

◀测井录井▶

doi:10.11911/syztjs.2019136

基于声波时差数据波动性识别异常压实地层的方法

席鹏飞¹, 杨明合¹, 郭王恒¹, 石建刚²

(1. 中国石油钻井工程重点实验室长江大学研究室, 湖北武汉 434000; 2. 中国石油新疆油田分公司工程技术研究院, 新疆克拉玛依 834000)

摘要: 受地层岩性等因素影响, 声波时差数据会出现波动, 从而影响利用声波时差识别异常压实地层的精度。为此, 依据正常压实泥岩段地层声波时差与井深的关系, 采用了密度聚类法剔除无效数据点, 并应用小波理论及概率分析方法分析了正常压实地层声波时差数据的波动性。研究发现: 密度聚类法能有效去除异常数据点, 对于正常压实地层声波时差数据波动性的描述, 一次小波分解即可满足要求; 分布拟合检验表明, 声波时差数据的波动性符合 tLocation-Scale 概率分布; 通过构建异常压实地层的概率计算公式, 可定量识别异常压实地层。计算结果表明, 通过计算声波时差波动概率能定量识别异常压实地层, 提高识别异常压实地层的精度, 避免分析的盲目性、随意性。对于其他基于泥页岩正常压实曲线的常规测井资料, 均可采用该方法进行分析。

关键词: 异常压实; 声波时差; 波动性; 概率分布

中图分类号: TE27⁺¹ 文献标志码: A 文章编号: 1001-0890(2019)06-0111-05

A Method for Identifying Abnormally Compacted Strata Based on the Fluctuation of Interval Transit Time Data

XI Pengfei¹, YANG Minghe¹, GUO Wangheng¹, SHI Jiangang²

(1. Yangtze University Laboratory, CNPC Key Laboratory for Petroleum Drilling Engineering, Wuhan, Hubei, 434000, China;
2. Engineering Technology Research Institute of CNPC Xinjiang Oilfield Branch, Karamay, Xinjiang, 834000, China)

Abstract: The fluctuation of interval transit time data can be caused by factors such as variable formation lithology which affects the accurate identification of abnormally compacted strata by interval transit time. Therefore, the fluctuation of interval transit time data in normally compacted sections was analyzed according to the relationship between the interval transit time and well depth at normally compacted mudstone section, by eliminating the invalid data points using a density data clustering method, made possible by means of wavelet theory and the probability analysis method. Research results showed that the density clustering method could effectively remove the anomalous data points. One wavelet decomposition could meet the requirement to describe the fluctuation of interval transit time data at the normally compacted section; the distribution fitting test indicated that the fluctuation of interval transit time data was consistent with tLocation-Scale probability distribution. Thus the abnormally compacted section could be quantitatively identified by constructing a probability calculation formula. The calculation results showed that the abnormally compacted section could be quantitatively identified by calculating the fluctuation probability of interval transit time, improve the identification accuracy of abnormally compacted section, and avoid the blindness and randomness of analysis. Similar analyses can be performed for other conventional logging data analyses based on the normal compaction curve of shale.

Key words: abnormally compacted; interval transit time; fluctuation; probability distribution

声波时差法是计算异常压实地层压力的一种常用方法, 其核心是在确定正常压实段声波时差趋势线基础上, 准确计算异常声波时差点偏离趋势线的程度^[1-3]。由于声波时差受地层岩性等因素的影响会出现波动, 影响了利用声波时差识别异常压实地层的精度。现场往往是结合实钻及测井等资料识别

收稿日期: 2019-06-01; 改回日期: 2019-10-24。

作者简介: 席鹏飞 (1993—), 男, 河南南阳人, 2017 年毕业于长江大学工程技术学院石油工程专业, 在读硕士研究生, 主要从事钻井工艺技术研究。E-mail: 201772129@yangtzeu.edu.cn。

