Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Hybrid Approach to Managing Oilfield Development Processes under Uncertainty

Batyr Orazbayev
Gumilyov Eurasian National University
Kulman Orazbayeva
Esil University
Arailym Zhienbek
Gumilyov Eurasian National University
Valentina Makhatova
Khalel Dosmukhamedov Atyrau University
Sandugash Iskakova
Atyrau Oil and Gas University
Balbupe Utenova
Atyrau Oil and Gas University
Lyailya Kurmangaziyeva
Kurmangazieval475@gmail.com

Khalel Dosmukhamedov Atyrau University

Article

Keywords: hybrid approach, uncertainty, fuzzy model, expert knowledge, genetic algorithm, oil field

Posted Date: November 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11299486/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Article

Hybrid Approach to Managing Oilfield Development Processes under Uncertainty

Batyr Orazbayev[1,2], Kulman Orazbayeva[3], Arailym Zhienbek[1], Valentina Makhatova[4], Sandugash Iskakova[5], Balbupe Utenova[5*], Lyailya Kurmangaziyeva[4*]

[1] Institute of Digital Sciences and Artificial Intelligence, Gumilyov Eurasian National University, Astana, 010008, Kazakhstan
[2] Higher School of Information Technology and Engineering, Astana International University, Astana, 020000, Kazakhstan
[3] Faculty of Applied Sciences, Esil university, Astana, 010005, Kazakhstan
[4] Faculty of Physics, Mathematics and Information Technologies, Kh. Dosmukhamedov Atyrau University, Atyrau, 060000, Kazakhstan
[5] Faculty of Information Technologies, Atyrau Oil and Gas University, Atyrau, 060027, Kazakhstan

[*]Corresponding Authors: Email: Kurmangazieval475@gmail.com, b.utenova@aogu.edu.kz

Abstract
Effective management of oil field development processes is a complex challenge due to geological heterogeneity, uncertain petrophysical properties, unstable reservoir pressure, and fluctuating production and injection rates. Economic and technological factors further complicate decision-making. Conventional optimization methods based on deterministic models and standalone artificial intelligence techniques cannot fully account for uncertainties in initial data and expert estimates. This paper proposes a novel hybrid approach that combines artificial intelligence, fuzzy logic, neuro-fuzzy modeling, and evolutionary optimization. The approach integrates fuzzy models for representing linguistic uncertainty, the adaptive ANFIS model for learning from field data, and a genetic algorithm for optimizing field operating modes. A mathematical model for multi-criteria management is developed, considering cumulative oil production, oil recovery factor, produced water volumes, and operational costs. Membership functions and an inference engine are used to formalize expert knowledge and account for uncertainty in technological parameters. Optimization is performed using a genetic algorithm to identify effective solutions under interconnected parameters and nonlinear constraints. The results demonstrate that integrating fuzzy logic, neuro-fuzzy modeling, and evolutionary optimization improves the robustness of the control system to uncertain information. The proposed approach enhances decision-making, improves forecasting of field development indicators, and supports rational control of reservoir operating modes under changing field conditions.

Keywords: hybrid approach, uncertainty, fuzzy model, expert knowledge, genetic algorithm, oil field**.**

Introduction

Oil field development (OFD) is a complex, multi-level technological process whose efficiency is determined by a large number of interrelated geological, hydrodynamic, technological, and economic factors. Modern fields are characterized by a high degree of reservoir heterogeneity, variations in the petrophysical properties of rocks, reservoir pressure instability, varying reservoir permeability, the presence of residual oil zones, and continuous changes in the operating modes of production and injection wells [1, 2]. These features significantly complicate the decision-making process in OFD management and reduce the effectiveness of traditional optimization methods.

Most existing management models assume the availability of precise quantitative data, whereas real field development processes are accompanied by various types of uncertainty [3]. Therefore, current attention is focused on the use of artificial intelligence technologies, data mining, and decision support methods, as described, for example, in [4–6]. This approach makes it possible to analyze large volumes of production information, identify hidden patterns in field performance, and forecast changes in technological indicators with high accuracy under uncertain conditions.

Sources of uncertainty in OFD include limited geological and field data, measurement errors, ambiguity in the interpretation of geophysical survey results, the stochastic nature of filtration processes, changes in reservoir fluid properties, and the subjectivity of expert estimates by specialists. As a result, the

exclusive use of deterministic models often leads to a decrease in forecasting quality and the derivation of suboptimal operating regimes.

One of the most promising areas for solving this problem is the application of soft computing technologies, which integrate fuzzy logic, artificial neural networks, evolutionary algorithms, and probabilistic information processing methods [7–9]. The theory and practice of soft computing are described by Rajesh et al. in [7]. Issues related to the application of soft computing technologies in engineering and intelligent methods in geophysics are investigated by the authors of [8, 9]. Unlike classical mathematical methods, these technologies allow for effective handling of incomplete, inaccurate, and contradictory data, ensuring robust solutions under uncertainty.

Hybrid intelligent systems that combine the advantages of various artificial intelligence methods are of particular interest [10, 11]. Bajaj et al. in [10] considered general issues of creating and applying hybrid intelligent systems. Oscar and Patricia in [10] proposed the development of hybrid intelligent systems based on the integration of fuzzy logic, neural networks, and metaheuristic algorithms. Fuzzy logic provides the formalization of expert knowledge and allows technological parameters to be described using linguistic variables [12–15]. Artificial neural networks possess the ability to automatically learn from production data and identify complex nonlinear relationships between the parameters of the field development process [16]. Genetic algorithms effectively solve large-scale multi-criteria optimization problems, ensuring the search for a global optimum in the presence of numerous constraints [17]. The combined use of these methods makes it possible to compensate for the drawbacks of each and significantly increase the efficiency of the intelligent control system.

In recent years, the Adaptive Neuro-Fuzzy Inference System (ANFIS), which is a five-layer feedforward neural network combining the principles of artificial neural networks (ANNs) and fuzzy logic, has become widely used in modeling the technological processes of the oil and gas industry. These models combine the transparency of fuzzy inference systems with the learning capabilities of ANNs, allowing for the automatic adjustment of membership function parameters and the adaptation of the model to changing field operating conditions [18]. However, in most published studies, neuro-fuzzy models are used primarily for forecasting individual technological parameters, whereas issues of comprehensive field development management considering multi-criteria optimization remain insufficiently explored.

Furthermore, existing optimization methods are often focused on achieving a single objective, such as maximizing oil production or minimizing operational costs. In practice, the OFD process requires the simultaneous consideration of several conflicting criteria, including increasing the oil recovery factor, reducing product water cut, minimizing energy costs, maintaining reservoir pressure stability, and complying with technological constraints. Solving such problems requires the application of multi-criteria optimization methods, i.e., decision-making capable of accounting for conflicting objectives and the uncertainty of initial information.

This paper proposes a hybrid approach to OFD management based on the integration of fuzzy logic, the ANFIS neuro-fuzzy model, and a genetic algorithm. The proposed approach is designed for intelligent decision support in selecting optimal field operating modes under conditions of uncertainty in geological and technological information.

The scientific novelty of the research lies in the development of a hybrid intelligent management model that integrates a fuzzy inference mechanism, adaptive learning of a neuro-fuzzy model, and evolutionary multi-criteria optimization. Unlike existing approaches, the proposed model provides a comprehensive account of the uncertainty in technological parameters, expert knowledge, and production data, which allows for increased forecasting accuracy and efficiency in selecting optimal OFD operating modes.

The practical significance of the work is determined by the potential application of the developed approach in intelligent control systems for oil-producing enterprises, field digital twins, and automated decision support systems. The use of the proposed model will make it possible to increase the oil recovery factor, reduce operational costs, improve the quality of well stock management, and ensure the sustainable functioning of field development processes under conditions of high uncertainty.

The aim of the research is to develop and investigate a hybrid intelligent approach to OFD management, based on the integration of fuzzy logic methods, ANFIS neuro-fuzzy modeling, and genetic optimization, in order to enhance decision-making efficiency under uncertain conditions. To achieve the stated aim, the following objectives are addressed:

1. To develop a mathematical model of the development process for the studied oil field, taking into account the uncertainty of technological parameters.
2. To construct a fuzzy control model for the OFD process.
3. To develop an ANFIS neuro-fuzzy model for forecasting the field's technological indicators.

4. To develop a genetic algorithm for the multi-criteria optimization of operating modes and a hybrid algorithm for managing oil field development processes.

5. To conduct computational experiments and evaluate the effectiveness of the this hybrid approach compared to traditional management methods.

Methods

The object of the study in this work is the Kenkiyak oil field, located in the Aktobe region, which is one of the largest and most significant oil fields in western Kazakhstan. The deep horizons of this field are characterized by high reservoir pressures of up to 80 MPa and reservoir temperatures of about 100 °C, which significantly complicates its development. In addition, the limited volume of geological and field data, the ambiguity of its interpretation results, the stochastic nature of filtration processes, and changes in reservoir fluid properties cause uncertainty, leading to the inefficiency of existing models. In this regard, for the effective management of the Kenkiyak field development processes under uncertainty, it is necessary to apply modern development and production technologies using integrated artificial intelligence methods. To develop models of the development process for the studied field under uncertainty, the following data and information are used:

- available data on the values of: reservoir pressure; oil production and water injection rates; water cut, reservoir permeability, and porosity, which may have a stochastic nature;
- results of theoretical and geophysical studies, and geological and field data;
- expert information obtained in the form of fuzzy data from expert specialists.

Probability theory and mathematical statistics methods are used to collect and process stochastic data [19, 20]. The collection, formalization, and processing of fuzzy expert information and the results of the interpretation of geological and field data, as well as geophysical studies, are performed using expert assessment methods and the mathematical apparatus of fuzzy set theory [11–15, 21–25]. Expert assessment methods in oil field development combine the structured elicitation of professional judgments, multi-criteria decision analysis, and fuzzy logic to evaluate complex fields, asset performance, and risks.

The efficiency of field operation is simultaneously affected by the geological characteristics of the reservoir, the petrophysical properties of the reservoir, reservoir pressure, water cut, operating modes of production and injection wells, as well as economic constraints. Most of the listed factors vary over time and have a nonlinear relationship; consequently, the use of deterministic control methods fails to ensure robust optimal decision-making.

