

文章编号: 1000-0747(2018)02-0312-08 DOI: 10.11698/PED.2018.02.14

基于流线聚类人工智能方法的水驱油藏流场识别

贾虎, 邓力琿

(油气藏地质与开发工程国家重点实验室 西南石油大学, 成都 610500)

基金项目: 中国石油科技创新基金(2017D-5007-0202)

摘要: 以某碳酸盐岩油藏注水开发为例, 提出针对流线模拟结果的流场识别方法。在流线模拟计算完成后, 利用基于 Ocean 平台自行编写的插件将流线数据导出, 并通过 Python 编程语言进行后续数据处理及聚类分析, 直观反映不同开发阶段水驱油藏流场分布。采用密度峰值聚类算法作为流线聚类主要算法, 以轮廓系数算法作为聚类评价算法, 选取合理的聚类数, 并对不同聚类算法结果进行对比。当聚类数相同时, 密度峰值聚类算法比 K-means、层次聚类、谱聚类算法对不同类型流线区分能力更强且轮廓系数较高, 说明了算法的有效性。依据流线聚类结果可对流场进行量化处理, 有效识别油藏中无效注水循环通道以及具有开发潜力的区域, 同时可对同一注采井间流线进行细分, 描述注采井间水相驱动能力大小分布, 为注水优化、井网层系调整、深部调剖等方案决策提供依据。图 10 表 3 参 23

关键词: 水驱开发; 水驱效率; 流场识别; 流线模拟; 聚类算法; 人工智能

中图分类号: TE344; TE341

文献标识码: A

Oil reservoir water flooding flowing area identification based on the method of streamline clustering artificial intelligence

JIA Hu, DENG Lihui

(State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation in Southwest Petroleum University, Chengdu 610500, China)

Abstract: For the case of carbonate reservoir water flooding development, the flow field identification method based on streamline modeling result was proposed. The Ocean for Petrel platform was used to build the plug-in that exported the streamline data, and the subsequent data was processed and clustered through Python programming, to display the flow field with different water flooding efficiencies at different time in the reservoir. We used density peak clustering as primary streamline cluster algorithm, and Silhouette algorithm as the cluster validation algorithm to select reasonable cluster number, and the results of different clustering algorithms were compared. The results showed that the density peak clustering algorithm could provide better identified capacity and higher Silhouette coefficient than K-means, hierarchical clustering and spectral clustering algorithms when clustering coefficients are the same. Based on the results of streamline clustering method, the reservoir engineers can easily identify the flow area with quantification treatment, the inefficient water injection channels and area with developing potential in reservoirs can be identified. Meanwhile, streamlines between the same injector and producer can be subdivided to describe driving capacity distribution in water phase, providing useful information for the decision making of water flooding optimization, well pattern adjustment and deep profile.

Key words: water flooding; water flooding efficiency; flow field identification; streamline simulation; cluster algorithm; artificial intelligence

引用: 贾虎, 邓力琿. 基于流线聚类人工智能方法的水驱油藏流场识别[J]. 石油勘探与开发, 2018, 45(2): 312-319.

JIA Hu, DENG Lihui. Oil reservoir water flooding flowing area identification based on the method of streamline clustering artificial intelligence[J]. Petroleum Exploration and Development, 2018, 45(2): 312-319.

0 引言

中国陆相沉积的碎屑岩和海相沉积的碳酸盐岩油藏, 储集层非均质性较强, 且目前许多高含水砂岩油藏历经长期水驱开发后剩余油分布杂乱且分散, 难以有效认识水驱油藏动用规律, 导致注水开发效率低下。同时地层在经历长时间流体冲刷后易形成优势通道,

导致无效水循环^[1], 降低水驱效率。为了更好描述水驱油藏特征, 前人通过水驱特征曲线^[2]及井间连通性模型^[3]对水驱流场进行了表征。同时, 侯玉培等^[1,4-5]在流场识别方面做了大量研究, 建立了流场评价体系, 采取层次分析法对流场进行了评价, 但层次分析法需要人为设置评价权重, 因此方案决策存在不确定性。在地质模型基础上, 前人通过流线模拟对生产数据进行

拟合并分析了流体运移规律及流动轨迹^[6-10]。与根据孔渗饱分布确定流场强弱的常规方法相比,流线模拟可更准确且直观地显示出流体主要流通区域。此外,可依据注水对采油的贡献量对水驱油藏进行注水优化、水驱控制、井网优化^[11-18]。主流的油藏数值模拟软件如 Petrel RE 虽可准确计算流线分布并依据上述方法进行方案优化,但缺乏量化分析流场的方法,且离散的流线分布无法代表实际流场,也无法区分水相驱动能力不同的流动区域。通过对商业软件 Petrel RE 进行二次开发实现了流线聚类功能,可对水相驱动能力不同的流线进行区分并加以分析,为水驱流场识别提供了一种新的研究手段。

聚类方法又被称为无监督学习,是人工智能方法的分支之一。不同于通过注采井间配位对流线进行归类的分析方式,本文采用聚类算法^[19]对流线依据其空间位置及流线性质进行聚类,提取出流场潜在分布结构。聚类算法以最大化类与类间的差异并最小化类内差异为原则对流线进行聚类,能确保每一类流线均可最大程度表征其所在的流场,同时区分出不同类型的流线,该类无监督学习算法对经验公式与人为评价依赖度较小,因此对不同类型油藏均具有较好的适用性。

本文以某碳酸盐岩油藏为例,通过流线聚类方法,对水相驱动能力不同的流线进行区分并进一步对同一注采井间流线进行细分,开展流场精细化描述,找出潜在优势流场,为水驱流场调整与提高采收率方案决策提供科学依据与技术支撑。

