第 44 卷第 4 期　　河 北 科 技 大 学 学 报　　Vol. 44, No. 4
2023 年 8 月　　Journal of Hebei University of Science and Technology　　Aug. 2023

文章编号:1008-1542(2023)04-0323-12

基于 SSA-VMD-LSTM-NKDE 的短期风电功率概率预测

高晓芝[1],郭　旺[1,2],郭英军[1],宋静冉[1],孙鹤旭[1]

(1. 河北科技大学电气工程学院,河北石家庄　050018;2. 国网河北省电力有限公司任丘市供电分公司,河北任丘　062550)

摘　要:为进一步提高风电功率预测精度,提出一种基于麻雀搜索算法(SSA)优化 VMD 参数的组合预测方法。首先,使用麻雀搜索算法对 VMD 参数进行优化,并利用优化后的 VMD 对数据进行分解;其次,结合灰色关联分析法和熵权法对环境变量进行相关性分析,选择相关性最高的影响因素与分解得到的各模态分量组合作为 LSTM 预测模型的输入,获得更为精确的预测结果;最后,建立基于非参数核密度估计(NKDE)的风电功率概率预测模型,实现对风电功率预测结果不确定性的有效量化。结果表明,所提组合模型的 MAE,RMSE 和 MAPE 比 VMD-LSTM 模型的分别下降了 39.51%,33.22%和 40.39%。SSA-VMD-LSTM-NKDE 组合模型不仅能够有效提高确定性预测的精度,而且还能够实现对风电功率预测结果不确定性的有效量化,为风电功率预测提供了科学决策依据。

关键词:风能;麻雀搜索算法;变分模态分解;熵权法;灰色关联分析;组合预测模型

中图分类号:TM614　　**文献标识码:**A　　**DOI:**10.7535/hbkd.2023yx04001

Short-term wind power probabilistic forecasting based on SSA-VMD-LSTM-NKDE

GAO Xiaozhi[1], GUO Wang[1,2], GUO Yingjun[1], SONG Jingran[1], SUN Hexu[1]

(1. School of Electrical Engineering, Hebei University of Science and Technology, Shijiazhuang, Hebei 050018, China; 2. Renqiu Power Supply Branch, State Grid Hebei Electric Power Company Limited, Renqiu, Hebei 062550, China)

Abstract: In order to further improve the accuracy of wind power forecasting, a combined forecasting method based on sparrow search algorithm (SSA) optimizing variational mode decomposition (VMD) parameters was proposed. Firstly, the

收稿日期:2022-09-29;修回日期:2023-04-30;责任编辑:冯　民

基金项目:河北省省级科技计划(20314501D,19214501D)

第一作者简介:高晓芝(1981—),女,河北石家庄人,讲师,博士,主要从事微电网控制及电能质量改善技术方面的研究。

通信作者:孙鹤旭教授。E-mail:sun_hebust@163.com

高晓芝,郭旺,郭英军,等. 基于 SSA-VMD-LSTM-NKDE 的短期风电功率概率预测[J]. 河北科技大学学报,2023,44(4):323-334.

GAO Xiaozhi, GUO Wang, GUO Yingjun, et al. Short-term wind power probabilistic forecasting based on SSA-VMD-LSTM-NKDE[J]. Journal of Hebei University of Science and Technology, 2023, 44(4): 323-334.

SSA was used to optimize the VMD parameters, then the optimized VMD was used to decompose the data. Secondly, the entropy weight method and grey relational analysis were combined to analyze the correlation of environmental variables, and the combination of the most relevant influencing factors and the decomposed modal components were selected as the input of the LSTM prediction model to obtain more accurate prediction results. Finally, a wind power probability prediction model based on NKDE was established to effectively quantify the uncertainty of wind power prediction results. The results show that the MAE, RMSE and MAPE of the proposed combination model decrease by 39.51%, 33.22% and 40.39%, respectively, compared with the VMD-LSTM model. The SSA-VMD-LSTM-NKDE combination model can not only effectively improve the accuracy of deterministic prediction, but also effectively quantify the uncertainty of wind power prediction results, which provides scientific decision-making basis for wind power prediction.

Keywords: wind energy; sparrow search algorithm; variational mode decomposition; entropy weight method; grey relational analysis; combined prediction mode

在“双碳”目标驱动下,构建以新能源为主体的新型电力系统成为必然趋势。近年来,风能在全球电力生产结构中的占比持续增加,然而风能的不稳定性给电网的安全稳定运行带来了巨大挑战[1]。高精度的功率预测不仅能有效减少电力系统备用容量配置,提高风电机组的控制功率,缓解电网调峰压力,还能提供未来某一时刻风电功率的波动区间,为电力系统提供更全面的预测信息,以便更加灵活可靠地进行调度安排[2]。

近年来,循环神经网络(RNN)因具有记忆功能、深度表达能力,能够挖掘数据中的时序信息,在电力系统典型的时间序列预测中得到充分利用[3],作为高效循环神经网络代表的长短期记忆网络(LSTM),通过特殊化的结构设计可有效避免传统 RNN 在训练过程中出现的梯度消失以及梯度爆炸问题,从而更适用于电力系统相关数据预测[4-6]。文献[7]提出一种基于长短期记忆神经网络的风功率预测模型,结果证明长短期记忆神经网络模型的预测精度比支持向量机模型高,具有较高的应用价值。由于预测模型的精确程度与输入数据的质量息息相关[8],因此对数据进行合理的处理与分解能够提升预测模型的学习效果。国内外许多研究者从信号分解重构的角度进行研究,以降低其不平稳性,从而提高预测精度。常用的分解方法有经验模态分解(EMD)[9]、集合经验模态分解(EEMD)[10]和变分模态分解(VMD)[11]等。EMD 不仅解决了小波基选取与确定分解尺度困难的问题[12],还保留了小波多分辨的优势,可用于风速时间序列的分析。文献[13]提出一种基于 EMD 的短期风功率预测方法,利用 EMD 对历史数据进行分解降噪,从而达到提高预测精度的效果。但经验模态分解在形成固有模态函数时,会出现信息特征提取混淆的问题。EEMD 通过噪声辅助信号处理有效解决了这一问题。文献[14]利用 EEMD 将原始风速序列划分为几个固有模态函数,形成一个潜在的特征集,然后通过迭代过程自动生成更合适的子特征集,最后通过相应的机器学习模型进行预测。尽管 EEMD 改善了 EMD 分解时模态混叠现象,但在分解时添加白噪声的幅值难以确定,过低将无法抑制混叠现象,过高又会出现较多的伪分量[15]。VMD 可以将复杂序列分解为幅值频率可调的若干信号[16],克服了集合 EMD 频率特征不易分辨等问题,实现信号的准确分解,运算效率更高。文献[17]利用 VMD 非递归性的优势,对原始数据进行分解,通过与其他模型对比,表明 VMD 分解能够有效提取风功率序列的细节信息。需要指出的是,VMD 在分解前需人为设定一些参数值,为了避免取值不当带来的影响,可通过优化算法确定最优参数组合,从而提高预测精度。一些文献中已经提出了采用不同的优化算法,如粒子群优化算法(PSO)[18]、鲸鱼优化算法(WOA)[19]和麻雀优化算法(SSA)[20]来优化 VMD 参数,但研究成果大多应用在故障诊断的特征参数提取方面,在功率预测研究中应用较少。相比 PSO 容易早熟收敛,WOA 易陷入局部极值,导致精度不高等问题,SSA 搜索精度高、收敛速度快、稳定性好,而且克服了容易陷入局部最优的问题。

