

DOI: 10.19783/j.cnki.pspc.210424

基于动态贝叶斯网络的电价区间预测

王洪涛^{1,2}, 邹斌¹

(1. 上海大学机电工程与自动化学院, 上海 200444; 2. 宁德师范学院信息与机电工程学院, 福建 宁德 352100)

摘要: 在高比例可再生能源参与市场竞争的背景下, 电价波动更为剧烈。为了对电价区间进行预测, 提出了动态贝叶斯网络(Dynamic Bayesian Network, DBN)的电价区间预测方法。该方法以风电发电量、总发电量和总用电量的预测值以及历史电价真实值为输入数据, 以贪婪搜索算法确定 DBN 的网络结构, 以最大似然估计法(Maximum Likelihood Estimate, MLE)估计 DBN 网络参数, 建立 DBN 模型。然后以风电发电量、总发电量和总用电量的预测值为推理证据, 采用联合树推理得到电价预测的离散值和后验概率, 实现电价的区间预测。最后将所提方法与电价真实值、对比方法进行比较, 验证了所提方法的有效性。所提方法不仅能得到电价的预测区间, 而且能给出对应的概率, 对提高市场成员的收益、规避价格风险具有指导意义。

关键词: 电价预测; 区间预测; 动态贝叶斯网络; 联合树推理; 向前向后算法; 改进 k-means 聚类; 平均差异度

Prediction interval forecasts of electricity price based on dynamic Bayesian networks

WANG Hongtao^{1,2}, ZOU Bin¹

(1. School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200444, China;

2. School of Information Mechanical and Electrical Engineering, Ningde Normal University, Ningde 352100, China)

Abstract: Given the high proportion of renewable energy participating in market competition, the fluctuation of electricity price will be more severe. In order to predict the range of electricity price, a dynamic Bayesian network (DBN) interval prediction method is proposed. In this method, the predicted data of wind power generation, total power generation and total electricity consumption, and the real value of electricity price, are taken as input data. The network structure of DBN is determined by a greedy search algorithm, and the network parameters of DBN are estimated by maximum likelihood estimated (MLE). The DBN model is established. Then, with the predicted value of wind power generation, total power generation and total electricity consumption as the reasoning evidence, the discrete value and a posteriori probability of electricity price prediction are obtained using union tree reasoning, and the interval prediction of electricity price is realized. Finally, the proposed method is compared with the real value of electricity price and the comparison method to verify the effectiveness of the proposed method. The proposed method can not only get the predicted range of electricity price, but also give the corresponding probability. This has guiding significance for increasing the income of market members and avoiding price risk.

This work is supported by the Natural Science Foundation of Fujian Province (No. 2019J01845).

Key words: electricity price forecasting; interval prediction; dynamic Bayesian network; associative tree reasoning; forward and backward algorithm; improved k-means clustering; mean degree of difference

0 引言

在高比例可再生能源参与竞争的电力市场中(如北欧电力市场), 现货市场中电价受可再生能源发电成本低、出力波动性强^[1]以及输电线路阻塞等

多方面因素影响, 容易出现电价剧烈波动和跳跃, 给电价的点预测带来较大的困难^[2-3]。为了给市场成员的短期交易、风险价值计算^[4]以及报价策略等提供更为丰富的电价信息和决策自由度, 有必要对电价区间进行预测^[5-6]。

区间预测是指在给定置信水平下, 预测未来时刻目标量可能出现的上下界^[7-8]。以神经网络

基金项目: 福建省自然科学基金项目资助(2019J01845)

(Neural Network, NN)为代表的区间预测是目前的主要方法^[9-10]。

2011年 Abbas Khosravi 提出了边界估计(Lower Upper Bound Estimation, LUBE)的区间预测方法^[11], 该方法将 NN 的单个输出调整为两个输出, 直接输出预测区间的上界和下界, 通常以预测区间覆盖概率(Prediction Intervals Coverage Probability, PICP)、预测区间平均带宽(Prediction Intervals Normalized Average Width, PINAW)、累积带宽偏差(Accumulated Width Deviation, AWD)为评价指标, 通过赋予各评价指标适当的权重系数, 构造综合目标函数, 利用综合目标函数最小优化网络参数, 实现区间预测。

其后国内外研究者提出了不同类型的 NN 进行区间预测, 并以启发式算法优化 NN 的网络参数。文献[12]提出了自适应模糊神经推理与模拟退火算法相结合的区间预测方法, 文献[13-14]将多层感知机与粒子群(Particle Swarm Optimization, PSO)算法相结合, 对风电功率、负荷进行区间预测。文献[15]将极限学习机与 PSO 相结合对风电功率进行区间预测, 文献[16]将小波神经网络与改进人工蜂群相结合进行风电区间预测。文献[17]采用支持向量机预测电价的上界与下界, 并采用 PSO 对支持向量机的超参数进行优化。文献[18]提出基于残差神经网络对电价区间预测的方法。

国内期刊上关于电价区间预测的报道尚不多见, 但关于风电功率、光伏出力、负荷的区间预测已有报道。文献[19]采用 PSO 优化核极限学习机(Kernel Extreme Learning Machine, KELM)对风电功率进行区间预测。文献[20]采用 PSO 优化 LUBE 的输出权重对光伏出力进行区间预测, 以均方根宽度代替 PINAW, 改进了综合目标函数, 降低了综合目标函数中权重系数选取的难度。文献[21]提出改进区间预测评价指标的 PSO 优化 LUBE 方法, 对负荷区间进行预测, 能较好地包络波动性负荷。

以上文献均以综合目标函数(见下文式(13))最小优化网络参数, 进而实现目标量的区间预测。但综合目标函数的分目标函数(即评价指标, 如: PICP、PINAW、AWD)的权重系数依赖于经验, 得到的预测区间的覆盖概率是基于总体平均的概念, 即: 各时刻电价预测区间的覆盖概率视为相等, 无法给出具体时刻预测区间对应的概率。由于市场中的价格风险主要是由电价非正常跳跃和价格尖峰引起的, 对市场成员而言, 预测出各时刻的电价区间, 并给出其相应的概率, 对提高收益、规避价格风险更具有指导意义。

基于以上分析, 本文以北欧现货市场中丹麦西部地区(DK1)和东部地区(DK2)的风电发电量、总发电量、总用电量的预测数据和历史电价数据为基础, 采用平均差异度为判据的改进 k-means 聚类方法对数据离散化, 得到离散样本集。然后以离散样本集为基础, 建立动态贝叶斯网络(Dynamic Bayesian Network, DBN)模型; 以未来时刻风电发电量、总发电量和总用电量的离散值(聚类类别)为推理证据, 推理未来电价的离散值和后验概率。将各离散值对应的聚类中心与后验概率的乘积之和作为电价预测的均值, 实现电价点预测。分别将离散值对应的连续区间的上下界与后验概率的乘积之和作为预测区间的上下界, 实现电价的区间预测。最后将所提方法与电价真实值、文献[19]的区间预测方法对比, 验证了所提方法的有效性。

