Peer reviewed Journals

Sci Chronice

RESEARCH

Regime-Conditioned Bayesian Digital Twins for Annular Multiphase Flow: Unified Modeling, Inference, and Control from Sparse Surface Measurements

Santiago Arévalo[1] and Julián Restrepo[2]

[1]Universidad de la Amazonia, Avenida Circunvalar 45–21, Departamento de Ingeniería de Sistemas, Florencia, Caquetá, Colombia
[2]Universidad Mariana, Calle 18 No. 34–52, Departamento de Ingeniería de Sistemas, Pasto, Nariño, Colombia

Abstract

Multiphase flow in wellbore annuli and near-well conduits is a dominant determinant of pressure integrity, circulation stability, and the detectability of hazardous transients such as gas influx. In operational settings, the downhole state is only partially observed through surface sensors that are delayed, noisy, and confounded by uncertain rheology and evolving geometry. A central difficulty is that the effective constitutive behavior of gas–liquid mixtures is topology dependent: the same superficial velocities can yield distinct momentum losses depending on whether the mixture is dispersed, intermittent, or stratified, and regime transitions can occur abruptly under modest boundary changes. This paper proposes a unified digital-twin formulation that couples mechanistic conservation laws with a probabilistic regime process and a physically constrained closure-learning layer. The key contribution is a regime-conditioned Bayesian state-space model in which a discrete latent regime variable modulates the uncertain closure terms governing friction, slip, and compressibility residuals, thereby aligning state estimation, regime inference, and risk assessment under one probabilistic semantics. A constrained variational filtering method is developed to assimilate sparse surface pressure and flow signals while enforcing admissibility constraints on void fraction and mixture properties. The framework further embeds an influx source model for probabilistic kick detection and connects the posterior to a regime-aware stochastic model predictive control layer for pressure-window regulation. Simulation studies across inclination changes, sensor latency, frictional mismatch, and influx scenarios demonstrate improved calibration and more stable control relative to regime-agnostic probabilistic baselines, while remaining computationally compatible with real-time decision support.

1. Introduction

Annular multiphase flow arises in drilling and production whenever gas and liquid coexist in the wellbore, whether due to intentional aeration, gas liberation from solution during pressure drawdown, or unplanned influx from the formation [1]. The operational consequences span routine performance and low-probability high-impact events. On the routine side, frictional losses and hydrostatic head determine pump requirements, cuttings transport, and the stability of circulating density along complex trajectories. On the safety side, transient multiphase dynamics govern whether bottomhole pressure remains within a narrow window bounded by fracture pressure above and formation pore pressure below. In deepwater and long-reach contexts, this window can be small in absolute terms relative to the magnitudes of hydrostatic and frictional contributions, so modest uncertainty in slip, compressibility, or friction can flip a safe plan into an unsafe excursion. This motivates digital twins that do more than forward simulate: they must infer downhole state from limited measurements, quantify uncertainty in a decision-relevant way, and support control actions under that uncertainty [2].

The inverse problem faced in practice is structurally ill-conditioned. Surface measurements such as standpipe pressure, casing pressure, flow-in, flow-out, and choke position constrain downhole states only indirectly through wave propagation and delayed transport. Multiple downhole configurations, including different spatial distributions of void fraction and different slip conditions, can generate similar surface pressure traces, especially when the system is subject to time-varying boundary conditions such as pump ramps, connections, or choke changes. Moreover, many industrial hydraulic models depend on regime-dependent closures. When a model chooses a regime incorrectly, it selects incompatible correlations for friction, interfacial exchange, and slip, and the resulting state estimate can drift even if the numerical solver is stable [3]. This creates a feedback loop: inaccurate state leads to inaccurate regime selection, which further degrades the state.

A second structural complication is that annular multiphase dynamics are hybrid. The conservation laws governing transport and pressure are continuous, but the constitutive behavior changes discontinuously across regime transitions. A digital twin that ignores the hybrid structure often absorbs regime effects as process noise or bias, leading to posteriors that are either overconfident and wrong or so diffuse that they are not decision useful. Conversely, a pipeline that performs regime classification as a separate upstream task can become inconsistent if the classifier's output does not match the closure family used by the downstream hydraulic model [4]. In settings where the same signals must support both routine control and rare-event detection, such inconsistency is particularly problematic: a model might explain a pressure anomaly by switching regimes when the anomaly is actually an influx signature, or explain an incipient regime transition as an influx.

This paper advocates a different organizing principle: treat regime as a latent variable inside the same probabilistic model that represents conservation-law dynamics, closure uncertainty, and measurement processes. Doing so makes regime inference, state estimation, and event detection mutually constrained. The model cannot declare a regime that is inconsistent with observed pressure evolution unless it pays likelihood and prior penalties, and it cannot explain sustained discrepancies by drifting closures without impacting the inferred probability of an influx source. The result is a digital twin whose outputs are posterior distributions over both continuous downhole states and discrete regime trajectories, enabling risk-aware decisions rather than point-estimate heuristics [5].

There is empirical evidence that regime structure is learnable from features in two-phase flow data, and that classification accuracy can be high under carefully defined datasets and feature sets. For example, a study on vertical gas–liquid regime identification found that a cross-validated nearest-neighbor classifier could achieve 98% test accuracy with an optimized neighborhood size, separating several major flow regions in a way that is consistent with established phenomenology [6]. Such results suggest that regime information is present in observables, but they do not address how to integrate that information into a physically consistent estimator or controller in the presence of uncertain closures, delays, and unmodeled disturbances. The present paper leverages the idea that regimes can be inferred, but embeds it within a hybrid Bayesian state-space model that retains mechanistic structure and produces calibrated uncertainty.

The thesis of this work is that a practical annular digital twin should satisfy three requirements simultaneously [7]. It should retain conservation-law structure so that inferred states remain physically coherent across the wellbore, it should represent regime as a stochastic process that modulates closure families in a way that is consistent with topology-dependent behavior, and it should quantify uncertainty so that both alarm logic and control actions can be posed as risk management rather than thresholding. Meeting these requirements leads to a regime-conditioned Bayesian digital twin whose core is a hybrid latent-variable model, coupled to constrained online inference and to stochastic model predictive control.

The contributions are technical rather than survey-oriented. A reduced mechanistic backbone is formulated in a way that is differentiable and amenable to probabilistic inference under delay. Closure uncertainty is represented as a regime-conditioned operator acting through physically meaningful channels such as friction and slip, with constraints that enforce positivity, boundedness, and numerical stability [8]. A variational filtering method is developed for online posterior inference over continuous and discrete states in a hybrid dynamical system, including a nonnegative influx source model for probabilistic kick detection. Finally, a regime-aware stochastic MPC layer uses the posterior predictive distribution to compute control actions that satisfy chance constraints on pressure windows and that hedge against regime uncertainty. The overall system is designed to be deployable under realistic computational budgets and sensor imperfections, while avoiding brittle dependence on any single correlation set or classifier output.

2. Physics-Informed Hybrid State-Space Model

A digital twin for annular multiphase flow must reconcile modeling fidelity with inference tractability. Full three-dimensional multiphase CFD can represent topology and interfacial dynamics, but it is computationally infeasible for online estimation and control [9]. Conversely, purely empirical models can be fast but may extrapolate poorly outside the training envelope and can violate conservation. The approach taken here is to adopt a one-dimensional conservation-law backbone along the well path, represent unresolved physics via structured closures, and embed the result in a probabilistic state-space model.

Let $s \in [0, L]$ denote measured depth along the well trajectory, with inclination $\theta(s)$ and annular cross-sectional area $A(s)$. Let $P_w(s)$ denote wetted perimeter and $D_h(s) = 4A(s)/P_w(s)$ the hydraulic diameter. Consider a gas phase and a liquid phase with densities $\rho_g(p, T)$ and $\rho_\ell(p, T)$ and viscosities $\mu_g(p, T)$ and $\mu_\ell(p, T)$, with temperature $T(s, t)$ treated either as known from a thermal model or as slowly varying relative to pressure transients [6]. Let $\alpha(s, t)$ denote gas void fraction and $u_g(s, t)$, $u_\ell(s, t)$ denote phase velocities averaged over their occupied sub-areas. De-

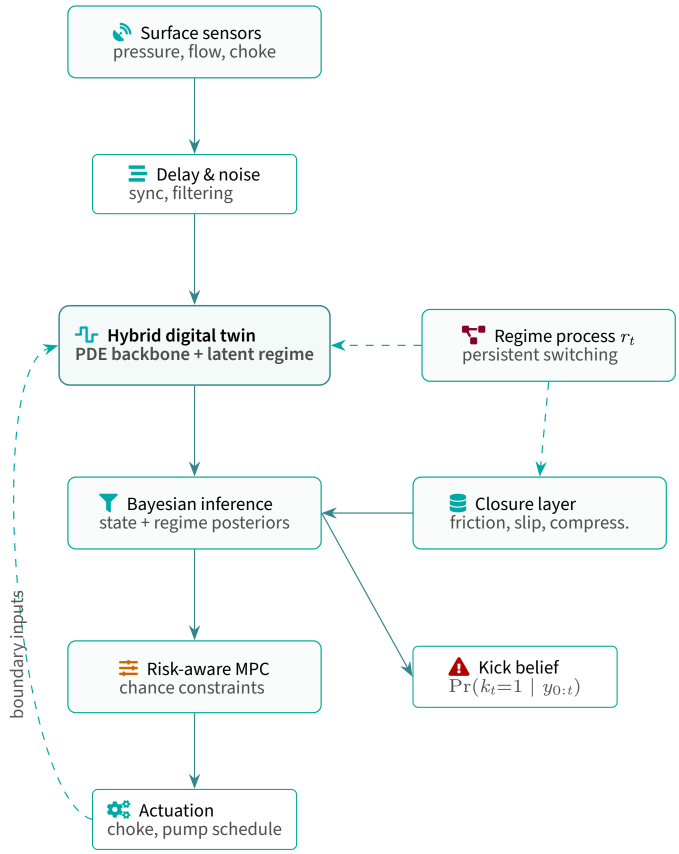

Figure 1. End-to-end architecture of the regime-conditioned Bayesian digital twin: delayed surface signals are assimilated into a hybrid physics backbone with a latent regime process, producing joint posteriors over downhole state and regimes that drive probabilistic kick belief and risk-aware MPC for pressure-window regulation.

fine phase volumetric flow rates $Q_g = A\alpha u_g$ and $Q_\ell = A(1-\alpha)u_\ell$, and define mixture flow rate $Q_m = Q_g + Q_\ell$. The phase mass balances can be written as

$$\frac{\partial}{\partial t}(A\alpha\rho_g) + \frac{\partial}{\partial s}(\rho_g Q_g) = \Gamma_g(s,t), \tag{1}$$

$$\frac{\partial}{\partial t}(A(1-\alpha)\rho_\ell) + \frac{\partial}{\partial s}(\rho_\ell Q_\ell) = \Gamma_\ell(s,t). \tag{2}$$

where Γ_g and Γ_ℓ represent volumetric mass sources per unit length arising from influx, dissolution, or phase change. In drilling hydraulics, a gas kick can be represented as a localized positive Γ_g supported on a spatial region near the influx point and active during a short time window [10]. Even when influx is absent, compressibility can make the apparent storage terms significant in gas-bearing sections, making transient inference sensitive to both equation-of-state modeling and closure accuracy.

