

基于小波变换和深度学习的短期天然气负荷预测研究

田文才¹, 乔伟彪¹, 周国峰¹, 刘 伟²

(1. 华北水利水电大学 环境与市政工程学院, 河南 郑州 450046; 2. 重庆大学 资源与安全学院, 重庆 400044)

摘 要: 随着天然气在能源消耗中占比越来越大, 如何准确预知未来的天然气消耗量, 对天然气资源合理规划具有重大意义。针对此问题, 提出一种基于小波变换和深度学习的短期天然气负荷预测模型。首先对所收集的天然气负荷数据利用不同小波变换进行分解, 之后对其进行归一化处理; 其次利用深度学习算法对数据进行训练与预测; 然后利用小波重构对预测的数据分别进行整合; 最后以平均绝对百分误差、平均绝对误差和均方根误差为评价指标, 评价不同小波变换的预测结果, 计算最优小波变换的最优阶数和层数。结果表明, Fk小波变换第22阶第6层相对于其他小波变换和直接利用LSTM进行预测具有更高的预测精度。

关键词: 深度学习; LSTM; 小波; 阶数; 层数

中图分类号: TE82

文献标志码: A

doi:10.3969/j.issn.1672-6952.2021.05.016

Research on Short-Term Natural Gas Load Forecasting Based on Wavelet Transform and Deep Learning

Tian Wencai¹, Qiao Weibiao¹, Zhou Guofeng¹, Liu Wei²

(1. School of Environmental and Municipal Engineering, North China University of Water Resources and Electric Power, Zhengzhou Henan 450046, China; 2. School of Resources and Safety Engineering, Chongqing University, Chongqing 400044, China)

Abstract: As natural gas accounts for an increasing proportion of energy consumption, how to accurately predict the future natural gas consumption is of great significance to the rational planning of natural gas. For this problem, a short-term natural gas load forecasting model based on wavelet transform and deep learning was proposed. First, the collected natural gas load was decomposed by using different wavelets, and then normalized it. Secondly, the data was trained and predicted by using the deep learning algorithm Long Short-Term Memory (LSTM); then the predicted data was separately integrated by using wavelet reconstruction. Finally, the average absolute percentage error, average absolute error and root mean square error were used as evaluation indicators to evaluate the prediction results of different wavelets, and the optimal order and number of layers of the optimal wavelet were calculated. The examples show that the 22nd-order 6th layer of Fk wavelet transforms has higher prediction accuracy than other wavelets transforms and direct use of LSTM for prediction.

Keywords: Deep learning; LSTM; Wavelet; Orders; Number of layers

随着全球环境日益恶化, 寻找清洁能源来代替煤、石油等重污染能源已经迫在眉睫, 天然气作为一种清洁能源越来越受到人们的关注。2017年全球天然气消耗量为 $3.62 \times 10^{12} \text{ m}^3$, 其增速为 2.2%, 预计到 2022 年全球天然气消耗量将达到 4.00×10^{12}

m^3 。我国天然气消耗量在 2023 年将达到全球总量的 40%, 国内天然气产量和进口量不断增加^[1-6]。只有对天然气消耗量进行准确的预测, 提前做好相应设施的规划、建设, 才能对其实现最大程度上的利用, 更好地保障人们的生活质量。