综上,本文拟综合考虑风速、风向和温度等多种环境因素的影响,针对在传统 VMD 算法中因参数设置不当导致信号欠分解或过分解的问题,引入 SSA 算法对 VMD 的关键参数进行优化,通过灰色关联度分析法和熵权法筛选出相关性最高的环境影响因素,提出一种基于 SSA-VMD-LSTM 的组合预测模型,并采用非参数核密度估计(NKDE)分析 SSA-VMD-LSTM 模型的预测误差分布,然后计算给定置信度下的预测区间,最后通过实际数据验证所提模型的有效性和优越性。

1　基于 SSA 优化 VMD 参数

1.1　VMD 算法原理

VMD 算法本质上是一个变分问题，能很好地处理非线性数据问题。该模型通过交替方向乘子法和迭代更新计算最优解，将复杂信号分解为多个振幅和频率可调的信号（IMF）。IMF 可用作 FM-AM 信号，其表达式为

$$u_k(t)=A_k(t)\cos[\varphi_k(t)]\ , \tag{1}$$

式中 $A_k(t)$ 是瞬时振幅函数，相位函数 $\varphi_k(t)$ 是非递减的，因此有 $\omega_k(t)=\varphi'_k(t)\geqslant 0$，定义 $\omega_k(t)$ 为 $u_k(t)$ 的瞬时频率。

利用 Hilbert 变换计算模态函数的解析信号，得到单边频谱。将分解后的信号与其中心频率混合，并将每个模态频谱转换为基频频率[21]。根据调制信号梯度方向的平方 L^2 范数对各模态的带宽进行估计，其表达式为

$$\min_{\{u_k\}\cdot\{\omega_k\}}\left\{\sum_{k=1}^{K}\left\|\partial_t\left[(\delta(t)+\frac{\mathrm{j}}{\pi t})^{*}u_k(t)\right]\mathrm{e}^{-\mathrm{j}\omega_k t}\right\|_2^2\right\}, \tag{2}$$

$$\text{s.t.}\qquad \sum_{k=1}^{K}u_k=f, \tag{3}$$

式中：$\{u_k\}$ 为分解得到的 K 个模态分量；$\{\omega_k\}$ 为各分量的中心频率；$*$ 为卷积符号；$\delta(t)$ 为狄克拉函数。

引入二次惩罚因子 α 和拉格朗日乘法算子 λ，将约束变分问题转变为非约束变分问题[22]：

$$L(\{u_k\},\{\omega_k\},\lambda)=\alpha\sum_{k=1}^{K}\|\partial_t\left[(\delta(t)+\frac{\mathrm{j}}{\pi t})u_k(t)\right]\mathrm{e}^{-\mathrm{j}\omega_k t}\|_2^2+\|f(t)-\sum_{k=1}^{K}u_k(t)\|_2^2+\langle\lambda(t),f(t)-\sum_{k=1}^{K}u_k(t)\rangle。 \tag{4}$$

利用交替方向乘子迭代算法更新 u_k^{n+1}，ω_k^{n+1} 和 λ_k^{n+1} 求解上式“鞍点”，即为约束变分模型的最优解。

1.2　麻雀搜索算法优化 VMD 参数

由 VMD 算法特点可知，模态个数 K 和二次惩罚因子 α 需在分解前进行预设，如果参数设置不当，可能会出现信号欠分解或过分解问题，从而影响预测精度。如果采取局部寻优的方式，即只考虑单个影响因素，将另一个参数设置为常数，就会忽略两者之间的交互作用。因此，本文采用 SSA 算法对 VMD 的参数进行寻优，消除人为设定参数的主观影响。SSA 是一种受麻雀觅食和反捕食行为启发的新型群体优化算法[23-24]。在觅食过程中，一部分麻雀是发现者身份，另一部分是追随者身份，二者可以根据能量储备量的多少相互转换，但所占比例不变，发现者具有较好适应度值且具备较高的储备量，其任务是寻找食物，为整个种群提供觅食方向，追随者则是通过发现者来获取食物。迭代过程中，发现者的位置更新规则如下：

$$X_{i,j}^{n+1}=\begin{cases}X_{i,j}^{n}\times\exp(\dfrac{-i}{\alpha\times\mathrm{iter}_{\max}}), & R_2<S,\\ X_{i,j}^{n}+q\times\boldsymbol{L}, & R_2\geqslant S,\end{cases} \tag{5}$$

式中：$X_{i,j}^{n}$ 是第 i 个麻雀在第 j 维的位置；n 为迭代次数；q 是服从正态分布的随机数；$\boldsymbol{L}$ 是一个元素全为 1 的矩阵；$\alpha\in(0,1]$，是一个随机数；$\mathrm{iter}_{\max}$ 是最大迭代次数；R_2 表示预警值，取值范围为[0，1]；S 表示安全值，取值范围为[0.5，1]。追随者的位置更新规则如下：

$$X_{i,j}^{n+1}=\begin{cases}q\times\exp(\dfrac{X_{\mathrm{worst}}^{n}-X_i^{n}}{i^2}), & i>\dfrac{c}{2},\\ X_p^{n+1}+|X_{i,j}^{n}-X_p^{n+1}|\times\boldsymbol{A}^{+}, & i\leqslant\dfrac{c}{2},\end{cases} \tag{6}$$

式中：X_p 是目前发现者占据的最优位置；X_{worst} 是当前全局最差位置；$\boldsymbol{A}$ 是一个满足 $\boldsymbol{A}^{+}=\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}})^{-1}$ 的矩阵；c 为麻雀个数。当意识到危险时，麻雀种群会做出反捕食行为，意识到危险的麻雀在整个种群中随机产生，其数学模型如下：

$$X_{i,j}^{n+1}=\begin{cases}X_{\mathrm{best}}^{n}+\beta\times|X_{i,j}^{n}-X_{\mathrm{best}}^{n}|, & f_i>f_g,\\ X_{i,j}^{n}+Q\left[\dfrac{|X_{i,j}^{n}-X_{\mathrm{worst}}^{n}|}{(f_i-f_\omega)+\varepsilon}\right], & f_i=f_g,\end{cases} \tag{7}$$

式中：X_{best} 是当前全局最优位置；β 是一个随机数，作为步长控制参数；ε 是避免分母为零的最小常数；$Q\in[-1,1]$，是一个随机数，表示麻雀移动的方向，也是步长控制参数；f_i 是此时麻雀个体的适应度值；f_g 是此时全局最佳的适应度值；f_{worst} 是最差的适应度值。

