基于遗传算法优化Seq2Seq模型的建筑空调电能预测研究

李　鹏[1,2]，肖　辉[2]，李秀花[3]，赵　恒[4]，邵　頲[5]

（1. 中国电力工程顾问集团华东电力设计院有限公司，上海 200063；2. 同济大学电子与信息工程学院，上海 201804；3. 中国海诚工程科技股份有限公司，上海 200031；4. 上海硕物天成信息科技有限公司，上海 200052；5. 上海华建工程建设咨询有限公司，上海 200041）

摘要：介绍了一种建筑空调电能预测方法。空调电能占建筑总能耗的50%以上，空调电能的预测有利于可再生能源的安装容量预测和电力系统的负荷预测。采用遗传算法自适应搜索Seq2Seq模型的最优超参数，由此生成输出序列。通过与已有常见预测方法进行对比，证实了GA-Seq2Seq模型在MAE、RMSE和R^2等指标中表现优异。此外，还研究了GA-Seq2Seq模型的超参数和分类标签设置效果。

关键词：遗传算法；Seq2Seq模型；空调电能；能耗预测

Research on energy prediction of building air conditioning based on genetic algorithm optimization Seq2Seq model

LI Peng[1,2], XIAO Hui[2], LI Xiuhua[3], ZHAO Heng[4], SHAO Ting[5]

(1. East China Electric Power Design Institute Co., Ltd., of CPECC, Shanghai 200063, China; 2. Department of Electronic and Information Engineering, Tongji University, Shanghai 201804, China; 3. China Haisum Engineering Co.,Ltd., 200031, China; 4. Shanghai Shuowu Tiancheng Information Technology Co.,Ltd., Shanghai 200052, China; 5. Shanghai Huajian Architecture, Engineering & Consulting Co., Ltd., Shanghai 200041, China)

Abstract: This paper introduces a method of electric energy prediction for building air conditioning. The air-conditioning electric energy accounts for more than 50% of the total energy consumption of buildings, and the prediction of air-conditioning electric energy is beneficial to the installed capacity prediction of renewable energy and the load prediction of power system. In this paper, genetic algorithm is used to search the optimal hyperparameters of Seq2Seq model and generate the output sequence. In this research, it is confirmed that GA-Seq2Seq model has excellent performance in MAE, RMSE, R^2 and other indicators by comparing with the existing common forecasting methods. The hyperparameter and classification label setting effects of GA-Seq2Seq model are also studied.

Keywords: genetic algorithm; Seq2Seq model; electric energy of air condition; energy consumption prediction

0　引言

空调负荷是建筑物的主要能耗，通常占总能耗的50%以上。建筑物空调能耗的预测决定了建筑总能耗的预测[1]，从而影响了零碳建筑可再生能源的安装容量，对电力系统负荷预测与调度也起到一定作用；另一方面，掌握了空调能耗的预测数据，便于开展节能研究，有利于提高能源利用效率。

建筑物空调负荷与季节、天气、人员活动等因素密切相关，而人员活动因素则反映为时间、工作日等信息，因此建筑物空调负荷同时具备规律性和随机性。本研究利用某综合办公建筑的空调电能数据集，通过大数据资料获得建筑物空调负荷的规律。

1 预测算法

由于空调负荷在建筑能耗中的重要作用，其预测研究得到了较多重视。文献[2]采用深度学习算法对大型建筑中央空调能耗开展了预测，能够提高预测准确度。文献[3]基于长短期记忆（Long Short-Term Memory，LSTM）模型对空调能耗进行预测，预测效果优于传统方法。文献[4]基于Catboost和极端随机树方法，开展了办公建筑空调能耗预测，并指出了能耗峰值的时段。文献[5]采用软件模拟了办公建筑的空调能耗，并结合多组数据评价模型的准确性，得出了对空调能耗影响较大的几个因素。从上述文献可以看出，当前空调能耗预测研究偏重于空调自身的冷热能耗，而相对较少关注空调电能的预测。电能是更为通用的能源，通过对电能的预测，可以更加准确地反映空调系统所消耗能源对于电力系统预测的意义。