收稿日期: 2021-03-29 修回日期: 2021-04-06

基金项目: 重庆市基础研究与前沿探索项目(cstc2018jcyjAX0441)。

作者简介: 田文才(1998-), 男, 硕士研究生, 从事供热、供燃气、通风及空调方面的研究; E-mail: thl980302@163.com。

通信联系人: 乔伟彪(1985-), 男, 博士, 讲师, 从事天然气管网系统智能化调度方面的研究; E-mail: wbq0408@163.com。

天然气负荷预测可以分为年负荷预测、月负荷预测和日负荷预测,其预测时间越长,预测所产生的误差就越大^[7]。截至目前,学者们在天然气负荷预测方面进行的研究可分为物理方法和统计方法两种。物理方法是利用温度、风速、云量等有关的物理信息建立预测模型,并对天然气消耗量进行预测,其预测误差较大。统计方法是利用大量的历史数据建立系统的输入和输出之间的关系,对天然气消耗量进行预测,如人工神经网络(Artificial Neural Network, ANN)模型、灰色预测模型、多层感知机(Multilayer Perceptron, MLP)、支持向量机模型和 Elman 神经网络模型等^[8-14]。但是,这些方法对处理少量数据具有优越性,在处理大量历史数据时需要提前对数据进行筛查,因此应用范围受到限制^[15],而且其预测精度也有待进一步提高。为了进一步提高预测精度,国内外许多学者采用了组合模型预测方式^[16]。乔伟彪等^[17]提出了基于混沌理论和 Volterra 自适应滤波器的天然气负荷预测模型,其预测结果与人工神经网络、傅里叶级数预测模型的预测误差相比得到了提高。P.Brinks 等^[18]提出了基于结构校准支持向量机回归预测模型(Structure-Calibrated Support Vector Regression, SC-SVR),其平均绝对百分比误差为 2.36%,但是其预测结果没有达到令人满意的效果。

近些年来,针对具有时间序列的数据处理问题,小波分析得到了快速发展,其在处理多尺度空间和时间上的数据特征和趋势具有很强的泛化能力。深度学习算法中的 LSTM(Long Short-Term Memory)可以对数据进行更深层次的学习,发现数据中的内在规律,避免主观因素的影响,保证预测结果的可靠性。但是,目前大多学者只利用单个小波对数据进行分解,没有利用不同小波与 LSTM 算法相结合。为了提高天然气负荷预测精度,本文提出了一种基于小波变换和深度学习的短期天然气负荷预测模型。利用不同小波变换对天然气负荷数据进行分解,使用 LSTM 算法进行训练与预测,分别以平均绝对百分比误差(Mean Absolute Percentage Error, MAPE)、平均绝对误差(Mean Absolute Error, MAE)和均方根误差(Root Mean Square Error, RMSE)作为评价指标对预测结果进行比较,指出其最优小波的最优阶数和最优层数。研究结果对国家政策的制定和天然气公司运行规划具有重要意义。

1 基本理论及仿真模型

1.1 小波理论

小波变换(Wavelet Transform, WT)可以利用

伸缩平移方法对时间序列信号实现在高频处对时间细分;低频处对频率细分^[16],从而使信号的特征和趋势更加明显,克服 Fourier 变换的弊端,在非平稳的时间序列方面是一种很有效的方法。

1.1.1 小波函数 小波函数经母小波的伸缩和平移可以构成新的子小波,假设小波函数 $\lambda(x)$,即 $\int_{-\infty}^{+\infty} \lambda(x) dx = 0$ (低通滤波器),将小波函数 $\lambda(x)$ 进行伸缩和平移构成一簇函数,其表达式为:

$$\lambda_{(a,b)}(t) = \frac{1}{\sqrt{a}} \lambda\left(\frac{t-b}{a}\right) \quad a, b \in R; a \neq 0 \quad (1)$$

式中, $\lambda_{(a,b)}$ 为子小波; t 为平移参数; a 为尺度因子, $a > 0$; b 为平移因子; R 为实数。

1.1.2 小波分解与重构 小波快速分解算法有 Mallat 算法和 ATrous 算法。Mallat 算法作为一种纯数字算法,具有设计简单、计算快速等特点,在处理时间序列数据上具有很大优势^[19]。因此,根据天然气负荷消耗量的序列长度,采用 Mallat 算法,即塔式算法。其表达式为:

$$\begin{cases} y_{j+1} = D_{xj} \\ x_{j+1} = A_{xj} \end{cases} \quad j = 1, 2, 3, \dots, J \quad (2)$$

式中, A 、 D 分别为高通滤波器和低通滤波器; J 为分解层数。

将 x_0 定义为原始信号,将其分解为 $A_1, A_2, A_3, \dots, A_J$ 和 $D_1, D_2, D_3, \dots, D_J$, A_J 和 D_J 分别为原始信号 x_0 在对应小波分解的高频分量和细节分量。其分解过程如图 1 所示。

