applied sciences

MDPI

Article

Quantitative Evaluation and Screening of Complex Structural Traps in the Southern Junggar Basin Using a Support Vector Machine

Lianghou Xu [1], Zhichao Pang [2], Chunming Jia [2], Wei Wang [3], Gangqiang Chen [1,*] and Xiaoquan Li [1]

[1] College of Earth Sciences and Engineering, China University of Petroleum-Beijing at Karamay, Karamay 834000, China; 2024216819@st.cupk.edu.cn (L.X.)
[2] Research Institute of Exploration and Development, PetroChina Xinjiang Oilfield Company, Karamay 834000, China
[3] Urumqi Branch of Research Institute, Bureau of Geophysical Prospecting Inc., CNPC, Urumqi 830011, China
* Correspondence: 2017592004@cupk.edu.cn; Tel.: +86-990-6633522

Abstract

Complex structural traps in the southern Junggar Basin require quantitative evaluation despite limited drilling-based labels. This study examined a support vector machine (SVM) workflow for 49 traps, including 16 labeled and 33 unlabeled traps. Eleven geological variables were considered, with grouped feature selection and model optimization performed within nested leave-one-out cross-validation (LOOCV). The final six-variable linear SVM achieved an accuracy of 0.5625 and a weighted F1-score of 0.5382 under strict nested LOOCV, matching the majority-class baseline accuracy (0.5625); nested leave-one-structural-belt-out validation yielded 0.4375, below the corresponding majority-class baseline (0.5625). None of the four Class II traps were correctly identified, while a source-rock-thickness-only SVM achieved higher accuracies under both validation schemes (0.6875 and 0.6250). Feature-selection stability was moderate (mean pairwise Jaccard similarity = 0.6732). The final model assigned 18 of 33 unlabeled traps to Class I, with 25 and 26 categorical agreements with fuzzy comprehensive evaluation and gray relational analysis, respectively. Seven illustrative traps were examined to interpret geological agreement and disagreement; their descriptive SVM scores do not constitute independently validated prospect rankings. The findings demonstrate the limitations of multivariable SVM screening with extremely small labeled datasets and the value of cautious geological cross-method interpretation.

Keywords: complex structural traps; trap evaluation; trap prioritization; southern margin of the Junggar Basin; support vector machine; extremely small sample size; leave-one-out cross-validation; feature-selection stability; model re-optimization

Check for updates

Academic Editor: Nikolaos Koukouzas

Received: 28 September 2026
Revised: 26 October 2026
Accepted: 27 October 2026
Published: 28 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

As hydrocarbon exploration extends into deep and subtle plays [1], complex structural traps increasingly involve diverse accumulation mechanisms, difficult structural interpretation, and high drilling costs. Reliable trap evaluation is therefore essential for reducing exploration risk. Traps are fundamental units of hydrocarbon accumulation and preservation [2] and are commonly evaluated within a petroleum-system framework by integrating structural, source-rock, reservoir, migration, and preservation-related factors to assess hydrocarbon potential and geological risk.

Trap evaluation has progressively developed from qualitative expert assessment to quantitative, probabilistic, fuzzy, and data-driven approaches [3–8]. Geological risk probability methods and Monte Carlo simulation introduced probabilistic treatment of exploration uncertainty [4], while fuzzy comprehensive evaluation, gray relational analysis, and artificial neural networks enabled the integration of multiple geological factors [5–7]. Petroleum-system and accumulation-dynamics studies subsequently incorporated the temporal and spatial relationships among hydrocarbon generation, expulsion, migration, and accumulation [8]. More recently, machine-learning methods, including support vector machines (SVMs), have been increasingly applied to geological risk assessment and other petroleum-geoscience problems. These quantitative approaches provide alternatives to conventional expert-weighted or predefined functional formulations, but their reliability depends strongly on the amount and representativeness of available labeled data.

Earlier petroleum-prospect studies in the western Sverdrup Basin provide important precedents for the quantitative integration of geological evidence. Chen et al. mapped petroleum favorability using fuzzy integration [9], and Chen and Osadetz combined multivariate and Bayesian statistical methods for geological risk mapping and prospect evaluation [10]. Chen et al. subsequently applied an SVM-based, two-category approach to geological risk evaluation in a structural petroleum play [11]. These studies demonstrate complementary ways of combining geological attributes and statistical models in exploration assessment. Building on this foundation, the present study considers three-class evaluation of individual structural traps in the southern Junggar Basin, with particular attention to nested model validation, geological-variable selection stability, and the geological interpretation of agreement and divergence between SVM and established assessment methods.

Beyond structural-play risk mapping, SVM and related approaches have been applied to small-sample petroleum-geoscience problems, including kerogen-type discrimination [12], low-contrast oil-pay identification [13,14], and reservoir-parameter prediction [15,16]. However, quantitative structural-trap evaluation remains challenging where only a small number of drilled traps have reliable labels. In such settings, selected geological variables and SVM configurations may change as additional labeled traps become available, motivating explicit examination of feature-selection stability and model re-optimization.

Methodological research provides a basis for examining both model evaluation and the consistency of selected variables in small-sample studies. Ambroise and McLachlan demonstrated that selecting variables before cross-validation can bias estimated classification performance [17]. Varma and Simon [18] and Cawley and Talbot [19] further examined how model selection can affect cross-validation-based performance estimates. Nogueira et al. [20] developed a statistical framework for evaluating the stability of feature selection when the training data change. These studies motivate the nested cross-validation and fold-wise feature-selection stability analyses adopted in this work.

These questions are particularly relevant to the southern Junggar Basin, where deeply buried structural targets, strong tectonic deformation, and costly drilling complicate exploration decisions. In the study area, 16 of 49 identified traps have exploration-based labels. This limited training set motivates a compact SVM workflow that explicitly examines feature selection and model configuration. At the same time, established geological evaluation methods can offer a reference for examining whether the model classifications are consistent with recognized source-rock, reservoir, burial, and preservation conditions.

Accordingly, this study develops an SVM-based quantitative workflow for structural-trap evaluation in the southern Junggar Basin. Its objectives are to (1) construct and evaluate an SVM classification model using the available labeled traps; (2) assess the stability of selected geological variables and the adjustment of model configuration as labeled

samples accumulate; and (3) compare the resulting SVM evaluations with conventional trap-evaluation methods and interpret the similarities and differences in a petroleum-geological context.

2. Regional Geological Setting and Exploration History

2.1. Exploration History of the Southern Junggar Basin

Petroleum exploration in the southern Junggar Basin began in 1909 and has continued for more than a century. The exploration focus has progressively shifted from shallow targets to deeper objectives and from the upper and middle assemblages to the lower assemblage [21].

During 1909–1995, exploration initially focused on upper-assemblage plays in the Neogene Shawan and Taxihe formations. Early exploration was guided mainly by surface oil seeps, mud volcanoes, and shallow drilling. The Dushanzi oil well drilled in 1909 marked the beginning of regional petroleum exploration, and the discovery of the Qigu oilfield in 1958 confirmed the hydrocarbon potential of the shallow Jurassic succession [22].

During 1996–2008, the widespread application of two-dimensional seismic data and advances in drilling technology supported exploration of deeper targets. Discoveries including the Hutubi gas field, Kaindike, and Mahe fields [23] shifted exploration toward the middle assemblage, including Paleogene Anjihaihe mudstones and Cretaceous Donggou reservoirs.

During 2008–2019, exploration progressively advanced into the deep lower assemblage, including the Cretaceous Qingshuihe Formation and Jurassic intervals. Early wells such as Xihu 1, Dushan 1, and Dafeng 1 encountered hydrocarbon shows, but the limited ability of two-dimensional seismic data to delineate deep and structurally complex traps constrained exploration success. The subsequent application of prestack depth migration improved the imaging and delineation of deep structures, including the Gaoquan anticline. In 2019, Well Gaotan 1 encountered prolific oil and gas flow from the Qingshuihe Formation, marking a major exploration breakthrough in the lower assemblage [24].

Since 2020, exploration has increasingly emphasized integrated play evaluation, systematic trap delineation, and production-capacity development. Wells Hutubi 1, Tianwan 1, and Qingbei 1 have subsequently yielded high-rate oil and gas from the Qingshuihe Formation [25], further supporting the hydrocarbon potential of the deep southern margin.

Overall, exploration in the southern Junggar Basin has evolved from relatively shallow, isolated well-based exploration toward integrated evaluation of deep and structurally complex plays. The shift toward the lower assemblage has increased the importance of identifying and evaluating numerous structurally variable traps under limited drilling information. Consequently, systematic and reproducible trap evaluation has become increasingly important for exploration screening in the deep southern margin.

2.2. Distribution and Geological Characteristics of Structural Traps

The southern margin of the Junggar Basin is a typical foreland fold-and-thrust belt characterized by intense structural deformation and complex structural styles, and the identified traps are predominantly structural. A total of 49 structural traps have been identified in the study area across eight structural belts (Figure 1). Differences in stress regime, detachment-layer development, and source–reservoir configuration result in pronounced spatial variations in trap development. Detailed information on the 16 labeled traps used for quantitative evaluation is summarized in Table 1, while the complete information for all 49 identified traps is provided in Supplementary Table S1.

