

Article

A Traceable Event-Window Evidence-Fusion Workflow for Screening Injection–Production Responses in Mature Waterflood Reservoirs

Jian Li ¹, Peng Cao ¹, Yunfeng Xu ^{2,3,*}, Congying Qiao ¹, Yuhui Zhou ^{2,3}, Yuanhao Zheng ^{2,3}, Hongtu Qian ¹ and Qiaoyu Ge ¹

- ¹ CNPC R&D (DIFC) Company Limited, Dubai P.O. Box 415747, United Arab Emirates; lijian@cnpcrd.com (J.L.); caopeng@cnpcrd.com (P.C.); qiaocongying@cnpcrd.com (C.Q.); qianhongtu@cnpcrd.com (H.Q.); geqiaoyu@cnpcrd.com (Q.G.)
² School of Petroleum Engineering, Yangtze University, Wuhan 430100, China; zhyhtree@yangtzeu.edu.cn (Y.Z.); zy17731528753@gmail.com (Y.Z.)
³ Western Research Institute, Yangtze University, Karamay 834000, China
* Correspondence: xuyunfeng@yangtzeu.edu.cn

Abstract

Mature waterflood reservoirs contain long injection–production histories, yet routine surveillance remains challenged by delayed producer responses and interference from multiple injectors. Existing diagnostic plots, CRM-type analyses, and data-driven models are useful, but they often require preselected patterns, aggregate long-period behavior, or provide limited event-level traceability. This study presents an event-window evidence-fusion workflow (EWEF) for screening candidate injection–production responses without treating the screened objects as causal proof. EWEF uses material injection-rate perturbations as observational triggers and constructs auditable event-window evidence for candidate injector–producer objects. The workflow includes dynamic-record checking, injection-event extraction, candidate-pair construction, pre/post response-window quantification, lag scanning of oil-rate and water-cut responses, static-prior scoring from geometry and transmissibility proxies, relative multi-injector attribution, and rule-model disagreement routing. A transparent rule layer assigns engineering review classes, and a random-forest model trained on rule-derived weak labels serves only as a consistency-review gate. The workflow was applied to three anonymized field waterflood samples comprising 54 dynamic wells, 153 injection perturbation events, 297 candidate injector–producer evidence objects, and 30 final review-list objects. In the weak-label consistency model, lagged oil-rate correlation, water-cut shift, oil-rate shift, static prior, and water-cut lag correlation were recurrent influential input variables. In one representative event, the six-month post-event window showed that oil rate increased by 5.49 m³/d, whereas water cut decreased by 0.52 percentage points relative to the pre-event baseline. Lag scanning yielded modest correlations, with maximum absolute values of approximately 0.34; these correlations are treated as descriptive screening evidence rather than as statistically confirmed connectivity indicators. EWEF therefore provides a traceable workflow for screening and prioritizing field-review candidates. It does not establish causal interwell connectivity. Confirmation requires independent tracer, pressure, intervention, logging, or simulation evidence, and field-specific calibration remains necessary.



Academic Editor: Alberto Maria Gambelli

Received: 14 June 2026

Revised: 9 July 2026

Accepted: 15 July 2026

Published: 21 July 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](https://creativecommons.org/licenses/by/4.0/)

[Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: waterflood surveillance; injection-production response; event-window analysis; interwell connectivity; capacitance-resistance model; mature waterflood reservoir; uncertainty routing

1. Introduction

Identifying injection–production responses remains a routine surveillance task in mature waterflood reservoirs, but reliable screening is difficult. Production responses are seldom synchronous with injection adjustments, and observed signals may be obscured by reservoir heterogeneity, operational interruptions, multi-injector interference, and high water cut. Recent waterflood studies have therefore extended beyond field-scale performance tracking toward injection–production relationship analysis, interwell connectivity characterization, and data-assisted surveillance workflows [1–4]. Reservoir-flow simulation and development-measure evaluation studies across waterflood, fracturing, and geothermal settings further show that subsurface responses depend on multiphase flow, heterogeneity, operational measures, and fracture morphology [5–9].

Capitance-resistance models, INSIM-derived injection-efficiency workflows, probabilistic connectivity analyses, and related proxy models have improved the scalability of waterflood evaluation and optimization [10–18]. These methods are useful for pattern-level diagnosis and operational management, and several operationally tuned variants can already condition the analysis on time windows, injection-rate changes, or prior connectivity assumptions. The present workflow is therefore not positioned as a replacement for CRM-type or INSIM-type models, nor as proof that event-window construction is methodologically superior to those approaches. Its narrower purpose is to preserve a traceable chain from an individual injection perturbation to a bounded producer response and to pass that evidence object to field review or later CRM, tracer, pressure, logging, or simulation validation.

Machine learning and data-driven proxy models provide another route for production prediction, injector-efficiency evaluation, and waterflood and WAG decision support [19–24]. These approaches can represent nonlinear relationships in field dynamic data and support rapid surveillance or optimization studies. Nevertheless, prediction accuracy alone does not resolve the specific surveillance question addressed in this study. A model may reproduce production trends but still provide limited traceability. It may not identify the injection event, response window, lag time, neighboring injector, or uncertainty source that supports a candidate injection–production interpretation.

The unresolved practical gap is therefore not the absence of another waterflood prediction model. Rather, the gap lies in the limited availability of reproducible evidence objects that link material injection perturbations to bounded production responses while retaining attribution uncertainty. Recent interwell-connectivity studies have increasingly incorporated complex-well data integration, data-driven connectivity estimation, probabilistic uncertainty quantification, and physics-informed machine learning [14–18,25]. These developments highlight the need for surveillance workflows that are both data-scalable and interpretable, while remaining conservative in their physical claims.

This study proposes an event-window evidence-fusion workflow (EWEF) for screening injection–production responses in mature waterflood reservoirs. The workflow treats each injection perturbation as an observational trigger and constructs event-window evidence for candidate injector–producer objects. Unlike connectivity coefficients or production forecasts, the output of EWEF is a ranked review list rather than a causal declaration. This positioning is consistent with recent efforts in connectivity characterization and data-driven

interwell simulation, while keeping final field interpretation dependent on independent evidence such as tracer tests, pressure data, production logging, field interventions, or full-physics simulation results [25–27].

This work contributes to waterflood surveillance in four practical aspects. First, the workflow converts raw dynamic records into event-conditioned evidence objects, rather than relying only on continuous trend fitting or preselected well-pattern diagnostics. Second, it fuses static and dynamic evidence so that short-term response correlations are interpreted together with well-pattern plausibility. Third, it explicitly represents multi-injector interference instead of forcing a single-injector explanation when attribution is ambiguous. Fourth, it uses a weak-label machine learning model as a consistency-review layer and routes rule-model disagreement to field review. The proposed workflow is therefore intended as an engineering surveillance architecture for traceable screening and prioritization of candidate injection–production responses, not as a fundamental replacement for established CRM, simulation, tracer, or pressure-based methods. The following sections describe the construction of event-window evidence, the rule-model fusion procedure, and the anonymized field application used to evaluate the workflow.

2. Data and Methods

EWEF is formulated as an event-conditioned evidence-fusion workflow. Let i denote an injection well, j denote a production well, and e denote a detected injection perturbation. The event-conditioned analytical object is an injector–producer–event unit. For reporting and final routing, event-level descriptors are retained within a candidate injector–producer evidence object. The candidate-object counts in Table 1 therefore represent unique injector–producer objects with event-window evidence, not the Cartesian product of all perturbation events and all producers. The feature vector contains static plausibility, event magnitude, window-response shifts, lagged response correlations, multi-injector context, and data-quality information, as summarized in Equation (1). Rule labels, model probabilities, and disagreement status are derived from this feature vector in the routing step. The final output is not a binary declaration of connectivity, but a ranked review list with traceable evidence:

$$U_{ije} = (i, j, e), \quad x_{ije} = [S_{ij}, |\Delta q_e|, \Delta \bar{y}_o, \Delta \bar{y}_l, \Delta f_w, \rho_o, \tau_o^*, \rho_w, \tau_w^*, C_{ij}, z_q] \quad (1)$$

Table 1. Anonymized sample scale and candidate injector–producer evidence objects.