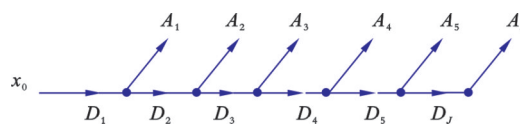


图 1 小波分解

小波重构算法为:

$$x_0 = D_1 + D_2 + \dots + D_J + A_J \quad (3)$$

利用小波函数对所收集的数据进行分解,得到的高频分量和细节分量可以初步学习天然气负荷变化特征和趋势^[20]。

1.2 长短期记忆网络(LSTM)

长短期记忆网络(Long Short-Term Memory, LSTM)是一种可以处理序列数据的循环神经网络。

LSTM 是 RNN 的一个变种,在处理时间序列数据方面具有很大的优势,其在不需要改变算法的情况下可以延长模型的记忆时间,提高模型的预测精度^[22],并且还可以在不同时刻改变设计连接间的权重系数。LSTM 的特点是通过增加输入门限、输出门限和遗忘门限使自循环参数的权重不断发生

变化,在很大程度上避免梯度消失和梯度膨胀问题^[15]。LSTM模型结构如图2所示。

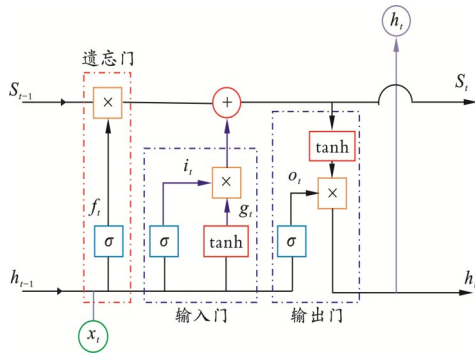


图2 LSTM模型结构

从图2可以看出,LSTM模型结构图由输入门、输出门和遗忘门组成。输入数据由本单元此刻输入状态、上一个单元的输出状态和本单元的上一个时刻状态组成。其中,遗忘门用来选择需要遗忘的信息,输入门用来选择需要保存的信息,而输出门利用Sigmoid函数乘以当前状态下的tanh值作为需要输出的信息。LSTM单元更新过程如式(4)~(9)所示。

$$g_t = \tanh(W_{cx}x_t + W_{ch}h_{t-1} + b_c) \quad (4)$$

$$i_t = \sigma(W_{ix}x_t + W_{ih}h_{t-1} + b_i) \quad (5)$$

$$f_t = \sigma(W_{fx}x_t + W_{fh}h_{t-1} + b_f) \quad (6)$$

$$s_t = f_t \otimes c_{t-1} + i_t \otimes g_t \quad (7)$$

$$o_t = \sigma(W_{ox}x_t + W_{oh}h_{t-1} + b_o) \quad (8)$$

$$h_t = o_t \otimes \tanh(s_t) \quad (9)$$

式中, i_t, g_t 为输入门状态; o_t, f_t 分别为输出门和遗忘门状态; h_t 和 s_t 分别为当前LSTM单元的输出和LSTM的当前状态; σ 为Sigmoid的激活函数; $\otimes$ 为LSTM单元中的元素按位置相乘; $W_{cx}, W_{ch}, W_{ix}, W_{ih}$ 为输入门权重矩阵; W_{fx}, W_{fh} 为遗忘门权重矩阵; W_{ox}, W_{oh} 为输出门权重矩阵。因为LSTM有输入门限、输出门限和遗忘门限以及独立的记忆单元,所以可以很好地读取、更新和储存距离较长的历史信息^[23]。

