

Evaluating aeration savings and ammonium exceedances through complete-state matching in a wastewater treatment benchmark

Aymen Zegnoui^{1,2*} , Abdelhalim Benmansour²

¹ University Centre of Maghnia, El-Chouhada Street, El-Zawiya Road, P.O. Box 600, 13300 Maghnia, Algeria

² Materials and Renewable Energy Research Unit (URMER), University Aboubekr Belkaid Tlemcen, BP 119, 13000 Tlemcen, Algeria

* Corresponding author's e-mail: a.zegnoui@cu-maghnia.dz

ABSTRACT

We combined complete-state matching with direct ammonium excursion analysis to assess the effluent consequences of interventions that save aeration energy. Fixed dissolved oxygen (DO) references of 2.00, 1.90 and 1.80 mg O₂/L were compared in Benchmark Simulation Model No. 2 (BSM2) from 42 restored plant states, with identical influent forcing within each 24–48 h comparison. Separate reruns and metric recomputation checked numerical reproducibility. Annual simulations shared a common initial state over days 245–609. At 48 h, median BSM2 aeration-energy indicator (AE) fell by 1.497% and 2.963% at 1.90 and 1.80 mg O₂/L, while flow-weighted effluent ammonium rose by 4.776% and 10.230%. Both responses occurred in all 42 contexts. Annual AE decreased from 4225.433 to 4164.423 and 4104.719 kWh/d, ammonium increased from 0.474291 to 0.501401 and 0.532616 mg N/L, and time above 4 mg N/L rose from 35.75 to 41.50 and 48.75 h. Total nitrogen and the effluent quality index nevertheless improved, and the annual ammonium 95th percentile remained below 4 mg N/L. Under a study-specific post hoc non-aggravation rule, 21 of 168 candidate–horizon comparisons failed for NH₄, with two additional BOD₅ failures (23 overall). Combining matched plant states with direct excursion analysis revealed ammonium penalties that favorable aggregate indicators did not capture. These effects are limited to one BSM2 realization, its influent sequence, the retained PI configuration and the three tested DO references; their magnitudes do not establish electricity savings or ammonium risk at a real treatment plant.

Keywords: complete-state matching, ammonium exceedances, aeration energy, wastewater treatment, benchmark simulation, dissolved oxygen.

INTRODUCTION

Aeration supports carbon oxidation and nitrification and accounts for a substantial share of electricity demand in activated-sludge treatment. Oxygen supply must therefore be adjusted with regard to both energy use and effluent quality. A DO controller can track its reference accurately while that reference remains poorly suited to the current biological demand (Åmand et al., 2013; Rieger et al., 2014). Model-based optimization and full-scale ammonia-based control studies

have examined how this choice affects energy consumption and ammonium removal (Åmand and Carlsson, 2012; Medinilla et al., 2020).

For a treatment plant considering changes to its aeration strategy, a useful starting point is to assess the energy benefit of a lower DO reference alongside its effect on ammonium during vulnerable operating conditions. Such an assessment can help define the performance requirements and safeguards to investigate when developing a subsequent control strategy. Assessing the sustainability of water management requires more than

a single process indicator. Oussama et al. (2026) discuss groundwater recharge in Adrar, Algeria, in relation to local hydrogeology, environmental conditions and health. In wastewater treatment, life-cycle and global impact assessments extend evaluation beyond energy and effluent indicators (Barbu et al., 2017; Meneses et al., 2015). Chen et al. (2021) incorporate life-cycle impact indicators into simulated control of DO and chemical dosing, linking control decisions to broader environmental objectives.

Benchmark models provide a common setting for comparing aeration policies under reproducible influent conditions. BSM2 includes primary clarification, biological nitrogen removal, secondary clarification, sludge treatment and liquid returns (Gernaey et al., 2014; Jeppsson et al., 2007). Its MATLAB–Simulink implementation allows control actions in the aeration tanks to be evaluated across the plant (Vrečko et al., 2006). The default PI controller maintains DO at 2.00 mg O₂/L in the second aerobic compartment, with linked oxygen-transfer coefficients in the other two aerobic compartments (Alex et al., 2008; Nopens et al., 2010).

Published studies have adjusted aeration through ammonia feedback, feedforward control and coordination with sludge-retention-time control (Schraa et al., 2019; Vrečko et al., 2006a). In BSM2, ammonia–DO cascades and nitrogen–energy supervision have been assessed through annual, bimonthly and weekly profiles and through the frequency and duration of effluent-limit violations (Vrečko et al., 2006b; Revollar et al., 2020a; 2020b). Data-driven DO selection has also been investigated, and AI-based aeration control has been reviewed (Monday et al., 2024; Qambar and Al Khalidy, 2022). Recent contributions address full-scale nitrogen–energy control and the effects of non-ideal sensors and actuators (Ruano et al., 2024; Santín et al., 2025). Table 1 summarizes selected control evaluations and their relation to this study.

The energy–ammonium trade-off is established in the literature. Our contribution is to evaluate it by restoring complete plant states, comparing ammonium excursions directly between paired interventions, and checking the calculations at branch and annual scales. This design shows where changes in mean concentration or EQI give a different picture from changes in event duration, frequency, area or peak. Numerical reruns test repeatability, while separately

implemented post-processing checks the derived results. The fixed DO references serve as controlled interventions within this evaluation; reducing a fixed reference is not proposed as a new control method.

We applied this evaluation to fixed DO references of 2.00, 1.90 and 1.80 mg O₂/L using two complementary simulation designs. Tier A comprised 42 matched plant states, each exposed to identical influent conditions over 24–48 h. Tier B followed the official annual evaluation period, days 245–609, from a common state. The DO reference was the only intervention; PI tuning, process equations and other controls were retained. We evaluated aeration energy (AE), EQI and mean effluent concentrations together with ammonium excursion duration, frequency, area and peak.

The aim was to determine whether the lower references saved aeration energy without creating or enlarging ammonium excursions under a study-specific post hoc comparison rule. The short simulations resolve the response in individual operating contexts, while the annual runs show whether the same ordering persists over the benchmark year. Both draw on one BSM2 realization.

MATERIALS AND METHODS

BSM2 plant configuration

We used the BSM2 closed-loop model from the IWA task-group BSM2_vR2019b release, implemented locally in MATLAB R2024a/Simulink. The plant has an average dry-weather influent flow of 20,648.36 m³/d, a 900 m³ primary clarifier, a 12,000 m³ activated-sludge reactor and a 6000 m³ secondary clarifier. Two 1500 m³ non-aerated compartments precede three 3000 m³ aerobic compartments. Biological reactions follow Activated Sludge Model No. 1 (ASM1), and secondary settling is represented by a ten-layer non-reactive model (Gernaey et al., 2014). The closed-loop model file and its initialization script were unchanged from that release. Local input preparation and source provenance are described below and in Supplementary Table S9.

The sludge line comprises thickening, co-digestion of primary and thickened secondary sludge using Anaerobic Digestion Model No. 1 (ADM1), dewatering and reject-water storage. Thickener and dewatering liquors return upstream of primary clarification (Gernaey et al.,

Table 1. Evaluation of selected aeration-control strategies and relation to the present paired evaluation

Study	Strategy evaluated	Evaluation design	Relation to the present study
Rieger et al. (2014)	NH ₄ -based feedforward and feedback aeration	Process-oriented comparison of ammonia response, delays, aeration demand and controller tuning	Establishes NH ₄ as supervisory information; the present setpoints remain fixed
Revollar, Meneses, et al. (2020)	Cascade S _{NH,SP} with dynamic assessment	Annual, weekly and bimonthly plant-wide BSM2 assessment using energy, effluent and environmental indicators	Adds complete-state matching and direct NH ₄ excursion comparisons to evaluation across time scales
Revollar, Vilanova, et al. (2020)	N/E-index hierarchical DO supervision with an event safeguard	BSM2 comparison using global indicators and effluent-violation frequency and duration	Uses an efficiency target and a DO safeguard; the present criterion compares NH ₄ excursions with PI2
Santín et al. (2015)	Two-level variable-DO supervision	BSM1 simulations across supervisory controllers and tunings, assessed with effluent and operating-cost indices	Demonstrates conditional DO supervision; the present study characterizes fixed references
Schraa et al. (2019)	ABAC coordinated with SRT control	Model-based comparison using NH _x removal, nitrification consistency and aeration energy	Coordinates aeration and biological capacity; sludge-wasting control is not modified here
Amand and Carlsson (2012)	Optimal aeration and rapid NH ₄ feedback	Model-based comparison using effluent NH ₄ and aeration demand	Provides a control-oriented reference for the energy–NH ₄ trade-off
Medinilla et al. (2020)	Full-scale ABAC	Field before–after analysis normalized for ammonia loading and air-transfer conditions	Supports operational relevance; the present AE remains a BSM2 indicator
Ruano et al. (2024)	Neuro-fuzzy predictive aeration control	Full-scale evaluation of nitrogen removal and energy use under operating variability	Demonstrates adaptive field control; no adaptive controller is evaluated here
Santín et al. (2025)	Hierarchical ammonium control with non-ideal sensors and actuators	Simulation with filtering, adaptation and measurement-error prediction	Highlights sensing and actuation constraints relevant to later implementation
Present study	Fixed references of 2.00, 1.90 and 1.80 mg O ₂ /L	Forty-two matched 24–48 h contexts and the annual day-245–609 trajectory; AE, EQI, effluent indicators and NH ₄ excursion burden	Combines complete-state matching, numerical verification and NH ₄ excursion comparisons

Note: Comparisons are qualitative because the studies differ in platform, horizon, manipulated variable, comparator and performance definition.