Sample	Dynamic Wells	Injectors	Producers	Time Steps	Injection Events	Candidate Pair Evidence Objects	Final Review-List Objects
Sample A	28	11	17	143	87	187	12
Sample B	21	8	13	148	43	104	12
Sample C	5	2	3	148	23	6	6

Note. Candidate pair evidence objects represent unique injector–producer objects after event-window evidence construction and routing; they are not the Cartesian product of all perturbation events and all producers.

The workflow has seven ordered operations: dynamic-record checking, event extraction, candidate-pair construction, response-window quantification, lag-response scanning, multi-injector attribution, and disagreement-gated routing. This sequence keeps static screening, dynamic response interpretation, and model-based review logically separated. Static plausibility is used before dynamic interpretation, lag evidence is interpreted within the event-conditioned analysis, and the machine learning layer is used only after the rule-derived evidence vector was constructed. Table 2 summarizes the modules, inputs, checks,

and outputs. In Equation (1), i denotes an injector, j denotes a producer, e denotes an injection perturbation event, x_{ije} denotes the evidence vector for the injector–producer–event unit, and z_q denotes the data-quality and operating-state label associated with the event window.

Table 2. Method modules, inputs, checks, and outputs.

Module	Input	Main Operation/Check	Output
Data checking	Raw dynamic records	Unify time steps; label shut-in, restart, missing, and abnormal intervals	Interpretable rate, water-cut, pressure, and state matrices
Event extraction	Injector-rate series and state labels	Detect material increases, decreases, shut-ins, and restarts using amplitude, persistence, and event-spacing rules	Injection perturbation inventory
Static prior	Well-pattern geometry and kh proxy	Score distance, transmissibility proxy, and local adjacency	Static plausibility score for candidate links
Response window	Event time and producer series	Compare pre-event and post-event oil, liquid, water-cut, and pressure responses	Window-response descriptors
Lag scan	Injection-rate changes and producer-response changes	Scan delayed correlations over a finite lag range	Best lag and lagged correlation evidence
Attribution	Candidate injector set for each producer	Estimate relative explanatory contribution under lag assumptions	Dominant-injector or multi-injector context
Disagreement gate	Rule labels and random-forest weak-label review	Route conflicts and low-certainty objects to review	Field-review list

2.1. Anonymized Data Basis

The workflow was applied to three anonymized waterflood samples derived from field dynamic records. The input records include injection rate, injection pressure, production liquid rate, oil rate, water cut, pressure where available, operating-state records, well locations, and petrophysical or transmissibility proxies. Confidential block names and well names were removed before manuscript preparation. Samples are denoted as Sample A, Sample B, and Sample C; wells are referred to only as anonymous injectors and producers. Table 1 and Figure 1 summarize the scale of the anonymized evidence base.

2.2. Dynamic-Record Checking and State Labeling

Raw dynamic records were aligned to a common monthly time axis and checked before event extraction. For producers, oil rate $q_{o,j}(t)$, liquid rate $q_{l,j}(t)$, and water cut $f_{w,j}(t)$ were checked jointly. Records with physically inconsistent combinations, such as $q_{o,j}(t) > q_{l,j}(t)$, or water-cut fluctuations during zero-liquid-rate intervals, were marked as abnormal. For injectors, injection rate $q_{inj,i}(t)$, injection pressure where available, and operating-state labels were interpreted together to distinguish actual operational adjustment from missing data, shut-in, or restart effects.

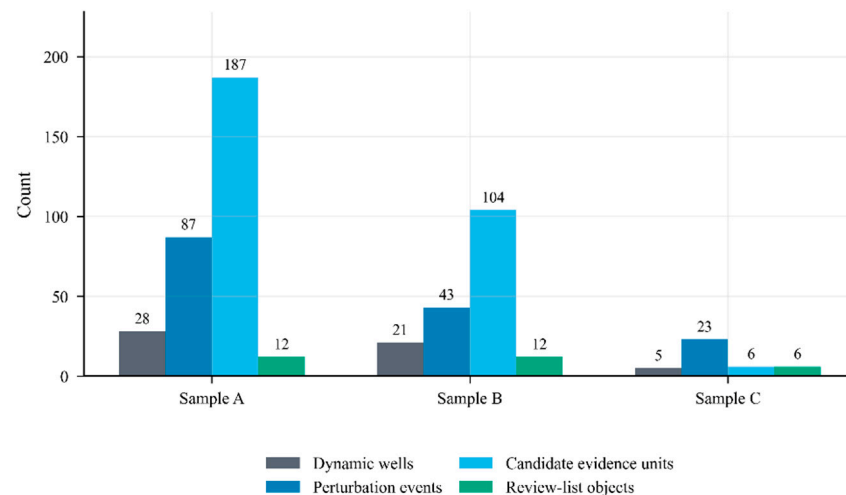


Figure 1. Scale of the anonymized evidence base. The workflow processed dynamic wells, injection perturbation events, candidate injector–producer evidence objects, and final review-list objects.

Missing or abnormal intervals were not directly interpolated for response interpretation. They were assigned state labels such as short missing interval, shut-in interval, shut-injection interval, restart, or long invalid interval. These labels are propagated into later evidence construction so that an apparent production response is not interpreted without knowledge of the operating context. This preprocessing step makes subsequent event-window calculations traceable to operating-state information.

2.3. Injection-Perturbation Event Extraction

For injector i at time step t , the first-order injection perturbation is defined by the rate difference between two adjacent reporting periods. A candidate event is retained when the absolute rate change exceeds both a relative materiality threshold and an engineering absolute threshold, as given in Equation (2). To reduce false events caused by short operational noise, the rate change must persist for at least n_e time steps or be supported by a restart, shut-in, or operating-state transition:

$$\Delta q_{inj,i}(t) = q_{inj,i}(t) - q_{inj,i}(t-1), \quad |\Delta q_{inj,i}(t)| \geq \max(\eta \bar{q}_{inj,i}, Q_{\min}) \quad (2)$$

Each retained event e is stored as $e = \{i, t_e, q_{prev}, q_{curr}, \Delta q, |\Delta q|, s_e\}$, where s_e is the event class: injection increase, injection decrease, shut-in, or restart/new injection. Adjacent events associated with the same operational adjustment are merged when their separation is shorter than a minimum spacing rule. The parameters η , $Q_{\min}$, n_e , and the spacing rule in Equation (2) define the relative rate-change threshold, the absolute engineering threshold, the minimum persistence requirement, and the minimum separation between adjacent events, respectively. They are reported symbolically because the asset-specific numerical values are tied to confidential operating practice. In qualitative sensitivity terms, relaxing η , $Q_{\min}$, n_e , or the spacing rule would generally retain more perturbation events and expand the downstream candidate and review-list objects, whereas tightening them would suppress short operational noise but may remove weak yet potentially relevant responses. Transfer to another field therefore requires recalibration and sensitivity checking rather than direct reuse of the present settings.

When multiple retained events for the same injector–producer pair have overlapping response windows, the workflow does not force a single-event interpretation. Events that belong to the same operational adjustment are merged by the spacing rule; otherwise, the

evidence object is flagged as overlapping-event uncertainty and retained for field review rather than being used as a clean single-event attribution.

