

The Next Cost Wave Is a Simulation Problem

What an Agent-Based Model Reveals About Offshore Well Decommissioning

Why P&A campaigns blow their budgets — and why the fix is batching, not bigger contingency

Ilan Gleiser · Erica Crandon

August 2026

AlphaIQ — See the process. Change the outcome.

-69% P50 well-work cost, rigless vs rig	127 P50 campaign days saved	12/12 Validation checks passed	-74% Fleet emissions, rigless vs rig
---	---	--	--

1. The Problem: Gantt Charts Cannot See Feedback

Decommissioning is the bill at the end of the party. More than a hundred thousand offshore wells worldwide need permanent plug and abandonment (P&A) over the next three decades, and in mature basins the liability already rivals new-development budgets. Unlike drilling, P&A produces no revenue to absorb an overrun: every day a vessel waits on weather, every failed barrier test, every stuck tool lands directly on a provision that auditors and regulators read line by line.

The standard tools were built for a different problem. Campaigns are still planned on deterministic schedules and factored cost lines — durations treated as known, weather as a fixed allowance, crews as interchangeable. That machinery works when tasks are independent. It fails structurally when outcomes are dominated by *feedback*: a campaign that slips samples more storms, which slips it further; a crew that repeats one operation gets faster, but only if the plan lets it repeat.

Two couplings do most of the damage. Weather risk compounds with duration — waiting time is bought at the vessel's day rate, so the marginal cost of slip is convex. And crew learning compounds with batching — the well-documented geometric learning curve only accrues across an unbroken sequence of identical operations, which means the campaign *strategy* decides which learning curves exist at all. A static plan carries neither effect. Our simulations put their combined worth at roughly two-thirds of campaign cost and nearly half the schedule.

So we built the campaign as an agent-based model. In *welldecom*, wells, vessels, an operator and a regulator are software agents following their own rules in a shared environment with six-hour weather resolution. Cost and schedule are not inputs — they emerge. This post walks through the model, its validation, and what it says about the single biggest decision in P&A: how to structure the campaign.

2. Six Terms That Carry the Story

Plug and abandonment (P&A). The permanent sealing of a wellbore at end of life: cement barriers across the reservoir and annuli, wellhead and conductor removed, seabed cleared. Regulators typically require completion within a fixed window after production ceases.

AB1 / AB2 / AB3. The industry's abandonment state machine: reservoir permanently isolated (AB1); all barriers installed and verified (AB2); wellhead and conductor removed, site cleared, decommissioning complete (AB3). In the model these states are earned by verified work, never scheduled.

PWC versus section milling. Two technologies for the annular barrier. Section milling grinds out casing from a heavy rig — an 8-to-20-day operation with hot risks. Perforate-wash-cement (PWC) sets the same barrier in a single 2-to-5-day trip from a light intervention vessel, removing the rig from the critical path.

Spread rate, WOW, NPT. The spread rate is the all-in daily cost of a vessel and its services (\$180k/day for a light vessel, \$350k/day for a semisubmersible here). WOW is waiting on weather; NPT is non-productive time from mechanical or downhole events. Both are paid at (or near) the full spread rate.

Learning rate (L). The Crawford curve parameter: each doubling of repetitions multiplies task duration by L. A standardised PWC tool set runs near $L = 0.85$; conservative milling near 0.95. Lower is faster learning — and the difference compounds.

P10 / P50 / P90. Percentiles of a Monte Carlo ensemble. P50 is the planning base case; P90 anchors contingency; the P10–P90 band is the honest width of the answer — provided you say what the band varies (more on that in Section 5).

