

基于竞争学习机制的 LSTM 风电多目标区间预测

任鹏 付文杰 申洪涛 陶鹏 张洋瑞

(国网河北省电力有限公司营销服务中心 河北 石家庄 050000)

摘要 为了进一步提升区间预测的综合效果,提出一种基于竞争学习机制的 LSTM 风电多目标区间预测方法。提出基于 LSTM 的上下界估计模型来构建风电区间多目标预测模型,研究多目标系统中的估计误差与预测区间平均宽度的关系;进一步考虑预测误差,引入一种新的偏最小二乘评价指标;通过引入竞争学习机制,提出一种改进的非支配快速排序遗传算法,有效实现了多目标优化。最后通过两个实际风电数据集进行实验验证,结果表明提出方法具有较高的预测精度。

关键词 风电预测 长短期记忆网络 区间预测 遗传算法

中图分类号 TP3 TM715 文献标志码 A DOI:10.3969/j.issn.1000-386x.2024.06.044

MULTIPLE OBJECTIVE INTERVAL PREDICTION OF LSTM WIND POWER BASED ON COMPETITIVE LEARNING MECHANISM

Ren Peng Fu Wenjie Shen Hongtao Tao Peng Zhang Yangrui

(State Grid Hebei Marketing Service Center, Shijiazhuang 050000, Hebei, China)

Abstract In order to further improve the comprehensive effect of interval prediction, a multiple objective interval prediction method of LSTM wind power based on competitive learning mechanism is proposed. The upper and lower bounds estimation model based on LSTM was proposed to construct the multiple objective prediction model of wind power interval, and the relationship between the estimation error and the average width of prediction interval in the multiple objective system was studied. Further considering the prediction error, a new partial least squares evaluation index was introduced. In addition, by introducing competitive learning mechanism, an improved non dominated quick sort genetic algorithm was proposed, which effectively realized multiple objective optimization. Two real wind power data sets were used to verify the proposed method. The results show that the proposed method has high prediction accuracy.

Keywords Wind power forecasting Long and short-term memory network Interval forecasting Genetic algorithm

0 引言

风力发电作为一种清洁、经济的发电资源,在现代电力系统中的应用越来越广泛^[1]。风力发电的大规模普及,可以减少 SO₂ 和 NO 等污染气体的排放,并且缓解能源短缺问题^[2]。然而,风电的大规模普及也给电力系统的供电侧带来了很大的不确定性,使得系统运行更具挑战性。

近年来,概率预测方法被广泛研究,以有效地量化风电不确定性不同于传统的点预测方法,作为上述方

法中最直观的一种,区间预测引起了人们的广泛关注。传统的区间预测方法,如模糊推理^[3]、贝塔分布函数^[4]和高斯过程^[5]有特定的数据分布假设。作为一种非参数方法,最近提出的上下界估计(Lower and Upper Bound Estimation, LUBE)不需要对数据分布做任何假设,直接从输入构造预测区间,可以有效地解决上述问题^[6]。一般而言,可以在单目标和多目标框架下解决润滑问题。Li 等^[7]提出了一个单目标 LUBE 框架,在覆盖概率约束下使预测区间的平均宽度最小,以获得更好的训练性能。李鉴博等^[8]利用经验模态分解和样本熵技术对风电序列进行分解和分组,然后将

ELM 应用于多目标框架。孙勇等^[9]应用径向基函数神经网络(RBF)在多目标框架下构造预测区间,利用 k 均值聚类算法确定 RBF 的中心。但是上述研究都是基于浅层学习模型。然而,在处理时间序列预测问题时,可能很难映射出动态的投入产出关系,这将限制 LUBE 的性能。近年来,深度神经网络由于具有很强的非线性映射能力,引入了风电预测领域,但现有的研究大多集中在点预测上,对于带有极强随机性的区间预测研究极少。薛阳等^[10]提出一种基于双向长短期记忆网络(BiLSTM)和自助法(Bootstrap)的风电功率区间预测方法,突出了风电功率数据间关联程度的重要性。熊鸣^[11]提出一种基于 LSTM 和非参数核密度估计的风电功率概率区间预测,采用渐进积分均方误差准则的窗宽优化方法实现非参数核密度的估计。任文凤等^[12]了构建 LSTM-VMD 组合区间预测模型,实现模型预测效果的提升。但是上述方法由于对于区间预测综合效果包括训练时间和预测精度等具有进一步探索的价值。

为了进一步提升预测效果,提出了一种基于竞争学习机制的 LSTM 风电多目标区间预测方法。首先引入一种新的偏最小二乘评价指标,然后将基于 LSTM 的上下界估计模型来构建风电区间预测模型,进一步通过引入竞争学习机制,提出了一种改进的 NSGA-II 多目标优化算法,有效提升了区间预测的速度与精度。基于真实风电数据集的一系列实验验证了提出模型的有效性。

1 基本理论

与传统的点预测方法不同,区间预测的是用预测区间来检测不确定性。传统的区间预测方法分两步构建预测区间:首先用回归模型对未来风电进行点预测,然后根据数据分布得到上下界。然而,由于点预测存在预测误差,假设的数据分布在实际中可能不适用,因此传统方法无法获得高质量预测区间。