2.4. Response-Window Quantification

For event e at time t_e , EWEF defines a pre-event baseline window and a delayed post-event response window. The response window is allowed to shift by a tested delay τ so that delayed production responses are not forced into a same-month interpretation. In Equation (3), L_{pre} and L_{post} denote the baseline and response-window lengths, τ denotes the tested delay, and t_e denotes the event time:

$$W_e^- = [t_e - L_{pre}, t_e), W_e^+(\tau) = (t_e + \tau, t_e + \tau + L_{post}] \quad (3)$$

For a production response variable $y_{j,m}(t)$, with m representing oil rate, liquid rate, water cut, or pressure, the raw window response and standardized response are computed by Equation (4). In Equation (4), $\Delta y_{j,m,e}$ represents the pre/post event-window shift, μ_{pre} and μ_{post} are the window means, and $\sigma_{j,m}$ is the valid-interval scale factor used for standardization. Standardization allows response strength to be compared across wells and variables with different fluctuation scales:

$$\Delta \bar{y}_{j,m}(\tau) = \text{mean}_{t \in W_e^+(\tau)} y_{j,m}(t) - \text{mean}_{t \in W_e^-} y_{j,m}(t), R_{j,m}(\tau) = \frac{\Delta \bar{y}_{j,m}(\tau)}{\sigma_{j,m} + \varepsilon} \quad (4)$$

The reported event-window statistics use six monthly time steps before and after the event. This choice is an operational window length rather than a universal reservoir-response constant. Wider context windows are shown only for visual interpretation. A different asset should test whether the final review-list count and response classes remain stable under shorter and longer windows; a longer response window may capture delayed responses but also increases exposure to common field trends and unrelated operations, whereas a shorter window may miss slow responses.

2.5. Static Prior and Lag-Response Scanning

Static plausibility is introduced as a prior rather than a truth label. For injector i and producer j , the prior combines normalized distance, a normalized effective kh or transmissibility proxy, and a local adjacency indicator, as written in Equation (5). Here, D_{ij} , K_{ij} , and A_{ij} are scaled to comparable ranges, and the weights w_D , w_K , and w_A sum to one. These weights are field-calibration parameters rather than universal constants. The term prevents a remote pair with an incidental short-term correlation from being promoted without geometric or transmissibility plausibility, but it should not be interpreted as proof of an actual flow path. In heterogeneous reservoirs, the available kh proxy may be an incomplete representation of true transmissibility, and the prior may partly act as a proximity filter; therefore, prior-weight sensitivity must be tested before the workflow is transferred or used for decisions:

$$S_{ij} = w_D(1 - D_{ij}) + w_K K_{ij} + w_A A_{ij}, w_D + w_K + w_A = 1 \quad (5)$$

Because injection response is rarely instantaneous, EWEF scans finite response lags. For $\tau = 0, 1, \dots, \tau_{\max}$, the lagged response correlation and selected lag are computed by Equation (6), where $\tau_{\max}$ is the maximum operationally plausible delay specified during field calibration. Correlations are computed over valid dynamic intervals after excluding shut-in, missing, and abnormal states. The best lag is not interpreted as a standalone result or as a physical travel time. It is only one component of the evidence vector and must be consistent with response direction, event-window magnitude, static prior, and

data-quality labels. Table 3 lists the main evidence variables used in the rule-model fusion. Because monthly flow-rate series may contain temporal autocorrelation, common field trends, and synchronized operations, the Pearson correlations are used as descriptive screening descriptors rather than confirmatory statistics. Formal significance testing would require object-specific valid-month counts, an effective-degrees-of-freedom correction, or dependence-preserving block bootstrap/permutation tests:

$$\rho_{ij,m}(\tau) = \text{corr}[\Delta q_{inj,i}(t - \tau), \Delta y_{j,m}(t)], \tau_{ij,m}^* = \text{argmax}_{\tau} |\rho_{ij,m}(\tau)| \quad (6)$$

Table 3. Core evidence variables used in rule-model fusion.

Variable Group	Examples	Interpretation in EWEF
Static prior	Distance, effective kh, prior score	Physical plausibility of a candidate response link
Window response	Oil-rate shift, liquid-rate shift, water-cut shift, pressure shift	Local pre/post response to the perturbation event
Lag evidence	Best lag, oil-rate correlation, water-cut correlation	Delayed dynamic consistency
Interference context	Neighboring injector count, normalized contribution scores	Risk that the response is not uniquely attributable
Uncertainty evidence	Rule confidence, model probability, disagreement status	Priority for field review and manual validation

2.6. Multi-Injector Attribution and Disagreement Gate

A producer may be affected by several injectors within overlapping response windows. For producer j , let $N(j)$ be the set of plausible neighboring injectors selected from the static-prior and operating-context checks. A simplified attribution model is used to estimate relative explanatory contribution, and the normalized contribution C_{ij} is interpreted as relative dynamic contribution under the available records and lag assumptions, not as direct proof of a physical channel:

$$\Delta y_j(t) = \alpha_0 + \sum_{i \in N(j)} \alpha_{ij} \Delta q_{inj,i}(t - \tau_{ij}^*) + \varepsilon(t), C_{ij} = \frac{|\alpha_{ij}|}{\sum_{k \in N(j)} |\alpha_{kj}|} \quad (7)$$

The coefficients in Equation (7) are estimated from valid standardized rate-change intervals, and the contribution scores are used only as relative screening evidence. In Equation (7), β_{ij} denotes the fitted dynamic-response coefficient for injector i and producer j , and C_{ij} denotes the normalized relative contribution among the neighboring injectors of producer j . If the contribution denominator is small or the neighboring-injector contributions are highly dispersed, the object is not assigned to a single injector without field review. A transparent rule layer maps x_{ije} into engineering review classes, including effective support, water-cut risk, weak response, delayed response, multi-injector influence, and review class. A random-forest classifier trained on these rule-derived weak labels is then used as a consistency-review model rather than as independent ground truth [28].

The disagreement gate is defined by comparing the rule label $g_{rule}(x_{ije})$ with the model label $g_{RF}(x_{ije})$ and model confidence $p_{RF}(x_{ije})$. In Equation (8), g_{rule} is the rule-derived review class, g_{RF} is the weak-label model class, p_{RF} is the model confidence, and $H(C_j)$ is the entropy-based ambiguity indicator for dispersed multi-injector contributions. A candidate object is routed to field review when rule-model disagreement, low model confidence, or high multi-injector contribution entropy is observed. These routing criteria

are conservative review triggers rather than proof of connectivity. Because complete independent tracer, pressure-transient, intervention-outcome, or full-physics connectivity labels are not available for every candidate object, the workflow is evaluated as a screening and prioritization method rather than as a causal-confirmation method. The weak-label model should therefore be read conservatively. Its purpose is not to validate the rule layer from independent truth labels; instead, it provides a second, nonlinear mapping of the same evidence vector so that rule-model disagreement, low confidence, or dispersed attribution can be routed to manual field review:

$$d_{ije} = \begin{cases} 1, & g_{\text{rule}}(x_{ije}) \neq g_{\text{RF}}(x_{ije}) \text{ or } p_{\text{RF}}(x_{ije}) < p_{\text{min}} \text{ or } H(C_j) > H_{\text{min}} \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

3. Results

The three samples cover different well-pattern scales and dynamic complexity. Sample A is the largest case, with 28 dynamic wells, 87 injection perturbation events, and 187 candidate pair evidence objects. Sample B contains 21 dynamic wells, 43 events, and 104 candidate pair evidence objects. Sample C is a small-pattern case, with five dynamic wells, 23 events, and six candidate pair evidence objects. Figure 2 provides field-scale production context, Figure 3 summarizes the perturbation-event inventory, and Figure 4 shows the static-prior and geometric evidence. This range demonstrates application across different pattern sizes, although it does not by itself establish computational scalability or field-wide generalization.

