

基于机器学习的原油管输能耗预测方法研究

徐磊^{1,2}, 侯磊^{1,2*}, 李雨^{1,2}, 张鑫儒^{1,2}, 白小众³, 雷婷^{1,2}, 朱振宇^{1,2}, 刘金海³, 谷文渊³, 孙欣³

1 中国石油大学(北京)机械与储运工程学院, 北京 102249

2 中国石油大学(北京)油气管道输送安全国家工程实验室/石油工程教育部重点实验室, 北京 102249

3 国家管网集团北方管道有限责任公司锦州输油气分公司, 锦州 121000

* 通信作者, houleicup@126.com

收稿日期: 2020-10-23

国家重点研发计划项目“油气长输管道及储运设施检验评价与安全保障技术”(2016YFC0802100)资助

摘要 准确的短期能耗预测是原油管道能耗管理的重要依据, 有助于能耗目标设定、调度优化和机组组合。原油管道能耗主要体现在泵机组上消耗的电能, 因此, 有必要对原油管道电耗展开准确预测。传统预测方法通常忽略数据噪声干扰, 对数据非线性特征的研究也不够深入, 上述因素使原油管道能耗预测变得复杂。因此, 提出一种将分解技术、分层抽样、改进粒子群算法和反向传播神经网络相结合的混合预测模型, 模型由数据预处理、优化、预测和评价4个部分组成。采用数据分解技术去除冗余噪声, 提取数据的主要特征; 采用分层抽样对数据集进行划分, 避免随机抽样引起的样本偏差; 将改进粒子群算法优化后的反向传播神经网络作为预测器。针对我国3条原油管道, 对提出的模型展开准确性评价, 平均绝对百分误差分别为4.02%、3.58%和3.88%。研究表明, 相比几种主流机器学习和SPS软件内的能耗预测模块, 提出的预测模型具有较高的预测精度和较强的泛化能力, 能被用于原油管道短期电耗预测。

关键词 能耗预测; 原油管道; 分解技术; 机器学习; 反向传播神经网络

Research into prediction of energy consumption of crude oil pipelines based on machine learning

XU Lei^{1,2}, HOU Lei^{1,2}, LI Yu^{1,2}, ZHANG XinRu^{1,2}, BAI Xiaozhong³, LEI Ting^{1,2}, ZHU Zhenyu^{1,2}, LIU Jinhai³, GU Wenyuan³, SUN Xin³

1 College of Mechanical and Transportation Engineering, China University of Petroleum-Beijing, Beijing 102249, China

2 National Engineering Laboratory for Pipeline Safety/MOE Key Laboratory of Petroleum Engineering, China University of Petroleum-Beijing, Beijing 102249, China

3 Jinzhou Oil and Gas Transmission Branch, National Pipe Network Group, Northern Pipeline Co., Ltd., Jinzhou 121000, China

Abstract Accurate short-term energy consumption prediction is crucial to energy management of crude oil pipelines. Based on the prediction results of energy consumption, some vital decisions such as energy consumption target setting, scheduling optimization and unit combination can be implemented effectively. The energy consumption of crude oil pipelines covers all aspects of the pipeline transportation system, among which the electricity consumption of the pump units is the most extensive. The electricity

引用格式: 徐磊, 侯磊, 李雨, 张鑫儒, 白小众, 雷婷, 朱振宇, 刘金海, 谷文渊, 孙欣. 基于机器学习的原油管输能耗预测方法研究. 石油科学通报, 2020, 04: 567-577

XU Lei, HOU Lei, LI Yu, ZHANG XinRu, BAI Xiaozhong, LEI Ting, ZHU Zhenyu, LIU Jinhai, GU Wenyuan, SUN Xin. Research into prediction of energy consumption of crude oil pipelines based on machine learning. Petroleum Science Bulletin, 2020, 04: 567-577. doi: 10.3969/j.issn.2096-1693.2020.04.050

consumption of the pump units by far accounts for the major part of the energy consumption of the crude oil pipelines. Therefore, it is necessary to accurately predict the energy consumption of the pump units, so as to have an overall assessment for the energy consumption of the pipeline system. At present, there is a wealth of methods that can be used to predict the energy consumption of crude oil pipelines. In traditional prediction methods, there are many limitations that make the prediction results deviate from the actual energy consumption. Generally speaking, the neglect of noise interference and the lack of in-depth research on the nonlinear characteristics of the data are the most common problems. The above factors complicate the energy consumption prediction of crude oil pipelines and make the prediction accuracy unsatisfactory. In order to solve the shortcomings of the traditional prediction methods, a novel hybrid prediction method is proposed for the short-term energy consumption prediction. The proposed hybrid method is based on the decomposition technique, stratified sampling, a modified particle swarm algorithm and a back-propagation neural network. The proposed model consists of four parts: the data preprocessing module, the optimization module, the prediction module and the evaluation module. The decomposition technique is adopted to eliminate the redundant noise and extract the major features of the original data. The stratified sampling method is used to divide the data set to avoid the sampling bias of random sampling. The back-propagation neural network optimized by the modified particle swarm optimization algorithm is regarded as a predictor. Based on three crude oil pipelines located in China, the proposed prediction model is evaluated by comparing the predicted results with the actual data. The mean absolute percentage errors of the evaluation indicators are 4.02%, 3.58% and 3.88% respectively. Compared with several popular machine learning methods and the prediction modules in SPS software, the proposed prediction method has excellent prediction accuracy and generation ability, which can be used for short-term energy consumption prediction of crude oil pipelines.

Keywords energy consumption prediction; crude oil pipeline; decomposition technique; machine learning method; back-propagation neural networks

doi: 10.3969/j.issn.2096-1693.2020.04.050

0 引言

原油是重要的战略储备物资,主要通过铁路、水路、公路和管道4种方式运输。其中,管道运输因其运量大、封闭安全、易于管理等优点而被广泛应用。原油管道是将油田生产的原油输送至炼厂、港口或铁路转运站,具有管径大、输量大、运输距离长、分输点少的特点^[1]。我国原油管道2020年的规划里程为32 000 km,2025年在此基础上要新增15.63%的里程。目前,我国所产原油80%以上为凝点较高的含蜡原油和黏稠的重质原油,输送过程耗能很大^[2]。其中,泵机组的电耗是总能耗的主要部分,通过对泵机组电耗预测,能够做出一些关键决策,如能耗目标设定、批量调度和机组组合。依据能耗预测时间间隔长短,可分为短期预测、中期预测和长期预测,三者分别是指一天到一周、一周到一月、一月到一年的时间间隔^[3]。本研究的能耗数据是以天为间隔进行采集的,因此,可视作短期能耗预测。

在实际管输过程中,能耗相关数据存在信息冗余、噪声干扰及非线性等特征,加大了预测难度。近年来,随着软计算技术的快速发展,机器学习方法已在能耗预测领域中得到广泛应用,该方法具有较强的容错性,较好的预测性能^[4-6]。图1统计了2000年1月至2020年6月机器学习方法在能耗预测领域所发表论文的数量趋势,数据来源于Web of Science数据库。通