通信作者: 杨明合, yangminghe76@qq.com。

基金项目: 国家科技重大专项“准噶尔盆地致密油开发示范工程”(编号: 2017ZX05070-05) 资助。

异常压实地层、计算异常压实地层的压力。该方法的识别精度取决于个人的经验及对地层的认识程度,不可避免地存在随意性和盲目性,因而异常压实地层的识别精度不高。为提高识别异常压实地层的精度、准确预测异常压实地层的压力,除需要对声波时差数据进行预处理外,还要进一步研究声波时差数据自身包含的规律。为此,笔者采用密度聚类法来剔除声波时差数据中的无效数据,基于小波理论分析了正常压实段内声波时差沿正常压实趋势线的波动性^[4-6]。通过分布拟合检验,确定了声波时差波动性的概率分布函数,从概率的角度辨识异常压实地层,避免了实际计算中的随意性,该方法同样适用于其他类似测井资料(如密度)的分析。

1 正常压实段声波时差数据的分布拟合

对于正常压实地层,井深与声波时差的自然对数呈线性关系^[7]:

$$\ln \Delta t = -kH + \ln \Delta t_0 \quad (1)$$

式中: Δt 为地层声波时差, $\mu\text{s}/\text{m}$; H 为井深, m ; Δt_0 为初始井深处的地层声波时差, $\mu\text{s}/\text{m}$; k 为地层压实系数。

1.1 剔除无效数据点

由于受地层、井径和测量仪器等因素的影响,声波时差数据中含有较多无效数据,笔者采用密度聚类(dbSCAN)法对声波时差数据进行预处理^[8-9]。图1为新疆油田玛湖凹陷A井 501.00~4 190.00 m 井段 21 297 个声波时差数据聚类分析结果,图中 580 个噪声点和 769 个边界点为无效数据,其余为核心点,聚为 2 簇。初步判断第 1 簇核心点位于正常压实地层,对应 508.00~2 965.00 m 井段。

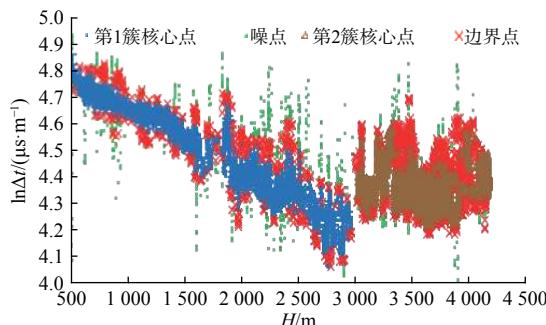


图1 A井声波时差数据密度聚类分析结果

Fig. 1 The results of density clustering analysis by interval transit time in Well A

1.2 正常压实段声波时差数据的波动性分析

图2所示为图1中正常压实井段(第1簇核心点)的散点图。从图2可以看出,该井段声波时差的自然对数与井深明显呈线性关系,拟合关系式为: $\ln \Delta t = -2.34 \times 10^4 H + 4.90, R^2 = 0.9170$ 。

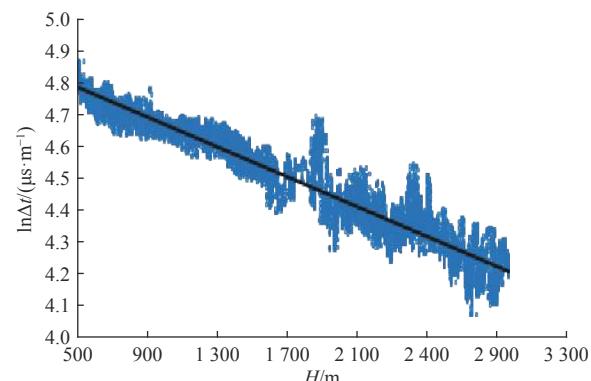


图2 正常压实井段 $\ln \Delta t$ 沿井深的分布

Fig. 2 The distribution of $\ln \Delta t$ along well depth at normally compacted section

由图2可知,随井深 H 增大, $\ln \Delta t$ 在趋势线两侧波动显著,显然受地层因素影响较大。这种波动可用小波分析进行描述^[10-11]。依据小波分析理论,数据点沿其趋势线方向的波动可用如下公式描述:

$$y = a_0 + \sum_{i=1}^n [a_i \sin(\omega_i x) + a'_i \cos(\omega'_i x)] \quad (2)$$

