Supplementary Material

# Tables and Figures

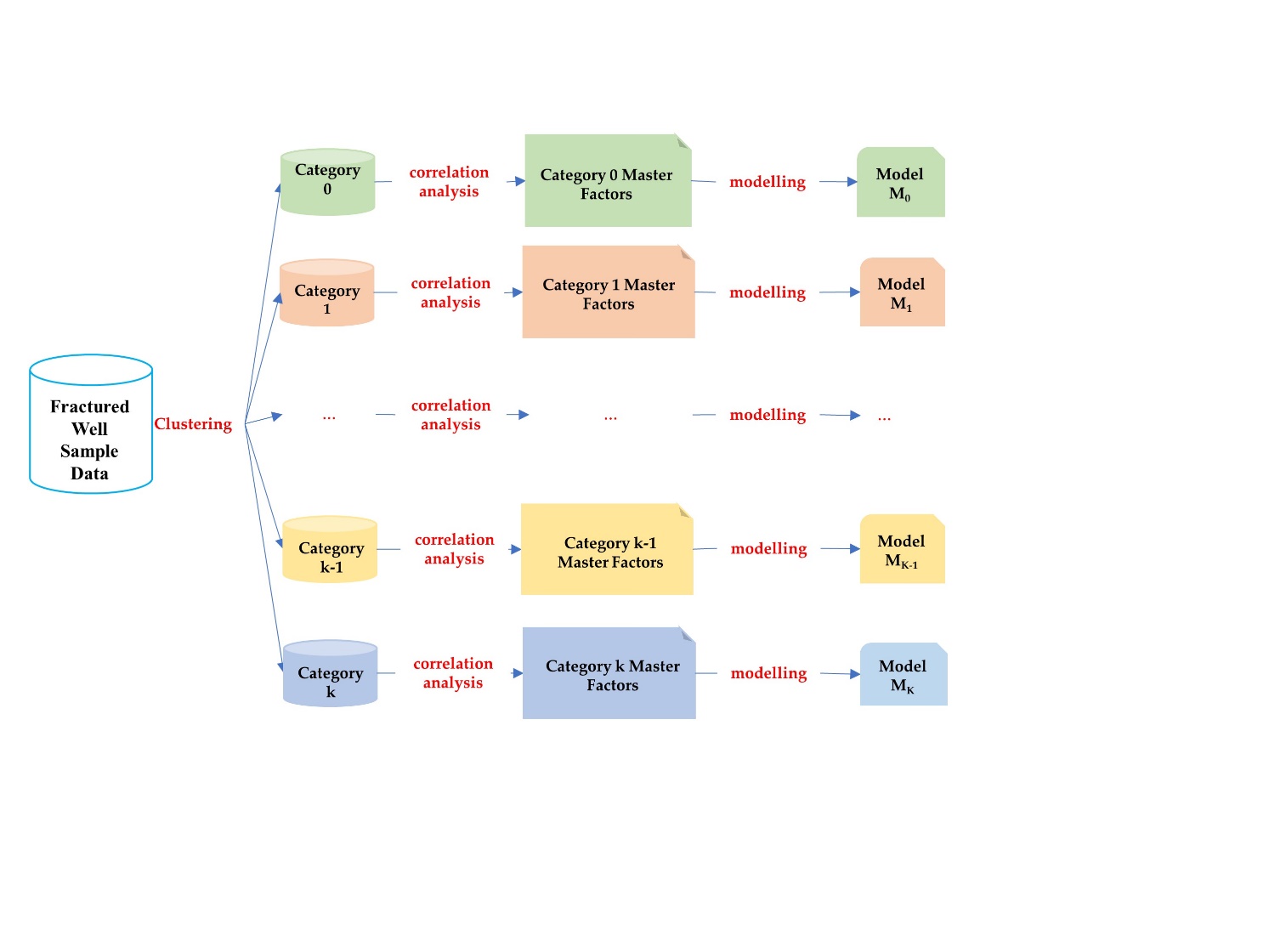


Figure 1. Modeling schematic of the fractured well capacity prediction model

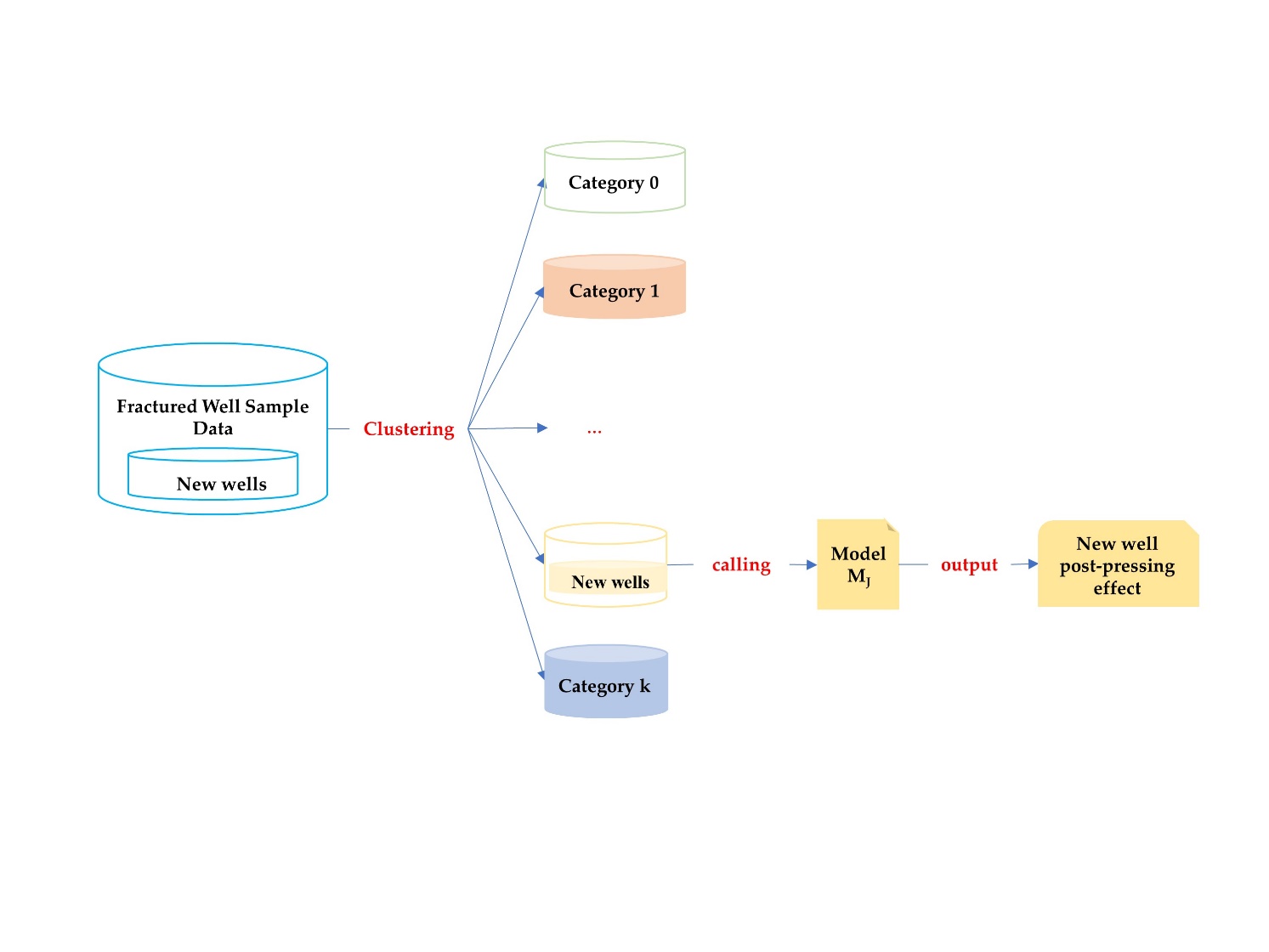
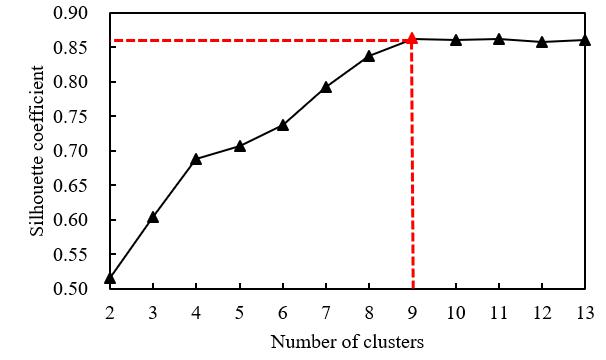
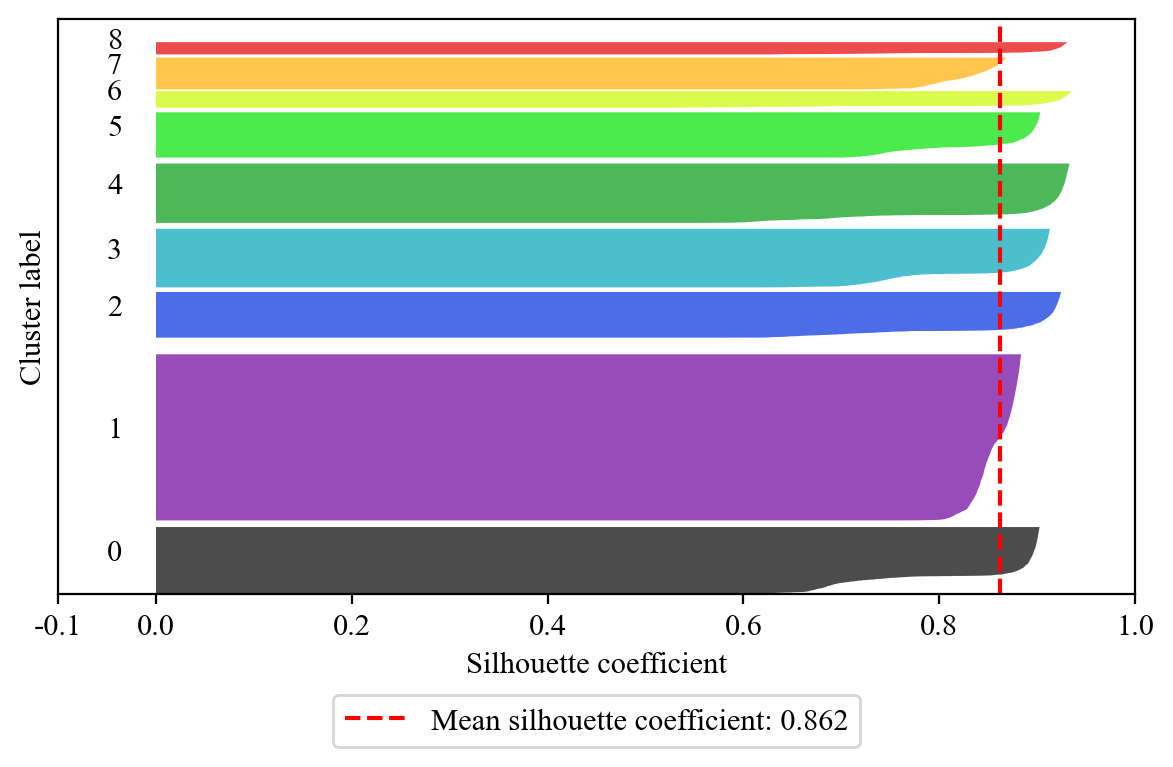


Figure 2. Schematic diagram of post-fracturing capacity intelligence model application