1 电价的影响因素和 DBN 模型

1.1 电价的影响因素

在可再生能源未参与市场竞争的情况下, 市场电价主要受供求关系、机组报价、输电线路阻塞等因素的影响, 具有非线性和非平稳的特点^[22-23]。当可再生能源参与市场竞争时, 由于可再生能源的发电成本低, 并且出力具有不确定性、间歇性和波动性, 更容易引起电价的剧烈波动, 因此在电价预测时不仅需要考虑发电量与用电量的供求关系对电价的影响, 还需要考虑可再生能源发电量对电价的影响。

在电力市场中, 市场交易中心向市场成员披露的信息不仅包括历史电价数据, 还包括交易的供需信息、发电量预测、用电量预测、可再生能源发电量预测等各类信息(如北欧电力市场), 为了提高收益、规避价格风险, 市场成员需要综合各类信息后决定报价策略。考虑到电价受多方面因素的影响, 并且多个因素之间具有相关性, 而 DBN 是基于概率不确定性的推理网络, 为因果信息的学习和推理提供了有效途径, 能对各时间片上的状态变量进行预测, 已被用于光伏发电概率预测^[24]、负荷预测^[25]和经济指标预测^[26]等方面。

因此本文根据影响电价主要因素的专家知识^[4,27-28], 以风电发电量、总发电量和总用电量作为电价的解释变量, 构建 DBN 模型, 通过证据推理预测电价区间。

1.2 DBN 的基本原理

静态贝叶斯网络定义为: $BN=(G, \theta)$, 其中 G 为随机变量集 X_1, X_2, \dots, X_n 上的有向无环图, 即网络结构; θ 为网络参数, 即条件概率表。BN

模型表达了多维随机变量的联合概率分布^[29-31]。

$$P(x_1, x_2, \dots, x_T) = \prod_{i=1}^T P(x_i | Pa(x_i)) \quad (1)$$

式中, $Pa(x_i)$ 表示随机变量 x_i 的父结点。

DBN 是在静态贝叶斯网络的基础上增加了时间属性, 将 BN 在时间轴上展开, 描述含有时间因素的随机过程。DBN 由 (B_0, B_{\rightarrow}) 两部分组成, 其中 $B_0 = (G_0, \theta_0)$ 为初始网络, G_0 表示初始网络的结构, θ_0 表示初始网络的参数, 它定义了初始状态下随机变量的联合概率分布。 $B_{\rightarrow} = (G_{\rightarrow}, \theta_{\rightarrow})$ 为转移网络, G_{\rightarrow} 表示转移网络的结构, θ_{\rightarrow} 表示转移网络的参数, 它定义了相邻两个时间片之间随机变量的状态转移概率 $P(x_{t+1} | x_t)$ 。

如果确定了 B_0 和 B_{\rightarrow} , 则将 B_{\rightarrow} 沿时间轴展开, 得到的 DBN 示意图如图 1 所示。图 1 中随机变量的下标表示时刻, 上标表示第 i 个随机变量。DBN 模型中从 $t=0$ 至 $t=T$ 时, 随机过程在时间轨迹上的联合概率分布为

$$P(x_1^1, x_2^1, \dots, x_T^1) = P_0(x_0^1) \prod_{i=1}^T P_{\rightarrow}(x_i^1 | x_{i-1}^1) = \prod_{i=1}^N P_0(x_i^0 | Pa(x_i^0)) \cdot \prod_{i=1}^T \prod_{i=1}^N P_{\rightarrow}(x_i^1 | Pa(x_i^1)) \quad (2)$$

式中: x_i^0 、 x_i^1 表示第 i 个变量在 0 和 1 时刻的取值; $P_0(x_i^0)$ 、 $P_{\rightarrow}(x_i^1)$ 表示其父结点。

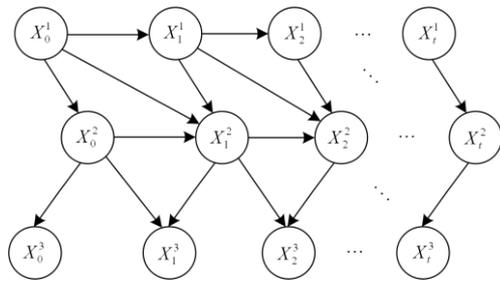


图 1 DBN 示意图

Fig. 1 DBN schematic

1.3 DBN 建模

DBN 建模的本质是在已知随机过程在 $1:t$ 时刻的样本 $\{X_{1:t}\}$ 的情况下, 通过学习得到随机过程在时间轨迹上的联合概率分布。DBN 建模的主要步骤为以下 3 步。

1) 获得时间序列的离散样本集

时间序列的离散化方法主要有等宽离散化、等频离散化和 k-means 聚类方法^[32-33], 但确定离散化区间个数和聚类个数的主观性强, 对离散化结果影响大。本文以文献[34]提出的平均差异度为判据, 确

定聚类中心个数 k 和初始聚类中心。首先定义样本距离、样本平均差异度和样本集总体的平均差异度。

(1) 两个样本间距离

设样本数据集 $D = [x_1, x_2, \dots, x_T]^T$, $x_i = [x_{i1}, x_{i2}, \dots, x_{iL}]$, D 的维数是 $T \times L$, 定义两个样本向量 x_i 与 x_j 之间的欧氏距离为

$$d(x_i, x_j) = \left[\sum_{l=1}^L |x_{il} - x_{jl}|^2 \right]^{1/2} \quad (3)$$

(2) 样本 x_i 的平均差异度

$$d_i = \sum_{j=1}^T d(x_i, x_j) / T \quad (4)$$

(3) 样本集的平均差异度

$$MD = \sum_{j=1}^T d_j / T \quad (5)$$

确定初始聚类中心的步骤是: 首先将 d_i 最大的样本作为第 1 个聚类中心 $c(1)$, 并在样本集中剔除 $c(1)$, 形成新的样本集。然后新的样本集中找 d_i 最大的样本, 若该样本与已选定的初始聚类中心的差异度均大于 MD , 则该样本作为第 2 个聚类中心 $c(2)$, 否则找第二大的平均差异度样本进行判断, 直至选出 $c(2)$ 。如此循环, 直至将满足条件的聚类中心 $c(q)$ 筛选出来。以 $\{c(1), c(2), \dots, c(q)\}$ 为初始聚类中心, 聚类中心的个数为 q , 再按常规的 k-means 聚类方法对样本数据集 D 中的各列作 k-means 聚类, 聚类后得到离散数据集 \bar{D} 。

2) DBN 的结构学习

以离散数据集 \bar{D} 为基础, 采用 BIC 评分函数^[35], 在 DBN 可能的结构中搜索最优的网络结构。采用贪婪搜索算法, 从空的网络结构开始, 通过 3 种基本算子:

- (1) 增加边: 在网络中两个结点间添加一条新边;
- (2) 删除边: 删除网络中两个结点间的一条边;
- (3) 反转边: 将网络中已有的一条边的方向反转。

