

引用格式: 刘洪林, 周尚文, 李晓波. PCA-OPLS 联合法快速评价页岩气井储量动用程度[J]. 油气藏评价与开发, 2023, 13(4): 474-483.

LIU Honglin, ZHOU Shangwen, LI Xiaobo. Application of PCA plus OPLS method in rapid reserve production rate prediction of shale gas wells[J]. Petroleum Reservoir Evaluation and Development, 2023, 13(4): 474-483.

DOI: 10.13809/j.cnki.cn32-1825/te.2023.04.009

## PCA-OPLS 联合法快速评价页岩气井储量动用程度

刘洪林<sup>1,2,3</sup>, 周尚文<sup>1,2,3</sup>, 李晓波<sup>1,2,3</sup>

(1. 中国石油勘探开发研究院, 北京 100083; 2. 中国石油非常规油气重点实验室, 北京 100083;  
3. 国家能源页岩气研发(实验)中心, 河北 廊坊 065007)

**摘要:** 川南地区已经钻探完成了上千口页岩气井, 地质评价、钻探和生产过程中产生了大量高维度的数据, 利用生产获得的各种数据预测页岩气储量动用程度, 是指导页岩气勘探开发的一项重要工作。耦合主成分分析法(PCA)与正交偏最小二乘法(OPLS)分析方法是形成一种快速、准确的储量动用程度预测新方法。通过结合主成分分析和正交偏最小二乘法, 建立了PCA-OPLS联合法, 借助昭通页岩气井样本, 检验了新方法在储量动用程度评价中的应用效果。实例应用表明, 页岩气井储量动用预测准确率满足要求, 可以用于预测新钻井的储量动用程度, 从而证明PCA-OPLS联合法能达到快速、准确预测储量动用程度的目的。该方法操作简单, 准确率高, 在页岩气储量动用评价中具有较好的应用前景。

**关键词:** 储量动用程度预测; PCA-OPLS法; 多元非线性回归; 精度检验; 页岩气

中图分类号: TE122

文献标识码: A

### Application of PCA plus OPLS method in rapid reserve production rate prediction of shale gas wells

LIU Honglin<sup>1,2,3</sup>, ZHOU Shangwen<sup>1,2,3</sup>, LI Xiaobo<sup>1,2,3</sup>

(1. Research Institute Petroleum Exploration & Development, Petrochina, Beijing 100083, China; 2. PetroChina Key Laboratory for Unconventional Oil and Gas, Beijing 100083, China; 3. National Energy Shale Gas Research and Development(Experiment) Center, Langfang, Hebei 065007, China)

**Abstract:** In southern Sichuan, thousands of shale gas wells have been drilled, generating a vast amount of high-dimensional data during geological evaluation, drilling, and production processes. Predicting reserve recovery ratios through data exploration and analysis is essential for guiding the exploration and development of shale gas resources. To achieve this goal, a novel approach is introduced, which couples principal component analysis(PCA) and orthogonal partial least square(OPLS) analysis, enabling rapid and accurate prediction of reserve production degree. The new method is put to the test using Zhaotong shale gas well samples to evaluate its effectiveness in predicting reserve recovery ratios. The results show that the average accuracy of reserve recovery ratio prediction using PCA-OPLS method surpasses the anticipated result, that this algorithm can swiftly and precisely predict recovery ratios. With its advantages of simplicity, high accuracy, and promising application prospects, this method holds great potential for efficiently evaluating the production and recovery ratios of shale gas reserves.

**Keywords:** reserve recovery ratio prediction; PCA-OPLS method; multiple nonlinear regression; accuracy test; shale gas

在中国四川盆地威远、长宁、涪陵等地区已经实现了页岩气资源的规模开发, 随着勘探开发工作的深入, 确定可动用储量逐渐成为页岩气开发评价的重要工作, 也是进一步调整开发方案的重要基础<sup>[1-3]</sup>。

与常规油气藏类似, 页岩气井的储量动用程度为某一生产时间动用的地质储量与已经投入开发地质储量的比值, 单井可采储量可以通过数值模拟法、生产动态分析法等方法获得。页岩气井的单井控制储量

收稿日期: 2022-04-08。

第一作者简介: 刘洪林(1973—), 男, 硕士, 高级工程师, 主要从事非常规油气勘探开发工作。地址: 北京市海淀区学院路20号中国石油勘探开发研究院, 邮政编码: 100083。E-mail: liuhonglin69@petrochina.com.cn

基金项目: 中国石油前瞻性基础性研究项目“煤层气新区新层系新领域战略与评价技术研究”(2121DJ2303); 中国石油科学与技术开发项目“页岩气提高储量动用技术研究”(kt2021-11-02-01)。

可以通过单井控制面积和资源丰度进行计算,该方法属于页岩气储量动用的直接算法,由于页岩储层的储量动用受储层物性、含水饱和度、动用范围、地层压力等多因素影响,页岩气的产出与储层地质参数、工程参数等紧密相关,因此,也可以通过分析各参数与单井可采储量与单井控制储量的关系,利用多元回归的数学方法进行页岩气储量动用的预测。正交偏最小二乘法(OPLS)是一种新型的多元统计数据分析方法,它由TRYGG Johan等人于2002年首次提出,是一种多因变量对多自变量的回归建模方法,其最大特点是可以去除自变量 $X$ 中与分类变量 $Y$ 无关的数据变异,使分类信息主要集中在一个主成分中,模型变得简单和易于解释,其模型预测效果优于PLS(偏最小二乘法),该方法后来被推广到微生物基因分析领域<sup>[4-10]</sup>。国内有学者将主成分分析法(PCA)与多元非线性回归法结合应用在储层敏感性和煤层含气量预测方面,并取得了良好的效果<sup>[4,11]</sup>;在孔隙度预测方面,优选基于地震数据中与孔隙度相关性较大的5种地震属性,建立回归方程,预测孔隙度的精度相对较高,可以得到目标区的孔隙度分布情况<sup>[12]</sup>。也有学者采用Boosting-PLS(根据错误率采样的偏最小二乘法)进行页岩总含气量的预测,提取最能反映含气性的密度、中子、铀、钍等测井信息,通过模型自动学习训练样本中输出曲线响应值与总含气量的响应关系,利用所得响应关系建立总含气量预测模型<sup>[13]</sup>。在农业领域,采用PCA和PLS用于大豆品质评价<sup>[14]</sup>。目前也有学者在研究页岩气产能主控因素中使用PLS提取关键因素<sup>[15]</sup>,但是由于页岩气地质、工程和生产要素之间具有较高的相关性,如含气量与总有机碳(TOC)之间、加砂量和压裂液量之间等相关性较强,因此在因素分析中容易导致参数重复赋权。在国内开展的储量动用研究工作中,多为常规油气或致密气,未见在页岩气储量动用评价中的应用<sup>[16-21]</sup>,通过将PCA方法中降维优势和OPLS多变量预测优势结合起来尝试用于页岩气储量动用程度分析,可以为页岩气储量动用程度评价提供一种新方法,提高储量动用程度分析速度和预测效果。

