基于深度学习分位数回归模型的风电功率概率密度预测

李 彬,彭曙蓉,彭君哲,黄士峻,郑国栋
(长沙理工大学 电气与信息工程学院,湖南 长沙 410114)

摘要:针对风电功率预测问题,在现有预测方法和概率性区间预测的基础上,提出基于深度学习分位数回归的风电功率概率预测方法。该方法采用 Adam 随机梯度下降法在不同分位数条件下对长短期记忆神经网络(LSTM)的输入、遗忘、记忆、输出参数进行估计,得出未来 200 h 内各个时刻风电功率的概率密度函数。根据美国 PJM 网上的风电功率实际数据的仿真结果表明,所提方法不仅能得出较为精确的点预测结果,而且能够获得风电功率完整的概率密度函数预测结果。与神经网络分位数回归相比,其精度更高,且在同等置信度下的预测区间范围更小。

关键词:LSTM;核密度估计;风电功率概率预测;LSTM 分位数回归;概率密度分布
中图分类号:TM 614 **文献标识码**:A DOI:10.16081/j.issn.1006-6047.2018.09.003

0 引言

截至 2016 年年底,中国风电并网容量已达到 148.6 GW[1],然而随着风电在电网中比例的提高,风电的随机性、波动性等缺点也被逐步放大,大规模发展风电给电网带来了巨大挑战。提前精确预测风电功率,可以更好地指导电网发电、调度等工作,以及针对风电爬坡和其他对电网具有较大威胁的风电事件做好预防和消除工作[2]。

统计学习方法和物理方法是目前风电功率预测的 2 种主要方法。其中物理方法主要是利用数值天气预报 NWP(Numerical Weather Prediction),目前基于该方法的预测精度均有限。相较于物理方法,统计学习方法在短期风电功率预测上预测精度较高。统计学习方法又分为点预测(确定性预测)和区间预测(不确定性预测),点预测的预测方法有 IAFSA-BPNN(Improvement Artificial Fish Swarm Algorithm-BP Neural Network)[3]等。文献[4]提出了一种在 3 种不同优化准则下多预测组合模型的权重确定方法,并利用风电场实际数据验证了预测精度得到了有效提高。文献[5]提出了一种在改进局域 Volterra 自适应滤波器基础上的混沌时间序列预测模型。

然而确定性预测不能对风电功率不确定性进行定量描述的缺陷很难克服。在含风电的电网规划、运行和安全稳定分析领域中需要对风电的波动区间进行较精确的估计,仅仅得到单个点的预测值是不够的[6]。而且对风电的概率分布以及置信区间的预测可用于优化风电场的出力可靠性[7]、指导配电网无功规划[8]、进行实时市场调度[9],进一步给电网的发展规划带来更多的信息。如文献[10]采用极限学习机 ELM(Extreme Learning Machine)模型对风电功率进行概率区间预测。

回归分析 RA(Regression Analysis)是确定 2 种及以上变数间相互依赖的定量关系的一种统计分析方法。通常的回归分析方法又称为均值回归分析,均值回归很难代表响应变量服从非对称分布或者散布较大时的情况。Koenker 等提出的分位数回归 QR(Quantile Regression)可以较好地克服均值回归的不足。Taylor 研究了一种神经网络非线性分位数回归模型[11],文献[12-13]将径向基函数(RBF)神经网络分位数回归模型应用于电力负荷概率密度预测,并在考虑实时电价的因素下采用支持向量机(SVM)分位数回归模型对电力负荷进行了概率密度预测。文献[14]提出一种采用因子分子法对 24 h 内小时级的风电功率序列进行降维处理后,再输入神经网络分位数回归预测模型进行风电曲线概率预测的方法。

基于以上分析,为了能得到更加准确的风电功率预测结果,给出更加准确、范围更小的预测区间和更符合风电功率的概率密度分布,提出一种长短期记忆神经网络 LSTM(Long Short Term Memory neural networks)分位数回归概率密度预测方法。该方法可以预测未来风电功率区间以及概率分布,给电网运行带来指导作用。以美国 PJM 网(http://www.pjm.com/markets-and-operations/ops-analysis.aspx)上 MIDATL 地区的风电功率数据作为研究数据,能够得到一定置信度下的未来风电功率预测区间。

1 LSTM 回归

深度学习在机器学习中的应用非常广泛,1997 年 Hochreiter 和 Schmidhuber 提出了一种 LSTM,能够很好地解决序列的长期依赖问题[15],其主要结构如图 1 所示,LSTM 可以由时间展开表示成这种链状结构,LSTM 的重复模块中有 4 个神经网络层。

收稿日期:2018-03-22;**修回日期**:2018-07-22
基金项目:湖南省教育厅创新平台开放基金资助项目(17K001)
Project supported by Hunan Provincial Department of Education Innovation Platform Open Fund(17K001)

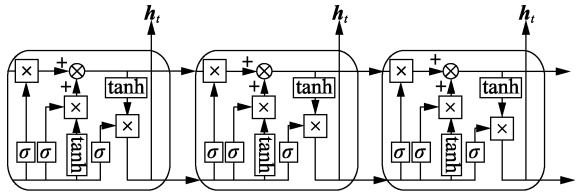


图1 LSTM结构

Fig.1 LSTM structure

图1中一个矩形框称作一个元胞(cell),LSTM有4个门,第一层神经元为遗忘门(forget gate)的Sigmoid控制层,如式(1)所示。

$$\boldsymbol{f}_t=\sigma(\boldsymbol{W}_{\mathrm{f}}[\boldsymbol{h}_{t-1},\boldsymbol{x}_t]+\boldsymbol{b}_{\mathrm{f}}) \tag{1}$$

