

分类号： TP341

密 级： 公开

学校代码： 10363

学 号： 2220340116



安徽工程大学

Anhui Polytechnic University

硕士学位论文

题 目 基于预测控制的永磁同步电机
改进滑模控制研究

论 文 作 者 王胜鹏

指 导 教 师 孙驷洲

学 科（专业） 控制科学与工程

研 究 方 向 控制理论与控制工程

论文提交日期： 2025 年 5 月 10 日

分类号：TP242.6

单位代码：10363

密 级：公 开

学 号：2210310105

题 目：基于核函数选择和参数同步优化的风速短期预测

题 目：Short-term wind speed prediction based on kernel
function selection and parameter synchronization

学生姓名：王胜鹏

导 师：孙驷洲（副教授）

专业学位类别：控制科学与工程

研究方向（领域）：控制理论与控制工程

论文答辩日期：2025 年 05 月 10 日

安徽工程大学学位论文原创性声明

本人郑重声明：我恪守学术道德，崇尚严谨学风。所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已明确注明和引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品及成果的内容。论文为本人亲自撰写，我对所写的内容负责，并完全意识到本声明的法律后果由本人承担。

学位论文作者签名：王胜鹏

日期：2025 年 5 月 10 日

安徽工程大学学位论文版权使用授权书

学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅或借阅。本人授权安徽工程大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

保密 ☐，在 ____ 年解密后适用本版权书。

本学位论文属于

不保密 ☒。

学位论文作者签名：王胜鹏

指导教师签名：孙驹洲

日期：2025 年 5 月 10 日

日期：2025 年 5 月 10 日

基于核函数选择和参数同步优化的风速短期预测

摘 要

随着社会对能源需求的日益增长,以及化石能源消耗和污染问题的加剧,国家开始大力发展可再生能源。风能就是一种可再生清洁能源,是化石能源的替代品之一。随着可再生、分布广且清洁的风能发电得到快速发展,在过去的几十年里,风力发电量迅速增加,并且风电在电网中渗透率不断提高,这有利于环境保护和经济可持续发展。然而,风电具有随机波动性的特点,大规模风电并网后,由于受到外界因素影响,风电的不确定性对电网的安全稳定性产生严重影响。风速是影响风力发电的最直接因素,风速的准确预测是解决该问题的有效方法之一,因此,准确的风速短期预测对电网电能调度和安全稳定运行非常重要。

最小二乘支持向量机(LSSVM)具有鲁棒性强、适应性好以及善于处理小样本的优点,广泛应用于风速预测等时间序列变量的预测。LSSVM 比支持向量机(SVM)和其他基于人工神经网络的模型具有更好的风速预测性能,然而,输入变量、核函数的类型及其核参数影响 LSSVM 的回归性能。因此,许多研究人员努力从这些因素出发,提高基于 LSSVM 的风速预测模型的预测性能。受前人文献的启发,本文构建了一种基于核函数选择和参数同步优化的风速短期预测模型,主要工作和研究内容如下:

(1) 针对风速随机波动性大和非线性强,本文采用“分解-重构”策略,构建了一种基于核函数选择和参数同步优化的新型风速组合预测模型(VPBL)。在所构建的 VPBL 中,首先,通过中心频率观测法(CFO)优化设置变分模态分解(VMD)的分解层数,将原始历史风速样本自适应地分解为多个振幅较小的分解模态,从而降低了 LSSVM 的预测难度。通过相空间重构(PSR)方法将每个分解模态构造为 LSSVM 的相关输入矩阵,进行分解模态预测。在每个分解模态预测中,利用 CEBA 实现了 PSR 和 LSSVM 的核函数的选择和参数的同步优化。各分解模态的预测值用于重构最终风速预测结果;

(2) LSSVM 的风速预测性能依靠核函数种类及其参数,而单一核函数种类无法适应不同分解模态的特性。针对该问题,本文构建了新型核函数选择策略,为 LSSVM 在每个分解模态预测时,利用改进蝙蝠算法从典型核函数集中选择最优

核函数类型并优化其参数，集成不同核函数各自的优势，以提高 LSSVM 风速预测模型的整体预测性能；

(3) 传统遗传算法(GA)、粒子群算法(PSO)和和谐搜索算法(HS)能够很好地处理复杂问题，然而，这些算法对初始种群和初始化参数敏感，且缺乏鲁棒性。与传统的 PSO、GA 和 HS 相比，蝙蝠算法(BA)既能在搜索最优解的过程中采用频率调谐方法增强算法种群的多样性，又能采用脉冲响度和频率自适应调整的策略，实现在全局优化和局部优化之间动态调整搜索强度，提高了 BA 的优化性能。因此，本文使用 BA 来优化新建风速预测模型的参数。但是 BA 算法与其优化算法一样，也存在容易陷入局部最优的不足。针对上述的不足，本文对传统 BA 算法进行改进，获得增强型蝙蝠算法，即 CEBA。利用混沌函数初始化算法初始种群；引入动态扰动系数和正切随机扰动机制对种群个体进行扰动，使种群个体进入更广阔的可求解空间进行寻优；将平均总体位置调整机制引入 BA 算法中，利用种群个体位置总体平均值引导蝙蝠个体进入其他搜索空间进行寻优，降低局部最优的可能性；

(4) 最后，利用内蒙古某风电场实测风速数据进行测试和评估新建的风速预测模型 VPBL 的有效性，实验结果表明：与其典型的风速预测模型和近几年开发的风速预测模型相比，VPBL 的风速预测精度都有显著提高。

关键词：变分模态分解、混沌增强蝙蝠算法、最小二乘支持向量机、核函数选择、风速预测。

Short-term wind speed prediction based on kernel function selection and parameter synchronization optimization

ABSTRACT

With the increasing demand for energy and the intensification of fossil energy consumption and pollution problems, the country has begun to vigorously develop renewable energy. Wind energy is a renewable and clean energy and one of the alternatives to fossil energy. With the rapid development of renewable, widely distributed and clean wind power generation, wind power generation has increased rapidly in the past few decades, which is beneficial to environmental protection and economic sustainable development, and the penetration rate of wind power in the power grid is constantly increasing. However, wind power has the characteristics of random fluctuation. After large-scale wind power is connected to the grid, the uncertainty of wind power due to external factors has a serious impact on the security and stability of the grid. The higher wind speed prediction accuracy is one of the effective methods to solve this problem. Wind speed is the most direct factor affecting wind power generation. Therefore, accurate short-term wind speed prediction is very important for power dispatching and safe and stable operation of power grid.

Least squares support vector machine (LSSVM) has the advantages of strong robustness, good adaptability and good handling of small samples. It is widely used in the prediction of time series variables such as wind speed prediction. LSSVM has better wind speed prediction performance than support vector machine (SVM) and other ANN-based models, however, the input variables, type of kernel function and its kernel parameters have a great influence on the regression performance of LSSVM. Therefore, many researchers have made efforts to improve the prediction performance of LSSVM-based models based on these factors. Inspired by previous literatures, this paper develops a short-term wind speed prediction method based on kernel function selection and parameter synchronization optimization. The main work and research contents are as follows.

(1) Aiming at the large random fluctuation and strong nonlinearity of wind speed, this paper adopts the "decomposition-reconstruction" strategy to develop a new combined

wind speed prediction model (VPBL) based on kernel function selection and parameter synchronization optimization. In the newly built VPBL, firstly, the number of decomposition layers of variational mode decomposition (VMD) is optimized by the center frequency observation (CFO) method, and the original historical wind speed samples are adaptively decomposed into several decomposition modes with small amplitudes, thus reducing the prediction difficulty of LSSVM. Each decomposition mode is constructed as a correlation input matrix of LSSVM by phase space reconstruction (PSR) method, and the decomposition mode prediction is performed. In each decomposition mode prediction, the selection of kernel function and the simultaneous optimization of parameters are realized by CEBA. The predicted values of each decomposition mode are used to reconstruct the final wind speed prediction results.

(2) The wind speed prediction performance of LSSVM depends on the kernel function type and its parameters, and a single kernel function type cannot adapt to the characteristics of different decomposition modes; Aiming at this problem, this paper constructs a new kernel function selection strategy, which uses the improved bat algorithm to select the optimal kernel function type from the typical kernel function set and optimize its parameters for each decomposition mode prediction of LSSVM, and integrates the advantages of different kernel functions to improve the overall prediction performance of LSSVM wind speed prediction model.

(3) Traditional genetic algorithm (GA), particle swarm optimization (PSO) and harmony search algorithm (HS) can handle complex problems well, however, these algorithms are sensitive to initial population and initialization parameters, and lack robustness. Compared with the traditional PSO, GA and HS, BAT algorithm (BA) can not only enhance the diversity of algorithm population by using frequency tuning method in the process of searching for the optimal solution, but also adopt the strategy of adaptive adjustment of pulse loudness and frequency to dynamically adjust the search intensity between global optimization and local optimization, which improves the optimization performance of BA. Therefore, this paper uses BA to optimize the parameters of the newly built wind speed prediction model. However, BA algorithm, like other optimization algorithms, also has the shortcomings of easily falling into local optimum. In view of the above shortcomings, this paper improves the traditional BA algorithm and obtains the enhanced bat algorithm, namely CEBA. The initial

population of the algorithm is initialized by using chaotic functions; dynamic disturbance coefficients and tangent random disturbance mechanisms are introduced to perturb the individuals in the population, enabling them to enter a broader solvable space for optimization; the average overall position adjustment mechanism is incorporated into the BA algorithm, guiding the bat individuals to enter other search spaces for optimization by using the overall average position of the population individuals, reducing the possibility of local optimum.

(4) Finally, the effectiveness of the newly develop wind speed prediction model VPBL is tested and evaluated by using the measured wind speed data of a wind farm in Inner Mongolia Province. The experimental results show that compared with other typical wind speed prediction models and wind speed prediction models that can be develop in the last few years, the wind speed prediction accuracy of VPBL is significantly improved.

KEYWORDS: Variational mode decomposition, chaotic enhanced bat algorithm, least squares support vector machine, kernel function selection, wind speed prediction.

目 录

摘 要	I
ABSTRACT	III
第 1 章 绪论	1
1.1 研究背景及意义	1
1.2 国内外研究现状	3
1.2.1 风速预测研究现状	3
1.2.2 研究现状分析	8
1.3 研究内容	8
第 2 章 基于最优变分模态分解和相空间重构的风速预处理	10
2.1 风速时间序列特性	10
2.2 最优变分模态分解	11
2.2.1 变分模态分解	11
2.2.2 最优变分模态分解	13
2.3 相空间重构	13
2.4 本章小结	15
第 3 章 基于最优核函数选择的 LSSVM 模型	16
3.1 机器学习	16
3.1.1 机器学习的基本结构	16
3.1.2 经验风险最小化	17
3.2 最小二乘支持向量机	18
3.2.1 支持向量机	18
3.2.2 最小二乘支持向量机	19
3.2.3 核函数选择	20
3.3 基于 LSSVM 的风速预测	21
3.4 本章小结	22
第 4 章 蝙蝠算法的改进	24
4.1 蝙蝠算法	24
4.1.1 标准蝙蝠算法	24

4.1.2 标准蝙蝠算法的不足.....	25
4.2 蝙蝠算法的改进策略.....	26
4.2.1 基于混沌方法增强 BA 种群多样性	26
4.2.2 提高 BA 搜索能力的策略	28
4.3 CEBA 优化性能的验证	30
4.4 本章小结.....	34
第 5 章 基于 OVMD-PSR-CEBA-LSSVM 的风速预测	35
5.1 构建 OVMD-PSR-CEBA-LSSVM 模型	35
5.2 预测实验结果及对比分析.....	36
5.2.1 数据来源.....	36
5.2.2 评价指标.....	38
5.2.3 风速预处理结果.....	38
5.2.4 实验结果对比.....	41
5.2.5 假设检验.....	50
5.3 本章小结.....	51
第 6 章 结论与展望.....	53
6.1 研究结论.....	53
6.2 工作展望.....	53
参考文献.....	55
攻读硕士期间取得的学术成果.....	59
致谢.....	60

第 1 章 绪论

1.1 研究背景及意义

现阶段我国经济发展的目标已经从高速增长转变为高质量增长^[1]，党的二十大报告中指出：“推动经济社会发展绿色化、低碳化是实现高质量发展的关键环节。加快推动产业结构、能源结构、交通运输结构等调整优化。”经济的高速发展和高质量发展离不开能源，我国对能源需求大幅增长^[2]。根据国家统计局最新发布的《中国统计年鉴》，我国的能源消费总量由 1978 年的 57144（万吨标准煤）增长到 2024 年的 597000（万吨标准煤），46 年之间的平均递增速率达到 19.00%，且相较于 2021 年度的 498314（万吨标准煤）的消耗，同比增长 19.8%。根据 2024 年最新的数据显示，我国近二十年的能源消费总量如图 1-1 所示，我国能源消耗总量逐年上涨，能源需求越来越大。在我国能源结构中，煤炭石油天然气占据了绝对的地位。化石能源开采会对地貌环境造成严重破坏，同时影响到地表和水域生态系统。煤炭燃烧不充分会造成空气污染，产生二氧化碳加剧温室效应，引起气候变化导致海平面上升等一系列问题。此外，化石能源属于不可再生资源，并且储量有限，随着对其需求的不断增加，化石能源枯竭的局面终将会到来。因此，采用绿色可再生的能源来代替化石能源势在必行^[3]。

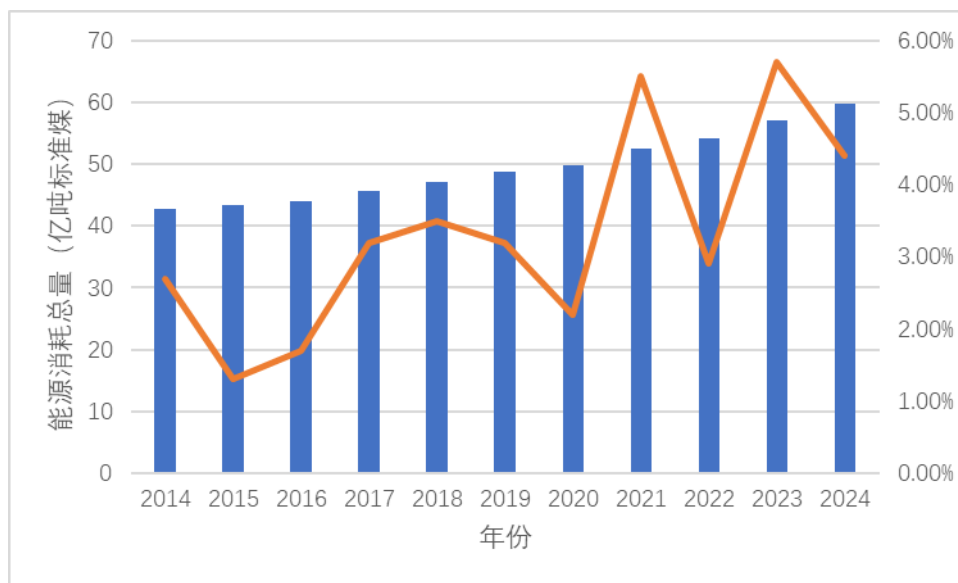


图 1-1 2000 年-2024 年中国能源消耗总量

可再生能源因其可持续性和环保特性，已成为当前国内外专家学者研究能源

问题的核心焦点，他们致力于推动这类能源的开发与利用。风能是一种重要且可再生、清洁的能源形式，在实现能源结构转型和达成碳中和目标方面扮演着至关重要的角色^[4]。据预测，至 2030 年，中国风电发电量将占据全国总发电量的 20% 以上。图 1-2 直观展示了 2014 年至 2024 年间中国风电装机总量的显著增长趋势：国家能源局发布的《2024 年全国电力工业统计数据》显示，截至 2024 年底，全国风电累计装机达到 52068 万千瓦，标志着中国风电行业已迈入一个快速发展的全新阶段^[5]。

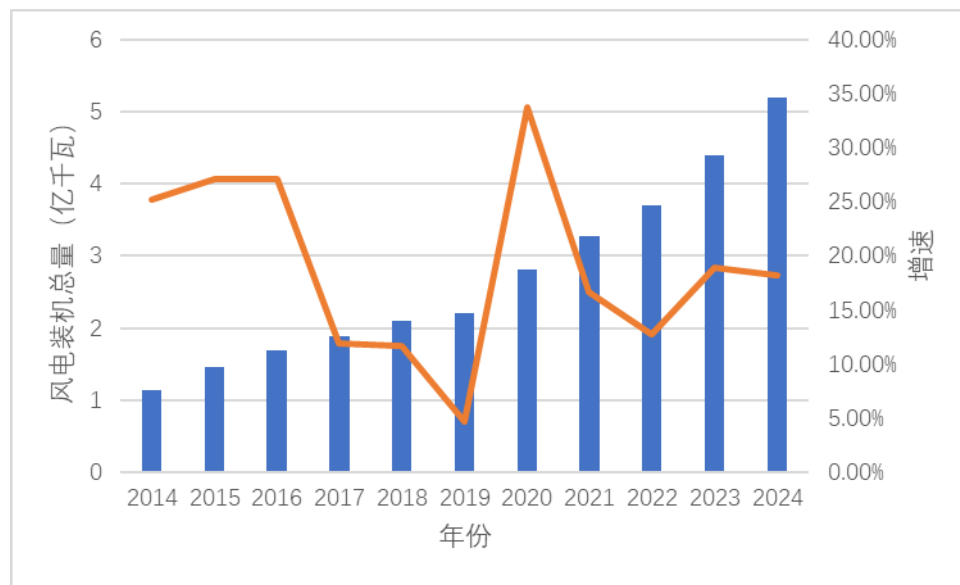


图 1-2 2014 年-2024 年中国风电装机总量

风速随季节和天气条件的变化而波动，通常在春季和秋季时较高，夏季则相对较低^[6]。此外，诸如风暴、气旋等气象事件会显著影响瞬时风速，展现出风速的随机性和间歇性特征。这些特性在大规模风电并网后，给电力系统的电能质量和安全稳定运行带来了挑战。通过风速预测，可以进一步推导出风电功率的预测结果，并分析风电并网时的各项关键参数。这一过程有助于电网运营者及时采取调峰调频措施。一旦预测到风速的变化，电网便能灵活调整其他发电源的输出功率，并相应地调节电网负荷，以应对风电的波动性，从而减少电力系统的不稳定性，避免电压和频率的大幅波动，确保电网能够平稳运行。

风速预测在风电厂的运营管理中扮演着至关重要的角色，不仅能够帮助风电厂科学合理地规划发电量^[7]，还能指导发电计划的精细安排，从而有效提升发电效率并显著降低能源成本，进而最大化发电效率和经济收益。具体而言，当风速预测显示风速即将增大时，风力发电厂可以迅速响应，增加发电量，充分利用更

为强劲的风能资源，实现高效发电。相反，如果预测到风速将减弱，风电厂则能够适时调整策略，减少发电量，避免在风速不足的情况下有不必要的资源损耗和设备磨损。高精度的风速预测方法对于电网调度、风电场的日常运行以及整个电力系统的稳定性具有直接且深远的影响。精确的风速预测能够确保电网调度中心提前做出合理安排，平衡供需，有效避免电力过剩或短缺的情况，维护电网的安全稳定运行。同时，风电场也能根据预测结果优化设备运行策略，减少因风速波动带来的发电不稳定问题，进一步提升整体运营的经济性和可靠性。因此，不断提升风速预测技术的精度和可靠性，是推动风电行业持续健康发展的关键所在。

利用风向变化趋势、风速大小及具体方向等详尽参数，能够更为精确地预估风机的输出功率，这一预测能力为电力调度中心提供了重要的决策依据，使其能够依据这些预测数值对电能调度策略进行合理且及时的调整。精确的风速预测成果能够有效削弱风电因自身不确定性而对电网运行产生的负面影响，显著提升风电的并网效率，并进一步削减风电场的运营成本。鉴于此，风速预测技术已然成为全球范围内科研工作者们广泛关注的焦点。研究和开发出具备高度准确性的风速预测模型，对于推动风电产业的稳健发展具有不可估量的重要意义，不仅能够增强风电的竞争力，还能促进能源结构的绿色转型。

1.2 国内外研究现状

1.2.1 风速预测研究现状

依据风速预测方法的核心运作机制，可以将这些方法划分为物理方法、统计分析法、人工智能法以及组合预测模型等几个主要类别。短期风速预测致力于提供涵盖未来几分钟、几小时乃至几天内的风速预估信息^[8]。为了满足电力系统多样化的运行需求，风速预测按照时间尺度可以划分为^[9]：

（1）超短期预测：主要应用于风电机组的即时控制以及电力系统的负荷跟踪，确保风电机组能够根据风速变化即时调整运行状态^[10]，优化发电效率。同时也可以用来评估风电场输出的电能质量，发现潜在的问题，为电力系统的稳定运行提供保障。

（2）短期预测：提前数小时的电力预测对电网调度部门而言至关重要，有助于调整调度计划以实现经济的电网运行，同时也能有效优化系统的备用容量配置

[11]。系统的备用功率需求往往取决于预测误差的最大值。当运行人员掌握了次日每小时精确的预测数据时，能够更科学地调度电厂和平衡功率，优化次日的机组组合策略，从而显著降低机组组合的成本。

(3) 中长期预测：主要用于电力系统的宏观管理以及能源市场的交易策略制定，风能作为能源市场的一种新兴能源，本身因为风的间歇性和随机性而有较大的波动性^[12]，所以做好中长期预测可以为电力市场的参与者提供决策支持^[13]，有助于对风力资源进行长期评估，为风电场的规划和建设提供科学依据。

通过上述分类，风速预测技术能够全面覆盖电力系统的不同运营需求，从即时控制到长期规划，为风电行业的可持续发展提供坚实的技术支撑。

在风速预测与风电场发电量预估这一研究领域内，科研人员持续致力于探索高效且可靠的工具与技术，并在此过程中构建了众多不同的预测方法^[14]。这些预测方法和技术可以按照是否纳入特定位置的地形信息作为输入参数来进行分类。据此，主流的方法可被归结为以下三种类型：

