基于气田开发评价的指标协同处理与项目排序

李海涛，方一竹，陈艳茹，邹 禺

（中国石油天然气股份有限公司西南油气田分公司勘探开发研究院，成都 610000）

摘要：针对气田项目多指标项目排序问题，提出一种融合层次分析法（AHP）和模糊综合评价（FCE）的评价方法。建立定量分析模型，采用专家咨询和一致性检验确定指标权重。通过K-S（Kolmogorov-Smirnov）检验以及对指标的核心影响因素分析，确定数据样本分布，对于服从正态、对数正态分布或者均匀分布的指标分别采用改进的 Z 值法和改进的min-max法进行归一化处理，并用对数的方法剔除异常值，统一指标量级实现对项目进行定性与定量评价。使用层次分析法实现天然气项目的多指标排序，并结合聚类分析和轮廓系数分析生成项目推荐等级，构建综合系数优选图和内部收益率与设计规模指标的协同分析图，为气田开发方案决策和项目评级提供参考。

关键词：气田开发评价；指标协同排序；改进 Z 值法；改进min-max法

中图分类号：TE32 **文献标志码：**A **文章编号：**1671－1807(2026)03－0039－10

随着市场需求的持续增长和能源结构的优化调整，对天然气项目进行合理评估和排序在优化资源配置、提高投资效益以及确保国家能源安全方面尤为重要。国内学者主要聚焦于层次分析法（analytic hierarchy process，AHP）在特定场景的应用。万有余等[1]针对青海油田大型压裂施工管理，通过层次分析法分析压裂工序权重，构建了压裂管理制度，有效提升了施工效率；陆亚秋等[2]以涪陵页岩气田为例，采用层次分析法和均方差法相结合确定11个指标权重，建立了页岩气开发选区评价方法计算公式，并确定分类阈值；梁裕如等[3]针对延安气田增压方案，构建基于技术经济与安全指标的层次模型，发现“集中增压＋12站区域增压”为最优解。国外研究则侧重于方法整合与改进。Amiri[4]创新融合模糊AHP与TOPSIS（逼近理想解排序）法构建页岩气田土壤修复植物选择体系，其三角模糊数处理增强了不确定性分析能力；Al-Dhafeeri[5]开发了基于AHP的重油开采风险评估框架，通过多维度因子关联提升决策透明度；He等[6]针对地下储气库选址评价中的不确定性因素（如盖层等），开发了一套基于AHP-FCE（层次分析法-模糊综合评价）的综合评价方法，建立了包含6个指标和44个子指标的完整评价体系，并验证了其有效性。

应用层次分析法之前，数据的预处理至关重要。传统层次分析法依赖专家打分确定权重，若原始数据范围差异过大，易导致权重分配偏差，针对油气藏邻域多指标数据存在量纲差异的问题，直接进行指标处理与综合量化缺乏可比性。Yang等[7]针对隧道掘进机刀座性能评价和选择的多准则决策问题，提出了一种基于模糊层次分析法和模糊综合评价的混合评价方法，并对备选方案数据进行定性和定量处理，将定性指标转化为隶属度向量，并对定量指标进行min-max、线性归一化处理，最后对备选方案进行综合评价和排序。刘乔寿等[8]针对联邦学习中通信开销过大的问题，利用 Z 值法计算每个梯度值与平均值之间的距离，提出了基于 Z 值法的动态稀疏压缩算法，实现了降低通信成本的同时保证模型性能。刘传斌等[9]针对专家评分尺度差异导致评价结果不公的问题，建立了基于相对逆序数的评议数据影响力参数的数据处理方法，解决了量化专家评分对最终排序结论的影响。

鉴于此，提出一种融合层次分析法和模糊综合评价气田项目多指标排序方法。首先基于文献调研筛选评价指标并检验其统计分布特征，再采用改进 Z 值法和改进min-max法进行归一化处理，对含异常值指标使用对数变换进行处理。随后构建层

收稿日期：2025-07-30

基金项目：2024年度中国石油青年科技专项计划（2024DQ03028）

作者简介：李海涛（1978—），男，四川成都人，博士，高级工程师，研究方向为油气田开发规划；通信作者方一竹（1989—），女，四川成都人，硕士，高级工程师，研究方向为油气田开发规划；陈艳茹（1988—），女，重庆人，硕士，高级工程师，研究方向为开发规划；邹禺（1997—），女，四川内江人，硕士，助理工程师，研究方向为开发规划。

次分析递阶模型，通过专家赋权与一致性检验确定指标权重，分析核心指标关联性，并对权重前两位的指标进行协同效应解析，构建指标协变关系分析图，探析其与项目综合评分的相互作用规律。最后通过聚类分析划分项目推荐等级，建立综合评价排序体系。综上，提出一种融合层次分析法和模糊综合评价的气田项目多指标排序方法，通过改进的数据预处理方法和指标权重确定方法，有效地解决数据量纲差异和权重分配偏差问题，实现气田开发项目的综合评价和等级推荐，并揭示项目选择规律，以期为气田开发方案决策和目标优选提供有效参考。

1　多层次模糊评价分析

1.1　评价方法

气藏开发优先级评价需考虑经济、技术不同类别参数，定量化研究这些参数时，存在影响因素多、主控因素不明确、参数敏感性未知、评价指标未量化等问题，因此，这类问题常常采用层次分析法进行研究[10-14]。层次分析法是将定性问题定量化处理常见的方法，核心思想是通过构建不同影响因素的多层次分析结构模型，利用定性指标模糊量化的方法计算因素的权重值，然后得到不同因素的影响程度及总目标的最终权重[15-17]。

西南油气田在四川盆地发育海相碳酸盐岩与页岩等多元储层类型，目前探明率较低，展现出巨大的勘探开发潜力。根据规划目标，至 2030 年该区域天然气年产量将突破 600 亿～800 亿 m^3，着力建设成为国内领先的现代化天然气工业基地。为实现这一战略目标，亟须系统开展气藏开发规划、经济效益评估及优化策略研究，实施科学统筹布局，确保在产量持续增长的同时达成资源高效开发与经济效益双优目标。针对区内多类型气藏项目特征，基于井点实钻数据构建定量评价模型，通过经济指标与技术指标的多维耦合分析，建立覆盖全要素的综合评价体系。具体思路是：①在综合分析研究气藏特征的基础之上，优选评价指标；②将影响因素进行分类筛选，通过条理化、层次化归类，建立层次结构模型；③在评价指标标准化的基础之上，构造判断矩阵并确定指标权重；④进行因素排序及一致性检验；⑤对评价结果进行验证分析[18]（图 1）。