为了预测电力负荷，科研工作者采用了很多智能算法，其中Seq2Seq模型是应用较为广泛的方法[6]。文献[7]利用LSTM与Seq2Seq模型预测短期电力负荷，与传统预测方法相比，预测准确度显著提升。文献[8]利用长短序列双通道输入与Seq2Seq模型预测短期电力负荷，该模型的平均绝对误差（MAE）、均方根误差（RMSE）、平均绝对百分比误差（MAPE）和R^2指标更好。负荷预测模型中，文献[9]提出的多元自适应回归样条（MARS）模型比人工神经网络（ANN）和逻辑回归（LR）模型给出了更准确、更稳定的预测。文献[10]考虑到数据量庞大，采用聚类分析结合Seq2Seq模型进行负荷预测，取得了更好的结果。如果将气候因素考虑进实际预测中，会取得更好的预测效果[11]。不过，空调电能预测包括短期预测，预测任务常常集中在峰值和谷值时段，Seq2Seq模型无法提前获知重要数据的规律[12]，从而造成预测结果不准确。为了解决这个问题，本研究提出将遗传算法与空调电能负荷预测的Seq2Seq模型相结合。

结合每日空调电能负荷数据，首先将Seq2Seq模型的数据分成高、中、低三部分；然后考虑每日气候数据，并分类标注日负荷的特征负荷值；最后采用遗传算法在Seq2Seq模型中寻找最优超参数，以提高空调电能预测准确度。

2 数据处理

2.1 Seq2Seq模型和遗传算法

Seq2Seq是一种网络模型，由编码器和解码器两部分构成。Seq2Seq模型在时间序列多步预测领域取得了良好的效果。采用LSTM作为编码器和解码器，在电力系统预测中具有广泛应用，可以有效解决梯度消失问题和反向传播过程中出现的长期依赖问题[13]。在本研究中，Seq2Seq模型通过输入24弯矩负荷来预测未来4弯矩负荷，具有很好的拟合性。本研究采用的Seq2Seq模型单层结构如图1所示。

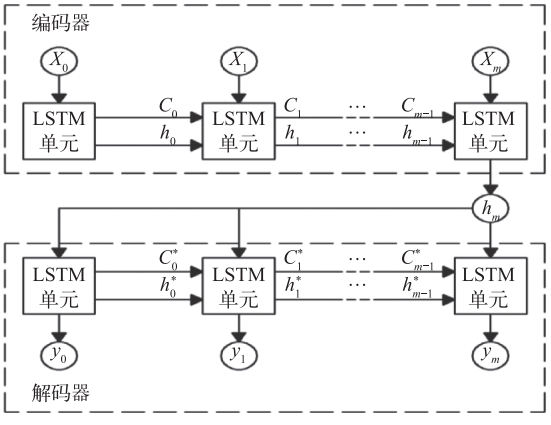


图 1 Seq2Seq 模型单层结构

Fig. 1 Structure of single-layer Seq2Seq model

图1中，Seq2Seq模型的输入序列$X_0 \sim X_m$作用在编码器上，输出序列$y_0 \sim y_m$作用在解码器上。其中，m时刻，隐藏节点h_m的参数和存储单元C_m作为状态向量传递给解码器，解码器据此生成输出序列$y_0 \sim y_m$。

Seq2Seq模型开始训练时，需要设置超参数，设置最优超参数对模型性能的提高具有重要意义[14]。遗传算法在优化搜索方面表现良好，能够自适应搜索最优超参数，应用在深度学习中，可对模型进行优化。因此考虑依靠遗传算法来优化Seq2Seq模型。

2.2 数据准备

本研究掌握了上海市某综合办公建筑2020年1月2日—2021年3月30日的完整空调电能数据集，并通过公开渠道获取了同时间段、同地点的气象数据。这些气象数据包括温度、湿度、可见度和风向等，可以和数据集中的电能数据相互对照，用于优化算法。分析数据集中空调电能数据，采用高度分段，如图2所示。