2014). Each checkpoint therefore includes the biological and hydraulic states, PI memory, settler layers, sludge-line and digester states, and sensor states. These states were restored together from the serialized operating point. Here, complete-state matching refers to the plant, controller and subsystem states represented in that object; it does not imply preservation of the solver's internal integration state. Figure 1 shows the layout, and Table 2 defines the intervention.

The intervention was limited to the fixed oxygen reference presented to the existing PI loop. Process equations, plant topology, influent forcing and controller tuning were retained. The plant topology was redrawn from the official IWA BSM2 layout and technical description (Alex et al., 2008).

Dissolved-oxygen PI loop

The lower loop regulates measured oxygen in reactor 4. Using the benchmark sign convention, the tracking error is

$$e(t) = S_{O,ref}(t) - S_{O,4,m}(t) \quad (1)$$

A state-space representation of the PI with back-calculation anti-windup is

$$v(t) = K_p e(t) + x_i(t) \quad (2)$$

$$K_L a_4(t) = \text{sat}[v(t); 0, 360] \quad (3)$$

$$\frac{dx_i}{dt} = \frac{K_p}{T_i} e(t) + \frac{K_L a_4(t) - v(t)}{T_t} \quad (4)$$

$$\begin{aligned} K_L a_3(t) &= K_L a_4(t), \\ K_L a_5(t) &= 0.5 K_L a_4(t) \end{aligned} \quad (5)$$

The MATLAB configuration used $K_p = 25 \text{ m}^3 \text{ g(COD)}^{-1} \text{ d}^{-1}$, $T_i = 0.002 \text{ d}$, $T_t = 0.001 \text{ d}$ and a baseline $S_{O,ref}$ of 2 g(COD)/m^3 , numerically equal to $2 \text{ mg O}_2/\text{L}$ (Alex et al., 2008). The $0.10 \text{ mg O}_2/\text{L}$ decrements were selected to examine the local dose response around the benchmark reference; the three-point grid characterizes these tested references rather than an optimum. Because x_i is part of the operating state, it was restored rather

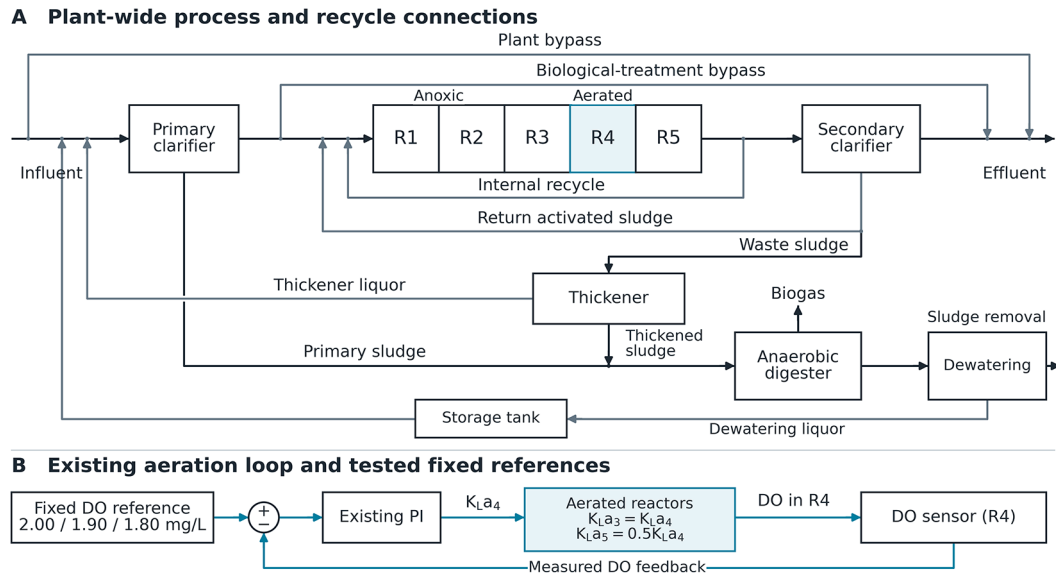


Figure 1. Simplified BSM2 plant layout and the existing aeration loop used for the fixed-reference comparisons

Table 2. BSM2 configuration and study boundary

Item	Specification
Annual preparation	Local states_ bsm2.mat loaded; days 0–245 with dyninfluent; complete day-245 operating point retained
Annual evaluation	Days 245–609 with identical dyninfluent future (364 d)
Activated sludge	2 × 1500 m ³ anoxic + 3 × 3000 m ³ aerobic reactors
Secondary clarifier	6000 m ³ , ten-layer non-reactive settler
Sludge line	Thickening, ADM1 digestion, dewatering and liquid returns
DO references	2.00 (PI reference), 1.90 and 1.80 mg O ₂ /L
Only intervention	DO reference in reactor 4; all other model and control settings unchanged
OCI	Not evaluated; no total operating-cost inference

than reset at each checkpoint. $K_L a_4$ was bounded at 360 d⁻¹ and $K_L a_5$ remained equal to 0.5 $K_L a_4$. At the stored reporting points, no annual branch reached the $K_L a_4$ bound.

Matched contexts and annual simulations

Tier A comprised 42 checkpoints between days 252.510 and 493.354, grouped into six temporal blocks: 252.510–286.865, 296.375–325.521, 331.990–372.375, 383.365–410.469, 416.396–450.385 and 458.458–493.354 d. Seven anchors were selected in each block, with 21 CENTRAL and 21 TAIL labels across the panel and at least 48.75 h between anchors. Selection covered central and stressed hydraulic conditions using current influent flow and three robust descriptors over the preceding six hours: median, interquartile range and Theil–Sen slope. This structured design provided balanced temporal

coverage within the selected window; no alternative checkpoint sets were tested. At each anchor, PI2 (2.00), DO190 (1.90) and DO180 (1.80 mg O₂/L) started from the same complete state and received the same subsequent influent for 48 h. Outcomes at 24 h were extracted from those trajectories. Each branch contained 193 reporting points at 15 min intervals. These are 42 temporally related states from one deterministic realization, not independent biological replicates; the 24 h and 48 h assessments also share trajectories.

Anchor selection preserved the temporal blocks and CENTRAL/TAIL balance, with unique, sufficiently spaced anchors and an available 48 h follow-up at each anchor. Checkpoint metadata linked each restored state to its source simulation. Each branch was checked for the prescribed checkpoint, DO reference, future influent and reporting grid, as well as the required output variables and final-state record.

For Tier B, the closed-loop model was initialized with locally generated states and run with the dynamic BSM2 influent from day 0 to day 245. The complete operating point and random-number-generator state at day 245 were then restored for each of the three DO references. All branches followed the same influent realization to day 609, giving 34,945 timestamps and 34,944 intervals of 15 min over 364 d. The annual comparison thus began from the same plant condition. Figure 2 and Table 3 summarize the design.

Model initialization used states generated locally with the BSM2 steady-state model. The public 14-day sensor-noise sequence was extended to cover the 609-day simulation period by cyclic repetition, omitting the duplicate endpoint of each cycle. These initialization states and the prescribed noise sequence define the BSM2 realization evaluated in this study. The initialization procedure, input provenance and numerical differences from the states supplied with the public release are documented in Table S9.

Tier A estimates local setpoint effects from 42 matched complete states. Tier B follows the same ordered dose response over the annual BSM2 evaluation period. Each rerun saved its trajectories and final state before accessing the production reference. The metric discrepancy in Figure

2 was obtained with the original post-processing formulas and does not validate the BOD₅ definition; the bypass correction is described under Numerical verification.

Performance indicators

Here, NH₄ denotes effluent ammonium nitrogen represented by the BSM2 soluble ammonium-nitrogen state S_{NH} and reported in mg N/L. TN denotes total nitrogen, TKN total Kjeldahl nitrogen, AE the BSM2 aeration-energy indicator (kWh/d), and EQI the BSM2 effluent quality index. KLa is the volumetric oxygen-transfer coefficient (d⁻¹); true and measured DO refer to the process state and sensor output, respectively. AE and EQI were calculated according to the official BSM2 evaluation definitions (Vrečko et al., 2009). For an evaluation interval T of duration t_{obs} days, mean AE was calculated as:

$$AE_T = \frac{S_{O,sat,15}}{1.8 \times 1000 t_{obs}} \int_T \sum_{i \in \{3,4,5\}} V_i K_L a_{i,15}(t) dt \quad (6)$$

where: $S_{O,sat,15} = 8 \text{ g O}_2/\text{m}^3$, V_i is the volume of aerated reactor i , and $K_L a_{i,15}$ is the oxygen-transfer coefficient corrected to 15 °C.