第二、三层分别为输入门(input gate)和tanh层,分别见式(2)、(3)。输入门的Sigmoid层决定要更新哪些信息;tanh层创造了一个新的候选值$\tilde{\boldsymbol{C}}_t$,该值可能被加入元胞状态中。

$$\boldsymbol{i}_t=\sigma(\boldsymbol{W}_{\mathrm{i}}[\boldsymbol{h}_{t-1},\boldsymbol{x}_t]+\boldsymbol{b}_{\mathrm{i}}) \tag{2}$$

$$\tilde{\boldsymbol{C}}_t=\tanh(\boldsymbol{W}_{\mathrm{C}}[\boldsymbol{h}_{t-1},\boldsymbol{x}_t]+\boldsymbol{b}_{\mathrm{C}}) \tag{3}$$

再将元胞的旧状态$\boldsymbol{C}_{t-1}$更新为$\boldsymbol{C}_t$,见式(4)。

$$\boldsymbol{C}_t=\boldsymbol{f}_t*\boldsymbol{C}_{t-1}+\boldsymbol{i}_t*\tilde{\boldsymbol{C}}_t \tag{4}$$

最后,通过一个如式(5)所示包含Sigmoid层的输出门(output gate)将式(4)通过一个tanh层之后(使得输出值在-1~1之间),再与输出门相乘,这样就将遗忘和记忆参数带至最后的输出。

$$\boldsymbol{O}_t=\sigma(\boldsymbol{W}_{\mathrm{O}}[\boldsymbol{h}_{t-1},\boldsymbol{x}_t]+\boldsymbol{b}_{\mathrm{O}}) \tag{5}$$

$$\boldsymbol{h}_t=\boldsymbol{O}_t\tanh(\boldsymbol{C}_t) \tag{6}$$

其中,$\boldsymbol{W}_{\mathrm{f}}$、$\boldsymbol{b}_{\mathrm{f}}$分别为遗忘门的权重和偏置;$\boldsymbol{W}_{\mathrm{i}}$、$\boldsymbol{b}_{\mathrm{i}}$分别为输入门的权重和偏置;$\boldsymbol{W}_{\mathrm{C}}$、$\boldsymbol{b}_{\mathrm{C}}$分别为更新值的权重和偏置;$\boldsymbol{W}_{\mathrm{O}}$、$\boldsymbol{b}_{\mathrm{O}}$分别为输出门的权重和偏置;$\sigma(\cdot)$为Sigmoid激活函数;$\tanh(\cdot)$为双曲正切激活函数;$[\boldsymbol{h}_{t-1},\boldsymbol{x}_t]$表示行数相等的矩阵或向量进行列合并;"$*$"表示对应元素相乘。

2 条件概率密度预测

2.1 线性分位数回归模型

考虑样本$\boldsymbol{Y}=[\boldsymbol{Y}_1\ \boldsymbol{Y}_2\ \cdots\ \boldsymbol{Y}_N]$和$\boldsymbol{X}=[\boldsymbol{X}_1\ \boldsymbol{X}_2\ \cdots\ \boldsymbol{X}_N]$,其中$\boldsymbol{Y}$为响应变量,$\boldsymbol{X}$为相应的解释变量,$N$为样本量,求线性回归的模型参数可以通过求解式(7)所示的目标函数来得到,其中$\boldsymbol{\beta}$为回归系数。

$$\min\sum_{i=1}^{N}(\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{\beta}) \tag{7}$$

式(7)中参数向量$\boldsymbol{\beta}$的估计可以考虑转化为求解式(8)所示的优化问题。

$$\min_{\boldsymbol{\beta}}\sum_{i=1}^{N}\rho_\tau(\boldsymbol{Y}_i-\boldsymbol{X}_i'\boldsymbol{\beta})=\min_{\boldsymbol{\beta}}\Big[\sum_{i\,|\,\boldsymbol{Y}_i\geqslant\boldsymbol{X}_i'\boldsymbol{\beta}}\tau\,|\,\boldsymbol{Y}_i-\boldsymbol{X}_i'\boldsymbol{\beta}\,|+\sum_{i\,|\,\boldsymbol{Y}_i<\boldsymbol{X}_i'\boldsymbol{\beta}}(1-\tau)\,|\,\boldsymbol{Y}_i-\boldsymbol{X}_i'\boldsymbol{\beta}\,|\Big] \tag{8}$$

其中,$\boldsymbol{X}_i'(i=1,2,\cdots,N)$为在分位数条件下的解释变量;$i\,|\,\boldsymbol{Y}_i\geqslant\boldsymbol{X}_i'\boldsymbol{\beta}$表示第$i$个响应变量实际值大于等于线性回归估计值;分位数$\tau\in(0,1)$;$\rho_\tau(u)=u[\tau-I(u)]$,$I(\cdot)$为示性函数,如式(9)所示。

$$I(u)=\begin{cases}1 & u<0\\ 0 & u\geqslant 0\end{cases} \tag{9}$$

由式(8)得出,在不同的分位数τ下可以得到不同的参数估计$\boldsymbol{\beta}(\tau)$,就能够测算出在不同分位数τ下解释变量对响应变量的条件分位数的影响。τ在(0,1)内连续取值时,就可以得到响应变量的条件分布,然后得到条件密度,最终得出条件密度预测。