## 1 PCA-OPLS联合法的建立

### 1.1 主成分分析法(PCA法)

PCA法是一种数据降维的数学方法,该方法把

数据投射到样本散布最开的正交向量上,使投射平均均方误差尽可能小。该方法将数据进行特征分解,构造新变量(称为主成分),并保证各变量之间正交。方差越大主成分含原变量信息量越大。提取较少几个主成分就可以包含原数据的信息,将高维数据降到低维。PCA法处理不但能减少每组样本的特征个数,而且确保这些特征相互独立,使样本规律更清晰,便于回归模型的挑选,并加快算法速度。PCA法的降维步骤为:

1) 建立原始样本矩阵 $X$ ,样本数量要求不小于自变量数量。

2) 对原始数据进行标准化处理。

$$x_{ij} = \frac{x_{ij} - \bar{x}_j}{\sqrt{\text{Var}(x_j)}} \quad (i=1, 2, \dots, n; j=1, 2, \dots, p) \quad (1)$$

式中: $\bar{x}_j$ 为矩阵 $X$ 矩阵中每列数据的平均值; $\text{Var}(x_j)$ 中每列数据的方差。

3) 计算样本相关系数矩阵 $R$ 。

4) 计算相关系数矩阵的特征值和相应的特征向量,得到特征值 $\lambda=(\lambda_1, \lambda_2, \dots, \lambda_p)$ ,特征向量 $a_i=(a_{i1}, a_{i2}, \dots, a_{ip})$ 。

5) 根据影响因素累积贡献大小选择前 $k$ 个因素,计算累积贡献率的公式为

$$\eta = \frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^p \lambda_i} \quad (2)$$

式中: $k$ 为特征向量的数量; $\lambda_i$ 为矩阵 $R$ 的特征值; $\eta$ 为累积贡献率。

通常取 $\lambda > 1$ 的前 $k$ 个特征向量,认为前 $k$ 个影响因素已包含数据的原始信息。前 $k$ 个特征向量组成的矩阵 $U$ ,即主成分得分系数矩阵。

6) 利用式(3),将 $p$ 个影响因素降维成 $k$ 个影响因素,即:

$$X' = X_0 \cdot U \quad (3)$$

式中: $X'$ 为PCA法处理后的样本矩阵。

### 1.2 正交最小偏二乘分析方法(OPLS法)

多元回归分析是对2个或2个以上具有相关关系的变量变化的一般关系进行分析,它主要是研究自变量 $X$ 和因变量 $Y$ 之间的相互关系,建立变量间经验公式,以便达到预测或控制的目的。多元非线性回归就是寻找能表示特征值 $x_1, x_2, \dots, x_n$ 和预测值 $y_i$

之间的非线性函数关系式:

$$y_i=f(x_1, x_2, x_3, \dots, x_n) \quad (4)$$

式中: $y_i$ 为预测值; $x_1, x_2, \dots, x_n$ 为特征值。

通过最小二乘法可求得模型中相应的参数。在变量系统中提取若干对系统具有最佳解释能力的新综合变量,并求出当前主成分数量时总的模型预测指数 $Q_2$ 和变量投影重要性值(VIP),若 $0.5 < Q_2 < 0.8$ 模型预测能力一般,若 $Q_2 > 0.8$ ,模型具有较好预测能力。同时根据VIP值大小进行排序,若参数VIP值 $> 1$ ,参数具有重要性。变量投影重要性指标VIP值用以评价自变量对因变量的影响能力,其定义式为:

$$VIP = \sqrt{\frac{p}{R_d(Y; t_1, t_2, \dots, t_r)} \sum_{k=1}^r R_d(Y; t_k) w_{kj}^2} \quad (5)$$

式中: $R_d(Y; t_k)$ 和 $R_d(Y; t_1, t_2, \dots, t_r)$ 是单个主成分 $t_k$ 对产量 $Y$ 的影响能力和所有主成分 $t_1, \dots, t_r$ 对 $Y$ 的累计解释能力, $p$ 为自变量数量。自变量 $x_j$ 在解释因变量 $Y$ 的重要性,可以用变量投影重要性指标 $VIP_j$ 来测度。 $VIP_j$ 值越大,说明 $x_j$ 对 $Y$ 的影响能力就越大,即各影响因素对产能的贡献程度越大。

### 1.3 PCA-OPLS联合法实现流程

PCA-OPLS联合法实现流程主要有以下3个步骤:①页岩气样本数据集采集:包括各种地质、储层、压裂和生产参数,并计算储量动用程度。②样本实验集和验证集划分:对实验集样本进行PCA降维处理,降维后得到的特征彼此之间相互独立,不存在线性相关的主成分,重新命名主成分为新变量指标,便于探索新变量与预测值之间的数学关系,用最小二乘法对实验集样本数据综合分析,建立新变量与预测值之间的数学模型。③模型精度检验:用验证集样本对模型进行精度检验,形成预测模型并应用于新区的预测。

## 2 实验分析及应用

### 2.1 自变量和因变量的建立

以昭通国家级页岩气示范区为例,开展PCA-OPLS联合法的实验分析。昭通国家级页岩气示范区主体处于盆外复杂构造残留拗陷区(图1),浙江油田分公司在2a内实现了集中评价与产能试验,探明了千亿立方米级储量的大型浅层页岩气田。目的层龙马溪组埋深500~1500m,五峰组一龙一,

亚段优质页岩发育厚度3~5m、有机质丰度TOC介于2.58%~3.21%、 $R_o$ 介于1.99%~3.08%、含气性介于3.30~5.51m<sup>3</sup>/t、孔隙度介于3.98%~5.41%、脆性矿物含量介于51%~75%、地层压力系数介于1.25~1.62<sup>[22-25]</sup>。目前已经有27口井投入生产,单井平均测试产量7.03×10<sup>4</sup>m<sup>3</sup>/d,取得了较好的效果,采用单井可采储量与控制储量比值法计算储量采出程度平均为9.2%~39.6%。