1 流线聚类

流线模拟以流线表示流体长期冲刷而形成的优势通道,模拟地层流体的流动轨迹。本文通过流线模拟方法表征流场,以油藏流体在一定时间步长内均沿流线方向流动为基本假设,流线轨迹通过流线追踪算法^[20]获取,在获取流线轨迹后通过求解沿流线方向的一维质量守恒方程,即可得到流体沿流线的饱和度分布,并显示油藏流体流动规律。

流线模拟结果表征了流体性质在离散流线轨迹上的分布,因此需对流线信息加以提取才可得到流场分布的具体描述。本文通过流线聚类的方法将流场分为不同流动区域,识别出不同类型流场分布以及该类流场特性,通过这些信息便可针对性地进行后续油藏开发方案设计。流线聚类步骤包括特征参数提取,聚类分析及聚类评价。

1.1 特征参数提取

聚类算法需依据样本特征对其进行聚类,因此需提取出具有区分度且便于实际应用的流线特征,本文提取的流线特征及公式见表1。其中流线油水体积比由 V_{ow} 表示,若该值较大,则代表沿该流线方向未被注入水波及到的油较多。流线油水流速比由参数 v_{ow} 表示,该值越大,则表示沿该流线方向水驱油能力越强。

表1 流线特征及公式

流线特征描述	公式
流线中点位置 x 轴坐标	$X_a = \sum_{b=1}^{n_a} x_{ab} / n_a$
流线中点位置 y 轴坐标	$Y_a = \sum_{b=1}^{n_a} y_{ab} / n_a$
油水体积比	$V_{ow} = \sum_{b=1}^{n_a} S_{oab} / \sum_{b=1}^{n_a} S_{wab}$
油水流速比	$v_{ow} = \sum_{b=1}^{n_a} v_{oab} / \sum_{b=1}^{n_a} v_{wab}$

因为流线在 z 轴方向跨度不大,故选择流线中点 x , y 轴坐标表征流线位置。同时,由于流线模拟计算的数值耗散问题,部分流线可能出现油水两相流速均接近 0 的情况,因此无法通过流体流速直接表征流线位置的水相驱动能力。此外,对于水驱后期的流场,水相渗透率往往较高,而流体饱和度为线性增加的属性,因此对驱动能力区分效果不明显,故本文采用油水体积比及油水流速比表征流线位置水相驱动能力,该方式可对水相饱和度较低而油相饱和度较高的流线进行细分,同时对水相饱和度大于一定程度的流线进行一定的粗化,从而最大程度反映不同水相驱动能力的流场分布。

参数选取的意义在于将流线依据其空间位置、驱替效率、波及效率进行聚类,将具有类似性质的流线归为一类,从而识别出具有开发价值的油藏区域,为后期注水优化、井网层系调整、深部调剖等方案决策提供科学依据和技术支撑。

在进行特征提取后,需对每条流线的特征参数进行归一化处理以便之后的聚类分析。

1.2 聚类分析

聚类分析是一种用于寻找数据之间内在结构的技术。聚类算法有密度峰值、K-means、谱聚类、层次聚类算法等,其中由于密度峰值聚类算法对数据结构区分能力较强且其结果不具有随机性,因此本文采用此法。密度峰值聚类算法^[19]的基本原理为,依据与样本性质类似的样本数量的多少作为样本密度,选取局部区域内密度最大点作为聚类中心,同时确保其与其他

密度更大的样本之间的“距离”相对更大。密度峰值聚类算法流程图见图 1。

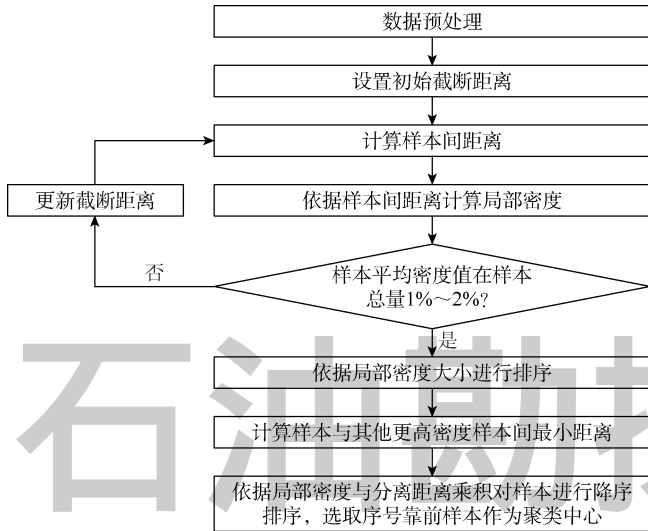


图 1 密度峰值聚类算法流程图

样本 i 的局部密度计算公式为^[19]：

$$\rho_i = \sum_j \chi(d_{ij} - d_c) \quad (i = 1, 2, \dots, n_p) \quad (1)$$

$$\chi(d_{ij} - d_c) = \begin{cases} 1 & (d_{ij} - d_c < 0) \\ 0 & (d_{ij} - d_c \geq 0) \end{cases}$$

设 d_c 可使样本平均局部密度为总样本数的 1%~2%^[19], d_{ij} 通过欧几里得距离公式计算。(1) 式计算的结果是离散值, 且易影响算法稳定性, 因此笔者采用指数函数的形式计算局部密度:

$$\rho_i = \sum_{j \neq i} e^{-\left(\frac{d_{ij}}{d_c}\right)^2} \quad (i=1, 2, \dots, n_p) \quad (2)$$

同时, 需保证样本与其他比该样本密度更大的样本有较大距离, 设 ρ_{q_i} 为样本局部密度降序排列的样本序号, 即 $\rho_{q_1} > \rho_{q_2} > \dots > \rho_{q_{n_p}}$, 则可定义样本分离距离 δ_{q_i} 为:

$$\delta_{q_i} = \begin{cases} \min_{q_j, j < i} d_{q_i, q_j} & (i \geq 2) \\ \max_j d_{q_i, j} & (i=1) \end{cases} \quad (3)$$