2.2 LSTM分位数回归模型

在式(7)所反映的线性回归模型中,限定了解释变量和响应变量之间只能是线性关系。但是在现实中变量之间更多的是非线性关系。Taylor提出了神经网络分位数QRNN(Quantile Regression Neural Network)回归模型:

$$Q_Y(\tau\,|\,\boldsymbol{X})=f(\boldsymbol{X},\boldsymbol{W}(\tau),\boldsymbol{V}(\tau)) \tag{10}$$

其中,$\boldsymbol{W}(\tau)=(w_{ij}(\tau))_{i=1,2,\cdots,I;j=1,2,\cdots,J}$为输入层与隐含层之间的连接权重,$I$为输入层节点数目,$J$为隐含层节点数目;$\boldsymbol{V}(\tau)=(v_{jk}(\tau))_{j=1,2,\cdots,J;k=1,2,\cdots,K}$为隐含层与输出层之间的连接权重,$K$为输出层节点数目。

$$\begin{aligned}\min_{\boldsymbol{W},\boldsymbol{V}}\Big\{\sum_{i=1}^{N}\rho_\tau[\boldsymbol{Y}_i-f(\boldsymbol{X}_i,\boldsymbol{W},\boldsymbol{V})]+\lambda_1\sum_{i,j}w_{ij}^2+\lambda_2\sum_i v_i^2\Big\}=\\ \min_{\boldsymbol{W},\boldsymbol{V}}\Big[\sum_{i\,|\,\boldsymbol{Y}_i\geqslant f(\boldsymbol{X}_i,\boldsymbol{W},\boldsymbol{V})}\tau\,|\,\boldsymbol{Y}_i-f(\boldsymbol{X}_i,\boldsymbol{W},\boldsymbol{V})\,|+\\ \sum_{i\,|\,\boldsymbol{Y}_i<f(\boldsymbol{X}_i,\boldsymbol{W},\boldsymbol{V})}(1-\tau)\,|\,\boldsymbol{Y}_i-f(\boldsymbol{X}_i,\boldsymbol{W},\boldsymbol{V})\,|\Big]+\\ \lambda_1\sum_{i,j}w_{ij}^2+\lambda_2\sum_i v_i^2\end{aligned} \tag{11}$$

其中,λ_1、λ_2为惩罚参数。

分位数神经网络回归可以参考式(8),将其对参数$\boldsymbol{W}(\tau)$、$\boldsymbol{V}(\tau)$的估计转化为求解式(11)的优化问题。式(11)中λ_1、λ_2可以防止模型在训练过程中出现过拟合现象,可以使用Adam随机梯度下降法对式(11)进行求解,估计出参数矩阵$\boldsymbol{W}(\tau)$、$\boldsymbol{V}(\tau)$。

$$\begin{aligned}f_{\mathrm{cost}}=\sum_{i=1}^{N}\rho_\tau[\boldsymbol{Y}_i-f(\boldsymbol{X}_i,\boldsymbol{W},\boldsymbol{b})]=\\ \sum_{i\,|\,\boldsymbol{Y}_i\geqslant f(\boldsymbol{X}_i,\boldsymbol{W},\boldsymbol{b})}\tau\,|\,\boldsymbol{Y}_i-f(\boldsymbol{X}_i,\boldsymbol{W},\boldsymbol{b})\,|+\\ \sum_{i\,|\,\boldsymbol{Y}_i<f(\boldsymbol{X}_i,\boldsymbol{W},\boldsymbol{b})}(1-\tau)\,|\,\boldsymbol{Y}_i-f(\boldsymbol{X}_i,\boldsymbol{W},\boldsymbol{b})\,|\end{aligned} \tag{12}$$

其中,$\boldsymbol{W}=\{\boldsymbol{W}_{\mathrm{f}},\boldsymbol{W}_{\mathrm{i}},\boldsymbol{W}_{\mathrm{C}},\boldsymbol{W}_{\mathrm{O}}\}$为LSTM的权重参数;$\boldsymbol{b}=\{\boldsymbol{b}_{\mathrm{f}},\boldsymbol{b}_{\mathrm{i}},\boldsymbol{b}_{\mathrm{C}},\boldsymbol{b}_{\mathrm{O}}\}$为网络结构的偏置项权重。

根据LSTM回归模型的结构和式(12)所示的神经网络分位数回归方法,可以建立一种LSTM分位数回归模型,并将其对网络结构参数估计的

BPTT(Back Propagation Through Time)算法过程的代价函数转化为如式(13)所示的分位数回归的目标函数。最终可以将参数估计看作式(13)所示的优化问题,并用 Adam 随机梯度下降法求解该优化问题。

$$\min_{\boldsymbol{W},\boldsymbol{b}}\left[f_{\text{cost}}+\frac{\lambda}{2}|(\hat{\boldsymbol{W}},\hat{\boldsymbol{b}})|^2\right] \tag{13}$$