The main objective of the research is to develop a hybrid intelligent control model that integrates fuzzy logic, neuro-fuzzy modeling, and evolutionary optimization methods. The proposed approach is designed to select optimal operating modes for an oil field in the presence of uncertainty in the initial information.

Let the vector of input technological parameters be of the form $x = (x_6)$, where x_1 is the reservoir pressure, MPa; x_2 is the oil production rate, m³/day; x_3 is the water injection rate, m³/day; x_4 is the water cut of the produced fluids, %; $x_5 -$ is the reservoir permeability, mD; x_6 is the reservoir porosity, %.

The vector of controlled variables is defined as: $u = (u_1, u_2, \dots, u_5)$, where $u_i, i = \underline{1,5}$, characterize the controllable technological interventions: adjustment of pumping equipment capacity; modification of the waterflooding regime; regulation of bottomhole pressure; redistribution of flows among wells; and modification of the well stock operating regime.

The results of the system's operation are described by the vector of output indicators: $y = (y_1, y_2, y_3, y_4)$, where y_1 is cumulative oil production; y_2 is the oil recovery factor; y_3 is the water cut; y_4 is operational costs. Then, in a general form, the control process is described by the operator: $y = F(x, u, t)$, where $F(\cdot)$ is the nonlinear model of the controlled object; t is the field development time.

Since the reservoir parameters are not fully known, the actual model of the object contains uncertainty: $y = F(x, u) + \Delta(x)$, where $\Delta(\mathbf{x})$ characterizes the aggregate uncertainty of the model.

Architecture of the hybrid intelligent system

The proposed hybrid control system integrates three main intelligent methods: a fuzzy inference system; the ANFIS neuro-fuzzy network; and a genetic algorithm. Thus, the field development process control scheme can be represented by the following modules:

Production data → Fuzzy system → ANFIS → Genetic algorithm → Optimal control actions.

The first module is designed for the collection and preprocessing of production information, as well as the conversion of quantitative values of technological parameters into linguistic evaluations. For example,

reservoir pressure is low, normal, high; water cut is low, medium, high; oil production rate is low, medium, high. At this stage, membership functions are used.

The second module is ANFIS, designed to model the relationship: $\hat{y} = f(x, W)$, where $\hat{y}$ is the predicted value of the technological indicator; W is the network parameters.

During the training process, the model parameters are automatically adjusted based on production data. The third module is designed to implement multi-criteria optimization and determine the control action values $\mathbf{u}*$ that ensure the optimal value of the objective function $J(\mathbf{u})\rightarrow$max.

As a result, an optimal oil field development regime is formed. The architecture of the proposed hybrid intelligent system for OFD process management is shown in Fig. 1.

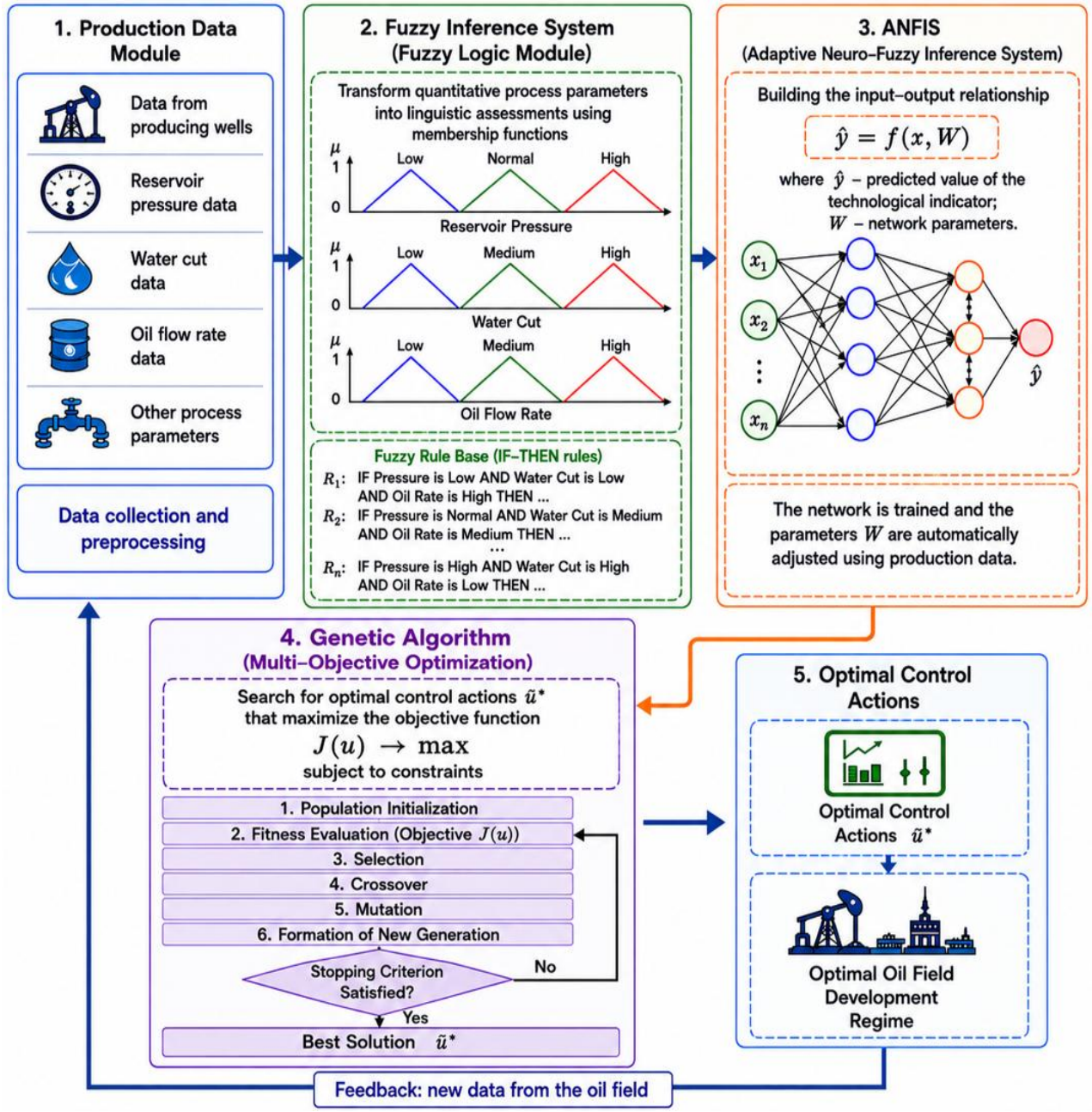


Fig.1 Architecture of the hybrid intelligent system for OFD process management.

Mathematical model of the OFD process

Oil field development is considered as a multi-criteria control problem.

The criteria vector has the form $F = (F_4)$, where the local criteria are: $F_1 = Q_o$ is cumulative oil production; $F_2 = K_r$ is the oil recovery factor; $F_3 = -W$ is the water cut of produced fluids; $F_4 = -C$ is operational costs. Then, the maximization of the vector can be written as: max $F = (Q_o, K_r, -W, -C)$.

To obtain a single optimal solution, a weighted sum of criteria is used: $J = \sum_{i=1}^{4} \omega_i F_i$, where $\sum_{i=1}^{4} \omega_i = 1$, and ω_i are the importance coefficients of the local criteria.
If it is required to increase the priority of $F_1 = Q_o$, i.e., oil production, then $\omega_1 > \omega_2 > \omega_3 > \omega_4$. is selected.
The development process is constrained by the following technological conditions:

- `reservoir pressure constraint:` $P_{min} \le P \le P_{max}$;
- `water cut constraint:` $W \le W_{max}$;
- `well productivity constraint:` $Q_i \le Q_i^{max}$, where $i = 1, \dots, n$ is the well number;
- `injection well stock constraint:` $Q_{inj} \le Q_{inj}^{max}$, where the subscript *inj* stands for injection;

Q_{inj} is the injection volume.

In the presence of uncertainty in technological parameters, a fuzzy description of the constraints is introduced.

For each constraint, a membership function $\mu_i(x): X \to [0,1]$.
For example, for reservoir pressure

$$\mu_P(P) = \{0, P < P_{min}, \frac{P - P_{min}}{P^* - P_{min}}, P_{min} \le P \le P^*, 1, P \ge P^*,$$

where P^* is the preferred operating pressure.

Similarly, membership functions for the remaining technological parameters are defined. The overall degree of satisfaction of all constraints is calculated as $\mu(x) = min(\mu_1, \mu_2, \dots, \mu_m)$.

The use of the maximin principle makes it possible to select a solution that ensures the maximum degree of simultaneous satisfaction of all technological constraints, $max\ \mu(x)$.
Thus, the optimization problem takes the form: $max\{J(u),\ \mu(x)\}$, which corresponds to the formulation of the multi-criteria control problem under uncertainty.

The proposed mathematical model combines a deterministic description of the field development process with a fuzzy representation of the uncertainty in technological parameters. This formulation allows for taking into account both quantitative production data and expert knowledge when selecting optimal operating modes.

Fuzzy control model for the OFD process

The efficiency of OFD process management is largely determined by the completeness and reliability of the initial data. However, under real operating conditions, many parameters are characterized by uncertainty caused by reservoir heterogeneity, measurement errors, variability in the petrophysical properties of the reservoir, and the subjectivity of expert estimates. Therefore, this hybrid approach utilizes the fuzzy logic framework, which allows for the formalization of expert knowledge and the accounting of uncertainty in technological information.

The most significant technological parameters have been selected as the input variables of the system: $X = \{P,\ Q_o,\ Q_{inj},\ W,\ k,\ \phi\}$, where P is the reservoir pressure, MPa; Q_o is the oil production rate, m³/day; Q_{inj} is the water injection rate, m³/day; W is the water cut, %; k is the reservoir permeability, mD; ϕ is the reservoir porosity, %.

The output variable is the integral assessment of the operating mode efficiency: $Y = E$, where E is the quality of the selected field development mode.