从而得到不同样本的局部密度及分离距离分布 (见图 2)。

聚类中心应同时具有较大的局部密度与分离距离值, 因此图 2 内偏离横、纵坐标较远处的点可作为聚类中心点, 第 i 个样本的中心点评价公式为:

$$\gamma_i = \rho_i \delta_{q_i} \quad (4)$$

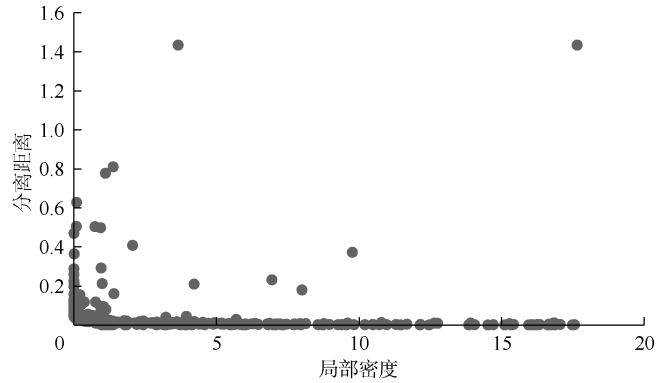


图 2 样本局部密度与分离距离分布图

若要将总体样本分为 n_c 类, 即可选取依据 γ 降序排列在前 n_c 的样本作为聚类中心点。Rodriguez 等^[21]通过对人工随机生成数据集分析发现, 若以降序排序后的 γ 值序号作为横轴, $\lg \gamma$ 值作为纵轴, 则 $\lg \gamma$ 曲线从聚类中心过渡到非聚类中心的曲线趋势有明显跳跃, 可为聚类中心选取提供参考, 因此本文取 $\lg \gamma$ 值排序曲线趋势出现跳跃的位置作为总样本类数, 即聚类数 n_c 。确定聚类中心后, 将非聚类中心样本与距离最近的聚类中心归为一类。

1.3 聚类评价

聚类数对聚类效果有着极大影响: 聚类数过小将粗化数据, 丢失有效信息; 聚类数过大将致使聚类结果无法有效压缩, 造成分析困难。同时, 不同聚类算法如密度峰值、K-means、谱聚类、层次聚类算法等结果往往不尽相同, 因此, 有必要对聚类结果进行评价, 从而选择合理的聚类算法及聚类数。Liu Y 等^[22]在已知样本真实类别数的情况下, 分析了不同聚类评价算法如 Calinski-Harabasz、指数、Dunn、轮廓系数、Davies-Bouldin、Xie-Beni、SD、S-Dbw 等是否具有识别出真实样本聚类数的能力。轮廓系数与 S-Dbw 算法同其他聚类评价算法相比, 对不同的样本类型均能有效识别出样本实际的类别数, 然而 S-Dbw 算法对离散点较敏感, 结果易出现不连续情况。因此本文采用轮廓系数算法^[23], 计算样本的轮廓系数, 根据已评价聚类后的样本与同类样本间平均距离以及样本与非同类样本间平均距离, 对聚类效果进行评价, 轮廓系数值为 -1~1, 趋近于 1 时代表聚类效果相对较优。

2 流场表征

图 3 为某碳酸盐岩油藏内孔隙度及渗透率分布, 由于溶孔溶洞的存在, 部分区域孔隙度及渗透率较高, 为流体提供了较好的渗流条件。图 4 为用 Petrel RE 软

件生成的该油藏注水开发6年后通过流线模拟得到的流线分布。图4a为依据流线起点进行分类的结果，图4b为油相饱和度在流线上的分布。在流线分布图中，流线越密集则流动强度越大，由图4流线模拟结果可