近年来,LUBE 已成功地应用于风电、未来负荷和太阳能发电的区间预测。LUBE 仅需一个步骤就可以生成预测区间,其中两个输出为下限和上限。具体而言,LUBE 模型可以看作是从输入到预测区间的直接映射,而不需要任何数据分布的假设。此外,准确度和效率是 LUBE 研究的两个主要问题:高精度的预测区间意味着可以有效地量化不确定性,而高效率的预测区间则意味着能为系统操作员提供有价值的信息。因此,如何评价预测区间的准确性和效率是 LUBE 领域的一个重要课题。

1.1 预测区间覆盖率

PICP 是预测区间的一个重要评价指标,它表示未来风电大小在预测的上下限中的概率。显然,较大的 PICP 意味着构建的预测区间可以更准确地表示未来风电值的不确定性。PICP 定义如下:

$$P_{\text{PIC}} = \frac{1}{N} \sum_{i=1}^N a_i \quad (1)$$

式中: N 为试验数据样本数; a_i 为二进制值参数,计算公式如式(2)所示。

$$a_i = \begin{cases} 1 & y_i \in [y_{i,l}, y_{i,u}] \\ 0 & y_i \notin [y_{i,l}, y_{i,u}] \end{cases} \quad (2)$$

式中: y_i 是测试样本 i 的值; $y_{i,l}$ 和 $y_{i,u}$ 是预测的上下界。为了保证预测区间的准确性,通常要求 PICP 大于预先设定的置信水平 $(1 - \alpha)$ 。

1.2 预测区间归一化平均宽度

虽然 PICP 是预测区间精度的一个关键指标,但是却忽略了预测区间的效率,因为用一个区间较大的预测区间可以很容易地得到一个较大的 PICP,甚至接近 100%。基于上述考虑,引入 PINAW 来评估:

$$W_{\text{PINAW}} = \frac{1}{NW} \sum_{i=1}^N (y_{i,u} - y_{i,l}) \quad (3)$$

式中: W 是预测区间取值范围。标准化的目的是忽略风力的大小。

1.3 基于覆盖宽度的准则

一般而言,在预测预测区间时,PICP 和 PINAW 是两个相互冲突的问题:增加覆盖率必然会使宽度变大,而缩小宽度会导致覆盖率降低。为了评价预测区间的整体性能,将 PICP 和 PINAW 结合起来,得到 CWC,其计算公式如式(4)所示。

$$C_{\text{CW}} = W_{\text{PINAW}} (1 + \gamma (P_{\text{PIC}}) e^{-\eta (P_{\text{PIC}} - \mu)}) \quad (4)$$

式中: γ 是二进制参数,定义如式(5)所示。

$$\gamma = \begin{cases} 0 & P_{\text{PIC}} > \mu \\ 1 & P_{\text{PIC}} \leq \mu \end{cases} \quad (5)$$

式中: γ 和 μ 是 C_{CW} 的两个控制参数。 γ 反映了预测区间的覆盖概率要求,可根据预先确定的置信水平 $(1 - \alpha)$ 进行计算, η 是当获得的预测区间不能满足覆盖概率要求时的惩罚系数。

2 估计误差

当采用现有的 PICP 索引时,测试数据样本被转换成二进制变量来说明它们是否在预测范围内。由于连续值比二进制变量能传递更多的信息,如果采用 PICP,风电数据集提供的数据信息就没有得到充分利

用。此外,现有的研究大多是研究覆盖率,而忽略了超过区间的误差,这是不合理的,因为超出预测区间可能会对决策过程产生重大影响。

随着风电的大规模普及,电力系统的供电侧存在着很大的不确定性,使得组合问题的求解更加困难。在相关文献中,鲁棒电力系统机组组合问题被认为是处理风电不确定性的一种有效方法,通过在最坏情况下做出决策来确保供电可靠性。尽管具有鲁棒性的电力系统机组组合可以处理预测区间中的风电不确定性,但如果未来的风力发电不在预测区间范围内,风险仍然存在。如图 1 所示,当风电功率超过上限时,风电系统将无法整合。另一方面,当未来的风力发电量小于预测区间的下限时,供电系统可能没有足够的电量来满足用户的用电需求,从而导致甩负荷。因此,超过预测区间的预测范围误差会对鲁棒电力系统机组组合问题的操作风险产生巨大影响。因此,为保证供电可靠性,降低运营成本,在预测区间时应考虑风电的估算误差。文献[13]利用 Winkler 评分 (WS) 来评价预测区间的性能,计算方法是预测区间的宽度和估计误差的加权和。

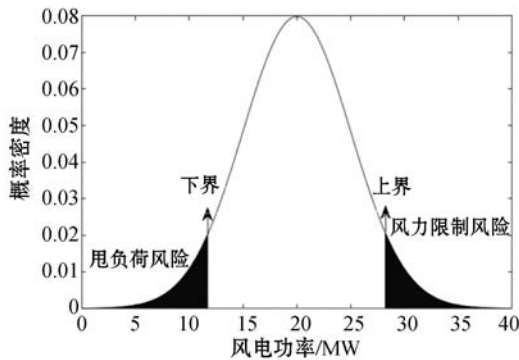


图 1 甩负荷和风力限制风险

通常,在一个预先确定的置信水平下,具有较小绝对值 WS 的预测区间的质量较高。然而,WS 无法区分平均宽度和估计误差对结果的影响。在鲁棒电力系统机组组合问题等实际应用中,操作人员需要知道预测区间的估计误差,从而对操作风险进行评估。此外,为了降低操作风险,在训练区间预测模型时,需要独立最小化预测区间的估计误差。为了克服上述缺陷,本文提出了一种新的评价指标,即预测区间估计误差 (PIEE),如式(6)所示。

$$E_i = \begin{cases} (y_i - y_{i,u}) & y_i \geq y_{i,u} \\ 0 & y_{i,l} \geq y_i \leq y_{i,u} \\ (y_{i,l} - y_i) & y_{i,l} \geq y_i \end{cases} \quad (6)$$

