DOI: 10.7500/AEPS20211208003

# 基于时序混合密度网络的超短期风电功率概率预测

董骁翀<sup>1</sup>,孙英云<sup>1</sup>,蒲天骄<sup>2</sup>,王新迎<sup>2</sup>,李 烨<sup>2</sup>

(1. 华北电力大学电气与电子工程学院,北京市 102206; 2. 中国电力科学研究院有限公司,北京市 100192)

摘要:风电概率预测能够为新型电力系统安全运行提供关键的边界条件。提高预测精度是风电概率预测研究的关键问题,并且提高隐式模型的可解释性有益于人工智能模型的推广应用。因此,文中提出了时序混合密度网络,提取风电时序数据的局部矩信息作为输入通道,采用时序卷积网络提取多时间尺度的概率特征,并使用混合Beta分布构建概率预测信息。算例结果表明,局部矩通道能有效提高模型训练的收敛性,并且由时序混合密度网络提取的混合分布参数具有一定的可解释性,其预测结果相比现有模型具有更高的精度。

关键词:风电功率;概率预测;混合密度网络;时序卷积网络;最大似然估计;可解释性

## 0 引言

随着中国"碳达峰·碳中和"目标的提出,风力发 电将成为绿色低碳能源结构转型的重要支撑。由于 风力发电受一次能源风能的影响,其出力存在不确 定性。因此,精确可靠的概率预测是量化风电出力 不确定性的必要工具,也是推动新型电力系统发展 的重要技术<sup>[1]</sup>。

概率预测方法根据表达形式可分为预测区间 (prediction interval, PI)、分位数回归(quantile regression,QR)和概率密度函数(probability density function,PDF)。PI预测包含对应置信区间的上下 边界<sup>[2-3]</sup>。QR预测采用分位数信息描述预测概率分 布信息<sup>[4-5]</sup>。PI和QR预测仅提供区间信息,不能连 续描述概率预测信息。PDF预测可使用参数法对 预测值的概率分布进行建模,也可通过非参数法的 核密度估计或混合分布模型建立 PDF<sup>[6-7]</sup>。

依据建模方法,风电概率预测可分为统计学方 法和人工智能方法。统计学方法对历史信息与预测 概率分布之间的函数关系进行显式建模,主要包含 高斯过程模型<sup>[8]</sup>、马尔可夫链模型<sup>[9]</sup>、时间序列模 型<sup>[10]</sup>、Copula函数模型<sup>[11]</sup>等。随着风电概率预测中 影响因素的多元复杂化,传统统计学方法难以有效 量化相关因素。而深度神经网络具有良好的高维非 线性映射能力,基于人工智能方法的概率预测近年 来得到了广泛关注<sup>[12]</sup>。通过最大后验估计 (maximum a posteriori estimation, MAP)或最大似 然估计(maximum likelihood estimation, MLE)对神 经网络进行训练可以实现 PDF 预测<sup>[13]</sup>。贝叶斯神 经网络在给定先验分布后通过变分推理训练实现 MAP,但先验分布参数较难确定<sup>[14]</sup>。MLE可作为 神经网络损失函数引导网络拟合 PDF 的参数<sup>[15]</sup>。

为了有效解决参数法适用性较差的缺点,混合 密度网络(mixture density network, MDN)被提出用 于拟合混合分布,灵活描述预测 PDF<sup>[16]</sup>。MDN 能 够有效结合各类网络结构实现风电概率预测。文献 [17] 融合卷积神经网络和门控循环单元提取时空特 征,使用MDN 拟合混合高斯分布实现短期风电功 率概率预测。文献[18]使用MDN 拟合混合Beta分 布,相比混合高斯分布能有效防止密度泄漏。文献 [19]采用包含注意力机制的MDN模型融合多元气 象信息进行概率预测。现有研究中MDN对子分布 参数的提取依靠对网络输出参数的直接拆分,难以 说明混合分布中参数提取的原理及在风电概率预测 中的意义,缺乏可解释性。另外,风电功率的概率特 征存在持续性,能为概率预测提供可靠信息,而现有 风电概率预测研究中缺少对多时间尺度概率特征的 提取与融合<sup>[20]</sup>。