For momentum, a complete two-fluid formulation includes separate phase momentum balances with interfacial forces. For inference, it can be advantageous to work with a mixture momentum balance augmented by slip and interfacial residuals, because pressure is commonly shared between phases in long slender conduits and because surface pressure signals constrain mixture momentum more directly than individual phase momenta. Define mixture density $\rho_m = \alpha\rho_g + (1-\alpha)\rho_\ell$ and mixture superficial velocity $v_m = Q_m/A$. A reduced mixture momentum balance can be expressed as

$$\frac{\partial}{\partial t}(\rho_m v_m) + \frac{\partial}{\partial s}\left(\rho_m v_m^2 + p\right) = -\rho_m g \sin\theta(s)[11] - \frac{\tau_w P_w}{A} + \mathcal{I}_{\text{slip}} + \mathcal{R}_{\text{mod}}. \tag{3}$$

where τ_w is wall shear stress, $\mathcal{I}_{\text{slip}}$ denotes residual terms arising from slip and phase segregation not captured by the mixture form, and $\mathcal{R}_{\text{mod}}$ aggregates modeling discrepancy including neglected acceleration terms, eccentricity effects, and numerical diffusion. The decomposition is deliberate: τ_w and $\mathcal{I}_{\text{slip}}$ are precisely where regime dependence enters, and thus where regime-conditioned closure learning is most valuable.

To close the system, relationships between α, v_m, and slip must be specified. A widely used reduced closure is drift-flux, in which the gas velocity is written as a mixture velocity plus a drift term. One can express

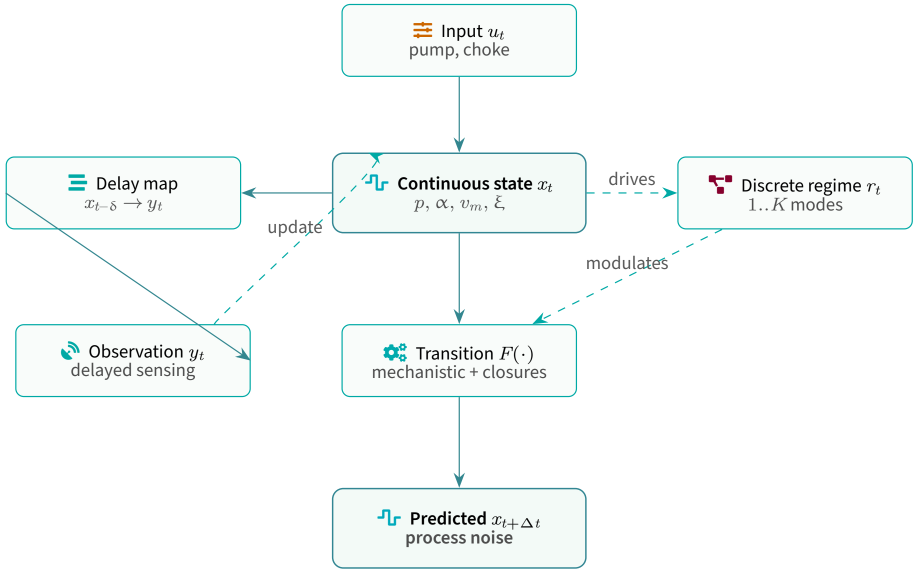

Figure 2. Single-step hybrid state-space structure: the continuous annular state evolves through a mechanistic transition with regime-conditioned closures, while delayed surface measurements update the filtered belief; regime dynamics are persistent and state-dependent, enabling joint inference over x_t and r_t.

Table 1. Key notation used in the hybrid state-space formulation.

Symbol	Description	Units
s	Measured depth along the well trajectory	m
$p(s,t)$	Pressure field along the annulus	Pa
$\alpha(s,t)$	Gas void fraction (gas volume fraction)	–
$v_m(s,t)$	Mixture superficial velocity	m s^{-1}
Q_g, Q_ℓ	Gas and liquid volumetric flow rates	m^3/s
ρ_m	Mixture density $\rho_m = \alpha\rho_g + (1-\alpha)\rho_\ell$	kg/m^3

[12]

$$u_g = C_0 v_m + V_d, \qquad Q_g = A\alpha u_g, \tag{4}$$

where C_0 is a distribution parameter and V_d is a drift velocity capturing buoyancy-driven relative motion and segregation. In inclined and horizontal sections, C_0 and V_d become strongly dependent on topology, because stratification can dominate and the assumption of dispersed gas breaks down. Similarly, wall shear stress can be related to a friction factor f through [13]

$$\tau_w = \frac{1}{8} f \rho_m v_m^2, \tag{5}$$

but the effective f depends on rheology, eccentricity, and regime. For non-Newtonian drilling fluids, apparent viscosity depends on shear rate, which itself depends on v_m and geometry. A common representation is Herschel–Bulkley, with shear stress $\tau = \tau_y + K\dot{\gamma}^n$, leading to apparent viscosity $\mu_{\text{app}} = (\tau_y + K\dot{\gamma}^n)/\dot{\gamma}$. In an annulus, $\dot{\gamma}$ varies radially, and eccentricity concentrates shear in the narrow side. Capturing these effects exactly is complex, so the mechanistic backbone uses a baseline rheological model and a baseline friction correlation, while a learnable correction term accounts for systematic deviations in an inference-consistent way.

The hybrid nature of the system suggests introducing a discrete regime variable that modulates closures. The regime variable does not need to correspond exactly to a named regime taxonomy; it needs to partition the closure space into a small number of modes that are internally coherent [14]. This is motivated by empirical observations that horizontal and high-inclination two-phase flow can exhibit a richer set of regimes and transitions than near-vertical flow, and that data-driven classifiers can resolve major and minor regions in horizontal maps, including a cross-validated nearest-neighbor model reported with 97.4% accuracy under a particular dataset and workflow [15]. In the present framework, that richness is reflected not by hard-coded maps but by a latent discrete process $r_t \in \{1, \dots, K\}$ that influences the closure terms for slip and friction.

Spatial discretization converts the PDE system into a large-scale dynamical system. Partition the well into N cells of length Δs_i with cell-averaged variables. Let $p_i(t)$, $\alpha_i(t)$, and $v_{m,i}(t)$ denote cell averages. Let x_t

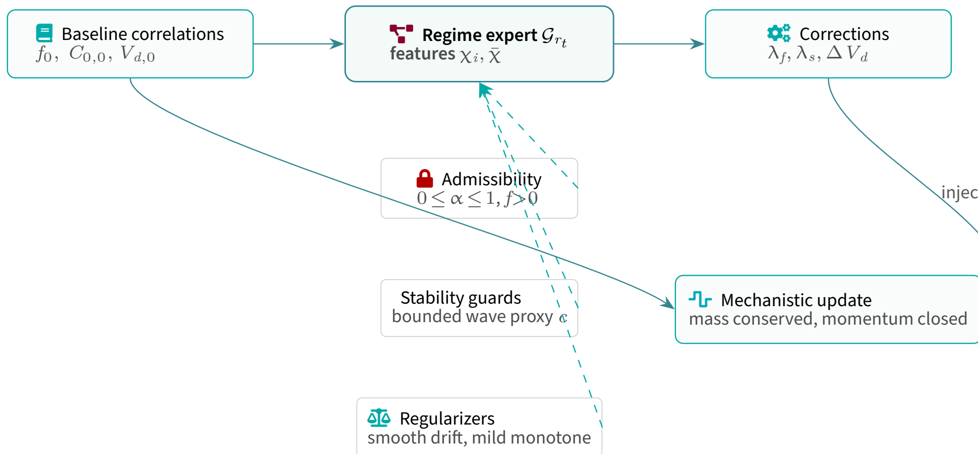

Figure 3. Regime-conditioned closure construction: baseline correlations are corrected by a regime expert that outputs bounded multipliers for friction and slip channels; admissibility and stability constraints regularize the closure so state evolution remains physically coherent and numerically well behaved.

Table 2. Illustrative flow regimes and their qualitative impact on closures.