其中,$\hat{\boldsymbol{W}}(\tau)$、$\hat{\boldsymbol{b}}(\tau)$分别为带分位数条件的 $\boldsymbol{W}$、$\boldsymbol{b}$。求解出参数 $\hat{\boldsymbol{W}}(\tau)$、$\hat{\boldsymbol{b}}(\tau)$后,代入式(14)中可以得到 $\boldsymbol{Y}$ 的条件分位数估计。

$$\hat{Q}_{\boldsymbol{Y}}(\tau|\boldsymbol{X})=f(\boldsymbol{X},\hat{\boldsymbol{W}}(\tau),\hat{\boldsymbol{b}}(\tau)) \tag{14}$$

当 τ 在(0,1)上连续取值时,条件分位数曲线 $\hat{Q}_{\boldsymbol{Y}}(\tau|\boldsymbol{X})$就被称为条件分布(累计),文献[16]从分布函数 $F(F^{-1}(\tau))=\tau$ 出发推导出条件密度预测,如式(15)所示。

$$P(\hat{Q}_{\boldsymbol{Y}}(\tau|\boldsymbol{X}))=\frac{\mathrm{d}\tau}{\mathrm{d}\hat{Q}_{\boldsymbol{Y}}(\tau|\boldsymbol{X})} \tag{15}$$

接下来对式(15)执行 $\boldsymbol{X}$ 条件化、τ 离散化操作,最后采用密度估计就可以得到 $\boldsymbol{Y}$ 的条件密度预测 $P(\hat{Q}_{\boldsymbol{Y}}(\tau|\boldsymbol{X}))$。

LSTM 回归对非线性的时间序列具有很好的拟合能力,在道路运输相关的交通流速预测[17]、医学相关的蛋白质二级结构序列预测[18]等问题上都有较好的预测效果。概率密度估计主要是为了得到概率密度曲线,使得电网工作人员能更好地了解未来风电功率波动范围,获得更多的有用信息。利用分位数回归可以将二者有机结合起来,在不同分位数下进行 LSTM 回归可以得到多个点预测的结果,进而继续采用高斯核进行概率密度估计,得到概率密度函数,具体程序流程图见图 2。

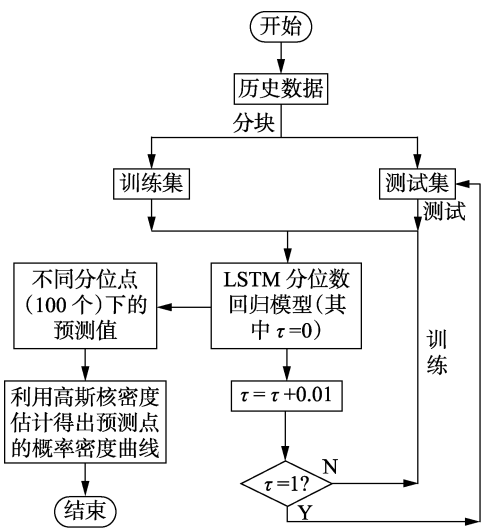


图 2　程序流程图

Fig.2　Program flowchart

2.3　模型评价

风电功率点预测模型常用的评价指标有平均绝对误差(MAE)、均方误差(MSE)、均方根误差(RMSE)。从风电功率特性出发,给出了一种比较 2 个风电功率点预测模型的检验方法——DM(Diebold-Mariano)检验。该检验方法在假设两模型误差相等并服从正态分布的基础上做假设检验,当 DM 检验绝对值大于某个阈值时证明 2 个模型的预测结果有显著差异[19]。

然而上述评价指标并不能用来评价概率预测的结果。考虑风电功率的随机性强、波动范围大的情况,提出一种非参数模型的风电概率区间预测方法,并给出参考风电功率特性的概率区间评估指标[20],采用如下 2 个评价指标。

a. 可靠性指标。

在置信度 1−α 下,一共有 N_{interval}个预测区间,可靠性评价为:

$$\begin{cases} I_i^{\alpha}=[L_i^{\alpha},U_i^{\alpha}] \\ \xi_i^{\alpha}=\begin{cases}0 & P^i\notin I_i^{\alpha}\\ 1 & P^i\in I_i^{\alpha}\end{cases} \quad i=1,2,\cdots,N_{\text{interval}} \\ R_{\text{cover}}=\dfrac{1}{N_{\text{interval}}}\displaystyle\sum_{i=1}^{N_{\text{interval}}}\xi_i^{\alpha} \end{cases} \tag{16}$$

其中,L_i^{α} 为在上述置信度下第 i 个预测区间的下界;U_i^{α} 为对应的上界;I_i^{α} 为对应区间;P^i 为对应实际点的值;R_{cover}为可靠性指标;ξ_i^{α} 表示第 i 个真实值是否落在预测区间内,即风电功率预测区间对真实值的覆盖率。

风电功率实际落在预测区间内的概率应该等于或接近事先给定的置信度,可靠性指标越接近置信区间值则代表该预测模型越可靠。

b. 敏锐性指标。

可靠性指标不能全面体现概率预测结果的好坏,因为区间很大时可靠性通常会更高,相应的区间宽度过大会导致能提供的有用信息较少,所以还需要敏锐性指标来共同判断区间预测结果的好坏(区间平均宽度):

$$\begin{cases} \delta_i^{\alpha}=U_i^{\alpha}-L_i^{\alpha} \\ \delta_{\text{mean}}^{\alpha}=\dfrac{1}{N_{\text{interval}}}\displaystyle\sum_{i=1}^{N_{\text{interval}}}\delta_i^{\alpha} \end{cases} \tag{17}$$