1.3 WT-LSTM算法流程

步骤1 收集研究对象的历史负荷数据,利用Matlab对数据进行处理。

步骤2 预处理的数据分别利用Coif、Sym、Fk小波函数进行变换,使其变换成高通分量数据和低通分量数据。

步骤3 把小波分解过的数据利用Matlab自带归一化函数进行归一化处理。

步骤4 把所收集数据的80%划分为训练集,剩下的20%的数据划分为测试集。

步骤5 利用构建的LSTM预测模型进行预测,把预测后的结果进行反归一化处理。

步骤6 对预测数据分别按阶、层进行重构,重构的数据分别与LSTM的预测结果和原始数据进行比较,并且计算其误差。

步骤7 误差结果分别以MAPE、MAE和RMSE进行表示,并且对不同小波函数的不同阶层的误差进行比较,通过对其分别进行误差比较,寻找最优小波的最优阶数和层数。

2 数据来源及处理

2.1 数据来源

原始天然气负荷数据由中国石化天然气分公司提供,选取了安阳华润燃气的天然气总负荷,从2017年1月1日到2019年9月2日共计975个日负荷数据,作为本次模型的原始数据。模型中的滑动时间窗口设置为10,所以最终参加训练与预测的数据为965个,把965个原始数据的80%(772个)划分为训练集,剩下的20%(193个)划分为测试集。安阳华润燃气的原始天然气原始负荷如图3所示。

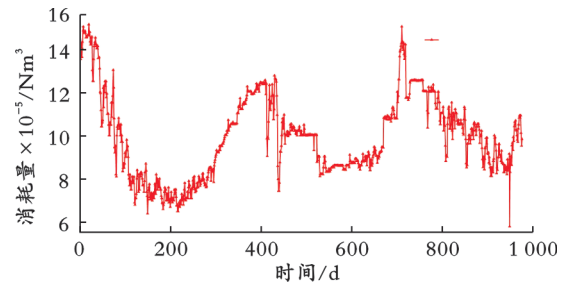


图3 安阳华润燃气原始负荷

从图3可以看出,天然气负荷具有明显的波动性,其日负荷具有随机性和非线性,并且每年成周期性变化,其随时间的变化趋势非常明显。而小波具有强大的时间和空间尺度处理能力,可使研究对象的历史燃气负荷特征更加明显,从而可以获得更高的预测精度。

2.2 数据处理

天然气负荷数据具有明显的波动性、周期性、时序性等特点,而小波分析对于处理时间尺度和空间尺度上的数据具有强大的能力,可以对数据的特征差异进行放大,从而利于预测模型进行学习。利用Coif小波、Sym和Fk小波等对天然气负荷数据进行分析。利用Matlab2018b内置工具箱选取Coif小波、Fk小波和Sym小波对数据进行分解^[24]。通过分析发现,Fk小波相对于其他两种小波可以更好地展现负荷的特征和趋势。Fk小波第22阶分解如图4所示。