1.2　评价指标确定

西南油气田主力气藏呈现多维异质性特征，传统开发决策采用的单一指标无法有效表征地质-工程-经济复合系统的非线性关联。以内部收益率为例，其指标权重仅反映经济维度，掩盖了龙王庙组

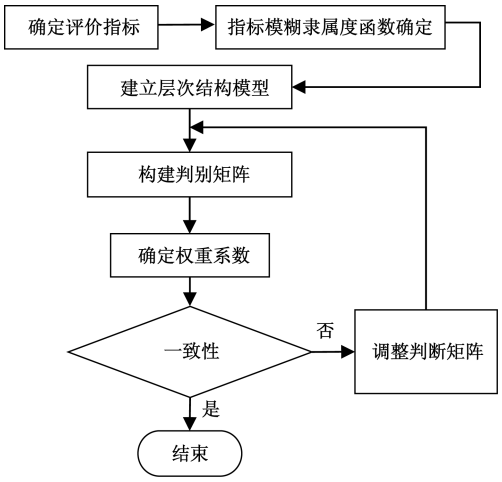


图 1　层次分析技术流程

水侵防控的技术成本、元坝高压开发的安全投入等隐性制约要素[19-23]。因此，指标集的确定要考虑开发实际生产需要，包含地质资源潜力、开发技术可行性、经济效益、风险因素、生产动态以及环境影响等影响因素，综合评估各方面影响，再根据实际情况进行校正[24-26]。

研究立足西南油气田项目实际生产情况，选取多个指标重点构建包含以下核心指标的评估体系，即设计规模、单井累计产量、亿立方米产能进尺、亿立方米产能投资、内部收益率、单位完全成本、净现值。同时利用层次分析法与模糊综合评价法对气田开发区块进行排序，以实现项目综合排序的目的。研究成果可为西南油气田开发选区提供有力借鉴，对实现中国西南地区气田大规模开发具有重要意义。

为了确保评价指标的客观真实性，避免指标权重分散，将影响气田排序方案的因素分为 2 个一级指标和 7 个二级指标，建立综合排序方案的三层次评价指标结构体系（图 2）。

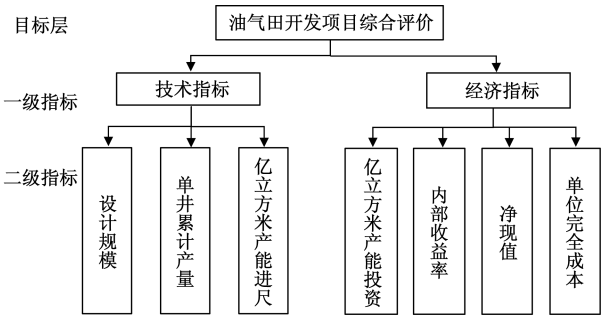


图 2　优选评价的三层次评价指标结构体系

2　评价指标权重集的确定

2.1　统一指标量级

数据标准化处理是一种将不同量纲、不同尺度

的数据进行统一处理的方法,可以消除数据间的差异,提高数据的可比性和可解释性[9]。若未标准化,高量级指标(如完全成本)会掩盖低量级指标(如累计产量)的贡献,导致综合评价结果失真。因此对于参数量纲不同的数据,直接比较不能体现各气藏指标间差异,且没有实际意义。综上,为使各项参数具有可比性,需对各项参数的取值进行标准化处理,同时考虑到不同分布特性的指标数据处理的方式不同,则需使用不同的标准化方法对数据进行处理。

2.1.1 确定指标分布

对各指标的参数特性进行分析,从而确定各指标分布。由于亿立方米产能进尺、单位完全成本、净现值指标的样本量较大,可优先考虑使用柯尔莫哥洛夫-斯米尔诺夫(Kolmogorov-Smirnov,K-S)检验指标的正态性,K-S检验作为一种对两个(连续)分布进行比较的非参数统计方法,能够针对单个样本是否来自某一特定分布(如正态分布等)的问题进行检验。假设检验的问题如下:H_0:样本来自的总体服从某分布;H_1:样本来自的总体不服从某分布。K-S检验统计量 $D=\max|F_n(x)-F_0(x)|$。其中:$F_n(x)$为待检验分布的分布函数;$F_0(x)$为某特定分布。检验结果见表1。

由上述检验结果可知,3个指标K-S、S-W的P值均大于0.05,呈现显著性,则亿立方米产能进尺、单位完全成本、内部收益率指标均服从正态分布。

考虑到其他二级指标数据样本数较少,直接使用正态性检验偏差较大,因此通过分析影响各二级指标的各因素间是否存在交叉影响以及影响程度是否均衡来判定其分布特性。设计规模的核心影响因素是资源禀赋,受地质构造、有机质丰度、热演化历史等复杂因素控制,其分布具有空间异质性和尺度依赖性,且气田开发需满足最小经济规模,导致规模分布存在天然下限。结合USGS和IHS数据库的工程实践数据,气田规模显著右偏,得出设计规模指标是符合对数正态分布的;单井累计产量受渗透率、孔隙度、裂缝发育程度等参数的空间差异影响,但三者之间不存在交叉影响,即服从均匀分布。

净现值由多期现金流折现求和构成,而现金流受气价波动、产量衰减、成本变动等正数随机变量的综合影响。这些变量通常服从几何级增长或指数衰减过程,其对数转换后呈正态性。同时,净现值作为多期乘性效应的累积结果,根据中心极限定理表现为长尾特征,即净现值服从对数正态分布。如果气田投资项目较多,且这些项目的规模和成本控制相对稳定,那么根据中心极限定理,许多独立同分布的随机变量的和倾向于服从正态分布,得到亿立方米产能投资近似正态分布。考虑到投资成本往往不会取负值,且通常有长尾特性,即存在一些非常高成本的项目,那么此时项目的亿立方米产能投资是服从对数正态分布的。