其中,δ_i^{α} 为第 i 个预测区间的宽度;$\delta_{\text{mean}}^{\alpha}$为敏锐性指标。预测风电功率的敏锐性指标越小代表得到的风电功率区间越小,从而可以得到更多未来时间段的有效风电功率信息。

单一的敏锐性指标和可靠性指标都不能全面反映风电功率概率预测模型的好坏,只有综合敏锐性指标与可靠性指标才可以充分反映概率区间预测结

果的优劣。

3 算例仿真

美国PJM网上MIDATL地区2014年8月1日至2015年9月1日的风力发电量如图3所示。以2014-01-08T04:00:00至2015-01-01T13:00:00时间段内的3 716个时间点为训练样本,提前2 h预测之后200个时间点的风电功率。本文仿真的实验计算机条件是CPU酷睿i7-7700、内存8 GB、GPU1050Ti 4 GB。通过Keras深度学习框架将每个分位数下的LSTM迭代100个轮次(epochs),LSTM结构为64个门结构。训练之前按式(18)将数据进行归一化,这样可以提高模型的准确率。

$$\boldsymbol{X}_t=\frac{\boldsymbol{X}_t-\boldsymbol{X}_{\min}}{\boldsymbol{X}_{\max}-\boldsymbol{X}_{\min}} \tag{18}$$

其中,$\boldsymbol{X}_t$为t时刻的样本向量;$\boldsymbol{X}_{\min}$、$\boldsymbol{X}_{\max}$分别为所有样本的最小值和最大值。

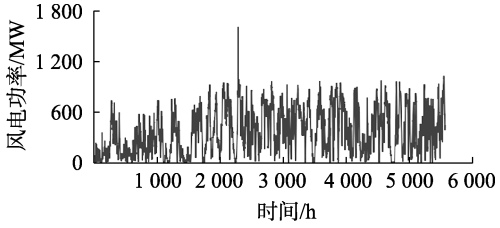


图3 美国风电功率数据

Fig.3 American wind power data

根据上述内容,用QRLSTM(Quantile Regression Long Short Term Memory neural networks)回归预测模型得到从2015-01-01T16:00:00到2015-01-09T21:00:00总共200个时间点的预测结果,每个时间点间隔1 h。为了体现QRLSTM回归预测模型的预测准确度,对比所用的QRNN回归预测模型所得的预测结果和预测区间,QRNN选取隐含层节点数为64,惩罚参数为1。

QRLSTM的可靠性指标为84.16%,敏锐性指标为124.54,QRNN对应指标分别为79.12%和191.99,指标数据与图4对应的QRLSTM与QRNN预测结果一致。QRLSTM可靠性指标比QRNN高5.04%。虽然理论上可靠性应该与置信度相等,且上述2个模型的可靠性都没有达到置信度90%,但是比较接近置信度的结果也是可以接受的预测效果。QRNN比QRLSTM的敏锐度指标高出54.08%,如此大的差距充分体现了QRLSTM的优势,不仅预测的可靠性高于QRNN,且能在预测结果中提供更多的有效信息。

通过2个模型指标数据和图4的对比明显看出真实值很大概率落在QRLSTM回归的预测区间内。相比QRNN回归模型,真实值落在QRNN预测区间内的概率明显要小很多,且预测的区间宽度比

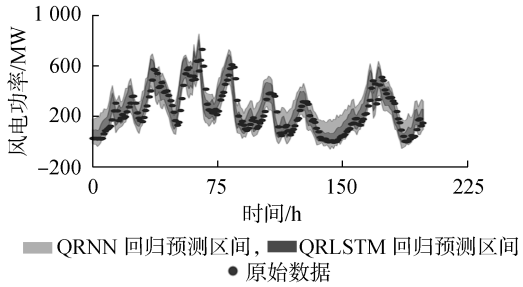


图4 QRLSTM与QRNN风电功率预测结果对比

Fig.4 Comparison of wind power forecasting results between QRLSTM and QRNN

QRNN的预测区间小很多,这充分说明了本文提出的QRLSTM回归模型可很好地预测风电功率的波动性,且可以预测较长时间的风电功率波动性。

采用QRLSTM回归方法可以得到预测点的概率密度曲线,从预测点0~200中随机抽取第4、47、90、104、138、168这6个时间点的概率密度函数分布区间见附录中图A1,其中第104个时间点的概率密度函数见图5。从预测出的概率密度函数可以看出,QRLSTM可以预测出风电功率的完整概率密度分布,且真实值都落在该密度函数中间。以上示例说明该方法能够给出未来预测时间点概率密度曲线。

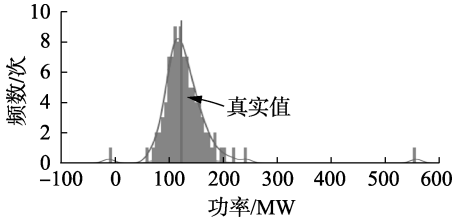


图5 QRLSTM风电功率预测概率密度函数

Fig.5 Probability density function of wind power forecasting for QRLSTM

附录中表A1展示了QRLSTM与QRNN预测模型前24个预测点的预测区间和预测区间范围的差值。从数据中也可以看出,QRLSTM预测模型预测出的概率区间比QRNN预测出的概率区间范围小很多。