Membership functions

Each input variable is represented by a set of linguistic terms. For example, for reservoir pressure, the following set is used: $T(P) = \{Low,\ Medium,\ High\}$.

Gaussian membership functions are applied to describe the terms $\mu(x) = exp\left(-\frac{(x-c)^2}{2\sigma^2}\right)$, where c is the center of the membership function; σ is the width coefficient of the function.

For some parameters, triangular membership functions can be used.

$$\mu(x) = \{0, x \le a, \frac{x - a}{b - a}, a < x \le b, \frac{c - x}{c - b}, b < x < c, 0, x \ge c,$$

where x is the value of the considered parameter (input variable) for which the degree of membership in the fuzzy set is calculated; a is the left boundary of the membership function; b is the peak of the triangle, where the degree of membership is maximum. For $x \ge c$ the degree of membership again becomes zero. On the

interval from a to b, the function value increases linearly from 0 to 1; at point b, the maximum is reached ($\mu = 1$); on the interval from b to c, the function decreases linearly from 1 to 0.

The use of various types of membership functions allows for the most adequate description of actual technological dependencies.

Fuzzy rule base. Expert knowledge is formalized using Mamdani-type rules. A typical rule has the form

$$R_i: \textbf{IF } P = High \textbf{ and } Q_o = High \textbf{ and } W = Low, \textbf{THEN } E = Excellent.$$

... ...

$$R_j: \textbf{IF } P = Low \textbf{ and } W = High, \text{ THEN } E = Poor.$$

The complete knowledge base contains $N = \prod_{i=1}^{m} n_i$ rules, where m is the number of input variables; n_i is the number of terms for each variable.

Rule aggregation. The degree of activation of each rule is determined by the minimum operator $w_i = min(\mu_1, \mu_2, \dots, \mu_n)$. Alternatively, the product can be used: $w_i = \prod_{j=1}^{n} \mu_j$.

Defuzzification. After aggregating all active rules, the resulting fuzzy set is converted into a numerical value.

The center of gravity method is used: $E = \frac{\int y\,\mu(y)\,dy}{\int \mu(y)\,dy}$, where E is the resulting (crisp) value of the output variable after defuzzification; y is the value of the output variable (integral assessment of the operating mode efficiency); $\mu(y)$ is the membership function of the output variable, indicating the degree of membership of the value y to the corresponding fuzzy set, $0 \le \mu(y) \le 1$; dy is the infinitesimal increment of the variable y used in integration. It signifies that the integral is computed over the entire range of values of the output variable.

The obtained value serves as input data for the neuro-fuzzy model.

ANFIS neuro-fuzzy model

Despite the high efficiency of classical fuzzy systems, the quality of control significantly depends on the choice of membership functions and the rule base. To automatically tune these parameters, the Adaptive Neuro-Fuzzy Inference System (ANFIS) is used in this work.

The trainable model is defined by the expression: $\hat{y} = f(x, W)$, where x is the input vector; W is the set of model parameters; $\hat{y}$ is the predicted technological indicator.

ANFIS consists of five sequential layers.

First layer. The values of the membership functions are calculated: $O_i^{(1)} = \mu_i(x)$.

Second layer. The rule weights are calculated: $w_i = \prod_{j=1}^{m} \mu_{ij}(x_j)$.

Third layer. Normalization is performed: $\bar{w}_i = \frac{w_i}{\sum_{k=1}^{N} w_k}$.

Fourth layer. For the Sugeno model, the output of each rule is defined by a linear function:

$$f_i = a_i P + b_i Q_o + c_i Q_{inj} + d_i W + e_i k + g_i \phi + r_i.$$

Then $O_i^{(4)} = \bar{w}_i f_i$.

Fifth layer. The overall output of the model is: $\hat{y} = \sum_{i=1}^{N} \bar{w}_i f_i$.

Thus, the network automatically captures the nonlinear relationship between the technological parameters and the field development performance indicators.

Network training. The model parameters are determined by minimizing the quadratic error function:

$$J(W) = \frac{1}{2} \sum_{k=1}^{M} (\hat{y}_k)^2.$$

Parameter updating is carried out using a hybrid algorithm $W^{(1)} = W^{(t)} - \eta \nabla J(W)$, where $W^{(1)}$ is the new value of the model parameter vector after completing the next training step (iterations $t + 1$). These are the updated parameters to be used in the subsequent iteration; W_t is the current value of the model parameter vector at iteration t. The vector W includes all tunable parameters of the model, such as: neural network weights; membership function parameters (centers, widths, etc.) in ANFIS; η is the learning rate. It is a coefficient that determines the magnitude of the change in the model parameters per iteration. Typically, $0 < \eta < 1$. If the value of η is too large, the algorithm may fail to converge; if it is too small, the training process

will be very slow; $\nabla J(W)$ is the gradient of the error function with respect to the parameters W. It represents a vector of partial derivatives of the loss function $J(W)$:

$$\nabla J(W) = \left[\frac{\partial J}{\partial w_1}, \frac{\partial J}{\partial w_2}, \dots, \frac{\partial J}{\partial w_n}\right]^T.$$

The gradient indicates the direction of the steepest increase in the error function. Therefore, it is subtracted during parameter updating to minimize the model error.

Physical meaning of the formula. At each training iteration, the model parameters are adjusted in the direction opposite to the gradient of the error function:

$$New\ parameters = Current\ parameters - Learning\ rate \times Error\ gradient.$$

Thus, the algorithm gradually decreases the value of the error function $J(W)$, thereby improving the model's accuracy.

For ANFIS. When considering the hybrid learning algorithm for ANFIS, the given formula is applied only to update the premise parameters (i.e., the membership function parameters). The consequent parameters of the Sugeno rules are typically calculated using Least Squares Estimation (LSE). Therefore, the complete hybrid algorithm consists of two phases:

Forward pass is calculation of the consequent parameters using the least squares method.

Backward pass is updating of the membership function parameters using the formula:

$$W^{(1)} = W^{(t)} - \eta \nabla J(W).$$

It is precisely this combination of methods that is referred to as the *hybrid learning algorithm for ANFIS.*

The least squares method is applied to the linear coefficients, while the membership function parameters are adjusted using the backpropagation method. This combined algorithm ensures a high training speed and model stability.

Genetic algorithm for multi-criteria optimization

After forecasting the technological indicators using ANFIS, the search for the optimal field operating mode is performed.

Each individual in the population is described by the vector: $z = (P, Q_o, Q_{inj}, W, k, \phi)$.

Fitness function. The objective function is defined as $F = \omega_1 Q_o + \omega_2 K_r - \omega_3 W - \omega_4 C$, where Q_o is oil production; K_r is the oil recovery factor; W is the water cut; $C -$ is operational costs; ω_i are the weight coefficients.

Constraints. Optimization is carried out subject to the following conditions:

$$P_{min} \le P \le P_{max}, Q_o \le Q_o^{max}, W \le W_{max}, Q_{inj} \le Q_{inj}^{max}.$$

Selection. The probability of selecting an individual is determined by the expression: $p_i = \frac{F_i}{\sum_{j=1}^{N} F_j}$.

Crossover. Offspring are generated according to the rule: $z' = \alpha z_1 + (1 - \alpha) z_2$, where z' is the offspring (new solution) obtained as a result of the crossover of two parent individuals. It is a new parameter vector that inherits the traits of both parents; α is the blending coefficient (crossover coefficient). It determines the proportion of traits the offspring inherits from the first and second parents. Its value lies in the range $0 < \alpha < 1$. If $\alpha = 0.5$, the offspring is the exact average of the two parents; if $\alpha > 0.5$, the offspring is more similar to the first parent; if $\alpha < 0.5$, the offspring is more similar to the second parent.

z_1 is the first parent, representing the parameter vector (chromosome) of the first solution; z_2 is the second parent, representing the parameter vector (chromosome) of the second solution.

Mutation. The mutation operation is described by the expression: $z_i' = z_i + \delta$, where z_i' is the value of the i-th parameter (gene) after mutation. It is the new value obtained after introducing a random change to the original gene; z_i is the original value of the i-th parameter (gene) before the mutation operation is performed. It belongs to the current individual (chromosome); δ is a random perturbation (mutational change) added to the gene value. It can be either positive or negative, allowing the parameter value to be randomly increased or decreased. $\delta \sim N(0, \sigma^2)$ denotes the normal (Gaussian) distribution of the random variable δ with an expected value $\mu = 0$, indicating the absence of systematic bias (on average, the change is zero); a variance of σ^2, which

determines the spread of possible changes around zero; and a standard deviation $\sigma = \sqrt{\sigma^2}$, characterizing the average magnitude of the random change.

The mutation operation modifies an individual solution parameter by adding a small random deviation: *New gene value = Original gene value + Random perturbation.*

Mutation enables the genetic algorithm to explore new regions of the solution space, prevents premature convergence of the population, and facilitates the search for the global optimum.

Stopping criterion. The optimization terminates when the following condition is met: $| F_{t+1} - F_t | < \varepsilon$, or after reaching the maximum number of generations $t = t_{max}$, where F_{t+1} is the value of the objective or fitness function at the next optimization iteration $(t + 1)$. It characterizes the quality of the best found solution after the next step of the algorithm; F_t is the value of the objective (or fitness) function at the current iteration t. This value is used for comparison with the result of the next iteration; ε is the specified allowable error (convergence threshold). It is a small positive number that determines the minimum change in the objective function value at which the algorithm is considered to have converged. Typically, ε > 0, for example, $\varepsilon = 10^{-3},\ 10^{-4}\ or\ 10^{-6}$, depending on the required solution accuracy.

The criterion $| F_{t+1} - F_t | < \varepsilon$ means that the change in the objective function value between two consecutive iterations has become so small that further optimization practically does not improve the solution. In this case, the algorithm terminates, and the found solution is accepted as optimal (or near-optimal).

Hybrid algorithm for OFD process management

The hybrid approach integrates the three considered methods (based on the fuzzy model, the ANFIS neuro-fuzzy model, and the genetic algorithm) into a single intelligent system for OFD process management. The operational scheme of the proposed hybrid algorithm for OFD process management is shown in Fig. 2.

