

[风电发电]

基于GRA-IPSO-GRU的短期风电功率预测研究

文/方 严 万俊杰
安科瑞电气股份有限公司

考虑到风力发电功率具有波动性和不稳定性的特点，提出了一种基于GRA-IPSO-GRU的短期风电功率预测模型以保证电力系统的安全稳定运行。将预测模型应用到实际案例后的对比结果表明，所提预测模型的准确度更高，稳定性更强，为实现微电网能量管理中的发电优化调度、负荷管理提供了一定的依据。

近年来，我国风电装机容量增长迅速，目前风电装机规模居世界首位。随着碳达峰、碳中和目标的提出，未来新能源装机量将迎来爆发式增长。

目前的风电功率预测方法主要分为三种，分别为物理方法、时间序列方法和人工智能方法。物理方法主要是以数字天气预报为基础的预测方法；时间序列方法是通过统计学方法进行处理，常用的时序模型有自回归滑动平均模型、差分自回归滑动平均模型和自回归条件异方差模型；常用的人工智能算法包括人工神经网络、支持向量机等。本文采用门控循环单元神经网络模型，充分利用了风电场功率的时序相关性进行预测，并通过改进的粒子群算法对网络参数进行优化。

灰色关联分析法

灰色关联分析法（Grey Relational Analysis, GRA）是一种以小样本和贫信息为基础的灰色系统分析方法。灰色关联分析法适用于任何有规律或无规律的样本，且在分析时计算量较小，因而得到广泛应用。具体步骤如下：首先确定评价对象和影响因素，并建立评价矩阵。假设有 m 个影响因素（ $i=1,2,\dots,m$ ），每个影响因素有 n 个数据（ $k=1,2,\dots,n$ ），则可建立评价矩阵为

$$X = \{x_i(k)\} = \begin{pmatrix} x_1(1) & x_1(2) & \cdots & x_1(k) & \cdots & x_1(n) \\ x_2(1) & x_2(2) & \cdots & x_2(k) & \cdots & x_2(n) \\ \vdots & \vdots & & \vdots & & \vdots \\ x_i(1) & x_i(2) & \cdots & x_i(k) & \cdots & x_i(n) \\ \vdots & \vdots & & \vdots & & \vdots \\ x_m(1) & x_m(2) & \cdots & x_m(k) & \cdots & x_m(n) \end{pmatrix} \quad (1)$$

步骤一：确定参考序列和比较序列，将短期风力发电功率数据列记为参考序列 $X_j(k)$ ， $k=1,2,\dots,n$ ， $j=1,2,\dots,h$ ，其中 n 表示风力发电功率数据列的长度， h 表示风机的数量。将短期风力发电功率影响因素数据列记为比较序列 $X_i(k)$ ， $k=1,2,\dots,n$ ， $i=1,2,\dots,m$ ，其中 n 为影响因素数据列的长度， m 表示影响因素的个数。

步骤二：对数据进行无量纲化处理，由于各影响因素数据之间的物理意义和计量单位有所不同，会产生不同的量纲，从而导致数量级差悬殊，故应对数据进行无量纲化（均值化）处理

$$X'_j(k) = X_j(k) / \bar{X}_j \quad (2)$$

$$X'_i(k) = X_i(k) / \bar{X}_i \quad (3)$$

式中， $X_j(k)$ 为第 j 个参考序列的第 k 个数据； $\bar{X}_j$ 为第 j 个参考序列的平均值； $X_i(k)$ 为第 i 个比较序列的第 k 个数据； $\bar{X}_i$ 为第 i 个比较序列的平均值。

步骤三：计算参考序列和比较序列之差的绝对

值

$$\Delta_{ji}(k) = |X_j'(k) - X_i'(k)| \quad (4)$$

式中, $X_j'(k)$ 和 $X_i'(k)$ 分别为第 j 个无量纲参考序列和第 i 个无量纲比较序列。

步骤四: 求二级的 $\Delta_{\max}$ 和 $\Delta_{\min}$

$$\Delta_{\max} = \max_i \max_k \Delta_{ji}(k) \quad (5)$$

$$\Delta_{\min} = \min_i \min_k \Delta_{ji}(k) \quad (6)$$

式中, $\Delta_{\max}$ 和 $\Delta_{\min}$ 分别为二级的最大和最小绝对差。

步骤五: 计算各影响因素的关联系数

$$\xi_i(k) = (\Delta_{\min} + \rho \Delta_{\max}) / \Delta_{ji}(k) + \rho \Delta_{\max} \quad (7)$$