通过以上 3 种算子的不断操作, 得到新的网络结构, 然后用 BIC 评分函数对新的网络结构进行评分, 将评分最高的网络结构作为 DBN 结构学习的结果。

3) DBN 的参数学习

在确定 DBN 结构之后, 从 \bar{D} 中学习网络参数, 得到结点之间的条件概率表, 即: B_0 和 B_{\rightarrow} 的条件概率表 θ_0 和 θ_{\rightarrow} 。在完全观测的数据集 \bar{D} 中, 采用最大似然估计法^[36]寻找到 $\theta_{i,j,k}^0$ 和 $\theta_{i,j,k}^{\rightarrow}$ 的最优估计 $\hat{\theta}_{i,j,k}^0$ 和 $\hat{\theta}_{i,j,k}^{\rightarrow}$ 。

通过以上 3 个步骤可以得到 $B_3=(G_3, \Theta)$ 和 $B_2=(G_2, \Theta)$, 从而建立 DBN 模型。

2 DBN 推理

2.1 DBN 推理过程

DBN 推理是在已建立了 $1:t$ 时刻的 DBN 基础上, 添加解释变量在 $t+1$ 时刻的样本(推理证据), 推理出 $t+1$ 时刻待预测变量的离散化状态(聚类类别)和后验概率。

如前文所述, 将风电发电量 W 、总发电量 P 、总用电量 L 作为电价 EP 的解释变量, 假设由 1.3 节已经得到了时间轨迹上的联合概率分布 $P(W_{0:t}, P_{0:t}, L_{0:t}, EP_{0:t})$ 。如果添加 $t+1$ 时刻 W_{t+1} 、 P_{t+1} 和 L_{t+1} 的样本作为推理证据, 则通过 DBN 推理就能得到后验概率 $P(EP_{t+1} | W_{0:t+1}, P_{0:t+1}, L_{0:t+1}, EP_{0:t})$, 逐步递推就可以得到 $t+k$ 时刻电价 EP_{t+k} 的后验概率, DBN 证据推理过程如图 2 所示。

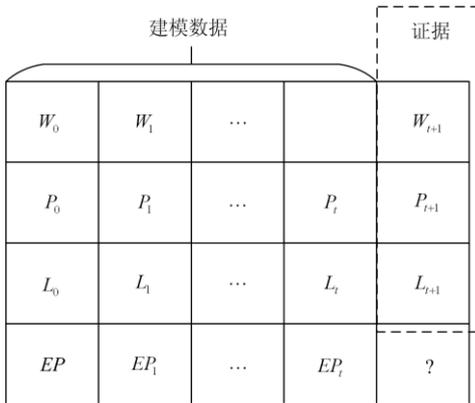


图 2 DBN 推理
Fig. 2 DBN inference

DBN推理有精确推理和近似推理两种方法^[26], 本文采用DBN的联合树推理算法。联合树推理的基本思想是利用DBN中相邻的两个时间片内的节点构造1.5DBN联合树, 然后输入证据, 利用向前向后算法进行推理, 推理结束后, 边缘化 $t+k$ 时刻(k 表示预测步长)的预测变量, 就能计算出预测变量取不同离散值时对应的后验概率。

利用贝叶斯网络工具箱FullBNT-1.0.7, 实现精确推理的具体步骤如下。

- 1) 将DBN模型转换为1.5DBN联合树。
- 2) 输入证据, 采用向前向后算法进行推理。
- 3) 计算电价在各时刻取不同离散值时的后验概率 $P(EP_{t+k} = q)$ 。
- 4) $t+k$ 时刻电价的预测均值^[37]为

$$\overline{EP_{t+k}} = \sum_{q \in \mathcal{Q}} P(EP_{t+k} = q) R(EP_{t+k} = q) \quad (6)$$

式中: t 表示当前时刻; k 表示预测步长, $k=1, 2, \dots, N$; $c(\cdot)$ 表示离散值为 q 时的聚类中心; q 和 N_q 分别表示离散值和聚类中心的个数, $q=1, 2, \dots, N_q$; $\text{Pr}(\cdot)$ 为后验概率。

5) $t+k$ 时刻电价预测区间的下界和上界为

$$L_{t+k} = \sum_{q \in \mathcal{Q}} l(EP_{t+k} = q) R(EP_{t+k} = q) \quad (7)$$

$$U_{t+k} = \sum_{q \in \mathcal{Q}} u(EP_{t+k} = q) R(EP_{t+k} = q) \quad (8)$$

式中: $l(\cdot)$ 、 $u(\cdot)$ 表示离散值为 q 时聚类中心对应区间的下界和上界。

2.2 DBN 电价区间预测的步骤

Step1 原始数据采集。采集北欧电力现货市场公开的2017年丹麦DK1、DK2的风电发电量、总发电量和总用电量预测值以及历史电价的每小时数据。

Step2 时间序列离散化。采用改进的k-means聚类离散化方法, 若时间序列的连续值 $x_i(t)$ 与第 q 个聚类中心的距离最小, 则对应的离散值为 q , 由此构造离散数据集 \bar{D} 。

Step3 DBN结构学习。根据离散数据集 \bar{D} , 以 BIC为评分函数, 采用贪婪搜索算法, 得到初始网络结构 G_0 和前后两个时间片之间的转移结构 G_{\rightarrow} 。

Step4 DBN参数学习。根据离散数据集 \bar{D} , 采用最大似然估计法, 估计出 $\delta_{i,j,k}^0$ 和 $\delta_{i,j,k}^{\rightarrow}$ 。

Step5 建立DBN模型。根据 G_0 、 $\delta_{i,j,k}^0$ 和 G_{\rightarrow} 、 $\delta_{i,j,k}^{\rightarrow}$, 建立DBN模型。

Step6 DBN转化为联合树。调用联合树推理引擎: `engine=jtree_dbn_inf_engine(DBN)`, 将DBN模型转换为1.5DBN联合树。

Step7 证据推理。输入推理证据, 调用证据推理模块 `[engine,L]=enter_evidence(engine,evidence)`, 采用向前向后算法完成证据推理。

Step8 计算后验概率。调用子函数 `Pr= marginal_nodes(engine, nodes,t)`, 得到电价离散值对应的后验概率。

Step9 计算电价预测的均值。以电价的后验概率, 按式(6)计算 $t+k$ 时刻电价的预测均值 $\overline{EP_{t+k}}$ 。

Step10 计算电价预测区间。按式(7)、式(8)计算 $t+k$ 时刻电价预测区间的下界 L_{t+k} 和上界 U_{t+k} 。

2.3 区间预测评价指标

区间预测效果的评价指标^[19]主要有3个, PICP从实际观测值处于预测区间上界和下界之间的概率来评价; PINAW从预测区间上界和下界之间的宽度来评价; 当实际观测值落在预测区间之外时, AWD从偏离预测区间上界(或下界)的程度来评价, PICP、PINAW和AWD的表达式如式(9)一式(12)所示。

1) PICP

$$PICP = \frac{1}{N} \sum_{k=1}^N \varepsilon_k \quad (9)$$

式中: N 为预测样本数; ε_k 为变量, 如果 $t+k$ 时刻的预测目标值 EP_{t+k} 处于 $[L_{t+k}, U_{t+k}]$ 之内, 则 $\varepsilon_k = 1$, 否则 $\varepsilon_k = 0$ 。若PICP明显小于95%的置信水平, 则说明区间预测结果不可靠。

