

基于稳健损失函数的 Bi-LSTM 短期电力负荷预测方法研究

谢金润¹ 薛杰仁²

1 南京工程学院电力工程学院沈国荣学院, 江苏南京, 211167;

2 河海大学电气与动力工程学院, 江苏南京, 211100;

摘要: 短期电力负荷预测在电力系统调度与能源管理中具有重要意义。针对传统长短期记忆网络 (LSTM) 在存在异常尖峰时预测精度下降的问题, 本文提出一种基于稳健损失函数的双向长短期记忆网络 (Huber Bi-LSTM) 预测方法。该方法在网络输出层引入 Huber 损失函数, 通过在小误差区采用二次惩罚、大误差区采用线性惩罚, 降低异常点对梯度更新的影响, 提升模型鲁棒性。以某地区 5 日总负荷数据为例, 采用滑动时间窗构造输入特征, 对未来 24 小时负荷逐点预测。实验结果表明, 所提方法在归一化域 sMAPE 稳定在 1.97% 左右, 峰谷拟合精度较传统 MSE 损失 Bi-LSTM 明显提高, 并有效抑制了异常值引起的预测波动。该方法在保证整体趋势预测精度的同时, 增强了模型在异常数据下的稳定性, 具有较好的工程应用价值。

关键词: 短期电力负荷预测; 双向长短期记忆网络 (Bi-LSTM); 稳健损失函数; 异常值抑制

DOI: 10.69979/3060-8767.25.10.019

1 背景介绍

随着电力系统规模的不断扩大和用电结构的日益复杂, 负荷预测在电网运行、规划以及电力市场交易等环节中的作用愈加突出^[1]。短期负荷预测 (Short-Term Load Forecasting, STLF) 能够为机组调度、需求响应和储能优化提供及时参考, 是保障电力系统安全性与经济性的重要基础。近年来, 随着分布式能源和可再生能源比例的上升, 负荷曲线呈现出更强的波动性和不确定性, 对预测方法的精度与稳定性提出了更高要求^[2]。

在传统方法中, ARIMA 和小波分析多数用于线性或局部频率分析, 支持向量机 (SVM)、随机森林 (RF) 等机器学习模型虽能提升预测能力, 但难以有效建模非线性长时依赖关系^[3]。LSTM 在处理复杂时序方面表现优异, 在电力负荷预测中已有诸多成功应用^{[4][5]}。

现实负荷数据通常包含设备故障、通信丢包或极端天气下的异常尖峰, 这些异常值会对传统基于均方误差 (MSE) 的训练策略产生负面影响, 引起峰谷预测偏差与波动振荡^[6]。一些工作尝试引入异常检测、数据清洗或多模型融合等办法来增强鲁棒性^[7], 但这些方法往往依赖复杂特征工程与额外计算开销, 不适合实时短期预测系统。相比之下, 损失函数的替换是一种结构简洁又易实现的优化手段, 其中 Huber 损失函数的稳定性受到了研究青睐^[8]。

基于此, 本文提出在 Bi-LSTM 回归输出层采用 Hu-

ber 损失函数, 通过在小误差区间保持二次惩罚, 大误差区间转为线性惩罚, 有效抑制异常样本对训练的影响。同时, 结合滑动时间窗输入策略和双向结构, 提升模型对负荷上下文变化的捕捉能力, 实现未来 24 小时逐点预测。

2 算法设计

本文提出的预测系统以历史负荷时间序列为输入, 通过滑动时间窗抽取固定长度的历史片段作为模型的输入特征; 核心预测模块采用双向长短期记忆网络从前后两个方向提取时序依赖特征; 在回归输出层引入 Huber 损失函数, 以在小误差区保持二次惩罚、大误差区转为线性惩罚, 从而削弱异常点对参数更新的过度影响。

2.1 数据预处理与输入构造

预处理阶段, 模型对训练段负荷数据进行线性归一化, 将数值映射到 [0, 1] 区间, 以避免不同量纲导致的训练不稳定; 再用训练段的归一化参数对验证集和测试集进行变换, 防止信息泄漏。设原始负荷序列为 $(y_t)_{t=1}^T$, 训练段最小值与最大值分别为 $y_{\min}$ 与 $y_{\max}$, 归一化公式为:

$$\tilde{y}_t = \frac{y_t - y_{\min}}{y_{\max} - y_{\min}}$$

W 的滑动时间窗构造输入-输出对。对任意时刻 t , 输入向量由前 W 个归一化观测组成, 输出为时刻 t 的归一化负荷:

$$x_t = [\tilde{y}_t - W, \tilde{y}_t - W + 1, \dots, \tilde{y}_t - 1]^T, \quad y_t^* = \tilde{y}_t$$