通过收集昭通地区页岩气井的脆性矿物含量、储层厚度、总有机碳、孔隙度等20项自变量参数组成矩阵 $X$ ,通过单井可采储量和控制储量计算获得单井储量动用程度,定义为因变量 $Y$ ,详见表1。根据样本数量不小于自变量数量的要求,采用其中20个样本作为实验样本,即采用最小样本数量进行实验,目的在于探索最小样本情况下的方法精度,其余7个样本作为验证样本。然后利用公式1进行变量的标准化,使得自变量 $X$ 中具有不同数量级参数的方差都能得到有效表达。

### 2.2 PCA法降维

在采用OPLS进行建模之前,先将标准化处理后的实验样本进行主成分分析,以降低样本数据的维度、简化识别模型。主成分分析的结果见表2,前5个主成分的方差贡献率分别为30.03%、20.10%、14.00%、10.30%和7.15%。由此得知,前5个主成分的累计贡献率已达81.58%,能够很好地表征原样本的信息。

### 2.3 主成分矩阵和主成分得分系数矩阵 $U$

由PCA分析获得的成分矩阵表3可以看出:主成分PCA1主要与资源丰度、优质储层厚度等反映资源品质的指标相关,定义为资源品质指标;主成分PCA2主要与累计产气、IP30等反映气井产量的指标相关,定义为产量指标;主成分PCA3主要与主压裂液量、加砂量等反映了改造程度的指标有关,定义为改造程度指标;主成分PCA4主要与吸附气含量、含气饱和度等反映储层质量指标相关,定义为储层质量指标;主成分PCA5主要与压裂排量参数相关,定义为压裂施工指标。同时,通过PCA分析获得5个主成分的得分系数矩阵见表4,得分系数矩阵是计算主成分的基础数据。

以页岩气井实验样本的主成分PCA1为例说明主成分计算方法。

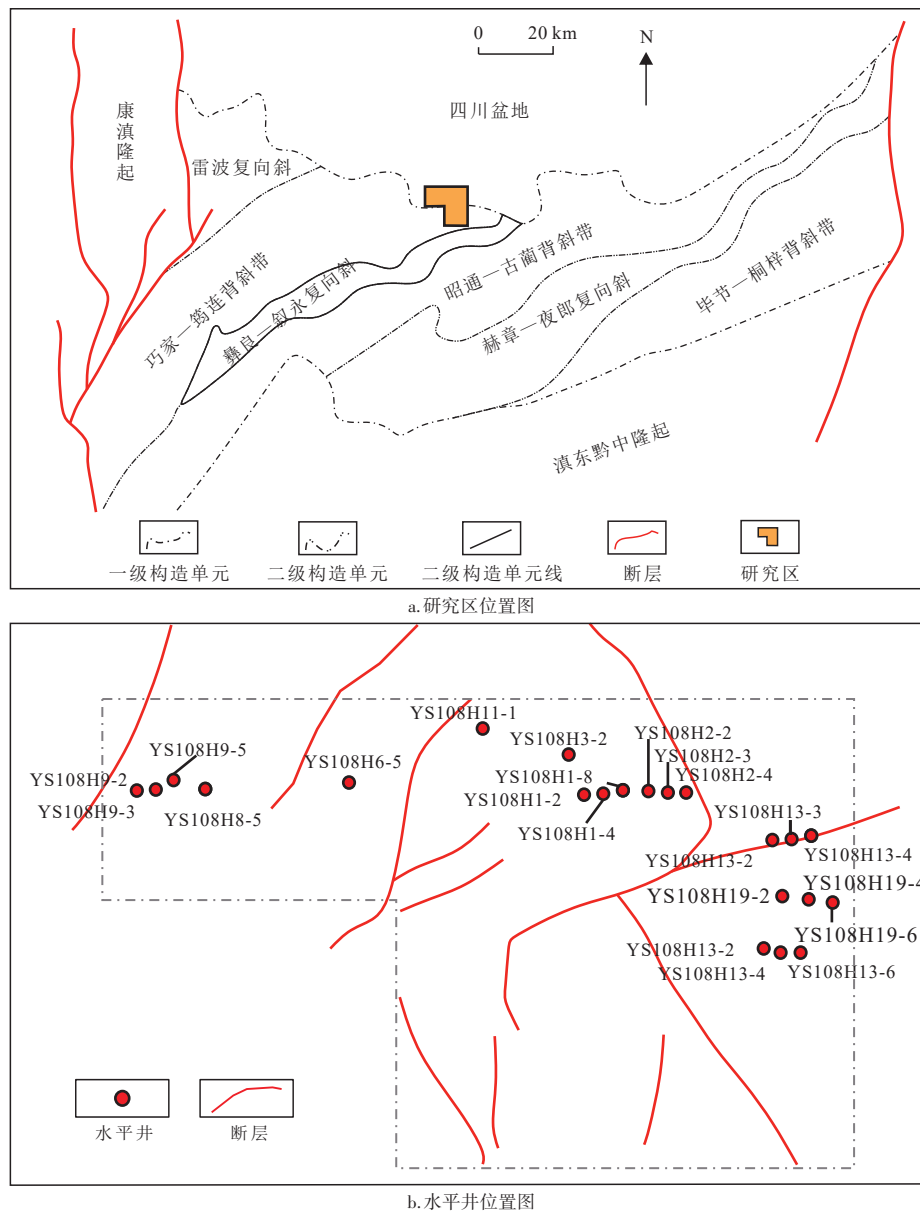


图1 昭通浅层页岩气田的位置

Fig. 1 Location of shallow shale gas field of Zhaotong

$PCA1=0.837 \times \text{优质储层厚度} + 0.795 \times \text{资源丰度} - 0.738 \times \text{井口压力} - 0.663 \times \text{总有机碳} + 0.149 \times \text{累计产气} + 0.141 \times IP30 + 0.463 \times \text{孔隙度} + 0.459 \times \text{含气量} - 0.012 \times \text{主压裂用量} - 0.026 \times \text{改造段数} - 0.107 \times \text{加砂量} - 0.097 \times \text{吸附气含量} - 0.168 \times \text{含气饱和度} + 0.141 \times \text{水平段长} + 0.637 \times \text{单井储量} - 0.320 \times \text{脆性矿物含量} - 0.156 \times \text{最大排量} - 0.042 \times \text{最小排量} - 0.133 \times \text{井深}$