过统计分析可得,机器学习方法在能耗预测领域的研究呈现逐年递增的趋势。随着“智慧管道”概念的提出,姜昌亮^[7]、黄维和^[8-9]、宫敬^[10-11]、张劲军^[12]、吴长春^[13]、董绍华^[14-15]等众多行业学者围绕智慧管道开展了大量研究工作,以数据驱动为依托的机器学习方法在油气管道行业的应用与发展势在必行^[16-22]。

在油气管输行业,若干学者对基于神经网络的油气管道能耗预测展开了相关研究。ZENG^[23]利用神经

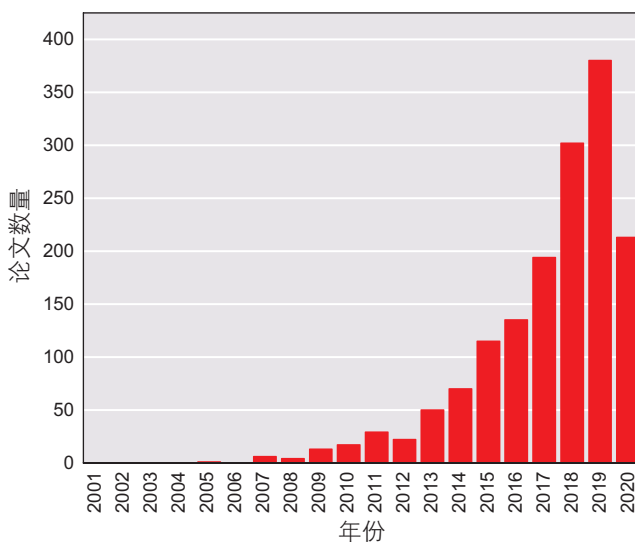


图1 机器学习方法在能耗预测领域发表论文数量

Fig. 1 Number of papers occupied by machine learning method in the field of energy consumption prediction

网络对成品油管道电耗进行预测，并与线性回归和支持向量机展开对比，研究表明，所提出的神经网络模型有助于批次调度和能耗目标的设定。吴倩^[24]提出一种基于BP神经网络的原油管道能耗预测方法，研究表明，与线性回归和灰色模型相比，建立的模型预测精度较高，适用于预测多种原油管道的耗电量和耗油量。温馨^[25]将神经网络、支持向量机和极限学习机分别用于成品油管道能耗预测，研究表明，神经网络学习能力强、收敛能力强，拥有较高的预测精度。侯磊^[26]利用BP神经网络建立输油管道能耗预测模型。预测结果表明，该模型预测偏差不超过4%，为输油管道能耗预测提供了一种新思路。高山卜^[27]建立基于改进的BP神经网络原油管道能耗预测模型，得到该模型的误差在3%以内，且模拟值能够反映真实值的变化趋势。林冉^[28]以某条输油管道几年来输量及生产油耗、电耗数据为基础，用人工神经网络的方法建立了管道输量与生产油耗、电耗的预测模型。分析表明，该模型的计算结果相对偏差在±5%以内，满足工程实际需要，能够用该模型来预测热油管道的生产油耗和电耗。

但以上方法对数据非线性和数据噪声考虑均不够充分，预测精度仍有一定的提升空间。图2~4统计了各类机器学习方法在能耗预测领域的应用。通过调研可得，智能优化算法和数据分解技术已被运用到风能、太阳能、核能等领域，用来解决数据非线性难以准确拟合、数据噪声难以高效去除的问题，但在油气行业运用甚少^[29-32]。因此，本研究提出一种数据分解算法(CEEMDAN)、改进粒子群优化算法(IPSO)和反向传播神经网络(BPNN)

相结合的混合预测模型。采用自适应噪声完整集成经验模态分解(CEEMDAN)去除冗余噪声，提取原始数据的主要特征，采用IPSO优化反向传播神经网络。结合国内3条原油管道，对提出的模型展开准确性评价。

1 基本方法

1.1 自适应噪声完整集成经验模态分解

实际现场采集的数据通常含有噪声，常规方法无

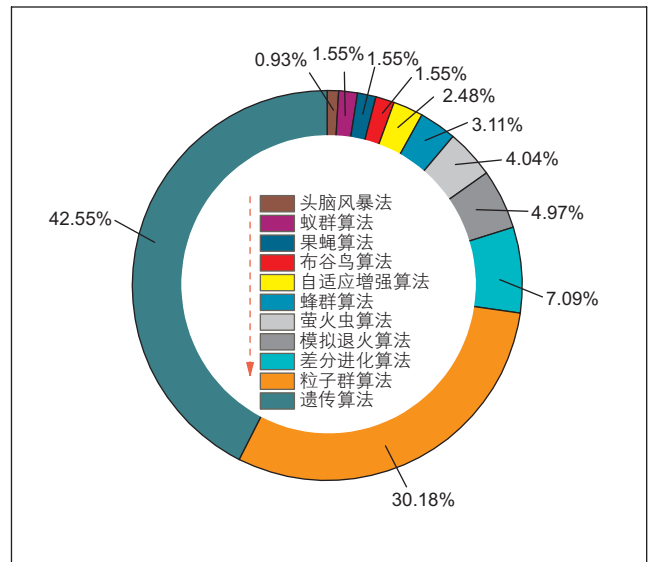


图3 能耗预测领域的优化算法统计

Fig. 3 Statistics of optimization algorithms in the field of energy consumption prediction

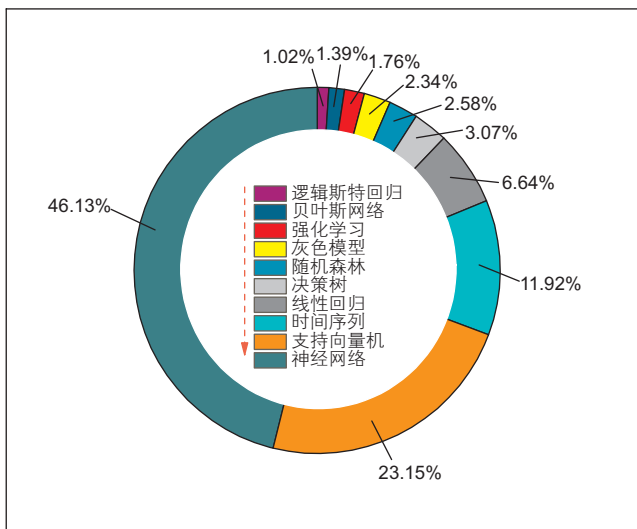


图2 能耗预测领域的基准模型统计

Fig. 2 Statistics of benchmark models in the field of energy consumption prediction

