

DOI:10.19651/j.cnki.emt.2209103

基于混合神经网络的短期电力负荷预测方法

任 轩 汪庆年 尚 宝 姜宏伟 常 乐
(南昌大学信息工程学院 南昌 330036)

摘要: 电力负荷预测对电网的经济运行至关重要,为提高短期负荷预测精度并降低混合神经网络模型的训练时间,提出了一种基于多层感知器(MLP)的基础子网、简单循环单元(SRU)与主成分分析(PCA)的短期电力负荷预测模型。首先,考虑影响电力负荷变化的各种因素,建立负荷预测输入特征集;其次,利用PCA对输入网络的部分特征进行变换并降维;最后,将经过PCA处理后得到的全新数据信息作为模型的输入,并结合Adam梯度下降算法进行训练,输出负荷预测的结果。通过仿真实验结果表明,包含SRU的混合模型在全部测试集样本上的MAPE为2.126%,远低于仅有子网的单一模型与包含DNN的混合模型,而与包含LSTM的混合模型相比,训练时间却降低了22.74%,同时PCA的应用也使得模型的收敛速度加快,极大地减小了训练轮数。

关键词: 短期负荷预测;主成分分析;基础子网;简单循环单元;混合模型

中图分类号: TM715;TP18 **文献标识码:** A **国家标准学科分类代码:** 470.40

Short-term load forecasting method based on hybrid neural network

Ren Xuan Wang Qingnian Shang Bao Jiang Hongwei Chang Le
(School of Information Engineering, Nanchang University, Nanchang 330036, China)

Abstract: Load forecasting is crucial to the economic operation of the power grid. In order to improve the accuracy of short-term load forecasting and reduce the training time of the hybrid neural network, a short-term load forecasting method based on basic network with multilayer perceptron (MLP), simple recurrent units (SRU) and principal component analysis (PCA) is proposed. Firstly, the method considers various power load influencing factors to establish input feature sets of the load forecasting task, and uses PCA to transform and reduce the historical load and temperature characteristics which are the original inputs of the network. Then, the method uses new data information obtained after PCA as the inputs of the hybrid neural network model, and trains the network with Adam gradient descent algorithm. Finally, the outputs of the proposed model are load forecasting results. The results of the experiments show that the MAPE of the hybrid model including SRU on all test set samples is 2.126%, which is much lower than that of the single model with only basicnet and the hybrid model including DNN, and compared with the hybrid model including LSTM, the training time is reduced by 22.74%, and the application of PCA also accelerates the convergence of the model, which greatly reduces the number of training epochs.

Keywords: short-term load forecasting; principal component analysis; basicnet; simple recurrent units; hybrid model

0 引 言

电力系统负荷预测对电网的安全正常运行至关重要,同时也关系到系统的经济运行。随着电力工业逐渐市场化,负荷预测更是成为电力公司的核心业务之一。

针对短期负荷预测问题,到目前为止国内外的研究人员提出了多种方法,大体上可以分为两大类,一类是以数学理论为主的传统经典的方法,以多元线性回归^[1]、ARIMA^[2]和卡尔曼滤波法^[3]等为代表,此类方法以纯数学

推导为主,缺点是难以反映原始序列中非线性因素的影响;另一类则是以人工智能理论为主,如文献[4]应用专家系统预测方法实现电力负荷预测,但是专家系统作为一种较早的人工智能范式,本质上还是通过人预先去学习理解数据之间的关系,从而实现预测。支持向量机(SVM)是日前在分类与回归分析中常用的机器学习算法,也被应用于负荷预测领域^[5-6]。文献[7]提出一种结合相似日的GOA-SVM负荷预测模型,即利用蝗虫优化算法(GOA)优化SVM模型,结果显示预测精度比遗传算法(GA)与粒子

群算法(PSO)优化后的 SVM 模型更高。文献[8]提出了一种基于自编码器与果蝇优化算法(FOA)优化极限学习机(ELM)的短期电力负荷预测方法,并通过实验证明该方法的预测精度满足要求。但是 SVM 与 ELM 均不擅长处理大规模的数据,因此相关研究中采用的训练样本与测试样本数量也比较少。