2.1.2 确定样本分布参数

确定各指标样本分布后,再通过极大似然估计法计算样本均值、样本方差。考虑到不同分布总体有不同的计算方法,故需对不同分布总体的均值、方差计算方法进行分类讨论。

(1)设总体有分布 $f(x;\theta_1,\theta_2,\cdots,\theta_k)$,$X_1,X_2,\cdots,X_n$ 是从正态总体中 $N(\mu,\sigma^2)$ 中抽出的样本,则样本的似然函数 $f(x_1;\theta_1,\theta_2,\cdots,\theta_k)\cdots f(x_n;\theta_1,\theta_2,\cdots,\theta_k)$,记为$L(x_1,x_2,\cdots x_n;\theta_1,\theta_2,\cdots,\theta_k)$,简写为$L$,有

$$L=\prod_{i=1}^{n}\left\{(\sqrt{2\pi\sigma^2})^{-1}\exp\left[-\frac{1}{2\sigma^2}(X_i-\mu)^2\right]\right\} \tag{1}$$

再解下面方程组:

$$\begin{cases}\dfrac{\partial \ln L}{\partial \mu}=\dfrac{1}{\sigma^2}\sum\limits_{i=1}^{n}(X_i-\mu)=0\\[2ex]\dfrac{\partial \ln L}{\partial(\sigma^2)}=-\dfrac{n}{2\sigma^2}+\dfrac{1}{2\sigma^4}\sum\limits_{i=1}^{n}(X_i-\mu)^2=0\end{cases} \tag{2}$$

得到样本分布的均值和方差:

$$\mu=\frac{1}{n}\sum_{i=1}^{n}X_i \tag{3}$$

$$\sigma^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2 \tag{4}$$

对 σ^2 进行修正,得到方差的无偏估计 s^2 为

表1 亿立方米产能进尺正态性检验结果

指标	样本数	平均值	标准差	偏度	峰度	K-S检验		S-W检验	
						统计量D值	P	统计量W值	P
亿立方米产能进尺	30	2.044	0.719	0.302	−1.094	0.129	0.233	0.944	0.120
单位完全成本	30	1 042.53	141.08	0.78	1.53	0.1	0.64	0.93	0.06
内部收益率	30	8.64	4.87	−0.72	1.27	0.11	0.41	0.96	0.22

$$s^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2 \tag{5}$$

(2)设总体有分布 $f(x;\theta_1,\theta_2,\cdots,\theta_k)$,

$$f(x;\theta_1,\theta_2,\cdots,\theta_k)=\begin{cases}\dfrac{1}{b-a}, & \text{若 } a\leqslant x\leqslant b\\ 0, & \text{其他}\end{cases} \tag{6}$$

设 $X_1,X_2,\cdots,X_n$ 为从均匀分布 $U(a,b)$ 的总体中抽出的样本,满足 $a\leqslant X_i\leqslant b,i=1,2,\cdots,n$,并且

$$L=\prod_{i=1}^{n}f(X_i;a,b)=(b-a)^n \tag{7}$$

为使似然函数最大,确保样本落在区间 $[a,b]$ 中,取 $a=\min(X_1,X_2,\cdots,X_n)$,$b=\max(X_1,X_2,\cdots,X_n)$,得到均值、方差如下:

$$E(X)=\frac{1}{b-a}\int_a^b x\,\mathrm{d}x=\frac{1}{2}(a+b) \tag{8}$$

$$D(X)=E(X^2)-(EX)^2=\frac{1}{12}(b-a)^2 \tag{9}$$

2.1.3　数据归一化

确定指标分布后,由于指标数据的量级不同,不具有可比性,因此还需要对指标进行归一化处理。本文主要使用下面 3 种归一化方法对数据进行处理。

(1)改进的 min-max 法。对于服从区间 $[a,b]$ 上均匀分布 $X\sim[a,b]$ 的指标,要使其尽可能落在 $[0,1]$ 区间上,步骤如下。

第 1 步:使用极大似然估计出均匀分布 X 的参数 a、b。

第 2 步:考虑到均匀分布中,对于任意一个连续型随机变量,都可以通过它的累计分布函数线性变换成一个服从 $[0,1]$ 区间上的均匀分布 Y,则有线性关系式 $Y=pX+q$ 成立。根据均值和方差的线性有

$$\begin{cases}E(Y)=pE(X)+q=\dfrac{1}{2}\\ D(Y)=p^2D(X)=\dfrac{1}{12}\end{cases} \tag{10}$$

式中:$E(Y)$ 为 $[0,1]$ 区间上的均匀分布 Y 的均值;$D(Y)$ 为 $[0,1]$ 区间上的均匀分布 Y 的方差;p 和 q 为待求解的一次函数系数。解方程组得到

$$\begin{cases}p=\sqrt{\dfrac{1}{12D(X)}}\\ q=\dfrac{1}{2}-\sqrt{\dfrac{1}{12D(X)}}E(X)\end{cases} \tag{11}$$

结合式(8)、式(9),则服从区间 $[a,b]$ 上均匀分布 X 与 $[0,1]$ 区间上的均匀分布 Y 的关系为

$$Y=\sqrt{\frac{1}{(b-a)^2}}X+\frac{1}{2}-\sqrt{\frac{1}{(b-a)^2}}\frac{a+b}{2} \tag{12}$$

(2)改进的 Z 值法。对于服从区间 $[\mu,\sigma]$ 上正态分布的数据 $X\sim[\mu,\sigma]$,要使其尽可能落在 $[0,1]$ 区间上,步骤如下。

第 1 步:使用极大似然估计出正态分布 X 的参数 a、b。

第 2 步:对正态分布 X 进行标准化,使其服从标准正态分布 $Y\sim[0,1]$。

第 3 步:为保持数据分布特性不变,使用 3σ 法则对数据进行压缩,使数据落在 3σ 区间 $[\mu-3\sigma,\mu+3\sigma]$ 内的概率为 99.7%。

$$\begin{aligned}&p\left(\mu'-3\sigma'\leqslant\frac{Y'-\mu'}{\sigma'}\leqslant\mu'+3\sigma'\right)\\&=\int_{\mu'-3\sigma'}^{\mu'+3\sigma'}f(x)\,\mathrm{d}x=\phi(3)-\phi(-3)\\&=0.997\,4\end{aligned} \tag{13}$$

$$[\mu'-3\sigma',\mu'+3\sigma']=[0,1] \tag{14}$$

式中:$\phi(z)$ 为累积分布函数,z 为 $[-3,3]$ 的概率为 $\phi(3)-\phi(-3)$;μ'、σ' 为压缩后的正态分布 Y' 的均值、方差,即 $Y'\sim[\mu',\sigma']$,取 $\mu'=\frac{1}{2}$,得到 $\sigma'=\frac{1}{6}$,则服从区间 $[\mu,\sigma]$ 上正态分布的数据 X 与变换后的正态分布 Y' 的关系为

$$Y'=\left(\frac{X-\mu}{\sigma}+\frac{1}{2}\right)\frac{1}{6} \tag{15}$$

第 4 步:对于落在 $[0,1]$ 区间外的异常值,考虑将对数变换应用于指标数据。对数变换通过压缩高值区域,扩展低值区域,从而减少右偏,增强指标正态性。处理方法如下:计算指标数据最大值与当前元素的差加 1 的对数,然后将这个对数结果除以指标数据最大值的对数。

$$\mu_i=\frac{\ln[\max(D_i)-(D_i+1)]}{\ln[\max(D_i)]},\quad i=1,2,\cdots,n \tag{16}$$