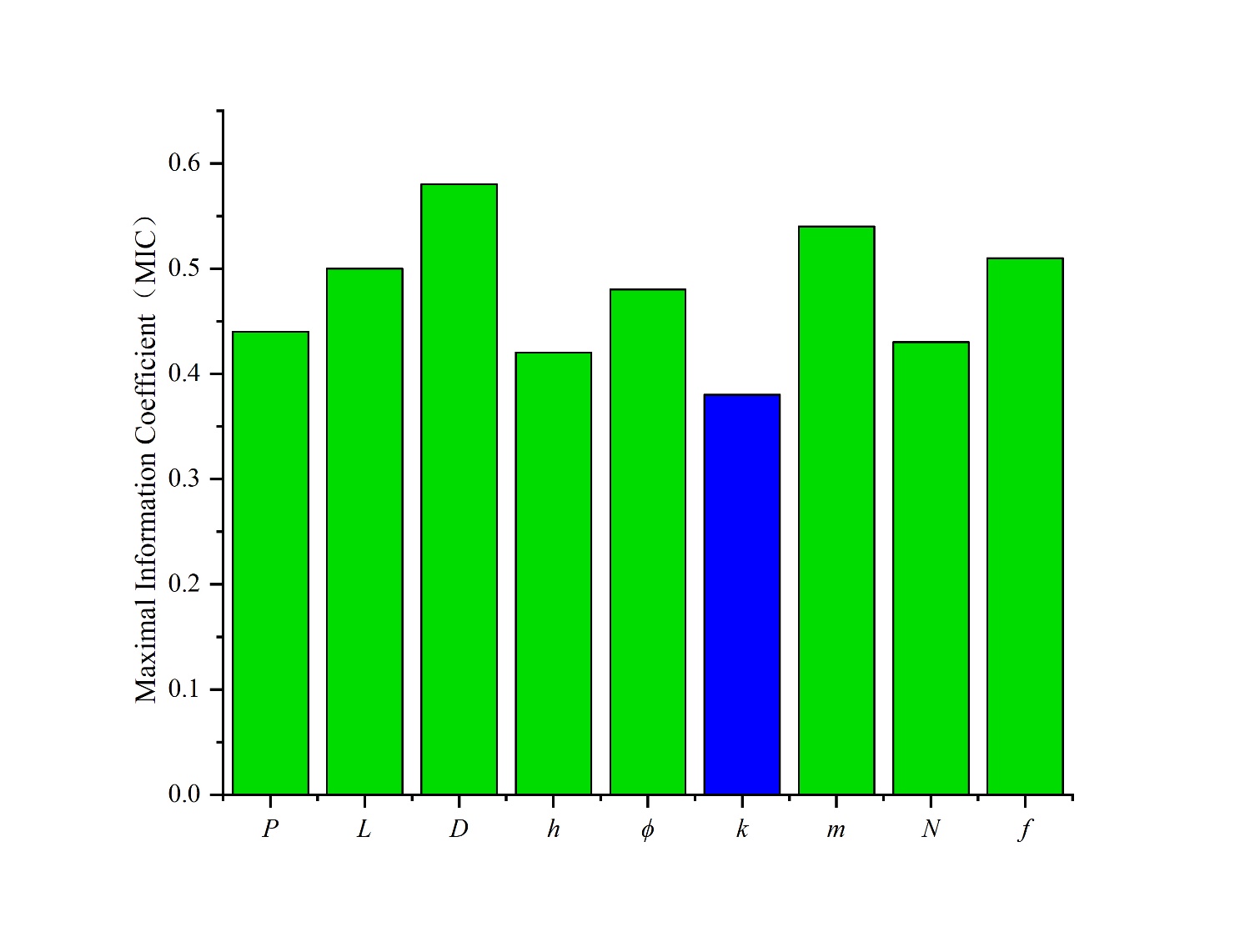


(A) The silhouette coefficient score situation for each cluster.

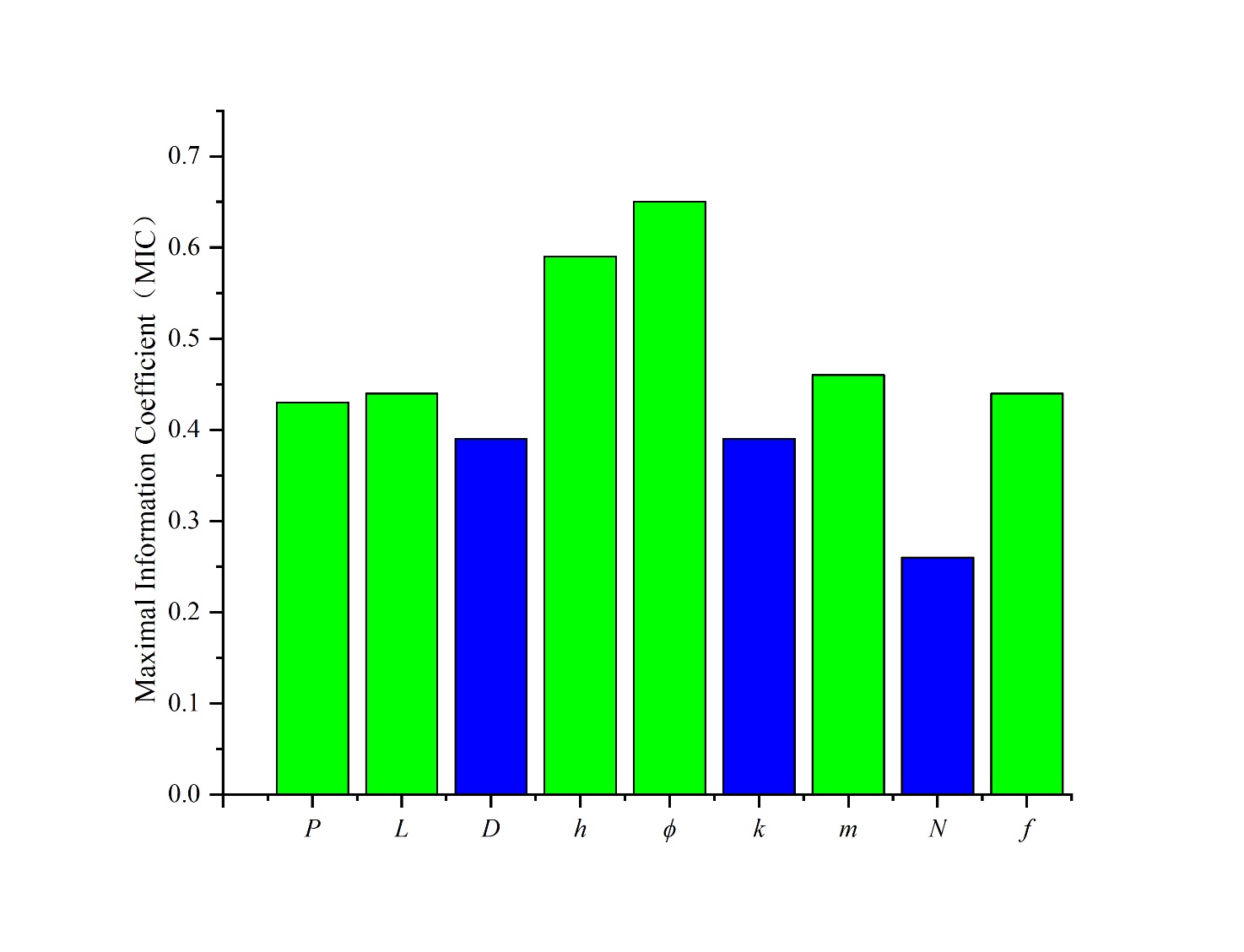


(B) Silhouette coefficient at k=9

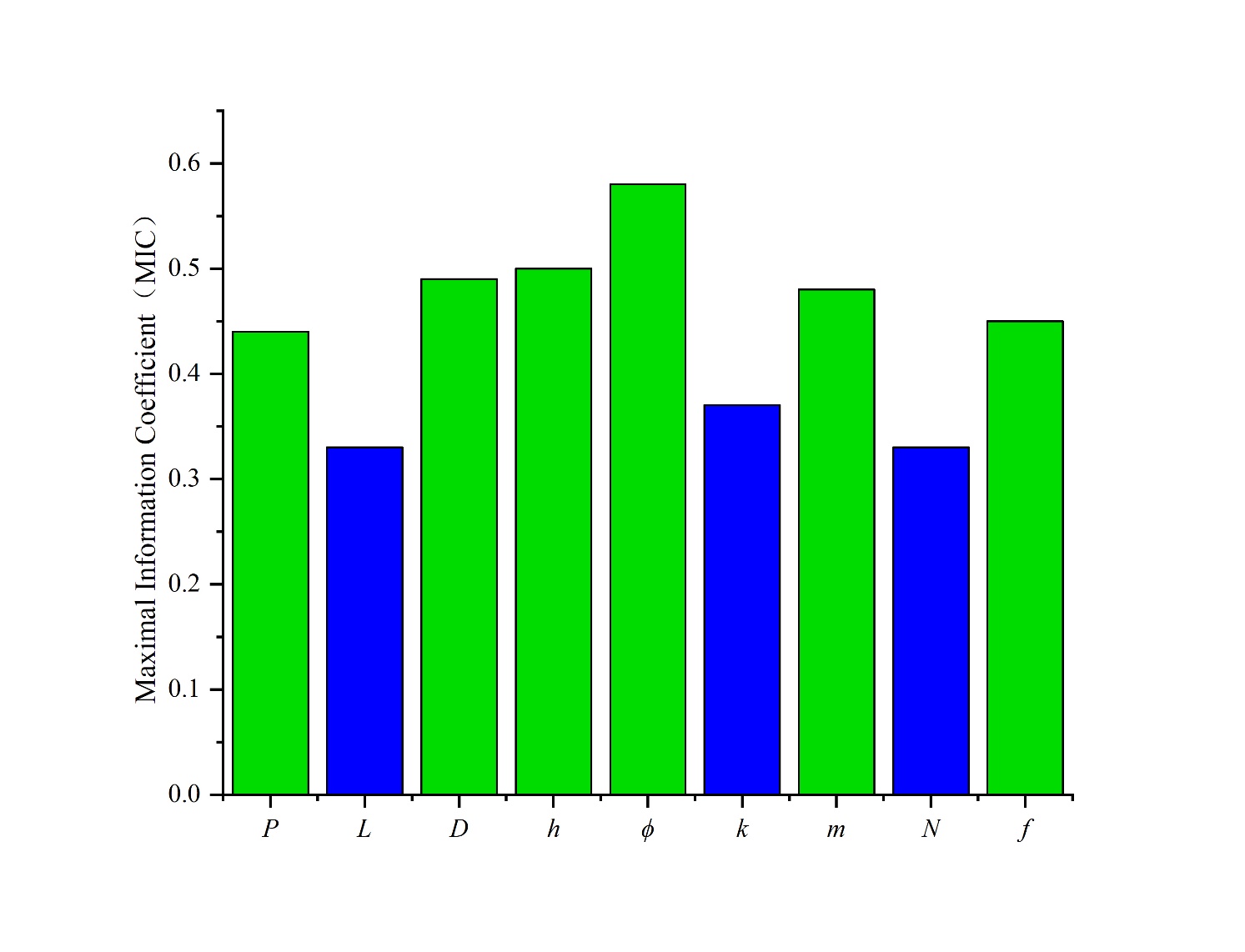
Figure 3. Evaluation of Clustering Results