(1) 物理预测方法

采用物理方法进行风速预测时，主要依赖数据进行解析，同时融入风电场所处区域的地形、地貌等环境特征以及天气状况、温度等气象因素进行综合考量。这种方法的一个显著特点是，在预测过程中无需依赖风电场过往的风速历史数据^[15]。然而，物理预测方法对输入变量的精确度有着极高的要求，特别是地形和地貌等关键参数的准确性。此外，物理预测方法更适用于风电场的长期风速预测，能够为电力调度决策、电力系统建模以及风电场的扩容规划提供科学依据。

随着风电场风机安装容量的持续增长，风电在电网中的占比将日益增大^[16]。然而，风力发电的穿透功率需保持在最大安全值以内，一旦超限，风电机组并网将对电网安全稳定运行构成严重威胁。为解决此问题，科学预测风电场输出功率变得至关重要，这需依据风力、风速及风向的动态变化趋势来进行，以增强风力发电的可控性。当前，数值天气预报系统^[17](NWP)融合了物理与统计方法，但其预测精度受限于 NWP 本身的能力。然而，随着风力发电机组容量的持续攀升，对风力发电预测精准度的需求也随之提升到了新的高度。传统的发电量计算数学模型难以有效捕捉风速的急剧波动，导致计算得出的预测值与实际观测值之间存在显著的偏差。文献[18] 构建了一种集合变分模态分解算法与深度自适应滤波的

自适应信号处理框架,对大规模风电集群进行功率预测。有利于提升电网对于风电的消纳能力,保障风能安全、稳定、高效地发电。文献[19]结合风速-功率曲线对风电场实测数据中的异常值进行分类梳理,提出了基于 CNN-LSTM-Attention 的改进 LSTM 算法对 NWP 风速进行修正,提高预测精度的有效性。文献[20]提出一种非参数核密度估计和 NWP 相结合的方法,并对预测风速误差进行校正,建立基于风速误差校正和 ALO-LSSVM 组合的风电功率预测模型。

(2) 统计预测方法

风速预测的统计方法是一种基于大量历史风速数据和其他相关气象信息,运用数学统计原理和机器学习算法来构建预测模型的技术手段^[21]。这种方法通过深入分析历史数据中的规律和模式,尝试捕捉风速变化的随机性、周期性和趋势性特征,从而为风电场的规划、运营以及电力系统的调度提供科学、准确的预测结果。在风速预测的统计方法中,历史数据扮演着至关重要的角色。这些数据不仅包含了风速本身的时间序列信息,还可能涉及温度、湿度、气压等多种气象因素。通过对这些数据的综合分析和处理,统计预测方法能够建立起一个能够反映风速变化规律的数学模型。这个模型可以根据当前的气象条件和历史数据,对未来一段时间内的风速进行预测。尽管风速预测的统计方法具有广泛的应用前景,但也面临着一些挑战。一方面,预测结果的准确性受到输入数据质量和数量的影响。如果历史数据存在缺失、异常或噪声等问题,可能会导致预测模型出现偏差或不稳定。另一方面,极端天气事件(如风暴、龙卷风等)的风速变化往往具有突发性和不可预测性,这给统计预测方法带来了额外的难度。

然而,尽管存在这些挑战,风速预测的统计方法仍然在风电行业发挥着重要作用。通过不断优化预测模型、提高数据质量和数量、引入先进的机器学习算法等手段,可以进一步提高风速预测的准确性和可靠性,为风电场的规划、运营以及电力系统的调度提供更加科学、准确的决策支持。

(3) 人工智能方法

人工智能技术在风速预测领域得到了广泛应用,其核心在于利用先进的算法建立精确的风速预测模型,以捕捉输入变量与输出变量间复杂的非线性关系。这些方法主要包括神经网络^[22]、小波分析^[23]和支持向量机(SVM)^[24]等。其中,人工神经网络作为一种统计方法,在风速和功率预测中尤为常见。由输入层、一个

或多个隐藏层以及输出层构成，各层间通过神经元连接，能够利用训练和学习过程对输入与输出间的非线性关系进行建模。这一过程无需依赖明确的数学表达式，展现了自学习、自组织和自适应的能力。

而 SVM 则是在统计学习理论上发展的一种通用学习方法，将样本空间中的非线性可分问题转化为 Hilbert 空间中的线性问题。SVM 的性能高度依赖于输入数据及参数选择，其优势在于利用“核技巧”处理线性到非线性的转换，并在简化样本处理方面展现出独特效能。此外，最小二乘向量机(LSSVM)在 SVM 基础上，通过约束方程替代不等式约束，降低了数据处理难度并加速了机器学习过程，因此在处理非线性问题时表现出更高的计算效率^[25]。有研究表明，结合 NWP 和非参数核密度估计的方法，能有效矫正 LSSVM 在风速预测中存在的相位滞后问题，进一步提升了预测的准确性。

单一风速预测模型往往受限于其建模能力，难以达到较高的预测精度。为了克服这一局限性，研究者们提出了构建风速预测组合模型的方法^[26]，该方法结合了风速预处理技术、参数优化算法与人工智能算法，旨在融合不同预测手段的优势，从而显著提升风速预测的准确度。具体而言，文献[27]将人工神经网络(ANN)与 NWP 预报及其他气象变量相结合，提出了两种混合模型用于风速预测；文献[28]利用长短期记忆(LSTM)、卷积神经网络(CNN)等不同类型的神经网络，构建了基于机器学习的风场预测模型。而文献[29]则对含有注意力机制的双向长短期记忆网络(BILSTM)进行了改进，利用贝叶斯优化对模型超参数进行优化，提高预测精度。一般而言，组合模型因其能够综合多种预测方法的优点，往往展现出比单一模型更优的预测性能和更强的泛化能力，能够实现更为全面且精确的建模。风速组合预测模型主要分为：

(1) 基于权重分配的组合模型

通过赋予两种或更多种单一预测模型以特定的权重，并将它们加权组合起来，可以构建一个集合了不同预测模型特点的组合模型^[30]。这种组合预测模型在学术研究中备受关注，其建模的核心挑战在于精确地确定每个模型的加权系数，以确保组合后的模型在稳定性和预测能力上达到最优状态。

(2) 基于数据预处理的组合模型

在风速预测组合模型中，将原始风速分解为多个子序列以降低其非平稳性，

能有效地提高风速预测精度。文献[31]提出一种基于两层分解和 PCA-BP-RBF 神经网络的混合模型,两层分解能更好地解决模态混叠和端点效应等问题,采用 PCA-BP 方法对模型输入数据进行过滤,消除了冗余和不相关的信息,降低了 RBF 模型的复杂度,提高了预测模型的性能。文献[32]将分解后的风速子序列分别输入到 LSTM 模型中进行风速多步预测,最后将各模型输出结果进行叠加获得预测风速。文献[33]通过 VMD 将风速序列分解为一系列相对独立且平稳的子序列来提高预测精度,缺点是 VMD 计算量大,处理长时间信号时分解结果可能不稳定,且正则化系数、分解层数和惩罚因子的选择对分解结果有较大影响。文献[34]在混合模型中使用小波变换技术对风速数据进行预处理,并使用 GA 算法对支持向量机参数进行调整,研究表明,该模型的预测精度优于其他模型。文献[35]通过小波包分解(WPD)处理原始风速后,在相空间中将各分量重构为具有稳定性的风速子序列,合理剔除异常噪声数据,有利于提高训练数据平稳性,凸显数据的细节特征。

(3) 基于参数优选的组合模型

优化组合模型的参数能有效调控其复杂度,从而提升模型对数据的拟合度、泛化能力和准确性。通过人工智能算法对参数进行优化,能够显著增强组合模型的稳定性,使其对噪声和干扰具备更强的抵御能力,进而提高其鲁棒性和稳定性。文献[36]提出了一种基于随机布谷鸟搜索算法(RCSA)优化的人工神经网络(ANN)模型用于短期风速预测,该算法通过引入随机因子提升了性能,相较于传统 ANN 模型,RCSA-ANN 模型的预测误差显著降低。文献[37]则利用鲸鱼优化算法(WOA)的高收敛速度和精度,优化了 LSSVM 模型的惩罚因子和核函数宽度,从而提高了其计算速度和预测精度。文献[38]采用改进蚁群优化算法(IACO)对 LSSVM 回归模型的关键参数进行优化,结果显示 IACO-LSSVM 模型具备更高的预测精度,验证了该方法的有效性和可行性。文献[39]则融合 VMD、相空间重构、改进的北方苍鹰优化算法(INGO)和共享权重门控记忆网络(SWGMN),构建了一种超短期风速混合预测模型,该模型不仅预测精度高,而且鲁棒性强。文献[40]则提出了一种基于混沌初始化、高斯分布局部吸引点、早熟判断和扰动算子的改进量子粒子群优化算法(QPSO),对组合模型的参数进行同步优化,显著提升了混合模型的预报性能。此外,改进蝙蝠算法^[41]、混合 Copula 优化算法^[42]、Jaya 算法^[43]

等也是常用于优化组合模型参数的人工智能算法。

1.2.2 研究现状分析

根据现有研究,国内外在风电场风速预测领域已进行了大量探索,尽管提出的预测模型在一定程度上提升了预测效果,但仍存在诸多不足,风速预测的准确率仍有很大的提升空间。首先,风速信号中携带的噪声增加了风速的不确定性,这些噪声可能会改变风速信号的频谱特性,使得预测模型难以精确捕捉信号特征,进而降低预测的准确性和可靠性。其次,单一预测模型存在固有的局限性,而多种模型组合形成的新模型,其预测能力在很大程度上依赖于风速预处理过程中特征提取的准确性和完整性。最后,风速预测模型的优化算法存在早熟收敛,容易陷入局部最优的问题。

综上分析,高准确度的风速数据缺失处理、降噪能力强的预处理方法、精确稳定的预测模型,对风速预测和风电并网有着重要意义。因此,构建了一种基于最优变分模态分解、相空间重构、混沌增强蝙蝠算法和 LSSVM 的新型混合模型。

1.3 研究内容

为了提高风速预测的准确性,首先使用最优变分模态分解对原始风速数据进行预处理;其次,使用风速数据训练 LSSVM,提高 LSSVM 的预测能力;最后使用改进的蝙蝠算法对 PSR 和 LSSVM 选择合适的核函数以及优化参数,提高其对特定子序列的预测能力。并利用内蒙古某风电场的实测历史风速数据对本文构建风速预测模型进行了测试和评估。本文章节安排如下:

第 1 章:绪论。首先,阐述了风速预测的研究背景与研究意义;其次,研究短期风速预测的国内外研究现状;最后,阐述了本文的研究内容和论文章节内容。

第 2 章:基于最优变分模态分解和相空间重构的风速预处理。采用了 OVMD 方法来分解原始风速数据,获得不同频率和幅值更小的分解模态序列,以降低 LSSVM 风速预测难度。VMD 的分解结果主要受信号分解层数 K 等影响,因此,采用中心频率法,计算了各个分解子序列的中心频率,并据此确定最优的分解层数 K 。在此基础上,通过相空间重构技术确定 LSSVM 的输入变量矩阵。

第 3 章:基于最优核函数选择的 LSSVM 模型。首先对 SVM 和 LSSVM 的基本理论进行了阐述。利用改进蝙蝠优化算法,以综合评价指标(CEI)值为适应

度函数从典型核函数集中选取每个分解子序列的最优化核函数,典型核函数集包括径向基函数 (RBF)、指数径向基函数 (ERBF)、多项式核函数 (Poly)、双曲正切核函数(Sigmoid)、高斯核函数 (Gaussian)、Morlet 小波基函数(Morlet)、墨西哥帽小波(Mexihat)、线性核函数(Linear)。采用滚动预测策略进行风速预测。通过选取各个分解子序列的 LSSVM 核函数,可以有效地提高整体预测方法的预测精度。

第 4 章: 蝙蝠算法的改进。为了增强 BA 算法全局搜索能力,降低 BA 算法陷入局部最优解的可能性,对传统 BA 算法进行改进。首先,混沌初始化 BA 的种群;其次,引入动态扰动系数和正切随机扰动机制对种群个体进行扰动,使种群个体进入更广阔的可求解空间进行寻优;最后,将平均总体位置调整机制引入 BA 算法中,利用种群个体位置总体平均值引导蝙蝠个体。为了验证改进后的 BA 的有效性,采用标准的 CEC15 测试集进行验证改进的 CEBA。

第 5 章: 基于核函数选择和参数同步优化的风速短期预测。基于“分解-重构”的策略,构建了基于 OVMD、PSR、CEBA 和 LSSVM 的新型风速短期预测模型,利用 CEBA 算法对 LSSVM 进行核函数选择,以及对 PSR 和 LSSVM 中的参数同步优化。通过某风电场的历史风速数据进行仿真实验,验证结果明: 并与其他组合预测模型对比,新建的风速预测模型 VPBL 预测精度和预测稳定性更高。

第 6 章: 总结与展望。总结全文的研究工作,分析创新点并指出新建风速预测组合模型还需要改进部分,并展望今后研究的方向。

第 2 章 基于最优变分模态分解和相空间重构的风速预处理

风速预测对于风力发电行业至关重要，风力发电机通过风速预测来改变自身的运行状态，确保在风速较高时最大化电力输出，从而提高能源利用效率。同时，预测结果还有助于电力系统更合理地调度其他能源，确保电力供应的稳定性和可靠性，进而降低运营成本。此外，风速预测还能增强电网的稳定性，通过预测风力发电的波动性，电网运营商可以更好地管理电网负荷，避免过载和电压波动等问题，确保电网的安全运行。风电场采集的气象数据会存在较大的波动性和非线性，从而影响到风速预测的精度和可靠性。因此，需要提高风速预测模型的精准度，分析风电场风速信号的特性，首先解释最优变分模态分解工作原理，利用中心频率观测法人工设置 VMD 的分解数，将原始风速样本自适应地分解为几个振幅较小的子序列。并利用相空间重构法对分解的风速子序列进行重构，建立预测模型输入变量矩阵，增加预测准确性。

2.1 风速时间序列特性

风速时间序列特性复杂多变，包括趋势性、季节性、非平稳性、随机性以及多种环境因素（如温度、气压、风向）的关联性。这些特性共同作用于风速值，使得风速预测成为一项具有挑战性的任务，但同时也为风力发电等应用提供了关键的数据基础和预测依据。通过时间序列分析等方法，可以揭示风速时间序列中的潜在规律和趋势，从而提高风速预测的准确性和可靠性。通过收集内蒙古某一风电场四个季节每隔一小时的风速数据用于风速时间序列特性分析。图 2-1 为四个季节风速曲线图。

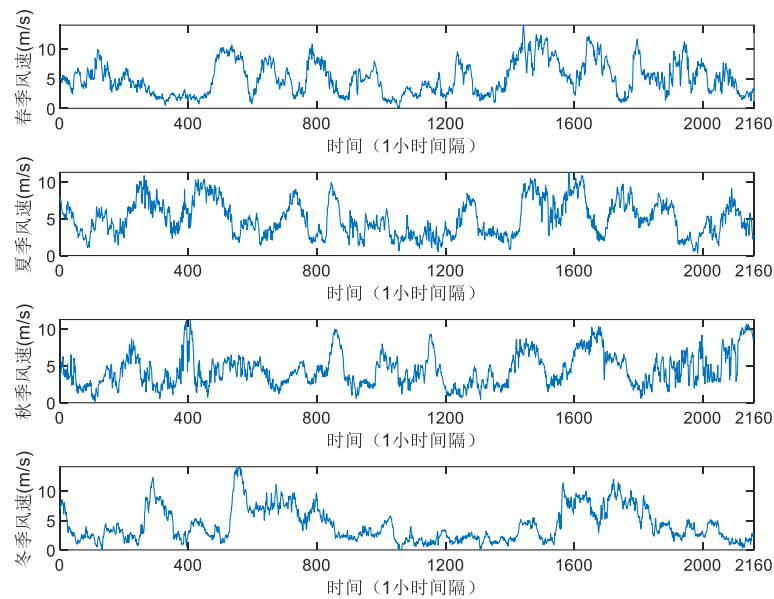


图 2-1 四季风速曲线图

表 2-1 风速时间序列特性

风速时间序列	最小值	均值	最大值	方差
春季	0.13	3.99	14.11	6.72
夏季	0.25	5.45	13.20	6.83
秋季	0.31	5.96	12.04	6.44
冬季	1.77	6.56	12.35	4.19

根据风速时间序列数据分析，风速呈现明显的季节性变化。（1）平均风速，冬季风速最高 6.56 (m/s)，其次是秋季 5.96 (m/s)、夏季 5.45 (m/s)和春季 3.99 (m/s)，表明冬季和秋季风力较强，而春季风力较弱。（2）风速方差，夏季 $6.83(\text{m/s})^2$ 和春季 $6.72(\text{m/s})^2$ 的波动性较大，而冬季 $4.19(\text{m/s})^2$ 风速变化最为平稳，说明春夏季节风速不稳定，冬季风速较为稳定。（3）风速极值，春季的最大风速达到全年最高 14.11 (m/s)，表明春季可能出现极端强风天气，而秋季和冬季的极值相对较低。这些特征帮助模型更好地捕捉风速的变化规律和趋势，从而提高预测的准确性。同时为风能资源利用和气象灾害防范提供了重要参考。

2.2 最优变分模态分解

2.2.1 变分模态分解

VMD 是一种自适应信号处理方法，利用迭代方法，通过计算变分模态的最优解，将时间序列信号分解为几个本征模态函数(IMF)^[44]。VMD 利用变分原理

分解时间序列，能够更好的处理一些混沌性强和非平稳的序列，以此来减少模态混叠问题。在分解过程中，VMD 自适应地捕获每个模态的最优中心频率的有限带宽，以产生给定信号的有效分解子序列，该子序列是在每个模态的乘积等于原始信号的约束下进行的。其分解过程是通过优化一个变分问题来实现，以此找到最佳的模态分解，让分解后的模态尽量相互独立。VMD 分解的具体步骤如下：

(1) 通过希伯特变换(Hilbert Transform, HT)计算每个模态函数 $\mu_k(t)$ 解析信号，把中心频率 $\omega_k(t)$ 的相应指数混叠，以此获取每个模态的频谱，如式(2-1)所示：

$$\left[\left(\delta(t) + \frac{j}{\pi t} \right) * \mu_k(t) \right] e^{-j\omega_k t} \quad (2-1)$$

(2) 构造变分问题，通过高斯平滑原理对 $\mu_k(t)$ 信号进行带宽解调，该算法的核心过程是变分问题的构造与求解，表达式中的 $\delta(t) + \frac{j}{\pi t}$ 是希尔伯特变换的冲激响应，需要构造的变分问题如式(2-2)所示：

$$\begin{cases} \min \left\{ \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * \mu_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \\ s. \quad t. \quad \sum_{k=1}^K u_k = f \end{cases} \quad (2-2)$$

式中 u_k 是第 k 个模态分量； ω_k 是对应分量的中心频率； f 是原始信号； $\delta(t)$ 是狄拉克函数。

(3) 通过引入拉格朗日乘子 $\lambda(t)$ 与二次惩罚因子 α ，将约束变分问题转化为无约束变分问题如式(2-3)所示：

$$L(\{\mu_k\}, \{\omega_k\}, \lambda) = \alpha \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * \mu_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \left\| f(t) - \sum_{k=1}^K \mu_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_{k=1}^K \mu_k(t) \right\rangle \quad (2-3)$$

式中 μ_k 为第 k 个模态分量， ω_k 是第 k 个变模态函数的有限带宽中心频率， $e^{j\omega_k t}$ 是中心频率复平面向量描述； $f(t)$ 是原始信号； α 是惩罚因子； λ 是拉格朗日乘子； $\hat{\mu}_k^{n+1}$ ， $\hat{f}(\omega)$ ， $\hat{\lambda}^n(\omega)$ 分别是 μ_k^{n+1} ， $f(\omega)$ ， $\lambda^n(\omega)$ 的傅里叶变换。

(4) 利用交替方向乘子法(Alternating Direction Method of Multipliers, ADMM)连续更新各模态分量及中心频率，式(2-4)通过频域闭式解更新模态分量，分母项 $1 + 2\alpha(\omega - \omega_k^n)^2$ 强制频谱能量集中至当前中心频率附近；式(2-5)则通过计算模态能量的加权平均频率，将中心频率动态调整至频谱能量重心位置。二者交替迭代，结合惩罚因子 α 的带宽约束与拉格朗日乘子 λ 的残差修正，逐步分离出窄带、中心频率明确的模态分量，实现信号的自适应分解。

$$\hat{\mu}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i < k} \hat{\mu}_i^{n+1}(\omega) - \sum_{i > k} \hat{\mu}_i^n(\omega) + \frac{\hat{\lambda}^n}{2}}{1 + 2\alpha(\omega - \omega_k^n)^2} \quad (2-4)$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{\mu}_k^{n+1}(\omega)|^2 d\omega}{\int_0^\infty |\hat{\mu}_k^{n+1}(\omega)|^2 d\omega} \quad (2-5)$$

(5) 继续重复循环，直到式(2-6)中的收敛条件被满足时，停止迭代，否则，返回继续更新 μ_k 、 ω_k 和 λ 。

$$\sum_k \frac{\|\hat{\mu}_k^{n+1} - \hat{\mu}_k^n\|_2^2}{\|\hat{\mu}_k^n\|_2^2} < \varepsilon \quad (2-6)$$