式中:n 为指标数据量;$D=[D_1,D_2,\cdots,D_n]$ 为出现异常值的指标数据列;$\mu=[u_1,u_2,\cdots,u_n]$ 为消除异常值后的指标数据列。

(3)改进的对数正态归一法。针对服从对数正态分布的净现值、亿立方米产能投资指标,即

$$\ln X\sim N(\mu,\delta) \tag{17}$$

第 1 步:使用对数变换将所有数据替换成以 e 为底的对数。

第 2 步:后续处理步骤同方法(2)类似,计算得到两个指标的归一化数据。

2.2 构造判断矩阵

在确定各层次、各因素之间的权重时,如果只是定性的结果,则常常不容易被接受,因而需使用专家打分法,即不把所有因素放在一起比较,而是两两相互比较。因此,采用相对尺度来尽可能减少性质不同的诸因素相互比较的困难,提高准确度,涉及判断矩阵标度含义如下:标度 1 代表两个因素同等重要;标度 2 代表因素 i 相比因素 j 比较重要;标度 3 代表因素 i 相比另一个因素 j 非常重要;取数字 1~3 及其倒数作为标度,分别代表同等重要、比较不重要、非常不重要[27-29](表 2)。

表 2 标度取值

因素 i 相比因素 j 的重要程度	标度
非常重要	3
比较重要	2
同等重要	1
比较不重要	1/2
非常不重要	1/3

根据专家意见和现场经验判断结果,一级指标判断矩阵为

$$\boldsymbol{A}=(a_{ij})_{n\times n}=\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{pmatrix} \tag{18}$$

式中:n 为一级指标数据个数;$a_{ij}=\dfrac{1}{a_{ji}}$。

二级指标技术和经济指标判断矩阵为 $\boldsymbol{B}^{(1)}$、$\boldsymbol{B}^{(2)}$。

$$\boldsymbol{B}^{(1)}=(b_{ij}^{(1)})_{n\times n}=\begin{pmatrix}b_{11}^{(1)} & b_{12}^{(1)} & \cdots & b_{1n}^{(1)}\\ b_{21}^{(1)} & b_{22}^{(1)} & \cdots & b_{2n}^{(1)}\\ \vdots & \vdots & & \vdots\\ b_{n1}^{(1)} & b_{n2}^{(1)} & \cdots & b_{nn}^{(1)}\end{pmatrix} \tag{19}$$

$$\boldsymbol{B}^{(2)}=(b_{ij}^{(2)})_{m\times m}=\begin{pmatrix}b_{11}^{(2)} & b_{21}^{(2)} & \cdots & b_{1n}^{(2)}\\ b_{21}^{(2)} & b_{22}^{(2)} & \cdots & b_{2n}^{(2)}\\ \vdots & \vdots & & \vdots\\ b_{n1}^{(2)} & b_{n2}^{(2)} & \cdots & b_{mm}^{(2)}\end{pmatrix} \tag{20}$$

式中:n 为二级技术指标数据个数;m 为二级经济技术指标数据个数;$b_{ij}^{(1)}=\dfrac{1}{b_{ji}^{(1)}}$;$b_{ij}^{(2)}=\dfrac{1}{b_{ji}^{(2)}}$。

2.3 指标权重确定与一致性检验

将判断矩阵的各行向量进行几何平均,再进行归一化,得到一级指标与二级指标权重和特征向量。然后通过计算各级矩阵的一致性指标与一致性比率,对一级指标和二级指标的判断矩阵进行一致性检验。若满足检验,即判断矩阵具有满意的一致性,即计算得到的权重为各级指标的真实权重;若不满足检验,即判断矩阵一致性较差,需要重新调整判断矩阵。

2.4 *K*-means(*K* 均值)聚类分析

考虑到油气田开发的特征数据维度高、样本量大,且在实际生产数据归一化处理中,*K*-means 算法以其低计算复杂度迅速处理大规模高维数据,符合工程实践对时效性的要求[30-34]。通过标准化或归一化手段消除量纲差异,*K*-means 在数值型数据的聚类效果上表现稳定,非常适合于油气田开发的多指标融合分析。*K*-means 算法生成的簇中心能够直观地量化表征不同气田子类的特征,便于与 AHP 评价指标(如资源丰度、开发难度等)进行对接,构建出一个清晰的层次化评价体系。相较于层次聚类[$O(n^3)$复杂度]或 DBSCAN(密度聚类算法),*K*-means[$O(nkt)$复杂度]更适用于处理油气田开发的海量项目数据,满足实时或准实时的分析需求。其次,采用轮廓系数法确定最佳聚类数目。轮廓系数通过衡量簇内凝聚度与簇间分离度来评估聚类质量,其值域为[−1,1],值越接近 1 表明聚类效果越好[35]。通过计算不同 k 值对应的平均轮廓系数并基于实际生产需要及决策考量,选择相应的 k 值,可确保分类结果的客观性和合理性。

综上,使用轮廓系数法和 *K*-means 聚类对项目的综合评价结果进行量化分级,并划分为 k 个推荐等级,其中$[R_{k-1},R_k)$代表最优推荐项目,$[R_1,R_2)$推荐优先级最低。

考虑到气井的综合评价值一般直接反映在各一级指标的两个主要参数上,因此,考虑绘制两项参数归一化数据的变化趋势曲线。绘制方法如下。

(1)对项目在两个参数上的总得分使用 *K*-means方法得到的聚类结果,找到 k 个分类中的最小值。

(2)将最小值升序排列,舍弃数值最小的数,将其余 $k-1$ 个数由小到大分别记为 $m_1,m_2,\cdots,m_k$。

(3)分别确定两个主要参数的真实权重,记为 λ_1、λ_2。

(4)将两个主要参数的归一化数据极值设置为横纵坐标最大值,即 x 轴数值区间为$[0,x_{\max}]$,y 轴数值区间为$[0,y_{\max}]$。

(5)绘制直线 $y_i=k_ix+b_i$,$i=1,2,\cdots,k-1$,其中 $k_i=-\dfrac{\lambda_1}{\lambda_2}$,$b_i=\dfrac{m_i}{\lambda_2}$。

