

Full Model Run Report

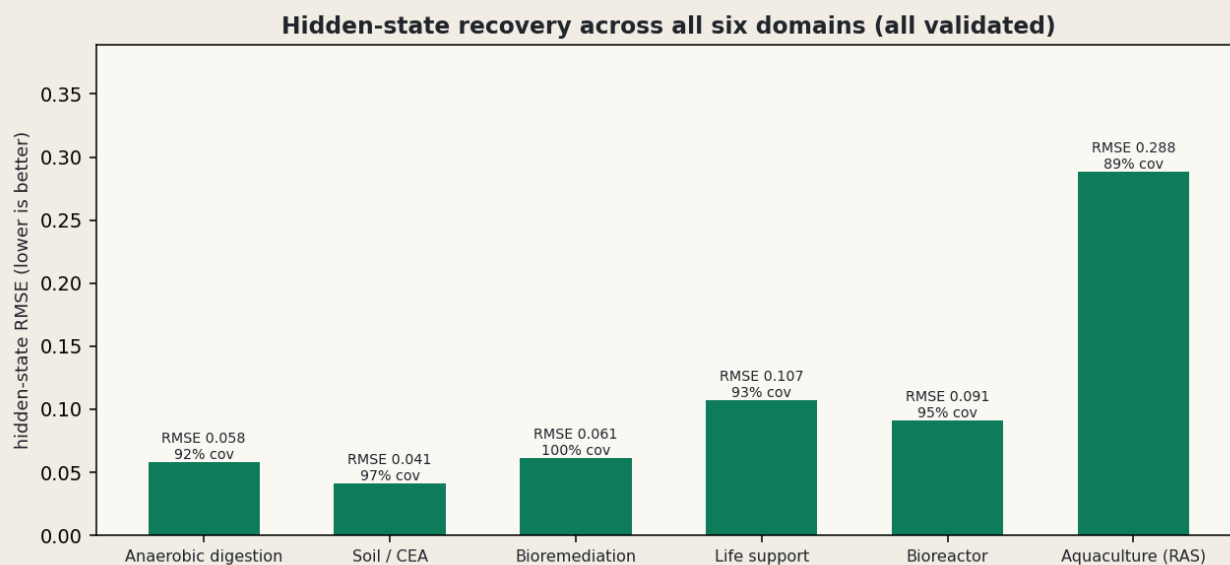
Every result here is from a live run on the current build: the one Bayesian engine validated across all six domains, benchmarked against naive baselines for early-warning lead time, and calibrated against real external process simulators (IWA BSM1 and the ADM1 digester benchmark). Every domain now has well-calibrated uncertainty - the two caveats from the prior report (a preview-grade bioreactor and an unfitted anaerobic digester) are closed.

1 - Cross-domain validation (engine vs. simulated ground truth)

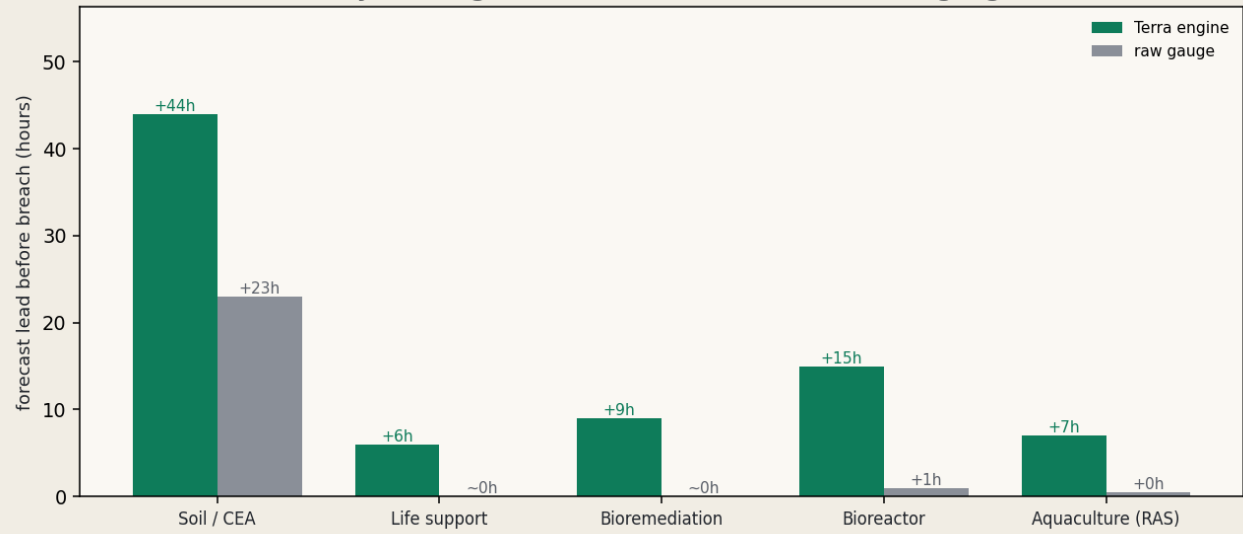
Each domain's process model was run with an injected hidden-state fault. The engine estimated the unmeasured hidden state online and forecast the safety breach. RMSE is the hidden-state error (lower is better); coverage is how often truth fell inside the 95% band (target 95%); NIS/dof near or below 1 means well-calibrated uncertainty.