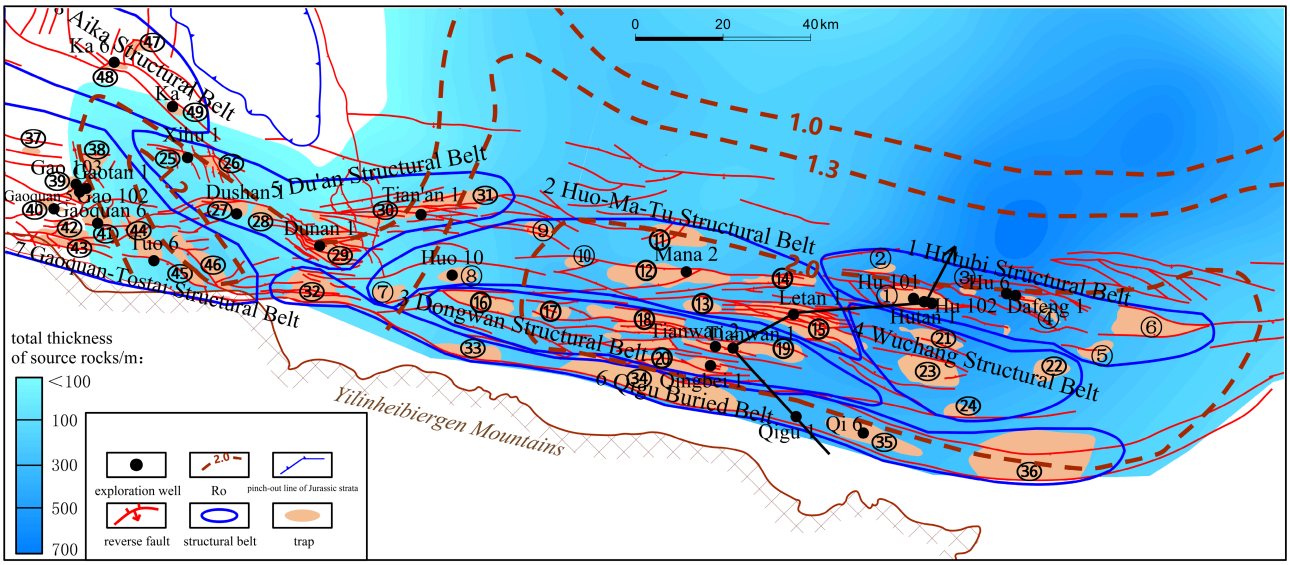


Figure 1. Trap distribution map of the southern Junggar Basin.

During the Early to Middle Jurassic, the basin was characterized by relatively low-relief topography and widespread shallow-water deltaic deposition. During the Middle to Late Jurassic, enhanced subsidence in the south produced a north-high/south-low structural framework [26], providing the regional setting for lower-assemblage reservoir and seal development. Figure 2 shows a representative north–south seismic section across the middle segment of the southern margin, crossing the Qigu, Dongwan, Huo-Ma-Tu, and Hutubi structural belts and illustrating the along-strike variation in deformation style within the foreland thrust system [27] and the lower-assemblage accumulation framework [28]. The Qigu buried belt is dominated by fault-nose traps over a broad range of burial depths. Well Qigu 1 encountered commercial oil flow from shallow Jurassic reservoirs, whereas deeper traps remain prospective lower-assemblage targets. The Dongwan structural belt contains large lower-assemblage anticlinal traps with substantial closure. Wells Qingbei 1 and Tianwan 1 have yielded high-rate gas from the Qingshuihe Formation, and Jurassic source-rock conditions are among the most favorable in the study area. The Huo-Ma-Tu structural belt is dominated by fault-nose traps. Well Fengtan 1 encountered commercial gas flow from the Cretaceous, whereas the complex fault system associated with multiple detachment levels [29] produces compartmentalization and complex gas–water relationships in some traps, as indicated by Well Letan 1. The Hutubi structural belt is characterized by relatively gentle deformation, limited fault damage, and favorable preservation conditions; Wells Hutubi 1, Hu 101, Hu 102, and Hu 103 have all yielded significant hydrocarbon production. Overall, the lower assemblage along this section comprises a vertical source–reservoir–seal configuration of Lower Jurassic source rocks, Qingshuihe and Middle–Upper Jurassic reservoirs, and regional mudstone seals of the Tugulu Group. Trap effectiveness is strongly influenced by the sealing capacity of deep faults and by compartmentalization associated with detachment layers.

Table 1. Geological characteristics of the 16 labeled structural traps used for quantitative evaluation.

Structural Belt	Trap ID	Trap Type	Trap Area (km^2)	Burial Depth (m)	Closure Height (m)	Source-Rock Thickness (m)	Source-Rock Maturity (%)	Gas Intensity (10^8 m^3/km^2)	K_1q Reservoir Lithology	K_1q Sedimentary Facies	K_1q Reservoir Thickness (m)	Fault-Damage Degree	Class
Hutubi structural belt	TR1	Anticlinal trap	84.78	7380	600	375	2.0%	60	Fine sandstone	Braided river delta front	34	Weak	I
	TR4	Anticlinal trap	22.46	6950	200	365	2.0%	60	Fine sandstone	Braided river delta front	43	Weak	I
Huo-Ma-Tu structural belt	TR8	Anticlinal trap	26.5	7500	400	310	1.4%	45	Siltstone	Braided river delta front	26	Weak	I
	TR12	Fault-nose trap	153.91	8250	500	340	2.0%	70	Medium-fine sandstone	Braided river delta front	35	Weak	I
	TR15	Fault-nose trap	178.4	7450	600	355	2.0%	80	Fine sandstone	Braided river delta front	40	Weak	II
Dongwan structural belt	TR19	Anticlinal trap	128.12	8200	900	380	2.0%	100	Argillaceous siltstone	Braided river delta front	24	Weak	I
	TR20	Anticlinal trap	91.7	7450	2200	375	2.0%	100	Gravelly sandstone	Braided river delta front	22	Weak	I
Du'an structural belt	TR25	Anticlinal trap	3.97	5340	100	135	1.1%	20	Argillaceous siltstone	Braided river delta plain	50	Weak	II
	TR27	Anticlinal trap	27.58	6090	1200	165	1.1%	16	Pebbly fine sandstone	Braided river delta front	36	Moderate	II
	TR29	Anticlinal trap	69.83	6730	800	160	1.1%	21	Fine sandstone	Braided river delta front	17	Strong	III
	TR30	Anticlinal trap	94.68	6750	500	220	1.1%	36	Pebbly sandstone	Braided river delta front	15	Moderate	II

Table 1. *Cont.*

Structural Belt	Trap ID	Trap Type	Trap Area (km^2)	Burial Depth (m)	Closure Height (m)	Source-Rock Thickness (m)	Source-Rock Maturity (%)	Gas Intensity (10^8 m^3/km^2)	K_1q Reservoir Lithology	K_1q Sedimentary Facies	K_1q Reservoir Thickness (m)	Fault-Damage Degree	Class
Qigu buried belt	TR35	Anticlinal trap	79.64	7400	350	360	1.9%	80	Siltstone	Braided river delta front	25	Weak	I
Gaoquan–Tostai structural belt	TR39	Fault-nose trap	27.02	5800	400	100	0.9%	35	Gravelly sandstone	Fan delta front	20	Weak	I
	TR40	Anticlinal trap	10.47	5850	200	85	0.8%	22	Sandy conglomerate	Fan delta plain	18	Weak	III
	TR41	Anticlinal trap	14.43	6520	250	115	0.9%	25	Sandy conglomerate	Fan delta plain	16	Weak	I
Aika structural belt	TR49	Anticlinal trap	6.57	5015	100	115	0.8%	16	Pebbly sandstone	Braided river delta front	10	Weak	III

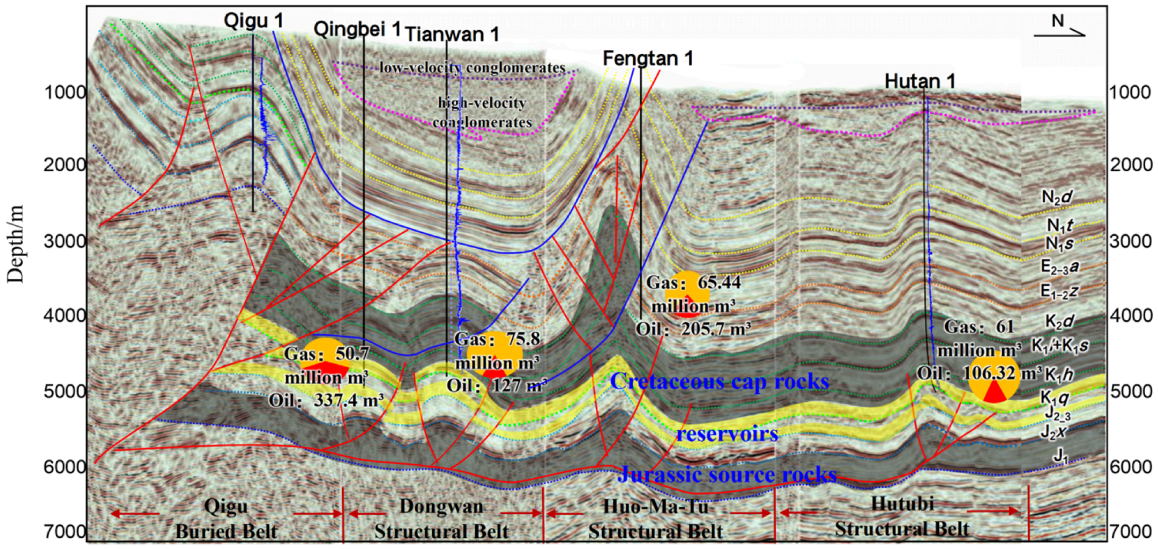


Figure 2. North–south interpreted seismic section across the middle segment of the southern Junggar Basin (location indicated by the black arrow in Figure 1).

The Wuchang transfer zone has a favorable source–reservoir configuration and limited fault damage, making it an important area for deep exploration. The Du'an structural belt has relatively poor source-rock conditions and heterogeneous reservoirs; its traps are mainly Class II or Class III, and their effectiveness depends strongly on the accurate assessment of fault sealing. The Gaoquan–Tostai structural belt lacks substantial local source rocks, and hydrocarbons are thought to be supplied mainly by Jurassic source rocks in the Sikeshu Sag [30]; accumulation is therefore controlled by the interaction between migration pathways and fault sealing. The Aika structural belt, adjacent to the Chepaizi Uplift, contains well-preserved anticlines and favorable fault-sealing conditions in its eastern part but depends largely on external hydrocarbon supply, with accumulation similarly controlled by fault sealing [31].