针对以上问题,本文提出了一种基于时序混合 密度网络(temporal mixture density network, TMDN)的超短期风电概率预测模型。TMDN在输 入层采用局部矩通道的预处理方法,从风电时序数 据中提取局部概率特征作为网络的输入信息。并且 TMDN采用多层时序卷积结构提取多时间尺度的

收稿日期: 2021-12-08; 修回日期: 2022-02-15。

上网日期: 2022-04-11。

国家重点研发计划资助项目(2020YFB0905900);国家自然 科学基金资助项目(51777065)。

概率特征,通过混合Beta分布推演概率预测,有效 提高了模型的可解释性。算例将本文所提方法与现 有方法进行了对比,表明TMDN具有更高的预测 精度。

#### 1 混合密度网络

MDN采用神经网络映射输入数据与混合分布 参数之间的关系,从而实现对条件概率分布的拟 合<sup>[21]</sup>。MDN融合了人工智能与非参数估计方法的 优势,混合分布通过多个子分布参数的灵活组合具 有更强的拟合能力,并且神经网络可以映射高维非 线性关系。

MDN的基本结构如附录A图A1所示,以混合 Beta分布为例,采用全连接网络映射混合分布参数。输入层输入条件信息,输出层输出混合Beta分 布的权重参数 $\omega$ 与形状参数 $\alpha$ 、 $\beta$ 。混合Beta分布的 概率密度函数为:

$$P(y) = \sum_{i=1}^{m} \omega_i(x) \frac{\Gamma(\alpha_i(x) + \beta_i(x))}{\Gamma(\alpha_i(x))\Gamma(\beta_i(x))} \cdot y^{\alpha_i(x) - 1} (1 - y)^{\beta_i(x) - 1}$$
(1)

式中: $P(\cdot)$ 表示概率密度函数;y为目标值;x为输入 值;m为混合分布数量; $\omega_i(x) \in \mathbf{R}^+$ 为第i个 Beta分 布 的 权 重 参 数 ,其 需 要 满 足  $\sum \omega_i(x) = 1$ ;  $\alpha_i(x) \in \mathbf{R}^+$ 和 $\beta_i(x) \in \mathbf{R}^+$ 为第i个 Beta分布的形状 参数; $\Gamma(\cdot)$ 表示 Gamma 函数。

Beta 分布为定义在(0,1)区间上的有界分布, 相比 Gaussian 分布, Beta 分布可有效防止密度泄漏。 为了满足混合分布参数的固有约束, MDN 的输出 层可通过 Softmax 函数保证权重参数 $\omega_i(x)$ 为正实 数且和为1,即

$$\omega_i(x) = \operatorname{Softmax}(h_i(x)) = \frac{\exp(h_i(x))}{\sum_i \exp(h_i(x))} \quad (2)$$

式中: $h_i(x)$ 为MDN输出层信息。

为了保障形状参数 $\alpha_i(x)$ 和 $\beta_i(x)$ 为正实数,可 将数据通过Softplus激活函数,即

$$Softplus(x) = \ln(1 + \exp(x))$$
(3)

混合 Beta 分布的期望值 E(y)可作为确定性预 测值,即

$$E(y) = \sum_{i=1}^{m} \omega_i(x) \frac{\alpha_i(x)}{\alpha_i(x) + \beta_i(x)}$$
(4)

当使用 MDN 进行风电功率概率预测时,取 L 步的风电功率实测值作为输入,预测 K 步后的风电 功率概率分布,该问题可以用式(5)表示。

$$f(\boldsymbol{x}_{t},\boldsymbol{\theta}) = P(\boldsymbol{y}_{t+K}) \tag{5}$$

射函数。 假设有 n组风电实测数据作为 MDN 训练数据, 由 输 入 值  $X = [x_1, x_2, \dots, x_j, \dots, x_n]$ 以及 目标 值  $Y = [y_1, y_2, \dots, y_j, \dots, y_n]$ 构成训练集,是典型的监 督学习问题。概率预测模型以 X 作为 MDN 的输 入,希望由 MDN 输出参数构成的混合 Beta 分布在 目标值 Y处概率值最大。设 Y 中的目标值均为独 立分布, MDN 的损失函数  $l(\theta|X, Y)$  可由对数似然 函数 ln  $P(Y|\theta, X)$ 表示,即