近年来,深度学习开始成为人工智能领域的热点,并广泛应用在自然语言处理、计算机视觉等方面,同样也开始应用在负荷与能源预测领域。典型的经典网络如多层感知器(MLP)^[9]、深度置信网络(DBN)^[10]开始被应用在电力负荷预测上。文献[11]提出一种基于深度残差网络(ResNet)^[12]的短期负荷预测模型,并取得了较高的预测精度。文献[13-14]使用长短期记忆网络(LSTM)进行短期负荷预测,文献[15]则搭建了一种基于卷积神经网络(CNN)和 LSTM 的混合预测模型。其中,得益于对时间序列数据的优异学习能力,LSTM 和门控循环单元(GRU)^[16]这类循环神经网络(RNN)应用得最为广泛。但是,不论是深层网络还是此类基于门控的 RNN 的训练代价都很大,面对海量数据样本时表现得更为明显。因此在保证深度学习模型预测精度的同时,应当合理地降低模型的训练代价。

基于以上背景,本文利用无监督学习算法 PCA 对输入网络的特征进行投影变换并降维,进而起到特征增强的作用,再构建基于 MLP 结构的子网与 SRU 组成的混合神经网络模型。SRU^[17]作为 LSTM 与 GRU 的优化改进网络,解除了内部门状态对前一时刻状态的依赖,使得内部单元可以并行计算,从而加快训练速度。本文以美国新英格兰地区的实际负荷数据为算例,通过仿真实验证明,该方法的预测精度优于传统神经网络模型,而与 LSTM 模型对比,SRU 在保持精度的同时,训练时间更短。

1 模型原理

1.1 多层感知器

多层感知器(MLP)是深度学习中的典型网络,是前馈神经网络(FNN)中的一种,隐藏层特别深时又叫作深度神经网络(DNN)。MLP 的结构如图 1 所示,第一层为输入层,最后一层为输出层,中间为隐藏层,每个神经元都与前后两层的所有神经元密集连接。其中,前后两层的数学关系为其中,前后两层的数学关系为:

$$a^{[l]} = g^{[l]}(W^{[l]}a^{[l-1]} + b^{[l]}) \quad (1)$$

式中: $W^{[l]}$ 和 $b^{[l]}$ 分别为 l 层的权重和偏置,为该层的激活函数, $a^{[l-1]}$ 为该层的输入,即 $l-1$ 层的输出, $a^{[l]}$ 为该层的输出。

1.2 简单循环单元

对于时间序列数据,RNN 相比于其他类型的网络具有更强的学习能力,但是也面临梯度消失等问题,且对于长期依赖问题表现欠佳。虽然之后诞生了 LSTM 与 GRU,但是此类具有门控机制的 RNN 虽然克服了普通 RNN 的缺

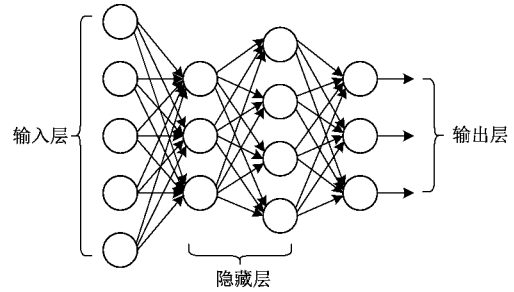


图 1 MLP 结构图

陷,但内部结构却更为复杂,同时其内部计算单元依然有着时序上的依赖,故无法并行计算。而简单循环单元(SRU)的出现则解决了这个问题,SRU 通过解除内部计算对上一时刻隐状态的依赖,并对整体的门结构进行改进,进而确保网络具有 LSTM 的性能,其内部结构如图 2 所示。

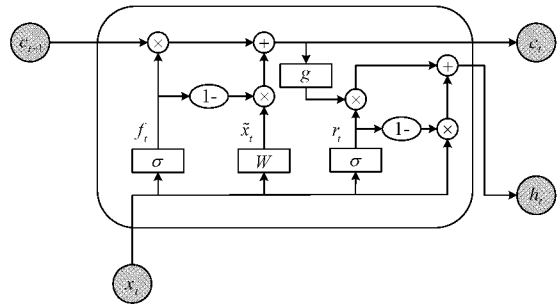


图 2 SRU 结构图

$$\tilde{x}_t = Wx_t \quad (2)$$

$$f_t = \sigma(W_f x_t + b_f) \quad (3)$$

$$r_t = \sigma(W_r x_t + b_r) \quad (4)$$

$$c_t = f_t \odot c_{t-1} + (1 - f_t) \odot \tilde{x}_t \quad (5)$$

$$h_t = r_t \odot g(c_t) + (1 - r_t) \odot x_t \quad (6)$$

上述公式即为 SRU 的内部计算原理, f_t 与 r_t 分别为遗忘门与重置门, x_t 为当前时刻的输入, h_t 为当前时刻的隐状态输出, c_{t-1} 和 c_t 分别为上一时刻和当前更新后的单元状态; W 、 W_f 与 W_r 分别为权重矩阵, b_f 与 b_r 为各自对应的偏置向量, σ 表示 sigmoid 函数, \tanh 为双曲正切函数。其中, $g(\cdot)$ 为非线性函数,一般也设置为 \tanh 函数。