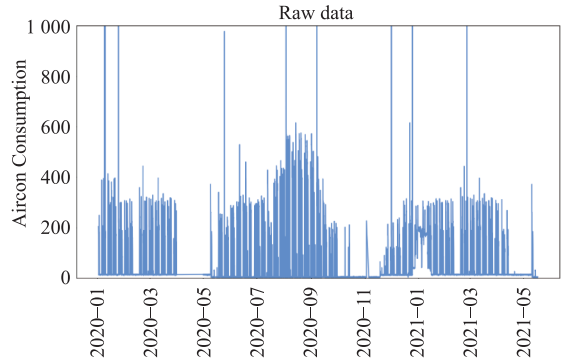


图 2 数据集空调电能数据图

Fig. 2 Original HVAC load data

2.3 特征选择

图2中，数据集中存在数据缺失值和离群值，需要进行数据处理，以提高空调电能模型预测准确度。采用将上一个时刻赋值给缺失值或离群值的方法，通过Python库中fill函数填充数据集。经过数据处理后，该数据集空调电能数据如图3所示。

由图3可知，经过数据处理后，该数据集不仅有效降低了数据维度，还去除了冗余和不相关特征。数据处理过程既提高了模型的预测准确度，也提高了数据集的泛化能力。考虑了空调电能与气象数据的关系，因此选取空调电能、温度、湿度、风力、风向、云层和能见度作为模型的输入，见表1。

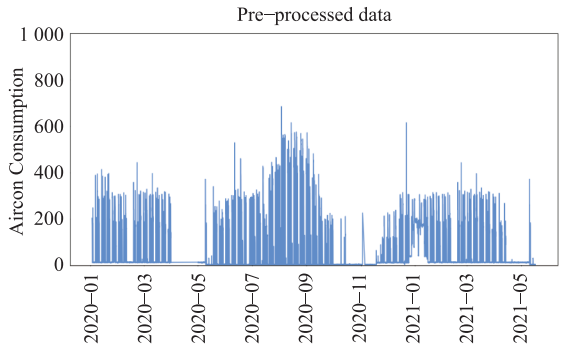


图 3 数据处理后空调电能数据图

Fig. 3 HVAC load data after data processing

表 1 模型输入参数

Tab. 1 Partial input features of the model

序号	输入参数
1	空调电能 /（kW·h）
2	温度 /℃
3	湿度（%）
4	风力 /（m/s）
5	云层（%）
6	能见度 /km

3 能耗预测

目前，很多文献提出了电能预测方法，其中GA-LSTM、BiLSTM 、CNN-LSTM和RNN等都已经得到了验证[15-17]。为了验证研究方法的优越性，与以上四种方法进行比较，以检查主要预测指标。

3.1 数据集分类

基于遗传算法（Genetic Algorithm，GA）优化Seq2Seq模型，简称GA-Seq2Seq。GA-Seq2Seq的建筑空调电能预测方法是否可行，可通过本研究掌握的数据集进行验证。将数据集分为训练集和测试集，其中2020年1月2日—2021年1月1日一整年的空调电能数据为训练集，2021年1月2日—2月1日的数据为验证集。其中，训练集采用GA-Seq2Seq模型进行训练，测试集用于验证模型的准确性。

3.2 模型预测对比

采用数据归一化方法，将所有特征值缩小到

0～1之间[18]，以便消除各特征值之间的尺度差异。本研究的模型输入特征有温度、湿度、可见度和风向等，预测值选择每日10:00。

采用Python程序完成，选择Adam作为优化器，并选择MSE均方误差损失函数为140次迭代的损失函数。选取的预测结果评价指标为MAE、RMSE和R^2，各预测方法的预测结果的对比如图4所示，各预测方法预测结果评价指标对比见表2。