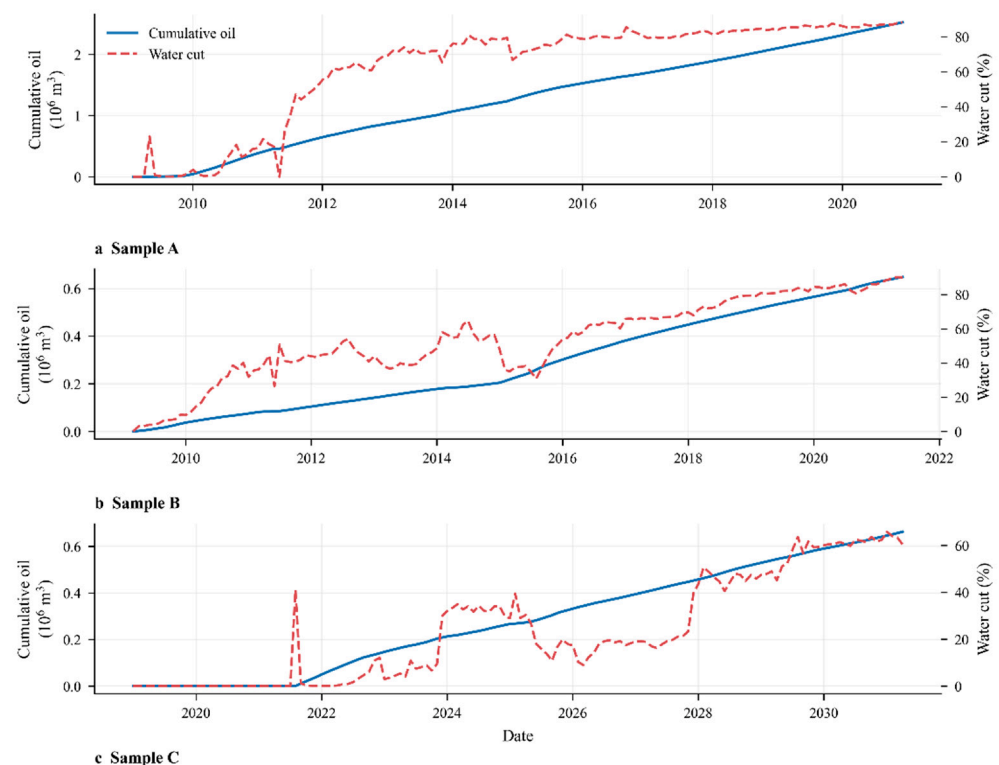


Figure 2. Field-scale cumulative oil and water-cut dynamics for the three samples. These curves provide development context but are not used as substitutes for local event-window attribution.

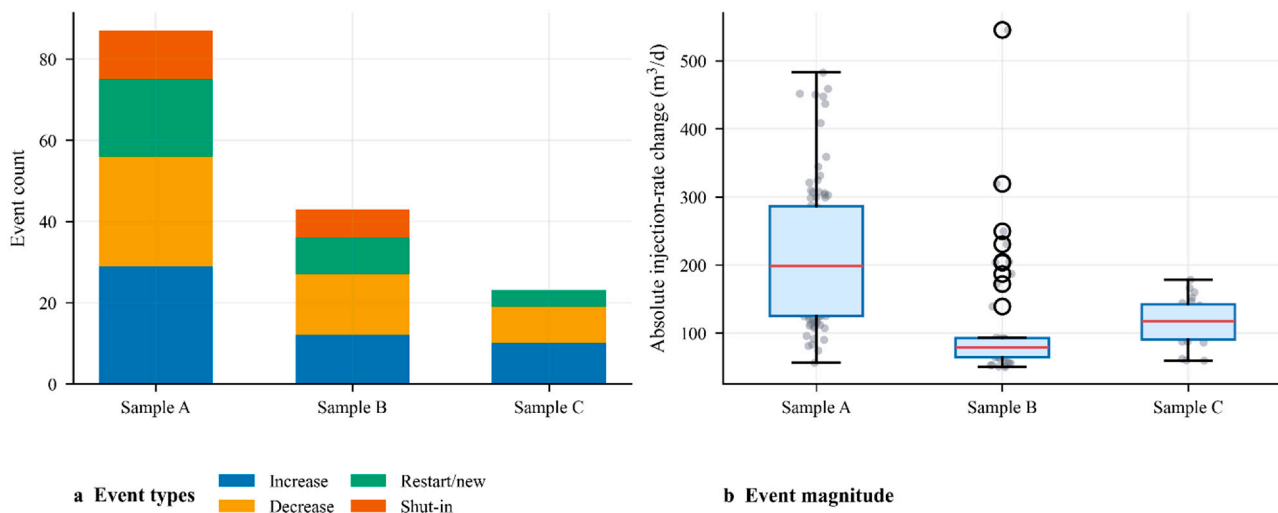


Figure 3. Injection perturbation inventory. Panel (a) summarizes event types; panel (b) shows the distribution of absolute injection-rate changes retained as material perturbations.

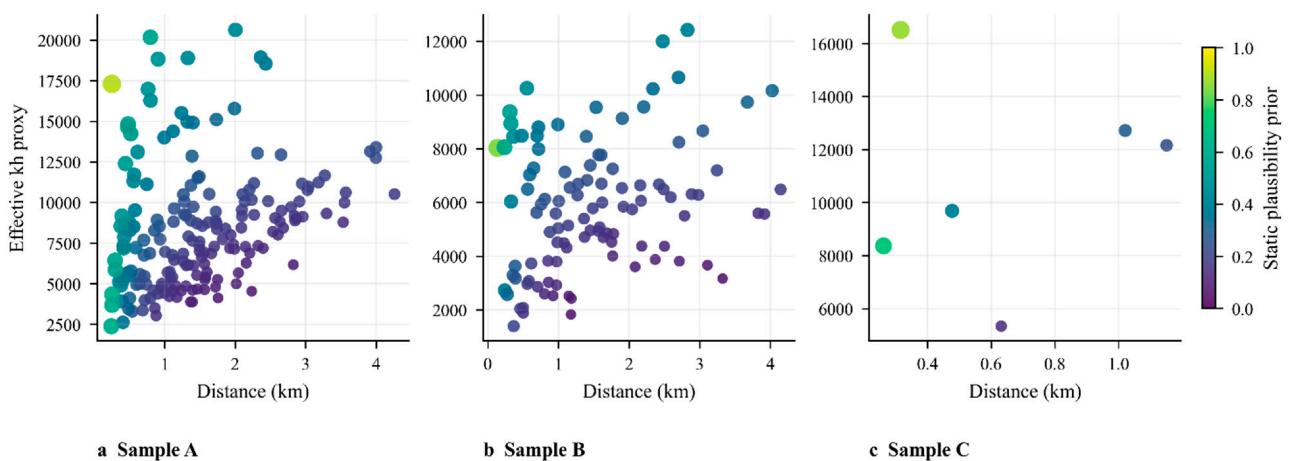


Figure 4. Static prior and geometric evidence. Candidate links are displayed by injector–producer distance, effective kh proxy, and normalized prior score.

3.1. Injection Perturbation Inventory

Across all samples, EWEF extracted 153 injection perturbation events. Sample A includes 29 rate-increase events, 27 rate-decrease events, 19 restart/new-injection events, and 12 shut-ins. Sample B includes 12 rate-increase events, 15 rate-decrease events, nine restart/new-injection events, and seven shut-ins. Sample C includes 10 rate-increase events, nine rate-decrease events, and four restart/new-injection events. The mean absolute rate-change magnitudes are $216.1 \text{ m}^3/\text{d}$, $109.8 \text{ m}^3/\text{d}$, and $116.9 \text{ m}^3/\text{d}$ for Samples A, B, and C, respectively. These numbers describe event magnitude but should be interpreted together with the threshold and persistence rules described in the method.

3.2. Representative Event-Window Response

A representative anonymous object in Sample A illustrates how the event-window design works, as shown in Figure 5 and summarized in Table 4. The injection perturbation magnitude is $309.8 \text{ m}^3/\text{d}$. Within the six-month post-event window, oil rate increased by $5.49 \text{ m}^3/\text{d}$, whereas water cut decreased by 0.52 percentage points relative to the pre-event window. This example is used to show the structure of the evidence object; it is not used as standalone proof of reservoir connectivity.