所有 PIEE 的计算公式如式(7)所示。

$$E_{PIE} = \frac{1}{NW} \sum_{i=1}^N E_i \quad (7)$$

基于上述 PIEE 指数,可以用预测区间的平均宽度单独测量预测区间的估计误差,这样,决策者就可以对预测区间之外的操作误差进行评估并有效降低误差。虽然 PIEE 可以用来衡量预测区间的准确性,但是预测区间的效率却被忽略了。例如,当使用非常宽的 PI 时,可以很容易得到一个较小的 PIEE。然而,这是毫无意义的,因为当预测区间范围非常大时,没有关于未来风电不确定性的有用信息传达给系统运营商。因此,预测区间的范围也是 LUBE 的重要因素,通常由 PINAW 进行评估。基于以上考虑,将 PIEE 与 PINAW 一起进行优化,从而获得准确率较高的预测区间。

3 基于 LSTM 的区间预测模型

RNN 通过隐藏层之间的连接,可以有效地捕捉输入的序列特征,然而,RNN 存在梯度消失和梯度爆炸的问题。作为 RNN 的一种改进模型,LSTM 可以较好地解决消失梯度和爆炸梯度问题,LSTM 的模型结构如图 2 所示。与经典的全连接神经网络不同,LSTM 可以展开为具有相互关联的隐藏层的各种神经网络。在时间步骤 t 中,当前输入 X_t 与先前的隐藏层状态 h_{t-1} 结合,共同计算当前隐藏层的状态 h_t ,因此,先前输入 $\{X_{t-n}, X_{t-(n-1)}, \dots, X_{t-1}\}$ 包含的信息存储在当前隐藏层中。由于目前的输出 O_t 是根据 X_t 和 h_t 得到的,上一步的输入将对 O_t 会有较大影响,因此 LSTM 可以作为获取风电时间序列特征的有效模型。

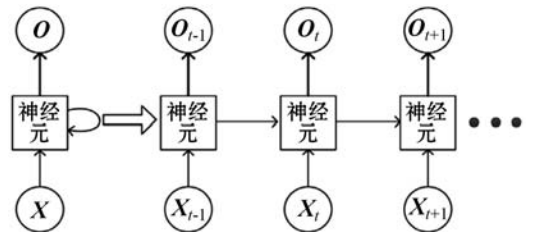


图 2 LSTM 结构图

为了解决 RNN 的消失梯度和爆炸梯度问题,在 LSTM 中引入了神经元结构,如图 3 所示。在神经元中有各种各样的门,来决定添加或删除当前单元的状态信息。每个门都由一个位乘法运算和一个 Sigmoid 激活函数组成,用来选择性地让信息通过。输入门决定是否将先前的隐藏层状态和当前的输入添加到当前单元状态。遗忘门决定在计算当前单元状态时是否应该使用前一个单元状态。输入门和输出门的输出见式(8)和式(9)。

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (8)$$

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (9)$$

式中: σ 是激活函数; h_{t-1} 是上一步的神经元状态; x_t 是当前输入; W_i 、 W_f 、 b_i 和 b_f 表示权重矩阵和偏差向量。

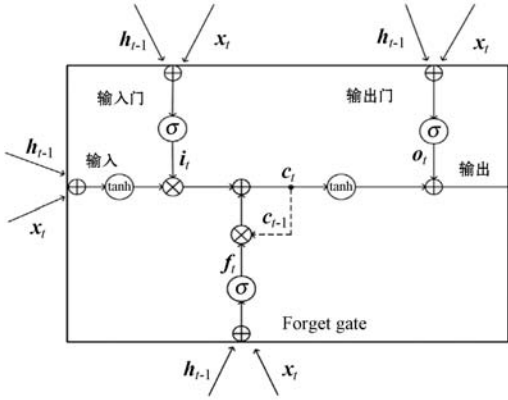


图3 LSTM神经元的结构

$$\hat{c}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (10)$$

式中: \hat{c}_t 是用来计算 c_t 的临时变量, b_c 是偏差向量。

因此,当前神经元状态 c_t 的计算如下:

$$c_t = f_t c_{t-1} + i_t \hat{c}_t \quad (11)$$

根据当前神经元的输出状态,决定是否使用下一个神经元的输出状态:

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (12)$$

因此,神经元的输出如下。

$$h_t = o_t \tanh(c_t) \quad (13)$$

最后,基于 LSTM 的区间预测模型的结构如图 4 所示。



图4 LSTM的区间预测模型结构

4 改进后的快速非支配排序遗传算法

遗传算法是一种基于自然选择和进化规则的元启发式算法。由于其快速的收敛速度和全局搜索能力,广泛地应用于各个领域,其中 NSGA-II 是最有效的多目标优化算法之一。通过引入精英原理和快速排序机制,证明了 NSGA-II 计算速度快,具有较强的全局搜索能力。近年来,NSGA-II 已应用于解决电力系统机组组合问题、经济调度、最优潮流和预测风电间隔等工程问题,本文通过引入竞争学习策略对 NSGA-II 做进一步的改进。