式中 ε 是收敛准则的判据；

研究发现，VMD 的分解结果主要受信号分解数 K 和拉格朗日乘子更新步长 τ 的影响^[44]。当 K 过大时，所有分解模态的中心频率会收敛甚至重叠。当 K 太小时，一些分量可能被划分到相邻的子序列中，或者甚至被丢弃。因此，在 VMD 中应采用最佳参数 K 和 τ ，这是由中心频率法确定的。对于每个分解模态，计算中心频率，并使用残差指数(REI)的最小值来计算参数 τ ，如式(2-7)所示：

$$R = \min \frac{1}{N} \sum_{t=1}^N \left| \sum_{k=1}^K [u_k(t) - \sum_{k=1}^K u_k] \right| \quad (2-7)$$

式中 N 表示分解数据的长度。

2.2.2 最优变分模态分解

为了削弱风速时间序列的非线性和非平稳性，利用 OVMD 将经验风速样本分解为多个子序列，以便通过具有最优核函数的 LSSVM 进行更好的预测。为获得最优分解结果，采用 CFO 方法计算各分解子系列的中心频率，确定分解模态的总分解数 K 。OVMD 需要预先确定两个关键参数：模态数 K （分解层数）和拉格朗日乘子更新步长 γ ，这两个参数对 VMD 的噪声抗扰性和分解效率有显著影响。当 K 值较小时，可能导致数据分割不充分。 K 值过大则可能引入过多噪声并引发模态混叠等问题。因此，本文采用中心频率法来确定合适的模态数 K 。

2.3 相空间重构

风速序列具有混沌时间序列的特征。PSR 是处理和分析混沌时间序列信号的重要方法，使用坐标延迟将一维时间序列信号转换为多维^[45]。在本研究中，选择

PSR 来确定 LSSVM 的输入矩阵, 然后, 分析风速数据内的内部关系^[46]。

混沌现象在自然界及人类社会中广泛存在, 表现为在确定性系统中出现的不规则性和随机性特征。混沌理论主要描述非线性系统的复杂多样性及其在不同尺度上的行为。可以通过坐标延迟将一维时间序列扩展到多维空间。

在时间序列分析中, 时间延迟 t 和最小嵌入维数 m 是构建相空间结构的核心参数。为了精确选定这两个关键参数, 研究采用了互信息法和 Cao 方法^[20]进行估算。

互信息, 作为信息论中的一个核心概念, 用于量化信源的信息含量。鉴于混沌系统固有的非线性特性, Fraser 与 Sweeney 利用互信息来评估系统内部的非线性相关性。凭借互信息在量化非线性特征方面的优势, 本研究采用该方法来确定合适的时间延迟 t 。

Cao 方法则是对虚假最近邻(False Nearest Neighbors, FNN)技术的一种改进。混沌时间序列可以视为高维混沌轨迹在低维空间中的投影, 这一过程中, 原本相距甚远的点可能会因投影效应而变得邻近, 从而产生伪近邻现象。Cao 方法通过剔除这些伪近邻, 旨在恢复混沌轨迹的真实结构, 由于其计算高效、能够有效区分随机信号与确定性信号, 并且适用于小样本数据处理, Cao 方法在相空间重构领域得到了广泛应用。因此, 本研究选择 Cao 方法来确定最小嵌入维数 m 。

对于给定的风速系列 $x=\{x_i | i=1,2,\dots,N\}$ 相应的相空间轨迹如式(2-8)所示:

$$X_j = [x_j, x_{j+\tau}, \dots, x_{j+(d-1)\tau}] (j=1, 2, \dots, N-(d+1)\tau) \quad (2-8)$$

式中 d 和 t 分别表示嵌入维数和延迟时间, N 表示风速序列的样本长度。

此外, 相应的重建矩阵和输出向量可以分别通过式(2-8)获得, 如式(2-9)和(2-10)所示:

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_{N-(d-1)\tau} \end{bmatrix} = \begin{bmatrix} x_1 & x_{1+\tau} & \cdots & x_{1+(d-1)\tau} \\ x_2 & x_{2+\tau} & \cdots & x_{2+(d-1)\tau} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N-(d-1)\tau} & x_{N-d\tau} & \cdots & x_N \end{bmatrix}_{(N-(d-1)\tau) \times d} \quad (2-9)$$

$$Y = \begin{bmatrix} x_{2+(d-1)\tau} \\ x_{3+(d-1)\tau} \\ \vdots \\ x_{N+1} \end{bmatrix}_{(N-(d-1)\tau) \times 1} \quad (2-10)$$

PSR 将原本相空间上的一个时间序列重构扩展到 N 个样本上, 影响 PSR 的

因素主要是嵌入维数 m 和延迟时间 τ 。延迟时间 τ 如果太小，会使相邻的两个坐标分量过于接近，辨识度不够高从而无法提供两个独立的坐标分量，延迟时间 τ 如果太大，两个坐标分量就会处于独立关系，使得吸引子轨迹在坐标分量上的投影毫无相关性。嵌入维数 m 过低会使得信息无法全部表示出来， m 过高会大大增加计算。为了确定 LSSVM 输入变量的维数，通常使用相空间重构^[47]或 PACF^[48]，后者需要主观的观察 PACF 图来确定维数，缺乏自适应性。因此为了自适应地确定 LSSVM 的输入矩阵维数，采用 PSR 方法计算 LSSVM 的输入维数，假设维数为 6 时重建输入变量矩阵，在多步预测过程中，以六个连续的风速时间序列作为输入变量来预测下一个风速值。此外，在预测之前，将分解的子序列线性归一化到[0, 1]范围内，以降低预测难度。预测值最终在预测后进行反归一化。其输入变量矩阵的构造如图 2-2 所示。

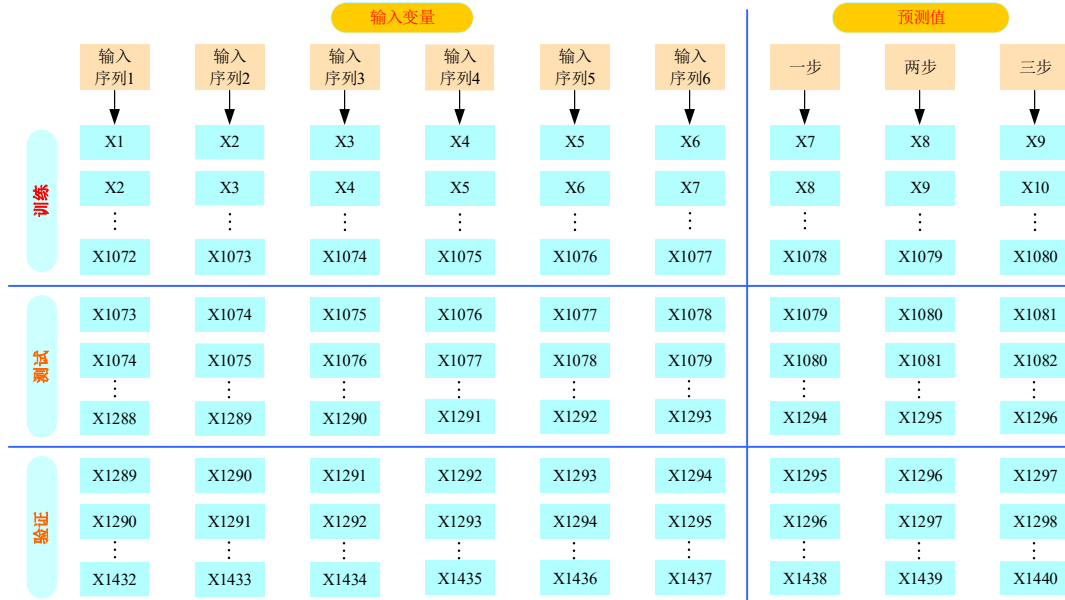


图 2-2 LSSVM 输入变量矩阵的构造

2.4 本章小结

本章主要对风速时间序列特性进行分析和针对风速时间序列的非线性问题进行风速数据预处理，为后续研究风速时间序列和构建预测模型奠定了理论基础。通过最优变分模态分解对数据进行预处理，采用 CFO 方法计算子序列中心频率，确定分解数 K 。通过相空间重构后，能进一步提高预测的准确性。

第3章 基于最优核函数选择的 LSSVM 模型

SVM 是一种机器学习算法，基于结构风险最小化原则和核函数技术，通过找到一个最优超平面或非线性决策函数来进行划分，并且让所有数据到这个超平面的距离最短。LSSVM 是对 SVM 的改进版，通过最小二乘法来解决 SVM 的寻找最优超平面的问题，降低计算难度。LSSVM 是处理小样本和非线性样本的有效方法，因此适用于短期风速预测。核函数及其参数的选择影响 LSSVM 的预测精度，通过采用最优核函数选择策略，为每个子序列的 LSSVM 模型选择最合适的核函数，从而有效提高预测精度。

3.1 机器学习

3.1.1 机器学习的基本结构

机器学习是一种先进的计算技术，利用特定的算法和统计模型，使计算机系统能够从大量数据中自动学习并提取有用的信息和知识^[49]。这种学习过程可以是监督的，即使用带有标签的数据来训练模型，使其能够预测新数据的标签或目标值；也可以是无监督的，即在没有任何标签的情况下，从数据中发现隐藏的结构、模式和关联。机器学习的核心优势在于其能够处理大规模的数据集，并从中提取出有价值的信息。随着数据量的不断增加和计算能力的不断提升，机器学习正在成为推动各行各业发展的重要力量。其基本模型如图 3-1 所示。

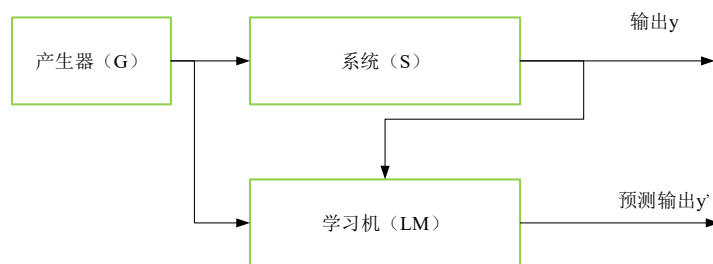


图 3-1 机器学习基本模型

由图 3-1 可知机器学习由产生器、系统和学习机组成。其中产生器是用来产生随机向量 x ，这些向量由分布函数 $F(x)$ 产生并且相互独立；系统主要作用是对输入 x 利用分布函数 $F(y|x)$ 进行操作，并根据内部的映射关系得到输出 y 与学习机生成的预测输出进行比较；学习机从已知函数参数集 $f(x, \omega)$ ， $\omega \in \Omega$ 中寻优，其中 Ω 是参数集合， $f(x, \omega)$ 是学习机得到的逼近函数，逼近能力取决于 $\{f(x, \omega)\}$ 。

机器学习的目标是找出逼近能力最好的逼近函数 $f(x, \omega_0)$ ，使得预测期望风险最小，这种方法是由训练集所影响的，训练集由联合概率 $F(x, y) = F(x)F(y|x)$ 随机产生的 n 个独立分布的数据组成。

$$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n) \quad (3-1)$$

使用最优逼近函数 $f(x, \omega_0)$ ，最小化预测期望风险，如式(3-2)所示：

$$R(\omega) = \int L(y, f(x, \omega)) dF(x, y) \quad (3-2)$$

式中 $L(y, f(x, \omega))$ 为 $f(x, \omega)$ 对结果 y 预测产生的损失大小，损失函数在回归逼近问题如式(3-3)所示：

$$L(y, f(x, \omega)) = (y - f(x, \omega))^2 \quad (3-3)$$

3.1.2 经验风险最小化

经验风险最小化(Empirical Risk Minimization, ERM)是指在机器学习中，通过选择一个在训练集上平均损失最小的模型来拟合数据^[50]，作为训练完成的模型。其中“经验风险”是模型在训练数据集上的平均损失，用来度量模型在整个训练数据集上预测效果的好坏。要确定最小化期望风险，需要通过大数定理进行估算来确定 $R_{emp}(\omega)$ ，如式(3-4)所示：

$$R_{emp}(\omega) = \frac{1}{n} \sum_{i=1}^n L(y_i, f(x_i, \omega)) \quad (3-4)$$

式中 $R_{emp}(\omega)$ 代表经验风险，可以通过给定的样本数据进行计算。因此，ERM 的核心目标是确定使 $R_{emp}(\omega)$ 最小的参数 ω ，具体的求解过程如下：

- (1) 对决策函数 $f(x, \omega)$ 的参数 ω 进行预设或假设，以确定一系列可接受的模型。
- (2) 针对每一个预设的模型，利用训练数据集来估算其经验风险（即分类误差）。
- (3) 从所有模型中，选择出经验风险最小的那个分类器（即决策函数）作为最优模型。

随后，通过经验风险最小化所求得的结果，可以作为期望风险 $R(\omega)$ 的一个近似或替代。众多学习方法均采用了这一经验风险最小化的核心理念。

在最小二乘方法中，如果将式(3-3)所定义的损失函数代入到式(3-4)中，就可以得到相应的经验风险函数，如式(3-5)所示：

$$R_{emp}(\omega) = \frac{1}{n} \sum_{i=1}^n L(y_i, f(x_i, \omega))^2 \quad (3-5)$$

在最大似然方法中，将式(3-1)带到式(3-4)中得到经验风险函数，如式(3-6)所示：

$$R_{emp}(\omega) = \frac{1}{n} \sum_{i=1}^n \log p(x_i, \omega) \quad (3-6)$$

综上所述，以最小经验风险来替代最小期望风险的前提是具备无限多的样本数据。然而，在实际情境中，所能获取的样本容量往往是有限的。因此，用经验风险来直接代替期望风险虽然在理论上可行，但在实际操作中缺乏足够的实证支持。

3.2 最小二乘支持向量机

3.2.1 支持向量机

SVM^[50]是一种广泛应用于模式识别和机器学习领域的监督学习算法，是一种在机器学习领域广泛使用的算法，特别是在分类问题中表现出色。基于统计学习理论，通过构建一个最优超平面来实现对数据的分类，将数据映射到高维特征空间，通过构建一个最优超平面来实现数据的划分或拟合。这个超平面被称为决策边界，能够最大化不同类别之间的间隔。SVM 的目标是找到一个超平面，使得距离该超平面最近的样本点的间隔最大，这样的超平面对未知的工程问题有很好的分类预测能力。

SVM 的基本思想在于通过一个隐含的非线性映射函数，将输入向量转换到一个高维特征空间。在这个高维空间中，SVM 构造了一个最优分类超平面，而是通过核函数来计算高维空间中向量的内积，而不需要明确知道这个非线性映射函数的具体形式，从而巧妙地避免了由于维度增加所带来的计算复杂性。这种方法使得 SVM 能够有效地处理非线性可分问题，而无需直接面对高维空间中的复杂计算。

SVM 数据处理过程如下：对于给定 n 维数据集 $(x_1, y_1), (x_2, y_2), \dots, (x_i, y_i) \in \mathbb{R}^n$ ，决策面如式(3-7)所示：

$$f(x) = \omega \bullet g(x) + b \quad (3-7)$$

式中 ω 为权重系数， $g(x)$ 为非线性映射函数， b 为阈值。

为了满足结构风险最小化原则，最优分类超平面要达到条件 $y_i(\omega \bullet g(x_i) + b) \geq 1$ ，引入非负松弛变量 ξ_i ，优化问题可以转化如式(3-8)所示：

$$\begin{cases} \min \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n \xi_i, C > 0 \\ s.t. \quad y_i [(\omega \cdot g(x_i) = b)] \geq 1 - \xi_i, \xi_i \geq 0 \end{cases} \quad (3-8)$$

式中 C 为惩罚因子。引入拉格朗日算法后，原始的优化问题可以被转化为对偶形式进行求解，如式(3-9)所示：

$$\begin{cases} \min \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n y_i y_j \alpha_i \alpha_j K(x_i, x_j) - \sum_{i=1}^n \alpha_i \\ s.t. \quad \sum y_i \alpha_i = 0, 0 \leq \alpha_i \leq C \end{cases} \quad (3-9)$$

式中 $K(x_i, x_j)$ 是核函数。

3.2.2 最小二乘支持向量机

LSSVM 是 SVM 的一种变体，继承了 SVM 的优点，在优化目标函数时采用最小二乘损失函数，通过线性映射将输入空间映射到高维特征空间，在该空间中利用核函数对数据进行分类或回归，从而将 SVM 的优化问题转化为求解一个等价的线性方程问题，使得 LSSVM 的计算效率更高^[52]。

对于给定的训练集输入 $X = \{x_1, x_2, \dots, x_n\}$ ，输出 $Y = \{y_1, y_2, \dots, y_n\}$ ，LSSVM 算法如式(3-10)所示：

$$y(x) = \omega^T \theta + b \quad (3-10)$$

式中 θ 表示一个非线性映射函数，能够将输入 X 与输出 Y 之间的复杂非线性关系转化为 Y 与 $\theta(X)$ 之间的线性关系， b 为偏差向量，而 ω^T 为权重向量。基于风险最小化原则，该问题可以转化为约束优化问题如式(3-11)所示：

$$\begin{cases} \min_{\omega, b, e} J(\omega, e) = \frac{1}{2} \|\omega\|^2 + \frac{1}{2} \gamma \sum_{i=1}^n e_i^2 \\ s.t. \quad y_i (\omega^T \varphi(x_i) + b) = 1 - e_i, i = 1, 2, \dots, n \end{cases} \quad (3-11)$$

式中 e_i 为误差项，用于衡量模型预测与实际观测之间的差异； γ 为正则化因子，用于平衡经验风险和结构风险； $\psi(\cdot)$ 为非线性映射函数，将输入数据映射到高维特征空间，以线性可分的方式处理原本非线性可分的问题。通过引入 Lagrange 乘子 α ，构建一个 Lagrange 函数如式(3-12)所示：

$$L(\omega, b, e, \alpha) = J(\omega, e) - \sum_{i=1}^n \alpha_i (y_i (\omega^T \varphi(x_i) + b) - 1 + e_i) \quad (3-12)$$

对式(3-12)求偏导，得到回归函数如式(3-13)所示：

$$y(x) = \sum_{i=1}^n \alpha_i K(x, x_i) + b \quad (3-13)$$

3.2.3 核函数选择

LSSVM 是处理小样本和非线性样本的有效方法，因此，可以有效应用于短期风速预测^[54]。核函数的选择在很大程度上决定了模型处理不同类型数据和处理问题的能力，并直接影响模型的输出结果，因此，核函数的类型及其参数会影响 LSSVM 的回归性能^[52]。基于不同核函数构建 LSSVM 获得不同的回归性能，因此，核函数的类型是构建高性能 LSSVM 的重要因素，主要典型核函数^[52] 如式 (3-14)~(3-21)所示：

(1) 径向基函数 (RBF)

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\delta^2}\right) \quad (3-14)$$

(2) 指数径向基函数 (ERBF)

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|}{2\delta^2}\right) \quad (3-15)$$

(3) 多项式核函数 (Poly)

$$K(x_i, x_j) = (x_i \cdot x_j + r)^d \quad (3-16)$$

(4) 双曲正切核函数(Sigmoid)

$$K(x_i, x_j) = \tanh(a \cdot x_i \cdot x_j + r) \quad (3-17)$$

(5) 高斯核函数 (Gaussian)

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{\delta^2}\right) \quad (3-18)$$

(6) Morlet 小波基函数(Morlet)

$$K(x_i, x_j) = \cos\left[1.75 \frac{\|x_i - x_j\|}{\sigma}\right] \cdot \exp\left(-\frac{\|x_i - x_j\|^2}{2\delta^2}\right) \quad (3-19)$$

(7) 墨西哥帽小波(Mexihat)

$$K(x_i, x_j) = \frac{2}{\sqrt{3}} \pi^{-\frac{1}{4}} \left[1 - \frac{\|x_i - x_j\|^2}{\delta^2}\right] \cdot \exp\left(-\frac{\|x_i - x_j\|^2}{2\delta^2}\right) \quad (3-20)$$

(8) 线性核函数 (Linear)

$$K(x_i, x_j) = x_i \cdot x_j \quad (3-21)$$

利用具有最优核函数的 LSSVM 获得的结果来重建最终的风速预测结果，如图 3-1 所示。

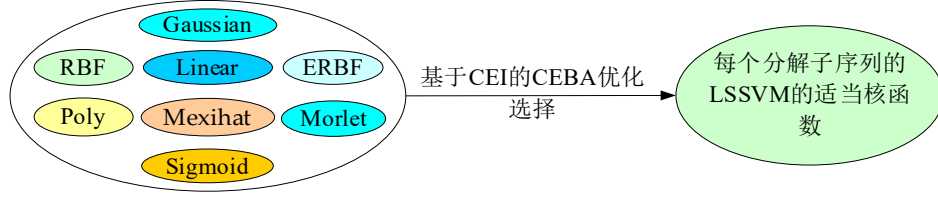


图 3-1 核函数选择

由于预测模型的预测性能受输入变量的影响，并且基于一种核函数的 LSSVM 不能让所构建的 VPBL 中的所有分解子序列获得令人满意的预测精度。为了提高所构建的 VPBL 的预测性能，对于利用 OVMD 获得的每个分解子序列，应用最优核函数选择策略为 LSSVM 选择合适的核函数。

计算基于不同核函数下的 LSSVM 的预测误差，包括均方根误差(RMSE)、平均绝对百分比误差(MAPE)、平均绝对误差(MAE)和标准差(SDE)。获得的预测误差被归一化，如式(3-22)所示：

$$Index_i^* = \frac{Index_i - \min_{1 \leq i \leq N}(Index)}{\max_{1 \leq i \leq N}(Index) - \min_{1 \leq i \leq N}(Index)} \quad (3-22)$$

式中 N 代表核函数的数量， $Index_i^*$ 表示 RMSE、MAPE 和 MAE 的标准化评估指数。

根据预测误差指数，基于第 i 个核函数的 LSSVM 的 CEI 值如式(3-23)所示：

$$CEI_i = (0.4 \times RMSE_i^* + 0.3 \times MAPE_i^* + 0.3 \times MAE_i^*) \quad (3-23)$$