2.2 Bi-LSTM 模型与稳健损失函数

Bi-LSTM 是一种能够同时捕捉前向和后向时序依赖的循环神经网络结构。第一层 Bi-LSTM 输出整个时间窗的特征序列，第二层 Bi-LSTM 仅取最后一个时间步的隐藏状态作为整个输入窗的压缩表示：

$$\hat{y}_t = f_{\Theta}(x_t)$$

其中 $f_{\Theta}(\cdot)$ 表示由网络参数 Θ 定义的非线性映射。

为增强模型对异常值的容忍度，回归端采用 Huber 损失函数，其定义为：

$$L_{\delta}(y, \hat{y}) = \begin{cases} \frac{1}{2}(y - \hat{y})^2, & \text{if } |y - \hat{y}| \leq \delta \\ \delta \cdot |y - \hat{y}| - \frac{1}{2}\delta^2, & \text{otherwise} \end{cases}$$

其中 $\delta > 0$ 为分段拐点，小误差区等价于均方误差（MSE），大误差区等价于平均绝对误差（MAE），既保留了 MSE 在小误差下的平滑特性，又具备 MAE 的稳健性。批量训练下，目标函数为：

$$\mathcal{J}(\Theta) = \frac{1}{N} \sum_{t=1}^N L_{\delta}(f_{\Theta}(x_t), y_t^*) + \lambda \|\Theta\|_2^2$$

其中 N 为训练样本数， λ 为权重衰减系数。本文默认设置 $\delta = 0.02$ ，在结果分析中对其敏感性进行讨论。

2.3 训练、预测与评估

训练阶段采用 Adam 优化器，学习率分段衰减，并结合 Dropout 抑制过拟合。预测阶段按滑窗推进逐点输出未来 24 小时的归一化预测值，再通过构造公式的逆变换恢复到原始量纲。

为全面衡量模型性能，本文使用均方根误差（RMSE）、平均绝对误差（MAE）及对称平均绝对百分比误差（sMAPE）：

$$\begin{aligned} \text{RMSE} &= \sqrt{\frac{1}{T} \sum_{t=1}^T (y_t - \hat{y}_t)^2} \\ \text{MAE} &= \frac{1}{T} \sum_{t=1}^T |y_t - \hat{y}_t| \\ \text{sMAPE} &= \frac{2}{T} \sum_{t=1}^T \frac{|y_t - \hat{y}_t|}{|y_t| + |\hat{y}_t| + \varepsilon} \end{aligned}$$

其中 ε 为避免零分母的微小常数。

3 实验分析

3.1 数据集与实验环境

本文实验所用数据为某区域的实际电力负荷数据，采样间隔为 15 分钟，选取连续 5 日数据作为建模样本。实验平台为 MATLAB R2022b，CPU 为 Intel i7 处理器，

内存 32GB，操作系统为 Windows10。

原始数据示例如图 1 所示，可以看出负荷曲线呈现明显的日周期性特征，且早晚高峰波动较大，存在一定异常波动点，这对传统基于 MSE 损失的模型会带来一定预测偏差。

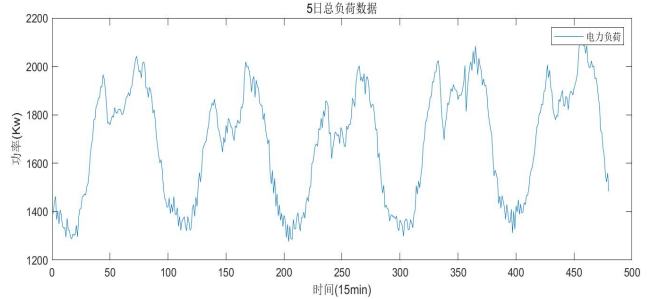


图 1 5 日总负荷曲线

3.2 预测结果对比

图 3-2 给出了基于 Huber 损失 Bi-LSTM 模型的预测曲线与真实值对比结果。可以看出，该模型能够较好地跟随真实负荷变化趋势，在早晚高峰等波动剧烈时段也能保持较好的拟合效果，峰值位置与幅度较为接近，谷值预测误差较小。