在竞争学习策略中,每个个体是通过从精英策略中学习进化的。具体而言,首先根据 NSGA-II 中的密度估计方法计算每个精英策略的密度,然后从目前的

种群中随机选择一个单独的 P_a ,根据轮盘方法选择一个精英策略 P_s :一个染色体的密度越大,它被选中的概率就越大。最后, P_a 的位置更新公式为:

$$P_a \leftarrow P_a + \eta \text{Rand}(0,1) (P_s - P_a) \quad (14)$$

式中: η 是控制参数。通过引入上述竞争性学习策略,个体的染色体可以通过学习性能更好地解进行更新。优化算法如算法 1 所示。

算法 1 INSGA-II

1. 输入相关算法参数
2. 初始化种群, $P(1)$ 代表初始化种群大小
3. **While** IT 小于等于最大迭代次数 **do**
4. 得到适应度值, $P(IT)$ 代表在第 IT 次迭代中的种群数
5. **for** $i=0:1:P(IT)$ **do**
6. 构造训练子集上单个 i 的上下界
7. 根据式(3)和式(7)计算每个 i 的 PINAW 和 PIEE
8. **end for**
9. 得到总种群数
10. 结合 $P(IT)$,得到一个总体种群数 P_u
11. 对比择优
12. **for** $i=1:1:P_u$ **do**
13. **for** $j=i+1:1:P_u$ **do**
14. 比较选择 i 和 j 之间的较优者
15. **end for**
16. **end for**
17. 竞争性学习策略
18. **for** $i=1:1:P_u$ **do**
19. **if** 个体支配时间 $i=0$ **then**
20. 将个体 i 作为一个精英策略
21. **end if**
22. **end for**
23. 快速排序
24. 基于快速排序机制将种群分为不同的层次
25. 根据密度机制得到一个新的种群 P_u
26. 获取子代种群
27. 将选择、交叉、变异和竞争学习应用于 P_u
28. **end while**
29. 输出 P_u 内的最优解

算法的具体步骤如下:

步骤 1 将原始数据集分为训练、验证和测试子集,其中训练集用于调整 LSTM 的参数,验证集用于确定 LSTM 的结构,测试集用于验证模型的泛化能力。为了减轻不良数据对区间预测的负面影响,采用了文献[14]中的 3σ 原则。

步骤 2 总体初始化:对于每个个体 i ,LSTM 中的参数 W 、 U 、 V 、 b_i 和 b_o 通过随机均匀分布进行抽样初始化。

步骤3 适应值:每个 i 都可以看作是 LUBE 模型的解决方案。然后根据式(3)和式(6)计算 PIEE 指数和 PINAW 指数。

步骤3.1 假设 IT 表示当前的迭代时间。设为 1。

步骤3.2 对于每个个体,令 $t=1, E_{PIE}=0, W_{PINA}=0$ 。

步骤3.3 假设 D_{train} 中的样本总数为 n , LSTM 的输入是之前的 m 个观测值。根据 LSTM 的前向传播过程,构造了 y_{t+m} 的上下界。如果 $y_{t+m,u} \leq y_{t+m} \leq y_{t+m,l}$, 令 E_{t+m} 为 0; 否则 $E_{t+m} = \min\left\{\frac{1}{w}(|y_{t+m,u} - y_{t+m}|), \frac{1}{w}(|y_{t+m,l} - y_{t+m}|)\right\}$ 。

步骤3.4 预测区间的宽度的计算公式为 $d_{t+m} = \frac{1}{w}(y_{t+m,u} - y_{t+m,l})$ 。

步骤3.5 $E_{PIE} = E_{PIE} + E_{t+m}, W_{PINA} = W_{PINA} + d_{t+m}$ 。

步骤3.6 如果 $t = n - m$ 进行步骤 4, 否则 $t = t + 1$, 返回步骤 3.3。

步骤4 得到总体种群数:目前的种群是结合精英策略,以获得一个总体种群数 P_u 。(当 $IT=1$ 时,精英策略数目设为 0)。

步骤5 优势检查:检查每个个体之间的优势关系。

步骤6 快速排序:根据优势关系将种群分为不同的层次。然后,采用文献[15]中基于密度得到了一个新的群体 P_n 。

步骤7 获取子代群体:对 P_n 进行选择、交叉、变异和竞争学习,得到一个后代群体 P_o 。

步骤8 迭代: $IT = IT + 1$, 如果 IT 大于最大迭代时间,算法停止。否则,返回步骤 3.2。

步骤9 记录 P_u 的结果。

步骤10 记录 P_o 的结果。

步骤11 记录 P_n 的结果。

5 实验与结果分析

为了验证本文提出的基于 LSTM 的 LUBE 的有效性,本节对两个实际风电数据集案例进行了研究。实验在 Intel i7-7700 3.6 GHz CPU 计算机上进行的,编程语言采用 C++。该算法在人群中使用 60 条染色体,进行 800 次迭代。

5.1 案例 1

5.1.1 数据集和参数设置

案例研究 1 中使用的风力发电数据集来自爱尔兰国家电网,时间段为 2017 年 1 月至 2017 年 12 月,数据集的时间间隔为 15 分钟,即每天采集 96 个风电样本,数据集如图 5 所示。