Regime index r	Qualitative description	Dominant closure sensitivity
1	Dispersed bubbly / quasi-homogeneous mixture	Sensitive to drift-flux distribution parameter C_0 and mixture compressibility
2	Intermittent slug / churn structures	Strong sensitivity in slip and effective friction, pronounced hysteresis
3	Stratified or segregated configurations in inclined sections	Interfacial shear and gravity-induced segregation dominate momentum exchange
4	Annular or mist-like transport at high velocities	Wall shear and entrainment control pressure losses and void profiles
5	Transitional or mixed-pattern regions	All closure channels have elevated uncertainty and variability

denote the concatenated state vector

$$x_t = \left[p_{1:N}(t),\ \alpha_{1:N}(t),\ v_{m,1:N}(t),\ \xi_{1:N}(t)\right], \tag{6}$$

where $\xi_{1:N}(t)$ denotes auxiliary closure states such as friction multipliers, slip multipliers, or compressibility residuals. Boundary inputs u_t include pump rate, choke opening, and any controlled backpressure at the surface [16]. The numerical time advance over timestep Δt is written as

$$x_{t+\Delta t} = F(x_t, u_t, r_t; \phi, \psi) + w_t, \tag{7}$$

where ϕ denotes known parameters and geometry, ψ denotes learnable closure parameters, and w_t denotes process noise capturing unmodeled disturbances. The mapping F is built from a stable finite-volume or finite-difference discretization of the conservation laws, augmented by closure terms that depend on r_t and on the closure state ξ. The additive noise w_t is not merely a mathematical convenience; it represents uncertainty in friction due to cuttings concentration, uncertainty in gas compressibility due to temperature gradients, and unmodeled phenomena such as micro-annulus effects and sensor-induced boundary fluctuations.

A measurement model maps the latent state to observed surface signals [17]. Let y_t denote measurements at time t, such as surface pressure channels and flow rates. A generic observation model is

$$y_t = H(x_{t-\delta}, u_{t-\delta}; \kappa) + b + v_t, \tag{8}$$

where δ represents sensor and transport delay, κ denotes known sensor mapping parameters, b denotes bias terms, and v_t denotes observation noise. The delay can differ across channels and can be time varying. The digital twin accommodates this by maintaining a fixed-lag window of past states and applying delayed updates, rather than assuming instantaneous observation [18].

The regime process is modeled as a Markov chain whose transition probabilities depend on state and input, allowing persistence and hysteresis. A logit parameterization is

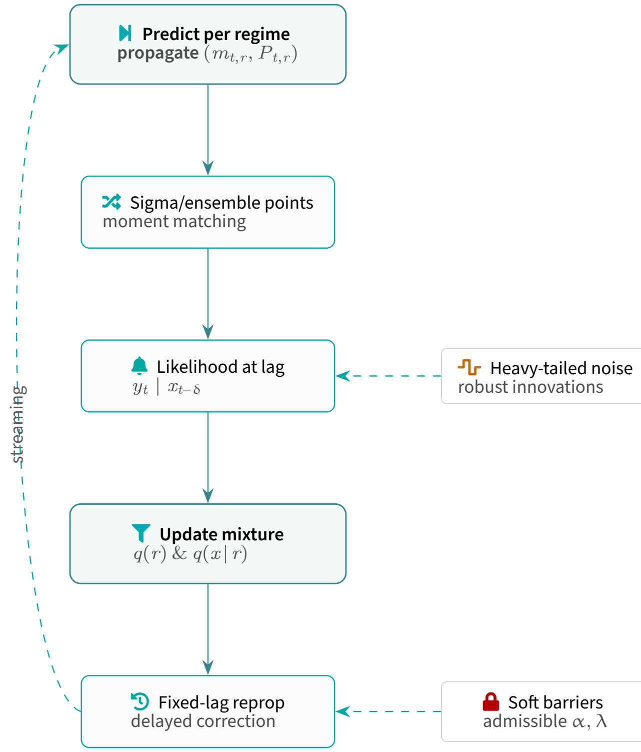

Figure 4. Constrained variational filtering loop: regime-conditioned prediction is paired with delayed likelihood evaluation and mixture updates, then a fixed-lag repropagation reconciles late-arriving information; robust noise and admissibility barriers stabilize inference under sensor imperfections and closure uncertainty.

Table 3. Main components of the latent state vector in the discretised model.

Block	Cor	Physical meaning	Dimension
Pressure field	$p_{1:}$	Cell-averaged annular pressure along the well path	N
Void fraction field	$\alpha_{1:}$	Gas volume fraction profile, encoding topology and compressibility	N
Mixture velocity field	$v_{m,}$	Mixture superficial velocity governing momentum transport	N
Closure state	$\xi_{1:}$	Latent friction and slip multipliers and related closure variables	N
Influx state (optional)	$(k_{t,}$	Kick activity indicator and associated influx rate	2

$$\Pr(r_{t+\Delta t} = j \mid r_t = i, x_t, u_t) = \exp\big(\eta_{ij} + \zeta_{ij}^\top \varphi(x_t, u_t)\big) \Big/ [19] \sum_{k=1}^{K} \exp\big(\eta_{ik} + \zeta_{ik}^\top \varphi(x_t, u_t)\big) \tag{9}$$

where η_{ij} encode baseline persistence and φ collects features such as mixture velocity, void fraction gradient, inclination, and boundary transients. The persistence structure is important: real flows do not switch regimes at every timestep even when operating near a boundary, and inference should respect that by penalizing implausibly rapid switching.

The above defines the backbone of a hybrid state-space model. The remaining challenge is to specify closure representations that are flexible enough to capture regime-dependent effects but constrained enough to preserve physical plausibility and numerical stability. That challenge is addressed by regime-conditioned closure operators, which are developed next [20].

3. Regime-Conditioned Closure Operators with Admissibility Constraints

Closure uncertainty dominates the gap between idealized conservation-law models and field behavior. In annular multiphase flow, the largest discrepancies commonly arise from effective friction under evolving rheology and geometry, and from slip and segregation effects that alter mixture density and momentum loss. A digital twin that treats these discrepancies as unstructured noise often becomes either unstable or uninformative. The strategy here is to represent closures as structured latent fields with regime-conditioned dynamics, constrained so that inferred states remain admissible.

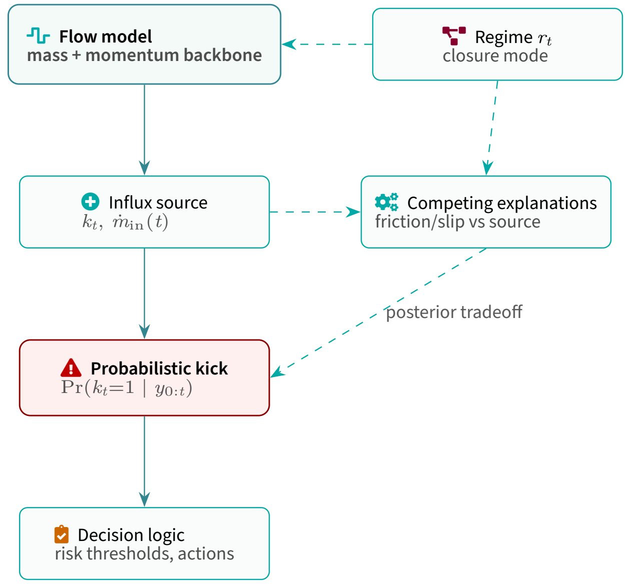

Figure 5. Integrated influx identification: a sparse nonnegative source model competes with regime-conditioned closure adjustments to explain persistent pressure/flow discrepancies, yielding a calibrated posterior kick probability that plugs directly into operational decision logic.

Table 4. Conceptual layers of the digital twin architecture.

Layer	Description	Representative elements	Section
Conservation backbone	One-dimensional gas–liquid conservation laws along the annulus	Mass and mixture-momentum balances, finite-volume discretisation	§2
Closure layer	Regime-conditioned corrections to friction, slip, and compressibility	Friction multipliers, drift-flux parameters, closure-state dynamics	§3
Regime process	Discrete latent process describing topology-dependent modes	State-dependent Markov chain with logit transition parameterisation	§2, §3
Inference layer	Online posterior inference over states, regimes, and influx	Constrained variational filtering with fixed-lag smoothing	§4
Control layer	Risk-aware pressure regulation using the posterior predictive	Stochastic MPC with chance constraints and risk measures	§5

The closure representation begins by separating baseline physics from corrections. Let the mechanistic backbone compute a baseline friction factor f_0 and baseline drift-flux parameters $(C_{0,0}, V_{d,0})$ using standard correlations and rheological models. The digital twin introduces correction multipliers $\lambda_{f,i}(t)$ and $\lambda_{s,i}(t)$ in each cell. Friction is modified as $f_i = \lambda_{f,i} f_{0,i}$, and drift terms are modified as $C_{0,i} = C_{0,0,i}\exp(\lambda_{s,i})$ and $V_{d,i} = V_{d,0,i} + \Delta V_{d,i}$, with $\Delta V_{d,i}$ represented by a bounded function of state. The exponential mapping for C_0 ensures positivity and avoids sign flips that would be physically nonsensical [21]. The friction multiplier is constrained to be positive and bounded above to prevent numerical blow-up.

These correction fields are not arbitrary. They are modeled as outputs of a regime-conditioned operator driven by local features. Let $\chi_i(t)$ denote a feature vector built from cell variables and known geometry, for example $\chi_i = [p_i, \alpha_i, v_{m,i}, \theta_i, D_{h,i}, \mu_{\mathrm{app},i}]$ along with time-derivative proxies such as $\Delta v_{m,i}/\Delta t$ to capture transients. A regime-conditioned neural closure can be written as

$$\lambda_i(t) = \mathcal{G}_{r_t}\big(\chi_i(t), \bar{\chi}(t)\big), \tag{10}$$

where λ_i collects $(\lambda_{f,i}, \lambda_{s,i}, \Delta V_{d,i})$ and $\bar{\chi}$ denotes global features such as pump rate and choke setting. To ensure admissibility, $\mathcal{G}_r$ is constructed with constrained outputs. A simple construction uses unconstrained network outputs $\tilde{\lambda}_{f,i}$ and $\tilde{\lambda}_{s,i}$ and maps them via

$$\begin{aligned} \lambda_{s,i} &= \lambda_{s,\min} \\ [22] \quad &+ \big(\lambda_{s,\max} - \lambda_{s,\min}\big) \\ &\times [?]\sigma(\tilde{\lambda}_{s,i}) \end{aligned} \tag{11}$$