其他主成分PCA2、PCA3、PCA4、PCA5计算方法与其相同,最终得到PCA降维后的矩阵,见表5,包含5个自变量的矩阵X'。

#### 2.4 OPLS法分析

将表2中20个样品资源品质指标、产量指标、改造程度指标、储层质量指标、压裂施工指标等作为自变量,计算的储量动用程度作为因变量,通过SIMICA13.0软件建立计算储量动用程度与自变量的校正OPLS模型,进行偏最小二乘法分析,求出当前5个主成分时的模型拟合精度指数 $R^2$ 为0.8912和模型预测指数 $Q_2$ 值为0.8693,说明模型具有较好预测能力,最后使用7个验证样本数据进行精度检验,若

表1 页岩气井实验样本的主要参数  
Table 1 Key parameters of testing samples from shale gas wells

井号	脆性 矿物/ %	储层 厚度/ m	总有机碳/ %	孔隙度/ %	总含 气量/ ( $\text{m}^3 \cdot \text{t}^{-1}$ )	含气饱 和度/ %	吸附 含气量/ ( $\text{m}^3 \cdot \text{t}^{-1}$ )	资源丰度/ ( $10^8 \text{ m}^3 \cdot \text{km}^{-2}$ )	单井 储量/ $10^4 \text{ m}^3$	水平 段长/ m	压裂 液量/ $\text{m}^3$	加砂量/ t	最小 施工排量/ ( $\text{m}^3 \cdot \text{min}^{-1}$ )	最大 施工排量/ ( $\text{m}^3 \cdot \text{min}^{-1}$ )	井口 压力/ MPa	IP30/ $10^4 \text{ m}^3$	单井 可采储量/ $10^4 \text{ m}^3$	测试 日产量/ $10^4 \text{ m}^3$	生产 天数/ d	储量 动用程度/ %
YS108H1-2	62	72	2.3	6.3	5.5	54.3	3.1	9.0	4.5	1408	34569	1381	3.5	10.7	31.7	13.31	0.67	18.60	1976	14.9
YS108H1-4	63	71	2.2	5.9	5.3	52.1	2.5	9.2	4.5	1404	36982	1315	8.0	11.3	29.7	10.99	1.54	24.23	2147	34.2
YS108H1-8	58	70	2.5	6.1	5.6	51.4	3.1	8.8	4.6	1478	36563	1288	9.3	10.8	15.8	8.92	1.08	26.08	2147	23.5
YS108H2-2	58	69	2.4	6.3	5.6	56.0	2.5	8.9	4.5	1433	38322	2096	12.2	14.2	23.6	9.09	0.81	14.89	1695	18.0
YS108H2-3	58	70	2.7	6.0	5.5	52.5	3.5	8.8	6.0	1900	18759	1055	12.8	13.5	27.3	8.12	0.55	35.14	1290	9.2
YS108H2-4	56	70	2.4	5.9	5.8	53.4	2.9	8.8	5.1	1630	41819	2306	13.3	14.0	16.5	15.61	1.26	29.03	1693	24.7
YS108H2-5	55	70	2.8	5.9	5.5	52.0	3.4	8.9	3.6	1154	29240	1443	9.8	14.2	21.6	8.04	0.89	24.75	1701	24.7
YS108H3-2	56	72	2.6	5.9	5.5	50.5	3.3	9.0	4.8	1491	37770	1576	10.2	12.7	20.9	15.85	1.10	17.50	2280	22.9
YS108H6-5	58	62	2.8	5.2	4.5	48.5	3.2	6.7	3.7	1515	34376	987	6.0	11.0	15.7	8.28	0.87	15.79	2257	23.5
YS108H8-5	54	62	2.7	4.6	5.0	46.0	3.2	6.8	4.0	1665	32217	1893	10.0	13.2	11.6	2.12	0.51	10.87	1871	12.8
YS108H9-2	56	60	2.6	4.6	4.7	44.3	3.0	6.8	3.8	1570	36756	1630	10.3	12.1	11.1	4.13	0.47	12.23	1533	12.4
YS108H9-3	56	61	2.7	4.5	4.9	45.5	2.6	6.7	4.0	1650	44452	2113	10.0	12.3	12.9	4.08	0.83	11.25	1540	20.8
YS108H9-5	56	61	2.7	4.7	4.9	44.7	3.1	6.7	3.3	1350	31031	1442	10.2	12.1	13.7	3.72	0.33	13.19	1543	10.0
YS108H11-1	49	69	2.3	4.9	4.9	52.0	3.2	8.0	3.9	1383	26093	1209	9.3	11.4	13.0	9.16	1.09	14.00	2183	27.9
YS108H13-2	54	71	2.6	5.7	4.9	53.8	2.8	8.4	3.3	1100	37900	2250	11.4	13.0	21.0	9.74	0.82	17.70	1065	24.8
YS108H13-3	54	71	2.7	5.7	5.2	53.8	2.5	8.3	3.4	1140	40497	2387	11.1	12.7	22.0	11.68	1.00	16.64	1035	29.4
YS108H13-4	55	53	2.5	5.1	4.2	54.4	2.8	5.4	2.5	1330	46500	2750	11.3	12.2	19.1	8.02	0.99	32.50	1035	39.6
YS108H19-2	55	54	3.0	5.3	4.6	55.3	2.5	5.4	3.0	1560	42618	1879	11.2	13.3	29.2	6.27	0.36	22.90	1432	12.0
YS108H19-4	68	53	2.7	5.4	4.2	54.5	3.3	5.4	2.5	1340	37693	1725	12.1	14.1	29.5	5.97	0.38	27.40	1426	15.2
YS108H19-6	67	53	2.8	5.1	4.4	56.4	2.2	5.4	1.8	970	25031	1079	11.8	13.1	14.4	6.95	0.31	37.30	1426	17.2

$R^2 < 0.80$ , 可以通过重新选择实验样本数据构建新的OPLS预测模型, 通过计算获得所有样本数据预测结果绘制交汇图和参数重要性VIP图(图2)。由图2可知储量动用程度预测值与计算值相关系数 $r$ 达到0.851 1, 具有较高的预测精度, 预测误差越小, 越能够满足预测要求。

OPLS法计算得到各参数的VIP值并进行排序, 按照排序最终判断影响页岩气储量动用程度的主控因素大小, 结果是改造程度指标>资源品质指标>储层质量指标>产量指标>压裂施工指标, 其中改造程