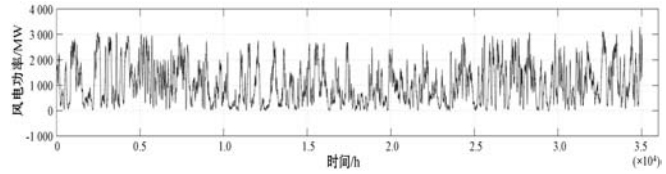


图5 风电样本数据集

根据文献[15]的划分方法,本文将爱尔兰国家电网一年的风电数据分为训练、验证和测试三个子集,其中前 10 个月的数据用于 LSTM 的训练,验证和测试子集分别为 1 个月的数据。由于验证集可以看作是训练集的一部分,因此训练集约占整个数据集的 91.5%。

在估计权重和偏差之前,首先要确定 LSTM 的结构。由于输出包括上限和下限,因此将输出参数的数目设为 2。为了确定输入参数的数目,分析了风电数据集的时间序列特征。首先,图 6 中的样本自相关函数表明风电序列之间存在明显的相关性。在时间序列分析中,如自回归模型,当前观测值 W_t 被构造为前 k 个观测值的线性组合 $(W_{t-1}, W_{t-2}, \dots, W_{t-k})$ 。虽然 LSTM 是一个非线性模型,但自相关分析可以提供过去观测值的相关性来表示当前观测值。通常,用偏自相关函数来确定 k 的值。从图 7 可以看出,当滞后为 11 个观测值时,偏自相关函数近似为零,因此,将 LSTM 的输入设为 W_t 的前 10 个观测值。

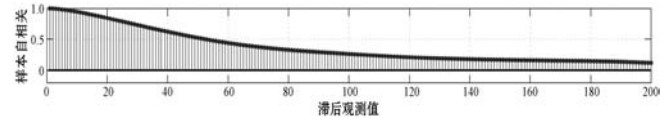


图6 样本自相关函数

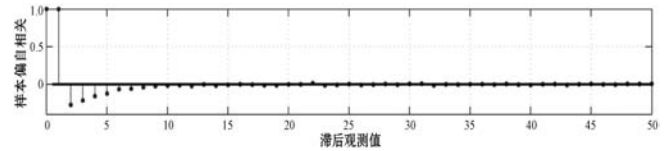


图7 部分样本自相关函数

LSTM 的结构对预测区间的质量有着重要的影响,隐藏节点少会导致表示能力差,而隐藏节点过多会导致泛化能力低下,因此,选择合适的 LSTM 结构是构建高质量预测区间的关键步骤。本文应用文献[17]中的交叉验证方法来确定 LSTM 的结构,为了简化 LSTM 结构,隐藏节点的数量在 1~30 之间。因此,从 30 个候选对象中构造出最佳结构,然后通过训练集和验证集对每个候选对象进行训练和验证。为了评价 Pareto 最优解在验证集上的性能,采用了目前广泛使用的指标——超体积。超体积是 Pareto 最优解较前沿的综合评价指标,衡量了由 Pareto 最优解共同支配的目标空间部分的大小。不同隐藏节点数的 LSTM 的交叉验证结果如图 8 所示,结果表明,具有 14 个隐藏节点的 LSTM 可以获得最高的超体积,因此, LSTM 的结

构设为 10-14-2。

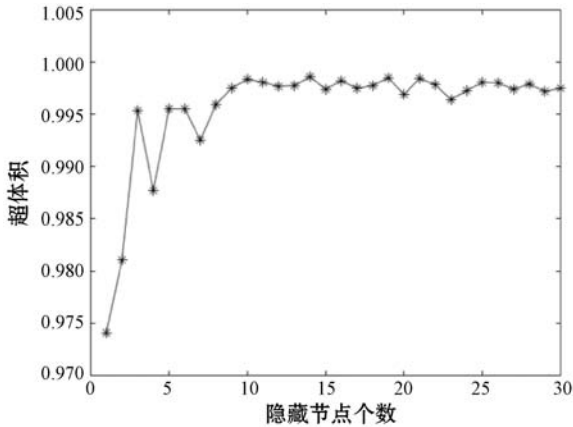


图8 不同隐藏节点数的 LSTM 的交叉验证结果

5.1.2 上下界估计模型的性能

PICP 和 PIEE 从不同的角度对预测区间的精度进行了评价: PICP 表示未来风电值在预测区间内的概率, 而 PIEE 是用来测量预测区间的估计误差。基于 LSTM 的 LUBE 模型在测试集上获得的总体 Pareto 最优解如图 9 所示, 其中每个点表示一对目标, 可以发现 PINAW 和 PIEE 是相互冲突的。为了减小估计误差, 应增加预测区间的宽度, 反之亦然。如果预测区间宽度减小, 则会从预测区间中提出更多的观测值, 此外, 现有的观测结果, 在预测区间之外的观测值与预测的下界和上界之间的差异也将变得更大。因此, PIEE 和 PINAW 本身就存在冲突, 需要进行权衡。