 $l(\theta|X, Y) = \ln P(Y|\theta, X) =$ 

$$\ln \prod_{j=1}^{n} P(y_j | \theta, x_j) = \sum_{j=1}^{n} \ln P(y_j | \theta, x_j) \quad (6)$$

为了使 $P(Y|\theta, X)$ 具有极大值,MDN可以通过 MLE训练网络获得最优网络参数 $\theta^*$ 。

$$\theta^* = \arg \max l(\theta | X, Y) \tag{7}$$

#### 2 时序混合密度网络

#### 2.1 时序数据局部矩通道

卷积神经网络的多通道具有提取不同特征的作用,但特征意义较为模糊。本文通过局部矩计算对时序数据进行预处理,提取概率特征明确的局部矩信息构建通道,为MDN拟合预测概率分布提供有效的数据支撑。

局部矩信息可以反映局部数据的概率特征,通 过局部矩信息的变化特征能够有效推演概率预测。 定义时序数据的局部 n 阶矩为在一个长度为 k 的窗 口内计算 n 阶矩。并通过窗口在时序数据上滑动, 计算不同时刻点的局部矩。时序数据局部矩的计算 过程如图 1 所示。图中: x 为窗口内 x 的均值; µ1 和 µ2 分别为局部一阶矩与局部二阶中心矩。





由图1可知,时序数据经过局部矩计算后将得 到一组新的时序数据。为了保证数据维度的一致 性,在求解时需对时序数据两端进行重复填充,即使 用时序数据的边缘值扩张数据尺寸。局部矩的求解 与卷积的不同之处在于:局部矩的滑动窗口计算是 固定的解析运算,不包含可学习的参数,具有明确的 概率特征意义。不同局部矩信息表征着不同的概率 特征作为 MDN 的输入通道。为了保障输入数据的 量纲一致性,取局部二阶中心矩的算数平方根 µ̂2 作 为输入数据。

时序数据的局部矩反映了时序数据在概率特征 上的变化情况。局部一阶矩描述的是时序数据在窗 口内的平均值,其可以刻画非平稳时间序列的均值 变化趋势。而局部二阶中心矩描述的是非平稳时序 数据在窗口内的方差,可以刻画波动变化趋势。

# 2.2 时序卷积网络

时序卷积网络(temporal convolutional network, TCN)融合传统卷积网络结构提取时序数据特 征<sup>[22]</sup>。相较于循环神经网络,TCN能够以非循环形 式处理时序数据,在避免梯度消失/爆炸问题的同时 有着良好的并行计算能力。TCN可以依据学习问 题的形式灵活地设计网络结构,提取出针对目标的 时序特征。本文模型设计中采用因果卷积运算:

$$g_{l} * x_{l} = \sum_{l=0}^{J-1} g_{ll} x_{l-dl}$$
(8)

式中:g<sub>i</sub>为滤波器的输出,g<sub>u</sub>表示其步长为*l*时的元 素;"\*"表示卷积运算;*J*为卷积核的长度;*d*为膨胀 系数。从式(8)中可知,因果卷积的输出仅依赖之前 的时序数据,具有严格的因果关系。

风电出力的概率特征具有持续性,因此未来风 电出力的概率特征与历史特征存在相关性。若风电 出力在短时间内存在较大波动,该波动特性在未来 出力中具有一定持续性。但短时间内得到的概率特 征难以表征更长时间尺度上的稳定性及趋势变化。 因此,风电功率概率预测需要综合考虑多时间尺度 的概率特征对预测点带来的影响。本文在结构设计 中采用融合多层 TCN输出信息的结构,见图2。