Term	One-line definition
P&A	Permanent sealing and removal at end of well life
AB1 → AB2 → AB3	Reservoir isolated → barriers verified → conductor out, site cleared
PWC	Single-trip rigless annular barrier (2–5 days)
Section milling	Rig-based casing removal for the same barrier (8–20 days)
Spread rate / WOW / NPT	Daily vessel cost; weather waiting; non-productive time
Crawford learning rate L	Duration multiplier per doubling of repetitions (0.85–0.95)

3. Methodology: Four Stochastic Engines, One Clock

Durations are three-point estimates, not points. Every phase on every well draws from a PERT distribution over (optimistic, most-likely, pessimistic) days — mean $(a + 4m + b)/6$ — scaled by measured depth. The milling estimate (8, 10, 20) is not just longer than PWC's (2, 3, 5); it is more right-skewed, and the skew is where budgets die.

Crews learn geometrically, per activity. The N -th repetition of an activity by the same crew takes $T_1 \cdot N^b$, with $b = \ln L / \ln 2$. The counter is tracked per crew *and per activity family*: barrier-setting proficiency does not transfer to milling. This bookkeeping is why batching wins — an unbroken thirty-plug sequence rides one curve all the way down, while a sequential campaign restarts every curve at every phase change.

Risk arrives as discrete events, priced in days and dollars. Stuck tubing, lost circulation, swarf blockages, cut misruns: Bernoulli trials at task start, each adding a triangular duration penalty and a tooling cost at the spread rate. Phase-0 diagnostics — logging the well before intervention — cut downhole event probabilities six-fold, from 0.30 to 0.05.

Weather is a process, not an allowance. Significant wave height follows a seasonal AR(1) process in log space at six-hour resolution, peaking mid-January. Transit, station-keeping and heavy lifts each have their own sea-state ceiling per vessel, and lifts need a clean six-hour window to start. Autocorrelation produces the multi-day storms and multi-day windows that a flat allowance cannot represent. One design detail we insist on: the weather field is drawn *before* any task draw, so two strategies compared on the same seed face identical seasons — a controlled experiment, not an accident of the random stream.

How Small Shocks Become Big Outturns — Four Coupling Pathways

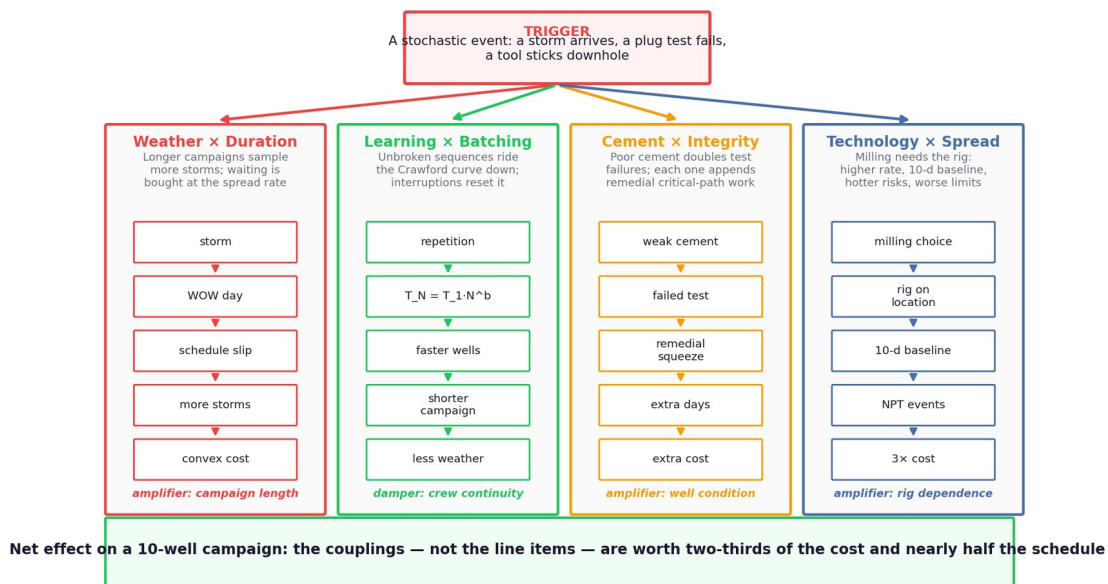


Figure 1: The four coupling pathways. Three amplify shocks; one — learning under batching — damps them. Campaign design is the choice of which pathways you feed.