In summary, traps in the study area exhibit pronounced zonation parallel to the structural belts and segmentation in the east–west direction. The north–south distribution is controlled mainly by progressive deformation of multiple thrust systems, whereas east–west variations reflect differences in source-rock maturity, detachment-layer distribution, and major strike-slip accommodation zones. Deep lower-assemblage traps are dominated by anticlinal and fault-nose types and generally occur at great burial depths with favorable source–reservoir configurations, making them important exploration targets. Accumulation studies in the basin interior indicate that Indosinian–Yanshanian strike-slip faults can act as vertical hydrocarbon-migration conduits, whereas Himalayan low-angle thrust faults provide preferential lateral migration pathways; effective spatial coupling between faults and sand bodies is important for hydrocarbon enrichment [32]. Multi-stage fault interactions should therefore be considered when evaluating trap effectiveness in the structurally complex southern margin.

3. Conventional Trap-Evaluation Methods

Three conventional geological evaluation methods were applied to the 49 structural traps to provide reference assessments for comparison with the SVM. Fuzzy comprehensive evaluation and gray relational analysis used the same six geological variables as the final SVM: burial depth, Jurassic source-rock thickness, Jurassic source-rock maturity, Jurassic gas-generation intensity, reservoir thickness, and fault-damage degree. Both methods employed a common geological weighting scheme while retaining their respective

evaluation procedures. Geological risk probability evaluation followed a multiplicative formulation integrating source-rock, reservoir, trap, and preservation factors.

3.1. Geological Risk Probability

The geological risk formulation was used to obtain a dimensionless composite score $P = P_s\ P_r\ P_t\ P_p$ for each of the 49 traps. The source-rock component was defined as P_s = 0.3*s_thickness* + 0.3*s_maturity* + 0.4*s_gas*, where the three terms are normalized Jurassic source-rock thickness, maturity, and gas-generation intensity. The coefficients were predefined geological weights, assigning equal contributions to source-rock thickness and maturity (0.3 each) and a relatively higher contribution to gas-generation intensity (0.4). They were not estimated or optimized using the labeled trap outcomes. The reservoir component *P_r* was normalized reservoir thickness, the trap component *P_t* was normalized trap area, and the preservation component *P_p* was set at 1.0, 0.5, and 0.1 for weak, moderate, and strong fault damage, respectively. Continuous factors were min–max normalized over the 49-trap inventory. The composite P score was used to rank traps according to relative geological favorability. For descriptive visualization, the ranked traps were divided into high- (ranks 1–16), intermediate- (17–33), and low-score (34–49) groups. These groups represent relative P-score positions rather than the drilling-based Classes I–III. Because the score includes multiplicative normalized factors, its numerical values were interpreted as a relative evaluation index rather than a calibrated occurrence probability.

3.2. Fuzzy Comprehensive Evaluation

Fuzzy comprehensive evaluation was conducted using six geological variables retained through full-sample SVM feature selection: burial depth, Jurassic source-rock thickness, Jurassic source-rock maturity, Jurassic gas-generation intensity, reservoir thickness, and fault-damage degree. The corresponding weight vector was (0.150, 0.100, 0.100, 0.125, 0.150, 0.375), reflecting the authors' geological judgment regarding the relative importance of the evaluation factors. The weights were normalized to sum to unity. For continuous variables, class-reference centers were determined from the respective medians of the labeled training traps. Piecewise-linear semi-trapezoidal membership functions were constructed using these centers, with equal membership assigned where class centers coincided. Fault-damage degrees of weak, moderate, and strong were associated with Class I, II, and III memberships, respectively. Weighted aggregation yielded the membership vector $B = (B_\mathrm{I}, B_\mathrm{II}, B_\mathrm{III})$, and each trap was assigned to the class with the highest membership. A descriptive evaluation index was calculated as $D = B_\mathrm{I} + 0.5B_\mathrm{II}$. For the leave-one-out cross-validation (LOOCV) assessment, normalization parameters and class-reference centers were recalculated using the 15 training traps in each fold. The final evaluation of all 49 traps was performed using the reference centers derived from the complete set of 16 labeled traps.

3.3. Gray Relational Analysis

Gray relational analysis used the same six geological variables and weights (0.150, 0.100, 0.100, 0.125, 0.150, 0.375). Continuous features were min–max normalized using the labeled training traps; weak, moderate, and strong fault damage were assigned preservation scores of 1.0, 0.5, and 0.1, respectively. A class-specific reference sequence was formed from the component-wise mean of the normalized attributes in the corresponding labeled class. With a distinguishing coefficient $\rho = 0.5$, variable-specific relational coefficients were $\xi_cj = \rho/(|z_j - r_cj| + \rho)$, and the class relational grade was $\gamma_c = \Sigma_j\ w_j\ \xi_cj$. Each trap was assigned to the class with the highest γ_c, and the Class I relational grade γ_I was used for ranking. In each leave-one-out fold, normalization and reference sequences were

calculated from the other 15 labeled traps. The final 49-trap evaluation used the full 16-trap reference set.

3.4. Results of the Three Geological Evaluation Methods

The geological risk index was evaluated for all 49 traps, with p values ranging from 0 to 0.686888. The descriptive rank-based grouping comprised 16 high-, 17 intermediate-, and 16 low-score traps (Figure 3a). Among the 33 unlabeled traps, TR34, TR36, TR21, TR23, TR22, TR32, and TR14 had the seven highest P scores (Supplementary Table S2). Because these groups represent relative score positions, not outcome-based trap classes, a classification accuracy was not calculated for the geological risk index. Its high sensitivity to unfavorable multiplicative factors should be considered when interpreting rank differences.

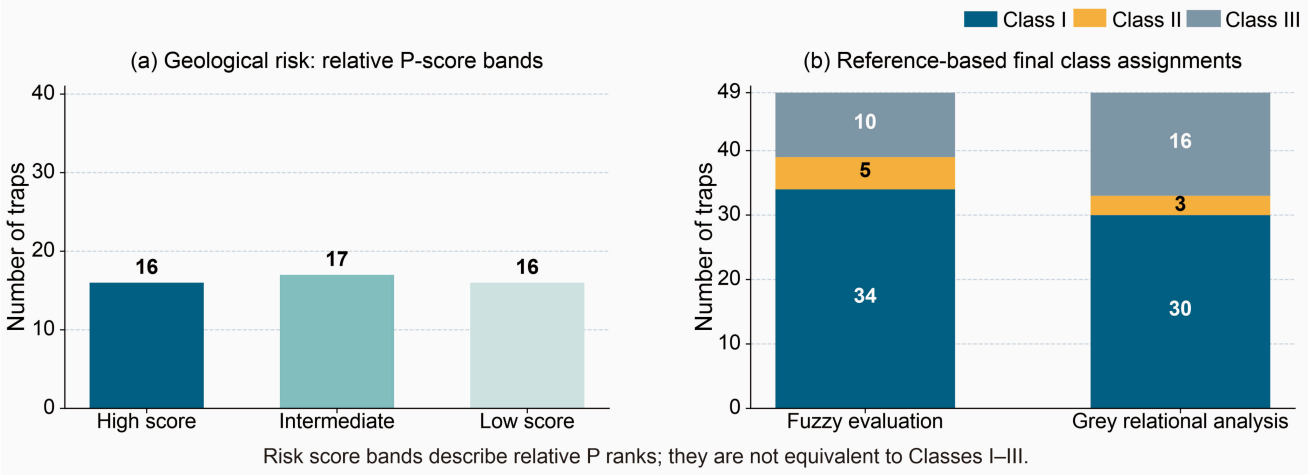


Figure 3. Results for 49 traps: (**a**) geological-risk relative P-score groups (16 high, 17 intermediate, 16 low); (**b**) Class I–III assignments by fuzzy comprehensive evaluation and gray relational analysis. The risk groups are not equivalent to observed outcome classes.

Using the six-variable fuzzy formulation, leave-one-out evaluation of the 16 labeled traps yielded 14 correct class assignments (accuracy = 0.8750; weighted F1-score = 0.8604; macro F1-score = 0.8556; balanced accuracy = 0.8333). The Class I, II, and III recalls were 9/9, 2/4, and 3/3, respectively. The final evaluation assigned 34, five, and 10 of the 49 traps to Classes I, II, and III (Figure 3b); among the 33 unlabeled traps, the corresponding counts were 23, three, and seven. Rankings by D are reported in Supplementary Table S2.

The six-variable gray relational analysis correctly classified 10 of the 16 labeled traps under leave-one-out evaluation (accuracy = 0.6250; weighted F1-score = 0.5655; macro F1-score = 0.4563; balanced accuracy = 0.5926). Recalls for Classes I, II, and III were 7/9, 0/4, and 3/3, respectively. The final evaluation assigned 30, three, and 16 of the 49 traps to Classes I, II, and III (Figure 3b); the 33 unlabeled traps included 22, one, and 10 in these classes. The Class I relational grades γ_I and final assigned classes are given in Supplementary Table S2. The results reflect similarity to the geological attribute patterns represented by the labeled reference classes.

4. Quantitative Trap Evaluation Using a Support Vector Machine

4.1. Data Preprocessing and Feature Selection

The study comprised 49 structural traps, including 16 labeled traps (nine Class I, four Class II, and three Class III) and 33 unlabeled traps. Eleven original geological variables were considered as candidate predictors. Continuous variables were standardized using statistics estimated from the corresponding training partition. Trap type, reservoir lithology, and reservoir depositional facies were treated as nominal variables and represented by one-

hot encoding; categories not observed during fitting were handled as unknown at prediction time. Fault-damage degree was treated as an ordinal variable (weak = 0, moderate = 1, and strong = 2) and standardized within the training pipeline. The three observed class labels were encoded as 0, 1, and 2. Unlabeled traps were excluded from supervised model fitting. The Z-score transformation is defined as follows:

$$x^* = \frac{x - \mu}{\sigma}, \quad (1)$$

where x^* is the standardized value, x is the original feature value, μ is the mean estimated from the training data, and σ is its standard deviation.