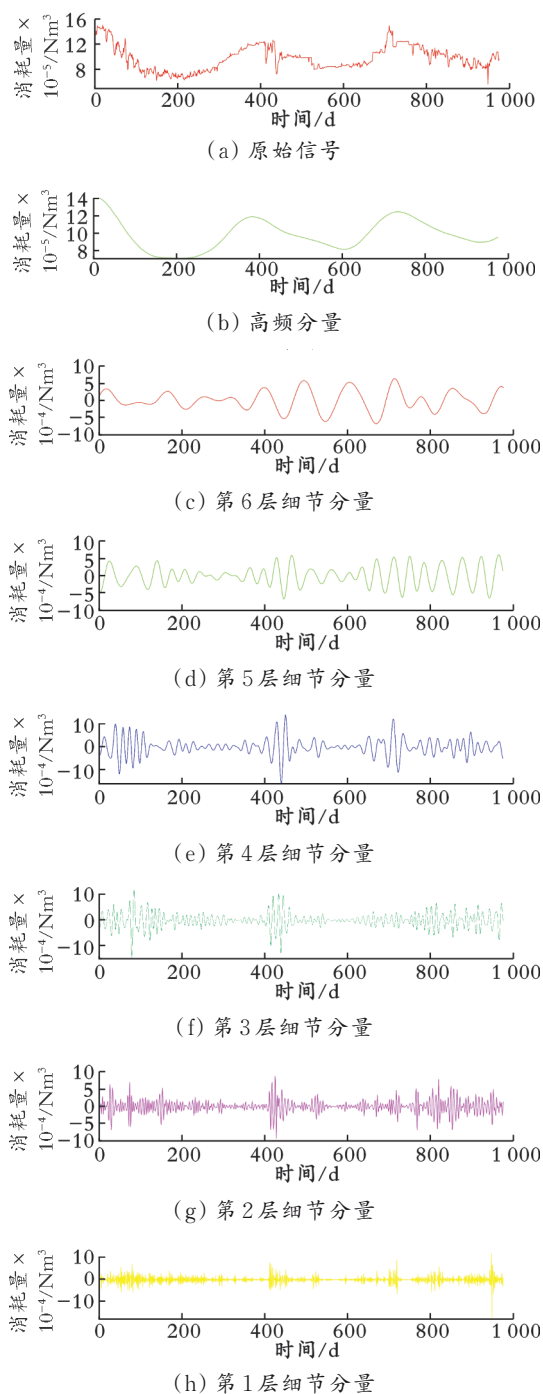


图4 Fk小波第22阶分解图

从图4可以看出,第6层的高频分量很好地反映天然气负荷消耗趋势,其变化规律具有周期性,夏季天然气负荷消耗量相对于冬季较少。第6层的细节分量可以反映天然气负荷的稳定部分。第4层和第5层细节分量具有明显的波动性,其波动值越大表示消耗量越大,即天然气消耗量增幅越大。原始信号具有很大的波动性,可以在一定程度上表示持续的峰-谷时间^[21],当原始数据出现波峰或者波谷时,第1层、第2层和第3层的细节分量都会对应地出现波峰或者波谷。

针对原始数据波动性很大、可能出现量级差异

等问题,利用小波分解之后对其数据进行归一化处理。通过归一化处理消除数据的量级误差,提高数据质量及预测精度。

利用最大最小法对数据进行归一化处理,其表达式为:

$$x_k = (x_k - x_{\min}) / (x_{\max} - x_{\min}) \quad (10)$$

式中, x_k 为当前数据; $x_{\min}$ 为输入数据最小值; $x_{\max}$ 为输入数据最大值。

2.3 误差评价指标

为评价所提出模型预测精确度,选择平均绝对百分比误差(MAPE)、平均绝对误差(MAE)和均方根误差(RMSE)作为此模型的检验标准,其表达式为:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \quad (11)$$

$$\text{MAE} = \sqrt{\frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i|} \quad (12)$$

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

式中, Y_i 为数据中第*i*个数据真实值; $\hat{Y}_i$ 为数据中第*i*个数据预测值; n 为数据的总数。

3 预测结果及仿真分析

本文所采用的WT-LSTM模型输入层为10,LSTM层为20个隐藏单元和1个全连接层;训练参数MaxEpochs设置为250,InitialLearnRate设置为0.01,LearnRateDropPeriod设置为125,LearnRateDropFactor设置为0.1,重复10次实验,然后取平均值。

对安阳华润燃气负荷利用LSTM进行了直接预测,同时分别利用Fk小波、Coif小波和Sym小波进行分解,如表1所示。由表1可知,Sym小波最优阶数和层数是7阶第8层,其MAPE、MAE和RMSE分别是0.81%、87.77和10 510.65;Coif小波最优阶数和层数为5阶第5层,其MAPE、MAE和RMSE分别是0.82%、88.75和11 819.61;Fk小波最优阶数和层数是22阶第6层,其MAPE、MAE和RMSE分别是0.71%、82.19和9 222.59。安阳华润燃气原始负荷、直接预测负荷和不同小波的组合预测负荷如图5所示。从表1和图5还可以看出,其最优小波是Fk小波,最优阶数和层数是第22阶第6层。

为了验证所得结论,选取青岛新奥燃气的总负

荷进行比对,选取从2017年1月1日到2019年9月2日共计975个日负荷数据作为模型的原始数据,直接利用LSTM进行预测的MAPE、MAE和RMSE分别为6.28%、149.04和31 782.80。其每个小波最小误差对应的MAPE、MAE和RMSE如表2所示。青岛新奥燃气的原始负荷、直接预测和组合预测负荷如图7所示。

