积,将历史时刻的气象数据和负荷数据按照时间进行分组再卷积,图3中贡献度低于0.4的气象数据不作为模型输入。

经过筛选对比验证后,将湿度、温度、小时数(保留时序特性)、电负荷、热负荷作为模型特征变量组合,同时,为了使模型能更好地捕捉负荷的时序变化特性,本文通过对比选择合适的时间滞后的特征变量组合作为模型输入。

本文将历史滞后3h至历史滞后9h的特征变量组合作为模型输入进行对比,以夏季典型周为例,运用本文模型预测电-热负荷,其预测结果的MAPE值如图4所示。显然,历史滞后7h的MAPE值均最小,精度最高,因此选用历史滞后7h的特征变量组合作为模型输入。

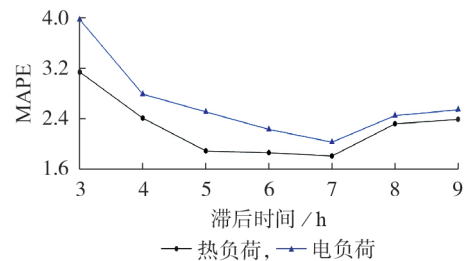


图4 不同历史滞后时刻预测的电-热负荷MAPE值
Fig.4 MAPE values of electricity and thermal load forecasted at different historical lag times

2.2 预测模型结构

改进的双重注意力机制的电-热负荷预测模型包括输入变量、改进的SE注意力机制层、3层GCNN、改进的多头注意力机制层、2层GRU网络、全连接层,其结构图如图5所示,图中: $x'_{11} - x'_{1c}$ 为特征变量; Y_{ele} 、 Y_{heat} 分别为预测的电负荷和热负荷。

电-热负荷预测的具体步骤为:首先,将由历史7h的特征变量(湿度、温度、小时数、电负荷、热负荷)构成的特征向量输入改进的SE注意力机制层中,量化各小时特征变量的重要程度,为各小时的变量赋权,获得加权特征向量;其次,将加权特征向量输入分组卷积层中,提取各时序变量中的重要信息;然后,在此基础上,进行展平操作,再将展平后的一维向量输入改进的多头注意力机制层中,为各输入变量赋权,获得加权特征向量;最后,通过GRU网络训练学习负荷变化的时序性,输出电-热负荷预测结果,并对该结果进行评估。

模型的超参数和训练时间对训练结果有重要影响,尤其是GRU网络层的设置直接影响模型的线性

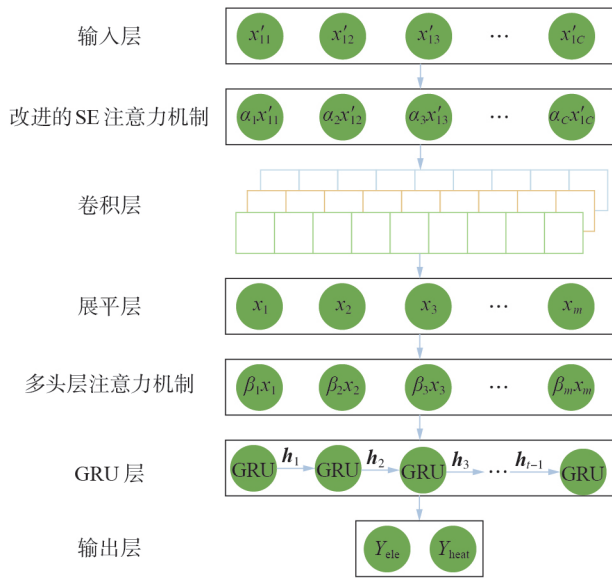


图 5 预测模型结构图

Fig.5 Structure diagram of forecasting model

拟合能力。本文通过网格搜索算法对超参数进行寻优,最终本文实验中的参数设置为:GCNN为3层,步长均为2,各层卷积核数分别为8、16、32;GRU网络的双层神经元数分别为64、128,学习率为0.01,训练周期为100次,迭代次数为600次。