从图2中可知,顶层TCN能够提取[t-7,t]时 段内的概率特征,而底层TCN仅能感知[t-1,t] 时段内的概率特征。堆叠多层TCN可以扩大网络 在时序数据上的感受野,提取时间尺度更长的概率 特征。为了融合多时间尺度的概率特征,模型将每 层TCN输出的尾部特征值作为一个子分布的参数 特征。不同 Beta分布的参数特征来源于不同的 TCN层,从而MDN的预测结果可以融合多时间尺



度的概率特征构成混合Beta分布。

### 2.3 TMDN结构

结合时序数据局部矩通道与TCN,本文提出了 基于TMDN的风电超短期概率预测模型,网络结构 如图3所示。



首先,时序数据经过局部矩计算构建局部矩通 道,结合原始时序数据作为TMDN的输入信息。其 次,TMDN结构整合TCN层、权重正则化(weight normalization,WN)层、激活函数LeakyReLU作为时 序卷积模块提取概率特征。TCN层用于压缩时序 维度并提取概率特征。WN层将网络权重分解为标 量和向量进行更新,加速收敛<sup>[23]</sup>。LeakyReLU激活 函数被用于提高网络的非线性拟合能力。最后,多 个时序卷积模块提取的输出信息作为多时间尺度的 概率特征被提取至输出层。经过一维卷积网络压缩 至3个通道,并将相应通道的权重和形态参数输入对 应的激活函数获得满足条件的混合Beta分布参数。

MDN将以最大似然估计为目标进行训练。训练 MDN时取负对数似然(negative log-likelihood, NLL)函数作为损失函数。

$$L = -\ln \sum_{i=1}^{m} \omega_i(x) P(y|\alpha_i(x), \beta_i(x))$$
(9)

采用log-sum-exp方法对损失函数进行改写: L=

$$-\ln\sum_{i=1}^{m}\exp\left\{\ln\omega_{i}(x)+\ln\Gamma(\alpha_{i}(x)+\beta_{i}(x))-\ln\Gamma(\alpha_{i}(x)-1)\ln\gamma+(\alpha_{i}(x)-1)\ln\gamma+(\alpha_{i}(x)-1)\ln\gamma+(\beta_{i}(x)-1)\ln(1-\gamma)\right\}$$
(10)

## 2.4 评价指标

本文采用平均绝对误差(mean absolute error, MAE) $I_{MAE}$ 和均方根误差(root mean square error, RMSE) $I_{RMSE}$ 作为确定性预测评价指标。并采用平 均覆盖误差(average coverage error, ACE)和预测区 间平均宽度(predict interval average width, PIAW) 分析概率预测的可靠性和锐度性能,连续等级概率 分数(continuous ranked probability score, CRPS)  $I_{CRPS}$ 被用于综合评价概率预测精度<sup>[1]</sup>。

$$I_{\rm MAE} = \frac{1}{N} \sum_{s=1}^{N} |y_s - \bar{y}_s|$$
(11)

$$I_{\rm RMSE} = \sqrt{\frac{1}{N} \sum_{s=1}^{N} (y_s - \bar{y}_s)^2}$$
(12)

式中:N为测试样本数量; $y_s$ 为风电功率实测值; $\bar{y}_s$ 为确定性预测值。

$$I_{\text{PICP},r} = \frac{1}{N} \sum_{s=1}^{N} \chi_{sr}$$

$$\chi_{sr} = \begin{cases} 1 & y_s \in [P_{sr}^{\min}, P_{sr}^{\max}] \\ 0 & y \notin [P_{sr}^{\min}, P_{sr}^{\max}] \end{cases}$$
(13)

$$I_{\text{ACE},\tau} = (I_{\text{PICP},\tau} - \tau) \times 100\%$$
(14)

$$I_{\text{PIAW, r}} = \frac{1}{N} \sum_{s=1}^{N} (P_{sr}^{\text{max}} - P_{sr}^{\text{min}})$$
(15)