从表2和图6中可以看到,Fk小波22阶第6层预测误差最小,其MAPE为1.13%,MAE为64.48,RMSE为5 301.71,相对于直接预测MAPE、MAE和RMSE的预测精度均有很大的提高,从另一面印证了Fk小波在22阶第6层的所预测的可靠性。

表 1 安阳华润燃气预测误差			
预测方法及阶层	MAPE/%	MAE	RMSE
SLSTM 直接预测	3.51	181.33	49 906.45
Sym2阶6层	1.60	123.58	21 215.74
Sym3阶6层	1.15	104.08	15 503.20
Sym4阶9层	1.12	104.13	15 602.73
Sym5阶6层	0.89	93.14	12 411.02
Sym6阶6层	0.97	96.24	13 531.50
Sym7阶8层	0.81	87.77	10 510.65
Sym8阶6层	0.90	92.75	12 554.18
Coif1阶9层	1.58	123.14	21 394.27
Coif2阶9层	1.24	108.49	17 131.71
Coif3阶5层	0.99	97.13	14 305.03
Coif4阶6层	1.00	91.67	12 602.55
Coif5阶5层	0.82	88.75	11 819.61
Fk4阶3层	1.76	130.85	24 691.33
Fk6阶6层	1.10	102.25	14 920.89
Fk8阶6层	1.02	98.16	13 402.37
Fk14阶7层	0.77	85.41	10 212.18
Fk18阶6层	0.77	85.46	10 128.55
Fk22阶6层	0.71	82.19	9 222.59

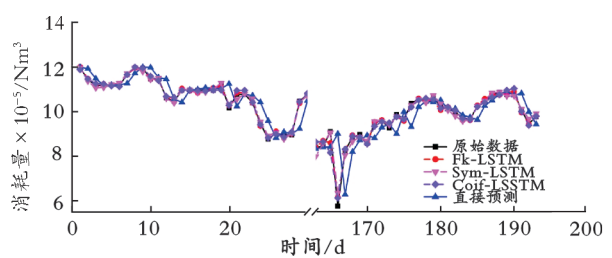


图5 安阳华润燃气原始负荷、直接预测负荷和不同小波组合预测负荷

表 2 青岛新奥燃气负荷预测误差			
小波类型及阶层	MAPE/%	MAE	RMSE
直接预测误差	6.28	149.04	31 782.80
Sym 8阶5层	1.14	71.25	6 615.22
Coif 5阶6层	1.35	69.92	6 210.37
Fk 22阶6层	1.13	64.48	5 301.71

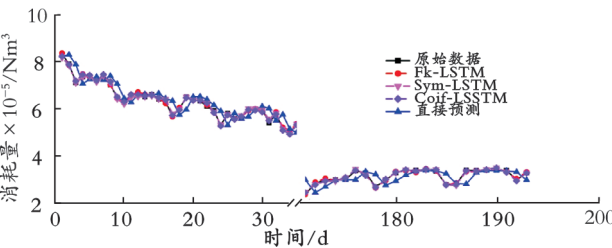


图6 青岛新奥燃气的原始、直接预测和不同小波组合预测负荷

4 结 论

所提出模型的最优小波为Fk小波,最优阶数为第22阶,最优层数为第6层,其预测精度为99.29%,达到了预期效果。此外,在保证预测精度的同时还指出了最优小波基的阶数和层数,可以在数据处理时直接对相应的阶数和层数进行分解,减少预测时间,为天然气负荷预测提供了一种直接处理数据的方法,同时对天然气在线工程的运用具有一定的参考价值。

参 考 文 献

[1] Ye Z H, Li Y K, Yan Q. Global natural gas consumption history and energy consumption pattern [M]. De Gruyter, Sydney:ESSE,2017:241-246.

[2] 黄献智,杜书成. 全球天然气和LNG供需贸易现状及展望[J]. 油气储运,2019,38(1):12-19.

