

基于ELSTM算法的区域风电发电量预测方法的研究

朱尤成, 王金荣, 徐 坚

(国电电力云南新能源开发有限公司, 云南 昆明 650051)

摘要: 介绍了当前风电发电量预测的常用手段及面临的挑战, 阐述了不同时间尺度下风电发电量预测的区别, 将中长期发电量的预测手段和方法作为重点分析的对象, 提出了基于ELSTM算法的区域性风电发电量预测方法, 对其原理进行了分析。对国内云南某区域7个气象站中的气象要素数据、区域内风电场的气象要素数据及发电量数据进行算法的验证, 测试过程中构建了BP-NN全参数、ELSTM全参数、BP-NN风场参数和ELSTM风场参数四种对比模型, 验证了其算法的有效性, 明确了该算法具备工程应用价值, 并能够针对区域中长期风电发电量有效地进行预测, 进而为区域性的风电大规模消纳提供良好的基础。

关键词: 风电; 发电量; 预测; 算法; 中长期

Research on prediction method of regional wind power generation based on ELSTM

ZHU Youcheng, WANG Jinrong, XU Jian

(Guodian Power Yunnan New Energy Development Co., Ltd., Kunming 650051, China)

Abstract: The current common methods and challenges of wind power generation prediction are introduced; The differences of wind power generation prediction in different time scales are expounded; The improved RNN regional wind power generation prediction method based on LSTM is put forward and its principle is analyzed. The algorithm by the real data of a certain region of Yunnan in China is validated and is compared with the conventional RNN algorithm, which verifies the effectiveness of the algorithm; It makes clear that the algorithm has engineering application value and can effectively predict the regional wind power generation, and provides a good foundation for the large-scale consumption of regional wind power.

Key words: wind power; power generation; prediction; algorithm; medium and long term

0 引言

近年来, 风力发电在我国的发展速度十分迅猛, 风电累计装机容量从2007年的589万kW增长到2017年的16 400万kW, 年均增长率达到39.47%。随着煤炭资源的日益紧张, 我国的风电并网装机容量还会进一步上升。随着国内风电资源的不断利用, 如何实现区域性风电发电量

的预测成为了各大院校和科研机构所重点关注的问题。风电场的选址影响因素不仅包含了当地的气候, 还包含了整个区域内其他能源的占比。同时, 风电预测的分类标准种类也存在多种方式, 不仅包含了建模对象、模型原理这种从机理角度出发的方法, 还包含了时间尺度和空间范围这种从时空角度出发的方式。其中时间尺度由于受到

预测能力和技术水平的限制,其各个国家的划分标准也都不相同,但大体上可分为超短期预测、短期预测和中长期预测。目前在短期预测上,国内外的专家学者已经进行了大量的算法研究,并在实际应用中取得了良好的效果,如神经网络、非参数回归,进一步地还包括径向基神经网络、广义自回归神经网络以及ELMAN神经网络等方法。以上方法一定程度上均能够解决短期预测的问题,然而风电的大规模并网给电力系统安全稳定运行所带来的影响也与日俱增。

要同时保证风电的发展和电力系统的安全运行就不得不找到其中的平衡点,这也就导致了区域内严重的弃风现象,进而造成了巨大的能源浪费。因此,可以说中国风电健康发展的关键点就在于如何解决大规模的风电消纳问题。而大规模风电消纳的问题一直也是世界性的难题,欧美国普遍采用分布式电源入网的方法增强电网系统的调峰能力,但我国的新能源分布情况比较集中,并且规模较大,因此很难就地消纳。所以我国一般会采用将多余的风电送到负荷中心,通过配合区域间资源互补的优势进行集中消纳,但这就导致了跨区域风电输送的电量非常大,而短时的电力交易会对电网的调度系统产生非常大的影响,所以需要制订长期的电力交易计划。因此,区域的中长期发电量预测成为了解决这一问题的重要关键点所在。本文基于某区域内历史多年气象真实数据、历史发电量数据进行预测模型的建立,并同时通过建立区域内风电场站年度发电量预测模型,来预测未来一年的发电量,并和实际的发电量进行了对比。

1 方法的提出

1.1 循环神经网络

循环神经网络(RNN)是一种强大且成熟的深度神经网络算法,在对具有长期依赖性的时序数据处理上效果显著,近年来在自然语言处理(Natural Language Processing, NLP)中(如语音识别、机器翻译等)取得了巨大成功,并得到了

广泛应用,同时也受到了专业人士的高度认可。

RNN内部包含三个部分:输入单元(Input Units)、输出单元(Output Units)和隐含单元(Hidden Units),其训练时采用BPTT算法在其模型上进行梯度计算,其数学表达式为

$$h_t = f(W_{hx}x_t + W_{hh}h_{t-1} + b_h) \quad (1)$$

$$y_t = g(W_{yh}h_t + b_y) \quad (2)$$

式中, W_{hx} 、 W_{hh} 、 W_{yh} 分别为输入层到隐含层的权重矩阵、隐含层自循环的权重矩阵及隐含层到输出层的权重矩阵; b_h 和 b_y 分别为隐含层和输出层的偏置量。