式中: $I_{PICP,\tau}$ 为置信度  $\tau$ 下的预测区间覆盖率(predict interval coverage probability, PICP); $\chi_{s\tau}$ 为示性函数;  $P_{s\tau}^{max}$ 和 $P_{s\tau}^{min}$ 分别为采样点 s在置信度  $\tau$ 下功率区间的 上界和下界; $I_{ACE,\tau}$ 为 $I_{PICP,\tau}$ 与置信度  $\tau$ 之差, $I_{ACE,\tau}$ 越 接近于0说明预测区间可靠性越高; $I_{PIAW,\tau}$ 为置信度  $\tau$ 下的PIAW。PIAW 评价概率预测表达不确定信 息的锐度性能,PIAW 与 ACE 共同组成评价体系, 当 ACE 相同时 PIAW 越小表明预测结果的锐度性 能更好。

$$I_{\rm CRPS} = \frac{1}{N} \sum_{s=1}^{N} \left( E(|Y - y_s|) - \frac{1}{2} E(|Y - Y'|) \right) \quad (16)$$

式中:  $Y 和 Y' 为从概率密度函数 P(y) 中采样得到的独立随机变量; <math>E(\cdot)$ 表示求期望值。 $I_{CRPS}$ 越小表明概率预测综合性能越好。

#### 3 算例分析

## 3.1 算例概况

算例采用澳大利亚White Rock风电场实测数 据进行验证分析,风电场额定容量为175 MW,数据 分辨率为15 min<sup>[24]</sup>。取2019年1月1日至2020年 12月31日数据作为训练集,2021年1月1日至 2021年6月30日数据作为测试集。

算例中的模型通过深度学习框架 Pytorch 1.8.0 搭建。计算机 CPU为 Intel Core i9 7900X(3.30 GHz), 内存128 GB,显卡为NVIDIA TITAN V(显存12 GB)。 算例将 TMDN模型与4种概率预测模型从确定性 预测与概率预测两个方面进行对比。

对比模型包括基于 Gaussian 分布和 Beta 分布的两种深度自回归(deep autoregressive recurrent, DeepAR)模型、基于混合高斯分布的 MDN 模型(用 FCN-GMM 模型表示)和基于混合 Beta 分布的 MDN模型(用FCN-BMM模型表示)。

#### 3.2 模型训练分析

算例采用4h风电实测数据作为输入,单步预测 未来15min至4h的风电功率。TMDN模型中3层 TCN的卷积核大小分别为1×2、1×2与1×4,局部 矩窗口宽度为3。模型训练参数设置训练次数为 500次,批数量(batch size)为500,学习率为0.001。 为了对比模型的训练效率,5种模型的训练时间见 附录A表A1。

算例在训练中设置了局部矩通道的对比模型。 模型1不附加局部矩通道,将原始数据直接输入模 型训练。模型2附加局部一阶矩通道和局部二阶矩 通道。两种模型在20次训练中损失函数的置信区 间与均值见附录A图A2。从图A2可知,模型2在 训练中损失函数的置信区间更窄,且训练收敛后模 型2的损失函数均值更小。因此,局部矩通道能够 有效提高网络的收敛性。

为了说明 TMDN 模型结构提取多时间尺度概 率特征的有效性,算例将 TMDN 模型与 FCN-BMM 模型的混合 Beta分布权重进行对比,取 [t-2,t]时段的局部二阶矩信息表征短时间尺度 内风电功率波动情况,测试集分布权重与局部二阶 矩信息的关系见附录A图A3。依据 TMDN模型设 置的卷积核大小和分布权重 $\omega_1(x), \omega_2(x), \omega_3(x)$  分别对应前 30 min、1 h和4 h内的概率特征。从图 A3(a)中可知, $\omega_1(x)$ 与风电功率波动呈现明显的正 相关,当风电波动较大时,概率预测偏向于 30 min 内的概率特征。 $\omega_2(x)$ 与风电功率波动呈现负相 关,当风电功率波动较小时,风电概率预测更偏向于 1 h内的概率特征。而当风电功率较为平稳时,  $\omega_3(x)$ 权重较大,风电概率预测偏向于更长时间尺 度的概率特征。从图 A3(b)可知,FCN-BMM模型 提取的权重与风电波动特征关系较为模糊,且不同 权重对应的时间尺度在模型中也无法有效对应。因 此,TMDN模型在风电概率预测中能够提取并融合 多时间尺度的概率特征,有效引导风电概率预测,具 有更好的可解释性。