(6)绘制直线 $\lambda_1x=\lambda_2y$。

3　算例研究

为验证评价结果的有效性，选取西南油气田下某 30 个气藏项目为例(表 3)进行计算分析。

基于西南油气田区块数据得到的经济技术指标的样本分布情况，对今后的同品质的气藏开发有一定的参考价值，可以更深层次从整体指标的分布、分布参数有一个量化的认识，对其他项目也有意义。再利用前文介绍的确定分布参数方法得到指标的具体分布如下：①设计规模服从 $X\sim \ln N(17.35,\ 22.55)$；②净现值服从 $X\sim \ln N(18.3,48.47)$；③亿立方米产能投资服从均匀分布 $X\sim \ln N[2.27,6.37]$；④单井累计产量服从均匀分布 $X\sim U[0.73,6.15]$；⑤亿立方米产能进尺服从 $X\sim N[2.04,0.72]$；⑥单位完全成本服从 $X\sim N[1\,042.53,141.08]$；⑦内部收益率服从 $X\sim N[8.64,4.87]$。

然后对各级评级指标使用改进 Z 值法、改进 min-max 法和改进的对数归一化方法进行标准化处理，将所有指标数据归一到[0,1]上，不仅保留了数据随机分布信息也不会丢失指标信息。以内部收益率为例(表 4)，使用改进 Z 值法进行归一化处理。

为量化评估不同归一化方法的性能，以内部收益率指标为例，对比了传统 min-max 法、传统 Z 值法以及本文提出的改进 Z 值法的处理效果。传统 Z 值法 $\left(Z=\frac{X-\mu}{\delta}\right)$ 虽能消除量纲，但处理后的数据均值为 0，标准差为 1，且理论上无界，该方法虽能消除量纲，但不利于后续基于距离的 K-means 聚类分析。传统 min-max 法 $\left(Z=\frac{X-X_{\min}}{X_{\max}-X_{\min}}\right)$ 将数据线性压缩至[0,1]区间，但极度依赖样本最大值和最小值，对异常值非常敏感，一个异常值就会压缩正常数据的分布区间，导致分布形态失真。基于指标总体分布的参数进行归一化，并对异常值进行对数处理，旨在将数据映射到[0,1]区间的同时，最大限度地保留其总体的分布形态。这确保了基于这些指标进行的 AHP 权重计算和 K-means 聚类更能反映气田项目的真实差异。表 5 给出内部收益率指标经不同方法处理后的关键统计特征对比。

综上所述，与传统方法相比，本文中采用的改进归一化策略在保证数据可比性的基础上，显著优化了对于数据原始分布特征的保留效果，为后续基于距离和多指标综合的排序与聚类分析奠定了更可靠的数据基础。下面构造判断矩阵以及确定指标权重，最后得到两个一级指标的最大特征值、判断矩阵的权重向量和一致性指标与比率，以及两个二级指标的一致性指标与比率，结果如下。

表 3　指标参数原始值

项目	技术指标			经济指标			
	设计规模/亿 m^3	单井累计产量/亿 m^3	亿立方米产能进尺/[万 m·(亿 m^3)]	亿立方米产能投资/[亿元·(亿 m^3)$^{-1}$]	内部收益率/%	单位完全成本/[元·(km^3)$^{-1}$]	净现值/亿元
X1	11.00	6.15	1.20	4.44	10.49	865.00	6.20
X2	16.00	6.15	0.96	4.15	13.85	841.00	20.69
X3	9.00	4.62	1.08	5.06	6.56	962.00	−3.80
X4	5.00	4.62	1.31	5.16	7.29	945.00	−1.12
X5	5.00	6.09	1.60	6.37	4.13	1 049.00	−9.61
X6	10.00	6.15	1.06	3.85	9.37	840.00	3.48
X7	10.00	1.31	2.69	3.20	12.42	1 182.00	19.06
X8	40.00	1.25	3.26	3.29	11.51	1 236.00	60.35
X9	10.00	1.17	3.06	3.44	11.18	1 142.00	21.64
X10	10.00	1.22	3.14	3.53	10.58	1 167.00	19.71
X11	5.00	1.25	2.75	3.45	12.73	1 090.00	8.93
X12	5.00	1.18	2.88	3.47	12.10	1 118.00	12.86
X13	60.00	1.28	2.18	2.54	17.40	1 079.00	179.44
X14	115.00	1.33	2.65	2.97	14.40	1 156.00	186.22
X15	30.00	1.00	2.64	2.27	15.27	1 098.00	62.15
X16	35.00	4.62	1.46	3.96	8.56	926.00	3.40
X17	15.00	6.15	1.19	4.01	9.08	871.00	3.52
X18	30.00	4.89	1.68	5.16	6.34	955.00	−10.20
…	…	…	…	…	…	…	…
X30	3.00	0.93	1.85	3.56	4.39	1 065.00	−1.22

表 4　不同数据处理方法对比

项目	原始内部收益率/%	传统 min-max 法	Z 值法	改进的 Z 值法
X1	10.49	0.70	0.38	0.55
X2	13.85	0.85	1.07	0.64
X3	6.56	0.54	−0.43	0.45
X4	7.29	0.57	−0.28	0.46
X5	4.13	0.43	−0.93	0.38
X6	9.37	0.66	0.15	0.52
X7	12.42	0.79	0.78	0.60
X8	11.51	0.75	0.59	0.57
X9	11.18	0.73	0.52	0.57
X10	10.58	0.71	0.40	0.55
X11	12.73	0.80	0.84	0.61
X12	12.10	0.77	0.71	0.59
X13	17.40	1.00	1.80	0.73
X14	14.40	0.87	1.18	0.65
X15	15.27	0.91	1.36	0.67
X16	8.56	0.62	−0.02	0.50
X17	9.08	0.64	0.09	0.51
X18	6.34	0.53	−0.47	0.44
…	…	…	…	…
X30	4.39	0.44	−0.87	0.39