3 算例分析

3.1 对比模型和评估指标

为验证本文模型的有效性,选用前2000 h的负荷数据构成训练集,预测未来一周7 d共168 h的负荷。

为了便于对比分析,分别建立GCNN^[26]、GCNN-GRU^[26]、传统的SE注意力机制^[27]、改进的SE注意力机制、多头特征注意力机制以及Transformer^[28]进行对比。同时,为验证本文模型的科学性和稳定性,选择夏、冬2季来进行验证。夏季热负荷需求低,热负荷曲线峰谷差值不大,电负荷消耗较大,峰谷差值也较大,负荷曲线波动明显;冬季电负荷消耗量略小于夏季,但冬季的热负荷需求多,且峰谷差值较大,负荷曲线波动同样明显。由于夏季电负荷和冬季热负荷的波动性以及不确定性较强,其季节特性与负荷

以及负荷之间的耦合关系更为复杂。

为了验证本文模型的预测性能,采用MAPE、均方根误差(root mean square error, RMSE)和平均绝对误差(mean absolute error, MAE)对预测结果进行评价。其中,MAPE主要体现模型整体误差水平,MAE表示预测值与实际值的相似程度, RMSE表示极端异常值对预测结果的影响,三者的计算公式如附录B式(B1)~(B3)所示。

3.2 夏季负荷序列预测

选用夏季某一周的负荷数据进行验证分析,图6为电负荷和热负荷的预测结果。由图可知,各模型均能预测负荷变化趋势,在负荷剧烈波动时,预测会出现偏差。在图6(b)中的150—160 h时段,其他模型的预测结果明显与实际值偏离较多,而本文模型的预测结果与实际负荷曲线更贴合,偏差更小,这说明本文模型有良好的预测效果。

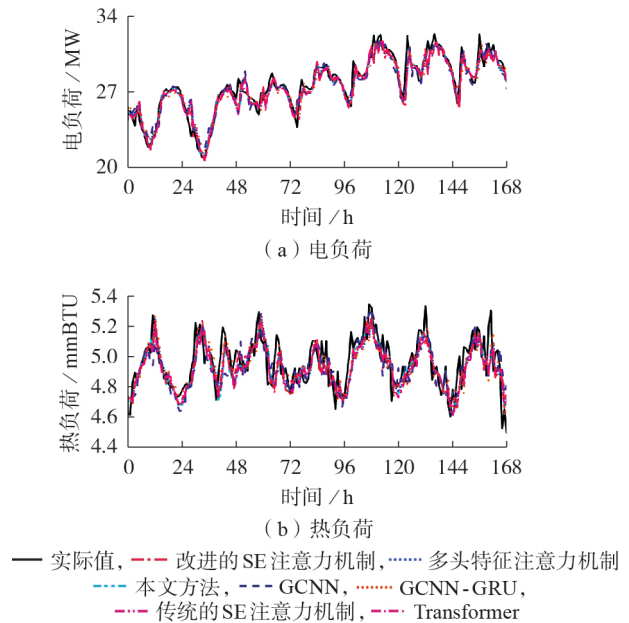


图 6 夏季电-热负荷序列预测结果

Fig.6 Forecasting results of electricity and thermal load series in summer

表1为各模型的夏季负荷预测效果对比。由表可知:GCNN和GRU的组合模型误差基本低于

表 1 不同模型的夏季负荷预测效果对比

Table 1 Comparison of summer load forecasting effect among different models

模型	MAPE		RMSE		MAE	
	电负荷 / %	热负荷 / %	电负荷 / MW	热负荷 / mmBTU	电负荷 / MW	热负荷 / mmBTU
GCNN	2.56	2.01	0.958	0.1200	0.690	0.0979
GCNN-GRU	2.36	1.92	0.872	0.1270	0.649	0.0967
传统的SE注意力机制	2.18	1.87	0.894	0.1163	0.591	0.0931
改进的SE注意力机制	2.14	1.84	0.887	0.1160	0.583	0.0922
多头特征注意力机制	2.15	1.86	0.851	0.1180	0.610	0.0936
Transformer	2.05	1.82	0.840	0.1154	0.573	0.9160
本文模型	2.03	1.81	0.836	0.1150	0.564	0.0904

GCNN模型,这说明GRU能更好地学习负荷的时间序列特性,与GCNN相结合增强了模型的拟合能力,组合模型能够在一定程度上提升预测精度;相较于传统的SE注意力机制,改进的SE注意力机制的精度更高,验证了其有效性;相较于GCNN-GRU,本文提出的2种注意力机制的精度都有所提升,改进的SE注意力机制和多头特征注意力机制的电负荷MAPE值分别减小9.32%、8.90%,热负荷MAPE值分别减小4.17%、3.13%,这说明本文所提2种注意力机制均能有效地对输入变量进行权重分配,提升模型精度,改进的SE注意力机制提升的精度略高于多头特征注意力机制;与Transformer模型相比,本文模型的预测精度较高;相较于GCNN-GRU,本文模型的电负荷和热负荷MAPE值分别减小13.98%、5.73%,这说明本文方法结合2种注意力机制的优势增强了模型的拟合能力,可挖掘输入变量耦合性,预测效果更好。