2) PINAW

$$PINAW = \frac{1}{NR} \sum_{i=k}^N [U_{t+k} - L_{t+k}] \quad (10)$$

式中, R 为目标值的变化范围, 即目标值最大值和最小值的差值。

3) AWD

$$AWD_k = \begin{cases} \frac{L_{t+k} - EP_{t+k}}{U_{t+k} - L_{t+k}} & \text{if } EP_{t+k} < L_{t+k} \\ 0 & \text{if } EP_{t+k} \in [L_{t+k}, U_{t+k}] \\ \frac{EP_{t+k} - U_{t+k}}{U_{t+k} - L_{t+k}} & \text{if } EP_{t+k} > U_{t+k} \end{cases} \quad (11)$$

$$AWD = \frac{1}{NR^*} \sum_{i=k}^N AWD_k \quad (12)$$

式中: R^* 用于对累积带宽偏差进行归一化处理; AWD 的值越小, 表明预测区间的质量越高。

由式(9)一式(12)可知, 区间预测的PICP值越大、PINAW和AWD的值越小, 区间预测的效果越好, 但3个评价指标之间是相互矛盾的, 文献[19]以综合目标函数(式(13))最小优化KELM的权重, 实现风电功率的区间预测。

$$\min_{\beta} F = \sum_{i=1}^m [\gamma_i | PICE_i^{(\alpha_i)} | + \varphi_i | PINAW_i^{(\alpha_i)} | + \lambda_i | AWD_i^{(\alpha_i)} |] \quad (13)$$

式中: $PICE = |PICN - PICP|$, 表示预测区间覆盖概率偏差, $PICN = 100 \times (1 - \alpha_i) \%$; γ_i 、 φ_i 、 λ_i 表示 PICE、PINAW 和 AWD 的权重系数; i 表示迭代次数; m 为最大迭代次数。

为了公平比较本文方法和文献[19]的PSO-KLEM方法的区间预测效果, 以PICP、PINAW和AWD作为区间预测的评价指标, 检验区间预测的精度。

3 算例分析

3.1 数据说明

DBN建模数据: 采用北欧电力市场Nord Pool现货市场公开的2017年1月1日至1月7日丹麦DK1(西部地区)和DK2(东部地区)风电发电量、总发电量和总用电量预测值以及电价的每小时数据^[38]。

建模数据中, 电价采用的是历史真实数据, 风电发电量、总发电量和总用电量采用的是预测数据。未使用风电发电量、总发电量和总用电量历史真实数据的原因是: 由于现货市场中, 在 $t+k$ 时刻(未来时刻)的风电发电量、总发电量和总用电量的真实值是未知的, 市场成员只能利用上述各量的预测值和电价的历史数据作为电价预测的依据。因此本文将风电发电量、总发电量和总用电量的预测数据以及历史电价作为DBN的建模数据。

输入DBN模型的推理证据是2017年1月8日DK1、DK2地区的风电发电量、总发电量和总用电量24h的预测数据, 经DBN模型推理后, 输出结果为1月8日每小时电价的离散值和后验概率, 利用式(6)一式(8)得到电价均值和区间预测的上界和下界。

PSO优化KELM方法的主要参数设置为: RBF径向基核函数的带宽 $\sigma=0.02$ 、惩罚系数 $C=10$, PSO算法的种群规模为25, 惯性权重系数 ω 按线性递减策略调整, $\omega \in [0.1, 0.9]$, 最大迭代次数为200, $\alpha_i=0.05$, 综合目标函数中的权重系数 $\gamma_i=0.8$, $\varphi_i=1$, $\lambda_i=0.2$ 。

3.2 DBN建模结果

按照1.3节给出的DBN建模步骤, 采用改进的k-means聚类离散化方法得到了21个聚类中心, 电价的离散状态与聚类中心、聚类区间的对应关系见表1。表1中聚类区间的下界和上界分别为离散状态下连续样本集中样本的最小值和最大值。由表1的聚类中心和对应区间可知, DK1区域的电价中出现了负电价。

以数据集 \bar{D} 为基础, 调用贝叶斯网络工具箱FullBNT-1.0.7中的DBN结构和参数学习子函数, 建立的DBN网络结构图如图3所示。

图3中初始网络 G_0 中的 W_0 、 P_0 、 L_0 、 EP_0 分别代表 $t=0$ 时刻风电发电量、总发电量、总用电量和电价。 G_0 中 W_0 、 P_0 和 L_0 均为 EP_0 的父结点, 意味着这三者与电价存在因果关系; W_0 是 P_0 的父结点, 二者之间构成因果关系。图3中的转移网络 G_s 为马尔科夫过程, 将 G_0 与 G_s 相结合, 得到在时间轴上展开的DBN结构, 其每一时刻的片内结构表征

表 1 DK1 电价的离散状态、聚类中心和对应区间

Table 1 Discrete state, clustering center and coverage interval of electricity price in DK1

离散状态	聚类中心/(欧元/MWh)	对应区间/(欧元/MWh)	离散状态	聚类中心/(欧元/MWh)	对应区间/(欧元/MWh)
1	-5.68	[-6.27,-5.09]	12	32.78	[32.50,32.99]
2	0.06	[-0.05,0.12]	13	34.00	[33.53,34.32]
3	10.88	[10.55,11.10]	14	35.16	[34.59,35.93]
4	15.42	[13.75,16.43]	15	37.90	[36.88,39.01]
5	17.59	[16.88,18.13]	16	42.55	[42.05,43.04]
6	21.22	[20.90,21.81]	17	44.73	[44.54,44.99]
7	27.22	[25.86,28.11]	18	50.08	[50.05,50.11]
8	29.16	[28.64,29.56]	19	56.10	[56.05,56.15]
9	29.97	[29.65,30.28]	20	69.70	[69.68,69.72]
10	30.71	[30.37,31.05]	21	72.96	[71.99,73.92]
11	31.42	[31.09,31.98]	—	—	—

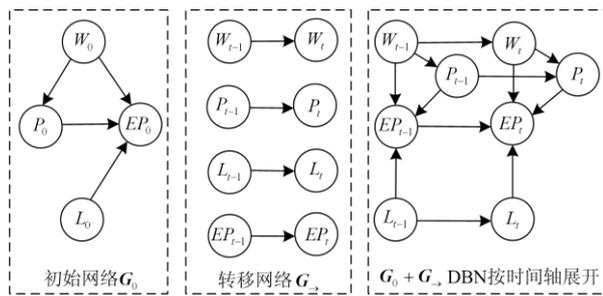


图 3 动态贝叶斯网络结构图
Fig. 3 DBN structure chart

了各结点(随机变量)之间的条件相依关系, 相邻两个时间片之间, 各结点状态按照马尔科夫状态转移矩阵进行转移。

通过极大似然估计法可以得到 DBN 的参数(条件概率表), 图 3 中 t 时刻电价 EP_t 有 4 个父结点, EP_t 的条件概率为 $P(EP_t | W_t, P_t, L_t, EP_{t-1})$ 。图 4 给出了条件概率 $P(EP_t = 1 | W_t, P_t, L_t = 1, EP_{t-1} = 1)$ 的三维概率图, 其含义是: 在 $L_t = 1, EP_{t-1} = 1$ 的条件下, $EP_t = 1$ 的概率由二维随机变量 (W_t, P_t) 的离散状态决定。由图 4 可知, t 时刻风电发电量、总发电量、总用电量以及 $t-1$ 时刻电价的离散值共同决定了 t 时刻电价的取值(离散的)和对应的概率。