看出流体流动区域主要集中于渗透率、孔隙度均较高，即溶孔溶洞较为发育的区域，如I4井至P4、P6井间区域，这些区域内渗流条件较好，因此易形成注入水流动的优势通道。

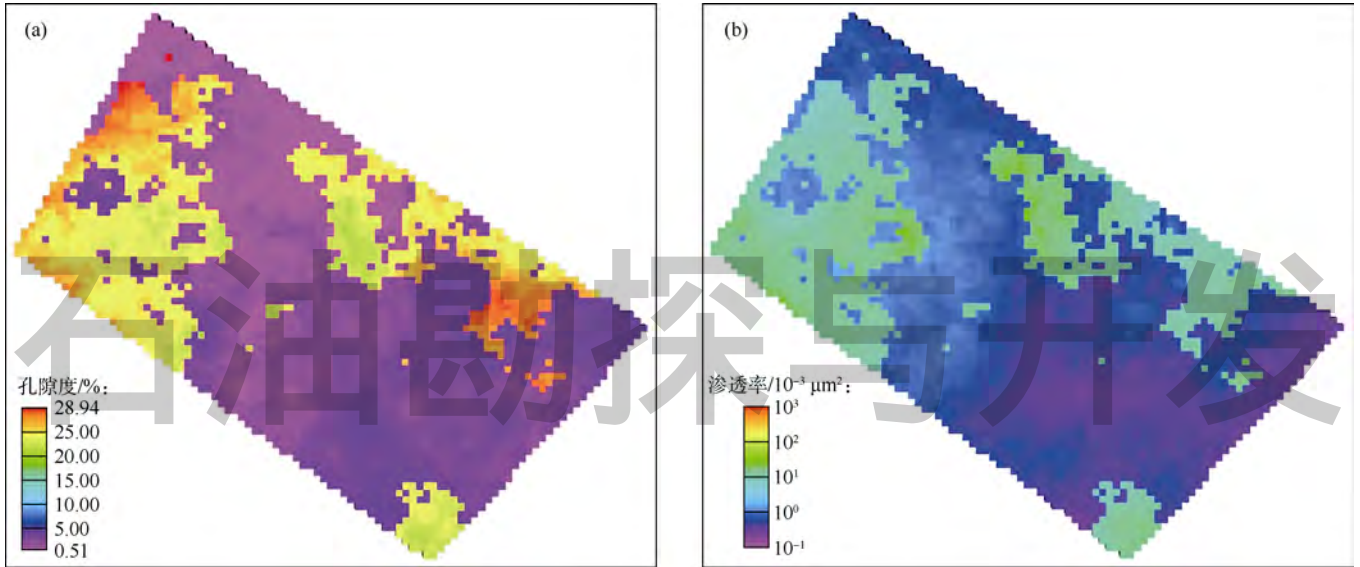


图3 某碳酸盐岩油藏孔隙度、渗透率分布

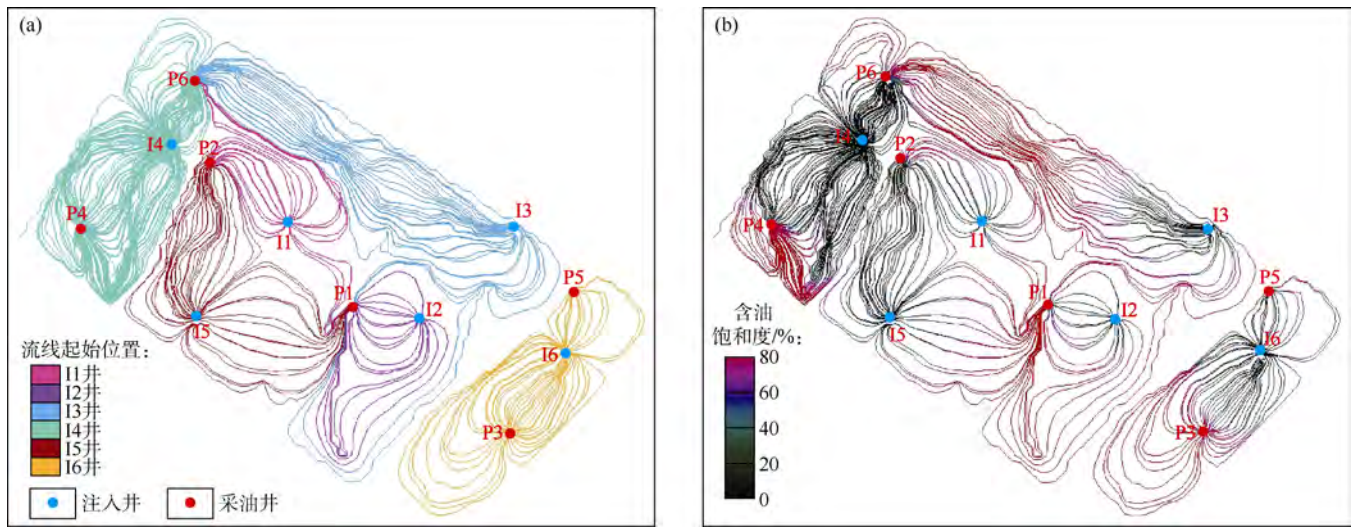


图4 通过Petrel RE软件生成的开发6年后流线分类(a)及流线油相饱和度分布(b)

2.1 聚类参数及算法

在获取了流线模拟结果后，笔者通过自行开发基于Ocean平台的插件实现了与Petrel RE的对接，能导出表示流线油藏流体、水相流体以及油相流体流动速率的属性数据，表示流线油相、水相饱和度分布的属性数据以及流线位置数据。

插件将Petrel RE流线数据导出至txt或excel文件，之后即可通过具有快速开发能力的Python编程语言进行后续数据处理及聚类分析。依据表1公式提取出流

线特性后，进行密度峰值聚类计算，当设定截断距离为0.01时样本平均局部密度为1.4053%。由于流线聚类结果对聚类数极为敏感，不同聚类数下聚类结果有较大差异(见图5)，因此有必要通过合理的方法选取聚类数。

密度峰值聚类算法通过计算流线样本局部密度与距离属性后依据二者的乘积进行排序，并选择lgγ出现跳跃处序号作为聚类数。图6中红色虚线处为跳跃较为明显位置，表示聚类数为14，但除该点外，样本序