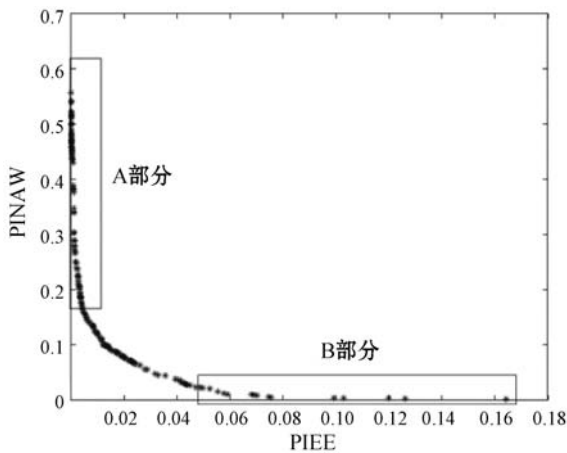


图9 Pareto 最优解

此外, A 部分中的 PIEE 相对较小的, 可以看出, PIEE 的轻微下降会使 PINAW 显著增加。这种现象是合理的, 因为较小的 PIEE 意味着大多数观察值都在 PI 内, 增加预测区间的宽度只会减少少量观测值的估计误差。因此, PINAW 应该大幅度增加, 从而可以改善 PIEE。相比之下, 在 PIEE 相对较大的 B 部分, PINAW 的小幅增加会使 PIEE 得到显著改善。因为 PIEE 意味着大多数观察结果都不在预测区间范围内, 当预测区间的宽度稍大时, 大多数观测值的估计误差将减小,

因此, PIEE 明显减少。

在本案例中, 将 2017 年的最后一个月的数据作为测试集, 因此, 总共使用了 96×31 个测试样本。采用模糊推理方法选择相应的 Pareto 最优解, 其中 PINAW 和 PIEE 分别为 0.009 78 和 0.121。如图 10 所示, 预测的下限和上限几乎能够覆盖所有的观察结果。因此, 本文对整个数据集的划分可能不会导致过拟合问题。

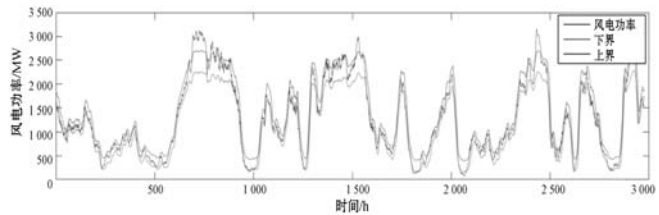


图10 测试子集的上下界

5.1.3 算法性能

为了证明本文算法的有效性, 将 NSGA-II^[5], FS-MOPSO^[8] 作为基准算法。在训练过程中通过不同算法获得的 Pareto 最优解如图 11 所示。结果表明, FS-MOPSO 获得的 Pareto 最优解具有较高的 PINAW 和 PIEE 值, 这意味着 FS-MOPSO 并没有较强的全局搜索能力, 很容易陷入局部最优。为了全面评估不同 Pareto 最优解的多样性和收敛性, 采用了目前广泛使用的指数超体积, 其中参考点设为 (1, 1)。然后得到三种算法的超体积, 分别为 0.994 6、0.992 7 和 0.883 9, 其中 INSGA-II 的值最大。从实验结果可以看出, INSGA-II 在构建预测区间时可以获得收敛速度和多样化的 Pareto 最优解。

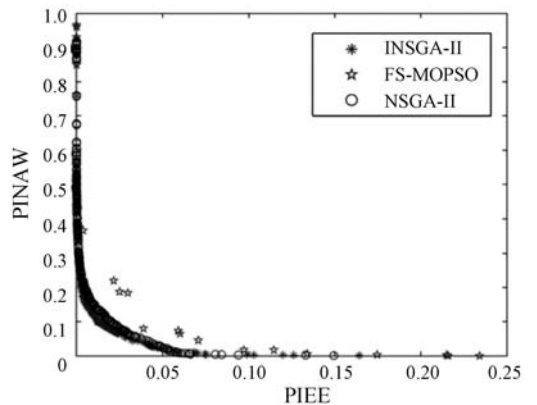


图11 Pareto 最优解

为了训练 LUBE 模型, INSGA-II、NSGA-II 和 FS-MOPSO 算法的运行时间分别为 6 363 s、4 705 s 和 4 489 s。尽管 INSGA-II 的训练时间相对较长, 但其性能最佳。在训练过程中, 进行了数百万次前向传播操作以计算 PIEE 和 PINAW 指数。然而, 模型训练只需要一个前向传播, 就可以构建实现的风速间隔。因此, 一旦模型训练完毕, LUBE 模型就可以在很短的时间

内在实际风速预测应用中构建 PI。此外,由于 LSTM 获得风速的时间序列特征,只要发电的时间序列特征保持不变,LUBE 模型仍然有效,这意味着 LUBE 模型不需要经常重新训练。因此,INSGA-II 可以作为解决 LUBE 问题的有效算法。

5.1.4 与基准预测模型比较

本节例说明提出的方法在处理区间预测任务时的性能。将 BiLSTM、LSTM-NPKD 和 LSTM-VMD 作为对比方法。不同预测模型得到的 Pareto 最优解如图 12 所示。研究发现,提出的方法优于其他预测模型。

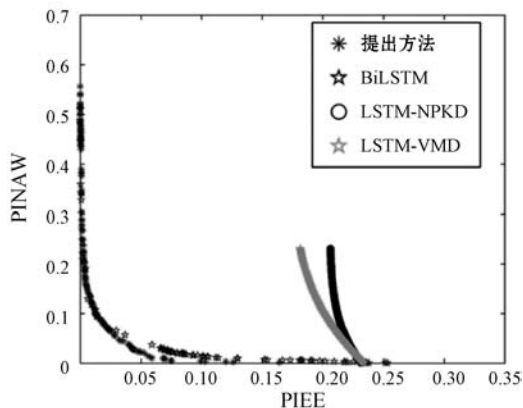


图 12 不同预测模型得到的 Pareto 最优解

当采用 INSGA-II 算法时,本方法、BiLSTM^[10]、LSTM-NPKD^[11]和 LSTM-VMD^[12]的执行时间分别为 3 103 s、2 910 s、1 877 s 和 1 821 s。尽管提出的方法需要更多时间进行训练,但它在这四种模型中表现最佳。如上文所述,LUBE 模型可以在很短的时间内训练构建预测区间。此外,LUBE 模型不需要重新训练。