度指标、资源品质指标的VIP值大于1.0, 为模型的重要影响参数。

为了比较方法间的准确性, 利用SPSS软件采用神经网络、多元线性回归方法进行拟合, 对比PCA-OPLS联合法、神经网络法、多元线性回归方法之间的精度见表6, 可知PCA-OPLS联合法平均误差1.44%, 神经网络法平均误差为4.86%, 直接回归法平均误差为6.13%。与神经网络、多元线性回归法相比, PCA-OPLS联合法可大幅度提高预测值的准确率。

### 3 新方法验证应用

为进一步验证PCA-OPLS联合法的可靠性和可推广性, 采用威远地区的页岩气井数据作为新验证样品进行进一步分析, 威远地区钻井较多, 其中威202井区和威204井区已经投产490口井。按照上述应用流程, 先采用PCA对数据进行降维, 然后建立OPLS预测模型分析方法, 获得预测拟合与计算储量动用率相关性图(图3), 结果表明威远地区预测效果仍然较好, 相关系数分别为0.923 4和0.864 2。大样本数据分析结果明显好于昭通地区的小样本数据预测精度, 说明更多的样本数据对于模型的迭代和分析能够提供更多的信息, 避免个别异常样本的干扰, 提高模型的预测精度。

表2 页岩气井实验样本的总方差解释

Table 2 Interpretation of total variance of testing samples from shale gas wells

成分	初始特征值			提取载荷平方和		
	总计	方差百分比	% 累积	总计	方差百分比	% 累积
1	6.00	30.03	30.03	6.00	30.03	30.03
2	4.02	20.10	50.16	4.02	20.10	50.13
3	2.81	14.00	64.25	2.81	14.00	64.13
4	2.07	10.30	74.64	2.07	10.30	74.43
5	1.43	7.15	81.79	1.43	7.15	81.58
6	0.93	4.67	86.47			
7	0.68	3.42	89.89			
⋮	⋮	⋮	⋮			
15	0.02	0.09	99.93			

表3 页岩气井实验样本的成分矩阵

Table 3 Composition matrix of testing samples from shale gas wells

标准化的X变量	主成分				
	PCA1	PCA 2	PCA 3	PCA 4	PCA 5
Z-score(优质储层厚度)	0.837	0.080	-0.101	0.308	-0.220
Z-score(资源丰度)	0.795	0.382	0.019	0.370	-0.108
Z-score(井口压力)	0.738	0.075	-0.155	-0.108	0.049
Z-score(总有机碳)	-0.663	-0.234	-0.039	-0.144	0.242
Z-score(累计产气)	0.149	0.902	-0.016	-0.119	-0.018
Z-score(IP30)	0.141	0.901	-0.012	-0.111	-0.012
Z-score(孔隙度)	0.463	0.752	0.152	-0.009	0.133
Z-score(含气量)	0.459	0.636	0.188	0.400	0.103
Z-score(主压裂用液量)	-0.012	0.022	0.966	0.074	-0.008
Z-score(改造段数)	-0.026	0.000	0.962	0.131	-0.057
Z-score(加砂量)	-0.107	0.126	0.916	0.012	0.257
Z-score(吸附气含量)	-0.097	0.460	-0.134	0.722	-0.062
Z-score(含气饱和度)	-0.168	0.291	0.117	-0.686	0.311
Z-score(水平段长)	0.141	-0.206	0.245	0.680	0.150
Z-score(单井储量)	0.637	0.134	0.121	0.670	0.016
Z-score(脆性矿物含量)	-0.320	0.025	-0.321	-0.566	-0.062
Z-score(最大排量)	-0.156	0.145	0.030	0.020	0.889
Z-score(最小排量)	0.042	-0.184	-0.015	-0.024	0.841
Z-score(井深)	-0.133	0.514	0.128	0.124	0.577

表4 页岩气井实验样本的主成分得分系数矩阵  
Table 4 Principal component score coefficient matrix of testing samples from shale gas wells

标准化的X变量	成分得分系数				
	资源品质指标	产量指标	改造程度指标	储层质量指标	压裂施工指标
Z-score(优质储层厚度)	0.837	0.080	-0.101	0.308	-0.220
Z-score(资源丰度)	0.795	0.382	0.019	0.370	-0.108
Z-score(井口压力)	0.738	0.075	-0.155	-0.108	0.049
Z-score(总有机碳)	-0.663	-0.234	-0.039	-0.144	0.242
Z-score(累计产气)	0.149	0.902	-0.016	-0.119	-0.018
Z-score(IP30)	0.149	0.902	-0.016	-0.119	-0.018
Z-score(孔隙度)	0.463	0.752	0.152	-0.009	0.133
Z-score(含气量)	0.459	0.636	0.188	0.400	0.103
Z-score(主压裂用液量)	-0.012	0.022	0.966	0.074	-0.008
Z-score(改造段数)	-0.026	0.000	0.962	0.131	-0.057
Z-score(加砂量)	-0.107	0.126	0.916	0.012	0.257
Z-score(吸附气含量)	-0.097	0.460	-0.134	0.722	-0.062
Z-score(含气饱和度)	-0.168	0.291	0.117	-0.686	0.311
Z-score(水平段长)	0.141	-0.206	0.245	0.680	0.150
Z-score(单井储量)	0.637	0.134	0.121	0.670	0.016
Z-score(脆性矿物含量)	-0.320	0.025	-0.321	-0.566	-0.062
Z-score(最大排量)	-0.156	0.145	0.030	0.020	0.889
Z-score(最小排量)	0.042	-0.184	-0.015	-0.024	0.841
Z-score(井深)	-0.133	0.514	0.128	0.124	0.577

表5 页岩气井实验样本的PCA处理后的矩阵  
Table 5 PCA processed matrix of testing samples from shale gas wells