The 16 labeled traps were classified into three exploration-outcome categories based on available drilling and testing information and integrated geological interpretations (Supplementary Table S7). Class I represents traps with documented effective oil or gas productivity; Class II includes traps with hydrocarbon indications but without confirmed effective hydrocarbon accumulation or sustained productivity; and Class III represents dry holes or traps interpreted to have lost effective hydrocarbon preservation. The classification integrates drilling outcomes and geological evidence rather than relying on a single numerical production threshold.

Reservoir thickness, lithology, and depositional facies were consistently compiled for the Lower Cretaceous Qingshuihe Formation (K_1q) across all 49 traps to maintain a common stratigraphic reference for the candidate predictors.

Given the limited labeled dataset (n = 16) and 11 candidate geological variables, feature selection was implemented using grouped analysis of variance (ANOVA) within the SVM pipeline. The F-statistic was computed from the training data after standardization and one-hot encoding; each original geological variable received the maximum F-statistic of its encoded columns. Selection of a nominal variable retained the entire associated one-hot group. The number of original geological-variable groups retained (k) was optimized within inner cross-validation. In the final model fitted to all 16 labeled traps, six groups were selected: burial depth, Jurassic source-rock thickness, Jurassic source-rock maturity, Jurassic gas-generation intensity, reservoir thickness, and fault-damage degree.

Feature selection and SVM parameter optimization were repeated within each outer training fold during strict nested LOOCV. The final model was then fitted using all 16 labeled traps with the six geological variables selected by the full-sample optimization.

4.2. Model Construction and Hyperparameter Optimization

A support vector machine (SVM) [33] was employed for three-class structural-trap evaluation. The support vector classifier (SVC) implemented in scikit-learn handles multi-class classification using a one-versus-one strategy. Both linear and radial basis function (RBF) kernels were considered to examine different decision-boundary structures.

Model optimization covered the regularization parameter $C \in \{0.01, 0.1, 1, 10, 100\}$ and the number of retained original geological-variable groups $k \in \{3, 4, 5, 6, 7\}$. For the RBF kernel, the kernel coefficient γ was searched over {scale, auto, 0.01, 0.1, 1}. After excluding redundant γ combinations for the linear kernel, the search space contained 150 distinct configurations (Table 2). Feature selection and hyperparameter optimization were performed within the SVM modeling pipeline (Figure 4) using leave-one-out cross-validation (LOOCV) (Figure 5).

Table 2. Hyperparameter search space for grouped geological-variable selection and SVM optimization (150 nonredundant configurations).

Hyperparameter	Search Range	Description
Number of features k	{3, 4, 5, 6, 7}	Number of original geological-variable groups retained by grouped ANOVA
Kernel function	{linear, rbf}	Linear or radial basis function
C	{0.01, 0.1, 1, 10, 100}	Regularization coefficient
γ	{'scale', 'auto', 0.01, 0.1, 1}	RBF kernel only (not applied to the linear kernel)
Evaluation metric	Accuracy	Inner-LOOCV model-selection criterion

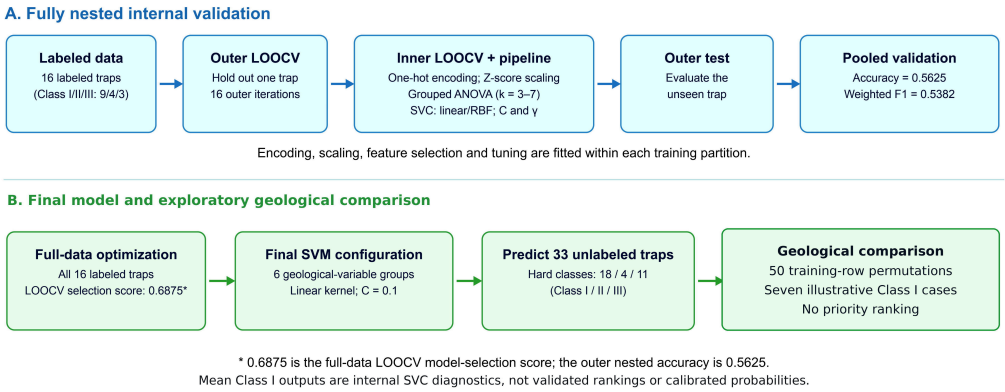


Figure 4. SVM workflow showing strict nested LOOCV, full-sample model selection, exploratory hard classification, and non-ranked geological comparison. The full-sample model-selection score is distinct from the nested-validation estimate.

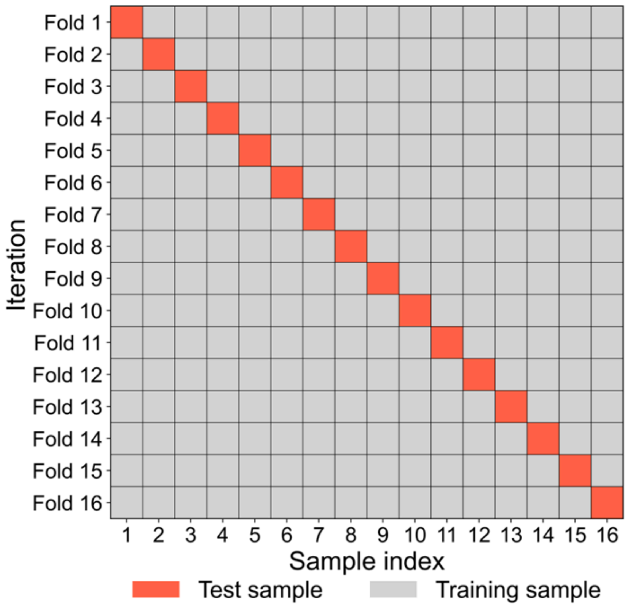


Figure 5. Schematic diagram of leave-one-out cross-validation.

All modeling experiments were conducted in Python 3.11.15 using scikit-learn 1.8.0. Class imbalance was handled by setting class_weight = "balanced" in SVC.

4.3. Model Performance Evaluation

Model performance was evaluated using strict nested leave-one-out cross-validation (LOOCV) to account for feature selection and hyperparameter optimization. The procedure

consisted of 16 outer iterations, each holding out one labeled trap for evaluation. Within each outer training set of 15 traps, preprocessing, grouped ANOVA feature selection, and SVM hyperparameter optimization were performed using an inner LOOCV search over 150 candidate configurations. Accuracy was used as the inner model-selection criterion, and the selected model generated a hard class prediction for the held-out trap. Predictions from all 16 outer iterations were pooled to calculate accuracy, weighted F1-score, macro F1-score, and balanced accuracy.

The nested-validation procedure was distinguished from the final model-selection process. After evaluating the complete modeling workflow, all 16 labeled traps were used to select the final SVM configuration for exploratory classification of the 33 unlabeled traps. The cross-validation score obtained during this full-sample optimization was treated as a model-selection criterion rather than an independent estimate of predictive performance.

Under strict nested LOOCV, nine of the 16 labeled traps were assigned to the correct class (accuracy = 0.5625; weighted F1 = 0.5382; macro F1 = 0.4078; balanced accuracy = 0.4815). Class I precision and recall were 0.875 and 0.778, respectively, while none of the four Class II traps were correctly identified (Table 3). The results illustrate the difficulty of separating the three classes with the currently available labeled samples.

Table 3. Class-wise SVM performance from pooled predictions of strict nested LOOCV (16 labeled traps; 9/4/3 by class).

	Precision	Recall	F1-Score	Number of Samples
Class I	0.875	0.778	0.824	9
Class II	0.000	0.000	0.000	4
Class III	0.286	0.667	0.400	3
Overall	Accuracy = 0.5625	Weighted F1 = 0.5382	Macro F1 = 0.4078	16

The full-sample LOOCV optimization identified eight SVM configurations with the same highest classification accuracy of 0.6875 (11/16). The configuration retaining six geological-variable groups with a linear kernel and $C = 0.1$ was selected as the first tied configuration in the predefined parameter-grid order, not because it achieved a higher cross-validation score. The eight configurations produced five distinct classification vectors for the 33 unlabeled traps. Six traps (TR16, TR28, TR31, TR32, TR36, and TR46) changed class across these configurations, whereas the remaining 27 retained consistent assignments (Supplementary Tables S9 and S10). The full-sample selection accuracy should not be interpreted as the strict nested-validation performance.

Using the selected final configuration, the SVM assigned 18 of the 33 unlabeled traps to Class I, four to Class II, and 11 to Class III (Table 4). As an additional descriptive sensitivity analysis, the final SVM was fitted with SVC(probability=True) across 50 permutations of the 16 training-row positions, without changing the samples, labels, predictors, or model parameters. The resulting mean Class I outputs are reported in Table 4 for reference. These outputs primarily reflect sensitivity to the internal probability-estimation procedure rather than geological or sampling uncertainty and are not used to establish exploration-priority rankings. Table 5 presents seven illustrative Class I traps for geological comparison, listed by trap ID rather than predicted score.

Table 4. Hard SVM classifications obtained from the first tied full-sample configuration, descriptive mean Class I outputs across 50 training-row permutations, and selected-feature range information for the 33 unlabeled traps. * The mean outputs are internal SVC diagnostics, not calibrated probabilities, confidence intervals, or exploration-priority rankings; the last column gives the number of selected variables outside the labeled-sample range. Six trap classifications vary under alternative equally scoring parameter configurations (Table S10).