式中, ρ 为分辨系数, $\rho \in [0, 1]$, 一般取 $\rho = 0.5$ 。

步骤六: 计算各影响因素 $X_i(k)$ 相对风力发电功率 $X_j(k)$ 的熵权关联度

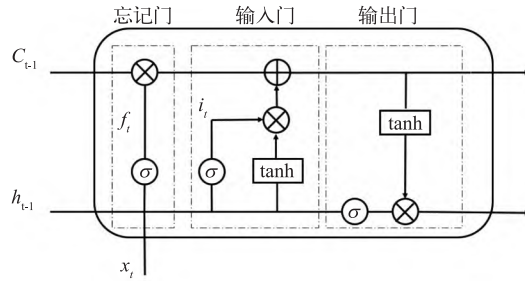
$$r_i = \frac{1}{n} \sum_{k=1}^n \xi_i(k) \quad (8)$$

通过灰色关联分析法计算, 可得到风力发电功率的影响因素相对风力发电功率的灰色关联度。灰色关联度的大小代表各影响因素对风力发电功率的影响程度, 灰色关联度值越大, 两者之间的关联越大; 反之, 两者之间的关联越小。

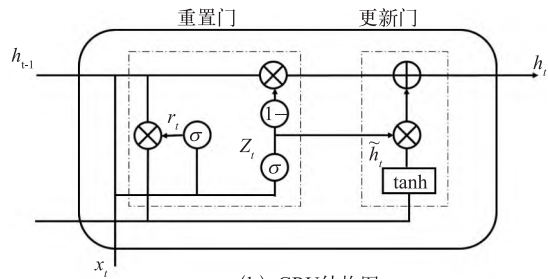
门控循环单元

门控循环单元 (Gate Recurrent Unit, GRU) 是长短期记忆 (Long Short Term Memory, LSTM) 神经网络的一种变体。它保持了LSTM的效果, 同时结构也更加简单, 且不容易出现过拟合的现象。LSTM是一种改进的循环神经网络, 改进之处就是用LSTM细胞单元结构代替了RNN的隐藏层单元, 并且有效地解决了RNN中存在的长期依赖的问题。GRU在LSTM的基础上对门的设计进行了改进, 训练参数有所减少, 但是内部结构本质上并未改变, 不仅具有和LSTM一样的预测准确度, 而且还大大减少了网络训练时间。

在图1中可以看出, GRU将LSTM中的忘记门和输入门优化成更新门, 将输出门更改为重置门。更新门的作用是用来控制记忆单元, 而重置门是将单元状态和输出合并为隐藏状态, 并且靠隐藏状态传输信息。LSTM中信息流添加到记忆单元并且与遗忘门相互独立, 而在GRU中, 信息流的更新是与更新门在一起完成的, 因此GRU与LSTM相比, 减少



(a) LSTM单元结构图



(b) GRU结构图

图1

了结构的复杂度, 运算速度也有所提升, 公式为

$$z_t = \sigma(W_z[h_{t-1}, x_t]) \quad (9)$$

$$r_t = \sigma(W_r[h_{t-1}, x_t]) \quad (10)$$

$$\tilde{h}_t = \tanh(W[\tilde{r}_t * h_{t-1}, x_t]) \quad (11)$$

$$h_t = (1 - z_t) h_{t-1} + z_t \tilde{h}_t \quad (12)$$

式中, z_t 为更新门的状态, 决定了信息的取舍; r_t 为重置门的状态, 决定了信息被忘记的程度, $\tilde{h}_t$ 为当前神经元的待定输出, 相当于对当前时刻的状态进行了记忆; h_t 为当前神经元的输出; x_t 为当前神经元的输入; h_{t-1} 为上一个神经元的输出; σ 为sigmoid函数; W 为权重参数; $\tanh$ 双曲正切函数。

基于IPSO-GRU的短期风电功率预测模型

1. 粒子群算法及其改进

粒子群优化 (Particle Swarm Optimization, PSO) 算法是一种模拟鸟类觅食行为的群智能算法。鸟类在空中飞行时, 如果有一只鸟脱离队伍飞向栖息地, 其余的鸟类也会跟随这只鸟一起飞向栖息地。鸟类寻找栖息地的过程与一个特定的问题寻找解的过程十分相似。