Only reactors 3–5 contribute to AE. The corresponding EQI was:

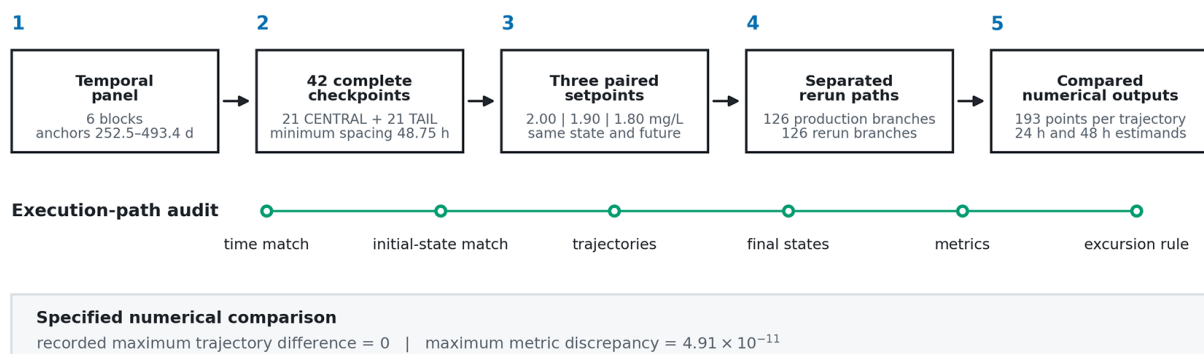


Figure 2. Matched-context experimental design and numerical verification

Table 3. Multi-scale experimental design

Component	Comparison unit	Run structure	Horizon/grid
Tier A: contexts	42 checkpoints × 3 setpoints	126 productions + 126 separate computational reruns	24 and 48 h; 15 min
Tier A: matching	Complete plant state	Same future influent within each checkpoint	Only SO ₄ ref differs
Tier B: annual	1 day-245 state × 3 setpoints	3 productions + 3 separate computational reruns	Days 245–609; 15 min
Tier B: preparation	Local initialization states loaded; days 0–245 dyninfluent	One canonical complete state	No action-specific re-initialization
Numerical verification	Separate output paths	Production reference loaded only after rerun outputs were saved	Specified array-level comparison

$$\text{EQI}_T = \frac{1}{1000 t_{\text{obs}}} \int_T Q_e(t) W_e(t) dt. \quad (7)$$

$$W_e(t) = 2\text{TSS}_e(t) + \text{COD}_e(t) + 30S_{\text{NKj},e}(t) + 10S_{\text{NO},e}(t) + 2\text{BOD}_{5,e}(t)$$

In this expression, Q_e is total effluent flow. The settler effluent and both bypass streams were combined on a load basis before the effluent variables were calculated (Vrečko et al., 2009). Here, $W_e(t)$ denotes the weighted sum of effluent pollutant concentrations shown in Equation 7. EQI is expressed in kg pollution units/d. The stored effluent signal is already the combined stream. Recalculation therefore uses it directly and applies the separate bypass BOD₅ weighting without adding the bypass flows a second time.

BOD₅ assigns factors of 0.25 to treated biodegradable material and 0.65 to bypass material, using $\text{SS} + \text{XS} + 0.92(\text{XBH} + \text{XBA})$ for the biodegradable concentration. With the combined effluent signal, BOD₅ equals 0.25 times its biodegradable concentration plus 0.40 times the total bypass biodegradable load divided by total effluent flow. This retains the separate bypass weighting; EQI uses the resulting BOD₅.

Means for effluent NH₄, TN, COD, BOD₅ and TSS were weighted by effluent flow. We used the BSM2 evaluation thresholds of 4 mg N/L for NH₄, 18 mg N/L for TN, 100 mg/L for COD, 10 mg/L for BOD₅ and 30 mg/L for TSS; these are benchmark evaluation values rather than universal discharge limits. Annual empirical 95th percentiles were calculated from the 34,944 interval-start samples on the 15 min grid, and exceedance fractions were obtained by dividing exceedance duration by 8736 h. Initialization sensitivity was examined by recalculating annual AE, NH₄ and TN contrasts after omitting the first 48 h and the first seven days. Excursion durations, episode counts, excess areas, maxima and KLa saturation checks refer to the 15 min reporting grid. Shorter excursions, inter-sample peaks and brief saturation cannot be excluded by these records. The NH₄ event analysis contains 168 candidate comparisons (42 contexts × 2 reduced references × 2 horizons); extending the same comparison to five pollutants gives 840 rows.

Annual energy equivalents were calculated by multiplying the difference in mean AE by 364 d. AE represents aeration demand in BSM2. Its conversion into blower electricity, total operating costs or a net plant energy balance was outside this evaluation. OCI, pumping and mixing energy,

blower efficiency, electricity tariffs, sludge production, biogas recovery and greenhouse-gas emissions were not evaluated as outcomes.

For pollutant $c(t)$ at effluent flow $Q(t)$, the flow-weighted mean over evaluation interval T was

$$\bar{c}_Q = \frac{\int_T Q(t)c(t) dt}{\int_T Q(t) dt}. \quad (8)$$

For the NH₄ threshold $C_{\text{thr}} = 4$ mg N/L, duration was expressed in h and integrated excess in mg N·h/L, with time in Equation 9 measured in hours:

$$D_{\text{exc}} = \int_T \mathbb{1}[c(t) > c_{\text{thr}}] dt, \quad (9)$$

$$A_{\text{exc}} = \int_T \max[c(t) - c_{\text{thr}}, 0] dt$$

We formulated a study-specific non-aggravation criterion post hoc so that aeration savings or improvements in aggregate effluent quality could not compensate for new or enlarged pollutant excursions relative to PI2 under the stated allowances. This analytical criterion does not define regulatory compliance. Numerical tolerances were 1.4×10^{-8} for NH₄, 2.8×10^{-8} for TN, 1.1×10^{-7} for COD, 2.0×10^{-8} for BOD₅ and 4.0×10^{-8} for TSS, applied in the respective reported units. The duration allowance was one 15 min reporting interval, and an allowance of one episode accommodated splitting or merging at an excursion boundary. These allowances were numerical conventions for the comparison, not operational margins. The candidate flow-weighted mean had to remain below the relevant BSM2 threshold plus the numerical tolerance. If PI2 had no excursion, the candidate also had to show none. If PI2 already had an excursion, the candidate could add at most 15 min of duration and one episode, with no increase in integrated area or peak excess beyond numerical tolerance. A comparison passed only if the mean and all excursion conditions were satisfied. Relative changes in the mean were reported separately. Applying this rule gives 21 NH₄ failures and two additional BOD₅ failures; it does not affect the 42/42 directions of the paired NH₄ means or the annual burden calculations.

We examined NH₄ classification sensitivity over a fixed grid of 720 combinations. The grid crossed three numerical tolerances (0, 1.4×10^{-8} and 10^{-6}), four duration allowances (0, 0.25, 0.50 and 1.00 h), three episode allowances (0, 1 and 2),

five area allowances (0, 0.10, 0.25, 0.50 and 1.00 mg N·h/L), and four peak allowances (0, 0.05, 0.10 and 0.25 mg N/L). A new excursion in a candidate whose matched PI2 branch had none remained a failure in every scenario. Table 4 summarizes the indicators and their roles in the comparison. The sensitivity grid was specified after the trajectories had been generated and therefore constitutes a post hoc analysis. Neither the study nor its primary rule was independently preregistered.

Numerical verification

Simulations used the variable-step ode45 solver, with relative and absolute tolerances of 1×10^{-5} and 1×10^{-8} and automatic maximum-step selection. The stored FixedStep value of 0.01 d was inactive in this mode. Each of the 126 Tier-A branches and the three annual branches was rerun in a separate output path using the same MATLAB R2024a/Simulink implementation, numerical options, checkpoint, influent and reporting grid. Annual reruns also restored the random-number-generator state.

Each rerun saved its trajectories, metrics and final state before accessing the production outputs. Comparisons covered timestamps, dimensions, numeric classes, specified trajectories, complete final states, official metrics and event masks, using the same model file throughout. These comparisons assess repeatability within the recorded implementation. Because identical reruns can reproduce a shared implementation error, agreement alone does not exclude software errors or validate the model against an alternative implementation or experimental observations.

Separate Python calculations reconstructed 252 Tier-A metric rows from numeric extracts of

the 126 production trajectories. Reconstruction of BOD₅ from treated and bypass loads identified an omitted bypass contribution in the original short-horizon post-processing. The correction changed BOD₅ and EQI at N022 over 48 h for all three references and added two BOD₅ rule failures; the 21 NH₄ failures were unchanged. The calculations reproduced all 48 median/IQR entries of Table 5 and all 75 numerical entries of Table 8 at their displayed precision, all 23 rule failures and all 720 NH₄ sensitivity outcomes. The 756 KLa records were also checked against the stored trajectories. These calculations verify post-processing of existing outputs; they do not rerun BSM2 or provide third-party model validation. Calculation details and executable scripts are provided in EEET_Verification_Package.zip.