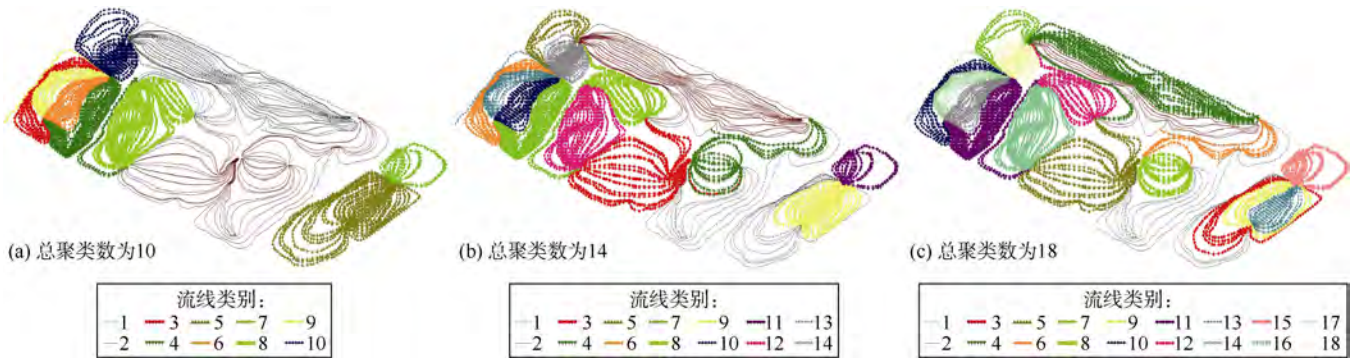


图5 不同聚类数下流线聚类结果

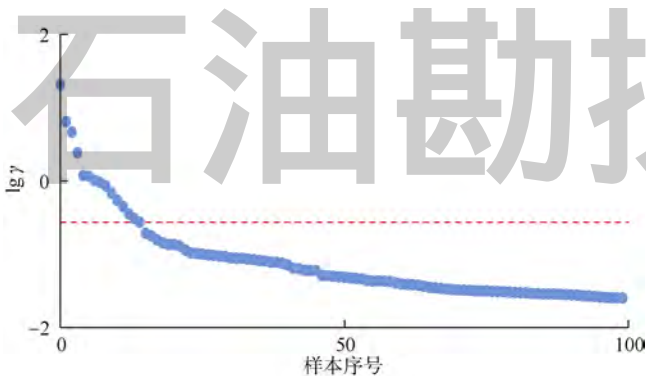


图6 不同样本序号下lgγ分布

号为4时同样出现非常明显的跳跃,因此需进一步通过聚类评价算法对聚类结果合理性进行评价。本文采用聚类评价能力较好的轮廓系数聚类评价算法计算不同聚类数下的轮廓系数得分,并以该得分对水驱开发6年后的流线聚类效果进行评价(见图7)。

图7为依据轮廓系数算法得出的流线聚类结果在不同聚类数下的得分,得分越接近1代表聚类效果越

好。其中红色圆点表示聚类数为14,可观察到该点处轮廓系数得分最高,因此合理聚类数为14。以水驱开发6年后油藏流线分布为例,在取聚类数均为14的情况下,对比密度峰值聚类算法与K-means、层次聚类、谱聚类算法间的聚类结果。由图8可见,4种聚类算法计算结果存在差异。

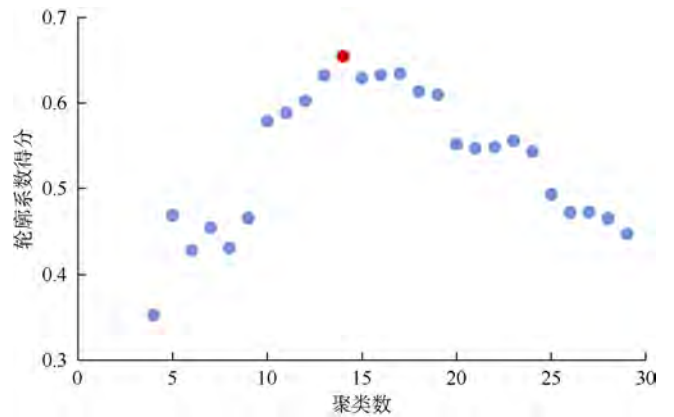


图7 不同聚类数时密度峰值聚类的轮廓系数得分

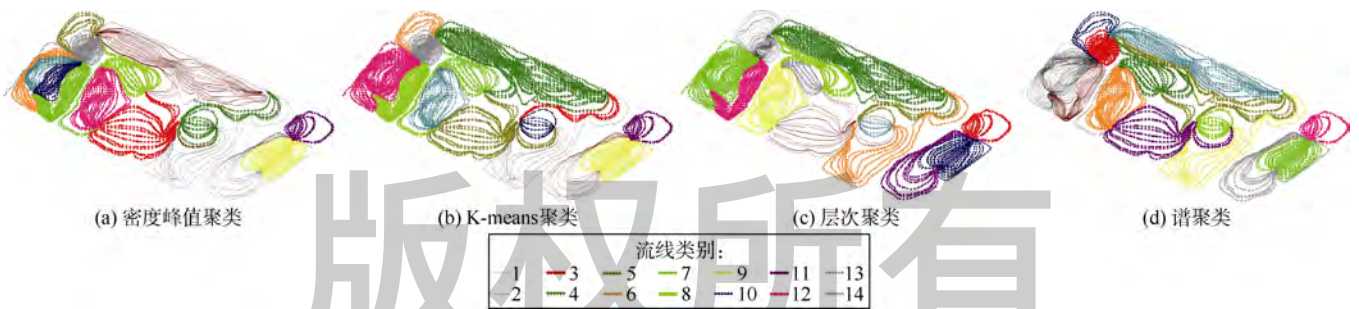


图8 不同聚类算法结果对比图

密度峰值、K-means、层次聚类、谱聚类4种聚类算法计算结果对应的轮廓系数分别为0.851 914,0.853 310,0.850 220,0.816 990。虽然K-means算法轮廓系数最高,但K-means算法计算结果具有随机性,不利于对流场进行确定性分析。其次为密度峰值聚类、层次聚类以及谱聚类。层次聚类算法轮廓系数略低于密度峰

值算法,但该算法受异常点影响较大,稳定性较差。此外,谱聚类算法的轮廓系数以及对流线的区分效果较其他算法均较低,因此本文选择密度峰值聚类算法作为流线聚类的主要算法。

2.2 聚类结果

油水流速比与油水体积比可表示流线方向水驱油

能力及流线方向未被注入水波及到的油相体积，因此本文以油水流速比与油水体积比的平均数作为水驱评价系数。该系数越大，则代表水相驱替效果越好，按该系数由大至小将流线聚类结果从 1 到 14 排序，结果见图 9。与 Petrel RE 流线模拟结果（见图 4）相比，本文的流线聚类结果对流场进行了有效区分，有利于找出具有开发潜力的区域，如 I3 至 P6 井间的第 2 类流线；还可识别出影响注水效率的无效注水循环区域，如 I4 至 P6 井间的第 14 类流线以及 I6 至 P5 井间的第 11 类流线；同时可对属于相同注采井间的流线进行细分，将其分为具有不同驱动能力的边流与主流：I4 至 P6 井间，第 5 类流线为边流，第 14 类为主流，I4 至 P4 井间，第 6、第 7 类为边流，第 10、第 13 类为主流。