#	Mechanism	What it does in the model
1	PERT durations	Right-skewed task times, depth-scaled, per phase per well
2	Crawford learning	$T_1 \cdot N^b$ per crew per activity; batching preserves the sequence
3	Bernoulli risk events	Discrete NPT with triangular penalties and tooling costs
4	Barrier rulebook	30 CFR §250.1715 if-then plug placement; integrity tests pass/fail
5	AR(1) metocean	Seasonal 6-hour Hs process gating transit, DP and lifts
6	Cost coupling	Every day and event priced at the spread rate into an 11-line WBS

4. Architecture: Four Agents, One Discrete-Event Engine

The pipeline runs from well schematics to percentiles. Well stock and the regulatory plug program compile into a campaign plan; the metocean field is drawn; the discrete-event engine executes the plan against the weather, the learning curves and the risk draws; and a Monte Carlo layer repeats the whole campaign 10,000 times from recorded seeds to produce P10/P50/P90 with scope-labelled bands.

welldecom — Simulation Pipeline

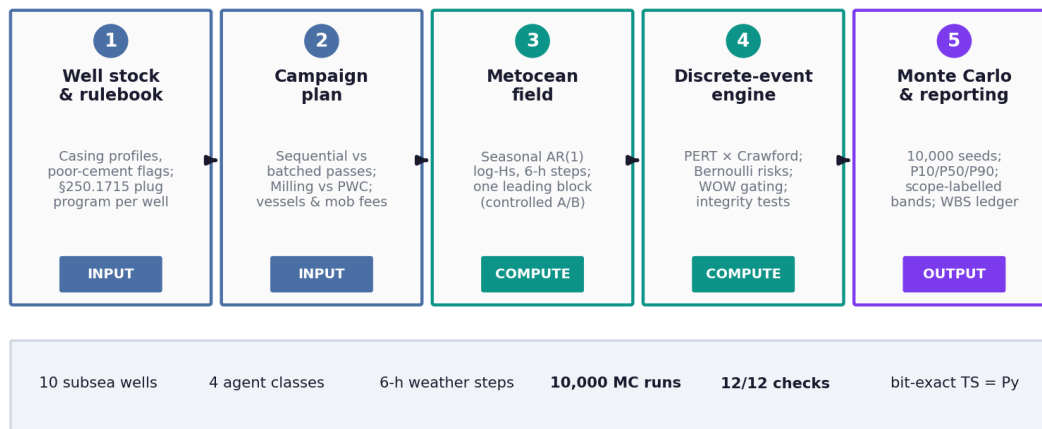


Figure 2: The simulation pipeline. Every run is reproducible from one recorded master seed; the TypeScript and Python implementations agree bit-for-bit.

Four agent classes interact through five channels. Wells hold state machines and cement quality; vessels hold spread rates, weather limits and fuel burn; the operator holds strategy, technology and routing; the regulator holds deadlines, audits and the carbon price. Barrier verification is the channel where geology bites back: every plug faces a 15,000-lb / 1,000-psi integrity test, and failures append remedial cement to the critical path — twice as often on wells with poor primary cement.