[3] Jie G, Chen J, Hui S, et al. Current state and prospect of China natural gas industry[J]. China Oil & Gas, 2016, 23(1): 22-28.

[4] Balitskiy S, Bilan Y, Strielkowski W, et al. Energy efficiency and natural gas consumption in the context of economic development in the European Union[J]. Renewable and Sustainable Energy Reviews, 2016, 55:156-168.

[5] Fadiran G, Adebisuyi A T, Fadiran D. Natural gas consumption and economic growth: Evidence from selected natural gas vehicle markets in Europe[J]. Energy, 2019, 169:467-477.

[6] 王中元,罗东坤,刘璘璘. 我国城市燃气发展阶段及其主要特征[J]. 油气储运,2016,35(2):115-123.

[7] Kumar P, Dhal S, Kumar N. Neural network based forecasting model for natural gas consumption[J]. Communications on

- Applied Electronics, 2017, 7(11):1-8.
- [8] Baldacci L, Golfarelli M, Lombardi D, et al. Natural gas consumption forecasting for anomaly detection[J]. Expert Systems with Application, 2016, 62(15):190-201.
- [9] Furuoka F. Natural gas consumption and economic development in China and Japan: An empirical examination of the Asian context[J]. Renewable & Sustainable Energy Reviews, 2016, 56:100-115.
- [10] Akpinar M, Yumusak N. Day-ahead natural gas forecasting using nonseasonal exponential smoothing methods[C]// IEEE International Conference on Environment & Electrical Engineering, Milan:IEEE Press, 2017:1-4.
- [11] Hribar R, Papa G, Silc J. Prediction of natural gas consumption using empirical models [C]// 2017 International Conference on Smart Systems and Technologies (SST), Osijek: IEEE Press, 2017.
- [12] Özmen A, Yilmaz Y, Ozmen, Weber G W, et al. Natural gas consumption forecast with MARS and CMARS models for residential users[J]. Energy Economics, 2018, 70:357-381..
- [13] Sen D, Gunay M E, Tunc K. Forecasting annual natural gas consumption using socio-economic indicator for making future policies[J]. Energy, 2019, 173:1106-1118.
- [14] Beyca O F, Ervural B C, Tatoglu E, et al. Using machine learning tools for forecasting natural gas consumption in the province of Istanbul[J]. Energy Economics, 2019, 80:937-949.
- [15] 杨梅,李忠,吴昊. 基于多尺度时间特征的 LSTM 短期负荷预测[EB/OL].(2020-11-23)[2020-12-19]. <https://doi.org/10.14107/j.cnki.kzgc.20200542>.
- [16] 余凤,徐晓钟. 基于优化小波 BP 神经网络的燃气短期负荷预测[J]. 计算机仿真. 2015, 32(1), 372-376.
- [17] 乔伟彪,陈保东,卢泓方. 基于混沌理论和 Volterra 自适应滤波器的天然气负荷预测[J]. 中国科学(技术科学), 2015, 45(1), 91-102.
- [18] Brinks P. Potential-analysis of grey energy limits for residential buildings in Germany[J]. Energy and Buildings, 2016, 127: 571-579.
- [19] 乔伟彪,陈保东. 基于 Haar 小波变换和 ARIMA-RBF 的天然气时负荷预测[J]. 石油化工高等学校学报, 2015, 28(4): 75-80.
- [20] 刘春霞,李军,党伟超,等. 基于遗传算法优化小波神经网络的短期天然气负荷预测[J]. 计算机系统应用, 2020, 29(4): 175-180.
- [21] 涂建维,高经纬,李召,等. 基于长短时记忆网络的结构智能控制算法研究[J]. 华中科技大学学报(自然科学版), 2019, 47(12):110-115.
- [22] 周帆,赵荣普,杨文镏,等. 长、短期电力负荷大数据下的智能预测实例分析[J]. 电气工程与自动化, 2019, 36(6):27-30.
- [23] 罗航,王厚军,龙兵. 基于“松散型”小波神经网络的时间序列预测研究[J]. 计算机工程与应用, 2008, 44(8): 16-19.

(编辑 陈 雷)