rl core

Beyond the Chatbot:

Defining Industrial Agency in the Water Sector

Andrew Patterson, March 2026

The current discourse around "Agentic AI" in the water sector is increasingly dominated by the capabilities of Large Language Models (LLMs). We see narratives of agents that can query GIS maps, check crew availability, and draft work orders. While these linguistic models are proficient at narrative tasks and information retrieval, a fundamental architectural gap exists when these same models are proposed for real-time industrial control.

This gap is often obscured by a common misconception—the tendency to assume that because a model can describe a process fluently, it can manage the process reliably. In reality, managing a chemical dosing valve or a setpoint trajectory is not a linguistic problem; it is a numerical optimization problem. To improve stability and efficiency in critical infrastructure, we must distinguish between the "Chatbot" that describes the plant and the "Control Agent" that operates within its physics.

Architectural Limits: Distinguishing Fluency from Control

The core of the issue lies in the mismatch between architecture and objective. An LLM is a probabilistic engine designed for "next token prediction." It collapses a massive probability space to generate the most statistically likely sequence of text.

When a linguistic model is tasked with control, it is essentially attempting to "talk" a valve into a position based on statistical patterns found in its training data. This is an exceptional tool for summarizing a standard operating procedure (SOP), but it is a dangerous architecture for managing a high-stakes control loop. On the other hand, a control policy does not predict text; it optimizes for specified performance rewards.

Feature	Large Language Models (LLMs)	Control Agents (RL)
Primary Logic	Probabilistic Next Word Prediction	Performance Reward Optimization
Performance Verification	External / Heuristic-based	Internal / Objective-based
Data Input	Unstructured Text / RAG	High-Frequency Time-Series (Telemetry/ Sensor Data)
Output Type	Natural Language / Code	Continuous Numerical Setpoints
Grounding	Linguistic Patterns	Physical Dynamics & Rewards
Constraint Handling	Soft Prompting (Unreliable), Expert Systems	Hard Action Limits & Penalized Constraints

Framing Reinforcement Learning (RL) as a modern evolution of control theory helps clarify this distinction. Where traditional PID control handles the "fast-twitch" regulation of a valve, and Model Predictive Control (MPC) attempts to follow a rigid physics model, RL directly optimizes setpoints by adapting to current process dynamics. Expecting a linguistic model to manage these intricacies is analogous to asking a poet to tune a PID loop: the vocabulary may be correct, but the underlying logic is not grounded in the continuous numerical reality of the plant.

Process Resilience: Adapting to Environmental Drift

In complex industrial environments like wastewater treatment, a major challenge to automation is "drift." Traditional control strategies, particularly Model Predictive Control (MPC), rely on a rigid physics model of the facility. While effective at commissioning, these models inevitably degrade as equipment wears, sensors drift, and environmental conditions—such as seasonal influent temperature or chemical potency—shift.

Maintaining an MPC controller requires periodic, high-cost model refreshes. In the interim, the system becomes a liability, operating on outdated assumptions that no longer match the physical reality of the plant.

True industrial agency requires process resilience—the ability to treat process variability not as an error but as new data. An RL policy continuously recalibrates its understanding of the environment, "drifting" alongside the plant to ensure the control policy remains grounded in the current state of the hardware.

Notably, this approach does not require extensive historical datasets to begin delivering value. Because RL encapsulates both the optimization routine and the data acquisition process—the "explore/exploit" cycle—the system can be deployed on assets with limited data history. By setting tight initial guardrails, operators can allow the agent to gather data and identify optimal strategies safely and autonomously from day one.

Asset Specialization: Autonomous Policy Discovery

A recurring theme in the "Agentic AI" narrative is the pursuit of a generalist model—a single entity capable of reasoning across an entire plant. In critical infrastructure, however, generalization is often a vulnerability. In high-stakes environments, "generalization" often results in the prioritization of "average" fleet behavior over the localized physics of a single asset.

A model that generalizes across multiple plants often lacks the resolution to distinguish between nominal operation and a unit-specific anomaly. Furthermore, the increased complexity of a general-purpose architecture creates a larger attack surface and makes the system significantly harder to verify. True reliability is found in asset specialization, where a policy is explicitly optimized for the unique dynamics, mechanical wear, and sensor noise of the exact hardware it manages.

Online Reinforcement Learning enables the autonomous discovery of these unit-specific control policies. While the underlying software package is general, the resulting control behaviors are highly specialized. When deployed across multiple assets—such as a fleet of four scrubber units—the system does not execute identical logic across the fleet. Instead, it identifies the optimal operating policy for each unit individually. Each controller adapts to the specific mechanical wear, sensor drift, and chemical potencies of the asset it manages.

This discovery process is driven by the operator's definition of Key Performance Indicators (KPIs). The operator specifies the desired outcomes—such as chemical cost reduction and efficiency targets—and the online RL agent identifies the control strategy required to achieve them in that specific physical context. This eliminates the traditional overhead of Model Predictive Control (MPC), where each unit requires independent modeling and tuning. Unlike linguistic models that require hand-crafted context to simulate specialized knowledge, an online RL agent builds its policy directly from the telemetry of the asset it supervises. The result is a system where every unit in the facility is managed by its own dedicated controller, optimized for its unique physical state.

Operational Oversight: Constraint-Based Governance

Industrial facilities are frequently governed by "set-and-forget" logic. While a human operator possesses the expertise to optimize a process, the cognitive load required to manage persistent setpoint adjustments across dozens of units is prohibitive. Supervisory automation via RL acts as a capacity multiplier, handling the hourly adjustments necessary to track shifting influent conditions.

This is not a replacement for human judgment, but a shift in the nature of oversight toward operational governance. In this model, the operator maintains absolute control through **hard setpoint guardrails**, including:

- **Absolute Action Ranges:** Hard limits for setpoints that the agent cannot physically violate (e.g., pump speed restricted between 50-100 rpm).
- **Delta/Ramp Limits:** Restricting the rate of change for control actions (e.g., setpoints can only shift by ±5% per hour).
- **Temporal Controls:** Limiting the frequency of adjustments (e.g., setpoint changes occurring only once every 15 minutes).

While setpoints are governed by these hard mathematical bounds, downstream process variables (such as H2S outlet concentration) are managed as **penalized constraints**. Because physics and environmental upsets can cause process variables to shift regardless of the control action, the agent is trained to prioritize these targets through heavy penalties in the optimization routine. This ensures the system remains focused on regulatory compliance even in the face of unpredictable plant dynamics.

The transition from generative AI to agentic control in the water sector requires a fundamental shift in how we value "intelligence." Linguistic models offer a powerful interface for reporting, but they lack the numerical grounding required for physical process control. True industrial agency is achieved through specialized, adaptive policies that prioritize performance reward optimization over linguistic prediction—equipping operators with a relentless supervisory capacity that ensures every unit in the facility is operating at its physical limit, every hour of every day.