PSO算法可以描述为: 在一个 D 维搜索空间中, 一个种群由 s 个粒子组成, 可由如下矩阵表示

$$\mathbf{Z} = (Z_1, Z_2, Z_3, \dots, Z_s) \quad (13)$$

第 i 个粒子的位置可以表示为

$$Z_i = (Z_{i1}, Z_{i2}, Z_{i3}, \dots, Z_{iD})^T \quad (14)$$

式中, Z_{iD} 为第*i*个粒子在第*D*维的位置。

$$V_i = (V_{i1}, V_{i2}, V_{i3}, \dots, V_{iD})^T \quad (15)$$

式中, V_{iD} 为第*i*个粒子在第*D*维的速度。

随着每一次的迭代进行, 粒子都通过个体的最优值和全局的最优解来更新自身的速度和位置, 更新公式为

$$V_{id}^{k+1} = \omega V_{id}^k + c_1 r_1 (P_{id}^k - Z_{id}^k) + c_2 r_2 (P_{gd}^k - Z_{id}^k) \quad (16)$$

$$Z_{id}^{k+1} = Z_{id}^k + a V_{id}^{k+1} \quad (17)$$

式中, v_{id}^{k+1} 为第*i*个粒子在第*k*+1次迭代中第*d*维的速度; Z_{id}^{k+1} 为粒子*i*在第*k*+1迭代中第*d*维的位置; *k*为当前迭代次数; ω 为惯性权重; $d=1, 2, 3, \dots, D$; $i=1, 2, 3, \dots, s$; c_1 和 c_2 为学习因子, 代表粒子本身的信息和对外交流的信息; $r_1 \in \text{rand}(0, 1)$, $r_2 \in \text{rand}(0, 1)$; a 为约束因子; P_{jd}^k 为某个种群在第*d*维的全局极值点的位置。

为了防止粒子跳出搜索的范围, 应将粒子的位置变化范围设置在 $-Z_{\max} \sim Z_{\max}$ 之间, 速度的变化范围设置在 $-V_{\max} \sim V_{\max}$ 之间。

惯性权重系数 ω 的表达式为

$$\omega = \omega_{\max} - k \left(\frac{\omega_{\max} - \omega_{\min}}{N} \right) \frac{\pi}{4}, (k=1, 2, 3, \dots, N) \quad (18)$$

式中, $\omega_{\max}$ 为初始惯性权重系数的最大值; $\omega_{\min}$ 为结束迭代时惯性权重系数的最小值; *k*为当前迭代次数; *N*为最大的迭代次数。

PSO通过迭代计算确定粒子的全局最优解。然而, 粒子群算法在收敛的过程中会出现很多的问题, 比如停滞或者收敛准确度比较低的问题, 因此, 为了提高粒子群算法的局部和全局搜索能力, 本文对粒子群算法进行了改进, 改进部分如下。

1) 本文采用一种非线性权重方法改善标准粒子群算法的不足, 这种算法可描述为

$$\omega = \omega_{\max} - (\omega_{\max} - \omega_{\min}) \tan\left(\frac{k}{N}\right) \cdot \frac{\pi}{4} \quad (19)$$

当迭代次数*k*较小时, ω 与 $\omega_{\max}$ 近乎相等, 保证了算法的全局搜索能力, 随着迭代次数*k*的增加, ω 以非线性方式递减, 保证了算法的局部搜索能力, 同时调节全局搜索和局部搜索能力, 使得算法更加灵活, 非线性权重与迭代次数关系如图2所示。