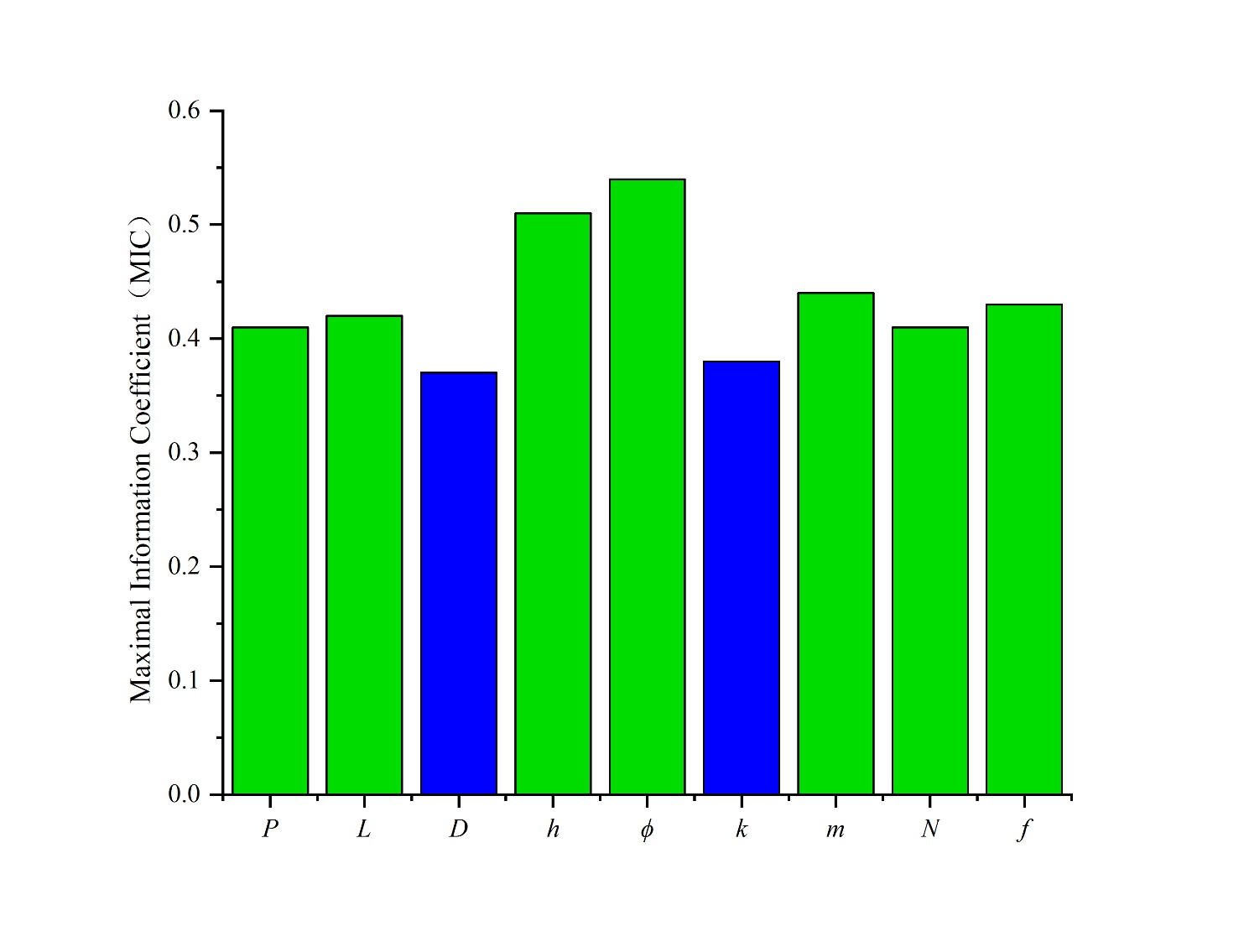
(A) Model 0



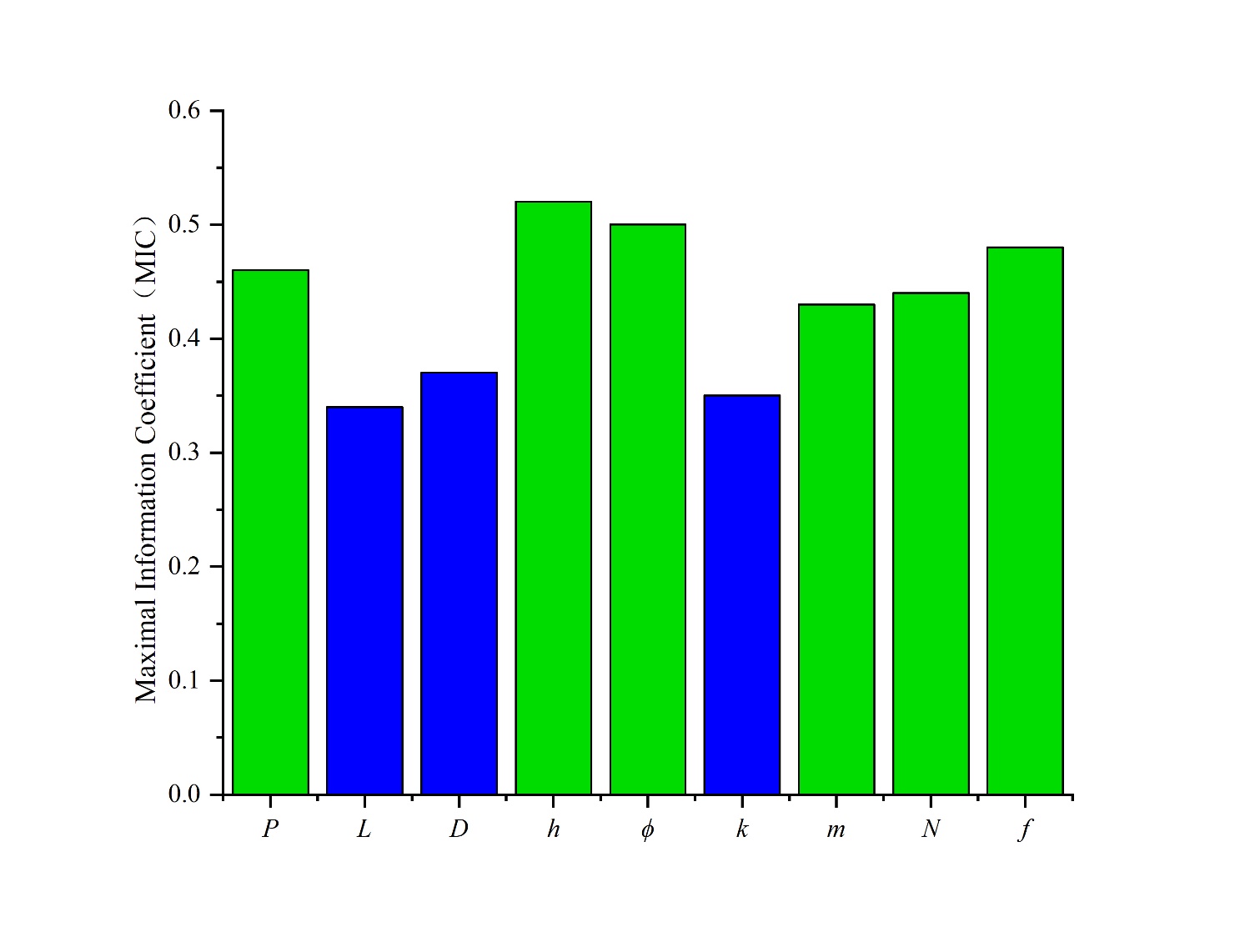
(B) Model 1



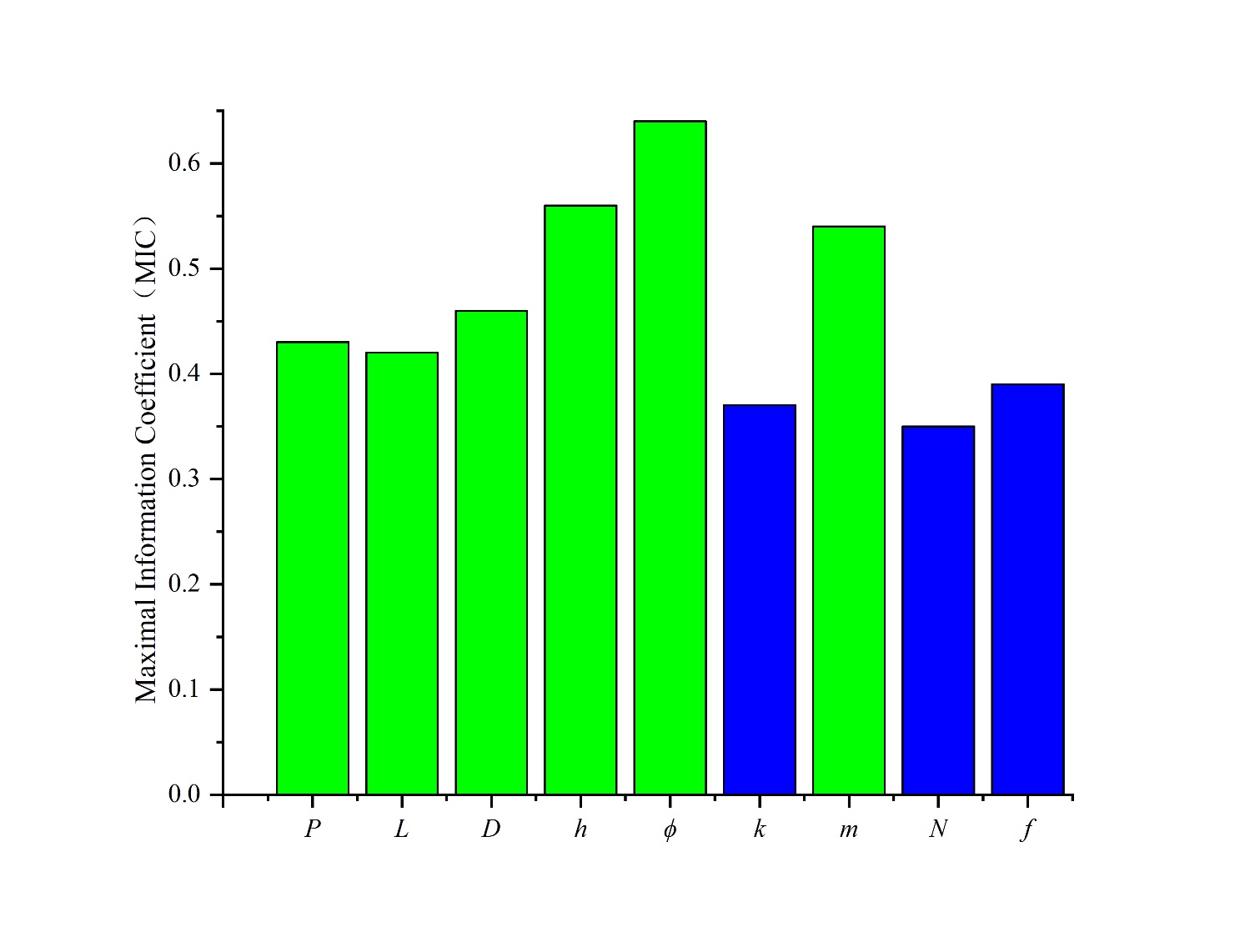
(C) Model 2



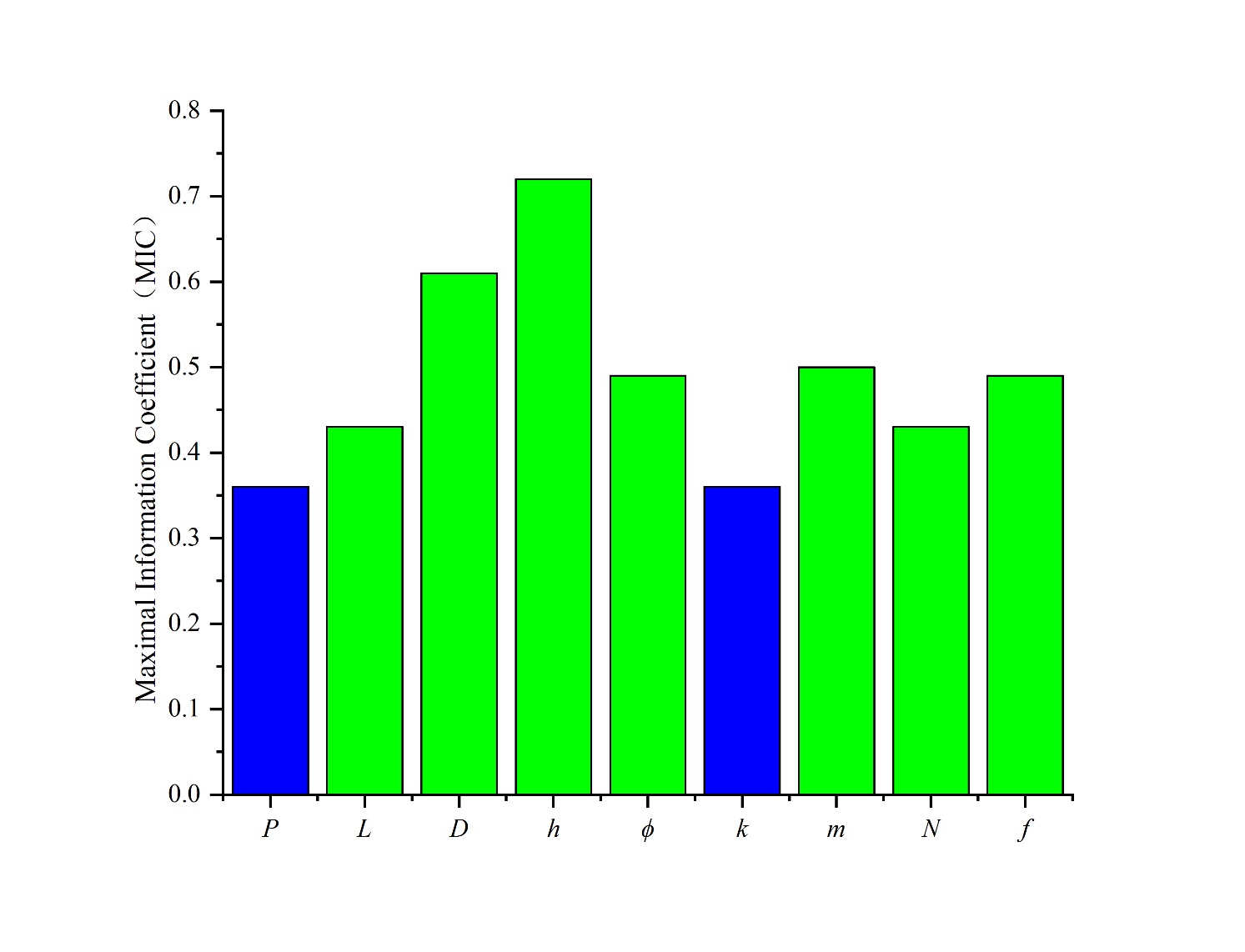
(D) Model 3