对于每个分解子序列，具有最小 CEI 值的 LSSVM 的第 i 个核函数被选定为核函数组合中的最佳核函数。

3.3 基于 LSSVM 的风速预测

考虑到不同的风速时间序列具有不同的非线性特性，没有一种 LSSVM 核函数适用于所有不同特性的风速时间序列。设计一种核函数选择策略，以从 LSSVM 的典型核函数集中选择最佳核函数，该核函数集包括 RBF、ERBF、Poly、Sigmoid、Gaussian、Morlet、Linear、Mexihat。当 CEI 值最小时，该核函数为 LSSVM 在子序列预测时的最优核函数，能使 LSSVM 更好地捕捉分解子序列的动态特性，并采用滚动预测策略进行风速预测。通过这种最优核函数选择策略，无论是分解后的子序列还是整个预测方法的预测精度都能得到有效地提高。在核函数选择过程中，每个分解子序列 $S_i(1-1000)$ 用于训练具有不同核函数的预测器 LSSVM， $S_i(1001-1296)$ 用于确定具有最小 CEI 值的 LSSVM 预测值以确定最佳核函数。利

用具有最优核函数的 LSSVM 进行预测各个子序列的预测值 $\hat{y}_k$ (1297-1440) 进行积累来获得最终的风速预测结果，具体流程图如图 3-3 所示。

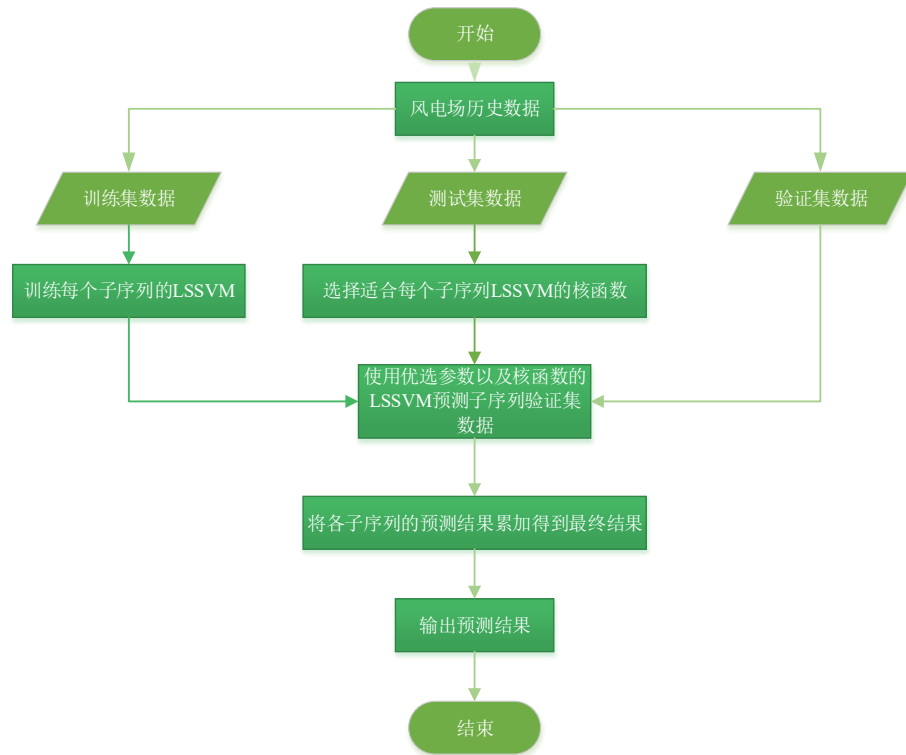


图 3-3 LSSVM 预测流程图

基于 LSSVM 的风速预测流程如下：

- (1) 将风电场的历史风速数据划分为训练集、测试集和验证集，其中训练集用于训练每个子序列的 LSSVM，而测试集则用来为每个子序列的 LSSVM 选择最优核函数，验证集用来验证模型的预测性能。
- (2) 通过训练样本训练 LSSVM，并利用 CEBA 算法优化 LSSVM 的参数以及选取核函数，得到最优的模型。
- (3) 利用已经训练完成的 LSSVM 模型，对验证集中的数据进行预测，将每个子序列的预测结果积累，从而得出最终的预测结果。

3.4 本章小结

本章首先阐述了机器学习机 SVM 和 LSSVM 的工作原理，并针对 LSSVM 的风速预测性能与核函数及其参数以及输入风速的特性有关，单一核函数无法适应不同风速分解子序列的问题，本章构建了一种新型的 LSSVM 核函数选择策略，通过蝙蝠算法对每个分解子序列选择最优的 LSSVM 核函数。

LSSVM 是从机器学习损失函数角度对 SVM 进行改进,在 LSSVM 优化问题的目标函数中使用二范数,并利用等式约束条件代替 SVM 中的不等式约束条件,使得 LSSVM 方法的优化问题的求解变为通过 Kuhn-Tucker 条件得到的一组线性方程组的求解,采用滚动预测策略进行风速预测。通过这种最优核函数选择,可以有效地提高分解结果和整个预测方法的预测精度。

第4章 蝙蝠算法的改进

蝙蝠算法(BA)是一种寻找全局最优解的优化算法,受到蝙蝠利用超声波寻找食物时来进行定位机理的启发,应用于优化问题找到最优解,然而标准的蝙蝠算法存在早熟收敛和局部最优问题。随着迭代次数的增加,种群多样性下降,更新最优解的能力下降,算法容易陷入局部最优解。因此,为了提高 BA 的优化性能和降低算法早熟收敛的可能性,对 BA 进行改进。首先,在 BA 中引入混沌映射方法初始化种群,即采用具有良好遍历能力的逻辑自映射函数对种群进行初始化,增加初始种群的遍历性和多样性;其次,在标准 BA 中引入了动态扰动系数、正切随机优化和平均种群位置方法,增强了全局搜索能力。为了验证改进后 BA 的有效性,采用标准的 CEC15 测试集进行测试。

4.1 蝙蝠算法

4.1.1 标准蝙蝠算法

BA 模拟自然蝙蝠在飞行和觅食过程中的回声定位行为,搜索优化问题的解空间以获得最优解,但需要满足如下所示的理想规则^[55]:

(1) 蝙蝠在飞行过程中利用回声定位感知距离,能自动识别猎物 and 障碍物进行捕食或捕食性飞行。

(2) 在捕食飞行过程中,波长和响度 A_0 会自动发生变化。蝙蝠会根据自身与猎物的距离自动调整超声波脉冲频率 r 和脉冲频率 f 。在非狩猎飞行过程中,蝙蝠以速度 v_i 在空间位置 x_i 随机飞行。

(3) 假设脉冲响度从最大值 A_{max} 逐渐降低到最小值 A_{min} 。

在 BA 中,每个蝙蝠个体代表优化问题的可行解,而解空间中的猎物代表最优解。在第 $(t+1)$ 次迭代中,更新第 i 个蝙蝠的频率、速度和位置,如式(4-1)~(4-3)所示:

$$r_i = r_{\min} + \beta \times (r_{\max} - r_{\min}) \quad (4-1)$$

$$v_i(t+1) = v_i(t) + r_i \times (x_i(t) - x^*) \quad (4-2)$$

$$x_i(t+1) = v_i(t) + x_i(t) \quad (4-3)$$

式中 β 表示 $(0, 1)$ 内的随机值, x^* 表示当前迭代中的全局最优解,从式(4-3)中可以看出,只有频率是影响每个个体在每次迭代中步长的关键因素。

当 BA 执行局部优化时, 在第 t 次迭代时从最优解集中随机选择个体, 并对其应用随机扰动变量。该算法在随机解附近的空间中搜索新解, 如式(4-4)所示:

$$x_{new} = x_{old} + \lambda \times A(t) \quad (4-4)$$

式中 λ 表示(0, 1)内的均匀随机扰动变量, $A(t)$ 表示第 t 次迭代时的个体响度。

在觅食过程中, 响度和脉冲发射随着蝙蝠与猎物的距离而变化。在觅食开始时, 蝙蝠会发出高音量和低频脉冲。发现猎物后, 响度降低, 脉率增加, 狩猎过程的如式(4-5)所示:

$$\begin{cases} A_i(t+1) = \theta \times A_i(t) \\ r_i(t) = r^0 [1 - \exp(-\eta t)] \end{cases} \quad (4-5)$$

式中 η 是常数, θ 表示(0, 1)内的随机变量, $A_i(t)$ 和 $r_i(t)$ 分别是第 t 次迭代时的响度和脉冲发射率。

算法 I 中列出了标准 BA 的伪代码。

算法 I: 标准 BA 的伪代码

1. 在 BA 中初始化参数, 包括群体大小 N_{max} , 维度编号 D , 频率范围 $[f_{min}, f_{max}]$, 迭代次数 T_{max} 等。
 2. 生成初始速度 v_i , 位置 x_i , 脉冲响度 A_i , 和频率 r_i 等。
 3. 使用初始的 x_i 计算适应度函数 $fit(x_i)$ 。
 4. **While** $t \leq T_{max}$ **do**
 5. 根据等式(4-2) ~ (4-3)更新单个脉冲频率 r_i , 速度 v_i , 和位置 x_i 。
 6. **if** $rand > r_i$ **then**
 7. 使用随机游走法根据等式(4-4)获得新的种群最优解。
 8. **endif**
 9. 使用新的个体 x_i 计算适应度函数。
 10. **if** ($f_{new} \leq fit(x_{new}(i))$) & ($rand < A_i$) **then**
 11. 更新种群个体 $x_i(t+1)$ 。
 12. **endif**
 13. 根据等式(4-5)获得响度 $A_i(t+1)$ 和脉冲频率 $r_i(t+1)$ 。
 14. 比较和更新, 得到当前全局最优解。
 15. **endwhile**
 16. **endwhile**
 17. 输出全局最优解。
-

4.1.2 标准蝙蝠算法的不足

BA 是一种新的启发式优化算法, 模拟蝙蝠在飞行和觅食过程中的回声定位。

蝙蝠每秒钟不断调整频率和响度,以寻找食物和避开障碍物。在蝙蝠捕食过程中,蝙蝠发出的超声波数量和响度与蝙蝠和猎物的距离成反比。随着蝙蝠与猎物距离的减小,蝙蝠发出的超声波数量和响度逐渐减小。

从算法 I 中的伪代码可以看出,由个体种群得到的新可行解需要在第 7 行中通过随机游走方法满足第 6 行中的条件“ $rand > r_i$ ”。从对标准 BA 工作机制的分析中可以得出一些不足之处,如下所示:

(1) 标准 BA 的初始种群是随机产生的,遵循非均匀分布,缺乏变异机制,容易导致标准 BA 的早熟收敛和局部最优。

(2) 从算法 I 中的第 6 行可以看出,当满足“ $rand > r_i$ ”条件时,可行解被更新。从公式(4-5)来看,随着迭代次数的增加, $r_i(t)$ 逐渐接近 $r_0(t)$,而传输频率 r_i 逐渐增大,更新最优解的能力下降,不利于种群的多样性,算法容易陷入局部最优解。

(3) 响度 $A_i(t)$ 和脉冲发射率 $r_i(t)$ 更新是单调的,使得 BA 在迭代优化过程中缺乏灵活性,不能根据历史种群状态进行动态调整,不利于种群个体的多样性。

(4) 从算法 I 和式(4-5)可以看出,标准 BA 的搜索能力取决于随机变量 λ ,在(0, 1)内,当个体陷入局部最优时,算法很难跳出,影响了标准 BA 的优化能力。

4.2 蝙蝠算法的改进策略

针对上述问题,改进了标准 BA 的种群多样性和搜索能力,以增强局部搜索和全局优化能力,防止 BA 在快速收敛过程中陷入局部收敛。因此,本节从两个方面对标准 BA 进行了修改,获得增强型蝙蝠算法(CEBA),增强 BA 中种群的多样性和提高搜索能力。

4.2.1 基于混沌方法增强 BA 种群多样性

启发式智能优化算法基本都会遇到早熟收敛。学者们提出了不同的解决方案来防止优化算法的早熟收敛。混沌是一种特殊的非周期、不规则的运动状态,具有多样性、遍历性和随机性等特点。混沌映射将变量按照一定的规则变换到解空间,增强了优化算法中种群的多样性,是解决早熟收敛问题的典型方法^[53]。本文将混沌映射应用于启发式优化算法,以提高可行解的多样性,降低过早收敛的可能性。

由于混沌映射的不可重复性,基于混沌映射的算法与原有算法相比一般可以

获得更好的全局搜索能力^[54]，这将有助于算法逃离局部最优，避免早熟收敛。为了克服标准 BA 的第一个缺点，在标准 BA 中也引入了混沌映射方法^[55]。采用具有良好遍历能力的逻辑自映射函数对种群进行初始化，以提高初始种群的多样性和遍历性，逻辑自映射函数如式(4-6)所示：

$$L_{i+1}^d = 1 - 2 \times L_i^d \quad (4-6)$$

式中 L_i^d 表示 $(-1, 1)$ 内的变量。

根据逻辑自映射函数，将 BA 中的个体转换为 $(-1, 1)$ ，如式(4-7)所示，然后初始化种群个体，如式(4-8)所示：

$$L_i^d = \frac{2(x_i^0 - a_i^d)}{b_i^d - a_i^d} - 1 \quad (4-7)$$

$$x_i^0 = \frac{(b_i^d - a_i^d) \times L_i^d + (b_i^d + a_i^d)}{2} \quad (4-8)$$

式中 b_i^d 和 a_i^d 表示最大值和最小值。

从式(4-5)中可以看出。脉冲频率 r_i 和响度 A_i 是单调的。在 BA 中，这种变量单调性不利于产生可行解的多样性。用混沌正弦映射函数代替标准 BA 的发射脉冲频率 r_i 和响度 A_i 。改善了脉冲频率和响度的多样性，正弦映射，用于调整脉冲速率，如式(4-9)所示：

$$\begin{cases} r_{i+1}^d = \tau \times (r_i^d) \times \sin(\pi \times r_i^d) \\ A_{i+1}^d = \eta \times (A_i^d) \times \sin(\pi \times r_i^d) \end{cases} \quad (4-9)$$

从图 4-1 可以看出，正弦映射函数表现出较好的混沌多样性和遍历性。

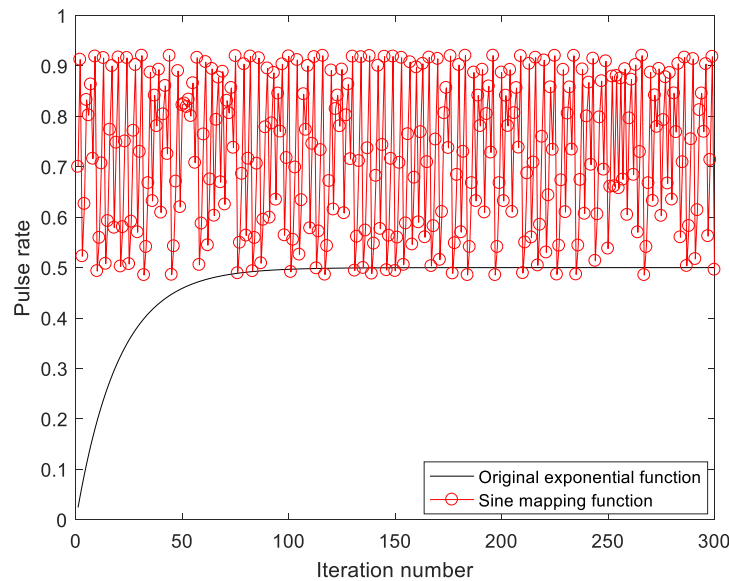


图 4-1 正弦映射函数曲线与指数函数曲线

4.2.2 提高 BA 搜索能力的策略

为了增强全局搜索能力，避免陷入局部最优，在标准 BA 中引入了动态扰动系数、正切随机优化和平均种群位置方法。

(1) 动态扰动系数

从标准 BA 及其伪代码可以看出，位置 $x_i(t+1)$ 根据式(4-9)更新。在式(4-9)中，根据 $x_i(t)-x^*$ 更新个体速度 $v_i(t)$ ，其是当前个体位置和上一代最佳位置之间的距离。位置更新受到距离 $x_i(t)-x^*$ 的约束，这使得 BA 容易陷入局部最优，因此，不容易获得全局最优解。为了避免陷入局部最优，在当前个体位置引入动态扰动系数，对当前种群的个体位置进行扰动，使个体跳出其位置进入邻域空间搜索可行解，从而增强了全局优化的灵活性和搜索能力。动态扰动系数 ρ 和位置更新 $x_i(t+1)$ 如式(4-10)和(4-11)所示：

$$\rho = 1 + \cos\left(\frac{-\pi t}{2 \times T_{\max}} - \frac{\pi}{2}\right) + \gamma \times \text{betarnd}() \quad (4-10)$$

$$x_i(t+1) = \rho \times x_i(t) + v(t) \quad (4-11)$$

式中 t 、 $T_{\max}$ 分别为当前迭代次数和最大迭代次数，动力扰动系数 ρ 如图 4-2(a) 所示。

从图中可以看出，对当前个体位置存在较大扰动，使得种群 $x_i(t)-x^*$ 的约束降低，促进个体进入相邻区域，增强了 BA 算法的全局搜索能力。随着迭代次数的增加，对当前种群的扰动减小，BA 算法搜索局部最优解的能力增强。 γ 为扰动偏差因子，根据文献[56]设为 0.5， $\text{betarnd}()$ 表示为 β 分布随机数生成函数， $\gamma \times \text{betarnd}()$ 可对系数 ρ 施加动态扰动，增强了 BA 搜索全局最优解的能力。

(2) 正切随机优化机制

从式(4-4)可以看，BA 在局部搜索阶段通过当前最优解和当前脉冲响度进行更新。在(4-4)中， λ 表示(0,1)内的均匀随机扰动变量，这使得 BA 容易陷入盲搜索状态，不能更好地挖掘算法的局部寻优能力。为了充分利用 BA 的局部搜索能力，用正切随机探索代替式中的均匀随机扰动变量 λ 。式(4-4)可以改进为式(4-12)并如图 4-2(b)所示：

$$x_{\text{new}} = x_{\text{old}} + \tan((r - 0.5) \times \pi) \times A(t) \quad (4-12)$$

式中 r 表示(0,1)内均匀分布的随机数, $\tan((r-0.5)\times\pi)$ 是正切随机变量。

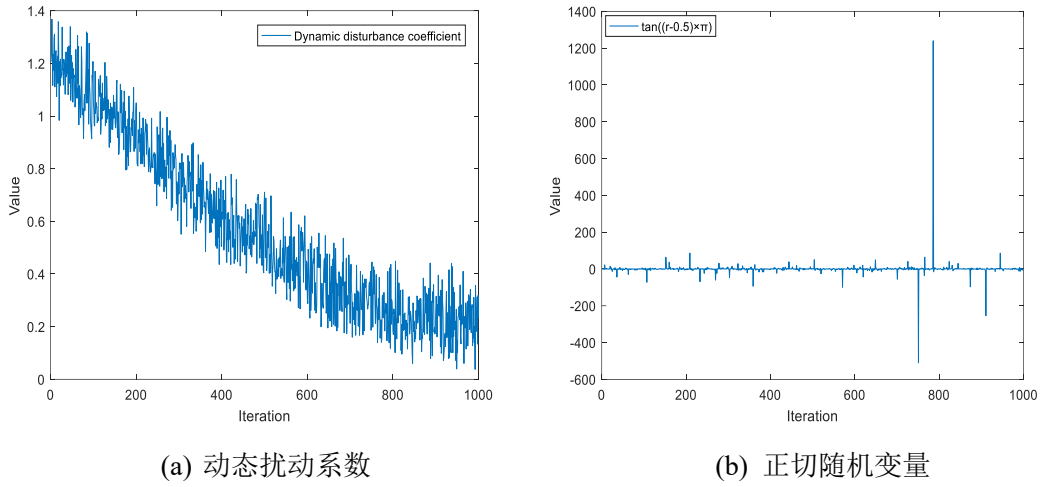


图 4-2 动力扰动系数、正切随机变量与迭代的关系

从图 4-2(b)可以得出以下结论。

1) $\tan((r-0.5)\times\pi)$ 的大部分值分布在 0 附近,使得算法围绕当前最优解进行搜索,增强了算法的局部寻优能力。

2) $\tan((r-0.5)\times\pi)$ 的罕见大值可以使算法跳出局部最优解,进入其他空间搜索最优解,这使得 BA 的局部优化策略更加灵活,而不是盲目随机搜索。

(3) 平均种群位置

标准 BA 依靠“更好更近”的个体来引导搜索最优解。随着迭代次数的增加,算法不断搜索最优解,种群个体不断收敛。

如果当前全局最优解与种群最优解之间的距离较大,则其他种群个体受其影响,容易导致早熟收敛。为了解决这一问题,在标准 BA 中引入了平均位置法,利用种群的平均位置来引导个体蝙蝠。在算法初始阶段,种群个体的平均位置较大。随着迭代次数的增加,算法逐渐收敛,平均位置逐渐减小。利用种群的平均位置来调整变异步长,在早期迭代中可以提高算法的寻优能力,在后期迭代中可以加速算法的收敛,可以有效地降低算法陷入过早局部最优的可能性。第 i 个个体的速度更新如式(4-13)和(4-14)所示:

$$m(t) = \left(\frac{1}{N} \sum_{i=1}^N x_{i,1}(t), \frac{1}{N} \sum_{i=1}^N x_{i,2}(t), \dots, \frac{1}{N} \sum_{i=1}^N x_{i,D}(t) \right) \quad (4-13)$$

$$v_i(t+1) = v_i(t) + (m(t) - x_i(t)) \times r_i \quad (4-14)$$

式中 N 和 D 分别表示种群数量和种群维数。

4.3 CEBA 优化性能的验证

为了测试改进 BA 的优化性能,采用标准 CEC15 测试集^[57]中的函数将 CEBA 与其他常见的优化算法进行比较,包括 PSO^[58]、灰狼优化算法(GWO)^[59]和 QPSO^[60],并根据相应文献设置算法参数。CEC15 标准测试集如表 4-1 所示。