从上述计算过程可以看出,SRU 也有门控机制,但是解除了对 h_{t-1} 的依赖,且在计算隐状态 h_t 时引入了高速公路网络(highway network)^[17],如式(6)所示,这能够使得训练更深层的 SRU 网络时不出现梯度消失或模型退化等问题。

1.3 PCA 原理

PCA 的思想是通过一系列的数据变换操作,将 n 维特征映射到 k 维上,这 k 维是全新构造出来的特征数据,而不是简单地从 n 维特征中去除其余 $n-k$ 维特征,并且全新的 k 维特征线性不相关。因此,PCA 常被用于数据预处理,降低输入数据的维度,提高网络的训练效率^[18]。PCA 的具体步骤如下:

1) 数据中心化。假设有 m 个样本, 每个样本有 n 个特征, 则样本数据可以用 $m \times n$ 的矩阵 \mathbf{X} 表示, 即:

$$\mathbf{X} = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{m1} & \cdots & x_{mn} \end{bmatrix} \quad (7)$$

接着计算每列特征的均值, 再用原样本矩阵 \mathbf{X} 的每个元素减去对应列的均值, 继而得到矩阵 \mathbf{Y} 。

2) 计算 \mathbf{Y} 的无偏样本协方差矩阵 \mathbf{R} , \mathbf{R} 的计算方式如下:

$$\mathbf{R} = \frac{1}{m-1} \mathbf{Y}^T \mathbf{Y} \quad (8)$$

再求得 \mathbf{R} 的特征值 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ 及其对应的特征向量 u_1, u_2, \cdots, u_n 。

3) 计算主成分的方差贡献率和累计方差贡献率, 确定主成分个数。

$$\begin{cases} \eta_i = \frac{\lambda_i}{\sum_{i=1}^n \lambda_n} \times 100\% \\ \eta_{\Sigma}(k) = \sum_{i=1}^k \eta_i \end{cases} \quad (9)$$

式中: η_i 为第 i 个主成分的方差贡献率, $\eta_{\Sigma}(k)$ 为前 k 个主成分的累计方差贡献率。通常当累计方差贡献率达到 85%~95% 时, 则前 k 个主成分即可保留原始数据的绝大部分特征。

4) 计算前 k 个主成分的最终样本值, 即

$$\mathbf{Z}_{m \times k} = \mathbf{Y}_{m \times n} \mathbf{U}_{n \times k} \quad (10)$$

式中: $\mathbf{U}_{n \times k}$ 为前 k 个主成分对应的特征向量构成的矩阵, 即 $\mathbf{U}_{n \times k} = [u_1, u_2, \cdots, u_k]$ 。 $\mathbf{Z}_{m \times k}$ 为降维后得到的样本矩阵, 特征由 n 维降至 k 维。

2 特征工程与数据处理

2.1 负荷特征的选取

文中选用美国 ISO-NE 公司提供的 2010 年~2013 年的负荷数据作为实际算例, 在仅有历史负荷数据的基础上, 添加了温度, 季节, 是否工作日, 是否节假日这 4 类特征。所选取的特征如表 1 所示。所有特征均未考虑到待预测日中的实际负荷, 未产生时间泄露, 可以做到日前负荷预测。

2.2 数据预处理及精度评价指标

1) 数据归一化

深度学习模型中的输入数据通常维数众多且量纲不同, 数值差距大, 在模型训练前通常需要对原始样本数据预处理, 将数据值保持在 $[0, 1]$ 范围内, 本文采用简单的最大值归一化方法, 即

$$x^* = \frac{x}{x_{\text{peak}}} \quad (11)$$

式中: x 为原始数据, x_{peak} 为该维度数据中的峰值, x^* 为归一化后得到的数据。

表 1 负荷预测特征集

影响因素	特征描述	维度
历史负荷	每隔 1 周对应相同时刻的负荷	12
	过去 7 天与预测点相同时刻的负荷	7
温度	每隔 1 周对应相同时刻的温度	12
	过去 7 天与预测点相同时刻的温度	7
	预测点的温度	1
日期因素	季节的独热编码	4
	是否工作日的独热编码	2
	是否节假日的独热编码	2