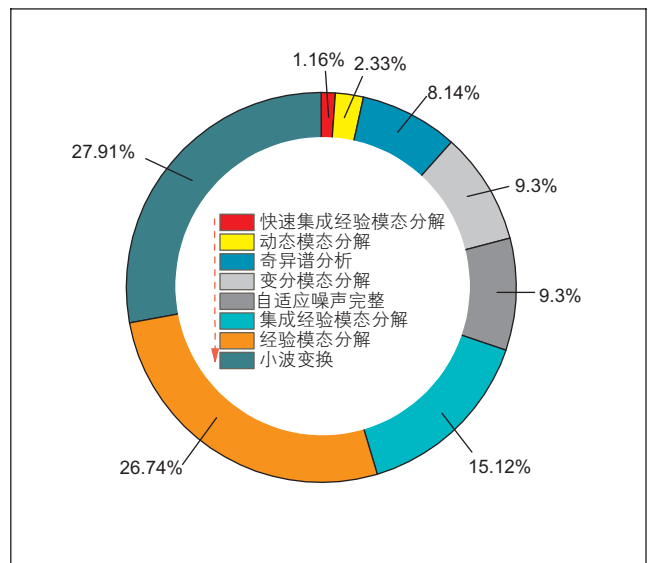


图4 能耗预测领域的分解算法统计

Fig. 4 Statistics of decomposition algorithms in the field of energy consumption prediction

法提取数据的主要特征。前人学者对此做了诸多研究, Huang^[33]提出经验模态分解(EMD)方法将原始信号分解为多个内涵模态分量(IMF), 但模型和EMD无法很好的融合。针对此问题, Wu和Huang^[34]提出集成经验模态分解(EEMD)方法, EEMD在稳定性方面有明显的改善, 但很难完全抵消所增加的噪声。为了提高EEMD的性能, Torres^[35]提出一种高级的自适应噪声完整集成经验模式分解算法(CEEMDAN)。CEEMDAN算法不仅能够解决EMD中模态混合的问题, 而且通过加入成对白噪声, 可以高效消除噪声, 实现较好预测效果。因此, CEEMDAN算法作为一种实用的数据预处理工具, 能够提高预测性能。

1.2 反向传播神经网络

人工神经网络(ANN)目前被广泛应用于各行业。其中, 神经元是人工神经网络的基本处理单元, 在输入节点处接收非线性信息, 经内部处理后在输出节点处生成响应^[36]。反向传播神经网络(BPNN)典型的三层前馈神经网络结构如图5所示。在反向传播神经网络中, 神经元分层排列, 包括输入层、隐藏层和输出层。人工神经网络将多个神经元相互连接, 每个连接都有相应的权重。假设 W_{ij} 是神经元 i 和 j 之间的连接权重, X_i 是神经网络的输入向量, 则两个连续层($k-1, k$)的神经元 j 的输出值由公(1)式确定

$$Y_j = \Theta \left(b_j + \sum_{i=1}^m W_{i,j} X_i \right) \quad i \in [1, m] \text{ and } j \in [1, n] \quad (1)$$

其中 b_j 是阈值, Θ 是激活函数, m 是 $k-1$ 层中神经元

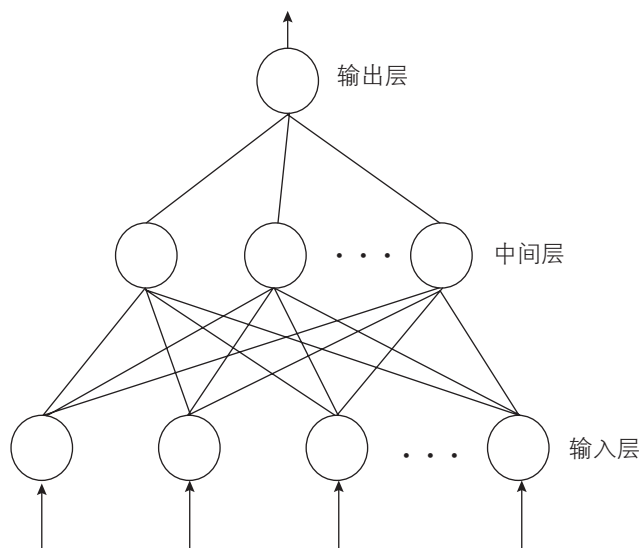


图5 三层前馈神经网络结构图

Fig. 5 Structure diagram of 3-layers feed-forward neural network

的数量, n 是 k 层中神经元的数量。

1.3 改进粒子群算法

粒子群优化算法最早是由Eberhart和Kennedy^[37]于1995年提出, 能够解决复杂约束优化问题, 同时又有较快的收敛速度。然而, 传统的粒子群优化算法有时会陷入局部最优, 因此提出改进粒子群优化算法(IPSO)来解决该问题。改进粒子群优化算法基于分布函数确定粒子位置, 与传统的粒子群优化算法相比, 改进粒子群优化算法需要调整的参数少, 工作效率高、不容易陷入局部最优^[38]。根据改进粒子群优化算法, 计算种群的最佳位置如下:

$$M_{\text{best}} = \frac{1}{M} \sum_{i=1}^M P_i \quad (2)$$

$$P_i(t+1) = \varepsilon P_i(t) + (1-\varepsilon) P_g(t) \quad (3)$$

$$x_i(t+1) = P_i(t+1) + \alpha |M_{\text{best}} - x_i(t)| \ln\left(\frac{1}{u}\right) \quad (4)$$

其中 M_{best} 表示粒子的平均历史最佳位置, M 是粒子群的大小, P_i 是第 i 个粒子的历史最佳位置, $P_i(t)$ 是在第 i 个粒子在时间 t 的最佳位置, $P_g(t)$ 是在时间 t 的全局最佳位置, $x_i(t)$ 是第 i 个粒子在时间 t 的位置, α 是创新参数, 值不大于1, ε 和 u 为在 $(0, 1)$ 之间的均匀分布。结合式(3)和等式(4), 可得到后代粒子的位置。

1.4 分层抽样方法

本研究生产数据来源于东北3条原油管道, 记为A、B、C, 以天为周期, 采集2019年01月01日至2019年6月29日的数据, 共计180组数据, 管道A的部分数据如表1所示。在这项研究中, 对于输入参数的选取需要满足两个条件, 参数是变量且与电耗间存在相关性, 因此, 选取管输过程中的排量、平均进压、平均汇压、平均出压、平均泵出、平均进温、平均汇温、平均出温、地温、下站进压和下站进温等11个参数作为预测模型的输入参数。电耗是预测值, 所以选取电耗作为输出参数。

训练集和测试集按4:1的比例进行划分。由于传统的随机抽样方法可能会导致测试集的分布规律与原始数据集分布规律有较大偏差, 使预测结果不客观。因此, 考虑采用分层抽样方法来划分数据集^[39-40]。在原油运输过程中, 输量对电耗影响最大, 根据输量的分布规律可知, 某些特定范围内的数据量较少, 采用抽样时可能会遗漏该范围内的数据, 为了保证每一数据范围内的数据都能够按比例被抽取, 因此, 依据数据的分布规律, 将采集的数据划分为4个区间。以管