表 4-1 标准 CEC15 测试集

序号	函数	维数	种类
F1	Rotated Bent Cigar Function	10/30	单峰
F2	Rotated Discus Function	10/30	单峰
F3	Shifted and Rotated Weierstrass Function	10/30	单峰
F4	Shifted and Rotated Schwefel's Function	10/30	多模态
F5	Shifted and Rotated Katsuura Function	10/30	多模态
F6	Shifted and Rotated HappyCat Function	10/30	多模态
F7	Shifted and Rotated HGBat Function	10/30	多模态
F8	Shifted and Rotated Expanded Griewank's plus Rosenbrock's Function	10/30	多模态
F9	Shifted and Rotated Expanded Scaffer's F6 Function	10/30	多模态
F10	Hybrid Function 1 (N=3)	10/30	混合
F11	Hybrid Function 2 (N=4)	10/30	混合
F12	Hybrid Function 3 (N=5)	10/30	混合
F13	Composition Function 1 (N=5)	10/30	混合
F14	Composition Function 2 (N=3)	10/30	复合
F15	Composition Function 3 (N=5)	10/30	复合

采用 Wilcoxon 符号秩检验 (W 检验) 和 t 检验对不同算法和 CEBA 的优化结果进行比较和分析。 t 检验是一种定量参数检验,假设不同算法的优化结果服从正态分布, t 检验指数比 W 检验更准确。然而,W 检验是一种非参数检验,不要求算法的优化结果服从分布^[61]。在 W 检验和 t 检验的统计结果中,“+”、“ $\approx$ ”和“-”分别表示算法优化结果显著优于对比算法、无显著差异和显著劣于对比算法。表 4-2 列出了标准 CEC15 测试集的低维(10 维)和高维(30 维)测试的 100 次独立运行的平均值。

表 4-2 高低维 CEC15 标准测试集运行结果

函数	指标	低维 CEC15 测试函数集 (10 维)					高维 CEC15 测试函数集 (30 维)				
		CEBA	BA	QPSO	PSO	GWO	CEBA	BA	QPSO	PSO	GWO
F1	Best	5.2872E+03	1.3211E+04	9.2581E+03	5.3754E+04	7.9853E+04	7.4529E+05	8.7087E+06	8.3329E+06	9.4846E+06	9.7622E+06
	Mean	2.0321E+05	4.1278E+05	7.0615E+05	7.3202E+05	7.6575E+05	5.1990E+06	1.1365E+07	1.0829E+07	1.4054E+07	1.5066E+07
	Std	1.9262E+05	4.8372E+05	4.3709E+05	5.2488E+05	6.027E+05	1.3917E+06	5.1572E+06	4.3489E+06	6.8769E+06	8.0511E+06
	t-test	/	+	+	+	+	+	+	+	+	+
	W-test	/	+	+	+	+	+	+	+	+	+
F2	Best	8.3658E+03	6.5471E+04	5.7102E+04	7.6921E+04	7.8549E+04	3.5732E+06	5.9031E+06	5.7675E+06	6.9089E+06	9.7284E+06
	Mean	1.5608E+06	2.9307E+06	3.9145E+06	8.9921E+06	7.5534E+06	6.2361E+07	1.2298E+08	1.0724E+08	1.3569E+08	3.0768E+08
	Std	6.6793E+05	2.7911E+06	3.0170E+06	6.3301E+06	6.8867E+06	4.0481E+07	4.9228E+07	4.8128E+07	6.4438E+07	8.6892E+07
	t-test	/	+	+	≈	+	+	+	+	≈	+
	W-test	/	+	+	+	+	+	+	+	≈	+
F3	Best	6.8938E+00	9.5256E+00	1.0212E+01	1.1295E+01	1.2548E+01	2.0220E+01	2.0534E+01	2.0368E+01	2.0603E+01	2.0675E+01
	Mean	1.8788E+01	1.8211E+01	1.9548E+01	1.9677E+01	1.9217E+01	2.0399E+01	2.0602E+01	2.0443E+01	2.0639E+01	2.0759E+01
	Std	1.8993E+00	2.0379E+00	2.2997E+00	2.4872E+00	2.1133E+00	5.1278E-02	6.4473E-02	6.2257E-02	8.0332E-02	8.7827E-02
	t-test	/	+	+	+	+	+	+	+	+	+
	W-test	/	+	+	+	+	+	+	+	+	+
F4	Best	3.3476E+00	4.5441E+00	6.0122E+00	7.7660E+00	8.5661E+00	3.9910E+01	5.8655E+01	5.4655E+01	6.3328E+01	7.9349E+01
	Mean	7.1327E+00	1.0748E+01	1.3990E+01	1.5908E+01	1.7119E+01	7.0209E+01	8.9788E+01	8.3459E+01	9.2697E+01	1.0127E+02
	Std	1.9656E+00	2.3831E+00	3.7877E+00	4.0869E+00	4.5119E+00	8.8227E+00	1.0662E+01	9.9061E+00	1.3711E+01	1.9322E+01
	t-test	/	+	+	+	+	+	+	+	+	+
	W-test	/	+	+	+	+	+	+	+	+	+
F5	Best	7.1229E+00	1.5965E+01	2.0994E+01	2.8968E+01	3.0319E+01	8.2457E+02	1.2286E+02	1.0692E+03	1.3269E+03	1.4566E+03
	Mean	1.0278E+02	2.0949E+02	2.8784E+02	3.8783E+02	5.6559E+02	2.5729E+03	4.3673E+03	4.8757E+03	4.6544E+03	4.9983E+03
	Std	7.7354E+01	9.2898E+01	1.0221E+02	6.8927E+01	8.9297E+01	4.2811E+02	5.0698E+02	4.6588E+02	5.2877E+02	5.6601E+02
	t-test	/	+	+	+	+	+	+	+	+	+
	W-test	/	+	+	+	+	+	+	+	+	+
F6	Best	7.0354E+01	9.5619E+01	1.1993E+02	1.7881E+02	1.9338E+02	5.8665E+04	7.5448E+04	7.0037E+04	8.5298E+04	1.1733E+05
	Mean	1.1928E+03	1.5794E+03	1.8977E+03	6.2338E+03	7.8755E+03	5.6634E+05	8.6081E+05	7.9324E+05	9.4298E+05	1.1899E+06
	Std	9.4008E+02	9.9691E+02	1.0887E+03	1.3928E+03	1.4647E+03	3.8750E+05	4.7619E+05	4.3667E+05	4.8912E+05	7.1522E+05
	t-test	/	+	+	+	+	+	+	+	+	+
	W-test	/	+	+	+	+	+	+	+	+	+
	Best	4.6690E-01	7.3557E-01	6.4756E-01	8.7958E-01	1.0447E+00	6.6687E+00	9.7927E+00	8.7013E+00	1.0458E+01	1.2819E+01
	Mean	9.6771E-01	1.1974E+00	1.5648E+00	2.5872E+00	3.2117E+00	1.1237E+01	1.5446E+01	1.2671E+01	1.8538E+01	2.9155E+01

F7	Std	3.7719E-01	4.0558E-01	3.6318E-01	4.6576E-01	5.8876E-01	1.5429E+00	2.6592E+00	1.5368E+00	3.9598E+00	5.1772E+00
	<i>t</i> -test	/	+	+	+	+	+	+	+	+	+
	W-test	/	+	+	+	+	+	+	+	+	+
F8	Best	3.1177E+00	6.6167E+00	7.2269E+00	8.3227E+00	2.1451E+01	1.6952E+04	5.4559E+04	2.7899E+04	6.4339E+04	7.2119E+04
	Mean	4.3571E+02	5.5098E+02	7.7217E+02	9.3376E+02	1.2788E+03	1.7433E+05	4.9917E+05	3.3822E+05	6.9376E+05	7.4539E+05
	Std	3.5722E+02	4.1037E+02	4.1277E+02	4.1983E+02	5.5098E+02	8.2118E+04	2.1373E+05	1.6976E+05	3.1088E+05	4.5322E+05
	<i>t</i> -test	/	+	+	+	+	+	+	+	+	+
	W-test	/	+	+	+	+	+	+	+	+	+
F9	Best	1.0012E+02	1.0014E+02	1.0011E+02	1.0015E+02	1.0019E+02	1.0302E+02	1.0485E+02	1.0465E+02	1.0522E+02	1.0611E+02
	Mean	1.0022E+02	1.0026E+02	1.0029E+02	1.0038E+02	1.0056E+02	1.0392E+02	1.0854E+02	1.0694E+02	1.0784E+02	1.0993E+02
	Std	6.1036E-02	7.7899E-02	7.1697E-02	8.5677E-02	1.0549E-01	3.7723E-01	5.8992E-01	4.7996E-01	6.0688E-01	7.4217E-01
	<i>t</i> -test	/	+	+	+	+	+	+	+	+	+
	W-test	/	+	+	+	+	+	+	+	+	+
F10	Best	2.4588E+02	2.7154E+02	2.7269E+02	3.5677E+02	4.4766E+02	1.0475E+04	3.5869E+04	2.8776E+04	3.7885E+04	4.9439E+04
	Mean	4.0154E+02	5.8944E+02	5.5960E+02	8.3788E+02	8.9778E+02	1.3429E+05	4.0211E+05	3.7463E+05	4.2214E+05	4.3550E+05
	Std	1.0659E+02	1.5996E+02	2.0226E+02	3.3607E+02	3.5054E+02	7.7855E+04	1.3447E+05	1.2287E+05	1.4468E+05	1.7228E+05
	<i>t</i> -test	/	+	+	+	+	+	+	+	+	+
	W-test	/	+	+	+	+	+	+	+	+	+
F11	Best	3.6649E+00	5.6778E+00	5.0937E+00	7.7578E+00	8.2887E+00	3.0733E+02	3.1211E+02	3.0831E+02	3.1578E+02	3.2998E+02
	Mean	4.2236E+01	7.0972E+01	6.0922E+01	8.3393E+01	9.9113E+01	3.5766E+02	3.7764E+02	3.6797E+02	3.8122E+02	4.0776E+02
	Std	7.0718E+01	9.7618E+01	8.0212E+01	8.3501E+01	7.6988E+01	9.6580E+01	1.3447E+02	1.2176E+02	1.4055E+02	1.5991E+02
	<i>t</i> -test	/	+	0	+	+	+	+	0	+	+
	W-test	/	+	+	+	+	+	+	+	+	+
F12	Best	1.0076E+02	1.0091E+02	1.0089E+02	1.0227E+02	1.0274E+02	1.0422E+02	1.0646E+02	1.0446E+02	1.0594E+02	1.0617E+02
	Mean	1.0236E+02	1.0245E+02	1.0231E+02	1.0506E+02	1.0656E+02	1.0637E+02	1.0789E+02	1.0705E+02	1.0811E+02	1.0802E+02
	Std	5.6768E-01	6.0766E-01	5.9866E-01	7.7855E-01	1.0794E+00	7.2278E+01	9.5667E-01	9.5577E-01	1.0966E+00	1.1267E+00
	<i>t</i> -test	/	+	+	+	+	+	+	+	+	+
	W-test	/	+	+	+	+	+	+	+	+	+
F13	Best	2.2312E+01	2.4443E+01	2.3863E+01	2.8761E+01	3.5366E+01	7.8985E+01	9.4667E+01	8.8224E+01	1.0057E+02	1.0284E+02
	Mean	2.7671E+01	3.0379E+01	2.9298E+01	3.5227E+01	4.2773E+01	9.7566E+01	1.1358E+02	1.0644E+02	1.0988E+02	1.1158E+02
	Std	1.8339E+00	2.0227E+00	2.0032E+00	2.6922E+00	2.9042E+00	4.6632E+00	6.0557E+00	5.4663E+00	6.2007E+00	6.1198E+00
	<i>t</i> -test	/	+	+	+	+	+	+	+	+	+
	W-test	/	+	+	+	+	+	+	+	+	+
F14	Best	1.7966E+02	2.6786E+02	1.8055E+02	3.3716E+02	3.2273E+02	3.1454E+04	3.1487E+04	3.1455E+04	3.1687E+04	3.2443E+04
	Mean	2.8976E+03	3.2177E+03	2.8374E+03	3.1812E+03	4.0237E+03	3.3342E+04	3.3698E+04	3.4679E+04	3.4736E+04	3.4644E+04
	Std	9.6739E+02	1.4467E+03	9.8062E+02	1.0127E+03	1.5636E+03	9.8996E+02	1.0356E+03	1.1433E+03	1.1687E+03	1.1976E+03

	<i>t</i> -test	/	+	0	0	+	+	+	+	+	+
	W-test	/	+	0	+	+	+	+	+	+	+
F15	Best	1.0006E+02	1.00107E+02	1.0052E+02	1.0303E+02	1.0478E+02	1.0232E+02	1.0878E+02	1.0559E+02	1.1247E+02	1.1598E+02
	Mean	1.0366E+02	1.0639E+02	1.0509E+02	1.0896E+02	1.1122E+02	1.0668E+02	1.2565E+02	1.1655E+02	1.3557E+02	1.3783E+02
	Std	1.7956E+00	3.4337E+00	2.5983E+00	3.8924E+00	4.0933E+00	1.6997E+00	2.8778E+00	2.3377E+00	4.4688E+00	3.9965E+00
	<i>t</i> -test	/	+	+	+	+	+	+	+	+	+
	W-test	/	+	+	+	+	+	+	+	+	+
Total	<i>t</i> -test	/	15/0/0	13/2/0	14/1/0	15/0/0	/	15/0/0	14/1/0	15/0/0	15/0/0
	W-test	/	15/0/0	14/1/0	15/0/0	15/0/0	/	15/0/0	15/0/0	15/0/0	15/0/0

从表 4-2 可以看出, CEBA 在 15 个测试函数的低维和高维测试中取得了比 BA、PSO、QPSO 和 GWO 更好的优化结果。在低维运算中, t 检验统计表明, CEBA 算法分别有 15、13、14 和 15 个函数优于 BA、PSO、QPSO 和 GWO 算法。在高维运算中, CEBA 算法分别有 15、14、15 和 15 个函数优于 BA、PSO、QPSO 和 GWO 算法。W 检验结果表明, 在低维运算中, CEBA 算法分别有 15、14、15 和 15 个函数优于 BA、PSO、QPSO 和 GWO 算法。在高维运算中, 与 BA、PSO、QPSO 和 GWO 算法相比, CEBA 算法优于 15 个函数。通过混沌映射产生初始化种群, 动态扰动种群, 平均更新种群位置, 提高了 CEBA 算法的全局寻优能力。

4.4 本章小结

本章首先分析了传统蝙蝠算法工作原理和算法存在的不足, 并进行改进。为了克服标准 BA 早熟性, 在标准 BA 中引入了混沌映射方法。针对蝙蝠算法迭代后期种群变异差和容易陷入局部最优解的问题, 在标准 BA 中引入了动态扰动系数、切线随机优化和平均总体位置等方法, 以提高全局搜索能力, 避免陷入局部最优状态。再利用 CET15 测试集进行测试 BA 和 CEBA 的优化性能, 实验结果表明改进后 CEBA 算法的收敛速度和寻优能力均有效提升。

第5章 基于 OVMD-PSR-CEBA-LSSVM 的风速预测

风速数据具有非线性强、波动性大且伴随着时间上的间歇性，单一 LSSVM 无法获得较高的风速预测精度，需要信号分解技术对风速数据进行预处理，降低原始风速数据的波动性和非线性；此外，核函数类型和核参数对 LSSVM 回归性能影响很大，通过改进型 CEBA 算法对 LSSVM 进行最优核函数及其参数优化选择，构建风速预测组合模型 OVMD-PSR-CEBA-LSSVM(VPBL)，实现核函数选择，相空间参数和核参数同步优化的风速短期预测。

5.1 构建 OVMD-PSR-CEBA-LSSVM 模型

本文采用“分解-重构”的策略构建风速预测模型，即对风速数据样本进行分解成多个不同子序列、各子序列预测建模、各子序列预测结果重构获得最终的风速预测结果。本文所构建的风速预测模型结构图如图 5-1 所示。

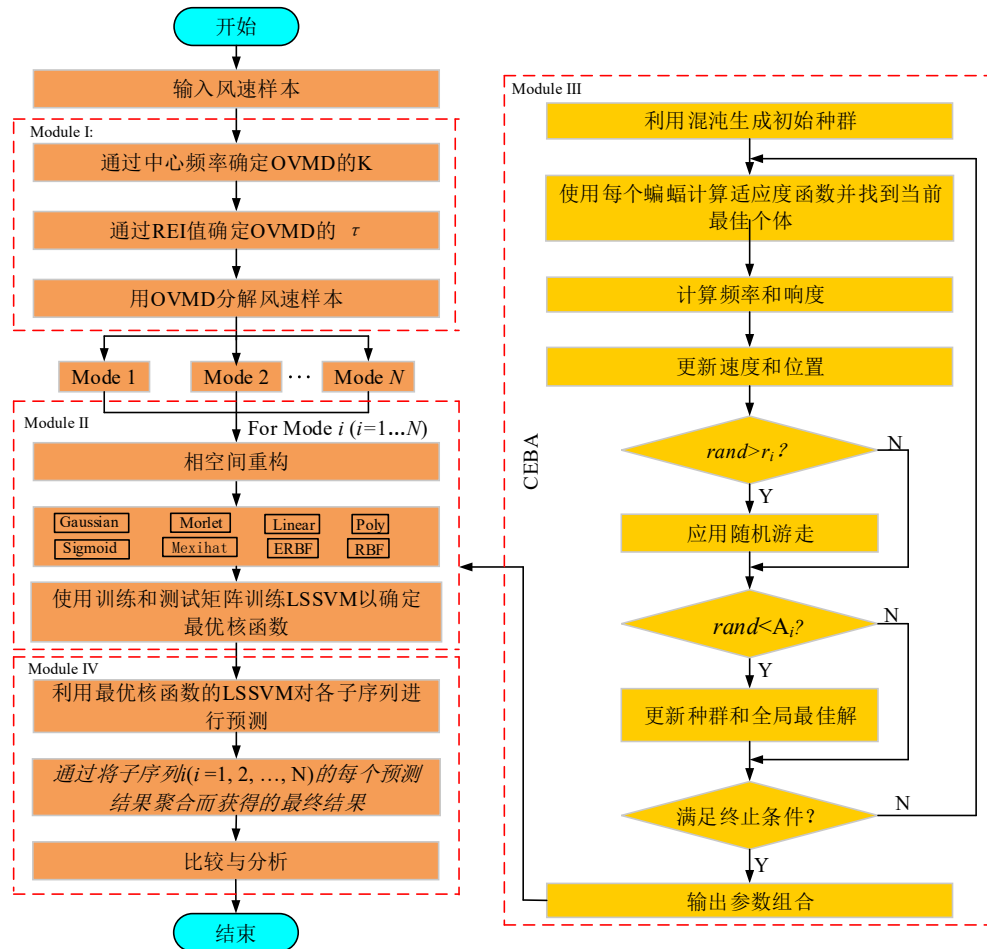


图 5-1 VPBL 模型结构图

步骤 1：首先将风速数据样本分为训练样本、测试样本和验证样本，其中，

训练样本训练预测模型；测试样本用于优化选择分解子序列的 LSSVM 最优核函数；验证样本用于验证新建预测模型的有效性；

步骤 2：利用 OVMD 方法将原始风速时间序列分解为相对稳定性好、波动性较小的分解模态子序列。然后，应用 PSR 方法将每个分解子序列构造成其对应的特征输入矩阵并进行线性归一化，作为 LSSVM 的输入变量进行风速预测，其中线性归一化如式(5-1)所示：

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (5-1)$$

式中 x 是给定风速序列， $x_{\min}$ 和 $x_{\max}$ 分别为给定风速序列中的最小值和最大值。

步骤 3：考虑到不同的风速分解子序列具有不同的非线性特性，LSSVM 基于单一核函数很难适用于所有不同特性的风速分解子序列，采用核函数选择策略并利用 CEBA 优化算法从 LSSVM 的典型核函数集中选择最佳核函数及其参数，该核函数集包括 RBF、ERBF、Poly、Sigmoid、Gaussian、Morlet 和 Linear 核函数。对每个分解模态子序列的 LSSVM 选择最优核函数，无论是分解模态子序列还是整个预测模型的预测精度都能得到有效的提高。

步骤 4：对传统蝙蝠算法存在早熟性收敛的不足进行改进，将混沌算法、动态扰动系数、正切随机优化机制和平均种群位置法引入蝙蝠算法中，提高种群的初始遍历性，调整蝙蝠算法的个体飞行策略使个体进入更广阔的搜索空间，平均位置法避免种群个体聚集的可能性，实现算法全局最优；利用改进的 CEBA 对 LSSVM 的核函数选择及其参数以及 PSR 的参数进行同步优化；

步骤 5：对各分解子序列分量的预测结果进行累积获得最终的风速预测结果，并对预测结果进行反归一化；最后，利用统计指标评估预测模型有效性。

5.2 预测实验结果及对比分析

在本节中，执行了四种评估方法，包括混沌增强蝙蝠算法的验证、所构建的 VPBL 与现有预测模型的比较以及 Diebold-Mariano (DM) 检验，以进一步分析所构建的 VPBL 的预测性能。

5.2.1 数据来源

内蒙古地区风能资源非常丰富，在全区 118.3 万平方公里的土地上，风能总储量达到 8.98 万千瓦，风能技术可开发利用量为 1.5 亿千瓦，占全国可利

用风能储量的 40%。内蒙古风速的季节变化和日变化基本上与生产和生活用电规律相吻合，且地域辽阔，人口稀少，大部分地区为平坦的草场，十分适宜建设大型风电场。本章选取了国家气象中心记录的内蒙古某风电场历史风速数据进行测试新建风速预测模型和其它预测模型，并用统计指标对所得实验结果进行比较和分析，验证新建风速预测模型的有效性。