综上所述,本文模型能够充分挖掘负荷之间以及负荷与输入特征变量间的耦合关系,对于夏季场景能够预测负荷波动趋势,精度较高,比其他模型有更好的预测效果。

3.3 冬季负荷序列预测

为了进一步验证本文模型能捕捉负荷波动特性,选取冬季某一周的负荷数据进行验证。冬季电负荷消耗较小,但热负荷消耗较大,负荷曲线波动较大。分别采用GCNN、GCNN-GRU、改进的SE注意力机制、传统的SE注意力机制、多头注意力机制、Transformer这6种模型进行负荷预测分析对比。

电负荷和热负荷实际值与不同模型预测曲线的对比如附录A图A4所示。在图A4(a)中的150—160h时段,GCNN和GCNN-GRU对负荷波动趋势的预测结果偏离实际值明显,而添加注意力机制的模型能较好地预测负荷变化趋势。

本文模型、Transformer、改进的SE注意力机制、传统的SE注意力机制、多头特征注意力机制、GCNN-GRU、GCNN模型的精度对比如附录A表A1所示。由表可知:冬季热负荷的预测精度低于电负荷,这是由于冬季热负荷特性更为复杂;GCNN-GRU组合模型的精度相较于GCNN提升较多;相较于GCNN-GRU,在添加注意力机制的3种指标模型中,改进的SE注意力机制的预测精度高于传统的SE注意力机制,改进的SE注意力机制和多头特征注意力机制的电负荷的MAPE分别减小9.52%、8.23%,热负荷的MAPE分别减小12.33%、15.07%;本文方法结合2种机制后的电负荷和热负荷预测精度均最高;2种改进的注意力机制在冬季热负荷特性波动的场景下仍然适用,且精度提升程度比夏季更高;相较于GCNN-GRU,本文模型的电负荷和热负荷的MAPE分别减

小11.69%、25.75%。

注意力机制的运用以及对输入特征赋予的权重突出了重要信息对模型的影响,弱化了非关键信息对模型的影响,解决了神经网络无差别对待输入特征变量的问题,提升了模型精度。上述夏、冬2种场景消融实验的对比说明,改进的SE注意力机制为各分组赋予权重,多头特征注意力机制为各输入变量赋予合适的权重,本文所提的2种注意力机制均能提升预测精度,更好地挖掘负荷与输入特征的依赖关系,使得模型更好地预测负荷变化趋势,提升预测精度。

4 结论

本文提出一种基于改进的双重注意力机制的电-热负荷短期预测方法,主要结论为:

1)本文提出改进的SE注意力机制,该机制自适应地为输入信息分配注意力权重,以评估不同输入对期望输出的贡献度,并突出重要输入信息的影响;

2)本文通过引入改进的多头特征注意力机制实现对GRU神经网络预测性能的改善,同时通过对模型训练过程中注意力权重变化趋势及最终结果的可视化在一定程度上实现了模型的可解释性;

3)6组消融实验对比说明了本文所提的2种注意力机制均能提升预测精度,添加了注意力机制的模型优于未添加注意力机制的模型,本文模型的效果比单一注意力机制的效果更好,本文模型对于夏、冬2季具有不同负荷特性的场景都能达到较好的拟合效果,其MAPE均低于3%,结果具有普适性。

后续笔者将进一步挖掘综合能源系统中的耦合特性。

附录见本刊网络版(<http://www.epae.cn>)。

参考文献:

- [1] 王杰,何山,王维庆,等. 基于混合灰色理论的风速预测研究[J]. 太阳能学报,2018,39(12):3544-3549.
WANG Jie, HE Shan, WANG Weiqing, et al. Study on wind speed prediction based on mixed grey theory[J]. Acta Energetica Solaris Sinica, 2018, 39(12): 3544-3549.
- [2] 杨丽徒,张永锋,许向伟,等. 半参数回归分析法在电力负荷预测中的应用[J]. 郑州大学学报(工学版),2010,31(3):29-32.
YANG Lixi, ZHANG Yongfeng, XU Xiangwei, et al. Application of semi-parametric regression analysis method in electric power load forecasting[J]. Journal of Zhengzhou University (Engineering Science), 2010, 31(3): 29-32.
- [3] LAOUAFI A, MORDJAOUI M, LAOUAFI F, et al. Daily peak electricity demand forecasting based on an adaptive hybrid two-stage methodology[J]. International Journal of Electrical Power & Energy Systems, 2016, 77: 136-144.
- [4] 郑瑞晓,张妹,肖先勇,等. 考虑温度模糊化的多层长短时记忆神经网络短期负荷预测[J]. 电力自动化设备,2020,40(10):181-186.
ZHENG Ruixiao, ZHANG Shu, XIAO Xianyong, et al. Short-