#### 3.3 预测结果对比

概率预测模型经过训练后,采用测试集进行验证。基于5种模型的概率预测方法在2021年1月 12日06:00至12:00未来15min的概率预测结果如 图4所示,展示了置信度为10%至90%的置信区间 与实测值。从图中可知,基于5种模型的概率预测 方法均能较好地描述风电出力的不确定性。预测时 段内风电场出力呈现上爬坡状态,06:00时风电场出 力为3.14 MW,09:00时风电场出力为71.40 MW。



Fig. 4 Confidence interval of probabilistic forecasting

为了分析概率预测模型在不同风电出力情形中的预测效果,06:00和09:00时刻的预测概率分布图见图5,图中每张子图的上、下小图分别表示06:00和09:00时刻的预测概率分布图。06:00时刻风电出力较小,由于Gaussian分布是无界分布,DeepAR-Gaussian模型和FCN-GMM模型预测的PDF存在密度泄漏。而09:00时刻5种模型预测的PDF能够较好地包含在有效区间内。独立的Gaussian分布在期望逼近0时,其PDF出现更为明显的密度泄漏。而Beta分布是严格定义在区间(0,1)中的有界分布,能够在有效区防止密度泄漏现象。因此,在概率

分布描述的范围上, Beta分布相比 Gaussian 分布具有一定优势。

为了定量对比预测模型,对比5种模型在超短 期内多个时间尺度的MAE、RMSE及CRPS来衡量 确定性预测与概率预测的精度,如附录A表A2所 示。从表A2中可知,MDN模型相比DeepAR模型 在概率预测上更为准确,其中TMDN模型在各个时 间尺度中各项指标均达到了最优值,在确定性预测 与概率预测中均有着更好的精度。

本文进一步验证了概率预测模型的可靠性与锐 度性能。5种预测模型未来1h预测的ACE指标及



Fig. 5 Probability density distribution of forecasting

PIAW 指标见附录 A 表 A3 和表 A4。从表 A3 中可 知, DeepAR-Gaussian 模型、DeepAR-Beta 模型的 ACE指标均大于 0,表明其 PICP 指标大于对应的置 信度。MDN 模型相比 DeepAR 模型有着更低的 ACE指标,可靠性更高。TMDN模型的 ACE指标 最贴近于 0,在置信度为 40% 至 90% 时均获得了最 小值。TMDN模型的 ACE指标在 30% 置信度下达 到最大负值(-2.23%),在 40% 置信度下达到最大 正值(0.74%)。

从附录A表A4中可以看出,5种预测模型的 PIAW指标随着置信度的升高均呈现上升趋势。 DeepAR模型的PIAW指标高于MDN模型。混合 分布相比独立分布有着更强的灵活性,给概率预测 结果带来更好的锐度性能。TMDN模型的PIAW 指标在20%至90%置信度下均达到了最小值。从 ACE和PIAW两个指标综合分析可知,TMDN模 型有着更高的可靠性与锐度性能,能够更精确地反 映概率预测信息。

## 4 结语

本文提出了基于 TMDN 的超短期风电功率概 率预测模型。模型融合时序数据的局部矩特征作为 MDN 的输入,并结合 TCN 融合多时间尺度的风电 概率特征进行预测。通过算例分析得到如下结论:

1)时序数据的局部矩信息能够提取局部概率特征,融合局部矩信息作为网络输入能够提高 MDN 对 MLE 的拟合能力,使模型在训练中具有更好的收敛性;

2) TMDN模型中的多层 TCN结构能够融合多时间尺度的概率特征,包括短时间尺度的波动性与 长时间尺度的稳定性特征,相比传统 MDN模型具 有可解释性;

3) TMDN 模型与 DeepAR-Gaussian、DeepAR-Beta、FCN-GMM、FCN-BMM 模型相比,进行超短 期风电功率概率预测具有更高的精度。

本文在概率预测研究中仅考虑了实测功率信息,后续工作将融合数值天气预报信息探索风电功 率的多步概率预测问题。 附录见本刊网络版(http://www.aeps-info.com/ aeps/ch/index.aspx),扫英文摘要后二维码可以阅读 网络全文。

# 参考文献

[1]万灿,宋永华.新能源电力系统概率预测理论与方法及其应用 [J].电力系统自动化,2021,45(1):2-16.