3.3 DBN 推理与区间预测结果分析

建立 DBN 模型后, 输入 $t-k$ 时刻的证据向量 $[W_{t-k}, P_{t-k}, L_{t-k}]^T$, 通过 DBN 推理得到 EP_{t+k} 的后验概率。图 5 为 2017 年 1 月 8 日 DK1 预测电价的后验概率三维图, 图中给出了 k h 电价处于不同离散状态下的后验概率, 后验概率定量地给出了电价取不同离散值时发生的概率, 为市场成员提供了丰富的概率信息。

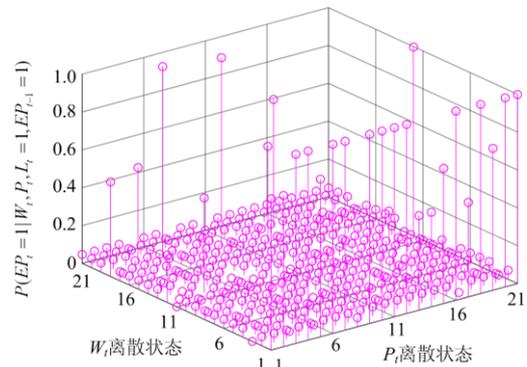


图 4 DBN 条件概率

Fig. 4 DBN conditional probability

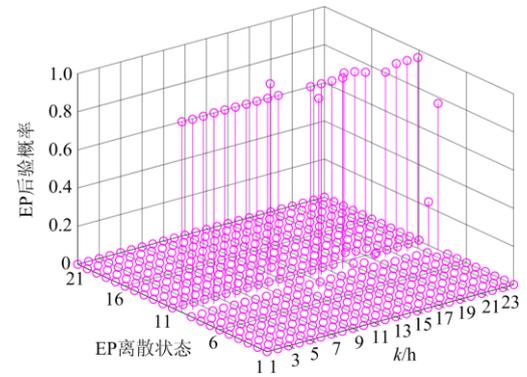


图 5 DK1 电价 24 h 后验概率

Fig. 5 Posterior probability of 24 h electricity price in DK1

图 6 给出了 11:00 时电价的离散状态和后验概率图, 当预测电价的离散状态为 14 时(对应的区间 [34.59,35.93]), 其后验概率为 0.94。

图 7 为 DK2 在 2017 年 1 月 8 日的每小时电价后验概率三维图, 电价的离散状态数为 32, 由平均差异度的 k-means 聚类方法确定。图 8 给出了 13:00

时电价的离散状态和后验概率图, 当预测电价的离散状态为 13 时, 其后验概率为 0.936。DK2 区域电价的离散状态与聚类中心、对应区间见表 2。

由电价的后验概率和离散状态对应的聚类中心, 根据式(6)可计算出各时刻电价的均值。图 9 为 DK1 地区 2017 年 1 月 8 日每小时电价均值, 由图 9 可见: 各时刻的电价预测均值与真实电价相差不大, 在 10:00—13:00 时, 电价剧烈波动, 电价预测均值仍具有较高的预测精度。

由图 6、图 8 和图 9 可知, 通过 DBN 模型对电价进行推理预测, 不仅得到了电价的预测区间(以离散值表示), 而且能够给出对应的概率, 并进一步计算出各时刻电价的预测均值, 这为市场成员确定报价策略提供了丰富的电价预测信息。

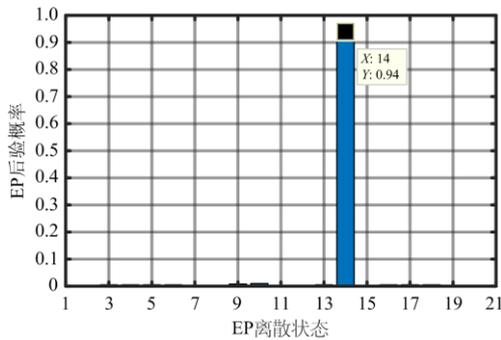


图 6 DK1 地区 11:00 时电价的后验概率
Fig. 6 Posterior probability of electricity price at 11:00 in DK1 region

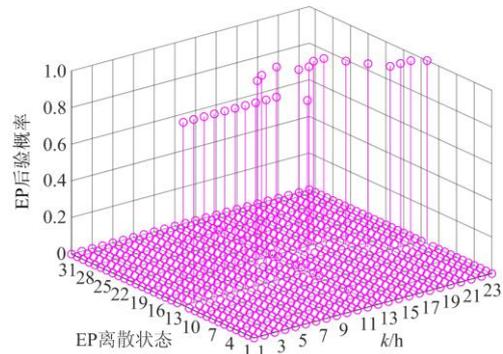


图 7 DK2 电价 24 h 后验概率

Fig. 7 Posterior probability of 24 h electricity price in DK2

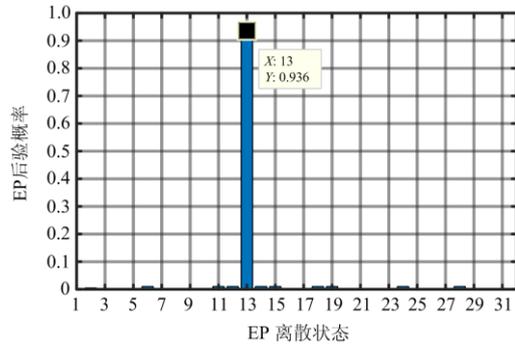


图 8 DK2 地区 13:00 时电价的后验概率
Fig. 8 Posterior probability of electricity price at 13:00 in DK2 region

表 2 DK2 电价的离散状态、聚类中心和对应区间

Table 2 Discrete state, clustering center and coverage interval of electricity price in DK2

离散状态	聚类中心/(欧元/MWh)	对应区间/(欧元/MWh)	离散状态	聚类中心/(欧元/MWh)	对应区间/(欧元/MWh)
1	-5.68	[-6.27,-5.09]	17	34.09	[33.89,34.32]
2	0.04	[-0.05,0.11]	18	34.81	[34.50,35.16]
3	10.38	[9.49,11.1]	19	35.62	[35.22,36.28]
4	13.15	[12.55,13.75]	20	37.15	[36.78,37.61]
5	15.47	[15.27,15.67]	21	39.17	[38.92,39.42]
6	16.45	[16.03,16.88]	22	41.47	[41.26,41.68]
7	17.95	[17.77,18.13]	23	42.09	[42.04,42.20]
8	21.22	[20.09,21.81]	24	43.21	[43.04,43.38]
9	26.28	[26.24,26.31]	25	45.04	[44.54,45.83]
10	27.99	[27.48,28.28]	26	49.49	[49.37,49.61]
11	29.08	[28.64,29.49]	27	50.08	[49.93,50.02]
12	30.01	[29.65,30.23]	28	55.95	[55.62,56.14]
13	30.65	[30.37,31.09]	29	57.70	[57.50,54.90]
14	31.79	[31.32,32.24]	30	59.99	[59.92,60.04]
15	32.74	[32.33,33.02]	31	69.70	[69.28,70.12]
16	33.41	[33.27,33.52]	32	72.96	[71.99,73.92]