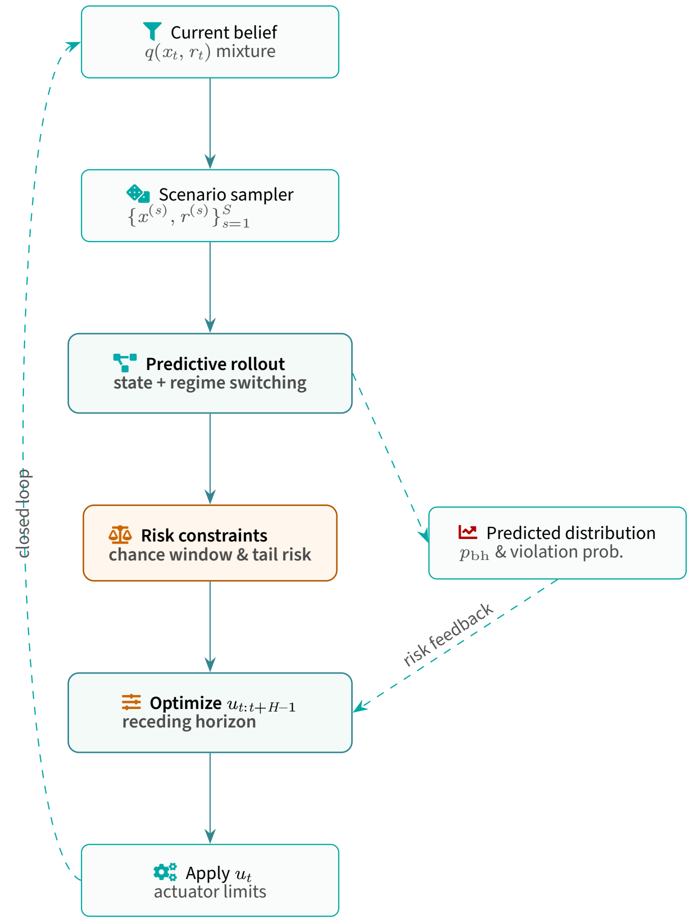

Figure 6. Regime-aware stochastic MPC: samples drawn from the joint posterior over state and regime are rolled out through the hybrid predictor to evaluate chance constraints on the pressure window and tail-risk penalties; the optimized action is applied in a receding-horizon loop with uncertainty-aware risk reporting.

with sigmoid $\sigma(\cdot)$ and conservative bounds $(\lambda_{f,\min}, \lambda_{f,\max})$, $(\lambda_{s,\min}, \lambda_{s,\max})$. The drift correction $\Delta V_{d,i}$ can be bounded similarly to prevent unrealistic relative velocities. These bounds can be set broadly enough to allow adaptation while excluding pathological values.

A central objective is to prevent the closure operator from violating conservation. The closure should not inject or remove mass; it should act through momentum and slip terms that are already present in the physical model. This is enforced by designing F so that closure affects only the wall shear and slip-related constitutive relations, leaving the discrete mass balance structure intact. When numerical schemes are conservative, this means that total mass changes are attributable only to boundary fluxes and explicit source terms such as influx, not to closure artifacts [23].

Numerical stability imposes additional constraints. Hyperbolic conservation laws have characteristic speeds that should remain bounded for a given physical system. In practice, unstable closures can introduce effective compressibility that yields unrealistically fast pressure waves, causing oscillations and filter divergence. To regularize this, the digital twin computes an implied wave speed proxy from local mixture compressibility. If ρ_g follows an equation of state and the liquid is nearly incompressible, the mixture bulk modulus can be approximated by a weighted harmonic relation, leading to an effective sound speed c_i that depends on α_i and on the gas compressibility [24]. While the precise acoustic model may be approximate, it is still useful as a stability guard. A differentiable penalty enforces that c_i remains within conservative bounds:

$$\mathcal{P}_c = \sum_t \sum_{i=1}^{N} \Big[\max\big(0,\ c_i(t) - c_{\max}\big)\Big]^2$$

$$[25] + \sum_t \sum_{i=1}^{N} \Big[\max\big(0,\ c_{\min} - c_i(t)\big)\Big]^2 \qquad (12)$$

The bounds can be selected based on phase properties and operational expectations, and can be widened if necessary to avoid undue bias.

A second stability-related constraint concerns the relationship between friction and flow rate. Although multiphase friction can be nonmonotone near transitions, gross nonphysical behavior such as decreasing

Table 5. Regime-conditioned closure channels and associated constraints.

Channel	Baseline model	Regime-dependent modifier	Constraint
Wall friction factor f	Correlation based on hydraulic diameter and rheology	Cell-wise multiplier $\lambda_{f,i}(t)$ driven by features and regime r_t	$\lambda_{f,i} \in [\lambda_{f,\min}, \lambda_{f,\max}]$
Drift-flux parameter C_0	Standard drift-flux relation	Exponential correction $C_{0,i} = C_{0,0,i} \exp(\lambda_{s,i})$	Positivity and bounded magnitude for $C_{0,i}$
Drift velocity V_d	Inclination- and density-difference-based estimate	Additive correction $\Delta V_{d,i}$ from a neural operator	Bounds on relative velocity to avoid unphysical slip
Closure state ξ_t	Static or slowly varying parameters	Mean-reverting dynamics toward regime-conditioned targets	Smooth temporal evolution with rate parameter γ
Transition logits	Static regime transition matrix	State- and input-dependent logits $\eta_{ij} + \zeta_{ij}^\top \varphi(x_t, u_t)$	Encouraged persistence and controlled switching frequency

Table 6. Stability and admissibility regularisers used in learning and inference.

Penalty	Symbol	Target property	Implementation
Wave-speed bounds	$\mathcal{P}_c$	Effective acoustic speed c_i remains within conservative limits	Squared penalties on violations of $[c_{\min}, c_{\max}]$ across cells and time
Friction monotonicity	$\mathcal{P}_m$	Avoid grossly non-monotone friction with respect to $\lvert v_m \rvert$	Penalise negative $\partial \lambda_{f,i} / \partial \lvert v_{m,i} \rvert$ via automatic differentiation
Regime separation	$\mathcal{P}_{\text{sep}}$	Prevent collapse of all regimes to identical closures	Exponential penalty on small pairwise distances between ψ_r and $\psi_{r'}$
Void-fraction bounds	–	Maintain $0 \le \alpha_i \le 1$ and avoid near-singular states	Sigmoid parameterisation of α_i and soft barriers near pathological limits
Closure shrinkage	–	Keep regime-specific parameters close to a shared reference	Hierarchical Gaussian priors with mean reversion towards μ_ψ

friction multiplier with increasing mixture velocity across broad ranges can destabilize inference by creating spurious equilibria. A soft monotonicity regularizer can be implemented by penalizing negative derivatives of $\lambda_{f,i}$ with respect to $|v_{m,i}|$:

$$\mathcal{P}_m = \sum_t \sum_{i=1}^{N} \Big[\max\big(0, -\partial \lambda_{f,i}(t) / \partial |v_{m,i}(t)|\big) \Big]^2. \qquad (13)$$

This penalty is computed via automatic differentiation within the closure network [26]. It does not force strict monotonicity, but it discourages the closure from exploiting noise in a way that conflicts with basic physical intuition over the operating range.

The discrete regime variable r_t modulates closure through a mixture-of-experts architecture. Each regime has its own parameter set ψ_r, but parameters are tied by a hierarchical prior to share statistical strength and avoid overfitting. A convenient hierarchical structure is

$$\psi_r \sim \mathcal{N}(\mu_\psi, \Sigma_\psi), \qquad \mu_\psi \sim \mathcal{N}(0, \sigma_\mu^2 I), \qquad (14)$$

with Σ_ψ diagonal or low-rank for computational simplicity [27]. This shrinks regime-specific parameters toward a common mean unless data supports separation. To avoid regime collapse, where all regimes become identical and the discrete variable becomes irrelevant, a separation regularizer encourages distinct closure behaviors:

$$\mathcal{P}_{\text{sep}} = \sum_{r<r'} \exp\Big(-\frac{1}{\tau} |0\psi_r - \psi_{r'}|0^2 \Big), \qquad (15)$$

where τ tunes the strength. Because this regularizer is soft, it allows regimes to remain similar if the system truly exhibits little topology dependence under a given dataset, but it prevents degeneracy in which regimes are unused.

The closure fields can also be endowed with their own temporal dynamics, reflecting that friction multipliers and slip corrections should evolve smoothly unless operating conditions change abruptly [28]. This is represented by a latent evolution equation

$$\xi_{t+\Delta t} = \xi_t + \Delta t \, \mathcal{D}_{r_t}(\xi_t, x_t, u_t) + \epsilon_t, \qquad (16)$$

Table 7. Probabilistic influx-source submodel embedded in the twin.

Quantity	Symbol	Prior / model	Role in the hybrid system
Activity indicator	k_t	Bernoulli(π_t) with state- or scenario-dependent π_t	Switches between no-influx and active-influx regimes
Influx rate	$\dot{m}_{\text{in}}(t)$	Lognormal prior when $k_t = 1$, zero otherwise	Sets strength of the gas source entering the annulus
Spatial weight	$\omega(s)$	Prescribed profile over a suspected depth interval	Localises influx to a specific region in the gas mass balance
Activation probability	π_t	Calibrated or learned function of operating conditions	Encodes prior belief on kick likelihood at each time step
Gas mass source	$\Gamma_g(s,t)$	$\Gamma_g(s,t) = \dot{m}_{\text{in}}(t)\,\omega(s)$	Couples influx model into the conservation-law backbone

Table 8. Observation channels and their treatment in the measurement model.