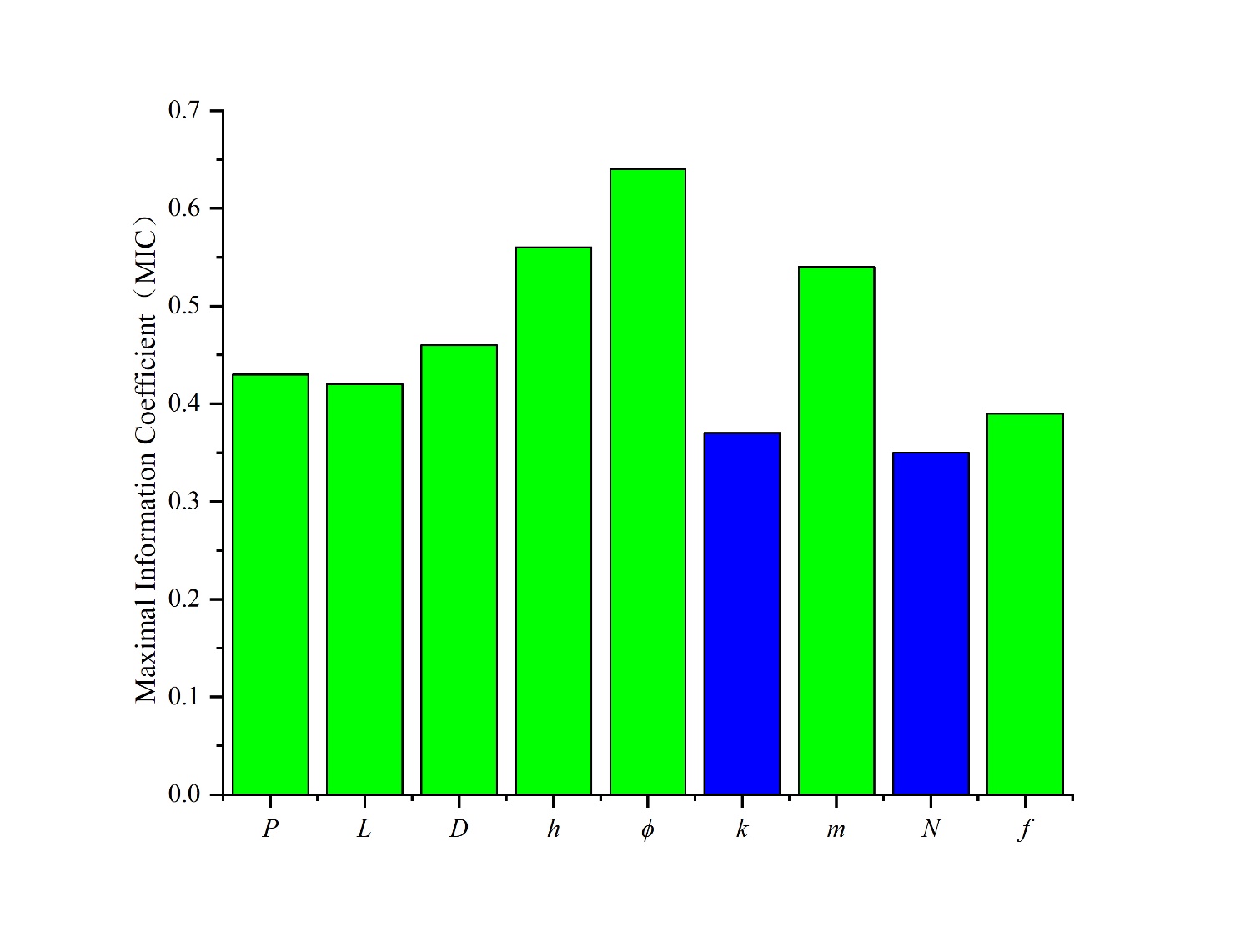
(E) Model 4



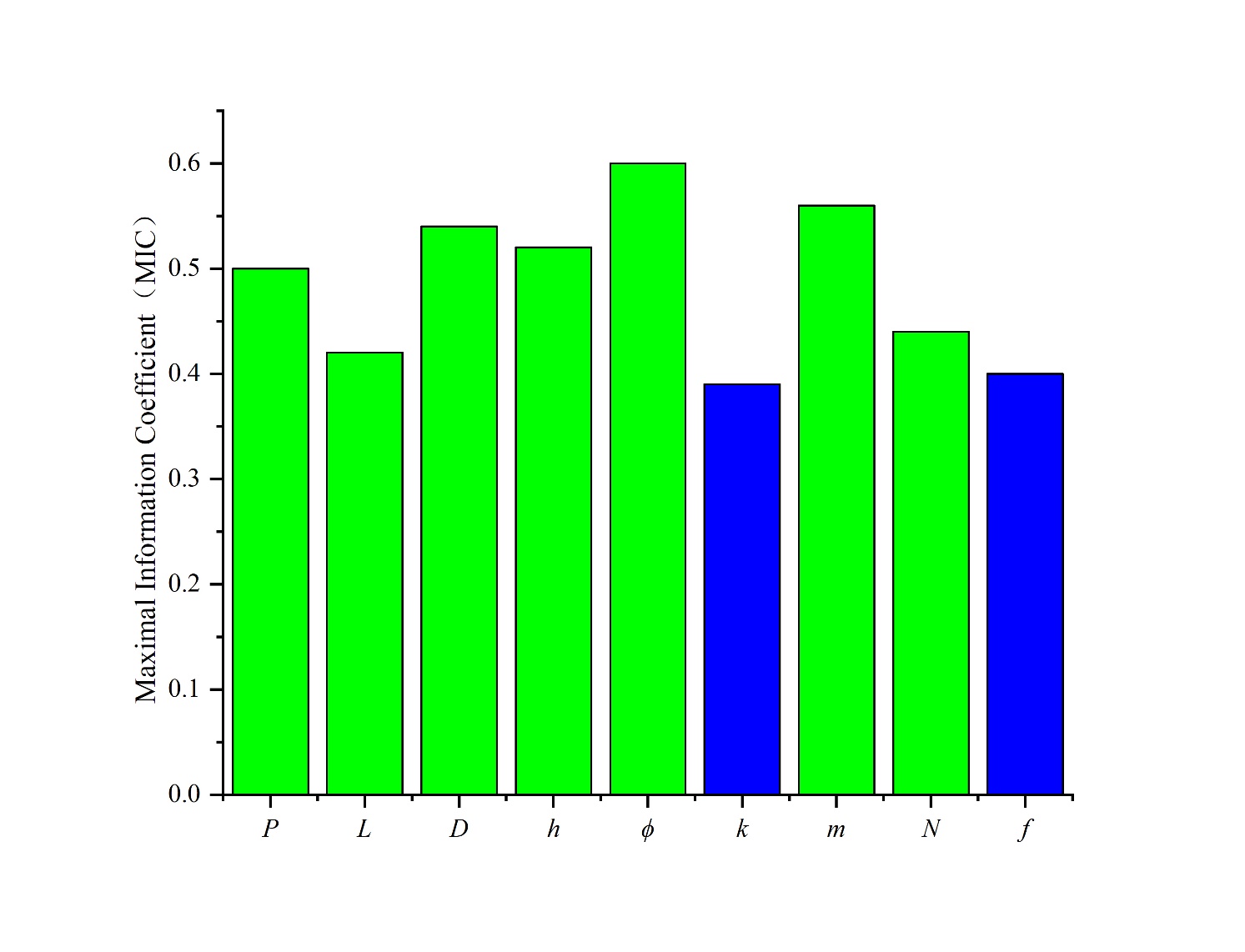
(F) Model 5



(G) Model 6



(H) Model 7



(I) Model 8

Figure 4. Correlation between main control factors and production capacity in different intervals