Trap ID	SVM Hard Class	Mean Class I Output *	Out-of-Range Features
TR2	I	0.6585	0
TR3	I	0.6685	0
TR5	I	0.6521	0
TR6	II	0.5454	0
TR7	I	0.5890	0
TR9	I	0.6545	0
TR10	I	0.6862	1
TR11	I	0.6732	0
TR13	I	0.6982	0
TR14	I	0.6781	0
TR16	I	0.5896	0
TR17	I	0.6962	1
TR18	I	0.6941	0
TR21	I	0.6961	0
TR22	I	0.6845	1
TR23	I	0.6947	1
TR24	I	0.7026	2
TR26	II	0.4385	0
TR28	III	0.4640	0
TR31	III	0.5507	1
TR32	II	0.5295	0
TR33	I	0.7030	1
TR34	I	0.6304	0
TR36	II	0.5642	1
TR37	III	0.3675	4
TR38	III	0.4486	0
TR42	III	0.3335	0
TR43	III	0.3170	0
TR44	III	0.3384	0
TR45	III	0.5042	0
TR46	III	0.4539	1
TR47	III	0.2816	4
TR48	III	0.3051	3

Table 5. Seven illustrative SVM-assigned Class I traps for geological comparison, listed by trap ID rather than score or exploration priority. All seven remain Class I under the eight equally scoring full-sample configurations (Table S9), but this does not establish reliable prospect ranking or robustness to label uncertainty. † Number of selected geological variables outside the labeled-sample range.

Trap ID	SVM Hard Class	Out-of-Range Variables †	Illustrative Geological Comparison
TR13	I	0	Cross-method Class I agreement
TR17	I	1	Class I agreement; out-of-range input
TR18	I	0	Favorable source rock; thin reservoir
TR21	I	0	Source-rock and preservation agreement
TR23	I	1	Source and preservation; out-of-range input
TR24	I	2	Favorable source rock; two out-of-range inputs
TR33	I	1	Contrast with geological risk score

To examine the influence of geological setting and model complexity, the complete SVM workflow was further evaluated using nested leave-one-structural-belt-out cross-

validation (LOGO). In each outer iteration, all labeled traps from one structural belt were held out, while preprocessing, feature selection, and hyperparameter optimization were performed using the remaining training traps. Several predefined single- and two-variable SVM models were also evaluated under the same validation schemes using a 30-configuration linear/RBF kernel search, training-fold standardization, and balanced class weights. A majority-class classifier was included as a simple reference baseline. The resulting nested-LOOCV and nested-LOGO accuracies are compared in Table 6.

Table 6. Performance of the full multivariable SVM and predefined simpler geological-variable SVM baselines under nested LOOCV and nested LOGO; the majority-class predictor is included as a non-SVM reference. Results are calculated from pooled held-out predictions.

Model	LOOCV Accuracy	LOGO Accuracy	LOGO Weighted F1	LOGO Balanced Accuracy
Full multivariable SVM	0.5625	0.4375	0.4032	0.4074
Burial depth	0.5625	0.3750	0.4256	0.3426
Jurassic source-rock thickness	0.6875	0.6250	0.5655	0.5926
Jurassic source-rock maturity	0.5000	0.3125	0.2898	0.4074
Burial depth + source-rock thickness	0.6250	0.5625	0.4993	0.5556
Maturity + gas-generation intensity	0.6250	0.5625	0.5241	0.5556
Majority-class (always Class I)	0.5625	0.5625	0.4050	0.3333

Under strict nested LOGO, the complete multivariable SVM correctly classified seven of 16 labeled traps (accuracy = 0.4375), compared with nine of 16 under nested LOOCV (0.5625). The source-rock-thickness-only SVM achieved accuracies of 0.6875 under nested LOOCV and 0.6250 under nested LOGO, respectively, under the same two validation schemes. With labeled traps available from only seven structural belts and uneven class distributions among them, these results highlight the importance of additional labeled examples from different geological settings.

The simple-variable baselines in Table 6 provide a geological reference for the multivariable SVM. Their results indicate that Jurassic source-rock thickness carries a substantial part of the discriminative signal in the currently labeled dataset.

A separate sensitivity assessment then examined only the SVM classifier. The primary complete-workflow scheme reselected geological-variable groups and optimized the SVM parameters within each outer training partition. In the fixed-feature scheme, the six variables identified using all 16 labeled traps were held constant, but standardization, SVM fitting, and kernel/hyperparameter optimization were still conducted within the corresponding training partitions. Table 7 contrasts the resulting pooled outer-fold predictions.

Table 7. Sensitivity of SVM validation results to the feature-selection procedure: strict nested LOOCV and a fixed six-feature assessment.

SVM Validation Scheme	Pooled Accuracy	Pooled Weighted F1	Pooled Macro F1	Balanced Accuracy
Strict nested LOOCV (fold-wise feature selection)	0.5625	0.5382	0.4078	0.4815
Fixed globally selected six features (inner SVM tuning)	0.6875	0.6998	0.6317	0.6481

Table 7 shows an accuracy of 0.5625 and weighted F1 of 0.5382 for strict nested LOOCV, compared with 0.6875 and 0.6998 for the fixed-six-feature assessment. The fixed subset was

selected from all 16 labeled traps before evaluation; therefore, this second scheme is used to examine sensitivity to the feature-selection procedure, while strict nested LOOCV provides the main internal performance estimate.

To assess the sensitivity of the modeling workflow to individual geological inputs and class labels, controlled one-at-a-time perturbation experiments were conducted by repeating feature selection, hyperparameter optimization, and nested validation (Supplementary Table S6). These hypothetical scenarios affected pooled classification performance and, in some cases, the hard class assignments of unlabeled traps. Additional sensitivity analyses examined alternative interpretations of a historical label (Supplementary Table S8). These experiments characterize dependence on individual records within the available small dataset rather than providing calibrated uncertainty intervals.

4.4. Feature-Selection Stability Under the Extremely Small-Sample Condition

For the SVM feature-selection stability analysis, the original geological-variable groups retained in each of the 16 outer nested-LOOCV training folds were recorded. Selection frequency was defined as the number of folds in which a group was retained (Table 8 and Figure 6). Pairwise Jaccard similarities were calculated from the 16 selected subsets, giving 120 unique fold pairs (Figure 7).

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}, \tag{2}$$

where A and B represent the sets of geological features selected in two different outer validation folds.

Table 8. Selection frequencies of the 11 original geological-variable groups across 16 nested LOOCV outer folds.

Feature	Selected in Outer Folds	Selection Frequency
Jurassic source-rock thickness	14/16	0.875
Jurassic source-rock maturity	11/16	0.688
Jurassic gas-generation intensity	15/16	0.938
Burial depth	16/16	1.000
Reservoir thickness	11/16	0.688
Fault-damage degree	8/16	0.500
Trap area	0/16	0.000
Closure height	0/16	0.000
Reservoir depositional facies	0/16	0.000
Reservoir lithology	0/16	0.000
Trap type	0/16	0.000

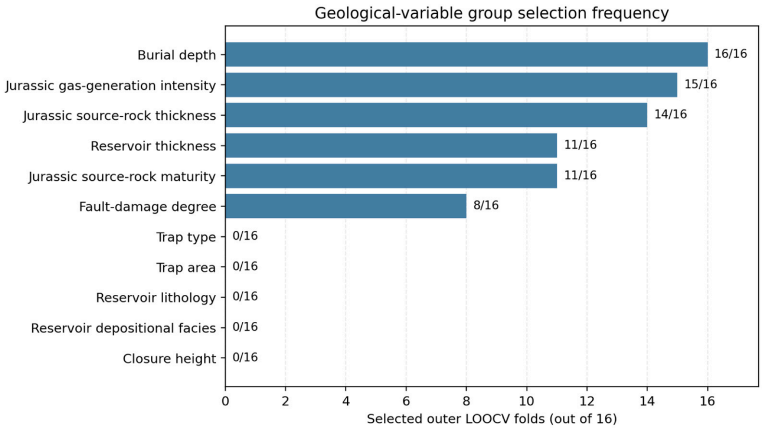


Figure 6. Selection frequencies of 11 original geological-variable groups across 16 outer folds of strict nested LOOCV.

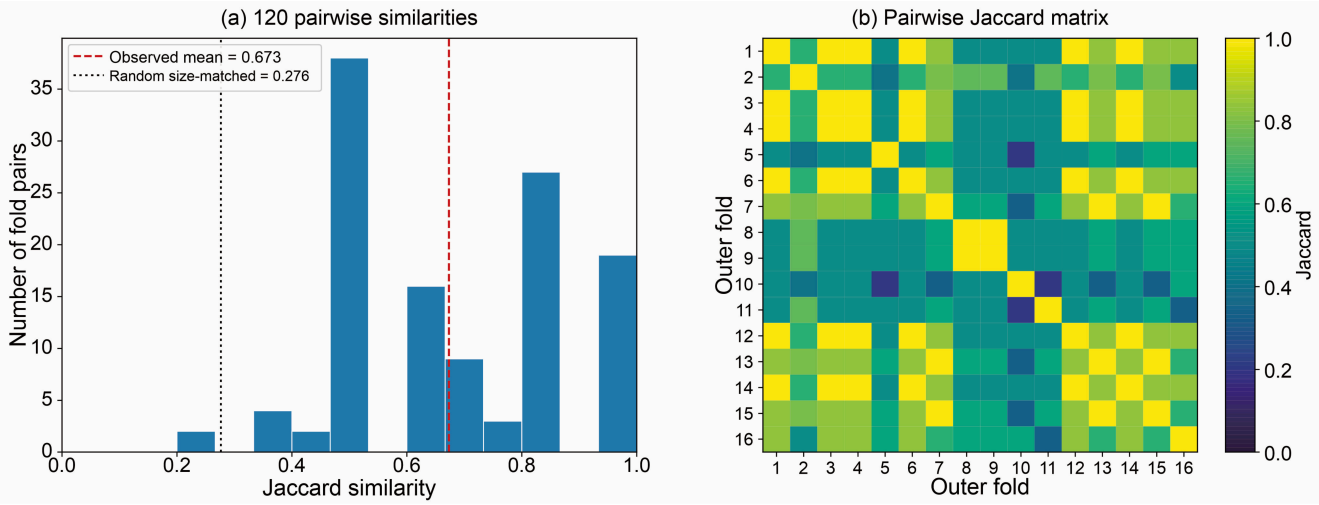


Figure 7. Pairwise Jaccard similarity of the original geological-variable groups selected in 16 strict nested-LOOCV outer folds: (**a**) distribution of 120 pairwise similarities and (**b**) corresponding 16 × 16 similarity matrix. Dashed and dotted lines indicate the observed mean and size-matched random-set reference, respectively.