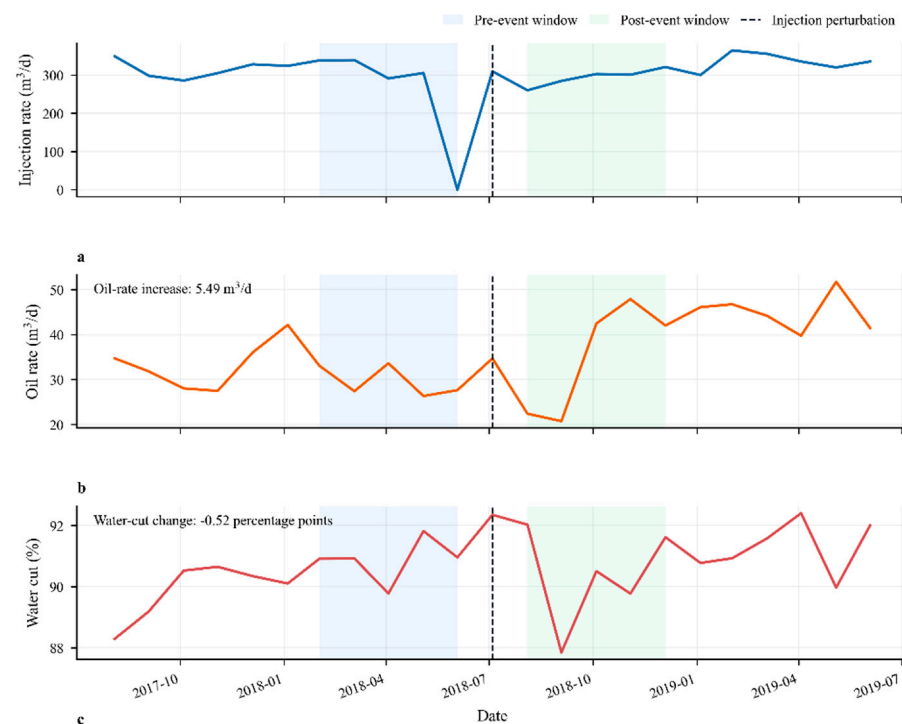


Figure 5. Representative event-window response for an anonymous injector–producer pair. (a) Injection-rate history and perturbation timing; (b) producer oil-rate response; (c) producer water-cut response. The dashed line marks the injection perturbation; shaded intervals indicate pre-event and post-event response windows.

Table 4. Representative anonymous event-window evidence.

Metric	Value	Interpretation
Injection perturbation magnitude	309.8 m^3/d	Material operational change retained as an event
Oil-rate shift	+5.49 m^3/d	Oil-rate increase relative to the pre-event baseline in the illustrative object
Water-cut shift	−0.52 percentage points	Water-cut decrease relative to the pre-event baseline; no immediate water-cut deterioration in this example
Event-window length	Six months before and after the event	Operationally interpretable comparison window

Note. Positive and negative signs in the Table denote changes relative to the pre-event baseline.

The interpretation of this example depends on the event timing, the pre-event baseline, the post-event response, and the water-cut behavior. In high-water-cut reservoirs, an oil-rate increase accompanied by a strong water-cut increase would not be classified as simple effective support. The decrease in water cut in this object indicates that no immediate water-cut deterioration was observed in the selected post-event window.

3.3. Lag Response and Multi-Injector Attribution

The lag scan in Figure 6 shows that response evidence depends on the selected delay. For the anonymous lag-scan object in Sample A, the largest observed oil-rate correlation occurs at lag 7 time steps with $r = 0.34$, while the largest observed water-cut correlation

occurs at lag 8 time steps with $r = -0.34$. These correlations are modest, with a maximum absolute magnitude of approximately 0.34. The lag profile is irregular and should not be read as a statistically significant peak, a stable delay estimate, or direct evidence of physical connectivity without additional testing. It may be influenced by data density, window construction, common field trends, synchronized operations, and temporal autocorrelation. The lag results are therefore interpreted as screening evidence only and must be combined with event-window shifts, static-prior evidence, and operating-state information.

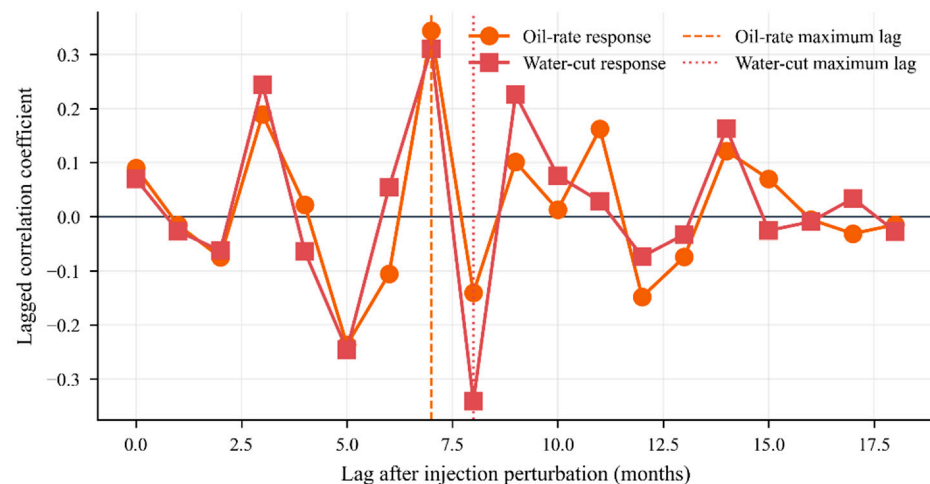


Figure 6. Lag-response scan for an anonymous candidate object. Oil-rate and water-cut correlations are evaluated across a finite delay range and used as screening evidence.

Multi-injector interference is a recurring issue in the larger samples. For one anonymous producing well, the attribution module identified seven neighboring injectors in the response context, with the largest normalized contribution equal to 0.19. Figure 7 illustrates this dispersed attribution pattern. Such objects should be routed to field review before any injection adjustment is assigned to one well pair.

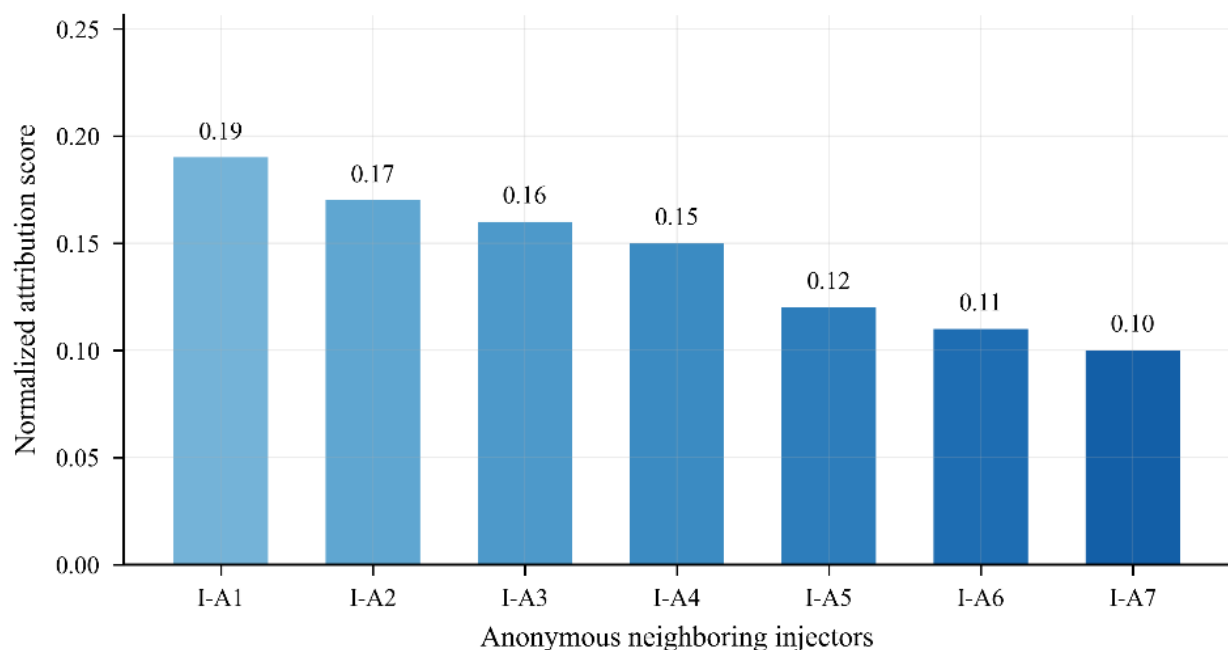


Figure 7. Multi-injector attribution for an anonymous production object. Dispersed normalized scores indicate that no single neighboring injector dominates the available dynamic evidence.