这些现象进一步说明流线聚类对不同水相驱动能力流线具有较好区分效果。不同类别流线的流场性质见表 2。

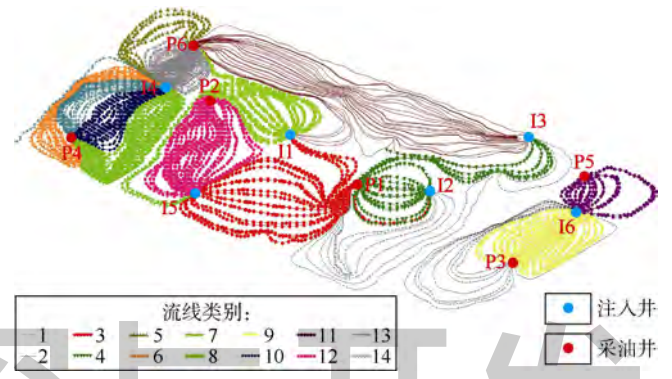


图 9 开发 6 年流线聚类图

表 2 不同流线类别的流场性质

流线类别	流体总流量/ ($m^3 \cdot d^{-1}$)	流线上油相平均 饱和度/%	油相总流量/ ($m^3 \cdot d^{-1}$)	流线上水相 平均饱和度/%	水相总流量/ ($m^3 \cdot d^{-1}$)	油水 体积比	油水 流速比	水驱评价 系数
1	298.48	69	244.44	31	54.03	1.75	3.32	2.535
2	1 040.86	64	792.62	36	248.24	1.78	2.92	2.350
3	318.71	62	226.33	38	92.39	1.37	1.91	1.640
4	121.09	54	80.65	46	40.44	1.08	1.77	1.425
5	143.22	49	86.48	51	56.74	0.87	1.43	1.150
6	857.28	38	395.95	62	461.32	0.60	0.82	0.710
7	308.17	39	132.76	61	175.41	0.65	0.75	0.700
8	382.06	34	167.29	66	214.77	0.30	0.40	0.350
9	370.40	14	72.62	86	297.78	0.14	0.20	0.170
10	44.53	22	10.00	78	34.53	0.17	0.14	0.155
11	101.91	17	19.72	83	82.19	0.12	0.07	0.095
12	558.54	16	75.72	84	482.82	0.10	0.05	0.075
13	277.29	7	30.05	93	247.24	0.06	0.07	0.065
14	440.85	2	12.10	98	428.74	0.01	0.01	0.010

由表 2 可见，不同类流线的油水流速比与油水体积比相差较明显，说明流线聚类算法对不同类别的流场具有很好的区分能力。同时，通过该流线聚类分布即可有针对性地对分析水相驱动效果较差区域，如第 10 至 14 类流线区域，这些区域往往位于注采井间连线处（见图 9），易发生水窜，影响注水波及系数。因此应选择合理方式调整注水制度，或是对该区域进行深部调剖，调整流体流向。此外，流线聚类方法可依据水驱评价系数识别出注采井间不同类型的流线，如 I4 井至 P4 及 P6 井区域。该区域因位于溶洞发育区域，因此流量较大，流线分布较密集，第 5、第 6、第 7、第 10、第 13、第 14 类流线均位于该区域，其中水相驱动能力较低的区域为第 10、第 13、第 14 类流线，其水驱评价系数分别为 0.155，0.065，0.010，说明该流线

区域内主要流通流体为水相，注采井间形成了优势通道并降低了注入水波及效率，而该区域内第 5、第 6、第 7 类流线水驱效果较好，水驱评价系数分别为 1.150，0.710，0.700，说明在同一注采井间存在水相驱动能力相差较大的流场分布。因此仅对该井进行注水优化无法进一步改善 I4 井的驱油效果，可考虑对 I4 井注入调剖剂，对具有较大水流量区域进行分流，增大边流处注水流量以提高波及效率。流线聚类方法同时也可用于寻找油藏内具有开发潜力的区域，如本次聚类结果中第 1、第 2、第 3 类流线，由于其注采井间无裂缝或溶洞将注采井直接相连，流体流动相对缓慢，因此水驱过程相对平缓，是具有较好开发潜力的区域。