Burial depth was selected in all 16 outer folds, Jurassic gas-generation intensity in 15, Jurassic source-rock thickness in 14, Jurassic source-rock maturity and reservoir thickness in 11 each, and fault-damage degree in eight (Table 8). Trap area, closure height, trap type, reservoir lithology, and depositional facies were not selected in these outer folds. The frequently selected variables are consistent with the geological importance of burial and source-rock conditions in the southern Junggar Basin.

The mean Jaccard similarity across 120 feature-set pairs was 0.6732 (standard deviation 0.2040; median 0.6000; range 0.2000–1.0000; Figure 7). A size-matched random-set reference based on the 11 original candidate variables yielded an expected similarity of 0.2764. The higher observed similarity indicates a recurring core of selected geological variables, particularly burial depth and source-rock characteristics, although the complete feature subset varied among folds.

Ties were systematic rather than occasional in the inner-LOOCV searches: all 16 outer folds had between two and 26 parameter configurations with equal highest inner-validation accuracy (Supplementary Table S11). The implemented selection rule chose the first configuration in the specified grid order when scores tied. Thus, neither the fold-wise selected hyperparameters nor the final unlabeled-trap classifications should be interpreted as uniquely determined by the available labeled samples.

4.5. Feature and Hyperparameter Re-Optimization with Increasing Labeled-Sample Size

To examine how the SVM configuration evolves with new labeled information, retrospective cumulative training sets of eight, 12, and 16 traps were assembled from the drilling and evaluation chronology in Supplementary Table S1. The initial eight traps were TR29, TR35, TR49, TR8, TR25, TR27, TR39, and TR40. Four subsequently documented traps (TR41, TR1, TR15, and TR30) were added at the 12-sample stage, followed by TR19, TR20, TR4, and TR12 at the 16-sample stage. The corresponding Class I/II/III counts were 3/2/3, 5/4/3, and 9/4/3, respectively.

At each stage, grouped ANOVA variable selection and SVM hyperparameters were re-optimized by a 150-configuration LOOCV search (Table 9). The within-stage model-selection accuracy for the initial eight-trap subset reached 1.000; however, this result was obtained on only eight observations during hyperparameter selection and provides no reliable evidence of out-of-sample predictive skill. The preferred kernel changed from RBF for

eight and twelve traps to linear for sixteen, and the selected number of variables increased from three to six. These changes reflect sample-composition and model-selection sensitivity, not an established learning curve or demonstrated improvement in generalization.

Table 9. SVM feature selection and hyperparameter optimization across cumulative retrospective chronological training subsets of 8, 12, and 16 labeled traps.

Labeled Samples	k (Groups)	Kernel	C	γ	LOOCV Selection Accuracy *
8	3	RBF	0.1	1.0	1.0000
12	3	RBF	0.1	0.01	0.6667
16	6	Linear	0.1	N/A	0.6875

* The reported LOOCV values are within-stage model-selection accuracies, not strict nested-validation estimates. The three subsets contain 3/2/3, 5/4/3, and 9/4/3 Class I/II/III traps, respectively. N/A denotes not applicable.

In a retrospective next-batch check, the eight-trap model correctly classified one of the next four traps, whereas the 12-trap model correctly classified all four of the following traps. Crucially, all four traps in the latter batch were Class I, so a trivial majority-class predictor assigning Class I to every trap would also achieve 4/4. These sequential checks therefore do not establish a predictive advantage for the SVM. Periodic re-optimization as independently verified labels accumulate is a methodological recommendation to be tested with more representative forward or external validation, not an empirical performance improvement demonstrated by these checks.

5. Comparison and Geological Interpretation of Trap-Evaluation Results

Four geological evaluation approaches were compared for the 33 unlabeled traps, focusing on (1) categorical agreement among SVM, fuzzy comprehensive evaluation, and gray relational analysis; (2) differences in geological formulations, including the separate relative score provided by geological risk probability; and (3) geological interpretation of concordant and divergent assessments. The continuous geological risk score P is not equivalent to a Class I–III assignment. Likewise, SVM internal Class I outputs, fuzzy index D, and gray relational grade γI are different, method-specific quantities and cannot be treated as directly comparable hydrocarbon-occurrence probabilities.

5.1. Overview of Evaluation Results

The final SVM assigned Class I to 18 of the 33 unlabeled traps, compared with 23 and 22 Class I assignments by fuzzy comprehensive evaluation and gray relational analysis, respectively. As summarized in Table 10, SVM and fuzzy evaluation agreed on 25 of the 33 class assignments, SVM and gray analysis on 26, and fuzzy and gray evaluation on 30; all three classifications agreed for 25 traps. The seven illustrative traps in Table 5 are all assigned to Class I by these three methods, but their agreement does not establish their relative priority or independently confirm hydrocarbon accumulation. Geological risk scores remain available for all 49 traps in Supplementary Table S2 and are discussed through specific geological contrasts rather than a unified exploration-priority ranking.

Table 10. Categorical agreement among SVM, fuzzy comprehensive evaluation, and gray relational analysis for the 33 unlabeled traps; comparisons are based on hard class assignments rather than score rankings.

Categorical Comparison (33 Unlabeled Traps)	Agreements	Disagreements	Agreement (%)
SVM vs. fuzzy comprehensive evaluation	25	8	75.8%
SVM vs. gray relational analysis	26	7	78.8%
Fuzzy comprehensive evaluation vs. gray relational analysis	30	3	90.9%
All three methods (same assigned class)	25	8	75.8%

Note: Each pairwise row reports the number of traps assigned to the same (or different) Class I–III category. The final row counts traps for which all three methods agree. Agreement is descriptive and is not independent evidence of hydrocarbon occurrence. Geological risk probability is excluded because its continuous *P* score does not assign outcome classes; complete method-specific scores are available in Supplementary Table S2.

5.2. Geological Analysis of Consensus Targets

A first type of agreement concerns favorable source-rock and preservation conditions. TR21 and TR23, both in the Wuchang structural belt, have Jurassic source-rock thicknesses of 380 and 400 m, maturities of approximately 2.0%, and gas-generation intensities of 75 and 85 $\times$ 10^8 m^3/km^2, respectively. Both have weak recorded fault damage and reservoirs approximately 50 and 25 m thick, respectively. They are assigned to Class I by the SVM, fuzzy evaluation, and gray relational analysis, while their geological risk *P* scores are also relatively favorable. TR24 in the same structural belt has similar source-rock attributes (390 m thickness, approximately 2.0% maturity, and gas-generation intensity of 80 $\times$ 10^8 m^3/km^2) and weak fault damage; it is likewise assigned to Class I by all three methods. However, its smaller trap area (17.66 km^2) contributes to a lower multiplicative geological risk score than those of TR21 and TR23. These cases illustrate common geological interpretations despite differences in assessment mechanisms, not a validated order of exploration priority.

5.3. Geological Interpretation of Inter-Method Discrepancies

TR18 illustrates how the same trap can receive a favorable class assignment but a lower relative score under a different aggregation rule. It has approximately 360 m of Jurassic source rock, maturity near 2.0%, gas-generation intensity of 100 $\times$ 10^8 m^3/km^2, and weak fault damage, but its reservoir is only 20 m thick. It was classified as Class I by the SVM, fuzzy evaluation, and gray relational analysis, while its geological risk score (p = 0.051295) placed it in the intermediate relative group. This difference is consistent with the direct multiplicative contribution of reservoir thickness in the geological risk formulation.

TR34 and TR36 provide a contrasting example. They have relatively high geological risk *P* scores (0.686888 and 0.374504) among unlabeled traps, supported in part by their large trap areas (242.14 and 160.63 km^2) and reservoir thicknesses (45 and 40 m). Both were assigned to Class I by fuzzy and gray methods, whereas the SVM assigned TR34 to Class I and TR36 to Class II. The geological risk formulation incorporates trap area, whereas the six-variable SVM, fuzzy, and gray methods instead include burial depth. This difference helps explain the contrasting interpretations without implying that the risk-score order is an independently validated prospect ranking.

A further contrast is provided by TR2, TR3, and TR5 in the Hutubi structural belt. All three are assigned to Class I by the SVM, fuzzy evaluation, and gray relational analysis. Their relatively small trap areas (9.82, 5.15, and 24.40 km^2, respectively) limit their multi-

plicative *P* scores: TR2 and TR5 fall in the intermediate relative band and TR3 in the low band. Gray relational grades emphasize closeness to the mean geological pattern of labeled Class I traps; therefore, a favorable relational assessment need not imply a high score under the source–reservoir–trap–preservation product.

TR33 is assigned to Class I by the SVM, fuzzy evaluation, and gray relational analysis. Its reservoir thickness is 20 m, and its geological risk score (P = 0.044319) falls in the intermediate relative group. This contrast illustrates differences between geological-class assignment and the multiplicative effects of reservoir-related factors in the geological risk score, without supporting a definitive exploration-priority ranking.

5.4. Implications for Geological Screening

The nested-validation results do not support using the multivariable SVM as an independent prospect-prioritization tool. Its LOOCV accuracy matched the majority-class baseline (0.5625), its LOGO accuracy was lower (0.4375 versus 0.5625), and a source-rock-thickness-only SVM performed better under both schemes. Moreover, the model failed to correctly classify any of the four Class II traps. Accordingly, Table 5 presents seven illustrative examples in ascending trap-ID order rather than as an exploration-priority ranking. Four of these seven cases have at least one selected variable outside the labeled-sample range, further limiting any generalization of their inferred class assignments. TR21 and TR23 illustrate source-rock and preservation agreement across evaluation approaches; TR18 combines favorable source conditions with a thin reservoir; and TR33 highlights differences between class assignment and multiplicative geological risk scores.