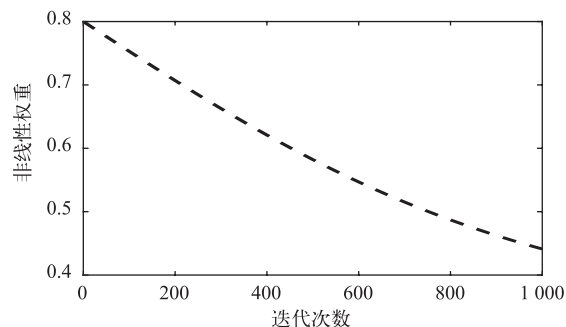


图2 非线性权重与迭代次数关系图

2) 遗传算法中变异的思想可以很好地防止粒子陷入局部极值, 在粒子群算法中加入变异因子, 使得粒子每次更新之后可以以一定变异率初始化粒子, 增大搜索范围。

将改进后的PSO算法对10维sphere函数进行极值寻优, PSO中惯性权重 $\omega=0.8$, 学习因子 $c_1=1.5$, $c_2=1.5$, IPSO中惯性权重最大值 $\omega_{\max}=0.8$, 最小值 $\omega_{\min}=0.2$, IPSO算法和PSO算法中最大迭代次数均设为1 000, 种群规模 $N=5$ 。sphere函数三维图如图3所示。

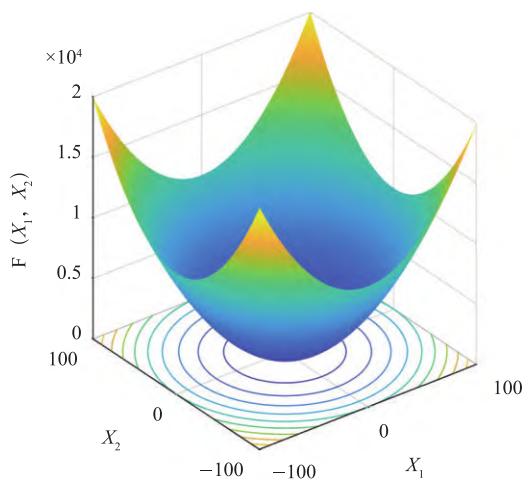


图3 sphere函数三维图

图4为PSO与IPSO对10维sphere函数极值寻优, 图4中横坐标代表算法的迭代次数, 纵坐标代表适应度值。从图4中可以清晰看出, PSO算法大约在60次迭代时就早早地陷入了局部最优解, 而改进的IPSO算法极大地增强了寻优能力, 在1 000次迭代时依然具有寻找更优解的能力, 其中曲线下降越快, 表明算法的收敛速度越快; 曲线越低, 表明算法的收敛准确度越高。

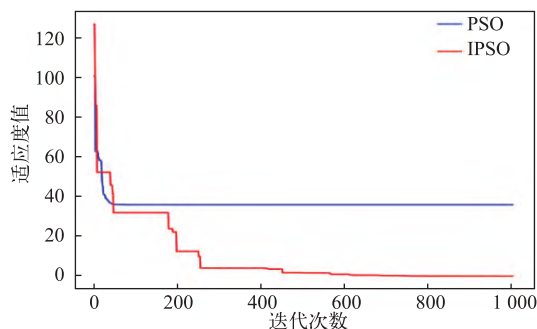


图4 PSO和IPSO的适应度变化曲线图

2. IPSO-GRU预测模型

GRU模型关键参数的选取对模型预测准确度的影响非常大, 所以对于参数的选择很重要。粒子群算法具有控制参数少、易于实现和运行速度快等优点, 在处理非线性的问题上具有一定的优势, 可用于选取GRU的模型参数。

在本文的GRU模型中, 学习率为 ε , 两层隐藏层神经元数量 L_1 、 L_2 和迭代次数 k 对模型的性能有着很大的影响。将这四个参数作为粒子群算法寻优的特征, 利用IPSO算法对模型进行优化。在GRU的隐藏层中, 每一层的输出作为下一层的输入, 最后通过全连接层输出数据。具体建模流程如图5所示, 具体步骤如下。

1) 将数据集按照9:1的比例划分出训练集和测试集, 并对风电功率数据进行标准化处理, 其中标准化处理的公式为

$$x^* = \frac{x - \mu}{\sigma} \quad (20)$$

式中, μ 为所有样本数据的均值; σ 为所有样本数据的标准差。

2) 初始化模型的参数。包括种群规模、学习因子、迭代次数以及粒子位置和速度的取值范围。

3) 根据初始化神经元个数 L_1 、 L_2 和学习率 ε 、迭代次数 k 的取值建立GRU网络模型, 将训练集和测试集分别用于模型的训练与测试, 将训练集预测结果的均方误差作为粒子的适应度值。