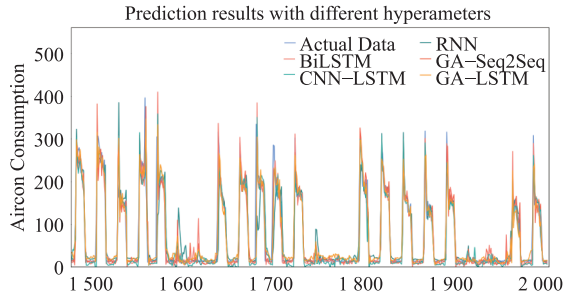


图 4　各预测方法预测负荷曲线

Fig. 4 Comparison of forecast load curves of each forecasting method

表 2　各预测方法预测结果评价指标对比

Tab. 2 Comparison of evaluation indexes of prediction results of each forecasting method

预测模型	MAE	RMSE	R^2
GA-Seq2Seq	0.0159 6	0.031 9	0.887 6
GA-LSTM	0.0167 1	0.032 4	0.884 3
BiLSTM	0.0165 5	0.029 2	0.905 8
CNN-LSTM	0.0175 9	0.033 6	0.875 3
RNN	0.0193 6	0.036 4	0.853 4

由表2可知，GA-Seq2Seq模型在MAE、RMSE和R^2等指标中都比其余预测方法表现更好。

3.3　模型优化

采用遗传算法对Seq2Seq模型进行优化，以找到最优超参数。将遗传算法的迭代次数设置为20次，并将每代种群的数量设置为20个，每个DNA设置为2个。GA-Seq2Seq的四种最优超参数与实际空调电能负荷的对比优化结果如图5所示，四种最优超参数预测评价指标对比见表3。

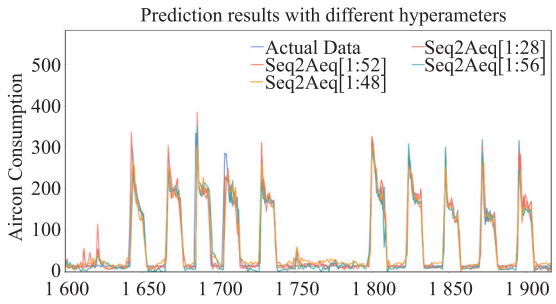


图 5　四种最优超参数预测负荷结果对比

Fig. 5 Comparison of the results of 4 optimal hyperparameters for load forecasting

表 3　四种最优超参数预测评价指标对比

Tab. 3 Comparison of 4 optimal hyperparameter prediction evaluation indexes

超参数 [层数：隐藏值]	MAE	RMSE	R^2
[1:52]	0.015 96	0.031 9	0.887 6
[1:56]	0.016 55	0.029 5	0.905 0
[1:48]	0.016 72	0.032 8	0.889 2
[1:28]	0.018 48	0.033 3	0.877 3

由表3可知，当层数为1、隐藏值为52时，GA-Seq2Seq模型的性能最佳。

对比了分类标签的设置效果。是否使用分类标签的空调电能预测曲线与实际空调电能曲线对比如图6所示，四种空调电能预测评价指标对比见表4。

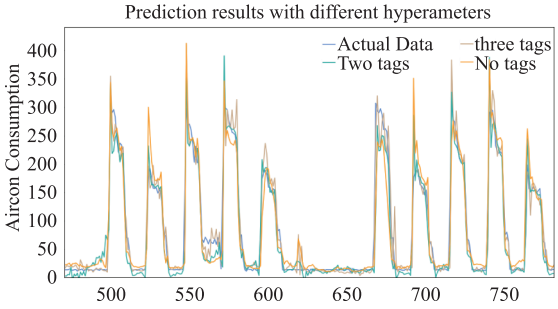


图 6　分类标签使用情况预测负荷结果对比

Fig. 6 Classification label usage forecast load comparison chart

表 4　四种空调电能预测评价指标对比

Tab. 4 Comparison of 4 kinds of HVAC energy prediction and evaluation indicators

标签数量	MAE	RMSE	R^2
0	0.016 72	0.032 0	0.897 3
1	0.016 70	0.029 2	0.887 8
2	0.016 56	0.031 9	0.905 5
3	0.015 90	0.032 0	0.887 8