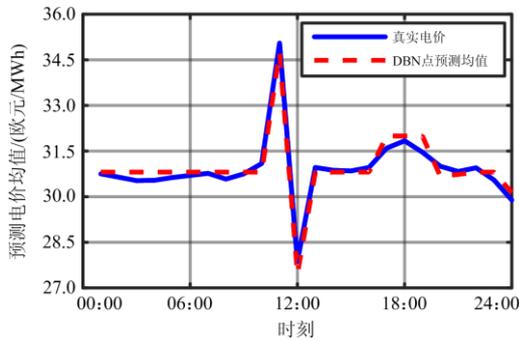


图 9 DK1 电价每小时电价均值

Fig. 9 Average hourly electricity price of DK1

利用式(7)、式(8)得到的区间预测结果如图 10 所示,文献[19]中 PSO-KELM 方法的区间预测结果如图 11 所示。由图 10 可见,在 01:00—09:00 时,真实电价较为平稳,预测区间的下界和上界也较为平直,预测区间的宽度基本相同;在 10:00—13:00 时,电价剧烈波动,预测区间下界和上界随之起伏,仍能覆盖真实电价;在 20:00 时,预测区间虽未能覆盖真实电价,但偏离区间上界的程度较小。

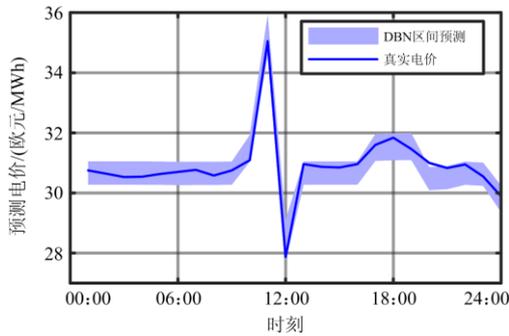


图 10 DBN 方法的 DK1 电价区间预测结果

Fig. 10 DBN method DK1 electricity price interval prediction results

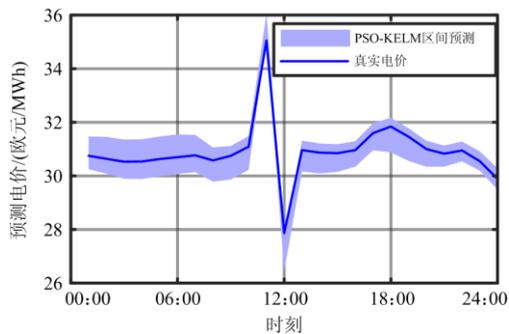


图 11 PSO-KELM 方法的 DK1 电价区间预测结果

Fig. 11 PSO-KELM method DK1 electricity price interval prediction results

在图 11 中,01:00—10:00 时,以较宽的预测区间覆盖真实电价;在 12:00 时,真实电价为 24 h 内的最小值,预测区间未能覆盖真实电价;其后预测区间虽能覆盖真实电价,但预测区间上界与下界之间的宽度较大。对比图 10 和图 11 中各时刻的区间预测结果可知,两种方法都能较好地覆盖真实电价,但图 10 的预测区间更窄(两种方法预测区间的 PINAW 指标比较见表 3),说明 DBN 方法的电价区间预测质量更高。

图 12、图 13 给出了本文方法与文献[19]方法对 DK2 地区电价的区间预测结果。对比图 12 和图 13 可知:DBN 的区间预测在 13:00 时未能覆盖真实电价,PSO-KELM 的区间预测在 03:00、04:00 两时刻未能覆盖真实电价,DBN 的 PICP 指标为 95.83%,比 PSO-KELM 的 PICP 指标 91.67%高,两种方法的 PINAW、AWD 指标见表 3。

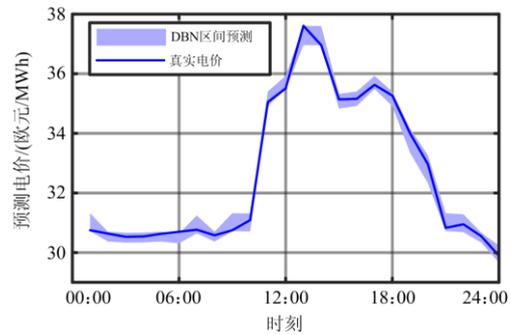


图 12 DBN 方法的 DK2 电价区间预测结果

Fig. 12 DBN method DK2 electricity price interval prediction results

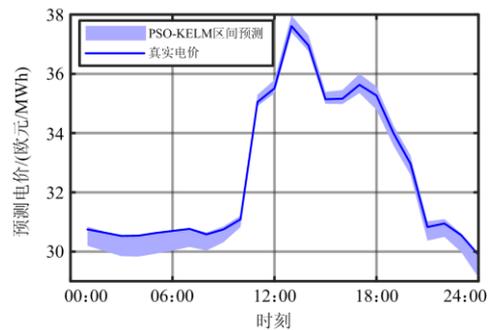


图 13 PSO-KELM 方法的 DK2 电价区间预测结果

Fig. 13 PSO-KELM method DK2 electricity price interval prediction results

由表 3 可见,DK1 的电价区间预测评价指标中,DBN 方法和 PSO-KELM 方法的 PICP 相等,但 DBN 方法的 PINAW、AWD 指标更小。DK2 的电价区间预测指标中,DBN 方法比 PSO-KELM 方法的 PICP

值大, 比 PINAW 值和 AWD 值小, 说明 DBN 方法的区间预测质量更高。

表 3 两种方法的区间预测评价指标

Table 3 Comparison of interval prediction evaluation index between the two methods

区域	PICP/%		PINAW/%		AWD/%	
	DBN	PSO-KELM	DBN	PSO-KELM	DBN	PSO-KELM
DK1	95.83	95.83	10.27	12.02	0.42	0.64
DK2	95.83	91.67	9.13	10.03	0.77	0.83

本文提出的 DBN 方法预测电价区间是以风电发电量、总发电量和总用电量作为电价的解释变量, 通过 DBN 确定了电价与风电发电量、总发电量和总用电量之间的因果关系(以条件概率的形式表示), 通过证据推理给出电价的离散状态和对应概率, 可同时得到电价的点预测和区间预测结果, 预测效率高。采用 DBN 推理模型能够刻画解释变量与电价的因果关系, 既给出了电价预测的均值, 同时也给出了预测区间的下界、上界以及对应的概率, 为市场成员提供的电价信息量更全面、更丰富。

PSO-KELM 方法也能够给出各时刻电价预测区间, 并以评价指标 PICP 给出覆盖概率, 但不能给出各时刻预测区间对应的概率信息。对市场成员而言, 各时刻预测区间对应的概率是需要关注的, 尤其是在高比例可再生能源参与市场竞争、电价波动剧烈的情况下, 获得电价预测区间对应的概率更有现实意义。另外该方法在采用式(13)作为综合目标函数时, 权重系数 γ_i 、 ϕ_i 、 λ_i 的选取依赖于主观经验, 对区间预测的精度影响较大。