The 18 SVM Class I assignments are exploratory classifications rather than independently verified discoveries or stable exploration priorities. Although descriptive Class I outputs are preserved in Table 4 to make the numerical analysis transparent, their means across row-order permutations must not be interpreted as calibrated hydrocarbon-occurrence probabilities, independent resampling evidence, uncertainty intervals, or a defensible seven-trap ranking. Cross-method agreement can motivate additional geological examination, but shared input information prevents it from serving as independent confirmation. Prospect prioritization requires further geological, drilling, and testing evidence.

5.5. Overall Comparison of Evaluation Methods

The four evaluation methods provide complementary perspectives on geological favorability because they differ in their mathematical formulations, treatment of geological indicators, and interpretation of evaluation outputs. Geological risk probability combines source-rock, reservoir, trap, and preservation conditions multiplicatively, making the composite score sensitive to unfavorable individual factors. Fuzzy comprehensive evaluation integrates indicator-specific class memberships through weighted aggregation, whereas gray relational analysis emphasizes the overall similarity between individual traps and class-specific reference patterns. In contrast, the SVM learns classification boundaries from the available labeled geological examples, rather than relying on a prescribed geological aggregation formula.

The comparison of the 33 unlabeled traps indicates substantial descriptive agreement among the three class-assignment methods (Table 10). However, different evaluation formulations can emphasize source-rock, reservoir, burial, and fault-related preservation conditions differently, even for traps receiving the same broad class. The geological risk probability method is discussed through its continuous relative P score rather than class agreement. The results therefore support interpretation of geological similarities and differences, not a validated prospect ranking.

These results do not establish the superiority of the SVM over conventional geological evaluation methods. The relatively small labeled dataset and geological dependence among traps limit the reliability of the learned classification boundaries, while the conventional approaches remain sensitive to their predefined weights, membership functions, and reference patterns. Moreover, agreement among the methods should not be treated as independent validation of hydrocarbon accumulation because some geological inputs and reference information are shared. The principal value of the integrated comparison is therefore to identify geological patterns supported by different evaluation mechanisms, highlight discrepant assessments requiring further investigation, and provide complementary evidence for preliminary trap screening. Independent drilling and geological verification remain necessary before the relative reliability of these approaches can be established.

6. Model Limitations and Transferability

6.1. Model Limitations

The evaluation is based on an extremely small labeled dataset and should be considered in the context of its sampling characteristics and geological uncertainty.

The predictor selection and unlabeled classifications are sensitive to individual training records and to the choice among equally scoring SVM configurations. The controlled perturbations changed nested-validation performance and predictions, including hypothetical changes to individual training labels (Supplementary Table S6); alternative treatment of a historically interpreted label also changed classifications (Table S8). Across the eight full-sample tied models, six of 33 unlabeled traps received different hard classes (Tables S9 and S10). These diagnostics demonstrate limited robustness, not calibrated uncertainty intervals. Additional independent observations are required to resolve this instability.

(1) The 16 labeled traps are unevenly distributed among structural belts and outcome classes. This influences the evaluation of less-represented classes, particularly Class II, and motivates continued acquisition of labeled examples from additional structural settings.

(2) Although the final SVM retained six geological variables, the feature subsets changed among nested training folds. Other geological controls may become important as the number and diversity of labeled traps increase.

(3) Drilled targets were not selected randomly, so the labeled dataset reflects prior exploration priorities. Some historical exploration-outcome labels also incorporate geological preservation interpretations that may overlap with fault-damage degree as a model predictor, potentially introducing dependence between the labels and predictor information. Newly evaluated traps and more balanced drilling outcomes would help characterize the range of geological conditions represented in each class.

(4) The retrospective 8/12/16-trap experiment demonstrates sensitivity of selected variables and model configuration to sample composition, not improved predictive skill with additional data. In particular, the eight-sample model-selection score of 1.000 is not an independent validation estimate, and the 4/4 subsequent check equaled the trivial all-Class-I predictor. Re-optimizing the model as verified labels accumulate is therefore recommended for future applications, together with independent evaluation.

(5) The K_1q reservoir thickness, lithology, and depositional facies were compiled as common reference-horizon predictors across the trap inventory. For traps whose documented hydrocarbon-producing intervals differ from K_1q, including examples in the Huo-Ma-Tu, Qigu, and Aika structural belts, these K_1q reservoir attributes do not describe the actual producing interval. This stratigraphic mismatch can weaken the geological relevance and transferability of the reservoir predictors. Separately, the descriptive SVC Class I outputs vary with training-row order, and several unlabeled traps fall outside

the labeled-variable ranges; neither these outputs nor cross-method agreement provide validated exploration priorities.

The conventional reference methods use geological weighting or class-reference rules that differ from the SVM decision mechanism. However, the six variables used in fuzzy comprehensive evaluation and gray relational analysis were selected using the complete labeled dataset. Their fixed-feature LOOCV results should therefore be interpreted as conditional internal comparisons rather than unbiased estimates of the complete feature-selection workflow. Agreement among the methods provides a basis for comparative geological interpretation but does not independently establish predictive reliability.

The stability of model-based classifications and their geological interpretation remain an open question requiring independent exploration evidence.

6.2. Model Transferability and Applicability Conditions

The applicability of the proposed SVM is constrained by the limited number and uneven geological distribution of labeled samples. The present dataset contains only 16 labeled traps from seven structural belts, including nine Class I, four Class II, and three Class III traps. The complete multivariable SVM achieved an accuracy of 0.5625 under strict nested LOOCV, compared with 0.4375 under nested leave-one-structural-belt-out validation. This difference indicates limited transferability across geological settings.

Furthermore, 12 of the 33 unlabeled traps have at least one selected geological variable outside the range represented by the labeled training data. Such out-of-range values provide a practical warning of potential extrapolation, although they do not constitute a formal statistical test of data drift. For application to additional traps or other regions, differences in geological-variable ranges, previously unrepresented categorical values, and changes in class composition should be examined before transferring the model. Additional representative labels and independent regional validation would be required, followed by renewed feature selection and hyperparameter optimization where appropriate. The current workflow should therefore be regarded as a preliminary geological screening approach rather than a directly transferable predictive model.

Future work will examine the applicability of this workflow to other structural belts and sedimentary basins with contrasting petroleum-system characteristics.

7. Conclusions

(1) An SVM-based quantitative evaluation workflow was developed for 49 complex structural traps in the southern Junggar Basin, including 16 labeled traps (nine Class I, four Class II, and three Class III). The representative class-balanced linear SVM (C = 0.1) retained six geological variables and classified the 33 unlabeled traps into 18 Class I, four Class II, and 11 Class III traps. This configuration was the first of eight equally scoring full-sample optima, so the classification counts depend partly on parameter tie resolution.

(2) Strict nested LOOCV yielded an accuracy of 0.5625 and a weighted F1-score of 0.5382, matching the majority-class baseline accuracy. Under nested LOGO, accuracy fell to 0.4375, below the majority-class baseline of 0.5625. A source-rock-thickness-only SVM performed better under both schemes (0.6875 and 0.6250), and the multivariable SVM correctly identified none of the four Class II traps. Thus, the available data do not demonstrate a predictive advantage for the multivariable model.

(3) Burial depth, Jurassic gas-generation intensity, and source-rock thickness occurred most frequently in nested feature selection (mean pairwise Jaccard similarity = 0.6732). The retrospective 8/12/16-label sequence showed that model configurations depended on the available training subset, but its within-stage scores and four-trap forward checks did not demonstrate an improvement over simple baselines. Future use should therefore

include periodic re-optimization when independently verified new labels become available, evaluated through more representative external or forward-time validation.

(4) The SVM, fuzzy comprehensive evaluation, and gray relational analysis showed categorical agreement on 25 or more of the 33 unlabeled traps in pairwise comparisons, but this agreement does not constitute independent validation because the approaches share geological inputs. Among eight tied full-sample SVM configurations, five distinct prediction vectors were obtained and six unlabeled-trap classes varied, further limiting the reliability of individual assignments.

(5) Seven SVM-assigned Class I traps were retained as illustrative cases of cross-method geological agreement and divergence, not as a score-ordered priority subset. Although these seven cases were consistently assigned to Class I by the eight tied full-sample configurations, this does not establish independence from training-label uncertainties or provide validated drilling recommendations. Additional labels and independent geological, drilling, or testing evidence are needed before robust exploration prioritization is possible.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/app162110669/s1. Table S1: Geological characteristics, drilling information, and class labels for the 49 structural traps; Table S2: Detailed geological risk, fuzzy comprehensive evaluation, and gray relational analysis results; Table S3: Geological-variable definitions and categorical encoding; Table S4: Class-specific LOOCV performance of fuzzy and gray evaluations; Table S5: LOOCV confusion matrices for fuzzy and gray evaluations; Table S6: Sensitivity to individual input perturbations; Table S7: Available exploration evidence for the 16 historical labels; Table S8: Sensitivity to alternative treatment of the TR29 label; Table S9: Equally scoring full-sample SVM configurations and their unlabeled class counts; Table S10: Unlabeled traps whose hard classes vary among tied configurations; Table S11: Counts of tied inner-LOOCV configurations in 16 outer validation folds.

Author Contributions: Conceptualization, G.C.; methodology, L.X.; software, W.W.; validation, X.L.; formal analysis, X.L.; investigation, C.J., W.W. and L.X.; resources, Z.P. and C.J.; data curation, L.X., W.W. and X.L.; writing—original draft preparation, L.X.; writing—review and editing, G.C. and L.X.; visualization, L.X.; supervision, G.C., Z.P. and C.J.; project administration, G.C. and Z.P.; funding acquisition, G.C. All authors have read and agreed to the published version of the manuscript.