Channel	Measured variable	Latency handling	Noise model
Standpipe pressure	Surface pump-side pressure trace	Fixed-lag window with delayed state update at time $t - \delta_{\text{sp}}$	Approximately Gaussian additive noise with heavy-tailed robustness
Casing / annulus pressure	Surface annulus pressure at wellhead	Channel-specific delay δ_c and possible bias term	Gaussian with bias correction and online innovation checks
Flow-in rate	Pump output flow into the annulus	Treated as both input and noisy observation with delay	Jointly modelled with small variance to reflect actuator authority
Flow-out rate	Return flow at the surface	Delayed signal constrained by transport time through the annulus	Heavy-tailed noise to tolerate transient mismatches and cuttings effects
Choke / backpressure	Choke position or applied backpressure	Directly observed control variable, optionally filtered	Low-variance measurement used to condition boundary states

where ξ collects closure states and ϵ_t is noise. The operator $\mathcal{D}_r$ can be as simple as a mean-reverting term toward the instantaneous closure network output, which stabilizes adaptation:

$$\xi_{t+\Delta t} = (1-\gamma)\xi_t + \gamma\,\hat{\xi}_{r_t}(x_t, u_t) + \epsilon_t, \qquad (17)$$

with $\gamma \in (0,1)$ controlling the adaptation rate and $\hat{\xi}_r$ the network output. This makes closure adaptation neither instantaneous nor frozen, and it helps prevent the closure from absorbing short-lived measurement noise.

The above design yields closures that are expressive yet constrained. They can adjust friction and slip in regime-dependent ways while preserving mass conservation, enforcing positivity, and maintaining stability proxies. The remaining challenge is inference: computing the posterior distribution over the high-dimensional continuous state and the discrete regime process from delayed noisy measurements, while optionally learning closure parameters from data [?]. That inference problem is addressed next.

4. Constrained Bayesian Inference with Regime and Influx Source Identification

Online inference in a hybrid multiphase model is difficult because it combines nonlinear continuous dynamics, a discrete regime process, delayed observations, and uncertain closures. The target is a filtered posterior at each time step that provides both point estimates and uncertainty quantification. The approach here uses a constrained variational filtering framework that yields a tractable approximation to the posterior while respecting admissibility constraints.

Let $x_{0:T}$ denote the continuous state trajectory and $r_{0:T}$ denote the regime trajectory. Let $y_{0:T}$ denote observations and $u_{0:T}$ denote known inputs. The joint

Table 9. Components of the stochastic model predictive control cost.

Component	Expression (schematic)	Weight	Interpretation
Pressure tracking	$\left(p_{\mathrm{bh}}(x_\tau) - p^\star_{\mathrm{bh}}\right)^2$	ω_p	Penalises deviations of bottom-hole pressure from a target value or profile
Control effort	$\|0u_\tau - u_{\mathrm{ref}}\|0^2$	ω_u	Limits aggressive choke and pump adjustments and favours smooth actions
Influx risk	$\Pr(k_\tau = 1 \mid x_\tau, r_\tau)$	ω_k	Discourages actions that increase the probability or severity of kicks
Regime penalty	$\rho(r_\tau)$	ω_r	Encodes preferences against regimes associated with higher uncertainty or risk
Terminal cost	$\ell_T(x_{t+H})$	–	Shapes the horizon-end state towards safe, well-pressurised configurations

Table 10. Representative simulation scenarios used for evaluation.

Scenario	Description	Dominant uncertainties	Performance focus
Benign pump and choke ramps	Routine transients with no influx under varying rates	Rheology, compressibility, and moderate friction mismatch	State-estimation accuracy and uncertainty calibration
Friction-mismatch tests	Time-varying friction multipliers along the well	Cuttings loading and eccentricity-induced friction variation	Robustness of closure learning and bias reduction
Inclination change / hybrid structure	Sections with varying inclination and topology transitions	Regime occupancy and transition hysteresis	Regime inference quality and stability of the hybrid model
Sensor latency and dropout	Channel-specific delays and intermittent missing data	Time-varying latency and observation availability	Fixed-lag smoothing and resilience of the filter
Gas influx events	Localised gas kicks of different onset times and magnitudes	Influx rate and depth, interaction with closures	Detection timeliness, false-alarm control, and regulated response

model factorizes as [29]

$$p(x_{0:T}, r_{0:T}, y_{0:T} \mid u_{0:T}) = p(x_0, r_0) \times \prod_t [30] p(x_{t+\Delta t} \mid x_t, u_t, r_t) \times p(r_{t+\Delta t} \mid r_t, x_t, u_t) \times p(y_t \mid x_{t-\delta}, u_{t-\delta}) \quad (18)$$

The transition density $p(x_{t+\Delta t} \mid x_t, u_t, r_t)$ is induced by the deterministic integrator F and by process noise w_t. Because multiphase systems can experience occasional large disturbances, a heavy-tailed noise model improves robustness [31]. One representation is a scale mixture of Gaussians:

$$w_t \mid \tau_t \sim \mathcal{N}(0, \tau_t Q), \qquad \tau_t \sim \mathrm{InvGamma}(a, b), \quad (19)$$

which yields Student-t marginal behavior for w_t. This allows the filter to accommodate occasional mismatches without permanently inflating the covariance.

To handle regime and continuous uncertainty jointly, a natural approach is a Rao–Blackwellized representation in which the posterior is a mixture over regimes with regime-conditioned continuous distributions. The variational approximation adopts the structure [32]

$$q(x_t, r_t) = q(r_t)\, q(x_t \mid r_t), \quad (20)$$

where $q(r_t)$ is categorical and $q(x_t \mid r_t)$ is Gaussian with mean $m_{t,r}$ and covariance $P_{t,r}$. This yields a mixture of K Gaussians for x_t, weighted by regime probabilities. The mixture representation is compact and can be propagated online.

Prediction proceeds by propagating each regime-conditioned Gaussian through the nonlinear transition model. Because direct analytic propagation is generally unavailable, a sigma-point or ensemble method is used [33]. For each regime r, a set of samples $\{x_t^{(s)}\}$ is drawn from $\mathcal{N}(m_{t,r}, P_{t,r})$ using a deterministic cubature rule or Monte Carlo. Each sample is advanced by the integrator with regime-conditioned closure:

$$x_{t+\Delta t}^{(s)} = F(x_t^{(s)}, u_t, r; \phi, \psi) + w_t^{(s)}. \tag{21}$$

The predicted mean and covariance are then obtained by moment matching. This is an approximation, but it is stable when the timestep is small and the dynamics are not excessively stiff, and it is compatible with differentiable implementations when learning closures.

Regime probabilities are predicted using the transition model:

$$\hat{q}(r_{t+\Delta t} = j) = \sum_{i=1}^{K} q(r_t = i)\, \mathbb{E}_{q(x_t \mid r_t = i)}\big[\Pr(r_{t+\Delta t} = j \mid r_t = i, x_t, u_t)\big]. \tag{22}$$

The expectation can be approximated by the same samples used for the continuous propagation [34]. This couples regime prediction to the inferred continuous state, enabling state-dependent switching.

Measurement updates incorporate delayed observations. Because y_t depends on $x_{t-\delta}$, the filter maintains a fixed-lag window of length L_δ that covers the maximum expected delay. When y_t arrives, it updates the approximate posterior for the delayed state, then repropagates forward through the window to update the current state. This fixed-lag smoother avoids the brittleness of assuming known exact delay while remaining computationally bounded.

Let $z = t - \delta$ denote the effective state time for the observation [35]. For each regime component, the measurement likelihood is computed as

$$\ell_r(y_t) = \mathbb{E}_{q(x_z \mid r_z = r)}\big[p(y_t \mid x_z, u_z)\big]. \tag{23}$$

If the observation model is approximately Gaussian, then the likelihood can be computed by linearization around the predicted mean, yielding a Kalman-like update. Otherwise, likelihood can be approximated by sampling. The regime weights are updated by Bayes' rule:

$$q(r_z = r) \propto \hat{q}(r_z = r)\, \ell_r(y_t), \tag{24}$$

and the continuous posteriors are updated by the corresponding Gaussian conditioning step [36].

Admissibility constraints are enforced by parameterization and by soft barriers. Void fraction is represented by an unconstrained variable $\tilde{\alpha}_i$ with $\alpha_i = \sigma(\tilde{\alpha}_i)$, guaranteeing $0 \le \alpha_i \le 1$. Densities are computed from pressure via equations of state that guarantee positivity. Closure multipliers are bounded via sigmoid mappings as described previously. In addition, barrier penalties discourage states that approach numerical singularities, such as α near 1 with vanishing liquid holdup in sections where liquid continuity is required for pressure transmission. These penalties enter the variational objective rather than hard projecting states, which helps maintain differentiability and avoids discontinuous updates.

Kick detection is integrated by augmenting the state with an influx source process [37]. Let $\dot{m}_{\text{in}}(t)$ denote influx mass rate, and let k_t denote an activity indicator. The model uses a sparse nonnegative prior:

$$\begin{aligned} k_t &\sim \text{Bernoulli}(\pi_t) \\ \dot{m}_{\text{in}}(t) \mid k_t = 0 &= 0 \\ \dot{m}_{\text{in}}(t) \mid k_t = 1 &\sim \text{LogNormal}\big(\mu_{\text{in}}, \sigma_{\text{in}}^2\big) \end{aligned} \tag{25}$$

and the source term in the gas mass balance is written as $\Gamma_g(s,t) = \dot{m}_{\text{in}}(t)\,\omega(s)$ with spatial weight function ω localized near a candidate influx depth. If influx depth is unknown, a small set of candidate depths can be represented by a categorical variable that is inferred jointly, but the present formulation focuses on the common operational case where influx is suspected in a known interval.

The posterior over k_t provides a probabilistic kick indicator [38]. The digital twin does not need to commit to a hard threshold; instead, it outputs $\Pr(k_t = 1 \mid y_{0:t})$ and the posterior distribution of $\dot{m}_{\text{in}}(t)$. This posterior view aligns with how operators reason under uncertainty and supports risk-aware control. It is also compatible with empirical evidence that machine learning methods can detect kick symptoms from available signals with high accuracy across both static and circulating scenarios, including supervised and unsupervised approaches reported with accuracies at or above 90% and with decision-tree performance reported at 96% and 98% for particular influx types and conditions [35]. The present framework differs in that it embeds detection into a physically constrained probabilistic model, making the alarm semantics consistent with hydraulic evolution and regime uncertainty.