式中: y 为原始信号; a_0 对应 y 的低频信号; ω_i 和 ω'_i 为相位; a_i 和 a'_i 为振幅, $\sum_{i=1}^n [a_i \sin(\omega_i x) + a'_i \cos(\omega'_i x)]$ 为对应 y 的高频信号。

若把式(1)看作式(2)中的低频信号,即 $a_0 = -kH + \ln \Delta t_0$;令 $y = \ln \Delta t$,那么正常压实井段 $\ln \Delta t$ 的波动性用小波理论近似描述为:

$$\ln \Delta t = (-kH + \ln \Delta t_0) + \Delta \quad (3)$$

其中, $\Delta = \sum_{i=1}^n [a_i \sin(\omega_i x) + a'_i \cos(\omega'_i x)]$, Δ 为 $\ln \Delta t$ 沿着趋势线的波动项。

实际上,若设低频信号为直线,可利用点到直线的距离公式,直接导出式(3)。对于图2中的任意一点,若其位于趋势线上方,则该点到趋势线的距离为:

$$\Delta_i = [\ln \Delta t_i - (-kH_i + \ln \Delta t_0)] \cos(\arctan k) \quad (4)$$

式中: Δ_i 为第 i 个点到趋势线的距离; H_i 为第 i 个点所在的井深, m 。

当趋势线斜率 k 很小时, $\cos(\arctan k) \approx 1$, 则式

(4) 可改写为:

$$\ln \Delta t_i = (-kH_i + \ln \Delta t_0) + \Delta_i \quad (5)$$

若点 i 位于趋势线下方, 则该点到趋势线的距离为 $-\Delta_i$ 。

根据式(5), 可以得到波动项 Δ 沿井深的分布, 如图 3 所示。很明显, Δ 在水平方向上同样呈现出波动性, 可用小波理论对 Δ 进一步进行分解。统计图 3 中波动项 Δ 的分布频率, 结果如图 4 所示。图 4 中, Δ 的分布集中于均值附近, 故对 $\ln \Delta t$ 应用一次小波分解即可满足要求。

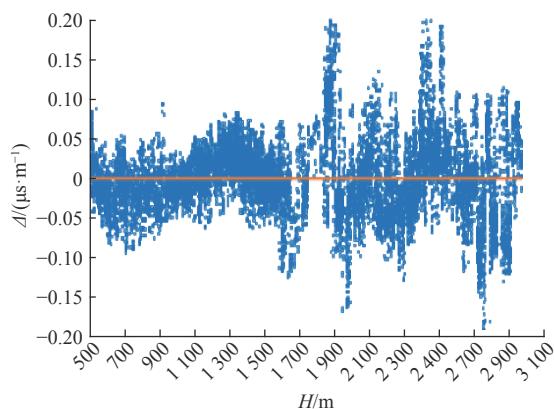


图 3 正常压实段波动项 Δ 沿井深的分布

Fig. 3 Distribution of fluctuation item Δ along the depth of well at normally compacted section

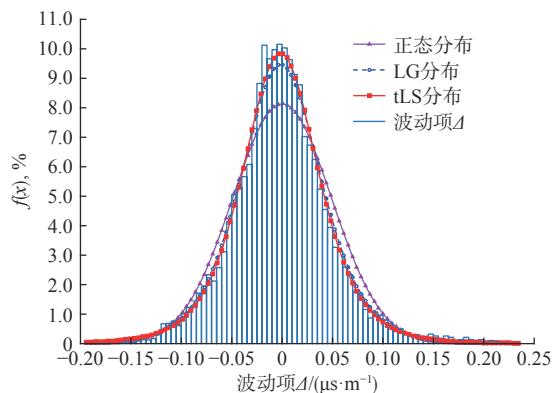


图 4 波动项 Δ 分布的拟合

Fig. 4 The fitting of the fluctuation item Δ distribution

1.3 正常压实段声波时差数据频率分布

从图 4 可以看出, 波动项 Δ 集中于均值附近, 表现为尖峰特征, 其峰值大于正态分布的峰值。对这种集中明显的情况, 一般使用 Logistic(以下简记为 LG) 分布或 t Location-Scale(以下简记为 tLS) 分布进行拟合, 其结果明显优于正态分布^[12-13]。

tLS 分布和 LG 分布的概率密度函数分别为:

$$f(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sigma \sqrt{\nu\pi}} \left[1 + \frac{(x-\mu)^2}{\nu} \right]^{-\frac{\nu+1}{2}} \quad (6)$$

$$f(x) = \frac{e^{-\frac{x-\mu}{\sigma}}}{\sigma^2 [1 + e^{-\frac{x-\mu}{\sigma}}]^2} \quad (7)$$