A portable MATLAB procedure used the local model files and 43 serialized checkpoints. Configuration checks covered all 129 cases without running simulations. Four additional replay executions assessed three selected cases: N001_DO200 twice, N001_DO190 once and ANNUAL_DO200 once. All 49 stored numeric fields agreed exactly with the production and rerun references for the two Tier-A cases and with the production reference for annual DO200. Python comparisons also found exact agreement of the stored final-state representation and random-number-generator state (Table S2). These four executions supplement the 129 production–rerun pairs and do not constitute a new replay of all 129 cases.

During these replays, Simulink restored the block states but discarded the initial solver state. Agreement of the sampled outputs and stored final states therefore does not establish restoration of solver internals or equivalence to uninterrupted integration. The annual replay reproduced

Table 4. Outcomes and interpretation hierarchy

Outcome	Unit	Calculation	Decision role
AE	kWh/d	Paired change; annual mean	Resource outcome; never compensates NH ₄ failure
EQI	kg pollution units/d	Official load-based BSM2 calculation	Aggregate environmental index
NH ₄	mg N/L	Flow mean, p95, maximum, duration, episodes, area, peak	Pollutant-specific outcome treated as non-compensable under the study rule
TN	mg N/L	Flow mean, p95 and threshold burden	Plant-wide nitrogen outcome
COD/BOD ₅ /TSS	mg/L	Flow mean; p95 for TSS; predefined event masks	Secondary effluent outcomes
DO/KLa	mg/L; d ⁻¹	True/measured DO; $K_{L,a,3/4/5}$	Controller tracking and effort
Final state	Complete structure	All numeric leaves	Complete-state numerical comparison
OCI	—	Not evaluated	Outside the reported outcome set

12 of 14 interval columns exactly. The remaining BOD₅ and EQI differences arose from floating-point operation grouping: applying the original grouping to the same raw output reproduced all 14 columns exactly. Table S2 and the replay verification note distinguish the native comparison from this recalculation.

Descriptive analysis

For Tier A, each comparison was the difference between a reduced-reference branch and PI2 started from the same complete state. We report medians, interquartile ranges (IQRs), observed ranges and direction counts for the 42-state panel. The anchors follow one temporal realization, so these are descriptive summaries; no inferential p-values, confidence intervals or signed-rank tests were calculated. A post hoc analysis compared temperature, influent flow and NH₄ margin under matched PI2 between contexts that passed or failed the excursion rule. This analysis characterized the observed contexts without fitting a predictive model or selecting an operating cutoff. Annual results were used to quantify excursions, examine initialization sensitivity and compare the responses across time scales. Uncertainty across alternative influent realizations was not quantified. Resampling these linked states as if they were independent would not establish that uncertainty.

RESULTS

Agreement between reruns and controller behavior

The production–rerun comparison showed exact agreement of the specified trajectories and complete final states for all 126 Tier-A branches and the three annual branches. Metric reconstruction with the original post-processing formulas agreed within 4.91×10^{-11} . This agreement does not validate the BOD₅ formula, which required correction of the bypass contribution. Direct reconstruction from the stored production trajectories reproduced KLa means within 5.12×10^{-13} d⁻¹ and the $K_L a_4$ saturation fractions exactly. The trajectories also satisfied $K_L a_3 = K_L a_4$ and $K_L a_5 = 0.5 K_L a_4$ exactly. Across the 24/48 h branches, global $K_L a_4$ maxima were 233.119/236.920 d⁻¹ for DO200, 228.607/232.414 d⁻¹ for DO190 and 224.176/227.978 d⁻¹ for DO180. At the stored

reporting points, none of the 42 contexts reached either 360 d⁻¹ or the 99% proximity level of 356.4 d⁻¹ at either horizon, and no annual branch reached 360 d⁻¹. No $K_L a_4$ saturation was observed on this grid; brief saturation between reporting points cannot be excluded.

Short-horizon dose response across 42 matched contexts

At 48 h, AE decreased in all 42 contexts by a median of 1.497% under DO190 and 2.963% under DO180. Absolute differences ranged from −75.226 to −46.123 kWh/d and from −148.696 to −91.236 kWh/d, respectively. Flow-weighted NH₄ increased in every context, with median changes of +4.776% and +10.230%. Absolute increases ranged from +0.002486 to +0.075862 and from +0.005248 to +0.162535 mg N/L. A higher mean did not necessarily cause a failure of the excursion rule. Table 5 gives the medians and interquartile ranges; Figures 3 and 4 show the paired responses.

Every context shifted toward lower AE and higher mean NH₄ under the reduced references. Excursion-rule failures occurred only in some contexts and do not represent the full set of NH₄ increases. Lines join outcomes obtained from the same complete plant state. For each reduced reference, AE decreased and NH₄ increased in 42/42 contexts.

Ammonium excursion classifications

Across the five pollutants, the analysis comprised 420 PI2 self-comparisons (42 contexts × 2 horizons × 5 pollutants) and 840 candidate comparisons (42 × 2 reduced references × 2 horizons × 5 pollutants), for a total of 1260. All PI2 self-comparisons passed. Of the 840 candidate comparisons, 817 passed and 23 failed (2.74%). The failures comprised 21 NH₄ comparisons across seven anchors and two BOD₅ comparisons at a further anchor, N022. Each failure concerned a different candidate–horizon pair, giving 23/168 failed pairs overall (13.69%). Within the NH₄ subset, 21/168 comparisons failed (12.5%): 10 for DO190 across six contexts and 11 for DO180 across seven. Table 6 reports NH₄ outcomes only.

Table 6 groups the NH₄ failures into 16 instances of material worsening, three new exceedances and two instances of small rule-based worsening. These descriptive categories do not determine pass/fail or establish operational severity. All 504 candidate comparisons for TN,

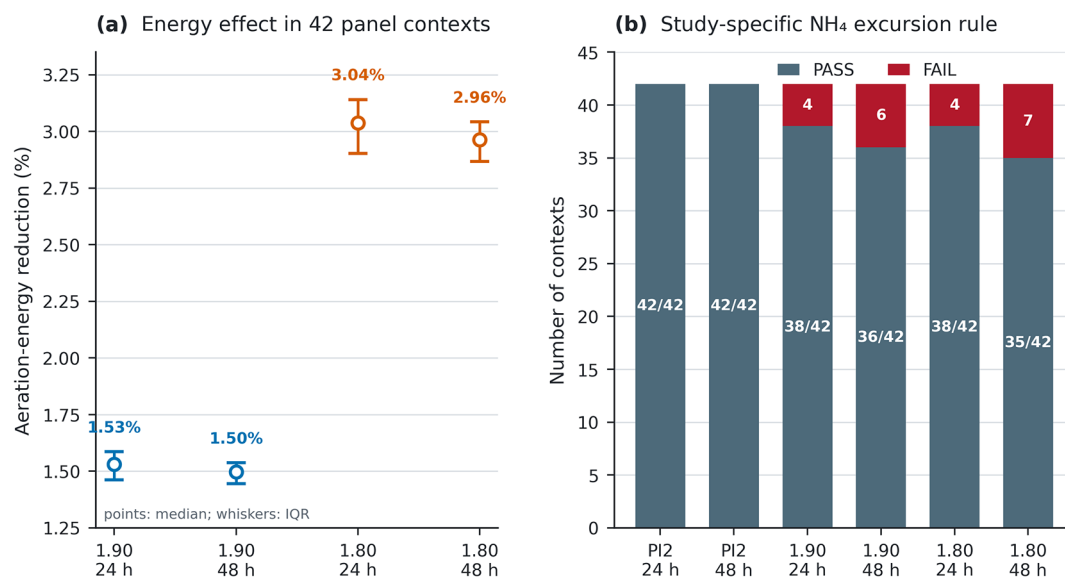


Figure 3. Short-horizon aeration-energy reductions and study-specific NH_4 excursion-rule outcomes across 42 state-matched contexts

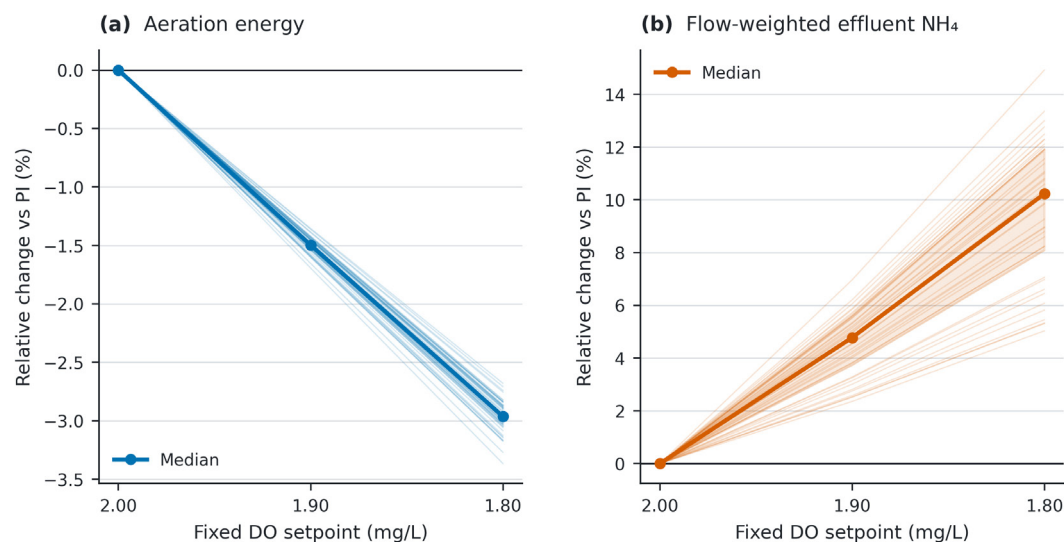


Figure 4. Paired 48 h responses for aeration energy and flow-weighted NH_4

COD and TSS passed; 166/168 BOD_5 comparisons passed. In the sensitivity analysis, the exact no-allowance rule, the study rule and a temporal relaxation allowing 1 h and two additional episodes each produced 21/168 NH_4 failures.