由图6和表4可知，采用3个标签的方法效果最佳，而不使用分类标签的方法评价指标最差。

综上所述，GA-Seq2Seq模型的超参数优化可以提高预测准确度；通过设置分类标签，也可以提高预测评价指标。

4 结束语

为了能够预测空调电能，参考了空调能耗预测和电力预测方法，提出采用基于遗传算法优化的Seq2Seq模型。采用遗传算法自适应搜索Seq2Seq模型的最优超参数，据此生成输出序列，主要结论如下。

1）对比了GA-LSTM、BiLSTM、CNN-LSTM、RNN等已经得到验证的预测方法，对于空调预测研究而言，GA-Seq2Seq模型在MAE、RMSE和R^2等指标中都比其他预测方法表现优异。

2）对GA-Seq2Seq模型进行优化，以找到最优超参数，在此过程中，当层数为1、隐藏值为52时，GA-Seq2Seq模型的性能最佳。

3）对比了分类标签的设置效果，GA-Seq2Seq模型采用3个标签的方法效果最佳，而不使用分类标签的方法评价指标最差。

空调电能预测不仅对建筑自身能耗预测有价值，且有利于电力系统的负荷预测与调度，而且对于未来零碳建筑的实施具有重要作用[19]。零碳建筑能源系统为了匹配各类能源的储能参数，势必需要实现准确的能耗预测。为了更好地服务零碳建筑，未来将专注于包括空调电能在内的各类建筑系统的配置与预测。

参考文献

[1] 李鹏，肖辉，邵颋，等．民用建筑电气负荷计算研究[J]．电气应用，2023，42(10)：93-100．

[2] 冉均均，袁磊．基于深度学习的大型建筑中央空调能耗预测系统[J]．自动化技术与应用，2022，41(12)：1-4．

[3] 廖文强，王江宇，陈焕新，等．基于长短期记忆神经网络的暖通空调系统能耗预测[J]．制冷技术，2019，39(1)：45-50，54．

[4] 唐政磊，刘百慧．基于机器学习的办公建筑暖通空调系统能耗预测及优化调度[J]．中国新技术新产品，2023(2)：62-64

[5] 杨柳，侯立强，李红莲，等．空调办公建筑能耗预测回归模型[J]．西安建筑科技大学学报(自然科学版)，2015，47(5)：707-711．

[6] 徐先峰，王世鑫，龚美，等．引入BL-Seq2Seq模型进行负荷预测[J]．计算机仿真，2021，38(8)：103-107，501．

[7] 李建芳，纪鑫，张海峰，等．基于LSTM与Seq2Seq模型的短期电力负荷预测方法[J]．电子设计工程，2023，31(24)：150-153，158．

[8] 李国栋，王春红，刘嘉城，等．基于L-S-Seq2Seq的双通道超短期电力负荷预测[J]．电力信息与通信技术，2023，21(10)：56-65．

[9] NALCACI G，ÖZMEN A，WEBER G W．Long-term load forecasting：models based on MARS，ANN and LR methods[J]． Central European Journal of Operations Research，2019 (27)：1033-1049．

[10] MASOOD Z，GANTASSI R，CHOI Y．A multi-step time-series clustering-based Seq2Seq LSTM learning for a single household electricity load forecasting[J]．Energies，2022，15(7)：2623．

[11] IMANI M．Electrical load-temperature CNN for residential load forecasting[J]．Energy，2021，227：120480．

[12] CHEN Zhe，CHEN Yongbao，XIAO Tong，et al．A novel short-term load forecasting framework based on time-series clustering and early classification algorithm[J]．Energy and Buildings，2021，251：111375．

[13] 程冬梅．LSTM研究现状综述[J]．信息系统工程，2022(1)：149-152．

[14] YANG L，SHAMI A．On hyperparameter optimization of machine learning algorithms：Theory and practice[J]．Neurocomputing，2020，415：295-316．