式中: $\Gamma(x)$ 为伽马函数; μ 为位置参数; σ 为尺度参数; ν 为形状参数。

在 Matlab 软件的 Distribution Fitting 模块中, 分别用 tLS 分布和 LG 分布的概率密度函数拟合波动项 Δ 的分布规律, 结果见图 4。由图 4 可知, 相较于正态分布, tLS 分布和 LG 分布更加接近波动项 Δ 的分布频率。拟合确定的 tLS 分布和 LG 分布概率密度函数的参数值见表 1。

表 1 tLS 分布和 LG 分布概率密度函数的参数

Table 1 Probability density function of the tLS distribution and the LG distribution

函数分布	μ	σ	ν
tLS	-1.56×10^{-3}	3.85×10^{-2}	4.96
LG	-1.25×10^{-3}	2.64×10^{-2}	

1.4 波动项 Δ 频率分布的拟合检验

分布拟合检验基于假设检验, 通过检验给定分布函数与样本分布的符合程度, 从而对该分布做出取舍^[14]。数据点数 n 大于 2 000 时, 常用皮尔逊卡方 (χ^2) 和柯尔莫哥洛夫-斯米诺夫 (K-S) 检验^[15-16], 笔者采用 χ^2 进行拟合检验。

皮尔逊检验统计量 χ^2 为:

$$\chi^2 = \sum_{i=1}^k \frac{n}{p_i} \left(\frac{f_i}{n} - p_i \right) = \sum_{i=1}^k \frac{(f_i - np_i)^2}{np_i} \quad (8)$$

式中: n 为试验次数; k 为样本空间划分的子集(或区间)的个数; p_i 为第 i 个子集上给定分布的概率; f_i 为第 i 个子集上样本出现的频数。

一般当总体分布未知时, 检验假设 H_{01} : 总体分布函数为 tLS, 即 $X \sim F(\mu, \sigma, \nu)$; H_{02} : 总体分布函数为 LG, 即 $X \sim L(\mu, \sigma)$ 。

当试验次数 n 充分大 ($n=13\,099 \geq 50$), 则当 H_0 为真时, 统计量 χ^2 近似服从自由度为 $k-1$ 的 χ^2 分布^[16], 故在给定显著性 α 的条件下, H_0 的拒绝域为:

$$\chi^2 \geq \chi_{\alpha}^2(k-1) \quad (9)$$

如果把波动项 Δ 包含的所有数据看作样本空间, 那么按照样本值从小到大的顺序, 可以把样本空间

$[-0.19, 0.23]$ 划分为 k 个区间(子集)。图 5 给出了当 $k=50$ 时, tLS 分布、LG 分布及正态分布在各区间上概率 p_i 与频率 f_i/n 的对应关系。

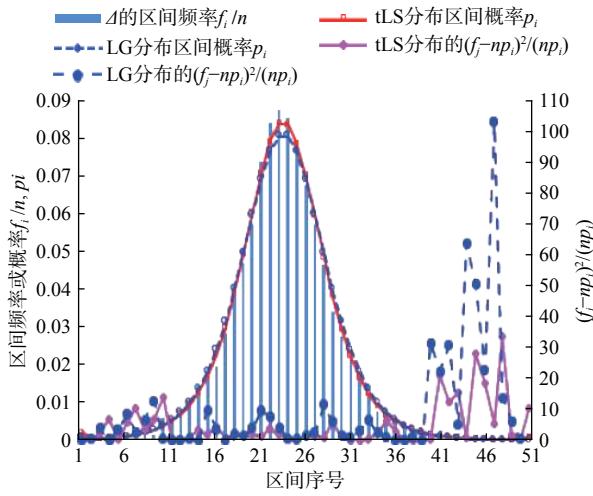


图 5 $k=50$ 时 2 种分布与样本的频率拟合

Fig. 5 Frequency fitting of two kinds of distributions and samples when $k=50$

从图 5 可以看出, 第 i 个区间上 tLS 分布和 LG 分布的 $(f_i - np_i)^2 / (np_i)$ 值在区间中心区域小、边缘大。进一步分析发现, 在边缘区间内, 波动项 Δ 的频数 f_i 都很小(<50), 相对样本总量可以忽略, 在计算 χ^2 时需忽略这些将导致 χ^2 明显增大的边缘区间, 即有效样本区间为 11~41 区间。