2) 精度评价指标

为评价深度学习预测模型结果的优劣, 本文选取回归分析中常用的平均绝对百分比误差(MAPE)与均方根误差(RMSE)作为预测结果的评估指标, 其中 MAPE 可以从总体上评估单日或全部测试集预测值的准确度, 而 RMSE 则对预测异常值相当敏感, 如果对单个或某几个点的预测偏差极大, 将对 RMSE 产生较大影响。两种评估指标的数学表达式如下:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \quad (12)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (13)$$

式中: y_i 与 \hat{y}_i 分别为待测日第 i 个采样点负荷的真实值和预测值, n 为待测日采样点个数。

2.3 主成分分析

根据表 1 中的输入特征, 选取与负荷和温度有关的特征进行 PCA 降维, 采取独热编码的日期因素则加以保留。以对应凌晨零点的负荷特征为例, 部分主成分的特征值、方差贡献率与累计方差贡献率如表 2 所示。

表 2 方差贡献率与累计方差贡献率

主成分	方差贡献率/%	累计方差贡献率/%
1	60.832 9	60.832 9
2	16.797 9	77.630 8
3	3.981 6	81.612 4
4	2.078 8	83.691 2
5	1.858 7	85.549 9
6	1.273 5	86.823 4
7	1.147 9	87.971 3
8	0.952 4	88.923 7
9	0.924 3	89.848 0
10	0.880 5	90.728 5
16	0.006 7	95.199 6
⋮		
29	0.001 8	99.145 5

由表 2 可知,前 3 个主成分的方差贡献率较大,累计起来已超过 80%,前 10 个主成分的方差贡献率则达到了 90%,到第 29 个主成分时,累计贡献率已达到 99%。

3 模型实现细节与超参数设置

3.1 模型概述

本文提出的混合神经网络的大体结构如图 3 所示。先利用 PCA 将各子网的输入特征进行降维,然后将降维后得到的低维数据信息输入到 24 个基本子网,再将 24 个子网的输出送入 SRU 网络进行再训练,即对粗预测序列的较准。如果仅让各子网单独输出预测结果,那么整个输出序列缺乏整体性,预测曲线往往不够平滑。

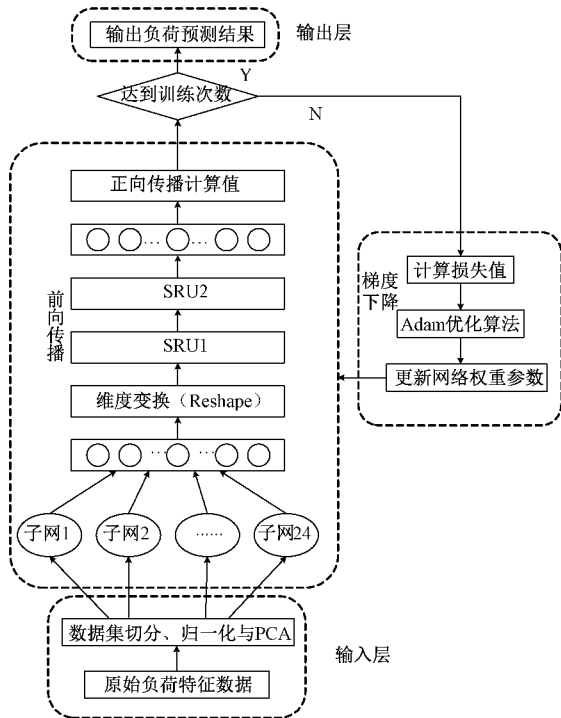


图 3 混合模型结构示意图

3.2 模型实现细节与超参数设置

选择合理的激活函数与损失函数、以及各种超参数的设置对模型的最终性能以及泛化能力至关重要,因此必须通过反复试验确定其数值。本文先不尝试使用 PCA,而是通过大量对比实验,先确定混合神经网络的各种超参数、各层激活函数、损失函数与梯度下降优化算法。

1)对于单个基础子网,其结构如图 4 所示。

历史负荷与温度输入层后面接着 2 个隐藏层,每层神经元的个数设置为 10;而工作日和季节的输入层的后面有一个隐藏层,神经元个数为 6;输出拼接之后,后续还有两个隐藏层,神经元数量均设置为 16,节假日因素则在倒数第一个隐藏层前输入进网络,并与前一层的输出相拼接。