Across all 720 scenarios, the median was 21 and the range was 10–21 failures; 567/720 scenarios retained all 21. Allowing 0.25 mg $\text{N}\cdot\text{h}/\text{L}$ in area and 0.10 mg N/L in peak reduced the count to 19. The broadest tested margins—1 h,

Table 5. Paired relative effects across 42 complete-state contexts, median % [IQR]

Horizon	Action	AE	EQI	NH_4 mean	TN mean
24 h	DO190	−1.531 [−1.587, −1.461]	−0.269 [−0.349, −0.058]	+4.104 [+2.728, +5.060]	−0.750 [−0.830, −0.663]
24 h	DO180	−3.037 [−3.141, −2.903]	−0.544 [−0.709, −0.096]	+8.816 [+5.883, +10.764]	−1.537 [−1.698, −1.360]
48 h	DO190	−1.497 [−1.537, −1.445]	−0.449 [−0.571, −0.236]	+4.776 [+3.747, +5.553]	−1.203 [−1.277, −1.079]
48 h	DO180	−2.963 [−3.042, −2.867]	−0.895 [−1.154, −0.447]	+10.230 [+8.113, +11.904]	−2.444 [−2.595, −2.199]

Note: Percentages are relative to the PI2 branch started from the same complete checkpoint under identical influent conditions.

Table 6. NH₄ excursion-rule classifications by setpoint and horizon

Action	24 h fail	48 h fail	Unique contexts	Failed rows	Archived descriptive label
DO190	4/42	6/42	6	10	7 material; 1 new; 2 small rule-based
DO180	4/42	7/42	7	11	9 material; 2 new; 0 small rule-based
Total	8 rows	13 rows	7 unique	21	16 material; 3 new; 2 small rule-based

two episodes, 1.00 mg N·h/L in area and 0.25 mg N/L in peak—reduced it to 10. Classification was therefore stable under the tested numerical and temporal allowances but changed when larger area and peak increments were permitted. Paired mean NH₄ increased in all 168 candidate–horizon comparisons regardless of the rule.

The two BOD₅ failures occurred at N022 for DO190 and DO180 over 48 h. With both treated and bypass loads included, each of the three branches had one 15 min exceedance of 10 mg/L. The maxima were 10.229631 mg/L for PI2, 10.229968 mg/L for DO190 and 10.230335 mg/L for DO180. Both candidates increased peak excess and integrated excess relative to PI2 beyond the numerical tolerance and therefore failed the area/peak conditions. These failures represent small enlargements of an excursion already present in the matched reference. The rule does not establish their operational significance. Table S10 gives the corrected values.

Total nitrogen and aggregate effluent quality

TN decreased in all 42 contexts at both reduced references. At 48 h, median changes were −1.203% for DO190 and −2.444% for DO180, with absolute differences of −0.207325 to −0.105922 and −0.425139 to −0.216143 mg N/L. EQI improved in 38/42 contexts for each reference. COD and BOD₅ increased throughout the panel, although their median relative changes remained below 0.011% and 0.050%. TSS changes were smaller and varied in sign. All 21 NH₄ excursion failures coincided with lower AE and TN. EQI also improved in seven of those comparisons and worsened in 14, with no ties. The seven improvements show that a lower aggregate effluent index can accompany a failure to preserve NH₄ excursion performance.

Operating conditions associated with excursion failures

NH₄ excursion failures were concentrated in the later, colder blocks. For DO180 at 48 h, blocks 1–3 had no failures (0/21), while blocks 4, 5 and 6 had two, four and one. Median influent temperature at the checkpoint was 10.17 °C in the seven failed contexts and 14.96 °C in the 35 passing contexts; median flow was 24,496 versus 20,095 m³/d. NH₄ under matched PI2 showed the clearest descriptive separation. Failed contexts had a median maximum of 4.527 mg N/L and a median mean of 1.254 mg N/L, compared with 1.242 and 0.349 mg N/L in passing contexts. These are associations within the selected panel.

All five contexts in which matched PI2 already exceeded 4 mg N/L failed under DO180. Two others, N034 and N037, developed new exceedances from PI2 maxima of 3.943 and 3.808 mg N/L. Three failures were labelled CENTRAL and four TAIL, indicating that the sampling label alone did not distinguish the vulnerable contexts. Lower temperature, higher flow and less NH₄ margin under PI2 accompanied failures, but no predictive model or operating cutoff was derived. Table 7 and Figures 5–7 show these exploratory comparisons:

- Entries are descriptive medians, except for the CENTRAL/TAIL counts. Temperature is measured at the checkpoint; flow is averaged over the following 48 h. The contexts are deterministic and temporally structured.
- Excursion failures cluster in late, cold contexts with limited NH₄ margin under matched PI2. The display is exploratory and does not define a control rule.
- Initial flow does not fully separate pass and fail contexts. Interpretation requires the plant state, temperature, subsequent load and NH₄ margin under matched PI2.
- Spearman coefficients summarize rank associations in this fixed panel; they are not population parameters or a fitted policy.

Table 7. Secondary exploratory description of DO180 48 h contexts

Variable	7 fail contexts	35 pass contexts	Interpretation
Influent temperature at checkpoint, °C	10.17	14.96	Cold states are over-represented
Mean influent flow, m ³ /d	24,496	20,095	Higher hydraulic loading in fail group
Matched PI NH ₄ max, mg N/L	4.527	1.242	Baseline NH ₄ margin differs markedly
Matched PI NH ₄ mean, mg N/L	1.254	0.349	Baseline nitrification stress differs
CENTRAL/TAIL	3 / 4	18 / 17	Sampling label is not sufficient

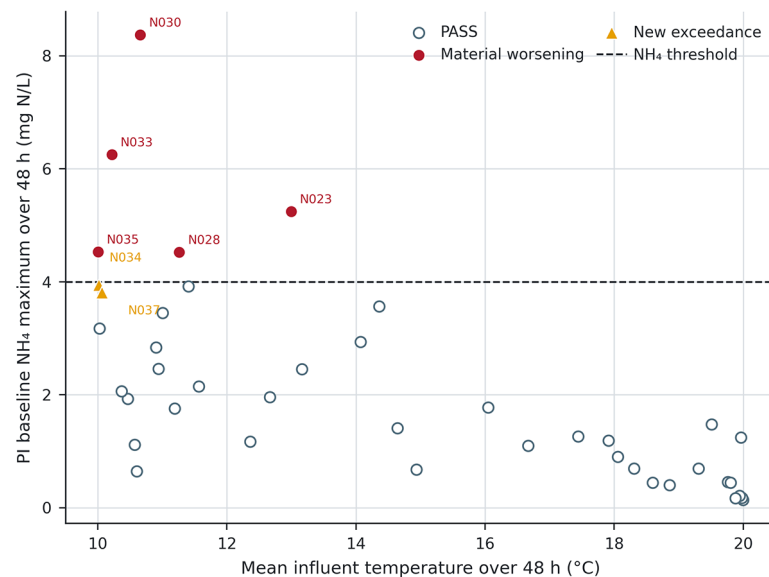


Figure 5. Secondary exploratory description of the physical context of 48 h NH₄ excursion risk. The horizontal axis shows mean influent temperature over the 48 h evaluation horizon, whereas Table 7 reports influent temperature at the checkpoint