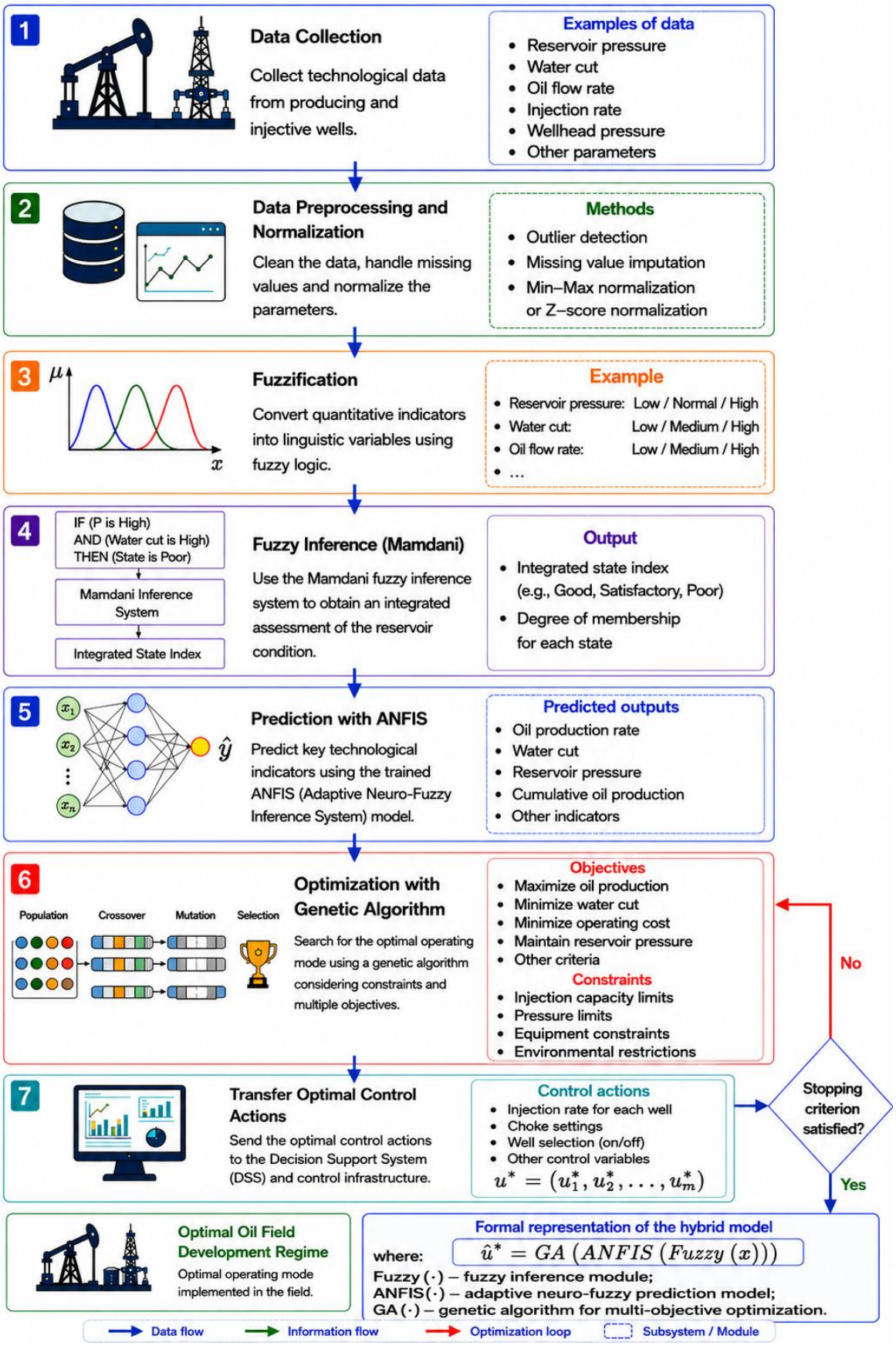


Fig.2 Operational scheme of the hybrid algorithm for OFD process management.

Thus, the operation of the hybrid algorithm for OFD process management can be represented by the following sequence:

1. Collection of technological data from production and injection wells.
2. Data preprocessing and parameter normalization.

3. Conversion of quantitative indicators into linguistic variables using fuzzy logic.
4. Generation of an integral assessment of the field state using the Mamdani fuzzy inference system.
5. Forecasting of technological indicators using the trained ANFIS neuro-fuzzy network.
6. Search for the optimal field operating mode using a genetic algorithm, taking into account technological constraints and efficiency criteria.
7. Transmission of optimal control actions to the decision support system.

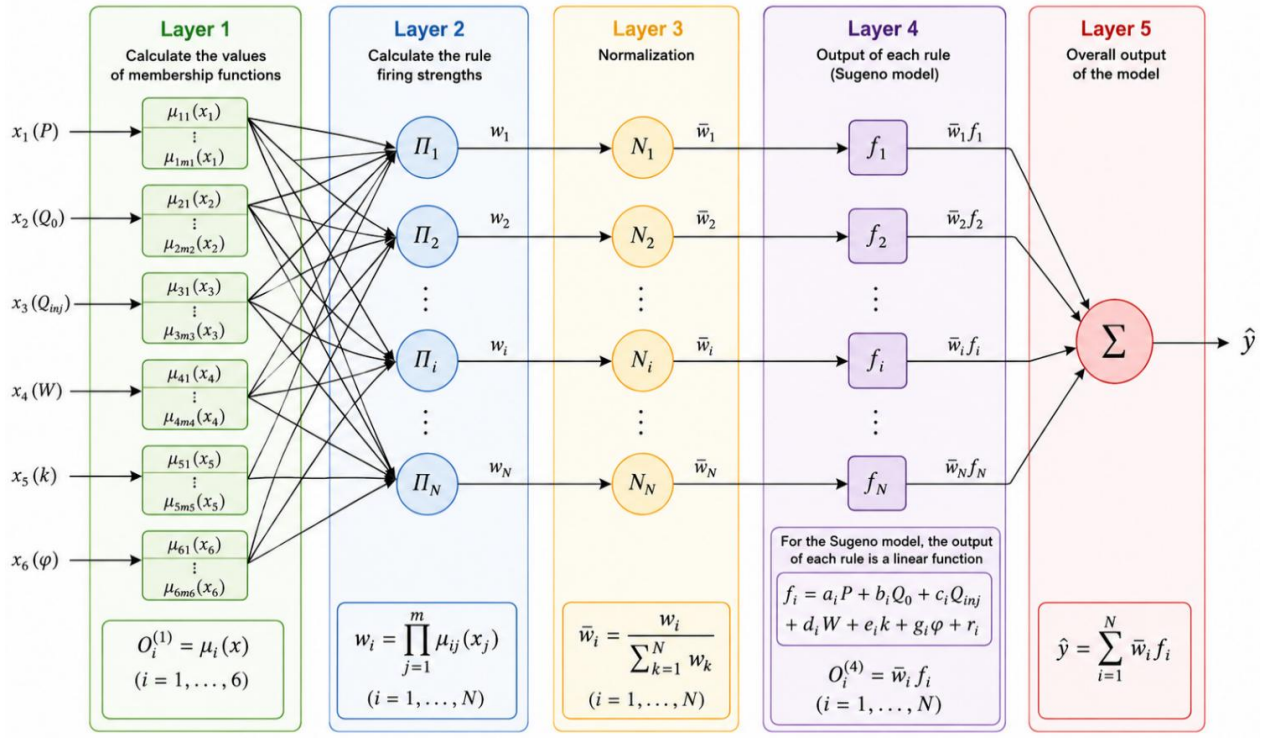
Formally, the hybrid model can be represented by the operator $u^* = GA\big(ANFIS(Fuzzy(x))\big)$, where $Fuzzy\ (\cdot)$ is the fuzzy inference module; $ANFIS\ (\cdot)$ is the adaptive neuro-fuzzy forecasting model; $GA\ (\cdot)$ is the genetic algorithm for multi-criteria optimization.

This hybrid approach combines the advantages of expert knowledge, adaptive machine learning, and evolutionary optimization. Fuzzy logic enables the formalization of uncertainty in the initial data; ANFIS allows for the automatic adaptation of the model to changing field development conditions; and the genetic algorithm ensures an efficient search for globally optimal operating modes. The integration of these methods forms an intelligent control system characterized by high robustness to uncertainty and capable of making effective decisions under the complex geological and technological conditions of oil field development.

Results

Architecture of the Adaptive Neuro-Fuzzy Inference System (ANFIS)

The constructed ANFIS architecture, comprising the five sequential layers described in the previous section, is presented in Fig. 3.



x_1 is reservoir pressure; x_2 is oil production rate; $x_3$is water injection rate; x_4- water cut; x_5- reservoir permeability; x_6- reservoir porosity; m1=m2=m3=m4=m5=m6=3; Total number of rules N=3^6

Fig.3 Architecture of the ANFIS adaptive neuro-fuzzy model for managing the oil field development process.

Results of the computational experiment

To evaluate the effectiveness of this hybrid approach, a computational experiment was conducted simulating the oil field development management process under conditions of uncertainty in geological and technological

parameters. The study was performed based on the production data of the Kenkiyak oil field, the characteristics of which correspond to the typical operating conditions of the fields in Western Kazakhstan.

The main input parameters of the model included reservoir pressure P, oil production rate Q_o, volume of injected water Q_{inj}, water cut W, permeability k and reservoir porosity ϕ. For each parameter, variation ranges were defined to reflect actual production operating conditions. The initial values of the oil field's technological parameters are presented in Table 1.

Table 1 Initial values of the oil field technological parameters.

Parameter	Symbol	Unit of measurement	Min	Max	Average
Reservoir pressure	P	MPa	16	28	22
Oil production rate	Q_o	m³/day	40	210	125
Water injection rate	Q_{inj}	m³/day	50	250	145
Water cut	W	%	15	85	47
Permeability	k	mD	45	620	280
Porosity	ϕ	%	10	29	19

To form the training and test samples, 1000 records of production data were used. Of these, 70% were used for model training, 15% for validation, and 15% for testing.