当显著性水平 $\alpha=0.05$ 时, 根据式(8)计算出 tLS 分布的 $\chi^2=51.1$, 而 $\chi^2_{0.05}(50-1)=66.339$, 不满足式(9), 接受 H_{01} , 认为波动项 Δ 的分布符合 tLS 分布。同理, 计算出 LG 分布的 $\chi^2=84.9$, 满足式(9), 则拒绝 H_{02} , 即认为波动项 Δ 的分布不符合 LG 分布。

2 异常压实段概率分析

对于异常压实段的声波时差, 其 $\ln \Delta t$ 将会明显偏离 tLS 分布的位置参数 μ , 记 Δt_i 与位置参数 μ 的距离为 D_i , 即:

$$D_i = |\Delta t_i - \mu| \quad (10)$$

对于第 i 点声波时差 Δt_i , D_i 越大, Δt_i 与位置参数 μ 的距离越大, 也就是 $\ln \Delta t_i$ 到“由式(1)确定趋势线”的距离就越大, 即 Δt_i 属于正常压实段的可能性越小。据此可以构建 Δt_i 属于正常压实段概率的计算公式。从 tLS 概率密度函数图(见图 6)可以看出, 对于任意 Δt_i , 可以用图中阴影面积 S 来表示 Δt_i 距离 μ 的远近程度。结合概率定义, 那么 Δt_i 属于正常压实段的概率 p_i 可表示为:

$$p_i = 2 \int_{-\infty}^{\mu - D_i} f(x) dx = 2P[-\infty < x \leq (\mu - D_i)] = 2F(\mu - D_i) \quad (11)$$

式中: $f(x)$ 、 $P(x)$ 和 $F(x)$ 为 tLS 分布的概率密度函数、概率函数和概率分布函数, 其参数可从上文求出。

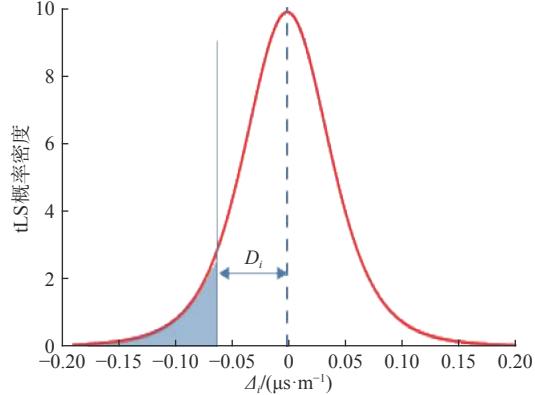


图 6 正常压实段波动项 Δ_i 的 tLS 分布

Fig. 6 tLS distribution of fluctuation item Δ_i at normally compacted section

当 $D_i = 0$ 时, $\Delta t_i = \mu$, $\ln \Delta t_i$ 落在趋势线上, 此时 $p_i=1$, 即 Δt_i 属于正常压实段的概率为 1。

A 井整个井深范围内的 Δt_i 、 p_i 及 tLS 分布的 3σ 直线 ($D_i \in [0, 3\sigma]$ 时认为 Δt_i 波动正常) 如图 7 所示。