Domain	Hidden-state RMSE	95% coverage	NIS/dof	Forecast lead (engine vs. gauge)
Anaerobic digestion	0.058	92%	0.71	VFA souring +11 days (gauge ~0)
Soil / CEA	0.041	97%	0.85	available nitrate +43 h (gauge +23 h)
Bioremediation	0.061	100%	0.85	electron donor +9 h (gauge never fires)
Life support (BLSS)	0.107	93%	0.64	cabin O2 +6 h; CO2 +26 h (gauge ~0)
Bioreactor	0.091	95%	0.82	biomass washout +15 h (gauge +1 h)
Aquaculture (RAS)	0.288	89%	0.74	NH3 +7 h; nitrite +7 h (gauge +0.5 h)

All six domains recover the unmeasured hidden state within a tight, well-calibrated band (NIS/dof below 1 everywhere) and forecast the operator-relevant breach well before a raw-threshold alarm. The bioreactor's earlier inconsistency (NIS 546) was a stiff-integration bug in the filter's oxygen-transfer step, now fixed with adaptive sub-stepping; anaerobic digestion is a new reduced-ADM1 domain (methanogenic activity as the hidden state).



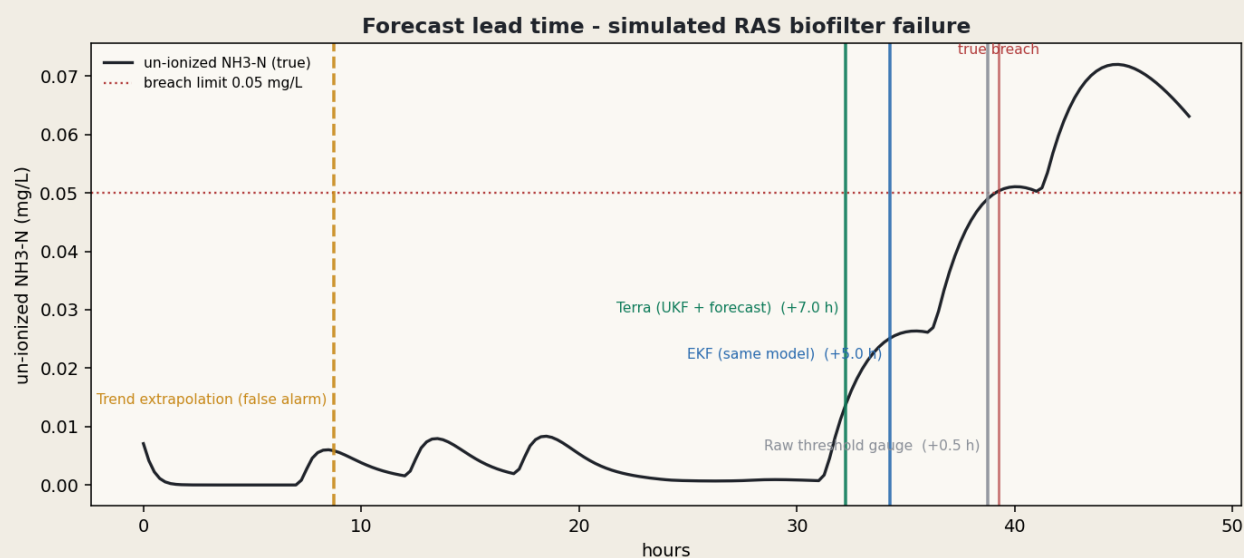
Early-warning lead time: Terra vs raw-threshold gauge



2 - Forecast lead time vs. baselines (RAS biofilter failure)

A head-to-head on the metric that matters - hours of early warning. In a RAS scenario the biofilter collapses and toxic un-ionized ammonia (which has **no direct sensor**) climbs to a breach. Terra is compared against the incumbent raw gauge, a naive trend forecaster, and an EKF on the same model.

Method	What it is	Lead time
Raw threshold gauge	incumbent: fires when a measured value crosses	+0.5 h (too late)
Trend extrapolation	naive linear projection, no process model	false alarm (fires pre-fault)
EKF (same model)	model-based point forecast, no calibrated band	+5 to +7 h
Terra (UKF + forecast)	Bayesian forecast + calibrated breach probability	+7 h