选取连续 10 天风速数据(10 分钟间隔采样)，共计 1440 个风速数据，前 1000 个风速数据用于训练预测模型；随后 296 个风速数据用作测试数据，用于选取 LSSVM 合适的核函数；最后 144 个风速数据用于评估所新建的风速预测模型，风速数据如图 5-2 所示。表 5-1 列出了经验风速时间序列的统计描述。由图 5-2 和表 5-1 可知，风速时间序列表现出非常强的随机波动性。

表 5-1 经验样本的统计描述 (m/s)

风速时间序列		最小值	均值	最大值	标准差
样本 A	所有样本 (1-1440)	4.24	8.95	15.15	2.03
	训练样本 (1-1000)	4.24	8.71	15.15	2.11
	测试样本 (1001-1296)	5.64	9.78	13.51	1.60
	验证样本(1297-1440)	4.79	8.93	11.84	1.86
样本 B	所有样本 (1-1440)	3.01	6.64	12.67	1.54
	训练样本 (1-1000)	3.12	6.58	10.6	1.47
	测试样本 (1001-1296)	3.01	6.20	10.6	1.39
	验证样本 (1297-1440)	5.40	8.02	12.67	1.62

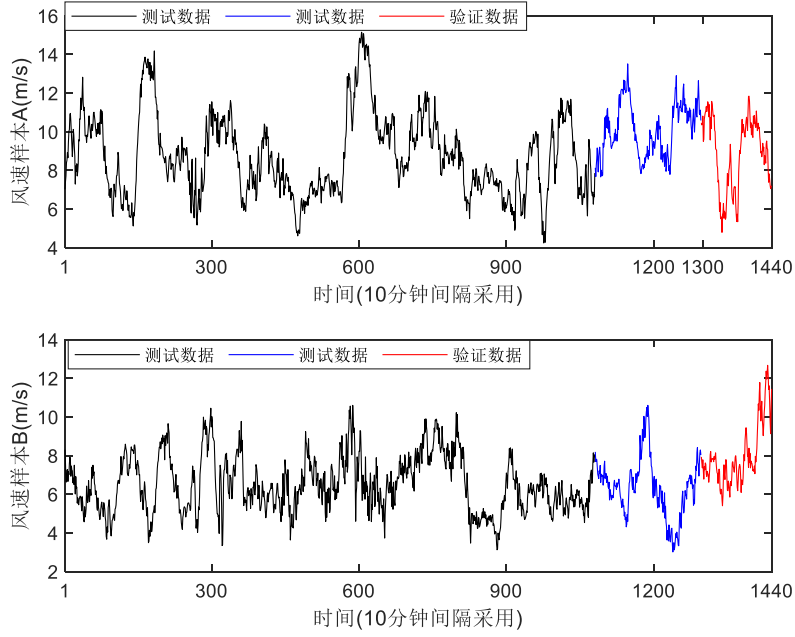


图 5-2 原始经验风速时间序列

5.2.2 评价指标

为了评估所构建的基于改进的 CEBA 同步优化方法的 OVMD-PSR-LSSVM 的预测性能，构建了几种单一和组合预测方法，使用统计指标进行比较和分析，包括均方根误差(RMSE)、平均绝对百分比误差(MAPE)、平均绝对误差(MAE)用来衡量预测值和实际值之间的误差，而皮尔逊相关系数(R)，用来衡量预测值与实际值之间的线性相关强度，分别表示如式(5-2)、(5-3)、(5-4)和(5-5)所示：

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (v(i) - v'(i))^2} \quad (5-2)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |v(i) - v'(i)| \quad (5-3)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{|v(i) - v'(i)|}{v(i)} \quad (5-4)$$

$$R = \frac{(\frac{1}{N}) \sum_{i=1}^N (v(i) - \bar{v}(i))(v'(i) - \bar{v}'(i))}{\sqrt{(\frac{1}{N}) \sum_{i=1}^N (v(i) - \bar{v}(i))^2} \times \sqrt{(\frac{1}{N}) \sum_{i=1}^N (v'(i) - \bar{v}'(i))^2}} \quad (5-5)$$

式中， N 表示风速的数量， $v(i)$ 、 $\bar{v}(i)$ 、 $v'(i)$ 和 $\bar{v}'(i)$ 分别是风速实测值、实测值均值、风速预测值和预测值均值。

5.2.3 风速预处理结果

为了削弱风速时间序列的非线性和非平稳性，利用 OVMD 将经验风速样本

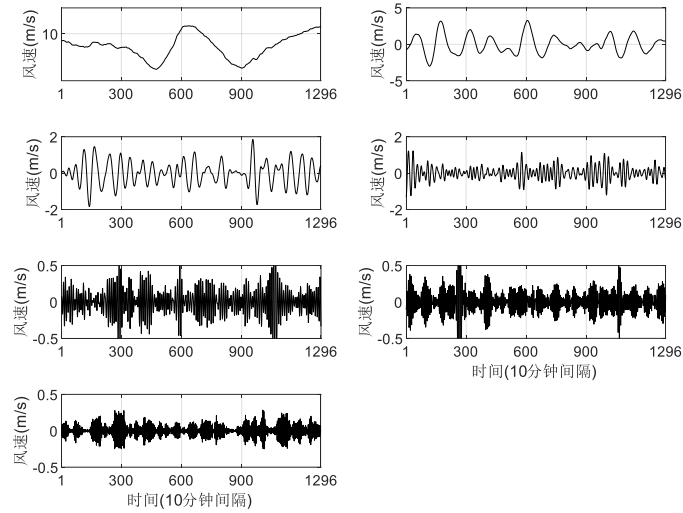
分解为多个模态，以便通过具有最优核函数的 LSSVM 进行更好的预测。为获得最优分解结果，采用 CFO 方法计算各分解模态的中心频率，计算不同 K 值下各分解模式的中心频率 ω_k ，一旦出现相似频率，此时的 K 值被确定为最佳 K 值，不同 K 值下各模态的中心频率如表 5-2 所示。

表 5-2 风速样本不同分解数 K 下的中心频率

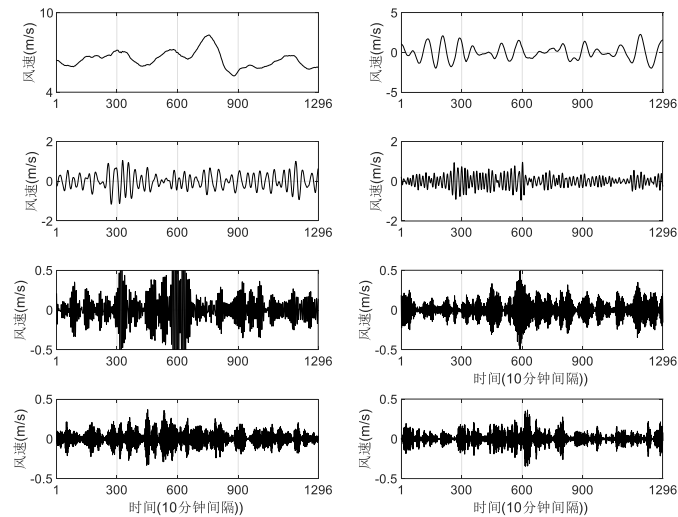
K	模态 1	模态 2	模态 3	模态 4	模态 5	模态 6	模态 7	模态 8	模态 9	模态 10
A	3	0.3379	45.5672	1267.551						
	4	0.3208	44.657	485.5443	1189.423					
	5	0.3117	43.425	411.145	740.245	1112.572				
	6	0.2835	42.786	398.137	711.556	791.398	1167.718			
	7	0.223	40.989	291.432	694.547	695.688	904.433	1210.274		
	8	0.179	40.272	280.897	651.588	651.899	786.622	1137.209	1298.286	
	9	0.177	40.776	79.049	598.669	598.733	724.282	837.555	994.347	1277.328
	3	0.386	47.898	1373.673						
	4	0.357	46.559	584.775	1119.812					
B	5	0.336	44.782	491.473	812.742	1212.337				
	6	0.312	43.988	437.437	734.542	821.774	1180.660			
	7	0.288	42.713	334.277	691.732	733.365	934.344	1194.675		
	8	0.223	42.446	289.225	594.227	673.347	807.427	948.287	1235.68	
	9	0.219	42.232	287.896	584.235	626.683	795.762	946.441	1089.236	1188.097
	10	0.192	41.122	89.242	548.224	597.584	728.257	838.686	904.274	1023.082
										1228.355

从表 5-2 可以看出，对于样本 A，当分解数 K 为 8 时，模态 2 和模态 3 的中心频率相似，这意味着风速时间序列样本被过度分解，因此，分解数 K 被设置为 7。对于样本 B，通过相同的方法将风速数据样本 B 的总分解数 K 确定为 8。OVMD 可以有效地提取代表原始风速时间序列变化趋势的主导分量，由于分解后的分量相对非线性和波动性小，因此可以大大提高模型的预测精度。

另外，将惩罚因子 α 、Tau、Tol、DC 分别设为 2000、0、1e-6、0。因此，经验风速时间序列的分解结果如图 5-3 所示。



(a) 风速数据样本 A 的 OVMD 分解结果



(b) 风速数据样本 B 的 OVMD 分解结果

图 5-3 基于 OVMD 的风速分解结果

OVMD 可以有效地提取代表原始风速时间序列变化趋势的主导分量。由于分解后的分量相对稳定，因此可以有效地提高模型的预测精度。

不同的风速预处理方法，包括小波变换分解(WT)和经验模态分解(EMD)，分别集成到所新建的模型中，用于评估 OVMD 的分解性能。在风速预处理阶段，分别应用 OVMD、WT 和 EMD 将风速样本分解为不同的子序列，然后利用分解后的子序列分别进行预测。这些预测模型与所构建的模型之间的唯一区别是应用了不同的风速预处理方法，其测试结果如表 5-2 所示。

表 5-2 基于不同分解方法的 VPBL 预测结果

样本		A			B		
模型	指标	VPBL(WT)	VPBL(EMD)	OVMD	VPBL(WT)	VPBL(EMD)	OVMD
一步	RMSE	0.3181	0.3637	0.2436	0.2979	0.2885	0.223
	MAPE	3.2008	3.8168	2.4902	3.2649	3.2626	2.3023
	MAE	0.2707	0.3188	0.2101	0.2586	0.2503	0.1942
	R	0.9864	0.9813	0.9916	0.9836	0.9846	0.9931
两步	RMSE	0.3212	0.3787	0.2627	0.3109	0.291	0.2546
	MAPE	3.3578	4.0171	2.7056	3.2012	3.2853	2.5946
	MAE	0.2797	0.3343	0.2315	0.2685	0.251	0.2208
	R	0.9854	0.9798	0.9907	0.9864	0.9837	0.9808
三步	RMSE	0.3378	0.4063	0.2998	0.3274	0.3105	0.2719
	MAPE	3.4414	4.1755	3.0154	3.3694	3.5797	2.7999
	MAE	0.2864	0.3509	0.2587	0.2857	0.2759	0.2329
	R	0.984	0.9769	0.9871	0.9848	0.9828	0.9895

从表中可以看出，可以得出以下一些结论：

与其他预处理方法的分解效果相比，OVMD 在风速分解方面具有更好的能力。在 1 步、2 步和 3 步预测中，基于 OVMD 的预测模型的 RMSE 值分别为 0.2436 m/s、0.2627 m/s 和 0.2998 m/s，与 1 步、2 步和 3 步预测中基于 WT 的预测模型相比分别降低了 0.0745 m/s、0.0585m/s 和 0.038 m/s，与 1 步、2 步和 3 步预测中基于 EMD 的预测模型相比分别降低了 0.1201m/s、0.116 m/s 和 0.1065m/s。对于风速样本 B 也可以得到类似的结论。风速时间序列的随机波动非常明显，信号预处理方法有助于克服这种波动。所构建的与 OVMD 相结合的模型优于所构建的具有信号预处理方法 WT 和 EMD 的模型。此外，采用 WT 的模型的预测精度与采用 EMD 的模型相似。由此可见，OVMD 可以较好地消除噪声成分对回归性能的负面影响，从而提高 LSSVM 的预测能力，信号预处理方法在风速预测模型中是不可或缺的。

5.2.4 实验结果对比

在本节中，在一步、两步和三步范围内进行了四个比较实验。在实验一中，进行了 LSSVM 的最优核函数选择策略的验证；在实验二中，用 VPBL 其他单一预测模型进行比较；实验三中，使用不同算法对 VPBL 优化效果进行比较；实验四中，用 VPBL 和其他组合模型的预测结果对比；

（1）实验一（基于不同核函数的 LSSVM）

在实验中，构建了具有不同核函数的 VPBL，为每个分解的子序列选择最佳的核函数，预测结果如表 5-3 所示。从表 5-3 可以看出，对于不同的风速样本和预测步长，预测模型的预测性能是不同的。对于风速样本 A，在一步预测中预测精度从 1 到 8 的预测模型是 Poly、Gaussian、Sigmoid、Morlet、RBF、ERBF、Mexihat、和 Linear，在两步预测中预测精度从 1 到 8 的预测模型是 Poly、Gaussian、Sigmoid、Morlet、RBF、ERBF、Mexihat、和 Linear，在三步预测中预测精度从 1 到 8 的预测模型是 Poly、Gaussian、Sigmoid、Morlet、RBF、ERBF、Mexihat、和 Linear，此外，采用滚动预测策略来训练和验证样本，以优化核函数的选择。

表 5-3 基于不同核函数的 VPBL 的预测结果

样本	模型	一步				两步				三步			
		RMSE	MAPE	MAE	R	RMSE	MAPE	MAE	R	RMSE	MAPE	MAE	R
A	VPBL(RBF)	0.3137	3.1579	0.2671	0.9859	0.3236	3.3728	0.28	0.9851	0.3267	3.3541	0.2878	0.9851
	VPBL(Poly)	0.2809	2.846	0.2422	0.9885	0.3101	3.1754	0.2674	0.9867	0.315	3.2724	0.2723	0.9858
	VPBL(ERBF)	0.317	3.171	0.2698	0.9855	0.3325	3.3811	0.287	0.9845	0.3445	3.6018	0.3038	0.9835
	VPBL(Mexihat)	0.3172	3.3489	0.2797	0.9861	0.3246	3.3695	0.2863	0.9852	0.3597	3.6184	0.3105	0.9816
	VPBL(Gaussian)	0.3081	3.1626	0.2706	0.9864	0.3248	3.2607	0.2763	0.986	0.3445	3.6018	0.3038	0.9835
	VPBL(Morlet)	0.3116	3.178	0.2696	0.9864	0.327	3.3661	0.2865	0.9846	0.3361	3.4582	0.2958	0.9838
	VPBL(Sigmoid)	0.3091	3.1304	0.2625	0.9864	0.311	3.2252	0.27	0.987	0.3569	3.7023	0.3122	0.9815
	VPBL(Linear)	0.3208	3.2931	0.2826	0.9856	0.3274	3.3348	0.2834	0.9855	0.3597	3.6184	0.3105	0.9816
	Proposal	0.2436	2.4902	0.2101	0.9916	0.2627	2.7056	0.2315	0.9907	0.2998	3.0154	0.2587	0.9871
B	VPBL(RBF)	0.2978	3.2293	0.252	0.9834	0.3566	3.6896	0.3089	0.9822	0.3641	3.7464	0.3205	0.981
	VPBL(Poly)	0.3157	3.4591	0.2706	0.9818	0.3427	3.7122	0.2869	0.9774	0.4267	4.945	0.385	0.9669
	VPBL(ERBF)	0.31	3.444	0.2674	0.9818	0.3528	3.9334	0.3067	0.9764	0.3697	4.0757	0.3148	0.9756
	VPBL(Mexihat)	0.3156	3.4791	0.2675	0.9812	0.353	3.8786	0.3055	0.9776	0.393	4.319	0.3391	0.9718
	VPBL(Gaussian)	0.3281	3.589	0.2797	0.9807	0.3691	3.8128	0.3203	0.9822	0.3736	4.1601	0.3227	0.9749
	VPBL(Morlet)	0.3134	3.48	0.2689	0.9814	0.3294	3.7014	0.2839	0.9797	0.3783	4.2249	0.3302	0.9727
	VPBL(Sigmoid)	0.3303	3.7294	0.2915	0.9788	0.3437	3.8368	0.297	0.9781	0.3641	4.0278	0.3127	0.9752
	VPBL(Linear)	0.3519	3.5268	0.2978	0.9843	0.3731	3.7222	0.3197	0.9818	0.4534	4.6106	0.3913	0.9718
	Proposal	0.223	2.3023	0.1942	0.9931	0.2546	2.5946	0.2208	0.9808	0.2719	2.7999	0.2329	0.9895

通过最小预测指标 CEI 确定两组风速样本不同分解子序列在 LSSVM 中核函数的最优选择，如表 5-4 所示。从表 5-4 可以看出，风速样本 A 被分解成 7 组子序列，并且每个子序列的最优核函数不同，子序列 IMF1~IMF7 的频率从低到高逐渐增加。对于 IMF1，无论是单步预测还是多步预测水平，LSSVM 的最优核函数都是线性函数。对于其他 IMF，最优核函数随预测步长而变化。在最优预测结果中，通过改进的 CEBA 算法对每个分解子序列的预测结果进行聚合，得到基于核函数选择策略的最终预测结果。当执行最优核函数选择策略时，预测精度都得到了提高。

表 5-4 每个分解子序列的最优核函数

样本	IMF	一步	两步	三步	样本	IMF	一步	两步	三步
A	1	Linear	Linear	Poly	B	1	Linear	Linear	Linear
	2	Poly	Poly	RBF		2	RBF	ERBF	Poly
	3	RBF	Morlet	ERBF		3	RBF	Morlet	Gaussian
	4	Gaussian	ERBF	Poly		4	Poly	RBF	RBF
	5	RBF	RBF	RBF		5	ERBF	Gaussian	Poly
	6	ERBF	RBF	Poly		6	RBF	RBF	RBF
	7	RBF	RBF	Poly		7	RBF	Gaussian	Mexihat
						8	Poly	RBF	RBF

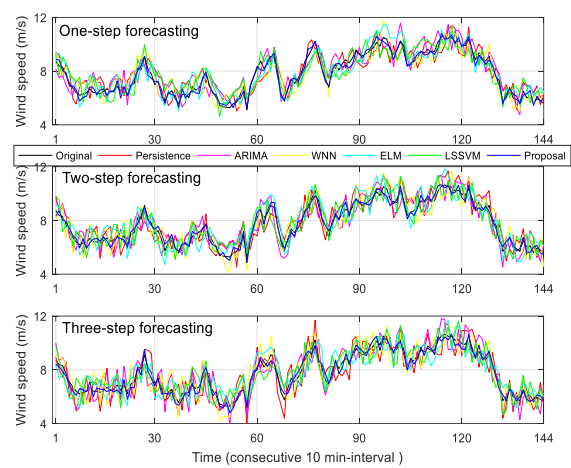
预测结果表明,线性核函数的 LSSVM 在低频信号预测方面表现出显著优势,而在预测高频子序列时,其他核函数在捕捉风速样本的非线性信号方面比线性核函数有更好的预测效果。所构建的 LSSVM 核函数选择策略可以根据最小 CEI 值自适应地确定每个分解子序列的最优核函数,从而显著提高 LSSVM 的预测性能。

(2) 实验二（所构建的模型与其他单一预测模型比较）

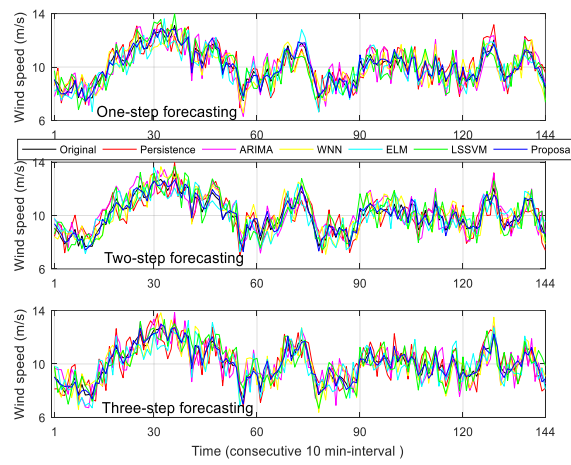
在本节中,五种单一个体预测方法,包括人工智能方法 ELM、WNN、LSSVM 以及统计方法 Persistence 和 ARIMA。Persistence 是一种众所周知的统计方法,通常被用作评估新开发的预测模型的基准方法^[54],该模型也被用作本实验中的基准方法。单一预测方法的预测结果以及预测误差分别如表 5-5、图 5-4 和图 5-5 所示。

表 5-5 单一方法和组合模型预测结果

样本	模型	一步				两步				三步			
		RMSE	MAPE	MAE	R	RMSE	MAPE	MAE	R	RMSE	MAPE	MAE	R
A	Persistence	0.9076	9.1109	0.7722	0.9001	0.922	9.5643	0.8066	0.885	0.941	9.6281	0.8154	0.8952
	ARIMA	0.819	8.3425	0.7044	0.9103	0.8801	9.092	0.7638	0.9142	0.9088	9.4377	0.8012	0.897
	WNN	0.7842	8.0691	0.6902	0.9165	0.8737	8.9376	0.7536	0.9097	0.8741	8.9771	0.7547	0.9073
	ELM	0.8039	8.4042	0.7089	0.9173	0.8898	8.8761	0.7625	0.9119	0.8986	8.9458	0.7652	0.8946
	LSSVM	0.7774	8.1573	0.696	0.9237	0.8391	8.41	0.7101	0.9254	0.8832	8.7987	0.7577	0.909
	Proposal	0.2436	2.4902	0.2101	0.9916	0.2627	2.7056	0.2315	0.9907	0.2998	3.0154	0.2587	0.9871
B	Persistence	0.7562	7.781	0.6656	0.9219	0.7747	7.78	0.6565	0.92	0.7915	8.22	0.6979	0.9192
	ARIMA	0.73565	7.4956	0.6377	0.9295	0.7446	7.7347	0.6436	0.9269	0.7672	7.9108	0.6628	0.9213
	WNN	0.7057	7.2022	0.6133	0.9359	0.7328	7.4291	0.6328	0.9287	0.7565	7.7745	0.6522	0.9247
	ELM	0.7061	7.1406	0.611	0.9343	0.7213	7.1466	0.6107	0.9308	0.7281	7.2356	0.6253	0.9302
	LSSVM	0.6562	6.5599	0.5594	0.9391	0.6869	7.1142	0.6028	0.9381	0.7395	7.4901	0.6295	0.9286
	Proposal	0.223	2.3023	0.1942	0.9931	0.2546	2.5946	0.2208	0.9808	0.2719	2.7999	0.2329	0.9895