表 5　内部收益率指标不同归一化方法结果对比

统计特征	原始数据	传统 min-max 法	传统 Z 值法	改进的 Z 值法
取值范围	[−5.96，17.40]	[0.00，1.00]	[−3，1.8]	[0.12，0.73]
平均值	8.64	0.63	0.00	0.50
标准差	4.87	0.21	1.00	0.13
偏度	−0.72	−0.72	−0.72	−0.72
峰度	1.27	1.27	1.27	1.27

注：传统方法的统计量基于样本计算；本文方法基于归一化后的结果计算。

(1)一级判断矩阵：$\boldsymbol{A}=\begin{bmatrix}1 & 1.5\\ 0.67 & 1\end{bmatrix}$。

(2)二级判断矩阵：$\boldsymbol{B}_1=\begin{bmatrix}1 & 0.5 & 0.58 & 1.38\\ 2 & 1 & 2 & 1.75\\ 1.72 & 0.5 & 1 & 2\\ 0.73 & 0.57 & 0.5 & 1\end{bmatrix}$，$\boldsymbol{B}_2=\begin{bmatrix}1 & 1.88 & 1.62\\ 0.53 & 1 & 1.75\\ 0.62 & 0.57 & 1\end{bmatrix}$。

(3)两个一级指标的判断矩阵为 $\boldsymbol{W}_1=[0.19\quad 0.38\quad 0.27\quad 0.16]^{\mathrm{T}}$，$\boldsymbol{W}_2=[0.46\quad 0.31\quad 0.23]^{\mathrm{T}}$，特征值为 $\lambda_{\max}^1=4.08$，$\lambda_{\max}^2=3.06$。

(4)判断矩阵 $\boldsymbol{A}$ 的一致性比率为 0，判断矩阵 $\boldsymbol{B}_1$ 的一致性比率为 0.03，判断矩阵 $\boldsymbol{B}_2$ 的一致性比率为 0.03，均小于 0.1，所有判断矩阵均通过一致性检查，说明判断矩阵可靠，特征向量即为权值。

从评价模型建立及计算结果可以看出，针对油气田开发项目综合评价选取的几个影响因素可以看到，内部收益率、设计规模为综合评价主控因素。通过使用层次分析法计算项目综合得分。部分结果见表 6。

在进行 K-means 聚类前，为确定最佳聚类数目 k，并考虑计算了 k 从 2～6 时的平均轮廓系数(图 3)。当 $k=6$ 时，平均轮廓系数达到最大值 0.635，表明将 30 个气藏项目分为 6 类具有最佳的聚类结构，但基于实际业务需求和研究目的，将项目分为

表 6　油气田开发项目气藏实例计算结果

项目	技术指标			经济指标				综合得分 R
	设计规模	单井累计产量	亿立方米产能进尺	亿立方米产能投资	内部收益率	单位完全成本	净现值	
X1	0.51	1.00	0.70	0.49	0.55	0.66	0.47	0.61
X2	0.58	1.00	0.76	0.53	0.64	0.68	0.51	0.67
X3	0.46	0.78	0.73	0.38	0.45	0.57	0.44	0.53
X4	0.34	0.78	0.67	0.36	0.46	0.59	0.45	0.51
X5	0.34	0.99	0.60	0.00	0.38	0.49	0.43	0.45
X6	0.49	1.00	0.73	0.57	0.52	0.68	0.46	0.62
X7	0.49	0.32	0.35	0.65	0.60	0.37	0.50	0.48
X8	0.78	0.31	0.21	0.64	0.57	0.33	0.61	0.52
X9	0.49	0.30	0.26	0.62	0.57	0.41	0.51	0.47
X10	0.49	0.30	0.24	0.61	0.55	0.39	0.50	0.46
X11	0.34	0.31	0.33	0.62	0.61	0.46	0.48	0.46
X12	0.34	0.30	0.30	0.62	0.59	0.43	0.49	0.45
X13	0.86	0.31	0.47	0.72	0.73	0.47	0.92	0.65
X14	1.00	0.32	0.36	0.67	0.65	0.40	0.94	0.64
X15	0.72	0.27	0.36	0.74	0.67	0.45	0.62	0.57
X16	0.75	0.78	0.64	0.56	0.50	0.61	0.46	0.61
X17	0.57	1.00	0.70	0.55	0.51	0.65	0.46	0.62
X18	0.72	0.82	0.59	0.36	0.44	0.58	0.43	0.56
…	…	…	…	…	…	…	…	…
X30	0.23	0.26	0.55	0.61	0.39	0.48	0.45	0.61

注：权重条件为设计规模为 0.185 1、单井累计产量为 0.124 7、亿立方米产能进尺为 0.135 1、亿立方米产能进尺为 0.111 0、内部收益率为 0.229 7、单位完全成本为 0.162 2、净现值为 0.096 2。