5.2 案例 2

为了避免存在数据偏差,本文基于开放式电力系统数据的风电数据集进行了另一个案例研究。如图 13 所示,该数据集是 2017 年 1 月至 2017 年 12 月的风电观测数据,收集时间间隔为一小时,整个数据集进一步分为训练、验证和测试子集:前 10 个月的数据用于训练 LSTM,验证和测试子集分别为 1 个月的数据。根据自相关和部分自相关分析,将 LSTM 的输入设为前 27 个已实现风电观测值。然后,基于交叉验证方法,将隐藏层神经元数目设为 18 个。因此,LSTM 的结构设置为 27-18-2。

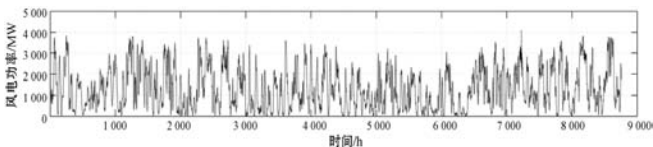


图 13 风电功率数据集

PIEE 与 PINAW 固有的矛盾如图 14 所示。PIEE 可以通过减少 PINAW 和来增加,反之亦然。此外,如

A 部分所示,当 PIEE 相对较小时,PIEE 的轻微减少会使 PINAW 显著增加。B 部分表明,当 PIEE 相对较大时,PINAW 的小幅增加会使 PIEE 得到显著改善。在这个案例研究中,使用 2 431 个样本来测试本文方法的性能。基于模糊推理方法,选取相应的 Pareto 最优解为(0.008 99,0.164 7)。如图 15 所示,预测的上界和下界可以以较高概率覆盖风电观测值。因此,可以得出结论,本文提出的 LUBE 模型适用于不同的风电数据集。

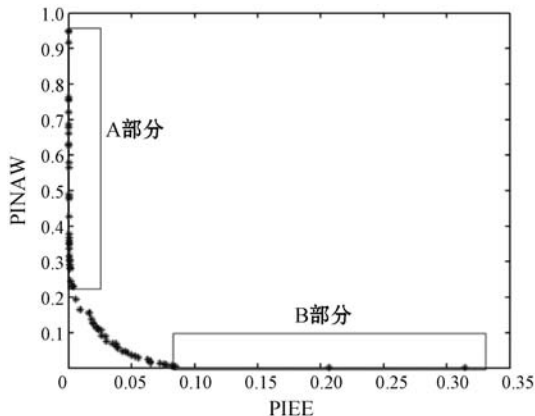


图 14 PIEE 与 PINAW 固有矛盾

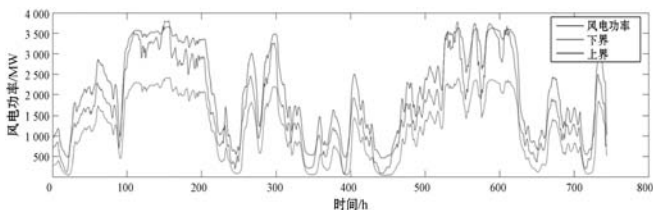


图 15 案例 2 的上下界

6 结 语

为了进一步提升区间预测的效果,提出了一种基于竞争学习机制的 LSTM 风电多目标区间预测方法。最后通过两个实际风电数据集进行实验验证,实验结果表明:预测区间估计误差和预测区间归一化平均宽度之间存在内在的冲突,预测区间估计误差稍有减小,则预测区间归一化平均宽度会显著增大预测区间估计误差较小,当预测区间估计误差较大时,预测区间归一化平均宽度的小增量将使预测区间估计误差显著提高。另外,所提出的求解算法能够提供多样性和收敛性的 Pareto 最优解,并且证明了提出方法能够有效提升预测效果。

参 考 文 献

[1] 姜兆宇,贾庆山,管晓宏. 多时空尺度的风力发电预测方法综述[J]. 自动化学报,2019,45(1):51-71.