采用SSA对VMD参数进行优化时，应先确定适应度函数，本文以样本熵作为SSA优化的适应度函数，熵值越小，分量频率越明显，模态混叠现象越弱，其定义为

$$\mathrm{SampEn}(\mathrm{data},q,r)=\ln B^{q}(r)-\ln B^{q+1}(r)\ ,\quad(8)$$

式中：data为随机信号时间序列；q 为嵌入维度；r 为相似性容限；B 为2个向量在相似容限 r 下匹配 q 或 $(q+1)$ 个实数的概率。

SSA优化VMD参数流程如图1所示。

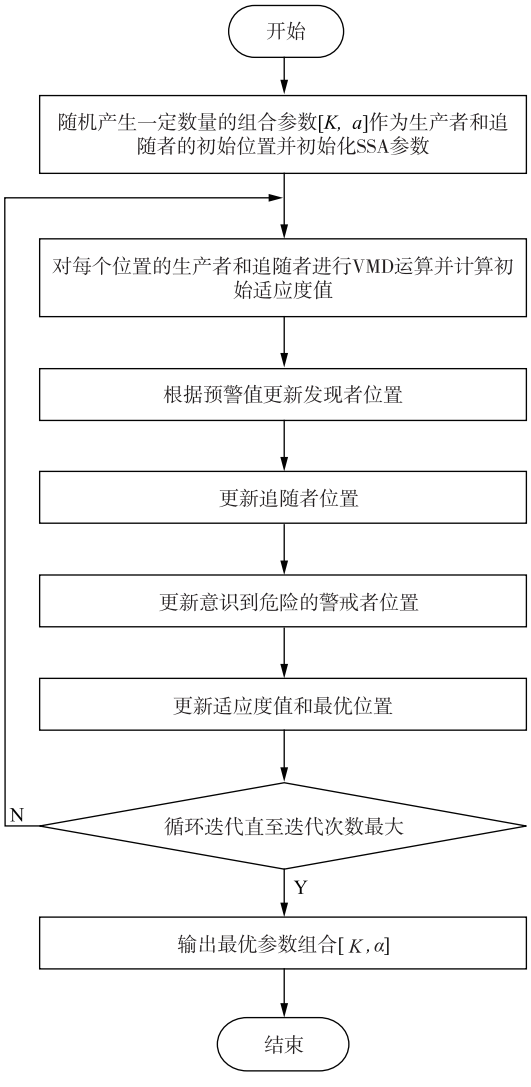

图1 麻雀搜索算法优化VMD参数流程

Fig. 1 Process of VMD parameters optimized by SSA

2 熵权-灰色关联度分析

由于使用单一的关联度分析指标不能准确刻画风电功率的规律性及相关性[25]，本文采用灰色关联度分析法和熵权法[26]组合分析不同环境变量与风电功率间的相关性，从而得出影响风电功率的主要因素和次要因素，选取主要影响因素作为预测模型的输入。

将风电功率数据作为参考序列记为 $X_0(k)=[X_0(1),X_0(2),\cdots,X_0(n)]$，环境特征数据作为比较列记为 $X_i(k)=[X_i(1),X_i(2),\cdots,X_i(n)]$。熵权-灰色关联度分析步骤如下。

步骤1 变量的无量纲化。选用均值化处理，将数量级大的序列均值归一化到1的量级附近，见式(9)。

$$X'_i=\frac{nX_i}{\sum_{k=1}^{n}X_i(k)}。\quad(9)$$

步骤2 求差序列、最大差和最小差。

$$\Delta_i(k)=|X'_0(k)-X'_i(k)|\ ,\quad(10)$$

$$\Delta_{max}=\max_i\max_k|X'_0(k)-X'_i(k)|\ ,\quad(11)$$

$$\Delta_{min}=\min_i\min_k|X'_0(k)-X'_i(k)|\ 。\quad(12)$$

步骤3 计算关联系数。ρ 为分辨系数，一般取0.5。

$$\xi_i(k)=\frac{\Delta_{min}+\rho\Delta_{max}}{\Delta_i(k)+\rho\Delta_{max}}\ 。\quad(13)$$

步骤4 计算熵权-灰色关联度。

$$R_i=\sum_{k=1}^{n}W_i\xi_i(k)\ ,\quad(14)$$

式中各评价指标熵权 $W_i=\dfrac{1-E_i}{m-\sum_{i=1}^{m}E_i}$，其中计算指标熵值 $E_i=-\dfrac{1}{\ln n}\left(\sum_{k=1}^{n}P_i\ln P_i\right)$，计算指标比重 $P_i=\dfrac{X'_i(k)}{\sum_{k=1}^{n}X'_i(k)}$。

3　基于 SSA-VMD-LSTM-NKDE 的概率预测

3.1　非参数核密度估计理论

非参数核密度估计(NKDE)是概率论中用来估计未知参数概率密度函数(PDF)的一种非参数化估计方法。根据风电场相关历史数据信息，通过核密度估计就可以得到变量分布，不需要对样本分布进行任何假设，常被用于风电或光伏预测等领域的概率预测分析。

针对一组样本数为 n 的风电功率预测误差数据 $p=\{p_1,p_2,\cdots,p_n\}$，其密度函数为 $f(p)=\dot{F}(p)$，由 NKDE 方法对 p 的分布估计为

$$f_n(p)=\frac{F(p+h)-F(p-h)}{2h}, \tag{15}$$

式中 h 为窗宽，用来决定误差数据分布的区间划分。

当样本量足够大且窗宽趋于 0 时，基于非参数核密度估计的风电功率预测误差的概率密度函数为

$$\hat{f}(p,h)=\frac{1}{nh}\sum_{i=1}^{n}K\left(\frac{p-p_i}{h}\right), \tag{16}$$

式中：p_i 为第 i 个样本；函数 $K(p,h)$ 为核函数。

在 NKDE 方法中，窗宽和核函数是 2 个关键参数。其中，窗宽 h 用来决定误差数据分布的区间划分，设置过小，易受单个样本点的影响，导致局部波动特征明显从而影响估计模型的分布情况；选择过大，会使所得的 PDF 过于平滑，导致预测误差的概率密度分布与实际情况不符。核函数的选取也会直接影响 NKDE 方法的密度曲线光滑程度及拟合程度，从而影响结果的精确度。常用于 NKDE 的核函数如表 1 所示。

表 1　常用的核函数

Tab. 1　Common kernel functions

核函数	计算式
高斯函数	$\frac{1}{\sqrt{2\pi}}\exp(-\frac{u^2}{2})$
三角函数	$(1-\lvert u\rvert)\ \lvert u\rvert\leqslant 1$
Epanechnikov	$\frac{3}{4}(1-u^2)\lvert u\rvert\leqslant 1$

3.2　概率预测步骤

由于风电功率具有波动性，为降低原始风电功率数据对预测精度产生的影响，本文采用 VMD 对原始数据进行分解；为避免在分解时出现信号欠分解或过分解现象，采用 SSA 优化 VMD 参数，从而降低预测误差。结合灰色关联度分析法和熵权法对环境变量进行相关性分析，然后将相关性最高的影响因子和 VMD 分解得到的各模态分量组合作为 LSTM 预测输入，利用 NKDE 方法对 SSA-VMD-LSTM 模型的误差进行分析，得到相应的概率密度曲线，最后综合确定性预测结果和误差分布结果，确定出某一置信区间下的预测区间。风电功率概率预测具体步骤如下：

步骤 1　对原始风电数据进行归一化处理；

步骤 2　利用 SSA 算法优化 VMD 参数，获得最优参数组合；

步骤 3　利用 SSA-VMD 将风功率序列分解为 K 个模态分量；

步骤 4　利用熵权-灰色关联度分析法对环境特征变量做相关性分析；

步骤 5　将相关性最高的影响因子与 VMD 分解得到的模态分量分别进行组合；

步骤 6　分别对步骤 5 中的组合序列建立 LSTM 预测模型，并将各序列的预测值进行整合，得到最终确定性预测值；

步骤 7　统计确定性预测误差，采用 NKDE 方法确定概率密度函数；

步骤 8　设置一定比例的置信度$(1-\alpha)$，得到该置信度下误差波动区间$[F_{\alpha/2},F_{(1-\alpha)/2}]$；

步骤 9　结合相应的确定性预测结果，构造概率预测区间。

SSA-VMD-LSTM-NKDE 模型的预测流程如图 2 所示。

4　算例分析

4.1　评价指标

平均绝对误差(MAE)，是实际值和预测值之间绝对误差的平均值，其表达式为

$$\mathrm{MAE}=\frac{1}{n}\sum_{i=1}^{n}|y_i-\hat{y}_i|\ ,\tag{17}$$

式中：n 为样本数量；y_i 为预测值；$\hat{y}_i$ 为实际值。

均方根误差（RMSE）是误差平方的期望值，用来表示预测值和实际值二者偏差的标准差，表示样本的离散程度，其表达式为

$$\mathrm{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}\ ,\tag{18}$$

式中：n 为样本数量；y_i 为预测值；$\hat{y}_i$ 为实际值。从式（18）可以看出，RMSE 存在最小值零点，从极小值向两侧随着误差的增大 RMSE 趋于无穷大。