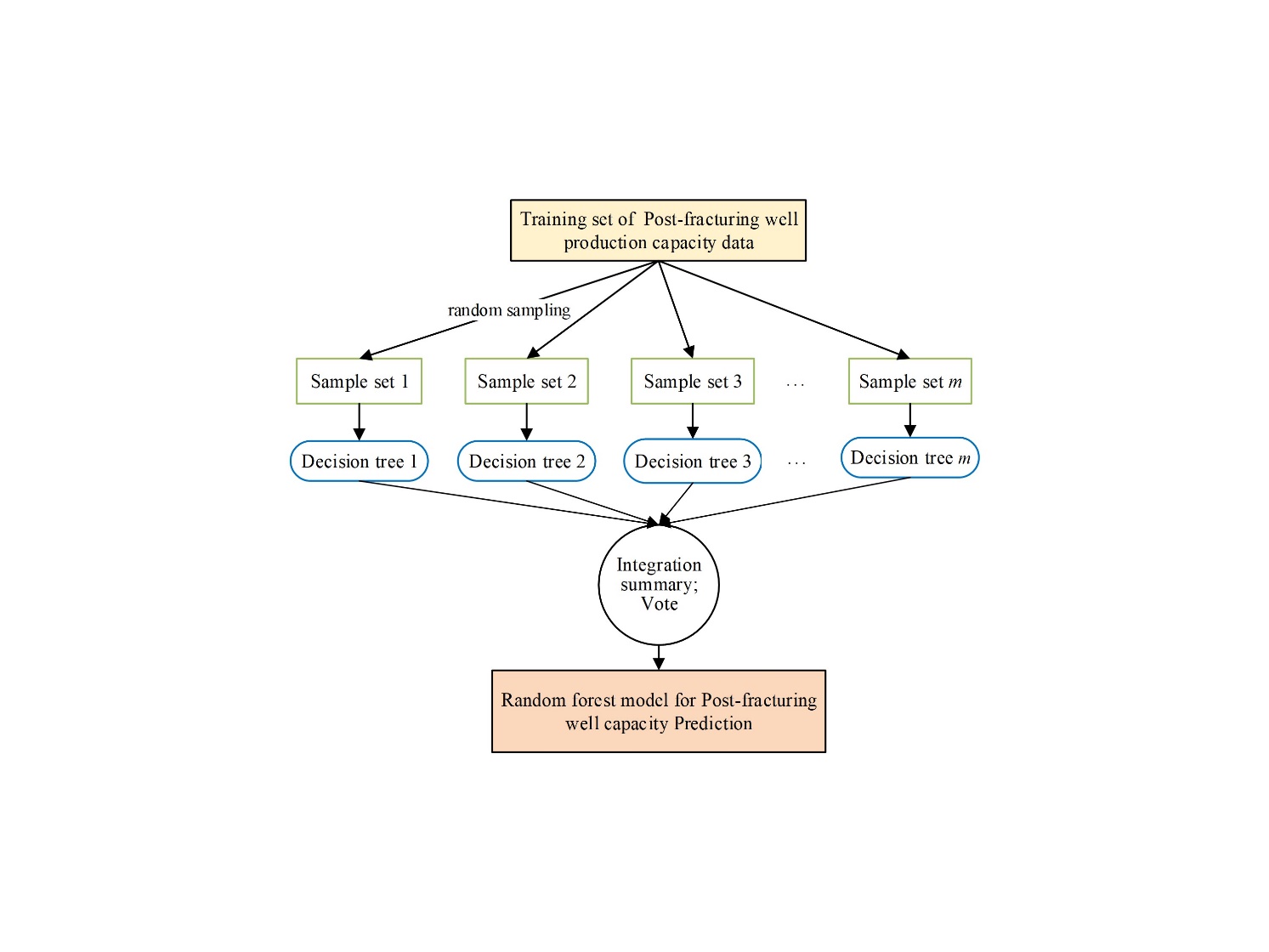


Figure 5 Process of random forest training

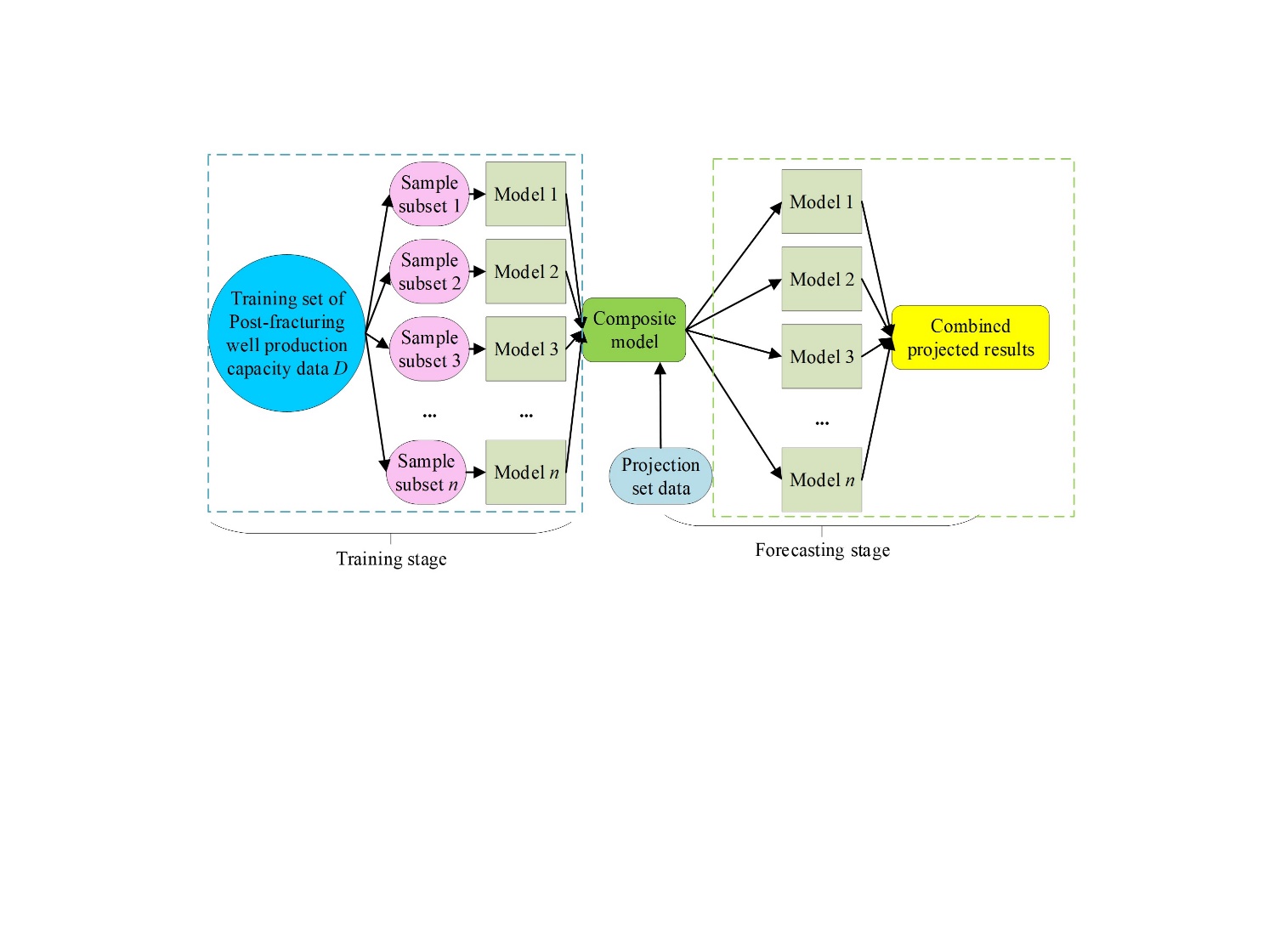


Figure 6 Principle diagram of Bagging regression algorithm

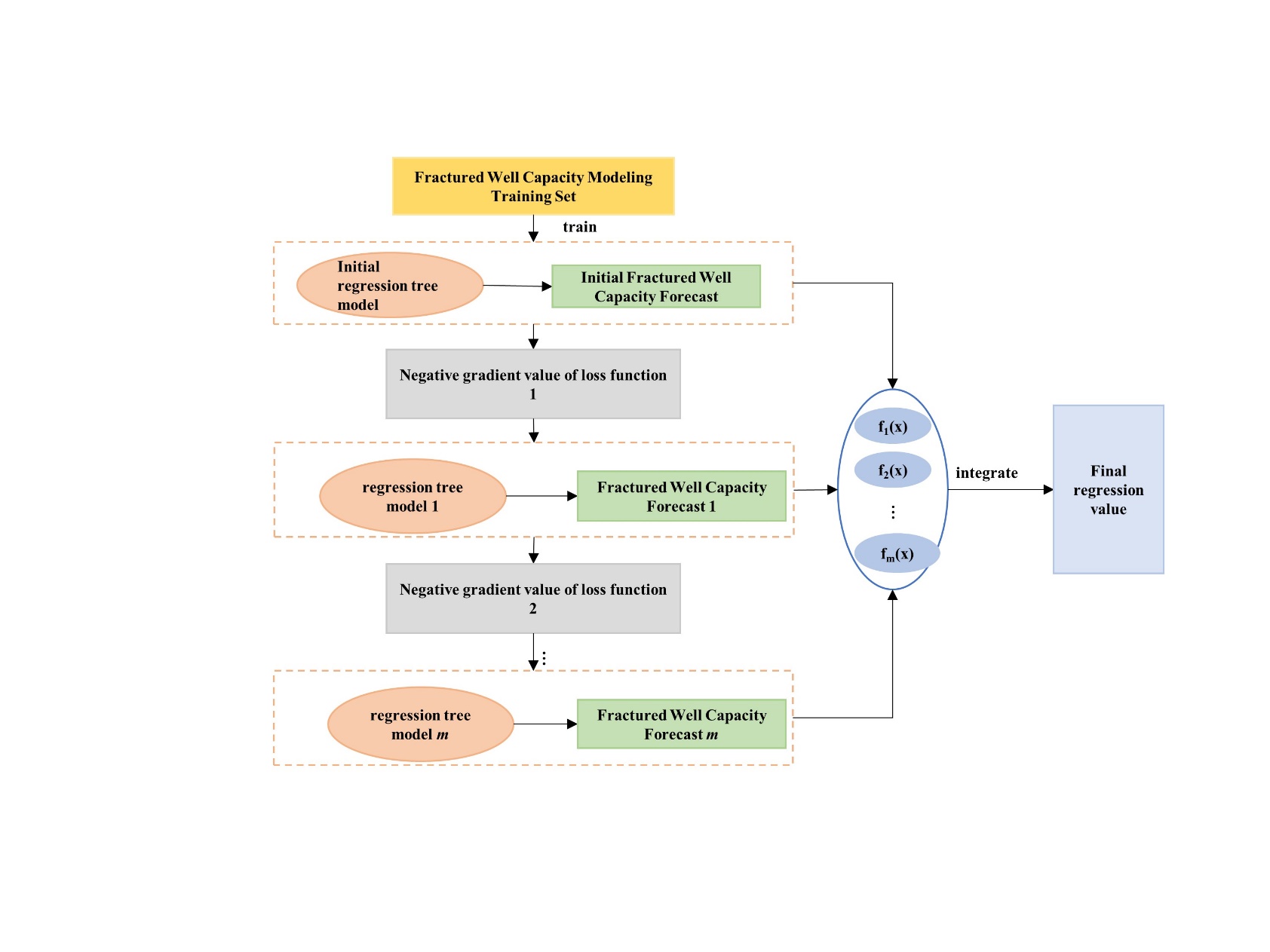
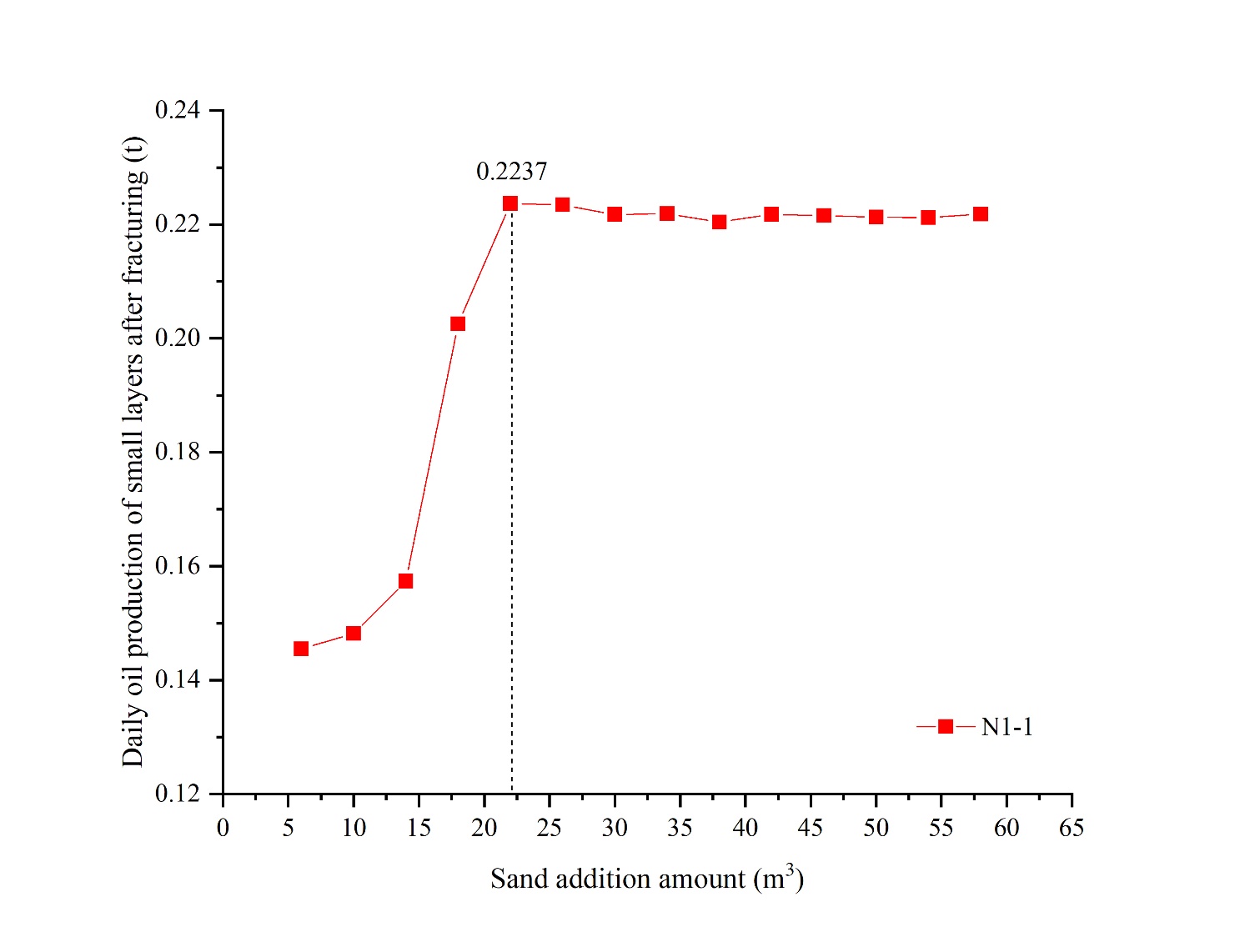
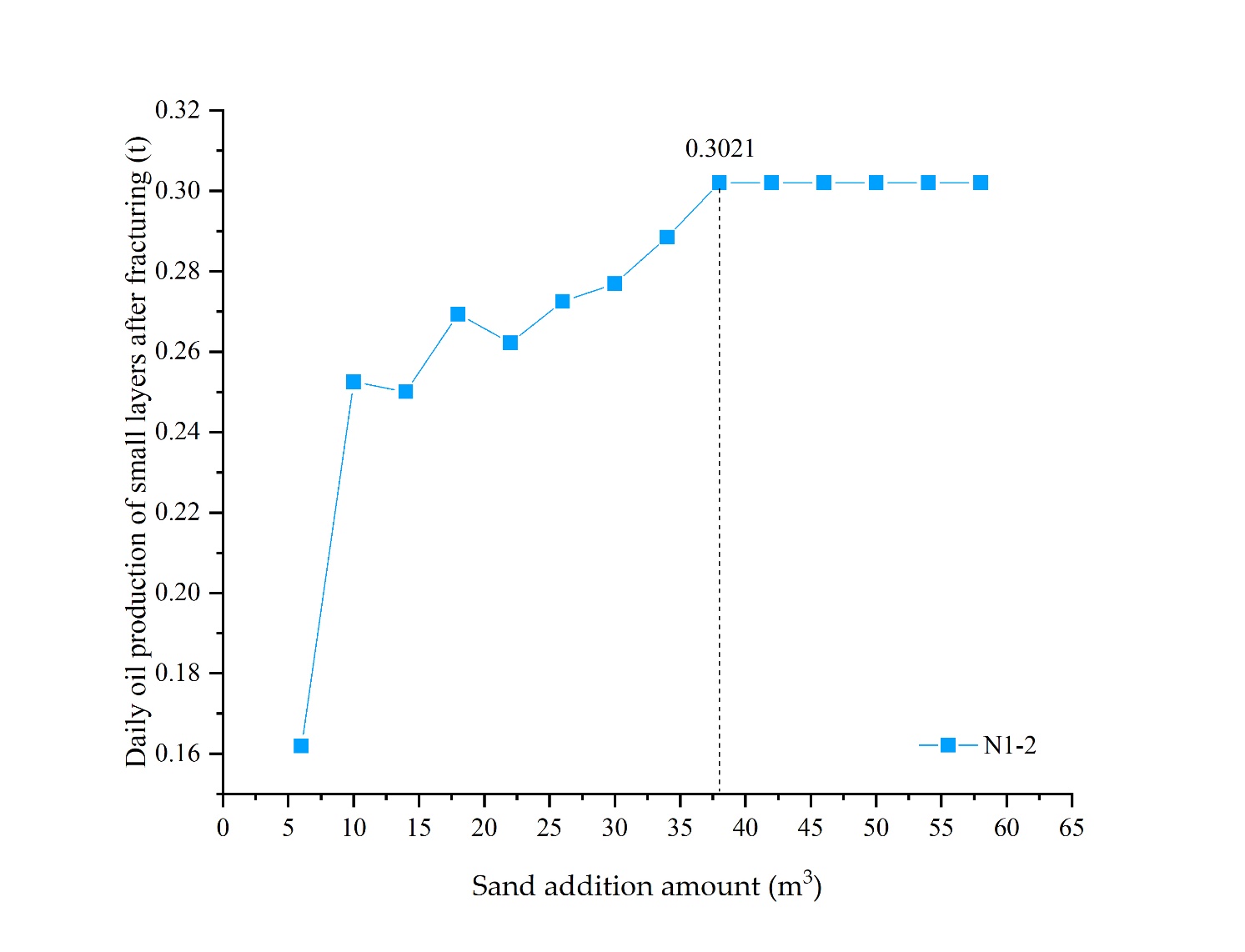