- term load forecasting of multi-layer long short-term memory neural network considering temperature fuzziness[J]. *Electric Power Automation Equipment*, 2020, 40(10):181-186.
- [5] 刘波,秦川,鞠平,等. 基于XGBoost与Stacking模型融合的短期母线负荷预测[J]. *电力自动化设备*, 2020, 40(3):147-153.
LIU Bo, QIN Chuan, JU Ping, et al. Short-term bus load forecasting based on XGBoost and Stacking model fusion[J]. *Electric Power Automation Equipment*, 2020, 40(3):147-153.
- [6] 张超,陈杰睿,冯平. 基于混沌理论的电力系统负荷预测应用[J]. *计算机与数字工程*, 2018, 46(11):2165-2169, 2174.
ZHANG Chao, CHEN Jierui, FENG Ping. Power system load forecasting based on chaos theory[J]. *Computer & Digital Engineering*, 2018, 46(11):2165-2169, 2174.
- [7] FAN G F, PENG L L, HONG W C. Short term load forecasting based on phase space reconstruction algorithm and bi-square kernel regression model[J]. *Applied Energy*, 2018, 224:13-33.
- [8] BARMAN M, DEV CHOUDHURY N B, SUTRADHAR S. A regional hybrid GOA-SVM model based on similar day approach for short-term load forecasting in Assam, India[J]. *Energy*, 2018, 145:710-720.
- [9] VAN HOUDT G, MOSQUERA C, NÁPOLES G. A review on the long short-term memory model[J]. *Artificial Intelligence Review*, 2020, 53(8):5929-5955.
- [10] RAFI S H, NAHID-AL-MASOOD, DEEBA S R, et al. A short-term load forecasting method using integrated CNN and LSTM network[J]. *IEEE Access*, 2021, 9:32436-32448.
- [11] 朱凌建,荀子涵,王裕鑫,等. 基于CNN-BiLSTM的短期电力负荷预测[J]. *电网技术*, 2021, 45(11):4532-4539.
ZHU Lingjian, XUN Zihan, WANG Yuxin, et al. Short-term power load forecasting based on CNN-BiLSTM[J]. *Power System Technology*, 2021, 45(11):4532-4539.
- [12] 王金玉,胡喜乐,闫冠宇. 基于VMD的CNN-BiLSTM-Au的短期负荷预测[J]. *吉林大学学报(信息科学版)*, 2023, 41(6):1007-1014.
WANG Jinyu, HU Xile, YAN Guanyu. Short-term load prediction of CNN-BiLSTM-Att based on VMD[J]. *Journal of Jilin University(Information Science Edition)*, 2023, 41(6):1007-1014.
- [13] WANG J Z, ZHANG L Y, LI Z W. Interval forecasting system for electricity load based on data pre-processing strategy and multi-objective optimization algorithm[J]. *Applied Energy*, 2022, 305:117911.
- [14] 肖白,高文瑞,李道明,等. 基于 3σ -CEEMDAN-LSTM的空间负荷预测方法[J]. *电力自动化设备*, 2023, 43(3):159-165.
XIAO Bai, GAO Wenrui, LI Daoming, et al. Spatial load forecasting method based on 3σ -CEEMDAN-LSTM[J]. *Electric Power Automation Equipment*, 2023, 43(3):159-165.
- [15] 王凌云,周翔,田恬,等. 基于多维气象信息时空融合和MPA-VMD的短期电力负荷组合预测模型[J]. *电力自动化设备*, 2024, 44(2):190-197.
WANG Lingyun, ZHOU Xiang, TIAN Tian, et al. Combination forecasting model of short-term power load based on multi-dimensional meteorological information spatio-temporal fusion and MPA-VMD[J]. *Electric Power Automation Equipment*, 2024, 44(2):190-197.
- [16] 邹智,吴铁洲,张晓星,等. 基于贝叶斯优化CNN-BiGRU混合神经网络的短期负荷预测[J]. *高电压技术*, 2022, 48(10):3935-3945.
ZOU Zhi, WU Tiezhou, ZHANG Xiaoxing, et al. Short-term load forecasting based on Bayesian optimization CNN-BiGRU hybrid neural network[J]. *High Voltage Technology*, 2019, 48(10):3935-3945.
- [17] 崔树银,汪昕杰. 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[J]. *电力自动化设备*, 2022, 42(5):32-39, 81.
CUI Shuyin, WANG Xinjie. Multivariate load forecasting in integrated energy system based on maximal information coefficient and multi-objective Stacking ensemble learning[J]. *Electric Power Automation Equipment*, 2022, 42(5):32-39, 81.
- [18] AOUAD M, HAJJ H, SHABAN K, et al. A CNN-sequence-to-sequence network with attention for residential short-term load forecasting[J]. *Electric Power Systems Research*, 2022, 211:108152.
- [19] 赵兵,王增平,纪维佳,等. 基于注意力机制的CNN-GRU短期电力负荷预测方法[J]. *电网技术*, 2019, 43(12):4370-4376.
ZHAO Bing, WANG Zengping, JI Weijia, et al. A short-term power load forecasting method based on attention mechanism of CNN-GRU[J]. *Power System Technology*, 2019, 43(12):4370-4376.
- [20] 汤健,侯慧娟,陈洪岗,等. 基于BI-GRU改进的Seq2Seq网络的变压器油中溶解气体浓度预测方法[J]. *电力自动化设备*, 2022, 42(3):196-202, 217.
TANG Jian, HOU Huijuan, CHEN Honggang, et al. Concentration prediction method based on Seq2Seq network improved by BI-GRU for dissolved gas in transformer oil[J]. *Electric Power Automation Equipment*, 2022, 42(3):196-202, 217.
- [21] 田浩含,张智晟,于道林. 基于改进LSTM的区域综合能源系统多元负荷短期预测研究[J]. *电力系统及其自动化学报*, 2021, 33(9):130-137.
TIAN Haohan, ZHANG Zhisheng, YU Daolin. Research on multi-load short-term forecasting of regional integrated energy system based on improved LSTM[J]. *Proceedings of the CSU-EPSA*, 2021, 33(9):130-137.
- [22] WU K H, WU J, FENG L, et al. An attention-based CNN-LSTM-BiLSTM model for short-term electric load forecasting in integrated energy system[J]. *International Transactions on Electrical Energy Systems*, 2021, 31(1):1-15.
- [23] PASCANU R, MIKOLOV T, BENGIO Y, et al. On the difficulty of training recurrent neural networks[C]//*Proceedings of the 30th International Conference on International Conference on Machine Learning*. Atlanta, GA, USA: ACM, 2013:1310-1318.
- [24] AUS. Campus metabolism[DB/OL]. [2021-10-12]. <http://cm.asu.edu/>.
- [25] NSRDB Data viewer[DB/OL]. [2021-10-12]. https://maps.nrel.gov/nsrdb-viewer/?aL=x8Cl3i%255Bv%255D%3Dt%26Jea8x6%255Bv%255D%3Dt%26Jea8x6%255Bd%255D%3D1%26VRLt_G%255Bv%255D%3Dt%26VRLt_G%255Bd%255D%3D2%26mcQtmw%255Bv%255D%3Dt%26mcQtmw%255Bd%255D%3D3&bL=clight&cE=0&lR=0&mC=4.740675384778373%2C22.8515625&zL=2.
- [26] 宋一格,王宁,李宏昌,等. 基于分组卷积与双注意力机制的河流水面污染图像分类[J]. *计算机系统应用*, 2022, 31(9):250-256.
SONG Yige, WANG Ning, LI Hongchang, et al. Image classification of river water surface pollution based on grouped convolution and dual attention mechanism[J]. *Computer Systems and Applications*, 2022, 31(9):250-256.
- [27] 莫文彬,伊力哈木·亚尔买买提. 通道注意力机制的局部遮挡人脸表情识别[J]. *计算机应用与软件*, 2024, 41(2):145-151.
MO Wenbin, ILHAM Yarmaiti. Local occlusion facial expression recognition based on channel attention mechanism[J]. *Computer Applications and Software*, 2024, 41(2):145-151.