Funding: Key Research and Development Project of Xinjiang Uygur Autonomous Region, Development of Evaluation Technology for Oil and Gas Accumulation in Ultra Deep Complex Fault Zones of Basins (2024B01015-3).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The geological data and exploration information supporting this study are provided in the Supplementary Materials. Additional original exploration records are subject to institutional and proprietary restrictions. The analysis code is planned for public release upon publication, subject to the necessary institutional authorization.

Conflicts of Interest: Author Wei Wang was employed by the company CNPC Geophysical Company Limited. Authors Zhichao Pang and Chunming Jia were employed by PetroChina Xinjiang Oilfield Company. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

SVM	Support Vector Machine
LOOCV	Leave-One-Out Cross-Validation

RBF	Radial Basis Function
LOGO	Leave-One-Group-Out Cross-Validation
ANOVA	Analysis of Variance
SVC	Support Vector Classifier

References

1. Luo, Q.; Wang, Q.; He, X.; Shang, F.; Zhang, H.; Wen, F.; Qiu, Z. Nontraditional Hydrocarbon Resources: New Physical Domains with Great Potential for Petroleum Exploration and Development. *Nat. Gas Explor. Dev.* **2024**, *47*, 1–11. [CrossRef]
2. Hou, L.; Yang, F.; Yang, C.; Yang, H.; Li, Y. Principles and Methods for Quantitatively Evaluating the Effectiveness of Conventional Petroleum Zones and Traps. *Acta Pet. Sin.* **2021**, *42*, 1126–1141. [CrossRef]
3. Liu, G. *Petroleum Geology*, 4th ed.; Petroleum Industry Press: Beijing, China, 2008.
4. Cai, X.; Hou, Q. Geological Evaluation Method of Hydrocarbon Traps and Its Application. *Pet. Geol. Oilfield Dev. Daqing* **1994**, 1–5. [CrossRef]
5. Yang, C.; Wang, M.; Nie, X. The Advantage of Comprehensive Evaluation Method by Blurred Mathematics in Trap Evaluation. *J. Southwest Pet. Inst.* **1999**, *21*, 41–42.
6. Zhang, L.; Zhang, J. Application of Grey Association Analysis in Trap Evaluation. *Pet. Geol. Oilfield Dev. Daqing* **1995**, *14*, 12–19.
7. Dong, Q.; Zhou, Y.; Jin, X. Artificial Neural Network Algorithm of Geological Trap Appraisal. *Fault-Block Oil Gas Field* **1998**, *5*, 14–15.
8. Hao, F.; Zou, H.; Jiang, J. Dynamics of Petroleum Accumulation and Its Advances. *Earth Sci. Front.* **2000**, *7*, 11–21.
9. Chen, Z.; Osadetz, K.G.; Embry, A.F.; Hannigan, P.K. Hydrocarbon favourability mapping using fuzzy integration: Western Sverdrup Basin, Canada, Canadian Arctic Archipelago. *Bull. Can. Pet. Geol.* **2002**, *50*, 492–506. [CrossRef]
10. Chen, Z.; Osadetz, K.G. Geological Risk Mapping and Prospect Evaluation Using Multivariate and Bayesian Statistical Methods, Western Sverdrup Basin of Canada. *AAPG Bull.* **2006**, *90*, 859–872. [CrossRef]
11. Chen, Z.; Liu, Y.; Osadetz, K. Geological Risk Evaluation Using the Support Vector Machine with Examples from the Late Triassic-Early Jurassic Structural Play in Western Sverdrup Basin, Canadian Arctic Archipelago. *Bull. Can. Pet. Geol.* **2012**, *60*, 142–157. [CrossRef]
12. Zhao, J.; Ge, X.; Fan, Y.; Liu, J.; Chen, Y.; Xing, L. A Genetic Algorithm–Driven Support Vector Machine to Discriminate the Kerogen Type Using Conventional Geophysical Logging Data. *AAPG Bull.* **2023**, *107*, 1837–1849. [CrossRef]
13. Bai, Z.; Tan, M.; Shi, Y.; Guan, X.; Wu, H.; Huang, Y. Log Interpretation Method of Resistivity Low-Contrast Oil Pays in Chang 8 Tight Sandstone of Huanxian Area, Ordos Basin by Support Vector Machine. *Sci. Rep.* **2022**, *12*, 1046. [CrossRef]
14. Zhou, X.; Li, Y.; Song, X.; Jin, L.; Wang, X. Thin Reservoir Identification Based on Logging Interpretation by Using the Support Vector Machine Method. *Energies* **2023**, *16*, 1638. [CrossRef]
15. Sun, Y.; Qi, Y.; Chen, L.; He, Y.; Ji, H.; Shi, Y.; Zhao, B.; Liu, Z. Small-Data-Driven Machine Learning for Well Logging Characterization of Pore Structure in Low-Permeability Sandstone Reservoirs. *J. Palaeogeogr.* **2026**, *28*, 369–380. [CrossRef]
16. Xie, D.; Li, H.; Li, J.; Bai, P. Research of KPCA-SVM Predicting Model in Hydrocarbon Reservoir. In Proceedings of the 2025 5th International Conference on Artificial Intelligence and Industrial Technology Applications (AIITA), Xi'an, China, 28–30 March 2025; pp. 428–433.
17. Ambroise, C.; McLachlan, G.J. Selection Bias in Gene Extraction on the Basis of Microarray Gene-Expression Data. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 6562–6566. [CrossRef]
18. Varma, S.; Simon, R. Bias in Error Estimation When Using Cross-Validation for Model Selection. *BMC Bioinform.* **2006**, *7*, 91. [CrossRef]
19. Cawley, G.C.; Talbot, N.L.C. On Over-Fitting in Model Selection and Subsequent Selection Bias in Performance Evaluation. *J. Mach. Learn. Res.* **2010**, *11*, 2079–2107.
20. Nogueira, S.; Sechidis, K.; Brown, G. On the Stability of Feature Selection Algorithms. *J. Mach. Learn. Res.* **2018**, *18*, 1–54. [CrossRef]
21. Lei, D.; Zhang, J.; Chen, N.; Xiang, B. Conditions for Gas Pooling in the Lower Assemblage in the Southern Margin of the Junggar Basin and the Exploration Prospect of Large Hydrocarbon Fields. *Nat. Gas Ind.* **2012**, *32*, 16–22. [CrossRef]
22. Hu, H.; Zhang, Y.; Zhuo, Q.; Jia, C.; Guo, Z. Hydrocarbon Charging History of the Lower Petroleum System in the Southern Junggar Basin: Case Study of the Qigu Oil Field. *Nat. Gas Geosci.* **2019**, *30*, 456–467. [CrossRef]
23. Chen, Y.; Wang, X.; Fang, S.; Zhang, J.; Zhao, M.; Liu, S.; Bai, Z. Hydrodynamic Characteristics of Hydrocarbon Accumulation in E1-2z of Huo-Ma-Tu Structural Belt, Southern Junggar Basin, NW China. *J. China Univ. Pet. (Ed. Nat. Sci.)* **2013**, *37*, 30–36. [CrossRef]
24. Du, J.; Zhi, D.; Li, J.; Yang, D.; Tang, Y.; Qi, X.; Xiao, L.; Wei, L. Major Breakthrough of Well Gaotan 1 and Exploration Prospects of Lower Assemblage in Southern Margin of Junggar Basin, NW China. *Pet. Explor. Dev.* **2019**, *46*, 216–227. [CrossRef]

25. Pang, Z.; Li, J.; Wang, J.; Gao, Y.; Wei, L.; Zhang, D. Major Discovery in Well QB1 and Exploration Prospects of the Lower Assemblage in Dongwan Structural Belt, Southern Margin of Junggar Basin. *Acta Pet. Sin.* **2025**, *46*, 2245–2258. [CrossRef]
26. Ye, H.; Zhu, F.; Wang, G.; Shi, W.; Kang, X.; Dong, G.; Na, Z.; Wang, R. Paleogeomorphy Restoration of Permian-Jurassic and Its Hydrocarbon Implications in Junggar Basin, NW China. *Lithol. Reserv.* **2025**, *37*, 122–132. [CrossRef]
27. Zhu, M.; Wang, X.; Xiao, L. Structural Characteristics and Evolution in the Southern Margin of Junggar Basin. *Xinjiang Pet. Geol.* **2020**, *41*, 9–17.
28. Guo, W.; Jiang, Y.; Bian, B.; Yuan, B.; Zhang, X.; Wu, K.; Sun, W.; Li, T. Complex Tectonic Deformation Characteristics, Formation and Evolution of the Mountain Front in the Middle Part of the Southern Margin of Junggar Basin. *Northwest. Geol.* **2023**, *56*, 62–74. [CrossRef]
29. Hou, R.; Wei, L.; Wang, Z. Structural Evolution Analysis of Huomatu Anticline Belt in the Southern Margin of Junggar Basin. *China Manganese Ind.* **2017**, *35*, 5–8.
30. Liu, G.; Li, J.; Zhu, M.; Qi, X.; Feng, Y.; Yuan, B.; Pang, Z. Controlling Factors and Favorable Area Prediction of Cretaceous Qingshuihe Formation in Gaoquan Area of the Southern Junggar Basin. *Earth Sci.* **2024**, *49*, 3529–3546. [CrossRef]
31. Qie, R.; Yao, A.; Wu, K.; Li, Q. Current and Historical Sealing Ability of Aika Fault Belt in West Wing of Chepaizi Uplift. *China Energy Environ. Prot.* **2017**, *39*, 67–72.
32. Yin, Z.; Zhu, F.; Zhou, Z.; Wang, L.; Liu, X.; Na, Z.; Wang, Y.; Huang, D. Main Controlling Factors of Hydrocarbon Accumulation and Enrichment of Jurassic Xishanyao Formation in Monan Slope, Junggar Basin. *Lithol. Reserv.* **2025**, *37*, 33–46. [CrossRef]
33. Cortes, C.; Vapnik, V. Support-Vector Networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.