通过以上的分析和比较, 说明 DBN 方法建立的推理预测模型具有物理意义明确、预测机理清晰、预测结果信息量丰富的特点, 通过与 PSO-KELM 方法的预测结果进行对比, 验证了所提方法的有效性。

4 结论

在高比例可再生能源参与市场竞争的背景下, 电力市场中电价受发电成本和供求关系等因素的影响更容易发生剧烈波动, 为了给市场成员提供更丰富的电价预测信息, 提出基于 DBN 的电价区间预测方法, 得到以下结论:

1) 以平均差异度为判据, 自动确定 k-means 初始聚类中心和个数, 减少了离散化过程中数据特征的丢失, 提高了 DBN 预测模型的精度。

2) 所提方法能够有效地实现电价的区间预测, 避免了分目标函数 PICP、PINAW、AWD 权重系数选择的主观性, 与 PSO-KELM 区间预测方法相比,

所提方法的区间预测质量更高。

3) 所提方法能获得每个预测时刻电价的区间和对应的概率, 并进一步给出预测电价的均值和预测区间的上下界, 提高了预测效率。各时刻电价的预测区间和对应的概率为市场成员提供了丰富、具体的电价预测信息, 对市场成员提高收益、规避价格风险具有指导意义。

参考文献

- [1] HUANG W, ZHANG N, KANG C, et al. From demand response to integrated demand response: review and prospect of research and application[J]. Protection and Control of Modern Power Systems, 2019, 4(1): 1-13.
- [2] 张显, 王锡凡. 短期电价预测综述[J]. 电力系统自动化, 2006, 30(3): 92-101.
ZHANG Xian, WANG Xifan. Review of electricity price forecasting[J]. Automation of Electric Power Systems, 2006, 30(3): 92-101.
- [3] 贾雪枫, 李存斌. 考虑短期负荷影响的 DeepESN 电力市场实时电价预测研究[J]. 智慧电力, 2021, 49(1): 64-70.
JIA Xuefeng, LI Cunbin. Real-time electricity price forecasting of electricity market using DeepESN considering short-term load impact[J]. Smart Power, 2021, 49(1): 64-70.
- [4] 肖云鹏, 王锡凡, 王秀丽, 等. 面向高比例可再生能源的电力市场研究综述[J]. 中国电机工程学报, 2018, 38(3): 663-674.
XIAO Yunpeng, WANG Xifan, WANG Xiuli, et al. Review on electricity market towards high proportion of renewable energy[J]. Proceedings of the CSEE, 2018, 38(3): 663-674.
- [5] 黄剑平, 陈皓勇, 林镇佳, 等. 需求侧响应背景下分时电价研究与实践综述[J]. 电力系统保护与控制, 2021, 49(9): 178-187.
HUANG Jianping, CHEN Haoyong, LIN Zhenjia, et al. A summary of time-of-use research and practice in a demand response environment[J]. Power System Protection and Control, 2021, 49(9): 178-187.
- [6] 聂涌泉, 顾慧杰, 彭超逸, 等. 激励性含风电竞标日前市场出清电价机制的建模[J]. 电力系统保护与控制, 2020, 48(19): 25-34.
NIE Yongquan, GU Huijie, PENG Chaoyi, et al. Research on modeling the incentive electricity pricing mechanism in day-ahead electricity market clearing containing wind power bidding[J]. Power System Protection and Control, 2020, 48(19): 25-34.
- [7] 彭曙蓉, 李彬, 彭君哲, 等. 基于 CEEMDAN 分解的

- 深度学习分位数回归电价预测[J]. 电力系统及其自动化学报, 2020, 32(4): 138-144.
- PENG Shurong, LI Bin, PENG Junzhe, et al. Electricity price forecasting based on CEEMDAN decomposition and deep learning quantile regression[J]. Proceedings of the CSU-EPSC, 2020, 32(4): 138-144.
- [8] 夏成文, 许凯帅, 鲍玉昆, 等. 基于单值聚类分析的区域居民概率负荷预测研究[J]. 电力信息与通信技术, 2021, 19(1): 1-10.
- XIA Chengwen, XU Kaishuai, BAO Yukun, et al. Research on probabilistic load forecasting for regional residential users based on single value cluster analysis[J]. Electric Power Information and Communication Technology, 2021, 19(1): 1-10.
- [9] NOWOTARSKI J, WERON R. Recent advances in electricity price forecasting: a review of probabilistic forecasting[J]. Renewable & Sustainable Energy Reviews, 2018, 81: 1548-1568.
- [10] 魏华栋, 陶媛, 蔡昌春, 等. 基于改进长短期记忆神经网络的短期负荷预测[J]. 电测与仪表, 2020, 57(19): 93-98.
- WEI Huadong, TAO Yuan, CAI Changchun, et al. Short-term load forecasting based on improved long short-term memory neural network[J]. Electrical Measurement & Instrumentation, 2020, 57(19): 93-98.
- [11] KHOSRAVI A, NAHAVANDI S, CREIGHTON D, et al. Lower upper bound estimation method for construction of neural network-based prediction intervals[J]. IEEE Transactions on Neural Networks, 2011, 22(3): 337-346.
- [12] KHOSRAVI A, NAHAVANDI S, CREIGHTON D. Prediction interval construction and optimization for adaptive neurofuzzy inference systems[J]. IEEE Transactions on Fuzzy Systems, 2011, 19(5): 983-988.
- [13] QUAN H, SRINIVASAN D, KHOSRAVI A. Short-term load and wind power forecasting using neural network-based prediction intervals[J]. IEEE Transactions on Neural Networks and Learning Systems, 2014, 25(2): 303-315.
- [14] WANG J D, FANG K J, PANG W J, et al. Wind power interval prediction based on improved PSO and BP neural network[J]. Journal of Electrical Engineering & Technology, 2017, 12(3): 989-995.
- [15] WAN C, XU Z, PINSON P, et al. Optimal prediction intervals of wind power generation[J]. IEEE Transactions on Power Systems, 2014, 29(3): 1166-1174.
- [16] SHEN Y, WANG X, CHEN J. Wind power forecasting using multi-objective evolutionary algorithms for wavelet neural network-optimized prediction intervals[J]. Applied Sciences-Basel, 2018, 8(2): 185-198.
- [17] SHRIVASTAVA N A, KHOSRAVI A, PANIGRAHI B K. Prediction interval estimation of electricity prices using PSO-tuned support vector machines[J]. IEEE Transactions on Industrial Informatics, 2015, 11(2): 322-331.
- [18] CHAWEEWAT P, SINGH J G. An electricity price interval forecasting by using residual neural network[J]. International Transactions on Electrical Energy Systems, 2020, 30(9): 12506-12521.
- [19] 杨锡运, 关文渊, 刘玉奇, 等. 基于粒子群优化的核极限学习机模型的风电功率区间预测方法[J]. 中国电机工程学报, 2015, 35(增刊 1): 146-153.
- YANG Xiyun, GUAN Wenyuan, LIU Yuqi, et al. Prediction intervals forecasts of wind power based on PSO-KELM[J]. Proceedings of the CSEE, 2015, 35(S1): 146-153.
- [20] 黎敏, 林湘宁, 张哲原, 等. 超短期光伏出力区间预测算法及其应用[J]. 电力系统自动化, 2019, 43(3): 10-16.
- LI Min, LIN Xiangning, ZHANG Zheyuan, et al. Interval prediction algorithm for ultra-short term photovoltaic output and its application[J]. Automation of Electric Power Systems, 2019, 43(3): 10-16.
- [21] 徐诗鸿, 张宏志, 林湘宁, 等. 基于改进评价指标的波动性负荷短期区间预测[J]. 电力系统自动化, 2020, 44(2): 156-163.
- XU Shihong, ZHANG Hongzhi, LIN Xiangning, et al. Improved evaluation index based short-term interval prediction of fluctuation load[J]. Automation of Electric Power Systems, 2020, 44(2): 156-163.
- [22] 殷豪, 曾云, 孟安波, 等. 基于奇异谱分析的短期电价预测[J]. 电力系统保护与控制, 2019, 47(1): 115-122.
- YIN Hao, ZENG Yun, MENG Anbo, et al. Short-term electricity price forecasting based on singular spectrum analysis[J]. Power System Protection and Control, 2019, 47(1): 115-122.
- [23] 冯长有, 王锡凡, 王秀丽, 等. 电价分布及分类预测模型[J]. 电力系统自动化, 2009, 33(6): 25-30.
- FENG Changyou, WANG Xifan, WANG Xiuli, et al. Electricity price distribution and classified forecasting model[J]. Automation of Electric Power Systems, 2009, 33(6): 25-30.
- [24] 董雷, 周文萍, 张沛, 等. 基于动态贝叶斯网络的光伏发电短期概率预测[J]. 中国电机工程学报, 2013, 33(增刊 1): 38-45.
- DONG Lei, ZHOU Wenping, ZHANG Pei, et al. Short-term photovoltaic output forecast based on dynamic Bayesian network theory[J]. Proceedings of the CSEE, 2013, 33(S1): 38-45.