3.4. Feature Importance and Disagreement Routing

The random-forest consistency-review model provides a descriptive check on the evidence vector. Across the three samples, high-ranking variables in Figure 8 include oil-rate lag correlation, event-window water-cut shift, event-window oil-rate shift, static prior, and water-cut lag correlation. Because the model is trained on rule-derived weak labels, these rankings should not be interpreted as independent evidence of causal connectivity or as proof that the rule layer is physically correct. High importance assigned to variables used by the rule layer is expected in this setting and does not by itself strengthen the evidential argument. The feature importance plot is therefore used only to audit which evidence components drive the weak-label consistency-review layer.

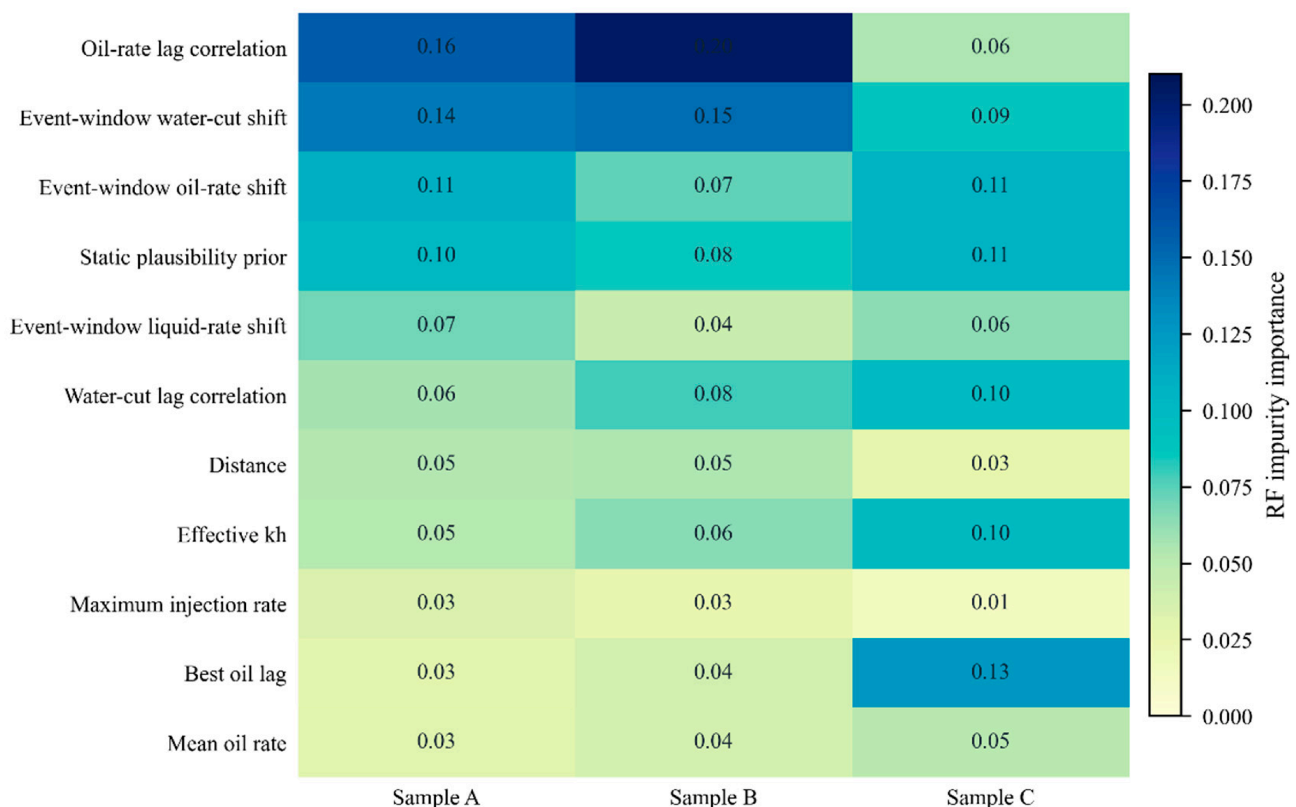


Figure 8. Feature importance in the random-forest consistency-review model. The model is trained on rule-derived weak labels and is interpreted only as a descriptive consistency-review layer.

Rule-model routing produced 24 agreement-routed review-list objects and six disagreement-routed review-list objects across the three samples, as shown in Table 5 and Figure 9. Table 5 summarizes the overall rule-label distribution, model-label distribution, final routing, and mean model confidence for all 297 candidate pair objects, rather than only the representative example. The mean model confidence values are 0.77, 0.84, and 0.82 for Samples A, B, and C, respectively. Sample A and Sample B each contain three disagreement-routed objects, while Sample C contains no disagreement-routed object in the final review list. Because Sample C has only six candidate pair evidence objects and all six are routed to the final review list, it should be treated as a small-pattern sanity check rather than as statistical evidence of screening selectivity. Without independent confirmation of at least a subset of the routed objects, the false-positive and false-negative behavior of the final review list remains unknown.

Table 5. Rule-model label distribution and final review routing.

Sample	Candidate Pair Objects	Rule-Label Distribution	Model-Label Distribution	Final Review-List Routing	Mean Model Confidence
Sample A	187	ES: 12; WCR: 22; WR: 66; MI: 74; RC: 13	ES: 10; WCR: 21; WR: 68; MI: 75; RC: 13	12 total; 9 agreement-routed; 3 disagreement-routed	0.77
Sample B	104	ES: 1; WCR: 17; WR: 40; MI: 41; RC: 5	WCR: 15; WR: 40; MI: 44; RC: 5	12 total; 9 agreement-routed; 3 disagreement-routed	0.84
Sample C	6	ES: 1; WR: 4; RC: 1	ES: 1; WR: 4; RC: 1	6 total; 6 agreement-routed; 0 disagreement-routed	0.82

Note. ES = effective support; WCR = water-cut risk; WR = weak response; MI = multi-injector influence; RC = rule-label review class. The final review list is a routed priority list and can include objects selected because of agreement, disagreement, or small-sample review policy.

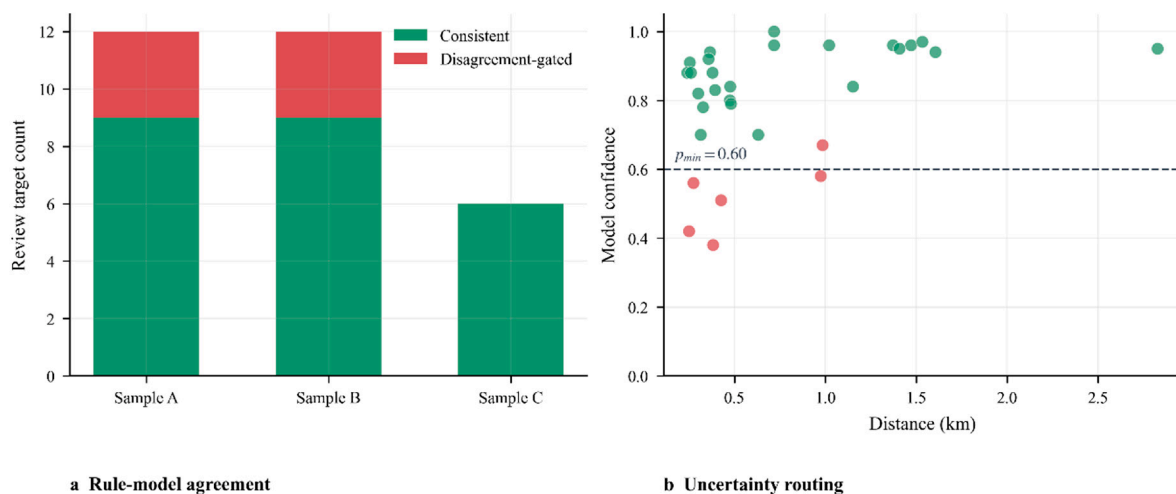


Figure 9. Rule-model agreement and uncertainty routing. Disagreement is retained as a field-review signal rather than averaged into a single unqualified prediction.