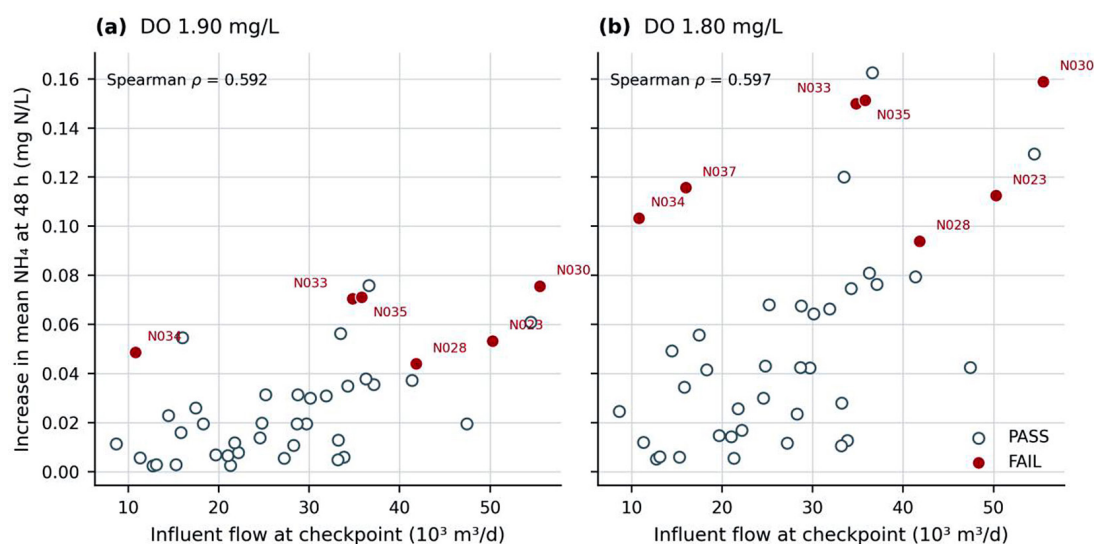


Figure 6. NH₄ sensitivity to initial influent flow across matched contexts

Annual response to the fixed references

The annual branches retained the same ordering over days 245–609. Relative to DO200,

mean AE decreased by 1.444% and 2.857% under DO190 and DO180, corresponding to approximately 22,208 and 43,940 kWh over 364 d. Flow-weighted NH₄ increased by 5.716% and

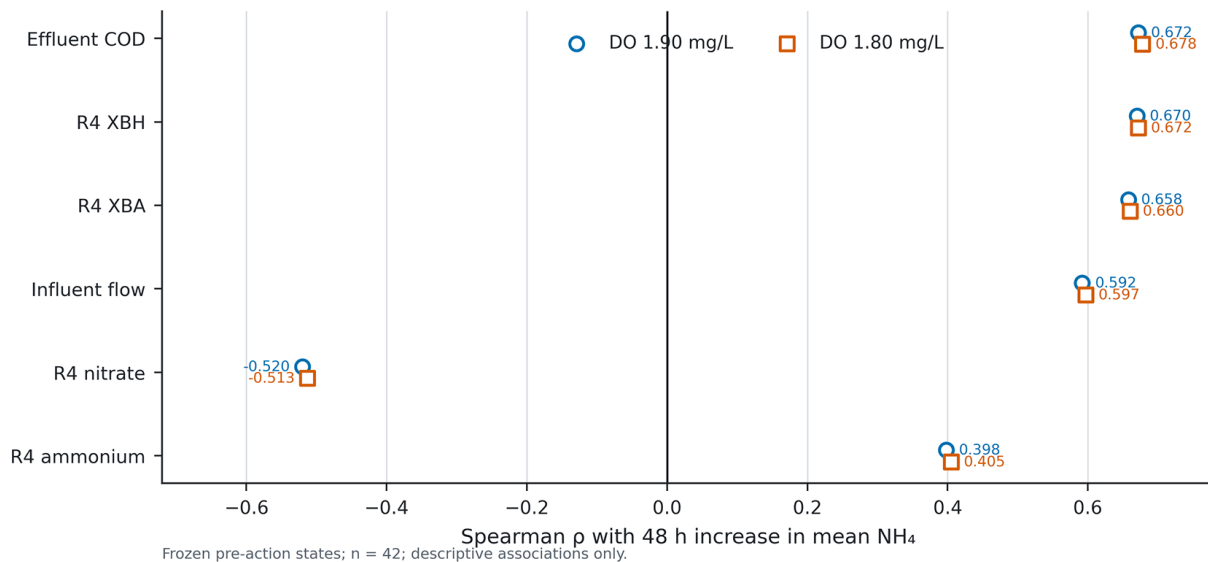


Figure 7. Exploratory rank associations between initial physical variables and 48 h effects

12.297%. Time above 4 mg N/L rose from 35.75 h to 41.50 and 48.75 h, with increases in episode count and integrated excess. Although the exceedance fraction rose from 0.409% to 0.475% and 0.558%, the empirical NH_4 95th percentile remained below 4 mg N/L for all three references. TN decreased, while changes in COD, BOD₅ and TSS were small. Table 8 gives the annual values, Figure 8 shows the main relative changes, and Table S5 provides the full set of relative changes.

Relative changes are measured against DO200 over days 245–609. Lower AE, EQI and TN coincide with higher mean NH_4 and greater excursion burden.

Agreement across evaluation horizons

The short branches and annual simulations gave the same ordering of the three references. Annual AE changes were close to the 48 h panel medians, while the NH_4 and TN responses were somewhat larger over the year. Removing the first 48 h scarcely altered the annual contrasts: AE $-1.444\%/-2.856\%$, NH_4 $+5.719\%/+12.304\%$ and TN $-1.715\%/-3.463\%$ for DO190/DO180. After removing seven days, the corresponding changes were AE $-1.444\%/-2.856\%$, NH_4 $+5.727\%/+12.321\%$ and TN $-1.714\%/-3.461\%$. The annual ordering therefore persisted beyond the initial transition. The second 0.10 mg O_2/L reduction produced slightly less additional AE saving

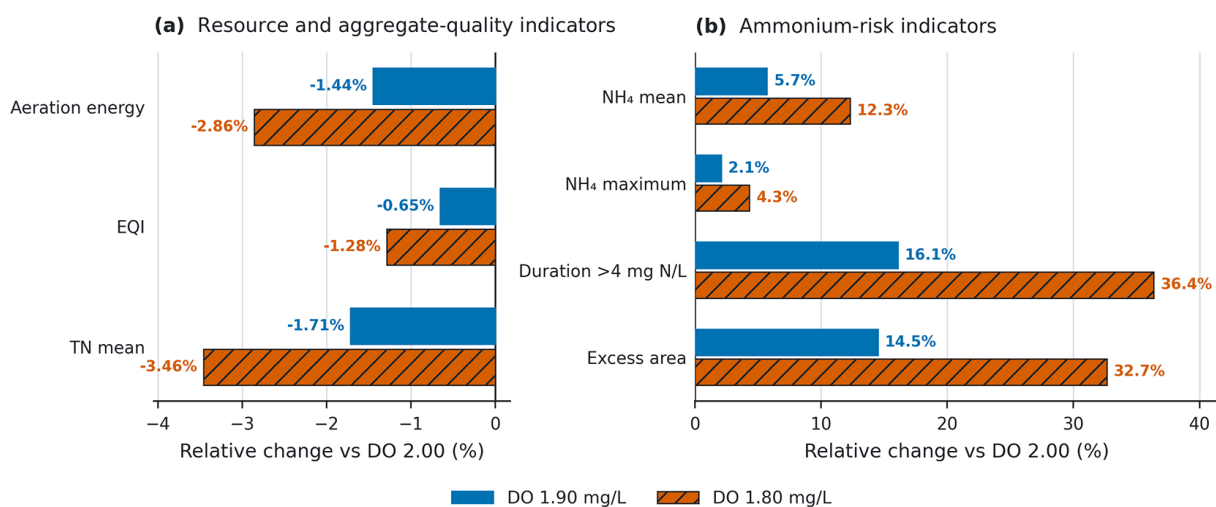


Figure 8. Annual fixed-DO dose response

Table 8. Annual plant-wide BSM2 results using the official evaluation definitions, days 245–609; excursion and maximum values refer to the 15 min reporting grid

Indicator	DO200	DO190	DO180
AE, kWh/d	4,225.433	4,164.423	4,104.719
EQI, kg pollution units/d	5,576.665	5,540.453	5,505.059
NH ₄ mean, mg N/L	0.474291	0.501401	0.532616
NH ₄ maximum, mg N/L	8.370566	8.544851	8.731611
NH ₄ empirical p95, mg N/L	1.542	1.650	1.775
NH ₄ time >4, % of year	0.409	0.475	0.558
NH ₄ duration >4, h	35.75	41.50	48.75
NH ₄ episodes >4	11	13	14
NH ₄ excess area, mg N·h/L	48.199	55.190	63.938
NH ₄ peak excess, mg N/L	4.370566	4.544851	4.731611
TN mean, mg N/L	13.529427	13.297743	13.061576
TN empirical p95, mg N/L	16.752	16.515	16.277
TN time >18, % of year	1.176	0.939	0.710
COD mean, mg/L	49.016567	49.019416	49.022611
BOD ₅ , mg/L	2.790649	2.791409	2.792262
TSS mean, mg/L	15.171307	15.171159	15.170986
TSS empirical p95, mg/L	19.737	19.736	19.736
TSS time >30, % of year	0.343	0.343	0.343
True DO mean, mg/L	2.000412	1.900415	1.800419
Measured DO mean, mg/L	1.999838	1.899840	1.799843
$K_L a_4$ mean, d ⁻¹	126.762978	124.932684	123.141569
$K_L a_5$ mean, d ⁻¹	63.381489	62.466342	61.570785
$K_L a_4$ maximum, d ⁻¹	252.280081	247.377674	242.849943
TN duration >18, h	102.75	82.00	62.00
TN excess area, mg N·h/L	83.584	66.252	53.429
OCI	Not evaluated	Not evaluated	Not evaluated

Table 9. Concordance between 48 h matched-context medians and annual changes

Outcome	DO190 48 h	DO190 annual	DO180 48 h	DO180 annual
AE	–1.497%	–1.444%	–2.963%	–2.857%
EQI	–0.449%	–0.649%	–0.895%	–1.284%
NH ₄ mean	+4.776%	+5.716%	+10.230%	+12.297%
TN mean	–1.203%	–1.712%	–2.444%	–3.458%

and a larger additional NH₄ increase than the

first. Table 9 compares the short and annual responses.