平均绝对百分误差（MAPE），是衡量预测准确性的统计性指标，其表达式为

$$\mathrm{MAPE}=\frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{p_i-\hat{p}_i}{p_i}\right|\ ,\tag{19}$$

式中：n 为样本数量；y_i 为预测值；$\hat{y}_i$ 为实际值。从式（19）可知，MAPE 越接近 0 值，代表预测误差越小。

风电功率概率预测方法的评价指标主要包含可靠性和敏锐度 2 个方面。目前较为常见的用来评价可靠性的指标为预测区间覆盖率（PICP），其值越大表示实际风电功率点落入预测区间的个数越多，但过大的区间覆盖率不具有研究价值；用预测区间平均宽度（PINAW）来反映敏锐度，即区间预测结果的平均带宽，在同等置信水平下其值越小越好。

PICP 计算公式：

$$\mathrm{PICP}=\left(\frac{1}{N}\sum_{i=1}^{N}c_i\right)\times 100\%\ ,\tag{20}$$

式中：N 为样本总数；c_i 是区间上限 U_i 和区间下限 L_i 对实际值的可靠性指标，若真实值落入$[L_i,U_i]$内，则 $c_i=1$，反之 $c_i=0$。

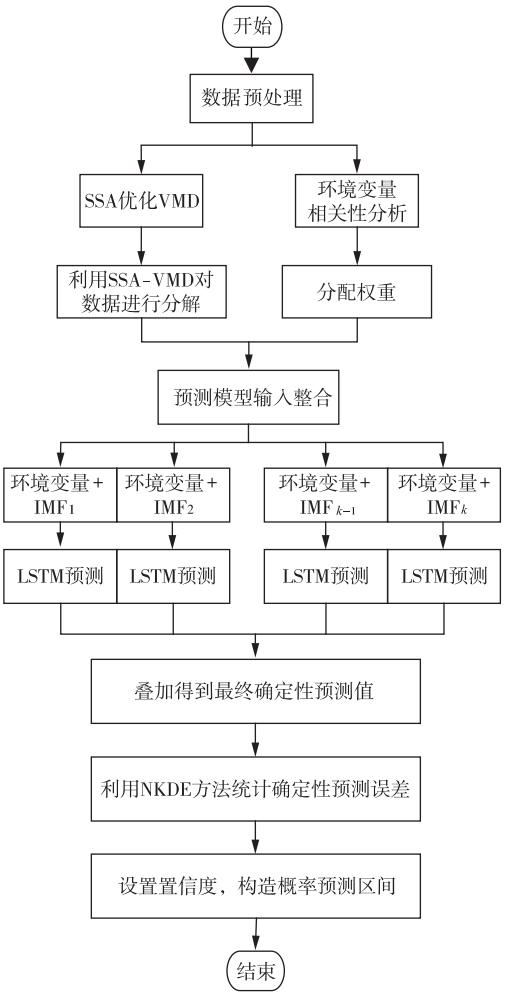

图 2 SSA-VMD-LSTM-NKDE 模型预测流程图

Fig. 2 Flow chart of SSA-VMD-LSTM-NKDE model prediction

PINAW 计算公式：

$$\mathrm{PINAW}=\frac{1}{N}\sum_{i=1}^{N}(U_i-L_i)\ 。\tag{21}$$

通过预测区间平均宽度的指标，能够有效避免因预测区间过于强调可靠性而导致预测区间的范围过大，致使不能有效地获取决策信息。当 PICP 一定时，PINAW 越小，说明预测区间越窄，预测效果越好，但过于保守的锐度指标不利于决策者对不确定性进行评估，从而导致系统运行的经济性降低。

温克勒得分（Score）：

$$\mathrm{Sr}_i=\begin{cases}-2\partial v_i-4(L_i-P_i), & P_i<L_i,\\ -2\partial v_i, & L_i\leqslant P_i\leqslant U_i,\\ -2\partial v_i-4(P_i-U_i), & P_i>U_i,\end{cases}\tag{22}$$

$$\mathrm{Score}=\sum_{i}^{N}\mathrm{Sr}_i\ ,\tag{23}$$

式中：$v_i=U_i-L_i$。

Score 是一项综合评价预测区间覆盖率和宽度的指标，可以用来反映整体的预测效果，它是非正的，结果越接近 0 值，表明预测效果越好。

4.2 基于 SSA-VMD 的风电功率分解结果

实验选取河北省某风电场 2018-07-01 至 2018-07-31 的风电机组实测数据，时间间隔 15 min，共 3 000

个样本点。由于风电功率波动、数据采集误差以及随机噪声的存在严重影响了预测模型的精确度，因此需要对原始数据进行分解。本文采用 VMD 对原始数据进行分解，从而降低数据的复杂度，但在传统的 VMD 算法中，K 和 α 的取值不当会导致原始信号欠分解或过分解，因此本文提出用 SSA 优化 VMD 参数模型。图 3 为 VMD 分解结果，$K=4$，$\alpha=2\ 000$；图 4 为 SSA 优化 VMD 后分解结果，$K=8$，$\alpha=2\ 371$。