井号	自变量X'					因变量Y
	资源品质指标	产量指标	改造程度指标	储层质量指标	压裂施工指标	储量动用程度/%
YS108H1-2	38 745.19	242 578.90	29 169.37	-26 625.29	38 684.12	14.9
YS108H1-4	31 815.66	200 651.80	32 300.94	-21 119.02	31 753.22	34.2
YS108H1-8	25 669.95	163 450.99	32 491.84	-16 325.98	25 599.78	23.5
YS108H2-2	25 977.16	166 748.52	33 369.81	-16 560.60	25 913.70	18.0
YS108H2-3	23 559.73	148 623.42	15 293.25	-15 499.28	23 498.18	9.2
YS108H2-4	45 494.63	283 912.81	34 990.12	-31 269.70	45 425.26	24.7
YS108H2-5	23 072.32	147 564.03	25 219.65	-15 067.94	23 006.75	24.7
YS108H3-2	46 243.02	288 485.75	31 464.01	-32 137.25	46 174.13	22.9
YS108H6-5	23 799.00	151 848.49	30 773.49	-14 974.36	23 739.38	23.5
YS108H8-5	5 442.66	40 631.93	29 577.75	-909.97	5 380.57	12.8
YS108H9-2	11 374.61	76 894.37	33 755.54	-5 246.87	11 313.41	12.4
YS108H9-3	11 123.02	76 088.64	41 027.27	-4 541.59	11 062.52	20.8
YS108H9-5	10 214.87	69 443.30	28 294.55	-4 880.08	10 155.26	10.0
YS108H11-1	26 495.12	167 691.19	22 035.12	-17 699.01	26 425.29	27.9
YS108H13-2	27 887.31	178 252.07	32 591.70	-18 339.07	27 820.61	24.8
YS108H13-3	33 636.45	213 393.80	34 506.93	-22 589.42	33 570.31	29.4
YS108H13-4	22 637.13	147 494.82	41 224.96	-13 616.19	22 590.07	39.6
YS108H19-2	17 589.50	115 829.81	38 711.51	-9 716.63	17 549.24	12.0
YS108H19-4	16 724.85	110 252.57	33 984.25	-9 539.17	16 685.64	15.2
YS108H19-6	19 808.92	127936.21	21 583.31	-12 956.80	19 758.71	17.2

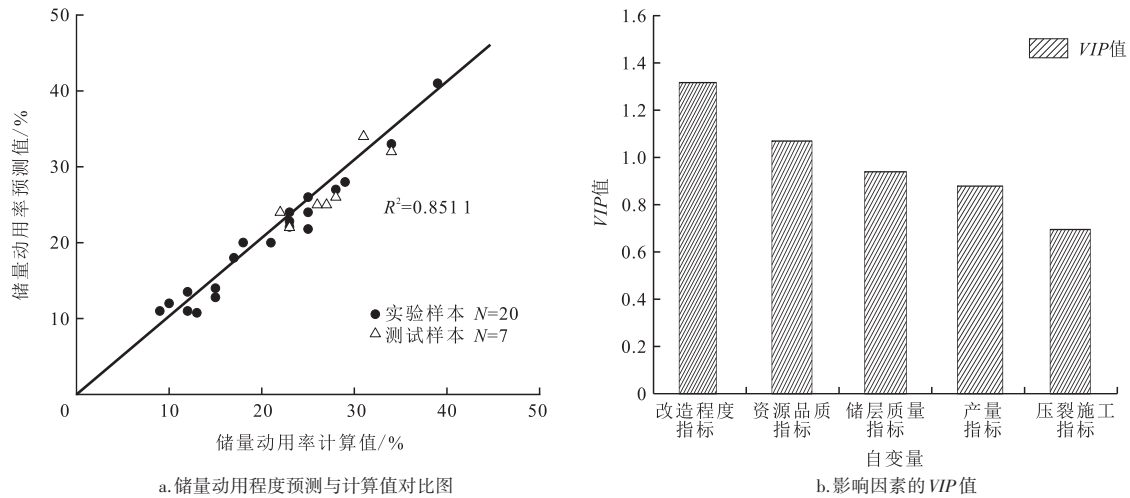


图2 储量动用程度预测结果及影响因素VIP值分布

Fig. 2 Prediction results and VIP value distribution of reserve recovery ratio

表6 不同方法预测结果与计算结果比较

Table 6 Comparison of prediction results and calculation results of different methods

井号	实际储量 动用程度/%	预测值/%			预测误差/%		
		PCA-PLS联合法	神经网络法	多元线性回归法	PCA-PLS法	神经网络法	多元线性回归法
YS108H1-2	14.9	14.0	24.0	18.7	0.9	9.1	3.8
YS108H1-4	34.2	33.0	25.0	19.2	1.2	9.2	15.0
YS108H1-8	23.5	22.8	21.9	19.0	0.7	1.6	4.5
YS108H2-2	18.0	20.0	26.4	21.3	2.0	8.4	3.3
YS108H2-3	9.2	11.0	4.7	18.7	1.8	4.5	9.5
YS108H2-4	24.7	26.0	26.2	23.3	1.3	1.5	1.4
YS108H2-5	24.7	24.1	23.8	17.9	0.6	0.9	6.8
YS108H3-2	22.9	24.0	26.6	19.4	1.1	3.7	3.5
YS108H6-5	23.5	22.1	17.3	18.2	1.4	6.2	5.3
YS108H8-5	12.8	10.8	17.8	18.3	2.0	5.0	5.5
YS108H9-2	12.4	11.0	14.0	20.4	1.4	1.6	8.0
YS108H9-3	20.8	20.0	18.2	25.7	0.8	2.6	4.9
YS108H9-5	10.0	12.0	15.6	18.4	2.0	5.6	8.4
YS108H11-1	27.9	27.0	20.8	17.7	0.9	7.1	10.2
YS108H13-2	24.8	21.8	26.6	23.4	3.0	1.8	1.4
YS108H13-3	29.4	28.0	27.4	25.5	1.4	2.0	3.9
YS108H13-4	39.6	41.2	26.3	31.4	1.6	13.3	8.2
YS108H19-2	12.0	13.5	18.5	24.5	1.5	6.5	12.5
YS108H19-4	15.2	12.8	20.4	21.5	2.4	5.2	6.3
YS108H19-6	17.2	18.0	18.6	17.6	0.8	1.4	0.4
平均值					1.44	4.86	6.13

## 4 结论与认识

1) 通过对样本进行PCA法降维处理,再用OPLS法建立计算值与预测值之间的回归拟合数学模型,建立了PCA-OPLS联合法预测页岩气储量动用程度的新方法。

2) PCA-OPLS联合法的优点在于利用PCA降维法保证样本信息损失最小减少的前提下,大大减少样本矩阵维度和计算工作量,同时发挥OPLS模型多变量预测优势,确定影响储量动用程度的核心参数。