DISCUSSION

Paired assessment of aeration savings and ammonium performance

Restoring each complete plant state allowed the three references to be compared without differences inherited from earlier operating histories. The lower references reduced AE and raised mean NH₄ in every matched context. Separate reruns and metric recomputation supported the

numerical repeatability of these contrasts. The annual simulations retained the same ordering: AE fell by 1.444% and 2.857%, but the duration of NH_4 exceedances increased by 16.084% and 36.364%. The second 0.10 mg O_2/L reduction produced a larger additional ammonium penalty for a slightly smaller additional energy saving (Table 9). The sampled oxygen-transfer coefficients remained below the saturation limit. Tracking the reduced reference nevertheless did not establish that it preserved ammonium performance.

This energy–ammonium trade-off is consistent with the annual BSM2 results of Revollar, Meneses, et al. (2020). Their ammonia–DO cascade reduced electricity use per treated volume by 7.6%, TN by 16.2% and EQI by 3.7%, while mean ammonium increased from approximately 0.474 to 1.052 mg N/L. Lower TN and EQI therefore coexisted with reduced ammonium removal. Their energy saving cannot be compared directly with our 1.44–2.86% AE reductions: their controller adjusted DO to an ammonium target, and their electricity indicator included aeration, pumping and mixing. Our paired design characterizes these responses within individual operating contexts around 2.00 mg O_2/L , with the DO reference as the only changed control input.

The control actions included in a comparison also determine how its energy benefit can be interpreted. Schraa et al. (2019) coordinated ammonia-based aeration with SRT control, allowing biological capacity and oxygen supply to be managed together. Here, sludge-wasting control was retained, and each short comparison began with the same biomass and other plant states. Do et al. (2026) examined a broader combination of aeration scheduling, DO control, airflow and selector configuration in a BioWin–CapdetWorks study initialized with plant data. Those scenarios assessed several operational changes together. Restricting the intervention to the DO reference gives the paired contrasts a clear attribution within BSM2, while coordinated biological and aeration changes remain outside the tested application.

Information added by direct excursion comparisons

Examining the pollutant contributions to annual EQI explains why direct ammonium assessment can yield a different interpretation from aggregate effluent quality. Relative to DO200, the weighted nitrate contribution fell by 53.654

units under DO190 and 109.102 under DO180. These reductions outweighed increases of 17.358 and 37.318 units in the TKN contribution, which includes ammonium; the combined TSS, COD and BOD_5 changes were only +0.084 and +0.178 units. Net EQI therefore fell by 36.212 and 71.606 units. Rieger et al. (2014) describe how limiting aeration can favor denitrification while relaxing complete nitrification. Our lower TN and higher NH_4 are consistent with that trade-off, but the effluent balances alone do not separate changes in the underlying reaction rates.

The comparison also depends on temporal resolution. Revollar, Meneses, et al. (2020) examined weekly and bimonthly profiles as well as annual means to identify changes concealed by annual aggregation. Here, excursions are evaluated directly on the 15 min reporting grid against the corresponding PI2 trajectory. NH_4 exceeded 4 mg N/L for only 0.409–0.558% of the year, so the 95th percentile remained below the threshold even as excursion duration, frequency and integrated excess increased (Table 8). Comparing duration, episode count, integrated excess and peak between matched branches identifies whether a reduced reference creates an exceedance or enlarges an existing one. It also distinguishes a rise in the mean from a failure of the excursion rule: the former occurred in every context, whereas NH_4 rule failures occurred in 21 candidate–horizon comparisons.

The non-aggravation criterion determines how these trade-offs are classified. All 21 NH_4 -failed comparisons saved AE and reduced TN, and seven also improved EQI. The eco-efficiency assessment of Revollar et al. (2021) considers a wider set of objectives, including energy use, biogas, effluent quality and emissions across seasonal operating windows. Such an assessment can identify useful compromises across the plant. Our rule retains each NH_4 failure when gains occur in AE, TN or EQI. The paired classification therefore makes the ammonium consequence of such a compromise explicit under the stated allowances.

Using paired assessments to inform control development

The paired outcomes identify conditions in which an aeration reduction warrants closer assessment: failures were concentrated in colder, later contexts. Revollar et al. (2021) also incorporate seasonal temperature effects when

selecting control actions. In our DO180 48 h comparison, the seven failed contexts had lower median temperature, higher flow and substantially less NH_4 margin under PI2 than the 35 passing contexts (Table 7). However, neither initial flow nor the CENTRAL/TAIL label separated all outcomes. The NH_4 means and maxima used in this comparison came from the matched PI2 trajectories over the evaluation horizon; they are retrospective descriptions, not measurements that would all be available when a controller chooses a reference.

Ammonia-based control offers a way to respond to changing treatment demand. Rieger et al. (2014) distinguish feedback used to limit aeration or contain ammonia peaks from feedforward action intended to anticipate influent disturbances. Revollar et al. (2020) reported effluent-violation frequency and duration and incorporated a safeguard against prolonged operation at excessively low DO in their N/E-index supervisor. Their safeguard acts during control; our excursion rule assesses the resulting effluent against a reference trajectory. The present results indicate why both functions matter: an adaptive policy needs to choose when oxygen can be reduced, and its evaluation needs to establish whether the resulting NH_4 events meet the stated evaluation criteria. The fixed-reference comparisons supply benchmark cases for evaluating a subsequent safeguard, while the present study does not establish a validated switching threshold.

Field evidence gives a further perspective on the energy result. Medinilla et al. (2020) evaluated ABAC before and after installation using metered blower consumption normalized for ammonia loading and air-transfer conditions, and estimated annual energy-cost savings of about $9 \pm 1\%$. Our AE values describe modelled oxygen-transfer demand and do not contain that plant-specific electrical information. A practical assessment would also need to address the sensing and actuation imperfections considered in hierarchical ammonium control by Santín et al. (2025).

In a model calibrated against plant data, restoring the same initial state could support comparisons of proposed DO adjustments across challenging operating scenarios. Direct excursion metrics could then test whether energy gains coincide with new or more persistent ammonium exceedances. Causal ammonium measurements, loading and temperature could be examined as inputs to conditional oxygen reduction. These comparisons

would help specify monitoring and recovery tests; operating thresholds and expected savings would still require plant-specific measurements.

Scope and limitations

Complete-state matching removes differences inherited before each short intervention, including differences in biomass, hydraulics and controller memory. The annual runs then allow the consequences of maintaining each reference to develop over the benchmark year. Their agreement with the short comparisons, including after omitting the first 48 h or seven days, supports persistence within this BSM2 realization. The 42 checkpoints are nevertheless temporally structured states from one trajectory, and the annual comparison uses the same realization. Separate computational reruns confirm repeatability in the recorded MATLAB/Simulink implementation; they provide neither independent biological replication nor cross-platform validation.

The primary excursion rule was formulated post hoc during the analysis. Its NH_4 sensitivity grid yielded 10–21 failures across 720 combinations, with 567 retaining all 21. Numerical and temporal allowances did not change this count in the reported scenarios, whereas larger area and peak allowances did. Classification therefore depends on the chosen margins; the paired NH_4 mean changes and measured annual excursion burdens do not. For the listed study, moderate-area/peak and broadest-margin scenarios, the count changes from 21 to 19 to 10. This count is a result of the chosen decision rule, not a universal measure of ammonium risk.

The results apply to the selected BSM2 climate and influent sequence, checkpoints and PI configuration. Alternative checkpoint panels and other plant configurations were not tested, and brief saturation between reporting points cannot be excluded. The exploratory context analysis establishes no causal predictor or operating threshold. Only three fixed references were evaluated, and AE was the energy outcome; OCI, measured blower electricity, tariffs and a complete energy or emissions balance were outside the study. These boundaries leave open how much of the identified aeration saving could be retained by an adaptive controller while preserving ammonium excursion performance. No alternative influent realization, parameter set, solver or BSM2 implementation was evaluated. Robustness to a different selection

of checkpoints and uncertainty across realizations therefore remain unquantified.

Ma et al. (2024) link aeration control to N₂O emissions in an alternating anoxic–oxic sequencing batch reactor. Such emissions were not evaluated here, so AE savings alone do not establish an overall environmental benefit.

CONCLUSIONS

Complete-state paired comparisons, direct excursion analysis and structured computational verification revealed ammonium penalties alongside gains in AE, TN and EQI. Both reduced DO references lowered AE and raised flow-weighted NH₄ in all 42 contexts and over the annual evaluation. Annual AE decreased by 1.444% and 2.857%, while mean NH₄ increased by 5.716% and 12.297% and time above 4 mg N/L by 16.084% and 36.364%. The annual NH₄ 95th percentile remained below the threshold. Comparing excursions with means and EQI therefore added information that aggregate improvements alone did not provide. The contribution lies in this paired assessment and its verification; fixed-reference reduction is an established intervention.