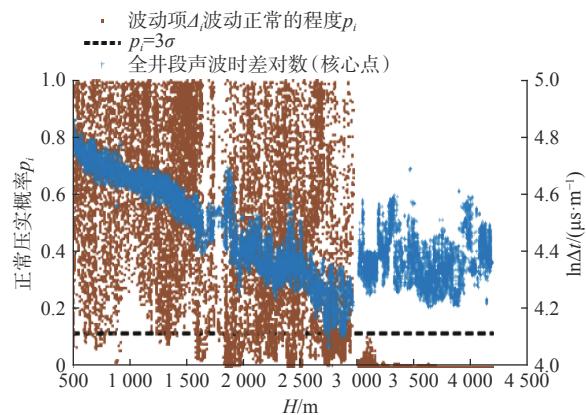


图 7 A 井 $\ln \Delta t$ 及 p_i 沿井深的分布

Fig. 7 Distributions of $\ln \Delta t$ and p_i along well depth in Well A

从图 7 可以看出: 500.00~2 950.00 m 井段绝大部分 p_i 大于 0.12, 位于 3σ 直线上方。若以 3σ 作为判断标准, 则该段属于正常压实段; 井深 2 950.00 m 以深井段的 p_i 小于 0.02, 属于正常压实段的概率基本为 0, 为异常压实段。实钻资料显示, 该井在井深 2 940.00 m 左右压力明显抬升, 地层压力系数由

1.04 升至 1.71, 证明该方法识别异常压实段的可靠性较好。

同一区块 B 井整个井深范围内的 Δt_i 、 p_i 及 tLS 分布的 3σ 直线如图 8 所示。由图 8 可知, 在井深 3 030.00 m 左右概率 p_i 开始降低, 对应地层压力开始升高; 井深 3 200.00 m 以深井段, p_i 绝大部分在 3σ ($\sigma=0.047$) 直线以下, 说明该地层属于正常压实段的概率小于 14%, 判断为异常压实段。

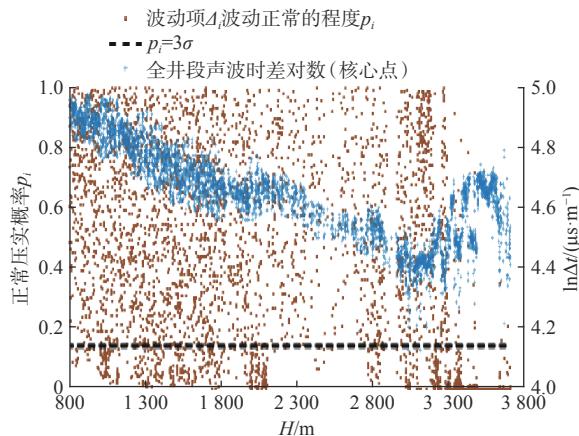


图 8 B 井 $\ln\Delta t$ 及 p_i 沿井深的分布

Fig. 8 Distributions of $\ln\Delta t$ and p_i along well depth in well B

3 结论与建议

1) 受地层岩性等因素影响, 声波时差数据会出现波动而影响识别异常压实地层的精度。对于去除异常数据后的正常压实段的声波时差数据, 其波动性可用小波理论进行分析, 分布拟合检验表明, 这种波动性可用 tLocation-Scale 分布进行描述。

2) 基于 tLocation-Scale 分布, 构建识别异常压实段的概率计算公式, 从概率的角度辨识异常压实地层, 并用概率量化了该辨识的可靠程度, 避免了计算的盲目性、随意性。