采用密度峰值聚类算法对不同注水开发时刻流线

进行聚类的结果见图 10。

图 10 反映了流线聚类变化的历史过程, 对应时刻的流场聚类性质见表 3。表 3 中不同类别流线的油水流

动比及油水体积比表明, 流线聚类方法对不同时刻的流场均可进行较好的区分, 不同类流线间油水流速比与油水体积比差异较大, 区分效果均较好。

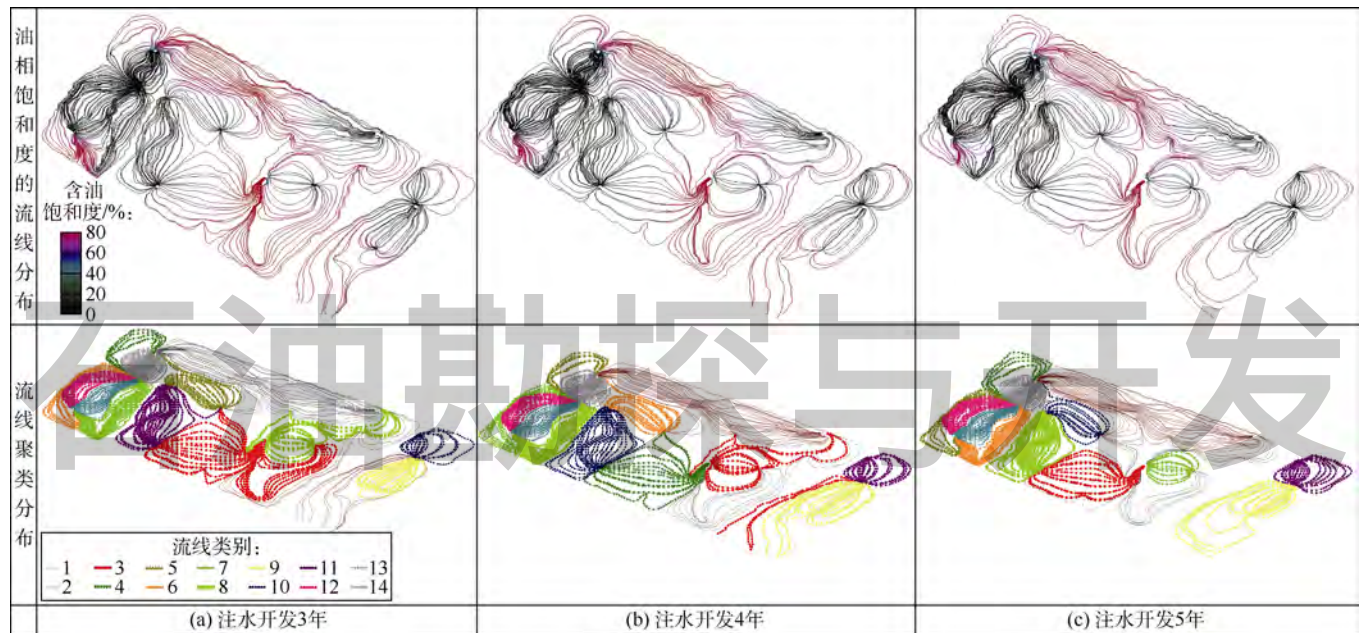


图 10 不同注水开发时刻流线聚类分布

表 3 不同注水时间下不同类别流线性质

流线类别	注水开发 3 年				注水开发 4 年				注水开发 5 年			
	油相总流量/ ($\text{m}^3 \cdot \text{d}^{-1}$)	水相总流量/ ($\text{m}^3 \cdot \text{d}^{-1}$)	油水 体积比	油水 流速比	油相总流量/ ($\text{m}^3 \cdot \text{d}^{-1}$)	水相总流量/ ($\text{m}^3 \cdot \text{d}^{-1}$)	油水 体积比	油水 流速比	油相总流量/ ($\text{m}^3 \cdot \text{d}^{-1}$)	水相总流量/ ($\text{m}^3 \cdot \text{d}^{-1}$)	油水 体积比	油水 流速比
1	138.82	27.15	1.88	3.50	114.35	10.07	2.89	8.19	146.70	29.18	1.92	3.77
2	682.93	226.78	1.66	2.51	684.87	204.64	1.68	3.05	685.48	217.88	1.56	2.75
3	200.62	87.94	0.92	1.29	63.85	66.79	1.30	2.32	190.99	95.00	0.95	1.40
4	59.34	44.35	0.60	1.05	209.39	89.60	1.12	1.72	56.04	41.84	0.56	0.96
5	84.81	114.47	0.39	0.62	73.90	113.40	0.78	1.61	77.57	110.96	0.41	0.60
6	108.89	275.33	0.20	0.30	50.56	39.09	0.86	1.45	101.33	297.98	0.21	0.32
7	30.97	56.86	0.13	0.09	91.31	122.76	0.47	0.69	33.71	61.50	0.14	0.09
8	39.84	438.98	0.11	0.05	122.28	281.93	0.28	0.41	40.27	461.25	0.12	0.06
9	47.07	103.84	0.10	0.07	63.20	121.52	0.11	0.06	44.04	112.82	0.10	0.07
10	20.67	129.44	0.11	0.05	56.50	465.01	0.11	0.05	21.31	118.35	0.12	0.05
11	38.19	113.11	0.10	0.04	33.10	117.00	0.10	0.05	38.85	113.63	0.10	0.04
12	9.19	431.70	0.03	0.02	10.18	386.34	0.03	0.03	8.91	434.43	0.03	0.02
13	12.17	371.81	0.01	0.01	16.22	404.91	0.01	0.01	13.21	402.57	0.01	0.01
14	15.20	573.75	0	0	14.94	533.84	0	0	14.33	526.19	0	0

3 结论

本文为水驱流场识别提供了一种新的研究方法, 以流线聚类的方式将不同类型流场进行区分, 同时依据密度峰值聚类算法及轮廓系数聚类评价算法确定合理聚类数, 确保聚类结果准确性。密度峰值聚类方法结果较稳定, 且对流场区分效果较好, 可作为流线聚类主要算法。

油水流速比及油水体积比可表征流场特征, 依据聚类结果可找出潜在优势流场分布, 有效识别油藏中无效注水循环通道以及具有开发潜力的区域, 在本文流线聚类结果中, 不同类流场间的油水流速比及油水体积比差异较大, 表明区分效果较好; 同时, 流线聚类方法可将同一注采井间的流线细分为主流及支流, 进一步为注水优化、井网层系调整、深部调剖等方案提供科学决策和技术支撑。流线聚类方法在不同注水

时刻均可对不同类型流场进行较好的区分，表明流线聚类算法对流场评价具有较好的适用性。

致谢：感谢斯伦贝谢公司在中国石油工程设计大赛期间为西南石油大学提供的 Petrel RE 软件及 Ocean 平台。

符号注释：

a ——流线； b ——位于流线上的点； d_c ——截断距离，无因次； d_{ij} ——样本 i, j 间的距离，无因次； d_{q_i, q_j} ——样本 q_i 至样本 q_j 的距离，无因次； i, j ——样本序号； q_i, q_j ——依据局部密度降序排列的第 i, j 个样本； n_c ——聚类数； n_a ——流线 a 上的总点数； n_p ——样本总量； S_{oab}, S_{wab} ——第 a 条流线上 b 点处油相、水相饱和度，%； V_{ow} ——油水体积比，无因次； v_{oab}, v_{wab} ——第 a 条流线上 b 点处油相、水相流动速率， m^3/d ； v_{ow} ——油水流速比，无因次； x_{ab}, y_{ab} ——第 a 条流线上 b 点处 x, y 轴坐标； X_a, Y_a —— a 流线中点位置 x, y 轴坐标； γ ——聚类中心评价系数，无因次； δ_q ——样本的分离距离，无因次； ρ ——样本密度，无因次。