Agent Network — Four Classes, Five Interaction Channels

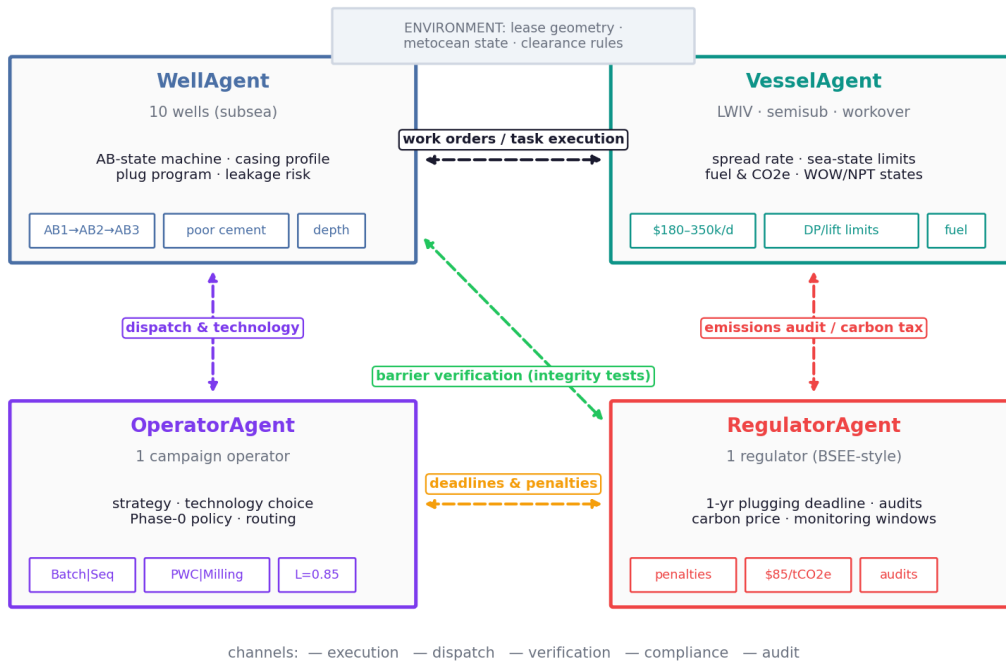


Figure 3: Agent classes and interaction channels. Nothing in the model reads a schedule; the schedule is what the interactions produce.

Component	Count / type	Role
WellAgent	10 subsea wells	AB-state machine, casing profile, plug program, leakage risk
VesselAgent	LWIV, semisub, workover	Execution units with weather limits, spread rates, emissions
OperatorAgent	1	Strategy, technology, Phase-0 policy, nearest-well routing
RegulatorAgent	1 (BSEE-style)	Plugging deadline, penalties, emissions audit, carbon price
Engine	6-h discrete-event	Weather-gated task execution, WOW/NPT accounting
Monte Carlo	10,000 iterations	Spawned seed streams, percentiles, scope-labelled bands

5. What the Model Found

5.1 Rigless batching cuts the bill by two-thirds

The headline: across 10,000 simulated campaigns of ten subsea wells, the batched rigless PWC programme lands at a P50 well-work cost of \$31.3M against \$102.5M for sequential rig-based milling — a 69% reduction — and finishes in 155 days instead of 282. The distributions barely touch: the rigless P90 sits far below the rig campaign's P10. This is not one lucky draw; it is the whole ensemble.