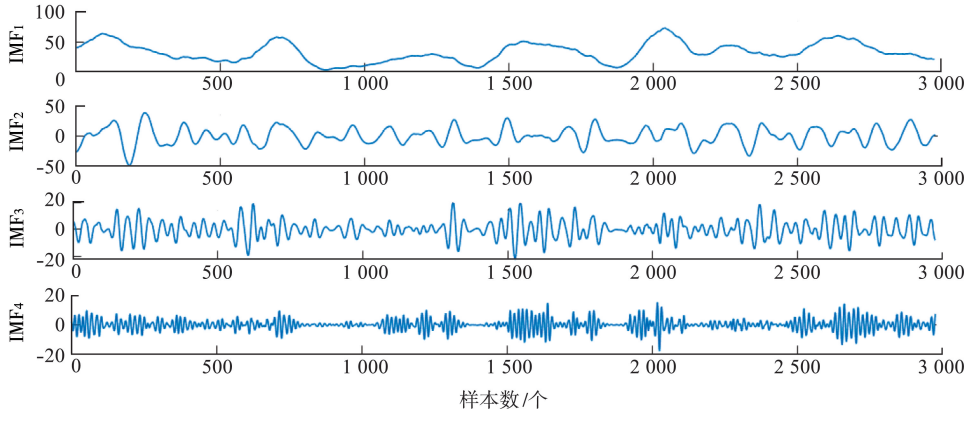

图 3　VMD 分解风功率结果

Fig. 3　Decomposition results of wind power by VMD

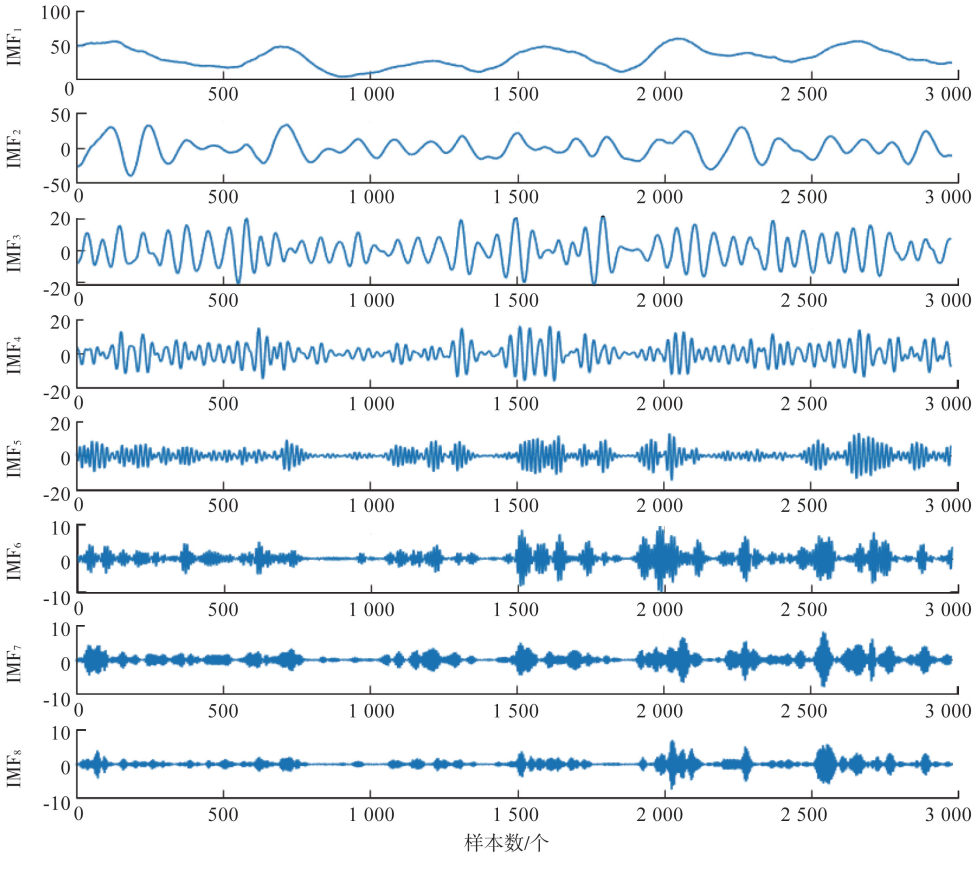

图 4　SSA-VMD 分解风功率结果

Fig. 4　Decomposition results of wind power by SSA-VMD

从图中对比可以看出，经过优化后分解数据的非线性降低，局域变化波动趋于平缓，通过将随机性较强的原始风电数据分为低频和高频子序列，有助于提高模型预测精度，其中低频子序列能够反映原数据的整体波动趋势，高频子序列能够体现对应时刻的细节信号和噪声信号。将图 4 中的模态分量进行 Hilbert 变换即可得到其边际谱，如图 5 所示。

由图 5 可知，经 SSA-VMD 分解得到的 8 个模态分量的频率区分较为明显，且分解后的每个 IMF 既保持了原始风电功率序列的特征又抑制了模态混叠。由此可见，经 SSA 优化后的 VMD 更有利于提高风电功率预测精度。图 6 为不同优化算法对 VMD 参数进行优化时适应度值对比结果，从中可以看出，粒子群算法在第 17 次迭代时出现最小适应度值 0.674 1，萤火虫算法在第 14 次迭代时出现最小适应度值 0.673 4，鲸鱼优化算法在第 12 次迭代时出现最小适应度值 0.672 9，SSA 算法随迭代次数增加迅速收敛，在第 8 次迭代时达到最小适应度函数值 0.672 5，验证了 SSA 算法的优越性。

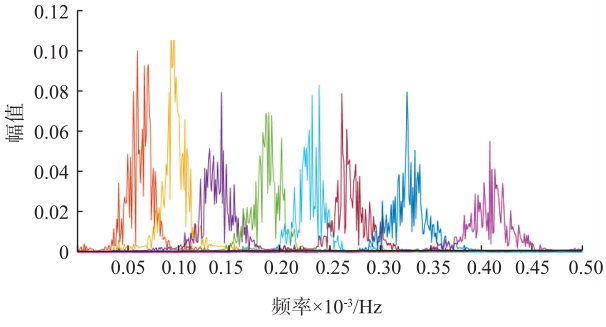

图 5 SSA-VMD 分解 IMF 边际谱

Fig. 5 SSA-VMD decomposition IMF spectrum

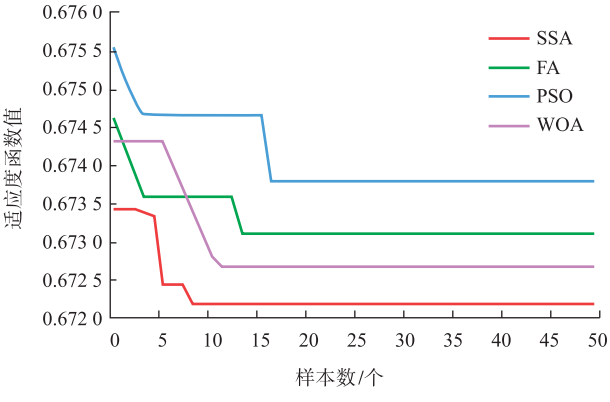

图 6 不同优化算法适应度曲线变化趋势

Fig. 6 Variation Trend of fitness curve of different optimization algorithms

4.3 基于相关性分析的预测结果

本文使用河北省某实际风电场的数据，结合风力发电特征，选取风速、空气密度、风向、温度和气压作为风电功率的影响因子，按熵权-灰色关联度分析法的计算步骤，求得各个气象因素与风电功率的相关程度分别为风速 0.981、空气密度 0.853、风向 0.871、气压 0.612、温度 0.537。可见，风速对风电功率预测的影响程度最大，其次是风向、空气密度、气压、温度。因此选取风速作为预测模型的输入。图 7 为考虑风速影响和剔除风速影响时的预测结果对比图，从中可以看出剔除风速后，预测误差明显增大。

4.4 不同模型预测结果对比分析

实验选取 3 000 个样本点，其中的前 90%用于训练，后 10%用于测试。在 LSTM 中，输入层时间步数表示进行功率预测的历史数据序列长度，其数值的确定需要考虑模型训练的有效性以及历史信息的完备性[27]；输入层维数则是各个 IMF 分量的输入变量；隐藏层数目即是 LSTM 层的个数，增加隐藏层数可提高模型的非线性拟合能力，但同时也使模型更复杂，预测时间随之增加，甚至引发过拟合问题，因此本文将隐藏层数选择的范围控制在 2 层内；LSTM 输出为功率值，故输出变量维数为 1。学习率初始值设置为 0.005，梯度阈值设置为 1。为验证该模型的有效性，本文分别搭建 BP，SVM，ELM，LSTM，EMD-LSTM 和 VMD-LSTM 模型进行对比，为保证客观性，各预测模型的参数均通过多次尝试实验选取最佳值。误差指标如表 2 所示，对比结果如图 8—图 10 所示。