WAN Can, SONG Yonghua. Theories, methodologies and applications of probabilistic forecasting for power systems with renewable energy sources [J]. Automation of Electric Power Systems, 2021, 45(1): 2-16.

- [2] KHODAYAR M, WANG J H, MANTHOURI M. Interval deep generative neural network for wind speed forecasting [J]. IEEE Transactions on Smart Grid, 2019, 10(4): 3974-3989.
- [3]杨茂,董昊.基于数值天气预报风速和蒙特卡洛法的短期风电 功率区间预测[J].电力系统自动化,2021,45(5):79-85.
  YANG Mao, DONG Hao. Short-term wind power interval prediction based on wind speed of numerical weather prediction and Monte Carlo method [J]. Automation of Electric Power Systems, 2021, 45(5): 79-85.
- [4] WAN C, LIN J, WANG J H, et al. Direct quantile regression for nonparametric probabilistic forecasting of wind power generation[J]. IEEE Transactions on Power Systems, 2017, 32 (4): 2767-2778.
- [5] YU Y X, HAN X S, YANG M, et al. Probabilistic prediction of regional wind power based on spatiotemporal quantile regression [J]. IEEE Transactions on Industry Applications, 2020, 56(6): 6117-6127.
- [6] 刘燕华,李伟花,刘冲,等.短期风电功率预测误差的混合偏态 分布模型[J].中国电机工程学报,2015,35(10):2375-2382.
  LIU Yanhua, LI Weihua, LIU Chong, et al. Mixed skew distribution model of short-term wind power prediction error[J].
  Proceedings of the CSEE, 2015, 35(10): 2375-2382.
- [7] HE Y Y, LI H Y. Probability density forecasting of wind power using quantile regression neural network and kernel density estimation [J]. Energy Conversion and Management, 2018, 164: 374-384.
- [8] YAN J, LI K, BAI E W, et al. Hybrid probabilistic wind power forecasting using temporally local Gaussian process [J]. IEEE Transactions on Sustainable Energy, 2016, 7(1): 87-95.
- [9] CARPINONE A, GIORGIO M, LANGELLA R, et al. Markov chain modeling for very-short-term wind power forecasting [J]. Electric Power Systems Research, 2015, 122: 152-158.
- [10] 丁藤,冯冬涵,林晓凡,等.基于修正后ARIMA-GARCH模型的超短期风速预测[J].电网技术,2017,41(6):1808-1814. DING Teng, FENG Donghan, LIN Xiaofan, et al. Ultra-shortterm wind speed forecasting based on improved ARIMA-GARCH model[J]. Power System Technology, 2017, 41(6): 1808-1814.
- [11] SUN M C, FENG C, ZHANG J. Conditional aggregated probabilistic wind power forecasting based on spatio-temporal correlation[J]. Applied Energy, 2019, 256: 113842.
- [12] WANG Y, ZOU R M, LIU F, et al. A review of wind speed and wind power forecasting with deep neural networks [J].

Applied Energy, 2021, 304: 117766.