4. Discussion

EWEF should be viewed as a field-surveillance screening workflow rather than a replacement for established waterflood analysis. Compared with conventional surveillance, its distinctive feature is not the use of oil rate, liquid rate, or water cut themselves, but the conversion of operational changes into event-window evidence objects. Each candidate interpretation can be traced back to a perturbation event, a response window, a lag scan, static plausibility, attribution context, and rule-model agreement status.

Compared with CRM and related connectivity or optimization methods [10–18], EWEF is deliberately less aggressive in its physical claim. It does not infer a permanent connectivity coefficient as the primary result. Instead, it asks whether available raw records support an operational response interpretation and whether the response is dominated by one injector or confounded by multiple injectors. Table 6 summarizes this positioning relative to common alternatives. The distinction from CRM-type approaches should therefore be understood as one of output form, auditability, and review routing, not as a demonstrated superiority over operationally tuned CRM, INSIM, probabilistic CRM, or simulation-based workflows. The present study does not include a direct CRM benchmark on the same anonymized samples; such a benchmark would be valuable for quantifying agreement and disagreement between EWEF review-list objects and CRM-derived connectivity or allocation indicators.

Table 6. Method positioning relative to common alternatives.

Method Family	Main Strength	Limitation for Local Surveillance	Role of EWEF
Diagnostic plots	High interpretability and field familiarity	Often manual and focused on selected wells or patterns	Enumerates event-conditioned candidates while preserving evidence traceability
Waterflood curves	Field-scale recovery and water-cut trend analysis	Limited local injector–producer attribution	Uses local event windows rather than only field-level trends
CRM/connectivity models	Quantitative interwell connectivity	May aggregate long-period signals and obscure event-specific evidence	Provides event-level candidates before detailed CRM, tracer, or simulation review
Black-box ML/proxy models	Flexible nonlinear prediction	May provide limited field-review traceability	Uses ML as a weak-label consistency layer with explicit disagreement routing
Full-physics simulation	Physics-rich scenario evaluation	Data and calibration intensive for rapid surveillance	Prioritizes candidates for further simulation or intervention review

4.1. Engineering Implications

For injection management, the main implication is that positive production response and water-control risk must be assessed together. An increase in liquid rate after injection enhancement is not sufficient evidence of effective support if oil rate does not increase or water cut rises rapidly. Conversely, a delayed but directionally consistent response should not be discarded only because same-date correlation is weak.

The field-review list provides a practical route for prioritizing engineering attention. High-confidence agreement cases can be reviewed efficiently, while disagreement cases can be routed to operation-record checking, pressure review, layered injection–production inspection, production logging, tracer follow-up, or staged injection tests. This design preserves engineering interpretability while using machine learning to support consistency checking in the screening process.

4.2. Limitations and Future Work

The study has several limitations. First, the available evidence comes from anonymized field dynamic records and derived evidence lists; complete independent tracer, pressure-transient, intervention-outcome, or full-physics connectivity labels are not available for every candidate object, so false-positive and false-negative rates cannot be estimated from the present data. Second, the random-forest model is trained on rule-derived weak labels, so it should be interpreted as a consistency-review tool rather than as an independent classifier; feature importance from this model does not provide independent physical validation. Third, the static prior uses simplified geometry and kh proxies rather than a full geological model, and it may act partly as a proximity filter in highly heterogeneous settings. Fourth, the event thresholds, response-window length, lag range, and prior weights are field-calibration parameters; because asset-specific values cannot be disclosed, the present paper cannot report a quantitative threshold-sensitivity curve or prove that the retained 153 events are conservative rather than permissive. Fifth, lag correlations are modest and

may be affected by temporal autocorrelation, common trends, synchronized operations, and window-construction effects.

Future work should combine EWEF with pressure-transient information, layered injection profiles, production logging, tracer data, numerical-simulation updates, blind field-intervention validation, and explicit negative-control pairs such as distant or geologically implausible injector–producer combinations. A larger anonymized benchmark would allow uncertainty quantification for event thresholds, response-window length, lag range, static-prior weights, rule-model disagreement rates, and agreement with CRM-type connectivity indicators. Such a benchmark should report how retained-event counts, candidate-pair counts, and final review-list counts change under relaxed and tightened perturbation thresholds and prior-weight settings. Until such validation is available, EWEF should be used to screen and prioritize candidate responses rather than to confirm causal connectivity.

5. Conclusions

This study presented EWEF, a traceable event-window evidence-fusion workflow for screening candidate injection–production responses in mature waterflood reservoirs. The workflow converts raw dynamic records into candidate injector–producer evidence objects containing static plausibility, pre/post response shifts, lagged response evidence, multi-injector context, rule labels, model probabilities, and disagreement status.

In the anonymized application, the workflow processed 54 dynamic wells, 153 injection perturbation events, 297 candidate injector–producer evidence objects, and 30 final review-list objects across three samples. Event-window analysis provided localized response evidence that field-scale cumulative curves alone cannot reveal, and lag scanning showed that response evidence depends on delay and should be interpreted jointly with event-window and static-prior evidence. Because the lag correlations are modest, irregular, and not corrected here for temporal autocorrelation, they should be read as descriptive screening descriptors rather than as statistically confirmed lag peaks.

The random-forest consistency-review model indicated that oil-rate lag correlation, water-cut shift, oil-rate shift, static prior, and water-cut lag correlation were influential evidence, but it did not provide independent causal validation. Because the model is trained on rule-derived weak labels, high feature importance for rule-input variables is expected and should be interpreted only as an audit of the weak-label routing layer. The disagreement gate retained uncertainty for engineering review, making the workflow more suitable for surveillance triage than an unqualified black-box prediction.

The current results support EWEF as a traceable workflow for screening and prioritizing candidate injection–production responses. They do not establish the accuracy of the final review list against independent ground truth, and the screening quality remains unknown without validation of at least a subset of routed objects. Causal confirmation and injection-adjustment decisions should rely on additional field evidence such as tracer, pressure, logging, intervention, CRM comparison, or simulation results.

Author Contributions: Conceptualization, P.C. and Y.X.; methodology, J.L. and P.C.; software, C.Q. and Y.Z. (Yuanhao Zheng); validation, H.Q., Y.Z. (Yuanhao Zheng), and Q.G.; formal analysis, Y.Z. (Yuhui Zhou); investigation, H.Q.; data curation, Y.Z. (Yuhui Zhou) and Q.G.; writing—original draft, J.L.; writing—review and editing, J.L. and C.Q.; visualization, Y.Z. (Yuanhao Zheng); supervision, Y.X.; project administration, J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the China National Petroleum Corporation (CNPC) under Grant No. 2025ZG60.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: Authors Jian Li, Peng Cao, Congying Qiao, Hongtu Qian and Qiaoyu Ge were employed by the CNPC R&D (DIFC) Company Limited. 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. This research was funded by the China National Petroleum Corporation (CNPC). Any opinions, findings, and conclusions expressed in this paper are those of the authors and do not necessarily reflect the views of CNPC.