[15] 周璇，林家泉．基于改进长短时记忆网络的地

面空调能耗预测[J]．北京航空航天大学学报，2023，49(10)：2750-2760．
[16] 王东风，刘婧，黄宇，等．结合太阳辐射量计算与CNN-LSTM组合的光伏功率预测方法研究[J]．太阳能学报，2024，45(2)：443-450．
[17] 程换新，黄震．基于改进PSO优化RNN的短期电力负荷预测模型[J]．电子测量技术，2019，42(20)：94-98．
[18] LEMUEL C P V，KARL A S A，JUSTINE S S M，et al．Performance analysis of artificial neural network models for hour-ahead electric load forecasting[J]．Procedia Computer Science，2022，197：16-24．
[19] 李鹏，肖辉，邵颀．低碳办公建筑储能系统优化配置研究[J]．电气应用，2023，42(2)：1-8．

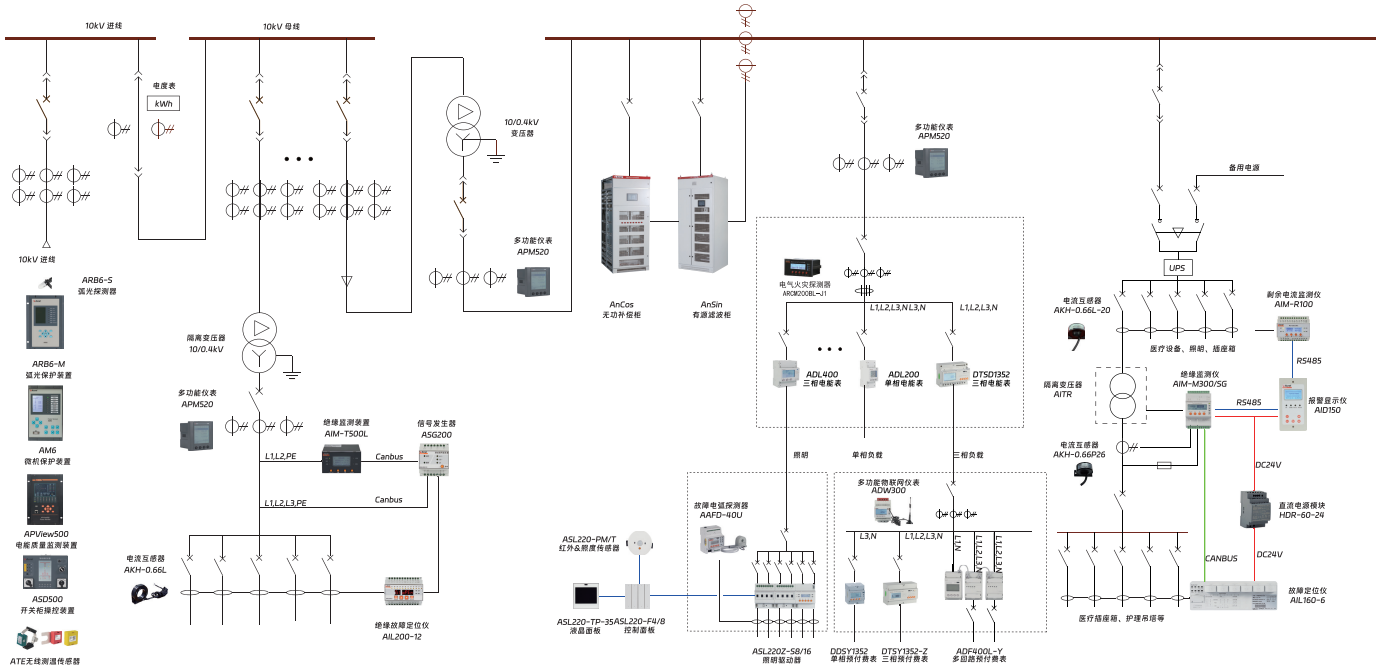
作者简介

李鹏（1986—），男，博士，高级工程师，研究方向为变电电气一次系统、智能照明控制、零碳建筑能源系统。

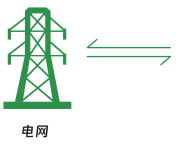
肖辉（1969—），女，教授，研究方向为智能照明控制、模式识别与人工智能。

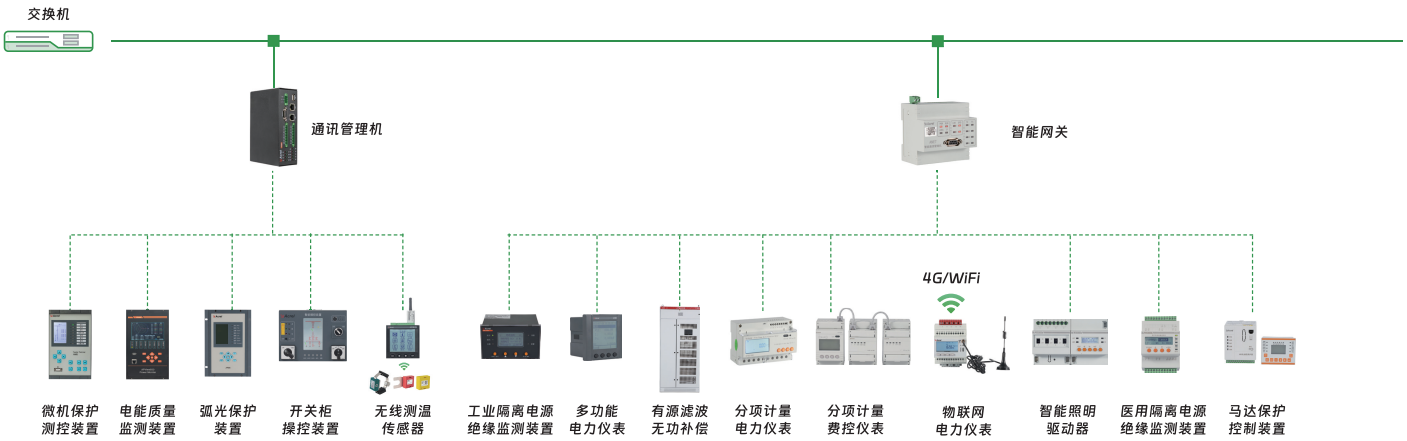
（收稿日期：2024-10-20）

AcrelEMS企业



电力保护监控　　工业隔离电源系统　　智能照明控制系统　　分项计量及综合计费系统　　医疗隔离电源系统





变电站监控保护　　电能质量治理及分项计量　　安全、节能控制

地　　址：上海市嘉定区育绿路253号