2)SRU 网络共 2 层,第一层的输出维度为 64,即单元数为 64,第二层设定为 32。SRU 的后面还有一层全连接

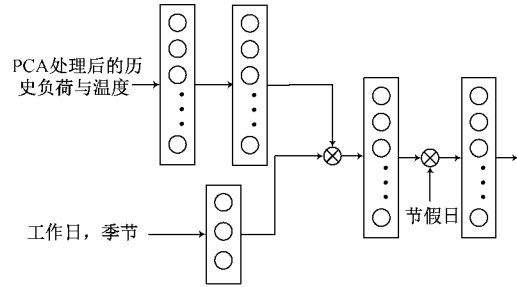


图 4 基础子网的结构

层,神经元个数也为 64。网络最后的输出层设定为 24 个神经元。

3)激活函数与损失函数

各子网的隐藏层均采用的是非线性激活函数 SELU,其表达式如下:

$$\text{SELU}(x) = \lambda \begin{cases} x, & x > 0 \\ \alpha(c^x - 1), & x \leq 0 \end{cases} \quad (14)$$

式中: λ 与 α 都是可调参数,一般给定 $\lambda = 1.0577$, $\alpha = 1.6733^{10}$, SELU 函数是 ReLU 激活函数的改进版本,克服了 ReLU 函数的一些缺陷,非常与采用截断正态分布的网络权重初始化方式结合使用,本文采用 lecun_normal 初始化方式。而 SRU 的激活函数选用默认的 tanh 函数,各层则采用 sigmoid 函数。

通过与回归分析中常用的各类损失函数进行对比,最终选择如下激活函数:

$$L = \frac{1}{NH} \sum_{i=1}^N \sum_{h=1}^H \frac{|y_{(i,h)} - \hat{y}_{(i,h)}|}{y_{(i,h)}} + \frac{1}{2N} \sum_{i=1}^N \max(0, \max_h \hat{y}_{(i,h)} - \max_h y_{(i,h)}) + \max(0, \min_h y_{(i,h)} - \min_h \hat{y}_{(i,h)}) \quad (15)$$

式中: N 为天数, H 为每天的负荷采样点总数, $y_{(i,h)}$ 与 $\hat{y}_{(i,h)}$ 分别为第 i 天的第 h 个负荷采样点对应的负荷真实值与预测值, $\max_h y_{(i,h)}$ 和 $\min_h y_{(i,h)}$ 表示第 i 天实际负荷的最大值和最小值,同理可知 $\max_h \hat{y}_{(i,h)}$ 和 $\min_h \hat{y}_{(i,h)}$ 。

可以明显看出,式(15)前一项即为前文介绍的评价指标中的 MAPE,后一项可以看作是一个惩罚项,因为在训练初期模型的输出往往与真实负荷曲线差距较大,该项通过对模型施加惩罚,从而能够使训练初期损失下降得更快。

4)优化训练策略

模型的训练采用小批量训练 (Mini-Batch) 策略与 Adam 优化算法,小批量训练是深度学习中最常使用的优化方法,该方法是将训练集中的样本分成若干份,并且在分割前可以选择将训练集样本顺序打乱,从而使得一次训练可以迭代多次,而 Adam 算法是一种自适应学习率的梯度下降算法,其计算流程如下:

计算当前时间步 t 时刻的梯度:

$$g = \frac{1}{m} \nabla_{\theta} \sum_{i=1}^m L(f(x_i, \theta), y_i) \quad (16)$$

式中: m 为样本个数, ∇_{θ} 计算符指代损失函数 $L(\cdot)$ 对网络中参数 θ 的梯度, g 为计算后得到的梯度。

更新有偏一阶矩估计和二阶矩估计:

$$\begin{aligned} s &\rightarrow s' = \beta_1 s + (1 - \beta_1) g \\ r &\rightarrow r' = \beta_2 r + (1 - \beta_2) g^2 \end{aligned} \quad (17)$$

式中: β_1 与 β_2 为矩估计的衰减速率, 分别为 0.9 和 0.999, s 与 r 为矩估计, 在 $t=0$ 时设定为 0, 即它们的初始值为 0。

修正一阶矩估计和二阶矩估计的偏差:

$$\begin{aligned} \hat{s} &= \frac{s}{1 - \beta_1^t} \\ \hat{r} &= \frac{r}{1 - \beta_2^t} \end{aligned} \quad (18)$$

更新学习参数:

$$\theta \rightarrow \theta' = \theta - \alpha \frac{\hat{s}}{\sqrt{\hat{r}} + \delta} \quad (19)$$

式中: α 为全局学习率, 默认为 0.001, δ 是一个很小的常数, 用于数值稳定从而避免分母为 0, 默认设定为 10^{-8} 。

相较于传统梯度下降算法, Adam 具有惯性保持和环境感知的能力, 通常被认为对超参数的选择相当鲁棒, 尽管有时学习率需要从建议的默认值修改。因为该方法记录了梯度的一阶矩, 即过往梯度与当前梯度的滑动平均, 每次迭代时即可实现梯度的稳定过渡, 可以适应不同的目标函数。