表 1 管道 A 的部分数据

Table 1 Partial data of Pipeline A

时间	排量/t	进压/ MPa	汇压/ MPa	出压/ MPa	泵出/ MPa	进温/ °C	汇温 /°C	出温 /°C	地温 /°C	下站进压/ MPa	下站进温 /°C	机组电耗/ (10 ³ kW/d)
20190106	12 059	1.71	1.55	4.39	5.75	36.8	42.2	44.1	12.1	0.18	40.5	13 439
20190123	12 604	1.84	1.66	4.52	5.84	35.3	40.2	42.2	10.5	0.22	39.1	13 580
20190209	13 645	1.32	1.16	4.91	4.99	34.7	40.1	41.4	9.9	0.27	38.4	13 335
20190307	13 473	1.72	1.51	4.56	4.64	35.1	41.8	42.8	9.8	0.22	39.3	10 676
20190405	13 103	1.69	1.49	5.37	5.46	35.6	41.9	43.3	12.7	0.25	40.5	14 521
20190510	12 122	1.49	1.31	4.63	4.71	35.0	42.8	44.0	17.4	0.28	40.9	11 535
20190522	16 405	2.20	1.99	5.44	5.57	35.6	41.9	43.2	19.3	0.37	41.5	15 586
20190607	16 931	1.55	1.35	4.95	5.07	35.8	41.9	43.3	21.1	0.31	41.2	15 308

道 A 为例，两种抽样方法得到的测试集与初始样本的偏差如表 2 所示。3 条管线的随机抽样和分层抽样的平均绝对百分比误差(MAPE)见图 6。对于 A、B、C 管道，随机抽样的 MAPE 分别为 26.44%、19.20%、17.46%，分层抽样的 MAPE 分别为 2.16%、2.16%、

1.05%。结果表明，分层抽样得到的测试集与初始样本有较好的一致性。因此，本文采用分层抽样来划分训练集和测试集。

2 预测模型框架

根据分解技术(CEEMDAN)和改进粒子群算法的 BP 神经网络(IPSO-BPNN)建立混合预测模型。采用 CEEMDAN 去除冗余噪声，提取原始数据的主要特征；采用 IPSO 拟合非线性特征，优化 BPNN 网络结构。结合国内 3 条原油管道，对比机器学习方法和 SPS 能耗预测模块，对所提出混合模型(CEEMDAN-IPSO-BPNN)的预测效果进行准确性评价，混合预测系统的建立流程如图 7 所示。

表 2 基于管道 A 随机抽样和分层抽样的平均绝对百分比误差
Table 2 Mean absolute percentage error of random sampling and stratified sampling based on pipeline A

分类	随机抽样相对误差/%	分层抽样相对误差/%
3.0	12.11	1.91
4.0	8.71	2.16
5.0	54.50	2.39
6.0	30.45	2.16

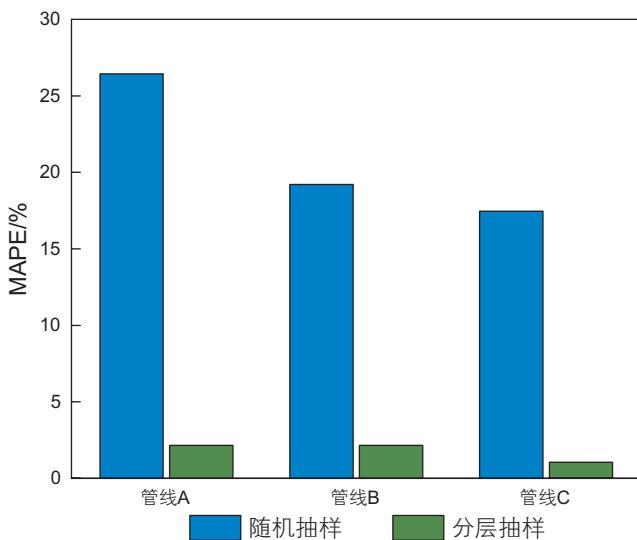


图 6 3 条管线随机抽样和分层抽样平均绝对百分比误差
Fig. 6 Mean absolute percentage error of random sampling and stratified sampling based on the 3 pipelines

3 预测结果分析

实验过程通过 Python 3.6.6 语言实现，处理器为 Intel Xeon E5-2643 v4，计算机内存为 96.0 GB。以管道 A 为例，CEEMDAN-IPSO-BPNN 混合模型的主要实验参数如表 3 所示。将 CEEMDAN-IPSO-BPNN 混合模型与主流机器学习方法和 SPS 能耗模块的预测性能分别展开对比，全面评价 CEEMDAN-IPSO-BPNN 的预测性能。

3.1 精度评估指标

通过大量文献调研，选取相对误差(RE)、决定系数(R²)、平均绝对百分比误差(MAPE)、均方根误差(RMSE)、平均绝对误差(MAE)、泰尔不等式系数(TIC)、分数偏差(FB)、预测结果统计 1(U1)、预测结果统计 2(U2)和方向精度(DA)10 个主流指标来评价模