References

- Hao, T.; Zhong, L.; Liu, J.; Zhang, X.; Zhang, L. Waterflood efficiency assessment using injection-production relationship analysis method. *Math. Geosci.* **2023**, *55*, 201–228. [\[CrossRef\]](#)
- Setiyawan, H.; Evelyn, K.; Hisan, R.F.; Fardiansyah, I. Subsurface waterflood management improvement: The art of reservoir connectivity characterization. In Proceedings of the SPE Improved Oil Recovery Conference, Tulsa, OK, USA, 22–25 April 2024. [\[CrossRef\]](#)
- Laochamroonvorapongse, R. Advancement in waterflood and EOR performance assessment and optimization with capacitance-resistance model in the largest oil field, Thailand. In Proceedings of the International Petroleum Technology Conference, Bangkok, Thailand, 1–3 March 2023. [\[CrossRef\]](#)
- Nikmardan, N.; Rafiei, Y.; Ameri, M.J. Novel integrated approach for waterflood optimization in mature multilayer reservoirs with advanced well completions using capacitance resistance model. *Arab. J. Sci. Eng.* **2024**, *49*, 8833–8855. [\[CrossRef\]](#)
- Debossam, J.G.S.; de Freitas, M.M.; de Souza, G.; Souto, H.P.A. Picard-Newton Method for Water-Alternating-Gas Injection Simulation in Heterogeneous Reservoirs. *Processes* **2026**, *14*, 20. [\[CrossRef\]](#)
- Dong, W.; Jiao, J.; Xie, S.; Lyu, C.; Cui, G.; Meng, J. Cumulative production curve method for the quantitative evaluation on the effect of oilfield development measures: A case study of the nitrogen injection pilot in Yanling oilfield, Bohai Bay Basin. *Pet. Explor. Dev.* **2016**, *43*, 672–678. [\[CrossRef\]](#)
- Li, Y.; Cong, Z.; Fu, X.; Jiang, X.; Ranjith, P.G.; Li, T. A novel simulation approach for far-field temporary plugging fracturing in horizontal wells. *Pet. Sci.* **2026**. [\[CrossRef\]](#)
- Cong, Z.; Liu, J.; Li, Y.; Ranjith, P.G.; Li, X. The numerical modeling and simulation of CO₂ fracturing flooding in low-permeability reservoirs. *Rock Mech. Rock Eng.* **2026**. [\[CrossRef\]](#)
- Cong, Z.; Liu, J.; Ranjith, P.G.; Li, X.; Zhu, T. Numerical investigation of the influence of fracture morphology on enhanced geothermal system (EGS) development using the discrete element method. *Energy Convers. Manag.* **2026**, *348*, 120669. [\[CrossRef\]](#)
- Mamghaderi, A.; Aminshahidy, B.; Bazargan, H. Prediction of waterflood performance using a modified capacitance-resistance model: A proxy with a time-correlated model error. *J. Pet. Sci. Eng.* **2021**, *198*, 108152. [\[CrossRef\]](#)
- Simha, S.; Ujjwal, M.; Modi, G. Practical considerations when using capacitance resistance modelling CRM for waterflood optimization. In Proceedings of the SPE/IATMI Asia Pacific Oil & Gas Conference and Exhibition, Virtual, 12–14 October 2021. [\[CrossRef\]](#)
- Zhao, H.; Xu, L.; Guo, Z.; Liu, W.; Zhang, Q.; Ning, X. A new and fast waterflooding optimization workflow based on INSIM-derived injection efficiency with a field application. *J. Pet. Sci. Eng.* **2019**, *179*, 1186–1200. [\[CrossRef\]](#)
- Liu, W.; Zhao, H.; Sheng, G.; Li, H.A.; Xu, L.; Zhou, Y. A rapid waterflooding optimization method based on INSIM-FPT data-driven model and its application to three-dimensional reservoirs. *Fuel* **2021**, *292*, 120219. [\[CrossRef\]](#)
- Liu, S.; Cao, L.; He, H.; Yang, T.; Zhou, D. An improved interwell connectivity model to obtain interwell connectivity information by using complex well data. *Simulation* **2021**, *97*, 135–143. [\[CrossRef\]](#)
- Dastgerdi, E.J.; Shabani, A.; Zivar, D.; Jahangiri, H.R. Estimation of underground interwell connectivity: A data-driven technology. *J. Taiwan Inst. Chem. Eng.* **2020**, *116*, 144–152. [\[CrossRef\]](#)
- Thiam, M.; Nakhaee, A. Reservoir interwell connectivity estimation from small datasets using a probabilistic data driven approach and uncertainty quantification. *Geoenery Sci. Eng.* **2023**, *230*, 212154. [\[CrossRef\]](#)
- Cheng, H.; Vyatkin, V.; Osipov, E.; Zeng, P.; Yu, H. LSTM based EFAST global sensitivity analysis for interwell connectivity evaluation Using Injection and Production Fluctuation Data. *IEEE Access* **2020**, *8*, 67289–67299. [\[CrossRef\]](#)
- Liu, X.; Qu, X.; Jiang, M.; Huang, J.; Chen, C. A new method to quantify interwell connectivity using performance data and intelligence algorithm. In Proceedings of the International Petroleum Technology Conference, Dhahran, Saudi Arabia, 13–15 January 2020. [\[CrossRef\]](#)
- Zhong, Z.; Sun, A.Y.; Wang, Y.; Ren, B. Predicting field production rates for waterflooding using a machine learning-based proxy model. *J. Pet. Sci. Eng.* **2020**, *194*, 107574. [\[CrossRef\]](#)

20. Deng, L.; Pan, Y. Data-driven proxy model for waterflood performance prediction and optimization using Echo State Network with Teacher Forcing in mature fields. *J. Pet. Sci. Eng.* **2021**, *197*, 107981. [[CrossRef](#)]
21. Shabani, A.; Moosavi, M.S.; Zivar, D.; Jahangiri, H.R. Data-driven technique for analyzing the injector efficiency in a waterflooding operation. *Simulation* **2020**, *96*, 701–710. [[CrossRef](#)]
22. Xiong, X.; Lee, K.J. Data-driven modeling to optimize the injection well placement for waterflooding in heterogeneous reservoirs applying artificial neural networks and reducing observation cost. *Energy Explor. Exploit.* **2020**, *38*, 2413–2435. [[CrossRef](#)]
23. Xu, Y.; Zhao, H.; Pathegama Gamage, R.; Chen, Q.; Zhou, Y.; Rao, X. Advanced attention-based spatial-temporal neural networks for enhanced CO₂ water-alternating-gas performance prediction and history matching. *Phys. Fluids* **2024**, *36*, 097116. [[CrossRef](#)]
24. Xu, Y.; Liu, W.; Zhao, H.; Rao, X.; Liu, B. Application of the FlowNet model for history matching and production optimization in water alternating gas enhanced oil recovery. *Phys. Fluids* **2024**, *36*, 083323. [[CrossRef](#)]
25. Nagao, M.; Datta-Gupta, A.; Onishi, T.; Sankaran, S. Reservoir connectivity identification and robust production forecasting using physics informed machine learning. In Proceedings of the SPE Reservoir Simulation Conference, Galveston, TX, USA, 28–30 March 2023. [[CrossRef](#)]
26. Li, Y.; Onur, M. INSIM-3P: An interwell data-driven physics-based simulator for three-phase flow. In Proceedings of the SPE Reservoir Simulation Conference, Galveston, TX, USA, 25–27 March 2025. [[CrossRef](#)]
27. Tiong-Hui, W.; Cheng, S.T.Z. Decoding interwell tracer data using novel analytical approach to optimise waterflood management in a complex oil rim field, in Offshore East Malaysia. In Proceedings of the Offshore Technology Conference Asia, Kuala Lumpur, Malaysia, 20–23 March 2018. [[CrossRef](#)]
28. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[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.