4 仿真实验与结果分析

4.1 实验环境与数据集划分

本文的编程环境为 Anaconda, 编程语言为 Python3.6, 采用 Keras 深度学习框架, 并以 Tensorflow 为后端。

以 2010 年~2012 年 3 年的数据作为训练集, 考虑到构建输入特征所需的最长 12 周以前的历史负荷, 以及构建验证集所需的 10% 的样本数量, 最终的训练样本为 911 天, 2013 年全年 365 天作为测试样本。

4.2 PCA 的应用与主成分选取

首先保留协方差矩阵所有的特征值与对应的特征向量, 并将 39 维的历史符合与温度特征向特征向量上进行投影, 得到 39 个主成分, 并输入到混合神经网络模型进行训练。即仅利用 PCA 进行数据变换, 先不剔除部分贡献率低的主成分分量, 并和利用原始数据输入的方法进行对比, 图 5 为两模型在验证集上的损失下降对比图。为方便观测, 图中的纵轴仅包含部分区间。由图 5 可知, 未经过 PCA 处理的模型收敛得更慢, 经过接近 450 代训练后才能收敛, 且模型在训练过程中表现不稳定, 验证集损失波动幅度大, 鲁棒性较差, 虽然损失有时也能低于使用 PCA 的模型, 但是总体效果不好。

然而, 经过 PCA 处理的模型收敛速度大大加快, 并且验证集 loss 上的波动幅度明显更小。由此可知, 在本任务

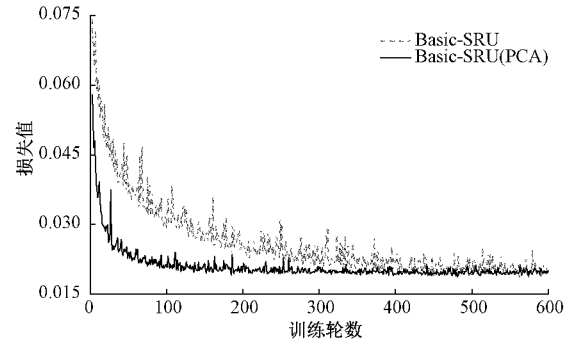


图 5 验证集损失下降对比

中神经网络更擅长学习经过 PCA 处理后的主成分分量与样本标签间的关系, 即使保留所有主成分分量不进行降维, 也能显著减小训练代价, 并提升模型的鲁棒性。

考虑到最后几个主成分的方差贡献率很小, 因此最终选择累计方差贡献率为 99% 时的所有主成分作为输入。

4.3 预测实验与结果分析

为评估模型的预测效果以及泛化能力, 还需要利用测试集样本对模型进行测试, 并与其他模型进行对比。将仅有 24 个子网且未使用 PCA 的预测模型记为模型 1; 各子网与普通 DNN 组合而成的混合模型 (DNN 即为有较多隐藏层的 MLP) 记为模型 2; 各子网与 LSTM 组成的混合模型记为模型 3; 模型 4 即为本文模型。本文模型的超参数设置在前文已有叙述, 模型 3 中 LSTM 的各层单元数与 SRU 完全一致, 模型 2 中 DNN 的隐藏层为 10, 神经元数目交替设置为 20 和 24。使用了 PCA 降维的模型训练轮数设定为 280, 未使用 PCA 的模型训练轮数设置为 500, 所有模型的子网部分均一样, 且都采用相同的优化训练方法。

首先从测试集样本中随机选取一工作日进行测试, 各模型的预测曲线与逐点误差如图 6~7 所示。由图可知, 各模型在工作日上的预测结果均不错, 不过仅有子网的模型相邻点之间的 MAPE 有较大的起伏, 即预测曲线不够平滑, 输出缺乏一定的整体性, 总体的 MAPE 也是最高的。而其他混合模型的预测曲线则较为平滑, 其中 SRU 与 LSTM 的预测效果差不多, 在该日的各时间段表现都比较平衡, 总体的 MAPE 水平也都低于 DNN。