To implement the proposed approach, the parameters of the fuzzy system, the ANFIS neuro-fuzzy network, and the genetic algorithm were tuned, as presented in Table 2.

Table 2 Parameters of the proposed hybrid model.

Component	Parameter	Value
Fuzzy system	Number of input variables	6
Fuzzy system	Number of linguistic terms	3
Fuzzy system	Number of rules	729
ANFIS	Membership functions	Gaussian
ANFIS	Number of epochs	100
ANFIS	Learning algorithm	Hybrid
Genetic algorithm	Population size	100
Genetic algorithm	Number of generations	200
Genetic algorithm	Crossover probability	0.85
Genetic algorithm	Mutation probability	0.05

The use of three linguistic terms for each variable ("low", "medium", "high") allowed the formation of a knowledge base consisting of $N = 3^6 = 729$ fuzzy rules.

Scheme of the computational experiment. The computational experiment was conducted in the following stages:

1. Collection and preprocessing of production data.
2. Formation of the fuzzy knowledge base.
3. Training of the ANFIS neuro-fuzzy network.
4. Forecasting of technological indicators.
5. Search for optimal operating modes using the genetic algorithm.
6. Analysis of the robustness and sensitivity of the obtained solutions.

The scheme of the computational experiment is presented in Fig. 4.

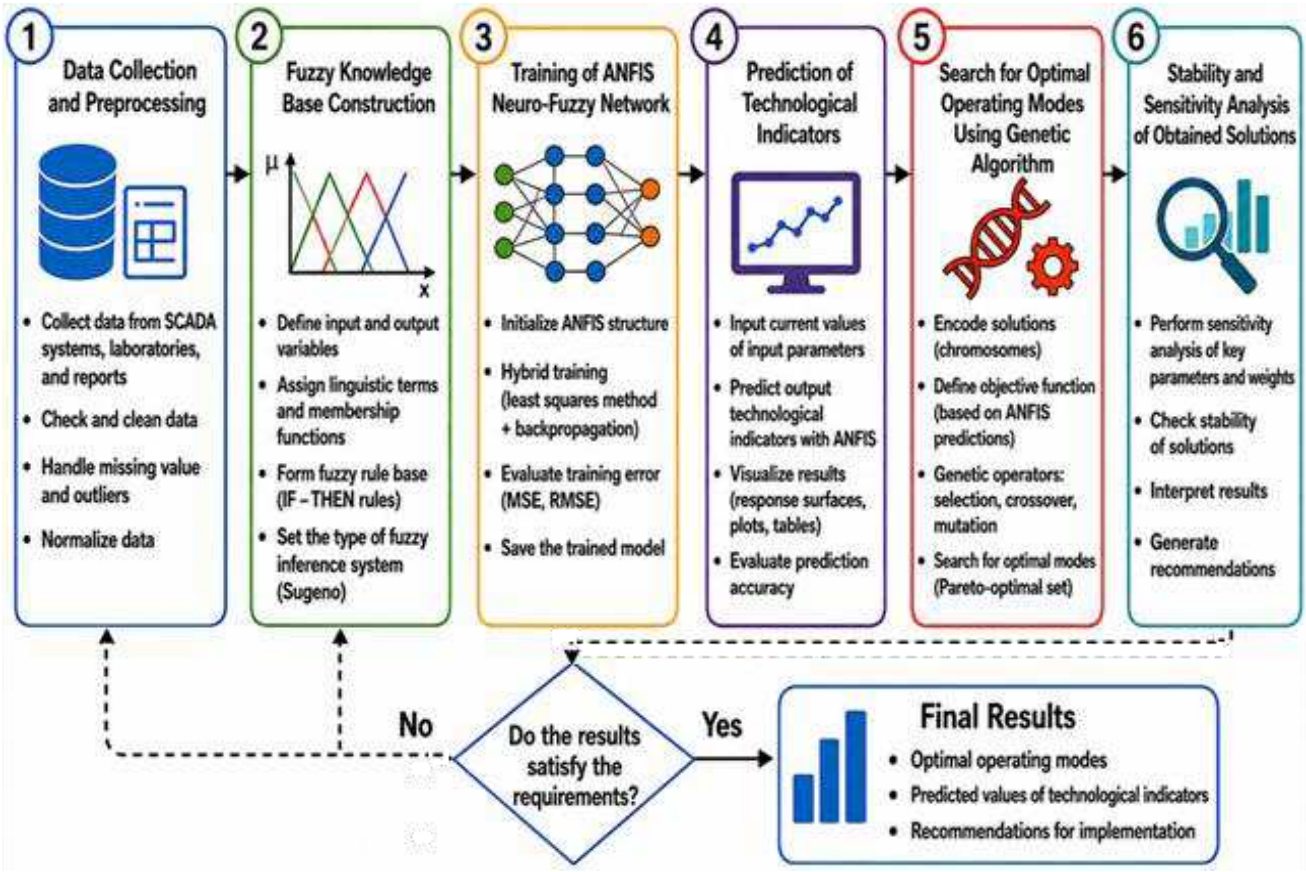


Fig. 4 Operational scheme of the proposed hybrid intelligent control system for oil field development.

ANFIS modeling results

After training the ANFIS, an evaluation of the forecasting quality for the oil recovery factor was conducted. The error function was defined by the expression:

$$MSE = \frac{1}{N}\sum_{i=1}^{N} (y_i - \hat{y}_i)^2,$$

where MSE is the mean squared error, an indicator of the ANFIS model's forecasting quality, characterizing the average value of the squared deviations between the actual and predicted values. The lower the MSE value, the higher the model's accuracy; N is the number of observations (experimental data, examples, or test samples) used to calculate the error; y_i is the actual (experimentally measured) value of the oil recovery factor for the i-th observation; $\hat{y}_i$ is the predicted (calculated by the ANFIS model) value of the oil recovery factor for the i-th observation; i is the observation index, taking values from 1 to N.

During the training process, a steady decrease in the error was observed. The main results of the ANFIS model training are presented in Table 3.

Table 3 Results of the ANFIS model training.

Epoch	MSE
0	0.512
10	0.238
20	0.149
40	0.091
60	0.061
80	0.044
100	0.037

The obtained results demonstrate the high efficiency of the hybrid learning algorithm. After 100 epochs, the mean squared error decreased more than 13-fold. Figure 5 presents the model's learning curve, illustrating the training process: the mean squared error versus the number of epochs.

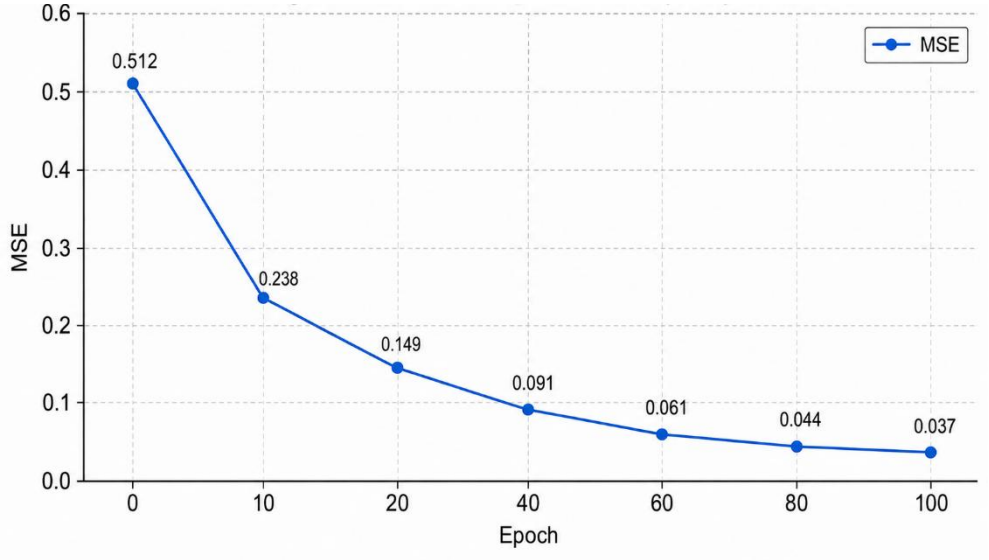


Fig. 5 ANFIS model learning curve, convergence of the training process.

The obtained graph demonstrates a rapid decrease in the error during the initial stages of training, followed by a subsequent stabilization of the optimization process.

Forecasting of the oil recovery factor. The forecasting results on the test set are presented in Table 4.

Table 4 Forecasting results using the ANFIS model

Example	Actual value of K_r	Predicted value of K_r	Error (%)
1	0.321	0.317	1.25
2	0.354	0.361	1.98
3	0.387	0.381	1.55
4	0.412	0.406	1.46
5	0.438	0.445	1.60

The average *relative forecasting error* does not exceed 2%, confirming the high accuracy of the constructed model.

Figure 6 presents a comparison of the actual and predicted values of the oil recovery factor using the ANFIS model.

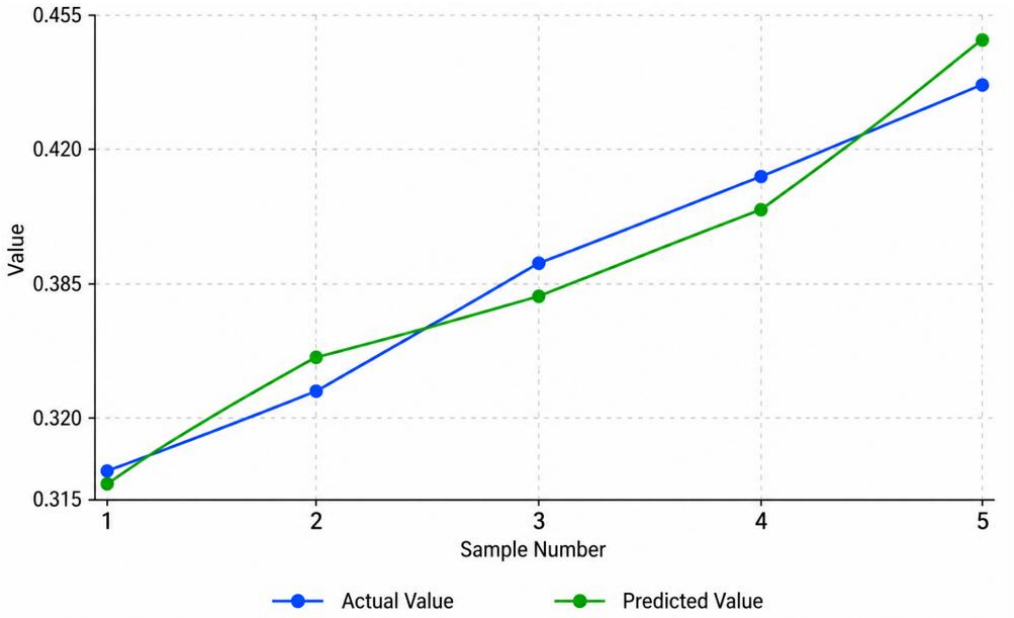


Fig. 6 Comparison of the actual and predicted values of the oil recovery factor using the ANFIS model.