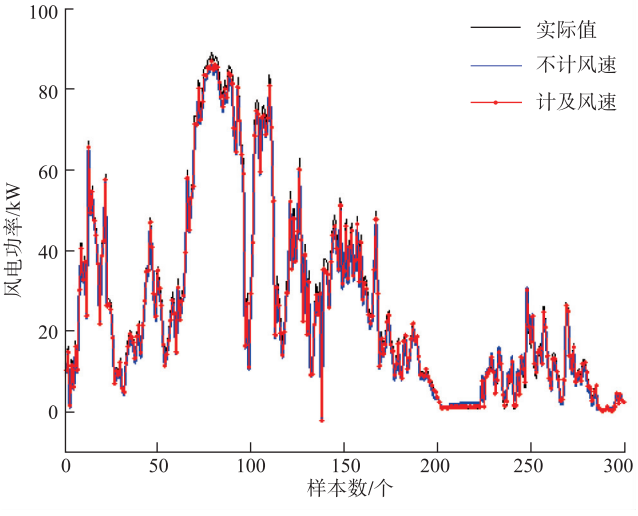

图 7 预测结果对比图

Fig. 7 Comparison of prediction results

表 2 不同模型误差评估表

Tab. 2 Error evaluation of different models

预测模型	MAE	RMSE	MAPE/%
BP	8.310 4	9.874 3	11.237 0
ELM	7.835 1	8.267 7	9.043 7
SVM	6.082 3	7.310 4	7.490 6
LSTM	4.531 9	5.994 2	6.755 3
EMD-LSTM	3.021 4	4.170 9	4.910 6
VMD-LSTM	2.189 6	3.287 6	3.274 5
SSA-VMD-LSTM	1.324 4	2.195 3	1.963 4

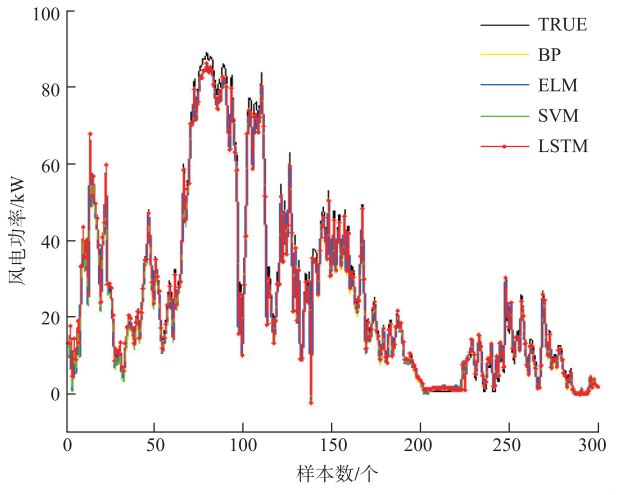

图 8　单一模型预测结果对比图

Fig. 8　Comparison of prediction results of single model

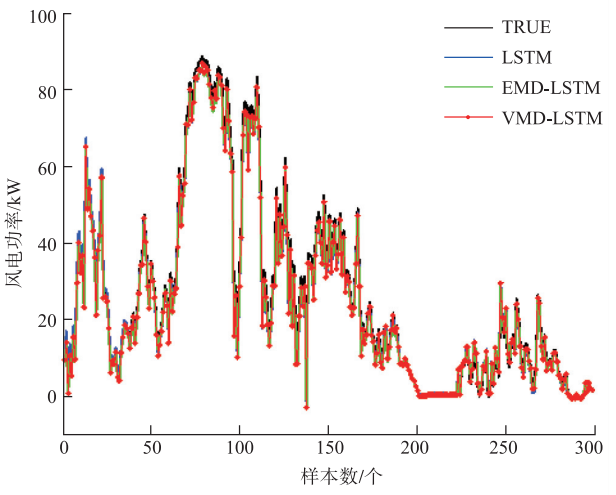

图 9　未优化模型预测结果对比图

Fig. 9　Comparison of prediction results of unoptimized model

从图 8 预测曲线和表 2 误差指标对比来看，LSTM 的预测效果更好，相比于 BP，ELM 和 SVM 模型，MAE，RMSE 和 MAPE 均有所下降，这是由于 LSTM 模型具有长时记忆特性，可以利用数据的时间序列性，有效避免了风电数据的波动性。图 9 为 LSTM，EMD-LSTM 和 VMD-LSTM 的模型预测结果对比图，从图 9 预测曲线和表 2 误差指标对比来看，经过数据分解后预测效果明显比未对数据进行分解的预测效果好，说明 LSTM 的精确程度与输入数据的质量息息相关。与 EMD-LSTM 相比，VMD-LSTM 模型的 MAE，RMSE 和 MAPE 分别下降了 27.53%，21.18%和 33.32%，表明 VMD 分解降噪效果比 EMD 好，更适合处理时序信息的预测工作。图 10 为 VMD-LSTM 和 SSA-VMD-LSTM 模型预测结果对比图，从图 10 预测曲线和表 2 误差指标对比来看，经过参数优化后再进行 VMD 分解的预测精度比直接用 VMD 进行分解的预测精度要高，SSA-VMD-LSTM 模型的 MAE，RMSE 和 MAPE 比 VMD-LSTM 的分别下降了 39.51%，33.22%和 40.39%。实验表明，采用 SSA 对 VMD 算法中的参数进行寻优，能够有效降低因参数设置不当导致信号欠分解或过分解问题对预测效果的影响。

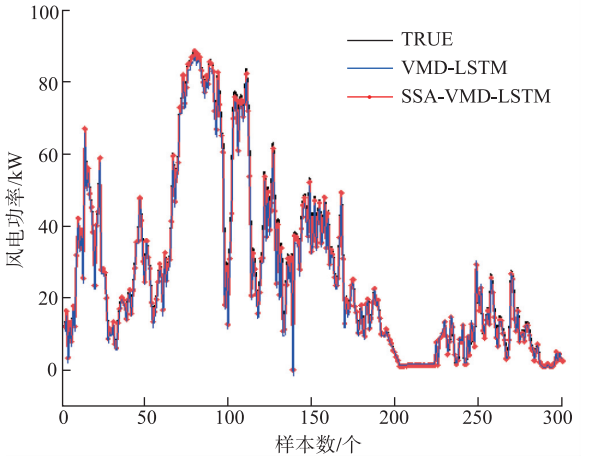

图 10　优化后模型预测结果对比图

Fig. 10　Comparison of prediction results of optimized models

4.5　概率预测结果

利用 NKDE 方法对 SSA-VMD-LSTM 模型的确定性预测误差进行分析，采用渐进积分均方误差(AMISE)来计算窗宽 h，当使用高斯核函数时，窗宽的计算可简化为 $h=1.06\sigma n^{-1/5}$。为验证 NKDE 方法的优越性，将其与假定 Gamma 分布的参数估计方法进行对比分析，基于 SSA-VMD-LSTM 模型预测误差的 NKDE 方法和基于 Gamma 分布估计方法的 PDF 如图 11 所示。

由图 11 可知，采用 NKDE 方法得到的 PDF 具有更好的拟合效果，更符合实际风电功率预测的误差分布。采用 Gamma 分布得到的 PDF 在波峰附近的拟合度明显不够，虽然 Gamma 概率密度曲线相比 NKDE 方法的概率密度曲线平滑，但与实际误差分布差别较大，因此采用 NKDE 方法对确定性预测误差分布的描述更加准确。通过对图 11 中采用 NKDE 方法拟合得到的 PDF 求积分，可得到对应的累积概率密度函数(CDF)，从而获得给定置信度下的误差分布边界，具体如图 12 所示。从中可以得出，在不同置信度下风电功率预测误差的上下界，本文取分位点为 $\alpha/2=5\%$和 $\alpha/2=10\%$，即可得到置信度为 90%和 80%下的风电功率波动区间，不同置信度下的风电功率概率预测结果如图 13 所示。