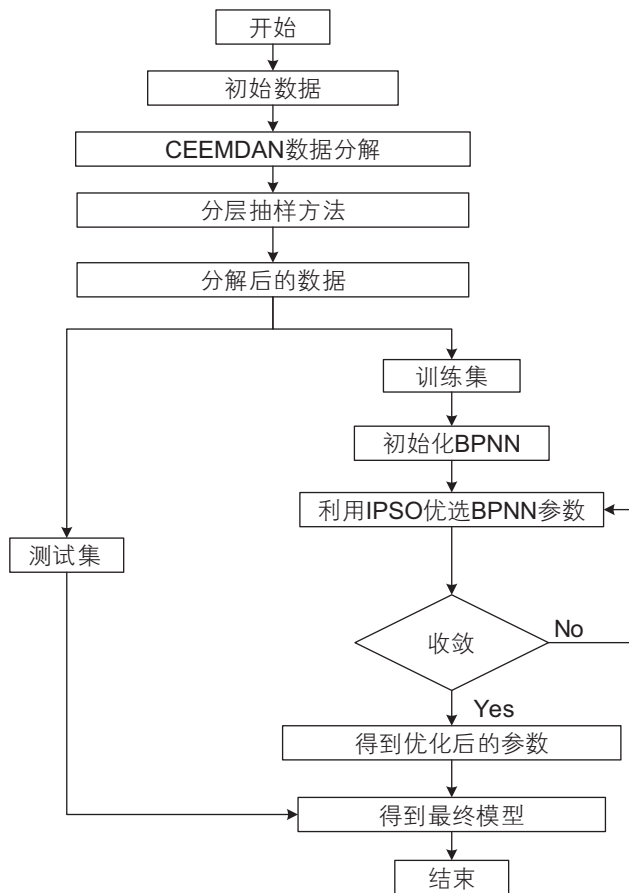


图7 CEEMDAN-IPSO-BPNN混合预测模型的流程图
Fig. 7 Flow chart of CEEMDAN-IPSO-BPNN

型的预测性能^[41-44], 表达式如表4所示。

3.2 与常规机器学习方法对比

将建立的混合预测模型CEEMDAN-IPSO-BPNN

表4 模型预测性能评价的十个主流指标

Table 4 Ten mainstream indicators for performance evaluation

指标	表达式	指标	表达式
RE	$RE(\%) = \frac{1}{N} \sum_{i=1}^N \left(\frac{y'_i - y_i}{y_i} \right) \times 100\%$	TIC	$TIC = \sqrt{\frac{1}{N} \times \sum_{i=1}^N (y'_i - y_i)^2} / \left(\sqrt{\frac{1}{N} \times \sum_{i=1}^N y_i^2} + \sqrt{\frac{1}{N} \times \sum_{i=1}^N y_i'^2} \right)$
R ²	$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - y'_i)^2}{\sum_{i=1}^N (y_i - \bar{y}_i)^2}$	FB	$FB = 2(\bar{y} - \bar{y}') / (\bar{y} + \bar{y}')$
MAPE	$MAPE(\%) = \frac{1}{N} \sum_{i=1}^N \left \frac{y'_i - y_i}{y_i} \right \times 100\%$	U1	$U1 = \sqrt{\frac{1}{N} \times \sum_{i=1}^N (y_i - y'_i)^2} / \left(\sqrt{\frac{1}{N} \times \sum_{i=1}^N y_i^2} + \sqrt{\frac{1}{N} \times \sum_{i=1}^N y_i'^2} \right)$
RMSE	$RMSE = \sqrt{\frac{1}{N} \times \sum_{i=1}^N (y'_i - y_i)^2}$	U2	$U2 = \sqrt{\frac{1}{N} \times \sum_{i=1}^N ((y_{i+1} - y'_{i+1}) / y_i)^2} / \sqrt{\frac{1}{N} \times \sum_{i=1}^N ((y_{i+1} - y'_i) / y_i)^2}$
MAE	$MAE = \frac{1}{N} \sum_{i=1}^N y'_i - y_i $	DA	$DA = \frac{1}{1} \times \sum_{i=1}^N w_i, w_i \begin{cases} 1, & \text{if } (y_{i+1} - y_i)(y'_{i+1} - y_i) > 0 \\ 0, & \text{otherwise} \end{cases}$

表3 实验参数值

Table 3 Experimental parameters

算法	参数	数值
CEEMDAN	最大迭代次数	100
	噪声标准差	1
IPSO	种群规模	10
	最大迭代次数	10
	加速因子	0.6
	精度	10 ⁻⁸
BPNN	输入层神经元数	105
	隐含层神经元数	19
	输出层神经元数	1
	学习速率	0.06583
	正则化系数	0.08144

与GA-BPNN、PSO-BPNN、IPSO-BPNN、EMD-IPSO-BPNN和EEMD-IPSO-BPNN等5个模型展开预测性能对比, 评价混合模型的预测性能。对3条管道的10个指标分别取平均值得到表5和图8, 为了全面评价所提出混合模型的预测性能, 开展了对比研究。

(1)通过GA-BPNN与PSO-BPNN的预测结果对比发现, PSO-BPNN的预测性能优于GA-BPNN, 证明PSO对于预测性能的提高优于GA, 相比主流优化算法, PSO对于预测性能的提高幅度较大。

(2)通过PSO-BPNN与IPSO-BPNN的预测结果对比发现, IPSO-BPNN的预测性能优于PSO-BPNN, 证明改进粒子群算法在一定程度上提高了粒子群算法的性能, 有效解决了优化算法容易出现局部最优解的问题。

表 5 3 条管道上十个评价指标的平均值

Table 5 Average values of ten evaluation indicators for the 3 pipelines

模型	<i>RE</i>	<i>R</i> ²	<i>MAPE</i>	<i>RMSE</i>	<i>MAE</i>	<i>TIC</i>	<i>FB</i>	<i>U1</i>	<i>U2</i>	<i>DA</i>
GA-BPNN	-1.673	0.765	11.23	3341.7	2454.9	0.063	-0.0048	0.063	0.341	0.806
PSO-BPNN	-0.973	0.801	9.783	3073.4	2138.1	0.058	-0.0031	0.058	0.328	0.814
IPSO-BPNN	0.263	0.845	7.447	2688.2	1691.8	0.050	0.0013	0.050	0.271	0.843
EMD-IPSO-BPNN	-0.161	0.949	5.609	1531.3	1139.8	0.032	-0.0008	0.032	0.149	0.814
EEMD-IPSO-BPNN	-0.073	0.956	5.315	1434.1	1064.5	0.029	-0.0013	0.029	0.143	0.806
CEEMDAN-IPSO-BPNN	0.305	0.973	3.828	1112.3	788.30	0.023	0.0013	0.023	0.113	0.843