4) 根据粒子的初始适应度值确定局部最优解(P_{best})和全局最优解(G_{best})的位置, 并且将每个粒子的最优位置作为历史最佳位置, 根据式(17)、式(18)更新粒子的速度和位置, 计算对应的粒子适应度值, 并比较局部与全局的最优解, 使得预测的准确率更高。

5) 判断粒子是否达到最大迭代次数, 如果达到了最大迭代次数, 则输出最优参数; 若未达到最大迭代次数, 则返回步骤4)继续执行, 直到达到迭代次数为止。

6) 根据IPSO优化获得的最优参数构建IPSO-GRU预测模型, 结合模型的评价指标对模型进行分析。

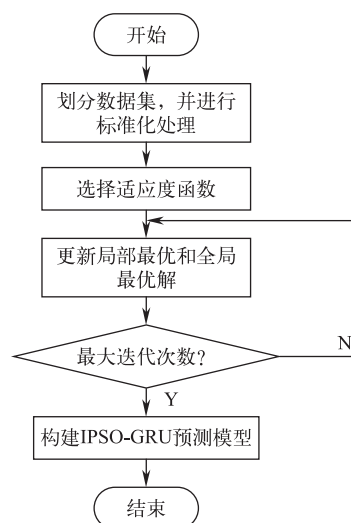


图5 IPSO-GRU建模流程图

3. 模型评价指标

为验证所提模型的优势, 选用了均方根误差(RMSE)、平均绝对误差(MAE)和拟合度(R^2)这三个指标作为模型的评价指标。各指标的计算公式为

$$RMSE = \sqrt{\frac{\sum_{i=1}^S (\hat{Y}_i - Y_i)^2}{S}} \quad (21)$$

$$MAE = \frac{1}{S} \sum_{i=1}^S |\hat{Y}_i - Y_i| \quad (22)$$

$$R^2 = 1 - \frac{\sum_{i=1}^S (\hat{Y}_i - Y_i)^2}{\sum_{i=1}^S (\bar{Y}_i - Y_i)^2} \quad (23)$$

式中, $\hat{Y}_i$ 为预测值; Y_i 为真实值; $\bar{Y}_i$ 为平均值; S 为测试数据样本总数。

实验验证与分析

本文数据集选择KDD Cup 2022风力发电数据集, 选择前10 000组数据作为实验数据, 前9 000组作为训练数据, 后1 000组作为测试数据, 具体功率

曲线如图6所示,数据精度为15 min,影响因素数据为风速 X_1 、风向与涡轮机舱位置的夹角 X_2 、温度 X_3 、涡轮机舱内的温度 X_4 、机舱方向 X_5 以及三个叶片螺距角 X_6 、 X_7 和 X_8 。

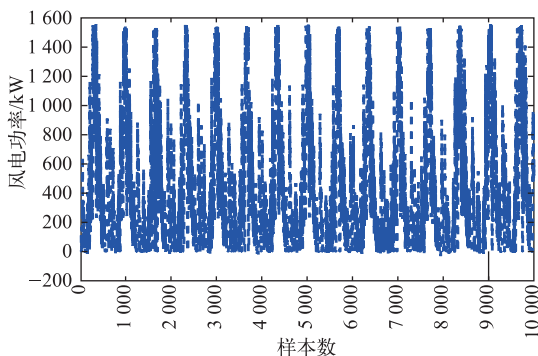


图6 风力发电功率曲线图

1. 主要影响因素

根据灰色关联分析法筛选出影响风力发电功率的主要影响因素。首先确定参考序列和比较序列,然后对数据进行无量纲化处理,接着计算参考序列和比较序列之差的绝对值,最后计算各影响因素相对风电功率的灰色关联度。各影响因素的灰色关联度柱状图如图7所示。

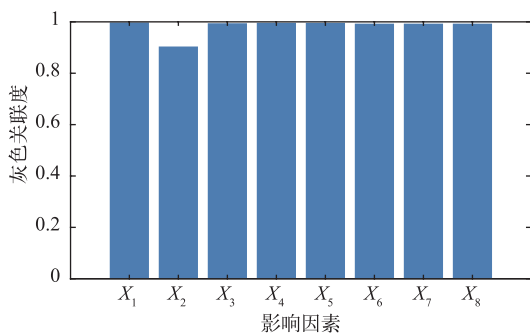


图7 各影响因素灰色关联度柱状图

由图7可知,各影响因素相对于风电功率的灰色关联度值均大于0.8,其中大于0.8的称为强相关影响因素,因此,本文将这8个影响因素都作为模型的输入。

2. 模型参数确定

为了验证IPSO-GRU模型的预测准确度,选取了LSSVM、LSTM、GRU和PSO-GRU模型,在相同的测试集上比较各模型的准确度,其中LSSVM模型参数 $\sigma=10$, $C=10$;LSTM和GRU模型的隐藏层第一层神经元数量为200,第二层神经元数量也为