Learning of closure parameters ψ can be performed offline or online. Offline learning uses historical or simulated data to maximize the marginal likelihood of observations under the model [39]. Online adaptation can improve performance under changing rheology, but it must be constrained for safety. The variational framework supports online learning by stochastic gradient ascent on an evidence lower bound with stability penalties:

$$\mathcal{L} = \mathbb{E}_q\Big[\log p(x_{0:T}, r_{0:T}, k_{0:T}, y_{0:T} \mid u_{0:T}) - \log q(x_{0:T}, r_{0:T}, k_{0:T})\Big]$$

$$- \beta_c \mathcal{P}_c$$

[40] $$- \beta_m \mathcal{P}_m$$

$$- \beta_{\text{sep}} \mathcal{P}_{\text{sep}} \tag{26}$$

To prevent drift, parameter updates are bounded, and a mean-reverting prior discourages large deviations from validated baseline parameters. In deployments where online learning is disallowed, the same inference algorithm runs with fixed ψ, still producing regime-conditioned posteriors and probabilistic kick indicators [41].

The inference machinery outputs a posterior over current downhole pressure profile, void fraction profile, closure multipliers, regime probabilities, and influx probability. The next step is to use these posteriors to compute control actions that regulate pressure and mitigate risk. This motivates a regime-aware stochastic control formulation.

5. Risk-Aware Control from Regime-Conditioned Posteriors

Pressure-window regulation in multiphase drilling is a control problem under severe uncertainty. Actions such as choke adjustments and pump-rate changes influence both the continuous hydraulic state and the probability of regime transitions, and they can either suppress or exacerbate influx growth [42]. A controller that ignores regime uncertainty can perform well in a narrow operating region but can destabilize when regime-dependent friction and compressibility shift the system response. Conversely, a controller that is overly conservative can degrade operational efficiency and still fail if it does not properly account for rare-event dynamics.

The proposed control layer is a stochastic model predictive controller that uses the digital twin as a probabilistic predictor. Let u_t denote the control input at time t, including choke opening and pump rate, subject to actuator bounds and slew-rate limits. Let x_t and r_t denote the latent state and regime [43]. Define a target bottomhole pressure $p^\star_{\text{bh}}$ or a target pressure trajectory. Define a control horizon of H steps. The objective is to choose a sequence $u_{t:t+H-1}$ minimizing expected cost under the posterior predictive distribution:

$$J(u_{t:t+H-1}) = \mathbb{E}\Big[\sum_{\tau=t}^{t+H-1} \ell(x_\tau, r_\tau, u_\tau) + \ell_T(x_{t+H}) \;\Big|\; y_{0:t}, u_{0:t-1}\Big]. \tag{27}$$

A suitable stage cost includes pressure regulation, control effort, and risk terms tied to influx probability and undesirable regime occupancy:

$$\ell(x, r, u) = \omega_p\big(p_{\text{bh}}(x) - p^\star_{\text{bh}}\big)^2$$

[44] $$+ \omega_u |0u - u_{\text{ref}}|0^2$$

$$+ \omega_k \Pr(k = 1 \mid x, r)$$

[45] $$+ \omega_r \rho(r) \tag{28}$$

where $\rho(r)$ is a regime penalty encoding that some regimes may be less controllable or associated with larger uncertainty in closures, and $\Pr(k = 1 \mid x, r)$ is the predicted influx activity probability given the state. The inclusion of $\Pr(k = 1 \mid x, r)$ in the objective integrates detection with response, discouraging actions that are likely to worsen an influx even if they momentarily reduce pressure error.

Safety constraints are critical. The bottomhole pressure must remain within a window $[p_{\min}, p_{\max}]$ with high probability. This is naturally expressed as a chance constraint:

$$\Pr\Big(p_{\min} \le p_{\text{bh}}(x_\tau) \le p_{\max}\Big) \ge 1 - \epsilon, \qquad \tau = t, \ldots, t + H, \tag{29}$$

where ϵ is a tolerated risk level [46]. Because the predictive distribution is typically a mixture over regimes and may be non-Gaussian, enforcing chance constraints exactly is difficult. A practical approximation uses scenario-based constraints. Draw S samples from the current posterior over (x_t, r_t), propagate each sample forward under candidate control sequences using the mechanistic integrator and regime transition model, and enforce that the fraction of violating scenarios is below ϵ:

$$\frac{1}{S}\sum_{s=1}^{S} \mathbf{1}\Big(p^{(s)}_{\text{bh}}(\tau) \notin [p_{\min}, p_{\max}]\Big) \le \epsilon, \qquad \tau = t, \ldots, t + H. \tag{30}$$

This yields a conservative approximation when S is sufficiently large, and it is compatible with mixture posteriors without requiring analytic forms.

Risk measures beyond chance constraints can further stabilize behavior [47]. Conditional value-at-risk (CVaR) penalizes tail outcomes more strongly than mean-squared error. Let Z denote the random variable representing maximum pressure excursion over the horizon, such as $Z = \max_\tau \max(0, p_{\text{bh}}(x_\tau) - p_{\max}, p_{\min} - p_{\text{bh}}(x_\tau))$. CVaR at level α can be represented as an optimization-friendly form:

$$\text{CVaR}_\alpha(Z) = \min_\zeta \Big[\zeta + \frac{1}{1-\alpha}\mathbb{E}\big[(Z - \zeta)_+\big]\Big]. \tag{31}$$

Including CVaR_α in the cost encourages control actions that reduce worst-case excursions under the posterior, which is desirable when the cost of violating the pressure window is high.

The regime process complicates prediction because control actions can influence regime transition probabilities by changing mixture velocity and void fraction. The predictive simulator therefore propagates both continuous states and regimes. In practice, this can be implemented by sampling regime transitions for each scenario according to the state-dependent transition probabilities [?]. Over short horizons, regime persistence often dominates, and the predictive distribution can be approximated by conditioning on the most likely regime or on a small set of high-probability regime sequences. This reduces computational cost without ignoring regime uncertainty entirely.

Optimization of the control sequence can be performed by direct shooting with stochastic objectives. If the integrator and closures are differentiable, gradient-based methods can be used with reparameterized noise and with a continuous relaxation for regime transitions during optimization. If gradients are unreliable due to discrete switching, derivative-free methods can be used over a low-dimensional parameterization of the control sequence, such as piecewise-constant choke settings and pump rates [48]. Because the control is applied in receding-horizon fashion, warm starts from the previous solution significantly reduce computation.

A key property of the regime-conditioned twin is that it can predict different pressure responses under different regimes for the same control action. This enables the controller to choose actions that are robust across regimes. For example, an action that appears safe under a dispersed-gas closure might violate the upper pressure bound under an intermittent closure with higher effective friction; the scenario-based chance constraint will reveal this and steer the controller to a safer alternative. Similarly, if an inferred regime is close to a transition boundary, the controller can avoid aggressive pump ramps that increase the probability of switching into a regime with larger uncertainty [49].

The output of the control layer is not merely a recommended action but a quantified risk profile. The system can report the predicted probability of window violation under the recommended action, the predicted distribution of bottomhole pressure, and the predicted change in influx probability. This supports human-in-the-loop decision making and enables conservative overrides when the inferred uncertainty is large.

The next section evaluates the framework in simulation and discusses deployment pathways, emphasizing robustness to modeling mismatch and sensor imperfections.

6. Evaluation and Deployment Pathways

Evaluation focuses on whether the regime-conditioned Bayesian digital twin improves decision-relevant performance relative to regime-agnostic baselines under conditions that emulate field uncertainties [50]. Because direct downhole ground truth is rarely available in operations, simulation is used as a controlled environment where true pressure and void fraction profiles are known, while observations are restricted to realistic surface channels with delay and noise.

The simulation environment uses a higher-fidelity multiphase model as the data-generating process. This ground-truth simulator differs from the digital twin in three ways to represent model mismatch. The frictional behavior includes time-varying multipliers driven by a latent cuttings-loading process, with spatial variation that increases in high-inclination segments. Slip behavior includes topology-dependent drift parameters that vary with inclination and superficial velocity ratio [51]. Compressibility includes a temperature gradient that affects gas density and bulk modulus. Measurements are generated with additive noise, bias, and delay, and boundary conditions include pump ramps, choke changes, and periods of connection-like flow interruptions.

Scenarios are designed to separate benign transients from hazardous ones. Benign scenarios include step changes in pump rate and choke adjustments with no influx, under varying friction mismatch. Hazard scenarios include gas influx events of varying onset times and magnitudes, injected at a specified depth interval [52]. Influx rates are chosen to create subtle early signatures that could be confused with friction changes. The evaluation also includes sensor faults such as intermittent dropouts and temporarily increased noise, to test robustness of the heavy-tailed noise model and fixed-lag smoothing.

Baselines include a regime-agnostic probabilistic filter that shares the same mechanistic backbone but uses a single closure model without a discrete regime variable, and a deterministic observer based on the mechanistic model with fixed tuned parameters. The regime-agnostic probabilistic filter represents a common compromise: it quantifies uncertainty but cannot attribute discrepancies to topology changes except by inflating process noise. The deterministic observer represents operational practice where a tuned model is used for forward prediction and deviations are interpreted qualitatively [53].

Metrics are computed for state accuracy, uncertainty calibration, detection timeliness, and control performance. State accuracy is evaluated for bottomhole pressure and for spatially averaged void fraction in key well segments. Calibration is evaluated by empirical coverage of nominal credible intervals, for example whether a 90% interval contains the true value approximately 90% of the time. Detection timeliness is evaluated by the delay between true influx onset and the time when the posterior $\Pr(k_t = 1 \mid y_{0:t})$ exceeds a probability threshold chosen to yield a target false alarm rate under benign scenarios. Control performance is evaluated by the fraction of time bottomhole pressure falls outside the window, the magnitude of excursions, and the control effort required.

In benign transient scenarios, the deterministic observer exhibits systematic bias when friction multipliers drift because it cannot adapt. The regime-agnostic probabilistic filter adapts partially by absorbing mismatch into process noise and closure drift, but it tends to broaden its posterior to maintain consistency, reducing informativeness [54]. The regime-conditioned digital twin reduces bias while maintaining sharper uncertainty because it can reallocate probability mass across regimes when the pressure response changes in a way consistent with topology dependence. This prevents the filter from treating regime shifts as unexplained noise and improves calibration by aligning uncertainty with modeled structure.