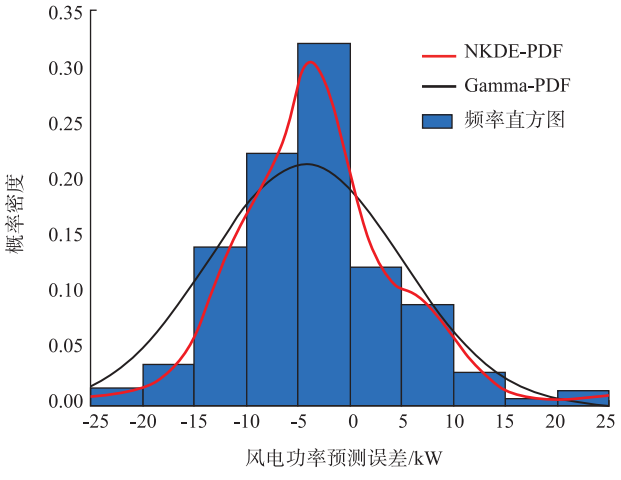

图 11 风电功率预测误差 PDF

Fig. 11 PDF of wind power prediction error

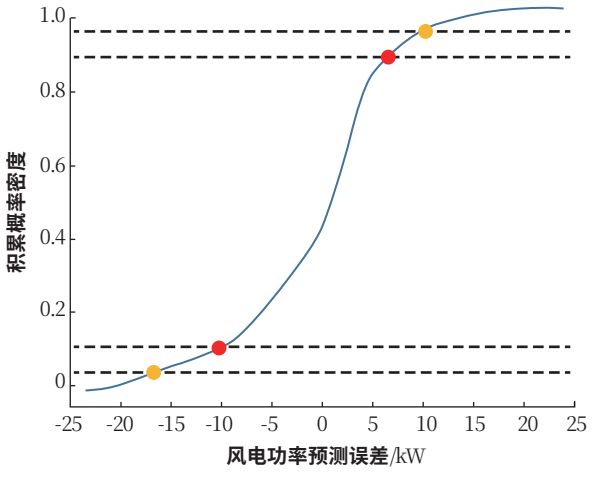

图 12 风电功率预测误差 CDF

Fig. 12 CDF of wind power prediction error

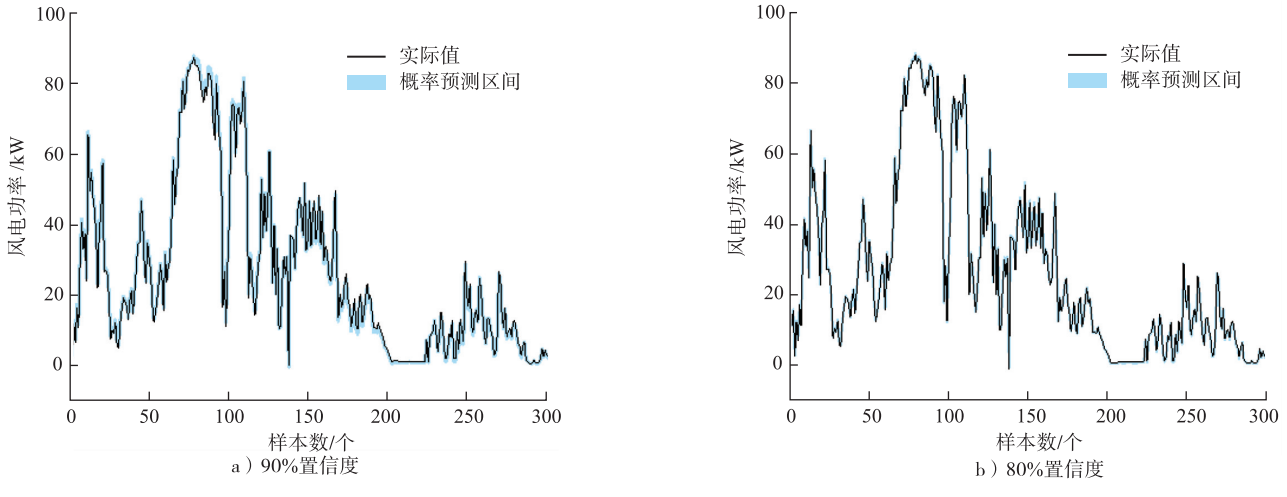

图 13 不同置信度下的风电功率概率预测结果

Fig. 13 Wind power probability prediction results under different confidence levels

在图 13 中，对比不同置信度下的风电功率概率预测结果可以看出，随着置信度的降低，实际风电功率点落在概率预测区间内的概率越小。为了能够更加直观地对参数化估计方法和非参数化估计方法的概率预测性能进行评价，选取 PICP，PINAW 和 Score 3 种常用于风电功率概率预测的评价指标来衡量概率预测的效果，概率预测结果评价指标如表 3 所示。

表 3 概率预测结果对比

Tab. 3 Comparison of probability prediction results

指 标	方法	90%置信度	80%置信度
PICP	NKDE	98.631	93.337
	Gamma	96.134	91.266
PINAW	NKDE	4.214	3.442
	Gamma	5.833	5.058
Score	NKDE	−0.735	−0.874
	Gamma	−0.987	−1.103

由表 3 可知，在置信度一定的条件下，NKDE 的覆盖率均高于 Gamma 分布方法的覆盖率，且在置信度为 90%和 80%时都满足覆盖率大于预设置信度下的覆盖率。但随着置信度的降低，预测区间覆盖率也降低，由此可见，若预测区间太宽，会出现过于保守现象。通过 PINAW 指标能够有效衡量预测区间的敏锐性，避免了

因预测区间过于强调可靠性而导致预测区间过宽，失去参考价值。本文所用的 NKDE 方法比 Gamma 分布方法的预测区间平均宽度窄，说明在一定置信度下，采用 NKDE 方法的概率预测结果更贴近真实值，预测效果更好，且在置信度为 90%和 80%时的 PINAW 值均符合短期预测中预测区间平均宽度，应在 30%～50%的要求。但在相同置信度下，PICP 随 PINAW 的增大而增大。当 PINAW 一定时，PICP 越大越好；当 PICP 一定时，PINAW 越小，表明预测区间更逼近实际功率值，预测效果越好。仅有 PICP 和 PINAW 不能全面评价概率预测性能，因此引入 Score 评价指标。采用 NKDE 方法的 Score 指标的绝对值均小于采用 Gamma 分布方法的 Score 指标的绝对值。结果表明，采用 NKDE 方法的概率预测效果更好。

5　结　论

针对 VMD 分解方法参数人为设置难以达到最优的问题，提出了一种基于 SSA-VMD-LSTM-NKDE 的组合风电预测模型。结合灰色关联度分析法和熵权法对环境变量进行相关性分析，选择相关性最高的影响因素与 SSA-VMD 分解得到的各模态分量组合作为 LSTM 模型的输入，并建立了基于 NKDE 的风电功率概率预测模型，实现对风电功率预测结果不确定性的有效量化。通过算例对比分析，得出以下结论。

1）采用 SSA 优化 VMD 参数可有效搜索得到最优参数组合。算例分析表明，与 PSO 和 WOA 等参数优化算法相比，SSA 算法更易于跳出局部最优，收敛速度更快。

2）采用熵权-灰色关联度分析风功率环境变量的相关性方法能够客观有效地对风电功率的影响因素进行评估，得出风速是影响因子最高的因素的结论。算例表明，是否考虑风速输入影响因素将直接影响风电功率预测精度。

3）采用 SSA-VMD-LSTM 组合预测模型能够有效提升短期风电功率预测精度，与未优化的预测模型相比，所提方法的 MAE，RMSE 和 MAPE 均有所降低，验证了该组合预测模型的有效性。

4）在风电功率点预测的基础上建立了基于 NKDE 的风电功率概率预测模型。与假定 Gamma 分布的参数化方法对比结果表明，基于 NKDE 的风电功率概率预测结果更可靠和更敏锐，而且在给定置信度下均满足短期风电功率预测的要求，具有一定的实际工程意义。

本文不足之处在于仅针对单风电场功率预测展开研究，下一步将针对大规模风电集群开展区域多风电场功率预测研究。

参考文献/References：

[1]　王伟胜，王铮，董存，等. 中国短期风电功率预测技术现状与误差分析[J]. 电力系统自动化，2021，45(1)：17-27.
WANG Weisheng，WANG Zheng，DONG Cun，et al. Status and error analysis of short-term forecasting technology of wind power in China [J]. Automation of Electric Power Systems，2021，45(1)：17-27.