3) 通过实例应用表明PCA-OPLS联合法在昭通、威远页岩气井储量动用预测准确率满足要求,可以用于预测新钻井的储量动用程度。威远地区大样

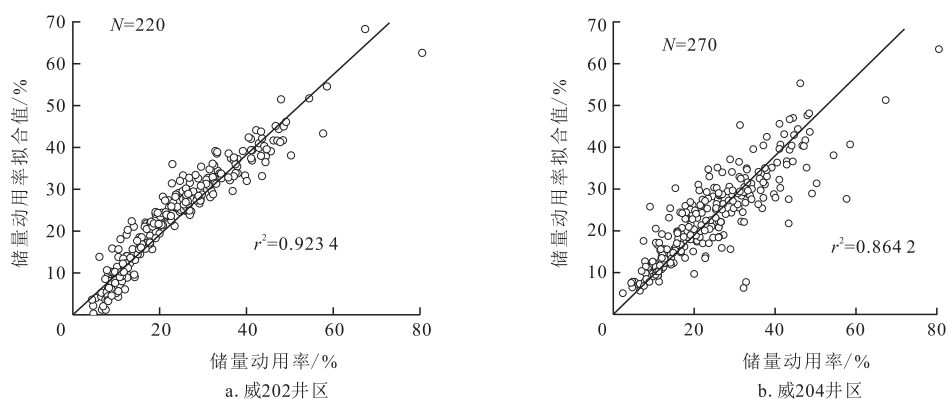


图3 威远地区储量动用程度拟合结果与计算结果对比

Fig. 3 Comparison between fitted and calculated results of reserve recovery ratio in Weiyuan area

本对比表明,当样本量越大,预测结果误差越小,综合预测法使用精度更高,效果更好。

#### 参考文献

- [1] 马新华,李熙喆,梁峰,等.威远页岩气田单井产能主控因素与开发优化技术对策[J].石油勘探与开发,2020,47(3):555-563.  
MA Xinhua, LI Xizhe, LIANG Feng, et al. Dominating factors on well productivity and development strategies optimization in Weiyuan shale gas play, Sichuan Basin, SW China[J]. Petroleum Exploration and Development, 2020, 47(3): 555-563.
- [2] 车世琦.涪陵气田平桥区块页岩气选区评价[J].石油地质与工程,2022,36(4):48-54.  
CHE Shiqi. Selection evaluation of shale gas in Pingqiao block of Fuling gas field[J]. Petroleum Geology & Engineering, 2022, 36(4): 48-54.
- [3] 张雪峰,徐最,康弘男,等.达坂城次凹二叠系页岩气地质特征与勘探潜力[J].石油地质与工程,2022,36(5):46-51.  
ZHANG Xuefeng, XU Zui, KANG Hongnan, et al. Geological characteristics and exploration potential of Permian shale gas in Dabancheng sub-sag[J]. Petroleum Geology & Engineering, 2022, 36(5): 46-51.
- [4] 王沛,欧阳传湘,陈宏生,等.应用PCA和多元非线性回归快速预测储层敏感性[J].断块油气田,2018,25(2):232-235.  
WANG Pei, OUYANG Chuanxiang, CHEN Hongsheng, et al. Application of PCA and multiple nonlinear regression to rapid prediction of reservoir sensitivity[J]. Fault-Block Oil & Gas Field, 2018, 25(2): 232-235.
- [5] 祁鹏.基于偏最小二乘的高炉铁水硅含量预测研究[D].包头:内蒙古科技大学,2010.  
QI Peng. Partial least squares prediction of silicon content in blast furnace[D]. Baotou: Inner Mongolia University of Science & Technology, 2010.
- [6] 吕娜.雪衣藻响应和适应环境胁迫分子调节的多组学研究[D].广州:华南理工大学,2014.  
LYU Na. Molecular regulation of snow alga *Chlamydomonas nivalis* in response to stress conditions using multi-omics study [D]. Guangzhou: South China University of Technology, 2014.
- [7] 卢丽,丁博,刘青,等.多元元素含量结合化学计量学技术在鳕鱼物种鉴别和产地溯源中的应用[J].食品安全质量检测学报,2021,12(24):9436-9444.  
LU Li, DING Bo, LIU Qing, et al. Application of multi-elements content analysis combined with chemometrics in species identification and origin tracing of cods[J]. Journal of Food Safety and Quality, 2021, 12(24): 9436-9444.
- [8] 李霜.基于代谢组学的数据标准化方法的综合评价[D].重庆:重庆大学,2018.  
LI Shuang. The comprehensive assessment of normalization methods in MS-based metalomics data[D]. Chongqing: Chongqing University, 2018.
- [9] 李君岩.偏最小二乘回归在油田固井质量预测中的应用研究[D].大庆:东北石油大学,2014.  
LI Jungang. Application of partial least squares regression in oilfield cementing quality prediction[D]. Daqing: Northeast Petroleum University, 2014.
- [10] 柯朝甫,张涛,武晓岩,等.代谢组学数据分析的统计学方法[J].中国卫生统计,2014,31(2):357-359.  
KE Chaopu, ZHANG Tao, WU Xiaoyan, et al. Statistical methods for Metabolomics data analysis[J]. Chinese Journal of Health Statistics, 2014, 31(2): 357-359.
- [11] 郝天轩,柳猛.基于多元非线性回归理论的煤层瓦斯含量预测[J].煤炭技术,2014,33(9):1-3.  
HAO Tianxuan, LIU Meng. Prediction of coal seam gas content based on multivariate nonlinear regression[J]. Coal Technology, 2014, 33(9): 1-3.
- [12] 王昕旭.偏最小二乘回归在孔隙度预测中的应用[J].地球物理学进展,2015,30(6):2807-2813.  
WANG Xinxu. Partial least-squares regression for the application in porosity prediction[J]. Progress in Geophysics, 2015, 30(6): 2807-2813.
- [13] 冯明刚,严伟,朱林奇,等.基于Boosting-PLS回归的页岩总含气量预测方法[J].煤炭技术,2018,37(10):23-25.  
FENG Minggang, YAN Wei, ZHU Linqi, et al. Method to predict total gas content in shale reservoir combining boosting-PLS algorithm and conventional logging data[J]. Coal Technology, 2018, 37(10): 23-25.
- [14] 朱如凯,李梦莹,杨静儒,等.细粒沉积学研究进展与发展方向[J].石油与天然气地质,2022,43(2):251-264.  
ZHU Rukai, LI Mengying, YANG Jingru, et al. Advances and trends of fine-grained sedimentology[J]. Oil & Gas Geology,