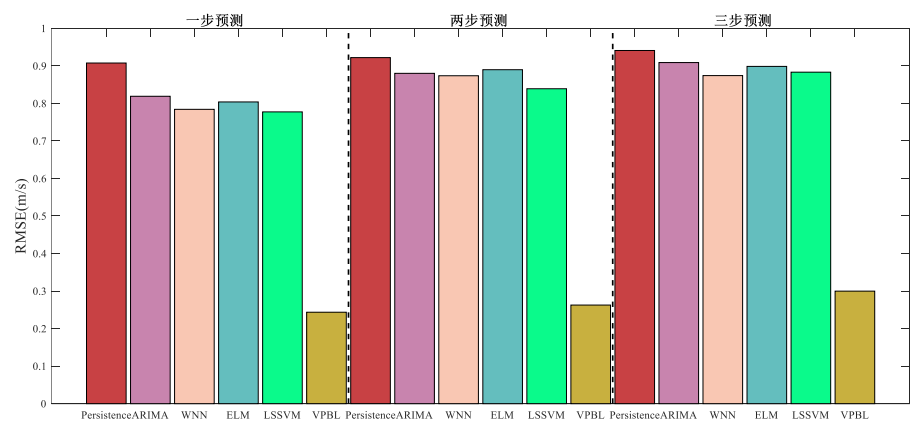


(a) 不同模型对样本 A 的预测结果

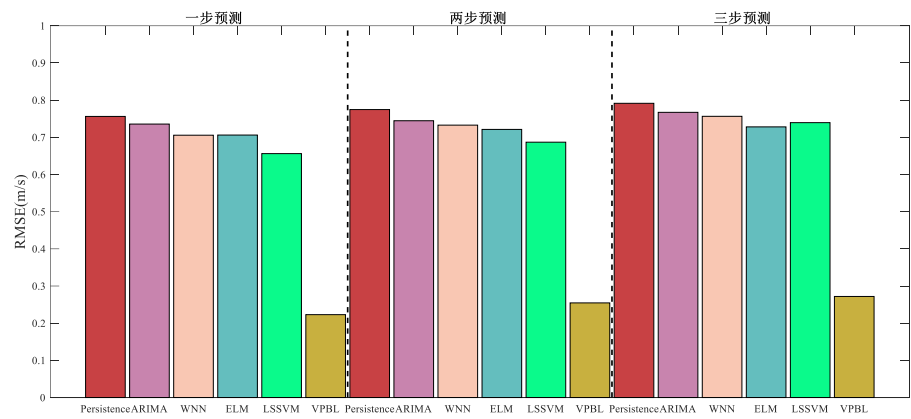


(b) 不同模型对样本 B 的预测结果

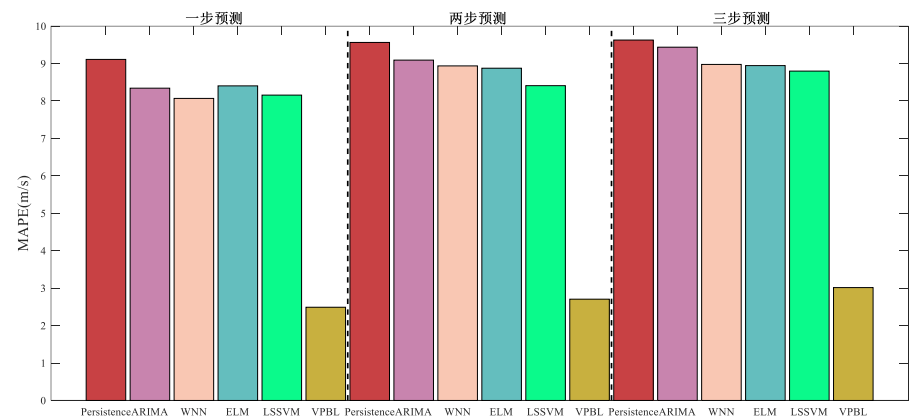
图 5-4 不同模型的预测结果



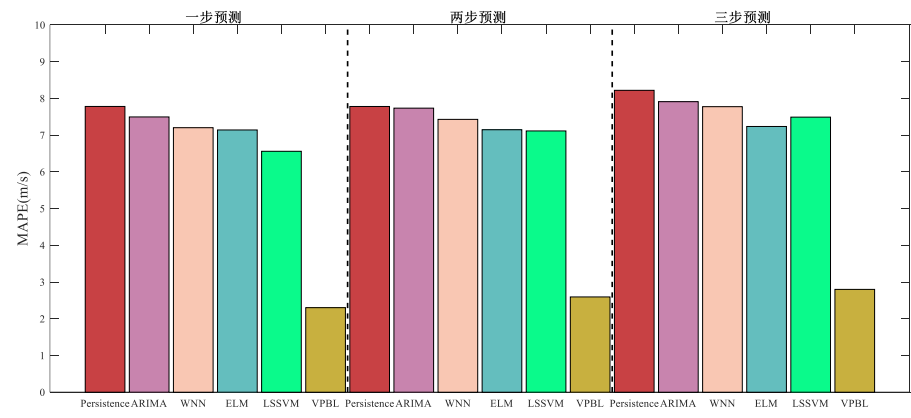
(a) 风速数据 A 的预测误差 RMSE



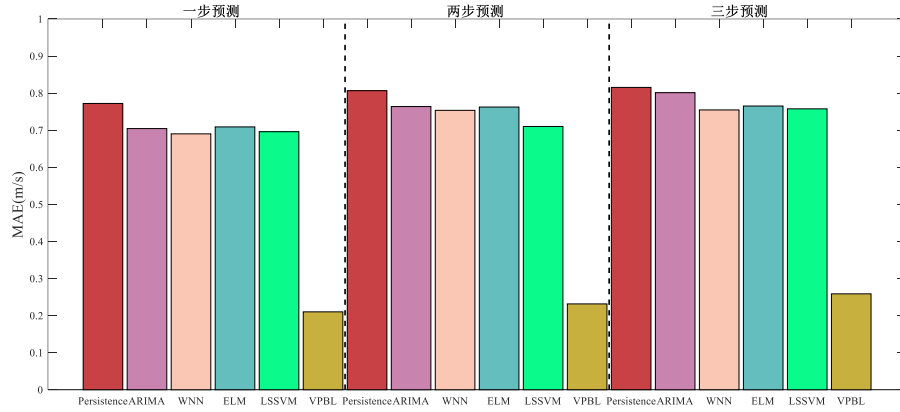
(b) 风速数据 B 的预测误差 RMSE



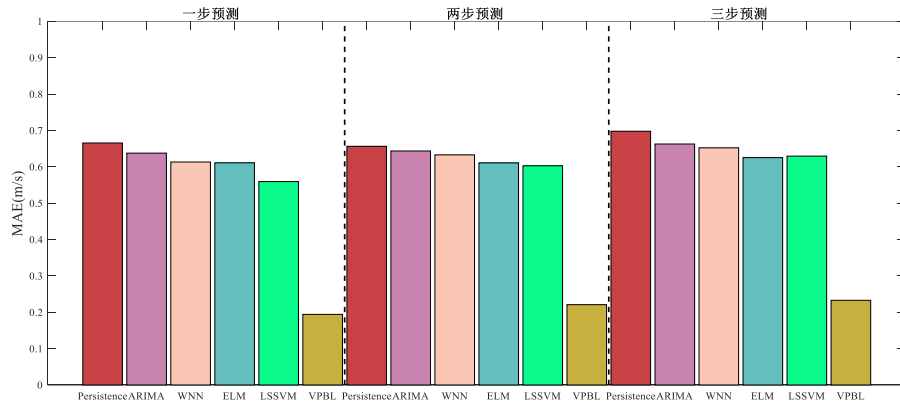
(c) 风速数据 A 的预测误差 MAPE



(d) 风速数据 B 的预测误差 MAPE



(e) 风速数据 A 的预测误差 MAE



(f) 风速数据 B 的预测误差 MAE

图 5-5 不同模型的预测误差

从表 5-5 中列出的统计误差，通过对实验结果可以得到：对于风速样本 A，（1）与统计方法 Persistence 和 ARIMA 相比，人工智能方法 WNN、ELM 和 LSSVM 获得了更好的预测精度。（2）预测精度最高到预测精度最低的单一方法依次为 LSSVM、WNN、ELM、ARIMA 和 Persistence。（3）在预测模型中，对于 1 步、2 步和 3 步预测，所构建的预测模型获得了最高的预测精度，最小的 RMSE 值分别为 0.2436 m/s、0.2627 m/s 和 0.2998 m/s。从风速样本 B 的预测结果中可以得出相同的结论。

造成这一结果的原因是，单一预测模型直接利用原始风速样本进行风速预测，而基于信号分解的组合预测模型首先将原始风速时间序列分解为几个相对稳定的子序列，然后利用子序列进行风速预测，可以有效降低预测模型的预测难度。人工智能方法 WNN、ELM 和 LSSVM 在风速预测中优于统计方法 Persistence 和 ARIMA，因为人工智能方法可以更好地捕捉隐藏在历史风速数据中的非线性信

息。此外，不仅对于风速数据 A，而且对于风速数据 B，预测精度肯定会随着预测步长的增加而降低。与其他没有信号预处理方法的模型相比，基于信号预处理方法的风速预测模型无论是单步预测还是多步预测，预测精度都有有效提高。

根据式(5-6)计算构建的 VPBL 相对于其他比较预测模型的改进百分比^[66]。结果如表 5-6 所示：

$$P_{index} = \frac{|index_1 - index_2|}{index_1} \times 100\% \quad (5-6)$$

式中 $index_1$ 和 $index_2$ 分别代表比较预测模型和建议的 VPBL 的统计误差指数。

表 5-6 构建的 VPBL 模型较其他比较模型的改善百分比 (%)

样本	A					B				
比较模型	Persistence	ARIMA	WNN	ELM	LSSVM	Persistence	ARIMA	WNN	ELM	LSSVM
一步	P_{RMSE}	73.16	70.26	68.66	69.7	68.94	70.51	69.69	68.4	68.42
	P_{MAPE}	72.67	70.15	69.47	70.37	69.14	70.41	69.28	68.03	67.76
	P_{MAE}	72.79	70.17	69.81	70.36	69.56	70.82	69.55	68.34	68.22
	P_R	10.17	8.93	7.35	8.1	8.19	7.72	6.84	6.11	6.29
两步	P_{RMSE}	71.51	70.15	68.69	70.48	69.93	67.14	65.81	65.26	64.7
	P_{MAPE}	71.71	70.24	67.83	69.52	69.73	66.65	66.46	65.08	63.69
	P_{MAE}	71.3	69.69	67.4	69.64	69.28	66.37	65.69	65.11	63.84
	P_R	11.94	8.37	7.06	8.64	8.9	6.61	5.82	5.61	5.37
三步	P_{RMSE}	68.14	67.01	66.06	66.64	65.7	65.65	64.56	64.06	62.66
	P_{MAPE}	68.69	68.06	65.74	66.29	66.42	65.94	64.61	63.99	61.3
	P_{MAE}	68.27	67.71	65.86	66.19	65.72	66.63	64.86	64.29	62.75
	P_R	10.27	10.04	8.59	10.34	8.8	7.65	7.4	7.01	6.37

从表 5-6 可以看出，VPBL 的预测性能优于基准方法 Persistence 和其他单独的预测方法。例如，在样本 A 的一步、两步和三步预测中，所构建的 VPBL 的 P_{RMSE} 值比 Persistence 分别降低 73.16%、71.51%和 68.14%，相应的 P_{MAPE} 值分别为 72.67%、71.71%和 68.69%。在样本 B 的一步、两步和三步预测中，所构建的 VPBL 的 P_{RMSE} 值比 Persistence 分别降低 70.51%、67.14%和 65.65%，相应的 P_{MAPE} 值分别为 70.41%、66.65%和 65.94%。上述比较和分析表明，构建的 VPBL 比基准法和其他单独的预测方法更有效。

(3) 实验三（不同算法对 VPBL 优化效果比较）

采用 PSO、GA、BA 和 CEBA 优化算法优化选取 VPBL 的参数，用于多步风速预测。所构建的 VPBL 模型经不同算法优化后的预测结果如表 5-7 所示。

表 5-7 不同算法对 VPBL 的预测结果

样本	模型	A				B			
		VPBL(BA)	VPBL(GA)	VPBL(PSO)	VPBL(CEBA)	VPBL(BA)	VPBL(GA)	VPBL(PSO)	VPBL(CEBA)
一步	RMSE	0.3381	0.3383	0.3317	0.2436	0.2584	0.2544	0.2636	0.223
	MAPE	3.3219	3.4659	3.3513	2.4902	2.946	2.8808	2.9471	2.3023
	MAE	0.288	0.2943	0.2865	0.2101	0.224	0.2197	0.2296	0.1942
	R	0.9839	0.9842	0.9857	0.9916	0.9874	0.9881	0.9869	0.9931
两步	RMSE	0.37	0.3768	0.3624	0.2627	0.2898	0.2888	0.2929	0.2546
	MAPE	3.7265	3.7515	3.7412	2.7056	3.2403	3.1694	2.9555	2.5946
	MAE	0.3163	0.3209	0.3156	0.2315	0.2487	0.2459	0.2551	0.2208
	R	0.9814	0.9796	0.9823	0.9907	0.9846	0.9855	0.988	0.9808
三步	RMSE	0.3948	0.3895	0.393	0.2998	0.3231	0.3094	0.316	0.2719
	MAPE	3.9912	4.0163	4.0789	3.0154	3.5166	3.5819	3.5021	2.7999
	MAE	0.343	0.3365	0.3264	0.2587	0.2766	0.2757	0.2745	0.2329
	R	0.9797	0.9783	0.9784	0.9871	0.9815	0.9824	0.9828	0.9895

从表 5-7 中可知：

(1) 从风速样本 A 的 1 步预测到 3 步预测，与基于 PSO 的 VPBL 相比，所构建的 VPBL 的 RMSE 值分别降低了 0.0881 m/s、0.997 m/s 和 0.0932m/s。与基于 BA 的 VPBL 相比，RMSE 值分别降低了 0.0945 m/s、0.1073 m/s 和 0.095m/s。与基于 GA 的 VPBL 相比，RMSE 值分别降低了 0.0947 m/s、0.1141m/s 和 0.897 m/s。

(2) 对于风速样本 B，与基于 PSO 的 VPBL 相比，在 1 步、2 步和 3 步预测下，所构建的 VPBL 的 RMSE 值分别降低了 0.0406 m/s、0.0383m/s 和 0.0441 m/s。与基于 BA 的 VPBL 相比，RMSE 值分别降低了 0.0354 m/s、0.0352 m/s 和 0.0512m/s。与基于 GA 的 VPBL 相比，RMSE 值分别降低了 0.0314 m/s、0.0342 m/s 和 0.0375m/s。

基于 CEBA 算法的 VPBL 在统计指标方面优于基于 BA、PSO 或 GA 的 VPBL，这表明改进的 CEBA 算法能够优化风速预测中的 VPBL 参数，降低欠拟合或过拟合的概率，提高 VPBL 的预测性能。

(4) 实验四（VPBL 和其他组合模型的结果对比）

与文献[62]、[63]和 [64]中预测模型 WT-GA-SVM、EEMD-GA-BPNN 和 EWT-CSA-LSSVM 相比较，评估和分析所构建 VPBL 模型的有效性。例如，在参考文献[62]中，Liu 采用 WT 将风速数据分解成不同子序列，其次，利用 SVM 对不同子序列进行预测，并利用 GA 算法对 SVM 参数优化选取，最后将不同子

序列预测结果进行积累得到最终的风速预测结果；在参考文献[63]中，Wang 等人提出了一种基于 EEMD、GA 和 BPNN 的短期风速预测组合模型，在组合模型中，EEMD 用于将训练风速样本分解为多个模态函数(IMFs)和一个残差分量，然后将分解后的分量作为人工神经网络的输入变量，通过遗传算法优化人工神经网络参数，通过聚合每个分解分量的预测结果获得最终预测结果，预测结果验证了所构建的模型预测性能优于 WNN-GA-BPNN 模型。在文献[64]中，Hu 采用 EWT 将风速数据分解成不同类型子序列，然后，利用 LSSVM 对不同子序列预测结果进行积累得到最终的风速预测结果，不同的预测模型预测结果如表 5-8 所示。

表 5-8 VPBL 和其他组合模型的预测结果

样本		A				B			
模型		WT-GA-SVM	EEMD-GA-BP	EWT-CSA-LSSVM	VPBL	WT-GA-SVM	EEMD-GA-BP	EWT-CSA-LSSVM	VPBL
一步	RMSE	0.4489	0.4638	0.3987	0.2436	0.3246	0.3509	0.2896	0.223
	MAPE	4.5658	4.7196	4.1757	2.4902	3.6356	4.0655	3.0192	2.3023
	MAE	0.3857	0.4036	0.3499	0.2101	0.2834	0.317	0.2536	0.1942
	R	0.9722	0.9721	0.9783	0.9916	0.9802	0.9771	0.9878	0.9931
两步	RMSE	0.4961	0.497	0.4485	0.2627	0.3911	0.3828	0.3262	0.2546
	MAPE	5.0267	5.1299	4.6707	2.7056	4.4064	4.2887	3.4361	2.5946
	MAE	0.4273	0.4303	0.3924	0.2315	0.341	0.3367	0.288	0.2208
	R	0.9647	0.9656	0.9718	0.9907	0.9725	0.9728	0.9849	0.9808
三步	RMSE	0.5213	0.5338	0.4898	0.2998	0.4467	0.4382	0.4029	0.2719
	MAPE	5.4346	5.6227	5.0718	3.0154	5.029	4.8644	4.1647	2.7999
	MAE	0.4565	0.4713	0.4311	0.2587	0.3896	0.3739	0.3528	0.2329
	R	0.9607	0.9609	0.9678	0.9871	0.9651	0.9652	0.9782	0.9895

与组合模型 WT-GA-SVM、EEMD-GA-BP、EWT-CSA-LSSVM 相比，所构建的 VPBL 模型有更高的预测精度。例如，在样本 A 中，EWT-CSA-LSSVM 在 1 步、2 步和 3 步预测时的 RMSE 误差分别为 0.3987m/s、0.4485 m/s 和 0.4898 m/s，大于所构建的预测模型的 RMSE 误差。对于风速样本 A，与 WT-GA-SVM、EEMD-GA-BP 和 EWT-CSA-LSSVM 相比具有更好的预测效果。与 WT-GA-SVM、EEMD-GA-BP 和 EWT-CSA-LSSVM 相比，该模型在 1 步预测时的 RMSE 误差分别减少了 0.2053 m/s、0.2202 m/s 和 0.1551 m/s，在 2 步预测时的 RMSE 误差分别减少了 0.2334 m/s、0.2343 m/s 和 0.1858 m/s，在 3 步预测时的 RMSE 误差分别减少了 0.2215 m/s、0.234 m/s 和 0.19 m/s。

对于风速样本 B，与 WT-GA-SVM、EEMD-GA-BP 和 EWT-CSA-LSSVM 相比，该模型对 1 步预测的 RMSE 误差分别降低了 0.0706 m/s、0.0969 m/s 和 0.0666 m/s，对 2 步预测的 RMSE 误差分别降低了 0.1041 m/s、0.0958 m/s 和 0.0716 m/s，对 3 步预测的 RMSE 误差分别降低了 0.1465 m/s、0.138 m/s 和 0.131 m/s。因此，

与其他用于单步预测或多步预测的比较预测模型相比,所构建的模型获得了最佳的预测性能。

该模型优于其他比较预测模型,因为 OVMD 中的分解参数采用中心频率法调整,PSR 中的参数、核参数和核函数的选择采用改进的 CEBA 同步优化,可以有效地提高 VPBL 的预测性能。

5.2.5 假设检验

DM 检验是一种典型的假设检验方法,是一种统计推断方法,不同于传统的实验数据分析^[65]。DM 检验针对时间序列数据的特性,通过引入自相关校正的方差估计方法,有效修正了损失差异序列的标准误计算,从而使其在动态时间序列预测场景中表现出更强的适用性。DM 测试在数学上如式(5-7)^[65]所示:

$$DM = \frac{[\sum_{i=1}^n \frac{(L(e_1(t)) - L(e_2(t)))}{n}] \times S^2}{\sqrt{\frac{S^2}{n}}} \quad (5-7)$$

式中 $L(e_i(t))$ 表示预测误差的损失函数。 $e_1(t)$ 和 $e_2(t)$ 分别是所构建的 VPBL 和比较模型的预测误差。 S^2 代表 $L(e_1(t)) - L(e_2(t))$ 的估计方差。

通过将绝对 DM 值与正态分布 $N(0,1)$ 的临界值 $Z_{\alpha/2}$ 进行比较,其中 α 代表显著性水平,如果 DM 值不在临界值 $[-Z_{\alpha/2}, Z_{\alpha/2}]$ 的范围内,则 H_0 将被拒绝,这意味着所构建的新风速预测方法 VPBL 的预测结果与比较模型之间存在很大差异。对于风速预测结果, H_0 和 H_1 可以在数学上如式(5-8)所示:

$$\begin{cases} H_0 : E(L(e_1(t))) = E(L(e_2(t))) \\ H_1 : E(L(e_1(t))) \neq E(L(e_2(t))) \end{cases} \quad (5-8)$$