Analysis of the graph demonstrates that the predicted values of the oil recovery factor practically coincide with the actual values. The maximum deviation does not exceed 2%, and the mean relative error is approximately 1.57%, indicating the high accuracy of the ANFIS model. Both curves exhibit the same growth trend, confirming the model's ability to adequately reproduce the patterns of oil recovery factor variations and its applicability for forecasting under various field development regimes. The nearly perfect overlap of the curves demonstrates the model's capability to accurately capture the nonlinear patterns of the field development process.

Comparison with existing methods

To evaluate the effectiveness of the proposed hybrid ANFIS+GA+Fuzzy method, a comparison was conducted with the following established methods: fuzzy system (Fuzzy); artificial neural network (ANN); genetic algorithm (GA); and the ANFIS neuro-fuzzy system.
The following criteria were used for comparison:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2};\ MAE = \frac{1}{N}\sum_{i=1}^{N} \mid y_i - \hat{y}_i \mid;\ R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}.$$

The comparison results are presented in Table 5.

Table 5 Comparison of the forecasting accuracy of existing methods and the proposed hybrid method

Method	RMSE	MAE	R^2
Fuzzy	0.084	0.071	0.891
ANN	0.062	0.050	0.925
GA	0.058	0.046	0.931
ANFIS	0.041	0.033	0.962
Proposed hybrid method	0.029	0.022	0.983

The obtained results demonstrate that this proposed hybrid approach yields the lowest RMSE and MAE values, as well as the highest coefficient of determination, confirming the effectiveness of the proposed method. For example, compared to the classical fuzzy system, the forecasting error decreased by approximately $\frac{0.084-0.029}{0.084} \times 100 = 65.5\%$. At the same time, the coefficient of determination increased from 0.891 to 0.983.

Thus, the proposed hybrid method, by integrating fuzzy logic, ANFIS, and a genetic algorithm, provides a significant improvement in forecasting quality and the efficiency of oil field development process management.

Sensitivity analysis and comparative evaluation of the effectiveness of the hybrid approach

Following the evaluation of forecasting accuracy, the robustness of the proposed hybrid model to changes in optimization priorities and the uncertainty of technological parameters was assessed. The objective of the sensitivity analysis was to identify the factors that have the greatest impact on the efficiency of oil field development process management, as well as to evaluate the robustness of the obtained optimal solutions.

The main sensitivity factors considered were: the weight coefficient for oil production, ω_1; the weight coefficient for the oil recovery factor, ω_2; the weight coefficient for the water cut, ω_3; and the weight coefficient for operational costs, ω_4. The vector of weight coefficients is expressed as $\omega = (\omega_1, \omega_2, \omega_3, \omega_4)$, where the sum of the weight coefficients equals 1.

The baseline calculation scenario corresponded to the values ω = (0.40, 0.30, 0.20, 0.10). Each coefficient was varied within a range of ±20%, while the remaining coefficients were automatically normalized to maintain the condition $\sum_{i=1}^{4} \omega_i = 1$.

To quantitatively evaluate the sensitivity, a relative coefficient was used: $S_i = \frac{\Delta F/F}{\Delta\omega_i/\omega_i}$.

Here S_i is the relative sensitivity coefficient of the objective function to the change in the i-th weight coefficient; $F$is the initial value of the objective function; $\Delta F$is the change in the objective function after varying the parameter; ω_i is the initial value of the i-th weight coefficient; $\Delta\omega_i$is the change in this weight coefficient. The coefficient S_i characterizes the relative change in the objective function per unit relative change in the corresponding model parameter. The larger the value of S_i, the stronger the influence of the considered parameter on the control quality.

Influence of the weight coefficients. The calculation results are presented in Table 6.

Table 6 Sensitivity analysis of optimization criteria

Weight coefficient	Variation	Change in objective function	Sensitivity
ω_1	±20 %	±13.8 %	High
ω_2	±20 %	±8.1 %	Moderate
ω_3	±20 %	∓6.3 %	Moderate
ω_4	±20 %	∓4.4 %	Low

The results presented in this table show that the weight coefficient for oil production, ω_1, has the greatest influence on the optimal solution. This is explained by the fact that oil production is the primary optimization criterion and directly determines the value of the objective function.

The weight coefficient for the oil recovery factor, ω_2 has a moderate influence. Its increase leads to the selection of more efficient waterflooding regimes that ensure an increase in ultimate oil recovery; however, it is accompanied by a certain decrease in current production. The weight coefficients for water cut and operational costs have a less pronounced influence on the final solution. This indicates the high robustness of the developed algorithm to changes in secondary criteria.

Analysis of the influence of input parameter uncertainty. To evaluate the robustness of the model, computational experiments were conducted at various levels of uncertainty in the initial data.

Let $x_i = x_i^0 + \delta_i, where\ x_i^0$ is the nominal value of the parameter; δ_i is the random deviation. This means that the actual value of the parameter equals the nominal value plus a random measurement error. It was assumed that the random deviation of the i-th parameter follows a normal distribution with a zero mean and a variance of σ^2, i.e. $\delta_i \sim N(0, \sigma^2)$, where σ characterizes the level of measurement uncertainty. The zero mean indicates the absence of systematic measurement error, whereas the parameter σ characterizes the level of random uncertainty and the measurement accuracy.

In the experiments, four noise levels were considered: 0%; 5%; 10%; and 15%. The results of the robustness evaluation of the proposed hybrid model are presented in Table 7.

Table 7 Robustness of the proposed hybrid model.

Noise level	RMSE	MAE	R^2
0 %	0.029	0.022	0.983

5 %	0.031	0.024	0.979
10 %	0.035	0.028	0.972
15 %	0.041	0.033	0.964

Analysis of the model's robustness to noise demonstrated that as the level of random perturbations increased from 0% to 15%, the RMSE value increased from 0.029 to 0.041. Despite the decrease in forecasting accuracy, the error growth remains moderate, indicating the high robustness of the ANFIS model and its ability to maintain adequate forecasting quality under conditions of input data uncertainty. Furthermore, as can be seen from Table 7, even at an uncertainty level of 15%, the coefficient of determination remains above 0.96, confirming the high stability of the hybrid model.

Figure 7 presents a graph showing the dependence of RMSE on the level of input data uncertainty. It can be seen from Fig. 7 that the error increases almost linearly with increasing uncertainty. At the same time, the rate of error growth remains relatively small, confirming the high robustness of the developed algorithm.

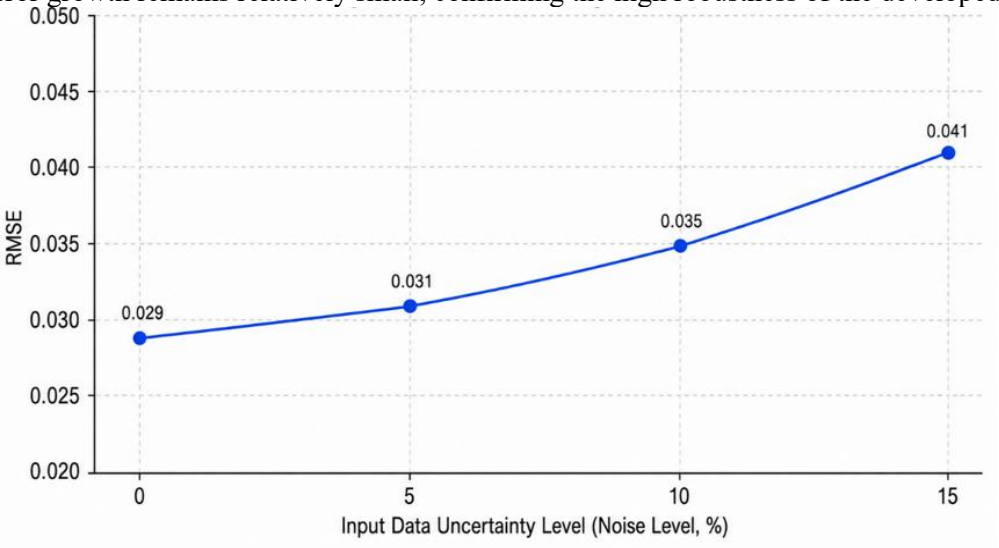


Fig. 7 Effect of input data uncertainty on the forecasting accuracy of the proposed hybrid model

Pareto analysis of multi-criteria optimization

To investigate the trade-offs between conflicting criteria, the Pareto principle was applied.
The optimal set is defined as:

$$P = \{x \in X \mid \nexists x' : F(x') \succ F(x)\},$$

where P is the Pareto set (of Pareto-optimal solutions); x is the considered solution; X is the set of all feasible solutions satisfying the constraints; $x \in X$ means that the solution x belongs to the set of feasible solutions; $\mid$ means "such that"; $\nexists$ means "there does not exist"; x' is any other feasible solution; $F(x)$ is the vector of values of all criteria for the solution x. $F(x') \succ F(x)$ means that the solution x' dominates the solution x, that is: it is no worse across all criteria and strictly better in at least one criterion.

Consequently, the Pareto set consists of all feasible solutions for which there does not exist another feasible solution that outperforms them across all criteria simultaneously. In the vector of objective criteria $F(x)$:

$$F(x) = (Q_o,\ K_r,\ -W,\ -C),$$

Q_o is the oil production volume (or oil production rate), which must be maximized; K_r is the oil recovery factor, which also must be maximized; is the water cut of the produced fluids, which must be minimized; C is the total operational costs, which also must be minimized.