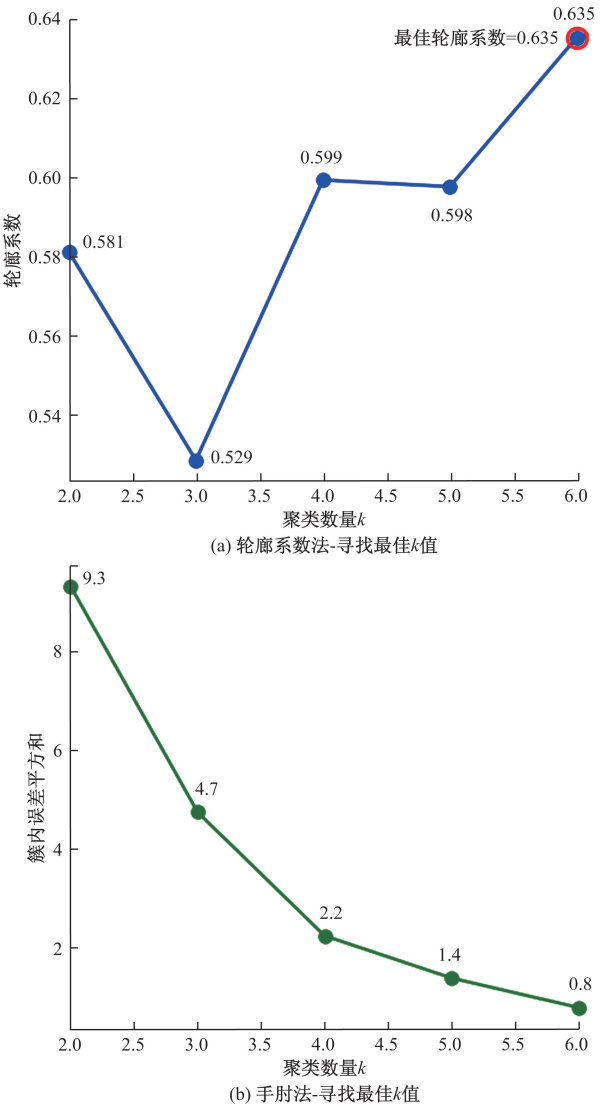


图 3　不同聚类数 k 对应的平均轮廓系数及簇内误差平方和

4 个推荐等级，此决策基于以下考虑：①业务解释性。4 级分类（非常推荐、推荐、一般推荐、暂不推荐）更符合气田开发项目管理的实际应用场景，便于决策者理解和操作。②样本数考量。当前样本数为30 个项目，分为 6 类可能导致某些类别样本数过少（<3 个项目），统计意义不足。③实用性优先。在聚类质量可接受的前提下（$k=4$ 时轮廓系数为 0.62，仍高于 0.5），选择更简洁的分类体系有利于实际应用。因此，最终确定聚类数 $k=4$。

然后使用 K-means 对综合得分结果聚类，得到项目的 k 个评价推荐等级：推荐评价值 $R<0.4$，一般推荐评价值 R 介于[0.4，0.5)，推荐评价值 R 介于[0.5，0.6)，非常推荐评价值 $R\geq0.6$。考虑到气井的综合评价值一般直接反映在内部收益率和设计规模两个参数上，因此绘制两个指标的归一化数据随评价值变化的曲线，首先使用 K-means 聚类得到 3 个参数值，即 $m_i(i=1,2,3)$ 分别为 0.16，0.21，0.28，然后分别确定内部收益率、设计规模真实权重在经济指标真实权重中的占比，记为 λ_1、λ_2，即 $\lambda_1=0.6\times0.387=0.23$，$\lambda_2=0.4\times0.463=0.19$，最后绘制直线：① $y_i=kx+b_i$，$i=1,2,3$，其中 $k=-\frac{\lambda_1}{\lambda_2}$，$b_i=\frac{m_i}{\lambda_2}$；② $\lambda_1 x=\lambda_2 y$。

得到内部收益率-设计规模协变关系分析图（图 4），图中的数字为每个项目所对应的项目编号。

图 4 中，虚线分界线是对加权数据的聚类分类结果，黑色实线分界线是指标偏好分界，上方的项目更偏重设计规模，下方的项目更偏重内部收益率。实线分界线是一条以原点为起点，斜率为 λ_1/λ_2 的直线，越靠近直线的项目，其内部收益率得分与设计规模得分越接近。位于分界线上方的项目设计规模真实得分高于其内部收益率真实得分，更偏重设计规模指标；位于分界线下方的项目设计规模真实得分小于其内部收益率真实得分，更偏重内部收益率指标。实线上的项目代表在设计规模和内部收益率两个方面表现均衡。选取 X1～X30 井的评价值绘制与内部收益率和设计规模的关系曲线，结合表 6、图 4 可以看出：整体趋势为评价值越大内部收益率和设计规模越高，气井综合评价值越高，即评价值与内部收益率、设计规模呈现很好的一致性。

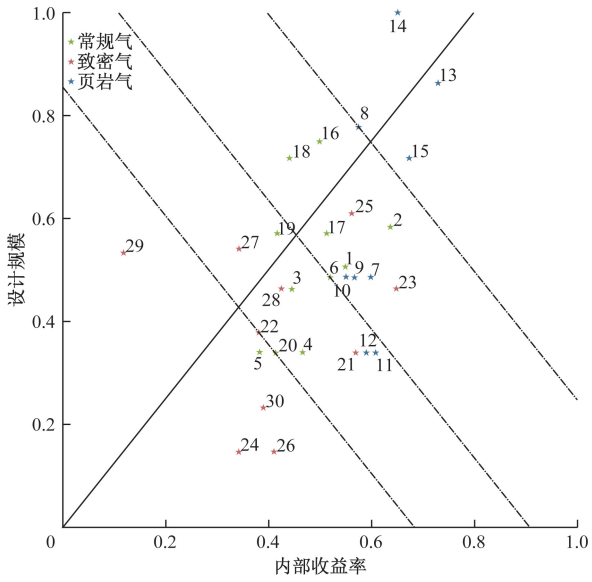


图 4　内部收益率-设计规模散点图

4　结论

（1）改进的指标量级统一方法有效提升数据可比性。针对油气藏项目多指标数据量纲差异的问

题,通过 K-S 检验分析指标数据的分布特征,采用改进 Z 值法与改进 min-max 法的双重标准化策略,有效地将不同分布特征的指标数据统一到[0,1]区间,在消除极端值影响的同时保留数据分布特征,实现指标量纲统一与数据可比性提升。

(2)指标协变关系分析图揭示项目选择规律。通过绘制内部收益率与设计规模指标协同分析图,实现开发方案决策可视化,并结合西南油气田区域特征,采用 K-means 聚类生成项目推荐等级,揭示项目优选规律。

(3)对于油气田各开发项目,基于层次分析法建立储层评价模型的方法可以很好地认识气藏,实现储层的定量综合评价,方法具有一定推广性。下一步可结合人工智能,通过计算机学习进一步开展更加复杂、精确的储层定量分类评价研究。

参考文献

[1] 万有余,林海,郭得龙,等. 层次分析法制定大型压裂管理制度提升非常规储层压裂效果[J]. 石油钻采工艺,2024,46(4):479-491.

[2] 陆亚秋,王进,曹梦茜. 基于改进的层次分析法的页岩气开发选区评价方法[J]. 油气藏评价与开发,2021,11(2):70-77.

[3] 梁裕如,何鹏,韩建红,等. YQ2 井区增压集输方案多层次模糊评价优选[J]. 油气储运,2022,41(12):1430-1437.

[4] AMIRI M P. Project selection for oil-fields development by using the AHP and fuzzy TOPSIS methods[J]. Expert Systems with Applications,2010,37:6218-6224.

[5] AL-DHAFEERI M. An analytic hierarchy process for technical risk assessment selection[D]. Manchester: The University of Manchester,2018.

[6] HE Y,ZHAO G,TANG Y,etal. Available target evaluation of underground gas storage based on AHP-FCE methods[J]. Geoenergy Science and Engineering,2024,242:213-255.

[7] YANG M,JI Z Y,ZHANG L,et al. A hybrid comprehensive performance evaluation approach of cutter holder for tunnel boring machine[J]. Advanced Engineering Informatics,2022,52:101546.