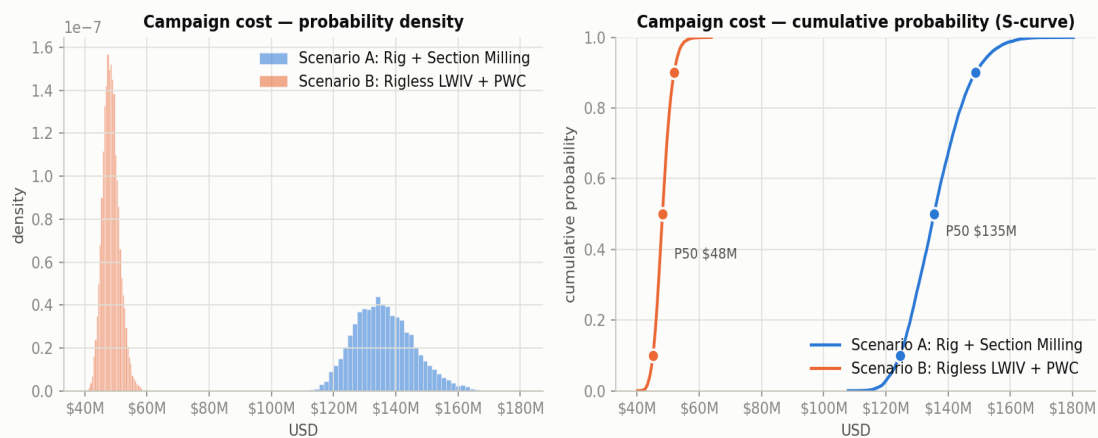


Figure 4: Full-scope campaign cost, 10,000 iterations per scenario — probability density (left) and cumulative S-curves with P10/P50/P90 markers (right).

5.2 The batched campaign looks stalled when it is winning

A finding for the compliance office, not the spreadsheet: under batching, no well reaches final abandonment for months — reservoir barriers go in across the whole stock, then Phase 2, then conductors — and then all ten wells complete within weeks. An interim audit reading “zero wells finished” would misread an on-schedule campaign as a failing one. Brief the regulator on the completion-curve shape before the first progress report, not after.

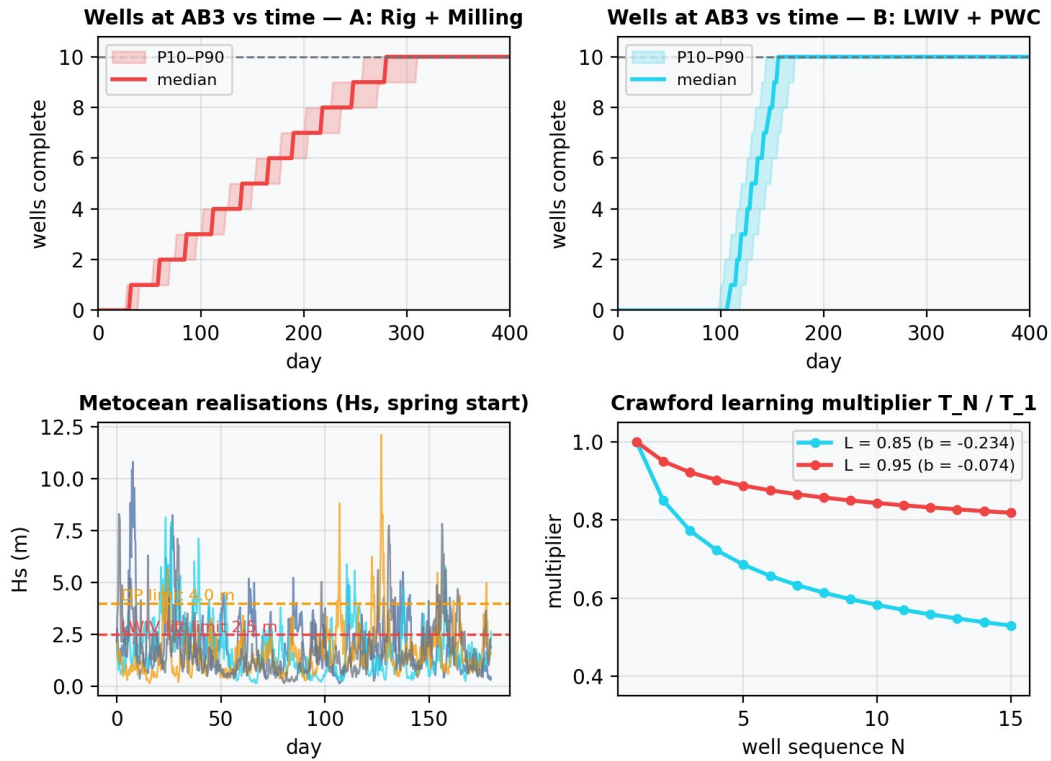


Figure 5: Campaign progress ensembles (200 realisations). Sequential execution climbs steadily; batching plateaus then sprints. Below: the metocean realisations and learning curves driving the dispersion.