[28] 黄旭锐,于丰源,杨波,等. 基于Transformer网络和多任务学习的园区综合能源系统电-热短期负荷预测方法[J]. 南方电网技术,2023,17(1):152-160.

HUANG Xurui, YU Fengyuan, YANG Bo, et al. Short-term electric-thermal load forecasting method for park-level integrated energy system based on transformer network and multi-task learning[J]. Southern Power System Technology, 2023, 17(1):152-160.

作者简介:

余 强(1984—),男,教授,博士研究生导师,主要研究方向为配电网规划、运行与控制,光伏/光热发电技术以及电动农机电机驱动控制等(**E-mail**:yuqiang_21st@sina.com);

韩静娴(1997—),女,硕士研究生,主要研究方向为配电网规划、运行与控制(**E-mail**:herzeleid0602@cau.edu.cn)。

(编辑 王锦秀)

Collaborative forecasting of electricity-thermal load based on improved dual squeeze and excitation and multi-head feature attention mechanism

YU Qiang¹, HAN Jingxian¹, YANG Ziliang¹, SONG Jidong¹, YANG Dechang¹, QI Haijie², YU Peng³

(1. College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China;

2. State Grid Smart Grid Research Institute Co., Ltd., Beijing 102200, China;

3. State Grid Shandong Electric Power Research Institute, Jinan 250003, China)

Abstract: Loads are diverse and coupled in the integrated energy system, in order to improve the accuracy of load forecasting, a short-term electricity-thermal load collaborative forecasting model based on an improved dual attention mechanism and group convolutional neural network-gated recurrent unit is proposed. Each input channel is weighted by the improved squeeze and excitation attention mechanism, and grouped for the convolution. The convolution results are weighted by the multi-head feature attention mechanism and the load is forecasted by the input gated recurrent unit model. The example simulative results show that the mean absolute percentage errors of the proposed model are less than 3%.

Key words: integrated energy system; load forecasting; group convolutional neural network; gated recurrent unit; improved squeeze and excitation attention mechanism; multi-head feature attention mechanism

(上接第185页 continued from page 185)

Frequency response characteristics modeling method of virtual power plant considering multiple flexible resources

SHI Zheng¹, ZHU Chao², XU Yin¹

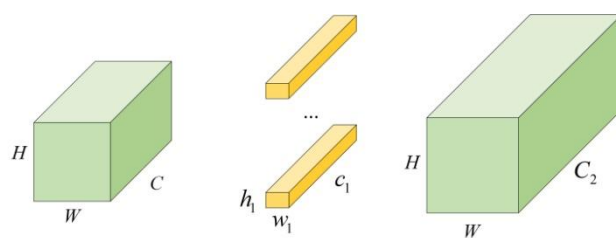
(1. School of Electrical Engineering, Beijing Jiaotong University, Beijing 100044, China;

2. Economic and Technical Research Institute of State Grid Zhejiang Electric Power Co., Ltd., Hangzhou 310002, China)

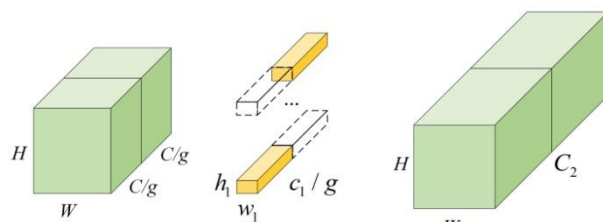
Abstract: In order to adapt to the double-high development trend of DC receiving-end regional power grid, a frequency response characteristics modeling method of virtual power plant (VPP) considering multiple flexible resources is proposed for the problem of emergency frequency control of large power grid assisted by VPP. The “centralized-decentralized” hierarchical control architecture for VPP participating in emergency frequency control is designed. The frequency response characteristics of multiple flexible resources are analyzed, and the system frequency response model and equivalent model of VPP are constructed. Based on the successive linear programming method, the VPP frequency response model parameter estimation method is proposed, aiming to balance the accuracy and computational efficiency of the model. Based on PSD-BPA, the accuracy of the proposed method is verified by simulation test of three-area DC receiving-end regional power grid, the parameter estimation results are conservative, which can prevent the system from misjudging in emergency frequency control decision.