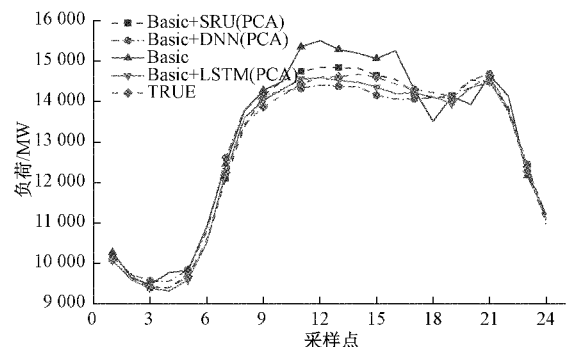


图 6 各模型工作日预测曲线对比图

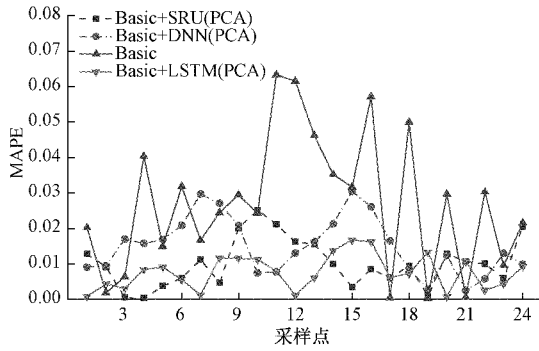


图 7 各模型工作日预测误差对比图

图 8~9 绘制了各模型在某一休息日兼节假日的预测曲线与逐点误差对比图。结合两图可知,休息日的日间负荷峰值相比于工作日远小于夜间负荷值,但是各模型的预测值都大于真实值,各模型的主要误差也集中在中午时段,其中仅有基础子网的模型的误差最大,某些点的 MAPE 甚至超过了 7%。混合模型中表现较好的依旧是 SRU 与 LSTM,在误差较大的区间,其 MAPE 也能控制在 3% 以下。

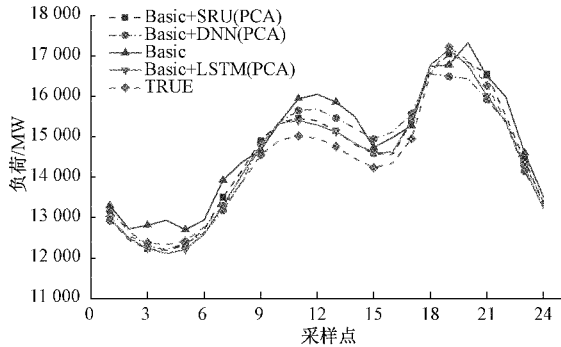


图 8 各模型休息日预测曲线对比图

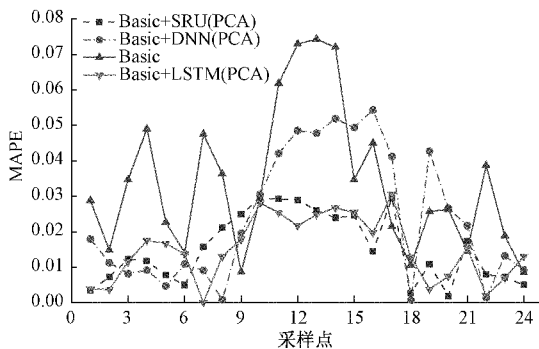


图 9 各模型休息日预测误差对比图

为了更全面地评估模型的泛化能力,利用全部测试集样本对各模型进行测试,测试集样本包括 2013 年全年 365 个样本。表 3 给出了各模型在测试集上的表现以及在训练集上的训练时间。

由表 3 可知,混合模型的 MAPE 和 RMSE 均比仅有 24 个子网(模型 1)的模型低,说明了混合模型比单一模型

表 3 各模型在全部预测集上的预测结果

模型	MAPE/%	RMSE/MW	训练用时/s
模型 1	2.815	621.91	316.23
模型 2(PCA)	2.432	524.85	225.36
模型 3(PCA)	2.076	471.16	326.39
模型 4(PCA)	2.126	476.01	252.17

的预测性能更好,同时可以发现,仅由 24 个子网组成的模型在全年测试集上的 MAPE 为 2.821%,误差在 3% 以下,这充分说明文中选取的输入特征、以及针对各点建立子网的措施合理有效。在 3 种混合模型中,本文提出的混合模型的预测精度达到了 97.87%,相比于 DNN,SRU 的 MAPE 降低了 0.306%,RMSE 降低了 9.31%,而与 LSTM 模型相比,MAPE 和 RMSE 仅增大了 0.05% 与 4.85 MW,但训练时间减少了 22.74%,说明了 SRU 能够实现加速,减小模型的训练代价。同样地,在上一节 PCA 实验部分已经证明,使用了 PCA 的模型能够显著减小训练轮数,进而减小网络的训练代价。