200,训练次数为100,学习率为0.005;PSO-GRU中惯性权重 $\omega=0.8$,学习因子 $c_1=1.5$, $c_2=1.5$,IPSO-GRU中惯性权重最大值 $\omega_{\max}=0.8$,最小值 $\omega_{\min}=0.2$,IPSO-GRU和PSO-GRU中最大迭代次数均设为1 000,种群规模 $N=5$,维度为4,隐藏层两层神经元、训练次数和学习率的优化区间分别为[1,200],[1,200],[100,500],[0.0001,0.01]。

3. 实验结果与分析

本文将LSSVM、LSTM、GRU、PSO-GRU和IPSO-GRU预测模型的预测结果与真实值进行了对比分析,如图8~图12所示。可以看出,IPSO-GRU模型的预测结果最接近于真实值。对5个模型预测结果的评价指标值进行对比分析,结果见表1。由表1可以看出,IPSO-GRU模型的RMSE和MAE值均为最小,说明该模型预测结果的误差最小; R^2 值最大,说明该模型预测结果的拟合度最高,综上所述,采用IPSO优化GRU模型使得模型达到了更高的预测准确度。

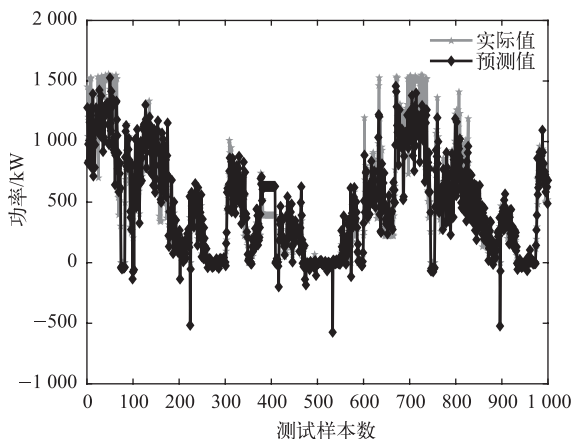


图8 LSSVM模型实际值与预测值关系图

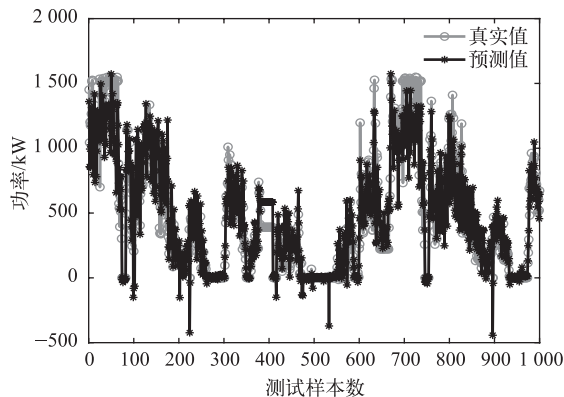


图9 LSTM模型实际值与预测值关系图

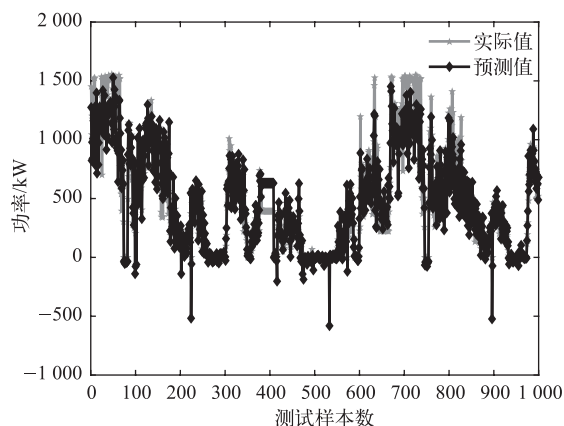


图10 GRU模型实际值与预测值关系图

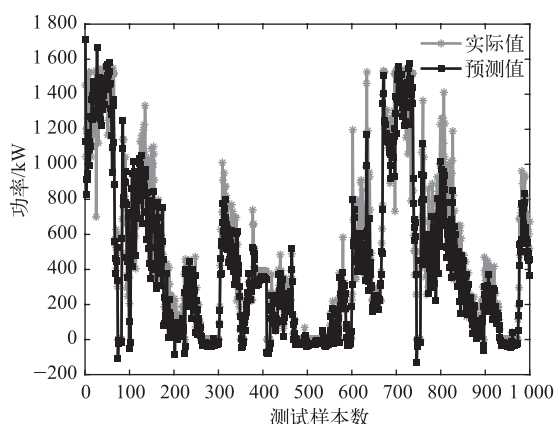


图11 PSO-GRU模型实际值与预测值关系图

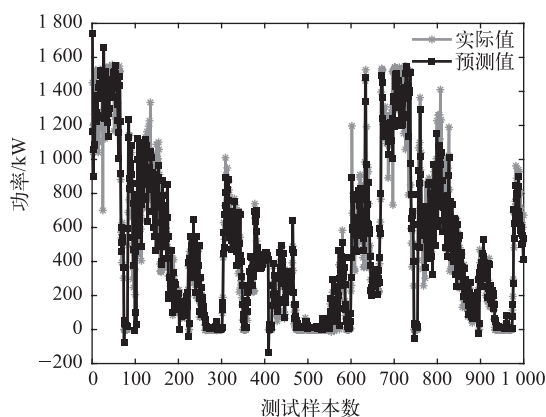


图12 IPSO-GRU模型实际值与预测值关系图

表1 测试数据误差对比分析

预测模型	RMSE	MAE	R_2
LSSVM	173.23	120.44	0.865
LSTM	165.21	114.42	0.876
GRU	152.65	101.92	0.894
PSO-GRU	137.84	94.96	0.928