图6对比了另外2种点预测的结果和LSTM分

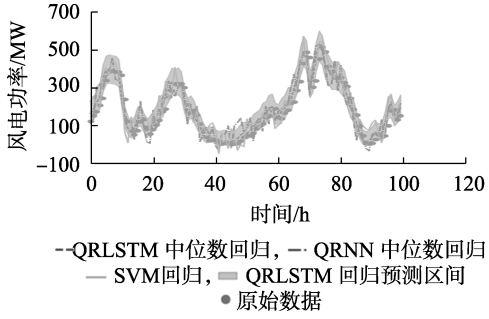


图6 QRLSTM区间预测与点预测结果对比

Fig.6 Comparison between QRLSTM interval prediction and point prediction

位数回归的区间预测结果,可看出相比于点预测,区间预测的可信度更高,在实际使用预测结果时可以更好地规避误差带来的风险。QRLSTM、QRNN、SVM 这 3 种点预测模型平均绝对误差分别为 34.8、52.8、60.09,可以看出 QRLSTM 中位数预测的准确率比 QRNN 和 SVM 都要高。

4　结论

本文提出一种 QRLSTM 预测模型,其可靠性指标比 QRNN 高 5.04%,敏锐性指标比 QRNN 低 54.08%,充分体现了 QRLSTM 的优势,能够提供更多的有效信息给电力行业使用者。

该方法不仅在预测精度上有提高,而且能获得风电功率概率密度函数,可以为电力系统增强风电消纳能力提供更多的有用信息,便于风电的发展,促进环保和解决能源不足问题。

附录见本刊网络版(http://www.epae.cn)。

参考文献:

[1] 国家能源局. 2016 年风电并网运行情况[EB/OL]. (2016-08-20)[2017-12-19]. http://www.nea.gov.cn/201701/26/c_136014615.htm.

[2] 钱政,裴岩,曹利宵,等. 风电功率预测方法综述[J]. 高电压技术,2016,42(4):1047-1060.
QIAN Zheng, PEI Yan, CAO Lixiao, et al. Review of wind power forecasting method[J]. High Voltage Engineering, 2016, 42(4): 1047-1060.

[3] 张颖超,王雅晨,邓华,等. 基于 IAFSA-BPNN 的短期风电功率预测[J]. 电力系统保护与控制,2017,45(7):58-63.
ZHANG Yingchao, WANG Yachen, DENG Hua, et al. IAFSA-BPNN for wind power probabilistic forecasting[J]. Power System Protection and Control, 2017, 45(7): 58-63.

[4] 张露,卢继平,梅亦蕾,等. 基于不同优化准则的风电功率预测[J]. 电力自动化设备,2015,35(5):139-145.
ZHANG Lu, LU Jiping, MEI Yilei, et al. Wind power forecasting based on different optimization criterions[J]. Electric Power Automation Equipment, 2015, 35(5): 139-145.

[5] 王兰,李华强,吴星,等. 基于改进局域 Volterra 自适应滤波器的风电功率混沌时间序列预测模型[J]. 电力自动化设备,2016,36(8):40-44.
WANG Lan, LI Huaqiang, WU Xing, et al. Wind power chaotic time series prediction model based on improved local Volterra adaptive filter[J]. Electric Power Automation Equipment, 2016, 36(8): 40-44.

[6] 吴问足,乔颖,鲁宗相,等. 风电功率概率预测方法及展望[J]. 电力系统自动化,2017,41(18):167-175.
WU Wenzu, QIAO Ying, LU Zongxiang, et al. Methods and prospects for probabilistic forecasting of wind power[J]. Automation of Electric Power Systems, 2017, 41(18): 167-175.

[7] 叶瑞丽,郭志忠,刘瑞叶,等. 基于置信区间估计及储能装置优化配置的风电场可靠出力研究[J]. 电力自动化设备,2017,37(5):85-91.
YE Ruili, GUO Zhizhong, LIU Ruiye, et al. Reliable power output based on confidence interval estimation and optimal ESS configuration of wind farm[J]. Electric Power Automation Equipment, 2017, 37(5): 85-91.

[8] 谢俊,王璐,傅旭华,等. 考虑风电功率概率分布不确定性的含风电配电网无功规划方法[J]. 电力自动化设备,2016,36(6):40-47.
XIE Jun, WANG Lu, FU Xuhua, et al. Reactive power planning with consideration of wind power probability distribution uncertainty for distribution network[J]. Electric Power Automation Equipment, 2016, 36(6): 40-47.

[9] 江岳文,温步瀛. 结合风电功率超短期预测值偏差的实时市场调度[J]. 电力自动化设备,2015,35(3):12-17.
JIANG Yuewen, WEN Buying. Real-time market dispatch based on ultra-short-term forecast error of wind power[J]. Electric Power Automation Equipment, 2015, 35(3): 12-17.

[10] WAN Can, XU Zhao, PINSON P. Probabilistic forecasting of wind power generation using extreme learning machine[J]. IEEE Transactions on Power Systems, 2014, 29(3): 1033-1044.

[11] TAYLOR J W. A quantile regression neural network approach to estimating the conditional density of multiperiod returns[J]. Journal of Forecasting, 2000, 19(4): 299-311.