- [13] MURPHY K P. Probabilistic machine learning: an introduction (adaptive computation and machine learning)[M]. London, UK: MIT Press, 2020.
- [14] LIU Y Q, QIN H, ZHANG Z D, et al. Probabilistic spatiotemporal wind speed forecasting based on a variational Bayesian deep learning model [J]. Applied Energy, 2020, 260: 114259.
- [15] SALINAS D, FLUNKERT V, GASTHAUS J, et al. DeepAR: probabilistic forecasting with autoregressive recurrent networks [J]. International Journal of Forecasting, 2020, 36 (3): 1181-1191.
- [16] MEN Z X, YEE E, LIEN F S, et al. Short-term wind speed and power forecasting using an ensemble of mixture density neural networks[J]. Renewable Energy, 2016, 87: 203-211.
- [17] AFRASIABI M, MOHAMMADI M, RASTEGAR M, et al. Advanced deep learning approach for probabilistic wind speed forecasting [J]. IEEE Transactions on Industrial Informatics, 2021, 17(1): 720-727.
- [18] ZHANG H, LIU Y Q, YAN J, et al. Improved deep mixture density network for regional wind power probabilistic forecasting
   [J]. IEEE Transactions on Power Systems, 2020, 35(4): 2549-2560.
- [19] ZHANG H, YAN J, LIU Y Q, et al. Multi-source and temporal attention network for probabilistic wind power prediction [J]. IEEE Transactions on Sustainable Energy, 2021, 12(4): 2205-2218.
- [20] 孙亚南,黄越辉,孙谊媊,等.基于运行数据的短期风电功率预 测误差互补特性探析[J].电力系统自动化,2021,45(21): 215-223.

SUN Yanan, HUANG Yuehui, SUN Yiqian, et al. Operation data based analysis on complementary characteristics of shortterm power prediction error for wind power[J]. Automation of Electric Power Systems, 2021, 45(21): 215-223.

- [21] CHRISTOPHER B. Mixture density networks [R]. Birmingham, UK: Aston University, 1994: 1-26.
- [22] CHEN Y T, KANG Y F, CHEN Y X, et al. Probabilistic forecasting with temporal convolutional neural network [J]. Neurocomputing, 2020, 399: 491-501.
- [23] SALIMANS T, KINGMA D P. Weight normalization: a simple reparameterization to accelerate training of deep neural networks [EB/OL]. [2021-10-30]. https://arxiv.org/abs/ 1602.07868.
- [24] Aneroid Energy. Wind energy [EB/OL]. [2021-10-30]. https://anero.id/energy/wind-energy.

董骁翀(1995—),男,博士研究生,主要研究方向:人工智能在电力系统中的应用。E-mail:dream\_dxc@163.com

孙英云(1975—),男,通信作者,博士,教授,主要研究方向:新能源接入、电力系统优化与控制。 E-mail: sunyy@ncepu.edu.cn

蒲天骄(1970—),男,教授级高级工程师,主要研究方向: 电力系统自动控制、电力人工智能技术。E-mail:tjpu@epri. sgcc.com.cn

(编辑 鲁尔姣)

#### Ultra-short-term Probabilistic Forecasting of Wind Power Based on Temporal Mixture Density Network

DONG Xiaochong<sup>1</sup>, SUN Yingyun<sup>1</sup>, PU Tianjiao<sup>2</sup>, WANG Xinying<sup>2</sup>, LI Ye<sup>2</sup>

(1. School of Electrical and Electronics Engineering, North China Electric Power University, Beijing 102206, China;2. China Electric Power Research Institute, Beijing 100192, China)

**Abstract:** Probabilistic forecasting of wind power can provide critical boundary conditions for the safe operation of new power systems. Improving forecasting accuracy is the key problem of the research for probabilistic forecasting of wind power, and improving the interpretability of implicit models is beneficial to the promotion and application of artificial intelligence models. Therefore, a temporal mixture density network is proposed, which extracts the local moment information of time series data of wind power as input channels. The temporal convolutional network is used to extract the multi-time-scale probabilistic features, and the mixed Beta distribution is used to construct the probabilistic forecasting information. The results of the case study show that the local moment channel effectively improves the convergence of the model training, and the mixed distribution parameters extracted by the temporal mixture density network have a certain interpretability. Compared with the existing models, the forecasting results of the temporal mixture density network have better accuracy.

This work is supported by National Key R&D Program of China (No. 2020YFB0905900) and National Natural Science Foundation of China (No. 51777065).

**Key words:** wind power; probabilistic forecasting; mixture density network; temporal convolutional network; maximum likelihood estimation; interpretability