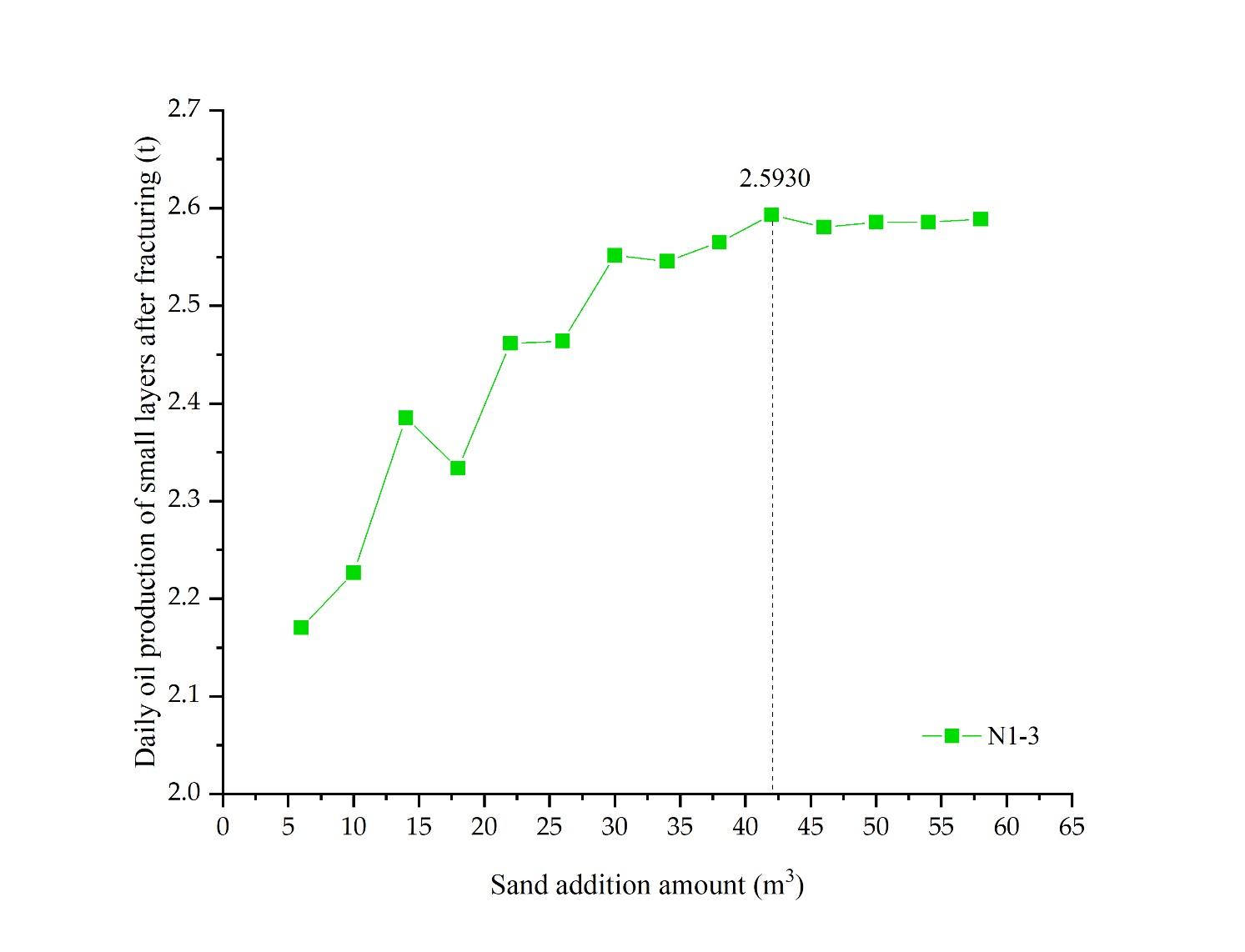
Figure 7. Establishment process of post fracturing production capacity prediction model