5.3 The start date is worth more than most negotiations

Timing is a first-order lever. An identical work scope started in winter waits 53 days on weather at P50, against 25 for a summer start — and every one of those days is bought at the spread rate. Phase-0 diagnostics are the other cheap win: P50 non-productive time nearly halves, from 11.5 to 6.3 days, paying for the logging run on the first avoided stuck string.

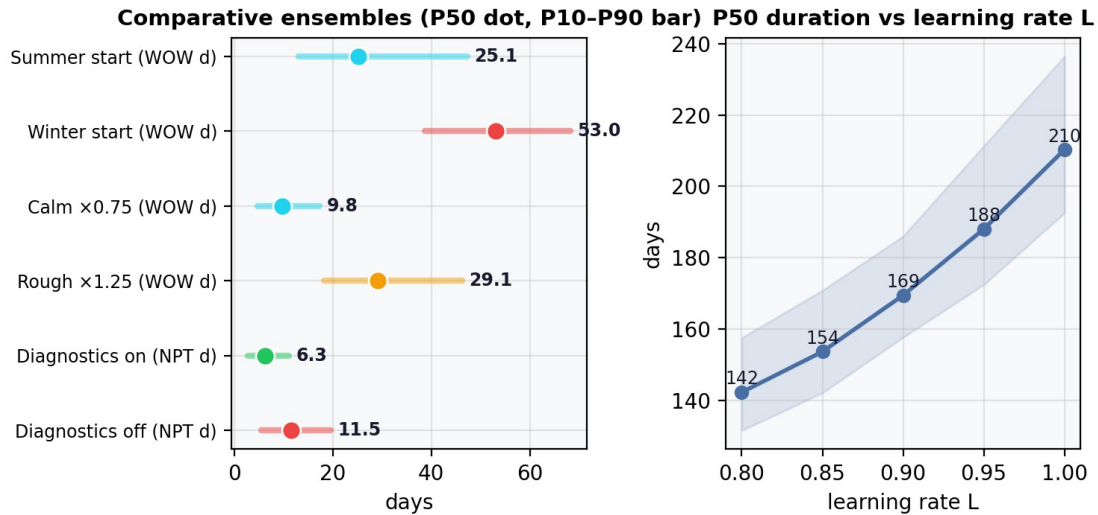


Figure 6: Comparative ensembles (P50 dot, P10–P90 bar) for seasonality, storm scale and diagnostics; and P50 duration versus crew learning rate — monotonic, steep, and the argument for standardised tooling.

5.4 Weather risk compounds — the model shows it, a Gantt chart cannot

The rig campaign does not just cost more per day; it is unluckier per campaign. Being on location 82% longer, it samples more storms: 43 P50 WOW days against 20 for the rigless programme, with a heavier right tail. Weather risk multiplies with duration rather than adding to it — which is why the model's duration advantage for rigless execution (45%) exceeds what static comparisons publish (42%), and why the difference is a feature of reality, not an error of the model.