[12] 何耀耀,许启发,杨善林,等. 基于 RBF 神经网络分位数回归的电力负荷概率密度预测方法[J]. 中国电机工程学报,2013,33(1):93-98.
HE Yaoyao, XU Qifa, YANG Shanlin, et al. A power load probability density forecasting method based on RBF neural network quantile regression[J]. Proceedings of the CSEE, 2013, 33(1): 93-98.

[13] 何耀耀,刘瑞,撖奥洋. 基于实时电价与支持向量分位数回归的短期电力负荷概率密度预测方法[J]. 中国电机工程学报,2017,37(3):768-776.
HE Yaoyao, LIU Rui, HAN Aoyang. Short-term power load probability density forecasting method based on real time price and support vector quantile regression[J]. Proceedings of the CSEE, 2017, 37(3): 768-776.

[14] 李丹,任洲洋,颜伟,等. 基于因子分析和神经网络分位数回归的月度风电功率曲线概率预测[J]. 中国电机工程学报,2017,37(18):5238-5248.
LI Dan, REN Zhouyang, YAN Wei, et al. Month-ahead wind power curve probabilistic prediction based on factor analysis and quantile regression neural network[J]. Proceedings of the CSEE, 2017, 37(18): 5238-5248.

[15] LI X, WU X. Constructing long short-term memory based deep recurrent neural networks for large vocabulary speech recognition[C]// IEEE International Conference on Acoustic, Speech and Signal Processing. Brisbane, QLD, Australia: IEEE, 2015: 4520-4524.

[16] 许启发,徐金菊,蒋翠侠,等. 基于神经网络分位数回归的 VaR 金融风险测度[J]. 合肥工业大学学报(自然科学版),2014,37(12):1518-1522.
XU Qifa, XU Jinju, JIANG Cuixia, et al. Financial risk measure VaR based on quantile regression neural network[J]. Journal of Hefei University of Technology (Natural Science), 2014, 37(12): 1518-1522.

[17] MA X, TAO Z, WANG Y, et al. Long short-term memory neural network for traffic speed prediction using remote microwave sensor data[J]. Transportation Research Part C Emerging Technologies, 2015, 54: 187-197.

[18] HANSON J, YANG Y, PALIWAL K, et al. Improving protein disorder prediction by deep bidirectional long short-term memory recurrent neural networks[J]. Bioinformatics, 2017, 33(5): 685.
[19] CHEN Hao, WAN Qiulan, WANG Yurong. Refined diebold-mariano test methods for the evaluation of wind power forecasting models [J]. Energies, 2014, 7: 4185-4198.
[20] PINSON P, NIELSEN H A, MØLLER J K, et al. Non-parametric probabilistic forecasts of wind power: required properties and evaluation[J]. Wind Energy, 2007, 10(6): 497-516.

作者简介：


李 彬

李 彬(1993—)，男，湖南常德人，硕士研究生，主要研究方向为智能电网中的信息处理(**E-mail**：495681419@qq.com)；

彭曙蓉(1975—)，女，湖南长沙人，副教授，博士，主要研究方向为智能信息处理(**E-mail**：173764138@qq.com)。

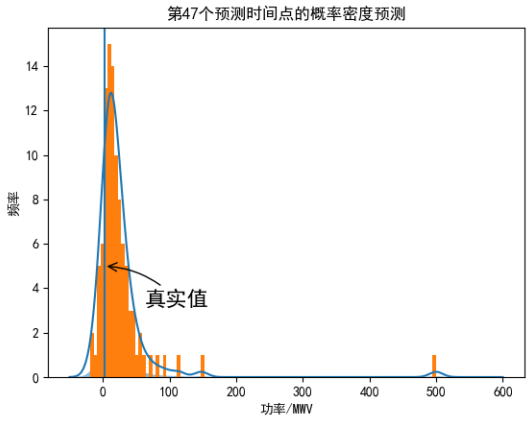
Wind power probability density forecasting based on deep learning quantile regression model

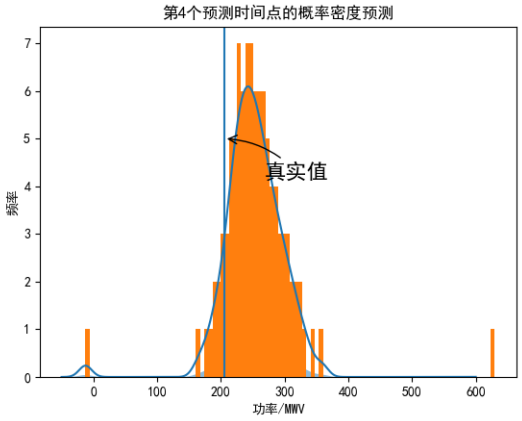
LI Bin, PENG Shurong, PENG Junzhe, HUANG Shijun, ZHENG Guodong
(School of Electrical & Information Engineering, Changsha University of Science & Technology, Changsha 410114, China)