1.2 基于LSTM的RNN优化学习算法

传统的RNN算法中每一步的输出不仅取决于当前步的网络状态,还会受到前面若干步的网络状态的影响。随着时间步不断迭代,该算法会导致梯度逐渐消失。为了解决这一问题,Hochreiter和Schmidhuber最早提出了长短期记忆LSTM神经网络模型,LSTM神经网络是一种特殊的RNN,其主要作用是了解决传统算法中的长时间依赖问题,LSTM通过采用新的记忆单元(Memory Units)代替传统RNN模型中的隐含单元,让每个记忆单元中都包含一个或多个记忆细胞,记忆单元内部包含四个元素:输入门(Input Gate)、输出门(Output Gate)、遗忘门(Forget Gate)和记忆细胞(Memory Cell)。

其节点内部示意图如图1所示。

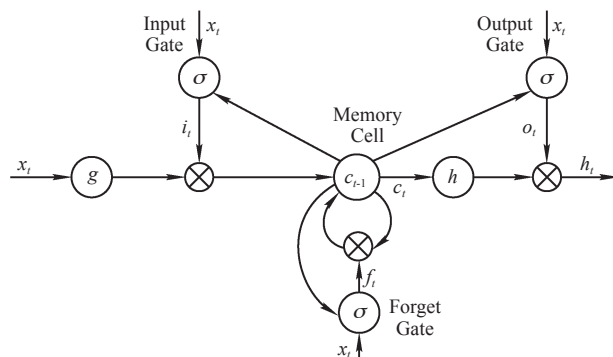


图1 LSTM内部节点示意图

Fig. 1 Schematic diagram of LSTM internal nodes

图1中, x_t 、 h_t 分别为 t 时刻LSTM网络的输入和输出, $t=1,2,3,\dots$ 。其输出 h 可通过以下公式迭代计算得出

$$i_t = \sigma(W_{ix}x_t + W_{ih}h_{t-1} + W_{ic}c_{t-1} + b_i) \quad (3)$$

$$f_t = \sigma(W_{fx}x_t + W_{fh}h_{t-1} + W_{fc}c_{t-1} + b_f) \quad (4)$$

$$o_t = \sigma(W_{ox}x_t + W_{oh}h_{t-1} + W_{oc}c_{t-1} + b_o) \quad (5)$$

$$c_t = f_t c_{t-1} + i_t \varphi(W_{cx}x_t + W_{ch}h_{t-1} + b_c) \quad (6)$$

$$h_t = o_t \varphi(c_t) \quad (7)$$

式中, i_t 、 f_t 、 o_t 和 c_t 分别为输入门、输出门、记忆细胞和输出门的输出; W_{ix} 、 W_{ih} 和 W_{ic} 分别为输入信息、上时刻输出信息和记忆细胞到输入门的权重矩阵; W_{fx} 、 W_{fh} 和 W_{fc} 分别为输入信息、上时刻输出和记忆细胞到遗忘门的权重矩阵; W_{ox} 、 W_{oh} 和 W_{oc} 分别为输入信息、上时刻输出信息和记忆细胞到输出门的权重矩阵; b_i 、 b_f 、 b_o 和 b_c 分别代表了四个元素的偏置量。

2 实验验证与结果分析

本文研究数据采用中国云南某区域内的7个气象站和区域内风电场的气象要素,其中包括风速、气温、大气压力和湿度;总计32维区域内气象特征数据进行风电发电量负荷预测。训练样本数据如图2所示。

2.1 算法步骤

为进一步提升LSTM算法的准确性,

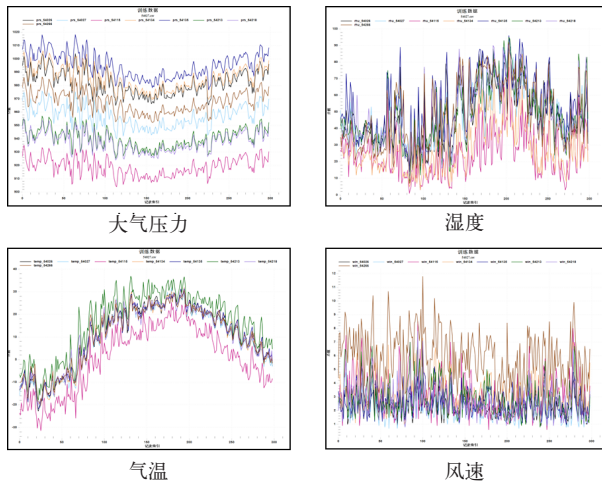


图2 32维共计11 648采样点的训练数据示意图

Fig. 2 Training data of 11 648 sampling points in 32 dimensions

本文引入经验模态分解 (Empirical Mode Decomposition, EMD) 方法对传统的LSTM算法进行优化。

(1) 数据平稳化

气象特征数据的信号类型属于非平稳信号,在此类信号中提取特征信息是建立预测模型的关键,本文采用经验模态分解 (EMD) 方法对信号进行平稳化处理,其具体步骤如下:

1) 通过信号中的极大值和极小值点拟合其包络线,得到 $X_{\max}(t)$ 和 $X_{\min}(t)$ 。

2) 计算上下包络线的平均值,计算公式为

$$m_1(t) = \frac{X_{\max}(t) + X_{\min}(t)}{2} \quad (8)$$

3) 对原始信号和包络线上下平均值作差,得到余下信号 $d_1(t)$,对于平稳信号,可以认为 $d_1(t)$ 就是它的第一个模态函数 (IMF);但是对于非平稳信号,则往往得到的第一个模态函数并不准确,那么就需要再次进行筛选。

4) 对余下信号 $d_1(t)$ 不断进行步骤1)~步骤3)的处理,直到筛分门限值SD小于门限值,则停止筛选。这样就可以得到一个合适的第一阶模态分量 $c_1(t)$,也就是第一个模态函数(IMF),SD的数学表达式为

$$SD = \sum_{t=0}^T \left[\frac{|d_{k-1}(t) - d_k(t)|^2}{d_{k-1}^2(t)} \right] \quad (9)$$

5) 通过对原始信号和第一阶模态分量求差,得到第一阶残差量 $r_1(t)$,再重复步骤1)~步骤4), n 次后得到第 n 阶模态函数 $c_n(t)$ 和符合标准的残差量 $r_n(t)$,则原始信号经EMD分解后的数学表达式为

$$X(t) = \sum_{n=1}^N c_n(t) + r_n(t) \quad (10)$$

由于各个气象要素的物理单位都不相同,而且相差较大,会严重影响预测结果的准确度,在数据预处理的第二步对数据进行了归一化处理,即

$$n(x) = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (11)$$

式中, x 为原始数据; $n(x)$ 为归一化的数据。

将所有数据项均转为0~1之间。

(2) 模型训练

如前文所述,本算法选择使用LSTM模型对数据特征进行自适应训练。根据训练数据的特性,本文所采用的神经网络接受32维的数据输入;串联3个LSTM层,每层的输出维度也同为32维,激活函数采用 $\tan h$ 函数,之后串联一个全连接层产生预测输出,激活函数采用线性函数;为了避免过拟合,在每层神经网络中应用Dropout规则随机选择20%的权值不进行迭代;考虑到优化器的自适应性能,本文选择Adam优化器,目标函数采用均方误差(Mean Squared Error, MSE)函数。

(3) 预测

通过前述的训练过程,模型已经得到了稳定收敛的权值,接下来将同样经过平稳化和归一化处理的待预测数据输入训练好的模型,将模型的输出作为预测结果。

最后将预测结果进行了反归一化处理,得到预测后的负荷值。

2.2 结果分析

为了全面验证算法的实际效果,本文还同时构建了传统BP神经网络模型,并同时风电场采集到的气象数据作为独立建模参数,分别构建了BP-NN全参数模型、LSTM全参数模型、BP-NN风场参数模型和LSTM风场参数模型,并和下一年的真实发电数据进行了测试对比,其对比效果如图3所示。

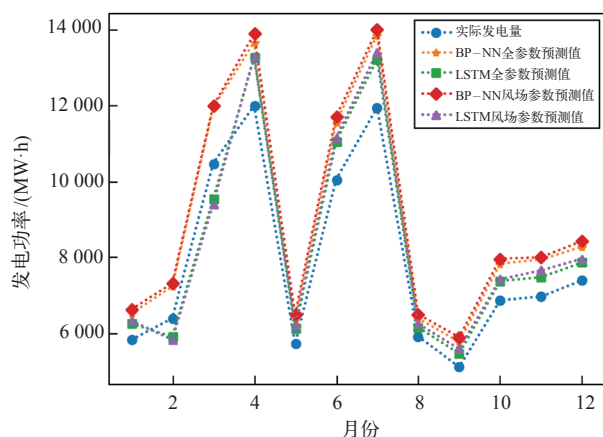


图3 各种模型预测结果

Fig. 3 Prediction results of various models

最后采用平均绝对百分比误差MAPE和平均绝对误差MAE来评估结果的准确性,并进行误差分析,其模型预测误差分析表如下表所示。

由上表可以发现,整体上来看采用LSTM算法构建的模型平均绝对百分比误差和平均绝对误差均比传统的BP-NN模型要小。从参数选择上来说,采用区域范围内全部气象要素建立的预测模型误差要比采用风电场采集的气象要素建立的预测模型误差要小。由此可见,选择LSTM算法对整个区域内全气象参数进行模型预测能够对发电量预测准确度的提高提供很大帮助。