[8] 刘乔寿,皮胜文,原炜锡. 基于 Z-Score 动态压缩的高效联邦学习算法[J]. 计算机应用研究,2024,41(7):2093-2097.

[9] 刘传斌,余乐安,刘斌,等. 基于数据标准化的专家群组评分聚合方法比较[J/OL]. 系统工程理论与实践,1-21[2025-06-27]. https://link.cnki.net/urlid/11.2267.N.20250224.1429.010.

[10] 徐文军,张莎莎,吴财芳,等. 基于层次分析法的煤储层构造复杂程度定量评价[J]. 河南理工大学学报(自然科学版),2019,38(2):20-26.

[11] 邓雪,李家铭,曾浩健,等. 层次分析法权重计算方法分析及其应用研究[J]. 数学的实践与认识,2012,24(7):93-100.

[12] 张小东,张硕,孙庆宇,等. 基于 AHP 和模糊数学评价地质构造对煤层气产能的影响[J]. 煤炭学报,2017,42(9):2385-2392.

[13] 马云东,胡明东. 改进的 AHP 法及其在多目标决策中的应用[J]. 系统工程理论与实践,1997,17(6):41-45.

[14] 张旭,颜其彬,李祖兵. 陆相碎屑岩储层定量评价的新方法:以河南某油田为例[J]. 天然气地球科学,2007,18(1):141-148.

[15] 米伟伟,谢小飞,曹红霞,等. 鄂尔多斯盆地东南部二叠系山 2-盒 8 段致密砂岩储层特征及主控因素[J]. 岩性油气藏,2022,34(6):101-117.

[16] 王永骁,付斯一,张成弓,等. 鄂尔多斯盆地东部山西组 2 段致密砂岩储层特征[J]. 岩性油气藏,2021,33(6):12-20.

[17] 胡昌蓬,徐大喜. 页岩气储层评价因素研究[J]. 天然气与石油,2012,30(5):38-42.

[18] 周锋,黄仕林,李晓明,等. 基于层次分析法的致密气藏储层分类定量评价:以新场气田蓬莱镇组气藏为例[J]. 油气藏评价与开发,2024,14(3):468-474,483.

[19] 张烈辉,胡勇,李小刚,等. 四川盆地天然气开发历程与关键技术进展[J]. 天然气工业,2021,41(12):60-72.

[20] 陈蟒蛟,谭开俊,文龙,等. 四川盆地中二叠统天然气成藏特征及巨大勘探前景[J]. 地学前缘,2023,30(1):11-19.

[21] 谢增业,杨春龙,李剑,等. 四川盆地致密砂岩天然气成藏特征及规模富集机制:以川中地区上三叠统须家河组气藏为例[J]. 天然气地球科学,2021,32(8):1201-1211.

[22] 吕文雅,苗凤彬,张本键,等. 四川盆地剑阁地区须家河组致密砾岩储层裂缝特征及对天然气产能的影响[J]. 石油与天然气地质,2020,41(3):484-491,557.

[23] 郑志红,李登华,白森舒,等. 四川盆地天然气资源潜力[J]. 中国石油勘探,2017,22(3):12-20.

[24] 气田开发主要生产技术指标及计算方法:SY/T 6170-2012[S]. 北京:石油工业出版社,2012.

[25] 李兴彦,刘俊杰,安小絮,等. 气田开发方案经济评价研究及应用[J]. 中国石油和化工标准与质量,2013,33(8):146.

[26] 高倩. 气田开发投资经济评价方法研究[D]. 北京:中国石油大学,2008.

[27] 董凤娟,冯兵,何光渝,等. 应用层次分析法确定地下储气库最优设计方案[J]. 油气储运,2007(10):18-21,62.

[28] 李东晖,田玲钰,聂海宽,等. 基于模糊层次分析法的页岩气井产能影响因素分析及综合评价模型:以四川盆地焦石坝页岩气田为例[J]. 油气藏评价与开发,2022,12(3):417-428,454.

[29] SAATY T L. Decision making with the analytic hierarchy process[J]. International Journal of Services Sciences, 2008, 1(1): 83-98.
[30] 郭秀英，陈义才，张鉴，等. 页岩气选区评价指标筛选及其权重确定方法：以四川盆地海相页岩为例[J]. 天然气工业，2015，35(10)：57-64.
[31] 周天伦，曾超，范晨，等. 基于快速聚类-信息量模型的汶川及周边两县滑坡易发性评价[J]. 中国地质灾害与防治学报，2021，32(5)：137-150.
[32] 余果，李海涛，方一竹. 基于组合预测理论的天然气产量峰值模型[J]. 天然气勘探与开发，2025，48(1)：87-96.
[33] 余果，李海涛，方一竹. 峰值预测理论在四川盆地致密气的应用[J]. 天然气勘探与开发，2024，47(2)：89-96.
[34] 余果，刘海峰，李海涛，等. 川南页岩气峰值产量预测[J]. 天然气勘探与开发，2023，46(1)：97-104.
[35] PETER R J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis[J]. Journal of Computational & Applied Mathematics, 1987, 20: 53-65.

Collaborative Optimization of Evaluation Metrics and Project Prioritization for Gas Field Development

LI Haitao, FANG Yizhu, CHEN Yanru, ZOU Yu

(Research Institute of Exploration and Development, Southwest Oil and Gasfield Company, PetroChina Company Limited, Chengdu 610000, China)

Abstract: An evaluation method combining analytic hierarchy process (AHP) and fuzzy comprehensive evaluation (FCE) was proposed to solve the multi-index project ranking problem of gas field projects. A quantitative analysis model was established, and the weights of indicators were determined by expert consultation and consistency test. Through K-S(Kolmogorov-Smirnov) test and analysis of the core influencing factors of the indicators, the distribution of data samples was determined. The improved *Z*-value method and improved min-max method were used to normalize the indicators that are subject to normal or lognormal distribution, and uniform distribution, to achieve the qualitative and quantitative evaluation of the projects with unified index levels. AHP is used to realize the multi-index ranking of natural gas projects, and the project recommendation ranking is generated by combining cluster analysis. The comprehensive coefficient optimization diagram and the cooperative analysis diagram of internal rate of return and design scale indicators are constructed to provide reference for the gas field development plan decision-making and project rating.

Keywords: gas field development evaluation; index collaborative sorting; improved *Z*-value method; improved min-max method