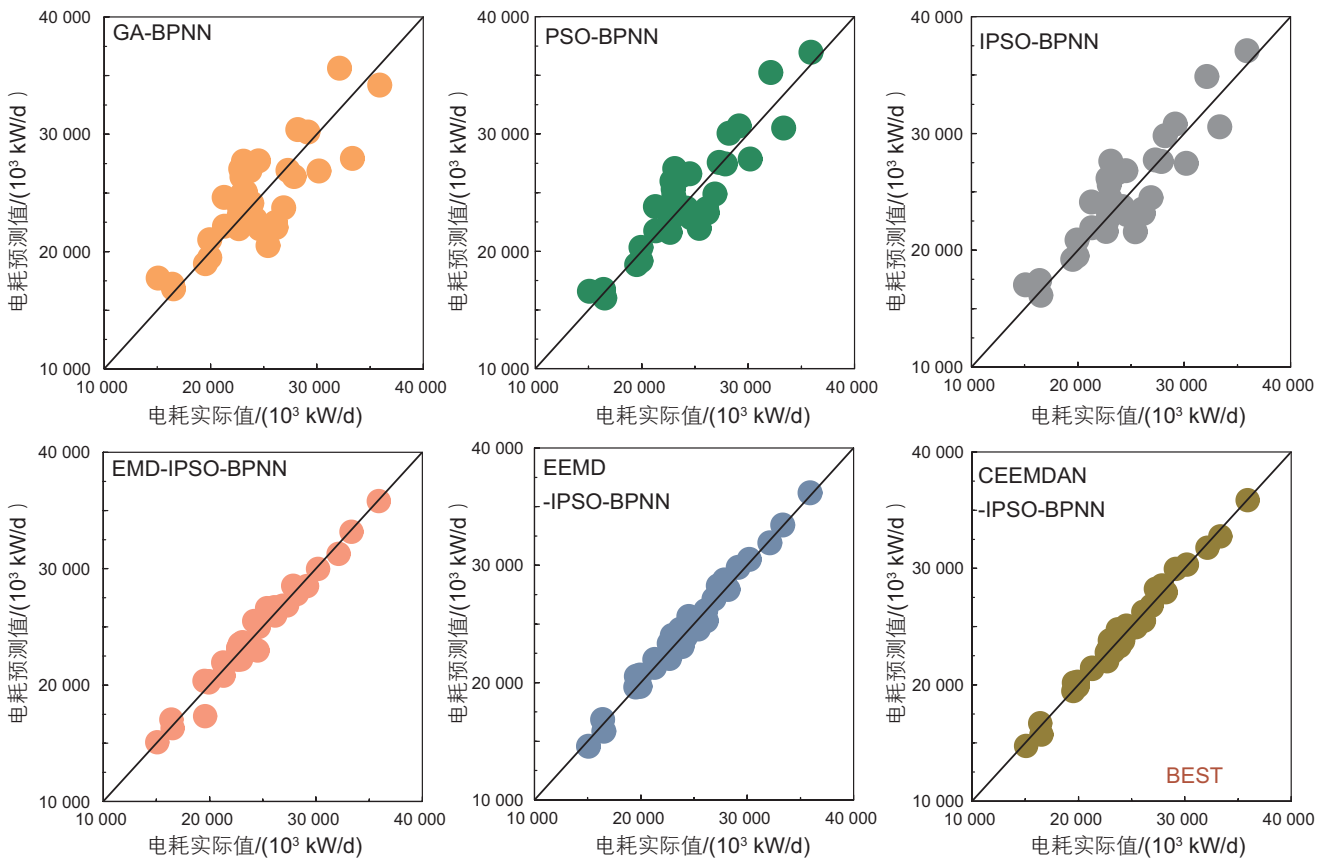


图 8 建立的模型和对比模型的预测值

Fig. 8 Predicted values of the proposed model and the comparison models

(3) 通过IPSO-BPNN与EMD-IPSO-BPNN、EEMD-IPSO-BPNN和CEEMDAN-IPSO-BPNN预测结果对比发现,IPSO-BPNN预测效果相对最差,证明分解算法能够在一定程度上提高模型的预测性能。

(4)通过CEEMDAN-IPSO-BPNN与EMD-IPSO-BPNN和EEMD-IPSO-BPNN预测结果展开对比发现,相比EMD和EEMD数据分解技术,CEEMDAN具有更好的分解效果,能够深入地捕获数据特征,提高模型预测结果。

因此,通过与5个机器学习模型的对比可知,建立的混合模型CEEMDAN-IPSO-BPNN具有最好的预

测性能。

3.3 与SPS能耗预测软件对比

SPS是款成熟的瞬态水力模拟软件,能够模拟单一流体、批次流体和单相混合流体在管道中的输送过程。利用软件内置的管道、泵、压缩机、阀和控制器等,SPS能建立对应的数学模型,通过计算流量、压力、密度、温度和其它一些沿线随时间变化的参数来开展仿真研究。在油气储运领域,SPS软件使用广泛,主要功能包括停输再启动分析、管输能耗分析、泵和压缩机操作计划分析、水击泄压系统设计、批次计划、

不同操作方案和不同设计的经济性分析等。

为保证预测结果具有可对比性,本研究以上文提到的管道A、B和C为例,运用SPS软件分别建立管输能耗预测模型。将预测结果与CEEM-DAN-IPSO-BPNN的预测结果进行对比,深入评价CEEMDAN-IPSO-BPNN模型的预测性能。

首先建立站间管道系统模型,通过计算泵功率得到每天耗电量,模型采用设备Pump来模拟中间站的输油泵,采用General Pipe来模拟站间管道,管道输送介质采用出站掺混后的原油,泵的特性曲线、管道内径、管线长度以及掺混原油特性等数据通过现场获取。模型以进口控制压力、出口控制流量作为边界条件,计算泵机组每天耗电量。以管道A为例,图9为建立的SPS能耗预测简易流程,预测得到的部分结果如表6所示。对3条管道模拟得到的10个指标取平均值得到统计表7,图11为CEEMDAN-IPSO-BPNN与SPS预测结果相对误差绝对值的箱型图。

从结果上来看,SPS模拟得到的MAPE为10.72%,对造成现有误差的原因进行了思考,主要包括:(1)管道服役时间长,管道内壁产生磨损,粗糙度难以估计,与设计规范的推荐值有差别;(2)管道输送介质含蜡量较高,管道在长时间运行过程中会积蜡,并且管道的清管效果未知,影响了管道的实际水力半径;(3)站内管线复杂,模拟中只能粗略考虑站内的摩阻损失,直接影响了模拟数据的精确度;(4)总传热系数难以准确

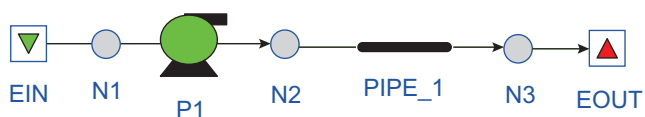


图9 基于管道A建立的SPS能耗预测流程

Fig. 9 The SPS energy consumption prediction process based on pipeline A

表6 基于管道A的部分预测值

Table 6 Partial predicted value based on pipeline A

日期	机组耗电/(kWh/d)	模拟耗电/(kWh/d)
20190102	13 479	12 173
20190126	14 011	13 066
20190303	9225	9942
20190326	11 059	11 805
20190418	12 856	12 780
20190519	12 068	12 988
20190604	15 417	13 625
20190619	30 189	27 425

确定;(5)采集传感器的精度不准;(6)泵的实际运行特性跟厂家提供的特性曲线有偏差;(7)现场数据为一天的平均值或累计值,而不同时刻的数据存在波动性,仿真数据是根据一天的累计输量和平均压力稳态模拟所得到的,这与真实情况存在差别。

通过表7和图10分析可得,针对相同数据集,CEEMDAN-IPSO-BPNN相较于主流商业软件SPS拥有更好的预测性能。表明机器学习预测效果对于数据的依赖性要弱于SPS软件,运用SPS展开预测,若想取得理想的预测效果,需要详细的参数信息,在实际生产运行中部分参数较难获取。而机器学习能通过数据分解技术和非线性映射功提高模型的预测性能,也间接表明了基于数据驱动的机器学习方法在油气储运行业的优势和潜力。

表7 建立的模型预测值和SPS能耗模块预测值的对比

Table 7 Comparison of the predicted values between the proposed model and the SPS energy consumption module

日期	CEEMDAN-IPSO-BPNN	SPS软件
RE	0.305	-2.915
R ²	0.973	0.827
MAPE	3.828	10.715
RMSE	1112.3	2752.7
MAE	788.30	2472.6
TIC	0.023	0.056
FB	0.0013	-0.022
UI	0.023	0.056
U2	0.113	0.640
DA	0.843	0.604