邮　　编：201801

传　　真：021-69158303

订货电话：400-8208615

服务电话：800—820-6632

http://www.acrel.cn

E-mail：ACREL001@vip.163.com

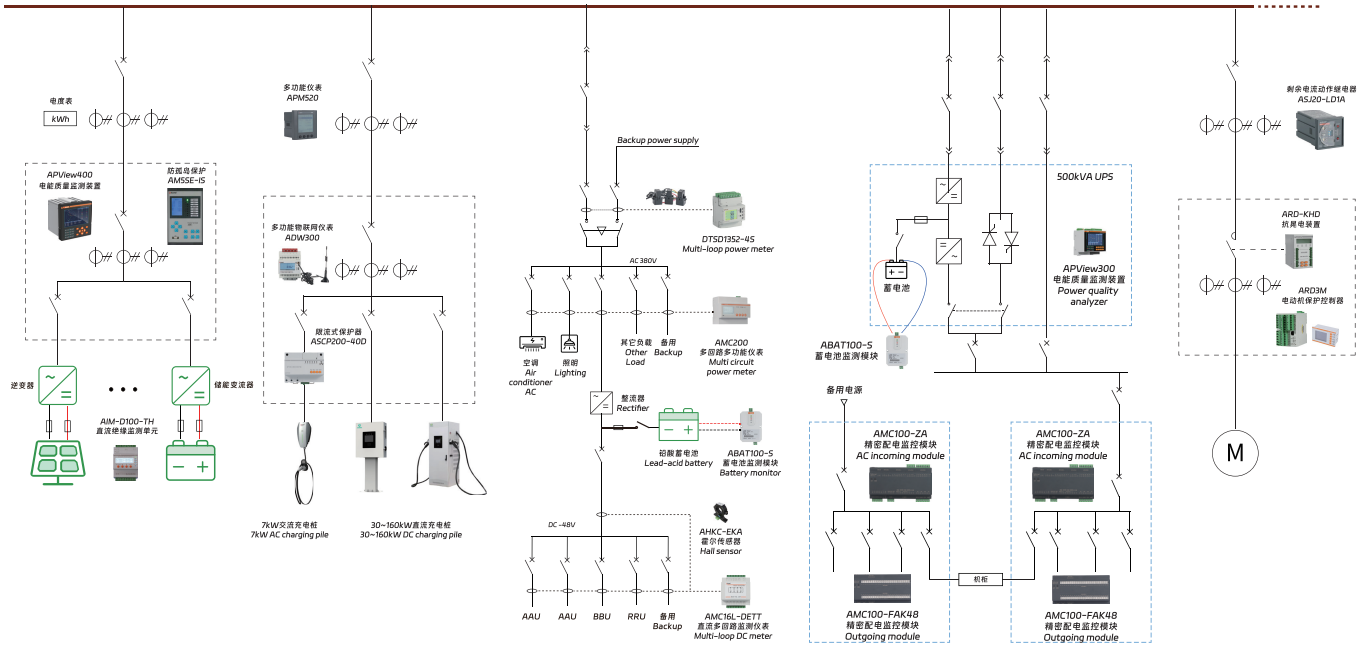
公众号

安科瑞电气股份有限公司

广告 | AD

[股票代码：300286. SZ]

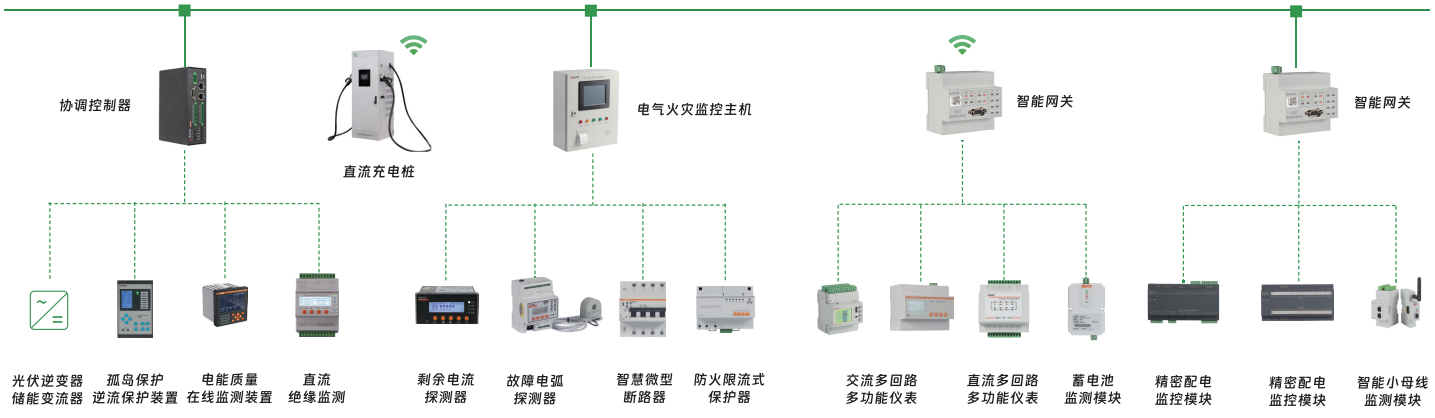
微电网能效管理平台



光伏储能管理系统　有序充电系统　铁塔基站智慧用电　数据中心精密配电　马达控制系统

AcrelEMS
企业微电网能效管理平台

虚拟电厂或
碳排放监管平台



新能源策略控制　安全用电管理　铁塔基站智慧用电　数据中心精密配电

直通车

安科瑞电气股份有限公司

E选型

安科瑞电气股份有限公司