- [2] Alawida M, Samsudin A, The J S, et al. A new hybrid digital chaotic system with applications in image encryption[J]. *Signal Processing*, 2019, 160:45 – 58.
- [3] Meysam A C, Shahram J, Narhe N K. A novel keyed parallel hashing scheme based on a new chaotic system[J]. *Chaos Solitons and Fractals*, 2016, 87:216 – 225.
- [4] Shu J L, Xuan Q M, Yuan L C, et al. On the security of a chaotic encryption scheme: Problems with computerized chaos infinite computing precision[J]. *Computer Physics Communications*, 2003, 153(1):52 – 58.
- [5] Boesgaard M, Vesterager M, Pedersen T, et al. Rabbit: A new high-performance stream cipher [C]//International Workshop on Fast Software Encryption, 2003:307 – 329.
- [6] 刘建东. 扩展整数帐篷映射与动态散列函数[J]. *通信学报*, 2010, 31(5):51 – 59.
- [7] 刘嘉辉, 张宏莉. 基于可扩展精度的 Logistic 混沌随机序列的并行计算方法[J]. *中国科学技术大学学报*, 2011, 41(9):837 – 846.
- [8] 刘建东, 张啸, 赵晨, 等. 动态整数帐篷映射模型及其性能分析[J]. *计算机科学*, 2016, 43(11):226 – 229.
- [9] Ouannas A, Khennaoui A, Bendoukha S, et al. The dynamics and control of the fractional forms of some rational chaotic maps[J]. *Journal of Systems Science & Complexity*, 2020, 33(3):584 – 603.
- [10] 黄峰, 冯勇. 二维混沌映射图像加密安全性分析及改进算法[J]. *哈尔滨工业大学学报*, 2007(9):1411 – 1414.
- [11] Sun Y F, Lv Z W. Digital image encryption with chaotic map lattices[J]. *Chinese Physics B*, 2011, 20(4):136 – 142.
- [12] 张勇, 陈滨. Logistic 映射的有限字长研究[J]. *电子科技大学学报*, 2006(3):292 – 294, 316.
- [13] 盛利元, 全俊斌. 计算机迭代下混沌序列的周期研究[J]. *计算机应用*, 2010, 30(7):1802 – 1804, 1808.
- [14] 刘建东. 基于整数耦合帐篷映射的单向 Hash 函数及其性能分析[J]. *计算机研究与发展*, 2008(3):563 – 569.
- [15] 张啸, 刘建东, 商凯, 等. 基于耦合动态整数帐篷映射格子模型的轻量级 Hash 函数[J]. *北京石油化工学院学报*, 2016, 24(1):43 – 48.
- [16] 陈飞, 刘建东, 胡辉辉, 等. 二维整数帐篷映射模型设计及安全性仿真分析[J]. *计算机工程与应用*, 2019, 55(1):103 – 108, 173.
- [17] Nesa N, Ghosh T, Banerjee I. Design of a chaos-based encryption scheme for sensor data using a novel logarithmic chaotic map[J]. *Journal of Information Security and Applications*, 2019, 47:320 – 328.
- [18] Amigó J M, Kocarev L, Szczepanski J. Theory and practice of chaotic cryptography [J]. *Physics Letters A*, 2007, 366(3):211 – 216.
- [19] Rukhin A, Soto J, Nechvatal J, et al. A statistical test suite for random and pseudorandom number generators for cryptographic applications [J]. *Applied Physics Letters*, 2015, 22

(7):1645 – 1679.

~~~~~  
(上接第 311 页)

- [ 2 ] 苏柯文, 张永明. 基于改进鸟群算法优化聚类的风电场等值建模[J]. *计算机应用与软件*, 2021, 38(1):266 – 271.
- [ 3 ] 刘军, 张彬彬, 赵婷. 基于模糊评价的风电场有功功率分配算法[J]. *电工技术学报*, 2019, 34(4):786 – 794.
- [ 4 ] Prajapati V, Mahajan V. Congestion management of power system with uncertain renewable resources and plug in electrical vehicle[J]. *IET Generation Transmission & Distribution*, 2019, 13(6):59 – 71.
- [ 5 ] 张金环, 王超群, 张彤, 等. 基于高斯混合分布模型的风电功率预测误差统计分析研究[J]. *智慧电力*, 2020, 48(7):59 – 64.
- [ 6 ] Dou F, Cheng J, Wang W, et al. Power system reserve scheduling with wind farm integration considering robust security constraints [J]. *Journal of Physics: Conference Series*, 2020, 16(1):102 – 111.
- [ 7 ] Li C, Tang G, Xue X, et al. The short-term interval prediction of wind power using the deep learning model with gradient descend optimization[J]. *Renewable Energy*, 2020, 155(6):96 – 108.
- [ 8 ] 李鉴博, 樊小朝, 史瑞静, 等. 基于互补式集合经验模态分解和 IPSO-LSSVM 的短期风功率预测 [J]. *水力发电*, 2020, 46(11):95 – 100.
- [ 9 ] 孙勇, 李宝聚, 孙志博, 等. 融合 RBF 神经网络和集对分析的风电功率超短期预测 [J]. *昆明理工大学学报(自然科学版)*, 2020, 45(5):49 – 58.
- [ 10 ] 薛阳, 张宁, 俞志程, 等. 基于 BiLSTM 和 Bootstrap 方法的风电功率区间预测 [J]. *可再生能源*, 2020, 38(8):1059 – 1064.
- [ 11 ] 熊鸣. 基于 BP 神经网络与非参数核密度估计的短期风电功率概率区间预测 [J]. *北京信息科技大学学报(自然科学版)*, 2020, 35(4):51 – 56.
- [ 12 ] 任文凤, 冯志亮, 杜艳丽. 1 种改进长短期记忆神经网络的风电功率预测 [J]. *北华大学学报(自然科学版)*, 2020, 21(6):830 – 835.
- [ 13 ] Hu J, Lin Y, Tang J, et al. A new wind power interval prediction approach based on reservoir computing and a quality-driven loss function[J]. *Applied Soft Computing*, 2020, 92(6):106 – 127.
- [ 14 ] Wang R, Deng C, et al. Deep learning method based on gated recurrent unit and variational mode decomposition for short-term wind power interval prediction[J]. *IEEE transactions on neural networks and learning systems*, 2019, 62(6):112 – 123.
- [ 15 ] Li C, Tang G, Xue X, et al. The short-term interval prediction of wind power using the deep learning model with gradient descend optimization[J]. *Renewable Energy*, 2020, 155(2):114 – 131.