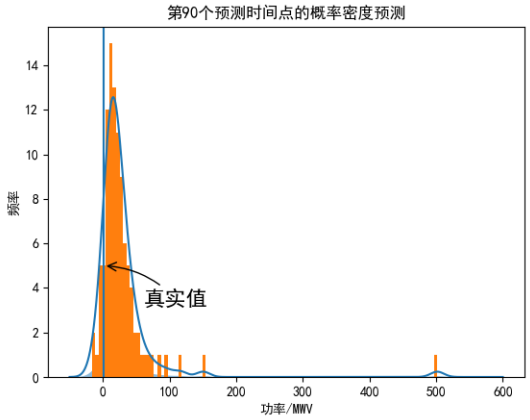
Abstract: Aiming at the problem of wind power forecasting, a deep learning quantile regression based method is proposed based on the existing forecasting methods and probabilistic interval forecasting. The method adopts Adam random gradient descent method to estimate the input, forgetting, memory and output parameters of LSTM (Long Short Term Memory neural networks) under different quantiles. The probability density function of wind power at each moment within 200 h in the future is obtained. According to the actual wind power data from PJM network in the United States, the simulative results show that the proposed method can obtain accurate point forecasting results and complete probability density function forecasting results of wind power, compared with the quantile regression of neural network, it has higher accuracy and smaller range of forecasting interval under the same confidence.

Key words: LSTM; nuclear density estimation; wind power probability prediction; LSTM quantile regression; probability density distribution

附录


第47个预测时间点的概率密度预测
频率
功率/MWV
真实值


第4个预测时间点的概率密度预测
频率
功率/MWV
真实值


第90个预测时间点的概率密度预测
频率
功率/MWV
真实值

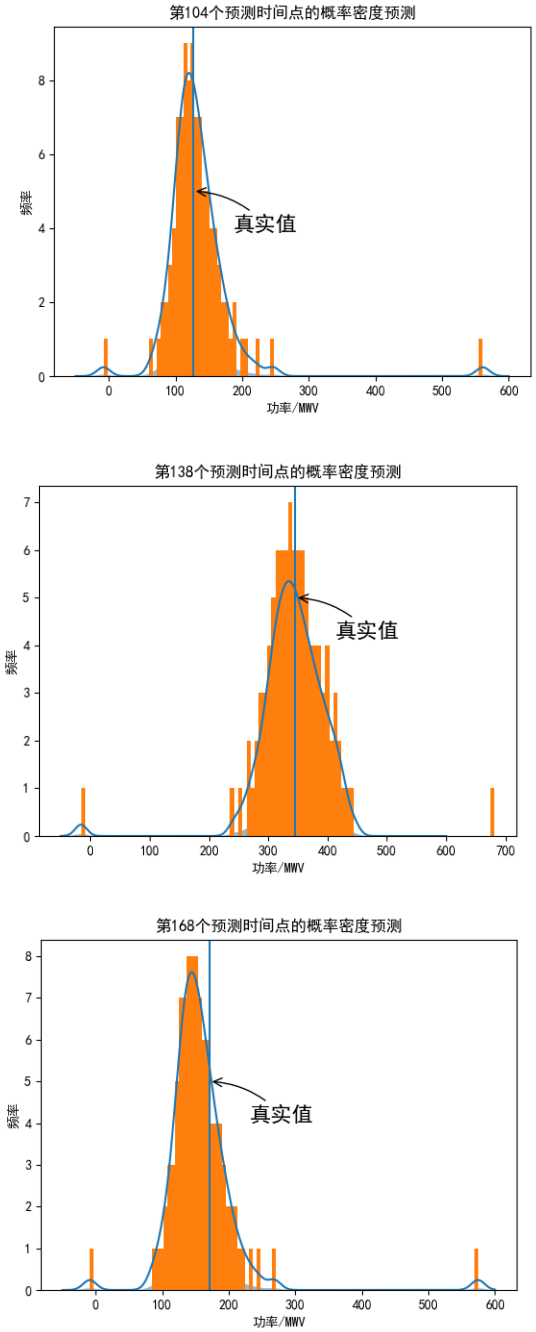


图 A1　QRLSTM 风电功率预测概率密度函数

Fig.A1 Probability density function of wind power forecasting for QRLSTM

表 A1 RQLSTM 与 RQNN 预测区间比较

Table A1 Predicted interval comparison between RQLSTM and RQNN

RQLSTM 预测区间	RQNN 预测区间	范围差值
[113.3,227.5]	[92.0,276.8]	71
[131.6,250.3]	[109.0,296.2]	68
[160.2,258.2]	[135.5,326.3]	65
[160.2,323.0]	[164.5,358.9]	63
[276.4,422.1]	[244.1,447.0]	57
[316.5,467.0]	[281.8,488.0]	56
[322.0,473.0]	[286.9,493.6]	56
[317.2,467.7]	[282.4,488.7]	56
[315.9,466.3]	[281.2,487.4]	56
[300.2,448.9]	[266.5,471.4]	56
[187.6,318.6]	[161.0,355.0]	63
[88.6,196.1]	[69.1,250.5]	74
[47.7,143.2]	[31.3,206.6]	80
[31.8,122.2]	[16.6,189.4]	82
[48.9,144.8]	[32.5,207.9]	80
[87.6,194.9]	[68.2,249.4]	74
[96.6,206.3]	[76.5,258.9]	73
[63.5,163.9]	[45.9,223.6]	77
[34.8,126.2]	[19.5,192.7]	82
[52.6,149.6]	[35.9,211.9]	79
[72.0,174.8]	[53.8,232.7]	76
[94.5,203.6]	[74.5,256.7]	73
[131.9,250.6]	[109.2,296.4]	68
[142.7,264.0]	[119.2,307.8]	67