Geminus

TIER 1 U.S. REFINER

How a Tier 1 U.S. Refiner Captured 1–2% More Revenue Per FCC While Modernizing RTO

Physics-informed AI surrogates replaced 90+ minute solver cycles with sub-5-second optimization. Deployment time dropped 10X, transforming the economics of real-time optimization across a multi-site fleet.

1-2% REVENUE INCREASE PER UNIT	**>5 sec** PER OPTIMIZATION RUN (VS. 90+ MIN)	**8-12 weeks** KICKOFF TO VALIDATED SURROGATE
10X FASTER DEPLOYMENT VS. TRADITIONAL RTO	**95%+** ACCURACY VS. RIGOROUS SIMULATOR	**Days-Weeks** RETRAIN AFTER UNIT REVAMP (VS. MONTHS)

THE CHALLENGE

RTO That Can't Keep Up With the Refinery It Runs

This Tier 1 U.S. refiner operates some of the most economically sensitive FCC units in the country. Tiny shifts in operating conditions change the mix of fuels and chemicals the units produce, and the difference between optimal and conservative operation is worth millions per unit per year.

The refiner's existing RTO infrastructure simply couldn't keep up. A single rigorous optimization run took 90+ minutes. Every feed change, catalyst swap, or constraint update triggered weeks of SME work on the underlying models. Operators didn't trust the simplified approximations embedded in the RTO layer, so the unit ran extremely conservatively, **leaving yield and margin on the table.**

- → Detailed FCC kinetic models too computationally heavy for real-time use (90+ min per run)
- → Simplified approximations diverged from what operators saw in the plant, eroding trust
- → Changes in conditions triggered weeks of SME effort
- → Purely data-driven (RL) alternatives need 6+ months of historical process data plus full retrains after any unit change, making them impractical for an evolving fleet

Geminus

TIER 1 U.S. REFINER

THE GEMINUS APPROACH

Physics-Informed AI: Simulator Fidelity at Operating Speed

Geminus doesn't replace the kinetic simulator the engineering team already trusts. Instead, Geminus operationalizes it. By creating physics-informed surrogates directly on the refiner's existing FCC kinetic model, Geminus preserves full reaction and thermodynamic fidelity while executing in milliseconds.

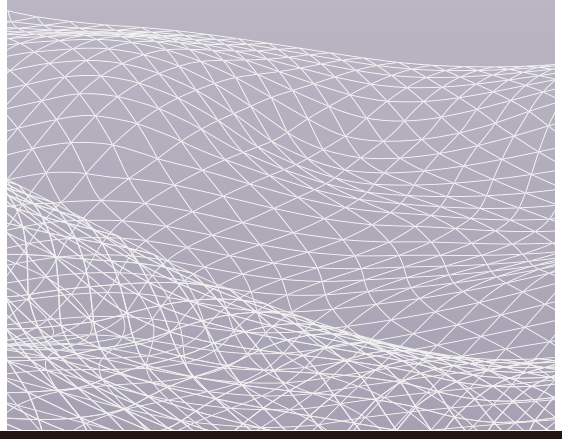
Three hard problems required original solutions:

HIGH-FIDELITY KINETICS AT MILLISECOND SPEED: Surrogates capture the full kinetic envelope of the FCC reactor across feed composition, reactor severity, catalyst-to-oil ratio, steam, and coke balance. Predictions land within 95%+ accuracy of the rigorous simulator.

OPTIMIZER-READY BY DESIGN: Smooth gradients, stable convergence, architecturally enforced physical consistency and predictable extrapolation in unseen conditions were tested explicitly so the surrogate behaves correctly inside RTO loops. The result is recommendations operators can trust and act on.

MODEL SUSTAINMENT IN DAYS, NOT MONTHS: When the simulator changes after a revamp, catalyst change, or new constraint, Geminus retrains the surrogate in hours, with with RTO updated and validated in days or weeks. Traditional first-principles RTO rebuilds take months. Purely data-driven approaches require 6+ months of new process data before training can even start.

> "This wasn't possible before. Not with the tools we had, not with the people we had. Now it's running every day, and we're more profitable for it."
>
> **— PROCESS ENGINEER, TIER 1 U.S. REFINER**



Geminus

TIER 1 U.S. REFINER

WHY THIS IS DIFFERENT

Physics-Informed, Not Data-Hungry

Unlike reinforcement learning or purely data-driven optimization, Geminus surrogates inherit physics directly from your validated simulator. No multi-month data collection. No retraining from scratch after every catalyst or feed change. The chemistry, thermodynamics, and constraints your engineers already trust are preserved by design.

Real-Time, Optimizer-Ready.

Query the surrogate from your existing RTO infrastructure, or run it with the Geminus optimizer at roughly 4–5 seconds per scenario on a single core, with full Jacobian computations included. Surrogates are explicitly designed for stable behavior inside refinery RTO loops, with smooth gradients, feasibility-aware classification, and validation against both constrained and unconstrained baselines.

Robust Beyond the Historical Envelope.

Since surrogates inherit physics from the simulator, they perform reliably in the regions outside the historical operating envelope where margin opportunities and constraint violations tend to live. Whereas purely data-driven RL approaches degrade in those regions and require fresh data collection plus full retrains to recover, Geminus models hold their accuracy and keep running.

SCALE & EXPANSION

Already Operating. Not a Pilot Story.

MULTI-SITE FCC ROLLOUT

Closed-loop optimization is deployed across multiple FCC units at one of the largest U.S. integrated refiners. Surrogates were benchmarked against the kinetic simulator and validated under both constrained and unconstrained conditions before scaling.

FROM ADVISORY TO CLOSED-LOOP

Surrogates integrate into the refiner's existing optimization platforms (platform-agnostic). Operator-validated recommendations support feed/catalyst variability, constraint management, and severity decisions in real time.

BROADER REFINERY OPTIMIZATION.

The same Physics-informed AI framework extends to any refinery unit with a validated process simulator. Hydroprocessing, distillation, gas treating, and utilities all qualify wherever a fidelity-speed gap is blocking real-time decisions.

Identify your unit. See the impact in 90 days.

Complimentary simulator review included. Zero Ops disruption. Little to no training required. Start Your Pilot [link to contact]