[2]　孙荣富，张涛，和青，等. 风电功率预测关键技术及应用综述[J]. 高电压技术，2021，47(4)：1129-1143.
SUN Rongfu，ZHANG Tao，HE Qing，et al. Review on key technologies and applications in wind power forecasting[J]. High Voltage Engineering，2021，47(4)：1129-1143.

[3]　SHAO Haijian，DENG Xing，JIANG Yingtao. A novel deep learning approach for short-term wind power forecasting based on infinite feature selection and recurrent neural network[J]. Journal of Renewable and Sustainable Energy，2018，10(4). DOI：10.1063/1.5024297.

[4]　QIN Yong，LI Kun，LIANG Zhanhao，et al. Hybrid forecasting model based on long short term memory network and deep learning neural network for wind signal[J]. Applied Energy，2019，236：262-272.

[5]　ZHANG Yongzhi，XIONG Rui，HE Hongwen，et al. Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries[J]. IEEE Transactions on Vehicular Technology，2018，67(7)：5695-5705.

[6]　康田雨，覃智君. 基于超参数优化和双重注意力机制的超短期风电功率预测[J]. 南方电网技术，2022，16(5)：44-53.
KANG Tianyu，QIN Zhijun. An ultra-short-term wind power forecasting method based on hyperparameter optimization and dual-stage attention mechanism[J]. Southern Power System Technology，2022，16(5)：44-53.

[7]　王炜，刘宏伟，陈永杰，等. 基于 LSTM 循环神经网络的风力发电预测[J]. 可再生能源，2020，38(9)：1187-1191.
WANG Wei，LIU Hongwei，CHEN Yongjie，et al. Wind power forecast based on LSTM cyclic neural network[J]. Renewable Energy Resources，2020，38(9)：1187-1191.

[8]　史加荣，赵丹梦，王琳华，等. 基于 RR-VMD-LSTM 的短期风电功率预测[J]. 电力系统保护与控制，2021，49(21)：63-70.
SHI Jiarong，ZHAO Danmeng，WANG Linhua，et al. Short-term wind power prediction based on RR-VMD-LSTM[J]. Power System Protection and Control，2021，49(21)：63-70.

[9] DU Pei,WANG Jianzhou,YANG Wendong,et al. A novel hybrid model for short-term wind power forecasting[J]. Applied Soft Computing,2019,80:93-106.

[10] LIN Guancen,LIN Aijing,CAO Jianing. Multidimensional KNN algorithm based on EEMD and complexity measures in financial time series forecasting[J]. Expert Systems with Applications,2021,168. DOI: 10.1016/j.eswa.2020.114443.

[11] HU Huanling,WANG Lin,TAO Rui. Wind speed forecasting based on variational mode decomposition and improved echo state network [J]. Renewable Energy,2021,164:729-751.

[12] ABEDINIA O,LOTFI M,BAGHERI M,et al. Improved EMD-based complex prediction model for wind power forecasting[J]. IEEE Transactions on Sustainable Energy,2020,11(4):2790-2802.

[13] 王佶宣,邓斌,王江. 基于经验模态分解与 RBF 神经网络的短期风功率预测[J]. 电力系统及其自动化学报,2020,32(11):109-115.
WANG Jixuan,DENG Bin,WANG Jiang. Short-term wind power prediction based on empirical mode decomposition and RBF neural network[J]. Proceedings of the CSU-EPSA,2020,32(11):109-115.

[14] CHEN Yaoran,DONG Zhikun,WANG Yan,et al. Short-term wind speed predicting framework based on EEMD-GA-LSTM method under large scaled wind history[J]. Energy Conversion and Management,2021,227. DOI: 10.1016/j.enconman.2020.113559.

[15] NAIK J,SATAPATHY P,DASH P K. Short-term wind speed and wind power prediction using hybrid empirical mode decomposition and kernel ridge regression[J]. Applied Soft Computing,2018,70:1167-1188.

[16] DRAGOMIRETSKIY K,ZOSSO D. Variational mode decomposition[J]. IEEE Transactions on Signal Processing,2014,62(3):531-544.

[17] 盛四清,金航,刘长荣. 基于 VMD-WSGRU 的风电场发电功率中短期及短期预测[J]. 电网技术,2022,46(3):897-904.
SHENG Siqing,JIN Hang,LIU Changrong. Short-term and mid-short-term wind power forecasting based on VMD-WSGRU[J]. Power System Technology,2022,46(3):897-904.

[18] NA Wu,ZHENG XIAN Yang,SHU XIAN Fan. Extraction method and application of traveling wave natural frequencies based on PSO and VMD[C]//2020 Chinese Automation Congress (CAC). Shanghai:IEEE,2020:28-34.

[19] 张萍,张文海,赵新贺,等. WOA-VMD 算法在轴承故障诊断中的应用[J]. 噪声与振动控制,2021,41(4):86-93.
ZHANG Ping,ZHANG Wenhai,ZHAO Xinhe,et al. Application of WOA-VMD algorithm in bearing fault diagnosis[J]. Noise and Vibration Control,2021,41(4):86-93.

[20] XUE Jiankai,SHEN Bo. A novel swarm intelligence optimization approach:Sparrow search algorithm[J]. Systems Science & Control Engineering,2020,8(1):22-34.

[21] SUN Zexian,ZHAO Mingyu. Short-term wind power forecasting based on VMD decomposition,ConvLSTM networks and error analysis [J]. IEEE Access,2020,8:134422-134434.

[22] NAIK J,DASH S,DASH P K,et al. Short term wind power forecasting using hybrid variational mode decomposition and multi-kernel regularized pseudo inverse neural network[J]. Renewable Energy,2018,118:180-212.

[23] YAN Pengcheng,SHANG Songhang,ZHANG Chaoyin,et al. Research on the processing of coal mine water source data by optimizing BP neural network algorithm with sparrow search algorithm[J]. IEEE Access,2021,9:108718-108730.

[24] WUMAIER T,XU Chang,GUO Hongyu,et al. Fault diagnosis of wind turbines based on a support vector machine optimized by the sparrow search algorithm[J]. IEEE Access,2021,9:69307-69315.

[25] HAN Min,ZHANG Ruiquan,QIU Tie,et al. Multivariate chaotic time series prediction based on improved grey relational analysis[J]. IEEE Transactions on Systems,Man,and Cybernetics:Systems,2019,49(10):2144-2154.

[26] YANG M S,NATALIANI Y. A feature-reduction fuzzy clustering algorithm based on feature-weighted entropy[J]. IEEE Transactions on Fuzzy Systems,2018,26(2):817-835.

[27] 胡威,张新燕,李振恩,等. 基于优化的 VMD-mRMR-LSTM 模型的短期负荷预测[J]. 电力系统保护与控制,2022,50(1):88-97.
HU Wei,ZHANG Xinyan,LI Zhenen,et al. Short-term load forecasting based on an optimized VMD-m RMR-LSTM model[J]. Power System Protection and Control,2022,50(1):88-97.