In scenarios featuring abrupt transitions induced by pump ramps in inclined sections, the regime-agnostic filter often responds with large covariance inflation, which can temporarily obscure other anomalies. The regime-conditioned twin instead increases posterior entropy over regimes while keeping continuous uncertainty controlled, producing posteriors that remain decision useful. This behavior depends on regime persistence modeling [55]. If self-transition probabilities are too low, regime chattering increases and can destabilize control; if they are too high, the model may resist necessary switching and accumulate bias. In the evaluation, moderate persistence combined with state-dependent transition logits yields stable regime inference across operating points.

In influx scenarios, early detection depends on whether the estimator attributes small discrepancies to benign uncertainty or to an active source. The regime-conditioned twin benefits from explicit source modeling: it can explain sustained deviations by a nonzero influx rate rather than by unbounded friction corrections. This yields earlier rise in $\Pr(k_t = 1 \mid y_{0:t})$ at matched false alarm rates. The regime-agnostic probabilistic filter can misattribute early influx signatures to friction drift, delaying detection until the mismatch becomes pronounced [56]. The deterministic observer can detect anomalies quickly in some cases but provides no calibrated probability, leading to either overly sensitive alarms or missed events depending on thresholding.

Control evaluation uses the stochastic MPC layer with chance constraints. In benign scenarios, the controller maintains pressure within bounds with fewer aggressive choke moves than a controller based on point estimates, because it hedges against uncertainty and avoids actions with large tail risk. In influx scenarios, the controller's objective includes predicted influx probability, which steers it toward actions that stabilize pressure while suppressing influx growth. Because prediction accounts for regime uncertainty, the controller avoids action sequences that are safe only under a single assumed regime [57]. This reduces the frequency and severity of pressure-window violations in scenarios where regime switching is likely.

Computational feasibility is assessed by measuring per-step inference and control time under representative discretizations. The regime-conditioned filter maintains K Gaussian components, each over N cells. Complexity scales with KN for linear operations and with KS for sample propagation, where S is the ensemble size per regime. Stable performance is obtained with modest K and S when the integrator is vectorized and when closure networks are small [58]. Control optimization is more expensive, but it can be executed at a slower rate than filtering, and it benefits from warm starts and from reusing scenario samples across candidate control updates.

Deployment in operational environments requires additional safeguards and integration considerations. Sensor streams can be inconsistent, delayed, or temporarily unavailable. The fixed-lag inference approach naturally tolerates delay variability, but long dropouts reduce identifiability. In such periods, the system should increase uncertainty and, if used for control, the controller should shift toward conservative actions or yield authority to a fallback policy [59]. A practical integration architecture therefore includes a supervisory layer that monitors posterior calibration proxies, such as innovation statistics and regime entropy, and triggers fallback when anomalies indicate filter divergence.

Online learning of closures can be valuable when fluid properties change, but it must be constrained. One deployment strategy is to permit parameter adaptation only during confirmed benign circulation, defined by low inferred influx probability and by stable operating conditions. Another is to use dual models, where a validated static model drives control and an adaptive model runs in shadow mode to propose updates that are accepted only after consistency checks. The hierarchical prior and bounded parameter updates reduce risk, but operational governance should still treat online adaptation as a controlled process with auditability [15].

Interpretability matters for operator trust. Although the regime variable is latent and not necessarily a named regime, the system can provide descriptive summaries derived from posterior state fields, such as whether inferred void fraction gradients indicate segregation or whether inferred slip multipliers indicate strong relative motion. Such summaries can be displayed alongside classical indicators, providing context without claiming exact matches to external flow maps. The most important interpretability output is uncertainty: credible intervals and risk probabilities provide a transparent basis for decisions, and they can be validated statistically against future measurements even without downhole ground truth.

Overall, the evaluation indicates that embedding regime as a latent variable within a physically constrained probabilistic model improves robustness to mismatch and supports more stable risk-aware control [60]. The next section summarizes the main conclusions and outlines directions for extension.

7. Conclusion

This paper introduced a regime-conditioned Bayesian digital twin for annular gas–liquid flow that unifies mechanistic conservation laws, topology-dependent closures, online inference under delay, probabilistic influx detection, and risk-aware pressure control. The central idea is that regime should not be treated as an external label or as a post hoc classification output, but as a latent discrete process that modulates closure structure within a single probabilistic semantics. This alignment makes state estimation, regime inference, and event detection mutually consistent: the model cannot change regimes without affecting likelihood, cannot drift closures arbitrarily without incurring stability penalties and priors, and cannot explain sustained discrepancies without increasing the posterior probability of an influx source when the data supports that explanation.

Technically, the digital twin combines a one-dimensional conservation-law backbone with regime-conditioned closure operators that act through friction and slip terms while preserving mass conservation and enforcing admissibility constraints [61]. A constrained variational filtering method was developed to maintain a tractable mixture representation over regimes with regime-conditioned continuous posteriors, incorporating heavy-tailed noise for robustness and fixed-lag smoothing for delayed observations. The inference model was augmented with a sparse nonnegative influx source process, yielding probabilistic kick indicators and posterior influx-rate estimates that can be integrated directly into decision logic. A stochastic model predictive control layer then used posterior predictive simulations to enforce chance constraints on pressure windows and to minimize risk-sensitive objectives that hedge against regime uncertainty and suppress influx growth.

Simulation studies under friction mismatch, inclination changes, sensor latency, and influx disturbances supported that the regime-conditioned formulation improves posterior calibration and stabilizes control relative to regime-agnostic probabilistic baselines and deterministic observers, while remaining compatible with real-time computational budgets under modest numbers of regimes and ensemble samples. Deployment considerations emphasize safeguards against divergence, conservative fallback policies under sensor dropouts, and constrained strategies for any online closure adaptation. Extensions of the framework can incorporate richer rheology and cuttings transport states, more expressive regime transition models, and systematic validation against field datasets where partial downhole sensing is available for calibration and benchmarking [62].

References

[1] M. M. Rajabi, B. Rohde, N. Maguire, R. S. Stave, and R. Tapper, "Successful implementations of tophole managed pressure cementing and managed pressure drilling in the caspian sea," in SPE/IADC Managed Pressure Drilling and Underbalanced Operations Conference and Exhibition, SPE, 3 2012.

[2] "Documentation of the oil and gas supply module (ogsm)," 1 1996.

[3] T. N. Ofei, S. Irawan, and W. Pao, "Cuttings-liquid frictional pressure loss model for horizontal narrow annular flow with rotating drillpipe," IOP Conference Series: Materials Science and Engineering, vol. 78, pp. 012009–, 4 2015.

[4] S. A. Holditch, H. Xiong, J. Rueda, and Z. Rahim, "Using an expert system to select the optimal fracturing fluid and treatment volume," in SPE Gas Technology Symposium, SPE, 6 1993.

[5] S. Rosenberg, K. Kotow, J. Wakefield, S. Sampietro, and B. Hulett, "Riserless casing drilling for mitigation of shallow hazards - a paradigm shift in deepwater well construction," in IADC/SPE International Drilling Conference and Exhibition, SPE, 3 2022.

[6] K. Manikonda, A. R. Hasan, C. E. Obi, R. Islam, A. K. Sleiti, M. W. Abdelrazeq, and M. A. Rahman, "Application of machine learning classification algorithms for two-phase gas-liquid flow regime identification," in Abu Dhabi International Petroleum Exhibition and Conference, p. D041S121R004, SPE, 2021.

[7] A. Duguid and R. McBeth, "Drake f-76 in-situ abandonment of a high arctic offshore completion and facilities," in Annual Technical Meeting, PETSOC, 6 1997.

[8] D. V. Eliseev, M. Y. Golenkin, A. Senkov, A. Latypov, I. Bulygin, A. Ruzhnikov, M. Buyanov, and A. Kashlev, "New vision: Ic taml5 wells on caspian offshore. reasons, implementation and results," in SPE Russian Petroleum Technology Conference and Exhibition, SPE, 10 2016.

[9] M. W. Massey, "Marine well containment company progress," in Offshore Technology Conference, OTC, 4 2012.

[10] A. Sowiżdżał, M. Starczewska, and B. Papiernik, "Future technology mix—enhanced geothermal system (egs) and carbon capture, utilization, and storage (ccus)—an overview of selected projects as an example for future investments in poland," Energies, vol. 15, pp. 3505–3505, 5 2022.

[11] R. Stork, "Computer model simplifies gas field production forecasting and development planning," The APPEA Journal, vol. 27, pp. 335–343, 6 1987.

[12] M. Yu, A. K. Vajargah, S. Z. Miska, R. Majidi, and M. E. Ozbayoglu, "Taking the proper action to gas influx during constant bottom-hole pressure technique of managed pressure drilling," in Offshore Technology Conference, OTC, 5 2013.

[13] W. C. Chin and X. Zhuang, "Parent-child, multilateral well and fracture flow interactions," 1 2020.

[14] C. Winter and A. Edgington, "Computer analysis of gas field deliverability and development economics," Journal of Canadian Petroleum Technology, vol. 3, pp. 180–186, 12 1964.

[15] K. Manikonda, R. Islam, C. E. Obi, A. R. Hasan, A. K. Sleiti, M. W. Abdelrazeq, I. G. Hassan, and M. A. Rahman, "Horizontal two-phase flow regime identification with machine learning classification models," in International Petroleum Technology Conference, p. D011S021R002, IPTC, 2022.

[16] "Contracts for field projects and supporting research on enhanced oil recovery. progress review no. 82, quarterly report, january–march 1995," 6 1996.

[17] I. Silvestrov, E. Hemyari, A. Bakulin, Y. Luo, A. Aldawood, F. Poletto, Y. Liu, Y. Du, A. Egorov, and P. Golikov, "Processing of seismic-while-drilling data from the drillcam system acquired with wireless geophones, top-drive, and downhole vibrations sensors," in SPE Middle East Oil & Gas Show and Conference, SPE, 12 2021.