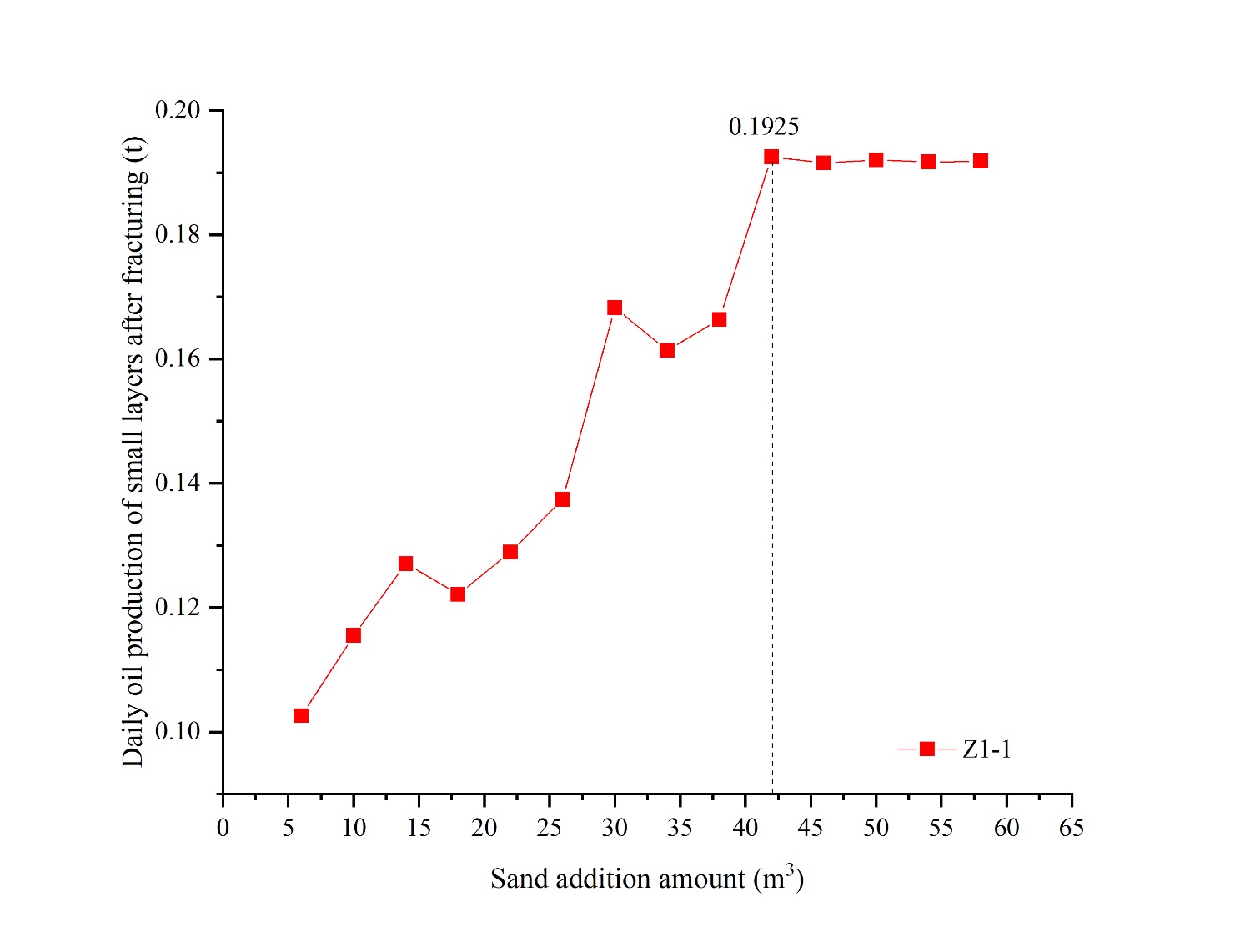
(A)N1-1



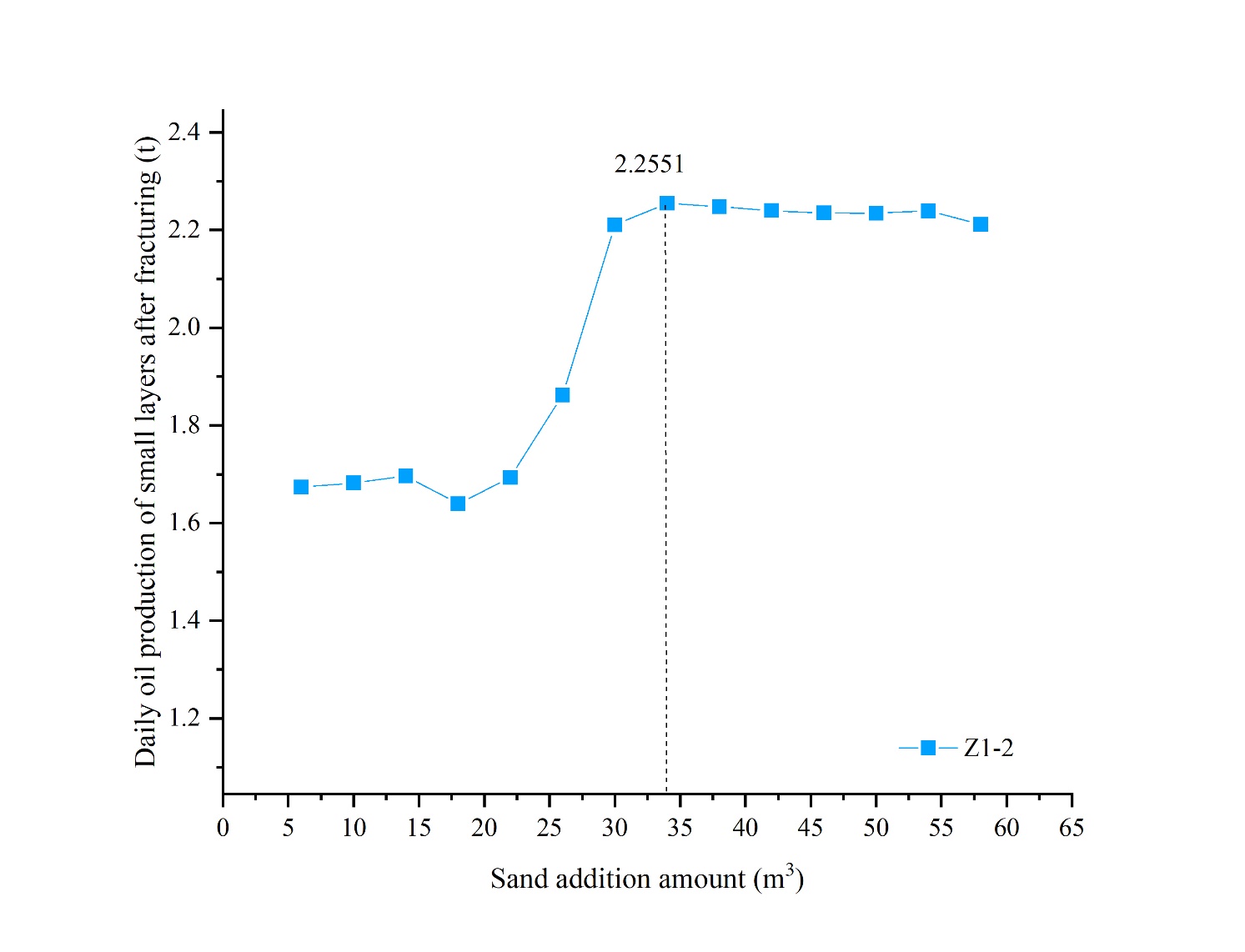
(B)N1-2



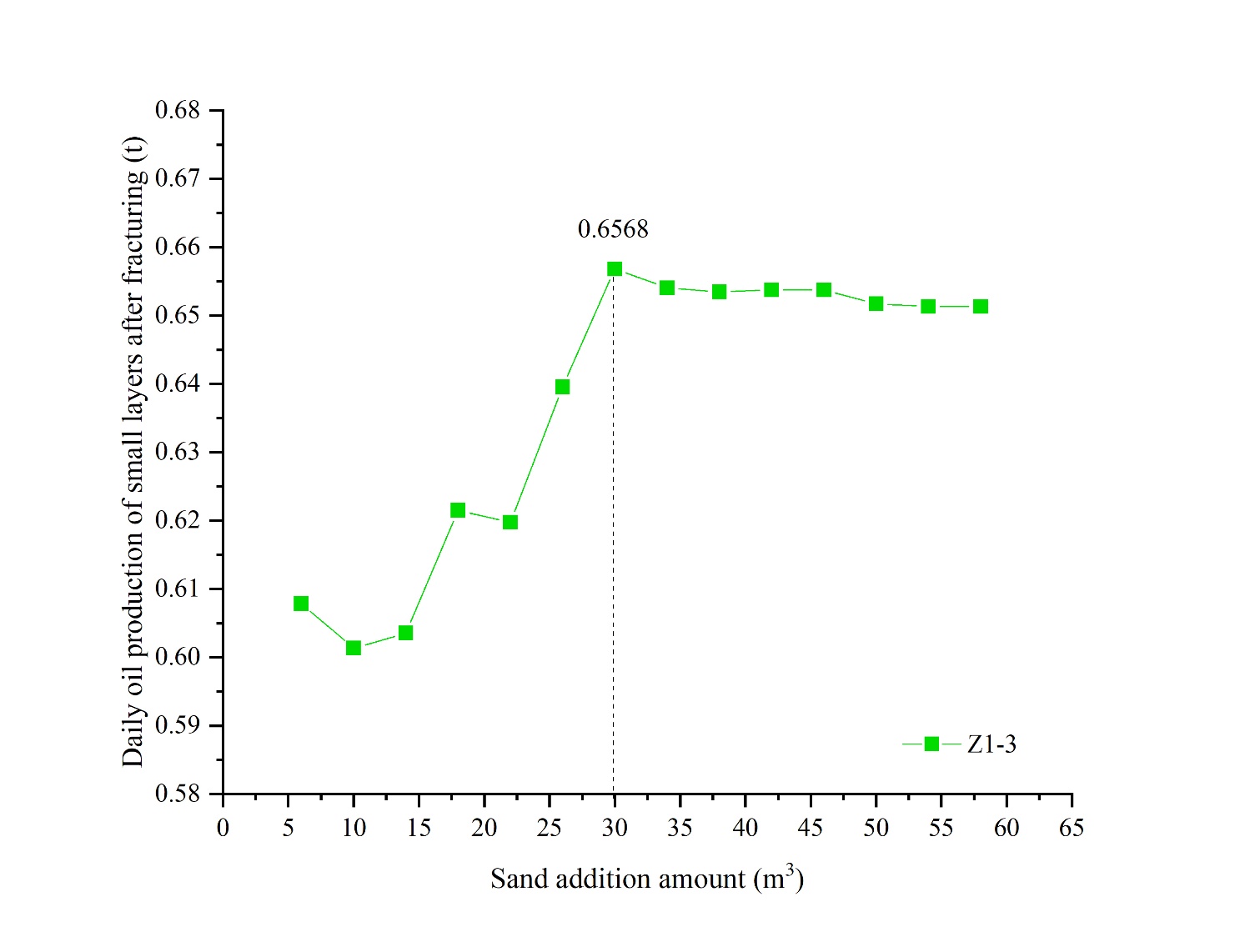
(C) N1-3



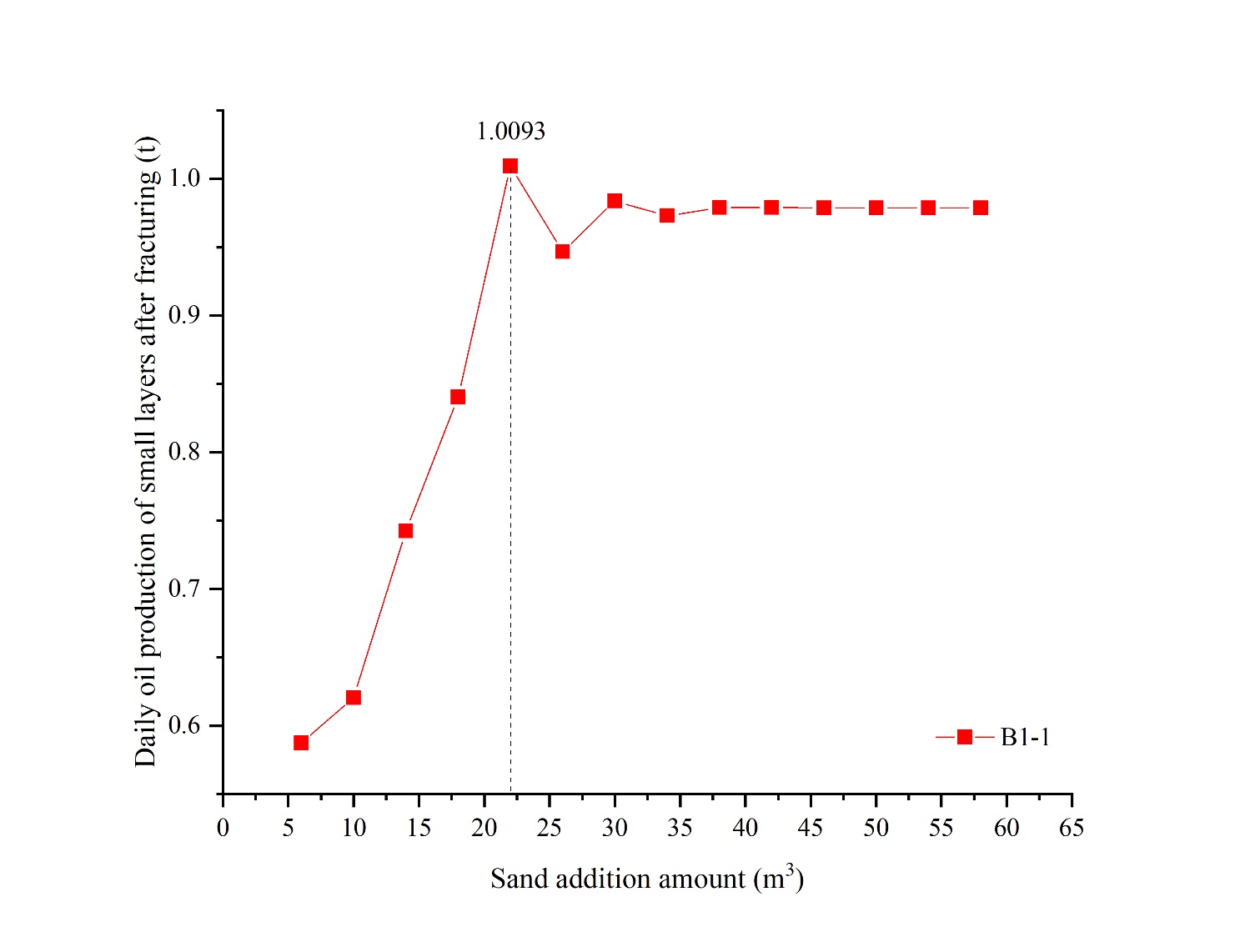
(D) Z1-1



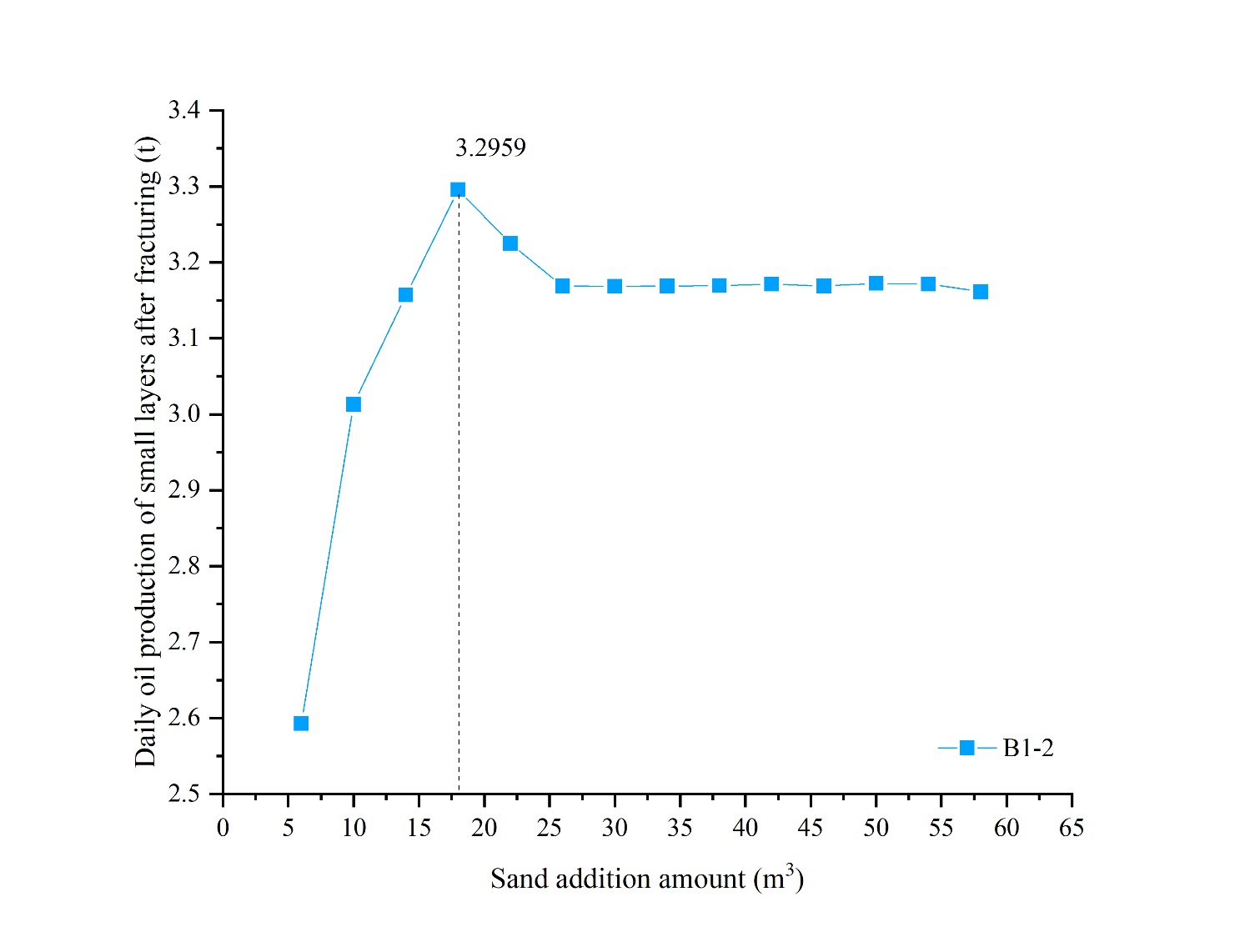
(E) Z1-2



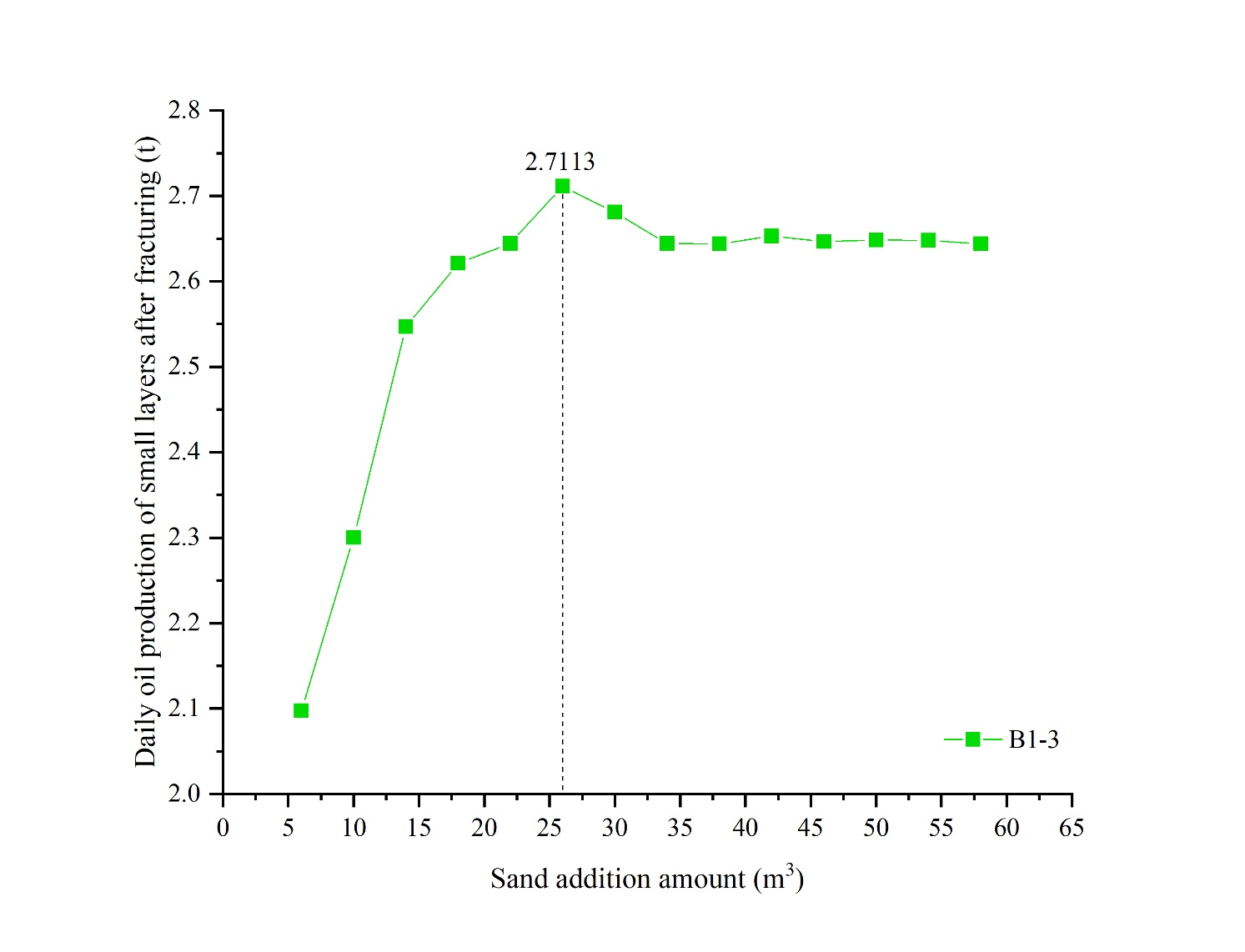
(F) Z1-3



(G) B1-1



(H) B1-2



(I) B1-3

Figure 8. Proppant volume optimization based on maximum production

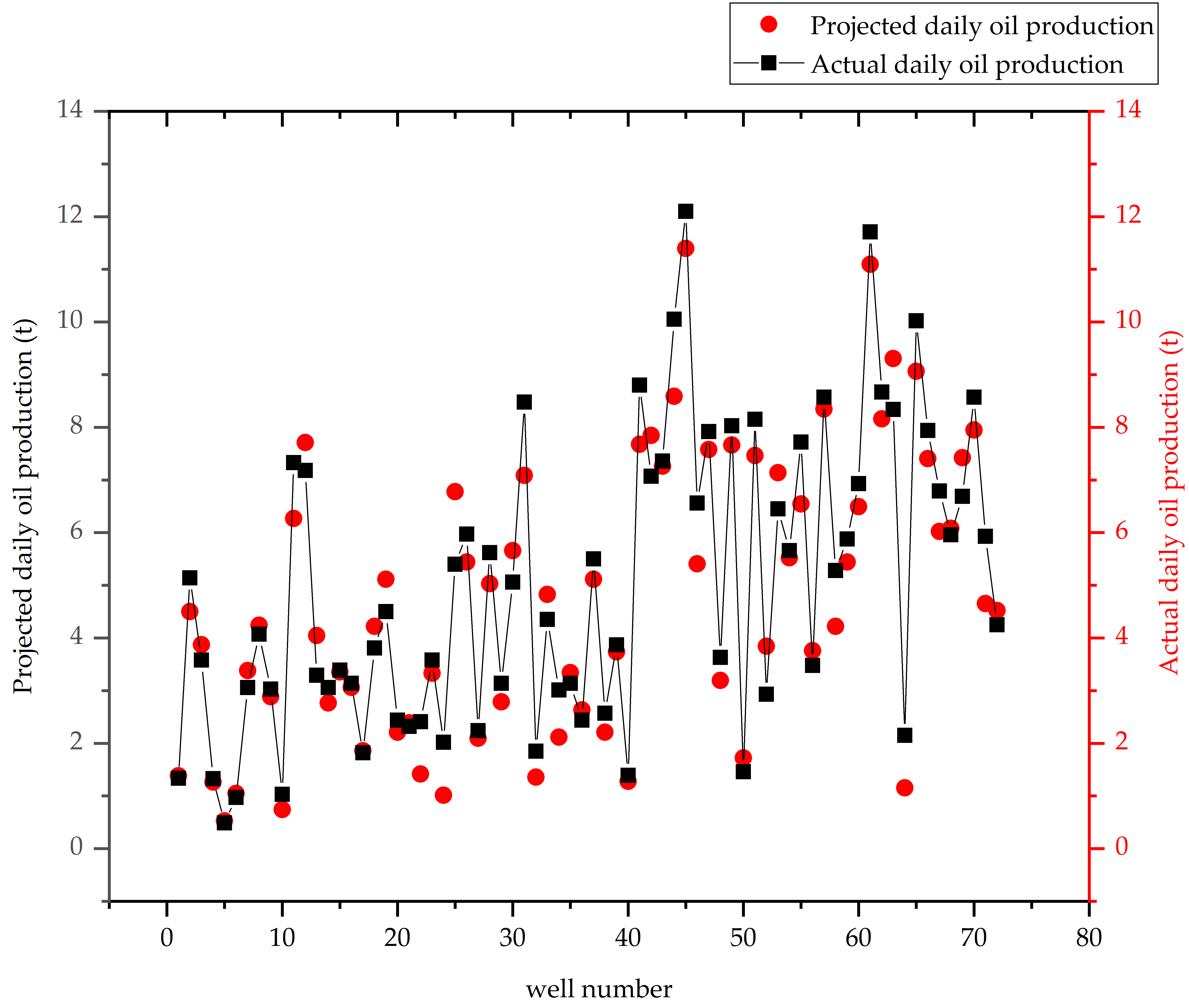


Figure 9. Predicted daily oil production vs. actual daily oil production

Table 1. Data composition of the data sample set

|  |  |  |  |
| --- | --- | --- | --- |
| displacement method | Fracturing wells  / well | fracturing technology | Fracturing wells  / well |
| Water flooding | 1878 | conventional fracturing | 921 |
| polymer flooding | 2379 | Multi-Fracture Fracturing | 3267 |
| Alkaline-Surfactant-Polymer | 1011 | Selective Fracturing | 1080 |
| total | 5268 | total | 5268 |

Table 2. Data Range of Dataset

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Data range | Present formation pressure  / MPa | Well Spacing  /m | Depth in the middle of the oil layer  /m | Sandstone thickness  /m | Effective thickness  /m | Porosity  /% |
| minimum value | 5.21 | 100 | 798 | 0.2 | 0.1 | 20 |
| maximum values | 20.48 | 300 | 1192 | 19.3 | 12.4 | 53 |
|  | Permeability  /μm2 | Fracture uncot  /strip | Total fracturing fluid volume  /m3 | Proppant volume  /m3 | Pre-fracture water content  /% | Daily oil production from small layer before fracking  /t |
| minimum value | 0.004 | 1 | 32 | 4.5 | 82 | 0.001 |
| maximum values | 1.5 | 2 | 111 | 120 | 99 | 8.79 |

Table 3. Dummy variable coded transformational relationships between the replacement method and the fracturing process

|  |  |  |
| --- | --- | --- |
| Data parameters | Data categories before conversion | Post-conversion data |
| Displacement method | Alkaline-Surfactant-Polymer | 00 |
| Polymer flooding | 01 |
| Water flooding | 10 |
| Fracturing technology | Selective Fracturing | 00 |
| Conventional fracturing | 01 |
| Multi-Fracture Fracturing | 10 |