The "minus" signs in $-W$ and $-C$ are used because, when constructing the Pareto front, all criteria are usually converted to a single form – maximization. That is: maximizing Qo; maximizing K_r; maximizing $-W$ is equivalent to minimizing W; maximizing $-C$ is equivalent to minimizing C.
As *the criterion for selecting the final solution*, the Bellman-Zadeh maximin criterion is applied:

$$max\left\{min\left(\mu_{Q_0}(x), \mu_{K_r}(x), \mu_W(x), \mu_C(x)\right)\right\},$$

where $\mu_{Q_0}(x), \mu_{K_r}(x), \mu_W(x), \mu_C(x)$ are the membership functions describing the degrees of satisfaction for the oil production volume (Q_o), oil recovery factor (K_r), water cut (W) and total operational costs (C); $min(\cdot)$

is the minimum degree of satisfaction among all criteria; $max\{\cdot\}$ is the selection of the solution for which this minimum degree of satisfaction is maximized.

Thus, according to this criterion for selecting the final solution, the degrees of satisfaction for all criteria are calculated for each solution, the worst (minimum) degree of satisfaction is determined, and the solution with the maximum value of this minimum degree is selected from all available solutions. As a result, a Pareto-optimal solution is selected in which the least satisfied criterion achieves the maximum possible degree of satisfaction. This approach ensures the most balanced compromise solution, providing an optimal trade-off among oil production, recovery factor, water cut, and operational costs.

Figure 8 shows the Pareto front obtained by the genetic algorithm.

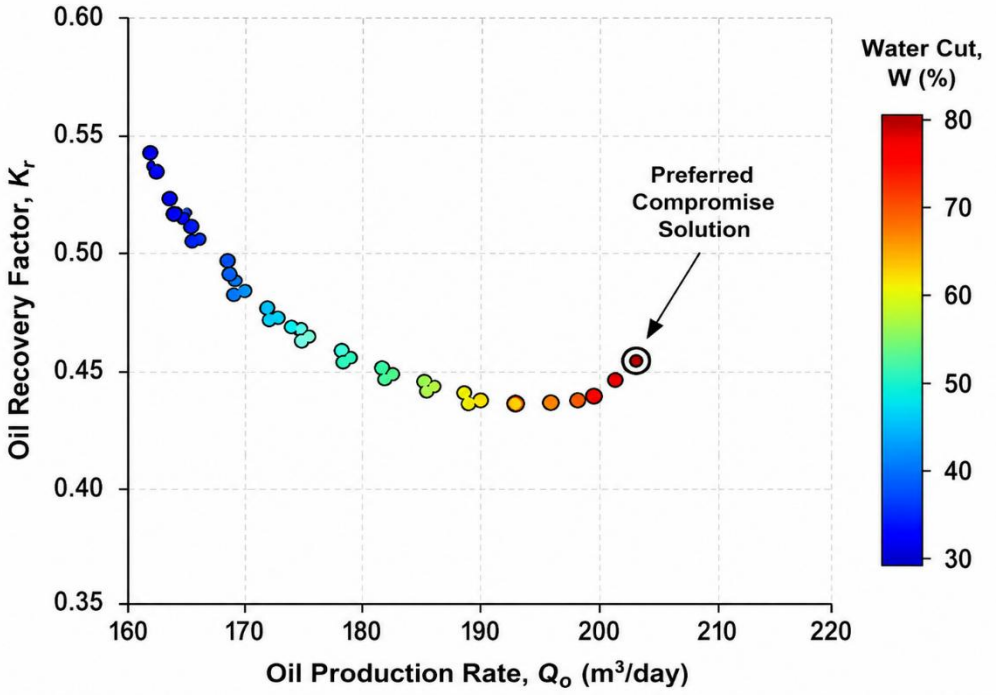


Fig. 8 Pareto front obtained using the proposed hybrid optimization algorithm.

It can be seen from Figure 8 that the proposed method generates a dense Pareto front containing a large number of non-dominated solutions. This provides the decision maker with the opportunity to select an operating regime depending on current production requirements.

Comparison of the effectiveness of various optimization methods. For the final evaluation of the effectiveness of the developed hybrid optimization algorithm, its results were compared with those of established optimization methods.

Table 8 Comparison of optimization results of the compared methods

Method	Oil production Q_o (m³/day)	Recovery factor K_r	Water cut W (%)	Operational costs C (%)	Computation time (s)
Deterministic optimization	183	0.412	52.6	100	**5.8**
Fuzzy Logic	188	0.425	50.9	98.7	8.6
ANN	191	0.431	49.8	97.2	10.9
ANFIS	195	0.444	47.6	95.5	12.7
Proposed hybrid method	**201**	**0.462**	**45.3**	**92.8**	13.4

The results demonstrate that the proposed approach provides the best values across nearly all the evaluated indicators, with the exception of computation time.

Discussion

The obtained results confirm that the integration of fuzzy logic, ANFIS neuro-fuzzy modeling, and a genetic algorithm allows for the effective consideration of various types of uncertainty inherent in oil field development processes. Unlike existing methods based on deterministic, fuzzy, and ANFIS models, as well as

artificial neural networks, the proposed hybrid approach possesses the ability to adapt to changing production conditions through learning from actual data and the utilization of expert knowledge.

The results of the computational experiments confirm that this hybrid approach, based on the integration of fuzzy logic, the ANFIS neuro-fuzzy network, and a genetic algorithm, provides more effective management of oil field development processes compared to traditional methods. Unlike existing approaches, the proposed approach accounts for the uncertainty of geological and technological information, which allows for obtaining more robust and practically applicable solutions.

The conducted comparison with existing methods has shown that the combination of artificial intelligence methods provides a synergistic effect. The fuzzy system formalizes the expert knowledge of specialists and allows for taking into account qualitative characteristics that are difficult to describe using analytical dependencies. In turn, ANFIS automatically adapts the parameters of the membership functions and the rule base based on production data, significantly improving forecasting accuracy. The genetic algorithm ensures an efficient search for the global optimum in a multidimensional solution space and avoids getting trapped in local extrema, which is one of the main drawbacks of gradient optimization methods.

The ANFIS training results demonstrate the model's high capability to approximate complex nonlinear relationships between field development parameters and technological indicators. The reduction of the training mean squared error to a value of 0.037 and the achievement of a coefficient of determination $R^2 = 0.983$ confirm the high quality of forecasting. The obtained metrics surpass the results of individual fuzzy logic models, artificial neural networks, and classical evolutionary algorithms, which is explained by the combined use of adaptation and optimization mechanisms.

The sensitivity analysis of the developed model is of particular interest. The obtained results show that the weight coefficient characterizing the priority of oil production has the greatest influence on the control quality. This is expected, as increasing the weight of this criterion shifts the optimal solution towards the region of maximum production well performance. Conversely, increasing the weight coefficients associated with limiting the water cut and reducing operational costs forms more conservative operating regimes that ensure the long-term efficiency of field development.

The study of the model's robustness to input data uncertainty showed that even with an increase in the level of random perturbations up to 15%, the forecasting quality decreases insignificantly. The value of the coefficient of determination remains above 0.96, and the growth of the RMSE and MAE errors is smooth. This indicates the high robustness of the proposed algorithm and the possibility of its application under conditions of incomplete and inaccurate geological and field data, which is a characteristic feature of most operating oil fields.

The results of the Pareto analysis show that the proposed algorithm is capable of generating a set of compromise solutions, ensuring a balance between productivity, oil recovery, water cut, and costs. Such an approach is particularly important for decision-making under changing economic and technological constraints.

The results of comparing the effectiveness of the proposed hybrid optimization algorithm with the results of existing methods, presented in Table 8, show that the proposed algorithm allows:

- increasing daily oil production by 3.1% to 9.8%;
- increasing the oil recovery factor by 4.1% to 12.1%;
- reducing the water cut by approximately 4.6% to 13.9%;
- decreasing operational costs by approximately 5.5% to 7.2%.

Despite the increase in computation time, the additional computational resource costs are insignificant compared to the achieved improvement in management efficiency.

In total, the obtained results indicate the high effectiveness of this hybrid approach and confirm the prospects of its use in intelligent systems for oil field development management and digital twins of oil and gas facilities.

It should also be noted that there are some limitations to the conducted study. In this work, data from a single studied field and a limited set of technological parameters were considered. In the future, it is advisable to expand the model by taking into account the spatial heterogeneity of the reservoir, geomechanical processes, and multiphase flow, as well as by using deep learning methods, digital twins, and reinforcement learning algorithms. This will make it possible to improve forecasting accuracy and expand the scope of application of the developed approach.

Thus, the results of the study indicate that the integration of fuzzy logic methods, neuro-fuzzy networks, and evolutionary optimization is a promising direction for the development of intelligent technologies for oil field development management under uncertain conditions.

Conclusion

In this study, a hybrid intelligent approach to managing oil field development processes under uncertainty is proposed, based on the integration of fuzzy logic methods, the Adaptive Neuro-Fuzzy Inference System (ANFIS), and a genetic algorithm for multi-criteria optimization. Unlike traditional management methods, the developed approach allows for the simultaneous consideration of quantitative production data and qualitative expert knowledge, which significantly enhances the efficiency of management decision-making.

A mathematical model for intelligent control has been developed, describing the relationship between the technological parameters of field development and operational efficiency indicators. To formalize the uncertainty of geological and technological information, the mathematical apparatus of fuzzy sets was employed, ensuring an adequate representation of expert knowledge regarding the state of the controlled object.

An ANFIS neuro-fuzzy model was constructed, enabling the automatic adaptation of membership functions and rule base parameters based on field data. The conducted computational experiments demonstrated high forecasting accuracy for technological indicators, as evidenced by low RMSE and MAE error values and a high coefficient of determination ($R^2 = 0.983$).