Key words: virtual power plants; emergency frequency control; frequency response model; parameter estimation; multiple flexible resources

附录 A



(a) 卷积神经网络



(b) 分组卷积神经网络

图 A1 CNN 与 GCNN 的结构原理图

Fig. A1 Schematic diagram of the structure of CNN and GCNN

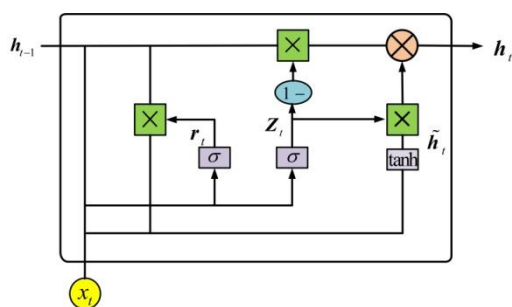


图 A2 GRU 结构

Fig. A2 The structure of GRU

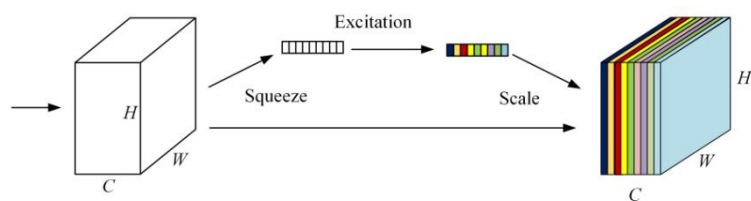
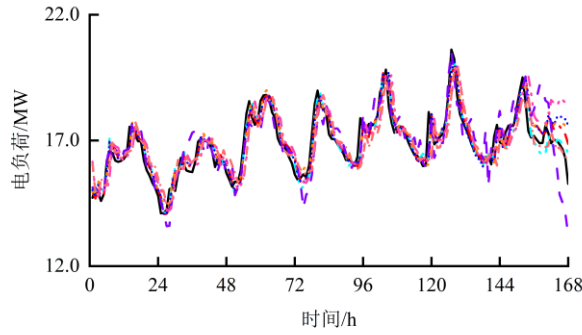
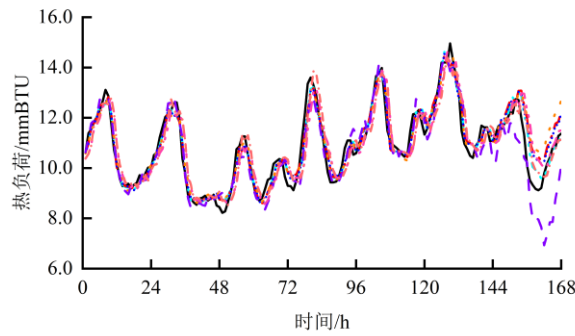


图 A3 传统 SE 注意力机制

Fig. A3 Traditional SE attention mechanism



(a)冬季电负荷预测结果图



(b)冬季热负荷预测结果图

— 实际值 - - - 改进SE 多头注意力 - · - · 本文方法
 - - - GCNN - · - · GCNN-GRU - · - · 传统SE - · - · Transformer

图 A4 冬季电热负荷序列预测结果图

Fig.A4 Winter electric-heat load series forecast results

表 A1 冬季负荷预测效果对比分析

Table A1 Comparative analysis of winter load forecasting effect

模型	MAPE		RMSE		MAE	
	电/%	热/%	电/ MW	热/ mmBTU	电/ MW	热/ mmBTU
GCNN	3.34	4.96	0.745	0.756	0.562	0.535
GCNN-G RU	2.31	3.65	0.524	0.548	0.392	0.389
传统 SE	2.27	3.34	0.483	0.526	0.371	0.355
改进 SE	2.09	3.20	0.478	0.493	0.351	0.346
多头注意 力	2.12	3.10	0.511	0.450	0.357	0.331
Transfor mer	2.18	2.82	0.496	0.481	0.365	0.326
本文方法	2.04	2.71	0.452	0.389	0.345	0.291

附录 B

$$E_{\text{MAPE}} = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y} - y_i}{y_i} \right| \quad (\text{B1})$$

$$E_{\text{RMSE}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y} - y_i)^2} \quad (\text{B2})$$

$$E_{\text{MAE}} = \frac{1}{n} \sum_{i=1}^n |\hat{y} - y_i| \quad (\text{B3})$$

式中： n 为预测样本的总数， y_i 为预测点 i 的实际值， \hat{y} 为预测值。上式中的值越低，则证明模型的精度越高