Table 4 Tree depth, tree number and maximum R2 determination coefficient of 9 types of fracturing well productivity prediction models (Random Forest)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model categories | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Regression tree depth/layer | 15 | 19 | 18 | 19 | 17 | 16 | 15 | 19 | 16 |
| Number of regression trees/tree | 850 | 100 | 850 | 950 | 150 | 350 | 150 | 900 | 100 |
| Test set coefficient of determinationR2 | 0.90 | 0.82 | 0.83 | 0.85 | 0.89 | 0.84 | 0.83 | 0.86 | 0.82 |
| Training set coefficient of determinationR2 | 0.97 | 0.96 | 0.99 | 0.97 | 0.98 | 0.98 | 0.99 | 0.96 | 0.97 |

Table 5 Tree depth, tree number and maximum R2 determination coefficient of 9 types of fracturing well productivity prediction models (Bagging)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model categories | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Regression tree depth/layer | 19 | 19 | 20 | 20 | 18 | 20 | 19 | 19 | 20 |
| Number of regression trees/tree | 200 | 100 | 800 | 1000 | 200 | 100 | 350 | 600 | 750 |
| Test set coefficient of determinationR2 | 0.90 | 0.91 | 0.85 | 0.89 | 0.91 | 0.85 | 0.85 | 0.88 | 0.82 |
| Training set coefficient of determinationR2 | 0.98 | 0.94 | 0.97 | 0.96 | 0.98 | 0.95 | 0.96 | 0.95 | 0.94 |

Table 6. Maximum values of model tree depth, number of trees, and corresponding R2 coefficients of determination for 9 types of fractured well production capacity prediction models (GBRT)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model categories | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Regression tree depth/layer | 7 | 8 | 10 | 7 | 6 | 14 | 7 | 5 | 9 |
| Number of regression trees/tree | 800 | 950 | 900 | 1000 | 850 | 950 | 850 | 950 | 550 |
| Test set coefficient of determinationR2 | 0.91 | 0.90 | 0.89 | 0.89 | 0.91 | 0.89 | 0.90 | 0.89 | 0.89 |
| Training set coefficient of determinationR2 | 0.93 | 0.92 | 0.93 | 0.92 | 0.92 | 0.91 | 0.94 | 0.93 | 0.92 |

Table 7. Accuracy of prediction models in practical applications

|  |  |  |
| --- | --- | --- |
| Model categories | MRE | RMSE |
| 0 | 7.94% | 0.26 |
| 1 | 11.33% | 0.35 |
| 2 | 17.12% | 0.42 |
| 3 | 14.66% | 0.45 |
| 4 | 10.97% | 0.44 |
| 5 | 9.92% | 0.57 |
| 6 | 12.32% | 0.58 |
| 7 | 13.18% | 0.60 |
| 8 | 9.47% | 0.85 |