式中 H_0 为原假设,代表预测误差无显著差异,而作为替代假设的参数 H_1 与 H_0 相反。

这一假设的验证为模型选择提供了严格的理论依据,尤其适用于多步预测或存在序列依赖性的复杂预测任务。表 5-9 列出了其它比较模型和构建的 VPBL 方法之间的 DM 测试结果的绝对值。从表 5-9 可以看出,对于风速样本 A 的一步预测结果,比较模型 VPBL(RBF)和构建的 VPBL 之间的绝对 DM 统计值为 3.2487,并且是最小的,并且大于值 $Z_{0.01/2} = 2.58$,这意味着所有的绝对 DM 统计值都不落在 $[-Z_{0.01/2}, Z_{0.01/2}]$ 的范围内,因此, H_0 将以 99% 的概率被拒绝。表 5-9 中列出的绝对 DM 统计值表明,所构建的模型 VPBL 在风速预测方面优于其他

比较模型和基准方法，DM 测试结果进一步验证了所构建的 VPBL 可以更好地增强风速预测性能。

表 5-9 DM 检测结果的绝对值

样本	模型	一步	两步	三步	样本	模型	一步	两步	三步
A	Persistence	11.9562	11.8929	12.3649	B	Persistence	12.1557	12.3837	12.5992
	ARIMA	11.1801	11.4731	11.6264		ARIMA	11.4839	11.3775	11.7114
	LSSVM	8.1227	9.1562	8.8117		LSSVM	8.3185	8.9844	9.1728
	ELM	8.7988	9.8701	8.8943		ELM	8.5497	8.4779	8.6841
	WNN	9.3112	8.9472	9.5481		WNN	8.9411	9.21881	9.1541
	VPBL(WT)	5.7147	5.5356	5.2672		VPBL(WT)	5.4811	5.6191	5.5144
	VPBL(EMD)	5.0844	5.6227	5.2101		VPBL(EMD)	5.4566	5.4799	5.6749
	VPBL(VMD)	4.9371	5.2219	5.3183		VPBL(VMD)	5.1512	5.5817	5.2482
	VPBL(BA)	5.1873	4.9378	5.3225		VPBL(BA)	4.8477	5.2117	5.0854
	VPBL(GA)	5.3218	5.2689	4.9844		VPBL(GA)	5.3129	5.0281	5.1158
	VPBL(PSO)	5.0115	5.2175	5.1148		VPBL(PSO)	5.1244	4.9814	4.8487
	VPBL(RBF)	3.4159	3.6216	3.5931		VPBL(RBF)	3.2487	3.1594	3.1898
	VPBL(Poly)	3.5148	3.7198	3.4897		VPBL(Poly)	3.5465	3.6147	3.4159
	VPBL(ERBF)	3.7538	3.8843	3.9534		VPBL(ERBF)	3.7758	3.5844	3.6547
	VPBL(Mexihat)	3.8958	3.6584	3.8931		VPBL(Mexihat)	3.6879	3.7899	3.6945
	VPBL(Morlet)	4.5418	3.9438	4.1584		VPBL(Morlet)	4.3227	4.2581	4.4149
	VPBL(Sigmoid)	4.2719	4.3328	4.3872		VPBL(Sigmoid)	4.5414	4.3215	4.2577
	VPBL(Linear)	5.5418	5.3571	5.2587		VPBL(Linear)	5.3747	5.5477	5.6841
	WT-GA-SVM	4.9115	5.1259	5.2487		WT-GA-SVM	4.5449	4.8896	5.0150
	EEMD-GA-BP	4.6584	4.5876	4.7147		EEMD-GA-BP	4.4843	4.7441	4.3491
	EWT-CSA-LSSVM	4.3218	4.1481	4.3589		EWT-CSA-LSSVM	4.3989	4.2071	4.4112

5.3 本章小结

为了验证本文所新建的风速预测组合模型的预测性能，利用内蒙古某个风电场的实测风速数据进行短期风速预测测试。在训练预测模型之前，原始风速数据通过 OVMD 分解成频率更低且波动幅度更小的多个子序列，并采用相空间重构技术对分解子序列从重构成 LSSVM 的输入变量矩阵，然后，利用 LSSVM 对重建的矩阵变量进行模型训练和预测。通过实验比较和分析以及 DM 测试，可以得出：

- （1）与单一风速预测模型相比，风速预测组合模型集成风速预处理技术、参数优化选取和风速预测方法，能够有效提高风速预测性能；
- （2）中心频率观测用于选择 OVMD 的最佳分解层数，风速数据得到了更好的分解；通过实验对比，其分析性能优于信号分解方法 WT 和 EMD；
- （3）采用混沌初始化种群、动态扰动系数、切线随机优化机制来提高蝙蝠算法的整体优化性能，使蝙蝠算法优化性能得到了有效提高；
- （4）与基于单一核函数的 LSSVM 预测方法相比，采用每个子序列最优核函数

的 VPBL 可以获得更好的风速预测效果；

(5) 与最近开发的预测模型 EWT-CSA-LSSVM^[62]、WT-GA-SVM^[63]和 EEMD-GA-BPNN^[64]相比，VPBL 获得更高的预测精度。

因此，所构建的 VPBL 是一种有效的风速预测模型。

第 6 章 结论与展望

6.1 研究结论

风速时间序列具有较强的波动性,给风力发电带来了较大的挑战,同时风电并网也影响了电力系统的安全性和可靠性,因此,构建有效的风速预测模型,提高风速预测精度,可以为风电场的稳定运行提供帮助。本文构建了一种基于 OVMD、PSR 和 LSSVM 的组合模型,并通过改进的 CEBA 进行核函数选择和参数优化。在训练过程之前,对原始风速样本进行 OVMD 分解,并用 PSR 方法重构分解子序列输入矩阵,然后将重构变量作为 LSSVM 的输入进行模型训练和预测。从前文的比较、分析和 DM 测试中,可以得出以下结论。

(1) 通过 CFO 方法确定了 OVMD 的适当分解数,OVMD 优于其他信号分解方法 WT 和 EMD,通过 OVMD 分解方法对原始风速数据进行处理,然后通过相空间重构得到处理后的风速数据,增强了风速预测精度。

(2) CEBA 是标准 BA 的改进版本,利用混沌初始化种群、动态扰动系数、正切随机优化机制和平均种群位置方法来提高 BA 的整体预测性能。通过对标准 CEC15 函数集的测试,以及高低维 CEC15 标准测试集运行结果,可以看出优化算法 CEBA 获得了比 BA、PSO、QPSO 和 GWO 更好的优化性能。

(3) 与基于不同核函数的 OVMD、PSR、CEBA 和 LSSVM 组合预测模型相比,所构建的 VPBL 利用了每个分解子序列的最优核函数,获得更精确的风速预测结果。

(4) 所构建的 VPBL 优于最近开发的预测模型 EWT-CSA-LSSVM^[62]、WT-GA-SVM^[63]和 EEMD-GA-BPNN^[64]。

6.2 工作展望

本文所构建的 VPBL 是一种有效的风速预测模型,通过风电场实测风速数据验证了构建的 VPBL 的有效性,可以拓宽现有的类似研究,然而,这项研究仍然存在一些局限性。今后的研究方向和重点为:

(1) 本文只从风速数据一个方面进行风速的预测研究,今后还要考虑环境因素的影响,将温度、湿度、风向等变量加入到预测模型中,以此来获得更精确的风速预测结果。

(2) 本文只选取了内蒙古某个风电场的风速数据，今后要收集多地风电场的风速数据，以此对风速预测模型进行更加全面的分析与比较。

参考文献

- [1] 司金冬, 吴熙, 郭其胜, 等. 面向高比例新能源消纳的地区电网柔性互联规划与运行技术综述[J]. 电网技术, 2024, 48(2): 1-18.
- [2] 杜效鹄, 周兴波, 周建平. 碳中和背景下我国电力碳排放水平分析[J]. 水力发电学报, 2023, 42(12): 1-11.
- [3] 占萌, 张亚耀, 马锐. 以新能源为主体的新型电力系统模型体系和同步稳定研究综述[J]. 北京师范大学学报(自然科学版), 2023, 59(06): 878-886.
- [4] 胡玮, 闵浩, 张国政, 等. 大规模风力发电对电力系统影响的分析[J]. 仪器仪表用户, 2025, 32(01): 139-141.
- [5] 洪烽, 贾欣怡, 梁璐, 等. 面向风电场频率支撑的混合储能层次化容量优化配置[J]. 中国电机工程学报, 2023, 43: 1-12.
- [6] 洪国庆, 吴国旻, 金宇清, 等. 电力系统风力发电建模与仿真研究综述[J]. 电力系统自动化, 2024, 48(17): 22-36.
- [7] 汪正军, 高静方, 赵冰, 等. 基于风速预测的风电场虚拟惯量协调控制技术[J]. 太阳能学报, 2022, 43(10): 138-143.
- [8] 彭天. 计及多时段趋势感知的风电集群功率超短期预测[D]. 东北电力大学, 2023. 000109.
- [9] Giorgi D G M, Ficarella A, Tarantino M. Assessment of the benefits of numerical weather predictions in wind power forecasting based on statistical methods[J]. Energy, 2011, 36(7): 3968-3978.
- [10] 关永锋, 喻敏, 胡佳. 基于自适应组合模型的超短期风速预测[J]. 电力系统保护与控制, 2022, 50(04): 210446.
- [11] 贺秋萍, 郭娜先, 阿成凤, 等. 短期风速预测模型研究进展综述[J]. 青海科技, 2023, 30(06): 140-143.
- [12] 张泰豪, 殷锋, 袁平. 风速及风功率预测方法综述[J]. 现代计算机, 2021, (08): 45-48.
- [13] 李建林, 张海军, 庞俊强. 基于神经网络和模式识别的中长期风速及发电量预测[J]. 环境技术, 2021, 39(05): 219-226.
- [14] 盛四清, 袁伟智. 基于二次分解和 BiLSTM-XGBoost-ARIMA 组合模型的短期风速预测[J/OL]. 华北电力大学学报(自然科学版), 2025, 01, 25(04): 1-13.
- [15] 李乾鑫, 张胜杰, 于天暝. 基于模糊认知图和粒计算的长期风速区间预测研究[J]. 吉林化工学院学报, 2023, 40(07): 71-76.
- [16] 韩子娇. 风—光—氢耦合系统协调控制与运行经济性研究[D]. 沈阳工业大学, 2024.
- [17] Sankhala K D, Rani I S, Srinivas D, et al. Validation of OceanSat-3 sea surface winds for their utilization in the NCMRWF NWP assimilation systems[J]. Advances in Space Research, 2025, 75(2): 1945-1959.
- [18] 闫琦. 考虑多位置 NWP 特征挖掘的风电集群功率短期预测[D]. 东北电力大学, 2023.
- [19] 刘世林. 基于 NWP 风速修正和融合模型的短期风电功率预测[D]. 西安理工大学, 2022..
- [20] 王斌, 魏成伟, 谢丽蓉, 等. 基于风速误差校正和 ALO-LSSVM 的风电功率预测[J]. 太阳能学报, 2022, 43(01): 58-63.
- [21] 蒋哲勇. 基于统计分析的确定性和概率性多时间尺度风速预测研究[D]. 南昌工程学院, 2023.
- [22] 胡锐, 乔加飞, 李永华, 等. 基于 WOA-VMD-SSA-LSTM 的中长期风电预测[J]. 太阳能学报, 2024, 45(09): 549-556.

- [23] 叶瑞丽, 郭志忠, 刘瑞叶, 等. 基于小波包分解和改进 Elman 神经网络的风电场风速和风电功率预测[J]. 电工技术学报, 2017, 32(21): 103-111.
- [24] 程启明, 陈路, 程尹曼, 等. 基于 EEMD 和 LS-SVM 模型的风电功率短期预测方法[J]. 电力自动化设备, 2018, 38(05): 27-35.
- [25] 唐非. 基于互补集成经验模态分解和 Stacking 融合的短期风速组合预测模型[J]. 太阳能学报, 2024, 45(07): 735-744.
- [26] He W, Yan L, Min X, et al. A combined wind speed prediction model based on data processing, multi-objective optimization and machine learning[J]. Energy Reports, 2023, 9 (S8): 413-421.
- [27] 张建平, 于新建, 陈栋, 等. 随机算法改进的 RCSA-ANN 模型及近海短期风速预测[J]. 空气动力学学报, 2022, 40(04): 110-116.
- [28] 毕贵红, 赵鑫, 李璐, 等. 双模式分解 CNN-LSTM 集成的短期风速预测模型[J]. 太阳能学报, 2023, 44(03): 191-197.
- [29] 尹元亚, 潘文虎, 赵文广, 等. 基于 CEEMDAN 和 BiLSTM-AM 的超短期风速预测方法[J]. 电测与仪表, 2024, 61(09): 77-84.
- [30] 周会友. 基于支持向量机的风电场风速预测方法研究[D]. 华北电力大学(北京), 2017.
- [31] Zhang Y G, Chen B, Pan G F, Zhao Y. A novel hybrid model based on VMD-WT and PCA-BP-RBF neural network for short-term wind speed forecasting[J]. Energy Conversion and Management, 2019, 9(195): 180-97.
- [32] 向玲, 刘佳宁, 苏浩, 等. 基于 CEEMDAN 二次分解和 LSTM 的风速多步预测研究[J]. 太阳能学报, 2022, 43(08): 334-339.
- [33] 张琰妮, 史加荣, 李津, 等. 融合残差与 VMD-ELM-LSTM 的短期风速预测[J]. 太阳能学报, 2023, 44(09): 340-347.
- [34] 柳玉, 曾德良, 刘吉臻, 等. 基于小波包变换的最小二乘支持向量机短期风速多步预测和信息粒化预测的研究[J]. 太阳能学报, 2014, 35(02): 214-220.
- [35] 王富强, 王东风, 韩璞. 基于混沌相空间重构与支持向量机的风速预测[J]. 太阳能学报, 2012, 33(08): 1321-1326.
- [36] 张建平, 于新建, 陈栋, 等. 随机算法改进的 RCSA-ANN 模型及近海短期风速预测[J]. 空气动力学学报, 2022, 40(04): 110-116.
- [37] Bo G, Keke L, Hong T Z, et al. Short-term forecasting and uncertainty analysis of wind power[J]. Journal of Solar Energy Engineering-Transactions of the ASME, 2021, 143(5): 054503.
- [38] Yang J X. A novel short-term multi-input-multi-output prediction model of wind speed and wind power with LSSVM based on improved ant colony algorithm optimization[J]. Cluster Computing-The Journal of Networks Software Tools and Applications, 2019, 22(2): S3293-S3300.
- [39] 付文龙, 章轩瑞, 张海荣, 等. 基于 INGO-SWGMN 混合模型的超短期风速预测研究[J]. 太阳能学报, 2024: 1-11.
- [40] Sun S Z, Wang Y, Meng Y, et al. Multi-step wind speed forecasting model using a compound forecasting architecture and an improved QPSO-based synchronous optimization[J]. Energy Reports, 2022, 8: 9899-9918.
- [41] Niu T, Wang J Z, Zhang K Q. Multi-step-ahead wind speed forecasting based on optimal feature selection and a modified bat algorithm with the cognition strategy[J]. Renewable Energy, 2018, 118: 213-229.
- [42] 黄宇, 张冰哲, 庞慧珍, 等. 基于混合 Copula 优化算法的风速预测方法研究[J]. 太阳

- 能学报, 2022, 43(10): 192-201.
- [43] 王逸文, 王维莉, 刘贤超, 等. 融合两阶段分解与 iJaya-ELM 的短期风速预测模型[J]. 电子测量与仪器学报, 2023, 37(07): 186-195.
- [44] 谢东良, 鄧伦海, 周康, 等. 基于 VMD-ORELM-EC 的超短期风速组合预测模型[J]. 合肥工业大学学报(自然科学版), 2024, 47(05): 703-711.
- [45] 高晟扬, 李法社. 基于 CEEMD-SE-PSR-BP 的短期风速预测[J/OL]. 太阳能学报, 2025, 01, 25: 1-7.
- [46] 田丽, 凤志民, 刘世林. 基于 CEEMD-PSR-FOA-LSSVM 的短期风电功率预测[J]. 可再生能源, 2016, 34(11): 1632-1638.
- [47] Fu W, Fang P, Wang K, et al. Multi-step ahead short-term wind speed forecasting approach coupling variational mode decomposition, improved beetle antennae search algorithm-based synchronous optimization and Volterra series model. Renewable Energy, 2021, 179, 1122-1139.
- [48] Wang J, Hu J, A robust combination approach for short-term wind speed forecasting and analysis: Combination of the ARIMA (Autoregressive Integrated Moving Average), ELM (Extreme Learning Machine), SVM (Support Vector Machine) and LSSVM (Least Square SVM) forecasts using a GPR (Gaussian Process Regression) model. Energy, 2015, 93, 41-56.
- [49] 王岩, 陈耀然, 韩兆龙, 等. 基于互信息理论与递归神经网络的短期风速预测模型[J]. 上海交通大学学报, 2021, 55(09): 1080-1086.
- [50] 尤晴. 基于 PSO 优化 LSSVM 的短期风速预测研究[D]. 哈尔滨工业大学, 2022. 002529.
- [51] Natarajan J Y ,Nachimuthu S D .Retraction Note: New SVM kernel soft computing models for wind speed prediction in renewable energy applications[J].Soft Computing,2024,(prepublish):1-1.
- [52] Yuan X, Chen C, Yuan Y, et al. Short-term wind power prediction based on LSSVM-GSA model. Energy Conversion and Management, 2015, 101, 393-401.
- [53] Altay EV, Alatas B. Bird swarm algorithm with chaotic mapping [J]. Artificial Intelligence Review, 2019, 49, 1-42.
- [54] Sun S, Fu J, Li A, et al. A new compound wind speed forecasting structure combining multi-kernel LSSVM with two-stage decomposition technique. Soft Computing, 2021, 25, 1479-1500.
- [55] Wei-Chiang Hong, Ming-Wei Li, Jing Geng, et al. Novel chaotic bat algorithm for forecasting complex motion offloating platforms. Applied Mathematical Modelling, 2019, 72, 425-443.
- [56] Sun S, Fu J, Wei L, et al. Multi-objective optimal dispatching for a grid-connected micro-grid considering wind power forecasting probability. IEEE Access, 2020, 8(17), 46981-46997.
- [57] Meng Y, Sun SZ, Wang Y, et al. A new compound structure combining DAWNN with modified water cycle algorithm-based synchronous optimization for wind speed forecasting. Energy Reports, 2022, 8, 12255-12271.
- [58] Liu J, Mei Y, Li X. An Analysis of the inertia weight parameter for binary particle swarm optimization. IEEE Transactions on Evolutionary Computation, 2016, 20(5), 666-681.
- [59] Hu P, Pan JS, Chu SC. Improved binary grey wolf optimizer and its application for feature selection. Knowledge-Based Systems, 2020, 195:105746.
- [60] Jordehi AR. Binary particle swarm optimization with quadratic transfer function: A new binary optimization algorithm for optimal scheduling of appliances in smart homes. Applied Soft Computing, 2019, 78:465-480.
- [61] Wang L, Ni H, Yang R, et al. An adaptive simplified human learning optimization algorithm[J]. Information Sciences, 2015, 320:126-139.

- [62] Liu D, Niu D, Wang H, et al. Short-term wind speed forecasting using wavelet transform and support vector machines optimized by genetic algorithm. *Renewable Energy*, 2014, 62, 592-597.
- [63] Wang S, Zhang N, Wu L, et al. Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and GA-BP neural network method, *Renewable. Energy*, 2016, 94, 629-636.
- [64] Hu J, Wang J, Ma K. A hybrid technique for short-term wind speed prediction. *Energy*, 2015, 563-574.
- [65] Hu, JM., Heng JN., Wen JM., et al. Deterministic and probabilistic wind speed forecasting with denoising- reconstruction strategy and quantile regression-based algorithm. *Renewable Energy*, 2020, 162, 1208-1226.
- [66] Zhang LF, Wang JZ, Niu XS, et al. Ensemble wind speed forecasting with multi-objective Archimedes optimization algorithm and sub-model selection. *Applied Energy*, 2021, 301, 117449.

攻读硕士期间取得的学术成果

Optimization of WPD-PSR-ELM Based on Improved Bat Algorithm for Wind Speed Forecasting(EI, 一作)

A Hybrid Structure Based on VMD-LSSVM and Improved Quantum Particle Swarm Simultaneous Optimization for Wind Speed Prediction(EI, 三作)

Multi-objective optimal dispatching of demand response-enabled microgrid considering uncertainty in renewable generations based on two-level iterative strategy(SCI, 三作)

致谢

行文至此，落笔为终。回首三年的硕士求学生涯，既有探索学术的充实与挑战，亦满载师长、亲友的关怀与支持。在此，谨以最诚挚的谢意，献给所有曾予我指引与力量的人。

首先，衷心感谢我的导师孙驷洲老师。感谢您在学术道路上的悉心指导，从论文选题的反复推敲、实验设计的严谨把关，到研究瓶颈时的耐心启发，您渊博的学识、严谨的治学态度与谦和的为师风范，令我受益终身。其次感谢我的父母与家人。求学二十余载，你们始终是最坚实的后盾。你们的理解、包容与无条件的支持，让我在迷茫时重拾勇气，在疲惫时找到归处。寸草之心，难报春晖，唯愿以未来的成长回报这份深情。最后我要感谢同窗侯浩天对我的重要帮助。日常进行学习与研究的时候，时刻给予我科研上的指导和意见。

感谢对本论文进行评审的各位老师，感谢你们在百忙之中抽出时间审阅论文。祝愿你们身体健康，工作顺心！