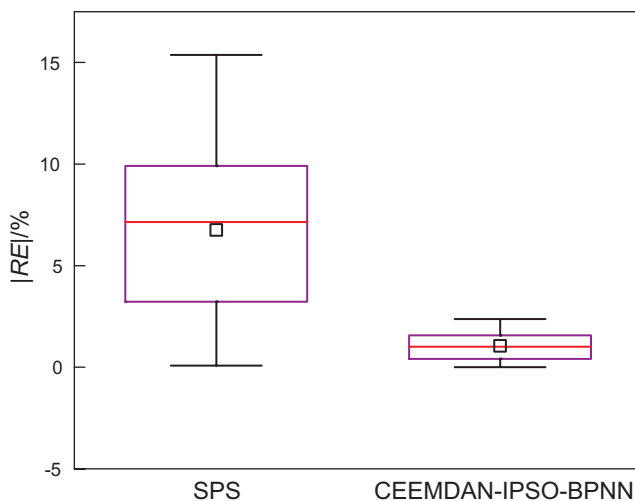


图10 混合模型与SPS预测相对误差绝对值箱型图

Fig. 10 Box plot for the absolute value of the relative error between the hybrid model and SPS

4 结论

采用分层抽样方法对初始数据进行划分,避免出现简单随机抽样遗漏某些特性的问题,保证预测结果的客观性和可靠性。提出了基于反向传播神经网络的改进粒子群优化算法,提高了粒子群的搜索能力、避免了局部最优解。将自适应噪声完整集成经验模态分解技术应用到油气管道领域,与其他分解技术相比,该分解技术能够高效消除预测过程中的冗余噪声,捕

捉原始数据集的主要特征。在此基础上,建立的混合预测模型综合了数据分解技术和优化算法的优势,对比5类主流机器学习方法和管输能耗预测软件SPS,建立的能耗预测模型平均绝对百分误差降低7.402%、5.955%、3.619%、1.781%、1.487%和6.887%,拥有较高的预测精度和较强的泛化能力,能有效指导能耗目标设定、调度优化和机组组合。提出的预测方法以机器学习为基础,解决了传统物理数学建模复杂等缺点,关注输入和输出参数的映射关系,在实际管道甚至是更加复杂的管道系统中也有一定的适用性。

参考文献

- [1] 杨筱蓓. 输油管道设计与管理[M]. 石油大学出版社, 1996. [YANG X H. Design and management of oil pipelines [M]. China Petroleum University Press, 1996.]
- [2] 孟迪. J油田含蜡原油常温输送技术及剪切效应模拟研究[D]. 西南石油大学, 2014. [MENG D. Simulation study on normal temperature transportation technology and shear effect of waxy crude oil based on J Oilfield [D]. Southwest Petroleum of University, 2014.]
- [3] AMASYALI K, EL-GOHARY N M. A review of data-driven building energy consumption prediction studies. *Renewable and Sustainable Energy Reviews* 2018, 81: 1192–1205.
- [4] HERRERA G P, CONSTANTINO M, TABAK B M, et al. Long-term forecast of energy commodities price using machine learning. *Energy*. 2019, 179: 214–221.
- [5] QU Z X, MAO W Q, ZHANG K Q. Multi-step wind speed forecasting based on a hybrid decomposition technique and an improved back-propagation neural network. *Renewable Energy* 2018, 133: 919–929.
- [6] HU Y S, LI J G, HONG M N, et al. Short term electric load forecasting model and its verification for process industrial enterprises based on hybrid GA-PSO-BPNN algorithm-A case study of papermaking process. *Energy* 2019, 170: 1215–1227.
- [7] 姜昌亮. 中俄东线天然气管道工程管理与技术创新[J]. 油气储运, 2020, 39(2): 121–129. [JIANG C L. Management and technological innovation in China–Russia Eastern Gas Pipeline Project[J]. *Oil & Gas Storage and Transportation*, 2020, 39(2): 121–129.]
- [8] 黄维和, 沈鑫, 郝迎鹏. 中国油气管网与能源互联网发展前景[J]. 北京理工大学学报(社会科学版), 2019, 21(1): 1–6. [HUANG W H, SHEN X, HAO Y P. Prospects of China’s oil and gas pipeline network and energy internet development [J]. *Journal of Beijing Institute of Technology (Social Sciences)*, 2019, 21(1): 1–6.]
- [9] 黄维和, 郑洪龙, 李明菲. 中国油气储运行业发展历程及展望[J]. 油气储运, 2019, 38(1): 1–11. [HUANG W H, ZHENG H L, LI M F. Development history and prospect of oil & gas storage and transportation industry in China [J]. *Oil & Gas Storage and Transportation*, 2019, 38(1): 1–11.]
- [10] 宫敬, 徐波, 张微波. 中俄东线智能化工艺运行基础与实现的思考[J]. 油气储运, 2020, 39(2): 130–139. GONG J, XU B, ZHANG W B. Thinking on the basis and realization of intelligent process operation of China–Russia Eastern Gas Pipeline [J]. *Oil & Gas Storage and Transportation*, 2020, 39(2): 130–139.]
- [11] 宫敬. 从旁接油罐到管网联运再到智能调控——中国输油管道工艺技术50年发展回顾与展望[J]. 油气储运, 2020, 39(8): 841–850. [GONG J. Review and outlook for development of oil pipeline technology in the past 50 years in China, from floating tank process to joint operation of pipeline network and further to intelligent control [J]. *Oil & Gas Storage and Transportation*, 2020, 39(8): 841–850.]
- [12] 苏怀, 张劲军. 天然气管网大数据分析方法及发展建议[J/OL]. 油气储运: 1–21[2020–08–30]. [SU H, ZHANG J J. Big data analysis method and development suggestion of natural gas pipeline network [J/OL]. *Oil & Gas Storage and Transportation*: 1–21 [2020–08–30].]
- [13] 吴长春, 左丽丽. 关于中国智慧管道发展的认识与思考[J]. 油气储运, 2020, 39(4): 361–370. [WU C C, ZUO L L. Understanding and thinking on the development of China’s intelligent pipeline [J]. *Oil & Gas Storage and Transportation*, 2020, 39(4): 361–370.]
- [14] 董绍华, 张河苇. 基于大数据的全生命周期智能管网解决方案[J]. 油气储运, 2017, 36(1): 28–36. [DONG S H, ZHANG H W. Solution of full-life-cycle intelligent pipeline network based on big data[J]. *Oil & Gas Storage and Transportation*, 2017, 36(1): 28–36.]
- [15] 董绍华. 中国油气管道完整性管理20年回顾与发展建议[J]. 油气储运, 2020, 39(3): 241–261. [DONG S H. Review of China’s oil and gas pipeline integrity management in the past 20 years and development suggestions [J]. *Oil & Gas Storage and Transportation*,