3) 该方法同样适用于其他基于泥页岩正常压实趋势线的常规测井资料(如密度、伽马及电阻率等)分析, 具有一定的推广应用价值。

参 考 文 献

References

- [1] CHILINGAR G V, SEREBRYAKOV V A, ROBERTSON J O Jr. Origin and prediction of abnormal formation pressures[J]. Developments in Petroleum Science, 2002, 50(1): 5–9.
- [2] MOUCHET J P, MITCHELL A. Abnormal pressures while drilling[M]. Boussens: Elf Aquitaine Edition, 1989.
- [3] 王鸿升. 异常高地层压力研究现状概述 [J]. 油气地球物理, 2015, 13(4): 58–63.
- [4] WANG Hongsheng. Summary of research status for abnormal high formation pressure[J]. Petroleum Geophysics, 2015, 13(4): 58–63.
戴威, 吴健. 应用小波降噪的测井数据预测异常压力 [J]. 化学工程与装备, 2018(9): 185–189.
- [5] DAI Wei, WU Jian. Prediction of abnormal pressure using logging data with wavelet denoising[J]. Chemical Engineering & Equipment, 2018(9): 185–189.
赖富强, 夏炜旭, 龚大建, 等. 基于小波高频属性的泥页岩裂缝测井识别方法研究 [J/OL]. 地球物理学进展. <http://kns.cnki.net/kcms/detail/11.2982.P.20190307.1305.028.html>.
- [6] LAI Fuqiang, XIA Weixu, GONG Dajian, et al. Logging identification method of mud shale fractures based on wavelet high frequency attribute[J/OL]. Progress in Geophysics. <http://kns.cnki.net/kcms/detail/11.2982.P.20190307.1305.028.html>.
- [7] 林卫星, 文劲宇, 艾小猛, 等. 风电功率波动特性的概率分布研究 [J]. 中国电机工程学报, 2012, 32(1): 38–46.
- [8] LIN Weixing, WEN Jinyu, AI Xiaomeng, et al. Probability density function of wind power variations[J]. Proceedings of the CSEE, 2012, 32(1): 38–46.
- [9] 陈庭根, 管志川. 钻井工程理论与技术 [M]. 东营: 石油大学出版社, 2000: 11–13.
- [10] CHEN Tinggen, GUAN Zhichuan. The theory and technology about drilling engineering[M]. Dongying: Petroleum University Press, 2000: 11–13.
- [11] 伍育红. 聚类算法综述 [J]. 计算机科学, 2015, 42(增刊 1): 491–499, 524.
- [12] WU Yuhong. General overview on clustering algorithms[J]. Computer Science, 2015, 42(supplement1): 491–499, 524.
- [13] 王俊. 面向大数据的 DBSCAN 聚类算法研究 [D]. 郑州: 解放军信息工程大学, 2017.
- [14] WANG Jun. Research on DBSCAN clustering algorithm oriented big data[D]. Zhengzhou: PLA Information Engineering University, 2017.
- [15] 马丽君, 孙根年, 王洁洁. 基于本底趋势线与小波函数的中国旅游成长及多周期分析 [J]. 旅游科学, 2009, 23(6): 21–27.
- [16] MA Lijun, SUN Genian, WANG Jiejie. An analysis of China's tourism growth and multicycle: Based on natural trend curve theory and wavelet analysis[J]. Tourism Science, 2009, 23(6): 21–27.
- [17] 王晓兰, 王明伟. 基于小波分解和最小二乘支持向量机的短期风速预测 [J]. 电网技术, 2010, 34(1): 179–184.
- [18] WANG Xiaolan, WANG Mingwei. Short-term wind speed forecasting based on wavelet decomposition and least square support vector machine[J]. Power System Technology, 2010, 34(1): 179–184.
- [19] 王明亮, 孙德宙. Logistic 分布预测林分直径结构的研究 [J]. 林业科学, 1998(5): 86–90.
- [20] WANG Mingliang, SUN Dezhou. The forecasting of diameter-based clustering by using of Logistic distribution in forest[J]. Forest Research, 1998(5): 86–90.
- [21] 杨茂, 董骏城. 基于混合分布模型的风电功率波动特性研究 [J]. 中国电机工程学报, 2016, 36(增刊 1): 69–78.
- [22] YANG Mao, DONG Juncheng. Study on characteristics of wind power fluctuation based on mixed distribution model[J]. Proceedings of the CSEE, 2016, 36(supplement1): 69–78.
- [23] 费鹤良. 分布拟合优度检验方法综述 [J]. 上海师范学院学报(自然科学版), 1982, 11(2): 129–142.
- [24] FEI Heliang. General overview on test method of distribution goodness of fit[J]. Journal of Shanghai Normal University(Natural Sciences), 1982, 11(2): 129–142.
- [25] 茅诗松. 统计手册 [M]. 北京: 科学出版社, 2003: 169–172.
- [26] MAO Shisong. Statistic brochure[M]. Beijing: Science Press, 2003: 169–172.
- [27] 盛骤, 谢式千, 潘承毅, 等. 概率论与数理统计 [M]. 北京: 高等教育出版社, 2008: 198–205.
- [28] SHENG Zhou, XIE Shiqian, PAN Chengyi, et al. Statistics and data analysis[M]. Beijing: Higher Education Press, 2008: 198–205.