参考文献：

- [1] 侯玉培. 整装油藏流场重整提高采收率的数值模拟研究[D]. 东营: 中国石油大学(华东), 2011.
HOU Yupei. Numerical simulation research of uncompartimentalized oilfields by flow field reforming to improve oil recovery[D]. Dongying: China University of Petroleum, 2011.
- [2] 王继强, 石成方, 纪淑红, 等. 特高含水期新型水驱特征曲线[J]. 石油勘探与开发, 2017, 44(6): 955-960.
WANG Jiqiang, SHI Chengfang, JI Shuhong, et al. New water drive characteristic curves at ultra-high water cut stage[J]. Petroleum Exploration and Development, 2017, 44(6): 955-960.
- [3] 赵辉, 康志江, 孙海涛, 等. 水驱开发多层油藏井间连通性反演模型[J]. 石油勘探与开发, 2016, 43(1): 99-106.
ZHAO Hui, KANG Zhijiang, SUN Haitao, et al. An interwell connectivity inversion model for waterflooded multilayer reservoirs[J]. Petroleum Exploration and Development, 2016, 43(1): 99-106.
- [4] 吕琦. 基于流线的油藏数值模拟研究[D]. 东营: 中国石油大学(华东), 2009.
LYU Qi. The research of numerical simulation of reservoir by streamline[D]. Dongying: China University of Petroleum, 2009.
- [5] 张乔良, 姜瑞忠, 姜平, 等. 油藏流场评价体系的建立及应用[J]. 大庆石油地质与开发, 2014, 33(3): 86-89.
ZHANG Qiaoliang, JIANG Ruizhong, JIANG Ping, et al. Establishment and application of oil reservoir flow-field evaluation system[J]. Petroleum Geology & Oilfield Development in Daqing, 2014, 33(3): 86-89.
- [6] ANWAR A R, MOHAMED B, MUSTAFA A, et al. An effective streamline history matching technique[R]. SPE 188073, 2017.
- [7] MILLIKEN W J, EMANUEL A S, CHAKRAVARTY A. Applications of 3D streamline simulation to assist history matching[J]. SPE Reservoir Evaluation & Engineering, 2001, 4(6): 502-508.
- [8] CHENG H, OYERINDE A S, DATTA-GUPTA A, et al. Compressible streamlines and three-phase history matching[R]. SPE 99465, 2007.
- [9] SINGH A P, MAUCEC M, KNABE S P. History matching using streamline trajectories[R]. SPE 172146, 2014.
- [10] BADDOURAH M A, ALTURKI A A, AL-ZAWAWI A S, et al. Streamlines guided assisted history matching methodology: A new approach[R]. SPE 178007, 2015.
- [11] THIELE M R, BATYCKY R. Water injection optimization using a streamline-based workflow[R]. SPE 84080, 2003.
- [12] SPYROU C E, ROSA A P L, KHATANIAN S K, et al. An approach to alternative waterflood designs and operations using streamline simulation: Application to an oil field in the North German Basin[R]. SPE 183879, 2017.
- [13] BUKHAMSEEN N Y, ALNAJEM A A, SAFFAR A H, et al. An injection optimization decision-making tool using streamline based fuzzy logic workflow[R]. SPE 186021, 2017.
- [14] BATYCKY R P, THIELE M R. Technology update: Mature flood surveillance using streamlines[J]. Journal of Petroleum Technology, 2016, 68(5): 22-25.
- [15] LI J, LEI Z, LI S, et al. Optimizing water flood performance to improve injector efficiency in fractured low-permeability reservoirs using streamline simulation[R]. SPE 182779, 2016.
- [16] MØYNER O, KROGSTAD S, LIE K A. The application of flow diagnostics for reservoir management[R]. SPE 171557, 2015.
- [17] MAQUI A F, ZHAI X, NEGREIRA A S, et al. A comprehensive workflow for near real time waterflood management and production optimization using reduced-physics and data-driven technologies[R]. SPE 185614, 2017.
- [18] AHONSI W, LAGER A, MUHAIRI L S A, et al. Tracer and streamline modelling for pattern validation and improved reservoir modelling[R]. SPE 181541, 2016.
- [19] RODRIGUEZ A, LAIO A. Clustering by fast search and find of density peaks[J]. Science, 2014, 344(6191): 1492-1496.
- [20] MATRINGE S F, GERRITSEN M G. On accurate tracing of streamlines[R]. SPE 89920, 2004.
- [21] RODRIGUEZ A, LAIO A. Supplementary materials for clustering by fast search and find of density[DB/OL]. [2017-12-20]. <http://science.sciencemag.org/content/suppl/2014/06/25/344.6191.1492.DC1>.
- [22] LIU Y, LI Z, XIONG H, et al. Understanding of internal clustering validation measures[R]. Sydney: IEEE, 2010.
- [23] ROUSSEEUW P. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis[J]. Journal of Computational and Applied Mathematics, 1987, 20(1): 53-65.

第一作者简介：贾虎（1983-），男，湖北武汉人，博士，西南石油大学副教授，主要从事油气井化学封堵、提高采收率与油藏工程方面的教学与科研工作。地址：四川省成都市新都区，西南石油大学油气藏地质与开发工程国家重点实验室，邮政编码：610500。E-mail: jiahuswpu@swpu.edu.cn
收稿日期：2017-08-21 修回日期：2018-01-03

（编辑 郭海莉）