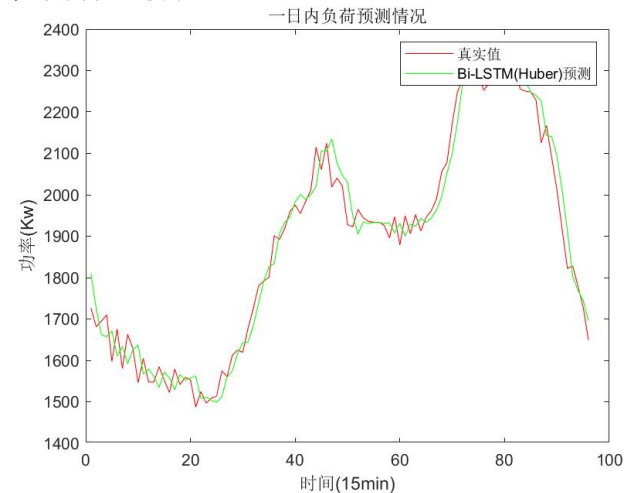


图 2 一日内负荷预测结果

为进一步评估模型性能，计算了 RMSE、MAE 与 sMAPE 三项指标，结果如下：

$$\text{RMSE} = 46.464, \quad \text{MAE} = 36.896, \quad \text{sMAPE} = 1.97\%$$

3.3 逐时误差分布分析

图 3-3 展示了预测结果的逐时误差分布。从下半部分的误差图可以看出，大部分预测点误差集中在 $\pm 50\text{kw}$ 以内，且未出现明显系统性偏差。个别高峰时刻出现的预测高估/低估现象，主要与原始负荷数据突变和采样噪声有关，但在 Huber 损失的抑制下，预测曲线波动幅

度明显低于传统 MSE 模型。

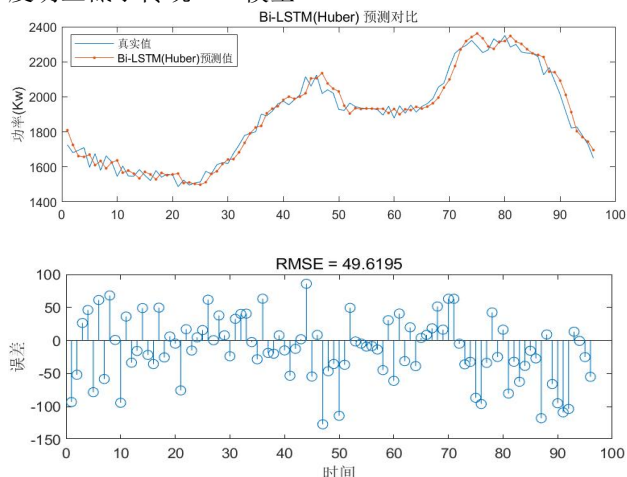


图3 预测误差分布

4 总结

本文针对短期电力负荷预测中负荷波动剧烈、对异常数据敏感性高等问题,提出了一种引入稳健损失函数的双向长短期记忆网络(Bi-LSTM)预测方法。该方法在传统 Bi-LSTM 架构中融入 Huber 损失函数,在小误差区采用二次惩罚、在大误差区采用线性惩罚,有效抑制了异常点对模型训练的不利影响,从而提升了预测的鲁棒性与泛化性能。实验结果表明,该方法在保持较高预测精度的同时,对异常波动具有更强的适应能力,为短期电力负荷预测提供了一种兼顾精度与稳定性的可行方案,具备一定的工程应用潜力。未来工作计划从两个方面展开:引入注意力机制(Attention)以增强模型对关键时段特征的关注度;或者进一步尝试扩展至多源信息融合预测场景,进一步利用气象、节假日等外部因素提升模型的综合预测性能。

参考文献

- [1] 杨佳泽,王灿,王增平. 新型电力系统背景下的智能负荷预测算法研究综述[J]. 华北电力大学学报(自然科学版), 2025, 52(03): 54-67.
- [2] 王海鹏. 注意力机制下多尺度时间卷积的短期电力负荷预测[D]. 南昌大学, 2022. DOI: 10. 27232/d. cnki. gnchu. 2022. 001523.
- [3] 肖荣鸽,夏海平,李雨泽,等. 基于 L1-SSA-SVM 模型的城镇燃气短期负荷预测[J]. 区域供热, 2024, (06): 150-158. DOI: 10. 16641/j. cnki. cn11-3241/tk. 2024. 06. 021.
- [4] 鲁娟,李明海,张柄涛,等. 基于 PSO-BI-LSTM 模型的短期电力负荷预测[J]. 建筑电气, 2025, 44(07): 38-42.
- [5] 徐亮,洪炎,苏静明,等. 基于多尺度深度可分离卷积和双向 Transformer 的短期电力负荷预测[J/OL]. 重庆工商大学学报(自然科学版), 1-12[2025-08-09]. <https://link.cnki.net/urlid/50.1155.N.20250703.1126.002>.
- [6] 张珂豪,刘庆杰,陈八,等. 基于 Informer-LSTM 方法的高邮台地电场数据异常检测研究[J]. 防灾科技学院学报, 2025, 27(02): 79-88.
- [7] 杨丹,陈俊. 基于相似日的多模型融合短期电力负荷预测[J]. 电脑与信息技术, 2025, 33(01): 6-9+54. DOI: 10. 19414/j. cnki. 1005-1228. 2025. 01. 028.
- [8] 康佳,姜宏伟. 基于非独立同分布样本 Huber 损失回归算法的泛化性能[J]. 兰州文理学院学报(自然科学版), 2025, 39(03): 1-7+23. DOI: 10. 13804/j. cnki. 2095-6991. 2025. 03. 020.