- 2022, 43(2): 251-264.
- [15] 郑爱维,梁榜,舒志国,等.基于大数据PLS法的页岩气产能影响因素分析——以四川盆地涪陵气田焦石坝区块为例[J].天然气地球科学,2020,31(4):542-551.  
ZHENG Aiwei, LIANG Bang, SHU Zhiguo, et al. Analysis of influencing factors of shale gas productivity based on large data technology: A case of Jiaoshiiba block in Fuling Gas Field, Sichuan Basin[J]. Natural Gas Geoscience, 2020, 31(4): 542-551.
- [16] 徐轩,朱华银,徐婷,等.多层合采气藏分层储量动用特征及判定方法[J].特种油气藏,2015,22(1):111-114.  
XU Xuan, ZHU Huayin, XU Ting, et al. Separated reserve producing characteristic and determination in multi-layer commingled producing gas reservoir[J]. Special Oil and Gas Reservoirs, 2015, 22(1): 111-114.
- [17] 朱桂良,孙建芳,刘中春.塔河油田缝洞型油藏气驱动用储量计算方法[J].石油与天然气地质,2019,40(2):436-442.  
ZHU Guiliang, SUN Jianfang, LIU Zhongchun. An approach to calculate developed reserves in gas drive fractured-vuggy reservoirs in Tahe oilfield[J]. Oil & Gas Geology, 2019, 40(2): 436-442.
- [18] 张歧,吕华,徐运动,等.含水饱和度对低渗砂岩气藏储量动用的影响[J].西部探矿工程,2012,24(4):85-87.  
ZHANG Qi, LYU Hua, XU Yundong, et al. Influence of water saturation on reserve production of low permeability sandstone gas reservoir[J]. West-China Exploration Engineering, 2012, 24(4): 85-87.
- [19] 贾爱林,王国亭,孟德伟,等.大型低渗-致密气田井网加密提高采收率对策——以鄂尔多斯盆地苏里格气田为例[J].石油学报,2018,39(7):802-813.  
JIA Ailin, WANG Guoting, MENG Dewei, et al. Well pattern infilling strategy to enhance oil recovery of giant low-permeability tight gas field: A case study of Sulige gas field, Ordos Basin[J]. Acta Petrolei Sinica, 2018, 39(7): 802-813.
- [20] 胡翌,刘维霞.利用储量动用质量评价储量动用状况的新方法[J].石油天然气学报,2011,33(9):60-63.  
HU Gang, LIU Weixia. A new method of evaluating reserves producing status by using reserves producing quality[J]. Journal of Oil and Gas Technology, 2011, 33(9): 60-63.
- [21] 胡翌.储量动用状况量化表征研究[J].成都理工大学学报(自然科学版),2012,39(1):100-103.  
HU Gang. Research on quantitative characterization of evaluating reserves employing status[J]. Journal of Chengdu University of Technology (Science & Technology Edition), 2012, 39(1): 100-103.
- [22] 刘亚龙,王海燕,王飞.太阳大寨区块浅层页岩气水平井优快钻井技术研究[J].石油和化工设备,2021,24(6):146-148.  
LIU Yalong, WANG Haiyan, WANG Fei. Study on optimal and fast drilling technology for shallow horizontal shale gas well in Taiyang Block[J]. Petro & Chemical Equipment, 2021, 24(6): 146-148.
- [23] 梁兴,单长安,张朝,等.昭通太阳背斜山地浅层页岩气“三维封存体系”富集成藏模式[J].地质学报,2021,95(11):3380-3399.  
LIANG Xing, SHAN Changan, ZHANG Zhao, et al. “Three-dimensional closed system” accumulation model of Taiyang anticline mountain shallow shale gas in the Zhaotong demonstration area[J]. Acta Geologica Sinica, 2021, 95(11): 3380-3399.
- [24] 梁兴,徐政语,张介辉,等.浅层页岩气高效勘探开发关键技术——以昭通国家级页岩气示范区太阳背斜区为例[J].石油学报,2020,41(9):1033-1048.  
LIANG Xing, XU Zhengyu, ZHANG Jiehui, et al. Key efficient exploration and development technologies of shallow shale gas: A case study of Taiyang anticline area of Zhaotong National Shale Gas Demonstration Zone[J]. Acta Petrolei Sinica, 2020, 41(9): 1033-1048.
- [25] 梁兴,张朝,单长安,等.山地浅层页岩气勘探挑战、对策与前景——以昭通国家级页岩气示范区为例[J].天然气工业,2021,41(2):27-36.  
LIANG Xing, ZHANG Zhao, SHAN Changan, et al. Exploration challenges, countermeasures and prospect of mountain shallow shale gas: A case study on the Zhaotong National Shale Gas Demonstration Area[J]. Natural Gas Industry, 2021, 41(2): 27-36.
- (编辑 徐佩)

(上接第466页)

- [20] BESKOK A, KAMIADAKIS G E. A model for flows in channels, pipes, and ducts at micro and nano scales[J]. Microscale Thermophysical Engineering, 1999, 3(1): 43-77.
- [21] BEHRANG A, KANTZAS A. A hybrid methodology to predict gas permeability in nanoscale organic materials; a combination of fractal theory, kinetic theory of gases and Boltzmann transport equation[J]. Fuel, 2017, 188: 239-245.
- [22] 李亚雄,刘先贵,胡志明,等.页岩气滑脱、扩散传输机理耦合新方法[J].物理学报,2017,66(11):230-240.  
LI Yaxiong, LIU Xiangui, HU Zhiming, et al. A new method for the transport mechanism coupling of shale gas slippage and diffusion[J]. Acta Physica Sinica, 2017, 66(11): 230-240.
- [23] 贾爱林,位云生,刘成,等.页岩气压裂水平井控压生产动态预测模型及其应用[J].天然气工业,2019,39(6):71-80.  
JIA Ailin, WEI Yunsheng, LIU Cheng, et al. A dynamic prediction model of pressure control production performance of shale gas fractured horizontal wells and its application[J]. Natural Gas Industry, 2019, 39(6): 71-80.
- [24] 孔祥言.高等渗流力学[M].合肥:中国科学技术大学出版社,1999.  
KONG Xiangyan. Advanced mechanics of fluids in porous media[M]. Hefei: University of Science and Technology of China Press, 1999.
- [25] 赵玉龙,梁洪彬,井翠,等.页岩气井EUR快速评价新方法[J].西南石油大学学报(自然科学版),2019,41(6):124-131.  
ZHAO Yulong, LIANG Hongbin, JING Cui, et al. A new method for quick EUR evaluation of shale gas wells[J]. Journal of Southwest Petroleum University (Science & Technology Edition), 2019, 41(6): 124-131.
- (编辑 余聪)