[18] T. J. Hammons, Geothermal Power Generation: Global Perspectives, Technology, Direct Uses, Plants, Drilling and Sustainability Worldwide. InTech, 6 2011.

[19] W. Sun, Z. Zhao, J. Xiong, Y. Rong, J. Bao, B. Liu, and Z. Chen, "A non-cross-linked <scp> co ₂ </scp> foam fracturing fluid with low friction," The Canadian Journal of Chemical Engineering, vol. 101, pp. 2510–2517, 10 2022.

[20] S. Z. Miska, N. Takach, K. Ashenayi, M. Yu, R. Ahmed, M. B. Pickell, L. Volk, L. Zhou, Z. Chen, A. Washington, and C. Redden, "Advanced cuttings transport study," 9 2003.

[21] "Final report on uncertainties in the detection, measurement, and analysis of selected features pertinent to deep geologic repositories," 7 1978.

[22] D. Luo and F.-S. Li, "Discussion on the sustainable development of environment in the context of shale gas development," DEStech Transactions on Environment, Energy and Earth Science, 6 2017.

[23] H. L. Ayala and E. Morgan, "Natural gas production engineering," Exploration and Production of Petroleum and Natural Gas, pp. 395–428, 9 2016.

[24] S. Blackburn, R. Sanders, C. Boyer, E. Lasseter, J. Stevenson, and R. Mills, "Demonstration of the enrichment of medium quality gas from gob wells through interactive well operating practices. final report, june–december, 1995," 12 1995.

[25] C. H. Atkinson, "Western gas sands project. status report, 1 july-31 july, 1979," 1 1979.

[26] D. S. Qudaihy, O. A. Faraj, F. N. Alnughaimish, A. H. Sunbul, D. E. Hembling, A. Ansari, and B. A. Voll, "Improving horizontal well productivity using novel technology and optimization of drilling fluids," SPE Drilling & Completion, vol. 20, pp. 205–208, 9 2005.

[27] S. Varnado, "Report of the workshop on advanced geothermal drilling and completion systems," 6 1979.

[28] R. Maglione and G. Robotti, "Field rheological parameters improve stand pipe pressure prediction while drilling," in SPE Latin America/Caribbean Petroleum Engineering Conference, SPE, 4 1996.

[29] H. Fitzgibbon, "Gas well automation," Journal of Petroleum Technology, vol. 16, pp. 873–876, 8 1964.

[30] O. Santos, J. Mendes, and P. Ribeiro, "The development and application of a software to assist the drilling engineer during well control operations in deep and ultra deep waters," in SPE Latin American and Caribbean Petroleum Engineering Conference, SPE, 4 2003.

[31] M. Falcone, L. Buss, and U. Fritsching, "Eulerian–lagrangian–eulerian simulations of two-phase minimum quantity lubrication flow in internal drill bit channels," Processes, vol. 10, pp. 600–600, 3 2022.

[32] L. Yang, A. Li, X. Cheng, and Y. Meng, "Under-balance drilling technology of slim directional well in region xiaoguai of karamay oilfield," in International Oil and Gas Conference and Exhibition in China, SPE, 11 2000.

[33] S. Panda, D. Mishra, and B. B. Biswal, "Determination of optimum parameters with multi-performance characteristics in laser drilling—a grey relational analysis approach," The International Journal of Advanced Manufacturing Technology, vol. 54, pp. 957–967, 11 2010.

[34] D. Baggenstos, "Reply on rc2," 12 2022.

[35] C. Obi, Y. Falola, K. Manikonda, A. Hasan, I. Hassan, and M. Rahman, "A machine learning approach for gas kick identification," SPE Drilling & Completion, vol. 38, no. 04, pp. 663–681, 2023.

[36] A. A. Sharkh, A. Fadaq, S. Kikuchi, M. A. Attas, H. E. Housani, E. Morrison, and K. Malash, "The challenge of improving water quality prior to discharge to the sea during well clean up and testing operations by the rig/barge," in Abu Dhabi International Petroleum Exhibition and Conference, SPE, 10 2002.

[37] M. Hohn, D. Patchen, M. Heald, K. Aminian, A. Donaldson, R. Shumaker, and T. Wilson, "Measuring and predicting reservoir heterogeneity in complex deposystems. the fluvial-deltaic big injun sandstone in west virginia. final report, september 20, 1991–october 31, 1993," 5 1994.

[38] F. Crea, A. Mehmood, K. Moyse, S. Saber, F. Toress, and M. A. Arnone, "Use of the near-balance drilling technology enables an operator to drill successfully the deepest exploratory well in pakistan: Case history," in SPE Annual Technical Conference and Exhibition, SPE, 10 2011.

[39] P. Pournazari, P. Ashok, E. van Oort, S. Unrau, and S. Lai, "Enhanced kick detection with low-cost rig sensors through automated pattern recognition and real-time sensor calibration," in SPE Middle East Intelligent Oil and Gas Conference and Exhibition, SPE, 9 2015.

[40] "The success story of optimization gas turbine compressor and generator fuel gas pressure to maximize production in x field," in Proceedings of Indonesian Petroleum Association, 46th Annual Convention & Exhibition, 2022, Indonesian Petroleum Association, 9 2022.

[41] V. K. Y. Kovalenko, Triaxial loading system as a tool for solving geotechnical problems of oil and gas production, pp. 317–326. CRC Press, 8 2012.

[42] M. P. Starrett, A. D. Hill, and K. Sepehrnoori, "A shallow-gas-kick simulator including diverter performance," SPE Drilling Engineering, vol. 5, pp. 79–85, 3 1990.

[43] R. Wemple, "Safety and emergency preparedness considerations for geotechnical field operations," 4 1989.

[44] H. Wu, X. Li, X. Gao, D. Chen, and Z. Li, "Development and application of water sealing technology for gas drainage boreholes.," ACS omega, vol. 7, pp. 733–743, 12 2021.

[45] V. Naik, H. Bahman, S. M. Al-Haddad, and F. Ali, "Production optimization from a short lateral horizontal well in a braided fluvial deposition system: A case study from greater burgan field, kuwait.," in SPE Kuwait Oil & Gas Show and Conference, SPE, 10 2017.

[46] M. Meng, J. Liang, Z. Kuang, J. Ren, Y. He, W. Deng, and Y. Gong, "Distribution characteristics of quaternary channel systems and their controlling factors in the qiongdongnan basin, south china sea," Frontiers in Earth Science, vol. 10, 6 2022.

[47] C. A. Ashburner, "Petroleum and natural gas in new york state," Transactions of the AIME, vol. 16, pp. 906–959, 12 1887.

[48] X. J. Zhang, S. Taoutaou, Y. Guo, Y. L. An, S. M. Zhong, and Y. Wang, "Engineering cementing solution for hutubi underground gas storage project," in SPE Unconventional Gas Conference and Exhibition, SPE, 1 2013.

[49] J. W. Tester, G. Morris, R. Cummings, and R. Bivins, "Electricity from hot dry rock geothermal energy: technical and economic issues," 1 1979.

[50] A. I. Bashirov, I. R. Galas, I. A. , and M. F. Nazyrov, "Controlled deep perforation by radial drilling of channels with the "perfobur" technical system to intensify the reservoir inflow," Kazakhstan journal for oil & gas industry, vol. 3, pp. 99–110, 6 2021.

[51] S. Bertram-Howery and P. Swift, “Potential for long-term isolation by the waste isolation pilot plant disposal system,” 6 1990.

[52] L. Chen and J. Lian, “Discussion on the development of intelligent drilling technology and equipment for gas drainage,” Journal of Electronic Research and Application, vol. 3, 7 2019.

[53] D. Wootan, R. Bolden, A. Bridges, N. Cannon, S. Chastain, B. Hey, R. Knight, C. Linschooten, A. Pitner, and B. Webb, “Summary report on the design of the retained gas sampler system (retained gas sampler, extruder and extractor),” 9 1994.

[54] C. Pang, Y. Shi, X. Wang, B. Li, and H. Luan, “Study on a high-efficiency mining technology system for gas outburst in coal seams—example of an h coal mine,” Minerals, vol. 12, pp. 795–795, 6 2022.

[55] M. A. Sultan, M. Ahmed, M. Ali, and G. Z. Babar, “Understanding the well control procedures for optimizing the well control system during drilling,” American Journal of Computing and Engineering, vol. 5, pp. 24–38, 11 2022.

[56] “Preliminary site characterization report, rulsion site, colorado,” 8 1996.

[57] K. Ogra, Y. Chandra, A. Pandey, V. Verma, A. Kumar, and R. K. Sinha, “Evolution of production logging technologies and capabilities: In quest to know the unknown, a brownfield case study, mumbai high, ongc western offshore,” in SPE Oil and Gas India Conference and Exhibition, SPE, 3 2012.

[58] K. Schmigel and S. D. Code, “Environmentally sensitive well successfully drilled underbalanced using recycling technology,” in IADC/SPE Managed Pressure Drilling & Underbalanced Operations, SPE, 3 2007.

[59] R. Akhmiyeva, “Influence of the oil industry on anthropogenic changes in natural landscapes,” SHS Web of Conferences, vol. 128, pp. 3003–03003, 12 2021.

[60] D. S. Qudaihy, F. Nughaimish, A. H. Sunbul, A. Ansari, D. E. Hembling, O. A. Faraj, and B. Voll, “New-technology application to extend the life of horizontal wells by creating uniform-flow-profiles: Production completion system: Case study,” in SPE/IADC Middle East Drilling Technology Conference and Exhibition, SPE, 10 2003.

[61] Q. Sun, Y. Liu, Z. Shen, Y. Han, Q. Wang, and P. Liu, “A combined thermal spallation and melting technology by plasma jet for deep and hard rock reservoirs,” SPE Journal, vol. 28, pp. 49–63, 8 2022.

[62] A. Hosseini, M. Najafi, and A. Morshedy, “Determination of suitable distance between methane drainage stations in tabas mechanized coal mine (iran) based on theoretical calculations and field investigation,” Journal of Mining Institute, vol. 258, pp. 1050–1060, 12 2022.