IPSO-GRU	116.44	80.15	0.938

结束语

本文建立了一种GRA-IPSO-GRU的短期风电功率预测模型。首先,采用灰色关联分析法筛选出了风电功率的主要影响因素;其次,引入非线性权重和遗传变异的思想对标准的粒子群算法进行改进,提高了算法的全局优化性能,同时利用改进的粒子群算法对GRU的模型参数进行寻优,有效地避免了参数选择盲目性的问题;最后将建立好的IPSO-GRU预测模型应用于实际案例中,并与其他模型的预测结果及误差进行对比分析,实验结果表明IPSO-GRU模型具有较高的预测准确度和较低的预测误差,在一定程度上为实现微电网能量管理中的发电优化调度、负荷管理提供一定的依据。EA

(收稿日期:2023-03-27)

参考文献

- [1] 邹金,朱继忠,赖旭,等.基于时空自回归移动平均模型的风电出力序列模拟[J].电力系统自动化,2019,43(3):101-107.
- [2] 曹俊波,周任军,邓学华,等.考虑优化ARIMA模型差分次数的风功率预测[J].电力系统及其自动化学报,2019,31(1):105-111.
- [3] 李力行,苗世洪,涂青宇,等.考虑异方差效应的风电不确定性建模及其在调度中的应用[J].电力系统自动化,2020,44(8):36-47.
- [4] 黄辰,吴峻青.基于人工神经网络的短期风电功率预测[J].华东电力,2014,42(7):1408-1410.
- [5] 王腾洲,李森文,黄宇轩,等.基于IWOA-SVM的风电功率预测[J].机械与电子,2022,40(5):9-12.
- [6] 吴静澈,赵心媛.基于灰色关联分析的广西海洋产业结构研究[J].北部湾大学学报,2021,36(6):25-30.
- [7] 吴芮,张守京,慎明俊.采用注意力DBN-GRU的滚动轴承剩余寿命预测方法[J].组合机床与自动化加工技术,2022(11):101-105.
- [8] 李爱国,覃征,鲍复民,等.粒子群优化算法[J].计算机工程与应用,2002(21):1-3,17.
- [9] 谭振龙,张伶俐,陈希,等.远海大容量机组海上风电集电系统技术经济研究[J].电气应用,2021,40(6):64-71.