To identify optimal operating modes for the oil field, a multi-criteria optimization genetic algorithm was developed, ensuring an efficient search for globally optimal solutions in the presence of numerous technological constraints. The application of the Pareto principle made it possible to generate a set of compromise solutions balancing oil production volume, oil recovery factor, water cut levels, and operational costs.

The results of the comparative analysis demonstrated that the proposed hybrid approach outperforms classical fuzzy logic methods, artificial neural networks, standalone ANFIS, and traditional genetic algorithms across the primary efficiency criteria. The obtained results indicate improved forecasting accuracy, an increased oil recovery factor, reduced operational costs, and enhanced robustness of the control system against the uncertainty of initial data.

The practical significance of the study lies in the potential application of the developed method in intelligent decision support systems, oil field digital twins, automated oil production management systems, and digital oil and gas enterprises. The implementation of this hybrid approach contributes to improved field operational efficiency, the rational use of natural resources, and the mitigation of production risks.

Promising directions for future research include the development of self-learning intelligent control systems, the integration of deep learning and reinforcement learning methods, the application of oil field digital twins, and the use of distributed computing to optimize the development processes of large oil and gas assets in real time.

Acknowledgements The authors acknowledge that all figures presented in this manuscript were created by the authors.

Funding This research received no external funding.

Author Contributions The authors confirm contribution to the paper as follows: study conception and design: B. Orazbayev, K. Orazbayeva; data collection: A. Zhienbek, L. Kurmangaziyeva; analysis and interpretation of results: B. Orazbayev,. B. Utenova, L. Author; draft manuscript preparation: B. Orazbayev, K. Orazbayeva, L. Kurmangaziyeva, S. Iskakova, V.Makhatova. All authors reviewed the results and approved the final version of the manuscript.

Data availability The anonymized dataset generated and analysed during the current study is available from the corresponding author upon reasonable request. The dataset contains no names, medical record numbers, contact information, dates of birth, or other direct personal identifiers. Access to the data will be considered for academic and research purposes, subject to compliance with applicable ethical and institutional requirements.

Competing Interests The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. Yagafarov, A.K., Kleshchenko, I.I., Zozulya, G.P. *Development of Oil and Gas Fields.* 396 (Tyumen: Tyumen Industrial University, 2022).
2. Philp, R.P. Composition and Properties of Petroleum. In: Wilkes, H. (eds) Hydrocarbons, Oils and Lipids: Diversity, Origin, Chemistry and Fate. *Handbook of Hydrocarbon and Lipid Microbiology. Springer*, *Cham*, 269–310 (2020).
3. Atkinson, C., Isangulov, R.A. Mathematical model of an oil and gas field development process. *European Journal of Applied Mathematics*. **21(3)**, 205–227. https://doi.org/10.1017/S095679251000001X (2010).
4. Corrigan, C.C., Ikonnikova, S.A. A review of the use of AI in the mining industry: Insights and ethical considerations for multi-objective optimization. *The Extractive Industries and Society*. **17**, 1–13. https://doi.org/10.1016/j.exis.2024.101440 (2024).
5. Duqueac, J., Godinhoac, A., Moreiraac, J., Vasconcelos, J. Data Science with Data Mining and Machine Learning A design science research approach. *Procedia Computer Science*. **237**, 245–252. https://creativecommons.org/licenses/by-nc-nd/4.0 (2024).
6. Sitorus, F., Cilliers, J.J., Pablo, R. Brito-Parada, P.R. Multi-criteria decision making for the choice problem in mining and mineral processing: Applications and trends. *Expert Systems with Applications*. **121**, 393-417. https://doi.org/10.1016/j.eswa.2018.12.00 (2019).
7. Rajesh, K., Chang, A.W., Sharma, T.K., Verma, T.K., Prakash, O., Anand, A. *Soft Computing: Theories and Applications.* 308 (Springer Nature Singapore, 2022). https://doi.org/10.1007/978-981-19-0707-4
8. Patel, K.K., Santosh, K.C., Oliveir,a C.G., Patel, A., Ghosh, A. *Soft Computing and Its Engineering Applications.* 317 (Springer Cham., 2024). https://doi.org/10.1007/978-3-031-88042-1
9. Alireza, H., Peter, S. *Application of Soft Computing and Intelligent Methods in Geophysics.* 278 (Springer Geophysics, 2018). https://doi.org/10.1007/978-3-319-66532-0
10. Bajaj, A., Madureira, A.M, Abraham, A. *Hybrid Intelligent Systems.* 469 *(Springer Cham.,* 2025). https://doi.org/10.1007/978-3-031-78925-0
11. Oscar, C., Patricia, M. Hybrid Intelligent Systems Based on Extensions of Fuzzy Logic. *Neural Networks and Metaheuristics.* 245 (Springer Nature Switzerland, 2023). https://doi.org/10.1007/978-3-031-28999-6
12. Chrysafiadi, K. Fuzzy Logic. In: Fuzzy Logic-Based Software Systems. *Learning and Analytics in Intelligent Systems, Springer, Cham.* **34**, 2–24. https://doi.org/10.1007/978-3-031-44457-9_1 (2023).
13. Orazbayev, B., Dyussembina, E., Uskenbayeva, G., Shukirova, A., Orazbayeva, K. Methods for Modeling and Optimizing the Delayed Coking Process in a Fuzzy Environment. *Processes*. **11(2)**, 1-19. https://doi.org/10.3390/pr11020450 (2023).
14. Orazbayev, B.B., Orazbayeva, K.N., Kurmangaziyeva, L.T., Makhatova, V.E. Multi-criteria optimisation problems for chemical engineering systems and algorithms for their solution based on fuzzy mathematical methods. *EXCLI Journal.* **14**, 984–998. https://doi.org/10.17179/excli2015-266 (2015).
15. Kalra, A. Introduction to Fuzzy Logic and Its Applications in Machine Learning. In: Mahajan, S., Rocha, Á., Pandit, A.K., Chawla, P. (eds) Smart Systems: Engineering and Managing Information for Future Success. *Information Systems Engineering and Management, Springer, Cham.* **22**, 1–15. https://doi.org/10.1007/978-3-031-76152-2_1 (2025).
16. Foukalas, F.A. Survey of Artificial Neural Network Computing Systems. *Cogn. Comput.* **17(4)**, 1-18. https://doi.org/10.1007/s12559-024-10383-0 (2025).
17. Alhijawi, B., Awajan, A. Genetic algorithms: theory, genetic operators, solutions, and applications. *Evol. Intel.* **17**, 1245–1256. https://doi.org/10.1007/s12065-023-00822-6 (2024).
18. Alakbari, F.S, Mohyaldinn, M.E, Ayoub, M.A, Muhsan, A.S, Hussein, I.A. A reservoir bubble point pressure prediction model using the Adaptive Neuro-Fuzzy Inference System (ANFIS) technique with trend analysis. *PLoS One.* **17(8)**, 1-20. https://doi.org/10.1371/journal.pone.0272790 (2022).
19. Bronshtein, I.N., Semendyayev, K.A., Musiol, G., Muehlig, H. Probability Theory and Mathematical Statistics. *Handbook of Mathematics. Springer, Berlin.* 743–794. https://doi.org/10.1007/978-3-662-05382-9_16 (2022).
20. Hose, D., Hanss, M. A universal approach to imprecise probabilities in possibility theory. *International Journal of Approximate Reasoning*. **133**, 133-158. https://doi.org/10.1016/j.ijar.2021.03.010 (2021).
21. Krivonogov, A.A., Philippovich, Y.N., Kesel, S.A. Application of the Method of Expert Assessments in Determining the Level of Criticality of the Information Security Vulnerability of Smart Contracts. *Software Engineering Application in Systems Design. Springer, Cham.* **596**, 3–27. https://doi.org/10.1007/978-3-031-21435-6_45 (2023).
22. Rostamian, A., Miranda, M.V., Mirzaei-Paiaman, A., *et al.* Analysis of different objective functions in petroleum field development optimization. *J Petrol Explor Prod Technol.* **14**, 2785–2805. https://doi.org/10.1007/s13202-024-01848-x (2024).

23. Chernyaeva, S., Korobova, L., Ivliev, M., Tolstova, I., Nikitin, B., Matytsina, I. Implementation of the Extrapolation Method of Expert Assessments in Selection Problems. *High-Performance Computing Systems and Technologies in Scientific Research, Automation of Control and Production. HPCST*. *Communications in Computer and Information Science, Springer, Cham.* **1304**, 147–161. https://doi.org/10.1007/978-3-030-66895-2_10 (2020).
24. Clark, T.D., Larson, J.M., Mordeson, J.N., Potter, J.D., Wierman, M.J. Fuzzy Set Theory. In: Applying Fuzzy Mathematics to Formal Models in Comparative Politics. *Studies in Fuzziness and Soft Computing, Springer, Berlin, Heidelberg.* **225**, 29–63. https://doi.org/10.1007/978-3-540-77461-7_2 (2018).
25. Mockor, J., Hynar, D. On Unification of Methods in Theories of Fuzzy Sets, Hesitant Fuzzy Set, Fuzzy Soft Sets and Intuitionistic Fuzzy Sets. *Mathematics*. **9(447)**, 1–26. https://doi.org/10.3390/math9040447 (2021).

Authors and Affiliations

Batyr Orazbayev[1,2], Kulman Orazbayeva[3], Arailym Zhienbek[1], Valentina Makhatova[4], Sandugash Iskakova[5], Balbupe Utenova[5*], Lyailya Kurmangaziyeva[4*]

Lyailya Kurmangaziyeva
Kurmangazieval475@gmail.com
Balbupe Utenova
b.utenova@aogu.edu.kz

[1] Institute of Digital Sciences and Artificial Intelligence, Gumilyov Eurasian National University, Astana, 010008, Kazakhstan
[2] Higher School of Information Technology and Engineering, Astana International University, Astana, 020000, Kazakhstan
[3] Faculty of Applied Sciences, Esil university, Astana, 010005, Kazakhstan
[4] Faculty of Physics, Mathematics and Information Technologies, Kh. Dosmukhamedov Atyrau University, Atyrau, 060000, Kazakhstan
[5] Faculty of Information Technologies, Atyrau Oil and Gas University, Atyrau, 060027, Kazakhstan