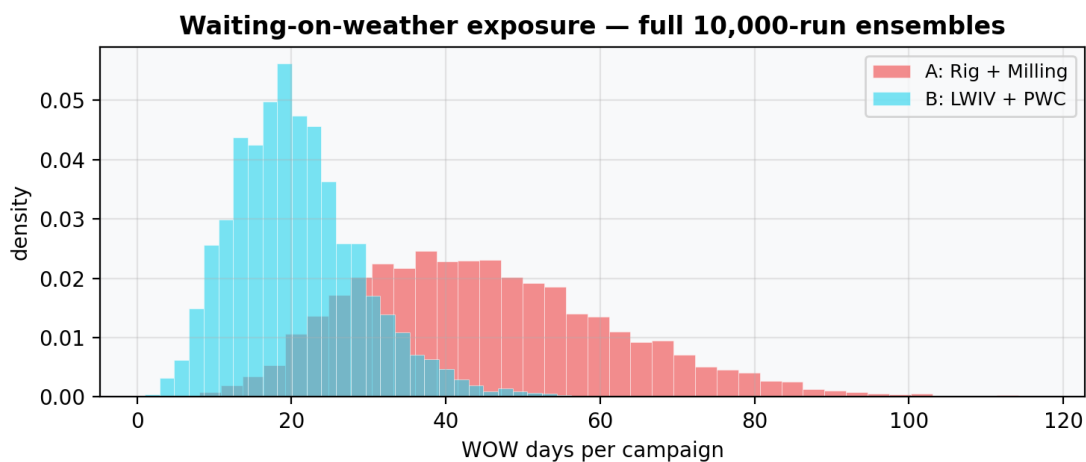


Figure 7: Waiting-on-weather days per campaign across the full ensembles. The rig campaign's distribution is both shifted and skewed — longer exposure samples more storms.

5.5 Validation: twelve checks, and a band that says what it varies

Trust is earned by protocol, not assertion. The model passes five verification checks (analytic limits reproduced to machine precision; bit-exact agreement between

independent Python and TypeScript implementations), six behavioural acceptance checks on independent ensembles, and a benchmark against the published strategic comparison: 3.2% MAPE on matched cost scope, inside a 5% gate. And one honesty rule we now apply everywhere: a seed-only Monte Carlo band prices execution luck, not input error. Stacking plausible input ranges — day rates $\pm 15\%$, learning rate ± 0.04 , carbon price \$50–130/t — widens the P10–P90 cost band by 2.1 $\times$. Every band we publish is labelled with its scope; the stacked one is what belongs in a budget.

Validation group	Example	Result
Verification (5 checks)	MC mean vs analytic PERT sum	PASS · 0.004% err
Cross-implementation	TypeScript vs Python, 48 fields	bit-exact
Behavioural acceptance (6)	Winter WOW $\gg$ summer WOW	PASS
Benchmark (matched scope)	MAPE vs published table	3.2% < 5% gate
Uncertainty honesty	Seeds-only vs stacked band	2.1$\times$ wider

6. What To Do With This

The couplings are the budget. In the reference campaign, the interactions a deterministic plan cannot see — weather with duration, learning with batching, cement quality with integrity testing — are worth two-thirds of the cost and nearly half the schedule. Decommissioning planning that ignores them is not conservative; it is wrong in a known direction.

Three takeaways for decision-makers:

1. Choose strategy before contingency. Batching by activity and moving Phase 2 off the rig changed the P50 by 69% in our ensembles; no realistic contingency percentage does that. Structure first, buffer second.

2. Buy information early. Phase-0 diagnostics and a deliberately chosen start date are the two cheapest interventions in the model — one halves NPT, the other halves weather waiting — and both are decisions made before a single vessel sails.

3. Demand bands that state their scope. Ask of any P10–P90: what does it vary? A seed-only band prices luck on fixed assumptions and reads 2× too confident. Ask for the stacked band, the recorded seed, and the tail from the big ensemble.

What's next: multi-vessel parallelism within passes, port and supply-chain constraints, basin calibration on hindcast metocean data and audited campaign outturns — and the same MAPE gate applied to real history. The model is built twice (Python and TypeScript), validated to 12/12, and reproducible from a single seed; the interactive dashboard runs the full engine in the browser with a fan-chart ensemble mode, in English and Portuguese.

Explore the live simulation — seven agent-based engines, ensemble mode, and the full validation suite — on the AlphaQ Pré-Sal ABM platform.

About the Authors

Ilan Gleiser and **Erica Crandon** build agent-based simulation and decision platforms at AlphaQ, where the welldecom model and the Pré-Sal ABM platform were developed. This post presents simulation results on stylised data; it is not investment, provisioning, or regulatory advice.