- 2020, 39(3): 241–261.]
- [16] 何蕾, 温凯, 吴长春, 等. 基于多种智能算法的腐蚀天然气管道可靠性评价方法[J]. 石油科学通报, 2019, 4(3): 310–322. [HE L, WEN K, WU C C, et al. A corroded natural gas pipeline reliability evaluation method based on multiple intelligent algorithms [J]. Petroleum Science Bulletin, 2019, 4(3): 310–322.]
- [17] 王雨墨, 李彦博, 李晓平, 等. 神经网络预测管道冲蚀速率研究进展[J]. 石油科学通报, 2020, 5(1): 114–121. [WANG Y M, LI Y B, LI X P, et al. Recent progress on ANN-based pipeline erosion predictions [J]. Petroleum Science Bulletin, 2020, 5(1): 114–121.]
- [18] 张河苇, 金剑, 董绍华, 等. 一种基于管道大数据的腐蚀因素相关性分析方法[J]. 科学通报, 2018, 63(8): 777–783. [ZHANG H W, JIN J, DONG S H, et al. A corrosion correlation analysis method based on pipeline big data [J]. Chinese Science Bulletin, 2018, 63(8): 777–783.]
- [19] 伍星光, 侯磊, 吴守志, 等. 基于改进SHIPP模型的大型原油库区事故风险预测[J]. 油气储运, 2020, 39(5): 519–529. [WU X G, HOU L, WU S Z, et al. Accident risk prediction of large crude oil tank area based on improved SHIPP model [J]. Oil & Gas Storage and Transportation, 2020, 39(5): 519–529.]
- [20] WEI N, LI C J, PENG X L, et al. Daily natural gas consumption forecasting via the application of a novel hybrid model. Applied Energy 2019; 250: 358–368.
- [21] LIN Z, LIU X L, LAO L Y, et al. Prediction of two-phase flow patterns in upward inclined pipes via deep learning[J]. Energy 2020; 210: 118541.
- [22] SU H, ZIO E, ZHANG J J, et al. A systematic data-driven Demand Side Management method for smart natural gas supply systems[J]. Energy Conversion and Management 2019; 185: 368–383.
- [23] ZENG C, WEUC, ZUO L, et al. Predicting energy consumption of multiproduct pipeline using artificial neural networks[J]. Energy 2014; 66: 791–798.
- [24] 吴倩. 原油管道运行能耗统计分析预测[D]. 中国石油大学(北京), 2012. [WU Q. Statistical analysis and forecast of energy consumption in crude oil pipeline operation[D]. China University of Petroleum (Beijing), 2012.]
- [25] 温馨. 长输成品油管道能耗预测方法研究[D]. 西南石油大学, 2016. [WEN X. Study on energy consumption forecasting method of long-term oil pipeline[D]. Southwest Petroleum University, 2016.]
- [26] 侯磊, 许新裕, 崔金山, 等. 基于BP神经网络的输油管道能耗预测方法[J]. 节能技术, 2009, 27(5): 401–406. [HOU L, XU X Y, CUI J S, et al. A Prediction Method of Energy Consumption for Oil Pipeline Based on BP Neural Network[J]. Energy Conservation Technology, 2009, 27(5): 401–406.]
- [27] 高山卜, 钱成文, 张沛, 等. 基于改进的BP神经网络管输能耗预测模型[J]. 油气储运, 2014, 33(8): 869–872. [GAO S P, QIAN C W, ZHANG P, et al. Prediction model of pipeline energy consumption based on improved BP neural network[J]. Oil & Gas Storage and Transportation, 2014, 33(8): 869–872.]
- [28] 林冉, 朱英如, 余绩庆, 陈由旺, 等. 基于神经网络的热油管道能耗预测模型[J]. 石油石化节能, 2012, 2(01): 6–9. [LIN R, ZHU Y R, YU J Q, et al. The predicting model for the energy-consumption of the hot oil pipeline based on artificial neural network [J]. Energy Conservation in Petroleum & Petrochemical Industry, 2012, 2(01): 6–9.]
- [29] HU J M, WANG J Z, MA K L. A hybrid technique for short-term wind speed prediction[J]. Energy 2015, 81: 563–574.
- [30] MEENAL R, SELVAKUMAR A. Assessment of SVM, empirical and ANN based solar radiation prediction models with most influencing input parameters[J]. Renewable Energy, 121: 324–343.
- [31] TANG L, YU L, WANG S, et al. A novel hybrid ensemble learning paradigm for nuclear energy consumption forecasting[J]. Applied Energy 2012, 93: 432–443.
- [32] AHMAD T, CHEN H, HUANG R, et al. Supervised based machine learning models for short, medium and long-term energy prediction in distinct building environment[J]. Energy 2018, 158(1): 17–32.
- [33] HUANG N E, SHEN Z, LONG S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis[J]. Proceedings A 1998, 454(1971): 903–995.
- [34] Wu Z, HUANG N E. Ensemble empirical mode decomposition: A noise-assisted data analysis method[J]. Adv Adaptive Data Anal 2009, 1(1): 1–41.
- [35] TORRES M E, CLOLMINAS M A, SCHLOTTHAUER G, et al. A complete ensemble empirical mode decomposition with adaptive noise[C]. In: Acoustics, speech and signal processing (ICASSP), 2011 IEEE international conference on. IEEE; 2011.
- [36] MOSHE L, VLADIMIR Y, ALLAN P, et al. Multilayer feedforward networks with a nonpolynomial activation function can approximate any function[J]. Neural Networks 1993, 6(6): 861–867.
- [37] KENNEDY J, EBERHART R. Particle swarm optimization neural networks, 1995[C]. proceedings, IEEE Int. Conf. 4 (1995): 1942–194.
- [38] ZHENG W B. Research on gas load forecasting based on BP combination model[D]. Qingdao University 2016.
- [39] HENS A B, TIWARI M K. Computational time reduction for credit scoring: An integrated approach based on support vector machine

- and stratified sampling method[J]. *Expert Systems with Applications* 2012, 39(8): 6774–6781.
- [40] YE Y, WU Q, HUANG J Z, et al. Stratified sampling for feature subspace selection in random forests for high dimensional data[J]. *Pattern Recognition* 2013, 46(3): 769–787.
- [41] LI C, ZHU Z J, YANG H F, et al. An innovative hybrid system for wind speed forecasting based on fuzzy preprocessing scheme and multi-objective optimization[J]. *Energy* 2019, 174(1): 1219–1237.
- [42] WANG J Z, WU C Y, NIU T. A novel system for wind speed forecasting based on multi-objective optimization and echo state network[J]. *Sustainability* 2019, 11(2): 1–34.
- [43] WANG J Z, YANG W D, DU P, et al. A novel hybrid forecasting system of wind speed based on a newly developed multi-objective sine cosine algorithm[J]. *Energy Conversion and Management* 2018, 163: 134–150.
- [44] AHAMD T, CHEN H, GUO Y, et al. A comprehensive overview on the data driven and large scale based approaches for forecasting of building energy demand: A review[J]. *Energy and Buildings* 2018, 165: 301–320.

(编辑 马桂霞)