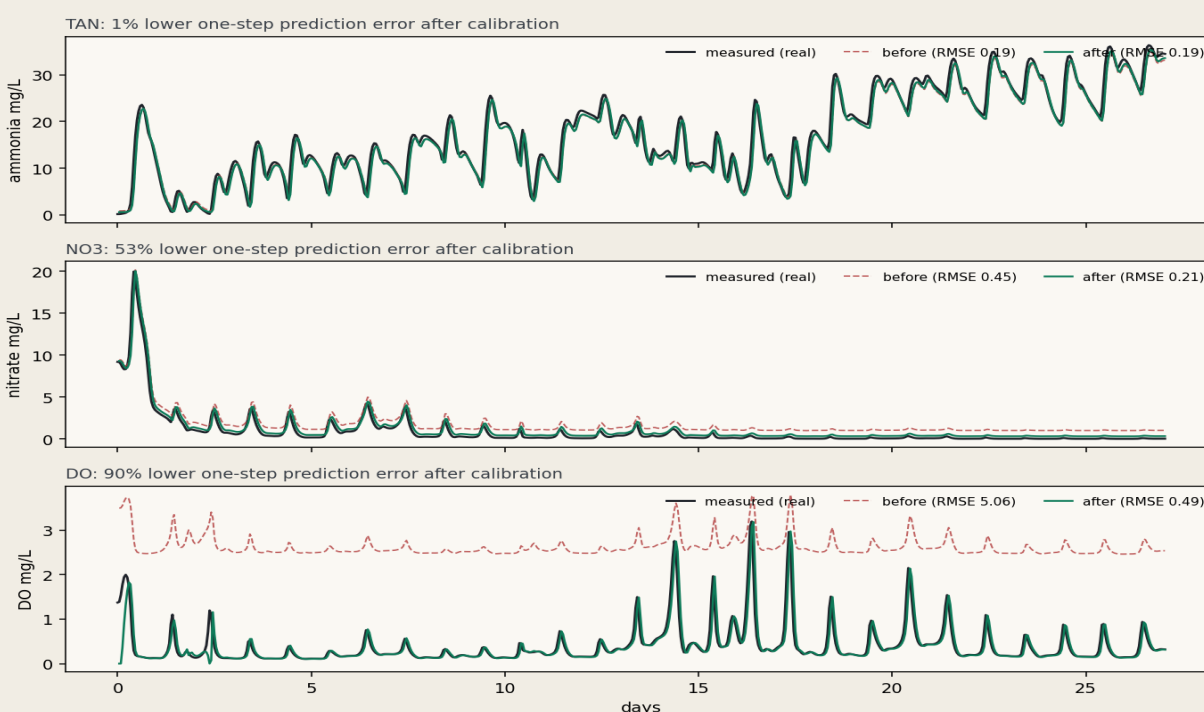
EKF and Terra land close together - the moat is the calibrated process model, not the filter. The naive forecaster's early alarm is a false positive during routine feeding; Terra's calibrated breach probability is what separates a real warning from noise.

3 - Real-data calibration (IWA BSM1 wastewater benchmark)

Uncalibrated, the engine diverges on real data whose scales differ from a domain's defaults. A one-shot calibration - now a one-click button in the console - fits the model to the customer's own log by minimising one-step prediction error. On the real IWA BSM1 benchmark:

Channel	Before (RMSE)	After (RMSE)	Improvement
Ammonia (TAN)	0.192	0.190	~0% (already tracking)
Nitrate (NO3)	0.452	0.213	53% lower
Dissolved oxygen (DO)	5.056	0.488	90% lower

Terra model vs real BSM1 data — before / after calibration



Terra prediction vs. real BSM1 measurements, before (dashed red) and after (green) calibration. 27 days of benchmark data.

4 - Real external simulators - what actually ran

Beyond Terra's own process models, the engine was pointed at real, published third-party simulators. This is the honest inventory.

Simulator	Domain	Status	Result
IWA BSM1 (bsm2-python)	Wastewater	RAN + calibrated	Standard activated-sludge benchmark. Calibrated: DO prediction error -90%, nitrate -53% (section 3).
ADM1 anaerobic digester (EXPOsan / QSDsan)	Biogas / AD	RAN + calibrated	120-day dynamic run. A new reduced-ADM1 domain was built (methanogenic activity as the hidden state) and calibrated to it: one-step prediction error cut from 163.7 to 0.77, acetate error -89%.
SIMOC closed-habitat (your run)	Life support	RAN	Inferred crop photosynthetic capacity settling to 0.75 as cabin CO2 climbed to ~1680 ppm.
WOFOST / PCSE	Agriculture	installed, ready	Wageningen crop-growth model. Import-verified; needs weather + crop params wired.
AquaCrop-OSPy	Agriculture	installed, ready	FAO crop-water model.
BioSTEAM / COBRApy	Bioreactor	installed, ready	Biorefinery process sim and genome-scale metabolic models.
PenSimPy (penicillin)	Bioreactor	blocked	C++ toolchain build fails in this environment; runs on a normal box.

Bottom line

The single engine recovers the hidden biological state and beats raw-threshold monitoring on early-warning lead time across **all six** validated domains, with well-calibrated uncertainty everywhere (NIS/dof below 1), and it calibrates to two real benchmark simulators (IWA BSM1 and the ADM1 digester). The prior report's two caveats are now closed: the bioreactor inconsistency was a solver-stiffness bug, fixed engine-wide with adaptive sub-stepping; anaerobic digestion is a purpose-built reduced-ADM1 domain that both validates and calibrates to real data. The clearest, best-validated commercial wedge remains recirculating aquaculture.

Reproduce it: scripts/run_demo.py and terra.validate (section 1), scripts/leadtime_shootout.py (section 2), scripts/calibrate_real.py and scripts/run_anaerobic.py (sections 3-4).