- [25] 庞传军, 张波, 余建明. 基于 LSTM 循环神经网络的短期电力负荷预测[J]. 电力工程技术, 2021, 40(1): 175-180, 194.
PANG Chuanjun, ZHANG Bo, YU Jianming. Short-term power load forecasting based on LSTM recurrent neural network[J]. Electric Power Engineering Technology, 2021, 40(1): 175-180, 194.
- [26] 刘志坚, 杨志华, 黄蓉. 基于社会经济指标影响的电力系统月度负荷组合预测模型[J]. 昆明理工大学学报(自然科学版), 2014, 39(5): 58-64.
LIU Zhijian, YANG Zhihua, HUANG Rong. Monthly load combined forecasting model of power system based on influences of social economic factors[J]. Journal of Kunming University of Science and Technology (Natural Science Edition), 2014, 39(5): 58-64.
- [27] 谢谦, 董立红, 库向阳. 基于 Attention-GRU 的短期电价预测[J]. 电力系统保护与控制, 2020, 48(23): 154-160.
XIE Qian, DONG Lihong, SHE Xiangyang. Short-term electricity price forecasting based on Attention-GRU[J]. Power System Protection and Control, 2020, 48(23): 154-160.
- [28] 任远. 基于回声状态网络的电力市场电价预测[J]. 电力系统保护与控制, 2016, 44(11): 111-115.
REN Yuan. Echo-state-network based electricity price forecasting in electric power market[J]. Power System Protection and Control, 2016, 44(11): 111-115.
- [29] MALONE B, KANGAS K, JARVISALO M, et al. Empirical hardness of finding optimal Bayesian network structures: algorithm selection and runtime prediction[J]. Machine Learning, 2018, 107(1): 247-283.
- [30] CHICKERING D M, HECKERMAN D, MEEK C. Large-sample learning of Bayesian networks is NP-hard[J]. Journal of Machine Learning Research, 2004, 5: 1287-1330.
- [31] 黄解军, 万幼川, 潘和平. 贝叶斯网络结构学习及其应用研究[J]. 武汉大学学报(信息科学版), 2004, 29(4): 315-318.
HUANG Jiejun, WAN Youchuan, PAN Heping. Bayesian network structure learning and its applications[J]. Geomatics and Information Science of Wuhan University, 2004, 29(4): 315-318.
- [32] 张钰莎, 蒋盛益. 连续属性离散化算法研究综述[J]. 计算机应用与软件, 2014, 31(8): 6-8, 140.
ZHANG Yusha, JIANG Shengyi. Survey on continuous feature discretisation algorithm for discretisation[J]. Computer Applications and Software, 2014, 31(8): 6-8, 140.
- [33] 杨恒岳, 刘青荣, 阮应君. 基于 k-means 聚类算法的分布式能源系统典型日冷热负荷选取[J]. 热力发电, 2021, 50(3): 84-90.
YANG Hengyue, LIU Qingrong, RUAN Yingjun. Selection of typical daily cooling and heating load of CCHP system based on k-means clustering algorithm[J]. Thermal Power Generation, 2021, 50(3): 84-90.
- [34] 李武, 赵娇燕, 严太山. 基于平均差异度优选初始聚类中心的改进 K-均值聚类算法[J]. 控制与决策, 2017, 32(4): 759-762.
LI Wu, ZHAO Jiaoyan, YAN Taishan. Improved K-means clustering algorithm optimizing initial clustering centers based on average difference degree[J]. Control and Decision, 2017, 32(4): 759-762.
- [35] 邸若海, 高晓光, 郭志高. 基于改进 BIC 评分的贝叶斯网络结构学习[J]. 系统工程与电子技术, 2017, 39(2): 437-444.
DI Ruohai, GAO Xiaoguang, GUO Zhigao. Bayesian networks structure learning based on improved BIC scoring[J]. Systems Engineering and Electronics, 2017, 39(2): 437-444.
- [36] 赵建军, 王毅, 杨利斌. 基于时间序列预测的威胁估计方法[J]. 浙江大学学报(工学版), 2014, 48(3): 398-403, 422.
ZHAO Jianjun, WANG Yi, YANG Libin. Threat assessment method based on time series forecast[J]. Journal of Zhejiang University (Engineering Science), 2014, 48(3): 398-403, 422.
- [37] IBARGUENGOYTIA P H, REYES A, ROMERO-LEON I, et al. Wind power forecasting using dynamic bayesian models[J]. Nature-Inspired Computation and Machine Learning, Pt II, 2014, 8857: 184-197.
- [38] POOL N. Historical market data: Elspot prices_2017_hourly_eur. xls[EB/OL]. (2011-01-01) [2020-9-20]. <https://www.nordpoolgroup.com/historical-market-data/>.

收稿日期: 2021-04-25; 修回日期: 2021-07-24

作者简介:

王洪涛(1978—), 男, 博士, 副教授, 研究方向为电力系统规划、电力系统风险评估与电力市场等; E-mail: taiyuanwang0801@126.com

邹斌(1965—), 男, 通信作者, 教授, 博士生导师, 研究方向为电力系统规划与电力市场等。E-mail: zoubin@shu.edu.cn

(编辑 周金梅)