表 不同模型预测结果误差对比

Tab. Comparison of prediction error of different models

模型	第二年	
	MAPE(%)	MAE/MW
BP-NN 全参数预测值	12	1 057.8
LSTM 全参数预测值	4.5	418.5
BP-NN 风场参数预测值	13.4	1177
LSTM 风场参数预测值	5.3	487.7

3 结束语

本文提出的基于ELSTM优化学习算法的区域性风电发电量预测模型,通过某区域内气象点和风电场的关键气象数据进行了预测仿真和测试,证实了该优化学习算法能够有效提高区域内的风电发电量预测准确度。相比于传统的BP神经网络算法,基于ELSTM优化算法能够解决长时间依赖的问题,同时还能够加快模型的收敛速度;通过两种算法不同输入参数的对比分析,验证了本文所提出模型的有效性。其结果显示准确度远远优于传统BP-NN神经网络,并在实际应用的参数选择上提供了一定的指导意义,具备较好的可实用性,同时也为以后的区域性风电场制订长期的电力交易计划奠定良好的基础。

参考文献

- [1] 彭文,王金睿,尹山青. 电力市场中基于Attention-LSTM的短期负荷预测模型[J]. 电网技

- 术, 2019(5): 1745-1751.
- [2] 李军, 杜雪. 稀疏高斯过程在短期风电功率概率预测中的应用[J]. 电机与控制学报, 2019, 23(8): 67-77.
- [3] 郭建英, 孙永全, 王铭义, 等. 基于计算机仿真的风电机组系统可靠性综合[J]. 机械工程学报, 2012, 48(2): 2-8.
- [4] 李刚, 吴潮, 赵建平. 基于改进果蝇神经网络的短期风电功率预测[J]. 测控技术, 2018, 37(7): 27-30, 35.
- [5] 王芝茗. 大规模风电调度技术[M]. 北京: 中国电力出版社, 2016.
- [6] 陈正洪. 风电功率预测预报技术原理及其业务系统[M]. 北京: 气象出版社, 2013.
- [7] 纪浩林, 彭亮. 基于聚类的超闭球算法短期风速预测研究[J]. 测控技术, 2016, 35(8): 138-141, 145.
- [8] 杜雪, 李军. 基于高斯过程的短期风电功率概率预测[J]. 测控技术, 2018, 37(3): 14-17.
- [9] 高阳, 钟宏宇, 葛延峰, 等. 基于GRNN全信息神经网络的超短期风速预测研究[J]. 测控技术, 2016, 35(4): 154-157.
- [10] 黄山园, 李晓明. 基于提高RNN算法聚类效果的研究[J]. 太原师范学院学报(自然科学版), 2012(2): 77-80.
- [11] 张尧, 沈海斌. 非饱和区扩展的RNN算法优化[J]. 传感器与微系统, 2018, 37(3): 41-43.
- [12] 周洪煜, 曾济贫, 王照阳, 等. 基于混沌DNA遗传算法与PSO组合优化的RNN短期风电功率预测[J]. 电力系统保护与控制, 2013(2): 150-155.
- [13] 王敏, 许建, 潘永春. 计及风电预测可靠性的含风电电力系统优化调度模型[J]. 广东电力, 2017, 30(4): 43-49.
- [14] 周任军, 孙嘉贛, 张浩, 等. 考虑双重市场的含风电电力系统双层随机优化调度[J]. 电力系统及其自动化学报, 2018, 30(1): 15-21.
- [15] 王龙. 风电并网的仿真研究[J]. 计算机与数字工程, 2012, 40(2): 144-145.
- [16] 宋弦, 姚刚, 伍士红, 等. 含风电发电系统的负荷优化调度灵活性研究[J]. 计算机与数字工程, 2019, 47(9): 2130-2136.
- [17] 常俊晓, 游文霞, 肖隆恩. 含风电的发电资源优化调度与仿真研究[J]. 计算机仿真, 2015, 32(4): 120-123.
- [18] 李天. 含风电场的电力系统功率预测与优化调度研究[D]. 北京: 华北电力大学, 2018.
- [19] 肖峰, 陈国初. 基于功率预测的含风电场电力系统经济优化调度研究[J]. 电力科学与工程, 2016(32): 9-14.

作者简介



朱尤成

朱尤成 (1972—), 男, 本科, 工程师, 主要研究方向为热工控制技术。

(收稿日期: 2020-07-28) **EA**