The post hoc rule identified 21 NH₄ failures among 168 candidate–horizon comparisons; seven coincided with improved EQI. Two additional BOD₅ failures gave 23 failed pairs overall. These counts depend on the stated margins. DO₂₀₀ had the lowest ammonium mean and excursion burden, with no KLa saturation observed on the 15 min reporting grid. Findings apply to one BSM2 realization, its influent sequence, the retained PI configuration and DO references of 2.00, 1.90 and 1.80 mg O₂/L. They establish neither universal real-plant ammonium risk nor metered blower electricity savings. Overall environmental benefit also requires assessment beyond AE and effluent indicators.

REFERENCES

- Alex, J., Benedetti, L., Copp, J. B., Gernaey, K. V., Jeppsson, U., Nopens, I., Pons, M.-N., Rosen, C., Steyer, J.-P., Vanrolleghem, P. A. (2008). *Benchmark simulation model no. 2 (BSM2) (Technical Report No. 3)*. IWA Task Group on Benchmarking of Control Strategies for Wastewater Treatment Plants. https://iwa-mia.org/wp-content/uploads/2022/09/TR3_BSM_TG_Tech_Report_no_3_BSM2_General_Description.pdf
- Åmand, L., Carlsson, B. (2012). Optimal aeration control in a nitrifying activated sludge process. *Water Research*, 46(7), 2101–2110. <https://doi.org/10.1016/j.watres.2012.01.023>
- Åmand, L., Olsson, G., Carlsson, B. (2013). Aeration control—A review. *Water Science and Technology*, 67(11), 2374–2398. <https://doi.org/10.2166/wst.2013.139>
- Barbu, M., Vilanova, R., Meneses, M., Santín, I. (2017). On the evaluation of the global impact of control strategies applied to wastewater treatment plants. *Journal of Cleaner Production*, 149, 396–405. <https://doi.org/10.1016/j.jclepro.2017.02.018>
- Chen, K., Wang, H., Valverde-Pérez, B., Zhai, S., Vezzaro, L., Wang, A. (2021). Optimal control towards sustainable wastewater treatment plants based on multi-agent reinforcement learning. *Chemosphere*, 279, Article 130498. <https://doi.org/10.1016/j.chemosphere.2021.130498>
- Do, H. A., Othman, M., Dang, H. T. T. (2026). Integrated CapdetWorks–BioWin modeling for energy optimization of a municipal wastewater treatment plant. *Ecological Engineering & Environmental Technology*, 27(6), 342–353. <https://doi.org/10.12912/27197050/221717>
- Gernaey, K. V., Jeppsson, U., Vanrolleghem, P. A., Copp, J. B. (Eds.). (2014). *Benchmarking of control strategies for wastewater treatment plants (Scientific and Technical Report No. 23)*. IWA Publishing. <https://doi.org/10.2166/9781780401171>
- Jeppsson, U., Pons, M.-N., Nopens, I., Alex, J., Copp, J. B., Gernaey, K. V., Rosen, C., Steyer, J.-P., Vanrolleghem, P. A. (2007). Benchmark Simulation Model No. 2: General protocol and exploratory case studies. *Water Science and Technology*, 56(8), 67–78. <https://doi.org/10.2166/wst.2007.604>
- Ma, X., Yang, W., Zhao, H., Tan, Q. (2024). Effects of aeration control strategies on nitrous oxide emissions in alternating anoxic–oxic sequencing batch reactor systems. *Environmental Research*, 260, Article 119591. <https://doi.org/10.1016/j.envres.2024.119591>
- Medinilla, V. R., Sprague, T., Marseilles, J., Burke, J., Deshmukh, S., Delagah, S., Sharbatmaleki, M. (2020). Impact of ammonia-based aeration control (ABAC) on energy consumption. *Applied Sciences*, 10(15), Article 5227. <https://doi.org/10.3390/app10155227>
- Meneses, M., Concepción, H., Vrečko, D., Vilanova, R. (2015). Life cycle assessment as an environmental evaluation tool for control strategies in wastewater treatment plants. *Journal of Cleaner Production*, 107, 653–661. <https://doi.org/10.1016/j.jclepro.2015.05.057>

12. Monday, C., Zaghloul, M. S., Krishnamurthy, D., Achari, G. (2024). A review of AI-driven control strategies in the activated sludge process with emphasis on aeration control. *Water*, 16(2), Article 305. <https://doi.org/10.3390/w16020305>
13. Nopens, I., Benedetti, L., Jeppsson, U., Pons, M.-N., Alex, J., Copp, J. B., Gernaey, K. V., Rosen, C., Steyer, J.-P., Vanrolleghem, P. A. (2010). Benchmark Simulation Model No. 2: Finalisation of plant layout and default control strategy. *Water Science and Technology*, 62(9), 1967–1974. <https://doi.org/10.2166/wst.2010.044>
14. Oussama, N., Goroneskul, M., Bouderdara, N., Mokhtari, M., Fidjah, A., Lyndiuk, O., Hrynovskyi, A., Hrynovskyi, O. (2026). Traditional clay-based groundwater recharge systems for sustainable water management and the One Health concept: The Adrar model in Algeria. *Ecological Engineering & Environmental Technology*, 27(9), 341–360. <https://doi.org/10.12912/27197050/226536>
15. Qambar, A. S., Al Khalidy, M. M. (2022). Optimizing dissolved oxygen requirement and energy consumption in wastewater treatment plant aeration tanks using machine learning. *Journal of Water Process Engineering*, 50, Article 103237. <https://doi.org/10.1016/j.jwpe.2022.103237>
16. Revollar, S., Meneses, M., Vilanova, R., Vega, P., Francisco, M. (2020a). Quantifying the benefit of a dynamic performance assessment of WWTP. *Processes*, 8(2), Article 206. <https://doi.org/10.3390/pr8020206>
17. Revollar, S., Vilanova, R., Vega, P., Francisco, M., Meneses, M. (2020b). Wastewater treatment plant operation: Simple control schemes with a holistic perspective. *Sustainability*, 12(3), Article 768. <https://doi.org/10.3390/su12030768>
18. Revollar, S., Meneses, M., Vilanova, R., Vega, P., Francisco, M. (2021). Eco-efficiency assessment of control actions in wastewater treatment plants. *Water*, 13(5), Article 612. <https://doi.org/10.3390/w13050612>
19. Rieger, L., Jones, R. M., Dold, P. L., Bott, C. B. (2014). Ammonia-based feedforward and feedback aeration control in activated sludge processes. *Water Environment Research*, 86(1), 63–73. <https://doi.org/10.2175/106143013X13596524516987>
20. Ruano, M. V., Ribes, J., Ruiz-Martínez, A., Seco, A., Robles, A. (2024). An advanced control system for nitrogen removal and energy consumption optimization in full-scale wastewater treatment plants. *Journal of Water Process Engineering*, 57, Article 104705. <https://doi.org/10.1016/j.jwpe.2023.104705>
21. Santín, I., Meneses, M., Pedret, C., Barbu, M., Vilanova, R. (2025). Improving hierarchical ammonium control operation in wastewater treatment with non-ideal sensors and actuators using filters, adaptive control and measurement error prediction. *Systems Science & Control Engineering*, 13(1), Article 2454700. <https://doi.org/10.1080/21642583.2025.2454700>
22. Santín, I., Pedret, C., Vilanova, R. (2015). Applying variable dissolved oxygen set point in a two-level hierarchical control structure to a wastewater treatment process. *Journal of Process Control*, 28, 40–55. <https://doi.org/10.1016/j.jprocont.2015.02.005>
23. Schraa, O., Rieger, L., Alex, J., Miletic, I. (2019). Ammonia-based aeration control with optimal SRT control: Improved performance and lower energy consumption. *Water Science and Technology*, 79(1), 63–72. <https://doi.org/10.2166/wst.2019.032>
24. Vrečko, D., Gernaey, K. V., Rosen, C., Jeppsson, U. (2006a). Benchmark Simulation Model No. 2 in Matlab–Simulink: Towards plant-wide WWTP control strategy evaluation. *Water Science and Technology*, 54(8), 65–72. <https://doi.org/10.2166/wst.2006.773>
25. Vrečko, D., Hvala, N., Stare, A., Burica, O., Stražar, M., Levstek, M., Cerar, P., Podbevšek, S. (2006b). Improvement of ammonia removal in activated sludge process with feedforward-feedback aeration controllers. *Water Science and Technology*, 53(4–5), 125–132. <https://doi.org/10.2166/wst.2006.098>
26. Vrečko, D., Volcke, E. I. P., Jeppsson, U., Gernaey, K. V., Vanrolleghem, P. A. (2009). Evaluation criteria description and example for BSM2. *IWA Task Group on Benchmarking of Control Strategies for Wastewater Treatment Plants*. <https://assets.pubpub.org/qs7o9dx/51603225626306.pdf>