5 结 论

本文提出一种新的短期负荷预测方法,利用更为轻量化的 SRU 网络替代传统 LSTM 网络,并且结合非监督学习算法 PCA 实现对输入特征的变换与降维,再结合 Adam 等其他深度学习优化技术。通过真实数据仿真实验表明,与其他模型相比,SRU 能够降低网络的训练时间,而 PCA 的应用可以有效加速网络的收敛,并且提高网络的鲁棒性。与传统模型相比,文中方法在实际工程应用中具有较好的应用价值。

在未来工作中,考虑额外加入预测电价因素实现短期负荷预测,看是否能进一步提高预测精度。同时,可以考虑将文中模型应用至中期负荷预测乃至其他能源预测领域。

参考文献

- [1] 谢毓广,刘童,陈凡,等.基于置信度理论和多元线性回归方程的电网负荷预测[J].电工技术,2018(13):42-45.
- [2] 刘士进,孙立华,郭鹏.基于 EEDM-Arima 算法的某地区电力系统负荷预测及算例分析[J].电子测量技术,2020,43(7):185-188.
- [3] 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.
- [4] EROSHENKO S A, POROSHIN V I, SENYUK M D, et al. Expert models for electric load forecasting of power system[C]. 2017 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering(EIConRus), IEEE, 2017: 1507-1513.
- [5] 赵佩,代业明.基于实时电价和加权灰色关联投影的

- SVM 电力负荷预测[J]. 电网技术, 2020, 44(4): 1325-1332.
- [6] 陈家瑞, 陈忠孝, 秦刚, 等. 基于 PSO 算法与 SVR 算法在企业直流配电网短期负荷预测的研究[J]. 国外电子测量技术, 2020, 39(12): 70-73.
- [7] BARMAN M, CHOUDHURY N B D, 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.
- [8] 张淑清, 要俊波, 张立国, 等. 基于改进深度稀疏自编码器及 FOA-ELM 的电力负荷预测[J]. 仪器仪表学报, 2020, 41(4): 49-57.
- [9] ASKARI M, KEYNIA F. Mid-term electricity load forecasting by a new composite method based on optimal learning MLP algorithm[J]. IET Generation, Transmission & Distribution, 2020, 14(5): 845-852.
- [10] ZHU Y, SUN C, ZHENG W, et al. Short-term Load Forecasting Based on Kmeans and FR-DBN Models[C]. 2020 10th International Conference on Power and Energy Systems(ICPES), IEEE, 2020: 145-150.
- [11] CHEN K, CHEN K, WANG Q, et al. Short-term load forecasting with deep residual networks[J]. IEEE Transactions on Smart Grid, 2019, 10(4): 3943-3952.
- [12] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [13] KONG W, DONG Z Y, JIA Y, et al. Short-term residential load forecasting based on LSTM recurrent neural network[J]. IEEE Transactions on Smart Grid, 2017, 10(1): 841-851.
- [14] KWON B S, PARK R J, SONG K B. Short-term load forecasting based on deep neural networks using LSTM layer [J]. Journal of Electrical Engineering & Technology, 2020, 15(4): 1501-1509.
- [15] ALHUSSEIN M, AURANGZEB K, HAIDER S I. Hybrid CNN-LSTM model for short-term individual household load forecasting[J]. IEEE Access, 2020, 8: 180544-180557.
- [16] JUNG S, MOON J, PARK S, et al. An attention-based multilayer GRU model for multistep-ahead short-term load forecasting [J]. Sensors, 2021, 21(5): 1639.
- [17] LEI T, ZHANG Y, WANG S I, et al. Simple recurrent units for highly parallelizable recurrence [C]. Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing Brussels: ACL, 2018: 4470-4481.
- [18] 卜迟武, 刘涛, 赵博, 等. CFRP/Al 蜂窝结构缺陷红外热图特征的多结构形态学-PCNN 识别[J]. 电子测量与仪器报, 2021, 35(9): 222-228.
- [19] KLAMBAUER G, UNTERTHINER T, MAYR A, et al. Self-normalizing neural networks [C]. Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017: 972-981.

作者简介

任轩(通信作者), 硕士研究生, 主要研究方向为深度学习、电力系统负荷预测。

E-mail: 514557634@qq.com

汪庆年, 教授, 主要研究方向为智能电机。

E-mail: 2661738618@qq.com

尚宝, 硕士研究生, 主要研究方向为光伏功率预测。

E-mail: 1649039714@qq.com

姜宏伟, 硕士研究生, 主要研究方向为储能系统优化配置。

E-mail: 2313484182@qq.com

常乐, 硕士研究生, 主要研究方向为电力系统优化调度。

E-mail: 736281799@qq.com